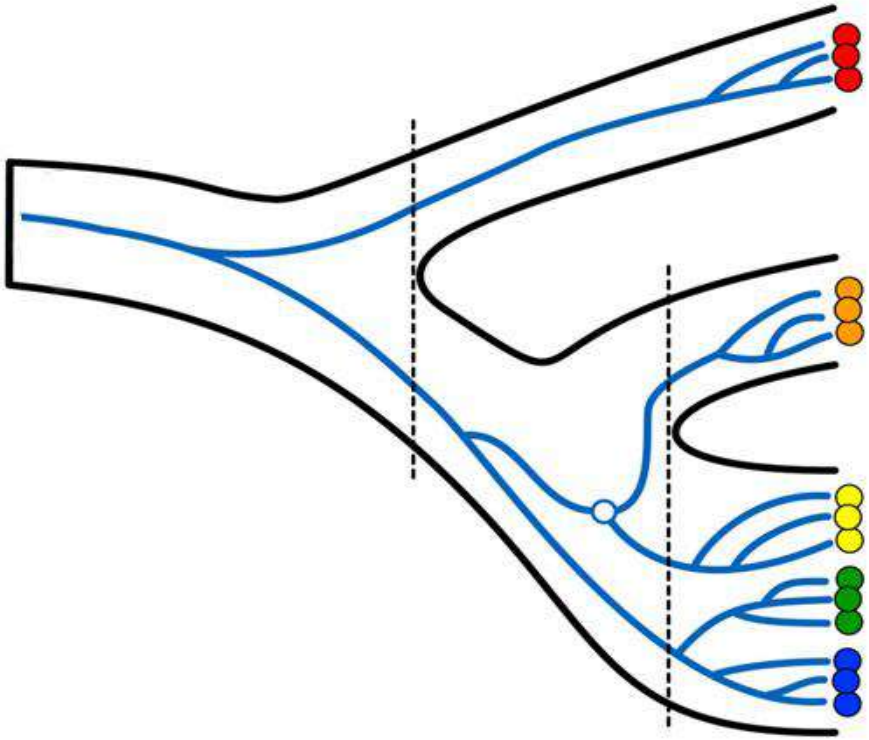
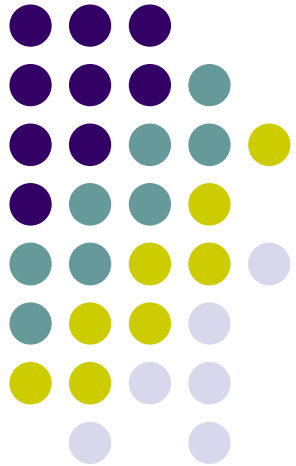
DNA-based species delimitation



Frederik Leliaert
Olivier De Clerck

Content

- Use of DNA in evolutionary and ecological studies
- Species concept
- Delineating species using DNA



Applications of molecular techniques

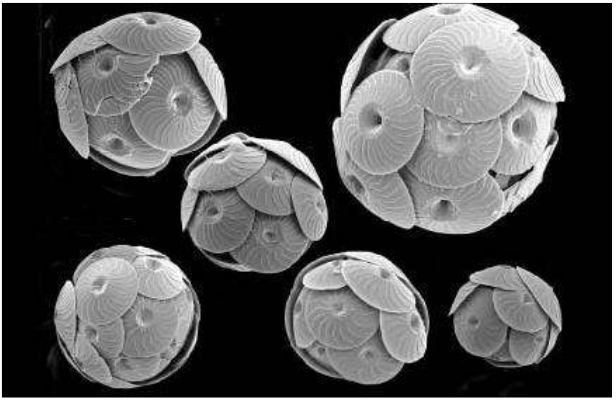
- DNA
  - Every living organism contains DNA
  - Every organism is characterized by a unique DNA signature
  - Through common descent DNA of closely of related organisms will be more similar than that of distantly related organisms

Applications of molecular techniques

- DNA to identify organisms
  - “barcode”
  - fast – accurate

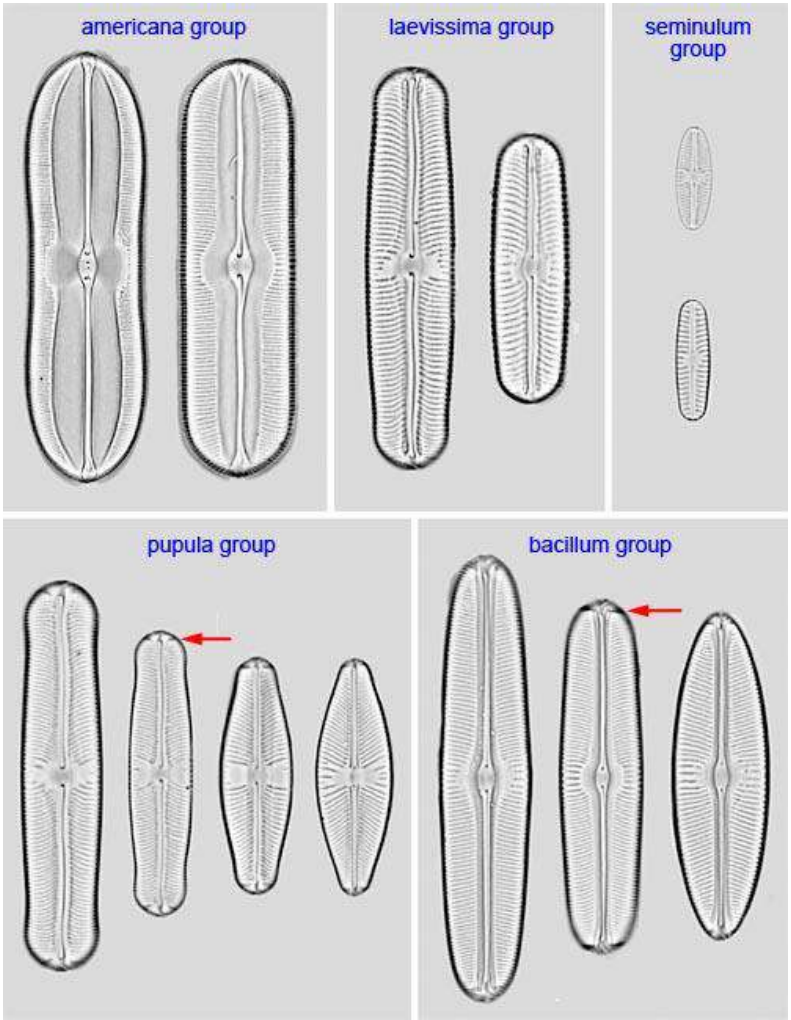
Applications of molecular techniques

- DNA to identify organisms
  - circumvents cryptic diversity
  - prominent in morphologically ‘simple’ organisms



←[coccolithophores]

[*Sellaphora* diatoms] →



Applications of molecular techniques

BBC NEWS | Watch One-Minute World News | News services: Your news when you want it

News Front Page
Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
Business
Health
Science & Environment
Technology
Entertainment
Also in the news
Video and Audio
Programmes
Have Your Say
In Pictures
Country Profiles
Special Reports

RELATED BBC SITES
SPORT
WEATHER
ON THIS DAY
EDITORS' BLOG

Last Updated: Friday, 21 December 2007, 23:38 GMT

E-mail this to a friend Printable version

Not one but 'six giraffe species'

Anna-Marie Lever
Science and nature reporter, BBC News

The world's tallest animal, the giraffe, may actually be several species, a study has found.

A report in BMC Biology uses genetic evidence to show that there may be at least six species of giraffe in Africa.

Currently giraffes are considered to represent a single species classified into multiple subspecies.

The study shows geographic variation in hair coat colour is evident across the giraffe's range in sub-Saharan Africa, suggesting reproductive isolation.

Giraffe populations have dropped by 30% over the past decade

"Using molecular techniques we found that giraffes can be classified into six groups that are reproductively isolated and not interbreeding," David Brown, the lead author of the study and a geneticist at the University of California Los Angeles (UCLA), told BBC News.

SEE ALSO
- Hungry Zimbabweans target giraffe
  22 Sep 07 | Africa
- Virus thwarts giraffe's romance
  01 Oct 07 | Devon
- Zoo starts hunt for new giraffes
  20 Feb 06 | Devon

RELATED BBC LINKS
- Science and Nature: Giraffe

RELATED INTERNET LINKS
- UCLA
- Wildlife Conservation Society
- BMC Biology

The BBC is not responsible for the content of external internet sites

TOP SCIENCE & ENVIRONMENT STORIES
- Night-sky image is biggest ever
- Phantom Eye 'spy plane' unveiled
- Higgs discovery rumour is denied

News feeds

MOST POPULAR STORIES NOW

MOST SHARED | MOST READ

1 Dozens killed in Libya 'massacre'
2 China police break up 'protests'

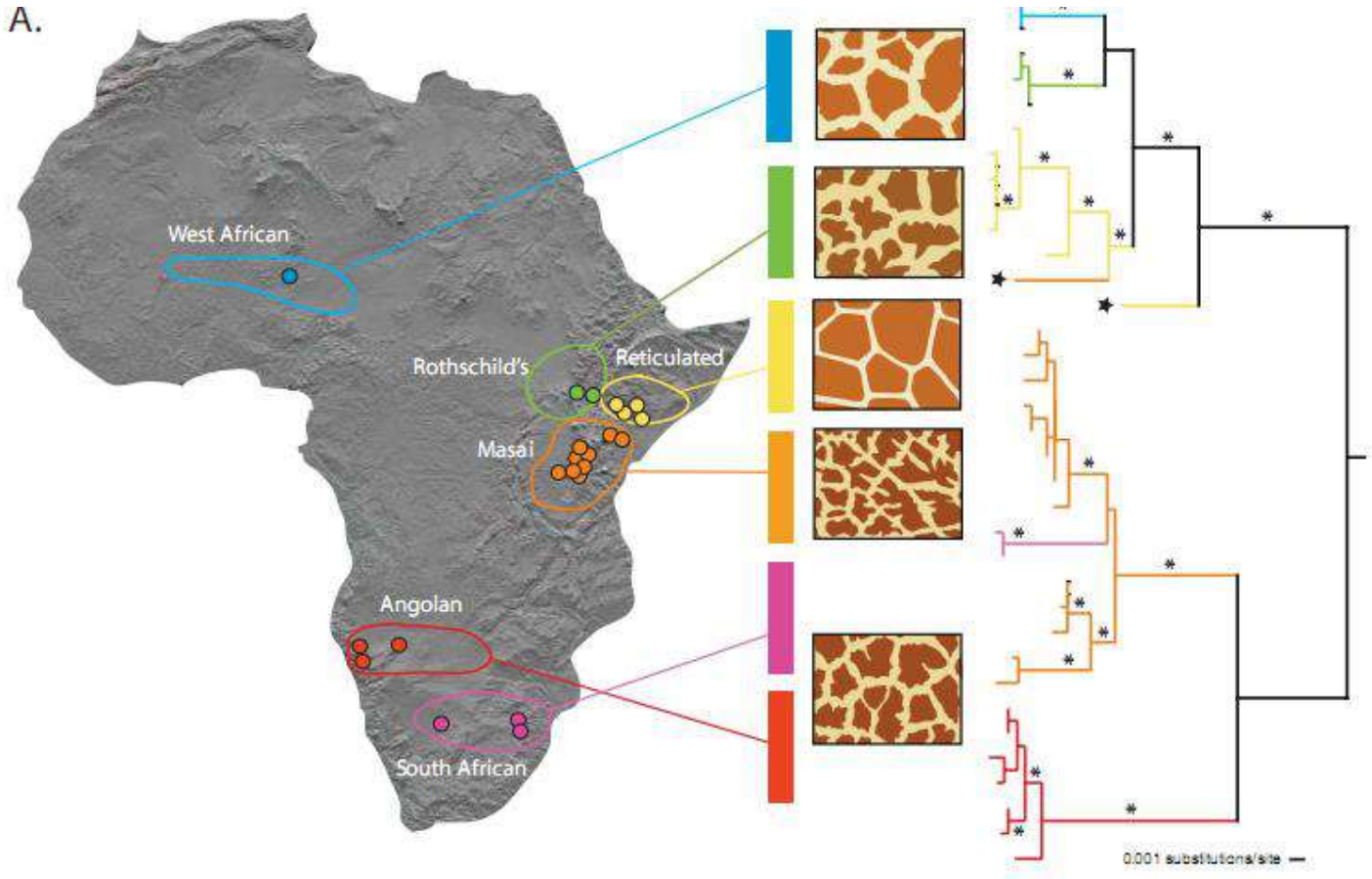
Applications of molecular techniques

Research article

Open Access

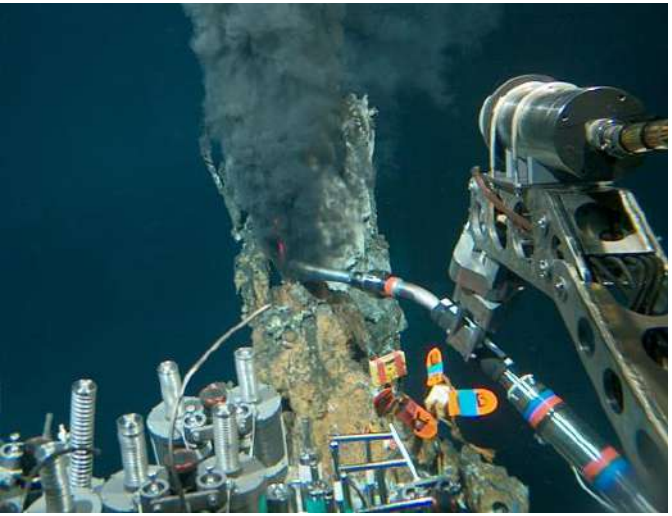
Extensive population genetic structure in the giraffe

David M Brown[1], Rick A Brenneman[2], Klaus-Peter Koepfli[1], John P Pollinger[1], Borja Milá[1], Nicholas J Georgiadis[3], Edward E Louis Jr[2], Gregory F Grether[1], David K Jacobs[1] and Robert K Wayne*[1]

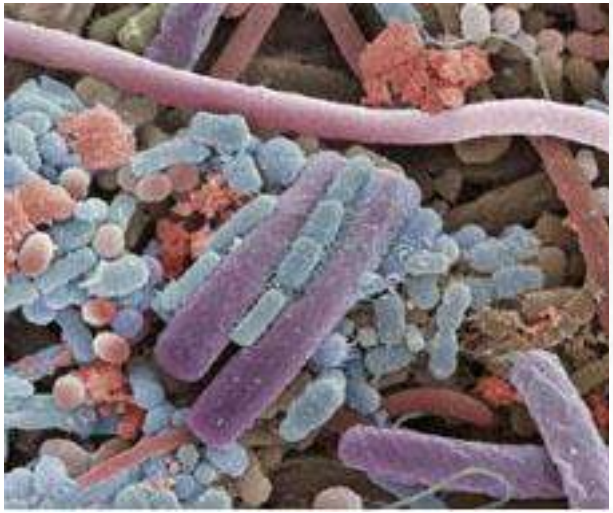


Applications of molecular techniques

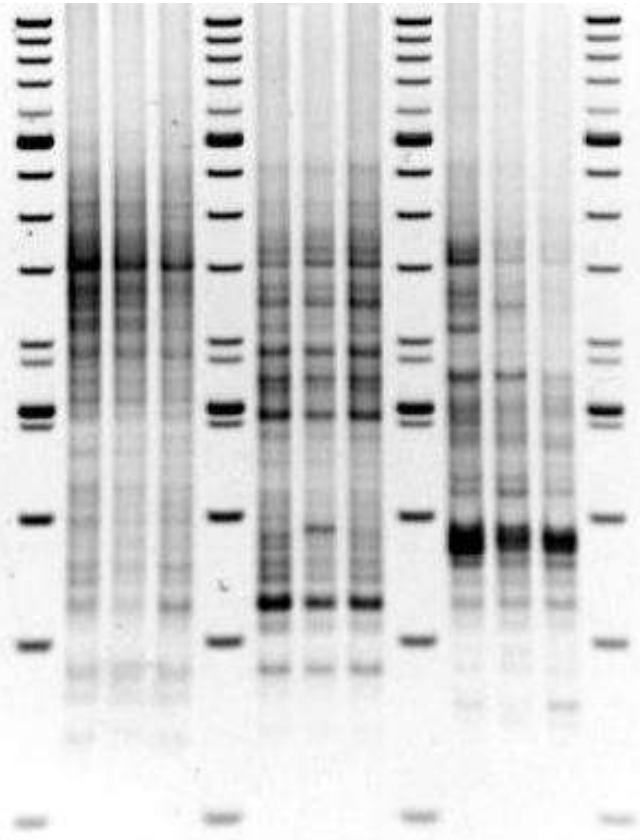
- DNA to identify organisms
  - circumvents non-culturable diveristy
    - e.g. microbial communties



[black smokers]



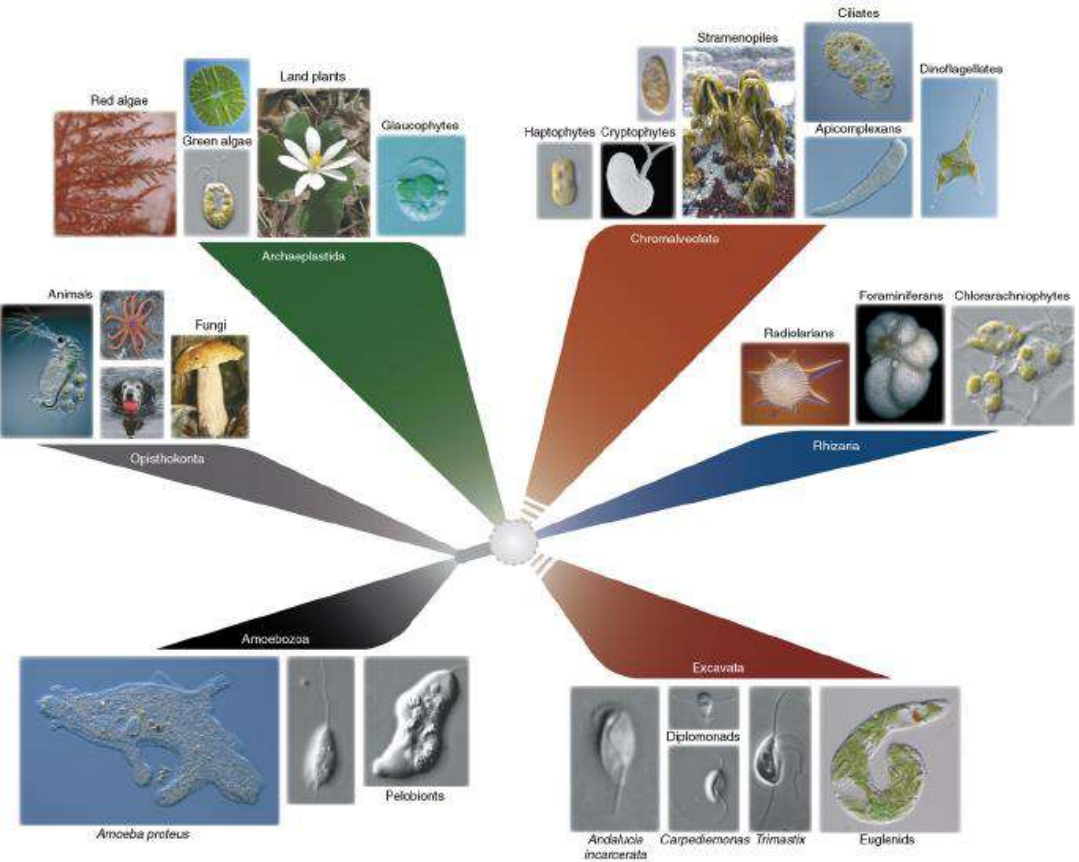
[bacteria TEM]

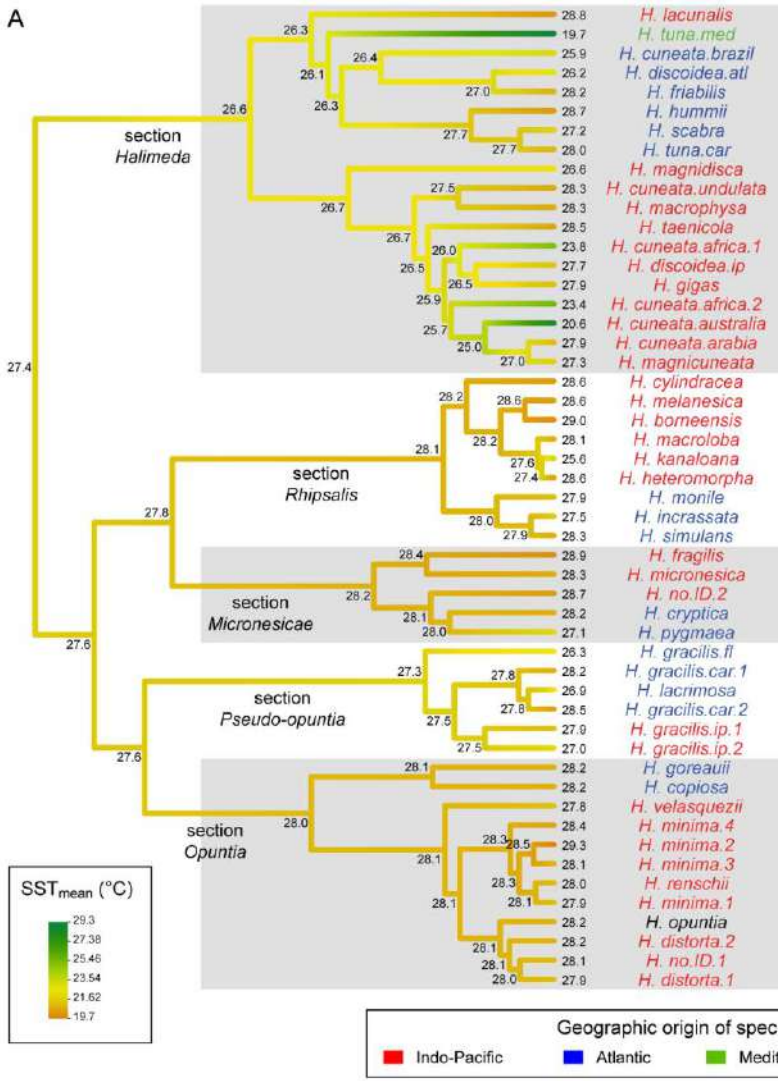


[DGGE gel]

Applications of molecular techniques

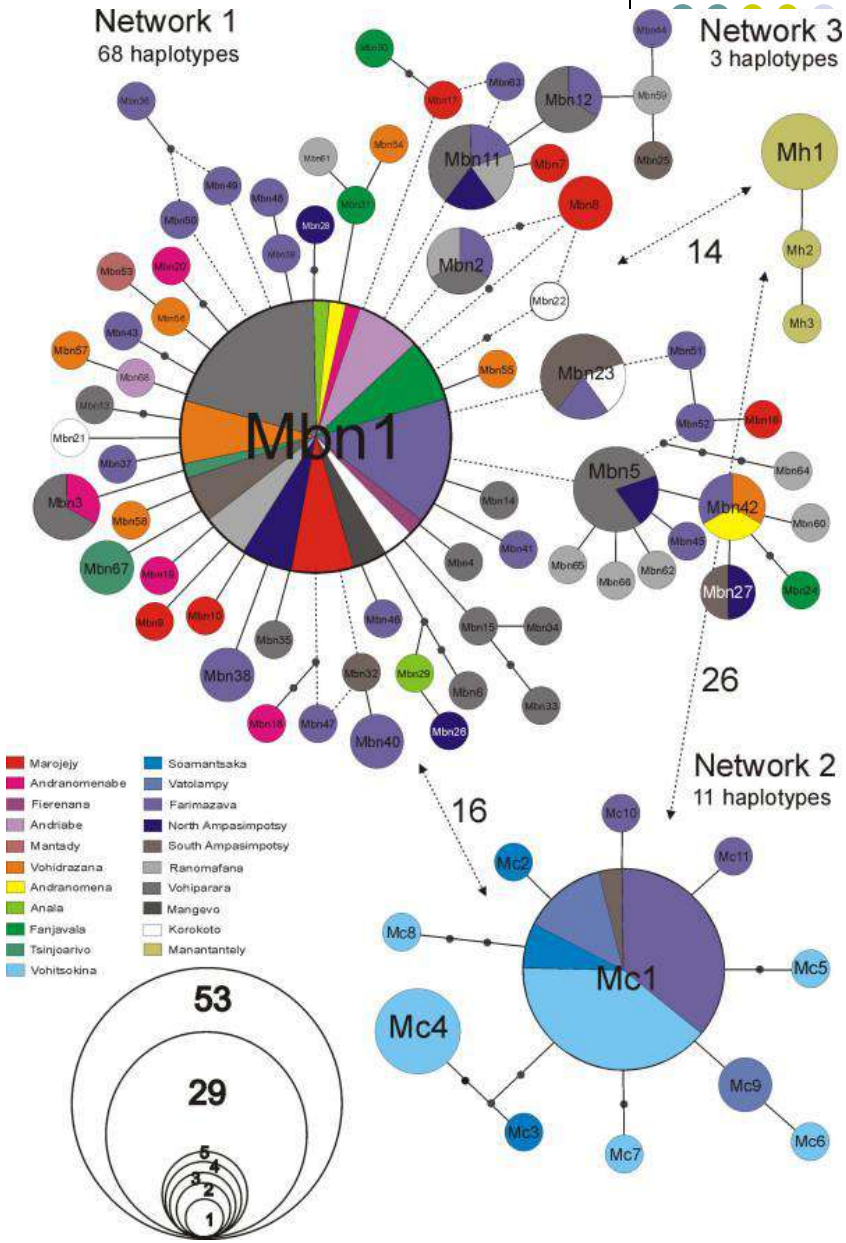
- DNA to identify relationships
  - from ancient relationships …
  - … to populations





Applications of molecular techniques

- DNA to identify relationships
  - from ancient relationships …
  - … to populations



Delineating species

- Higher taxonomic ranks are human constructs
- Species are real (more or less)

II. PHYLOGENY IS REAL, CLASSIFICATIONS ARE NOT

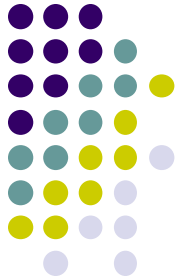
The phylogeny of life is real. There is a single evolutionary tree linking all organisms living and extinct (Darwin, 1859), and relationships are measurable entities that depend on the relative time since any pair of organisms last shared a common ancestor.

Biological classifications, then, are necessarily entirely human constructs. There is no single, true classification inherent in Nature that is there to be discovered. In this paper, I use the term 'classification' in two ways, as is common practice: first, *a classification* means an ordered list of species or higher taxa, and second, *classification* describes the process of achieving such an ordered list. The usages should be clear from context.

[Benton 2000]

Most biologists accept that species are real (Darwin, 1859; Huxley, 1940), allowing of course for the complications of incipient species, hybrids, sibling species, and the like. The 'reality' of species is different from that of the chemical elements in many ways, not least in the fact that they are linked by historical lines of descent (Darwin, 1859), and that their existence is finite in time.

Delineating species



- What is a species ?

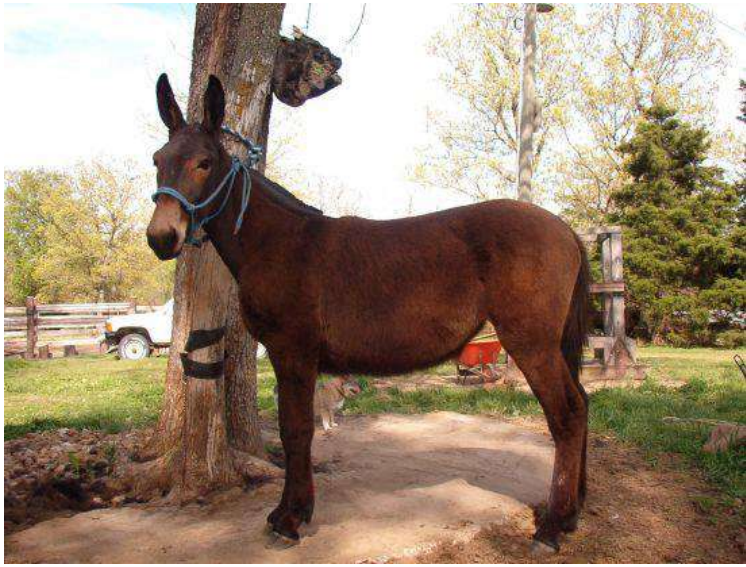
  - The simpler the question, the more difficult the answer …
  - Alternatively, the answer is difficult because the question is flawed

- It is surprisingly difficult to define the word "species" in a way that applies to all naturally occurring organisms, and the debate among biologists about how to define "species" and how to identify actual species is called the species problem.

- Ernst Mayr's biological species concept: a species comprises all the individual organisms of a natural population that generally interbreed at maturity in the wild and whose interbreeding produces fertile offspring.

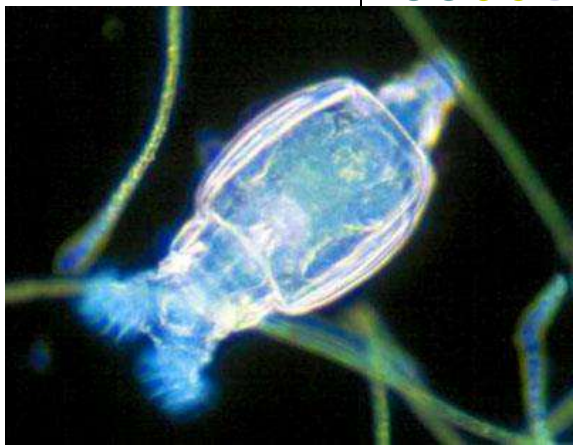
Delineating species

- What is a species ?
- Ernst Mayr's biological species concept: a species comprises all the individual organisms of a natural population that generally interbreed at maturity in the w and whose interbreeding produces fertile offspring.



Delineating species

- Problems with the biological species concept:
  - only works for sexual organisms
  - ignorance about the capability of morphologically similar groups of organisms to "potentially" interbreeding in nature

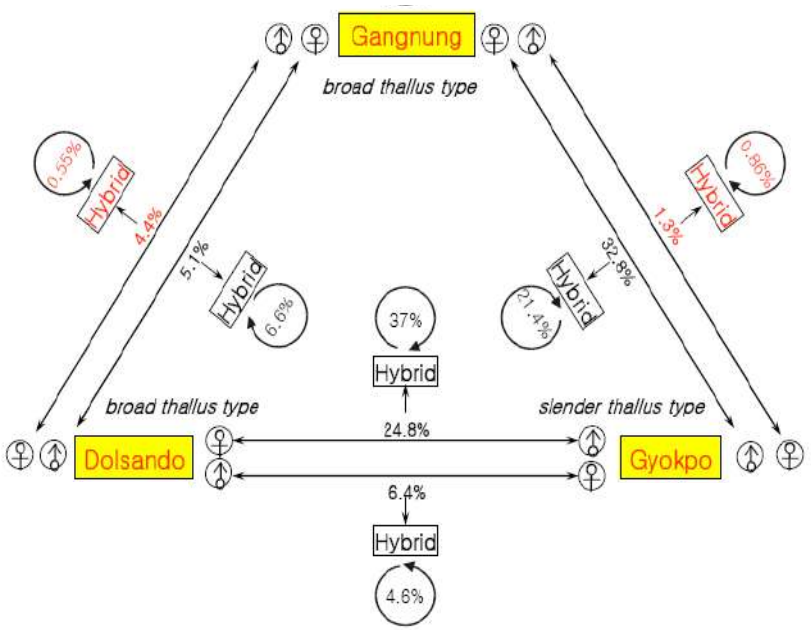


Bdelloid rotifers

Some red algae Stylonema

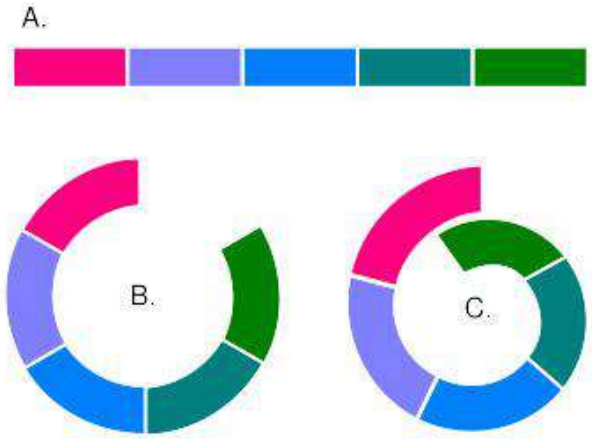
Delineating species

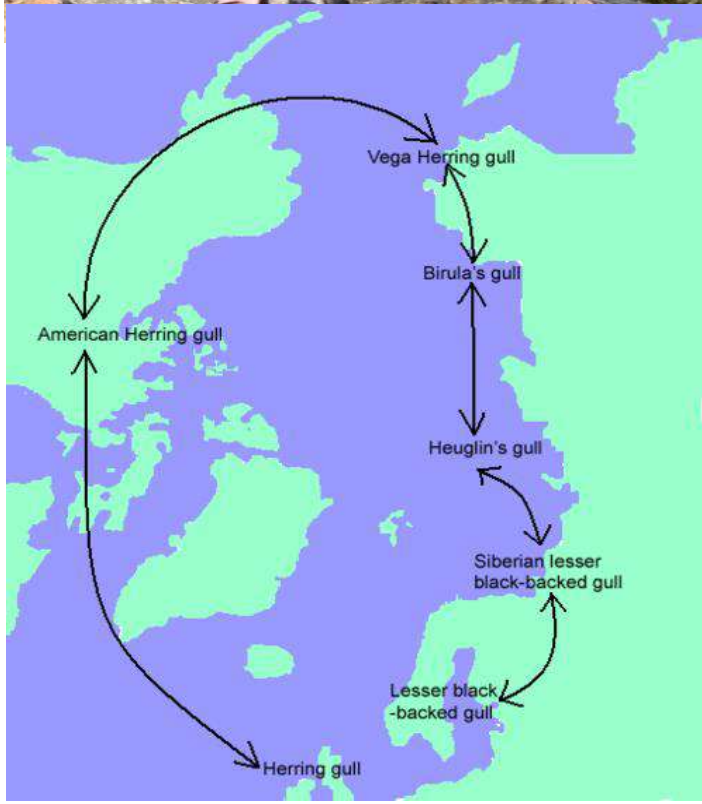
- Problems with the <u>biological species concept</u>:
  - reproductive isolation is a continuous character
    a matter of degree rather than presence/absence



Delineating species

- Problems with the biological species concept:
  - how to deal with ring species?





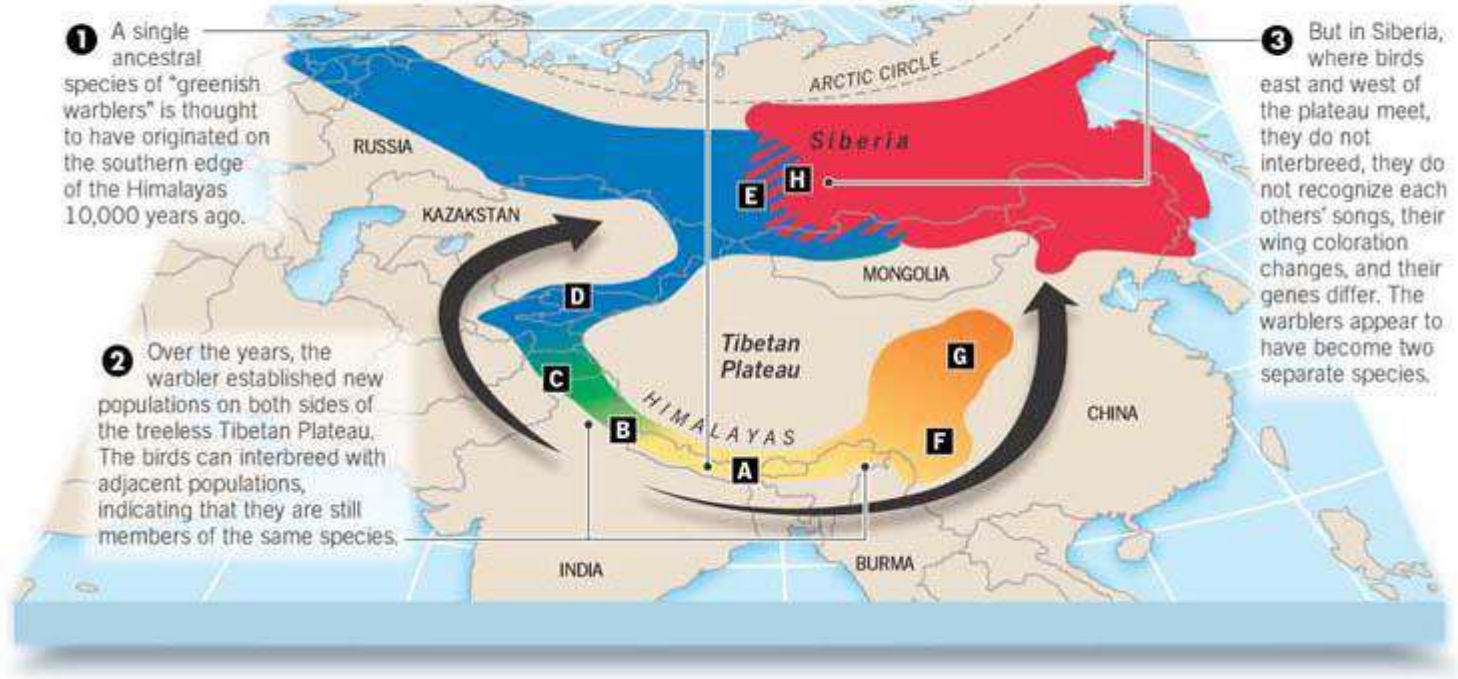
Delineating species

- Problems with the <u>biological species concept</u>:
  - how to deal with ring species?

Phylloscopus trochilus

Tracing the Evolution of Species

Biologists have discovered two populations of Eurasian songbirds in Siberia that show the strongest evidence yet of having evolved from a single ancestral species into two distinct ones. The map below shows the present ranges of the birds around the Tibetan Plateau, with gradations of color indicating where gradual changes have evolved between one subspecies and another.



Delineating species

- Problems with the biological species concept:
  - physical constraints ?

Delineating species

- What is a species ?

  - The simpler the question, the more difficult the answer …
  - Alternatively, the answer is difficult because the question is flawed

- Proliferation of species concepts

TABLE 1. A partial listing of species concepts and other ideas about

Name of species concept (alphabetically arranged)	Brief definition
Darwin's morphological concept	"Varieties" between which there are no or few morphological intermediates
Diagnostic ("phylogenetic") Concept	A species "is an irreducible (basal) cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent"
Ecological Concept	"A lineage which occupies an adaptive zone minimally different from that of any other lineage..."
Evolutionary concept	A lineage evolving separately and "with its own unitary evolutionary role and tendencies"
Genealogical concept	Species are mutually monophyletic in the genealogies at all (or at a consensus of) gene genealogies in the genome
Phenetic concept	Clusters of individuals circumscribed using multivariate statistical analysis
Polytypic Species	Taxa having many "types," i.e., geographic subspecies. Geographic populations are part of the same species if they intergrade in areas of overlap
Population concept	Populations are the real units of evolution, not species, because gene flow is generally weak. Morphological and genetic uniformity of species is explained by stabilizing selection acting separately in each population
Recognition concept	Species possess a shared fertilization system, known as "specific-mate recognition systems"
Taxonomy without species	Species are no more real than any other hierarchical level in the tree of life. Species and other taxonomic ranks should be replaced either by "rank-free taxonomy" (which can name each node in a bifurcating phylogeny—Mishler), or by genotypic clusters described according to their genetic divergence from other clusters (Hendry et al.)

Delineating species

19

A hierarchy of species concepts: the denouement in the saga of the species problem

R. L. Mayden

Contacting address: Department of Biological Sciences, P.O. Box 0344, University of Alabama, Tuscaloosa, AL 35487, USA

Table 19.1 Species concepts and standardized abbreviations

1. Agamospecies (ASC)
2. Biological (BSC)
3. Cohesion (CSC)
4. Cladistic (ClSC)
5. Composite (CpSC)
6. Ecological (EcSC)
7. Evolutionary Significant Unit (ESU)
8. Evolutionary (ESC)
9. Genealogical Concordance (GCC)
10. Genetic (GSC)
11. Genotypic Cluster Definition (GCD)
12. Hennigian (HSC)
13. Internodal (ISC)
14. Morphological (MSC)
15. Non-dimensional (NDSC)
16. Phenetic (PhSC)
17. Phylogenetic (PSC)
    1. Diagnosable Version (PSC_1)
    2. Monophyly Version (PSC_2)
    3. Diagnosable and Monophyly Version (PSC_3)
18. Polythetic (PtSC)
19. Recognition (RSC)
20. Reproductive Competition (RCC)
21. Successional (SSC)
22. Taxonomic (TSC)

Delineating species

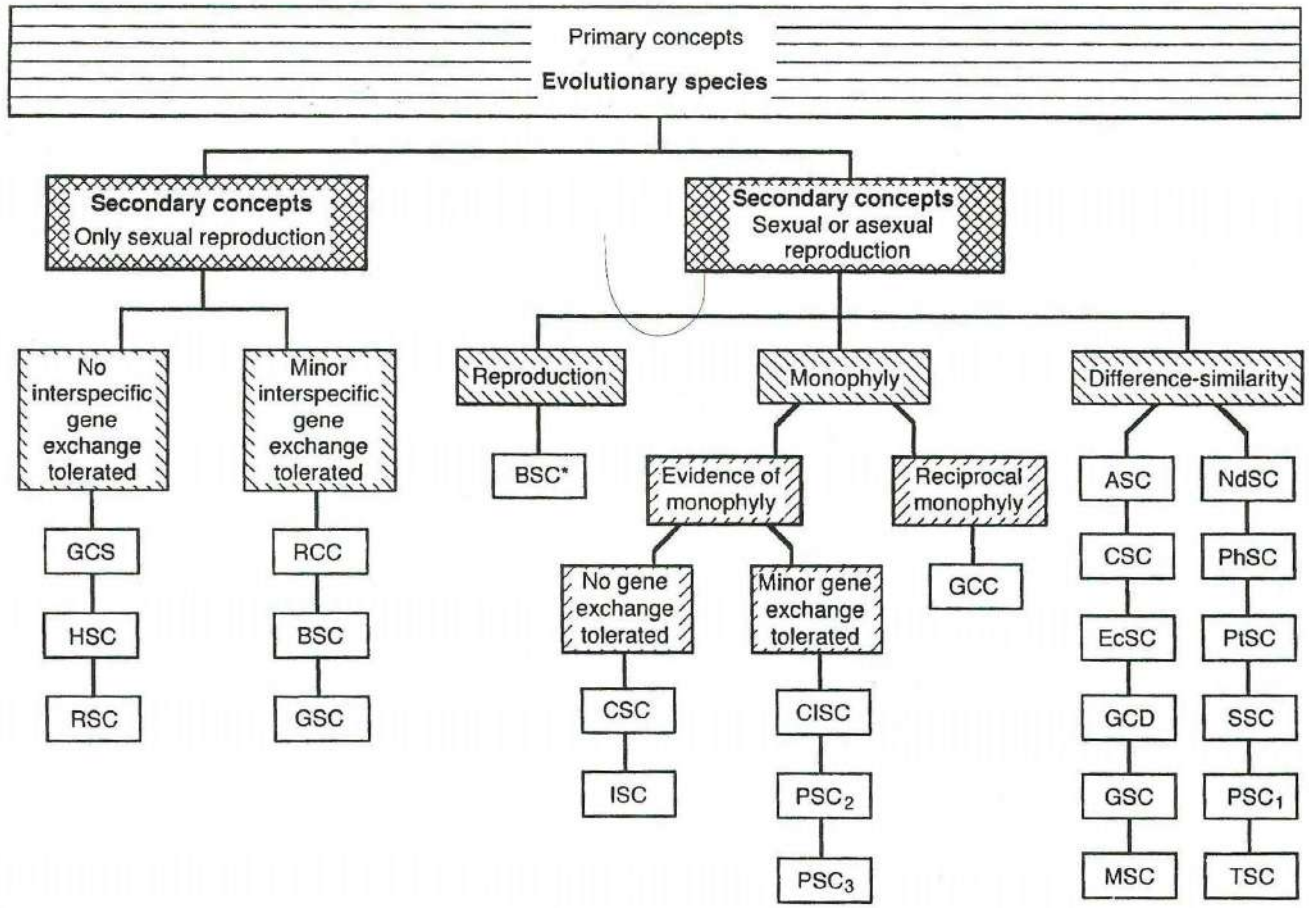


Figure 19.1 A hierarchy of primary and secondary species concepts. The non-operational Evolutionary Species Concept serves a primary concept of species. The operational secondary concepts form a hierarchy below this primary concept based on their toler or requirements for modes of reproduction, gene exchange, monophyly, and diagnosability. Because some concepts represent h versions of other concepts (mixed criteria) they may be depicted more than once in the hierarchy. Species concepts are listed a betically within any grouping. Asterisk denotes a version of BSC modified for asexual species. See Table 19.1 for concept abbrevia

Delineating species

- What is a species ?
- The general lineage concept

> *One should never quarrel about words, and never get involved in questions of terminology. One should always keep away from discussing concepts.*
>
> —Karl Popper, *Objective Knowledge: An Evolutionary Approach*

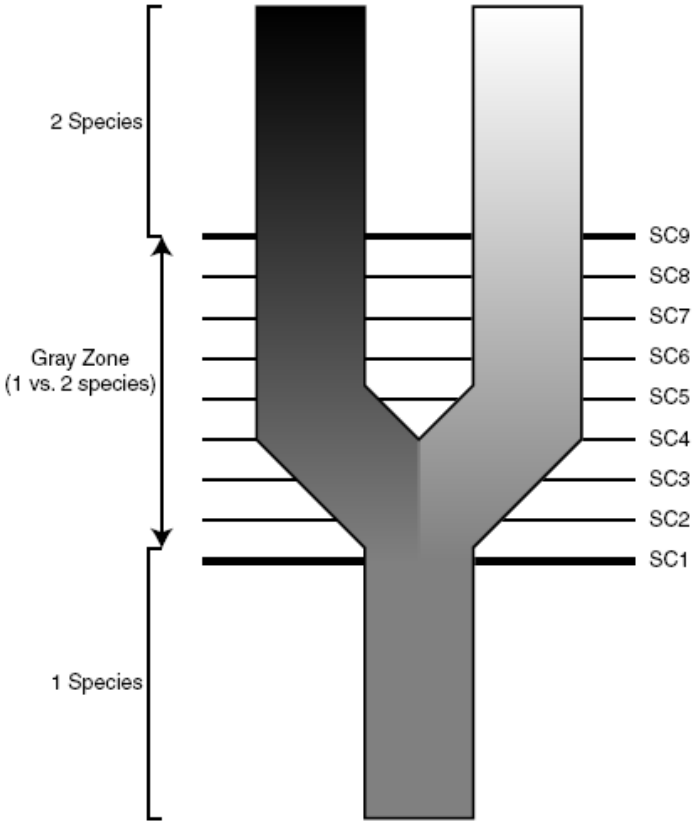
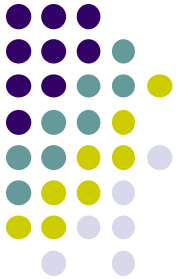


FIGURE 1. Lineage separation and divergence (speciation) and species concepts (after de Queiroz, 1998, 1999, 2005a). This highly simplified diagram represents a single lineage (species) splitting to form two lineages (species). The gradations in shades of gray represent the daughter lineages diverging through time, and the horizontal lines labeled SC (species criterion) 1 to 9 represent the times at which they acquire different properties (i.e., when they become phenetically distinguishable, diagnosable, reciprocally monophyletic, reproductively incompatible, ecologically distinct, etc.). The entire set of properties forms a gray zone within which alternative species concepts come into conflict. On either side of the gray zone, there will be unanimous agreement about the number of species. Before the acquisition of the first property, everyone will agree that there is a single species, and after the acquisition of the last property, everyone will agree that there are two. In between, however, there will be disagreement. The reason is that different contemporary species concepts adopt different properties (represented by the horizontal lines) as their species criteria—that is, as their cutoffs for considering a separately evolving lineage to have become a species.

[De Queiroz 2007]

Barcoding species



- If species are real and we agree on the ‘general lineage concept’ how can we delineate them ?

- The Barcoding option

  The value: quick & dirty

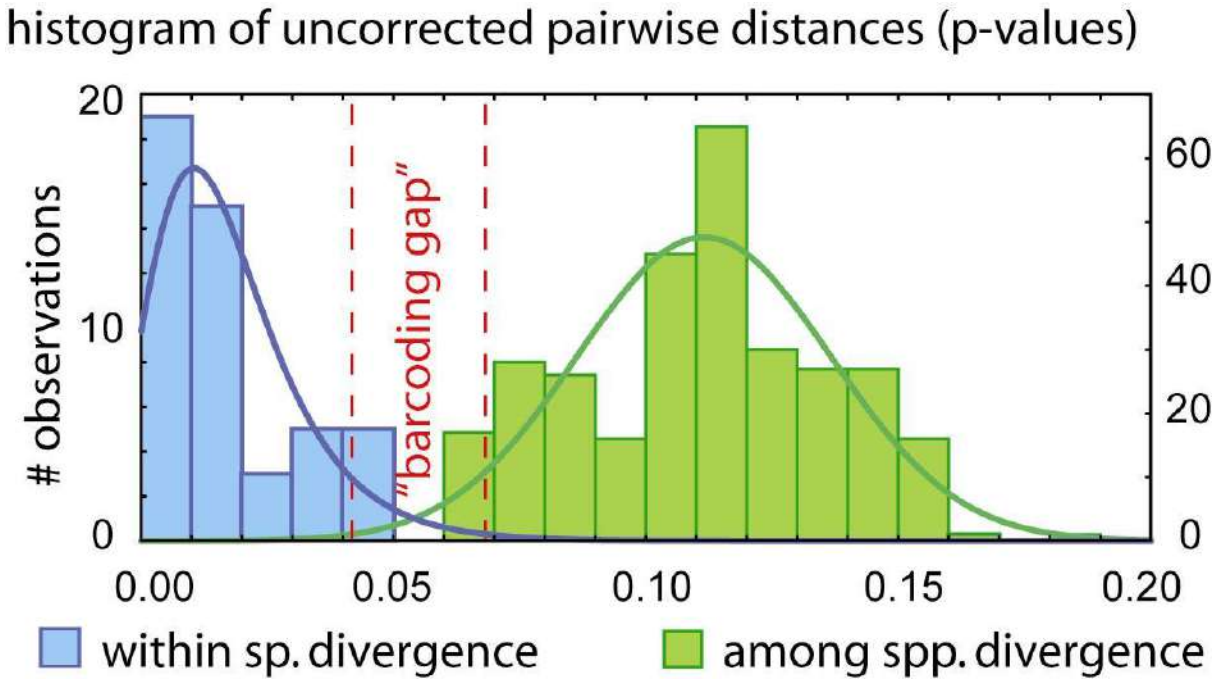
  - easy (even machines can do it)
  - cheap (compared to the salary of taxonomist)
  - fast (96-well plate takes 2 hours to sequence)

  The drawback:

  - you pay for simplicity in the long run
  - ignorance about speciation

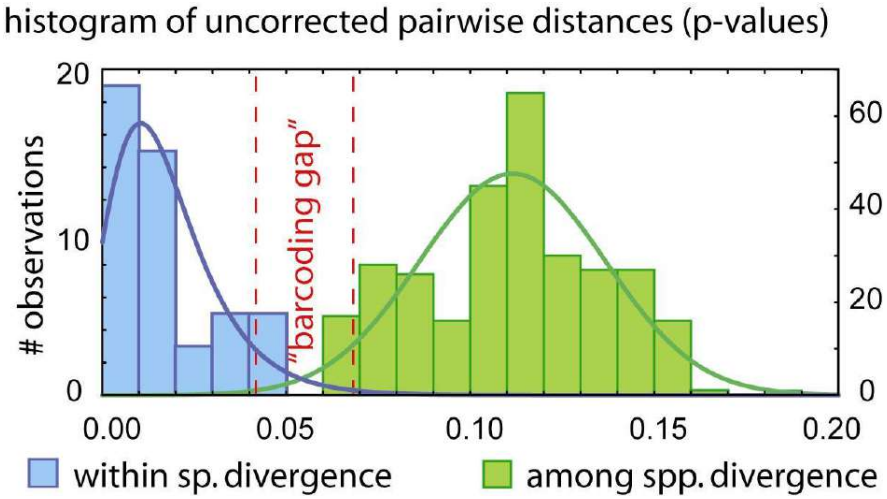
Barcoding species

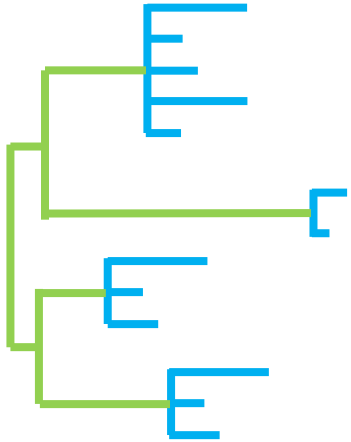
- The central tenet in barcoding



Barcoding species

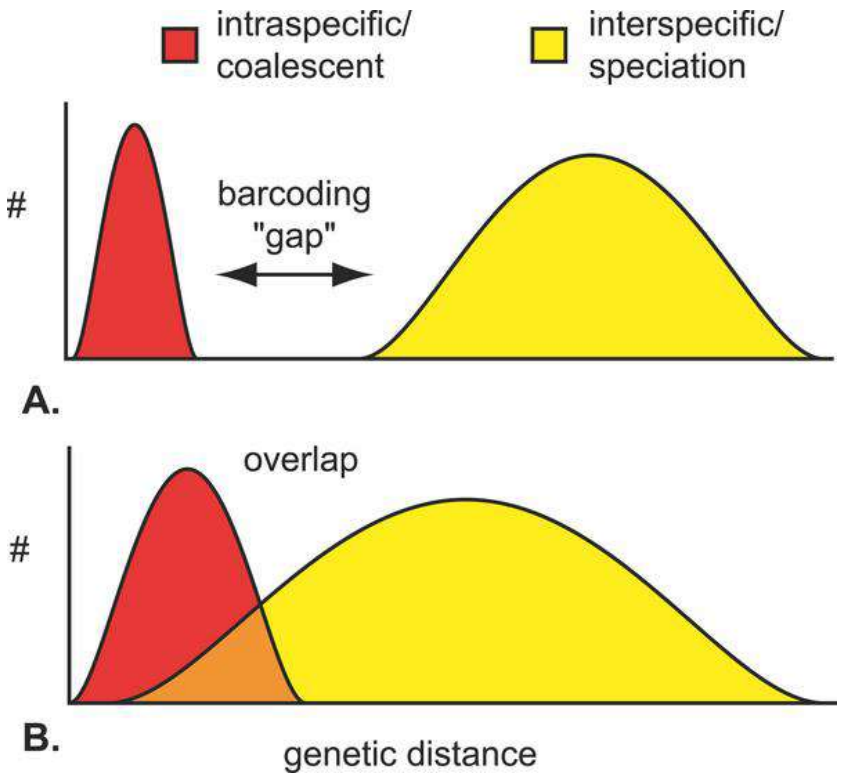
- The central tenet in barcoding





Barcoding species

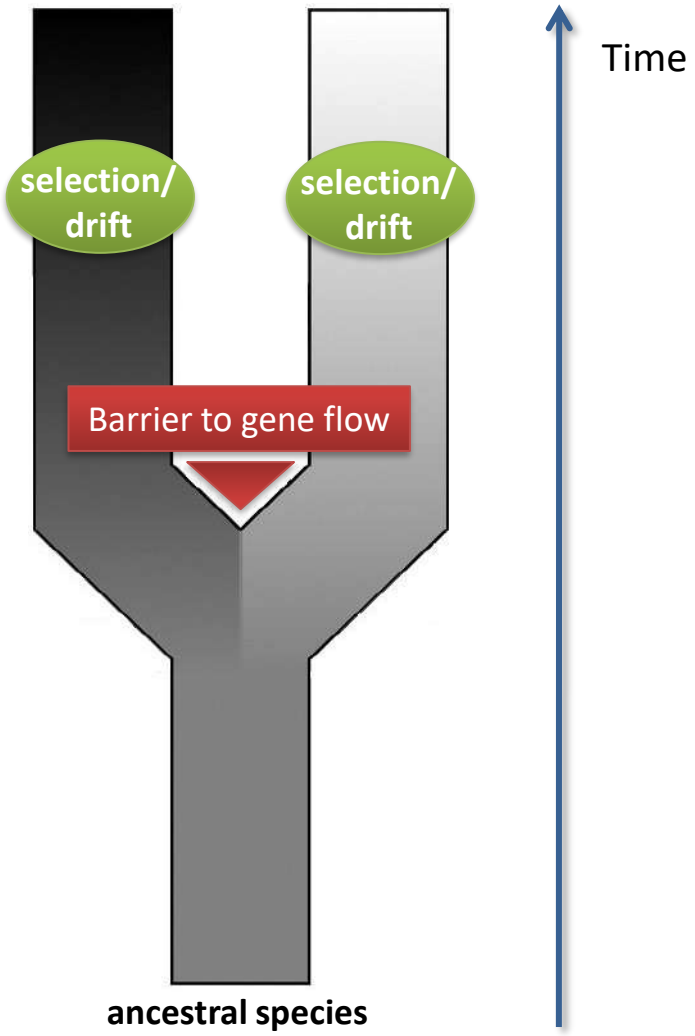
- The central tenet in barcoding



- Does a barcoding gap equals different species ?
- and vice versa ?
- How does sampling affect the gap ?

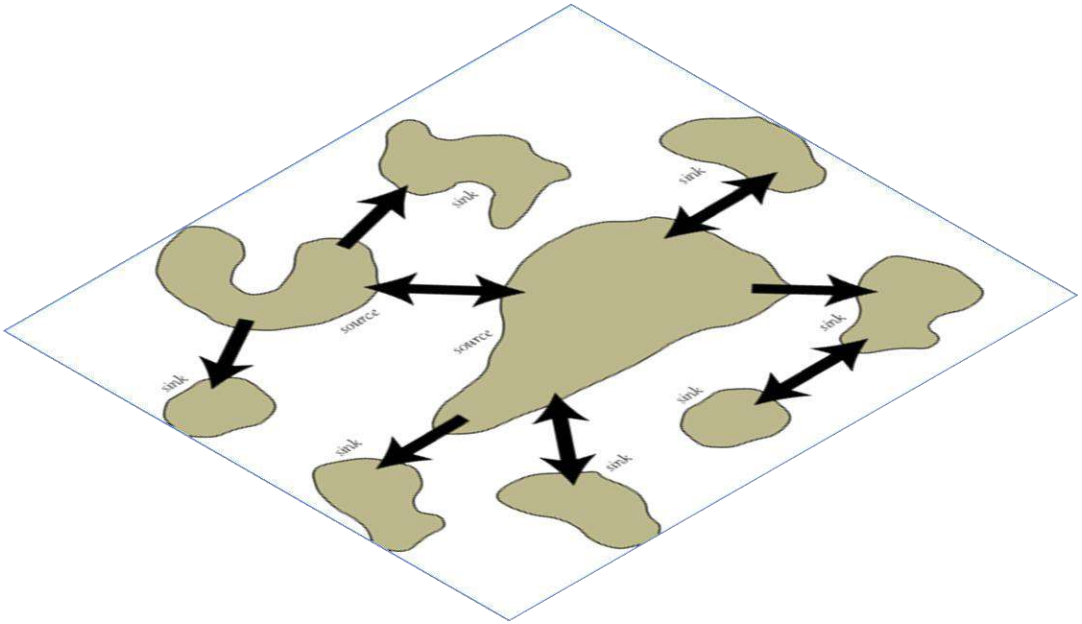
Beyond the barcoding gap - a more realistic view on speciation

separately evolving metapopulation lineages



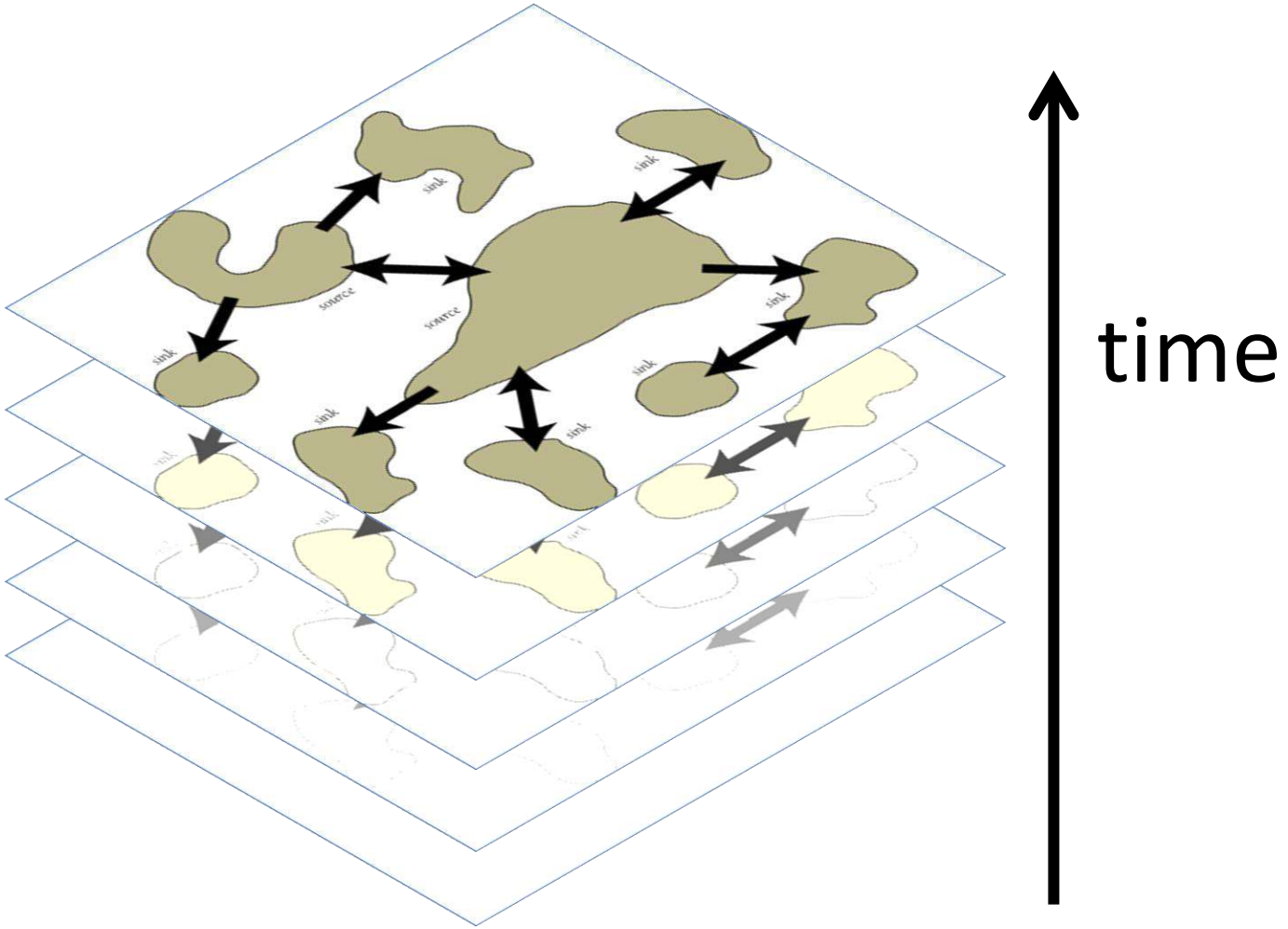
species =
separately evolving metapopulation lineage

Metapopulation = a group of spatially separated populations that are connected at some level

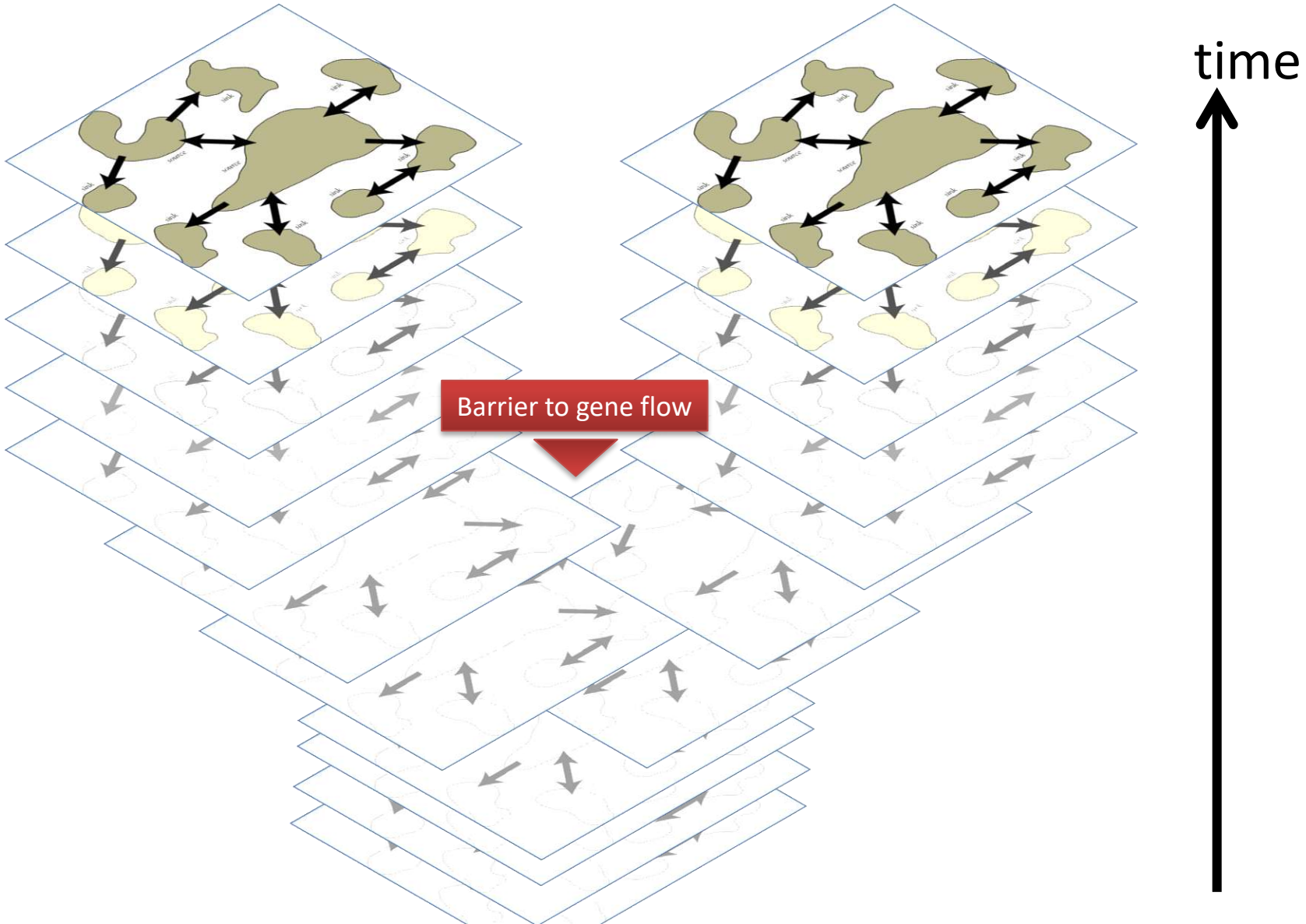


species =
separately evolving metapopulation lineage

Metapopulation extended through **time**



species =
separately evolving metapopulation lineages



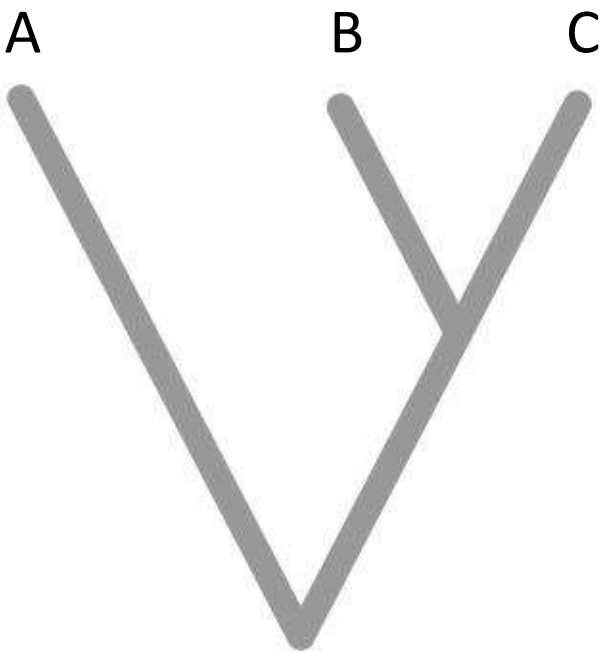
Phylogeny at the level of populations and species

Species are metapopulation lineages > new methods for DNA-based species delimitation

Genes: variable > informative below & above species level

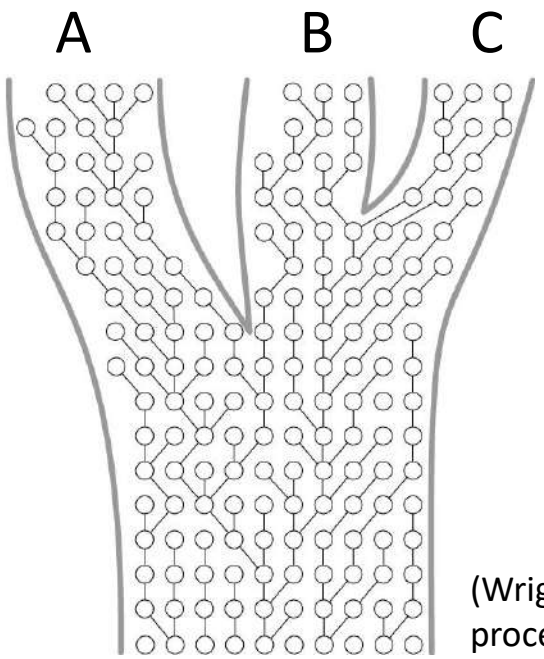
Gene trees

- Vital to understanding the process of speciation
- Span intraspecific and interspecific evolution



Species phylogeny

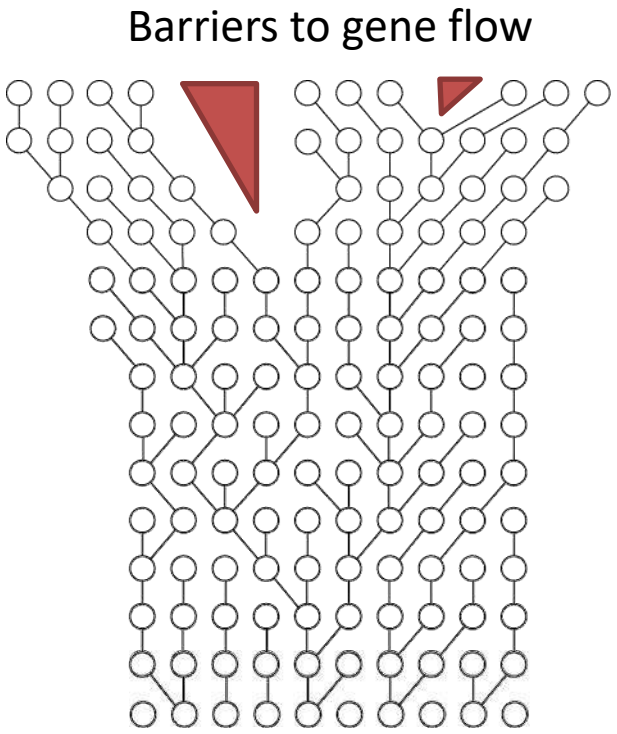
>>



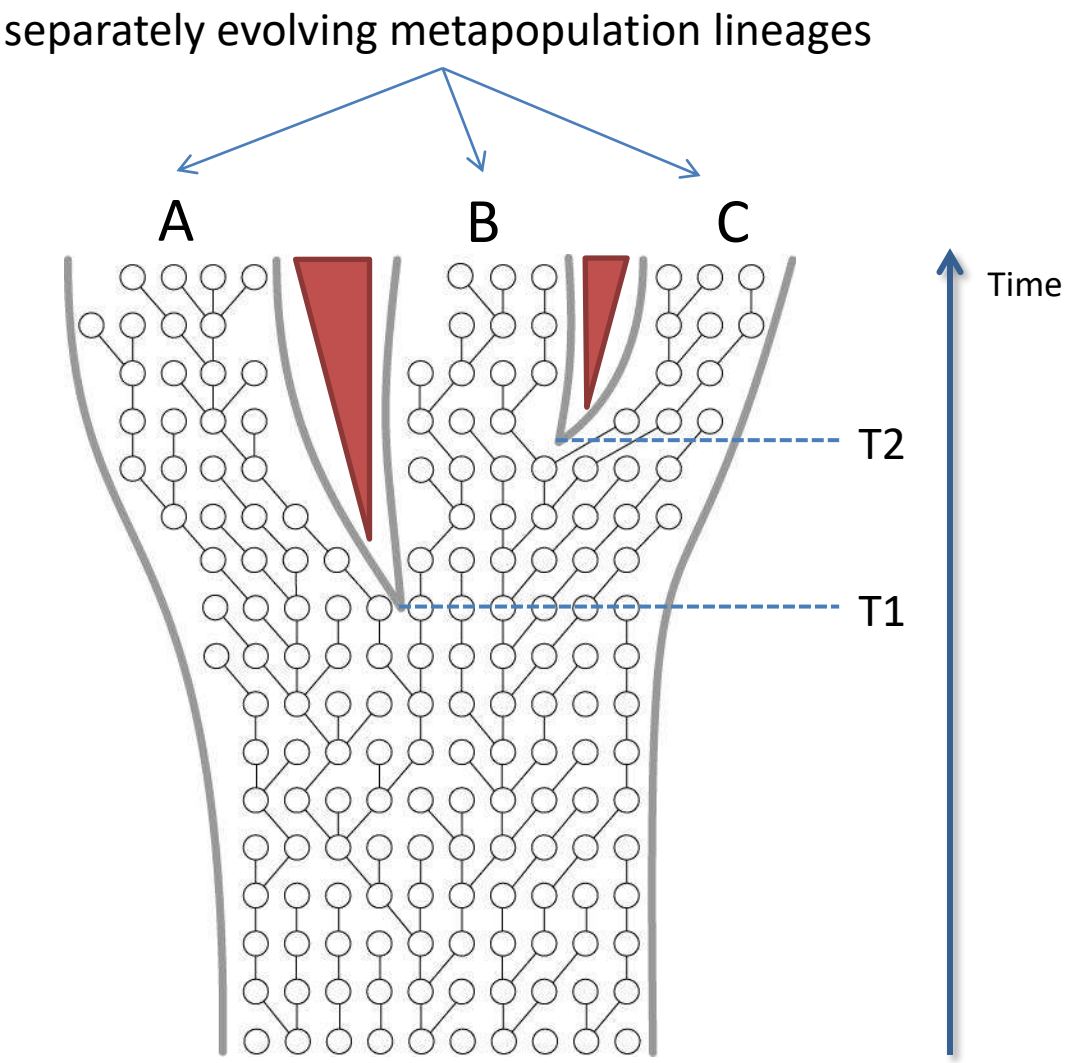
(Wright-Fisher process)

Population genetics:
coalescence process

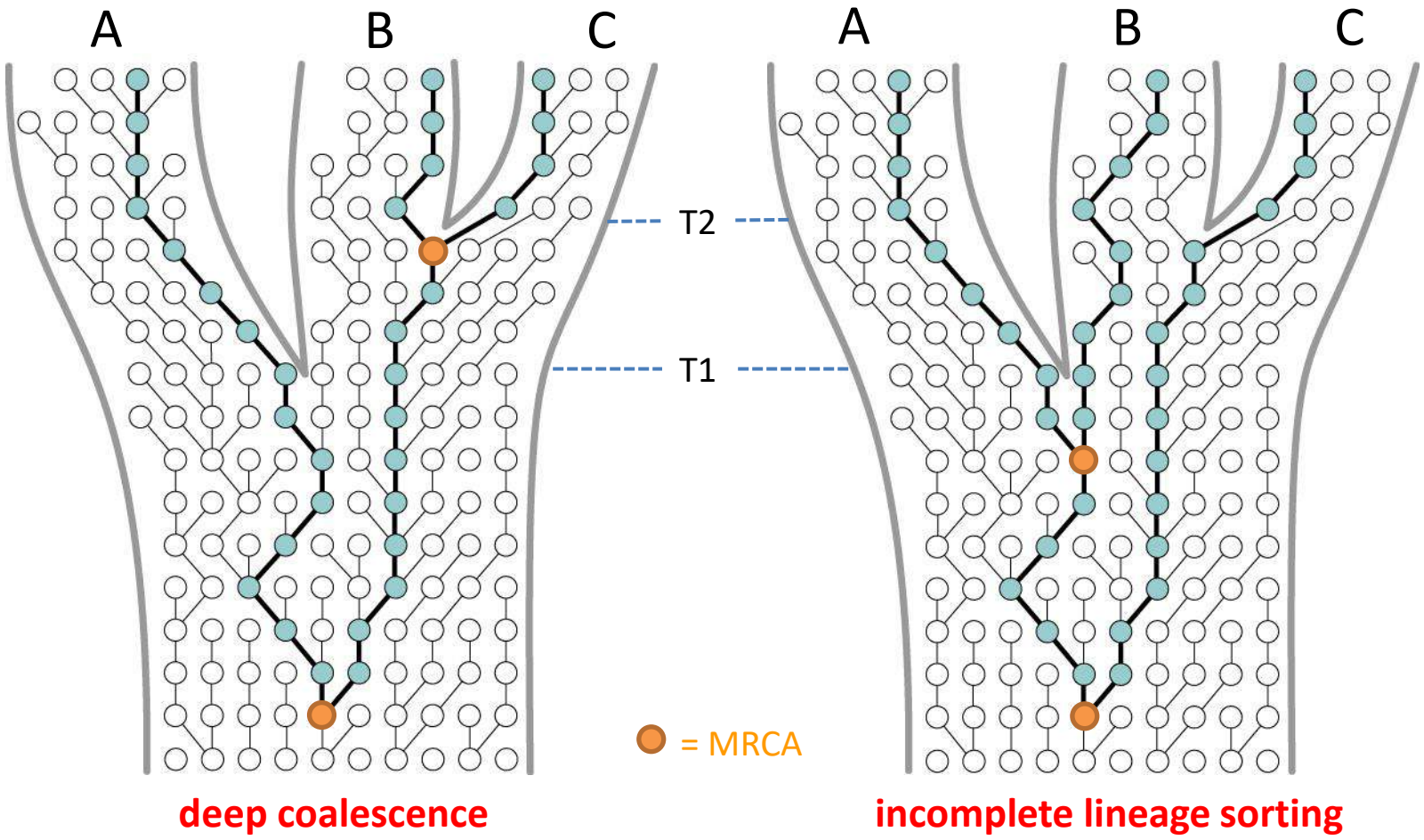
Phylogeny at the level of populations and species



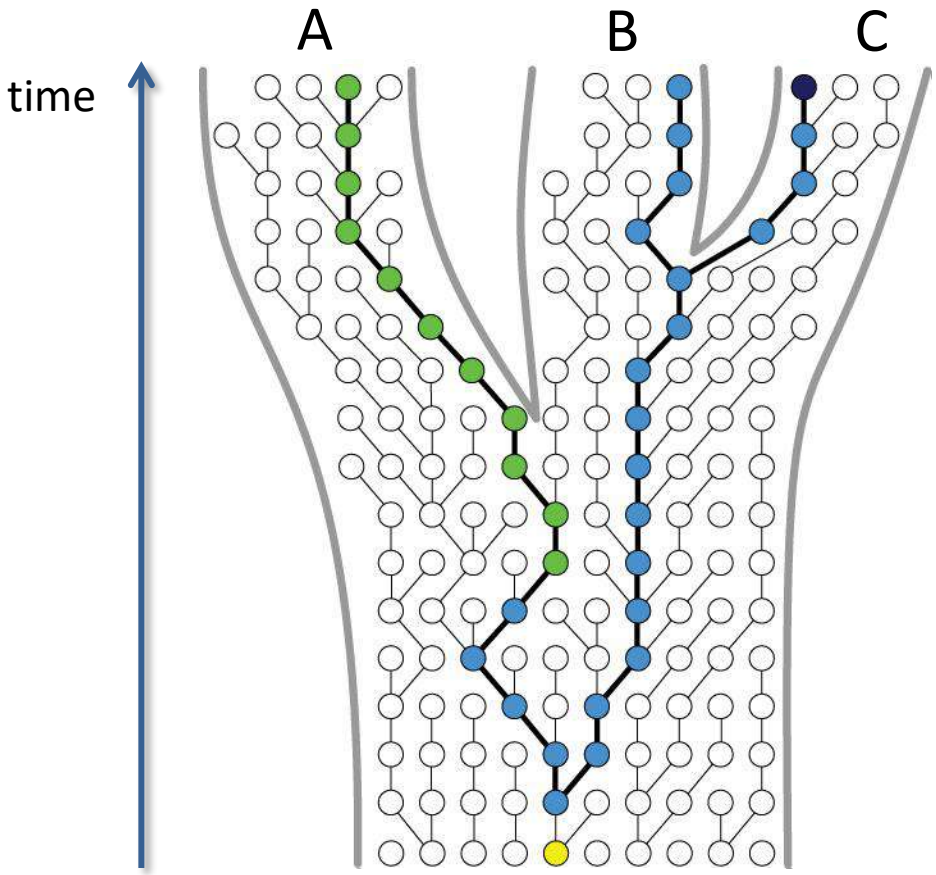
Phylogeny at the level of populations and species



Phylogeny at the level of populations and species

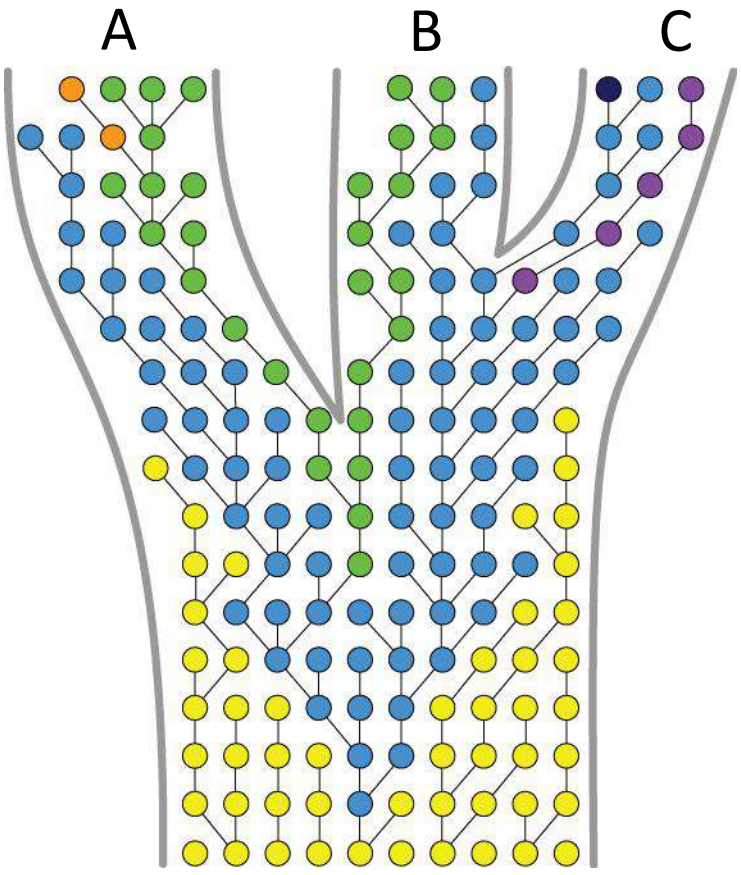


Phylogeny at the level of populations and species



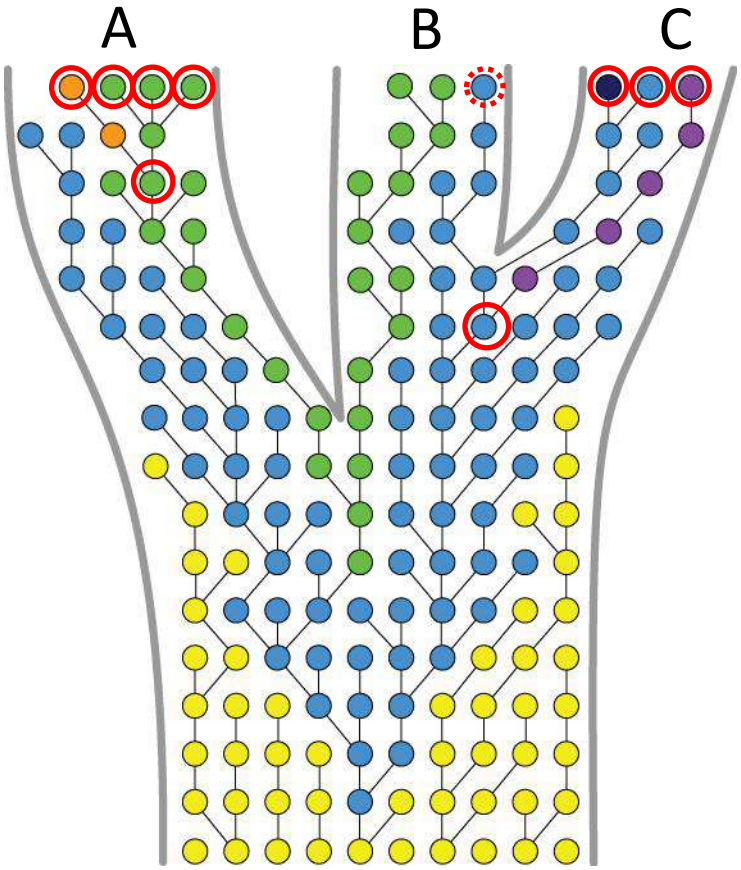
Phylogeny at the level of populations and species

Trans-species polymorphism

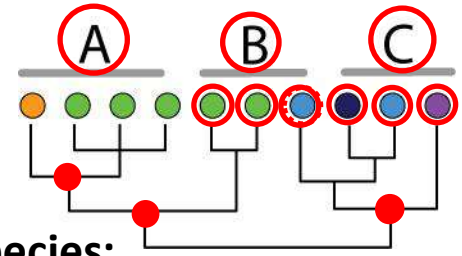


Phylogeny at the level of populations and species

Trans-species polymorphism



>>

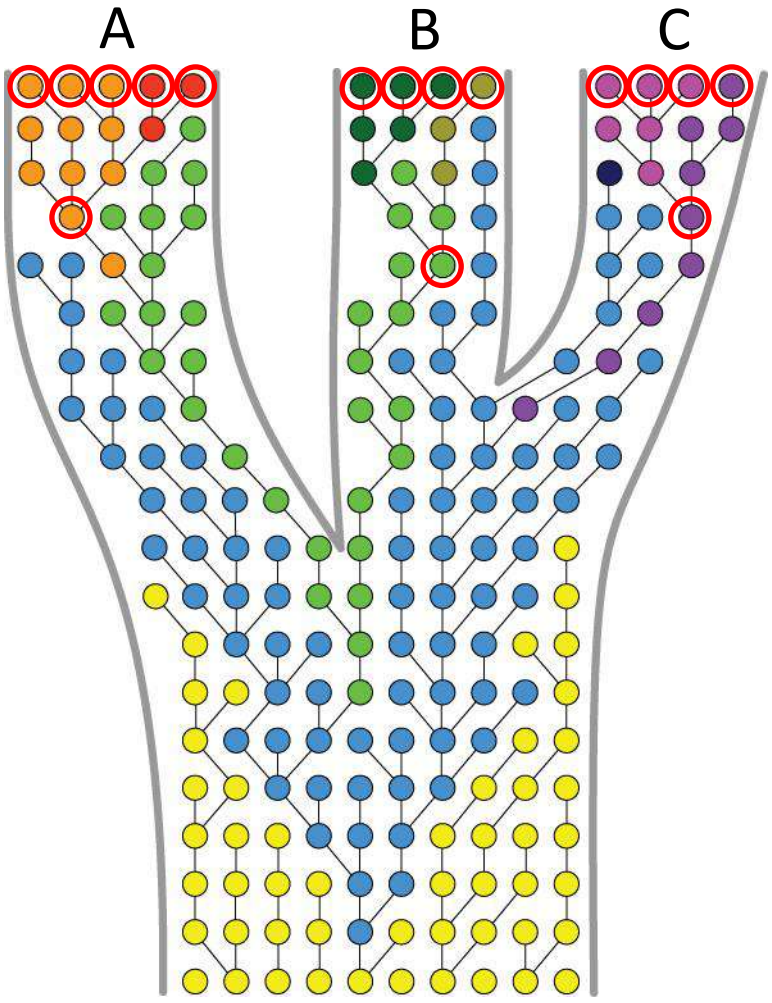


3 species:
Not reciprocally monophyletic

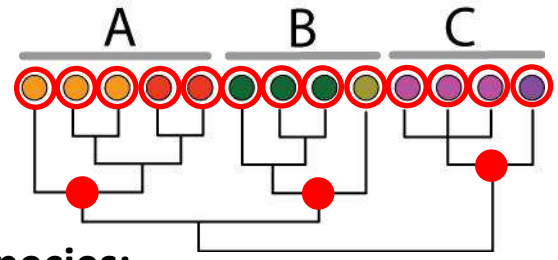
Phylogeny at the level of populations and species

Reciprocal monophyly

Fixation of alleles

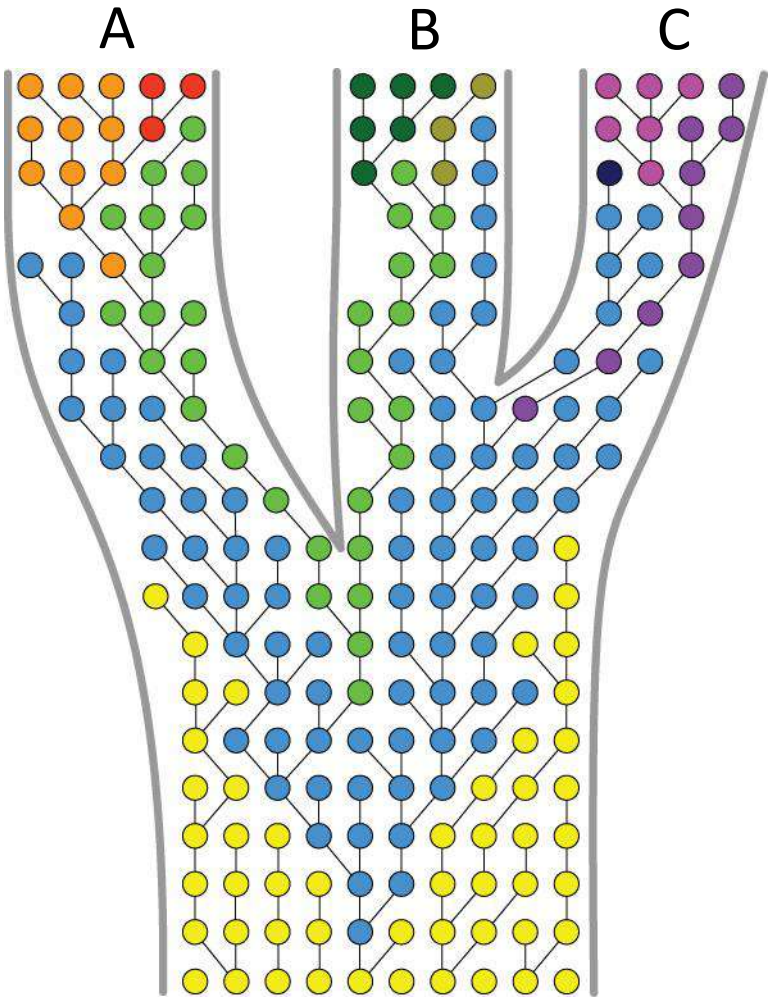


>>



3 species:
Reciprocally monophyletic

Phylogeny at the level of populations and species



- Gene genealogies below and above the species level are different in nature
- Young lineages reside within a zone where both processes meet

v
v

species delimitation

- population genetics
- phylogenetics

DNA-based species delimitation

Single locus coalescent-based approach

- General Mixed Yule Coalescent (GMYC) modeling

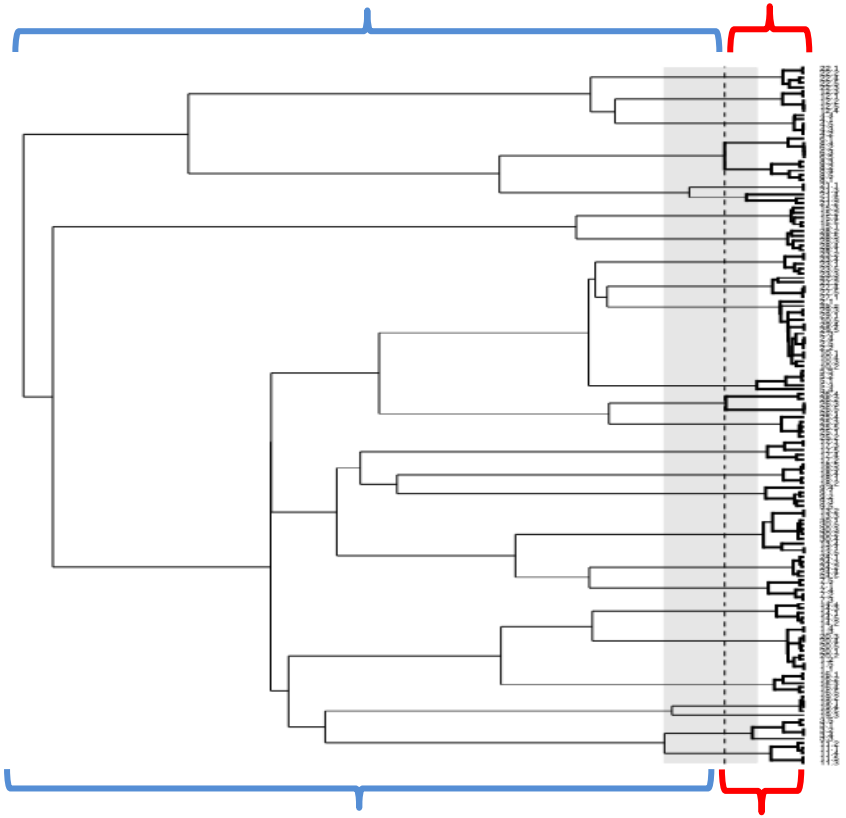
Multi locus coalescent-based approach

- Bayesian method (BP&P)

GMYC-model approach

- **Likelihood method**
- **Single locus DNA data**

Species-level (speciation & extinction) & **population-level (coalescence)** evolutionary processes



Different branching rates

[Pons, Barraclough et al. 2006, Monaghan et al. 2009]

GMYC-model approach

Raw data

Branching rates on ultrametric tree
(e.g. relaxed mol. clock tree)

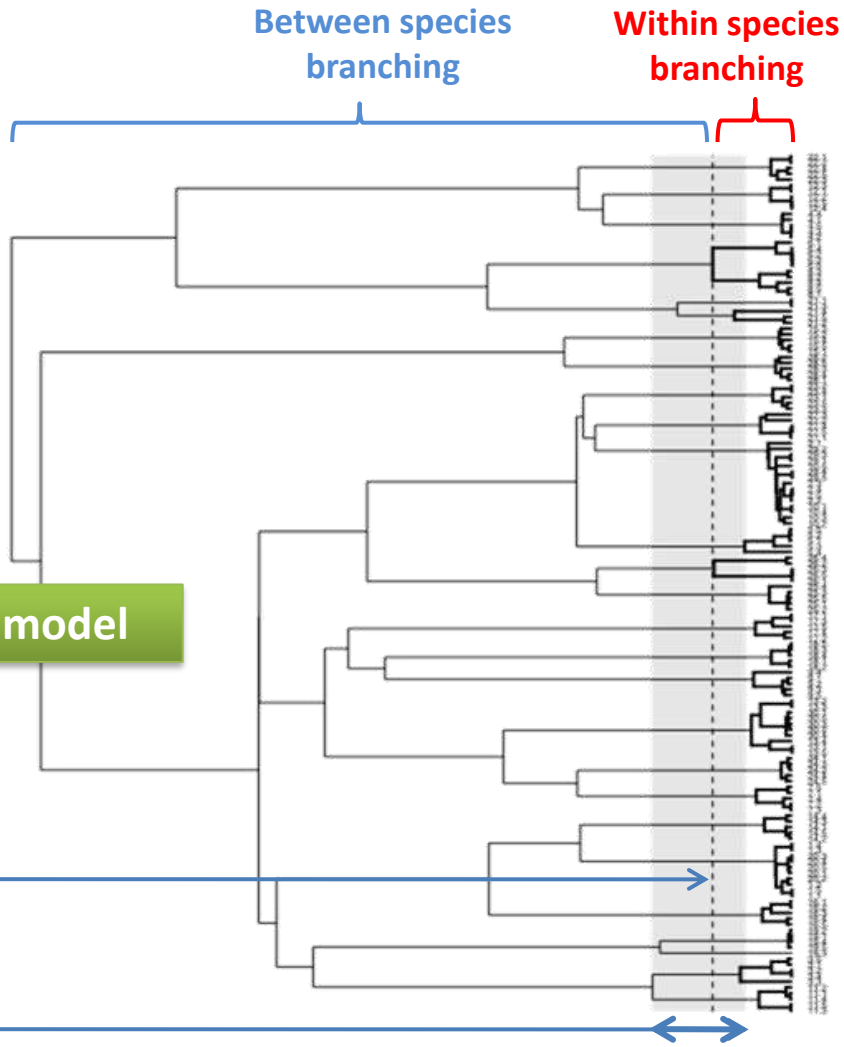
Statistical approach (LRT)

Null model: No separate lineages

Neutral coalescent model

Alternative model: Separately evolving lineages

General Mixed Yule Coalescent (GMYC) model



Location of switch from speciation to coalescent nodes

Statistical measure of confidence

Software: SPLITS package for R: http:// r-forge.r-project.org/projects/splits

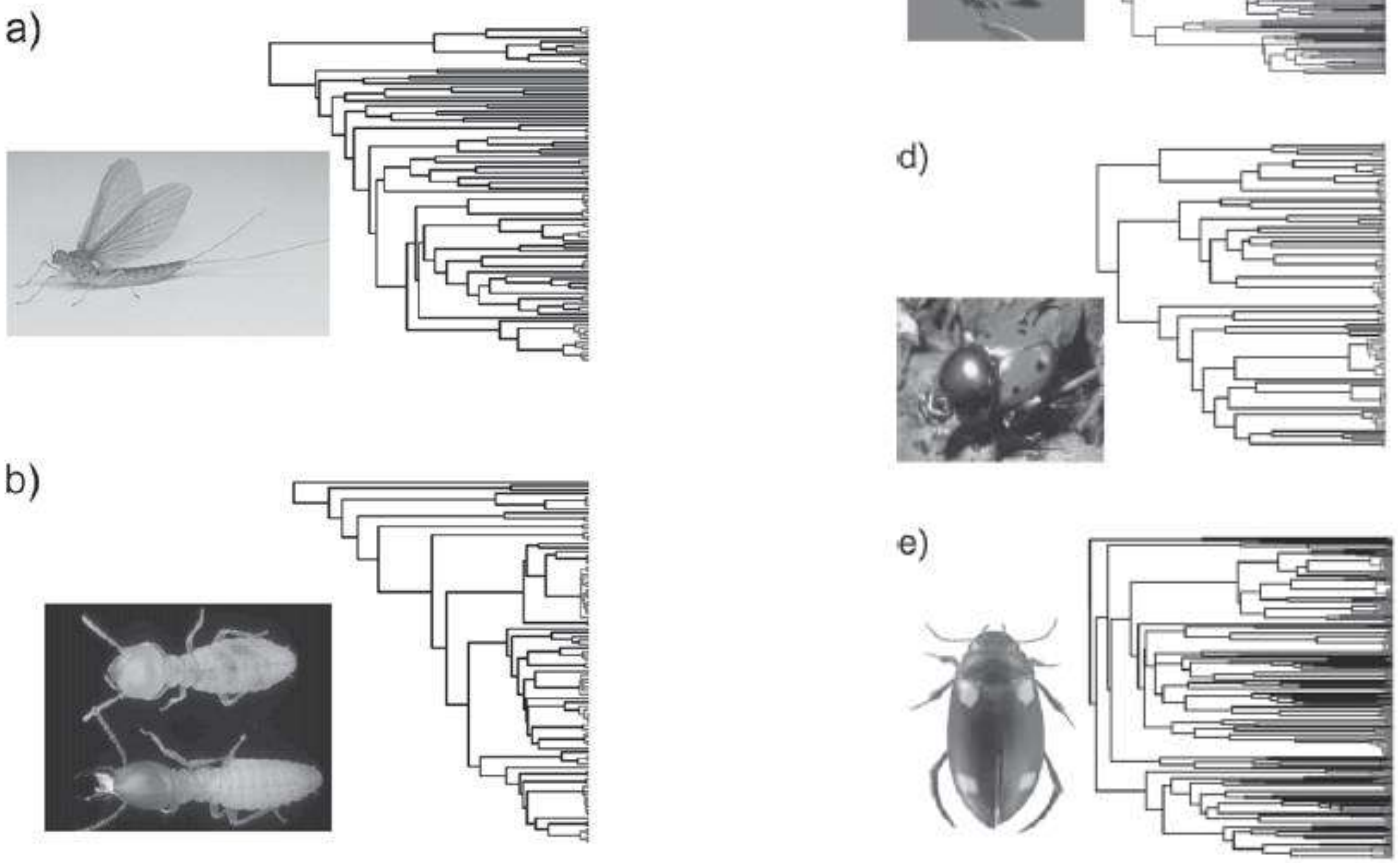
[Pons et al. 2006, Monaghan et al. 2009]

Syst. Biol. 58(3):298–311, 2009

DOI:10.1093/sysbio/syp027
Advance Access publication on July 1, 2009

Accelerated Species Inventory on Madagascar Using Coalescent-Based Models of Species Delineation

MICHAEL T. MONAGHAN[1,2,3,*], RUTH WILD[1,2], MIRANDA ELLIOT[1,2], TOMOC
MICHAEL BALKE[1,4], DAEGAN J.G. INWARD[1], DAVID C. LEES[1], RAVO RA
PAUL EGGLETON[1], TIMOTHY G. BARRACLOUGH[2], AND ALFRIED P. V

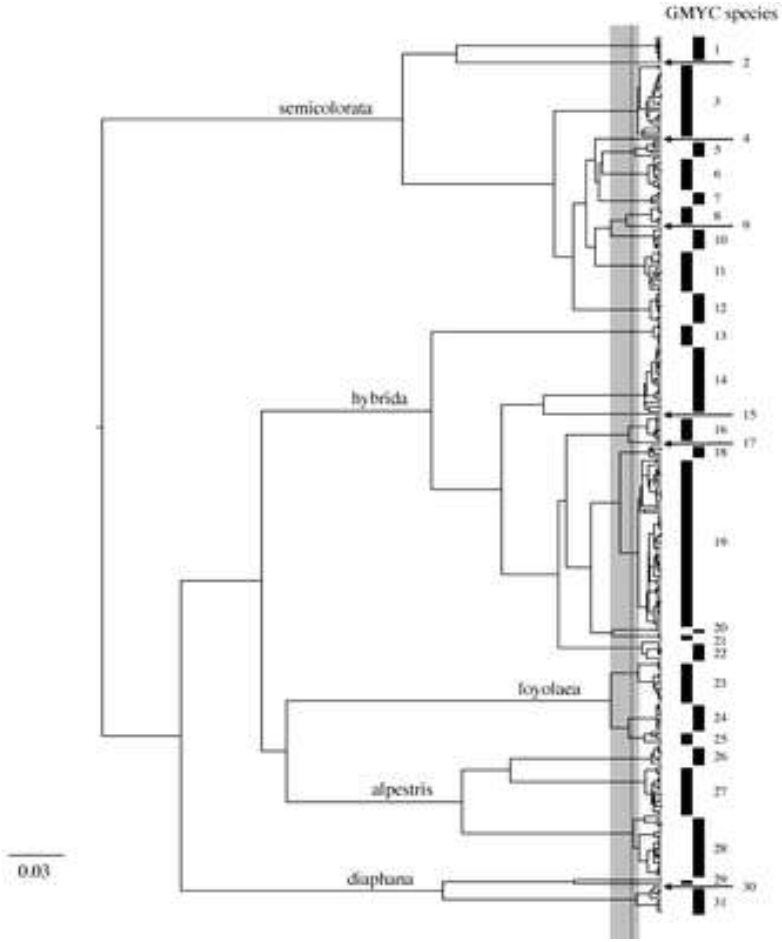


OPEN ACCESS Freely available online

PLoS one

Toward a DNA Taxonomy of Alpine *Rhithrogena* (Ephemeroptera: Heptageniidae) Using a Mixed Yule-Coalescent Analysis of Mitochondrial and Nuclear DNA

Laurent Vuataz[1,2]*, Michel Sartori[1], André Wagner[1], Michael T. Monaghan[3]



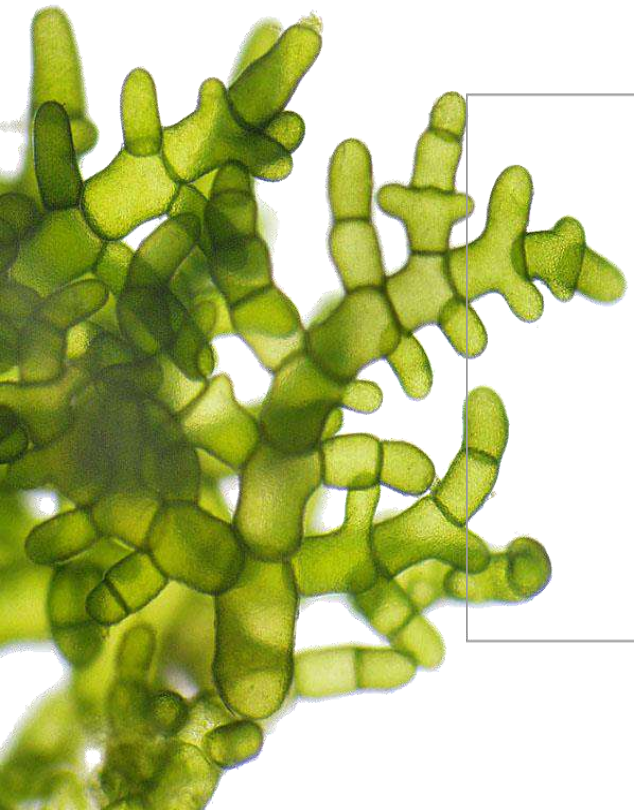
Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

DNA taxonomy in morphologically plastic taxa: Algorithmic species delimitation in the *Boodlea* complex (Chlorophyta: Cladophorales)

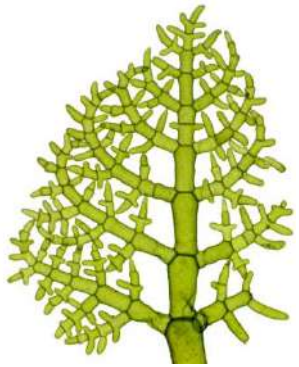
Frederik Leliaert [a,*], Heroen Verbruggen [a], Brian Wysor [b], Olivier De Clerck [a]



species 1
species 2
species 3
species 4
species 5
species 6
species 7
species 8
species 9
species 10
species 11
species 12
species 13
species 14
species 15
species 16
species 17
species 18
species 19

"Boodlea composita"

"Boodlea composita"



"Phyllodictyon anastomosans"

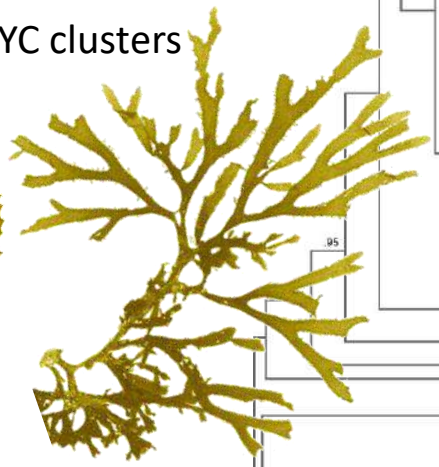
"Cladophoropsis vaucheriiformis"

Dictyota

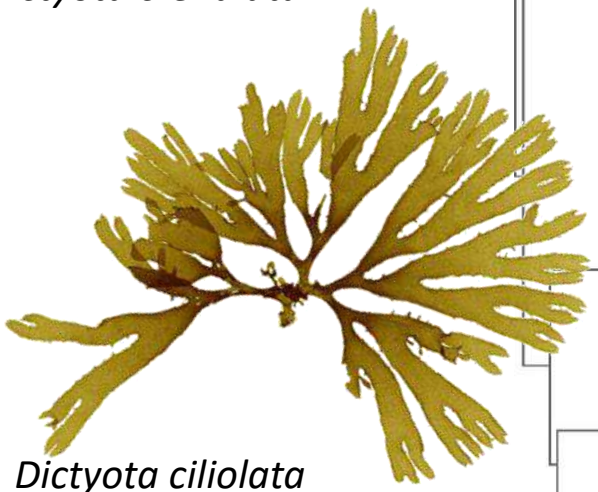
184 *psb*A sequences

$\log L_{GMYC} = 1672 > \log L_{nul} = 1649$,
$p < 0.001$

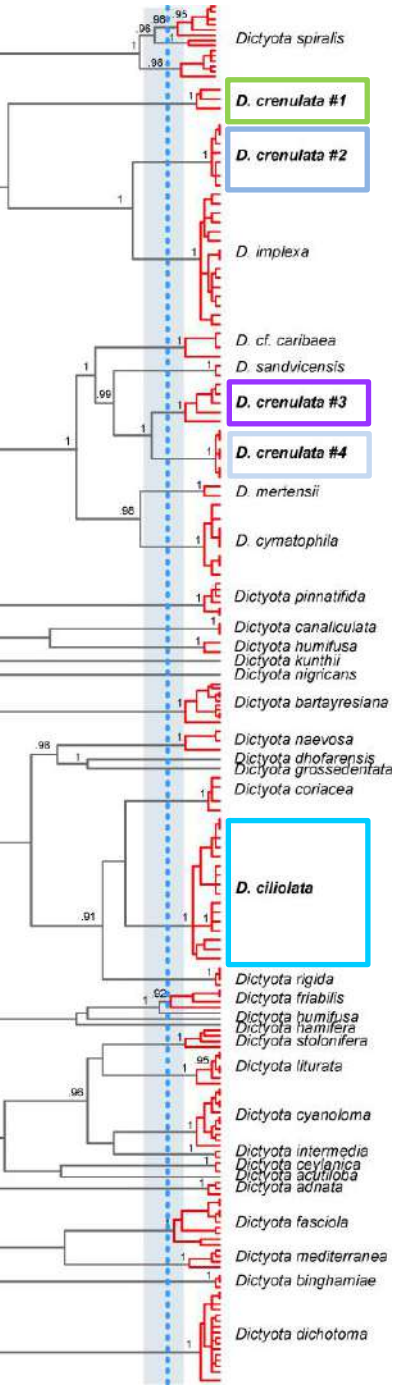
39 (34-43) GMYC clusters

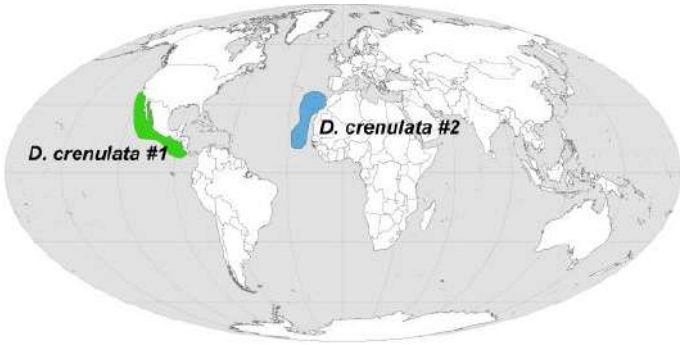


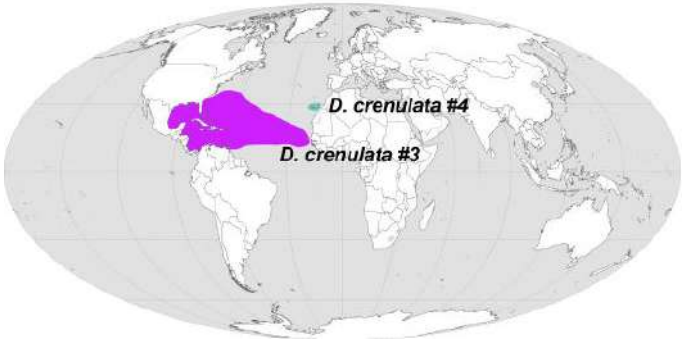
Dictyota crenulata

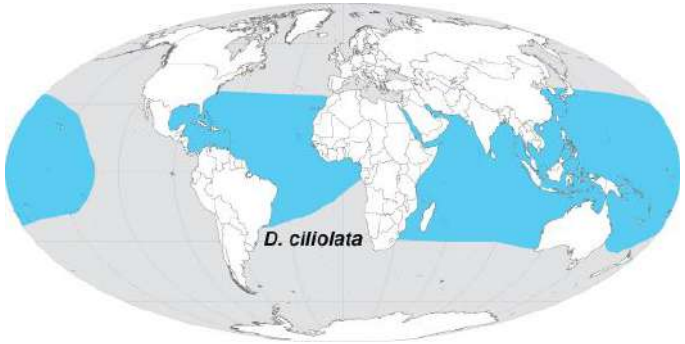


Dictyota ciliolata



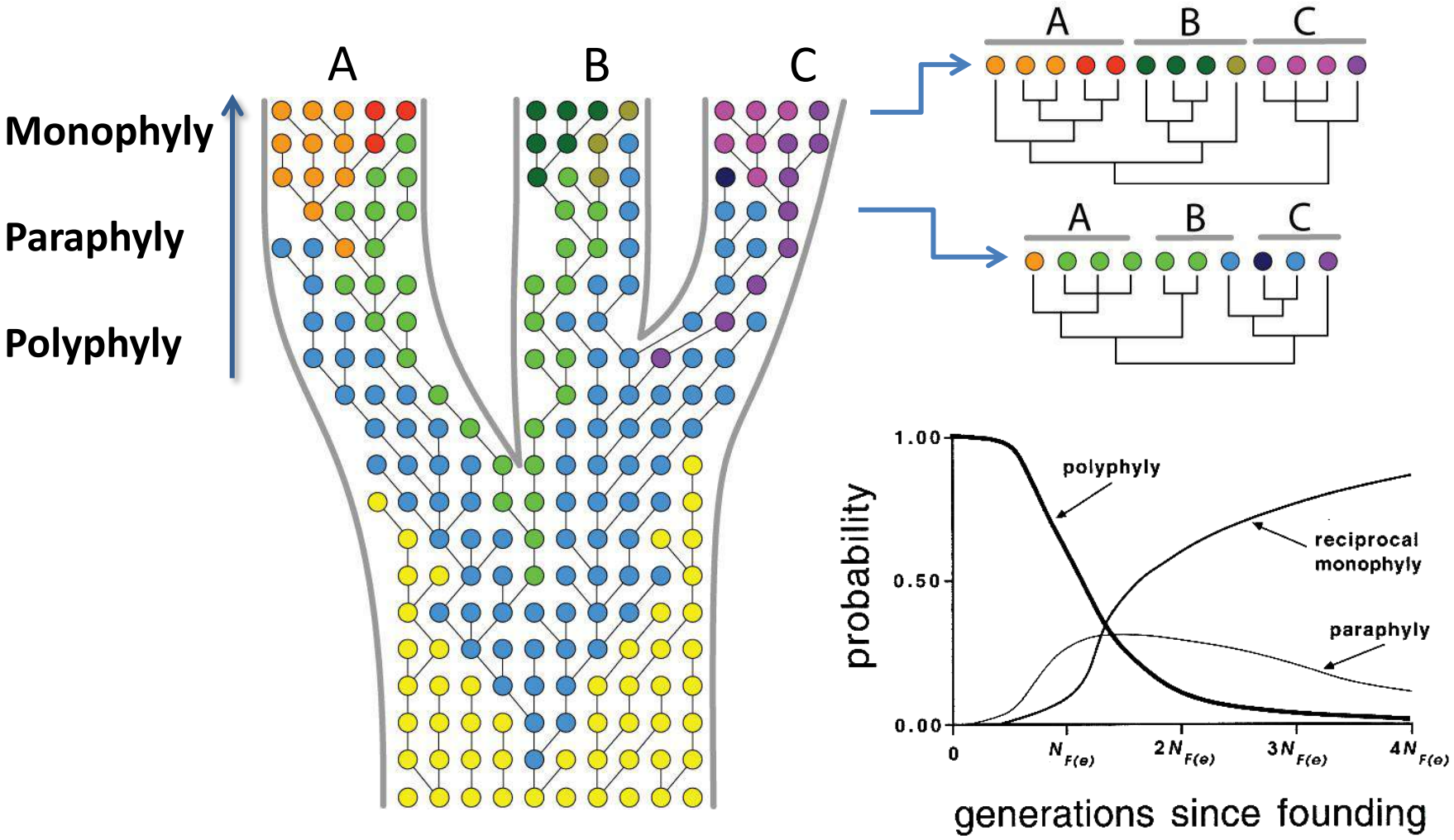






[Tronholm et al. unpublished]

Multi-locus methods



[Avise 2000]

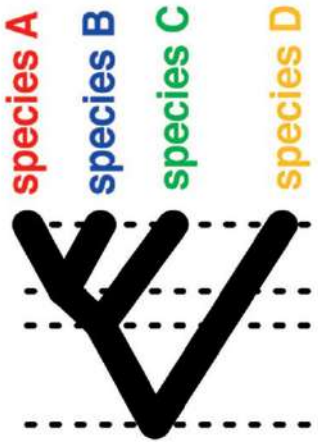
Multi-locus methods

species tree – gene tree discordance

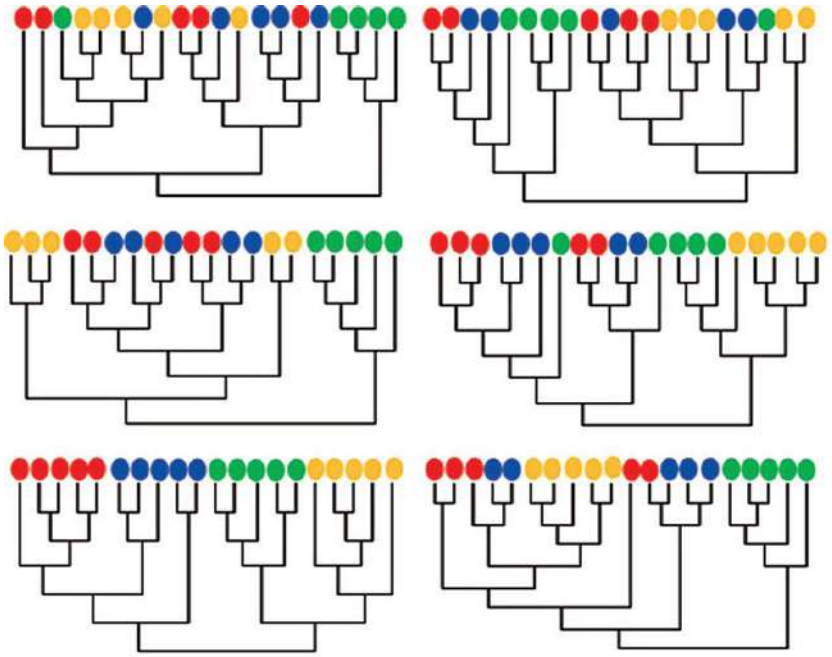
Early stages of speciation

- Retention of ancestral polymorphism
- Incomplete lineage sorting

→ different loci have their own gene trees that do not necessarily match the species tree



Species tree



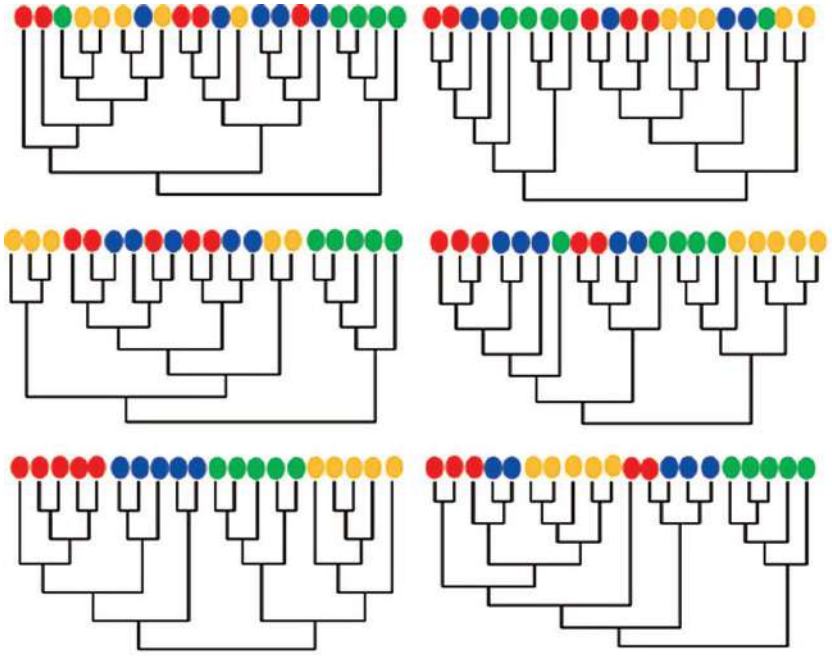
Gene trees of 6 different loci

[Knowles & Carstens 2007]

Multi-locus methods

- New species delimitation methods
  - **Multi-locus** data
  - Models combining
    - **Species phylogenies**
    - **Coalescent processes**

Despite the lack of monophyletic species, a **signal of species divergence persists in gene trees of unlinked loci**

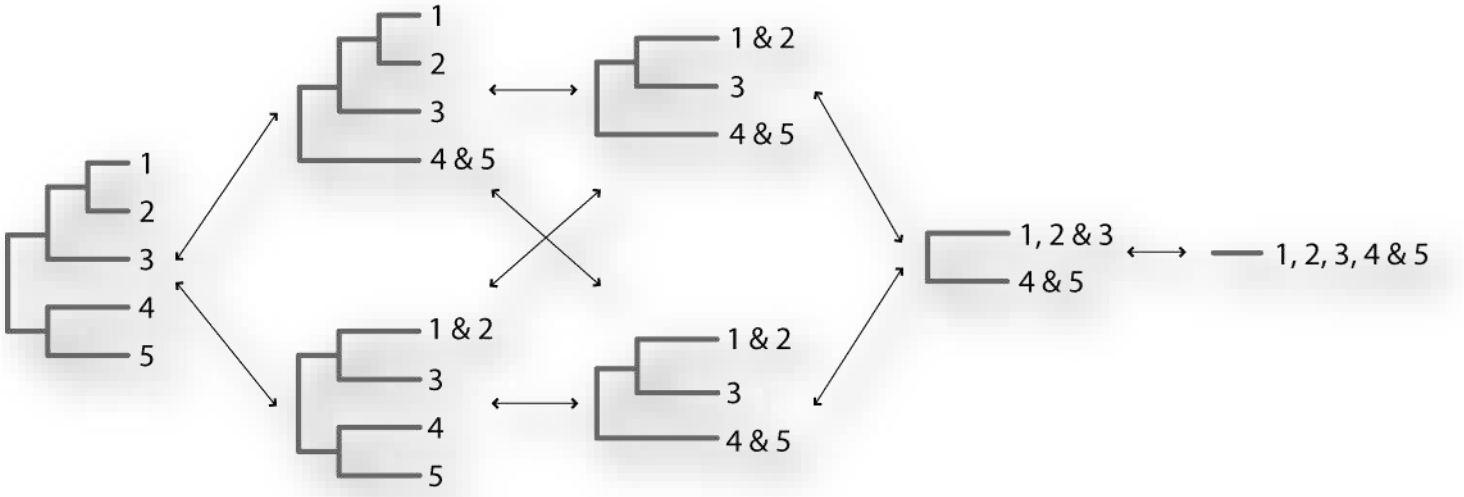


- Software
  - BP&P [Yang & Rannala 2010]: http://abacus.gene.ucl.ac.uk/
  - BROWNIE [O'Meara 2010]: http://www.brianomeara.info/brownie
  - SpedeSTEM [Ence & Carstens 2011]
  - Species delimitation – Geneious plugin [Masters, Fan & Ross 2010] http://www.biomatters.com

Multi-locus Bayesian species delimitation

Bayesian method (BP&P)

- Multilocus sequence data
- User-specified species tree
- Prior information:
  - population size (Θ)
  - divergence times (τ_0)
- Reversible-jump Markov chain Monte Carlo (rjMCMC): estimates posterior distribution for species delimitation models



[Yang & Rannala 2010]

PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2010) **277**, 3071–3077
doi:10.1098/rspb.2010.0662
Published online 2 June 2010

Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*)

Adam D. Leaché[1,2,*] **and Matthew K. Fujita**[1,3]

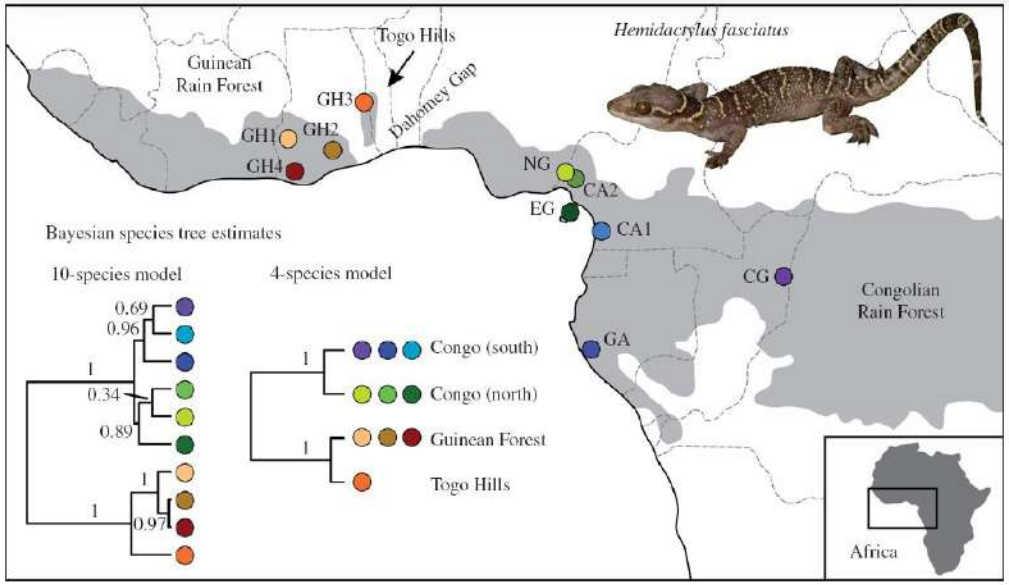


Figure 1. Populations of *Hemidactylus fasciatus* included in the study. The generalized distributions of the major rainforest fragments are shown in grey, which also approximate the distribution of *H. fasciatus*. Bayesian species trees are inferred using *BEAST, and numbers on nodes are posterior probability values.

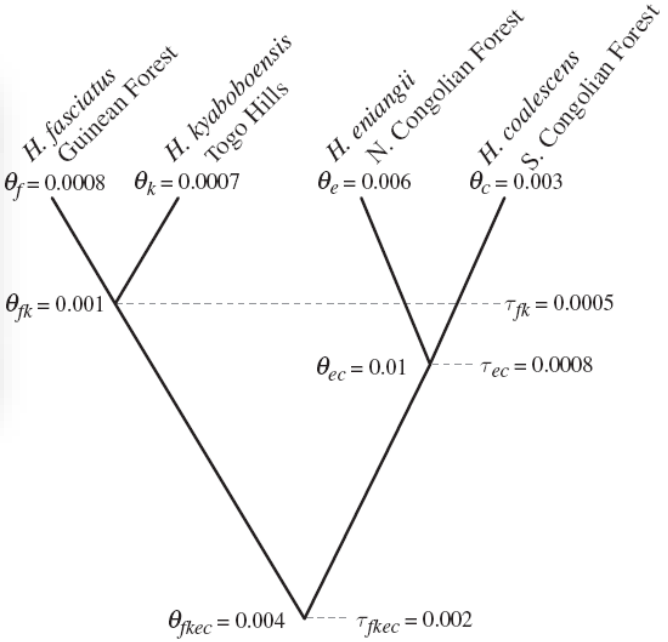


Figure 3. Bayesian species delimitation supports 4 species within *Hemidactylus fasciatus*. The posterior estimates (mean of the distribution) for θ and τ are provided on the species tree. The 10 species guide tree supports 5 species, with the extra species representing the population on Bioko Island (part of *H. eniangii* under the model proposed here).

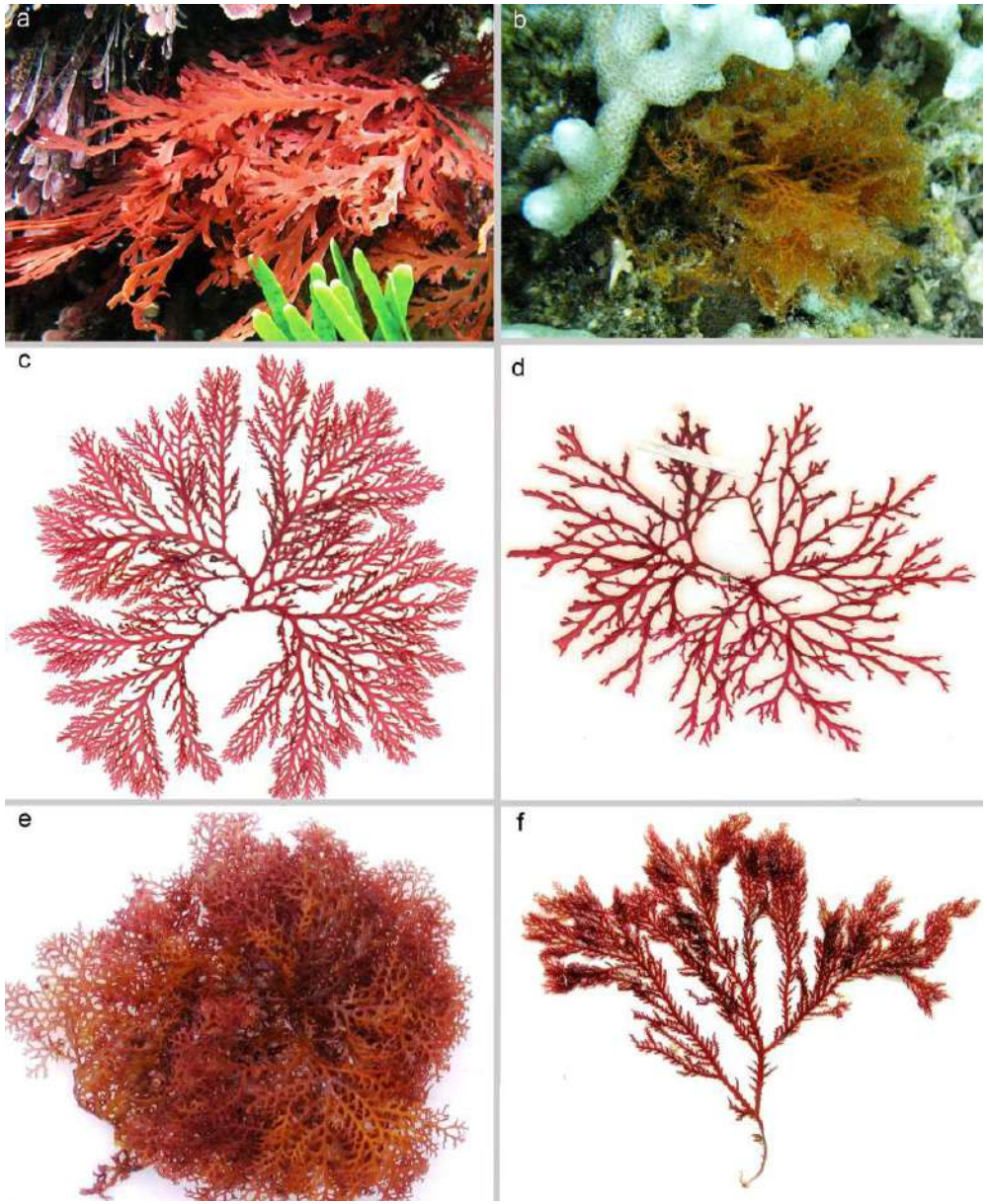
(i) Hemidactylus coalescens *sp. nov.*

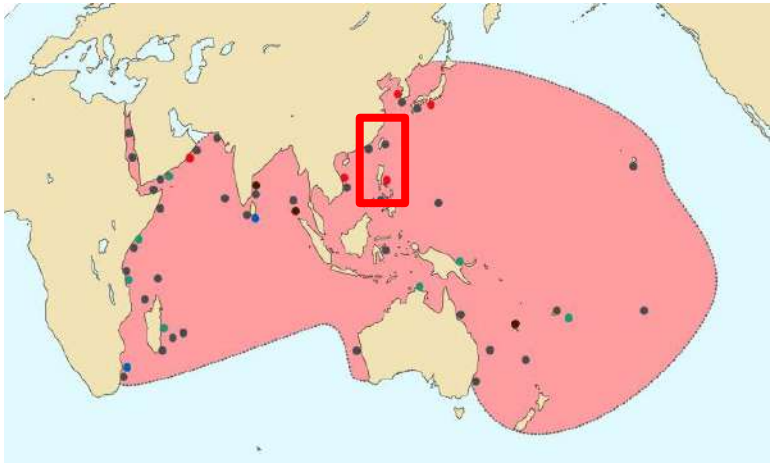
Holotype. Zoologisches Forschungsinstitut und Museum Alexander Koenig (ZFMK) 87680, adult male; Cameroon, Campo Region, Nkoelon, 2.3972° N, 10.04515° E, 85 m; collected by Michael F. Barej and Julia Wurstner, 27 October 2007. Paratype = ZFMK 87679.

Diagnosis. This species includes all populations that cluster with those from the southern portion of the Congolian rainforest included in this study (southern Cameroon, Gabon and Congo), with strong support in the Bayesian species delimitation model.

Etymology. This species is named after the coalescent process used to delimit the species.

Portieria





P. hornemannii
P. spinulosa
P. japonica
P. tripinnata
P. harveyii

Wiseman (1973):
P. hornemannii

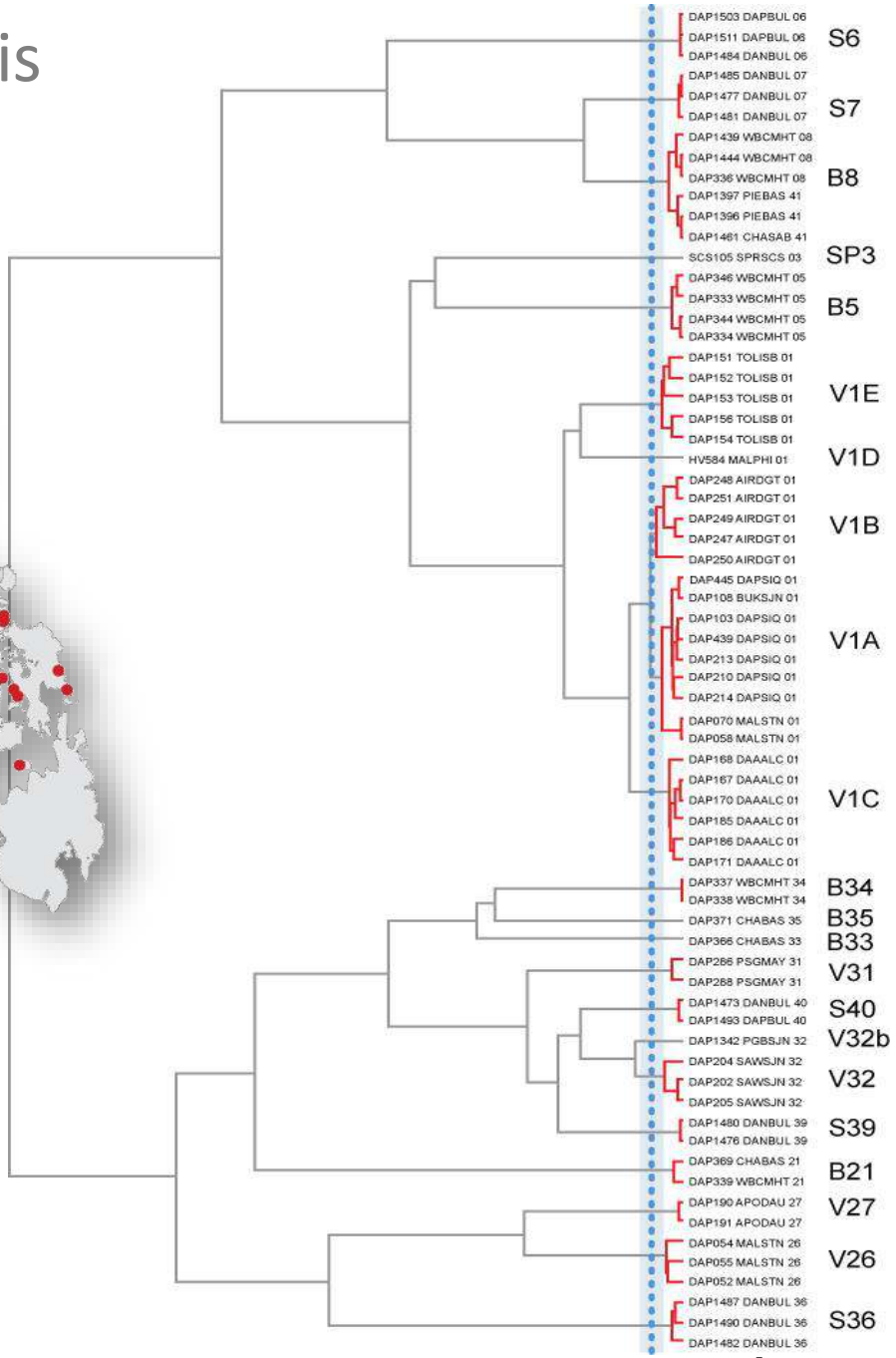
[Payo et al. unpublished]

Portieria: GMYC analysis

Philippines: 265 *cox* 2-3 spacer sequences

$\log L_{GMYC} = 462.92 > \log L_{nul} = 449.25$, $p < 0.001$

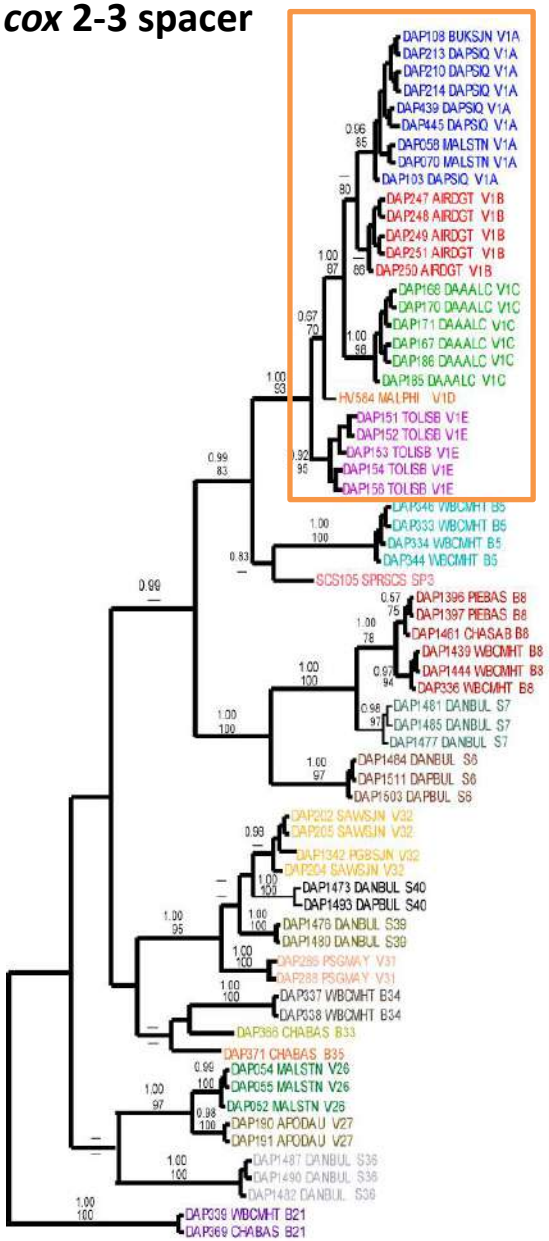
21 (19-24) **GMYC clusters**



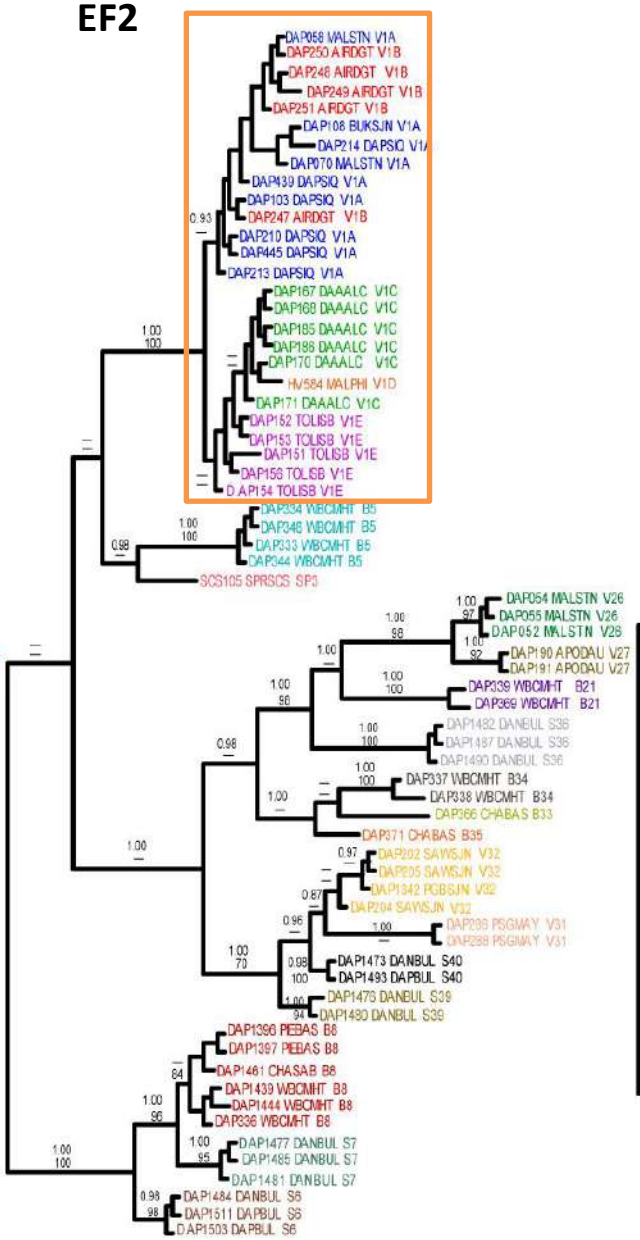
[Payo et al. unpublished]

Portieria: gene trees

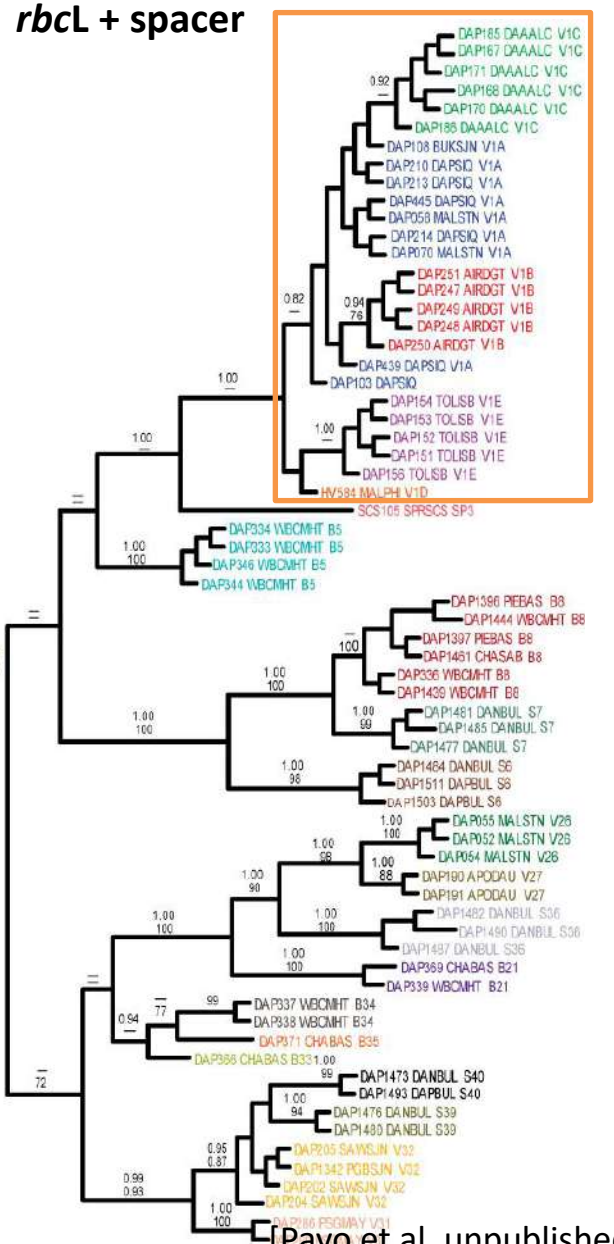
cox 2-3 spacer



EF2

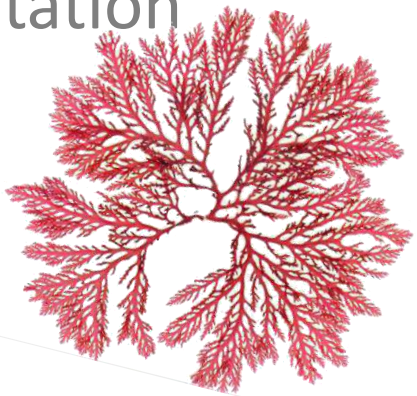


***rbc*L + spacer**



[Payo et al. unpublished]

Portieria: Multi-locus Bayesian species delimitation

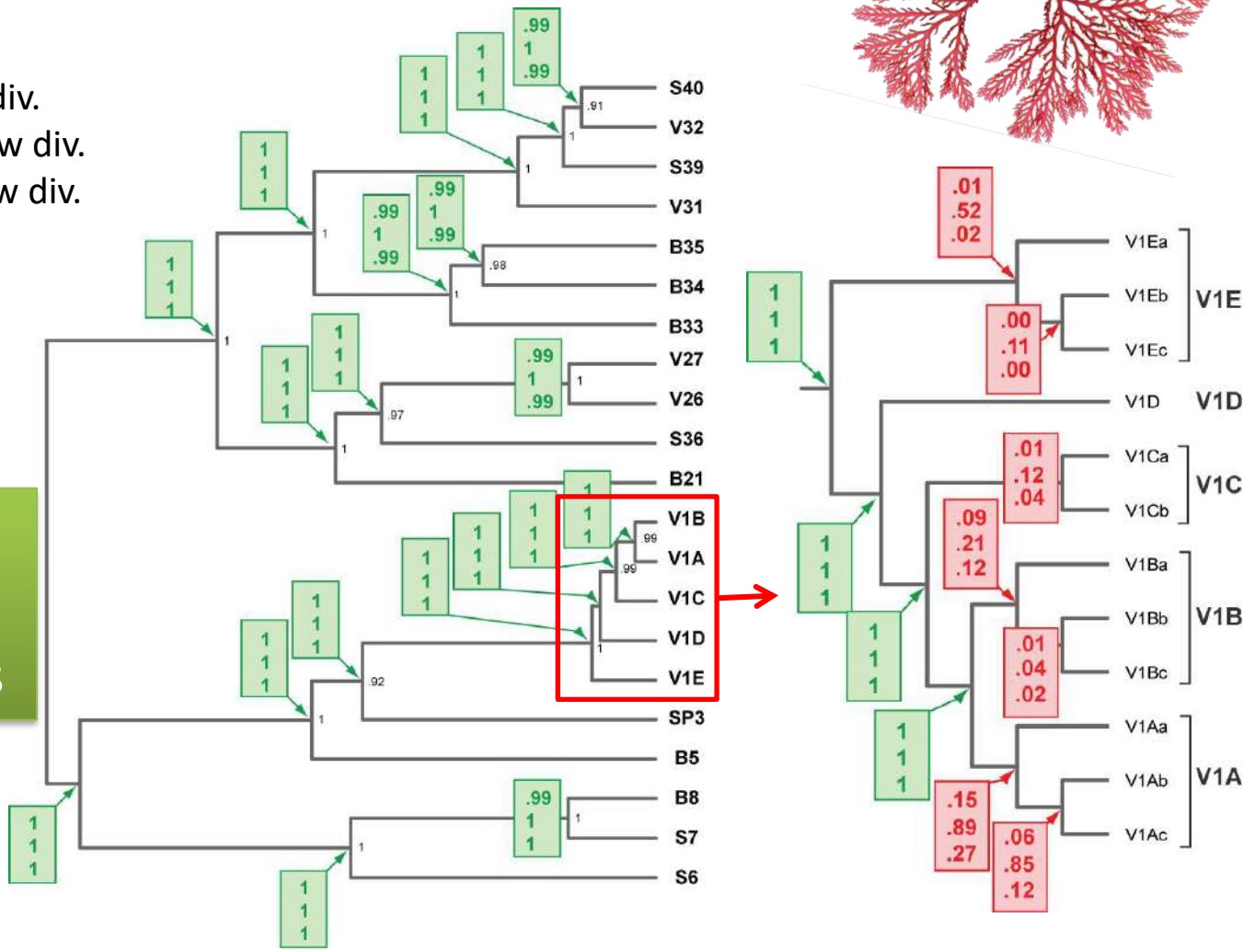


Species tree: *BEAST (cox / EF / rbcL)

Bayesian species delimitation: BP&P

3 prior combinations
- large pop size / deep div.
- small pop size / shallow div.
- large pop size / shallow div.

21 cryptic *Portieria* species in the Philippines



[Payo et al. unpublished]

Conclusions

Species delimitation

- Population genetic methods
- Phylogenetic methods

Probabilistic tests for species boundaries

- Statistical support
- Level of uncertainty

DNA barcoding

- Fast coalescence of mt and cp markers
- Effective in detecting species boundaries
- Useful for large scale analyses of species diversity

Further reading

Avise JC, Wollenberg K. 1997. **Phylogenetics and the origin of species**. Proceedings of the National Academy of Sciences. 94:7748-7755.

Sites JW, Marshall JC. 2003. **Delimiting species: a Renaissance issue in systematic biology**. Trends in Ecology & Evolution. 18:462-470.

De Queiroz K. 2007. **Species concepts and species delimitation**. Systematic Biology. 56:879-886

Knowles LL, Carstens BC. 2007. **Delimiting Species without Monophyletic Gene Trees**. Systematic Biology. 56:887-895.

Degnan JH, Rosenberg NA. 2009. **Gene tree discordance, phylogenetic inference and the multispecies coalescent**. Trends in Ecology & Evolution. 24:332-340.

Kuhner MK. 2009. **Coalescent genealogy samplers: windows into population history**. Trends in Ecology & Evolution. 24:86-93.

Yang Z, Rannala B. 2010. **Bayesian species delimitation using multilocus sequence data**. Proceedings of the National Academy of Sciences. 107:9264-9269.

Leaché AD, Fujita MK. 2010. **Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*)**. *Proceedings of the Royal Society B: Biological Sciences. 277:3071-3077.*