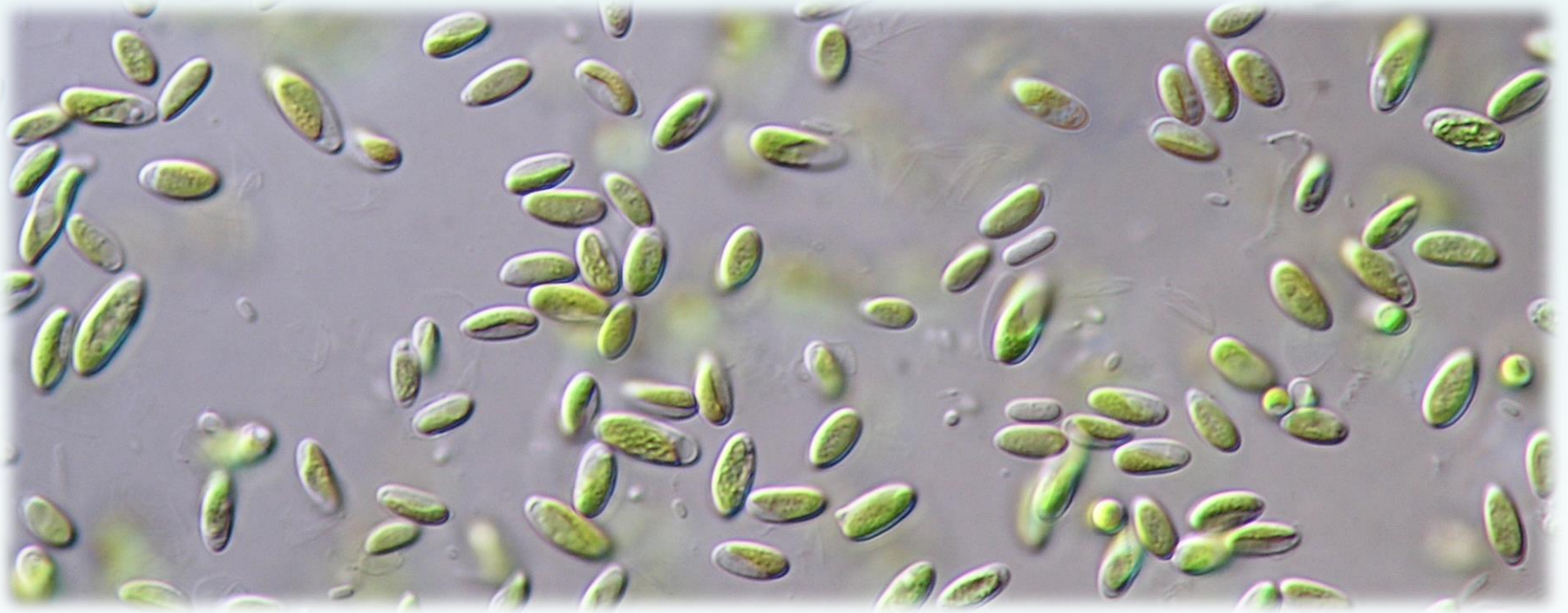
DNA-based taxonomy in ecologically versatile microalgae:

A re-evaluation of the species concept within the coccoid green algal genus *Coccomyxa*



Pavel Škaloud, Veronica Malavasi, Fabio Rindi, Sabrina Tempesta, Michela Paoletti, Marcella Pasqualetti

FACULTY OF SCIENCE
Charles University

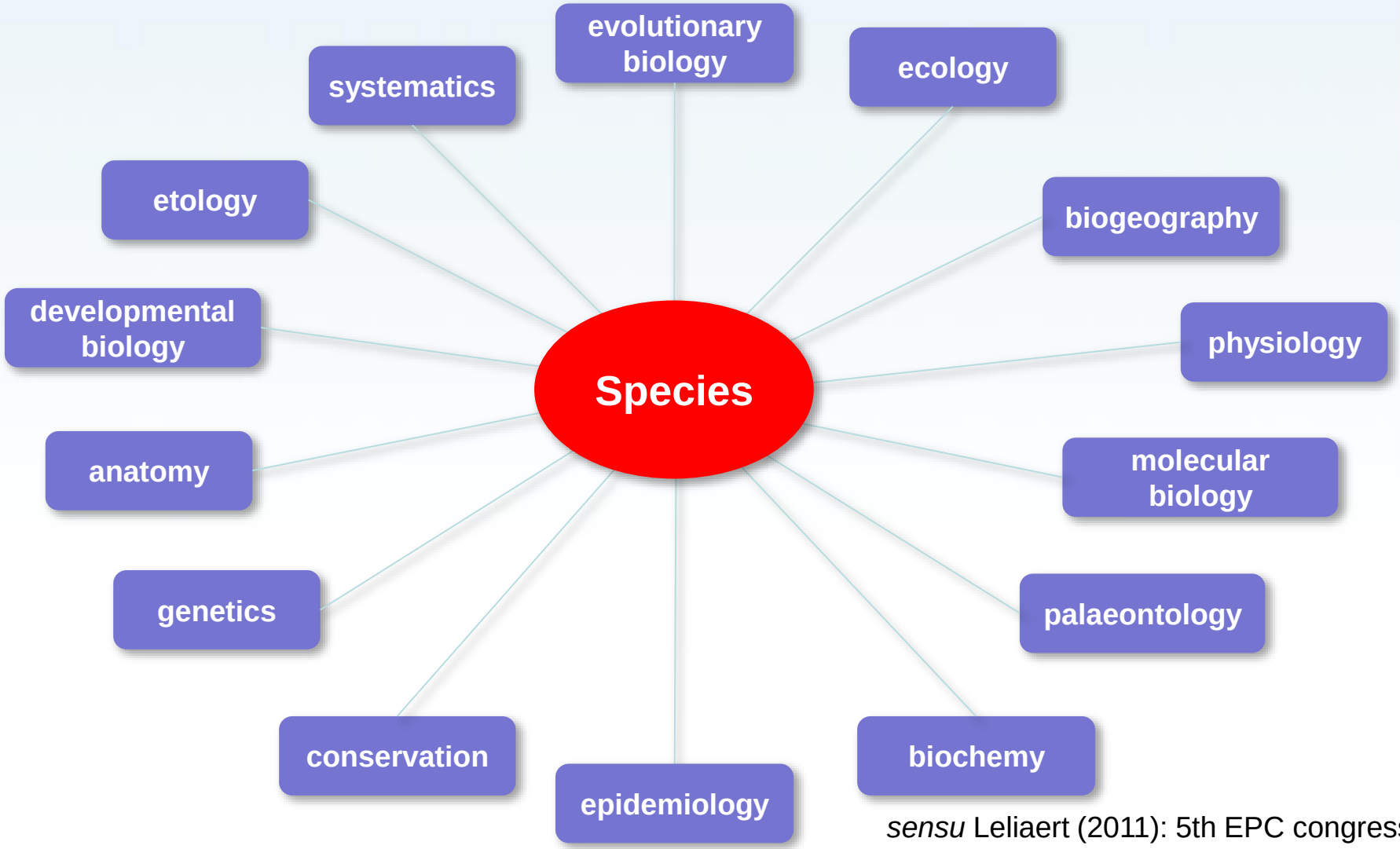
Algal speciation & evolution lab

Species

- Fundamental units of the systematics
  - Organisation of biodiversity to the well arranged system

Species

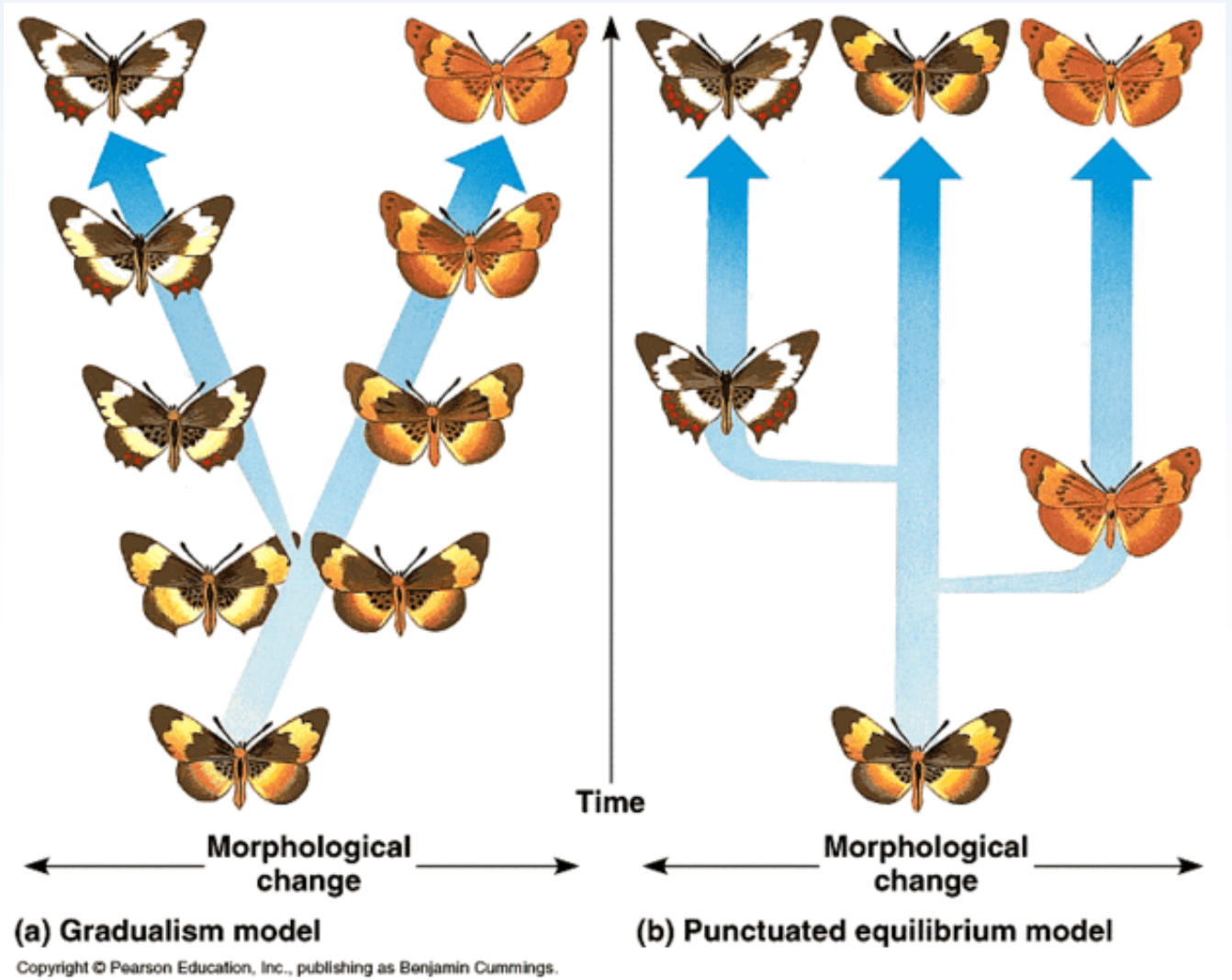
- Fundamental units in all biological disciplines
- It is of a great importance to delimit the species correctly



sensu Leliaert (2011): 5th EPC congress

Species delimitation

- Difficulty of simple species definitions
  - Species are not rigid, but evolving entities



Species concepts in protists

Biological species concept

Evolutionary significant unit

Cohesion species concept

Phylogenetic species concept

Polythetic species concept

Cladistic species concept

Recognition species concept

Genealogical concordance concept

Internodal species concept

Reproductive competition concept

Hennigian Species concept

Genetic species concept

Composite species concepts

Evolutionary species concept

Palaeospecies concept

Ecological species concept

Morphological species concept

Successional species concept

Non-dimensional species concept

Taxonomic species concept

Linnean species concept

Genotypic cluster definition

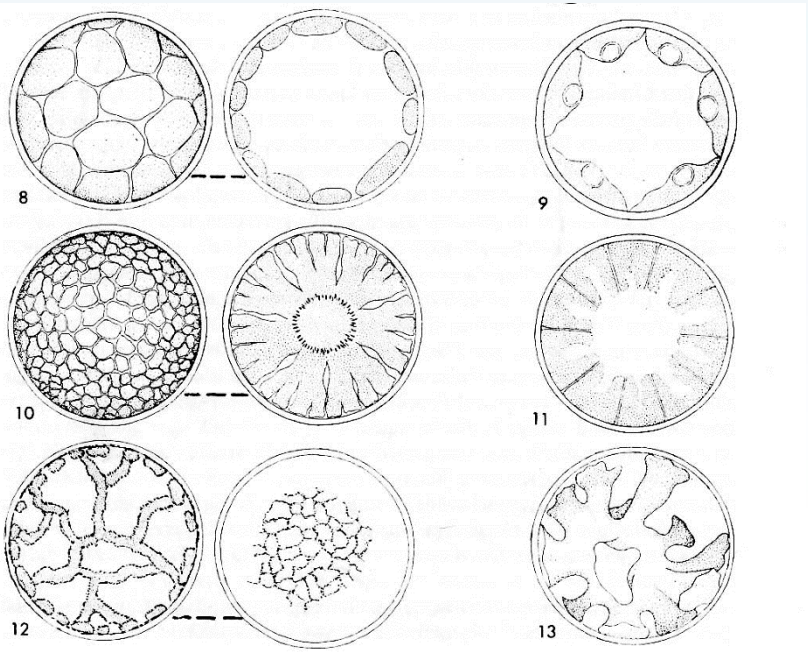
Phenetic species concept

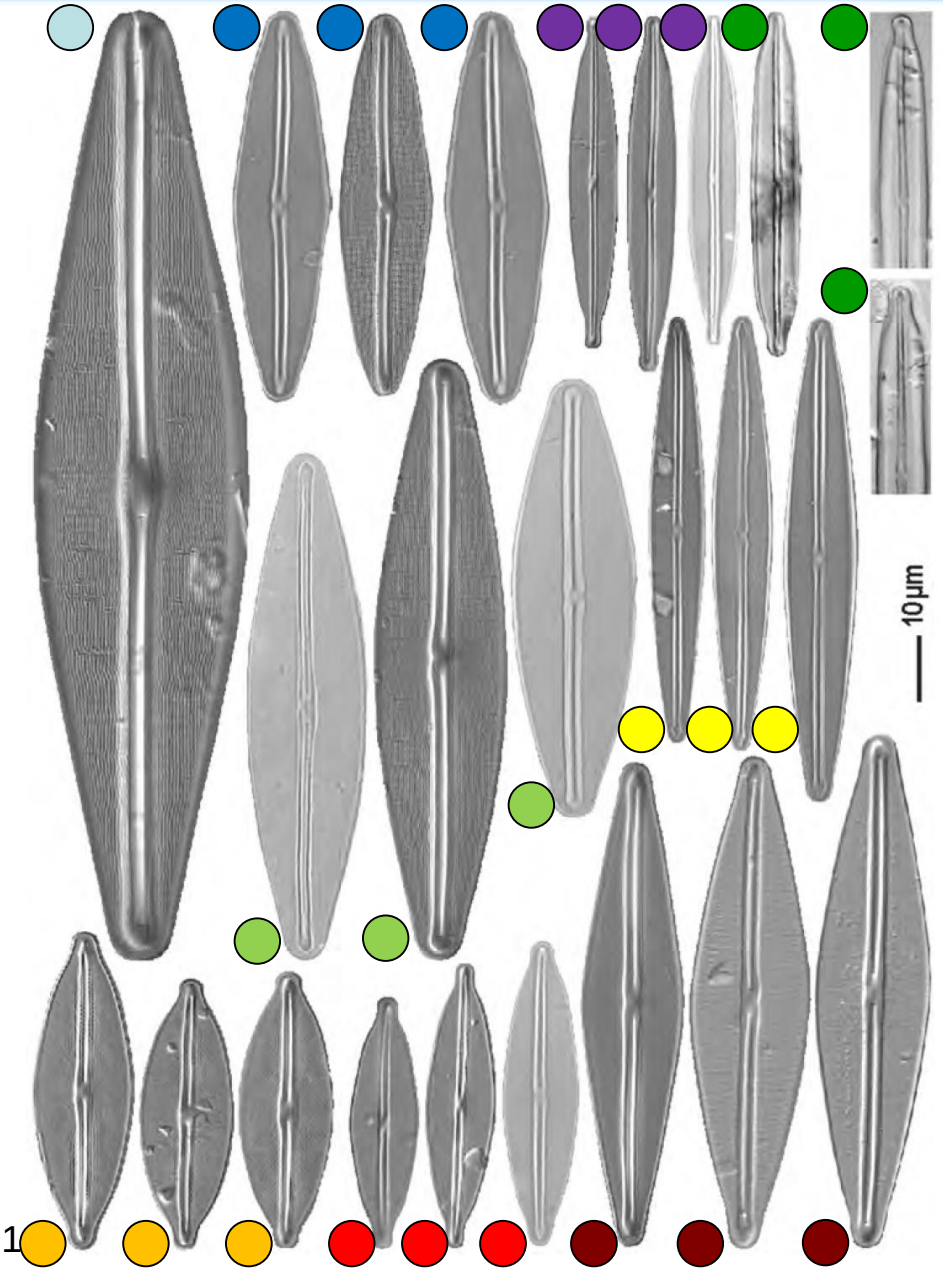
Agamospecies concept

Mayden (1997): In: *Species: the Units of Biodiversity*

Morphological species concept

- A tradition!
  - Still employed in some groups of algae

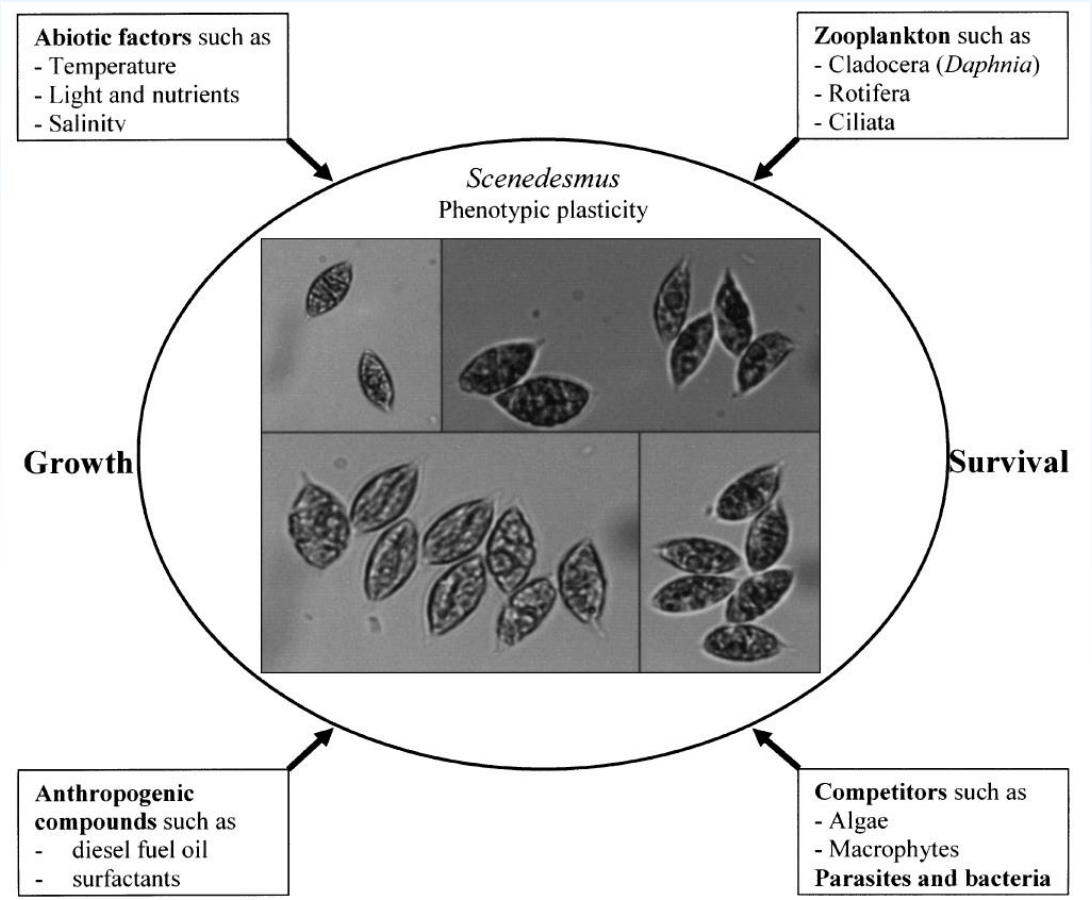




Beier & Lenge-Bertalot (2007): *Nova Hedwigia* 85: 73-91

Morphological species concept

- Species delimited based on investigation of natural samples
  - No data on phenotypic plasticity (*Scenedesmus*)



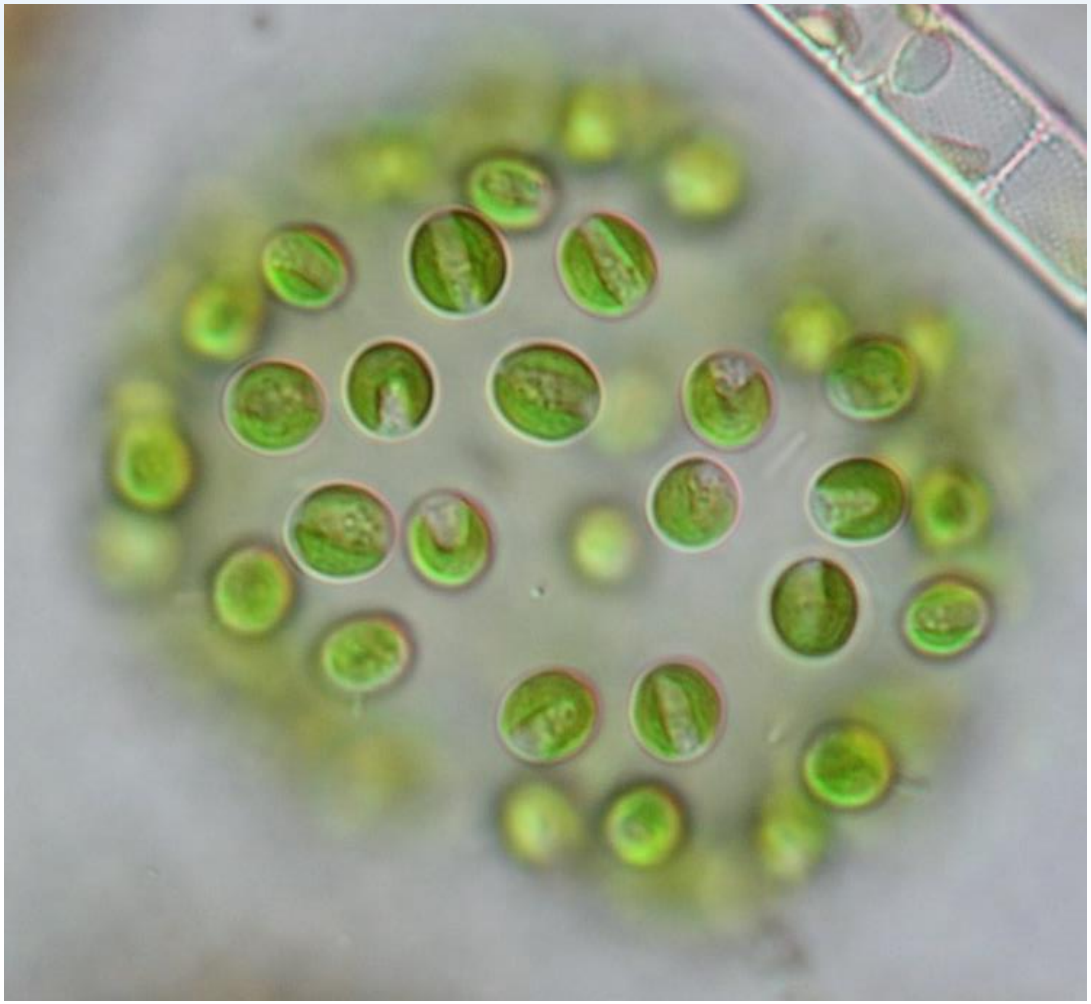
Lürling (2009): *Ann. Limnol. - Int. J. Lim.* 39: 85-101

Morphological species concept

- Species delimited based on investigation of cultured strains
  - Absence of traits manifested only in nature (*Micractinium*)

Morphological species concept

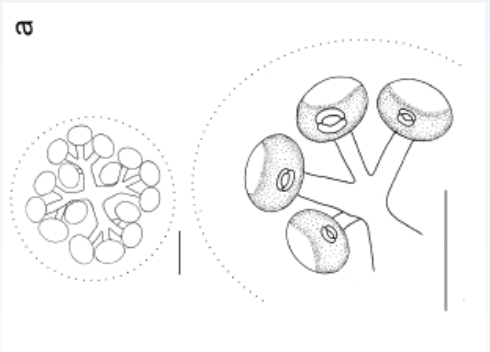
- Convergent morphological evolution
  - *Dictyosphaerium*



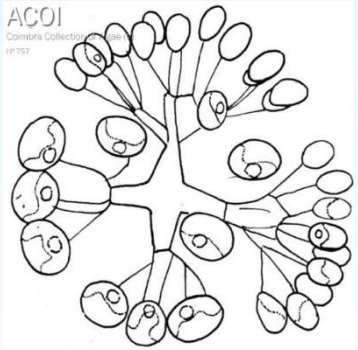
Morphological species concept

- Convergent morphological evolution
  - *Dictyosphaerium* (9 cryptic genera)

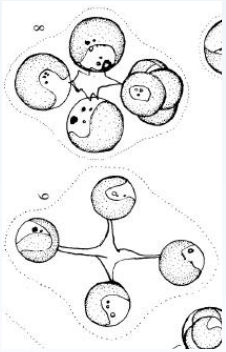
Dictyosphaerium



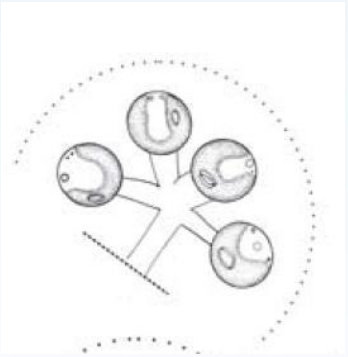
Hindakia



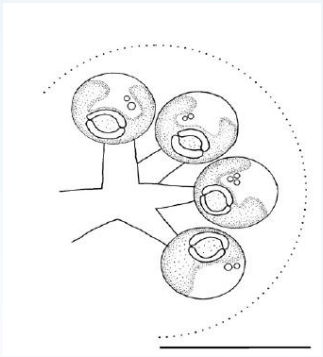
Mychonastes



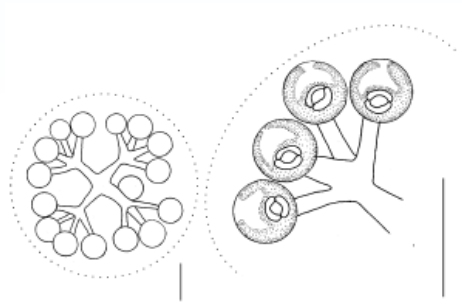
Chlorella



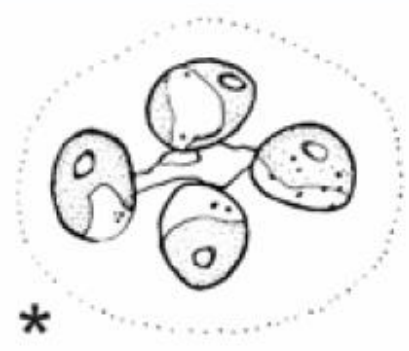
Heynigia



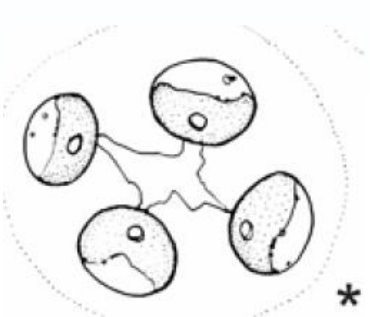
Mucidosphaerium



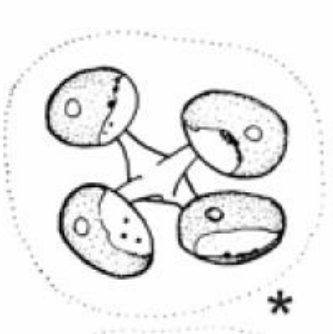
Compactochlorella



Kalenjinia

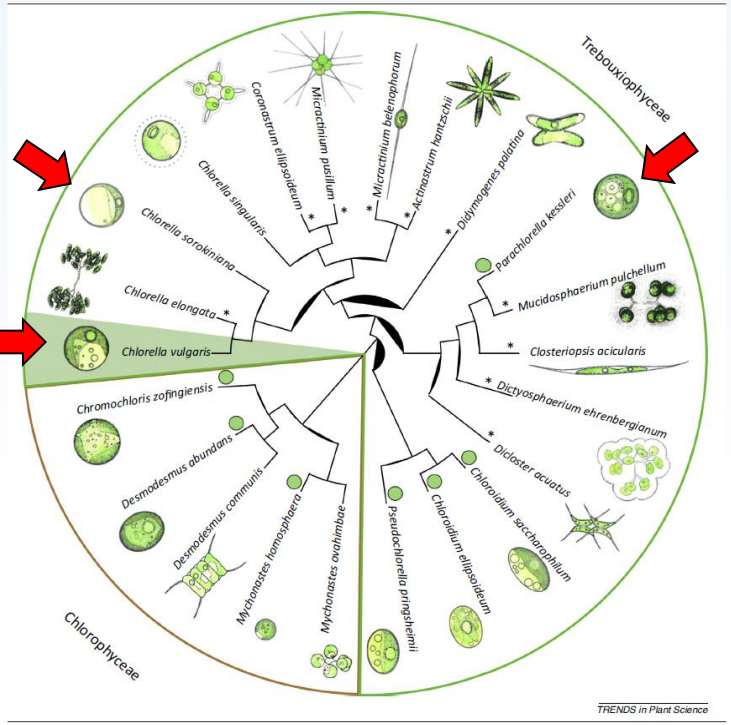


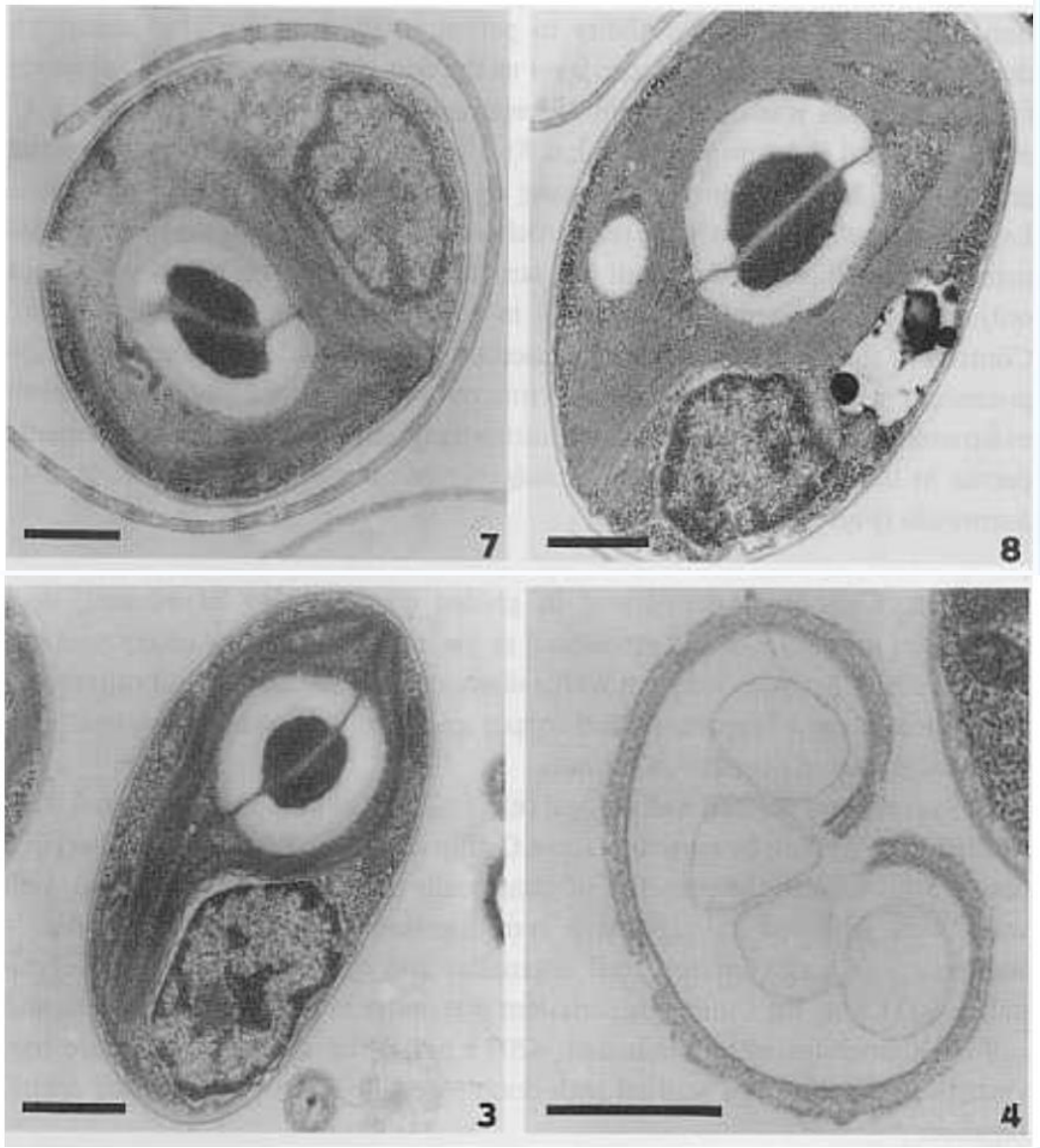
Masaia



Morphological species concept

- Ultrastructure
  - *Chlorella*

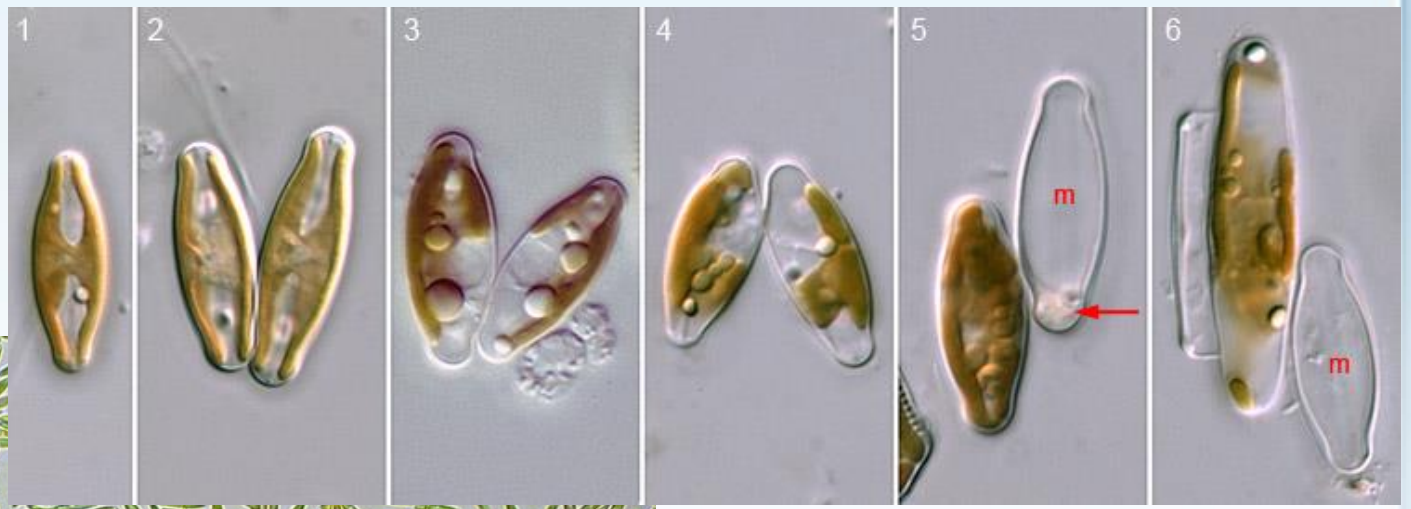




Němcová & Kalina (2000): *Algol. Stud.* 100: 95-105

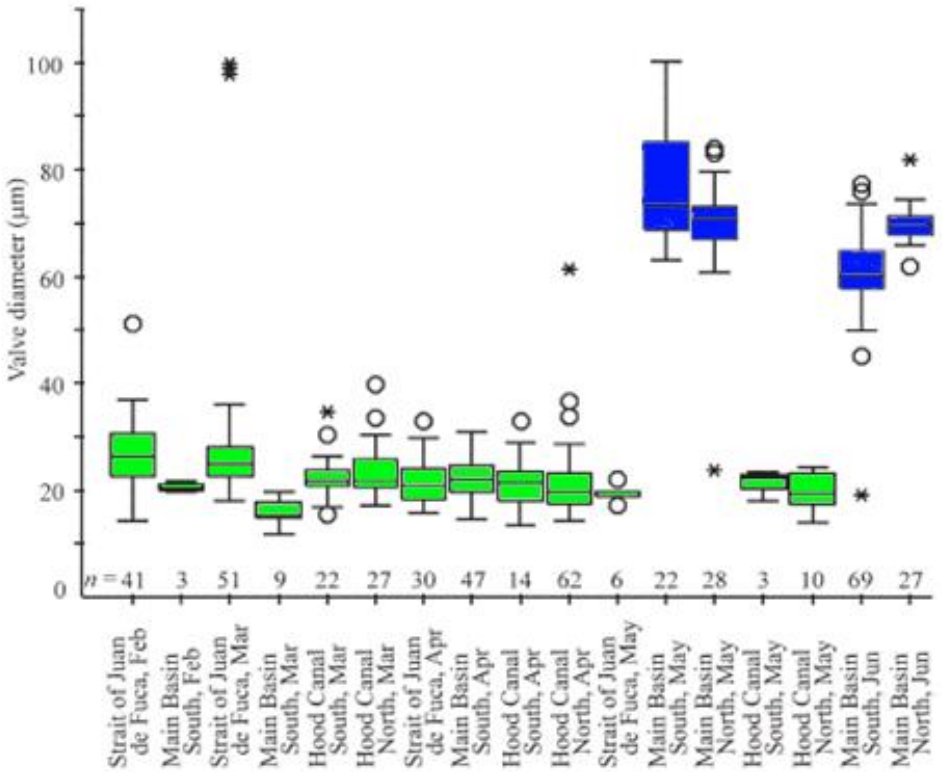
Biological species concept

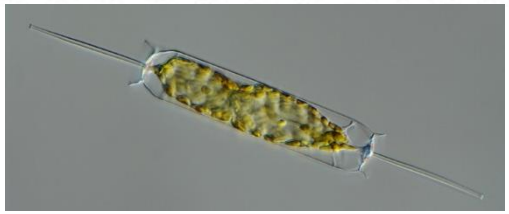
- Applicable only on sexually reproducing organisms
  - diatoms

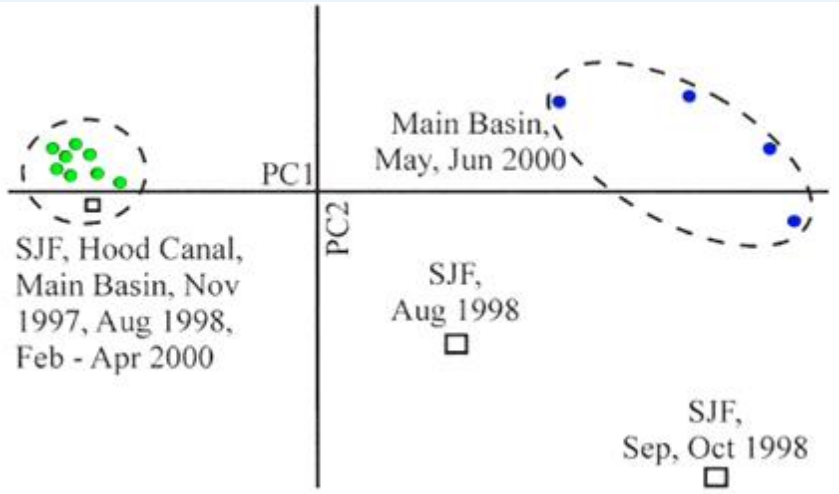


Biological species concept

- Laboratory crossing experiments test the incompatibility
  - ➢ a problem of temporaral isolation (*Ditylum*)



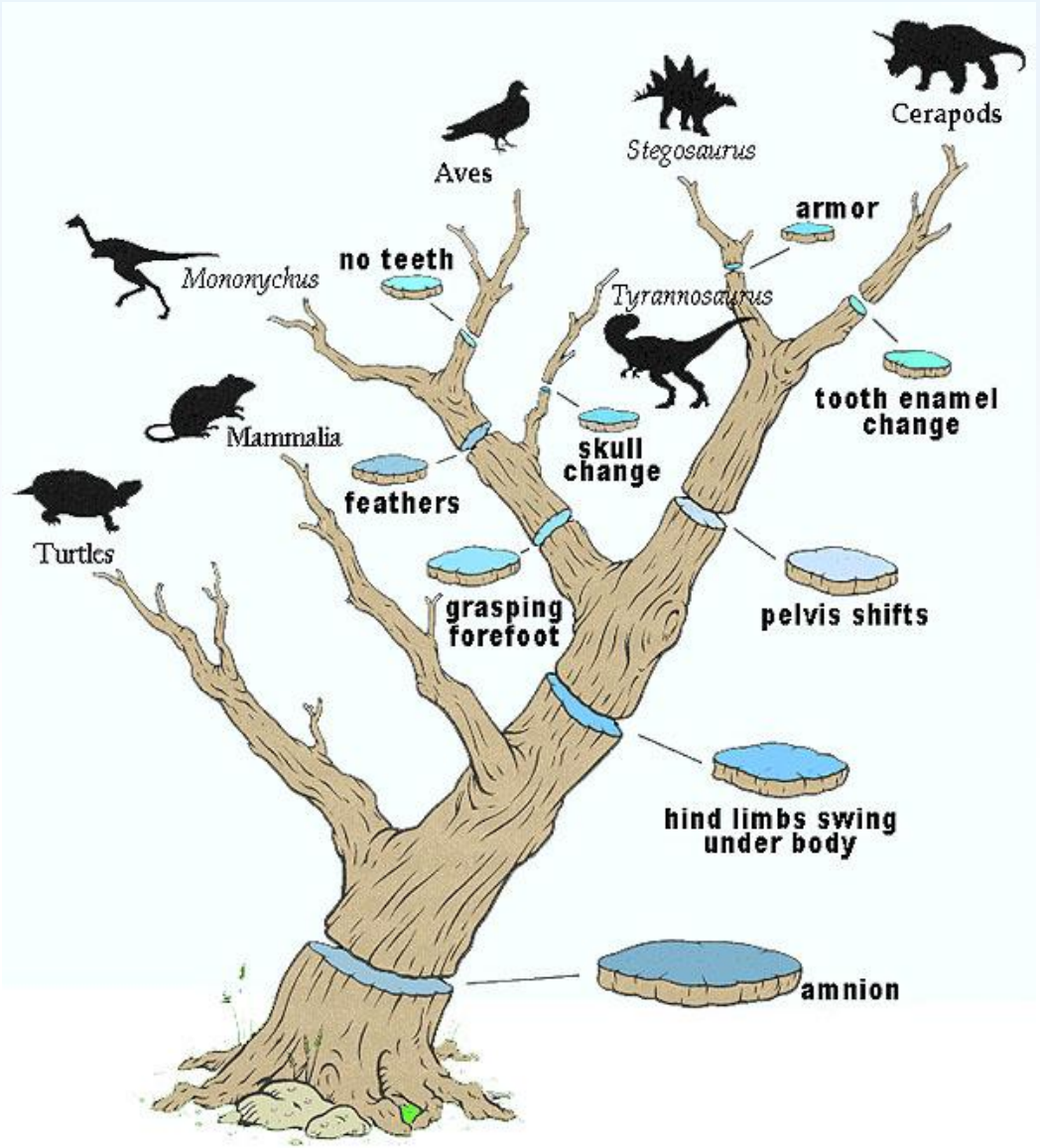




Rynearson et al. (2006): *Limnol. Oceanogr.* 51: 1249-61

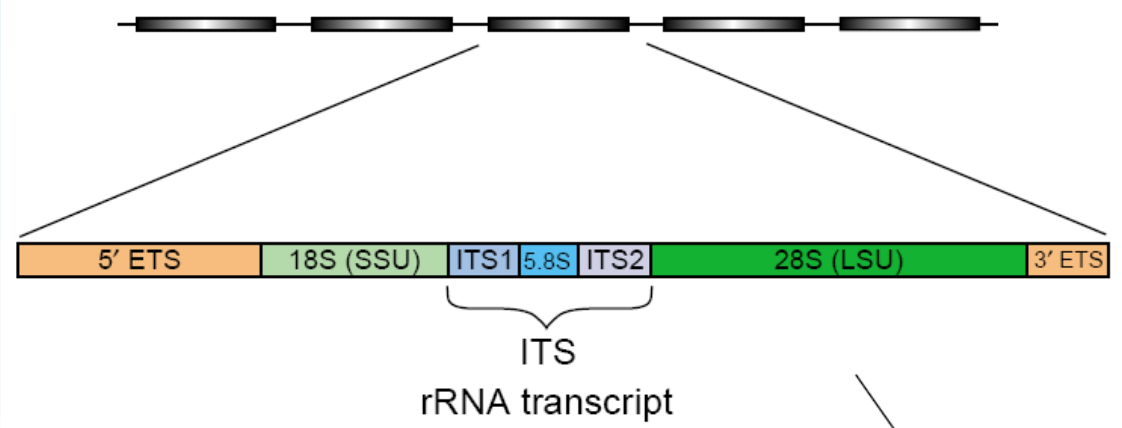
Phylogenetic species concept

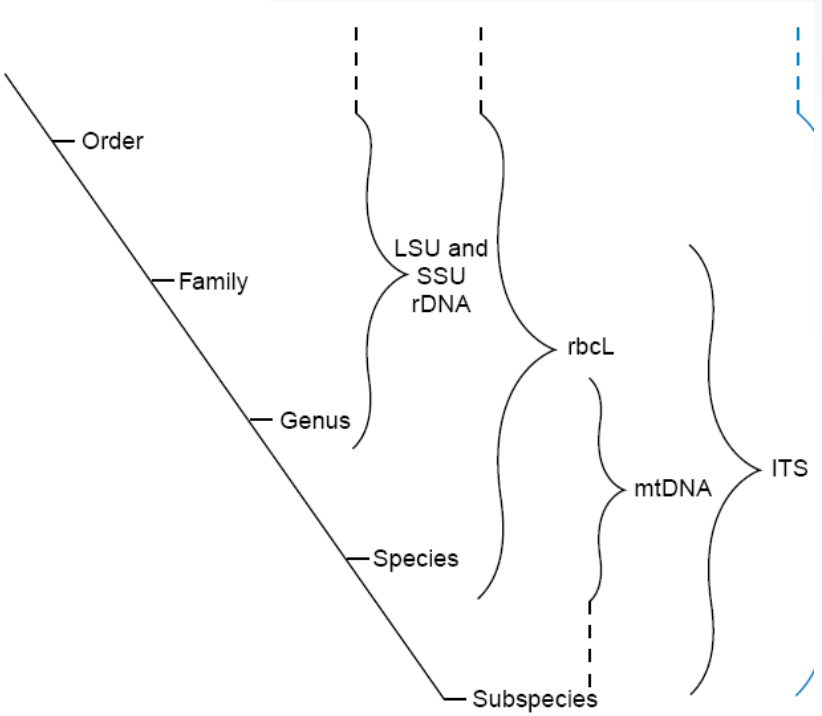
- Based on tree topology (monophyly, branch lengths, supports)



Phylogenetic species concept

- Ribosomal operon as the frequently used marker

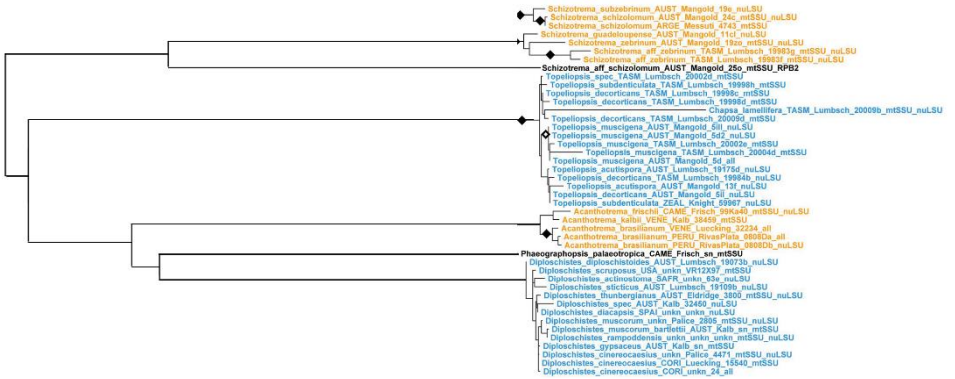


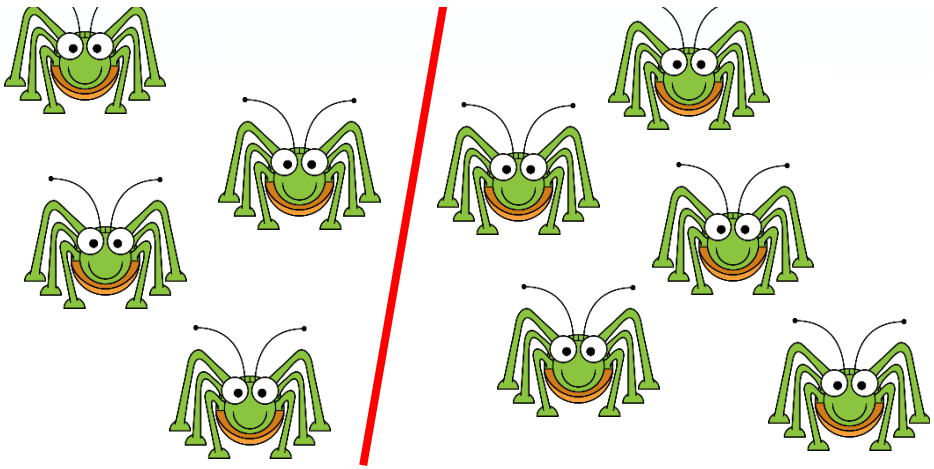


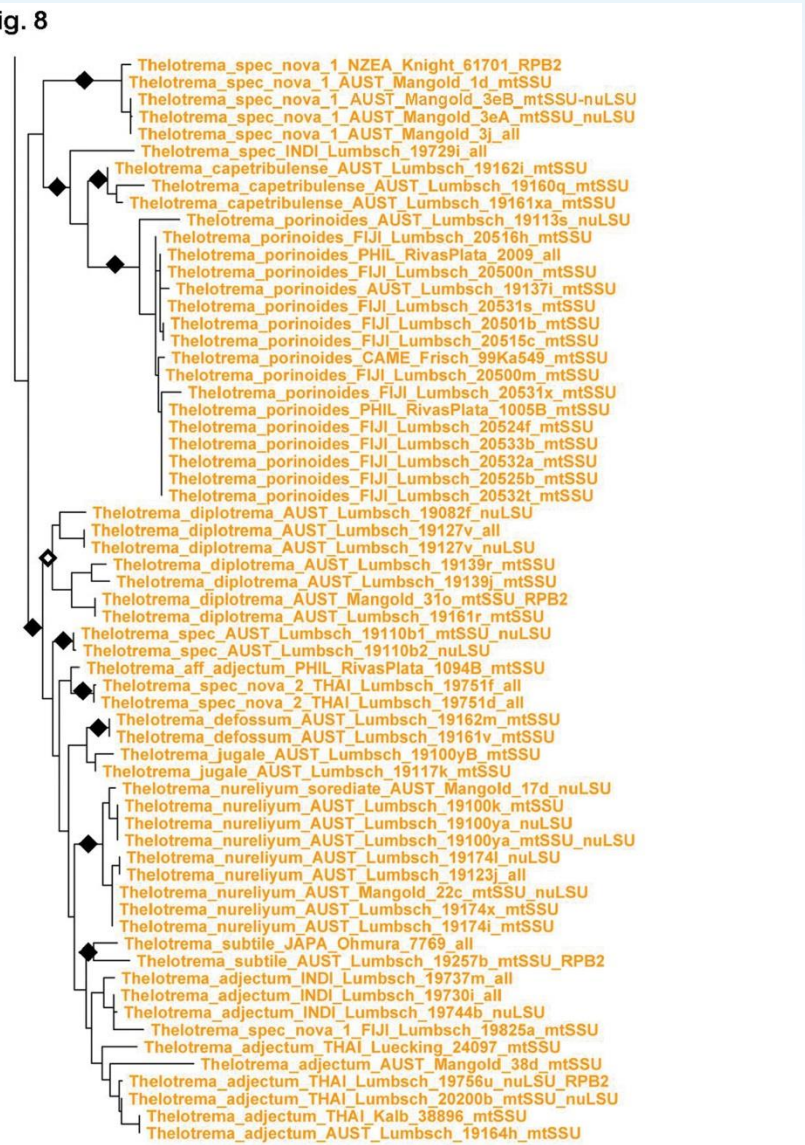
Coleman (2003): *Trends in Genetics* 19: 370-375

Phylogenetic species concept

- Based on tree topology (monophyly, branch lengths, supports)
  - Where to set species boundaries?
  - ***Objective criteria?***

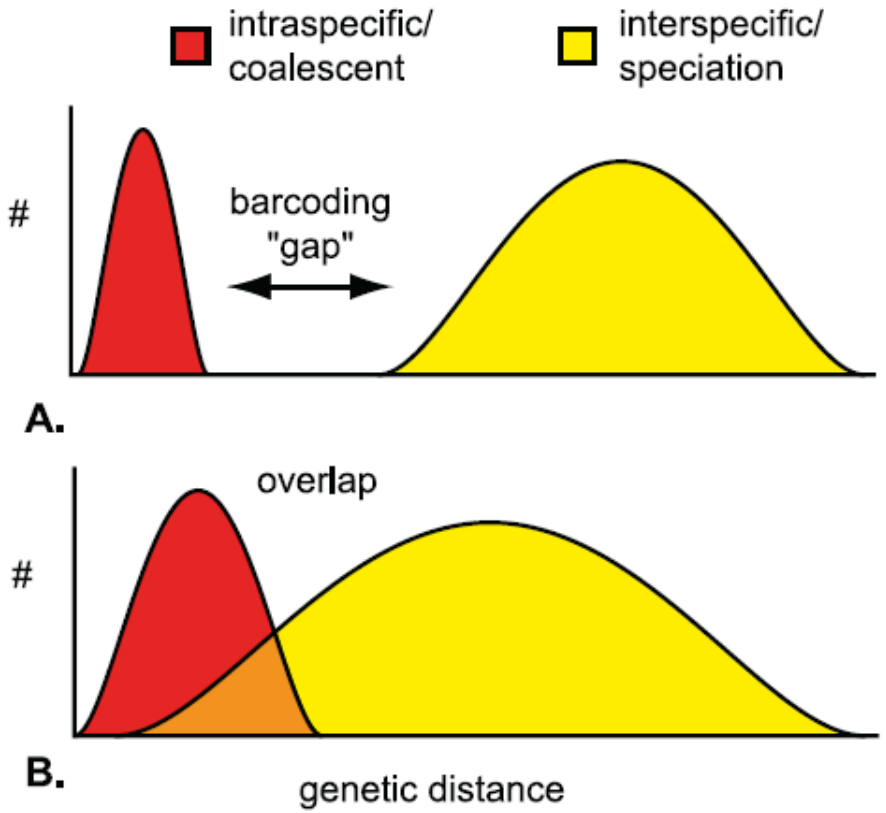


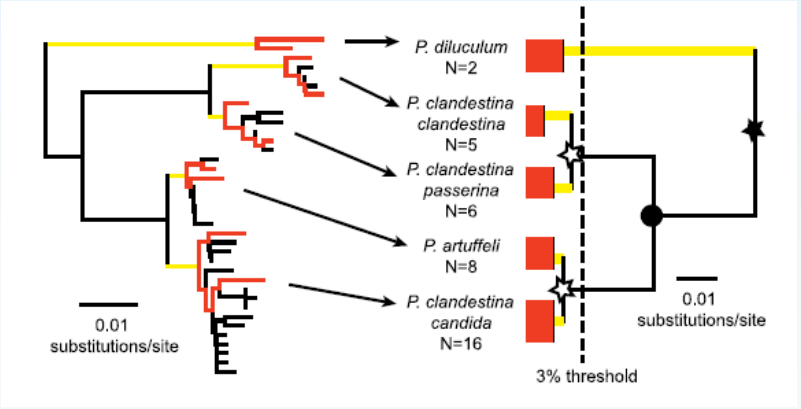


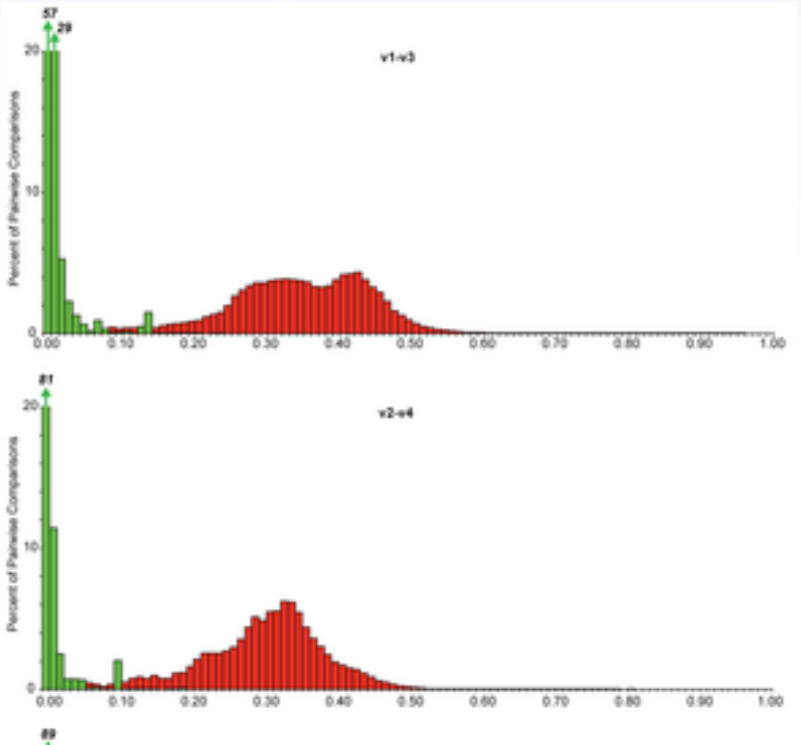


DNA barcoding

- „barcoding gap“



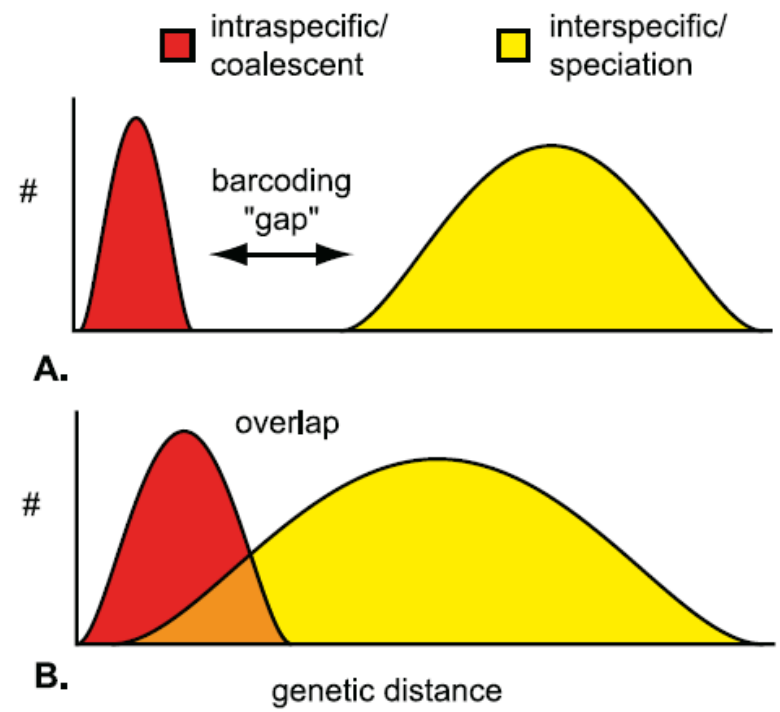


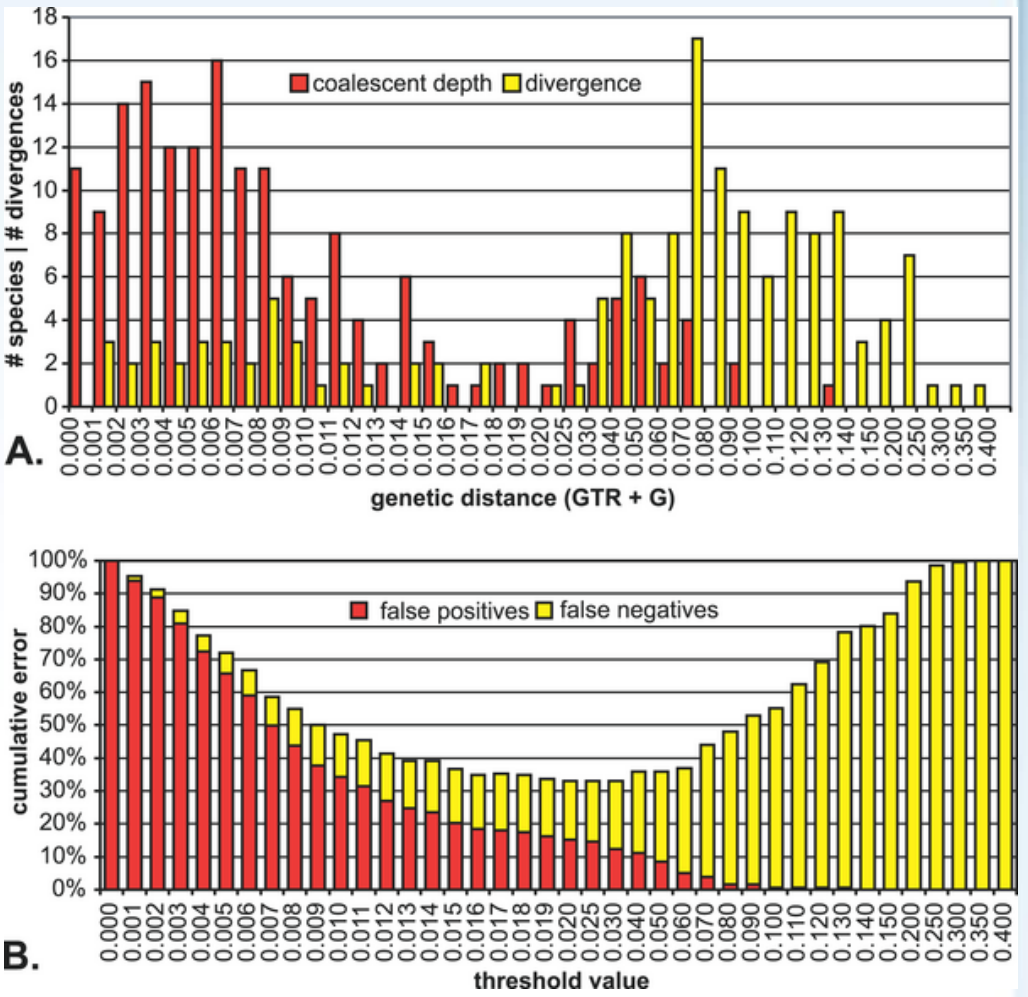


Meyer & Paulay (2005): *PLoS Biol.* 3: e422

DNA barcoding

- More sampling – less obvious barcoding gap

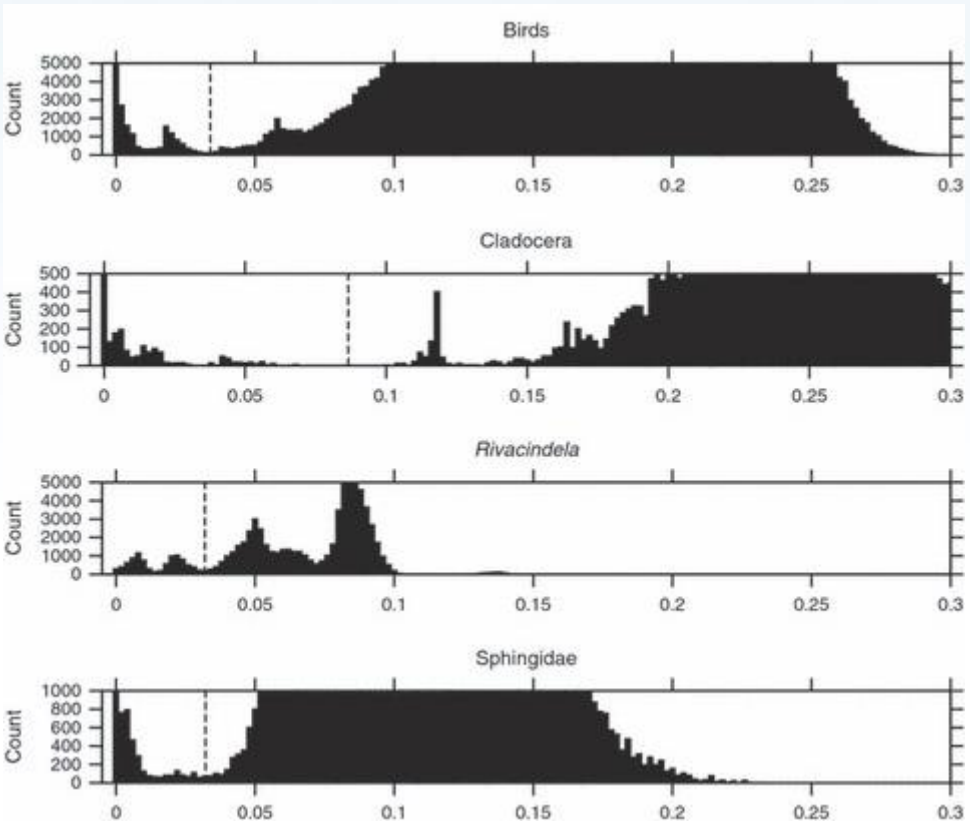




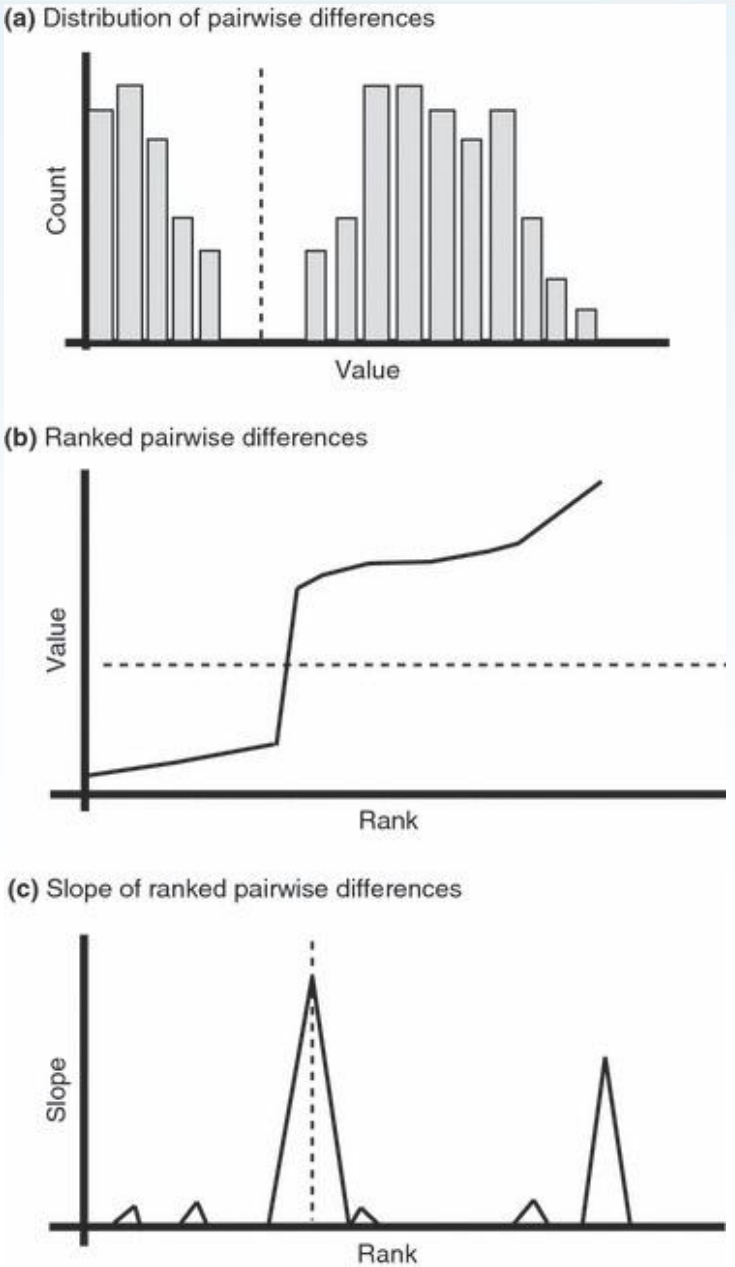
Meyer & Paulay (2005): *PLoS Biol.* 3: e422

DNA barcoding - ABGD

- Automatic Barcode Gap Discovery
  - Alignment as an input file
  - Scanning a range of intraspecific divergence to find the barcode gap

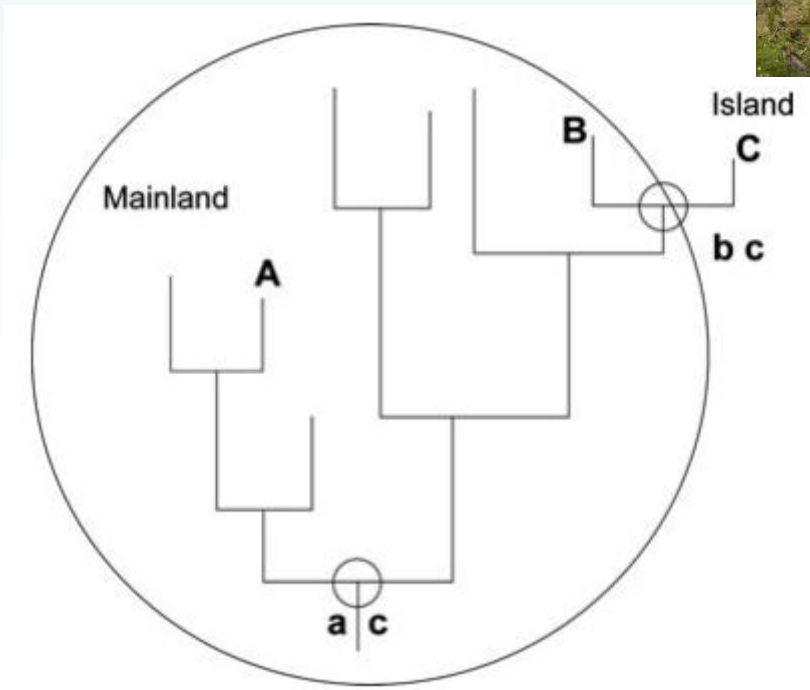


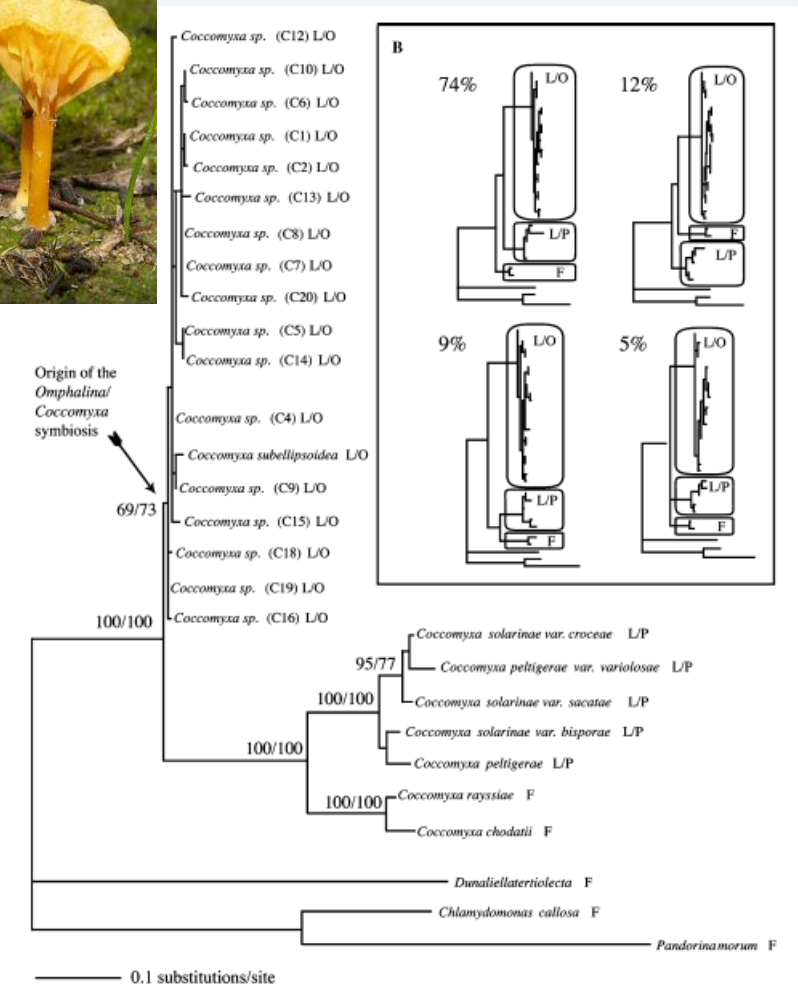
Puillandre et al. (2011): *Mol Ecol* 21: 1864



DNA barcoding

- Uneven mutation rates
  - tropics ~ temperate
  - islands ~ continents
  - free-living ~ symbionts



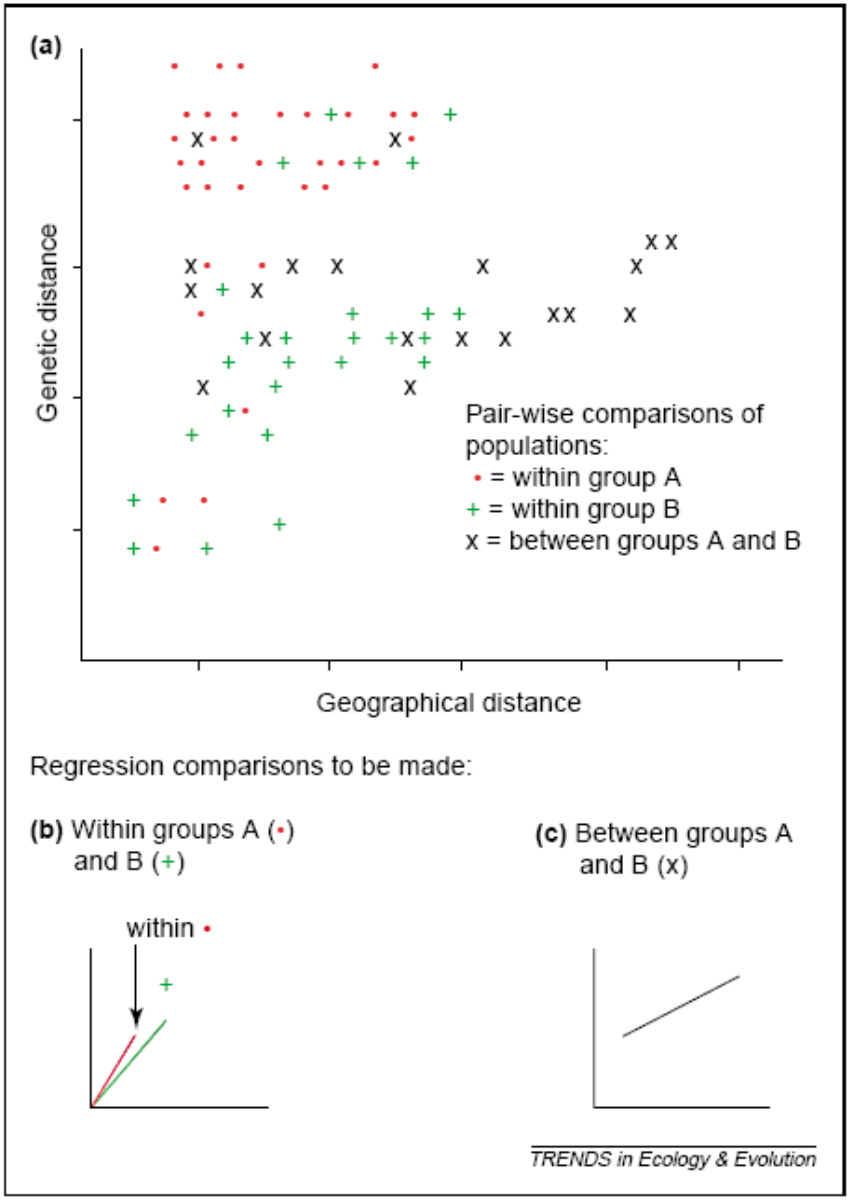


Wright et al. (2009): *Evolution* 63: 2275-2287

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

Good & Wake method

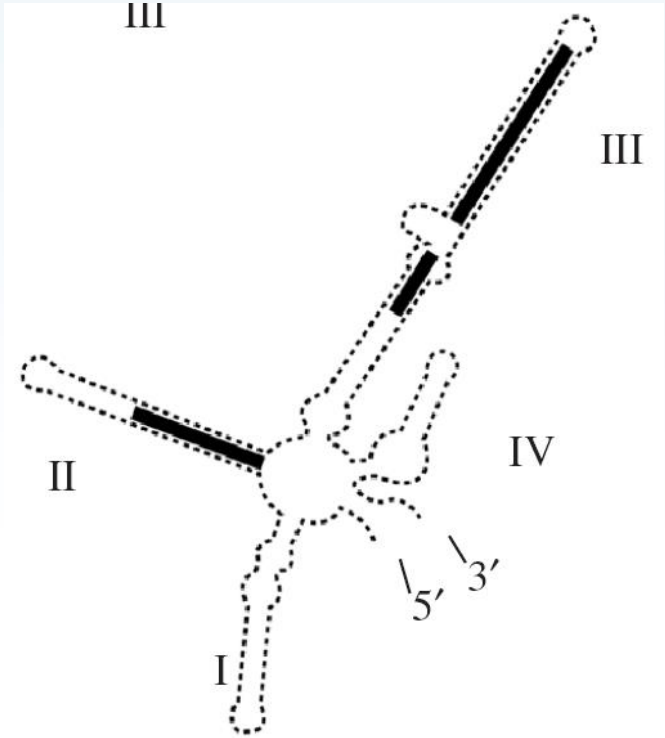
- test of *a priori* defined species
- linear regression of genetic and geographic dostances should in single species go through the graph origin (gene flow with isolation-by-distance)
- Different regression indicates the presence of two distinct, genetically isolated species

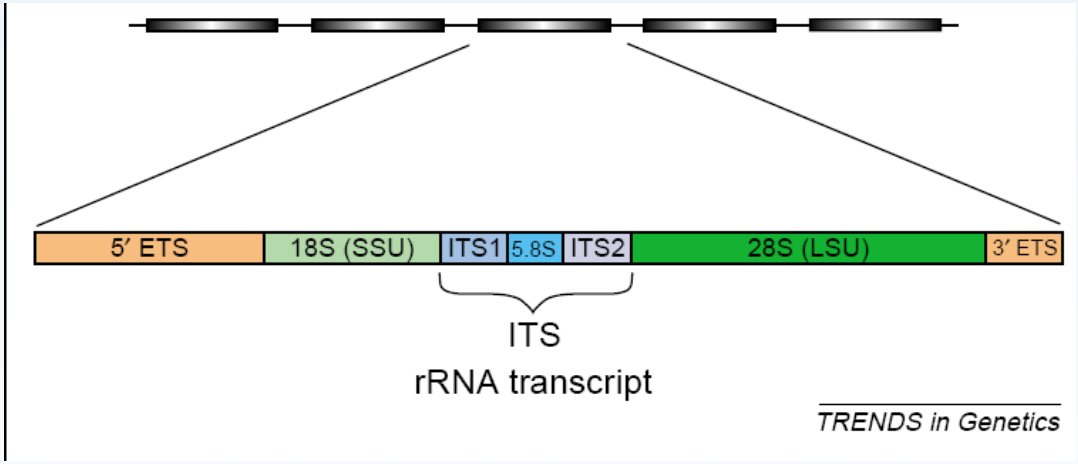


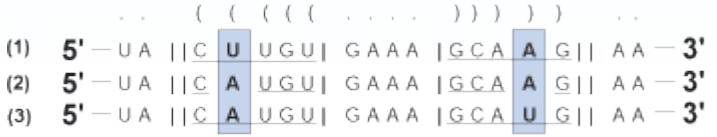
Sites & Marshall (2012): *Trends Ecol. Evol.* 18: 462-470

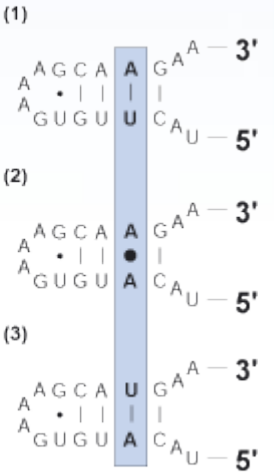
CBC concept

- Species delimited based on differences in conservative regions of the ITS2 molecule







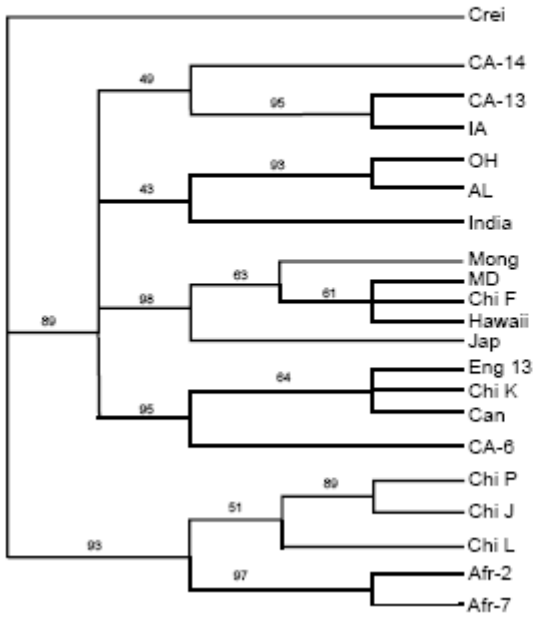


Coleman (2007): *Nucleic Acids Res.* 2007: 1-8

CBC concept

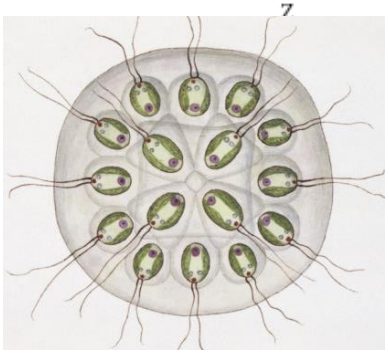
- Species delimited based on differences in conservative regions of the ITS2 molecule
  - ➢ correlation between CBC and sexual compatibility (*Gonium*)

A.



B.

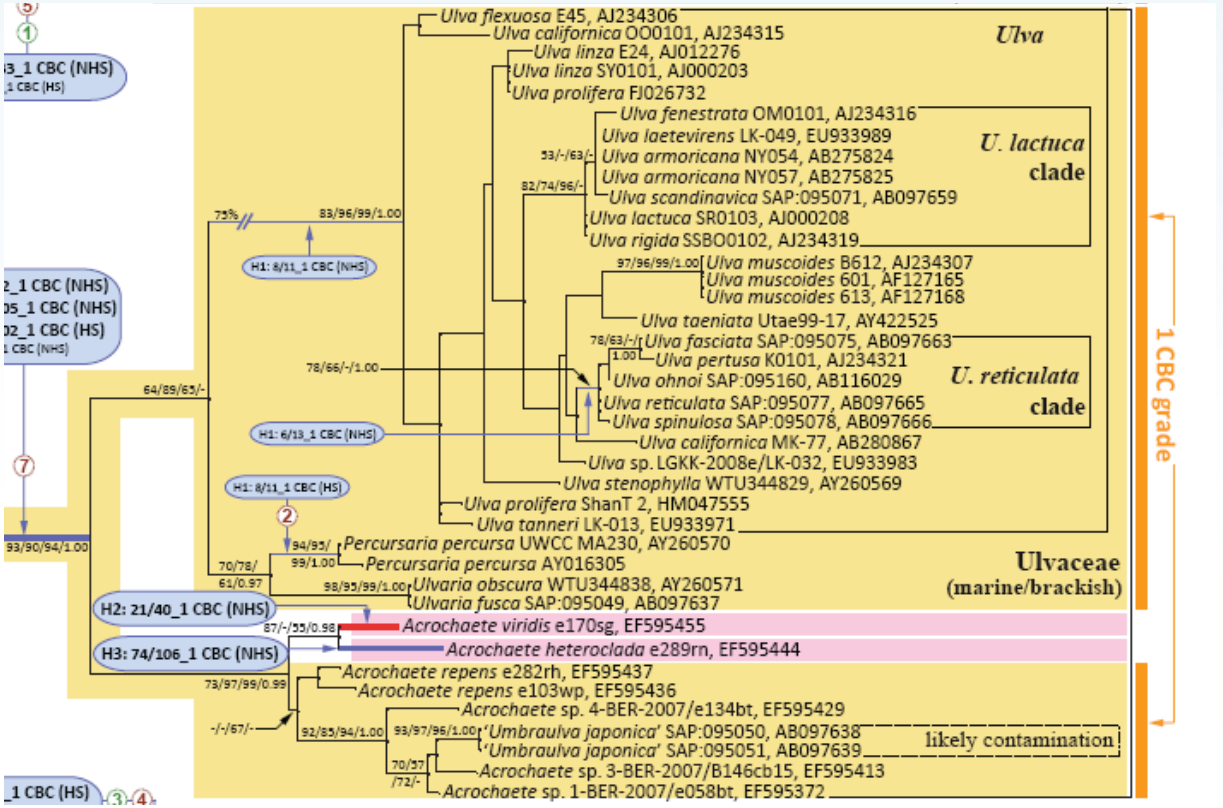
(+) / (-)	CA-13-x	CA-14-1	IA-2	AL-1	India	Mong	Jap-1	MD-4	Can-9	CA-7	Chi-J	Chi-L-1
CA-13-2	(Z)abc		(Z)		Z	z		(Z)				
CA-14-x	(Z)	(Z)abc										
IA-1			Zabc									
OH		Z	Za	Zabc	Zabc	Za		zab	Z		Za	
AL-2				Zabc	Zabc	Zabc	Zab				Zabc	
Mong		z	z	z	Za	Zabc	Zabc		z		Zab	z
Chi-F-2				Za	Zab	z	z	Zabc				
Jap-2				Z	Z	z	Zabc					
MD-3	(Z)					Zabc		Zabc			z	
Chi-K						Z			Zab		Za	
Eng-13	(Z)	(Z)			Z	Z			Zabc	(Z)		
Can-0	(Z)	(Z) a				Z		z	Zabc	(Z)		
CA-6		(Z)							Zab	Zabc	Za	
Afr -7-1				Z		z					Zabc	Zabc
Afr -2-3				Z		Za					Zabc	Za
Chi-P												
Chi-L-2											Zabc	Zabc



Coleman (2000): *Protist* 151: 1-9

CBC concept

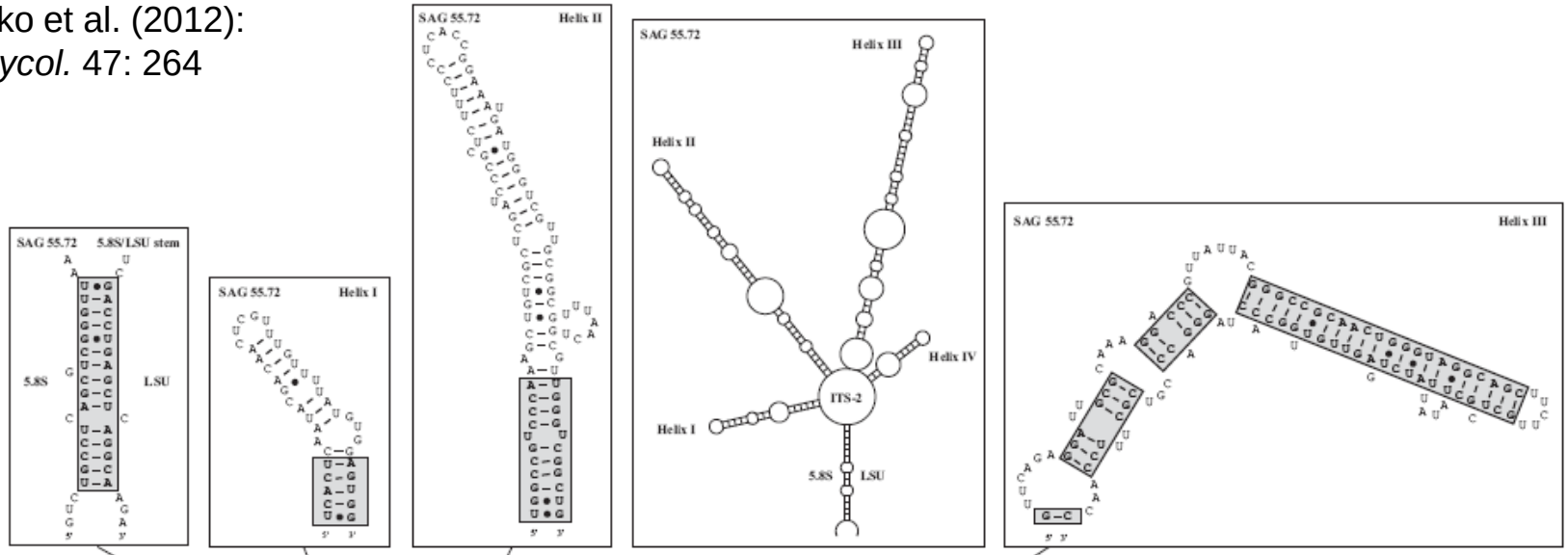
- Species delimited based on differences in conservative regions of the ITS2 molecule
  - ➢ not a universal concept
  - ➢ CBCs as a measure of genetic relation, not species marker



Caisová et al. (2011): *BMC Evol. Biol.* 11: 262

ITS2-barcode concept

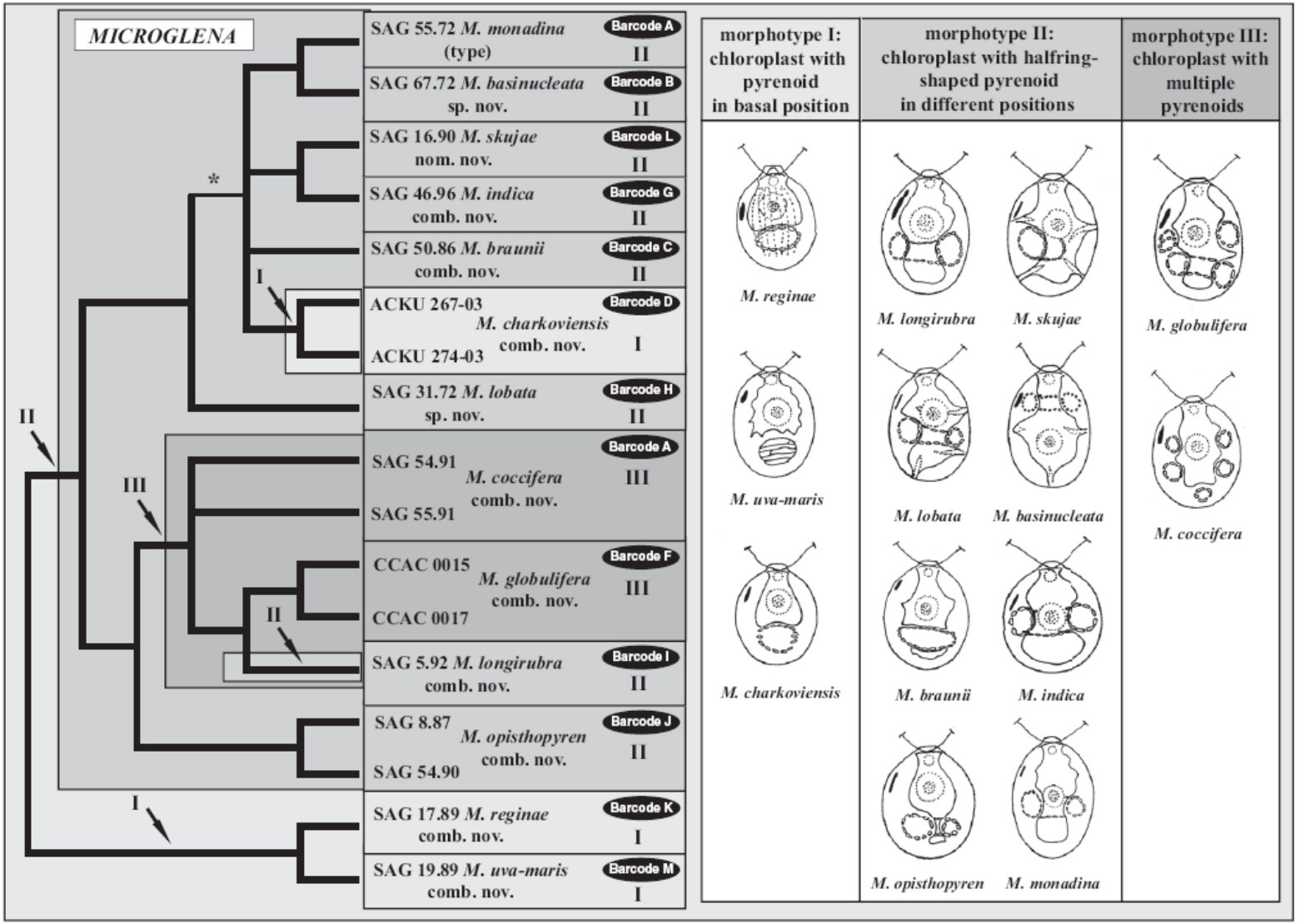
Demchenko et al. (2012):
Eur. J. Phycol. 47: 264



Species	- strain	5.8S	Helix I	Helix II	Helix III	LSU
		000000000000000 00000 00000	00000000000 0000000000	0000000000000000000000000000000000000000000000000000000000000000000000000000000	111111111111111111111111111111111111111111111111111111111
		000000000111111 11112 22222	22223333333 33344444444	4455555555556666666667777777778888888888999999999000000000011111111112222222223333333333444444	444455555555556
		123456789012345 67890 12345	67890123456 78901234567	8901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901	67890123456789
M. monadina	- SAG 55.72	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCCGUCCCA..UGGGUCGGCUG...G-----GG---AGCG--GG-CC--CG-GGCCGCAACUGGGUAGGCAGC--GCUGCUUAUCUAGUUGUGGGC-CG--GG-CC--CGCU---CC-----C...GACCUGAGCUAGGCA				
M. basinucleata	- SAG 67.72	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCUGUCCCA..UGGGUCAGCUG...GUGAG-GG-UUAACGCAGG--C--CG-GGCCGCAACUGGG--GGCAGC--GCUGCU--UCUAGUUGUGGCU-UG--G--CCUGCGUUGA-CC-CUCAC...GACCUGAGCUAGACA				
M. braunii	- SAG 50.86	UGCCUAGCUCGGGUU...UCAAU..AUUGG..UGGCUGUCCCA..UGGGUCAGCUG...GU-AGAGG--AGGCG-GGA-CC--CG-GGCCGCAAUGGGU-AGGCAGC--GCUGCUU-AUCUGUUGUGGCC-CG--GG-UCU-CGCUU--CCUCU-AC...GACCUGAGCUAGGCA				
M. charkoviensis	- ACKU 267-03	UGCCUAGCUCGGGUU...UCACC..GGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GU-GGAGGUUGGACG-CGG-CC--CG-GGCCGCAACUGGG--AGCAGC--GCUGCU--CUUAGUUGUGGCC-CG--GG-CCG-CGUUUGACCUCC-GC...GACCUGAGCUAGGCA				
	- ACKU 274-03	UGCCUAGCUCGGGUU...UCACC..GGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GU-GGAGGUUGGACG-CGG-CC--CG-GGCCGCAACUGGG--AGCAGC--GCUGCU--CUUAGUUGUGGCC-CG--GG-CCG-CGUUUGACCUCC-GC...GACCUGAGCUAGGCA				
M. coccifera	- SAG 54.91	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GUUAGAGG--AGACG-AGG-CC--CG-GGCCGCAAUUGGG--GGCAGC--GCUGCU--CUUAAUUGUGGCC-CG--GG-CCU-CGUCU--CCUCUAGC...GACCUGAGCUAGGCA				
	- SAG 55.91	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GUUAGAGG--AGACG-AGG-CC--CG-GGCCGCAAUUGGG--GGCAGC--GCUGCU--CUUAAUUGUGGCC-CG--GG-CCU-CGUCU--CCUCUAGC...GACCUGAGCUAGGCA				
M. glebulifera	- CCAC 0015	UGCCUAGCUCGGGUU...UCAUU..GAUGG..UGGCCGUCUCA..UGGGUCGGCUG...GUU-GAGG-UAAACG-AGG-CC--CG-GGCCGCAAUUGGG--AGCAGC--GCUGCU--CUUAGUUGUGGCC-CG--GG-CCU-UGUUUG-CUUC-AGC...GACCUGAGCUAGGCA				
	- CCAC 0017	UGCCUAGCUCGGGUU...UCAUU..GAUGG..UGGCCGUCUCA..UGGGUCGGCUG...GUU-GAGG-UAAACG-AGG-CC--CG-GGCCGCAAUUGGG--AGCAGC--GCUGCU--CUUAGUUGUGGCC-CG--GG-CCU-UGUUUG-CUUC-AGC...GACCUGAGCUAGGCA				
M. indica	- SAG 46.96	UGCCUAGCUCGGGUU...UCAUC..GAUGG..UGGUGGUCCCA..UGGGUCCGCUG...GUUAG-GG--AGGC--GGG-CC--CG-GGCCGCAACUGGGU-GGCAGC--GCUGCU-GCUUAGUUGUGGCC-CG--GG-CCU--GUCU--CC-CUAAC...GACCUGAGCUAGGCA				
M. lobata	- SAG 31.72	UGCCUAGCUCGGGUU...UCAUU..AUUGG..UGGCUGUCCCA..UGGGUCUGCUG...GU-AGAGG---GGCG-AGG-CC--CG-GGCCGCAAUGGGU-AGGCAGC--GCUGCUU-AUCUGUUGUGGCC-CG--GG-CCU-CGCU---CCUCU-AC...GACCUGAGCUAGGCA				
M. longirubra	- SAG 5.92	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GUU-GAGG--AGACG-AGG-CC--CG-GGCCGCAAUUGGGU-GGCAGC--GCUGCU-ACUUGAUUGUGGCU-CG--GG-CCU-CGUCU--CCUU-AGC...GACCUGAGCUAGGCA				
M. opisthopyren	- SAG 8.87	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GUUAGAGG---GGCG-AGG-CC--CG-GGCCGCAAUUGGG--GGCAGC--GCUGCU--CUUAAUUGUGGCC-CG--GG-CCU-CGCC---CCUCUAGC...GACCUGAGCUAGGCA				
	- SAG 54.90	UGCCUAGCUCGGGUU...UCACU..AGUGG..UGGCCGUCCCA..UGGGUCGGCUG...GUUAGAGG---GGCG-AGG-CC--CG-GGCCGCAAUUGGG--GGCAGC--GCUGCU--CUUAAUUGUGGCC-CG--GG-CCU-CGCC---CCUCUAGC...GACCUGAGCUAGGCA				
M. reginae	- SAG 17.89	UGCCUAGCUCGGGUA...UCACC..GGUGG..UGGCCAUCCCA..UGGGUCGGCCG...GGUAGAGAAC-GGUA-AGGACCU-AGGGGCCGCAAC-GGG--GGUAGC--GCUACU--CUU-GUUGUGGUUCCU-AGGUCCU-UACU-GUUUUCUACC...GACCUGAGCUAGGCA				
M. skujae	- SAG 16.90	UGCCUAGCUCGGGUU...UCAUC..GGUGG..UGGCCGUCCCA..UGGGUCUGCUG...GU-AG-GG--AGGCGCAGG-AC--CG-GGCCGCAAUUGGGU-GGCAGCUAGCUGCU-GCUUAAUUGUGGCC-CG--GU-CCUGUGUCU--CC-CU-GC...GACCUGAGCUAGGCA				
M. uva-maris	- SAG 19.89	UGCCUAGCUCGGGUU...UCACG..CGUGG..UGACAGUCCCA..UGGGUCUGUUG...GUAAGAGAUC-GGC--AGGACCUGAG-GGCCGCAACUGGGUAGGCAGC--GCUGUUUACUCAGUUGUGGGU-CUUAGGUCCU--GCC-GAUUUCUUAC...GACCUGAGCUAGGCA				

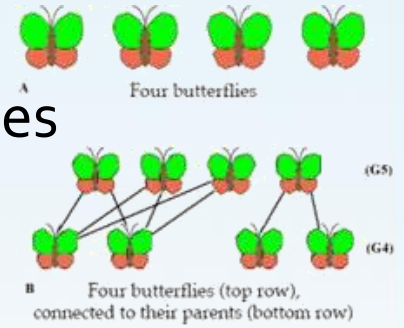
ITS2-barcode concept

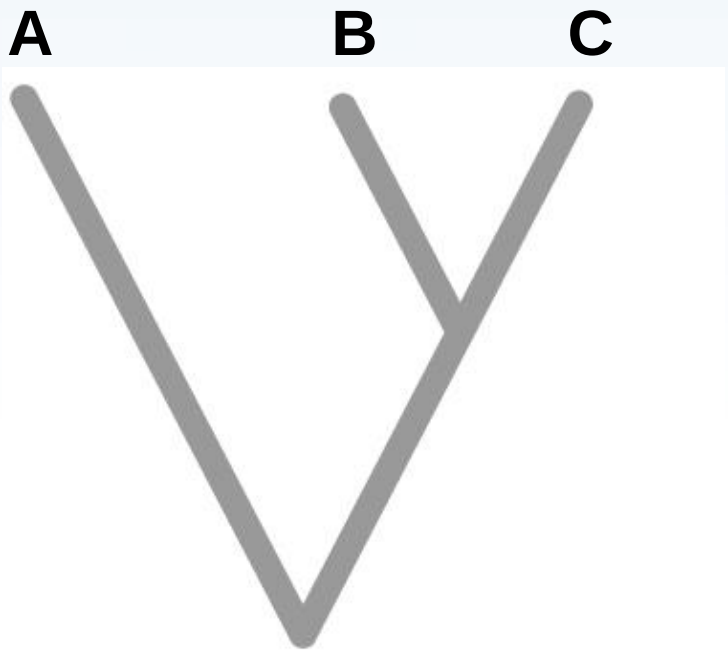
- Again, CBCs is a measure of genetic relation, not species marker



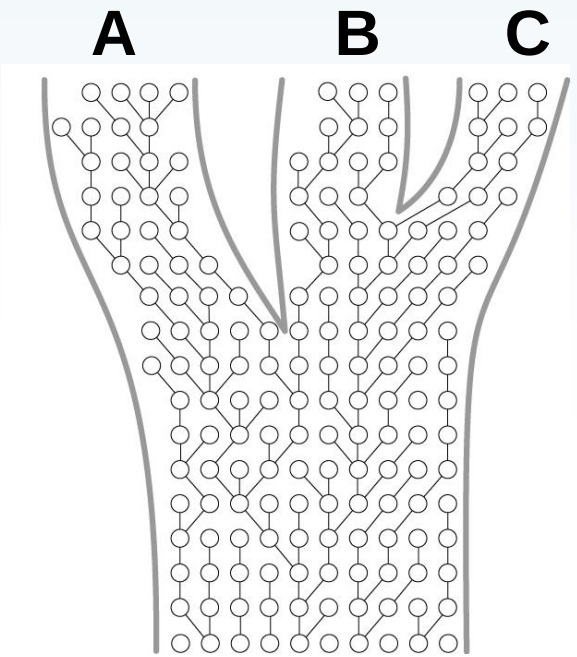
Coalescent theory

- Linking phylogenetics and population genetics
  - Identification of independently evolving lineages
  - GMYC, bPTP, BP&P, ….

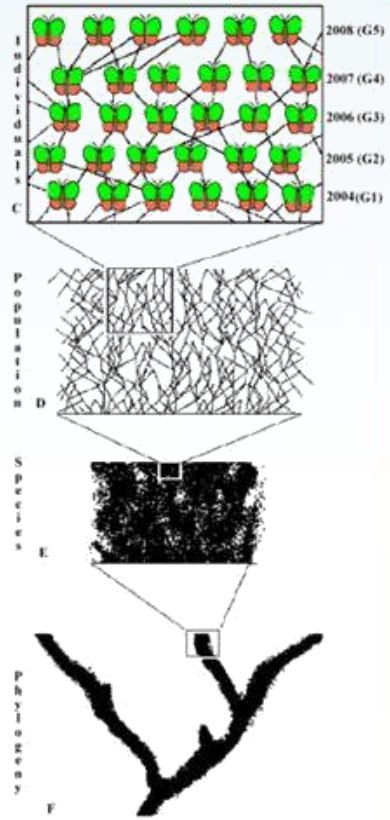




Phylogenetics

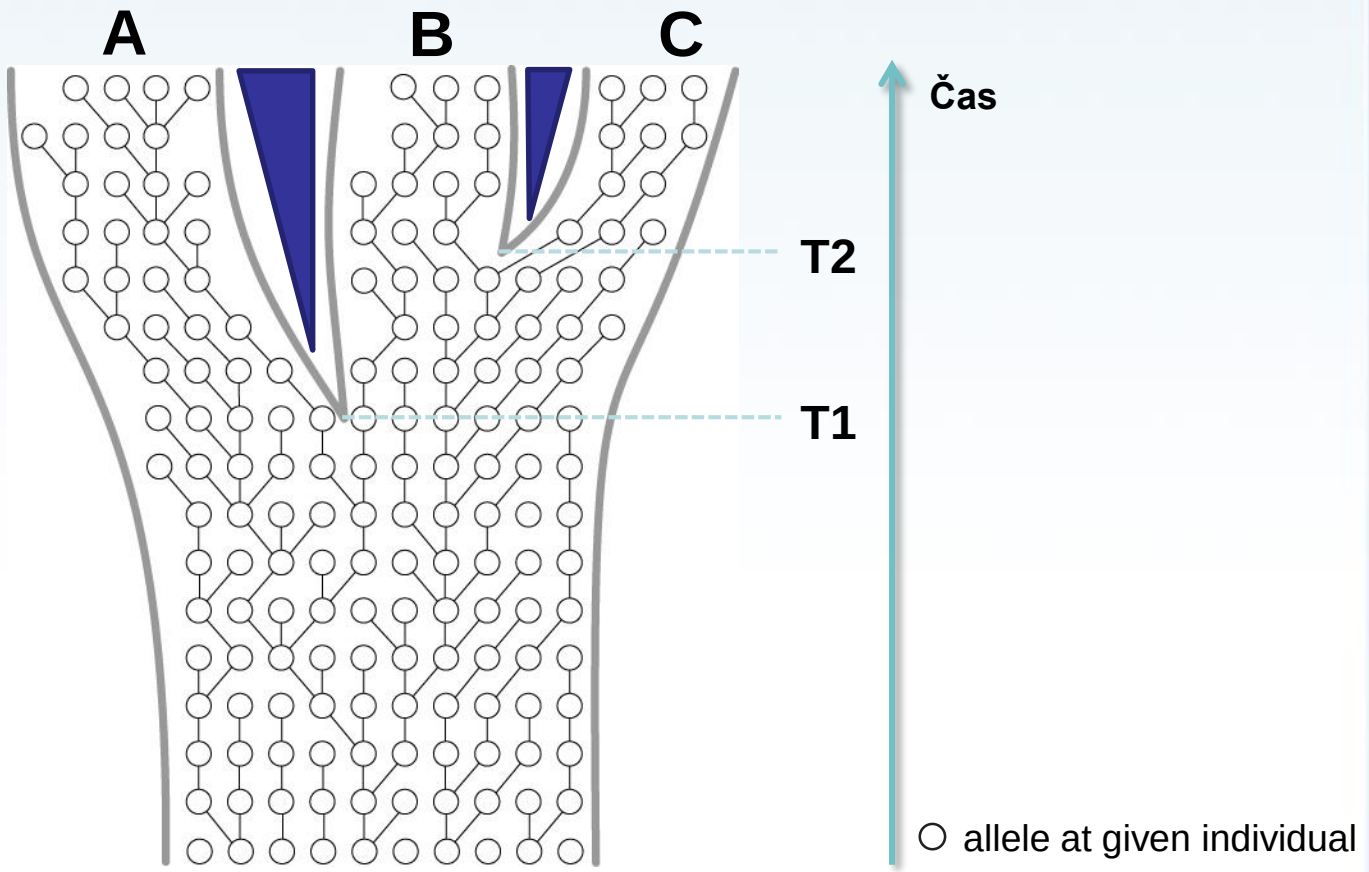


Population genetics



Coalescent theory

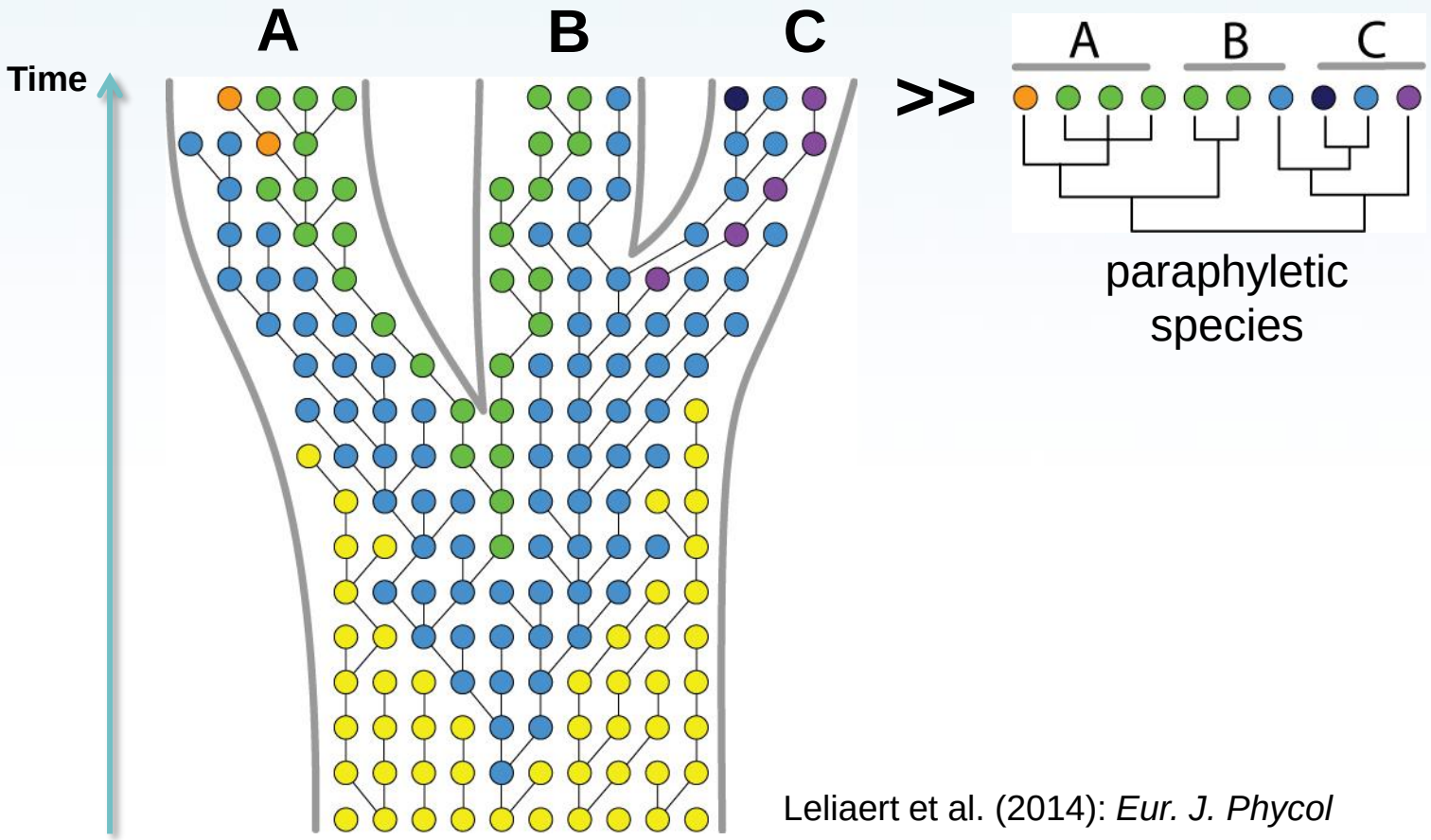
- Coalescence processes (Wright-Fisher)
  - allelic transfer to next generations
  - allelic frequences vary across generations



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

Coalescent theory

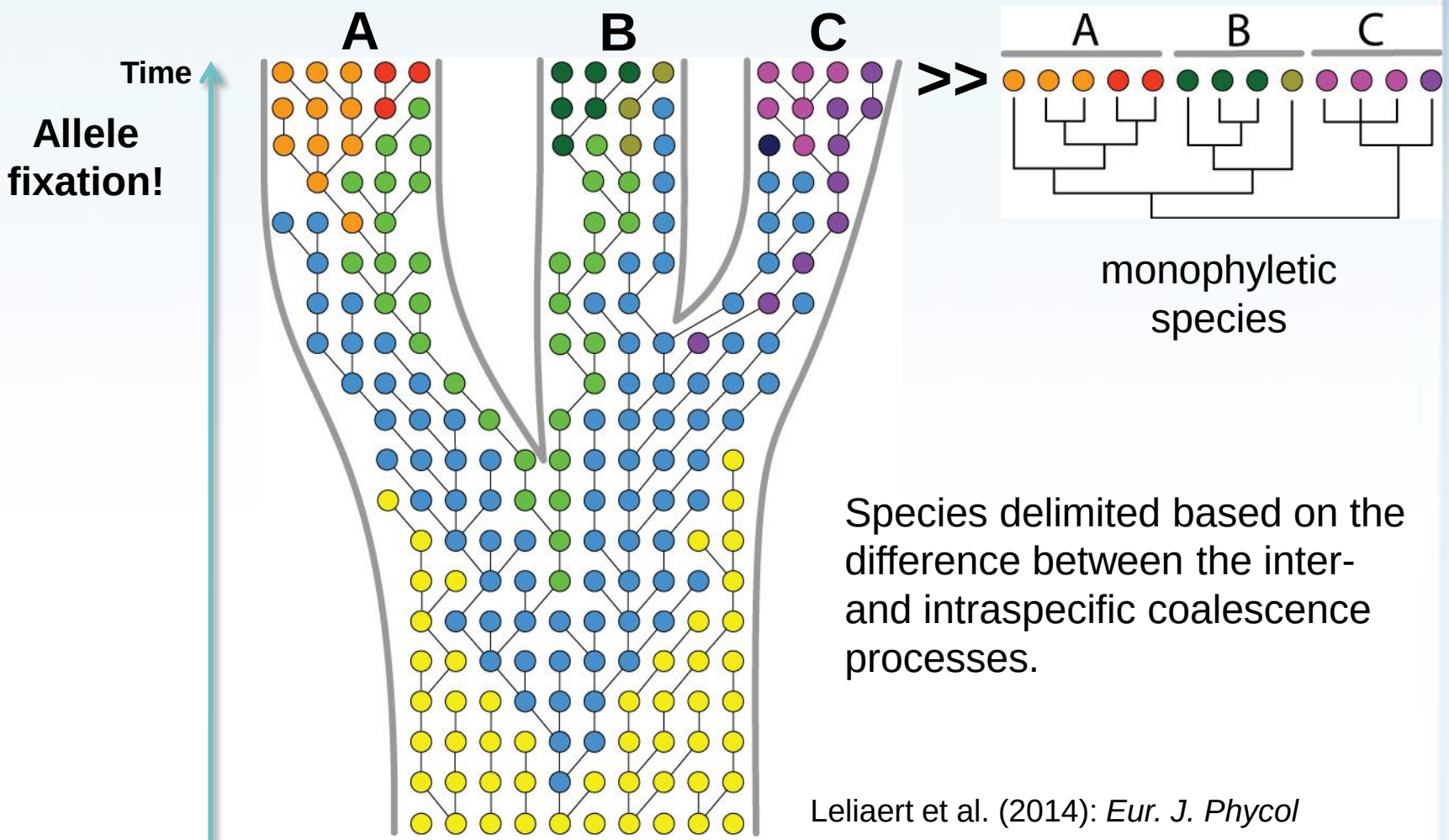
- Coalescence processes (Wright-Fisher)
  - during the coalescence, gene tree topologies resolve the species as polyphyletic, paraphyletic, and monophyletic



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

Coalescent theory

- Coalescence processes (Wright-Fisher)
  - during the coalescence, gene tree topologies resolve the species as polyphyletic, paraphyletic, and monophyletic

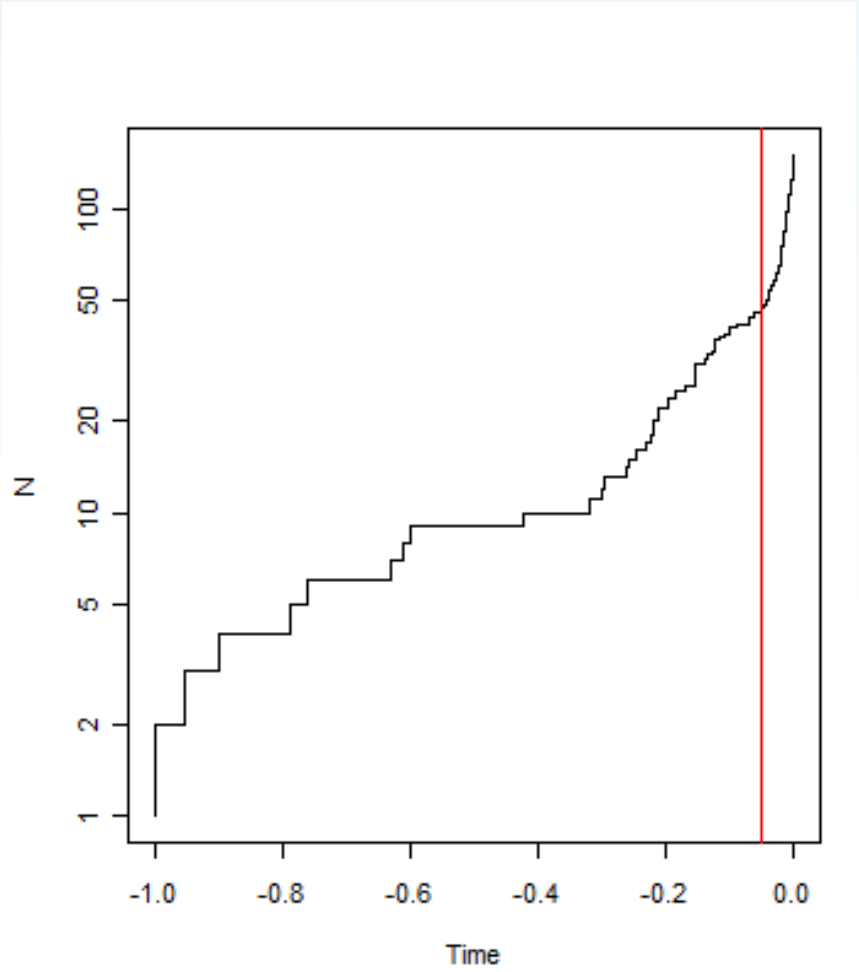


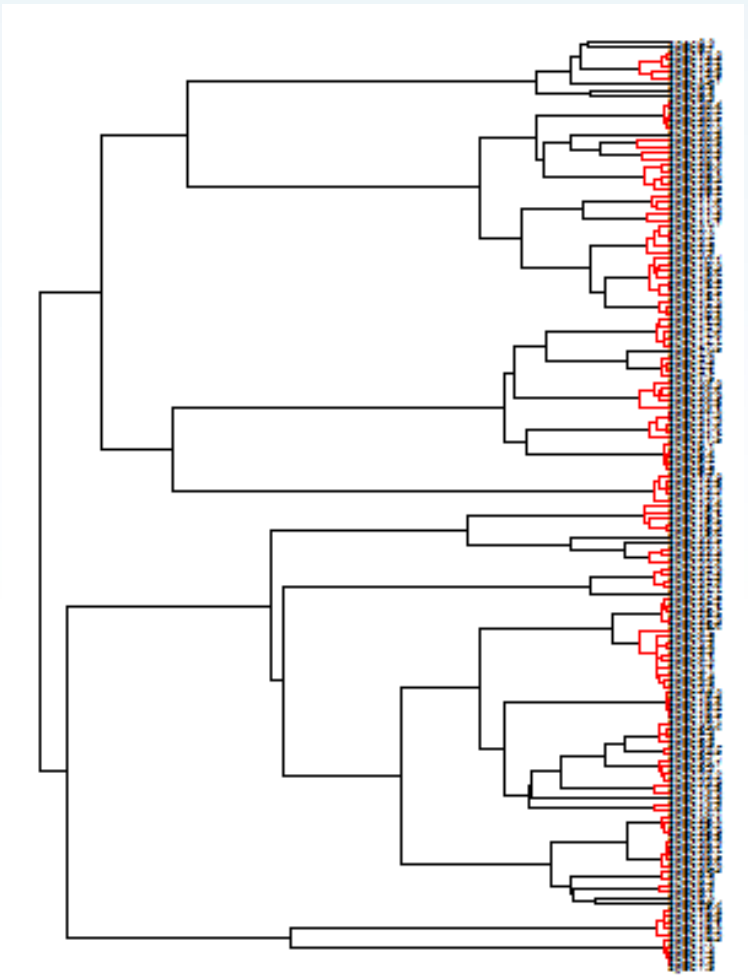
Species delimited based on the difference between the inter- and intraspecific coalescence processes.

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

GMYC method

- Different branching patterns within and among species
  - A combination of species diversification (Yule model) and intraspecific coalescence models

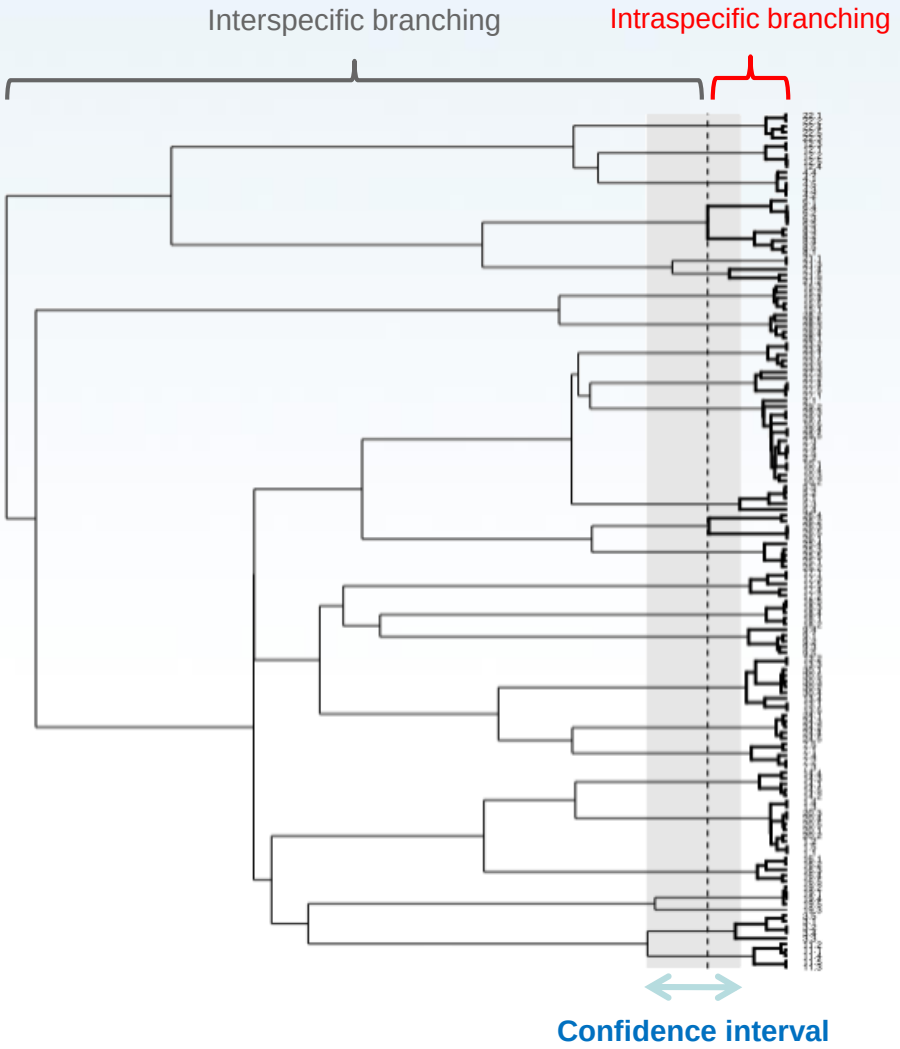




GMYC method

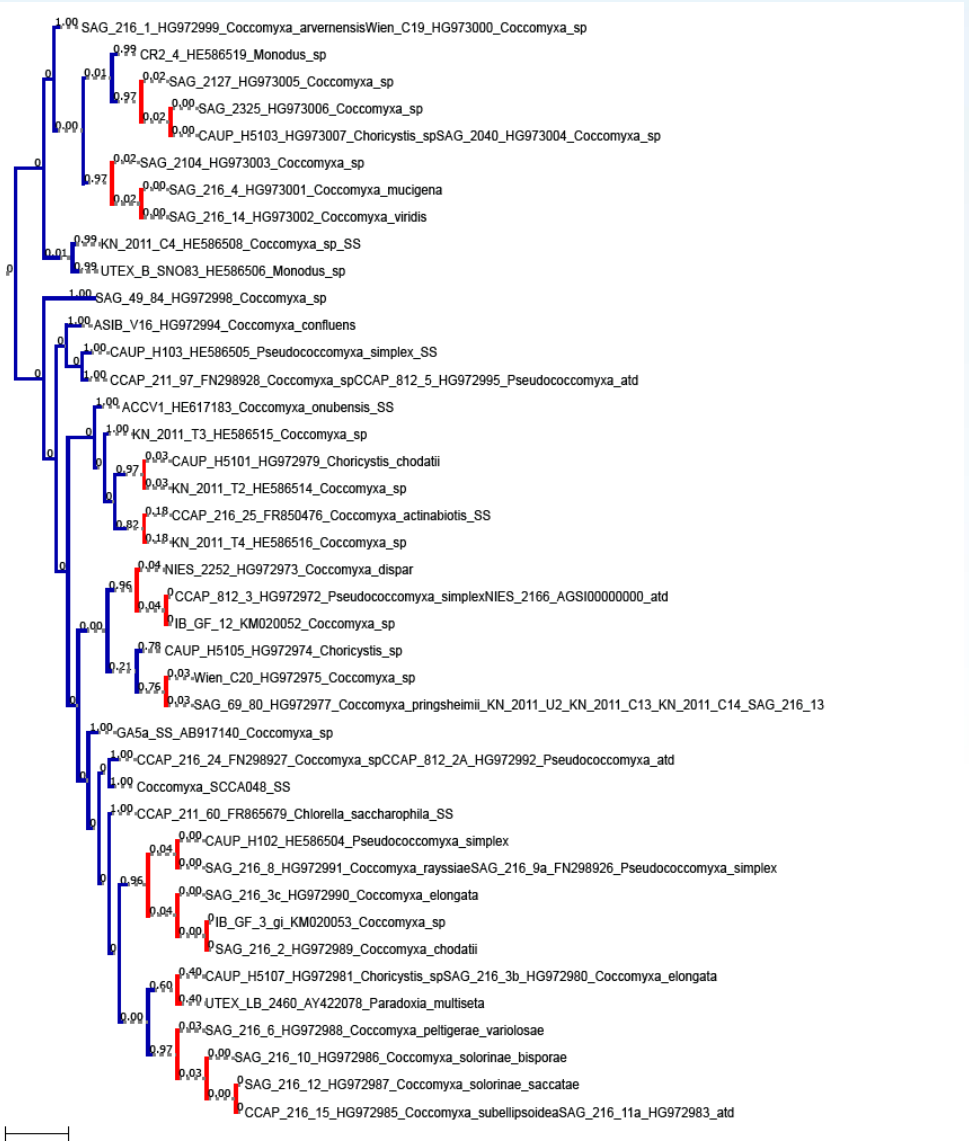
- Different branching patterns within and among species
  - ➢ statistical test, confidence interval

```
>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: 46
confidence interval: 42-56

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



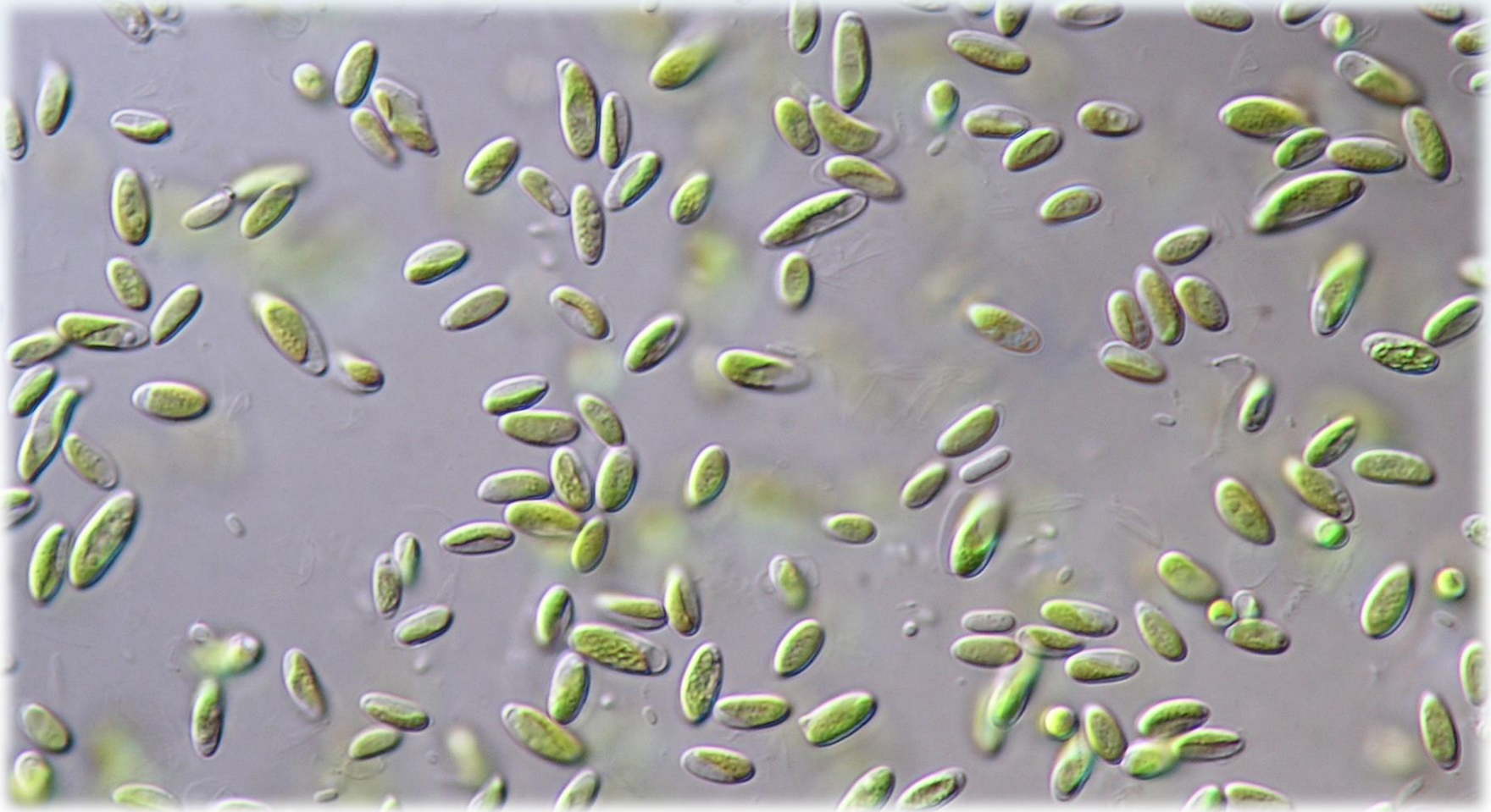
DNA-based species delimitation methods

Reference	Existing species	Species after delimitation	No. of samples	No. of loci	Discovery approaches	Validation approaches
Avila *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	None

Carstens et al. (2013): *Mol. Ecol.* 22: 4369-83

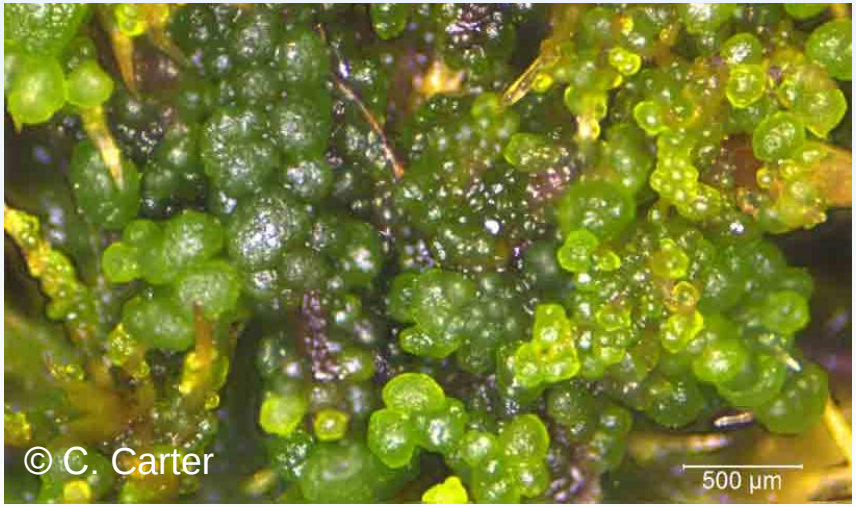
Species delimitation in *Coccomyxa*

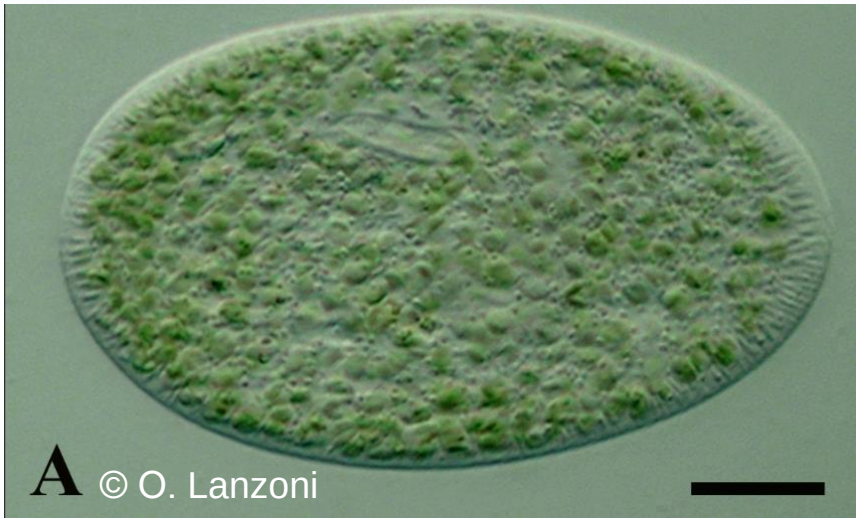
- *Coccomyxa* Schmidle (1901)
- Algaebase: actually 25 accepted taxa



Species delimitation in *Coccomyxa*

- *Coccomyxa* Schmidle (1901)
- Free-living, lichen photobionts, cilliate symbionts





Species delimitation in *Coccomyxa*

- *Coccomyxa* Schmidle (1901)
- Even parasites

Vol. 81: 231–240, 2008
doi: 10.3354/dao01956

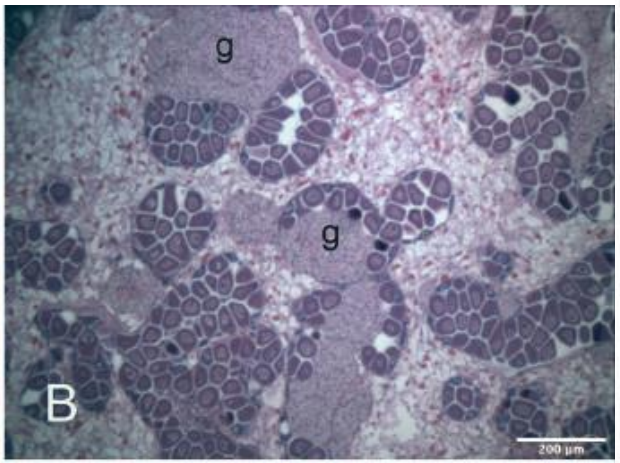
DISEASES OF AQUATIC ORGANISMS
Dis Aquat Org

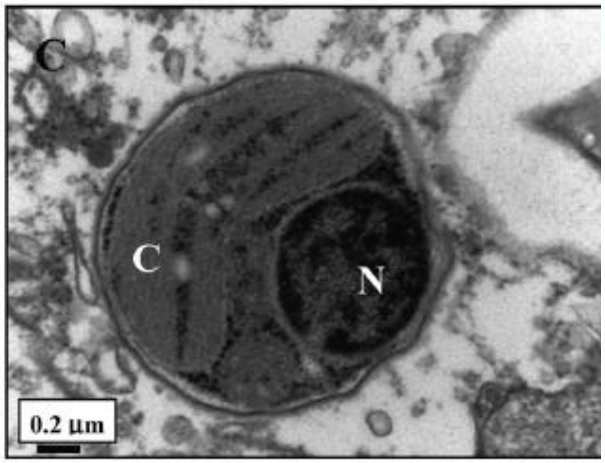
Published September 24

OPEN ACCESS

Phylogenetic and morphological characterisation of the green algae infesting blue mussel *Mytilus edulis* in the North and South Atlantic oceans

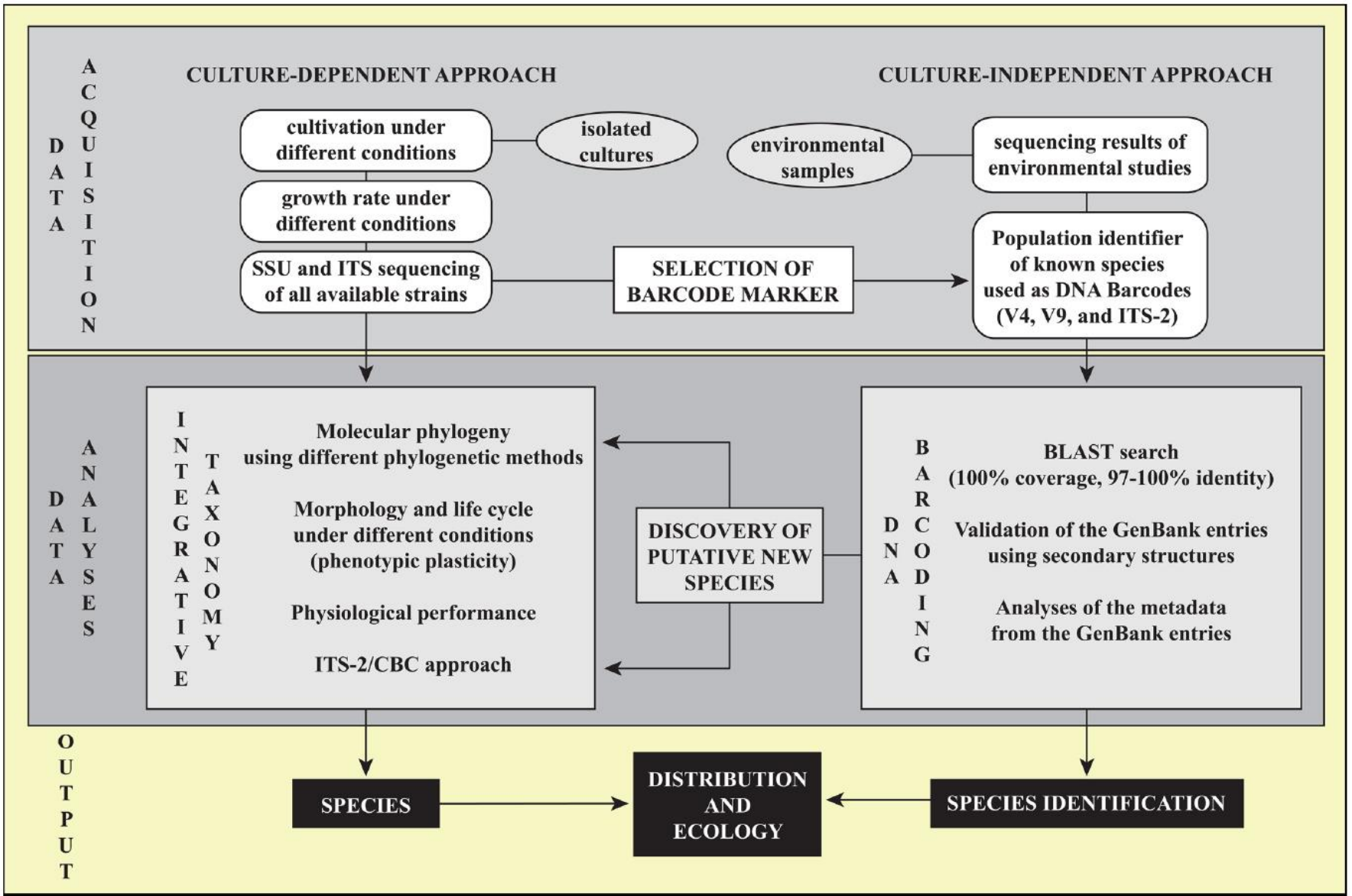
Francisco Rodríguez[1,5], Stephen W. Feist[2], Laure Guillou[1], Lisbeth S. Harkestad[3], Kelly Bateman[2], Tristan Renault[4], Stein Mortensen[3,*]





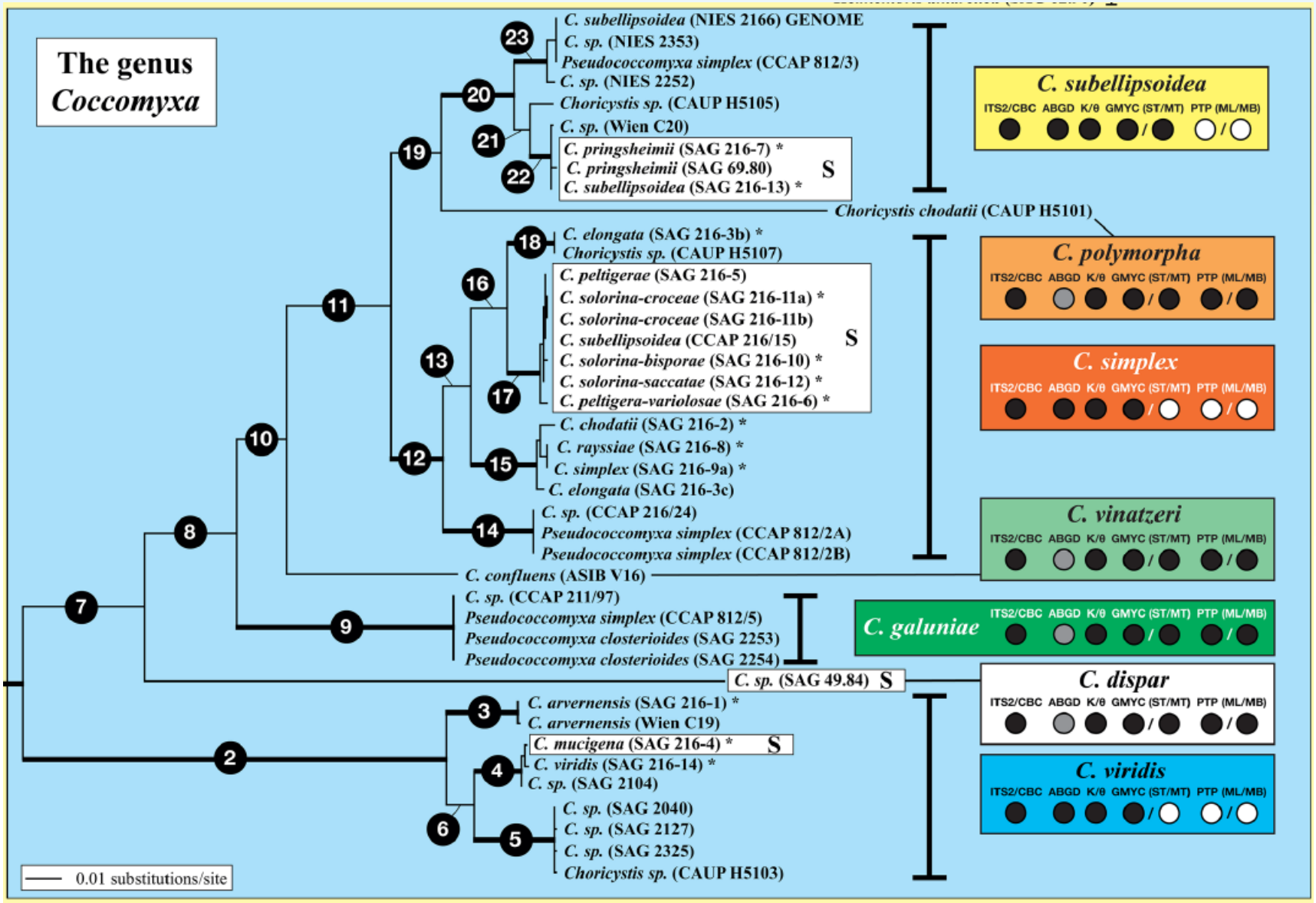
Coccomyxa

- Darienko et al. (2015): Reducing the number of species from 25 to 7

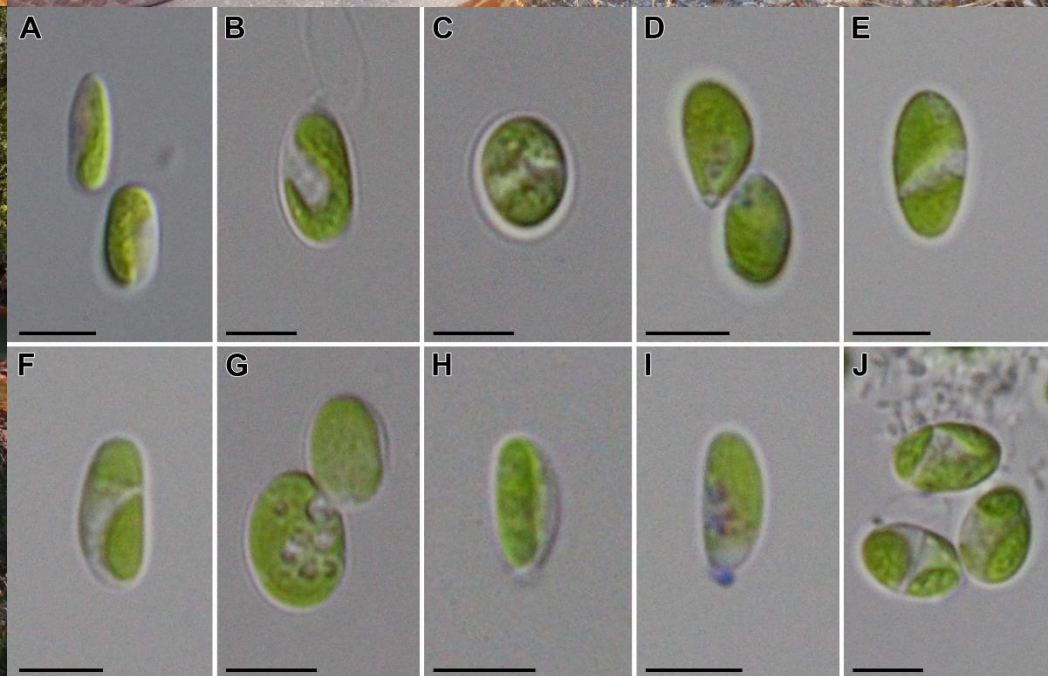


Coccomyxa

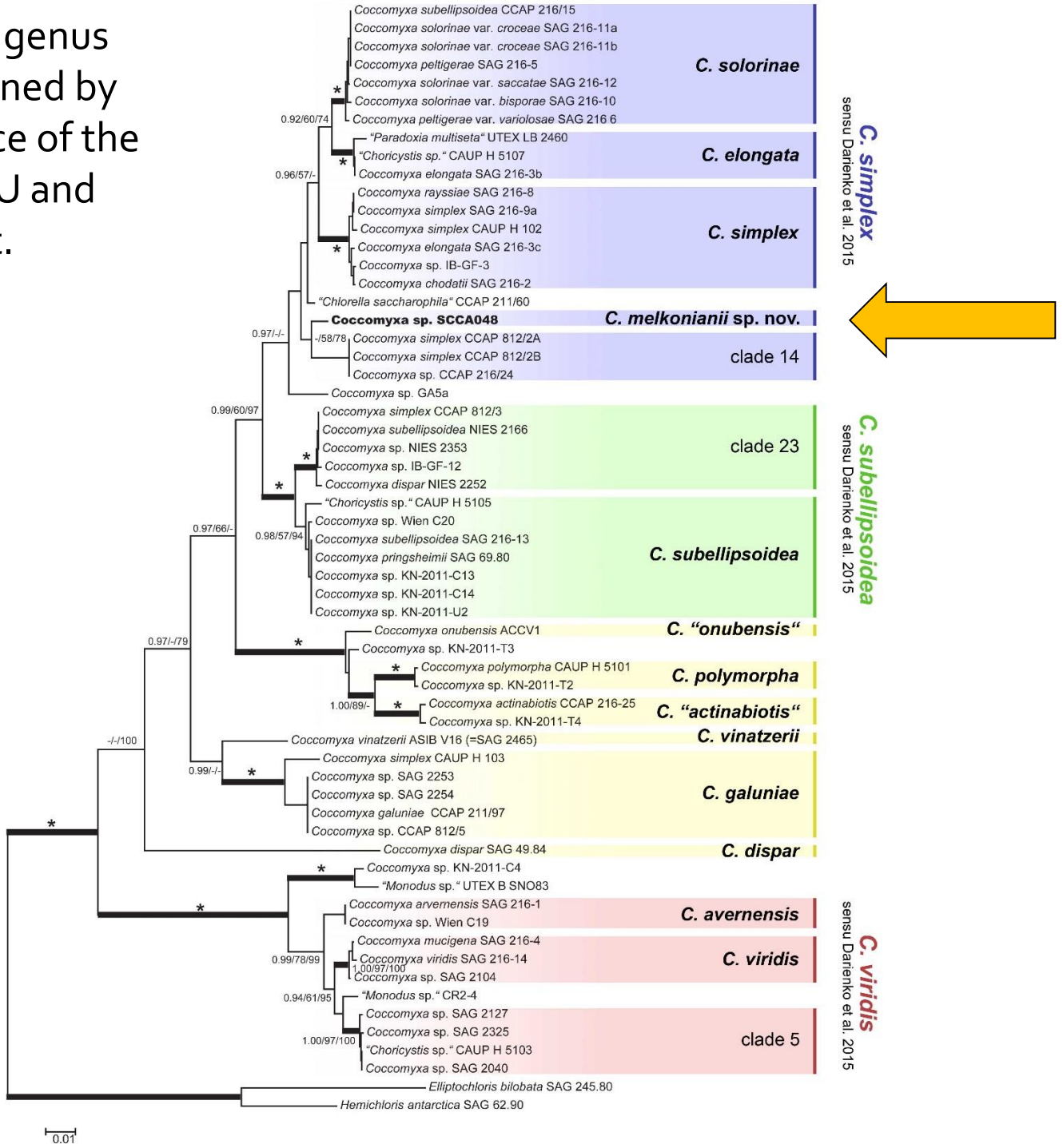
- Darienko et al. (2015): Reducing the number of species from 25 to 7

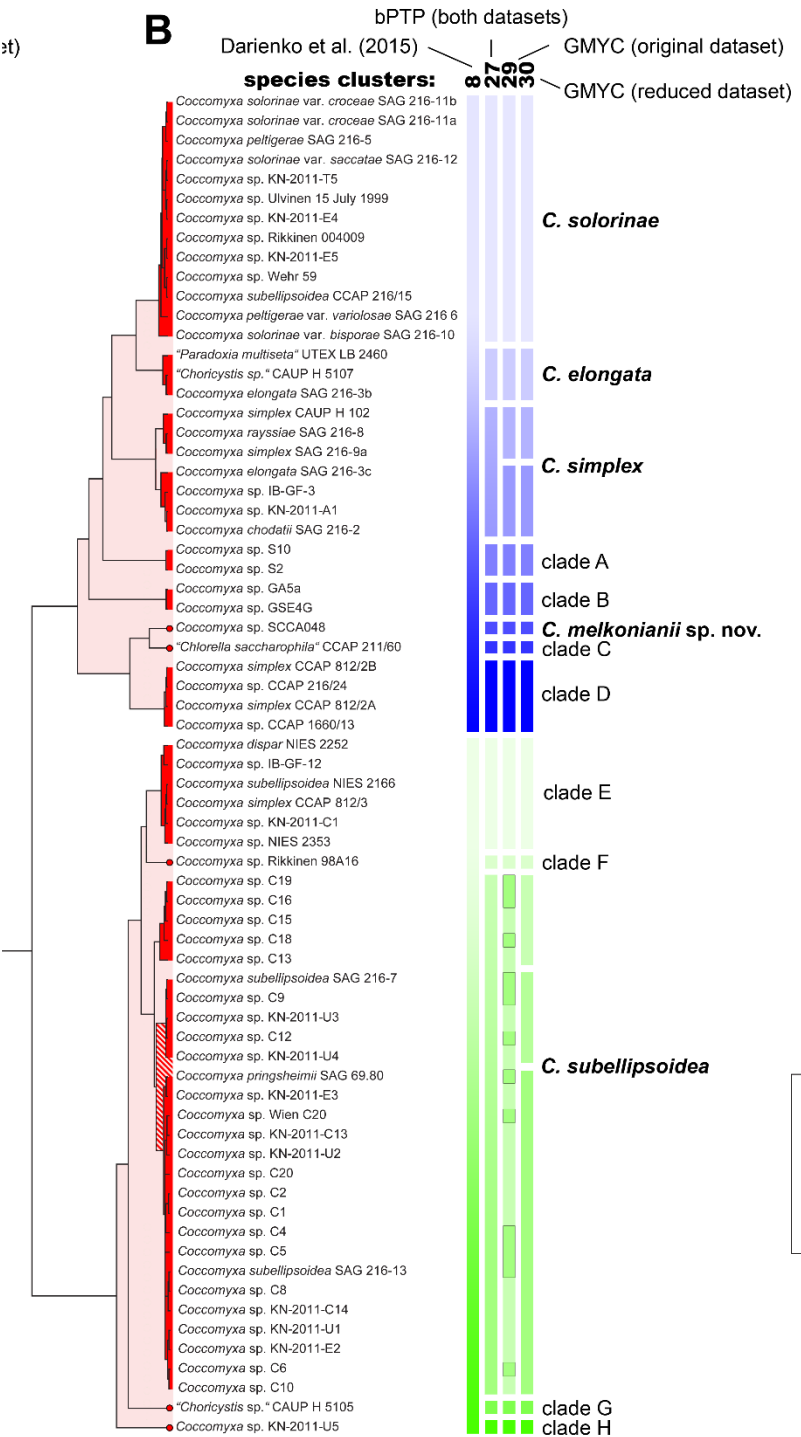


•SCCA048 – a recently isolated *Coccomyxa* from highly poluted Rio Irvi (high contents of Cd, Co, Fe, Mn, Zn)

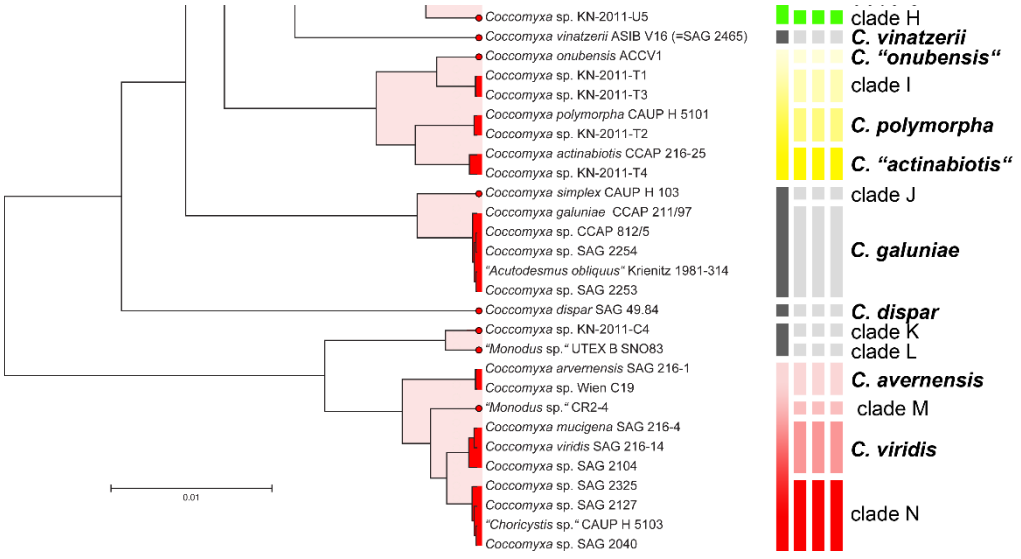


- Phylogeny of the genus Coccomyxa obtained by Bayesian inference of the concatenated SSU and ITS rDNA dataset.





- Three DNA species delimitation analyses performed (GMYC, bPTP, STACEY)
- Analysis of the broader dataset resulted in a very different species scenario (27-30 vs 7 species)

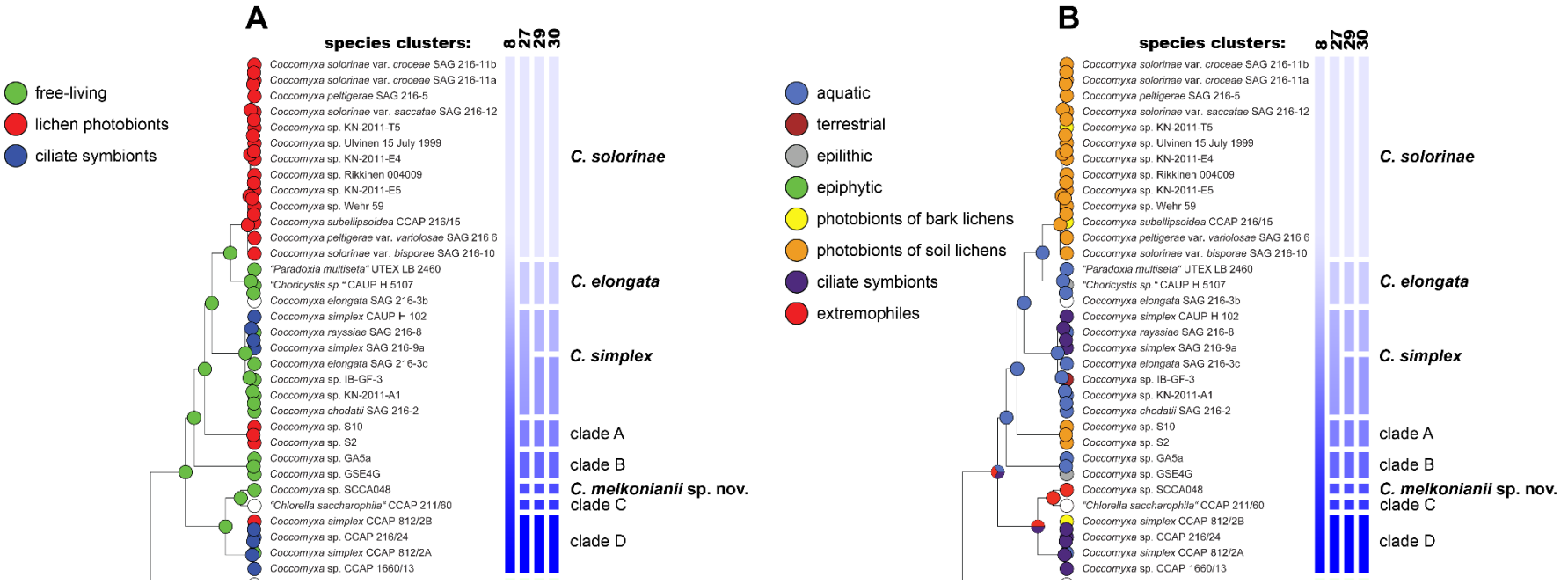


- High ecological significance of 27-30 species scenarios.

Table 3. Ecological differentiation of putative species clusters, based on the 98-taxa dataset.

Statistical test	Species scenario	Ecological traits		
		Living stage	Resilience	Habitat type
χ^2	8 species	0.034	0.011	**<0.001**
	27 species	**<0.001**	**<0.001**	**<0.001**
	29 species	**<0.001**	**<0.001**	**<0.001**
	30 species	**<0.001**	**<0.001**	**<0.001**
Mantel test	8 species	0.695	0.863	0.147
	27 species	**<0.001**	0.815	0.059
	29 species	0.054	0.807	0.652
	30 species	0.157	0.779	0.737

- E.g., differences in living stage can be well applied to delimit species within the *Coccomyxa simplex* s.l.



Species delimitation

- The modern DNA-based species delimitation methods represent a powerful tool to predict species boundaries.
- However, they are very sensitive to data sampling.
- To determine the species correctly, a combination of various methodological approaches should be applied.
- Ecological data can represent a very useful tool to delimit morphologically often undistinguishable species.
- Ecological characterization should receive much more attention than at present.

Thank you for your attention