Evolution des Eucaryotes

L'éclairage de la phylogénie moléculaire

Denis BAURAIN

Unité de Phylogénomique des Eucaryotes
Département des Sciences de la Vie / ULg

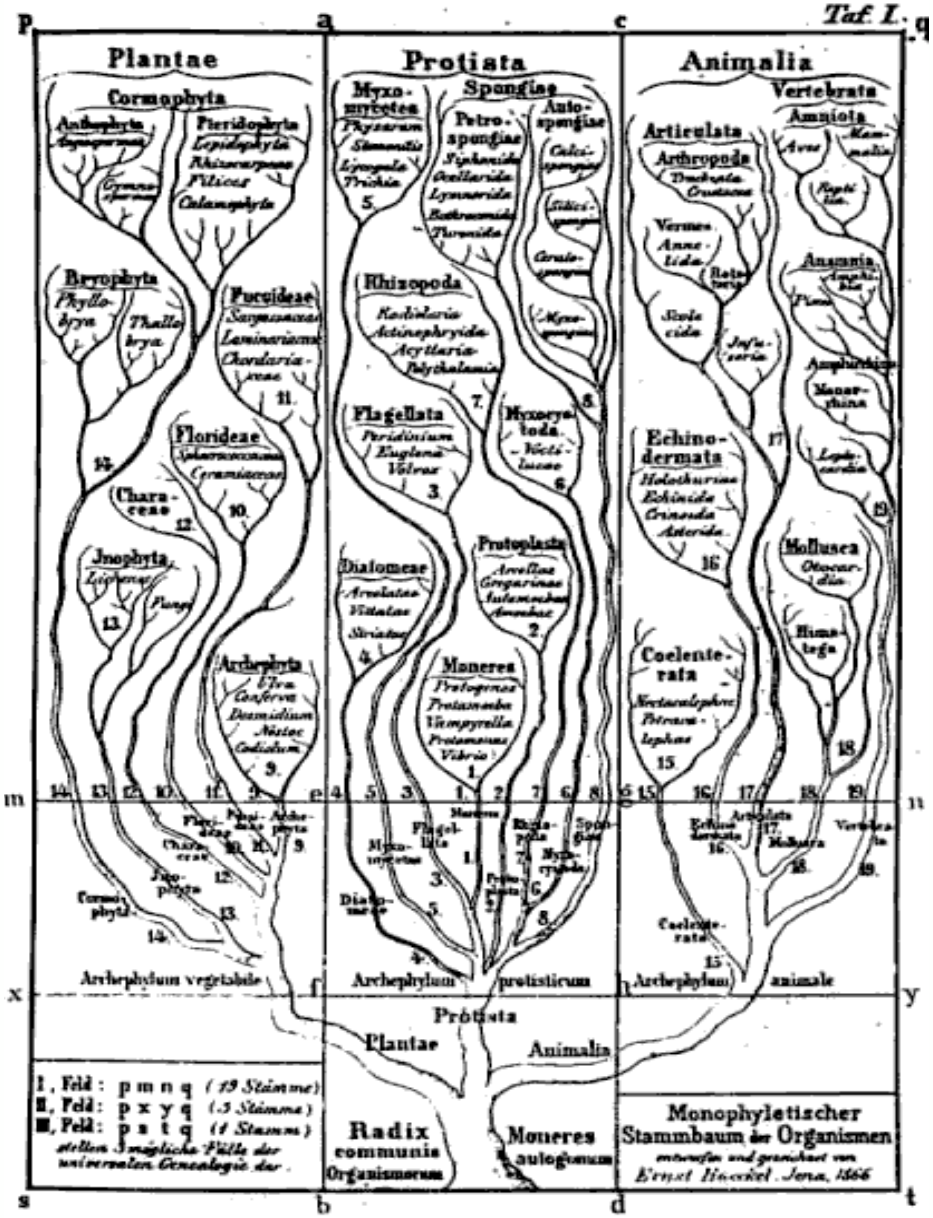
Collège Belgique / Palais des Académies
23 octobre 2014

Phylogénie moléculaire Grandeur et décadence des Archezoa

Phylogénie

Ch. Darwin

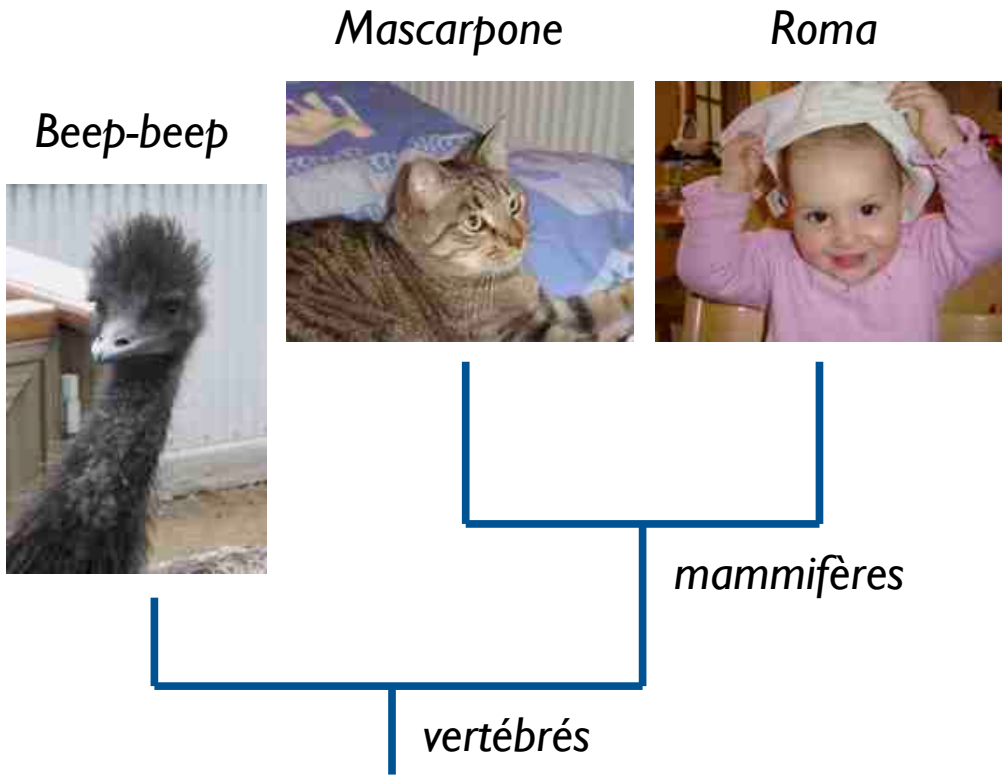
E. Haeckel



- domaine de la biologie étudiant les relations de parenté entre organismes actuels
- fondée sur la théorie de l'évolution des espèces de Ch. Darwin (1859)
- utilisant l'arbre à la fois comme métaphore (E. Haeckel, 1866) et outil de représentation

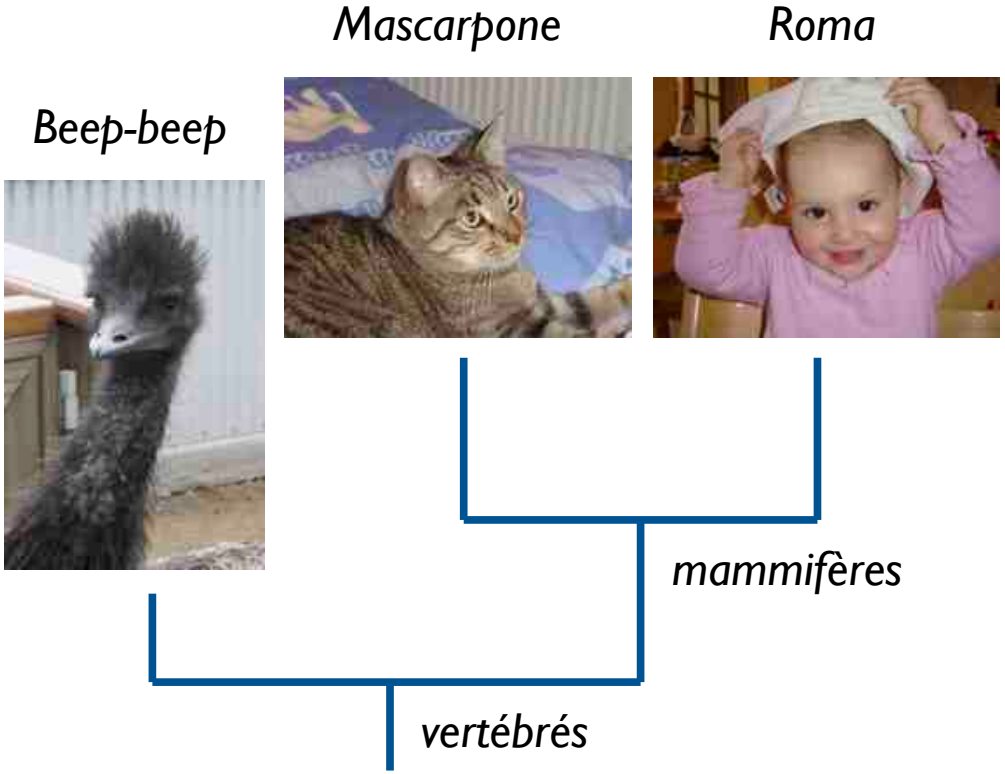
Phylogénie classique

- Caractères morphologiques
  - ouverts aux interpretations
  - pas toujours disponibles (ex. microbes)
  - souvent incommensurables entre lignées



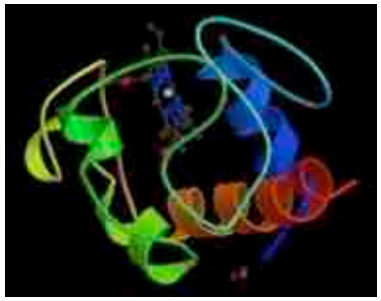
Phylogénie classique

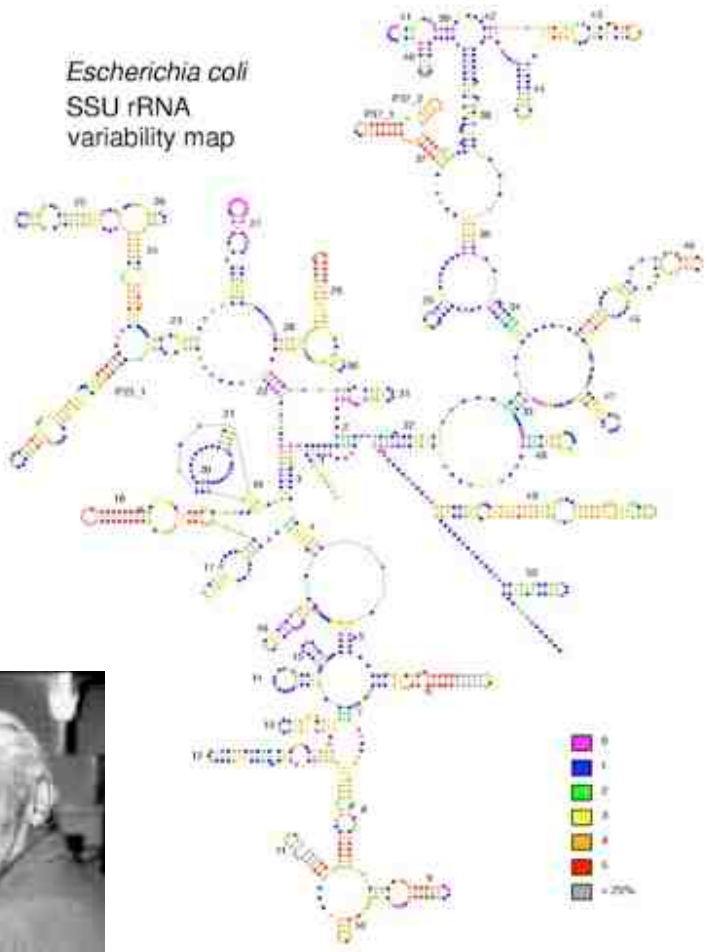
- Caractères morphologiques
  - ouverts aux interpretations
  - pas toujours disponibles (ex. microbes)
  - souvent incommensurables entre lignées



2014

Phylogénie moléculaire

- Caractères moléculaires
  - protéines
    - Fitch & Margoliash (1967)
    - 
      cytochrome *c* de cheval
  - DNA
  - SSU rRNA (16S/18S)
    - Carl Woese (1977)
    - molécule essentielle
    - structure universelle
    - très ancien




Carl Woese


Walter M. Fitch

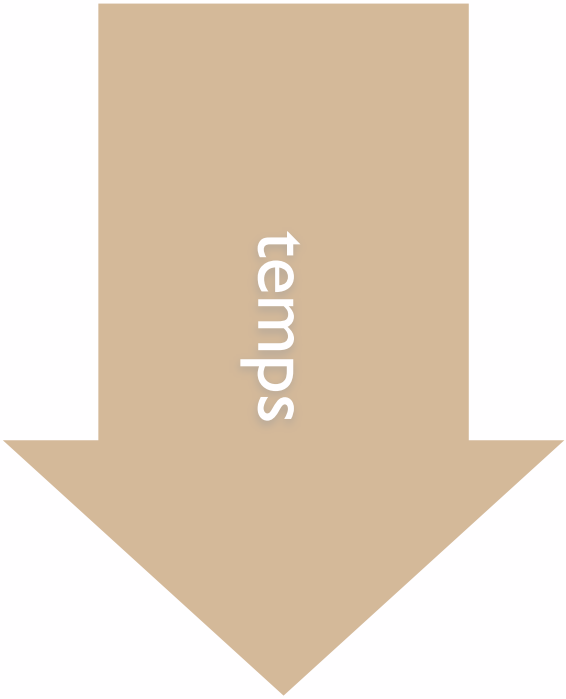
rRNA repris de Y. Van de Peer

Phylogénie moléculaire

principe de base / raisonnement de parcimonie

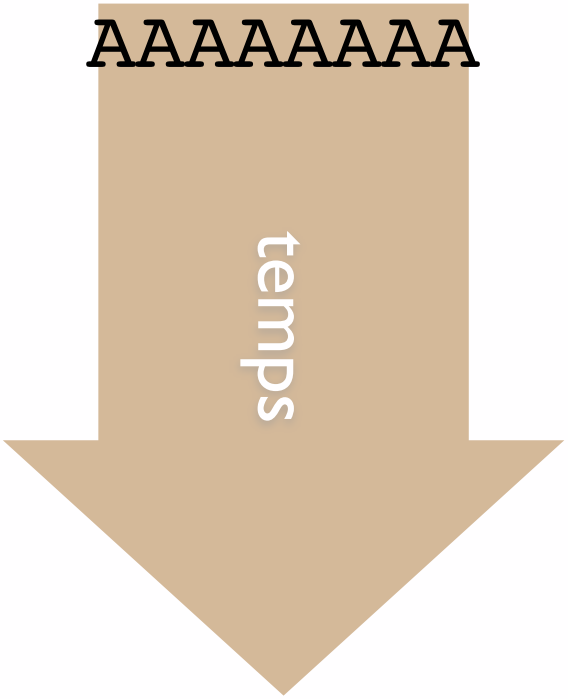
Phylogénie moléculaire

principe de base / raisonnement de parcimonie



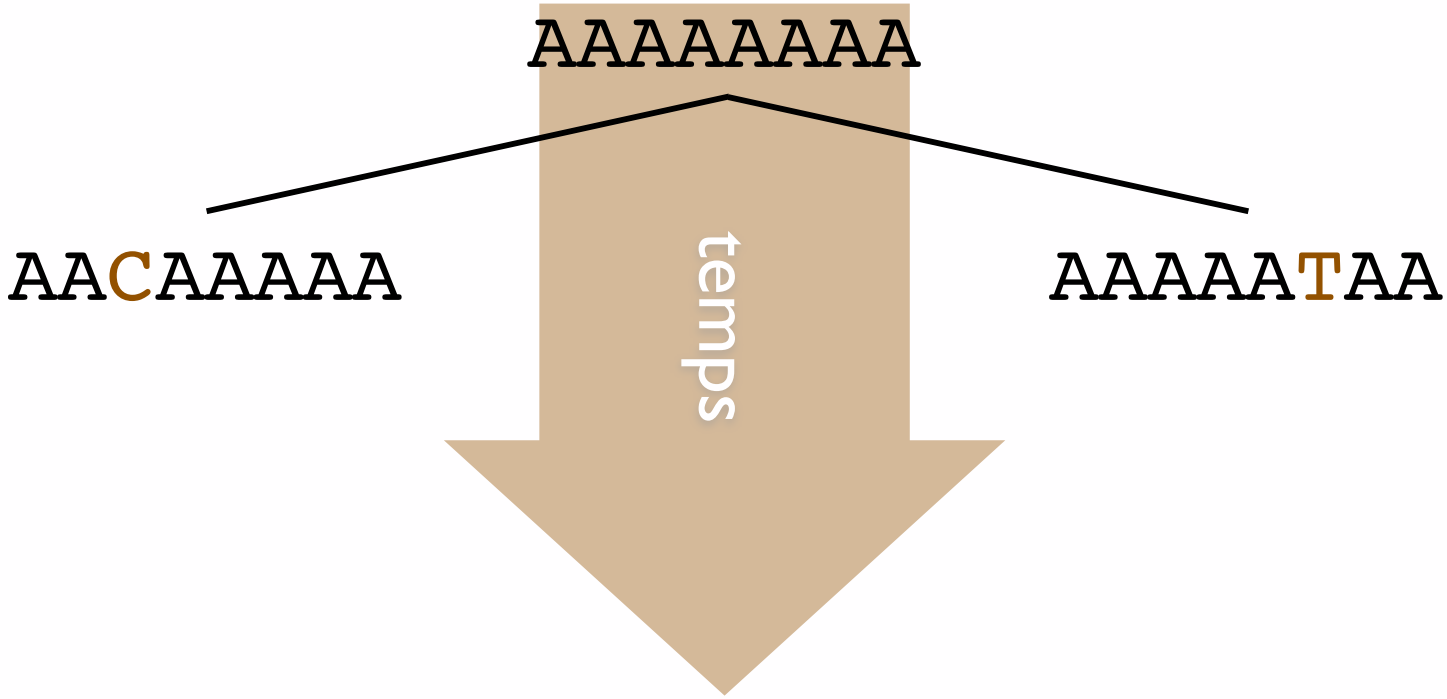
Phylogénie moléculaire

principe de base / raisonnement de parcimonie



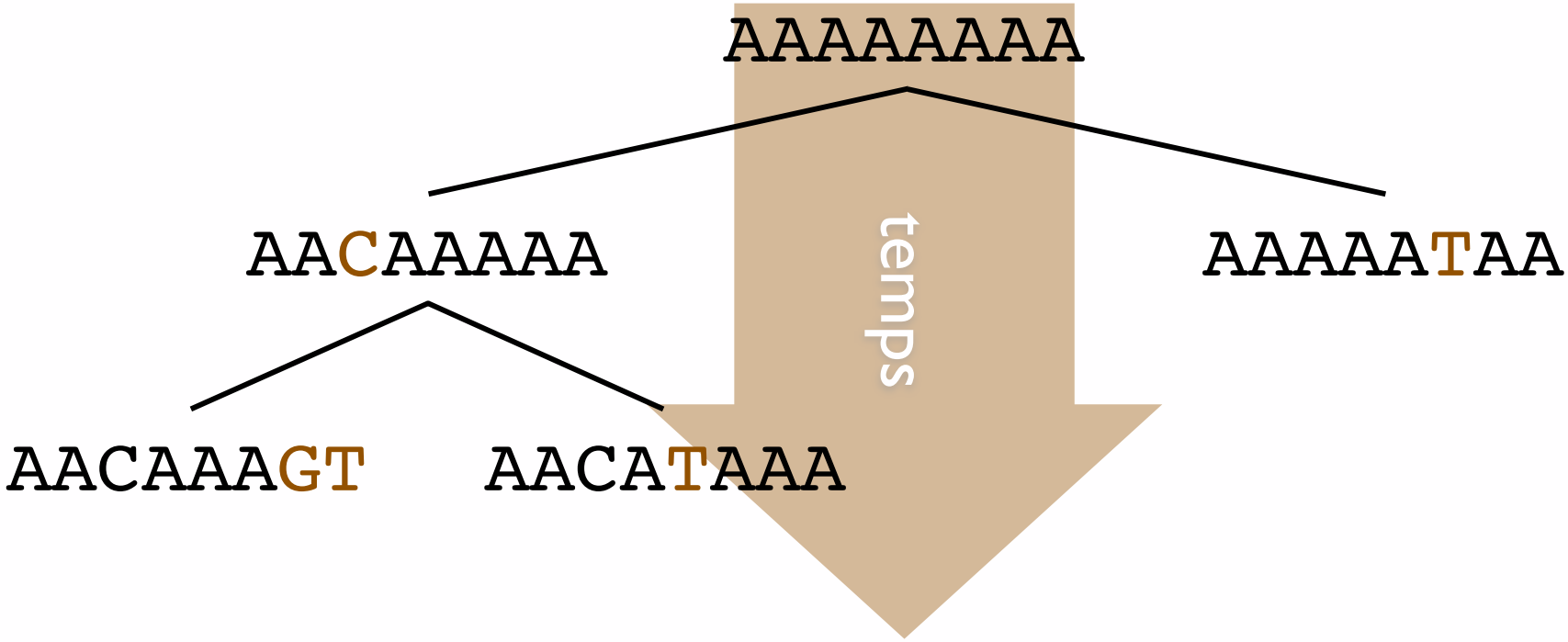
Phylogénie moléculaire

principe de base / raisonnement de parcimonie



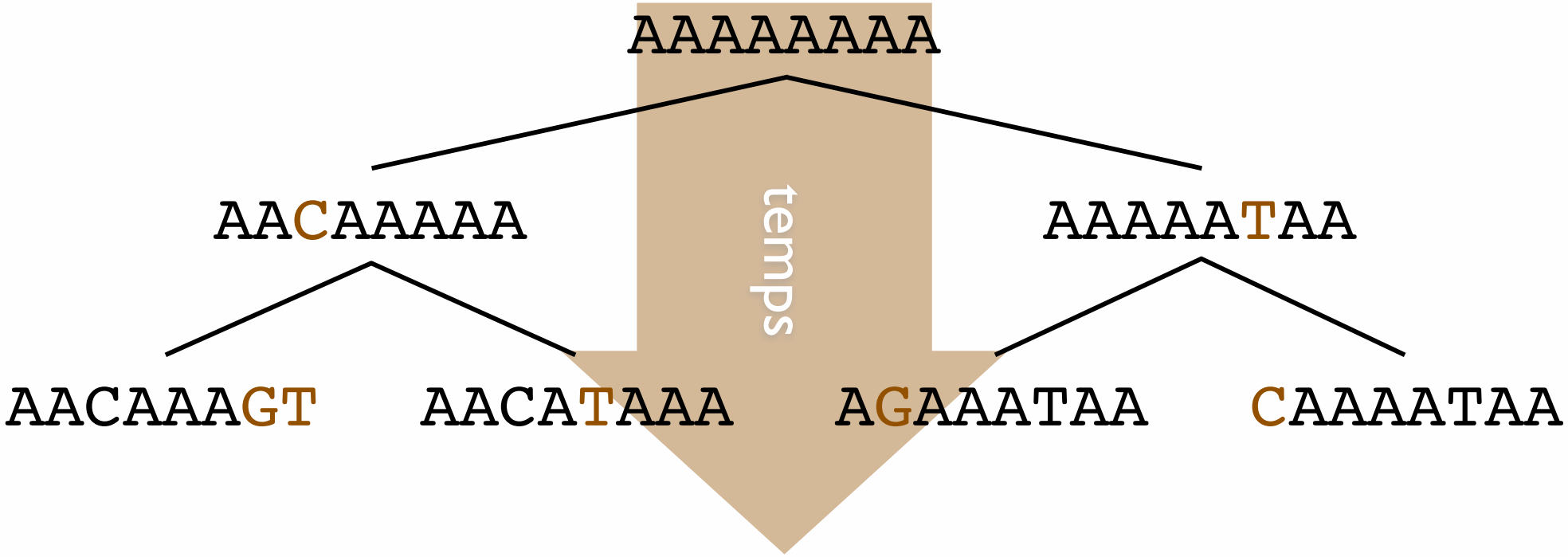
Phylogénie moléculaire

principe de base / raisonnement de parcimonie



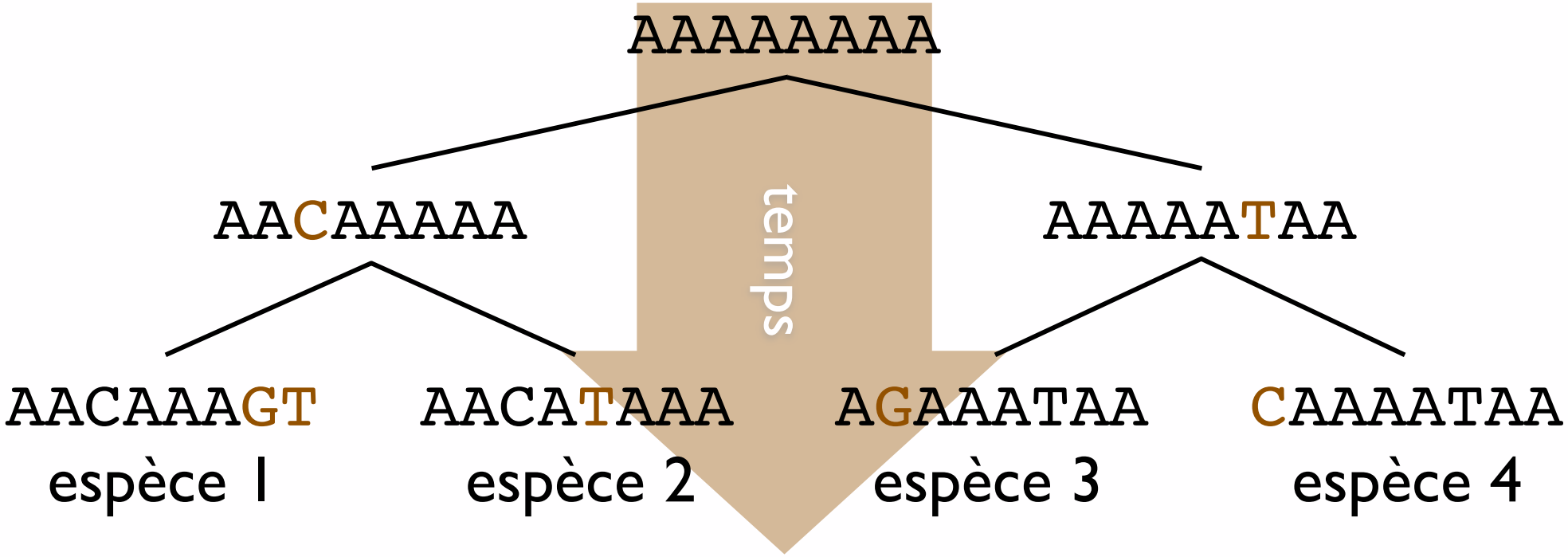
Phylogénie moléculaire

principe de base / raisonnement de parcimonie



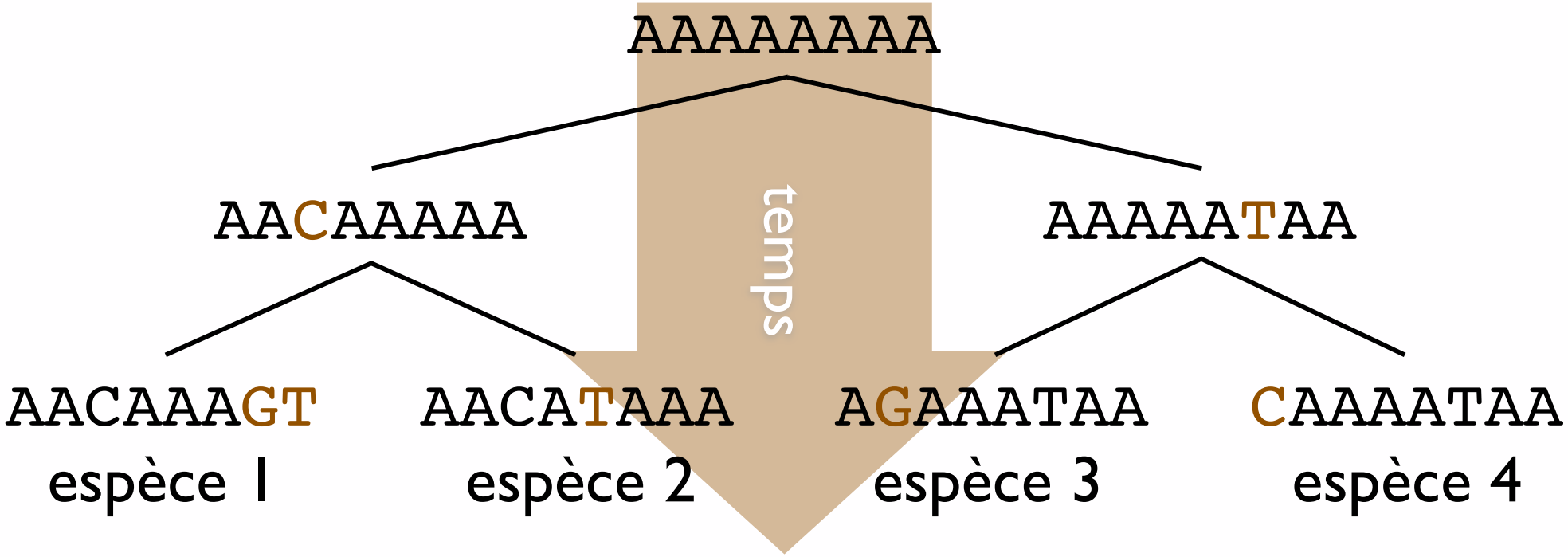
Phylogénie moléculaire

principe de base / raisonnement de parcimonie



Phylogénie moléculaire

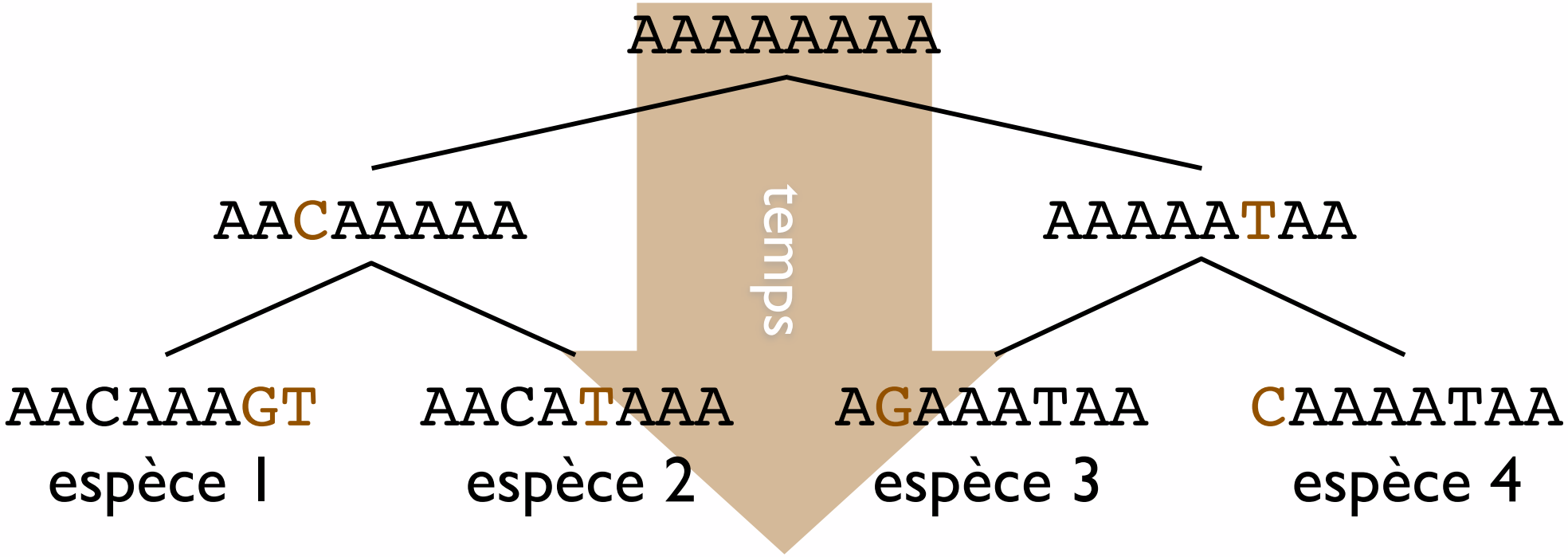
principe de base / raisonnement de parcimonie




```
1  AACAAAGT
2  AACATAAA
3  AGAAATAA
4  CAAAATAA
```

Phylogénie moléculaire

principe de base / raisonnement de parcimonie

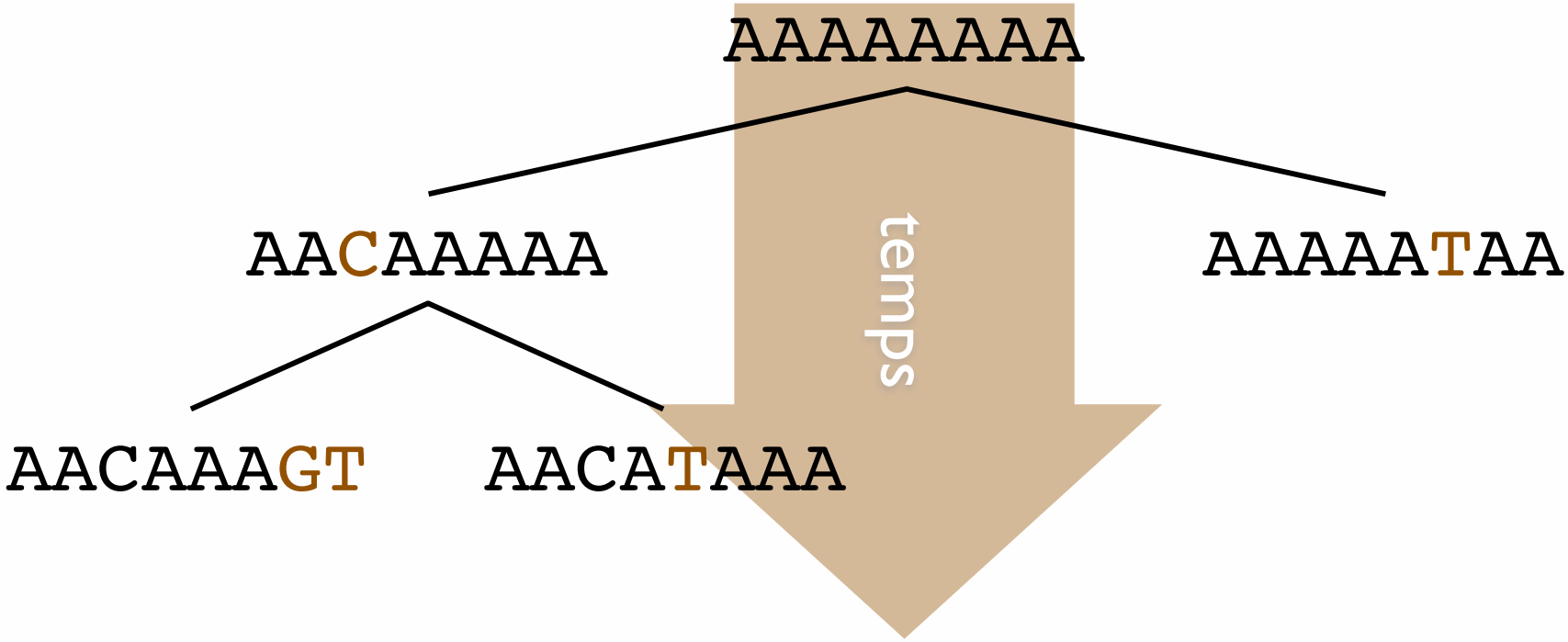


1 AACAAAGT
2 AACATAAA
3 AGAAATAA
4 CAAAATAA

Les substitutions sont assez rares.
La majorité des différences de séquence observées entre espèces sont donc héritées.

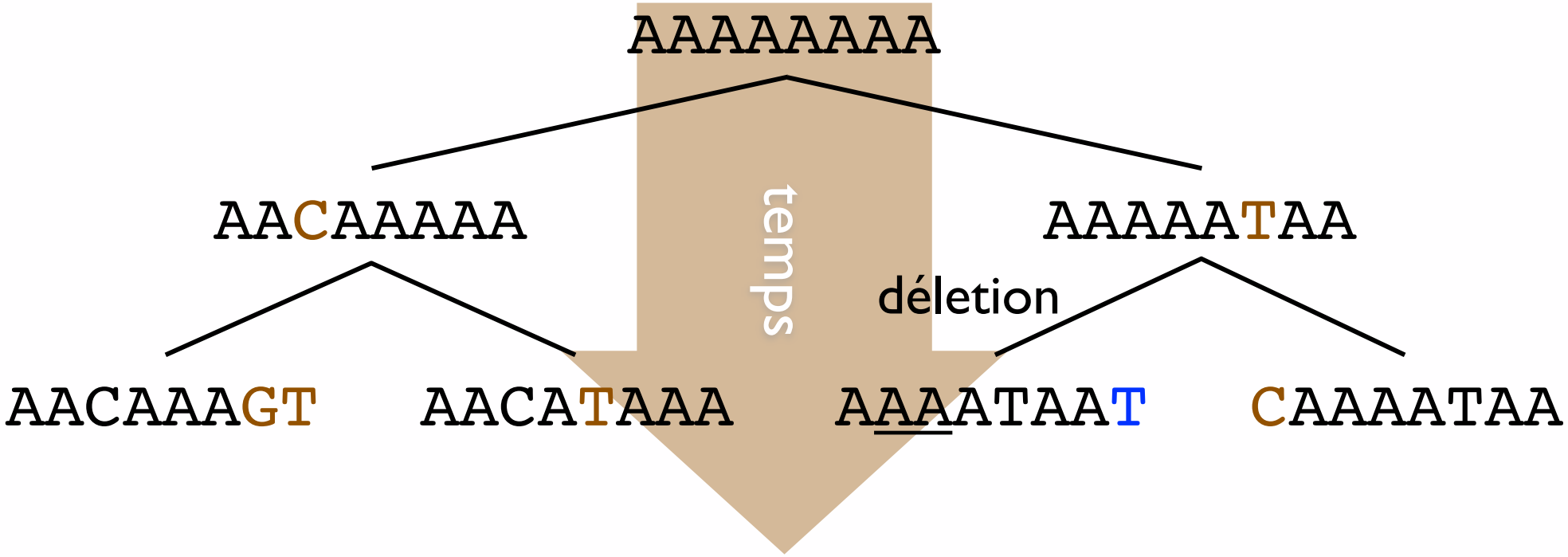
Phylogénie moléculaire

nécessité d'aligner les séquences



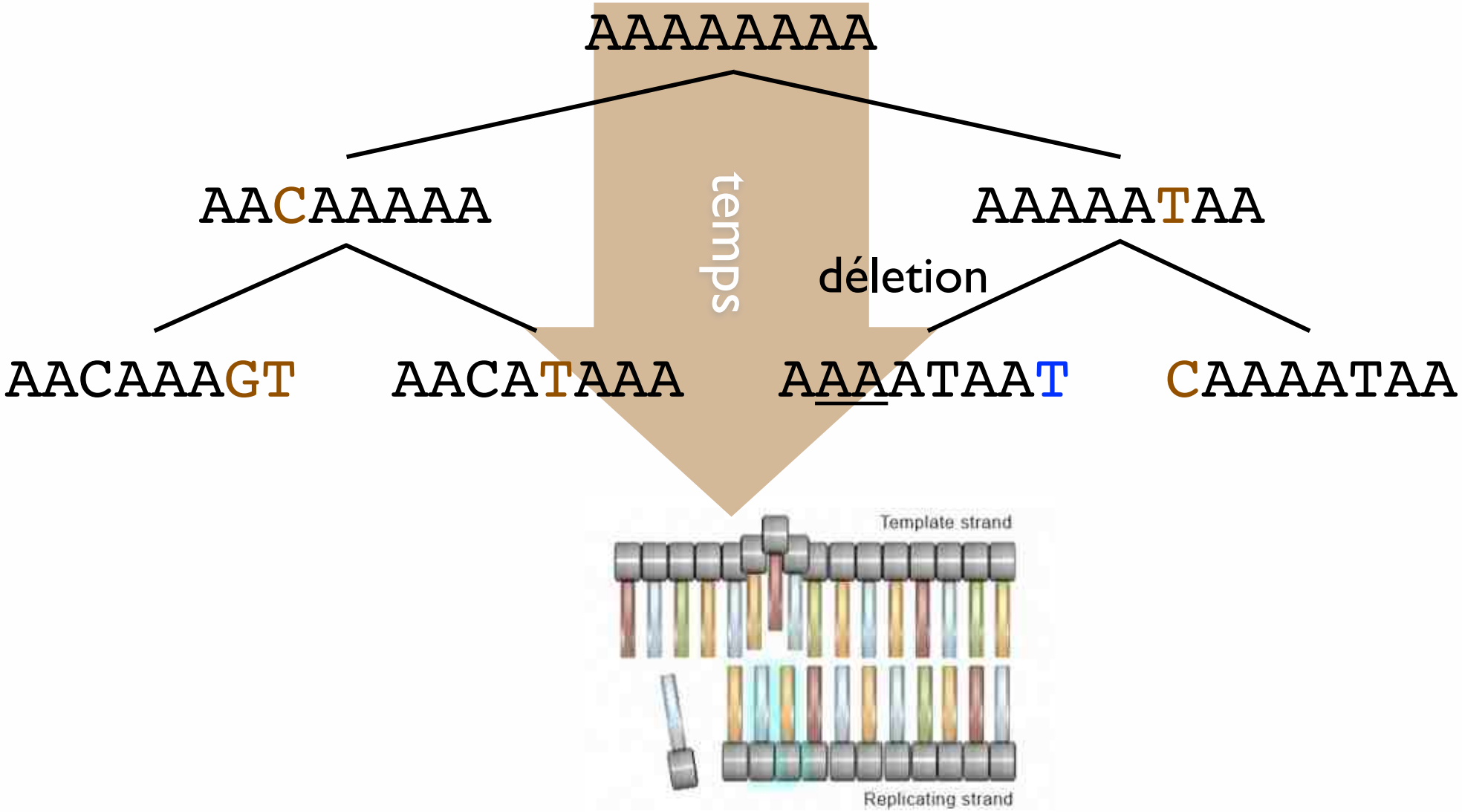
Phylogénie moléculaire

nécessité d'aligner les séquences



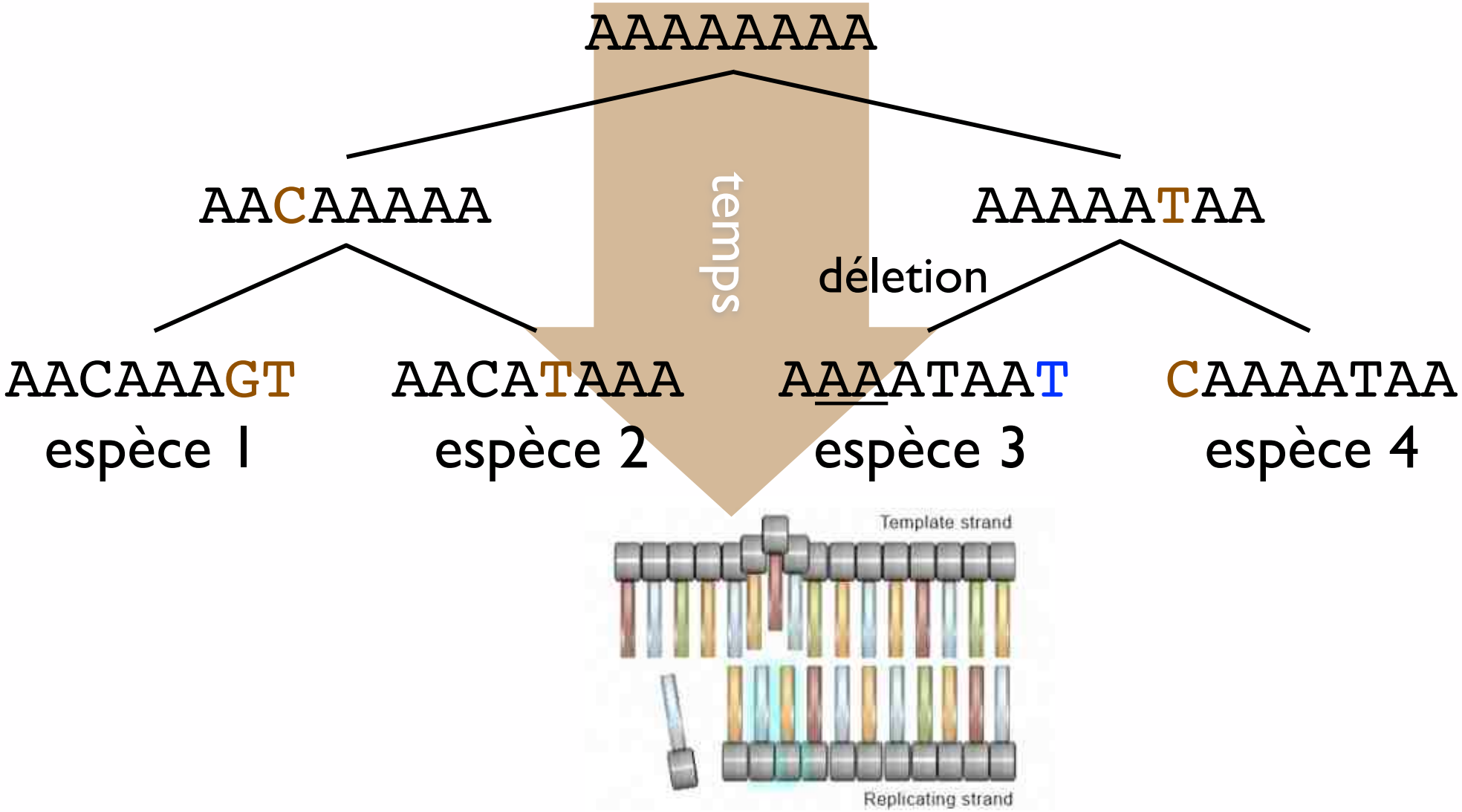
Phylogénie moléculaire

nécessité d'aligner les séquences



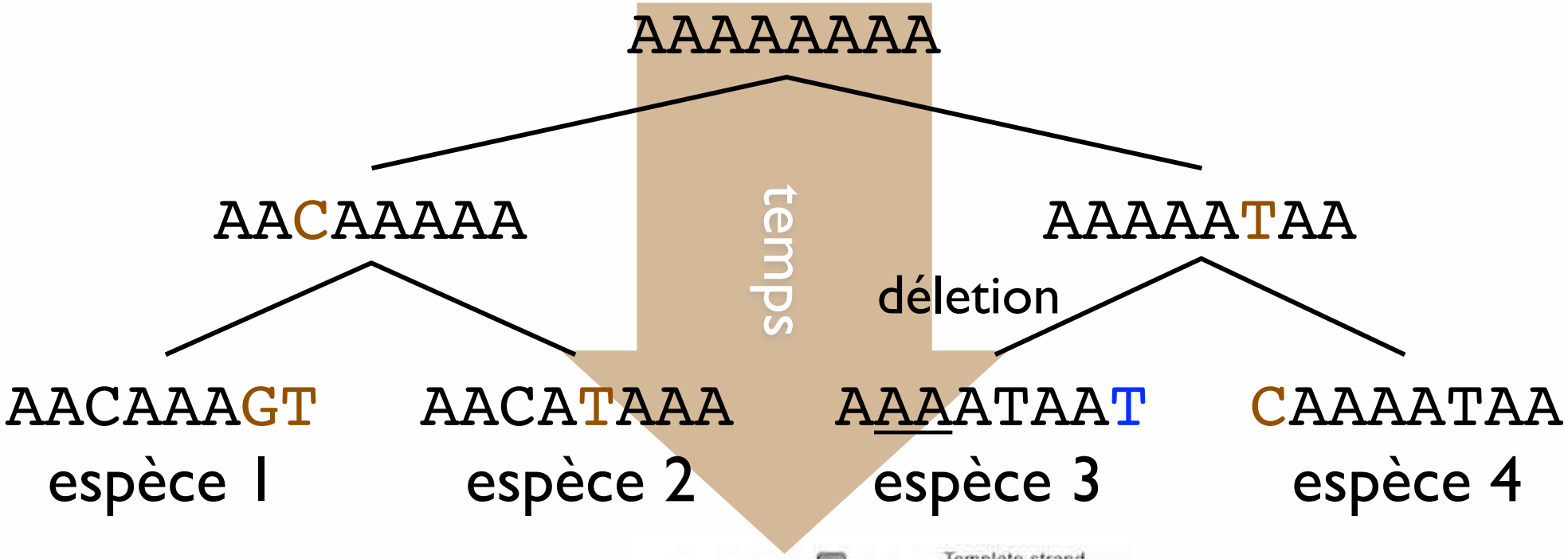
Phylogénie moléculaire

nécessité d'aligner les séquences

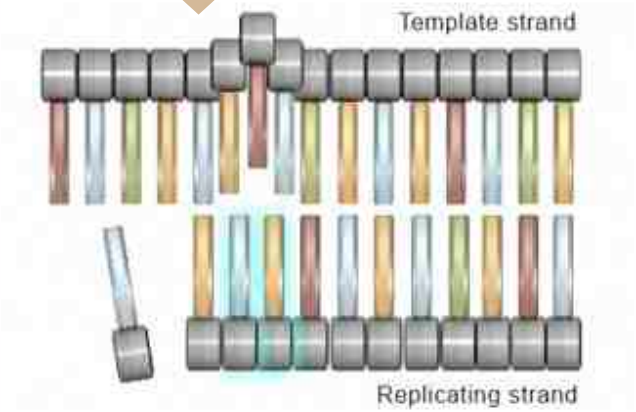


Phylogénie moléculaire

nécessité d'aligner les séquences

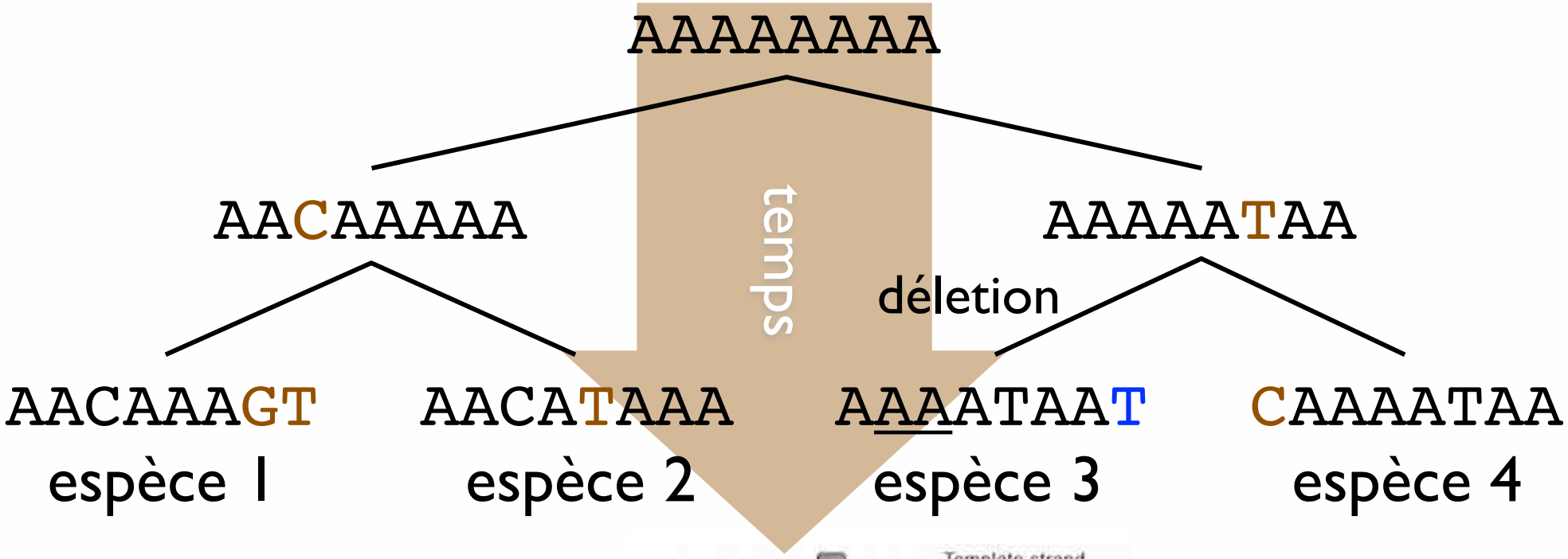


1 AACAAAGT
2 AACATAAA
3 AAAATAAT
4 CAAAATAA

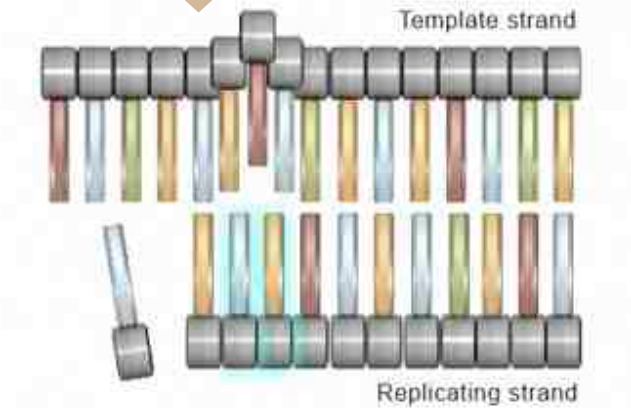


Phylogénie moléculaire

nécessité d'aligner les séquences



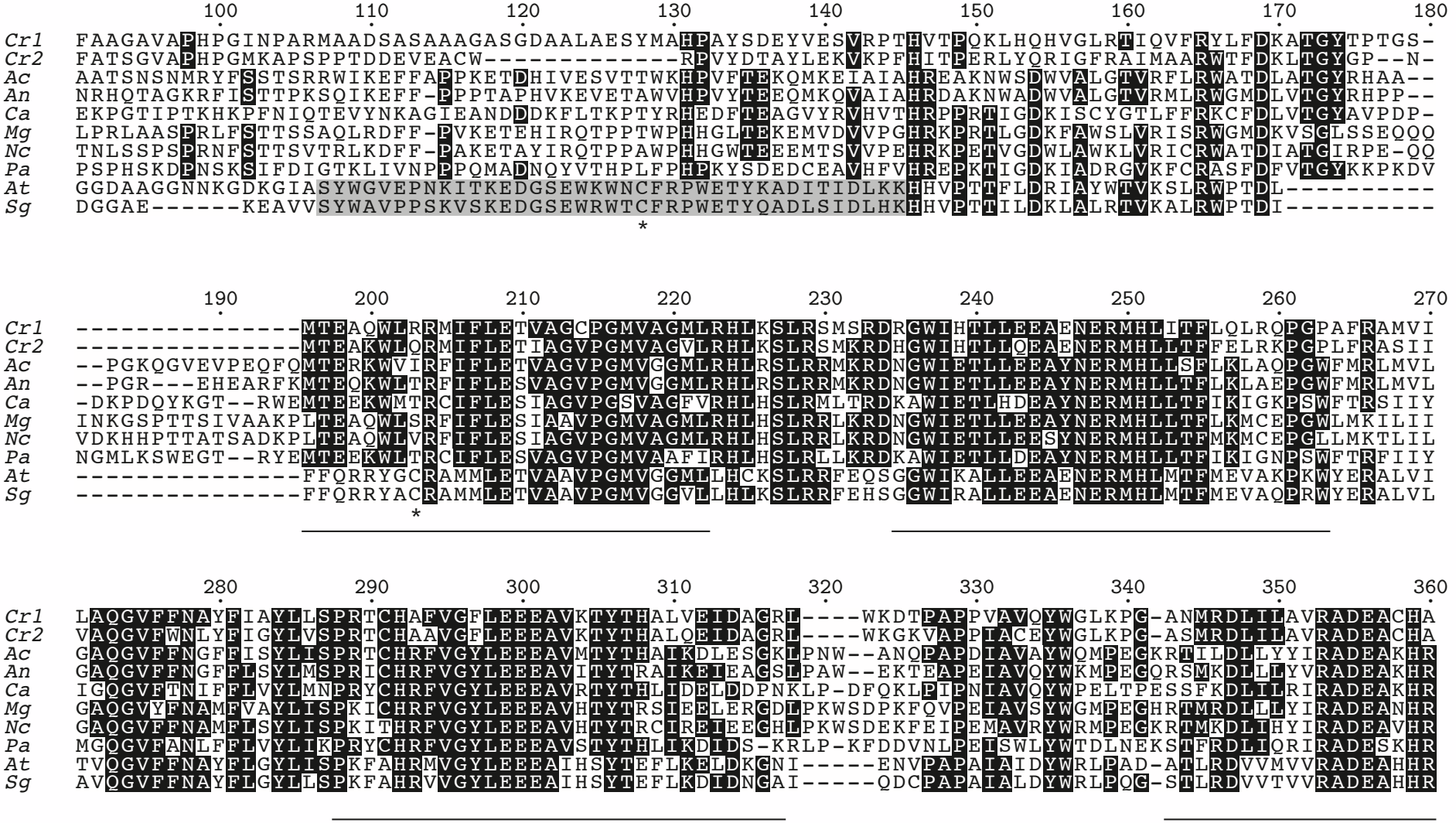

```
1  AACAAAGT
2  AACATAAA
3  AAAATAAT
4  CAAAATAA
```




```
1  AACAAAGT
2  AACATAAA
3  AA-AATAA
4  CAAAATAA
```

Phylogénie moléculaire

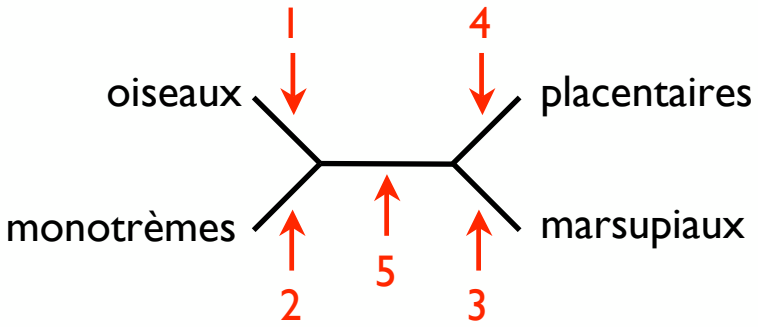
exemple : oxydase alternative

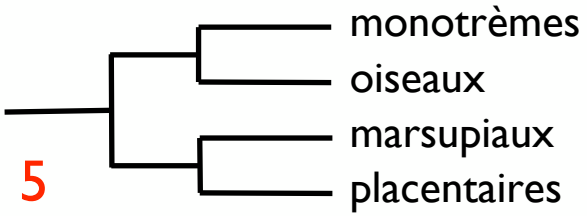


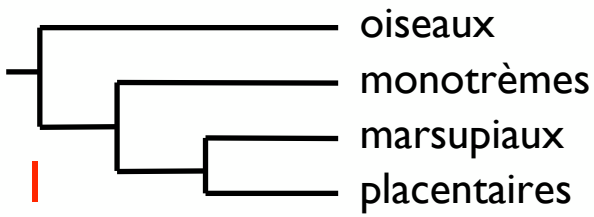
Baurain (2003) *Thèse de Doctorat*

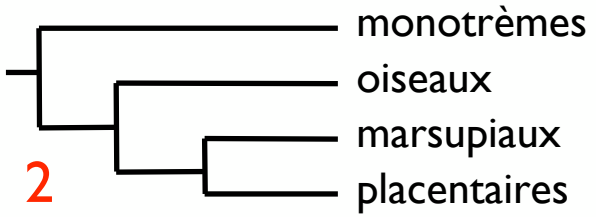
Phylogénie moléculaire

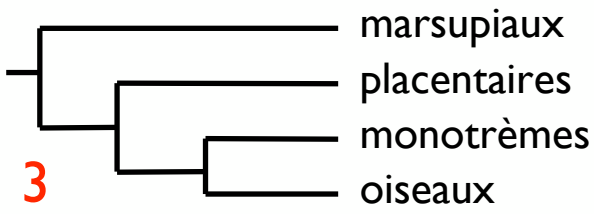
arbre non-enraciné, racine et groupe extérieur

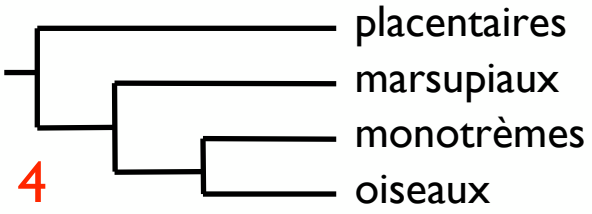






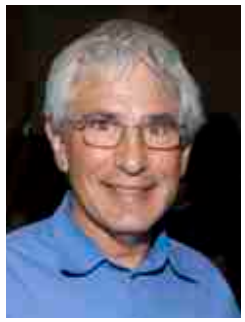






adapté de H. Philippe

Archezoa *vs.* 'Crown'



Mitch Sogin

T. Cavalier-Smith

SSU rRNA & facteurs d'élongation

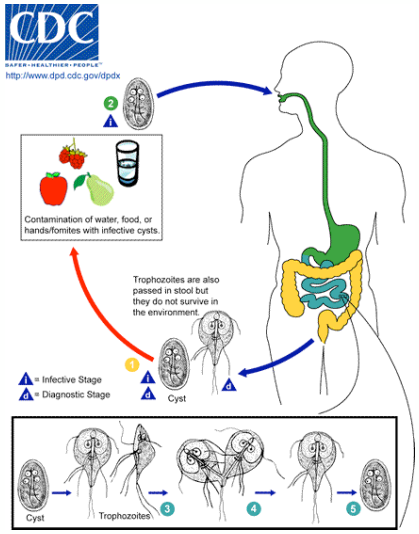
Sogin (1991) *Curr Opin Genet Dev* 1:457-463; Cavalier-Smith (1993) *Microbiol Rev* 57:953-994; Embley & Martin (2006) *Nature* 440:623-630

Que sont les Archezoa ?

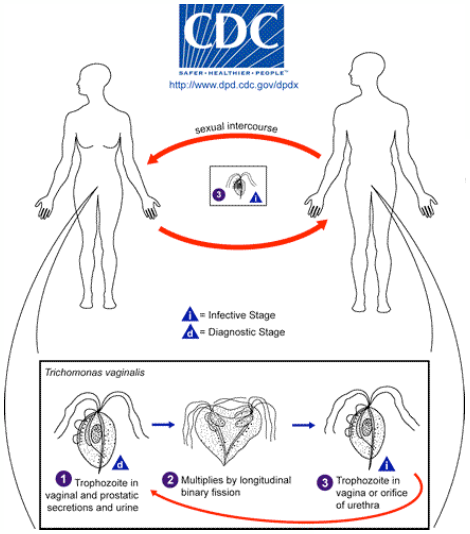
parasites
amitochondriaux

(ultrastructure simple,
pas de peroxysomes,
système endomembranaire
peu développé)

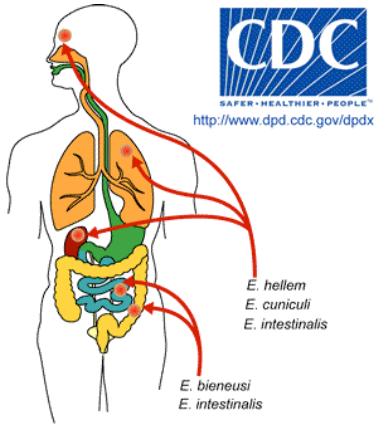
Giardia



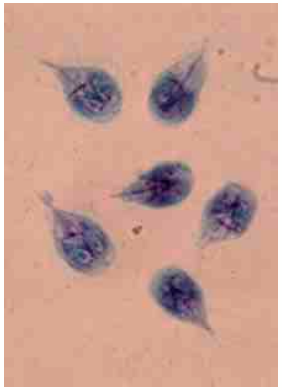
Trichomonas



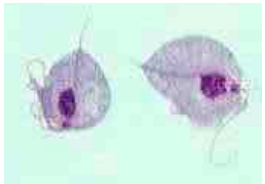
Encephalitozoon



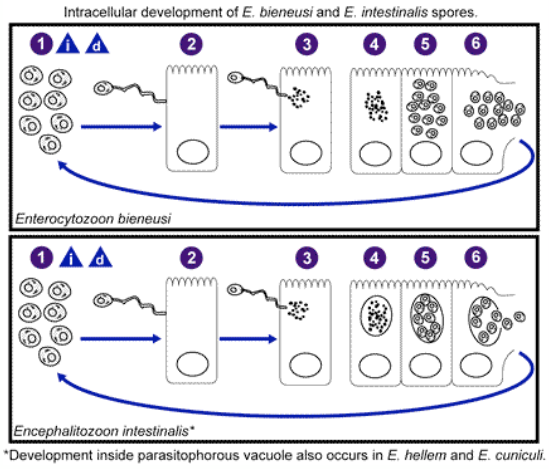
Giardia



Trichomonas



Encephalitozoon



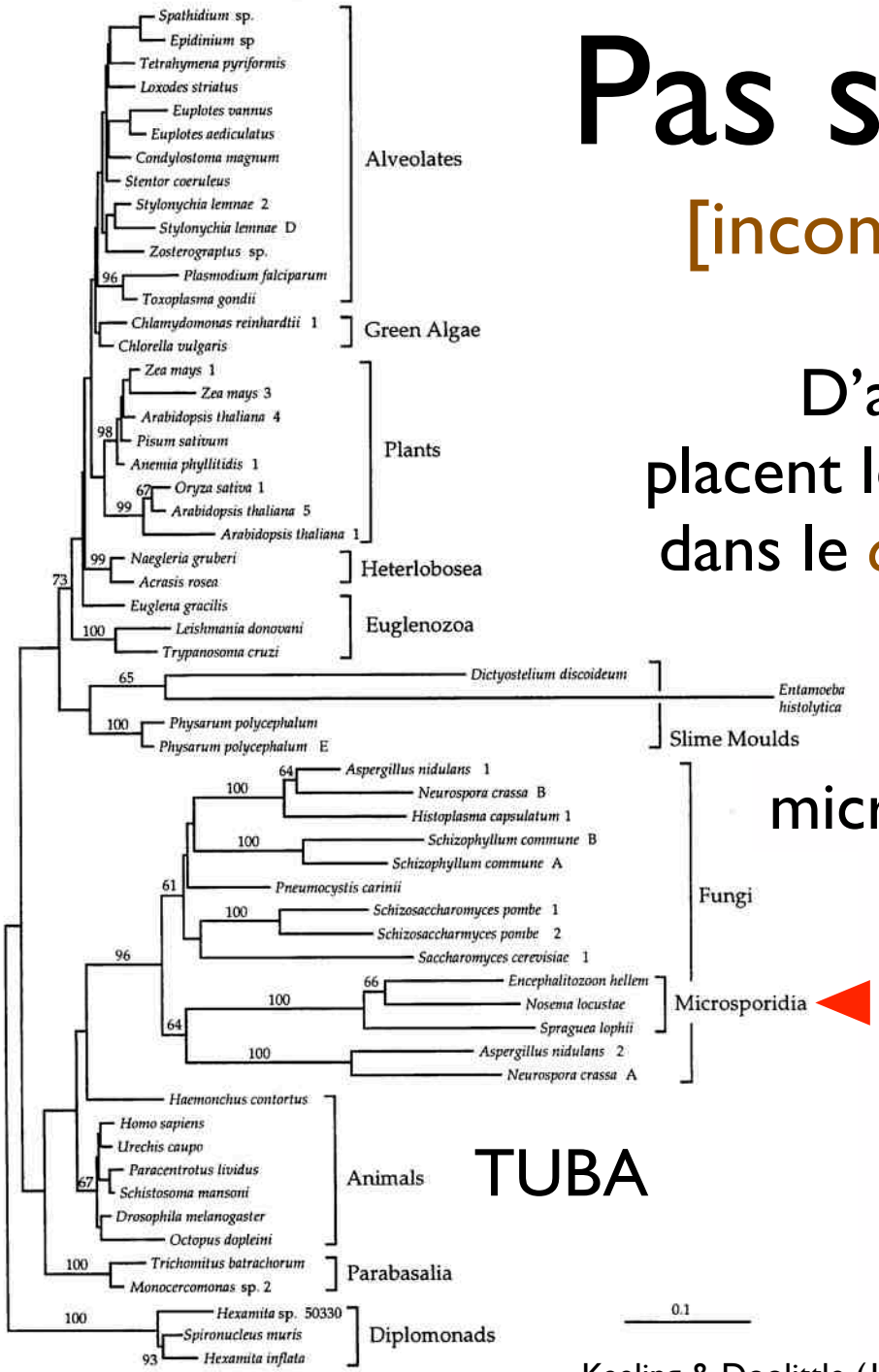
Pas si Archezoa (1)

[incongruence phylogénétique]

D'autres gènes placent les Archezoa dans le *crown* group.

Patrick Keeling

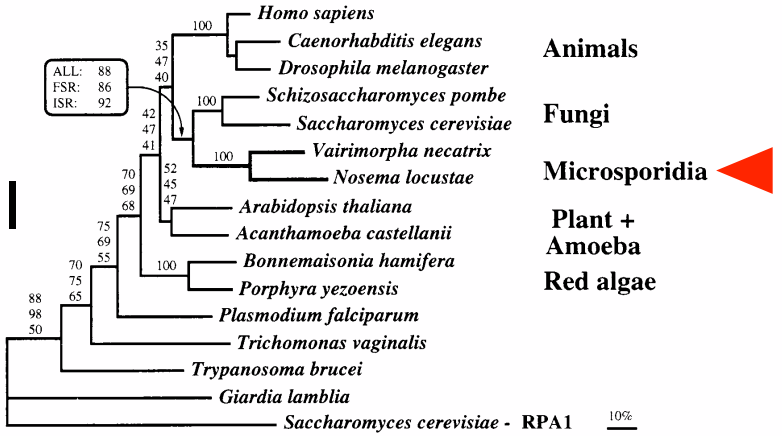
Martin Embley



TUBA

microsporidies = champignons dérivés

RBP1



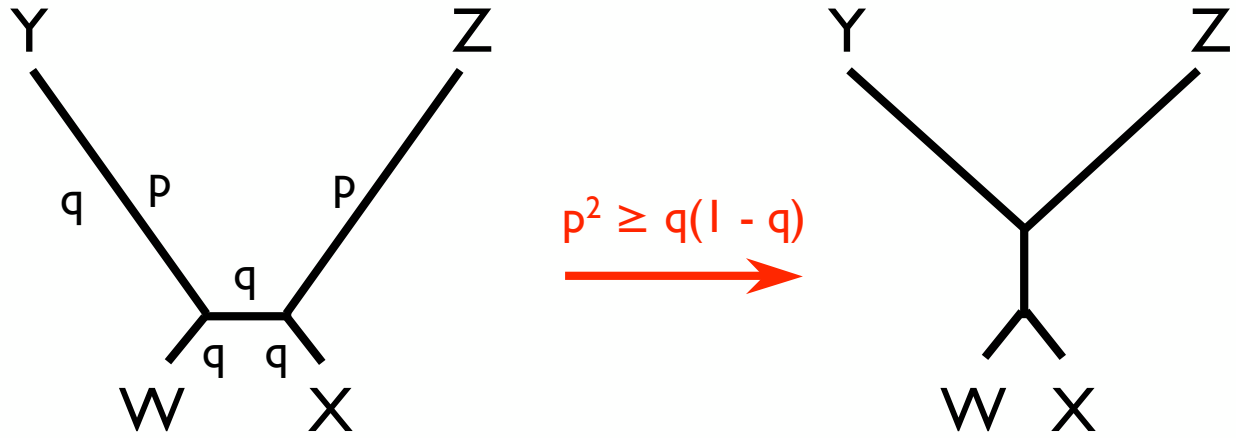
Keeling & Doolittle (1996) *Mol Biol Evol* 13:1297-1305; Hirt et al. (1999) *PNAS USA* 96:580-585

Joe Felsenstein

Hervé Philippe

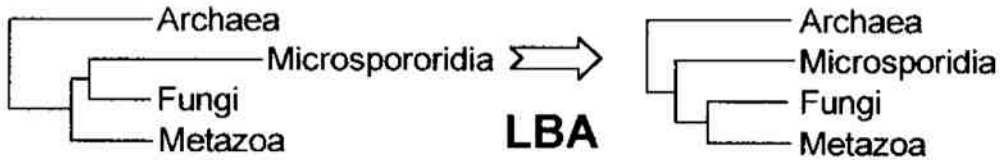
Pas si Archezoa (2)

[artéfacts de reconstruction]



si Y est un outgroup phylogénétiquement éloigné

L'attraction des longues branches (LBA) est responsable de la position artéfactuelle des Archezoa dans l'arbre des eucaryotes.



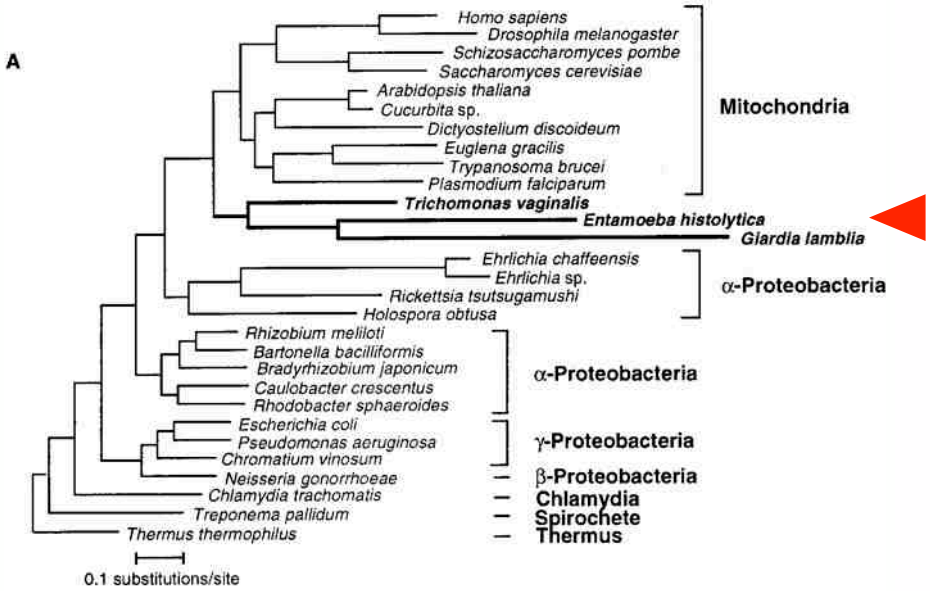
Felsenstein (1978) *Syst Zool* 27:401-410; Philippe et al. (2000) *Curr Opin Genet Dev* 10:596-601

Andrew Roger

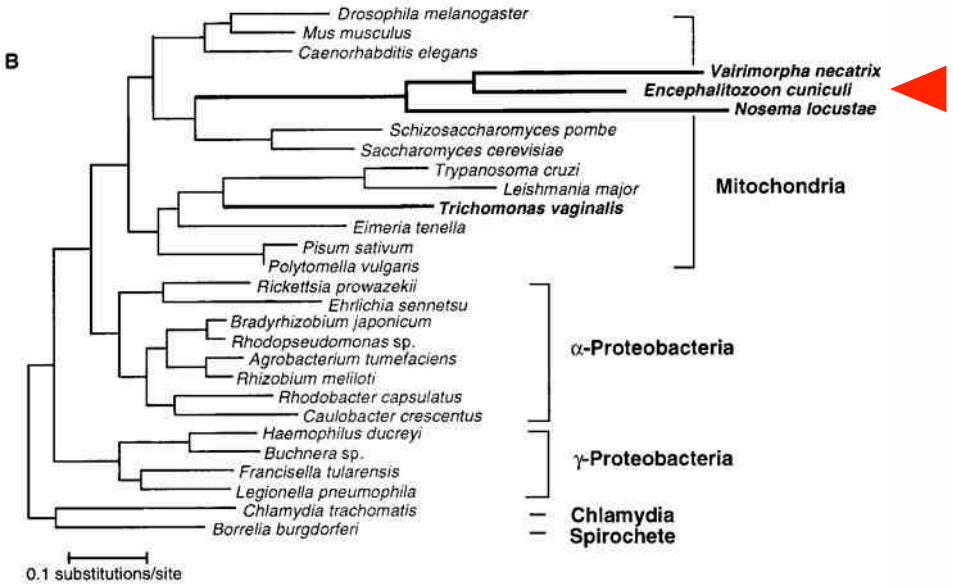
Pas si Archezoa (3)

[transfert de gènes]

CPN60



HSP70



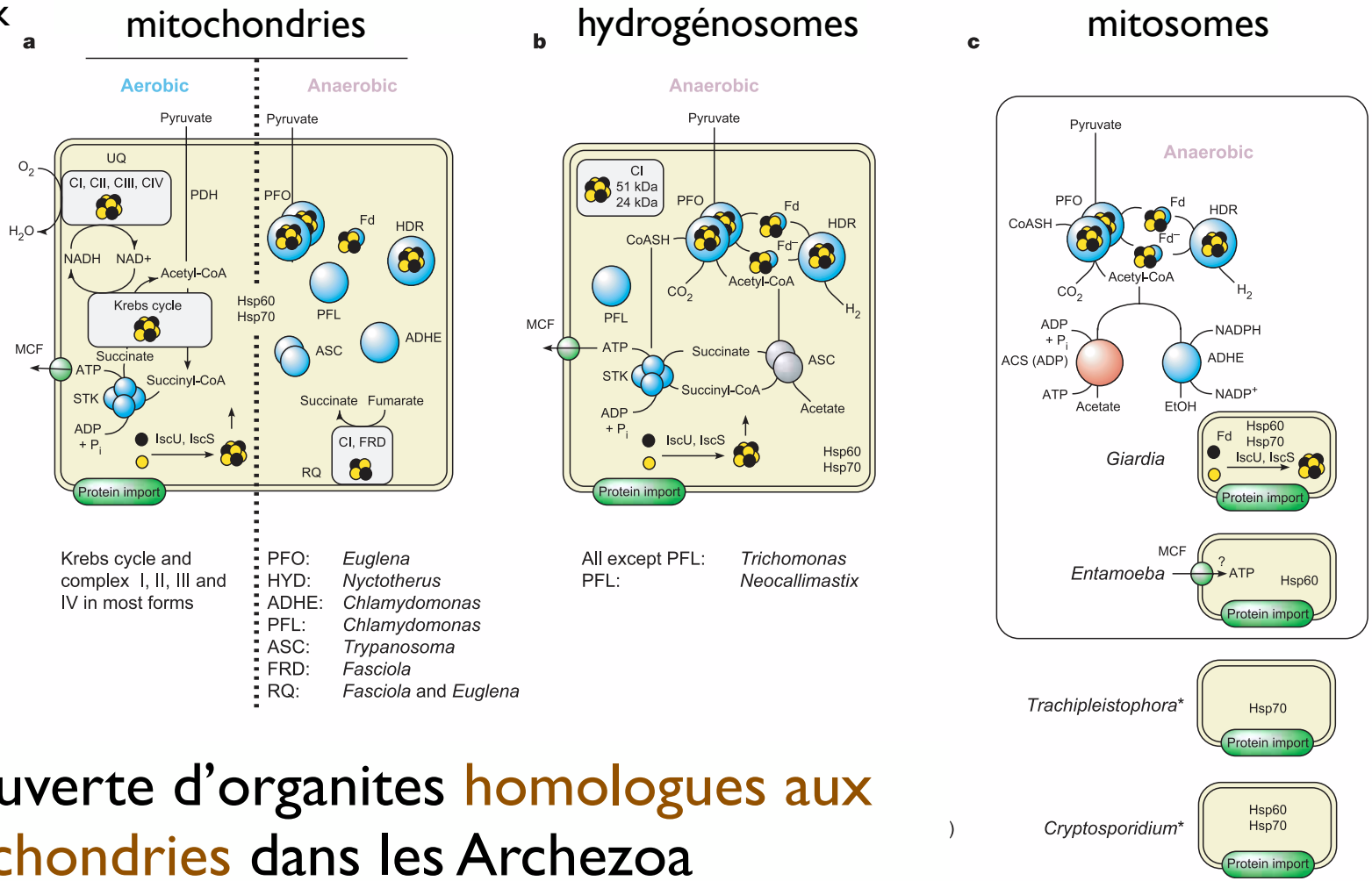
présence de gènes nucléaires d'origine mitochondriale dans les Archezoa

Roger (1999) *Am Nat* 154:146-163

C. Graham Clark

Pas si Archezoa (4)

[prédictions phylogénétiques]



découverte d'organites homologues aux mitochondries dans les Archezoa

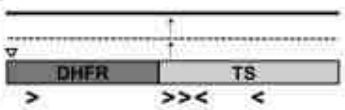
Tovar et al. (1999) *Mol Microbiol* 32:1013-1021; Embley & Martin (2006) *Nature* 440:623-630

Plan

1. Incongruence phylogénétique
   - *Phylogénomique*
2. Artéfacts de reconstruction
   - *Causes et remèdes*
3. Transfert de gènes (EGT/HGT)
   - *Orthologie, paralogie et xénologie*
4. Prédictions phylogénétiques
   - *Horloges moléculaires*

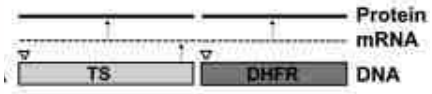
Les 6 Supergroupes

fusion des gènes DHFR et TS



‘unikontes’
vs. ‘bikontes’

Plantae
Glaucophytes
Red algae
Green algae
+ Plants
Chromalveolata
Alveolates
Stramenopiles
Haptophytes
Cryptophytes
Rhizaria
Radiolaria
Cercozoa
(inc. Foraminifera?)
Amoebozoa
(Lobose amoebae)
Mycetozoan slime moulds
Pelobionts,
Entamoebae
Euglenozoa,
Heterolobosea,
Jakobids
Oxymonads,
Trimastix
Malawimonas
Diplomonads,
Retortamonads,
Carpediemonas
Parabasalids
Excavata
Opisthokonta
Animals
Choano-
flagellates
Ichthyosporea
Nucleariid amoebae
Fungi (inc.
microsporidia)
Apusomonads
Collodictyonids
Centrohelid Heliozoa
racine ?
Eukaryotes
Prokaryotes
Eubacteria
Archaebacteria



~2002

Stechmann & Cavalier-Smith (2002) *Science* 297:89-91; Simpson & Roger (2004) *Curr Biol* 14:R693-696

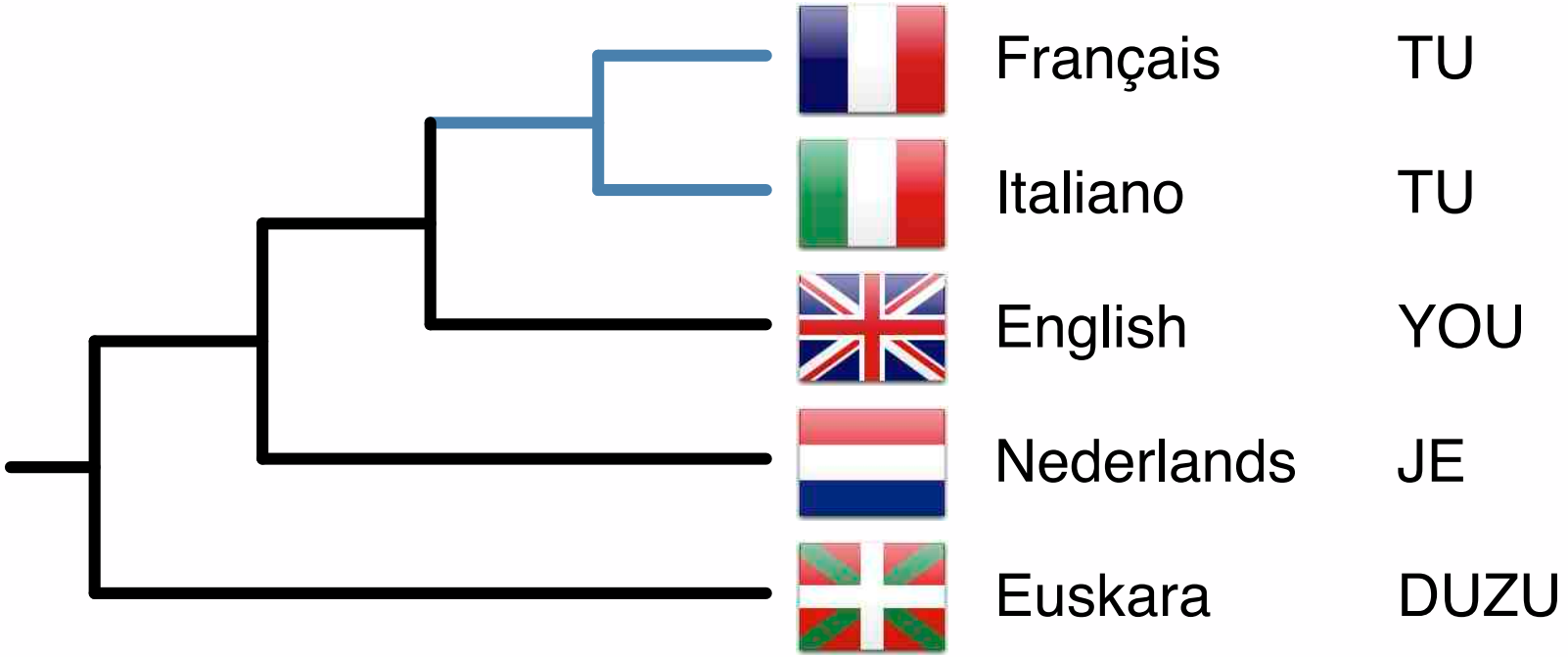
Incongruence phylogénétique
Phylogénomique

Incongruence

un exemple tiré de la linguistique

Potential relationships among European languages

based on the naive analysis of 1 word

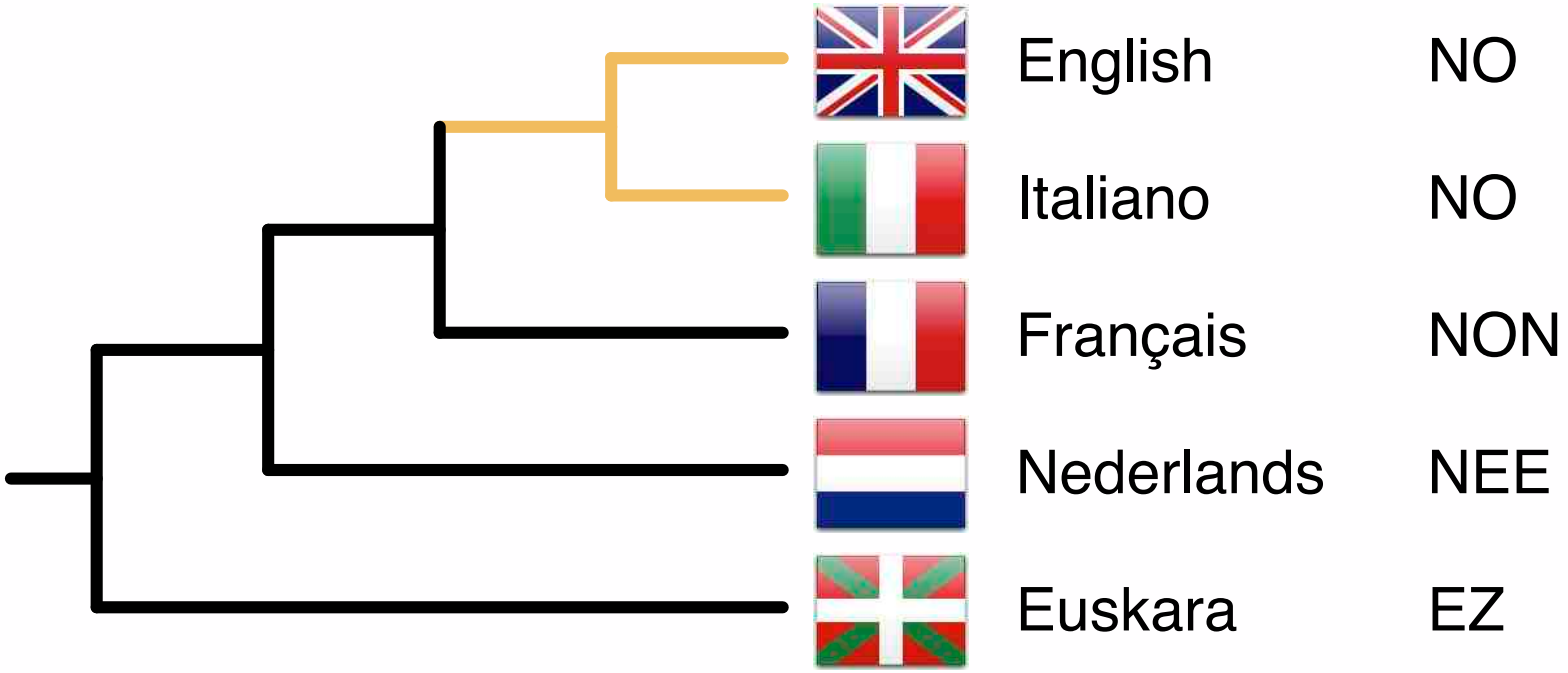


Incongruence

un exemple tiré de la linguistique

Potential relationships among European languages

based on the naive analysis of 1 word



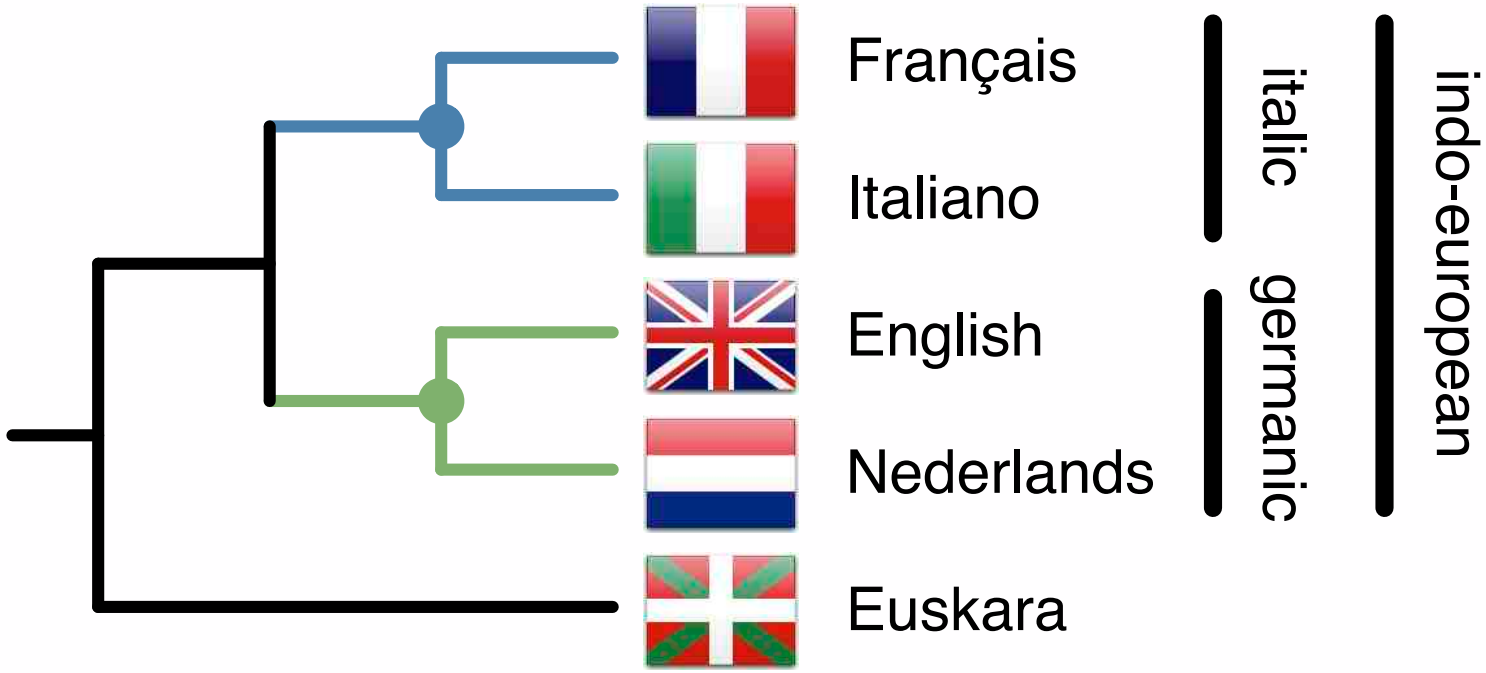
	français	italiano	english	nederlands	euskara
1	un	uno	one	een	bat
2	deux	due	two	twee	bi
3	trois	tre	three	drie	hiru
4	je	io	I	ik	I
5	tu	tu	you	je	duzu
6	qui ?	chi?	who?	wie?	nor?
7	oui	si	yes	ja	bai
8	non	no	no	nee	ez
9	mère	madre	mother	moeder	ama
10	père	padre	father	vader	aita
11	dent	dente	tooth	tand	hortz
12	coeur	cuore	heart	hart	bihotza
13	pied	piede	foot	voet	oinez
14	souris	topolino	mouse	muis	saguaren

Incongruence

un exemple tiré de la linguistique

Known relationships among European languages

strongly supported by the naive analysis of 14 words



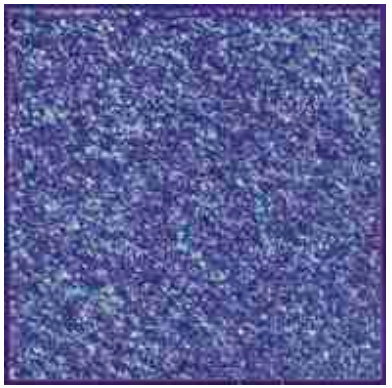
Phylogénomique

La révolution génomique a commencé en 1995.

Genomics

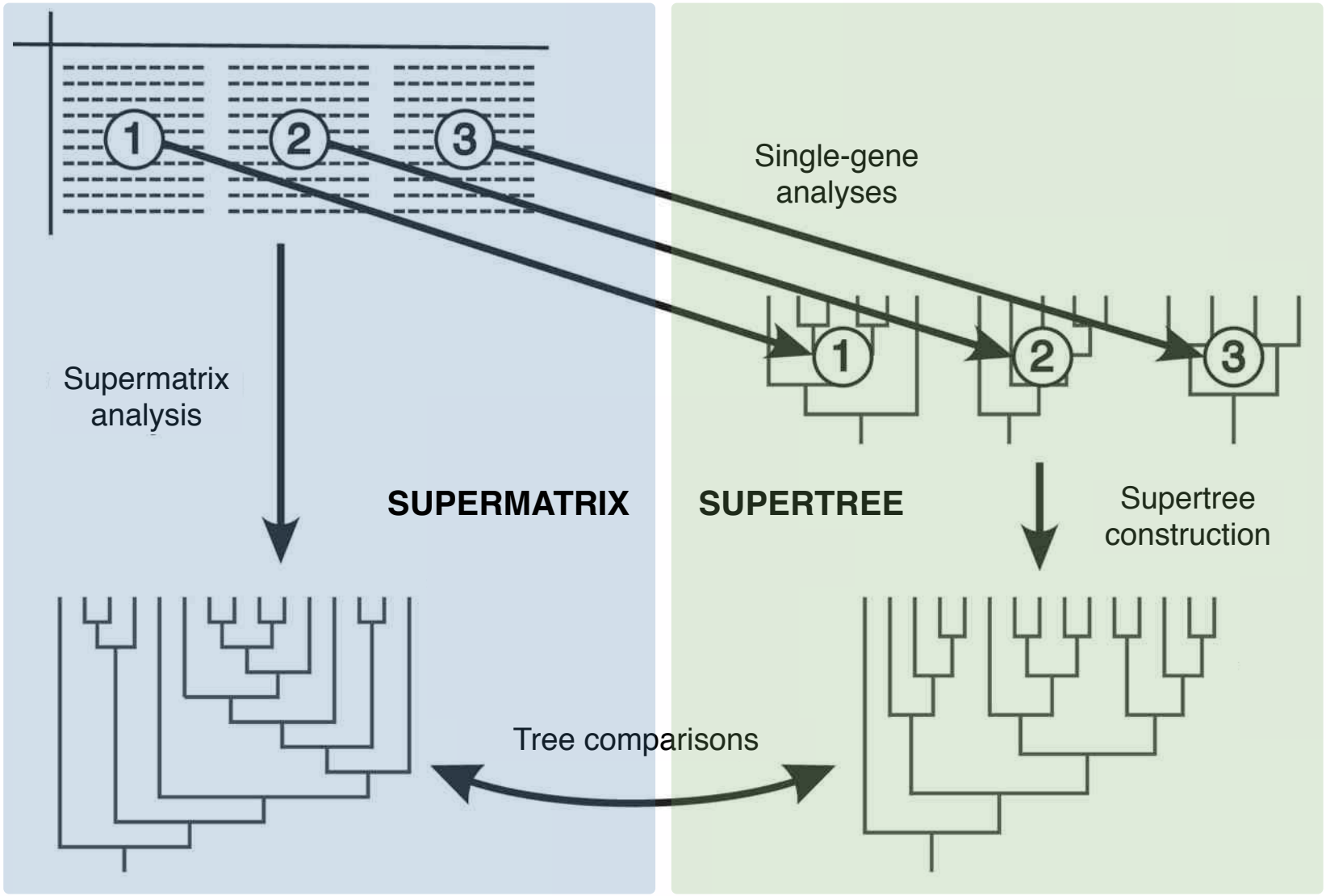
Automated Sanger sequencing machine (ABI 3730xl)

High-throughput experiments



High-density oligonucleotide microarray (Affymetrix array)

Phylogénomique



modifié de Bininda-Emonds (2010) *Palaeodiversity* 3 (Suppl.):99-106

Phylogénomique

concaténation d'alignements en supermatrices

	$\mathbb{S}^1$
S_1	A C G T C A A G
S_2	A C - T C C A G
S_3	A C - T C G A C

	$\mathbb{S}^2$
S_1	T G G - - T
S_3	A G C T C C
S_4	A G C T C G

	$\mathbb{S}^3$
S_1	C G G A C T A C G T
S_4	C C C T - - - - G G
S_5	C G T T C G A C G T

	$\mathbb{S}^1$	$\mathbb{S}^2$	$\mathbb{S}^3$
S_1	A C G T C A A G	T G G - - T	C G G A C T A C G T
S_2	A C - T C C A G	. . . . .	. . . . .
S_3	A C - T C G A C	A G C T C C	. . . . .
S_4	. . . . .	A G C T C G	C C C T - - - - G G
S_5	. . . . .	. . . . .	C G T T C G A C G T

Criscuolo A. (2006) *Thèse de Doctorat*

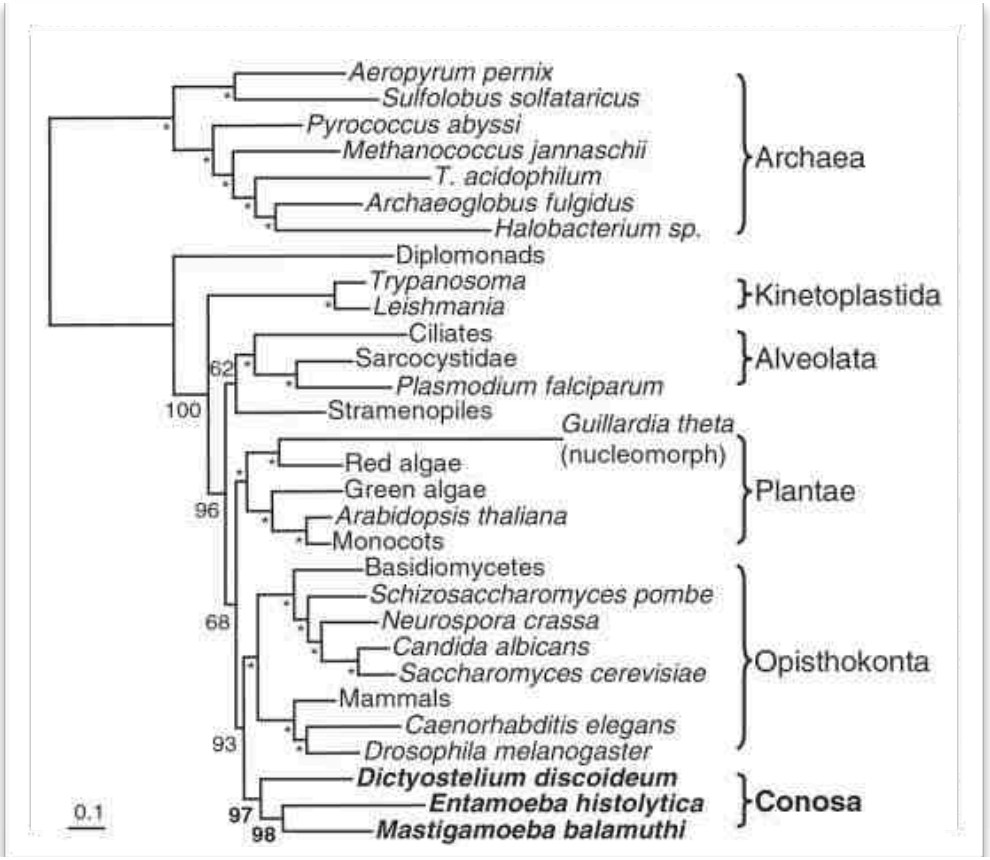
The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*

Eric Bapteste*, Henner Brinkmann†, Jennifer A. Lee‡, Dorothy V. Moore‡, Christoph W. Sensen§, Paul Gordon¶, Laure Duruflé*, Terry Gaasterland‡, Philippe Lopez*, Miklós Müller‡, and Hervé Philippe*‖

*Unité Mixte de Recherche 7622 Centre National de la Recherche Scientifique, Université Paris 6, 9 Quai Saint Bernard, Bât C, 75005 Paris, France; †Department of Biology, University of Konstanz, 78457 Konstanz, Germany; ‡The Rockefeller University, 1230 York Avenue, New York, NY 10021; §Department of Biochemistry and Molecular Biology, University of Calgary, 3330 Hospital Drive N.W., Calgary, AB, Canada, T2N 4N1; and ¶National Research Council, Institute for Marine Biosciences, 1411 Oxford Street, Halifax, NS, Canada B3H 3Z1

Communicated by William Trager, The Rockefeller University, New York, NY, December 11, 2001 (received for review October 10, 2001)

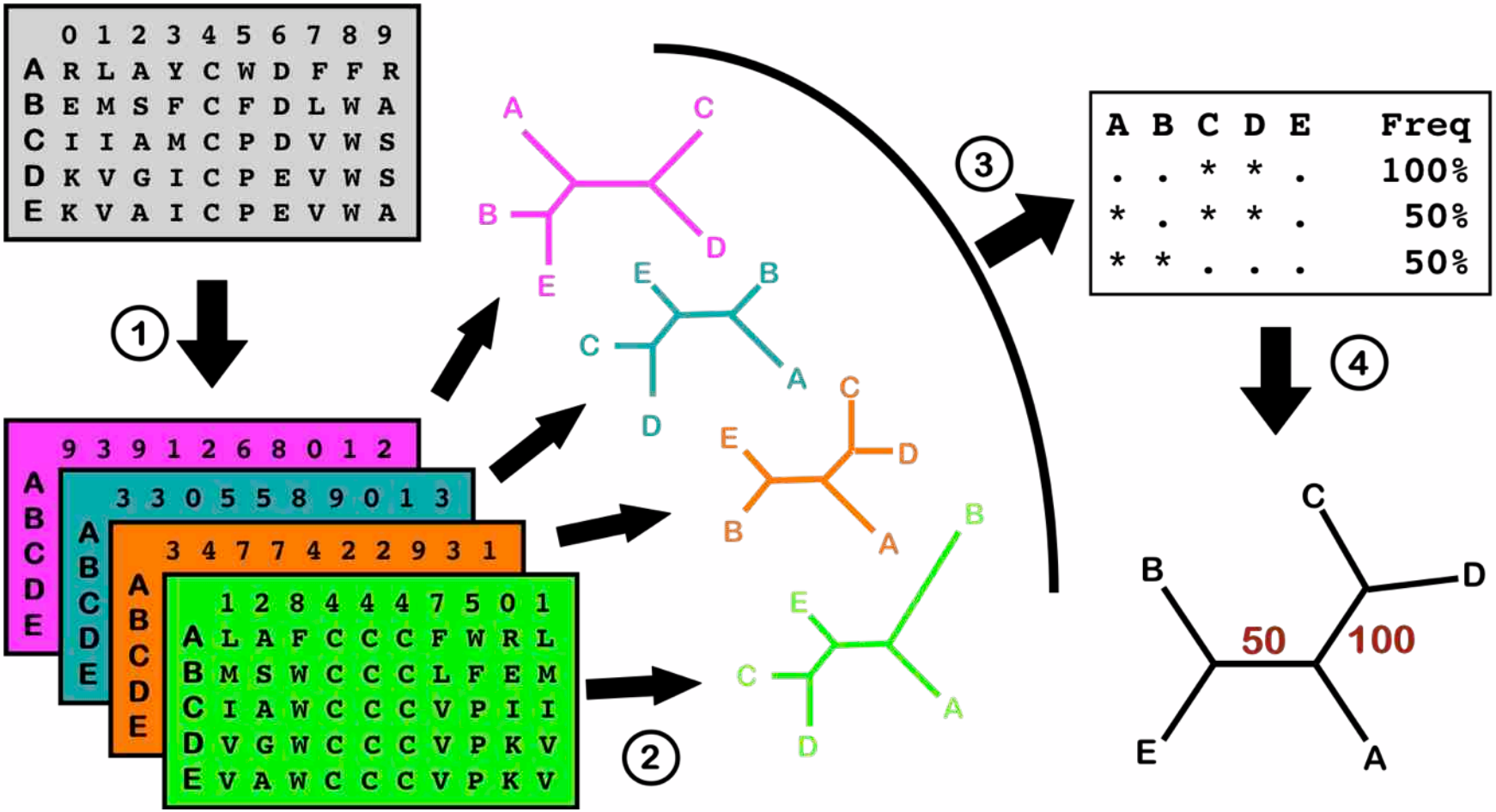
The phylogenetic relationships of amoebae are poorly resolved. To address this difficult question, we have sequenced 1,280 expressed sequence tags from *Mastigamoeba balamuthi* and assembled a large data set containing 123 genes for representatives of three phenotypically highly divergent major amoeboid lineages: Pelobionta, Entamoebidae, and Mycetozoa. Phylogenetic reconstruction was performed on ≈25,000 aa positions for 30 species by using maximum-likelihood approaches. All well-established eukaryotic groups were recovered with high statistical support, validating our approach. Interestingly, the three amoeboid lineages strongly clustered together in agreement with the Conosa hypothesis [as defined by T. Cavalier-Smith (1998) *Biol. Rev. Cambridge Philos. Soc.* 73, 203–266]. Two amitochondriate amoebae, the free-living *Mastigamoeba* and the human parasite *Entamoeba*, formed a significant sister group to the exclusion of the mycetozoan *Dictyostelium*. This result suggested that a part of the reductive process in the evolution of *Entamoeba* (e.g., loss of typical mitochondria) occurred in its free-living ancestors. Applying this inexpensive expressed sequence tag approach to many other lineages will surely improve our understanding of eukaryotic evolution.



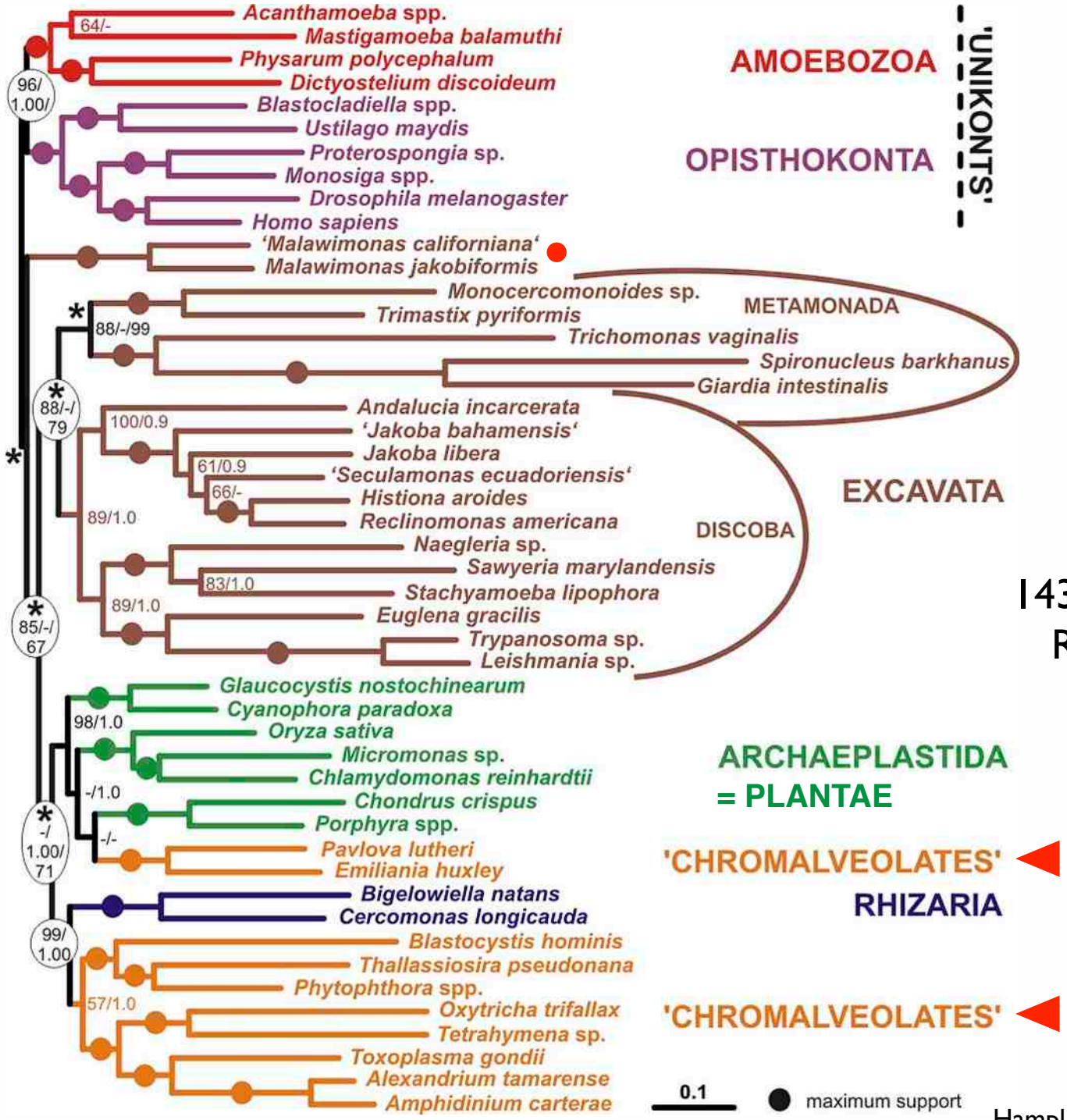
Unicellular amoebae are possibly the simplest eukaryotic

Phylogénomique

principe du bootstrap et robustesse des arbres



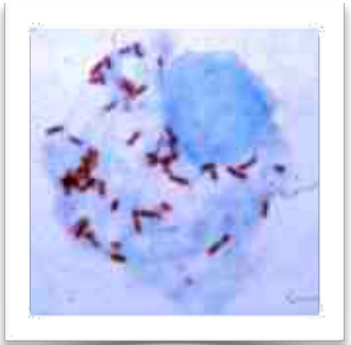
Felsenstein (1985) *Evolution* 39:783-791; Rodriguez-Ezpeleta (2007) *Thèse de Doctorat*



confirmation de l'existence de certains des supergroupes

143 gènes x 48 espèces
RAxML (WAG+Γ_4)

Hampl et al. (2009) *PNAS USA* 106:3859-3864



Rickettsia

42 gènes x 54 espèces
PhyloBayes (CAT+Γ_4)

confirmation de la position de la racine des eucaryotes

Derelle & Lang (2012) *Mol Biol Evol* 29:1277-1289

2

Artéfacts de reconstruction
Causes et remèdes

Artéfacts

une question de substitutions multiples

Artéfacts

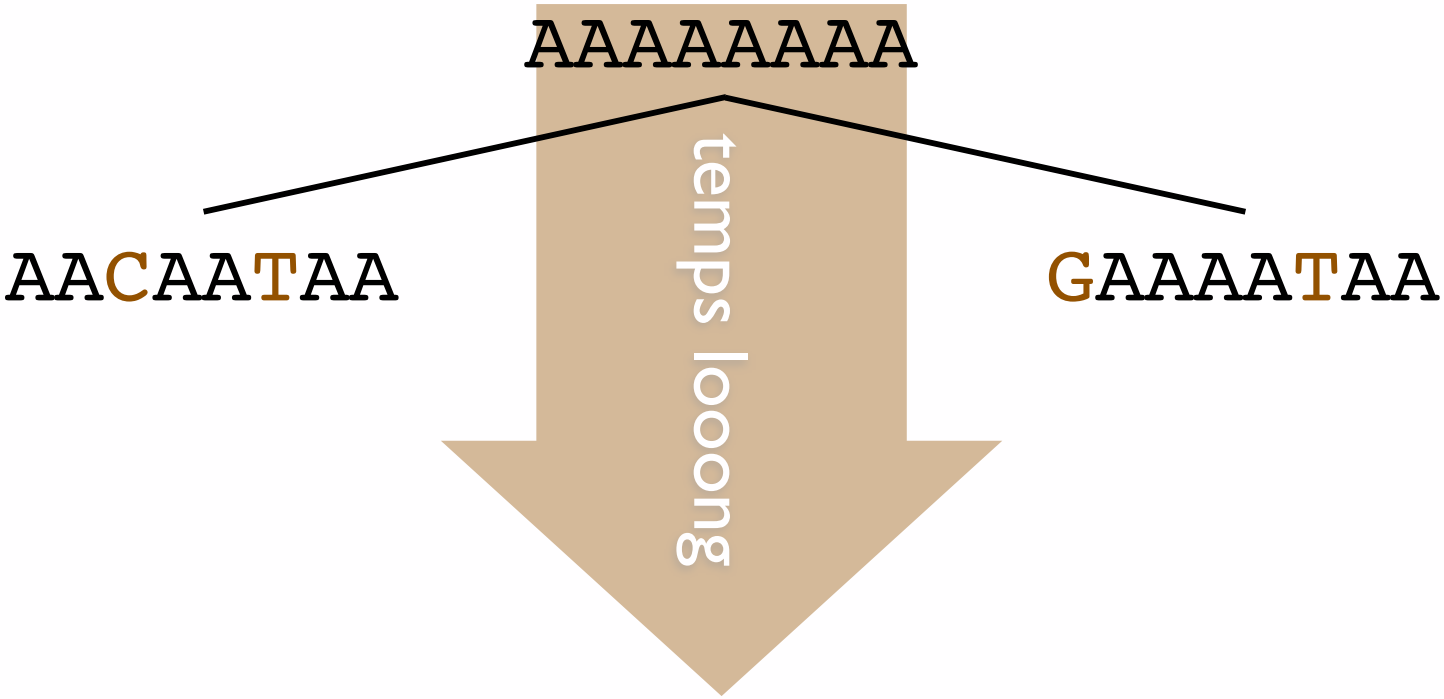
une question de substitutions multiples

Artéfacts

une question de substitutions multiples

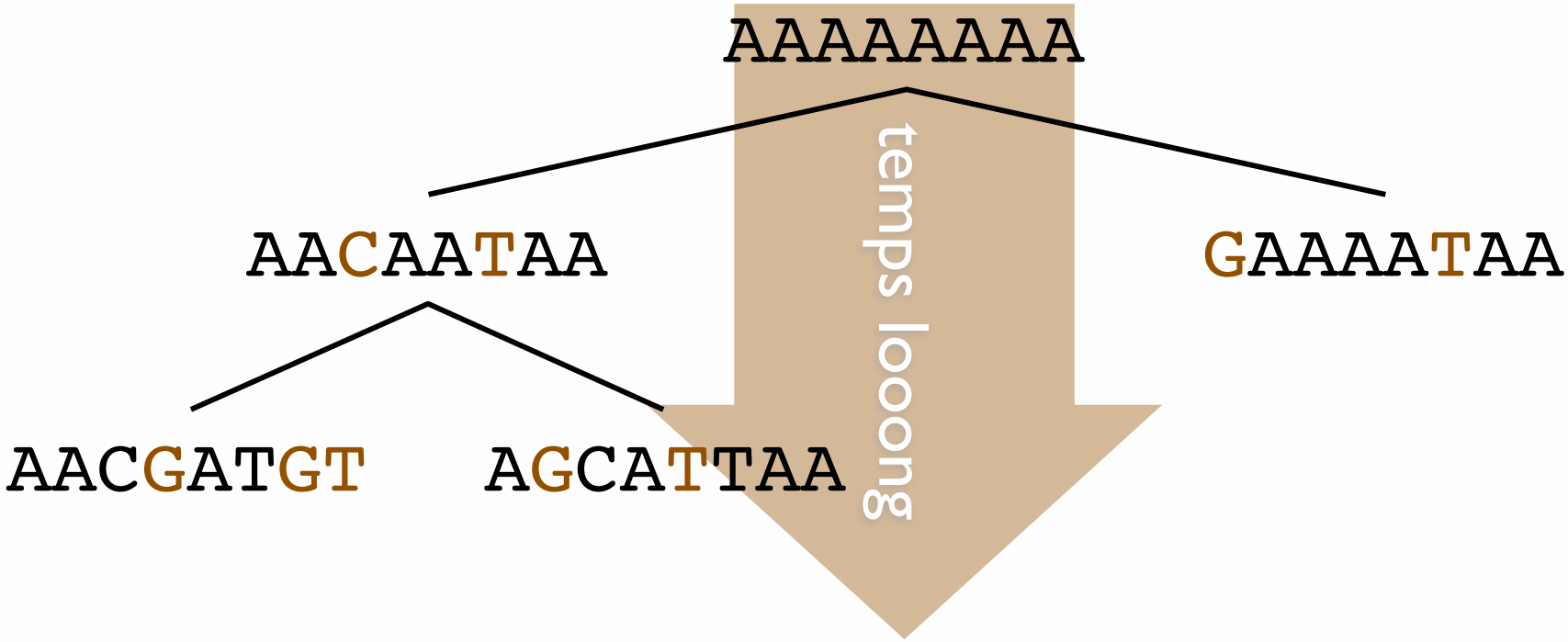
Artéfacts

une question de substitutions multiples



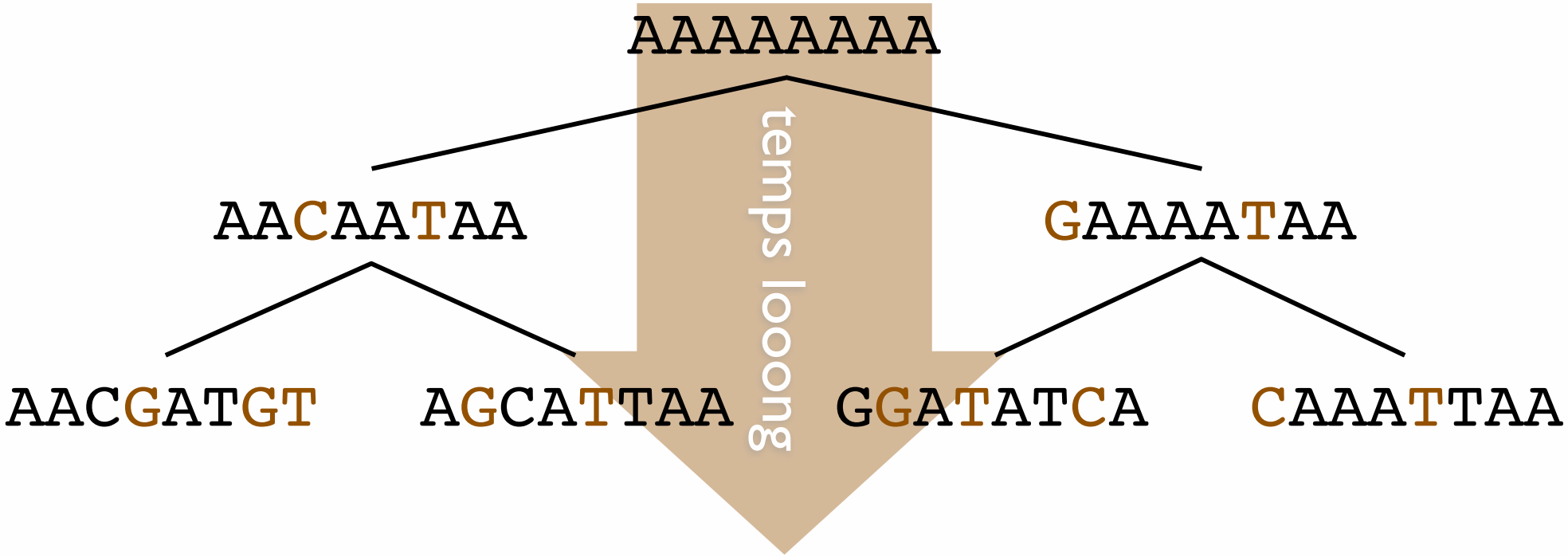
Artéfacts

une question de substitutions multiples



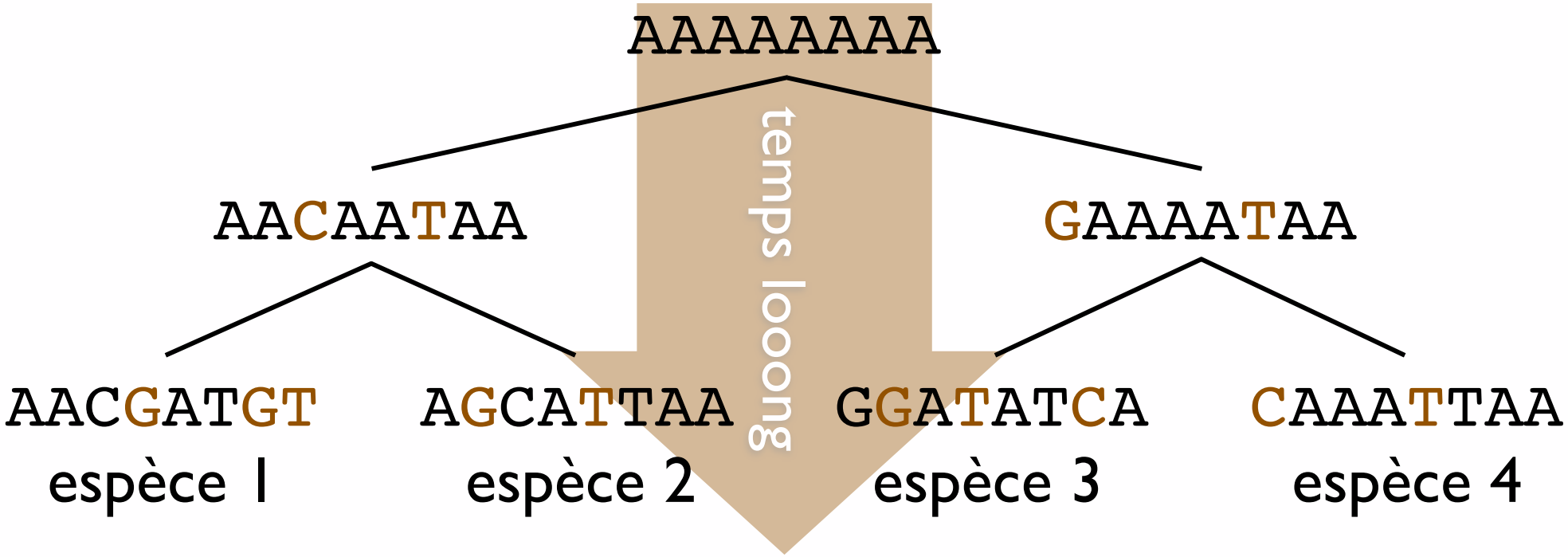
Artéfacts

une question de substitutions multiples



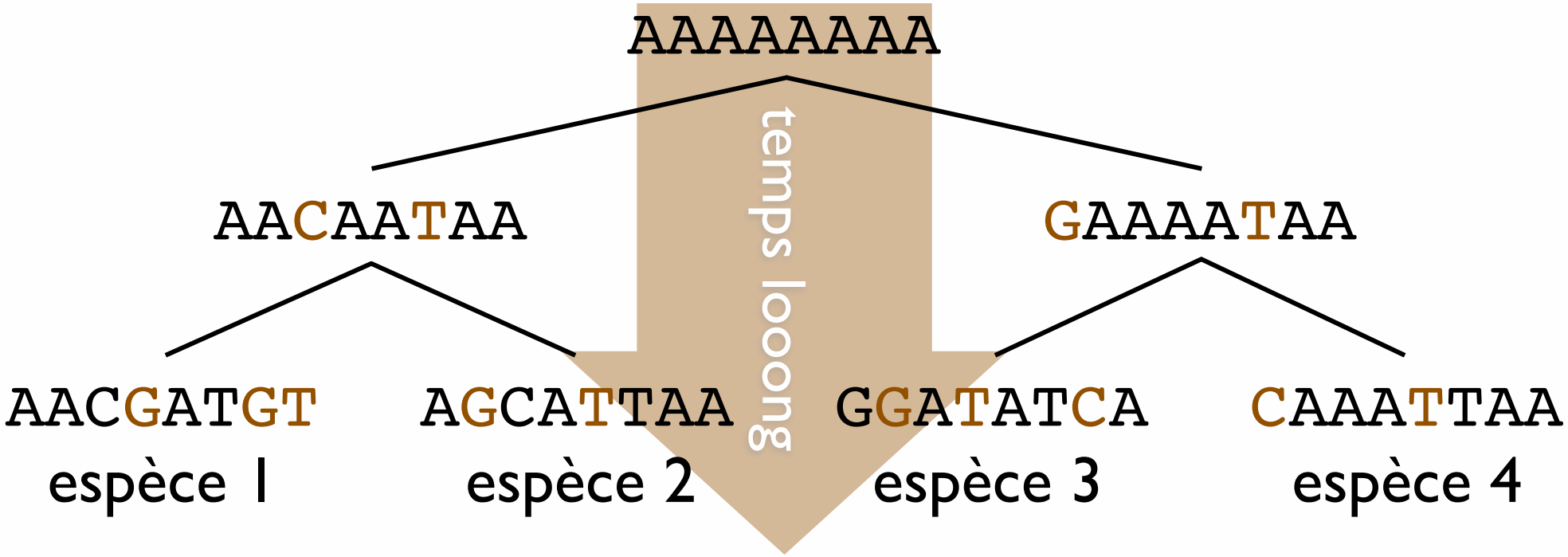
Artéfacts

une question de substitutions multiples



Artéfacts

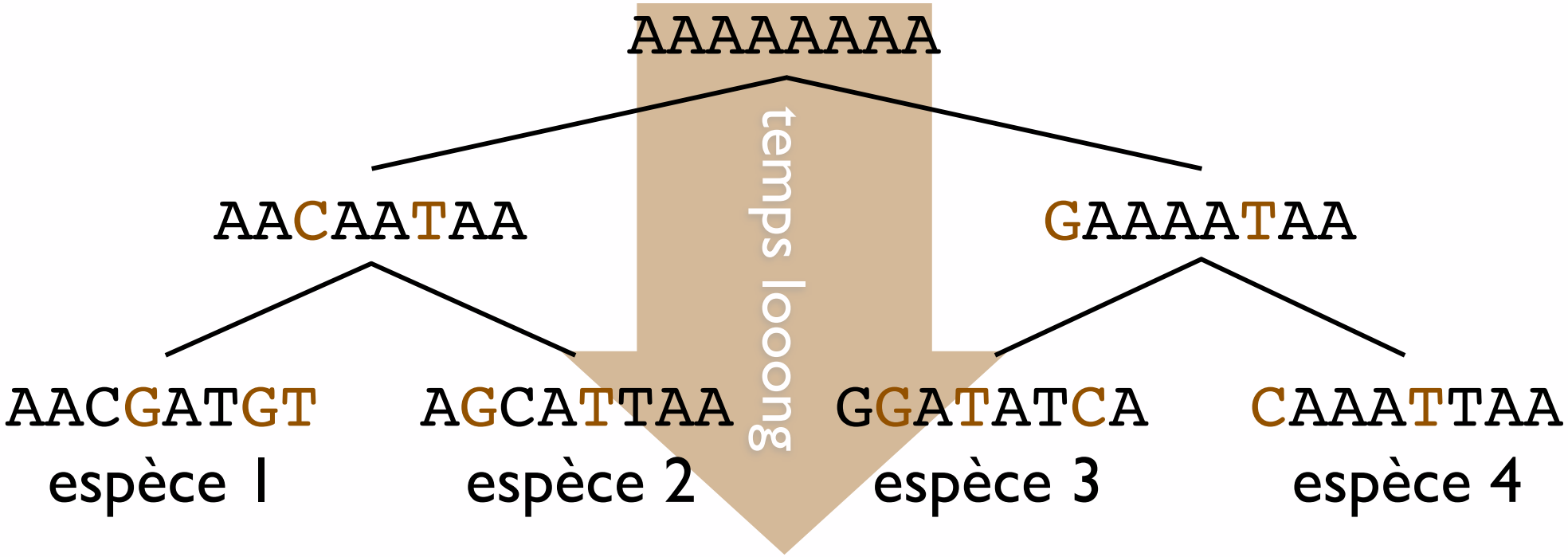
une question de substitutions multiples




```
1  AACGATGT
2  AGCATTAA
3  GGATATCA
4  CAAATTAA
```

Artéfacts

une question de substitutions multiples

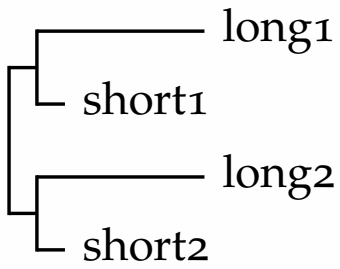


1 AACGATGT
2 AGCATTAA
3 GGATATCA
4 CAAATTAA

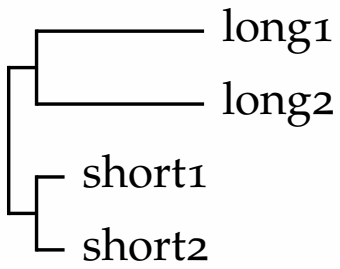
Lorsque l'homoplasie est trop importante, les sites informatifs deviennent trompeurs.
C'est le signal non-phylogénétique.

Artéfacts

attraction des longues branches / inconsistance

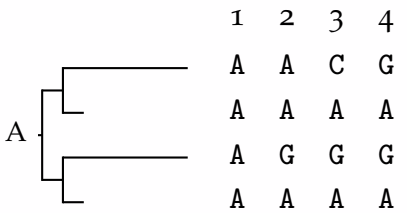


True



Recovered

1 both the same as short1 and short2
2 one the same and one different
3 both different and different from each other
4 both different but the same as each other



1	AAAA	parsimony uninformative
2	AAGA	parsimony uninformative
3	CAGA	parsimony uninformative
4	GAGA	parsimony misinformative

Puisque tous les sites sont soit non-informatifs, soit trompeurs, la parcimonie tend à regrouper les longues branches ensemble, et ce d'autant plus solidement qu'elle dispose de beaucoup de données.

adapté de P. G. Foster

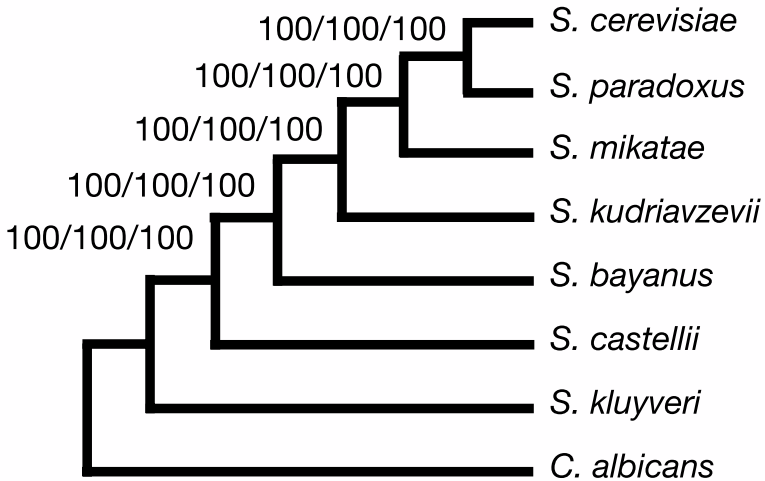
Evolution

Gee (2003) *Nature* 425:782

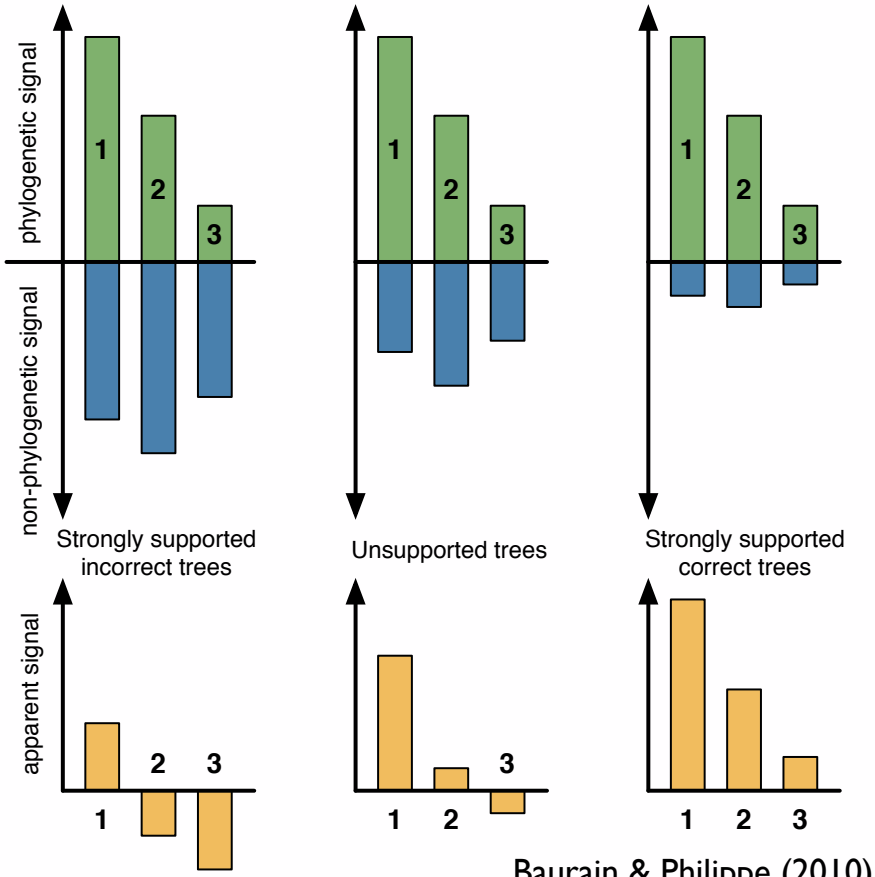
Ending incongruence

Henry Gee

Recovering the true evolutionary history of any group of organisms has seemed impossible. The availability of large amounts of genomic data promises an era in which the uncertainties are better constrained.



Rokas et al. (2003) *Nature* 425:798-804



Baurain & Philippe (2010)
In *Evolutionary Genomics and Systems Biology*

Opinion

TRENDS in Genetics Vol.22 No.4 April 2006

Full text provided by www.sciencedirect
SCIENCE DIRECT

Jeffroy et al. (2006) *Trends Genet* 22:225-231

Phylogenomics: the beginning of incongruence?

Olivier Jeffroy, Henner Brinkmann, Frédéric Delsuc and Hervé Philippe

Canadian Institute for Advanced Research, Centre Robert-Cedergren, Département de Biochimie, Université de Montréal, Succursale Centre-Ville, Montréal, Québec, Canada, H3C3J7

Artéfacts

Comment réduire le signal non-phylogénétique ?

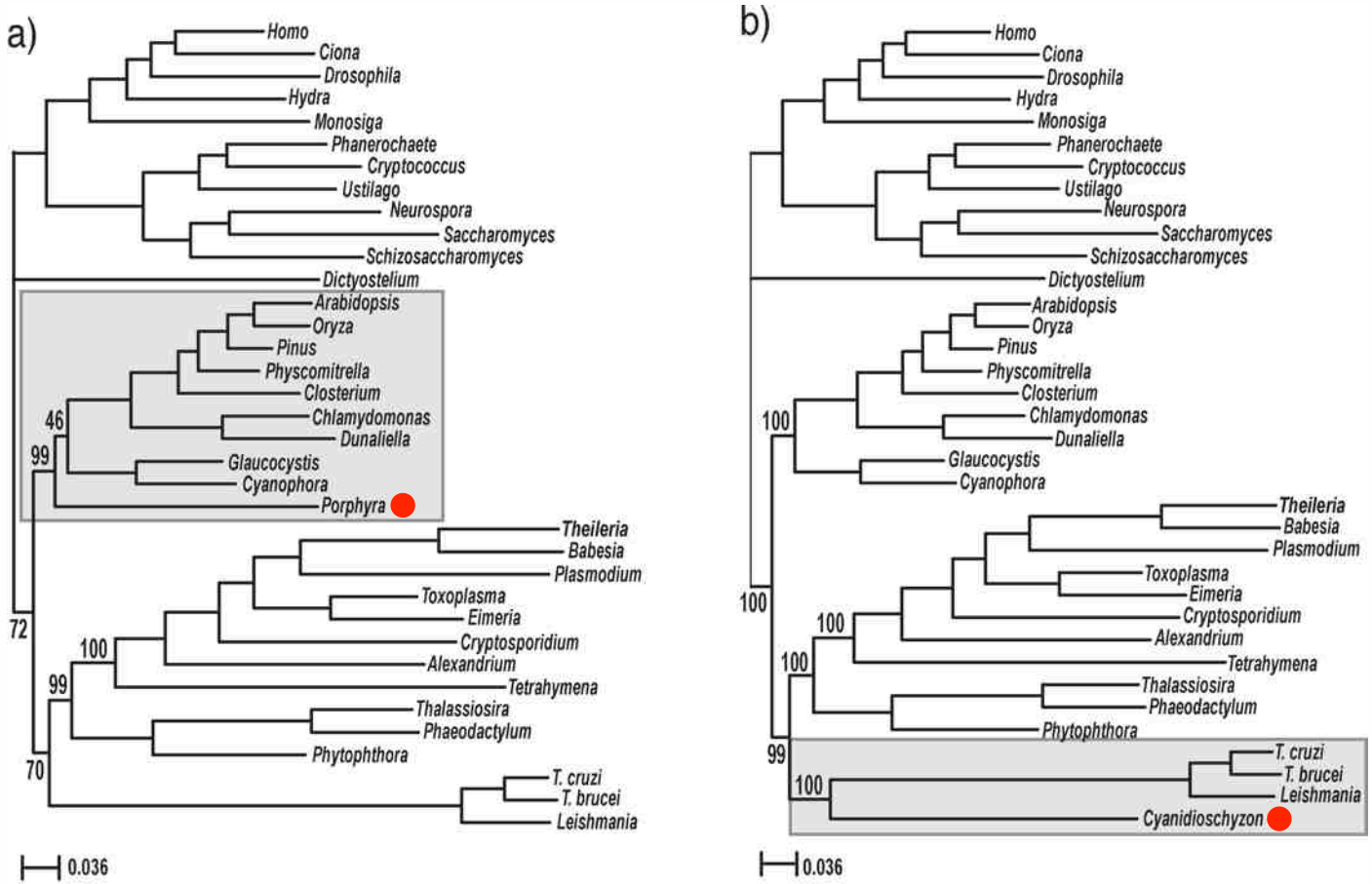
★ Combinaison de 3 approches :

1. Amélioration de l'échantillonnage taxonomique
   a. *Choix d'espèces à évolution lente*
   b. *Augmentation de la densité en espèces*
2. Elimination des sites à évolution rapide
3. Utilisation de modèles d'évolution sophistiqués

Philippe et al. (2011) *PLoS Biol* 9:e1000602

Remèdes

1a. choix d'espèces à évolution lente

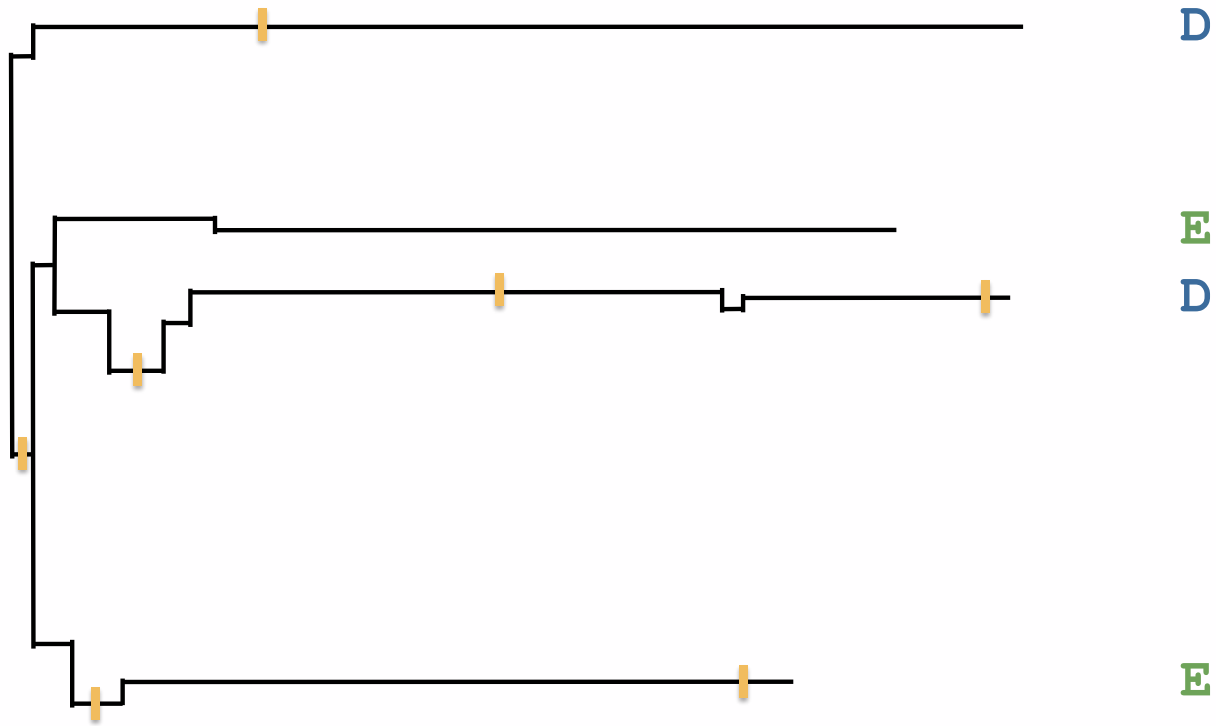


Le bootstrap évalue l'homogénéité du signal et non sa nature.

Rodriguez-Ezpeleta et al. (2007) *Syst Biol* 56:389-399

Remèdes

1b. augmentation de la densité en espèces

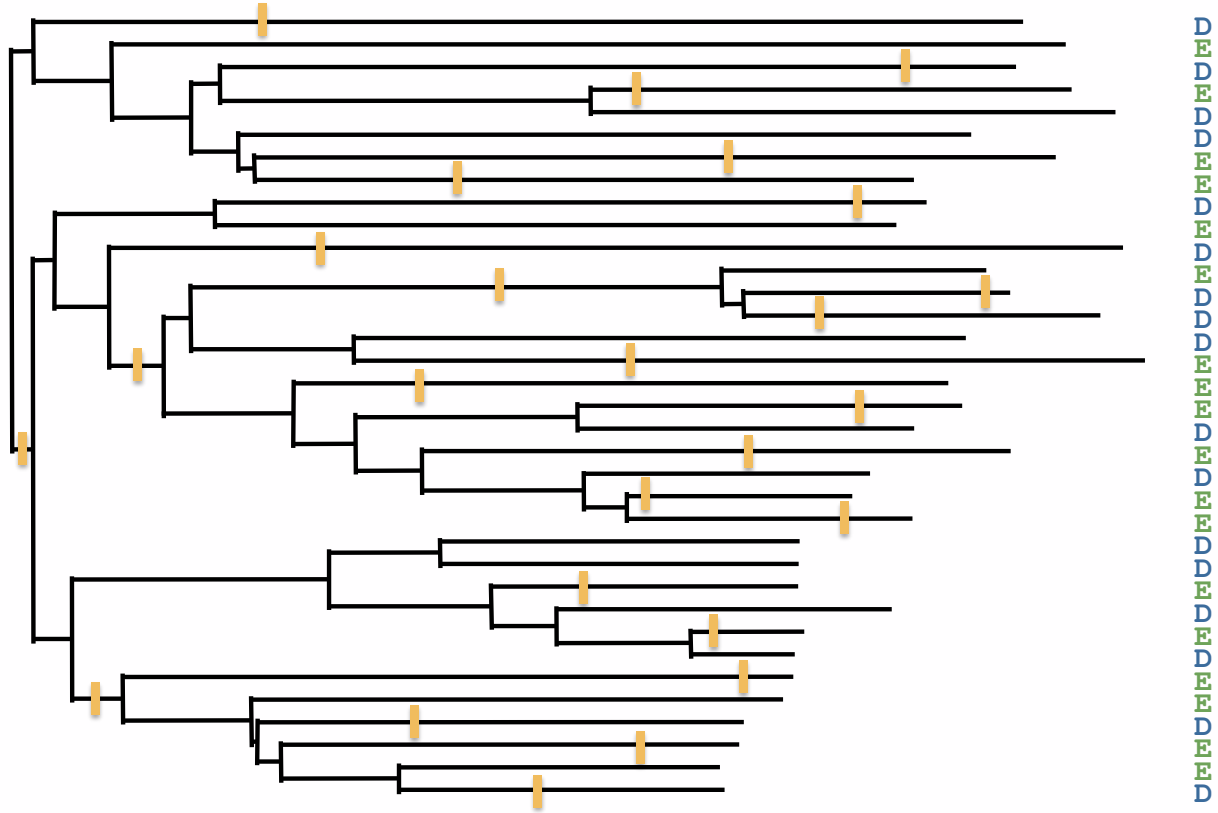


Un échantillonnage épars masque les substitutions multiples.

adapté de H. Philippe

Remèdes

1b. augmentation de la densité en espèces

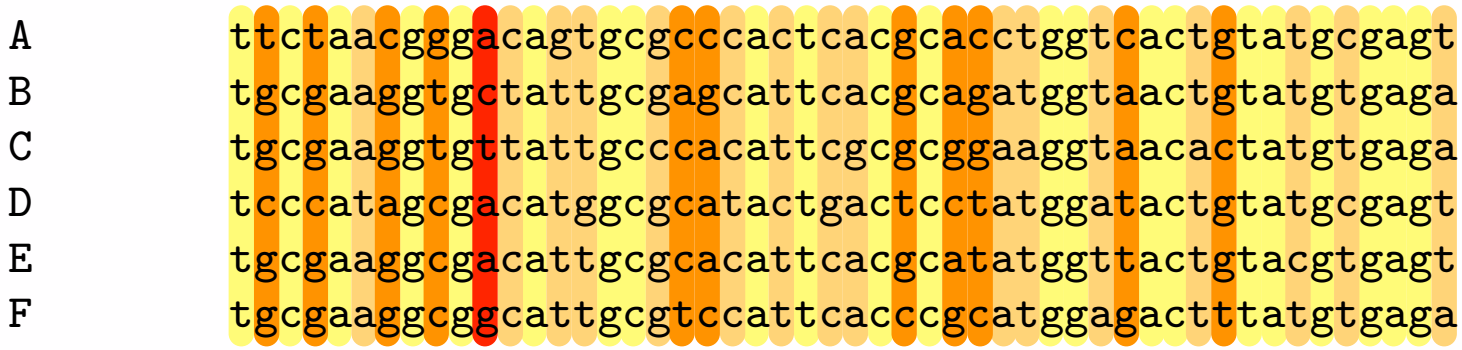


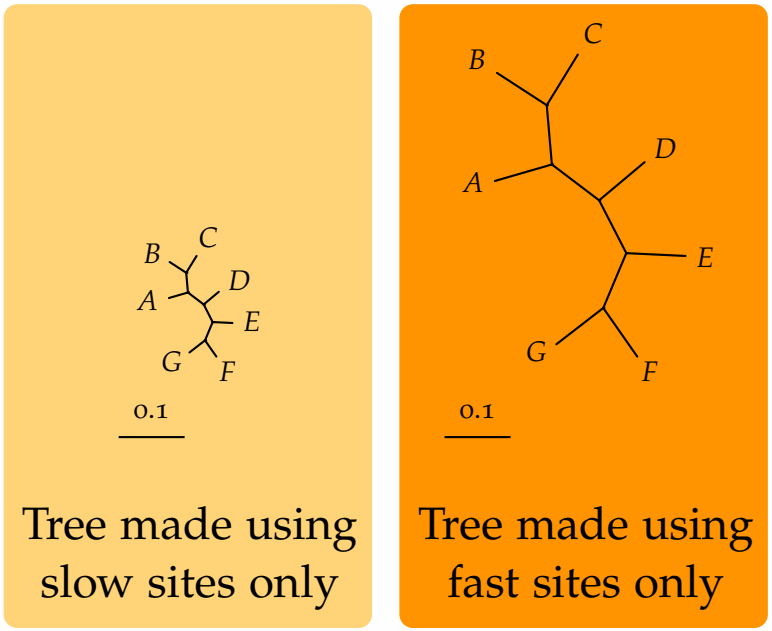
Un échantillonnage dense révèle les substitutions multiples.

adapté de H. Philippe

Remèdes

2. élimination des sites à évolution rapide



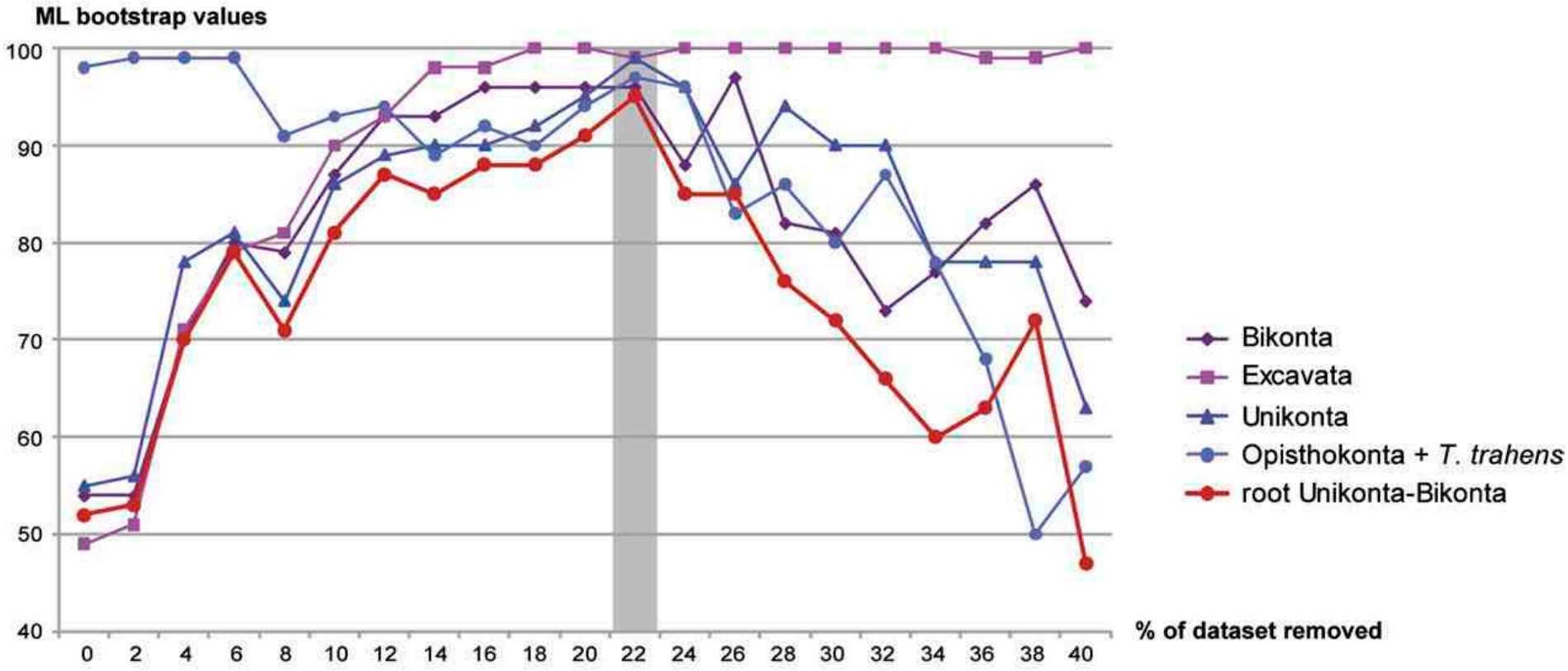


Les sites sous faible sélection sont plus saturés que ceux sous forte sélection.

adapté de P. G. Foster

Remèdes

2. élimination des sites à évolution rapide



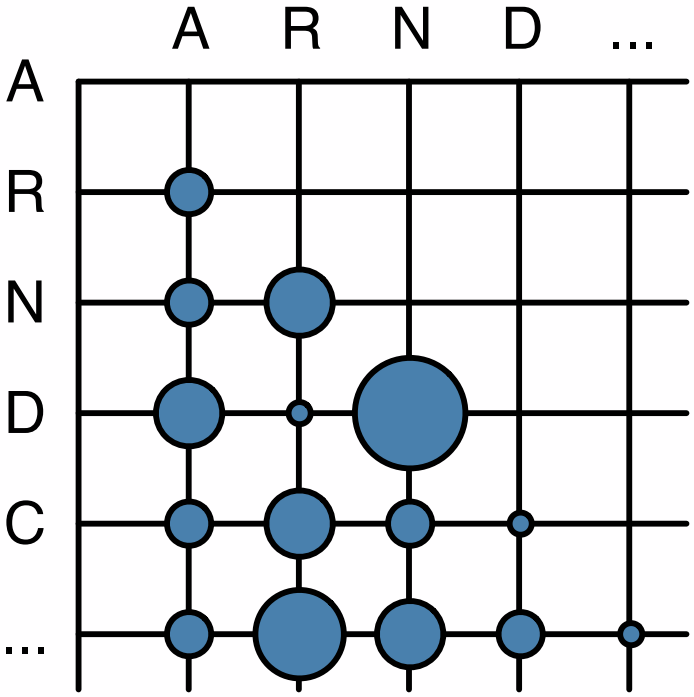
Le retrait des sites rapides améliore la robustesse de l'arbre.

Derelle & Lang (2012) *Mol Biol Evol* 29:1277-1289

Remèdes

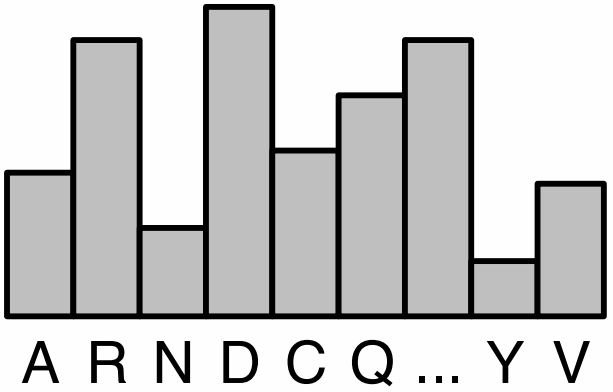
3. utilisation de modèles d'évolution sophistiqués

LG or WAG



1 global precomputed replacement matrix

+



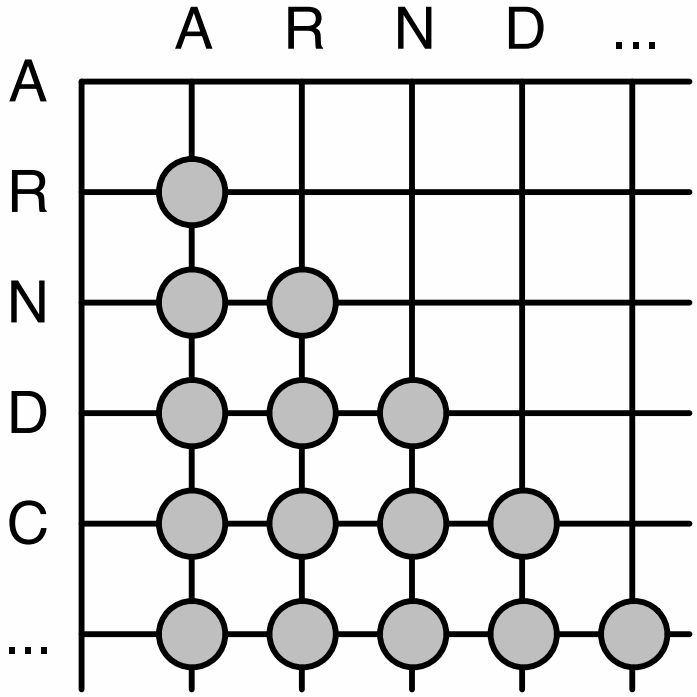
1 global compositional profile

Remèdes

3. utilisation de modèles d'évolution sophistiqués

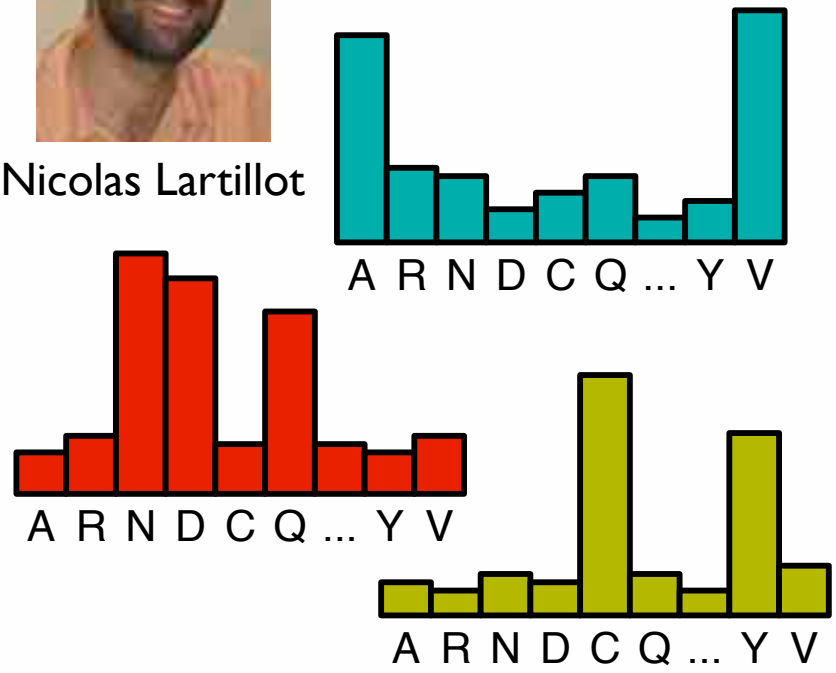
CAT

Nicolas Lartillot



1 global « flat »
replacement matrix

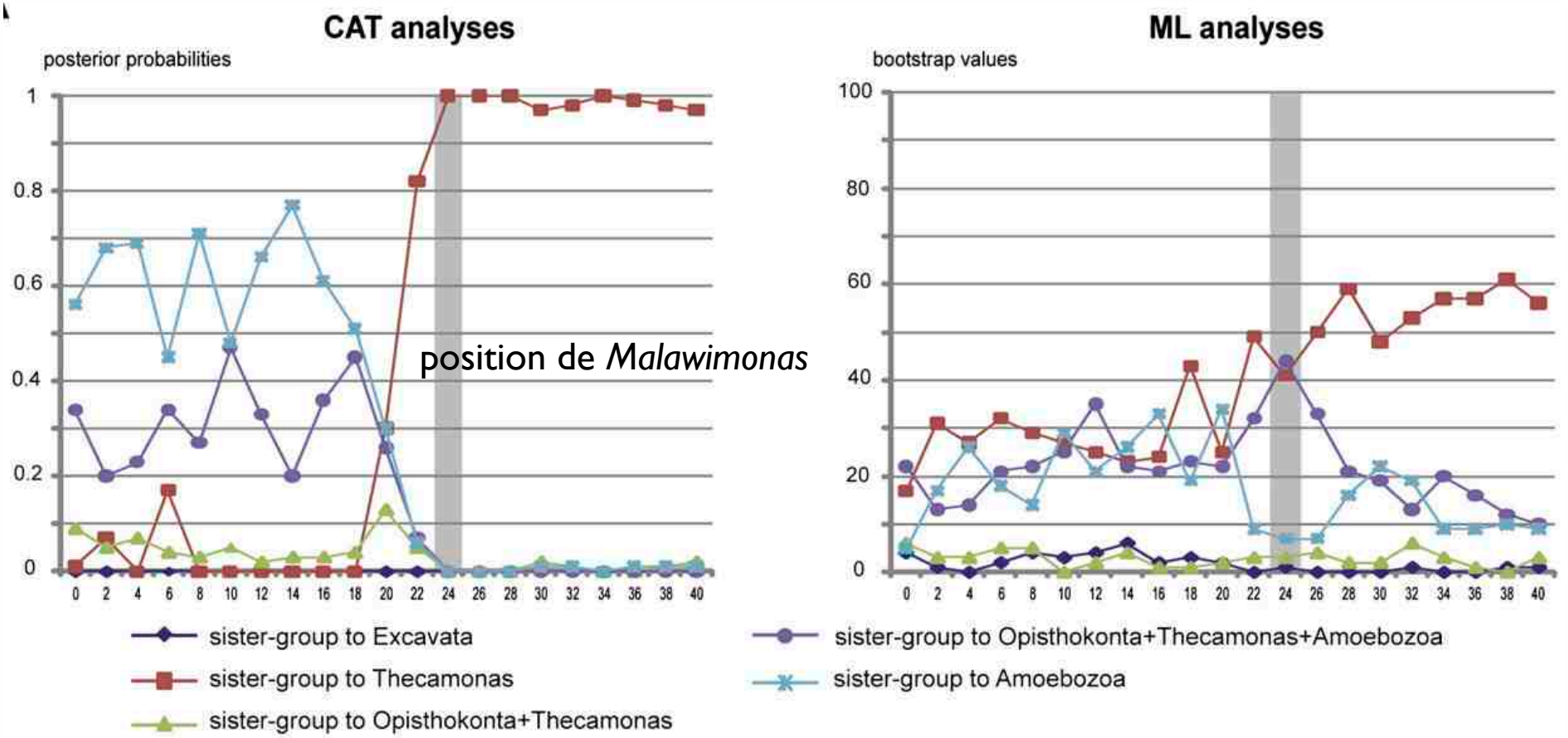
+



K distinct
compositional profiles

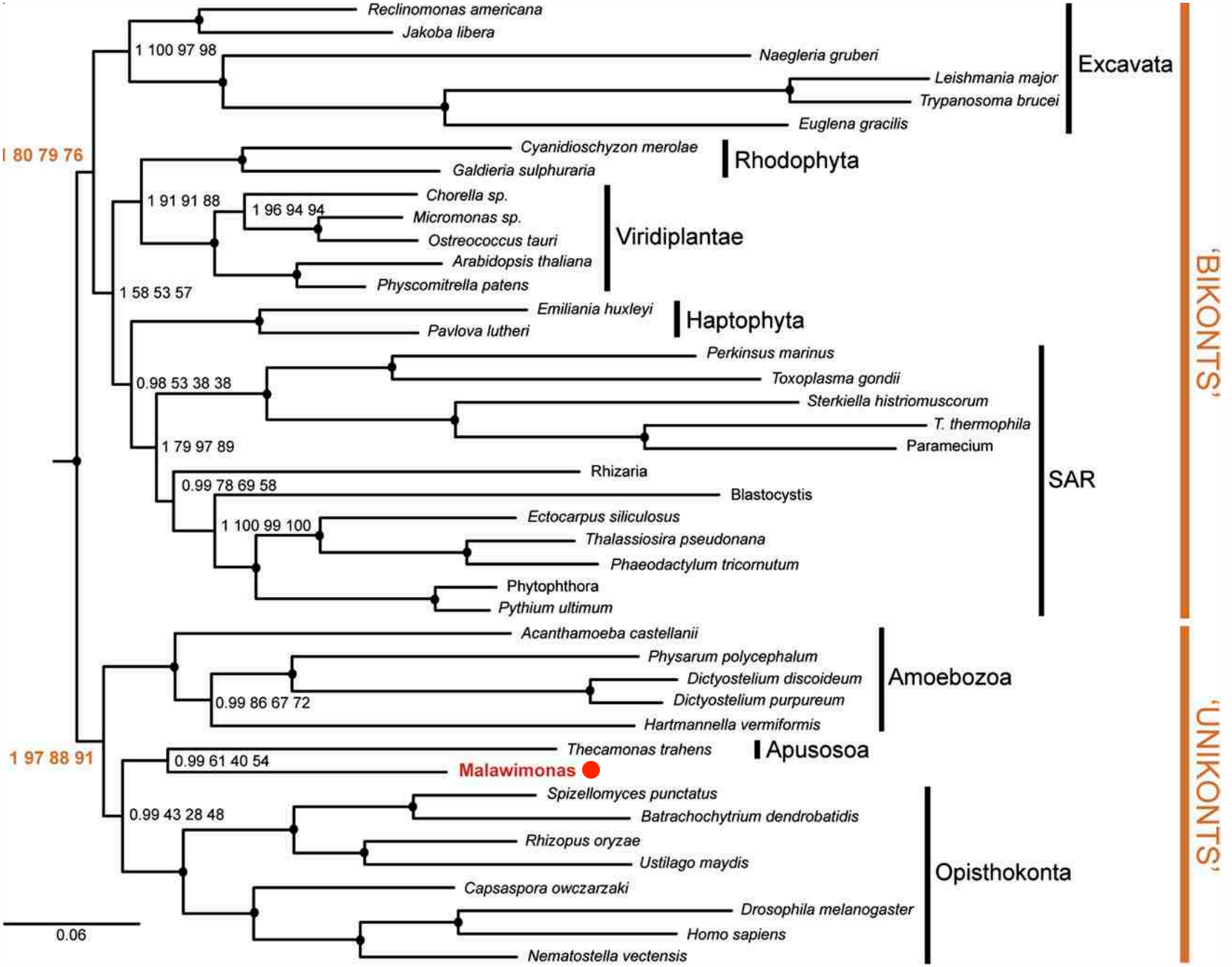
Remèdes

3. utilisation de modèles d'évolution sophistiqués



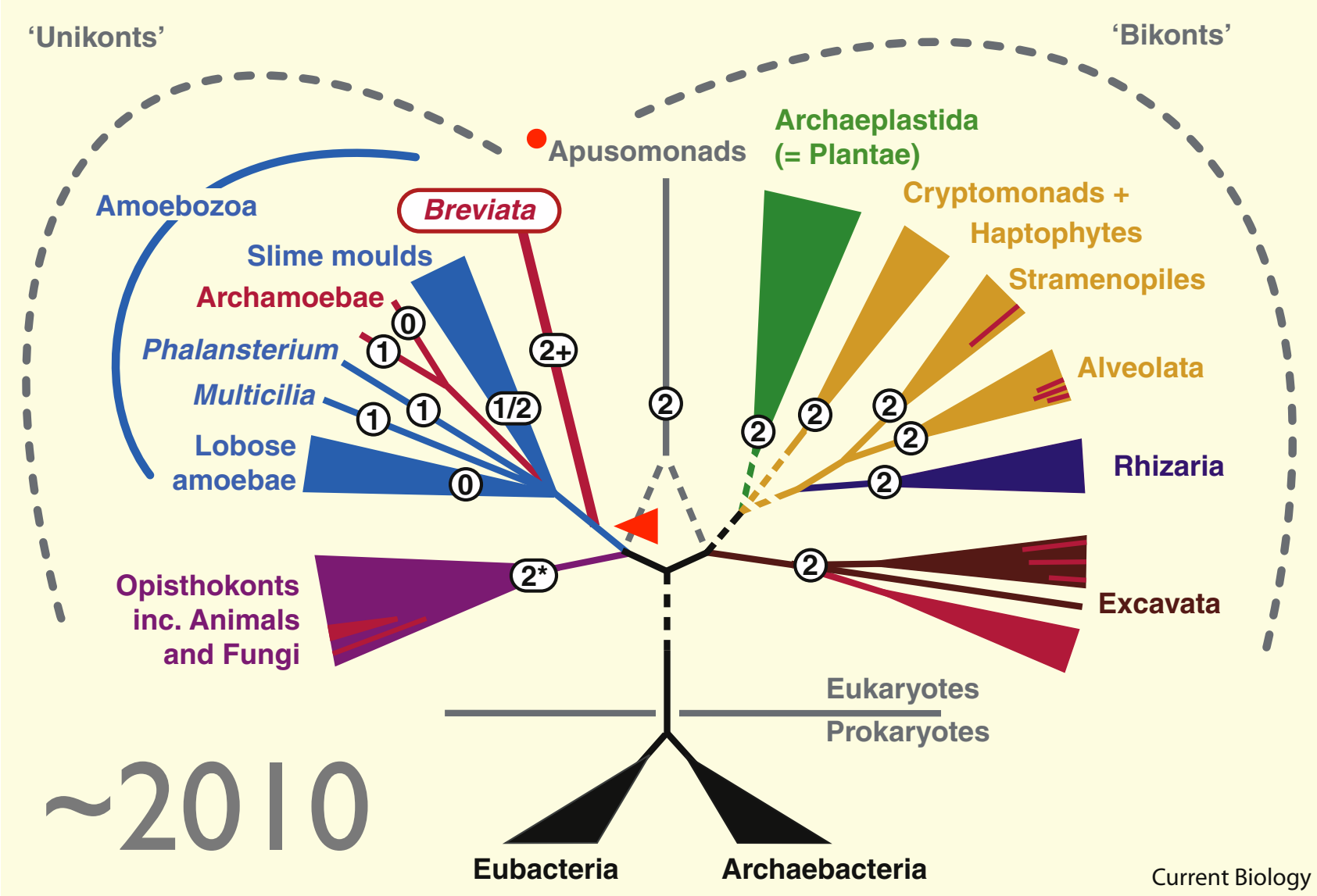
CAT améliore la détection des substitutions multiples.

Derelle & Lang (2012) *Mol Biol Evol* 29:1277-1289



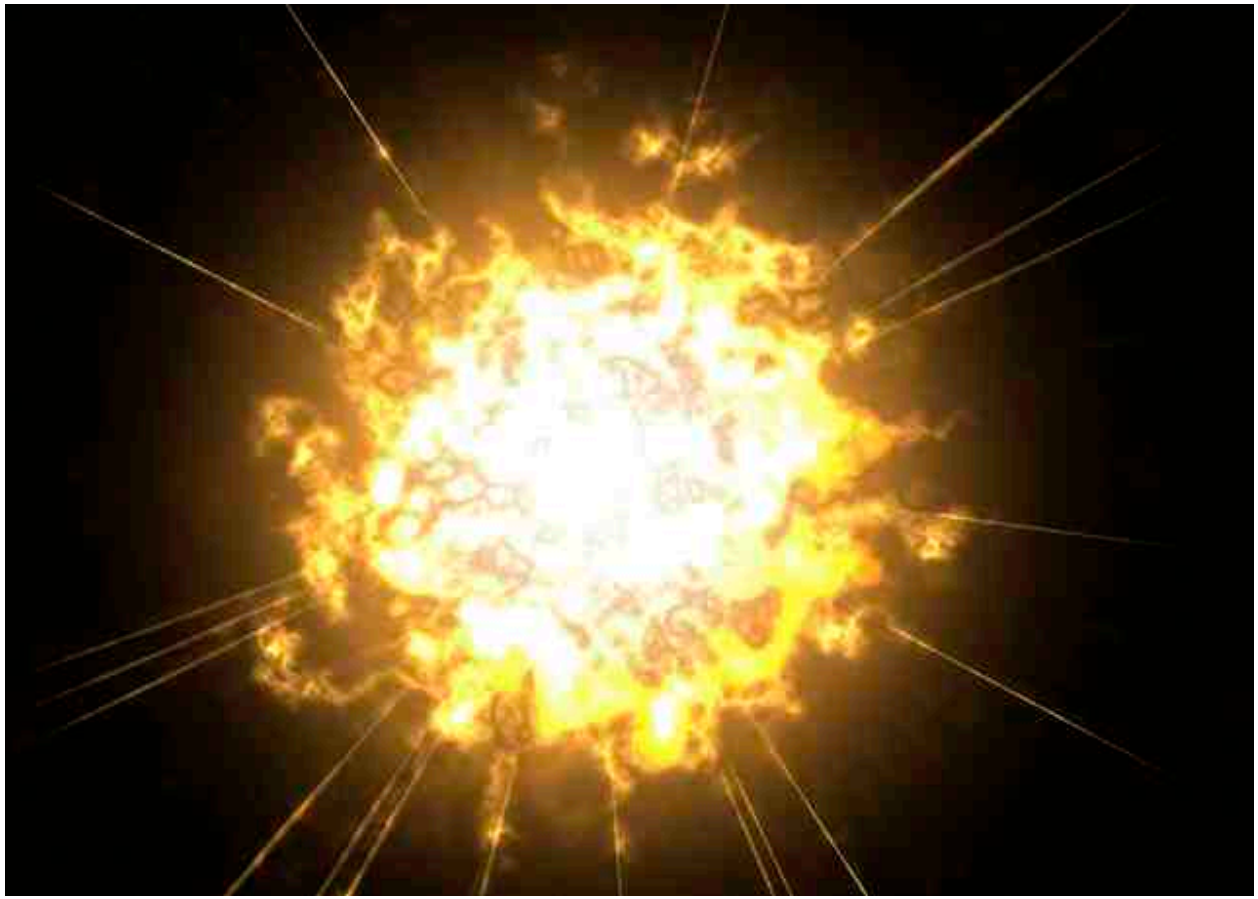
Derelle & Lang (2012) *Mol Biol Evol* 29:1277-1289

Explosion des groupes



modifié de Roger & Simpson (2009) *Curr Biol* 19:R165-7

Complication #1 : Big Bang



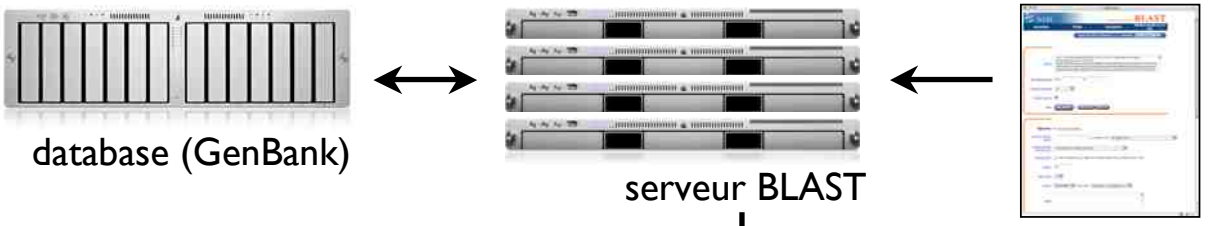
possible radiation explosive des Eucaryotes
aggravée par le signal non-phylogénétique

3

Transfert de gènes
Orthologie, paralogie et xénologie

Basic **L**ocal **A**lignment **S**earch **T**ool

- calcule la similarité entre deux séquences biologiques
- produit des alignements locaux : seule une portion de chaque séquence est alignée
- utilise des statistiques sophistiquées pour déterminer si un alignement pourrait avoir été obtenu par le seul hasard




```
S1 TTGACACCCTCCCAATTGTA
S2 ACCCCAGGCTTTACACAT

S1 TTGACACCCTCC-CAATTGTA ⇐global
       ||   ||    ||   |
S2 ACCCCAGGCTTTACACAT---    local
                           ⇓
S1 ---------TTGACACCCTCCCAATTGTA
             || ||||
S2 ACCCCAGGCTTTACACAT----------
```

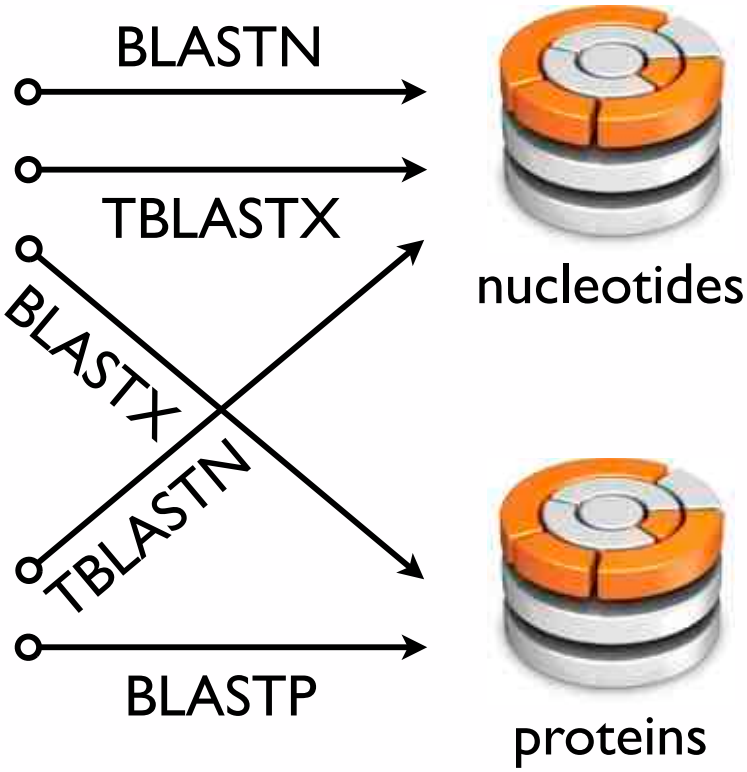
BLAST

à chaque recherche sa variante de BLAST

query

```
>DNA_seq
TATGGCAATTAAAATTGGTATCAATGGTTTTGG
TCGTATCGGCCGTATCGTATTCCGTGCAGCACA
ACACCGTGATGACATTGAAGTTGTAGGTATTAA
CGACTTAATCGACGTTGA
```

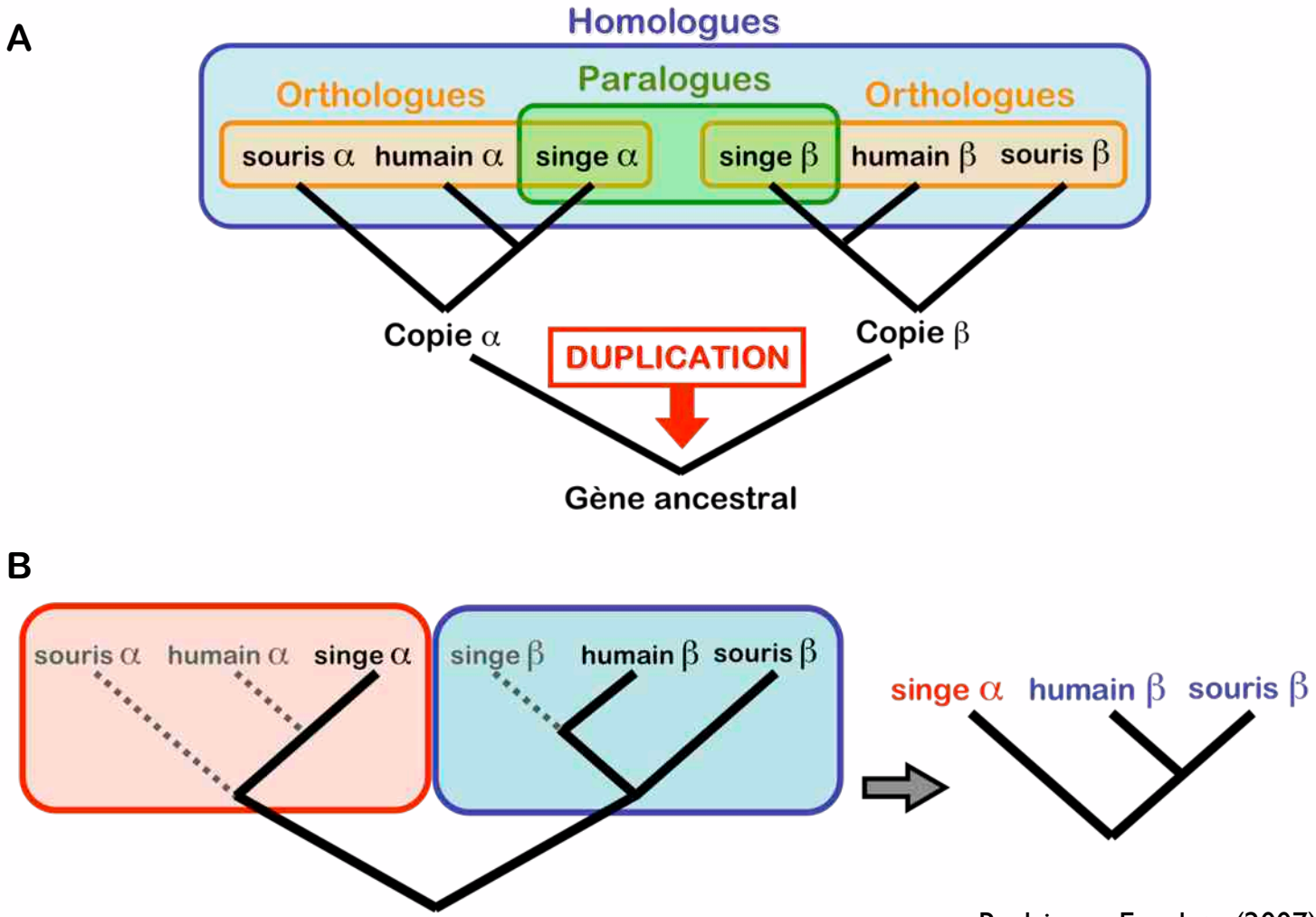
```
>protein_seq
MLQTAPMLPGLGPHLVPQLGALASASRLLGSIA
SVPPQHGGAGFQAVRGFATGAVSTPAASSPGHK
PAATHAPPTRLDLKPGAGSFAAGAVAPHPGINP
ARMAADSASAA
```



Les recherches de type protéique sont plus sensibles.

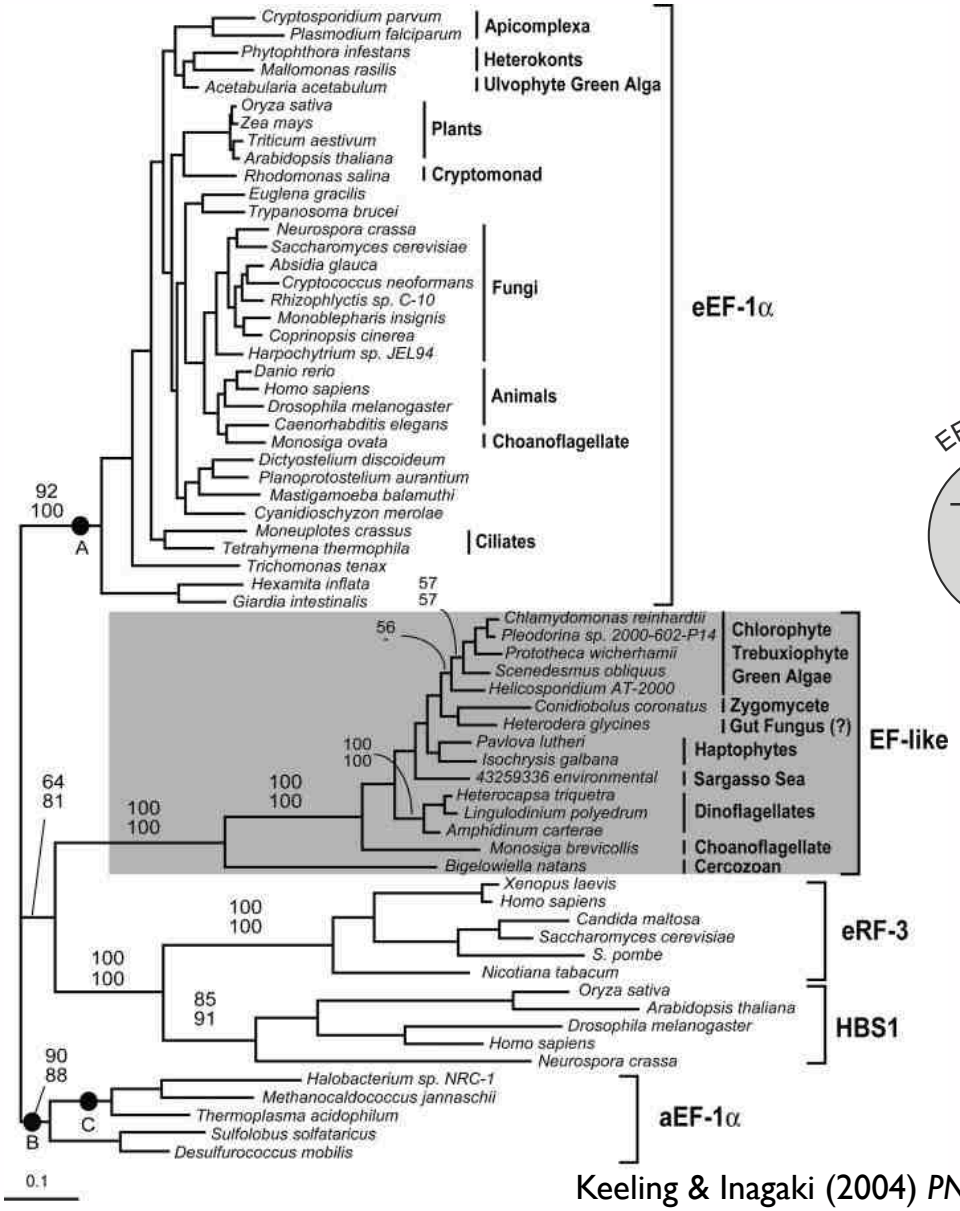
Orthologie et Paralogie

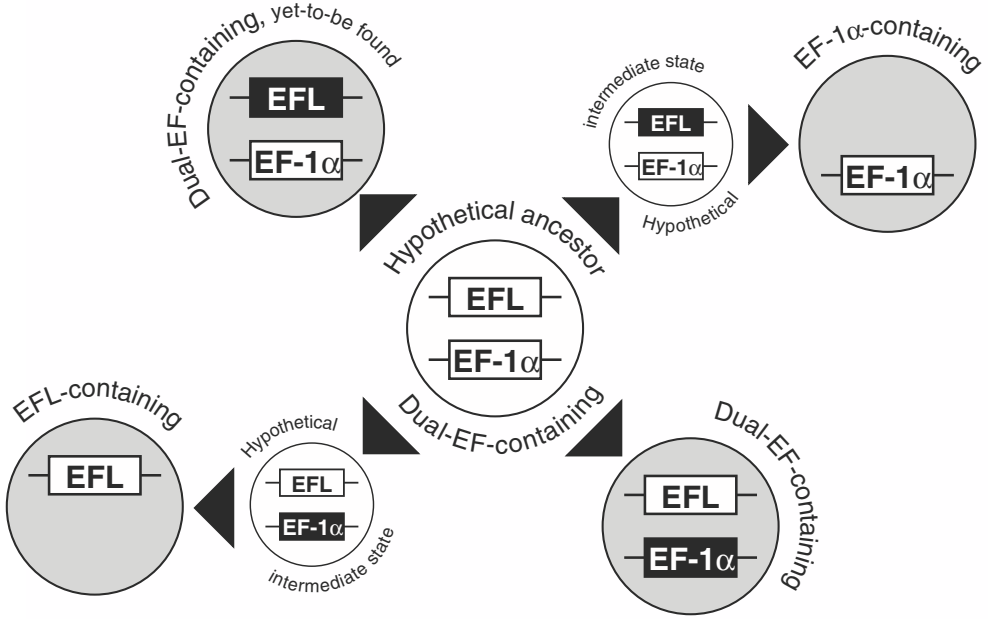
les dangers de la paralogie cachée



Rodriguez-Ezpeleta (2007) *Thèse de Doctorat*

Orthologie et Paralogie



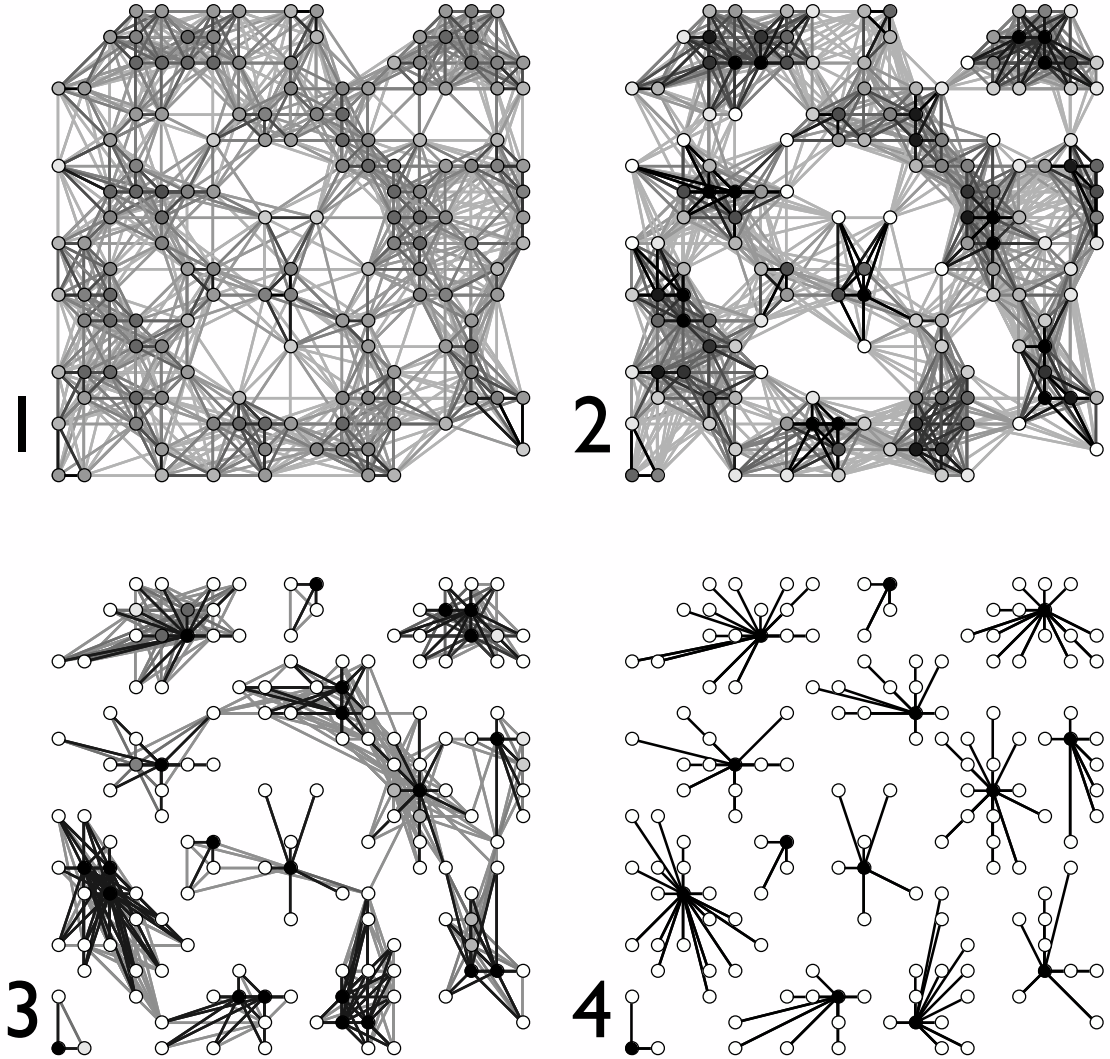


un exemple de paralogie cachée profonde :
le couple eEF-1a / EF-like

Keeling & Inagaki (2004) *PNAS USA* 101:15380-15385; Kamikawa et al. (2013) *BMC Evol Biol* 13:131

Orthologie et Paralogie

détermination des groupes orthologues



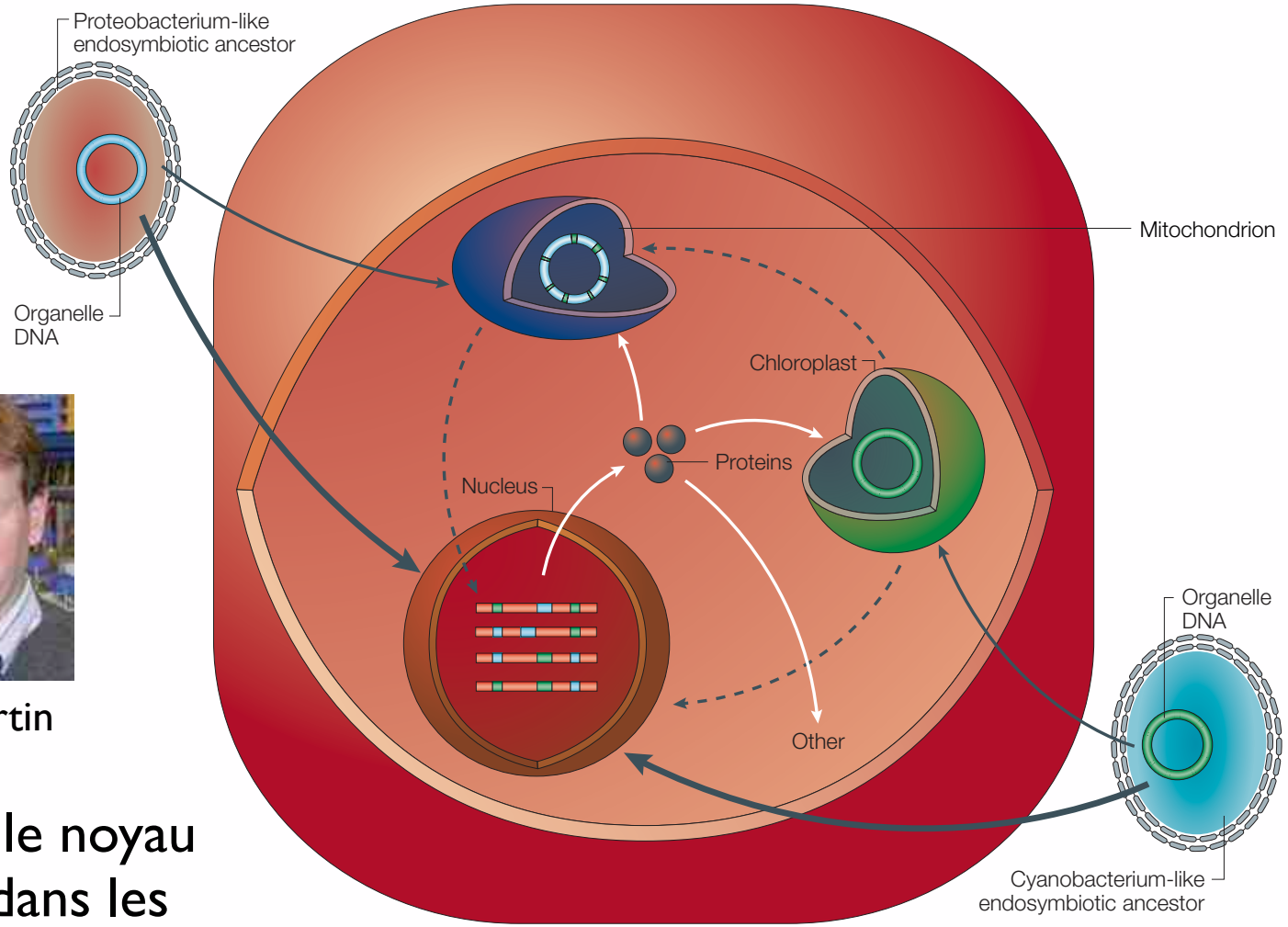
OrthoMCL objective le processus, mais n'est pas sans faille.

van Dongen (2000) *Thèse de Doctorat*

Endosymbioses primaires

Lynn Margulis

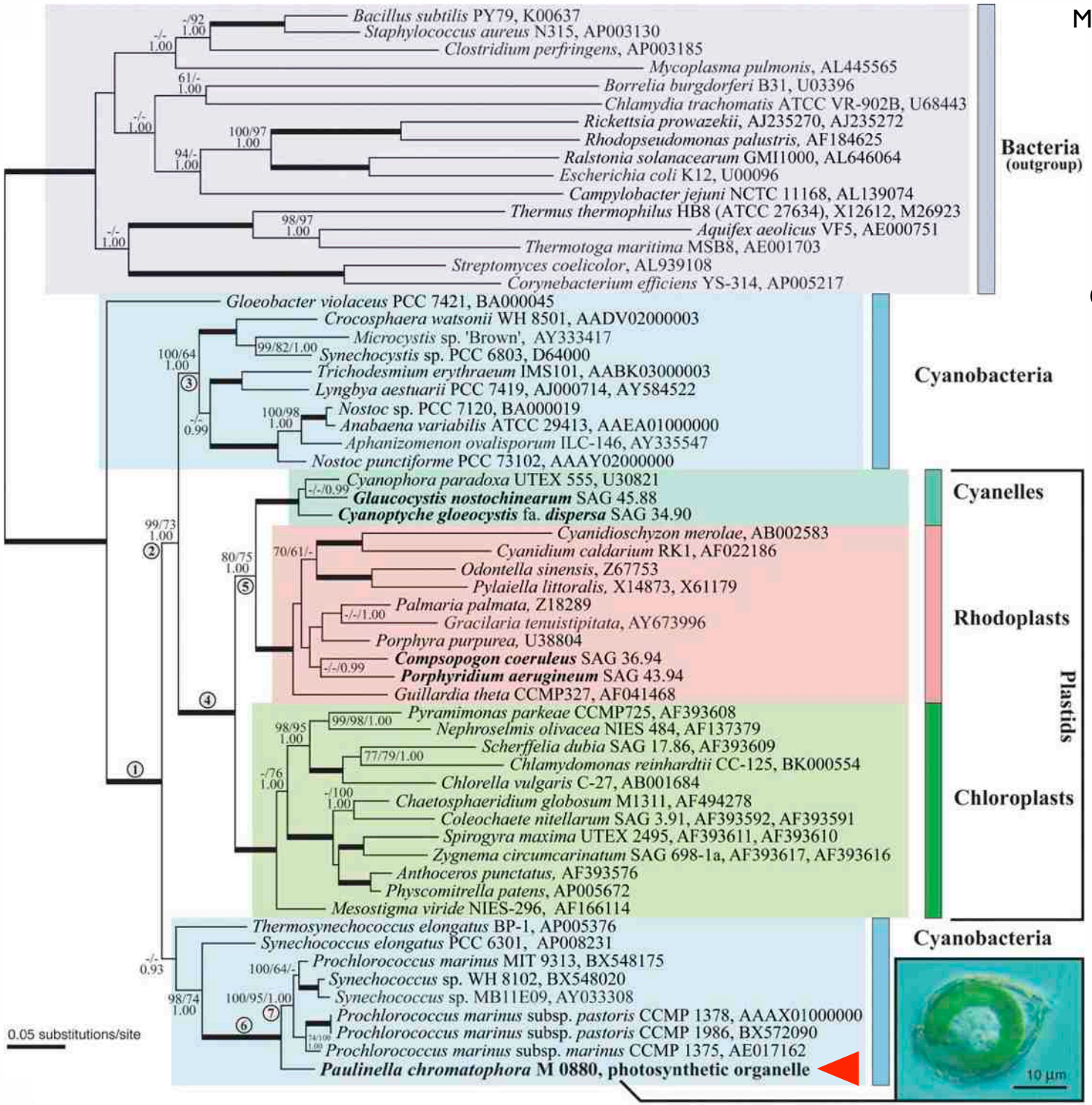
Bill Martin



relocalisation dans le noyau de gènes encodés dans les organites endosymbiotiques

Margulis (1968) *Science* 161:1020-1022; Timmis et al. (2004) *Nat Rev Genet* 5:123-135

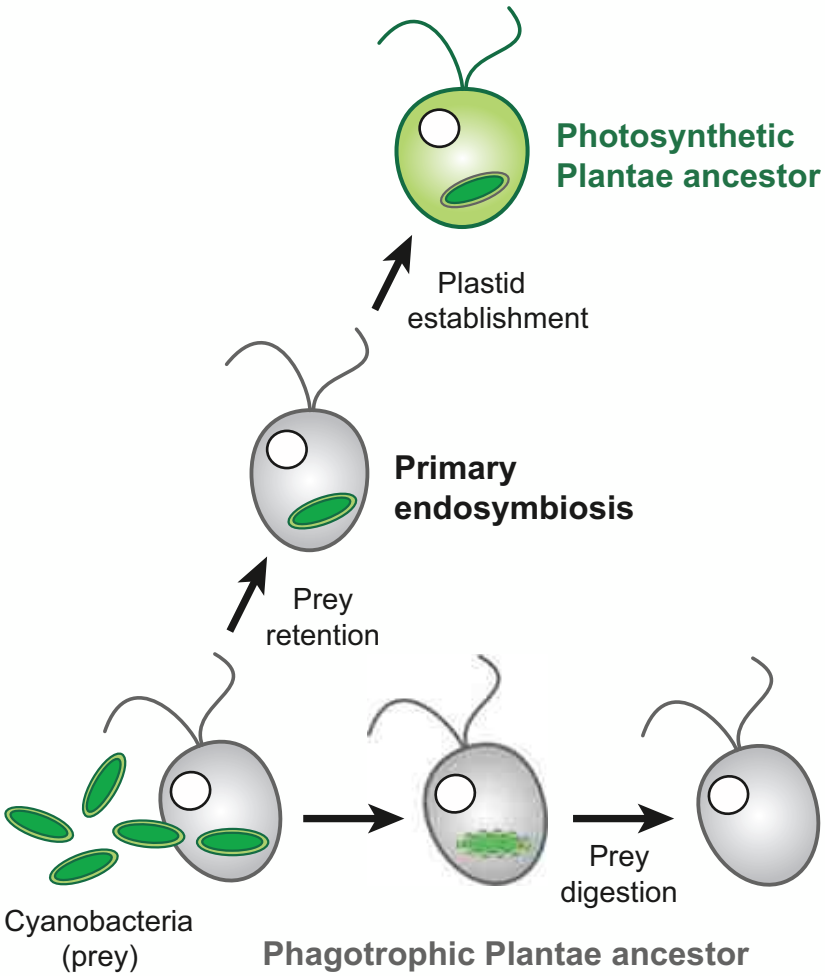
Marin et al. (2005) *Protist* 156:425-432



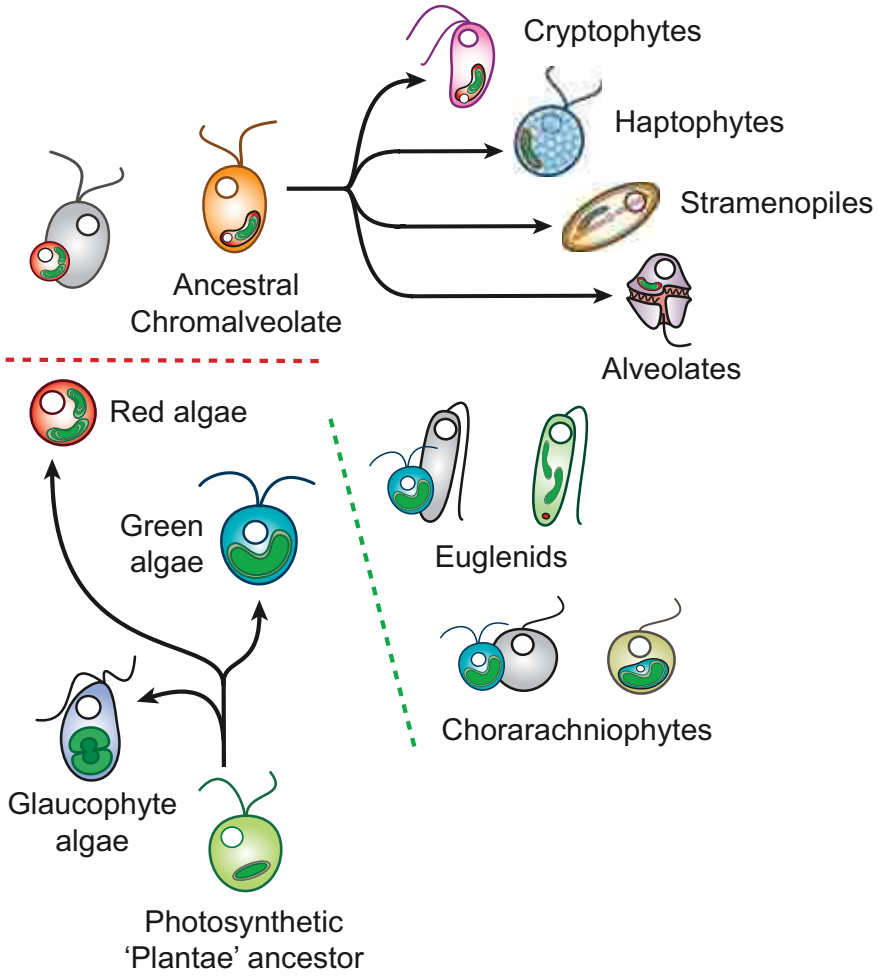
opéron rRNA

une seconde endosymbiose primaire : *Paulinella*

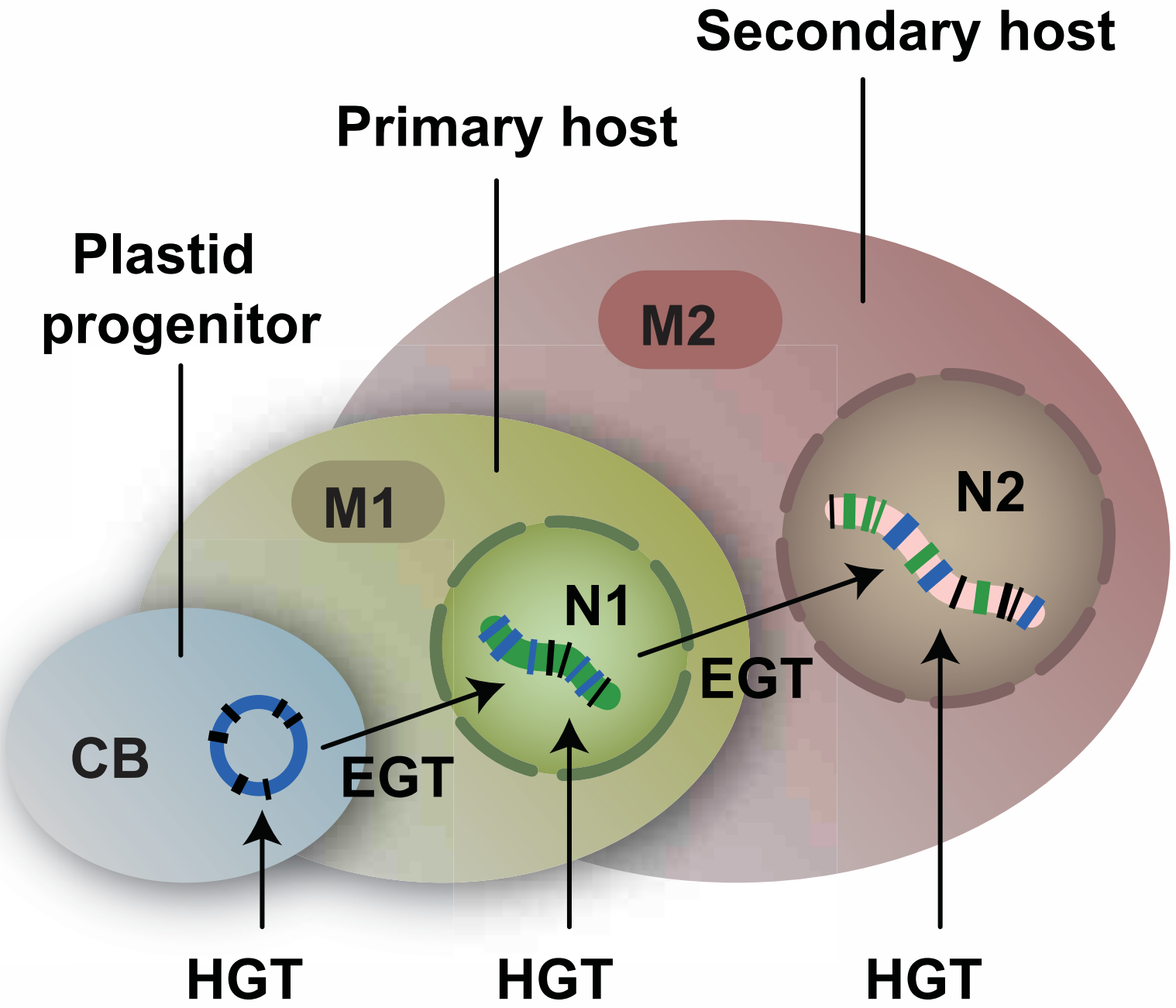
endosymbiose primaire Plantae



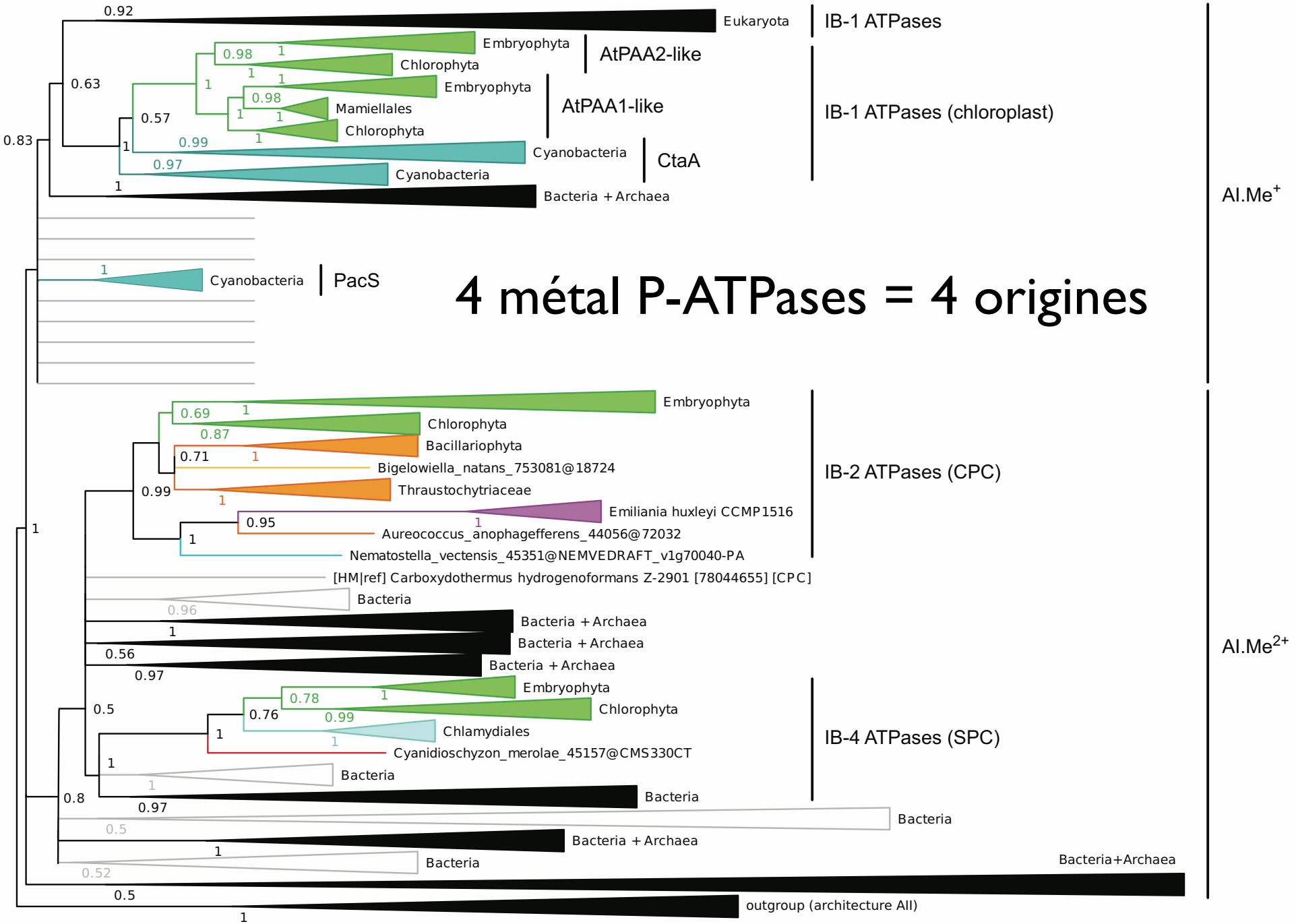
endosymbiose secondaire « Chromalveolés »



Gibbs (1978) *Can J Bot* 56:2883-2889; Reyes-Prieto et al. (2007) *Annu Rev Genet* 41:147-168

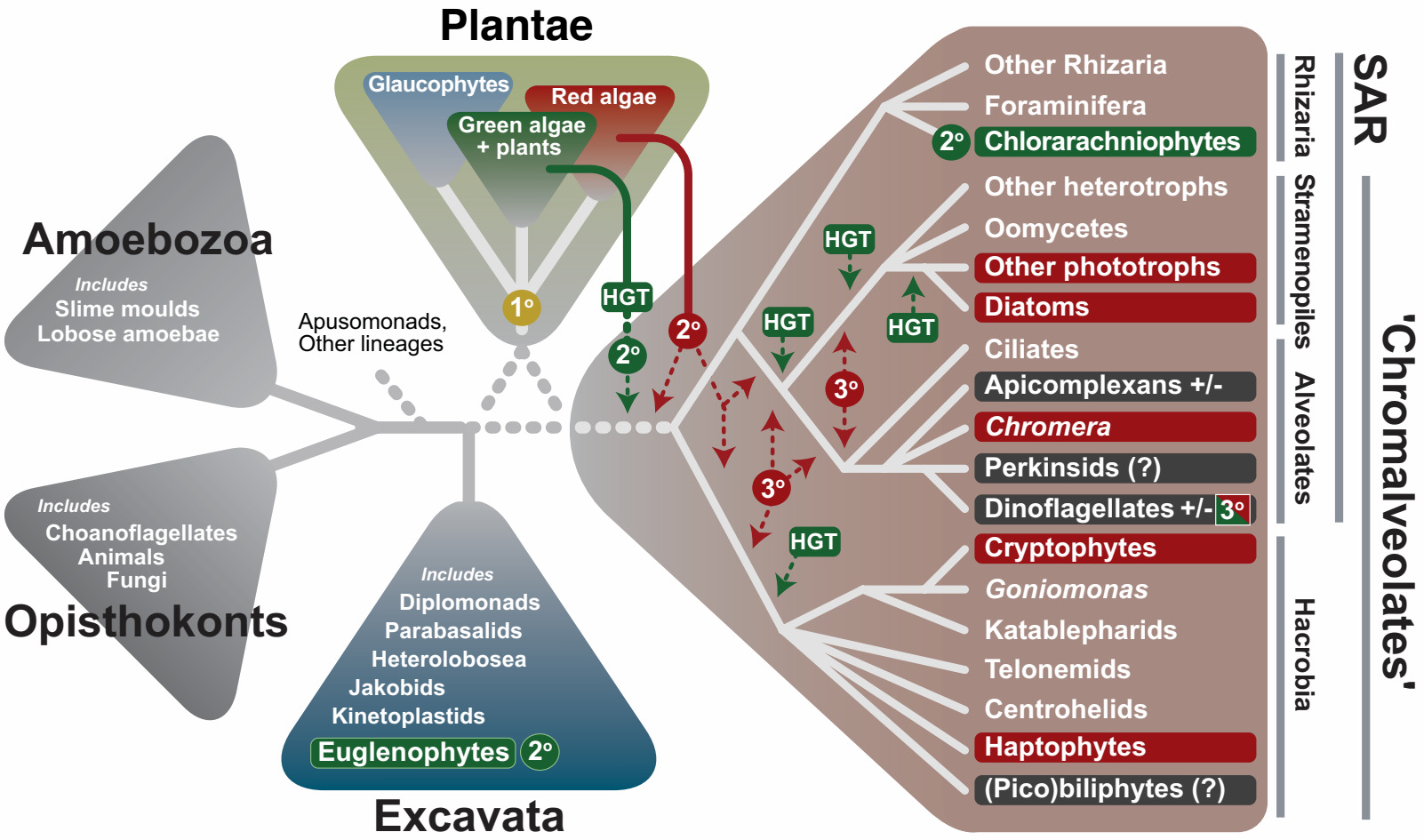


Elias & Archibald (2009) *BioEssays* 31:1273-1279



Hanikenne & Baurain (2013) *Front Plant Sci* 4:544

Complication #2 : EGT

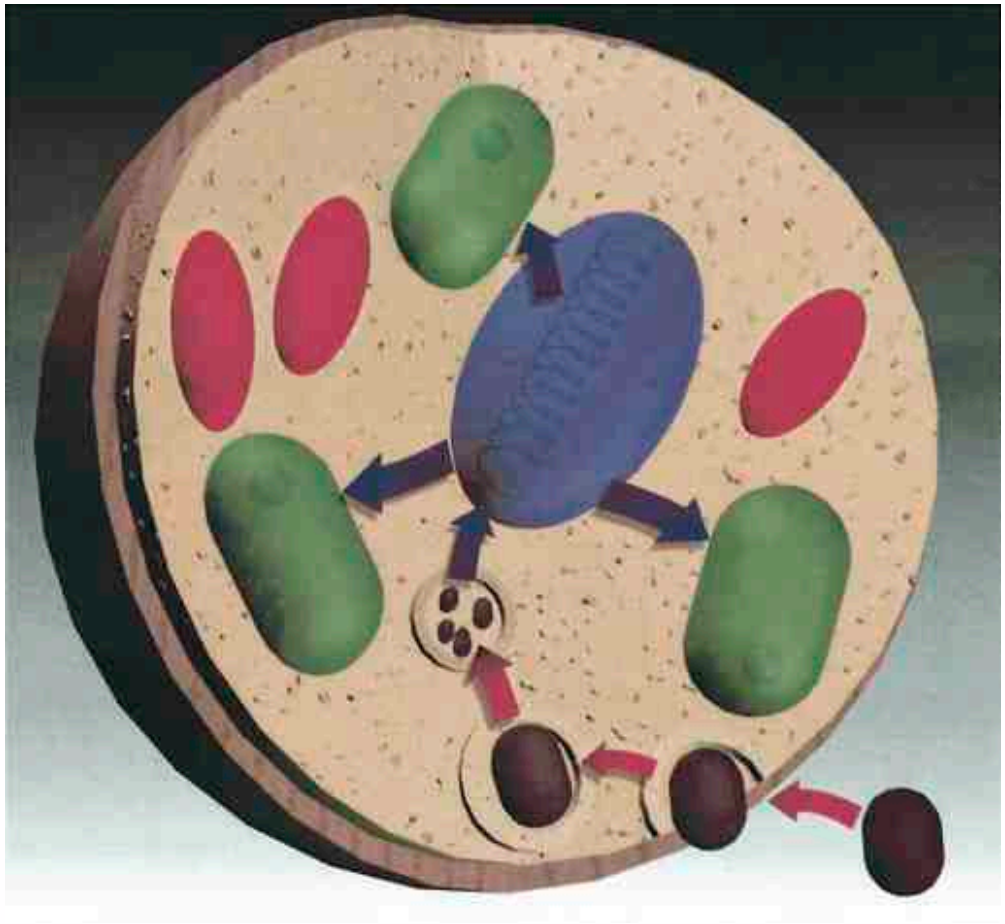


chimérisme complexe (mais biaisé) des génomes algaux
lié aux multiples transferts indépendants de plastes

modifié d'Elias & Archibald (2009) *BioEssays* 31:1273-1279; Petersen et al. (2014) *Genome Biol Evol* 6:666-684

Complication #3 : HGT

W. Ford Doolittle

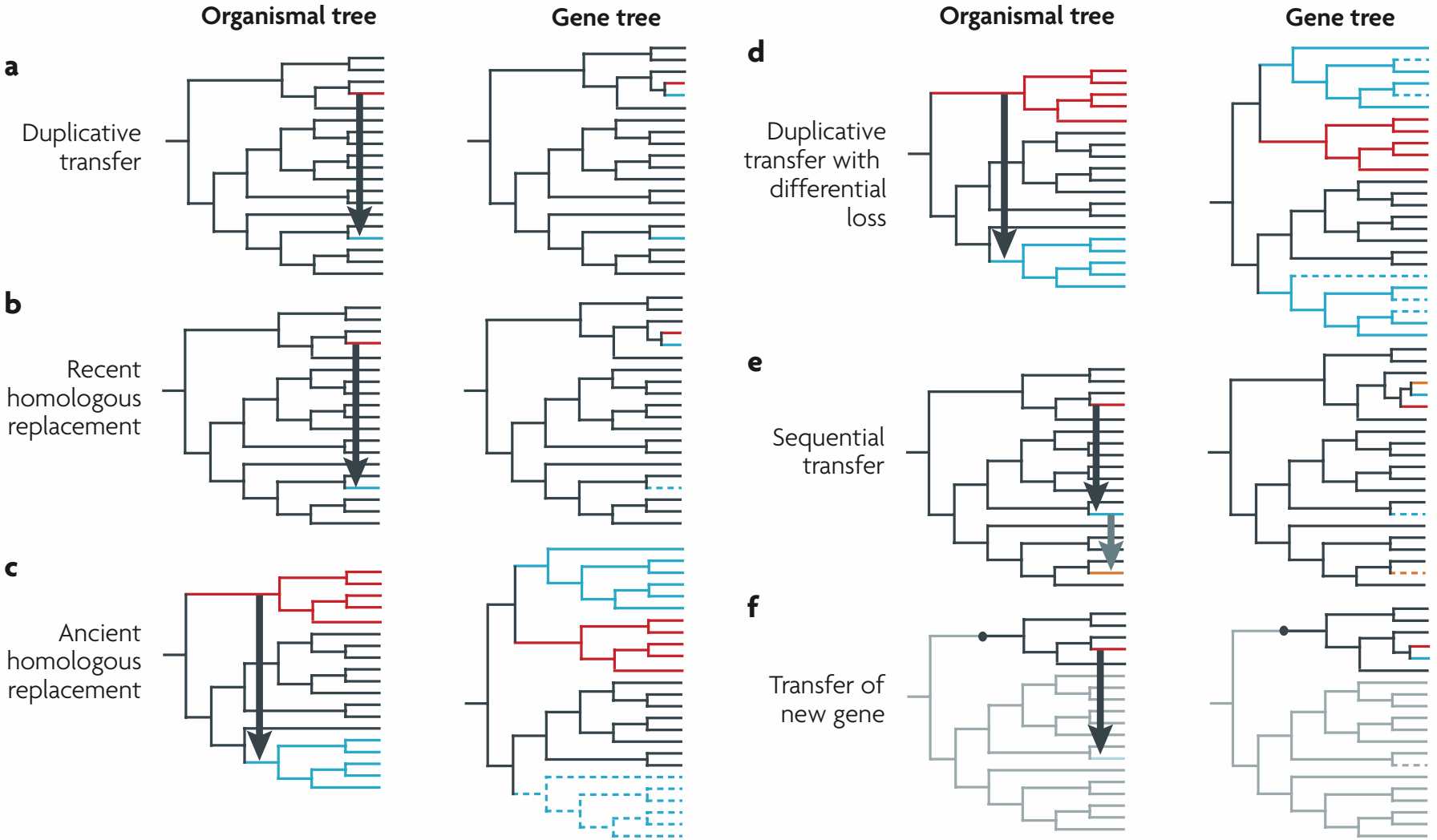


transfert horizontal de gènes
lié au « You are what you eat »

Doolittle (1998) *Trends Genet* 14:307-311

Transfert de gènes

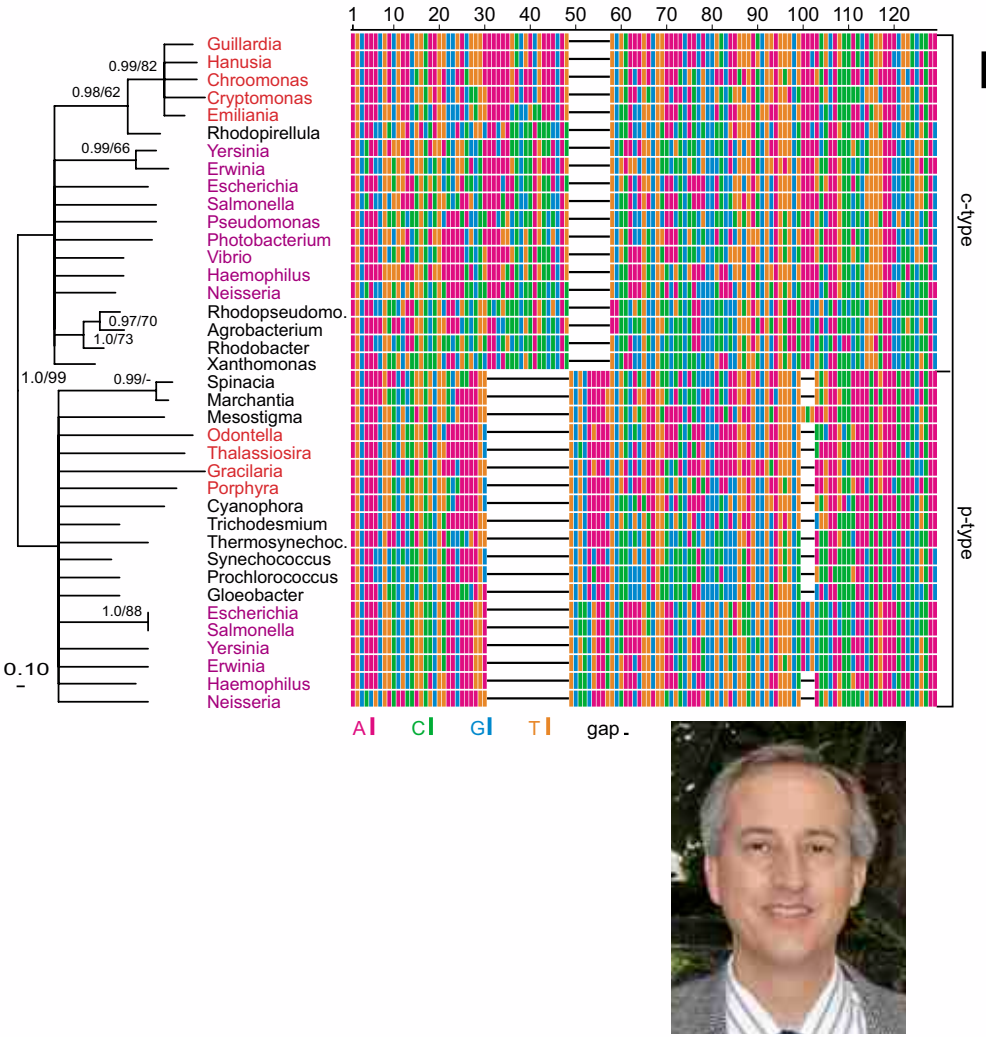
évolution des gènes ≠ évolution des organismes



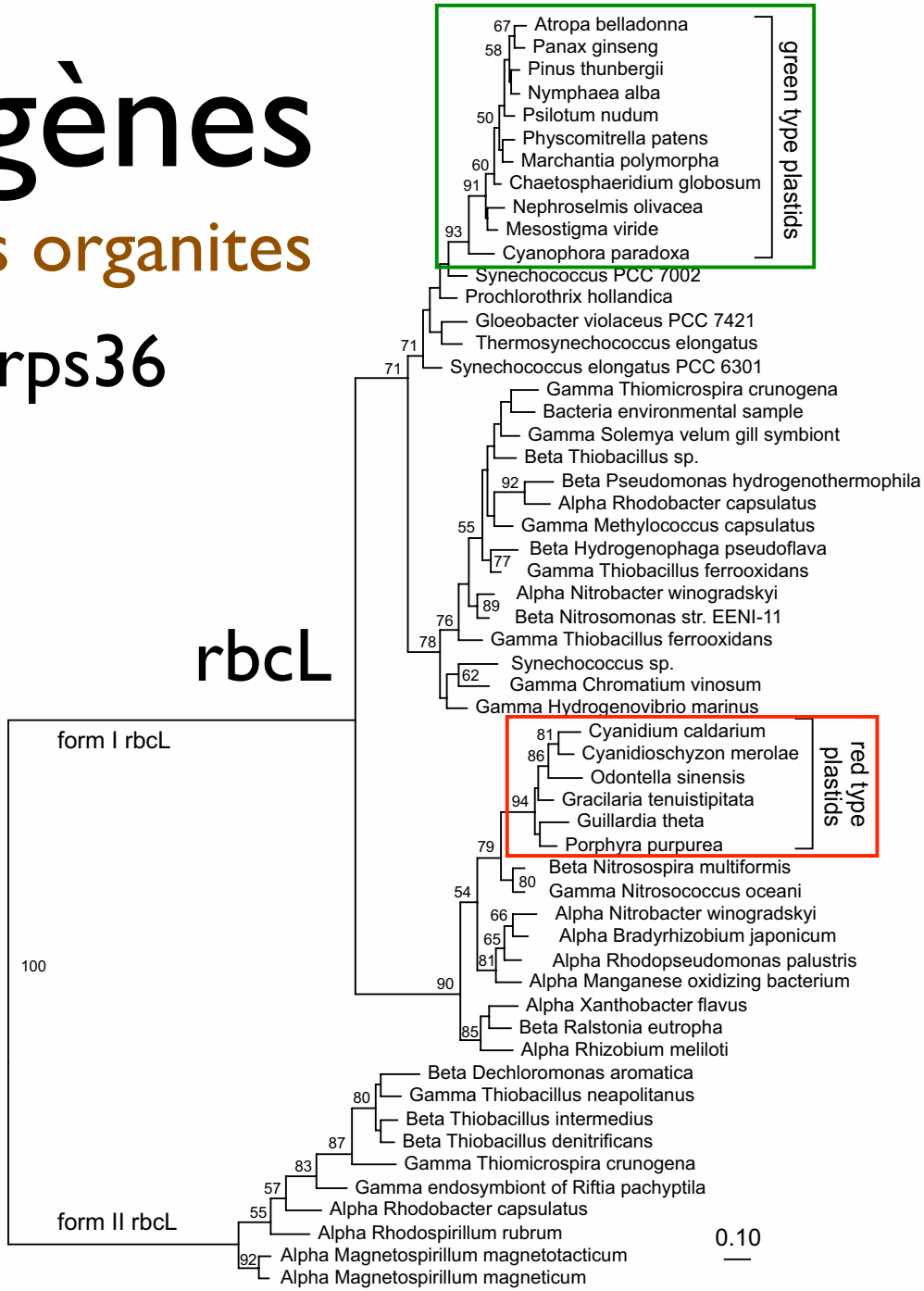
Keeling & Palmer (2008) *Nat Rev Genet* 9:605-618

Transfert de gènes

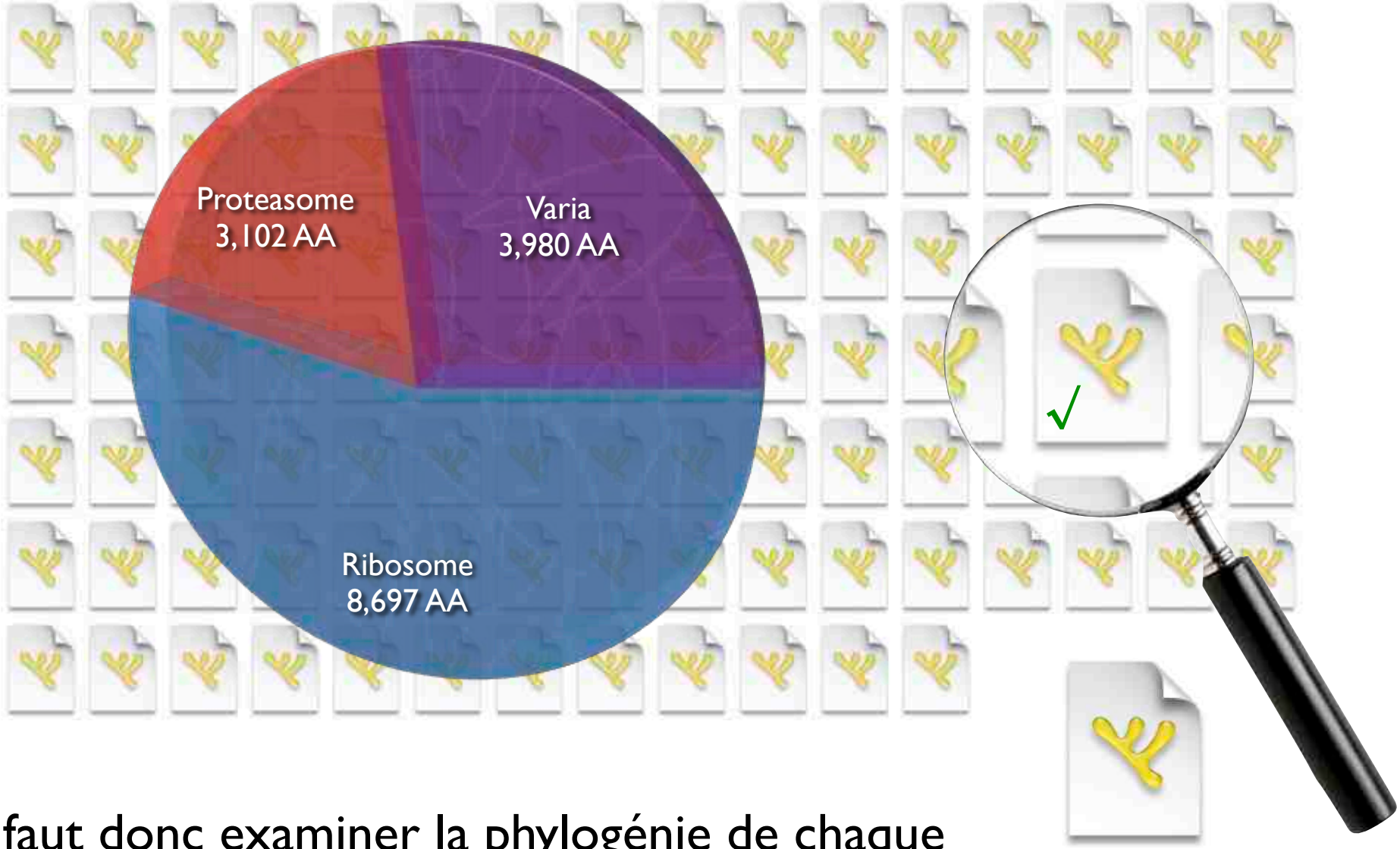
aussi dans les génomes des organites



Jeff Palmer

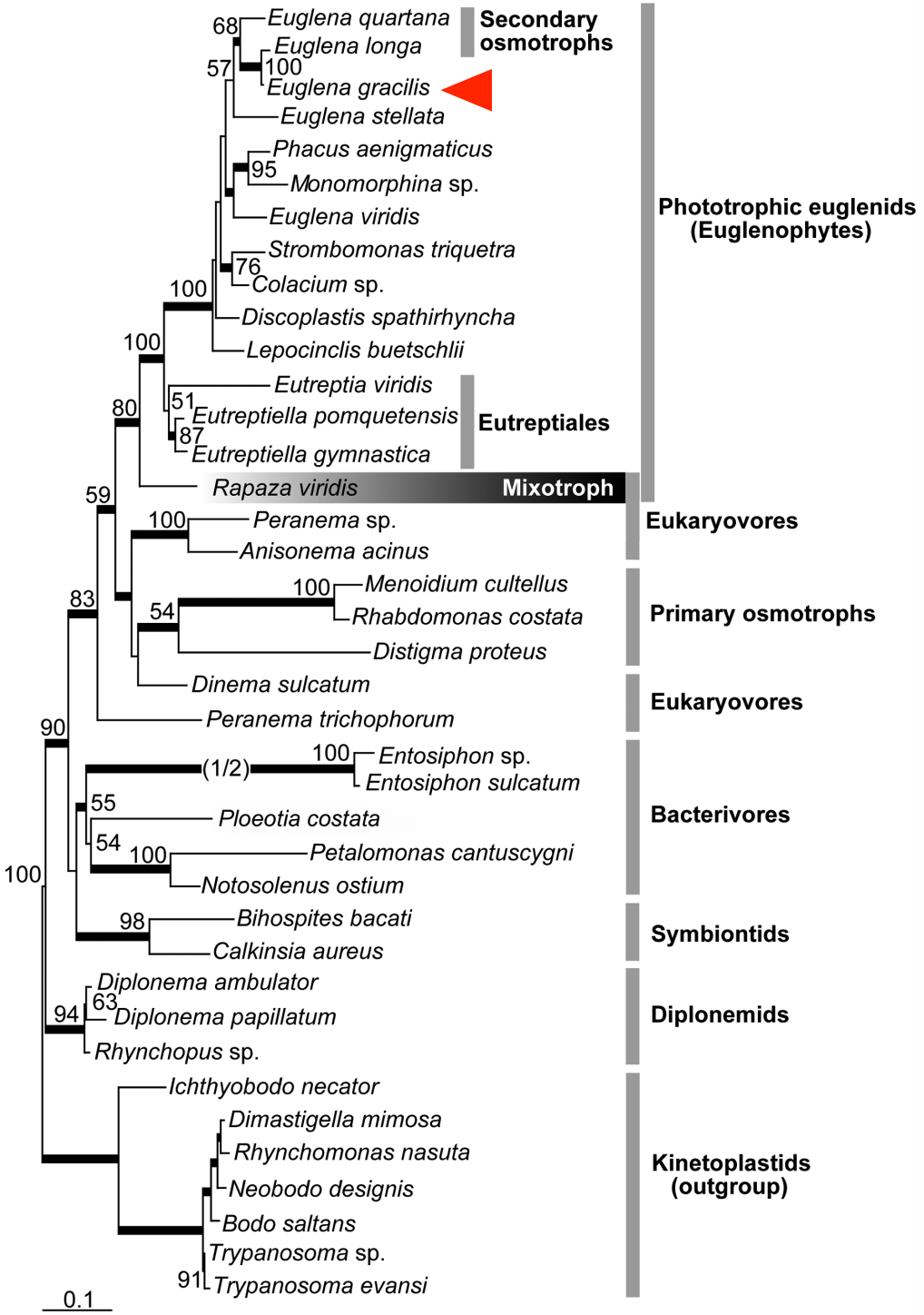


Rice et Palmer (2006) *BMC Biol* 4:31



Il faut donc examiner la phylogénie de chaque gène et la comparer à celle obtenue à partir de la supermatrice (analyse de congruence).

Baurain et al. (2010) *Mol Biol Evol* 27:1698-1709



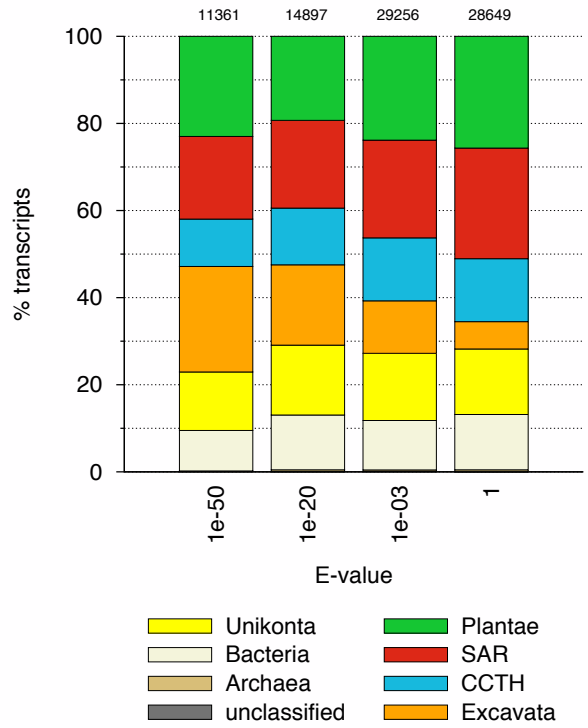
une algue « verte » au passé trouble : *Euglena gracilis* (Discoba, Excavata)

Yamaguchi et al. (2012) *BMC Evol Biol* 12:29

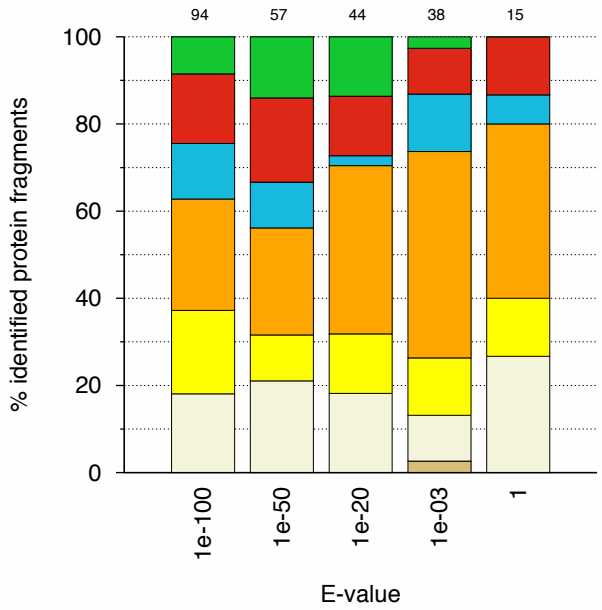
Transfert de gènes

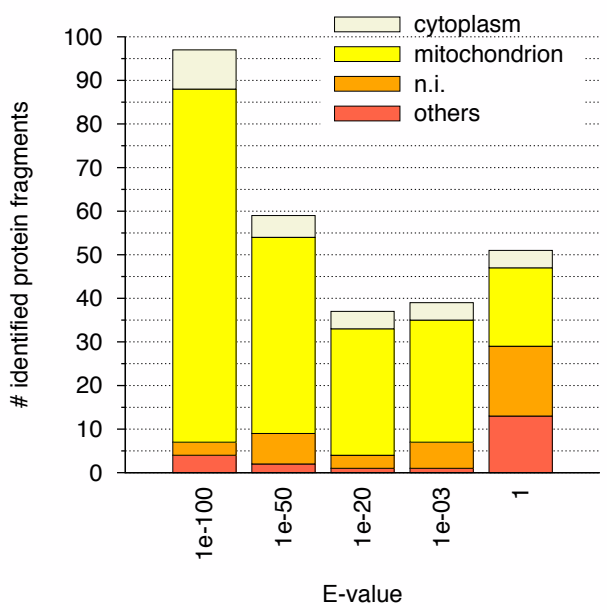
origines diverses des gènes chez *Euglena gracilis*

taxonomie des transcrits



taxonomie et localisation des protéines mitochondriales

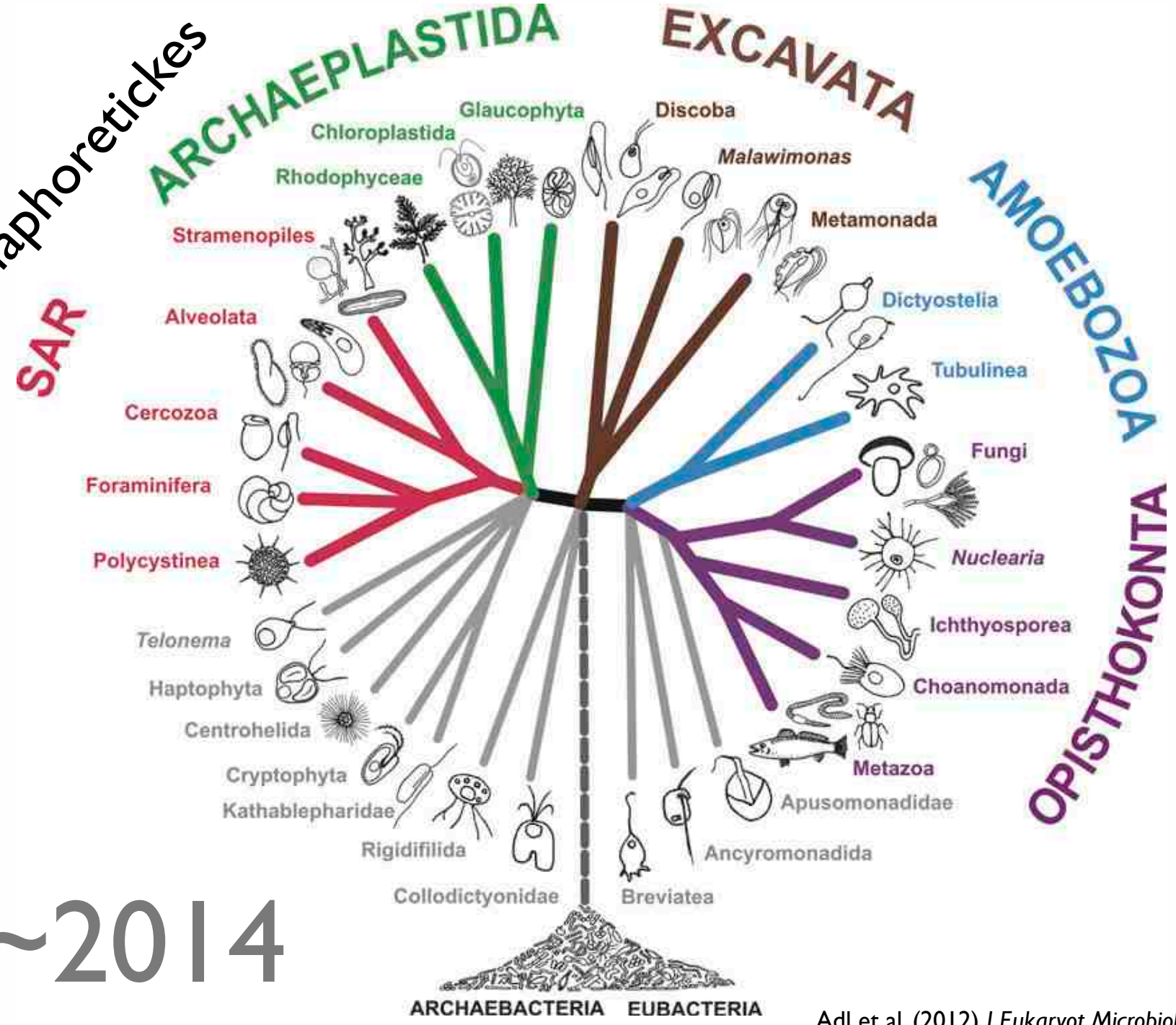




combinaison EGT/HGT

Perez et al. (2014) *Protist 2014 Conference*

Diaphoretickes



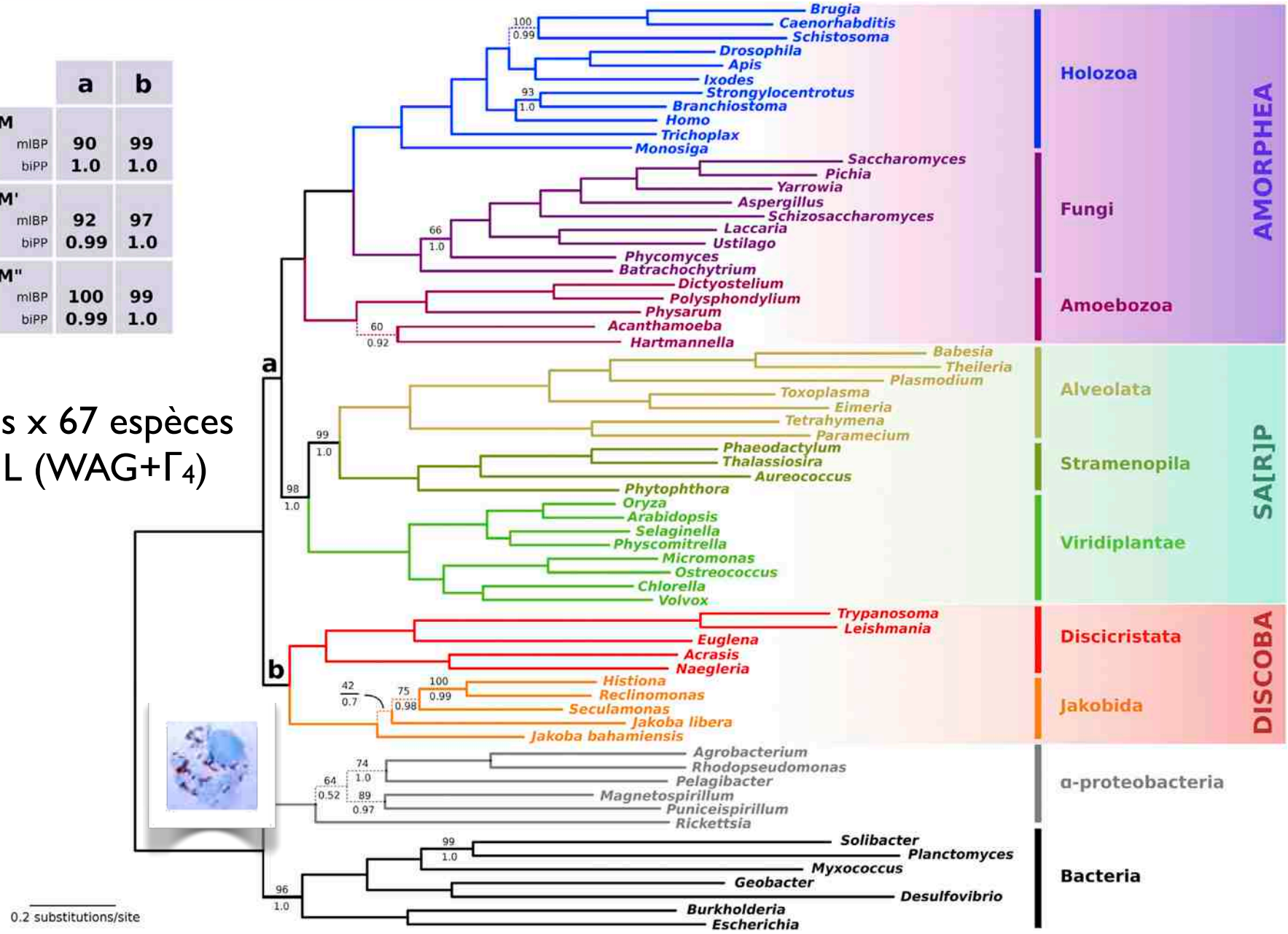
Amorphea

~2014

Adl et al. (2012) *J Eukaryot Microbiol* 59: 429-493

	a	b
M		
mlBP	90	99
biPP	1.0	1.0
M'		
mlBP	92	97
biPP	0.99	1.0
M"		
mlBP	100	99
biPP	0.99	1.0

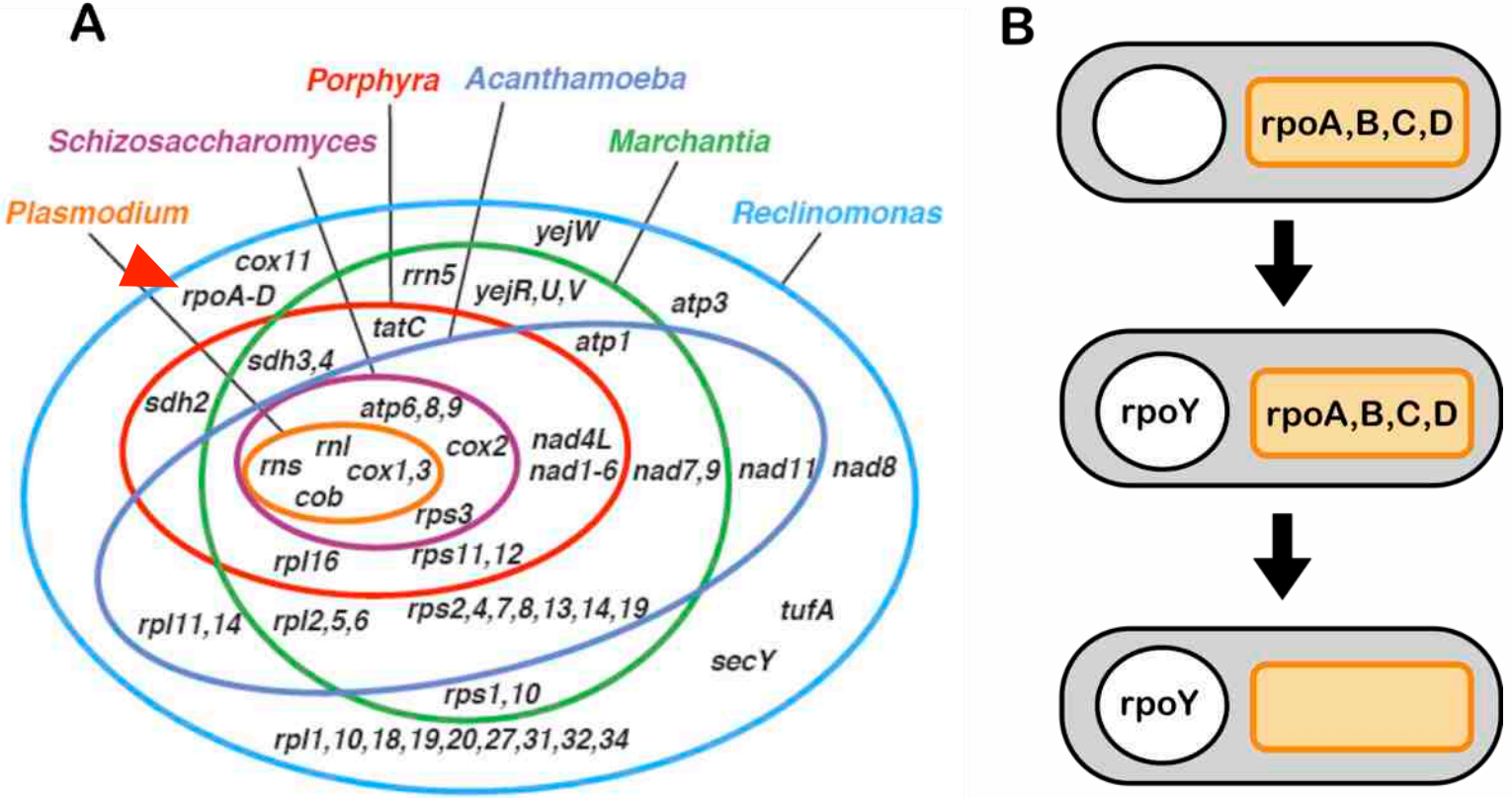
37 gènes x 67 espèces
RAxML (WAG+Γ_4)



position de la racine des eucaryotes dans les Excavata ?

He et al. (2014) *Curr Biol* 24:465-470

Discoba : Jakobida



Les jakobidés sont les seuls eucaryotes avec une ARN polymérase alpha-protéobactérienne. Tous les autres ont une enzyme (nucléaire) transférée depuis un bactériophage.

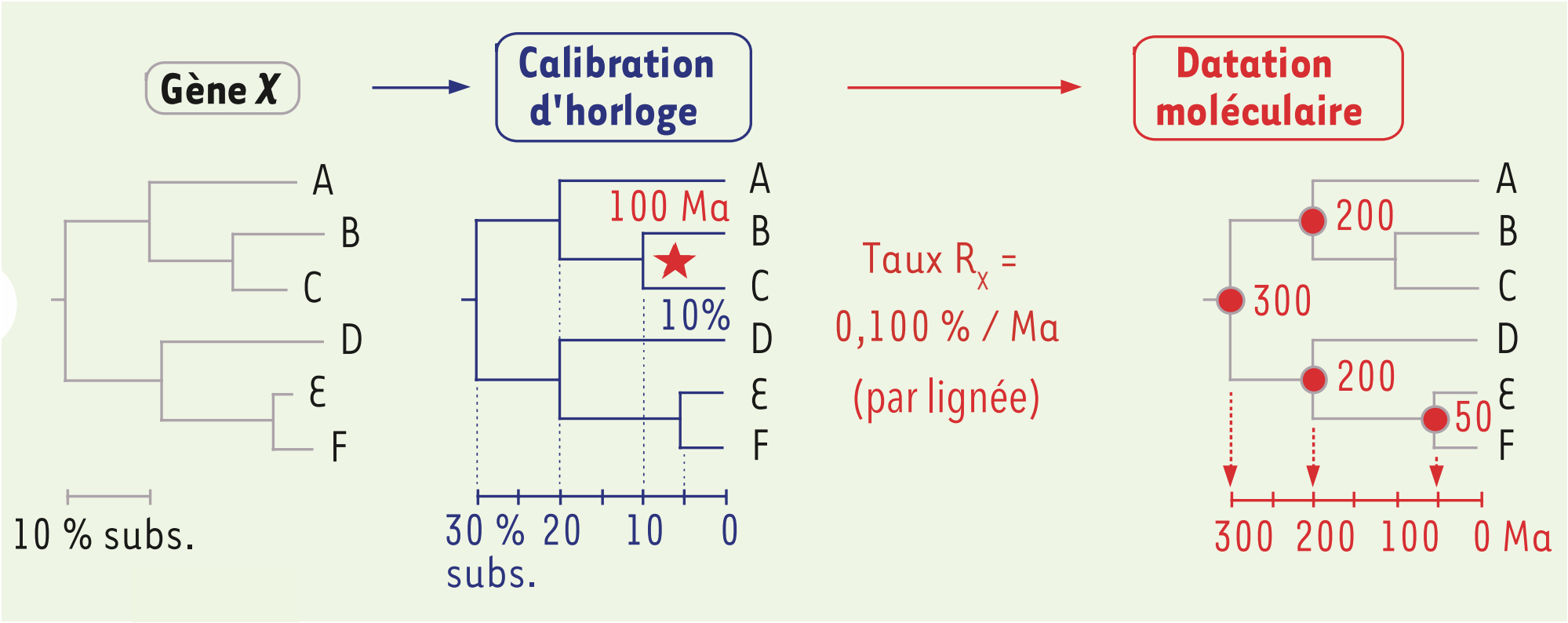
Lang et al. (1997) *Nature* 387:493-497 Rodriguez-Ezpeleta (2007) *Thèse de Doctorat*

4

Prédictions phylogénétiques
Horloges moléculaires

Horloges moléculaires

principe de base (vitesse d'évolution constante)

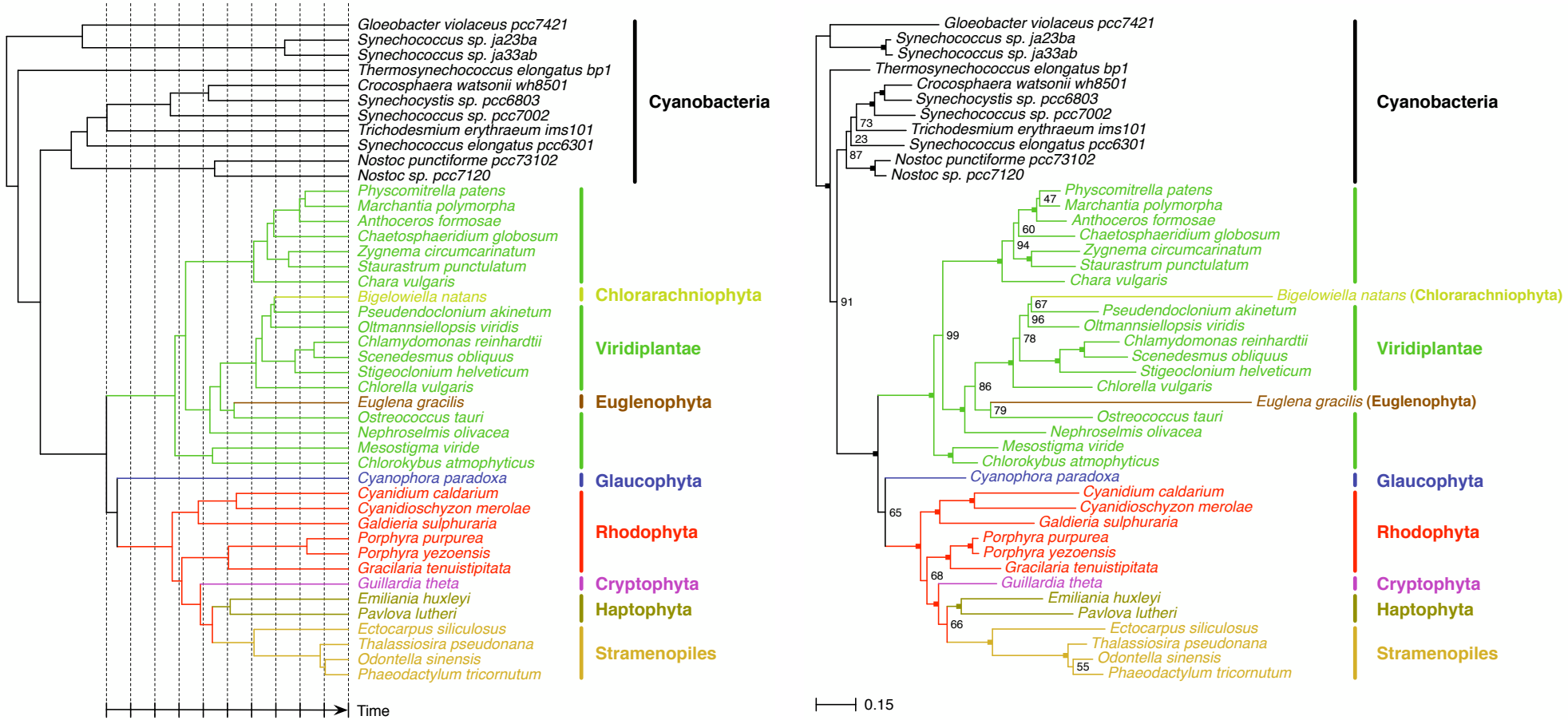


Si les différences de vitesses entre branches ne sont pas importantes, on peut contraindre l'arbre à être ultramétrique. Cela permet d'inférer la vitesse d'évolution à partir d'un seul point de calibration et de dater tous les autres noeuds en conséquence.

Douzery et al. (2006) *Med Sci* (Paris) 22:374-380

Horloges moléculaires

Hélas, les lignées n'évoluent pas à vitesse constante.



55 gènes x 44 espèces
PhyloBayes (CAT+Γ_4)

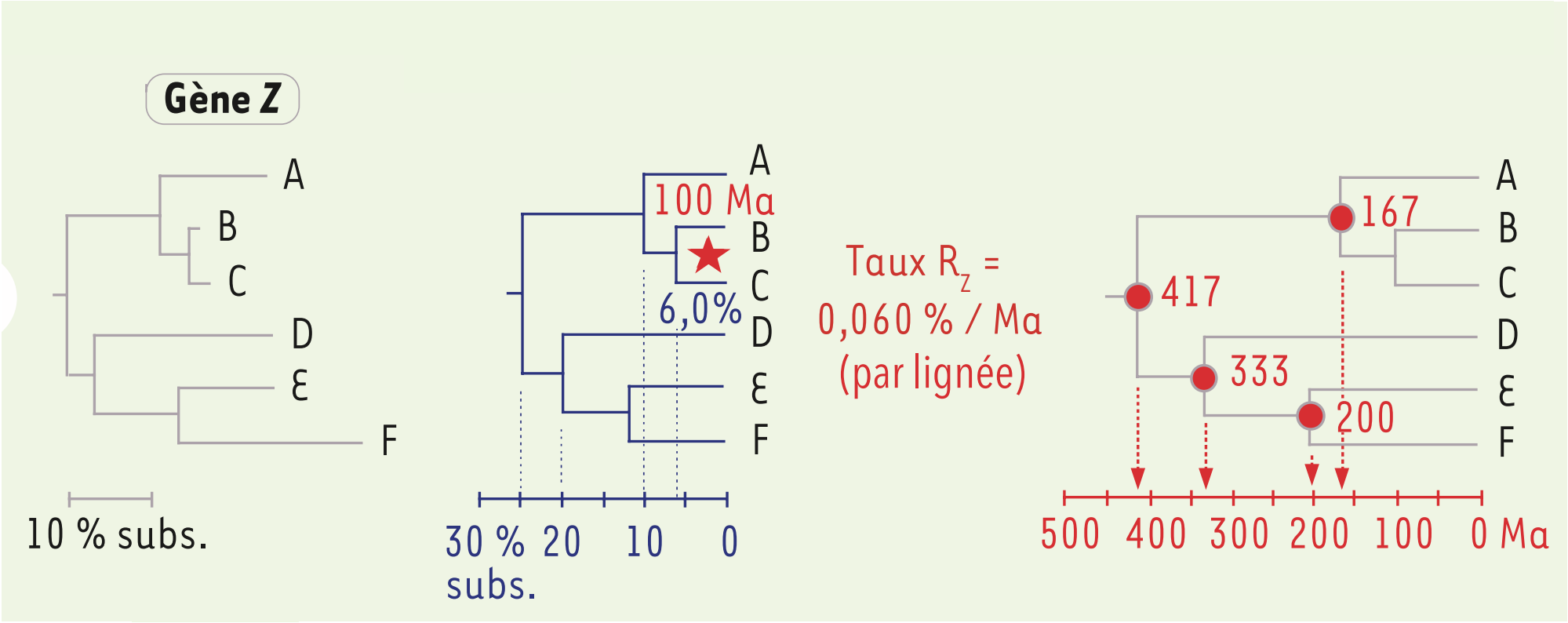
Baurain et al. (2010) *Mol Biol Evol* 27:1698-1709


distance = temps x vitesse

distance = temps x vitesse

Horloges moléculaires

impact de la vitesse d'évolution sur la datation



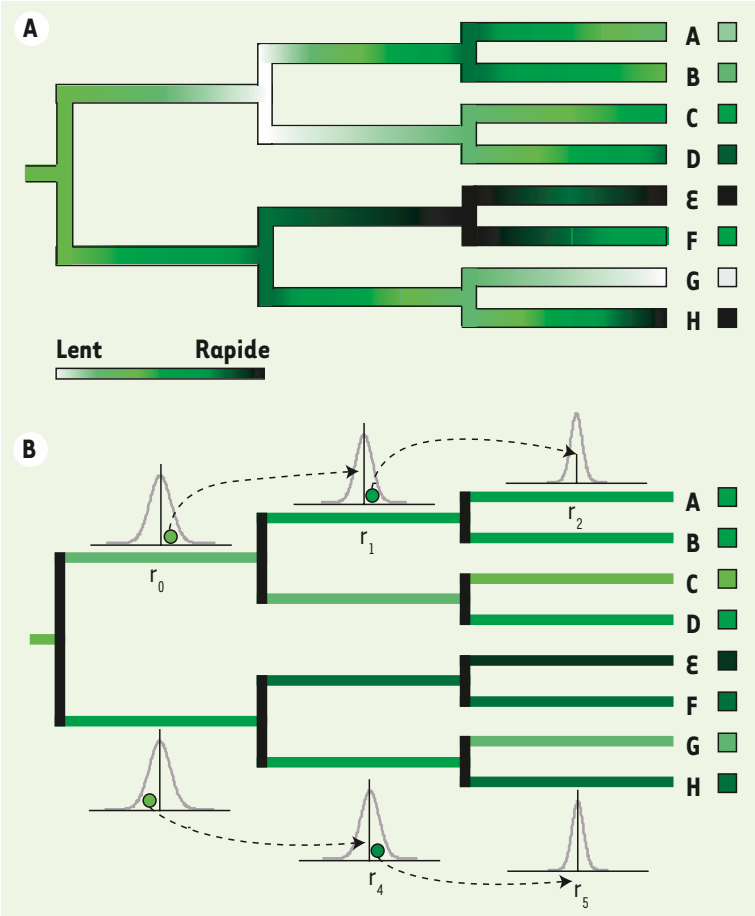
Si nous utilisons le gène Z, dont les vitesses d'évolution sont très différentes entre lignées (pas d'ultramétricité), les temps de divergence inférés seront incorrects puisque l'horloge n'est pas constante.

Douzery et al. (2006) *Med Sci* (Paris) 22:374-380

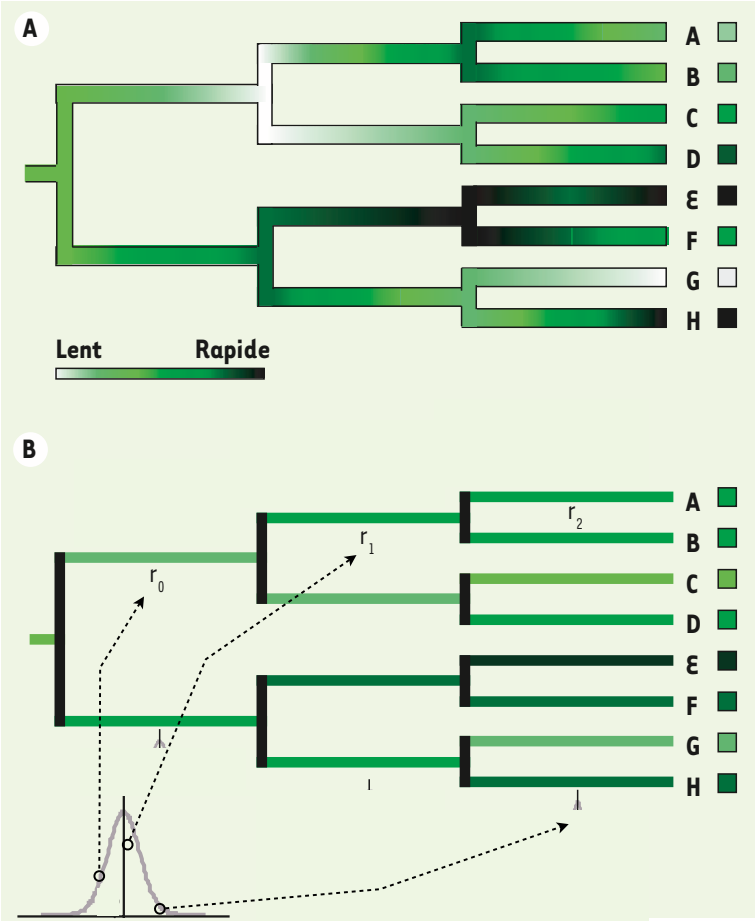
Horloges moléculaires

solution : les horloges moléculaires assouplies

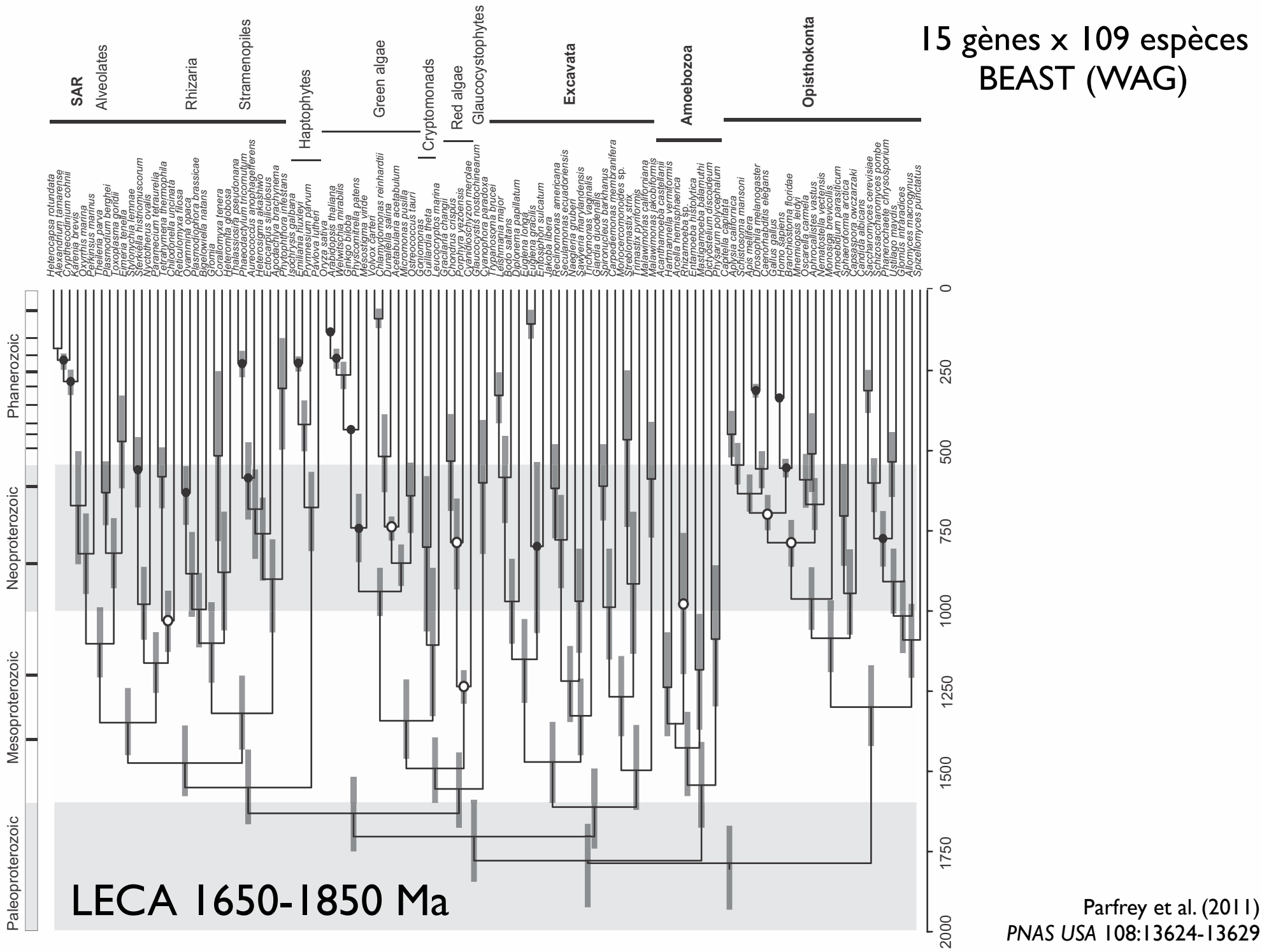
vitesses corrélées



vitesses non-corrélées



modifié de Douzery et al. (2006) *Med Sci* (Paris) 22:374-380



Parfrey et al. (2011)
PNAS USA 108:13624-13629

Taxon	Fossil	Eon*	Calibration†		Ref(s).
			Min	Dist	
Amniota	*Westlonthania*	Phan	328.3	4, 3	(54)
Angiosperms	Oldest angio pollen	Phan	133.9	2, 10	(55)
Ascomycetes	*Paleopyrenomycites*	Phan	400	4, 50	(56)
Coccolithophores	Earliest Heterococcolith	Phan	203.6	2, 8	(57)
Diatoms	Earliest diatoms	Phan	133.9	2, 100	(58)
Dinoflagellates	Earliest gonyaulacales	Phan	240	2, 10	(59)
Embryophytes	Land plant spores	Phan	471	2, 20	(60)
Endopterygota	Mecoptera	Phan	284.4	5, 5	(61)
Eudicots	Eudicot pollen	Phan	125	2, 1.5	(62, 63)
Euglenids	*Moyeria*	Phan	450	2, 40	(64)
Foraminifera	Oldest forams	Phan	542	2, 200	(65)
Gonyaulacales	Gonyaulacaceae split	Phan	196	2,10	(59)
Pennate diatoms	Oldest pennate	Phan	80	3, 5	(66)
Spirotrichs	Oldest tintinnids	Phan	444	2.5, 100	(67)
Trachaeophytes	Earliest trachaeophytes	Phan	425	4, 2.5	(68)
Vertebrates	*Haikouichthys*	Phan	520	3, 5	(69)
Animals	LOEMs, sponge biomarkers	Protero	632	2, 300	(70, 71)
Arcellinida	*Paleoarcella*	Protero	736	2, 300	(12)
Bilateria	*Kimberella*	Protero	555	2, 30	(72)
Chlorophytes	*Palaeastrum*	Protero	700	2.5, 300	(73)
Ciliates	Gammacerane	Protero	736	2.5, 300	(74)
Florideophyceae	Doushantuo red algae	Protero	550	2.5, 100	(75)
Red algae‡	*Bangiomorpha*	Protero	1174	3, 250	(11)

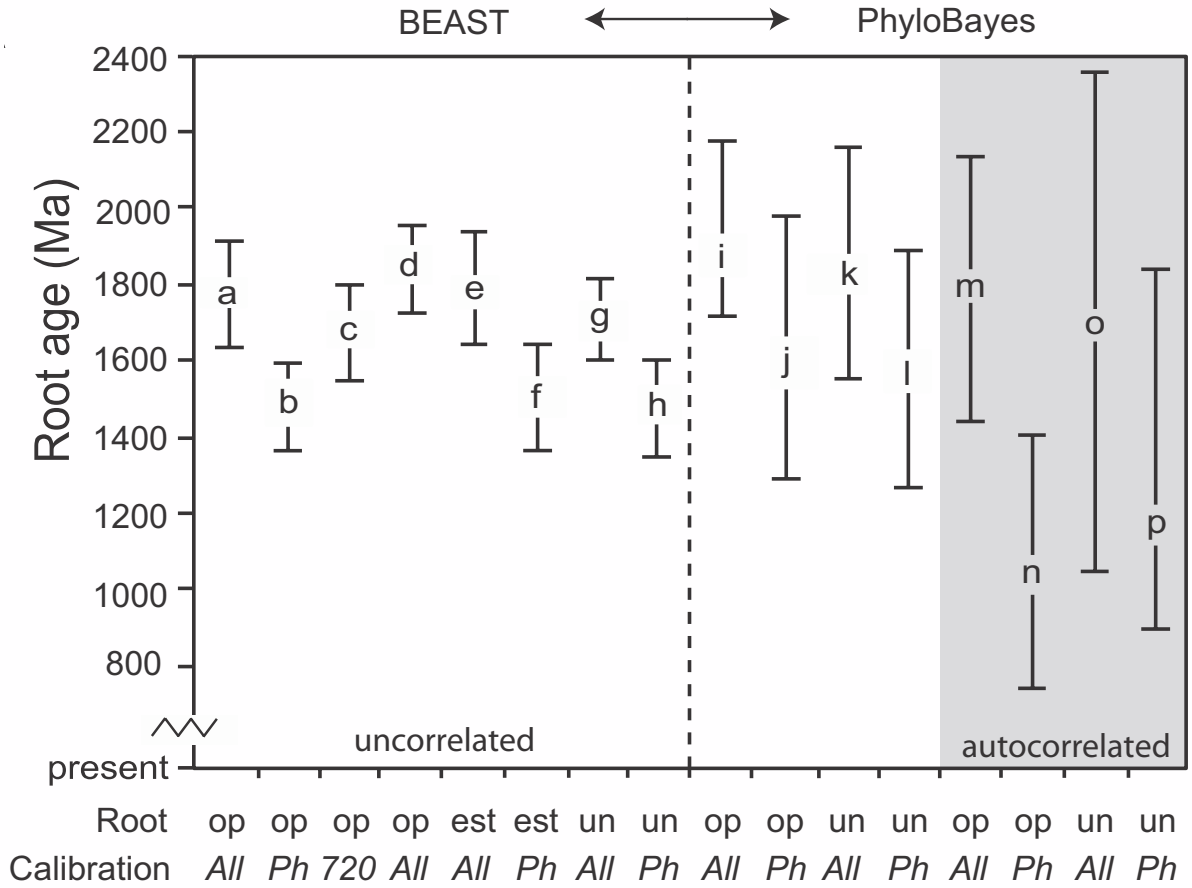
*Eon: Phan, Phanerozoic; Protero, Proterozoic. Proterozoic calibrations are excluded from *Phan* analyses.

†Calibration constraints are specified for BEAST using a gamma distribution with a minimum date in Ma based on the fossil record parameters as indicated: min, minimum divergence data; dist, gamma prior distribution (shape, scale). See Table S3 for details of PhyloBayes calibrations.

‡In the *All 720* analysis (*c*), the minimum age constraint for the red algae node is set to 720 Ma.

Horloges moléculaires

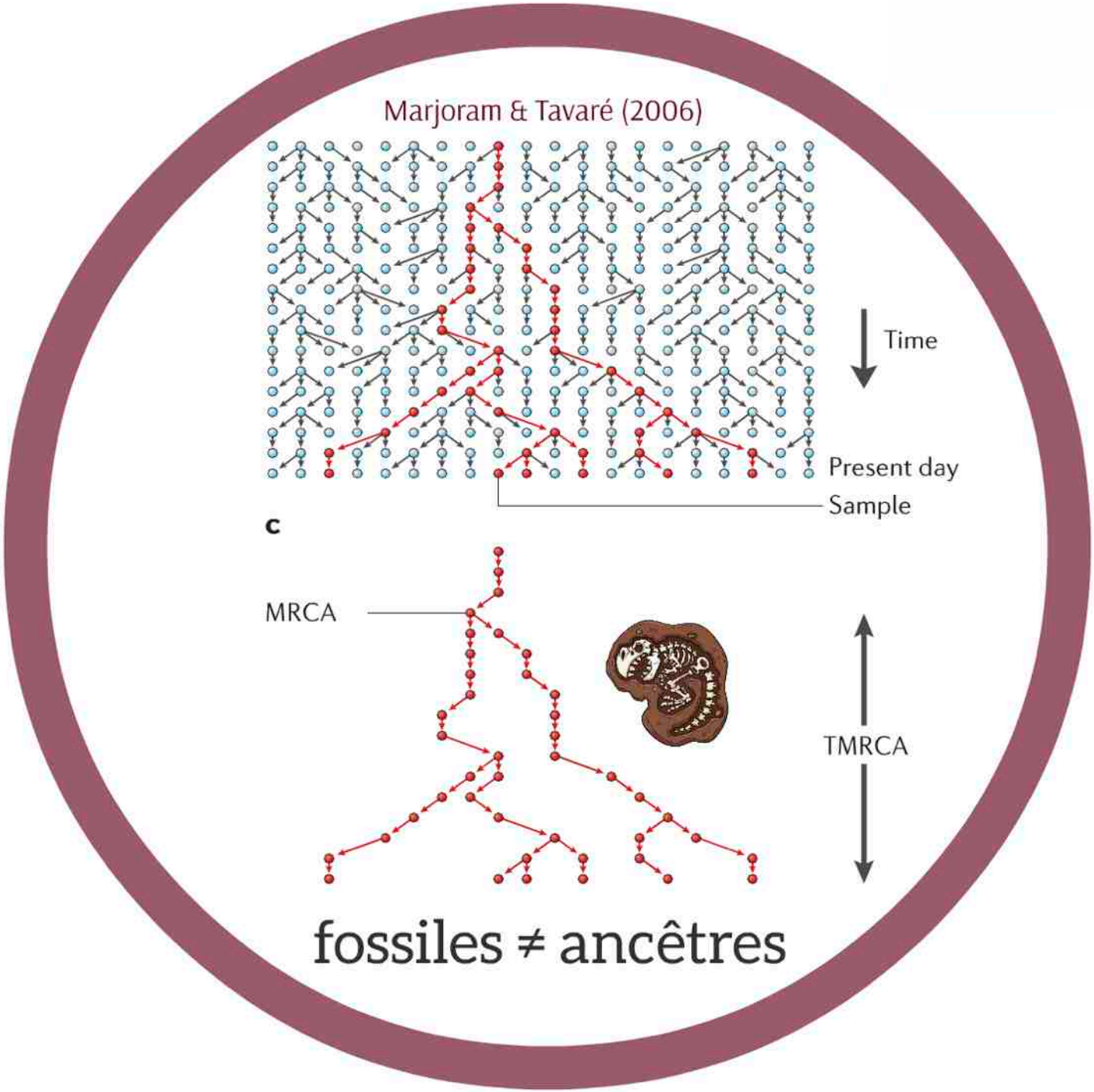
impact des choix méthodologiques sur la datation



LECA 1000-1900 Ma

Bangiomorpha

Parfrey et al. (2011) *PNAS USA* 108:13624-13629; Eme (2014) *Cold Spring Harb Perspect Biol* 6



fossiles ≠ ancêtres