

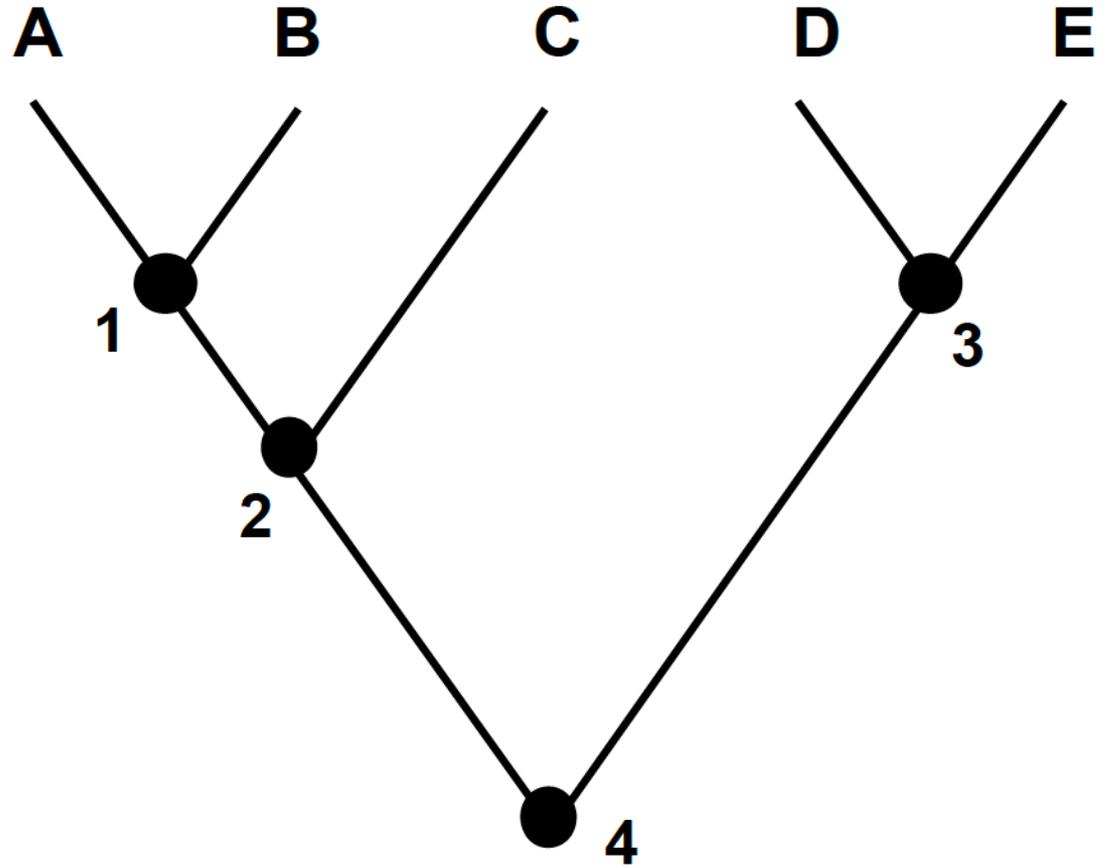
SV-K-2-2

Analyser des arbres phylogénétiques pour  
construire des scénarios évolutifs



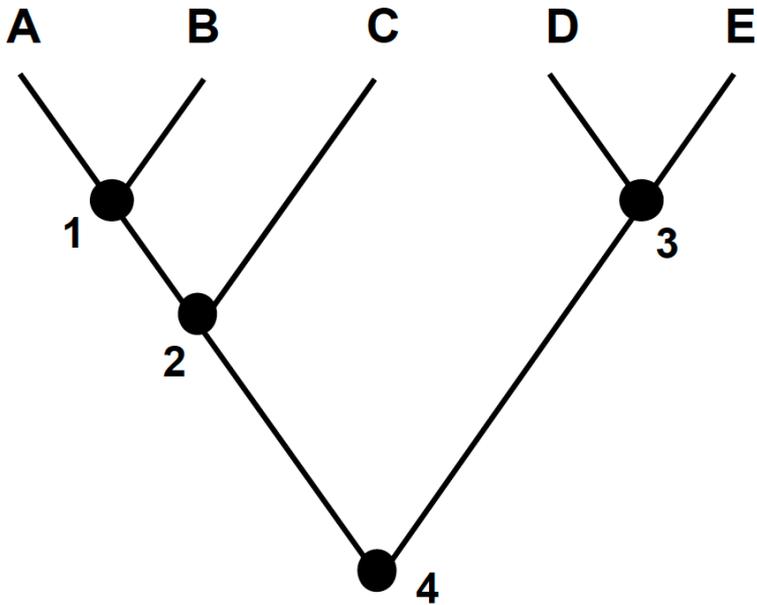
1. Soit  $X$  le groupe des taxons  $A$ ,  $B$  et  $C$ ,  $Y$  le groupe  $C$ ,  $D$  et  $E$  et  $Z$  le groupe constitué des taxons  $C$  et  $D$ . Qualifiez les groupes  $X$ ,  $Y$  et  $Z$ .

## Taxons :

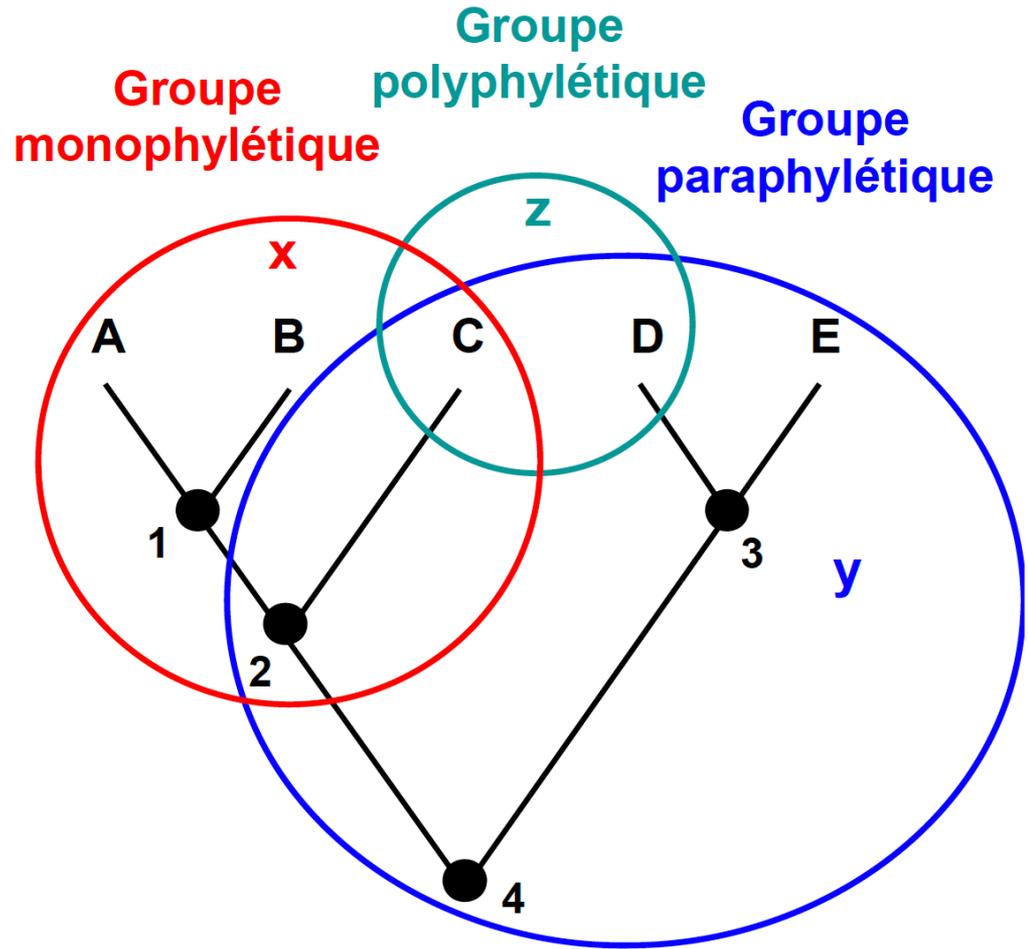


**Arbre fictif**

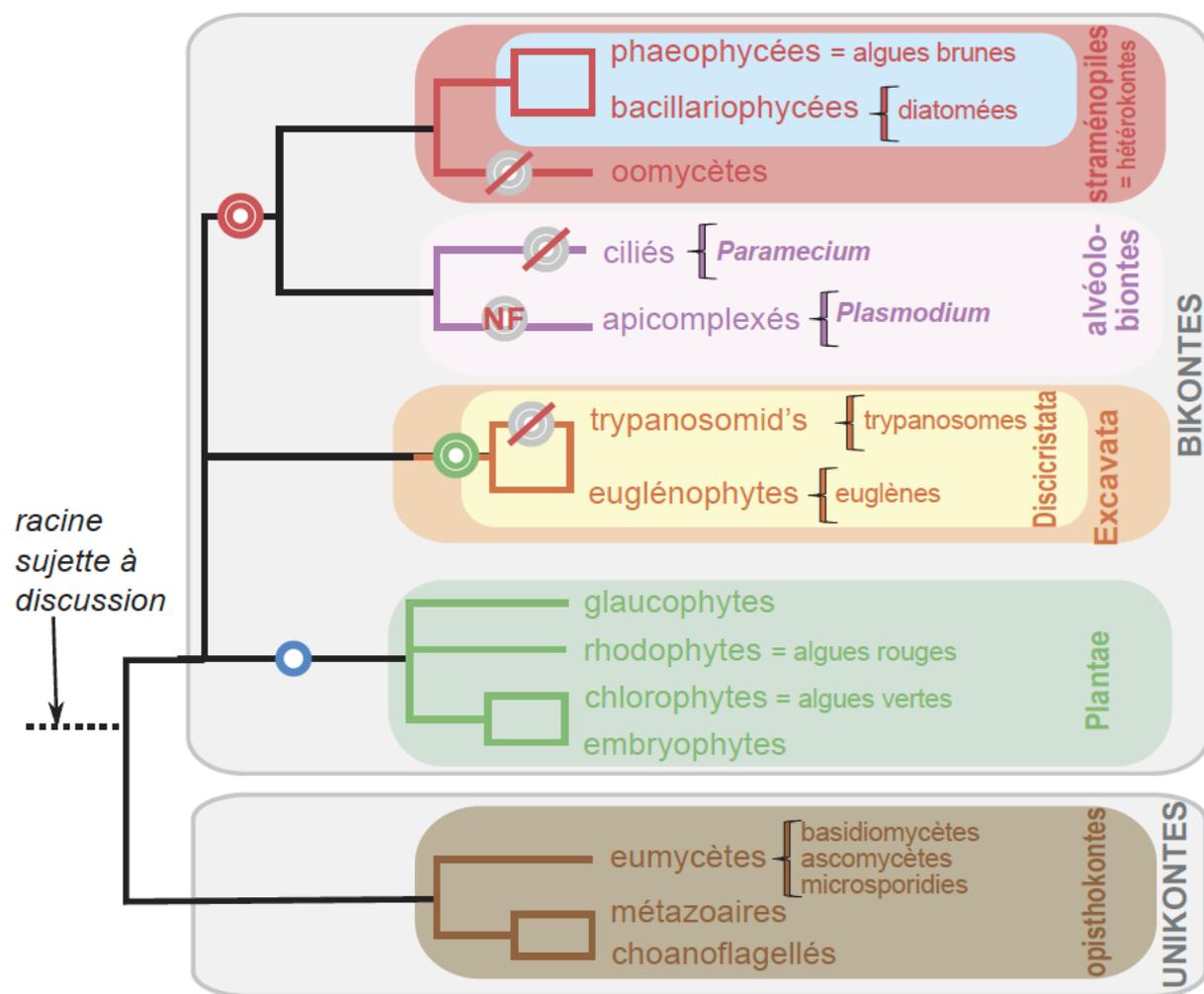
Taxons :



Arbre fictif



# Arbre simplifié des eucaryotes (modifié d'après Burki et coll, 2012, Pawlowski, 2013)



deux flagelles différents  
 autotrophie au carbone ou hétérotrophie  
 plastes à quatre membranes,  
 chlorophylles a et c,  
 β1-3 glucane (polyoside de réserve)

deux ou quatre flagelles, perdus chez  
 les rhodophytes et les angiospermes  
 autotrophie au carbone, plastes à deux  
 membranes, chlorophylles a et b,  
 amidon (polyoside de réserve)

un flagelle unique et propulseur  
 hétérotrophie au carbone  
 glycogène (polyoside de réserve)  
 chitine (polyoside de structure)

- endosymbiose primaire
- endosymbiose secondaire d'une algue verte
- endosymbiose secondaire d'une algue rouge

- ⊘ perte des plastes (sauf exceptions)
- ⊘ NF plastes non fonctionnels

Evolution réticulée

Régression et perte

Quelques apomorphies cellulaires caractéristiques de trois clades d'eucaryotes

Arbre des eucaryotes fondé sur l'analyse de 258 gènes concaténés. (modifié d'après Burki et coll., 2012, Pawlowski, 2013) in Tout-En-Un, Dunod, 2014

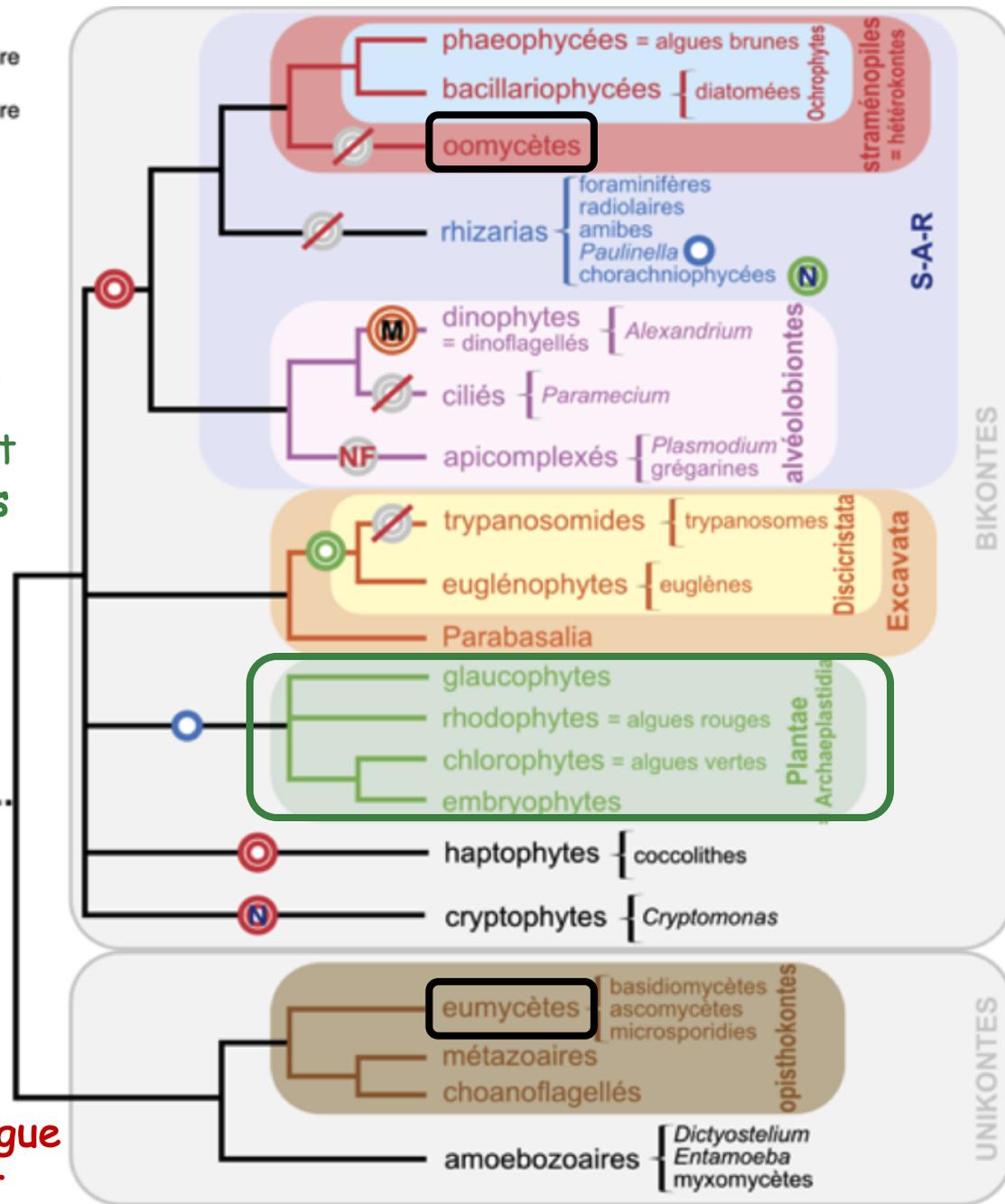
-  endosymbiose primaire
-  endosymbiose secondaire d'une algue verte
-  endosymbiose secondaire d'une algue rouge
-  endosymbiose tertiaire
- N** nucléomorphe
- M** événements multiples de pertes et gains
-  perte des plastes (sauf exceptions)
- NF** plastes non fonctionnels

Analyses moléculaires soutiennent hyp d'une homologie entre plastes des Plantae et CyanoB

3 arguments

- ✓ Monophylie de l'ensemble des Plantae
- ✓ Monophylie des plastes de Plantae
- ✓ Gènes plastidiaux des plantae = groupe frère avec les gènes de cyanobactéries

Analyses délicates à cause du dialogue symbiotique gènes nucl / gènes bact



## 2. Recherchez des arguments en faveur du modèle de l'endocytobiose

Image 1 : Ultrastructure d'une cyanobactérie, *Pseudanabaena*

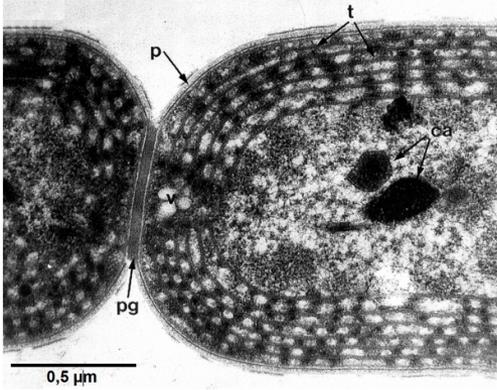


Image 3 : Ultrastructure d'une algue rouge, *Porphyridium rubrum*

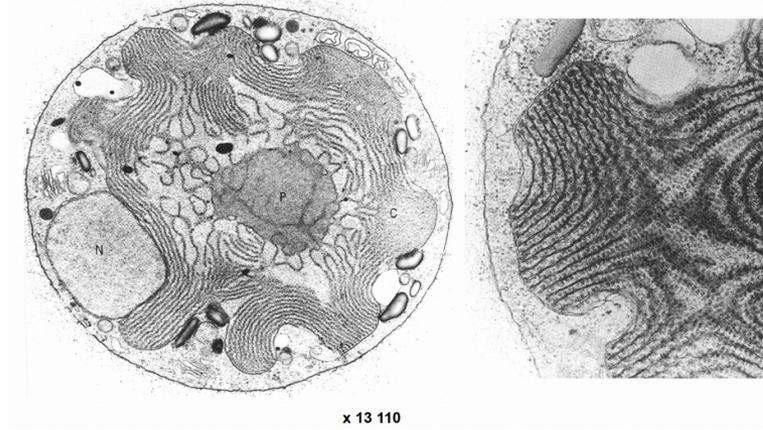


Image 5 : Espace périplastidial chez une algue brune

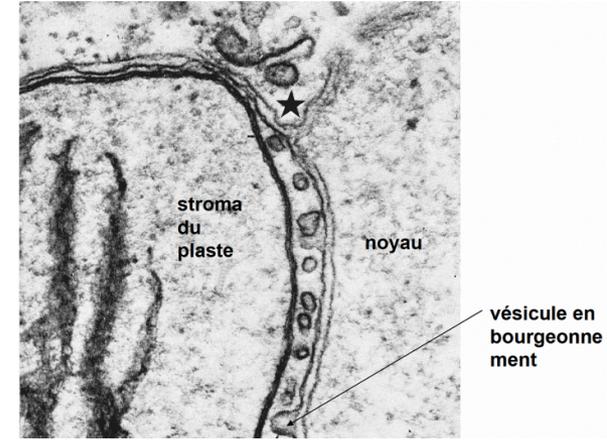


Image 2 : Une eubactérie se divisant dans le cytoplasme d'une archée

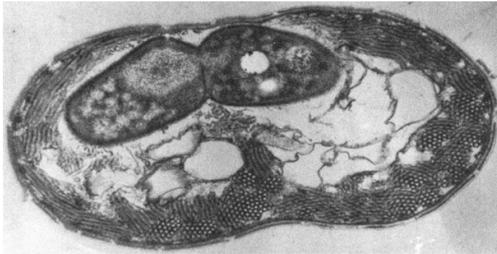
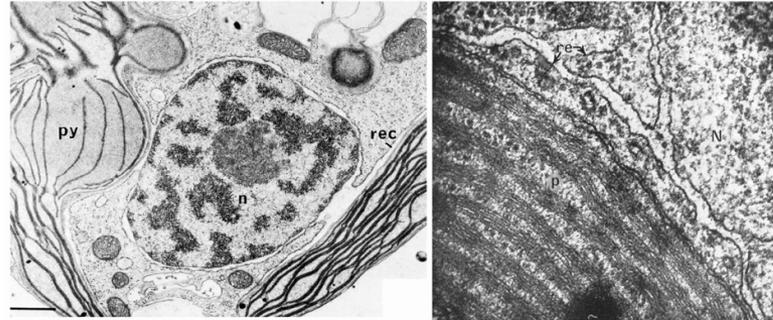


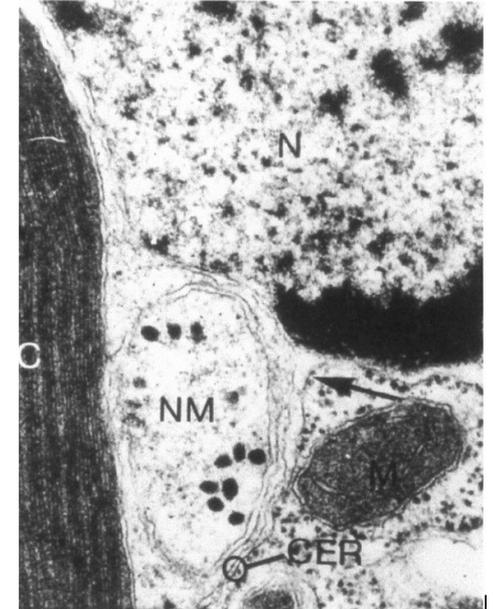
Image 4 : Ultrastructure de *Pleurochrysis pseudoroscoffensis* et *Elaschista fuciola*, 2 algues brunes



*Pleurochrysis pseudoroscoffensis*  
(MET, x 16650)

*Elaschista fuciola* (MET)

Image 6 : Plaste à nucléomorphe des Cryptophytes



## ***2. Recherchez des arguments en faveur du modèle de l'endocytobiose.***

Image 1 : Cyanobactéries compartimentées avec des thylakoïdes comme des plastes eucaryotiques.

Image 2 : Bactéries vivantes, car elles se divisent, incluses dans une archée.

Image 3 : Cellule eucaryotes munies d'organites compartimentés occupant tout le cytosol.

Image 4 : Continuité entre la membrane externe du noyau et le plaste (Elachista) et figure de bourgeonnement qui évoque une exocytose du plaste (Pleurochrysis).

Image 5 : Espace périplastidial avec une figure de bourgeonnement.

Image 6 : Présence d'un nucléomorphe évoquant un noyau vestigial

L'ensemble de ces images évoque une phagocytose de procaryotes au sein d'une cellule eucaryote, sans digestion, voire d'une endocytobiose secondaire d'un eucaryote à plaste par un autre eucaryote (cryptophyte de l'image 6)

### 3. Analysez le tableau de résultat de résultats ci-dessous et proposez une conclusion.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 <i>Saccharomyces cerevisiae</i>	-												
2 <i>Lemna minor</i>	0,29	-											
3 Lymphocyte B humain	0,33	0,36	-										
4 <i>Escherichia coli</i>	0,05	0,10	0,06	-									
5 <i>C. vibrioforme</i>	0,06	0,05	0,06	0,24	-								
6 <i>Bacillus firmis</i>	0,08	0,06	0,07	0,25	0,22	-							
7 <i>C. diphteriae</i>	0,09	0,10	0,07	0,28	0,22	0,34	-						
8 <i>Aphanocapsa</i>	0,11	0,09	0,09	0,26	0,20	0,26	0,23	-					
9 Chloroplaste ( <i>L. minor</i> )	0,08	0,11	0,06	0,21	0,19	0,20	0,21	0,31	-				
10 <i>M. thermoautotrophicum</i>	0,11	0,10	0,10	0,11	0,06	0,11	0,12	0,11	0,14	-			
11 <i>M. ruminantium</i>	0,11	0,10	0,10	0,12	0,07	0,13	0,12	0,11	0,12	0,51	-		
12 <i>Methanobacter sp.</i>	0,08	0,13	0,10	0,07	0,06	0,06	0,09	0,10	0,10	0,25	0,25	-	
13 <i>M. barkeri</i>	0,08	0,07	0,07	0,12	0,09	0,12	0,10	0,10	0,12	0,30	0,24	0,32	-

Cellules	Taux d'homologie entre séquences d'ARNr de chloroplastes de <i>Lemna minor</i> (Eucaryote) et diverses cellules	
<i>Saccharomyces cerevisiae</i>	0,08	« Faible » taux d'homologie
<i>Lemna minor</i>	0,11	
Lymphocyte B humain	0,06	
<i>Escherichia coli</i>	0,21	« Fort » taux d'homologie
<i>C. vibrioforme</i>	0,19	
<i>Bacillus firmis</i>	0,20	
<i>C. diphteriae</i>	0,21	
<i>Aphanocapsa</i>	0,31	
<i>M. Thermoautotrophicum</i>	0,14	Taux d'homologie « intermédiaire »
<i>M. ruminantium</i>	0,12	
<i>Methanobacter sp.</i>	0,10	
<i>M. barkeri</i>	0,12	

3. Analysez le tableau de résultat de résultats ci-dessous et proposez une conclusion.

Une réécriture du tableau permet de dégager 3 groupes. Le taux d'homologies des ARNr élevé entre bactéries et eucaryotes, particulièrement les cyanobactéries, permet de postuler que les chloroplastes eucaryotes présentent davantage de parenté avec les cyanobactéries qu'avec les autres types cellulaires.

Cellules	Taux d'homologie entre séquences d'ARNr de chloroplastes de <i>Lemna minor</i> (Eucaryote) et diverses cellules	
<i>Saccharomyces cerevisiae</i>	0,08	« Faible » taux d'homologie
<i>Lemna minor</i>	0,11	
<i>Lymphocyte B humain</i>	0,06	
<i>Escherichia coli</i>	0,21	« Fort » taux d'homologie
<i>C. vibrioforme</i>	0,19	
<i>Bacillus firmis</i>	0,20	
<i>C. diphtheriae</i>	0,21	
<i>Aphanocapsa</i>	0,31	
<i>M. Thermoautotrophicum</i>	0,14	Taux d'homologie « intermédiaire »
<i>M. ruminantium</i>	0,12	
<i>Methanobacter sp.</i>	0,10	
<i>M. barkeri</i>	0,12	

4. À partir de ces séquences, remplissez les cases blanches de la matrice de distance ci-dessous dont on rappelle le principe de construction.

1														
C	T	G	A	C	G	G	A	G	G	A	A	T	G	C
-	-	-	-	-	-	-	-	-	C	-	-	-	A	-
-	-	-	-	-	-	-	-	-	-	-	-	-	A	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16														
C	G	C	G	T	G	G	A	G	G	T	A	*	G	A
T	-	G	-	-	-	*	-	-	C	G	T	G	-	-
-	-	-	-	-	-	*	-	-	-	G	T	A	-	-
-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31														
A	G	G	C	C	C	A	C	G	G	G	T	C	C	T
-	-	-	-	T	-	T	T	C	C	-	-	T	G	G
C	-	-	-	T	T	G	T	-	-	-	-	T	G	-
-	-	-	-	T	-	G	-	-	-	-	-	T	G	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
46														
G	A	A	C	T	T	C	T	T	T	T	C	C	C	A
C	-	-	-	C	C	-	-	-	-	-	T	T	-	-
-	-	-	-	C	-	-	-	-	-	-	-	-	T	-
A	-	-	-	C	-	-	-	-	-	-	-	T	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
61														
G	A	G	A	A	G	A	A	G	*	*	*	C	A	A
-	G	-	-	G	T	-	T	A	C	C	A	A	-	-
-	G	-	-	G	-	-	-	T	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
76														
T	G	A	C	G	G	T	A	T	C	T	G	G	G	G
-	-	-	A	-	-	-	-	C	T	-	-	A	A	C
-	-	-	-	-	-	-	-	C	-	-	A	A	-	-
-	-	-	-	-	-	-	-	-	-	-	A	A	A	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
91									100					
A	A	T	A	A	G	C	A	T	C					
-	-	-	-	G	-	-	-	-	-					<i>Arabidopsis thaliana</i> (Chlorobionte)
-	-	-	-	-	-	-	-	-	-					<i>Anabaena</i> sp. (Cyanobactérie)
-	-	-	-	-	-	-	-	-	-					<i>Cyanophora paradoxa</i> (Glaucophyte)
-	-	-	-	-	-	-	-	-	-					<i>Porphyra purpurea</i> (Rhodophyte)



5. Utilisez cette matrice pour construire un arbre phylogénétique des plastes, en admettant que (1) Les séquences évoluent à la même vitesse dans toutes les branches de l'arbre et (2) que les pourcentages calculés reflètent les différences pour la séquence totale de l'ARN 16S. Quels enseignements en tirez-vous ? On rappelle ci-dessous le principe de construction de l'arbre phylogénétique.

La distance la plus faible est celle existant entre Porphyra et Arabidopsis d'où leur installation sur deux branches partant d'un même ancêtre commun et partageant la même distance totale les séparant soit  $12/2 = 6$ . La poursuite de la méthode permet de considérer cette branche comme un nouvel ensemble U1 qui permet de réécrire la matrice précédente en réévaluant une distance entre les représentants restants et le nouvel ensemble U1.

	Arabidopsis	Anabaena	Cyanophora	Porphyra
Arabidopsis	0	-	-	-
Anabaena	38	0	-	-
Cyanophora	20	29	0	-
Porphyra	12	31	17	0

Cette étape permet d'identifier la parenté plus importante de U1 avec Cyanophora et de constituer alors un nouvel ensemble U2 regroupant Cyanophora et U1, appuyé sur un ancêtre commun et se partageant la distance totale les séparant soit  $18,5/2 = 9,25$ .

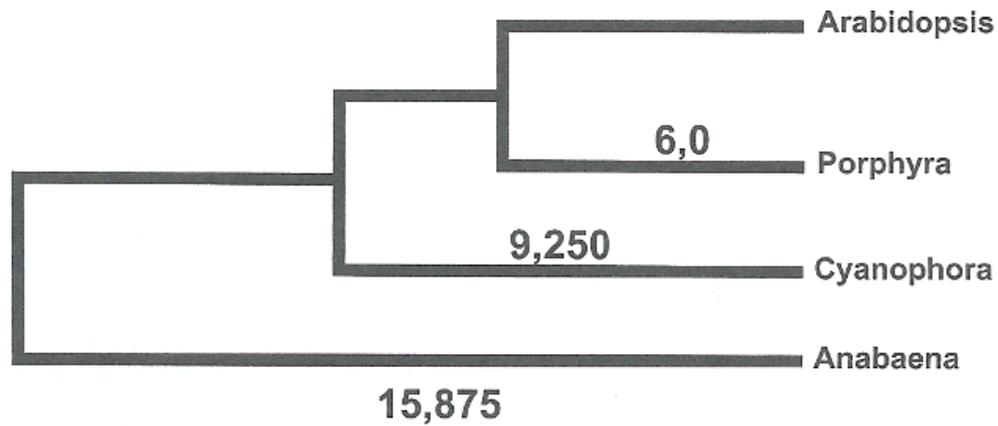
	U1	Anabaena	Cyanophora
U1	0,0	-	-
Anabaena	34,5	0,0	-
Cyanophora	18,5	29,0	0,0

	U2	Anabaena
U2	0,00	-
Anabaena	31,75	0,00

Cette étape permet d'identifier la parenté plus importante de U1 avec Cyanophora et de constituer alors un nouvel ensemble U2 regroupant Cyanophora et U1, appuyé sur un ancêtre commun et se partageant la distance totale les séparant soit  $18,5/2 = 9,25$ . il est alors possible de réécrire la matrice comme suit :

	U2	Anabaena
U2	0,00	–
Anabaena	31,75	0,00

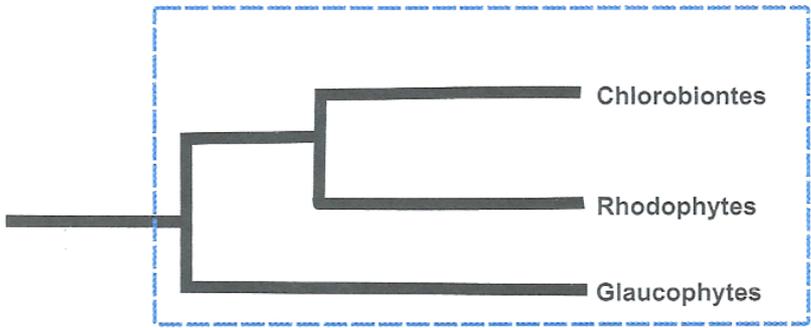
Anabaena et U2 peuvent donc être arrimés un ancêtre commun situé à une distance de l'un Anabaena et des autres membres formant U2 soit  $31,75/2 = 18,875$



6. En mettant en relation ces données issues de l'analyse des séquences nucléaires avec l'arbre que vous avez construit à partir des séquences d'ARNr 16S, déduire le caractère monophylétique ou polyphylétique des plastes de la lignée verte.

L'énoncé indique que les analyses des séquences moléculaires d'ADN et des protéines nucléaires conforte l'hypothèse d'une origine monophylétique de la lignée verte rassemblant chlorobiontes, Rhodophytes et Glaucophytes. Il est précisé également que les glaucophytes constituent le premier groupe à diverger à partir de l'ancêtre commun à ces 3 taxons. Le schéma ci-dessous est l'arbre de parenté correspondant.

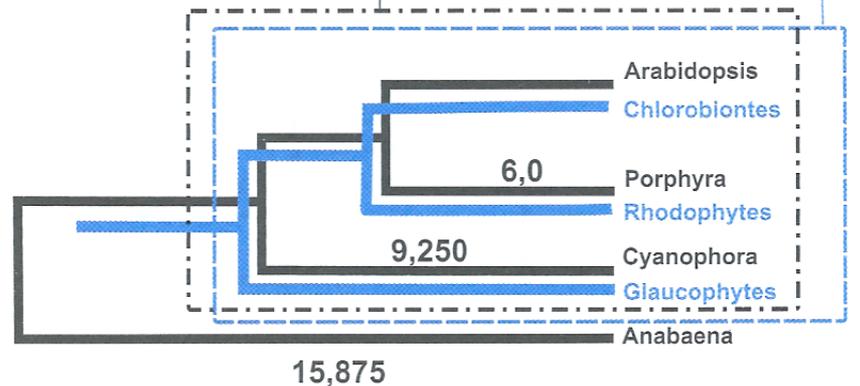
Caractère monophylétique de la lignée verte établie sur séquences protéiques et nucléiques nucléaires



Tous les représentants de la lignée verte sont positionnés sur des rameaux ayant une seule racine commune qu'ils ne partagent avec aucun autre représentant n'appartenant pas à cette lignée

Caractère monophylétique des plastes de la lignée verte

Caractère monophylétique de la lignée verte établie sur séquences protéiques et nucléiques nucléaires



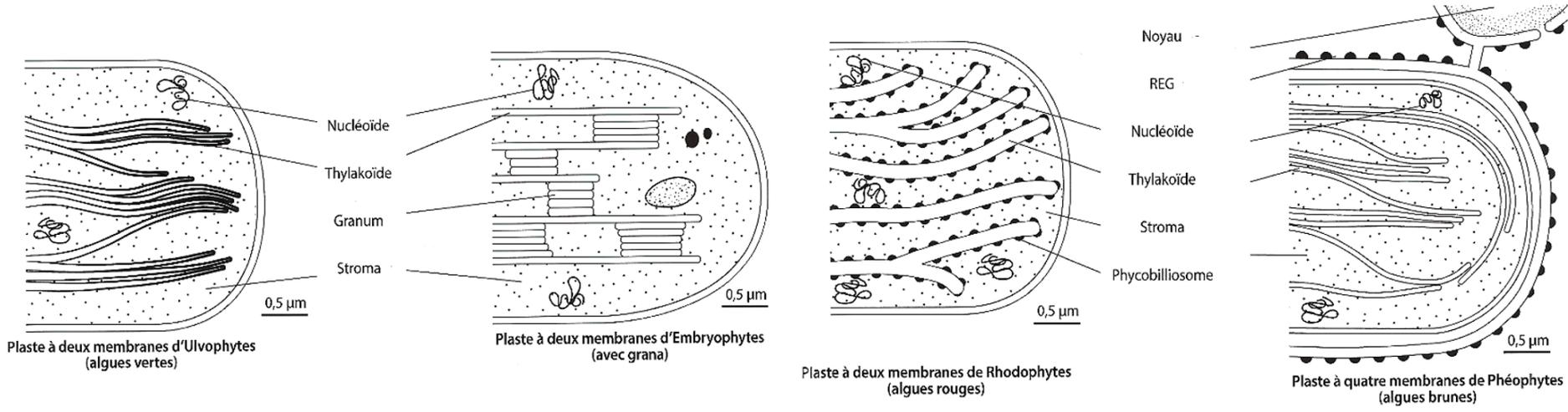
# II. ÉTABLISSEMENT D'UN CLADOGRAMME À PARTIR DES PLASTES EUCARYOTIQUES.

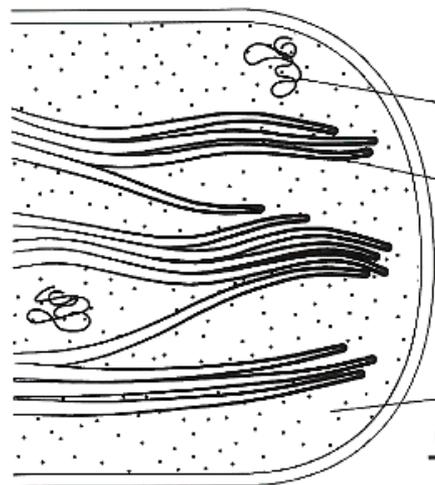
7. Proposez une méthodologie qui permette d'y parvenir en justifiant vos choix, en explicitant votre démarche et en argumentant la construction de votre représentation.

Comparaisons des principales caractéristiques des plastes primaires et secondaires eucaryotes (modifié d'après Archibald, 2008).

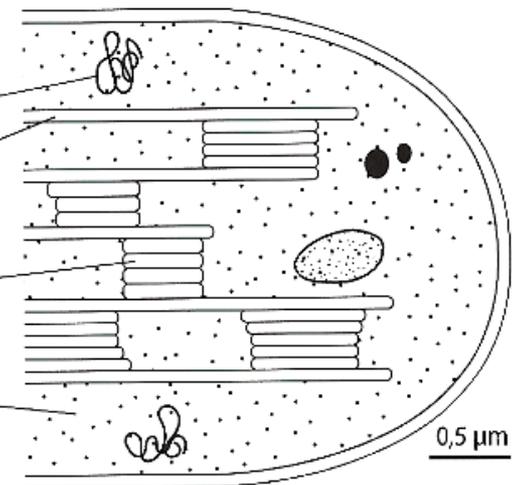
Phylum	Membranes	Thylacoïdes	Chlorophylle	Phycobilines	Caroténoïdes	Localisation
<b>cyanobactéries</b>	-	non accolés	a	PC, APC, PE	-	-
<b>Eucaryotes</b>						
<b>phéophycées</b>	4	Par 3	a et c	-	fucoxanthine	lumière du RER
<b>glaucophytes</b>	2	non accolés	a	PC, APC	-	cytoplasme
<b>rhodophytes</b>	2	non accolés	a (et d)	PC, APC, PE	-	cytoplasme
<b>chlorophytes</b>	2	par 2 à 6 (grana)	a et b	-	zéaxanthine	cytoplasme
<b>embryophytes</b>	2	associés, grana	a et b	-	zéaxanthine	cytoplasme

(APC = Allophycocyanine, PC = Phycocyanine et PE = Phycoérythrine).

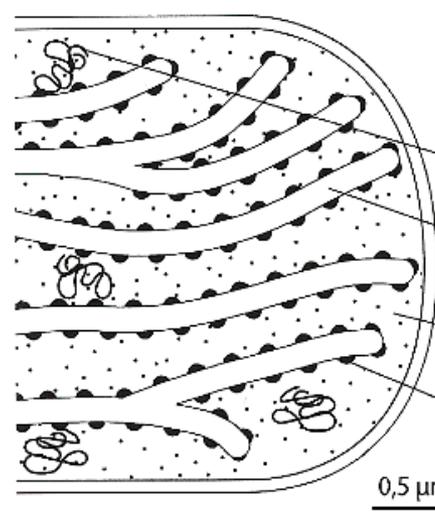




Plaste à deux membranes d'Ulvophytes  
(algues vertes)



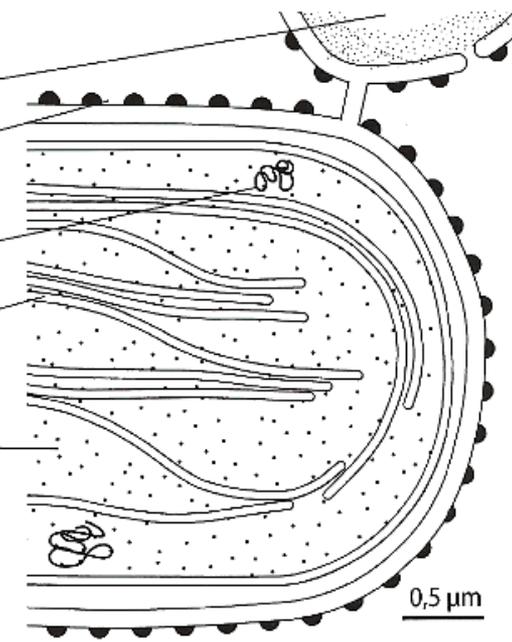
Plaste à deux membranes d'Embryophytes  
(avec grana)



Plaste à deux membranes de Rhodophytes  
(algues rouges)

- Nucléotide
- Thylakoïde
- Granum
- Stroma

- Noyau
- REG
- Nucléotide
- Thylakoïde
- Stroma
- Phycobilliosome



Plaste à quatre membranes de Phéophytes  
(algues brunes)

La construction du cladogramme est basée sur l'étude de **caractères homologues** dont on compare les états, primitifs ou dérivés.

La phylogénie est basée sur le partage de caractères à l'état dérivé (**synapomorphies**) permettant de définir un groupe monophylétique, seule entité à valeur taxonomique.

Les caractères « membrane et localisation » indiquent l'absence de plaste chez les cyanobactéries chez lesquelles le caractère « caroténoïdes » est absent.

Il faut polariser l'arbre en repérant les états primitifs et en déterminant un extragroupe, ensemble d'individus présentant tous les caractères considérés à l'état primitif (**symplesiomorphie**). Nous considérerons les cyanobactéries comme extragroupe.

Réécrivons le tableau en affectant la valeur 0 aux états primitifs et 1 aux états dérivés

Phylum	Membranes	Thylacoïdes	Chlorophylle	Phycobilines	Caroténoïdes	Localisation
<b>cyanobactéries</b>	0	0	0	0	0	0
<b>Eucaryotes</b>						
<b>phéophycées</b>	4	Par 3	a et c	-	fucoxanthine	lumière du RER
<b>glaucophytes</b>	2	0	0	PC, APC	0	cytoplasme
<b>rhodophytes</b>	2	0	a (et d)	0	0	cytoplasme
<b>chlorophytes</b>	2	par 2 à 6 (grana)	a et b	-	zéaxanthine	cytoplasme
<b>embryophytes</b>	2	groupés, grana	a et b	-	zéaxanthine	cytoplasme

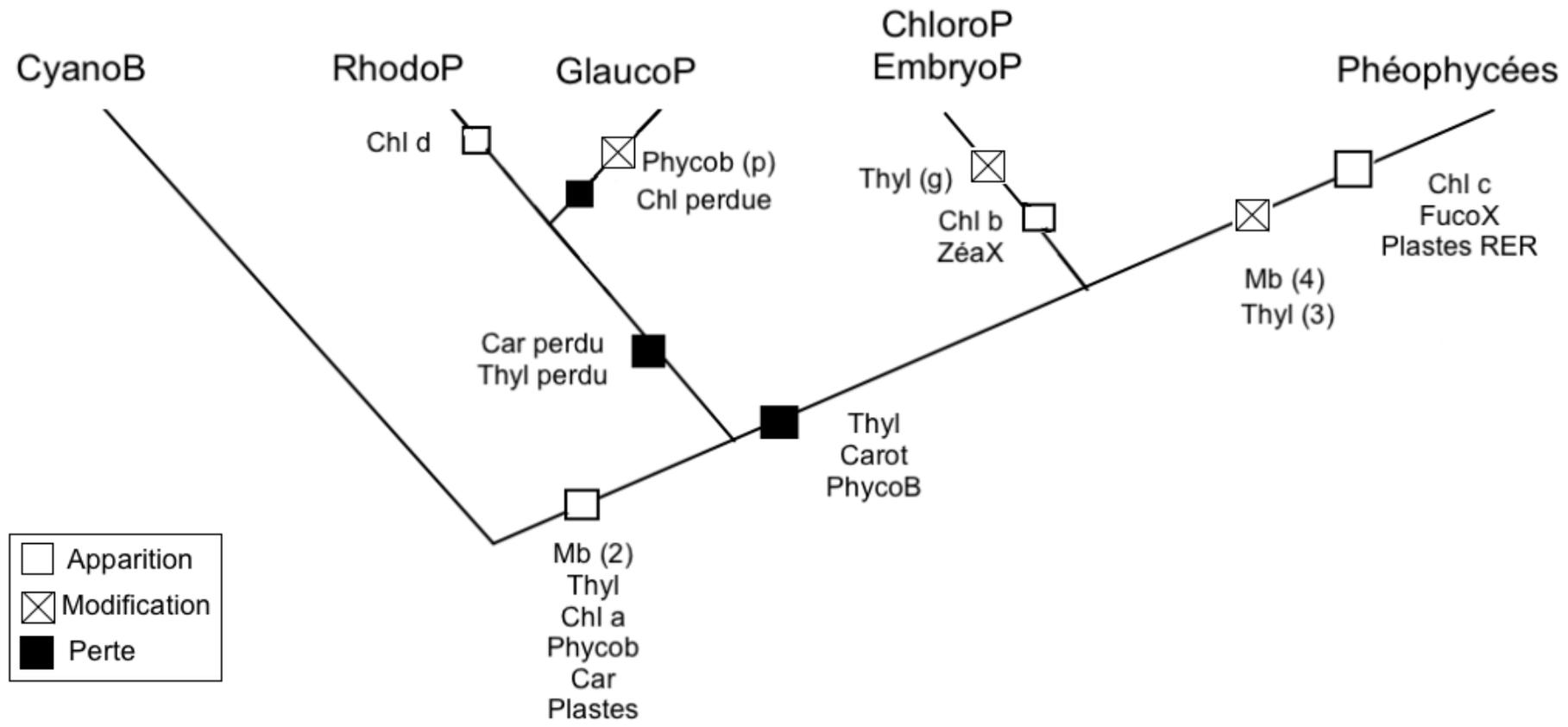
Pour un caractère donné, il peut y avoir plusieurs états dérivés. En affectant 1 à chacun d'entre eux, sans les distinguer, nous perdons de l'information. Ajoutons un chiffre ou une lettre en indice pour différencier chacun des états dérivés.

Phylum	Membranes	Thylacoïdes	Chlorophylle	Phycobilines	Caroténoïdes	Localisation
<b>cyanobactéries</b>	0	0	0	0	0	0
<b>Eucaryotes</b>						
<b>phéophycées</b>	1 <sub>(4)</sub>	1 <sub>(3)</sub>	1 <sub>(c)</sub>	1	1 <sub>(f)</sub>	1 <sub>(r)</sub>
<b>glaucophytes</b>	1 <sub>(2)</sub>	0	0	1 <sub>(p)</sub>	0	1 <sub>(cy)</sub>
<b>rhodophytes</b>	1 <sub>(2)</sub>	0	1 <sub>(d)</sub>	0	0	1 <sub>(cy)</sub>
<b>chlorophytes</b>	1 <sub>(2)</sub>	1 <sub>(g)</sub>	1 <sub>(b)</sub>	1	1 <sub>(z)</sub>	1 <sub>(cy)</sub>
<b>embryophytes</b>	1 <sub>(2)</sub>	1 <sub>(g)</sub>	1 <sub>(b)</sub>	1	1 <sub>(z)</sub>	1 <sub>(cy)</sub>

1<sub>(c)</sub> : Chlorophylle c ; 1<sub>(d)</sub> : chlorophylle d ; 1<sub>(p)</sub> : perte partielle ; 1<sub>(z)</sub> : zéaxanthine  
1<sub>(f)</sub> : fucoxanthine ; 1<sub>(r)</sub> : RER ; 1<sub>(cy)</sub> : cytoplasme ; g : groupés ;

Phylum	Membranes	Thylacoïdes	Chlorophylle	Phycobilines	Caroténoïdes	Localisation
<b>cyanobactéries</b>	0	0	0	0	0	0
<b>Eucaryotes</b>						
<b>phéophycées</b>	1 <sub>(4)</sub>	1 <sub>(3)</sub>	1 <sub>(c)</sub>	1	1 <sub>(f)</sub>	1 <sub>(r)</sub>
<b>glaucophytes</b>	1 <sub>(2)</sub>	0	0	1 <sub>(p)</sub>	0	1 <sub>(cy)</sub>
<b>rhodophytes</b>	1 <sub>(2)</sub>	0	1 <sub>(d)</sub>	0	0	1 <sub>(cy)</sub>
<b>chlorophytes</b>	1 <sub>(2)</sub>	1 <sub>(g)</sub>	1 <sub>(b)</sub>	1	1 <sub>(z)</sub>	1 <sub>(cy)</sub>
<b>embryophytes</b>	1 <sub>(2)</sub>	1 <sub>(g)</sub>	1 <sub>(b)</sub>	1	1 <sub>(z)</sub>	1 <sub>(cy)</sub>

1<sub>(c)</sub> : Chlorophylle c ; 1<sub>(d)</sub> : chlorophylle d ; 1<sub>(p)</sub> : perte partielle ; 1<sub>(z)</sub> : zéaxanthine ; 1<sub>(f)</sub> : fucoxanthine ;  
 1<sub>(r)</sub> : RER ; 1<sub>(cy)</sub> : cytoplasme ; g : groupés ;



### III. IDENTIFIER UN TRANSFERT HORIZONTAL DE GÈNES (THG)

#### III.1. Alignement de séquences protéiques et comparaison des arbres phylogénétiques chez les bactéries

La phylogénie de 10 bactéries appartenant aux genres *Salmonella*, *Escherichia*, *Yersinia*, *Buchnera*, *Wigglesworthia*, *Haemophilus*, *Pasteurella*, *Vibrio* et *Pseudomonas* est estimé à l'aide de séquences d'acides aminés de 4 protéines déduites des génomes la chaîne A de l'ATP synthétase ((ATPB), la biotine synthétase (BioB), la sous-unité A de l'ADN gyrase (GYRA) et la chaîne alpha de l'ARN polymérase ADN dépendante (RPOA). Trois groupes externes appartenant aux genres *Xanthomonas* et *Xylella* sont utilisés.

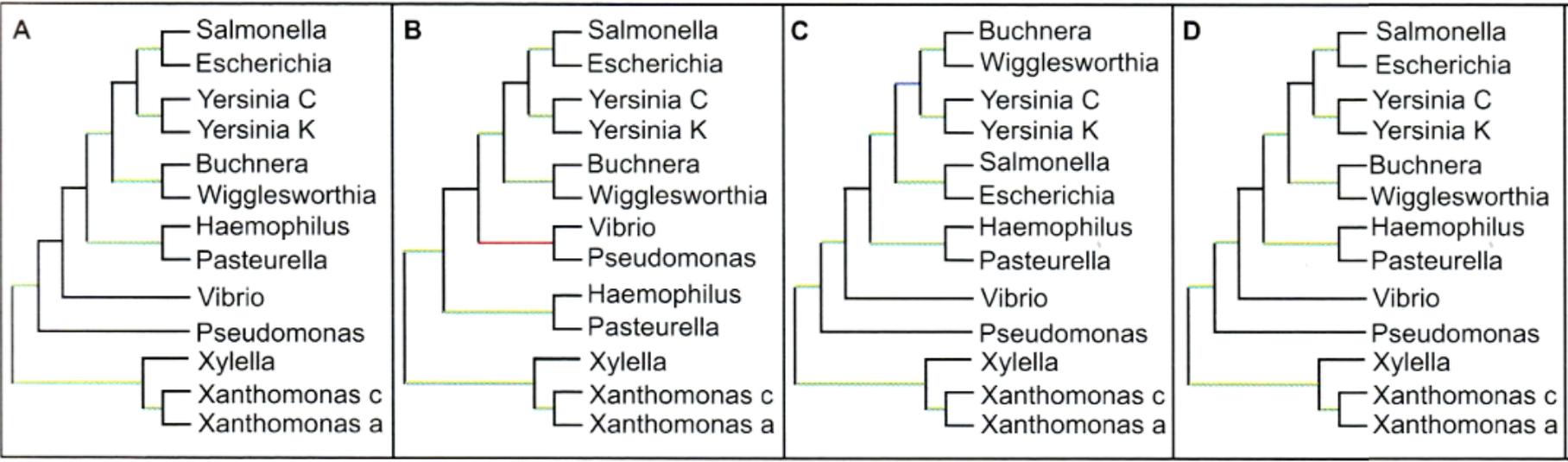
Les séquences homologues des 4 protéines sont alignées c'est-à-dire que les positions homologues de chaque séquence sont placées les unes en dessous des autres et représentent une colonne ou un site de l'alignement chaque colonne numérotée constitue un caractère moléculaire et chaque acide aminé constitue un état du caractère moléculaire.

	1	10	20	30	40	50	52														
1. Salmonella	VAKKATSGVPGK	FQTALIE	LIVGFVHGSVK	DM	-HGKSKLIA	PLAL	TIFVWVF														
2. Escherichia	VAKKATSGVPGK	FQTALIE	LVIGFVNGSVK	DM	-HGKSKLIA	PLAL	TIFVWVF														
3. Yersinia C	VAA SATSGVPGK	LQTAVE	LIIGFVDNSVR	DM	-HGKSKVIA	PLAL	TVFVWVL														
4. Yersinia K	VAA SATSGVPGK	LQTAVE	LIIGFVDNSVR	DM	-HGKSKVIA	PLAL	TVFVWVL														
5. Buchnera	VGKKITQGI	PGKLTALIE	LIFE FVNL	NVKS	SM	-QGKNALIA	PLSL	TVFIWVF													
6. Wigglesworthia	ISKKF	TIGTPGRI	QASIE	ILVE	FINS	NVKD	IFGH	NTNK	IIPPL	SLT	IFVWLF										
7. Haemophilus	VGKKA	TTGVP	PKMQC	LVE	IVVE	WVNG	I	VKEN	EF	-HGPR	NVVA	PLAL	TIFCWVF								
8. Pasteurella	VAKKAT	PGVPSKM	QC	FVE	IMVD	WD	G	IVKEN	EF	-HGPR	HAVG	PLAL	TIFCWVF								
9. Vibrio	MARKAT	AGVPGKL	QCAVE	MVVE	FVD	KNVK	D	TF	-HGR	NPLIA	PLAL	TIFCWVF									
10. Pseudomonas	A	AKKATSG	QPGGL	QNFVE	VMVE	FVD	TSVK	D	TF	-HGR	NPLIA	PLAL	TVFVWIF								
11. Xanthomonas a	A	TRNAT	AGVPGKW	QAFVE	IC	LE	FVDR	QAK	D	TY	-HGSS	SKLV	TP	IAIT	IFFWIL						
12. Xanthomonas c	A	TRNAT	AGVPGKW	QAFVE	IC	LE	FVDR	QAK	D	TY	-HGTS	SKLV	TP	IAIT	IFFWIL						
13. Xylella	G	TCKAT	SGI	PGKW	QAFVE	V	C	LE	FVDR	Q	V	K	D	TY	-HG	PS	SKLV	TP	IAIT	L	FFWIF

### III. IDENTIFIER UN TRANSFERT HORIZONTAL DE GÈNES (THG)

#### III.1. Alignement de séquences protéiques et comparaison des arbres phylogénétiques chez les bactéries

9. Identifiez les congruences. Comment les interpréteriez-vous ?



**Document 5 : arbre phylogénétique estimé pour les mêmes bactéries mais à partir de 3 protéines différentes**

**(A) Arbre d'ATPB (estimé à partir d'un alignement de 298 sites) ; (B) : Arbre de bio B (368 sites). (C) : Arbre de GYRA (1042 sites) (in Tout-En-Un, BCPST 2, Dunod 2022, modifié). Les branches bleue et rouge sont soutenues respectivement de manière faible et forte par une analyse de bootstrap.**

### III. IDENTIFIER UN TRANSFERT HORIZONTAL DE GÈNES (THG)

#### III.1. Alignement de séquences protéiques et comparaison des arbres phylogénétiques chez les bactéries

9. *Identifiez les congruences. Comment les interpréteriez-vous ?*

Les points communs ou **congruences** s'interprètent par le fait que les gènes qui codent ces protéines sont liés dans un même génome bactérien. Ils ont tous subi la même histoire évolutive, conduisant ainsi à inférer la même phylogénie. Les clades définis par les branches vertes se retrouvent ainsi dans tous les arbres.

*Salmonella + Escherichia,*

*Les 2 souches de Yersinia,*

*Buchnera + Wigglesworthia,*

*Le clade formé par les 6 bactéries précédemment mentionnées*

*Haemophilus + Pasteurella,*

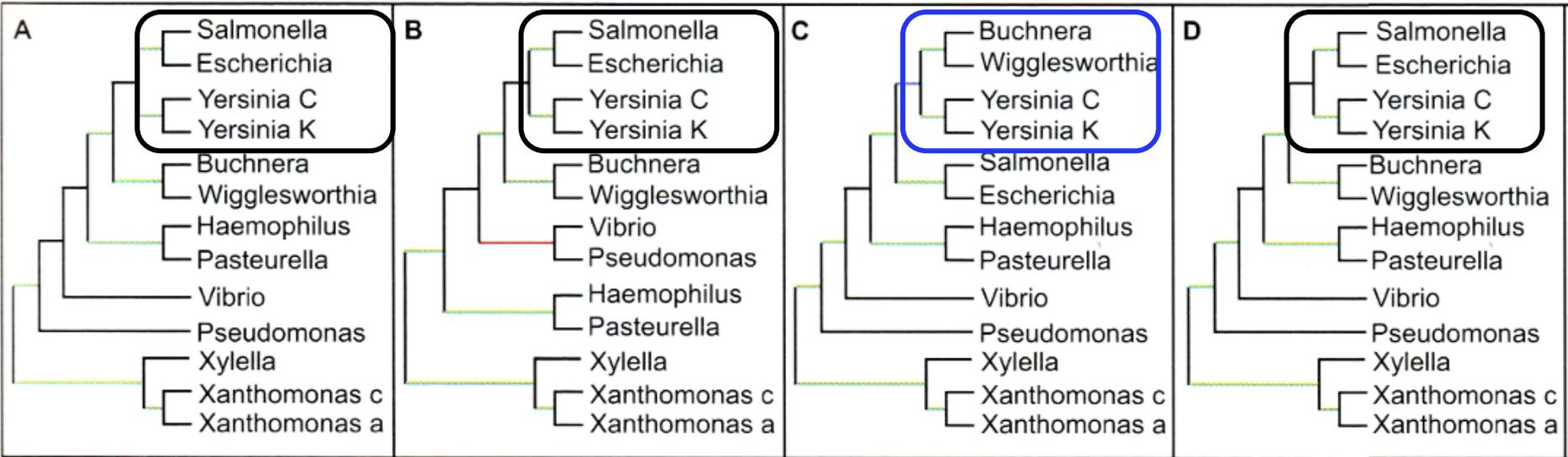
*Xanthomonas + Xylella.* Ces 3 bactéries forment le groupe frère de toutes les autres car elle sont été définies comme groupe externe.

10. Identifiez les incongruences.

Les différences topologiques ou **incongruences** sont les suivantes :

*Yersinia* est le groupe-frère de *Buchnera* + *Wigglesworthia* pour GYRA (branche bleue de la figure C) alors que c'est le groupe frère de *Salmonella* + *Escherichia* pour les 3 autres protéines.

*Vibrio* est le groupe-frère de *Pseudomonas* pour BIOB (branche rouge de la figure B) alors que c'est le groupe-frère du clade *Salmonella* + *Escherichia* + *Yersinia*, + *Buchnera* + *Wigglesworthia* + *Haemophilus* + *Pasteurella* pour les 3 autres protéines.



ATPB

BIOB

GYRA

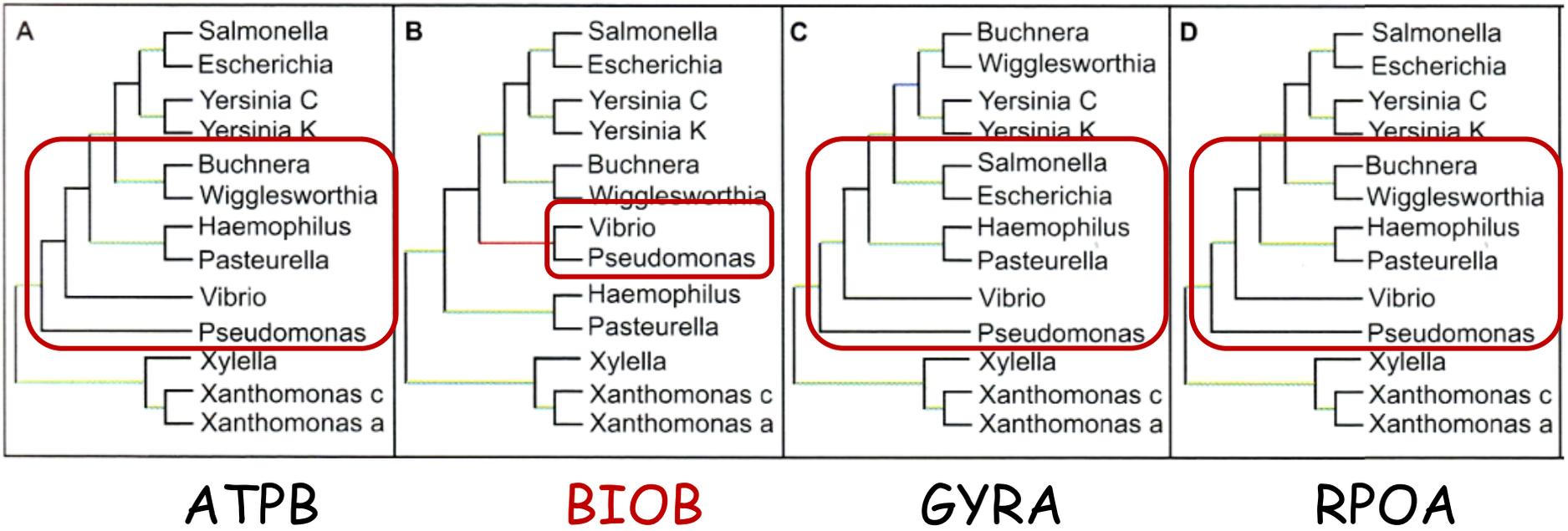
RPOA

10. Identifiez les incongruences.

Les différences topologiques ou **incongruences** sont les suivantes :

*Yersinia* est le groupe-frère de *Buchnera* + *Wigglesworthia* pour GYRA (branche bleue de la figure C) alors que c'est le groupe frère de *Salmonella* + *Escherichia* pour les 3 autres protéines.

*Vibrio* est le groupe-frère de *Pseudomonas* pour BIOB (branche rouge de la figure B) alors que c'est le groupe-frère du clade *Salmonella* + *Escherichia* + *Yersinia*, + *Buchnera* + *Wigglesworthia* + *Haemophilus* + *Pasteurella* pour les 3 autres protéines.



### III.2. Interprétation de l'incongruence entre arbres : l'hypothèse d'un THG

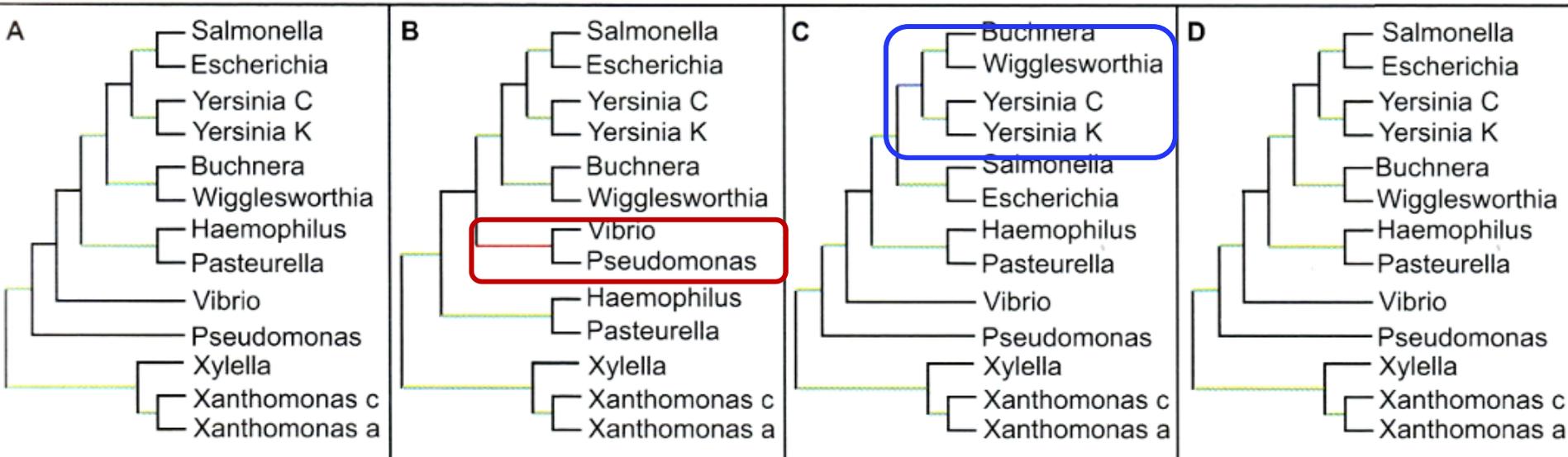
11. Comparez la phylogénie de GYRA et de BIOB. Laquelle des deux vous semble être la plus robuste ? Justifiez.

Pour la **phylogénie de GYRA** : la valeur de bootstrap est faible traduisant un signal phylogénétique faible. La branche est peu robuste. Ce n'est pas une véritable incongruence mais plutôt un manque d'information phylogénétique.

Pour la **phylogénie de BIOB** : la valeur de bootstrap est forte traduisant un signal phylogénétique fort. La branche est suffisamment robuste pour considérer une véritable incongruence révélant un évènement biologique.

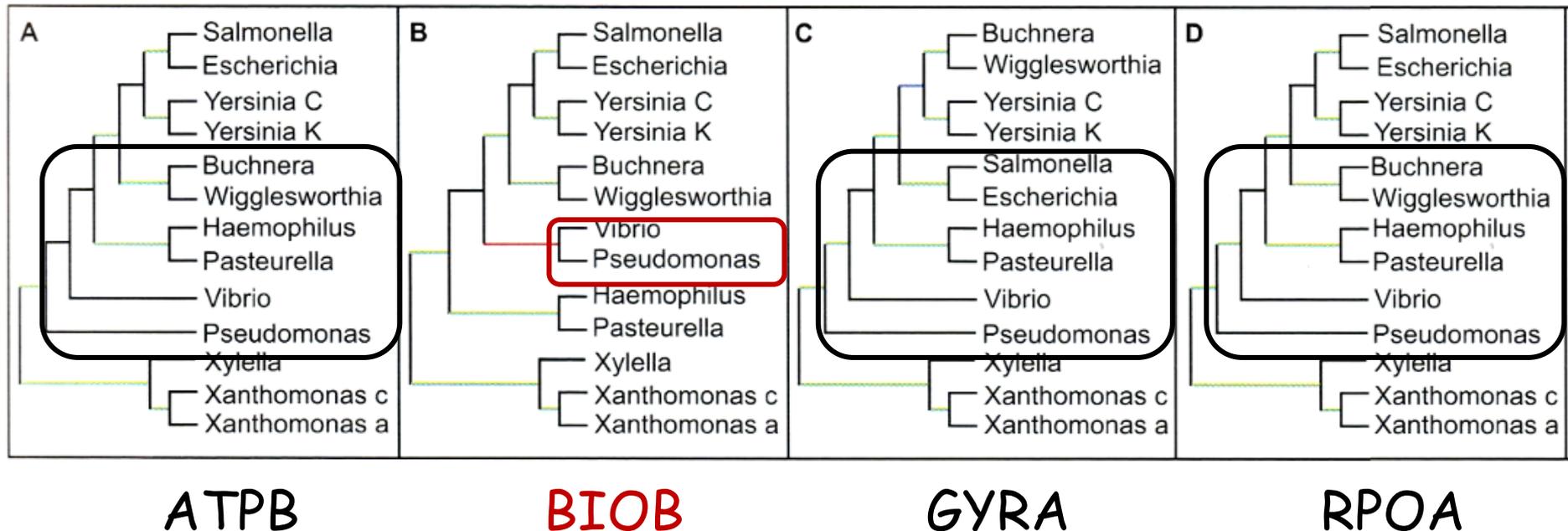
**BIOB**

**GYRA**



## 12. Discutez l'hypothèse d'un THG entre *Vibrio* et *Pseudomonas*.

Les phylogénies A (ATPB), C (GYRA) et D (RPOA) montrent une position plus profonde de *Pseudomonas* dans l'arbre tandis que *Vibrio* est un groupe-frère de *Salmonella* + *Escherichia* + *Yersinia*, + *Buchnera* + *Wigglesworthia* + *Haemophilus* + *Pasteurella*. On peut supposer que cet ordre de branchement majoritaire reflète la phylogénie des espèces étudiées. Le branchement minoritaire *Vibrio* + *Pseudomonas* suggère qu'un évènement génétique a affecté le gène qui code pour BIOB.

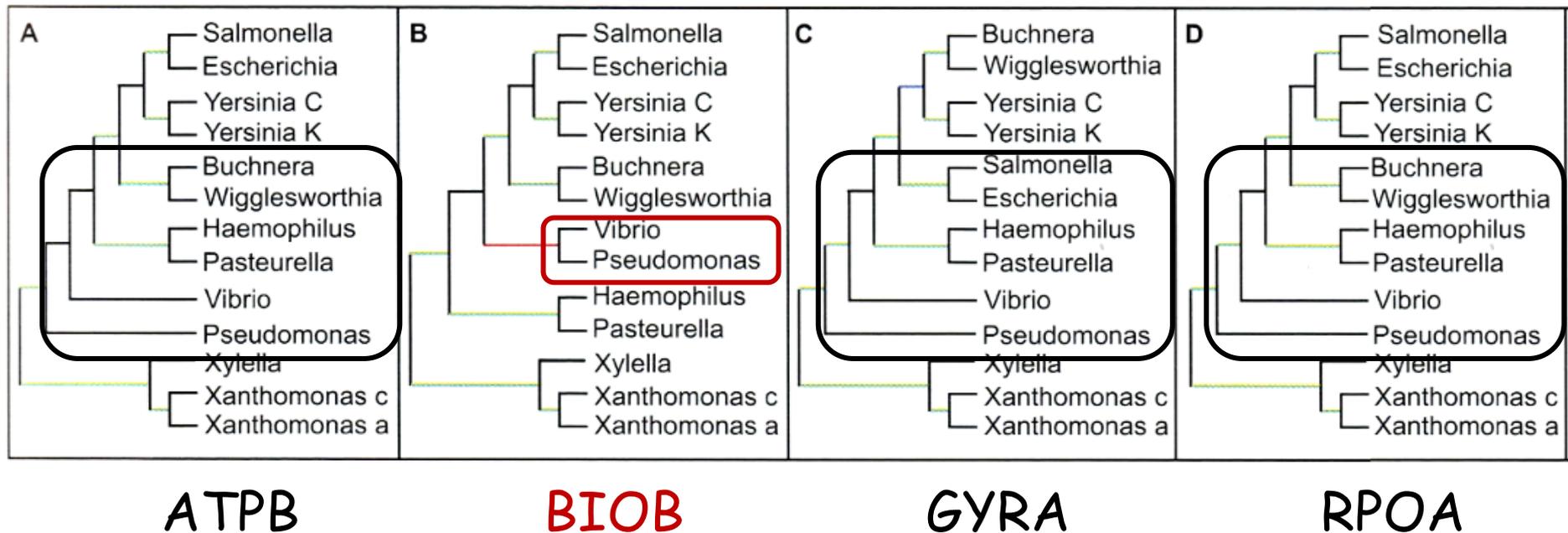


## 12. Discutez l'hypothèse d'un THG entre *Vibrio* et *Pseudomonas*.

Chez les bactéries la cause la plus fréquente d'incongruence entre un arbre de gènes (phylogénie estimée à partir de la protéine BIOB) et l'arbre des espèces (phylogénie estimée à partir des 3 autres protéines) est le transfert horizontal de gènes (THG). **L'incongruence est ici la position anormalement proche de *Vibrio* et *Pseudomonas*.**

Nous pouvons donc supposer qu'un ancêtre de *Vibrio* a donné une copie du gène codant BIOB à un ancêtre de *Pseudomonas*, remplaçant ainsi la copie présente dans ce dernier taxon. *Vibrio* serait le **donneur** et *Pseudomonas* le **receveur** conduisant à une protéine BIOB de *Pseudomonas* très apparentée à la protéine BIOB de *Vibrio*.

En absence de THG, *Pseudomonas* occuperait une position plus éloignée de *Vibrio* comme c'est le cas pour les 3 autres d'arbres d'ATPB, GYRA et RPOA.



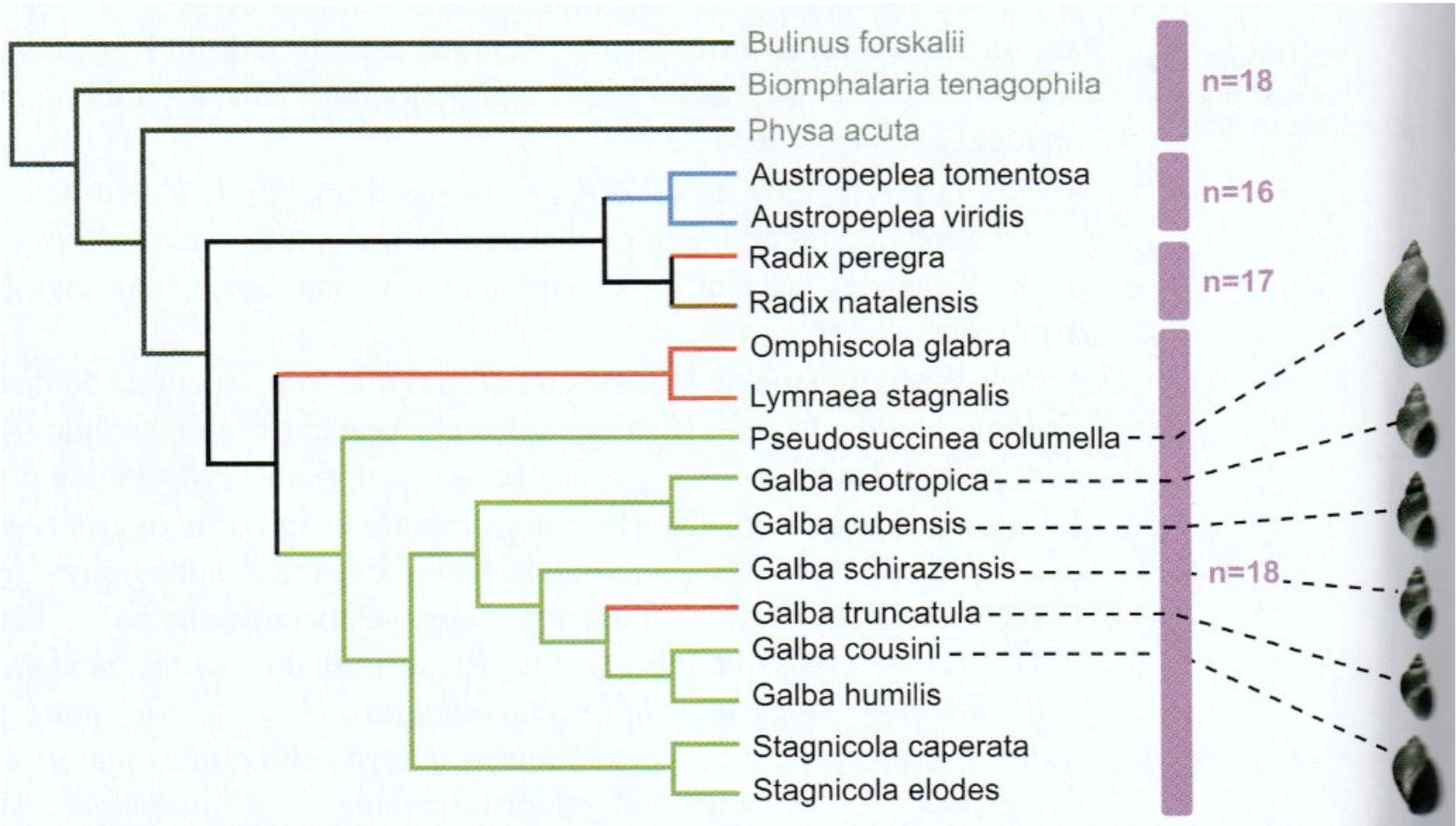
Remarques :

1. Pour reconstruire la phylogénie des espèces il faut d'abord identifier et éliminer les gènes qui ont été affectés par des événements de THG.
2. Dans le cadre de l'approche phylogénétique il faut combiner tous les alignements non affectés par les THG et réanalyser l'alignement ainsi obtenu. Ceci permet d'additionner les homologues i.e le signal phylogénétique de chacun des marqueurs moléculaires pris indépendamment.

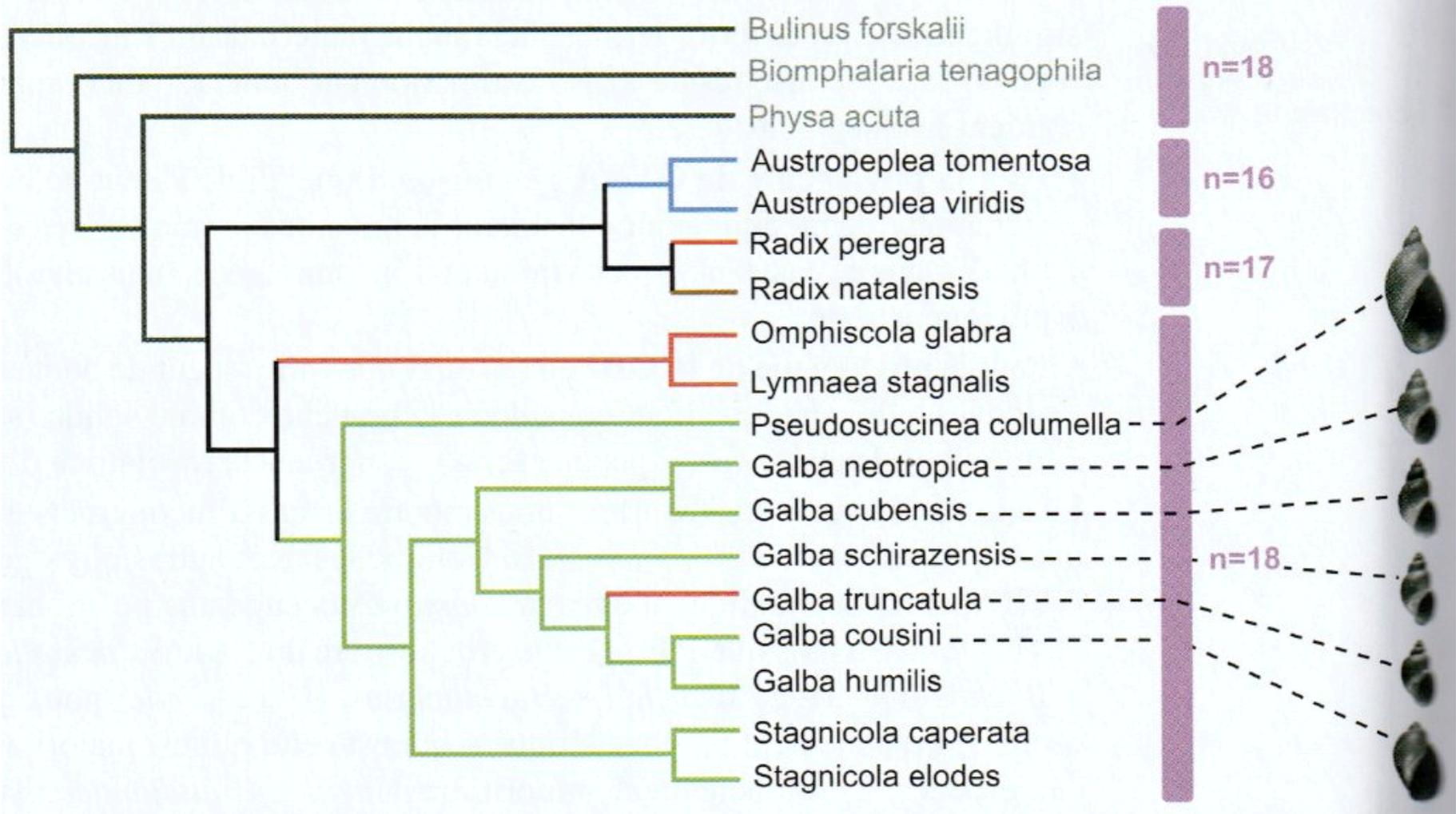
Si la fréquence des THG est trop importante, il serait plus approprié de reconstruire à l'évolution des taxons sous la forme d'un réseau plutôt que d'un arbre.

## IV. RETRACER UN SCÉNARIO DE COLONISATION GÉOGRAPHIQUE

13. Identifiez ce que proposent chacun de ces trois scénarios.



Document 6 : Arbre phylogénétique de 7 genres de limnées et trois groupes externes (en gris) inféré par un maximum de vraisemblance à partir des séquences d'ADN mitochondrial (ARN 16S) et nucléaire (espaceurs ribosomiques ITS1 et ITS2) (in Tout-En-Un, BCPST 2, Dunod 2022).



Couleur des branches = origine géographique des limnées :

Australasie (bleu)

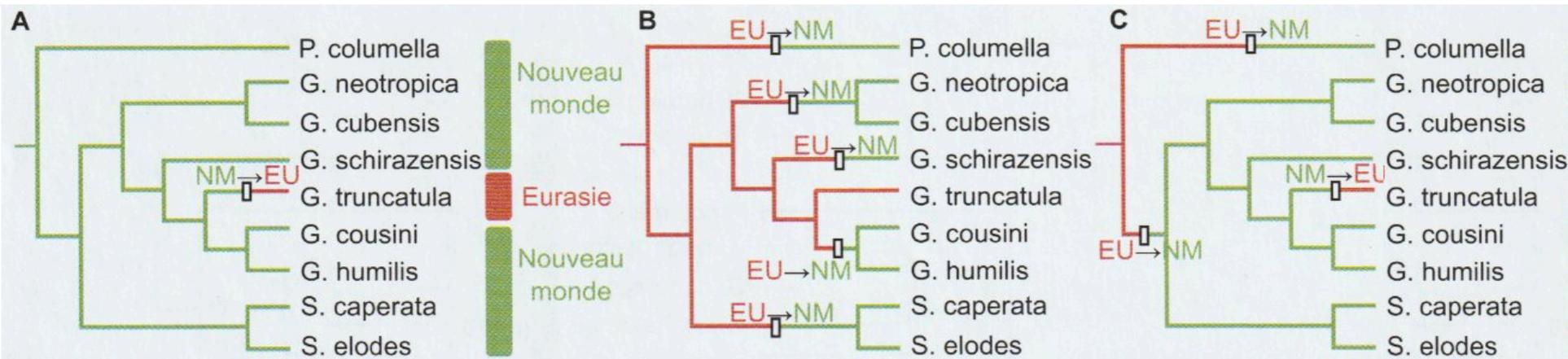
Eurasie (rouge)

Afrique (marron)

Amérique (vert).

## IV. RETRACER UN SCÉNARIO DE COLONISATION GÉOGRAPHIQUE

13. Identifiez ce que proposent chacun de ces trois scénarios.



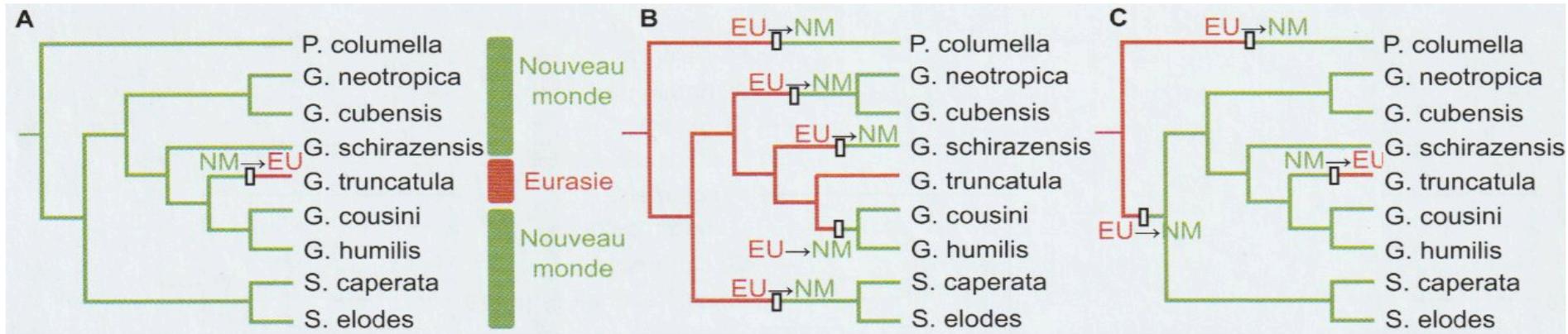
**(A)** Scénario dans lequel la distribution dans le nouveau monde (NM) est ancestrale. Un seul changement d'aire de répartition vers l'Eurasie (EU) permettrait d'expliquer ce qui est observé pour *Galba truncatula*.

**(B)** Scénario dans lequel la distribution en Eurasie est ancestrale. Cinq changements convergents d'aire de répartition vers le nouveau monde expliquerait la répartition actuelle des lignées

**(C)** Scénario dans lequel la distribution en Eurasie est ancestrale. Deux changements convergents d'aire de répartition vers le nouveau monde et une réversion vers l'Eurasie expliquerait la répartition délimitée

## IV. RETRACER UN SCÉNARIO DE COLONISATION GÉOGRAPHIQUE

14. Discutez leur validité et dites quel est celui qui vous paraît le plus parcimonieux.



**Scénario (A)** : un seul changement d'aire géographique pour *G. truncatula* ce qui constituerait un événement de migration du nouveau monde vers l'ancien monde.

**Le scénario (B)** : changements convergents d'aire géographique nécessitant un total de 5 migrations vers le nouveau monde.

**Scénario (C)** : alternatif en autorisant à la fois les changements convergents d'aires géographiques et un retour vers l'aire d'origine i.e une réversion, ce qui ne nécessiterait plus que 3 migrations.

Scénario A le plus parcimonieux : *G. truncatula* ne migre qu'une seule fois vers l'Eurasie