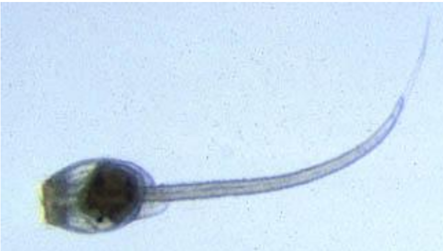
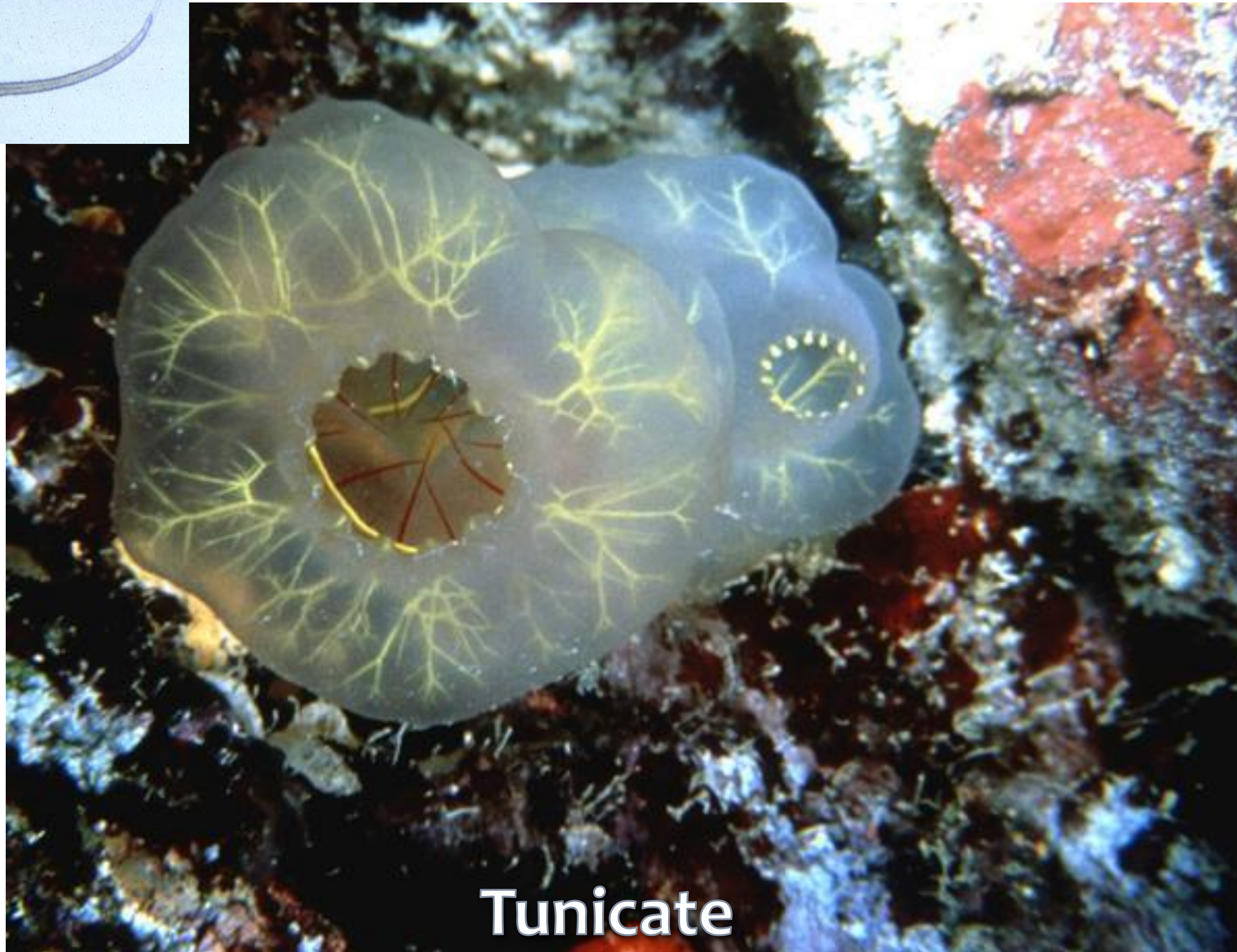
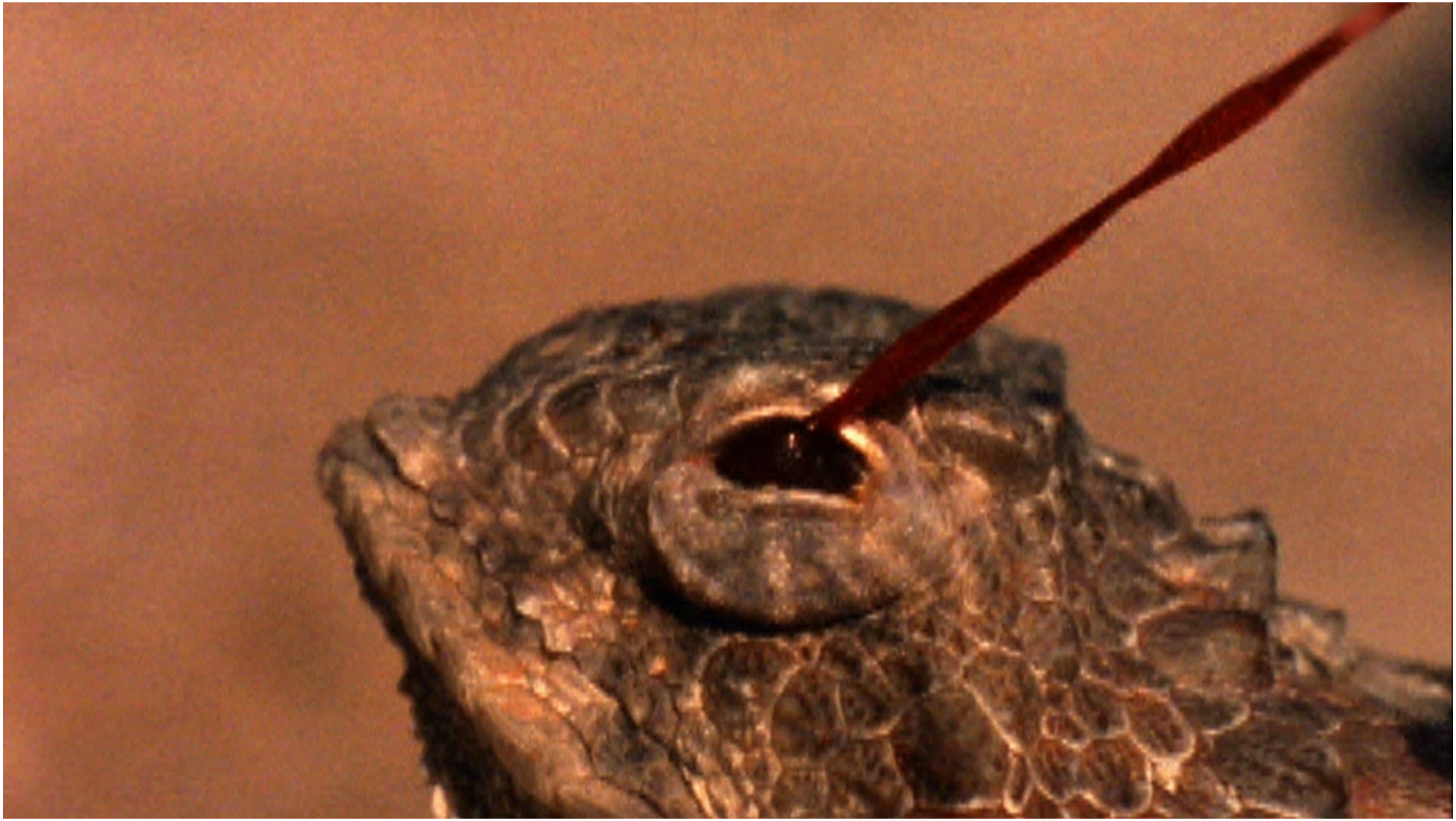
INFÉRENCE D'HISTOIRES ÉVOLUTIVES DE GÈNES PAR SPECIATION ET DUPLICATION

Manuel Lafond, LBIT

Université de Montréal




Tunicate



Phrynosoma


Turritopsis dohrnii

Introduction

- Arbres d’espèces, arbres de gènes
- Duplication et spéciation

Algorithmes et graphes

- Graphe de relation de gènes
- Reconstruction d’arbres à partir de graphes de relations

Conclusion

Gène RPGR :

Retinitis pigmentosa GTPase regulator

Participe à la pigmentation des yeux.

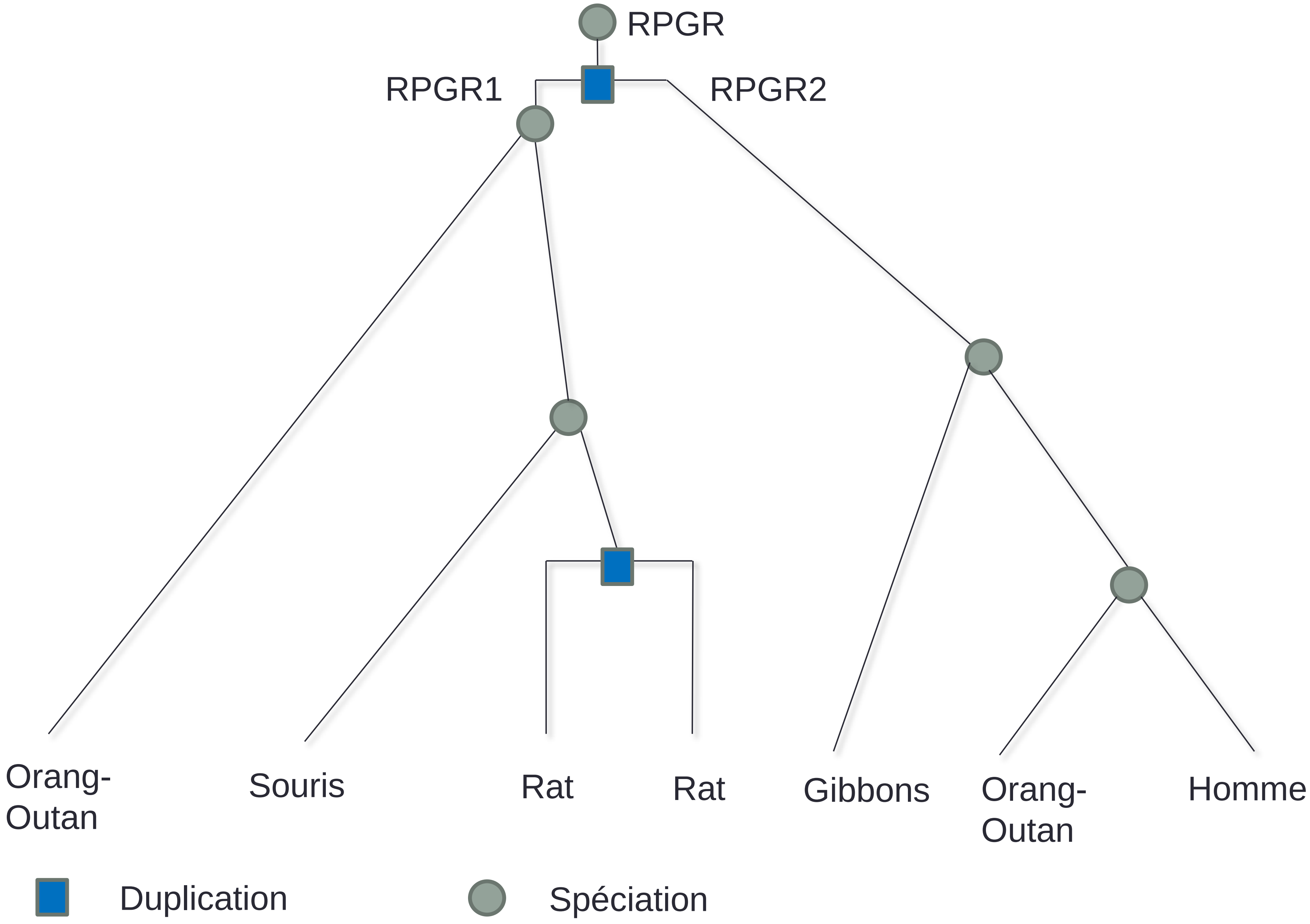
On s'intéresse à l'**histoire** du gène RPGR.

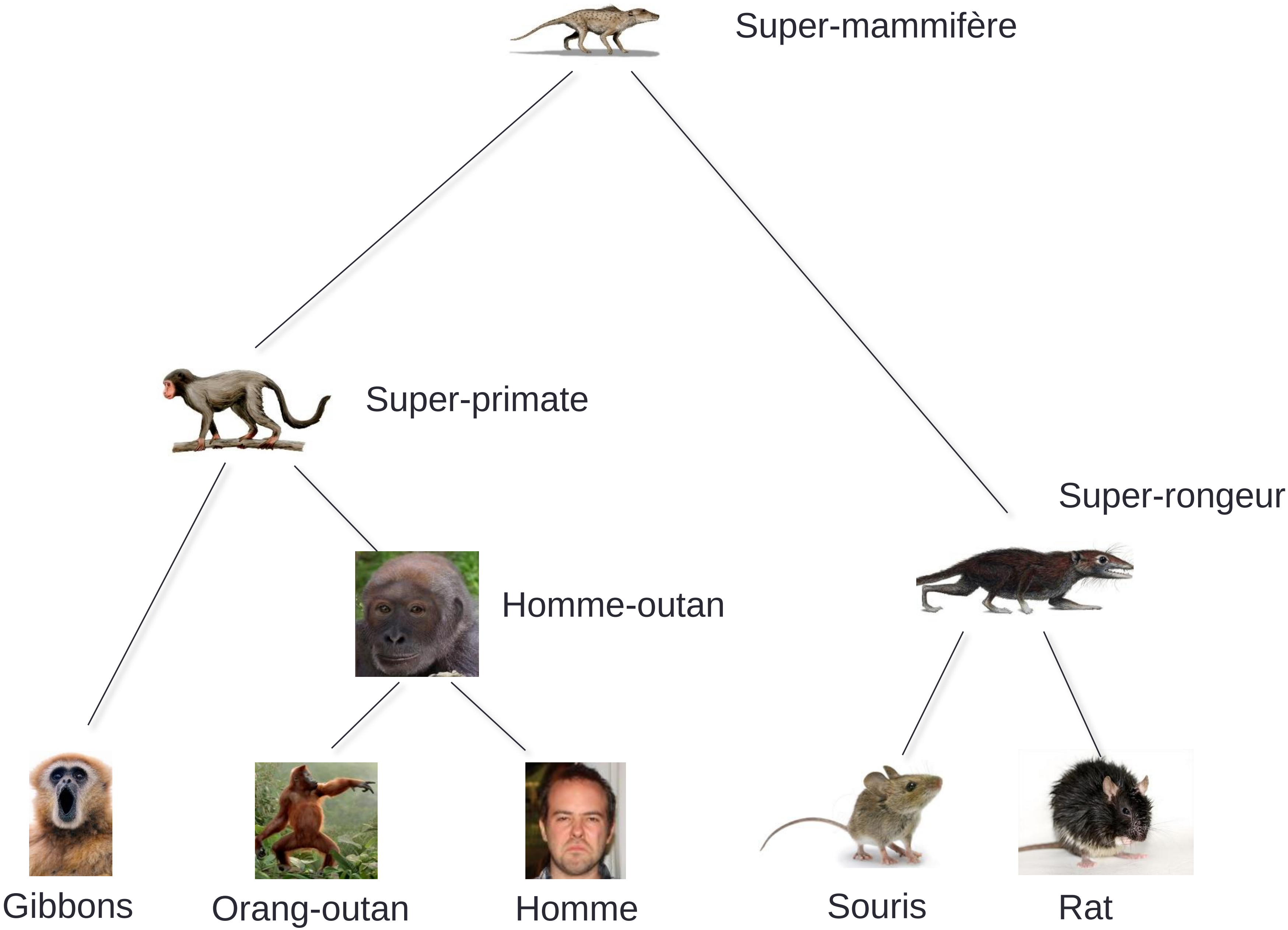
Presque tous les vertébrés ont une copie de ce gène.

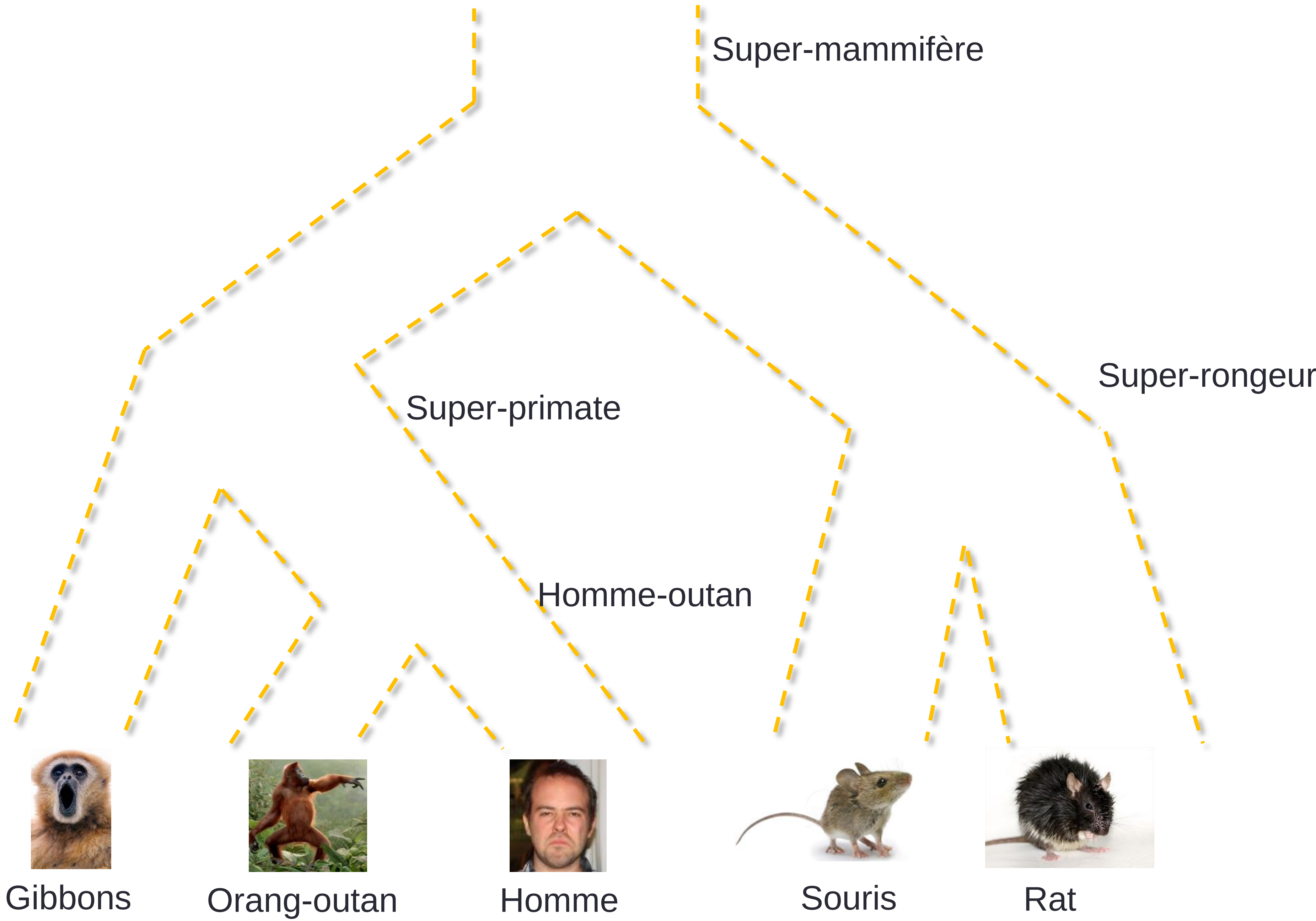
Qui est l'ancêtre de toutes ces copies ? Quelle est son histoire ?

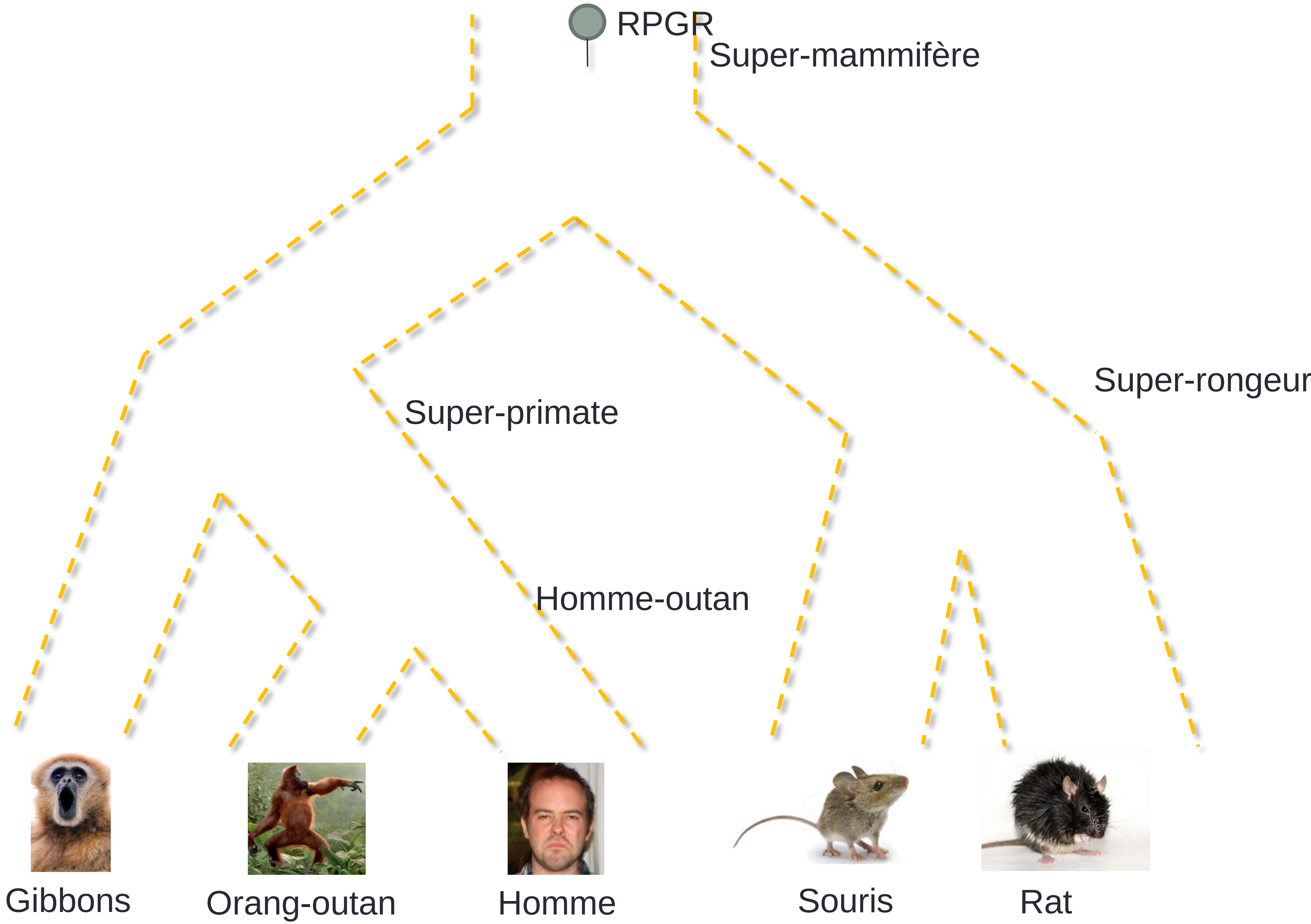
Un gène peut être :

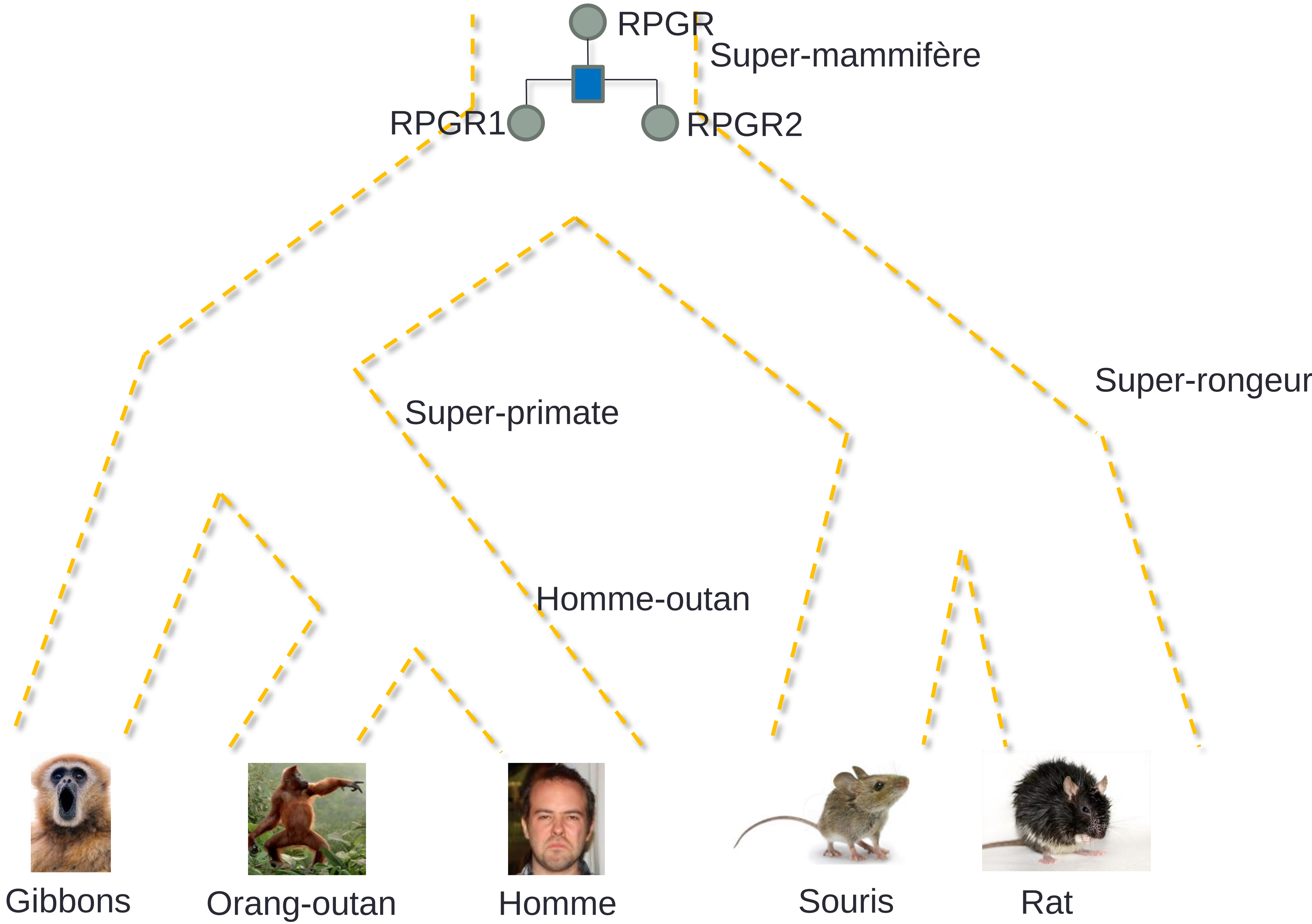
- **Transmis** aux espèces descendantes par **spéciation**
- **Dupliqué**
- **Perdu**

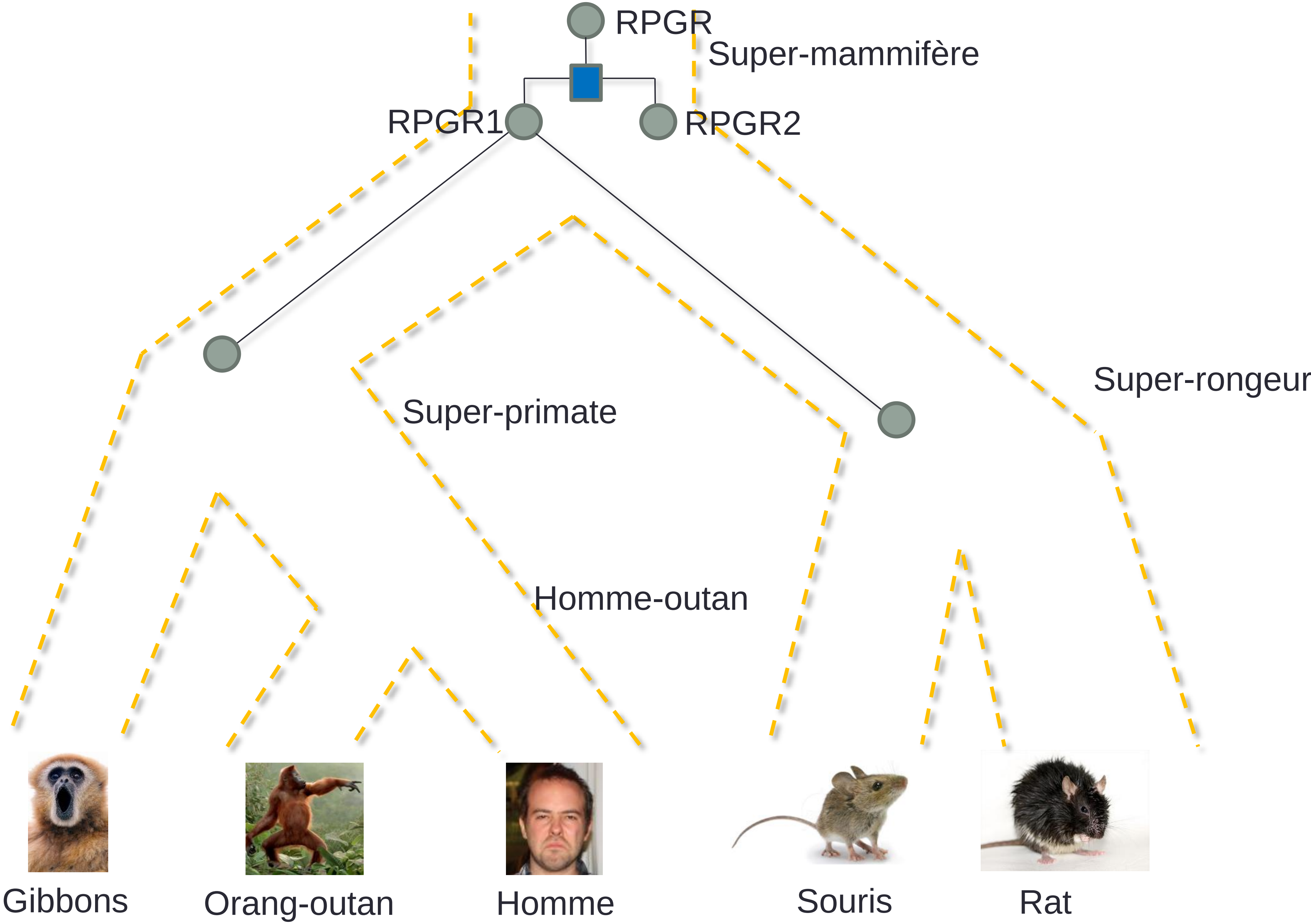

RPGR
RPGR1
RPGR2
Orang-Outan
Souris
Rat
Rat
Gibbons
Orang-Outan
Homme
Duplication
Spéciation


Super-mammifère
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

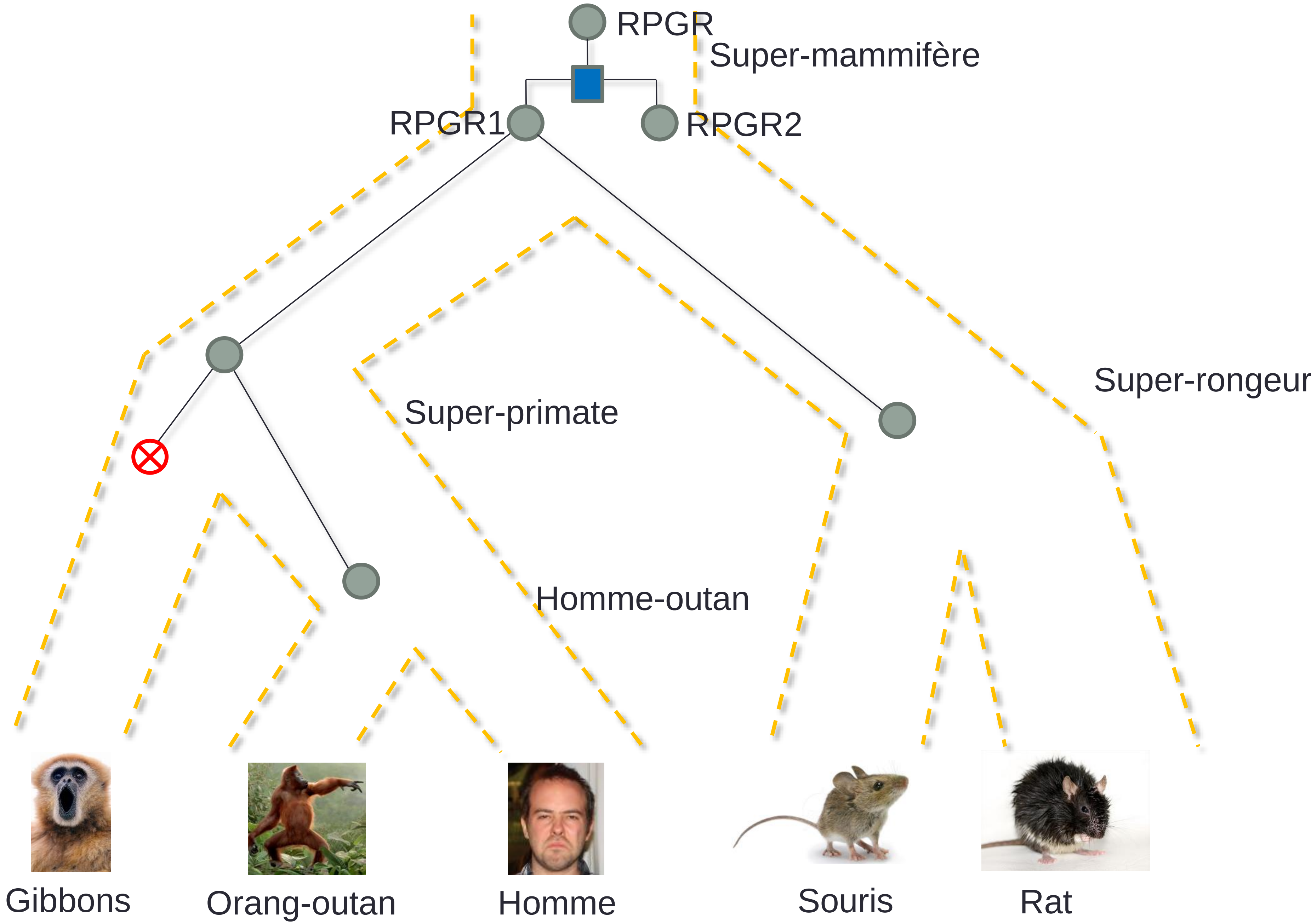

Super-mammifère
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

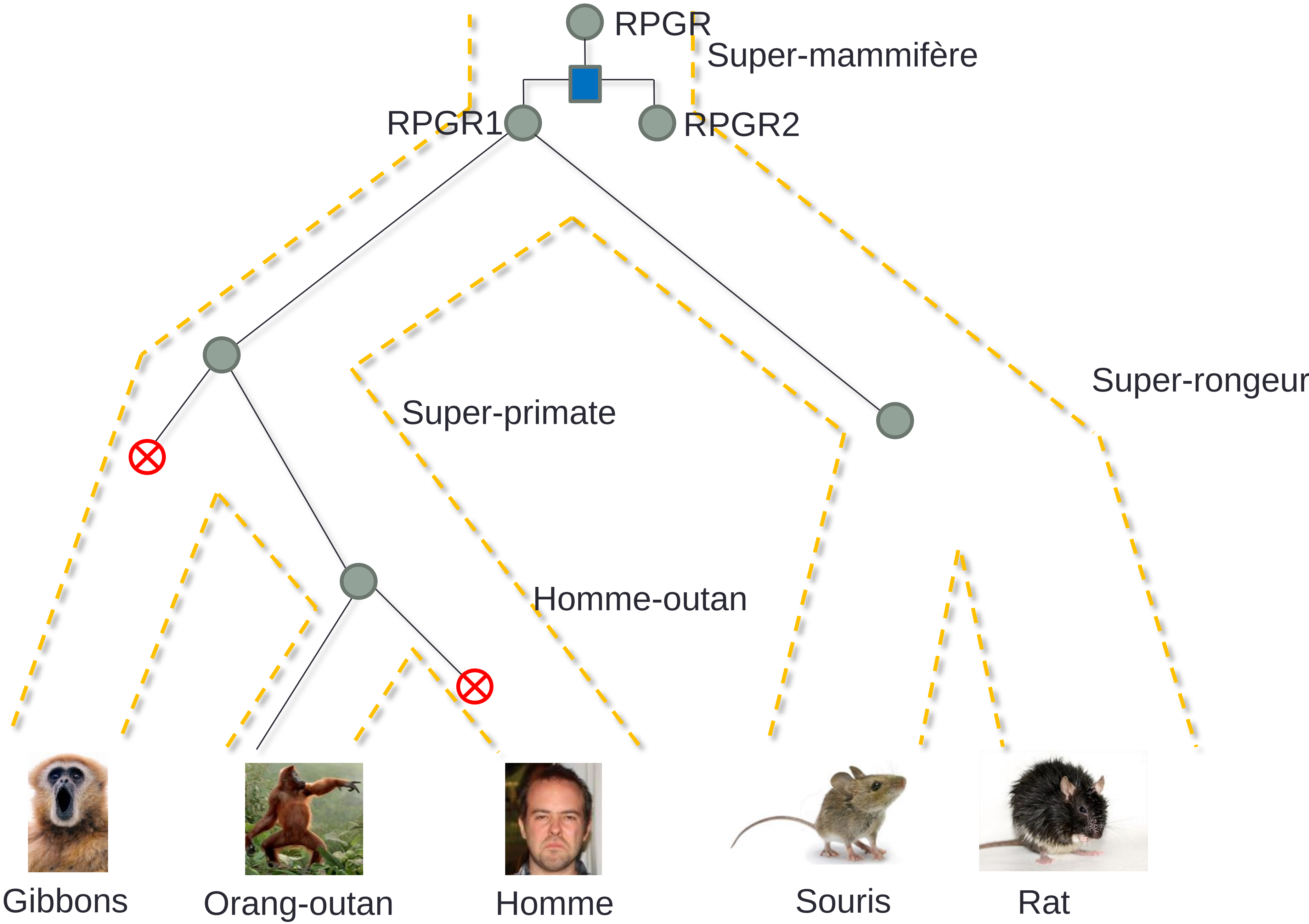

RPGR
Super-mammifère
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

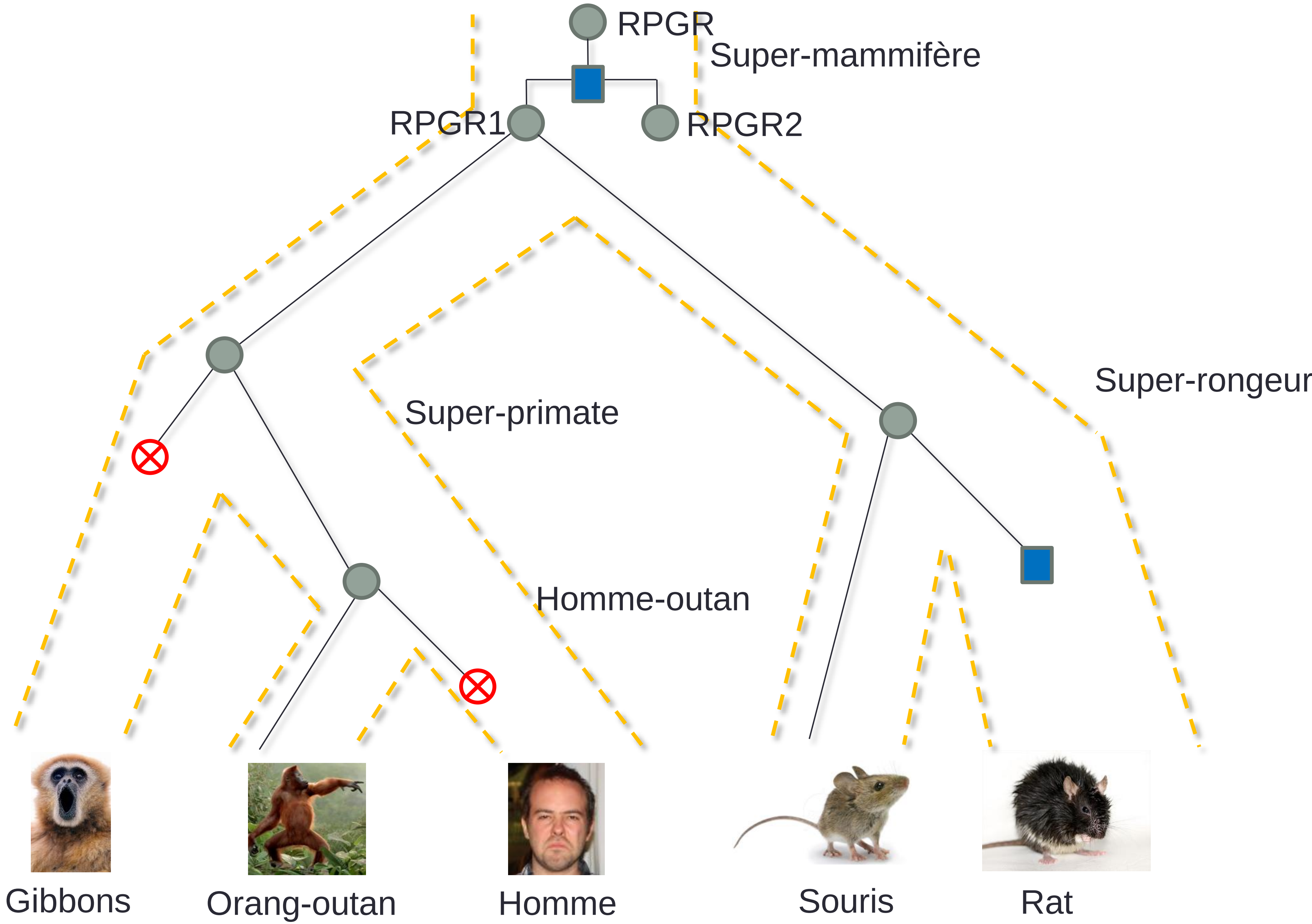

RPGR
Super-mammifère
RPGR1
RPGR2
Super-rongeur
Super-primate
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat



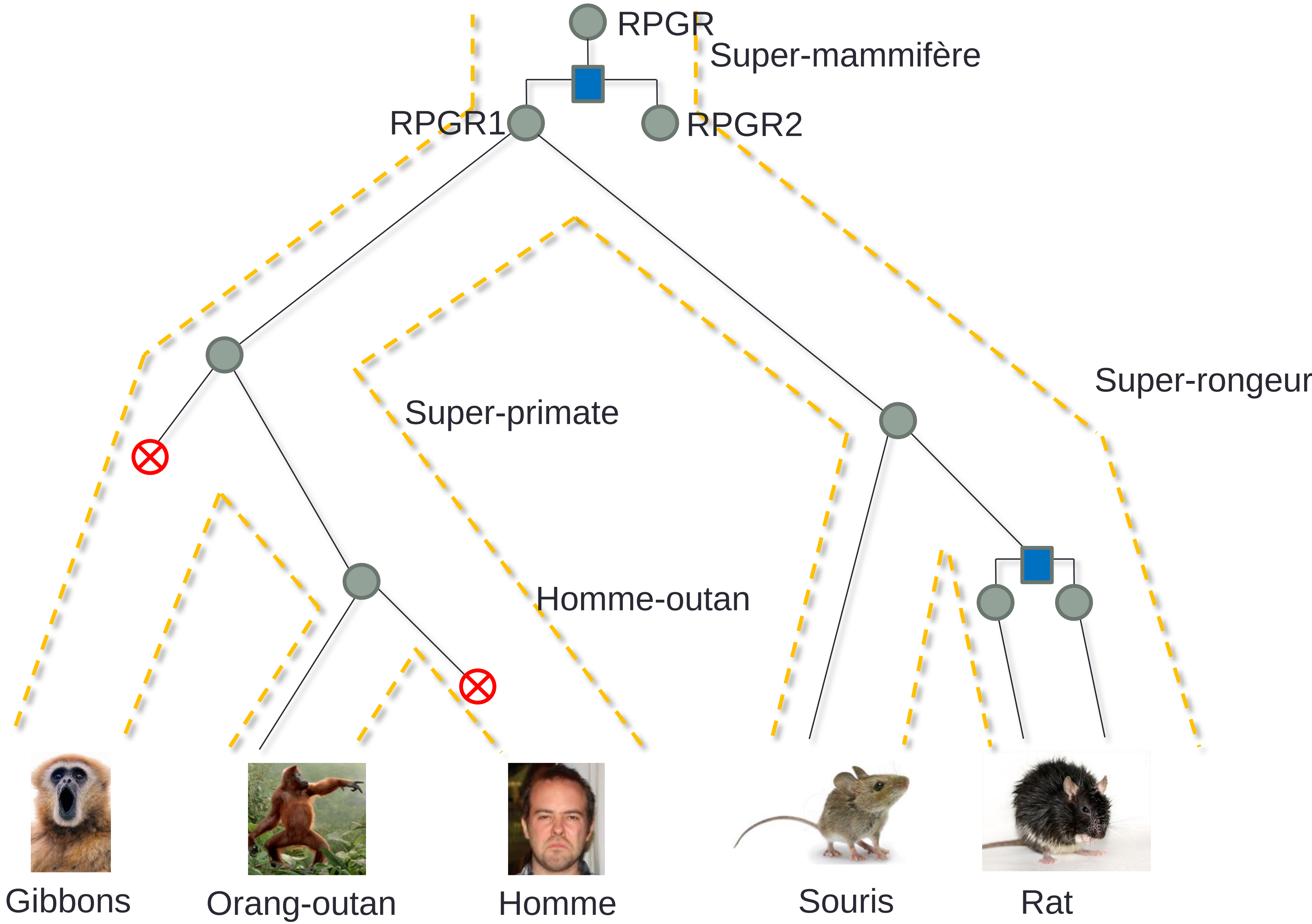
RPGR
Super-mammifère
RPGR1
RPGR2
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

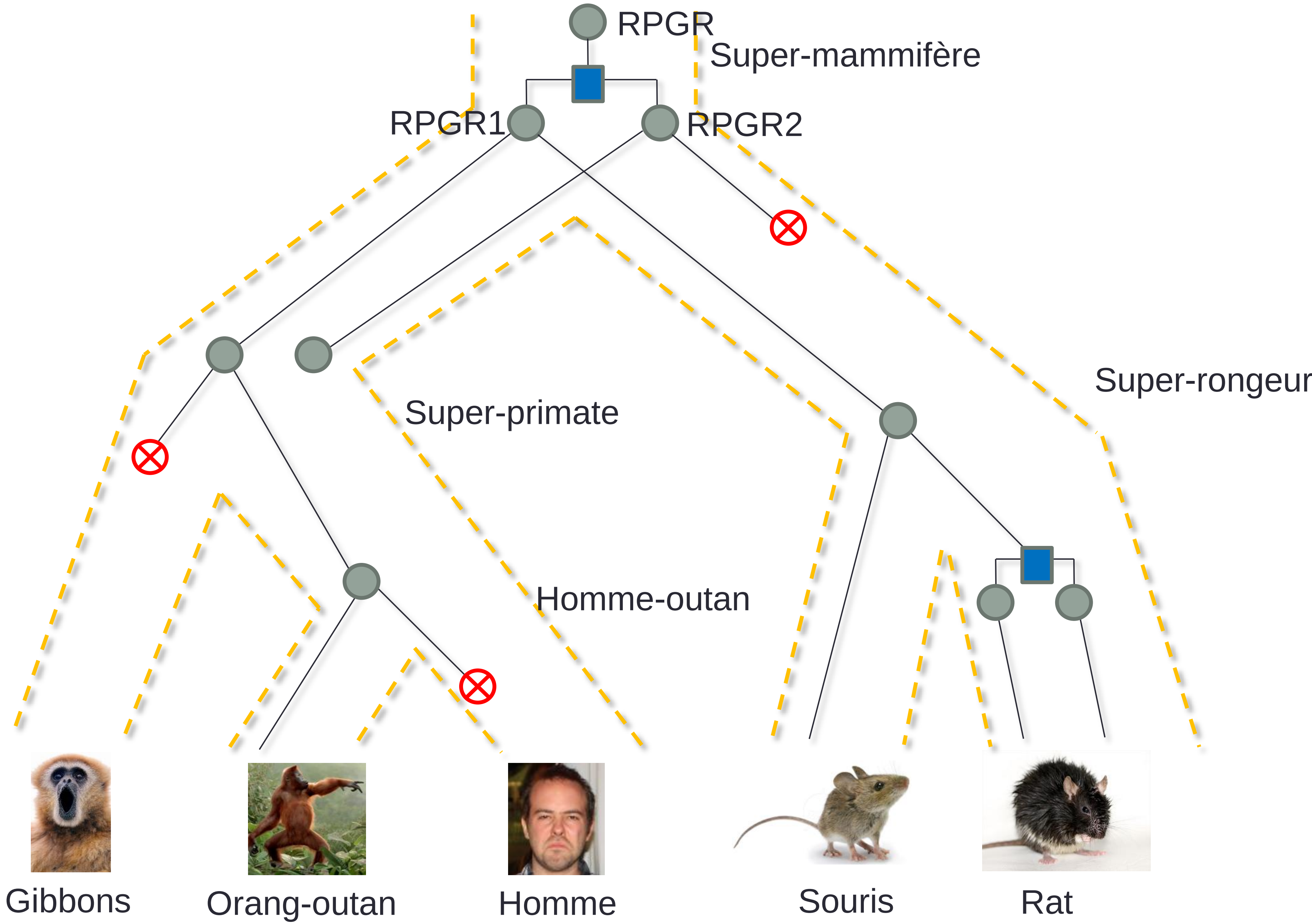

RPGR
Super-mammifère
RPGR1
RPGR2
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

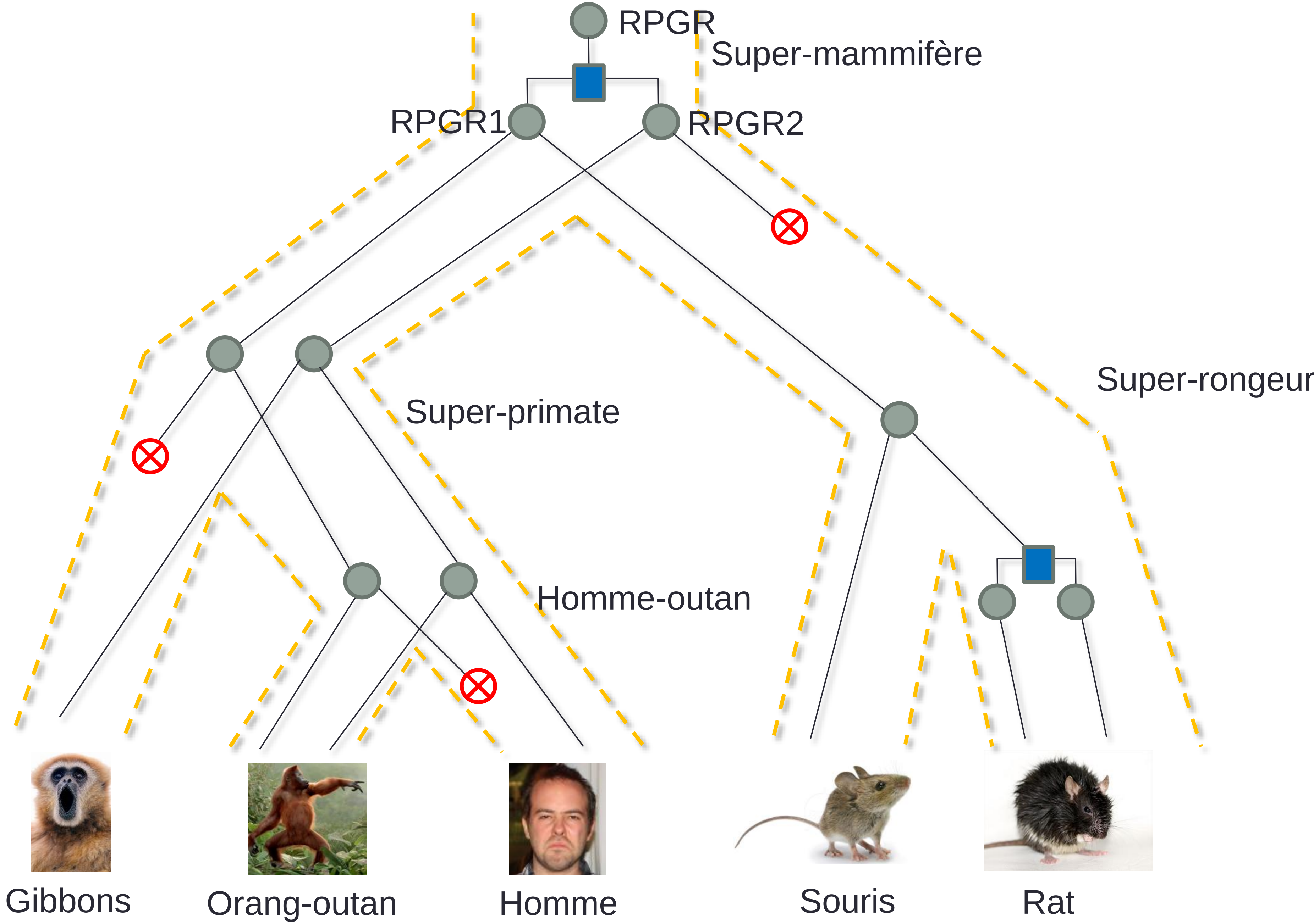

RPGR
Super-mammifère
RPGR1
RPGR2
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

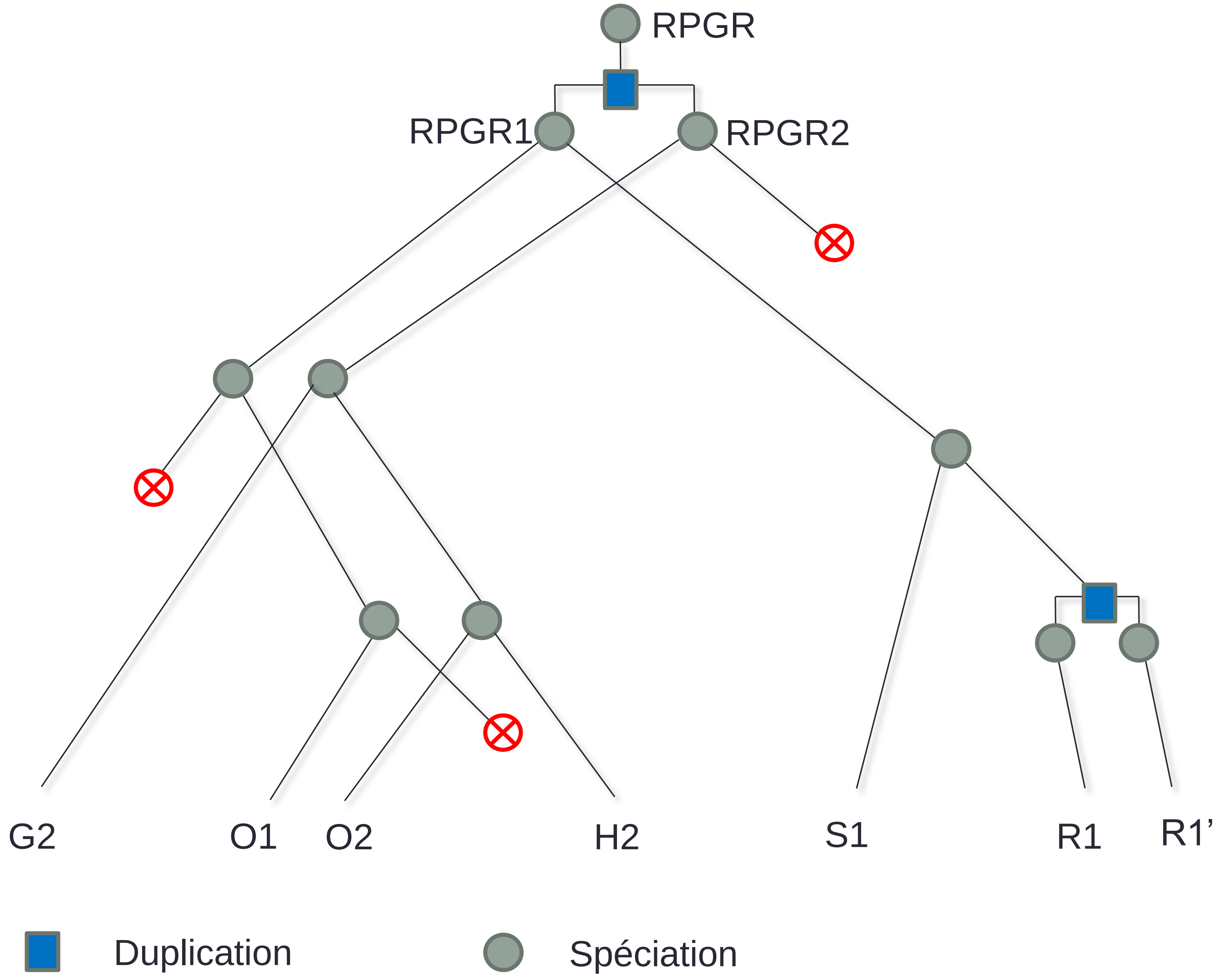


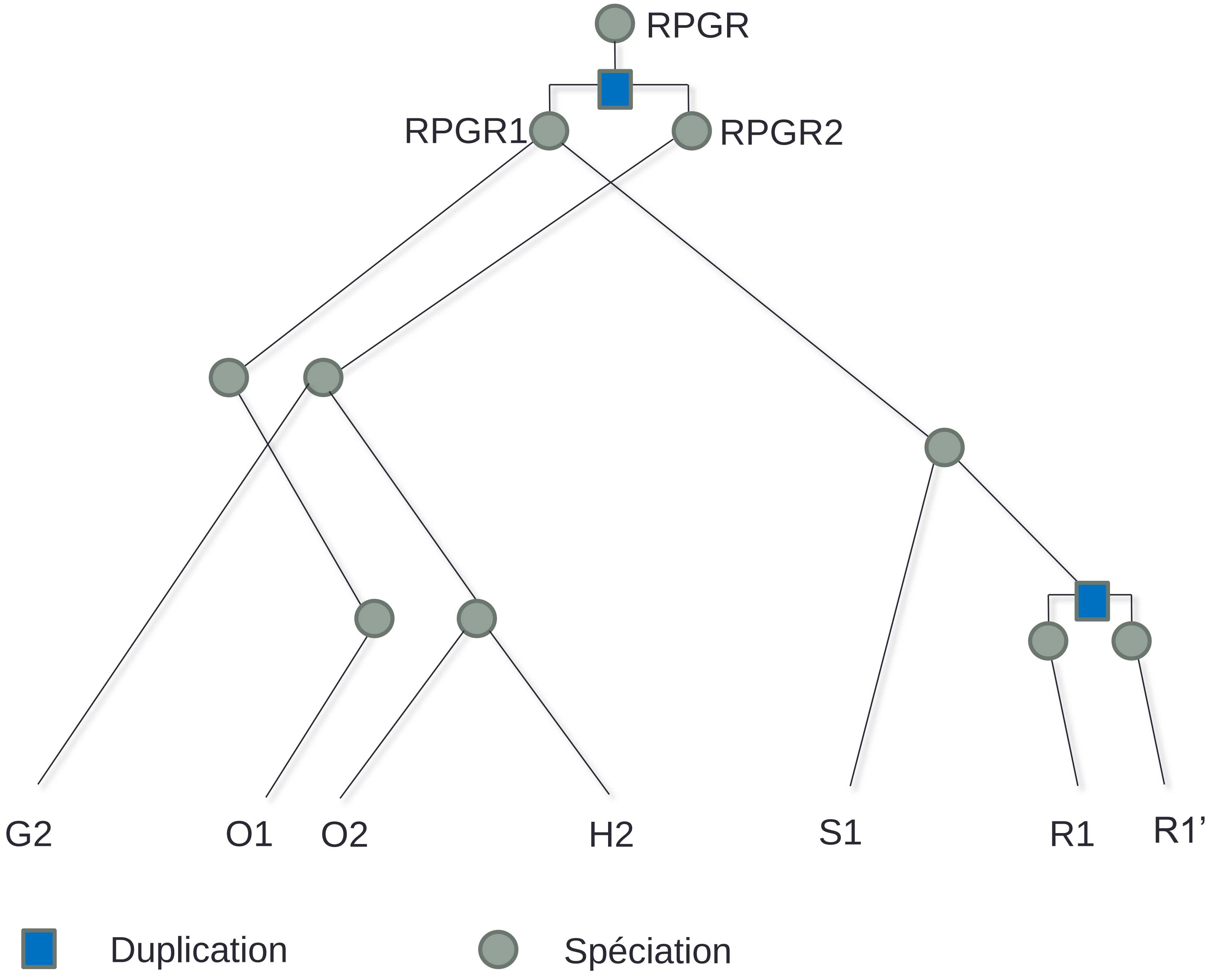
RPGR
Super-mammifère
RPGR1
RPGR2
Super-rongeur
Super-primate
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

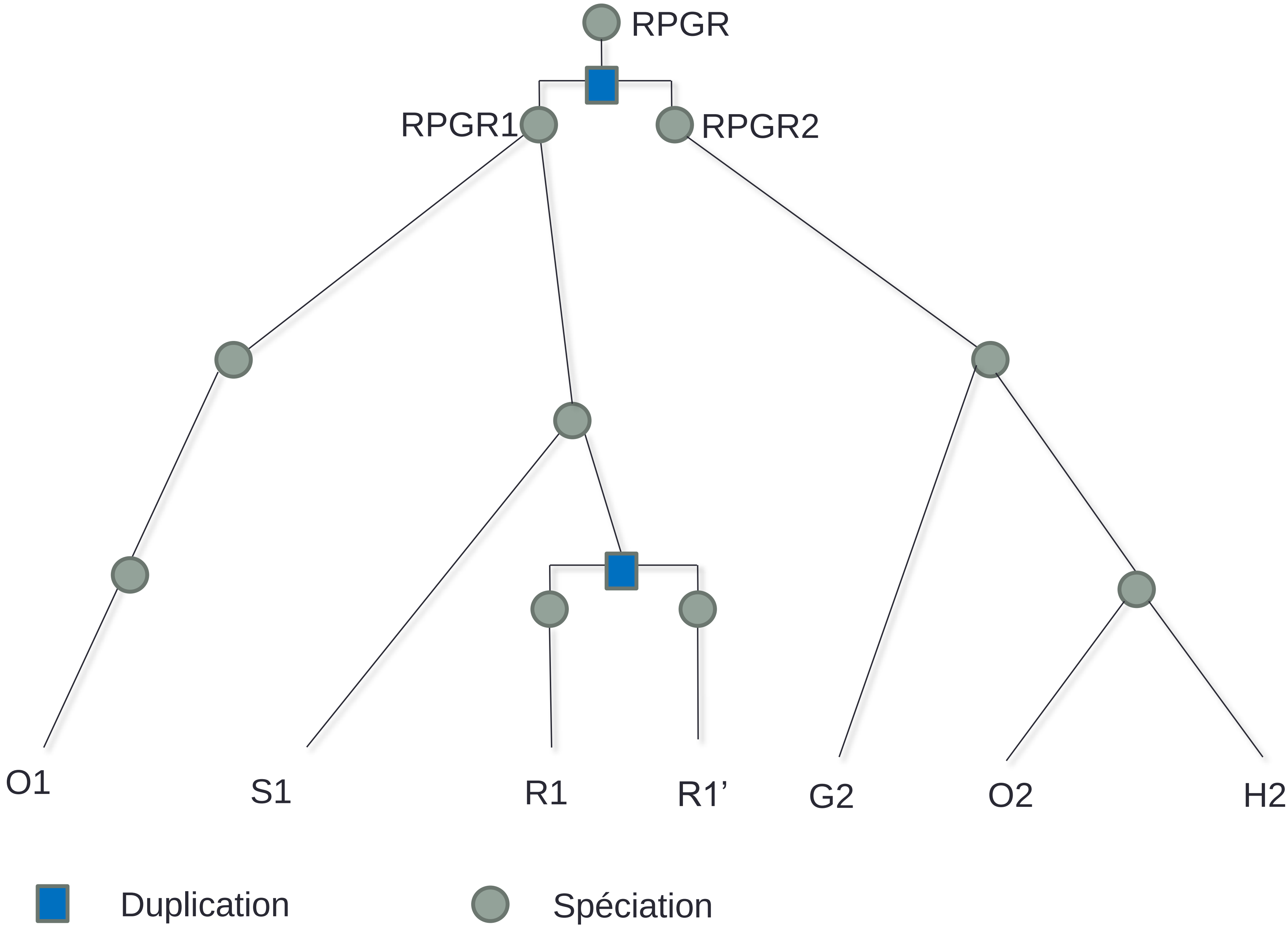

RPGR
Super-mammifère
RPGR1
RPGR2
Super-primate
Super-rongeur
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat

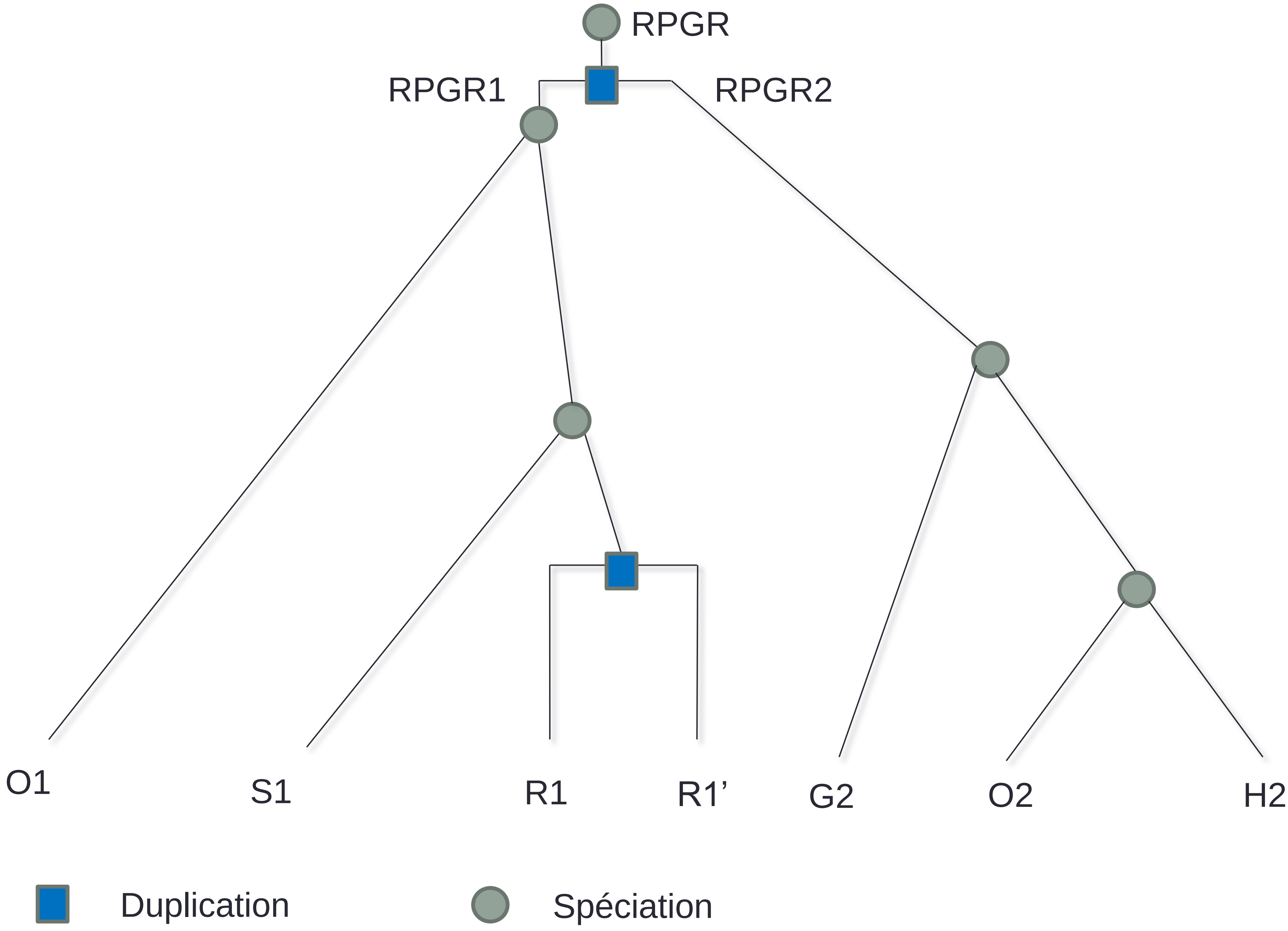

RPGR
Super-mammifère
RPGR1
RPGR2
Super-rongeur
Super-primate
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat


RPGR
Super-mammifère
RPGR1
RPGR2
Super-rongeur
Super-primate
Homme-outan
Gibbons
Orang-outan
Homme
Souris
Rat


RPGR
RPGR1
RPGR2
G2
O1
O2
H2
S1
R1
R1’
Duplication
Spéciation


RPGR
RPGR1
RPGR2
G2
O1
O2
H2
S1
R1
R1'
Duplication
Spéciation


RPGR
RPGR1
RPGR2
O1
S1
R1
R1'
G2
O2
H2
Duplication
Spéciation

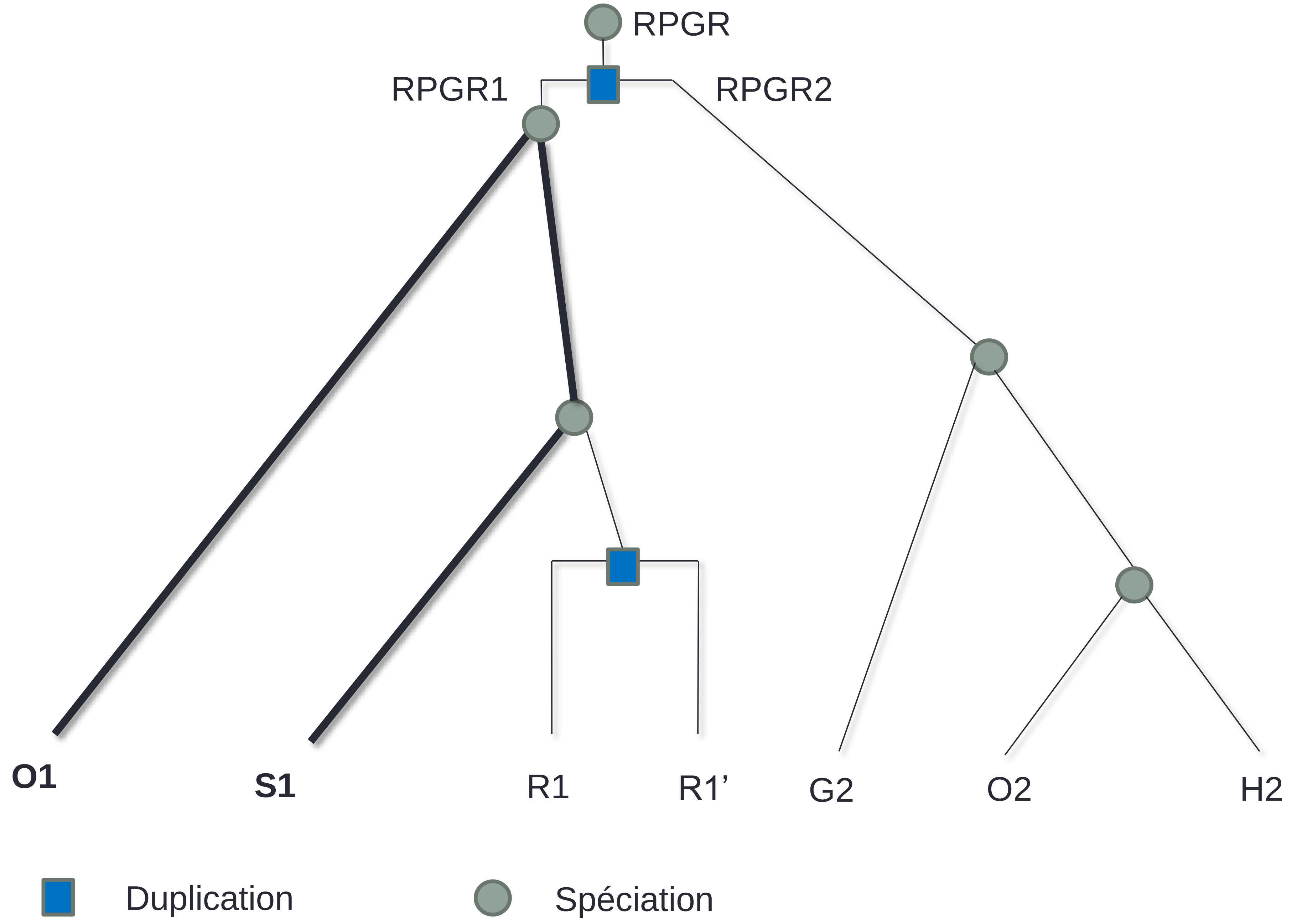

RPGR
RPGR1
RPGR2
O1
S1
R1
R1’
G2
O2
H2
Duplication
Spéciation

Orthologues et paralogues

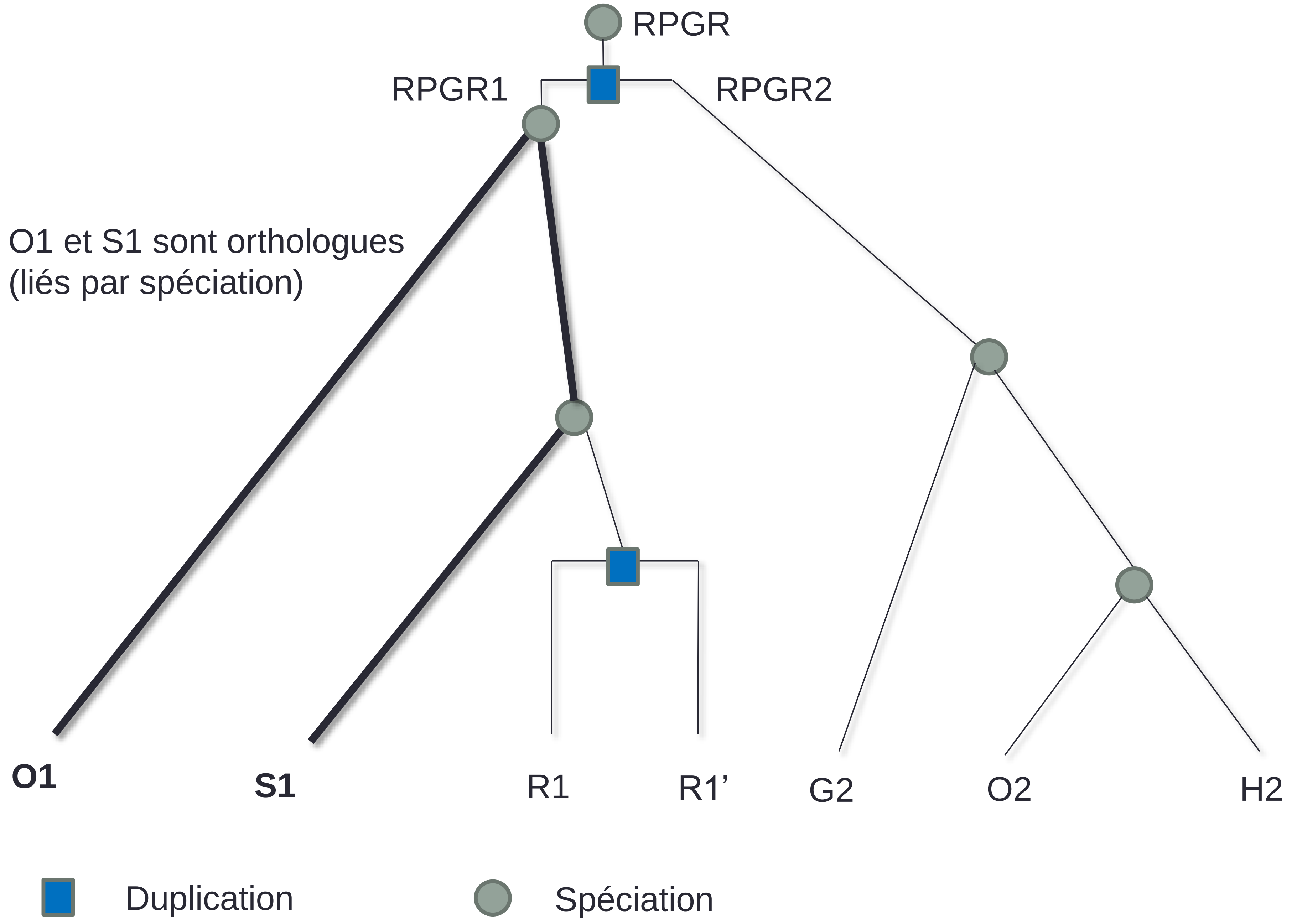
Deux gènes sont :

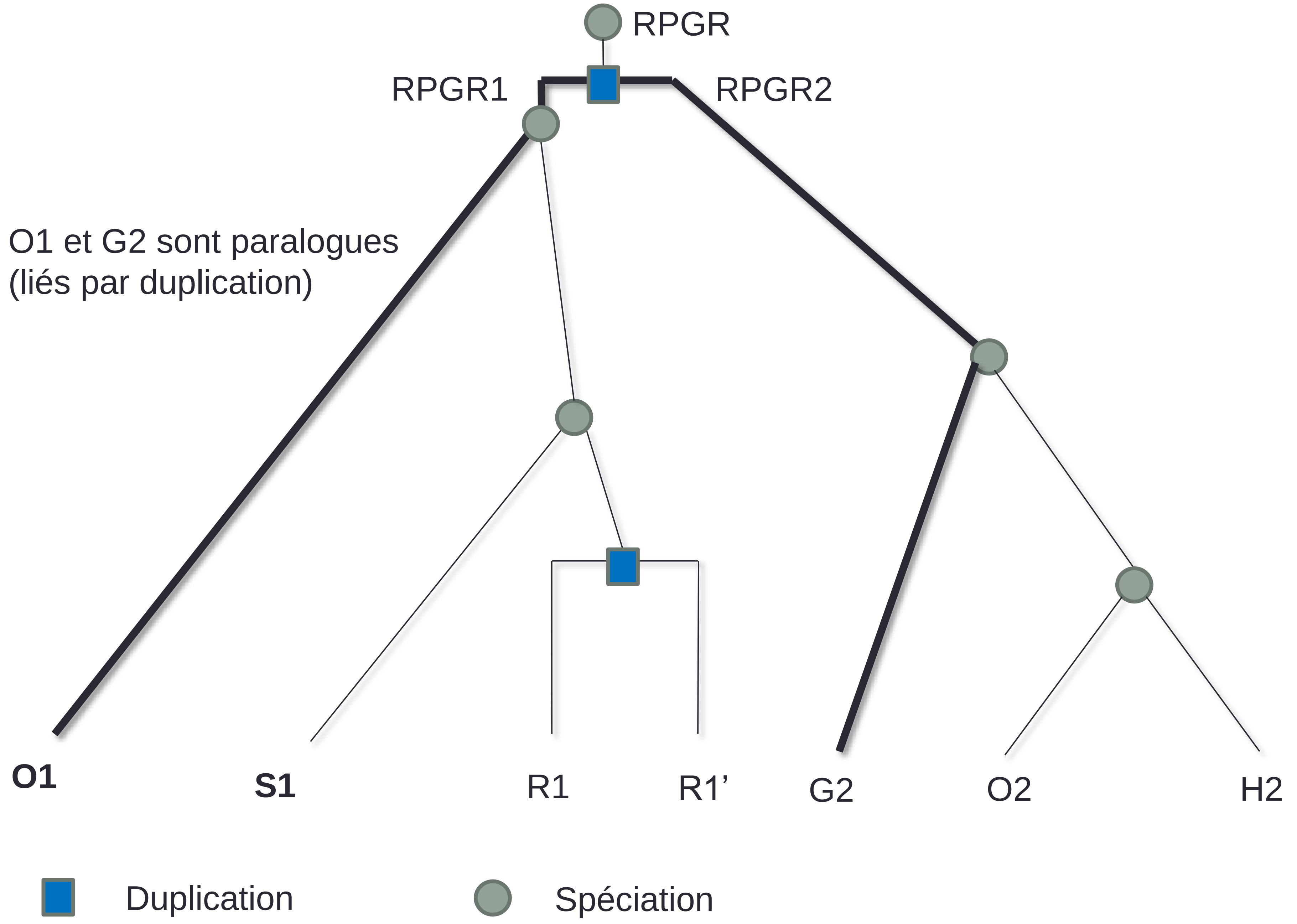
Orthologues si leur ancêtre commun a subi une **spéciation**

Paralogues si leur ancêtre commun a subi une **duplication**



RPGR
RPGR1
RPGR2
O1
S1
R1
R1'
G2
O2
H2
Duplication
Spéciation

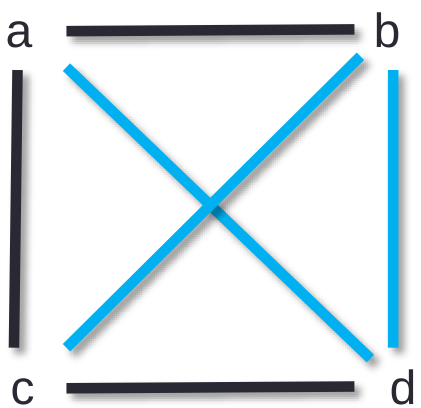

RPGR
RPGR1
RPGR2
O1 et S1 sont orthologues
(liés par spéciation)
O1
S1
R1
R1’
G2
O2
H2
Duplication
Spéciation


RPGR
RPGR1
RPGR2
O1 et G2 sont paralogues
(liés par duplication)
O1
S1
R1
R1’
G2
O2
H2
Duplication
Spéciation

Graphe d'orthologie/paralogie

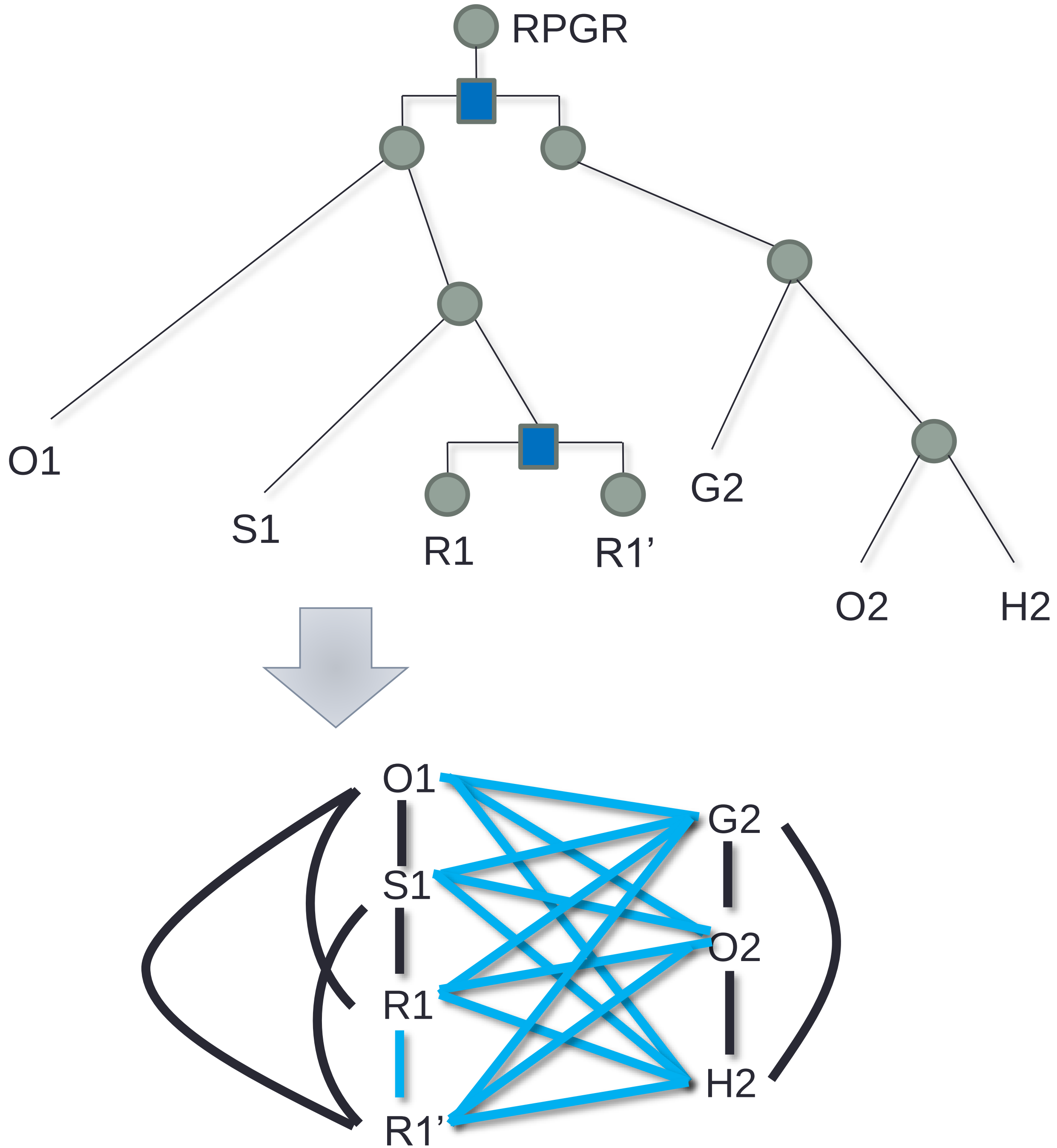
Orthologues = (a,b) (a, c) (c, d)

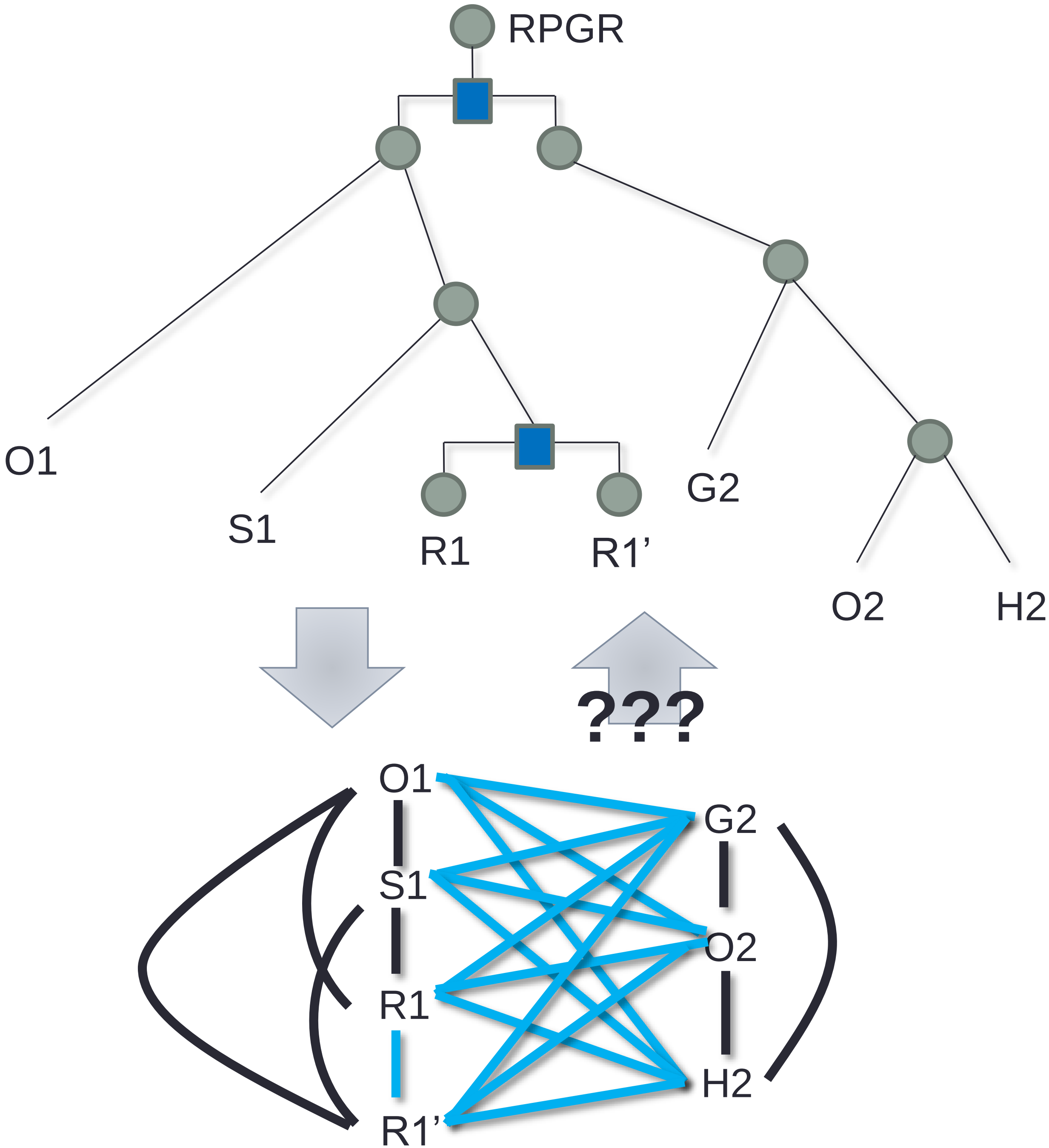
Paralogues = (a, d) (b, c) (b, d)

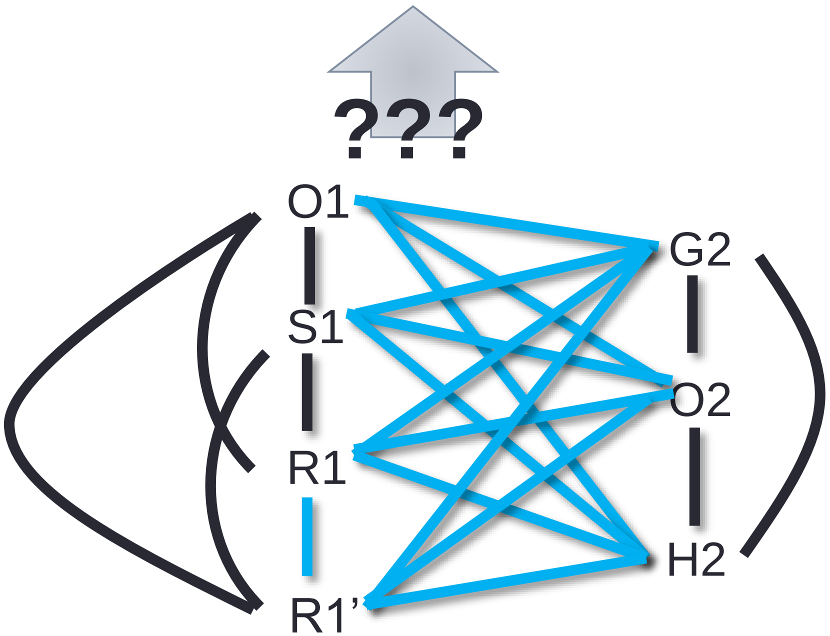


Orthologues

Paralogues


RPGR
O1
S1
R1
R1’
G2
O2
H2
O1
G2
S1
O2
R1
H2
R1’

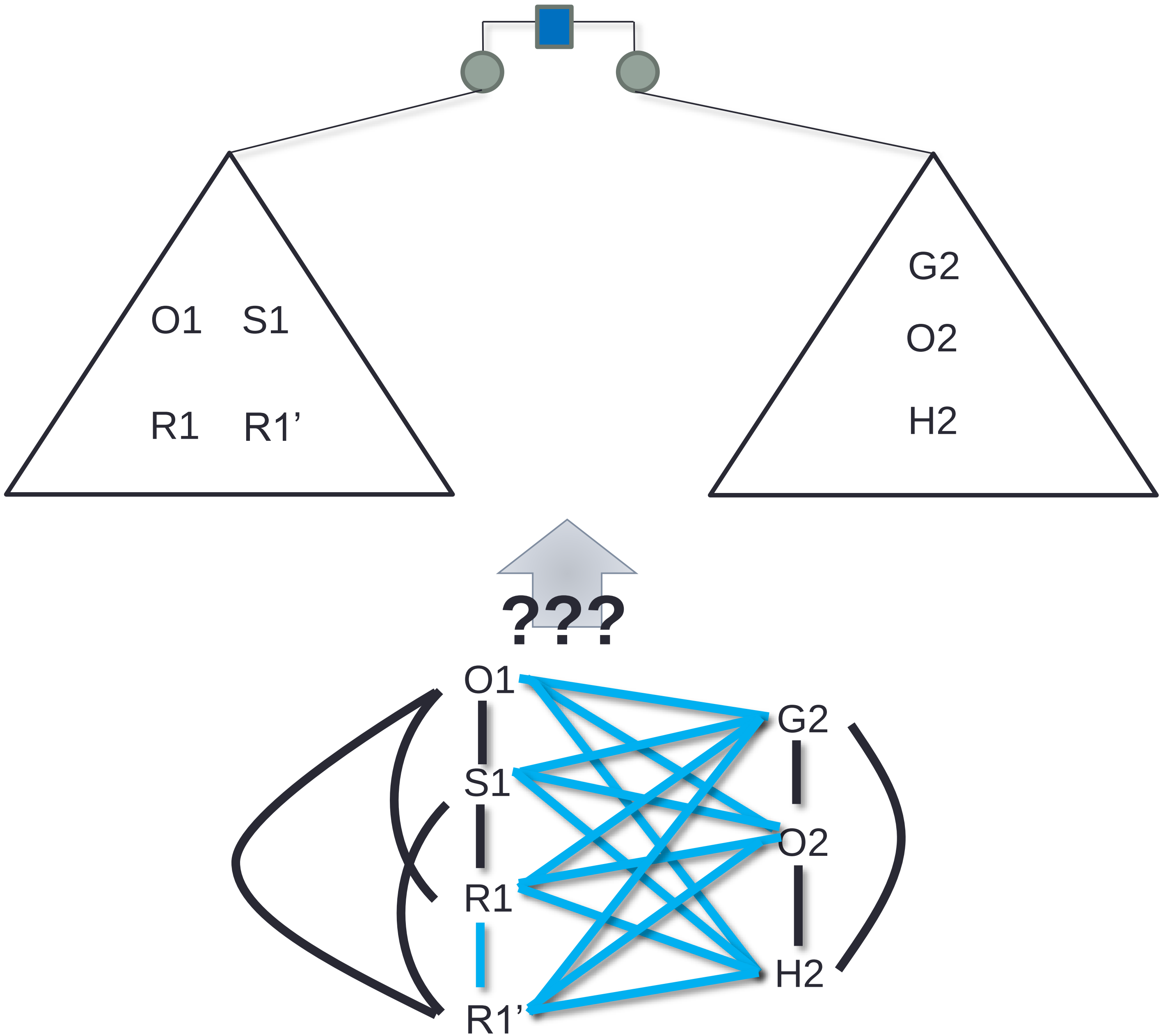

RPGR
O1
S1
R1
R1'
G2
O2
H2
???
O1
S1
R1
R1'
G2
O2
H2

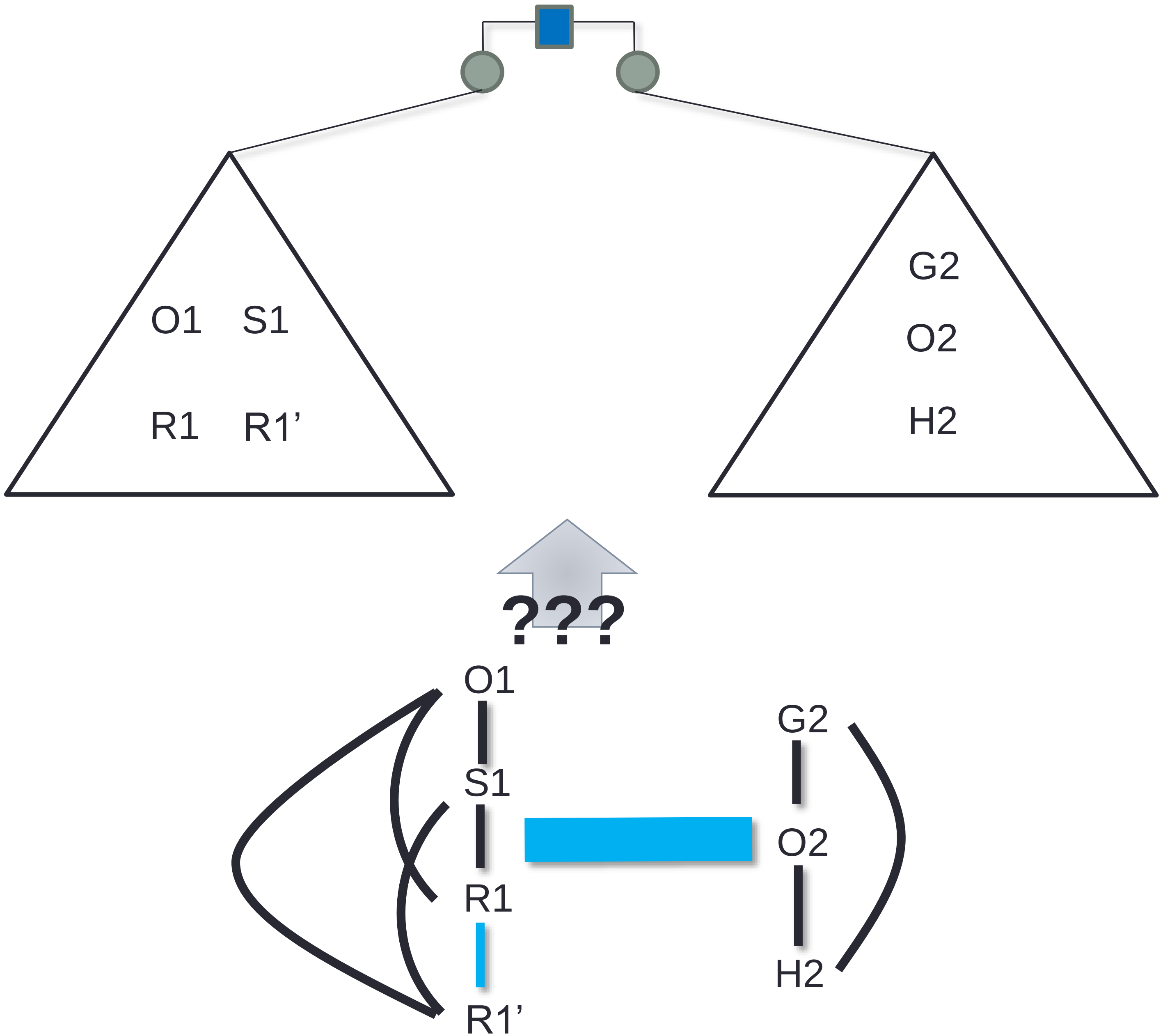


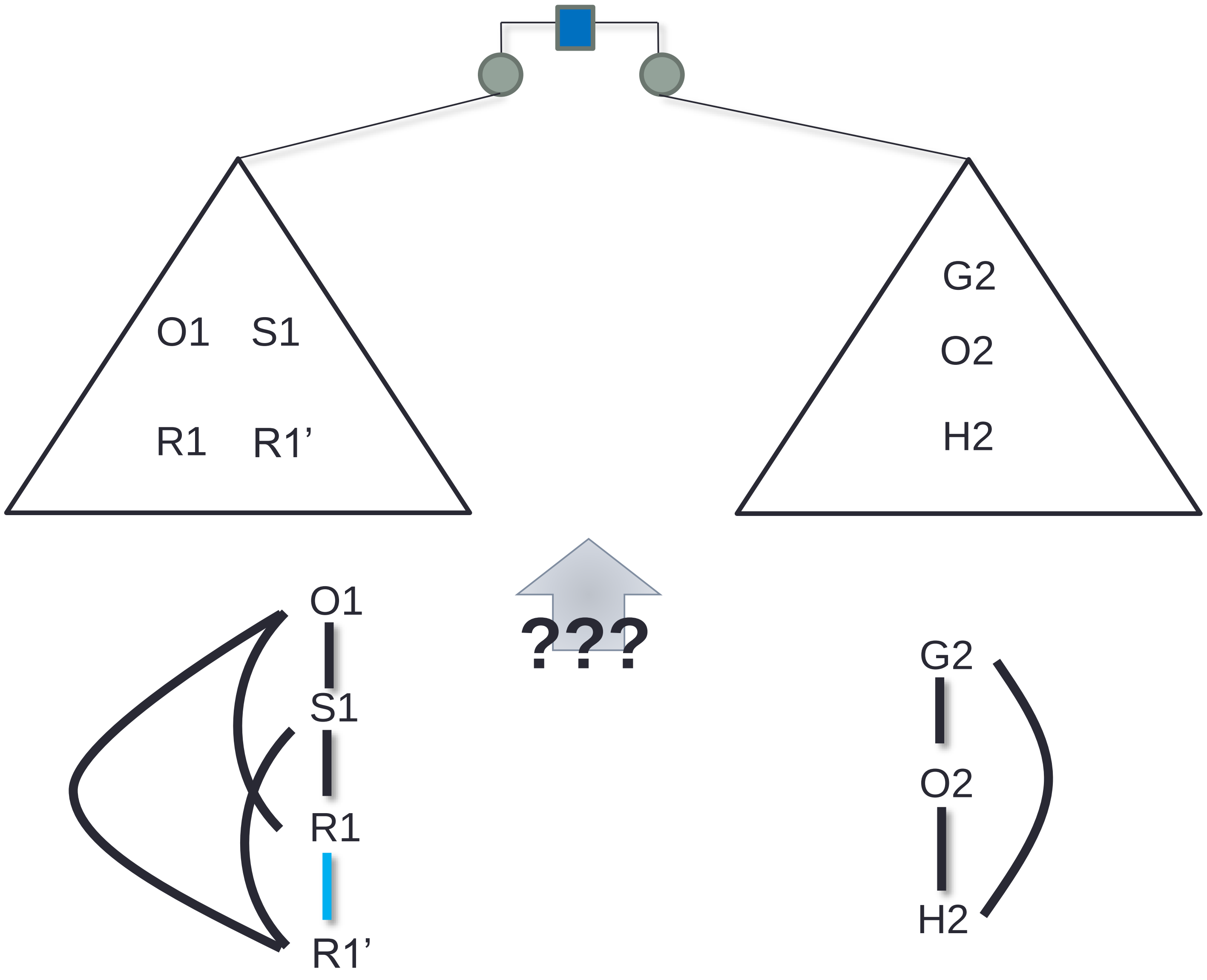
???
O1
S1
R1
R1'
G2
O2
H2

Problème 1 :

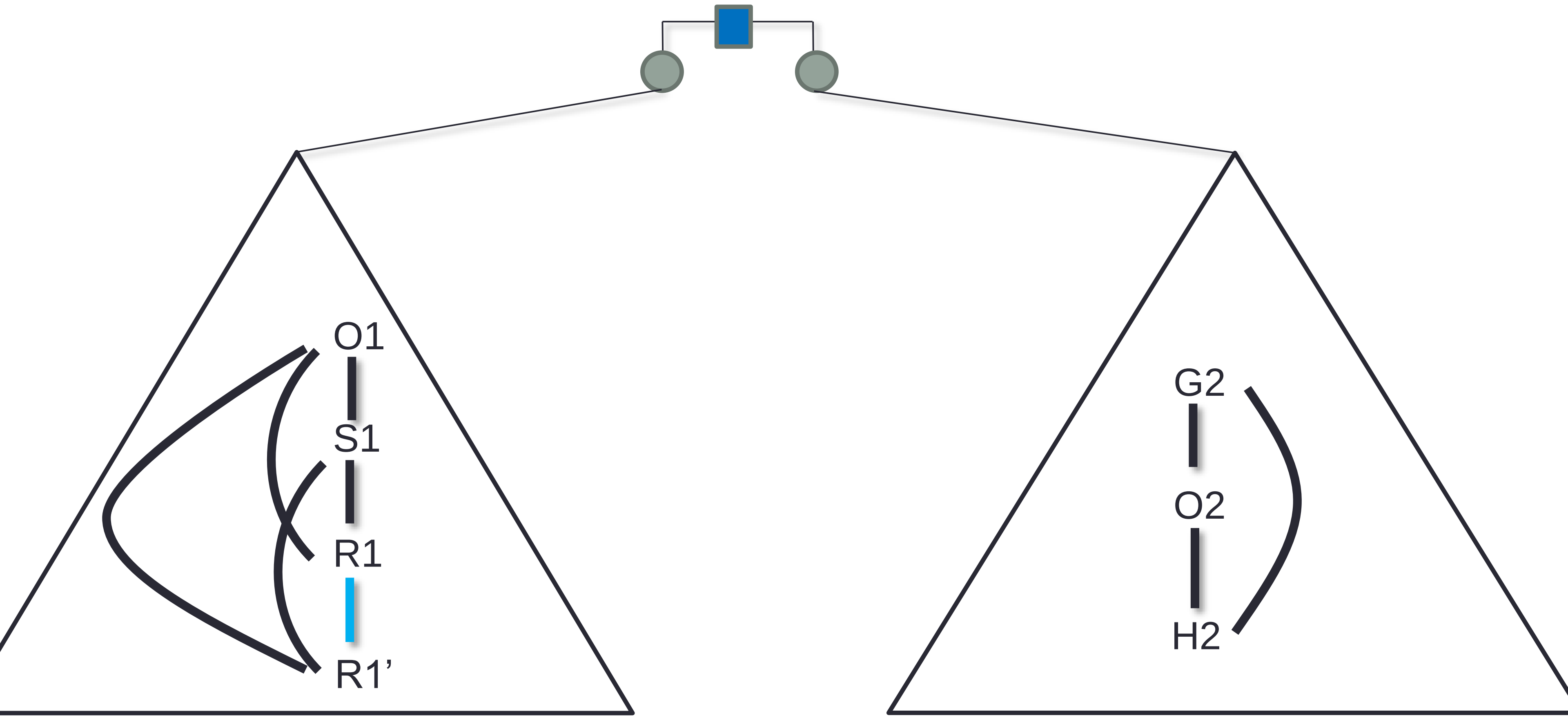
Étant donné un graphe de relations G est-ce qu’il existe un arbre qui satisfait ces relations ?

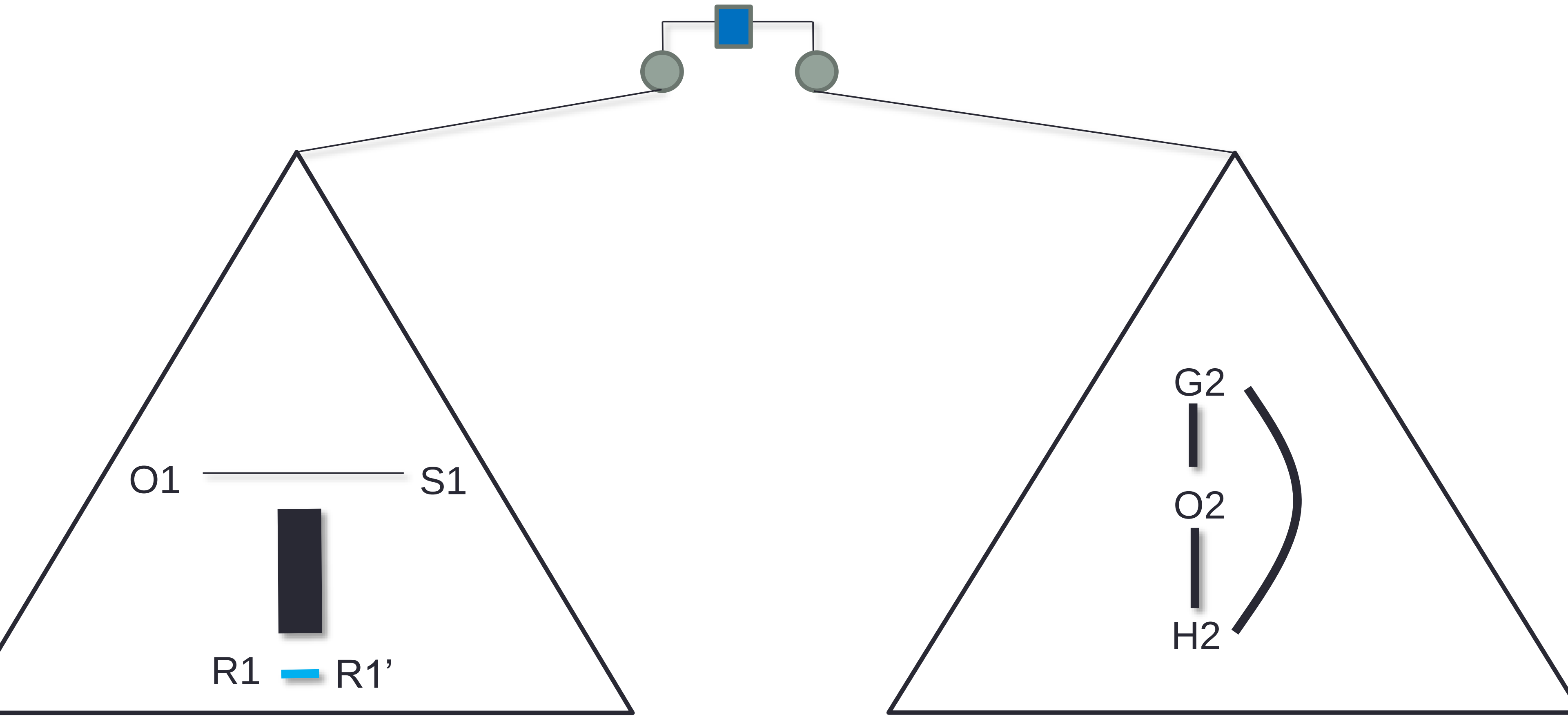

O1
S1
R1
R1'
G2
O2
H2
???
O1
S1
R1
R1'
G2
O2
H2

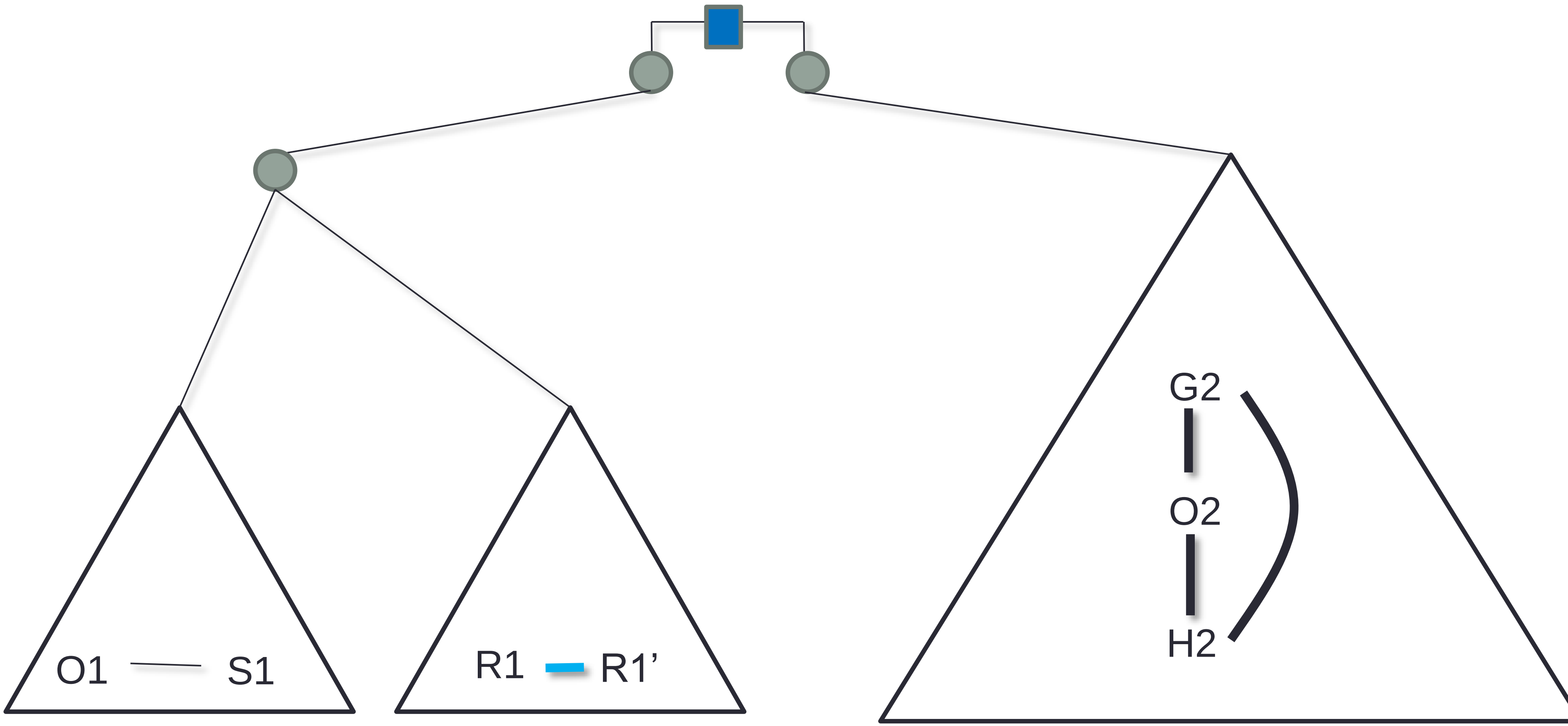

O1 S1
R1 R1'
G2
O2
H2
???
O1
S1
R1
R1'
G2
O2
H2



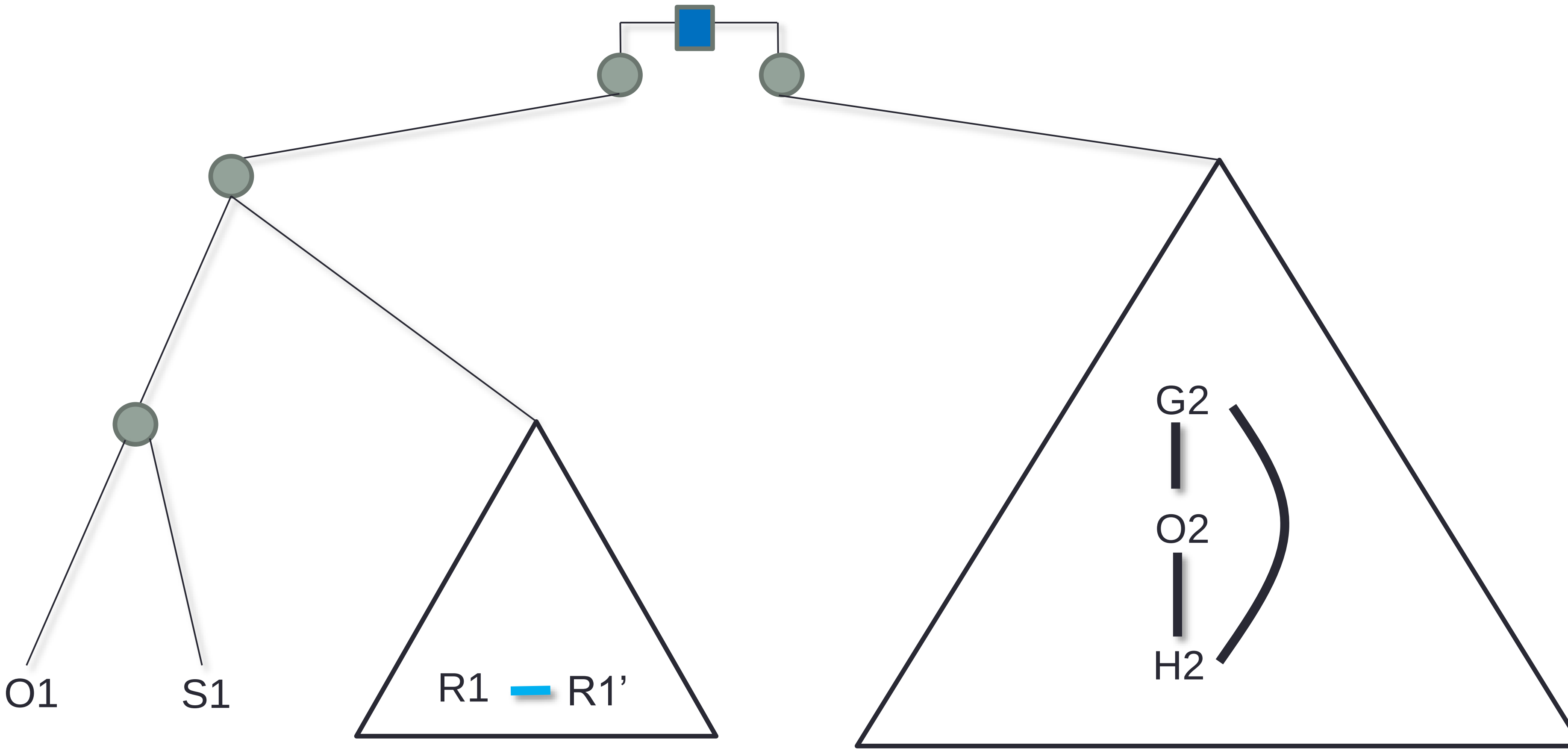
O1
S1
R1
R1'
G2
O2
H2
???
O1
S1
R1
R1'
G2
O2
H2


O1
S1
R1
R1'
G2
O2
H2

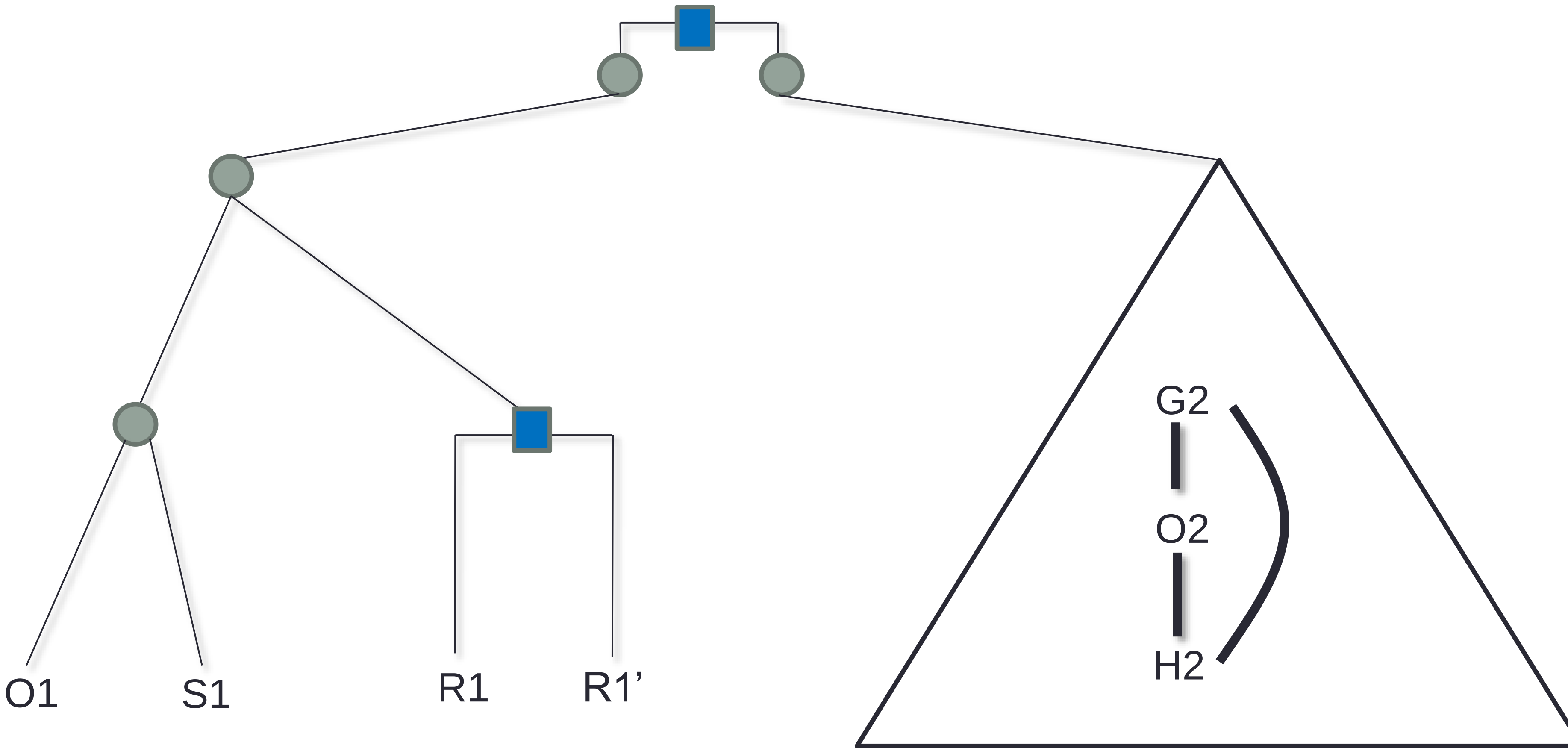

O1
S1
R1
R1’
G2
O2
H2



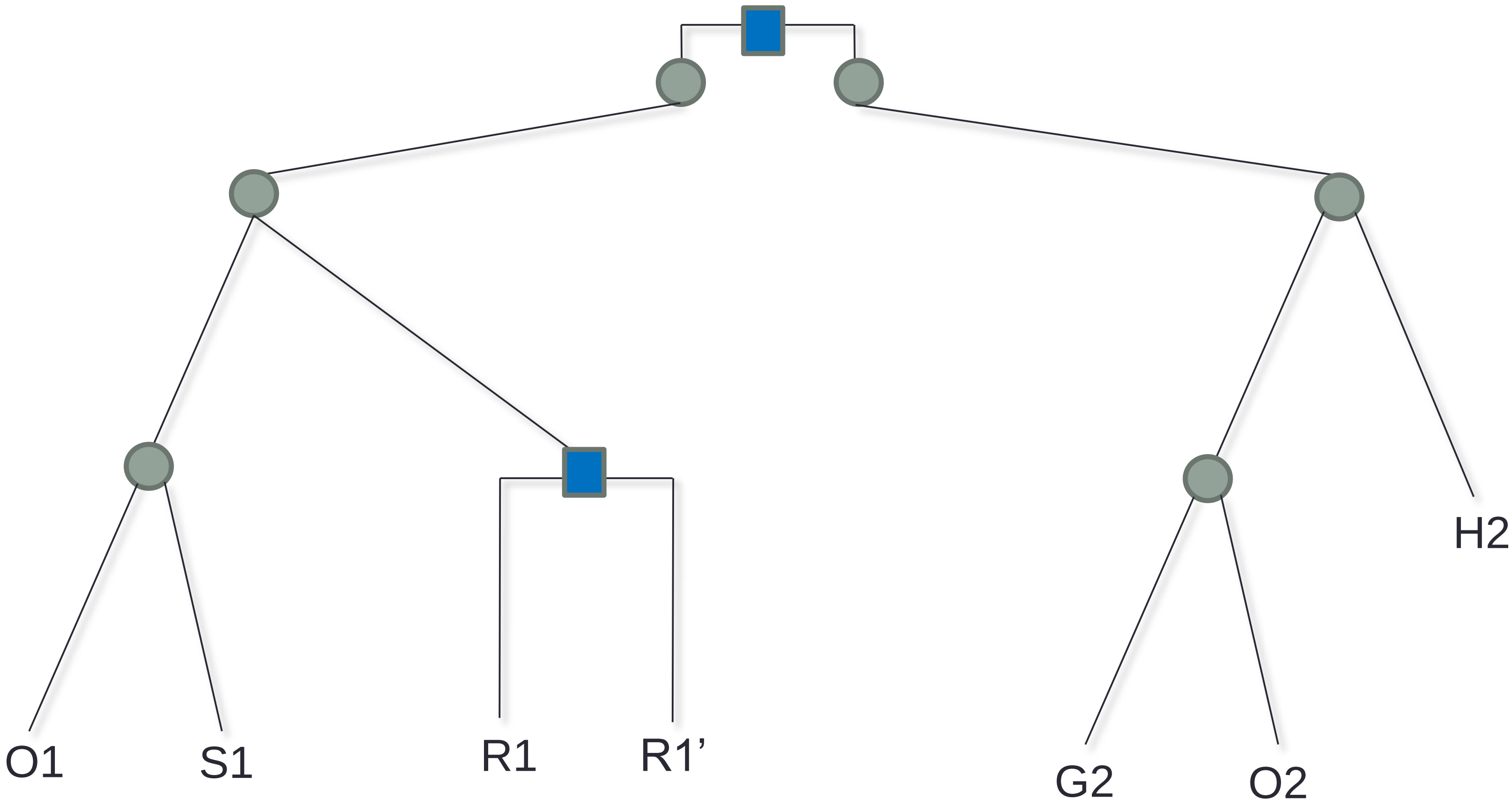
G2
O2
H2
O1
S1
R1
R1’



G2
O2
H2
O1
S1
R1
R1’



O1
S1
R1
R1’
G2
O2
H2


O1
S1
R1
R1’
G2
O2
H2

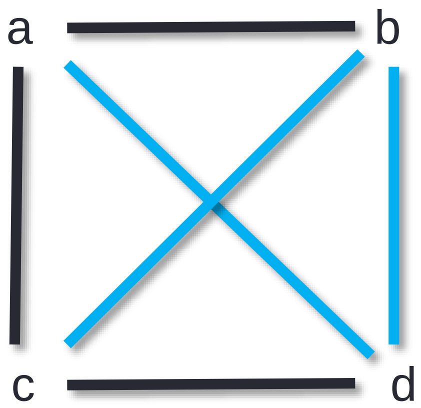
Lemme :

Si, à chaque étape, le graphe a une « séparation claire », alors on peut construire un arbre.

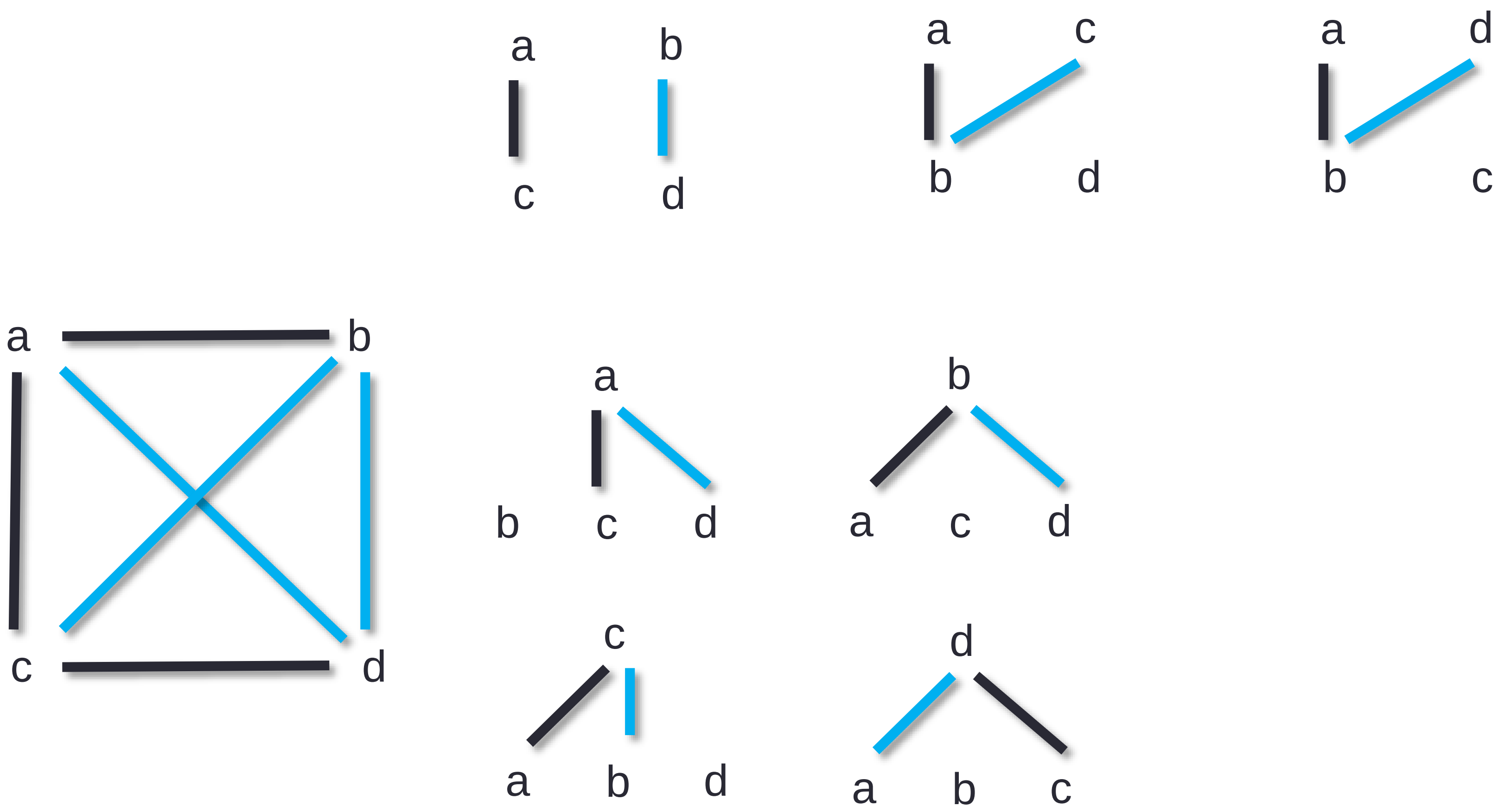
Et l’inverse ?

Si, à une certaine étape, le graphe n’a pas de « séparation claire », alors on ne peut pas contruire un arbre.

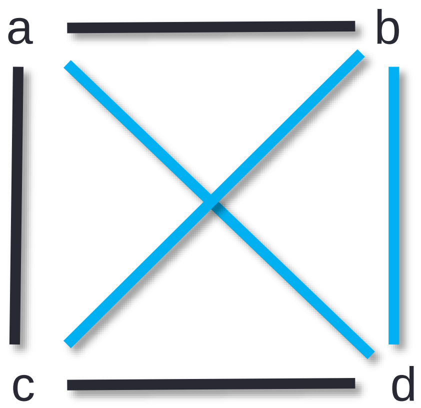
Pas de separation claire

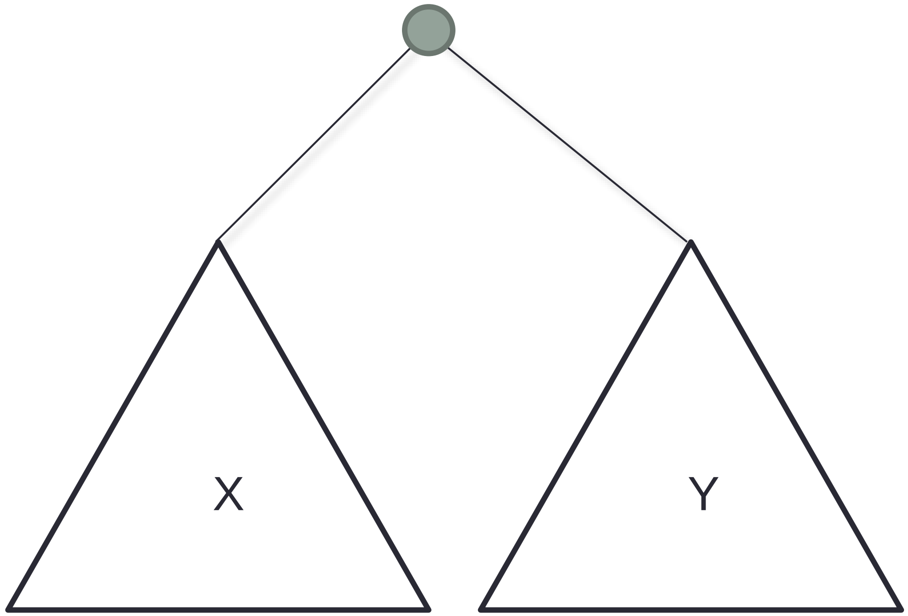


Pas de séparation claire

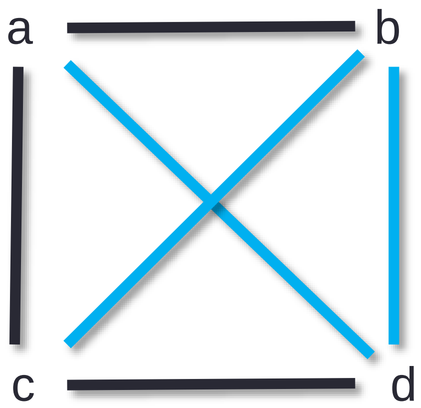


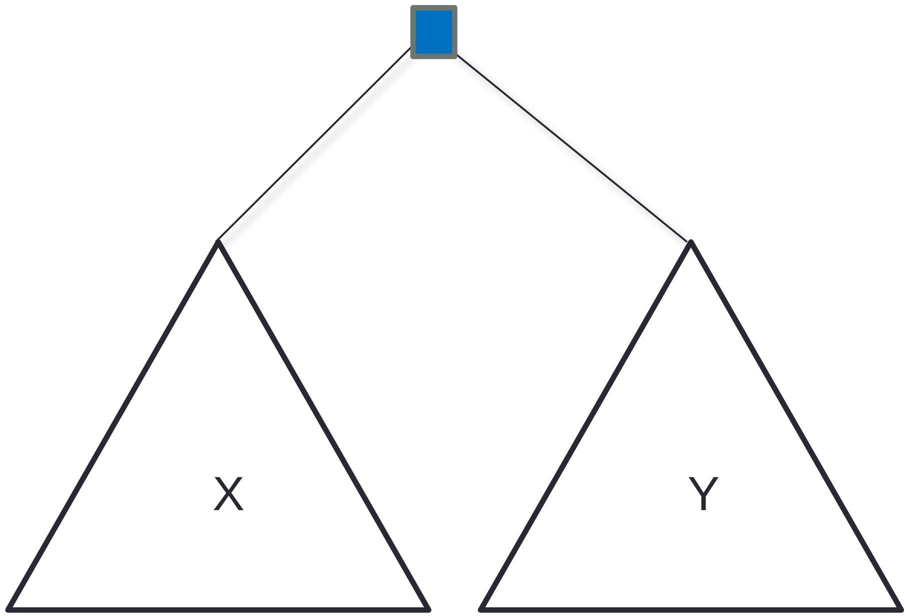
Pas de séparation claire



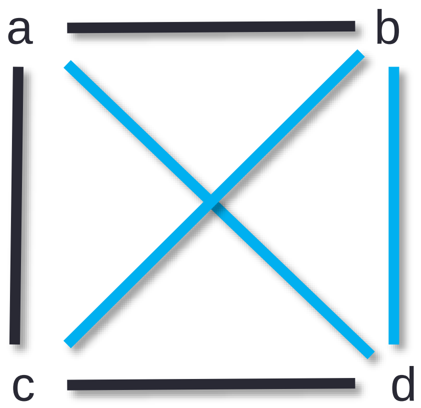


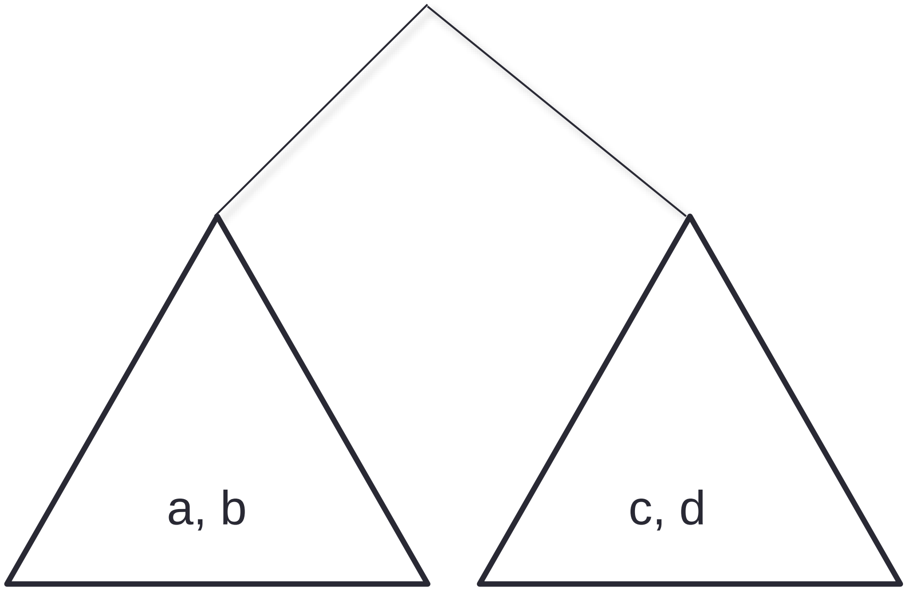
Pas de séparation claire



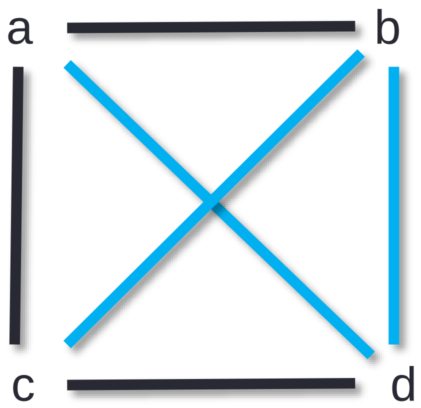


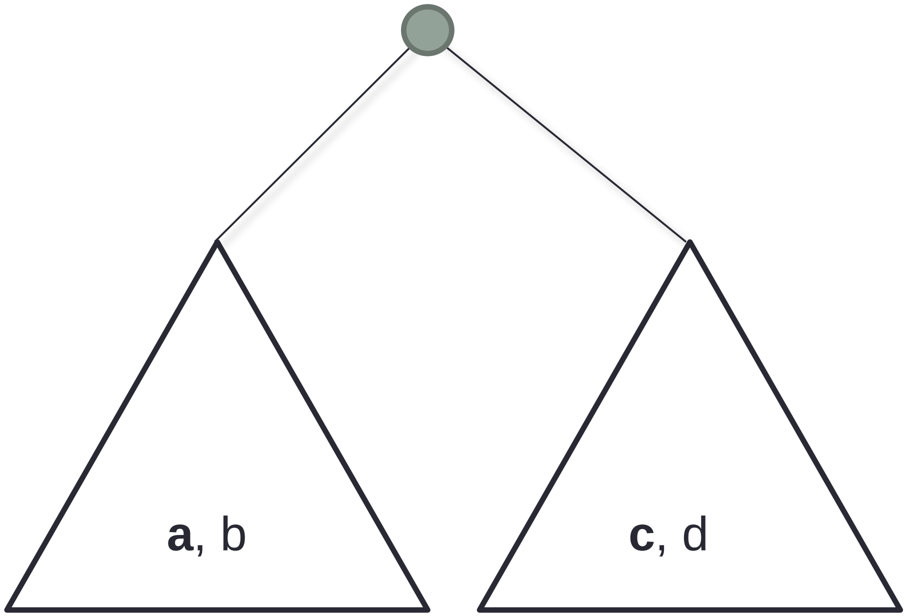
Pas de séparation claire



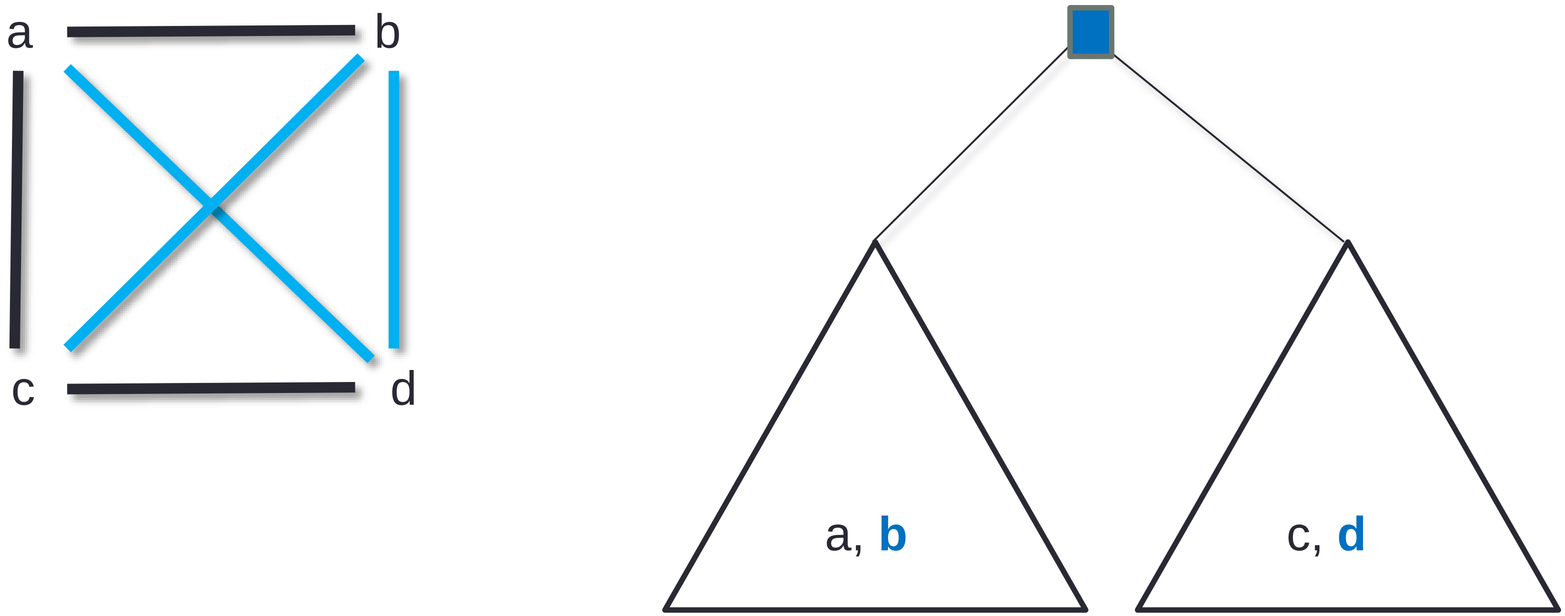


Pas de séparation claire



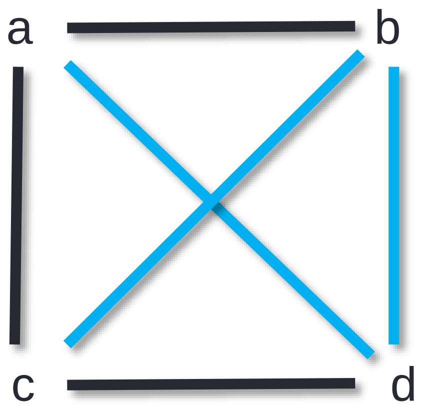


Pas de séparation claire



Lemme :

S’il y a un sous-graphe n’ayant pas de séparation claire, alors on ne peut pas construire l’arbre.



Théorème :

On peut construire un arbre à partir du graphe

si et seulement si

tout sous-graphe a une séparation claire

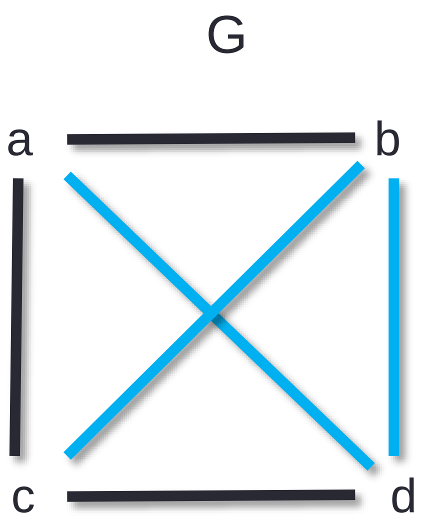
Théorème :

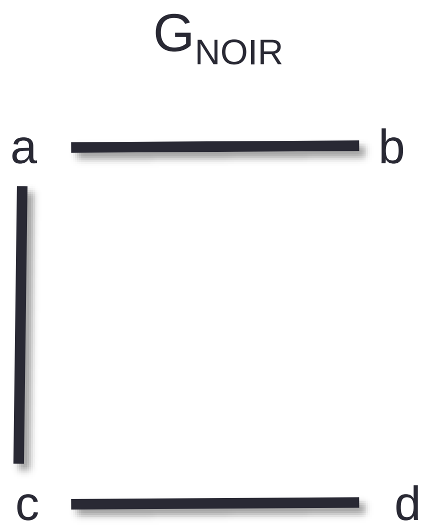
On peut construire un arbre à partir du graphe

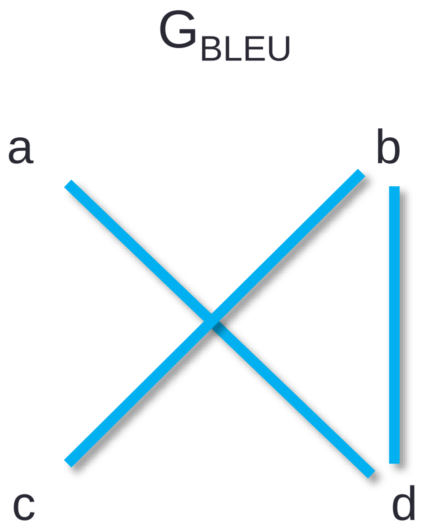
si et seulement si

tout sous-graphe a une séparation claire

Comment on vérifie ça, algorithmiquement ? Il y a un nombre exponentiel de sous-graphes.


G
a
b
c
d


G_{NOIR}
a
b
c
d


G_{BLEU}
a
b
c
d

Théorème :

On peut construire un arbre à partir du graphe

si et seulement si

tout sous-graphe a une séparation claire

Comment on vérifie ça, algorithmiquement ? Il y a un nombre exponentiel de sous-graphes.

Théorème :

Tout sous-graphe de G a une séparation claire

si et seulement si

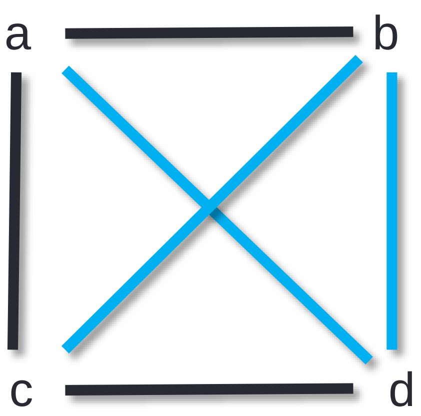
G_{NOIR} n'a aucune chemin induit de longueur 4

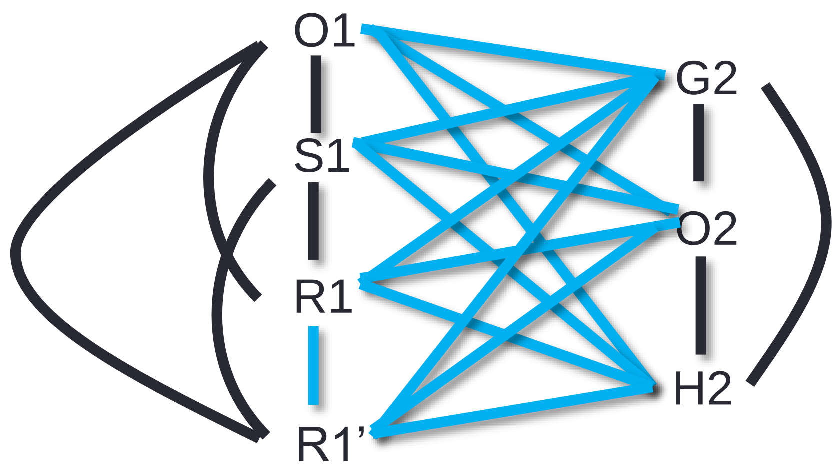
Théorème :

Tout sous-graphe de G a une séparation claire

si et seulement si

G_{NOIR} n'a aucune chemin induit de longueur 4



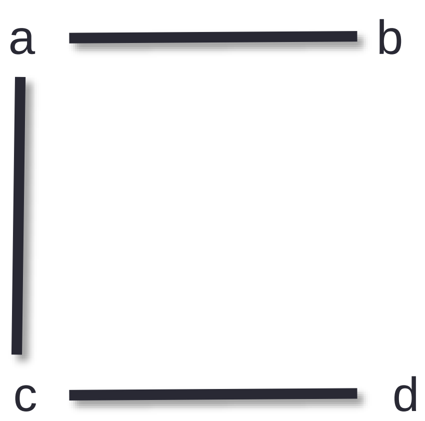


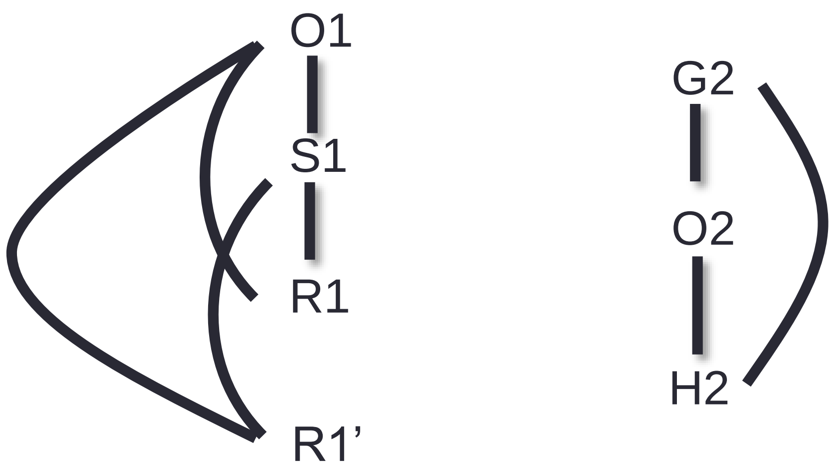
Théorème :

Tout sous-graphe de G a une séparation claire

si et seulement si

G_{NOIR} n'a aucune chemin induit de longueur 4





Problème 1 :

Étant donné un graphe de relations G est-ce qu’il existe un arbre qui satisfait ces relations ?

Problème 1 :

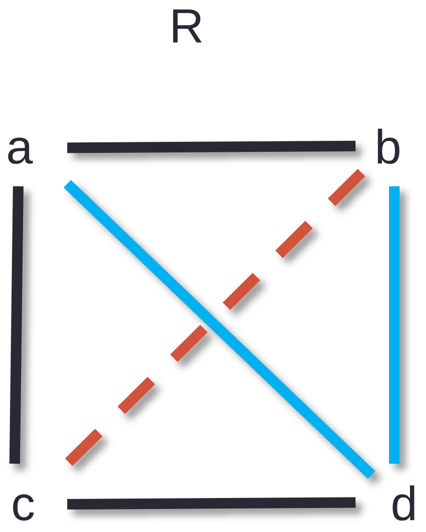
Étant donné un graphe de relations G est-ce qu’il existe un arbre qui satisfait ces relations ?

Problème 2 :

Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu’un tel arbre existe ?

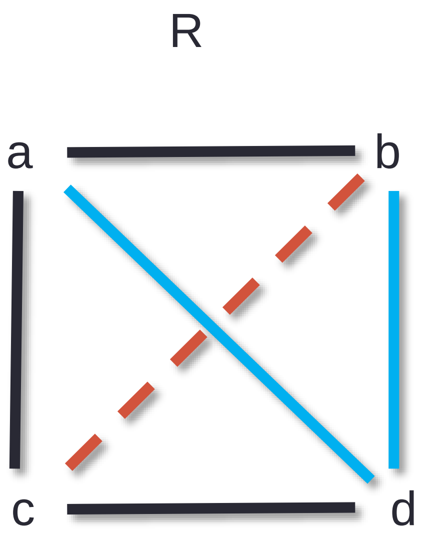
Problème 2 :

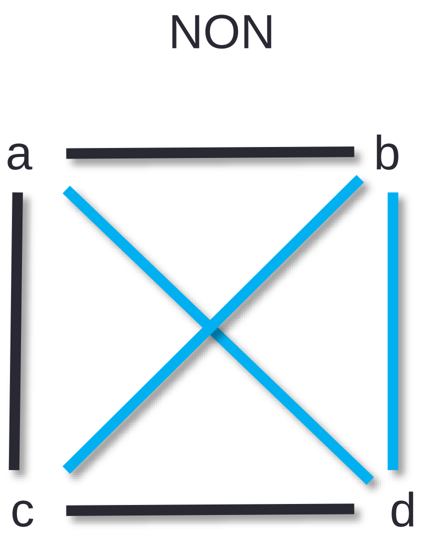
Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu'un tel arbre existe ?



Problème 2 :

Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu'un tel arbre existe ?

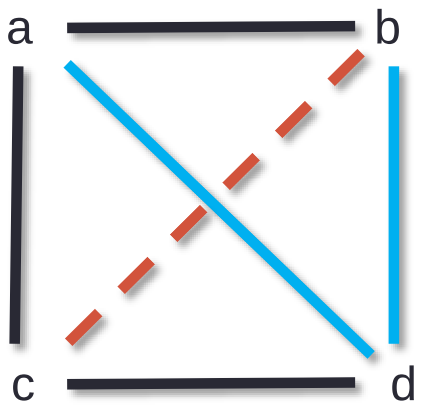




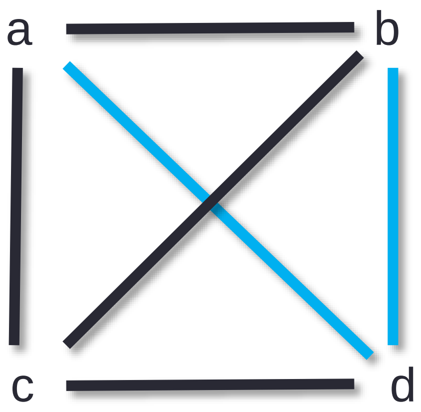
Problème 2 :

Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu'un tel arbre existe ?

R



OUI

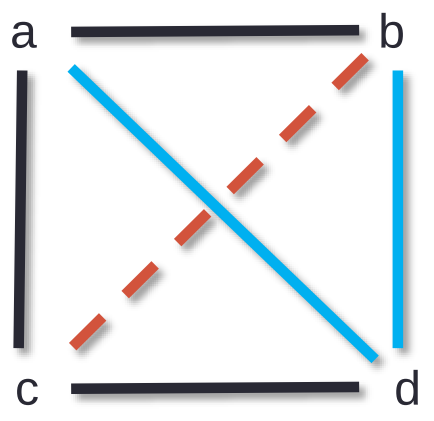


Problème 2 :

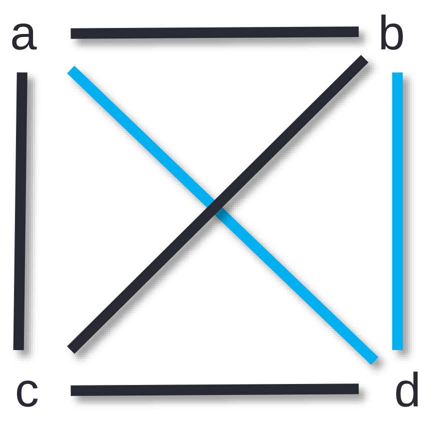
Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu’un tel arbre existe ?

C’est-à-dire, peut-on choisir une **coloration noir/bleu** pour les arêtes oranges telle que le graphe noir d’a **aucun chemin induit de longueur 4** ?

R



OUI



Problème 2 :

Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu’un tel arbre existe ?

C’est-à-dire, peut-on choisir une **coloration noir/bleu** pour les arêtes oranges telle que le graphe noir d’a **aucun chemin induit de longueur 4** ?

Il existe un algorithme pouvant faire ça en temps polynomial.

Problème 1 :

Étant donné un graphe de relations G est-ce qu’il existe un arbre qui satisfait ces relations ?

Problème 2 :

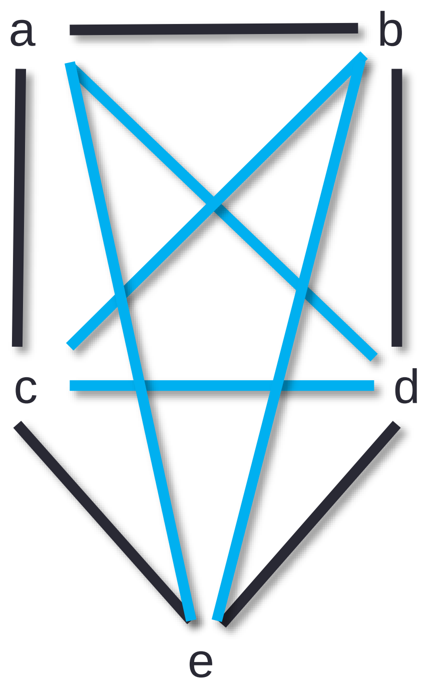
Et si certaines des relations de G sont inconnues (ou non-décidées), peut-on décider ces relations afin qu’un tel arbre existe ?

Problème 3 :

Et si G n’admet pas d’arbre qui satisfait ses relations, quel est le minimum de modifications qu’on doit apporter à G pour que ce soit le cas ?

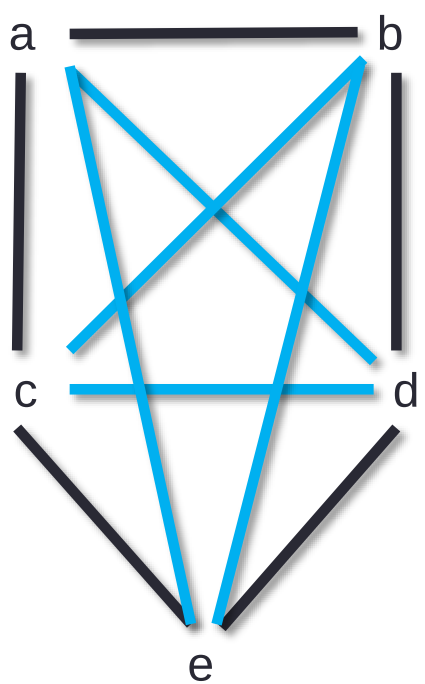
Problème 3 :

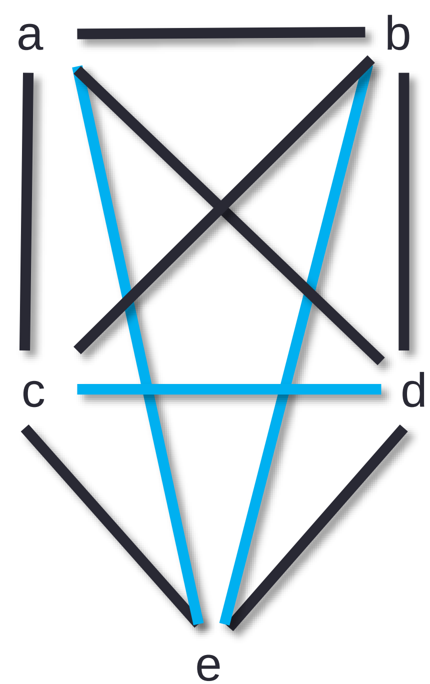
Et si G n’admet pas d’arbre qui satisfait ses relations, quel est le minimum de modifications qu’on doit apporter à G pour que ce soit le cas ?



Problème 3 :

Et si G n'admet pas d'arbre qui satisfait ses relations, quel est le minimum de modifications qu'on doit apporter à G pour que ce soit le cas ?

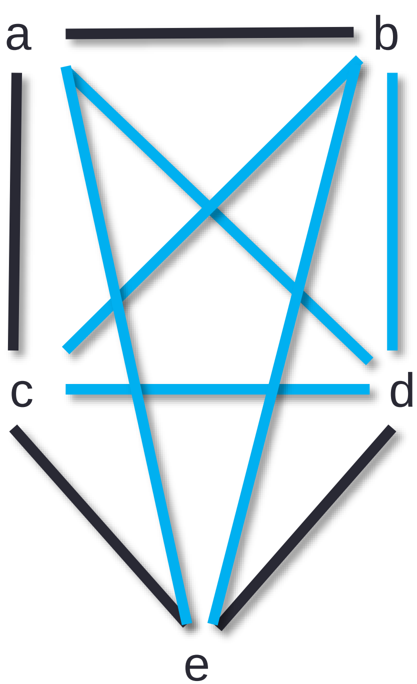


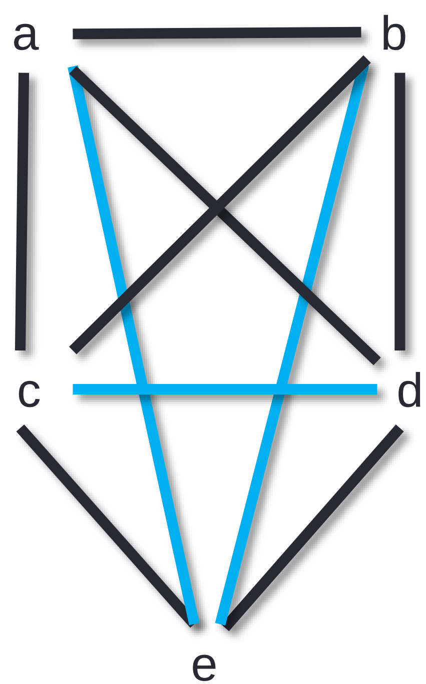


3 modifs

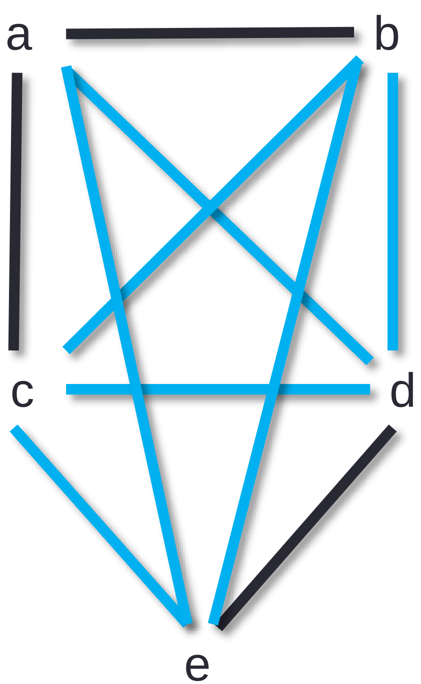
Problème 3 :

Et si G n'admet pas d'arbre qui satisfait ses relations, quel est le minimum de modifications qu'on doit apporter à G pour que ce soit le cas ?





3 modifs



1 modif

Problème 3 :

Et si G n’admet pas d’arbre qui satisfait ses relations, quel est le minimum de modifications qu’on doit apporter à G pour que ce soit le cas ?

NP-COMPLET !

Peut-on approximer une solution optimale (avec des garanties) ?

On ne sait pas…

Conclusion

Bio => Algo => Graphes

Mais il y a beaucoup d’autres problématiques.

Bio => Algo => (Prob & Stats | Groupes | Info théorique | …)