



Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique

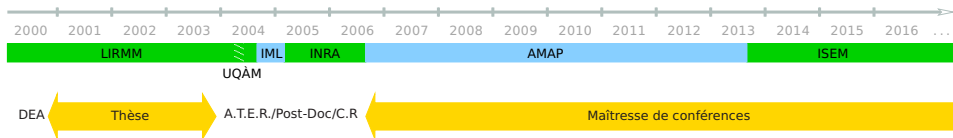
Sèverine Bérard

4 novembre 2016

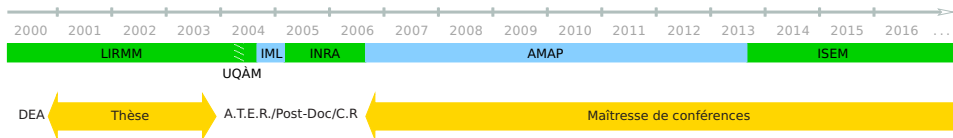
- 1 Mon parcours
- 2 Domaine de recherche
- 3 Reconstruction de génomes ancestraux : DeCo
- 4 Synthèse des contributions
- 5 Projets de recherche
- 6 Merci

- 1 Mon parcours
- 2 Domaine de recherche
- 3 Reconstruction de génomes ancestraux : DeCo
- 4 Synthèse des contributions
- 5 Projets de recherche
- 6 Merci

Déroulement carrière

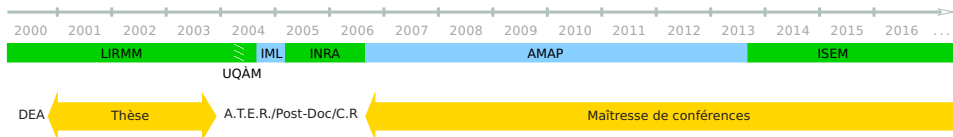


Déroulement carrière



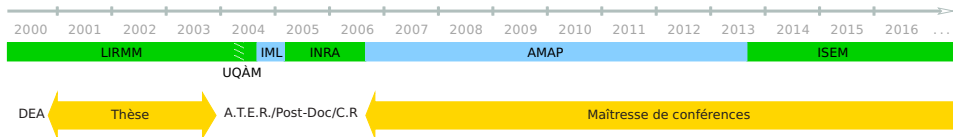
- ▶ Doctorat en informatique “*Comparaison de séquences répétées en tandem et application à la génétique*” (déc. 2003)

Déroulement carrière



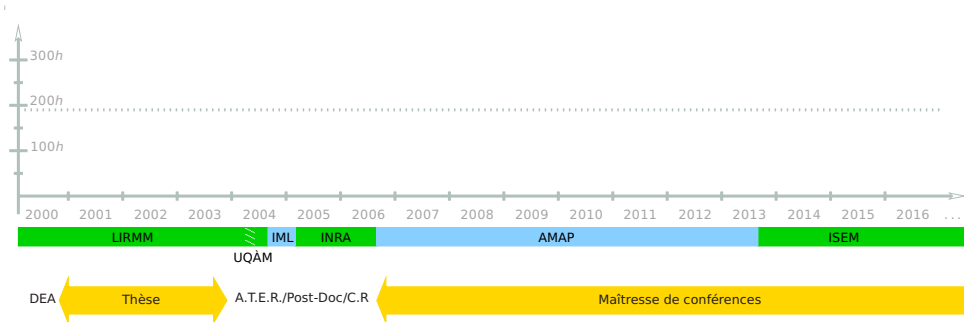
- ▶ Doctorat en informatique “*Comparaison de séquences répétées en tandem et application à la génétique*” (déc. 2003)
- ▶ Chargée de recherche au dépt. MIA de l’INRA de Toulouse (2005-2006)

Déroulement carrière



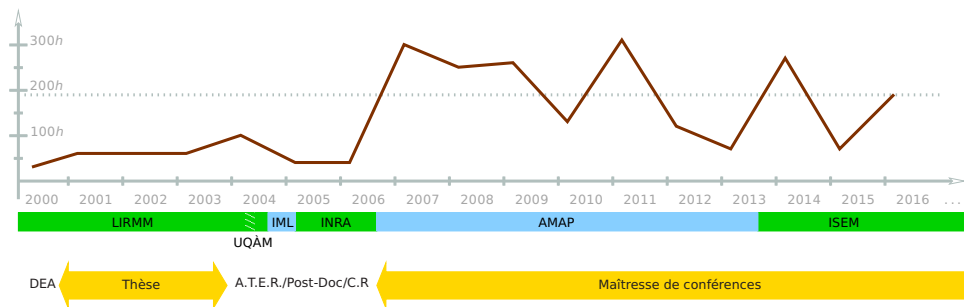
- ▶ Doctorat en informatique “*Comparaison de séquences répétées en tandem et application à la génétique*” (déc. 2003)
- ▶ Chargée de recherche au dépt. MIA de l’INRA de Toulouse (2005-2006)
- ▶ Maîtresse de conférences en informatique à l’UM dans des laboratoires de biologie depuis septembre 2006

Déroulement carrière



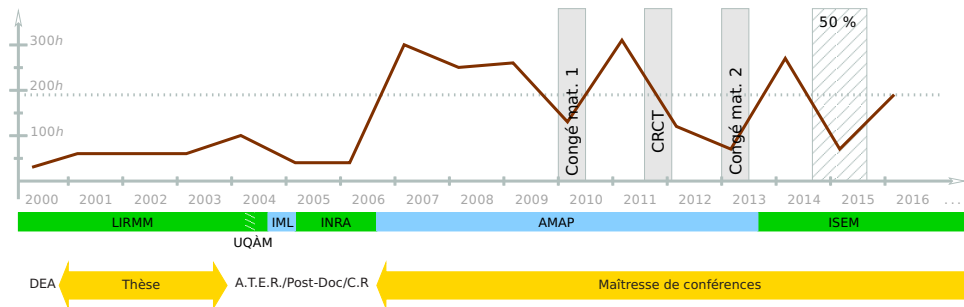
- ▶ Doctorat en informatique “*Comparaison de séquences répétées en tandem et application à la génétique*” (déc. 2003)
- ▶ Chargée de recherche au dépt. MIA de l’INRA de Toulouse (2005-2006)
- ▶ Maîtresse de conférences en informatique à l’UM dans des laboratoires de biologie depuis septembre 2006

Déroulement carrière



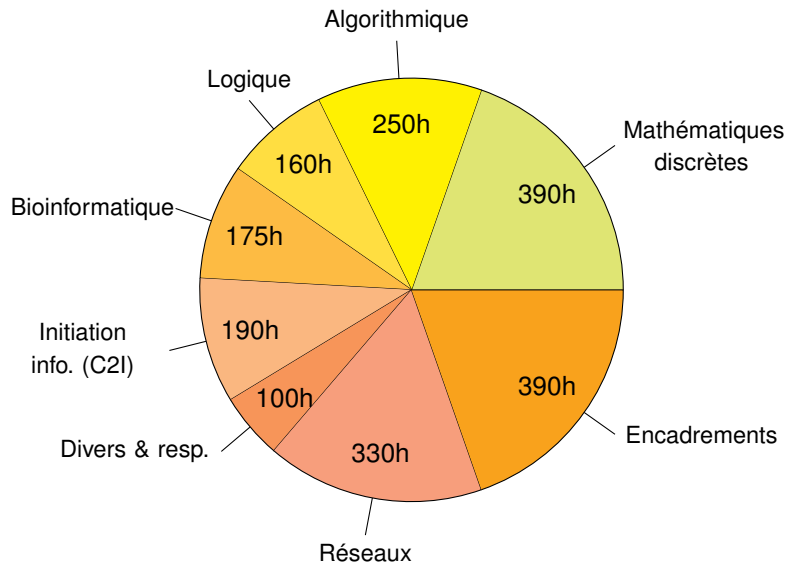
- ▶ Doctorat en informatique “*Comparaison de séquences répétées en tandem et application à la génétique*” (déc. 2003)
- ▶ Chargée de recherche au dépt. MIA de l’INRA de Toulouse (2005-2006)
- ▶ Maîtresse de conférences en informatique à l’UM dans des laboratoires de biologie depuis septembre 2006

Déroulement carrière



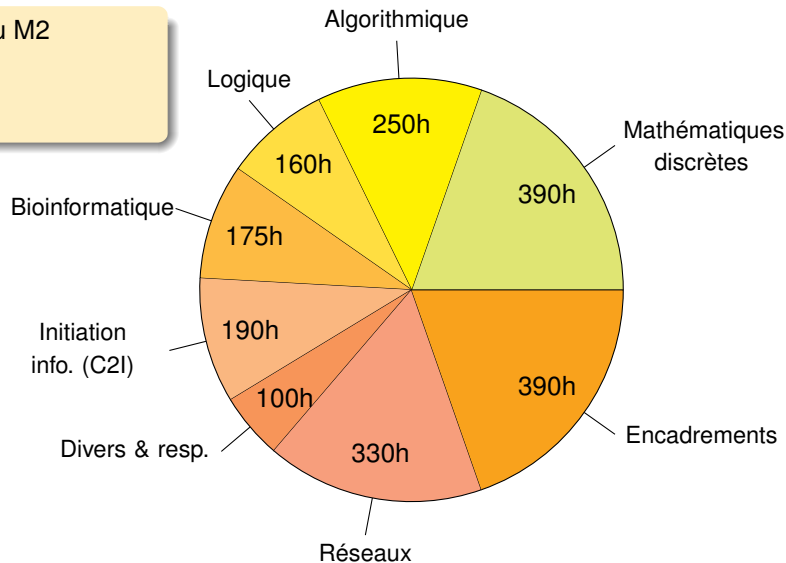
- ▶ Doctorat en informatique “*Comparaison de séquences répétées en tandem et application à la génétique*” (déc. 2003)
- ▶ Chargée de recherche au dépt. MIA de l’INRA de Toulouse (2005-2006)
- ▶ Maîtresse de conférences en informatique à l’UM dans des laboratoires de biologie depuis septembre 2006

Enseignements : 2000h en 10 ans



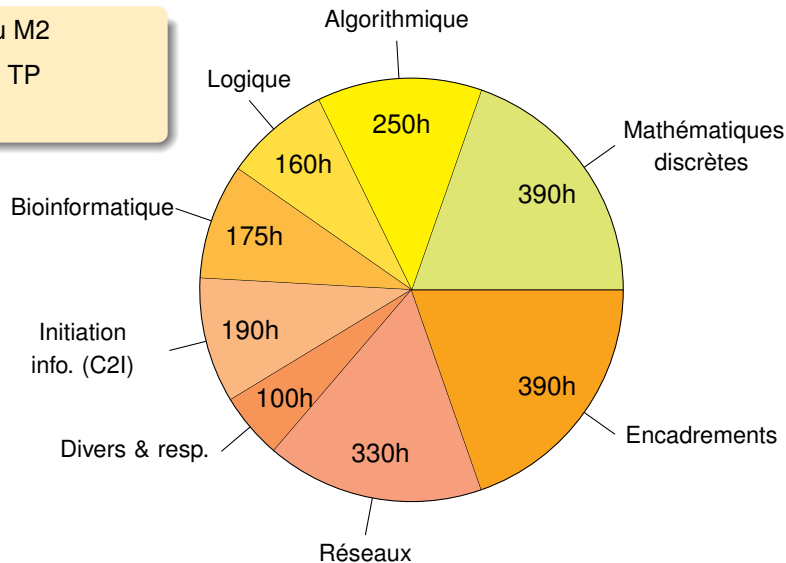
Enseignements : 2000h en 10 ans

► du L1 au M2



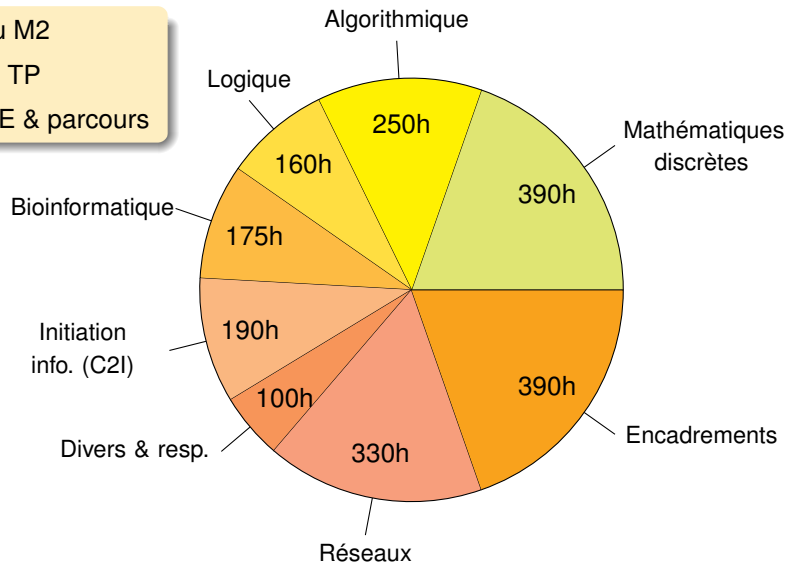
Enseignements : 2000h en 10 ans

- ▶ du L1 au M2
- ▶ CM, TD, TP



Enseignements : 2000h en 10 ans

- ▶ du L1 au M2
- ▶ CM, TD, TP
- ▶ Resp. UE & parcours



Groupe de Travail en Génomique Comparative (GTGC)



Groupe de Travail en Génomique Comparative (GTGC)



Animation scientifique de la communauté francophone

Groupe de Travail en Génomique Comparative (GTGC)



Animation scientifique de la communauté francophone

- ▶ Créé en 2005, co-animé avec Éric Tannier

Groupe de Travail en Génomique Comparative (GTGC)



Animation scientifique de la communauté francophone

- ▶ Créé en 2005, co-animé avec Éric Tannier
- ▶ Différents formats, partenaires, villes. En moyenne 50 participant(e)s



Groupe de Travail en Génomique Comparative (GTGC)



Animation scientifique de la communauté francophone

- ▶ Créé en 2005, co-animé avec Éric Tannier
- ▶ Différents formats, partenaires, villes. En moyenne 50 participant(e)s



- 1 Mon parcours
- 2 Domaine de recherche**
- 3 Reconstruction de génomes ancestraux : DeCo
- 4 Synthèse des contributions
- 5 Projets de recherche
- 6 Merci

Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique

Algorithme : Expliquer le titre de mon HDR

Données : Un titre

Résultat : Un diaporama d'explication

- 1 Choisir des *mots clés* du titre ;
- 2 **pour chaque** *mot clé* **faire**
- 3 | Choisir *images* ;
- 4 | Choisir *textes* ;
- 5 | *Diaporama* \leftarrow *images* \cup *textes* ;
- 6 **si** *satisfaisant* **alors**
- 7 | Retourner *Diaporama* ;
- 8 **sinon**
- 8 | Revenir ligne 1 ;

Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique

Algorithme : Expliquer le titre de mon HDR

Données : Un titre

Résultat : Un diaporama d'explication

- 1 Choisir des *mots clés* du titre ;
 - 2 **pour chaque mot clé faire**
 - 3 | Choisir *images* ;
 - 4 | Choisir *textes* ;
 - 5 | *Diaporama* \leftarrow *images* \cup *textes* ;
 - 6 **si satisfaisant alors**
 - 7 | Retourner *Diaporama* ;
 - 8 **sinon**
 - 9 | Revenir ligne 1 ;
-

► Un *algorithme* donne une/la solution à un *problème*

Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique

Algorithme : Expliquer le titre de mon HDR

Données : Un titre

Résultat : Un diaporama d'explication

- 1 Choisir des *mots clés* du titre ;
 - 2 **pour chaque mot clé faire**
 - 3 | Choisir *images* ;
 - 4 | Choisir *textes* ;
 - 5 | *Diaporama* \leftarrow *images* \cup *textes* ;
 - 6 **si satisfaisant alors**
 - 7 | Retourner *Diaporama* ;
 - 8 **sinon**
 - 9 | Revenir ligne 1 ;
-

- ▶ Un *algorithme* donne une/la solution à un *problème*
- ▶ Classes de problèmes (P, NP, ...)

Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique

Algorithme : Expliquer le titre de mon HDR

Données : Un titre

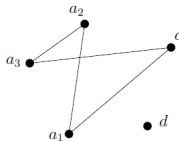
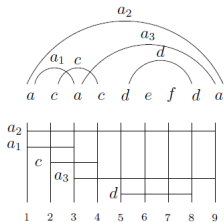
Résultat : Un diaporama d'explication

- 1 Choisir des *mots clés* du titre ;
 - 2 **pour chaque mot clé faire**
 - 3 | Choisir *images* ;
 - 4 | Choisir *textes* ;
 - 5 | *Diaporama* \leftarrow *images* \cup *textes* ;
 - 6 **si satisfaisant alors**
 - 7 | Retourner *Diaporama* ;
 - 8 **sinon**
 - 8 | Revenir ligne 1 ;
-

- ▶ Un *algorithme* donne une/la solution à un *problème*
- ▶ Classes de problèmes (P, NP, ...)
- ▶ "Types" d'algorithmes (exact, approché, polynomial, FPT, ...)

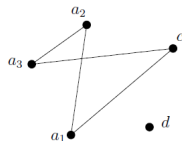
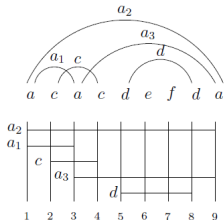
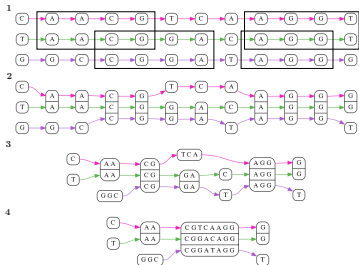
Histoires évolutives et autres comptes

Algorithmes et **graphes** pour la bioinformatique



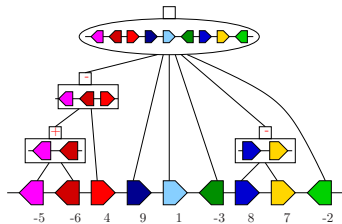
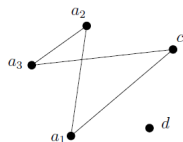
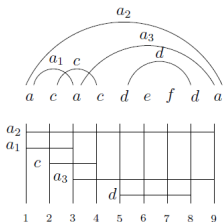
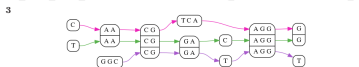
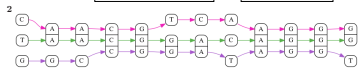
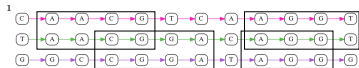
Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique



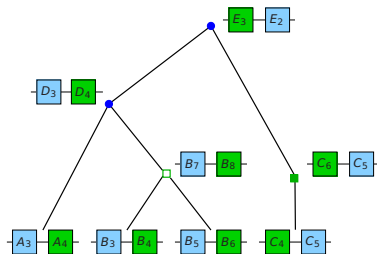
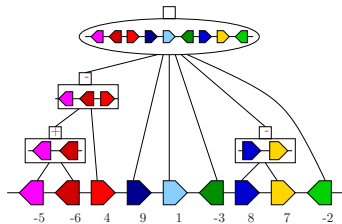
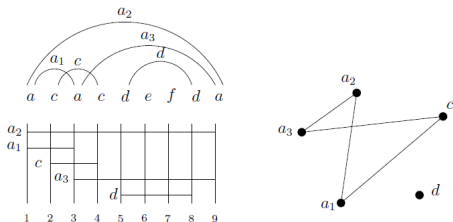
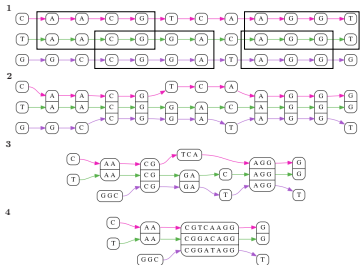
Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique



Histoires évolutives et autres comptes

Algorithmes et graphes pour la bioinformatique

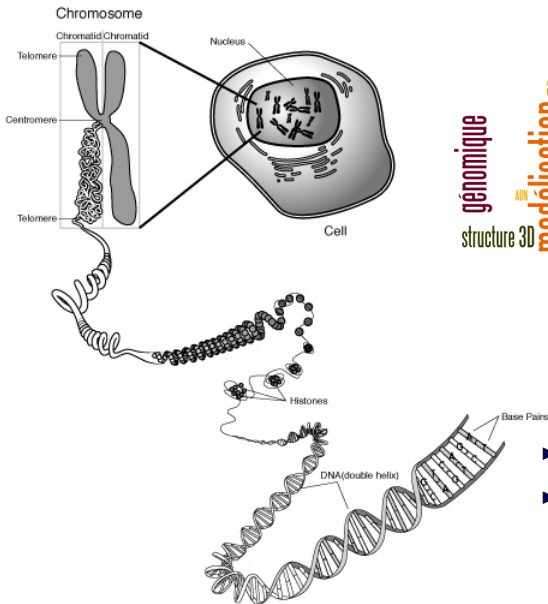






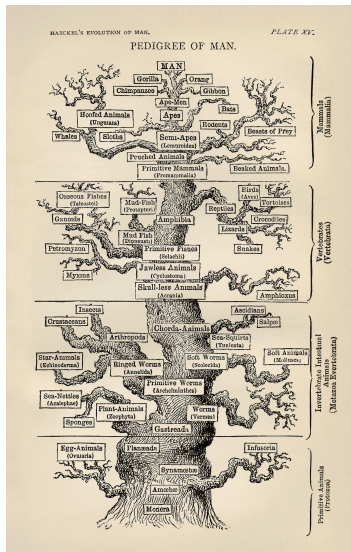
- ▶ Décrypter le vivant, l'évolution

La bioinformatique



- ▶ Décrypter le vivant, l'évolution
- ▶ Traces laissées par l'évolution dans les génomes

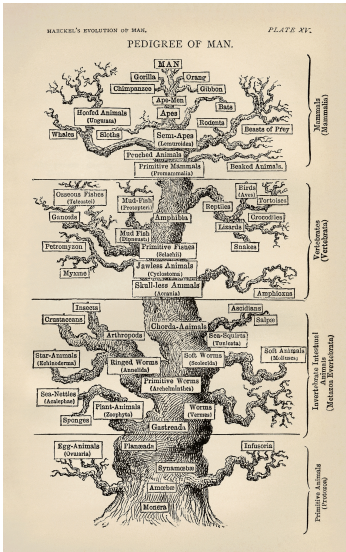
Évolution, phylogénie



La conception d'un arbre de vie au XIXe siècle, par Ernst Haeckel

L'évolution de l'Homme (1879)

Évolution, phylogénie

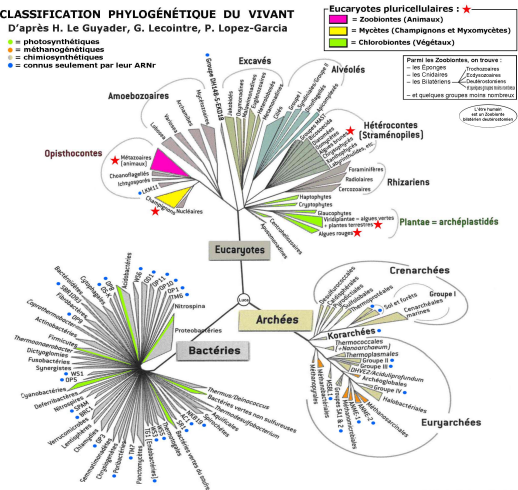


La conception d'un arbre de vie au XIXe siècle, par Ernst Haeckel

L'évolution de l'Homme (1879)

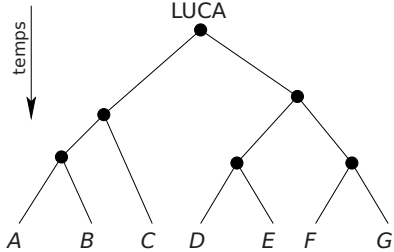
CLASSIFICATION PHYLOGÉNIQUE DU VIVANT

D'après H. Le Guyader, G. Lecointre, P. Lopez-García



La classification phylogénétique du vivant par Guillaume Lecointre & Hervé Le Guyader (2006) et d'après Purificación López-García & David Moreira (2008)

Évolution, phylogénie

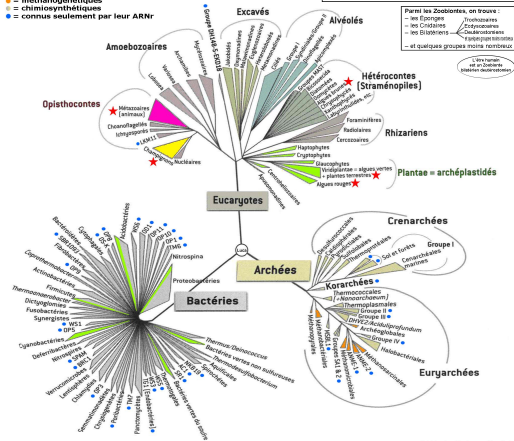


Représentation "informatique" d'un arbre phylogénétique

CLASSIFICATION PHYLOGÉNÉTIQUE DU VIVANT D'après H. Le Guyader, G. Lecointre, P. Lopez-Garcia

- = photosynthétiques
- = méthanogénétiques
- = chimiosynthétiques
- = connus seulement par leur ARNr

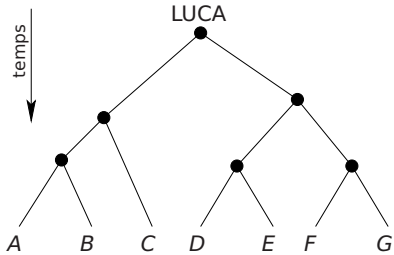
- Eucaryotes pluricellulaires : ★**
- = Zoobiontes (Animaux)
 - = Mycètes (Champignons et Myxomycètes)
 - = Chlorobiontes (Végétaux)



Spiridon Ion Cepleanu - Mer Nature

La classification phylogénétique du vivant
par Guillaume Lecointre & Hervé Le Guyader (2006) et d'après
Purificación López-García & David Moreira (2008)

Évolution, phylogénie



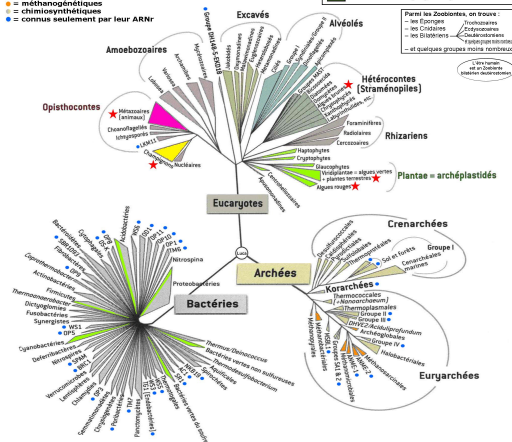
Représentation "informatique" d'un arbre phylogénétique

► Espèces actuelles/ancestrales

CLASSIFICATION PHYLOGÉNÉTIQUE DU VIVANT D'après H. Le Guyader, G. Lecointre, P. Lopez-Garcia

- = photosynthétiques
- = méthanogénétiques
- = chimiosynthétiques
- = connus seulement par leur ARNr

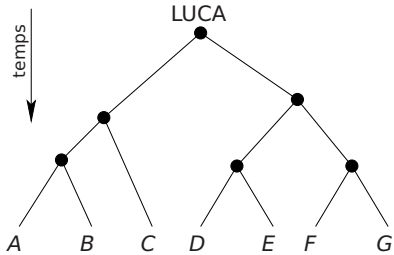
- Eucaryotes pluricellulaires : ★
- = Zoobiontes (Animaux)
 - = Mycètes (Champignons et Myxomycètes)
 - = Chlorobiontes (Végétaux)



Spiridon Ion Capelaru - Mer Nature

La classification phylogénétique du vivant
par Guillaume Lecointre & Hervé Le Guyader (2006) et d'après
Purificación López-García & David Moreira (2008)

Évolution, phylogénie



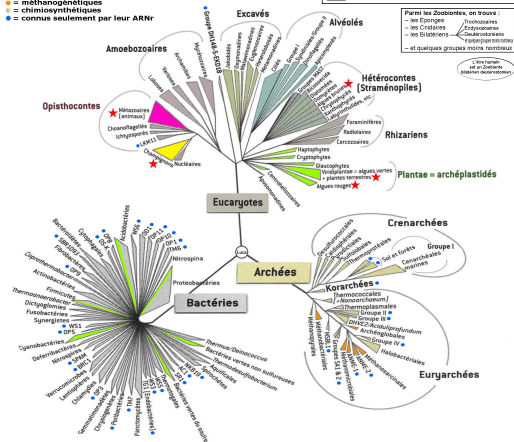
Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

CLASSIFICATION PHYLOGÉNÉTIQUE DU VIVANT D'après H. Le Guyader, G. Lecointre, P. Lopez-Garcia

- = photosynthétiques
- = méthanogénétiques
- = chimiosynthétiques
- = connus seulement par leur ARNr

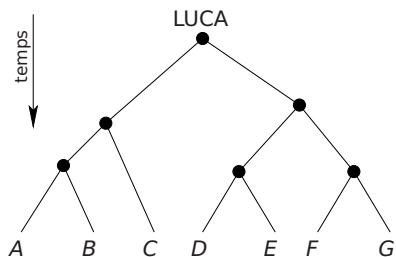
- Eucaryotes pluricellulaires : ★
- = Zoobiontes (Animaux)
 - = Mycètes (Champignons et Myxomycètes)
 - = Chlorobiontes (Végétaux)



Spiridon Ion Cepelaru - Mer Nature

La classification phylogénétique du vivant
par Guillaume Lecointre & Hervé Le Guyader (2006) et d'après
Purificación López-García & David Moreira (2008)

Modèle évolutif = ensemble d'événements évolutifs

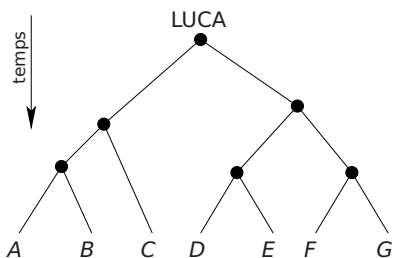


Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

Modèle évolutif = ensemble d'événements évolutifs

1 séquences insertion, délétion, substitution



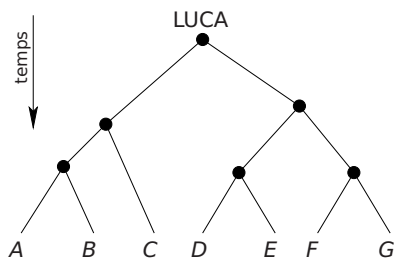
Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

Modèle évolutif = ensemble d'événements évolutifs

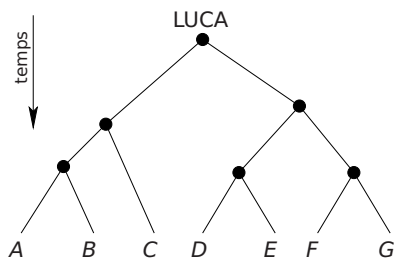
1 **séquences** insertion, délétion, substitution

2 **gènes** duplication (D), perte (L), transfert (T)



Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

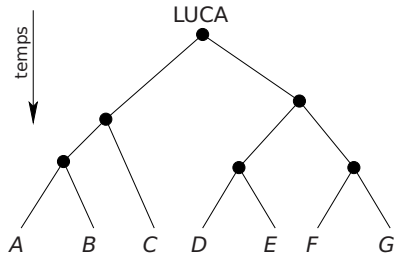


Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

Modèle évolutif = ensemble d'événements évolutifs

- 1 séquences** insertion, délétion, substitution
- 2 gènes** duplication (D), perte (L), transfert (T)
- 3 génomes** réarrangements génomiques :
inversion, DCJ, transposition, fusion/fission de chromosomes ...



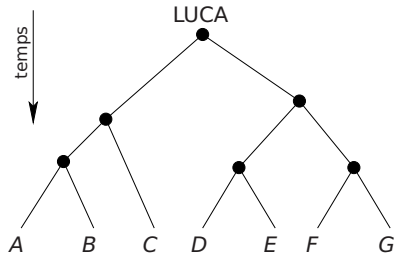
Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

Modèle évolutif = ensemble d'événements évolutifs

- 1 **séquences** insertion, délétion, substitution
- 2 **gènes** duplication (D), perte (L), transfert (T)
- 3 **génomés** réarrangements génomiques :
inversion, DCJ, transposition, fusion/fission de chromosomes ...

▶ Histoire évolutive

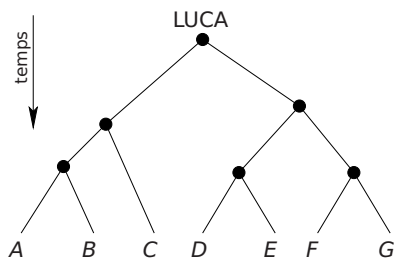


Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

Modèle évolutif = ensemble d'événements évolutifs

- 1 séquences** insertion, délétion, substitution
 - 2 gènes** duplication (D), perte (L), transfert (T)
 - 3 génomes** réarrangements génomiques :
inversion, DCJ, transposition, fusion/fission de chromosomes ...
- ▶ Histoire évolutive
 - ▶ Coût d'une histoire = somme des coûts des événements qu'elle contient



Représentation "informatique" d'un arbre phylogénétique

- ▶ Espèces actuelles/ancestrales
- ▶ Modèles évolutifs à différents niveaux

Modèle évolutif = ensemble d'événements évolutifs

- 1 séquences** insertion, délétion, substitution
- 2 gènes** duplication (D), perte (L), transfert (T)
- 3 génomes** réarrangements génomiques :
inversion, DCJ, transposition, fusion/fission de chromosomes ...

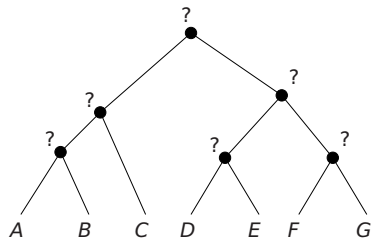
- ▶ Histoire évolutive
- ▶ Coût d'une histoire = somme des coûts des événements qu'elle contient

Principe de parcimonie

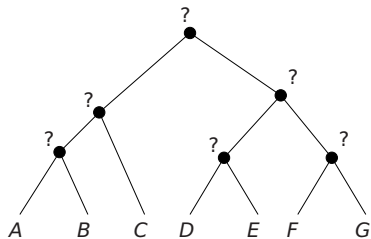
Les histoires de moindre coût sont les plus probables

- 1 Mon parcours
- 2 Domaine de recherche
- 3 Reconstruction de génomes ancestraux : DeCo**
- 4 Synthèse des contributions
- 5 Projets de recherche
- 6 Merci

Délimitons le problème

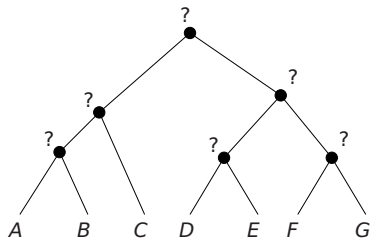


Délimitons le problème



Formalisation

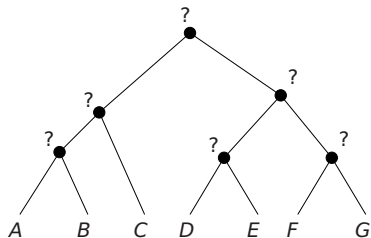
Délimitons le problème



Formalisation

- Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions

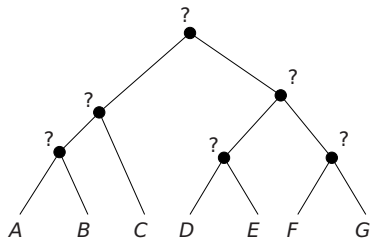
Délimitons le problème



Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

Délimitons le problème

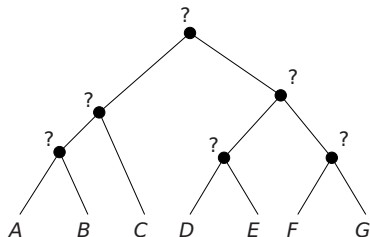


Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

Données disponibles

Délimitons le problème



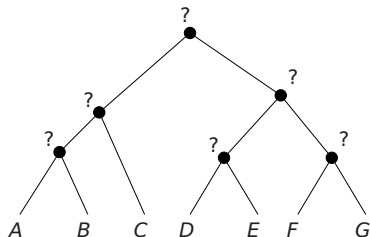
Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*

Délimitons le problème



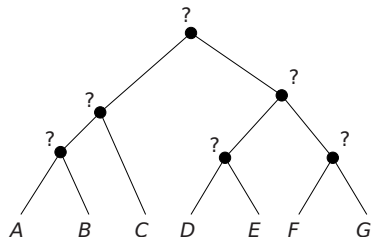
Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Délimitons le problème



Formalisation

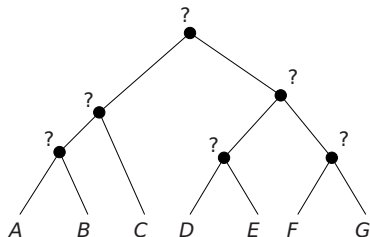
- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Problème

Délimitons le problème



Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

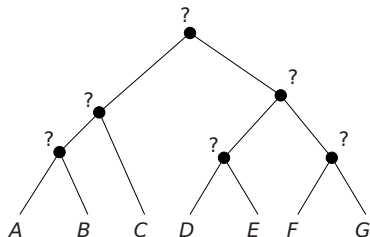
Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...

Délimitons le problème



Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

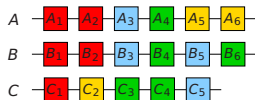
Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...
- ▶ qui "correspondent le mieux" au modèle évolutif

Délimitons le problème



Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

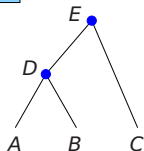
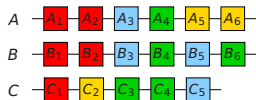
Données disponibles

- ▶ Des **espèces actuelles** + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...
- ▶ qui “correspondent le mieux” au modèle évolutif

Délimitons le problème



Formalisation

- ▶ Génome = un ensemble de gènes (ou marqueurs génomiques) et leurs positions
- ▶ Modèle évolutif

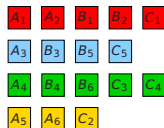
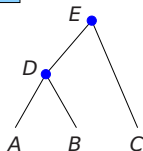
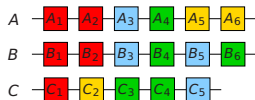
Données disponibles

- ▶ Des espèces actuelles + leur **phylogénie** → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...
- ▶ qui "correspondent le mieux" au modèle évolutif

Délimitons le problème



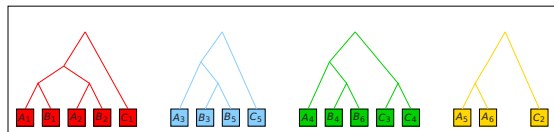
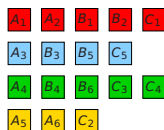
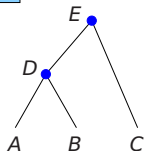
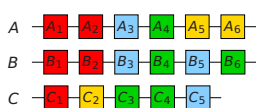
Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les **familles de gènes** + leurs phylogénies → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...
- ▶ qui “correspondent le mieux” au modèle évolutif

Délimitons le problème



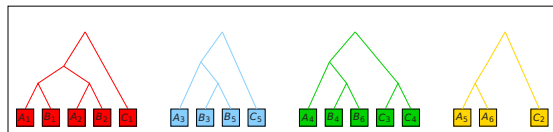
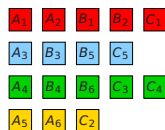
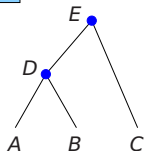
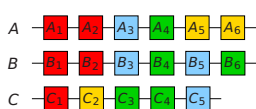
Données disponibles

- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs **phylogénies** → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...
- ▶ qui “correspondent le mieux” au modèle évolutif

Délimitons le problème



Données disponibles

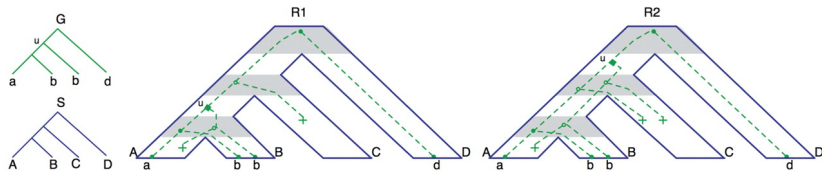
- ▶ Des espèces actuelles + leur phylogénie → *Arbre d'espèces*
- ▶ Les familles de gènes + leurs phylogénies → *Arbres de gènes*

Problème

- ▶ Déterminer les génomes ancestraux aux nœuds internes ...
- ▶ qui “correspondent le mieux” au **modèle évolutif** ⇒ critère à optimiser

Approche phylogénétique : la réconciliation

- Réconciliation des arbres de gènes avec l'arbre d'espèce

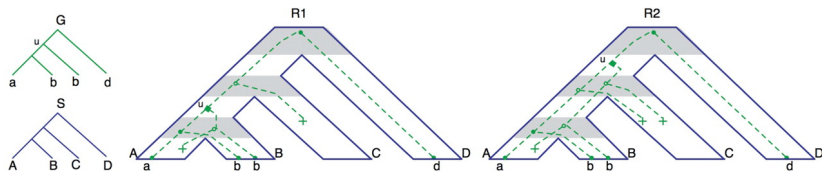


[Doyon *et al.*, 11]

- Données : arbres de gènes
- Modèle évolutif : DL

Approche phylogénétique : la réconciliation

- ▶ Réconciliation des arbres de gènes avec l'arbre d'espèce
- ▶ Fait apparaître les événements qu'ont subis les gènes

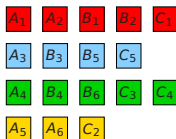
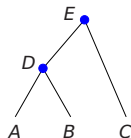


[Doyon *et al.*, 11]

- Données : arbres de gènes
- Modèle évolutif : DL

Approche phylogénétique : la réconciliation

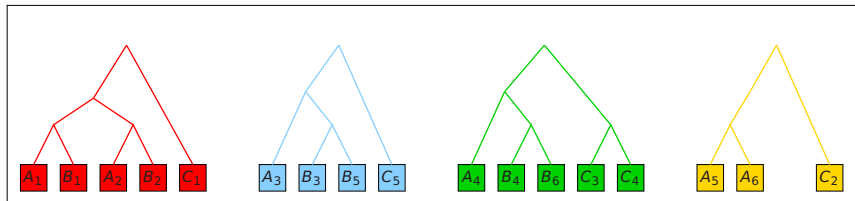
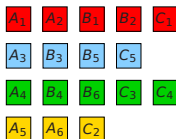
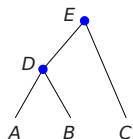
- ▶ Réconciliation des arbres de gènes avec l'arbre d'espèce
- ▶ Fait apparaître les événements qu'ont subis les gènes



- Données : arbres de gènes
- Modèle évolutif : DL

Approche phylogénétique : la réconciliation

- ▶ Réconciliation des arbres de gènes avec l'arbre d'espèce
- ▶ Fait apparaître les événements qu'ont subis les gènes

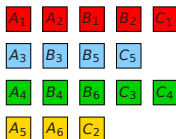
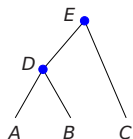


Réconciliation LCA
(*Least Common Ancestor*)

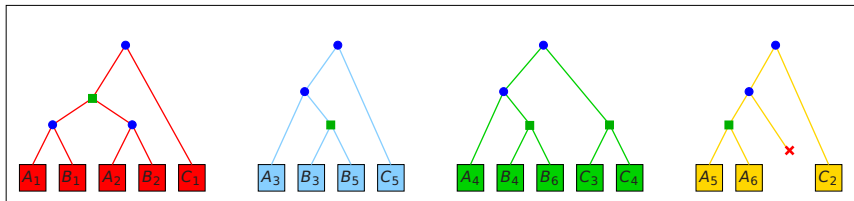
- Données : arbres de gènes
- Modèle évolutif : DL

Approche phylogénétique : la réconciliation

- ▶ Réconciliation des arbres de gènes avec l'arbre d'espèce
- ▶ Fait apparaître les événements qu'ont subis les gènes



- Spéciation
- Duplication de gène
- ✗ Perte de gène

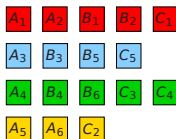
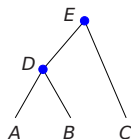


Réconciliation LCA
(*Least Common Ancestor*)

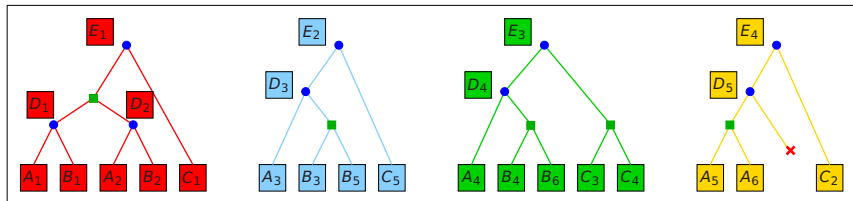
- Données : arbres de gènes
- Modèle évolutif : DL

Approche phylogénétique : la réconciliation

- ▶ Réconciliation des arbres de gènes avec l'arbre d'espèce
- ▶ Fait apparaître les événements qu'ont subis les gènes
- ▶ On obtient le contenu en gènes des génomes ancestraux



- Spéciation
- Duplication de gène
- ✗ Perte de gène

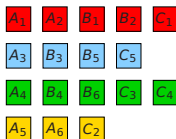
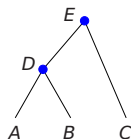


Réconciliation LCA
(*Least Common Ancestor*)

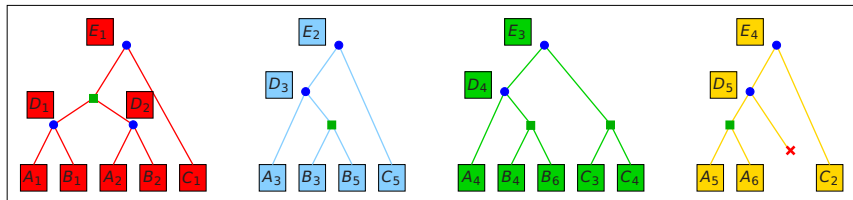
- Données : arbres de gènes
- Modèle évolutif : DL

Approche phylogénétique : la réconciliation

- ▶ Réconciliation des arbres de gènes avec l'arbre d'espèce
- ▶ Fait apparaître les événements qu'ont subis les gènes
- ▶ On obtient le contenu en gènes des génomes ancestraux



- Spéciation
- Duplication de gène
- ✗ Perte de gène



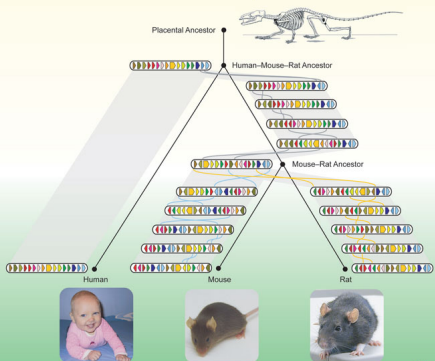
Réconciliation LCA
(*Least Common Ancestor*)

- Données : arbres de gènes
- Modèle évolutif : DL
- Résultats : gènes ancestraux
- Problème : pas d'ordre

GENOME RESEARCH

April 2004
Volume 14 Number 4

Rat Genome Special Issue



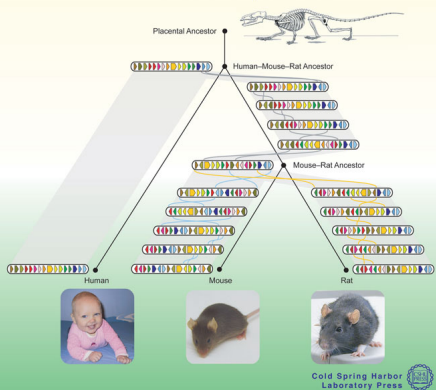


“Marqueurs uniques et universels”

- Génomes sous forme de permutations :

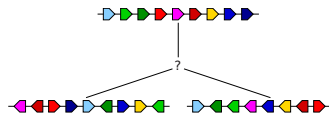
- Données : ordres de gènes
- Modèle évolutif : réarrangements

Rat Genome Special Issue



“Marqueurs uniques et universels”

- Génomes sous forme de permutations :
- Problème de la médiane

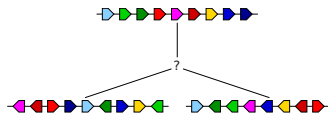


- Données : ordres de gènes
- Modèle évolutif : réarrangements



“Marqueurs uniques et universels”

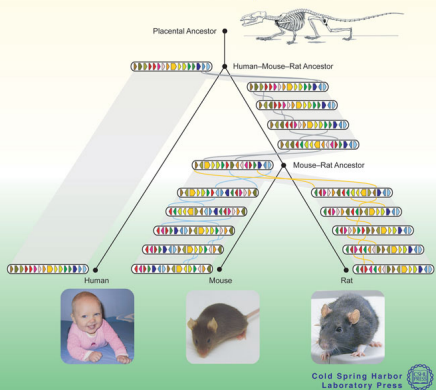
- Génomes sous forme de permutations :
- Problème de la médiane *NP-difficile* pour toutes les distances génomiques



- Données : ordres de gènes
- Modèle évolutif : réarrangements

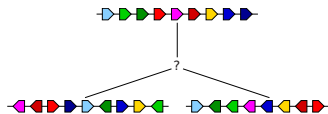


Rat Genome Special Issue



“Marqueurs uniques et universels”

- Génomes sous forme de permutations :
- Problème de la médiane *NP-difficile pour toutes les distances génomiques*



- Données : ordres de gènes
- Modèle évolutif : réarrangements
- Résultats : ordres ancestraux
- Problème : NP-difficile (pas réalisable à grande échelle)

Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'**adjacences** entre gènes

Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'**adjacences** entre gènes

Adjacence



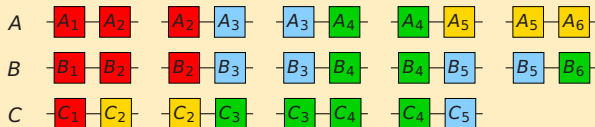
Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'adjacences entre gènes

Adjacence



Génomes = ensembles d'adjacences de gènes



Les adjacences donnent accès à la structure des génomes

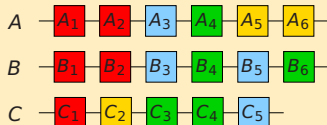
Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'**adjacences** entre gènes

Adjacence



Génomes = ensembles d'adjacences de gènes



Les adjacences donnent accès à la structure des génomes

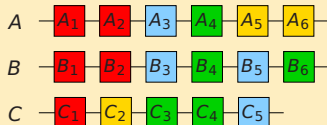
Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'**adjacences** entre gènes

Adjacence



Génomes = ensembles d'adjacences de gènes



Les adjacences donnent accès à la structure des génomes

Idées

- ▶ Combiner les informations des arbres de gènes et des ordres de gènes

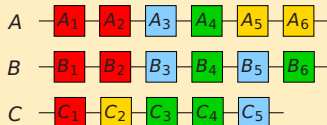
Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'**adjacences** entre gènes

Adjacence



Génomes = ensembles d'adjacences de gènes



Les adjacences donnent accès à la structure des génomes

Idées

- ▶ Combiner les informations des arbres de gènes et des ordres de gènes
- ▶ Hypothèse simplificatrice : les adjacences évoluent **indépendamment**

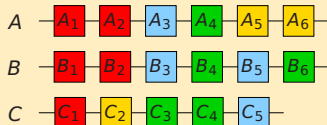
Notre approche : DeCo

DeCo reconstruit l'histoire évolutive des relations d'**adjacences** entre gènes

Adjacence



Génomes = ensembles d'adjacences de gènes



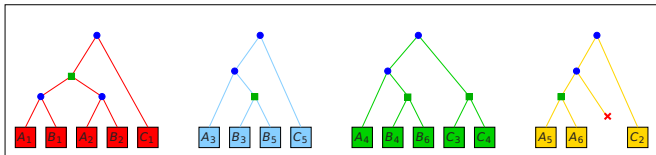
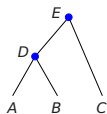
Les adjacences donnent accès à la structure des génomes

Idées

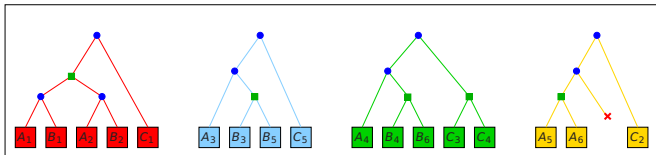
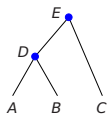
- ▶ Combiner les informations des arbres de gènes et des ordres de gènes
- ▶ Hypothèse simplificatrice : les adjacences évoluent **indépendamment**

Modèle simple

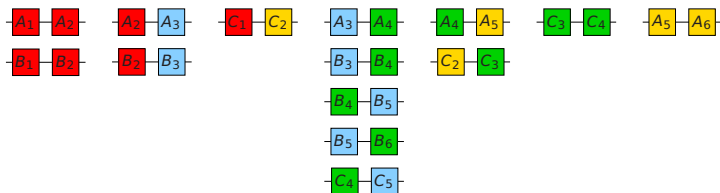
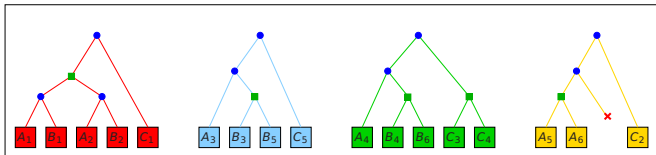
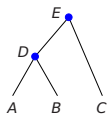
Présence ou absence d'une adjacence dans un génome



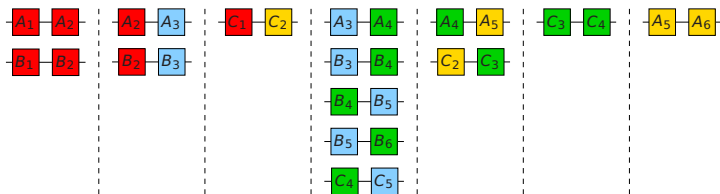
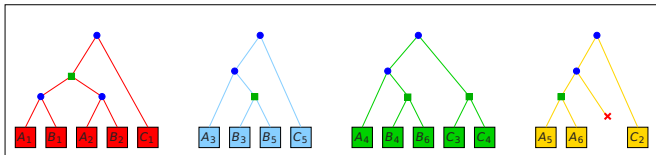
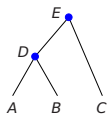
Données : arbre d'espèces, adjacences (\sim ordres de gènes) et arbres de gènes



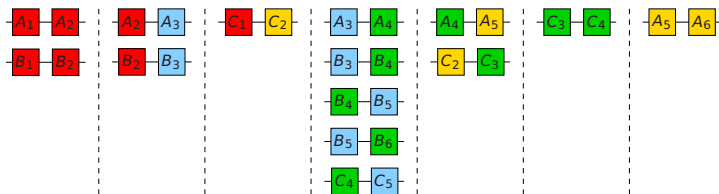
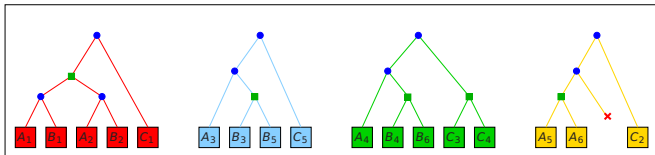
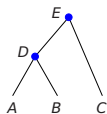
Données : arbre d'espèces, adjacences (\sim ordres de gènes) et arbres de gènes



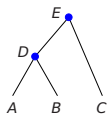
On groupe les adjacences qui peuvent avoir une origine évolutive commune en **classes** ~ familles d'adjacences



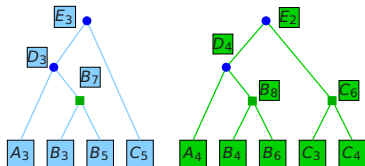
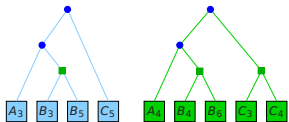
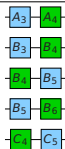
Comme les adjacences sont considérées comme **indépendantes**, on peut les traiter classe par classe



Exemple de reconstruction de l'histoire évolutive d'une classe

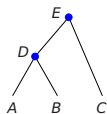


Classe "bleu-vert"

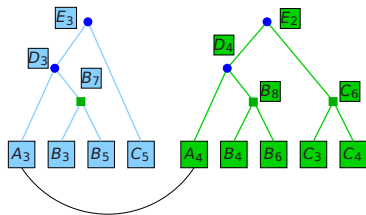
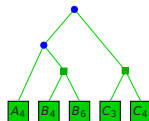
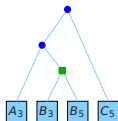
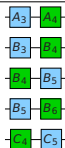


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

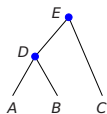


Classe "bleu-vert"

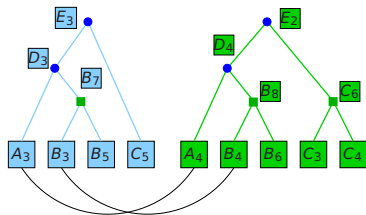
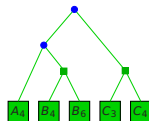
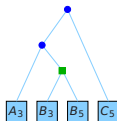


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

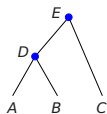


Classe "bleu-vert"

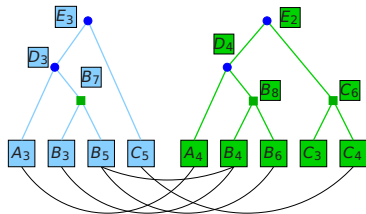
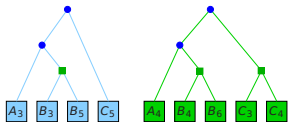
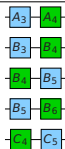


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

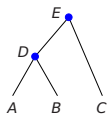


Classe "bleu-vert"

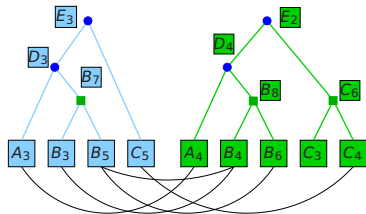
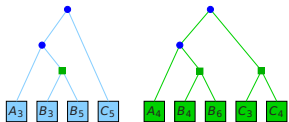
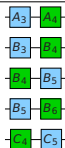


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

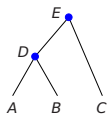


Classe "bleu-vert"

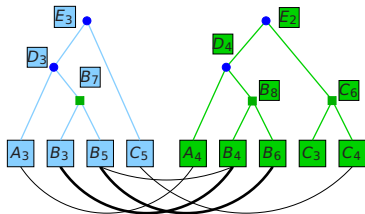
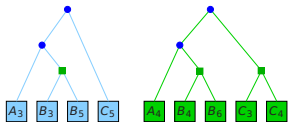
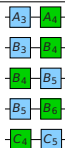


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

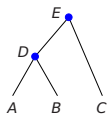


Classe "bleu-vert"

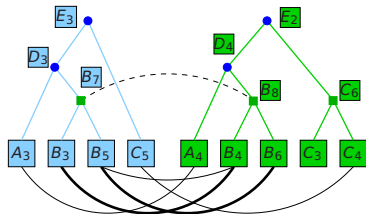
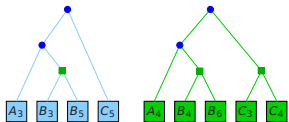
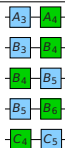


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

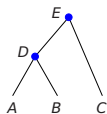


Classe "bleu-vert"

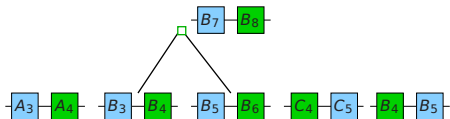
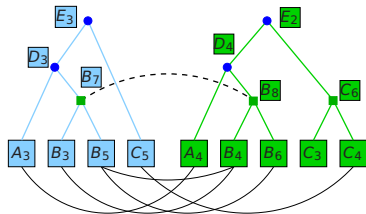
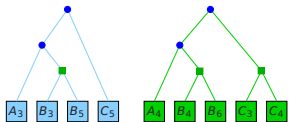
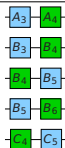


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

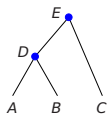


Classe "bleu-vert"

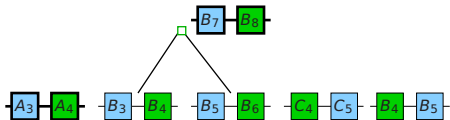
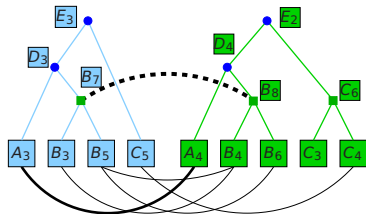
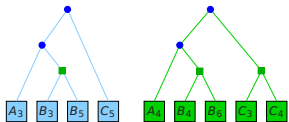
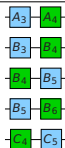


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

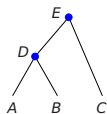


Classe "bleu-vert"

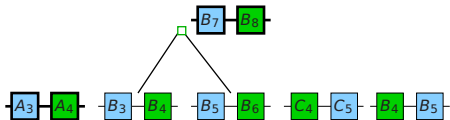
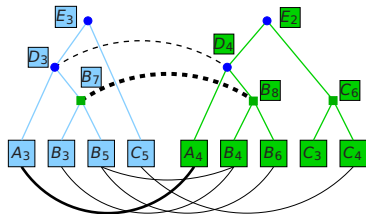
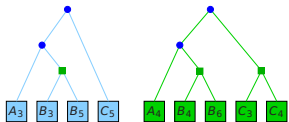
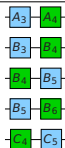


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

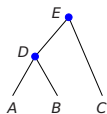


Classe "bleu-vert"

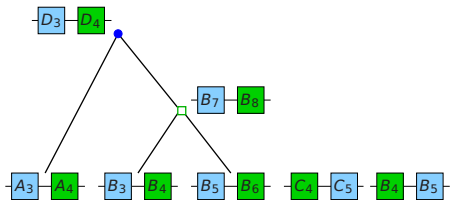
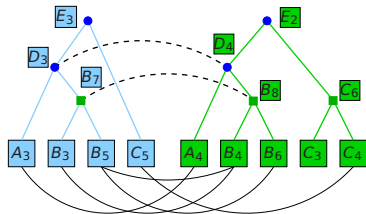
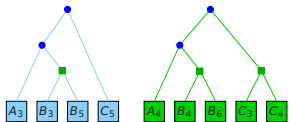


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

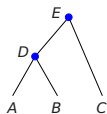


Classe "bleu-vert"

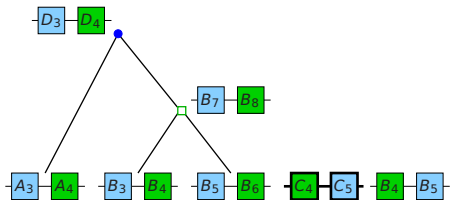
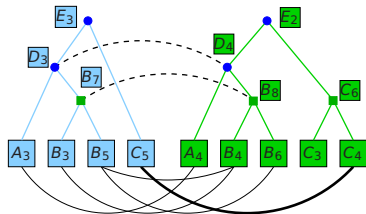
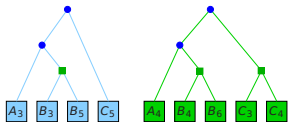
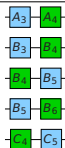


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

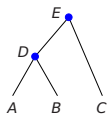


Classe "bleu-vert"

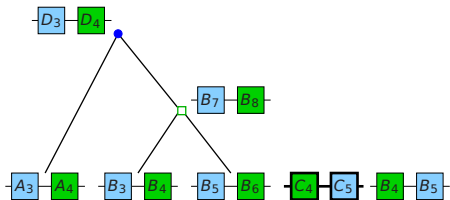
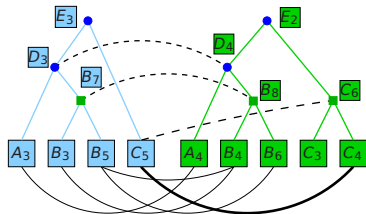
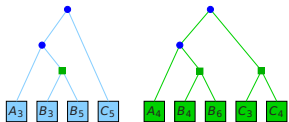
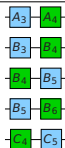


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

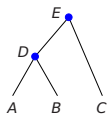


Classe "bleu-vert"

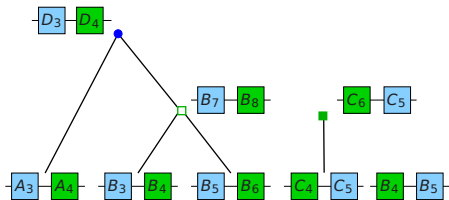
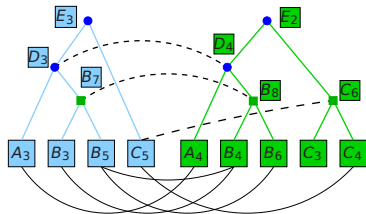
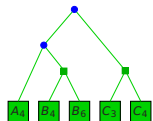
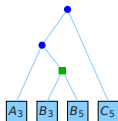
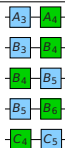


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

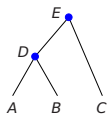


Classe "bleu-vert"

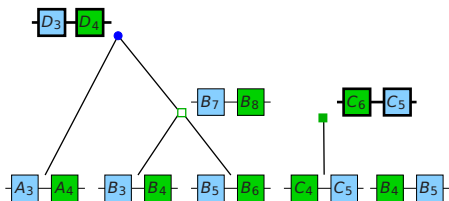
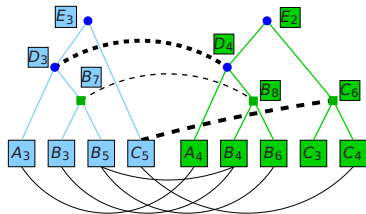
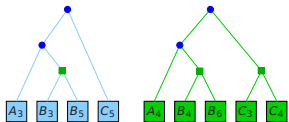
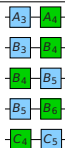


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

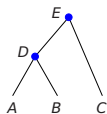


Classe "bleu-vert"

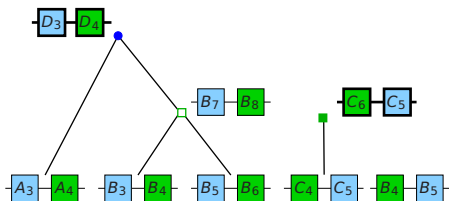
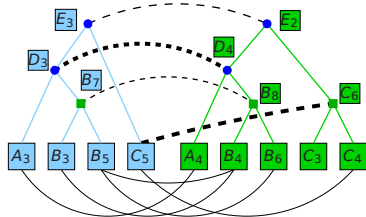
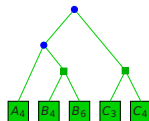
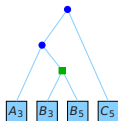
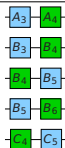


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

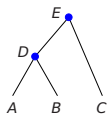


Classe "bleu-vert"

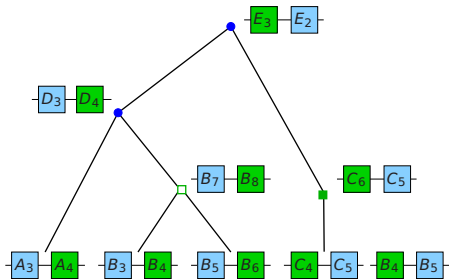
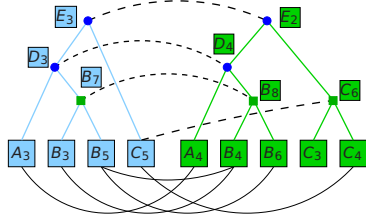
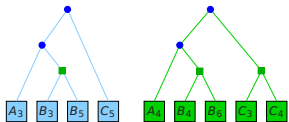
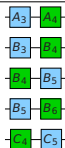


Exemple de reconstruction de l'histoire évolutive d'une classe

On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"

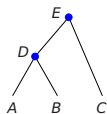


Classe "bleu-vert"

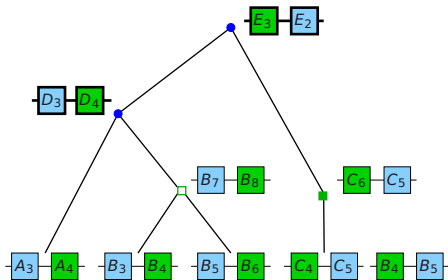
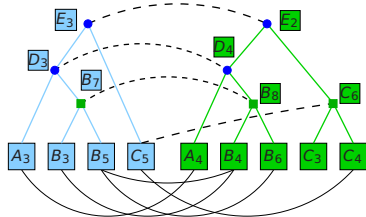
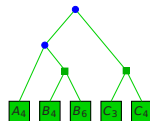
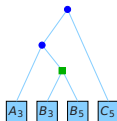
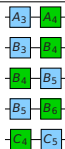


Exemple de reconstruction de l'histoire évolutive d'une classe

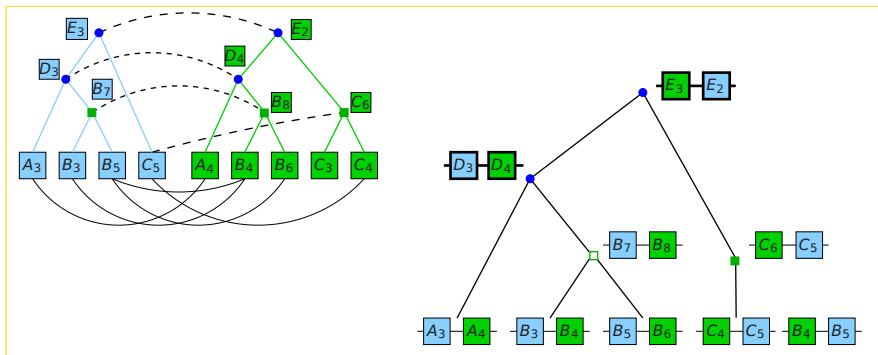
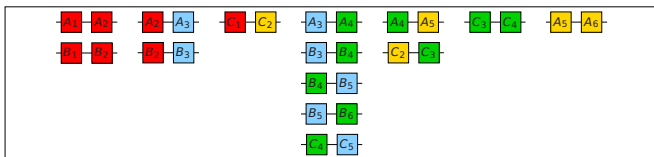
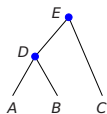
On cherche à construire l'**arbre évolutif des adjacences** de la classe "bleu-vert"



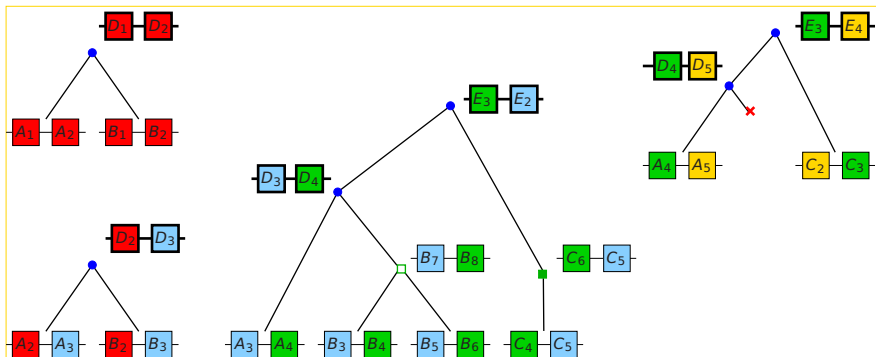
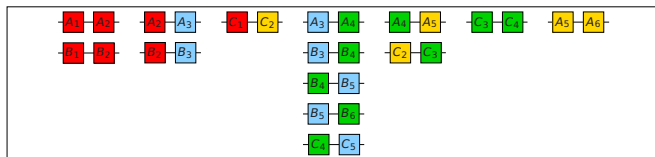
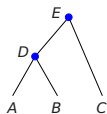
Classe "bleu-vert"



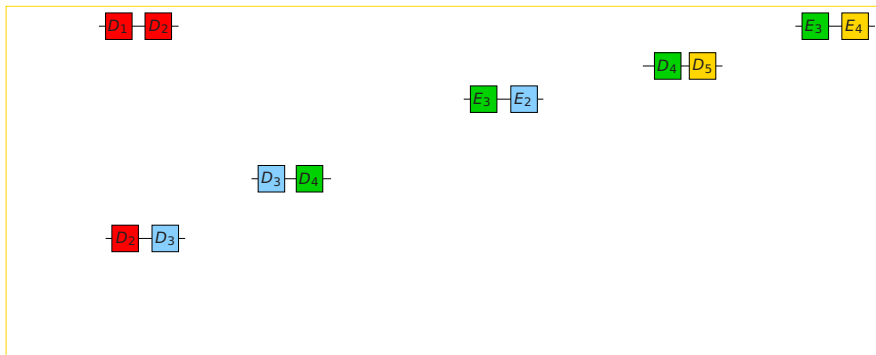
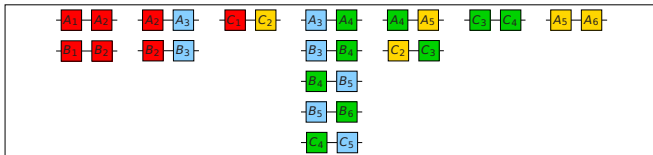
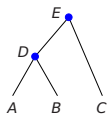
L'arbre évolutif d'adjacences donne accès aux adjacences ancestrales



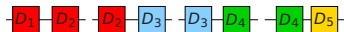
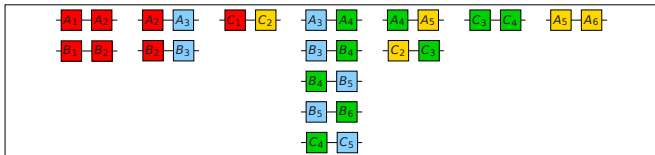
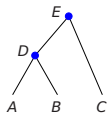
L'arbre évolutif d'adjacences donne accès aux adjacences ancestrales



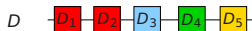
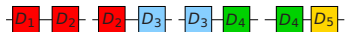
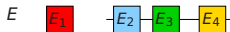
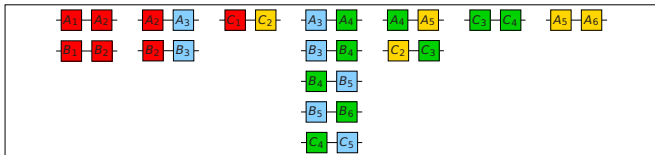
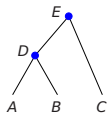
L'arbre évolutif d'adjacences donne accès aux adjacences ancestrales



L'arbre évolutif d'adjacences donne accès aux adjacences ancestrales



En regroupant les informations des classes on obtient la structure des génomes ancestraux



En regroupant les informations des classes on obtient la structure des génomes ancestraux

Reconstruire l'histoire évolutive des adjacences
c'est :

Reconstruire l'histoire évolutive des adjacences
c'est :

- ▶ inférer des adjacences ancestrales

Reconstruire l'histoire évolutive des adjacences
c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Reconstruire l'histoire évolutive des adjacences
c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Reconstruire l'histoire évolutive des adjacences c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Modèle évolutif

- Spéciation
- Duplication de gène
- ✗ Perte de gène
- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ▲ Création
- ≠ Cassure

Reconstruire l'histoire évolutive des adjacences c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Modèle évolutif

- Spéciation
- Duplication de gène
- ✕ Perte de gène
- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ▲ Création
- ≠ Cassure

Critère d'optimisation

Le coût d'une histoire évolutive d'adjacences

- ▶ est la somme des coûts des événements de création et de cassure

Reconstruire l'histoire évolutive des adjacences c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Modèle évolutif

- Spéciation
- Duplication de gène
- ✕ Perte de gène
- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ▲ Création
- ≠ Cassure

Critère d'optimisation

Le coût d'une histoire évolutive d'adjacences

- ▶ est la somme des coûts des événements de création et de cassure
- ▶ il peut être calculé de manière *récursive*

Reconstruire l'histoire évolutive des adjacences c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Modèle évolutif

- Spéciation
- Duplication de gène
- ✕ Perte de gène
- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ▲ Création
- ≠ Cassure

Critère d'optimisation

Le coût d'une histoire évolutive d'adjacences

- ▶ est la somme des coûts des événements de création et de cassure
 - ▶ il peut être calculé de manière *récursive*
-
- ▶ Pour chaque adjacence ancestrale potentielle on calcule 2 coûts :

Reconstruire l'histoire évolutive des adjacences c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Modèle évolutif

- Spéciation
- Duplication de gène
- ✗ Perte de gène
- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ▲ Création
- ≠ Cassure

Critère d'optimisation

Le coût d'une histoire évolutive d'adjacences

- ▶ est la somme des coûts des événements de création et de cassure
- ▶ il peut être calculé de manière *récursive*
- ▶ Pour chaque adjacence ancestrale potentielle on calcule 2 coûts :
 - celui d'une histoire évolutive qui la contient (c_1)

Reconstruire l'histoire évolutive des adjacences c'est :

- ▶ inférer des adjacences ancestrales
- ▶ et les événements qu'elles ont subis

Cette histoire *doit être en cohérence* avec les histoires des gènes

Modèle évolutif

- Spéciation
- Duplication de gène
- ✕ Perte de gène
- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ▲ Création
- ≠ Cassure

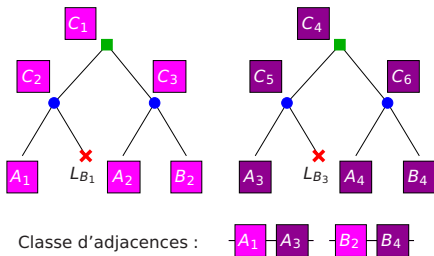
Critère d'optimisation

Le coût d'une histoire évolutive d'adjacences

- ▶ est la somme des coûts des événements de création et de cassure
- ▶ il peut être calculé de manière *récursive*

- ▶ Pour chaque adjacence ancestrale potentielle on calcule 2 coûts :
 - celui d'une histoire évolutive qui la contient (c_1)
 - celui d'une histoire évolutive qui ne la contient pas (c_0)

Programmation dynamique → "big data friendly"



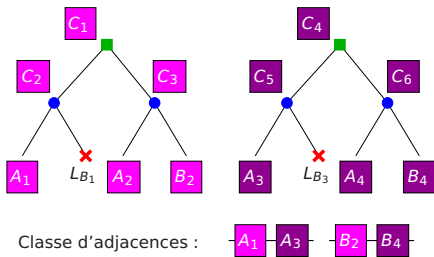
c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1							
L_{B_1}							
A_2							
B_2							
C_2							
C_3							
C_1							

► Matrice de programmation dynamique, coûts c_0 et c_1 :

→ $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents

→ $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents

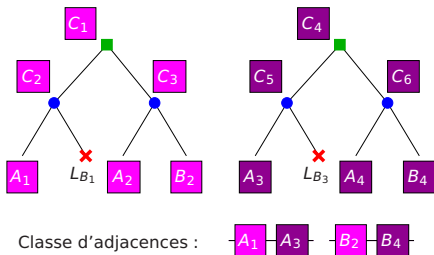
Programmation dynamique → "big data friendly"



c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1							
L_{B_1}							
A_2							
B_2							
C_2							
C_3							
C_1							

- ▶ Matrice de programmation dynamique, coûts c_0 et c_1 :
 - $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents
 - $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents
- ▶ Calcul des coûts entre nœuds *de même espèce*

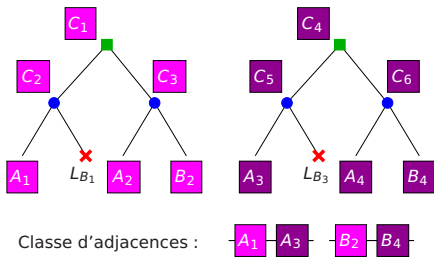
Programmation dynamique → "big data friendly"



c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1							
L_{B_1}							
A_2							
B_2							
C_2							
C_3							
C_1							

- ▶ Matrice de programmation dynamique, coûts c_0 et c_1 :
 - $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents
 - $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents
- ▶ Calcul des coûts entre nœuds *de même espèce*

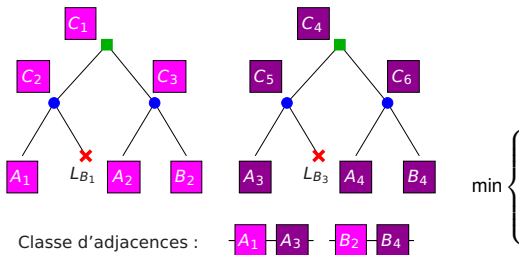
Programmation dynamique → "big data friendly"



c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1							
L_{B_1}							
A_2							
B_2							
C_2							
C_3							
C_1							

- ▶ Matrice de programmation dynamique, coûts c_0 et c_1 :
 - $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents
 - $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents
- ▶ Calcul des coûts entre nœuds *de même espèce*
- ▶ Calcul des coûts *selon les événements* associés aux nœuds
 - ⇒ *plusieurs cas*

Programmation dynamique → "big data friendly"



C_3, C_6 : nœuds de spéciation

$$c_1(C_3, C_6) =$$

min

- ▶ Matrice de programmation dynamique, coûts c_0 et c_1 :

→ $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents

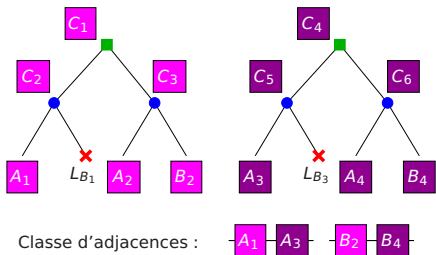
→ $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents

- ▶ Calcul des coûts entre nœuds *de même espèce*

- ▶ Calcul des coûts *selon les événements* associés aux nœuds

⇒ *plusieurs cas*

Programmation dynamique \rightarrow "big data friendly"



C_3, C_6 : nœuds de spéciation

$$c_1(C_3, C_6) =$$

$$\min \left\{ c_1(A_2, A_4) + c_1(B_2, B_4) \right.$$

► Matrice de programmation dynamique, coûts c_0 et c_1 :

$\rightarrow c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents

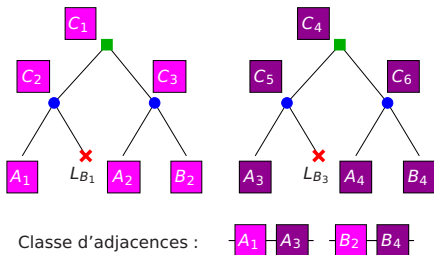
$\rightarrow c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents

► Calcul des coûts entre nœuds *de même espèce*

► Calcul des coûts *selon les événements* associés aux nœuds

\Rightarrow *plusieurs cas*

Programmation dynamique → "big data friendly"



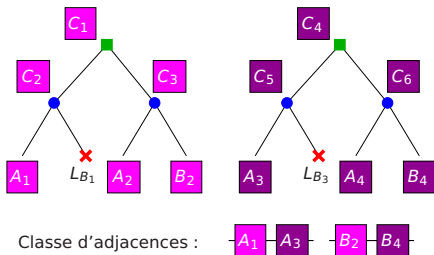
C_3, C_6 : nœuds de spéciation

$$c_1(C_3, C_6) =$$

$$\min \left\{ \begin{array}{l} c_1(A_2, A_4) + c_1(B_2, B_4) \\ c_1(A_2, A_4) + c_0(B_2, B_4) + C(\text{Cassure}) \end{array} \right.$$

- ▶ Matrice de programmation dynamique, coûts c_0 et c_1 :
 - $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents
 - $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents
- ▶ Calcul des coûts entre nœuds *de même espèce*
- ▶ Calcul des coûts *selon les événements* associés aux nœuds
⇒ *plusieurs cas*

Programmation dynamique → "big data friendly"



C_3, C_6 : nœuds de spéciation

$$c_1(C_3, C_6) =$$

$$\min \left\{ \begin{array}{l} c_1(A_2, A_4) + c_1(B_2, B_4) \\ c_1(A_2, A_4) + c_0(B_2, B_4) + C(\text{Cassure}) \\ c_0(A_2, A_4) + c_1(B_2, B_4) + C(\text{Cassure}) \end{array} \right.$$

▶ Matrice de programmation dynamique, coûts c_0 et c_1 :

→ $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents

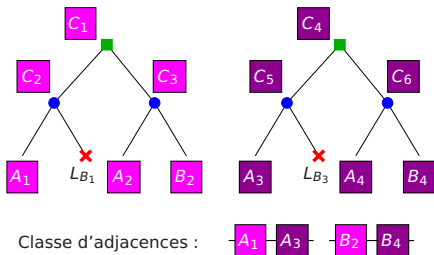
→ $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents

▶ Calcul des coûts entre nœuds *de même espèce*

▶ Calcul des coûts *selon les événements* associés aux nœuds

⇒ *plusieurs cas*

Programmation dynamique → "big data friendly"



C_3, C_6 : nœuds de spéciation

$$c_1(C_3, C_6) =$$

$$\min \left\{ \begin{array}{l} c_1(A_2, A_4) + c_1(B_2, B_4) \\ c_1(A_2, A_4) + c_0(B_2, B_4) + C(\text{Cassure}) \\ c_0(A_2, A_4) + c_1(B_2, B_4) + C(\text{Cassure}) \\ c_0(A_2, A_4) + c_0(B_2, B_4) + 2 * C(\text{Cassure}) \end{array} \right.$$

▶ Matrice de programmation dynamique, coûts c_0 et c_1 :

→ $c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents

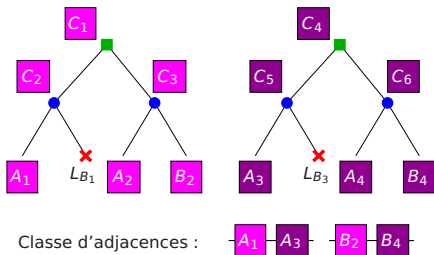
→ $c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents

▶ Calcul des coûts entre nœuds *de même espèce*

▶ Calcul des coûts *selon les événements* associés aux nœuds

⇒ *plusieurs cas*

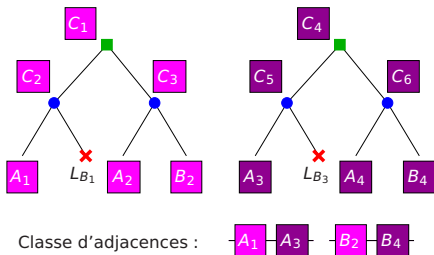
Programmation dynamique \rightarrow "big data friendly"



c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

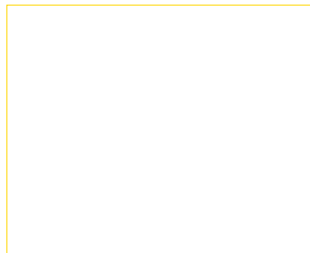
- ▶ Matrice de programmation dynamique, coûts c_0 et c_1 :
 - $\rightarrow c_0(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *ne sont pas* adjacents
 - $\rightarrow c_1(g_1, g_2)$ coût minimum d'une histoire évolutive où g_1 et g_2 *sont* adjacents
- ▶ Calcul des coûts entre nœuds *de même espèce*
- ▶ Calcul des coûts *selon les événements* associés aux nœuds
 - \Rightarrow *plusieurs cas*

Programmation dynamique \rightarrow "big data friendly"

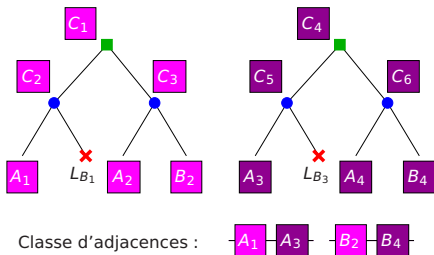


c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

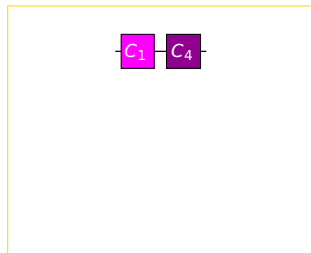


Programmation dynamique → "big data friendly"

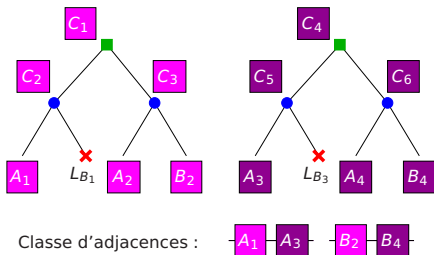


c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

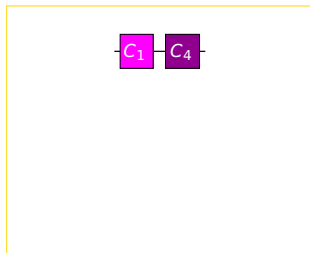


Programmation dynamique → "big data friendly"

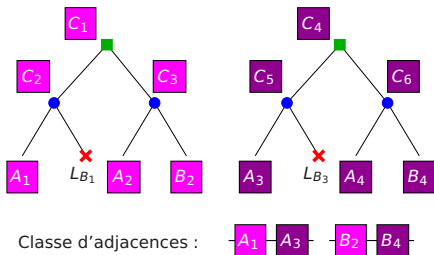


c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking



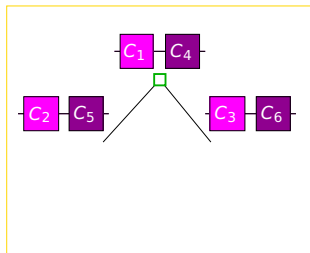
Programmation dynamique → "big data friendly"



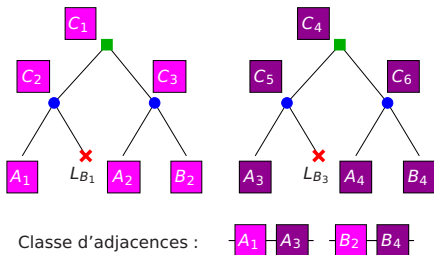
c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

- Duplication d'adjacence



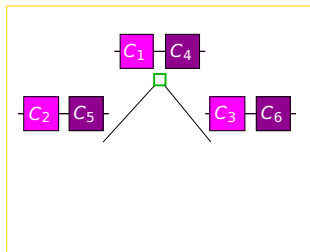
Programmation dynamique → "big data friendly"



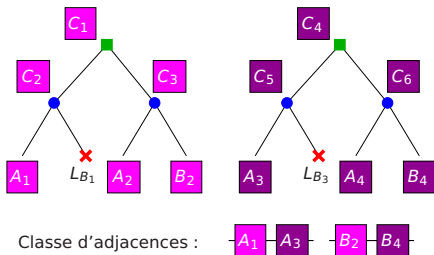
c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

- Duplication d'adjacence



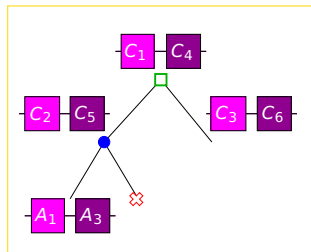
Programmation dynamique → "big data friendly"



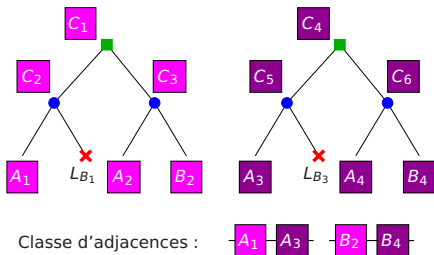
c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

- Duplication d'adjacence
- ⊗ Perte d'adjacence



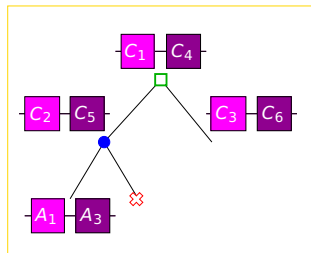
Programmation dynamique → "big data friendly"



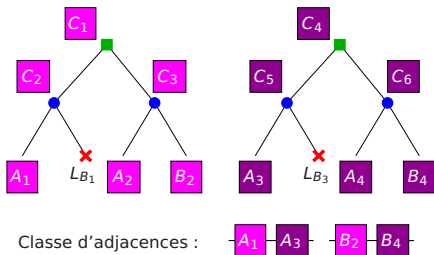
c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

- Duplication d'adjacence
- ⊗ Perte d'adjacence



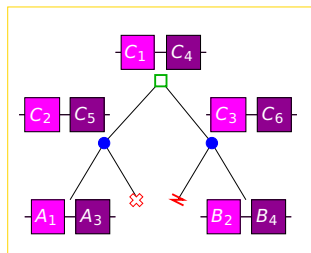
Programmation dynamique → "big data friendly"



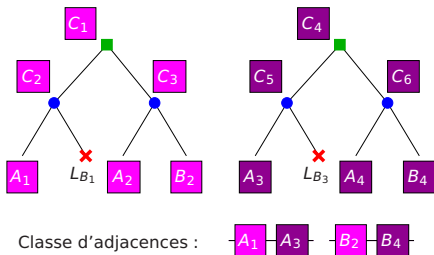
c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ≠ Cassure



Programmation dynamique \rightarrow "big data friendly"

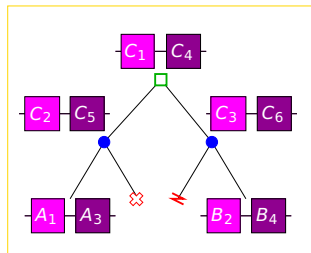


c_0/c_1	A_3	L_{B_3}	A_4	B_4	C_5	C_6	C_4
A_1	$\infty/0$		$0/\infty$				
L_{B_1}		$0/0$		$0/0$			
A_2	$0/\infty$		$0/\infty$				
B_2		$0/0$		$\infty/0$			
C_2					$1/0$	$0/1$	$1/0$
C_3					$0/1$	$1/1$	$1/1$
C_1					$1/0$	$1/1$	$2/1$

► Backtracking

- Duplication d'adjacence
- ⊗ Perte d'adjacence
- ⚡ Cassure

► Construction d'un arbre d'adjacence de coût minimum



Premier modèle mixant l'ordre des gènes et leurs phylogénies

Premier modèle mixant l'ordre des gènes et leurs phylogénies

- ▶ Permet de reconstruire la structure des génomes ancestraux

Premier modèle mixant l'ordre des gènes et leurs phylogénies

- ▶ Permet de reconstruire la structure des génomes ancestraux
- ▶ Méthodes antérieures limitées aux marqueurs uniques et universels et/ou ignorant duplications et pertes de gènes → **modèle plus général**

Premier modèle mixant l'ordre des gènes et leurs phylogénies

- ▶ Permet de reconstruire la structure des génomes ancestraux
- ▶ Méthodes antérieures limitées aux marqueurs uniques et universels et/ou ignorant duplications et pertes de gènes → **modèle plus général**
- ▶ Algorithme **exact** (généralisation de Fitch-Sankoff)

Premier modèle mixant l'ordre des gènes et leurs phylogénies

- ▶ Permet de reconstruire la structure des génomes ancestraux
- ▶ Méthodes antérieures limitées aux marqueurs uniques et universels et/ou ignorant duplications et pertes de gènes → **modèle plus général**
- ▶ Algorithme **exact** (généralisation de Fitch-Sankoff)
- ▶ Complexité en temps et en espace en $O(n^2)$: **rapide**

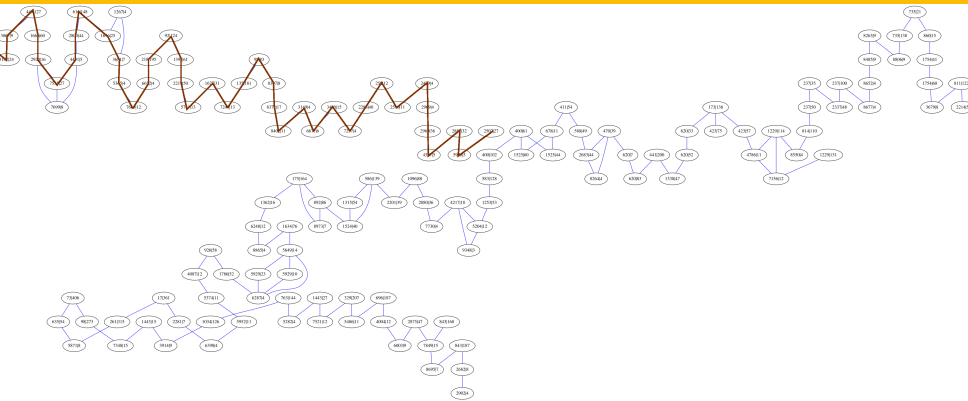
Premier modèle mixant l'ordre des gènes et leurs phylogénies

- ▶ Permet de reconstruire la structure des génomes ancestraux
- ▶ Méthodes antérieures limitées aux marqueurs uniques et universels et/ou ignorant duplications et pertes de gènes → **modèle plus général**
- ▶ Algorithme **exact** (généralisation de Fitch-Sankoff)
- ▶ Complexité en temps et en espace en $O(n^2)$: **rapide**
- ▶ Détection de duplications de plusieurs gènes simultanément

Premier modèle mixant l'ordre des gènes et leurs phylogénies

- ▶ Permet de reconstruire la structure des génomes ancestraux
- ▶ Méthodes antérieures limitées aux marqueurs uniques et universels et/ou ignorant duplications et pertes de gènes → **modèle plus général**
- ▶ Algorithme **exact** (généralisation de Fitch-Sankoff)
- ▶ Complexité en temps et en espace en $O(n^2)$: **rapide**
- ▶ Détection de duplications de plusieurs gènes simultanément
- ▶ **Possibilité de passer à l'échelle** : plusieurs dizaines d'espèces et milliers d'arbres de gènes

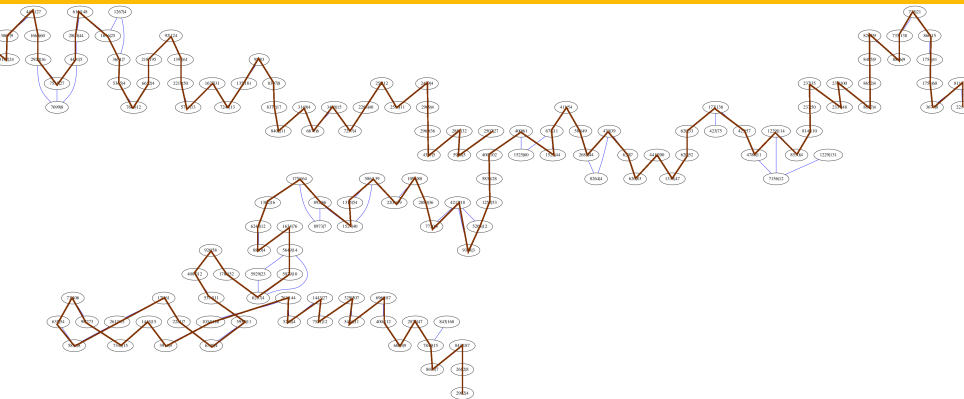
Résultats sur des jeux de données réels



Grphe (partiel) d'adjacences ancestrales

Ancêtre des hominoïdes (~ 16MA)

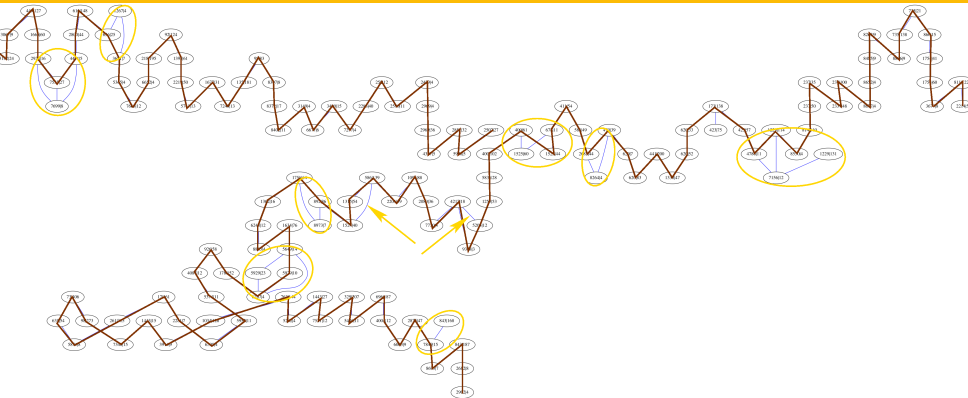
Résultats sur des jeux de données réels



Graphe (partiel) d'adjacences ancestrales

Ancêtre des hominoïdes (~ 16MA)

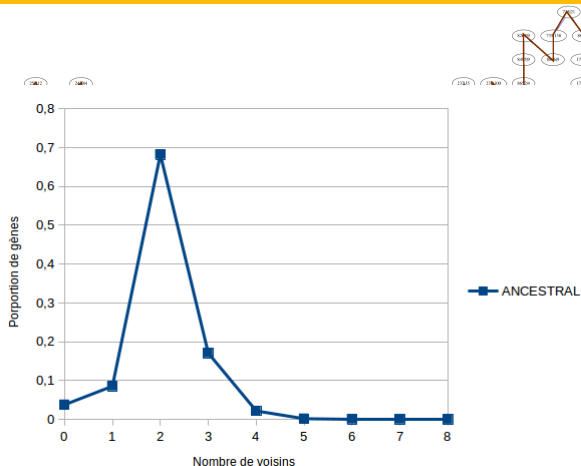
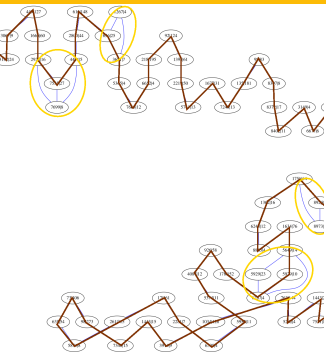
Résultats sur des jeux de données réels



Grphe (partiel) d'adjacences ancestrales

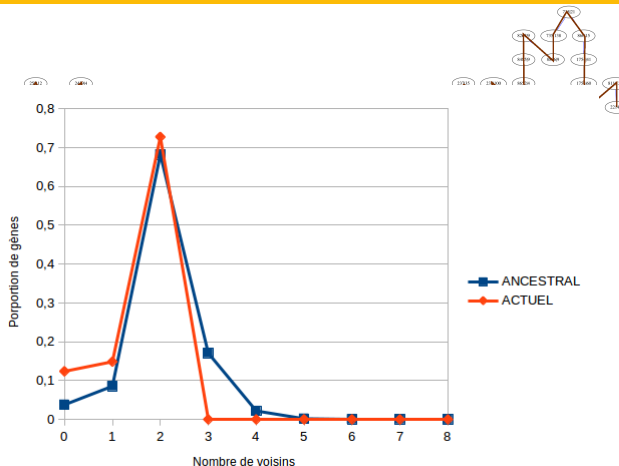
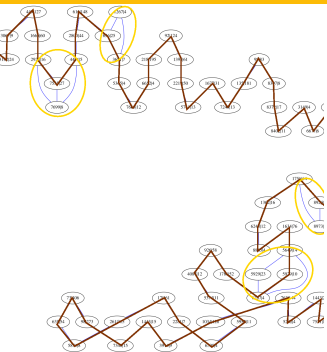
Ancêtre des hominoïdes (~ 16MA)

Résultats sur des jeux de données réels



Graphes (partiel) d'adjacences ancestrales
Ancêtre des hominoïdes (~ 16MA)

Résultats sur des jeux de données réels



Graphe (partiel) d'adjacences ancestrales
Ancêtre des hominoïdes (~ 16MA)

Non linéarité des génomes ancestraux

- ▶ Possibilité de linéariser par des méthodes *ad hoc*

Non linéarité des génomes ancestraux

- ▶ Possibilité de linéariser par des méthodes *ad hoc*
- ▶ Mais cela reflète aussi une certaine *incertitude* dans les données

Non linéarité des génomes ancestraux

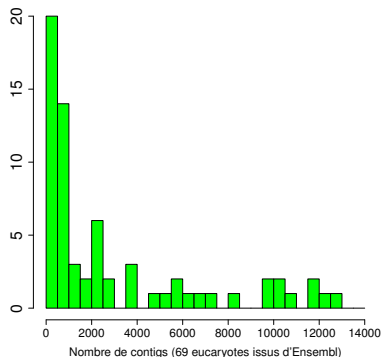
- ▶ Possibilité de linéariser par des méthodes *ad hoc*
- ▶ Mais cela reflète aussi une certaine *incertitude* dans les données
- ▶ Qualité des données :

Non linéarité des génomes ancestraux

- ▶ Possibilité de linéariser par des méthodes *ad hoc*
- ▶ Mais cela reflète aussi une certaine *incertitude* dans les données
- ▶ Qualité des données :
 - Histoires des gènes non exactes ?

Non linéarité des génomes ancestraux

- ▶ Possibilité de linéariser par des méthodes *ad hoc*
- ▶ Mais cela reflète aussi une certaine *incertitude* dans les données
- ▶ Qualité des données :
 - Histoires des gènes non exactes ?
 - Problème dans les assemblages : fragmentation

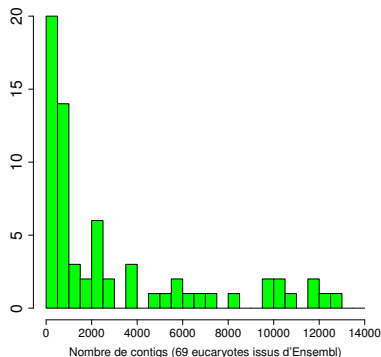


Non linéarité des génomes ancestraux

- ▶ Possibilité de linéariser par des méthodes *ad hoc*
- ▶ Mais cela reflète aussi une certaine *incertitude* dans les données
- ▶ Qualité des données :
 - Histoires des gènes non exactes ?
 - Problème dans les assemblages : fragmentation

ART-DeCo

Une adjacence non présente dans les données est pourvue d'une probabilité d'existence selon la fragmentation de l'espèce et *validée en fonction du contexte*



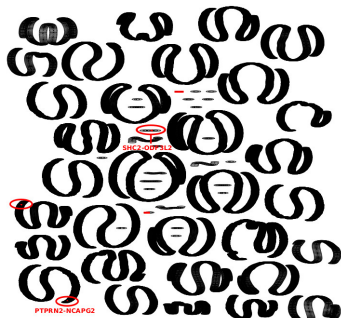
Non linéarité des génomes ancestraux

- ▶ Possibilité de linéariser par des méthodes *ad hoc*
- ▶ Mais cela reflète aussi une certaine *incertitude* dans les données
- ▶ Qualité des données :
 - Histoires des gènes non exactes ?
 - Problème dans les assemblages : fragmentation

ART-DeCo

Une adjacence non présente dans les données est pourvue d'une probabilité d'existence selon la fragmentation de l'espèce et *validée en fonction du contexte*

- ▶ Permet d'améliorer les assemblages ancestraux et actuels



Exemple avec le génome du cheval

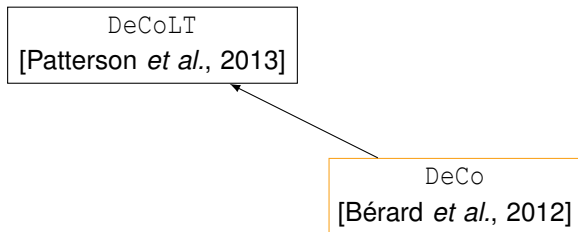
DeCo
[Bérard *et al.*, 2012]

- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)

DeCo "the mother of all"

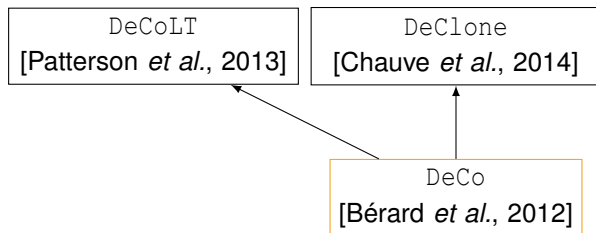
DeCoLT
[Patterson *et al.*, 2013]

DeCo
[Bérard *et al.*, 2012]



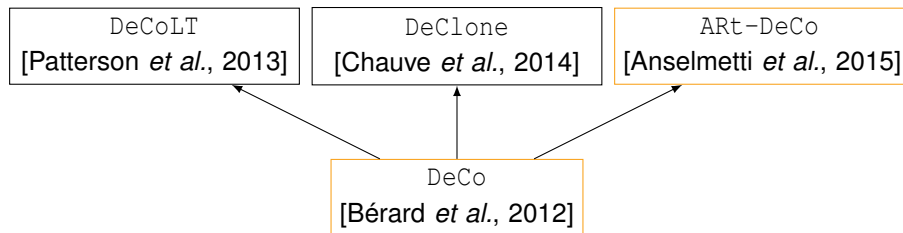
- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)

DeCo "the mother of all"



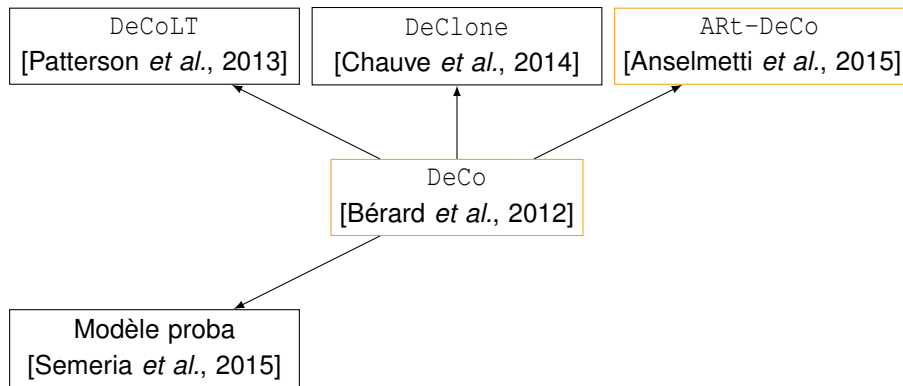
- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)
- ▶ DeClone : Cédric Chauve (SFU)

DeCo "the mother of all"



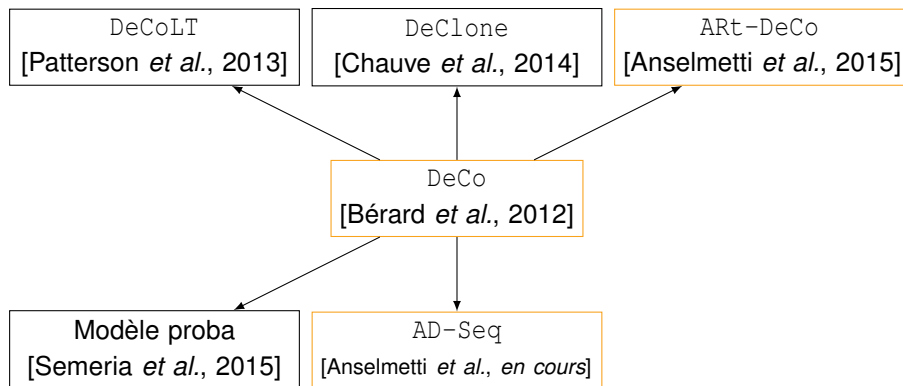
- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)
- ▶ DeClone : Cédric Chauve (SFU)
- ▶ ARt-DeCo : thèse Yoann Anselmetti (SB, Annie Chateau, Éric Tannier)

DeCo "the mother of all"



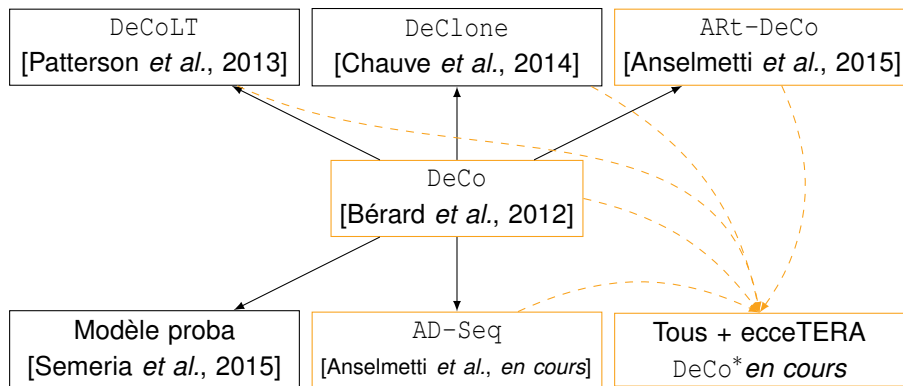
- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)
- ▶ DeClone : Cédric Chauve (SFU)
- ▶ ARt-DeCo : thèse Yoann Anselmetti (SB, Annie Chateau, Éric Tannier)

DeCo "the mother of all"



- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)
- ▶ DeClone : Cédric Chauve (SFU)
- ▶ ART-DeCo : thèse Yoann Anselmetti (SB, Annie Chateau, Éric Tannier) + AD-Seq

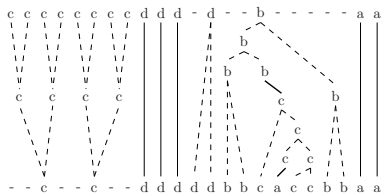
DeCo "the mother of all"



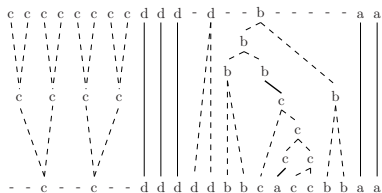
- ▶ DeCo : stage de M2 Coralie Gallien (SB, Éric Tannier)
- ▶ DeClone : Cédric Chauve (SFU)
- ▶ ART-DeCo : thèse Yoann Anselmetti (SB, Annie Chateau, Éric Tannier) + AD-Seq
- ▶ ecceTERA : Comprehensive gene tree-species tree reconciliation using parsimony [Jacox *et al.*, 2016] (Cédric Chauve + Céline Scornavacca, ISEM)

- 1 Mon parcours
- 2 Domaine de recherche
- 3 Reconstruction de génomes ancestraux : DeCo
- 4 Synthèse des contributions**
- 5 Projets de recherche
- 6 Merci

Alignement

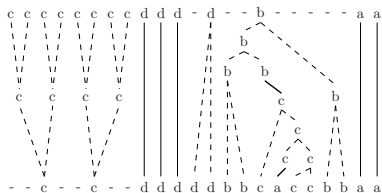


Alignement



Séquences répétées en tandem

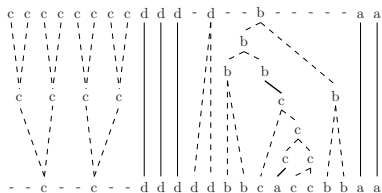
Alignement



Séquences répétées en tandem

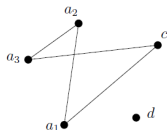
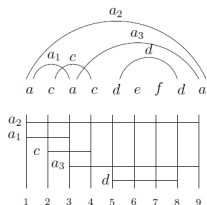
- Modèle évolutif particulier

Alignement



Séquences répétées en tandem

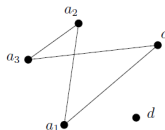
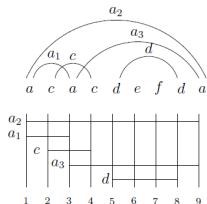
- ▶ Modèle évolutif particulier
- ▶ Amplifications et contractions



Séquences répétées en tandem

- ▶ Modèle évolutif particulier
- ▶ Amplifications et contractions
- ▶ Retrouver l'histoire évolutives des arches

Alignement



Séquences répétées en tandem

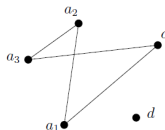
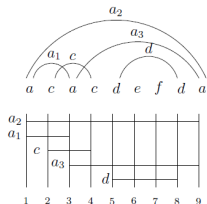
- ▶ Modèle évolutif particulier
- ▶ Amplifications et contractions
- ▶ Retrouver l'histoire évolutives des arches

Bactériophages

Alignement multiple de génomes complets de bactériophages colinéaires



Alignement



Séquences répétées en tandem

- ▶ Modèle évolutif particulier
- ▶ Amplifications et contractions
- ▶ Retrouver l'histoire évolutives des arches

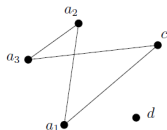
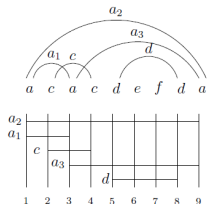
Bactériophages

Alignement multiple de génomes complets de bactériophages colinéaires

- ▶ Modèle évolutif particulier



Alignement



Séquences répétées en tandem

- ▶ Modèle évolutif particulier
- ▶ Amplifications et contractions
- ▶ Retrouver l'histoire évolutives des arches

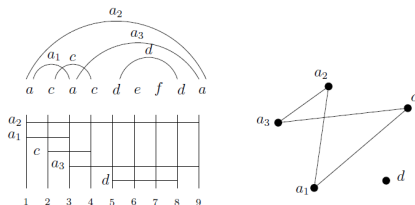
Bactériophages

Alignement multiple de génomes complets de bactériophages colinéaires

- ▶ Modèle évolutif particulier
- ▶ Recombinaison modulaire



Alignement



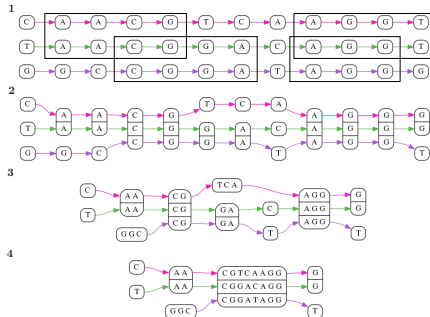
Séquences répétées en tandem

- ▶ Modèle évolutif particulier
- ▶ Amplifications et contractions
- ▶ Retrouver l'histoire évolutives des arches

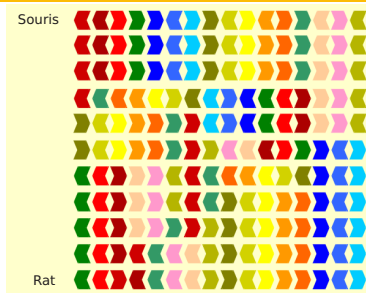
Bactériophages

Alignement multiple de génomes complets de bactériophages colinéaires

- ▶ Modèle évolutif particulier
- ▶ Recombinaison modulaire
- ▶ Graphe d'alignement/ordre partiel

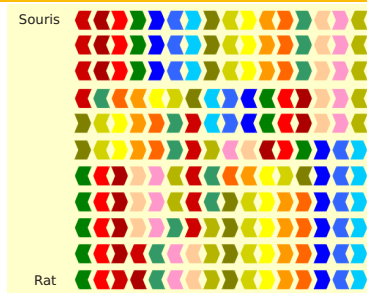


Génomique comparative



Scénarios évolutifs “parfaits”

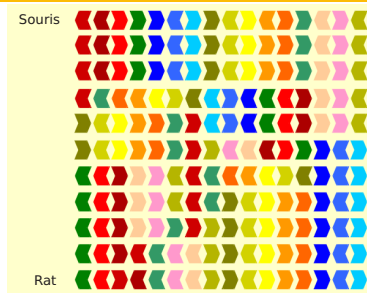
Génomique comparative



Scénarios évolutifs “parfaits”

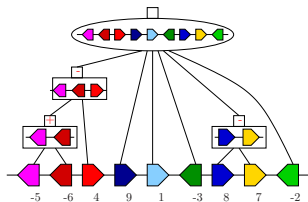
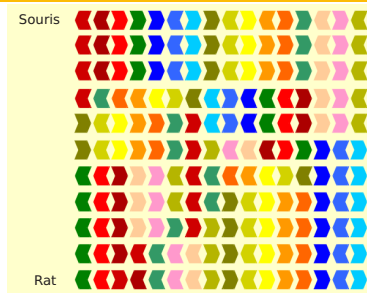
- Conservation de structures

Génomique comparative



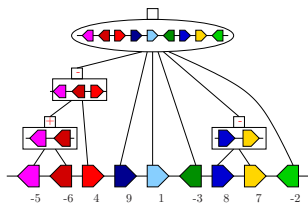
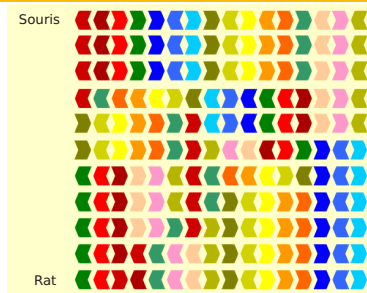
Scénarios évolutifs “parfaits”

- ▶ Conservation de structures
- ▶ Deux modèles évolutifs : inversions et DCJ



Scénarios évolutifs "parfaits"

- ▶ Conservation de structures
- ▶ Deux modèles évolutifs : inversions et DCJ
- ▶ PQ-arbres



Scénarios évolutifs "parfaits"

- ▶ Conservation de structures
- ▶ Deux modèles évolutifs : inversions et DCJ
- ▶ PQ-arbres
- ▶ Caractérisation des problèmes (P ou NP) et complexité paramétrée (algorithmes FPT)

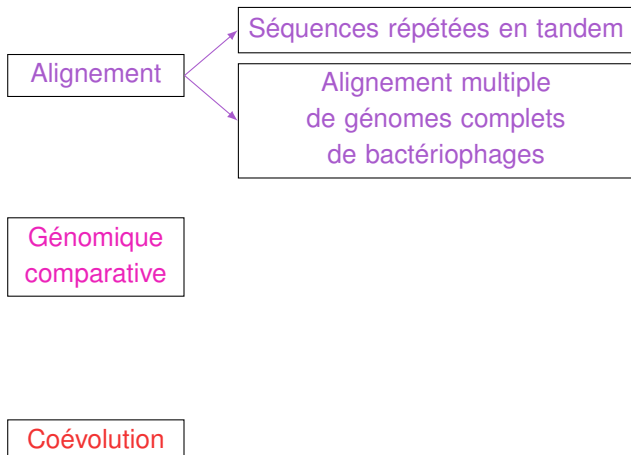
Contributions & collaborations principales

Alignement

Génomique
comparative

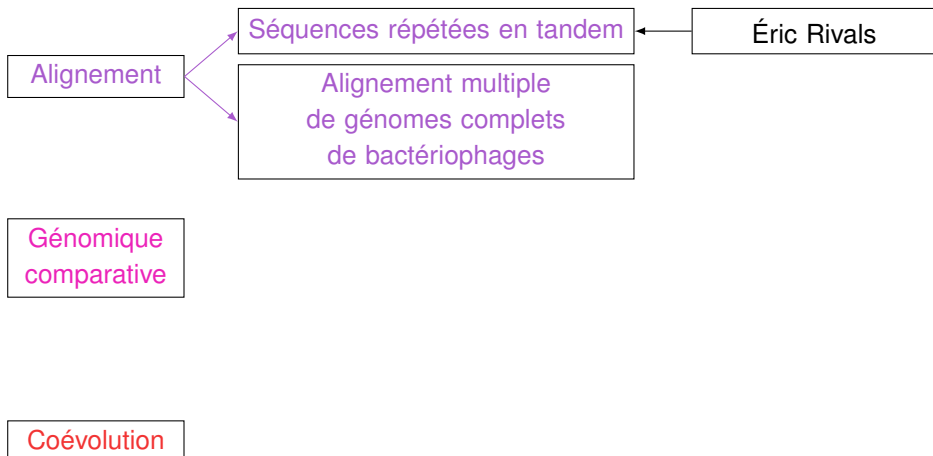
Coévolution

Contributions & collaborations principales



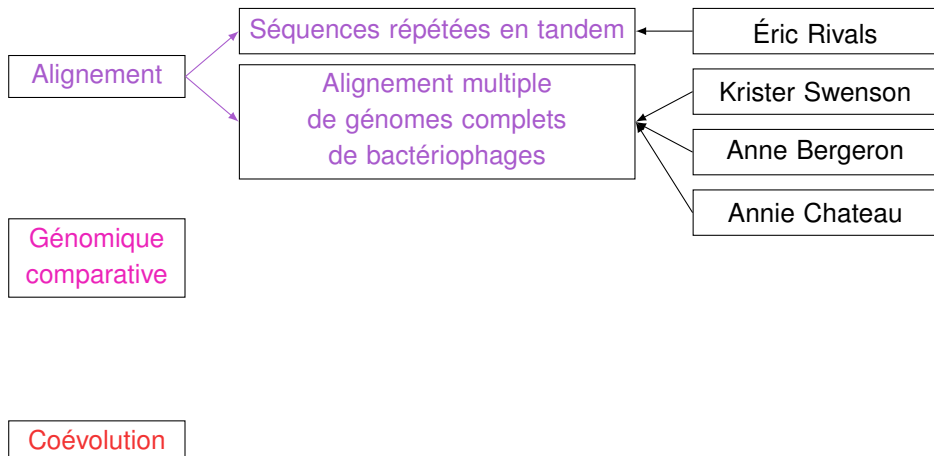
Mise au point de méthodes d'alignement spécifiques

Contributions & collaborations principales



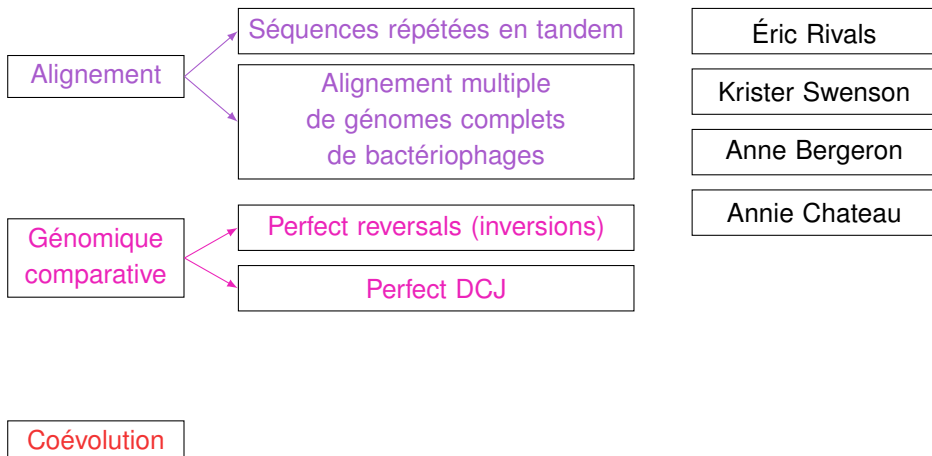
Mise au point de méthodes d'alignement spécifiques

Contributions & collaborations principales



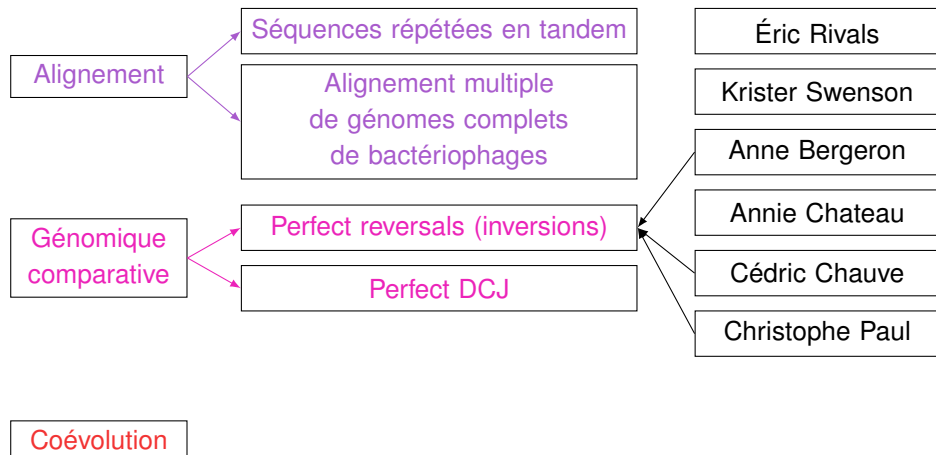
Mise au point de méthodes d'alignement spécifiques

Contributions & collaborations principales



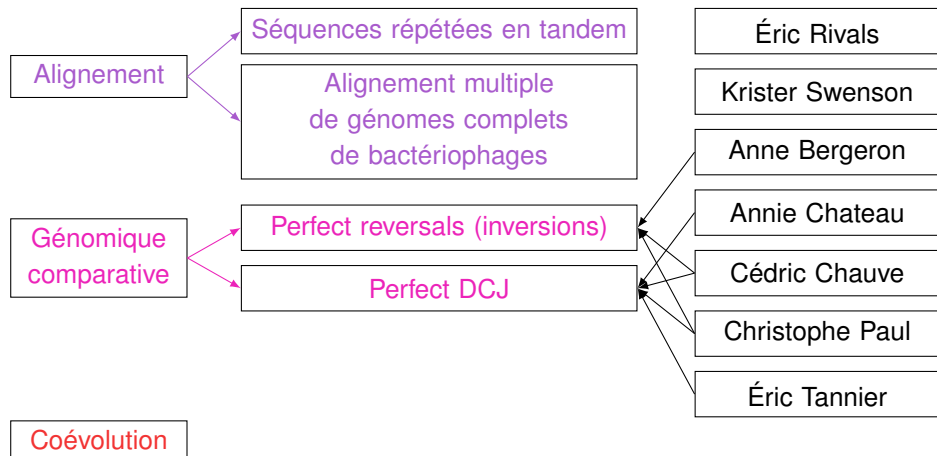
Scénarios évolutifs “parfaits” (conservation de structures), caractérisation des problèmes (P ou NP) et complexité paramétrée (algorithmes FPT)

Contributions & collaborations principales



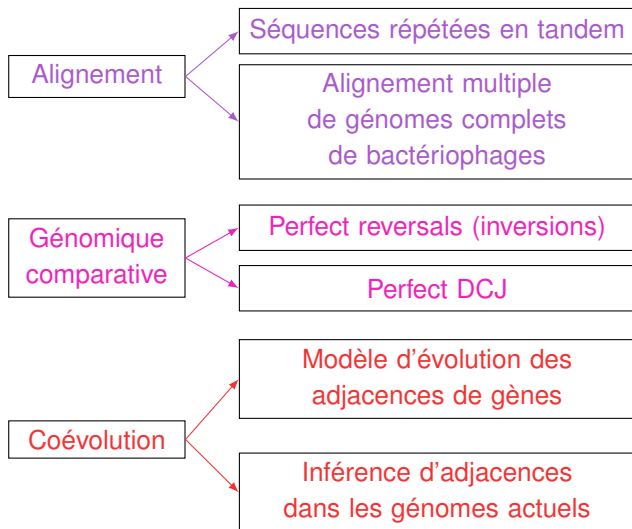
Scénarios évolutifs “parfaits” (conservation de structures), caractérisation des problèmes (P ou NP) et complexité paramétrée (algorithmes FPT)

Contributions & collaborations principales



Scénarios évolutifs “parfaits” (conservation de structures), caractérisation des problèmes (P ou NP) et complexité paramétrée (algorithmes FPT)

Contributions & collaborations principales



Éric Rivals

Krister Swenson

Anne Bergeron

Annie Chateau

Cédric Chauve

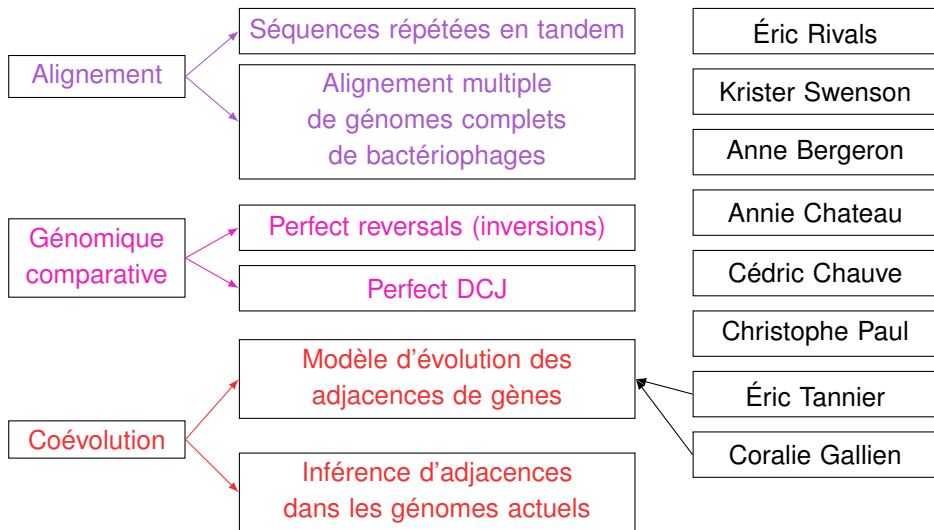
Christophe Paul

Éric Tannier

DeCo : structure des génomes ancestraux

Art-DeCo, AD-Seq, ... : amélioration d'assemblages

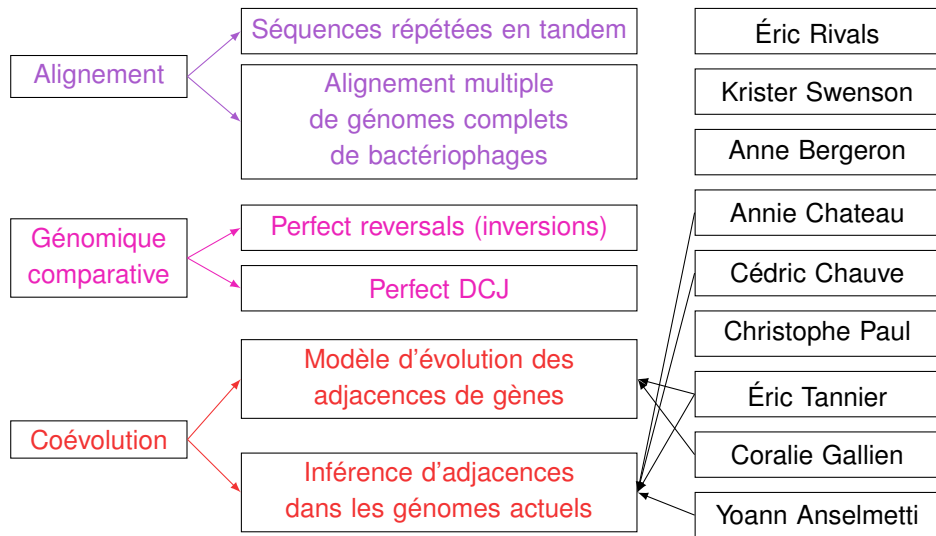
Contributions & collaborations principales



DeCo : structure des génomes ancestraux

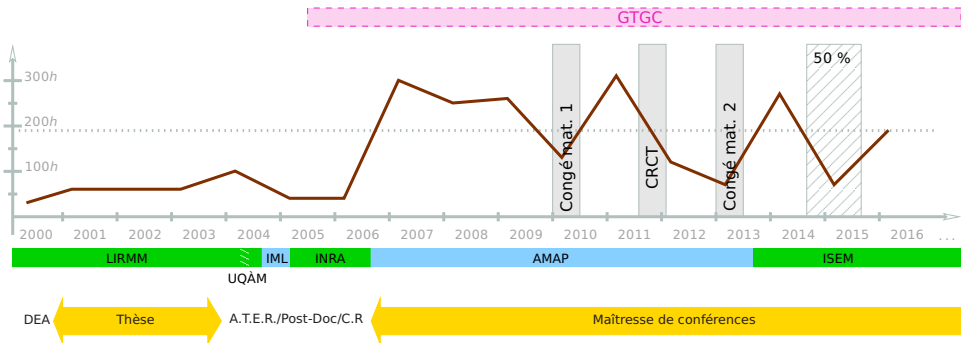
Art-DeCo, AD-Seq, ... : amélioration d'assemblages

Contributions & collaborations principales



DeCo : structure des génomes ancestraux

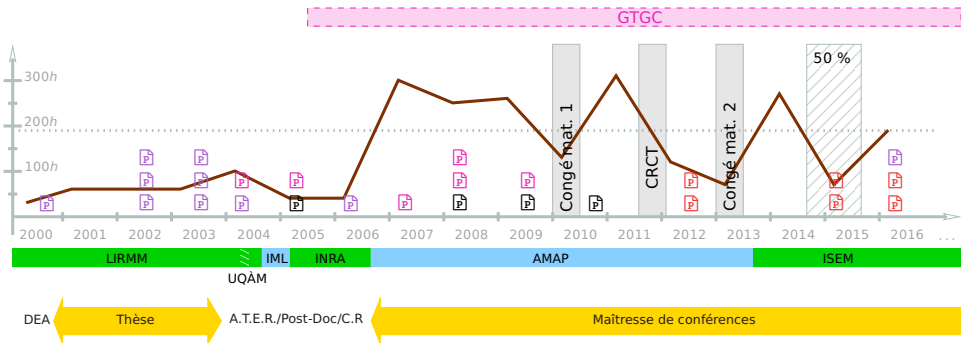
Art-DeCo, AD-Seq, ... : amélioration d'assemblages



Thèmes de recherche

- Alignement
- Génomique comparative
- Coévolution
- Autre

Légende

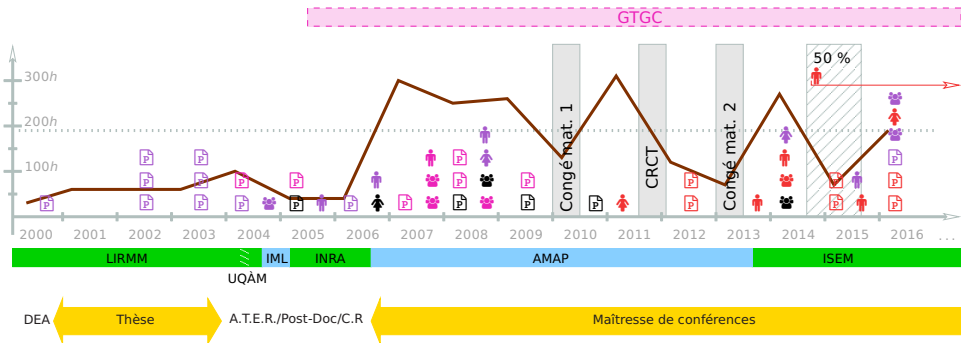


Thèmes de recherche

- Alignement
- Génomique comparative
- Coévolution
- Autre

Légende

- ▶ Publications :

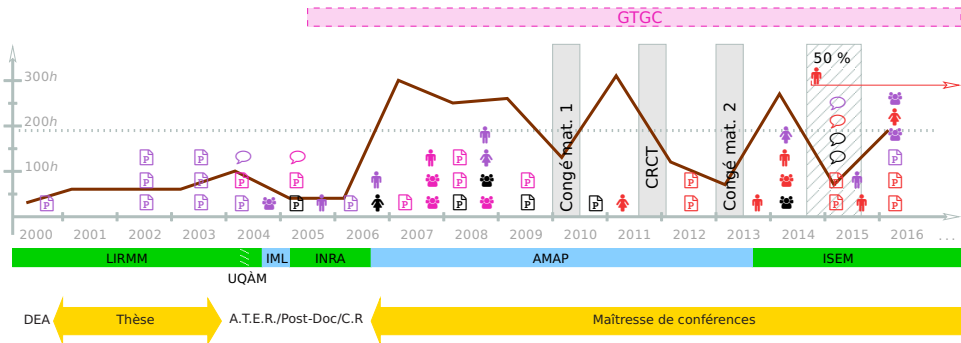


Thèmes de recherche

- Alignement
- Génomique comparative
- Coévolution
- Autre

Légende

- ▶ Publications :
- ▶ Étudiant(e)s : (groupes

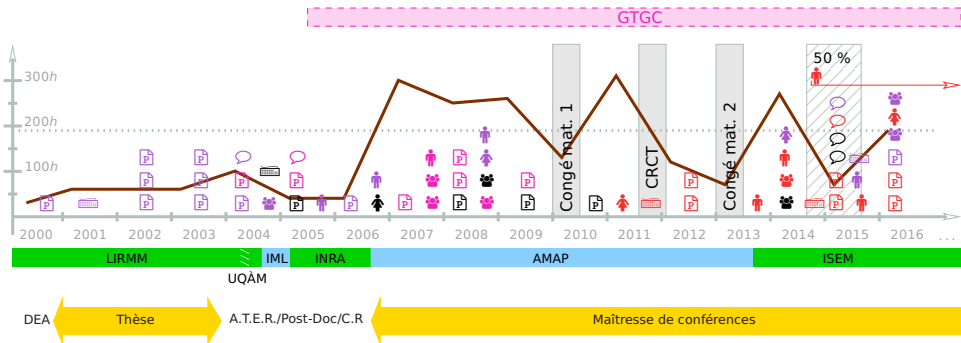


Thèmes de recherche

- Alignement
- Génomique comparative
- Coévolution
- Autre

Légende

- ▶ Publications :
- ▶ Étudiant(e)s : (groupes)
- ▶ Communications orales :

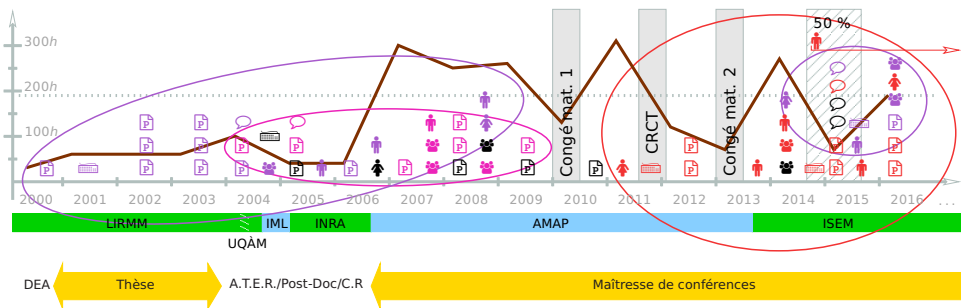


Thèmes de recherche

- Alignement
- Génomique comparative
- Coévolution
- Autre

Légende

- ▶ Publications :
- ▶ Étudiant(e)s : (groupes)
- ▶ Communications orales :
- ▶ Logiciels :



Thèmes de recherche

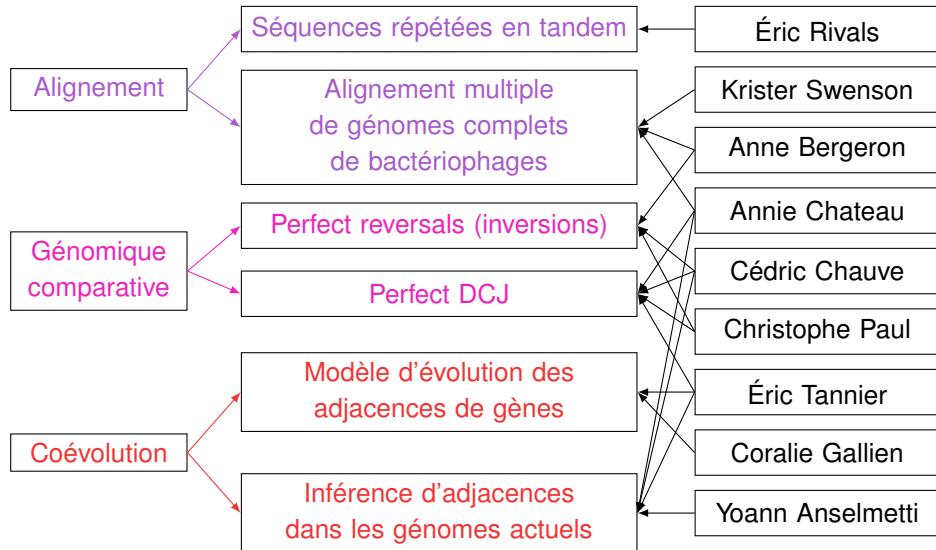
- Alignement
- Génomique comparative
- Coévolution
- Autre

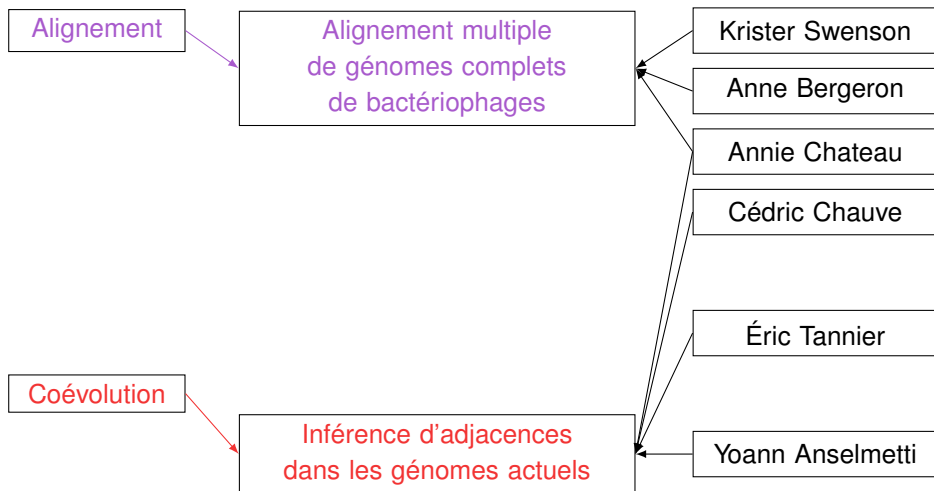
Légende

- ▶ Publications :
- ▶ Étudiant(e)s : (groupes)
- ▶ Communications orales :
- ▶ Logiciels :

- 1 Mon parcours
- 2 Domaine de recherche
- 3 Reconstruction de génomes ancestraux : DeCo
- 4 Synthèse des contributions
- 5 Projets de recherche**
- 6 Merci

Pour la suite





Amélioration des assemblages

Vers des études évolutives avec les réarrangements

Amélioration des assemblages

Vers des études évolutives avec les réarrangements

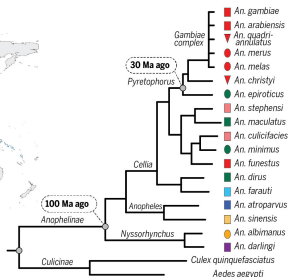
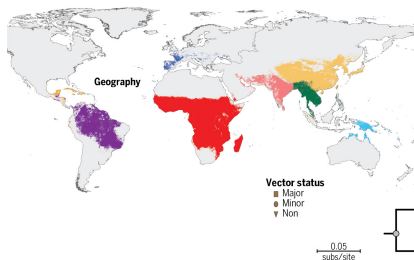
- ▶ Intégration des données de séquençage (AD-Seq)

RESEARCH ARTICLE SUMMARY

MOSQUITO GENOMICS

Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes

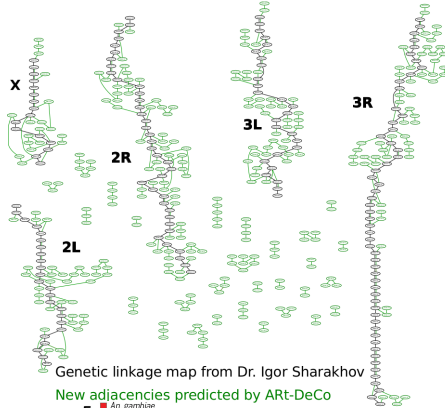
Daniel E. Neafsey,^{1†} Robert M. Waterhouse,^{2* et al.}



Amélioration des assemblages

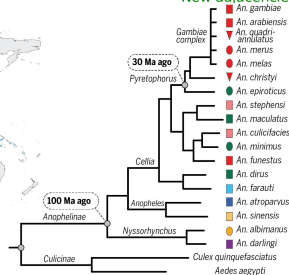
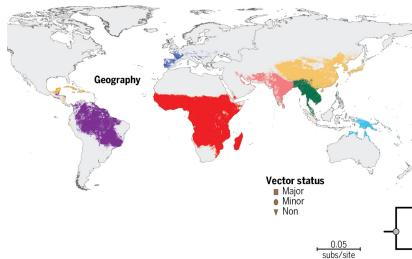
Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques



Genetic linkage map from Dr. Igor Sharakhov

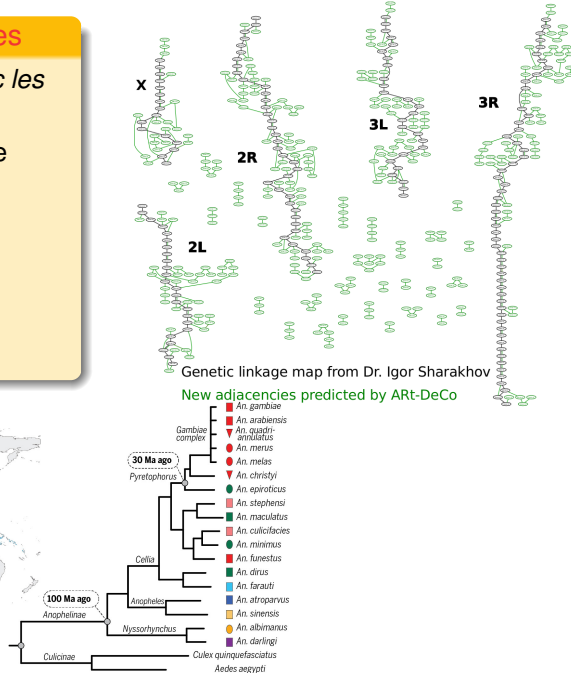
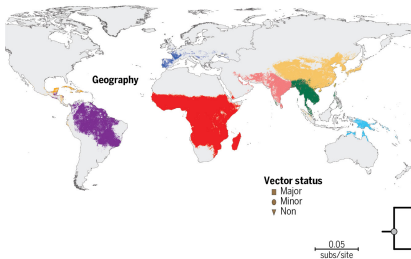
New adjacencies predicted by ART-DeCo



Amélioration des assemblages

Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes



Amélioration des assemblages

Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données



Collaborations :

Benoît Nabholz (ISEM)

Amélioration des assemblages

Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données



Collaborations :

Benoît Nabholz (ISEM)

Thomas Faraut (INRA, Toulouse)

Amélioration des assemblages

Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données

Bactériophages

Mise au point d'outils d'étude adaptés à leur mode d'évolution spécifique



ALPHA : un aligneur de génomes complets de bactériophages

Amélioration des assemblages

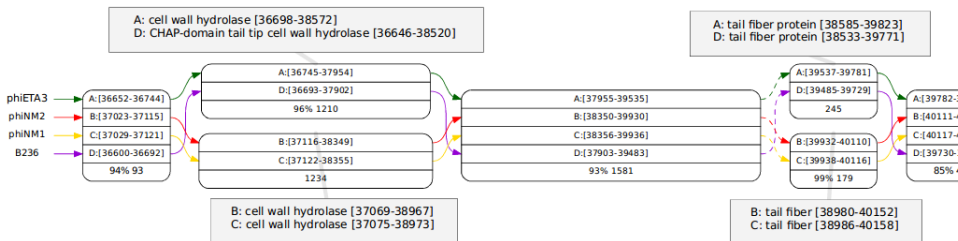
Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données

Bactériophages

Mise au point d'outils d'étude adaptés à leur mode d'évolution spécifique

- ▶ Transfert d'annotations à l'aide d'ontologies



Amélioration des assemblages

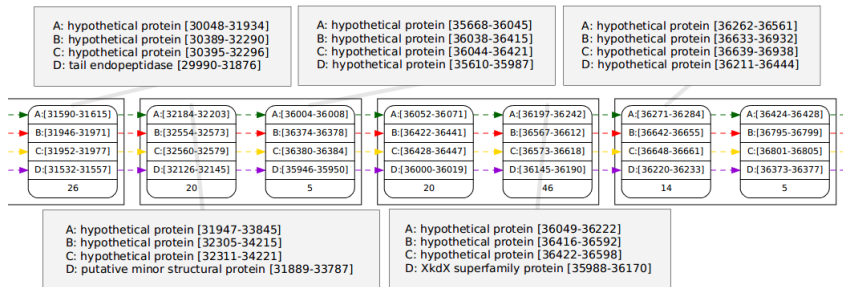
Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données

Bactériophages

Mise au point d'outils d'étude adaptés à leur mode d'évolution spécifique

- ▶ Transfert d'annotations à l'aide d'ontologies



Amélioration des assemblages

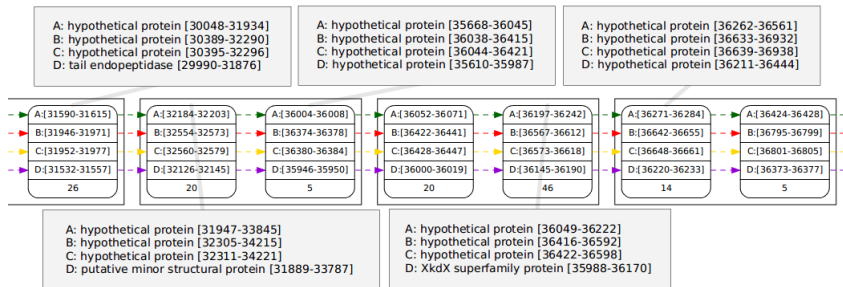
Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données

Bactériophages

Mise au point d'outils d'étude adaptés à leur mode d'évolution spécifique

- ▶ Transfert d'annotations à l'aide d'ontologies
- ▶ Génomes non-colinéaires (inversions, transpositions)



Amélioration des assemblages

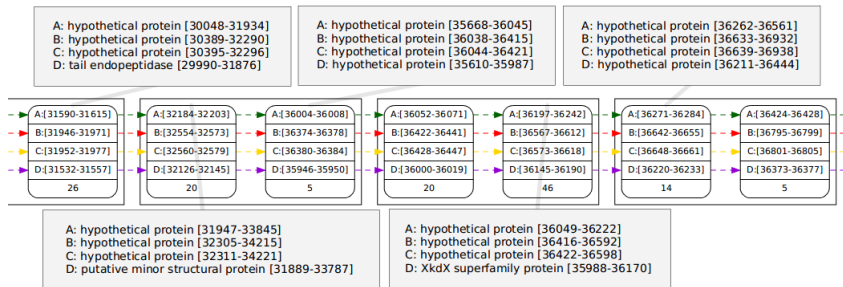
Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
- ▶ Cartes physiques
- ▶ Remise en question des adjacences existantes
- ▶ Nouveaux jeux de données

Bactériophages

Mise au point d'outils d'étude adaptés à leur mode d'évolution spécifique

- ▶ Transfert d'annotations à l'aide d'ontologies
- ▶ Génomes non-colinéaires (inversions, transpositions)
- ▶ Distance évolutive (identification/phylogénie)



Amélioration des assemblages

Vers des études évolutives avec les réarrangements

- ▶ Intégration des données de séquençage (AD-Seq)
 - ▶ Cartes physiques
 - ▶ Remise en question des adjacences existantes
 - ▶ Nouveaux jeux de données
-
- ▶ Stage M2 Info *janvier-juin 2017*
 - ▶ Demande de bourse de thèse *oct. 2017 - oct 2020*
 - ▶ Soutenance de thèse de Yoann *novembre 2017*

Bactériophages

Mise au point d'outils d'étude adaptés à leur mode d'évolution spécifique

- ▶ Transfert d'annotations à l'aide d'ontologies
 - ▶ Génomes non-colinéaires (inversions, transpositions)
 - ▶ Distance évolutive (identification/phylogénie)
-
- ▶ Visite d'Anne Bergeron à l'ISEM *février-mars 2017*
 - ▶ Rapprochement avec l'équipe de M. Hochberg et O. Kaltz (ISEM)
 - ▶ Post-doc à l'interface

Liens et convergence

- ▶ Revenir à un “bas niveau” de données

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

et dans 10 ans ?

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

et dans 10 ans ?

- ▶ Technologie de séquençage de molécules entières ?

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

et dans 10 ans ?

- ▶ Technologie de séquençage de molécules entières ?
 - ADN ancien toujours fragmenté (*e.g. Yersinia pestis*, peste noire)

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

et dans 10 ans ?

- ▶ Technologie de séquençage de molécules entières ?
 - ADN ancien toujours fragmenté (*e.g. Yersinia pestis*, peste noire)
 - Expériences d'évolution de phages

Liens et convergence

- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

et dans 10 ans ?

- ▶ Technologie de séquençage de molécules entières ?
 - ADN ancien toujours fragmenté (*e.g. Yersinia pestis*, peste noire)
 - Expériences d'évolution de phages
- ▶ Hétérogénéité des données

Liens et convergence

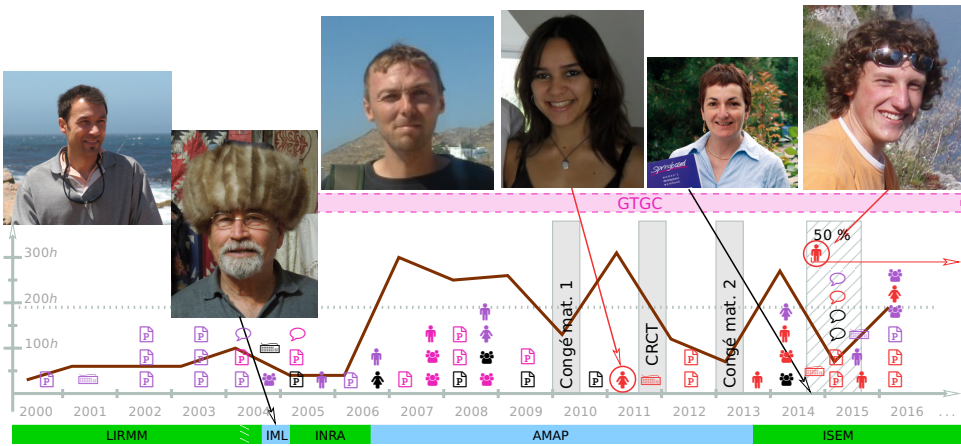
- ▶ Revenir à un “bas niveau” de données
- ▶ Intégrer plusieurs niveaux de données
- ▶ Produire des solutions concrètes à des problèmes biologiques

et dans 10 ans ?

- ▶ Technologie de séquençage de molécules entières ?
 - ADN ancien toujours fragmenté (*e.g. Yersinia pestis*, peste noire)
 - Expériences d'évolution de phages
- ▶ Hétérogénéité des données
- ▶ Attention à la qualité des données (contamination, ...)

Sommaire

- 1 Mon parcours
- 2 Domaine de recherche
- 3 Reconstruction de génomes ancestraux : DeCo
- 4 Synthèse des contributions
- 5 Projets de recherche
- 6 Merci**



Séverine Bérard