

1. Définition

La génomique comparative peut aider à l'identification de la fonction des gènes car deux gènes issus d'une même ancêtre commun ont des chances d'avoir la même de ancêtre un histoire évolutive.

2. Relations d'homologie entre les gènes

2.1. Homologie et identité de séquence

La comparaison de séquences avec des logiciels de comparaison de séquence type BLAST conduit à mesurer pourcentage d'identité entre les séquences nucléotidiques ou d'acides aminés pour les séquences d'acides aminés, on mesure également un pourcentage de similarité qui tient compte. Autre l'identité des acides aminés, du fait que les acides aminés appartiennent à une même famille selon que leur chaîne latérale est acide amines appartiennent à une même famille selon que leur chaîne latérale est acide basique polaire neutre. Non polaire cette identité et cette similarité sont quantifiables.

Si deux séquences présentent des identités ou similarités de séquence cela peut traduire :

- ✚ Qu'elles ont une même origine phylogénétique. On dit alors que les séquences sont homologues, c'est à dire qu'elles sont issues d'une séquence ancestrale commune (figure 7.1) contrairement à la notion d'identité ou de similarité (deux séquences peuvent être similaires à 55 %) la notion d'homologie n'est pas quantifiable (il existe ou il n'existe pas de séquences ancestrales communes).
- ✚ Qu'elles ont convergé bien qu'elles n'aient pas d'ancêtre commun ceci peut être du ou fait du hasard (dans des régions de faible complexité. C'est-à-dire avec des nucléotides ou acides aminés très répétés) ou d'une pression de sélection commune

On considère en général qu'il y a homologie entre deux séquences. Si leur similarité est supérieure à 30%.

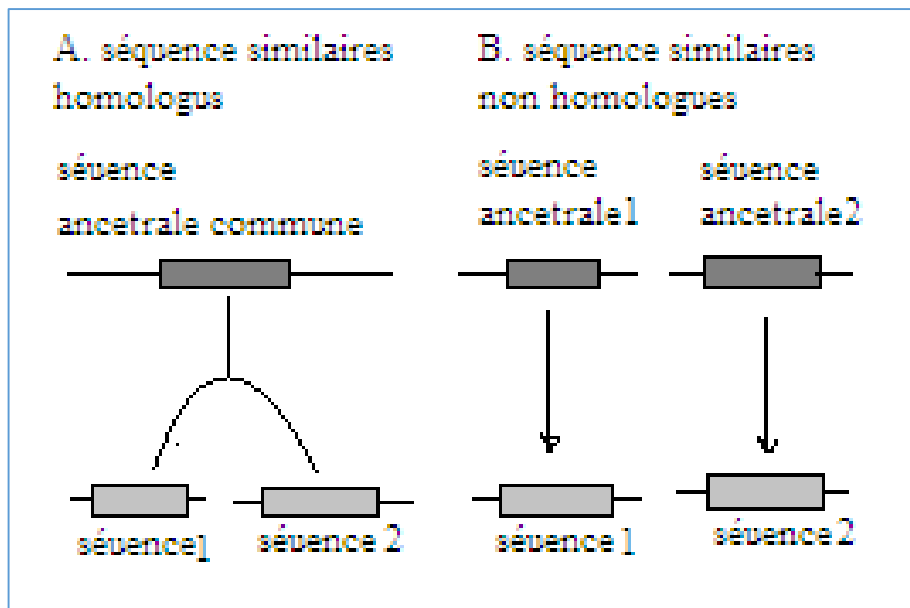


Figure 7.1 liens entre identité de séquences et homologie.

2.2. Homologie et origine ancestrale

Ou sient des gènes homologues, selon la nature de l'évènement évolutif

qui les a séparés, on décrit trois sous-classes (figure 7.2).

- A. **Les gènes orthologues:** ce sont des gènes ayant évolué à partir d'un même gène ancestral suite à des évènements de spéciation (séparation d'espèces) souvent la fonction des gènes orthologues est conservée au cour de l'évolution.
- B. **Les gènes paralogues:** ce sont des gènes issus d'évènement de duplication ou sien d'un ou plusieurs paralogues peut subir des modifications de séquence telles que sa fonction en deviennent différentes. Il peut également perdre sa fonction et devenir un pseudo-gène (gène inactif).
- C. **Les gènes xénologues:** ce sont des gènes ayant été acquis par transfert génétique horizontal (transfert de gènes entre bactéries) souvent les gènes xénologues confèrent à la bactérie receveuse une nouvelle fonction qu'elle ne possédait pas initialement.

On parle également de protéines orthologues, paralogues ou xénologues se l'on qu'elles sont codées par des gènes orthologues ou xéntologues.

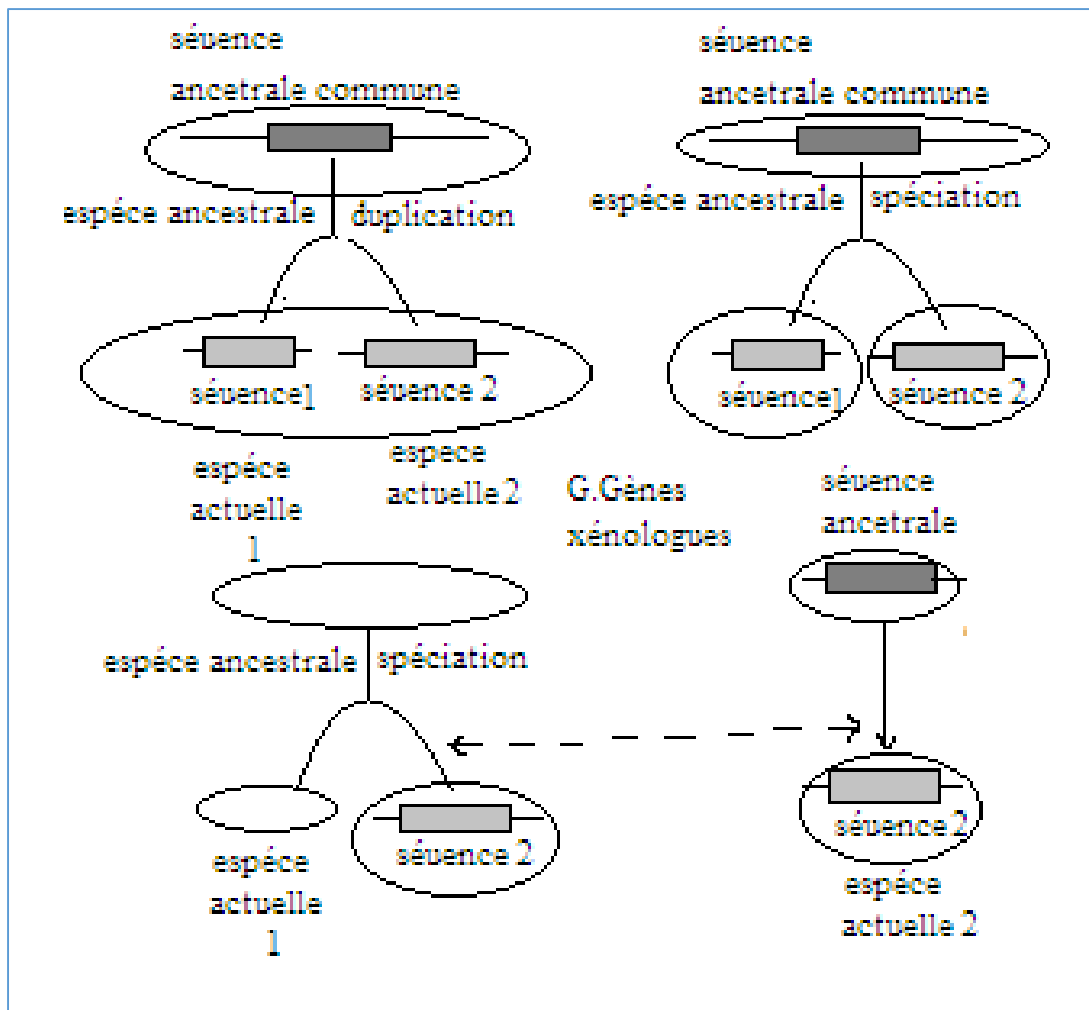


Figure 7.2 : les différentes classes.

3. Relation de synténie entre les génomes

A. Définitions de la synténie

La synténie est la présence de deux ou plusieurs gènes sur le même chromosome. En génomique comparative, ce terme fait référence à la conservation de l'ordre des gènes entre différentes espèces, indiquant une origine évolutive commune. Par exemple, si les gènes A, B et C sont dans cet ordre sur un chromosome chez deux espèces, on parle de synténie entre ces deux espèces.

4. Plateformes informatiques d'étude de la synténie

Les plateformes informatiques pour étudier la synténie comprennent des outils de bio-informatique qui effectuent des comparaisons génomiques pour identifier les blocs de gènes conservés. Ces plateformes utilisent des logiciels pour visualiser les alignements de séquences et intégrer des données afin d'analyser les relations évolutives des familles de gènes, ce qui peut

être utilisé pour l'annotation fonctionnelle et la compréhension de l'évolution des génomes. Elles s'intègrent souvent dans de plus grandes infrastructures de recherche, comme celles de l'AGAP Institut ou de France Génomique

5. La synténie comme aide à l'annotation

La synténie est l'utilisation de la conservation de l'ordre des gènes entre espèces pour faciliter l'annotation. En comparant les blocs de gènes conservés (blocs synténiques), on peut déduire des informations sur la fonction de gènes inconnus chez une espèce, en transférant les connaissances acquises sur leurs gènes orthologues chez d'autres espèces. Par exemple, si une série de gènes est conservée dans le même ordre et avec des fonctions similaires dans un organisme modèle et chez l'homme, on peut prédire les fonctions des gènes humains correspondants

6. La synténie comme marqueur d'événements chromosomiques

Le degré de synténie est inversement proportionnel au temps écoulé depuis la divergence à partir du génome ancestral. Les moteurs de cette divergence sont les remaniements génomiques de différentes natures :

- Fusions chromosomiques
- Cassures chromosomiques (échange de segments de séquences entre des chromosomes différents)
- Inversions chromosomiques
- Délétions (perte d'une région chromosomique)
- Duplication d'une région chromosomique

7. La comparaison de l'organisation chromosomique: (identification des relations d'homologue entre génomes)

Il existe plusieurs méthodes pour comprendre les relations d'homologue et la comparaison entre les génomes, et cela à l'aide d'un outil bioinformatique.

1-La bioinformatique: La bioinformatique est une composante essentielle de la génomique comparative à la fois pour le développement de bases de données, les développements méthodologiques (algorithmique et représentation des résultats) mais aussi pour la réalisation

d'analyses fines. Dans ce domaine on utilise le serveur web, leur architecture est divisée en 3 niveaux: la base de données relationnelle contenant les informations génomiques, l'application chargée du calcul des groupes de synténie, et une série d'interfaces web dédiées à l'exploration des résultats.

2- Quelques méthodes de comparaison

a. L'alignement réciproque de séquences

La méthode la plus généralement utilisée consiste à réaliser des alignements de séquences réciproques entre deux génomes (méthode du “Best Bidirectional Hits” (BBH) ou “meilleurs alignements réciproques”). L'idée sous-jacente de la méthode implique que les séquences de gènes orthologues présentent plus de similitudes entre elles que de similitudes avec une séquence du génome initial.

b. La méthode de regroupement

La méthode des “COGs” pour Cluster of Orthologous Groups” consiste à identifier, non seulement des paires de gènes orthologues entre deux génomes, mais plus généralement, des groupes de gènes orthologues entre plusieurs génomes selon les étapes suivantes :

- Toutes les séquences de gènes sont alignées entre les génomes considérés afin d'identifier l'ensemble des relations d'homologie.
- Puis, la détection des paralogues est réalisée par des alignements de la séquence d'un gène sur son propre génome.
- Enfin, la dernière étape consiste à rassembler les groupes de gènes (orthologues et paralogues) identifiés dans les deux premières étapes pour former des groupes de gènes homologues.

8. Les limites de synténie:

En pratique, la recherche de synténie nécessite préalablement d'identifier les familles de gènes orthologues entre tous les génomes comparés. Les principales limitations à la recherche de synténie sont en réalité liées à cette étape d'identification des orthologues. Les méthodes traditionnellement utilisées pour identifier les orthologues atteignent en effet leurs limites quand les séquences divergent. On peut considérer uniquement les orthologues qui limitent la

synténie, car cette dernière peut être dupliquée au sein d'un même chromosome, donc la formation ce qu'on appelle les paralogues.