



ALGORITHMES ET SEQUENCES D'ADN

VERVIER Jérôme

Département d'Informatique

On fait de plus en plus appel aux techniques et moyens informatiques pour faciliter et accélérer le travail de recherche à partir des informations récoltées par des expériences biologiques. La quantité d'information à analyser est immense, par exemple : le génome humain comporte 3,2 milliards de paires de bases (représentant près de 30.000 gènes)

Trois problèmes sont considérés dans ce travail :

- Cartographie de séquences d'ADN
- Réarrangement génomique
- Recherche de séquences d'ADN dans un génome

A chacun de ces problèmes s'offrent des solutions informatiques utilisant des algorithmes précis qui vont réaliser le plus rapidement possible (on l'espère) la recherche et/ou l'analyse des données. On va chercher à aller de plus en plus vite non seulement en perfectionnant les machines qui les exécutent mais aussi en perfectionnant ces algorithmes eux-même dans leur manière de travailler.

Bien souvent, plusieurs solutions (algorithmes) sont disponibles pour un même problème.



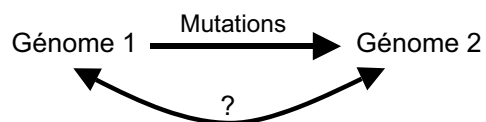
Cartographie de séquences d'ADN

Vu la taille et le nombre de nucléotides composant l'ADN humain, il est très pratique de réaliser une cartographie de ce dernier. On va donc chercher à connaître la position de séquences connues dans le génome humain pour faciliter la recherche.



Réarrangement génomique

Notre génome, avant d'arriver à ce qu'il est aujourd'hui, a subi des mutations : il est intéressant de savoir lesquelles peuvent s'être déroulées pour mieux comprendre notre organisme.



Recherche de séquences d'ADN

Connaissant une séquence de nucléotides précise, il est parfois nécessaire de rechercher tous les endroits dans une longue séquence d'ADN où cette courte séquence apparaît.

```

...ATGTCGATCGTAGCTAGGCTAGTTACGATCGA
CTGACTAGCACGTAGCTAGCTAGAGTCGGATCG
GATCGGATTTCGATCGTAGATCGGATTTCGGATCG
ATCGATCGGATCGATCGCTCATGATCAGCTCAGC
TAGACCTAGCTATGACTCTCGATCAGCGATCCTC
AGCTCTAGCTTAGCATCTAGCTACGGATCGACG
ATCGAGTTCGAGTTCGACGATCGATCGATCGATCG
ATCGATCGGATCGTACGTACGTAGCTAGCTAGC
GATCGATCATCGAGTCGATAGCTACGTACGTA...

```