

## 1- Introduction

L'information héréditaire définissant les caractéristiques biologiques d'un organisme eucaryote est portée par ses chromosomes et est codée dans l'ADN. Cet ADN se compose d'un enchaînement de quatre types de molécules appelées **nucléotides**: l'adénine (A), la guanine (G), la cytosine (C) et la thymine (T). Comme une série de lettres forme un mot, certaines séquences de nucléotides ont une signification fonctionnelle et constituent les **gènes**. La traduction des ADN codants en protéines passe par une phase de transcription de l'ADN en ARN. Les protéines ainsi synthétisées remplissent des rôles variés de structure, de transport, de communication, de défense, etc. D'autres gènes spécifient la synthèse de divers ARN impliqués dans la machinerie de traduction ou la régulation de l'expression génique. L'ensemble du patrimoine génétique d'un individu constitue son **génom**e. Toutes les cellules d'un organisme possèdent le même génome, mais cette information s'exprime différemment selon les tissus, sous l'effet de divers mécanismes de régulation. L'ADN peut subir des changements (mutations) sous l'effet, par exemple, d'agents extérieurs (UV, radioactivité, produits chimiques...). Les mutations ponctuelles modifiant un seul nucléotide sont appelées **polymorphisme SNP** (Single-Nucleotide Polymorphism). Les mutations créent des versions différentes d'un même gène (allèles) et peuvent parfois en modifier la fonction. Les mutations apparues dans les cellules germinales sont transmises à la descendance, contrairement aux mutations somatiques. L'ensemble des allèles constitue le **génotype**. L'interaction génotype et environnement conditionne les caractéristiques de l'organisme (**phénotype**).

## 2- Historique

Un bref regard en arrière souligne la vitesse avec laquelle notre connaissance des génomes s'est développée, ce qui peut laisser penser que les progrès futurs seront tout aussi spectaculaires.

- **1953** : Watson et Crick découvrent la structure de l'ADN
- **1956** : F. Sanger élabore le protocole de séquençage des protéines
- **1977** : F. Sanger élabore le protocole de séquençage de l'ADN
- **Années 80** : Technique de séquençage de l'ADN peu répandue,
- **1990** : lancement officiel du programme international du séquençage du génome humain
- **1991** : apparition de la première "puce à ADN" construite par la Société AFFYMETRIX

- **1995** : Séquençage du 1<sup>er</sup> génome procaryote *Haemophilus influenzae* (1,83 Mb)
- **1996** : Séquence du 1<sup>er</sup> génome eucaryote *Saccharomyces cerevisiae* (12 Mb)
- **1997** : séquençage du génome de la bactérie *Escherichia coli*
- **1998** : séquençage du génome du nématode *Caenorhabditis elegans*
- **2000** : séquençage des génomes de la drosophile, de la souris et de la plante *Arabidopsis*
- **2001** : premières données sur le séquençage complet du génome humain
- **2005** : 1740 projets génomes existent parmi lesquels les projets de génomes modèles.

De l'ensemble de ces avancées, on peut retenir quelques grandes idées parfois surprenantes :

L'homme partage **40%** de ses gènes avec les plantes, plus de **80%** avec les mammifères et **98,5%** avec les chimpanzés, et du point de vue génétique, les êtres humains sont à **99,9%** identiques. C'est-à-dire deux êtres humains diffèrent par 3 millions de bases nucléotidiques sur 3 milliards au total ; et cette faible différence (**0,1%**) autorise la détermination des empreintes génétiques à partir de l'ADN. Ces empreintes individuelles ont de multiples applications comme la recherche de paternité et la recherche en criminologie chez l'homme. Environ la moitié du génome humain est constitué de longues séquences répétitives sans fonction apparente. Seulement **5 à 10%** du génome correspond à des gènes qui peuvent être exprimés. L'étude du transcriptome ne s'adresse donc qu'à une toute petite partie du génome, mais certainement la plus importante car fonctionnelle.

### 3- Apparition de la génomique

La génétique (**étude des gènes**) consiste à étudier la transmission des caractères héréditaires d'une génération à l'autre. **La génétique directe** est l'étude génétique qui commence par l'observation d'une variation visible dans la morphologie ou la physiologie, c'est-à-dire elle commence par la recherche d'une différence génétique qui provoque une différence phénotypique.

Avec le développement des connaissances sur l'ADN et sur la façon dont il code l'information de la séquence d'acides aminés des protéines ainsi que le contrôle cellulaire de la production de celles-ci, une autre forme d'étude génétique devient possible c'est l'étude basée sur l'ADN ou bien **la génétique inverse**. Le programme de la génétique inverse débute par des

changements génétiques connus ensuite la recherche des modifications qui en résultent dans l'organisme.

Par exemple, à l'aide d'une séquence normale d'ADN, on peut lire la séquence d'acides aminés de la protéine synthétisée (correspondance code d'ADN/ AA codés), l'ADN peut ensuite être modifié de façon spécifique (en lui insérant une mutation aboutit à un blocage total de la production de la protéine ou à un changement de la protéine, c'est-à-dire modification ou suppression de son activité métabolique). Les conséquences d'une telle mutation sur le développement du phénotype peuvent ensuite être suivies de la même façon que les effets mutationnels dans le programme de génétique directe.

L'un des avantages de la génétique inverse est qu'elle permet la création sur mesure d'un grand nombre de mutations spécifiques.

Les génomes de nombreux organismes ont fait l'objet de plusieurs recherches au niveau moléculaire grâce aux applications des techniques appropriées, ce qui a donné lieu à une nouvelle discipline au sein de la génétique appelée **génomique**

#### 4- Définition de la génomique

La génomique est l'étude de l'ensemble des gènes d'une espèce (génome), de leur disposition sur les chromosomes (établissement de cartes génétiques et physiques), de la description fine de leur structure et de leurs fonctions. Elle comporte :

- 1) **la génomique structurale** qui étudie l'architecture et l'évolution des génomes et dont l'objectif principal est l'obtention de la séquence du génome ;
- 2) **la génomique fonctionnelle** qui vise à déterminer la fonction et l'expression des gènes en caractérisant l'ensemble des transcrits (transcriptome) et des protéines (protéome), ainsi que les éléments de leur régulation
- 3) **la protéomique** qui analyse le protéome (protéine). Elle contribue à l'annotation des génomes et à l'identification des séquences informatives. Ces trois approches sont complémentaires.

#### 5- Génomique Végétale

Elle permet d'identifier ce que l'on appelle les **gènes d'intérêt** intervenant dans :

- la résistance aux pathogènes (virus, bactéries, champignons),
- la qualité (saveur des fruits et des légumes, composition en acides gras ou en vitamines, qualité boulangère...),
- les performances agronomiques (précocité, architecture, économies d'intrants, rendement...).

La génomique permet aussi d'approfondir notre connaissance des principales voies métaboliques dans les plantes et de rationaliser la sélection de nouvelles variétés en combinant les caractères les plus performants dans des génotypes élites ; et d'exploiter la diversité existant entre tous les individus d'une même espèce et les formes sauvages apparentées.

### 6- Complexité des génomes des plantes

La diversité du règne végétal reflète la grande complexité des génomes des plantes :

- Taille parfois considérable,
- Duplications partielles ou complètes,
- Addition de génomes issus d'espèces apparentées, fruits d'une évolution naturelle depuis des millions d'année mais aussi des différentes formes de sélection (domestication) pratiquées par l'homme au cours des siècles.

Tout ceci nécessite la mise en place de méthodes et d'outils spécifiques ; c'est pour ça, les chercheurs ont choisi de travailler sur des espèces dites **modèles**, à génome de taille relativement réduite, comme l'arabette (*Arabidopsis thaliana*) et le riz.

L'intégration des données obtenues sur ces espèces modèles et leur extrapolation à l'étude des espèces d'intérêt agronomique, comme le blé, le maïs ou le colza, imposent l'utilisation de techniques de biologie moléculaire à haut débit et le développement d'outils bioinformatiques puissants.

### 7- Objectifs de la génomique végétale

- Séquencer les génomes
- Établir la cartographie des génomes
- Analyser un grand nombre d'organismes possibles
- Etudier l'organisation des gènes et des séquences régulatrices
- Etudier l'expression des gènes
- Etudier la structure et la fonction des protéines et des ARN pour lesquelles les gènes codent
- Utiliser ces informations dans les études du métabolisme (métabolome), des interactions entre les macromolécules biologiques (interactome), dans le développement...etc.
- Caractériser la fonction des milliers de gènes grâce à des approches globales (transcriptome : étude globale de l'expression du génome ; protéome et métabolome),

les réseaux d'interaction entre eux en réponse à divers stress abiotiques (exp sécheresse) ou biotiques (exp attaque par un champignon phytopathogène).

- Déterminer la fonction des gènes par la recherche de ses activités, de domaines fonctionnels ou de gènes partenaires.
- Établir les profils d'expression des gènes par comparaison entre tissus sains et pathologiques.

### **8- Champs d'applications de la génomique**

1- La génomique est utilisée dans le cas des maladies infectieuses :

- recherche de gènes de résistance aux antibiotiques,
- recherche et identification de gènes impliqués dans la pathogénicité

2- La génomique est utilisée dans le cas de maladies génétiques :

- recherche de gènes responsables de susceptibilité génétique aux maladies
- recherche de modifications de l'ADN à l'origine du cancer
- recherche de modification des profils d'expression des gènes, à l'origine de certains dysfonctionnements cellulaires.

3- La génomique est utilisée dans l'industrie :

- recherche de gènes impliqués dans les pathologies (cibles pharmaceutiques, marqueurs diagnostiques)
- recherche de gènes codant des molécules d'intérêt
- recherche de gènes responsables de résistances
- recherche de gènes permettant de comprendre les mécanismes clés (cancer, vieillissement, etc)