

# AI for Health

## Chroniques AI for Health - Série des « Omiques »

### La génomique : voyage au centre de nos gènes

Temps de lecture : 10 mn

L'**ADN** ? Ça vous parle ?

L'**ARN** ? Ça vous parle également ?

Les **protéines** aussi ?

Attention ça se corse : vous connaissez aussi la notion de **métabolites** ? Si c'est le cas, on vous tire notre chapeau c'est un 20/20 !

ADN, ARN, protéines et métabolites : voici 4 grands acteurs à l'origine de nos caractères apparents - *appelé phénotype* - à l'échelle moléculaire, cellulaire et de l'organisme. Pour un organisme humain par exemple, la couleur des yeux, le niveau de pilosité, les empreintes digitales ou la taille sont autant d'exemples de phénotypes !

L'étude des « **omiques** », c'est en fait un regroupement de disciplines qui s'intéressent à différentes étapes clés de la constitution d'un phénotype. La génomique liée à l'ADN notamment, la transcriptomique pour l'ARN, la protéomique pour les protéines et la métabolomique pour les métabolites.

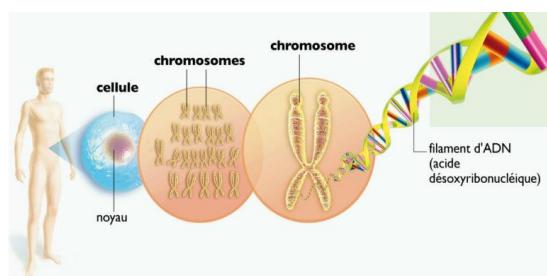
Bienvenue dans la première édition de la série des « omiques » où notre objectif est de vous aider à comprendre les concepts fondamentaux de l'étude de la chaîne de valeurs de la vie !

#### Partie 1 : l'ADN.

Vous êtes-vous déjà posé la question de comment, via une molécule de l'ordre du nanomètre (1 milliardième de mètre), des organismes ont pu voir le jour ? Plantes, animaux, bactéries, champignons, ... Autant de règnes du vivant avec une molécule à l'origine : de quoi faire tourner la tête !

L'ADN, c'est quoi ?

L'**Acide DésoxyriboNucléique**, de son petit nom « ADN », est une molécule qui ressemble à ceci :



Vous avez peut-être déjà vu ces images il y a (très) longtemps. Imaginez l'ADN comme une immense bobine de fil double. Déroulé, l'ADN s'étend sur 2 mètres environ. Enroulé, sous forme de chromosome, l'ADN tient dans le noyau d'une cellule.

# AI for Health

**Fun fact** : si on met bout à bout l'ensemble de l'ADN déroulé de nos cellules, on peut parcourir plus de 250 000 fois la distance Terre-Lune !

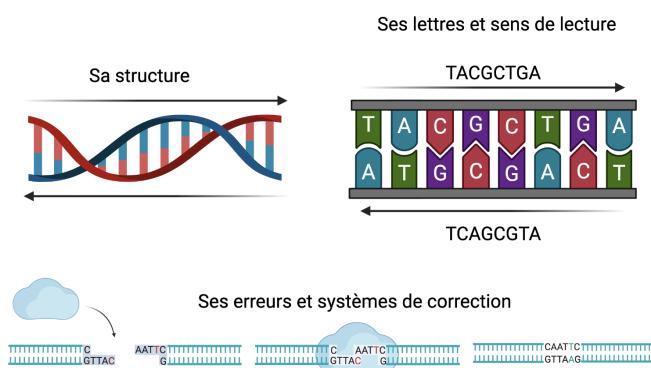
## Le langage de la vie

Mais comment l'ADN peut-il être à la base de notre développement ?

Voyons l'ADN et toute cette notion de vie comme un **langage**.

L'ADN a :

- **Ses lettres** : il est constitué d'autres molécules appelées "**bases azotées**" regroupées en paires : Adénine, Thymine, Cytosine et Guanine : ATCG. *Comme dans le film "Bienvenue à Gattaca"\** !
- **Ses règles et mots** :
  - A se lie à T et C à G
  - Un certain ordre des lettres a sa signification et entraînera des mécanismes particuliers



des erreurs comme des mutations soit par un échange de lettre, un oubli ou encore un ajout, des systèmes dits de réparation permettent de corriger les erreurs !

- **Son sens de lecture** : de gauche à droite sur un brin et de droite à gauche sur l'autre !

- **Ses variations** : en fonction de l'organisme étudié, les règles et mots peuvent changer (par exemple, l'ADN bactérien possède des spécificités que n'a pas celui de l'humain)

- **Ses corrections** : s'il y a

## Expression génique : le pouvoir du langage

Et que manque-t-il à un langage pour qu'il se diffuse et que des actions puissent se mettre en place ? **Il faut qu'il puisse s'exprimer**.

Nous l'avons vu, l'ADN est très long. Cependant, **moins de 2 %** seulement de sa séquence correspond à des parties dites "codantes pour des protéines" : **les gènes**. Dans le génome humain, 21 000 gènes environ ont été recensés.

Les gènes, du grec "genos" (naissance, origine), sont donc les séquences à l'origine de cette fameuse cascade dont on parlait au début : ADN (via les gènes), ARN, protéines puis métabolites.

\*Bienvenue à Gattaca est un film de SF de 1997 traitant notamment des modifications de l'ADN pour créer des humains dits parfaits. Gattaca est formé des 4 lettres A,T,G,C. - ndlr

# AI for Health

Pour boucler la boucle de la métaphore sur le langage, le processus à l'origine de tout cela est appelé **“expression génique”** et l'ensemble des gènes d'un organisme est appelé “génome”.

Partons enfin d'un postulat : chaque cellule possède le même ADN.

La même information de départ est présente dans toutes nos cellules et est strictement identique d'une cellule à l'autre. Alors, comment expliquer autant de différences structurelles et fonctionnelles entre nos cellules puis les différentes parties de notre corps ? Comment expliquer que des cellules de foie soient différentes de celles du cœur ? Comment un œil a-t-il cette forme avec ces propriétés et pas d'autres ? Pourquoi est-ce que je suis plus sensible aux allergies que d'autres ?

## La génomique

Nous entrons dans le cœur du sujet ! La science qui étudie et se sert du génome pour alimenter les connaissances, innover dans divers secteurs et développer des solutions d'avenir est appelée **génomique**. Cousine de la génétique, la génomique a un caractère d'étude des gènes comme un tout alors que l'autre science se focalise principalement sur l'étude de gènes en particulier, leur hérédité et leur association à des maladies ou traitements.

La génomique est divisée en **deux branches** complémentaires :

### 1. Génomique structurale

Tout langage doit posséder un lexique pour permettre à tout le monde de s'y retrouver. En génomique structurale, des techniques sont utilisées pour séquencer le génome, c'est-à-dire analyser la séquence qui compose les gènes et les annoter, donc donner des informations pratiques et des clés de lecture. Cela permet d'alimenter les connaissances générales sur le génome et de les affiner, mais c'est également une étape clé pour d'autres secteurs car cela constitue le point de départ de la compréhension fine de la fonction des gènes.

### 2. Génomique fonctionnelle

Cette partie, fortement liée à la génomique structurale, permet d'approfondir la connaissance du rôle des gènes en relation avec leurs fonctions. Elle permet également d'étudier la fonction des gènes, leur régulation et les interactions du résultat de leur **expression** : ARN et protéines.

La génomique permet, entre autres, de déterminer des différences dans le génome de certains individus par rapport à celui de l'espèce. Par exemple, en comparant le génome de M. ou Mme X au génome général de la population humaine - *Homo Sapiens Sapiens* - nous pouvons trouver des particularités qui

# AI for Health

dans plusieurs cas n'ont pas de conséquence sur le phénotype, mais dans d'autres, peuvent être révélateurs de certaines prédispositions à des maladies ou à l'inverse, de résistances à d'autres.

La génomique est à l'interaction de nombreuses autres disciplines de la biologie comme la biologie moléculaire, la génétique, l'immunologie,... En prenant en compte le contexte, l'état d'un patient et son profil génomique, il est possible de déterminer les facteurs de risques de développer telle ou telle maladie mais également d'aiguiller vers une thérapie plus ciblée.

Pour faire simple, **la génomique est une science de compréhension, d'orientation et de prédition** à toutes les étapes : recherche fondamentale, préclinique ou clinique. Elle permet d'analyser les "briques" du vivant pour comprendre ses fondations et anticiper sa structuration lors des étapes suivantes.

## IA en génomique : Deepmind, l'IA au service de la prédition

Prédire l'expression des gènes présente un fort intérêt pour la communauté scientifique. **Deepmind** a mis en place une architecture de réseaux neuronaux **"Enformer"** permettant une forte précision dans la prédition de l'expression des gènes à partir de séquences d'ADN. Cela permet de compléter les approches existantes pour mieux comprendre la régulation des gènes et les facteurs de causalité de certaines maladies.

## Focus Startup : WhiteLab Genomics, la startup qui innove grâce aux données génomiques

Une de nos startups membres, **WhiteLab Genomics**, agrège la connaissance des bases de données publiques issues des séquences de l'ADN pour trouver de nouvelles façons d'optimiser des thérapies innovantes comme la thérapie génique. Par le biais notamment de la création d'un atlas des gènes spécifiques de tissus ou organes, la startup participe à orienter des thérapies vers le bon tissu et au bon moment, tout en limitant les risques de rejet immunitaire.

Pour réaliser cela, les données utilisées sont notamment issues de bases de transcriptomique. La génomique vous a intéressé ? Ne manquez pas la suite de nos articles et notamment ceux de la série des omiques.

**Prochain épisode : "La transcriptomique : l'ARN passé au crible".**

Vous souhaitez nous contacter pour nous rejoindre, nous proposer une mission ou simplement vous renseigner ? **Ecrivez-nous à [ai@startupinside.com](mailto:ai@startupinside.com).** Nous sommes également preneurs de vos propositions de sujets d'articles !