**Le sujet de stage**

Le sujet de stage qui vous a été attribué est "Santé : Détection Prédictive de la Monosomie avec IA Explicable et Biomarqueurs Virtuels".

📌 Principales idées : Développer un modèle d’IA explicable (XAI) pour que les médecins comprennent comment l’algorithme identifie la monosomie (transparence des décisions IA).

Créer une base de données mondiale collaborative en blockchain où les chercheurs peuvent partager des découvertes sur la monosomie, garantissant sécurité et anonymisation des données génétiques.

Utiliser un modèle GAN (Generative Adversarial Networks) pour générer des biomarqueurs virtuels et améliorer la précision du diagnostic, même avec peu de données.

Innovations :

- Explicabilité IA (XAI) pour rendre le diagnostic transparent et fiable.

- Base de données blockchain pour partager des découvertes et améliorer le modèle en continu.

- GANs pour enrichir les données et augmenter la diversité des cas étudiés.

📌 Technologies utilisées : TensorFlow, SHAP (Explainable AI), GANs, Ethereum (Smart Contracts), BioPython, Flask.

La **monosomie** et la **chromosome** sont deux concepts liés mais distincts en génétique :

1. **Chromosome** :
   * C'est une structure filamenteuse présente dans le noyau des cellules, composée d'ADN et de protéines.
   * Les chromosomes contiennent les gènes, qui portent l'information génétique.
   * Les humains ont **46 chromosomes** (23 paires).
2. **Monosomie** :
   * C'est une anomalie chromosomique dans laquelle **il manque un chromosome** dans une paire.
   * Normalement, les cellules humaines ont **23 paires de chromosomes** (soit 46 chromosomes), mais en cas de monosomie, une paire ne possède qu’un seul chromosome au lieu de deux.
   * Exemple : **Le syndrome de Turner (45,X)** est une monosomie qui concerne la paire de chromosomes sexuels. Les femmes atteintes de ce syndrome n’ont qu’un seul chromosome X au lieu de deux (XX).

**Différence :**

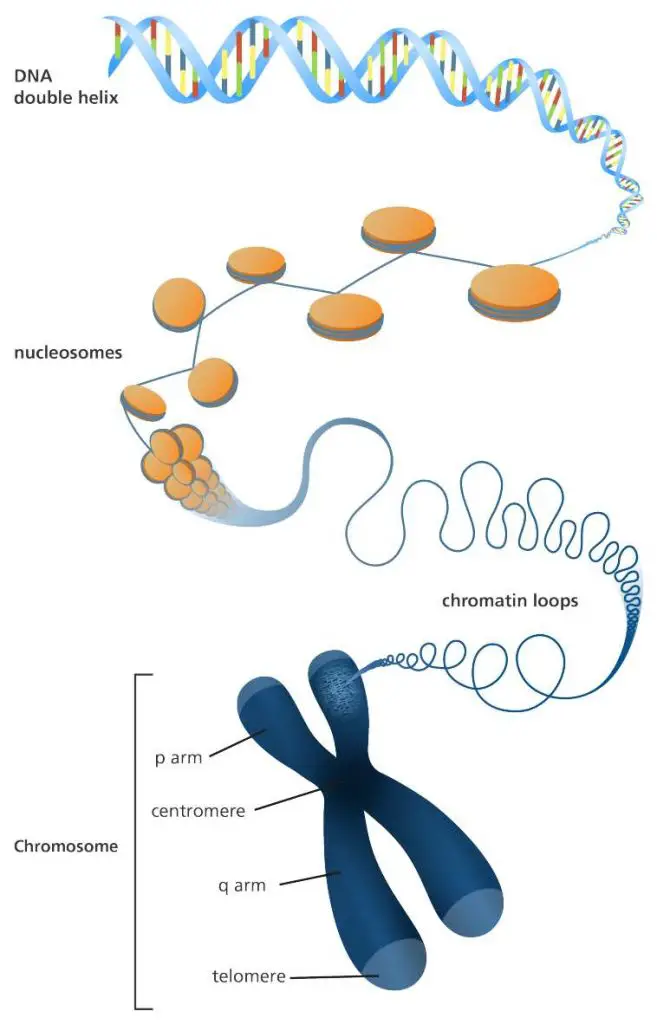
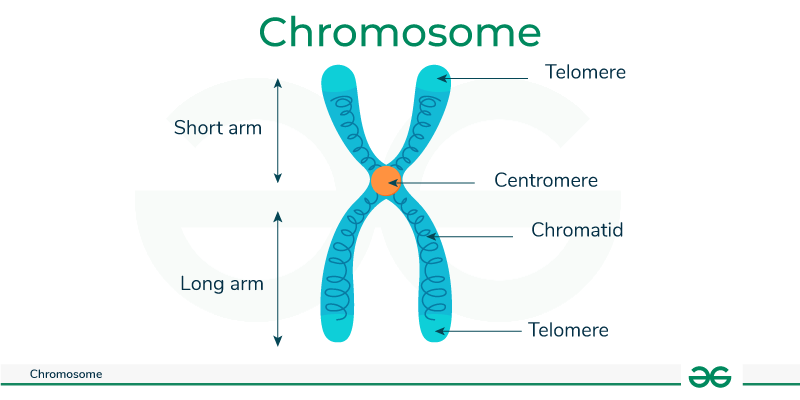
* **Chromosome** = une unité de structure contenant l'ADN.
* **1. C’est quoi un chromosome ? 🧬**
* Un **chromosome** est comme une boîte de rangement pour notre ADN. Imagine un long fil très fin (l’ADN) qui est enroulé sur lui-même plusieurs fois pour former un petit paquet : ce paquet, c’est un chromosome.
* Chaque cellule de ton corps (sauf certaines comme les globules rouges) contient **46 chromosomes**, organisés en **23 paires**. Tu en as reçu **23 de ta mère** et **23 de ton père**.
* Quand une cellule doit se diviser pour en créer une nouvelle, ces chromosomes se dupliquent et se séparent pour que chaque cellule fille ait le bon nombre d’informations génétique

**2. Comment est organisé un chromosome ? 📦**

Un chromosome, c’est **de l’ADN bien compacté**. Voici comment il est structuré :

* **L’ADN** 🧬 : C’est une longue molécule contenant toutes les instructions pour faire fonctionner notre corps.
* **Les nucléosomes** 🔵 : L’ADN ne peut pas rester en long fil, sinon il s’emmêlerait. Il s’enroule autour de petites protéines appelées **histones**, un peu comme un fil autour d’une bobine.
* **Les boucles de chromatine** 🔄 : Plusieurs nucléosomes se regroupent pour former des boucles.
* **Le chromosome** 📏 : Toutes ces structures sont encore plus compactées pour former un chromosome bien visible.

Un chromosome a une forme en **X** (sauf quand il est en train d’être copié). Il contient :

* **Un centromère** (le point au milieu).
* **Deux bras** :
  + **Le bras court (p arm)**
  + **Le bras long (q arm)**
* **Les télomères** (les extrémités) qui protègent le chromosome.

**3. Pourquoi avons-nous des chromosomes ? 🏗️**

Les chromosomes sont essentiels car ils stockent **les informations génétiques**. Ces informations sont organisées sous forme de **gènes**, qui sont des "recettes" pour fabriquer les protéines dont notre corps a besoin.

Chaque chromosome contient **des milliers de gènes**. Par exemple :

* Un gène peut déterminer la couleur de tes yeux.
* Un autre peut influencer ta taille.
* Certains gènes peuvent aussi être liés à des maladies.

**4. Comment l’ADN stocke-t-il les informations ? 📜**

L’ADN est comme **un livre de recettes** écrit avec **seulement 4 lettres** :

* **A** (Adénine)
* **T** (Thymine)
* **C** (Cytosine)
* **G** (Guanine)

Ces lettres s’associent toujours de la même façon :

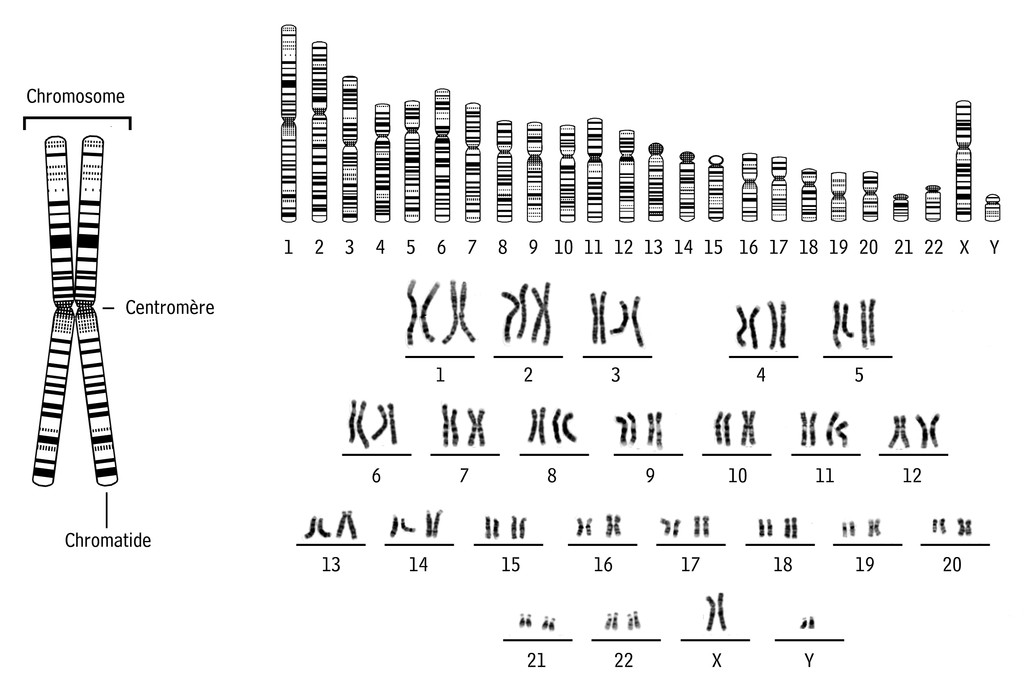
* **A va toujours avec T**
* **C va toujours avec G**

Imagine que ton ADN est une longue phrase écrite avec ces lettres. Chaque groupe de lettres donne une instruction à la cellule pour fabriquer une protéine spécifique.

**📋 Caryotype : Une Carte des Chromosomes**

La deuxième image représente un **caryotype**, c’est-à-dire une photographie de l’ensemble des chromosomes d’une cellule, classés en fonction de leur taille et de la position de leur centromère.

🔹 **Chaque cellule humaine possède 46 chromosomes**, regroupés en **23 paires** :

* **22 paires d’autosomes** (chromosomes non sexuels).
* **1 paire de chromosomes sexuels** (XX pour une femme, XY pour un homme).
* **Monosomie** = une **anomalie** dans laquelle un chromosome manque dans une paire.
* **📌 Qu'est-ce que la Monosomie ? 🧬❌**
* La **monosomie** est une anomalie génétique où **il manque un chromosome** dans une paire. Normalement, un humain possède **23 paires de chromosomes** (soit **46 chromosomes** au total), mais dans la monosomie, une paire n’a **qu’un seul chromosome au lieu de deux**.

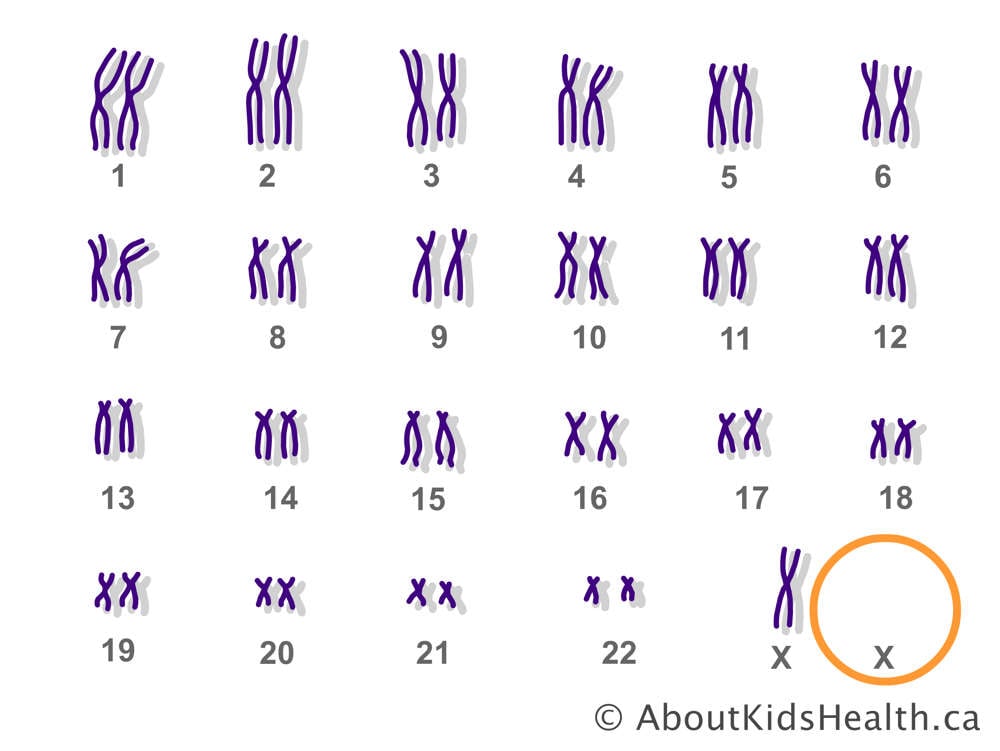
**🔬 Comment se produit une monosomie ?**

La monosomie survient lorsqu’une erreur se produit lors de la division cellulaire (méiose), ce qui entraîne **la perte d’un chromosome** dans la cellule fécondée. Cela peut être causé par :

* Une **non-disjonction** (lorsque les chromosomes ne se séparent pas correctement pendant la méiose).
* Une **délétion chromosomique** (une partie d’un chromosome est perdue).
* Un **transfert inégal de matériel génétique** lors de la fécondation.

**[Syndrome de Turner](https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.aboutkidshealth.ca%2Ffr%2Fsyndrome-de-turner&psig=AOvVaw0n63AL4axWc9F7-IakPKwl&ust=1739264551137000&source=images&cd=vfe&opi=89978449&ved=0CBcQjhxqFwoTCKDk1t3fuIsDFQAAAAAdAAAAABAw" \t "_blank)**

**Syndrome de Turner (45,X)**

* **Cause** : Perte totale ou partielle d’un chromosome X.
* **Sexe affecté** : Seulement les **femmes**.
* **Caractéristiques** :
  + Petite taille.
  + Absence de développement des caractères sexuels secondaires (seins, menstruations).
  + Infertilité.
  + Malformations cardiaques possibles.
  + Anomalies rénales.
  + Difficultés d’apprentissage dans certains domaines (ex. : orientation spatiale).
* **Diagnostic** : Caryotype (analyse des chromosomes).
* **Traitement** :
  + Hormone de croissance pour la taille.
  + Traitement hormonal substitutif pour le développement des caractères sexuels secondaires.

**Autres Maladies de Monosomie**

La **monosomie complète** est souvent létale avant la naissance, sauf pour la monosomie X (syndrome de Turner). Mais certaines **monosomies partielles** existent :

1. **Monosomie 4p (Syndrome de Wolf-Hirschhorn)**
   * Délétion partielle du bras court du chromosome 4.
   * Traits distinctifs du visage ("casque de guerrier grec").
   * Retard de développement sévère.
   * Convulsions.
   * Anomalies cardiaques.
2. **Monosomie 5p (Syndrome du Cri du Chat)**
   * Délétion partielle du bras court du chromosome 5.
   * Cri aigu du nourrisson (ressemblant à un miaulement de chat).
   * Microcéphalie (petite tête).
   * Retard mental et retard de croissance.
   * Difficultés motrices.
3. **Monosomie 18p**
   * Délétion du bras court du chromosome 18.
   * Petite taille, retard de développement.
   * Anomalies faciales et neurologiques.
4. **Monosomies partielles diverses**
   * Certaines personnes naissent avec des **délétions partielles** sur divers chromosomes, entraînant des troubles du développement et des malformations spécifiques.

**DATASET**

**Types de Données à Utiliser**

1. **Images Cytogénétiques** : Des images de cellules en métaphase peuvent être utilisées pour entraîner des modèles de détection de la monosomie.
2. **Données Génomiques** : Les séquences d'ADN et les profils d'expression génique peuvent fournir des informations cruciales pour identifier des biomarqueurs associés à la monosomie.
3. **Données Cliniques** : Les dossiers médicaux électroniques contenant des informations sur les diagnostics, les symptômes et les antécédents des patients peuvent enrichir le modèle prédictif.

**Chromosome karyotype Images :**

https://www.kaggle.com/datasets/aliabedimadiseh/chromosome-image-dataset-karyotype/code

<https://www.kaggle.com/datasets/arifmpthesis/bioimlab-chromosome-data-set-for-classification>

Articles:

<https://www.nature.com/articles/s41597-023-02003-7>

<https://www.nature.com/articles/s41467-021-25288-x>

<https://www.nature.com/articles/2401654>

<https://ieeexplore.ieee.org/abstract/document/9178721>