

Introduction / État de l'art

Le **caryotypage** consiste à préparer des cellules bloquées en métaphase, à colorer les chromosomes selon des techniques de bandes (G, R, Q), puis à photographier et disposer les 46 chromosomes humains (22 paires autosomiques + XX/XY) pour inspection visuelle [1].

Il permet de révéler des anomalies de **nombre** (ex. trisomie 21, Turner, Klinefelter) ou de **structure** (délétions, duplications, inversions, translocations), et reste un outil incontournable en diagnostic prénatal, postnatal et en hématologie.

Cependant, l'analyse est **longue**, **dépend fortement du biologiste** et de la qualité des préparations, ce qui entraîne une certaine **variabilité** entre opérateurs [2]. Pour gagner en rapidité, robustesse et **interprétabilité**, nous proposons de représenter chaque chromosome par un **profil mathématique de densité** (*statistique fonctionnelle*) afin de permettre des comparaisons directes entre courbes.. En parallèle, des approches d'IA existent (segmentation, classification) mais restent hétérogènes et peu standardisées [3].

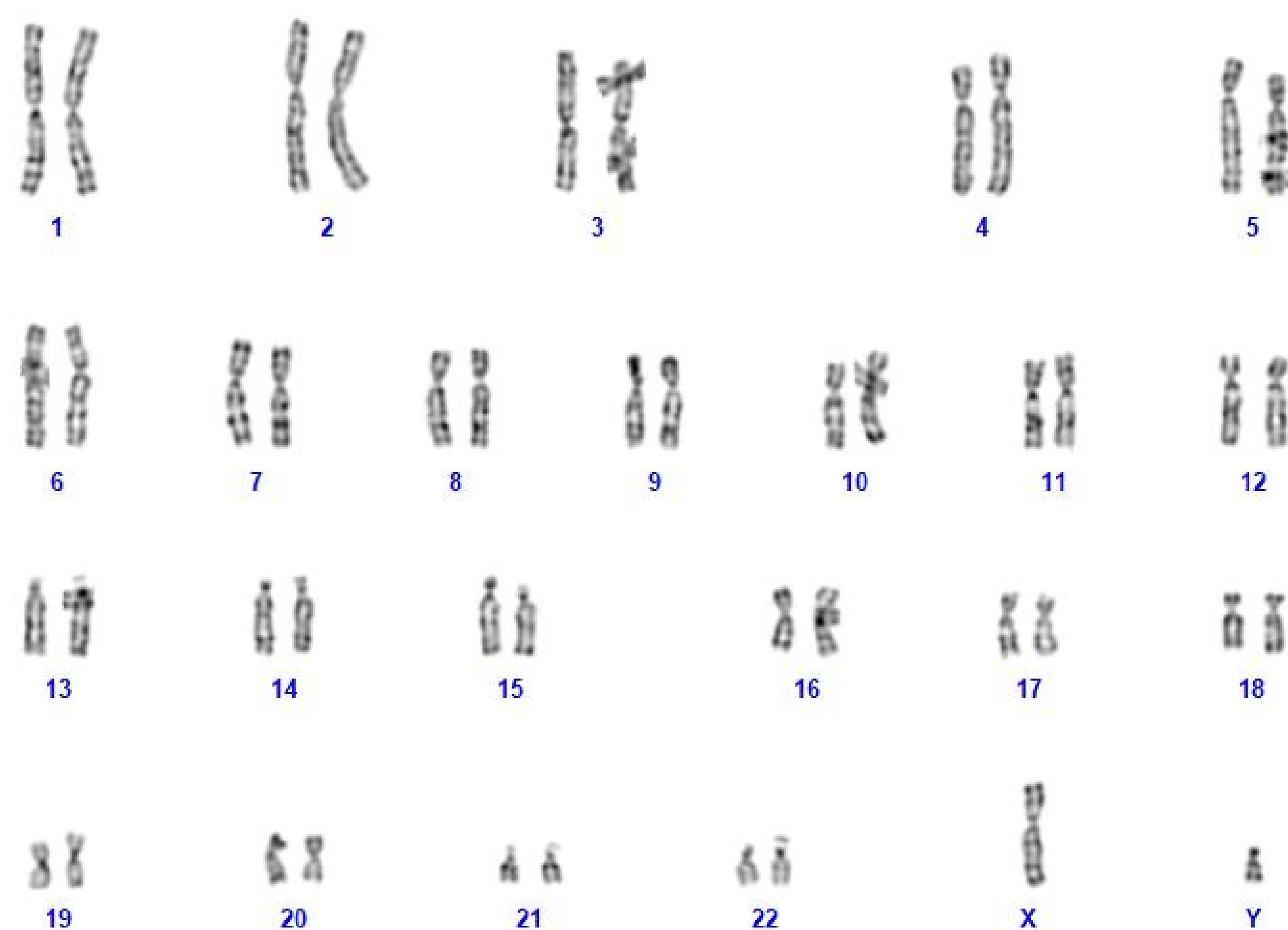


Figure 1. Caryotype humain sans anomalie notable (46,XY).

Problématique

L'analyse cytogénétique manuelle étant **chronophage** et **dépendante de l'expertise du biologiste**, comment développer une chaîne *image* → *profil fonctionnel* → *comparaison* qui permette :

- d'extraire automatiquement des **chromosomes droits et simples**,
- de repérer et traiter les **chevauchements éventuels** entre chromosomes,
- de construire des **profils de densité normalisés** sur [0–100],
- de comparer ces profils à l'aide de **distances robustes et interprétables**,
- et d'aboutir à un **score de normalité** par chromosome, établi à partir de plusieurs caryotypes d'un même patient.

Données

Sources. Fichiers binaires **Car.dat** (Genikon/Excilone) contenant les images de caryotypes.

Pré-traitements et sorties.

- **Extraction par chromosome** (1–22, X, Y) et sauvegarde **PNG**.
- **Base tabulaire** : identifiant image, classe chromosomique, vecteur $(X_1, \dots, X_{100})$ (profil de densité interpolé sur [0, 1]).

Avancées actuelles : extraction des chromosomes droits et profil de densité

Pipeline implémenté.

1. **Extraction** à partir des fichiers binaires **Car.dat**, en récupérant directement la **version droite** de chaque chromosome, classés 1–22, X, Y.
2. **Profil de densité** : obtenu en extrayant l'**axe médian** de chaque chromosome, puis en mesurant l'**intensité le long de cet axe** pour générer un signal de densité. Ce calcul est réalisé grâce au package **Banding-Pattern-Extraction** [4], avant **normalisation** et **interpolation uniforme** (100 points sur [0, 1]).
3. **Base structurée** prête pour les statistiques fonctionnelles (comparaison, clustering, détection d'anomalies).

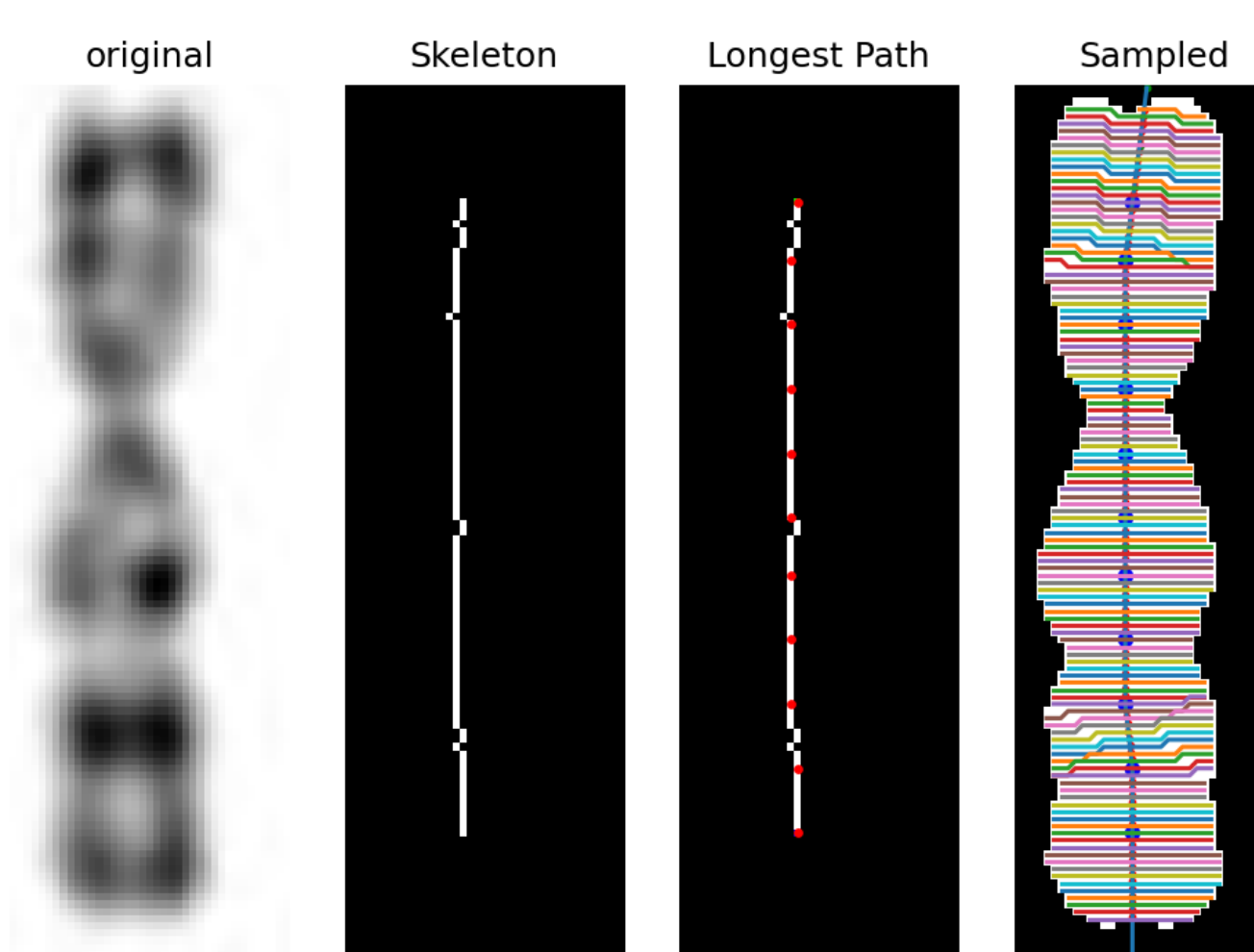


Figure 2. Extraction de l'axe central d'un chromosome

Objectifs

1. **Comparaison et alignement robustes** Définir des distances adaptées entre profils : norme L^2 (différences globales), corrélation $(1 - \rho)$, DTW (tolérance aux décalages locaux). Introduire une **pénalisation des compressions/dilatations** afin de limiter les déformations excessives, et **masquer automatiquement** les zones peu informatives.

2. **Amélioration du signal** Corriger l'**étalement du pigment** par des méthodes de déconvolution (Wiener, Richardson–Lucy) sous contraintes de lissage/non-négativité. Harmoniser les profils entre préparations en gérant l'**inversion éventuelle des contrastes** selon la teinture et en normalisant l'intensité pour une meilleure comparabilité.

3. **Statistiques fonctionnelles et morphométrie** Exploiter les outils de statistique fonctionnelle (profondeur de courbes, FPCA robuste, régions de densité élevée) pour repérer des profils atypiques. Compléter cette approche par des **indicateurs morphométriques** simples : ratio longueur/largeur, largeur en pixels le long de l'axe, puis intégrer ces informations avec les profils de densité.

4. **Détection des chevauchements** Mettre en place une étape de **repérage automatique des chromosomes chevauchants ou en contact**, afin de les **compter comme cas anormaux** et les intégrer directement dans la classification.

5. **Vers un score de normalité** Construire un **score global de normalité par chromosome**, combinant les indicateurs fonctionnels et morphométriques. Ce score sera calculé sur **plusieurs caryotypes d'un même patient** pour estimer la cohérence intra-patient et détecter des anomalies structurales ou des mosaïques de manière plus sensible.

Références

- [1] McGowan-Jordan J., Simons A., Schmid M. (2016). ISCN 2016 : International System for Human Cytogenomic Nomenclature. Karger.
- [2] Van Campen J. (2022). Karyotype NHS England.
- [3] Menaka D. and Ganesh Vaidyanathan S. (2022) Chromenet : a CNN architecture with comparison of optimizers for classification of human chromosome images. *Multidimensional Systems and Signal Processing*, 33:747-768.
- [4] Lukasuz (2024). *Banding-Pattern-Extraction* [code source]. GitHub : github.com/lukasuz/Banding-Pattern-Extraction.