

Décrypter le génome: le casse-tête du bio-informaticien !

Sandrine Moreira

sandrine.moreira@umontreal.ca

Doctorante en Bio-informatique

Centre Robert Cedergren

Université de Montréal

Qu'est-ce que la bio-informatique ?

La **bio-informatique** utilise
des **méthodes informatiques** pour
comprendre les **phénomènes** **Génomes** **vivant.**

De quoi à besoin le bio-informaticien ?



```
TAATTAAATCTAACAAATTTTCGTTACAGATTTCTTTACACACAAGTTATAC  
ACAAGTTAACTGGCTGTGGACAACCGTTTTTCACATCTGGACAGTTTTGT  
GGATAGATTTTGTAAGTCCTTGCTATCAGAGTGATTTTCTGATATTATAA  
TTCCGTGTCGAATAGAAATATAGCTGGGGAAAACATAAGTTATCCACAATA  
CATTTTTACTTTGTGGATAATTTTTTAACAGTGTGGATAACCTTATCC  
ATGGCTTTTCTATCTGTGGATAACTTTATAGCATCCATTTACATTACAT  
AAAAAAGGGGGGTACTAGTGCAATCAATTGAAGACATCTGGCAGGAAAC  
ACTGCAATCGTAAAAAAAATATGAGTAAACCTAGTTACGACACATGGA  
TGAAATCAACAACCGCTCATTCACCTGAAGGTAACACGTTTATTATTCA  
GCGCCCAATAATTTGTTCCGATTGGTTAGAGAAGAGTTACACGCAATT  
TATCGCTAACATTTTGCAAGAAATAACTGGTCGCTTATTTGATGTCCGCT  
TTATTGATGGCGAGCAGGAAGAAAACCTTTGAATACACTGTGATTAACCA  
AATCCAGCGTTAGATGAAGATGGCGTTGAAATGGAAAAACATATGCTTAA  
TCCGCGTTATGTTTTGATACTTTTGTCTATTGGTTCCAGGGAACAGATTTG  
CCCACGCAGCATCACTTGCAGTAGCCGAAGCACCAGCGAAAGCATATAAT  
CCACTCTTCAATTTATGGAGGAGTTGCCTCGGTAAAAACATTTAATGCA  
CGCAGTTGGCCACTATGTTCAACAACATAAAAGATAATGCGAAAGTAATGT  
ACCTTTCCAGCGAAAAATTCACCAATGAGTTTATTAGCTCTATTTCGTGAT
```

ADN



Ordinateur

Que fait le bio-informaticien de ses journées ?

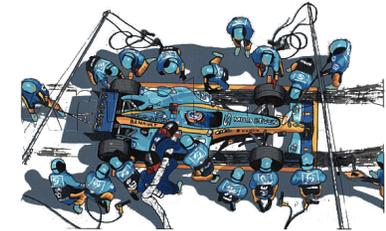
■ Bio-informatique appliquée

Utilisation de logiciels pour l'acquisition, le traitement et l'analyse de données biologiques



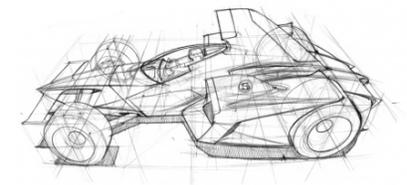
■ Genie Bio-informatique

Développement de logiciels basés sur des algorithmes établis

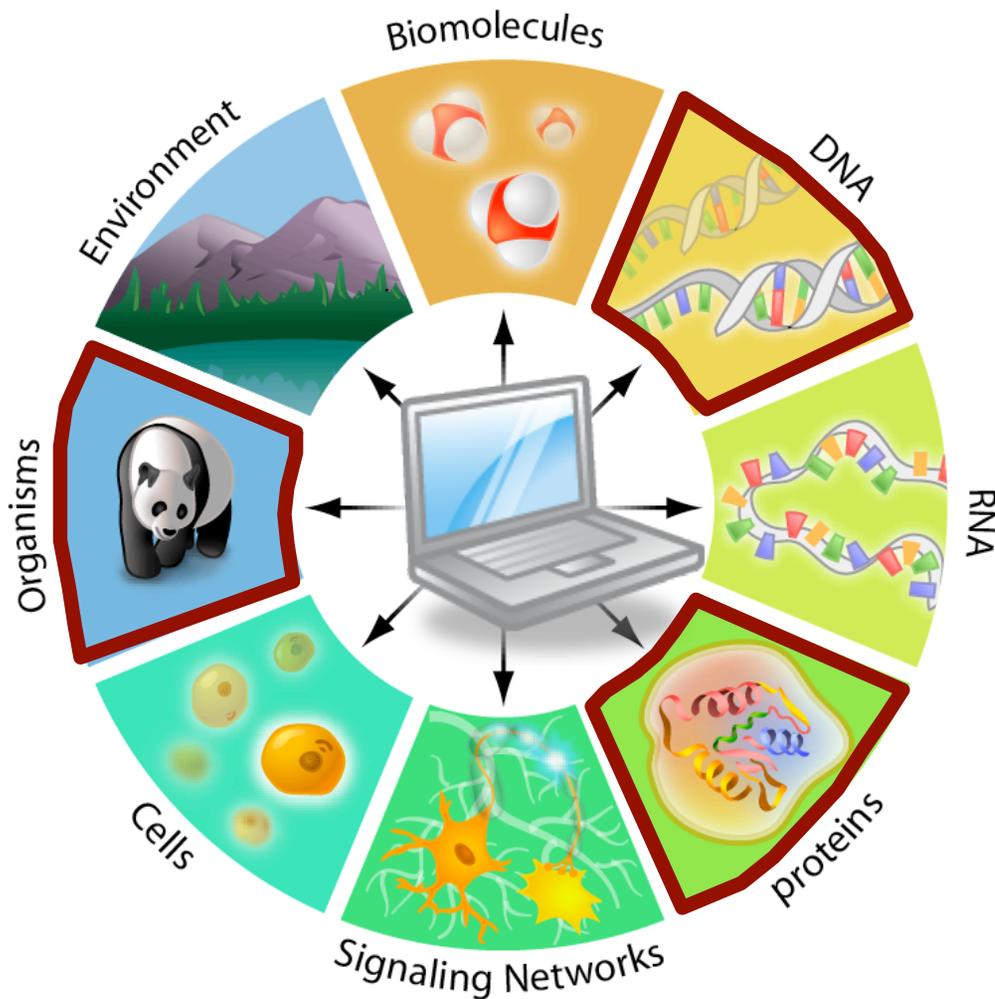


■ Bio-informatique théorique

Développement d'algorithmes pour résoudre des problèmes biologiques



Sur quels sujets travaille le bio-informaticien ?

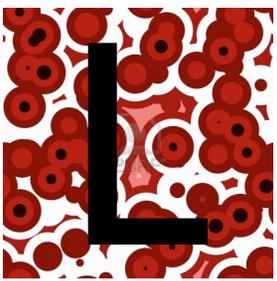


- Analyse d'une mutation sur l'ADN
- Structure des protéines

The background of the slide is a dark red color, populated with numerous red blood cells. Some cells are normal, appearing as smooth, biconcave discs. Others are sickle-shaped, which are elongated and curved, characteristic of sickle cell anemia. The cells are scattered across the frame, with some in sharp focus and others blurred in the background.

Des globules rouges qui ne tournent pas rond

Cas de l'anémie falciforme



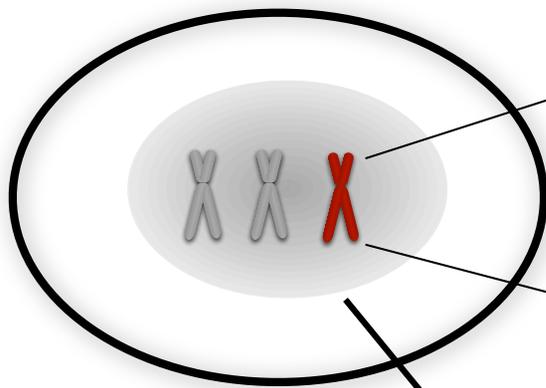
Les maladies génétiques

Cellule

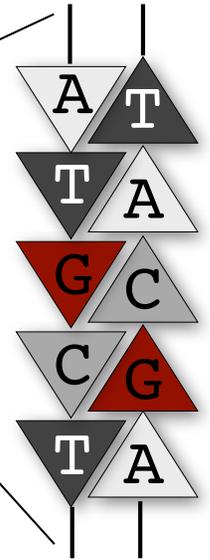
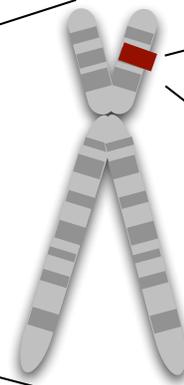
Chromosome

Gène

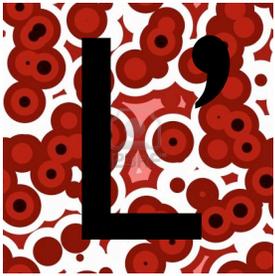
Séquence



Noyau



Les **maladies génétiques** sont liées à des modifications d'un **gène** qui l'empêchent de fonctionner normalement.



L'anémie falciforme

Globule rouge

Cellule du sang



Normal

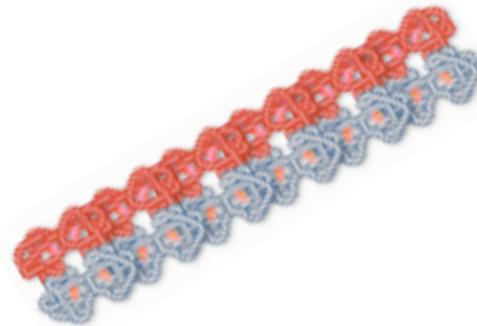


Malade



Hémoglobine

Molécule qui fixe l'oxygène



Gène de l'hémoglobine

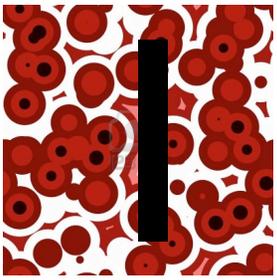


Mutation

L'**anémie falciforme** est dûe à une **mutation** du gène fabriquant l'**hémoglobine** qui cause sa polymérisation et donne une forme en **faucille** aux globules rouges.

The background of the slide is a dark red color, populated with numerous red blood cells. These cells are depicted in various orientations and sizes, some appearing as smooth, biconcave discs while others are slightly elongated or blurred, creating a sense of depth and movement. The lighting on the cells gives them a three-dimensional appearance with soft highlights and shadows.

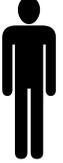
**Notre objectif: identifier la
mutation de l'anémie falciforme**



Identification d'une mutation

Gène

Séquence d'ADN


Normal

HbB

I

ATGC**G**TATGCTAGTCGTCAT

...


Malade

HbS

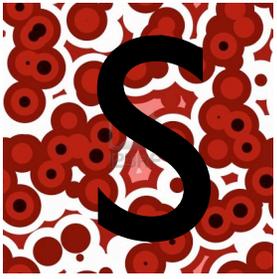
I

ATGC**T**TATGCTAGTCGTCAT

...

Mutation G > T

En **comparant** les séquences d'individus en **santé** et **malades**, les **différences** correspondent aux **mutations**.

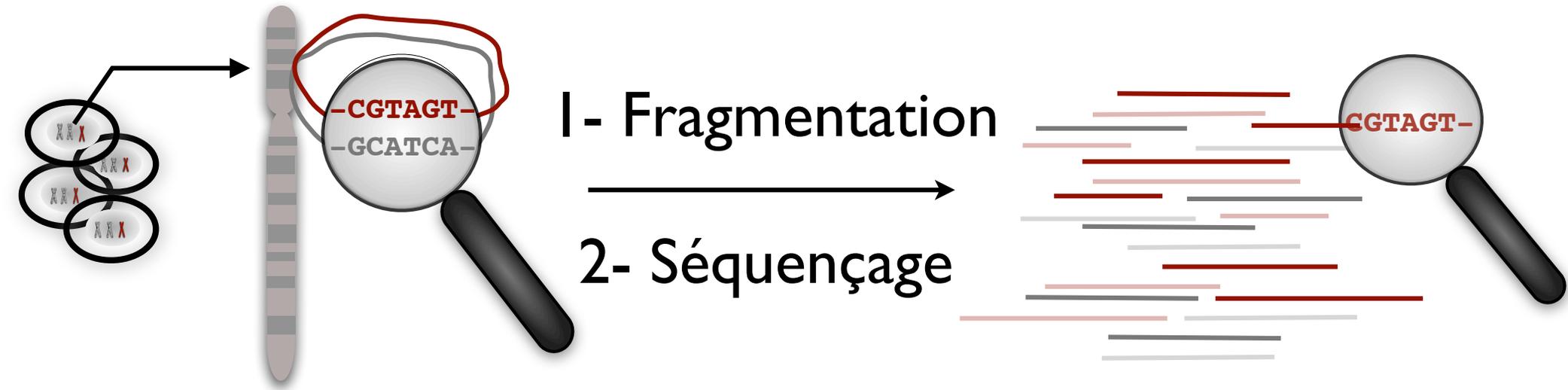


Séquençage de l'ADN

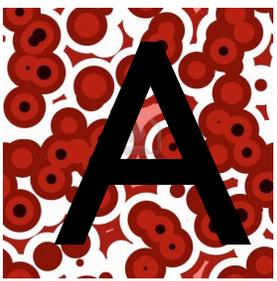
Cellules

ADN

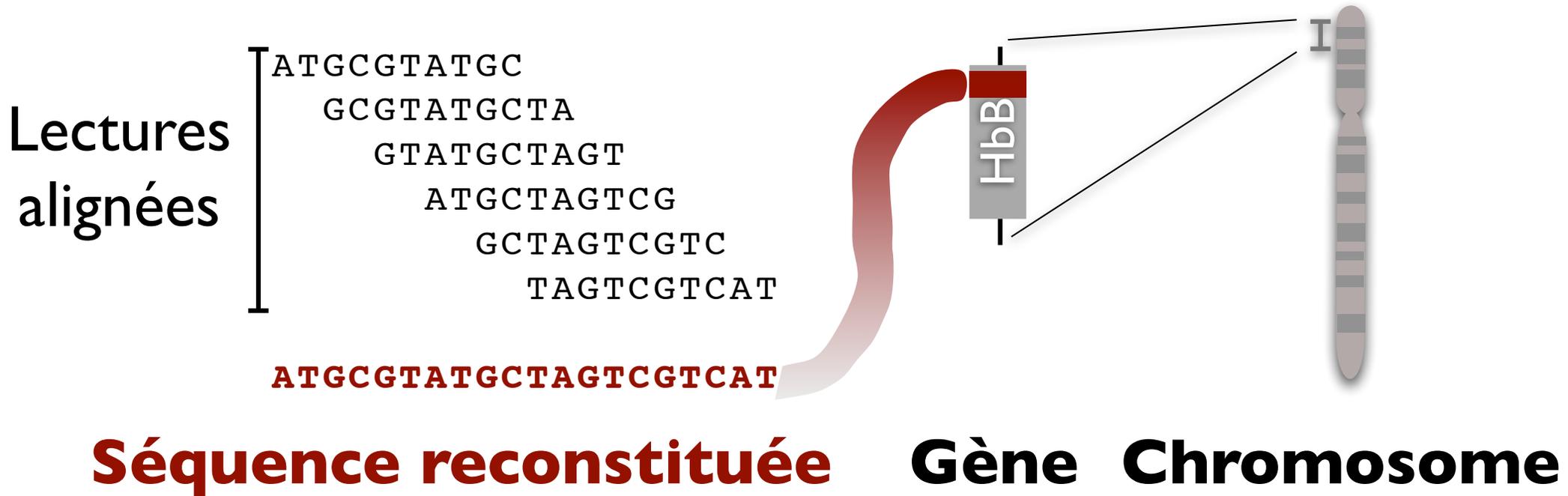
Lectures



Le **séquençage** consiste à déterminer la suite des lettres **A, T, G, C** des chromosomes. Les deux brins des **chromosomes** sont **fragmentés** en courtes séquences et lus par un **séquenceur**



Assemblage des lectures



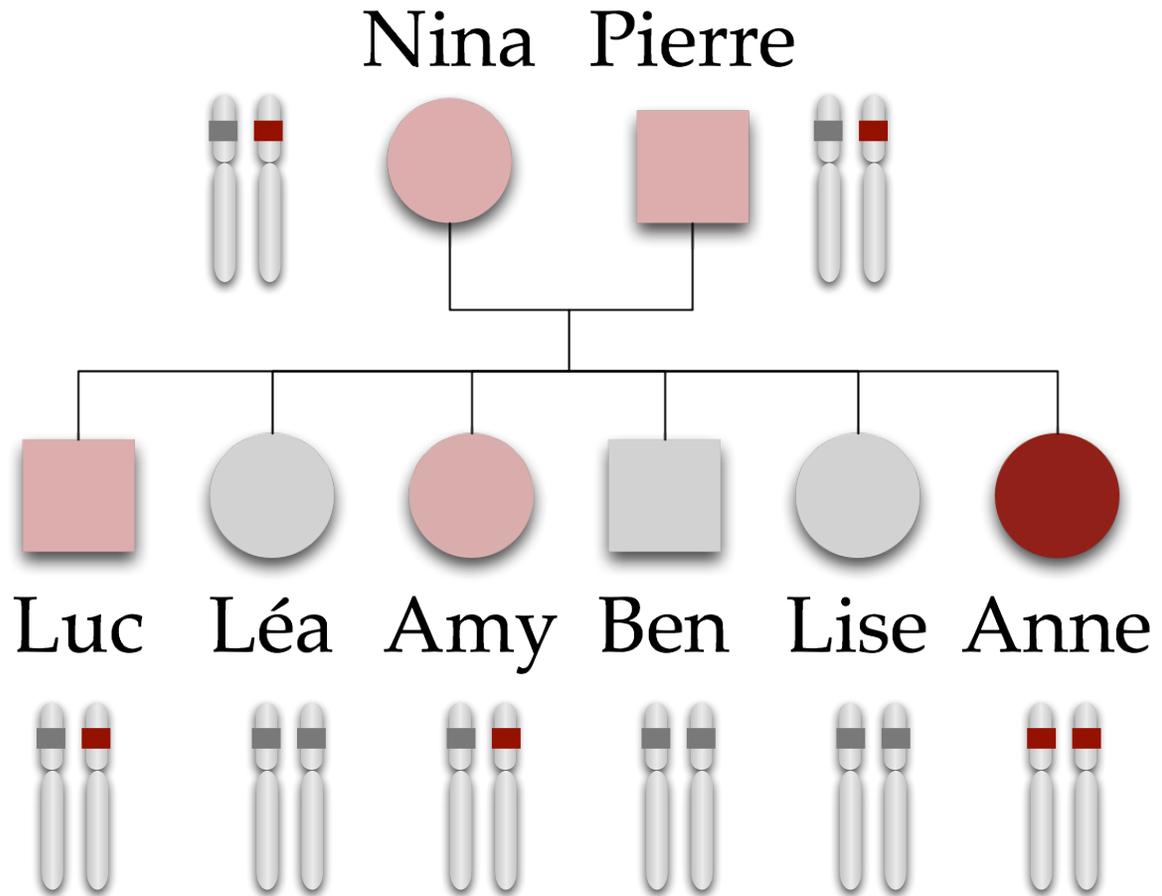
L'**assemblage** consiste à trouver les **zones chevauchantes** entres les **lectures** sorties du séquenceur pour les mettre bout-à-bout et reconstituer la **séquence** des gènes, et des chromosomes.

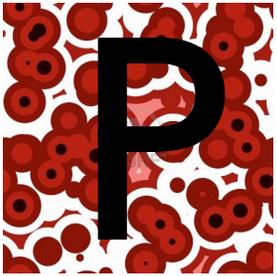


Votre mission: identifier la mutation de l'hémoglobine

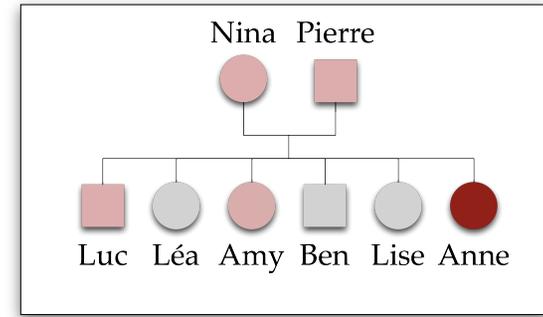
Tous les membres d'une famille ont donné leur ADN.

- Individus en santé
- Malades forme légère
- Malades forme grave

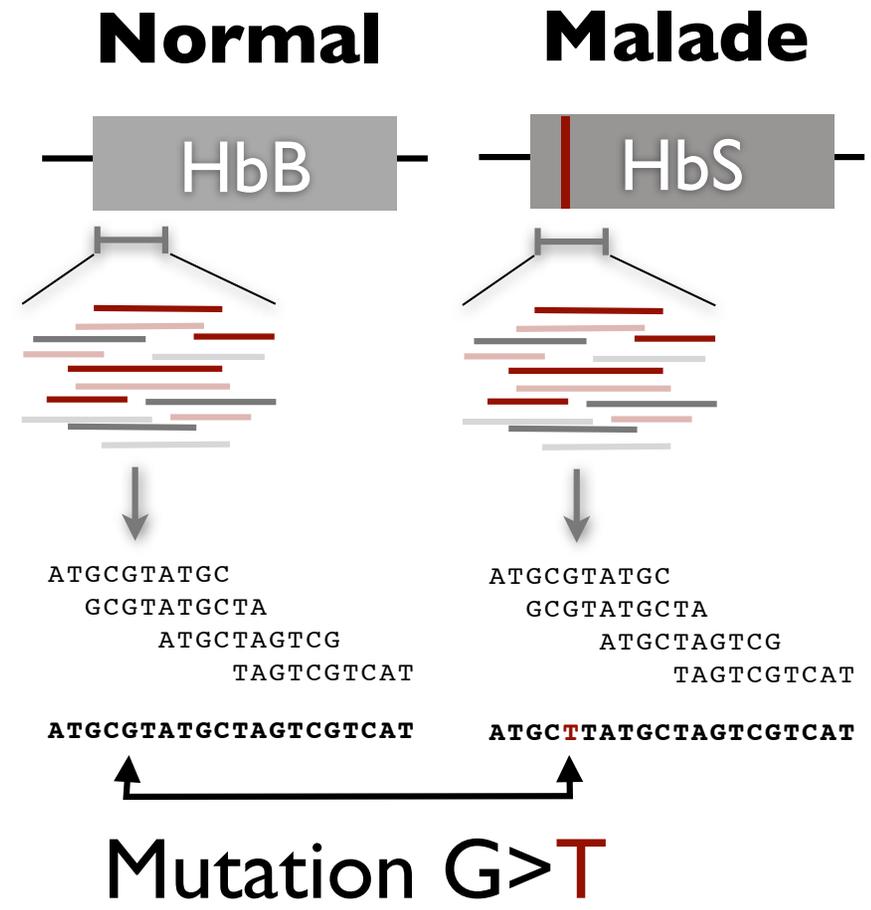


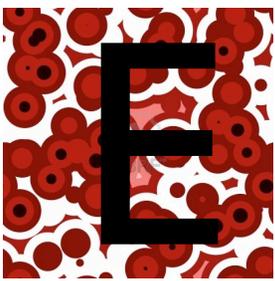


protocole



1. Nous avons séquencé le début des gènes de l'hémoglobine
2. Des enveloppes contiennent les lectures pour chaque membre de la famille
3. A vous de les **assembler** (à la main !) pour reconstituer la séquence
4. En **comparant** les séquences des membres de la famille, nous identifierons la **mutation**





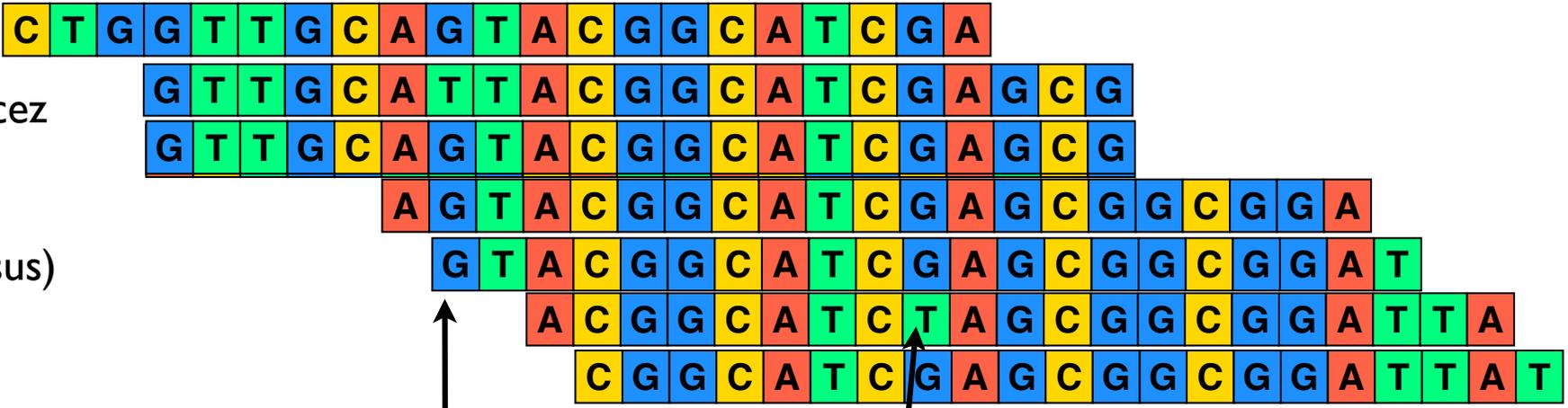
Exemple

1 10 20 30

G A C T G G T T G C A ^{T/G} T A C G G C A T C G A G C G G C G G A T T A T

X G A C T G G T T G C A T T A C G G C A T C

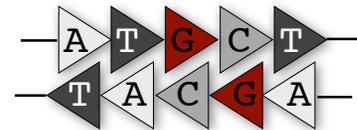
↑
Commencez par **ATG**
(avec une croix dessus)

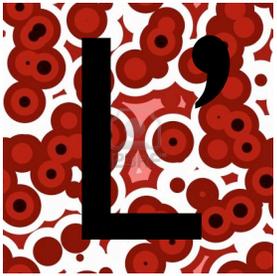


Polymorphisme ou SNP
(Single Nucleotide Polymorphism)
Mutation T/G

Erreur de séquence (1%)
! Léa difficile (3%)

• Un chromosome a **2 brins complémentaires**, écrits de chaque côté des bandes

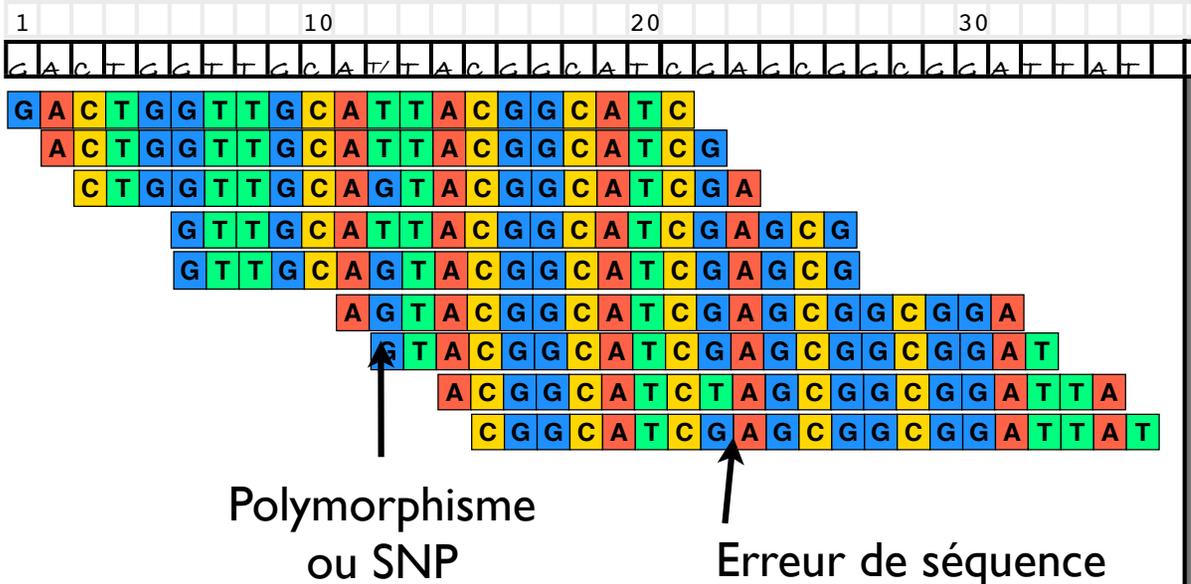




homme contre la machine !

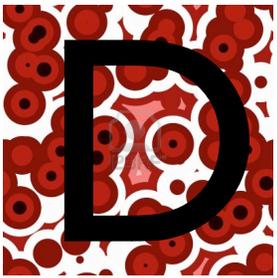
Assemblage «à la main»

- 20 séquences
- 21 bases
- Outil : mains + cerveau !

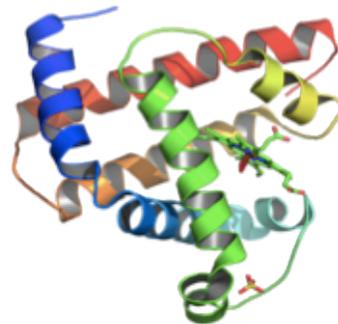
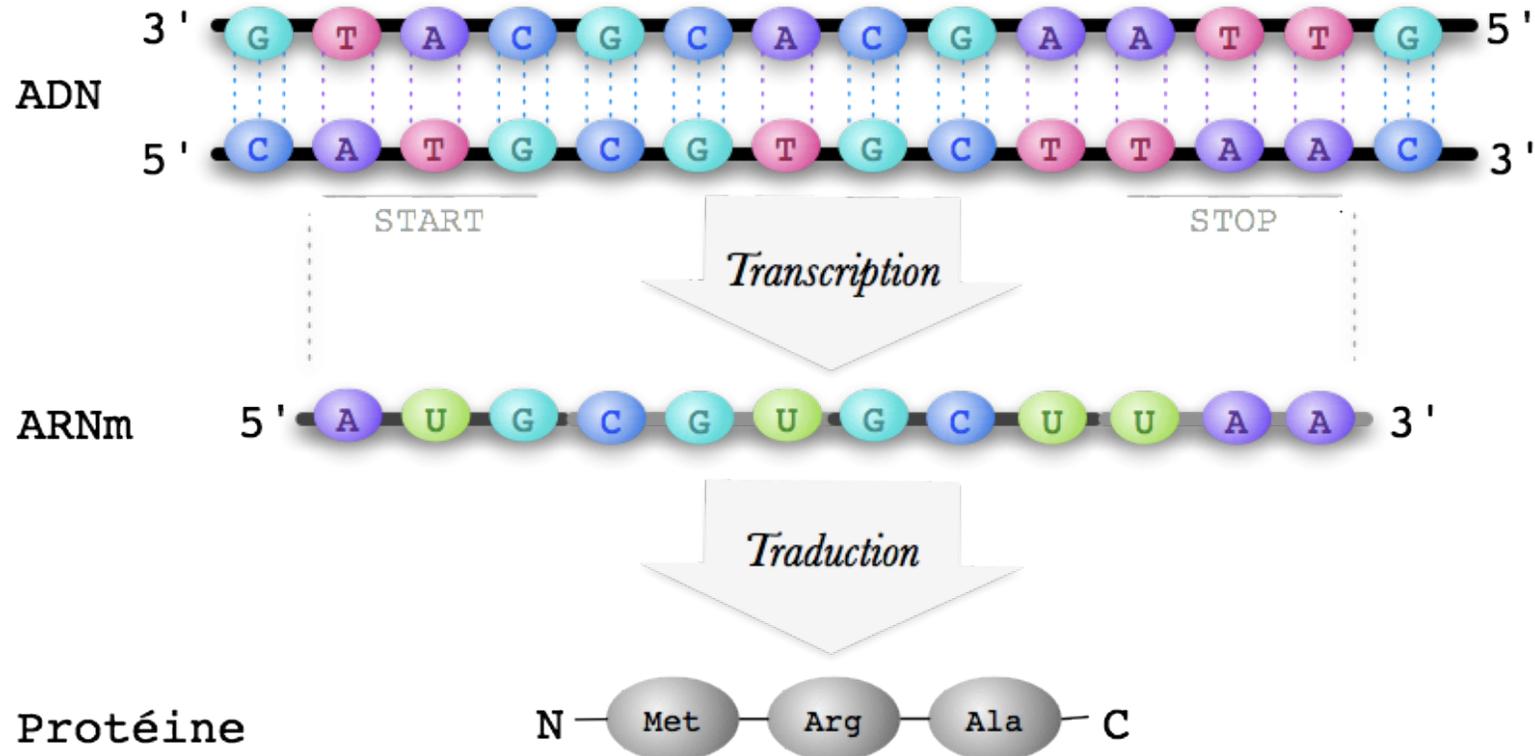


Assemblage par un outil bio-informatique

- 30000 séquences
- ~ 400 bases
- Outil : Mira
- Performances :
12 minutes
- Taille finale :
53000 bases



Du gène à la protéine



Les difficultés du métier

- Double compétence : biologie ET informatique
 - Bon niveau en biologie avec une large culture biologique (biochimie, génétique, physiologie, microbiologie, écologie, ...)
=> mémoire
 - Bon niveau en informatique (algorithmique, programmation, statistique, système, hardware, ...)
=> logique
- Double habileté : travail seul ET en équipe pluridisciplinaire
 - Capacité de concentration, autonomie
Travail toute la journée sur l'ordinateur
 - Bonnes capacités de communication
Interaction fréquentes avec biologistes et informaticiens

Merci !
Quel bon moment
nous venons de
passer !

