



UNIVERSITÉ DE LYON

Faculté de Médecine Lyon Est

# GENOME

Pr Jonathan LOPEZ



1

## Posez vos questions pendant le cours



- 1 Allez sur [wooclap.com](https://wooclap.com)
- 2 Entrez le code d'événement dans le bandeau supérieur

Code d'événement  
**PJINGO**



2

## 0 Plan

- Organisation du génome humain
- Gènes
- ADN intergénique

3

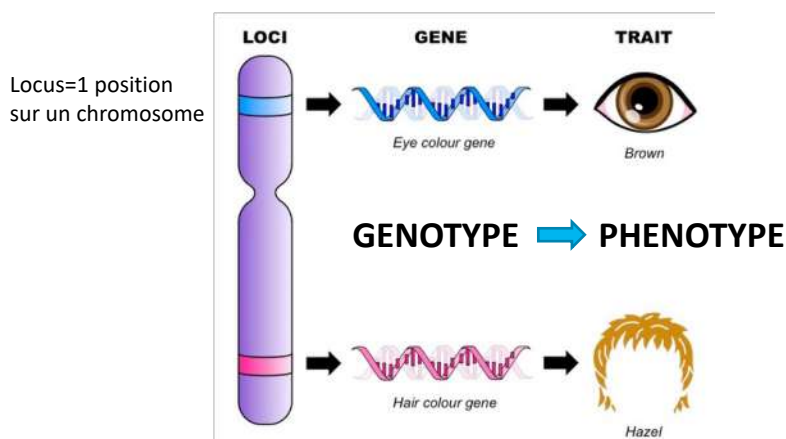
## 0 Plan

- Organisation du génome humain
  - Génotype vs phénotype
  - Polymorphisme génétique
  - Génome de référence
  - Composition du génome
- Gènes
- ADN intergénique

4

## 1 Génome : Génotype vs phénotype

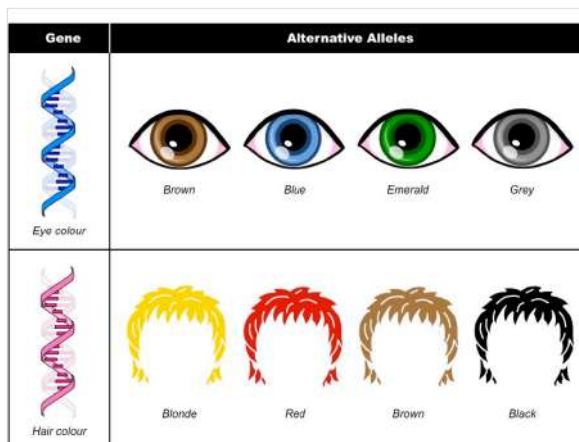
- Génotype = information génétique responsable de “traits phénotypiques”



5

## 1 Génome : Génotype vs phénotype

- La plupart des gènes existent dans différentes versions ou “allèles”



6

**1** Génome : **Génotype vs phénotype** Rappel

■ **Génome diploïde** (2n chromosomes)  
 ⇒ **2 copies de chaque gène (1 de chaque parent)**

**HOMOZYGOTE :**  
 Le même allèle sur les 2 chromosomes

**HETEROZYGOTE :**  
 1 allèle différent sur chaque chrom.

7 Faculté de Médecine Lyon-Sud

7

**1** Génome : **Génotype vs phénotype**

**1 allèle** (variant) peut être responsable de **différents phénotypes**

**“Expressivité”** = degré d’expression du phénotype (force du trait génétique)

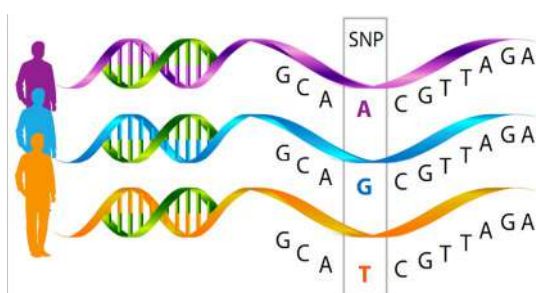
8 Faculté de Médecine Lyon-Sud

8

## 1 Génome : Polymorphisme génétique

Important

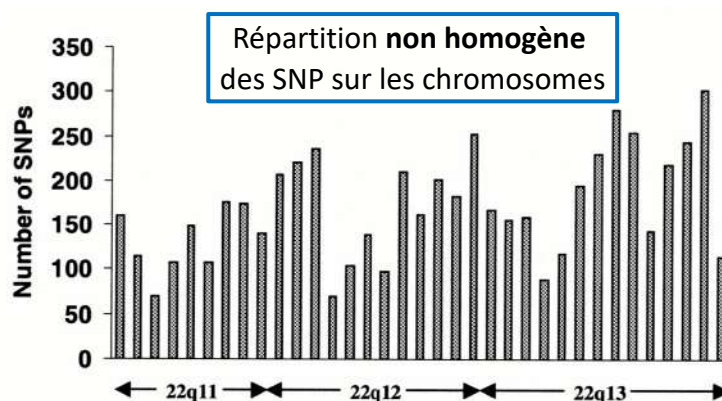
- **Single Nucleotide Variant** : variant génétique portant sur 1 nucléotide
- **Single Nucleotide Polymorphism (SNP "snip")** : SNV >1% dans la population saine
- **1 SNP tous les 100nts** dans le génome



9

## 1 Génome : Polymorphisme génétique

- **Single Nucleotide Polymorphism (SNP "snip")** :

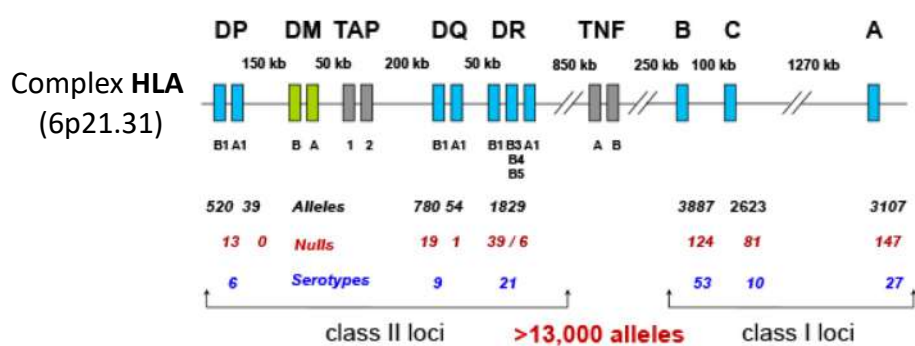


10

## 1 Génomique : Polymorphisme génétique

### ■ Single Nucleotide Polymorphism (SNP "snip") :

Certains loci sont **hypervariables**

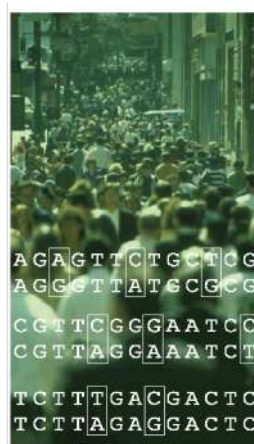


11

## 1 Génomique : Polymorphisme génétique

### ■ Single Nucleotide Polymorphism (SNP "snip") :

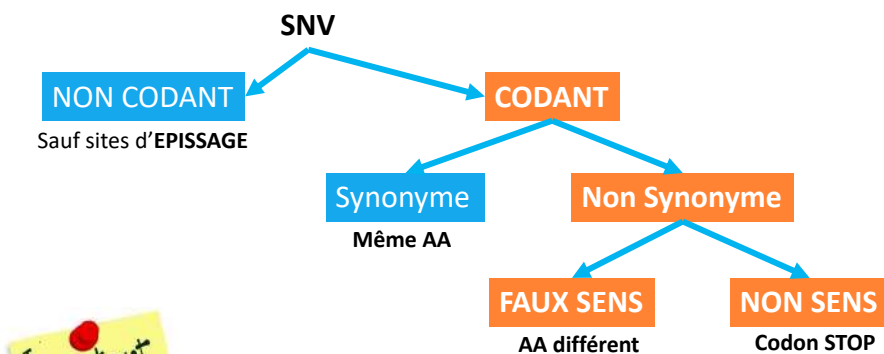
En comparant 2 individus  
 ⇔ 1 SNP tous les  
 1000/2000 bases



12

## ① Génome : Polymorphisme génétique

### ■ Impact fonctionnel d'un SNV



13

## ① Génome : Polymorphisme génétique

### ■ Impact fonctionnel d'un SNV

|               |                            |                                     |                                |                   |
|---------------|----------------------------|-------------------------------------|--------------------------------|-------------------|
| Variant bénin | Variant probablement bénin | Variant de signification incertaine | Variant probablement pathogène | Variant pathogène |
| Classe 1      | Classe 2                   | Classe 3                            | Classe 4                       | Classe 5          |

La plupart des variants ne sont pas pathogènes

(voir cours d'Alexandre Janin en MMOK)

14

## 1 Génome : Polymorphisme génétique Important

**Chromosome**

**Reference**  
 A B C D  
 ATCGGGTCATGTCA

**Sequence variants**

- Single nucleotide variant  
 A B C D  
 ATCGGGTCATATCA
- Small insertion <50pb  
 A B C D  
 ATCGGGTCATGACGTCA
- Small deletion  
 A B C D  
 ATCGGGTCAT

15 Faculté de Médecine

15

## 1 Génome : Polymorphisme génétique Important

**Normal**  
 mRNA AUGAAAGUUUUGGCAGUAAGUCCG  
 Protein Met Lys Phe Gly Ile Val Pro

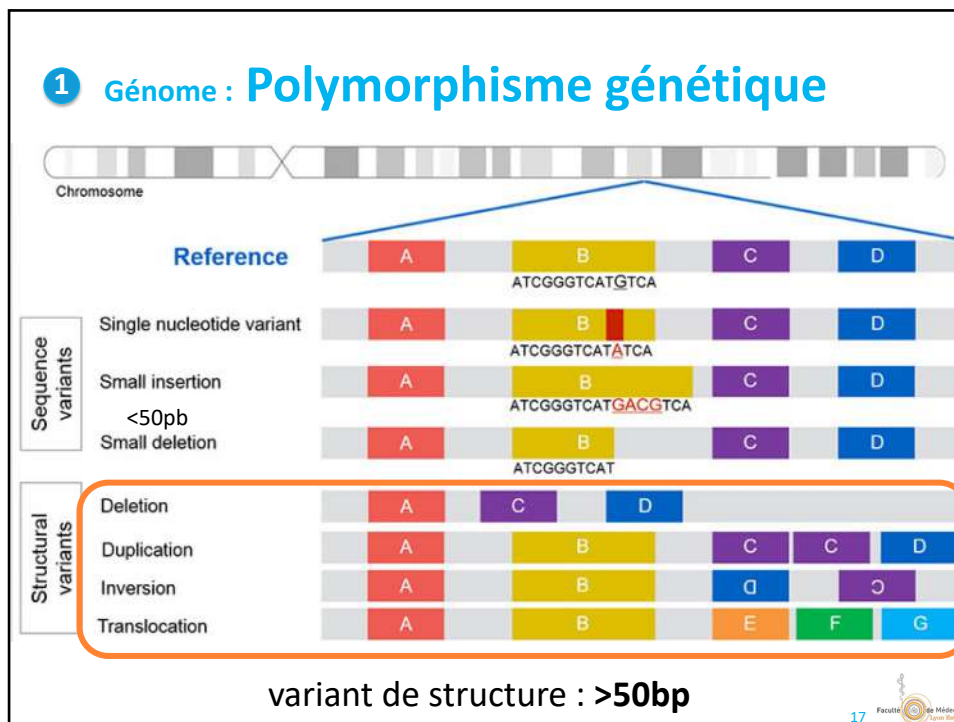
**Frameshift**  
 mRNA AUGAAAGUUUUGGCAGUAAGUCCG  
 Protein Met Lys Leu Ala STOP

**Insertions/délétions (≠ multiple de 3)**  
 ⇒ décalage du cadre de lecture ("frameshift")  
 ⇒ codon STOP prématuré ⇒ protéine tronquée

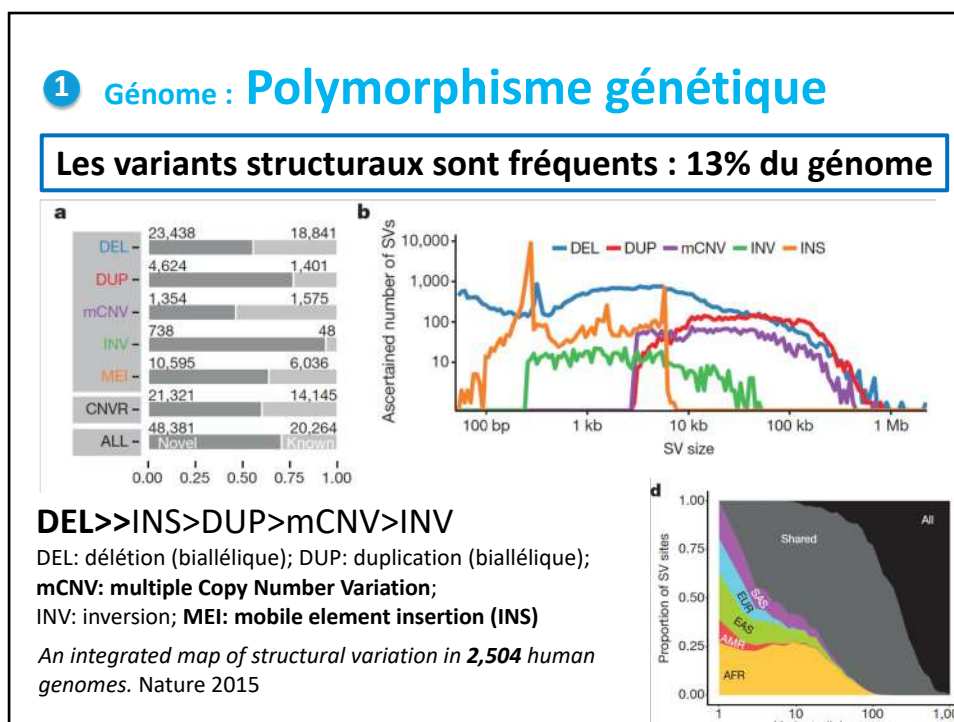
16 Faculté de Médecine

16





17



18

# 1 Génome : Polymorphisme génétique

## Copy Number Variation

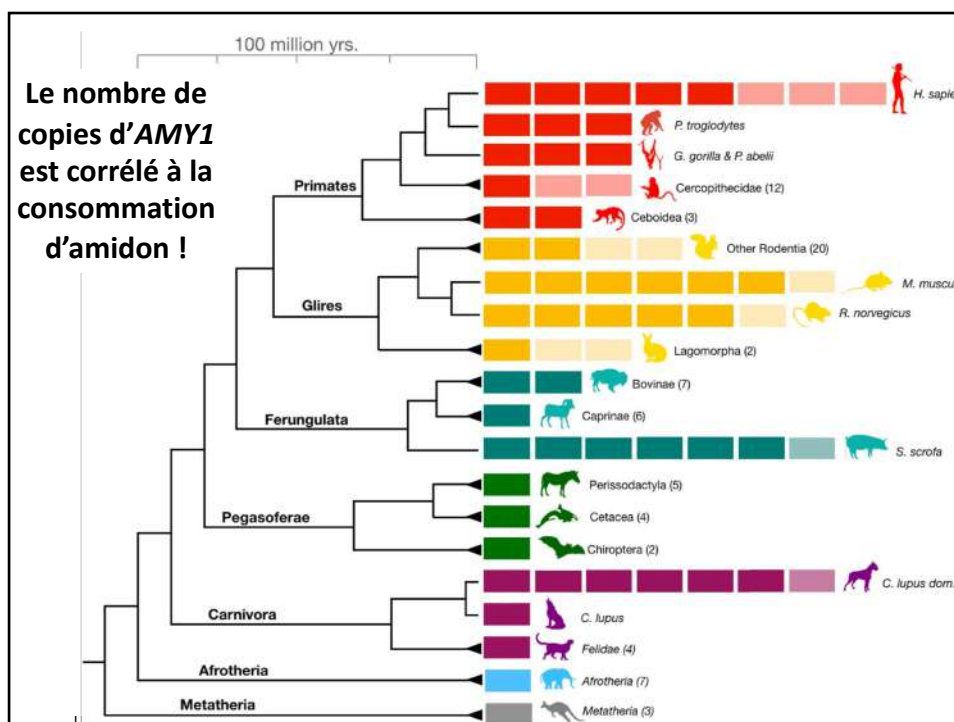
**Amylopectin**  
Polymers of  $\alpha$ -(1-4)-D-glycopyranosyl units with approximately 4%  $\alpha$ -(1-6) branching.

**Amylose**  
Polymer of  $\alpha$ -(1-4)-D-glycopyranosyl units

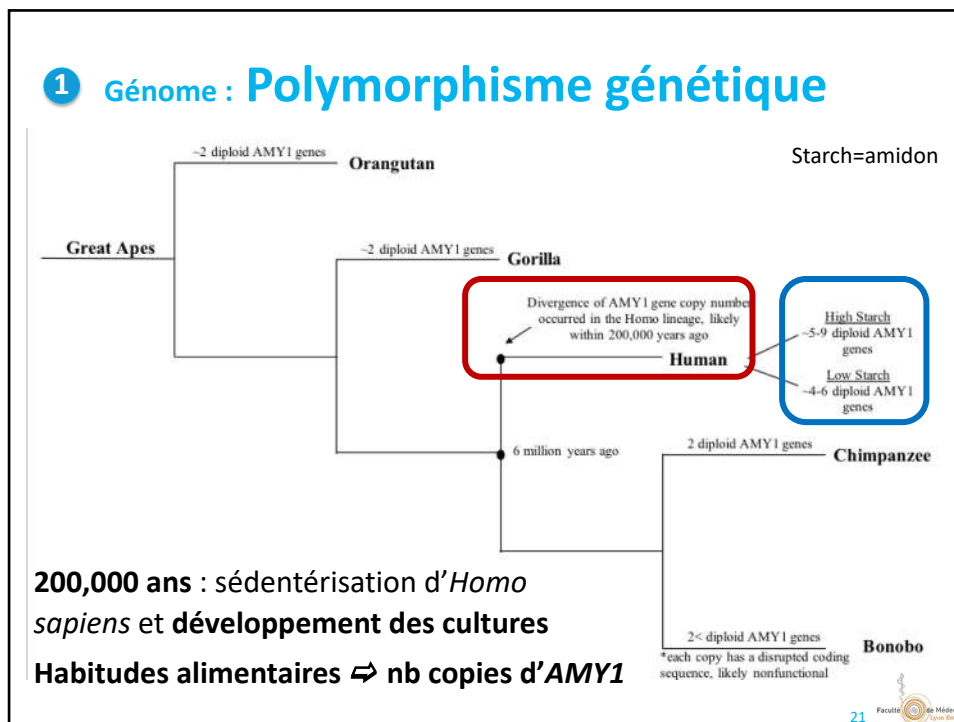
**Digestion de l'AMIDON**  
par l' $\alpha$ -amylase salivaire  
(codée par **AMY1**)

$\alpha$ -Amylase  
Amyloglucosidase

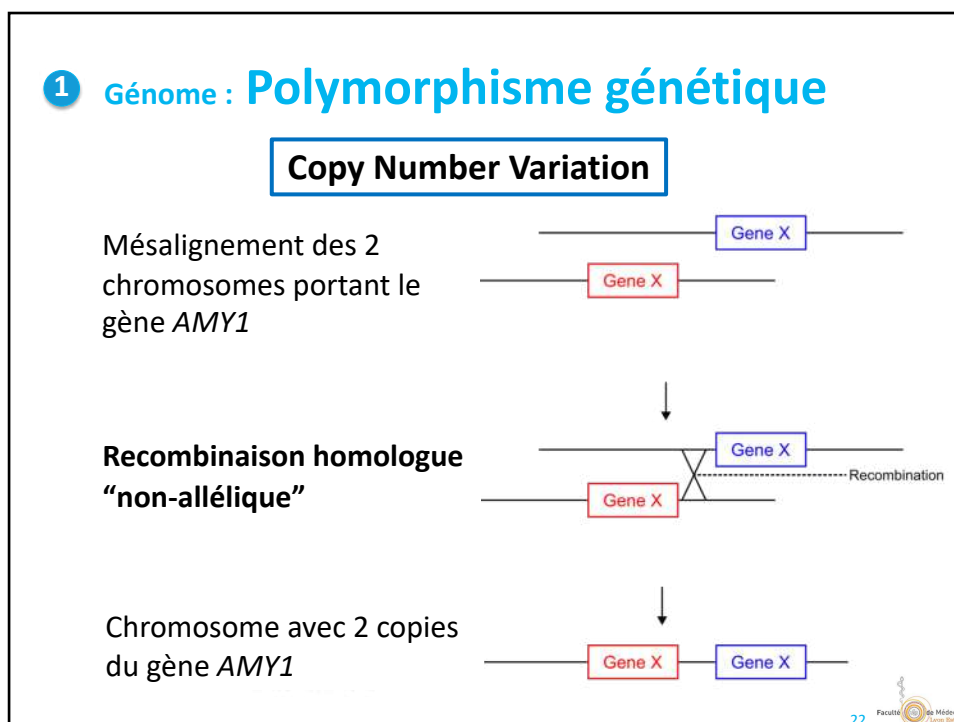
19



20



21



22

## 0 Plan

- Organisation du génome humain
  - Génotype vs phénotype
  - Polymorphisme génétique
  - **Génome de référence**
  - Composition du génome
- Gènes
- ADN intergénique

23

## 1 Génome : Génome de référence



- **13 ans (1990-2003), 3 milliards de \$**
- 20 équipes (USA, UK, Chine, Japon, Allemagne et France)
- **Génome humain haploïde de référence = mosaïque de quelques individus**

24

**1** Génome : **Génome de référence**

Chr: 20  
Position: 100298

T T GR:38  
REF ALT

En cas de SNP  $\Rightarrow$  l'allèle le plus fréquent est retenu

25 Faculté de Médecine  
Lyon 08

25

**1** Génome : **Génome de référence**

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y

- ◄ Region containing alternate loci
- Region containing fix patches
- Region containing novel patches

NNNNNNNNNNNNNNNN...

Certaines régions (télomères +++) étaient mal séquencées

26 Faculté de Médecine  
Lyon 08

26

## 1 Génome : Plan France Médecine Génomique

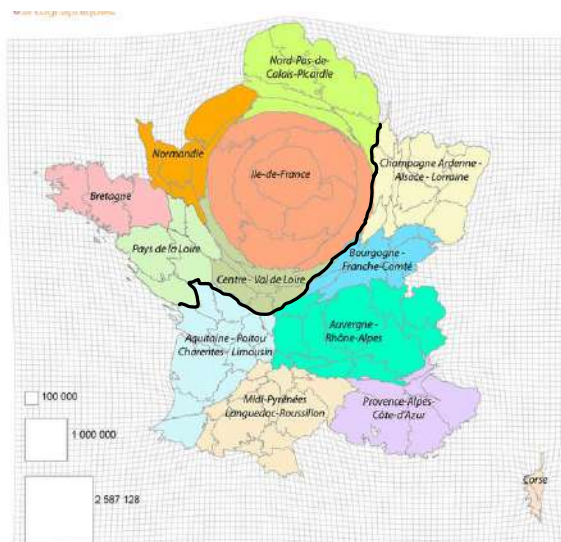
**SeqOIA**

LA PLATEFORME GÉNOMIQUE  
DE PARIS RÉGION

Sequencing, Omics,  
Information Analysis



Auvergne Rhône Alpes  
Genomique



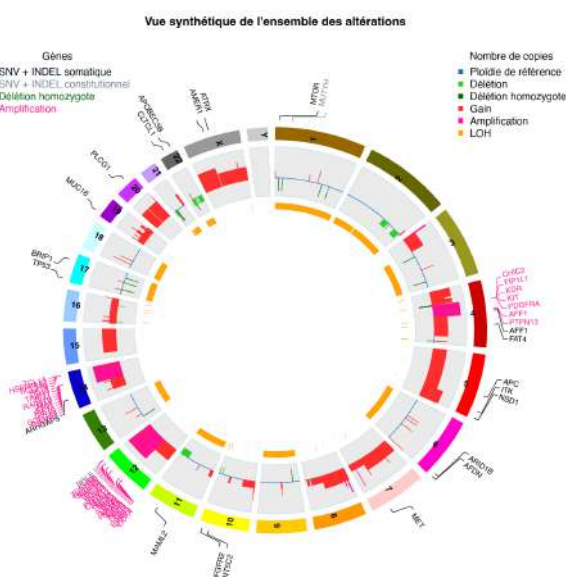
27

Faculté de Médecine  
Paris Saclay

27

## 1 Génome : Plan France Médecine Génomique

- "Séquençage h
- 48 WGS simult:
- Environ **1000€**
- Résultat "brut"  
<1 semaine



28

Faculté de Médecine  
Paris Saclay

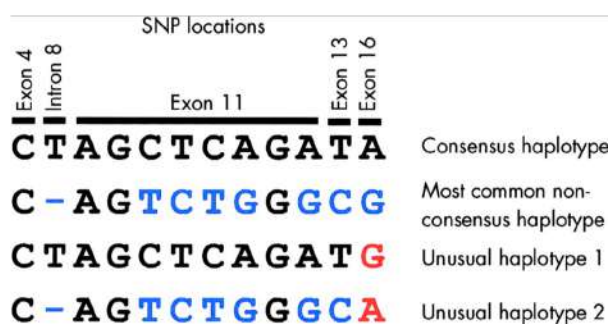
28



## 1 Génome : Haplotypes

Important

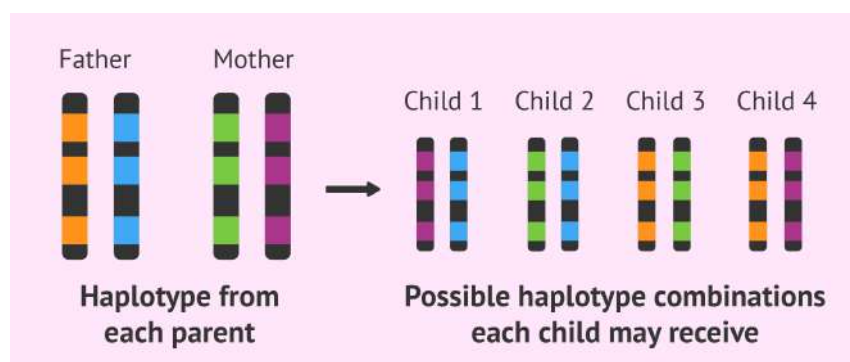
- Haplotype = groupe de SNPs séquentiels sur un même chromosome
- A l'intérieur de chaque haplotype **forte liaison entre les SNP** (très peu de recombinaisons, transmis en bloc)  
⇒ étude de la **transmission** intra-familiale



29

## 1 Génome : Haplotypes

- Haplotypage = identification des chromosomes parentaux transmis



30

## Posez vos questions pendant le cours



- 1 Allez sur [wooclap.com](https://wooclap.com)
- 2 Entrez le code d'événement dans le bandeau supérieur


Code d'événement  
**PJINGO**

31 

31

## 0 Plan

- Organisation du génome humain
  - Génotype vs phénotype
  - Polymorphisme génétique
  - Génome de référence
  - Composition du génome
- Gènes
- ADN intergénique

32 

32

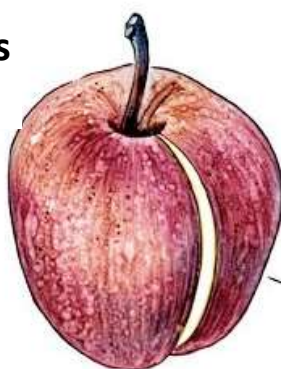


## 1 Génome : Composition du génome

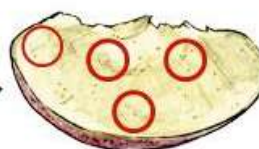
Important

- Seuls 1,5% du génome code pour des protéines (21000 gènes) ⇔ « exome »

**Génome**  
(3,2 milliards  
de bases)



**Exome = 1,5%**  
(48 millions de bases)



33 Faculté de Médecine  
Université de Lyon

33

## 3 Génome : Composition du génome

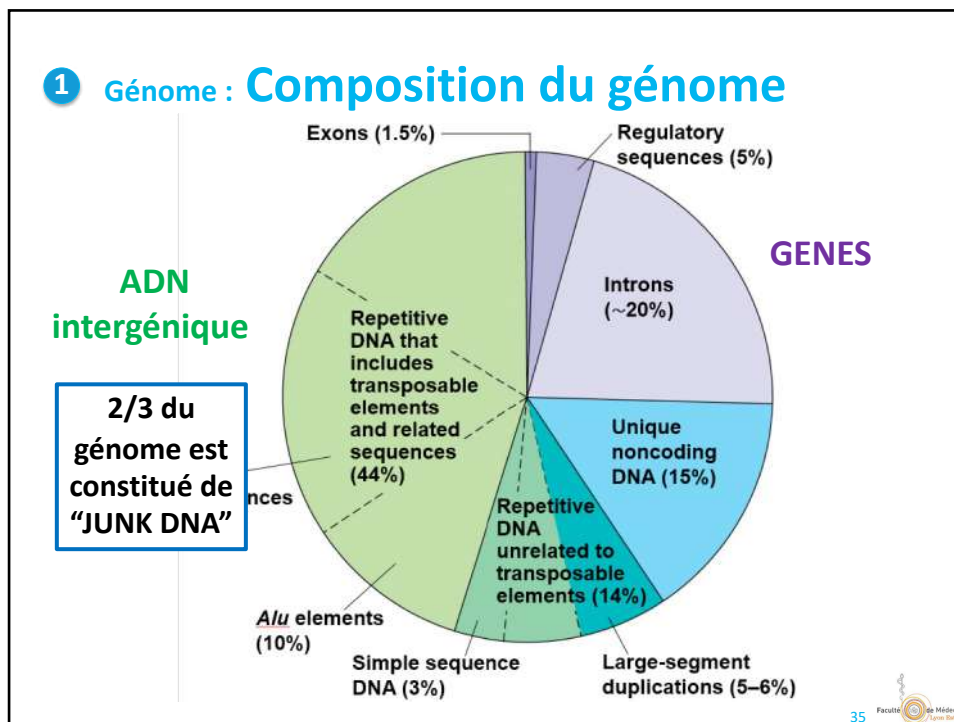
Taille de génomes **haploïdes**  
(en Millions de bases, Mégabases, Mb)

|                          |        |
|--------------------------|--------|
| • E. coli                | 4,5    |
| • S. cerevisiae (levure) | 14,4   |
| • C. elegans             | 100    |
| • D. melanogaster        | 165    |
| • H. sapiens             | 3,000  |
| • Arabidopsis thaliana   | 130    |
| • L. esculentum (tomate) | 1,000  |
| • N. tabacum (4x)        | 4,500  |
| • N. plumbaginifolia     | 2,300  |
| • B.napus (colza)        | 1,200  |
| • O. sativa (riz)        | 420    |
| • Z. mais                | 2,500  |
| • T. aestivum (blé)      | 16,000 |
| • Tulipa                 | 30,000 |

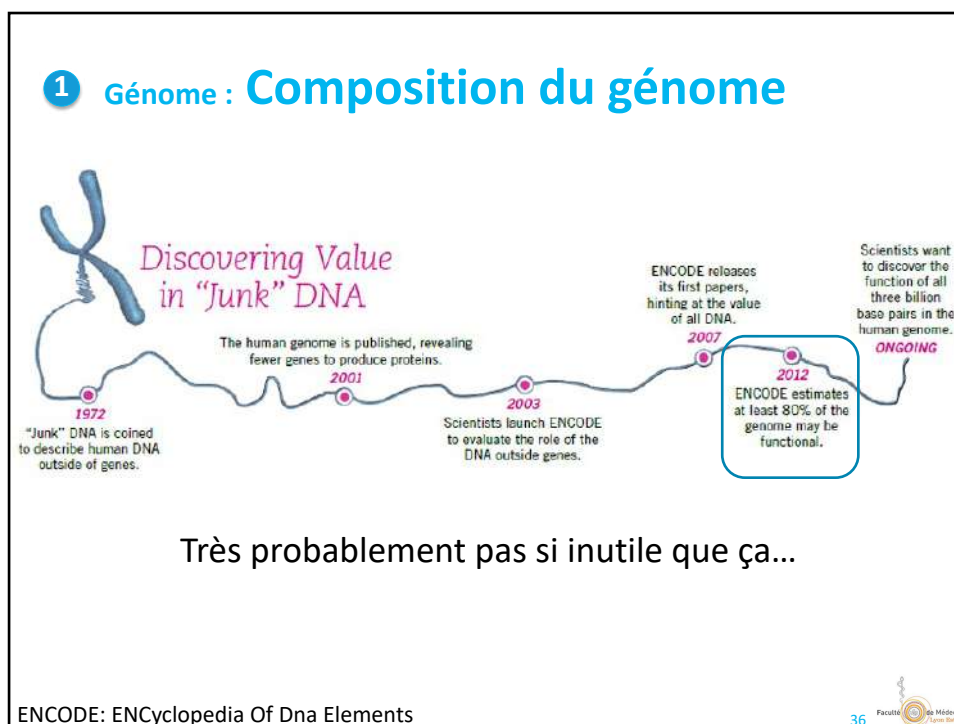
**Génome humain**  
**3 milliards (3 Gb)**  
**de nucléotides**

34 Faculté de Médecine  
Université de Lyon

34

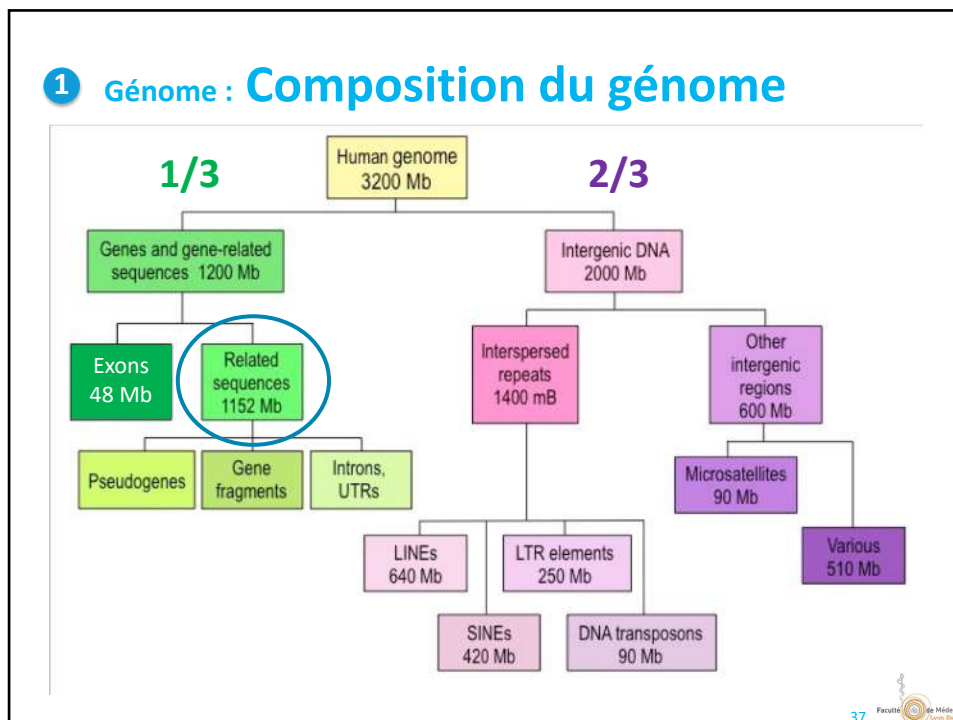


35



36

## 1 Génome : Composition du génome



37

## 1 Génome : Composition du génome

Table 2. Number of human genes in different databases as of July 2018[18]

|                      | Gencode <sup>[19]</sup> | Ensemble <sup>[20]</sup> | Refseq <sup>[21]</sup> | CHESS <sup>[22]</sup> |
|----------------------|-------------------------|--------------------------|------------------------|-----------------------|
| protein-coding genes | 19,901                  | 20,376                   | 20,345                 | 21,306                |
| lncRNA genes         | 15,779                  | 14,720                   | 17,712                 | 18,484                |
| antisense RNA        | 5501                    |                          | 28                     | 2694                  |
| miscellaneous RNA    | 2213                    | 2222                     | 13,899                 | 4347                  |
| Pseudogenes          | 14,723                  | 1740                     | 15,952                 |                       |
| total transcripts    | 203,835                 | 203,903                  | 154,484                | 328,827               |

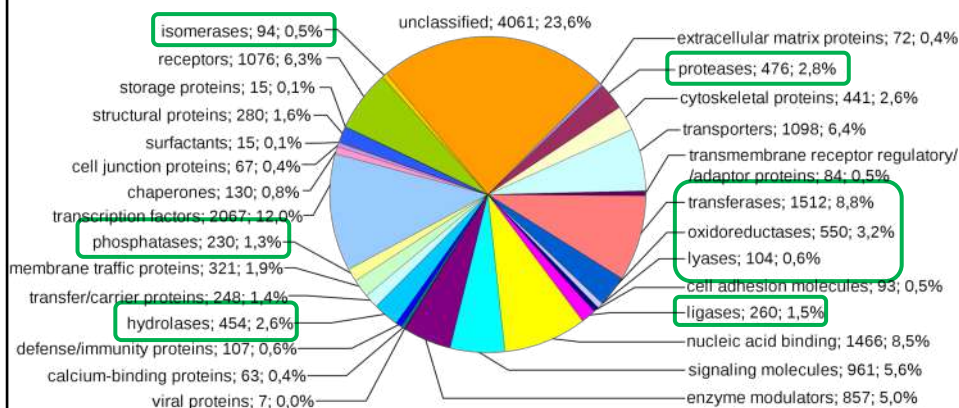
Gènes codants  $\approx$  gènes non codants > pseudogènes

Nb transcrits  $\gg$  nb de gènes (1 gène  $\Leftrightarrow$  +s transcrits)

38 Faculté de Médecine

38

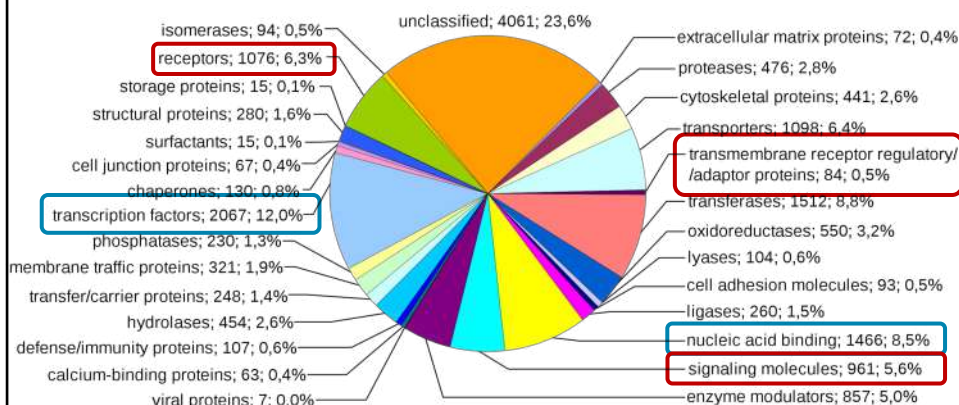
## 1 Génome : Composition du génome



Les gènes codent pour des protéines de fonctions différentes

39

## 1 Génome : Composition du génome




Les gènes codent pour des protéines de fonctions différentes

40

**1** **Géome : Composition du génome**


Les gènes sont souvent regroupés par homologie de séquence et fonction biologique  
 ⇒ **“CLUSTER”**

41 

41

**0** **Plan**

- Organisation du génome humain
- **Gènes**
  - Structure d'un gène eucaryote codant
  - Pseudogènes
- ADN intergénique

42 

42

## 2 Génome : Structure d'un gène eucaryote codant

Important

**Régions "cis" régulatrices**  
 ⇒ contrôle de la transcription  
 (voir cours d'Alexandre Janin + cours "régulation")

**Région transcrite**  
 ⇒ pré-ARNm

**Morcelée en Exons (codants) et Introns (non codants)**

© 2012 Pearson Education, Inc. Faculté de Médecine

43

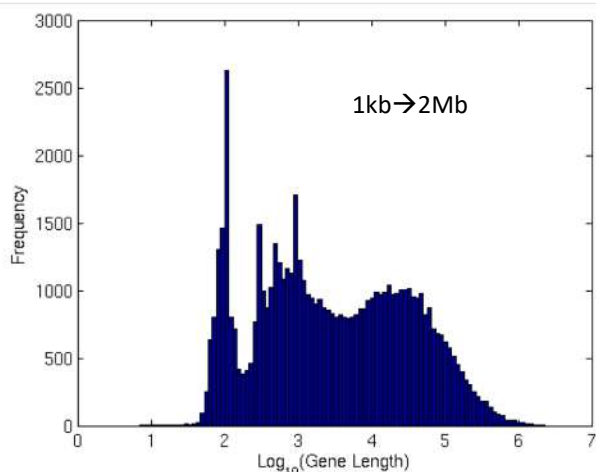
## 2 Génome : Structure d'un gène eucaryote codant

**1 seul brin est transcrit en ARN = brin sens OU antisens**

Faculté de Médecine

44

## 2 Génome : Structure d'un gène eucaryote codant



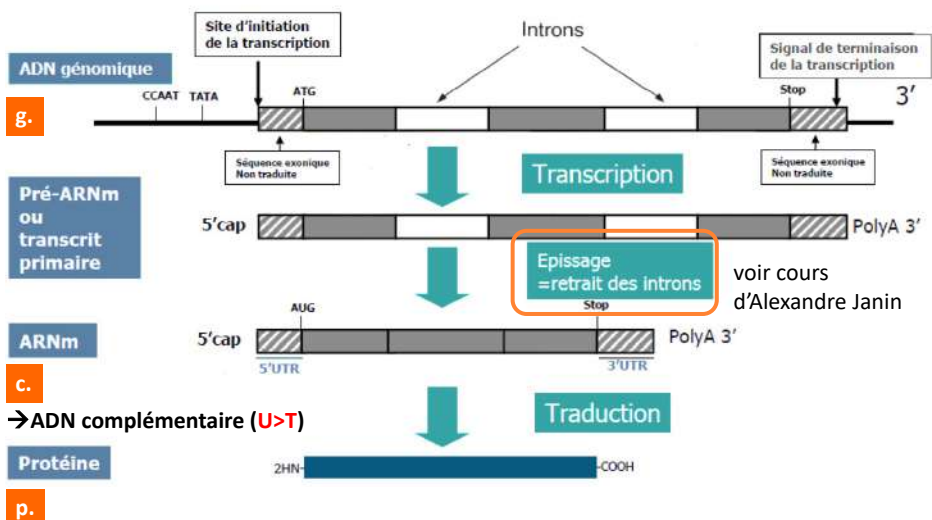
1 des + longs: *DMD*  
codant pour la **dystrophine**, 2,2Mb

Le + d'exons: *TTN*  
codant pour la **titine**,  
350 exons

45

## 2 Structure d'un gène eucaryote

Important



46



## ② Structure d'un gène eucaryote

Important

**Introns**

**Exons**

BRAF - B-Raf proto-oncogene, serine/threonine kinase | GRCh37 (Chr 7)

Overview of Transcript NM\_004333.5

Genome - chr7:140,624,512-140,624,355 (GRCh37) - 158 bps

Nucleotide Conservation

NM\_004333.5: Homo sapiens B-Raf proto-oncogene, serine/threonine kinase (BRAF), transcript variant 1, mRNA

Genome - chr7:140,453,196-140,453,039 (GRCh37)

Position chromosomique

ADNc → 3TCTATATAAAGAAGTACTTCTGGAGTGTCAATTTTATCCAGAGTATATTTCTTCATGAAGACCTCACAGTAAAAATAGG

Position sur le transcrit

Protéine → N I F L H E D L T V K I G

Position sur la protéine

47 Faculté de Médecine

47

## ② Structure d'un gène eucaryote

BRAF - B-Raf proto-oncogene, serine/threonine kinase | GRCh37 (Chr 7)

Overview of Transcript NM\_004333.5

Genome - chr7:140,453,215-140,453,058 (GRCh37) - 158 bps

Nucleotide Conservation

NM\_004333.5: Homo sapiens B-Raf proto-oncogene, serine/threonine kinase (BRAF), transcript variant 1, mRNA

Chr7:g.1404531136T>A : substitution d'1 nucleotide sur le gène (nomenclature g. rarement utilisée en routine)

c.1799T>A : modification d'1 nucléotide sur l'ADN complémentaire (essentiel d'indiquer quel transcrit est utilisé comme référence)

p.600Val>Glu ou p.600V>E : "La valine en position 600 de la protéine BRAF est remplacée par un glutamate"

48 Faculté de Médecine

48



## 2 Structure d'un gène eucaryote

**Gène** → EGFR

**Génome de référence** → GRCh38 (hg...)

**Version du transcrit** → NM\_005228.4

**Transcrits alternatifs** → List of transcripts including NM\_005228.4

| Build          | Transcript     | Exons | AAs  | Description  |
|----------------|----------------|-------|------|--|
| NM_201284.1    | NM_201284.1    | 16    | 705  | This variant (4) uses a different 3' terminal exon when ...  |
| NM_201282.1    | NM_201282.1    | 16    | 628  | This variant (2) uses a different 3' terminal exon when ...  |
| NM_201283.1    | NM_201283.1    | 10    | 405  | This variant (3, also known as ErbB1-S) uses a different...  |
| GRCh38 (hg...) | NM_005228.4    | 28    | 1210 | This variant (1) encodes the longest isoform (a).            |
|                | NM_001346899.1 | 27    | 1165 | This variant (7) lacks an in-frame exon in the 5' coding ... |
|                | NM_001346900.1 | 28    | 1157 | This variant (8) uses a novel 5' terminal exon compared...   |
|                | NM_001346898.1 | 27    | 1136 | This variant (6) has a 3' terminal exon that extends pas...  |
|                | NM_001346897.1 | 26    | 1091 | This variant (5) lacks an in-frame exon in the 5' coding ... |
|                | NM_001346941.1 | 22    | 943  | This variant (EGFRvIII, also known as delta-EGFR and d...    |
|                | NM_201284.1    | 16    | 705  | This variant (4) uses a different 3' terminal exon when ...  |
|                | NM_201282.1    | 16    | 628  | This variant (2) uses a different 3' terminal exon when ...  |
|                | NM_201283.1    | 10    | 405  | This variant (3, also known as ErbB1-S) uses a different...  |
| LRG_304        | LRG_304        | 28    | 1210 | Epidermal growth factor receptor, transcript variant 1       |

Transcript Details: **NM\_005228.4 (on GRCh38) - 28 exons - 1210 AAs**

This variant (1) encodes the longest isoform (a).

Buttons: Cancel, OK

49 Faculté de Médecine

49

## 2 Les séquences du concours...

**IMPORTANT** : L'**acide aminé** correspondant à chacun des codons est indiqué **sous la 1<sup>ère</sup> base** de celui-ci. Pour faciliter le repérage, les numérotations du 1<sup>er</sup> et du dernier nucléotide ou acide aminé sont indiquées pour chacune des lignes. Dans la séquence 2, les exons sont en majuscules et les bordures introniques en minuscules. (...) indique que certaines bases ne sont pas indiquées.

**Séquence 1**

```

1  GTCCGATGGGGCTCTGGTGGCGTGATCTGGCGGCCCCAGGCGTCAAGCACCCACACCCCTA 60
61  GAAGGTTCCGCAGCGACGTCGAGGCGCTCATGGTTGCAGGGGGCGCCCGCTTCAGTT 120
121 CA                                     : 180
181 CG ADNc/ARNm mature : pas d'intron : 240
241 CT                                     : 300
301 GCGTGGACCGCGCACGGCCCTTAGGTCCTCCGACAGGACGCAACCTCTCCCTGGCCC 360
361 TCATGGGCACCGTCAGCTCCAGGCGGTCTGGTGGCCGCTGCCACTGCTGCTGCTGCTGC 420
1  M G T V S S R R S W W P L P L L L L L L L L 20
421 TGCTGCTCTGGGTCCCGCGGGCGCCCGTGGCAGGAGGACGAGGACGGGACTACGAGG 480
21  L L L G P A G A R A Q E D E D G D Y E E 40
481 AGCTGGTCTAGCCTTTCGGTCCGAGGAGGACGCGCTGGCCGAGCACCCGAGCAGGAA 540
41  L V L A L R S E E D G L A E A P E H G T 60
541 CCACAGCCACCTTCCACCGCTGGCCCAAGSATCCGTGGAGGTTGCTGGCACCTACGTGG 600
61  T A T P H R C A K D P W R L P G T Y V V 80
    
```



**Séquence 2**

```

GTCCGATGGGGCTCTGGTGGCGTGATCTGGCGGCCCCAGGCGTCAAGCACCCACACCCCTA
GAAGGTTCCGCAGCGACGTCGAGGCGCTCATGGTTGCAGGGGGCGCCCGCTTCAGTT
AGACTGGCTCGGGGGCGGGGACG
IGATCCGCGCCCTTCACCGCC
XCACGCAAGGCTCAAGGCGCCGCG
GCGTGGACCGCGCACGGCCCTTAGGTCCTCCGACAGGACGCAACCTCTCCCTGGCCC
TCATGGGCACCGTCAGCTCCAGGCGGTCTGGTGGCCGCTGCCACTGCTGCTGCTGCTGC
1  M G T V S S R R S W W P L P L L L L L L L L 20
TGCTGCTCTGGGTCCCGCGGGCGCCCGTGGCAGGAGGACGAGGACGGGACTACGAGG
21  L L L G P A G A R A Q E D E D G D Y E E 40
AGCTGGTCTAGCCTTTCGGTCCGAGGAGGACGCGCTGGCCGAGCACCCGAGCAGGAA
41  L V L A L R S E E D G L A E A P E H G T 60
CCACAGCCACCTTCCACCGCTGGCCCAAGSATCCGTGGAGGTTGCTGGCACCTACGTGG
61  T A T P H R C A K
(...)tcatgctctcttgatggggccagGATCGTGGAGGTTGCTGGCACCTACGTGG
70  D P W R L P G T Y V V 80
    
```

**ADN/gène : présence d'introns**

50 Faculté de Médecine

50

## ② Structure d'un gène eucaryote

Important

**Attention!** la numérotation de l'ADNc (c.) commence à l'ATG  
= ne comprend pas le 5'UTR...

5'UTR numéroté en c.-

START traduction

51 Faculté de Médecine

51

## ② Structure d'un gène eucaryote

Important

**Attention!**  
Le codon START (ATG) n'est pas toujours localisé dans le premier exon

Exon 1 : non codant

NM\_002524.5: Homo sapiens NRAS proto-oncogene, (

...GGTCTTGCTGGTGTGAATGACTGAGTACAAACTG...  
M T E Y K L

52 Faculté de Médecine

52

## ② Structure d'un gène eucaryote

**Important**

**Attention!**

Le codon STOP (TAA, TAG ou TGA) est **toujours** localisé dans le dernier exon qui code aussi pour le 3'UTR (site polyA)

Codon STOP

53

## ② Structure d'un gène eucaryote

**Important**

Fin de l'exon | Début de l'intron

NM\_004333.5: Homo sapiens B-Raf proto-oncogene, serine/threonine kinase (BRAF), transcript variant 1, mRNA.  
c.90 c.100 c.110 c.120 c.130 c.138 c.138+10 c.138+20 c.138+30  
CGCCGCGCCGCGCCGCGCCCTCTTCGGCTGCGGACCGTCCGATTCGGGAGGAGGTAGTGTGGCGCCACCCTGCCGCGCTCC  
G A G A A A S S A A D P A I P E E  
30 35 40 45 46

c.138+1/+2  
GT=site donneur d'épissage

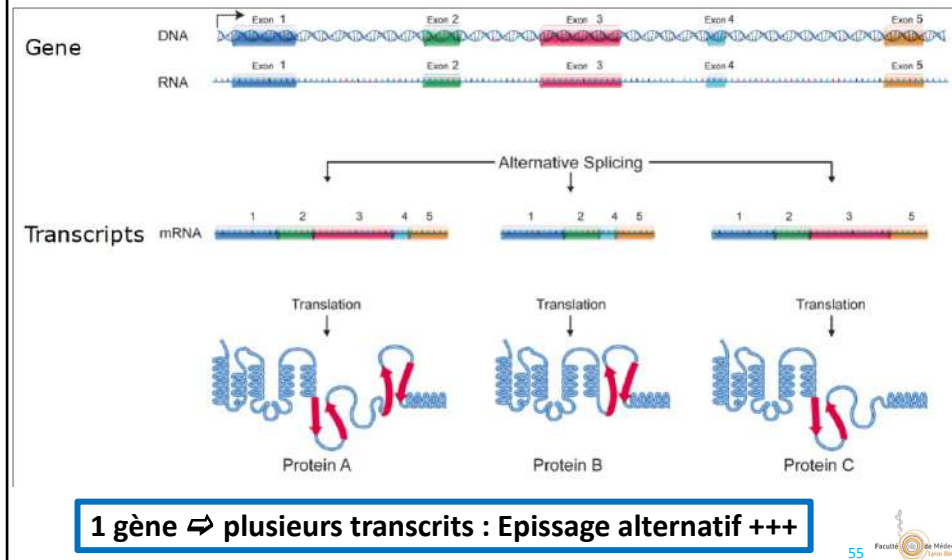
Fin de l'intron | Début de l'exon

NM\_004333.5: Homo sapiens B-Raf proto-oncogene, serine/threonine kinase (BRAF), transcript variant 1, mRNA.  
c.139-40 c.139-30 c.139-20 c.139-10 c.139 c.150 c.160 c.170  
ACTAAGTGATTTTTATTTAAATTTCTTTGAAATATTTT  
V W N I K Q M I K L T Q E  
47 50 55

c.139-2/-1  
AG=site accepteur d'épissage

54

## ② Epissage alternatif des pré-ARN

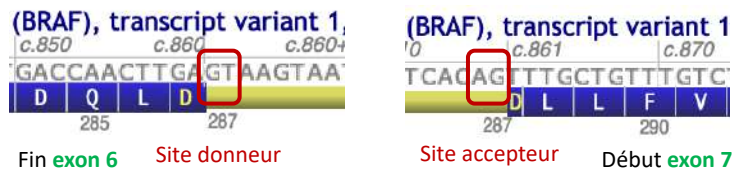


55

## ② Structure d'un gène eucaryote

Important

**Attention!** parfois un codon chevauche la jonction exon-exon



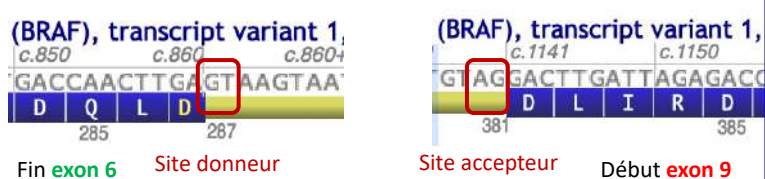
GA/.../T ⇒ GAT ⇒ D (Asp)

58

## ② Structure d'un gène eucaryote

Important

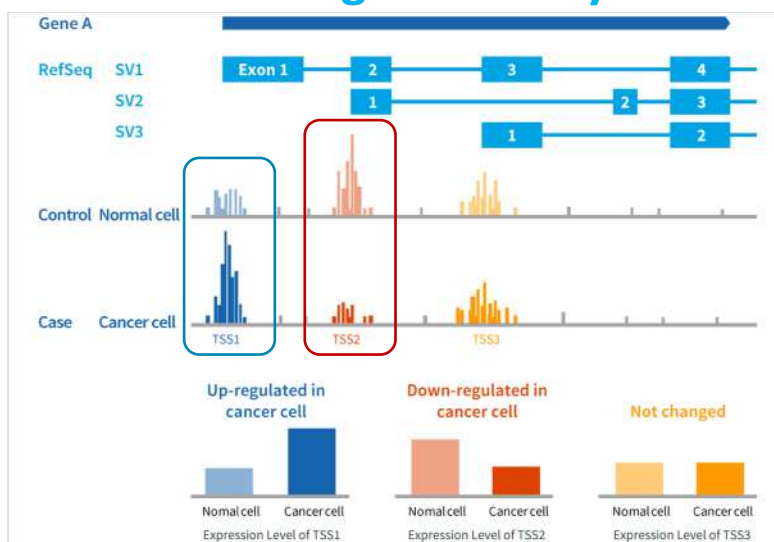
**Attention!** parfois un codon chevauche la jonction exon-exon si épissage alternatif  $\Rightarrow$  possible décalage du cadre de lecture...



GA/.../G  $\Rightarrow$  GAG  $\Rightarrow$  E (Glu)  
 puis ACT (Thr)  
 puis TGA (STOP)

59

## ② Structure d'un gène eucaryote



1 gène  $\Rightarrow$  plusieurs transcrits : Transcription Start Sites alternatifs

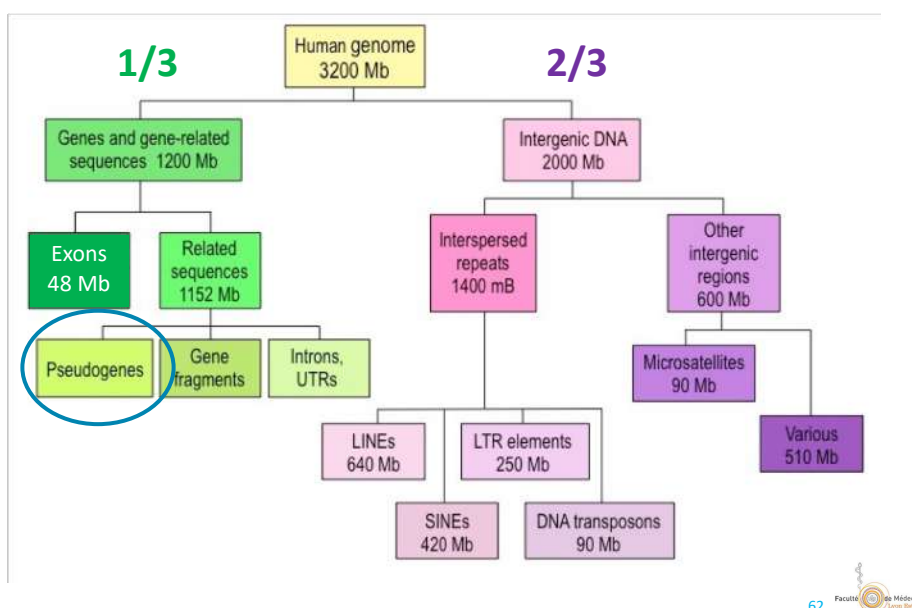
60

## 0 Plan

- Organisation du génome humain
- Gènes
  - Structure d'un gène eucaryote codant
  - Pseudogènes
- ADN intergénique

61

## 2 Génome : Composition du génome

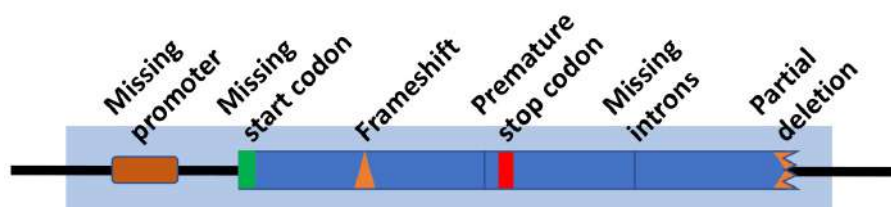


62

## 2 Génome : Pseudogènes

- Segments d'ADN présentant des **homologies de séquence avec des gènes actifs**
- **Perte de fonction par rapport au gène parental** ⇨ pas de protéine ou protéine non fonctionnelle

### Common defects of pseudogenes:

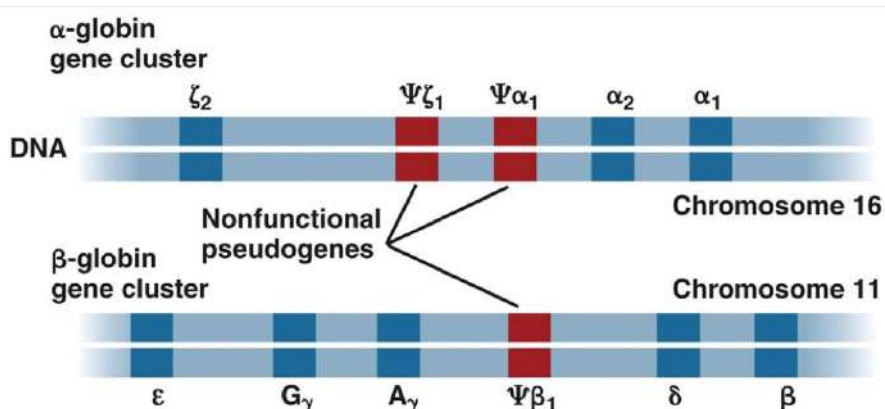


63 Faculté de Médecine

63

## 2 Génome : Pseudogènes

- Ex: clusters des gènes codant pour les **globines**



PRINCIPLES OF LIFE, Figure 12.11  
© 2012 Sinauer Associates, Inc.

Problème pour le séquençage et l'interprétation des variants...

64 Faculté de Médecine

64

## Posez vos questions pendant le cours



- 1 Allez sur [wooclap.com](https://wooclap.com)
- 2 Entrez le code d'événement dans le bandeau supérieur

Code d'événement  
**PJINGO**

65 Faculté de Médecine Lyon-Sud

65

## 0 Plan

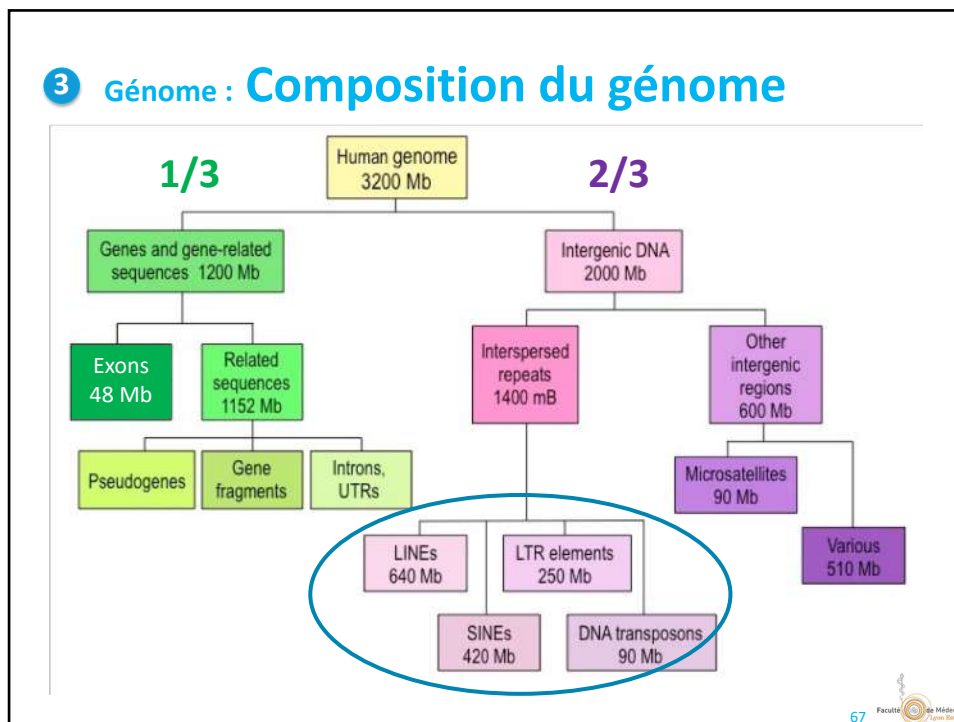
- Organisation du génome humain
- Gènes
- ADN intergénique
  - Séquences répétées
  - Microsatellites

66 Faculté de Médecine Lyon-Sud

66



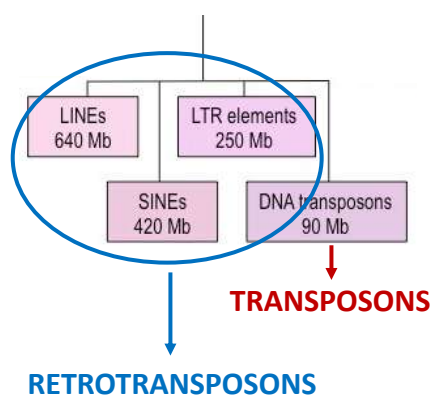
### 3 Génome : Composition du génome



67

### 3 Repeats : Séquences répétées dispersées

- 45% du génome humain
- **Dispersées** dans tout le génome
- 0,1 → 6kb
- **Moyennement répétitif**
- **Dérivent d'éléments "mobiles" ou "transposables"** capables de se déplacer
- Autonomes ou non



« gènes sauteurs »

68 Faculté de Médecine

68

### 3 Repeats : Séquences répétées dispersées

SCIENCES - GÉNÉTIQUE

## Ces virus qui ont colonisé notre génome : amis ou ennemis ?

Plus de 60 % de notre génome héberge de l'ADN issu d'entités virales qui ont, jadis, infecté de lointains ancêtres. Sans ce puissant moteur de l'évolution, nous ne serions pas là aujourd'hui à remonter le fil de cette improbable saga. Mais cet ADN trahit parfois ses hôtes, provoquant des maladies comme des cancers.

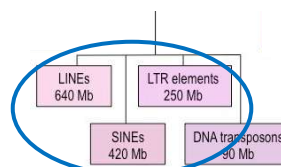
Par Florence Rosier

Publié le 26 février 2024 à 18h00, modifié le 22 avril 2024 à 12h41 · Lecture 10 min.

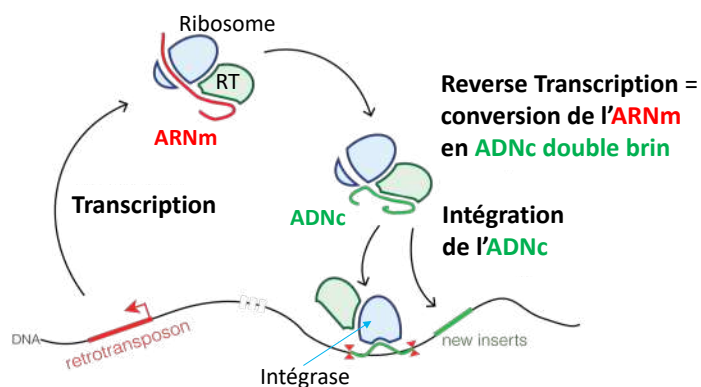


69

### 3 Repeats : Rétrotransposons

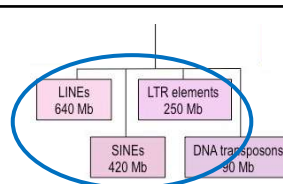


- Intermédiaire ARNm, rétrotranscription en ADNc puis intégration



70

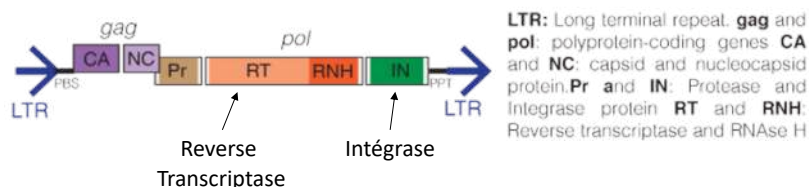
### 3 Repeats : Rétrotransposons



#### ■ LTR-rétrotransposons

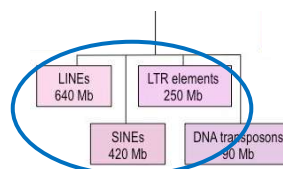
- contiennent des **Long Terminal Repeats**
- codent pour une **Reverse Transcriptase** et une **Intégrase** similaires à celles des rétrovirus.
- **Rétrovirus endogènes (HERV) : pas d'enveloppe (ENV) ⇔ non infectieux**

LTR-retrotransposon (5-20 kb)



71

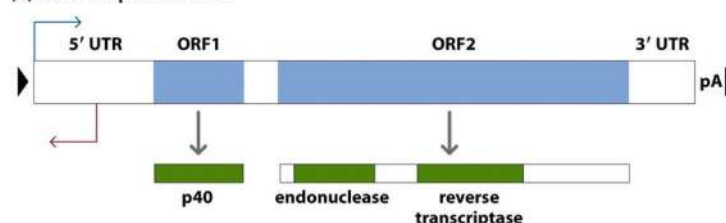
### 3 Repeats : Rétrotransposons



#### ■ Rétroposons : pas de LTRs

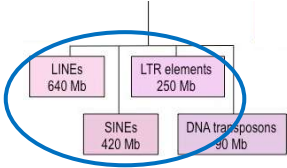
- **LINEs = Long Interspread Nuclear Elements**
  - $\approx 6\text{kb}$
  - codent pour une RT ( $\rightarrow$ ADNc) et une endonuclease ( $\rightarrow$ intégration)
  - transcrits par RNAPol II, 2 cadres de lectures non chevauchants

(A) LINE-1 repeat element



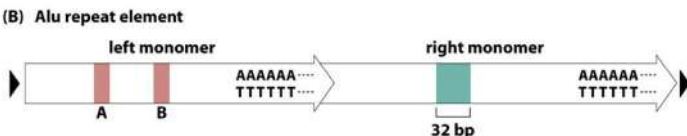
72

**3 Repeats : Rétrotransposons**



- **Rétroposons : pas de LTRs**
  - **SINEs = Short Interspersed Nuclear Elements**
    - 50-500bp
    - ne codent PAS pour une RT ⇨ NON autonomes
    - transcrits par RNAPol III
    - ARNt, ARNr 5S, ARN 7SL (**Alu**, B1)
    - Séquences Alu : 350bp, site *AluI*, 1.1 millions de copies

(B) Alu repeat element

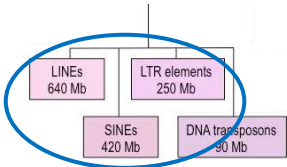


Promoteur RNAPol III

73 Faculté de Médecine

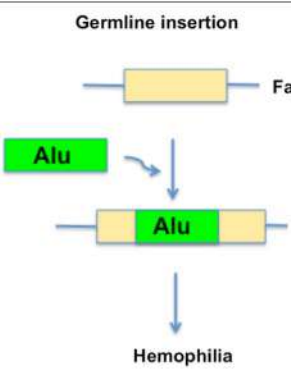
73

**3 Repeats : Rétrotransposons**



- **L'insertion d'éléments mobiles dans les gènes à l'origine de maladies**

Germline insertion

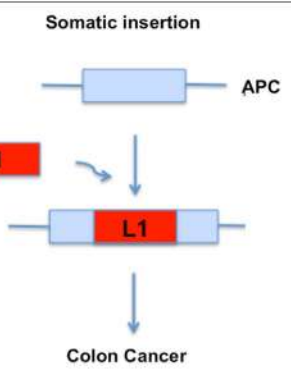


Factor VIII

Alu

Hemophilia

Somatic insertion



APC

L1

Colon Cancer

74 Faculté de Médecine

74

### 3 Repeats : Transposons

Transposon ADN code pour une **TRANSPOSASE**

Tandem Invert Repeat

**1/Liaison des transposases**

**2/Clivage**

**3/Transfert de la séquence capturée**

**Aucun transposon encore actif chez l'Homme**  
 (bactéries ou plantes +++)  
 ⇒ vestiges évolutifs  
 ⇒ rôle ?

75

### 3 Repeats : Transposons

**C gene (wild type)**

**c-m(Ds) (no Ac)**

**c-m(Ds) (+Ac)**

**c-m(Ac)**

**Phenotypes**

**Pigmented**

**Colorless**

**Spotted kernels**

**Spotted kernels**

**Chez le Maïs**

Figure 14-4 Introduction to Genetic Analysis, Ninth Edition © 2008 W.H. Freeman and Company

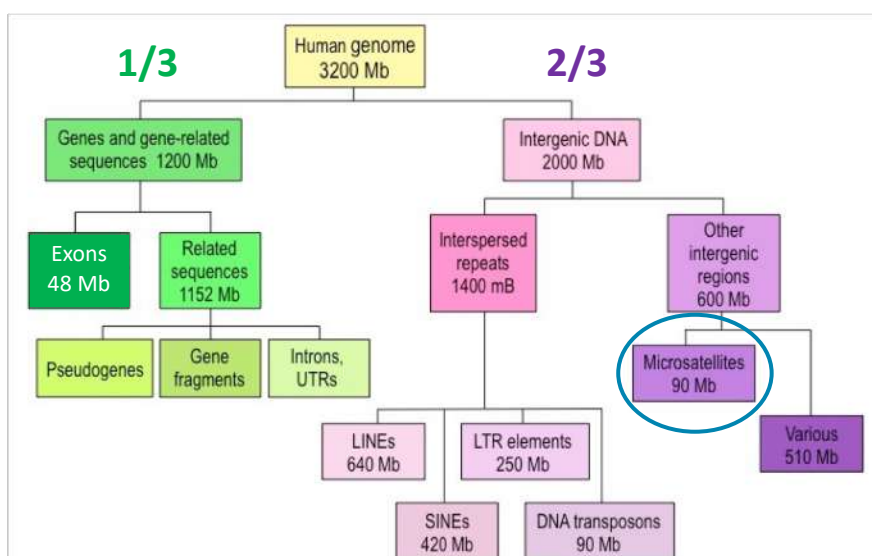
76

## 0 Plan

- Organisation du génome humain
- Gènes
- **ADN intergénique**
  - Séquences répétées
  - **Microsatellites**

77

## 2 Génome : Composition du génome



78

### 3 Repeats : Séquences répétées en tandem

- 10% du génome
- Dispersées dans le génome mais localisées (tjs aux mêmes endroits)
- Qqs pb à qqs centaines de pb
- Hautement répétitif

### 3 Repeats : Séquences répétées en tandem

- Plusieurs types
  - Satellites ⇨ au niveau des centromères
  - Minisatellites (Variable Number of Tandem Repeats ou VNTR) ⇨ 10-100pb répétées qqs milliers de fois  
Ex: motif TTAGGG répété dans les télomères

### 3 Repeats : Séquences répétées en tandem

- Microsatellites (Short Tandem Repeats ou STR)
  - ⇒ 1 à 8pb répétées en tandem de 5 à 50 fois

Repeat of ATAG  
(CA)<sub>n</sub> Dinucleotide repeat

Alleles  
#1 → CACACACACACACACACACA  
#2 → CACACACACACACACACACA  
#3 → CACACACACACACACACACACA

Genotypes  
1/1 2/2 3/3 1/2 1/3 2/3

Legend:  
→ Forward primer  
← Reverse primer  
— Flanking sequence

81

### 3 Repeats : Séquences répétées en tandem

- Microsatellites (Short Tandem Repeats ou STR)
  - Instabilité des microsatellites (MSI) dans certains cancers colorectaux par perte de protéines de réparation de l'ADN  
⇒ réponse à l'immunothérapie

a  
MLH1 PMS2 MSH2 MSH6

b  
Adenoma LY04-4  
Non-tumor LY04  
NR-21 BAT26 (blue) BAT25 (green) NR-24 NR-22

82



## Posez vos questions pendant le cours



- 1 Allez sur [wooclap.com](https://wooclap.com)
- 2 Entrez le code d'événement dans le bandeau supérieur

Code d'événement  
**PJINGO**

83 Faculté de Médecine Lyon-Sud

83

## Génome : évaluez-vous !



- 1 Allez sur [wooclap.com](https://wooclap.com)
- 2 Entrez le code d'événement dans le bandeau supérieur

Code d'événement  
**PJINGO**

84 Faculté de Médecine Lyon-Sud

84



Merci de votre attention