

# BIOLOGIE MOLÉCULAIRE

Intitulé du cours : Mutabilité, maintenance & dynamique du génome

Rédacteur : Milan (ex Archéus)  
Ronéo n° : 4 (dernière)



**Corporation des Carabins Niçois**

UFR Médecine  
28, av. de Valombrose  
06107 Nice Cedex 2  
<http://carabinsnicois.fr/>  
[roneo.c2n@gmail.com](mailto:roneo.c2n@gmail.com)

*Partenaires*



## Mot de présentation

Suite à quelques petits soucis d'organisation, vous vous retrouvez à quelques semaines de l'examen avec le dernier cours de biomol. Pas de panique cependant, c'est tranquille (et c'est pas comme si c'était une ronéo de physio non plus faut pas déconner). Ça se torche en 2-2, croyez-moi. Je vous ai mis l'intégralité de l'index de la biomol, comme ça vous avez tout sur une seule page 😊.



## Index

A : adénine	MAR : matrix attachment regions
ADN : acide désoxyribonucléique	MMR : mutation mismatch repair
AMPC : adénosine monophosphate cyclique	Mre11 : meiotic recombination 11
Aménorrhées : lit. « absence de règles »	MSI : microsatellite instability
Amniocentèse : lit. « piqure de l'amnios »	Mut(S/L/H) : mutator (S/L/H)
Aneuploïde : lit. « qui ne contient pas le bon nombre de K doubles »	NBS : nijmegen breakage syndrome
ARN : acide ribonucléique	NER : nucleotide excision repair
ARNm : ARN messenger	NGS : next-generation sequencing
ARNr : ARN ribosomique	NHEJ : non-homologous end joining
ARNt : ARN de transfert	N-ter : azote terminal
ATP : adénosine triphosphate	Octamère : lit. « 8 parties »
BER : base excision repair	ORF : open reading frame
BLM : bloom syndrome	PARP : poly (ADP-ribose) polymerase
C : cytosine	Pdb : paire de bases
CAP : catabolite activator protein	Rad50 : radiation sensitive 50
CNV : copy number variation	RISC : RNA induced silencing complex
CpG : Cytosine-phosphate-Guanine	RPA : replication protein A
CSB : cockayne syndrome B	SNP : single nucleotid polymorphism
Diploïde : lit. « aspect double »	Solénoïde : lit. « en forme de conduit »
DNP-PK : DNA-dependent protein kinase	STR : short tandem repeats
DNMT : DNA methyltransferase	T : thymine
E.Coli : escherichia coli	TC-NER : transcription-coupled NER
G : guanine	Télomère : lit. « fin de la partie »
GG-NER : global genome NER	U : uracile
Haploïde : lit. « aspect simple »	UTR : untranslated transcribed region
HNPCC : hereditary non polyposis colon cancer	VNTR : variable number of tandem repeats
IMG : interruption médicale de grossesse	XPC : xeroderma pigmentosum C
IRP : iron regulatory RNA-binding protein	XRCC : X-ray repair cross-complementing protein
IRE : iron response element	f : fréquence
K : chromosome	γ : gamète
LDL : low density proteins	γ-génèse : gamétogénèse

## Dédicaces

Je vous les mets ici mais faut les lire après avoir fini le cours ! J'espère que la biomol vous a plu, je suis un peu vieux maintenant pour vous donner des conseils sur comment bosser cette matière etc, demandez aux tuteurs ils sont là pour ça !

Dédicace à Louis Marais--Baglione, ancien tuteur de pharmaco, meilleur ami, énorme bg galbé comme on en voit rarement, qui me tanne depuis un moment pour être dans les dédis maintenant c'est fait ! (Il le mérite ceci dit).

Dédicace à tous les anciens tuteurs de biocell, une dynastie rarement égalée, tant par son charisme que sa grandeur, j'ai nommé le grand Médo, l'externe compétitif Duval, le petit Damien, Hugo, Tiff, Emma, Yamina l'Exilée et Aline la tata.

C'est fini pour moi, je vous souhaite bien du courage pour la dernière ligne droite, vous allez retourner cet exam !

## V – Mutabilité, maintenance & dynamique du génome



### A) Mutations & dynamiques du génome

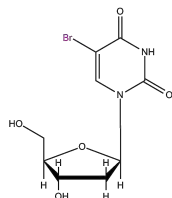
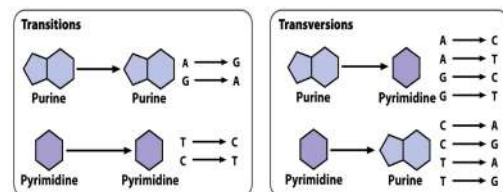
#### 1) Généralités

Une mutation est un changement dans la séquence d'ADN du génome d'une cellule. On distingue les mutations ponctuelles (substitutions et insertions ou délétions de petite taille) et les remaniements chromosomiques (délétion, duplication, insertion, inversion, translocation, etc). Certaines mutations sont liées aux erreurs inévitables de réplication, aux séquences répétées du génome qui favorisent les erreurs de réplication ou de *crossing-over*, ou liées aux modifications de bases. D'autres mutations sont liées à une exposition à un agent mutagène physique (radiations, rayons UV), chimiques (agents intercalants, analogues de bases...) ou pathogène (virus, bactéries). Les conséquences des mutations sont variables, selon qu'elles perturbent ou non le message génétique, constituant alors des polymorphismes, ou qu'elles sont somatiques ou germinales. À une échelle phylogénétique, les mutations ont constitué un moteur de sélection naturelle et favorisé l'évolution et la diversification des espèces, dont l'homme.

#### 2) Nature des mutations

##### a) Substitution, insertion & délétion

Parmi les substitutions, on distingue les transitions et les transversions. Une transition est une mutation qui remplace une purine ou une pyrimidine par une base de même nature, et une transversion remplace une purine par une pyrimidine ou inversement.



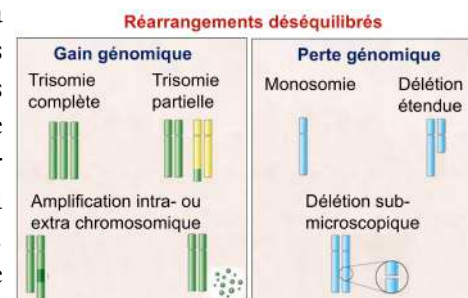
Les transitions sont les plus fréquentes et peuvent être causées par des agents mutagènes comme l'acide nitreux ou les analogues de bases tels que la 5-bromo-2-déoxyuridine (BrdU).

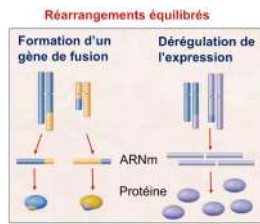
Lorsqu'elles affectent la séquence codante d'un gène, les mutations ponctuelles sont classées en trois catégories : les mutations silencieuses, faux-sens et non-sens.

Les insertions et les délétions peuvent entraîner un décalage du cadre de lecture de l'ARNm. Elles sont généralement plus sévères et sont notamment causées par des événements de transposition ou des erreurs de réplication au niveau de séquences répétées du génome.

##### b) Conséquences des remaniements chromosomiques

Les remaniements chromosomiques affectent le génome à une large échelle. Certains remaniements chromosomiques sont dits déséquilibrés car ils entraînent des gains ou des pertes de régions chromosomiques et donc des gènes que ces régions contiennent. Le gain génomique peut par exemple être lié à une trisomie partielle ou complète, ou à l'amplification intra ou extra-chromosomique d'une région. À l'inverse, la perte génomique peut être liée à une monosomie ou à une délétion chromosomique étendue ou de taille plus modeste (délétion sub-microscopique).



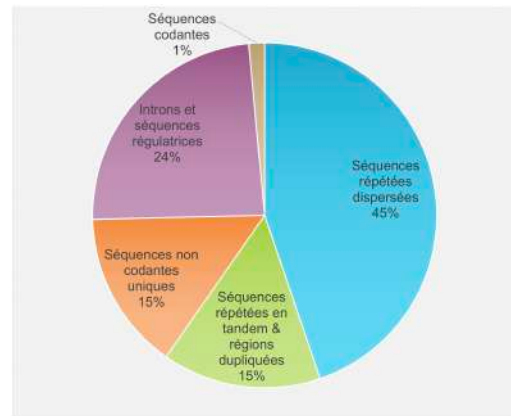


D'autres remaniements n'entraînent ni perte ni gain chromosomique et sont dits équilibrés (translocations équilibrées, inversions) mais juxtaposent des séquences d'ADN de gènes distants. Cette juxtaposition peut former un gène dit de fusion et une protéine modifiée possédant par exemple des propriétés oncogéniques (ex : gène de fusion Bcr-Abl) ou déréguler l'expression normale d'un gène.

### 3) Source des mutations

#### a) Erreurs de réplication

La source des mutations peut aussi venir d'erreurs de réplication liées à la *polymérase* ou aux séquences répétées du génome. Malgré la sélection stricte des bases par les *polymérases* et leur activité *proofreading*, la fidélité de la réplication est imparfaite. De plus, la nature des séquences du génome humain favorise en elle-même l'apparition de mutations, près de 60% du génome étant constitué d'autres séquences non codantes ou de séquences répétées. En effet, les régions codantes et non codantes des gènes ne représentent que 25% de l'intégralité du génome, le reste étant constitué d'autres séquences non codantes répétées.

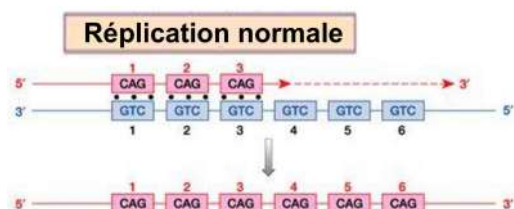


Les séquences répétées dites dispersées représentent 45% du génome et correspondent aux transposons et aux rétrotransposons qui favorisent les remaniements du génome et sa dynamique. Des séquences répétées en tandem représentent 5% du génome, correspondent aux minisatellites et aux microsatellites et favorisent les mutations et les insertions ou délétions de petite taille. Les régions génomiques dupliquées sont des familles de gènes apparentés issus d'un gène ancestral ayant évolué (duplication, mutation, transposition) et rendent compte de la dynamique passée du génome.

#### b) Mutations spontanées

##### α) Mutations spontanées liées aux erreurs de réplication

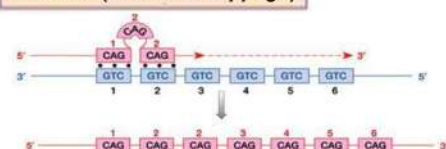
Les séquences répétées en tandem sont des séquences abondantes dans le génome et qui favorisent les erreurs de réplication. On distingue dans ces séquences répétées en tandem des minisatellites, constitués par la répétition d'un motif de 10 – 100 pnb et les microsatellites, constitués par la répétition d'un motif de 1 – 10 pnb. Les microsatellites sont des séquences formées de répétitions en tandem dont le motif répété est le plus souvent constitué de di-, tri- ou tétranucléotides (exemple : CAG CAG CAG...).



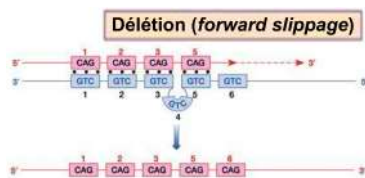
L'exemple ci-contre montre la répétition normale d'un microsatellite comprenant six répétitions du trinucléotide CAG. Dans certains cas, il peut se produire un mauvais alignement du brin parental et du brin fils en cours de réplication (dérapages réplicatifs) par formation de boucles sur l'un ou l'autre des brins.

Sur la figure de gauche, le glissement du brin fils forme une boucle contenant une répétition, ce qui conduit à sa réplication en excès et au final à une augmentation du nombre de répétition ( $n = 7$ ).

##### Insertion (backward slippage)



## Biologie moléculaire

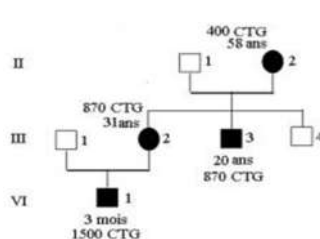
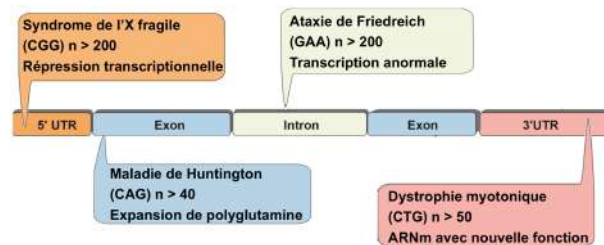


Sur la figure de droite, le glissement du brin parent forme là aussi une boucle contenant une répétition, ce qui conduit à son défaut de réplication et au final à une diminution du nombre de répétitions ( $n = 5$ ).

Ce phénomène est appelé instabilité des microsatellites (MSI) et peut conduire à l'apparition d'insertions ou de délétions (*indel*) d'un ou plusieurs nucléotides selon le motif répété. La conséquence va être un excès ou au contraire un défaut de réplication d'un nombre variable de motifs répétés, aboutissant à une augmentation ou à une diminution du nombre de répétitions.

Les maladies dites par expansion sont un groupe de maladies liées à l'augmentation de génération en génération du nombre de répétition du motif formant la séquence répétée. Certains gènes contiennent en effet dans leur séquence codante ou non codante des répétitions dont l'expansion va progressivement constituer une prémutation et au-delà d'un seuil critique, une mutation.

La figure ci-contre montre pour différentes maladies la localisation du motif répété dans le gène en cause, le nombre seuil de répétitions qui constitue une mutation et la conséquence de cette mutation. Comme la probabilité d'une erreur de réplication est de plus en plus élevée au fur et à mesure que le nombre de répétitions augmente, l'instabilité et l'expansion vont s'accroître d'une génération à l'autre.

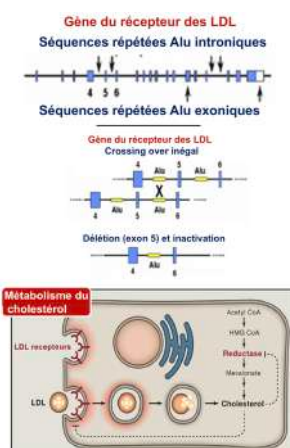


Le phénomène d'anticipation traduit l'augmentation des répétitions à chaque génération dont le nombre total est corrélé à la sévérité de la maladie ou à la précocité de son apparition. Cet arbre d'une famille atteinte de dystrophie myotonique de Steinert illustre ce phénomène, la maladie étant apparue à 58 ans chez la grand-mère maternelle, à 31 ans chez la mère et à 3 mois chez son fils.

### β) Mutations spontanées liées aux crossing-over inégaux en méiose

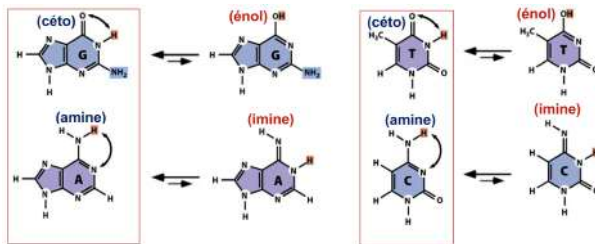
En méiose, les séquences répétées en tandem ou dispersées (exemple : séquence Alu) favorisent également l'instabilité des microsatellites ou les duplications ou délétions génomiques. Les crossing-over inégaux résultant d'un mauvais alignement entre séquences répétées entraînent une recombinaison pouvant soit aboutir à une duplication soit à une délétion.

À titre d'exemple, le gène qui code pour le récepteur des lipoprotéines de faible densité (appelées LDL) contient de nombreuses séquences répétées dispersées de type Alu. Ce récepteur permet de capter les LDL circulantes et après internalisation, le cholestérol libéré réduit sa propre synthèse endogène par rétro-contrôle négatif pour limiter le taux de cholestérol plasmatique. En cas d'alignement incorrect entre les séquences Alu encadrant par exemple l'exon 5 du gène, le *crossing-over* inégal entraîne la délétion de cet exon et la formation d'un récepteur non fonctionnel. Ainsi, certaines formes d'hypercholestérolémie familiale sont liées à un *crossing-over* inégal inactivant ce gène qui assure normalement la captation du cholestérol et la régulation de sa synthèse endogène.



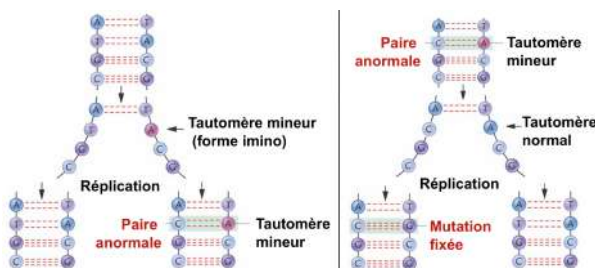
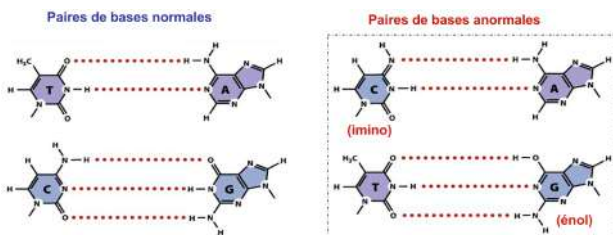


### γ) Mutations spontanées liées au phénomène de tautométrie



De façon spontanée, les bases peuvent subir une isomérisation de fonction appelée tautométrie<sup>†</sup> par déplacement d'un atome d'hydrogène et d'une double liaison. Cette isomérisation va convertir les groupes fonctionnels *céto* ou *amine* des bases et créer des formes tautomériques mineures des bases. Ce déplacement va ainsi convertir la fonction *céto* de la guanine et de la thymine en fonction *énol* et la fonction *amine* de l'adénine et de la cytosine en fonction *imine*.

Ces changements de groupe fonctionnel vont ainsi modifier les possibilités existantes de liaisons hydrogènes établies entre les bases et créer des paires de bases (A – C) ou (G – T) non canoniques.

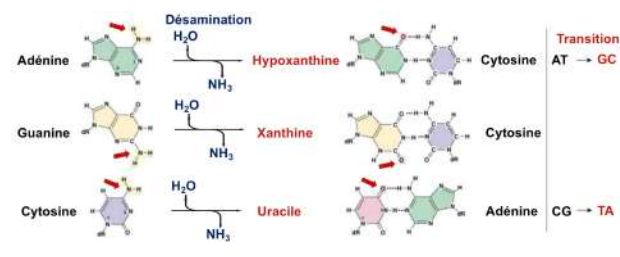
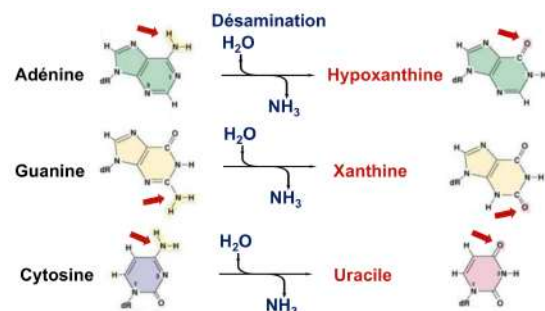


Si une base est présente dans l'ADN sous sa forme tautomérique mineure au moment de la réplication, un appariement anormal va se former et une mutation sera introduite sur le brin fils. En l'absence de détection et de réparation de cette paire de base erronée, même si le tautomère normal réapparaît, la réplication de la base de l'autre brin entraînera la fixation définitive de la mutation.

### δ) Mutations liées à la modification de bases par réactions chimiques inévitables

La dépurination correspond à la perte d'une adénine ou d'une guanine par rupture spontanée de sa liaison avec le désoxyribose, et dont le remplacement au hasard pourra introduire une mutation. La désamination correspond à la conversion spontanée ou parfois induite de la fonction *amine* d'une base en fonction *cétone* et qui peut également entraîner l'apparition d'une mutation.

Ce processus ne va concerner que l'adénine, la guanine, la cytosine, et la cytosine lorsqu'elle est sous-méthylée ; la thymine ne possédant pas de fonction *amine*. Ainsi, la désamination de l'adénine forme l'hypoxanthine, celle de la guanine forme la xanthine et celle de la cytosine forme l'uracile.

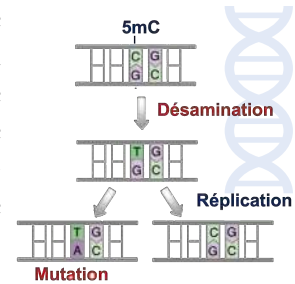


Lors de la réplication des bases ainsi désaminées, les nouvelles possibilités d'appariement entraînent la formation de paires de bases anormales, responsables après réplication de l'apparition de mutations.

<sup>†</sup> Étymologie : tautomérisation, du grec « ταυτός », qui signifie « même », et « μέρος », qui signifie « partie ».

# Biologie moléculaire

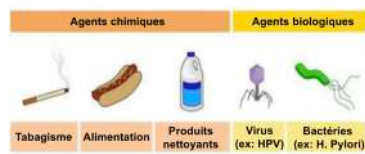
Le génome contient par ailleurs de dinucléotides CG dont la cytosine peut être méthyle (5mC) et désaminée, ce qui entraîne alors sa conversion en thymine. Dans ce cas, la réplication de cette thymine produira une paire de base TA et entraînera la disparition définitive du nucléotide de départ. Au niveau des régions activement transcrites, les dinucléotides CG à proximité des gènes sont peu méthylés et donc peu sensibles à cette désamination et peu enclins à disparaître.

[illegible]

En revanche, les dinucléotides des régions non transcrits contiennent fréquemment une cytosine méthylée dont la désamination entraîne progressivement leur disparition. Ces différences de méthylation expliquent l'abondance relative des dinucléotides CG à proximité des gènes et leur rareté dans les régions non codantes liées à leur disparition sélective dans ces régions.

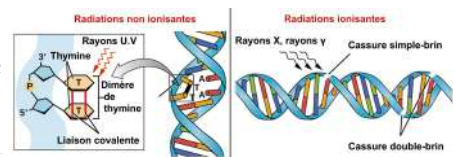
*ε) Mutations spontanées liées à l'exposition à des agents mutagènes*

Certains de ces agents exercent leur effet mutagène par le biais des dommages de l'ADN qu'ils provoquent, et d'autres induisent directement l'apparition des mutations. Ces agents peuvent être d'origine physique, comme par exemple les rayons UV, les rayons X utilisés en imagerie médicale, ou les radiations utilisées en radiothérapie.

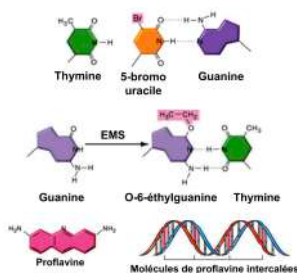


D'autres agents sont d'origine chimique, issus par exemple du tabagisme, d'agents conservateurs alimentaires (nitrates, nitrites), de la cuisson des aliments ou de produits de nettoyage ou cosmétiques. D'autres enfin sont d'origine biologique, comme les papillomavirus favorisant le cancer du col de l'utérus, la bactérie *Helicobacter pylori* favorisant le cancer gastrique, ou les transposons.

Les rayons UV exercent leur effet mutagène en induisant la formation de dimères entre thymines adjacentes de l'ADN, ralentissant la polymérase et favorisant les erreurs de réplication. Les radiations plus énergétiques (rayons X, rayons  $\gamma$  ou rayons cosmiques) exercent leur effet mutagène en induisant des cassures simple brin ou double brin de l'ADN.



Il existe de nombreuses classes d'agents mutagènes chimiques comme les analogues de bases, les agents alkylants, intercalants ou favorisant la désamination des bases et les radicaux libres. Les analogues de bases sont des composés dont la structure chimique est similaire aux bases de l'ADN avec lesquelles ils entrent en compétition durant la réplication, la *polymérase* ne pouvant les différencier.



La 5-bromouracile est un analogue de la thymine qui peut être introduit à sa place lors de la réplication, et qui s'appariera lors de la réplication suivante à la guanine (transition TA – CG). Les agents alkylants comme l'éthylméthanesulfonate modifient les bases et leurs propriétés d'appariement en leur ajoutant des groupes *alkyl* ( $-\text{CH}_3$ ,  $-\text{CH}_2-\text{CH}_3$ ...). Les agents intercalants (proflavine, bromure d'éthidium...) sont des agents qui s'insèrent dans l'ADN entre les paires de bases et peuvent entraîner des insertions ou délétions de paires de bases.

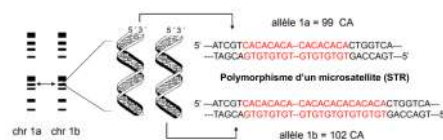
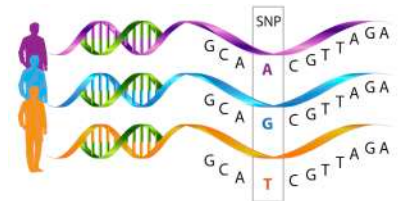
Les divers dommages induits par ces agents (remplacement ou modifications de bases, pontage entre brins de l'ADN, cassures de l'ADN) seront pris en charge par des systèmes de réparation.

#### 4) Différentes conséquences des mutations

##### a) Absence de conséquences

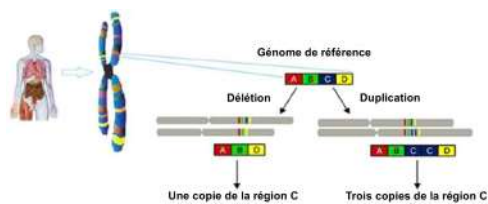
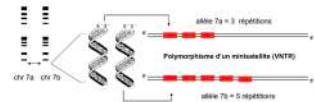
Parmi les différents types de variants nucléotidiques, certains n'ont aucune conséquence. Cette absence de conséquence d'un variant allélique fait parler de polymorphisme (allèle polymorphe) c'est-à-dire existant de façon normale sous différentes formes dans la population générale. On considère généralement qu'un variant est un polymorphisme lorsque  $f_{\text{population}} > 6\%$  et qu'il peut s'agir d'une mutation si  $f < 1\%$ . On va distinguer différents types de polymorphismes selon la structure du variant en question.

Les variations de séquence ponctuelles sont désignées par l'abréviation SNP et sont présentes pour  $f \approx 1/2000$  à 3000 nucléotides. Un polymorphisme peut également être constitué par une variation du nombre de répétitions des séquences répétées en tandem (minisatellites ou microsatellites) ou du nombre de copies d'un gène.



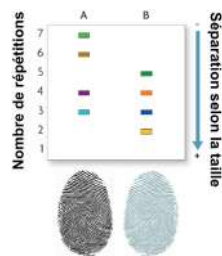
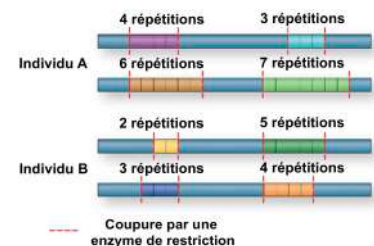
Les microsatellites sont nombreux dans le génome humain (> 50 000) et les variations du nombre de répétitions de leur motif sont appelées STR.

Les minisatellites sont concentrés au niveau des télomères et les variations du nombre de répétitions de leur motif sont appelées VNTR.



Les variations du nombre de copies d'un gène ou d'une région chromosomique sont appelées CNV et ont pour origine des événements de délétion ou de duplication. Les variations de combinaisons des différents polymorphismes entre individus constituent un moyen unique d'identification appelé « empreintes génétiques » et utilisées en génétique médicale ou en médecine légale.

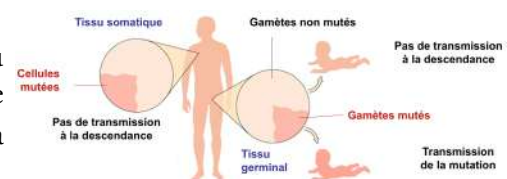
Sur la figure ci-contre sont représentées les variations du nombre de répétitions de séquences répétées en tandem qui existent chez deux individus au niveau de deux sites polymorphes. Des différences du nombre de répétitions de séquences répétées peuvent par exemple être mises en évidence par coupure de l'ADN au niveau de séquences encadrant les sites polymorphes.



Après coupure de l'ADN, les variations du nombre de répétitions se traduiront alors par des différences de longueur des fragments obtenus qu'on pourra ensuite séparer selon leur taille et visualiser. Lorsque des différences de longueur peuvent être mises en évidence par digestion de l'ADN à l'aide d'enzymes de restriction, on parle de polymorphisme de longueur des fragments de restriction (RFLP).

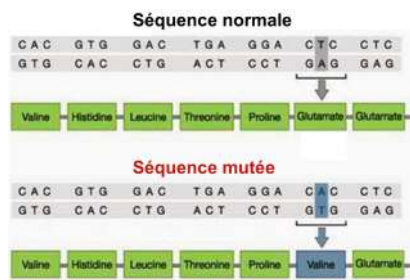
##### b) Transmission de la mutation

Une mutation dite somatique n'affectera que l'individu qui en est porteur, alors qu'une mutation dite germinale sera transmise à la descendance si un  $\gamma$  muté participe à la fécondation.





## c) Mutation délétère ou bénéfique pour l'individu



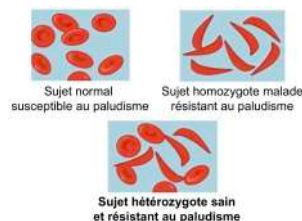
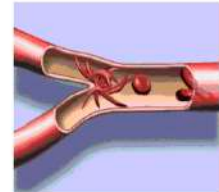
La drépanocytose<sup>†</sup> est la plus fréquente des maladies héréditaires et illustre bien ces variations des conséquences fonctionnelles des mutations. Elle est liée à la présence d'une mutation faux-sens du gène de la  $\beta$ -globine constituant l'hémoglobine et qui remplace un résidu glutamate de la globuline par une valine ( $\beta^6\text{Glu-Val}$ ).

À l'état homozygote, cette mutation est responsable de la polymérisation de l'hémoglobine donnant aux hématies une forme de faucille et entraîne une diminution de leur durée de vie (anémie falciforme). La perte de déformabilité de ces hématies entraîne leur blocage dans les capillaires et est notamment responsable d'accidents occlusifs et d'infarctus pouvant toucher tous les organes.

Hématies normales

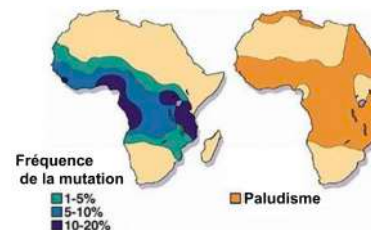


Hématie falciforme



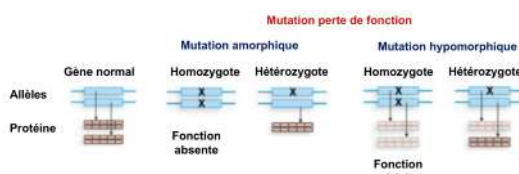
La mutation peut en revanche avoir un effet bénéfique protecteur contre le paludisme chez les sujets qui la portent à l'état hétérozygote ou homozygote. En effet, cette mutation réduit la durée de vie des globules rouges, ce qui perturbe le cycle de reproduction du parasite responsable du paludisme, celui-ci infectant les hématies pour se reproduire.

Au cours de l'évolution, le fait d'être hétérozygote a constitué un avantage sélectif permettant aux individus porteurs de survivre plus facilement dans les zones où le paludisme sévit le plus. Ceci explique pourquoi la mutation est aussi fréquente dans les zones où le paludisme sévit de façon endémique.



## d) Classification des mutations selon leur effet sur la fonction d'un gène &amp; de sa protéine

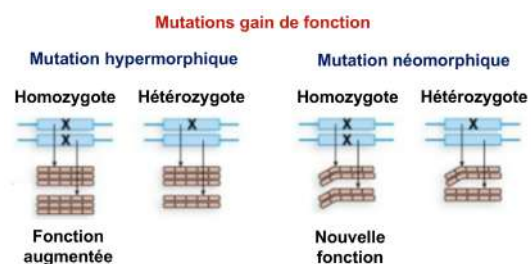
## α) Mutation perte de fonction



Les mutations appelées perte de fonction aboutissent à la formation d'une protéine de fonction absente (mutation amorphe) ou réduite (mutation hypomorphe). Elles correspondent généralement à des mutations récessives, l'allèle restant compensant la perte de fonction sauf s'il est insuffisant ou s'il est absent (haploinsuffisance).

## β) Mutation gain de fonction

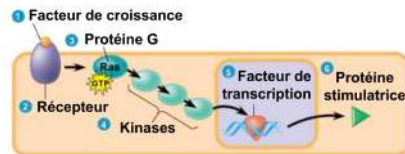
Les mutations dites gain de fonction aboutissent à la formation d'une protéine ayant une fonction augmentée (mutation hypermorphique) ou anormale (mutation néomorphe) et correspondent généralement à des mutations dominantes.



<sup>†</sup> Étymologie : drépanocytose, du grec « δρεπάνη », qui signifie « faux, serpe », et « κύτος », qui signifie « cellule ».

### γ) Le cancer

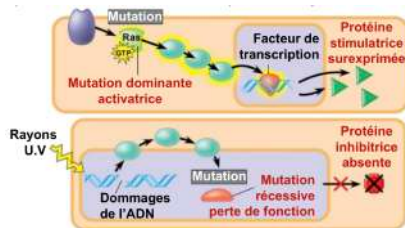
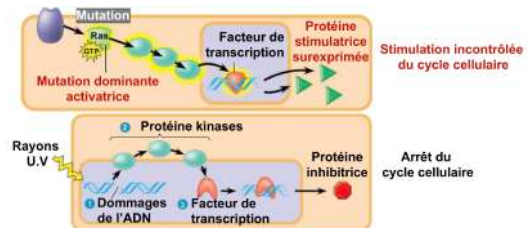
Le cancer est une pathologie liée à la fois à l'accumulation à l'état hétérozygote de mutations dominantes gain de fonction et à l'accumulation à l'état homozygote de mutations récessives perte de fonction. Les mutations gain de fonction vont affecter des gènes codant pour protéines stimulant la prolifération cellulaire, auxquelles elles vont conférer une augmentation de fonction ou une fonction nouvelle. Les mutations perte de fonction vont affecter des gènes codant pour des protéines inhibant la prolifération cellulaire, auxquelles elles vont conférer une fonction diminuée ou absente.



Stimulation régulée du cycle cellulaire

La stimulation normale du cycle cellulaire dépend de la liaison d'un facteur de croissance à son récepteur qui active une cascade de signification et la transcription d'une protéine stimulatrice.

Une mutation dominante activant par exemple la protéine Ras en l'absence de facteur de croissance entraînera l'activation incontrôlée du cycle cellulaire par surexpression de la protéine stimulatrice. De même, en cas de dommages de l'ADN, certains facteurs de transcription comme p53 assurent la production de protéines inhibitrices stoppant le cycle cellulaire pour permettre la réparation des lésions.

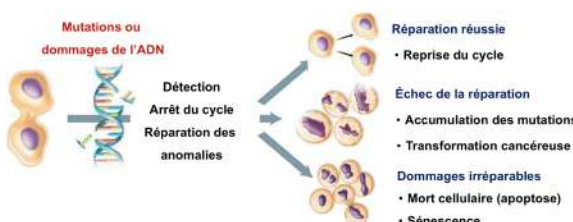


Une mutation récessive perte de fonction à l'état homozygote inactivant p53 empêchera la synthèse des protéines inhibitrices et entraînera la poursuite du cycle cellulaire malgré les dommages de l'ADN.

La cancérogénèse est un processus multi-étapes dépendant de la conjonction de l'activation de proto-oncogènes (version normale d'un oncogène) et de l'inactivation de gènes suppresseurs de tumeurs.

### 5) Systèmes de réparation de l'ADN

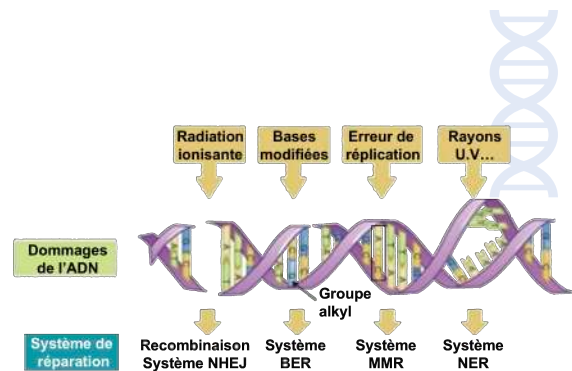
Il existe divers systèmes de maintenance du génome et de réparation de l'ADN. Ces systèmes permettent de minimiser l'apparition et l'accumulation des mutations. En plus des mécanismes de correction immédiate des erreurs de réplication, d'autres mécanismes liés au cycle cellulaire assurent le contrôle de l'intégrité du génome. Cette intégrité dépend de systèmes de détection des mutations ou des dommages de l'ADN qui activent des systèmes d'interruption du cycle cellulaire et des systèmes de réparation des anomalies.



Si la réparation réussit, la cellule reprendra le cycle cellulaire, si elle échoue et si les dommages autorisent la survie, elle accumulera les mutations favorisant sa transformation cancéreuse. Si la réparation échoue, et que les dommages sont incompatibles avec sa survie, elle déclenchera un programme de mort cellulaire par apoptose ou entrera en sénescence.

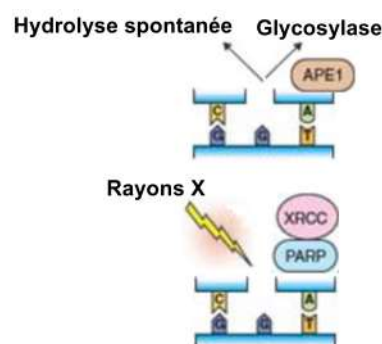
### a) Généralités

Les divers types de dommages sont pris en charge par un système de réparation spécifique. Le système de réparation par excision de base (BER) prend en charge les anomalies ne modifiant pas la structure de l'ADN comme les bases modifiées (désamination, alkylation...). Le système de réparation de mésappariements liés aux mutations (MMR) prend en charge notamment les mutations induites par les erreurs de réplication.



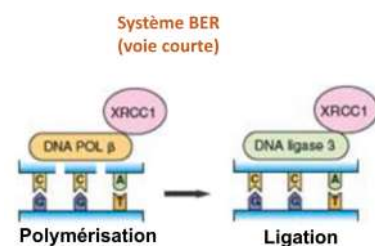
Le système de réparation par excision de nucléotide (NER) prend en charge les pontages entre brins qui modifient la structure de l'ADN (exemple : dimères de thymine). Les cassures double brin de l'ADN font intervenir soit la recombinaison homologue, soit le système de réparation non homologué par ligation des extrémités (NHEJ). Les systèmes BER, MMR et NER ont en commun de réparer les lésions de l'ADN en agissant sur un seul brin, chacun utilisant des mécanismes différents pour réparer des anomalies différentes.

### b) Système BER

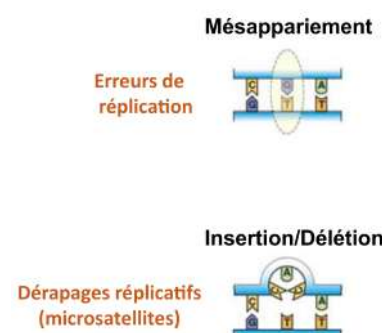


Le système BER permet de restaurer les sites abasiques créés par l'hydrolyse spontanée de bases ou de réparer les cassures simple brin de l'ADN créées par les rayons X. Le système BER peut aussi créer lui-même un site abasique pour supprimer une base anormale formée par exemple par désamination, alkylation, dépurination ou oxydation. Dans ce cas, il utilise au préalable une *glycosylase* spécifique de cette base modifiée qui la bascule hors de l'hélice et la supprime, formant le site abasique. Le site abasique créé par hydrolyse spontanée ou la *glycosylase* est reconnu par une *endonucléase* (APE1) qui supprime le sucre, formant elle-même une cassure simple brin.

Les cassures simple brin créées par les rayons X seront détectées par d'autres protéines comme XRCC et PARP. Quelles que soient la lésion initiale et la voie impliquées, la réparation finale du brin fait intervenir une *ADN polymérase* et une *ligase* pour réinsertion et ligation du/des nucléotides manquants.



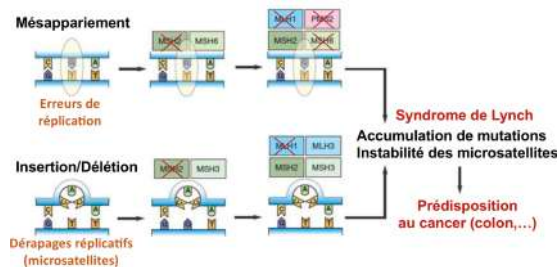
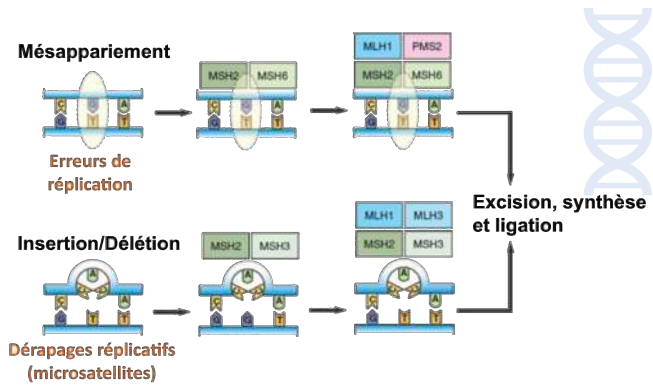
### c) Système MMR



Le système MMR prend en charge les mésappariements formés par les erreurs de réplication et les petites insertions ou délétions de nucléotides qui surviennent notamment au niveau des microsatellites. Ce système a été initialement décrit chez E.Coli où il est constitué de protéines MutS, MutL et MutH qui fonctionnent sous la forme de dimères. Chez les eucaryotes, le système s'est diversifié et spécialisé et comprend des homologues de MutS (MSH2, MSH3, MSH6), de MutL (MLH1, MLH2, MLH3, PMS1, PMS3) mais pas d'homologues de MutH.



L'hétérodimère MSH2-MSH6 (MutS $\alpha$ ) reconnaît essentiellement les substitutions et l'hétérodimère MSH2-MSH3 (MutS $\beta$ ) reconnaît les insertions et les délétions. Le dimère MLH1-PSM2 (MutL $\alpha$ ) est recruté sur les substitutions, tandis que le dimère MLH1-MLH3 (MutL $\gamma$ ) est recruté sur les insertions ou délétions. Un fragment d' $\approx 20$  nucléotides contenant l'erreur est excisé puis resynthétisé par une *ADN polymérase* et une ligase restaure la continuité du brin lésé.

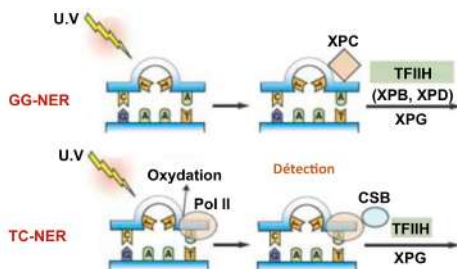
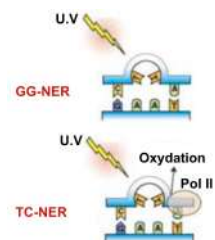


L'inactivation de constituants du système MMR est responsable d'une prédisposition héréditaire au cancer appelé syndrome de Lynch ou encore HNPCC. Ce syndrome est responsable d'un défaut de réparation des erreurs de réplication et d'une instabilité des microsatellites et prédispose à l'apparition de diverses formes de cancer (côlon).

L'augmentation du taux de mutations liée à l'inactivation du système facilite l'apparition de mutations activant des proto-oncogènes ou inactivant des suppresseurs de tumeur, faisant ainsi le lit du cancer.

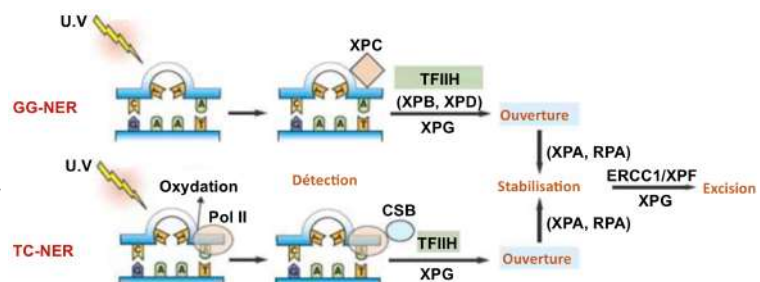
#### d) Système NER

Le système NER assure la réparation des lésions entraînant une distorsion de la double hélice induite par les rayons UVB (dimères de thymine) ou d'autres agents mutagènes. Il comprend une voie active en permanence (GG-NER) et l'autre activée spécifiquement par des lésions bloquant la transcription (TC-NER). Les deux voies comprennent quatre étapes, comprenant la détection de l'anomalie, l'ouverture de la double hélice autour de la lésion, l'incision de l'ADN de part et d'autre, et la resynthèse de l'ADN et la ligation.



Elles débutent respectivement par la reconnaissance de la lésion par la protéine appelée XPC ou par l'*ARN polymérase* et la protéine CSB. Elles font ensuite intervenir le complexe formant le facteur de transcription TFIIH dont les sous-unités XPB et XPD possèdent une activité *hélicase*, ainsi que l'*endonucléase* XPG.

Une fois l'hélice ouverte et les brins stabilisés par les protéines XPA et RPA, un court fragment contenant l'anomalie est excisé de part et d'autre par les *nucléases* ERCC1/XPF et XPG. La réparation s'achève par la resynthèse d'un fragment d'ADN et sa ligation.





## Point patho : pathologies du système NER



La maladie appelée Xeroderma Pigmentosum liée à l'inactivation de la voie globale de ce système est caractérisée par une hypersensibilité aux rayons UV favorisant l'apparition précoce de cancers cutanés.

La maladie appelée syndrome de Cockayne est liée à l'inactivation de la voie du système liée à la transcription et entraîne généralement un décès précoce des enfants atteints.

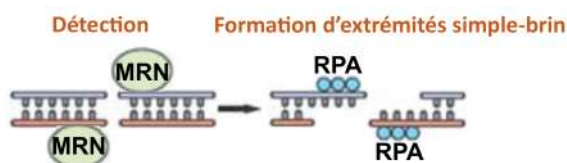
## e) Cassures double brin de l'ADN

Les cassures double brin de l'ADN représentent un danger sérieux pour l'intégrité du génome dont la réaction cellulaire est d'activer un point de contrôle (*check-point*) du cycle cellulaire.

## Cassure double-brin

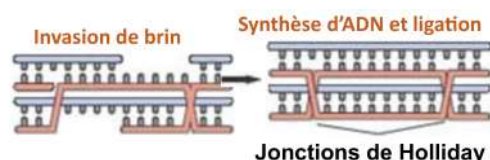
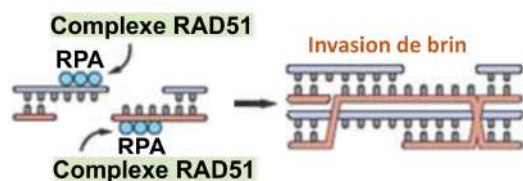


En effet, elles peuvent être responsables d'anomalies cytogénétiques majeures comme des translocations chromosomiques, des amplifications ou des délétions chromosomiques. Les cassures double brin de l'ADN peuvent être induites par les radiations ionisantes, les agents oxydants ou survenir au niveau d'une fourche de réplication bloquée. Elles sont détectées par une cascade de protéines aboutissant à l'activation de la recombinaison homologue ou du système de ligation non homologue des extrémités chromosomiques (NHEJ). L'avantage et le principe de la recombinaison homologue va être d'utiliser en mitose l'hélice d'ADN de la chromatide sœur comme matrice pour la reconstruction du segment d'ADN endommagé.



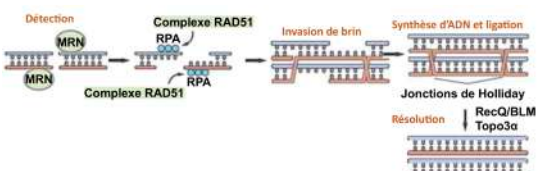
Elle débute par la reconnaissance de la lésion par le complexe MRN, formé des protéines Mre11, Rad50 et NBS. Des extrémités simple brin 3'-sortantes vont ensuite être créées grâce à une activité 5'-3' exonucléasique et seront recouvertes par les protéines RPA.

Les protéines BRCA2 et RAD52 entraînent la formation d'un complexe comprenant RAD51 et qui remplace les protéines RPA sur les fragments d'ADN simple brin. Le complexe RAD51 initie et guide un processus de recherche d'homologie entre brins et un brin de chaque duplex envahit l'autre duplex pour d'apparier avec son brin complémentaire.



Une *polymérase* assure la synthèse d'ADN par complémentarité et permet de restaurer les brins lésés à partir de leur extrémité 3' puis une *ligase* rejoint leurs extrémités. Au cours du processus, deux structures appelées jonctions de Holliday se sont formées au croisement des brins ayant envahi la chromatide homologue.

L'étape finale appelée « *résolution des jonctions* » permet de dénouer ces intersections grâce à des *hélicases* comme *RecQ* et *BLM*, en conjonction avec la *topoisomérase 3α*.





### Point patho :

L'importance de la réplication par recombinaison homologue est illustrée par l'existence de syndromes de prédisposition au cancer ou au vieillissement liés à l'inactivation de protéines de ce système.

Syndrome de Bloom



Syndrome de Werner

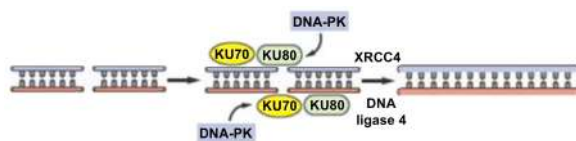
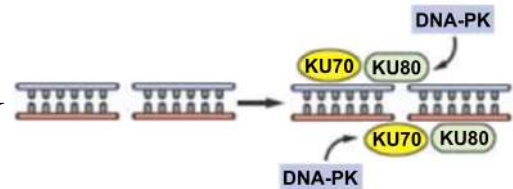


Le plus connu de ces syndromes est celui lié à l'inactivation des protéines BRCA1 & BRCA2, qui se traduit par une prédisposition héréditaire au développement précoce de cancer du sein et de l'ovaire. Enfin, l'inactivation des hélicases RecQ4 ou BLM se traduit par divers types de cancers, & celle de l'hélicase WRN entraîne un vieillissement accéléré & des pathologies liées à l'âge.

### f) Système NHEJ

Le système NHEJ est un système de réparation non fidèle qui consiste simplement à rejoindre bout à bout les fragments formés par la cassure sans tenir compte des pertes de matériel génétique.

Les extrémités double brin sont reconnues par l'hétérodimère formé des protéines Ku70 et Ku80. Puis, la protéine kinase dépendante de l'ADN appelée DNA-PK est recrutée par ces hétérodimères.



Les extrémités de l'ADN sont enfin directement reliées sans synthèse d'ADN après recrutement de la protéine XRCC4 et de l'ADN ligase 4.

### Récapitulatif

Le type, la source et les conséquences des mutations sont variables. Certaines sont spontanées, liées aux erreurs de réplication, aux séquences répétées du génome, à la tautomérie ou aux modifications spontanées des bases, et d'autres sont induites par des mutagènes. Les mutations peuvent être neutres, bénéfiques ou délétères, transmissibles ou non, et être responsables d'une perte ou d'un gain de fonction d'une protéine. Selon leur type et leur source, les mutations sont réparées par différents systèmes. Les systèmes BER, MMR et NER assurent respectivement la réparation des modifications spontanées de bases, des mésappariements ou insertions/délétions, et de lésions induisant une distorsion de l'ADN. La recombinaison homologue et le système de ligation des extrémités assurent respectivement de façon fidèle ou incomplète la réparation des cassures double brin de l'ADN. Chez l'homme, l'intégrité du génome est compromise dans différents syndromes liés à l'inactivation de protéines de l'un des systèmes de réparation des mutations et dommages de l'ADN.

## B) Dynamiques du génome & évolution

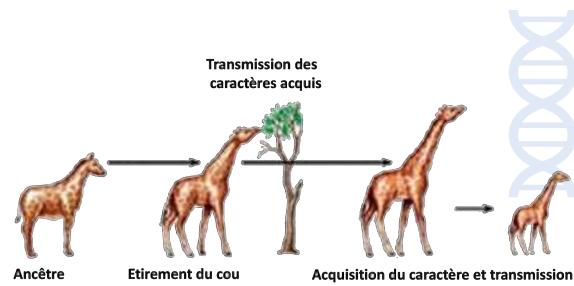
### 1) Théories de l'évolution

#### a) Lamarckisme

Le lamarckisme repose sur la transmission des caractères acquis. Pour Lamarck, l'usage intensif ou délaissé d'un organe chez un animal en développement modifierait cet organe, modification qui pourrait dans certains cas être transmise à la descendance.

## Biologie moléculaire

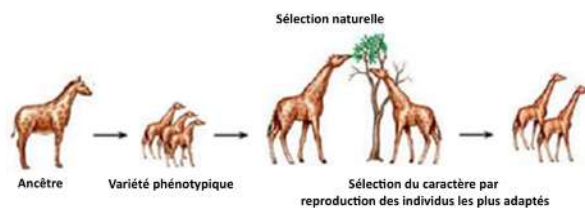
Ainsi, les girafes allongeraient leur cou en faisant systématiquement l'exercice de chercher à atteindre les branchages hauts, procréant ainsi progressivement des descendants au cou de plus en plus long. Selon cette théorie, l'évolution résulte de la transmission des caractères acquis au cours de la vie, l'usage ou le non usage d'un organe déterminant son développement ou sa disparition.



Cette théorie s'avère aujourd'hui partiellement valide, dans la mesure où les modifications épigénétiques peuvent être conservées au cours des générations cellulaires et sont transmissibles. L'épigénétique est un mécanisme par lequel certains caractères acquis par l'influence de l'environnement ou les habitudes de vie sont inscrites dans le génome puis transmises à la descendance.

## b) Darwinisme

Le darwinisme repose sur la théorie de la sélection naturelle. Pour Darwin, l'évolution repose sur une variabilité du caractère au sein de l'espèce, la sélection naturelle entraînant la conservation et la transmission du caractère le plus favorable à la survie.

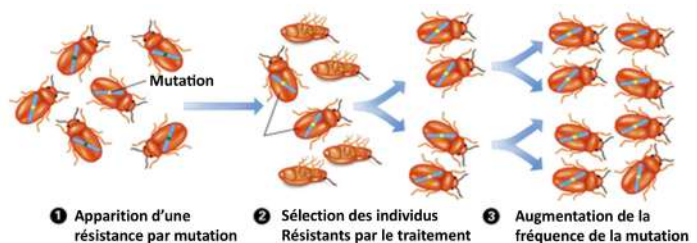


Ainsi, les girafes ayant un cou plus long auraient plus de descendants, étant capables en cas de disette d'atteindre plus facilement les feuillages des branches les plus hautes. La validité de cette théorie est liée à l'existence des mutations qui sont à la base de changements de caractères et constituent la source des variations phénotypiques au sein d'une population.

## 2) Les mutations : moteurs de l'évolution

Les mutations forment le mécanisme grâce auquel la sélection naturelle opère. La transmission de mutations non létales enrichit le pool génique d'une population et sa diversité, la sélection naturelle agissant ensuite en réduisant l'abondance des mutations défavorables en termes d'adaptations aux changements évolutifs, et en augmentant celle des mutations plus favorables.

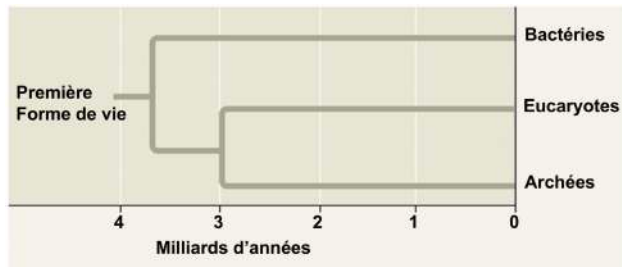
La résistance d'insectes aux pesticides ou de bactéries aux antibiotiques est en effet liée à l'apparition aléatoire d'une mutation conférant cette résistance, un traitement ultérieur entraînant ensuite la sélection des individus mutants résistants et la conservation de la mutation.



Ainsi, l'apparition de mutations liées à la mutabilité et la dynamique du génome permettent d'expliquer la diversité des espèces, et la sélection des individus permet ensuite d'expliquer leur évolution.

### 3) Analyse comparative de génomes

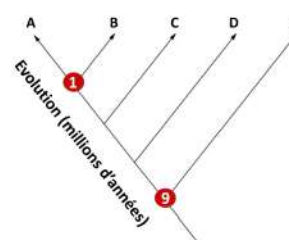
L'analyse comparative de génomes témoigne de leur dynamique. Cette comparaison fournit des preuves du rôle des mutations dans l'évolution. Les relations de proximité entre espèces en termes évolutifs sont représentés par un diagramme en forme d'arbre appelé arbre phylogénétique†.



Les 3 domaines des espèces vivantes (bactéries, archées et eucaryotes) seraient issus de l'évolution d'un ancêtre commun qui se serait produite il y a plusieurs milliards d'années. Ainsi, la comparaison entre génomes des procaryotes et des eucaryotes a pu mettre en évidence des différences expliquant leurs divergences anciennes.

Le développement des techniques de séquençage a permis d'obtenir la séquence complète du génome d'organismes procaryotes, eucaryotes unicellulaires et multicellulaires dont l'homme. Parallèlement, l'utilisation de la bio-informatique a permis d'analyser et de comparer entre eux ces différents génomes à la recherche de preuves de l'évolution et de mécanismes permettant de l'expliquer. Ainsi, la comparaison de la séquence d'un gène donné entre espèces phylogénétiquement proches et distantes a permis de montrer que leur évolution progressive est associée à l'accumulation des mutations. Les divergences de séquences sont d'autant plus grandes que les espèces sont éloignées et constituent une preuve moléculaire du rôle des mutations sur l'évolution (horloge moléculaire).

Sur ce concept, la connaissance du nombre de différences nucléotidiques dans un gène entre deux espèces permet d'en déduire le temps depuis lequel elles ont divergé. En définitive, l'évolution des eucaryotes serait liée à l'existence d'une dynamique particulière de leur génome favorisant les mutations et ayant permis leurs divergences évolutives.



Comparaison de séquence entre les espèces A à E

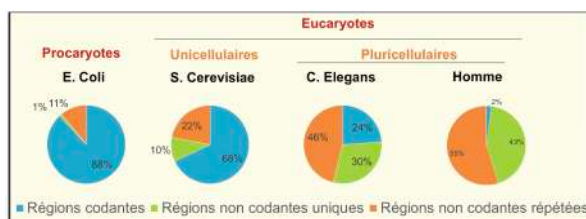
A) ...atccgattattgcacgatat...  
 B) ...atccgattttgcacgatat...  
 C) ...atccattttgcacgatat...  
 D) ...ttccraattttgcacgatat...  
 E) ...ttccraattttgcacgatat...

Une substitution par million d'années

### 4) Le contenu du génome eucaryote

#### a) Différences entre les génomes eucaryote & procaryote

Les génomes procaryote et eucaryote diffèrent dans la nature de leurs séquences. Ces séquences peuvent être séparées en séquences codantes permettant la synthèse des protéines, en séquences non codantes uniques et en séquences non codantes répétées.



La différence majeure entre les organismes les plus simples et les plus complexes ne réside pas dans le nombre de gènes de leur génome respectif, mais dans la nature des séquences qui les constituent. De façon paradoxale, il s'avère que plus un organisme est complexe, moins son génome contient de séquences codantes, et plus il contient de séquences non codantes.

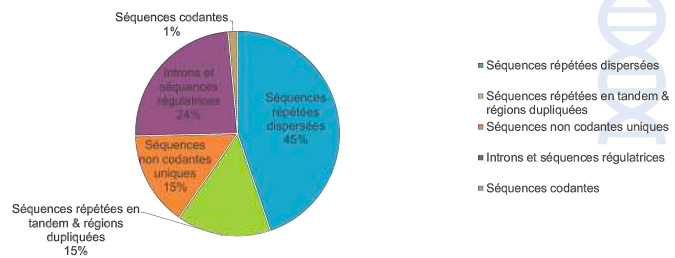
† Étymologie : phylogénétique, du grec « φυλον », qui signifie « tribu, race », et « génétique ».



## Biologie moléculaire

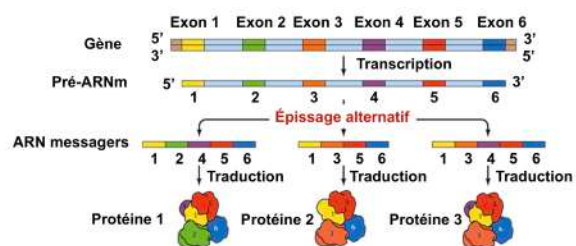
## b) Richesse du génome eucaryote en séquences non codantes uniques

Une part importante de ces séquences est représentée par les introns et par les séquences régulatrices qui modulent la transcription individuelle de chaque gène de façon spécifique.

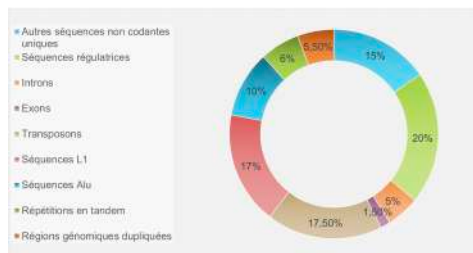


Les gènes procaryotes sont regroupés en opérons et dénués d'introns et les gènes eucaryotes sont régulés individuellement et leur séquence codante est morcelée par la présence des introns. L'existence de séquences régulatrices propres à chaque gène assure une finesse de la régulation de leur expression et un niveau de complexification supérieur aux fonctions eucaryotes.

La présence d'introns est à l'origine du phénomène d'épissage alternatif qui permet de produire plusieurs protéines à partir d'un seul gène et ainsi de diversifier le répertoire protéique des eucaryotes. Le nombre de protéines d'un organisme reflète davantage sa complexité que le nombre de gènes de son génome et chez l'homme, 20 000 à 30 000 gènes permettent la synthèse de 200 000 protéines différentes.



## c) Richesse du génome eucaryote en séquences non codantes répétées

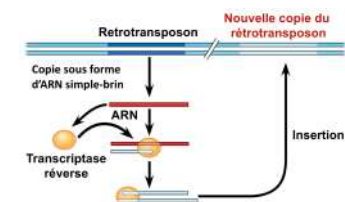
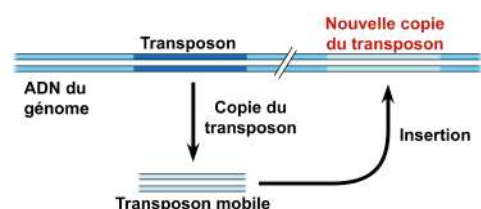


La majeure partie de ces séquences est représentée par des séquences répétées en tandem (minisatellites et microsatellites) et de larges régions du génome qui ont été dupliquées et qui forment aujourd'hui des familles de gènes.

## d) Éléments transposables

Les éléments transposables sont des moteurs de l'évolution. Ces éléments mobiles, aussi appelés gènes sauteurs, sont capables de se multiplier et de se déplacer dans le génome, et sont présents chez les eucaryotes, et à un degré moindre chez les procaryotes.

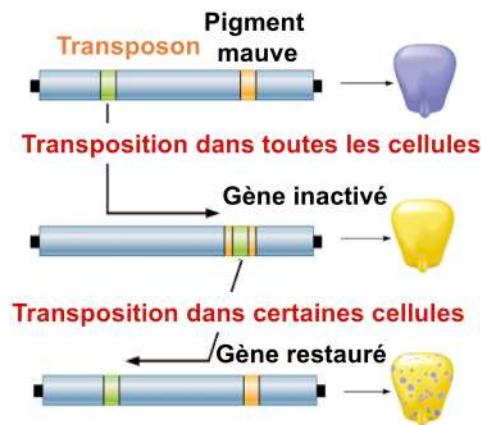
Le mécanisme de multiplication et de déplacement d'un transposon implique une étape de copie de la séquence sous forme d'ADN, suivie de l'insertion de cette copie à un autre endroit du génome.



Le mécanisme de multiplication et de déplacement d'un rétrotransposon implique la copie de sa séquence sous forme d'ARN, puis sa rétrotranscription en ADN ensuite inséré ailleurs dans le génome.

## Biologie moléculaire

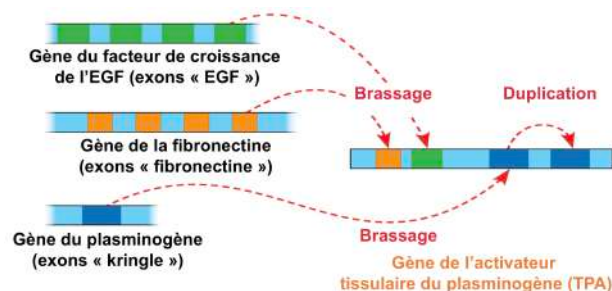
Ces éléments mobiles ont été découverts en 1947 par Barbara McClintock qui étudiait les modifications du maïs au cours de ses expériences de croisements. Elle découvrit d'abord l'existence de régions d'ADN capables de se déplacer, puis s'en servit pour expliquer l'apparition de variation de couleurs des grains de maïs.



L'absence de pigmentation d'un grain était liée à l'inactivation du gène codant pour la couleur par l'insertion du transposon dans sa séquence. Dans certaines cellules, ce gène pouvait ensuite être réactivé par un nouveau déplacement du transposon hors du gène de pigmentation. À une époque où on considérait le génome comme une entité statique, faite de gènes alignés le long des K, elle fut la première à mettre en évidence la dynamique du génome et les éléments responsables. En effet, les transposons peuvent, au cours de leur déplacement, inactiver des gènes ou entraîner des exons de gènes, voire des gènes entiers avec eux et augmenter la taille du génome en se multipliant.

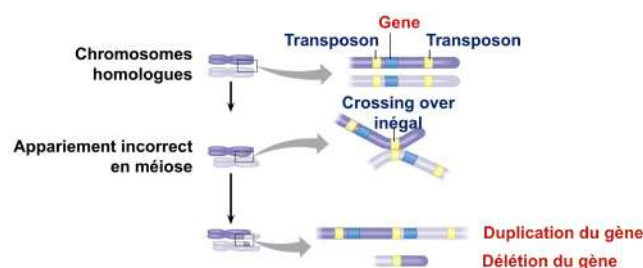
### α) Brassage d'exons

Les transposons peuvent inactiver un gène ou, en s'insérant dans ses séquences régulatrices, en augmenter ou en réduire l'expression. Au cours de leurs déplacements, des transposons encadrant par exemple un exon, peuvent l'entraîner et aller d'insérer dans la séquence codante d'un autre gène, créant ainsi un nouvel assortiment d'exons.



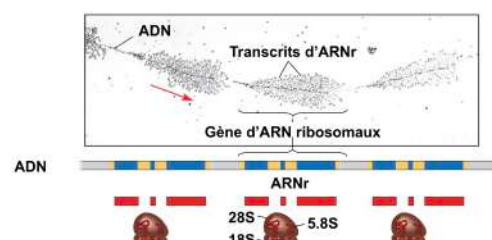
Ce phénomène appelé « *brassage d'exons* » est à l'origine de la création de gènes au cours de l'évolution par mixage d'exons issus de différents gènes pour en créer un nouveau.

### β) Crossing-over inégaux

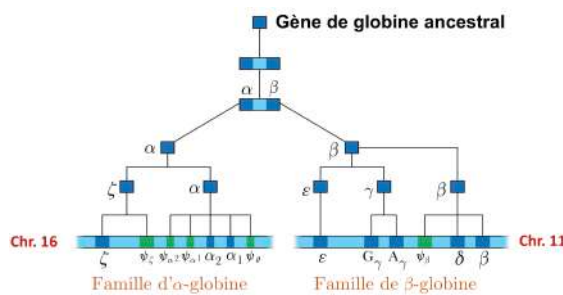


Du fait de leur existence en de multiples copies identiques dispersées dans le génome, ils peuvent au cours de la méiose favoriser la recombinaison non homologe ou des *crossing-over* inégaux. Ces *crossing-over* inégaux entre K impliquant les transposons peuvent être responsables de délétions et de duplications d'exons, de gènes, ou de régions entières du génome.

Ces événements sont à l'origine de l'existence de certains gènes en plus de deux copies dans le génome et qui peuvent être répétés en tandem de nombreuses fois comme les gènes codant les ARNr. Ils contribuent ainsi à la dynamique du génome et à l'évolution par ce type de mécanisme.

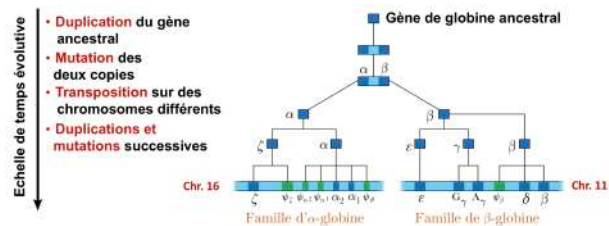


### γ) Familles multigéniques



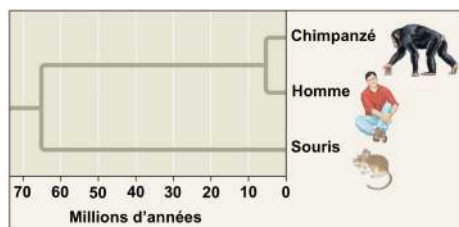
Les gènes identiques ou très similaires présents en de nombreuses copies dans le génome forment des familles dites multigéniques. Les gènes de globine appartiennent à une famille multigénique constituée de gènes similaires répartis sur deux K différents et qui forment les sous-familles des gènes d'α-globine et de β-globine.

Ces différents gènes sont issus d'un gène ancestral unique dont l'évolution par duplications, mutations et transposition sur deux K différents forme les familles actuelles de gènes de globine. Au cours de l'évolution, certains gènes dupliqués et initialement actifs ont été inactivés par mutation et sont ainsi appelés des pseudogènes ( $\psi$ ).



Les gènes actuellement actifs s'expriment à des périodes du développement différentes et les chaînes d'α-globine et de β-globine présentes à ces périodes s'associent pour former des hémoglobines différentes.

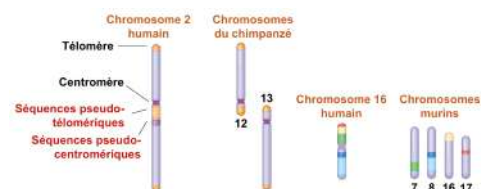
### 5) Preuves du rôle des mutations dans l'évolution



Ainsi, la comparaison entre génomes procaryote et eucaryote a permis de comprendre le rôle joué par la dynamique du génome eucaryote et notamment les transposons dans l'évolution. De même, la comparaison entre le génome d'espèces plus proches comme l'homme, le singe ou la souris a mis en évidence des événements génétiques plus récents à l'origine de leur divergence.

Ainsi, la différence du nombre de K entre l'homme ( $n = 46$ ) et ses proches parents comme le chimpanzé ( $n = 48$ ) témoigne de réarrangements qui seraient contemporains de leur divergence.

La mise en évidence sur le K 2 humain de séquences pseudo-télomériques à proximité du centromère est en faveur de la formation de ce K par la fusion de deux autres K. Les K 2 et 16 humains sont en effet issus respectivement de la fusion des K 12 et 13 du singe, et de réarrangements entre les K 7, 8, 16 et 17 de la souris.



### Récapitulatif

La dynamique du génome permet d'expliquer l'évolution et la diversification des espèces. La comparaison entre génomes d'espèces différentes fournit des preuves moléculaires de leur histoire évolutive et des mécanismes expliquant leurs évolutions respectives. Les génomes eucaryotes se distinguent par l'organisation de leurs gènes, contenant des introns, source de diversité protéique, et leurs séquences assurant une régulation fine de l'expression de leurs gènes. L'abondance de ces génomes en éléments transposables dynamiques est à l'origine de leur évolution et de la création de gènes par des mécanismes de transposition, de brassage d'exons, de délétions ou de duplications, de mutations et de réarrangements chromosomiques.