

Professeur : Claude Agnès Reynaud
Ronéo typeur : Maëlle Doveze
Ronéo lecteur : Amélie Le Duc

PLAN

- I. Historique
- II. Organisation des gènes
- III. Mécanismes responsables de la formation du répertoire
 - 1) Mécanismes de réarrangement
 - a. Réarrangement V (D) J
 - b. Les facteurs du réarrangement
 - c. Les étapes du réarrangement
 - d. Phénotype de l'inactivation des différents composants du mécanisme de réarrangement
 - 2) La réponse immune
 - Mutation somatique
 - Commutation isotypique (Switch)

I. Historique

La question de génération de la diversité des anticorps a été un véritable débat scientifique pendant des décennies.

→ **Quel est le mécanisme cellulaire qui peut expliquer que les Ac ont une reconnaissance quasiment illimitée (à 1 aa près) ?**

Les Ac peuvent aussi reconnaître des formes lévogyres et dextrogyres des aa.

Landsteiner a modifié la façon de percevoir la question dans les années 20-30 : il montre que les Ac sont capables de reconnaître des structures chimiques de synthèse. On est donc face à un système biologique qui est capable de reconnaître quelque chose qu'il n'a pas anticipé.

L'histoire des discussions scientifiques autour des Ig a suivi de très près l'évolution des idées dans différents domaines, ainsi que l'évolution des techniques.

- Dans les années 1950, les premières idées étaient que les Ac étaient capables de se modeler sur l'antigène et le reconnaître = **conception adaptative** (l'Ac s'adapte à l'Ag).
- Avec les débuts de la biologie moléculaire moderne (observation des premières mutations spontanées), on passe à une notion de **théorie sélective** : l'Ag sélectionne l'Ac le plus approprié parmi la diversité
 - ⇒ C'est la première fois qu'on évoque le fait que la diversité préexiste.

- Enfin on arrive à la **théorie de la sélection clonale** (Burnett 1967) : la diversité préexiste mais l'Ig est à la surface des lymphocytes B, et la cellule a la capacité de se multiplier après reconnaissance ; il s'en suit une diversification spontanée dont les mécanismes ne sont pas encore connus. Les lymphocytes vont donc présenter toute une diversité qui va être sélectionnée par l'Ag.
- Découverte de l'ADN : la question de la diversité des Ac va évoluer.
→ **Combien faut il de gènes pour coder ce répertoire quasiment infini des Ac ?**

Deux théories s'affrontent :

- **Les théories germinales** : toutes les spécificités sont dans le génome
- **Les théories somatiques** : il existe peu de gènes mais il se produit des choses au cours du développement du système lymphoïde qui va créer toute cette capacité de reconnaissance. Il y a différenciation par des mécanismes somatiques. Dans cette optique, on a :
 - Brenner et Milstein proposent un système de mutations au hasard (1966)
 - Il va y avoir proposition de mécanismes de recombinaison :
 - ↳ Kabat (1970) : à partir des séquences protéiques des Ig il a proposé le système du minigène hypothesis = le CDR serait greffé sur une structure de base.
 - ↳ somatic recombinaison (1967) = échange d'ADN entre les gènes qui ne se restreint pas au CDR.

Finalement, un petit peu de toutes ces hypothèses explique comment on fait le répertoire des Ig.

Parallèlement à ces 2 théories, il y a aussi eu une observation par Dreyer et Bennett : quand on séquence une Ig, on s'aperçoit qu'il y a 2 parties : une partie constante et une partie variable.

⇒ **Il faut 2 gènes différents pour faire un polypeptide** (concept révolutionnaire pour l'époque).

Le débat se termine avec l'explosion du génie génétique et une des premières séquences clonées a été la chaîne légère des Ig de souris.

Tonegawa analyse la configuration des gènes de la chaîne légère dans des myélomes et fait la comparaison avec des tissus non lymphoïdes. Il utilise le southern blot et la microscopie électronique.

- ↳ Il montre qu'il existe une combinatoire extrêmement importante fournie par un mécanisme unique qui est la **recombinaison au cours de la différenciation** d'une lignée cellulaire particulière.

II. Organisation des gènes

Grâce au séquençage du génome, on peut désormais compter les gènes.

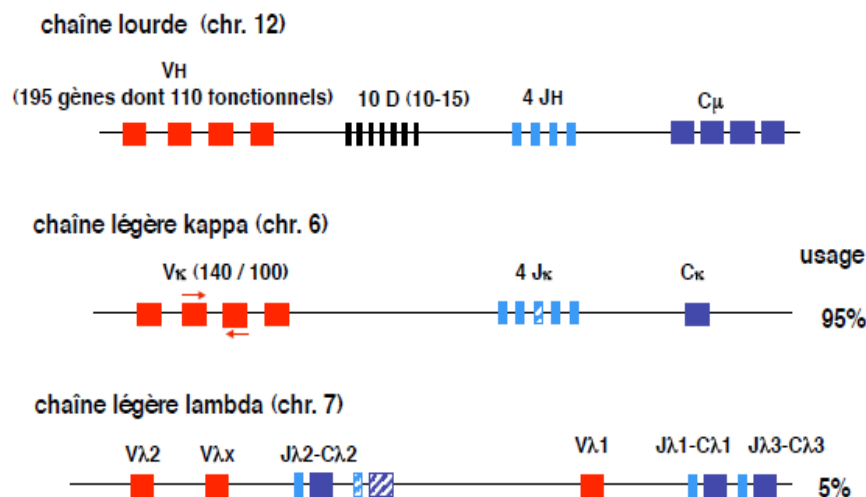
➤ Chez la souris

On se retrouve chez la souris avec 3 loci :

- ▲ un **locus pour la chaîne lourde** sur le chromosome 12 où on note 110 gènes fonctionnels, avec des éléments V, D, J et des parties constantes (ex : C μ , C α ...).
- ▲ 2 loci pour la chaîne légère :
 - un **locus kappa** sur le chromosome 6 (100 gènes fonctionnels) avec un bloc d'éléments J et une absence d'élément D, mais un répertoire V lambda assez important. Les gènes ont une orientation différente et en termes de réarrangement cela impose des résultats très différents :
 - ↪ Quand on est dans le même sens de transcription, le réarrangement sera une simple délétion.
 - ↪ Quand on est en polarité opposée, ce sera une inversion.
 - un **locus lambda** sur le chromosome 7 (3 gènes fonctionnels) avec une alternance d'éléments JC/JC/JC au niveau de la région constante. La découverte initiale de Tonegawa portait sur ce locus.

Il y a toujours une partie importante de gènes non fonctionnels, on a donc une coexistence entre des gènes fonctionnels et des gènes non fonctionnels.

La chaîne kappa est utilisée à 95% alors que la chaîne lambda est utilisée à 5%, on a donc un usage du répertoire qui est en gros proportionnel au nombre de gènes fonctionnels.

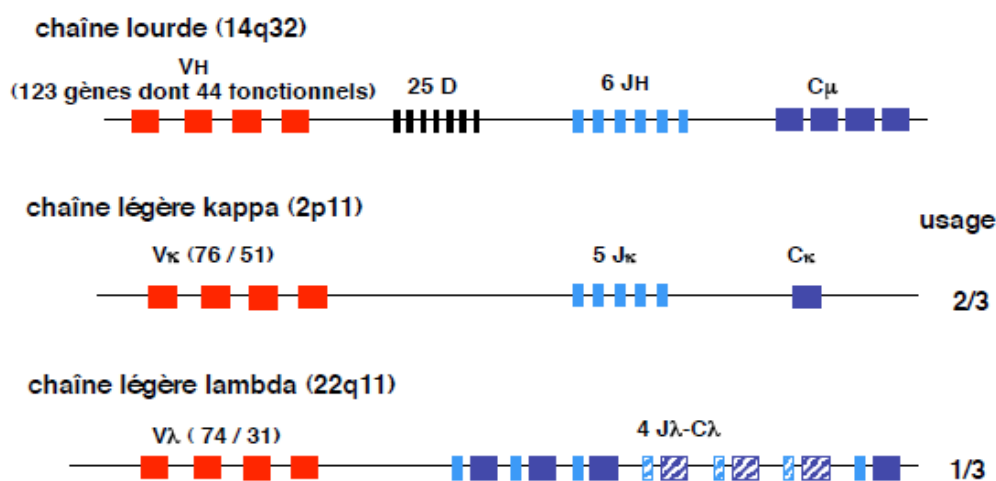


➤ Chez l'homme

Chez l'homme le répertoire fonctionnel est encore plus petit :

- ⤴ la **chaîne lourde** comprend entre 44 et 50 gènes fonctionnels avec des éléments V, D et J.
- ⤴ une **chaîne légère kappa** utilisée dans 1/3 du répertoire.
- ⤴ une **chaîne légère lambda** avec une alternance d'éléments JC/JC/JC comme chez la souris, utilisée dans les 2/3 du répertoire.

⇒ On est loin du millier de gènes proposé dans la théorie germinale.



Les gènes variables

3 caractéristiques :

- ⤴ On les classe en familles, d'après leur homologie de séquence.

Dans l'ancienne classification, Kabat les avait classés en sous groupes.

Par exemple, on a **7 familles VH** chez l'homme (= éléments V des gènes des chaînes lourdes d'Ig), mais le nombre de gènes fonctionnels dans une famille est extrêmement variable. Certaines familles ont un seul gène fonctionnel et certaines familles comme VH3 représentent à elles toutes seules la moitié du répertoire des gènes des Ig.

Ensuite on classe ces familles en groupe ou en clan. On a **3 groupes** :

- I : VH1 + VH5
- II : VH2 + VH4 + VH6
- III : VH3

Dans une famille on a des gènes fonctionnels mais aussi des gènes non fonctionnels.

▲ Existence de nombreux **pseudogènes**

On a 79 pseudogènes sur 123 VH chez l'homme.

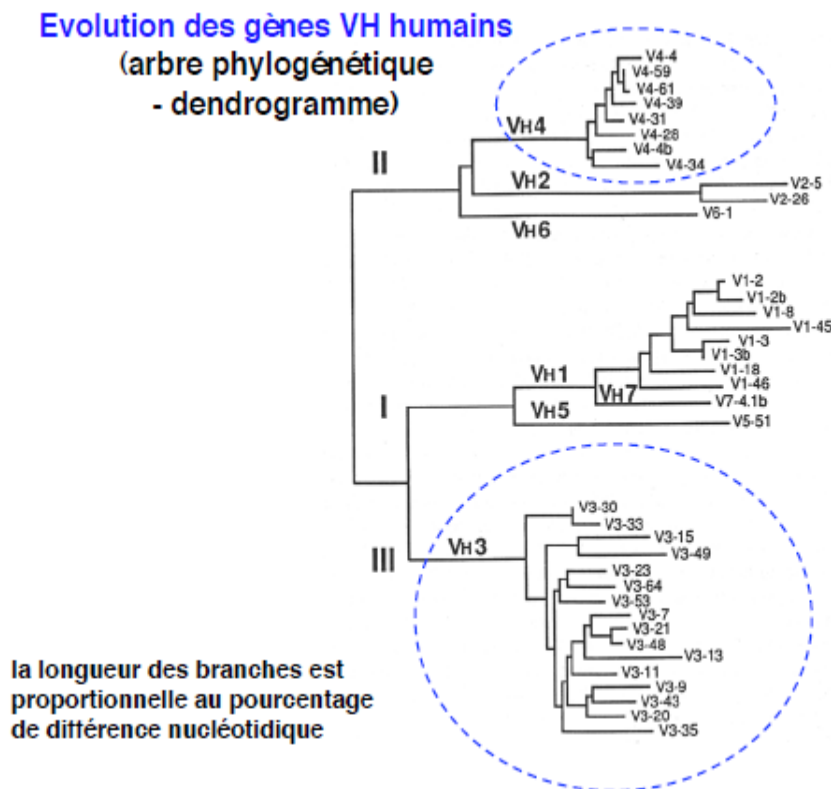
Il y a toute une discussion sur l'évolution des gènes : est-ce que les pseudogènes sont un réservoir pour des échanges de séquences mitotiques, qui pourraient faire qu'une partie de ces pseudogènes, par des événements d'échanges géniques, puissent réintégrer le groupe fonctionnel ?

▲ **Taux de réarrangement variable**

Même avec 44 gènes, ils ne sont pas tous égaux, c'est-à-dire que certains sont beaucoup plus utilisés que d'autres. Certains sont donc beaucoup plus présents dans le répertoire pré immun.

Par exemple, le gène 3-23 est particulièrement utilisé.

Arbre phylogénétique ou dendrogramme



Il sert à montrer la parenté des gènes.

On peut le faire soit sur l'évolution des gènes entre les espèces, soit dans les familles multigéniques pour montrer quelle est leur parenté.

Les branches relient chaque gène entre eux et la longueur des branches est proportionnelle à la divergence nucléotidique entre les différents gènes.

On a 3 super familles : les **clans** I, II et III.

Des sous groupes, c'est-à-dire **les familles** se sont clairement formées à l'intérieur de ces clans : VH1, VH2, ..., VH6.

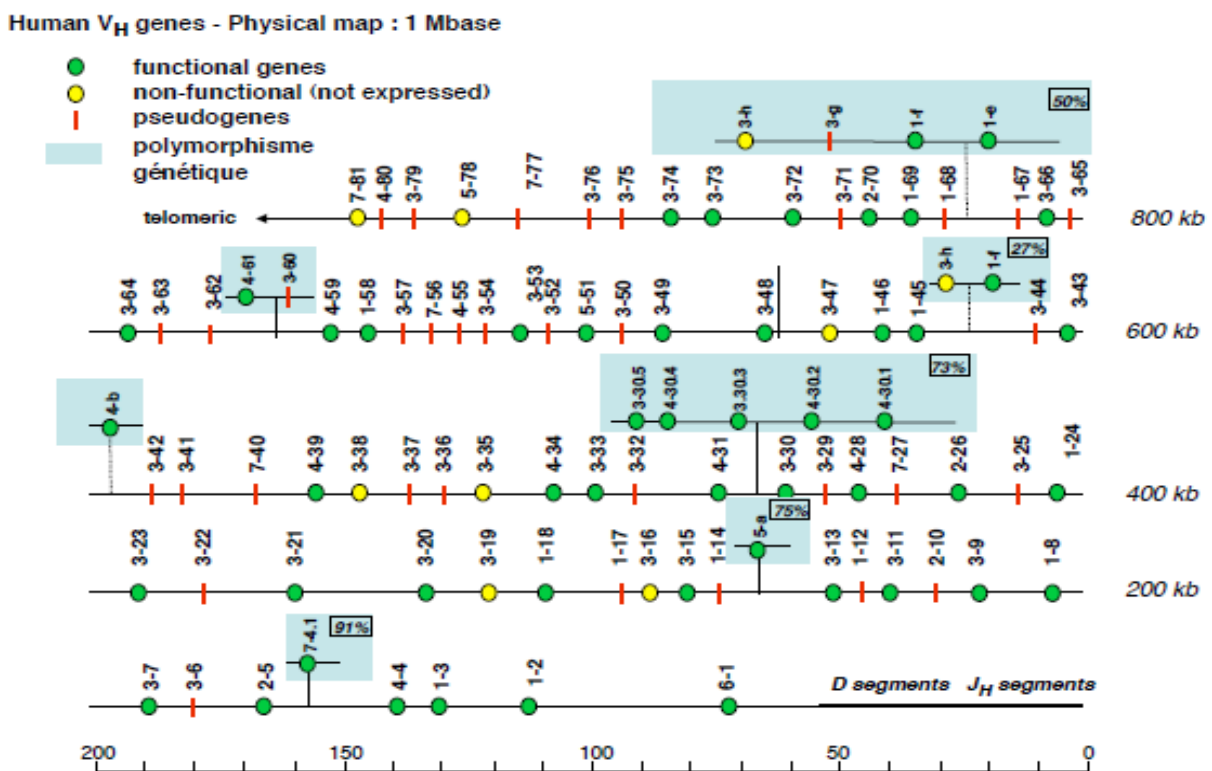
VH3 et VH4 sont les 2 plus grands sous groupes chez l'homme.

VH1 est le 3ème sous groupe.

On a ensuite des groupes extrêmement restreints avec seulement 1 ou 2 gènes fonctionnels.

⇒ C'est une sorte d'**histoire évolutive** de la façon dont ces gènes ont divergé, ce sont dupliqués et ont évolué les uns par rapports aux autres.

La carte des gènes VH chez l'homme (= fragments V des chaînes lourdes des gènes d'Ig)



Le séquençage des gènes VH chez l'homme fait environ 1 Méga base (Mba), chaque ligne fait 200 Kba.

Il y a une nomenclature très simple qui fait qu'on les appelle par le nom de leur famille et par leur position sur le chromosome.

Le premier c'est le gène VH6, on l'appelle 6-1 car c'est le premier sur le chromosome. Ensuite on a le gène 1-2 qui appartient à la famille VH1 et est en 2ème position sur le chromosome. (On part de la dernière ligne.)

La famille VH3 est dispersée sur totalité du locus.

Ronds verts : gènes fonctionnels

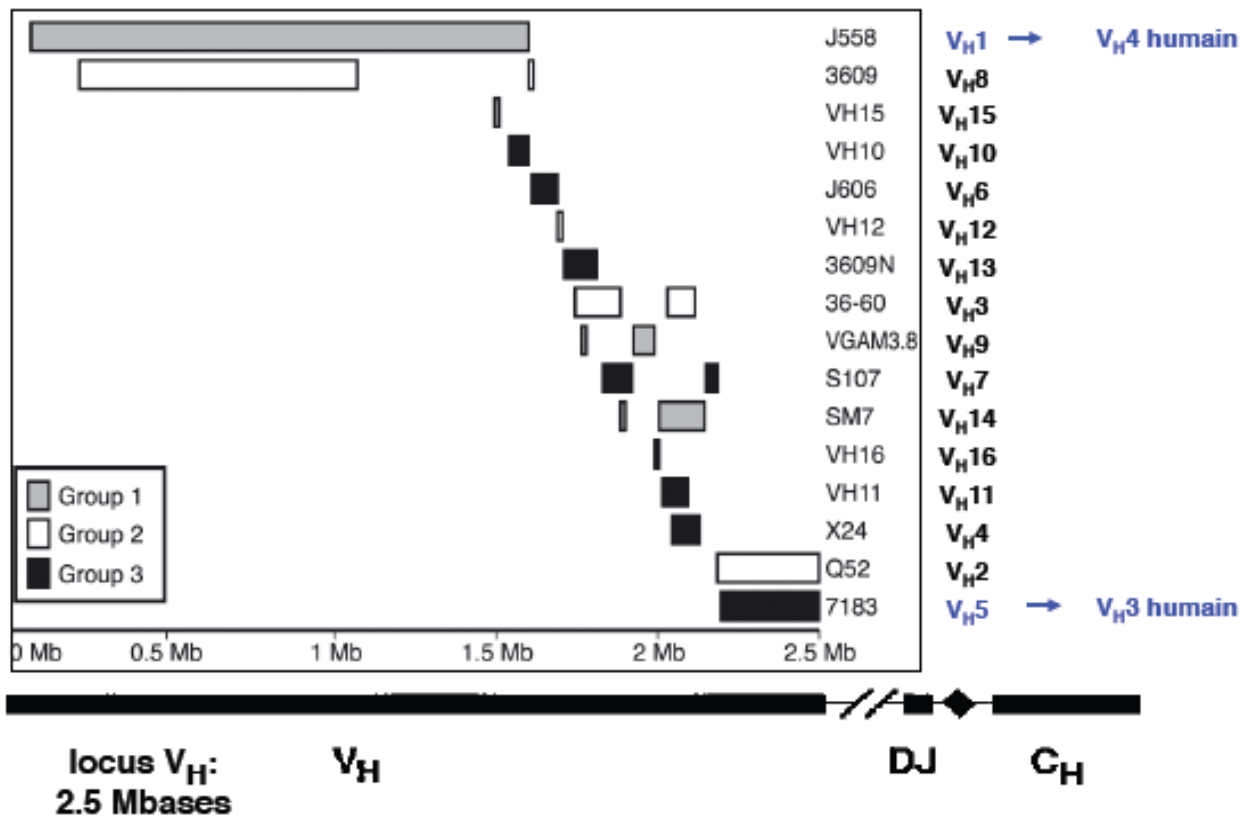
Ronds jaunes : gènes qui pourraient être fonctionnels mais qui ne sont jamais exprimés, ils ont sûrement des problèmes de promoteur.

Barres rouges : pseudogènes avec des défauts de séquence et qui ne peuvent pas être codants.

Boîtes bleues : polymorphisme génétique → nous n'avons pas tous les mêmes gènes V_H , même si les différences sont relativement modestes. En plus, certaines zones de variabilité correspondent à des gènes parfaitement fonctionnels, il y a donc des différences entre 2 individus jusqu'à une dizaine de gènes fonctionnels sur une quarantaine par polymorphisme génétique.

⇒ Ce sont des loci qui sont encore extrêmement variables.

En comparaison chez la souris : 16 familles



Une famille très grande correspond en gros à la famille V_H4 chez l'homme.
 Une autre famille est relativement grande et correspond au V_H3 .

Une chose est très frappante chez la souris, c'est récurrent dans le locus des Ig ; c'est que les familles ségrégent :

Les familles V_H1 et V_H8 sont mélangées mais sont toutes dans un coin.

Les familles V_H2 et V_H5 sont dans un autre coin.

Certaines familles sont un petit peu plus dispersées ou plus homogènes, des familles avec très peu de membres.

On observe une **segmentation** comme si les éléments de duplication se faisaient de façon plus proximale et comme si chez l'homme c'était plus des groupes de gènes divergents, des ensembles qui étaient dupliqués.

Les gènes de la famille D :

Chez la souris on a des gènes extrêmement semblables et qui sont côte à côte.

Chez l'homme on a un ensemble d'éléments de D1 à D6 qui sont relativement différents mais c'est le groupe ou les unités qui vont être dupliqués

Chez la souris, on a une extrême variabilité de la région D qui code pour le CDR3 des Ig.

La chaîne Kappa chez l'homme :

La totalité du locus qui comprend 40 gènes est dupliqué et inversé sur environ 400 Kba.

Donc quand on a séquencé une Ig kappa chez l'homme, c'est à peu près impossible de savoir si elle vient de ce gène là ou de son symétrique dans la duplication parce que ce sont des gènes quasiment identiques.

À la fin du locus il y a des petites différences, mais c'est une structure tout à fait étonnante avec sa duplication et son inversion qui portent sur à peu près 400 Kba.

Alors que chez la souris, il y a des gènes V kappa qui sont dans les 2 sens mais qui sont mélangés et distribués sur le génome.

⇒ Ce sont les techniques de séquençage qui ont permis d'obtenir le détail complet de ces organisations

III. Mécanismes responsables de la formation du répertoire

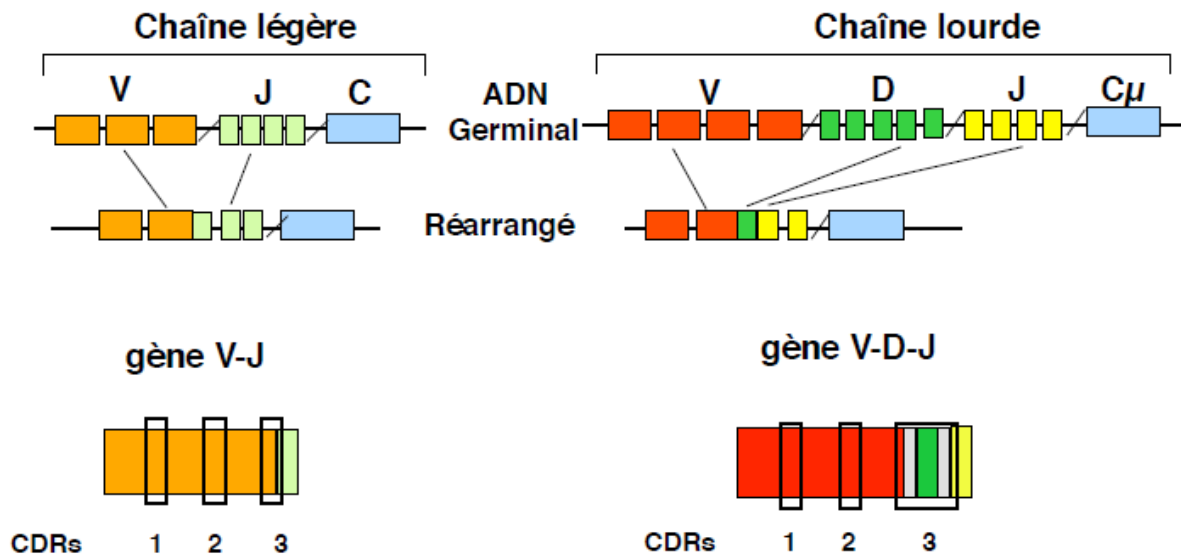
1) Mécanismes du réarrangement

a. Réarrangements VDJ

Le répertoire pré immun correspond au répertoire germinale, mais c'est essentiellement le produit d'un mécanisme de remaniement de l'ADN au cours du développement des lymphocytes B, qui s'appelle le réarrangement des gènes des immunoglobulines ou la **recombinaison VDJ**.

C'est une recombinaison entre 2 éléments de la chaîne légère : V et J.

Et c'est une recombinaison entre 3 éléments de la chaîne lourde : V, D et J.



Par exemple : pour la chaîne lourde
On a

- des régions V
- des régions D
- des régions J
- des régions constantes qui déterminent la variabilité isotypique de l'Ig (μ pour les Ig M, α pour les Ig A, γ pour les Ig G, δ pour les Ig D, ϵ pour les Ig E)

La cellule souche B contient tous les gènes V, D et J et toutes les parties constantes.

1er réarrangement : entre D et J avec perte de l'ADN entre les deux gènes choisis.

Par exemple D4 et J3 sont retenus et l'ADN entre les deux est excisé.

2ème réarrangement : on associe un fragment V à la paire DJ retenue avec perte de l'ADN entre le fragment V et la paire DJ.

Par exemple V2 s'associe à D4J3 avec perte de l'ADN entre V2 et D4.

La partie codante est donc V2D4J3 + la partie constante retenue.

Ce réarrangement donne **la partie variable de l'Ig**.

Il y a plusieurs aspects en termes de répertoire dans ce processus :

- ⤴ un aspect que l'on appelle la combinatoire, c'est-à-dire que si on a 10 gènes V et 5 gènes J, on a 50 combinaisons possibles alors qu'on a seulement 15 gènes.
- ⤴ Cette jonction entre les différents fragments a lieu à l'intérieur du CDR3.
- ⤴ De plus, ce que ce processus a d'unique, c'est qu'il est imprécis. Il y a donc une **variabilité de jonction** très importante créée lors de ce processus.
- ⤴ Il existe une enzyme spécifique qui va inclure des nucléotides au hasard, ce qui crée un élément de diversification considérable. On appelle ce processus **diversification de séquence N** car l'ajout peut être n'importe quels nucléotides, il n'y a pas de matrice.

Dans la plupart des livres d'immunologie, on trouve le répertoire produit par la combinatoire du réarrangement. C'est-à-dire que si on a 100 gènes VH, 10 gènes D et 4 gènes J, 100 gènes VL et 4 gènes J, et qu'on a une combinatoire VH/VL (= n'importe quelle chaîne lourde peut aller avec n'importe quelle chaîne légère), on a environ 2 millions de possibilités.

On voit ça souvent mais ce chiffre n'a aucun sens parce que 99% du répertoire est au niveau du CDR3 de la chaîne lourde, codé par le réarrangement V(D) J.

⇒ **Quel facteur peut-on mettre pour la variabilité de jonction ?**

La variabilité de jonction pour la chaîne légère est relativement modeste, la variation peut aller de 7 à 11 acides aminés.

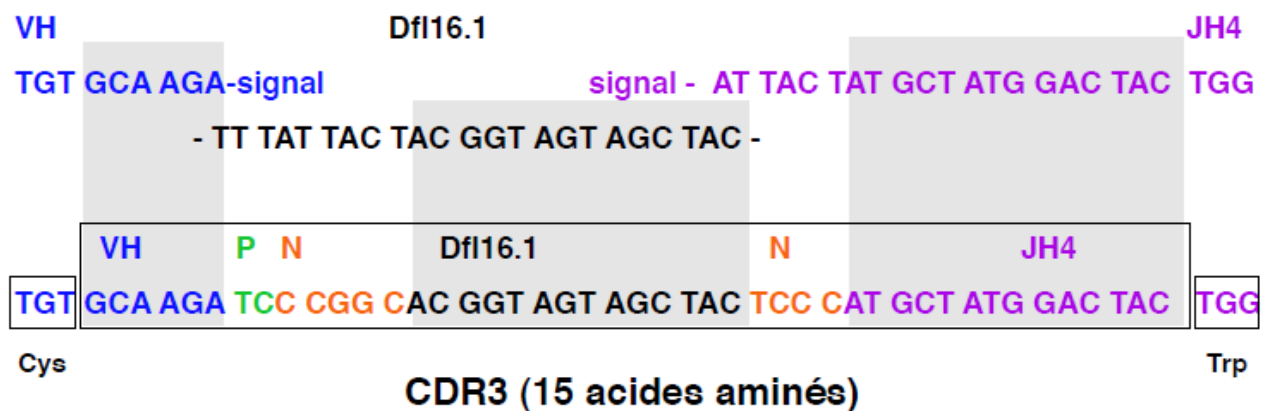
À l'inverse elle est considérable pour la chaîne lourde, vu que le processus de recombinaison peut produire une diversité qui peut aller de 6 à 20 aa. En admettant que la variation soit de 10 aa, cela représente une différence de 45 nucléotides.

Exemple : si on prend un gène VH, un gène D et un gène JH.

On a pris la totalité du **gène VH**, mais on a pris seulement une partie du **gène D** (car la jonction est imprécise), de même pour le **gène J**.

Pour ajouter des **séquences N**, on ajoute des **séquences P** (palindrome = ses 2 bases sont complémentaires des deux dernières).

⇒ Tout cela forme un CDR.



Si le CDR fait 10 aa, on a 10 positions où on peut avoir à peu près n'importe quel aa grâce à la recombinaison VDJ et à la diversification de séquence N. Donc si on a 10 positions avec 10 aa différents, ça fait **10¹⁰ combinaisons**.

D'où le fait que le répertoire des Ig est quasiment illimité.

⇒ Il est clair que simplement la jonction VDJ de la chaîne lourde est un facteur absolument dominant dans la création du répertoire pré immunitaire.

Si on regarde le répertoire des lymphocytes B naïfs (= ceux qui n'ont pas encore rencontré d'antigène et donc ne sont pas actifs), la séquence du CDR3 de la chaîne lourde est un marqueur de la cellule car elle est unique ; on peut dire que c'est un **marqueur de clonalité**.

Si on prend des LB naïfs, on ne trouvera quasiment jamais 2 séquences identiques. Donc quand on commence à voir des LB avec le même CDR3, c'est ce qu'on appelle la marque d'une **expansion clonale**. Cela peut être :

- ↳ soit un processus physiologique au cours de la réponse immune
- ↳ soit c'est un marqueur de tumeur, c'est-à-dire qu'il y a un clone B qui est en expansion pathologique. C'est un marqueur diagnostique du lymphome B, même dans la thérapie. C'est-à-dire qu'on peut évaluer la maladie résiduelle au cours de la chimiothérapie grâce à ce marqueur, il sert aussi à évaluer une possible rechute.

L'immunoscope ou CDR3 spectratyping

C'est une technique initialement de laboratoire mais qui a eu une vraie utilité en clinique. Cela correspond à une **mesure de la diversité clonale** d'une population lymphocytaire pour voir quel est le degré d'hétérogénéité des régions CDR3.

Cette technique utilise la PCR et ensuite une migration sur gel de produits de PCR.

- ↳ Elle nécessite des amorces, elle se fait au niveau de l'ARM m mais elle peut parfois se faire avec de l'ADN car cela permet d'avoir des amorces dans la région constante et de sélectionner la région constante que l'on veut.

On analyse les produits de PCR soit sur électrophorèse capillaire soit sur gel, ce qui permet de découper les bandes et de les séquencer.

⇒ Ainsi on évalue la diversification clonale.

Exemple :

On est au niveau du répertoire normal des cellules B.

On compare la diversité en termes d'acides aminés de la région CDR3 liée au gène VH3-15 entre des lymphocytes B naïfs et des lymphocytes B mémoire.

- Au niveau des LB naïfs, on a une distribution gaussienne ; c'est-à-dire une extrême variabilité.
- Au niveau de la population de cellules B mémoire d'enfants vaccinés, la diversité est quasiment totalement absente. On a donc des amplifications clonales physiologiques colossales au cours de la réponse immune ; à ne pas confondre avec un lymphome.

b. Les facteurs du réarrangement

1) Signaux de recombinaison

Ce sont des signaux au niveau de l'ADN donnant des informations aux enzymes.

Ils sont extrêmement précis, c'est-à-dire que chaque élément d'Ig a juste à côté de sa séquence codante 2 éléments :

- Un heptamère (7 nucléotides)
- Un nonamère (9 nucléotides)

L'heptamère d'un côté est complémentaire de l'heptamère de l'autre, c'est un palindrome (=séquence d'ADN pouvant se lire de la même façon dans les deux sens par rapport à un point central) : CAC est complémentaire de GTG.

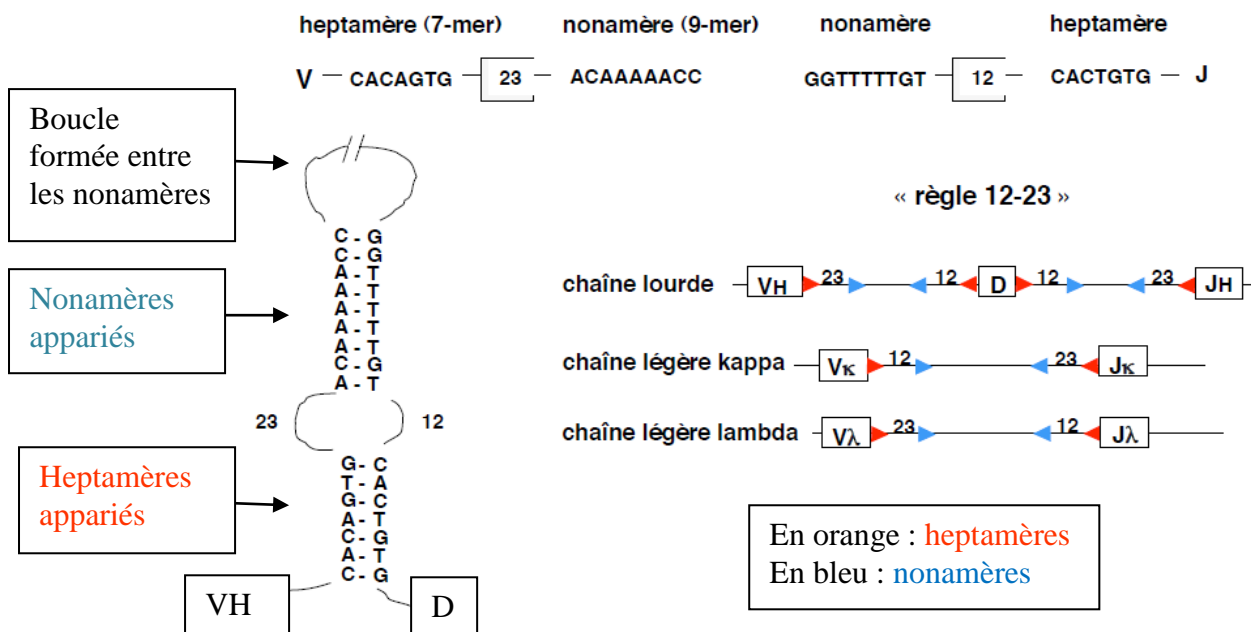
Le nonamère de la même façon est complémentaire du nonamère d'à coté.

Par contre il y a une règle absolue : pour que 2 segments d'Ig soient rapprochés par le réarrangement il faut que ces 2 signaux heptamère, nonamère soient espacés de 23 nucléotides d'un côté et de 12 de l'autre.

↳ On considère que 12 c'est un tour d'hélice et 23 c'est 2 tours d'hélice.

V et J peuvent être loin, c'est-à-dire qu'il peut y avoir un nombre important de nucléotides entre les 2 nonamères.

⇒ C'est la structure même de la reconnaissance des enzymes impliquées dans le réarrangement, qui demande la présence de ces 2 signaux avec ce double espacement.



Dans l'évolution du génome, ce ne sont pas toujours les mêmes espacements :

↳ Pour la chaîne lourde, on a le fragment de 23 nucléotides du côté du gène VH et le 12 du côté du gène D.

↳ Pour la chaîne légère kappa, on a le fragment 12 du côté du gène V et le 23 du côté du gène J ; et inversement pour la chaîne lambda.

Intérêt : on ne peut pas sauter le gène D lors du réarrangement.

Il y a une logique dans la façon dont les loci sont réorganisés qui fait que chaque élément suit un processus plus ou moins dirigé, évitant les fusions des gènes D ou leur oubli pendant le réarrangement.

⇒ Ce processus permet que le gène VH se retrouve à côté du gène D.

Il se produit la même chose entre les gènes D et J pour la chaîne lourde et entre les gènes V et J pour les chaînes légères.

2) Facteurs spécifiques aux cellules lymphoïdes

- Les 2 enzymes faisant le réarrangement s'appellent **RAG 1** et **RAG 2** (RAG = recombination activating genes).

Elles ont été découvertes par Baltimore.

Il s'était fait un substrat de réarrangement : un gène de résistance à un antibiotique proche des gènes d'Ig. S'il y a réarrangement, la cellule va acquérir la résistance à cet antibiotique. Il a mis ce substrat dans une cellule, il a transfecté cette cellule avec des morceaux de DNA génomique et certains ont réarrangé leur substrat.

Les gènes RAG 1 et RAG 2 sont petits et sont à 2 Kba de distance.

En transfectant l'ADN, Baltimore a donc pu récupérer les 2 gènes ensemble.

Si on inactive soit RAG 1 soit RAG 2, on n'a aucun LB et aucun LT.

A eux seuls ils sont suffisants pour induire le réarrangement, car il y a des protéines accessoires qui sont présentes dans toutes les cellules.

- ⇒ Ce sont les seuls gènes lymphoïdes spécifiques nécessaires pour faire le réarrangement des gènes des Ig.

- Il y a un autre facteur lymphoïde spécifique (qui donc n'est pas présent ailleurs que dans les lymphocytes), c'est la **Tdt** (terminal deoxynucleotidyl transferase).

Elle participe à la diversité du répertoire des Ig.

En cas d'absence, on a un répertoire qui est plus restreint, mais on aura quand même un réarrangement des gènes des Ig et on aura des LB et des LT.

- ⇒ C'est un facteur important mais facultatif.

3) Facteurs ubiquitaires qui contribuent à la réparation de l'ADN

Ces facteurs, importants en nombre, sont indispensables au réarrangement des gènes des Ig, même si leurs tâches sont beaucoup plus générales.

Si ces facteurs sont inactivés, on n'aura aucun LB et aucun LT.

- ▲ **DNA-PKcs** (Protéine kinase DNA dépendante avec une sous unité catalytique)

Elle a été identifiée grâce à une mutation spontanée chez la souris.

Son activité participe à la réparation des cassures double brin de l'ADN, or le réarrangement est une cassure double brin.

- ▲ **Artemis**

C'est une exo/endo nucléase.

- ▲ **Complexe ligase** (DNA ligase IV, XRCC4, Cernunnos/XLF)

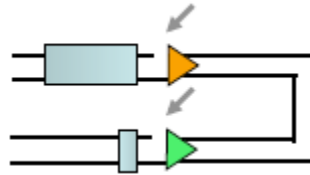
Ce complexe est nécessaire pour la ligation des extrémités de l'ADN.

- ▲ **Les protéines Ku70 et Ku 80**

c. Les étapes du réarrangement :

1. **Apparier les gènes D et J** (ou V et J, ou V et D) qui sont séparés par des distances dans le génome qui peuvent être considérables ; cela forme ce qu'on appelle une *synapse*. Cet appariement est rendu possible par le complexe RAG des signaux de recombinaison. On a besoin de RAG1 et RAG2 ensemble, séparément ils n'ont pas d'activité.

2. Le complexe RAG va faire une **cassure simple brin** de l'ADN des 2 côtés au niveau des heptamères.



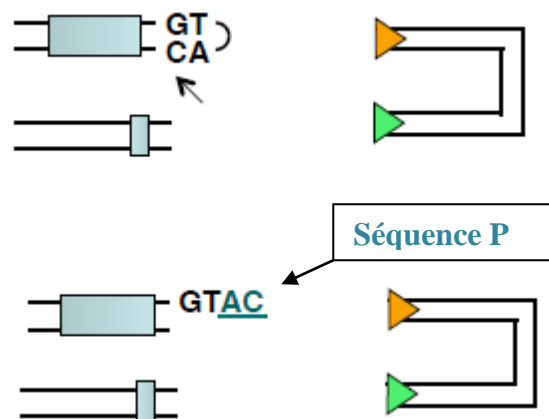
3. Il s'en suit une **cassure double brin** en formant une structure en épingle à cheveux.



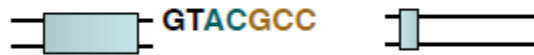
Le réarrangement est un acte invasif pour l'ADN, il y a alors tout un processus qui vise à le rendre le moins dangereux possible.

Cette structure en épingle à cheveux est un moyen de garder le contrôle sur cette cassure double brin.

4. Les **protéines Ku 70 et Ku 80** se lient aux extrémités d'ADN
5. L'ouverture de la structure en épingle à cheveux est faite par la **DNA-PKcs**.
6. **Artémis** coupe à différents endroits de l'épingle à cheveux. Elle coupe très souvent à 2 bases de l'extrémité, ces 2 bases libérées étaient initialement sur le brin opposé et donc sont complémentaires. Ce sont les *séquences P*.



7. On a ensuite une addition de bases par la **Tdt**. Ou parfois de la dégradation : une partie de la séquence codante est enlevée.



8. Religation des 2 côté par le complexe ligase

⇒ On finit par produire 2 choses :

→ Une jonction qui contient les éléments signaux, la jonction précise, qui va être perdue.

→ Une jonction codante, qui est le produit d'un processus extrêmement imprécis, et donc générateur de diversité.



jonction codante
(imprécise)



jonction signal
(précise)

Phase RAG dépendante : étapes 1 à 3

Phase dépendante des facteurs ubiquitaires : étapes 4 à 8

⇒ Il y a énormément de facteurs et d'activités qui travaillent au niveau de cette jonction, et toute une partie, au début avec le complexe RAG, qui protègent cette jonction pour que les 2 fragments d'ADN cassés soient sous contrôle.

d. Phénotype de l'inactivation des différents composants du mécanisme de réarrangement

➤ *Chez la souris*

Inactivation de RAG1 et RAG2 : phénotype SCID (= pas de LT et LB)

↳ Phénotype uniquement restreint au système immunitaire

Inactivation de la DNA-PKcs ou d'Artemis : on a un phénotype SCID et une radiosensibilité modérée

Inactivation des protéines Ku70 et Ku80 : on a un phénotype SCID + un retard de croissance + une apoptose des cellules neurologiques (ce sont les plus sensibles à l'absence de facteur de réparation aux cassures double brin, surtout au stade embryonnaire) + radiosensibilité

Inactivation de la ligase IV-XRCC4 : létalité embryonnaire tellement l'apoptose neuronale est importante

- ⇒ Le phénotype est de plus en plus grave car son incidence est croissante hors du système immunitaire.
- ⇒ On a un gradient de phénotype.

➤ *Chez l'homme*

RAG 1 et 2

Mutation inactivante : SCID

Mutation hypomorphe (= un tout petit peu d'activité avec très peu de clones et donc très peu de diversité) : SCID + auto-immunité forte et beaucoup d'Ig E

- ↳ Syndrome d'Omenn

Mutation d'Artemis : SCID + radiosensibilité modérée

Mutation de Cernunnos (= cofacteur ligase IV) ou de la ligase IV: SCID, microcéphalie, retard mental

- ↳ Ces facteurs n'ont pas tout à fait le même rôle chez l'homme et chez la souris, car leur inactivation ne donne pas exactement le même phénotype.

- ⇒ Pour résumer la fonction du réarrangement : le répertoire théorique souvent excède le nombre de LB de l'organisme ; on a cependant des réarrangements à de plus grande fréquences que d'autres (ex : gène 3-23), d'autres plus rares qui vont émerger au cours de la vie générant une nouvelle diversité de combinatoire du réarrangement.
- ⇒ Il y a beaucoup d'espèces pour lesquelles le réarrangement n'est pas générateur d'autant de diversité. Et donc elles utilisent des mécanismes de diversification additionnels pour former leur répertoire pré immun.

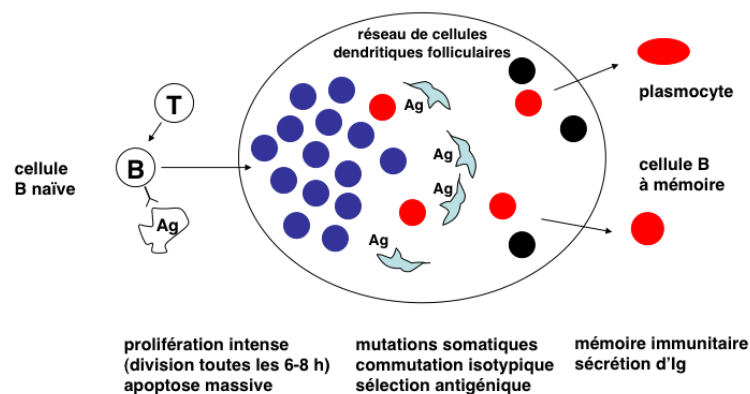
2) La réponse immunitaire

- Le **réarrangement V (D) J** que l'on a vu jusqu'à présent était impliqué dans **la formation du répertoire pré-immun**, c'est-à-dire avant la rencontre des cellules B avec l'antigène
- Au cours de **la réponse immunitaire**, une **2^{ème} étape de diversification des gènes des Ig** va avoir lieu au cours de 2 mécanismes :
 - **les mutations somatiques**
 - **les commutations isotypiques**
- La diversification des gènes au cours de la réponse immunitaire est un **phénomène « adaptatif »** unique au monde vivant, par lequel l'antigène va induire une **mutagenèse ciblée du récepteur** responsable de sa reconnaissance et de son élimination, afin qu'il reconnaisse mieux cet Ag

- Les premières observations datent des années 1970 grâce à l'analyse de séquences protéiques de chaînes légères λ de myélomes de souris.
- On a pu comparer les séquences pour démontrer que les hybridomes différaient par quelques acides aminés qui étaient retrouvés dans des régions hypervariables.
- Dans l'article paru dans Nature en 1970, les auteurs ont conclu très justement que la variabilité vue dans ces hybridomes suggère que **la diversité est générée par des mutants spontanés et sélectionnés progressivement, acides aminés par acides aminés.**

Comment tous ces mécanismes cellulaires se déclenchent-ils ?

- La rencontre d'une cellule B et de l'antigène, en présence d'une cellule T auxiliaire spécifique et de cellules dendritiques, va activer les cellules B
- Cette activation déclenche leur **prolifération intense** (elles vont se diviser toutes les 6-8 heures) mais aussi un phénomène d'**apoptose massif**
- Il existe des structures appelées **centres germinatifs**, dans la **rate**, dans les **ganglions** et dans les **organes lymphoïdes secondaires**
- Dans ces centres se trouve un **réseau de cellules**, les **cellules dendritiques folliculaires** sur lesquelles l'antigène va être retenu sous forme de complexe immun
- Les cellules B qui prolifèrent vont engendrer des **mutations somatiques** et la **commutation isotypique** au niveau de leur génome
- Les cellules B qui reconnaissent alors le mieux l'antigène vont l'internaliser et recevoir des signaux les faisant sortir du centre germinatif
- Ainsi, à la sortie du centre germinatif, les cellules B vont alors donner
 - soit des **plasmocytes** qui sécrètent des Ig \rightarrow c'est la fonction effectrice des cellules B
 - soit des **cellules B mémoires** qui peuvent rester des semaines, des mois, des années, voire des décennies chez l'homme



Que se passe-t-il au niveau moléculaire ?

- Petit point historique :
 - la maturation de l'affinité de la réponse immunitaire a d'abord été étudiée avec des haptènes (réactifs chimiques de toute petite taille) couplés à une protéine « carrier »
 - les haptènes sont incapables de déclencher une réaction immunitaire par eux-mêmes
 - un certain nombre d'entre eux (phényl-oxazolone, arsenate, phosphoryl-choline, dinitrophenol) génèrent des réponses récurrentes, basées sur l'existence d'un répertoire germinale mobilisant une combinaison V_H-V_L réarrangée, d'occurrence

- fréquente et présentant une bonne affinité pour l'haptène
- ainsi, l'immunisation des animaux entraîne la production d'hybridomes à différents stades de la réponse
- on peut donc analyser l'affinité des anticorps produits et caractériser les gènes des immunoglobulines
- le groupe Berek et Milstein a eu le prix Nobel avec l'invention des hybridomes et des anticorps monoclonaux
 - c'est le début du séquençage des ARN à l'époque
 - leur système modèle est la phényl-oxazolone = un petit haptène possédant un noyau benzénique qui mobilise de manière canonique un même gène V_H et un même gène V_L
 - ils ont comparé les séquences au cours de la réponse primaire, secondaire et tertiaire : il existait bien des mutations des immunoglobulines, dont 2 mutations clés au niveau de la chaîne légère
 - ces mutations, à elles seules, peuvent augmenter d'un facteur 10 l'affinité des anticorps
- **Conclusion** : l'augmentation de l'affinité est liée au changement de quelques mutations clés
- Aussi, au cours des immunisations successives, d'autres réarrangements des Ac de plus faible affinité initiale sont susceptibles d'entraîner une forte amélioration de l'affinité pour l'antigène et donc de la réaction immunitaire

a. Mutation somatique - Hypermutation

Qu'est-ce que l'hypermutation ?

- Il s'agit d'une **mutagenèse** de l'ADN ciblant la région variable → la région constante n'est pas touchée
- Cette mutagenèse s'enclenche à des stades de différenciation spécifiques des cellules B, lorsqu'elles sont activées par un antigène
- Son taux est en moyenne de 10^{-3} mutation/pb/division cellulaire donc 1/1000 base à chaque division (soit 100 000 fois plus que la mutagenèse spontanée de l'ADN à chaque division)



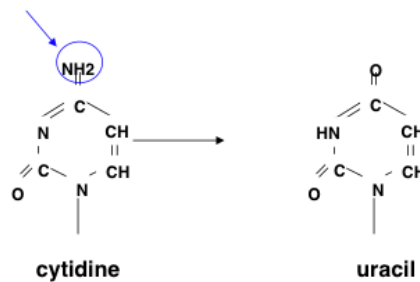
- La région ciblée se situe essentiellement à 2 kb en aval du promoteur et n'est pas spécifique des gènes V
- En effet, si on met n'importe quelle séquence entre le promoteur et l'enhancer des Ig, elle sera mutée de la même façon
- Par contre, elle dépend des éléments régulateurs de la transcription : les **enhancers** et les **promoteurs**
- Elle affecte toutes les bases ATGC, essentiellement par des mutations ponctuelles (transition/transversion), mais elle n'est pas tout à fait aléatoire

- Même si l'ensemble de la région est ciblée, il existe des **régions préférentiellement ciblées** : **les hotspots** de mutations
- Le **hotspot de mutation** correspond à une **glycine** ou à une **sérine** qui sont des acides aminés hydrophiles extrêmement représentés dans les **CDRs**
- Ils sont essentiels pour les interactions avec les différents antigènes qu'ils peuvent reconnaître
- Or la sérine peut être codée par 6 codons : AGT, AGC, TCN (N étant n'importe quel acide aminé)
- Il va donc y avoir une évolution des gènes V de façon que pour les CDRs :
 - dans les **hotspots** de mutation, les sérines soient codées par **AGT ou AGC**
 - dans les **frameworks** (séquence de base, résidus invariants), elles soient codées par **TCN** qui sont des combinaisons moins mutables
- On a alors une co-évolution du mécanisme d'hypermutation et de ses gènes cibles des immunoglobulines

Comment ce mécanisme se produit-il ?

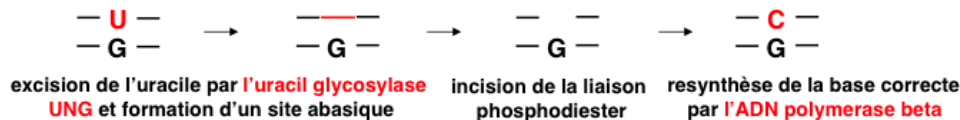
- *Brenner et Milstein, en 1966, proposent un mécanisme dans lequel l'ADN est clivé, puis réparé avec une enzyme → il a fallu 30 ans pour montrer que c'est à peu près ça qu'il se passe*
- Plusieurs éléments vont entrer en compte dans le mécanisme de l'hypermutation :
 - des **facteurs spécifiques**, qui vont cibler le locus des immunoglobulines
 - des **facteurs ubiquitaires**
 - le **ciblage** ? → *la question se pose car on ne possède pas l'équivalent des heptamères et des nonamères*
- Le **facteur spécifique de l'hypermutation** est le gène de l'AID (Activating Induced cytidine Deaminase) qui a été isolé en 1999
 - *M. Honjo essayait d'isoler les facteurs responsables de la commutation isotypique et pour cela faisait une soustraction, car à cette époque le séquençage systématique n'existait pas*
 - *Principe de la soustraction = on prend une cellule qui fait « quelque chose » et une autre qui ne fait pas « quelque chose », on se débarrasse de tous les ARN communs à ces 2 cellules pour ne garder que ceux qui sont spécifiques*
 - *A ce moment, avec un peu de chance, on met en évidence les facteurs impliqués dans ce « quelque chose »*
- En utilisant la méthode de la soustraction, M. Honjo a trouvé le gène de l'AID, qui **code pour l'enzyme de l'hypermutation**, mais aussi de la **commutation isotypique** et de la **conversion génique**
- L'AID est une **cytidine désaminase** → elle transforme les cytidines en uraciles

AID : cytidine deaminase (cytidine → uracile)



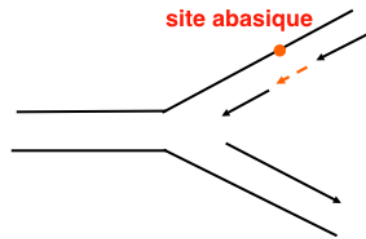
Que se passe-t-il quand un uracile est présent dans l'ADN ?

- En vérité c'est une **lésion très banale** car il peut y avoir une **erreur au cours de la réplication** : les dUTP peuvent être incorporés à la place des dTTP
- La déamination est par conséquent spontanée dans l'ADN
- Des **uraciles glycosylases** (Ung, Smug, MBD4...) sont capables d'**exciser l'uracile**, produisant ainsi un **site abasique** (*présence de la liaison phosphodiester sucre-carbone mais plus de base*)
- Une **ADN polymérase β** est capable de **resynthétiser la base correcte**
réparation fidèle par le mécanisme d'excision de base



Mécanisme moléculaire de l'hypermutation : une réparation détournée de sa fonction première

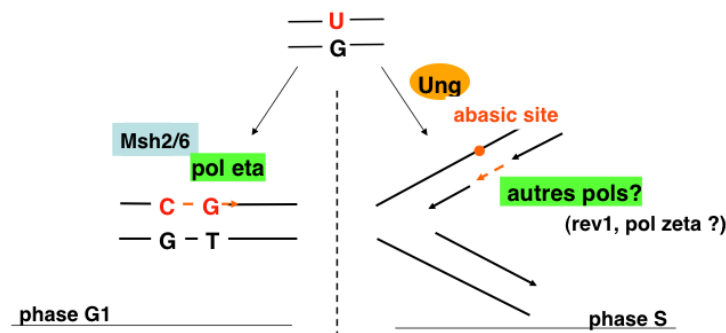
- Dans l'hypermutation somatique ce mécanisme de réparation ne se passe pas bien puisque des mutations apparaissent
- L'**AID cible les gènes des Ig, les désamine, mais on a ensuite une réparation mutagène**
- C'est un mécanisme complexe de « réparation » par des **facteurs ubiquitaires** permettant l'introduction de mutations au-delà des seuls sites de déamination
- Le détail de ce mécanisme n'est pas entièrement connu mais on sait qu'il **mobilise essentiellement 3 composants** :
 - l'**uracile glycosylase** mais à contre-emploi
 - les **enzymes du mismatch repair** (un système qui normalement s'occupe des erreurs, des mésappariements au moment de la réplication)
 - des **ADN polymérases translésionnelles** qui sont non-réplicatives et à très fort taux d'erreurs spontanées
- *Il existe un syndrome chez l'homme, le **syndrome XP-V** (variante du Xeroderma pigmentosum) qui correspond à l'inactivation d'une de ces polymérases → risque accru de cancers cutanés UV-induits, inactivation de la polymérase éta*
- Comme ces **polymérases translésionnelles** ne sont pas très exigeantes au niveau de leur substrat, elles sont **capables de franchir des lésions de l'ADN « non-informatives »** bloquant habituellement la progression de la fourche de réplication (sites abasiques, dimères de thymidine provoqué par les UVs, base oxydée...)



- Après le franchissement de la lésion, les polymérase répliquatives reprennent leur fonction.
- Ces polymérase sont donc **mutagènes face à l'ADN normal et anti-mutagènes face aux lésions de l'ADN** → ceci explique le syndrome XP-V

Modèle proposé pour le mécanisme d'hypermutation

- Lorsque la **polymérase éta** voit un dimère « TT », elle est capable de synthétiser un dimère « AA » en face et est donc **antimutagène**
- Lorsqu'elle est **absente**, d'autres enzymes sont recrutées et la **réparation est alors plus lente et moins fidèle**
- Les patients qui ont une absence de la polymérase éta ont un taux de cancers cutanés accru car le **taux d'erreurs de réparation de l'ADN est plus élevé**
- Dans le mécanisme d'hypermutation somatique tous ces processus sont détournés de leur fonctions pour faire des mutations



- Très schématiquement, le processus commence par la **désamination d'une cytidine au locus des Ig**
- Si on est en phase G1 :
 - au lieu d'être reconnu par une glycosylase, l'uracile va être reconnu comme **mismatch** par une enzyme du **mismatch repair** (complexe Msh 2/6)
 - l'enzyme va alors **resynthétiser un brin d'ADN autour de la lésion** en mobilisant la polymérase éta
 - cette polymérase étant mutagène, elle **introduit des mutations**
 - elle crée aussi des **mutations au-delà du site initial de désamination** → ainsi, on mute au-delà des régions ciblées par l'AID
- Si on est en phase S :
 - l'**uracile glycosylase** va **exciser l'uracile** formant alors un site abasique
 - les **polymérase translésionnelles**, qui en face d'un site abasique peuvent insérer n'importe quel nucléotide, vont être recrutées
 - ces polymérase vont franchir le site abasique en **créant une mutation**

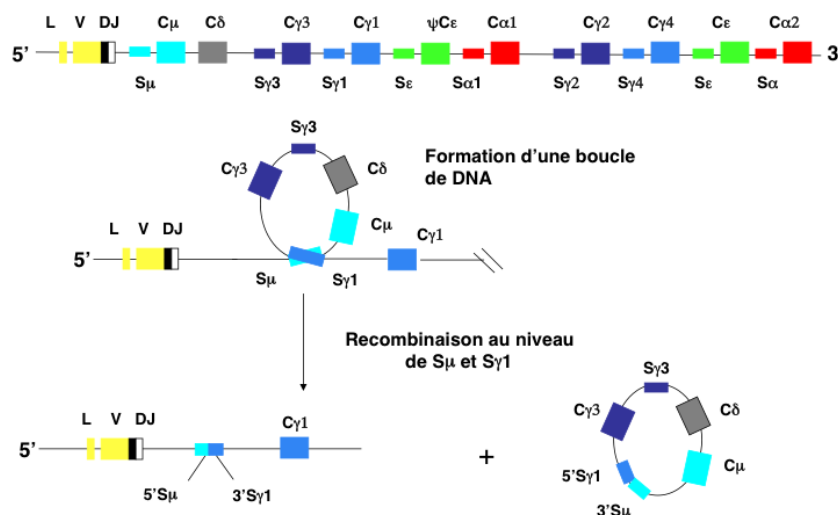
Quelques questions se posent encore aujourd'hui :

- Pourquoi la réparation des uraciles est-elle mutagène dans les cellules B (modifications spécifiques des facteurs impliqués) ?
- Comment se fait le ciblage au gène V, puisque contrairement au réarrangement, on a ni d'heptamère-nonamère, ni de signaux de reconnaissance ?

NB : On s'est aperçu que tout un panel de gènes était ciblé par l'hypermutation mais avec des fréquences moindres que pour les gènes des Ig. Ce phénomène est évidemment un facteur de survenue de lymphomes extrêmement important.

b. Commutation isotypique ou « isotype Switch »

- La commutation isotypique ou « isotype switch » correspond au 3^{ème} mécanisme de la diversité qui a lieu au cours de la réponse immunitaire
- Il correspond à un **mécanisme de recombinaison au niveau des régions constantes du locus des chaînes lourdes permettant un changement de fonctions effectrices** tout en gardant la spécificité de l'anticorps
- En amont de chaque gène codant les régions constantes (excepté en amont de C δ), il existe une **séquence de nucléotides répétés** → il s'agit de la **séquence switch**
- Ces **motifs répétés s'apparient 2 à 2** formant une boucle de DNA
- Cette **boucle de DNA** est ensuite **éliminée**
- On a donc une conservation de la spécificité du gène V mais avec des régions constantes différentes.



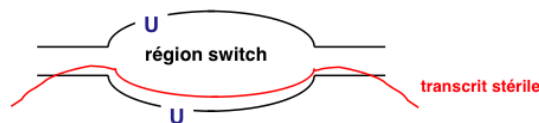
- Organisation des gènes au niveau constant : c'est un locus qui s'étend sur 200 kb, avec les isotypes μ et δ côte à côte, et plus loin 4 C γ chez la souris, 1 ϵ et 1 α
- Chez l'homme, il y a une duplication (et pas une amplification locale) de γ 3, γ 1, ϵ et α , avec la dérive possible de ces gènes en pseudogènes

Qu'est ce que les régions switch ?

- Les régions switch sont de petites régions de 1 à 2 kb (jusqu'à 10 kb pour certaines régions)
- Elles correspondent à des séquences **extrêmement répétées** : répétition de 10 pnb pour S μ , S α , S ϵ et environ 50 pnb pour S γ
- Ces séquences sont à la fois **riches en nucléotides G et en hotspots de mutations**
- *Un seul élément de 10 bases possède 3 hotspots (= site de reconnaissance pour l'AID), ce qui en fait une centaine sur toute la séquence Switch*

Mécanisme de la commutation isotypique

- L'**activation de la séquence Switch** se passe en **2 temps**
- Il faut tout d'abord **rendre ces régions accessibles** aux les enzymes de remaniement de l'ADN → ceci est permis par l'**activation d'une transcription**
- La transcription réalise alors ce qu'on appelle un **transcrit stérile**, car il ne sert à rien sinon à rendre la région accessible
- Chaque **chaîne lourde est activée** par un « cocktail » de **cytokines** qui lui est propre
- En raison de leur structure répétée, les transcrits « stériles » sont stables et restent fixés à leur région d'ADN correspondante formant des « **R-loop** »
- Les **hotspots** présents au niveau des séquences Switch sont par conséquent **exposés**, et activent alors l'**enzyme AID**



- L'**enzyme AID** crée une **cassure double-brin** suivie d'une **religation** → il s'agit donc d'un événement assez complexe et beaucoup de facteurs semblent y participer
- Au début, c'est toujours le facteur spécifique qui intervient : l'AID
- Puis, des facteurs ubiquitaires interviennent à leur tour :
 - l'uracile glycosylase et les molécules du mismatch repair pour les coupures double-brin
 - H2AX pour la formation de la synapse
 - les molécules du NHEJ : KU70, KU80, ligase IV, XRCC4
 - d'autres molécules de réparation : ATM, nbs1, 53BP1, ...

Les déficits immunitaires liés à un défaut de la commutation isotypique

- Le **syndrome hyper IgM** est connu depuis longtemps
- Les patients ont un trop plein d'IgM sériques, qui peuvent être normales, mais on ne trouve aucun autre isotype d'Ig dans le sang
- La cause peut être cellulaire et **concerner l'interaction entre les cellules T et les cellules B**
→ l'interaction entre les molécules CD40 ligand/CD40
- Pour CD40, seulement 3 cas sont avérés dans le monde
- Par contre, pour CD40 ligand, le gène se situe sur le chromosome X et la probabilité que la mutation ne soit pas silencieuse est donc bien plus élevée (98 sur 140)

- Des causes moléculaires sont aussi possibles :
 - des mutations de l'AID
 - un déficit en uracile glycosylase
 - des mutations dans les gènes du mismatch repair (*généralement ces patients ne vivent pas longtemps car ces molécules sont totalement fondamentales pour le maintien de l'intégrité du génome*)
 - d'autres causes clairement moléculaires et intrinsèques aux cellules B sont encore inconnues

Comment l'AID est-il ciblé au locus des Ig ?

- On a longtemps cru que les mutations somatiques étaient ciblées aux gènes des Ig, ce qui est faux
- Il existe une **vingtaine d'autres loci touchés par l'AID** mais à des **taux beaucoup plus faibles** que les Ig (10 à 50 fois)
- Certains de ces loci sont **faiblement mutés, comme bcl-6**
- D'autres sont **désaminés par l'AID et réparés, comme c-myc, Pax-5 et pim-1**
- Cependant, tous ces loci sont **susceptibles d'événement de translocation**, l'introduction de mutations après l'action de l'AID ou la réparation pouvant produire des cassures double-brin dans l'ADN dans des cellules à très forte prolifération
- Ces phénomènes de translocation peuvent être à l'**origine de la formation de lymphome par création d'oncogènes**
- *Dans les lymphomes, la réparation marche moins bien → toutefois, le ciblage par l'AID continue et on peut voire apparaître des mutations*
- **Chaque fois que la cellule B remanie son génome elle s'expose à une forte possibilité d'erreur**
 - dans le lymphome folliculaire, le gène bcl-2 est transloqué
 - dans le lymphome de Burkitt, le gène c-myc est transloqué
 - dans le lymphome B large diffus, le gène bcl-6 est transloqué
 - *d'autres translocations de bcl-6 sont possibles sur des oncogènes, ce qui est très probablement expliqué par le fait que ce gène est ciblé par l'hypermutation*

IV. Conclusion sur les mécanismes de formation du répertoire

- **Une recombinaison site-spécifique** = réarrangement → des signaux très précis ciblent des activités très précises
- **Une mutagenèse plus ou moins aléatoire** grâce aux hotspots de mutations lors des phénomènes de mutations somatiques = l'hypermutation
- **La commutation isotypique** → une recombinaison non homologue entre des séquences répétées (Switch)
- **Une recombinaison homologue non réciproque** qu'on appelle la conversion génique, responsable du répertoire immunitaire dans certaines espèces mais pas chez l'Homme
- *Remarque :* seul le premier existe sur les cellules T
- Tous ces mécanismes participent à la **formation du répertoire des cellules B et/ou à la maturation de la réponse immunitaire**
- Les **conséquences négatives** de cette extrême plasticité génomique sont représentées par les **événements de translocation** (et de mutation ?) à l'origine de nombreuses **tumeurs lymphoïdes** (90% des tumeurs lymphoïdes étant d'origine B chez l'Homme)