

# L'épigénétique

A préparer : TD1 : récupérer doc sur madoc, répondre aux rubriques.

Patient : profil pathologiq, phénotypique □ quels médicaments peuvent lui convenir ?

## I Intro

Capital génétique : l'info que l'on transmet à chaque division cellulaire.

Pour un mm capital génétique, toutes les cellules ne l'expriment pas de la mm façon.

- Certains gènes ne peuvent pas être lus : cela définit le type cellulaire.
- C'est ce qu'on appelle l'épigénétique.

Compréhension de ce « verrouillage de gène » : ça a pris du temps.

On a travaillé sur des jumeaux :

Mm patrimoine génétique, mais ils ont qd même des divergences phénotypiques, elles s'expriment physiquement (+ ou -+ important), mais surtout s'expriment par une sensibilité à telle ou telle pathologie.

Plus les jumeaux sont élevés ensemble, moins cette différence se voit.

- Facteurs épigénétiques influencés par facteurs environnementaux (alimentation, taux de pollution, exposition au soleil ...). Sans parler de mutation, c'est vraiment de la modulation d'expression des gènes.

1<sup>er</sup> clonage d'un animal : Dolly. Elle est viable mais ne ressemble pas à sa mère.

## II Historique

1<sup>er</sup> à avoir évoqué ces problèmes : Aristote, 350 av JC.

Il parle d'épigénèse.

Observation d'embryon de poulet. Il voit que les organes se différencient progressivement. Structures de plus en plus complexes. Il propose que cette complexification soit issue de forces extérieures.

Théorie avant : la préformation : l'individu est déjà préexistant et déjà tout fait. Il est juste tout petit.

15<sup>e</sup> 16<sup>e</sup> 17<sup>e</sup> : fluide émi par l'homme qui est super important. Avant : L'œuf de la femelle sert juste de receptacle.

Girolamo Fabrici et William Harvey. Soutiennent l'épigénèse et disent que les deux sont important : semence male ET semence femelle.

17<sup>e</sup> 18<sup>e</sup> : Malpighi, Wolff. Il met en évidence les spermatozoïdes. Mais comme la théorie préformationniste est TRES présente, il dit que les spz sont de tous petits individus.

Wolff fin 18<sup>e</sup>, dit qu'il y a des structures, et qu'elles se différencient vraiment en organes. C'est le 1<sup>er</sup> à parler sérieusement de morphogénèse.

19<sup>e</sup> : Théorie cellulaire. Provient d'une observation de liège. L'écorce est un alignement de carrés -> fait penser à l'organisation des chambres des moines, appelées les cellules. D'où le nom !

1. On commence alors à penser que toute chose est faite de cellule.
  2. Ces cellules sont des unités de base : pour la structure des tissus, pour la fonction, et donc la physiologie des êtres vivants.
  3. Ces cellules sont vivantes et peuvent apparaître à partir d'une autre cellule □ notion de division cellulaire.
- On doit donc avoir deux cellules (1 femelle et 1 male) qui fusionnent pour donner un embryon.  
Ça complète l'idée qui avait été donnée avant. Mendel avec ses petits pois.  
Flemming observe des chromosomes. Darwin et l'hérédité.

20<sup>e</sup> : confrontation généticiens, embryologistes : (1) -> l'information contenue dans les cellules n'est pas exprimée partout pareil. (2) -> Les tissus se différencient, puis ce qui ne sert plus est rejeté.

1942 : Conrad Waddington (sans les connaissances de l'ADN de Watson et Crick). Il s'intéresse au lien qu'il y a entre génotype et phénotype. Il dit que ce lien est l'épigénétique. Il y a une topologie. Elle est définie par des gènes. Différentes expressions de gènes, avec des degrés différents. Au sein de ce paysage, la bille emprunte différents chemins. Elle peut donc arriver en bas de la vallée avec des spécialisations différentes. Et plusieurs croisements successifs de ce type sont rencontrés. À la fin la cellule aura atteint un stade de différenciation final : cellulaire, nerveuse... Ce n'est pas irréversible. Avec l'aide de facteurs externes (de l'énergie ou autre), il est possible à la bille de remonter ou de passer d'une vallée à l'autre. C'est ce qu'on connaît maintenant avec les IPS (induced pluripotent stem cell).

1970-80 : épigénétique et hérédité.

Cellules en culture pour des manip. Comment on est capable de conserver la spécialisation d'une cellule en culture. Au bout de plusieurs passages, une cellule musculaire redevient fibroblaste.

On s'intéresse aussi aux cellules souches. Comment différencier une cellule à partir d'une pluripotente. Comment éteindre l'expression d'un gène.

Chez les sujets femelles, sur les deux chromosomes X, il y en a un sur les deux qui est pas du tout ou très peu exprimé.

On refait les travaux de Mendel avec les drosophiles. Certaines choses sont héréditaires, transmissibles, mais qui ne suivent pas les lois de Mendel. □ Phénomène supplémentaire qui vient modifier l'expression des gènes.

1975 : Méthylation de l'ADN : permet ou non la lecture des gènes. C'est un des phénomènes épigénétiques les plus importants. Redéfinition de l'épigénétique proposée par Waddington 40 ans plus tôt.

Changements de la fonction des gènes. Ces changements sont transmissibles au cours de la division cellulaire. (mitose/méiose). C'est pas de la modification de séquence (mutation).

### III Définitions et Rappels

Epigénétique :

Correspond à des modulations héritées d'expression de gènes, qui interviennent sans qu'il y ait de modification dans la séquence d'ADN. Ces modifications ne peuvent pas être imputées à des modifications de séquence... (diapo)

La forme la plus compactée de l'ADN : chromosome. Adn enroulé autour d'histones, tout ça très compacté.

Unité structurale de base de l'ADN : nucléosome : env. 200 nucléotides, enroulés autour de de l'histone (environ 146 pb), 11nm de diamètre. Histone : octamère. H2A, H2B, H3 et H4.

Si adn enroulé autour d'histone : difficile de lire.

Les 3 mécanismes majeurs de modification épigénétique :

1. Modification de l'ADN : la méthylation de l'ADN.
2. Modification des histones. : détermine la réorganisation du nucléosome : est-ce que le fragment d'ADN sera lisible ?
3. Implication des ARNs non codants. Il n'y a qu'1.5 à 2% du code génétique qui va donner des ARNm, mais 64% du génome est transcrit. □ ARN de transfert, small nuclear ARN, miRNA ...

## IV La méthylation de l'ADN.

« Rinks et holiday » ont observé des résidus cytosines qui peuvent être méthylés, en position 5 :

5-methyl cytosine. La méthylation nécessite des enzymes : DNA methyl transferase (DNMT). Conséquence de méthylation : silencing des gènes. Tout ça se fait dans les îlots CGP. C'est ce mécanisme qui va permettre de mettre en œuvre la spécificité d'une cellule.

Cette méthylation peut être héritée (des parents et d'une cellule à l'autre au cours des mitoses). Au cours de la division, ya une cellule mère qui a des groupements méthylés. Se divise et donne un brin à chacune des cellules filles. □ Fragment d'ADN hémi-méthylé chez la cellule fille (celui qui vient de la mère est méthylé, l'autre vient d'être synthétisé). La DNMT synthétise l'autre brin (conservation par rapport au brin parent). C'est une enzyme de maintien, de conservation, ou de maintenance. Elle maintient le profil des méthylation au cours de mitoses. Se fait en association avec PCNA. Et pleins de acteurs (compliqué) interviennent. Tous ces facteurs ouvrent encore un monde de la recherche.

En plus de la DNMT classique (DNMT1), il y a les *de novo* méthylTransférases DNMT3A et 3B. Celles-ci sont très sensibles aux facteurs environnementaux.

LSH et E2F6 : facteurs de transcription répresseurs, associés à ces deux enzymes. Si la stimulation perdure, la méthylation perdure, et est alors transmise de cellules en cellules.

Ya des processus de déméthylation de l'ADN. Si la stimulation des enzymes cesse, elles ne sont plus activées et arrête de méthyler l'ADN.

Ya aussi des facteurs de transcription qui protègent les gènes contre la méthylation (notamment les gènes essentiels à la survie de la cellule). C'est un processus passif.

Processus actif de deméthylation. Fait appel à une série d'enzymes : les TET enzymes.

5-methyl cytosine -> 5-hydroxymethyl cytosine -> 5-formyl cytosine 5-carboxy cytosine -> ... -> cytosine.

Gène la lecture -> Ne gêne plus la lecture.

En connaissant le profil de méthylation de l'adn, on sait quels genes sont actifs. On va aussi chercher à savoir à quel niveau de méthylation est arrivé le segment d'ADN. Certains suggèrent que ça donnerait un autre niveau de régulation. Ça pose encore des pbms au niveau technologique (en termes de sensibilité, de l'assurance avec laquelle on détecte telle ou telle forme de méthylation).

TET enzyme -> hypométhylation -> en cas de dérégulation : peut engendrer des pathologies.

Le profil de méthylation de certains gènes est transmissible aux enfants. On parle d'emprunte parentale. Parfois, celui de la mère est méthylé -> seul le gène qui provient du père est exprimé : profil ou emprunte paternelle.

Nombreuses pathologies avec des méthylation aberrantes.

- Hypométhylation : associée à instabilité génétique, surexpression oncogénique -> processus de cancérisation.
- Hyperméthylation : inactivation des certains gènes suppresseurs de tumeurs -> cancérisation.

Voir tableau pour des exemples de gènes inhibés par méthylation aberrantes.

## V Modifications de la structure de la chromatine/des protéines

ADN enroulé autour des histones -> si modification des histones : modif du nucléosome, de l'enroulement : modif de la lecture. Les extrémités NH2 des histones peuvent subir des modif. Celles des histones H3 et H4 sont les plus étudiées. (H2A et H2B ont aussi des modif, mais c'est bcp moins connu).

Initiation transcription : les facteurs de transcription déplacent l'histone en se fixant sur la tata box (par déstabilisation). Cette libération de l'espace permet la mise en place des activateurs de transcription, la venue de l'ARN polymérase. Tous ces facteurs et co-facteurs ne suffisent pas à expliquer le déplacement de l'histone. Il y a d'autres complexes ATP-dépendant qui se fixe sur l'ADN, provoque l'activation d'enzyme □ déplacement du nucléosome (dissociation Adn - histone).

Modifications des histones : Phosphorisations, acétylations, méthylations des queues des histones (arginines, Lysine)...

### Acétylation des histones

Se fait sur des groupements lysines. Greffe d'un groupement acétyle sur une lysine. C'est fait par des enzymes acétylases et les désacétylases font la réaction inverse. Lysine : charge positive. ADN : charge négative -> crée l'enroulement de l'ADN autour de l'histone : **hétérochromatine**. Quand il y a acétylation des lysines, elles perdent leurs charges positives : neutralisation des lysines -> perte de l'attraction ADN - histone. □ Rend l'ADN lisible : **euchromatine**.

Là on voit les processus uns par uns, mais il faut garder à l'esprit que dans la cellule, tout se fait en même temps.

Grande diversité d'acétylases et de déacétylases. S'il y a cette diversité, c'est qu'il y a une raison.

### La méthylation des histones

Lysines et Arginines sont les cibles. On ne sait pas si la méthylation active ou réprime les gènes. Elle peut faire les 2. Globalement la méthylation des Arg est activatrice. Mais pour les lysines, on ne sait pas. Ça dépend de leur position sur la queue des histones, de leur degré de méthylation (mono, di ou tri-méthylées). On essaie de cartographier ça. C'est le code histone.

### Autres modifications

Déglycosylation ... (voir tableau)

### Fonction des readers

Reader = lecteurs : molécules, enzymes ou cofacteurs qui interviennent dans chacune des étapes décrites avant. Toutes ces protéines partenaires sont à explorer pour définir les conditions exactes dans lesquelles un gène est activé ou inhibé.

Les *modifieurs* de la chromatine sont là pour faire le lien entre 2 partenaires d'intérêt. Toutes les machineries (transcription, réparations, recombinaison, réplication, RNA processing) sont dépendantes de ce qu'on vient de décrire.

## VI Les ARN non-codants

Correspond à 64 % du génome. 1.5% du génome pour le codant.

C'est tout un monde de découverte et d'étude. Ce sont en général de petits arn d'une 30aine de base. Il faut classer tout ça, répertorier leurs fonctions...

On retrouve ces ARN chez les eucaryotes et des micro-organismes arché.

Certains ARN non-codants impliqués dans l'expression de gènes. miRNA : prennent part à la régulation de certains ARNm. Les listes des tableaux proposés sont incomplètes, et les mécanismes sont encore souvent obscures.

Nécessaire de connaître le génome d'intérêt avant de se lancer dans l'analyse épi-génétique. □ Identification des bases de l'adn qui nous intéresse : séquençage (NGS, voir an dernier). Dans un contexte physio ou pathologique d'intérêt, une approche puce également. S'intéresser au transcriptome aussi. RNAseq (séquençage des ARNm). Ce qui est traduit à l'instant t : ribo seq : isole les ARN ribosomiaux et faire du séquençage derrière. On peut également isoler les ARN en cours de synthèse dans le noyau.

- Différentes techniques pour isoler ce qui nous intéresse : on séquence derrière.

Micro-Array pangénomiques : méthode d'hybridation pour identifier la présence de tel ou tel gène ds un organisme. C'est uniquement possible pour des espèces très utilisées en études (car les entreprises de production de puces ne font que celles-ci).

Comment établir le profil de méthylation ds un contexte particulier ?

cytosines méthylées au niveau des îlots CPG -> inactivation des gènes. Quelles sont les cytosines méthylées, ou sont-elles positionnées... -> compréhension des mécanismes génétiques.

Tableau résumant certaines de ces approches technologiques :

- Par **immunoprécipitation** des zones méthylées.
- Propriétés de certaines **enzymes de restriction** à « grignoter » l'adn au niveau de zones non méthylées : on ne garde donc que les méthylées.
- La 1<sup>e</sup> mise en place, celle de référence. Au **bisulfite de sodium** : molécule permettant de convertir cytosine non-méthylée en uracil.

Identifier les cytosines méthylées par immunoprécipitation :

ADN issus du noyau, coupés en fragments par sonication. Puis deux approches :

- on utilise une partie de l'échantillon, on dénature les petits fragments : marqués par un fluoroforme rouge p.ex. C'est le contrôle de la cellule. Pour l'instant, méthylé ou pas, on s'en fiche.
- La deuxième partie de ces échantillons, on les dénature : ils deviennent simple brins. Anticorps reconnaît les cytosines méthylées. On précipite les fragments qui contiennent les cytosines méthylées. On élimine tous les fragments non-méthylés (qui sont censés être actifs). À partir de cette préparation, on la marque avec une cyanine verte p.ex. Hybridation sur puce. Visible en vert -> Égalité d'expression entre le contrôle et la précipitation : ce sont les gènes méthylés. Une différence de couleur : Ce sont les fractions actives, non-méthylées. Une approche séquençage est également possible (à la place de l'approche puce). La technique utilisée dépend du potentiel du labo.

Utiliser des enzymes de restriction

Certaines peuvent cliver au niveau des îlots CPG. C'est là qu'on trouve les cytosines méthylées. BstUI : un exemple d'enzyme de restriction.

Qd les cytosines sont méthylées, elles sont insensibles au clivage par l'enzyme de restriction. On extrait l'adn, on fragmente par sonication -> des fragments doubles brins, méthylés ou non, tout mélangé ds le même tube. On lie des adaptateurs à ces fragments : ce sont des séquences d'ADN connues, qui permettent d'amplifier ces fragments. On fait agir l'enzyme de restriction, en prenant le pari qu'il y a un site de coupure au niveau des îlots CPG. □ Les fragments méthylés vont être coupés : éliminés. On ne garde que ceux qui sont méthylés. Là aussi, on enrichit en fragments méthylés, et on élimine ceux qui ne le sont

pas. On fait alors une PCR : seuls les méthylés sont intacts et ont conservé les deux adaptateurs : Amplification de l'ADN méthylé. On retrouve alors les deux approches dont on a parlé avant (puce ou séquençage).

#### 1<sup>e</sup> méthode mise en place : Bisulfite conversion-based analysis

Le bisulfite de sodium peut convertir les cytosines non-méthylées en uracil. On prend les fragments. bisulfite → toutes les cytosines non-méthylées sont transformées en uracil. Quand on les lit, au lieu d'avoir des G, on va avoir des T. Explication :

Avec des cytosines non-méthylées

CGATTAC

GCTAATG (complémentaire)

Bisulfite agit :

UGATTAU

TGATTAT

Avec des cytosines méthylées : il n'y a pas de changement.

Il faut donc connaître le profil normal de l'ADN pour pouvoir le comparer avec cette phase de lecture.

Celles qui sont restées C sont des cytosines méthylées. Celles qui ont changé sont les non-méthylées.

Amplification PCR → direct sequencing, NGS, PCR quantitative à haute résolution, permettant de déterminer les modifications de bases.

Voilà pour les approches d'identification du profil de méthylation.

Il y a des cytosines méthylées, mais il y a aussi des hydroxy-méthylées (dans le processus de déméthylation, grâce aux enzymes TET), Formyl-méthylées, Carboxy-méthylées. On commence à s'intéresser à ces profils aussi. On pense que c'est encore un autre niveau de signature, un autre niveau de codage. Les manipulations précédentes ne le font pas extrêmement bien (séparer les hydroxy-méthylées des méthylées) car le bisulfite traite indifféremment les différentes formes méthylées. Comment faire ?

1/ lecture de la séquence

2/ profil général de méthylation

3/ cible de hydroxy-méthylées uniquement : profil

#### oxydative bisulfite

Oxydation des groupements hydroxy-méthyl des cytosines grâce au périodate de potassium ( $K_2O_8$ ). Les hydroxy-méthyl sont transformées en cytosine formyl-méthylées. On fait agir le bisulfite. La formyl-cytosine sera lue comme une cytosine normale : transformée en uracil par le bisulfite : on les voit comme des T sur la séquence complémentaire. Les cytosines méthylées sont toujours protégées du bisulfite.

CTC<sub>m</sub>GATC<sub>hm</sub>

GAG C TAG (séquence complémentaire : celle qu'on lit)

KruO<sub>4</sub> :  
CTC<sub>m</sub>GATC<sub>fm</sub>  
GAG CTAG

Bisulfite :  
UTC<sub>m</sub>GATU  
TAG C TAT

#### TET assisted bisulfite

Ce sont des enzymes qui dé-méthylent les cytosines. Une 1<sup>e</sup> enzyme : glucosyl-transférase : greffe un glucose sur les 5-hydroxy-methyl cytosines. Les méthyl-cytosines ne sont pas affectées. Sous cette forme glycosylée, les cytosines ne sont pas reconnues par les TET. On fait agir TET : elles reconnaissent les méthyl-cytosines et les transforment en, carboxy-méthyl : la forme la plus dé-méthylée. On fait agir le bisulfite de sodium : reconnaît cytosines non-méthylées, et aussi les carboxy-méthylées → transforme ces formes en uracil : elles sont lues comme des T. Les seules lues comme des cytosines normales sont celles qui étaient d'abord 5-hydroxy-méthylées.

5hydroxymethyl cytosine : plusieurs transformations pour les isoler spécifiquement :

- anticorps spécifiques des 5hmC. Ok mais des contamination par des 5mC sont récurrentes... Trouver qqchose d'encore plus spécifique.
- bisulfite transforme les 5hmC en 5 methyl sulfonate cytosine. Change pas la lecture : toujours lue en cytosine. Mais on utilise cette propriété en designant des Anticorps qui reconnaissent spécifiquement cette 5-methyl sulfonate cytosine : c'est plus spécifique.

Comme dans toute approche immunoprécipitation, il ya des faux positifs... Autre série de techniques développée.

Greffe de glucose par glucosyl-transférase. On a donc le glucose a la place du groupmt OH.

- La protéine JBP.1 (J-binding protein 1) : reconnaît les 5hmC (partenaire naturelle dans le noyau. Elle reconnaît également la forme glucosylée. En ajoutant cette prot a la forme glucosylée, un mécanisme semblable à l'immunoprécipitation se déclenche → environnement riche en 5hmC glucosylée.
- Plutôt qu'une protéine partenaire, utiliser un complexe streptavidine - biotine. On est toujours avec le glucose qui protège le groupmt. On rajoute donc une biotine sur ce groupement (on oxyde une partie du glucose pour ensuite procéder à la greffe : soit sur un groupement OH, soit sur un groupement N<sub>3</sub>). Utiliser un mécanisme semblable à l'immunoprécipitation : isoler les fragments et les séquencer.

Quelques soit les approches de précipitation, il y a du séquençage au bout, et on compare par rapport au profil normal, ou méthylé de façon général (sans différencier les différentes formes de méthylation).

Ça génère des datas : analyse (voir tableau poly). Analyser les données avec plusieurs logiciels pour voir s'il y a une concordance des résultats.



## VII Approches technologiques à l'étude des modifications épigénétiques

Identification des bases, séquençage NGS

Comment l'ADN est organisé, comment analyser ça.

Déterminer où se situent les nucléosomes. Utiliser des Dnases pour couper l'ADN libre. des Mnases (micrococcal nucléales) : récupérer le nucléosome, qui lui est enroulé, protégé. Donc 2 préparations : 1 riche en nucléosome, 1 riche en ADN libre.

Une fois les histones isolés : dissociation en chauffant : localiser le positionnement des histones, à quelle fraction d'ADN étaient-ils associés, pour une cellule à un instant donné. ADN de cellule cancéreuse : très accessible : libre. Approche grossière, qui donne des infos importantes.

Des approches technologiques permettent de s'intéresser uniquement à la partie accessible :

Formaldéhyde : bloque les interactions ADN-nucléosome : fige le profil à un instant donné. Un peu de Dnase pour séparer les nucléosomes des fragments fixés. Densité différente □ fond du tube : nucléosomes ; surnageant : ADN libre.

D'autres méthodes d'études du nucléosome, plus précises :

Chromosome conformation capture : ADN dans le noyau, sous une certaine conformation. On fige cette organisation, par des molécules chimiques, ou par des UV. Irradiation □ fige les interactions protéines – ADN. Comme avant, on fait des digestions ciblées : identifier et récupérer les gènes associés aux histones. C'est possible grâce à des enzymes de digestion. Ligation : relier les fragments de gènes liés à l'histone.

3C : chromosome conformation capture. PCR quantitative, d'un segment linéaire.

4C : Circular chromosome conformation capture : PCR inverse, puis microarray ou NGS pour caractériser les séquences.

5C : carbon copy chromosome conformation capture : PCR pour ajouter des petits adaptateurs T3 et T7 : amplification PCR : microarray ou NGS.

Hi-C : anticorps anti-histones. chromosome immunoprecipitation.

### Analyse des histones

Isoler les protéines histones...

**Histone chip-seq** : complexes figés par V ou formaldéhyde. anticorps anti-facteur de transcription : récupérer les facteurs qui se sont liés à l'ADN et avoir le profil des gènes régulés par ce facteur. Digérer le nucléosome de la zone de liaison facteur de transcription – ADN.

Ça on l'a vu l'année dernière. Si on peut faire ça avec les facteurs de transcription, on peut aussi faire ça avec les histones. Une petite 20aine

d'anticorps permettent de se lier à des histones avec certains aa à des endroits donnés, et modifiés par méthylation... □ récupérer les ptits fragments d'ADN associés.

**Histone chip-chip** : on utilise les micro-array. Permet comparaison profil sain profil pathologique.

## VIII Les pathologies liées aux modifications épigénétiques

La seule analyse du génome comme on le faisait avant, ne suffit pas à expliquer bcp de pathologies. Liés à des désordres épi-génétiques et pas forcément des mutations dans le génome.

Syndrome de RETT : maladie rare liée au chsm X (plutôt les femmes sont atteintes), maladies neuro-dégénératives. mutation du gène codant pour la MeCP2 : on touche directement à la méthylation de l'ADN. tant qu'on avait pas conscience de tous ces mécanismes, on pouvait pas éclaircir les mécanismes.

ICF syndrome : Défaut de méthylation dans certaines zones : région très accessibles alors qu'elles ne le devraient pas.

Fragile X syndrome : hyper-méthylation d'un gène : défaut d'expression.  
Neuro-dégénérative (atteint les jeunes enfants)

On trouve au moins les causes : les gènes et les phénomènes impliqués.

Beckwith-Wiedemann syndrome : croissance exacerbée du fœtus, dûe à la perte de méthylation de différentes régions de l'insuline growth factor 2.

→ profils d'expression des gènes différents même pour des vrais jumeaux.