

UE5 Génétique  
Pr E. Tournier-Lasserve  
le 01/12/2017 de 13h30 à 15h30  
Ronéotypeuse : Alice de Martin  
Ronéoficheuse : Maud d'Aboville

## **Cours 9 : Génétique médicale et recherche**

*La professeure a accepté de relire la ronéo. Elle a fait un petit speech sur les partiels, je vous l'ai mis à la fin du cours. Bon courage !*

## **Introduction**

### **I- L'identification des gènes**

#### **A. Stratégies utilisées**

1. Approche gène-candidat
2. Approche sans à priori

#### **B. Exemple : le Moyamoya**

1. Présentation de la maladie
2. Facteurs génétiques et environnementaux
3. Étude de l'association Moyamoya / achalasie
  - a. *Analyse de liaison*
  - b. *Séquençage génome entier*

### **II- Les modèles animaux**

#### **A. Stratégies utilisées**

1. Mutation chimique aléatoire / Mutagenèse dirigée
2. Transgénèse insertionnelle classique
3. Ciblage génique par recombinaison homologe dans les cellules ES
4. Méthode CRISPR-Cas9

#### **B. Les différents modèles animaux**

1. Présentations des modèles animaux
2. L'exemple des angiomes caverneux cérébraux

## Introduction

- L'objectif de ce cours est de comprendre l'importance de la recherche en génétique médicale pour améliorer le diagnostic et le traitement des affections génétiques.
- Ce cours est donc une introduction aux méthodes utilisées pour identifier les gènes et les anomalies moléculaires impliquées dans les maladies génétiques mendéliennes aux causes inconnues.
- Nous verrons aussi les méthodes utilisées pour la création de modèles animaux et leur importance pour la compréhension des mécanismes de ces maladies, ainsi que la réalisation d'essais thérapeutiques précliniques.

## I- L'identification des gènes

### A. Stratégies utilisées

#### 1. L'approche « gène candidat »

Cette stratégie est utilisée si un chercheur a une idée « géniale » sur la nature du gène qui pourrait être impliqué dans la maladie étudiée, par exemple :

- si on connaît le **gène responsable de la fonction altérée** dans la maladie
- si un gène est **connu pour être responsable d'une maladie identique dans une autre espèce**.

**On teste alors ce gène chez les malades et on peut ainsi détecter une mutation pathogène.**

Cette situation est très **simple** mais elle est aussi très **rare**.

#### 2. L'approche « sans à priori »

**C'est la situation la plus courante** : on n'a pas d'idée sur la nature du gène muté, ou alors les idées que l'on avait se sont révélées fausses.

On utilise alors les **approches de « génétique inverse »**, c'est à dire sans à priori sur la nature du gène muté.

### B. Exemple : le Moyamoya

#### 1. Présentation de la maladie

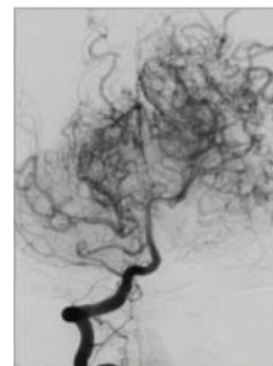
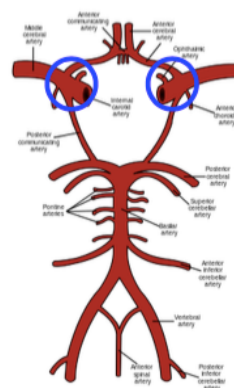
Moyamoya est une **angiopathie cérébrale** caractérisée par :

- **une sténose de la terminaison des artères carotides internes** (moyamoya en japonais)
- **un réseau de néo-vaisseaux anormaux** artériolaires en « volutes de fumée » (*photo*)
- des symptômes tels que des infarctus, des hémorragies cérébrales, des céphalées, de l'épilepsie.

C'est une **maladie rare**, qui est 10 fois plus fréquente au Japon qu'en Europe. Elle peut toucher les individus **dès l'enfance**.

En anapath, on observe :

- **une hyperplasie irrégulière fibrocellulaire de l'intima**



- une accumulation de cellules positives pour le **smooth muscle actin** (marqueur des cellules musculaires lisses)
- **une atrophie et déshabitation de la media**
- absence de cellules inflammatoires, et absence de lésions d'athérosclérose



Les mécanismes physiopathologiques du moyamoya sont pour l'instant inconnus, et il n'y a **pas de modèle animal adéquat**.

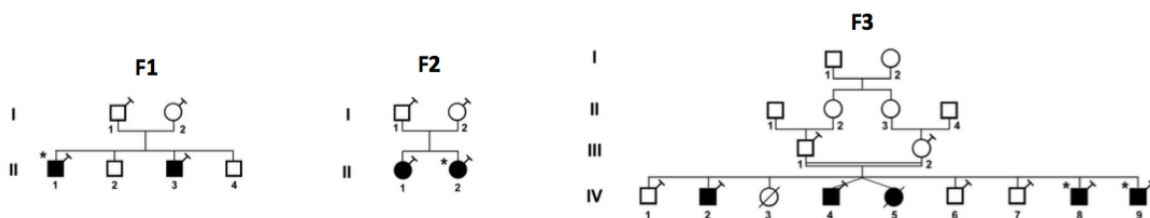
## 2. Facteurs génétiques et environnementaux

Il existe **2 formes** du Moyamoya :

- La maladie de moyamoya (MMD), qui est une **angiopathie cérébrale isolée** (aucune autre atteinte ni cause connue). On retrouve **10 % de cas familiaux**, avec 80 % de concordance chez les jumeaux monozygotes, avec 2 fois plus de femmes que d'hommes touchées. Cette maladie semble être donc **plus polygénique que mendélienne**.
- Moyamoya syndrome (MMS), qui est une **angiopathie moyamoya secondaire à une cause génétique ou acquise connue**, comme :
  - Une irradiation cervico-encéphalique
  - Les maladies mendéliennes avec faible pénétrance de l'angiopathie moyamoya, tels que NF1, la drépanocytose, le syndrome de Noonan ...
  - Les maladies mendéliennes avec forte pénétrance de l'angiopathie moyamoya : exemple de la maladie de moyamoya associée à une achalasie

**Le Moyamoya présente donc une hétérogénéité clinique et génétique.**

## 3. Étude de l'association Moyamoya / achalasie



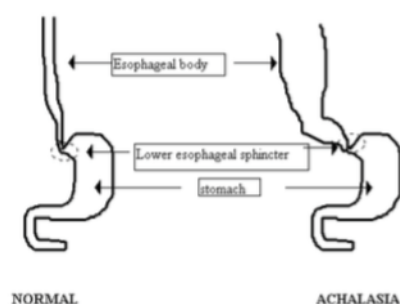
Ci-dessus 3 familles originaires du Maghreb où on retrouve cette association moyamoya et achalasie. Il y a de la **consanguinité** dans les 3 familles.

Le mode de ségrégation et la consanguinité sont en faveur d'une **transmission autosomique récessive**.

L'achalasie est une maladie rare caractérisée par une **absence de péristaltisme du bas œsophage et un défaut du sphincter à se relaxer**.

A l'imagerie on retrouve un **mégaœsophage** et un aspect typique d'achalasie lors de la mesure des pressions dans l'œsophage.

*Le moyamoya a une pénétrance d'environ 50%, contre 100 % pour l'achalasie, avec un début qui peut être très précoce.*



Après avoir posé une hypothèse sur le mode de transmission, il faut maintenant déterminer la stratégie pour trouver le gène responsable de la maladie : Quels sont les gènes candidats ? Doit-on utiliser l'approche « sans à priori » ? Comment identifier et localiser le gène responsable ?

a. analyse de liaison

Il s'agit d'une analyse de liaison entre des **marqueurs polymorphes dont on connaît la localisation chromosomique et le gène que l'on cherche à positionner**. En effet, 2 gènes situés à proximité l'un de l'autre seront transmis ensemble dans la descendance d'un sujet : on dit qu'ils sont **liés**.

On analyse alors la **vraisemblance d'une liaison entre 2 marqueurs** pour une distance théta de recombinaison donnée, exprimée sous la forme d'un **lod-score**.

**Le Lod-score est le log décimal du rapport entre la vraisemblance d'être lié à une valeur de theta donnée ET la vraisemblance de ne pas être lié :**

- Lod-score > 3 : 1000 fois plus de probabilité d'être lié (*1000 chances contre 1*)
- Lod-score < -2 : 100 fois moins de probabilité d'être lié

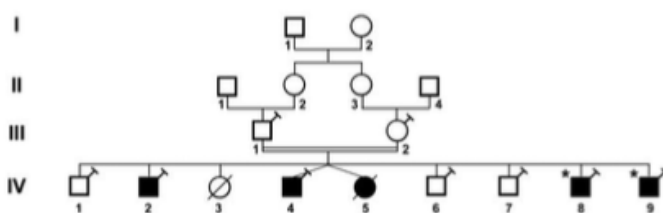
Cette approche permet à la fois d'**identifier les régions potentiellement liées**, mais aussi d'**exclure des régions** dans lesquelles le gène ne peut pas se trouver.

Les conditions pour l'analyse de liaison :

- Il faut disposer de **familles informatives** (nécessité de plusieurs membres par exemple) dans lesquelles le statut (le fait être malade ou sain) de chaque membre a été soigneusement défini.
- Il faut différents types de **marqueurs** et d'**approches technologiques**.

On peut utiliser différents types de marqueurs :

- **Les marqueurs microsatellites**, très informatifs, énormément utilisés jusqu'à ces dernières années.
- Ils sont maintenant remplacés par les **marqueurs de type SNP** (Single Nucleotide Polymorphisms) sur « puces » à ADN : ce sont des **marqueurs bialléliques co-dominants** correspondant à une variation d'une paire de base. Ils sont très nombreux et répartis sur l'ensemble du génome. On utilise la **technologie d'hybridation sur puce ADN** (Puces Affymetrix 250K contenant 250 000 SNP). On utilise une puce par sujet, qui permet le **criblage de l'ensemble de son génome**.



*les individus « informatifs » sont les individus avec une seringue : individus sur lesquels on a des informations concernant la maladie.*

On réalise donc une **analyse de liaison pangénomique** d'une des familles, la famille F3 : cette famille est potentiellement informative (beaucoup de sujets). On utilise les puces Affymetrix 250 K.

On trouve 5 régions suggestives d'une possible liaison avec un **lod-score maximum de 2,18** (ce qui équivaut à 150 chances pour 1).

496 gènes sont localisés dans ces régions : on a ainsi **exclu plus de 90% du génome** grâce à cette analyse.

*Comment identifier le gène causal parmi les 496 gènes localisés dans les régions candidates ?*

b. Séquençage exome entier

Le séquençage exome entier (whole exome ou WES) est une **révolution du séquençage haut débit**.

Il y a 20000 gènes dans le génome humain, l'exome entier représente 50 millions de paires de bases, le génome entier 3 milliards, l'exome entier chez un individu au hasard comporte 35000 variants par rapport à la séquence de référence, le génome entier 4 millions de variants chez chaque individu.

Il faut donc de **multiples filtres pour éliminer les variants non candidats** :

- **mode de transmission** : si l'on étudie une maladie autosomique récessive, les variants doivent être présents à l'état hétérozygote chez les parents sains et homozygote chez les enfants malades.
- **fréquence** : si la maladie est rare, on élimine les variants fréquents
- **nature de la mutation** (faux-sens, non sens...)

Le WES a permis l'identification de nombreux gènes impliqués dans des maladies mendéliennes : en 2012, il a permis l'identification de 100 gènes impliqués dans des maladies mendéliennes.

Il s'agit donc d'une **technique très puissante, mais qui a des limites** :

- Dans une proportion non négligeable de cas, on ne trouve pas le gène causal
- **La couverture est incomplète** : les petites insertions et délétions sont mal détectées
- **Les grands remaniements** (délétion ou duplication de grande taille) et les mutations situées dans **introns** ne sont pas détectées

Retour au cas :

On effectue un **séquençage exome entier des 2 cas index** (= le premier patient) des familles F3 et F2 :

- 1 seul gène était à la fois situé dans les régions candidates identifiées dans la famille F3, et muté dans les 2 familles, F3 et F2
- La mutation est de **type homozygote, conduisant à un codon stop**

*Quelle est l'étape suivante ? Y a-t-il d'autres arguments pour confirmer la nature causale de cette mutation ?*

		exon 8			
	Individu	Type	Position	Mutation	Statut
Famille A	III.1	Intron	1	c.1086+1G>A	het
	III.2	Intron	1	c.1086+1G>A	het
	IV.1	Intron	1	c.1086+1G>A	het
	IV.2	Intron	1	c.1086+1G>A	hom
	IV.3	Intron	1	c.1086+1G>A	hom
	IV.4	Intron	1	c.1086+1G>A	het
	IV.5	Intron	1	c.1086+1G>A	het
	IV.6	Intron	1	c.1086+1G>A	hom
	IV.7	Intron	1	c.1086+1G>A	hom
		exon 8			
	Individu	Type	Position	Mutation	Statut
Famille B	I.1	Nonsense	-	c.1045C>T	het
	I.2	Nonsense	-	c.1045C>T	het
	II.1	Nonsense	-	c.1045C>T	hom
	II.2	Nonsense	-	c.1045C>T	hom

On recherche une **co-ségrégation des variants avec les autres malades** dans F3 et F2, et on effectue un séquençage du gène dans F1 par Sanger. (tableau à droite)

On retrouve bien une mutation homozygote dans le même gène.

		exon 9			
	Individu	Type	Position	Mutation	Statut
Famille C	I.1	Frame Shift	-	c.1170delA	het
	I.2	Frame Shift	-	c.1170delA	het
	II.1	Frame Shift	-	c.1170delA	hom
	II.2	Frame Shift	-	c.1170delA	hom

**Les mutations de ce gène sont donc bien responsables du moyamoya et de l'achalasie.**

**Le gène muté GUCY1A3 code pour la sous unité alpha 1 de la guanylate cyclase soluble sCG** (récepteur principal du NO). C'est un acteur majeur de la relaxation des cellules musculaires lisses dans les vaisseaux et le tractus digestif (motilité).

Ce gène n'est cependant pas impliqué dans les autres formes « récessives » de la maladie de moyamoya (MMD). En effet un screening par Sanger de la totalité du gène dans 3 familles atteintes de MMD n'a révélé aucun variant délétère.

Applications de l'identification du gène :

- Cela a des **applications cliniques immédiates par l'arrêt de l'errance diagnostique** : On peut ainsi effectuer des tests diagnostiques, des conseils génétiques, des diagnostics prénataux...
- L'identification d'un gène ouvre la possibilité de **création d'un modèle animal**.
- On accède aussi à une **meilleure compréhension des mécanismes physiopathologiques**, ce qui est indispensable pour le développement d'approches thérapeutiques.

## II- Les modèles animaux

*Pourquoi utiliser des modèles animaux ?*

- On peut **valider d'une hypothèse** (gène candidat ou mutation candidate) grâce à un modèle animal
- On peut aussi **étudier la fonction normale d'un gène**
- Ce sont des modèles expérimentaux des maladies humaines, qui permettent une **meilleure compréhension des mécanismes physiopathologiques**, et qui permettent d'établir des **stratégies thérapeutiques** (élaboration et validation pré-clinique).

### A. Stratégies utilisées

#### 1. Mutation chimique aléatoire / mutagenèse dirigée

Il s'agit de « bombarder » d'**agents mutagènes, qui créent des mutations rapides au hasard**. *C'est très efficace mais ça crée trop de variants différents, il faut ensuite les trier, puis les recroiser avec des sujets sauvages : technique compliquée.*

Une collection de mutants est obtenue, on fait un **phénotypage de ces mutants** puis on identifie le gène muté (très long).

Exemples : Souris Apc<sup>Min</sup> (cancer intestin), souris Shaker (Myosin VIIA > surdité) ...

Exemple d'une colonie de souris ENU avec hémorragies cérébrales périnatales et porencéphalie :

**L'agent mutagène ENU** (Ethyl-N-Nitroso urea) est un agent alkylant qui entraîne des mutations ponctuelles dans l'ADN des spermatogonies. Il crée des mutations dans le gène *Col4a1*.

Les souris mutées présentent des **hémorragies cérébrales périnatales et une porencéphalie**.

Utilisation d'un modèle murin pour comprendre les mécanismes des collagénopathies *Col4a1* :

- Cela permet de décrire **l'histoire naturelle des lésions** avant même qu'elles ne se soient manifestées cliniquement chez la souris (stades précoces de la maladie), et d'**avoir accès aux tissus atteints** (tissus vasculaires, cerveau...).
- Ouvre la **possibilité d'analyses** histologiques, transcriptomiques, protéomiques etc. sur ces tissus qui sont très difficiles d'accès chez les patients.

#### 2. Transgénèse insertionnelle classique

Son but est l'**addition d'un gène donné afin de surexprimer ce gène**, avec ou sans modification génétique.

On effectue une **micro-injection d'un transgène dans des ovocytes fécondés réimplantés chez une souris**.

schéma :

on injecte l'ADN dans le plus gros pronucléus (en général le mâle). Mais l'ADN ne va pas forcément s'exprimer dans toutes les cellules.

=> on doit donc croiser les souris pour que toutes les cellules soient porteuses de la mutation.

Avantages :

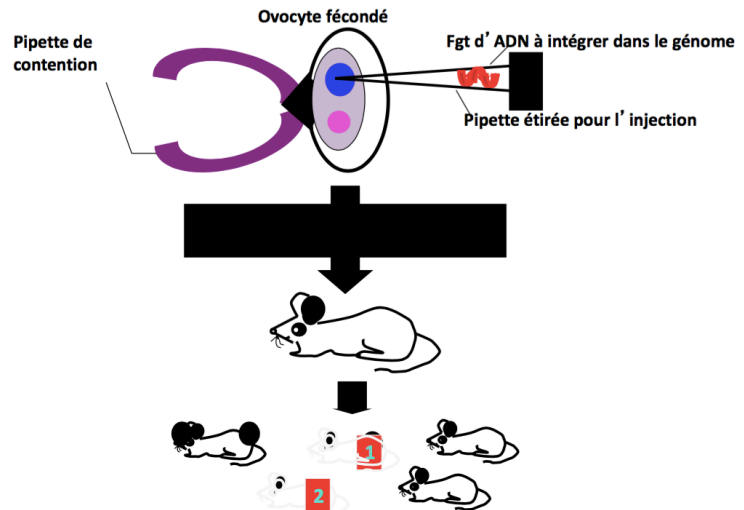
**Toute séquence d'ADN peut être injectée** (peut être un promoteur avec une séquence codante et un signal polyA...)

Inconvénients :

L'intégration du transgène est **rare** (3-10% selon la qualité de l'ADN et de l'expérimentateur), **aléatoire** dans le génome (site le plus souvent unique), et s'effectue en **multiples copies** (nombre variable, ce qui conditionne le niveau d'expression).

Elle peut aussi se faire dans un **gène essentiel**.

La première souris transgénique date de 1980/81 (micro-injection d'ADN dans le pronucléus du zygote).



### 3. Ciblage génique et recombinaison homologue

Le but de cette technique est **d'intégrer un gène donné au sein ou à la place de son homologue génomique chez la souris**. La recombinaison homologue permet d'échanger ou d'insérer une séquence d'ADN en un endroit précis du génome (mais événement très rare).

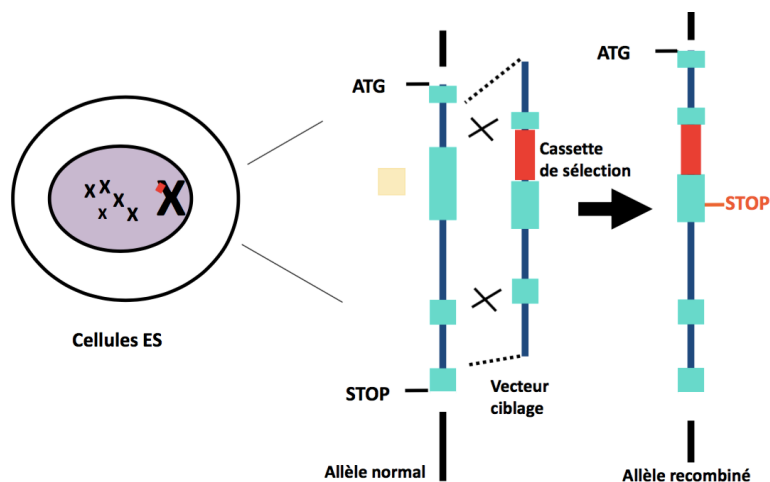
On peut ainsi :

- « invalider » un gène (**knock-out**)
- introduire une mutation dans un gène de la souris (**knock-in**)

Cette technique permet l'étude de la fonction d'un gène, la modélisation d'une maladie récessive ou d'une perte de fonction.

Principe :

On introduit une **modification génétique dans le génome de cellules totipotentes** (cellules ES) par recombinaison homologue et on réimplante des cellules dans un blastocyste de souris.



Histoire : (non cités par la prof)

- 1981: établissement de lignées de CS embryonnaires
- 1989: 1ères souris mutantes
- 2007: Prix Nobel attribué à MR Capecchi, M Evans, O. Smithies



#### 4. Méthode CRISPR-Cas9

Cette méthode est une révolution technologique, permettant un **temps d'analyse plus court**.

La Cas9 (CRISPR associated protein 9) est une **endonucléase** capable de couper chacun des 2 brins d'un ADN, permettant ainsi d'**éditer le génome** (remplacement d'un gène par un autre).

Elle est composée d'**ARN guide constant**, attaché à une séquence spécifique de l'ADN à cibler, et d'une **protéine Cas9** qui accueille cet ARN, guide le positionnement de cet ARN sur l'ADN-cible et découpe les deux brins d'ADN à la position correspondante.

#### B. Les différents modèles animaux

##### 1. Présentation des modèles animaux

- La souris :
  - **Leur physiologie est assez proche de celle de l'homme**
  - **L'élevage est facile**, le temps de génération est court (~ 21 jours)
  - Elles permettent une **connaissance complète du génome**
  - La manipulation de son génome est techniquement accessible, et les lignées sont « pures »
  
- La levure :
  - exemple : *Saccharomyces cerevisiae* (12 Mb, 6000 gènes)
  - permet l'étude de la **fonction de la frataxine** (ataxie de Friedreich)
  
- Le ver nématode :
  - exemple : *Caenorhabditis elegans* (100 Mb, 19.000 gènes)
  - modèles de **myopathie**
  
- Le poisson :
  - exemple : le **poisson zèbre** (embryons transparents : on peut utiliser des molécules fluorescentes observables à travers leur corps transparent) (morpholinos)
  - modèles de **cardiomyopathie**
  
- L'insecte :
  - exemple : la mouche du vinaigre (*Drosophila melanogaster*, 137 Mb, 14.000 gènes)
  - modèles de **maladie neurodégénérative**
  
- Le rat : transgénique et KO (depuis 2010)
  
- Les gros animaux (porc) : transgénique et plus récemment ciblage génique (*très chers, réservés à des recherches spécifiques*)

#### Quels modèles animaux utiliser pour l'étude d'une maladie génétique humaine ?

Les possibilités sont extrêmement diverses (technologie, espèce) : le design du modèle animal va être **conditionné par la maladie** (mode de transmission et nature des mutations) **et par la question posée**.

Il est très rare d'avoir un modèle animal récapitulant la maladie humaine à l'identique, d'où l'intérêt de **l'analyse combinée de différents modèles**.

Les modèles animaux sont bien sûr incontournables.

## 2. L'exemple des angiomes caverneux cérébraux

80% des cas sont sporadiques et peu symptomatiques (une seule lésion), 20 % des cas sont familiaux selon une transmission autosomique dominante, et symptomatiques dans 50 à 70 % des cas.

3 gènes ont été identifiés (MGC4607 et KRIT1 sur le Chr 7, PDCD10 sur le Chr 3).

Toutes les mutations des patients atteints (quelque soit le gène muté) aboutissent à un **codon STOP prématuré**, conduisant à des **mutations perte de fonction**.

On remarque qu'il n'y a aucune homologie de séquence entre les 3 protéines.

*Pourquoi faire des modèles animaux d'angiome caverneux ?*

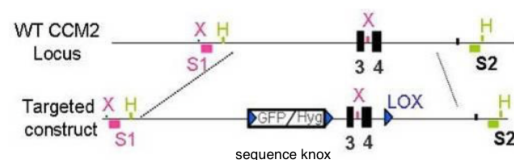
Pour comprendre le **rôle des gènes CCM dans l'angiogénèse cérébrale normale**, et comprendre comment la perte de ces gènes entraîne l'apparition d'angiomes.

On veut savoir quels types cellulaires requièrent la présence des protéines CCM pour un développement vasculaire normal au sein du SNC.

Il faut donc disposer de matériel biologique permettant l'**analyse de liaison à un stade très précoce** (ce qui n'est pas possible chez l'homme).

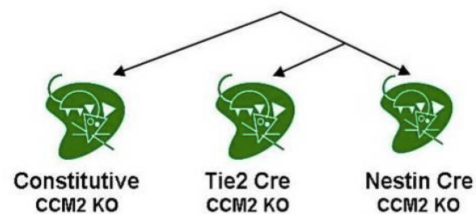
Stratégie utilisée pour invalider le gène CCM2 chez la souris :

On utilise la **recombinaison homologue en floxant les souris de 2 sites LOX autour du gène CCM2 et en les croisant avec des souris exprimant la recombinaise Cre** dans tous les tissus ou dans des tissus spécifiques.



On réalise 3 modèles de souris KO pour le gène CCM2 afin de comprendre la maladie :

- Un KO du gène dans **tous les tissus** de la souris
- Un KO du gène « tissu spécifique » au niveau des **cellules endothéliales**
- Un KO du gène « tissu spécifique » au niveau des **précurseurs des neurogliaux**



Résultats :

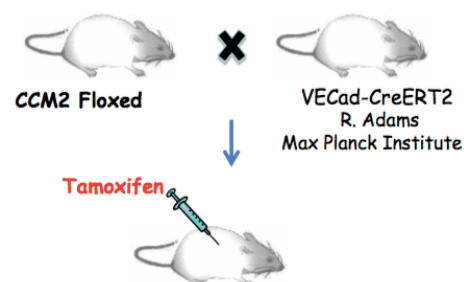
La délétion de CCM2 dans les **cellules endothéliales entraîne la mort précoce des embryons**, tandis que la délétion du gène dans les précurseurs neurogliaux n'entraîne pas de phénotype.

**CCM2 a donc un rôle essentiel dans l'endothélium vasculaire E10.**

La mort précoce des embryons ne permet pas d'analyser le développement des vaisseaux cérébraux.

*Nouvelle question: comment obtenir un modèle pertinent survivant après la naissance nous permettant l'analyse du développement de ces vaisseaux ?*

On effectue une délétion **spécifique et inducible** du gène CCM2 dans l'endothélium vasculaire, grâce à des **promoteurs inducibles** qui vont contrôler l'expression de Cre à différents moments : période néonatale, au moment du sevrage ou pendant l'embryogenèse.



**On invalide le gène par injection de tamoxifène.**

## **L'invalidation post-natale du gène dans les cellules endothéliales récapitule la maladie humaine.**

Remarque : Les lésions se développent dans les veines et pas dans les artères.

Cette technique permet le développement d'un modèle iCCM2 inductible et tissu spécifique.

On a ainsi créé un modèle robuste et pertinent d'angiomes caverneux, donnant des lésions très similaires aux lésions humaines, avec une pénétrance complète donnant des lésions complètes et un développement très rapide de la maladie : **grand potentiel pour d'éventuels essais thérapeutiques précliniques.**

Cette technique a mis en évidence qu'il y a une **restriction des lésions à certains lits vasculaires : lits veineux, rétine et SNC**. On a donc une altération de la morphogénèse du tube endothéliale dans les lésions CCM.

**Les effets de l'invalidation sont conditionnés par le « timing » de l'invalidation.**

Il reste toujours des questions, qui sont au coeur des travaux en cours et futurs, notamment sur les mécanismes impliqués dans la formation des lésions :

- Pourquoi seul le compartiment veineux est touché ?
- Quels sont les mécanismes conduisant à une morphogénèse altérée ?
- Quels sont les mécanismes expliquant « la fenêtre de temps » ?
- Quels sont les « Pathways impliqués »? (Wnt/Beta Cat pathway, Planar Cell Polarity pathway ...)

Le développement d'essais thérapeutiques précliniques sur l'angiome caverneux cérébral a pour objectifs d'empêcher le développement de nouvelles lésions et de favoriser la régression des lésions existantes. Il faut notamment analyser le nombre et la taille des lésions en histologie.

2 approches pour **identifier les agents thérapeutiques à tester** :

- approche basée sur l'**identification de cibles thérapeutiques candidates**
- approche de **criblage haut débit de composés sans à priori** sur des voies candidates

-----

questions tombables :

- *Le cours vise surtout à nous apprendre à se poser les bonnes questions et à comprendre les stratégies utilisées en recherche pour répondre aux questions d'examen (en bref pas tellement tombable mais méthodo à connaître).*

Partiels :

- *Tout l'examen est sous forme de QR*
  1. *!!! prenez le temps de bien lire les questions, ex : quels sont les modes de transmission possibles ? il faut tous les donner, pas seulement le plus probable*
  2. *bien réfléchir; ne pas partir sur une hypothèse*
- *Il y aura des cas de génétique clinique et moléculaire*
- *Toujours 3 questions : lire les 3 et commencez par la plus facile (il y a toujours une question plus facile)*
- *!! revoir bien les EDs pour les partiels*
- *Pas de questions pièges*

