

UE 11 : Principes de biologie Moléculaire et applications en génétique médicale – Fiche 2

RAPPEL PLAN

Cours 1 :

- I- Analyse du matériel génétique
 - 1- Extraction de l'ADN
 - 2- Extraction de l'ARN
 - 3- PCR
 - 4- Digestion enzymatique
- II- Biologie moléculaire et génétique médicale
- L'achondroplasie

Cours 2 (partie 1 = sans le NGS !):

III. Applications en génétique médicale

Partie 1 : Le syndrome de Wolfram

- A. 1ère étape : Exploration de WFS1
- B. Séquençage de l'ARNm
- C. Clonage Moléculaire
- D. Transformation : introduction du vecteur dans une bactérie
- E. Sélection, Isolement, amplification, extraction
- F. Carte de restriction
- G. Séquençage des inserts

Partie 2 : Autres techniques de Biomol

- A. Clonage d'expression
- B. PCR en temps réel

IV. Applications en génétique médicale

Partie 1 : Le syndrome de Wolfram

Rappel :

- Dans le noyau, les gènes sont transcrits en ARN (exons + introns)
- Les préARNm sont ensuite maturés en ARNm après épissage (exons)
- Les ARNm sont ensuite exportés dans le cytoplasme
- Dans le cytoplasme, les ARNm sont traduits en protéines

Il y'a des parties non codantes en amont (5'UTR), en aval (3'UTR) et les introns : on ne connaît pas leur fonction mais leur délétion entraîne des dégradations des ARNm (queue poly-A en 3'UTR qui stabilise les ARNm).

Même si on connaît la séquence entière du génome humain on ne connaît pas tous les gènes ni leurs fonctions.

Savoir traduire un ARNm en protéine n'implique pas connaître la fonction de celle-ci.

➔ Difficile de savoir si un variant identifié est bien responsable de la pathologie (certains variant ne sont pas pathogènes)

Vous savez désormais comment détecter une mutation ciblée et connue : PCR-RFLP, puis confirmation avec un séquençage (ex : Achondroplasie)

Que faire dans la situation, plus compliquée, où l'on doit rechercher plusieurs mutations dans un même gène ? Ou plusieurs gènes comme les maladies multigéniques ? Ou encore quand le séquençage est ininterprétable ou ne donne pas assez d'informations, comme dans le cas du Syndrome de Wolfram ?

C'est l'objet de la suite du cours :

Pour poser un diagnostic, on ne séquence que les séquences codantes (=exome)

On ne peut pas interpréter les mutations des **régions non codantes** car, ne donnant normalement pas d'acide aminé, on ne pourra pas faire le lien entre modification de l'ADN et altération des protéines. Or, même si on ne les cherche pas, des **mutations** dans ces régions peuvent avoir des **effets délétères** par ex en modifiant les séquences consensus d'épissage, en créant de **nouveau site accepteur d'épissage (=site cryptique d'épissage AG)** ... Si le préARNm est épissé différemment, l'ARNm est différent et la protéine aussi : elle peut être défectueuse !

Maintenant, mettez-vous dans la peau d'un médecin généticien. Un patient enfant (qu'on appellera Jimmy) arrive dans votre cabinet, il vous a été adressé par son médecin généraliste qui suspecte, après examen clinique, un Syndrome de Wolfram. On ne peut pas se contenter de signes cliniques il va falloir faire des examens complémentaires, ici en génétique. On va se préoccuper en particulier de WFS1 (gène codant, ARNm, protéine finale) car on sait que ce Syndrome y est associé.

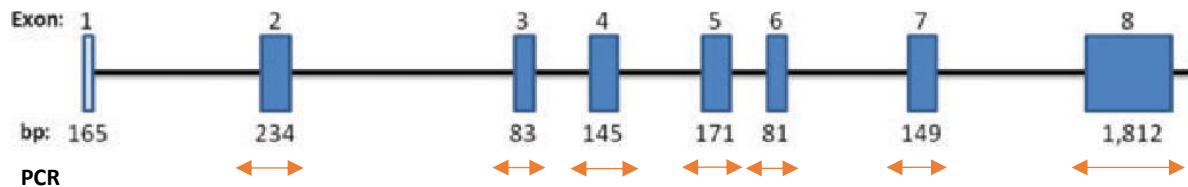
Syndrome de Wolfram :

- Diabète, atrophie optique, surdité, troubles neurologiques variables
- Pathologie **autosomique RECESSIVE ++**
- Gène responsable : **WFS1**. Petit gène de 8 exons, dont le 1^{er} est non codant (ATG se situe sur le 2^{ème})
- Code pour la **wolframine** protéine dont la fonction est inconnue. Elle aurait un rôle dans le flux calcique permettant la communication entre le Réticulum endoplasmique et la mitochondrie.

A. 1^{ère} étape : Exploration de WFS1

On réalise un séquençage sanguin en s'intéressant aux exons codants.

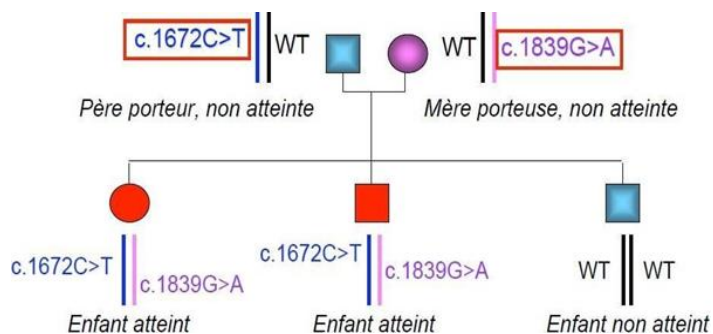
- 1) On va faire 7 amplifications par PCR des 7 exons codants (pas le 1 du coup) et de leurs jonctions avec les introns (5 nucléotides en amont et en aval de l'exons = site où se passe l'excision des introns). On place nos primers de part et d'autre des exons, on séquence l'exon et les 5 premiers/derniers nucléotides des introns : ce sont des sites d'épissage.



- 2) On vérifie ces 7 PCR sur gel et on contrôle les négatifs.
- 3) On les séquences avec la technique des didésoxyribonucléotides

a) Première situation : mutation dans l'ADN codant

On trouve à Jimmy un variant hétérozygote au niveau du 1672^{ème} nucléotide situé sur une séquence codante (mutation : substitution C->T notée c.1672 C>T) et sur l'autre allèle un variant c.1832 G>A



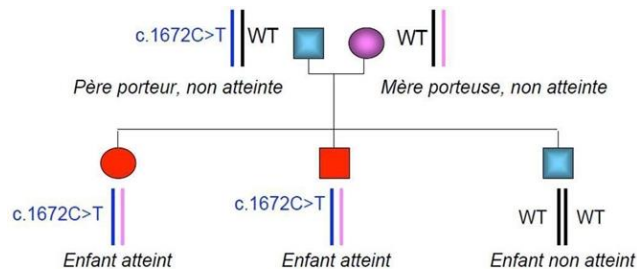
Explication : Le père et la mère sont **hétérozygotes mutés donc porteurs sains** (rappel : maladie récessive ++) mais n'ont pas le même variant (la même mutation).

Ici, l'enfant a hérité du variant 1672 du père, 1832 de la mère. Il n'est **pas homozygote pour la même mutation** mais **malade quand même** parce que ses 2 allèles codent pour une wolframine défectueuse. C'est un cas **d'hétérozygotie composite**.

- ✓ Il a tous les **signes cliniques**
- ✓ La littérature (=articles, études, docs scientifiques sur le sujet) possède les **preuves de la pathogénicité de ces variant** : ces mutations sur ces nucléotides sont déjà connues dans d'autres cas de patients. Et on avait déjà fait des tests montrant l'impact de cette mutation sur la protéine finale (cf. *vecteur d'expression plus loin dans le cours*)
- ✓ On a pu **justifier la transmission de la mutation** (on a trouvé l'allèle paternelle défectueux en séquençant le génome paternelle, pareil pour l'allèle maternelle)

➔ On peut poser le **diagnostic**

b) Deuxième situation : mutation dans l'ADN non codant



C'est moins « facile », si l'enfant à tous les **signes cliniques** mais après séquençage des 7 exons on a identifié le **variant du père**, on a un allèle muté c.1672C>T mais **rien sur l'allèle maternel**, l'allèle **semble être Wild Type** (=non muté=sauvage). Hors si elle c'était le cas l'enfant devrait être porteur sain (car c'est autosomique récessif la prof insiste) mais ce n'est pas le cas.

Hypothèses d'explication :

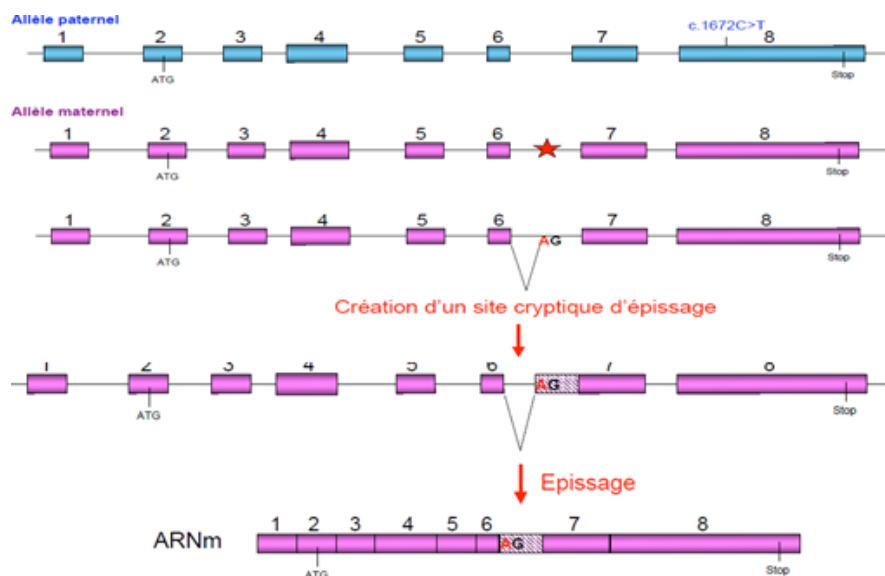
Si nous n'avons pas trouvé de mutation dans les séquences codantes il y'en a peut-être dans les séquences non codantes, les introns, que nous n'avons pas séquencés.

- ❖ Il y'a eu une **modification au niveau des sites d'épissage** entrainant la transcription d'un ARNm défailant donc une protéine défailante. On retombe sur une sorte d'hétérozygotie composite (2 mutations différentes mais pathologique au final).

Exemple : il suffit qu'il y ait un AG qui se crée par mutation, substitution, en plein milieu de l'intron ça va générer un **site cryptique d'épissage**. L'épissage ne se fera plus de la fin de l'exon 6 au début du 7 mais entre la fin du 6 et cet AG en plein milieu de l'intron : on gardera le bout entre le AG et l'exon 7, bout d'intron normalement non codant. Ce **bout d'intron** serait **traduit** et engendrerai une **protéine anormale**.

- ❖ Une **mutation** peut aussi faire « **sauter** » un **exon**. Par exemple on peut avoir un épissage de la fin de l'exon 6 au début de l'exon 8 (l'exon 7 est éjecté) ou même parfois du 4 au 7.

Ces zones codantes sont très peu prévisibles, on interprète mal leurs variations. De plus on a une base de données de cas mais les cas possibles sont trop nombreux.



Comment faire dans ce cas-là ?

On ne peut pas donner de diagnostic sans avoir clairement identifié les mutations chez le père ET chez la mère !

Important : Ici on a commencé par séquencer les exons : on fait à chaque fois cette étape car c'est :

-plus simple/rapide

-les exons sont plus courts que les introns

On va chercher plus loin. Nos hypothèses impliquent une modification génétique qui n'est pas analysable facilement au niveau de la séquence du gène. En effet, on pourrait séquencer tout le génome et non l'exome uniquement mais se serait **trop long** et même si on trouve un variant **on ne saurait pas ce qu'il implique sur la future protéine**, si on est « sur » qu'il soit pathogène. Par contre on peut observer une possible modification au niveau de l'ARNm reflet ++ de la protéine et donc meilleur reflet de l'activité cellulaire (la prof insiste++). On va donc séquencer l'ARNm.

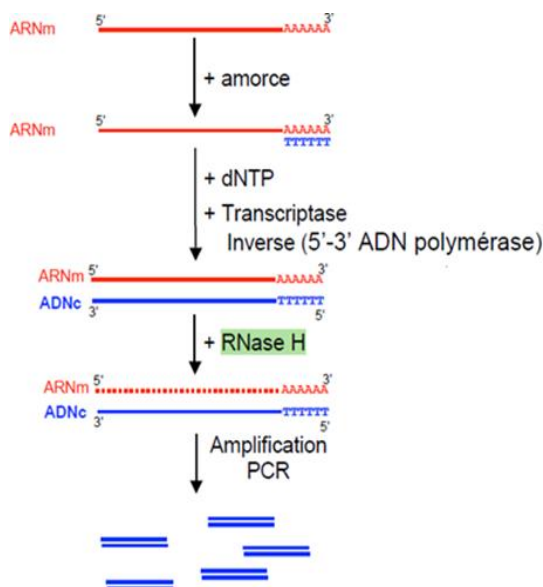
B. Séquençage de l'ARNm

1) Obtention de l'ARNm (cf.cours 1)

Même principe que l'ADN : on lyse, on récupère les ARNm du cytoplasme dans du **phénol à pH acide** (attention ! différence avec l'ADN). On isole les ARNm en utilisant la complémentarité de leur queue poly-A, caractéristique du type messager, avec un support solide d'oligoDT.

2) Transcription inverse

La Taq-polymérase est une ADN polymérase, elle a besoin d'une matrice d'ADN pour fonctionner. On va donc vouloir transformer notre brin d'ARNm en brin d'ADN. Comment ?



On ajoute à notre ARNm :

- ✓ une **amorce d'ADN** (une queue oligoDT)
- ✓ des **dNTP**
- ✓ la **TRANSCRIPTASE INVERSE** (=reverse transcriptase #coucouLesTéломèresde laBiomol 😊)

🔍 Transcriptase inverse :

5'-3' ADN polymérase **d'origine virale** qui synthétise un ADNc (ADN complémentaire) à partir d'une amorce d'ADN hybridée sur un ARNm.

NTB : Enzyme de restriction + Taq polymérase = origine bactérienne vs Transcriptase inverse = origine virale

On obtient un **hybride (=hétéroduplex) ARN/ ADNc**. Enfin on ajoute des **RNase H** (enzyme dégradant l'ARN uniquement s'il est hybridé à un ADN).

3) Amplification par PCR

Maintenant qu'on a un brin d'ADN simple on amplifie directement par PCR (pas besoin d'étape de dénaturation on a déjà du simple brin). Ici on suspectait qu'un fragment de l'intron 6 (entre exon 6 et 7) avait été conservé. On peut mettre des primers partout et recommencer nos 7 PCR, mais quand on « sait » comme ici (peut-être par la littérature ?) on va mettre 2 primers : **un au milieu de l'exon 6 et un au milieu de l'exon 7**. La PCR amplifiera les exons **et ce qu'il y a entre les deux**.

+ NTB : On va rajouter après notre amorce une certaine séquence, reconnaissable par une enzyme, pour qu'elle soit répliquée et ajoutée à nos amplicons (elle nous servira plus tard).

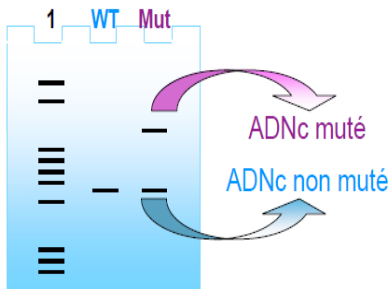
4) Vérification sur gel analytique

On compare la masse des deux **amplicons** (les 2 allèles de l'enfant : paternel et maternel) par électrophorèse. S'ils sont de **même taille** (même **nombre** de nucléotides et pas **type** de nucléotides : une substitution ne change pas significativement le poids du brin) ils migrent à la **même hauteur**. Ici, selon notre hypothèse, le variant d'épissage rajoutant une séquence codante, le produit PCR d'**ADNc maternel** devrait être **plus lourd** et donc **migrer moins loin** que l'ADNc du père.

NTB : on observe ça si le variant se trouve bien entre exon 6 et 7 si on n'observe rien on réessaie avec d'autres exons.

Résultat :

Analyse des produits PCR
après migration électrophorétique

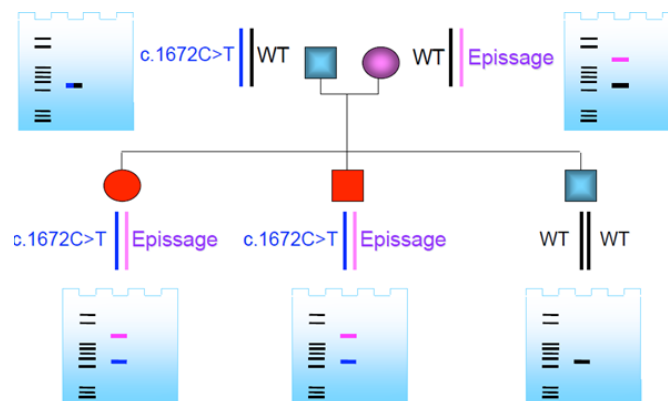


Piste WT (Wild Type) : correspondrait à la situation du **père**. La **mutation ponctuelle** 1672C>T au niveau de son exon 8 ne **modifie pas la taille** du fragment PCR car il s'agit juste d'une **substitution** de base.

Allèle maternel (Mut) : correspondrait à la situation de la **mère** et des deux **enfants atteints** (le brin muté est plus lourd et migre moins loin).

La migration électrophorétique nous permet de vérifier la présence d'un variant d'épissage qui rend l'ADNc plus lourd. Notre **hypothèse** semble donc être la bonne : il y'a bien un **variant d'épissage entre les exons 6 et 7 de l'ADNc chez la mère**. On va pouvoir faire une réaction de **séquence** des produits PCR obtenus précédemment pour **identifier ce variant d'épissage** !

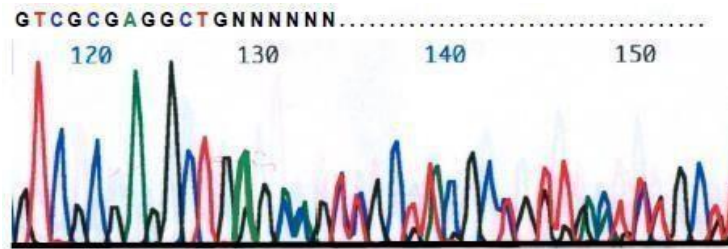
On est **obligés** de trouver précisément **les deux mutations** (via le séquençage) pour avoir le droit de poser un diagnostic car le syndrome est de **transmission récessive** et implique que **deux allèles soient mutés** !



5) Séquençage et identification de la mutation

Quand on séquence le gène de la Wolframine chez Jimmy, on n'a pas d'autre choix que de séquencer nos **deux allèles simultanément** car les produits PCR sont mélangés ! En effet chez les parents comme les enfants quand on extrait l'ARNm on obtient les deux allèles car ceux-ci sont ensemble en suspension dans le cytoplasme. Ici on aura l'allèle paternel et maternel (plus long) dans le même tube.

Résultat du séquençage :



A partir de la jonction exon 6 / 7 (~ position 130) la séquence devient illisible, on a une superposition de deux signaux.

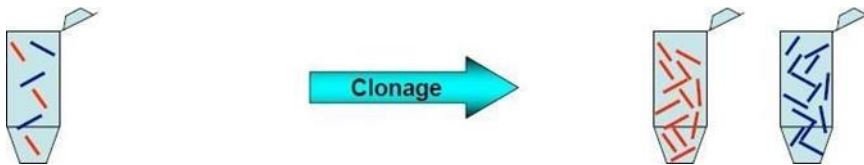
Ceci est dû au fait qu'on a un mélange de nos deux produits PCR de taille différente (à cause du variant d'épissage qui est plus long et qui décale la suite de la séquence). Ainsi, on a une superposition de la séquence de l'allèle paternel et de l'allèle maternel.

➔ Il faut **séparer les deux PCR** (1^{er} étape) puis les séquencer individuellement (2^{ème} étape). Pour cela on fait un **clonage moléculaire**.

C. Clonage Moléculaire

a) Principe

But = obtenir un **grand nombre de copies identiques et pures** d'une **séquence donnée d'ADN**



4étapes :

1. On crée un **ADN recombinant (=vecteur +insert)**.
 2. On introduit ces vecteurs dans la bactérie = **transformation bactérienne**
 3. On **cultive** les bactéries sur une boîte de Pétri (**amplification**) + **antibio (sélection)**
 4. On « **repique** » : on prélève des bactéries de chaque colonie pure et on les met chacune dans des **tubes différents** (on a **isolé** les 2 populations) on les fait proliférer de nouveau
- ➔ On **récupère des fragments purs en grande quantité**

Pour ce faire, on va utiliser des **vecteurs** (ADN circulaire double brin) dans lesquels on va insérer à chaque fois **un seul fragment d'ADN** appelé **insert** (notre produit PCR d'origine paternelle **ou** maternelle)

b) Les vecteurs

Il doit être :

- De **l'ADN circulaire double brin** / plasmidique
- De taille réduite pour avoir la place d'y insérer un fragment d'ADN étranger
- Avoir une **réplication épisomiale** (autonome indépendante du cycle du génome nucléaire de la cellule hôte)
- Posséder des **gènes de sélection** (généralement gène de résistance à l'Ampicilline = antibio), afin de trier les bactéries qui auront vraiment intégré le vecteur

2 types de vecteurs :

Vecteurs de clonage : pour isoler et amplifier un fragment d'ADN (ce qu'on utilise ici)

Vecteurs d'expression : pour transférer un gène dans une cellule eucaryote afin que la cellule exprime ces protéines, on pourra observer si la mutation qui influe la position d'une protéine dans la cellule, etc...

Le clonage moléculaire ne se fait qu'avec des cellules procaryotes (bactéries) !

Il y a **différents vecteurs de clonage selon la taille du fragment à cloner**, au-delà d'une certaine taille on n'utilise plus des bactéries mais des levures (YAC). Dans ce cours on s'intéressera uniquement au PLASMIDE.

Remarque : Les bactéries contiennent un unique chromosome libre dans le cytoplasme ET des petites molécules d'ADN circulaires appelées plasmides.

La séquence du plasmide est connue, ils sont commercialisés, on a :

- ✓ Un **polylinker (site multiple de clonage)** : courte **séquence connue** où il y a de nombreux **sites pour des enzymes de restriction** qui seront digérés afin d'introduire l'insert.
- ✓ Une **origine de réplication** : permet une réplication indépendante du chromosome de la cellule hôte
- ✓ Un **gène de sélection** : Il en existe plusieurs, mais le plus courant apporte une résistance à l'Ampicilline (antibio), donc si on étale les bactéries sur une boîte de Pétri avec de l'Ampicilline, seules celles avec le gène et donc le plasmide (donc intéressantes) vont se multiplier : on donne un avantage sélectif aux bactéries avec plasmide.

+(récap de la prof)+ Plasmide = polylinker + origine de Réplication + gène de sélection

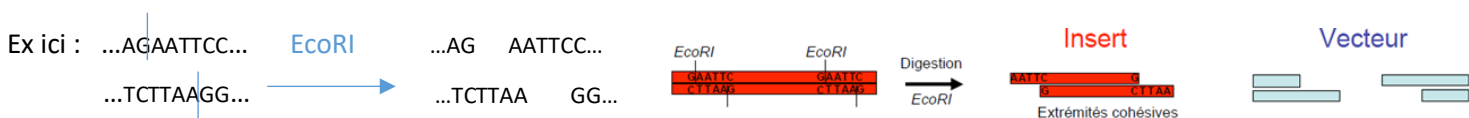
c) Création de l'ADN Recombinant

1. Linéarisation du plasmide, digestion du plasmide et de l'insert

On va essayer de digérer l'insert et le plasmide avec la même Enzyme de restriction. On la choisit en fonction de la manière dont on veut couper nos séquences il existe 3 **stratégies de clonage** : extrémités **franches**, **cohésives** ou **combinés**.

✚ Digestion du plasmide :

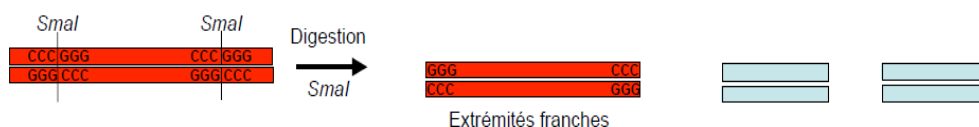
- Création **d'extrémités cohésives** uniquement : la meilleure solution (ex : enzyme EcoRI) car les bases complémentaires entre l'insert et le plasmide pourront s'hybrider spontanément par liaisons hydrogènes



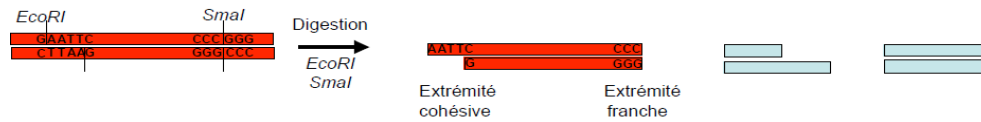
On a bien notre séquence palindromique, l'enzyme pour travail de **reconnaître** son site de digestion et de **couper** entre le G et la séquence AATTC au niveau des **2 brins** (++) diffère avec une endonucléase.

➔ Ici on choisit cette solution

- **Extrémités franches** uniquement (ex : SmaI) -> pas top, car le vecteur a tendance à se refermer sur lui-même



- Les deux à la fois : **on peut les combiner**



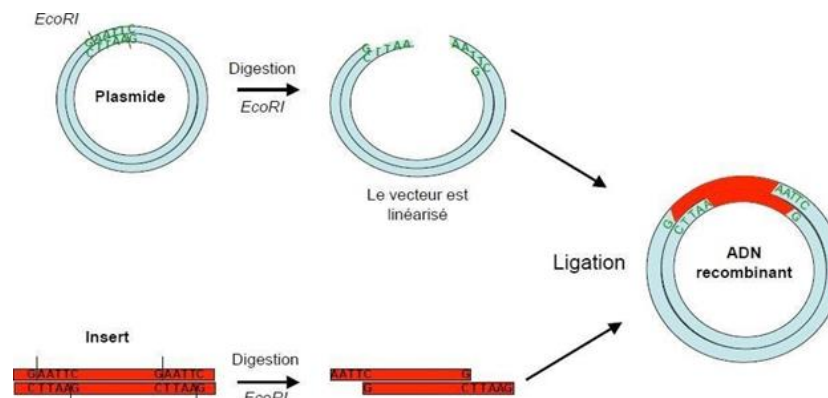
✚ Digestion du vecteur :

On a coupé le plasmide il est prêt à accueillir l'insert.

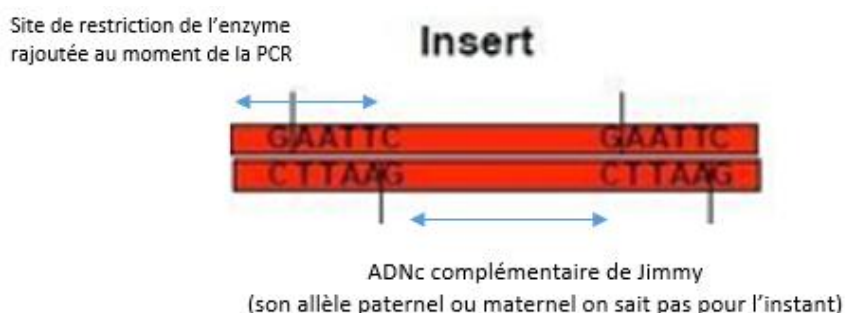
On va vouloir obtenir un insert avec des extrémités dont les nucléotides soient complémentaires au vecteur afin qu'ils s'apparient par complémentarité des bases. Mais comment faire ?

On sait que les extrémités simple brin libres de notre vecteur sont la séquence AATT sur un brin et TTAA sur le brin d'en face : obtenue après digestion d'EcoRI.

En fait on va s'arranger pour que notre **insert** possède le **site de restriction de l'enzyme choisie** (ici EcoRI) à ses extrémités, afin que l'enzyme puisse venir digérer l'insert et que les bords vecteurs/insert puisse se compléter par complémentarité (comme les pièces d'un puzzle).



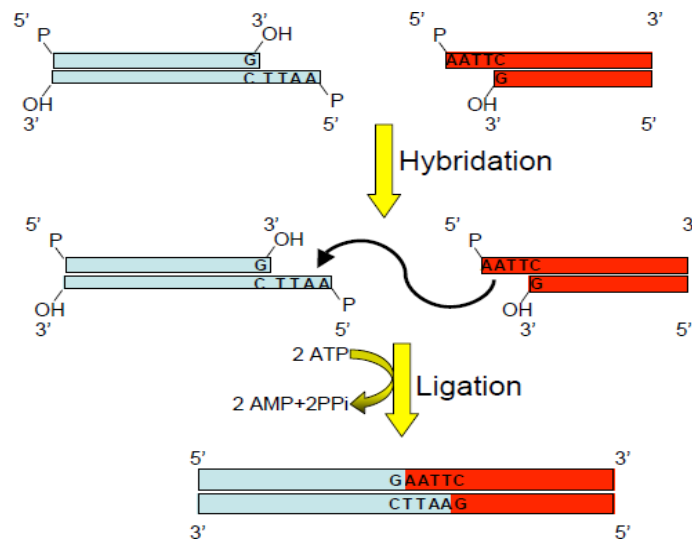
Pour cela, on se rappelle qu'au moment de la **PCR** on a **ajouté une séquence à l'amorce** afin que cette séquence soit amplifiée avec l'ADNc de la Wolframine (cf. NTB bas de page 4). Cette séquence était justement la suite de nucléotides GAATTC : site de restriction d'EcoRI. Attention normalement on n'a pas touché à l'information qu'on veut séquencer = ADNc de l'enfant !! On obtient après PCR :



2. Ligation du plasmide et de l'insert via la T4 DNA ligase

Par **complémentarité des bases** le brin 5'P-3'OH de l'insert s'hybride au brin 3'OH-5'P du vecteur -> création de **liaisons hydrogènes** entre deux brins face à face.

Puis afin de lier les brins 5'-3' de l'insert et 5'-3' du vecteur (liaison sur le même brin = liaison phosphodiester) on a besoin de la **T4 DNA ligase**. Enzyme qui catalyse la formation d'une liaison **phosphodiester** entre un 3'OH et un 5'Phosphate **en présence d'ATP** et d'ions divalents = liaison covalente entre deux fragments d'ADN.



RECAP :

Les liaisons hydrogènes :

- création : naturellement par complémentarité des bases
- destruction : par la chaleur, d'où l'étape de dénaturation de la PCR à 95°

Les liaisons phosphodiesteres = liaisons covalentes :

- création : ADN polymérase ou ligase (enzyme)
- destruction : la chaleur ne marche pas, il faut aussi une enzyme : nucléase (endo ou exo)

✚ L'étape de ligation selon les 3 stratégies de clonage :

- **Extrémités cohésives** : pas de soucis

Rien de particulier à faire la complémentarité des bases attire l'insert dans le plasmide (liaisons H) puis la T4 DNA ligase se charge des liaisons phosphodiesteres.

- **Extrémités franches** : étape supplémentaire de déphosphorylation (facultatif)

On rajoute une étape car la ligase risque de refermer le vecteur sur lui-même avant même d'avoir mis l'insert ce qui n'est pas intéressant. On ajoute des phosphatases au vecteur uniquement. Ce sont des enzymes bactériennes déphosphorylantes, elles remplacent les phosphates en 5' par un -OH. Ainsi la ligase ne peut plus agir sans utiliser l'insert, qui sera alors intégré.

Explication : La T4 DNA ligase n'a qu'une seule liaison phosphodiester à former dans le cas où l'insert n'est pas présent, alors qu'elle doit en faire deux si l'insert est intégré dans le vecteur.

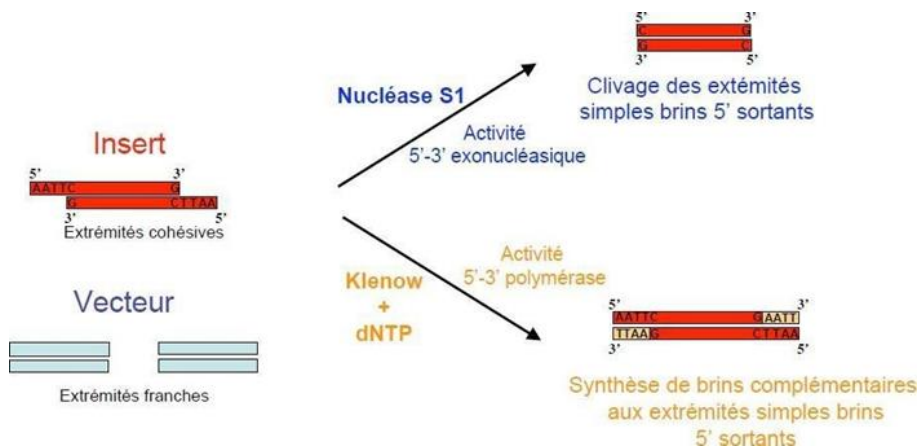
L'enzyme aura donc tendance à privilégier le cas le plus simple, c'est-à-dire celui où elle n'a qu'une seule liaison à faire, et donc où l'insert n'est pas intégré dans le vecteur !

- **Extrémité franches (vecteur) et cohésives (insert) ou l'inverse** :

On utilise **de préférence le même type d'enzymes de restriction** sur le **vecteur et sur l'insert** pour avoir le même type d'extrémité pour les deux.

Cependant, il existe des situations où **on ne peut pas** utiliser les mêmes enzymes de restriction pour le vecteur (pour lequel on dispose d'un grand nombre d'enzymes utilisables) et pour l'insert.

On a donc recours à une **stratégie de clonage** car le vecteur et l'insert ne présentent pas le même type d'extrémités (on peut avoir un insert à bords francs et un vecteur à bords cohésifs ou l'inverse, dans les deux cas les deux ne sont pas complémentaires et ne pourront pas s'apparier/s'hybrider).



Exemple, ici 2 solutions :

- Soit **on coupe** les extrémités 5' sortantes de l'insert :

Nucléase S1 : activité 5'-3' exonucléasique -> grignote les extrémités simple brin sortantes : l'enzyme s'arrête de grignoter les nucléotides dès qu'elle rencontre de l'ADN double brin !

- Soit **on allonge** l'extrémité 3' de l'insert pour faire du double brin partout :

Klenow + dNTP : activité 5'-3' polymérase -> comble l'ADN simple brin en synthétisant les brins complémentaires en présence de dNTP !

Point Vocab :

- Enzyme de restriction = ENDOnucléase (à l'intérieur) différence avec EXOnucléase (aux extrémités), attention aux pièges elle a insisté dessus
- Copier/synthétiser = polymérase
- Reformuler les liaisons = ligases.

Conclusion du clonage moléculaire :

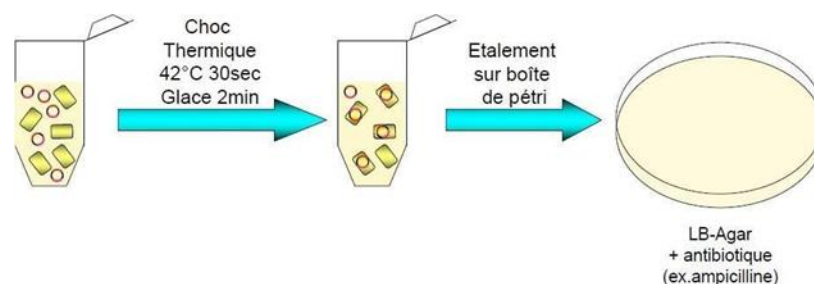
On se retrouve avec **un seul produit PCR dans un vecteur** : 2 molécules d'ADN ne pouvant pas s'intercaler dans le vecteur. On a ainsi créé un ADN recombinant afin **d'isoler UN produit PCR**, (que ce soit le maternel ou paternel on ne sait pas encore mais ce n'était pas le but, on verra plus tard en séquençant).

D. Transformation : introduction du vecteur dans une bactérie

On se sert de souche issue d'E.Coli. On place les ADN recombinants et les bactéries dans un tube. Pour que l'ADN recombinant puisse rentrer dans la cellule hôte on doit rendre les **bactéries compétentes** : leur **membrane** a été **perméabilisée** avec un **choc thermique** ou **électrique**.

On aura alors :

- Des bactéries ayant intégrées un ADN recombinant contenant **l'ADNc paternel**



- Des bactéries ayant intégrées un ADN recombinant contenant **l'ADNc maternel**
- Des bactéries ayant intégré le **vecteur sans insert** (vecteur qui s'est refermé tout seul)
- Des bactéries **sans ADN recombinant** = pas de gène de résistance situé sur le vecteur -> sensible aux antibiotiques

Introduction d'ADN dans une cellule **bactérienne** = **TRANSFORMATION**

Introduction d'ADN dans une cellule **eucaryote** = **TRANSFECTION**

Qu'est-ce qu'il nous reste à faire pour Jimmy ? Il va falloir maintenant sélectionner les bactéries qui nous intéressent (on élimine celles sans insert ou sans ADN recombinant), les isoler, les amplifier et extraire l'ADN Recombinant puis l'ADNc afin de le séquencer.

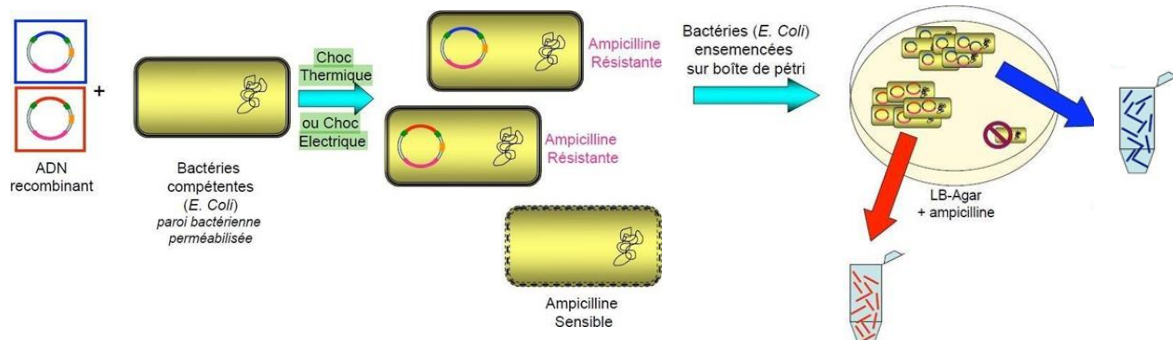
E. Sélection, isolement, amplification, extraction

a) Sélection

- Contiennent – ils un **Vecteur (avec ou sans insert)** ? -> sélection à l'**Ampicilline** (antibio)

On va étaler les bactéries sur une boîte de Pétri dans du LB-Agar + antibio : c'est là que le gène de résistance à l'Ampicilline entre en jeu. Il permet **d'éliminer les bactéries n'ayant pas intégré de vecteur** (elle insiste beaucoup sur ce point depuis le début du cours). Les boîtes sont stockées à 37°C, au bout de 16h on a des colonies.

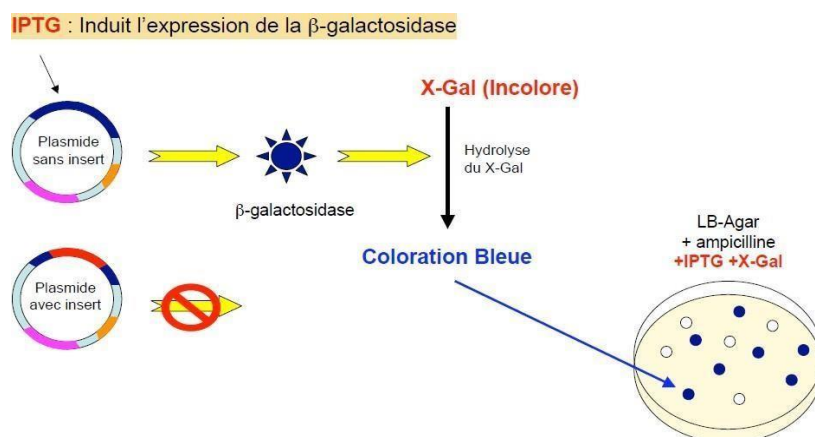
Il faut faire attention de pouvoir les séparer par la suite pour le repiquage, donc en tant que bon généticien t'en met pas trop dans la même boîte.



- Les vecteurs contiennent – ils un **insert** ? -> **sélection blanc/bleu**

Parfois malgré certaines précautions (comme la déphosphorylation pour les bords francs) on a des **vecteurs qui se referment sur eux-mêmes sans insert**.

Pour savoir s'il y'a un insert on va utiliser notre connaissance de la position des gènes du plasmide. Le **poly-linker** est au sein du **gène** codant pour la **Bêta-galactosidase**, l'insert étant intégré à ce niveau il va éteindre le gène ... s'il est bien présent !

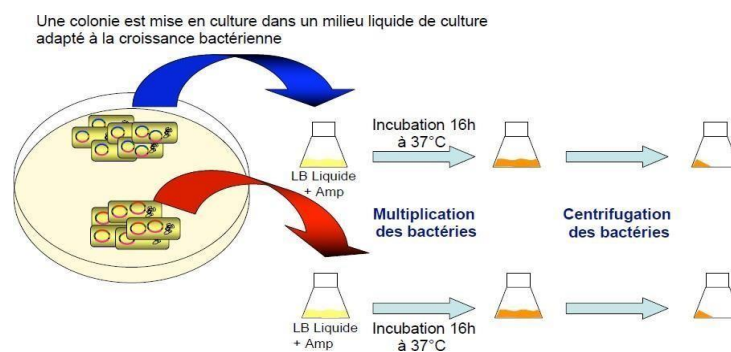


Donc, pour voir si la B-Galactosidase est exprimée on ajoute dans la boîte de Pétri : de l'**IPTG** (qui induit l'expression de la protéine) et **X-Gal** (substrat incolore qui devient bleu si la B-Galactosidase l'hydrolyse) : on ne repiquera que les colonies **BLANCHES** !

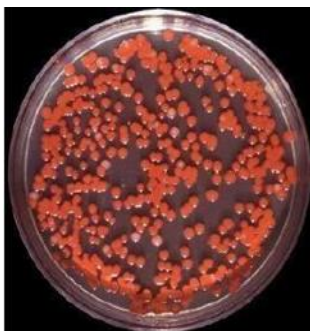
AVEC INSERT	SANS INSERT
Gène bêta-galactosidase inactivé	Gène bêta galactosidase activé
X-Gal incolore non hydrolysé	X-Gal hydrolysé colorant le milieu en bleu
COLONIES BLANCHES	COLONIES BLEUES

b) Isolement et amplification

On récupère donc **uniquement les colonies blanches** pour les repiquer en **milieu liquide à 37° pendant une nuit** avec de l'**ampicilline** pour pas que d'autres bactéries se développent, la multiplication de notre ADN d'intérêt va continuer.



Remarque : on veut séparer les deux populations d'ADN (paternel et maternel) en isolant les bactéries sur une boîte de pétri et en les étalant de façon à obtenir plein de petites colonies. Bien comprendre qu'**une bactérie = un ADN Recombinant = Un insert = ADNc paternel OU maternel**



UNE SEULE bactérie est à l'origine de toute une colonie

Un seul « **point** » correspond à **une colonie** et non pas à une bactérie ! Il s'agit en réalité d'une bactérie qui s'est multipliée pour donner une colonie : c'est la **même population**. Par contre **d'une colonie à l'autre** les bactéries ne contiendront **pas** forcément le **même ADN recombinant (paternel ou maternel)**.

On s'arrange aussi pour que les bactéries soient suffisamment éloignées les unes des autres pour que les clones soient purs et qu'on en ait un seul type par colonie !

c) Extraction

On récupère ensuite les bactéries par centrifugation. De ce culot bactérien on va extraire l'ADN, avec le même principe que pour l'extraire d'une cellule :

1) **Mise en suspension** des bactéries dans un tampon particulier : **NaOH** (lyse alcaline des protéines), détergents (**SDS**) pour la paroi bactérienne et **RNases** pour éliminer l'ARN bactérien.

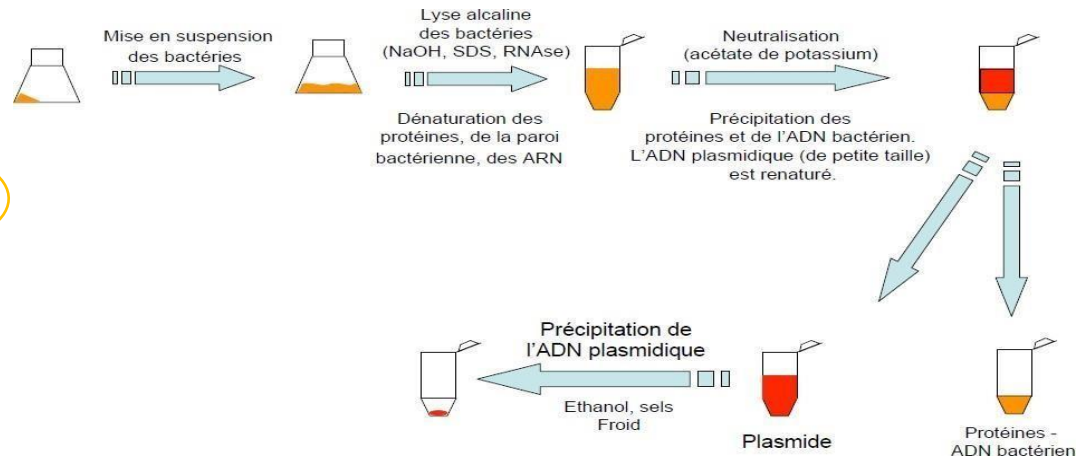
2) **Neutralisation** de la lyse (acétate de potassium)

3) **Précipitation** des protéines et de l'ADN bactérien. L'ADN plasmidique (de petite taille) est renaturé.

4) On **centrifuge** et on sépare protéines/ADN bactérien des plasmides (dans le surnageant).

5) **Précipitation** de l'ADN plasmidique : on récupère le surnageant à part et on rajoute éthanol et NaCl et on congèle à -20°C. Plus qu'à récupérer le culot d'ADN.

Cette partie osef un peu, lisez là quand vous revoyez le cours mais vous cassez pas la tête à l'apprendre par cœur !



Maintenant qu'on a notre ADN Recombinant, il ne nous reste plus qu'à isoler l'insert et le séquencer ! Mais avant, comme pour notre PCR, **on se doit de vérifier que la séquence que l'on veut séquencer (l'insert) est intacte** (imaginez si on séquence un insert abimé et qu'on fait à cause de ça un faux diagnostic !). Pour la PCR on faisait une électrophorèse pour vérifier que notre fragment faisait bien le bon nombre de paires de base. Ici on utilise les **cartes de restrictions**.

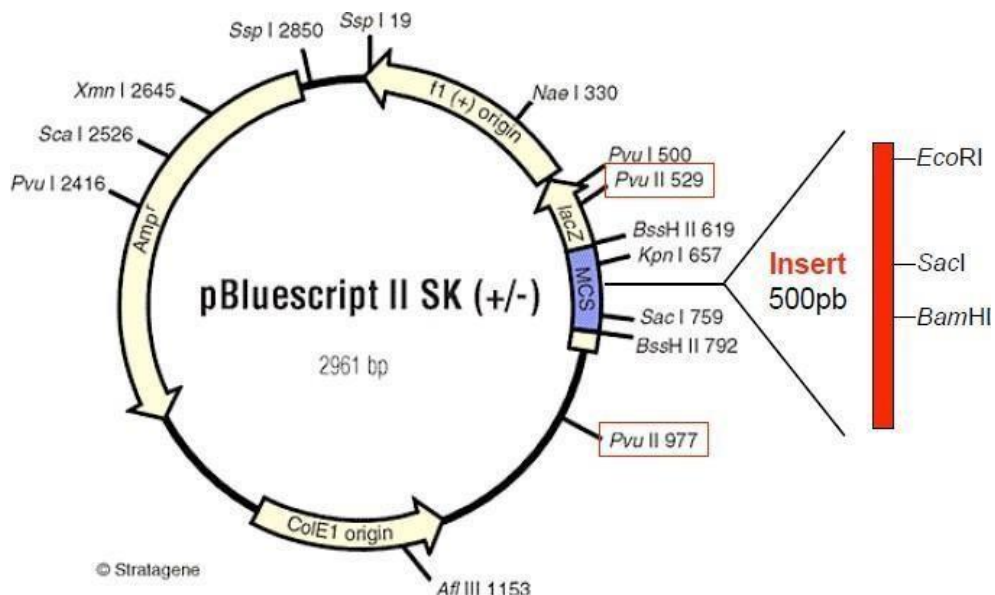
ON SE REVEILLE !! C'est le moment de finir ta dernière gorgée de café et de se concentrer un max guerrier -> c'est LE passage expérience de l'UE11 et c'est tombé chaque année en concours et c'est pas si difficile quand on saisit le truc

Une fiche méthodo t'attend sur le forum 😊

F. Carte de Restriction

Ces cartes sont en fait des **schémas** sur lesquels on retrouve la **séquence complète du plasmide** et où sont marqués tous les **sites de restriction** qui servent (on peut retrouver les autres dans la documentation du plasmide si jamais). On va alors **digérer le plasmide** avec différentes enzymes de façon à **vérifier qu'il y a bien l'insert** et qu'il est bien de la **bonne taille**, car même avec la sélection par B-Gal on n'est pas complètement surs.

Exemple du cours (carte ci-dessus) :

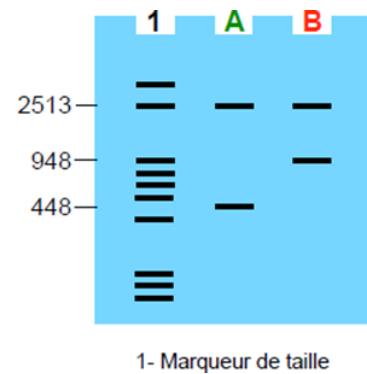


❖ Ici on a un vecteur qui fait 2961 pb

Si on coupe le vecteur avec **Pvu II**, l'enzyme agit en 529 et en 977.

On se retrouve alors avec un fragment de 2513 et un de 448 pb s'il n'y a pas d'insert (piste A), contre un fragment à 2513 et un à $448+500 = 948$ pb s'il y a l'insert (qui fait 500pb) (piste B).

On peut donc savoir s'il y a un **insert ou non** en faisant simplement **migrer sur gel** d'agarose : selon la **taille** observée on sait déterminer la présence ou non de l'insert.



❖ **Attention, vous pouvez avoir des sites de restriction dans l'insert.**

En concours, on vous donne un cas de figure en vous précisant la taille de l'insert et en vous disant si telle ou telle enzyme est présente ou non dans l'insert. On vous donne également un résultat de gel d'agarose et c'est à vous d'interpréter. Il faut donc toujours bien calculer les fragments qui sont coupés par telle ou telle enzyme.

Est-ce qu'on tient compte aux sites de restrictions dans l'insert ? Ce sera précisé. Par exemple ici « l'insert fait 500 pb et il contient un site de restriction à l'enzyme ScaI qui nous intéresse, cette enzyme coupe l'insert en 2 fragments de 200 et 300 pb ».

G. Séquençage des inserts

Rappel de la situation : on suspectait cliniquement un syndrome de Wolfram chez notre patient.

Le séquençage de l'ADN (l'exome) était illisible car il y avait un mélange de variant différents :

-> présence des 2 allèles (d'origine paternelle et maternelle confondus dans le cytoplasme et donc séquencés ensemble). Pour expliquer ce décalage on a suspecté la présence d'un variant d'épissage chez la mère qui décalerai tout au séquençage. Maintenant qu'on a purifié et isolé les 2 types d'ADNc (mère et père) on va pouvoir faire une PCR- séquençage classique.

Résultats :

-ADNc paternel : jonction exon 6/ exon 7 normale mais rappelez-vous qu'on avait identifié une substitution

-ADNc maternelle : présence dans cette jonction d'une séquence supplémentaire de 16 nucléotides qui introduit un site cryptique d'épissage

-> Maintenant que l'on a formellement séquencé : l'allèle paternel muté et les ARNm mutés issu du gène maternel on peut poser le diagnostic de syndrome de Wolfram !!

Petit rappel : pour poser un diagnostic 3 critères :

- ✓ Il a tous les **signes cliniques**
- ✓ La littérature (=études scientifiques) possède les **preuves de la pathogénicité de ces variant**
- ✓ On a pu **justifier la transmission de la mutation** (on a trouvé l'allèle paternel **et** maternel défectueux en séquençant le génome paternelle)

Ici le critère de littérature est acquis. Il faut comprendre qu'avec les nouvelles techniques de séquençage la difficulté n'est plus tellement de connaître la séquence des gènes et de détecter des variants par rapport à une séquence type mais de connaître la répercussion de ces mutations et de savoir si elles sont pathogènes ou non, leurs effets sur les protéines ...

Comment étudier l'impact sur la protéine ? -> Le clonage d'expression.

Partie 2 : Autres techniques de Biomol

A. Le Clonage d'expression

a) **Le but ?** Etudier artificiellement, dans les cellules eucaryote, l'expression d'une protéine.

b) **Principe ?** L'ADN recombinant (vecteur d'EXPRESSION (≠ vecteur de clonage) + insert) est multiplié dans une bactérie puis transféré dans une cellule eucaryote pour y être exprimé.

LE VECTEUR D'EXPRESSION

✓ Pour être multiplié dans la bactérie (comme le vecteur de clonage) :

- Un Poly linker ou site multiple de clonage
- une origine de réplication PROCARYOTE
- un gène de sélection PROCARYOTE



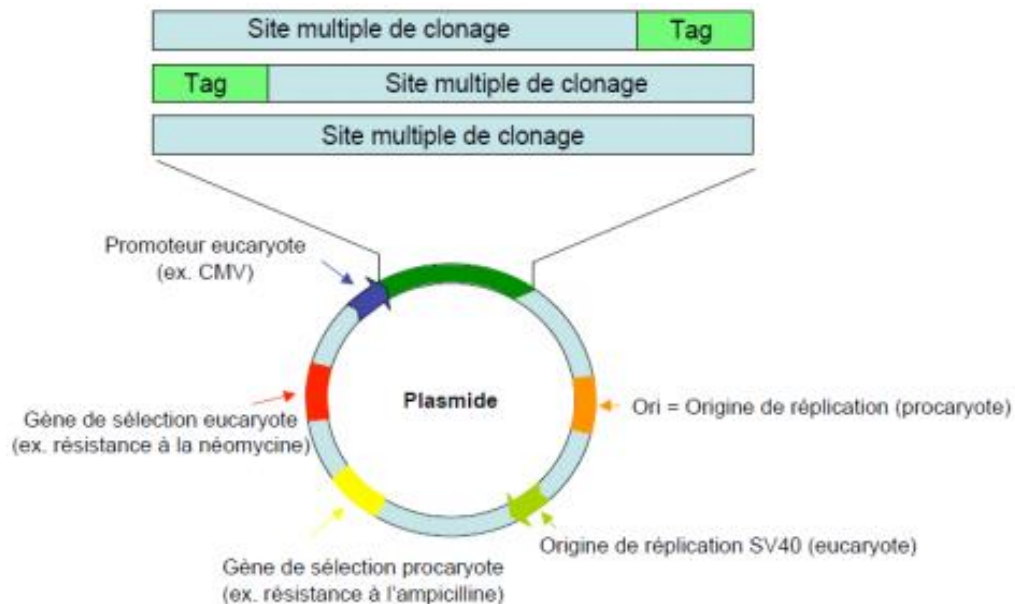
✓ Pour que la cellule eucaryote surexprime la protéine (transcription/traduction etc)

- une région promotrice EUCARYOTE (origine virale)
- une origine de réplication EUCARYOTE (SV 40, viral)
- un gène de sélection EUCARYOTE
- Tag = étiquette**



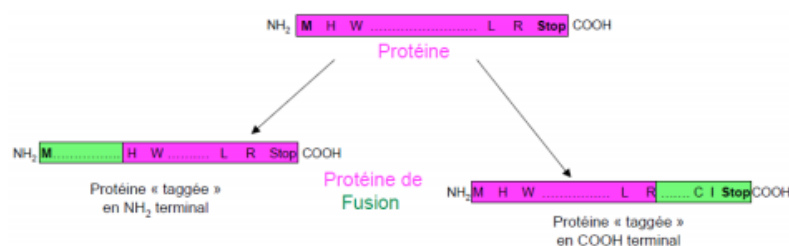
Le tag :

Pour suivre cette protéine au microscope une fois qu'elle aura été produite par la cellule eucaryote on peut utiliser des tags -> marqueurs fluorescents ou peptides antigènes.



On rajoutera ces **tags à l'insert** en **N-term** ou **C-term**. En **N-term** il faut faire attention de le mettre **après** la méthionine (**codon start**) pour qu'il soit transcrit + traduit et en **C-term** il faut le mettre **avant** le **codon stop** sinon il ne sera pas inclus dans la protéine.

Protéine de fusion = ADNc d'intérêt + étiquette



c) Les étapes ?

1^{ère} étape : Création de l'ADN Recombinant avec des enzymes de restriction (même étape que le clonage moléculaire)

2^{ème} étape : Transformation : l'ADN Recombinant est multiplié dans la bactérie

3^{ème} étape : TRANSFECTION : Transfert des ADN Recombinant dans une cellule eucaryote qui va exprimer le gène et produire la protéine.

Pour faire rentrer cette ADN recombinant dans la cellule eucaryote on peut utiliser différentes méthodes comme le électroporation etc.. Exemple : Transfection par phosphate de calcium :

L'ADN est mélangé avec une solution de chlorure de calcium puis on ajoute un tampon phosphate goutte à goutte. Il se forme un précipité de phosphate de calcium autour de l'ADN formant des « cristaux ». Ces cristaux de phosphate de calcium contenant l'ADN vont ensuite rentrer dans les cellules par endocytose ou phagocytose. L'ADN va être libéré dans le cytoplasme, il pourra s'intégrer ou non dans le génome nucléaire pour ensuite être transcrit en ARN messagers et exporté dans le cytoplasme pour la traduction.

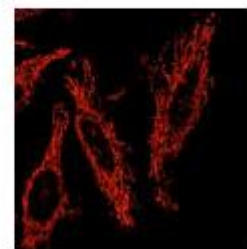
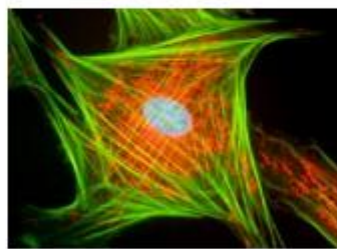
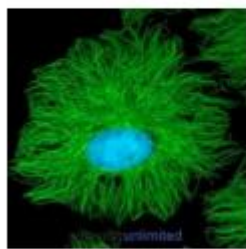
RAPPEL :

Introduction d'ADN dans une cellule **bactérienne**
= **TRANSFORMATION**

vs

Introduction d'ADN dans une cellule **eucaryote**
= **TRANSFECTION**

Les études de localisation utilisant des Tag fluorescents ou de l'immunofluorescence concernent uniquement les protéines

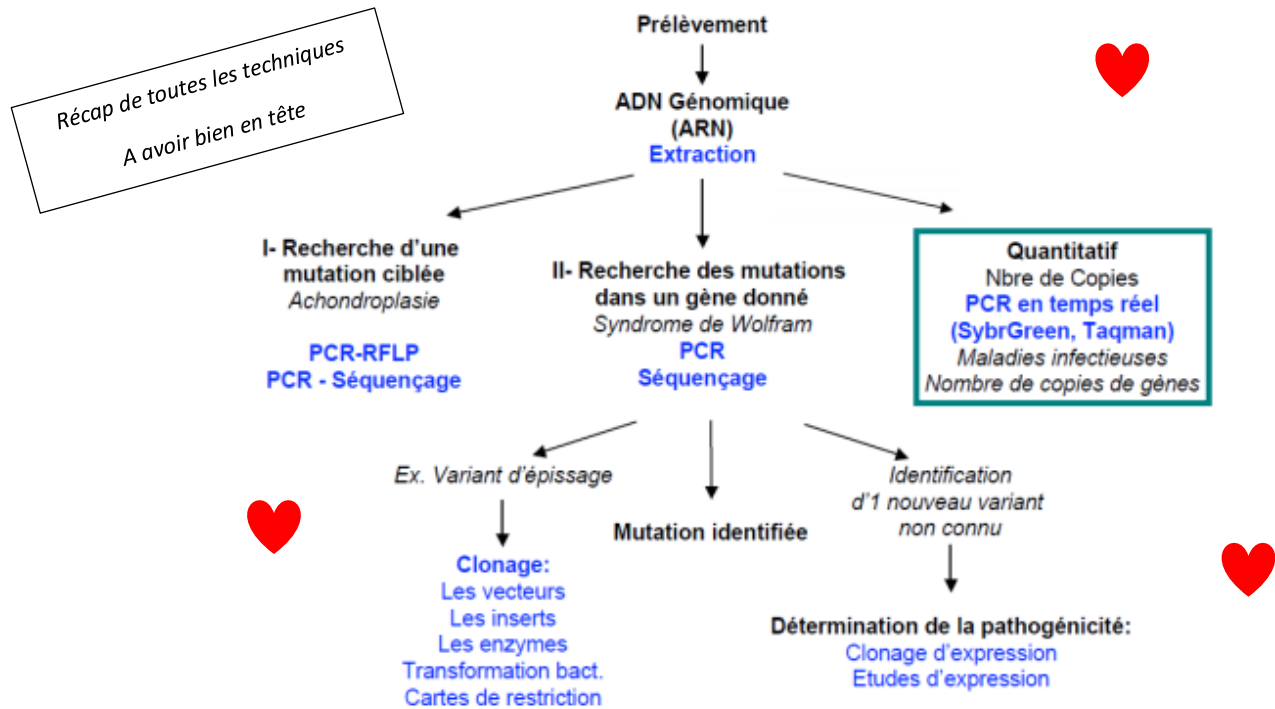


filaments de tubuline : vert / mitochondries : rouge / noyau : bleu

B. PCR en temps réel

Nous avons vu jusqu'à présent comment trouver une mutation : ciblée (achondroplasie) ou dans un gène donné (syndrome de Wolfram). Dans ces deux cas nous avons manipulé notre matériel génétique dans l'unique but final de séquencer un gène du patient -> but ? déterminer la nature de quelque chose (en l'occurrence des nucléotides), ces techniques avaient toutes un but QUALITATIF.

Mais dans certains cas, on peut être amenés à QUANTIFIER quelque chose, on fera donc des études de nombre (on ne s'intéresse plus à la nature des nucléotides). Notamment en virologie, si on veut déterminer la charge virale, il faut qu'on puisse savoir combien de nombre de copies du virus étaient dans notre échantillon (mais on se fiche de connaître leur séquence nucléotidique).



a) Principe

La PCR en temps réel utilise le même principe que la PCR avec les mêmes étapes, sauf que la mesure du nombre d'amplicons réalisées ne se fait PAS à la fin après 35-40 cycles mais en temps réel tout au long de la réaction.

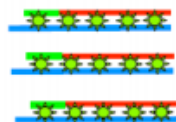
o PCR classique : Après 40 cycles, une certaine quantité de produits PCR est générée. Un agent intercalant est ajouté dans le milieu réactionnel, ce qui rend le système fluorescent sous lumière U.V. Technique non quantitative

o PCR en temps réel : On mesure la fluorescence après chaque cycle car l'agent intercalant est incorporé au fur et à mesure de l'avancement de la réaction, on ne l'ajoute pas à la fin ! Technique quantitative

PCR « classique »:
35-40 Cycles



PCR en temps réel:
Mesure de Fluorescence
à chaque cycle



Rappel :

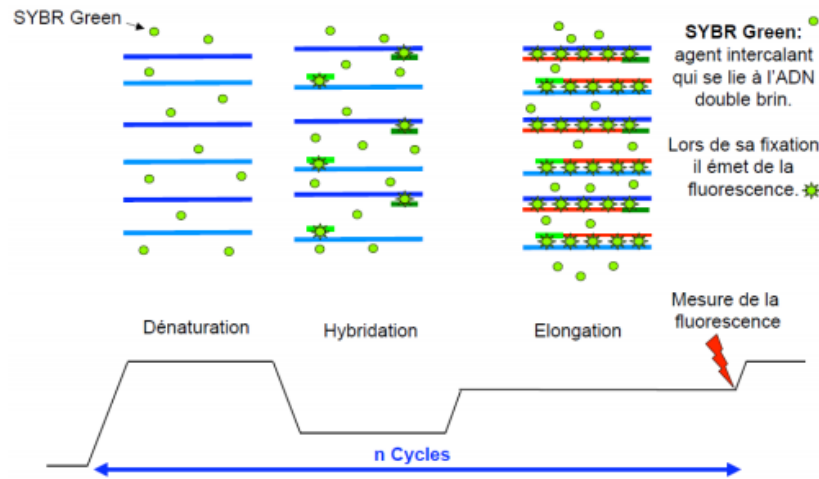
PCR = 3 étapes que l'on répète n cycles :

1. **Dénaturation** (95°C)
2. **Hybridation** (55°C)
3. **Elongation** (60°C)

b) SYBR Green

La mesure se fait grâce à un agent fluorescent (SYBR Green) dont l'émission de fluorescence est proportionnelle à la quantité d'amplicons produits.

Le SYBR Green est une petite molécule qui va s'intercaler dans l'ADN double brin et qui a la propriété d'émettre une fluorescence uniquement à partir du moment où elle est intercalée !



La mesure de la fluorescence se fait à la fin de chaque cycle après l'étape d'élongation.

Donc plus y a d'amplicons (qui sont justement des ADN double brin), plus le SYBR Green s'intercale et plus il y aura de fluorescence mesurée par l'appareil !

c) PCR Quantitative

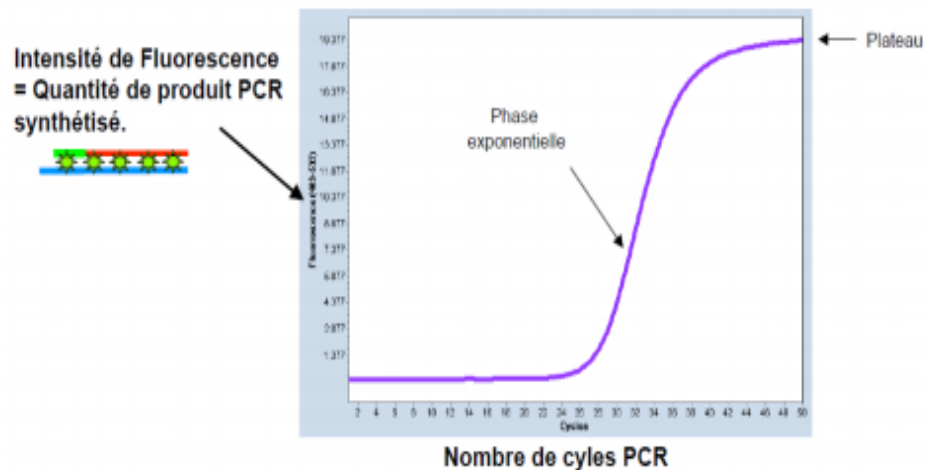
La principale application de la PCR en temps réel est la PCR quantitative.

La courbe représente l'intensité de fluorescence des produits PCR générés en fonction du nombre de cycles, et présente trois phases :

o **Phase de plateau** : pas assez de fluorescence pour avoir une détection.

o **Phase exponentielle** : au bout d'une vingtaine de cycles, les produits PCR sont générés exponentiellement jusqu'à saturation du système.

o **Phase de plateau** : (au bout de 40 cycles environ) le système est saturé car le milieu réactionnel est épuisé (plus de dNTP par exemple) et la Taq Polymérase fatiguée.

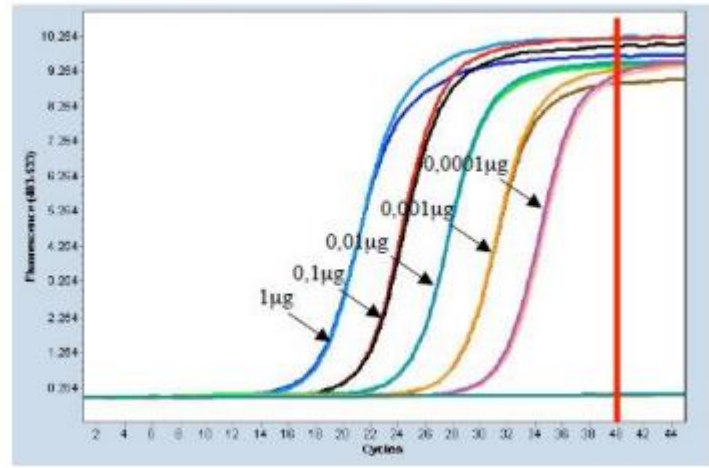


Plus on a d'ADN au départ, moins il faudra de cycles pour que la machine détecte la fluorescence.

Le moment où la machine commence à détecter = fin de la 1^{ère} phase plateau = début de la phase exponentielle correspond au cycle seuil Ct.

En mettant 1 microgramme d'ADN au départ, on obtient la courbe bleue (tout à gauche). Si on part avec une concentration de 0,1 microgramme (courbe rouge), soit 10 fois moins d'ADN, on voit qu'il faut plus de cycles pour que la fluorescence soit détectable : la courbe se décale vers la droite !

Mais, après 40 cycles, l'intensité des produits PCR générés est identique quelle que soit la quantité d'ADN génomique utilisée au départ !



L'écart entre le nombre de cycle seuil Ct des différentes courbes est proportionnel à la quantité de ADN placé dans la machine initialement (facteur 3,3 entre chaque courbe). La gamme étalon nous donnera une relation linéaire

En déterminant le cycle seuil (au point d'inflexion), on peut donc quantifier le milieu réactionnel et connaître exactement la quantité d'ADN de départ !

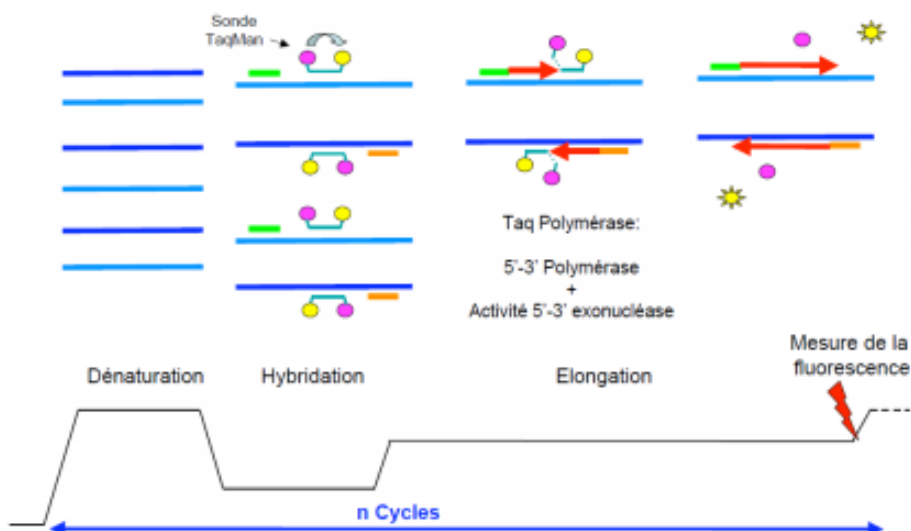
d) Technologies TaqMan

En dehors du SYBR Green, une autre technologie **plus spécifique** peut aussi être utilisée : les sondes TaqMan. C'est une seconde technique de PCR en temps réel où l'on rajoute un troisième fragment d'ADN (simple brin) dans la réaction en plus des deux primers de la PCR (d'où la + grande spécificité c'est une 3^{ème} assurance qu'on amplifie bien le gène qu'on veut).

La **sonde TaqMan** est un petit fragment **d'ADN simple brin** qui ira s'hybrider sur l'ADN à amplifier par **complémentarité**. Elle est constituée d'un **quencher** et d'un **fluorochrome** : le rôle du quencher est d'éteindre la fluorescence du fluorochrome quand ils sont physiquement proches.

Lors de l'élongation, la polymérase (en plus de synthétiser le brin fils de 5' en 3') va grignoter la sonde TaqMan (activité 5'-3' exonucléasique) pour pouvoir continuer son chemin. Ainsi, elle va séparer le quencher du fluorochrome qui pourra donc fluorescer et la machine détectera cette émission.

Là encore la mesure de la fluorescence se fait à la fin de la phase d'élongation, à la fin de chaque cycle. Et la quantité de fluorescence générée sera également proportionnelle à la quantité d'ADN au départ.



RECAP FINAL

Après prélèvement biologique et extraction du matériel génétique (ADN ou ARN) puis PCR et vérification, on s'oriente selon ce que l'on souhaite rechercher :

- ✓ Si on cherche une mutation ciblée, on utilise la PCR- RFLP qui doit être confirmée par une technique de séquençage (EX : achondroplasie)
- ✓ Si on cherche plusieurs mutations dans un gène donné au complet (EX : syndrome de Wolf RAM) on séquence le gène dans son ensemble grâce à une PCR-séquençage. 2 possibilités :
 - Soit on a identifié notre mutation (sur les 2 allèles dans Wolfram car récessif) et on peut délivrer le diagnostic.
 - Soit les variants ne sont pas situés sur les exons mais au niveau des introns. Possible création de sites cryptique dépistage. On effectuera un clonage moléculaire afin de séparer les allèles pour identifier exactement le variant dépistage en le séquençant.
Si on ne connaît pas ce variant on essaye de déterminer sa pathogénicité grâce au clonage d'expression.
- ✓ Si on cherche à quantifier le nombre de copies d'un gène on utilise la PCR en temps réel.
- ✓ Si on cherche des mutations dans plusieurs gènes on va utiliser le séquençage haut-débit. (cf. Fiche 3)

Dédicacess : **A toi**, qui a déjà bossé depuis environ 6 ½ mois maintenant, ne lâche pas, pour les efforts que tu as déjà fournis, pour la raison pour laquelle tu en rentré en PACES, et pour les merveilleuses études qui t'attendent

A la team Biomol, ma co-tuuut et nos Vieux Baptiste et Hugo que j'aurais sollicité même après son mandant pour relire ce DM (et comme c'est un <3 il nous l'a proposé)

A mes amies en or et combattantes de PACES Aubrée, Natacha, Linda, Andréa, petit coucou à Arthur et Camille aussi

Mes fillots et fillotes qui lâchent rien, Auxanne, Véronique, Côme, Sonia, Rémi et Natasha et mo-co parrain Lucas

A ma marraine de P2, Pauline qui verra très très probablement pas ce DM

A toute l'équipe du tut qui font un boulot formidable : petit + pour Lucas Cauvin et Virgile Santini (aka les 2 crânes d'œufs)

RIEN N'EST FICHU OU ACQUIS AVANT LE JOUR J

Signé : une 257 au S1 à qui on a bien fait comprendre que c'était fichu (eh ben non) ET qui crois beaucoup que TU peux réussir <3 TU ES CAPABLE DE BIEN PLUS QUE TU NE LE PENSE – CROIS EN TA REUSSITE

Restez chez vous, bossez au Soleil