

# Techniques d'analyses moléculaires

Master 1 Amélioration des plantes

Dr. D. TABTI

## Objectifs du cours :

- Connaitre les techniques utilisées en biologie moléculaire
- Connaitre une nouvelle notion : les QTLs
- Connaitre comment construire une carte de linkage
- Connaitre la sélection assistée par marqueurs

## Contenu de la matière :

### Partie Cours

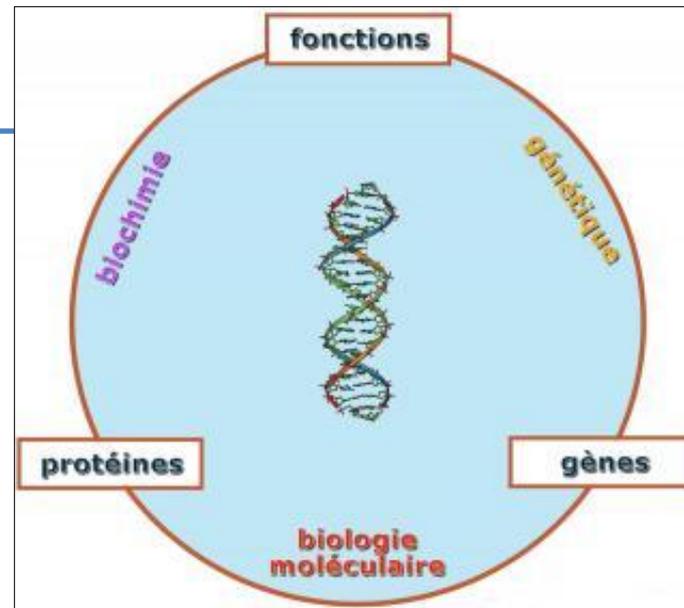
- Analyse du polymorphisme de l'ADN
- Introduction à l'analyse des QTLs et cartographie
- Sélection assistée par marqueurs
- Analyse du polymorphisme protéique
- Séquençage des macromolécules

### Partie TD: exposés

# Introduction

Biologie moléculaire : discipline qui explore la structure et la fonction des **macromolécules** biologiques, en explorant la structure, la biosynthèse et la fonction de l'**ADN** et de l'**ARN** au niveau moléculaire, et en explorant la manière dont elles interagissent les unes avec les autres et avec les **protéines**

Le domaine de recherche de la **biologie moléculaire** recoupe de plus en plus d'autres domaines de la biologie et de la chimie, en particulier la **génétique** (moléculaire) et la **biochimie**.



## Au sein du laboratoire de biologie moléculaire :

- **Produits** : amorces, enzymes (de restriction, taq-polymérase, ARNase, DNAase...), chloroforme, agarose, **BET**, ....
- **Consommables** : tubes eppendorf, gants, plaques PCR, embouts pour micropipettes, ..
- **Appareillages** : vortex, genogrinder, thermocycleur, machine à glace, bain marie, spectrophotomètre, hotte chimique, électrophorèse, centrifugeuse (réfrigérée), micropipettes , autoclave, ...

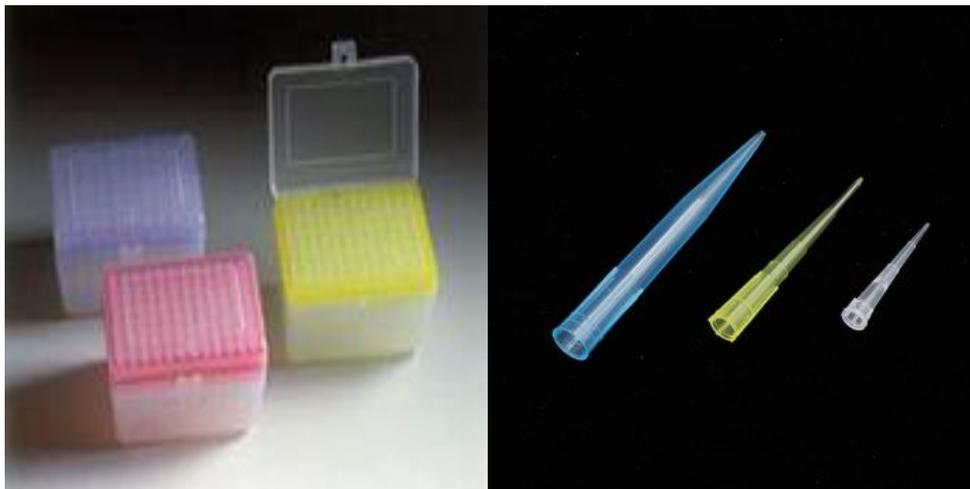
# Consommables



Tubes eppendorfs



Plaques PCR



embouts pour micropipettes

# Appareillages



Vortex



Genogrinder



Thermocycleur



Machine à glace



Bain marie

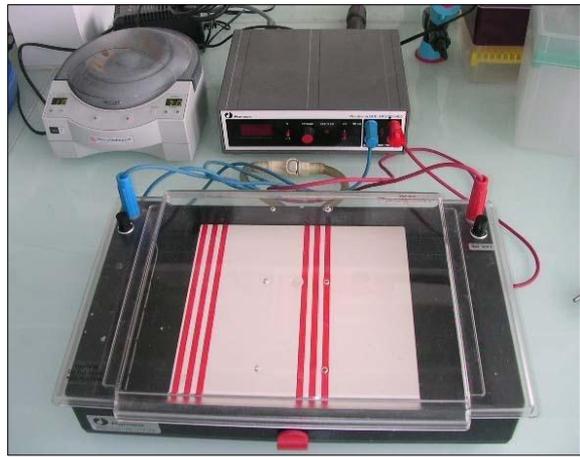


Spectrophotomètre

# Appareillages (Suite)



Hotte chimique



Electrophorèse



Centrifugeuse



Micro-pipettes



Autoclave

## **Au sein du laboratoire de biologie moléculaire :**

**Hygiène et sécurité** : toute activité dans un laboratoire peut nécessiter l'utilisation de produits chimiques potentiellement dangereux, pour soi et pour l'environnement. Il est donc important de suivre les consignes d'hygiène et de sécurité des bien et des personnes, et de bien respecter l'évacuation des déchets.

**Il faut impérativement travailler avec blouse et avec des gants.**

Certains solvants nécessitent d'être préparés et manipulés sous des hottes chimiques.

Dans un laboratoire de biologie moléculaire, le produit le plus dangereux est le **BET (bromure d'éthidium)**, classé parmi les CMR.

Les appareils utilisés peuvent aussi être une source de danger: la centrifugeuse qui tourne à haute vitesse, les autoclaves, les électrophorèse, les lampes UV...

**Travailler sans gaspillage des produits et manipuler soigneusement les appareils.**

# Chapitre 1: Analyse du polymorphisme de l'ADN et protéique

L'ADN = l'acide désoxyribonucléique est une macromolécule, constituant essentiel des chromosomes et porteur de caractères génétiques.

ADN = porteur de l'information génétique

L'extraction de haute qualité et de grande quantité d'ADN génomique à partir de tissus est au cœur de nombreux tests moléculaires et, plus récemment, l'application de séquençage pour caractériser la diversité des plantes, la récupération de l'ADN peut être considéré comme un outil fondamental de la science de la plante.

L'isolement de l'ADN remonte à la fin des années 1800 avec les travaux de Friedrich Miescher et ses collègues qui ont découvert en premier la présence de l'ADN dans les cellules.

Les étapes de base de l'extraction de l'ADN sont les suivantes :

1. La collecte et le stockage de tissus végétaux
2. La lyse des cellules végétales
3. La solubilisation des lipides et des protéines avec des détergents
4. La séparation de l'ADN des autres molécules
5. La purification de l'ADN séparé
6. La suspension dans un tampon approprié

# 1. Méthodes d'extraction d'ADN végétal

## Extraction d'ADN

```
graph TD; A[Extraction d'ADN] --> B[Kits commerciaux]; A --> C[Solutions préparées « recette maison »];
```

### Kits commerciaux

- Qiagen
- MP Biomedical
- Fisher Scientific
- ....

Solutions préparées  
« recette maison »

L'une des méthodes les plus utilisées chez les végétaux est dite la méthode CTAB (par rapport au composant principal du tampon de lyse qui est le bromure de cétyl-triméthyl-ammonium (Cetyl-Trimethyl-Ammonium Bromide)).

Le CTAB solubilise les membranes et les complexes avec l'ADN.

Cette méthode a été décrite en 1980, et elle reste populaire du fait que tous les composants peuvent être préparés au laboratoire et donc le coût par échantillon reste faible.

Il existe aussi la méthode SDS dont le tampon de lyse est dodécylsulfate de sodium (**Sodium Dodecyl Sulfate**).

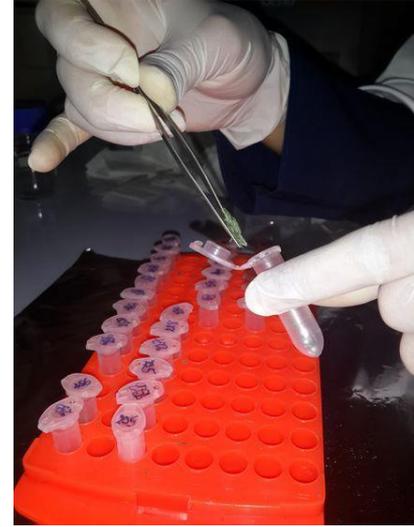
Au cours des dernières décennies, des kits commerciaux pour l'extraction rapide à partir de matériel végétal sont devenus d'une utilisation routinière.

Les kits commerciaux se sont révélés être très fiables dans la production de rendements élevés de l'ADN hautement purifié, et ainsi sont devenus la norme lors de la réalisation des dosages moléculaires sensibles. Ces kits restent trop chères par rapport aux méthodes anciennes.

# Exemple d'une méthode d'extraction classique : méthode CTAB

**Protocole d'extraction à partir de feuilles sèches:** ce protocole est valable pour beaucoup d'espèces végétales notamment *Medicago truncatula* et le pois chiche.

1. Récolter **de jeunes feuilles (3 à 4):** on choisit les jeunes feuilles parce que à ce stade, le taux des protéines et d'autres substances organiques est faible.
2. Si l'extraction se fait juste après la récolte des feuilles, on les met dans des tubes eppendorfs de 2ml avec 8 billes en métal et on met les tubes (couvercle ouvert) à l'étuve pour le séchage à 65°C pendant une nuit.
3. Si l'extraction ne se fait pas tout de suite, on peut conserver les feuilles dans des sachets en papier (étiquetés) et on les mets dans le gel de silice (gel desséchant). Le jour de l'extraction enlever les feuilles des sachets et les mettre dans les tubes eppendorf avec 8 billes et les passer pendant quelques minutes à l'étuve à 65°C (toujours le couvercle ouvert).



**Remarque:** les tubes doivent être numérotés avec un marqueur indélébile pour savoir la provenance de l'ADN extrait

4. Une fois les tubes sont enlevés de l'étuve, il faut les refermer immédiatement pour éviter qu'ils s'humidifient de nouveau.
5. Réduire les feuilles en poudre soit en les vortexant soit en utilisant le genogrinder.



6. Mettre les tubes dans un portoir.
7. Ajouter 1ml de **tampon d'extraction autoclavé** en utilisant une micropipette de 1 ml.



8. Incuber 20 minutes au bain marie allumé le matin à 65°C.



9. Ajouter 600  $\mu$ l de chloroforme et agiter vigoureusement à la main pendant 15 minutes à température ambiante.



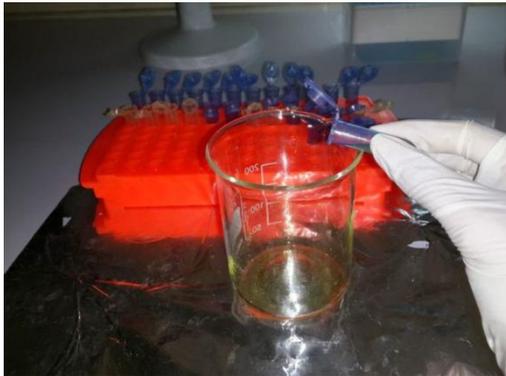
10. Centrifuger 10 minutes à vitesse maximale (13 000 tr/min)



11. Transférer la phase aqueuse supérieure dans un nouveau tube à la pipette sans entrainer l'interface.



12. Ajouter 600  $\mu$ l d'isopropanol et centrifuger immédiatement 20 secondes à vitesse maximale (cette étape est pour enlever le chloroforme et l'ADN va se rassembler en petite pelote).
13. Jeter le surnageant.



14. Ajouter 500  $\mu$ l d'éthanol 70 %.
15. Centrifuger 10 minutes vitesse maximale et jeter le surnageant.

15. Faire sécher sous la hotte ou sur la paillasse



16. Re-suspendre le culot dans 50  $\mu$ l d'eau distillée stérile

17. Rajouter 5  $\mu$ l d'ARNasa (2.5 mg/ml) et laisser une nuit à 4° c.

**Tampon d'extraction (pour 100 ml):**

- 2g de CTAB
- 10 ml de Tris-HCl 1M pH8
- 28 ml de NaCl 5M
- 4 ml d'EDTA 0.5 M pH8
- 0.5 ml  $\beta$ -méthylmercapto-éthanol (à ajouter après autoclavage).

**Le tampon peut être conservé pendant 3 mois.**

# Un kit commercial



## Dosage et analyse de la pureté

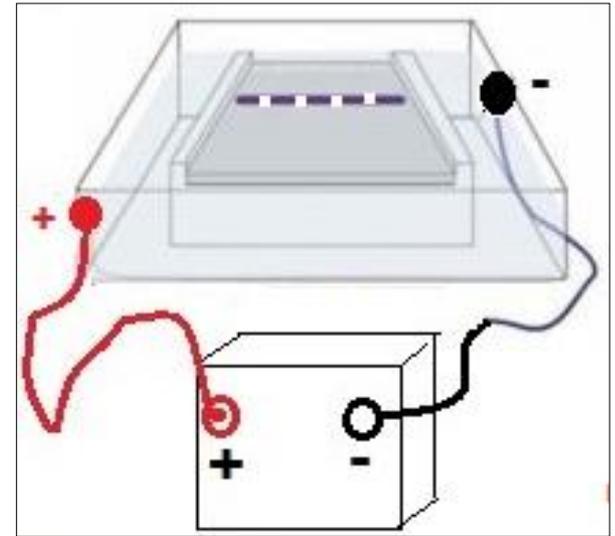
Le dosage peut se faire par 4 méthodes :

Électrophorèse sur gel d'agarose, le spectrophotomètre , le Qubit et **la qPCR**

1. L'électrophorèse est un outil important dans l'analyse qualitative et quantitative des acides nucléiques (ADN, ARN).

Son principe repose sur la séparation des molécules. Les critères de séparation utilisés sont **la taille** (poids moléculaire) et **la charge électrique**.

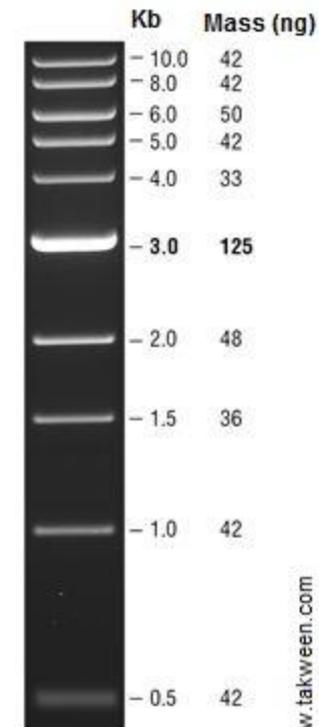
Comme les acides nucléiques sont chargés négativement (grâce au phosphate), le critère de la charge électrique n'est pas comptabilisé dans la séparation (toutes les molécules migrent vers l'anode +) donc les fragments de l'ADN sont séparés uniquement en fonction du poids moléculaire.



Le gel utilisé est le gel d'agarose (généralement à 1%): par utilisation de marqueur de taille « Low DNA ».

Le marqueur de taille nous permet, selon la migration de connaître la taille de notre ADN.

Exemple: 1 Kb DNA ladder (10 fragments)



1 kb DNA Ladder

# Comment préparer un gel d'agarose ?

Pour 200 ml:

2g d'agarose

200 ml de Tris Borate EDTA (TBE 0.5 X)

Une goutte de **BET**

- Mettre dans une fiole Erlenmeyer les 2 g d'agarose et les 200 ml de tampon « TBE ».
- Dissoudre complètement l'agarose dans le tampon en plaçant l'Erlenmeyer au four à micro-ondes et en agitant de temps à autre pour homogénéiser le mélange.



- Refroidir le mélange.
- Rajouter 2  $\mu\text{l}$  de BET.



- Couler lentement le gel dans le moule avec le peigne et laisser refroidir.



- Retirer le peigne et placer le gel dans la cuve en s'assurant qu'il est recouvert de tampon « TBE ».
- Les puits doivent être du côté de la cathode (-, fil noir), car l'ADN est chargé négativement et migrera vers l'anode (+, fil rouge)

Lancer la migration après avoir chargé les échantillons et les standards (Low DNA Mass Ladder) dans les puits selon le dosage suivant :

-5 $\mu$ l : 2  $\mu$ l d'ADN + 3  $\mu$ l de **Bleu de bromophénol** (tampon de charge) ou 3  $\mu$ l : 1  $\mu$ l d'ADN + 2  $\mu$ l de BBT

-1  $\mu$ l de tampon (BBT) + 4  $\mu$ l de Low DNA Mass Ladder



### **Tampon TBE (Tris borate EDTA) pH 8,3**

Pour 1000 mL d'eau distillée :

Tris.HCl 90 mmol.L<sup>-1</sup> : 10,89 g

Acide borique 90 mmol.L<sup>-1</sup> : 5,56 g

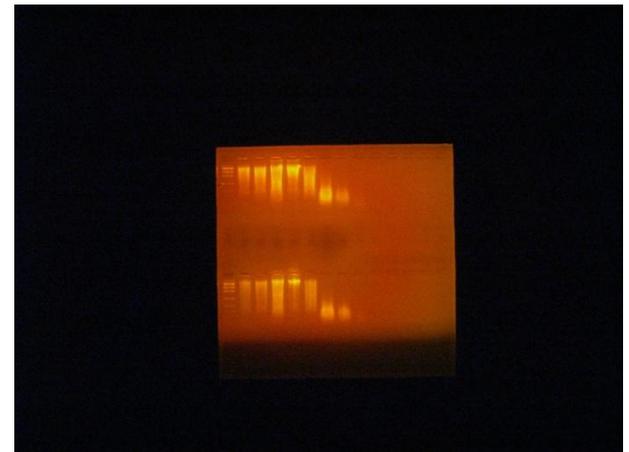
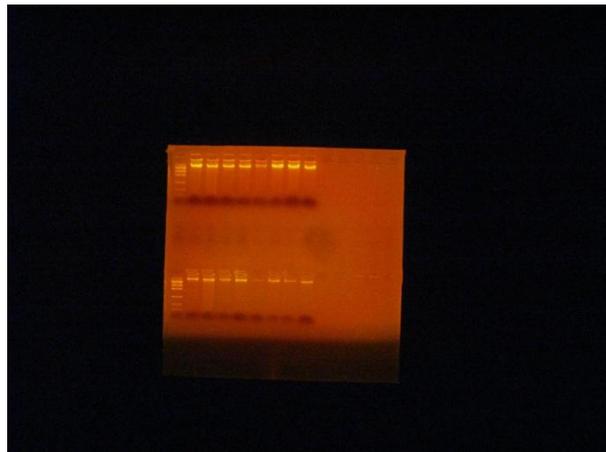
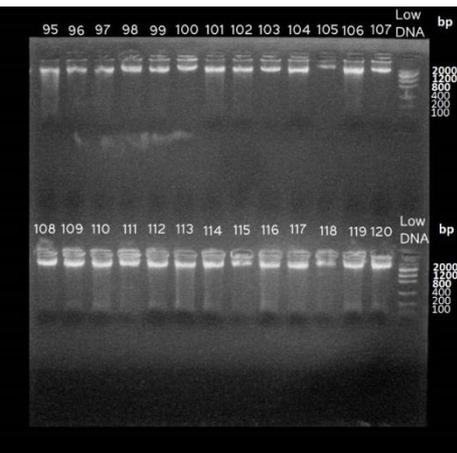
EDTA 2 mmol.L<sup>-1</sup> : 0,74 g

# Révélation de l'ADN

Après la migration d'électrophorèse, le gel est soumis à une lumière sous ultraviolet afin d'observer les bandes d'ADN fluorescentes.

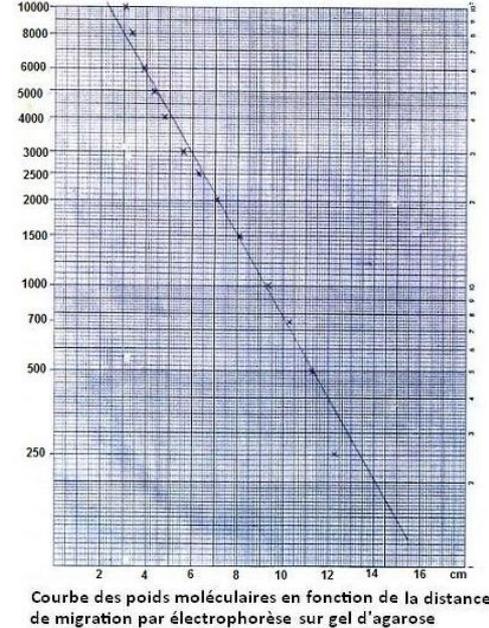
Le marqueur d'acide nucléique contenu dans le gel devient fluorescent avec une couleur rouge orangée.

L'estimation de la taille des fragments est faite à l'aide du marqueur de taille moléculaire (Low DNA Mass Ladder) utilisé simultanément dans un autre puits lors de la migration.



Pour déterminer le poids moléculaires inconnu des fragment d'ADN, on doit tracer la courbe des poids moléculaires de fragments d'ADN de poids connus (marqueurs de PM, DNA-ladder) en fonction de la distance parcourue) par les différents fragments.

La droite : la taille ou le poids = f (distance de migration) permet de déterminer la taille en paires de base d'un fragment d'ADN inconnu.



Courbe des poids moléculaires en fonction de la distance de migration par électrophorèse sur gel d'agarose

## 2. Spectrophotomètre (nano-drop)

Les acides nucléiques absorbent dans l'UV à  $\lambda_{\max} = 260 \text{ nm}$ . En mesurant l'absorbance d'une solution d'acide nucléique pure à 260 nm, on obtient sa concentration à l'aide de l'équation de la loi Beer-Lambert. En pratique, c'est une équation modifiée qui est employée :

$$c_m = (A.e) / l$$

$c_m$  = concentration massique d'acide nucléique en ng/ $\mu$ l

A = absorbance à 260 nm

e = coefficient d'extinction massique ou la constante en ng.cm/ $\mu$ l

l = trajet optique en cm (généralement 1 cm).

1 unité de DO = 50 ng/ $\mu$ l

En pratique, pour chaque séquence d'ADN/ARN, il n'est pas nécessaire de calculer le coefficient d'extinction ( $\epsilon$ ) spécifique. Des coefficients d'extinction estimatifs, acceptés par la communauté scientifique, sont généralement utilisés pour calculer les concentrations des acides nucléiques et qui sont suffisamment précises pour les expériences de biologie moléculaire. Ces coefficients d'extinction sont listés ci-après: ADN double brin =  $\sim 50 \text{ ng.cm}/\mu\text{l}$ ; ADN simple brin =  $\sim 33 \text{ ng.cm}/\mu\text{l}$ ; ARN =  $\sim 40 \text{ ng.cm}/\mu\text{l}$

Exemples:

1. On veut connaître la concentration d'un ADN en utilisant le spectrophotomètre. On a pris une quantité de l'ADN extrait, la densité optique affichée à 260nm est 0.025. Calculer la concentration de cet ADN.
2. 5 $\mu\text{l}$  d'ADN sont dilués dans 2 ml d'eau distillée stérile et sont passés au spectrophotomètre pour estimer la concentration de l'échantillon. Le chiffre indiqué à 260nm est de 0.020. Quelle est la concentration de cet échantillon ?

## Analyse de la pureté par spectrophotométrie

### **Le ratio A260 nm/280 nm:**

Ce ratio explique la pureté des acides nucléiques (évaluer la contamination de protéines dans une solution d'acides nucléiques).

D'une manière générale, la pureté d'une solution d'acide nucléique est considérée comme acceptable lorsque le ratio A260 nm/A280 nm est compris entre 1,8 – 2,0 pour l'ADN et entre 2,0 – 2,2 pour l'ARN.

### **Le ratio A260 nm/A230 nm:**

Ce ratio est un deuxième indicateur de pureté, il est généralement plus élevé que le ratio A260/A280. Un ratio A260/A230 compris entre 2,0 – 2,2 indique une bonne pureté des acides nucléiques. Lorsque ce ratio est significativement plus faible que 2,0 – 2,2, cela révèle la présence de contaminants absorbant à 230 nm dans la solution (le phénol, les carbohydrates, les peptides, l'acide humique, l'urée, l'EDTA, les polysaccharides ....). Leur présence est détectable grâce à ce ratio.

### 3. Le Qubit :

Son principe est la fluorométrie basée sur la mesure de la fluorescence des fluorophores spécifiques des acides nucléiques.

Mais des réactifs et standards dédiés sont nécessaires pour réaliser ces mesures sur le Qubit

contrairement à la spectrophotométrie, la fluorimétrie ne permet pas d'analyser la pureté et ni d'identifier des contaminants dans une solution d'acides nucléiques.



#### 4. La qPCR: la PCR quantitative ou PCR en temps réel

Son principe est de suivre la quantité d'ADN cycle après cycle grâce à un marqueur fluorescent dont l'intensité d'émission est proportionnelle à la quantité de produits.

On doit mettre le mélange réactionnel de la PCR une molécule émettrice de fluorescence (SyberGreen ou des sondes fluorescentes).

Le thermocycleur est branché à un ordinateur qui, avec un programme dédié va mesurer l'intensité de fluorescence cycle après cycle et il va tracer une courbe.

NB: Les auteurs recommandent d'utiliser en parallèle le NanoDrop et le Qubit (moins coûteux) ou le NanoDrop et la qPCR (plus coûteux), car ni le Qubit ni la qPCR ne permet la détection de contaminants présents dans les ADN purifiés. Les ratios A260/A230 et A260/A280 obtenus par le NanoDrop sont de bons indicateurs de pureté.

## Méthodes d'analyse du polymorphisme de l'ADN: marqueurs moléculaires

Le mot de **polymorphisme** signifie « plusieurs formes ».

Le polymorphisme génétique est l'existence, dans une population, de plusieurs états alternatifs de l'ADN, ou **allèles**, en une position définie du génome, ou *locus*. Cette définition a plusieurs aspects:

:

- (1) D'abord le trait doit être porté par les **chromosomes** et être transmissible.
- (2) Ensuite les allèles doivent être *homologues* pour leur position dans le génome.
- (3) Mais puisque le trait est transmissible, l'homologie de position implique aussi que les allèles soient homologues par descendance ; s'ils sont différents, c'est donc qu'une **mutation** (au moins) est intervenue.
- (4) Enfin, le polymorphisme génétique peut être défini à l'échelle de la plus petite unité composant l'ADN : le **site nucléotidique**. De ce fait, chaque variant nucléotidique peut déterminer le polymorphisme des niveaux supérieurs de structuration biologique (le gène, la protéine et le phénotype de l'individu).

Au sein du génome, différents phénomènes plus ou moins complexes sont responsables du polymorphisme:

1. Des mutations dues à un remplacement d'une ou plusieurs bases par d'autres sans le changement de la longueur de l'ADN (substitution)
2. Des insertions ou des délétions correspondant à l'ajout ou à la disparition de bases dans la séquence d'ADN (changement de la longueur de l'ADN) (pour le RFLP: perte ou gain d'un site de restriction)
3. Des réarrangements chromosomiques dus à des recombinaisons génétiques ou à l'insertion de transposons. Les recombinaisons génétiques sont la conséquence d'une cassure de l'ADN suivie d'un échange de matériel entre deux chromosomes.

Pour étudier le polymorphisme, des marqueurs sont utilisés:

- Marqueurs morphologiques
- Marqueurs moléculaires
- Marqueurs biochimiques (à voir dans le polymorphisme protéique)

**Les marqueurs morphologiques (phénotypiques):** Caractères morphologiques ou phénotypiques (couleur des poils, capacité à pousser en milieu carencé, présence ou absence de taches foliaires, etc...).

Ce sont des caractères faciles à mesurer mais peu polymorphes, rarement co-dominants (généralement dominants), sont influencés par l'environnement, sont dépendants souvent du stade de développement de la plante et souvent ne sont pas reproductible sous différents environnements.

Un marqueur moléculaire est un locus **polymorphe** qui renseigne sur le génotype de l'individu qui le porte.

Les marqueurs moléculaires permettent à la fois un diagnostic extrêmement **fin** de la variabilité et la mise en place de stratégies très rapides de création et sélection variétale.

Ces marqueurs présentent également différents avantages comparés aux marqueurs morphologiques et protéiques:

- très nombreux et très polymorphes
- neutres vis à vis de la sélection,
- couvrent le génome entier,
- indépendants de la partie de la plante prélevée, du stade de développement et des fluctuations environnementales.

Un marqueur génétique peut être défini comme **un gène ou une séquence polymorphe** d'ADN , facilement détectable grâce à son emplacement connu sur le chromosome. Ces marqueurs peuvent être utilisés pour différencier les individus, les espèces, les genres, les variétés,...

Ces marqueurs moléculaires deviennent aujourd'hui un outil essentiel d'amélioration des plantes et ouvrent de nouvelles perspectives pour le sélectionneur.

Un marqueur moléculaire, devrait idéalement présenter les caractéristiques suivantes:

1. Être **polymorphe**, c'est à dire posséder plus d'un allèle dans la population étudiée (variable entre individus).
2. Être **co-dominant**, ce qui signifie qu'un hétérozygote peut être différencié des homozygotes c'est-à-dire absence d'interaction intra locus.
3. Non **épistatique**, c'est-à-dire que son génotype peut être lu à partir de son phénotype quel que soit le génotype des autres locus (absence d'interaction inter locus).
4. Être **neutre**, c'est-à-dire que son polymorphisme est sans effet sur la variabilité phénotypique.

5. Être **insensible au milieu**, c'est-à-dire le génotype peut être inféré à partir du phénotype quel que soit le milieu.
6. **Reproductible** d'une expérience à une autre

Les marqueurs moléculaires peuvent se diviser en deux grands groupes, selon leur principe:

1. Les marqueurs basés sur l'hybridation moléculaire: RFLP
2. Les marqueurs basés sur l'amplification enzymatique *in vitro* (la PCR): SSR, AFLP, RAPD

# 1. Les marqueurs basés sur l'hybridation moléculaire: RFLP pour Restriction Fragment Length Polymorphism

## Polymorphisme de longueur des fragments de restriction

Le principe de base de cette méthode repose sur la comparaison de profils de coupure par les enzymes de restriction suite à l'existence d'un polymorphisme dans la séquence d'une molécule d'ADN par rapport à une autre.

Des mutations pouvant apparaître sur une séquence d'ADN reconnue par une enzyme de restriction provoquent des longueurs de fragments de restriction différents.

L'ADN des individus à comparer est donc digéré par une ou plusieurs enzymes de restriction.

Les produits de restriction sont ensuite séparés par électrophorèse sur gel d'agarose ou d'acrylamide en présence de marqueurs de taille ou de poids.

Le polymorphisme observé est utilisé comme critère d'identification

Par cette approche, deux individus pourront présenter des profils de restriction différents.

## Les enzymes de restriction:

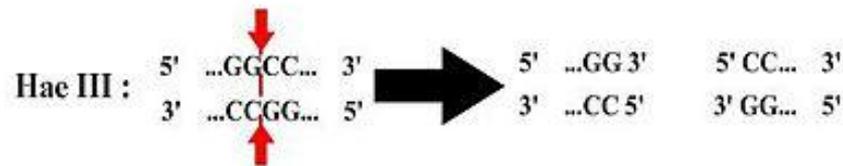
Ce sont des enzymes, principalement d'origine bactérienne, capables de couper l'ADN double brin à des endroits spécifiques (séquences nucléotidiques) appelés sites de restriction généralement de 4 à 6 paires de bases.

Ces sites sont de **nature palindromique**, c'est-à-dire que la lecture des deux brins complémentaires dans **des sens opposés donne la même séquence**.

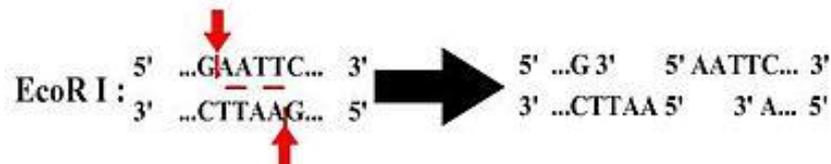
**Elles sont des endo-nucléases.**

Il y a deux catégories d'endo-nucléases : celles qui donnent des **extrémités franches** et celles qui donnent **des extrémités cohésives**.

Extrémités non cohésives :



Extrémités cohésives :



La nomenclature des enzymes de restriction est précise. Le nom comporte 3 ou 4 lettres correspondant entre autres à la bactérie à partir de laquelle on a extrait cette enzyme:

- La première lettre est en majuscule cela correspond à la première lettre du genre de la bactérie
- La seconde et troisième lettre sont en minuscule et correspondent aux deux premières lettres de l'espèce bactérienne
- La quatrième lettre correspond à la souche bactérienne, elle est écrite en majuscule mais n'est pas présente dans toutes les enzymes de restriction
- Enfin un chiffre romain donne le numéro d'ordre de caractérisation de l'enzyme

**EcoRI** : Enzyme de restriction d'*Escherichia* (genre) *coli* (espèce) souche RY317 la première à être caractérisée (I)

Types d'enzymes de restriction :

- I. Type I: elles reconnaissent une séquence d'ADN, puis se déplacent, elles s'arrêtent à 1000 à 5000 paires de bases plus loin
- II. Type II: coupe l'ADN au niveau de la séquence
- III. Type III: elles coupent une vingtaine de nucléotides plus loin que le site de reconnaissance.

Seules les enzymes de type II sont utilisées au laboratoire.

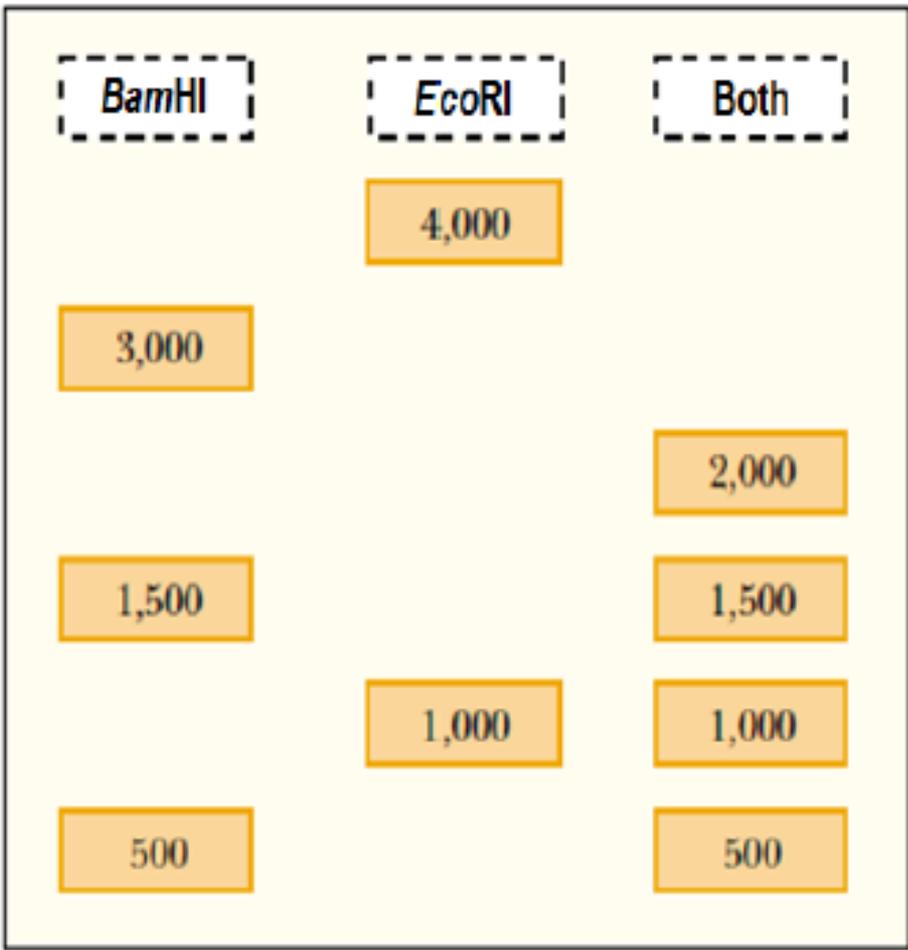
<b>Enzyme</b>	<b>Origine bactérienne</b>	<b>Site de restriction</b>
Alu I	<i>Athrobacter luteus</i>	AG/CT
Bam HI	<i>Bacillus amyloliquefaciens</i>	G/GATCC
Bgl II	<i>Bacillus blobiggi</i>	A/GATCT
Eco RI	<i>Escherichia coli</i>	G/AATTC
Hind III	<i>Haemophilus influenzae</i>	A/AGCTT
Hea III	<i>Haemophilus aegyptus</i>	GG/CC
Hpa II	<i>Haemophilus parainfluenzae</i>	C/CGG
Mbo I	<i>Moraxella bovis</i>	GA/TC
Pst I	<i>Providentia stuartii</i>	CTGCA/G
Taq I	<i>Thermophilus aquaticus</i>	T/CGA
Wba I	<i>Xanthomonas badrii</i>	T/CTAGA

Voici la séquence d'ADN suivante:

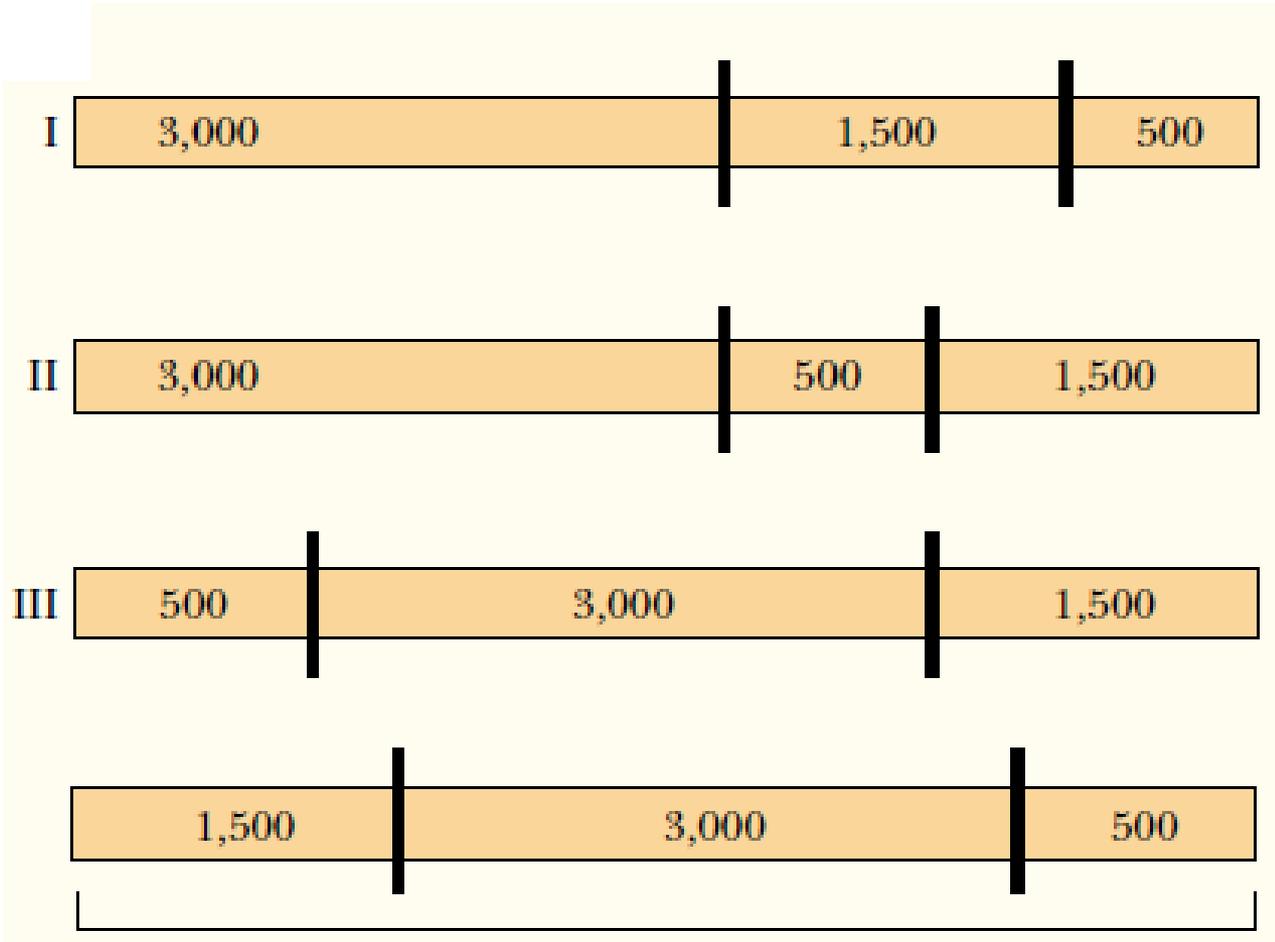
ATTTGGTGACACTATAGAATACACGGAATTCGAGCTCGATC  
TAAACCACTGTGATATCTTATGTGCCTTAAGCTCGAGCTAG

# Utilisation des endo-nucléases de restriction pour établir la mappe de restriction (Carte de restriction)

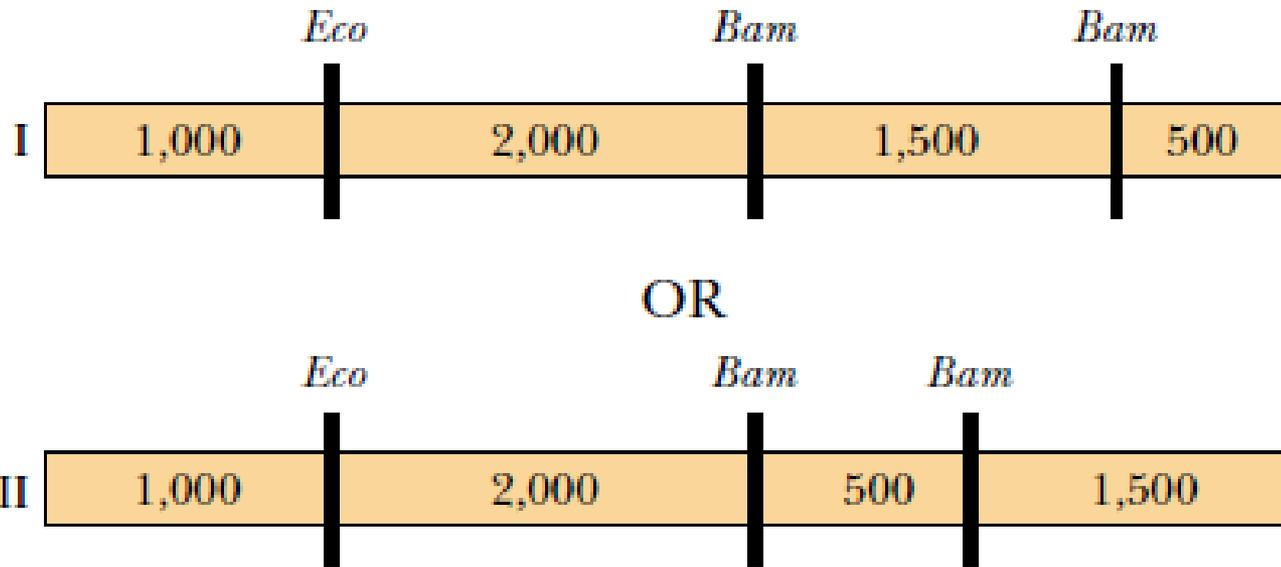
La figure au dessous montre la digestion d'un fragment d'ADN de 5 kb par deux enzymes de restriction (*BamHI* et *EcoRI*).



L'analyse des résultats de la digestion nous permet de déterminer les différentes possibilités de la localisation des sites de restriction pour chaque enzyme.



Les différentes possibilités de la localisation des sites de coupure par *Bam*HI.



Les différentes possibilités de localisation des sites de restriction des deux enzymes (*EcoRI*, et *BamHI*).

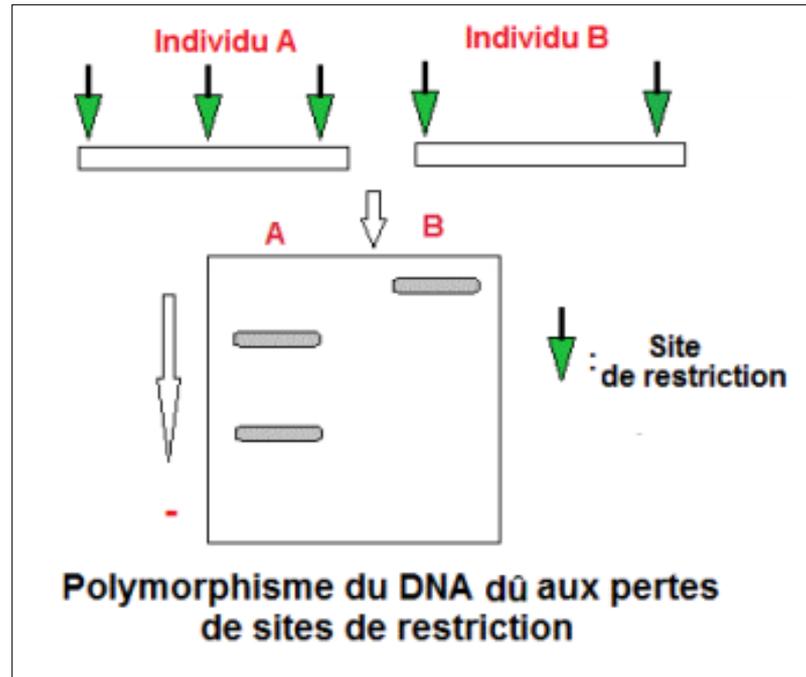
## causes du polymorphisme dans les RFLP

Les causes du polymorphisme sont :

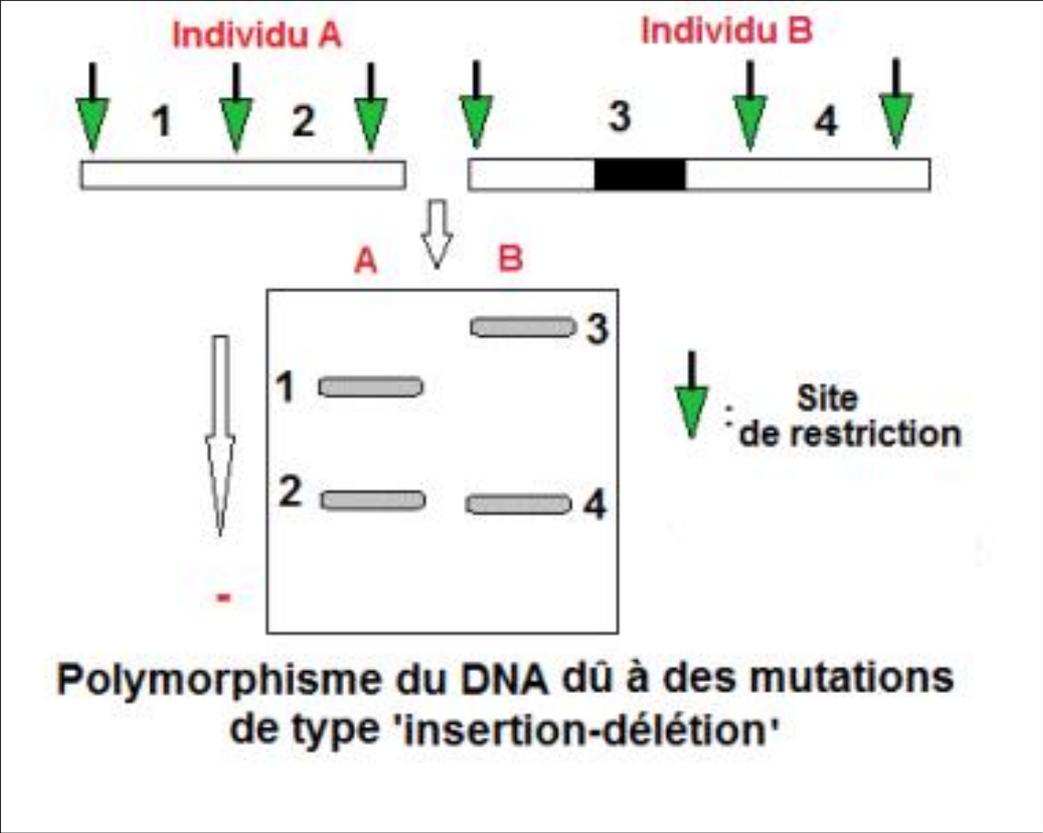
- perte ou gain d'un site de restriction
- mutation de type insertion-délétion

Si deux individus diffèrent par un ou plusieurs sites ou, même, par la distance séparant deux sites identiques consécutifs, il se crée une différence dans la longueur des fragments générés par l'enzyme de restriction.

# Polymorphisme du à la disparition d'un site de restriction

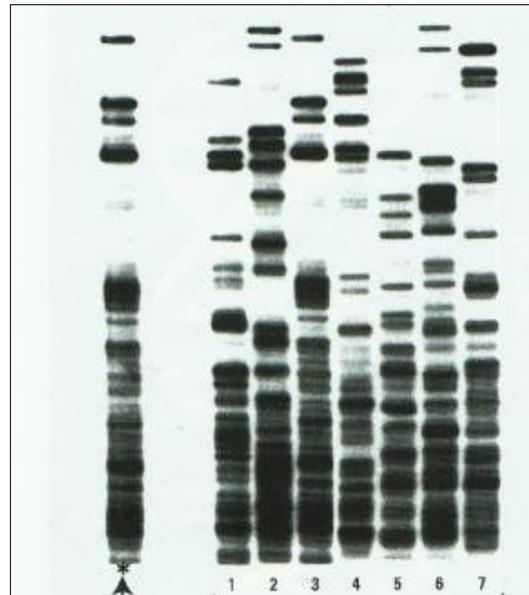


Polymorphisme du aux mutations de type 'insertion - délétion'.



La technique RFLP repose sur la digestion d'un ADN cible par une ou plusieurs enzymes de restriction spécifiques des sites de restriction portés par l'ADN.

Il est déraisonnable de chercher à caractériser un génome complet par les fragments de restriction obtenus puisque même pour un génome relativement petit, une ER donnera des centaines de fragments qui, par électrophorèse en gel d'agarose ne seront pas séparés et donneront une traînée: un seul voile continu (smear).



Pour être plus fiable, un RFLP doit être identifiable dans une région précise du génome.

Après électrophorèse, une hybridation est nécessaire (Southern blotting).

## Southern Blotting : Hybridation moléculaire ADN/ADN

Cette méthode a été initialement décrite par E.M. **Southern** 1975.

Blotting = Transfert

Elle consiste à détecter **spécifiquement** des fragments d'ADN transférés sur filtre par leur hybridation à une sonde (des séquences complémentaires marquées par radioisotope.)

**Sonde:** petit morceau d'ADN synthétisé dont la séquence correspond au moins en partie à celle d'un des brins de l'ADN.

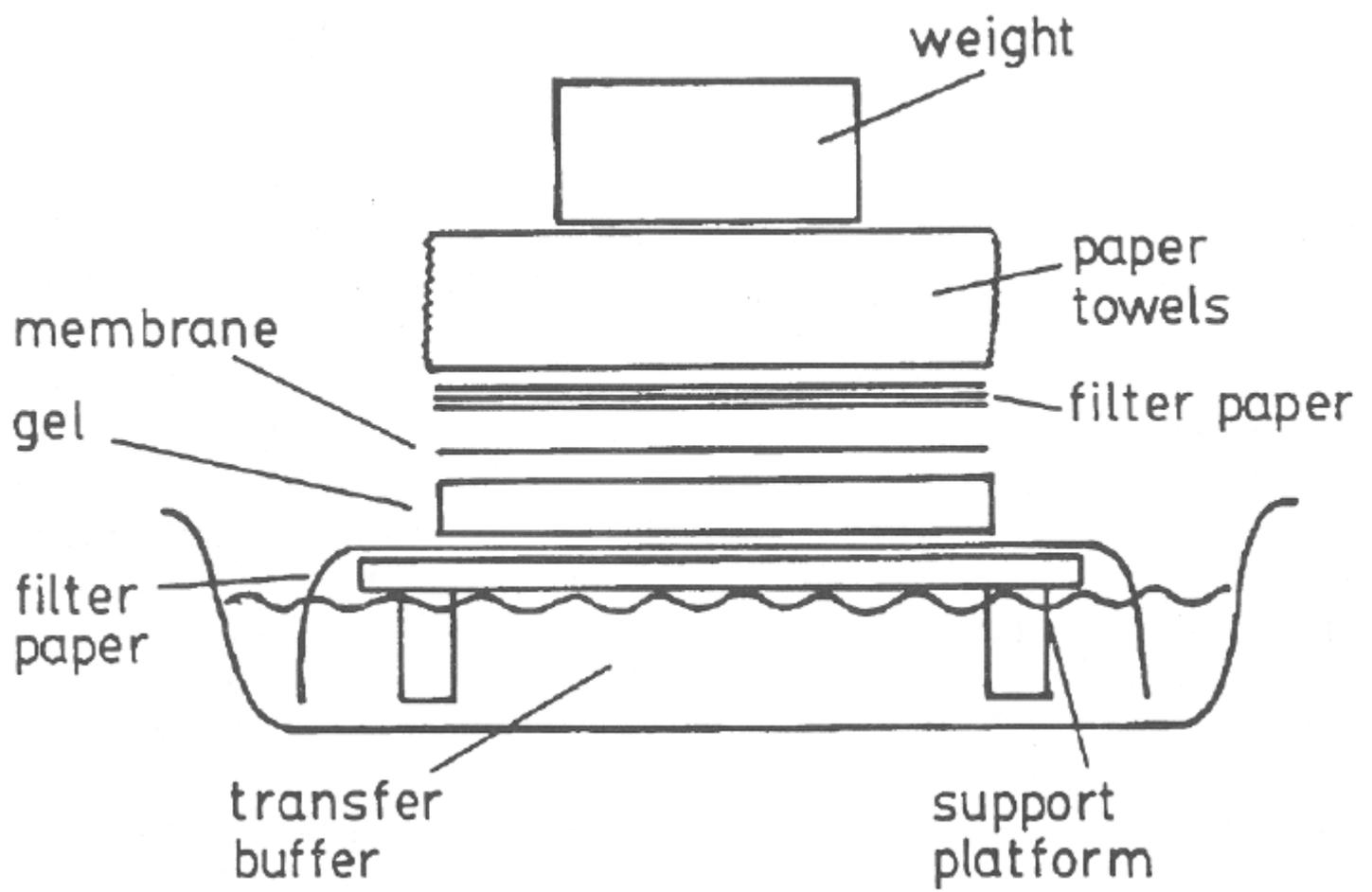
En marquant cette sonde par un atome radioactif ( $P^{32}$ ) ou par une molécule fluorescente liés covalentiellement à un des nucléotides de la sonde, on peut détecter la sonde et le morceau de l'ADN complémentaire qui lui est hybridé.

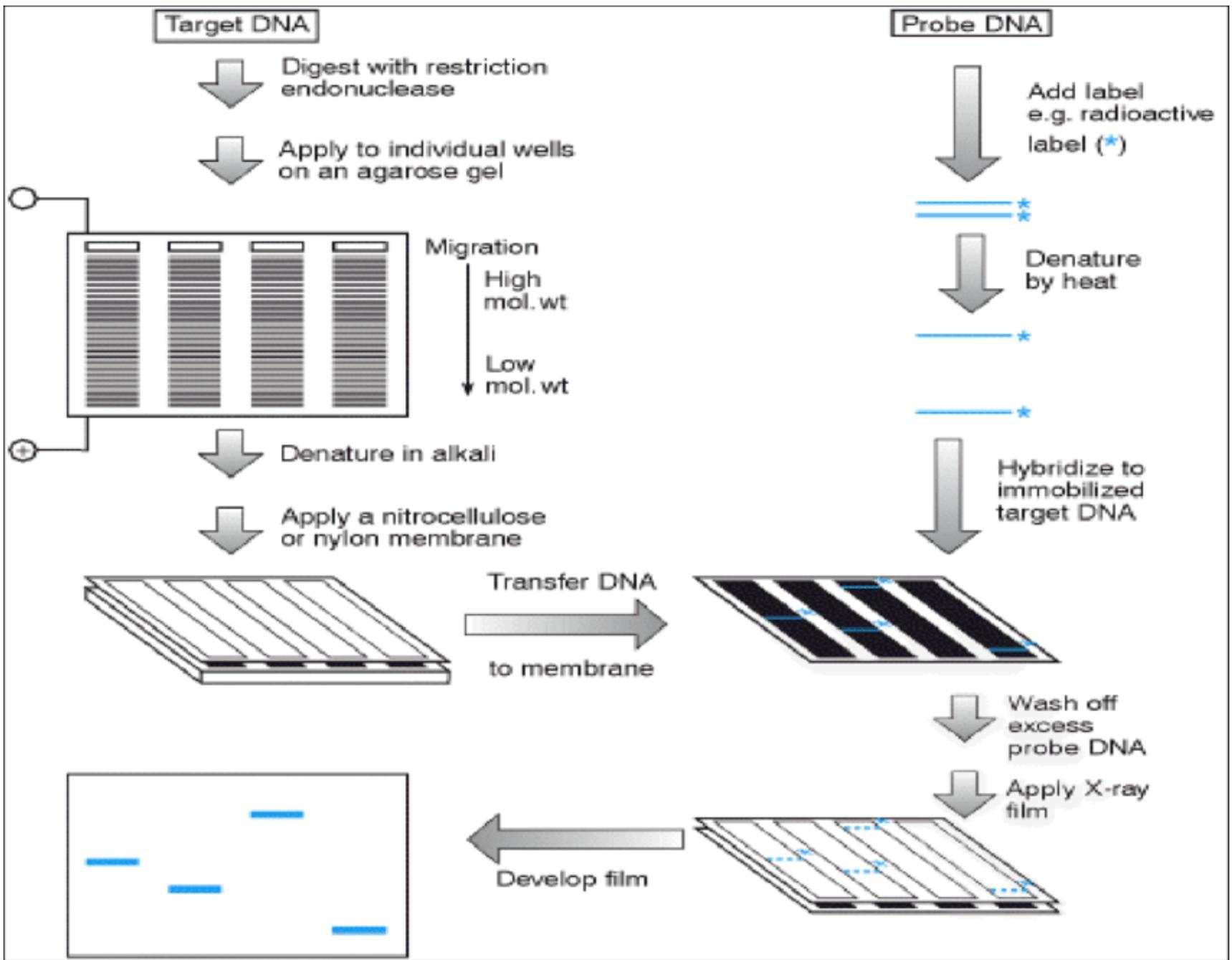
Les ADN cibles sont dénaturés puis transférés par capillarite sur une membrane de nitrocellulose ou de nylon.

Les fragments séparés sont hybridés avec un ADN sonde.

Une fois transféré sur une membrane de nylon ou de nitrocellulose et hybridé par un ADN sonde, l'ADN cible digéré est visualisé par autoradiographie.

Après mise en contact de la membrane avec la sonde marquée dans des conditions favorables à l'hybridation, des lavages sont effectués afin de faire apparaître après autoradiographie des bandes marquées correspondant à des fragments d'ADN complémentaires de la sonde.





## Récapitulatif des étapes

- 1- clivage enzymatique de restriction
- 2- électrophorèse sur gel d'agarose
- 3- dénaturation dans le gel des fragments de restriction
- 4- transfert sur support solide (membrane de nylon ou nitrocellulose) par capillarité (buvardage)
- 5- fixation de l'ADN sur la membrane par cuisson (mb. nitrocellulose) ou exposition U.V (mb. nylon)
- 6- hybridation avec une sonde marquée dénaturée
- 7- lavages
- 8- autoradiographie

## Les marqueurs moléculaires basés sur la PCR

Le développement de la technique PCR (*Polymerase Chain Reaction*) offre l'avantage d'analyser les marqueurs moléculaires en un temps court tout en utilisant des concentrations faibles d'ADN. Les marqueurs basés sur la technique PCR tendent à remplacer les systèmes classiques (les marqueurs morphologiques, iso-enzymatiques et RFLP) et deviennent très nombreux.

Le principe de la PCR étant fondé sur l'amplification de l'ADN, cette technique ne nécessite qu'une très petite quantité d'ADN. La réaction de polymérisation en chaîne a aussi l'avantage de ne pas dégrader définitivement l'ADN, contrairement à la RFLP.

## Réalisation d'une PCR

Une PCR classique se déroule dans un petit tube lui même placé dans un appareil programmable appelé thermocycleur. Ce dernier se contente de placer le tube aux températures voulues pendant les durées programmées et de recommencer en effectuant des cycles. Un cycle reproduit **trois températures différentes** pendant des durées différentes. La durée d'un cycle est de l'ordre de la minute.

Avant la réaction, tous les « acteurs » de la PCR sont introduits dans le même tube. Il s'agit de:

- l'ADN à amplifier,
- des oligonucléotides (ou amorces), spécifiques du segment d'ADN voulu (on parle de couple d'amorces)
- de l'ADN polymérase (ici la Taq polymérase)
- Les ions  $Mg^{2+}$  pour maintenir stable le pH du milieu
- et enfin du mélange des quatre désoxyribonucléotides constitutifs de l'ADN.

Tous sont ajoutés en large excès par rapport à l'ADN.

Etapas de la PCR:

-**Dénaturation** : l'ADN à amplifier est dénaturé à 95°C, les deux brins se séparent et servent de matrice.

-**Hybridation des amorces** : on baisse la température (50-60°C) pour permettre à l'amorce de s'hybrider spécifiquement aux extrémités de la séquence cible à amplifier.

-**Elongation ou polymérisation à 72°C** : à cette température optimale de l'activité de l'enzyme Taqpolymérase, l'amorce est allongée par complémentarité au brin matrice, ce qui produit deux nouvelles molécules d'ADN.

Après n cycle, on obtient  $2^n$  molécules d'ADN identiques.

Il faut un couple d'amorces : Forward et Reverse

Chaque amorce doit :

- Contenir 17 à 28 nucléotides (environ 20)
- % GC: plus de 50 % et moins de 68 %
- Extrémité 3' : de préférence se termine par GC/CG/CC/GG
- Th (température d'annealing) comprise entre 55 et 65 °c

$$T_h = T_m - 5$$

$$T_m = 4(G+C) + 2(A+T)$$

Température de fusion d'une amorce est la température à laquelle la moitié de l'ADN bicaténaire devient monocaténaire

- Des séquences qui ne soient pas complémentaires entre elles
- Éviter les palindromes
- Équilibre des bases :

20-30 % G; 20-30 % C; 20-25 % A; 20-25 % T;

## Comment dessiner une amorce

Soit la molécule d'ADN suivante :

5' A G A A T C C G T A G C.....T A G C A C T 3'  
3' T C T T A G G C A T C G .....A T C G T G A 5'

L'amorce forward : sens 5'....3' : c'est la plus facile à dessiner  
C'est celle qui va se fixer sur le brin anti sens

L'amorce reverse : anti-sens 5'.....3'  
C'est celle qui va se fixer sur le brin sens

exemple : TCTGGATGTT GTGAGGAAGG **AAGCTGAGAA**  
**CTGTGACTGC TTGCAAGG**AT TCCAAGTATG CCACTCCCTT  
GGTGGTGGTA CTGGATCTGG TATGGGTACG CTGTTGATCT  
CAAAGATCAG

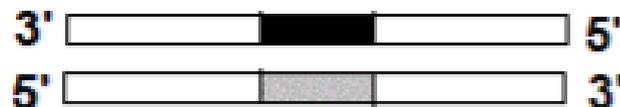
.....  
TTGTGGTCCC TAGGGCAAGC GGAC**CCTCGAT**  
**GAGTTCGGTG TTC**CCTTTCG TGTTGTTGCC

Les amorces doivent faire au moins 20 nucléotides  
séquences possibles:

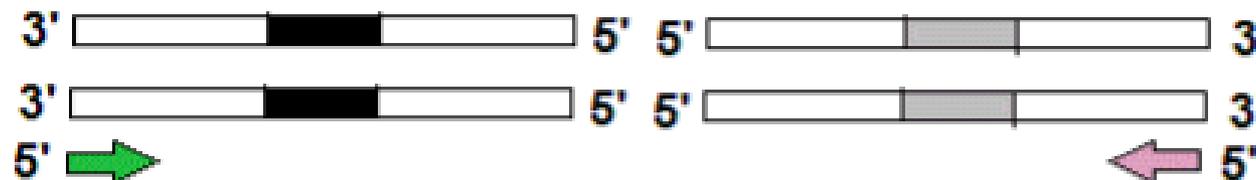
5'- AACTGTGACTGCTTGCAAGG -3'

5'- GAACACCGAACTCATCGAGG -3'

# Locus



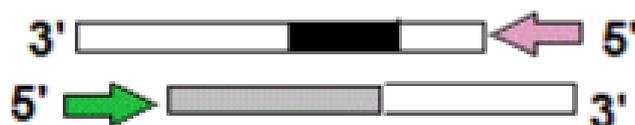
**Dénaturation**  
(90-95°C)



**Fixation des amorces (annealing)**  
(35-55°C)



**Polymérisation (extension)**  
(70-75°C)



**n cycles**  
(n = 25-50)



**Electrophorèse**  
**Deux amorces**

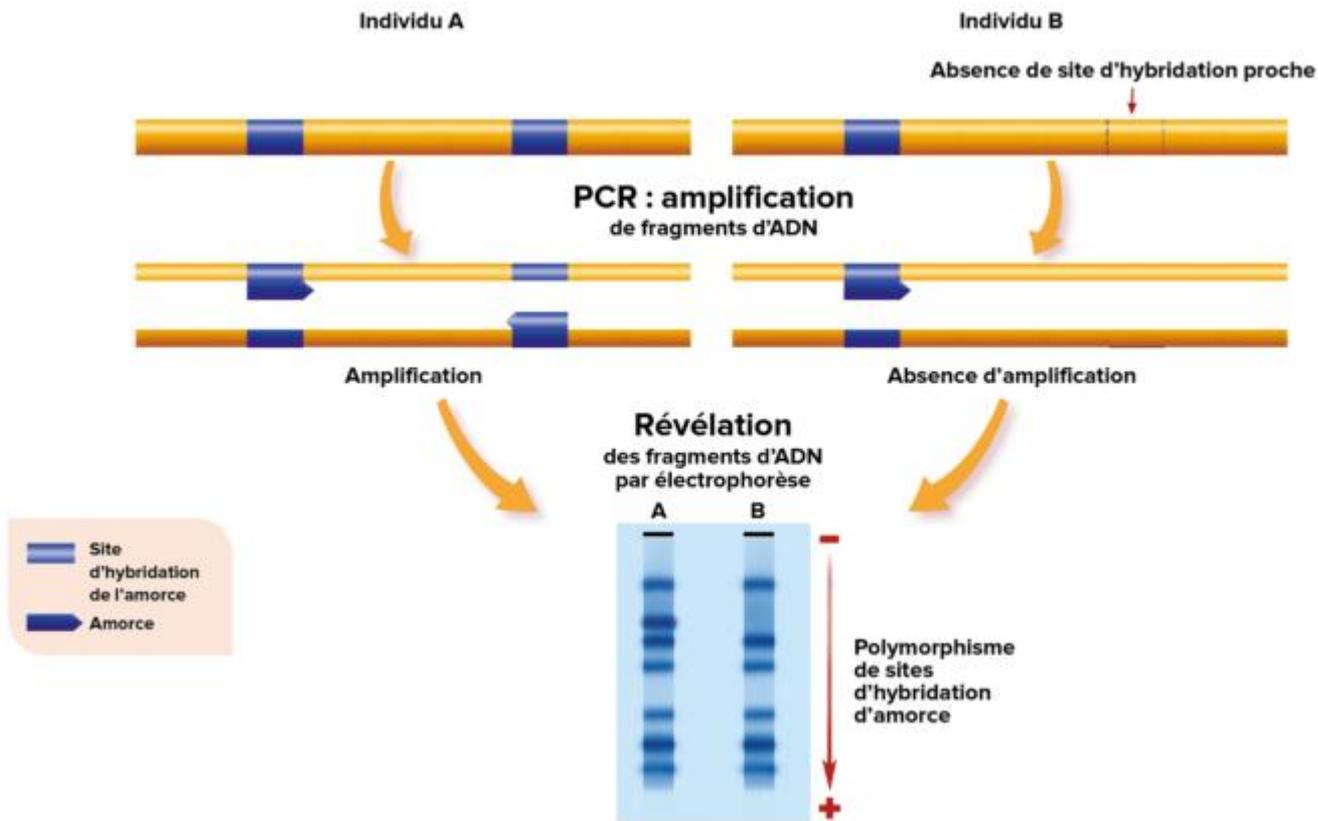
**Polymérisation en chaîne de l'ADN**

RAPD : **R**andom **A**mplified **P**olymorphism **D**NA

Polymorphisme d'amplification aléatoire de l'ADN

Cette technique consiste à réaliser une PCR en utilisant une amorce courte d'une dizaine de nucléotides, de séquence **arbitraire**. Cette amorce va s'hybrider **au hasard** dans le génome. C'est une technique qui ne nécessite pas obligatoirement de connaître la séquence de l'ADN cible. Les amorces utilisées seules ou par couples. Lorsqu'une seule amorce aléatoire est utilisée, l'amplification a lieu une fois, celle-ci se fixe sur 2 sites complémentaires peu distants (3000 pb au max)

# Les marqueurs RAPD



Amorce de séquence arbitraire ⇒ Marqueur RAPD

Le polymorphisme révélé est un polymorphisme de sites d'hybridation d'amorce. Les amorces constituent donc les marqueurs. Pour l'ensemble du génome, une dizaine de fragments sont amplifiés en moyenne puis séparés par électrophorèse.

Cette méthode ne nécessite pas de digestion par une enzyme de restriction, pas de transfert sur une membrane, pas de préparation de sonde radioactive. Elle est donc rapide et d'une faible technicité. Toutefois, sa reproductibilité est difficile à obtenir.

## Origines de polymorphisme de type RAPD :

Les causes du polymorphisme de type RAPD sont:

- Disparition par mutation (mutation ponctuelle, crossing-over, insertion, délétion, ...) d'un ou plusieurs sites de fixation de l'amorce.
- Eloignement, par insertion, des sites de fixation de l'amorce à une distance supérieure à 3000 pb. Il en résulte la disparition de fragments de DNA.
- Rapprochement, par délétion, des sites de fixation de l'amorce à une distance inférieure ou égale à 3000 pb. Il en résulte l'apparition de fragments surnuméraires.
- Rapprochement très accusé des sites de fixation de l'amorce ne donnant lieu qu'à de petits fragments de DNA n'apparaissant pas sur le gel après électrophorèse.

## AFLP (Amplified Fragment Length Polymorphism)

### POLYMORPHISME D'AMPLIFICATION DE FRAGEMENT D'ADN

La technique AFLP (Amplified Fragment Length Polymorphism, Vos et al., 1995) est fondée sur la mise en évidence conjointe de polymorphisme de sites de restriction et d'hybridation d'amorces arbitraires.

Cette technique utilise à la fois les enzymes de restriction et l'amplification PCR. On distingue les étapes suivantes :

-L'ADN génomique est coupé par deux enzymes de restriction.

-Des adaptateurs de **séquences connues** et spécifiques des enzymes de restriction utilisées, sont ajoutés aux extrémités des fragments de restriction générant ainsi une matrice pour l'amplification.

-Une première amplification (pré-amplification) est réalisée à l'aide d'amorces de séquences complémentaires à la séquence des adaptateurs et des sites de restriction.

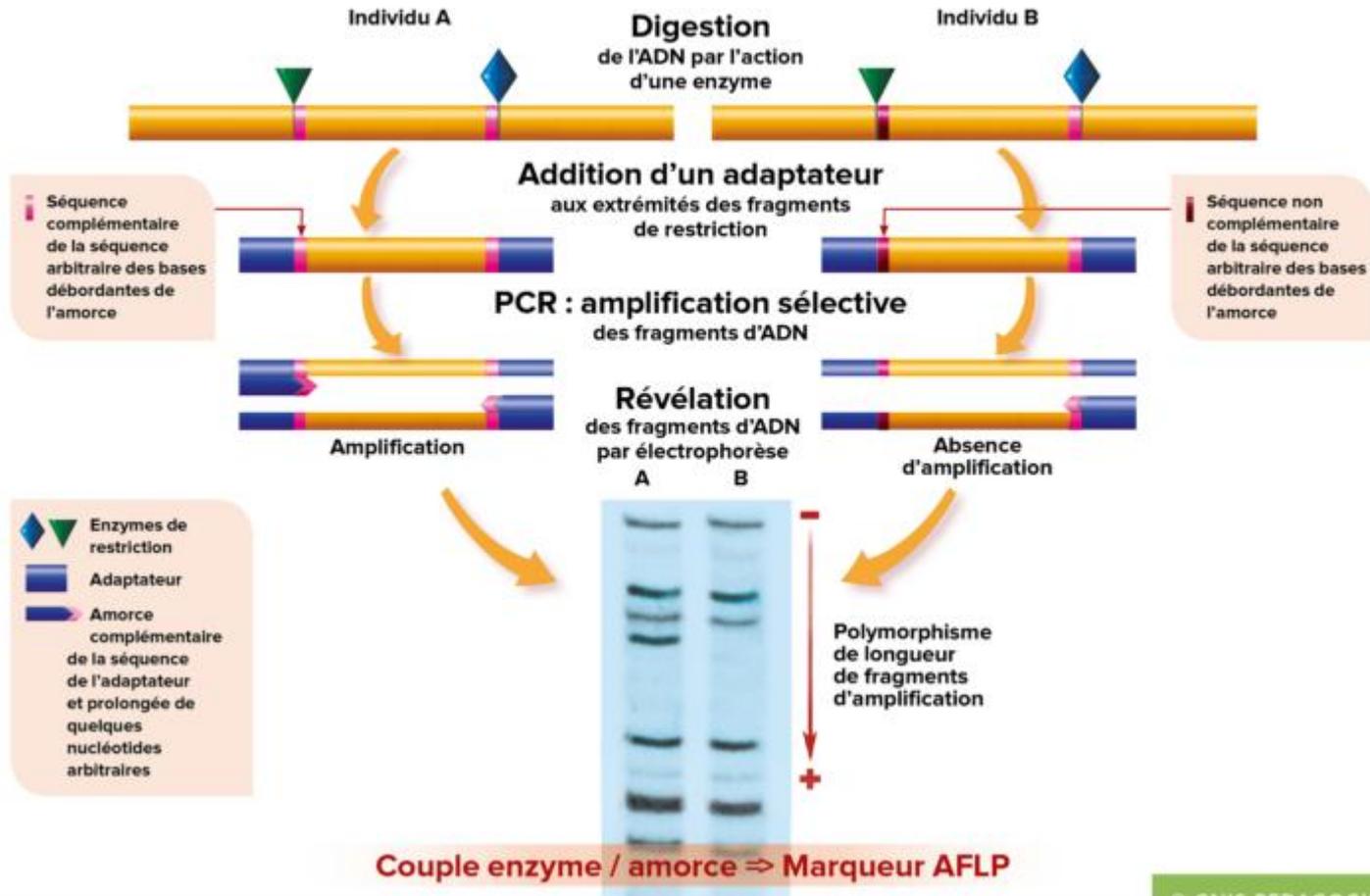
-La deuxième amplification (amplification sélective) utilise des amorces identiques aux premières mais **prolongées** à l'extrémité 3' de quelques nucléotides arbitraires (de 1 à 3 nucléotides).

-Ainsi, seuls sont amplifiés les fragments possédant les bases complémentaires de ces bases arbitraires. Ces amorces sélectives permettent de réduire le nombre de fragments amplifiés à une centaine.

- Ces fragments sont séparés par électrophorèse sur gel d'acrylamide dénaturant puis visualisés par coloration au nitrate d'argent ou révélés grâce à un marquage radioactif ou fluorescent réalisé lors de l'amplification sélective.

C'est la combinaison enzyme de restriction/amorce qui permet de révéler le polymorphisme entre les individus. Celle-ci constitue le marqueur AFLP. Le locus mis en évidence dépend de la séquence du site de l'enzyme de restriction et des bases arbitraires. Il existe de très nombreuses combinaisons enzyme/amorce. Il existe plusieurs centaines d'enzymes de restriction, mais une dizaine est généralement utilisée.

# Les marqueurs AFLP



## Marqueurs de type 'microsatellites' (SSR)

Les microsatellites ou SSR (Simple Sequence Repeats) (Morgante, Olivieri, 1993) sont des marqueurs moléculaires basés sur la PCR. C'est le **Polymorphisme de nombre d'unités de répétition**.

Sur le génome, il existe des séquences constituées d'unités répétées de 1 à 4 nucléotides. Ce sont les microsatellites. Les plus courants sont  $(A)_n$ ,  $(TC)_n$ ,  $(TAT)_n$  et  $(GATA)_n$ , les valeurs de  $n$  pouvant aller de quelques unités à plusieurs dizaines. On parle de séquences répétées en tandem ou SSR (Simple Sequence Repeats). L'intérêt de ces microsatellites réside dans leur polymorphisme. Celui-ci repose sur la variation du nombre d'unités de répétition, constituant le microsatellite.

Les SSR sont des séquences

**Les microsatellites sont des marqueurs multi-alléliques**

**Allèle 1**                      **9 répétitions CA**

---

...nnnnnnnnCACACACACACACACAnnnnnnnnnnn...

**Allèle 2**                      **13 répétitions CA**

---

...nnnnnnnnnnCACACACACACACACACACACAnnnnnnnnn...

**Allèle 3**                      **11 répétitions CA**

---

...nnnnnnnnnnnnCACACACACACACACACACAnnnnnnnnn...