

Bonjour!





# Une introduction à la biologie moléculaire: présentation et utilisation des techniques actuelles en recherche et en enseignement

Laboratoire Reproduction et Développement des Plantes  
Site Monod, ENS de Lyon

**Patrice Morel**, CR INRA

[patrice.morel@ens-lyon.fr](mailto:patrice.morel@ens-lyon.fr)





## Equipe Evolution et Développement de la Fleur :

Objectif de recherche menée :  
 “contribuer à une meilleure compréhension des origines moléculaires et de la diversification de la fleur et des organes qui la composent au cours de l'évolution”

RDP : <http://www.ens-lyon.fr/RDP/>

### Au RDP:

Des plantes modèles : *Arabidopsis* et *Petunia*.

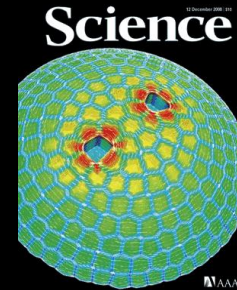
Des espèces d'intérêt agronomique : Rose et Maïs.



# La recherche au RDP:

## Recherche fondamentale :

augmentation de la connaissance  
→ publications scientifiques



## Recherche appliquée (objectifs à long terme) :

agir sur le développement des organes floraux:

Carpelles: développement des organes femelles.

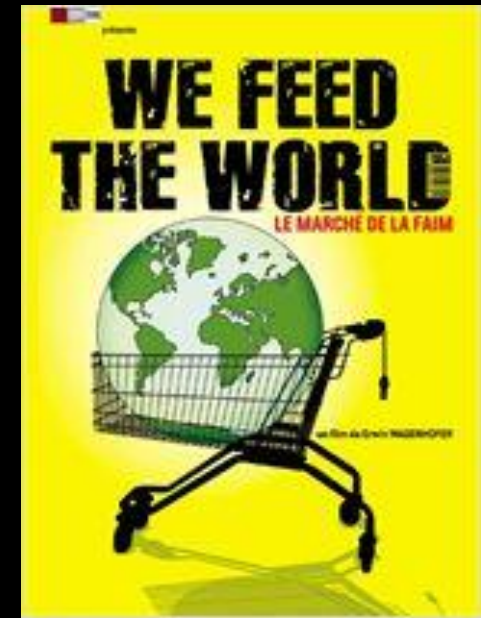


Pétales: projet *dodo*/fleur double

Petunia



Rose



La biologie moléculaire:  
présentation et utilisation des techniques  
actuelles en recherche et en enseignement

Deux parties étroitement imbriquées:

1. Projet de recherche

2. Description des techniques utilisées

Sauvage



Mutants



Projet de recherche: création et utilisation d'une collection de séquences flanquantes du transposon dTph1 chez *Petunia hybrida* W138

<http://www.ens-lyon.fr/RDP/Evolution-et-Developpement-de-la>

## **Plan:**

### **Introduction :**

- **Pourquoi utiliser *Petunia*?**
- **Pourquoi créer une collection de séquences flanquantes de *Petunia*?**

**I. Création d'une collection de séquences flanquantes d'un transposon endogène de *Petunia***

**II. Etude des gènes de la famille *AP1* de *Petunia* et utilisation de la collection de séquences flanquantes.**

# Introduction

*P. hybrida*



*A. thaliana*

## Pourquoi il est judicieux d'utiliser *Petunia* comme modèle...

- 3<sup>ème</sup> marché mondial en fleurs ornementales
- Facile à manipuler
- Facile à transformer (création d'OGMs),
- Possède un système de transposon **naturel** très actif,
- Le génome est presque entièrement connu,
- 3 mois seulement sont nécessaires de graine à graine,

Et comme on va le voir:

- > 230.000 séquences flanquantes de transposon différentes ont été obtenues...

→ Autant de mutants potentiels

<http://www.petuniaplatform.net/index.php?page=activities>



# Pourquoi il est judicieux de créer une collection de séquences flanquantes de transposon chez *Petunia*...

Exemple d'*Arabidopsis thaliana*.

<http://www.arabidopsis.org/>

Plante modèle de la biologie végétale.

Est au monde végétal ce que *Escherichia coli* est au monde bactérien

Au départ « forward genetic »: du phénotype au gène.



Isolement de centaines de mutants par simple criblage visuel:



*agamous*



*leunig*



*dwarf*



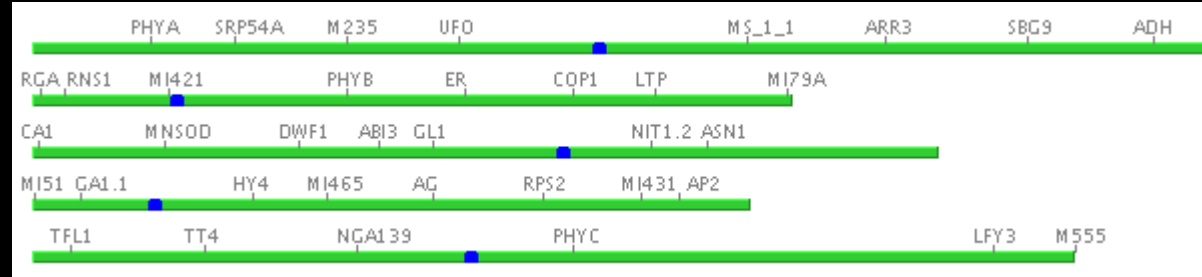
*pin*



*cauliflower*

# Séquençage de *Arabidopsis thaliana* : <http://www.arabidopsis.org/servlets/sv>

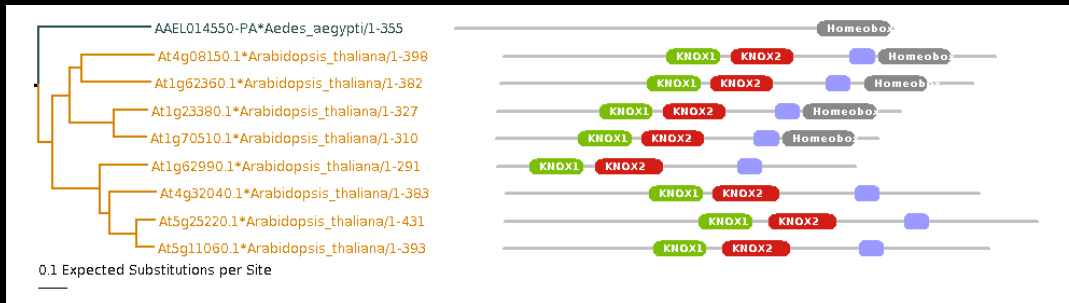
- 5 chromosomes
- 31.128 gènes



Analyses montrent que ces gènes peuvent :

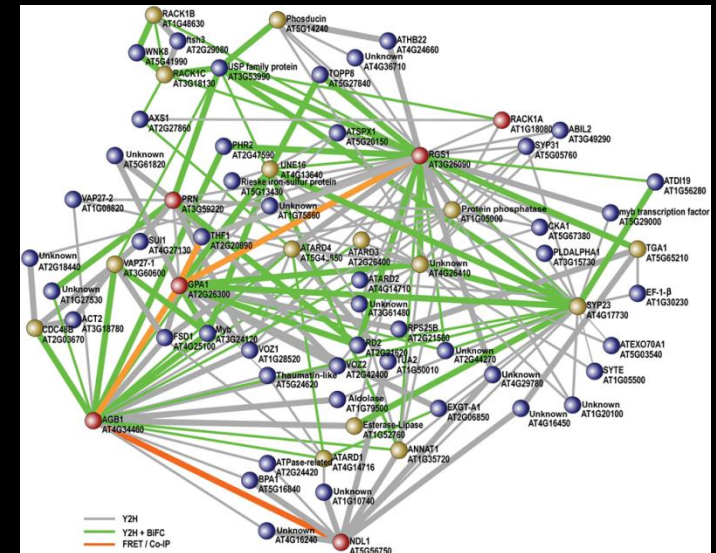
- être apparentés : <http://www.arabidopsis.org/browse/genefamily/index.jsp>
- coder pour des protéines qui interagissent...

[http://bar.utoronto.ca/interactions/cgi-bin/arabidopsis\\_interactions\\_viewer.cgi](http://bar.utoronto.ca/interactions/cgi-bin/arabidopsis_interactions_viewer.cgi)

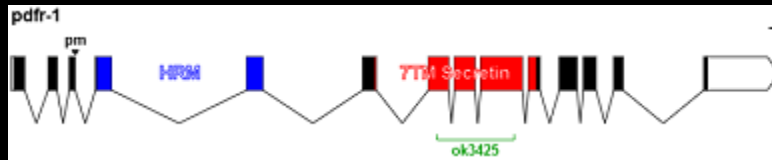


D'où le besoin de trouver des mutants :

- épistasie
- synergie
- antagonisme...



Ensuite « reverse genetic »: du gène au phénotype.



Quel effet?  
Quel phénotype?

But : créer un mutant pour chaque gène potentiel.

Utilisation d'*Agrobacterium tumefaciens*

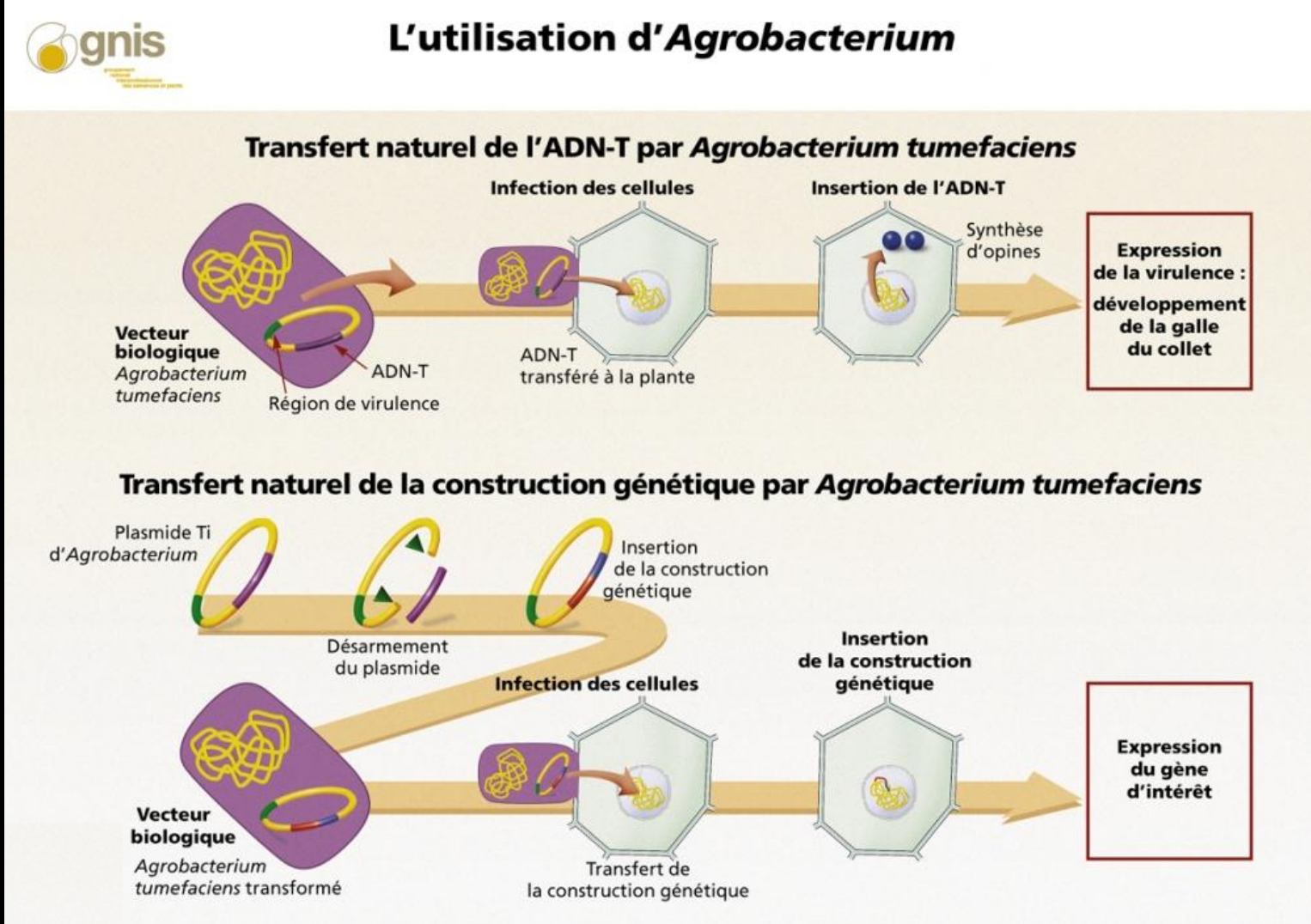
→ Maladie de la galle du collet.

Processus de transformation naturel... **OGM naturel!**

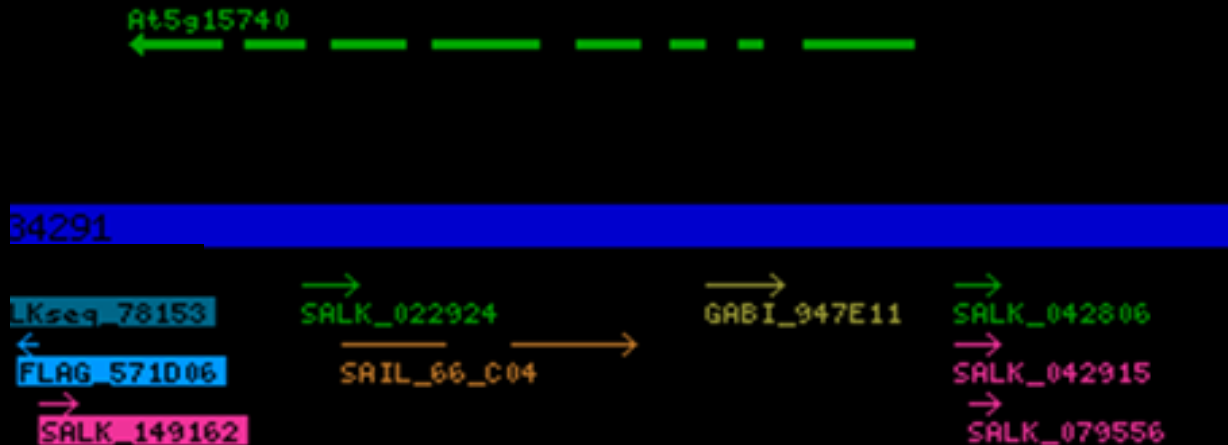


Copyright © APS Press

Possède un plasmide capable d'être transféré et inséré dans le génome végétal.



Problèmes : suivi de la construction avec un gène rapporteur (Km, BASTA)  
Génération de plantes transgéniques → confinement...



<http://signal.salk.edu/cgi-bin/tdnaexpress>

### A ce jour :

- 47.203 lignées homozygotes étiquetées, représentant 26.596 gènes
- Plusieurs collections: Salk Institute, Monsanto, INRA Versailles, U. Bielefeld
- Prix variable: d'une dizaine d'euros à quelques centaines
- Une seule borne connue : Left Border.

### Bilan :

- Temps: début en 2001
- Travail: plusieurs centaines de personnes pour générer les lignées  
deux instituts pour entretenir les lignées
- Coût: quelques 10 de millions de dollars
- Impact: système exogène → génération d'OGMs

# I. Création d'une collection de séquences flanquantes d'un transposon endogène de *Petunia*

Préambule

A. Utilisation du transposon dTph1

B. Principe de base: criblage des insertions dTph1 par PCR

C. Criblage de grandes populations: réunions d'aliquots des ADNs génomiques en pools selon une matrice 2D

D. Analyse simultanée des séquences flanquantes de transposon présentes dans plusieurs centaines de pétunias.

1. Définition des les séquences flanquantes?
2. Digestion des pools d'ADNs génomiques par des enzymes de restriction sélectionnées
3. Collage d'un adaptateur spécifique sur chaque site *MseI*
4. Amplification spécifique des séquences flanquantes couplée à l'identification des pools
5. Réunion de tous les produits PCR issus des différents pools en un super pool
6. Séquençage de masse de tous les produits PCR obtenus et annotation

E. Optimisation de la technique

F. A quoi à ça sert?

G. Comparaison Dead Star Facility (*Petunia hybrida*) avec SIGnAL T-DNA Express (*Arabidopsis thaliana*)

## Préambule

En 2007, on a décidé de générer pour *Petunia* une collection de mutants comparable à celle existant pour *Arabidopsis thaliana*

Au niveau mondial, le pétunia c'est :

- Petite communauté: quelques 100 de personnes
- Petits moyens: financiers ou matériels
- Peu de visibilité
- Mais mêmes besoins

D'où des impératifs :

- Personnel
- Coût
- Rapidité
- Fiabilité

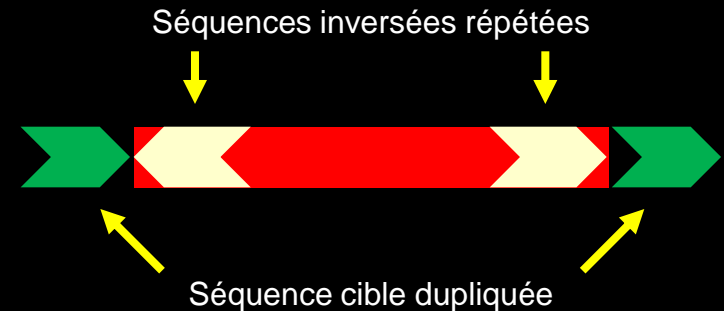
# A: Utilisation du transposon dTph1

## Transposon ou élément génétique transposable.

- Présents chez tous les organismes.
- Ont la capacité de se déplacer d'un endroit à un autre du génome.
- Souvent à l'origine de réarrangements chromosomiques (translocations, inversions, mutations...).
- Leur mobilité dépend de l'action d'une transposase interne (transposon autonome) , externe (transposon non autonome).

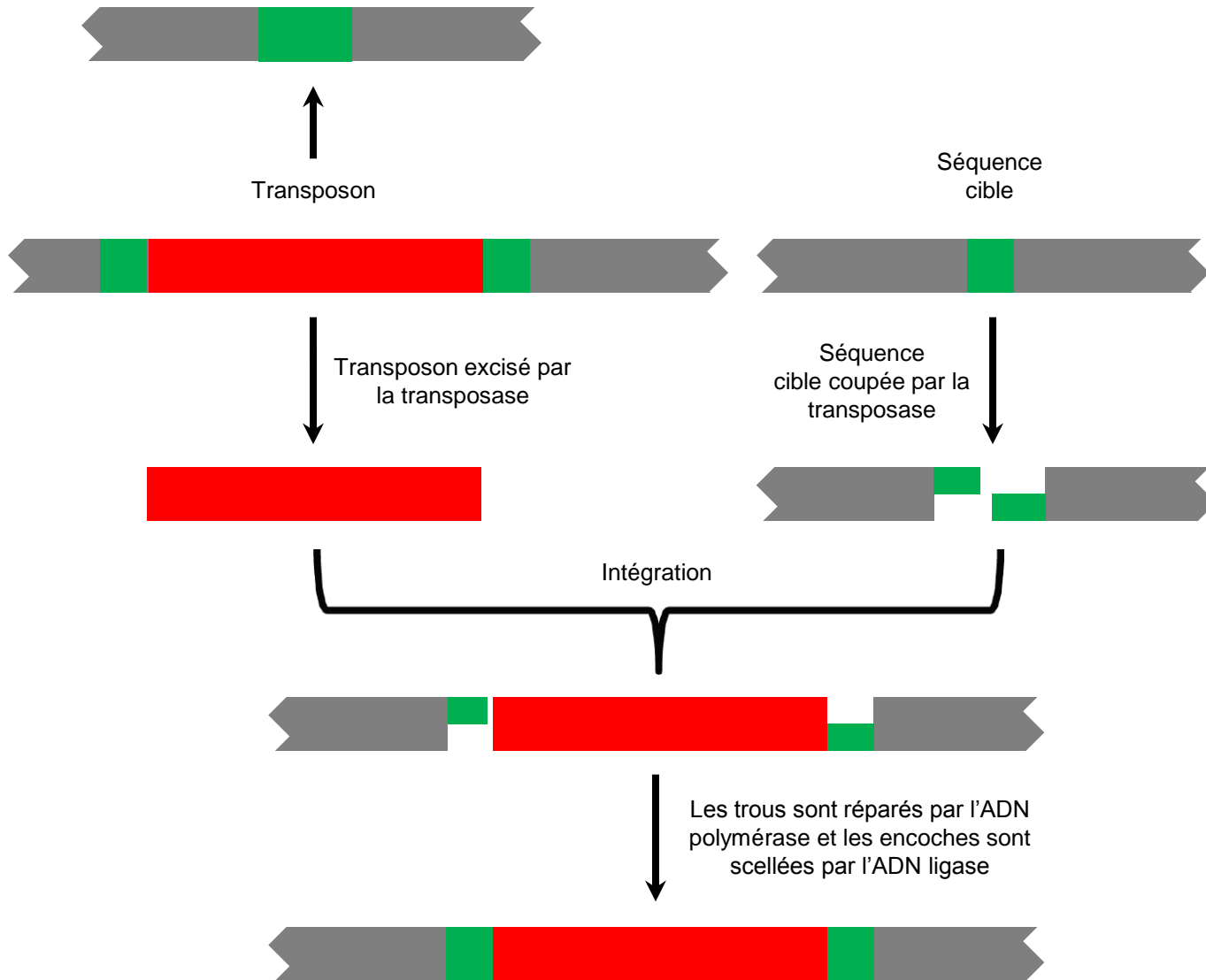
## Quelques informations concernant dTph1

- Transposon endogène
- Élément transposable non autonome
- Taille: 284 pb
- Répétitions inversées de 12 pb: GCTCCGCCCTG
- Duplication d'une séquence de 8 pb au point d'insertion
- Transposase agit en trans
- Pas encore clonée
- Sélection d'une lignée de *P. hybrida* très active pour l'activité de dTph1: W138
- dTph1: système de mutagenèse par insertion très efficace pour la génétique « forward » et la génétique « reverse ».
- (> 200 dTph1 copies/génome)





# Mécanisme d'excision de dTph1



## Visualisation de la mobilité de dTph1



Les portions colorées sont dues au transposon dTph1.

La lignée de *Petunia hybrida* W138 contient un locus *an1* muté à cause de l'insertion de dTph1 dans le promoteur. La mutation bloque la production de pigments d'anthocyanine et conduit à la formation d'une fleur blanche.

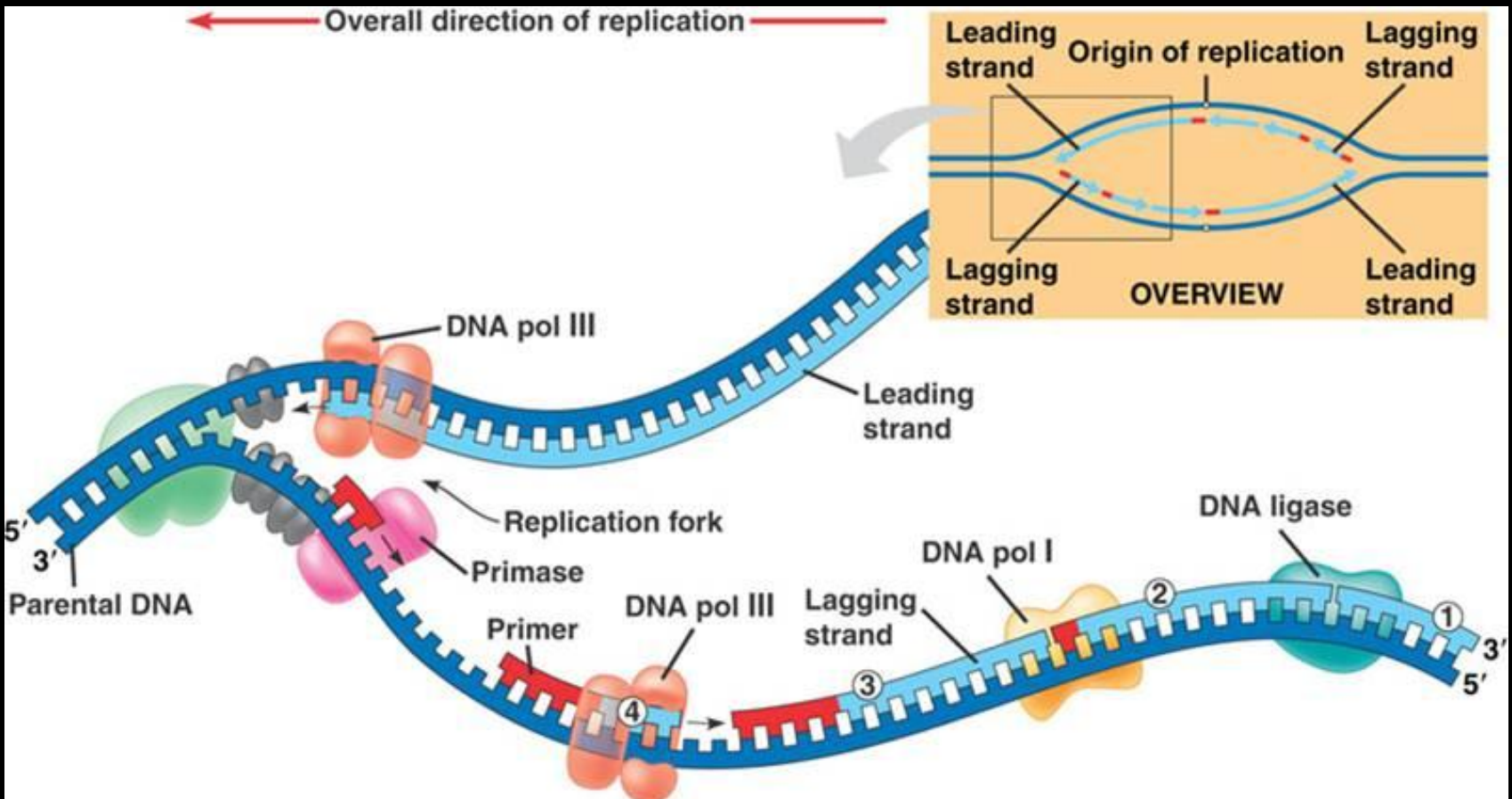
Durant le développement de la fleur, dTph1 s'excise dans certaines cellules et la production de pigment est restaurée. Le patron de coloration montre que les fleurs se développent à partir d'un petit nombre de cellules au centre du primordia floral.

La lignée W138 a été utilisée pour l'isolement de mutants car dTph1 s'excise et se réintègre dans d'autres gènes à une fréquence élevée.

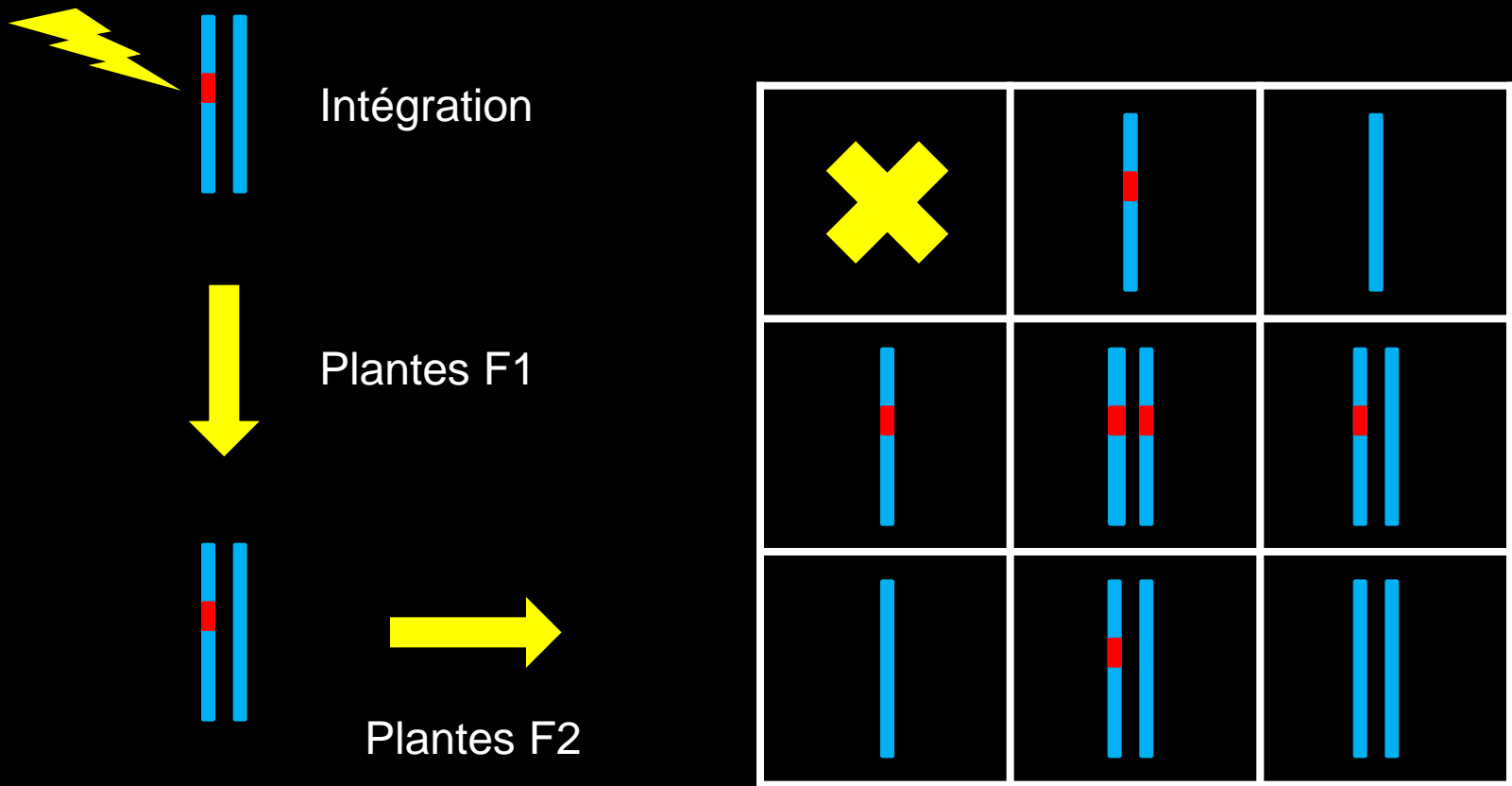
## Enzymes de modification utilisées en biologie moléculaire

**ADN polymérase.** Réplication ADN. Par rapport à une matrice simple brin. Elongation 5' vers 3'. Assure l'intégrité de la transmission du patrimoine génétique.

**ADN ligase.** Enzyme qui répare les brins brisés d'ADN en formant des liaisons phosphodiester covalentes.



dTph crée une mutation lors de son intégration.  
 Cette mutation va être stabilisée après multiplication.



En cas d'excision: duplication de la séquence cible et changement du cadre de lecture (+ 8 pb)

```

atg tgg atg atg ggt ...
Met Trp Met Met Gly
atg tgg atg att gga tga tgg gtc ...
Met Trp Met Ile Gly STOP
    
```

## **B. Principe de base: criblage des insertions dTph1 par PCR**

**PCR: polymerase chain reaction ou réaction de polymérisation en chaîne.**

Technique inventée par Kary Mullis en 1985.

Prix Nobel de chimie en 1993.

Nécessite:

Pour l'amplification proprement dite :

- Un ADN matrice
- Une ADN polymérase thermostable
- Deux amorces oligonucléotidiques (primers)
- Un mélange des 4 désoxyribonucléotides triphosphates (dNTP)
- Un thermo-cycleur

Pour l'observation et la révélation des fragments d'ADN amplifiés :

- Un appareil d'électrophorèse
- Un gel d'agarose
- Un marqueur de taille
- Un agent intercalant
- Une table UV

# L'amplification par PCR

Enzyme:

Taq polymerase

Enzyme capable de répliquer l'ADN.

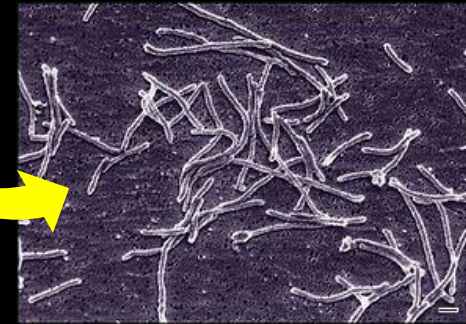
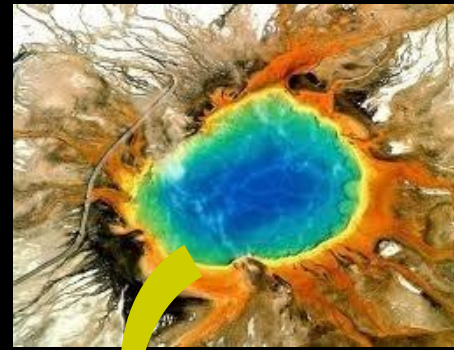
Isolée en 1969 dans le parc naturel du Yellowstone.

A partir de *Thermophilus aquaticus*, une bactérie thermophile vivant à proximité des sources d'eau chaude (de 50 à 80 °C).

Activité optimale à 72°C. Enzyme modifiée. OGM.

Fidèle

Vitesse d'élongation: 1.000 pb minute.



Amorce:

Longue de quelques dizaines de nucléotides. Obtenue par synthèse chimique. Sous forme simple brin 5' – 3'.

Hybridation spécifique avec la séquence complémentaire.

Elongée dans le sens 5' vers 3'.

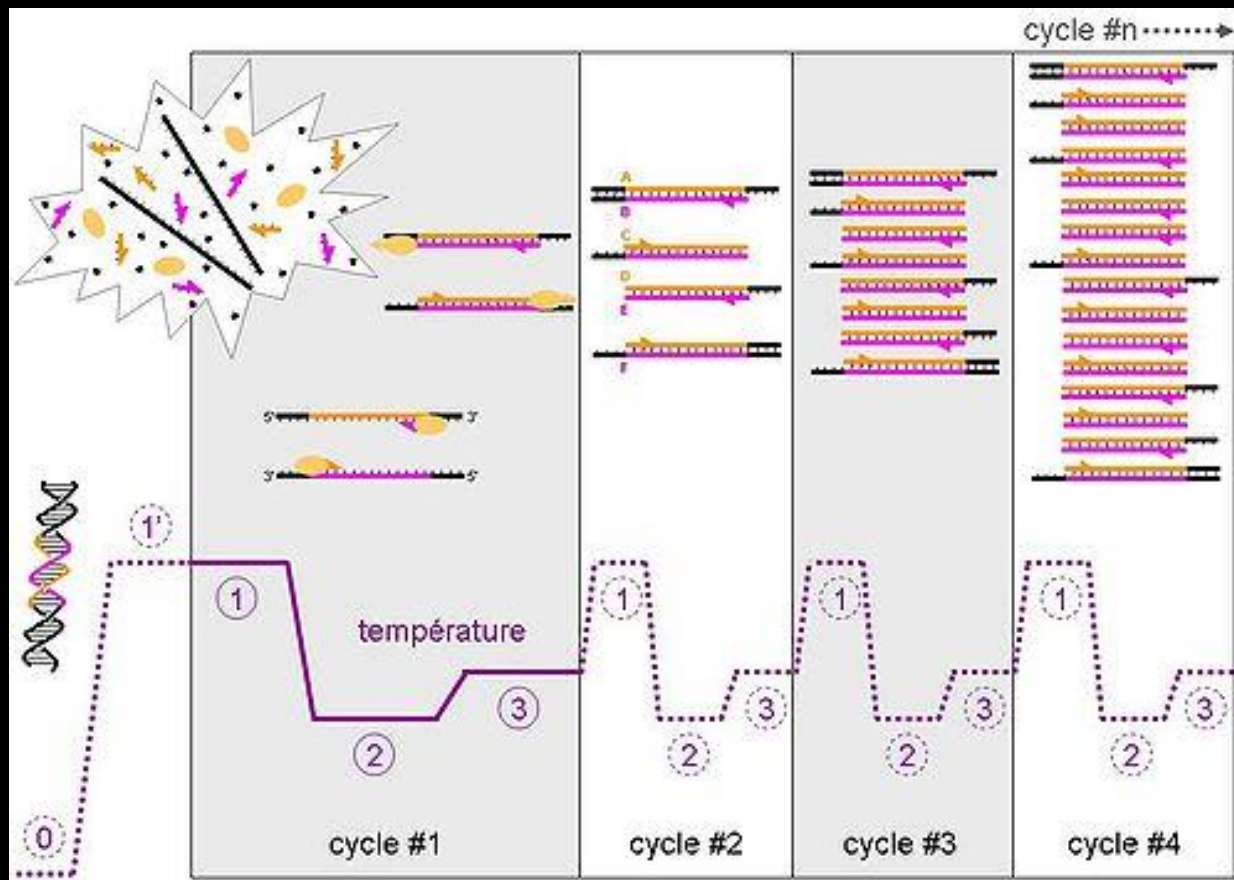
<http://www.prevx.com/ilenames/X2595202387671388793-X1/OLIGO.EXE.html>

Thermocycleur:

3.500 euros

Cycle complet en 2 heures.

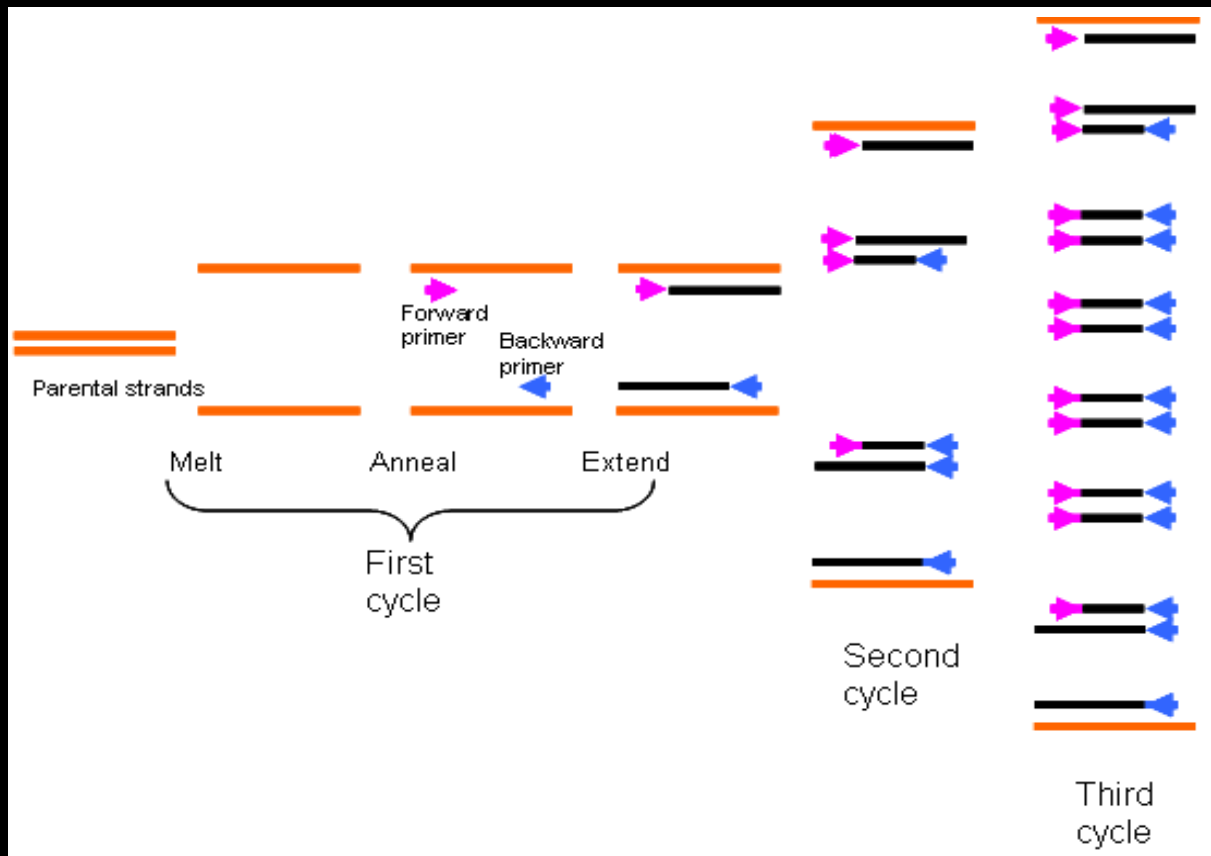




Cycle:

1. dénaturation. 30 s à 94°C
2. appariement. 30 s à 55°C
3. élongation. 30 s à 72°C

De 30 à 45 cycles



Cycles	1	2	3	4	5	10	20	30
Nbre total copies d'ADN	2	4	8	16	32	1.024	1.048.576	1.073.741.824
ADN: copies longues	2	4	6	8	10	20	40	60
ADN: copies cibles	0	0	2	8	22	1.004	1.048.536	1.073.741.764

C'est le conte du vizir et de l'échiquier.

<http://www.snv.jussieu.fr/bmedia/PCR/images/PCR-graphe.swf>



## La révélation.

Agarose.

Polymère non-ramifié à base d'agar-agar purifié

Séparation en fonction de la taille.

Gel d'agarose.

Réticulation fonction du % d'agarose dans le tampon.

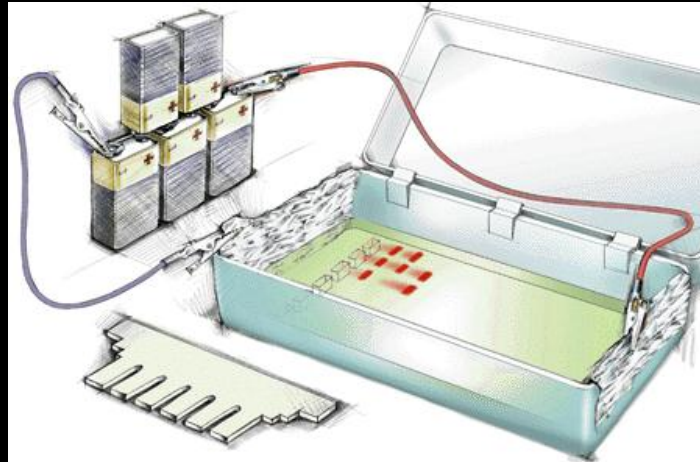
0,8 à 3% (g/ml). Selon la taille des fragments à séparer.

Electrophorèse en gel d'agarose

Courant électrique passe dans un tampon.

ADN chargé négativement.

Migre vers l'anode (+).

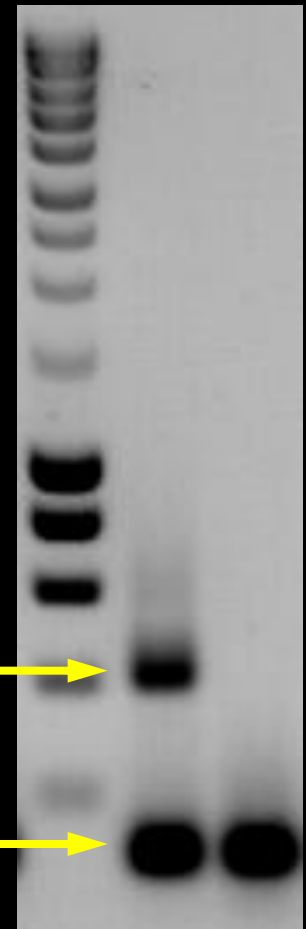


Révélation par un intercalant: bromure d'éthidium.

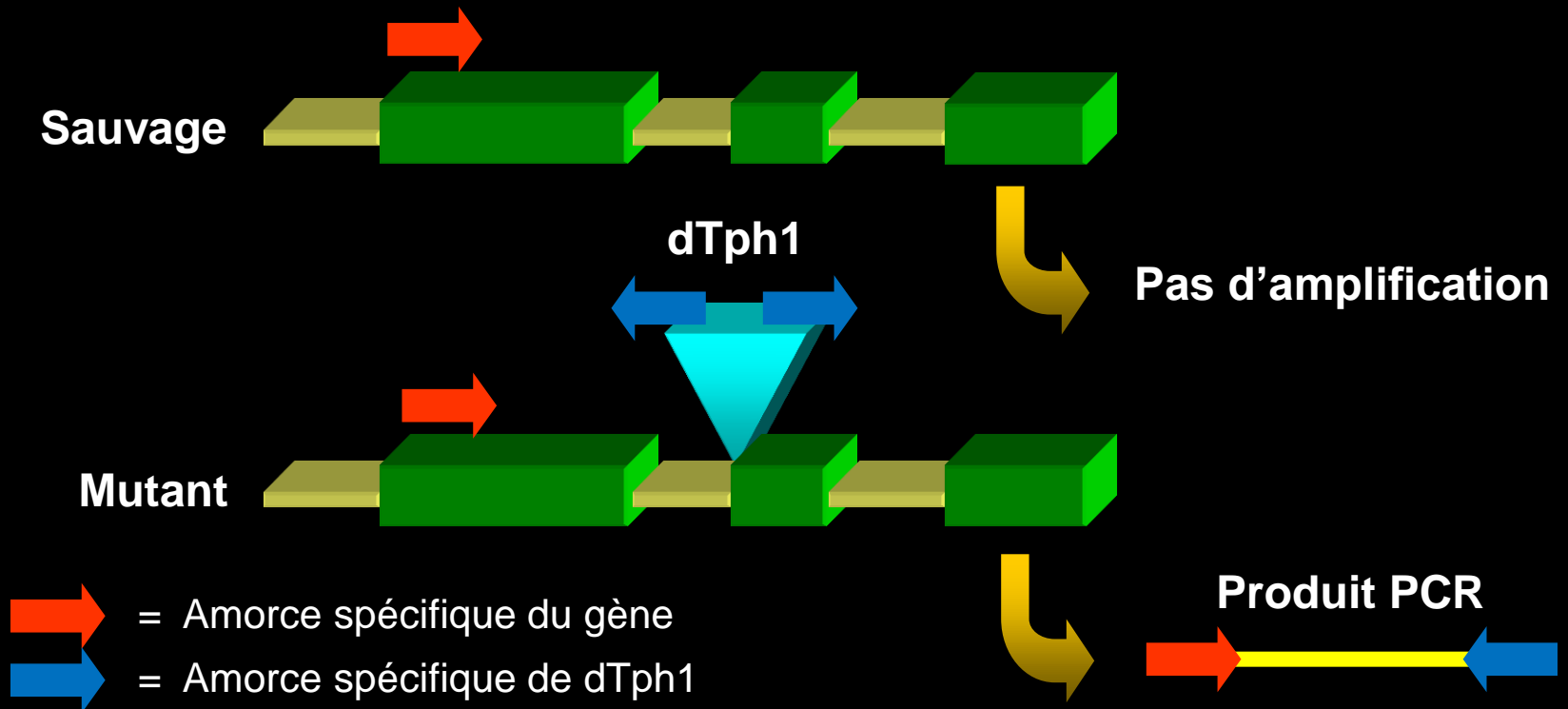
Dans la double hélice d'ADN.

Révélation sous UV. Et photographie.

M E+ E-



# Le criblage (recherche) des insertions dTph1 par PCR



## C. Criblage de grandes populations: réunions d'aliquots des ADNs génomiques en pools selon une matrice en 2 dimensions

	X01	X02	X03	X04
Y01	1	2	3	4
Y02	5	6	7	8
Y03	9	10	11	12
Y04	13	14	15	16

16 plantes (4 x 4)                      8 pools  
100 plantes (10 x 10)                  20 pools  
900 plantes (30 x 30)                  60 pools

→ Chaque plante apparaît dans 2 pools (un pool en X et un pool en Y)

→ Plante 6:              pool X02  
                              pool Y02

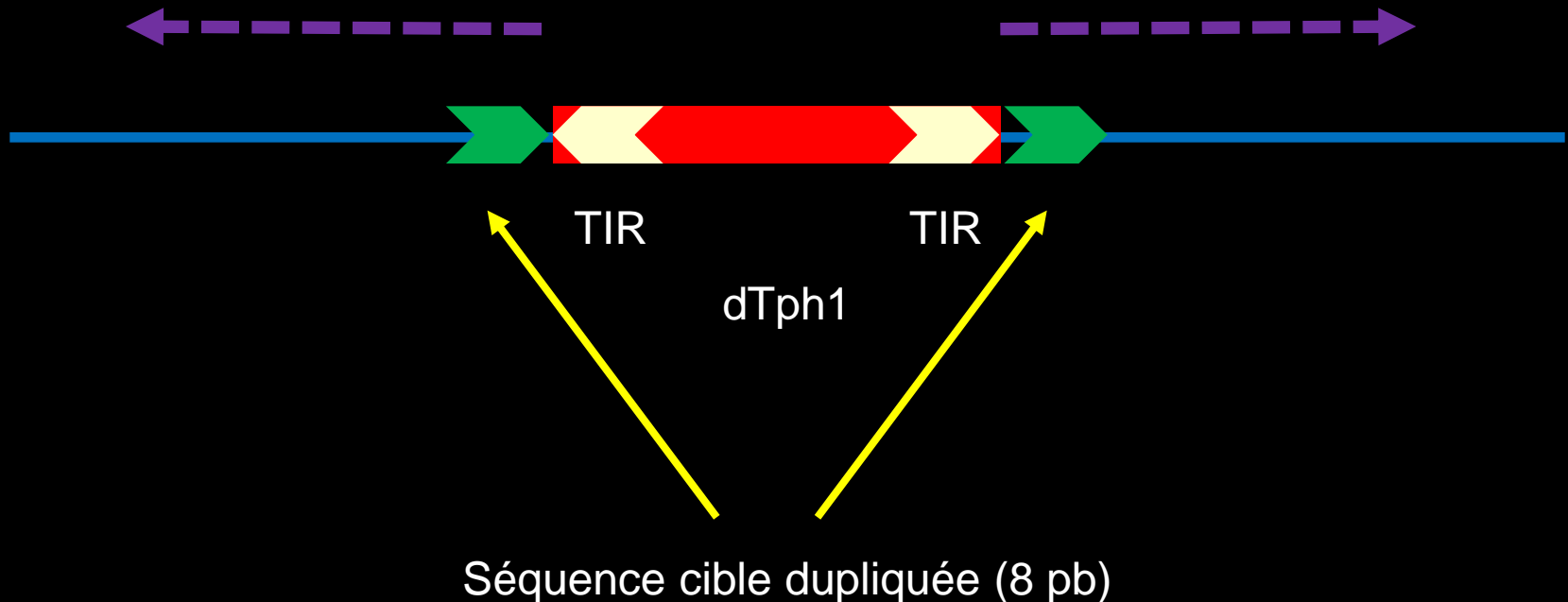
→ Ceci permet de réduire la taille de l'expérience: par exemple 60 pools/900 plantes

→ Tout signal issu d'une plante devrait donc apparaître dans 2 coordonats (X - Y)

→ L'identité de la plante parentale est donnée par l'intersection des 2 coordonats

## D. Analyse simultanée des séquences flanquantes de transposon présentes dans plusieurs centaines de pétunias.

### 1. Définition des séquences flanquantes?



Séquence flanquante:

2 pour chaque transposon:

une séquence flanquante droite

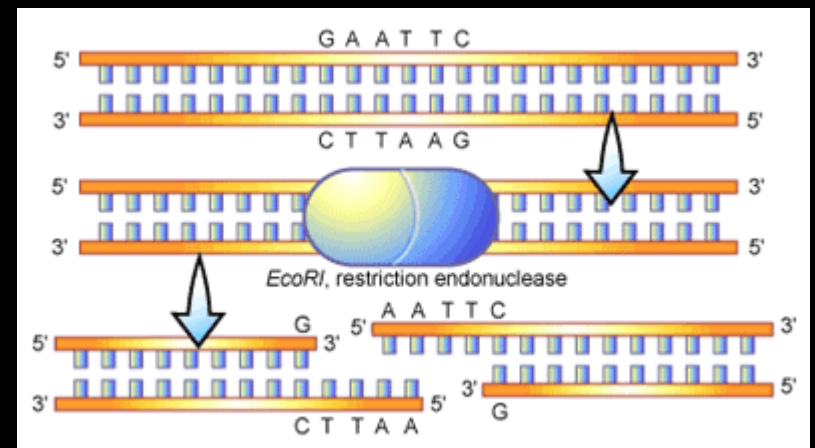
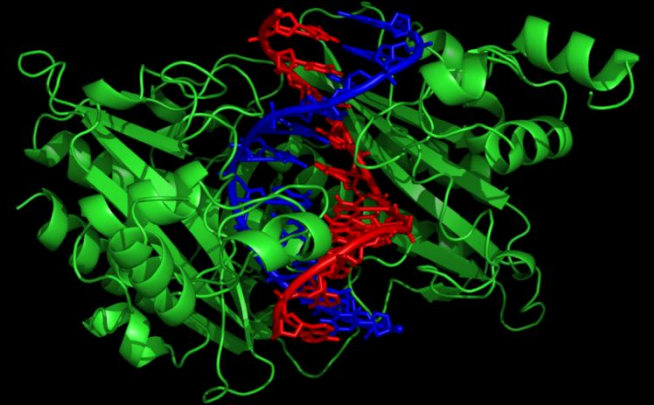
une séquence flanquante gauche

Commencent avec la séquence cible dupliquée

## 2. Digestion des pools d'ADNs génomiques par des enzymes de restriction sélectionnées

### Les enzymes de restriction:

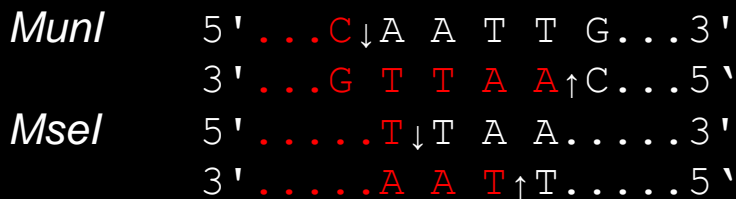
- Découvertes en 1962
- Endonucléases, d'origine bactérienne.
- Proviennent des systèmes de restriction-modification bactériens, systèmes de défense vis-à-vis des bactériophages. Les endonucléases digèrent l'ADN parasite et les méthylases protègent l'ADN de la bactérie.
- Permettent la différence entre le soi et le non soi.
- Coupent l'ADN au niveau de sites spécifiques, les sites de restriction, définis par une séquence nucléotidique particulière, de nature palindromique.
- Palindrome: phrase pouvant se lire dans les deux sens:  
Léon a erré à Noël  
La mariée ira mal.  
Élu par cette crapule  
La mère puce récupère mal
- Enzyme de restriction: palindrome sur les deux brins anti-parallèles.



## Les enzymes de restriction:

- Agissent sous forme de dimères.
- Plusieurs centaines connues et commercialisées.
- Toutes clonées, et produites par génie génétique, donc OGM.
- Coupent chaque fois qu'elles reconnaissent la bonne séquence.

Ont été utilisées:



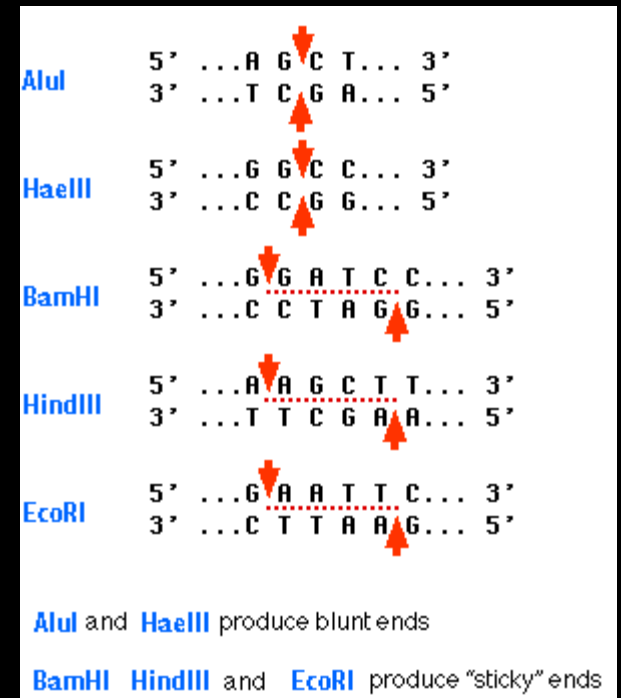
<https://www.neb.com/> (Biolabs)

<http://www.restrictionmapper.org/> (Programme de visualisation)

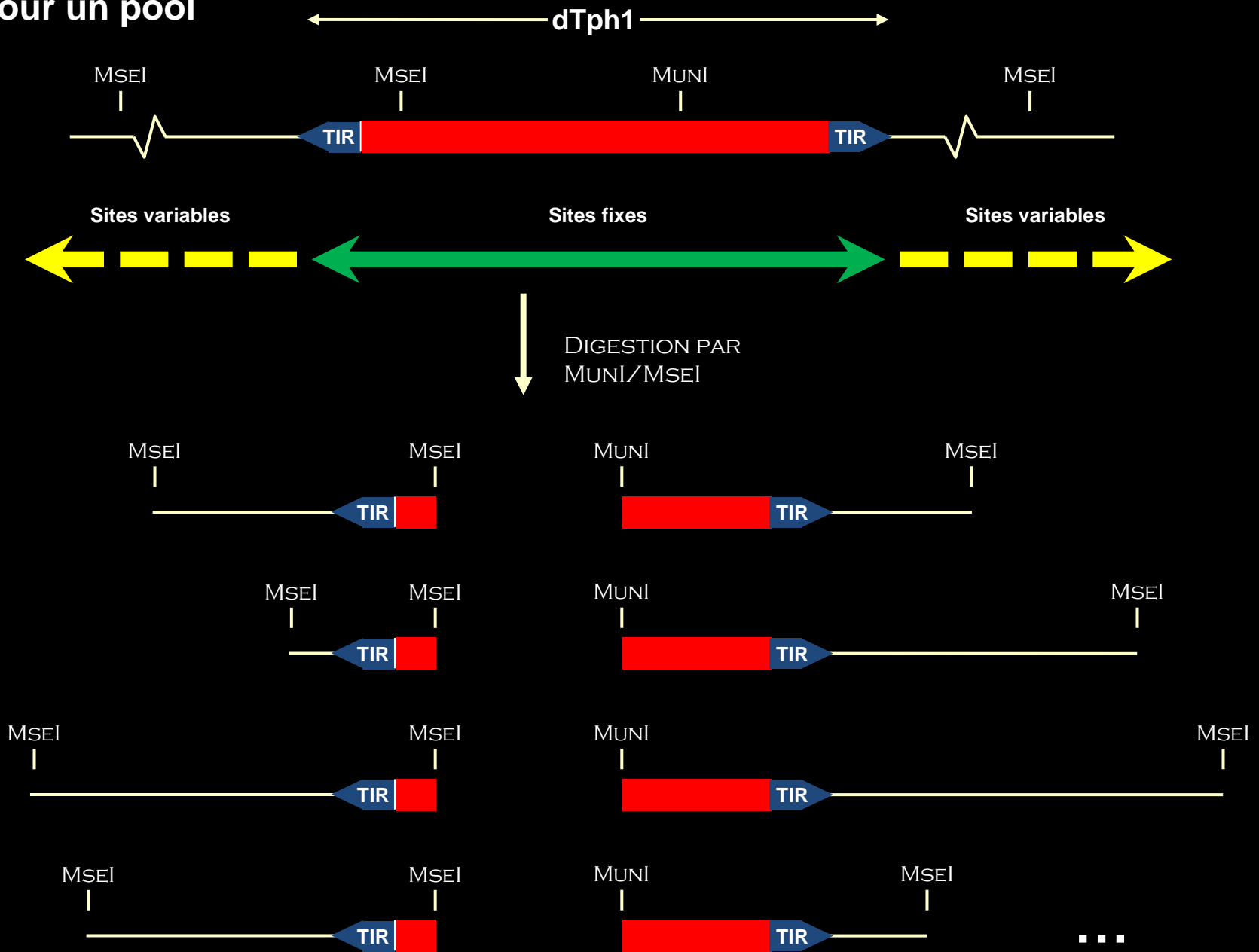
Fréquence de coupure statistique sur le génome

Site de 4 bases:      $(1/4)^4$  soit toutes les 256 pb

Site de 6 bases:      $(1/4)^6$  soit toutes les 4096 pb



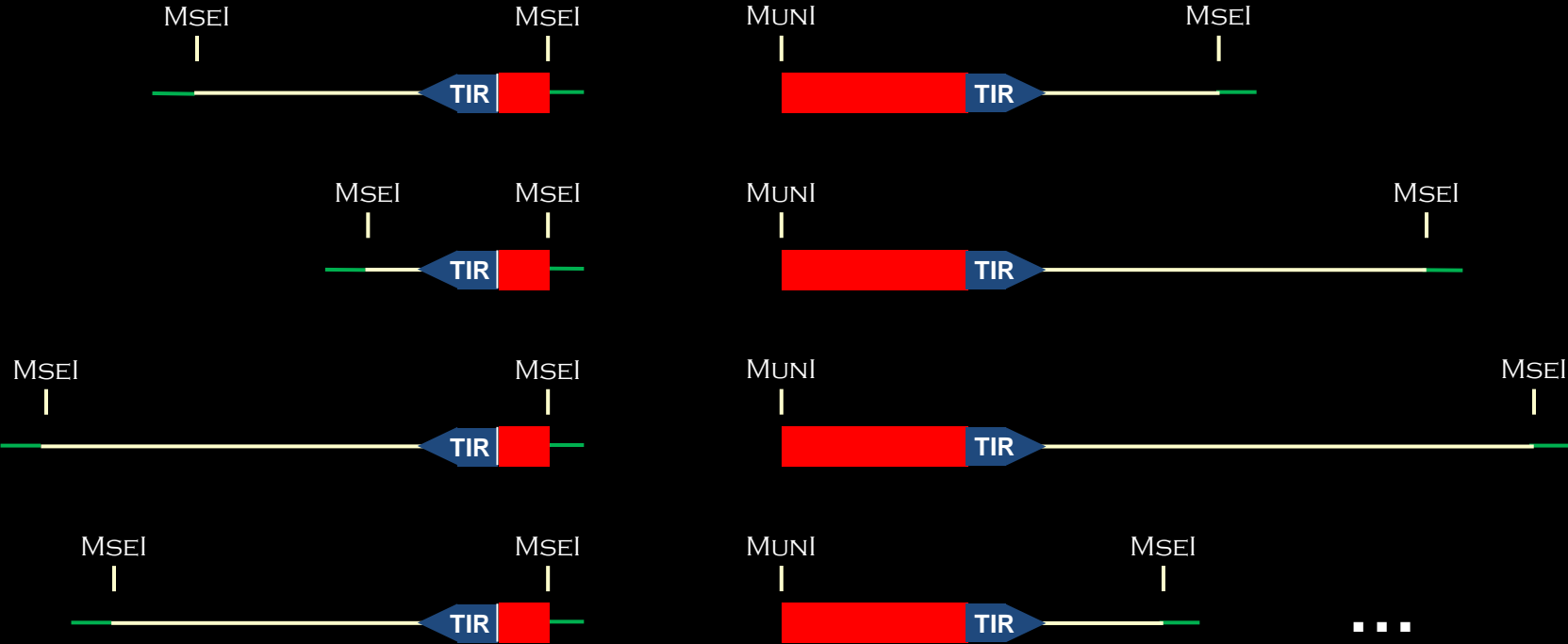
# Pour un pool





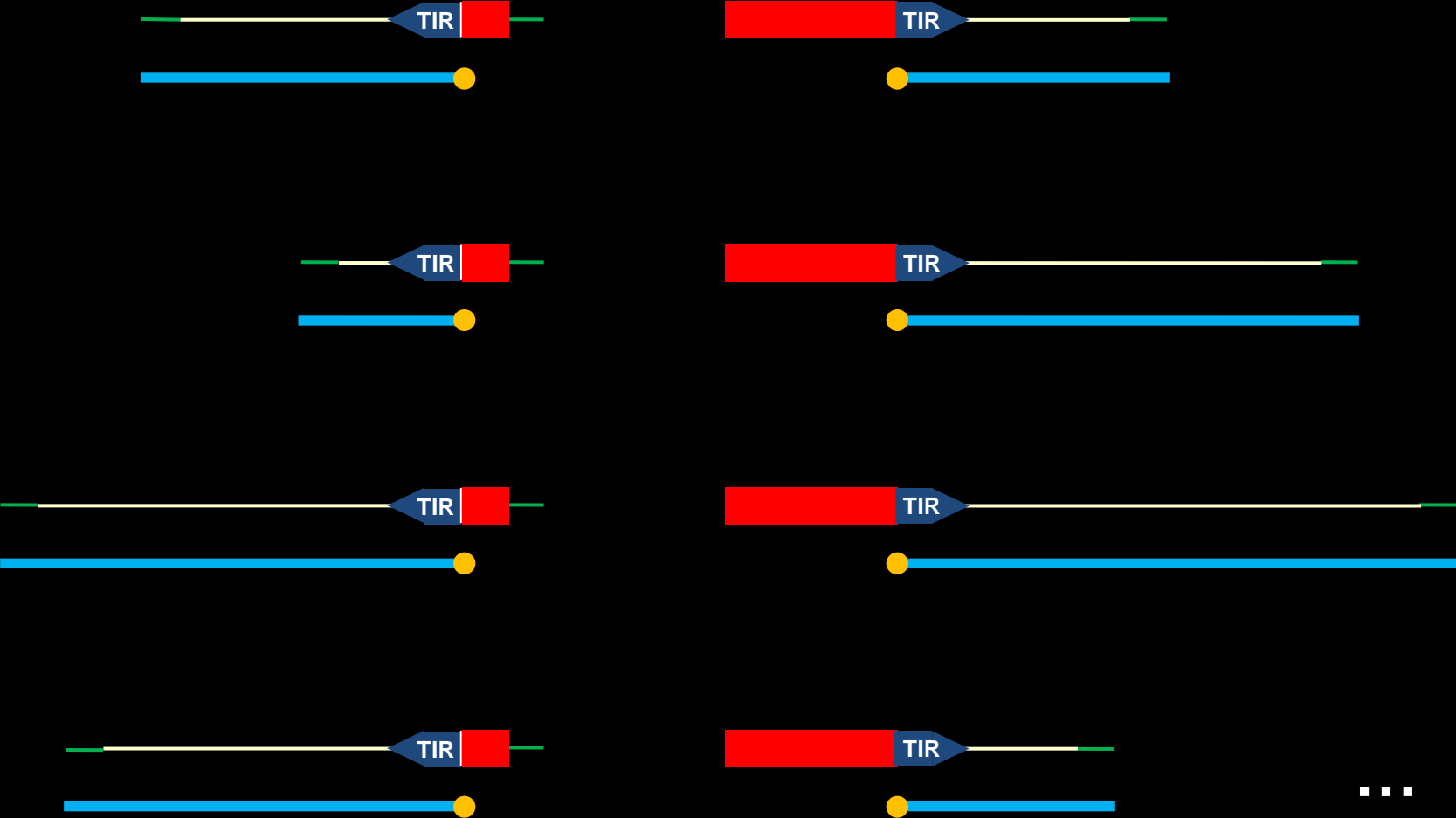


Toujours pour un pool. Mais il se passe la même chose dans chaque pool...





Pour un pool donné. Chaque fragment amplifié a une identité précisée par la séquence des 4 nucléotides de l'étiquette

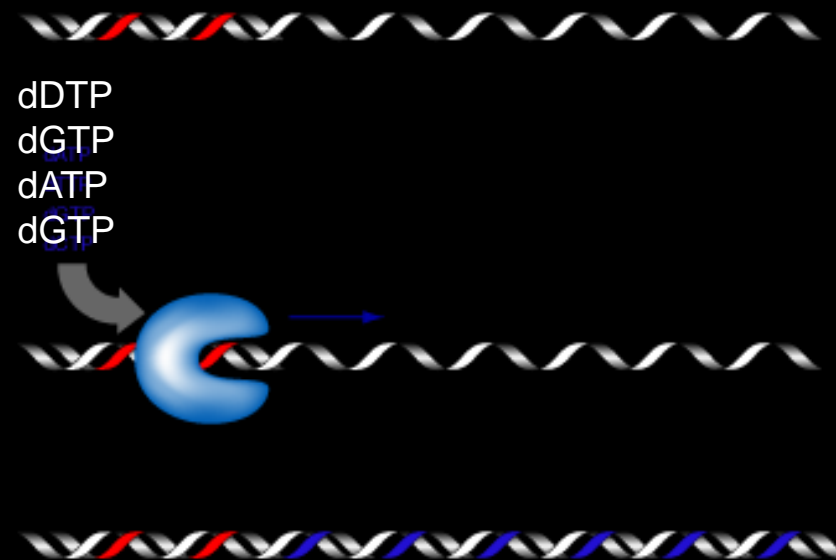


...

## 5. Réunion de tous les produits PCR issus des différents pools en un super pool



## 6. Séquençage de masse de tous les produits PCR obtenus et annotation

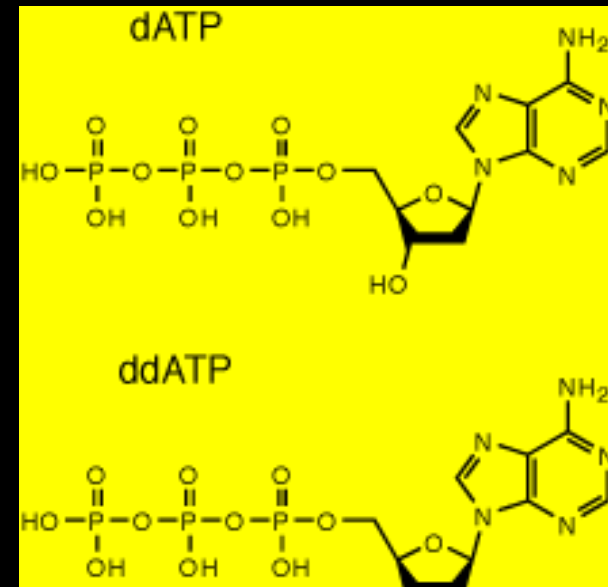
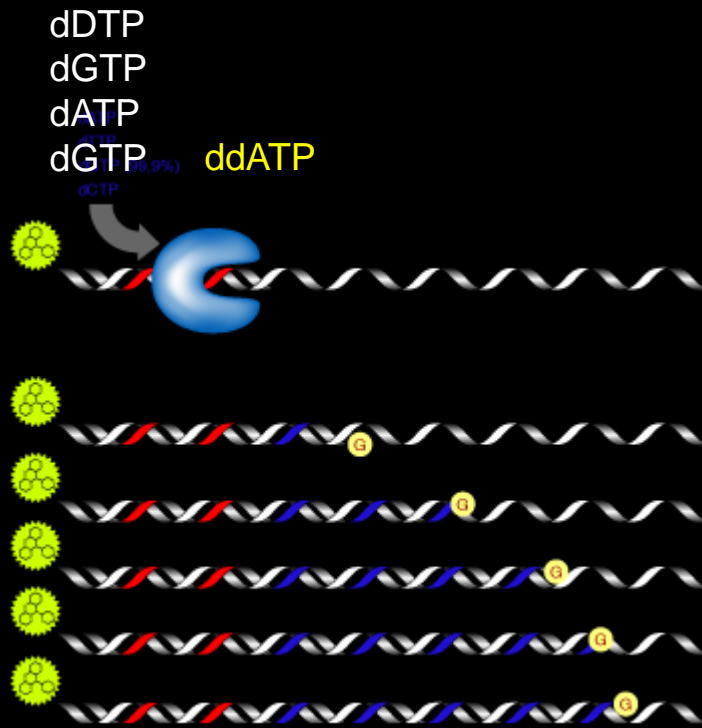


**Le séquençage:** détermination de l'enchaînement des 4 bases.  
Repose sur le mécanisme de polymérisation de l'ADN. La matrice (en blanc) est copiée par la polymérase qui allonge le brin complémentaire à partir de l'amorce

## Principe du séquençage par la méthode de Sanger.

On utilise un mélange de **désoxyribonucléotides** (dNTP) et des **didésoxyribonucléotides** (ddNTP) correspondants.

Les **didésoxyribonucléotides** (ici, le ddATP, en jaune) sont incorporés mais bloquent statistiquement l'allongement de la chaîne là où les dNTP sont normalement incorporés (ici, le dATP). Un traceur radioactif ou fluorescent (en vert clair) est attaché à l'extrémité de l'amorce de polymérisation et permet de détecter les fragments d'ADN synthétisés.



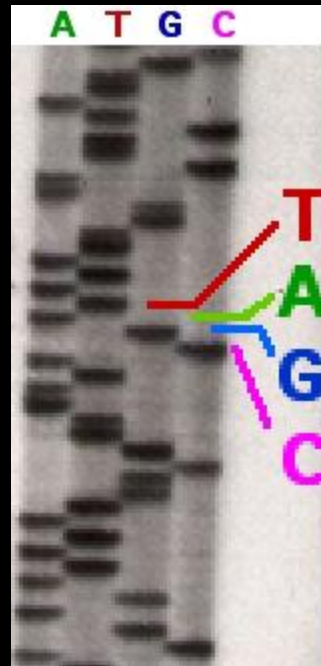
## Le séquençage radioactif.

Quatre réactions sont effectuées en parallèle, avec dans chacune, un didésoxyribonucléotide différent.

Ensuite les quatre réactions sont déposées côte à côte et mises à migrer simultanément sur un gel de poly-acrylamide.

L'ordre de chaque bande indique la position d'un nucléotide.

Les plus petits fragments migrant le plus loin, la séquence se lit de bas en haut.



Le séquençage avec des traceurs fluorescents.

Chacun des quatre didésoxyribonucléotides est associé à un traceur fluorescent différent.

T: rouge

A: vert

G: jaune

C: bleu

Les quatre réactions sont donc effectuées simultanément dans le même tube et révélées ensuite par passage devant un détecteur laser.

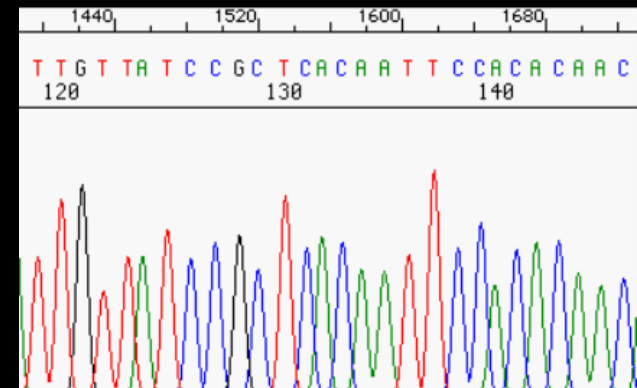
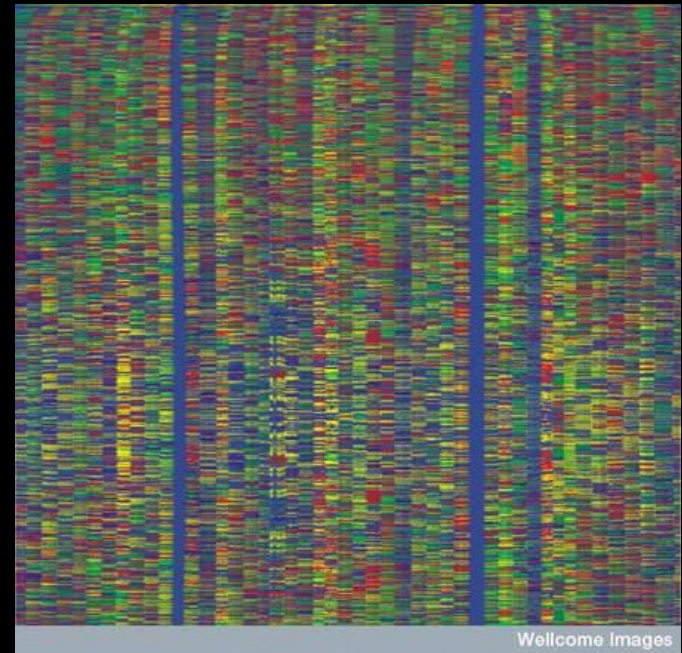
Avant (10 ans), on utilisait une séparation par migration sur gel de poly-acrylamide. Maintenant, les produits sont séparés sur des colonnes micro-capillaires.

Mais, *in fine*, on obtient le même type de profil chromatographique couplé à un fichier texte.

Depuis peu les séquenceurs de nouvelle génération: Illumina 454 ou HighSeq.

Séquences courtes: 100 bp

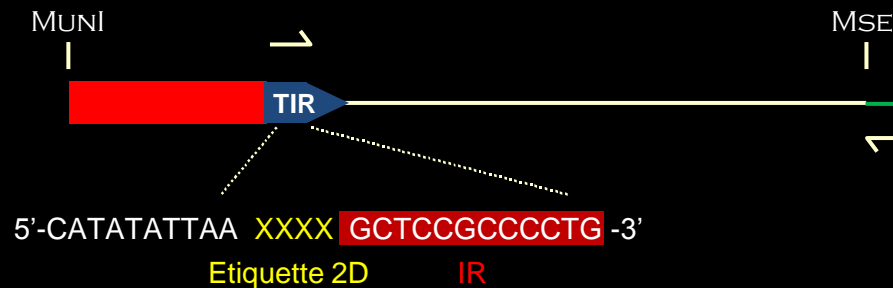
En grand nombre: milliers voire millions





## Séquençage de masse Illumina 454 sur le super pool.

On obtient la séquence des fragments amplifiés entre l'amorce TIR et l'amorce *MseI* (si elles ne sont pas trop éloignées).



300.000 séquences of 100 pb

Séquences arrivent sous forme de fichiers Text.

Avec un numéro d'identification spécifique.

Toutes les données sont traitées par ordinateur

Analysées et classées par un programme gratuit BioEdit

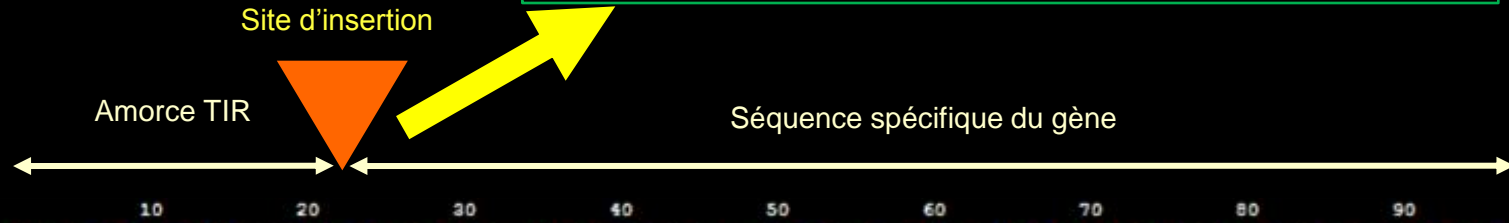
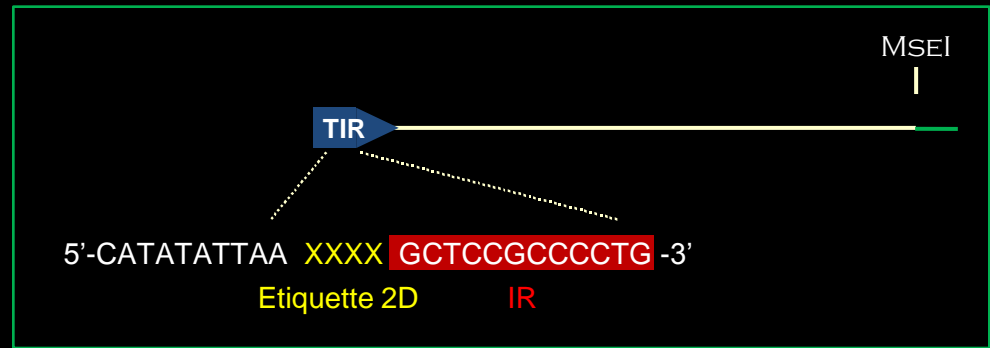
<http://www.mbio.ncsu.edu/bioedit/bioedit.html>

Disons "par ordre alphabétique"

Puis annotées manuellement.

# Annotation 2D

Empilement de toutes les séquences  
 Etiquette 01 à 30: dimension X  
 Etiquettes 31 à 60: dimension Y

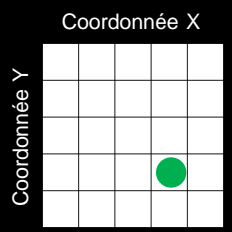
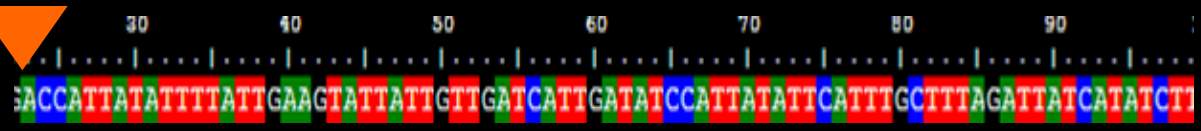


	10	20	30	40	50	60	70	80	90
Tag 01	00103432	TAAAGACGTAGCTCCGCCCTG	ACCATTATATTTTATTG	AAGTATTATTGTTGATCATTG	ATATCCATTATATTCATTTG	CTTTAGATTATCATAT			
Tag 01	00194560	TAAAGACGTAGCTCCGCCCTG	ACCATTATATTTTATTG	AAGTATTATTGTTGATCATTG	ATATCCATTATATTCATTTG	CTTTAGATTATCATAT			
Tag 01	00268486	TAAAGACGTAGCTCCGCCCTG	ACCATTATATTTTATTG	AAGTATTATTGTTGATCATTG	ATATCCATTATATTCATTTG	CTTTAGATTATCATAT			
Tag 41	00225597	TAAATCGGTAGCTCCGCCCTG	ACCATTATATTTTATTG	AAGTATTATTGTTGATCATTG	ATATCCATTATATTCATTTG	CTTTAGATTATCATAT			
Tag 41	00275063	TAAATCGGTAGCTCCGCCCTG	ACCATTATATTTTATTG	AAGTATTATTGTTGATCATTG	ATATCCATTATATTCATTTG	CTTTAGATTATCATAT			

Etiquette 2D

Séquences flanquantes  
 Cordonnées de la plante origine  
 (où était dTph1 et dont on a  
 récupéré les graines...)

Ici la plante est appelée: Lib06-M\_32011-#(X01-3,Y41-2)



→ → CA MARCHE !!!

## E. Optimisation de la technique

\* Génération d'une grande population de plantes de *Petunia* (7.000 plantes)

Lib2006 : 1.000 individus



Lib2008 : 3.540 individus



Lib2009 : 2.390 individus



- \* Utilisation des plantes en différentes collections ou librairies (2006, 2008, 2009)
- \* En parallèle à MseI, utilisation d'autres enzymes (Bfal...)
- \* Utilisation d'un double étiquetage, par addition séquentielle de code-barres.
  - 60 amorces différentes pour l'identification des pools
  - 15 amorces Illumina différentes (Librairies, enzymes)
- 60 x 15 = 900 # combinaisons de marquages différentes
- \* Utilisation d'un nouveau système de séquençage (Illumina HIGH Seq 2000).

→ 180.000.000 séquences.

## F. A quoi à ça sert?

Deux interrogations principales concernent cette collection de séquences flanquantes du transposon dTph1...

### Résultats

180.000.000 séquences produites

→ 230.000 points d'insertion différents déterminés (mutants?)

→ 1 insertion / 6,5 kb (en théorie)

→ Génome semi-saturé

### Qualité des résultats

1. Peut-on remonter d'une séquence à une plante ou à un lot de graines? (Ne pas oublier que l'on travaille sur l'ADN des parents)

Peut-on trouver des mutants?

Est-ce que les mutations se retrouvent entre les générations

2. La distribution des mutations est-elle homogène?



## G. Comparaison Dead Star Facility (*Petunia hybrida*) avec SIGnaL T-DNA express (*Arabidopsis thaliana*)

	<i>Arabidopsis</i>	<i>Petunia</i>
Durée du projet	13 ans	7 ans
Travail de génération	50-100 personnes	3-5 personnes
Travail de maintenance	10 personnes	3 personnes
Coût du projet	50.000.000 \$	50.000 \$
Impact	OGM	Non OGMs

**Gain d'efficacité : 100.000 !**

## II. Etude des gènes de la famille AP1 de *Petunia* et utilisation de la banque de séquences flanquantes.

### A. La famille APETALA1 (AP1) chez *Arabidopsis thaliana*

1. Fonction dans l'identité des organes floraux.
2. Fonction dans l'identité du méristème floral.

### B. La famille APETALA1 (AP1) chez *Petunia hybrida*

1. Le génome de *Petunia*
2. Recherche des séquences des membres de la famille AP1
  - a. Récupération des séquences des 4 gènes paralogues d'*Arabidopsis*.
  - b. Criblage des banques génomiques de *Petunia*.
  - c. La famille AP1 chez *Petunia*.
3. Isolement de mutants pour chacun des gènes de la famille AP1
  - a. Recherche dans la collection de séquences flanquantes de transposon et exploitation des résultats bruts.
  - b. Elimination des séquences de faible complexité.
  - c. Elimination des séquences en dessous du niveau du bruit de fond.
  - d. Récupération des séquences nucléotidiques correspondantes
  - e. Elimination des séquences redondantes
  - f. Ultime recherche blastN.
4. Les mutants obtenus pour les gènes de la famille AP1.

### C. Conclusions



# A: La famille APETALA1 (AP1) chez *Arabidopsis thaliana*

## 1. Fonction dans l'identité des organes floraux.

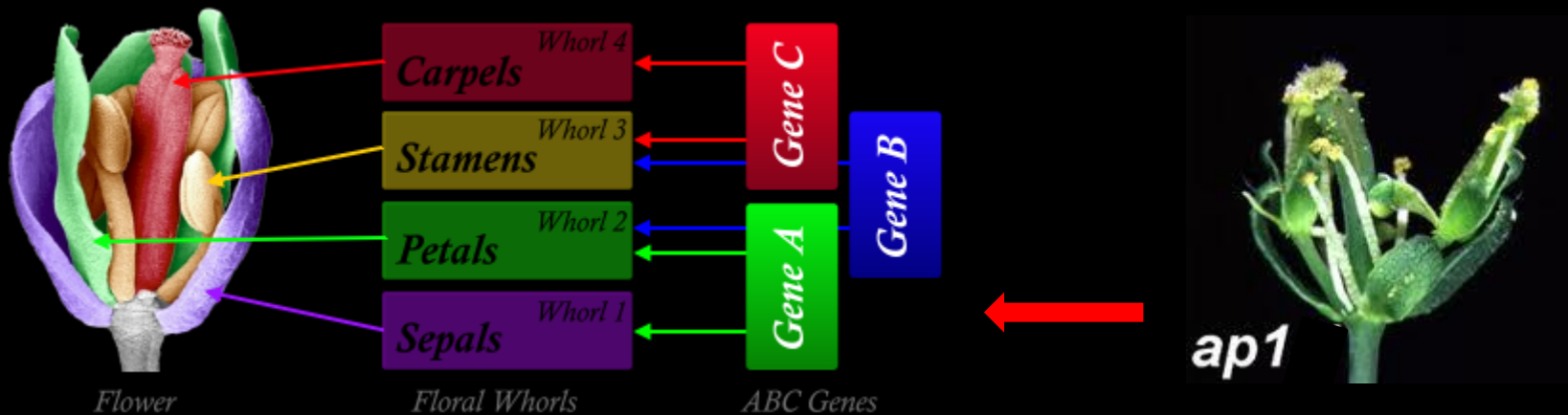
AP1 est requis pour la formation des sépales et des pétales.

Il code pour la fonction A dans le modèle ABC du développement floral.

Le phénotype mutant de type A n'est connu que chez *Arabidopsis*.

→ On peut donc avoir des doutes sur l'universalité de sa fonction.

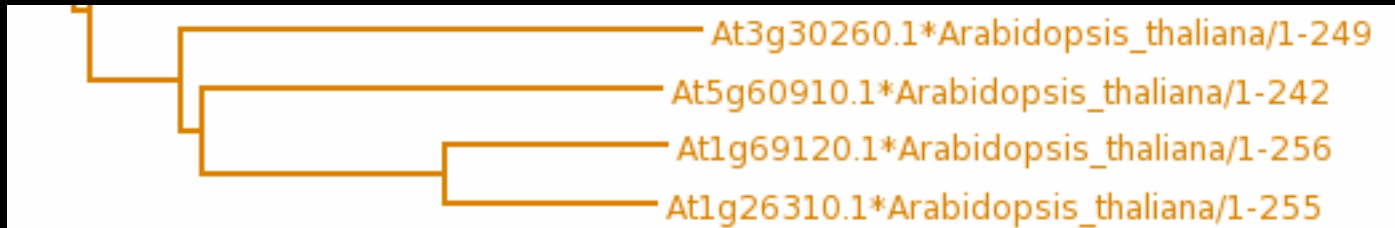
→ D'où l'intérêt d'étudier la fonction A chez *Petunia*.



Modèle ABC

→ Atelier avec Christophe Gaudin

## 2. Fonction dans l'identité du méristème floral.



Trois autres protéines appartiennent à la famille AP1  
(paralogues: issus d'une duplication):

- CAULIFLOWER (CAL)
- FRUITFULL (FUL)
- AGL79

*FUL* est nécessaire au développement du carpelle et du fruit

*CAL* est un gène d'identité du méristème mais il ne présente pas de phénotype par lui-même.

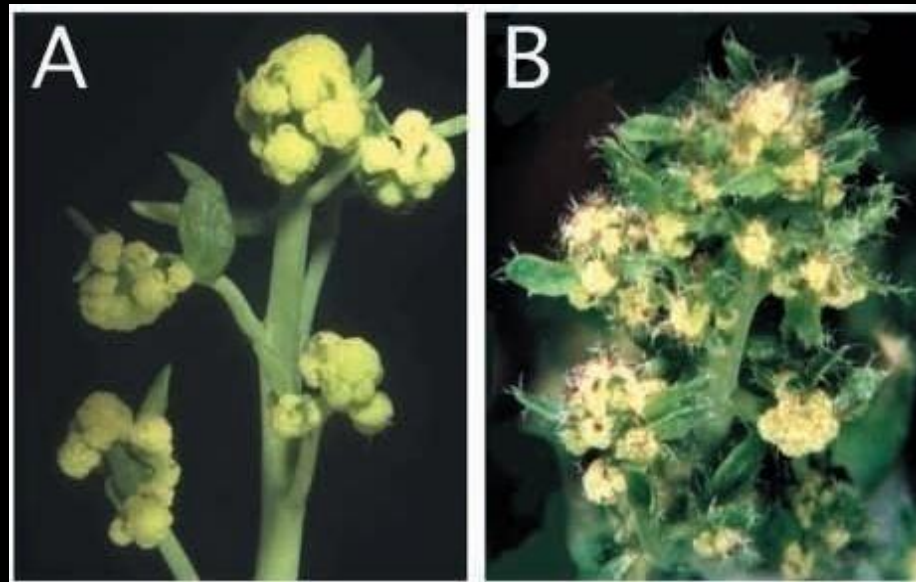
*AGL79* n'a pas encore été caractérisé.



Chez *Arabidopsis*, les combinaisons de mutants ont un effet sur l'identité du méristème floral.

*ap1 cal*: on obtient une structure en chou-fleur, avec une transformation totale du méristème floral en méristème d'inflorescence.

*ful*: accroît le phénotype du double mutant *ap1 cal*. Apparition de structure de type feuille caulinaire.



*ap1 cal*

*ap1 cal ful*

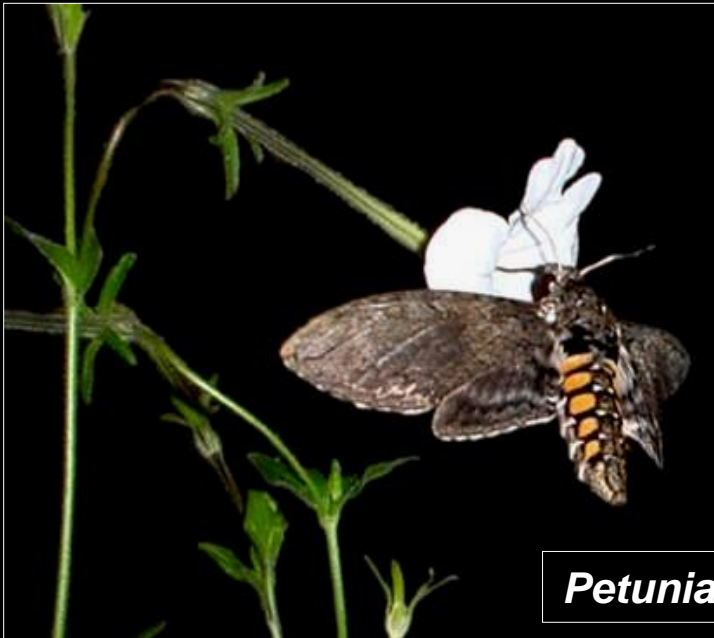
→ Retrouve-t-on les mêmes fonctions chez *Petunia* pour les orthologues des gènes de la famille AP1.

(Orthologues: issus d'une spéciation).

## B. La famille APETALA1 (AP1) chez *Petunia hybrida*

### 1. Le génome de *Petunia*

2 génomes de *Petunia* sont disponibles:



*Petunia axillaris*



*Petunia integrifolia*



华大基因  
BGI

*P. axillaris* : 2,891,906 séquences;  
1,340,831,961 bases  
*P. integrifolia* : 1,509,159 séquences;  
1,320,162,228 bases



Toutes les séquences obtenues ont été chargées sous une interface Bioedit  
<http://www.mbio.ncsu.edu/bioedit/bioedit.html>

Programme:

Gratuit

Facile d'utilisation

Très performant

## 2. Recherche des séquences des membres de la famille AP1

### a. Récupération des séquences des 4 gènes d'*Arabidopsis* (paralogues).

En allant sur le site d'*Arabidopsis*

<http://www.arabidopsis.org/>

En allant sur le site de US National Library of Medicine

<http://www.ncbi.nlm.nih.gov/pubmed>

Floral homeotic protein APETALA 1 [*Arabidopsis thaliana*]

NCBI Reference Sequence: NP\_177074.1

>gi|15222220|ref|NP\_177074.1| Floral homeotic protein APETALA 1 [*Arabidopsis thaliana*]

MGRGRVQLKRIENKINRQVTFKRRAGLLKKAHEISVLCDAEVALVVFHKGKLFESTDSCMEKILERYERYSYAERQLIAPESDVNTNWSMEY  
NRLKAKIELLERNQRHYLGEDLQAMSPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILRAQQEQWDQQN  
QGHNMPPLPPQQHQIQHPYMLSHQSPFLNMGGLYQEDDPMAMRRNDLELTLEPVYNCNLGCFAA

## b. Criblage des banques génomiques de *Petunia*

A partir de **BioEdit**

<http://www.mbio.ncsu.edu/bioedit/bioedit.html>

On recherche des homologues entre la séquence d'une protéine d'*Arabidopsis* et la banque de séquences de *Petunia* (orthologues).

Cette recherche se fait par un programme qui s'appelle **Blast**: Basic Local Alignment Search Tool

BLAST trouve des régions de similarité entre des séquences biologiques.

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

Il existe 5 possibilités de Blast

- blastn: Search a **nucleotide** database using a **nucleotide** query
- blastp: Search a **protein** database using a **protein** query
- blastx: Search a **protein** database using a **translated nucleotide** query
- Tblastn: Search a **translated nucleotide** database using a **protein** query
- Tblastx: Search a **translated nucleotide** database using a **translated nucleotide** query

→ On utilise **Tblastn**

Les séquences nucléotidiques de *Petunia* sont traduites dans les 6 cadres de lecture possibles (3 sur un brin d'ADN, 3 sur le brin complémentaire)

[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastn&BLAST\\_PROGRAMS=tblastn&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastn&BLAST_PROGRAMS=tblastn&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)

## Type de résultats obtenus:

Query= AP1 (256 letters)

Database: Peaxi.genome.v0.1.scaffolds.nrcontigs.fasta  
2,891,906 sequences; 1,340,831,961 total letters

Sequences producing significant alignments:	Score (bits)	E Value
Peaxi.v0.1.Scf25941 16.6	112	2e-024
Peaxi.v0.1.Scf3546 17.3	111	6e-024
Peaxi.v0.1.Scf8908 18.2	110	1e-023
Peaxi.v0.1.Ctg19820694 6.0	102	2e-021
Peaxi.v0.1.Scf14390 16.4	102	3e-021

>Peaxi.v0.1.Scf25941 16.6  
Length = 42584

Score = 112 bits (281), Expect = 2e-024  
Identities = 55/61 (90%), Positives = 59/61 (96%)  
Frame = +1

```
Query: 1      MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSSHKGKLFYSTD 60
          MGRGRVQ+KRIENKINRQVTFSKRR+GLLKKAHEISVLCDAEV L+VFS KGKLFY+TD
Sbjct: 15391 MGRGRVQMKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFYATD 15570
```

```
Query: 61      S 61
          S
Sbjct: 15571 S 15573
```



Il faut récupérer le bon contig (continuum de séquence) dans la base de données.

Ca se fait par une recherche du titre: CTRL F

```
>Peaxi.v0.1.Scf25941 16.6
TTATCATTATTATTATTGTTGTTGTTATTATTATTGTTATTATTTCCCGTTGTTTCCCTACATAAACATCGTTACAAAAATTTAAAGATTATATGTTT
AAATTCACGTGTCAAATGAACTCATTTGACTCTCGAAATCCGAACTATGACAACCAACTTAAAACGGAGGAAGTACTAAATATAGAAAGGCACGCTCGTG
AAAACATACTAAATAAAAAGTAAGACACATAAAATTTAAACGGATGAAATAAATGAATTTAAAAATTTATTCAAAGACTTTTTGAGGTCTGAGAGTCAAC...
```

Puis, on aligne la séquence protéique d'AP1 d'*Arabidopsis* avec la séquence nucléotidique génomique du gène correspondant de *Petunia*.

On utilise le programme **WISE2**

GeneWise compare une séquence protéique à une séquence nucléotidique génomique, en établissant les positions des exons et des introns.

<http://www.ebi.ac.uk/Tools/psa/genewise/>

On obtient:

Le début et la fin du gène sur la séquence nucléotidique proposée:

Gene 501 12613

La position des différents exons (et donc des introns), avec la phase correcte de traduction:

Exon 501 685	phase 0
Exon 10070 10148	phase 2
Exon 10615 10670	phase 0
Exon 10741 10840	phase 2
Exon 11272 11313	phase 0
Exon 12342 12383	phase 0
Exon 12503 12613	phase 0 → 7 exons

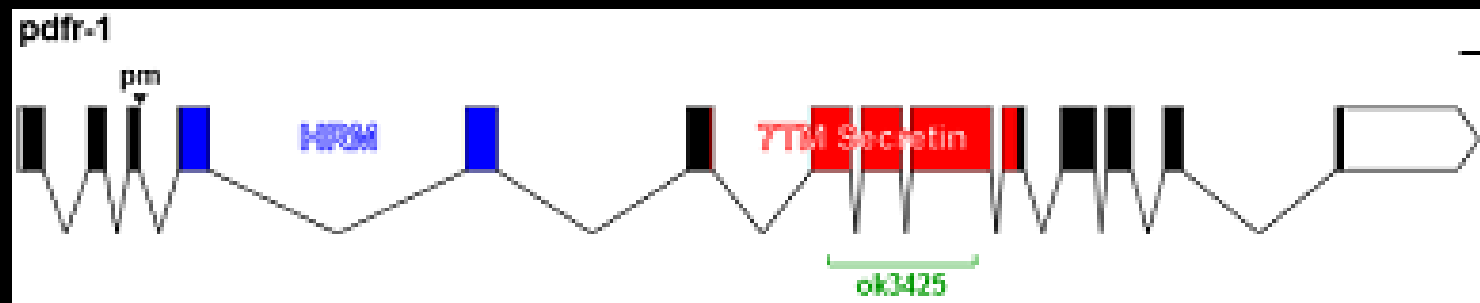
La séquence nucléotidique du gène correspondant à la protéine apparentée recherchée:

```
ATGGGGAGGAAGAAGGCTGAAATAAAGCTAATCCAAGACAAGAACAGTAGGCAAGTGACA
TTTTCAAAAAGAAGGAAAGGACTCATAAAGAAAAGCCAAAAGAACTCTCTATACTTTGTGAT
GCAGATGTTGCTGTTGTTGTCTTCTCTAACCGTGGCAGACTCTACGATTTCTCCAGCAAT
AACAGTCTGACGGAGATTGTTCAACGATACCACAGCCACGTGGAAGCAGAAAAAGAGATC
TCTGCCGAAGCTTCTGACGCTGAGCAGTCTAAGTATGCAAGCATCATGACAATAGGAGAA
ATGCTGCAAAACAATACAGAGACAACCTCGAGGAACCTGATGTTGATGATCTCAATGTGACT
GATCTTGTCCATTTGGAAAACCAAATACAAACTGCTCTAATGGAAGCTAGATCCAGGAAG
ACACATTTGATGATGGAATCCCTCAAGAGTCTTCATGAAAAGGAAAATCTGCTGAGAGAA
GAAAACAGAATTCTGGAAGACAAGATATCTATGACCACGAACACCAGAGAAGTAAATGGG
ATGGATTTGGATTTTACTAACCTTGCACCCATCCAACCAATTGGTGAAAAGCAGAGAGCA
ACTTTGAATTTTTTC
```

On peut ensuite traduire cette séquence pour obtenir la séquence de la protéine de *Petunia* paralogue de la protéine AP1 d'*Arabidopsis*.

Un dernier programme **WormWeb** permet d'obtenir une représentation graphique du gène déterminé:

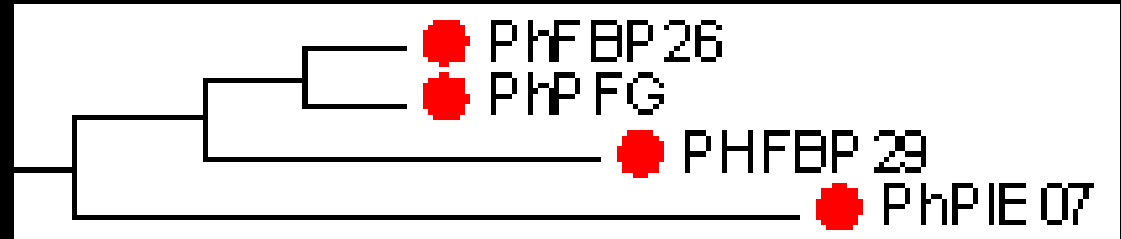
<http://wormweb.org/exonintron>



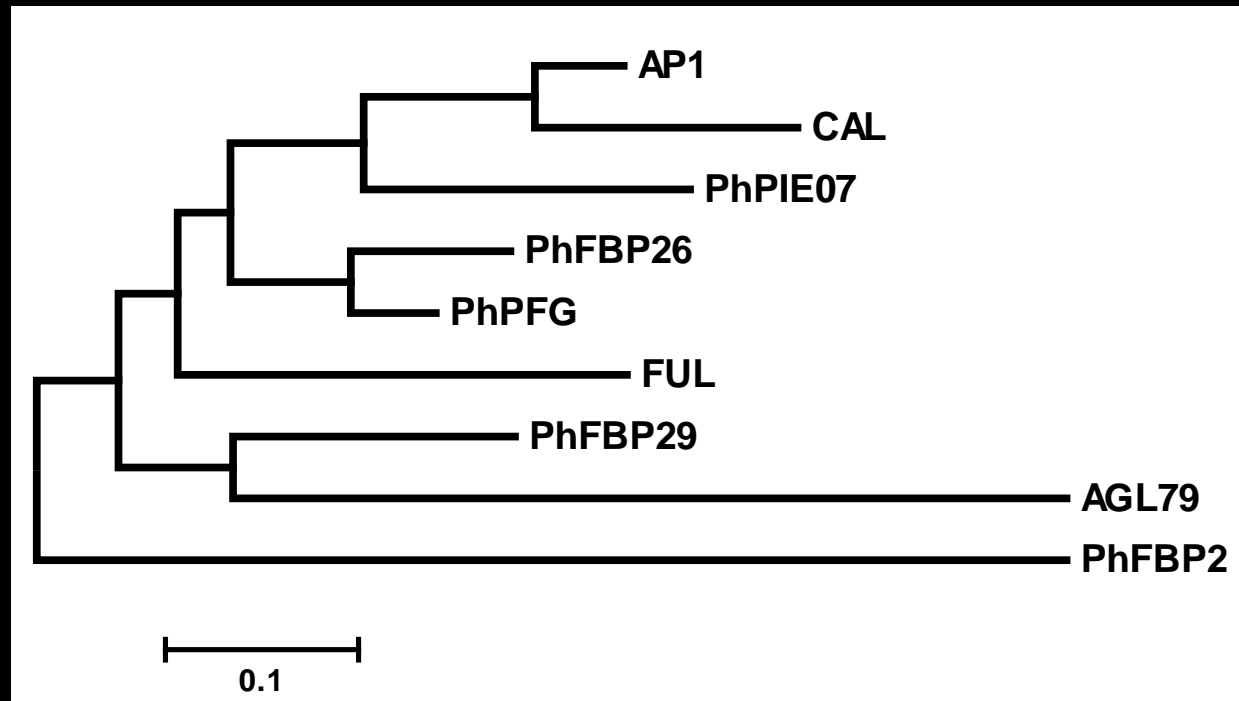
### c. La famille AP1 chez *Petunia*:

Elle est composée de quatre paralogues :

PIE07,  
PFG,  
FPG26,  
FBP29



Qui groupent plus ou moins avec leurs orthologues d'*Arabidopsis*



### 3. Isolement de mutants pour chacun des gènes de la famille AP1

#### a. Recherche dans la collection de séquences flanquantes de transposon et exploitation des résultats bruts:

Recherche faite grâce à blastN (homologie nucléotides/nucléotides) en utilisant les séquences nucléotidiques des gènes de la famille AP1

En fait deux séquences ont été utilisées pour chaque gène:

La séquence codante

La séquence génomique arbitrairement définie comme étant:

la séquence allant de l'ATG au STOP,  
avec 500 pb en amont de l'ATG (séquence 5'UTR),  
avec 200 pb en aval du STOP (séquence 3' UTR).

Nous recherchons les insertions présentes dans les exons ou à proximité immédiate des sites d'épissage → **mutants nuls**.

Les résultats primaires se présentent sous forme d'un listing:

Lib08-B_35087-#(X06-1,X24-1,X25-3,Y38-2,Y47-19,)	1,41E-29
Lib08-B_32891-#(X37-3,Y31-5,)	1,38E-28
Lib08-M_3658-#(X20-1,X38-2,Y05-1,Y06-1,Y13-1,Y17-1,Y35-1,)	1,38E-28
Lib09-M_85711-#(X096-1,Y40-1,)	1,07E-19
Lib09-M_19887-#(X072-3,Y06-1,Y17-2,Y20-1,)	1,07E-19
Lib06-L_199239-#(X2-1,Y3-1,Y8-2,Z3-1,Z5-1,Z7-1,)	1,07E-19
Lib09-B_62749-#(X108-1,X63-9,X69-1,X91-1,Y01-5,Y42-2,)	1,04E-18
Lib09-L_197307-#(X72-4,Y17-1,Y18-1,)	1,01E-17
Lib09-B_218752-#(X101103-1,Y13-1,Y14-1,)	6,22E-06
Lib08-M_289938-#(X28-4,Y12-1,Y14-1,Y22-8,Y36-1,)	6,22E-06



## d. Elimination des séquences redondantes

Lib06-B\_10331-#  
Lib06-B\_10405-#  
Lib06-B\_11902-#  
Lib06-B\_12070-#  
Lib06-B\_12265-#  
Lib06-B\_130653-#  
Lib06-B\_13619-#  
Lib06-B\_13790-#  
Lib06-B\_14105-#  
Lib06-B\_143803-#  
Lib06-B\_145288-#  
...

## e. Récupération des séquences nucléotidiques correspondantes



ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGATATCCATTATATTCATTTGCTTTAGATTATCATATCTT  
ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGATATCCATTATATTCATTTGCTTTAGATTATCATATCTT  
ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGACATCCATTATATTCATTTGCTTTAGATTATCATATCTT  
ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGATATCCATTATATTCATTTGCTTTAGATTATCATATCTT  
ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGATATCCATTATATTCATTTGCTTTAGATTATCATATCTT  
ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGATATCCATTATATTCATTTGCTTTAGATTATCATAT  
ACCATTATATTTTATTGAAGTATTATTGTTGATCATTGATATCCATTATATTCATTTGCTTTAGATTATCATATT

## f. Ultime recherche blastN.

On utilise les séquences flanquantes de transposon comme cribles contre les séquences nucléotidiques des quatre gènes de la famille AP1 de *Petunia*.

	% Homo	Fgt	MM	Gap	Pos Fgt	Pos Fgt	Pos Ctg	Pos Ctg	Début Gène	Ou Gène	Orientation	Library
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	69	0	0	1	69	313	245	306			-68Lib08-B_9182-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	69	0	0	1	69	313	245	306			-68Lib06-B_11902-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	71	0	0	1	71	989	1059	982	989		70Lib08-B_35087-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	52	0	0	1	52	4000	3949	3993			-51Lib09-B_62749-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	69	0	0	1	69	4745	4677	4738			-68Lib08-B_32891-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	69	0	0	1	69	4878	4946	4871	4878		68Lib06-L_1074-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	54	0	0	1	54	4885	4832	4878			-53Lib06-B_143803-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	54	0	0	1	54	5764	5711	5757			-53Lib09-M_19887-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	30	0	0	1	30	5855	5884	5848	5855		29Lib09-M_617966-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	30	0	0	1	30	5855	5884	5848	5855		29Lib08-M_464365-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	52	0	0	1	52	5855	5906	5848	5855		51Lib09-B_117870-
PhAGL6-Peaxi.v0.1.Scf9685_19.6	100.00	55	0	0	1	55	5855	5909	5848	5855		54Lib08-B_264587-

Permet de déterminer l'orientation de l'insertion

Le nombre de séquences pour une même insertion et donc leur origine

La généalogie entre les séquences et donc entre les plantes.

### Conclusions sur la transmission générationnelle des mutations:

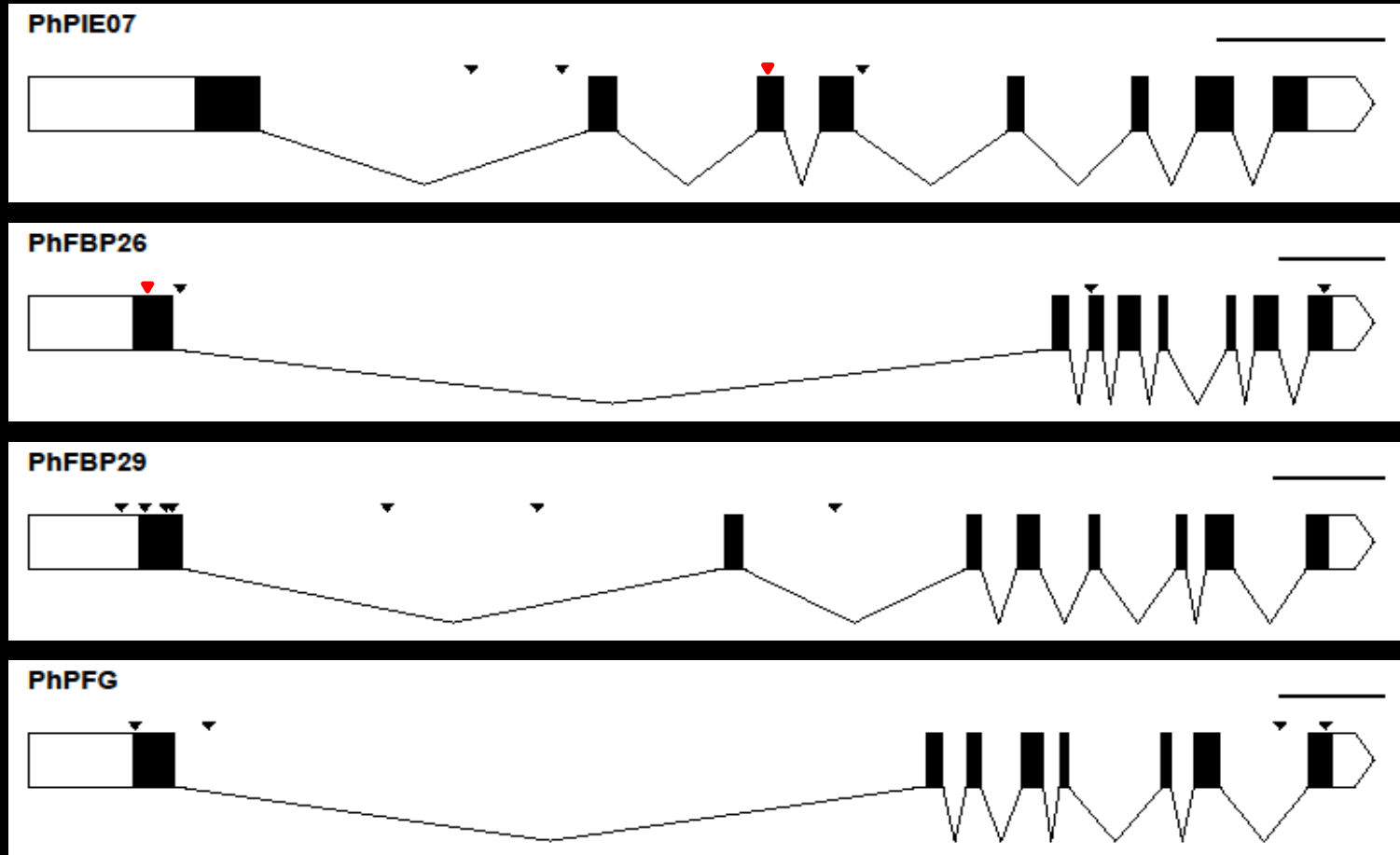
Probabilité de trouver une insertion dans la descendance : 83,69%

Probabilité de trouver une insertion dans la descendance avec les bons coordinats: 98,98%



## 4. Les mutants obtenus pour les gènes de la famille AP1.

On a des mutants KO pour chacun des quatre gènes.



Problème: aucun phénotype n'est visible dans les fleurs!!!!

Combinaisons de mutants : 6 doubles, 4 triples, 1 quadruple.  
Tous ont été construits. Tous produisent des fleurs. Avec des pétales.

### Phénotypes floraux



Sauvage



*pie07 pfg*



*triple fbp26h*



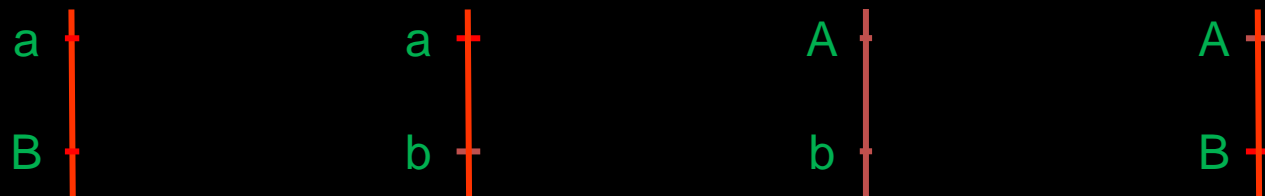
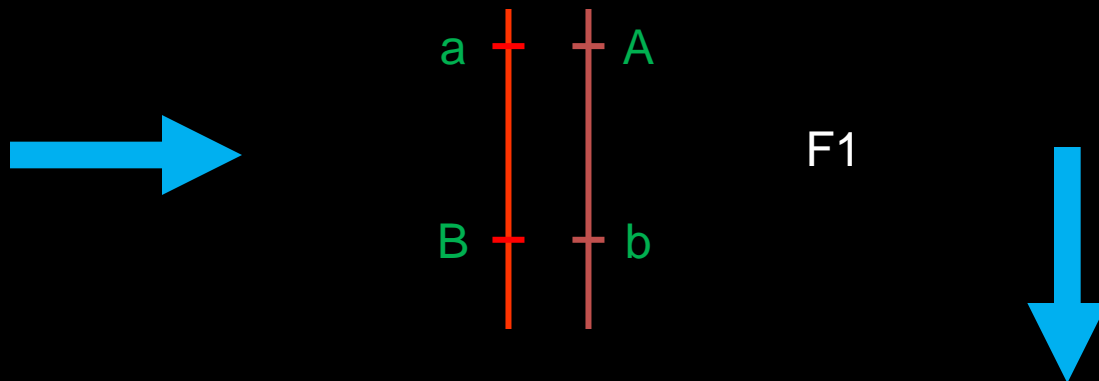
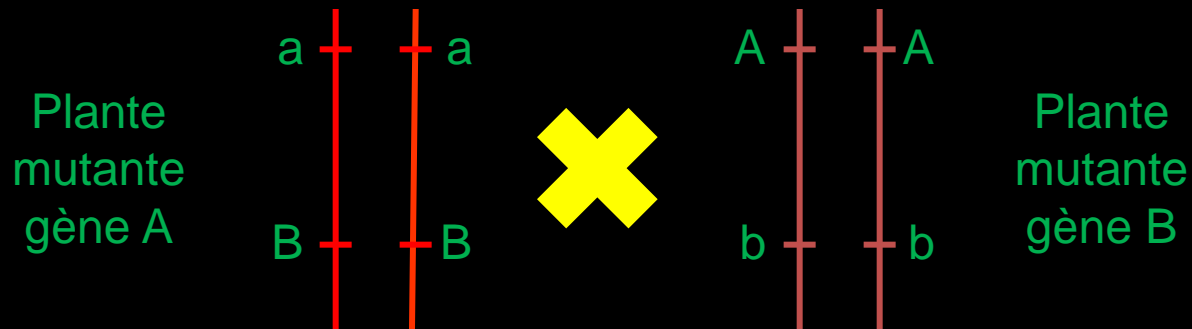
*triple fbp29h*



*triple pfg h*



*triple pie07 h*





a

B

a

b

A

b

A

B

a

B

a

a

B

B

a

a

b

B

A

A

b

B

A

a

B

B

a

b

a

a

B

b

a

a

b

b

A

a

b

b

A

a

B

b

A

b

a

A

B

b

a

A

b

b

A

A

b

b

A

A

B

b

A

B

a

A

B

B

a

A

b

B

A

A

b

B

A

A

B

B

## quadruple



Pas de phénotype de type AP1 (sans pétale).

Pas de perte de l'identité des organes. Même pour le quadruple.

Mais conversion homéotique des sépales en pétales.

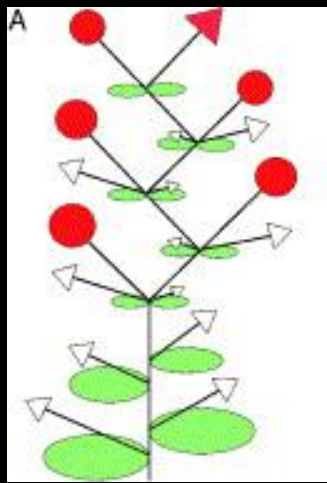
Des fleurs sauvages sont produites dans tous les cas, sauf pour le mutant triple-*pfgh* et le quadruple mutant.

→ On n'a donc pas de fonction pour les gènes *PIE07 (AP1)* dans l'identité du pétale comme chez *Arabidopsis*.

*Petunia*: plante sauvage



Fleur  
Bractée  
Noeud



Phyllotaxie en cyme



*pie07*  
*pfg*

*triple*  
*fbp26h*

*triple*  
*fbp29h*

*triple*  
*pie07h*

*triple*  
*pfgH*

*quadruple*



### Force du phénotype

Altérations phénotypiques sur la plante entière:

Avec apparition de bractée aux inter-nœuds

Diminution du nombre de fleurs

Formation d'une fleur terminale dans les cas les plus sévères

## Quadruple mutant:

- Fleur terminale.
- Les feuilles sont positionnées en spirale le long de la tige
- Perte de l'architecture en cyme typique (zig-zag) du *Petunia*
- Quand la fleur meurt, une nouvelle tige se ré-initie sur le côté. Et une nouvelle fleur apparaît 20 à 30 feuilles plus tard.



Fleur  
terminale



## C. Conclusions:

Chez *Petunia*, les membres de la famille PIE07 (AP1) ne sont pas nécessaires à la formation des pétales.

Ils ne sont pas non plus impliqués dans l'identité du méristème floral.

Par contre, ils ont une fonction dans l'identité du méristème d'inflorescence.

# MADNESS

## Thanks

MICHEL VANDENBUSSCH  
TOM GERATS  
JAN ZETHOF  
CHRISTOPHE TREHIN  
CHARLIE SCUTT  
IOAN NEGRUTIU

ET

JONATHAN LEGRAND  
SOPHY CHAMOT  
AURÉLIE VIALETTE

ONE STEP BEYOND...



<http://www.ens-lyon.fr/RDP/>  
<http://www.ens-lyon.fr/RDP/Evolution-et-Developpement-de-la>  
<http://www.mbio.ncsu.edu/bioedit/bioedit.html>  
<http://wormweb.org/exonintron>  
<http://exon.gatech.edu/eukhmm.cgi>  
<http://www.ebi.ac.uk/Tools/psa/genewise/>  
<http://www.ebi.ac.uk/Tools/msa/clustalw2/>  
<http://www.arabidopsis.org/>  
<http://signal.salk.edu/cgi-bin/tdnaexpress>  
<http://www.arabidopsis.org/servlets/sv>  
[http://solgenomics.net/organism/Solanum\\_lycopersicum/genome](http://solgenomics.net/organism/Solanum_lycopersicum/genome)  
<http://www.ncbi.nlm.nih.gov/pubmed>  
[http://pga.mgh.harvard.edu/web\\_apps/dna\\_utilities.html](http://pga.mgh.harvard.edu/web_apps/dna_utilities.html)  
<http://blast.ncbi.nlm.nih.gov/Blast.cgi>  
<http://biologylabs.utah.edu/jorgensen/wayned/ape/>  
<http://www.editpadlite.com/download.html>  
<http://serial-cloner.en.softonic.com/>  
[http://www.snapgene.com/products/snapgene\\_viewer/](http://www.snapgene.com/products/snapgene_viewer/)