



Institut d'Ecologie - Phytogénétique cellulaire

Développement d'outils de la mutagenèse ciblée par recombinaison  
homologue chez *Physcomitrella patens*

Thèse de doctorat

présentée à la

Faculté des Sciences de  
l'Université de Lausanne

par

Mikhail Chakhparonian

Diplômé en Biologie  
Université de Lausanne

**Jury**

Prof. François Marillier, Président  
Prof. Jean-Pierre Zrýd, Directeur de thèse  
Prof. Michel Laloue, Expert  
Dr Andrzej Stasiak, Expert  
Dr Didier Schaefer, Expert

LAUSANNE  
2001

## Remerciements

Au Professeur Jean-Pierre Zrýd de m'avoir donné les moyens d'effectuer ce travail dans son laboratoire et de m'avoir motivé, volontairement ou non, à m'initier à l'informatique, à l'enseignement, et à d'autres emplois, aussi fascinants qu'utiles. Je lui suis reconnaissant de m'avoir donné une vision nouvelle de la science et de ses buts.

Au Dr Didier Schaefer de m'avoir ouvert les yeux sur la vraie démarche scientifique, rigoureuse, exacte, réfléchie et, par-là, à la fois passionnante et ennuyeuse. Je lui suis redevable d'avoir gracieusement illustré une partie de mon travail: son sens de la photographie et de la réalisation de l'analyse Southern sont notoires.

Au Professeur Michel Laloue ainsi qu'au Dr Andrzej Stasiak d'avoir accepté la tâche de juger ce travail.

Aux collaborateurs du LPC qui m'ont aidé, que ce soit par action ou par omission, par leur exemple ou par leur contre-exemple, par leur encouragement (toujours bienvenu) ou par leur critique (toujours déplacée), à avancer sur mon chemin. Merci aux étudiants, L. Tchougounova, M. Bessire, J.-C. Wyss, qui ont contribué à ce travail par quelques expériences soigneusement effectuées, ainsi qu'au Dr Susan Vlach de m'avoir aidé lors de son séjour au LPC.

Aux collaborateurs des laboratoires du Professeur Michel Laloue et du Professeur Enzo Russo d'avoir coopéré dans l'exploration de la mousse *P. patens* et d'avoir effectué pour moi les analyses biochimiques et l'analyse de séquences.

A l'Office Fédéral de l'Education et de la Science et aux contribuables suisses et canadiens de leur soutien financier et administratif pendant différentes périodes de mon travail.

A mes parents, qui se sont contentés de mes appels téléphoniques irréguliers et qui m'ont apporté l'encouragement et l'envie de persévérer.

A mes amis, qui n'acceptent pas les plantes transgéniques mais qui me supportent quand même. A mon amie Julie Prussacova de s'être penchée avec moi sur quelques "pourquoi" très importants. A mon ami Laurent, sans qui tout cela aurait été beaucoup trop triste.

---

Life is much too important a thing ever to talk seriously about it.

*O. Wilde*

<b>RESUME .....</b>	<b>5</b>
<b>ABSTRACT.....</b>	<b>6</b>
<b>INTRODUCTION GENERALE .....</b>	<b>8</b>
1. LA MUTAGENÈSE CIBLÉE PAR RECOMBINAISON HOMOLOGUE EST L'OUTIL LE PLUS FIN DE LA GÉNÉTIQUE INVERSE .....	8
2. RECOMBINAISON HOMOLOGUE CHEZ LES PLANTES.....	9
2.1. <i>Recombinaison homologue chez les angiospermes</i> .....	9
2.2. <i>Recombinaison homologue chez Physcomitrella patens</i> .....	11
3. L'APPLICATION LA PLUS FINE DE LA MUTAGENÈSE CIBLÉE EST LA MODIFICATION PRÉCISE DES SÉQUENCES GÉNOMIQUES À UN NUCLEOTIDE PRÈS .....	12
<b>CHAPITRE I : EVALUATION DE L'EFFICACITÉ, DE LA PRÉCISION ET DE LA SPÉCIFICITÉ DE LA MUTAGENÈSE CIBLÉE PAR RECOMBINAISON HOMOLOGUE: LE GÈNE ZLAB1 DE LA FAMILLE MULTIGÉNIQUE CAB CODANT POUR LA PROTÉINE MAJEURE DE LIAISON DE LA CHLOROPHYLLE A/B ("CHLOROPHYLL A/B BINDING PROTEIN") .....</b>	<b>14</b>
1. INTRODUCTION .....	14
1.1. <i>Spécificité de la recombinaison homologue</i> .....	14
1.2. <i>ZLAB1 est un membre d'une famille de gènes codant pour la protéine majeure de liaison de la chlorophylle a/b ("chlorophyll a/b binding protein")</i> .....	16
2. RESULTATS.....	18
2.1. <i>Stratégie de la mutagenèse ciblée</i> .....	18
2.2. <i>Isolation des transformants intégratifs</i> .....	18
2.3. <i>Analyse moléculaire d'un échantillon des transformants intégratifs</i> .....	19
3. DISCUSSION .....	25
<b>CHAPITRE II : MISE EN ÉVIDENCE DE LA RÉPARATION DES SÉQUENCES INTERROMPUES ("GAP REPAIR") CHEZ P. PATENS .....</b>	<b>27</b>
1. INTRODUCTION .....	27
1.1. <i>Réparation des séquences interrompues</i> .....	27
1.2. <i>ZLAB1 et λ213 permettent d'étudier la réparation des brèches</i> .....	29
2. RESULTATS ET DISCUSSION .....	30
2.1. <i>Stratégie de transformation</i> .....	30
2.2. <i>Analyse moléculaire d'un échantillon des transformants intégratifs</i> .....	32
2.3. <i>Classification des événements et conclusions</i> .....	41
<b>CHAPITRE III : MUTAGENÈSE CIBLÉE DIRECTE DU GÈNE PPAPT CODANT POUR L'ADÉNINE PHOSPHORIBOSYLTRANSFERASE À L'AIDE DE VECTEURS DE REMPLACEMENT .....</b>	<b>44</b>
1. INTRODUCTION .....	44
1.1. <i>Disposition d'un vecteur de remplacement: paramètres à considérer</i> .....	44
1.2. <i>Caractéristiques du gène PpAPT</i> .....	45
2. RESULTATS.....	47
2.1. <i>Isolation de mutants APRT<sup>-</sup></i> .....	47
2.2. <i>Stratégie de la mutagenèse ciblée</i> .....	48
2.3. <i>Isolation des transformants intégratifs</i> .....	50
2.4. <i>Analyse moléculaire d'un échantillon des transformants intégratifs</i> .....	51
3. ANALYSE DU PHÉNOTYPE DES DIFFÉRENTS ALLELES DU GÈNE PPAPT OBTENUS PAR MUTAGENÈSE CIBLÉE	59
3.1. <i>Morphologie et développement des mutants DAP<sup>R</sup> en conditions standards</i> .....	59
3.2. <i>Explications possibles des altérations du phénotype des mutants DAP<sup>R</sup>-I</i> .....	62
3.3. <i>Transcription des allèles PpAPT mutés</i> .....	67
3.4. <i>Complémentation de l'allèle muté XΔX-R (APT)</i> .....	69
4. DISCUSSION .....	72
4.1. <i>L'efficacité de la mutagenèse ciblée</i> .....	72
4.2. <i>Types d'intégration des quatre vecteurs XΔX, BΔB, AΔB, AΔBneo</i> .....	75
4.3. <i>Conclusion</i> .....	78

<b>CHAPITRE IV. EXCISION DE MARQUEURS DE SELECTION A L'AIDE D'UN SYSTEME DE RECOMBINASE SPECIFIQUE D'UN SITE CRE/LOX .....</b>	<b>80</b>
1. INTRODUCTION .....	80
1.1. <i>Problèmes liés à l'utilisation de gènes marqueurs de sélection</i> .....	80
1.2. <i>Techniques permettant d'enlever les marqueurs de sélection</i> .....	81
1.3. <i>Recombinase spécifique d'un site</i> .....	82
1.4. <i>Système de recombinaison spécifique d'un site Cre/lox</i> .....	83
2. RESULTATS ET DISCUSSION .....	88
2.1. <i>Intégration d'un marqueur de sélection positive bordé de deux sites loxP parallèles dans le génome de P. patens</i> .....	88
2.2. <i>Excision du marqueur de sélection 35S nptII intégré dans les souches alox-17, alox-4N et alox-3RN</i> 91	
2.3. <i>Conclusion</i> .....	103
<b>CHAPITRE V: UTILISATION DE PROMOTEURS INDUCTIBLES ET CONSTITUTIFS CHEZ P. PATENS .....</b>	<b>107</b>
1. INTRODUCTION .....	107
1.1. <i>Expression constitutive</i> .....	107
1.2. <i>Expression conditionnelle</i> .....	108
2. RESULTATS ET DISCUSSION .....	112
3. PERSPECTIVES .....	120
<b>DISCUSSION .....</b>	<b>122</b>
1. LA MUTAGENÈSE CIBLÉE CHEZ P. PATENS : ETAT DES LIEUX .....	122
1.1. <i>De quels facteurs la mutagenèse ciblée dépend-elle?</i> .....	123
1.2. <i>Les modèles qui décrivent le mécanisme de l'intégration de l'ADN transformant chez P. patens</i> 126	
2. WHY ME ? OU POURQUOI LA MUTAGENÈSE CIBLÉE PAR RECOMBINAISON HOMOLOGUE EST-ELLE SI EFFICACE CHEZ P. PATENS ET NON CHEZ LES ANGIOSPERMES ? .....	128
2.1. <i>Les tentatives de développer la mutagenèse ciblée par recombinaison homologue chez les angiospermes</i> .....	128
2.2. <i>Pourquoi P. patens est-elle la seule plante terrestre qui intègre l'ADN exogène dans le génome préférentiellement par recombinaison homologue?</i> .....	130
3. RÔLE DE P. PATENS DANS LA BIOLOGIE VÉGÉTALE ET DANS LA BIOLOGIE DES EUKARYOTES PLURICELLULAIRES .....	131
<b>MATÉRIEL ET MÉTHODES.....</b>	<b>134</b>
<b>ANNEXE A: LISTE DES AMORCES PCR.....</b>	<b>136</b>
<b>ANNEXE B: CARTES DES PLASMIDES.....</b>	<b>137</b>
<b>ANNEXE C: ANALYSE DÉTAILLÉE DU RFLP DE DIFFÉRENTS ALLELES .....</b>	<b>143</b>
<b>BIBLIOGRAPHIE.....</b>	<b>149</b>
<b>CURRICULUM VITAE.....</b>	<b>162</b>

## Résumé

L'accomplissement des programmes mondiaux de séquençage du génome de l'homme et de la plante *A. thaliana* fait la une de tous les journaux. La fonction d'une grande partie des séquences nouvellement trouvées est inconnue. Une modification ciblée et précise de ces séquences in vivo permettrait d'étudier leur fonction. L'outil nécessaire à ces modifications est la mutagenèse ciblée par recombinaison homologue. Les techniques utilisées actuellement dans la biologie végétale telles que l'ARN "antisens" pour modifier l'expression des gènes étudiés ou les collections de mutants étiquetés chez *A. thaliana* ou *O. sativa*, sont longues, imprécises et stochastiques.

La mousse *P. patens* se démarque de tous les systèmes expérimentaux végétaux étudiés jusqu'ici par sa capacité d'intégrer l'ADN transformant préférentiellement par recombinaison homologue. Les outils de la mutagenèse ciblée sont à portée de main des biologistes végétaux. Dans le présent travail, nous avons démontré que tous les éléments nécessaires à la mutagenèse ciblée et précise sont réunis chez *P. patens*.

- La mutagenèse ciblée par recombinaison homologue est spécifique et précise chez *P. patens*: nous l'avons démontré en inactivant spécifiquement un gène membre d'une grande famille multigénique.
- Nous avons étudié l'intégration des séquences homologues interrompues ("gap repair"). Les séquences interrompues par une brèche sont majoritairement réparées alors que celles interrompues par une délétion le sont beaucoup moins fréquemment. Le fait que des mutations dans l'ADN transformant soigneusement choisies ne sont pas systématiquement corrigées lors de l'intégration nous permet de les introduire dans le génome.
- Les événements de conversion génique (remplacement simple) nécessaires au développement de la mutagenèse ciblée de précision constituent un quart des événements d'intégration des vecteurs de remplacement. Les événements de remplacement multiple ou d'insertion ont lieu avec des fréquences similaires.
- La mousse *P. patens* étant un organisme haplobiontique, la mutagenèse ciblée des gènes essentiels doit employer un système de régulation flexible. Nous avons démontré que plusieurs promoteurs hétérologues constitutifs ou inductibles sont actifs chez *P. patens*.
- La mutagenèse de précision exige des outils permettant d'enlever les séquences exogènes indésirables telles que les marqueurs de sélection. Nous avons montré que le système de recombinaison spécifique de site Cre/lox permet de les éliminer de façon précise et efficace.

Nous avons généré, par remplacement d'allèle ciblé, plusieurs mutants du gène codant pour l'enzyme adénine phosphoribosyltransférase. Ces mutants présentent un phénotype altéré en raison d'une perturbation du métabolisme de l'adénine. Les gènes d'*A. thaliana* *AtAPT1* et *AtAPT2* complètent génétiquement cette mutation. Cette étude mettant en jeu des allèles mutés in vitro montre que la mutagenèse de précision rend la génétique inverse efficace chez *P. patens* et que les gènes d'*A. thaliana* sont parfaitement fonctionnels chez *P. patens* - soulignant le fait que les mousses ont une biologie semblable à celle des angiospermes.

L'étude approfondie de la recombinaison homologue chez *P. patens* pourrait permettre de comprendre pourquoi ce mécanisme n'est pas efficace chez les autres plantes. La production d'un angiosperme capable d'intégrer l'ADN transformant par recombinaison homologue, sans que cela n'affecte son phénotype, ne sera pas tâche facile. Les biologistes peuvent dès aujourd'hui profiter des avantages uniques du modèle expérimental *P. patens*.

## Abstract

Sequencing of the genomes of the plant *A. thaliana* and of the human is one of the hottest subjects of today. Now that we know the sequence of a gene how can we find out its function? Functional studies rely on targeted modification of these known sequences in vivo and the most precise tool for these studies is gene targeting by homologous recombination. Methods such as antisense silencing or tagged mutant collections currently used for *A. thaliana* or *O. sativa* are stochastic thus incapable of producing precisely all desired mutations.

Integration of foreign sequences in *P. patens* genome by homologous recombination is unique in plant biology and vital for the development of gene targeting methods. Here, we have demonstrated that all the elements necessary for precise and direct gene targeting are available for plant biologists:

- Gene targeting is highly specific and precise in *P. patens* – we have demonstrated that by targeting one member of a large multigene family.
- We studied integration of homologous sequences interrupted by a gap or a deletion. Gaps are frequently repaired whereas deletions rarely are. The fact that sequences bearing a planned deletion are not systematically repaired means they can be introduced into the genome without modification by gap repair mechanisms.
- We have determined that as many as a quarter of all replacement vector integration events are simple gene conversions (replacements), the events that are necessary for precision mutagenesis. Replacement vectors also integrate by insertion or multiple replacement with similar frequencies.
- The moss *P. patens* is haplobiont and knocking out essential genes needs flexible regulation systems. We have shown that different constitutive and inducible promoters are available for *P. patens*.
- Precise mutagenesis requires tools to remove foreign sequences from transgene lines. We have demonstrated that selectable markers can be efficiently and precisely removed by Cre/lox site specific recombination system.

Several mutants of the gene coding for adenine phosphoribosyltransferase have been generated by allele replacement (gene conversion). The mutations confer an altered phenotype caused by abnormal adenine metabolism and can be complemented by homologous *A. thaliana* genes *AtAPT1* and *AtAPT2*. This study demonstrates that fine direct mutagenesis allows efficient reverse genetics in *P. patens* and underlines how close moss biology is to that of angiosperms.

Studies of homologous recombination in *P. patens* may help to elucidate the factors responsible for its efficiency in this moss. Making an angiosperm capable of integrating foreign DNA by homologous recombination without altering its phenotype is a challenging task. Meanwhile, biologists can take advantage of the unique experimental model based on *P. patens*, particularly for functional genomics.

Abréviations :

WT	type sauvage	ES	cellules de lignée embryonique de la souris
RTF	fréquence relative de transformation	CHO	cellules ovariennes du hamster chinois
NHEJ	ligature d'extrémités indépendante de l'homologie	PEG	polyéthylène-glycol
PCR	réaction de polymérisation en chaîne	BSA	albumine sérique bovine
35S CaMV	transcrit 35S du virus de la mosaïque du chou fleur ou son promoteur	RFLP	polymorphisme de longueur des fragments de restriction
<i>nptII</i> ( <i>hpt</i> ou <i>aph IV</i> )	néomycine (hygromycine) phosphotransférase	kb	mille paires de base
<i>Cab</i>	gène codant pour la protéine majeure de liaison de la chlorophylle a/b	bp	paire de base
<i>ZLAB1</i>	gène membre de la famille cab	ADNc, ADNg	ADN complémentaire et génomique
<i>PpAPT</i>	gène <i>APT</i> de <i>P. patens</i>		
<i>APT</i>	adénine phosphoribosyltransférase, gène (APRT - l'enzyme)		
DAP	2,6 diaminopurine	modèles :	
GUS	$\beta$ -glucuronidase	DSBR	réparation de ruptures double brin (selon Szostak)
GFP	protéine verte fluorescente	SDSA	appariement de brins dépendant de la synthèse
CAT	chloramphénicol acétyltransférase	SSA	appariement d'un brin simple
		OSI	invasion d'une extrémité

## Introduction générale

Le présent travail est consacré au développement d'outils de la mutagenèse ciblée de précision chez la mousse *Physcomitrella patens*.

Les paragraphes suivants passent en revue les divers aspects de la mutagenèse ciblée et son application dans la biologie végétale.

### 1. **La mutagenèse ciblée par recombinaison homologue est l'outil le plus fin de la génétique inverse**

La génétique classique recherche les gènes responsables des phénotypes connus. En revanche, la génétique inverse explore la contribution des gènes et de leurs modifications au phénotype de l'organisme (Berg, 1991). La fonction d'un gène peut être déduite de l'analyse de sa séquence dans certains cas. Ces études sont rarement précises et fiables et doivent toujours être confirmées par une analyse de la fonction du gène *in vivo*. Cette analyse repose sur des méthodes qui permettent de modifier *in vivo* le gène étudié de façon aléatoire ou ciblée (Bouchez and Höfte, 1998; Somerville and Somerville, 1999).

La mutagenèse par étiquetage à l'aide d'insertion d'éléments transposables ou de l'ADN-T est une des techniques permettant d'altérer la séquence des gènes endogènes de façon **aléatoire**. La mutagenèse par insertion des éléments de régulation ("enhancers" ou "silencers") ou de gènes rapporteurs dépourvus de promoteur permet de modifier et d'observer *in vivo* l'expression des gènes. Les individus dans lesquels le gène étudié est modifié peuvent être retrouvés dans les collections de mutants étiquetées à saturation. Le désavantage de ces techniques réside principalement dans l'impossibilité d'introduire les mutations ponctuelles ou planifiées avec précision; de plus, la production d'une collection comportant au moins une mutation de chaque gène d'un organisme (saturée à 100%) est une tâche fastidieuse (Bouchez and Höfte, 1998; Parinov and Sundaresan, 2000).

Un autre type d'études fonctionnelles repose sur la manipulation des produits de gènes à l'aide de leur sur-expression ou répression (Beertling, 1995; Potrykus and Spangenberg, 1995; Somerville and Somerville, 1999). Dans les deux cas, une cassette d'expression (du gène ou de sa copie antisens) doit être intégrée dans le génome. Une intégration non ciblée peut provoquer des mutations imprévues, donc indésirables, et l'expression de la cassette peut subir l'influence du contexte génomique inconnu ("effet de position"). Les cassettes de sur-expression sont, de surcroît, sujettes à des phénomènes de co-suppression stochastiques. L'expression antisens, elle, est contrôlée dans la majorité des cas par un promoteur exogène dont la régulation spatio-temporelle est différente de celle du gène que la cassette est censée réprimer. Elle peut aussi, dans le cas des gènes appartenant à une famille multigénique, inhiber l'expression de plusieurs gènes alors qu'un seul gène est étudié (Offringa and Hooykaas, 1995).

En revanche, la mutagenèse **ciblée** permet de modifier *in vivo* le gène étudié, de façon ciblée, précise et spécifique. Les techniques de mutagenèse **ciblée** se servent de la recombinaison homologue dans les cellules somatiques où elle est associée à la réparation de l'ADN

endommagé par des ruptures double brin; la recombinaison homologue est en effet induite par ces dernières. L'ADN transformant "ciblant" est introduite dans la cellule où il modifie ou remplace l'ADN endogène (une partie ou la totalité d'un gène ou ses séquences régulatrices) par recombinaison homologue de façon précise et ciblée. La fonction et la régulation des gènes peuvent alors être étudiées *in vivo* dans les mutants planifiés (Beertling, 1995).

*L'efficacité de la mutagenèse ciblée est définie par l'équilibre entre l'intégration par recombinaison homologue et l'intégration "illégitime"*

L'efficacité de la mutagenèse ciblée est caractérisée par la proportion des événements d'intégration ciblée par recombinaison homologue parmi tous les événements d'intégration de l'ADN transformant (exprimée en pourcentage). Elle est liée au choix de la voie majoritaire de la réparation des ruptures double brin d'ADN dans les cellules somatiques. En effet, la recombinaison entre l'ADN exogène et l'ADN chromosomique est stimulée par les ruptures double brin et les facteurs de la réparation sont directement impliqués dans l'intégration (Paques and Haber, 1999). La compétition entre deux voies de réparation, la recombinaison homologue et la ligature d'extrémités indépendante de l'homologie (NHEJ) définit l'aptitude d'un organisme à la mutagenèse ciblée. Haber fait l'hypothèse que les complexes protéiques Ku et Rad52p sont les "gardiens" respectifs des voies de NHEJ et de la recombinaison homologue (Haber, 2000). Les déterminants génétiques des deux voies de réparation des ruptures double brin sont connus (groupe épistatique de *RAD52* et de *Ku* (Paques and Haber, 1999)), mais il reste à découvrir quels sont les facteurs exacts responsables de l'équilibre entre les deux (Lee *et al.*, 1999).

Schaefer (Schaefer, 2001) répertorie l'efficacité de la mutagenèse ciblée chez différents organismes. Il souligne que l'intégration de l'ADN transformant se fait dans la majorité des eucaryotes par recombinaison illégitime, donc dans des sites aléatoires. Les organismes qui se démarquent par une efficacité de la mutagenèse ciblée utilisable (de 1 à 100%) sont les levures *S. cerevisiae* et *S. pombe*, les champignons filamenteux, dont *A. nidulans* ou *N. crassa*, le groupe de Trypanosomatidae et, chez les vertébrés, les cellules d'embryon de la souris (ES) et une lignée cellulaire de lymphocyte B du poulet.

Chez la plupart des végétaux, l'efficacité de la mutagenèse ciblée est pratiquement nulle à cause de la dominance de l'intégration illégitime de l'ADN transformant par NHEJ. Nous passons en revue la mutagenèse ciblée chez les plantes dans les paragraphes suivants.

## **2. Recombinaison homologue chez les plantes**

### **2.1. Recombinaison homologue chez les angiospermes**

Chez les plantes, les chercheurs mettent en évidence la recombinaison homologue entre des fragments d'ADN essentiellement par des études utilisant de gènes marqueurs. La plupart des expériences visent à démontrer que la recombinaison entre deux fragments du marqueur se recouvrant partiellement peut restituer un marqueur fonctionnel. Diverses études démontrent que la recombinaison extrachromosomique est la plus fréquente (Ohl *et al.*, 1995; Puchta and Hohn, 1996; Puchta and Meyer, 1995; Vergunst and Hooykaas, 1999). Jusqu'à 30 % des molécules incorporées dans la cellule se recombinent peu après le transfert chez *N. plumbaginifolia*. La co-intégration fréquente dans le même locus de deux ADN différents qui

partagent un fragment homologue indique que la recombinaison extrachromosomique est efficace (Hrouda and Paszkowski, 1994).

La fréquence des recombinaisons homologues intrachromosomiques est comprise entre  $10^{-4}$  et  $10^{-7}$  recombinaisons par division cellulaire (Lichtenstein *et al.*, 1995; Vergunst and Hooykaas, 1999). Dans ces études, les chercheurs observent des réarrangements complexes imprévus qui accompagnent la recombinaison intrachromosomique (Puchta and Hohn, 1996).

L'interaction entre deux fragments incomplets d'un marqueur (transformant et chromosomique) qui se recouvrent permet de sélectionner les événements de recombinaison homologue entre l'ADN transformant et l'ADN chromosomique. Celle-ci doit mener à l'intégration de l'ADN exogène dans la région homologue génomique et à la réparation du marqueur. La réparation du marqueur est donc directement sélectionnable si elle est précise.

Paszkowski et collaborateurs ont employé le gène de résistance à la kanamycine *nptII* comme substrat de recombinaison chez *N. tabacum* (Paszkowski *et al.*, 1988). Ils montrent les meilleurs résultats dans une souche portant environ 7 copies d'un fragment incomplet offrant 352 bp d'homologie avec l'ADN extrachromosomique. L'étude ne fournit pas d'analyse confirmant en détail la conformité de l'intégration aux prédictions, probablement à cause de sa complexité. Les auteurs comparent la fréquence de l'intégration ciblée avec celle observée pour l'intégration du vecteur portant le gène *nptII* entier et estiment que l'efficacité de la mutagenèse ciblée est d'environ 0.03%.

Halfter et collaborateurs ont utilisé le gène de résistance à l'hygromycine *hpt* comme substrat de recombinaison chez *A. thaliana* (Halfter *et al.*, 1992). L'intégration ciblée d'un vecteur portant la séquence ininterrompue de *hpt* sans promoteur ni signal de polyadénylation répare une cassette d'expression du gène *hpt* interrompu par une délétion de 19 bp préalablement introduite dans le génome. L'analyse par PCR démontre le remplacement de la copie défectueuse dans 4 clones sur  $3.4 \cdot 10^8$  protoplastes transformés. Selon le calcul des auteurs, l'efficacité de la mutagenèse ciblée est d'environ 0.03%.

Lee et collaborateurs ont étudié la recombinaison dans un locus génomique des gènes *SuRA* et *SuRB* codant pour deux isoenzymes de l'isolactate synthase, dont une mutation confère la résistance à l'herbicide sulfonylurée (Lee *et al.*, 1990). Ils recherchent l'intégration par recombinaison homologue d'une copie mutée du gène *SuRB* dans un des deux gènes cibles. L'analyse Southern a confirmé que 3 clones résistants à l'herbicide ont pu être le résultat d'une recombinaison selon leurs prédictions. L'analyse génétique de la descendance homozygote pour le locus muté a révélé la présence d'une copie intacte du locus de type sauvage et donc n'a pas confirmé cette conclusion. Ce type d'événement (appelé "hit and run") comprend probablement la correction du vecteur selon la copie homologue génomique suivie d'une intégration ectopique illégitime.

Offringa et collaborateurs montrent des événements similaires "hit and run" chez *N. tabacum* en utilisant le gène *nptII* comme substrat de recombinaison (Offringa and Hooykaas, 1995). Le fragment incomplet intégré préalablement dans le génome porte une délétion 3' du gène. Les chercheurs tentent de mettre en évidence sa réparation par recombinaison homologue avec un fragment extrachromosomique avec 3.6 kb d'homologie. L'analyse PCR confirme la réparation ciblée par recombinaison homologue de ce fragment dans 3 clones sur  $2.3 \cdot 10^6$  protoplastes transformés. Or, dans 2 d'entre eux, la copie réparée reste sur le vecteur qui s'est intégré ailleurs dans le génome. C'est un autre exemple des événements "hit and run".

Les efforts continus pour appliquer la mutagenèse ciblée aux plantes ont également eu pour résultat l'intégration ciblée dans un locus naturel, un facteur de transcription *TGA3* chez *A. thaliana* (Miao and Lam, 1995). Les chercheurs utilisent un vecteur de transformation qui porte le gène marqueur de sélection positive (35S *nptII*) entre deux fragments homologues aux locus TGA3 de 3 et 4 kb, respectivement. Les événements d'insertion illégitime sont contre-sélectionnés à l'aide d'un marqueur supplémentaire se trouvant sur le vecteur en dehors de l'homologie (GUS). Une analyse Southern confirme que dans 1 mutant sur 2580 cals transformants intégratifs, un allèle du locus TGA3 a été interrompu par remplacement. La régénération d'une lignée homozygote avec la mutation ciblée n'a pas été effectuée.

Kemplin et collaborateurs démontrent une intégration ciblée dans un autre locus de *A. thaliana*, *AGL5* (Kemplin *et al.*, 1997). Le vecteur utilisé porte 3 et 2 kb d'homologie en 5' et 3' d'un gène marqueur de sélection, respectivement. Les analyses génétiques et moléculaires permettent d'identifier 1 cas d'intégration ciblée sur 750 transformants intégratifs. Les auteurs ne rapportent pas la fréquence précise de cet événement.

L'efficacité de la mutagenèse ciblée mesurée comme rapport entre l'intégration ciblée et tous les événements d'intégration atteint donc à peine 0.03%. De surcroît, la structure du locus recombinant n'a été conforme aux prédictions que dans une seule étude (Halfter *et al.*, 1992). Les deux cas isolés d'intégration ciblée dans un locus naturel ne forment pas de base statistiquement valable pour une conclusion sur l'efficacité de la mutagenèse ciblée.

L'intégration de l'ADN exogène chez les angiospermes a donc lieu préférentiellement dans des locus aléatoires, par intégration "illégitime" opérée par la voie dominante de réparation de ruptures double brin - NHEJ. L'analyse de l'intégration des séquences homologues révèle, en effet, que la structure du locus recombinant ne peut pas toujours être prédite précisément (Britt, 1999; Gorbunova and Levy, 1997; Risseuw *et al.*, 1995; Salomon and Puchta, 1998; Takano *et al.*, 1997). L'importance des mécanismes d'intégration de l'ADN dans le développement de la mutagenèse ciblée sera discutée dans notre travail.

En somme, l'efficacité de la mutagenèse ciblée dans une approche directe chez les angiospermes est beaucoup trop faible pour permettre le développement d'un outil fiable et précis (Mengiste and Paszkowski, 1999; Morton and Hooykaas, 1995; Vergunst and Hooykaas, 1999).

## 2.2. Recombinaison homologue chez *Physcomitrella patens*

*P. patens* est une mousse (Funariales, Bryophyta). Les revues récentes de Reski et de Cove donnent une description détaillée du développement, de la génétique et de la biologie moléculaire des mousses (Cove *et al.*, 1997; Reski, 1998a). La biologie et la méthodologie de *P. patens* sont traitées dans les publications de notre laboratoire (Schaefer, 1994; Schaefer *et al.*, 1994; Schaefer *et al.*, 1991) ainsi que dans celles de Cove (Cove, 1992; Cove *et al.*, 1991; Cove and Knight, 1993) et Reski (Reski, 1999).

Schaefer démontre que l'intégration de l'ADN transformant dans le génome de *P. patens* a lieu par recombinaison homologue avec une efficacité de 66-100% par rapport à tous les cas d'intégration. Ses expériences ont permis de tirer les conclusions suivantes:

L'intégration de l'ADN exogène chez *P. patens* se fait par recombinaison homologue de façon ciblée indépendamment de locus;

L'efficacité de la mutagenèse ciblée est de 66-100% avec des fragments homologues de 2.3 à 3.6 kb.

Schaefer établit la valeur RTF<sup>1</sup> pour l'intégration de l'ADN exogène dans le génome de *P. patens* par recombinaison illégitime et par recombinaison homologue. La fréquence de l'intégration illégitime est d'environ 10<sup>-5</sup> régénérants; l'intégration est 10 fois plus fréquente par recombinaison homologue (10<sup>-4</sup>). De ce fait, une valeur RTF de 10<sup>-4</sup> pour un vecteur de transformation peut indirectement indiquer son intégration par recombinaison homologue, donc ciblée. L'intégration ciblée est également démontrée par une analyse moléculaire (PCR ou hybridation Southern) ainsi que par une analyse génétique qui démontre la liaison du vecteur ciblant (qui porte un gène marqueur) et du gène cible (Schaefer, 1994; Schaefer and Zryd, 1997).

L'efficacité de la recombinaison homologue dans les conditions testées est grandement supérieure à celle observée chez les autres plantes et comparable à celle de la levure *S. cerevisiae*. Cette caractéristique doit permettre le développement des outils de mutagenèse ciblée chez *P. patens* à l'instar de la méthodologie de la levure.

### **3. L'application la plus fine de la mutagenèse ciblée est la modification précise des séquences génomiques à un nucléotide près**

La possibilité d'isoler les événements d'intégration ciblée et précise a permis le développement de divers outils de la mutagenèse ciblée chez la levure *S. cerevisiae* (Orr-Weaver *et al.*, 1981; Rothstein, 1991) ou *S. pombe*, chez certains champignons et protistes (Beertling, 1995; Cruz, 1995; Shah-Mahoney *et al.*, 1997), chez la souris (cellules d'embryon, ES) (Bronson and Smithies, 1994; Moreadith and Radford, 1997; Morton and Hooykaas, 1995; Muller, 1999). Ces systèmes d'expérimentation reposent sur des vecteurs d'insertion et de remplacement, sur l'utilisation de la sélection positive ou négative, sur la possibilité de moduler l'expression des transgènes et sur la recombinaison spécifique de site (Morton and Hooykaas, 1995; Muller, 1999; Rothstein, 1991). Les objectifs principaux de la mutagenèse ciblée sont soit l'inactivation complète du gène ("**knock-out**") soit l'**introduction de mutations ponctuelles et précises**.

L'inactivation d'un gène peut être effectuée en une étape par insertion ou par remplacement.

La mutagenèse ponctuelle, ou de précision, du gène ciblé peut être effectuée par remplacement en une étape, lorsqu'il y a une possibilité de sélectionner les événements de remplacement; sinon, elle doit s'effectuer en plusieurs étapes par une combinaison d'événements d'insertion, de recombinaison homologue intrachromosomique et de remplacement. Cette stratégie permet théoriquement une modification précise de n'importe quelle séquence génomique à un nucléotide près.

Les travaux de Schaefer démontrent que l'inactivation d'un gène à copie unique se fait facilement en une étape par insertion. Le présent travail se fixe alors plusieurs objectifs, tous

---

<sup>1</sup> Fréquence relative de transformation (RTF) est calculée en divisant le nombre de transformants intégratifs par le nombre total de régénérants.

visant à enrichir la méthodologie de la mutagenèse ciblée chez *P. patens* en explorant les phénomènes biologiques sous-jacents. Ces objectifs se regroupent dans les chapitres complémentaires :

Dans les trois premiers chapitres, nous caractérisons la mutagenèse ciblée par recombinaison homologues chez *P. patens*:

Dans le chapitre I, nous confirmons l'efficacité, la spécificité et la fidélité de la mutagenèse ciblée du gène *ZLAB1*, membre de la famille multigénique *Cab* (codant pour la protéine majeure de liaison de la chlorophylle a/b, "chlorophyll a/b binding protein").

Dans le chapitre II, nous démontrons que les séquences interrompues par une brèche peuvent être réparées sur la matrice génomique homologue ("gap repair"), alors que les séquences interrompues par une délétion s'intègrent préférentiellement sans réparation. L'observation que les délétions sont introduites dans le génome sans être réparées nous permet de développer la mutagenèse ponctuelle.

Dans le chapitre III, nous déterminons comment l'efficacité de la mutagenèse et le mode d'intégration des vecteurs de remplacement dépendent de la forme du vecteur (linéaire ou surenroulée) et de la taille de la région d'homologie. Nous démontrons que l'intégration des vecteurs par remplacement simple nous permet désormais de développer les stratégies de la mutagenèse ciblée de précision. Nous confirmons cette conclusion en inactivant le gène *PpAPT* codant pour l'enzyme adénine phosphoribosyl transférase par remplacement ciblé et précis. Les mutants ciblés présentent un phénotype anormal qui est probablement provoqué par une perturbation du métabolisme de l'adénine. Le phénotype sauvage est restauré par une complémentation génétique par chacun des gènes homologues d'*A. thaliana AtAPT1* et 2. Cette étude démontre que la mutagenèse ciblée rend la génétique inverse réellement efficace chez *P. patens*, ce qui permet d'étudier la fonction d'une multitude de gènes végétaux sur le modèle de *P. patens*.

Dans les deux derniers chapitres, nous développons des outils additionnels permettant d'affiner les techniques de la mutagenèse ciblée:

Dans le chapitre IV, nous démontrons que la recombinase spécifique de site Cre/lox peut être utilisée pour exciser les marqueurs de sélection intégrés dans le génome des souches transgéniques.

Dans le chapitre V, nous comparons l'activité de différents promoteurs constitutifs et inductibles par les hormones glucocorticoïdes. Cette étude permet une régulation fine, et au besoin, inductible de transgènes. Le remplacement ciblé des éléments régulateurs de gènes endogènes par un de ces promoteurs exogènes analysés permettrait d'étudier leur régulation et, dans le cas des gènes essentiels, leur inactivation conditionnelle.

# **Chapitre I : Evaluation de l'efficacité, de la précision et de la spécificité de la mutagenèse ciblée par recombinaison homologue: le gène *ZLAB1* de la famille multigénique *Cab* codant pour la protéine majeure de liaison de la chlorophylle a/b ("chlorophyll a/b binding protein")**

## **Résumé**

L'analyse de l'intégration ciblée d'un vecteur d'insertion dans un gène membre d'une famille multigénique nous permet d'examiner la précision et la spécificité de la mutagenèse ciblée par recombinaison homologue. Le gène *ZLAB1* de la famille multigénique *Cab* codant pour la protéine majeure de liaison de la chlorophylle a/b est l'objet de cette étude. La spécificité est démontrée par le fait que le vecteur portant un fragment de 1 kb du gène *ZLAB1* s'intègre dans ce gène dans 25 % de tous les événements d'intégration. Nous démontrons la précision de l'intégration ciblée par l'analyse de la séquence du site d'intégration. De plus, notre analyse démontre indirectement qu'au moins une partie des autres événements d'intégration a lieu dans d'autres gènes membres de la famille *Cab*. Un autre gène *Cab* se trouve probablement dans le voisinage du gène *ZLAB1* et un événement d'intégration a provoqué la délétion de la région qui les sépare. Ceci indique que les gènes *Cab* de *P. patens* sont vraisemblablement groupés en cluster à l'instar de la famille *Cab* des angiospermes.

## **1. Introduction**

### **1.1. Spécificité de la recombinaison homologue**

La spécificité de la mutagenèse ciblée par recombinaison homologue dépend de la capacité de l'organisme à empêcher la recombinaison entre les séquences homologues non isogéniques (souvent appelées "homéologues").

Chez les systèmes tels que *E. coli* ou *S. cerevisiae*, les études de la spécificité de la recombinaison homologue s'appuient sur des analyses génétiques et structurales de l'initiation de ce processus pour démontrer que la spécificité de la recombinaison homologue est fortement dépendante de la spécificité de cette première étape. Ces études montrent qu'il existe plusieurs contraintes lors d'un événement de recombinaison non spécifique, c'est-à-dire entre des séquences non isogéniques.

La première contrainte est cinétique. Les études sur *E. coli* démontrent qu'une protéine, RecA, se polymérise sur l'ADN simple brin transformant et catalyse son interaction avec l'ADN double hélice ce qui amorce la recombinaison homologue; son homologue Rad51p a le même rôle chez la levure. Le complexe formé entre l'ADN simple brin et RecA catalyse son appariement avec la double hélice sur une longueur qui peut être de quelques kb; il effectue ensuite l'échange de brins qui mène à la formation de l'ADN hétéro-brin et d'un troisième brin. Le déplacement du troisième brin en direction 5' - 3' est appelé la migration de branches et il est accompagné de l'hydrolyse d'ATP (Stasiak, 1996). Les divergences entre les séquences en interaction créent une barrière énergétique primo à la formation des produits intermédiaires, secundo, lors de la migration des branches. Tertio, l'instabilité de l'ADN

hétéro-brin est d'autant plus grande que le nombre de nucléotides non complémentaires augmente. Ces trois contraintes cinétiques sont contrôlées par la protéine RecA (Bazemore *et al.*, 1997; Rao *et al.*, 1995). Des études effectuées *in vitro* sur des oligonucléotides démontrent que la constante d'équilibre dynamique de la réaction de la recombinaison homologue décroît en fonction du pourcentage de différence entre les séquences: elle diminue d'un facteur de 6 quand les mauvais appariements atteignent 14% (Bazemore *et al.*, 1997).

La deuxième contrainte est le contrôle de la spécificité de la recombinaison homologue par le système de réparation de mésappariements. Les travaux de Radman et collaborateurs démontrent que le système de réparation de mésappariements *mutHLS* d'*E. coli* empêche la recombinaison entre les séquences homologues divergeantes et par-là, sert de barrière génétique entre les espèces. En effet, les études des croisement par conjugaison ou par transduction entre *E. coli* et *S. typhimurium* montrent que la recombinaison entre l'ADN linéaire Hfr de *S. typhimurium* et le chromosome circulaire d'*E. coli* est plus fréquente dans les mutants du système *mutHLS*, d'un facteur de 5 à 1000 (Stambuk and Radman, 1998; Vulic *et al.*, 1997; Worth *et al.*, 1994). Datta et collaborateurs mettent en évidence que le contrôle de la spécificité de la recombinaison homologue chez *S. cerevisiae* est exercé par le système de réparation de mésappariements *msh2* (Datta, 1997).

Negritto et collaborateurs étudient deux gènes possédant une homologie de séquence nucléotidique de 83%, *SAM1* et *SAM2*, codant pour la S-adénosylmethionine synthase (EC:2.5.1.6) (Negritto *et al.*, 1997). Les chercheurs transforment la levure par plusieurs vecteurs portant des fragments d'homologie différents et sélectionnent les événements d'intégration dans le locus *SAM2*. Les fragments d'homologie contiennent soit le gène *SAM2* (isogénie), soit le gène *SAM1* (identité de 83%), soit des combinaisons chimériques des deux en dispositions différentes. L'intégration ciblée dans le locus *SAM2* est réduite d'un facteur d'environ 100 à 1000 avec le vecteur portant le gène *SAM1* par rapport au vecteur portant le gène *SAM2*; le facteur de réduction est moins important (entre 10 et 100) lorsqu'une des extrémités est isogénique au locus ciblé et le facteur de réduction est de 1 à 10 lorsque les deux extrémités sont isogéniques. L'inhibition est donc plus forte si les différences sont proches des extrémités de l'ADN transformant. La présence d'un deuxième locus homologue, *SAM1*, comme cible alternative capable de séquestrer le vecteur de transformation ne joue aucun rôle puisque l'efficacité de l'intégration ciblée dans une souche dépourvue de *SAM1* et la souche initiale est identique. Le même groupe de chercheurs étudie l'inhibition de la recombinaison homologue en utilisant les mêmes conditions dans un mutant de la réparation de mésappariements *msh2*; ils démontrent que le gène *SAM2* est ciblé avec les séquences du gène *SAM1* presque aussi fréquemment qu'avec celles de *SAM2* en absence du gène *msh2* fonctionnel. Ceci corrobore l'hypothèse que le contrôle de la fidélité de la recombinaison homologue s'effectue au niveau de la réparation de mésappariements.

L'identité entre l'ADN transformant et l'ADN cible joue un rôle critique chez la souris: la fréquence de l'intégration augmente d'un facteur de 4 à 5 lorsque l'ADN transformant est isogénique à l'ADN cible, comparé à l'ADN transformant non-isogénique provenant d'une autre souche de souris et comportant plus de 1% de différences de séquence nucléotidique (Deng and Capecchi, 1992). Une autre étude démontre une augmentation de l'efficacité de recombinaison homologue d'au moins un facteur 20 lorsque l'ADN transformant provient de la même souche et il est donc isogénique (Riele *et al.*, 1992). L'ADN non-isogénique comparé comporte des polymorphismes des sites de restriction d'au moins 5 enzymes et la différence entre les séquences estimée est en moyenne de 1 nucléotide pour 160. L'étendue d'identité parfaite semble être de prime importance selon les deux études. La mauvaise distribution de

divergences entre les deux ADN doit être la cause majeure de la dissociation ou de la non-formation du complexe de recombinaison homologue. Si, dans une étude, les séquences non homologues aux extrémités du fragment de ciblage ne semblent pas entraver la recombinaison homologue (Riele *et al.*, 1992), une autre étude montre que les dissimilarités entre les molécules ADN recombinantes inhibent la recombinaison homologue plus fortement si elles sont situées vers les extrémités des zones d'homologie de l'ADN transformant (Deng and Capecchi, 1992).

Ces études forment une base théorique pour les applications plus pratiques telles que la mutagenèse ciblée des gènes appartenant à une famille multigénique. Les systèmes expérimentaux ne possédant pas l'outil de la mutagenèse ciblée reposent souvent sur l'expression antisens du gène étudié afin d'altérer son expression native. Or, dans le cas de familles multigéniques, plusieurs membres de la famille suffisamment identiques risquent d'être inactivés. Dans certaines études, ceci peut être l'effet recherché; en revanche l'étude spécifique d'un gène sera compromise. Cet inconvénient peut être contourné dans les systèmes expérimentaux où la mutagenèse ciblée est efficace et spécifique.

L'établissement du seuil de spécificité est important pour les études d'inactivation de gènes lorsque ces gènes ne sont pas isolés. Dans ces cas, leur paralogue ou orthologue servirait à construire le vecteur de transformation. Si l'identité de séquence dépasse le seuil établi pour cet organisme, l'efficacité de la mutagenèse sera suffisante pour isoler les événements d'intégration ciblée. De plus, lorsqu'un gène hétérologue est introduit dans un organisme, la connaissance du seuil de spécificité permettrait d'estimer la probabilité que le gène hétérologue interagisse avec le gène homologue endogène (cf. chapitre III, page 70).

La spécificité de la recombinaison homologue chez *P. patens* n'est pas connue; nous ne pouvons pas estimer a priori le degré de différence entre deux séquences homologues suffisant pour inhiber la recombinaison de façon significative. Nous avons décidé d'effectuer la mutagenèse ciblée d'un gène membre d'une famille multigénique afin d'estimer directement la spécificité de la recombinaison homologue chez la mousse *P. patens*. Nous avons choisi le gène *ZLAB1*, membre d'une famille de gènes codant pour la protéine majeure de liaison de la chlorophylle a/b, comme objet de cette étude.

## 1.2. *ZLAB1* est un membre d'une famille de gènes codant pour la protéine majeure de liaison de la chlorophylle a/b ("chlorophyll a/b binding protein")

Le gène *ZLAB1* a été isolé et caractérisé par Z. Long et N. Nelson (Long *et al.*, 1989) à l'aide de criblage hétérologue d'une banque d'ADN génomique de la mousse *P. patens* avec l'ADNc du pois codant pour la principale protéine de liaison de la chlorophylle a/b (*Cab*) pFab31 comme sonde. Le gène *ZLAB1* code donc pour une protéine majeure de l'antenne collectrice du photosystème II.

Le clone génomique *ZLAB1* de 2544 bp isolé et analysé contient la région codante présumée ainsi que les régions 5' (d'environ 1.2 kb) et 3' (de 98 bp) du gène. La première analyse de l'organisation génomique des gènes *Cab* chez *P. patens* par hybridation Southern avec un fragment du clone *ZLAB1* indique l'existence d'une quinzaine de gènes homologues appartenant à la famille multigénique *Cab* (Long *et al.*, 1989). Nous avons obtenu par PCR et analysé la séquence des fragments d'autres membres de cette famille correspondant à la région entre 2172 et 2370 de la séquence publiée du *ZLAB1*. Cette analyse montre qu'il existe au

moins 13 membres de la famille *Cab* identiques à 86-94% à *ZLAB1* sur la longueur de la séquence nucléotidique analysée. La séquence des acides aminés déduites de cette région pour ces gènes est identique à 97-100% à celle de *ZLAB1* (Hoffmann *et al.*, 1999). Le gène *ZLAB1* est donc un bon candidat pour l'estimation de la spécificité de la mutagenèse ciblée par recombinaison homologue chez *P. patens*:

Les buts suivants ont été fixés dans cette étude:

- Vérifier l'efficacité de la mutagenèse ciblée dans *P. patens*. L'efficacité de la mutagenèse est la proportion (en pourcentage) des événements d'intégration ciblée parmi tous les événements d'intégration. Schaefer (Schaefer, 1994; Schaefer and Zryd, 1997) rapporte une efficacité supérieure à 90% pour 3 locus indéterminés de *P. patens*;
- Tester la spécificité de la mutagenèse ciblée de *P. patens* par inactivation spécifique d'un membre de la famille multigénique *Cab* qui n'est probablement pas létale au vu de la redondance de cette famille multigénique.
- Vérifier le modèle d'intégration d'un vecteur d'insertion en forme circulaire (surenroulé) ou linéaire (coupé dans la région d'homologie et présentant donc des ruptures d'ADN double brin). Prouver que la structure du site d'intégration peut être prédite précisément.

## 2. Résultats

Le présent travail a été effectué en collaboration avec le laboratoire du Professeur E. Russo (MPI, Berlin) et a fait objet d'une publication (Hoffmann *et al.*, 1999).

### 2.1. Stratégie de la mutagenèse ciblée

Un vecteur d'insertion pGLcab-a a été construit par D. Schaefer et A. Hoffmann (Hoffmann *et al.*, 1999). Le fragment d'homologie choisi est la région codante du gène *ZLAB1* tronquée des séquences codant pour 11 aa en 5' et des séquences codant pour 35 aa en 3'. L'intégration du vecteur par un événement de recombinaison dans la région homologue doit inactiver le gène *ZLAB1* puisqu'aucune copie du gène présente dans le site d'insertion ne sera fonctionnelle. Afin de sélectionner les transformants intégratifs, un marqueur de sélection 35S *hpt* conférant la résistance à l'hygromycine est placé sur le vecteur de transformation. L'inactivation du gène *per se* n'est pas l'intérêt principal de cette étude, d'autant plus qu'elle ne permettra probablement pas la sélection phénotypique des événements d'intégration ciblée. L'analyse moléculaire seule distingue les transformants intégratifs ciblés parmi tous les transformants. Nous évaluons la fidélité de la mutagenèse ciblée par l'analyse de la séquence du site d'intégration.

### 2.2. Isolation des transformants intégratifs

J'ai produit les transformants intégratifs à l'aide de quatre expériences de transformation. Les fréquences RTF sont présentées dans la Figure 1.

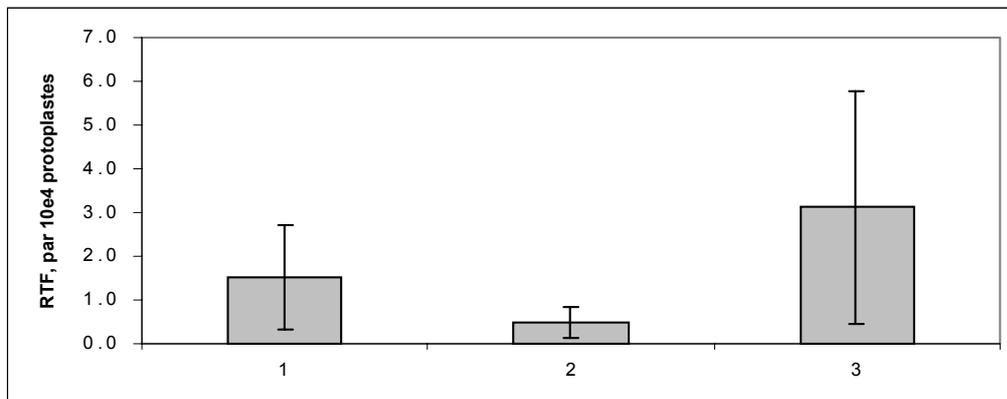


Figure 1. Fréquence relative de transformation (RTF) dans les expériences de la mutagenèse ciblée du gène *ZLAB1*

La valeur RTF moyenne est calculée au moyen d'un rapport du nombre de transformants intégratifs au nombre de protoplastes régénérés pour une expérience. Nous avons régénéré environ  $4 \cdot 10^5$  protoplastes par vecteur (total de 4 expériences, la barre d'erreur indique l'écart-type). Nous n'obtenons aucun transformant dans la transformation sans ADN, ni dans la transformation par le vecteur pGL2 qui ne porte pas de fragments homologues, en effet, la valeur RTF attendue est de l'ordre de  $10^{-5}$ ; cf. (Schaefer and Zryd, 1997);

1 - Le vecteur pGL 108 porte un fragment homologue de 2.3 kb (Schaefer and Zryd, 1997), une valeur RTF de l'ordre de  $2.35 \pm 1.44 \cdot 10^{-4}$  est attendue;

2 - Le vecteur pGL Caba sc est le vecteur d'intégration ciblée dans le gène *ZLAB1*, circulaire surenroulé;

3 - Le vecteur pGL Caba /BglIII est le vecteur pGLCaba coupé par BglIII dans la région d'homologie ( Figure 2, page 20).

Deux conclusions peuvent être tirées de la Figure 1:

- une augmentation de la RTF pour les deux formes du vecteur pGLcab-a (surenroulée et linéaire) par rapport au vecteur pGL2 suggère que ce vecteur-là s'intègre par recombinaison homologue. Cette augmentation est comparable mais inférieure à l'augmentation de la RTF entre les vecteurs pGL108 et pGL2;
- la forme linéaire du vecteur pGLcab-a transforme la mousse de façon plus efficace que la forme circulaire d'un facteur allant jusqu'à 25 (selon les expériences). Ceci est en accord avec l'observation dans tous les autres systèmes d'une augmentation de l'intégration des vecteurs portant une rupture double-brin.

### 2.3. Analyse moléculaire d'un échantillon des transformants intégratifs

Notre analyse doit répondre aux questions suivantes:

- l'intégration a-t-elle lieu dans le gène *ZLAB1* ?
- l'intégration a-t-elle lieu par insertion d'une ou de plusieurs copies du vecteur ?
- quelle est la structure du site d'intégration ?
- l'intégration est-elle précise ?

J'ai analysé par l'hybridation Southern l'ADN de 22 clones intégratifs choisis au hasard dont 12 obtenus avec le vecteur surenroulé et 10 avec le vecteur linéaire. L'analyse par PCR a été effectuée au laboratoire du Prof. E. Russo sur le même ADN, que je leur ai fourni.

Nous ne pouvons utiliser aucun fragment de la région codante du gène *ZLAB1* comme sonde pour l'analyse Southern puisque l'identité importante avec d'autres membres homologues de la famille *Cab* mène à l'apparition d'un grand nombre de signaux d'hybridation rendant l'analyse impossible.

Nous avons isolé un fragment de la séquence non traduite de la région 5'. Le fragment (appelé **cab5spec**) couvre 888 bp entre les sites de coupure XbaI et BstXI, positions -1112 et -225, respectivement, par rapport au site d'initiation de traduction ATG. Nous avons testé la spécificité de la sonde au gène *ZLAB1* par hybridation Southern avec l'ADN de *P. patens* de type sauvage (WT). En effet, un fragment de 2.5kb reconnu dans l'ADN WT digéré par EcoRI correspond au fragment génomique portant le gène *ZLAB1* cloné et caractérisé par le laboratoire de N. Nelson. L'intensité du signal témoigne de la spécificité de cette sonde :

- l'ADN digéré par EcoRI : 1 signal majeur + 2 signaux faibles;
- l'ADN digéré par XbaI : 1 signal majeur + 2 signaux faibles.

Les signaux faibles attestent d'une hybridation avec deux autres membres de la famille multigénique *Cab* dont les séquences 5' sont partiellement identiques à celles de *ZLAB1*.

Nous avons prédit le RFLP (polymorphisme de longueur des fragments de restriction) du gène *ZLAB1* intact ou inactivé par l'intégration ciblée du vecteur pGLcab-a pour EcoRI et XbaI (cf. Figure 2).

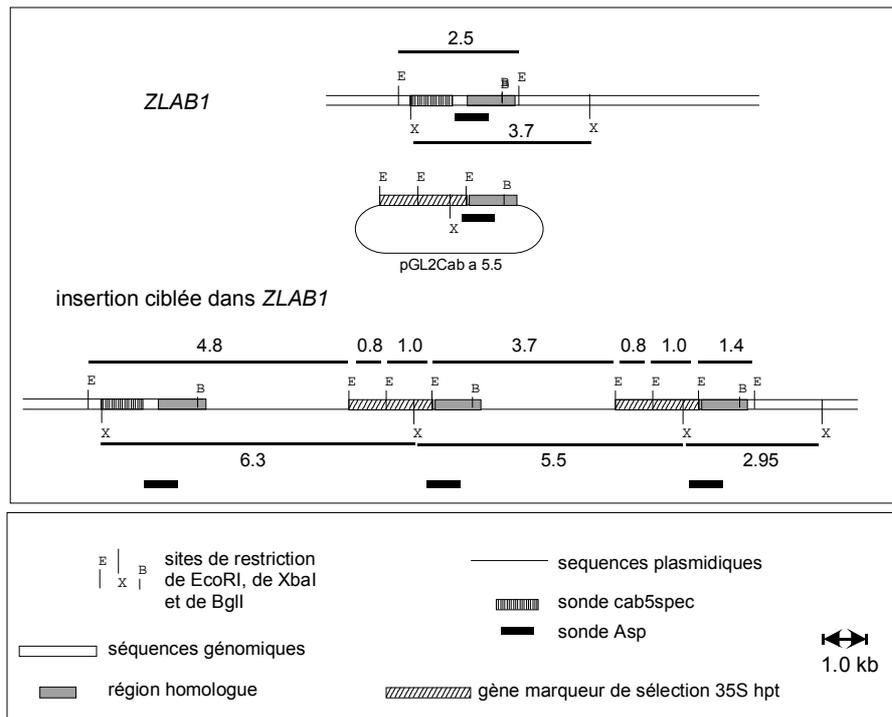


Figure 2. Insertion ciblée du vecteur pGLCab-a dans le gène *ZLAB1*.

Nous vérifions l'intégration ciblée dans le gène *ZLAB1* par le remplacement du fragment EcoRI de 2.5 kb par un fragment hybride de 4.8 kb, ainsi que par le remplacement du fragment XbaI de 3.7 kb par un fragment hybride de 6.3 kb. Ces fragments sont spécifiquement reconnus par la sonde cab5spec. Les fragments hybrides 3' EcoRI de 1.4 kb et XbaI de 2.95 kb, spécifiques de l'intégration ciblée, sont reconnus par la sonde AspI. Nous nous attendons à des signaux de faible intensité (EcoRI de 3.3 et 7.2 kb, XbaI de 3.6 et 4.4 kb) qui proviennent d'une hybridation avec deux autres gènes membres de la famille *Cab* (non montrés dans le schéma).

Nous vérifions l'orientation et le nombre des copies intégrées du vecteur par les fragments reconnus par la sonde pGL2 (séquences plasmidiques plus le marqueur de sélection). La détection de fragments EcoRI de 3.7, 1.0 et 0.8 kb et d'un fragment XbaI de 5.5 kb indique que l'intégration a lieu en répétition directes. La présence des fragments d'une autre taille indiquerait une intégration de copies partiellement digérées ou reliées en répétitions inversées. Si les copies du vecteur linéarisé avec BglII sont ligaturées en répétitions inversées, un fragment EcoRI de 2.7 kb et un fragment XbaI de 8.3 kb sont prédits (non montrés dans le schéma).

Notre analyse se fonde sur le RFLP des enzymes de restriction EcoRI ou XbaI de 12 clones issus d'une transformation par le vecteur surenroulé (série CABsc) et de 10 clones issus d'une transformation par le vecteur linéaire (série CABlin). Comme prédit, la sonde cab5spec s'hybride fortement avec un fragment de 2.5kb ou de 3.7kb dans l'ADN de WT digéré par EcoRI ou XbaI, respectivement. Il en est de même pour 9 clones de la série CABsc et pour 8 clones de la série CABlin. Nous en concluons que le gène *ZLAB1* n'est pas ciblé dans ces clones. En revanche, *ZLAB1* est ciblé selon nos prédictions (Figure 2) dans 3 clones de la série CABsc (CAB GT1-3) et 2 clones de la série CABlin (CABlin GT1 et 2). Le signal correspondant aux fragments WT du gène *ZLAB1* de 2.5 kb (EcoRI) ou 3.7kb (XbaI, si analysé) est absent; il est remplacé par un signal d'un fragment prédit de 4.8 kb (EcoRI) ou de 6.3 (XbaI). Les signaux correspondant à des membres proches de la famille *Cab* restent intacts dans les deux digestions analysées ce qui indique que ces gènes ne sont pas ciblés par l'intégration du vecteur. La Figure 3 montre un exemple de l'hybridation de la sonde cab5spec avec le gène *ZLAB1* intact de la souche WT et de 2 autres souches intégratives (CabNT1 et 2) et avec le gène *ZLAB1* interrompu par l'insertion ciblée du vecteur dans 4 souches.

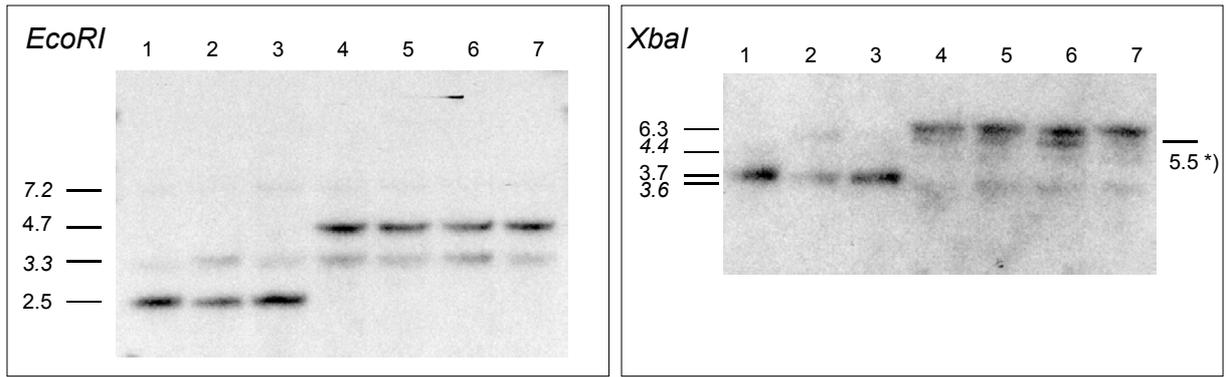


Figure 3. Mise en évidence de l'intégration ciblée dans le gène *ZLAB1* par hybridation avec la sonde *cab5spec*.

*EcoRI* - ADN de différentes souches digéré avec *EcoRI* ;

*XbaI* - ADN de différentes souches digéré avec *XbaI* ;

Les souches sont: 1 - WT, 2, 3 - intégration non ciblée; 4, 5, 6 - intégration ciblée du vecteur circulaire (CABGT1-3); 7 - intégration ciblée du vecteur linéaire (CABGTlin1).

\*) Le signal de 5.5 kb dans la digestion *XbaI* correspond au fragment plasmidique du vecteur qui s'hybride avec le plasmide contaminant la sonde, cf. l'hybridation avec la sonde pGL2.

Nous avons utilisé une sonde qui couvre une partie de la séquence non traduite en 5' plus une partie de la région codante pour l'extrémité N-terminale et une partie de l'intron (-124 à +560 par rapport au site d'initiation de la traduction ATG) présentée dans la Figure 2 (sonde **Asp**). Cette sonde est moins spécifique au gène *ZLAB1* parce qu'elle contient une partie de la séquence codante, similaire pour plusieurs gènes *Cab*. Par conséquent, nous détectons deux fragments supplémentaires (*EcoRI* de 2.3 et 9 kb et *XbaI* de 5 et 9 kb) correspondants à deux autres gènes *Cab* (Figure 4, digestions *EcoRI* et *XbaI*). De plus, nous prédisons la présence d'un fragment *XbaI* de 2.95 kb dans l'ADN des souches dans lesquelles *ZLAB1* a été ciblée. La Figure 4 (digestion *XbaI*) confirme la présence du fragment hybride *XbaI* de 2.95 kb formé en 3' de la position d'intégration du vecteur spécifique aux clones ciblés.

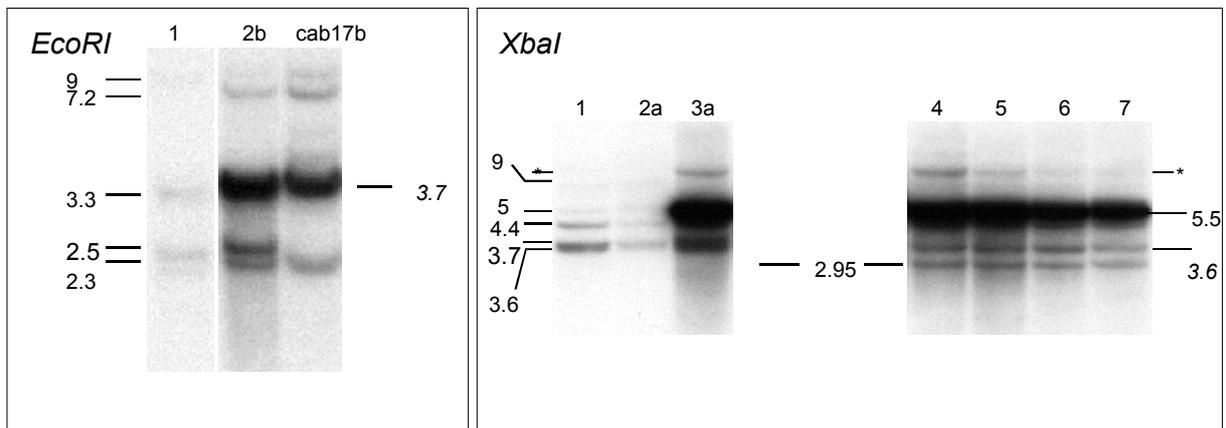


Figure 4. Mise en évidence des autres gènes *Cab* et du fragment hybride *XbaI* de 2.95 kb spécifique de l'intégration ciblée.

***EcoRI*** - Les souches sont: 1 - WT, 2b - intégration non ciblée (différente des autres figures).

Nous ne détectons pas le gène *ZLAB1* dans la souche *cab17b*. Cette souche est traitée plus loin, page 23.

***XbaI*** - Les souches sont: 1 - WT, 2a, 3a - intégration non ciblée (différentes des autres figures); 4, 5, 6 - intégration ciblée du vecteur circulaire (CABGT1-3); 7 - intégration ciblée du vecteur linéaire (CABGTlin1). Les copies du vecteur en répétitions directes sont détectées à 5.5 kb.

L'analyse PCR des 4 clones ici présentés confirme l'inactivation ciblée du gène *ZLAB1* mise en évidence par l'analyse Southern. De plus, la séquence partielle des fragments hybrides en 5'

et 3' amplifiés par PCR a été analysée (Figure 5); cette analyse a confirmé la précision de l'intégration du vecteur. Dans l'ensemble, ces résultats mettent en évidence l'inactivation ciblée et précise du gène *ZLAB1* conforme au modèle dans 3 sur 12 transformants obtenus avec le vecteur surenroulé et dans 2 sur 10 transformants obtenus avec le vecteur linéaire coupé dans la région d'homologie.

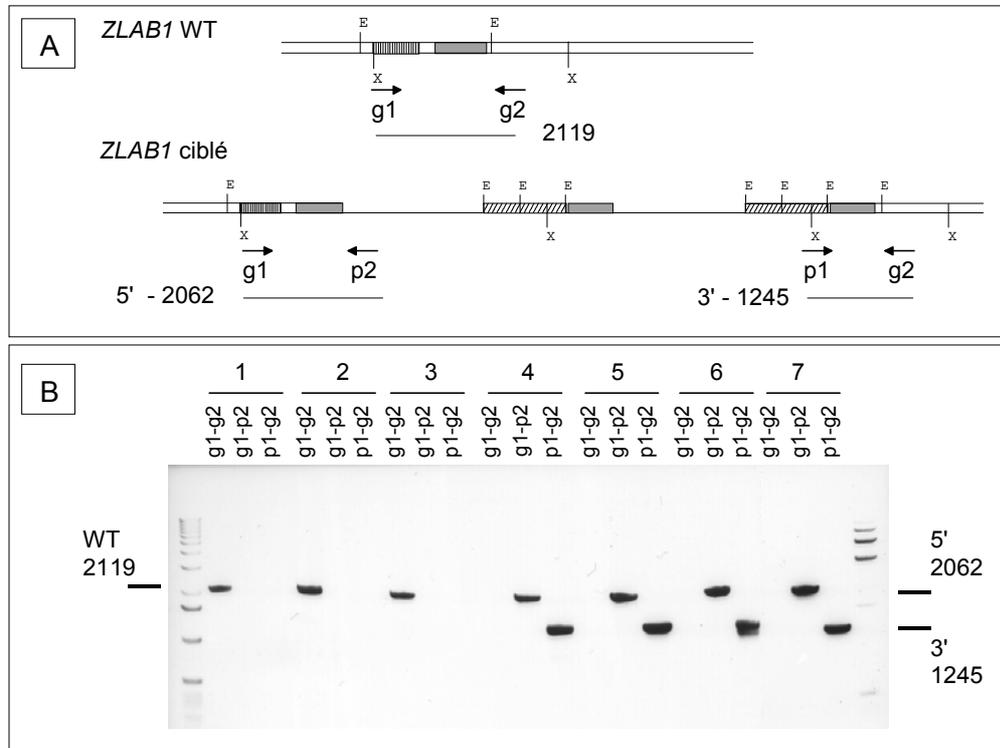


Figure 5. Mise en évidence de l'intégration ciblée par PCR.

A - La taille prédite des fragments amplifiés avec les amorces g1, g2 (spécifiques du gène *ZLAB1*), p1 et p2 (spécifiques du vecteur) dans la souche sauvage et les souches ciblées ou non (en bp). La séquence des amorces et l'analyse de séquence détaillée sont publiées (Hoffmann *et al.*, 1999).

B - La confirmation de la taille des fragments prédits par analyse PCR. Les souches sont : 1 - type sauvage, 2, 3 - souches non ciblées, 4 - 7 les souches ciblées présentés dans les figures précédentes.

Selon le modèle d'intégration proposé, l'intégration multiple du vecteur circulaire ne peut produire que des répétitions directes; en revanche, le vecteur linéaire pourrait former des répétitions inversées. Ceci est possible si plusieurs copies liées de façon aléatoire par ligation des bouts libres produisent des structures concaténées de haut poids moléculaire qui s'intègrent ensuite dans le génome. Si une copie du vecteur linéaire est fermée par ligature, et participe ensuite à la concaténation dépendante de la réplication ou de la recombinaison homologue extrachromosomique, seules les répétitions directes peuvent être produites.

J'ai déterminé le nombre et l'orientation des copies du vecteur intégrées par une analyse Southern utilisant pGL2 comme sonde. Nous prédisons, dans la Figure 2, page 20, leur détection par RFLP XbaI et EcoRI. La sonde pGL2 ne s'hybride qu'avec les séquences du vecteur intégré ainsi qu'avec les séquences hybrides entre le site d'intégration et le vecteur. Le nombre de copies est estimé d'après l'intensité du signal d'hybridation. La présence de plusieurs copies mène à l'apparition de signaux de forte intensité qui risquent de masquer les signaux d'hybridation avec les fragments hybrides et les autres gènes *Cab*. Dans cette analyse nous nous concentrons donc sur les signaux d'hybridation avec les répétitions du vecteur.

Le nombre de copies intégrées est de 2 à 30 (Figure 6). L'analyse des signaux d'hybridation des souches obtenues avec le vecteur linéaire ne révèle que l'intégration de copies directes puisque seuls les fragments du monomère du vecteur sont présents dans les digestions EcoRI et XbaI (Figure 6). L'analyse de l'intégration, ciblée ou non, du vecteur indique que seuls les fragments du monomère sont détectés; l'insertion de copies en répétitions directes a donc lieu.

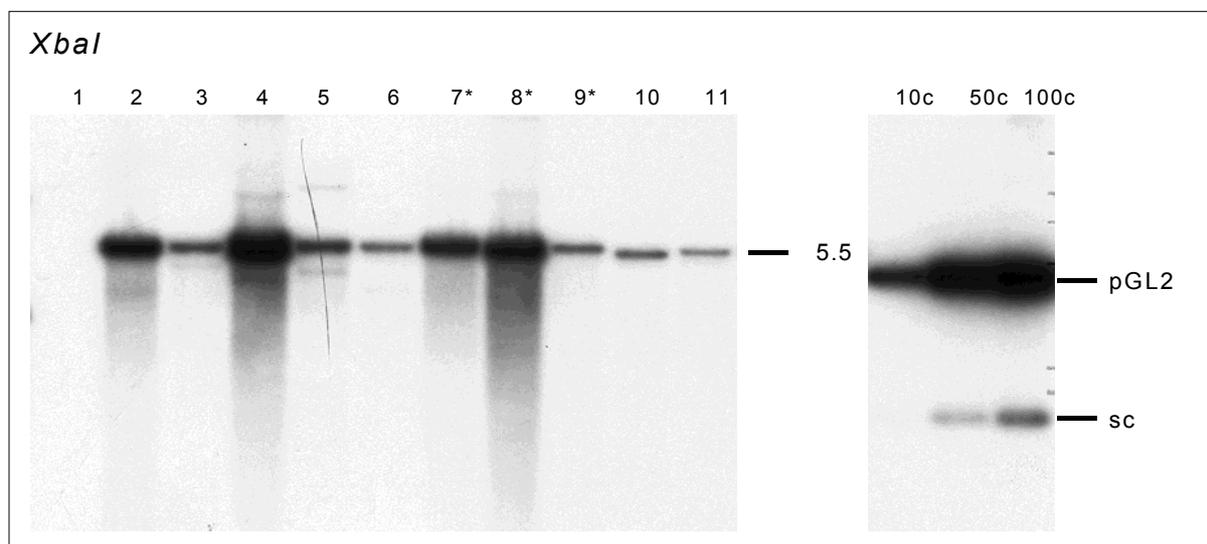


Figure 6. Hybridation de l'ADN de différentes souches digéré avec XbaI avec pGL2.

Les souches sont: 1 - WT, 2 - 6, 10, 11 - intégration non ciblée (CABNT); 7, 8 - intégration ciblée du vecteur circulaire (CABGT1-2); 9 - intégration ciblée du vecteur linéaire (CABGTlin1). 10c, 50c et 100c - le plasmide pGL2 chargé en quantités correspondant à 10, 50 et 100 copies par génome. Les souches à intégration ciblée sont marquées d'un astérisque. Les signaux de faible intensité sont les jonctions hybrides avec le site d'intégration. L'estimation du nombre se fait par comparaison de l'intensité du signal comparée à celle du signal du plasmide pGL2.

#### Délétion d'un fragment chromosomique

Nous avons obtenu un clone (*cab-17b<sup>#</sup>*) dans lequel les séquences 5' du gène *ZLAB1* sont éliminées par l'intégration du vecteur comme nous le proposons dans la Figure 8. Nos raisonnements sont les suivants:

- ✓ L'intégration du vecteur est confirmée par une analyse d'hybridation Southern Figure 4, page 21; les signaux du fragment EcoRI de 3.7 kb correspondent à une insertion multiple du vecteur.
- ✓ La même figure démontre que le fragment EcoRI de 2.5 kb du gène *ZLAB1* est absent: la sonde Asp ne le détecte plus.
- ✓ La Figure 7 confirme que la sonde spécifique des séquences 5' du gène *ZLAB1* ne détecte ni le gène *ZLAB1* intact ni les fragments attendus après une intégration ciblée.

<sup>#</sup> Transformations préliminaires avec le vecteur pGLcabB. Ce vecteur est différent du vecteur pGLcabA uniquement par l'orientation du fragment homologue. La Figure 8 présente le schéma du vecteur pGLcabB.

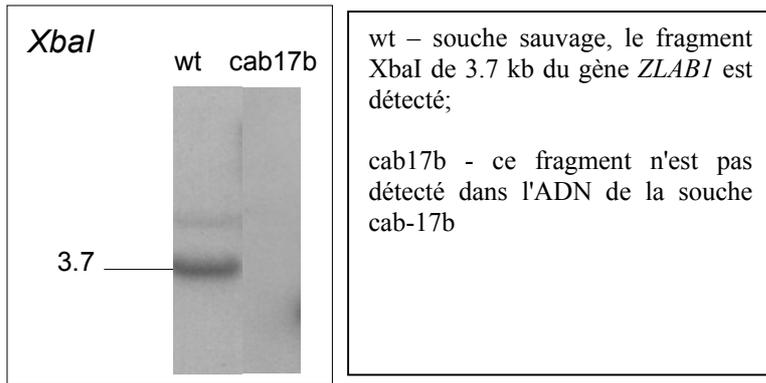


Figure 7. Démonstration de la disparition des séquences 5' du gène *ZLAB1* par hybridation de la digestion XbaI avec la sonde cab5spec.

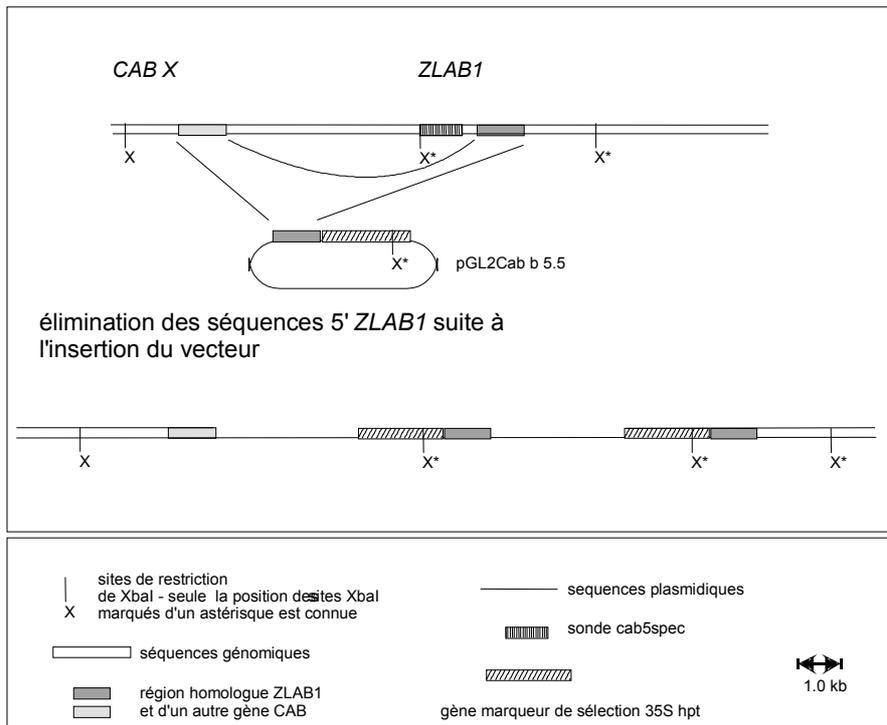


Figure 8. Explication de l'allèle cab-17b.

Les lignes en pointillé indiquent arbitrairement les régions qui ont subi un événement de recombinaison homologue entre *ZLAB1* et un autre gène de la famille *Cab* (*CAB X*). Le vecteur pGL2cabb est identique à pGLcaba mais le fragment homologue est inversé. Nous n'avons pas vérifié l'intégration du vecteur dans le gène recombiné. Notre analyse n'exclut pas la possibilité que le vecteur se soit intégré ailleurs dans le génome (par exemple, en amont en cis) et que l'événement de recombinaison intrachromosomique ait eu lieu ensuite, mais un tel événement est moins probable.

Chez plusieurs angiospermes, la famille multigénique *Cab* est produite par des duplications et par des inversions d'un ancêtre et par conséquent, les gènes sont groupés en clusters (Castresana *et al.*, 1987). Nous supposons que chez *P.patens*, les gènes *Cab* le sont aussi et nous postulons qu'un gène *CAB X* hypothétique se trouve en amont du gène *ZLAB1*. L'intégration ciblée a donc produit la délétion d'une partie d'un cluster *Cab* dans la souche cab-17b.

### 3. Discussion

Nous démontrons ici que la mutagenèse ciblée chez *P. patens* est ciblée et précise. En effet, le gène *ZLAB1*, membre de la famille *Cab*, est ciblée par le vecteur d'insertion circulaire dans 25% des cas (3 souches à l'intégration ciblée sur 12 analysées) et par le vecteur linéarisé dans 20% des cas (2 souches à l'intégration ciblée sur 10 analysées). La précision de l'intégration est confirmée par l'analyse de la séquence du site d'intégration. L'efficacité de ciblage du gène *ZLAB1* est plus faible que celle rapportée précédemment pour trois locus endogènes de *P. patens* (66-100%) (Schaefer and Zryd, 1997). Ceci peut être expliquée par plusieurs hypothèses:

1. Une efficacité moindre est spécifique au fragment homologue étudié. Ce phénomène est connu chez la levure (Rothstein, 1991) et il n'est pas exclu a priori pour la mousse *P. patens*. L'efficacité de la mutagenèse ciblée est liée à l'efficacité de la recombinaison homologue; celle-ci peut dépendre de l'intensité de transcription des gènes présents dans la région analysée. Une corrélation entre le taux de transcription et l'efficacité de la mutagenèse ciblée est supposée pour les locus  $\lambda 213$ ,  $\lambda 420$  et  $\lambda 108$  (Schaefer, données non publiées) quoiqu'aucune relation directe n'ait été établie à ce jour dans les organismes modèles. Le niveaux d'expression des gènes *Cab* est très variable chez les plantes et le taux de transcription du gène *ZLAB1* n'est pas connue;
2. Une meilleure efficacité de la mutagenèse ciblée par recombinaison homologue est associée à une meilleure accessibilité de la chromatine de différents régions. Le gène *ZLAB1* se situe peut-être dans une zone peu accessible;
3. Le vecteur peut être séquestré par les sites potentiels d'intégration homologue que présentent les autres membres de la famille multigénique. Notre étude ne les met pas en évidence en raison de la non disponibilité des sondes spécifiques à ces gènes (cf. le paragraphe ci-dessous). Cette hypothèse ne trouve tout de même pas de confirmation dans l'étude sur *S. cerevisiae* précitée (Negritto *et al.*, 1997): l'efficacité de la mutagenèse ciblée du gène *SAM2* avec divers vecteurs n'est pas corrélée avec la présence ou l'absence d'un deuxième gène homologue, *SAMI*. Une étude complémentaire pour vérifier cette hypothèse dans notre cas devrait se faire dans une souche dépourvue des autres gènes *Cab*; elle nous paraît difficile à réaliser.

Nous n'avons pas directement prouvé une intégration dans un autre gène membre de la famille *Cab* dans les souches non ciblées. Néanmoins, l'analyse initiale par PCR avec des amorces spécifiques au gène *ZLAB1* (position 5' à -77 de l'ATG) a indiqué que le gène *ZLAB1* est interrompu dans plus de souches que l'analyse Southern ne le démontre (Hoffman, A., données non publiées). Nous pensons que cette mauvaise interprétation est due à la faible spécificité des amorces et que les fragments observés sont amplifiés sur la matrice d'un autre gène *Cab* dans lequel le vecteur s'est intégré. En effet, l'analyse PCR utilisant une amorce 5' qui s'apparie à un fragment plus spécifique à *ZLAB1* (position 5' à -918 bp de l'ATG, région cab5spec) a été entièrement confirmée par l'analyse Southern. Nous pensons donc que dans une partie des souches non ciblées, l'intégration a lieu dans un gène de la même famille *Cab*. Dans le chapitre II, nous démontrerons en effet deux cas d'intégration des séquences homologues de *ZLAB1* dans un de ces gènes. L'identité entre ces gènes est très importante et l'intégration par recombinaison homologue est probablement favorisée dans des régions d'identité parfaite d'une longueur suffisante à l'intégration ciblée. Nous démontrons dans le

chapitre III qu'une longueur minimale requise pour une intégration ciblée est inférieure à 0.5 kb.

La présente étude mesure la spécificité en termes d'efficacité de la mutagenèse ciblée en présence de plusieurs cibles possibles que sont les autres membres de la famille multigénique *Cab*. Nous n'avons pas établi le seuil d'identité au-dessous duquel la recombinaison homologue n'est plus détectable. Pour ce faire, il faudrait estimer la mutagenèse ciblée du gène *ZLAB1* avec différents vecteurs portant des régions d'homologie avec une gamme croissante de différence de séquence. La diminution attendue de l'efficacité de cibler le gène *ZLAB1* par des régions homologues de plus en plus divergentes permettrait de trouver le seuil de la spécificité de la mutagenèse ciblée chez *P. patens*.

**La présente étude démontre qu'il est possible d'inactiver, de façon efficace et précise, un membre spécifique d'une famille multigénique.**

L'intégration du vecteur linéaire de transformation est plus fréquente (mesurée selon la valeur RTF) mais n'a pas produit plus d'événements de l'intégration ciblée que le vecteur circulaire. En effet, les ruptures double brin induisent toutes les voies d'intégration de l'ADN exogène mais ne modifient pas forcément leur fréquence relative ; nous en reparlerons dans la discussion générale (page 124).

Nous avons observé une intégration du vecteur dans deux séquences homologues: le gène *ZLAB1* en aval et un autre gène *Cab* en amont; cette intégration a provoqué la délétion du fragment génomique entre ces séquences. Nous pensons que même dans ce cas complexe et non prévu, l'intégration a lieu par recombinaison homologue précise. Une analyse de ségrégation génétique devrait nous permettre de prouver que l'intégration est ciblée.

Nous n'observons aucune altération de phénotype des souches avec le gène *ZLAB1* inactivé. Il n'est pas exclu pour autant que son inactivation produise des altérations que seule une analyse très fine pourrait déceler. Ce n'est pas surprenant compte tenu de la redondance de la famille *Cab* dont les fonctions peuvent se recouvrir. De plus, le gène *ZLAB1* n'est peut être pas le gène majoritairement transcrit pour la production des antennes collectrices du photosystème II de la mousse *P. patens*. Ces investigations intéressantes sont toutefois en dehors du cadre de la présente étude.

## Chapitre II : Mise en évidence de la réparation des séquences interrompues ("gap repair") chez *P. patens*

### Résumé

Nous avons étudié l'intégration des séquences homologues interrompues par une brèche ou par une délétion. Les brèches sont fréquemment réparées conformément à la matrice homologe génomique, mais en présence d'une matrice extrachromosomique, elles sont probablement réparées encore plus efficacement. La réparation est plus efficace lorsque l'intégration est ciblée mais elle peut être associée à une intégration illégitime ("hit and run"), quoique moins fréquemment. La réparation fidèle des brèches permet d'envisager le développement de la technique de récupération d'allèles ("allele rescue"). En revanche, les délétions ne sont réparées que dans une minorité des cas. Le fait que les séquences interrompues par une délétion ne sont pas systématiquement réparées lors de l'intégration permet de les introduire dans le génome, ce qui rend la mutagenèse ponctuelle possible.

### 1. Introduction

#### 1.1. Réparation des séquences interrompues

La récupération d'allèles ("allele rescue", (Rothstein, 1991)) est une méthode basée sur la réparation des séquences interrompues chez *S. cerevisiae* qui permet d'isoler une séquence génomique recherchée. Un vecteur navette (maintenu comme épisome chez *S. cerevisiae* et *E. coli*) porte la séquence d'intérêt interrompue par une brèche. Lors de la réparation de cette brèche à partir de la matrice génomique homologe, la séquence réparée est copiée sur le vecteur; ce vecteur peut être ensuite récupéré dans *E. coli*. La mise au point de la récupération d'allèles nécessite donc un procédé de récupération des éléments extrachromosomiques à partir de l'ADN total et une réparation efficace des séquences interrompues. Chez *P. patens*, les plasmides bactériens peuvent être répliqués comme éléments épisomiques (Ashton *et al.*, 2000; Schaefer, 1994), et la récupération d'un vecteur navette devrait être possible mais n'a pas encore été démontrée.

Le mécanisme de la réparation des séquences interrompues peut s'avérer un obstacle à la mutagenèse qui vise à introduire des délétions dans le génome. En effet, si l'intégration des vecteurs interrompus par une délétion est dépendante de la synthèse à partir de la matrice génomique homologe, le vecteur sera corrigé avant l'intégration.

Dans cette étude, nous explorons la réparation des séquences interrompues par une brèche ou une délétion chez *P. patens*. Nous allons revoir les concepts théoriques et les données expérimentales associés à la réparation des séquences interrompues.

Orr-Weaver et collaborateurs démontrent que l'intégration des vecteurs interrompus par une brèche dans 3 régions indépendantes (*HIS3*, *sup3+* et *rDNA*) est accompagnée de la réparation des brèches (de 300 à 1200 bp) dans 10 cas sur 10 chez *S. cerevisiae* (Orr-Weaver *et al.*, 1981). La structure finale après l'intégration est identique à celle qui serait obtenue avec le vecteur circulaire entier. Lorsque le vecteur possède des séquences de répllication autonome (ARS), il est réparé avec la même efficacité, la copie chromosomique servant de matrice. La réparation peut être accompagnée ou non d'un événement d'intégration. Les deux événements se produisent avec la même fréquence (Nickoloff *et al.*, 1989; Orr-Weaver and Szostack,

1983). En effet, la réparation des séquences interrompues chez la levure suit probablement le même mécanisme que la réparation des ruptures double brin décrit par le modèle DSBR de Szostak (Paques and Haber, 1999; Szostak *et al.*, 1983). Ce modèle prévoit que les jonctions Holliday peuvent être résolues de deux façons: avec ou sans crossing-over. L'intégration du vecteur réparé a lieu par un crossing-over. En revanche, la réparation de la brèche sans crossing-over laisse le vecteur comme un élément extrachromosomique. Le modèle DSBR de Szostak explique également le fait qu'une réparation complète et fidèle de la brèche est parfois accompagnée de modification des séquences voisines par conversion de gènes. D'après ce modèle, soit une extrémité de la brèche est dégradée et ensuite réparée conformément à l'information de la région homologue chromosomique, soit, lors de la formation de l'ADN hétérobrin, le brin en provenance du vecteur est réparé selon l'information chromosomique.

La réparation des séquences interrompues non intégrées peut être incomplète ou même absente. Lorsqu'une extrémité de la brèche est dégradée au-delà de la région d'homologie, la réparation se fait par ligature indépendante de l'homologie (NHEJ). La possibilité d'événements de ligature indépendants de l'homologie est confirmée par la mise en évidence d'une religature du vecteur sans aucune réparation (Orr-Weaver and Szostack, 1983). L'analyse de ces vecteurs nouvellement joints dans différents mutants de la levure a permis une compréhension plus profonde des déterminants génétiques de ce processus. En absence de réparation de ruptures double brin, les bouts collants sont joints de façon efficace alors que les bouts francs ne le sont pratiquement pas (Bartsch *et al.*, 2000). La ligature de bouts francs est régie par un mécanisme indépendant de la recombinaison homologue (Glasunov and Glaser, 1999).

Bartsch et ses collègues révisent à la baisse la fréquence des événements de réparation de brèches suivie d'une intégration (Bartsch *et al.*, 2000). Cette étude indique que le modèle de boucle en D migratrice décrit par Ferguson (Ferguson and Hollman, 1996) s'applique également à la réparation de brèche chez *S. cerevisiae*. Le rôle des deux extrémités de la brèche peut être, en effet, différent.

Les brèches sont aussi réparées de façon efficace lorsque le vecteur est introduit dans la levure *S. cerevisiae* à l'aide de *A. tumefaciens* et que la brèche est flanquée des séquences bordures gauche et droite de l'ADN-T (Risseuw *et al.*, 1996). La réparation enlève les séquences hétérologues que sont les séquences bordures et, de plus, peut s'étendre sur les régions adjacentes à la brèche. Les bordures peuvent être également ligaturées sans réparation de brèche avant l'intégration. Une fois la brèche reliée par simple ligature, elle n'est plus réparée. Le vecteur intégré porte donc une délétion. L'étude démontre que le type et la fréquence des événements de recombinaison dépendent de l'orientation du fragment homologue par rapport aux séquences bordures. La spécificité de l'intégration à l'aide de *A. tumefaciens* est probablement associée à sa protéine VirD2. Celle-ci est liée au bras droit de l'ADN-T; par-là, VirD2 peut le protéger de la dégradation. Cette protéine peut également participer activement à l'intégration de l'ADN-T.

Les études sur d'autres organismes révèlent que tous les modèles basés sur la levure *S. cerevisiae* ne décrivent pas la réparation des brèches. Chez *U. maydis* (Ferguson and Hollman, 1996), la réparation du vecteur n'est pas accompagnée d'une intégration dans la majorité de cas. De plus, les deux extrémités de la brèche ne jouent pas le même rôle dans le processus. Ferguson propose un modèle de boucle en D migratrice similaire aux modèles SDSA (appariement dépendant de la synthèse) et OSI (invasion d'une extrémité).

La réparation de brèche a également été démontrée dans les cellules ES de la souris. Les deux extrémités formant la brèche semblent jouer un rôle différent dans la réparation (Villemure *et*

*al.*, 1997). Le modèle proposé est donc similaire au modèle de boucle en D migratrice évoqué ci-dessus. Une seule extrémité peut initier la synthèse du fragment manquant selon le modèle OSI qui implique l'intégration d'une des extrémités par NHEJ.

La réparation de brèches chez *D. melanogaster* a été largement étudiée et révisée (Lankenau and Gloor, 1998). Elle suit l'excision d'un élément transposable P qui laisse une rupture double brin dans chacune des répétitions inversées terminales de 31 bp. Les deux extrémités double brin ainsi formées sont souvent sujettes à l'action d'exonucléases qui élargit la brèche. Les extrémités de la brèche peuvent "scanner" tout le génome en quête d'homologie. Lorsque celle-ci est trouvée, la brèche est réparée selon la matrice de la région homologe. Ce processus n'est pas accompagné de crossing-over et il n'y a donc pas d'échange de séquences flanquantes. De ce fait, le modèle SDSA décrit ce phénomène bien mieux que le modèle DSBR de Szostak qui prévoirait un événement de crossing-over dans une partie des cas.

Les études de la réparation de brèches et des mutations du vecteur conformément à la matrice chromosomique chez CHO (cellules ovariennes du hamster chinois) démontrent également que la réparation n'est pas forcément accompagnée d'une intégration ciblée (Adair *et al.*, 1990; Adair *et al.*, 1989; Adair *et al.*, 1998). Dans ce système, le vecteur peut être réparé selon la matrice chromosomique et s'intégrer ailleurs dans le génome. Ce type d'événement est appelé "hit and run". Il n'est pas fréquemment observé chez la levure car l'intégration de séquences homologues par recombinaison illégitime est très rare. En revanche, il a lieu lors d'événements d'intégration de l'ADN transformant homologue chez les végétaux (Vergunst and Hooykaas, 1999). Offringa et collaborateurs ainsi que le groupe de Lee ont observé des événements "hit and run" lors des tentatives d'introduire l'ADN transformant dans le génome par recombinaison homologue (nous les avons revues à la page 10).

Les études de la mutagenèse ciblée effectuées jusqu'ici sur la mousse *P. patens* ont utilisé des vecteurs circulaires ou linéaires coupés dans la région d'homologie ininterrompue. Leur intégration dans le génome de la mousse semble être régie par les mêmes mécanismes que ceux de la levure *S. cerevisiae*. La question légitime qui se pose est celle de savoir si les vecteurs interrompus par une brèche dans la région d'homologie peuvent être réparés à partir de la matrice génomique comme cela a lieu dans la levure *S. cerevisiae*.

Dans ce chapitre, nous étudions l'intégration et la réparation des séquences interrompues homologues de deux sites.

## 1.2. *ZLAB1* et $\lambda$ 213 permettent d'étudier la réparation des brèches

Nous avons décrit le gène *ZLAB1*, un membre de famille multigénique *Cab*, dans le chapitre I. Les vecteurs circulaires et linéaires coupés dans un site unique de la région d'homologie d'env. 1 kb s'intègrent dans ce gène avec une fréquence de 25% et 20% des transformants intégratifs, respectivement. L'inactivation du gène *ZLAB1* n'altère pas le phénotype de façon discernable.

Le locus  $\lambda$ 213 a été décrit par Schaefer (Schaefer and Zryd, 1997). Ce dernier travail démontre qu'un vecteur circulaire portant une région d'homologie de 2.3 kb s'intègre dans ce locus à copie unique avec une fréquence d'env. 67% (4/6) des transformants intégratifs. La même fréquence d'intégration du vecteur linéaire est démontrée (6/11, Schaefer D. G., données non publiées). L'insertion du vecteur dans le locus n'a pas altéré le phénotype des transformants de façon discernable.

Ces locus sont suffisamment étudiés pour être retenus pour la présente recherche.

## 2. Résultats et discussion

### 2.1. Stratégie de transformation

J'ai utilisé les vecteurs de transformation présentés dans la Figure 9: pGLcaba (Hoffmann *et al.*, 1999), (cf. aussi le chapitre précédent) et pHP213 (Schaefer and Zryd, 1997) ainsi que leurs dérivés pGLcaba $\Delta$ NgoMI et pHP213 $\Delta$ BglII (que j'ai construits, cf. la carte détaillée dans l'Annexe B, pages 137 et 138).

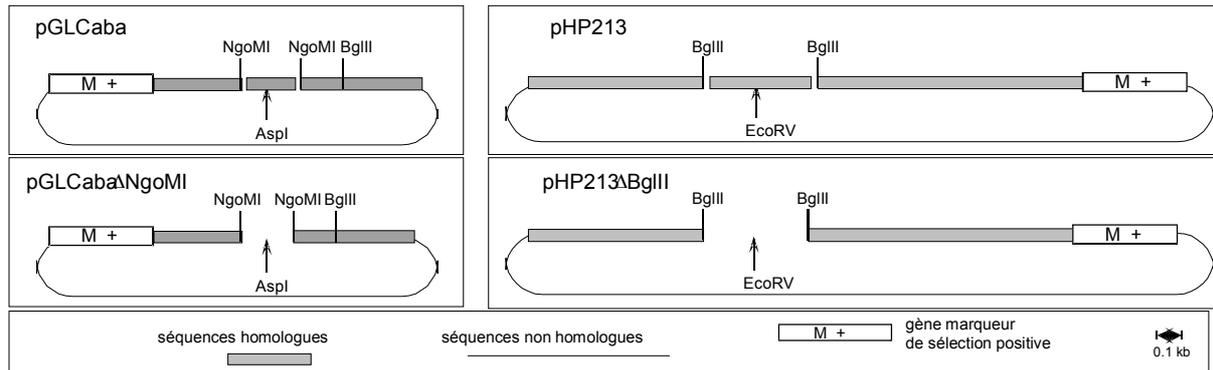


Figure 9. Les séquences interrompues utilisées pour étudier la réparation de brèches.

pGLcaba - la linéarisation du vecteur pGLcaba sur le site NgoMI libère un fragment de 0.2 kb ce qui produit une brèche de 0.2 kb dans le fragment de 1.0 kb homologue au locus *ZLAB1*. La longueur d'homologie est de 0.35 kb en 5' et de 0.5 kb en 3'. La délétion élimine le site AspI.

pGLcaba $\Delta$ NgoMI - correspond au vecteur pGLcaba dans lequel une délétion de 200 bp a été créée entre les sites NgoMI.

pHP213 - la linéarisation du vecteur pHP213 sur le site BglII libère un fragment de 0.4 kb ce qui produit une brèche de 0.4 kb dans le fragment de 2.3 kb homologue au locus  $\lambda$ 213. La longueur d'homologie est de 0.6 kb en 5' et de 1.3 kb en 3'. La délétion élimine le site EcoRI.

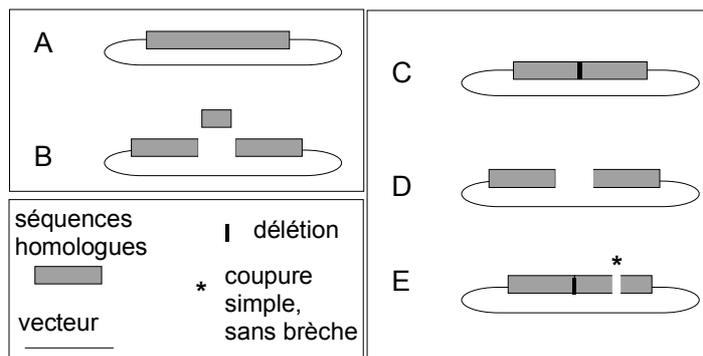
pHP213 $\Delta$ BglII - correspond au vecteur pHP213 dans lequel une délétion de 500 bp a été créée entre les sites BglII.

La flèche indique le site de restriction qui disparaît avec la création de la brèche.

Nous transformons *P. patens* avec chacun de ces vecteurs préparés en différentes configurations: circulaire, coupé dans un site créant une brèche et coupé dans un site sans créer de brèche.

La première analyse est la comparaison de l'efficacité de la transformation à l'aide de la valeur RTF (Figure 11). Nous vérifions ensuite la réparation des séquences interrompues dans les transformants intégratifs par des analyses moléculaires.

Nous remarquons que les vecteurs pGLcaba et pHP213 sont interrompus par une brèche (Figure 10, B). En revanche, les vecteurs pGLcaba $\Delta$ NgoMI et pHP213 $\Delta$ BglII sont interrompus soit par une brèche (Figure 10, D) soit par une délétion (Figure 10, C et E). Nous ferons une distinction entre la réparation d'une **brèche** et celle d'une **délétion**.



A - vecteur circulaire;  
 B - produit de la digestion du vecteur comportant le vecteur ouvert interrompu et le fragment correspondant à la brèche;  
 C - vecteur circulaire interrompu par une délétion;  
 D - vecteur ouvert interrompu par une brèche;  
 E - vecteur interrompu par une délétion coupé dans le fragment homologue en dehors de la délétion.

Figure 10. Explication de la configuration des vecteurs de transformation

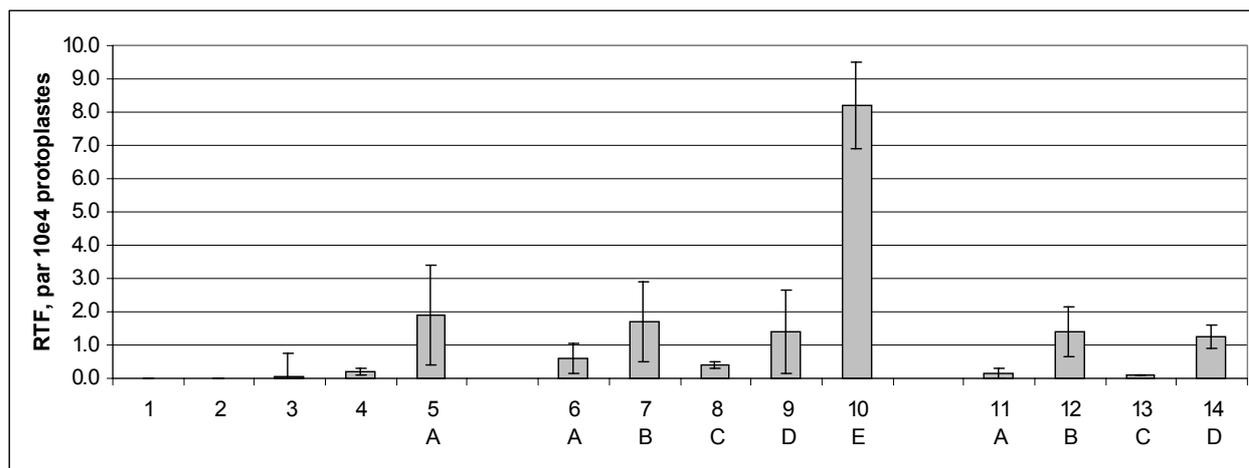


Figure 11. Caractérisation des transformations par des vecteurs portant des séquences interrompues.

La valeur RTF est calculée en divisant le nombre de transformants intégratifs par le nombre de protoplastes régénérés pour une expérience. Nous avons régénéré au total environ  $2 \text{ à } 6 \cdot 10^5$  protoplastes par vecteur, chaque vecteur a été utilisé dans 4 expériences, sauf le vecteur pHP23 (colonnes 3 et 4) qui en totalise 2). La barre d'erreur indique l'écart-type.

**Vecteurs de contrôle.** La lettre indique la configuration du fragment homologue comme ceci est schématisé dans la Figure 10 lorsque le vecteur porte un fragment homologue (1 - 4 aucune homologie):

- 1 - pas d'ADN. Aucun transformant n'est sélectionné;
- 2 - le vecteur pGL2 (circulaire surenroulé) porte le marqueur de sélection 35S *hpt* mais pas de fragment d'homologie. Nous attendons une valeur RTF très basse, de l'ordre de  $10^{-5}$ ;
- 3 - le vecteur pHP23 (circulaire surenroulé) porte le marqueur de sélection 35S *nptII* mais pas de fragment d'homologie. Nous attendons une valeur RTF très basse, de l'ordre de  $10^{-5}$ ;
- 4 - le vecteur pHP23 linéaire coupé (avec NdeI);
- 5 - le vecteur pGL 108 (circulaire surenroulé) porte un fragment homologue de 3.6 kb (Schaefer and Zryd, 1997), nous attendons une valeur RTF de l'ordre de  $2.35 \pm 1.44 \cdot 10^{-4}$ .

**Les vecteurs d'intégration dans les locus étudiés.** La lettre indique la configuration du fragment homologue comme ceci est schématisé dans la Figure 10:

- 6 A - le vecteur pGLcaba circulaire surenroulé porte un fragment homologue du gène ZLAB1 (chapitre I);
- 7 B - le produit de la digestion du vecteur pGLcaba avec NgoMI;
- 8 C - le vecteur pGLcabaΔNgoMI circulaire surenroulé porte une délétion dans le fragment homologue du gène ZLAB1;
- 9 D - le vecteur pGLcabaΔNgoMI digéré avec NgoMI;
- 10 E - le vecteur pGLcabaΔNgoMI digéré avec BglII;
- 11 A - le vecteur pHP213 circulaire surenroulé porte un fragment homologue du locus λ213 ((Schaefer and Zryd, 1997));
- 12 B - le produit de la digestion du vecteur pHP213 avec BglII;
- 13 C - le vecteur pHP213ΔBglII circulaire surenroulé porte une délétion dans le fragment homologue λ213;
- 14 D - le vecteur pHP213ΔBglII digéré avec BglII.

Nous tirons les conclusions suivantes de l'analyse des valeurs RTF :

- ✓ Les vecteurs homologues s'intègrent plus fréquemment que les vecteurs non-homologues, qu'ils soient linéaires ou circulaires surenroulés.
- ✓ Les vecteurs linéarisés s'intègrent plus fréquemment que les vecteurs circulaires surenroulés, qu'ils soient homologues ou non.
- ✓ Les vecteurs circulaires s'intègrent de façon efficace, que le fragment homologue soit interrompu par une délétion ou non (Figure 11, comparaison entre 6A et 8C; entre 11A et 13C).
- ✓ Les vecteurs linéaires s'intègrent de façon plus efficace que les vecteurs circulaires, d'un facteur de 3 à 20 ( Figure 11, comparaison entre 7B et 6A; entre 9D, 10E et 8C; entre 12B et 11A; entre 14B et 13A) même lorsqu'une brèche est créée par la coupure.
- ✓ Le vecteur pGLcabaΔNgoMI s'intègre mieux lorsque la coupure ne crée pas de brèche (Figure 11, comparaison entre 10E et 8C) que lorsque la coupure en crée une (Figure 11, comparaison entre 9D et 8C).
- ✓ Le vecteur ouvert interrompu s'intègre de façon plus efficace que le vecteur circulaire correspondant, que l'ADN transformant contienne le fragment manquant de la brèche ou non (Figure 11, la différence entre 7B et 6A est la même que celle entre 9D et 6A, la différence entre 12B et 11A est la même que celle entre 14D et 11A).

Ces conclusions nous rappellent que les formes linéaires de l'ADN transformant induisent la réparation des ruptures double brin par la recombinaison homologue et par NHEJ ce qui se traduit par des valeurs RTF plus élevées pour la transformation par des vecteurs linéaires. Lorsqu'un vecteur linéaire est interrompu par une brèche, l'intégration est moins efficace probablement parce que la brèche doit être soit reliée par ligation, soit réparée, les deux processus imposant une contrainte supplémentaire à l'intégration. Nous en discutons après avoir vérifié si les brèches et les délétions sont réparées lors de l'intégration.

## 2.2. Analyse moléculaire d'un échantillon des transformants intégratifs

La meilleure démonstration d'une réparation fidèle est l'analyse de la séquence intégrée. Néanmoins, une analyse plus rapide consiste à démontrer la restitution d'un site de restriction spécifique dans le fragment réparé.

Lorsque les produits de la digestion des vecteurs pGLcaba et pHP213 sont utilisés pour transformer la mousse *P. patens*, la présence des fragments libérés (inserts) qui correspondent à la brèche compliquent l'interprétation des résultats (l'étude sur *S. cerevisiae* dévoile une situation similaire (Orr-Weaver *et al.*, 1981)). En effet, le produit de la digestion peut contenir des molécules du vecteur circulaire (non détectables à l'analyse de la digestion ou formées par ligation d'une copie du vecteur par ligature). Dans cette situation, la brèche peut être réparée non seulement sur la matrice génomique mais aussi sur la matrice extrachromosomique (vecteurs circulaires) par recombinaison homologue extrachromosomique. De plus, la réparation peut se faire par re-ligature du vecteur avec l'insert dans la cellule de la mousse. Dans ce cas, la réparation peut avoir lieu pas forcément par une synthèse des fragments manquant mais par une ligation des inserts dépendante de l'homologie (un mécanisme proposé par Glasunov, (Glasunov and Glaser, 1999), dont nous discutons à la page 42).

Nous présentons l'analyse moléculaire détaillée des souches issues des transformations par les vecteurs interrompus (cf. la Figure 11): pHP213 digéré avec BglII (12B), pHP213 $\Delta$ BglII (13C et 14D) et pGLcaba $\Delta$ NgoMI (8C, 9D, 10E).

La vérification de la réparation se base sur l'hypothèse que plusieurs copies du vecteur s'intègrent en répétitions directes; cette hypothèse est soutenue par les résultats précédents. La digestion de l'ADN génomique libère alors les monomères du vecteur intégré. L'intensité du signal d'hybridation avec le vecteur (ou un de ces fragments utilisé comme sonde) est proportionnelle au nombre de copies intégrées. Les fragments hybrides entre le vecteur et le locus d'intégration sont analysés séparément, car le signal des fragments à copie unique est beaucoup plus faible que le signal correspondant aux monomères du vecteur présents en plusieurs copies.

Nous vérifions l'intégration ciblée par hybridation Southern avec une sonde spécifique à un fragment hybride entre le vecteur de transformation et le site d'intégration.

## Locus $\lambda 213$

**La vérification de l'intégration ciblée** se base sur la connaissance du RFLP BamHI de la souche sauvage et sur la prédiction du RFLP BamHI du locus  $\lambda 213$  interrompu par l'insertion ciblée du vecteur (Figure 12).

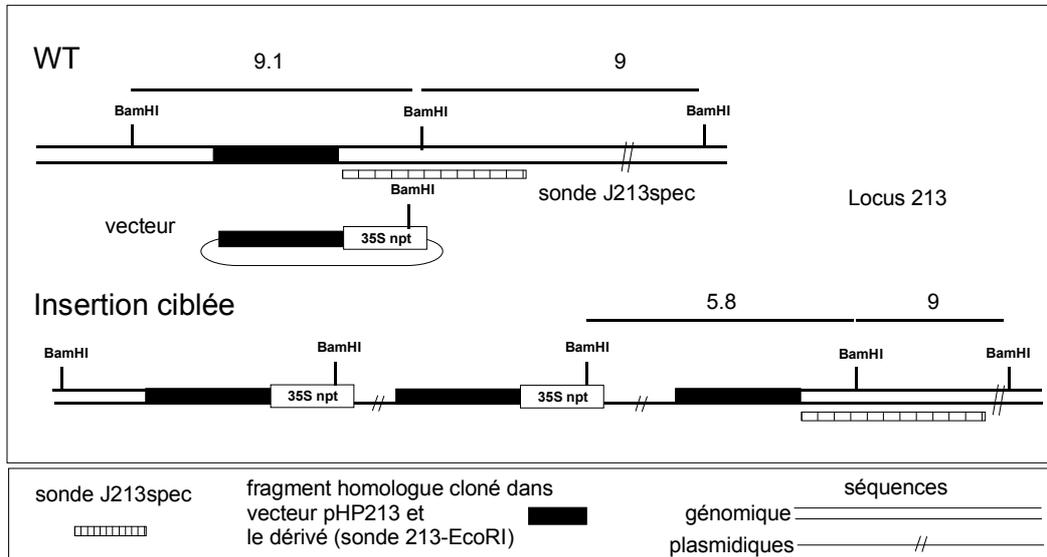


Figure 12. L'intégration ciblée du vecteur réparé dans le locus  $\lambda 213$  remplace le fragment BamHI de 9.1 kb par un fragment hybride de 5.8 kb.

La sonde Jspec213 (hachurée) est un fragment génomique de Sali-EcoRI d'environ 4.5 kb adjacente au fragment EcoRI de 2.3 kb qui a été utilisé comme région d'homologie dans les vecteurs de transformation pHP213 et pHP213 $\Delta$ BglII (en noir).

WT - L'hybridation de l'ADN WT de *P. patens* digéré par BamHI avec cette sonde présente 2 signaux de forte intensité (9.0 kb, 9.1 kb) et 1 signal d'intensité plus faible (1.5 kb) cf. Figure 13. Les signaux de 9.0 et 9.1 kb correspondent au locus  $\lambda 213$  et sont montrés ici; le signal faible (1.5 kb) provient d'une hybridation croisée avec un locus inconnu.

Insertion ciblée - l'intégration du vecteur provoque le remplacement du fragment de 9.1 kb par un fragment hybride de 5.8 kb. Les fragments de 9.0 kb et de 1.5 kb doivent rester intacts.

Notons que le fragment de 5.8 kb comprend une copie complète du fragment homologue. Si une copie non réparée s'intégrait en aval des répétitions du vecteur, la taille de ce fragment serait de 5.3 kb (le schéma ne le montre pas).

**La vérification de la réparation de la brèche** se base sur l'observation du fait que la réparation correcte du vecteur pHP213 $\Delta$ BglII restaure le vecteur pHP213. Le plasmide pHP213 digéré avec EcoRV a donc servi de contrôle positif de réparation et le plasmide pHP213 $\Delta$ BglII digéré avec EcoRV a servi de contrôle négatif de réparation. Les deux se distinguent facilement par le RFLP EcoRV (Figure 9, page 30, et ci-dessous dans la Figure 14 B, pistes R et NR, respectivement). Nous utilisons le fragment d'homologie porté par ces vecteurs comme sonde (213-EcoRI). La prédiction détaillée se trouve dans l'Annexe C, page 143.

Les contrôles R et NR ne sont pas présents sur le gel qui nous permet d'analyser l'intégration du vecteur pHP213 digéré par BglII. Nous tirons nos conclusions en analysant la taille des fragments correspondant (R - 4.7 kb et 2 kb; NR - 6.7 kb, cf. Figure 14 B, pistes R et NR, respectivement).

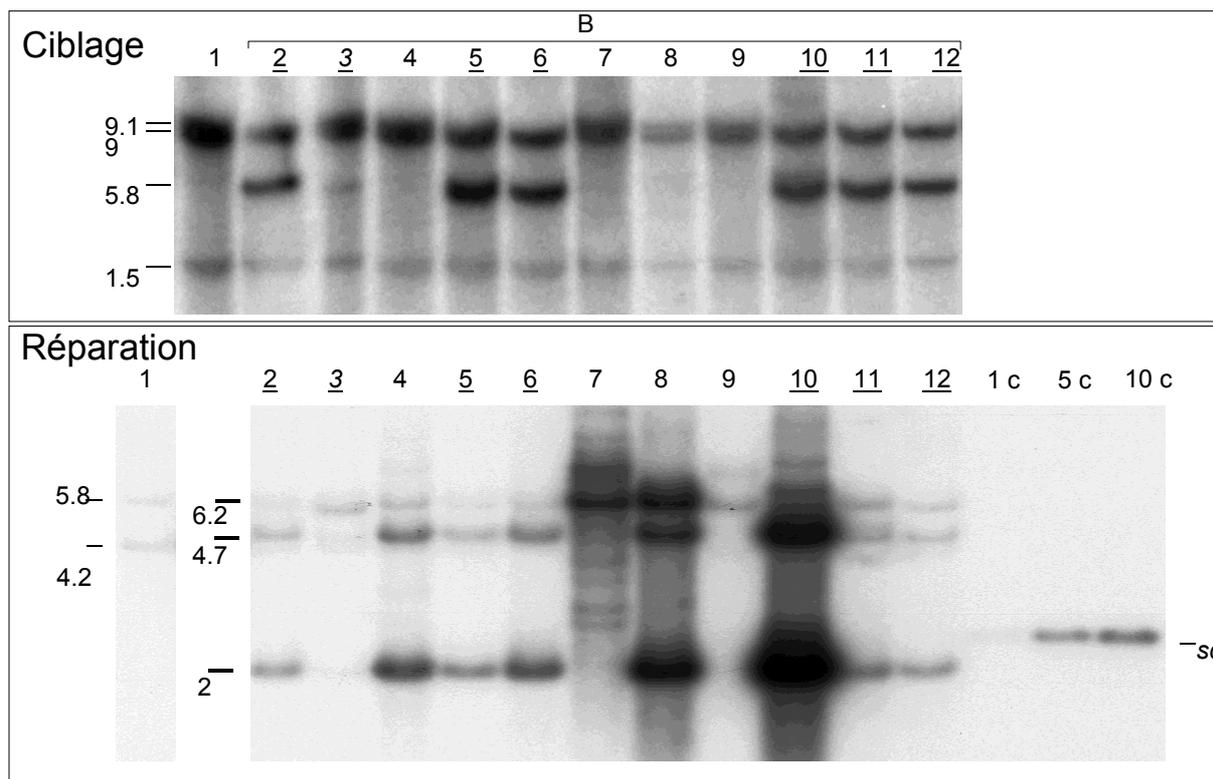


Figure 13. Analyse de l'intégration ciblée et de la réparation par hybridation Southern pour le locus  $\lambda$ 213.

**Les deux panneaux:**

pistes 1 - WT; 2 à 12 - souches transformées par le vecteur pHP213 digéré avec BglIII, condition Figure 10 B; Nous avons souligné les souches à intégration ciblée.

**Ciblage** - L'hybridation de l'ADN digéré avec BamHI avec la sonde Jspec213 vérifie l'intégration ciblée dans le locus  $\lambda$ 213.

pistes 1, 4, 7, 8, 9 - Le RFLP WT indique que le locus  $\lambda$ 213 n'est pas ciblé (3, en italique - est peut être chimérique car le doublet à 9 et 9.1 est présent en même temps qu'un signal de 5.8 kb de faible intensité);

pistes 2, 3, 5, 6, 10 à 12 - Le remplacement du fragment de 9.1 kb par le fragment hybride de 5.8 kb montre que le vecteur s'est intégré dans le locus  $\lambda$ 213. La copie intégrée en aval du site d'insertion est complète (sans délétion, montré dans la Figure 12, page 34).

**Réparation** - L'hybridation de l'ADN digéré avec EcoRV avec la sonde 213-EcoRI vérifie la réparation.

1c, 5c, 10c - Le plasmide pHP213 surenroulé (sc) est chargé en quantités correspondant à 1, 5 et 10 copies par génome;

Dans toutes les pistes, nous détectons les signaux de 4.2 et de 5.8 kb correspondant au fragments WT. La conservation de ces fragments après l'insertion des répétitions du vecteur dans les clones ciblés s'explique par l'hypothèse que les copies réparées se trouvent en amont et en aval des répétitions du vecteur.

pistes 7, 9 - l'intégration d'une partie des copies non réparées est révélée par le fragment de 6.2 kb;

piste 3 - réparation partielle (en italique - clone chimérique?);

pistes 2, 4-6, 8, 10-12 - l'intégration des copies réparées du vecteur (signaux 4.7 et 2.0 kb correspondent au vecteur entier). Nous pouvons soupçonner une intégration d'une très faible partie des copies non réparées de 6.2 kb, signal qui serait masqué par le signal de 5.8 kb (prononcé dans les pistes 4, 11, 12);

Nous détectons quelques signaux de 6.7 kb qui correspondent au vecteur réparé partiellement digéré (4.7 + 2 kb).

Pour le locus  $\lambda 213$ , les événements d'intégration du vecteur pHP213 digéré avec BglII (séquence homologue interrompue par une brèche et l'insert qui correspond à la brèche, Figure 11, B, page 31) se composent des classes suivantes:

Tableau 1. Les événements d'intégration et de réparation du vecteur pHP213 $\Delta$ BglII portant une séquence homologue interrompue.

Souche présentée dans la Figure 13, piste:	Structure du vecteur, voir Figure 10, page31		intégration	réparation §	≈ nombre de copies intégrées
2	brèche - B	ciblée	<i>conforme au modèle</i>	100%	5
3	brèche - B	<i>chimérique?</i>	<i>signaux faibles</i>	10%	1-3
4	brèche - B		<i>"hit and run", intégration illégitime</i>	90-100%	20
5	brèche - B	ciblée	<i>conforme au modèle</i>	100%	10
6	brèche - B	ciblée	<i>conforme au modèle</i>	100%	20
7	brèche - B		<i>intégration illégitime, complexe</i>	aucune	30
8	brèche - B		<i>"hit and run", intégration illégitime</i>	50%	40
9	brèche - B		<i>intégration illégitime</i>	aucune	5
10	brèche - B	ciblée	<i>conforme au modèle</i>	100%	>40
11	brèche - B	ciblée	<i>conforme au modèle</i>	90%	10
12	brèche - B	ciblée	<i>conforme au modèle</i>	90%	5-10

§ - La réparation est estimée en pourcentage des copies réparées dans le nombre total des copies intégrées selon l'intensité du signal.

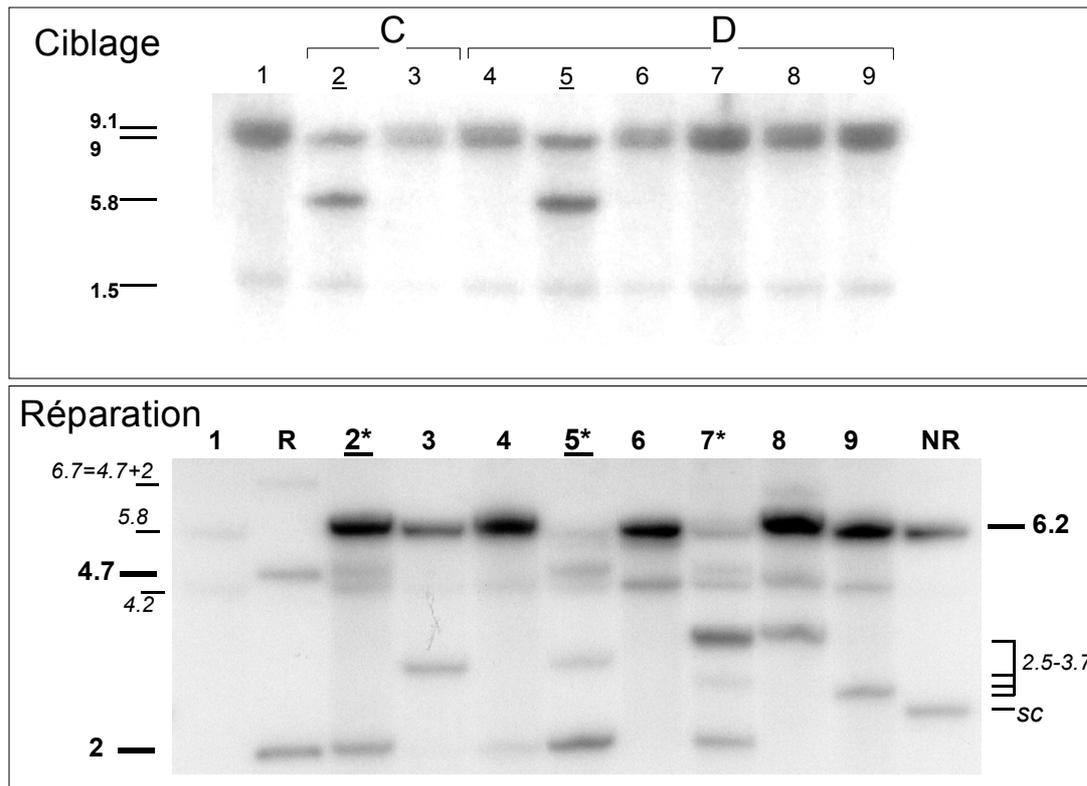


Figure 14. Analyse de l'intégration ciblée et de la réparation par hybridation Southern pour le locus  $\lambda 213$ .

**Les deux panneaux:**

pistes 1 - WT; 2 et 3 - souches transformées par le vecteur pHP213 $\Delta$ BglII circulaire, condition Figure 10 C; 4-9 souches transformées par le vecteur pHP213 $\Delta$ BglII interrompu, coupé avec BglII, condition Figure 10, D. Nous avons souligné les souches à intégration ciblée et nous avons marqué d'un astérisque les souches qui montrent une intégration de copies réparées du vecteur.

**Ciblage** - L'hybridation de l'ADN digéré avec BamHI avec la sonde Jspec213 vérifie l'intégration ciblée dans le locus  $\lambda 213$ .

pistes 1, 3, 4, 6 à 9 - le RFLP WT indique que le locus  $\lambda 213$  n'est pas ciblé;

pistes 2 et 5 - Le remplacement du fragment de 9.1 kb par le fragment hybride de 5.8 kb montre que le vecteur s'est intégré dans le locus  $\lambda 213$ . La copie intégrée en aval du site d'insertion est complète (sans délétion, montré dans la Figure 12, page 34).

**Réparation** - L'hybridation de l'ADN digéré avec EcoRV avec la sonde 213-EcoRI vérifie la réparation.

Contrôle NR - plasmide pHP213 $\Delta$ BglII digéré avec EcoRV chargé en quantité correspondant à 5 copies génomiques, sc - fraction non digérée;

Contrôle R - plasmide pHP213 digéré avec EcoRV;

pistes 2-4, 6-9 - l'intégration de copies non réparées du vecteur dans tous les clones (signal 6.2 kb correspond au contrôle NR);

pistes 2, 7 - l'intégration d'une partie des copies réparées du vecteur (signaux 4.7 et 2.0 kb qui correspondent au contrôle R);

piste 5 - toutes les copies intégrées sont réparées;

pistes 3, 4 - le fragment de 2 kb correspond au fragment du vecteur réparé et intégré par recombinaison illégitime, les copies réparées sont une partie mineure de toutes les copies;

pistes 3, 7, 8, 9 - les fragments d'une taille entre 2.5 et 3.7 kb sont les fragments hybrides formés par une intégration illégitime;

piste 5 - le signal à environ 3.5 kb n'est pas un fragment de deux copies du vecteur coupé dans le site BglII liées "head-to-head" puisque celui-ci devrait être accompagné par un autre signal "tail-to-tail", absent du tableau.

Le signal correspondant au fragment EcoRV de 5.8 kb du locus  $\lambda 213$  est masqué par le signal de forte intensité du monomère du vecteur pHP213 $\Delta$ BglII de 6.2 kb. Le signal correspondant au fragment EcoRV de 4.2 kb du locus  $\lambda 213$  est présent dans tous les clones. La conservation de ce fragment après l'insertion des répétitions du vecteur dans les clones ciblés (pistes 2 et 5) s'explique par l'hypothèse que les copies réparées se trouvent en amont et en aval des répétitions du vecteur. C'est en effet ces copies-là qui interagissent avec le chromosome lors de l'intégration.

Pour le locus  $\lambda 213$ , les événements d'intégration du vecteur pHP213 $\Delta$ BglIII portant une séquence homologue interrompue se composent des classes suivantes:

Tableau 2. Les événements d'intégration et de réparation du vecteur pHP213 $\Delta$ BglIII portant une séquence homologue interrompue.

Souche présentée dans la Figure 14, piste:	Structure du vecteur, voir Figure 10, page31		intégration	réparation §	≈ nombre de copies intégrées
2	délétion - C	ciblée	<i>conforme au modèle</i>	20%	15
3	délétion - C		<i>"hit and run", intégration illégitime</i>	10%	10
4	brèche - D		<i>"hit and run", intégration illégitime</i>	20%	15
5	brèche - D	ciblée	<i>un signal inattendu (cf. texte)</i>	100%	5
6	brèche - D		<i>intégration illégitime</i>	aucune	10
7	brèche - D		<i>"hit and run", intégration illégitime</i>	30%	5
8	brèche - D		<i>intégration illégitime</i>	aucune	15
9	brèche - D		<i>intégration illégitime</i>	aucune	15

§ - La réparation est estimée en pourcentage des copies réparées dans le nombre total des copies intégrées selon l'intensité du signal.

Nous discuterons de ces événements plus loin avec les résultats obtenus pour le gène *ZLAB1*.

### *ZLAB1*

**La vérification de l'intégration ciblée** se base sur la connaissance du RFLP EcoRI et XbaI de la souche sauvage et sur la prédiction du RFLP du gène *ZLAB1* interrompu par l'insertion ciblée du vecteur. Les digestions par les deux enzymes de l'ADN de tous les clones ont été hybridées avec la sonde spécifique au gène *ZLAB1* cab5spec (chapitre I).

**La vérification de la réparation de la brèche** se base sur l'observation du fait que la réparation correcte du vecteur pGLcaba $\Delta$ NgoMI restaure le vecteur pGLCaba. Par conséquent, l'intégration des copies du vecteur pGLcaba $\Delta$ NgoMI réparées correctement équivaut à l'intégration du vecteur pGLcaba. L'ADN de la souche CAB-NT1 obtenue avec ce dernier a donc servi de contrôle positif de réparation. Le plasmide pGLcaba $\Delta$ NgoMI digéré avec AspI a servi de contrôle négatif de réparation. Les deux se distinguent facilement par le RFLP AspI (Figure 9, page 30 et ci dessous dans la Figure 15 B pistes R et NR, respectivement). Nous utilisons le plasmide pGL2 comme sonde. La prédiction détaillée se trouve dans l'Annexe C, page 144.

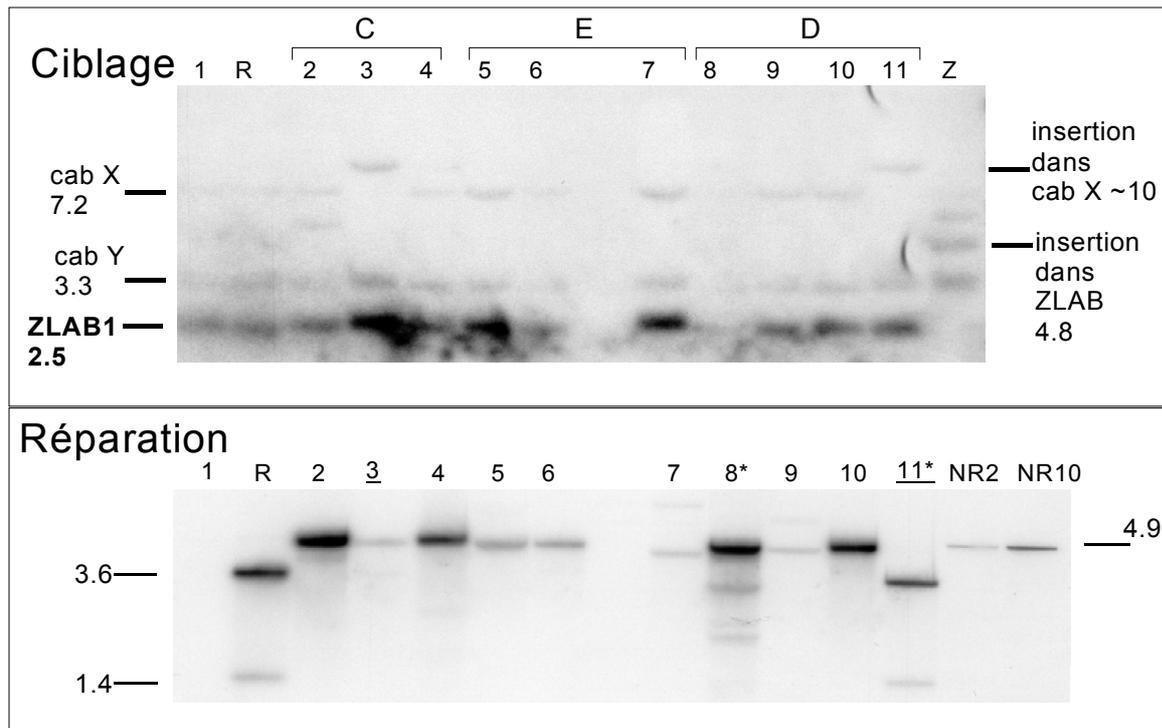


Figure 15. Analyse de l'intégration ciblée et de la réparation par hybridation Southern pour le locus *ZLAB1*.

**Les deux panneaux** pistes 1 - WT; 2 à 4 les souches transformées par le vecteur pGLCabaΔNgoMI circulaire, cf. condition Figure 10 C, page 31; pistes 5 à 7 - par le vecteur pGLCabaΔNgoMI interrompu par une délétion et coupé avec BglII, cf. condition Figure 10 E, page 31; 8 à 11 - par le vecteur pGLCabaΔNgoMI interrompu, coupé avec NgoMI, cf. condition Figure 10 D, page 31. Nous avons souligné les souches à intégration ciblée dans un gène *Cab* homologue et nous avons marqué d'un astérisque les souches qui montrent une intégration de copies réparées du vecteur;

**Ciblage** - L'hybridation de l'ADN digéré avec EcoRI avec la sonde cab5spec vérifie l'intégration ciblée dans *ZLAB1*.

piste 1 - WT; contrôle R - intégration illégitime du vecteur pGLCaba dans une souche du contrôle (CABNT1, cf. Figure 3, page 21); contrôle Z - intégration ciblée du vecteur pGLCaba dans le gène *ZLAB1* (CAB GT1, cf. Figure 3, page 21).

pistes 2 à 11 - le signal correspondant au fragment intact de l'allèle sauvage de 2.5 kb dans la digestion EcoRI indique que le locus *ZLAB1* n'est touché dans aucune souche (piste 8 est visible sur la surexposition).

pistes 1, 2, 4 à 10 - deux autres gènes membres de la famille *Cab* sont reconnus par la sonde cab5spec. Les signaux de faible intensité correspondent à des fragments EcoRI de 3.3 kb (d'un gène *Cab Y*) et 7.2 kb (d'un gène *Cab X*).

pistes 3 et 11 - l'intégration dans le gène *Cab X* remplace le fragment EcoRI de 7.2kb par un fragment de 10 kb.

**Réparation** - L'hybridation de l'ADN digéré avec AspI avec la sonde pGL2 vérifie la réparation du vecteur.

piste 1 - WT;

contrôle R - une souche de contrôle CABNT1 (chapitre I, Figure 3, page 21) qui porte des répétitions directes du vecteur pGLcaba dans un locus inconnu;

contrôle NR - plasmide pGLcabaΔNgoMI digéré avec AspI chargé en quantités correspondant à 2 ou 10 copies de vecteur par génome;

pistes 2 à 11 - les fragments hybrides formés par une intégration illégitime sont détectés par des signaux d'une faible intensité;

pistes 2 à 7, 9, 10 - seules les copies du vecteur pGLcabaΔNgoMI non réparées se sont intégrées;

piste 11 - toutes les copies du vecteur pGLcabaΔNgoMI intégré sont réparées;

piste 8 - une partie des copies du vecteur pGLcabaΔNgoMI intégré sont réparées. Des signaux correspondant à la taille du vecteur réparé (3.6 et 1.4 kb), accompagnés de signaux non prédits (env. 3.4, 2.6, 2.4 et 1.35 kb) sont détectables. Il est probable que la réparation ait lieu partiellement dans cette souche et qu'elle ne soit pas correcte dans certaines copies.

piste 7 - le signal est légèrement décalé vers le bas: une intégration d'une copie rognée et multipliée n'est pas exclue.

Pour le gène *ZLAB1*, les événements d'intégration du vecteur portant une séquence homologue interrompue se composent des classes suivantes:

Tableau 3. Les événements d'intégration et de réparation du vecteur pGLCabaΔNgoMI portant une séquence homologue interrompue.

<b>Souche présentée dans la Figure 15, piste:</b>	<b>interruption, voir Figure 10, page31</b>	<b>intégration</b>	<b>réparation</b>	<b>≈ nombre de copies intégrées</b>
2	délétion - C	<i>CAB X</i>	aucune	>40
3	délétion - C	<i>illégitime</i>	aucune	5
4	délétion - C	<i>illégitime</i>	aucune	30
5	délétion - E	<i>illégitime</i>	aucune	20
6	délétion - E	<i>illégitime</i>	aucune	15
7	délétion - E	<i>illégitime</i>	aucune	1-3
8	brèche - D	<i>illégitime</i> <i>présence de signaux non expliqués (cf. texte)</i>	10%	>40
9	brèche - D	<i>illégitime</i>	aucune	1-3
10	brèche - D	<i>illégitime</i>	aucune	>40
11	brèche -D	<i>CAB X</i>	100%	30

§ - La réparation est estimée en pourcentage des copies réparées dans le nombre total des copies intégrées selon l'intensité du signal.

Il est intéressant de noter que le vecteur pGLcabaΔNgoMI ne s'est pas intégré dans le gène *ZLAB1* mais dans un gène homologue *Cab* inconnu (*CAB X*). Toutes les copies intégrées dans le gène *CAB X* sont réparées. La séquence codante des gènes *Cab* étant bien conservée, la réparation prédite a pu avoir lieu selon la matrice du gène *ZLAB1* ou selon la matrice du site d'intégration (*CAB X*).

## 2.3. Classification des événements et conclusions

L'analyse de l'efficacité de l'intégration des séquences interrompues (mesurée par la valeur RTF, Figure 11 et conclusions, page 31) montre que l'interruption du vecteur par une brèche n'abolit pas l'intégration quoiqu'elle la rende moins efficace par rapport au vecteur linéaire correspondant. En effet, lorsque la brèche est réparée correctement, la synthèse du fragment manquant prend un temps supplémentaire. En revanche, lorsque le vecteur interrompu est ligaturé sans réparation, l'intégration est aussi peu efficace que celle d'un vecteur circulaire. L'analyse moléculaire démontre que les deux cas (réparation ou ligature) sont observées. Quant à la présence d'une délétion, elle n'affecte pas l'efficacité de l'intégration.

Nous regroupons les événements observés dans le Tableau 4 et présentons ensuite les conclusions que cette répartition nous permet de faire.

Tableau 4. Classification des événements observés lors de l'intégration des séquences interrompues par une brèche ou par une délétion.

Classe	Intégration	Réparation	Forme d'ADN				Nombre total d'événements
			<i>Brèche</i>		<i>Délétion</i>		
			<i>B</i>	<i>D</i>	<i>C</i>	<i>E</i>	
I	<u>ciblée</u>	avec réparation $\geq 50\%$	6	2	0	nt	<b>8</b>
II		avec réparation $< 50\%$	0	0	1	nt	<b>1</b>
III		sans réparation	0	0	1	nt	<b>1</b>
IV	<u>non-ciblée</u>	"hit and run": avec réparation $\geq 50\%$	2	0	0	0	<b>2</b>
V		"hit and run": avec réparation $< 50\%$	0	3	1	0	<b>4</b>
VI		sans réparation	2	5	2	3	<b>12</b>
<b>Nombre total d'événements :</b>			<b>10</b>	<b>10</b>	<b>5</b>	<b>3</b>	<b>28</b>

nt - non testé

### Réparation des séquences interrompues par une brèche (conditions B et D)

- ✓ La réparation d'une séquence interrompue par une brèche (conditions B ou D, cf. la Figure 10, page 31) est efficace et complète lorsque l'intégration est ciblée (classe I contre II et III).
- ✓ Lorsque l'intégration est illégitime, la réparation est moins fréquente et jamais complète (classes V et VI) si la matrice génomique seule est disponible (condition D). En revanche, si une matrice extrachromosomique est disponible (condition B, présence des copies réparées du vecteur, cf. page 32), la réparation peut être complète (classe IV) car elle peut

probablement avoir lieu par recombinaison extrachromosomique. Ce dernier mécanisme apparaît donc plus efficace que la réparation sur la matrice génomique. Nous tirons cette conclusion avec une certaine réserve car, comme nous l'expliquons à la page 32, les vecteurs interrompus de type B peuvent être réparés par ligature avec l'insert présent dans l'ADN transformant.

- ✓ Nous constatons que les cas de réparation sur la matrice **génomique** suivie d'une intégration illégitime (condition D, classe V) ressemblent aux événements nommés "hit and run": la correction du vecteur rapportée pour les cellules CHO (Adair *et al.*, 1998) ou l'intégration de l'ADN chez les végétaux (Vergunst and Hooykaas, 1999).

L'analyse du fragment réparé et de la matrice de réparation est nécessaire à la compréhension du mécanisme de celle-ci. L'analyse de l'événement "hit and run" nous le permet: une recombinaison par échange réciproque entre le génome et le vecteur répare le vecteur mais introduit une délétion dans le locus  $\lambda 213$ . En revanche, la réparation de brèche par la synthèse dépendante de l'homologie répare le vecteur mais laisse la matrice intacte. La conservation du RFLP EcoRV du locus  $\lambda 213$  après la réparation du vecteur confirme l'hypothèse d'une conversion de gènes sans échange réciproque de séquences. Nous ne pouvons pas exclure la possibilité que la correction du vecteur ait lieu après son intégration illégitime en *cis* ou en *trans*. La conversion selon ce mécanisme implique une recombinaison intrachromosomique, mais elle moins fréquente et donc moins probable. Offringa suppose que les deux mécanismes expliquent les cas "hit and run" qu'il observe lors de l'intégration des séquences interrompues chez *N. tabacum* (Offringa, 1992; Offringa *et al.*, 1993).

Nous remarquons qu'une partie des copies du vecteur interrompu échappe à toute réparation, que le fragment manquant soit présent dans l'ADN transformant ou non. Ces molécules sont re-circularisées et intégrées en répétitions directes. L'intégration des séquences interrompues non réparées témoigne d'un mécanisme de ligature efficace indépendant de la synthèse. La ligature n'est pas forcément dépendante de l'homologie puisque ces molécules s'intègrent de façon ciblée aussi bien qu'illégitime. Afin d'affirmer ou de contester une analogie avec le mécanisme de ligature dépendant de l'homologie postulé pour *S. cerevisiae* (Glasunov and Glaser, 1999), nous devons comparer l'efficacité de l'intégration d'un même vecteur coupé à l'intérieur et à l'extérieur de la région d'homologie.

#### Réparation des séquences interrompues par une délétion

Nous n'avons observé une réparation des séquences interrompues par une délétion (condition C ou E, cf. Figure 10, page 31) que dans une partie minoritaire des cas, et cette réparation n'est jamais complète (c'est-à-dire, toutes les copies ne sont jamais réparées, classes II et III).

L'étude de l'intégration ciblée dans le locus  $\lambda 213$  révèle que les copies du vecteur qui se situent en amont et en aval de l'intégration des répétitions directes sont réparées (la structure est proposée dans la Figure 12, page 34). Nous émettons alors l'hypothèse que dans ces cas, l'intégration d'un vecteur interrompu par une délétion se déroule comme suit: a) une concaténation de plusieurs copies du vecteur par recombinaison homologue extrachromosomique a lieu avant l'interaction avec la cible génomique b) ensuite, seules les copies qui entrent en réaction avec la région homologue d'intégration sont réparées. Nous mettons en évidence un cas similaire dans le chapitre III, page 58. Nous y émettons l'hypothèse que dans le cas en question, l'intégration d'un concaténat de plusieurs copies du vecteur est accompagnée de la correction de la copie dans laquelle la recombinaison a lieu.

L'observation du fait que les délétions ne sont pas toujours réparées même lors de l'intégration ciblée, nous laisse croire que la correction de l'information dans ce cas se fait selon la matrice génomique autant que selon la matrice de l'ADN exogène. La réparation des délétions est toutefois moins efficace que celle des brèches puisqu'elle n'est jamais complète et beaucoup moins fréquente lors d'une intégration non ciblée.

Le résultat pratique important est que le mécanisme de réparation des séquences interrompues n'est pas un obstacle majeur à la mutagenèse qui vise à introduire des délétions. En effet, si l'intégration des vecteurs interrompus par une délétion était entièrement dépendante de la synthèse conforme à la matrice génomique, ces délétions seraient immanquablement réparées et l'efficacité de la mutagenèse serait compromise.

La faisabilité de la mutagenèse directe avec des vecteurs de remplacement interrompus par une délétion dans la région d'homologie est étudiée dans le chapitre suivant.

# Chapitre III : Mutagenèse ciblée directe du gène *PpAPT* codant pour l'adénine phosphoribosyltransférase à l'aide de vecteurs de remplacement

## Résumé

La mutagenèse ciblée directe des gènes par remplacement d'allèles est nécessaire pour développer la mutagenèse ciblée ponctuelle de précision. Nous avons analysé l'intégration ciblée par recombinaison homologue des vecteurs de remplacement dans le gène codant pour l'enzyme adénine phosphoribosyltransférase (*PpAPT*) afin de définir la taille d'homologie optimale et la fréquence de différents types d'intégration. L'intégration a lieu par conversion génique (remplacement simple) dans un quart des cas; dans les autres cas l'intégration a lieu par insertion ou par remplacement multiple, avec des fréquences similaires. L'intégration est efficace (RTF atteint  $10^{-3}$ ) même lorsque la longueur d'homologie est de 1.2 kb. Nous avons ainsi remplacé l'allèle sauvage du gène de l'adénine phosphoribosyltransférase (*PpAPT*) par différents allèles codant pour une enzyme non fonctionnelle. Le phénotype anormal de ces différents allèles est probablement causé par une perturbation du métabolisme de l'adénine. Ces mutations sont génétiquement complémentées par les gènes homologues d'*A. thaliana* *AtAPT1* ou *AtAPT2*.

## 1. Introduction

Les stratégies de mutagenèse ciblée nécessitent l'isolation d'événements de conversion génique (remplacement) (Morton and Hooykaas, 1995; Rothstein, 1991). Dans ce chapitre, nous étudions l'intégration de différents vecteurs de remplacement afin de définir la taille optimale d'homologie ainsi que la structure du site d'intégration. L'intégration ciblée de différents types de vecteur a été étudié chez *S. cerevisiae* (Orr-Weaver *et al.*, 1981; Rothstein, 1991), chez la souris (cellules ES, (Hasty *et al.*, 1992; Hasty *et al.*, 1995; Hasty *et al.*, 1991b) et chez le hamster, (cellules CHO, (Adair *et al.*, 1990; Adair *et al.*, 1989; Adair *et al.*, 1998). Nous pouvons fonder nos prédictions sur les modèles proposés pour ces systèmes.

### 1.1. Disposition d'un vecteur de remplacement: paramètres à considérer

Nous testerons le vecteur de remplacement type (Figure 16) qui comprend:

1. un fragment homologue 5' (bras 5');
2. marqueur de sélection positive;
3. un fragment homologue 3' (bras 3');
4. plasmide de clonage.

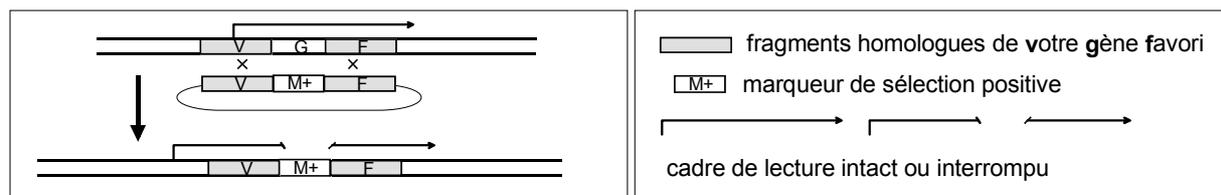


Figure 16. Intégration d'un vecteur de remplacement type (rectangle arrondi) par remplacement simple dans un site chromosomique (lignes doubles).

La relation entre la longueur des fragments homologues et l'efficacité de la mutagenèse ciblée est très différente selon les organismes et même les différents locus d'un même organisme. Chez la levure, la mutagenèse ciblée est mesurable et utilisable à partir de 0.1 kb (Koren *et al.*, 2000; Manivasakam *et al.*, 1995; Negritto *et al.*, 1997; Rothstein, 1991), tandis que chez la souris, les fragments homologues doivent être isogéniques et d'une taille de  $\geq 6$ kb (Deng and Capecchi, 1992; Riele *et al.*, 1992). Il existe une longueur optimale pour une mutagenèse efficace ou un seuil à partir duquel l'efficacité ne dépend plus de la longueur. La distribution de la longueur de l'homologie entre le fragment 5' et 3' pourrait influencer l'efficacité de la mutagenèse en générale et particulièrement la fréquence des événements de remplacement ou d'insertion en 5' ou 3'. L'influence de la forme du vecteur (circulaire ou linéaire) sur l'efficacité de la mutagenèse en général, et particulièrement sur le mode d'intégration doit être prise en compte.

Ces paramètres peuvent dépendre du locus ciblé et ne peuvent pas être prédit a priori chez *S. cerevisiae* (Rothstein, 1991). Ils peuvent être encore plus variables selon les organismes (Adair *et al.*, 1990; Adair *et al.*, 1989; Adair *et al.*, 1998; Hasty *et al.*, 1992; Hasty *et al.*, 1995; Hasty *et al.*, 1991b). Nous définissons ces paramètres pour l'intégration ciblée de vecteurs de remplacement dans le gène *PpAPT*.

## 1.2. Caractéristiques du gène *PpAPT*

Le gène *APT* codant pour l'enzyme adénine phosphoribosyltransférase (APRT, E.C. 2.4.2.7) se trouve dans tous les organismes (eubactéries, archéobactéries et les eucaryotes). L'APRT catalyse la formation de 5'-adénosine phosphate (AMP) à partir de l'adénine et de  $\alpha$ -D-5-phosphoribosyl N-pyrophosphate en présence d'ions de magnésium.

Chez la mousse *P. patens*, l'ADNc codant pour une enzyme APRT a été isolé par complémentation des mutants de *E. coli* APRT- (Schaefer *et al.*, en préparation). Le gène *PpAPT* a ensuite été isolé par criblage d'une banque d'ADN génomique. La séquence d'un fragment génomique XbaI-EcoRI de 4 kb a été déterminée, son analyse permettant de préciser la position et la longueur des introns (GenEMBL AC : xxx). L'analyse Southern de l'ADN génomique de la mousse *P. patens* avec ce fragment indique que la mousse ne possède qu'un gène *APT*. Il comprend 7 exons de 59 à 316 bp et la séquence codante est comprise entre le 2<sup>e</sup> et le 7<sup>e</sup> exons. La position des introns est similaire à celle des gènes homologues de *A. thaliana* *AtAPT1* et *AtAPT2*. La protéine de 185 aa déduite de la séquence d'ADNc présente 68% d'identité avec *AtAPRT1* et 61% avec *AtAPRT2* ((Moffatt *et al.*, 1994; Schaefer *et al.*, en préparation; Schnorr *et al.*, 1996).

Les mutants affectés dans l'activité APRT sont connus dans plusieurs organismes. Leur isolation est facilitée par leur résistance aux analogues des purines: 2,6-diaminopurine (DAP), 8-azaadénine ou 2-fluoroadénine. Chez le type sauvage, ces analogues sont convertis par l'action APRT en nucléotides toxiques correspondants. Chez *A. thaliana*, plusieurs mutants affectés dans l'activité APRT ont été isolés par la résistance à DAP (Gaillard *et al.*, 1998; Moffatt *et al.*, 1994; Moffatt *et al.*, 1991; Schnorr *et al.*, 1996).

Aucun mutant de *P. patens* dépourvu de l'activité APRT n'est connu à ce jour. Nous proposons de caractériser la fonction du gène *PpAPT* de la mousse *P. patens* à l'aide de la mutagenèse ciblée. Le gène *PpAPT* pourra être muté à l'aide de différents vecteurs de transformation. L'objectif essentiel du présent travail est de caractériser les types d'intégration de l'ADN transformant en fonction du vecteur utilisé. Nous présenterons également l'isolation et la caractérisation phénotypique de différents mutants obtenus.

Nous tentons de répondre aux questions suivantes :

- ✓ Quels sont les produits de l'intégration de vecteurs de remplacement chez *P. patens*?
- ✓ L'efficacité de la mutagenèse du locus APRT avec les vecteurs de remplacement est-elle différente de celle rapportée pour d'autres locus mutés à l'aide de vecteur d'insertion?
- ✓ Les événements de remplacement simple sont-ils suffisamment fréquents pour développer des stratégies se fondant sur les événements de remplacement?
- ✓ La longueur et la distribution de l'homologie entre le vecteur et la cible génomique affectent-elles l'efficacité de la mutagenèse?
- ✓ La forme des vecteurs de transformation (circulaire ou linéaire) influence-t-elle l'efficacité de la mutagenèse?
- ✓ L'inactivation complète du gène *PpAPT* privant *P. patens* de l'activité APRT est-elle possible avec les vecteurs de remplacement utilisés ?
- ✓ Quel est le phénotype de différents allèles mutés par l'intégration des vecteurs de transformation ?
- ✓ La mutation du gène *PpAPT* peut-elle être génétiquement complétée par des gènes d'*A. thaliana* (*AtAPT1* ou 2) ?

## 2. Résultats

### 2.1. Isolation de mutants APRT<sup>-</sup>

La résistance à DAP permet de sélectionner les mutants de mammifères ou d'*A. thaliana*. Cette approche n'est pas possible pour isoler les mutants de la levure *S. cerevisiae* puisque son enzyme APRT ne convertit pas DAP en analogue toxique de l'AMP (Alfonzo *et al.*, 1997). Nous avons dû vérifier la toxicité de DAP pour la mousse *P. patens* et la possibilité d'isoler des mutants résistants à DAP.

#### *Tests de toxicité de DAP pour P. patens*

Les protoplastes ou le protonéma fraîchement broyé de *P. patens* sont mis en culture sur le milieu PPNH4 en présence de différentes concentrations de DAP de 10 à 500 µM. La croissance du protonéma est inhibée en présence de DAP et complètement arrêtée si la concentration dépasse 100 µM. L'inhibition est proportionnelle à la concentration de DAP dans le milieu. Observées au microscope, les cellules ne s'allongent pas alors que la division peut encore continuer. Ce développement produit des filaments courts et rebondis qui ne colonisent pas le milieu, leur valant la comparaison au chou-fleur. Nous utilisons par conséquent la concentration de 150 µM à 300 µM pour inhiber la croissance du type sauvage. En présence de ces concentrations, les colonies de type sauvage ne dépassent pas le stade de quelques cellules et leur croissance est arrêtée si la concentration de DAP est maintenue. Les colonies non développées peuvent néanmoins reprendre une croissance normale à terme si elles sont transférées sur le milieu sans DAP.

#### *Résistance à DAP spontanée*

Etonnamment, une faible proportion des protoplastes régénérants de type sauvage (1 sur 10<sup>5</sup>, Schaefer D. G., communication personnelle) suppriment l'inhibition au terme de 3-4 jours et développent des colonies même en présence de DAP. La majorité de ces colonies sont compactes et colonisent peu le milieu. Les protoplastes isolés de ces clones (spDAP<sup>R</sup> pour résistance à DAP spontanée) peuvent redevenir sensibles à DAP (10 sur 14 testés). Une période de culture prolongée sur du milieu contenant DAP non renouvelé (>7 jours) provoque une apparition d'un nombre plus élevé de telles colonies DAP<sup>R</sup> (jusqu'à 10 pour 10<sup>5</sup> régénérants). Le transfert de ces colonies sur du milieu frais contenant DAP inhibe leur croissance à nouveau. Nous n'avons pas étudié ce phénomène en détail mais il est fort probable qu'à long terme, DAP soit dégradé dans le milieu ou encore que ces colonies arrivent à inhiber l'import de DAP. Pour la sélection de colonies effectivement résistantes à DAP, une période courte de croissance sur le milieu sélectif est donc préférable et un test répété est nécessaire.

#### *Mutants DAP<sup>R</sup> obtenus par mutagenèse UV*

Nous avons soumis le protonéma de type sauvage à la mutagenèse par UV en conditions correspondant à une létalité de 90-99%. Nous avons isolé les protoplastes du protonéma irradié et mis en culture en présence de DAP (250 µM). Parmi les protoplastes régénérants, environ 1 sur 10<sup>3</sup> (2 expériences: 1 sur 1200 et 3 sur 3000) développe des colonies en présence de DAP (mutDAPR - pour résistance à DAP après la mutagenèse). L'apparition de la

résistance à DAP peut ne pas être directement liée à une mutation du gène *PpAPT* ni même à une modification de l'activité APRT. De plus, nous observons la restauration de l'activité APRT dans la majorité des protoplastes issus des clones spDAPR. L'ADN de deux clones spDAPR ne présente pas de polymorphisme du gène *PpAPT* pour deux enzymes testées (données non présentées). En effet, il est possible que les protoplastes arrivent à inhiber la pénétration de DAP dans la cellule ou à rendre DAP importé inutilisable comme substrat à l'activité APRT. L'augmentation de la fréquence de l'apparition de clones DAP<sup>R</sup> après un traitement de mutagenèse d'un facteur de 100 indique que jusqu'à 99% de ces clones peuvent porter une mutation que nous n'avons pas caractérisée.

## 2.2. Stratégie de la mutagenèse ciblée

Nous avons construit quatre vecteurs de transformations AΔB, AΔBneo, BΔB et XΔX (Figure 17, la carte détaillée, l'auteur de la construction et la stratégie de clonage se trouvent dans l'Annexe B, pages 138 et 139). Trois vecteurs AΔB, BΔB et XΔX sont composés d'un fragment du locus du gène *PpAPT* portant une délétion, cloné dans un plasmide pUC21 ou pBluescript. Le quatrième vecteur AΔBneo diffère des trois précédents par la cassette de résistance à la néomycine 35S *nptII* insérée dans le site de la délétion BamHI du vecteur AΔB. L'isolation des événements d'intégration des vecteurs de transformation doit être possible soit à l'aide d'un gène marqueur présent sur le vecteur soit par la résistance à DAP provoquée par l'inactivation du gène *PpAPT*.

Nous avons utilisé le vecteur circulaire surenroulé ou linéarisé. Le vecteur linéarisé contient le fragment de remplacement homologue et le reste du plasmide non homologue.

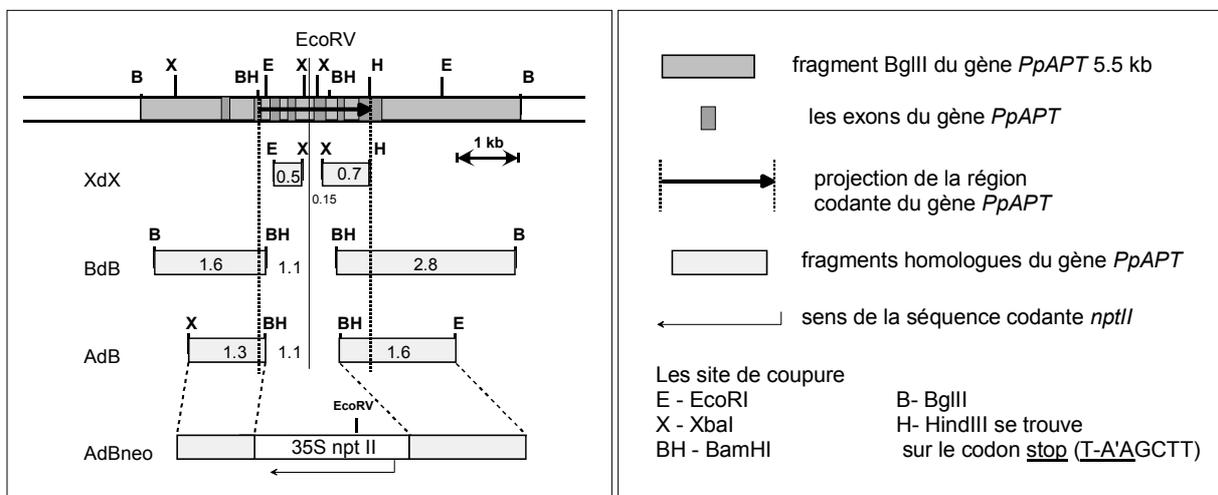


Figure 17. Vecteurs de transformation XΔX, BΔB, AΔB et AΔBneo.

Les fragments montrés sont clonés dans un plasmide, pBluescript pour AΔB et AΔBneo et pUC21 pour XΔX et BΔB. La forme linéaire de ses vecteurs comprend donc les fragments montrés et le plasmide de clonage linéarisés.

### *Intégration par simple remplacement - conversion génique*

Tous les vecteurs peuvent s'intégrer dans le génome par remplacement provoquant ainsi le remplacement de l'allèle sauvage par l'allèle interrompu par une délétion (0.15 ou 1.1 kb selon le vecteur; Figure 18, R). Le remplacement a lieu par interaction (une conversion de gènes ou

un double crossing-over) entre le gène *PpAPT* et le fragment homologue contenu dans le vecteur (le fragment de remplacement). Nous appellerons les allèles résultant du remplacement AΔB-R, AΔBneo-R, BΔB-R et XΔX-R\*. Ces allèles interrompus par une délétion ne codent pas pour une enzyme fonctionnelle parce que la délétion interrompt la séquence codante dans tous ces allèles (plus de détail à la page 67)

#### *Intégration par insertion*

L'insertion d'une ou de plusieurs copies du vecteur dans le fragment homologue peut avoir lieu en amont ou en aval de la délétion suite à un crossing-over (Figure 18, IN 5' et IN 3', respectivement). Dans ce cas nous parlerons des allèles AΔB-IN5, AΔBneo-IN5, BΔB-IN5 et XΔX-IN5 et AΔB-IN3, AΔBneo-IN3, BΔB-IN3 et XΔX-IN3, respectivement.

Une copie du fragment de remplacement linéaire s'intègre dans le locus homologue par insertion si un événement de recombinaison homologue est accompagné de NHEJ. Dans ce cas-ci, la fonctionnalité de l'allèle *PpAPT* muté ne peut être prédite avec précision. Plusieurs copies du fragment de remplacement peuvent former un concaténat (avec ou sans le reste du vecteur) par ligation et recombinaison extrachromosomique; celui-ci peut s'intégrer par 1) un crossing-over si le concaténat été refermé par ligation 2) par un double crossing-over dans un bras 5' ou 3', comme observé dans les systèmes ES ou CHO (Adair *et al.*, 1990; Adair *et al.*, 1989; Adair *et al.*, 1998; Hasty *et al.*, 1992; Hasty *et al.*, 1995; Hasty *et al.*, 1991b).

La fonctionnalité des allèles résultant de l'insertion de molécules fermées par recombinaison homologue peut être prédite comme pour le vecteur circulaire pour autant que le fragment de remplacement n'ait pas été endommagé par une activité exonucléasique. Pour qu'un événement d'insertion inactive le gène, le vecteur doit porter un fragment interne du gène tronqué en 5' et 3'. L'insertion d'un vecteur portant les régions du gène complètes en 5' et 3' restaure une copie du gène complet dans le fragment hybride en 5' ou 3'. Or, seul le vecteur XΔX porte un fragment interne, mais il est tronqué en 3' directement sur le codon d'arrêt TAA. L'insertion par recombinaison en 5' de la délétion doit interrompre le gène mais la fonctionnalité du gène dans le cas de l'insertion en 3' n'est pas claire. Une copie intacte de la région codante sera présente mais les séquences 3' non traduites seront remplacées par les séquences du vecteur. Il est possible que le transcrit de cet allèle soit instable ou mal traduit (nous le traitons à la page 67). L'insertion des trois autres vecteurs (AΔB, BΔB et AΔBneo) par recombinaison en amont ou en aval de la délétion va reconstituer une copie du gène *PpAPT* accompagnée de plusieurs copies du vecteur. Ces allèles pourront très probablement permettre une production normale de l'enzyme APRT et ne seront pas sélectionnés par le crible DAP<sup>R</sup>.

#### *Remplacement multiple*

Un événement de remplacement multiple remplace l'allèle sauvage par une forme concaténée de plusieurs copies du vecteur ou du fragment de remplacement (remplacement multiple). Nous pouvons anticiper le phénotype qui sera associé aux allèles AΔB-RM, AΔBneo-RM, BΔB-RM et XΔX-RM (Figure 18, RM pour remplacement multiple).

---

\* Le nom des allèles mutés n'est qu'un raccourci que nous utiliserons dans ce chapitre (par exemple, nous devrions désigner l'allèle XΔX-R par *ppapt-1*, XΔX-MR par *ppapt-2*, etc.)

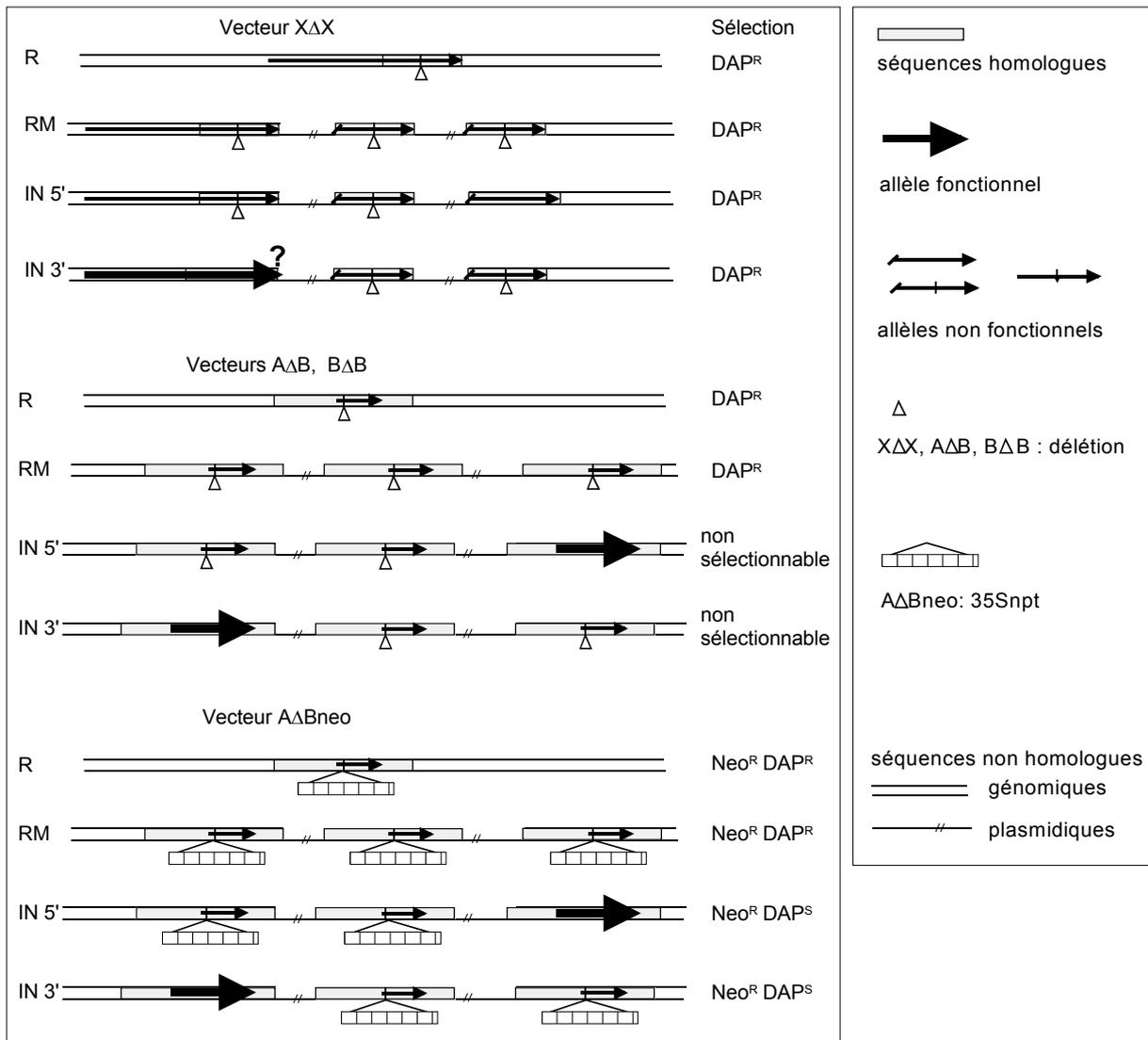


Figure 18. Différents types d'intégration par recombinaison homologue et la fonctionnalité hypothétique des allèles mutés.

Les fragments de remplacement peuvent être intégrés sans le vecteur de clonage (ligne fine);

R remplacement simple, RM remplacement multiple, IN 5' ou IN 3' insertion en 5' ou 3'. L'échelle n'est pas strictement respectée. Le schéma ne tient pas compte de la différence entre la longueur du fragment de remplacement des vecteurs AΔB, AΔBneo et BΔB.

### 2.3. Isolation des transformants intégratifs

Nous avons transformé la mouche *P. patens* avec les vecteurs schématisés dans la Figure 17, circulaires ou linéarisés. Le taux de survie a été entre 6 et 20% dans toutes les transformations ce qui atteste d'une qualité normale du traitement de transformation. Nous avons sélectionné les colonies résistantes à DAP (200  $\mu\text{M}$ ) dans les protoplastes régénérants transformés par XΔX, AΔB et BΔB. Nous avons sélectionné les colonies résistantes à la néomycine (G-418 à 25  $\text{mg L}^{-1}$ ) dans les protoplastes régénérants transformés par AΔBneo. Ces colonies Neo<sup>R</sup> ont été testées pour la résistance à DAP et environ  $\frac{2}{3}$  ont été DAP<sup>R</sup> (23 sur 30).

Nous caractérisons l'isolation des transformants intégratifs par la valeur RTF calculée par division du nombre des colonies résistantes dans le crible initial (DAP ou G-418) par le nombre de protoplastes régénérés (Figure 19).

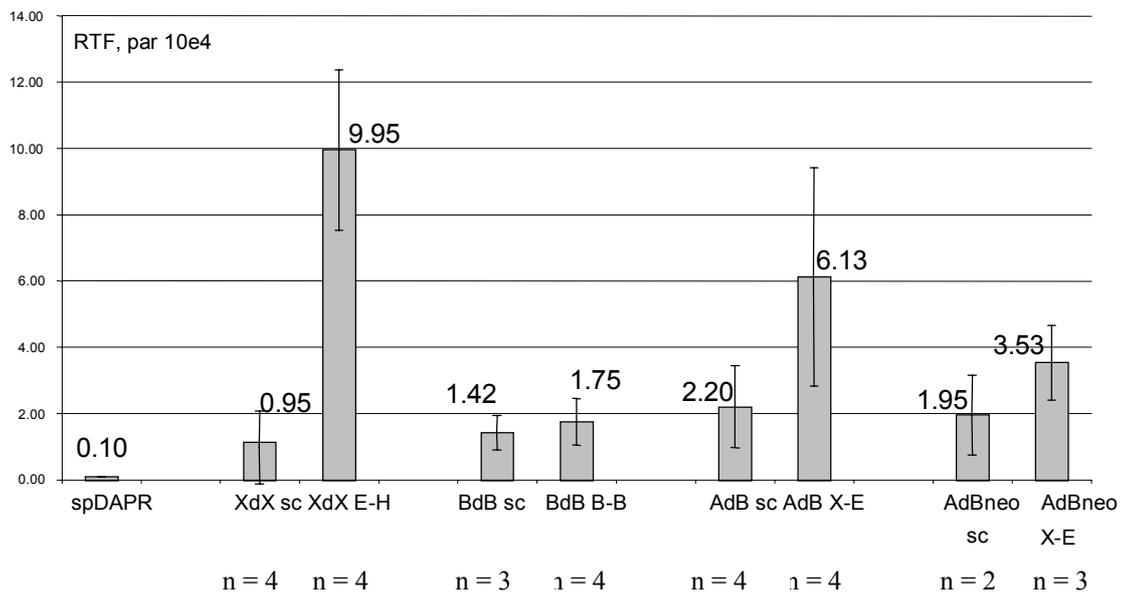


Figure 19. La valeur RTF exprime la fréquence moyenne d'apparition de colonies DAP<sup>R</sup> (AΔBneo -Neo<sup>R</sup>).

spDAPR - fréquence d'apparition de colonies DAP<sup>R</sup> spontanées attendues à une fréquence de 1 pour 10<sup>4</sup> protoplastes régénérants

sc - circulaire surenroulé;

XΔX E-H - linéarisé avec EcoRI et HindIII;

BΔB B-B - linéarisé avec BglIII;

AΔB X-E - linéarisé avec EcoRI et XbaI;

AΔBneo X-E - linéarisé avec EcoRI et XbaI;

Les barres d'erreur indiquent l'écartype. Le nombre d'expériences par vecteur est indiqué par n. Le nombre total de protoplastes régénérés par condition et par expérience est de 2 à 10 · 10<sup>4</sup>. La transformation avec chacun des vecteurs a produit des colonies DAP<sup>R</sup> à une fréquence plus élevée que 0.1 · 10<sup>-4</sup> (§) ce qui indique que presque la totalité des colonies DAP<sup>R</sup> est due à l'inactivation du gène *PpAPT* et non à une résistance spontanée.

## 2.4. Analyse moléculaire d'un échantillon des transformants intégratifs

Nous avons choisi un échantillon aléatoire de clones DAP<sup>R</sup>, Neo<sup>R</sup> DAP<sup>S</sup> et Neo<sup>R</sup> DAP<sup>R</sup> pour l'analyse moléculaire afin de vérifier l'inactivation du gène *PpAPT* par intégration ciblée et afin de déterminer le type d'intégration. La connaissance de la séquence nous permet de faire une prédiction détaillée du RFLP des allèles attendus, elle se trouve dans l'Annexe C, pages 145 - 147. Nous avons hybridé l'ADN de ces clones digéré avec EcoRV, NcoI ou XbaI avec la sonde A35 spécifique au locus APRT. La sonde A35 est le fragment BssHI-SmaI de 3.5 kb correspondant à quelque bp près au fragment XbaI - EcoRI qui se trouve dans vecteur AΔB. Nous avons également utilisé deux sondes spécifiques aux fragments hybrides en amont ou en aval du site d'intégration du vecteur XΔX pour analyser une partie des souches XΔX (données

§ La fréquence 0.1 · 10<sup>-4</sup> (cf. par exemple Figure 11, page 31 ou (Schaefer, 1994; Schaefer and Zryd, 1997)) est également obtenue avec des vecteurs non homologues. Une valeur RTF plus élevée, comme dans notre cas, confirme que l'intégration a lieu par recombinaison homologe.

non présentées, car ces sondes permettent les mêmes analyses que la sonde A35 sans détecter les répétitions du vecteur intégré).

Nous ne savons pas par quel mécanisme exact l'allèle WT est remplacé par un allèle portant une délétion ( $\Delta$ XbaI de 0.15 kb ou  $\Delta$ BamHI de 1.1 kb). Si le remplacement a lieu par conversion de gènes sans crossing-over, nous prédisons que la région remplacée est détruite lors du processus. Si le remplacement a lieu par un échange réciproque par un double crossing-over, la région remplacée se retrouve dans le vecteur de transformation. Le fragment portant cette région (dans notre cas XbaI de 0.15 kb et BamHI de 1.1 kb) s'intègre-t-elle ailleurs dans le génome? Pour répondre à cette question, nous avons comparé l'hybridation de l'ADN avec la sonde A35 avec l'hybridation du même ADN avec le fragment BamHI de 1.1 kb (sonde  $\Delta$ B) qui correspond à la région  $\Delta$ BamHI (Figure 17, page 48).

Les données suivantes ne tiennent pas compte des clones qui ont été sélectionnés comme DAP<sup>R</sup> mais qui se sont avérés des mutants spontanés ou chimériques DAP<sup>R</sup> + DAP<sup>S</sup> (3 clones pour 30 analysés), nous estimons leur proportion à environ 10%. Les mutants chimériques présentent un mélange des protoplastes mutants spontanés ou même de type sauvage qui se développent dans une colonie des protoplastes ciblés (souvent dans les couches supérieures de la colonies évitant par-là le contact avec DAP). Nous ajusterons les valeurs RTF présentées dans la Figure 19 en multipliant par un facteur de 0.9 pour les transformants DAP<sup>R</sup>. Elles seront commentées dans la Discussion. En revanche, tous les clones produits avec le vecteur  $\Delta$ Bneo sélectionnés comme Neo<sup>R</sup> sont de vrais transformants intégratifs. Cette différence est due au fait que l'allèle Neo<sup>R</sup> est dominant et que les tests de sélection de clones intégratifs en présence de la néomycine sont beaucoup plus stricts.

#### Types d'intégration du vecteur X $\Delta$ X

Nous avons pu observer les quatre types d'intégration attendus dans 14 clones sur 16 analysés (nous présentons un échantillon dans la Figure 20).

Les événements d'intégration sont les suivants:

R	RM	IN5	IN3	<i>cas complexes</i>
xa102	xeh101	xeh4 (1 copie)	xa101	xeh103
xeh6	xeh9	xeh5	xa103	xeh104
xeh2m	xeh2d		xeh105	
xa201, xa202, xa203				
<b>6</b>	<b>3</b>	<b>2</b>	<b>3</b>	<b>2</b>

*Nous avons obtenu les souches xa avec le vecteur circulaire et les souches xeh avec le vecteur linéarisé; les souches 201, 202, 203 ont été obtenues dans une expérience dans laquelle la forme de l'ADN n'a pas été rigoureusement contrôlée.*

*Cas complexes:* nous ne présentons pas le détail de l'analyse de l'intégration ciblée dans deux clones DAP<sup>R</sup> xeh103 et xeh104. Les fragments EcoRV du gène *PpAPT* (4.0 et 4.3) ont été remplacés par des fragments hybrides reconnus par la sonde A35 (*xeh103*: 6.5 et 7 kb; *xeh104*: 2, 2.5, 3.3, 4, 5 et 9 kb - aucun n'est reconnu par la sonde  $\Delta$ B). Ces fragments ne peuvent être expliqués précisément par aucun des événements (R, RM, IN5 ou IN3 du fragment ou du vecteur de remplacement intacts). Nous pensons que des concaténats chimériques des fragments partiellement digérés du vecteur X $\Delta$ X linéaire se sont formés par ligation intermoléculaire et se sont intégrés par un événement de remplacement complexe dans le gène *PpAPT*.

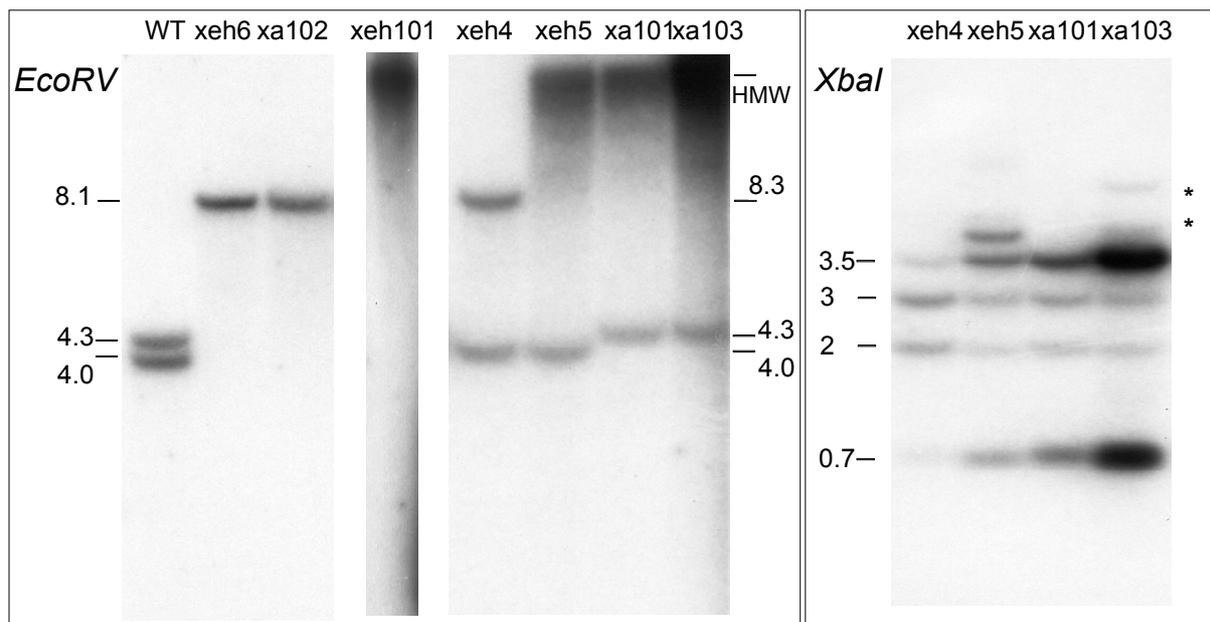


Figure 20. Démonstration de l'intégration ciblée dans le gène PpAPT au moyen d'une hybridation de l'ADN digéré avec EcoRV ou XbaI avec la sonde A35.

**EcoRV** - Les fragments du type sauvage (WT) sont de 4.3 kb (5') et 4.0 kb (3').

Souches xa102 et xeh6 - Le remplacement des fragments WT par un fragment de 8.1 kb indique que l'intégration a lieu par remplacement simple.

Souche xeh101 - Le remplacement des fragments WT par un fragment de haut poids moléculaire (HMW) indique que l'intégration a lieu par remplacement multiple.

Souches xeh4 et xeh5 - L'insertion en 5' remplace le fragment WT de 4.3 kb par un fragment hybride de 8.3 kb (1 copie) ou de haut poids moléculaire (HMW, plus d'une copie) laissant intact le fragment WT de 4.0 kb .

Souches xa101, xa103 - L'insertion en 3' remplace le fragment WT de 4.0 kb par un fragment hybride de HMW laissant intact le fragment WT de 4.3 kb.

**XbaI** - Les fragments XbaI sont de 2 et 3 kb indépendamment du type d'intégration. Cette analyse permet, en revanche, de vérifier le nombre et la structure des copies du vecteur intégré.

Souches xeh4, xeh5, xa101 et xa103 - La présence de fragments XbaI de 0.7 et 3.5 kb démontre que les copies multiples du vecteur entier sont intégrées en répétitions directes (des fragments de 0.5 ou 1.0 kb dévoilerait une intégration de copies du fragment de remplacement, répétitions directes ou inversées). Nous détectons des fragments partiellement digérés de 4.2 kb et plus (\* vu la différente intensité de ces signaux). Le nombre de copies du vecteur est: xeh4 - 1 copie; xeh5 - 5 à 10 copies; xa101 - 10 à 15 copies; xa103 - 15 à 20 copies.

#### Types d'intégration du vecteur BAB

Nous avons observé les deux types d'intégration attendus pour ce vecteur (remplacement simple et multiple) en analysant 9 clones (nous présentons un échantillon dans la Figure 21). Nous n'avons pas analysé en détail l'orientation des copies intégrées en répétition multiple.

Les événements d'intégration sont les suivants:

<b>R</b>	<b>RM</b>
bb102, bb104, bb105, bb2, bb3	ba 101, ba102, bb101, b103
<b>5</b>	<b>4</b>

Nous avons obtenu les souches ba avec le vecteur circulaire et les souches bb avec le vecteur linéarisé.

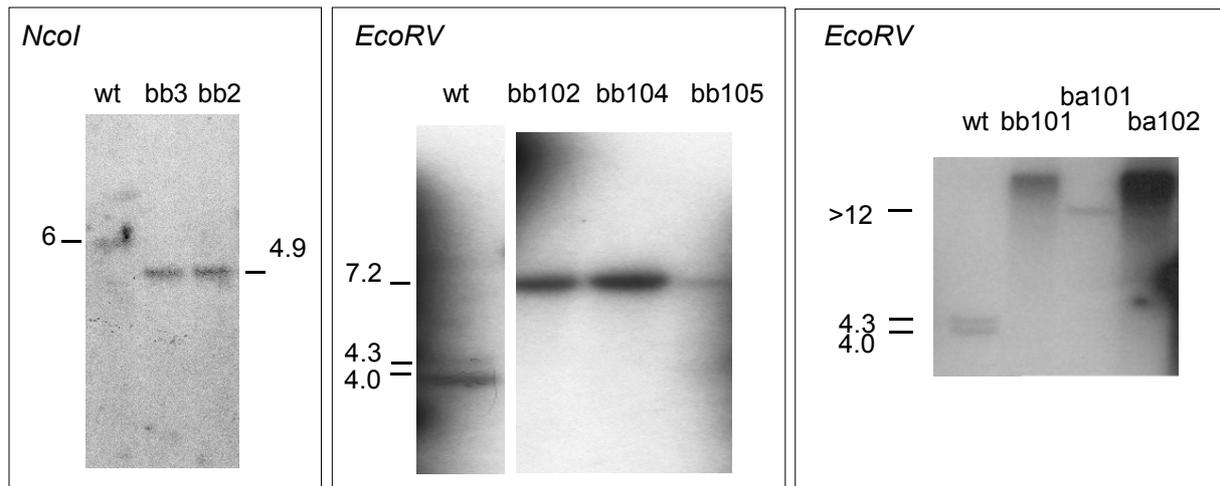


Figure 21. Démonstration de l'intégration ciblée dans le gène PpAPT au moyen d'une hybridation de l'ADN digéré avec *EcoRV* ou *NcoI* avec la sonde A35.

***NcoI*** - Le fragment du type sauvage est de 6 kb.

Souches bb2 et bb3 - le remplacement du fragment WT *NcoI* de 6 kb par un fragment hybride de 4.9 kb indique que l'intégration a lieu par remplacement simple.

***EcoRV*** - Les fragments du type sauvage sont de 4.3 et 4.0 kb.

Souches bb102, bb104 et bb105 - Le remplacement des fragments WT de 4.3 et 4.0 kb par un fragment hybride de 7.2 kb indique que l'intégration a lieu par remplacement simple.

Souches bb101, ba101 et ba102 - Le remplacement des fragments WT de 4.3 et 4.0 kb par un fragment de plus de 12 kb indique que l'intégration a lieu par remplacement multiple.

La sonde  $\Delta B$  reconnaîtrait les fragments hybrides non prédits qui seraient produits par une intégration de la région de la délétion *BamHI* de 1.1 kb. De tels fragments n'ont été détectés dans aucun des événements d'intégration par recombinaison homologue (non présenté) ce qui démontre que le fragment *BamHI* est définitivement supprimé.

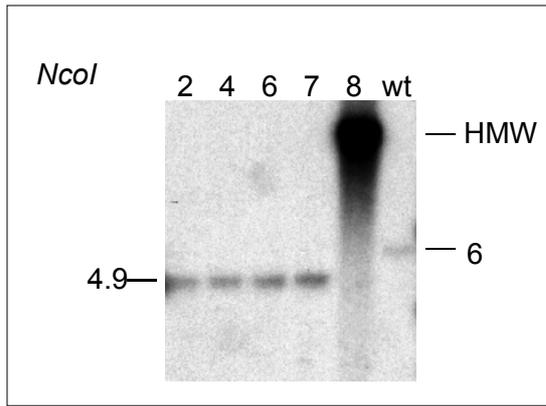
#### Types d'intégration du vecteur *AAB*

Nous avons observé les deux types d'intégration attendus pour ce vecteur (remplacement simple et multiple) en analysant 8 clones (nous présentons un échantillon dans la Figure 22). Nous n'avons pas analysé en détail l'orientation des copies intégrées en répétition multiple.

Les événements d'intégration sont les suivants:

<b>R</b>	<b>RM</b>
aex103, aex2, aex4, aex5, aex6, aex7	aex1, aex8
<b>6</b>	<b>2</b>

Nous n'avons pas analysé les souches obtenues avec le vecteur circulaire; nous avons obtenu les souches aex avec le vecteur linéarisé.



*NcoI* - Le fragment WT est de 6 kb.  
 Souches aex2 à aex7 - Le remplacement du fragment WT de 6 kb par un fragment hybride de 4.9 kb indique que l'intégration a lieu par remplacement simple;  
 Souche aex8 - Le remplacement du fragment WT de 6 kb par un fragment de haut poids moléculaire (HMW) indique que l'intégration a lieu par remplacement multiple. Aucun signal dans l'hybridation de cet ADN avec la sonde  $\Delta B$  démontre que ces clones n'ont aucune copie du fragment BamHI de 1.1 kb (non présenté).

Figure 22. Démonstration de l'intégration ciblée dans le gène PpAPT au moyen d'une hybridation de l'ADN digéré avec *NcoI* avec la sonde A35.

La sonde  $\Delta B$  reconnaît les fragments hybrides non prédits qui seraient produits par une intégration de la région de la délétion BamHI de 1.1 kb. De tels fragments n'ont été détecté dans aucun des événements d'intégration par recombinaison homologue.

#### Types d'intégration du vecteur $\Delta Bneo$

Nous avons pu observé les quatre types d'intégration attendus en analysant 9 clones. Les événements d'intégration du vecteur sont les suivants:

R	RM	IN5	IN3	RM et correction du vecteur
<b>Neo<sup>R</sup> DAP<sup>R</sup></b>	<b>Neo<sup>R</sup> DAP<sup>R</sup></b>	<b>Neo<sup>R</sup> DAP<sup>S</sup></b>	<b>Neo<sup>R</sup> DAP<sup>S</sup></b>	<b>Neo<sup>R</sup> DAP<sup>S</sup></b>
ana 101, anex101, anex103	anex106, anex105	anas106, anas108	anas107	anexs111
<b>3</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>1</b>

Nous avons obtenu les souches ana / anas avec le vecteur circulaire et les souches anex / anexs avec le vecteur linéarisé. La lettre **S** indique la sensibilité à DAP.

Nous présentons séparément l'analyse des souches Neo<sup>R</sup> DAP<sup>S</sup> et des souches Neo<sup>R</sup> DAP<sup>R</sup>.

*L'intégration par remplacement génère les souches DAP<sup>R</sup> Neo<sup>R</sup>*

L'analyse confirme notre hypothèse que toutes les 5 souches Neo<sup>R</sup> DAP<sup>R</sup> proviennent d'un événement de remplacement (simple ou multiple, Figure 23).

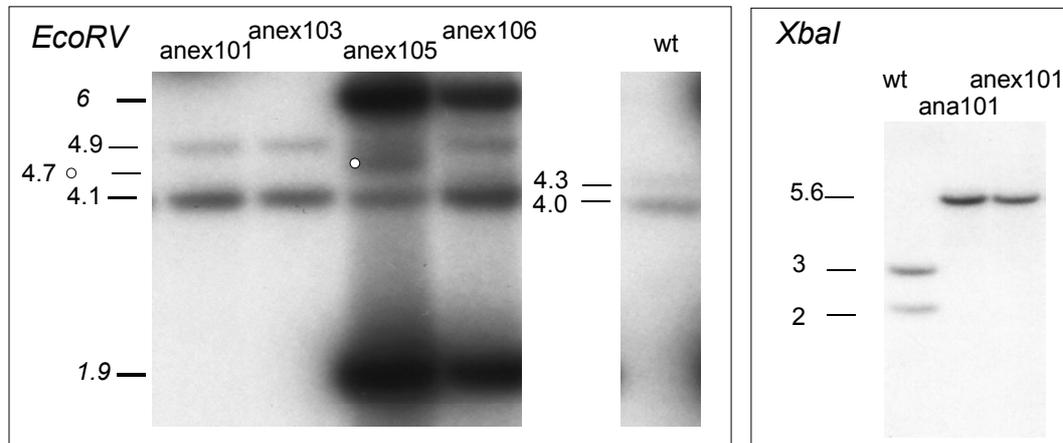


Figure 23. Démonstration de l'intégration ciblée dans le gène PpAPT au moyen d'une hybridation de l'ADN digéré avec EcoRV ou NcoI avec la sonde A35.

**EcoRV** - Les fragments du type sauvage sont de 4.3 et de 4.0 kb.

Souches anex 101, anex 103 - Le remplacement des fragments WT de 4.3 et de 4.0 kb par des fragments hybrides de 4.1 et 4.9 kb indique que l'intégration a lieu par remplacement simple.

Souches anex105 et anex106 - Le remplacement des fragments WT de 4.3 et de 4.0 kb par deux fragments hybrides de 4.1 et 4.9 kb accompagné des fragments EcoRV du vecteur de 2 et 6 kb indique que l'intégration a lieu par remplacement multiple. Plusieurs copies du vecteur s'intègrent en répétitions directes (anex105 - 20 à 30 copies; anex106 - 10 à 20 copies). Nous expliquons le fragment EcoRV de 4.7kb (marqué ° dans la souche anex105) par l'intégration de copies parallèles du fragment remplacement sans le reste du plasmide avec des copies du vecteur entier. Une figure détaillée se trouve dans l'Annexe C, page 147.

**XbaI** - Les fragments du type sauvage sont de 2 et 3 kb.

Souches ana101 et anex101 - Le remplacement des fragments WT de 2 et 3 kb par un fragment hybride de 5.6 kb indique que le remplacement a lieu par remplacement simple.

*L'intégration par insertion génère les souches Neo<sup>R</sup> DAP<sup>S</sup>*

Les fragments spécifiques de ces événements sont particulièrement évidents dans l'hybridation de l'ADN digéré avec EcoRV avec la sonde ΔB. Cette analyse confirme également que dans ces clones, la région remplacée ΔB n'est pas conservée et ne s'est pas intégrée ailleurs dans le génome.

Nous confirmons ainsi notre hypothèse que toutes les 4 souches Neo<sup>R</sup> DAP<sup>S</sup> proviennent d'un événement d'insertion ciblée (Figure 24).

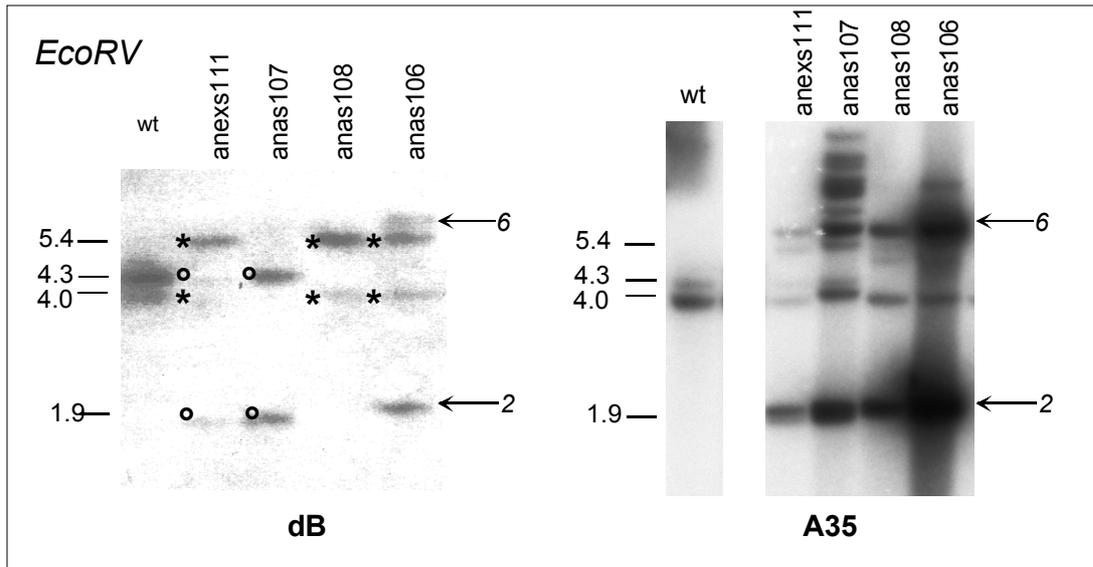


Figure 24. Démonstration de l'intégration ciblée dans le gène PpAPT au moyen d'une hybridation de l'ADN digéré avec EcoRV avec les sondes ΔB et A35.

**Sonde ΔB** - Les fragments WT reconnus sont de 4.3 kb (5') et de 4.0 kb (3'). Les signaux marqués de \* sont spécifiques d'une insertion 5' et ceux marqués de ° - d'insertion 3'.

Souche anas107 - Le remplacement du fragment WT 4.0 kb par un fragment de 1.9 kb indique que l'intégration a lieu par insertion en 3'.

Souches anas106 et anas108 - Le remplacement du fragment WT de 4.3 kb par un fragment de 5.4 kb indique que l'intégration a lieu par insertion 5'. Le clone anas106 possède un nombre très important de copies du vecteur intégré, bien plus que 20 (cf. A35). Dans ces conditions, les traces du plasmide pUC21 dans la sonde ΔB s'hybrident avec les fragments de 6 et 2 kb de ces répétitions (flèches).

Souche anexas111 - La présence des fragments WT et des deux jonctions 5' et 3' prédites indique que l'intégration est accompagnée de la correction de la copie du vecteur dans laquelle le crossing-over a lieu. Le paragraphe suivant le traite plus en détail.

**Sonde A35** - Le nombre de copies a été déterminé par hybridation de cet ADN avec la sonde A35.

anexas111 - 2 ou 3 copies; anas107 - 5 à 10 copies (digestion partielle); anas108 - 5 copies; anas106 - plus que 20 copies. Les signaux au-dessus de 6 kb sont probablement des fragments partiellement digérés. La présence de fragments EcoRV de 2 et 6 kb dans tous les clones indique que les copies multiples du vecteur entier se sont intégrées en répétitions directes. Cette conclusion doit encore être vérifiée car les répétitions inversées reconnues par la sonde A35 auraient pu être masquées (fragments EcoRV de 4 et de 12 kb attendus pour le vecteur, de 4 et de 5.4 kb pour le fragment de remplacement).

Un événement de correction du vecteur accompagne l'intégration dans la souche *anexs111*.

La souche *anexs111* est  $\text{Neo}^R \text{DAP}^S$ . Les fragments  $\text{EcoRV}$  WT § et les fragments hybrides pour l'intégration par insertion en 5' **et** en 3' (Figure 24) sont détectables. Ces fragments hybrides ne peuvent pas provenir d'une intégration illégitime car la sonde  $\Delta B$  ne reconnaît aucun fragment du vecteur. Pour expliquer la présence des deux fragments hybrides, nous proposons la structure de l'allèle *anexs111* basée sur la correction du vecteur avant l'intégration.

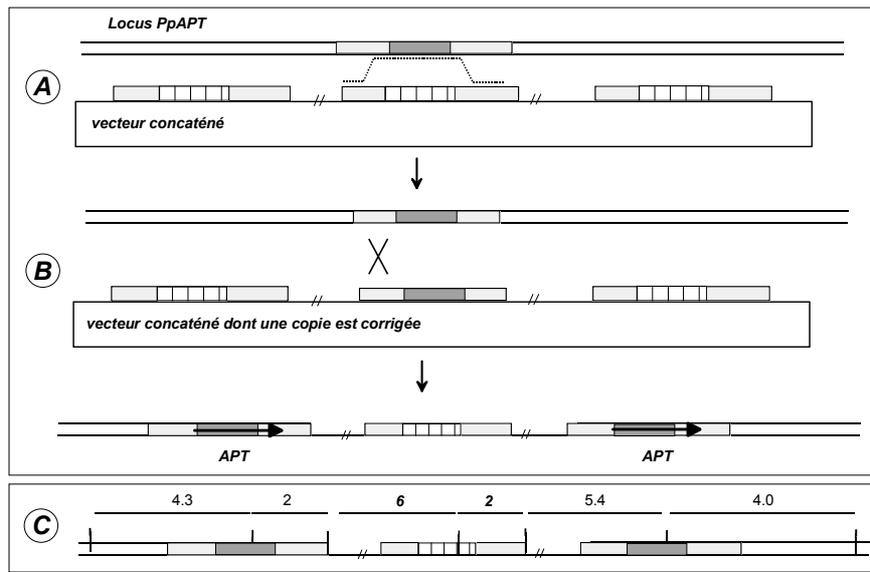


Figure 25. Explication de l'allèle *anexs111* et son RFLP  $\text{EcoRV}$ .

A - Le concaténat du vecteur interagit avec le site d'intégration. Une copie du vecteur est corrigée selon la matrice génomique.

B - L'intégration du vecteur corrigé produit deux copies du fragment entier du gène *PpAPT*.

C - Le RFLP de la souche *anexs111*.

Nous avons proposé un mécanisme similaire pour la réparation des séquences interrompues dans le chapitre II (pour le locus  $\lambda 213$ ). Nous postulons que la réparation a lieu préférentiellement dans les copies qui sont impliquées dans la recombinaison et, par conséquent, forment les jonctions en amont et en aval du site d'intégration.

§ Notez que la sonde  $\Delta B$  reconnaît comme prévu les fragments WT mais leur intensité est trop faible; le fragment de 4.0 kb est particulièrement mal visible car il s'hybride avec la sonde sur 0.3 kb seulement ( le signal du fragment de 2.0 kb est moins fort que celui du fragment de 6.0 kb aussi). Nous mettons mieux en évidence ces fragments avec la sonde A35.

### **3. Analyse du phénotype des différents allèles du gène *PpAPT* obtenus par mutagenèse ciblée**

#### **3.1. Morphologie et développement des mutants DAP<sup>R</sup> en conditions standards**

Cove et Schaefer donnent une description détaillée de la souche sauvage dans plusieurs publications (Cove, 1992; Cove *et al.*, 1991; Schaefer, 1994). Dans ces conditions standard (milieu PPNO3), l'observation du développement des clones DAP<sup>R</sup> révèle l'existence de deux classes de mutants que nous désignons par DAP<sup>R</sup>-I et DAP<sup>R</sup>-II dans ce chapitre. La morphologie et le développement des mutants DAP<sup>R</sup>-I est très différente de ceux de la souche sauvage (WT); en revanche, aucune différence entre les mutants DAP<sup>R</sup>-II et WT n'est discernable. Dans la Figure 26, nous résumons les différences entre le phénotype de la souche sauvage et celui d'un mutant DAP<sup>R</sup>-I (xeh6, page 53). Nos observations de la souche sauvage et des mutants DAP<sup>R</sup>-II sont conformes à la description standard (références ci-dessus). L'altération phénotypique des mutants DAP<sup>R</sup>-I se manifeste essentiellement par un développement inachevé des gamétophores (détail dans la Figure 26). Génétiquement, les mutants DAP<sup>R</sup>-I et DAP<sup>R</sup>-II sont différents par l'allèle du gène *PpAPT*; la classe DAP<sup>R</sup>-II est composée des mutants portant l'allèle XΔX-IN3 alors que la classe DAP<sup>R</sup>-I comprend tous les autres mutants DAP<sup>R</sup>. Nous le traiterons dans les paragraphes suivants (pages 62 et 67).

## Figure de la page suivante

**A. WT et Xeh6** - Le protonéma, partie filamenteuse de la colonie sur le milieu PPNO3 (le trait représente 20  $\mu\text{m}$ ).

Sur la photographie, le protonéma de la souche sauvage et celui du mutant DAP<sup>R</sup>-I est normal. Sur le milieu PPNO3, leur production est identique (sauf voir le point D).

**B.** - Aspect général de la colonie WT et xeh6 après 21 jours de croissance sur le milieu PPNO3 (le trait représente 1 mm).

**WT** - Les bourgeons se développent en gamétophores. Le gamétophore est une structure caulinaire (3D) "feuillée", d'une anatomie simple, sans vascularisation.

**Xeh6** - Les bourgeons se développent en gamétophores qui ne sont pas complètement formés. Les feuilles (cf. C) ne sont pas achevées, la base des feuilles a une forme de cal. Nous n'avons pas pu induire la formation de sporophytes sur les colonies xeh6 .

**C.** - Détail des feuilles sauvages et mutantes, milieu PPNO3.

**WT** (le trait représente 0.5 mm) - Les feuilles ont une structure très simple, d'une couche monocellulaire, sans vascularisation. Le pli central est bien marqué. Les feuilles se trouvent sur un gamétophore bien formé.

**Xeh6** (le trait représente 1 mm) - Les feuilles se développent mais n'ont pas la même forme que celles du type sauvage: la base a tendance à former un cal, la pointe et le pli central sont peu prononcés voire absents, la forme est arrondie et enroulée. L'ensemble des feuilles ne forment pas de gamétophore.

**D.** Croissance du protonéma d'un mois sur le milieu PPNO3 additionné de glucose. L'échelle est similaire à celle indiquée sous C.

**WT** - Environ 3-5% des initiales latérales de croissance se développent en bourgeon qui donne naissance à un gamétophore.

**Xeh6** - Pratiquement 90-100% des initiales de croissance latérales se différencient en "bourgeons" mais ne produisent jamais de gamétophores.

**E.** Développement de colonies portant des gamétophores après environ 1 mois de croissance sur le milieu PPNH4c (PPNO3 additionné d'une source d'azote  $\text{NH}_4^+$  et de glucose, cf. page 134). L'échelle est similaire à celle indiquée sous B.

**WT** - Les gamétophores sont normaux.

**Xeh6** - Aucun gamétophore ne se forme, les feuilles sont pratiquement absentes. Les bourgeons, plus nombreux (cf. sous D.), donnent naissance à des cals.

Le brunissement des colonies après une longue période de croissance est observé dans toutes les cultures mais il est particulièrement précoce chez xeh6, après 2-3 semaines sans changement de milieu contre 1-2 mois pour le WT (non montré). Le brunissement peut être dû à la mort cellulaire ou à une pigmentation. En effet, le nombre de cellules mortes est plus élevé dans les colonies sur PPNH4 que dans celles sur PPNO3. Nous ne connaissons pas les substances responsables de cette pigmentation (voir également page 64).

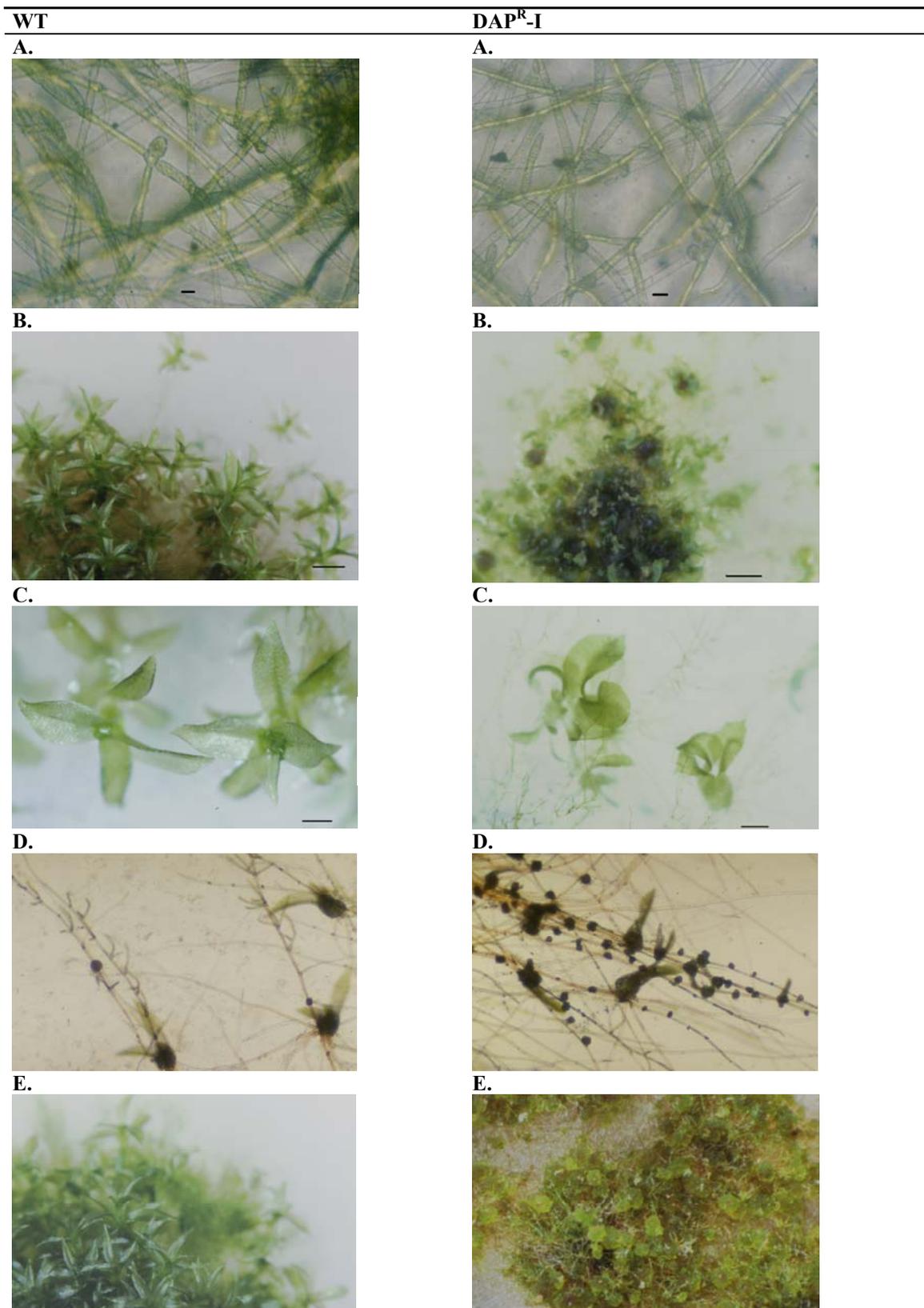


Figure 26, A-E. Morphologie et développement de la souche sauvage (WT) et d'un mutant DAP<sup>R</sup>-I (xeh6).

### 3.2. Explications possibles des altérations du phénotype des mutants DAP<sup>R</sup>-I

L'enzyme APRT codée par les gènes *APT* est impliquée dans le recyclage des purines (Figure 27). Dans les paragraphes suivants, nous tenterons d'expliquer comment la mutation du gène *PpAPT* provoque le phénotype observé chez *P. patens*.

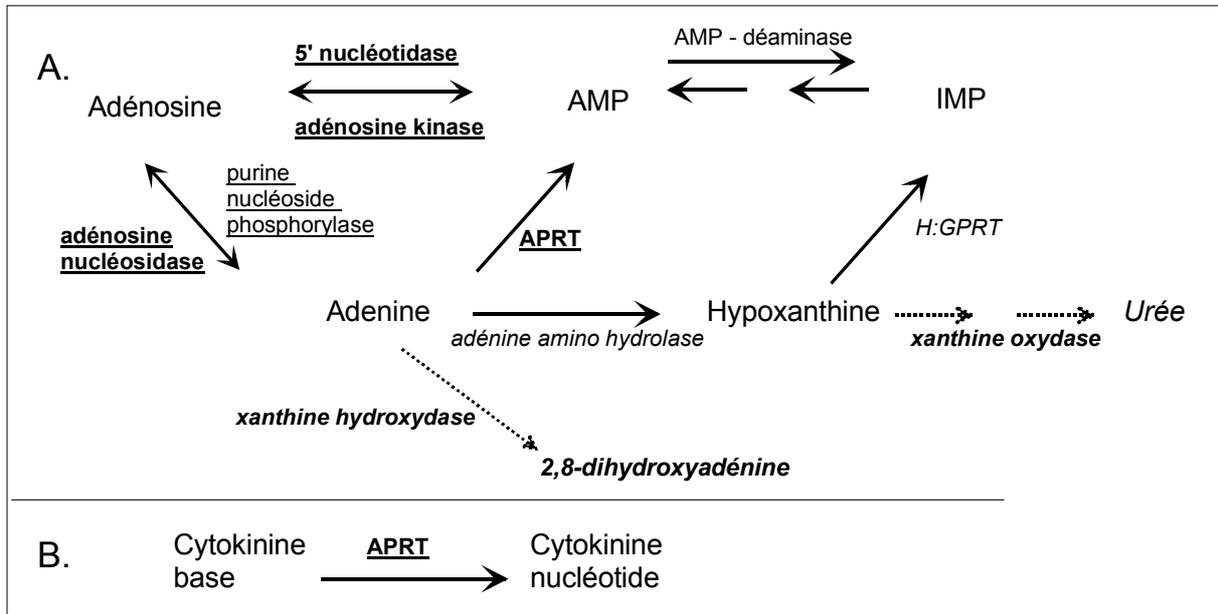


Figure 27. Métabolisme de l'adénine et activité APRT dans différents organismes (Lehninger *et al.*, 1993; Moffatt *et al.*, 1994; Schnorr *et al.*, 1996; von Schwartzberg *et al.*, 1998; Woods *et al.*, 1984).

En caractères soulignés: les enzymes impliquées dans le recyclage de l'adénine chez les angiospermes; **en caractères gras**: les enzymes également actives dans d'autres organismes; *en caractères italiques fins*: la voie de l'adénine amino hydrolase, active chez *S. cerevisiae*. Les flèches en pointillés expliquent le syndrome de déficience d'APRT chez l'homme.

**A.** - L'APRT appartient à un groupe d'enzymes phosphoribosyltransférases (PRT). Toutes les enzymes du groupe PRT catalysent le transfert du ribose 5'-phosphate sur la base azotée pour former un N-riboside-monophosphate. Le groupe phosphate est fourni par une molécule de 5'-phosphoribosylpyrophosphate (PRPP) qui se transforme en pyrophosphate. L'APRT catalyse la formation de 5'-adénosine phosphate (AMP) à partir d'adénine et de PRPP en présence d'ions de magnésium. Dans la majorité des eucaryotes et des procaryotes, c'est la voie majeure du recyclage de l'adénine. Une deuxième voie qui permet de convertir l'adénine en AMP passe par deux réactions séquentielles, catalysées par l'adénosine phosphorylase (ou purine nucléoside phosphorylase) et par l'adénosine kinase. Chez la levure *S. cerevisiae*, la conversion de l'adénine en AMP en absence d'activité APRT peut être catalysée par l'activité concertée de deux enzymes: adénine aminohydrolase (AAH, absente chez les mammifères ou *E. coli*) et hypoxanthine-guanine PRT (HPRT ou H:GPRT). IMP - inosine-5' monophosphate.

**B.** - Chez les plantes, l'activité APRT est probablement impliquée dans le métabolisme des cytokinines mais dans une mesure moindre que dans celui des purines. L'APRT végétale (codée par le gène *AtAPT2* chez *A. thaliana*) convertit les cytokinines bases en cytokinines nucléotides.

Chez les angiospermes, en plus de recycler les purines, l'APRT convertit les cytokinines bases en cytokinines nucléotides et par-là, peut jouer un rôle dans le métabolisme des cytokinines. L'activité APRT de *A. thaliana* est la mieux caractérisée (Gaillard *et al.*, 1998; Moffatt *et al.*, 1994; Moffatt *et al.*, 1991; Schnorr *et al.*, 1996). Elle est codée par trois gènes au moins, *AtAPT1*, 2 et 3. La comparaison des séquences génomiques indique que la position et pas la longueur des introns est conservée entre les deux premiers gènes. Le gène *AtAPT1* est majoritairement exprimé dans les feuilles alors que l'expression du gène *AtAPT2* est spécifique aux racines, aux tiges et aux fleurs mais absente dans les feuilles. L'expression du

gène *AtAPT2* est plus faible comparée à celle du gène *AtAPT1*. Les deux enzymes, APRT 1 et 2, respectivement, ont une affinité différente pour les substrats: l'adénine et la benzyladénine sont les substrats de l'enzyme APRT2 et seule l'adénine est le substrat de l'enzyme APRT1 d'après les valeurs constantes de Michaelis et de la vitesse maximale des réactions correspondantes. De même, plusieurs formes d'APRT ont également été trouvées chez le tabac et la tomate (Gaillard *et al.*, 1998) et probablement chez l'avoine (Itai *et al.*, 2000). Les auteurs de ces travaux supposent qu'au moins deux types de gènes *APT* sont généralement présents chez les plantes: l'un exprimé davantage dans les feuilles et l'autre dans les racines (Itai *et al.*, 2000). Leurs produits respectifs peuvent avoir des affinités différentes pour les purines et les cytokinines - une situation similaire aux enzymes *AtAPRT1* et 2 (Schnorr *et al.*, 1996). Le groupe de B. Moffat a isolé des mutants d'*A. thaliana* qui possèdent 1% de l'activité APRT du type sauvage. Ces mutants portent différents allèles du gène *AtAPT1* inactivés par des mutations ponctuelles. Les mutants absorbent la cytokinine N<sup>6</sup>-benzyladénine (BA) ajoutée au milieu de croissance mais ne la convertissent pas en N<sup>6</sup>-benzyladénosine 5'-monophosphate (BAMP) (Gaillard *et al.*, 1998; Moffatt *et al.*, 1994; Moffatt *et al.*, 1991; Schnorr *et al.*, 1996). Ces mutants d'*A. thaliana* présentent un phénotype anormal: la croissance est retardée, l'induction de cals est quasi impossible, et le développement du pollen n'est pas achevé. Ces études ne démontrent pas la cause de ces altérations, qui peut être la perturbation du métabolisme des cytokinines ou de celui des purines.

L'activité APRT de *P. patens* ne convertit pas les cytokinines bases en nucléotides (Schaefer *et al.*, en préparation). Il est surprenant de constater que le phénotype de la classe DAP<sup>R</sup>-I en culture sur le milieu PPNH4c ressemble en partie à celui des mutants surproduisant des gamétophores (*ove*) ou à celui de la souche sauvage sur du milieu enrichi en cytokinines (Ashton *et al.*, 1979a). Néanmoins, la teneur en cytokinines des mutants DAP<sup>R</sup>-I n'est pas différente de celle du WT. En outre, les expériences de culture croisée (cross-feeding) avec le WT n'a pas révélé de surproduction d'une forme active des cytokinines par la classe DAP<sup>R</sup>-I qui induirait la surproduction de bourgeons sur le WT.

Afin de vérifier que le phénotype DAP<sup>R</sup>-I est provoqué par l'altération du recyclage de l'adénine, nous avons testé l'effet sur le développement du protonéma de *P. patens*, de supplément de l'adénine dans le milieu de croissance. Nous avons observé le développement de la souche sauvage, d'une souche DAP<sup>R</sup>-I xeh6 (ou xa102) et d'une souche DAP<sup>R</sup>-II (donc DAP<sup>R</sup> et au phénotype normal) (Figure 28 - Figure 29).

L'adénine n'inhibe pas la croissance de la souche sauvage même en concentration de 1 mM (Figure 28, ligne supérieure marquée W). La présence d'adénine dans le milieu de culture n'affecte pas la morphologie de la souche sauvage jusqu'à une concentration de 500 µM. Même en présence d'une concentration en adénine de 1 mM, la croissance est quasi normale mais le développement du caulonéma et des bourgeons est plus lent.

L'adénine inhibe la croissance et la différenciation des gamétophores de la souche DAP<sup>R</sup>-I à partir d'une concentration de 60 µM (Figure 29). La formation de cals plutôt que de feuilles a lieu, et ceci proportionnellement à l'augmentation de la teneur en adénine. La toxicité de l'adénine est évidente à une concentration de 1.0 mM.

L'effet de l'adénine sur la souche DAP<sup>R</sup>-II ne se manifeste qu'à partir d'une concentration de 120 µM (Figure 29). Une concentration en adénine de 120 µM provoque chez cette souche des altérations similaires à celles observées chez les mutants DAP<sup>R</sup>-I en absence d'adénine. Des concentrations de 250, 500 et 1000 µM provoquent des altérations chez la souche DAP<sup>R</sup>-

II similaires à celles que nous observons chez les mutants DAP<sup>R</sup>-I en présence de concentrations de 60, 250 et 500  $\mu$ M, respectivement. Néanmoins, en présence d'une concentration de 500 $\mu$ M, les souches présentent déjà un phénotype similaire et la concentration de 1000 $\mu$ M est toxique pour les deux classes de mutants. Nous postulons que l'adénine provoque les mêmes altérations chez les mutants des deux classes mais une concentration plus élevée est nécessaire pour provoquer le même effet chez les mutants de la classe DAP<sup>R</sup>-II que chez les mutants de la classe DAP<sup>R</sup>-I.

Premièrement, nous formulons l'hypothèse que ces altérations du développement sont provoquées par l'accumulation de l'adénine qui n'est plus convertie en AMP par une activité APRT. La souche sauvage arrive à recycler l'adénine correctement même à des concentrations élevées ( $\approx$ 0.5 mM) en la convertissant en AMP par l'activité normale APRT. Le fait que de très hautes concentrations ( $>$ 1 mM) en adénine affectent même le type sauvage nous laisse penser qu'il existe un seuil de concentration à partir duquel l'adénine n'est plus recyclée correctement. Le fait que nous observons des altérations chez les mutants DAP<sup>R</sup>-I même en absence d'adénine dans le milieu nous indique que l'activité APRT est nulle chez ces mutants. Les résultats des analyses biochimiques (Tableau 5, WT et Xeh6, page 65, (Schaefer *et al.*, en préparation)) confirment cette hypothèse.

Deuxièmement, nous pensons que l'activité APRT des mutants DAP<sup>R</sup>-II est présente mais qu'elle est plus faible que celle de la souche sauvage. Elle est alors plus vite saturée par le surplus d'adénine ajoutée au milieu de croissance. L'analyse biochimique sur ces mutants est en cours. Afin de confirmer la différence supposée entre les niveaux d'activité APRT des mutants I et II, ces essais doivent être faits en présence de hautes concentrations (saturantes) d'adénine dans tous ces clones.

Nous ne savons pas quel est le mécanisme exact de l'influence de l'adénine non recyclée sur *P. patens*. Il est notable que chez l'homme, l'inactivation du gène APRT conduit à l'accumulation de 2,8-dihydroxyadénine, responsable de la formation de calculs rénaux. Ce composé est le produit de l'action de la xanthine oxidase sur l'adénine en absence de l'activité APRT (Figure 27) (Isselbacher *et al.*, 1994). Nous pensons qu'une accumulation de dérivés de l'adénine (peut-être insolubles et colorées (?)) comme dans le cas de la 2,8-dihydroxyadénine) provoque les altérations observées chez *P. patens*.

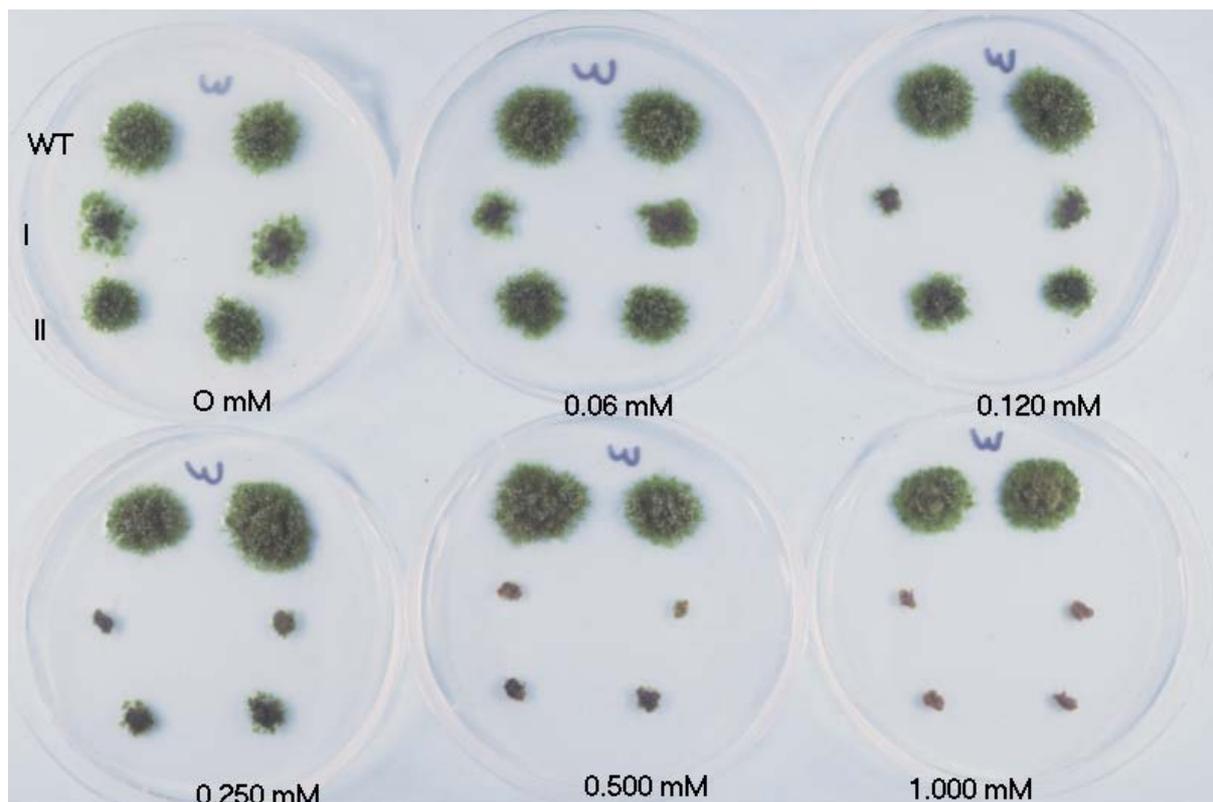


Figure 28. Croissance de la souche sauvage (W - haut), d'un mutant DAP<sup>R</sup>-I (I - milieu) et DAP<sup>R</sup>-II (II - bas) sur le milieu PPNO3 additionné d'adénine à des concentrations croissantes. L'effet observé est traité dans le texte.

Tableau 5. Métabolisme de l'adénine dans le protonéma de la souche sauvage et de différents mutants ((Schaefer *et al.*, en préparation))

Clone	Incorporation (%), §			Radioactivité dans les extraits, Bq	% de la radioactivité dans les métabolites		
	1h	2h	3h		Adénine	Adénylates	Autres
WT	26.9	40	47.5	48450	2.3	77.8	19.8
DAP <sup>R</sup> -I	3.9	5.3	6.3	4083	95	nd	5
<i>Xeh6</i>							
DAP <sup>R</sup> -II (en cours)							
<i>Apt1-C4</i>	35	64	68.4	62030	2.2	77.8	20
<i>Apt1-C5</i>	30.6	56	62.2	65250	5	83.2	11.7
<i>Apt2-C2</i>	24.2	43	52.6	58590	3.6	81.5	14.9
<i>Apt2-C4</i>	32.5	48.4	55	50340	9	71.2	19.7

§ - L'adénine marquée au tritium, d'une radioactivité connue (Ro), est ajoutée à une suspension aqueuse ( $\approx 80$  mg poids frais  $\text{ml}^{-1}$ ) de protonéma fraîchement broyé. La radioactivité (Rx) est mesurée dans 10 ml de suspension après 1, 2 et 3 heures d'incubation. Le rapport (Rx-Ro)/Ro, en %, sert de mesure de l'incorporation de l'adénine. La concentration de l'adénine marquée est de 19 nM, cette condition permet de détecter les activités APRT extrêmement faibles. Par conséquent, la différence entre l'activité APRT de la classe DAP<sup>R</sup>-I et celle de la classe DAP<sup>R</sup>-II n'est pas mise en évidence par cet essai. Un essai utilisant de concentrations saturantes de l'adénine est nécessaire afin de démontrer cette différence (en cours).

Les métabolites sont analysés par la chromatographie en phase liquide à haute pression (HPLC) dans les extraits du protonéma réalisés au moyen de l'acide perchlorique (Moffatt *et al.*, 1991). nd - non détectable. Les adénylates sont l'AMP, l'ADP et l'ATP.

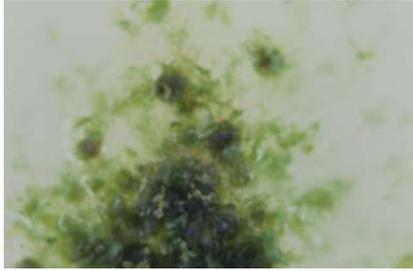
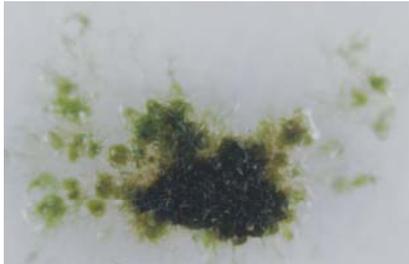
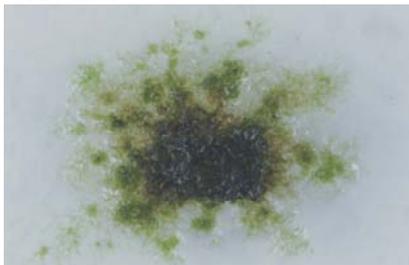
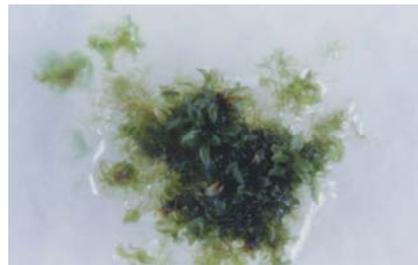
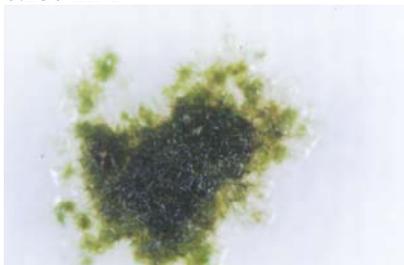
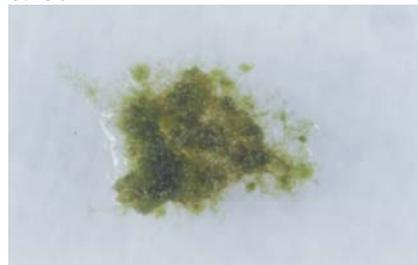
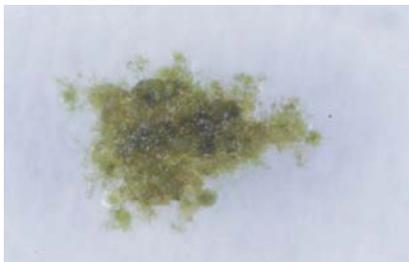
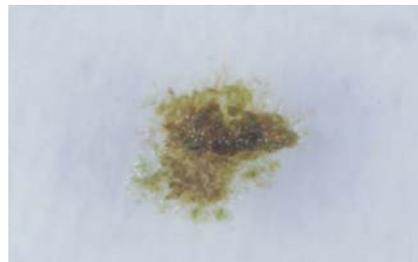
**DAP<sup>R</sup>-I****DAP<sup>R</sup>-II****Sans adénine****Sans adénine****0.060 mM****0.060 mM****0.120 mM****0.120 mM****0.250 mM****0.250 mM****0.500 mM****0.500 mM**

Figure 29. Croissance d'une souche DAP<sup>R</sup>-I (xeh6 ou xa102) et d'une souche DAP<sup>R</sup>-II (xa101 ou xa103) sur le milieu PPNO3 en présence de différentes concentrations d'adénine. L'effet de l'adénine sur ces souches est traité dans le texte.

### 3.3. Transcription des allèles *PpAPT* mutés

Nous avons vérifié la transcription des allèles mutés par analyse Northern (Figure 30).

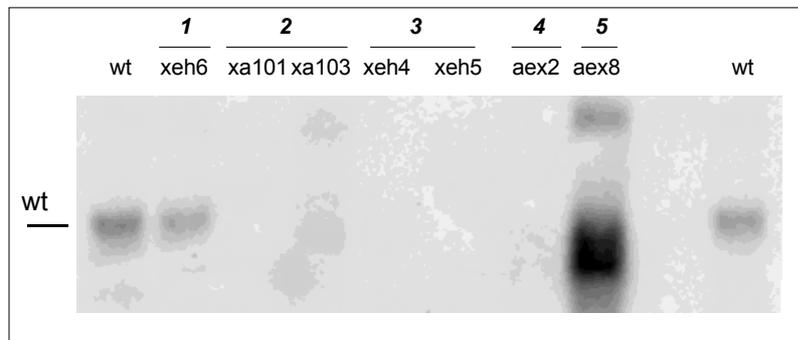


Figure 30. Hybridation de l'ARNm avec la sonde A35 qui couvre le fragment génomique XbaI-EcoRI de 3.5 kb (cf. figure Figure 17). La quantité d'ARN est identique dans toutes les pistes selon la coloration du gel au bromure d'éthidium, explications dans le texte.

- 1 - DAPR-I, XΔX-R
- 2 - DAPR-II, XΔX-IN3'
- 3 - DAPR-I, XΔX-IN5'
- 4 - DAPR-I, ΔΔB-R
- 5 - DAPR-I, ΔΔB-MR
- wt - souche sauvage

#### Type sauvage:

Fin de l'exon 4 et le début de l'exon 5.

```
DNA -----CGAAGTTTGTGCCAATGCGCAAGCCTCGTAAACTTCCAGGT....
protéine      K F V P M R K P R K L P G
```

```
.... AGGCCCTGTGATTAGGGAAGAGTATAGTCTAGAGTATGGTACCGATTGCATTGAG--
      P V I R E E Y S L E Y G T D C I E --
```

#### XΔX-R et son épissage possible (*alEx5*)

Fin de l'exon 4 et le début de l'exon 5.

```
DNA -----CGAAGTTTGTGCCAATGCGCAAGCCTCGTAAACTTCCAGGTAAGGGT....
protéine      K F V P M R K P R K L P G K G
```

```
... TCTAGAGTATGGTACCGATTGCATTGAG--
      Y G T D C I E --
```

Figure 31. Explication de l'épissage hypothétique de l'allèle XΔX-R.

Les lignes interrompues représentent les exons et les lignes pointillées les introns, que nous présentons pas en entier. Les sites d'épissage 5' et 3' (GT - AG) sont en gras. Dans l'allèle XΔX-R, l'exon 5' est tronqué de 32 bp jusqu'au site XbaI (souligné). Le fragment restant peut abriter un site 3' d'épissage; cet épissage ne change pas le cadre de lecture. En italique - l'utilisation hypothétique d'un autre site d'épissage 5'. La protéine produite alEx5 se trouve dans la Figure 32.

Un ARNm est transcrit de l'allèle XΔX-R malgré la délétion de 0.15 kb qui interrompt la séquence codante et joint le milieu de l'intron 4 avec le milieu de l'exon 5 (Figure 30, xeh6, ainsi que xa102, non présenté). Nous pensons que l'ARN de l'allèle XΔX-R ne code pas pour une enzyme fonctionnelle ce qui explique l'absence de l'activité APRT.

En effet, nous pourrions formuler, comme première hypothèse, que la disparition du site d'épissage 3' de l'intron 4 provoque la ligature de l'exon 4 à l'exon 6. Le cadre de lecture est conservé et l'exon 5 est complètement perdu. La taille du transcrit devrait diminuer de 147 bp (exon 5) mais la différence n'est pas aussi grande entre le transcrit sauvage observée et celui de xeh6. Nous proposons alors, comme seconde hypothèse, que l'épissage de l'intron 4 a lieu entre les sites d'épissage alternatifs (Figure 31). Le seul transcrit possible selon la règle GT-AG est plus court que celui du type sauvage de 33 bp, différence que nous ne détectons peut-être pas. Nous présentons sa traduction dans la Figure 32. Une analyse de séquence et une analyse Western sont nécessaires afin d'expliquer exactement le transcrit observé. Dans tous les cas, une partie de la séquence consensus du site de liaison des purines (VIREEYSL) est supprimée par la délétion XbaI et tous les ARNs épissés possibles codent probablement pour une protéine inactive.

Figure 32. Alignement des séquences peptidiques du produit des gènes *APT* de *P. patens*, *A. thaliana* (*AtAPT1*)

Pp	1	<b>MSASKDGD</b> PRIQYISDSIR <b>TIPDFPH</b> KGIMFRDVT <b>TTTTLLD</b> HKAFKDTIDIFVERYRDQKV
At apt1	1	<b>MATEDVQD</b> PRIAK <b>IASSIR</b> V <b>IPDFPK</b> PGIMFQDIT <b>TTTTLLD</b> TEAFKDTIALFVDRYKDKGI
dEx5	1	MSASKDGD <b>PRIQYISDSIR</b> TIPDFPH <b>KGIMFRDVT</b> TTTTLLD <b>HKAFKDTIDIF</b> VERYRDQKV
alEx5	1	MSASKDGD <b>PRIQYISDSIR</b> TIPDFPH <b>KGIMFRDVT</b> TTTTLLD <b>HKAFKDTIDIF</b> VERYRDQKV
Pp	61	DVIVG <b>IEARGFIF</b> GPPIAL <b>AIGAKFV</b> PMR <b>KPKLPG</b> PVIREEYS <b>LEYGTD</b> CIEMHVGAIE
At	61	SVVAG <b>VEARGFIF</b> GPPIAL <b>AIGAKFV</b> PMR <b>KPKLPG</b> KVISEEYS <b>LEYGTD</b> TIEMHVGAIVE
		<b>liaison de la purine ]</b>
dEx5	57	DVIVG <b>IEARGFIF</b> GPPIAL <b>AIGAKFV</b> PMR <b>KPKLP</b> .....
alEx5	57	DVIVG <b>IEARGFIF</b> GPPIAL <b>AIGAKFV</b> PMR <b>KPKLPGK</b> ..... YGTDCIEMHVGAIE
Pp	121	<b>KNERVVIV</b> DDLIATGGTL <b>GAAIKLLD</b> KVEAE <b>I</b> VE <b>CACLV</b> ELVDLKGREKLSGKPLYILVE
At	121	<b>EGERATI</b> DDLIATGGTL <b>AAAIR</b> LLER <b>VG</b> K <b>I</b> VE <b>CACV</b> IELPELKG <b>KE</b> KLGETSLFVLVK
		<b>[liaison PPRP ]</b>
dEx5	92	.....DKVEAE <b>I</b> VE <b>CACLV</b> ELVDLKGREKLSGKPLYILVE
alEx5	109	<b>KNERVVIV</b> DDLIATGGTL <b>GAAIKLLD</b> KVEAE <b>I</b> VE <b>CACLV</b> ELVDLKGREKLSGKPLYILVE
Pp	181	<b>Y</b> EGE-
At	181	<b>S</b> AA--
Ta	177	<b>A</b> DES <b>V</b>
dEx5	127	YEGE
alEx5	144	YEGE

(Schaefer *et al.*, en préparation), (Schnorr *et al.*, 1996).

Les lettres blanches sur un fond noir indiquent une identité d'acide aminé, sur le fond gris une similitude. Les peptides hypothétiques dEx5 et alEx5 sont alignés par rapport aux sites fonctionnels de liaison des purines et de PRPP, qui sont indiqués par les parenthèses carrées et mis en relief par des caractères gras sur un fond gris. Pp - *P. patens*, At - *A. thaliana*; dEx5 - peptide hypothétique de l'allèle *PpAPT* XΔX-R provenant de la disparition de l'exon 5; alEx5 - peptide hypothétique de l'allèle *PpAPT* XΔX-R provenant de l'épissage aberrant de l'intron 4, cf. texte.

Nous détectons des traces de différents transcrits dans les souches de l'allèle XΔX-IN3' (Figure 30, xa101 ou xa103). La particularité de cet allèle réside dans le fait que la seule copie intacte de la séquence codante du gène (la plus en amont et donc précédée du promoteur du gène *PpAPT*) n'a pas de signal de polyadénylation endogène; elle est tronquée dans l'exon 7 sur le codon STOP (HindIII, Figure 17 et Figure 18, pages 48 et 50). Les séquences non traduites 3' (3'UTR) sont remplacées par le vecteur de clonage pBluescript. La nature générale des régions de terminaison de transcription n'est pas suffisamment étudiée ((Lewin, 1999), page 792) ni le mécanisme exact de la formation des séquences d'ARNm 3' (Zhao *et al.*, 1999). Nous ne pouvons donc pas prédire comment la transcription se termine dans cet allèle du gène *PpAPT*. Néanmoins, il est connu que les 3'-UTR participent au contrôle de l'expression des gènes et que leur absence provoque une diminution de la traduction et de la transcription de l'ARN liée à son instabilité ainsi que la formation de transcrits anormalement longs (Conne *et al.*, 2000; Zhao *et al.*, 1999). La transcription de l'ARNm de type sauvage est en effet peu (xa103) ou pas (xa102) détectable et nous observons des transcrits anormalement longs dans l'ARN du clone xa103. En somme, nous postulons que la transcription de l'allèle XΔX-IN3 produit un ARNm qui code pour la protéine APRT normale mais que sa production est réduite par rapport à celle de la souche sauvage. En effet, cette observation confirme l'hypothèse selon laquelle le phénotype DAP<sup>R</sup>-II conféré par l'allèle XΔX-IN3 est dû à une activité APRT diminuée. Nous devrions confirmer cette hypothèse par l'analyse biochimique de l'activité APRT dans les suspensions du protonéma de ces clones (en cours).

Aucun ARNm du gène *PpAPT* n'est détecté dans les clones portant l'allèle XΔX-IN5 (Figure 30, xeh4 et xeh5). En effet, aucune copie fonctionnelle du gène n'est présente dans cet allèle; la copie en 5' est interrompue par une délétion et n'a pas de signal de poly-adénylation, les autres copies n'ont pas de promoteur (Figure 18, page 50).

Nous ne détectons qu'une trace de transcription de l'allèle AΔB-R (Figure 30, aex7, aucune taille définie). La délétion de 1.1 kb interrompt la séquence codante et élimine le premier site d'épissage 5'. En son absence, la séquence codante est interrompue par un codon stop du premier intron. Les codons stops dans tous les cadres de lecture excluent la production d'un ARNm codant. Nous ne pouvons pas expliquer le transcrit de la souche de l'allèle AΔB-MR (Figure 30, aex8). Cette souche porte plusieurs copies du vecteur, chacune possédant les séquences 5' et 3' du gène *PpAPT* interrompu par la délétion BamHI de 1.1 kb. Nous ne savons pas si tous les éléments 3' nécessaires à la terminaison correcte de la transcription sont présents dans ces copies. Par conséquent, la transcription d'un ARNm non fonctionnel et d'une taille anormale a probablement lieu à partir de ces séquences 5' multipliées, et c'est en effet ce que nous observons. Nous ne pouvons pas prédire la séquence exacte de cet ARNm sans analyses supplémentaires. Dans tous les cas, la délétion supprime les sites fonctionnels de la protéine traduite, celle-ci n'est donc pas active.

#### 3.4. Complémentation de l'allèle muté XΔX-R (*APT*)

Nous avons décidé de tenter la complémentation génétique de la mutation *APT*- de *P. patens* par expression des deux gènes, *AtAPT1* et *AtAPT2*.

### 3.4.1. Transformation de la mousse *Physcomitrella patens* en vue de l'obtention de transformants intégratifs exprimant le gène *AtAPT1* ou *AtAPT2*

Nous avons placé l'ADNc des gènes *AtAPT1* et le gène *AtAPT2* sous le contrôle du promoteur du gène actine1 du riz (McElroy *et al.*, 1991), actAtAPT1 et actAtAPT2, respectivement (cf. Annexe B, page 140). Nous avons co-transformé le mutant xeh6 (XΔX-R, DAP<sup>R</sup>-I) par chacun de ces vecteurs mélangé avec le vecteur d'insertion pGL108 (Schaefer and Zryd, 1997). Nous avons sélectionné les clones intégratifs par la résistance à l'hygromycine (pGL108). L'analyse PCR (Figure 33) avec des amorces spécifiques aux gènes *AtAPT1* (atapt1-5' et 3') ou *AtAPT2* (atapt2-5' et 3') a confirmé l'intégration des vecteurs actAtAPT1 et actAtAPT2 dans environ 70% des clones Hyg<sup>R</sup> intégratifs sur 20 (RTF ≈ 10<sup>-4</sup>). Nous avons ensuite analysé 2 clones transformés par le vecteur actAtAPT1 (apt1c4, apt1c5) et 2 clones par le vecteur actAtAPT2 (apt2c2, apt2c5). Nous n'attendons pas d'interaction entre le gène *PpAPT* (ADNg) et les séquences *AtAPT* (ADNc!) parce que la séquence codante du gène *PpAPT* ne possède que 65% d'identité avec *AtAPT1* (recouvrement de 343 bp) et 64% avec *AtAPT2* (recouvrement de 527 bp). L'intégration dans le locus λ108 doit être favorisée par le vecteur co-transformant pGL108 qui porte un fragment de 3.6 kb homologue et isogénique. Une amplification du fragment de 1.17 kb spécifique de l'allèle XΔX-R (m1-5' et m1-3') au lieu de 1.32 kb (WT) dans ces clones démontre que le locus *PpAPT* n'a pas été réparé sur la matrice des gènes *AtAPT*.

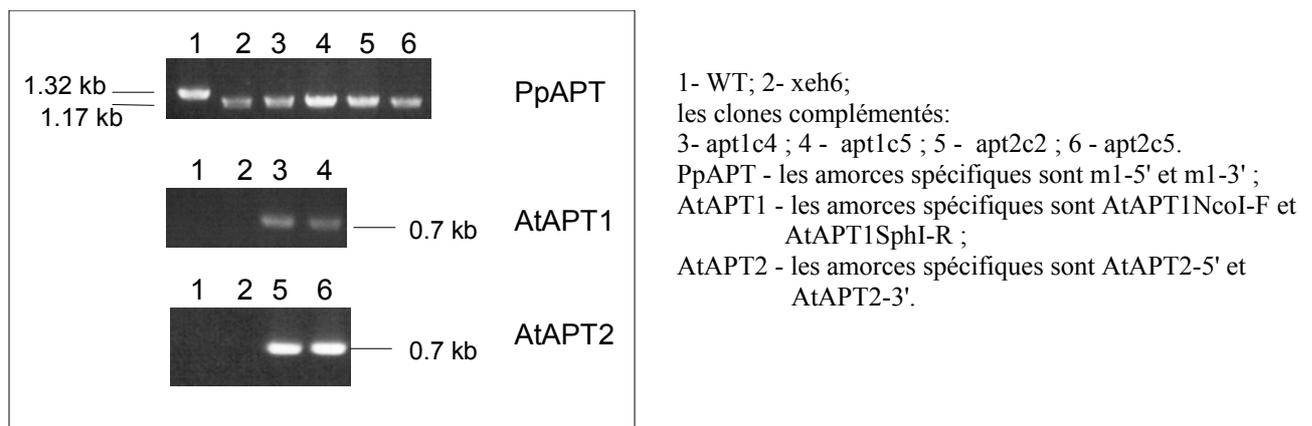


Figure 33 . Analyse moléculaire PCR de l'intégration des gènes *AtAPT1* et 2 dans les transformants intégratifs Hyg<sup>R</sup> de *P. patens*.

Afin de vérifier l'expression des gènes *AtAPT* chez *P. patens*, M. Laloue et collaborateurs ont mesuré l'assimilation de l'adénine (Tableau 5, page 65) et la conversion des cytokinines bases en cytokinines nucléotides (non montré) dans la suspension du protonéma des clones apt1c4, apt1c5, apt2c2, apt2c4 (Schaefer *et al.*, en préparation). L'assimilation de l'adénine est similaire dans le WT et apt1c4, apt1c5, apt2c2, apt2c4 mais pas dans le clone xeh6. En revanche, la conversion de cytokinines ne peut être mesurée que dans les clones apt2c2 et apt2c4 exprimant le gène *AtAPT2* (Schaefer *et al.*, en préparation). Cette analyse confirme que les enzymes APRT codées par les gènes de *A. thaliana* sont actives chez *P. patens* et conservent leur spécificité pour les substrats.

### 3.4.2. Analyse phénotypique des clones *apt1c4*, *apt1c5*, *apt2c2*, *apt2c4*

Le phénotype des souches *apt1c4*, *apt1c5*, *apt2c2* et *apt2c4* est identique à celui de la souche sauvage. La sporogénèse des souches *apt1c4* et *apt1c5* est normale. Les spores de ces souches sont normales et produisent des colonies identiques à celle de la souche sauvage. Nous n'avons pas testé les souches *apt2c2* et 4. Les deux gènes de *A. thaliana* complètent donc génétiquement la mutation du gène *PpAPT*. Ceci n'est pas surprenant puisque l'enzyme codée par ces deux gènes possède l'activité spécifique de l'enzyme PpAPRT, la conversion de l'adénine en AMP. Il est à noter que l'enzyme AtAPRT2 (qui peut également convertir les cytokinines bases en nucléotides correspondants) n'altère pas le phénotype des souches *apt2c2* et *apt2c4*. Nous n'avons pas vérifié si la production de l'enzyme AtAPRT2 altère la composition des cytokinines. Nous pensons que nous ne détecterons pas une telle influence en raison de la compartimentation différente de l'enzyme et des cytokinines bases: la forme majoritaire des cytokinines intracellulaires est nucléotidique. Les cytokinines bases forment la fraction majoritaire des cytokinines extracellulaires chez *P. patens* (Schulz *et al.*, 2000).

Nous pensons que le gène *PpAPT* peut aider à vérifier si le phénotype d'*A. thaliana* est dû à l'altération du métabolisme des cytokinines ou de celui des purines. Il faudrait tenter de compléter génétiquement un mutant d'*A. thaliana* déficient en activité APRT (Moffatt *et al.*, 1991) par le gène *PpAPT*: une complémentation génétique réussie indiquerait que l'altération du métabolisme des purines est responsable du phénotype observé.

### 3.4.3. Utilisation du gène *APT* comme marqueur de sélection et comme gène modèle de la mutagenèse ciblée

La sensibilité de la mousse à DAP en présence d'une enzyme APRT codée par un gène *APT* (endogène ou hétérologue) permet l'utilisation de ce gène comme marqueur de sélection négative. En plus, la sensibilité à l'adénine en absence de l'enzyme APRT permet l'utilisation d'un gène *APT* comme marqueur de sélection positive. Malheureusement, la déficience en activité APRT altère le phénotype de la souche utilisée. Nous discuterons de la possibilité de son utilisation à l'aide d'une expression conditionnelle dans le chapitre V.

La facilité de cibler le gène *PpAPT* et la possibilité de sélectionner par phénotype les différents allèles produits par la mutagenèse ciblée font de ce gène un modèle avantageux des études sur la mutagenèse ciblée, à l'instar de l'expérimentation sur les cellules CHO (Adair *et al.*, 1989). Nous pensons que les études de recombinaison homologue chez les végétaux dont nous discuterons à la fin, pourront aisément se servir de nos connaissances du gène *PpAPT*.

## 4. Discussion

### 4.1. L'efficacité de la mutagenèse ciblée

Nous évaluons l'efficacité de la mutagenèse ciblée au moyen du pourcentage du nombre des événements d'intégration ciblée parmi tous les événements d'intégration. La sélection de transformants DAP<sup>R</sup> employée dans cette étude ne permet d'isoler que les événements d'intégration qui affectent l'activité APRT du produit de l'allèle muté, c'est-à-dire que les événements d'intégration ciblée des vecteurs XΔX, AΔB et BΔB dans le gène *PpAPT*. L'analyse moléculaire ne fait donc que confirmer que seuls les allèles ne contenant pas de gène complet *PpAPT* sont sélectionnés. Nous ignorons par conséquent le nombre des événements d'intégration non ciblée, ou illégitime, de ces vecteurs. Seule l'analyse moléculaire de l'intégration du vecteur AΔBneo nous donne une indication sur l'efficacité de la mutagenèse ciblée; cette analyse démontre que le gène *PpAPT* est ciblé dans 100 % des clones analysés (9/9).

Nous pouvons, en revanche, évaluer la mutagenèse ciblée en termes de fréquence relative de transformation (RTF). Cette valeur nous permet de comparer la capacité *absolue* des vecteurs utilisés à s'intégrer par recombinaison homologue. Nous devrions toutefois corriger les valeurs RTF initiales présentés dans la Figure 19, page 51. Nous expliquons cette correction comme suit.

Premièrement, nous avons isolé des mutants spontanés. Ces souches DAP<sup>R</sup> apparaissent indépendamment de la transformation car ils n'intègrent pas l'ADN transformant et peuvent apparaître en présence de DAP sans le traitement transformant. Leur proportion est d'environ 10% de tous les clones DAP<sup>R</sup> produits par les transformations avec les vecteurs utilisés. Sachant donc que le gène *PpAPT* est ciblé dans 90% des clones analysés, nous corrigerons les valeurs RTF présentées dans la Figure 19 d'un facteur de 0.9 pour obtenir la proportion d'intégration ciblée dans tous les protoplastes régénérés après la transformation.

Deuxièmement, tous les événements d'intégration ciblée ne sont isolés que pour le vecteur XΔX. Comme l'indique le Tableau 7, page 75, environ la moitié de ces événements sont des remplacements et l'autre moitié des insertions. Nous trouvons une distribution similaire pour le vecteur AΔBneo. Si la distribution entre les événements d'insertion et de remplacement des vecteurs AΔB et BΔB est similaire à celle observée pour les vecteurs XΔX et AΔBneo, la valeur RTF obtenue pour les vecteurs AΔB et BΔB ne représente que la moitié de la fréquence de l'intégration ciblée de ces vecteurs dans le gène *PpAPT*.

Nous corrigeons les valeurs RTF présentées dans le Tableau 6 comme suit:

- Nous multiplions la valeur initiale par un facteur de 0.9 pour tous les vecteurs (sauf AΔBneo) compte tenu de l'apparition de clones DAP<sup>R</sup> non ciblés;
- Nous multiplions la valeur initiale par un facteur de 2 pour les vecteurs AΔB et BΔB puisqu'environ la moitié des événements d'intégration ciblée ne sont pas sélectionnables.

Tableau 6. La correction des valeurs RTF moyennes observées dans la transformation de *P. patens* par les vecteurs de remplacement.

Vecteurs	forme du vecteur	Taille d'homologie et taille du vecteur	RTF moyenne	RTF moyenne corrigée	RTF / mole de vecteur	RTF corrigée / mole de vecteur
pHP23	ou sc	-	0.1		0.017	
pGL2 §	lin		1.0		0.17	
XΔX	sc	0.5 + 0.7	0.95	0.86	0.18	0.17
F=x0.9	lin		<u>4.4</u>	9.95	8.96	1.92
BΔB	sc	1.6 + 2.8	1.42	2.56	0.45	0.81
F=x0.9x2	lin		<u>7.2</u>	1.75	3.15	0.55
AΔB	sc	1.3 + 1.6	2.20	3.96	0.58	1.04
F=x0.9x2	lin		<u>6.0</u>	6.13	11.03	1.61
AΔBneo	sc	1.3 + 1.6	1.95	1.95	0.66	0.66
	lin		<u>7.7</u>	3.50	3.50	1.18

§ (Schaefer and Zryd, 1997), chapitres I et II, intégration illégitime

Les valeurs correspondent au nombre de colonies sélectionnées (DAP<sup>R</sup> ou Neo<sup>R</sup>) par 10<sup>4</sup> protoplastes régénérants (RTF) présenté dans la Figure 19 et corrigé avec le facteur F. Nous divisons les valeurs RTF par le nombre de mole de l'ADN transformant pour tenir compte de la différente taille des vecteurs. sc - surenroulé; lin - linéaire (schéma des vecteurs, cf. Figure 17, page 48).

La comparaison de la valeur RTF avec celle obtenue avec les vecteurs non homologues nous indique indirectement que l'intégration a lieu par recombinaison homologue (voir également la discussion finale, page 123).

La différence entre l'efficacité de la mutagenèse des vecteurs circulaires n'est pas substantielle, mais semble dépendre de la taille de l'homologie. La taille optimale, que possèdent le vecteur AΔB et le vecteur AΔBneo, est de 1.3 + 1.6 kb.

L'efficacité est plus faible pour le vecteur BΔB portant un fragment d'homologie plus importante. Il ne paraît pas possible que cette faible efficacité soit due à la taille du vecteur même puisque le vecteur AΔBneo est le plus gros de tous les vecteurs. Chez la souris, la recombinaison a lieu plus fréquemment dans la région plus proche des extrémités (Deng *et al.*, 1993). Il est possible alors que l'efficacité moindre du vecteur BΔB soit provoquée par le fait que les régions qui se trouvent vers les extrémités du fragment homologue de ce vecteur soient spécifiquement moins aptes à la recombinaison. Ce sont en effet les régions qui sont absentes des autres vecteurs, mais nous ne les avons pas analysées suffisamment pour expliquer un tel effet. Nous pensons que la distribution de l'homologie autour du marqueur de sélection peut aussi expliquer la faible efficacité du vecteur BΔB. En effet, la distribution est particulièrement biaisée dans le vecteur BΔB:  $\frac{1}{3}$  -  $\frac{2}{3}$  contre environ  $\frac{1}{2}$  -  $\frac{1}{2}$  dans les autres vecteurs. Ce biais pourrait favoriser l'insertion par recombinaison dans la partie la plus grande du fragment de remplacement (3'). Etant donné que nous ne pouvons pas isoler les événements d'insertion dans notre crible DAP<sup>R</sup>, nous n'avons pas inclus dans les valeurs RTF la majorité des événements d'intégration du vecteur BΔB. Afin de prouver cette hypothèse, il faudrait étudier l'influence de la distribution de l'homologie autour du marqueur de sélection

positive sur l'intégration du vecteur comme nous le proposons plus loin (page 76, Figure 34). De plus, tous les événements d'intégration doivent être isolés.

Une relation entre la taille d'homologie portée par des vecteurs de remplacement a été établie, par exemple, pour la souris (Hasty *et al.*, 1991a). Dans cet organisme, la longueur totale des fragments homologues du vecteur est un facteur limitant l'efficacité de la mutagenèse ciblée. Ce facteur est plus important pour une intégration ciblée que la distribution de l'homologie entre le fragment 5' et 3'. Cette dernière étude démontre que la longueur de l'homologie entre 1.7 et 1.9 kb est un seuil critique pour le gène utilisé (*HPRT*), à partir duquel l'intégration ciblée devient détectable. L'efficacité augmente globalement d'un facteur de 190 si la taille d'homologie augmente de 1.3 à 6.8 kb. En revanche, environ 6 kb d'homologie est le seuil supérieur, à partir duquel l'efficacité n'augmente plus en fonction de la taille d'homologie, au moins pour ce gène.

Nous tirons les conclusions suivantes sur l'efficacité de la mutagenèse ciblée du gène *PpAPT* par des vecteurs de remplacement basés sur les fragments d'ADN génomiques :

- ✓ L'efficacité de mutagenèse des vecteurs de taille moyenne (4 à 8 kb) portant un fragment d'homologie entre 1.2 et 4.4 kb distribuée équitablement autour d'un marqueur de sélection est du même ordre de grandeur;
- ✓ L'insertion d'un gène marqueur de sélection hétérologue dans un fragment de remplacement ne diminue pas de façon importante l'efficacité de la mutagenèse par recombinaison homologue.

Nous estimons que le vecteur portant le fragment homologue le plus court atteint voire même dépasse le seuil à partir duquel la longueur de l'homologie n'augmente plus l'efficacité d'intégration du vecteur. Les résultats obtenus précédemment dans notre laboratoire (Schaefer *et al.*, en préparation) confirment cette hypothèse. Schaefer, Vlach et collaborateurs ont observé les valeurs RTF inférieures d'un facteur de 10 pour des vecteurs de remplacement construits à partir de l'ADNc de *PpAPT* (Schaefer *et al.*, en préparation). La longueur du fragment d'homologie le plus court encore efficace est de 260 + 418 bp, l'homologie en 5' étant interrompue par les introns 2 et 3 et en 3' par les introns 4, 5 et 6. L'analyse de la séquence du site d'intégration met en évidence que 50 bp suffit à la recombinaison dans le bras 5' du vecteur. Nous postulons que la longueur d'homologie s'approche du seuil au-dessous duquel l'intégration n'est plus détectable. Nous le traiterons dans la discussion finale.

Les conditions qui se démarquent par leur efficacité de mutagenèse sont les vecteurs XΔX et AΔB linéarisés (Figure 19, page 51 et Tableau 6, page 73). L'influence de ruptures double brin sur l'intégration de vecteurs sera débattue dans la discussion finale.

## 4.2. Types d'intégration des quatre vecteurs XΔX, BΔB, AΔB, AΔBneo.

Nous résumons l'analyse détaillée ci-dessus dans le tableau suivant:

Tableau 7. Distribution de différents types d'intégration en nombre absolu de clones analysés.

vecteur	remplacement	remplacement multiple	insertion 5'	insertion 3'	autre type d'intégration ciblée	Total
<i>XΔX</i>	6	3	2	3	2	16
<i>BΔB</i>	5	4	nd	nd	nd	9
<i>AΔB</i>	6	2	nd	nd	nd	8
<i>AΔBneo</i>	3	2	2	1	1	9
<b>Total</b>	<b>20</b>	<b>11</b>	<b>4</b>	<b>4</b>	<b>3</b>	<b>42</b>

nd - non isolé par notre crible

Chez la levure *S. cerevisiae*, l'intégration de vecteurs de type remplacement est moins efficace que celle de vecteurs d'insertion. Pratiquement la totalité (70-85%, (Orr-Weaver *et al.*, 1981)) des événements d'intégration des vecteurs de remplacement ont lieu par remplacement simple (Hastings *et al.*, 1993; Rothstein, 1991) et probablement plus par conversion de gène sans crossing-over que par un échange réciproque de séquences (Leung *et al.*, 1997).

Nous avons observé chez *P. patens* l'intégration des vecteurs de remplacement par remplacement multiple ou par insertion (Tableau 7). Ces événements ressemble davantage à ce qui est observé chez le hamster (cellules ovariennes, CHO) (Adair *et al.*, 1998) et chez la souris (cellules d'embryon, ES) (Hasty *et al.*, 1991b). Nous devons toutefois être prudents car l'intégration par recombinaison homologue dans ces systèmes expérimentaux est très peu efficace et la voie d'intégration majeure reste NHEJ.

Dans les cellules CHO ou ES, une relation est établie entre la taille d'homologie, la configuration du vecteur, la distribution de l'homologie entre les bras du vecteur, d'un côté, et le type d'intégration, de l'autre. Chez la souris, du moins, cette relation dépend des facteurs limitants que sont l'efficacité même de l'intégration par recombinaison homologue, la longueur minimale d'homologie, voire la nécessité de séquences isogènes recombinantes. Il est intéressant de constater des similitudes entre ces systèmes et la mousse *P. patens*, dans laquelle la recombinaison homologue est très efficace avec aussi peu que 0.50 kb d'homologie (Schaefer *et al.*, en préparation). Il se pourrait que ces similitudes sont liées à une similitude de l'organisation du génome, à la présence des introns (peu de gènes de *S. cerevisiae* en ont) ou à la structure de la chromatine, mais nous n'avons pas explorés ces phénomènes pour émettre une hypothèse précise.

Dans les paragraphes suivants nous tenterons de vérifier la relation entre le type d'intégration et les caractéristiques du vecteur : la taille du fragment homologue, la distribution de l'homologie entre les bras 5' et 3', et la forme (circulaire ou linéaire).

### Insertion ou remplacement selon la taille du fragment homologue

Le Tableau 7 nous permet de comparer le type d'intégration ciblée des vecteurs utilisés. La comparaison est particulièrement valable pour le vecteur XΔX puisque tous les quatre types d'événements ont pu être isolés par la sélection en présence de la DAP. Chacun des 4 types d'intégration génère environ  $\frac{1}{4}$  des événements d'intégration. Le remplacement et l'insertion comprennent environ la moitié des événements.

La comparaison des fréquences de différents types d'intégration du vecteur AΔBneo ne peut pas se baser sur le Tableau 7 parce que le nombre de clones DAP<sup>S</sup> et DAP<sup>R</sup> retenus pour l'analyse est arbitraire et ne reflète pas leur proportion respective au sein de tous les protoplastes Neo<sup>R</sup> (donc intégratifs). Environ  $\frac{2}{3}$  des clones Neo<sup>R</sup> se sont avérés DAP<sup>R</sup>. Sachant que le remplacement ciblé génère le phénotype DAP<sup>R</sup> et que l'insertion ciblée génère le phénotype DAP<sup>S</sup>, nous postulons donc que les événements d'intégration par remplacement constitue  $\frac{2}{3}$  de tous les événements d'intégration. Nous constatons que tous les types d'événements ont pu avoir lieu avec le vecteur AΔBneo indépendamment du mode de sélection.

La différence entre les vecteurs de notre étude réside essentiellement dans la longueur du fragment homologue. Dans les vecteurs XΔX, AΔB et AΔBneo, l'interruption de l'homologie (délétion ou insertion du gène marqueur) se trouve approximativement au milieu du fragment (cf. Figure 17). La distribution relativement homogène entre les types d'intégration ciblée possibles pourrait donc provenir de la distribution de l'homologie autour du marqueur de sélection. Afin d'étayer une telle hypothèse, nous devrions tester différentes versions du même vecteur. Les conditions testées doivent comprendre différentes distributions de la même taille d'homologie autour du marqueur (Figure 34, A) et différentes tailles d'homologie autour du même site d'insertion du marqueur (Figure 34, B).

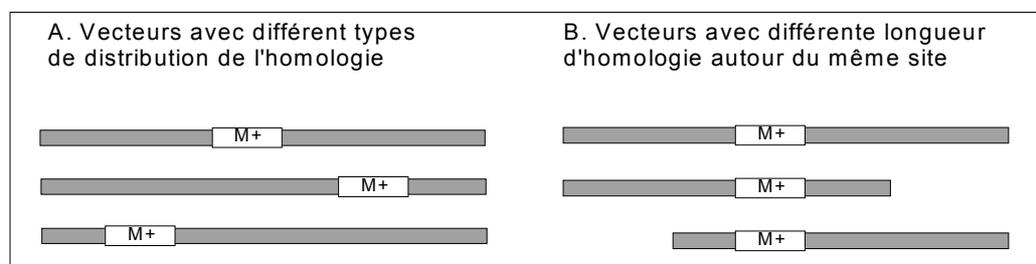


Figure 34. Conditions - tests pour étudier l'influence de la disposition du vecteur.

Thomas et collaborateurs testent une de ces conditions (B) chez la souris (cellules ES) (Thomas *et al.*, 1992). Un vecteur portant 9.1 kb (6.1 + 3 kb autour du gène marqueur) d'homologie au gène *HPRT* de la souris génère un événement de remplacement dans 100% des clones testés. Lorsque le fragment 3' est raccourci de 3 à 0.7 kb, la taille du fragment en 5' maintenue à 6.1 kb, le même événement se produit dans 13 / 18 clones seulement. La réduction de la taille du fragment en 3' provoque des événements de remplacement de la copie WT par des concaténats du vecteur (de type remplacement multiple) dans 3 clones (sur 18) et des événements plus complexes dans 2 autres clones. Les auteurs mentionnent qu'un vecteur avec 5.4 kb d'homologie mais distribuée également entre les fragments 5' et 3', ne génère que des événements de remplacement simple. Il est possible que la différence entre les types d'intégration de ces vecteurs provienne du fait que la longueur de 0.7 kb est insuffisante pour une recombinaison efficace chez la souris et que le concaténat du vecteur permet de

l'augmenter. La distribution entre les fragments 5' et 3' pourrait jouer un rôle dans le mécanisme d'intégration mais les auteurs ne favorisent pas cette hypothèse. En effet, une autre étude sur la souris (Hasty *et al.*, 1991a) démontre que le vecteur portant une distribution de l'homologie la plus asymétrique (5.6 + 0.472 kb) affiche le taux d'événements de remplacement le plus élevé, à savoir 7 sur 12 testés contre 2/5 ou 1/12 pour deux autres vecteurs portant 5.6 + 1.2 kb.

Nous ne pouvons faire la comparaison selon la condition exposée dans la Figure 34, B qu'entre le vecteur AΔB et BΔB. Nous constatons que le vecteur AΔB s'intègre plus fréquemment par remplacement simple que par remplacement multiple (6 contre 2). Le vecteur BΔB, avec une distribution de l'homologie moins équilibrée, s'intègre aussi souvent par remplacement simple que multiple (5 contre 4). Nous supposons que c'est la distribution de l'homologie entre les bras 5' et 3' qui influence le type d'intégration. Etant donné que la taille totale de l'homologie portée par le vecteur AΔB est inférieure à celle du vecteur BΔB, il est improbable que les raisonnements évoqués plus haut (trop peu d'homologie provoque le remplacement multiple) puissent s'appliquer à cette observation.

L'intégration de vecteurs de remplacement, basés sur l'ADNc du gène *PpAPT* produit les mêmes types d'intégration chez *P. patens* (Schaefer *et al.*, en préparation). Les bras 5' et 3' des vecteurs utilisés sont d'environ 0.3-0.4 kb d'homologie qui est interrompue par la présence des introns dans l'ADNg cible. Ces résultats nous montrent le seuil de la longueur d'homologie la plus petite encore suffisante pour une intégration à des fréquences détectables.

#### Vecteur linéaire et circulaire: insertion ou remplacement

Il n'y a pas de différence significative entre les types d'intégration de différentes formes d'un même vecteur chez le hamster (CHO) (Adair *et al.*, 1998). Les événements de remplacements du locus *APT* par une copie du fragment de remplacement ont lieu dans 50 et 78% des cas pour les vecteurs dont l'extrémité 5' ou 3', respectivement, est bloquée par le vecteur de clonage (donc par des séquences hétérologues). La majorité des événements restants sont la correction du vecteur sur la matrice du locus ciblé suivie d'une intégration illégitime du vecteur corrigé. Lorsque les deux extrémités sont bloquées (forme circulaire), les mêmes événements sont observés (49% remplacement simple, 51% correction et intégration illégitime). Des événements de remplacement multiple sont minoritaires (2 sur 102) et ne sont générés que par le vecteur avec l'extrémité 3' bloquée par une séquence hétérologue.

Tableau 8. Comparaisons possibles entre les types d'intégration pour les vecteurs circulaires ou linéaires

Type d'intégration	XΔX sc	XΔX lin	ΔΔBneo sc	ΔΔBneo lin
<i>R</i>	1	2	1	2
<i>RM</i>	-	3	-	2
<i>IN</i>	2	3	3	-

R - remplacement, RM -R multiple, IN insertion 5' ou 3' que nous ne pouvons tester que pour XΔX ou ΔΔBneo; sc - circulaire surenroulé; lin - produit de la digestion qui libère le fragment de remplacement.

Nous n'avons pas planifié nos expériences de sorte que nous puissions vérifier l'influence de la forme de l'ADN transformant sur son type d'intégration pour tous les vecteurs. Nous pourrions toutefois le vérifier pour le vecteur XΔX. Or, le nombre de clones analysés selon la forme d'ADN transformant est arbitraire (cf. page 52, nous avons analysé 6 pour sc et 8 pour lin) et n'est pas en relation exacte avec les valeurs RTF présentées dans la Figure 19, page 51 ( $0.9 \times 10^{-4}$  pour sc et  $9.9 \times 10^{-4}$  pour lin). Sans comparer de façon quantitative, nous concluons que les deux formes de ces vecteurs peuvent s'intégrer par remplacement simple ou multiple aussi bien que par insertion (Tableau 8).

#### 4.3. Conclusion

Nous avons démontré que le remplacement d'allèles (conversion génique) par simple copie est possible chez *P. patens*. Les événements de remplacement par copie unique ont pu être isolés en absence de contre-sélection à l'aide de marqueur de sélection négative. Il ont constitué environ un quart de tous les événements d'intégration pour deux vecteurs utilisés pour le même gène *PpAPT*.

L'intégration dans le gène *PpAPT* est très efficace, mesurée en termes RTF, qui atteint 1 pour 1000 protoplastes régénérés.

Les événements de remplacement multiples ressemblent à ceux connus chez la souris (ES) ou le hamster (CHO). Nous avons mis en évidence ces événements pour le vecteur linéaire ou circulaire, mais nous n'avons pas déterminé la relation exacte entre la forme du vecteur et le type d'intégration. Nous n'avons analysé qu'un des produits d'intégration (à savoir la copie génomique produite), nous ne distinguons donc pas la conversion de gène et l'intégration par une recombinaison réciproque par un double crossing-over.

La religature des fragments linéaires du vecteur à l'intérieur de la cellule accompagne l'intégration précise par recombinaison homologue. Les cas minoritaires d'intégration complexes (<10%) révèle l'existence d'une activité nucléasique qui provoquerait l'intégration de concaténats de fragments du vecteur partiellement digérés, peut être avec la participation de NHEJ. Néanmoins, une analyse rigoureuse de séquence devrait être nécessaire. Schaefer met en évidence l'existence d'une activité exonucléasique qui accompagne l'intégration des vecteurs non homologues dans certains cas (Schaefer, communication personnelle).

Dans le chapitre II, nous avons postulé que la correction du vecteur sur la matrice génomique homologue par réparation de brèches ou de délétions n'est pas assez efficace chez *P. patens*

pour constituer un obstacle à la mutagenèse par remplacement. L'efficacité de la mutagenèse observée le confirme pour le gène *PpAPT*. Au moins un événement de correction du fragment de remplacement sur la matrice du gène *PpAPT* a toutefois été isolé. Dans les conditions qui ont permis de l'isoler, cet événement a été minoritaire (1 sur 4) ce qui est en accord avec les résultats obtenus sur le locus  $\lambda 213$ .

L'efficacité de l'intégration des vecteurs de remplacement utilisés nous permet dorénavant le développement de la mutagenèse de précision. Nous démontrons ainsi la possibilité de la mutagenèse directe par remplacement ciblé de l'allèle sauvage par un allèle muté dont le phénotype peut être sélectionné. Dans le cas du gène *PpAPT* que nous présentons, nous avons planifié et produit différents allèles pour analyser ensuite le phénotype qu'ils confèrent. Nous avons trouvé que l'inactivation du gène *PpAPT* par une délétion introduite dans la région codante provoque un phénotype anormal. Le phénotype résulte probablement de la perturbation du métabolisme de l'adénine qui n'est plus correctement recyclée en absence de l'activité APRT. Le phénotype peut être complété par l'expression d'un gène homologue de *A. thaliana* (*AtAPT1* ou *AtAPT2*). Les gènes d'un angiosperme sont donc parfaitement actifs chez la mousse *P. patens*.

Pour sélectionner des allèles mutés dont le phénotype n'est pas discernable (voire inconnu), nous pouvons recourir à des gènes marqueurs placés dans les introns ou en dehors des éléments essentiels du gène (promoteur, 3' UTR). Nous pourrions rapidement vérifier cette stratégie sur le gène *PpAPT*. Nous construirions un vecteur test en introduisant un gène marqueur de sélection positive dans un fragment homologue du gène *PpAPT*, à l'intérieur d'un intron (le gène 35S *nptII* ou 35S *hpt* peut être introduit dans le site EcoRV de l'intron IV). Nous transformerions *P. patens* par ce vecteur test et sélectionnerions un événement de remplacement simple par la résistance à l'antibiotique et par une analyse moléculaire. Si le phénotype de la souche sélectionnée était normal, *APT+*, nous démontrerions que le marqueur n'influence pas l'épissage de cet intron. Au besoin, le marqueur pourrait être excisé par le système Cre/lox (cf. chapitre IV). Ensuite, *P. patens* serait transformée par un vecteur similaire portant la mutation ponctuelle étudiée dans la séquence codante du gène *PpAPT*. Nous sélectionnerions de nouveau un événement de remplacement simple par la résistance à l'antibiotique et par une analyse moléculaire. Le phénotype observé chez la souche sélectionnée serait attribuable uniquement à la mutation.

En revanche, les autres stratégies ("pop in / pop out", remplacement en deux étapes selon (Rothstein, 1991)) restent exclues tant qu'un marqueur de sélection négative adéquat ne sera pas disponible pour la mousse *P. patens*.

## Chapitre IV. Excision de marqueurs de sélection à l'aide d'un système de recombinaison spécifique d'un site Cre/lox

### Résumé

La mutagenèse de précision exige que les séquences exogènes indésirables puisse être éliminées du génome des souches génétiquement modifiées. Nous démontrons la possibilité d'exciser un gène marqueur de sélection intégré (*35S nptII*) à l'aide d'un système de recombinaison spécifique de site Cre/lox. Le système comprend les sites de reconnaissance loxP en orientation parallèle qui flanquent le gène à exciser et la recombinaison Cre du bactériophage P1 qui reconnaît spécifiquement les sites loxP et excise la région flanquée. Le gène *cre* codant pour la recombinaison est exprimée de façon transitoire sous le contrôle du promoteur 35S dans les protoplastes de la souche portant le marqueur et l'excision est vérifiée par une perte de résistance à G-418 et confirmée par une analyse moléculaire. L'excision précise d'une copie du gène flanqué de sites loxP a lieu dans 65% des protoplastes ayant exprimé la recombinaison, soit dans environ 5% des colonies testés. Lorsque la recombinaison Cre agit sur des répétitions directes du vecteur portant le gène bordé des sites loxP, l'excision a lieu avec efficacité similaire. Dans cette situation, nous avons observé une réduction du nombre de répétitions du vecteur.

### 1. Introduction

La mutagenèse ciblée précise exige que les séquences indésirables telles que les marqueurs de sélection, les vecteurs de clonage ou de transfert, puissent être éliminées du génome de la lignée transgénique. Dans les paragraphes suivants, nous traitons de l'élimination des gènes marqueurs de sélection à l'aide de la recombinaison spécifique de site.

#### 1.1. Problèmes liés à l'utilisation de gènes marqueurs de sélection

La présence d'un marqueur de sélection dans le gène muté par un événement d'intégration ciblée du vecteur pose plusieurs problèmes:

##### Interférence avec l'expression du gène d'intérêt

Les études sur la souris montrent que l'influence des marqueurs de sélection sur le gène qu'il interrompt peut masquer ou altérer le phénotype dû à la mutation étudiée. L'interférence peut avoir lieu à cause d'une expression trop forte due au promoteur du gène marqueur ou à cause des propriétés régulatrices de la structure du gène marqueur même. Fiering et les auteurs des travaux qu'il cite supposent que le gène marqueur, et le plus probablement ses éléments de régulation (promoteur ou enhancers), influence les éléments régulateurs des gènes locaux. Une transcription mal achevée du gène marqueur peut créer un transcrit fusionné avec une partie du gène qu'il interrompt ("read-through"). Selon l'orientation du gène marqueur, ce transcrit chimérique peut avoir un effet d'antisens ou de co-suppression. Il est possible que l'influence soit indirecte et s'étende sur d'autres gènes voisins ou distants (Fiering *et al.*, 1999).

Ces incertitudes ne permettent pas d'envisager une prédiction générale sur l'influence que le marqueur aura sur l'expression des gènes dans un cas donné de mutagenèse ciblée. Il est alors préférable d'éliminer les gènes marqueurs.

#### Impossibilité d'une utilisation ultérieure de la même cassette de résistance

Si un transformant a été sélectionné pour l'intégration d'une cassette de résistance à un antibiotique, il est évident qu'une deuxième transformation ne pourra plus recourir à la résistance au même antibiotique pour la sélection des transformants. Quatre gènes marqueurs de sélection positive seulement ont été testés et optimisés à ce jour pour *P. patens*: *nptII* et *hpt* codant pour néomycine et hygromycine phosphotransférases (Cove *et al.*, 1991; Schaefer *et al.*, 1991), *APT* codant pour adénine-phosphoribosyltransférase [chapitre III] (Schaefer *et al.*, en préparation) et *sul* codant pour dihydroptéroate synthase (résistance à sulfadiazine) (Kammerer and Cove, 1996).

#### Introduction d'une cible potentielle pour une intégration d'un élément homologue (promoteur, gène, terminateur)

Les vecteurs pour les transformations des souches transformées ne doivent pas contenir les éléments du gène marqueur déjà introduit dans *P. patens* pour éviter la recombinaison homologue. Ce point est particulièrement important car l'intégration de l'ADN transformant peut avoir lieu dans une région homologue dans presque 100% de cas chez *P. patens*.

#### Incertitude de la distribution dans l'environnement des mutants utilisés dans l'agriculture

L'opinion publique et la législation de plusieurs pays (Bryant and Leather, 1992; Fuchs *et al.*, 1993) exigent que les séquences exogènes telles que les marqueurs de sélection ou les vecteurs de clonage soient éliminées des organismes utilisés dans des applications biotechnologiques. La nécessité d'éliminer les marqueurs de sélection et d'autres séquences exogènes indésirables peut s'avérer très importante si le développement de la mousse *P. patens* attire l'intérêt d'investisseurs privés de l'industrie ou de l'agriculture.

## 1.2. Techniques permettant d'enlever les marqueurs de sélection

Suite à l'introduction par un vecteur de transformation, le marqueur de sélection se trouve dans une position très proche de la mutation introduite. L'élimination du marqueur à l'aide de croisements est alors peu efficace voire impossible puisque le marqueur et l'allèle muté vont suivre une ségrégation de gènes liés.

La mutagenèse à deux étapes de la levure *S. cerevisiae* permet l'élimination du gène marqueur à l'aide de l'utilisation d'un marqueur de sélection négative (Rothstein, 1991). Nous n'avons pas testé ces techniques sur la mousse *P. patens*, premièrement en raison de l'absence de marqueur de sélection négative adéquat (pour l'utilisation d'un gène *APT* comme marqueur cf. chapitre III, page 71).

La co-transformation dans certains cas peut fournir une solution à la séparation du transgène introduit et du marqueur de sélection (Yoder and Goldsbrough, 1994). Les deux doivent être introduits dans le transformant par deux vecteurs séparés dans une co-transformation. Pour stimuler une intégration indépendante, l'un portera un fragment de remplacement du gène étudié et l'autre le gène marqueur et un fragment homologue à un gène connu qui n'est pas lié au gène étudié. Ces deux vecteurs ne doivent contenir aucun fragment identique afin

d'empêcher leur recombinaison extrachromosomique. L'élimination du gène marqueur se fait ensuite par une série de croisements.

Le système de transposons de maïs *Ac/Ds* permet également d'éliminer les marqueurs de sélection (Ebinuma *et al.*, 1997; Goldsbrough *et al.*, 1993; Yoder and Goldsbrough, 1994). Le marqueur est bordé de deux éléments *Ds* non autonomes et le gène codant pour la transposase est introduit dans la même lignée simultanément ou ultérieurement, par une re-transformation ou par un croisement. Suite à l'action de la transposase, le marqueur est transposé ailleurs dans le génome et peut être éliminé par des croisements. D'un autre côté, une faible partie des éléments *Ds* excisés ne s'intègrent pas dans le génome et sont perdus, mais l'efficacité de cette méthode est faible.

Une stratégie directe consiste à éliminer le marqueur de sélection à l'aide d'un système de recombinaison spécifique d'un site. A cette fin, le marqueur bordé de deux sites de recombinaison spécifique est introduit dans le vecteur de transformation. Une fois que la souche mutante désirée est isolée, une recombinaison entre les sites bordures catalysée par une enzyme spécifique excisera le marqueur. Plusieurs systèmes de recombinases spécifiques d'un site existent et pourraient être testés.

### 1.3. Recombinase spécifique d'un site

La plupart des systèmes de recombinaison spécifique d'un site utilisés dans des hôtes hétérologues sont des systèmes à deux composantes: la recombinase elle-même et le site qui est spécifiquement reconnu par elle. Les quatre systèmes qui ont été testés chez les végétaux sont présentés dans le Tableau 9. Les enzymes de la classe des intégrases catalysent la réaction réversible d'intégration/excision présentée dans la Figure 35.

Tableau 9. Les systèmes de recombinaison spécifique d'un site fonctionnels chez les plantes (adapté de (Odell and Russell, 1994))

recombinase/site :	Cre/lox	FLP/FRT	R/RS	Gin/gix
<i>origine</i>	bactériophage P1	plasmide 2 $\mu$ m de <i>S. cerevisiae</i>	plasmide pSR1 de <i>Z. rouxii</i>	bactériophage Mu
<i>fonction naturelle</i>	maintien de la stabilité du phage; intégration dans le génome de l'hôte; circularisation du phage	maintien du nombre de copies du plasmide	maintien du nombre de copies du plasmide et de l'homologie exacte des répétitions bordures;	régulation d'expression de gènes
<i>enzyme</i>	Cre: 38.5 kDa, intégrase	FLP: 48 kDa, intégrase	R: 56 kDa, intégrase	Gin: 21.7 kDa, invertase
<i>site de reconnaissance</i>	loxP de 34 bp	FRT de 34 bp	RS de $\leq$ 58 bp	gix de 34 bp
<i>recombinaison dans les hôtes végétaux</i>	<i>A. thaliana</i> et <i>N. tabacum</i>	<i>A. thaliana</i> et <i>N. tabacum</i> ; cellules de <i>Z. mais</i> et <i>O. sativa</i>	<i>A. thaliana</i> et cellules de <i>N. tabacum</i>	cellules de <i>A. thaliana</i> et <i>N. tabacum</i>

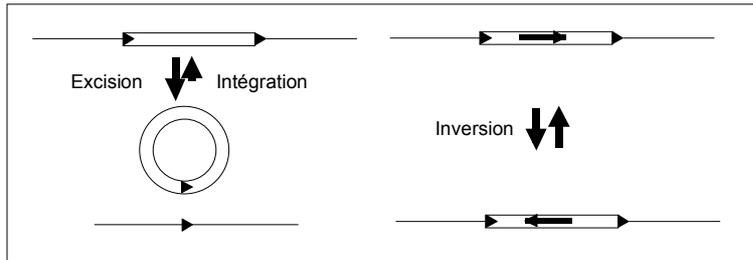


Figure 35. Les réactions catalysées par les intégrases.

Les triangles noirs orientés représentent les sites de reconnaissance.

Nous avons écarté le système Gin/gix de notre étude car l'enzyme naturelle n'est pas capable d'excision. Les plantes produisant l'enzyme Gin<sup>mut</sup>, mutée (capable de la réaction d'excision), de façon constitutive présentent des altérations sévères de phénotype.

Les trois enzymes restantes peuvent remplir la fonction d'intégrase dans une cellule végétale. Nous avons décidé de tester le système Cre/lox sur la mousse *P. patens*.

#### 1.4. Système de recombinaison spécifique d'un site Cre/lox

La première fonction naturelle du système Cre/lox est la résolution des structures dimères du phage P1 dans la cellule hôte *E. coli* (Odell and Russell, 1994). L'enzyme Cre de 38.5 kDa constitué de 343 aa est codée par le génome du phage. La séquence codante est de 1029 bp. L'enzyme peut catalyser la réaction entre différentes formes d'ADN. Il a été démontré que les sites loxP parallèles ou antiparallèles peuvent se trouver sur la même molécule d'ADN, sur des molécules différentes, sur une molécule d'ADN linéaire ou encore circulaire relaxée ou surenroulée. In vivo, l'enzyme agit de façon plus efficace sur les sites parallèles situés sur la même molécule avec surenroulement négatif. Cette affinité de l'enzyme pour le substrat monomoléculaire déplace l'équilibre dynamique de la réaction (Figure 35) vers l'excision.

Dans le génome du phage P1, les sites loxP sont séparés par environ 100 kb et la distance ne semble pas être le facteur limitant de la réaction (Odell and Russell, 1994). La structure des sites loxP comprend deux séquences bordures identiques de 13 bp en répétitions inversées séparées par une région asymétrique de 8 bp. Chacune des répétitions bordures est un site de liaison à l'enzyme. La région asymétrique de 8 bp définit l'orientation du site loxP. Lorsque les sites sont parallèles (dits en répétitions directes), le produit de la réaction est l'excision; lorsque les sites sont antiparallèles (dits en répétitions inversées), le produit de la réaction est l'inversion de la région bordée des sites loxP.

Les études sur des sites loxP mutés démontrent que les séquences bordures gauche et droite peuvent ne pas être identiques entre elles. L'équilibre dynamique est déplacé dans la réaction de recombinaison observée in vivo dans le génome de *N. tabacum* entre les sites mutés comme indiqué dans la Figure 36. La réaction est active si chaque site présente une séquence bordure WT et une séquence bordure mutée. En revanche, la réaction est inhibée entre un site WT et un site dont les deux séquences bordures sont mutées ce qui provoque le déplacement de l'équilibre dynamique vers l'intégration. La fidélité de la recombinaison n'est pourtant pas toujours respectée dans la réaction entre les sites lox mutés, et la restitution exacte des sites lox mutés en aval (Figure 36) n'est pas systématique (Odell and Russell, 1994), (Albert *et al.*, 1995).

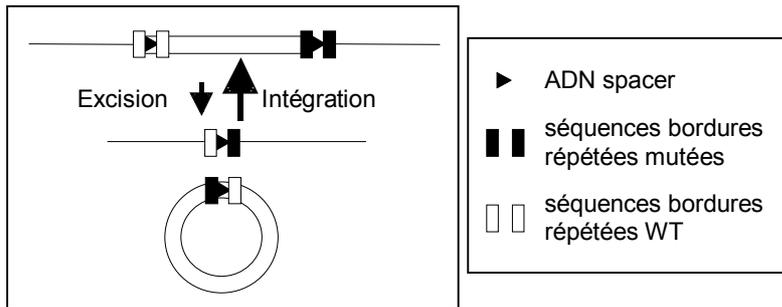


Figure 36. Déplacement de l'équilibre de réaction catalysée par la recombinase Cre entre les sites mutés.

#### 1.4.1. Fonctionnement du système Cre/lox dans des systèmes hétérologues

La recombinase Cre peut catalyser l'excision, l'intégration et l'inversion dans la levure, dans les cellules de mammifères, dans la souris ainsi que, chez les plantes, dans *N. tabacum*, *A. thaliana* et *Z. mays*. Il est facile d'utiliser des substrats extrachromosomiques afin de mettre en évidence la fonction de l'enzyme Cre dans un hôte hétérologue (Odell and Russell, 1994).

L'étude de Dale et collaborateurs sur *N. tabacum* (Dale and Ow, 1990) utilise différents plasmides substrats portant le gène de la luciférase (*luc*) et le vecteur pMM23 qui porte le gène *cre* placé sous le contrôle du promoteur 35S de CaMV (dans ce chapitre, nous l'appellerons p35SCre). Les chercheurs étudient la réaction dans les protoplastes de tabac co-transformés par le vecteur p35SCre et par un plasmide portant deux sites loxP ou avec deux plasmides qui en porte un chacun. Le gène *luc* est restitué par inversion (1) de la séquence codante bordée de sites loxP ou par concaténation (2) de deux plasmides portant chacun un fragment complémentaire avec un site loxP. L'excision (3) du gène *luc* d'un troisième plasmide est confirmée par analyse de séquence. Ainsi, l'étude a montré la possibilité de réaliser les trois réactions (1-3) catalysées par la recombinase Cre dans un système végétal.

#### Sites de reconnaissance endogènes dans les systèmes hétérologues

Il existe une possibilité que les génomes complexes contiennent des séquences qui pourraient être reconnues par les recombinases spécifiques d'un site (Odell and Russell, 1994). Chez *S. cerevisiae*, les événements de recombinaison avec des sites endogènes sont rares et ne sont probablement ni problématiques ni utiles. Les études effectuées à ce jour sur les animaux ne témoignent pas d'effet mutagène de l'expression du gène *cre* (ni d'ailleurs du gène FLP ou R). La vigilance lors de l'utilisation du système Cre/lox dans un hôte hétérologue doit tout de même être de mise. L'enzyme pourrait, en effet, reconnaître des sites n'ayant pas la même séquence mais une structure similaire (cf. même le site loxB chez *E. coli*). Les réactions entre ces sites et entre ces sites et les vecteurs de transformation, catalysée par la recombinase pourraient être mutagènes. La possibilité de ces réactions est corroborée par l'observation que certaines plantes transgéniques produisant la recombinase Cre présente un phénotype altéré (D. Schaefer, communication personnelle).

#### 1.4.2. Techniques développées

Le développement du système Cre/lox chez la souris est allié au développement de promoteurs inductibles ou spécifiques d'un tissu ainsi qu'à la mutagenèse ciblée (Muller, 1999; Odell and Russell, 1994; Ow and Medberry, 1995; Sauer, 1998; Vergunst and Hooykaas, 1999). Les principales techniques sont l'activation différée de gènes, l'inactivation conditionnelle de gènes, la production de réarrangements chromosomiques et l'excision de gènes marqueurs.

Le développement de telles techniques pour les végétaux est sensiblement en retard. La raison principale est l'impossibilité de la mutagenèse ciblée chez les angiospermes qui permettrait le remplacement d'un allèle par un allèle étiqueté à l'aide de sites loxP. De surcroît, l'utilisation de promoteurs conditionnels est peu développée chez les végétaux - par ailleurs, cette situation sera commentée dans le chapitre suivant. Chez les végétaux, la réaction d'intégration catalysée par la recombinaise Cre (Figure 35, page 83) est utilisée pour une insertion ciblée de transgènes portant un site loxP dans un locus préalablement étiqueté par un site loxP. Néanmoins, une étude de l'intégration de l'ADN-T portant un site loxP par recombinaison avec un site loxP préalablement introduit dans un chromosome chez *A. thaliana* révèle une complication imprévue liée à des événements d'intégration non ciblée (Vergunst *et al.*, 2000). En effet, le mécanisme d'intégration illégitime par NHEJ chez *A. thaliana* est tellement efficace que la majorité des événements d'intégration ont lieu ailleurs que dans le site loxP chromosomique. La recombinaison catalysée par l'enzyme Cre entre les sites loxP entraîne des événements de délétion, d'inversion et de translocation de segments chromosomiques. Ce phénomène peut servir d'outil intéressant pour cet organisme dont la carte complète du génome sera bientôt disponible mais rend précaire l'utilisation du système pour favoriser l'intégration ciblée. Cette étude est particulièrement révélatrice de la difficulté de la mutagenèse ciblée et de l'utilisation du système Cre/lox chez les végétaux.

##### Excision de gènes marqueurs de sélection

La réaction d'excision catalysée par la recombinaise Cre permet d'exciser les marqueurs de sélection. La perte de la résistance à l'antibiotique permet l'isolation des événements d'excision. L'absence de signaux d'hybridation ou d'amplification PCR et l'apparition d'un nouveau fragment recombinant prédit confirme la disparition du marqueur du génome. L'analyse de la séquence du site d'excision doit confirmer la précision de la recombinaison et l'absence d'éventuelles mutations induites. Sauer et collaborateurs démontrent une excision précise du gène *LEU2* bordé des sites loxP dans 98% de cellule de *S. cerevisiae* après 24 h d'induction du gène *cre* (Sauer, 1994). Une telle efficacité permet le développement d'une stratégie de "recyclage" de marqueurs chez la levure.

Chez les végétaux, le gène *cre* est introduit dans la souche porteuse d'un gène flanqué de sites loxP par une re-transformation stable ou par un croisement avec une souche exprimant le gène *cre* de façon stable (Odell and Russell, 1994). L'expression Cre transitoire est également possible. **L'efficacité du traitement (en %)** est la valeur obtenue par une division du nombre de plantes dans lesquelles le marqueur est excisé par le nombre de plantes ayant reçu le gène marqueur bordé des sites loxP et la recombinaise Cre. L'excision efficace a été démontrée chez *N. tabacum*, *A. thaliana* et *T. aestivialis* (Dale and Ow, 1991), (Russell *et al.*, 1992), (Bayley *et al.*, 1992), (Gleave *et al.*, 1999). Nous passons en revue ces études dans la discussion, voir le Tableau 11, page 104. L'innovation de l'étude du groupe de Gleave (Gleave *et al.*, 1999) est la

tentative de comparer l'efficacité de l'expression du gène *cre* transitoire d'un vecteur non intégré et de celle d'un vecteur intégré dans le génome. Les chercheurs suppose qu'en absence de sélection des événements d'intégration, le gène *cre* est perdu. Malheureusement, le gène *cre* s'intègre dans le génome dans la majorité des clones dans lesquels le marqueur bordé des sites loxP est excisé. Gleave et collaborateurs affirment qu'une expression transitoire d'un vecteur est difficile à mettre en œuvre chez la majorité des angiospermes puisque le vecteur peut facilement s'intégrer dans le génome par recombinaison illégitime. De plus, il est difficile d'établir une lignée à partir d'une seule cellule (protoplaste, par exemple) dans laquelle l'excision a eu lieu pour éviter l'expression chimérique du gène *cre*. Ces risques ne sont pas importants chez *P. patens* car l'intégration des vecteurs par recombinaison illégitime est extrêmement rare (1 pour 10<sup>5</sup> régénérants) alors que la réplication et l'expression des éléments épisomiques sont efficaces (le gène conférant une résistance à l'antibiotique est exprimé de façon transitoire dans 10 % des régénérants, le gène codant pour la protéine GFP dans 30% des régénérants) (Schaefer, 1994). La régénération de colonies clonales à partir d'un protoplaste dans des conditions physiologiques est aisée.

#### Réduction du nombre de copies du vecteur

La présence du marqueur de sélection en plusieurs copies dans le génome peut modifier les produits attendus de la réaction. Ceci peut être dangereux si les copies du marqueur se situent dans des locus différents. Une réaction entre eux peut amener à des événements de translocation, de délétion ou d'inversion de segments chromosomiques.

En revanche, ceci peut être un avantage s'il est présent dans le même locus sur les répétitions directes du vecteur intégré: dans ce cas-ci, l'excision du marqueur peut s'accompagner d'une réduction du nombre de copies du vecteur à une seule copie (Srivastava *et al.*, 1999). La mutagenèse d'un locus par intégration d'une seule copie du vecteur transformant est théoriquement plus avantageuse que celle de plusieurs copies. En effet, l'insertion de structures de haut poids moléculaire, que sont les répétitions du vecteur, risquent davantage d'apporter des modifications à la structure de la chromatine, indésirables pour une mutagenèse de précision. Il existe d'autres raisons pour éviter les répétitions de transgènes dans le génome: silençage par expression sens ou anti-sens, instabilité mitotique, méthylation d'ADN (Fagard and Vaucheret, 2000; Matzke *et al.*, 1995). En effet, chez *P. patens*, nous avons observé des événements de silençage, d'instabilité mitotique et de méthylation de l'ADN transformant intégré en répétitions multiples ((Kammerer and Cove, 1996), Schaefer, données non publiées). L'intégration d'un vecteur de transformation dans le génome de *P. patens* a lieu généralement dans un seul locus et en répétition directes. Dans cette étude, nous pouvons étudier la possibilité de réduire le nombre de copies intégrées à une seule copie par génome.

Lorsque les sites loxP naturels sont utilisés, il est dangereux d'appliquer la technique d'excision de marqueur plus d'une fois dans la même souche. L'excision du premier marqueur laissera un site loxP (loxP-1). L'utilisation d'un deuxième marqueur "floxé" dans la même souche introduira deux sites loxP supplémentaires (loxP-2 et -3). Si l'enzyme Cre est introduite dans la cellule, elle pourrait se servir des sites loxP-1 et loxP-2 ou bien des sites loxP-1 et loxP-3 comme substrat. Si les sites loxP-2 et -3 se trouvent sur un autre chromosome que le site loxP-1, la réaction entre eux mènera à des réarrangements chromosomiques (voir plus haut). Même s'ils se trouvent sur le même chromosome, l'excision ou l'inversion du segment chromosomique qui les sépare peut être délétère et nuire à l'analyse effectuée (à moins que ce ne soit l'effet recherché (Medberry *et al.*, 1995)). Malgré ce risque, cette technique peut être essayée plus d'une fois dans la même lignée car la recombinase agit

préférentiellement sur des substrats monomoléculaires et la réaction entre loxP-2 et loxP-3 doit être favorisée (cf. plus haut, (Odell and Russell, 1994)).

#### 1.4.3. Utilisation du système Cre/lox chez *P. patens*

Dans ce chapitre, nous cherchons à répondre aux questions suivantes:

- ✓ L'enzyme peut-elle catalyser l'excision d'un marqueur de sélection bordé de sites loxP chez la mousse *P. patens*?
- ✓ L'excision du marqueur a-t-elle lieu par une réaction de recombinaison précise entre les sites loxP flanquants?
- ✓ Quel est le produit de l'excision de plusieurs copies du marqueur "floxé", si elles sont intégrées en répétitions directes? Arriverons-nous à réduire le nombre de copies du vecteur intégrées par cette technique?
- ✓ Après l'excision, une réinsertion ectopique du marqueur peut-elle avoir lieu?
- ✓ Si un marqueur de sélection peut être complètement éliminé d'une souche de *P. patens*, la réaction est-elle suffisamment efficace pour permettre le développement d'une technique courante de laboratoire?
- ✓ L'expression transitoire de l'enzyme Cre est-elle mutagène pour la mousse *P. patens* ?

## 2. Résultats et discussion

Notre premier but a été la démonstration d'une activité du gène *cre* dans le contexte cellulaire de *P. patens*. Nous avons vérifié que l'expression transitoire du gène *cre* chez *P. patens* permet d'observer la réaction d'inversion. Comme substrat, nous avons utilisé le vecteur portant la séquence codante du gène rapporteur *luc* placée à l'envers en aval du promoteur 35S CaMV entre deux sites loxP antiparallèles (35S-loxP-*cul*-loxP) (Dale and Ow, 1990). L'inversion doit restaurer un gène *luc* fonctionnel et permettre son expression transitoire. Nous avons mis en évidence, dans des essais préliminaires, une activité luc induite par une co-expression transitoire des gènes Cre et 35S-loxP-*cul*-loxP (données non publiés). Nous avons alors décidé d'étudier directement l'excision d'un marqueur présent en différent nombre de copies dans un locus chromosomique.

### 2.1. Intégration d'un marqueur de sélection positive bordé de deux sites loxP parallèles dans le génome de *P. patens*

Nous avons introduit un marqueur bordé de sites loxP dans le génome de la mousse *P. patens* par une transformation de la souche WT avec le vecteur pAloxneo présenté (Figure 37, détail dans l'Annexe B, page 140). Le vecteur pAloxneo porte un fragment XbaI- EcoRI de 4 kb du gène *PpAPT* de *P. patens* (chapitre III) dans lequel une région interne de 1.1 kb est remplacée par le gène *nptII* sous le contrôle du promoteur 35S de CaMV (35S *nptII*). Le gène 35S *nptII* est bordé de deux sites loxP parallèles et orientées dans le sens du cadre de lecture du gène *PpAPT*. Le vecteur pAloxneo a donc une structure d'un vecteur de remplacement semblable à celle du vecteur pAΔBneo (chapitre III, et cf. Annexe B, page 139). Le gène NPT est entouré d'un fragment 5' du gène *PpAPT* de 1.3 kb et d'un fragment 3' du gène *APT* de 1.6 kb.

Nous avons produit des souches avec 1, 10 ou plus de copies du vecteur intégrées dans le génome. L'analyse par hybridation Southern et par PCR nous a permis de retenir trois souches porteuses du gène marqueur 35S *nptII* flanqué de sites loxP:

La structure présentée dans la Figure 37 est confirmée par l'analyse moléculaire présentée dans la Figure 38.

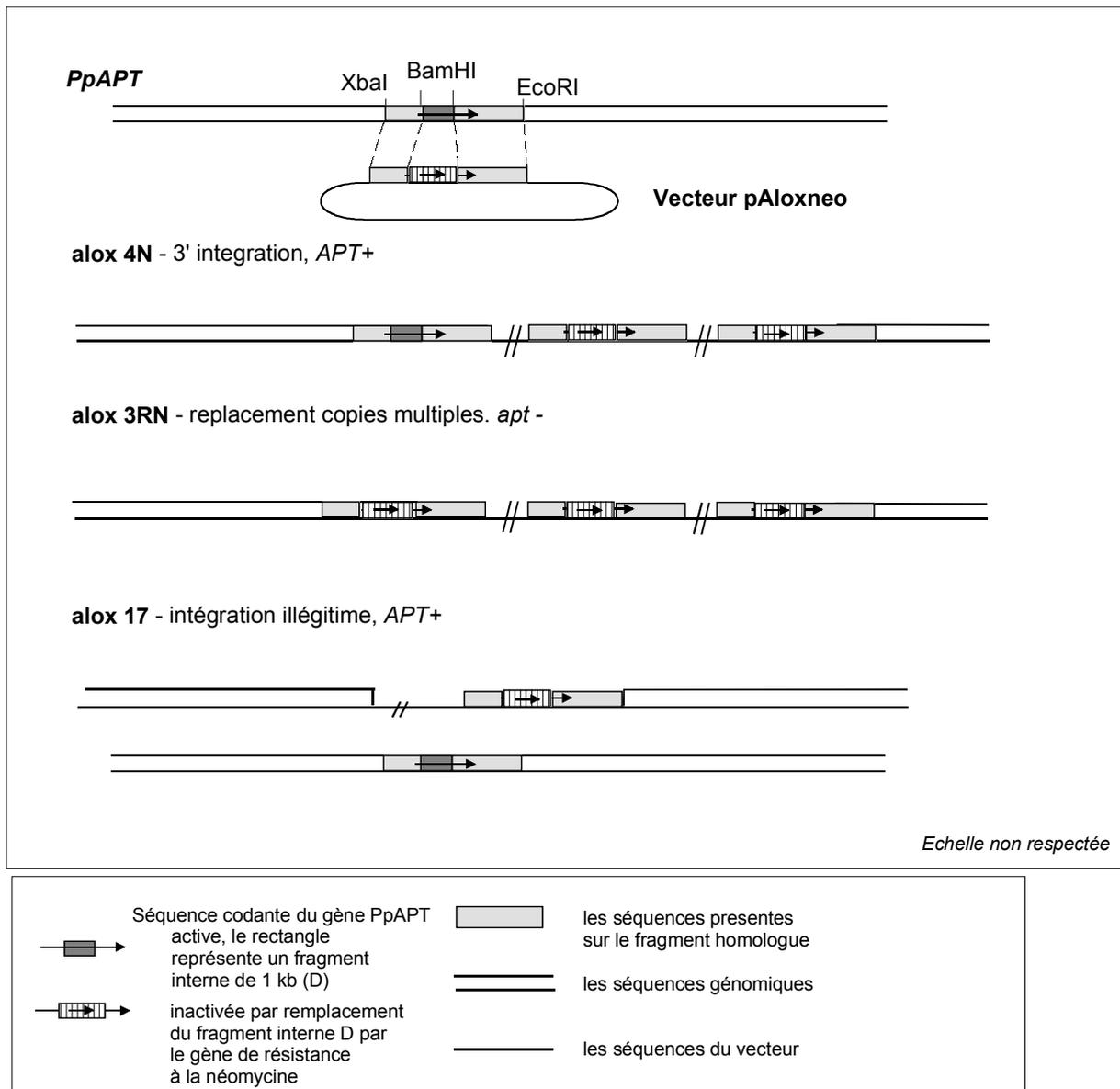


Figure 37. Structure du site d'intégration du vecteur pAloxneo dans les souches alox-17, alox-4N et alox-3RN.

**alox-4N** porte 5 à 10 copies du vecteur intégrées par insertion par recombinaison homologue en 3' dans le gène *PpAPT*. Comme attendu pour ce type d'intégration (chapitre III, page Figure 18), une copie du gène *PpAPT* fonctionnel est présente, le phénotype est par conséquent *APT*<sup>+</sup> Neo<sup>R</sup>.

**alox-3RN** porte plus de 10 copies du vecteur intégrées par remplacement multiple par recombinaison homologue dans le gène *PpAPT*. Le gène *PpAPT* est interrompu et le phénotype est *APT*<sup>-</sup> (DAP<sup>R</sup>) Neo<sup>R</sup>.

**alox-17** porte une copie du vecteur intégrée par recombinaison illégitime dans un locus inconnu (nous l'appellerons locus 17). Le gène *PpAPT* est intact, le phénotype est *APT*<sup>+</sup> Neo<sup>R</sup>.

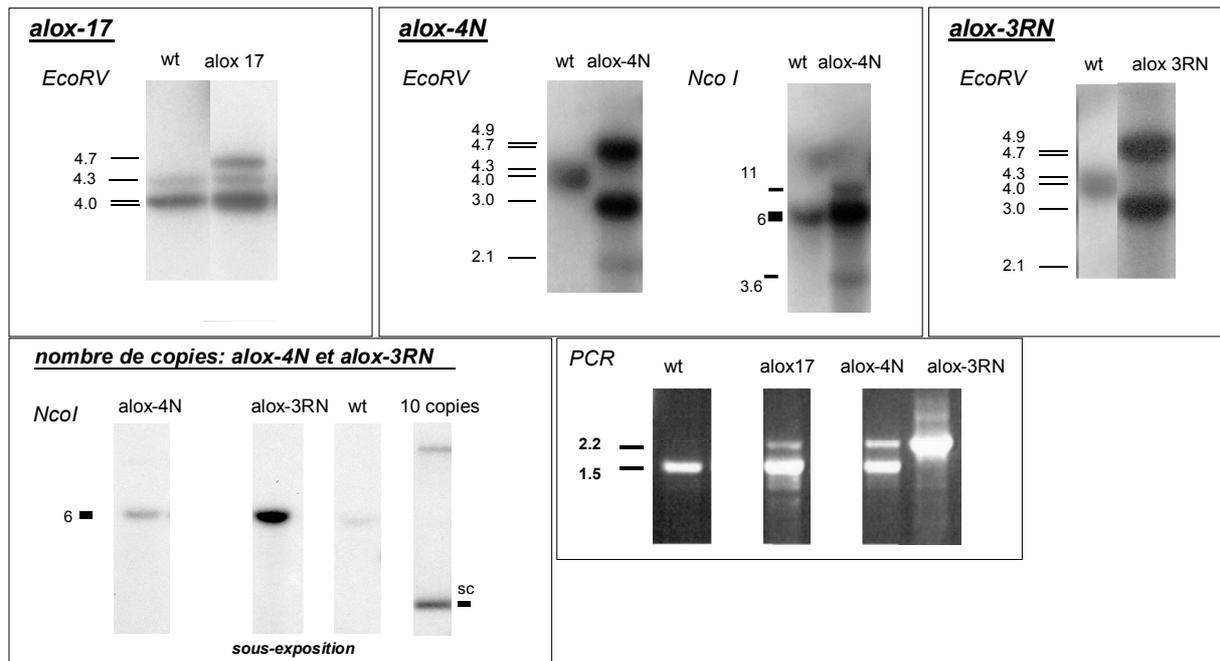


Figure 38. Analyse moléculaire des souches alox-17, alox-4N et alox-3RN.

EcoRV et NcoI - hybridation de l'ADN des souches alox-17, 3RN, 4N digéré avec EcoRV ou avec NcoI avec la sonde A35. sc - plasmide pAloxneo circulaire surenroulé. Le plasmide est chargé sur le gel en quantité correspondant à 10 copies par génome. WT - type sauvage, nous détectons deux fragments EcoRV de 4.3 et de 4.0 kb et un fragment NcoI de 6.0 kb.

PCR - analyse des produits d'amplification PCR avec les amorces ex2-5' et m1-3' sur la matrice de l'ADN des souches alox-17, 4N, 3RN. L'amplification produit un fragment de 1.5 kb sur la matrice du gène *PpAPT* sauvage et un fragment de 2.2 kb sur la matrice du fragment homologue du vecteur pAloxneo.

#### ***alox-17***

EcoRV - l'intégration n'a pas lieu dans le gène *PPAPT* dans la souche alox-17. Les deux fragments WT EcoRV de 4.3 et de 4.0 kb restent intacts et les fragments hybrides prédits pour l'intégration dans le gène *PPAPT* ne sont pas détectés. En revanche, les deux fragments EcoRV hybrides de 4.7 et de 4 kb environ confirment une intégration illégitime. L'absence du fragment EcoRV de 3 kb (apparaît lorsqu'une intégration de plus d'une copie a lieu, cf. alox-4N) et l'intensité des signaux attestent clairement d'une intégration d'une copie unique.

PCR - Les résultats de l'analyse par PCR confirment cette conclusion puisque les deux fragments de 1.5 kb et de 2.2 kb peuvent être amplifiés. Le phénotype est *APT+*, comme attendu.

#### ***alox-4N***

NcoI - Le nombre de copies intégrées est de 5 à 10 (comparé à 10 copies). Nous détectons tous les fragments NcoI attendus pour une insertion 3' (11, 6 et 3.6 kb) en surexposant le film.

EcoRV - La présence des fragments WT 3' de 4.3 kb et de la jonction 5' de 2.1 kb attendus confirment que l'intégration a eu lieu par recombinaison avec le fragment 3' du gène *PPAPT*. La jonction 3' de 4.9 kb également prédite dans les produits de cet événement d'intégration est masquée par le signal de forte intensité à 4.7 kb. L'intensité des deux signaux des fragments du vecteur (4.7 kb et de 3 kb) comparée à celle du signal d'une jonction (2.1 kb) indique qu'environ 5-10 copies se sont intégrées.

PCR - Les résultats de l'analyse par PCR confirment cette conclusion puisque les deux fragments de 1.5 kb et de 2.2 kb peuvent être amplifiés. Le phénotype est *APT+*, comme attendu.

#### ***alox-3RN***

NcoI - Le nombre de copies intégrées est de 10 à 20.

EcoRV - Les signaux d'hybridation avec les fragments de 4.7 et 3 kb de forte intensité confirment une intégration de plus de 10 (probablement  $\geq 20$ ) copies du vecteur. L'absence de la jonction 5' de 2.1 kb exclut l'intégration par recombinaison dans le fragment 3'. L'analyse phénotypique soutenue par l'analyse par PCR confirme que l'intégration du vecteur a certainement interrompu et inactivé le gène *PpAPT* par remplacement à copies multiples. Ceci a provoqué le phénotype *APT- DAP<sup>R</sup>*.

PCR - Aucune copie du gène *PpAPT* intact n'est pas détectable par PCR. Le seul fragment amplifié par PCR avec les amorces ex2-5' et m1-3' est de 2.2 kb provenant de l'amplification sur la matrice du vecteur intégré.

## 2.2. Excision du marqueur de sélection 35S *nptII* intégré dans les souches alox-17, alox-4N et alox-3RN

Nous effectuons l'expression transitoire de la recombinaise Cre en transformant les protoplastes des souches alox-17, alox-4N et alox-3RN par le vecteur d'expression de la recombinaise Cre (p35SCre, cf. Annexe B, page 140). Une période de culture sans aucune sélection d'environ 10 jours permet le développement de colonies isolées d'une taille de 1 à 5 mm en diamètre. Nous vérifions l'absence de l'intégration du gène *cre* dans ces colonies par hybridation Southern (traité plus bas, page 102).

L'efficacité du traitement à la recombinaise dépend de la "compétence" des protoplastes transformés, c'est-à-dire de l'efficacité de l'incorporation du vecteur d'expression dans les protoplastes et d'une transcription efficace du gène *cre*. Le vecteur d'expression p35SCre ne permet pas de sélectionner les colonies issues des protoplastes dans lesquels le vecteur p35SCre est incorporé et le gène *cre* correctement transcrit. Nous utilisons alors un contrôle parallèle afin d'estimer la "compétence" des protoplastes. Nous transformons un échantillon de protoplastes du même lot par le vecteur d'expression de GFP (le gène *rsGFP* sous le contrôle du promoteur 35S, le plasmide psmrsGFP (Davis and Vierstra, 1998)) et observons la protéine GFP produite. Le pourcentage des protoplastes qui produisent la GFP nous sert de mesure de compétence d'un lot de protoplastes donné. Nous l'appelons l'"indice GFP". Si un lot de protoplastes est caractérisé par un indice GFP de 10%, nous supposons que dans le même lot de protoplastes transformés par le vecteur d'expression de la recombinaise Cre, l'expression du gène *cre* a lieu dans au moins 10% des protoplastes régénérés.

L'excision du marqueur 35S *nptII* dans les souches analysées doit conduire à la perte de résistance à la néomycine (Neo<sup>S</sup>, analogue G-418 utilisé). Deux conditions de contrôle servent à vérifier que la perte de la résistance est due à l'excision par recombinaise Cre:

- ✓ les souches alox-17, alox-4N et alox-3RN sont transformées dans les mêmes conditions mais sans le vecteur d'expression de la recombinaise Cre;
- ✓ une souche anexs-111 est transformée dans les mêmes conditions avec ou sans le vecteur d'expression de la recombinaise Cre. Cette souche (chapitre III) porte plusieurs copies du marqueur 35S*nptII* sans aucun site loxP.

Si un de ces traitements produit des colonies sensibles à l'antibiotique, nous devons les analyser pour savoir comment la résistance a pu être perdue.

Un échantillon de colonies régénérées est testé pour la sensibilité à G-418. Le nombre de colonies testées et le nombre de colonies ayant perdu la résistance à G-418 après la transformation des quatre souches sont présentés dans le Tableau 10.

Tableau 10. La perte de la résistance à G-418 observée dans les colonies issues des transformations des souches alox-17, alox-4N, alox-3RN et de la souche de contrôle anexs111.

souche traitée	expérience *	marqueur	indice GFP **	Cre	nombre de colonies testées	nombre de colonies ayant perdu la résistance à G-418	Efficacité = $C / (A \cdot B)$
			<i>A</i>		<i>B</i>	<i>C</i>	<i>D</i>
<i>alox-3RN</i>	1	$\geq 20$ copies	2.5-4%	+	150	<u>0</u>	0%
	1		2.5-4%	-	150	<u>0</u>	
	2		19-20%	+	200	<u>0</u>	
	2		19-20%	-	200	<u>0</u>	
<i>alox-4N</i>	3	$\leq 10$ copies	20%	+	150	<u>6 + 1***</u>	30%
	3		20%	-	150	<u>0</u>	
	4		28-29%	+	150	<u>7 + 2***</u>	
	4		28-29%	-	150	<u>0</u>	
<i>alox-17</i>	5	1 copie	6%	+	200	<u>13</u>	65%
	5		6%	-	130	<u>2</u>	
	6		nt	+	300	<u>5 + 2***</u>	
<i>anexs-111</i>	7	2-3 copies	6-8%	+	220	<u>0</u>	
	7	sans loxP	6-8%	-	135	<u>0</u>	

\* par expérience nous entendons ici un lot de protoplastes d'une souche donnée

\*\* basé sur deux à trois estimations au microscope d'un échantillon

\*\*\* isolés comme sensibles, résistance recouvrée après remise en culture, cf. texte A, B, C, D - Le calcul de l'efficacité est traité à la page 103.

Ces résultats nous permettent de tirer les conclusions suivantes:

- ✓ Les marqueurs qui ne sont pas bordés de sites loxP sont stables. Ni le traitement de transformation ni l'action de l'enzyme Cre (souche anexs-111) ne provoquent de perte de résistance;
- ✓ Les marqueurs bordés de sites loxP restent stables dans les souches alox-4N et alox-3RN. Le traitement de transformation sans ADN ne provoque aucune perte de marqueur dans ces souches;
- ✓ Treize clones sur 300 dérivés de la souche alox-4N ont complètement perdu la résistance à G-418. En plus, nous avons initialement sélectionné 3 clones comme ayant perdu la résistance mais à l'analyse ultérieure, ces clones ont recouvré la résistance;
- ✓ Vingt clones sur 500 dérivés de la souche alox-17 ont perdu la résistance à G-418. Or, 2 clones sur 130 dérivés de la transformation sans ADN de la même souche (n1, n2) ont également perdu la résistance. Nous ne pouvons expliquer cette perte de résistance par une action de la recombinaison.

Aucun clone sur 350 dérivés de la souche alox-3RN n'a perdu la résistance à G-418. Nous proposons deux interprétations suivantes:

- ✓ La souche alox-3RN a un phénotype APRT<sup>-</sup> très prononcé. De plus, la croissance de cette souche sur du milieu normal PPNH4c est plus lente que celle du WT et les cellules restent plus compactes. L'isolation des protoplastes observée au microscope est d'une efficacité inférieure à celle de l'isolation des protoplastes WT, alox-17 ou alox-4N ( $0.3 \cdot 10^6$  contre  $1 \cdot 10^6$  de protoplastes par boîte de protonéma de 5-6 jours). La protéine GFP observée dans les protoplastes transformés avec ce vecteur d'expression apparaît en concentrations

moindres que dans les autres souches (observation microscopique). Nous n'avons détecté aucune production de la protéine GFP dans deux autres transformations des protoplastes de cette souche (non présentées). Nous en concluons que l'expression de l'enzyme Cre dans les protoplastes alox-3RN n'est pas de la même efficacité que dans les souches alox-17 et alox-4N.

- ✓ Le grand nombre de copies du marqueur à exciser (> que 20-30), ainsi que la structure plus encombrante que forment toutes les répétitions du vecteur peuvent rendre l'excision moins rapide. L'activité de la recombinase Cre atteinte dans ces conditions est peut-être insuffisante pour aboutir à l'excision complète des copies du marqueur.

Les clones dérivés de la souche alox-3RN conservent le phénotype Neo<sup>R</sup>, APRT<sup>-</sup>, DAP<sup>R</sup> ; ils ont également une croissance lente et les cellules plus compactes. Nous n'avons pas fait d'analyses moléculaires de ces souches.

Une analyse moléculaire doit confirmer que dans les souches qui ont perdu la résistance à G-418, une excision précise du marqueur de sélection 35S *nptII* a eu lieu.

### 2.2.1. Expression transitoire du gène cre dans la souche alox-17

#### Perte de la résistance après le traitement de contrôle

Il est de prime importance de comprendre pourquoi la résistance à G-418 a pu être perdue dans les clones n1 et n2 dérivés de la souche alox-17 indépendamment de l'enzyme Cre. L'ADN de ces souches digéré avec NcoI est hybridé avec la sonde A35 ou avec la sonde PNeo (digestion NcoI montrée dans la Figure 40, B). L'analyse montre que certains fragments résultant de l'intégration du vecteur pAloxneo sont perdus dans ces souches. Nous détectons les signaux d'hybridation de la sonde PNeo avec un fragment d'environ 1.2 kb dans la souche n1 et d'environ 5 kb (doublet?) dans la souche n2.

Au vu de ces résultats nous pensons que le locus d'intégration du vecteur pAloxneo dans la souche alox-17 est instable. Cette instabilité provient probablement d'événements de recombinaison homologue intrachromosomique dans le locus concerné qui peuvent provoquer une perte du vecteur menant à la sensibilité à G-418 dans 2 clones sur 130 testés (1.5%). Etant donné qu'une seule copie du vecteur est intégrée, il est possible que des régions homologues répétitives ou dédoublées soient présentes dans le locus avant l'intégration. Il n'est pas exclu pour autant qu'une duplication d'un fragment génomique soit associée à l'événement d'intégration du vecteur pAloxneo. Il n'existe pas d'importante identité d'une longueur ininterrompue entre le vecteur et le locus puisque les sondes A35 ou PNeo ne détectent pas d'autres signaux que ceux du vecteur et du gène *PpAPT*.

Nous devons tenir compte de l'instabilité du locus 17 lors de l'analyse des clones Neo<sup>S</sup> dérivés du traitement à la recombinase Cre de la souche alox-17. Nous devons nous attendre à des clones ayant perdu la séquence entière ou partielle du marqueur suite à des événements de recombinaison intrachromosomique d'une fréquence d'environ 2 %.

#### Perte de la résistance dans les clones dérivés de la souche alox-17 traitée à l'enzyme Cre

L'ADN des 20 clones dérivés de la souche alox-17 digéré avec NcoI est hybridé avec la sonde A35. Pour une partie de clones, nous analysons aussi l'ADN digéré avec EcoRV ou SmaI et hybridé avec la sonde A35 ou PNeo (Figure 40). L'analyse permet de répartir les clones dans 4 groupes selon les signaux détectés.

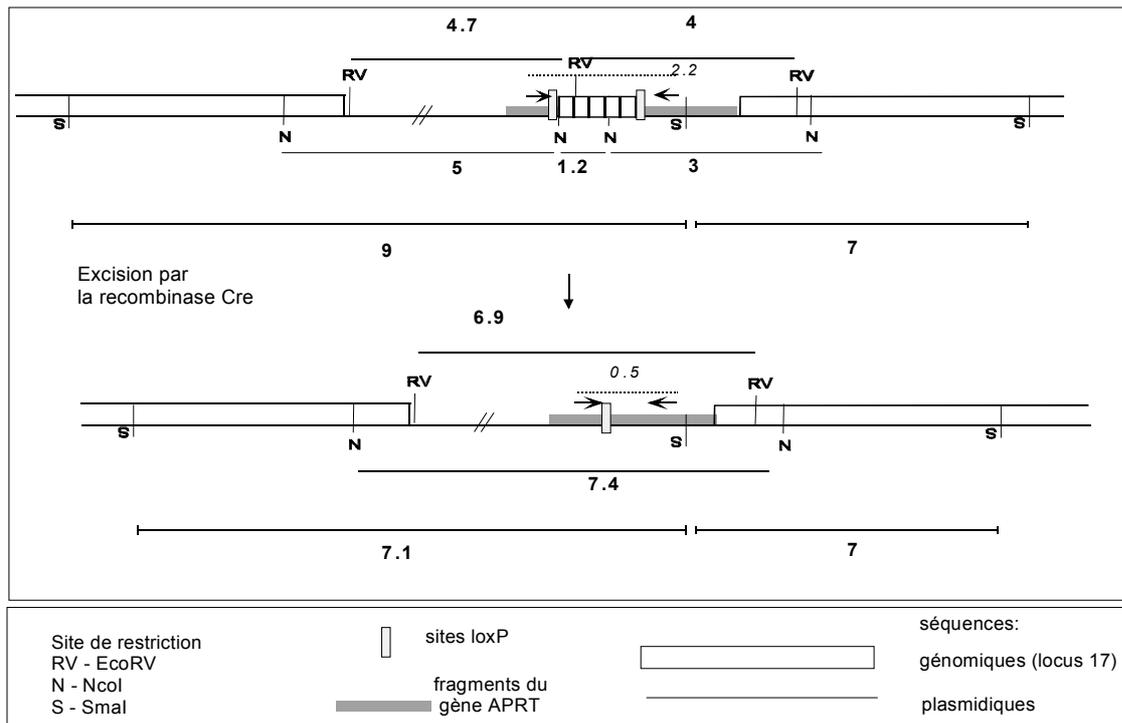


Figure 39. La structure du locus alox-17 avant et après l'excision du marqueur par recombinaison entre les sites loxP. Les flèches indiquent les sites des amorces en 5' et en 3'.

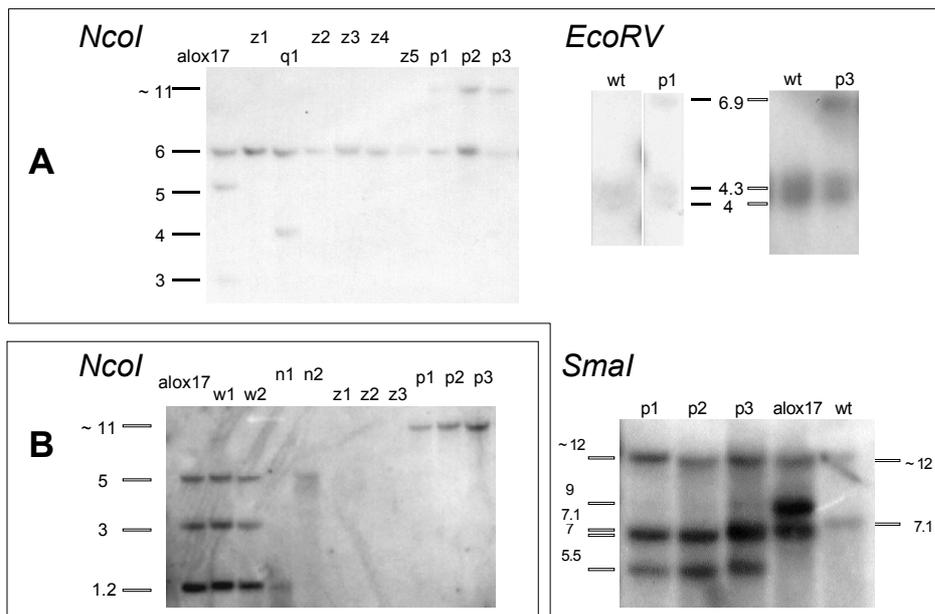


Figure 40. Analyse d'un échantillon des clones après l'expression transitoire de Cre, par hybridation Southern.

WT - souche sauvage, alox-17 souche initiale.

Groupe N - perte de la résistance par recombinaison intrachromosomique sans expression du gène *cre*.

Groupe W - perte de la résistance transitoire due au silençage.

Groupes Z et Q- perte de la résistance par un événement de recombinaison intrachromosomique qui enlève tout (Z) ou une partie (Q) du vecteur, indépendamment de l'action Cre.

Groupe P - perte de la résistance par excision du marqueur. Nous traitons dans le paragraphe suivant de la taille inattendue du fragment NcoI de 11 kb au lieu de 7.4 kb et du fragment SmaI de 5.5 kb au lieu de 7.1 kb.

A - digéré avec NcoI, EcoRV ou SmaI, la sonde est A35 (fragment BssHI-SmaI du gène *PpAPT*, cf. page 51); B - digéré avec NcoI, la sonde est PNeo (contient les séquences plasmidiques pBluescript et le fragment EcoRI du gène *35S nptII*).

L'analyse PCR est compliquée par la mauvaise amplification du fragment de 2.2 kb de la souche alox-17; cette amplification n'est pas reproductible. Néanmoins, le fragment de 0.5 kb qui est attendu après l'excision du marqueur est amplifié dans l'ADN de tous les clones testés (8) de la série *p* (non montré).

Nous répartissons les souches (20 souches) ayant perdu la résistance à l'antibiotique après l'expression transitoire du gène *cre* dans les 4 groupes suivants:

<i>Z</i>	<i>Q</i>	<i>P</i>	<i>W</i>
7	1	10	2

Nous présentons un échantillon des clones de chaque groupe dans la Figure 40 et nous les expliquons dans les paragraphes suivants.

*L'existence du groupe Z et d'un clone Q confirme que le locus 17 est instable*

Dans 7 clones sur 500 testés, la perte de tout le locus d'intégration du vecteur mène à la perte de la résistance (la Figure 40 montre les clones z1 à z4). Nous postulons que des événements de recombinaison intrachromosomique provoquent l'élimination de ces fragments dans le locus 17. Quant au clone q1, un événement de recombinaison intrachromosomique s'est peut être produit entre une partie du vecteur (fragment 3' de *PpAPT*) et un autre fragment local homologue, ce qui provoque la modification observée du RFLP. Nous n'avons pas analysé ce clone plus en détail.

Ces événements se produisent à une fréquence de 1.4%. Celle-ci correspond à la fréquence des événements du même type qui ont lieu sans la recombinaise Cre dans les clones n1 et n2 (1.5%). Nous en déduisons que ces événements ne sont pas stimulés par l'action de l'enzyme Cre et en sont indépendants. Nous croyons que l'instabilité observée est propre au locus 17, qu'elle soit occasionnée par l'événement d'intégration du vecteur pAloxneo ou non. Le fait que nous n'observons aucune perte de la résistance à G-418 dans les autres souches après le traitement de contrôle (sans la recombinaise Cre) confirme cette hypothèse. En effet, nous avons testé 130 clones pour la souche anexs-111, 300 clones pour la souche alox-4N, 350 clones pour la souche alox-3RN sans observer de perte de résistance.

*Le groupe P: l'excision du marqueur par recombinaison entre les sites loxP est catalysée par l'enzyme Cre mais elle est accompagnée d'événements de recombinaison intrachromosomique*

L'excision attendue du marqueur élimine deux sites NcoI et un site EcoRV appartenant au vecteur. Le nombre des fragments détectables avec les sondes utilisées doit diminuer. De plus, le fragment SmaI caractéristique du locus doit diminuer de taille d'environ 1.7 kb. La même diminution de 1.7 kb est attendue pour le fragment amplifié par PCR entre les sites d'appariement des amorces ex2-5' et m1-3'.

De prime abord, c'est ce que nous observons dans les clones p1-p10. Dans tous ces clones, les deux sites NcoI sont perdus. Par conséquent, le nombre des fragments détectables avec les sondes A35 et PNeo diminue comme prévu. La disparition du site EcoRV vérifiée dans 2 clones le confirme. De plus, la taille du fragment SmaI diminue dans les clones testés (3). La taille du fragment amplifié par PCR entre les sites d'appariement des ex2-5' et m1-3' sur le vecteur intégré diminue de 1.7 kb. Le fragment amplifié est donc de 0.5 kb. Le fragment

amplifié sur l'ADN d'un des clones a été sous-cloné et sa séquence analysée. Elle correspond parfaitement à l'excision attendue (voir plus bas).

Etonnamment, l'analyse révèle deux irrégularités:

- ✓ La taille prédite du fragment NcoI du locus modifié remplaçant les trois fragments du locus 17 de 5, 1.2 et 3 kb est de 7.5 kb ( $5 + 1.2 + 3 - 1.7$ ). Or, la taille observée du fragment détecté avec la sonde A35 dans tous les clones analysés (confirmée par la sonde PNeo dans 4 clones) est clairement de 10-11 kb.
- ✓ La taille du fragment SmaI dans le locus modifié est réduite comme prévu, mais de plus que de 1.7 kb! Au lieu de 7.2 kb, ce fragment mesure 5.5 kb.

Notre explication (Figure 41, page 97) se base sur les points importants suivants:

- ✓ La taille des fragments EcoRV correspond aux prédictions. Le site EcoRV du côté 5' du vecteur appartient à l'adaptateur du vecteur. Les sites NcoI et SmaI du côté 5' du vecteur ne lui appartiennent pas, ils se trouvent dans le locus 17;
- ✓ La taille et la séquence du fragment PCR, qui contient le site d'excision, amplifié sur le vecteur intégré correspond aux prédictions;
- ✓ L'instabilité du locus alox-17 est probablement indépendante de l'action de la recombinaison.

Nous postulons que les irrégularités proviennent des événements de recombinaison intrachromosomique. Il est possible qu'il y ait plus de répétitions homologues que nous montrons dans la Figure 41 mais nous n'en supposons que trois. Les événements de recombinaison homologue entre les fragments s1 et s2 sont probablement fréquents mais il ne sont pas sélectionnables *per se*. Nous les avons observés à une fréquence de  $\approx 1.5\%$  (des clones testés) uniquement parce qu'ils ont accompagné l'action de l'enzyme Cre. En revanche, les événements de recombinaison homologue entre les fragments s1 et s3 ou entre s2 et s3 (Figure 41, D), en étant probablement tout aussi fréquents (1.4%), peuvent être mis en évidence puisqu'ils mènent à la perte de la résistance à G-418.

*L'existence du groupe W indique que notre crible basé sur la perte de la résistance n'est pas strict*

Nous avons isolé les clones w1 et w2 comme sensibles à l'antibiotique au premier crible. A la première remise en culture, ces clones se sont montrés sensibles à G-418. Or, après le stockage à  $+4^{\circ}\text{C}$  pendant environ trois mois, le matériel de ces clones remis en culture se développe sur le milieu contenant de l'antibiotique. Le reprise de croissance après le stockage est d'abord plus lente que celle du contrôle positif (alox-17) mais après une deuxième re-mise en culture, ces clones ont complètement recouvert la résistance à l'antibiotique. D'après l'analyse Southern de ces clones, le locus 17 n'a subi ni l'excision du marqueur par la recombinaison Cre ni l'excision d'un segment génomique du locus 17 (Figure 40). Il est possible que la perte de résistance transitoire dans ces clones soit associée au silençage, mais n'ayant aucune preuve nous en restons aux hypothèses.

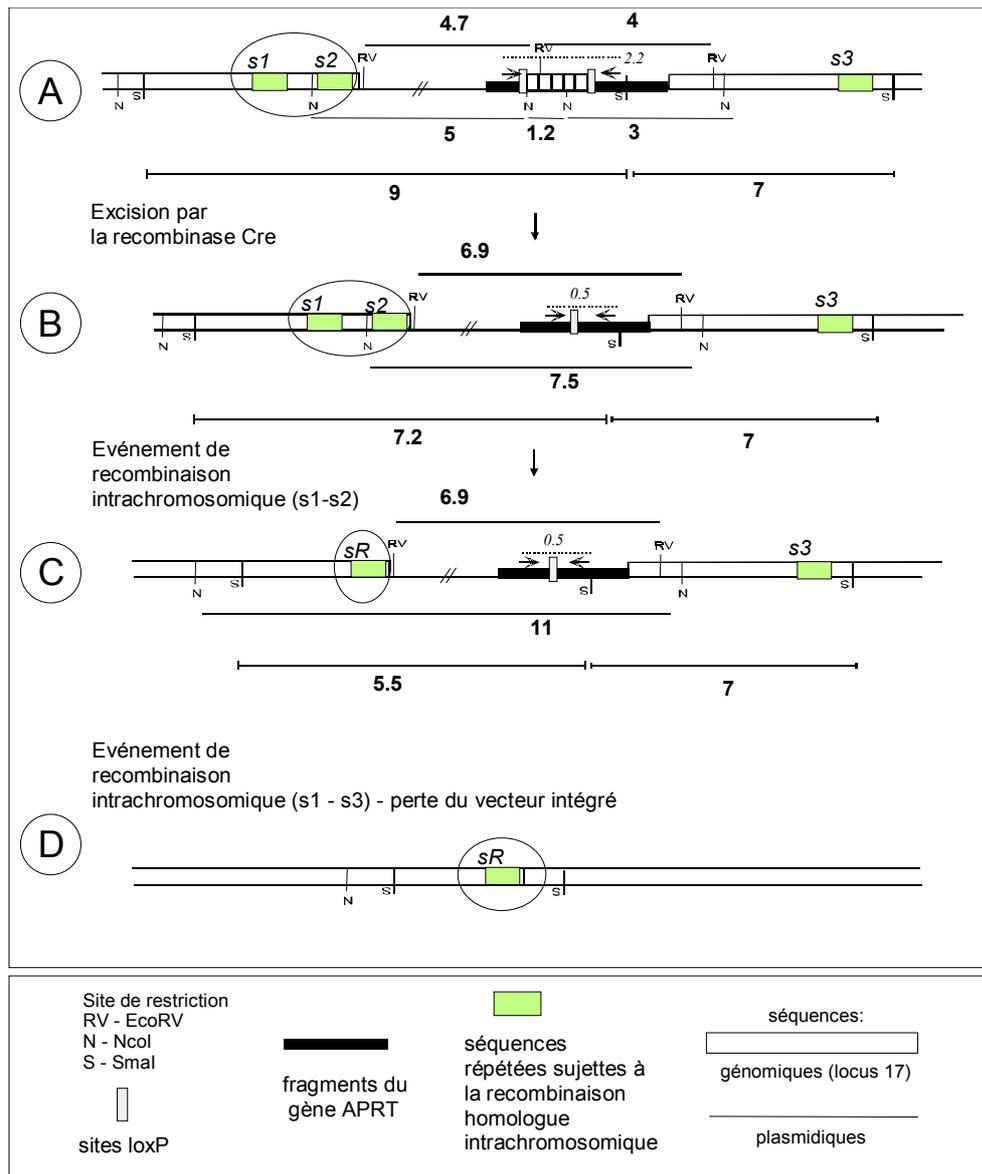


Figure 41. La recombinaison homologue intrachromosomique provoque une déviation du RFLP NcoI et SmaI prédit.

A - L'intégration d'une copie du vecteur pAloxneo dans le locus 17. Les sites putatifs de recombinaison intrachromosomique sont s1, s2 et s3.

B - La structure attendue après l'excision du marqueur.

C - La recombinaison entre les sites s1 et s2 mène à la perte du site NcoI qu'ils entourent (sR - site recombiné). Les fragments EcoRV détectés et le fragment amplifié par PCR n'en sont pas affectés, contrairement au RFLP NcoI et SmaI. Le RFLP EcoRV et l'amplification par PCR du fragment conformes aux prédictions suggèrent qu'une excision correcte a lieu sans que la séquence du vecteur même ou la séquence génomique en 3' soit altérée. Une augmentation de taille du fragment NcoI (de 7.4 kb prévu à 11 kb observé) peut être provoquée par une perte du site NcoI en 5' qui se trouve dans le locus 17. La disparition d'un fragment portant ce site doit en effet expliquer la diminution du fragment SmaI par rapport à la taille prévue. Etant donné que le locus est instable en lui-même, nous pouvons en déduire qu'un événement de recombinaison homologue intrachromosomique en 5' du site d'insertion du vecteur accompagne l'excision du marqueur par l'action de la recombinaise.

D - Un événement de recombinaison homologue intrachromosomique entre les sites s1 et s3. provoque une perte du vecteur intégré. Nous l'observons dans les souches du groupe W, dans lesquelles nous ne détectons plus les fragments du vecteur intégré.

### Souche alox-17 : conclusion

Nous avons montré que l'action de l'enzyme Cre sur les sites loxP dans la souche alox-17 excise le marqueur mais l'instabilité du locus empêche l'interprétation claire des résultats. Afin de vérifier si cette instabilité est due au locus même et afin de prouver qu'elle n'est pas provoquée par la présence des sites loxP, nous nous sommes concentrés sur l'analyse des clones dérivés de la souche alox-4N.

#### *2.2.2. Expression transitoire Cre dans la souche alox-4N*

Le traitement de contrôle de la souche alox-4N ne produit aucun clone Neo<sup>S</sup>. Nous concluons que le locus alox-4N n'est pas instable contrairement au locus 17. La perte de la résistance dans les clones dérivés après le traitement à l'enzyme Cre proviennent donc de l'excision du marqueur. La réaction que l'enzyme Cre peut toutefois catalyser dans la souche alox-4N peut avoir lieu entre n'importe quelle paire de sites loxP. La réaction entre les sites loxP distants ou flanquant plusieurs copies doit provoquer une diminution du nombre de copies du vecteur intégré. L'excision complète ne laisse qu'une copie du vecteur (Figure 42, A).

Nous avons hybridé avec la sonde A35 ou PNeo l'ADN de 12 clones dérivés de la souche alox-4N digéré avec NcoI ou EcoRV. La Figure 43 montre cette analyse. En plus, une excision précise du marqueur par recombinaison entre les deux sites loxP diminue la taille du fragment amplifié sur la matrice du vecteur entre les amorces ex2-5' et m1-3' d'environ 1.7 kb. Le fragment de 0.5 kb attendu est amplifié dans l'ADN de tous les clones Cre4N-01 à 12 (non présenté).

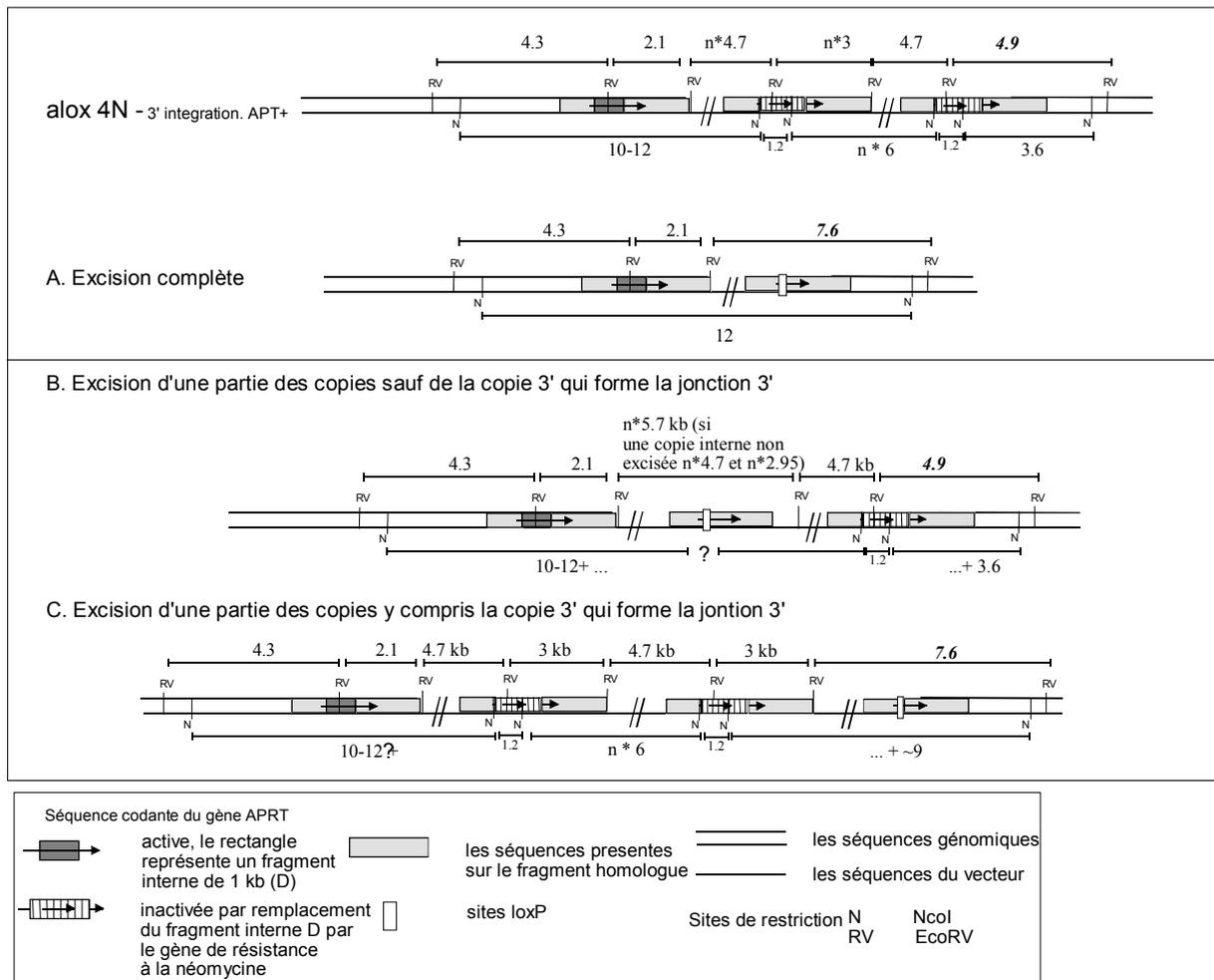


Figure 42. L'allèle alox-4N du locus *PPAPT* avant et après l'excision.

A - L'excision catalysée entre les sites loxP distants.

B et C - La réaction peut éliminer seules quelques copies du vecteur ou du marqueur. Nous vérifions l'excision du marqueur de la copie 3' qui forme la jonction 3' par l'analyse RFLP EcoRV et NcoI.

B - la copie 3' distante contient le marqueur 35S *nptII*. Le fragment EcoRV est de 4.9 kb et le fragment NcoI est de 3.6 kb.

C - le marqueur 35S *npt* est excisé de la copie 3' distante; le fragment EcoRV est de 7.6 kb et le fragment NcoI de 3.6 kb est remplacé par un fragment de plus de 9 kb;

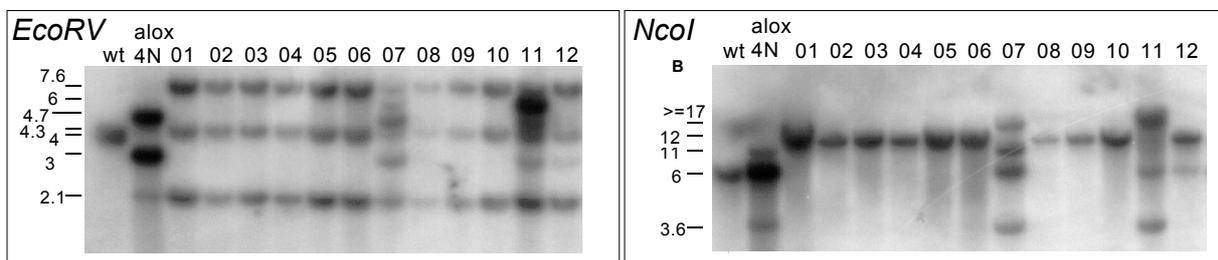


Figure 43. Démonstration de l'excision du marqueur par l'action de la recombinase Cre. Analyse d'hybridation Southern de l'ADN des clones Cre4N 01-12 digéré avec NcoI ou EcoRV et hybridé avec la sonde A35.

WT - souche sauvage; alox-4N - la souche traitée.

Les souches 01 à 06, et 08-10 montrent une excision exacte de toutes les copies comme le prévoit le schéma Figure 42, A.

Les souches 07, 11 et 12 ont été isolée par une perte transitoire de la résistance à G-418; la Figure 42 (B et C) explique le RFLP observé lorsqu'une partie des copies du marqueur et du vecteur est excisée.

L'ensemble de l'analyse nous permet de tirer les conclusions suivantes:

*Une perte complète de la résistance à G-418 de 9 clones est corrélée avec l'excision complète de toutes les copies du marqueur 35S nptII bordées de sites loxP.*

Dans 9 souches (Figure 43 les souches 01 à 06 et 08 à 10), l'enzyme Cre a catalysé une excision entre les deux sites loxP extrêmes (distants). Il est possible que l'excision ait lieu par plusieurs étapes intermédiaires: les marqueurs sont excisés laissant un site loxP par marqueur. La réaction est ensuite catalysée entre les sites restants ce qui réduit le nombre de copies du vecteur. Le produit final des deux enchaînements de réactions est une copie du vecteur avec un seul site loxP.

*Une perte partielle de la résistance à G-418 dans 3 clones est corrélée avec l'excision partielle des copies du vecteur et du marqueur.*

Trois clones (Figure 43, les souches 07, 11 et 12) sont isolés probablement grâce au silençage du marqueur lors de la sélection primaire. Les copies qui restent s'expriment lors de la deuxième sélection après une période de culture d'où une survie que nous détectons sur le milieu sélectif. Nous n'avons pas établi l'ordre des copies excisées ou non dans le locus alox-4N.

Nous trouvons deux types de fragments hybrides en 3': avec et sans le marqueur de sélection dans les clones Cre4N-07 et Cre4N-11 Ceci est suggéré par la présence de deux types de RFLP (Figure 42, B et C). En revanche, le clone Cre4N-12 ne possède qu'un type de la copie 3' distante (RFLP décrit par la Figure 42, C). En revanche, l'intensité du signal correspondant aux répétitions internes contenant 35S npt est plus faible que celle d'un fragment unique (*PpAPT* par exemple). En plus, le nombre de copies 35S *nptII* estimé à l'aide de l'intensité de signal ne correspond pas au nombre suggéré par la taille et le nombre de fragments EcoRV et NcoI. En effet, d'après l'analyse des fragments EcoRV, il ne peut y en avoir plus d'un mais d'après l'analyse des fragments NcoI, il doit y en avoir au moins deux. Ceci suggère que le nombre de copies contenant le 35S *nptII* varie selon les parties de la colonie.

La résistance partielle à G-418 pourrait être expliquée tout simplement par la contamination des trois clones par la souche initiale alox-4N. Une faible amplification par PCR avec les amorces ex2-5' et m1-3' du fragment de 2.2 kb dans certaines réactions pourrait suggérer la même explication. Néanmoins, si c'était le cas, cette contamination est extrêmement faible, vu l'intensité de signal d'hybridation correspondant aux fragments spécifiques à l'excision dans les souches Cre4N-07 et 11 et l'absence totale de signaux spécifiques à la souche alox-4N de l'ADN du clone Cre4N-12. L'amplification par PCR étant plus sensible à la présence d'infime quantité de l'ADN alox-4N, nous tirerons nos conclusions principalement de l'analyse Southern.

Autrement, il n'est pas exclu que l'enzyme Cre ait été active dans les protoplastes après leurs premières divisions. Ceci a pu provoquer un mosaïcisme dans le type de l'excision des copies du vecteur.

*L'excision n'est pas accompagnée d'événements non prédits*

Une observation bien plus importante est que la recombinase Cre n'a produit aucun événement de recombinaison imprévu détectable dans la souche alox-4N. Ceci soutient notre hypothèse que les événements de recombinaison intrachromosomique observés dans la souche alox-17 sont caractéristiques du locus d'intégration 17 et sont indépendants de l'action de l'enzyme Cre.

Les copies du vecteur excisées ne se sont pas intégrées ailleurs dans le génome puisque nous ne détectons aucun signal supplémentaire non expliqué dans aucun des clones.

### **2.2.3. Confirmation de la séquence de la région dans laquelle l'excision ne doit laisser qu'un site loxP**

Nous avons analysé la séquence du fragment de 0.5 kb amplifié entre les sites d'appariement des amorces ex2-5' et m1-3' (cf Figure 37, page 89) dans l'ADN d'un clone dérivé de la souche alox-17. Elle contient un site loxP dans l'orientation attendue. Il se trouve sur l'adaptateur utilisé pour le clonage du marqueur de sélection dans le fragment *PpAPT*. Aucune mutation n'a été constatée lors de la comparaison de la séquence obtenue avec les séquences du gène *PpAPT* ou de l'adaptateur utilisé.

### **2.2.4. Intégration du vecteur d'expression de l'enzyme Cre dans le génome de *P. patens* et l'absence d'effets mutagènes de l'expression du gène *cre***

Nous avons introduit le vecteur d'expression de l'enzyme Cre dans le génome de *P. patens* par une co-transformation avec le vecteur pGL108. Le vecteur pGL108 s'intègre de façon efficace dans le génome par recombinaison avec le locus  $\lambda$ 108, de fonction inconnue (Schaefer and Zryd, 1997). Ce vecteur porte un marqueur de sélection 35S *hpt* conférant la résistance à l'hygromycine. Lorsque la mousse est transformée par ce vecteur mélangé avec le deuxième vecteur p35SCre, celui-ci s'intègre dans le génome de façon stable. Nous supposons que l'intégration passe par la formation d'un produit intermédiaire de recombinaison extrachromosomique entre les vecteurs qui partagent environ 3 kb d'homologie au vecteur de base pUC.

Nous avons ainsi transformé la souche alox-17 et isolé les transformants intégratifs par la résistance à l'hygromycine à une fréquence RTF d'environ  $4 \cdot 10^{-4}$ . Les 35 transformants intégratifs isolés ont perdu la résistance à G-418. Un tiers environ de ces souches ont perdu la résistance seulement après 10 jours de culture sur PPNH4. Nous supposons que l'expression du gène *cre* intégré est responsable de cette perte différée.

L'ADN de 8 clones (108Cre17-3, 8, 13, 18, 23 et 28) digéré avec NcoI a été hybridé avec la séquence codante du gène *cre* isolée sous forme de fragment EcoRV-SphI du vecteur p35SCre (sonde **creCDS**). Les signaux d'intensité variable ont confirmé l'intégration du vecteur p35SCre en nombre de copies différent mais n'ont pas été analysés plus loin. Sachant que l'analyse moléculaire des clones issus de la souche alox-17 est compliquée par des événements de recombinaison homologue, nous n'avons pas tenté de déterminer dans quelle mesure la perte de résistance a été provoquée par l'action de la recombinase. Compte tenu des conclusions tirées sur la souche alox-17, nous supposons que l'activité de l'enzyme Cre est responsable de la perte de la résistance dans une partie de ces clones.

Les 8 clones intégratifs 108Cre17 ont servi de contrôle positif pour l'analyse d'une éventuelle intégration du vecteur p35SCre lors des essais d'expression transitoire. Le vecteur p35SCre peut, en effet, s'intégrer dans le génome des souches alox-17 ou alox-4N par recombinaison homologue avec le vecteur intégré pAloxneo car les deux vecteurs sont dérivés de pBluescript et possèdent donc 3 kb d'homologie. Nous postulons que cet événement est rare par rapport à la fréquence des souches isolées par perte de la résistance (1 pour  $10^4$  et 1 pour  $10^2$  régénérants, respectivement) et nous ne devons pas le détecter. Afin de le prouver, nous avons digéré l'ADN de 19 clones dérivés des essais d'expression transitoire du gène *cre* dans la souche alox-17 avec NcoI et hybridé avec la sonde **creCDS**. L'absence de signaux d'hybridation dans l'ADN de ces clones atteste que l'expression du gène *cre* est transitoire et que le vecteur d'expression est perdu en tous cas après cette période de culture (le matériel pour l'extraction d'ADN est collecté 1 mois après le traitement). Nous n'avons pas testé l'ADN des clones dérivés de la souche alox-4N avec cette sonde. Néanmoins, comme l'intégration de la cassette d'expression sans les séquences plasmidiques serait fort improbable, l'intégration du vecteur p35SCre dans ces clones se serait révélée par la présence de signaux inattendus d'hybridation avec le vecteur pBluescript. Nous n'en avons détecté aucun.

Nous avons observé la croissance des clones issus des traitements transitoires ainsi que celle des souches ayant intégré le vecteur d'expression Cre dans le génome, sur différents milieux (PPNH4, PPNO3) afin de détecter des éventuelles altérations morphologiques importantes. Ces observations n'ont pas mis en évidence de différences de morphologie et de développement du WT jusqu'à la formation de gamétophores. Nous n'avons pas testé la sporogénèse de ces souches.

### 2.3. Conclusion

L'expression transitoire du gène *cre* permet la production d'une enzyme active dans la cellule *P. patens*. Nous avons démontré que l'enzyme produite peut catalyser la réaction d'inversion d'un fragment bordé de sites loxP antiparallèles présent sur un vecteur extrachromosomique et la réaction d'excision d'un fragment bordé de sites loxP parallèles intégré dans le génome. Nous pouvons mettre cette dernière réaction à profit du développement d'un système de recyclage de marqueurs de sélection. Dans cette étude, nous démontrons une excision efficace du marqueur 35S *nptII*. Le marqueur présent en une copie a pu être excisé à une fréquence de 2% (10 clones obtenus sur 500 traités). Dans la souche testée, une instabilité du locus d'intégration indépendante de l'action de l'enzyme Cre a été observée. Le marqueur présent sur environ 10 copies du vecteur intégré en répétitions directes a pu être excisé complètement dans au moins 3% de clones (9 clones sur 300 testés). Des événements d'excision partielle ont pu être isolés par une résistance réduite à l'antibiotique (3 clones sur 200 traités). L'excision des séquences du vecteur qui se trouvaient flanquées des sites loxP restant après l'excision du marqueur a conduit à une réduction du nombre de copies du vecteur à une seule copie portant le dernier site loxP. Ce type d'événements a été attendu; Srivastava et collaborateurs ont déjà observé une situation similaire chez *T. aestivalis* (Srivastava *et al.*, 1999). Ni le marqueur ni les séquences du vecteur excisés ne se sont intégrés ailleurs dans le génome et ont probablement été éliminés par une activité nucléasique.

L'indice GFP (Tableau 10, page 92) nous permet d'estimer l'efficacité de l'excision comme nous la définissons à la page 85. Nous calculons d'abord le nombre des protoplastes dans lesquels l'enzyme a été produite en multipliant le nombre des protoplastes traités par l'indice GFP (Tableau 10, page 92, A\*B). L'efficacité du système de recombinase Cre (en %) est le rapport entre le nombre des colonies ayant perdu la résistance (C, Tableau 10, page 92, tenant compte de l'analyse moléculaire) au nombre A\*B. L'efficacité est comprise alors entre 14 et 65%. Un potentiel d'améliorer l'efficacité du traitement probablement existe et l'augmentation du taux de transcription du gène *cre* devrait y contribuer. Le taux d'expression du gène *cre* a été probablement un des facteurs limitants de la réaction d'excision du marqueur dans la souche alox-3RN. Nous supposons que la raison essentielle de l'échec de l'excision du marqueur de cette souche n'était pas tant le nombre de copies du vecteur que le phénotype de la souche. Ce phénotype de type APRT<sup>-</sup> est très prononcé; la difficulté d'isoler des protoplastes de cette souche diminue certainement sa compétence de transformation.

L'efficacité d'excision de marqueurs de sélection à l'aide du système de recombinase spécifique de site permet son utilisation courante chez la souris ou la levure. Chez les végétaux, l'efficacité de l'excision de marqueurs intégrés dans le génome sont rapportée dans le Tableau 11.

Tableau 11. Efficacité de l'excision de gènes marqueurs par le système Cre/lox chez les plantes.

réf.	plante	gène marqueur *	promoteur du gène cre	mode de transfert du gène cre	nombre de clones testés	nombre de cas d'excision attendu §	nombre de cas d'excision obtenu	efficacité
(Dale and Ow, 1991)	<i>N. tabacum</i>	<i>hpt</i>	35S	retransf	11	11	10	<b>91%</b>
(Dale and Ow, 1991)	<i>N. tabacum</i>	<i>hpt</i>	35S	X	78 (sur 316)	78	42	<b>54%</b>
(Bayley <i>et al.</i> , 1992)	<i>N. tabacum</i>	<i>luc</i>	35S	X	125	31	8	<b>26%</b>
(Bayley <i>et al.</i> , 1992)	<i>N. tabacum</i>	<i>luc</i>	35S	X	80	20	9	<b>45%</b>
(Russell <i>et al.</i> , 1992)	<i>N. tabacum</i>	<i>ALS</i>	35S	retransf	38	38	36	<b>95%</b>
(Russell <i>et al.</i> , 1992)	<i>N. tabacum</i>	<i>ALS</i>	35S	X	19-44 par X		chimères	qualitatif, ≠ en fonction du X
(Russell <i>et al.</i> , 1992)	<i>A. thaliana</i>	<i>ALS</i>	35S	X	env. 100 / X (F2) 400- 1000 /X (F3)		chimères	> <b>99.5%</b> dans F3
(Srivastava <i>et al.</i> , 1999)	<i>T. aestivalis</i>	<i>bar</i>	pUbi	X	env. 70	tous	tous	≅ <b>100%</b> ( <i>loxP-bar excisé détecté</i> )
(Gleave <i>et al.</i> , 1999)	<i>N. tabacum</i>	<i>nptII codA</i>	35S	retransf	7	7	7	<b>100%</b>
(Gleave <i>et al.</i> , 1999)	<i>N. tabacum</i>	<i>nptII codA</i>	35S	transit	a) 360 b) 413	?	a) 3 # b) 3 #	< <b>1% (?)</b>
ce travail	<i>P. patens</i>	<i>nptII</i>	35S	transit	a) 500 b) 300	6-30% donc a) 30-150 b) 18-90	a) 10 b) 9	<b>14-65%</b>

\* - Le nombre de copies du marqueur à exciser se limite à une copie intégrée à l'aide de *A. tumefaciens*, mais certaines études ne l'indiquent pas clairement (sauf (Srivastava *et al.*, 1999));

retransf -retransformation stable de la souche porteuse du marqueur par un vecteur d'expression du gène *cre*;

X - croisement de la souche porteuse du marqueur avec une souche exprimant le gène *cre*;

transit - expression transitoire du gène *cre*;

pUbi - le promoteur du gène d'ubiquitine de *Z. mays*;

§ Le nombre de cas d'excision attendu correspond à une efficacité de l'enzyme Cre de 100%. Cette efficacité devrait être atteinte dans toutes les cellules possédant les sites lox et l'enzyme Cre. Les testes faits sur F1 ont dû tenir compte de la ségrégation indépendante du locus du marqueur "floché" et du locus du gène *cre* lorsqu'aucun marqueur ne permet de sélectionner la descendance possédant les deux locus.

# - dont 2 ont intégré le gène *cre*.

L'efficacité de l'excision de marqueur à l'aide de la recombinaise Cre peut dépendre de plusieurs facteurs: l'organisme et le type de tissu utilisé, le taux d'activité de l'enzyme Cre, le mode de transfert et d'expression du gène *cre* ainsi que le nombre de copies du marqueur à exciser. Le mode de transfert et le taux d'activité de l'enzyme Cre jouent le rôle le plus important. La retransformation ou l'expression transitoire d'un vecteur extrachromosomique mais non le croisement sont les meilleures méthodes de transfert du gène *cre* dans les cellules cibles. Même si le taux d'activité de l'enzyme Cre peut influencer l'efficacité d'excision, le taux de transcription n'est pas forcément en proportion directe avec l'activité de l'enzyme (Russell *et al.*, 1992) et la transcription sous le contrôle du promoteur 35S a été satisfaisante (toutes les références). La question du rôle de la position génomique de la cible lox et de l'activité de l'enzyme Cre même reste ouverte car les résultats des études publiées sont contradictoires ((Russell *et al.*, 1992) contre (Odell *et al.*, 1990), (Odell and Russell, 1994)).

Notre étude démontre que cette efficacité est comparable pour les deux locus testés et pour 1 ou 10 copies du marqueur excisé. L'influence du mode de transfert ou du taux d'expression du gène *cre* n'a pas été comparée.

Lyznik et son groupe utilisent le gène *flp* (codant pour la recombinaise FLP) placé sous le contrôle du promoteur du gène *Gmhsp-17.5E* de soja, inductible par le choc thermique (Lyznik *et al.*, 1995). Le gène de la recombinaise et le gène *nptII* bordé des sites loxP sont introduits dans *Z. mais* par co-transformation et la réaction est induite dans les transformants par un choc thermique. Les chercheurs montrent que l'efficacité du système FLP/FRT est de 20-25%.

Luo et collaborateurs ont testé le système FLP/FRT chez *A. thaliana* pour exciser le gène *nptII*. Ils observent une excision dans toute la descendance du croisement entre les plantes qui portent le gène *nptII* bordé des sites loxP et celles qui expriment le gène *flp* sous le contrôle du promoteur 35S. L'efficacité est ici de 100%. Le groupe étudie, selon la même stratégie, l'excision du gène qui confère la stérilité mâle; ce gène code pour l'ARN anti-sens du gène *bcp1* spécifique du pollen et du tapétum. Dans ce dernier cas, ils recherchent la fertilité de la descendance. L'efficacité est tout aussi importante, de 98%; alors que 2% de la descendance est chimérique.

En somme, nous constatons que l'efficacité du système Cre/lox est comparable à celle du système FLP/FRT, comme l'affirme Odell dans sa revue (Odell and Russell, 1994).

Nous montrons ici que l'excision recherchée peut être réalisée chez *P. patens* dans 2-3% de clones traités, ce qui correspond, selon nos estimations à une efficacité du traitement de 14-65%. L'excision complète de plusieurs copies est accompagnée d'une réduction du nombre de copies du vecteur intégrées à une seule copie dans 100% de cas. Ceci est parfaitement dans la gamme des valeurs obtenues dans d'autres systèmes. Cette efficacité nécessite un criblage d'environ 200 clones traités. Du point de vue pratique, les clones désirés peuvent être sélectionnés en une période de quinze jours, le temps total de travail d'un laborant nécessaire ne dépassant pas 1 journée. La sélection des clones doit être suivie d'analyses moléculaires.

Notre étude ne met pas en évidence de problèmes liés à l'utilisation du système Cre/lox. Le système FLP/FRT pourrait également être testé mais nous n'attendons pas une amélioration significative de l'efficacité. En revanche, nous devons tester l'efficacité du traitement en fonction du taux d'expression de la recombinaise Cre. Nous prédisons une amélioration avec un promoteur plus fort que 35S CaMV, par exemple, pAct1 (McElroy *et al.*, 1991), que nous étudions dans le chapitre V.

Nous avons utilisé l'expression de la protéine GFP comme contrôle parallèle de la compétence des protoplastes. Néanmoins, un contrôle interne est nécessaire afin de mieux évaluer l'efficacité du traitement Cre. A cette fin, nous pourrions placer le gène codant pour la protéine GFP sur le même vecteur qui porte le gène *cre*. Nous estimerions alors l'efficacité de l'excision en pourcentage des protoplastes qui perdent le gène excisé parmi les protoplastes dans lesquels nous observerions la production de GFP.

Nous avons démontré dans nos essais d'expression transitoires que le vecteur ne s'intègre pas dans le génome et qu'il est perdu déjà après une brève période de culture. Nous avons également produit des souches intégratives portant le vecteur d'expression Cre. L'analyse de développement de la phase gamétophytique de ces transformants n'a pas révélé d'altérations morphologiques détectables due à la présence du gène *cre*. Afin de confirmer qu'une expression forte constitutive ne confère aucun phénotype particulier, cette analyse doit être complétée par une étude plus approfondie d'un transformant intégratif dans lequel l'expression du gène *cre* sera suivie à tous les stades de développement. L'étude de l'expression (Northern) doit être suivie de l'analyse Western et de l'activité de l'enzyme Cre.

Nous croyons que l'excision de marqueur à l'aide de l'expression transitoire du gène *cre* est préférable à l'expression constitutive. Même si nous ne mettons en évidence aucun effet mutagène apparent causé par l'expression constitutive, la probabilité d'un tel effet est nettement moindre lors de l'expression transitoire. Nous n'avons pas observé d'altérations du phénotype dans les clones dérivés des traitements transitoires.

L'insertion des sites loxP dans différents locus étudiés devrait également s'accompagner d'analyses phénotypiques afin de repérer d'éventuels événements d'instabilité qui pourraient être spécifiques au locus ou induite par l'insertion des sites loxP.

En somme, l'excision de marqueurs ainsi qu'une technique de réduction de nombre de copies du vecteur intégrées, basées sur le système de recombinase Cre peuvent facilement devenir une technique courante de l'arsenal de la mutagenèse ciblée. En effet, les recherches dans notre laboratoire utilisant la mutagenèse ciblée chez *P. patens* pourraient désormais se servir des marqueurs "floxés" pour construire les vecteurs de transformation. Les études qui cherchent à désactiver plusieurs gènes pourront nous aider à répondre à la question qui est de savoir si l'excision d'un deuxième marqueur bordé de sites loxP dans une souche portant déjà un site loxP pourrait être contrôlée afin d'éviter de dangereux remaniements de segments chromosomiques, qui ont une conséquence particulièrement grave dans un organisme haploïde.

# Chapitre V: Utilisation de promoteurs inductibles et constitutifs chez *P. patens*

## Résumé

La mutagenèse ciblée permet d'étudier la régulation des gènes en modifiant ou en remplaçant leur promoteur *in vivo*. La connaissance de l'activité de divers promoteurs hétérologues devient alors nécessaire. L'utilisation de promoteurs inductibles permet d'appliquer la mutagenèse ciblée à des gènes dont l'inactivation est létale. Nous examinons l'activité de différents promoteurs chimériques d'expression constitutive basés sur le promoteur 35S de CaMV et sur le promoteur du gène *actine1* du riz. Leur activité respective permet d'atteindre différents niveaux de l'expression constitutive (différence d'un facteur de 100). Nous montrons que le contexte de la traduction optimisé selon Kozak et la présence d'un intron en amont de la séquence codante augmentent l'expression transitoire du gène rapporteur GUS, comme on l'observe chez les angiospermes. Nous rapportons que le système chimérique inductible par les hormones glucocorticoïdes est fonctionnel chez *P. patens* et permet une induction de l'expression transitoire du gène rapporteur GUS atteignant un facteur de 50. Nous discutons des possibilités d'évaluer l'importance des différentes activités observées dans ces essais d'expression transitoire par rapport à d'autres gènes exprimés de façon stable dans le contexte cellulaire.

## 1. Introduction

L'analyse de processus biochimiques ou de développement à l'aide de plantes transgéniques nécessite des systèmes de régulation du taux de transcription des gènes étudiés. Différents niveaux d'expression et de régulation spatio-temporelle sont déterminés par le choix du promoteur et la structure locale de la chromatine. Les efforts visant à développer des systèmes de régulation portent sur les promoteurs (Potrykus and Spangenberg, 1995).

Dans ce chapitre, nous comparerons l'activité de différents promoteurs hétérologues constitutifs et conditionnels chez la mousse *P. patens*. Le potentiel de leur utilisation est d'autant plus important que les techniques de la mutagenèse ciblée chez *P. patens* permettent une manipulation de précision d'éléments régulateurs.

### 1.1. Expression constitutive

Les promoteurs les plus couramment utilisés dans la biologie végétale sont le promoteur du gène 35S du virus de la mosaïque du chou-fleur (CaMV), le promoteur du gène 19S de CaMV, le promoteur du gène *nos* de *A. tumefaciens* codant pour la nopaline synthase (pNOS), le promoteur du gène codant pour l'alcool déhydrogénase 1 de *Z. maïs* (pAdh1) ou le promoteur des gènes codant pour l'ubiquitine de *A. thaliana* ou de *Z. maïs*. Le plus utilisé est le 35S CaMV qui est actif dans pratiquement toutes les plantes dicotylédones et monocotylédones (Potrykus and Spangenberg, 1995).

McElroy et collaborateurs ont isolé le promoteur du gène *actine1* du riz (pAct1) pour développer une gamme de vecteurs d'expression pour les monocotylédones, dans lesquelles les promoteurs 35S CaMV ou pAdh1 ne sont pas suffisamment actifs (McElroy *et al.*, 1991).

Cette étude montre que l'expression transitoire du gène GUS<sup>§</sup> augmente (d'un facteur de 10) si l'intron 1 du gène Act1 est ajouté en amont du gène rapporteur. Un épissage efficace est nécessaire à la stimulation de l'expression observée. L'optimisation du contexte d'initiation de la traduction selon Kozak (Kozak, 1991) permet également une augmentation du taux d'expression (d'un facteur de 4). Ces deux effets sont additifs.

Plusieurs études démontrent que les promoteurs constitutifs pNOS, 35S et 19S de CaMV sont actifs chez *P. patens*. L'expression des gènes marqueurs de sélection *nptII* et *hpt* pilotés par ces promoteurs est suffisante pour conférer la résistance à l'antibiotique correspondant si *P. patens* est transformée par un vecteur portant ces gènes (Schaefer and Zryd, 1997; Schaefer *et al.*, 1991). Le promoteur 35S CaMV permet aussi une expression stable de gènes hétérologues provoquant une altération de développement de *P. patens* (Schaefer, 1994). Nous avons également montré dans le chapitre IV que le promoteur 35S confère une expression transitoire du gène *cre* suffisante.

Le promoteur pNOS a une activité très faible qui ne permet pas une transcription stable suffisante de gènes de résistance aux antibiotiques (Girod, P.-A., communication personnelle, ainsi que nos données non présentées).

Zeidler met en évidence que le promoteur pAct1 complet confère une bonne expression transitoire du gène GUS (Zeidler *et al.*, 1996). Dans des tests préliminaires, nous avons observé que le promoteur pAct1 permet une expression des gènes *hpt* et *nptII* nécessaire à une résistance stable ou transitoire aux antibiotiques (données non publiées, le vecteur pActHpt est présenté dans l'annexe B, page 142).

Dans ce chapitre, nous comparons l'expression transitoire du gène GUS placé sous le contrôle des promoteurs constitutifs 35S CaMV et pAct1. Nous testons l'influence du contexte d'initiation de la traduction défini par Kozak (Kozak, 1991) et de la présence de l'intron 1 du gène actine1 du riz (McElroy *et al.*, 1991) en amont du gène rapporteur.

## 1.2. Expression conditionnelle

Les promoteurs à expression conditionnelle permettent l'étude de l'expression de gènes létaux ou interférant avec le développement à un stade particulier, ainsi que la définition d'une corrélation exacte entre la régulation du transgène et le phénotype observé (discussion dans (Gatz and Lenk, 1998)).

Le système idéal d'expression conditionnelle doit posséder les propriétés suivantes:

- ✓ En absence d'induction, l'expression doit être la plus faible possible (activité basale);
- ✓ Le taux d'expression doit augmenter de façon significative sous l'effet de l'inducteur (activité induite). Nous évaluons cette augmentation par un facteur d'induction calculé au moyen du rapport de l'activité induite à l'activité basale;
- ✓ L'induction doit être spécifique au promoteur et ne doit avoir aucun effet secondaire;
- ✓ Il est souhaitable que le système inductible puisse être modulé par plusieurs inducteurs, les uns permettant de varier la cinétique de l'induction, les autres - la localisation de celle-ci.

---

<sup>§</sup> Nous désignerons ainsi le gène bactérien *uidA* codant pour l'activité de  $\beta$ -glucuronidase

En pratique, aucun système ne peut posséder toutes ces qualités. Par conséquent, si le produit du gène étudié n'agit qu'en fortes concentrations, une faible activité basale est acceptable. En revanche, si un niveau d'expression bas est recherché, aucune activité basale ne doit être présente mais un facteur d'induction élevé est moins important. De même, selon l'effet recherché, une activité résiduelle (restant après la levée de l'induction) peut être gênante.

Nous pouvons adopter deux stratégies dans les études sur les plantes: les promoteurs induits par un composé chimique spécifique endogène ou les promoteurs hétérologues induits par des substances qui ne se trouvent pas dans les cellules végétales. L'utilisation de promoteurs inducibles par des substances endogènes (des hormones ou des métabolites secondaires) est gênée par l'effet pléiotropique de celles-ci et par l'induction non spécifique ou croisée de ces promoteurs par d'autres facteurs (Gatz and Lenk, 1998; Moore *et al.*, 1998; Potrykus and Spangenberg, 1995).

Par exemple, les chercheurs du groupe de Cove (Knight *et al.*, 1995) démontrent une induction de l'expression du gène GUS piloté par le promoteur Em (du gène codant pour la protéine Em du blé) par l'acide abscissique et par le stress osmotique chez *P. patens*. Ce travail illustre le fait qu'un promoteur inducible par une substance régulatrice endogène peut être utilisé pour l'analyse de cette régulation; en revanche, il ne répond pas aux critères pratiques listés ci-dessus.

Nous nous tournons alors vers les systèmes de régulation hétérologue susceptibles d'être utilisés dans la biologie végétale.

#### Systemes de régulation hétérologues utilisés dans la biologie végétale

Les systèmes de régulation des organismes qui sont éloignés des plantes en termes d'évolution présentent peu de risque d'interférence avec les systèmes de régulation endogènes végétaux.

Puisque les systèmes de régulation des plantes ne doivent pas reconnaître les éléments régulateurs *cis* (promoteur), le gène codant pour le facteur d'activation en *trans* est introduit dans le génome de la plante étudiée. Les systèmes hétérologues nécessitent donc de construire une lignée produisant l'activateur *trans* de façon constitutive ou dans les tissus et à des stades de développements spécifiques. Pour que les résultats des études se servant de l'induction conditionnelle soient validés, la production de l'activateur *trans* ne doit pas altérer le phénotype de l'organisme étudié.

Les systèmes testés chez les plantes sont basés sur l'induction par la tétracycline (tc), par les hormones stéroïdes, par l'ion de cuivre ou par l'éthanol. Nous écartons de notre première étude l'expression inducible par l'ion de cuivre car l'interaction avec les mécanismes végétaux de détoxification et l'influence de l'ion de cuivre sur d'autres processus ne sont pas clairement définies.

#### *Expression régulée par la tétracycline*

Un groupe de chercheurs (Zeidler *et al.*, 1996) démontre l'efficacité du système de répression par tétracycline (tc), revu dans (Gatz and Lenk, 1998; Rossi and Blau, 1998), chez la mousse *P. patens*. L'expression stable ou transitoire du gène rapporteur GUS (estimée par la mesure de l'activité enzymatique) diminue d'un facteur supérieur à 100 en présence de tc dans le milieu en concentration de 1 mg/L. La coloration *in situ* par X-Gluc met en évidence une

répression efficace en présence de tc qui est levée au terme de 3 jours suite au transfert des plantes sur le milieu sans tc. La répression est plus lente après un transfert inverse (du milieu sans tc sur le milieu avec tc) mais les auteurs attribuent l'activité résiduelle à la dégradation lente des molécules de l'enzyme GUS. La tetracycline peut avoir un effet toxique sur les angiospermes (*A. thaliana*, *N. tabacum*) (Böhner *et al.*, 1999; Gatz and Lenk, 1998) et également sur *P. patens* (D. Schaefer, communication personnelle).

#### *Expression inductible par l'éthanol*

Le gène de l'alcool déhydrogénase I (*alcA*) du champignon *A. nidulans* est régulé par un facteur de transcription AlcR. Ce facteur est spécifique à la voie de la réponse cellulaire à l'éthanol et à des composés similaires et il se lie à la région du promoteur *alcA* (*palcA*) en leur présence. Le système fonctionne dans les cellules de tabac et de maïs (Caddick *et al.*, 1998). Caddick et collaborateurs mesurent l'expression du gène rapporteur *cat* placé sous le contrôle du promoteur *palcA* en présence de l'activateur AlcP sous le contrôle du promoteur 35S CaMV. L'activité basale du système est négligeable par rapport à l'activité induite par l'éthanol. Le facteur d'induction est d'environ 83. L'éthanol n'influence pas l'activité du promoteur 35S. Le promoteur *palcA*:35S est inactif dans les plantes sans l'activateur AlcR. Le système développé permet un bon facteur d'induction; l'inducteur est bon marché, facile à appliquer, biodégradable et peu toxique. Cette étude ne vérifie pas la spécificité de l'induction, le niveau de toxicité, l'importance de l'activité basale constatée, la cinétique à long terme et les concentrations endogènes de l'inducteur en différentes conditions. L'anoxie peut induire le système (Caddick *et al.*, 1998).

Les chercheurs de notre laboratoire (Schaefer, D. G., Wyss, J.-C. et Bessire, M., communication personnelle) ont effectué les essais de co-expression transitoire des vecteurs susmentionnés dans les protoplastes de *P. patens*. Dans ces essais, le gène GUS remplace le gène rapporteur *cat* et l'activité enzymatique est mesurée dans l'extrait brut des protoplastes transformés 24-72 heures après l'ajout de l'éthanol (Schaefer, 1994). Or, l'induction a lieu même sans ajout d'éthanol. Deux explications sont les plus vraisemblables: l'induction par l'éthanol endogène ou l'induction par l'éthanol utilisé dans les manipulations des cultures axéniques. Notre laboratoire n'a pas établi la cause exacte mais il est évident que le développement du système dans ces conditions pour une recherche fine chez *P. patens* est compromis.

#### *Expression régulée par les hormones glucocorticoïdes*

Une publication récente (Gatz and Lenk, 1998) révisé les systèmes de régulation conditionnelle développés pour les plantes à partir des promoteurs contenant des éléments de réponse aux hormones glucocorticoïdes et des domaines de liaison des récepteurs d'hormone (GR "glucocorticoid receptor"). Aucun système homologue n'étant connu chez les plantes, il ne devrait pas y avoir d'interaction avec des systèmes de régulation endogènes. De plus, les hormones glucocorticoïdes ne produisent pas d'effets pléiotropiques chez les plantes. Les chercheurs du groupe de Chua (Aoyama and Chua, 1997) ont développé un système permettant une expression de n'importe quel gène étudié inductible par l'hormone glucocorticoïde dexaméthasone (dx). Le système comprend un gène codant pour un facteur d'activation chimérique (GVG) et un promoteur (pUASGAL4, ou pUAS) qui est spécifiquement reconnu par celui-ci (Figure 44).

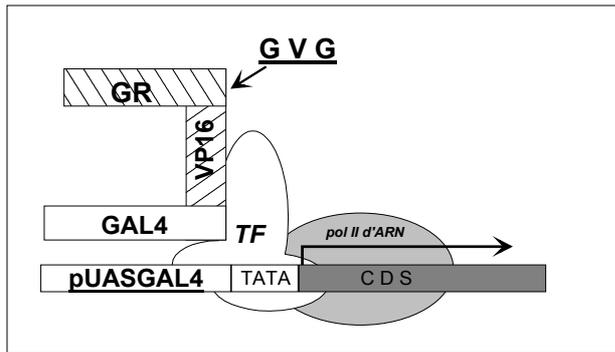


Figure 44. Système d'induction GVG construit par le groupe de Chua (Aoyama and Chua, 1997).

Le facteur d'activation chimérique GVG est composé d'un domaine de liaison de GR du rat, d'un domaine de transactivation de la protéine VP16 du virus H.simplex et d'un domaine de liaison à l'ADN du facteur de transcription GAL4 de *S.cerevisiae*. Le domaine d'activation VP16 interagit directement avec les facteurs généraux de transcription (TF) généralement conservés dans tous les eucaryotes. Le promoteur spécifiquement reconnu par le facteur GVG porte 6 copies de séquences d'activation des gènes *GAL* de la levure (*GAL UAS*) fusionnées au promoteur minimal, la région -46 à +9 du promoteur 35S qui porte la boîte TATA (*pUAS*). Le domaine de liaison à l'ADN de GAL4 du facteur GVG se lie aux séquences *GAL UAS* et active la transcription (CDS - séquence codante d'un gène). Cette liaison peut se produire en présence de glucocorticoïdes qui empêchera la liaison de la protéine GVG avec le complexe de protéines cellulaires comprenant la HSP90 ("heat shock protein").

L'étude de ce système (appelé GVG) chez *N. tabacum* et *A. thaliana* montre les avantages suivants:

- ✓ Le facteur d'induction est de plus de 100 ;
- ✓ Le facteur d'induction est proportionnel à la concentration de l'hormone permettant d'établir une relation dose - réponse;
- ✓ Différents dérivés de glucocorticoïdes permettent d'obtenir différents facteurs d'induction et différente cinétique d'induction;
- ✓ Il n'y pas d'effets secondaires dus à l'hormone ou à l'expression de l'activateur GVG.

Les points suivants n'ont pas été abordés de manière explicite dans cette étude:

- ✓ L'activité est indiquée en unité arbitraire sans référence;
- ✓ L'activité résiduelle rémanente persiste après la levée de l'induction.

La cinétique d'induction des systèmes conditionnels dépend de la durée du traitement inducteur, de l'export et de l'import de l'inducteur et de sa demi-vie. Afin de permettre un contrôle plus fiable de la répression, les auteurs du système GVG (Böhner *et al.*, 1999) ont développé un nouveau système (appelé TGV) fondé sur l'activation par l'hormone dx et sur la répression par tc. Les avantages du système TGV sont les suivants:

- ✓ L'hormone dx et le facteur d'activation portant le domaine de liaison aux glucocorticoïdes n'affectent pas le phénotype des plantes (observations, mesures de la fluorescence de la chlorophylle, taux de transcription du gène codant pour la glutathion S-transférase). Ceci complète l'étude précédente (Aoyama and Chua, 1997);
- ✓ L'induction par l'hormone dx est très efficace, d'un facteur atteignant 500 et l'activité basale dans le système TGV reste négligeable ( $\approx 0.1\%$  de 35S);
- ✓ La répression par tc est également efficace.

Les points faibles du système demeurent:

- ✓ La répression après la levée de l'induction par dx est lente voire absente en absence de tc;
- ✓ La répression par tc peut produire des effets sur la croissance des plantes.

Nous avons décidé de tester le système d'expression conditionnelle GVG construit par le Aoyama et ses collaborateurs (Aoyama and Chua, 1997) sur la mousse *P. patens*.

## 2. Résultats et discussion

Nous avons utilisé l'essai d'expression transitoire du gène GUS (Schaefer, 1994) afin d'évaluer le taux d'expression de différents promoteurs. Nous avons mesuré le produit (méthylumbellifère, MU) de l'action de l'enzyme GUS dans l'extrait des protoplastes transformés sur le substrat méthylumbelliféryl glucuronide (MUG).

Nous avons testé les promoteurs 35S (Jefferson *et al.*, 1987) et pAct1 (McElroy *et al.*, 1991; McElroy *et al.*, 1990) (expression constitutive), ainsi que le promoteur Em (expression induite par le stress osmotique et l'acide abscissique) (Knight *et al.*, 1995; Marcotte *et al.*, 1988) et pUAS (expression induite par les hormones glucocorticoïdes) (Aoyama and Chua, 1997). La construction des vecteurs est détaillée dans les références citées.

### Vecteurs d'expression constitutive et d'expression induite par le stress osmotique ou par l'acide abscissique

La figure ci-dessous explique la construction des vecteurs utilisés.

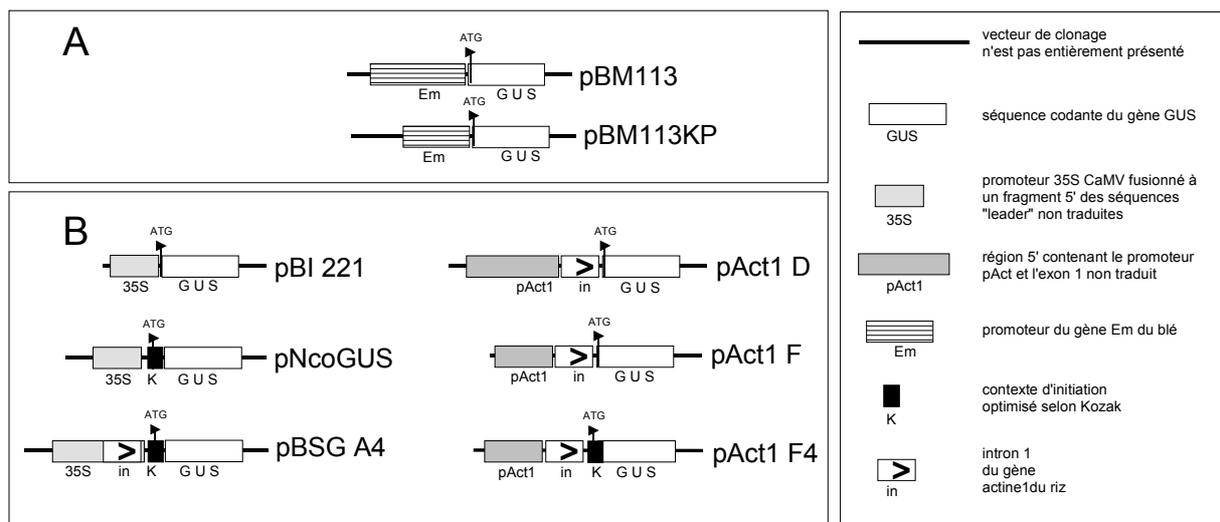


Figure 45. Les vecteurs d'expression constitutive du gène GUS.

A - Promoteurs inducibles par le stress osmotique (0.48 M et 0.33 M mannitol dans la solution de lavage et dans le milieu de régénération, respectivement) ou par l'acide abscissique (non testé) :

Le vecteur pBM113 porte le promoteur du gène Em du blé (pEm) complet;

Le vecteur pBM113kp porte le promoteur pEm dont la région 5' de 0.65 kb est supprimée.

B - Vecteurs qui utilisent les promoteurs constitutifs :

Les vecteurs pBI 221, pNcoGUS et pBCGA4 emploient le promoteur 35S ;

Les vecteurs pAct1-D, F et F4 emploient le promoteur pAct1.

Le contexte d'initiation optimisé selon Kozak est ACCATGG, le site NcoI contenant le codon d'initiation ATG est souligné.

Nous présentons la valeur moyenne  $\pm$  l'écart-type de plusieurs expériences. L'activité GUS mesurée dans les protoplastes transformés est variable selon l'expérience (l'écart-type est important), néanmoins, les résultats des expériences répétées dans notre laboratoire indiquent que la valeur moyenne est reproductible.

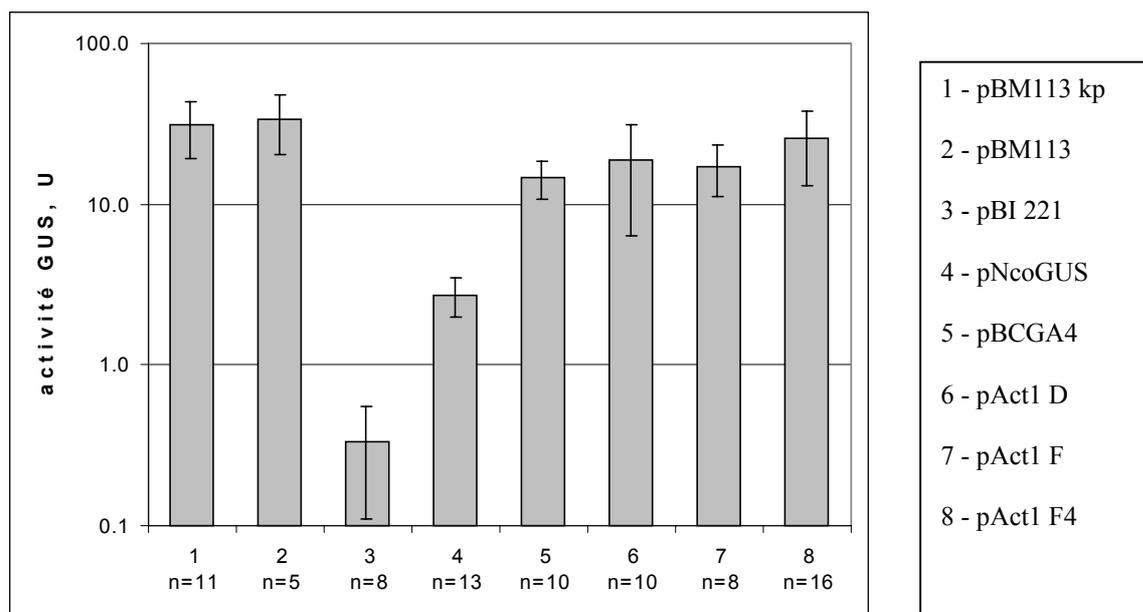


Figure 46. Mesures de l'activité GUS spécifique dans l'extrait brut des protoplastes transformés.

La barre d'erreur représente l'écart-type, n est le nombre d'expériences indépendantes. L'unité de mesure de l'activité GUS spécifique (U) est exprimée en nmole de métylumbelliféron MU produit par minute et par mg de protéines solubles (équivalent de BSA).

Nous ne détectons aucune activité GUS dans l'extrait des protoplastes transformés sans ADN (valeur zéro).

L'activité GUS spécifique mesurée dans notre essai d'expression des vecteurs pBM113 et pBM113kp est du même ordre de grandeur que celle induite par l'acide abscissique à  $10^{-4}$  M ( $31.4 \pm 27.3$  U) ou par le choc osmotique (mannitol à 0.44M,  $7.95 \pm 6$  U) rapportée par (Knight *et al.*, 1995). Dans notre essai, l'induction est provoquée par le choc osmotique dans les solutions de lavage et le milieu standard de régénération de protoplastes, contenant du mannitol à 0.48 et 0.33 M, respectivement (cf. Matériel et méthodes, page 134). Dans les protoplastes du riz, l'activité GUS spécifique est de 16.6 U après l'expression transitoire induite par l'acide abscissique à  $10^{-4}$  M (Marcotte *et al.*, 1988).

L'optimisation du contexte d'initiation de la traduction du gène GUS piloté par le promoteur 35S permet une augmentation de l'activité GUS d'un facteur de 9 (pNcoGUS comparé à pBI221). Le contexte optimisé provoque une augmentation d'un facteur de 1.4 pour le gène GUS piloté par le promoteur pAct1 (pAct1-F4 comparé à pAct1-F). Cette augmentation est du même ordre de grandeur (d'un facteur de 4) chez le riz et le maïs (McElroy *et al.*, 1991)).

L'introduction de l'intron 1 dans le gène GUS stimule la transcription. Nous observons une augmentation de l'activité GUS d'un facteur de 5.4 entre le vecteur pNcoGUS et pBCGA4. Cette augmentation est comparable à celle rapporté pour le riz et le maïs, qui est d'un facteur supérieur à 10 (McElroy *et al.*, 1991).

La délétion du fragment 5' de 0.8 kb en amont du promoteur pAct1 n'influence pas l'activité du promoteur de façon significative ni chez *P. patens* (pAct1-F comparé à pAct1-D), ni chez le riz (McElroy *et al.*, 1990).

Les vecteurs pBCGA4 et pAct-F4 nous permettent de comparer l'expression contrôlée par le promoteur 35S au promoteur pAct1. Les deux vecteurs portent l'intron 1 et pilotent le gène

GUS avec un contexte d'initiation de la traduction optimisé. Le promoteur pAct1 permet une activité GUS plus élevée, d'un facteur de 1.7. Le rapport est inversé chez le riz et le maïs (McElroy *et al.*, 1991), mais ces différences sont peu significatives.

Les auteurs de (McElroy *et al.*, 1991) soulignent deux informations intéressantes: ils observent une activité du promoteur pAct1 importante et l'effet stimulateur d'un intron de type monocotylédone sur la transcription chez les monocotylédones. Nous pouvons interpréter nos observations chez *P. patens* par l'hypothèse que la régulation chez la mousse *P. patens* est proche de celle des monocotylédones. Cette conclusion s'accorde avec celle de l'étude de Cove et collaborateurs (Knight *et al.*, 1995). Ces chercheurs, à l'aide de l'analyse de différentes mutations du promoteur Em du blé, de l'analyse des empreintes d'ADNase I et de l'analyse de la retardation sur gel, démontrent que les éléments fonctionnels du promoteur Em peuvent être utilisés dans la régulation de *P. patens*. Ils concluent que les éléments de régulation chez *P. patens* et chez les monocotylédones sont bien conservés.

La comparaison des promoteurs en utilisant les valeurs absolues de l'activité GUS spécifique observée dans l'extrait de différentes plantes est quelque peu difficile. Une des raisons principales est l'utilisation de techniques différentes d'expression et le choix du matériel végétal. Nous remarquons que la transformation par biolistique est la moins efficace. De plus, le Tableau 12 met en évidence une incohérence dans l'utilisation des unités de mesure dans certaines études qui ne permet pas de comparer ces activités.

L'activité GUS spécifique mesurée par le groupe de McElroy (McElroy *et al.*, 1991; McElroy *et al.*, 1990) est différente selon le type d'expérience. Nous pouvons comparer nos mesures de l'activité GUS spécifique induite par le promoteur pAct1 avec les valeurs observées dans l'expression piloté par le même promoteur dans les protoplastes du riz transformés avec PEG. Les valeurs que ce groupe de chercheurs observe dans d'autres expériences sont inférieures aux nôtres d'un facteur de 10 à 1000.

En revanche, l'activité des promoteur 35S et pAct1 rapporté par Zeidler (Zeidler *et al.*, 1996) ne peut pas servir de base de comparaison: les unités de mesure qu'il utilise ne sont pas explicites (activité spécifique mesuré par mg de protéines ou de l'ADN transformant). La citation des valeurs du travail de Jefferson (Jefferson *et al.*, 1987) est erronée.

La gamme de l'activité de différents promoteurs est une information importante pour des études sur *P. patens*. Les essais que nous rapportons indiquent que nous pouvons modifier l'expression constitutive de transgènes d'un facteur de 100, en employant un des promoteurs testés (Figure 46). Nous démontrons également que chez *P. patens*, la séquence définie par Kozak pour les eukaryotes forme le contexte optimal d'initiation de la traduction. Le rôle de l'épissage du premier intron dans l'activation de la transcription chez *P. patens* est aussi important que chez les angiospermes (Callis *et al.*, 1987; McElroy *et al.*, 1991).

Nous pourrions également étudier la régulation des gènes endogènes en remplaçant leurs éléments régulateurs par un des promoteurs connus et testés, grâce à la mutagenèse ciblée. La corrélation entre l'activité des promoteurs mesurée dans des essais d'expression transitoire et l'activité des gènes contrôlés par ces promoteurs dans un cas d'intégration stable est toutefois complexe. Nous discuterons de ces études à la fin de ce chapitre.

Tableau 12. Comparaison de l'activité GUS spécifique mesurée dans différents systèmes d'expression contrôlée par quelques promoteurs constitutifs.

Promoteur	Réf.	Matériel végétal	Condition et technique d'expression	Activité GUS spécifique moyenne	Unité de mesure
<b>Em</b>	(Marcotte <i>et al.</i> , 1988)	<i>O. sativa</i>	PEG §	<b>16.7</b>	U
	(Knight <i>et al.</i> , 1995)	<i>P. patens</i>	PEG §	<b>31.4</b>	U
	(Knight <i>et al.</i> , 1995)	<i>P. patens</i>	PEG <sup>M1</sup>	<b>7.95</b>	U
	ce chapitre	<i>P. patens</i>	PEG <sup>M2</sup>	<b>31.2</b>	U
<b>pAct1<sup>a</sup></b>	(McElroy <i>et al.</i> , 1991)	<i>O. sativa</i>	B	<b>0.045</b>	U
	(McElroy <i>et al.</i> , 1990)	<i>O. sativa</i>	PEG	<b>3.3</b>	U
	(McElroy <i>et al.</i> , 1991)	<i>Z. maïs</i>	B	<b>0.027</b>	U
	(McElroy <i>et al.</i> , 1991)	<i>N. tabacum<sup>e</sup></i>	B	<b>nd</b>	U
	ce chapitre	<i>P. patens</i>	PEG	<b>18.7</b>	U
<b>pAct1<sup>b</sup></b>	(Zeidler <i>et al.</i> , 1996)	<i>P. patens</i>	PEG	≈ <b>1000</b> (cf. texte)	nmole MU min <sup>-1</sup> mg <sup>-1</sup>
<b>35S<sup>c</sup></b>	(McElroy <i>et al.</i> , 1991)	<i>O. sativa</i>	B	<b>0.02</b>	U
	(McElroy <i>et al.</i> , 1991)	<i>Z. maïs</i>	B	<b>0.013</b>	U
	(McElroy <i>et al.</i> , 1991)	<i>N. tabacum<sup>e</sup></i>	B	<b>0.007</b>	U
	(Jefferson <i>et al.</i> , 1987)	<i>N. tabacum<sup>f</sup></i>	S	<b>0.283-0.577</b>	U
	ce chapitre	<i>P. patens</i>	PEG	<b>0.3</b>	U
<b>35S NcoI<sup>d</sup></b>	(Zeidler <i>et al.</i> , 1996)	<i>P. patens</i>	PEG	≈ <b>800</b> (cf. texte)	nmole MU min <sup>-1</sup> mg <sup>-1</sup>
	ce chapitre	<i>P. patens</i>	PEG	<b>2.7</b>	U

PEG - transformation des protoplastes par polyéthylène glycol; B - transformation des protoplastes par biolistique; S - mesures dans les transformants intégratifs (expression stable);

nd non détectable;

U - nmole MU min<sup>-1</sup> mg<sup>-1</sup> protéine si cette conversion est possible;

a - vecteur pAct1-F ou similaire (pAct1-R);

b - promoteur pAct1 complet;

c - promoteur 35S sans ajout d'intron ni optimisation du contexte Kozak;

d - promoteur 35S, le contexte d'initiation Kozak optimisé ;

e - culture de cellule en suspension;

f - mesuré dans différents organes des transformants intégratives;

§ - induction par acide abscissique 10<sup>-4</sup> M;

M1 et M2 - induction par choc osmotique (mannitol à 0.44, et à 0.33 M ou 0.48 M dans les solutions de transformation, respectivement, cf. pages précédentes).

Les vecteurs d'expression conditionnelle, inducible par les hormones glucocorticoïdes

Nous avons utilisé l'hormone synthétique dexaméthasone (dx) pour induire l'expression du promoteur pUAS en présence du facteur d'activation GVG. L'hormone n'a pas d'effets secondaires sur l'essai d'expression transitoire car elle n'affecte pas l'expression du promoteur pAct F4 (Figure 48, condition 1,  $\pm 30$  à  $150 \mu\text{M}$ ). Nous n'observons pas d'altération de croissance ou de développement de *P. patens* de souche sauvage sur le milieu PPNH4c ou PPNO3 additionné de  $100\mu\text{M}$  après 5 à 10 jours de culture (Vlach, S. J. communication personnelle).

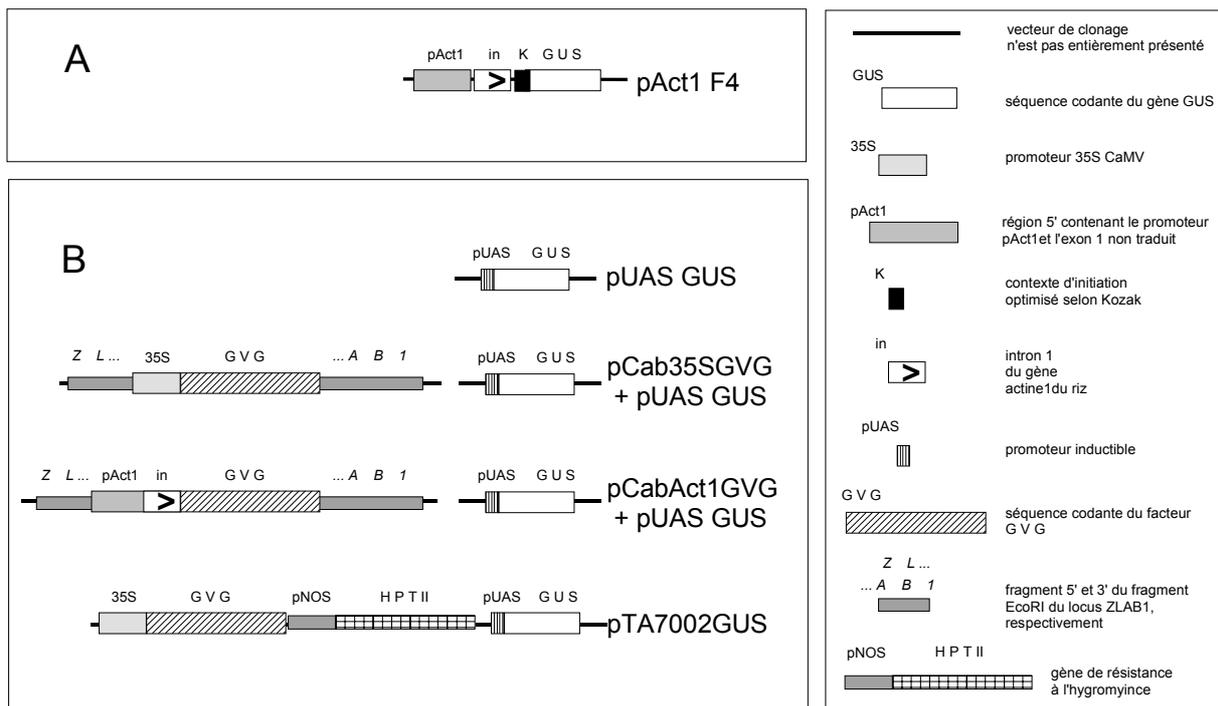


Figure 47. Les vecteurs permettant une expression conditionnelle du gène GUS.

A - Le promoteur constitutif utilisé comme contrôle positif est pAct1-F4.

B - Les vecteurs utilisés pour tester l'expression conditionnelle sont les suivants:

pUASGUS - porte le gène GUS sous le contrôle du promoteur inducible pUAS.

pCab35SGVG et pCabAct1GVG - portent le gène GVG sous contrôle du promoteur 35S ou pAct1, respectivement. Les séquences du fragment EcoRI du gène *ZLAB1* (chapitre I) permettent une intégration ciblée dans le locus *ZLAB1* par remplacement mais ces événements ne sont pas sélectionnés dans nos essais d'expression transitoire;

pTA7002GUS - porte le gène pUASGUS et le gène 35S GVG, ils sont séparés par le gène de résistance à l'hygromycine que nous n'utilisons pas dans les essais.

Afin de déterminer la teneur cellulaire en GVG nécessaire et suffisante à une expression conditionnelle du promoteur pUAS, nous avons testé l'expression du gène GVG<sup>§</sup> sous le contrôle des promoteurs 35S ou pAct1(+intron 1).

<sup>§</sup> Nous désignerons ainsi le gène chimérique codant pour le facteur GVG

Nous avons effectué les essais par co-expression transitoire des deux gènes, GVG (facteur d'expression piloté par 35S ou pAct1) et GUS (rapporteur piloté par pUAS) dans la souche sauvage.

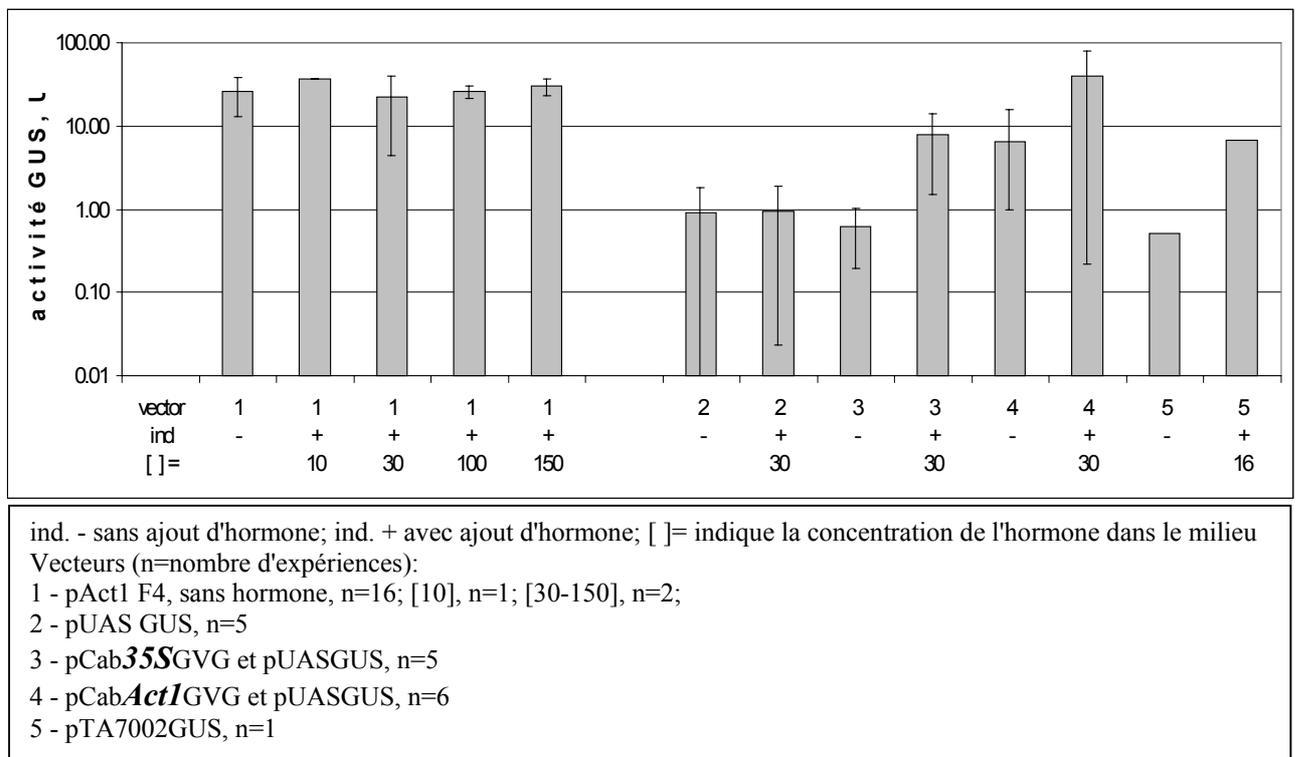


Figure 48. Mesures de l'activité GUS spécifique dans l'extrait brut des protoplastes transformés.

Les barres d'erreur représentent l'écart-type, si le nombre d'expériences permet de le calculer. L'unité de mesure de l'activité GUS spécifique (U) est exprimée en nmole de métylumbelliférol MU produit par unité de temps et par mg de protéines solubles (équivalent de BSA).

#### Activité basale du système d'expression conditionnelle

Nous pouvons comparer l'activité du promoteur pUAS (désigné pUAS) en absence du facteur GVG (Figure 48, conditions 2 ± hormone) à celle du promoteur constitutif le plus faible, 35S du vecteur pBI221 (Figure 45, condition 3, page 112). Il est tout à fait possible que des facteurs de transcription endogènes reconnaissent le domaine GAL4UAS; cette activité basale serait donc inévitable. Nous n'avons pas comparé cette activité à l'activité du promoteur minimal seul (boîte TATA sans le domaine GAL4UAS).

L'activité basale du promoteur pUAS est maximale en présence du facteur GVG piloté par le promoteur pAct1 (+ intron1) (Figure 47, condition 4-). Nous savons que l'activité du promoteur pAct1 mesurée dans les essais transitoires est plus forte que celle du promoteur 35S (Figure 45 page 112 et la discussion qui la suit). L'activité basale accrue pourrait être due à une production trop forte de l'activateur trans, GVG. Elle est moins importante si le facteur GVG est produit sous le contrôle du promoteur 35S, plus faible.

L'importance effective de cette activité basale ne peut être estimée qu'en comparaison avec le taux d'expression des gènes endogènes. Nous en discuterons plus loin.

### Facteur d'induction du système d'expression conditionnelle

Nous calculons le facteur d'induction par division de l'activité induite à l'activité basale atteinte *dans la même expérience* sans ajout d'hormone. Nous considérons comme non valable le calcul par une simple division de l'activité basale sur l'activité induite moyennes.

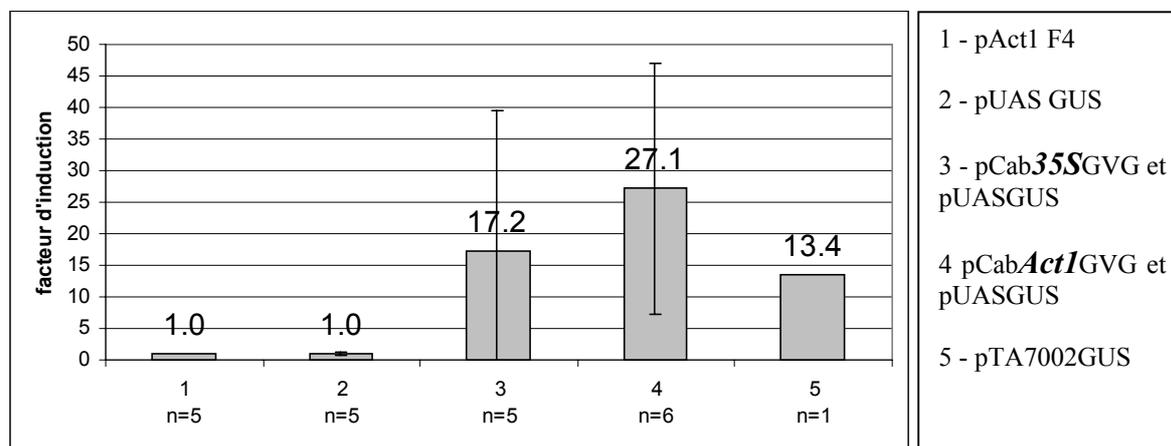


Figure 49. Facteur d'induction moyen de l'activité GUS spécifique mesurée dans l'extrait brut des protoplastes après l'expression transitoire.

La barre d'erreur indique l'écart-type; n= le nombre d'expériences indépendantes.

La co-expression du gène pUASGUS avec le gène GVG piloté par le promoteur pAct1 permet d'atteindre le facteur d'induction de  $27.1 \pm 19.8$ . Ce facteur est moins élevé si le gène GVG est piloté par le promoteur 35S. Nous pouvons donc moduler la régulation de l'expression pilotée par le promoteur pUAS en modifiant le taux d'expression de l'activateur GVG. Une expression plus forte de l'activateur nous permet une meilleure induction mais provoque également une activité basale plus élevée (cf. plus haut).

Plus généralement, nous pouvons modifier non seulement le taux d'expression mais aussi la régulation spatio-temporelle de l'expression de l'activateur GVG en le plaçant sous le contrôle d'un promoteur spécifique à un tissu ou à un stade de développement. La nécessité de la caractérisation d'autres promoteurs endogènes ou hétérologues devient évidente.

Plus le facteur d'induction est important, plus grande est la flexibilité de régulation que le système d'induction peut permettre. Nous pouvons comparer nos résultats aux essais d'expression conditionnelle chez *N. tabacum* ou *P. patens* résumées dans Tableau 13 (nous avons sélectionné les essais quantifiables). Les essais sur *N. tabacum* se fondent sur l'expression d'un gène rapporteur intégré dans un locus génomique inconnu.

Dans tous les cas, l'activité GUS spécifique basale est comparable (voire supérieure) à celle que nous mesurons dans nos essais. L'induction peut atteindre un facteur de 50 à 800, selon les études. Ces valeurs sont supérieures à celle que nous observons ( $27.1 \pm 19.8$ ). En somme, le système GVG inductible par les hormones glucocorticoïdes est efficace chez la mousse *P. patens* mais les éléments de régulation, tel que le gène exprimant le facteur GVG et peut-être le promoteur cible, peuvent être optimisés. Les efforts d'optimisation doivent principalement se concentrer sur la caractérisation d'un niveau optimal d'expression du facteur GVG.

Nous avons déjà mentionné que l'utilité d'un système de régulation dépend des gènes étudiés. Par conséquent, le facteur d'induction de  $26.2 \pm 18.2$  et l'activité basale faible mais présente, peuvent être acceptables ou non pour étudier tel ou tel gène. De surcroît, ces paramètres seront probablement révisés s'ils sont évalués dans le contexte génomique.

Tableau 13. Comparaison des systèmes de régulation conditionnelle développés pour *N. tabacum* et *P. patens*.

Réf.	Plante	Système	Inducteur, [ ]	Activité basale	Activité maximale induite	Facteur d'induction
(Caddick <i>et al.</i> , 1998)	<i>N. tabacum</i> , <i>S</i>	p35S:alcR + palcA:CAT	EtOH	0.38+/- 0.18 ng CAT	31.73 ± 1.24	<b>83.63</b>
(Aoyama and Chua, 1997)	<i>N. tabacum</i> , <i>S</i>	p35S:GVG + pUAS:luc	dx	1 u.a.	≥100 u.a.	≥ <b>100</b>
<i>l'unité de mesure de l'activité GUS spécifique est nmole min<sup>-1</sup> mg<sup>-1</sup> protéines (U)</i>						
(Böhner <i>et al.</i> , 1999)	<i>N. tabacum</i> , <i>S</i>	p35S: TGV + pUAS:GUS	dx (répresseur tc), 14 jours à 30 µM	0.004	2.079 ± 0.500	<b>520</b>
				0.005	0.767 ± 0.107	<b>153</b>
				0.004	1.842 ± 0.715	<b>461</b>
				0.005	2.151 ± 0.561	<b>430</b>
(Böhner <i>et al.</i> , 1999)	<i>N. tabacum</i> , <i>S</i>	±tTA + Ptax GUS	induction constitutive en présence de tTA,	0.012	10.26	<b>855</b>
(Böhner <i>et al.</i> , 1999)	<i>N. tabacum</i> , <i>S</i>	±tTA + pTop10 GUS	activateur	0.007	1.34	<b>191</b>
(Zeidler <i>et al.</i> , 1996)	<i>P. patens</i> , <i>S et T</i>	tTA + pTop10 GUS	tc, 1 mg/L	2-3	900 - 1100 *	≈ <b>500</b>
(Mett <i>et al.</i> , 1993)	<i>N. tabacum</i> , <i>S</i>	ACE1 + Cu- GUS	Cu <sup>++</sup> , 50 µM	0.020-0.1	1.2 - 1.6	≤ <b>50</b>
ce travail	<i>P. patens</i> , <i>T</i>	GVG + pUasGUS	dx, 16-150 µM	0.5 - 6.5	≤ 40	<b>26.2</b>

[ ] - concentration; S - transformants stables, T- expression transitoire, U - unités d'activité GUS ; \* valeurs absolues contestées, (cf. page 115), u.a. - unité arbitraire, tc - tetracycline, dx - dexaméthasone. Nous ne comparons pas les valeurs basale et induite absolues mesurées avec les différentes approches mais nous soulignons l'importance du facteur d'induction.

### 3. Perspectives

Pour continuer le développement du système de promoteurs inductibles chez *P. patens*, nous suggérons de:

- ✓ Produire une gamme de souches transgéniques stables produisant le facteur GVG sous contrôle de différents promoteurs. L'analyse se poursuivra sur les souches ayant intégré une copie du vecteur par remplacement dans différents locus.
- ✓ Analyser la relation entre le taux d'expression du facteur GVG et l'activité induite à l'aide d'essais d'expression transitoire GUS afin d'isoler des souches permettant un meilleur contrôle de l'expression transitoire;
- ✓ Analyser le phénotype des souches sélectionnées qui doit être normal.

L'analyse des essais transitoires ne permet pas une caractérisation complète d'un système de régulation pour des raisons suivantes:

- ✓ L'activité mesurée est le résultat d'expression de différents nombres de molécules du vecteur par cellule. La quantité d'ADN transformant dans la cellule est trop élevée et ne correspond pas à la quantité d'un gène étudié dans des conditions physiologiques;
- ✓ L'expression à partir des éléments extrachromosomiques ne caractérise que l'activité du promoteur même en dehors de tout contexte naturel chromosomique.
- ✓ L'induction de l'expression n'est mesurée que dans les protoplastes et dans des conditions non physiologiques (stress osmotique). Cette information donne peu d'indication sur l'expression dans les tissus formés et à différents stades de développement.

Il est donc impératif de tester le système de régulation avec un gène rapporteur ou endogène intégré en une copie dans le chromosome. L'utilisation du gène GUS peut être une première tentative qui permettrait une caractérisation quantitative.

Une bonne caractérisation se base sur un gène dont l'expression conditionnelle peut être évaluée par l'observation de phénotype. Dans cette optique, l'efficacité des systèmes de régulation conditionnelle est mise à l'épreuve à l'aide du gène isopentenyl-transférase (*ipt*) de la bactérie *A. tumefaciens* (revu dans (Gatz and Lenk, 1998), (McKenzie *et al.*, 1998)). L'enzyme *ipt* forme l'isopentenyl-AMP à partir de  $\Delta^2$ -isopentyl PPi et d'AMP. L'introduction de ce gène dans les plantes produit un niveau élevé de cytokinine dans les tissus transformés et mène à des altérations importantes de morphologie (par exemple, les parties aériennes ne s'allongent pas, la dominance apicale est abolie, les feuilles sont arrondies et petites, les racines ne se forment pas (Binns, 1994)). Même une faible expression de ce gène suffit à déclencher ces altérations. Ainsi, les auteurs de plusieurs études détectent une activité basale de différents promoteurs inductibles par le choc thermique lorsque ceux-ci contrôlent l'expression du gène *ipt* dans les plantes transgéniques (cité et traité dans la publication de McKenzie (McKenzie *et al.*, 1998)).

Un autre exemple est l'utilisation du promoteur inductible par l'ion de cuivre de *S. cerevisiae* (Mett *et al.*, 1993) pour contrôler l'expression du gène *ipt* dans le tabac (*Cu-ipt*). Les chercheurs (McKenzie *et al.*, 1998) n'ont observé aucune différence entre le type sauvage et le tabac portant le gène *Cu-ipt* sans induction. Par contre, ils observent une augmentation de la concentration en cytokinines, un retard de la sénescence des feuilles et un relâchement de la dominance apicale en différentes conditions d'induction. Les auteurs de cette étude associent

les modifications phénotypiques directement au fonctionnement du système de régulation conditionnelle.

Schaefer a observé un phénotype *ove* (sur-production de bourgeons et de gamétophores stériles (Ashton *et al.*, 1979a)), provoqué par l'expression du gène *ipt* dans *P. patens* (Schaefer, 1994). Ce phénotype est également provoqué par des concentrations de cytokinines élevées. De plus, une étude (Reutter *et al.*, 1998) montre que l'introduction du gène *ipt* complémente génétiquement la mutation "*budding*" et "*plastid division*" de deux mutants (P24, "*budding*" et PC22, "*plastid division*" et "*gametophore development*"). Ces chercheurs ont mesuré une concentration en cytokinines et en auxines élevée dans les transformants exprimant le gène *ipt*. La sévérité des altérations, par exemple le nombre de bourgeons induits, se trouve en corrélation avec la surproduction des cytokinines et donc avec l'expression du gène *ipt*. L'expression conditionnelle du gène *ipt* dans des transformants stables à l'aide du système GVG chez *P. patens* permettra de mieux le caractériser dans le contexte cellulaire et génomique naturel.

La mutagenèse de précision chez *P. patens* permet un remplacement des éléments de régulation de gènes endogènes dans leur contexte génomique. Nous proposons une première étude test sur le gène *PpAPT*. Ce gène présente les avantages suivants:

- ✓ Le gène *PpAPT* est probablement le mieux caractérisé de *P. patens*;
- ✓ Nous observons des phénotypes distincts associés à l'allèle sauvage, à l'allèle nul et à l'allèle 1% (cf. chapitre III, page 59, cet allèle correspond probablement à une faible expression du gène);
- ✓ Nous pouvons utiliser un hétérologue du gène *PpAPT* comme gène marqueur de sélection négative (présence de DAP) ou positive (en présence de l'adénine) dans une souche *APRT<sup>-</sup>*. Le problème réside dans le fait que la souche *APRT<sup>-</sup>* a un phénotype sévèrement altéré. Ceci exclut toute analyse phénotypique des mutations apportées. Une régulation conditionnelle de la mutation du gène *PpAPT* permettra de contourner ce problème: nous effectuerons la sélection dans la souche *APRT<sup>-</sup>* conditionnelle en absence d'induction et nous effectuerons les analyses phénotypiques en condition d'induction. L'analyse dans la souche sauvage étant préférable, nous proposons cette technique en attendant un meilleur candidat.

Le gène *PpAPT* placé entièrement sous le contrôle du système GVG permettra de continuer la caractérisation de ce système.

- ✓ Le phénotype de l'allèle nul est attendu en absence d'induction si l'activité basale est effectivement négligeable;
- ✓ Une condition d'induction doit exister permettant d'observer le phénotype sauvage. Ceci démontrera que l'activité stable induite peut atteindre au moins celle du gène *PpAPT* endogène;
- ✓ Une condition d'induction intermédiaire devra correspondre au phénotype de l'allèle 1%.

Nous pourrions démontrer une bonne flexibilité du système GVG si les trois objectifs sont atteints. Aucun phénotype secondaire ne devra être observé dans ces conditions d'induction.

## Discussion

### 1. La mutagenèse ciblée chez *P. patens* : état des lieux

Différents groupes de chercheurs ont déjà commencé à employer la mutagenèse ciblée pour étudier la fonction des gènes (Tableau 14).

Tableau 14. Mutagenèse ciblée de différents gènes chez *P. patens*.

Gène	Référence	Caractéristiques du fragment homologue (taille en kb)	Type de vecteur	RTF, 10 <sup>-4</sup> vecteurs circulaires §	RTF, 10 <sup>-4</sup> vecteurs linéaires §	Efficacité de la mutagenèse ciblée GT, %
<i>λ108</i>	(Schaefer and Zryd, 1997)	3.6	ADNg	I	2.9	100%
<i>λ420</i>		2.7	ADNg	I	0.82	100%
<i>λ213</i>		2.3	ADNg	I	0.5	5.8-6.7
	chapitre II	0.6 + 1.3	ADNg, brèche	I	0.1	1.23
						20%- 75%
						20%-50%
<i>ZLAB1</i>	(Hoffmann <i>et al.</i> , 1999); chapitre I	1.0	ADNg	I	0.47	0.7-2.7
	chapitre II	0.35 + 0.5	ADNg, brèche	I	0.07	0.8
						25%-50%
						0%-25%#
<i>PpAPT</i>	chapitre III	de 0.5 + 0.7 à 1.6 + 2.8	ADNg	R	0.86-3.9	3.2-11.0
	(Schaefer <i>et al.</i> , en préparation)	0.25 + 0.4	ADNc	R	0.06	0.63-0.68
						≈100%
						≈0 - 36%
<i>MCBI</i>	(Girod <i>et al.</i> , 1999)	0.74 + 0.58	ADNc	R		0.46
						2 / 55 ≈ 4 %
<i>Δ-6</i> désaturase	(Girke <i>et al.</i> , 1998)	1.1 + 1.2	ADNg	R		0.27
						5 / 5 = 100%
<i>ftsZ</i>	(Strepp <i>et al.</i> , 1998)	0.25 + 0.66	ADNc	R		0.43
						7 / 52 ≈ 14 %

§ - La valeur RTF est calculée par division du nombre de transformants intégratifs par le nombre total de régénérants. Nous présentons la gamme des valeurs moyennes.

# - intégration dans un locus homologue de la même famille ;

I - vecteur d'insertion et R - vecteur de remplacement ;

Les données non disponibles sont présentées par des cases vides.

L'analyse de l'ensemble des résultats nous permet de répondre aux questions suivantes:

- ✓ De quels facteurs l'efficacité de la mutagenèse ciblée dépend-elle?
- ✓ Pris avec les résultats des analyses moléculaires, ces résultats donnent-ils une indication sur les mécanismes de la recombinaison homologue chez *P. patens*?

- ✓ Pourquoi *P. patens* se démarque-t-elle des autres plantes? Quelles perspectives voyons-nous dans cette situation particulière et pour la mutagenèse ciblée dans la biologie végétale?

### 1.1. De quels facteurs la mutagenèse ciblée dépend-elle?

Le terme "*efficacité de la mutagenèse ciblée*" (Tableau 14, GT pour "gene targeting", page 122) indique la proportion des événements d'intégration ciblée parmi tous les événements d'intégration. Cette valeur caractérise l'aptitude du système expérimental à intégrer l'ADN *par recombinaison homologue* dans les conditions données. C'est une caractéristique *relative* (comparaison entre l'intégration ciblée et illégitime).

La valeur de la *fréquence relative de transformation* (Tableau 14, RTF) nous donne une information sur l'aptitude du système expérimental à intégrer l'ADN dans les conditions données. C'est une caractéristique *absolue* (est-ce que l'ADN s'intègre bien?). Bien évidemment, une comparaison des valeurs RTF pour différentes conditions permet de trouver des caractéristiques relatives. A titre d'exemple, une valeur RTF caractéristique de l'intégration d'un fragment homologue plus élevée que celle caractéristique d'un fragment non homologue atteste de l'efficacité de l'intégration par recombinaison homologue.

Nous pouvons par conséquent évaluer la mutagenèse ciblée en termes relatif et absolu.

Les paramètres principaux qui peuvent influencer la mutagenèse ciblée sont le type de vecteur, la longueur du fragment homologue et la forme du vecteur (linéaire ou circulaire surenroulé). Nous devrions évaluer la mutagenèse ciblée pour une série de vecteurs en variant le paramètre dont nous voulons mesurer l'importance et maintenant les autres paramètres constants.

#### *Type de vecteur (remplacement et insertion)*

Plusieurs études comparent l'efficacité absolue de l'intégration des vecteurs d'insertion ou de remplacement (valeur RTF). Les vecteurs d'insertion sont généralement plus efficaces chez la levure *S. cerevisiae*, (Hastings *et al.*, 1993; Leung *et al.*, 1997), d'un facteur allant jusqu'à 20 (Leung *et al.*, 1997). Chez la souris, les résultats sont un peu contradictoires. En absence de variations polymorphiques entre le vecteur et le locus chromosomique ciblé, les vecteurs d'insertion sont plus efficaces, d'un facteur de 5 à 10 (Hasty *et al.*, 1994). Dans une autre étude, l'efficacité des deux types de vecteurs a été jugée similaire (Deng and Capecchi, 1992). Hasty (Hasty *et al.*, 1994) explique les résultats de la deuxième étude par l'utilisation de fragments non isogéniques. Le même type d'étude (Adair *et al.*, 1998) sur les cellules CHO démontre que les vecteurs testés sont plus efficaces d'un facteur de 2 en configuration de remplacement plutôt que d'insertion. Pour les deux derniers systèmes expérimentaux, c'est principalement la structure des recombinants qui dépend du type du vecteur et non l'efficacité de la mutagenèse ciblée en général.

Nous n'observons pas de différence significative entre les vecteurs d'insertion ou de remplacement (Tableau 14, RTF ou GT). Il est notable qu'aucune étude ne permet de comparer rigoureusement l'efficacité des deux types, puisque nous n'avons pas étudié l'intégration des deux types de vecteurs dans un même locus. Nous pensons pouvoir effectuer

une telle étude comparative très rapidement en utilisant les vecteurs déjà testés en configuration "vecteur de remplacement" pour le gène *PpAPT*. En effet, ils peuvent être linéarisés à l'intérieur de l'homologie, dans le fragment 5' ou 3' ou encore sur le site de la délétion créant ainsi une brèche (vecteur d'insertion).

#### *Longueur du fragment homologue*

Nous pouvons vérifier l'influence de la longueur du fragment homologue sur l'intégration des vecteurs de remplacement dans le gène *PpAPT*. L'efficacité de la mutagenèse ciblée (Tableau 14, GT, page 122) est d'environ 100% avec un fragment homologue de 1.3 + 1.6 kb, basé sur l'ADNg. Lorsque le fragment est plus court (0.25 + 0.4 kb) et que l'homologie est interrompue (séquence ADNc), l'efficacité de l'intégration ciblée est toutefois suffisante pour que ces événements ( $\approx 10\%$ ) puissent être isolés. Il est à noter que la longueur de l'homologie ininterrompue est de 52 bp dans un bras du vecteur. Nous pensons que dans cette dernière étude, la qualité et la longueur de l'homologie deviennent un facteur limitant de l'intégration qui affecte non seulement l'efficacité de la mutagenèse ciblée, mais également la précision et le type des événements d'intégration. En effet, un cas d'intégration présente des délétions ou un gain de séquences dans les fragments hybrides. Nous discuterons des mécanismes qui peuvent être impliqués dans l'intégration de l'ADN chez *P. patens* plus loin (page 127).

En termes d'efficacité absolue (RTF), l'intégration des fragments isogéniques homologues plus longs (ADNg, entre 0.5+0.7 et 1.3+2.8 kb) est plus efficace que celle des fragments non isogéniques courts (ADNc, 0.25 + 0.4 kb). Nous n'avons pas trouvé le seuil minimal exact au-dessous duquel nous ne détectons plus l'intégration par recombinaison homologue, mais il se trouve vraisemblablement au-dessous de quelques centaines de bp. Le seuil maximal se trouve probablement autour de la longueur de 0.5 + 0.7 kb; au-dessus de cette longueur l'efficacité de l'intégration n'augmente plus de façon significative.

En somme, la mutagenèse ciblée est efficace de 4% à 100% avec aussi peu que 0.25 à 4 kb d'homologie (Tableau 14); ces valeurs sont plus proches de celles observées chez *S. cerevisiae* que de celles de tout autre système.

#### *Forme du vecteur : circulaire surenroulé et linéaire*

Les ruptures double brin sont l'inducteur de la recombinaison selon tous les modèles proposés pour les systèmes expérimentaux étudiés. Elles induisent les mécanismes de réparation des ruptures double brin sans pour autant modifier le rapport entre l'intégration par recombinaison homologue et illégitime. Chez la levure *S. cerevisiae*, les vecteurs d'insertion linéaires s'intègrent mieux que les vecteurs circulaires d'un facteur de 10 à 3000 (Orr-Weaver *et al.*, 1981; Rothstein, 1991). Chez le hamster (CHO) aussi, les vecteurs d'insertion linéaires s'intègrent mieux, d'un facteur plus modeste de 2 à 2.5. Les vecteurs de remplacement avec une coupure sur l'extrémité 5' ou 3' s'intègrent mieux d'un facteur de 3.5 à 4 par rapport aux mêmes vecteurs en forme circulaire (Adair *et al.*, 1998). Le vecteur d'insertion s'intègre plus fréquemment chez la souris (ES) d'un facteur de 5 à 12 (Hasty *et al.*, 1992). Chez la souris, l'intégration est plus fréquente, d'un facteur de 50 à 1000, lorsque des ruptures double brin sont induites *in vivo*, sans que le rapport entre les intégrations illégitime et ciblée change (Cohen-Tannoudji *et al.*, 1998; Smih *et al.*, 1995).

Nous estimons l'influence de la forme des vecteurs (linéaire ou circulaire surenroulé) sur la fréquence de l'intégration, séparément pour les vecteurs d'insertion et de remplacement. En effet, lorsqu'un vecteur circulaire interagit avec la cible chromosomique, nous ne pouvons prédire si le produit de la réaction sera l'insertion par un crossing-over, la conversion de la copie chromosomique ou extra-chromosomique ou un échange réciproque des séquences. Seuls les événements d'insertion ont pu être sélectionnés pour les vecteurs d'insertion de notre étude. En revanche, tous les événements ou seuls les événements de remplacement ont été sélectionnés pour les vecteurs de remplacement. De plus, les contraintes ne sont pas les mêmes lors de l'intégration des deux types de vecteurs. La rupture double brin produite par une coupure dans le fragment homologue d'un vecteur d'insertion présente deux extrémités adjacentes. Leur interaction avec le duplex d'ADN chromosomique est convergente, même si une brèche est formée. L'insertion peut passer par la formation d'une seule boucle D (ou un crossing-over). Au contraire, les ruptures double brin produites par une coupure aux extrémités du fragment homologue d'un vecteur de remplacement présentent deux extrémités non adjacentes. Leur appariement avec le duplex d'ADN chromosomique ne peut être résolu par une seule boucle D ou un seul crossing-over. La contrainte n'est donc pas la même que pour les vecteurs d'insertion (Adair *et al.*, 1998; Hasty *et al.*, 1994).

#### *Vecteurs de remplacement*

Les valeurs RTF présentées dans le Tableau 14 permettent de comparer la fréquence d'intégration de la forme circulaire et linéaire des vecteurs de remplacement du locus *PpAPT*. L'intégration de la forme linéaire est plus efficace d'un facteur de 2 à 10, rappelant l'augmentation observée pour le gène homologue dans le système CHO qui est d'un facteur de 3.5 à 4 (chapitre III et (Adair *et al.*, 1998)). La comparaison entre les deux systèmes n'est pas directe car le vecteur utilisé dans cette dernière étude avait une des extrémités 5' ou 3' bloquées par les séquences plasmidiques. Nous n'évaluons pas l'influence de la forme sur l'efficacité de la mutagenèse ciblée pour ces vecteurs (valeurs relatives, Tableau 14, GT) car les événements d'intégration illégitime n'ont été isolés pour aucun d'eux.

#### *Vecteurs d'insertion*

La valeur RTF a été grandement supérieure dans la transformation avec les vecteurs linéaires comparée à la transformation avec les vecteurs circulaires surenroulés. Ce phénomène n'est pas influencé par la présence d'une brèche entre les extrémités de la rupture. L'augmentation de la valeur RTF n'est pas forcément accompagnée d'une augmentation de l'efficacité de l'intégration ciblée dans la région d'homologie (chapitres I et II). En effet, la linéarisation du vecteur ne favorise pas uniquement l'intégration ciblée mais également l'intégration non ciblée. La rupture double brin stimule la machinerie de la réparation par recombinaison homologue tout autant que la réparation par NHEJ. Cette stimulation explique également le fait que le vecteur pHP23 sans homologie s'intègre mieux (valeur RTF) lorsqu'il est coupé.

Du point de vue pratique, l'introduction de ruptures double brin dans l'ADN transformant (in vivo ou en utilisant l'ADN linéaire) permet d'augmenter la fréquence de son intégration. Sans favoriser spécifiquement la recombinaison homologue et par là, l'efficacité de la mutagenèse ciblée, cette méthode permet d'augmenter sa fréquence absolue (valeur RTF) et de rendre l'isolation des événements d'intégration ciblée plus efficace (voire tout simplement possible) dans les systèmes où la recombinaison homologue est la voie mineure de réparation de

ruptures double brin (Gorbunova and Levy, 1997; Jasin, 1996). Chez *P. patens*, l'utilisation des vecteurs linéaires permet d'obtenir le même effet sans que ce soit une nécessité absolue.

## 1.2. Les modèles qui décrivent le mécanisme de l'intégration de l'ADN transformant chez *P. patens*

Est-il vraiment nécessaire d'étudier les mécanismes d'interaction entre l'ADN transformant et le chromosome? A notre avis, mis à part l'intérêt scientifique pur, l'importance pratique est évidente: la compréhension des mécanismes d'intégration d'ADN dans le génome de *P. patens* résoudra la nécessité d'une vérification systématique de la précision de l'intégration.

Nous n'avons pas planifié notre étude afin d'établir un modèle qui décrive le mécanisme de l'intégration de l'ADN transformant chez *P. patens*. Nous aimerions toutefois remarquer les caractéristiques suivantes:

- ✓ Un plasmide linéaire peut s'intégrer dans le chromosome selon un processus décrit comme un crossing-over (chapitres I-III);
- ✓ Un fragment homologue transformant interrompu par une brèche peut être réparé selon l'information présente dans la région homologue chromosomique dans une partie des cas d'intégration (chapitre II);
- ✓ Le vecteur peut être corrigé de façon conservatrice selon la matrice génomique, c'est-à-dire sans modifier la matrice (chapitre II et III);
- ✓ Nous avons observé des cas de recombinaison homologue intrachromosomique entre les séquences homologues situées en *cis* dans le chapitre IV (alox-17 instable);

L'intégration dans la majorité des cas est précise. Nous avons démontré la précision de l'intégration de l'ADN transformant chez *P. patens* par analyse de séquence (chapitre I, (Hoffmann *et al.*, 1999)) ou par le RFLP des sites d'intégration conforme aux prédictions. Les études récentes des autres groupes attestent d'une même précision du mécanisme (Girke *et al.*, 1998; Strepp *et al.*, 1998).

Nous observons donc chez *P. patens* les phénomènes qui ont servi à l'établissement des modèles DSBR (réparation de ruptures double brin) de Szostak, SDSA (appariement de brins dépendant de la synthèse) et SSA (appariement d'un brin simple) (Paques and Haber, 1999) qui décrivent les mécanismes précis d'intégration et de réparation de ruptures double brin. Nous pensons par conséquent que ces mécanismes sont très efficaces chez *P. patens*.

Chez les angiospermes, ces modèles ne peuvent pas expliquer le fait que dans certains cas seule une extrémité homologue est employée lors de la recombinaison. Pour les décrire, différents groupes de chercheurs proposent le modèle d'invasion d'une seule extrémité (OSI - "one side invasion") qui implique la participation de la ligature des extrémités indépendante de l'homologie (NHEJ - "non homologous end joining") (Britt, 1999; Gorbunova and Levy, 1999; Puchta, 1998; Puchta *et al.*, 1996). Ce modèle prévoit qu'une grande partie des événements d'intégration ne sera pas précise à cause de la recombinaison illégitime par NHEJ. En effet, plusieurs études démontrent que la majorité des cas d'intégration de l'ADN transformant ne sont pas précis et sont accompagnés d'insertion d'ADN de remplissage ("filler DNA"), de délétions dans le locus ciblé et d'intégration ectopique de copies additionnelles du vecteur (modifié ou non) (Britt, 1999; Gorbunova and Levy, 1997; Risseuw *et al.*, 1995; Salomon and Puchta, 1998; Takano *et al.*, 1997). Ces événements proviennent certainement

du mécanisme même de l'intégration d'ADN transformant dans le génome végétal et sont probablement indépendants de la stratégie de mutagenèse utilisée.

Chez *P. patens*, Girod et collaborateurs ont observé un cas d'intégration imprécise d'un côté du site d'insertion (Girod *et al.*, 1999). Ces chercheurs expliquent ce cas par le modèle OSI. Ils ne fournissent malheureusement pas d'analyse complète de cet événement et nous ne savons pas si les fragments imprévus sont dus à l'imprécision du mécanisme ou à celle de l'analyse.

Schaefer et collaborateurs analysent la séquence des produits d'intégration des vecteurs de remplacement, basés sur l'ADNc du gène *PpAPT* (Schaefer *et al.*, en préparation). Ils mettent en évidence une intégration des copies du vecteur modifiées aux extrémités dans un cas. Nous pensons que la recombinaison homologue est une étape limitante de l'intégration de ces vecteurs: l'intégration du bras 5' de ces vecteurs est sujette à des contraintes due à la courte taille d'homologie ininterrompue (52 bp). Une telle cinétique d'intégration favorise probablement des activités nucléasique (délétion) ou synthétique (gain de séquences) pendant ou avant la recombinaison.

Nous avons observé des cas minoritaires d'intégration **ciblée** d'ADN que nous ne pouvons pas décrire par les modèles précis tel DSBR ou SSA (chapitre II et III). Nous pensons donc que l'intégration par une voie impliquant NHEJ (comme le prévoit le modèle OSI) n'est pas prédominante chez *P. patens*. L'intégration illégitime 10 fois moins fréquente que l'intégration ciblée indique également que la participation de NHEJ dans l'intégration de l'ADN transformant est faible chez *P. patens*.

Des différences substantielles existent entre les organismes quant à l'implication des deux extrémités de l'ADN dans l'interaction entre les molécules homologues (les modèles SDSA, SSA, OSI). L'importance de la participation des extrémités créées par la rupture double brin reste une question ouverte chez *P. patens*. En effet, malgré leur effet stimulateur de recombinaison, il n'est pas démontré qu'elles initient la synthèse comme étape essentielle de la recombinaison. Chez la souris, Hasty propose un modèle décrivant l'intégration par recombinaison homologue avec ou sans crossing-over sans qu'une synthèse soit nécessaire (Hasty *et al.*, 1992). L'invasion de l'ADN cible se ferait par un brin simple du vecteur non rompu à l'intérieur de la région homologue. Ce modèle n'explique toutefois pas la réparation des brèches. Afin de se prononcer sur l'importance de tels mécanismes chez la mousse *P. patens*, il est nécessaire d'effectuer des expériences visant à démontrer le destin des extrémités lors de l'intégration de l'ADN transformant linéaire. Une approche rapide consisterait à étudier l'inclusion ou l'exclusion des séquences hétérologues attachées à une ou aux deux extrémités des fragments homologues linéaires dans les produits d'intégration.

*Quelle indication l'intégration de copies multiples peut-elle donner?*

Chez la levure *S. cerevisiae* les événements de remplacement ont lieu plutôt par conversion de gènes que par échange réciproque de séquence et l'intégration de plusieurs copies lors d'un remplacement n'est pas observée. En revanche, les événements d'insertion intègrent plusieurs copies du vecteur en répétitions directes (Hastings *et al.*, 1993; Orr-Weaver and Szostack, 1983; Orr-Weaver *et al.*, 1981; Rothstein, 1991). Chez la souris, les événements d'insertion intègrent majoritairement une seule copie du vecteur; par contre, il existe des événements de remplacement multiple (Hasty *et al.*, 1994; Hasty *et al.*, 1992). L'intégration d'une copie lors

d'un événement d'insertion indique probablement que l'appariement et l'invasion des extrémités du vecteur en quête d'homologie et la résolution de la structure intermédiaire d'intégration (boucle-D) sont des réactions plus rapides que la formation de concaténats (supposant que ces derniers se forment par recombinaison homologue extra-chromosomique). En revanche, si la vitesse d'intégration devient une étape limitante, les concaténats ont le temps de se former avant l'intégration. Chez la souris, cette situation peut avoir lieu avec les vecteurs de remplacement car ceux-ci ont probablement plus de contraintes lors de l'intégration à cause de la résolution des structures intermédiaires des deux extrémités ou des deux crossing-overs. L'intégration de plusieurs copies du vecteur de remplacement indiquerait alors que la vitesse d'intégration est supérieure à la vitesse de formation des concaténats.

Chez *P. patens*, les deux types de vecteurs (d'insertion ou de remplacement) peuvent s'intégrer en une ou plusieurs copies. La ligation des extrémités des vecteurs linéaires et la recombinaison extrachromosomique permettent une formation de concaténats chez *P. patens* (Ashton *et al.*, 2000; Schaefer *et al.*, 1994), ce travail]. Sur la base de l'hypothèse précédente, nous postulons que l'intégration des deux types de vecteurs est relativement efficace et que la vitesse de cette réaction est similaire à celle de la formation de concaténats puisque une ou plusieurs copies des deux types de vecteurs peuvent s'intégrer.

En remarque générale, nous mesurons la difficulté de comparer les événements d'intégration entre *P. patens*, d'un côté, et les cellules ES de la souris et CHO du hamster, de l'autre. Premièrement, la technique de transformation de ces cellules de mammifère est l'électroporation alors que nous utilisons la précipitation au PEG pour transformer la mousse. La quantité d'ADN transformant se trouvant dans la cellule est certainement différente et peut influencer le nombre de copies intégrées du vecteur. De surcroît, l'efficacité de l'intégration par recombinaison homologue est grandement inférieure chez ces cellules de mammifères et ces événements ne sont isolés que grâce à des méthodes sophistiquées de sélection. En effet, la participation de la voie NHEJ à l'intégration de l'ADN transformant est plus importante chez les mammifères. Nous nous contenterons donc d'une comparaison empirique.

## **2. Why me ? ou Pourquoi la mutagenèse ciblée par recombinaison homologue est-elle si efficace chez *P. patens* et non chez les angiospermes ?**

### **2.1. Les tentatives de développer la mutagenèse ciblée par recombinaison homologue chez les angiospermes**

Dans l'approche directe, la mutagenèse ciblée par recombinaison homologue chez les angiospermes est inefficace, comme nous l'avons revu dans l'Introduction. Différents groupes de chercheurs ont alors recherché un moyen de l'améliorer en testant plusieurs paramètres: la méthode de transfert de l'ADN dans les cellules végétales et le type de cellules transformées, la configuration du vecteur de transformation, ou encore l'utilisation de sélection positive et négative (Mengiste and Paszkowski, 1999; Offringa and Hooykaas, 1995; Thykjaer *et al.*, 1997; Vergunst and Hooykaas, 1999) Ces recherches n'ont pas apporté l'amélioration souhaitée.

Quant à la méthode de transfert, il a été particulièrement vexant de trouver que la bactérie *A. tumefaciens*, l'agent de transfert végétal par excellence, permet un transfert d'ADN qui s'intègre par recombinaison homologue chez *S. cerevisiae* (Bundock *et al.*, 1995; Risseuw *et al.*, 1996).

En ce qui concerne la phase du cycle cellulaire, il n'y pas d'études approfondies qui vérifient la compétence des cellules végétales en différentes phases du cycle cellulaire.

Une sélection positive et négative devrait permettre d'enrichir la population de transformants en événements d'intégration ciblée en supprimant l'intégration aléatoire, mais cette approche n'a encore permis d'obtenir aucune amélioration (Mengiste and Paszkowski, 1999; Thykjaer *et al.*, 1997).

Une autre solution consiste à augmenter l'efficacité de la mutagenèse ciblée par l'induction de la recombinaison homologue même. Différents chercheurs tentent d'augmenter la fréquence de la recombinaison à l'aide de l'introduction des ruptures double brin *in vivo* (Chiurazzi *et al.*, 1996; Puchta, 1999; Puchta *et al.*, 1996), et à l'aide de la réparation des brèches laissées par l'excision des éléments transposables (Xiao and Peterson, 2000). Nous pouvons émettre une critique générale sur ce progrès technique: seuls les sites prédéterminés, préparés à l'aide d'une insertion d'un site de coupure ou d'un élément transposable, pourront être ciblés ce qui réduit considérablement la généralisation de cette approche (Hohn and Puchta, 1999). De plus, l'induction de la réparation des ruptures double brin permet d'augmenter la fréquence absolue de tous les événements de recombinaison (homologue ou illégitime) et par conséquent ne diminue pas la participation de NHEJ à l'intégration de l'ADN transformant.

#### *Stimulation hypothétique de la recombinaison homologue*

Divers chercheurs supposent que la sur-expression des gènes codant pour les enzymes de la recombinaison homologue pourrait élever la fréquence de la mutagenèse ciblée. L'expression de la protéine bactérienne RecA nucléaire chez le tabac élève la fréquence de recombinaison homologue intrachromosomique d'un facteur de 10 (Reiss *et al.*, 1996). Cependant, il n'y pas de preuve que la protéine RecA hétérologue influence l'efficacité de la mutagenèse ciblée chez les plantes (Puchta *et al.*, 1996; Reiss *et al.*, 2000). Cette dernière étude ne démontre aucune augmentation significative de l'efficacité de la mutagenèse ciblée pour deux locus différents. L'action de la protéine RecA se limite à l'induction de l'échange de chromatides sœurs ainsi qu'à l'amélioration de la précision de la réparation des ruptures double brin. Les auteurs expliquent l'incapacité de la protéine RecA à induire la mutagenèse ciblée par le mode de transfert de l'ADN par *A. tumefaciens*. Son rôle est peut-être passif, celui d'empêcher l'interaction de l'ADN avec les complexes de réparation par NHEJ. Le mode de transfert (*A. tumefaciens*) en soi ne devrait pourtant pas être responsable de l'inefficacité de la mutagenèse ciblée chez les plantes puisque l'intégration ciblée de l'ADN transformant introduit à l'aide de *A. tumefaciens* est efficace chez la levure *S. cerevisiae* (Bundock *et al.*, 1995).

Shalev et collaborateurs démontrent une stimulation de la recombinaison extrachromosomique et intrachromosomique ainsi que des crossing-overs somatiques chez le tabac exprimant le gène bactérien codant pour une résolvasse RuvC (Shalev *et al.*, 1999). L'étude ne démontre pas d'amélioration de l'efficacité de la mutagenèse ciblée.

Dans l'approche inverse, la répression de la recombinaison illégitime peut être sélectivement inhibée. La répression de la voie NHEJ est importante pour la mutagenèse ciblée puisqu'il est supposé que sa participation active provoque des altérations même quand une intégration a lieu par recombinaison homologue.

La situation de la mutagenèse ciblée par recombinaison homologue dans la biologie végétale pourrait être débloquée par des études des facteurs qui régissent ce processus chez les plantes. Des mutants de recombinaison homologue d'*A. thaliana* sont connus mais pas les gènes

responsables. La fonction des orthologues végétaux des gènes impliqués dans la recombinaison homologue chez la levure reste inconnue. L'isolation d'un gène codant pour une protéine de type SMC ("structural maintenance of chromosomes") nécessaire à la recombinaison homologue chez *A. thaliana* (Mengiste *et al.*, 1999) confirme la suppression de la recombinaison homologue par la structure de la chromatine mais n'offre aucune solution pour développement de la mutagenèse ciblée.

L'accroissement de l'efficacité de la mutagenèse ciblée par ces méthodes reste encore à être démontré. Un point important sera la démonstration que l'expression ou la suppression des gènes impliqués dans la recombinaison homologue ou illégitime n'aura pas d'effets pléiotropiques. De façon plus subtile, l'expression de ces transgènes pourrait même interférer avec les mutations étudiées d'un cas à l'autre. Aucune étude de fonction des gènes utilisant ces techniques ne sera valable si les plantes permettant une mutagenèse ciblée efficace ont un phénotype anormal.

#### *Thérapie génique par correction de "faute d'orthographe"*

Strauss a proposé une méthode de correction de "fautes d'orthographe" génétiques à l'aide d'hybrides ADN-ARN (Cura, 1999). Testée sur le tabac et sur le maïs (Hohn and Puchta, 1999), la technique ne permet pas de mutagenèse précise car, étonnamment, les mutations produites ne correspondent pas systématiquement à la mutation portée par l'ADN mutagène. L'efficacité de la méthode est significative (chez le maïs par exemple: 1 mutation par 10<sup>4</sup> cellules recevant l'ADN mutagène) mais doit être améliorée. Même si les améliorations souhaitées la rendent plus efficace et plus précise, cette technique ne permet pas d'introduire de mutations autres que ponctuelles, d'un ou de plusieurs nucléotides, et directement sélectionnables. Elle pourrait compléter mais pas se substituer à la mutagenèse par recombinaison homologue.

## 2.2. Pourquoi *P. patens* est-elle la seule plante terrestre qui intègre l'ADN exogène dans le génome préférentiellement par recombinaison homologue?

L'hypothèse que *P. patens* serait plus proche des micro-organismes et trop éloignée des angiospermes est douteuse, puisque les données issues des projets de séquençage des EST, des gènes nucléaires ou des organites indiquent que l'homologie entre *P. patens* et les angiospermes est très élevée. L'analyse de l'usage des codons montre que *P. patens* est plus proche des dicotylédones que celles-ci ne le sont des monocotylédones. La taille du génome est également dans la gamme trouvée habituellement chez les angiospermes (Reski, 1999; Reski *et al.*, 1994).

Une hypothèse explique cette efficacité par la phase dominante haploïde de *P. patens*. En effet, les autres organismes dans lesquels la recombinaison homologue somatique est plus efficace que la recombinaison illégitime ont pour la plupart une phase haploïde: *S. cerevisiae*, *S. pombe*, *A. nidulans*, *D. discoideum*. Schaefer (Schaefer, 1994) suppose que les organismes haploïdes se protègent de l'intégration illégitime pouvant être létale pour eux. Un contrôle sévère de la recombinaison homologue chez les angiospermes, diploïdes, empêche des réarrangements des chromosomes portant des séquences alléliques homologues. Certains auteurs pensent que l'imprécision des mécanismes œuvrant chez les végétaux contribue à l'augmentation de la diversité génique. Dépourvues de mobilité, les plantes auraient

développé ce système pour augmenter leur adaptabilité au milieu changeant (Britt, 1999; Gorbunova and Levy, 1999).

Chez les angiospermes, diplobiontiques et possédant beaucoup d'ADN répétitif, une pression sélective s'exerce en faveur du mécanisme NHEJ car une recombinaison homologue fréquente mettrait en danger l'intégrité du génome en favorisant des recombinaisons entre l'ADN répétitif des chromosomes homologues (Schaefer, 1994). Dans le cas des régions homologues dispersées, la recombinaison homologue peut amener à des translocations ou à des délétions chromosomiques.

Chez *S. cerevisiae*, la recombinaison est la plus efficace dans les cellules diploïdes hétérozygotes pour le locus de type sexué *MAT* et dans les cellules haploïdes mutées ayant aboli la suppression des deux locus *HMRa* et *HMLα*. Elle est moins efficace dans les cellules haploïdes ou diploïdes homozygotes pour le locus de type sexué. La relation entre l'efficacité de la recombinaison homologue et la ploïdie est moins directe que les chercheurs ne le souhaiteraient et dépend plutôt de l'état d'un locus particulier chez cet organisme (Lee *et al.*, 1999).

Une autre vision consiste à expliquer cette dominance de la recombinaison homologue non par la phase haploïde de *P. patens* mais par sa phase gamétophytique. La régulation de la recombinaison pourrait être très différente entre les tissus sporophytique et gamétophytique, comme le propose Reski (Reski, 1999).

La phase du cycle cellulaire pourrait être une autre explication du phénomène. Le protonéma dont les protoplastes à transformer sont issus, se trouve à la transition G2/M la plupart de la journée dans les conditions de culture utilisées (Reski, 1999; Reski *et al.*, 1994). Pour un organisme haploïde, c'est la phase dans laquelle la réparation des ruptures double brin peut se faire par recombinaison homologue mitotique car les chromatides sœurs peuvent servir de matrice. Les cellules des angiospermes, diplobiontiques, ne sont pas aussi strictement synchronisées que celle de *P. patens*. Il se pourrait alors que la transition G2/M favorise la recombinaison homologue mitotique observée chez *P. patens*. Chez la levure, l'importance de la recombinaison homologue en fonction de l'avancement du cycle cellulaire n'est pas évidente (Galli and Schiestl, 1998; Haber, 2000).

En général, force est de constater que les facteurs responsables de l'efficacité de la mutagenèse ciblée par recombinaison homologue ne sont connus ni chez *P. patens* ni chez la levure *S. cerevisiae*.

### **3. Rôle de *P. patens* dans la biologie végétale et dans la biologie des eucaryotes pluricellulaires**

Dans le contexte actuel, *P. patens* paraît, en effet, irremplaçable. La mutagenèse ciblée est efficace et précise dans **les conditions naturelles**. Plus important encore, c'est un des rares systèmes **pluricellulaires eucaryotes** permettant une mutagenèse ciblée. Ces conditions ne sont réunies dans aucun autre système expérimental.

D'un côté, la levure *S. cerevisiae* possède la méthodologie de la mutagenèse ciblée la plus complète mais ne permet pas d'étudier les processus spécifiques aux organismes pluricellulaires. L'étude de Girod et collaborateurs de la fonction du gène MCB1 en est un exemple très convaincant (Girod *et al.*, 1999). Ces chercheurs montrent que la mutation de ce gène provoque une altération de développement chez *P. patens* mais non chez *S. cerevisiae*.

D'un autre côté, nous venons de démontrer que nous sommes loin de l'application de la mutagenèse ciblée aux angiospermes, les plantes qui présentent le plus d'intérêt pour la recherche biotechnologique. Parmi les systèmes expérimentaux pluricellulaires, on ne compte que la souris, le hamster ou la lignée des cellules de lymphocyte B du poulet. Certainement, ces systèmes resteront irremplaçables pour étudier les processus spécifiques aux animaux. L'efficacité de la mutagenèse dans ces systèmes est toutefois basse et la régénération des organismes mutants est lente et n'est possible que pour la souris.

A notre avis, la recherche de *P. patens* doit alors suivre au moins deux directions:

#### *Organisme modèle*

Les chercheurs doivent profiter de la mutagenèse ciblée chez *P. patens* pour étudier les processus biologiques communs à toute la biologie végétale, et plus généralement, à la biologie des organismes pluricellulaires. En effet, *P. patens* est suffisamment représentative du règne végétal et l'idée que les mousses peuvent être utilisées en tant que modèle n'est pas nouvelle (Cove *et al.*, 1997; Knight, 1994; Reski, 1998b; Reski, 1999; Schaefer, 1994; Schaefer, 2001; Wang and Cove, 1989). Ces auteurs montrent également que la méthodologie de *P. patens* est à la hauteur d'un organisme modèle. A ces revues s'ajoute le travail de Nishiyama et collaborateurs (Nishiyama *et al.*, 2000), qui ont employé la transformation par une banque d'ADN génomique étiquetée par des vecteurs navettes ("tagged mutagenesis and gene trap by shuttle mutagenesis") afin d'"étiqueter" ou de "capturer" les gènes par une intégration ciblée d'un gène de résistance ("étiquetage") ou d'un gène rapporteur sans promoteur ("capture"). Ces chercheurs ont déjà produit une collection de quelque cinq mille transformants dont 4% affiche un phénotype mutant et une collection de quelque cinq mille transformants exprimant le gène rapporteur (Nishiyama *et al.*, 2000). Ce type de technique n'existe que pour les organismes dans lesquels la recombinaison homologue est efficace, à l'instar de *S. cerevisiae* et il est plus efficace que l'étiquetage aléatoire pratiqué chez *A. thaliana* ou *O. sativa* (Jeon *et al.*, 2000; Parinov and Sundaresan, 2000), particulièrement pour étiqueter des gènes à copie unique de petite taille.

#### *Etudes de la mutagenèse ciblée par recombinaison homologue*

Il est illusoire de croire que la recherche sur *P. patens* toute seule répondra à plus de questions que la recherche sur *S. cerevisiae* ne l'a déjà fait. En revanche, la différence connue entre les voies de recombinaison de différents organismes laisse espérer que *P. patens* contribuera à trouver des modèles de recombinaison nouveaux et, peut être, spécifiques au règne végétal. De ce point de vue là, elle est la plante la mieux placée pour découvrir la recombinaison homologue. La compréhension de la "gestion" de l'intégration de l'ADN exogène chez *P. patens* peut se fonder sur des études comparatives, d'un côté, de *P. patens* et des angiospermes, et, d'un autre côté, de *P. patens* et de *S. cerevisiae*. Il faudra rechercher les homologues des gènes connus chez *S. cerevisiae* et étudier leur fonction chez *P. patens*. Des pistes parallèles doivent être empruntées: les mutants de *P. patens* incapables de recombinaison homologue, de NHEJ ou de la réparation de brèches sont le premier pas vers la découverte des facteurs spécifiques à *P. patens* et, peut-être, aux végétaux. L'analyse de l'intégration de l'ADN exogène dans différentes conditions (tissus, phase du cycle cellulaire,

phase de croissance et de développement) de *P. patens* et des angiospermes trouvera des facteurs supplémentaires.

Afin d'avoir les premières indications, nous pourrions tester l'efficacité de la recombinaison homologue chez *P. patens* dans différentes conditions:

- ✓ Nous transformerions les protoplastes extraits des cellules arrêtées en phase G1 ou G2 (utilisation des mutants qui doivent être isolés);
- ✓ Nous régènerions et transformerions les tissus naturellement diploïdes (sporophyte) ou artificiellement diploïdes (fusion somatique). Une aptitude différente à la mutagenèse ciblée (avec une même aptitude à la transformation et à la régénération ce qui reste à mettre en place) indiquerait l'importance de la ploïdie ou de la phase de développement (gamétophyte contre sporophyte). Nous proposons l'utilisation de promoteurs inductibles spécifiques au sporophyte (pollen, par exemple) afin d'induire les événements de recombinaison. Le jeu optimal de ces paramètres favorisant la recombinaison homologue sera probablement difficile à trouver.

Il est évident que la connaissance des facteurs responsables de la recombinaison homologue et de la voie NHEJ chez *P. patens* et les autres plantes aiderait dans cette démarche. Quant aux déterminants génétiques du choix de la voie de la réparation des ruptures double brin, ils ne sont connus chez aucun organisme. Les trouverons-nous chez *P. patens*?

En conclusion, nous émettons l'hypothèse que les études fondamentales sur *P. patens* expliqueront les facteurs qui rendent la mutagenèse ciblée efficace dans ce système mais n'avanceront peut-être pas le développement de ces techniques chez les autres plantes. Nos raisonnements sont les suivants:

Premièrement, plusieurs indices attestent que l'incapacité des angiospermes à intégrer l'ADN exogène de façon précise ciblée réside dans les mécanismes mêmes de l'intégration. Aucune condition physiologique normale n'a été trouvée dans laquelle l'intégration par recombinaison homologue serait favorisée. Même si les facteurs qui en sont responsables sont trouvés, il sera nécessaire de les modifier afin de changer les mécanismes d'intégration. A condition que ce soit réussi, les études à l'aide de la mutagenèse ciblée devront néanmoins se faire dans des plantes de type non sauvage, et ceci compromettra la validation des résultats.

Deuxièmement, l'efficacité de la mutagenèse ciblée par recombinaison homologue chez *P. patens* n'incite pas les chercheurs à mettre en place des méthodes sophistiquées d'isolation des cas d'intégration ciblée et précise. Or, au vu de ce qui précède, ces méthodes pourraient être la meilleure approche à suivre pour les angiospermes pour effectuer les études sur des organismes en conditions physiologiques.

Malgré les questions auxquelles nous cherchons toujours des réponses, la biologie se doit de profiter des connaissances déjà très importantes sur la mutagenèse ciblée chez *P. patens* qui permettent d'étudier tous les processus biologiques de *P. patens* en tant que mousse, plante et organisme pluricellulaire eucaryote.

## Matériel et méthodes

### *Souches bactériennes et clonage moléculaire*

Nous utilisons les souches bactériennes *E. coli* HB101 et DH5 $\alpha$  à des fins de clonage moléculaire. Les techniques de base de clonage moléculaire sont selon Sambrook (Sambrook *et al.*, 1989). Nous effectuons les réactions d'amplification par PCR dans un thermocycleur (Biometra) selon les indications du fournisseur, à savoir la matrice est 1 ng d'ADN plasmidique ou une fraction d'une colonie bactérienne porteuse du plasmide substrat dans un volume total de 50  $\mu$ l (Hi-Fi Expand System, instructions de Boehringer-Mannheim). Le cycle répété 33 fois se compose comme suit: 94° 15s, 55° 45 s, 68° 45s, (précédé de 4 min pour lyser les colonies à 94° et terminé par 5 minutes d'élongation à 72°). La liste des amorces mentionnés et utilisés dans ce travail se trouve dans l'Annexe A, page 136.

Les vecteurs de transformation de *P. patens* sont tous des plasmides isolés des cultures de *E. coli* et purifiés selon la méthode au polyéthylène-glycol (Sambrook *et al.*, 1989). Les plasmides utilisés sont listés dans l'Annexe B, page 137.

### *Souches de la mousse *P. patens* et conditions de culture*

La souche de type sauvage seule est à la base de cette étude (Schaefer *et al.*, 1994). Toutes les souches mutantes utilisées dans cette étude ont été obtenues dans le cadre de ce travail et sont décrites dans les chapitres correspondants.

La mousse *P. patens* est mise en culture comme décrit Schaefer (Schaefer and Zryd, 1997), sur du milieu minimal PPNO3 ou PPNH4 (additionné de tartrate d'ammonium 2.7 mM), comme décrit Cove (Ashton *et al.*, 1979b). Les deux milieux peuvent être additionnés de glucose 25 mM, chlorure de thiamine 1.5  $\mu$ M, acide paraaminobenzoïque 1.8  $\mu$ M et acide nicotinique 8  $\mu$ M, donnant PPNO3c et PPNH4c, respectivement. Pour l'observation phénotypique, nous utilisons les 4 types de milieux. Pour l'amplification du matériel et la conservation de lignées à moyen terme, nous employons PPNH4c. Le milieu solidifié contient 0.7% d'agar (Merck 1614). Les conditions de culture sont : jour/nuit 16/8 heures à  $\sim 80 \mu\text{E m}^{-2} \text{s}^{-1}$  (Ashton and Cove, 1977). L'isolation de protoplastes se toujours faite à partir de protonéma après 5 ou 6 jour de croissance sur le milieu PPNH4c, au besoin additionné de l'adénine (Fluka), de l'antibiotique (cf. plus bas, ajoutés après la stérilisation) ou de DAP (avant la stérilisation) (Schaefer and Zryd, 1997). Leur régénération se fait sur le milieu PPM solide dérivé du PPNH4c par ajout de 0.33M mannitol (Fluka) (Schaefer and Zryd, 1997).

### *Transformation de la mousse *P. patens**

Schaefer décrit les protocoles de la transformation des protoplastes de la mousse *P. patens* (type sauvage ou dérivés) à l'aide de transfert d'ADN au polyéthylène-glycol (PEG) (Schaefer, 1994; Schaefer *et al.*, 1994). Nous avons utilisé la version optimale (Schaefer and Zryd, 1997). Les paramètres clés sont comme suit: la quantité d'ADN transformant est de 15  $\mu$ g (10 $\mu$ g dans certaines transformations, le nombre de protoplastes transformés étant réduit à 200'000). En cas de co-transformation (deux vecteurs utilisés pour transformer le même protoplaste), le rapport de masse des deux vecteurs était habituellement de 1 : 1 (7.5  $\mu$ g : 7.5  $\mu$ g). La durée de choque thermique est de 5 minutes. La récupération de protoplastes en milieu liquide dure 24 heures avant la mise en culture sur le milieu solide. Les protoplastes sur le milieu solide régénèrent dans les mêmes conditions de culture que décrites ci-dessus sur le milieu PPM pendant 4-6 jours avant la mise en place de sélection. La sélection a été maintenue par changement de milieu solide additionné d'antibiotiques (hygromycine à 20 mg/L, Calbiochem, G-418 à 25-40 mg/L, Gibco (Schaefer and Zryd, 1997)) ou DAP (concentrations traitées dans la section des résultats, 250  $\mu$ M pour la sélection courante, Fluka) tous les 10 jours au moins. (Schaefer, 1994; Schaefer *et al.*, 1994; Schaefer *et al.*, 1991).

### *Expression transitoire des gènes rapporteurs (essais GUS, GFP)*

Nous employons le protocole de transformation décrit ci-dessus, ainsi que par Schaefer (Schaefer, 1994), pour les essais d'expression transitoire des gènes GUS ou GFP. Les protoplastes transformés sont maintenus en culture liquide (48 heures) avant les mesures. L'induction par dexaméthasone se fait par ajout de l'hormone (Riedel-deHäen, Allemagne) immédiatement après la transformation. Nous mesurons l'activité GUS selon Jefferson (Jefferson, 1987) dans l'extrait protéique brut au tampon SSE comme décrit par Schaefer (Schaefer, 1994). Le substrat fluorogénique était le méthylumbelliféryl glucuronide (MUG) et le produit mesuré méthylumbelliférone (MU) à l'aide de fluorimètre FluroscaanII®.

L'estimation de l'expression GFP s'est faite à l'aide de l'observation microscopique des cultures des protoplastes transformés (Davis and Vierstra, 1998). Nous comptons le nombre de protoplastes exprimant la protéine GFP, le niveau d'expression n'est estimé que qualitativement.

#### *Expression transitoire du gène codant pour la Cre recombinase*

Nous employons le même protocole de transformation (Schaefer, 1994) pour l'expression transitoire du gène codant pour la recombinase Cre du phage P1 (vecteur p35SCre, page 140). La régénération de protoplastes initiale se fait sans sélection. Le taux de transcription ou le taux en protéine Cre ne sont pas mesurés.

#### *Mutagenèse à l'aide d'irradiation UV*

Nous irradions le protonéma de 4-5 jours par une dose de 50 mJoule dans l'appareil de fixation d'ADN par UV, GeneLink (BioRad), ce qui correspond à environ 80 - 95 % de létalité dans les protoplastes régénérés. Le protonéma est mis à l'obscurité pendant 12 heures au moins pour inhiber l'activité de photolyase. Les protoplastes sont ensuite isolés (Schaefer and Zryd, 1997) du protonéma irradié et mis en culture sur le milieu PPM.

#### *Extraction de l'ADN et de l'ARN*

Nous extrayons l'ADN ou l'ARN total à partir de 2 à 20 grammes de tissus gamétophytiques constitués essentiellement de protonéma de 5 à 10 jours. L'ADN est extrait selon la méthode de CTAB (cétylemethylammonium bromide) comme décrit par Rogers (Rogers and Bendlich, 1988), l'ARN selon le protocole standard de Sambrook, avec adaptation aux cellules végétales (Sambrook *et al.*, 1989).

#### *Diagnostic par PCR*

Nous effectuons l'amplification par PCR dans un thermocycleur (Biometra) selon l'instruction du fournisseur (Gibco, Taq) avec 50 ng d'ADN génomique comme matrice, dans un volume total de 50 µl. Le cycle répété 33 fois se compose comme suit: 94° 15s, 55° 45 s, 72° 45s. Le produit est analysé sur gel d'agarose selon les procédures standards (Sambrook *et al.*, 1989). La liste des amorces utilisées dans ce travail se trouve dans l'Annexe A, page 136. Nous avons commandé l'analyse de séquence auprès de la maison Microsynth GmbH Sequencing Group (Balgach, Suisse).

#### *Analyse d'hybridation Southern et Northern*

Nous effectuons les analyses Southern et Northern selon les procédures standards (Sambrook *et al.*, 1989) et les indications du fournisseur (Zeta Probe, Biorad). Nous utilisons 2 à 3 µg par échantillon d'ADN et 10 à 15 µg d'ARN. Les sondes décrites dans les sections correspondantes sont marquées par amorces aléatoires à [ $\alpha$ -<sup>32</sup>P] dCTP selon le fournisseur des réactifs (Pharmacia, 1999). L'hybridation (16 heures, 62 °C, 3xSSC, tampon phosphate - SDS) et le lavage (0.1 SSC 65°C) correspondent à 99% d'homologie, selon Sambrook (Sambrook *et al.*, 1989) et les indications du fournisseur (Zeta Probe, Biorad). Les reconstructions génomiques correspondant à un nombre de copies du fragment reconnu par la sonde se basent sur une estimation de la taille du génome haploïde de  $4.6 \times 10^8$  bp selon Reski (Reski *et al.*, 1994).

## Annexe A: liste des amorces PCR

Liste des amorces PCR:

Nom	séquence
m1-5' (locus PpAPT)	atg ttc cga gat gtg acg acg
m1-3' (locus PpAPT)	gac aac ttc tca cga ccc tgc
PpAPT <sub>ex2</sub> SaII-5'	ctg taa att gtg tgt cga cag g
AtAPT1NcoI-5' (ou AtAPT <sub>NcoI</sub> F)	cat gcc atg gcg act gaa gat gtg ca
AtAPT1SphI-3' (ou AtAPT <sub>SphI</sub> R)	aca tgc atg cat gtt tct tgt taa gca gcc gac
AtAPT2-5'	atg ttt gcc gtg gag aat gg
AtAPT2-3'	tta taa ggt taa ttc atc aaa ctg g
AtAPT2NcoI-5'	gag tga gcc atg gtt gcc gtg g
AtAPT2HindIII-3'	ttg aat aga aaa agc ttt tat aag g
hptSpeI-5'	gaa cac tag tat atg aaa aag cct gaa ctc acc
CaMV <sub>ter</sub> NotI-3'	ttc gag cgc gcc cgc cac tgg
35SGVG-5'BstEII	ccc aca gat ggt gac cga gcc
e9 <sub>ter</sub> -3'	tca tga tca gat tgt cgt ttc c
nospr-5'	aaa ctg aag gcg gga aac gac
noster-3'	gat cta gta aca tag atg gac acc
UAS <sub>noster</sub> -5'	ggt gtc atc tat gtt act aga tc
3a-3' (bord)	gct gca agg cga tta agt tgg
GUS-5'SpeI	Vlach, S.J., disponible au laboratoire
GFP-5'SpeI	Vlach, S.J., disponible au laboratoire

## Annexe B: cartes des plasmides

L'information sur la construction des plasmides suit la carte seulement si elle n'est pas disponible ailleurs.

Les plasmides de base:

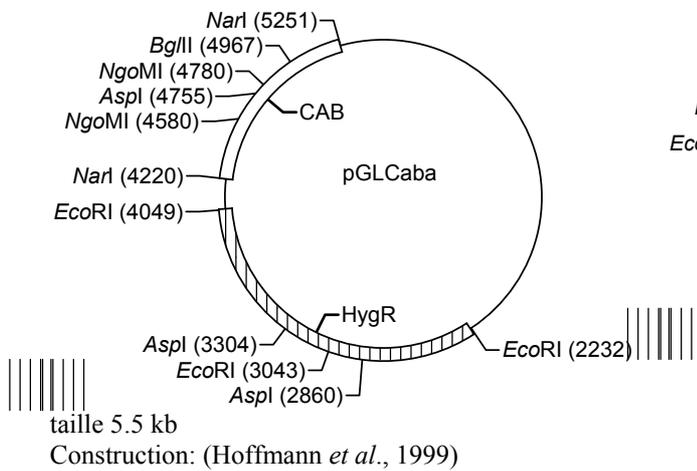
- ✓ pGL2 (Schaefer *et al.*, 1991), pHP23b (Paszkowski *et al.*, 1988), pBluescript, pUC21 et pBSM13+ (Stratagene, Suisse), pZLAB1 (Long *et al.*, 1989), pCOR102 et pCOR105 (McElroy *et al.*, 1991), pMM23 (p35SCre) (Dale and Ow, 1990), pTA7002 (Aoyama and Chua, 1997).

Les plasmides suivants sont dérivés de ces derniers et leur construction a été effectuée et décrite ailleurs:

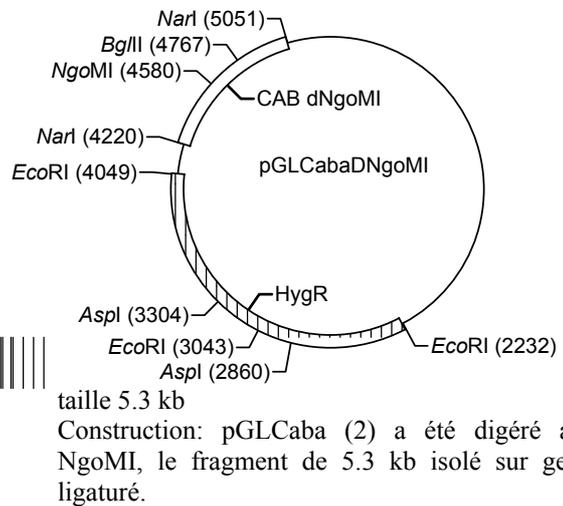
- ✓ pGL108 et pHP213 (Schaefer and Zryd, 1997), pGLCaba (Hoffmann *et al.*, 1999), pUVAN6.0 (Schaefer *et al.*, en préparation).

Nous présentons la carte des plasmides utilisés (les plasmides sont numérotés pour les renvois entre les constructions).

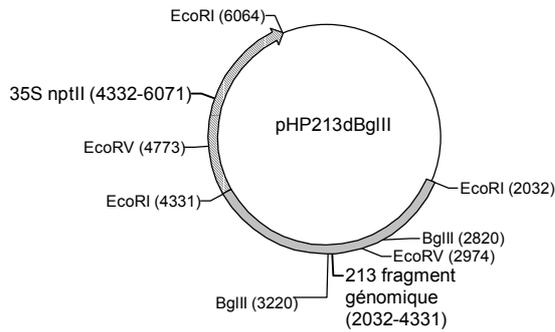
### 1. pGLCaba



### 2. pGLCabaΔNgoMI

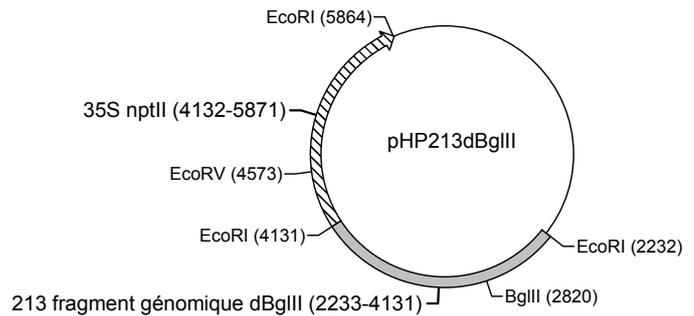


### 3. *pHP213*



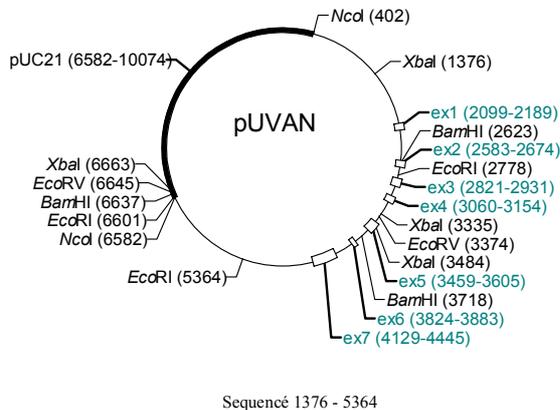
taille 6.67 kb  
Construction: (Schaefer and Zryd, 1997)

### 4. *pHPΔBglIII*



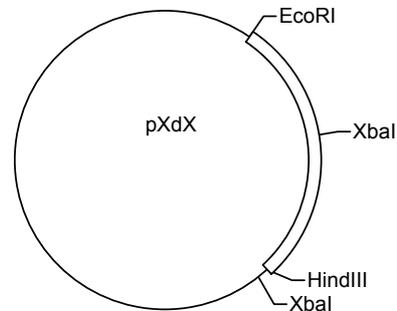
taille 6.27 kb  
Construction: plasmide pHP213 (3) digéré par BglIII, le fragment linéaire de 6.27 kb isolé sur gel et ligaturé.

### 5. *pUVAN 6.0*



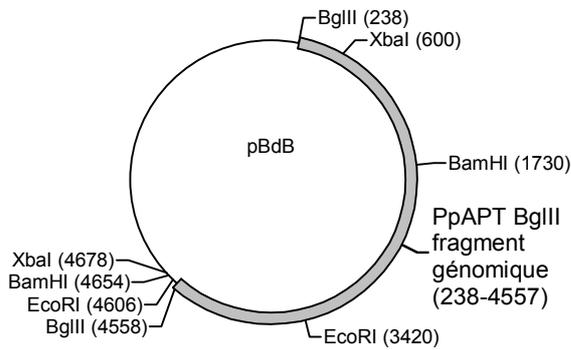
taille: 9.8 kb  
Construction : R. Buser, O. Blanc (Schaefer *et al.*, en préparation)  
Cette carte souligne le vecteur de clonage pUC21. Les lignes fines représentent le fragment NcoI du gène *PpAPT* et les boîtes indiquent les exons, numérotés.

### 6. *pXΔX*



taille 4.2 kb  
Construction: le fragment EcoRI-HindIII du gène *PpAPT* est cloné dans le vecteur pUC21δ (pUC21 dont le fragment entre les SalI -XbaI du polylinker est enlevé et les bouts remplis ligaturés) pour produire le plasmide pX. Ce dernier est partiellement digéré par XbaI, le fragment de 4.2 kb est isolé sur gel et ligaturé.  
*XΔX* - l'intron 2 et l'intron 4 (entre les exon 4 et exon 5) sont incomplets: l'intron 2 est tronqué en amont et l'intron 4 en aval. Le signal STOP TAA (HindIII) est suivi des séquences du vecteur pUC21δ.

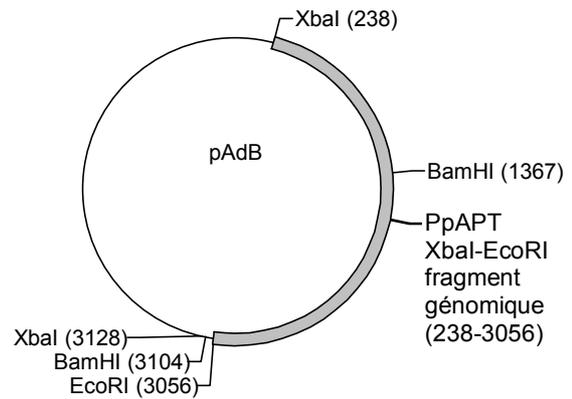
### 7. *pBAB*



taille 7.4 kb

Construction: le fragment BglIII contenant le gène *PpAPT* est sous-cloné dans le vecteur pUC21 pour produire le plasmide pB (non montré, construit par O. Blanc et R. Buser). Ce dernier est digéré par BamHI. Les fragments linéaires d'env. 4.3 et 3.1 kb sont isolés et ligaturés. L'orientation correspondant au vecteur pBΔB confirmée par une digestion EcoRI et XbaI.

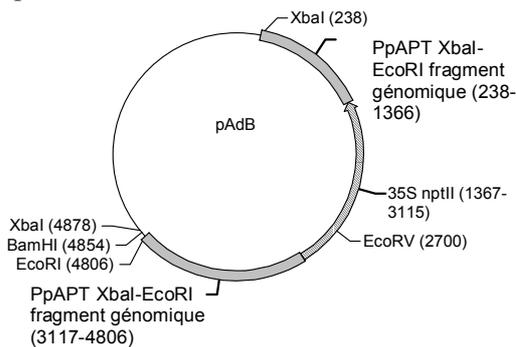
### 8. *pAΔB*



taille: 6 kb

Construction: Laloue et collaborateurs (Schaefer *et al.*, en préparation). Le fragment génomique XbaI-EcoRI contenant le gène *PpAPT* est cloné dans pBluescriptKS.

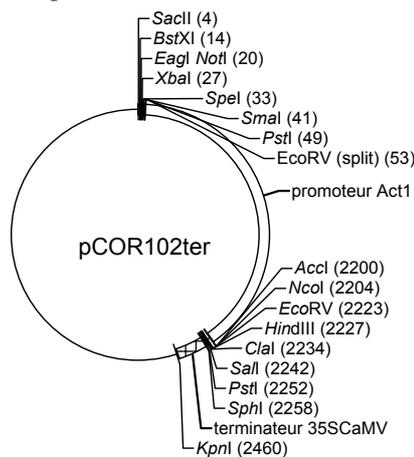
### 9. *pAΔBneo*



taille 7.6 kb

Construction: Laloue et collaborateurs (Schaefer *et al.*, en préparation). Le fragment EcoRI contenant la cassette 35S *nptII* du plasmide pHP23b cloné par remplissage de bouts francs dans le site BamHI du plasmide pAΔB, dans l'orientation antiparallèle au gène *PpAPT*

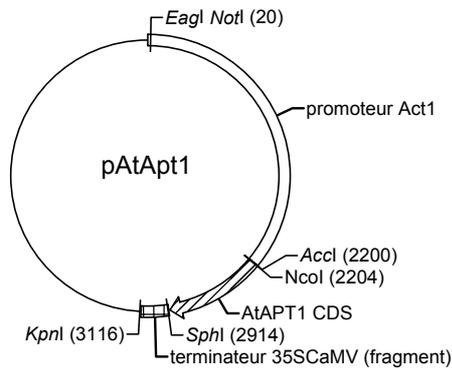
### 10. *pCOR102ter*



taille 5.4 kb

Construction: le terminateur CaMVter est découpé du plasmide pDH51 sous forme d'un fragment SalI-KpnI et inséré dans les sites correspondants du vecteur pCOR102 (McElroy *et al.*, 1991).

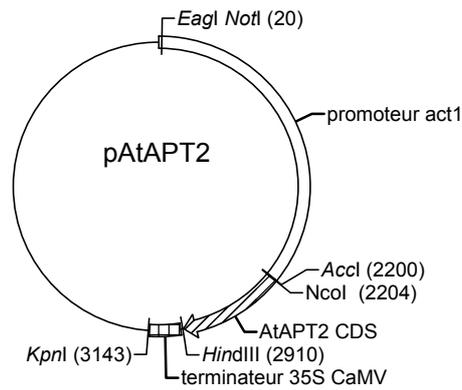
### 11. *pActAtAPT1*



taille: 6 kb

Construction: la séquence codante du gène *AtAPT1* est amplifiée par PCR à partir d'une banque d'ADNc de *A.thaliana* (don de P.-A. Girod) avec des amorces *AtAPT1*NcoI-5' et *AtAPT1*SphI-3'. Le fragment est coupé avec NcoI et SphI pour produire les extrémités collantes correspondantes et ligaturé avec le vecteur linéaire pCOR102ter (10) digéré avec NcoI et SphI.

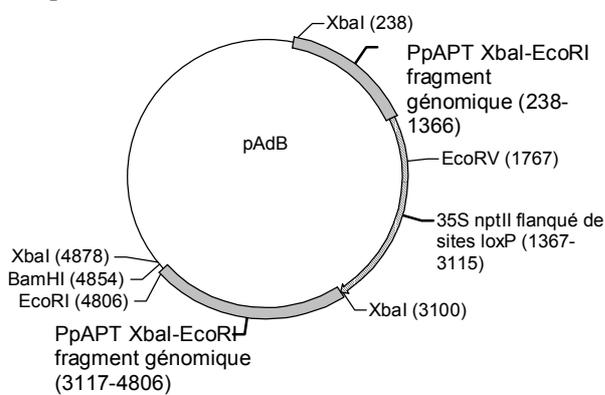
### 12. *pActAtAPT2*



taille: 6 kb

Construction: la séquence codante du gène *AtAPT2* est amplifiée à partir du plasmide *pAPTII* (Laloue et collaborateurs (Schaefer *et al.*, en préparation), ce plasmide contient l'ADNc *AtAPT2* cloné dans pBluescript) avec les amorces *AtAPT2*NcoI-5' et *AtAPT2*HindIII-3'. Le fragment est digéré avec NcoI et HindIII pour produire les extrémités collantes correspondantes et ligaturé avec le vecteur pCOR102ter linéaire coupé avec NcoI et HindIII.

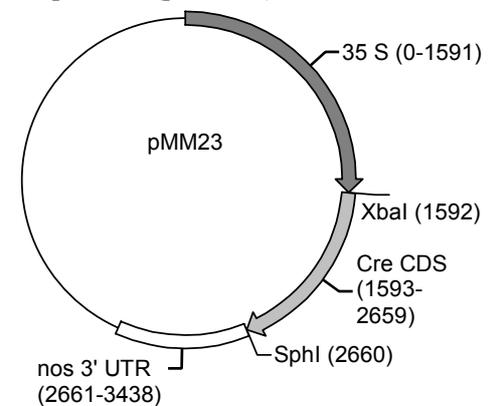
### 13. *pAloxneo*



taille 7.6 kb

Construction: Laloue et collaborateurs (Schaefer *et al.*, en préparation). Le fragment EcoRI contenant la cassette 35S*nptII* flanqué de sites loxP parallèles est cloné par remplissage de bouts francs dans le site BamHI du plasmide pAΔB, dans l'orientation parallèle au gène *PpAPT*

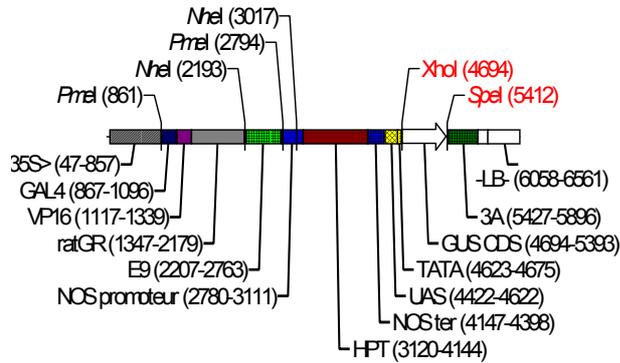
### 14. *pMM23 (p35SCre)*



taille 6.0 kb

Construction: (Dale and Ow, 1990)

### 15. pTA7002GUS

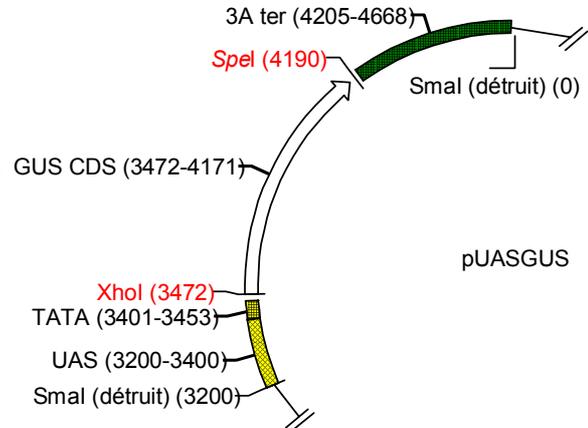


taille: 19 kb, la carte ne montre que les séquences qui se trouvent entre les séquences bordures gauche (LB) et droite (position 0, non montrée) du vecteur pTA7002 (Aoyama and Chua, 1997). Les éléments du plasmide sont expliqués en détail dans le chapitre V.

Construction: la séquence codante du gène GUS est amplifiée par PCR avec les amorces gus5'-XhoI et gus3'-SpeI à partir du plasmide pActF4GUS, et digéré par XhoI et SpeI pour produire les extrémités collantes correspondantes. Ce fragment, isolé sur gel, est ligaturé avec le plasmide pTA7002 (Aoyama and Chua, 1997) linéaire digéré par XhoI et SpeI en aval du promoteur UASGal4.

### 16. pUASGUS

(pGAL4UASGUS, les séquences pBSM13 sont partiellement présentées)

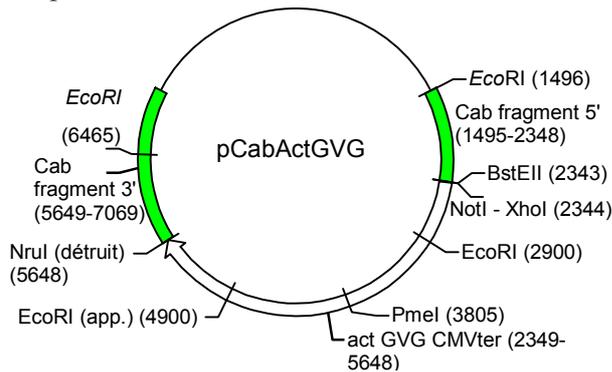


taille 3.9 kb

Construction: le fragment UAS-GUS-3a ter est amplifié par PCR avec les amorces UAS5'-nos et 3a-3' sur la matrice du plasmide pTA7002GUS et cloné dans le site SmaI du vecteur pBSM13 (Stratagène)

NB: le plasmide pTA7002GFP et pUASGFP sont obtenus selon les mêmes procédures avec la séquence codante GFP amplifié avec les amorces gfp5'-XhoI et gfp3'-SpeI sur la matrice du plasmide psmRSGFP (Davis and Vierstra, 1998).

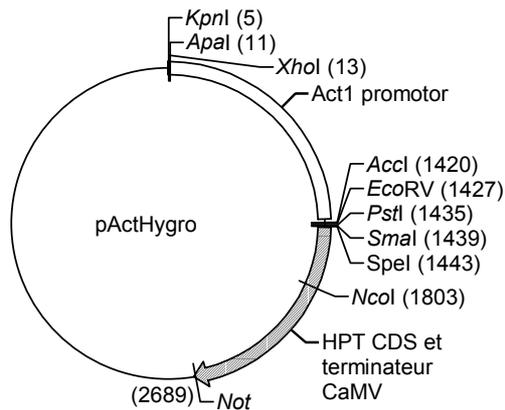
### 17. pCabActGVG



taille: 9.2 kb

Construction: le fragment XhoI-PmeI de 0.8 kb du plasmide pCab35SGVG-NX contenant le promoteur 35S est remplacé par le fragment XhoI-SpeI de 1.4 kb du plasmide pCOR105 (McElroy *et al.*, 1991) contenant le promoteur du gène act1 du riz. Pour cela, le fragment linéaire de 7.8 kb découpé avec XhoI et PmeI du plasmide pCab35SGVG (pages suivante) est isolé sur gel et ligaturé avec le fragment de 1.4 kb découpé avec XhoI et PmeI du plasmide pCOR105.

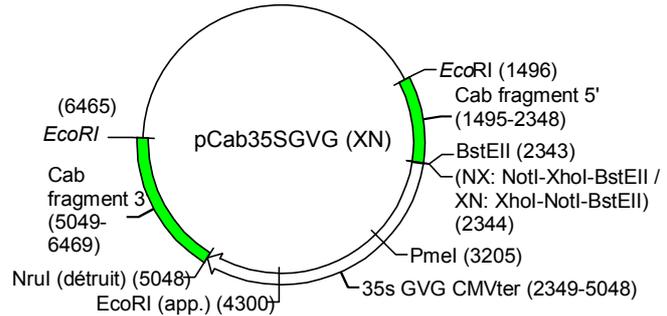
18. *pActhpt*



taille: 5.7 kb

Construction: la séquence codante du gène *hpt* (contenant le codon start ATG) suivi du terminateur CaMV est amplifiée par PCR avec les amorces *hpt*5'-SpeI et CaMV3'-NotI, isolée sur gel et digérée par SpeI et NotI pour produire les extrémités collantes correspondantes. Ce fragment est cloné dans les sites SpeI et NotI du vecteur pCOR105 (McElroy *et al.*, 1991). La mousse est transformée selon le protocole standard avec le vecteur *pActhpt*. Le nombre de protoplastes résistants est compris entre 5 et 10% des protoplastes régénérés d'une transformation avec le vecteur *pActhpt*. Cette valeur est similaire à celle observée généralement avec les gènes de résistance pilotés par le promoteur 35S en présence de concentrations habituelles de l'antibiotique de sélection (Schaefer, 1994; Schaefer *et al.*, 1994) Nous avons isolé des clones intégratifs (illégitimes car les vecteurs n'ont pas d'homologie avec le génome de *P.patens*) avec une fréquence RTF  $\leq 10^{-5}$ .

19. *pCab35SGVG (XN / NX)*



taille: 8.6 kb

Construction: la cassette d'expression 35S GVG (piloté par le promoteur nos) (Aoyama and Chua, 1997) est amplifiée avec les amorces 35SGVG-5'BstEII et e9-3' sur la matrice du plasmide pTA7002, isolée sur gel et digérée avec BstEII pour produire une extrémité collante. Ce fragment est ligaturé avec le fragment linéaire de pZLAB1 (Long *et al.*, 1989) de 6 kb coupé par BstEII et NruI pour produire le plasmide *pCab35SGVG* (utilisé dans les essais d'expression transitoire). Le polylinker NotI-XhoI-BstEII est introduit par clonage du linker `gtgacctcgagactagtgcggccg` (brin supérieur 5-3')/`gtgaccgcgccgccaactagtctcgag` (brin inférieur 5-3') dans le site BstEII du plasmide pour des clonages ultérieurs.

## Annexe C: analyse détaillée du RFLP de différents allèles

### Chapitre II

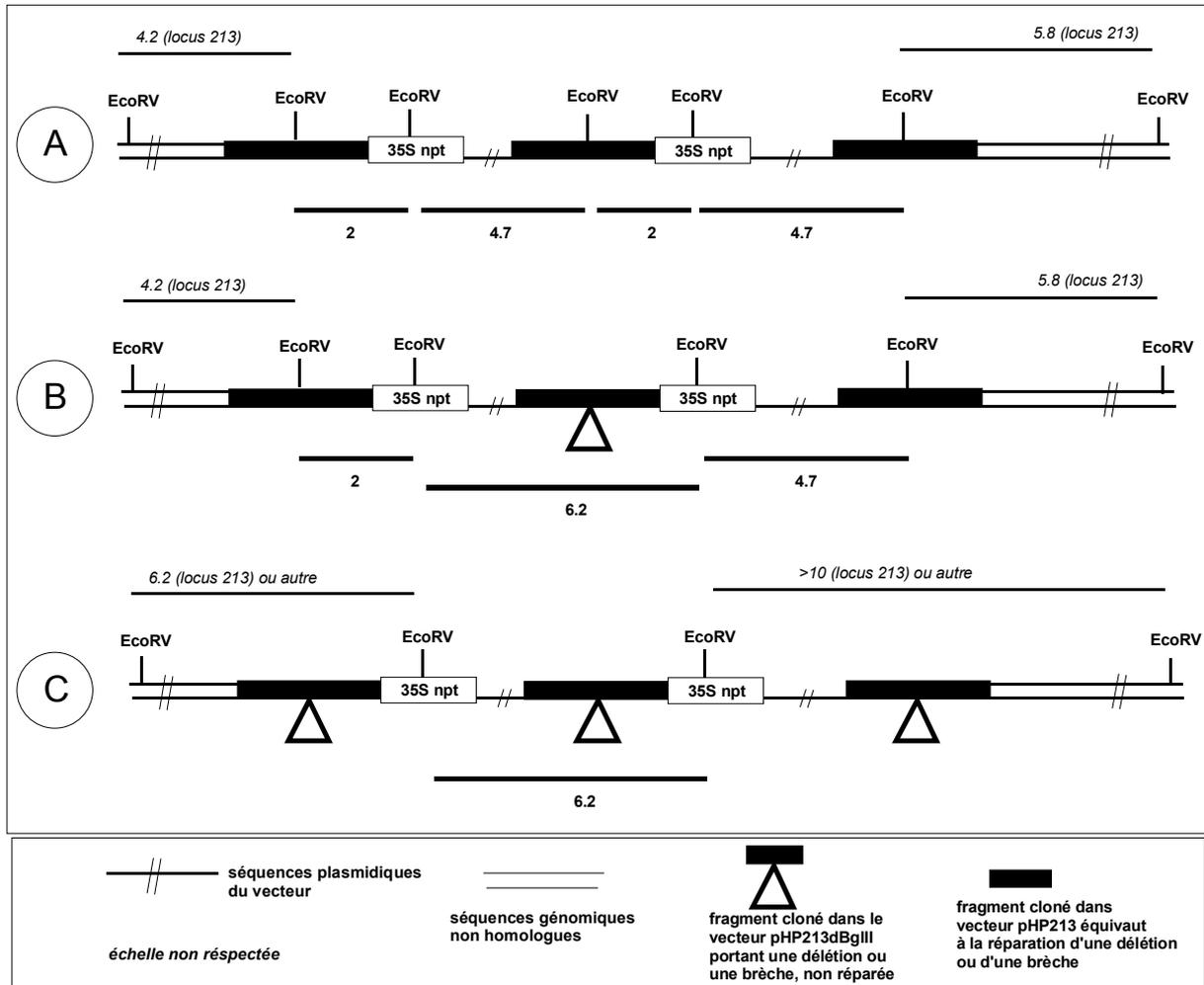


Figure 50. Schéma d'intégration du vecteur pHP213ΔBgIII avec et/ou sans réparation dans le locus  $\lambda$ 213.

Seuls les fragments hybrides seront différents suite à l'intégration dans un autre locus. Le fragment homologue porté par le vecteur pHP213ΔBgIII (rectangle noir) est utilisé comme sonde.

A - Intégration avec une réparation complète de toutes les copies des vecteurs

B - Intégration avec une réparation d'une partie des copies des vecteurs. Les copies internes ne sont pas réparées.

C - Intégration sans réparation.

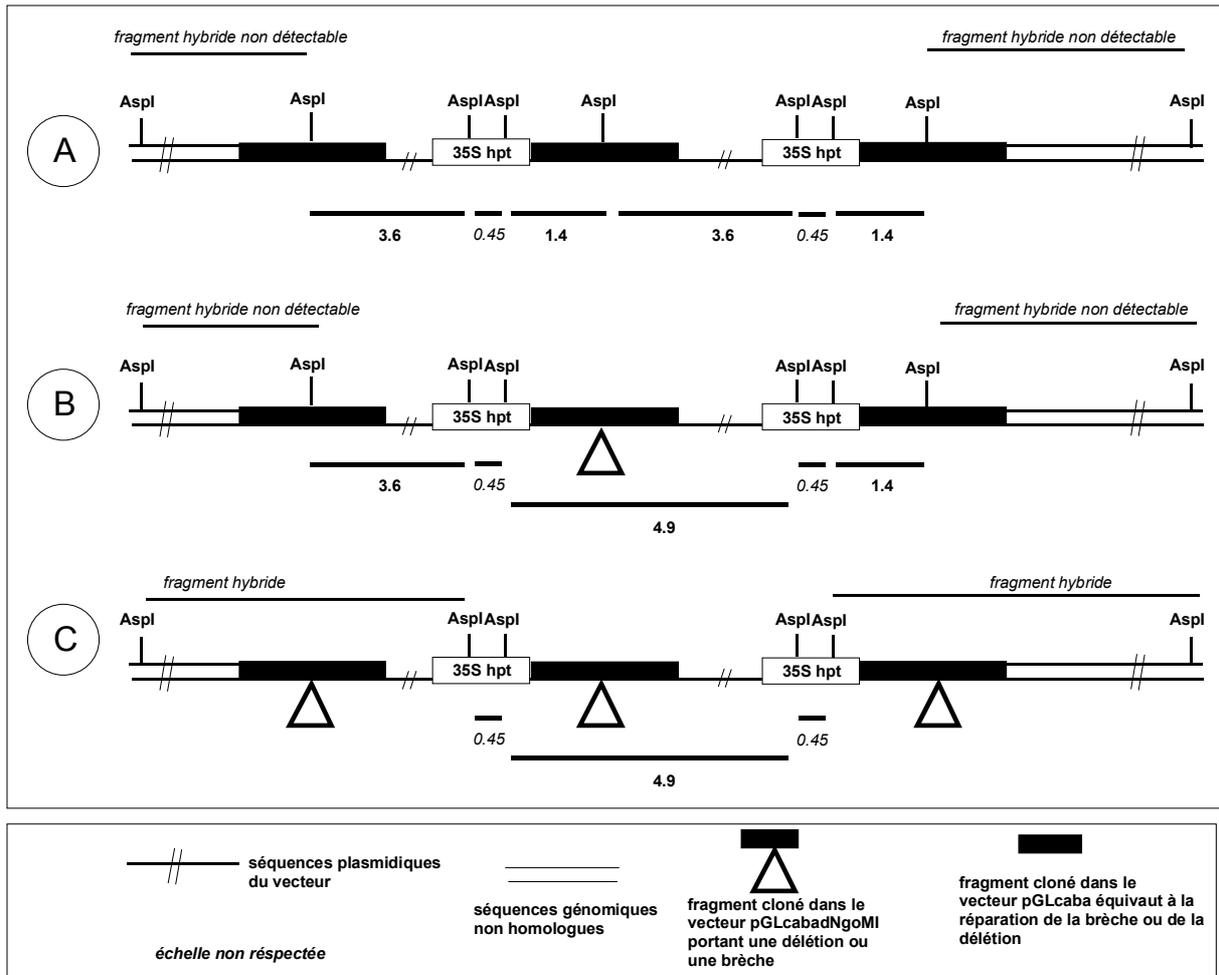


Figure 51. Schéma d'intégration du vecteur pGLCabaΔNgoMI avec et/ou sans réparation dans un gène *Cab*.

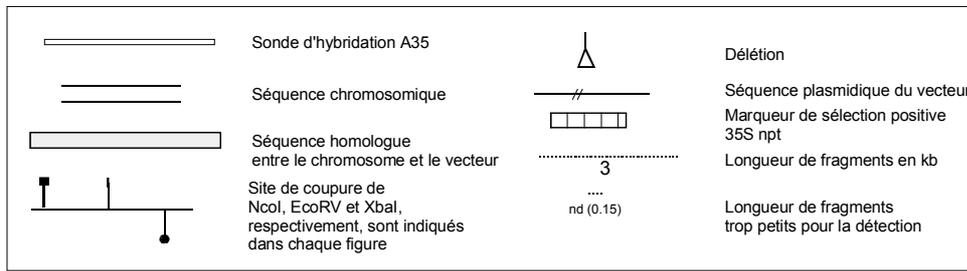
Seuls les fragments hybrides seront différents suite à l'intégration dans un autre locus. Le vecteur de clonage pGL2 (ligne fine et le marqueur de sélection) est utilisé comme sonde.

A - Intégration avec une réparation complète de toutes les copies du vecteur

B - Intégration avec une réparation d'une partie des copies du vecteur. Les copies internes ne sont pas réparées.

C - Intégration sans réparation.

Chapitre III



La signalétique ci-dessus s'applique aux figures suivantes.

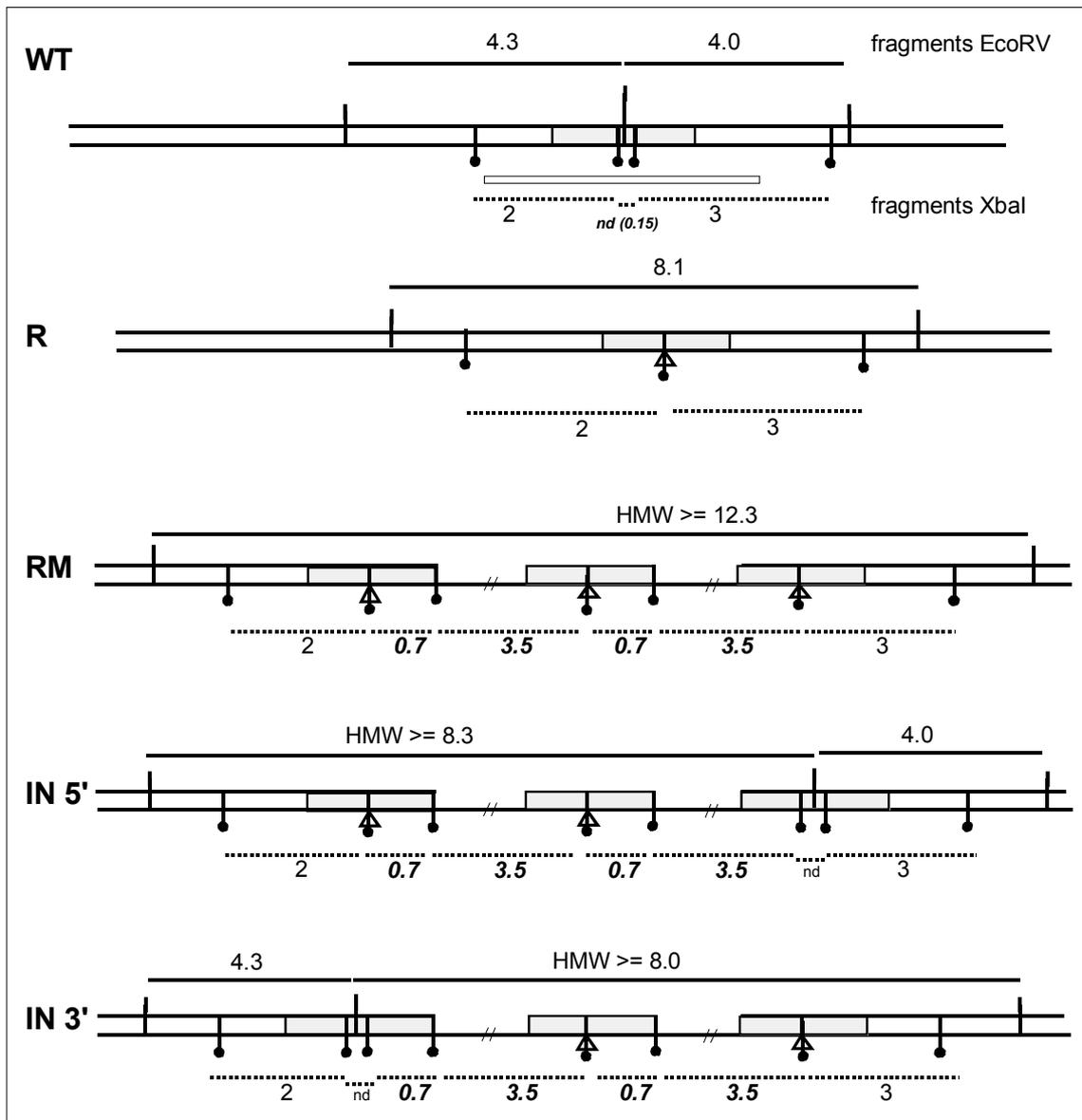


Figure 52. Le RFLP des allèles XAX et de l'allèle WT. Les répétitions du vecteur sont en italique.

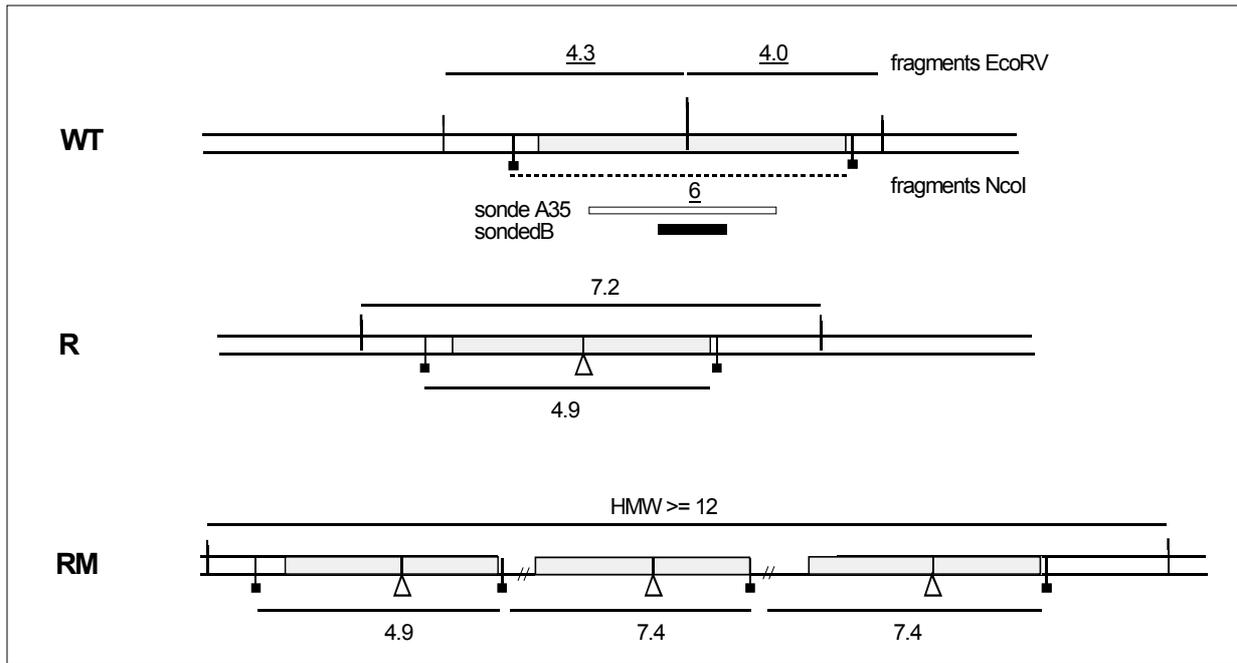


Figure 53. Le RFLP des allèles B $\Delta$ B et de l'allèle WT.

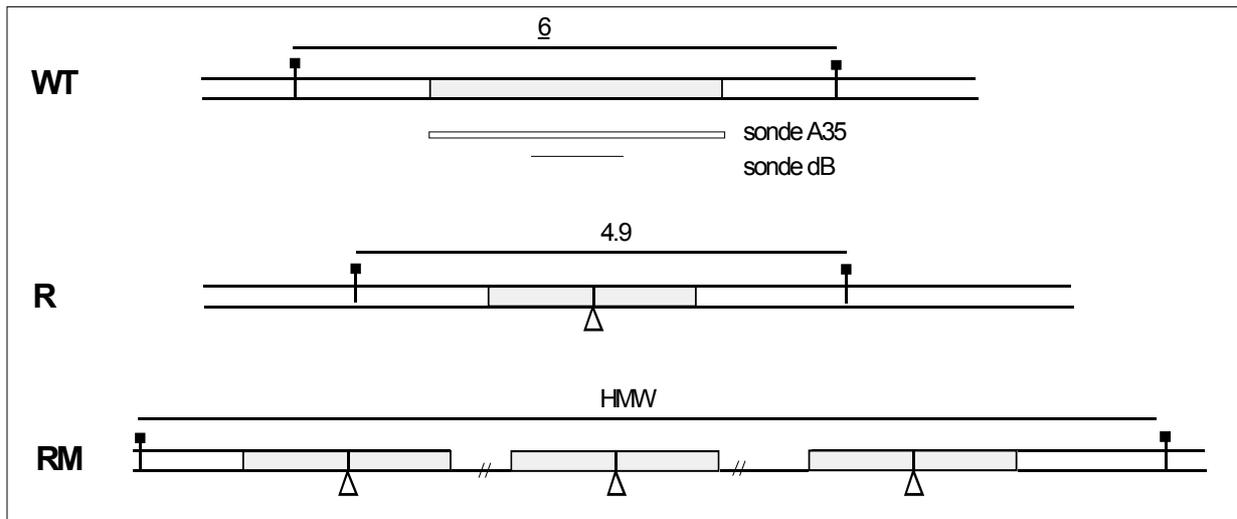
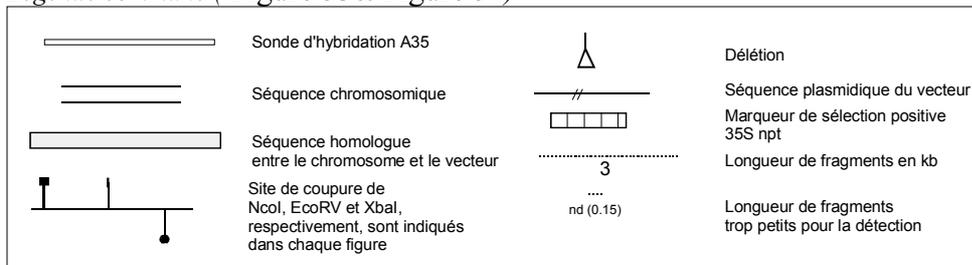


Figure 54. Le RFLP des allèles A $\Delta$ B et de l'allèle WT.

*Légende commune (Figure 53 & Figure 54)*



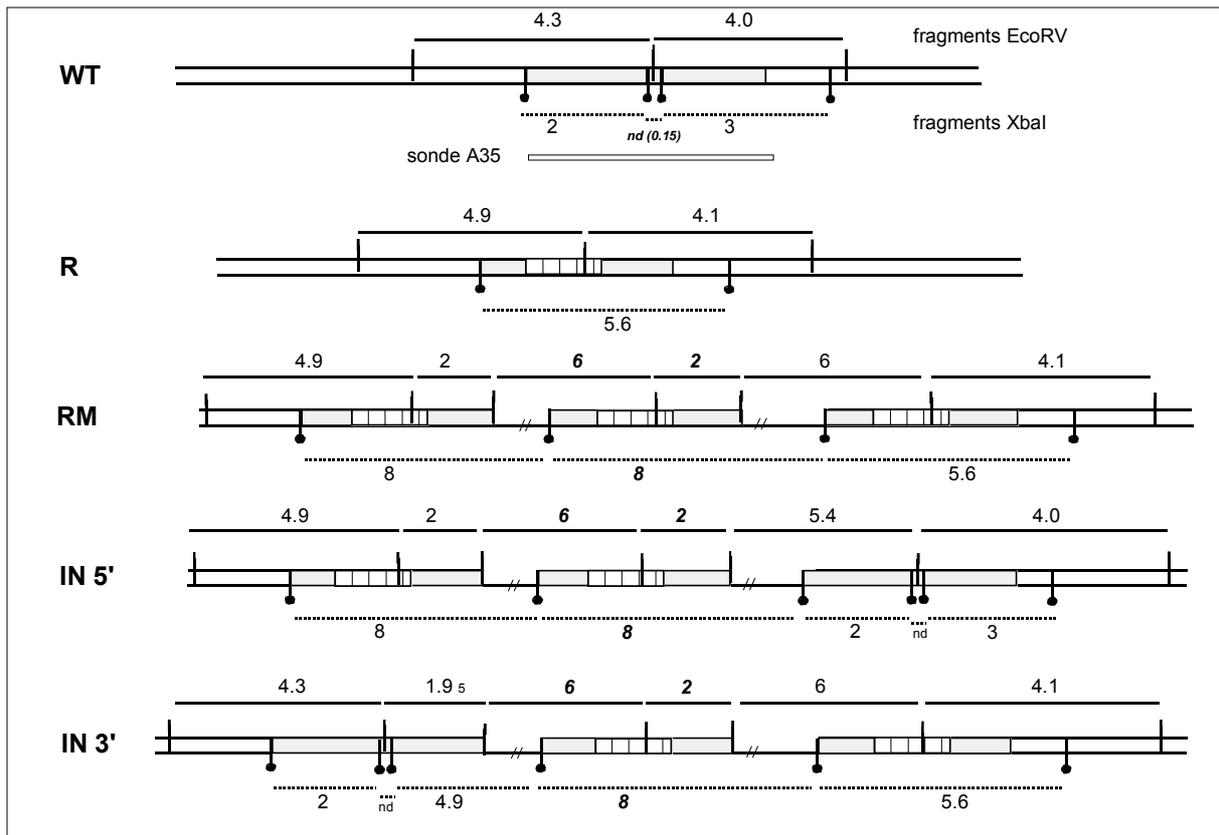


Figure 55. Le RFLP des allèles  $\Delta\Delta$ Bneo et de l'allèle WT.

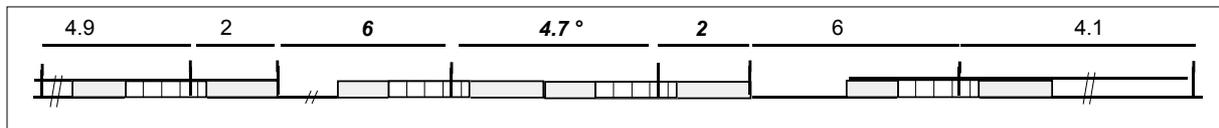
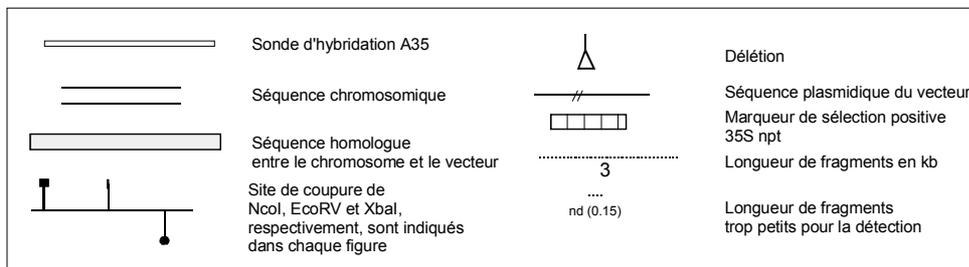


Figure 56. L'allèle anex105.

Le fragment de remplacement s'intègre en même temps et dans la même orientation que le vecteur entier en répétitions directs, formant un fragment EcoRV de 4.7 kb. Le fragment attendu de 4.7 kb est marqué de °.

Légende commune (Figure 55 & Figure 56)



Chapitre IV

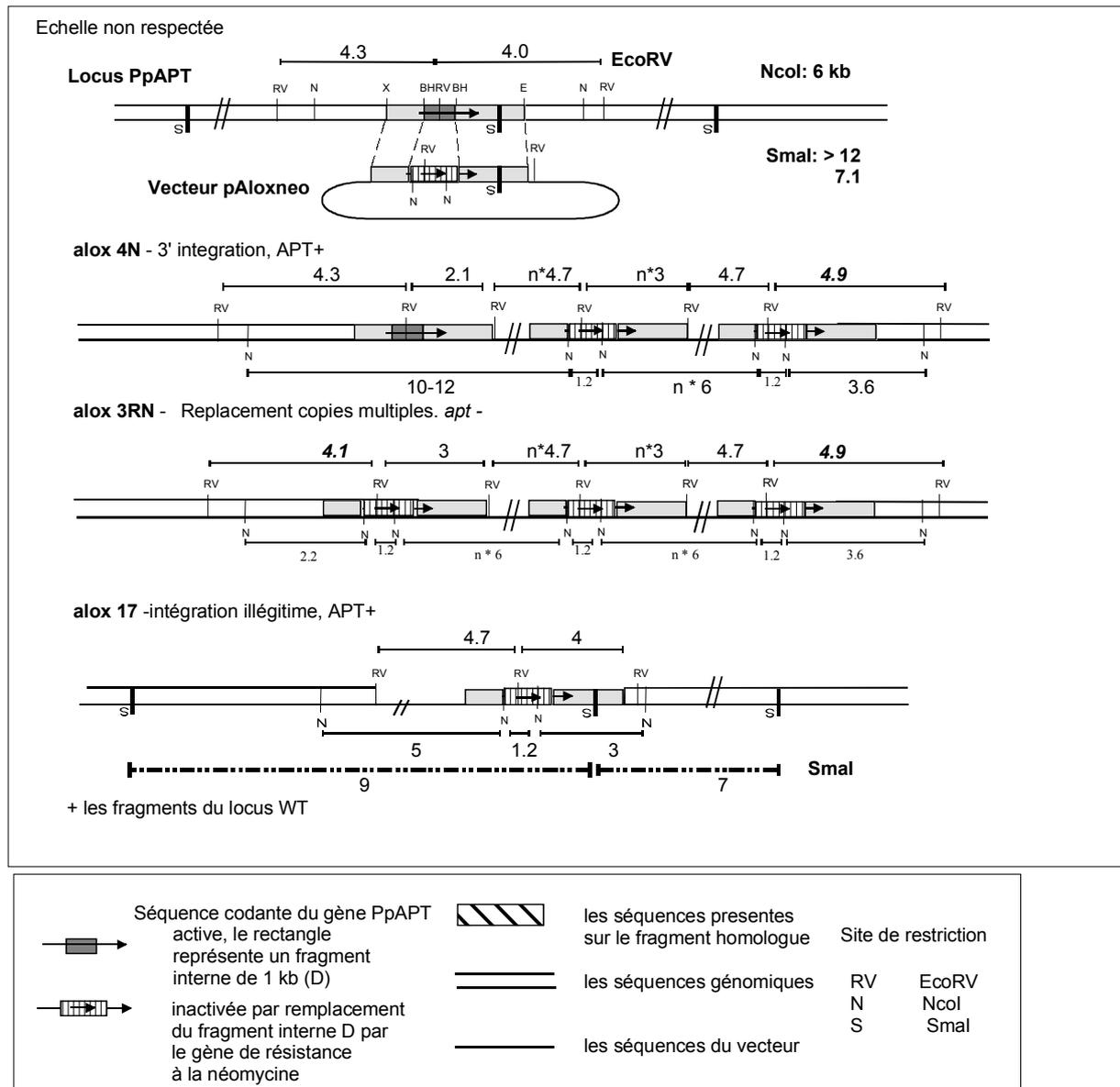


Figure 57. RFLP du locus d'intégration du vecteur pAloxneo dans les souches alox17, alox 4N et alox 3RN.

## Bibliographie

- Adair, G.M., Nairn, R.S., Wilson, J.H., Scheerer, J.B. and Brotherman, K.A. (1990) Targeted gene replacement at the endogenous *APRT* locus in CHO cells. *Somatic Cell & Molecular Genetics*, **16**, 437-41.
- Adair, G.M., Nairn, R.S., Wilson, J.H., Seidman, M.M., Brotherman, K.A., MacKinnon, C. and Scheerer, J.B. (1989) Targeted homologous recombination at the endogenous adenine phosphoribosyltransferase locus in Chinese hamster cells. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 4574-8.
- Adair, G.M., Scheerer, J.B., Brotherman, A., McConville, S., Wilson, J.H. and Nairn, R.S. (1998) Targeted recombination at the Chinese hamster APRT locus using insertion versus replacement vectors. *Somatic Cell & Molecular Genetics*, **24**, 91-105.
- Albert, H., Dale, E.C., Lee, E. and Ow, D.W. (1995) Site-Specific Integration of DNA into Wild-Type and Mutant lox Sites Placed in the Plant Genome. *The Plant Journal*, **7**, 649-659.
- Alfonzo, J.D., Sahota, A. and Taylor, M.W. (1997) Purification and Characterisation of Adenine Phosphoribosyltransferase from *Saccharomyces cerevisiae*. *Biochimica et Biophysica Acta*, **1341**, 173-182.
- Aoyama, T. and Chua, N.-H. (1997) A Glucocorticoid Mediated Transcriptional Induction System in Transgenic Plants. *The Plant Journal*, **11**, 605-612.
- Ashton, N.W., Champagne, C.E.M., Weiler, T. and Verkoczy, L.K. (2000) The Bryophyte *Physcomitrella patens* Replicates Extrachromosomal Transgenic Elements. *New Phytologist*, **146**, 391-402.
- Ashton, N.W. and Cove, D.J. (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants in the moss *Physcomitrella patens*. *Mol. Gen. Genet.*, **154**, 87-95.
- Ashton, N.W., Cove, D.J. and Featherstone, D.R. (1979a) The isolation and physiological analysis of mutants of the moss *Physcomitrella patens*, which over-produce gametophores. *Planta*, **144**, 437-442.
- Ashton, N.W., Grimsley, N.H. and Cove, D.J. (1979b) Analysis of gametophytic development in the moss *Physcomitrella patens*, using auxin and cytokinin resistant mutants. *Planta*, **144**, 427-435.
- Bartsch, S., Kang, L.E. and Symington, L.S. (2000) RAD51 is required for the repair of plasmid double-stranded DNA gaps from either plasmid or chromosomal templates. *Molecular & Cellular Biology*, **20**, 1194-205.
- Bayley, C.C., Morgan, M., Dale, E.C. and Ow, D.W. (1992) Exchange of Gene Activity in Transgenic Plants Catalysed by the Cre-lox Site-Specific Recombinase. *Plant Molecular Biology*, **18**, 353-361.

- Bazemore, L.R., Folta-Stogniew, E., Takahashi, M. and Radding, C.M. (1997) RecA tests homology at both pairing and strand exchange. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 11863-8.
- Beertling, W.M. (1995) Gene Targeting. *Gene Targeting*. CRC Press, pp. 2-44.
- Berg, P. (1991) Reverse Genetics: Its Origins and Prospects. *Biotechnology*, **9**, 342-344.
- Binns, A.N. (1994) Cytokinin Accumulation and Action: Biochemical, Genetic, and Molecular Approaches. *Annual Reviews*, 173-196.
- Böhner, S., Lenk, I., Rieping, M., Herold, M. and Gatz, C. (1999) Transcriptional Activator TGV Mediates Dexamethasone-inducible and Tetracycline-inactivatable Gene Expression. *The Plant Journal*, **19**, 87-95.
- Bouchez, D. and Höfte, H. (1998) Functional Genomics in Plants. *Plant Physiology*, **119**, 725-732.
- Britt, A.B. (1999) Molecular Genetics of DNA Repair in Higher Plants. , **4**, 20-25.
- Bronson, S.K. and Smithies, O. (1994) Altering mice by homologous recombination using embryonic stem cells. [Review] [56 refs]. *Journal of Biological Chemistry*, **269**, 27155-8.
- Bryant, J. and Leather, S. (1992) Removal of Selectable Marker Genes from Transgenic Plants: Needless Sophistication or Social Necessity. *Trends in Biotechnology*, **10**, 274-275.
- Bundock, P., den Dulk-Ras, A., Beijersbergen, A. and Hooykaas, P.J.J. (1995) Trans-Kingdom T-DNA Transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO Journal*, **14**, 3206-3214.
- Caddick, M.X., Greenland, A.J., Jepson, I., Krause, K.P., Qu, N., Riddell, K.V., Salter, M.G., Schuh, W., Sonnewald, U. and Tomsett, A.B. (1998) An Ethanol Inducible Gene Switch for Plants Used to Manipulate Carbon Metabolism. *Nature Biotechnology*, **16**, 177-180.
- Callis, J., Fromm, M. and Walbot, V. (1987) Introns Increase Gene Expression in Cultured Maize Cells. *Genes & Development*, **1**, 1183-1200.
- Castresana, C., Staneloni, R., Malik, V.S. and Cashmore, A.R. (1987) Molecular Characterisation of Two Clusters of Genes Encoding the Type I Cab Polypeptides of PSII in *Nicotiana plumbagnifolia*. *Plant Molecular Biology*, **10**, 117-126.
- Chiurazzi, M., Ray, A., Viret, J.F., Perera, R., Wang, X.H., Lloyd, A.M. and Signer, E.R. (1996) Enhancement of somatic intrachromosomal homologous recombination in *Arabidopsis* by the HO endonuclease. *Plant Cell*, **8**, 2057-66.

- Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D. and Jaisser, F. (1998) I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Molecular & Cellular Biology*, **18**, 1444-8.
- Conne, B., Stutz, A. and Vassali, J.-D. (2000) The 3' Untranslated Region of Messenger RNA: a Molecular "Hotspot" for Pathology? *Nature Medicine*, **6**, 637-641.
- Cove, D.J. (1992) Regulation of development in the moss, *Physcomitrella patens*. In Brody, S., Cove, D.J., Ottolenghi, S. and Russo, V.E.A. (eds.), *Developmental Biology. A molecular genetic approach*. Springer Verlag, Heidelberg, pp. 179 - 193.
- Cove, D.J., Kammerer, W., Knight, C.D., Leech, M.J., Martin, C.R. and Wang, T.L. (1991) Developmental genetic studies of the moss *Physcomitrella patens*. In Jenkins, G.I. and Schuch, W. (eds.), *Molecular biology of plant development*. Company of biologists, Cambridge, UK, pp. 31 - 43.
- Cove, D.J. and Knight, C.D. (1993) The moss *Physcomitrella patens*, a model system with potential for the study of plant reproduction. *The Plant Cell*, **5**, 1483-1488.
- Cove, D.J., Knight, C.D. and Lamparter, T. (1997) Mosses as Model Systems [Review]. *Trends in Plant Science*, **2**, 99-105.
- Cruz, A.K. (1995) Gene Targeting in Trypanosomatids. *Gene Targeting*. CRC-Press, Inc., pp. 65-81.
- Cura, T. (1999) Repairing the Genome's Spelling Mistakes. *Science*, **285**, 316-318.
- Dale, E.C. and Ow, D.W. (1990) Intra- and Intermolecular Site-Specific Recombination in Plant Cells Mediated by Bacteriophage P1 Recombinase. *Gene*, **91**, 79-85.
- Dale, E.C. and Ow, D.W. (1991) Gene transfer with subsequent removal of the selection gene from the host genome. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 10558-62.
- Datta, A.H.M.L.M.J.-R.S. (1997) Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9757-62.
- Davis, S.J. and Vierstra, R.D. (1998) Soluble Derivatives of Green Fluorescent Protein for Use in *Arabidopsis thaliana*. *Plant Molecular Biology*, **36**, 521-527.
- Deng, C. and Capecchi, M.R. (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Molecular & Cellular Biology*, **12**, 3365-71.
- Deng, C., Thomas, K.R. and Capecchi, M.R. (1993) Location of Crossovers during Gene Targeting with Insertion and Replacement Vectors. *Molecular and Cellular Biology*, **13**, 2134-2140.

- Ebinuma, H., Sugita, K., Matsnaga, E. and Yamakado, M. (1997) Selection of Marker-Free Transgenic Plants Using the Isopentyl Transferase Gene. *PNAS*, **94**, 2117-2121.
- Fagard, M. and Vaucheret, H. (2000) (Trans)Gene Silencing in Plants: How Many Mechanisms? *Annual Review in Plant Molecular Biology*, **51**, 167-194.
- Ferguson, D.O. and Hollman, W.K. (1996) Recombinational Repair of Gaps in DNA is Asymmetric in *Ustilago maydis* and Can Be Explained by a Migrating D-Loop Model. *PNAS*, **93**, 5419-5424.
- Fiering, S., Bender, M.A. and Groudine, M. (1999) Analysis of Mammalian Cis-regulatory DNA Elements by Homologous Recombination. *Analysis of Gene Expression*. Academic Press, pp. 42-66.
- Fuchs, R.L., Ream, J.E., Hammond, B.G., Naylor, M.W., Leimgruber, R.M. and Berberich, S.A. (1993) Safety Assessment of the Neomycine Phosphotransferase II (NPTII) Protein. *Biotechnology*, **11**, 1543-1547.
- Gaillard, C., Moffatt, B.A., Blacker, M. and Laloue, M. (1998) Male sterility associated with APRT deficiency in *Arabidopsis thaliana* results from a mutation in the gene *APT1*. *Molecular & General Genetics*, **257**, 348-53.
- Galli, A. and Schiestl, R.H. (1998) Effects of DNA Double-Strand and Single-Strand Breaks on Intrachromosomal Recombination Events in Cell-Cycle Arrested Yeast Cells. *Genetics*, **149**, 1243-1250.
- Gatz, C. and Lenk, I. (1998) Promoters that Respond to Chemical Inducers. *Trends in Plant Science*, **3**, 352-358.
- Girke, T., Schmidt, H., Zahringer, U., Reski, R. and Heinz, E. (1998) Identification of a novel delta 6-acyl-group desaturase by targeted gene disruption in *Physcomitrella patens*. *Plant Journal*, **15**, 39-48.
- Girod, P.A., Fu, H.Y., Zryd, J.P. and Vierstra, R.D. (1999) Multiubiquitin chain binding subunit MCB1 (RPN10) of the 26S proteasome is essential for developmental progression in *Physcomitrella patens*. *Plant Cell*, **11**, 1457-1471.
- Glasunov, A.V. and Glaser, V.M. (1999) The influence of mutation *rad57-1* on the fidelity of DNA double-strand gap repair in *Saccharomyces cerevisiae*. *Current Genetics*, **34**, 430-7.
- Gleave, A.P., Mitra, D.S., Mudge, S.R. and Morris, B.A.M. (1999) Selectable marker-free transgenic plants without sexual crossing: transient expression of cre recombinase and use of a conditional lethal dominant gene. *Plant Molecular Biology*, **40**, 223-235.
- Goldsbrough, A.P., Lastrella, C.N. and Yoder, J.I. (1993) Transposition Mediated Repositioning and Subsequent Elimination of Marker Genes from Transgenic Tomato. *Biotechnology*, **11**, 1286-1292.

- Gorbunova, V. and Levy, A.A. (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Research*, **25**, 4650-7.
- Gorbunova, V. and Levy, A.A. (1999) How plants make ends meet: DNA double-strand break repair [Review]. *Trends in Plant Science*, **4**, 263-269.
- Haber, J.E. (2000) Partners and Pathways - Repairing a Double-Strand Break [Review]. *TIG*, **16**, 259-264.
- Halfter, U., Morris, P.C. and Willmitzer, L. (1992) Gene targeting in *Arabidopsis thaliana*. *Molecular & General Genetics*, **231**, 186-93.
- Hastings, P.J., McGill, C., Shafer, B. and Strathern, N. (1993) Ends-In vs. Ends-Out Recombination in Yeast. *Genetics*, **135**, 973-980.
- Hasty, P., Crist, M., Grompe, M. and Bradley, A. (1994) Efficiency of insertion versus replacement vector targeting varies at different chromosomal loci. *Molecular & Cellular Biology*, **14**, 8385-90.
- Hasty, P., Rivera-Perez, J. and Bradley, A. (1991a) The Length of Homology Required for Gene Targeting in Embryonic Stem Cells. *Molecular and Cellular Biology*, **11**, 5586-5591.
- Hasty, P., Rivera-Perez, J. and Bradley, A. (1992) The Role and Fate of DNA Ends for Homologous Recombination in Embryonic Stem Cells. *Molecular and Cellular Biology*, **12**, 2464-2474.
- Hasty, P., Rivera-Perez, J. and Bradley, A. (1995) Gene conversion during vector insertion in embryonic stem cells. *Nucleic Acids Research*, **23**, 2058-64.
- Hasty, P., Rivera-Perez, J., Chang, C. and Bradley, A. (1991b) Target Frequency and Integration Pattern for Insertion and Replacement Vectors in Embryonic Stem Cells. *Molecular and Cellular Biology*, **11**, 4509-4517.
- Hoffmann, A.H., Codòn, A.C., Ivascu, C., Russo, E., Knight, C., Cove, D., Schaefer, D., Chakhparonian, M. and Zryd, J.-P. (1999) A Specific Member of the *Cab* Multigene Family Can Be Efficiently Targeted and Disrupted in the Moss *Physcomitrella patens*. *Mol. Gen. Genet.*, **261**, 92-99.
- Hohn, B. and Puchta, H. (1999) Gene therapy in plants. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 8321-8323.
- Hrouda, M. and Paszkowski, J. (1994) High fidelity extrachromosomal recombination and gene targeting in plants. *Molecular & General Genetics*, **243**, 106-11.
- Isselbacher, K.J., Braunwald, E., Wilson, J.D., Martin, J.B., Fauci, A.S. and Kasper, D.L. (eds.) (1994) *Harrison's Principles of Internal Medicine*. McGraw-Hill, Inc.

- Itai, R., Suzuki, K., Yamaguchi, H., Nakanishi, H., Nishizawa, N.-K., Yoshimura, E. and Mori, S. (2000) Induced Activity of Adenine Phosphoribosyltransferase in Iron-Deficient Barley Roots: a Possible Role for Phytosiderophore Production. *Journal of Experimental Botany*, **51**, 1179-1188.
- Jasin, M. (1996) Genetic manipulation of genomes with rare-cutting endonucleases. [Review] [34 refs]. *Trends in Genetics*, **12**, 224-8.
- Jefferson, R.A. (1987) Assaying Chimeric Genes in Plants: the GUS Gene Fusion System. *Plant Molecular Biology Reports*, **5**, 387-405.
- Jefferson, R.A., Kavanagh, T.A. and Bevan, M. (1987) GUS Fusions, beta-glucuronidase as a Sensitive and Versatile Gene Fusion Marker in Higher Plants. *The EMBO Journal*, **6**, 3901-3907.
- Jeon, J.S., Lee, S., Jung, K.H., Jun, S.H., Jeong, D.H., Lee, J., Kim, C., Jang, S., Yang, K., Nam, J., An, K., Han, M.J., Sung, R.J., Choi, H.S., Yu, J.H., Choi, J.H., Cho, S.Y., Cha, S.S., Kim, S.I. and An, G. (2000) T-DNA insertional mutagenesis for functional genomics in rice. *Plant Journal*, **22**, 561-570.
- Kammerer, W. and Cove, D.J. (1996) Genetic Analysis Of the Effects Of Re-Transformation Of Transgenic Lines Of the Moss *Physcomitrella patens*. *Molecular & General Genetics*, **250**, 380-382.
- Kempin, S.A., Liljegren, S.J., Block, L.M., Rounsley, S.D., Yanofsky, M.F. and Lam, E. (1997) Targeted disruption in *Arabidopsis* [letter]. *Nature*, **389**, 802-3.
- Knight, C.D. (1994) Studying plant development in mosses: the transgenic route. , **17**, 669 - 674.
- Knight, C.D., Sehgal, A., Atwal, K., Wallace, J.C., Cove, D.J., Coates, D., Quatrano, R.S., Bahadur, S., Stockley, P.G. and Cuming, A.C. (1995) Molecular Responses to Abscisic Acid and Stress are Conserved between Moss and Cereals. *The Plant Cell*, **7**, 499-506.
- Koren, P., Svetec, I.K., Mitrikeski, P.T. and Zgaga, Z. (2000) Influence of Homology Size and Polymorphism on Plasmid Integration in the Yeast *CYCI* DNA Region. *Curr. Genet.*, **37**, 292-297.
- Kozak, M. (1991) Structural Features in Eukaryotic mRNAs that Modulate the Initiation of Translation [minireview]. *The Journal of Biological Chemistry*, **266**, 19857-19870.
- Lankenau, D.H. and Gloor, G.B. (1998) In vivo gap repair in *Drosophila*: a one-way street with many destinations. [Review] [78 refs]. *Bioessays*, **20**, 317-27.
- Lee, K.Y., Lund, P., Lowe, K. and Dunsmuir, P. (1990) Homologous Recombination in Plant Cells after *Agrobacterium* Mediated Transformation. *The Plant Cell*, **2**, 415-425.

- Lee, S.E., Paques, F., Sylvan, J. and Haber, J.E. (1999) Role of yeast *SIR* genes and mating type in directing DNA double-strand breaks to homologous and non-homologous repair paths. *Current Biology*, **9**, 767-70.
- Lehninger, A.L., Nelson, D.L. and Cox, M.M. (1993) *Principles of Biochemistry*. Worth Publishers, New York.
- Leung, W., Malkova, A. and Haber, J.E. (1997) Gene targeting by linear duplex DNA frequently occurs by assimilation of a single strand that is subject to preferential mismatch correction. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 6851-6.
- Lewin, B. (1999) *Genes*. Oxford University Press, New York.
- Lichtenstein, C.P., Paszkowski, J. and Hohn, B. (1995) Intrachromosomal Recombination between Genomic Repeats. In Paszkowski, J. (ed.) *Homologous Recombination and Gene Silencing in Plants*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 95-122.
- Long, Z., Wang, S.-Y. and Nelson, N. (1989) Cloning and Nucleotide Sequence Analysis of genes Coding for the Major Chlorophyll-Binding Protein of the Moss *Physcomitrella patens* and the Halotolerant Alga *Dunaliella salina*. *Gene*, **76**, 299-312.
- Lyznik, L.A., Hirayama, L., Rao, K.V., Abad, A. and Hodges, T.K. (1995) Heat-inducible Expression of FLP Gene in Maize Cells. *Plant Journal*, **8**, 177-186.
- Manivasakam, P., Weber, S.C., McElver, J. and Schiestl, R.H. (1995) Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, **23**, 2799-800.
- Marcotte, W.R.J., Bayley, C.C. and Quatrano, R.S. (1988) Regulation of a Wheat Promoter by Abscissic Acid in Rice Protoplasts. *Nature*, **335**, 454-457.
- Matzke, M., Matzke, A.J.M. and Scheid, O.M. (1995) Inactivation of Repeated Genes - DNA-DNA Interaction? In Paszkowski, J. (ed.) *Homologous Recombination and Gene Silencing in Plants*. Kluwer Academic Press, Dordrecht, pp. 271-308.
- McElroy, D., Blowers, A.D., Jenes, B. and Wu, R. (1991) Construction of Expression Vectors Based on the Rice Actin 1 5' Region for Use in Monocot Transformation. *Molecular and General Genetics*, **231**, 150-160.
- McElroy, D., Zhang, W., Cao, J. and Wu, R. (1990) Isolation of an Efficient Actin Promoter for Use in Rice Transformation. *The Plant Cell*, **2**, 163-171.
- McKenzie, M.J., Mett, V., Reynolds, P.H.S. and Jameson, P.E. (1998) Controlled Cytokinin Production in Transgenic Tobacco Using a Copper-Inducible Promoter. *Plant Physiology*, **116**, 969-977.
- Medberry, S.L., Dale, E., Qin, M. and Ow, D.W. (1995) Intra-chromosomal rearrangements generated by Cre-lox site-specific recombination. *Nucleic Acids Research*, **23**, 485-90.

- Mengiste, T. and Paszkowski, J. (1999) Prospects for the precise engineering of plant genomes by homologous recombination [Review]. *Biological Chemistry*, **380**, 749-758.
- Mengiste, T., Revenkova, E., Bechtold, N. and Paszkowski, J. (1999) An SMC-like protein is required for efficient homologous recombination in *Arabidopsis*. *EMBO Journal*, **18**, 4505-12.
- Mett, V.L., Lochhead, L.P. and Reynolds, P.H.S. (1993) Copper-Controllable Gene Expression System for Whole Plants. *PNAS*, **90**, 4567-4571.
- Miao, Z.H. and Lam, E. (1995) Targeted disruption of the TGA3 locus in *Arabidopsis thaliana*. *Plant Journal*, **7**, 359-65.
- Moffatt, B., McWinnie, E.A., Agarwal, S.A. and Schaff, D.A. (1994) The Adenine Phosphoribosyltransferase - Encoding Gene of *Arabidopsis thaliana*. *Gene*, **143**, 211-216.
- Moffatt, B., Pethe, C. and Laloue, M. (1991) Metabolism of Benzyladenine is Impaired in a Mutant of *Arabidopsis thaliana* Lacking Adenine Phosphoribosyltransferase Activity. *Plant Physiology*, **95**, 900-908.
- Moore, I., Galweiler, L., Grosskopf, D., Schell, J. and Palme, K. (1998) A Transcription Activation System For Regulated Gene Expression in Transgenic Plants. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 376-381.
- Moreadith, R.W. and Radford, N.B. (1997) Gene targeting in embryonic stem cells: the new physiology and metabolism. [Review] [81 refs]. *Journal of Molecular Medicine*, **75**, 208-16.
- Morton, R. and Hooykaas, P.J.J. (1995) Gene replacement. *Molecular Breeding*, **1**, 123-132.
- Muller, U. (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. [Review] [147 refs]. *Mechanisms of Development*, **82**, 3-21.
- Negritto, M.T., Wu, X., Kuo, T., Chu, S. and Bailis, A.M. (1997) Influence of DNA sequence identity on efficiency of targeted gene replacement. *Molecular & Cellular Biology*, **17**, 278-86.
- Nickoloff, J.A., Singer, J.D., Hoekstra, M.F. and Heffron, F. (1989) Double-Strand Breaks Stimulate Alternative Mechanisms of Recombination Repair. *J. Mol. Biol.*, **207**, 527-541.
- Nishiyama, T., Hiwatashi, Y., Sakakibara, K., Kato, M. and Hasebe, M. (2000) Tagged Mutagenesis and Gene-Trap in the Moss, *Physcomitrella patens* by Shuttle Mutagenesis. *DNA Research*, **7**, 9-17.
- Odell, J., Caimi, P., Sauer, B. and Russell, S. (1990) Site-Directed Recombination in the Genome of Transgenic Tobacco. *Molecular and General Genetics*, **223**, 369-378.

- Odell, J.T. and Russell, S.H. (1994) Use of site specific recombination system in plants. In Paszkowski, J. (ed.) *Homologous Recombination and Gene Silencing in Plants*. Kluwer Academic Publishers Group, Dordrecht, The Netherlands, pp. 219-271.
- Offringa, R. (1992) Gene Targeting in Plants Using the *Agrobacterium* Vector System. *Clusius Laboratory, Institute of Molecular Plant Sciences*. University of Leiden, Leiden, NL, p. 151.
- Offringa, R., Franke van Dijk, M.E.I., Marcel, J.A., de Groot, J.A., van den Elzen, P.J.M. and Hooykaas, P.J.J. (1993) Nonreciprocal Homologous Recombination between *Agrobacterium* Transferred DNA and a Plant Chromosomal Locus. *PNAS*, **90**, 7346-7350.
- Offringa, R. and Hooykaas, P. (1995) Gene Targeting in Plants. *Gene targeting*. CRC Press, Inc.
- Ohl, S., Offringa, R., van den Elzen, P.J.M. and Hooykaas, P.J.J. (1995) Gene Replacement in Plants. In Paszkowski, J. (ed.) *Homologous Recombination and Gene Silencing in Plants*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 191-218.
- Orr-Weaver, T.L. and Szostack, J.W. (1983) Yeast Recombination: the Association between Double-Strand Gap Repair and Crossing-Over. *PNAS*, **80**, 4417-4421.
- Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J. (1981) Yeast Transformation: A Model System for the Study of Recombination. *PNAS*, **78**, 6354-6358.
- Ow, D.W. and Medberry, S.L. (1995) Genome Manipulation Through Site-Specific Recombination. *Critical Reviews in Plant Sciences*, **14**, 239-261.
- Paques, F. and Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. [Review] [575 refs]. *Microbiology & Molecular Biology Review (Washington, DC)*, **63**, 349-404.
- Parinov, S. and Sundaresan, V. (2000) Functional Genomics in *Arabidopsis*: Large Scale Insertional Mutagenesis Complements the Genome Sequencing Project. *Current Opinion in Biotechnology*, **11**, 157-161.
- Paszkowski, J., Baur, M., Bogucki, A. and Potrykus, I. (1988) Gene Targeting in Plants. *The EMBO Journal*, **3**, 2717-2722.
- Potrykus, I. and Spangenberg, G. (eds.) (1995) *Gene Transfer to Plants*. Springer Verlag, Berlin Heidelberg.
- Puchta, H. (1998) Repair of Genomic Double-Strand Breaks in Somatic Plant Cells by One-Sided Invasion of Homologous Sequences. *The Plant Journal*, **13**, 331-339.
- Puchta, H. (1999) Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics*, **152**, 1173-1181.

- Puchta, H., Dujon, B. and Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 5055-60.
- Puchta, H. and Hohn, B. (1996) From Centimorgans to Base Pairs - Homologous Recombination in Plants [Review]. *Trends in Plant Science*, **1**, 340-348.
- Puchta, H. and Meyer, P. (1995) Substrate Specificity of Plant Recombinases Determined in Extrachromosomal Recombination Systems. In Paszkowski, J. (ed.) *Homologous Recombination and Gene Silencing in Plants*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 123-156.
- Rao, B.J., Chiu, S.K., Bazemore, L.R., Reddy, G. and Radding, C.M. (1995) How specific is the first recognition step of homologous recombination? [Review] [39 refs]. *Trends in Biochemical Sciences*, **20**, 109-13.
- Reiss, B., Klemm, M., Kosak, H. and Schell, J. (1996) RecA Protein Stimulates Homologous Recombination in Plants. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 3094-3098.
- Reiss, B., Schubert, I., Kopchen, K., Wendeler, E., Schell, J. and Puchta, H. (2000) RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by *Agrobacterium*. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 3358-63.
- Reski, R. (1998a) Development, Genetics and Molecular Biology of Mosses [Review]. *Botanica Acta*, **111**, 1-15.
- Reski, R. (1998b) *Physcomitrella* and *Arabidopsis*: the David and Goliath of the Reverse Genetics. *Trends in Plant Science*, **3**, 209-210.
- Reski, R. (1999) Molecular genetics of *Physcomitrella* [Review]. *Planta*, **208**, 301-309.
- Reski, R., Faust, M., Wang, X.H., Wehe, M. and Abel, W.O. (1994) Genome analysis of the moss *Physcomitrella patens* (Hedw.) B.S.G. *Molecular & General Genetics*, **244**, 352-9.
- Reutter, K., Atzorn, R., Hädeler, B., Scmuelling, T. and Reski, R. (1998) Expression of the Bacterial *ipt* Gene in *Physcomitrella* Rescues Mutations in budding and in plastid division. *Planta*, **206**, 196-203.
- Riele, H., Maandag, E.R. and Berns, A. (1992) Highly Efficient Gene Targeting in Embryonic Stem Cells Through Homologous Recombination with Isogenic DNA Constructs. *PNAS*, **89**, 5128-5132.
- Risseuw, E., Franke-van Dijk, M.E. and Hooykaas, P.J. (1996) Integration of an insertion-type transferred DNA vector from *Agrobacterium tumefaciens* into the *Saccharomyces cerevisiae* genome by gap repair. *Molecular & Cellular Biology*, **16**, 5924-32.

- Risseeuw, E., Offringa, R., Franke-van Dijk, M.E. and Hooykaas, P.J. (1995) Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus. *Plant Journal*, **7**, 109-19.
- Rogers, S.O. and Bendlich, A.J. (1988) Extraction of DNA from Plant Tissues. In Gelvin, S.B., Schilperoort, R.A. and Verma, D.P.S. (eds.), *Plant Molecular Biology Manual*. Kluwer Academic Press, Dordrecht, The Netherlands.
- Rossi, F.M.V. and Blau, H. (1998) Recent Advances in Inducible Gene Expression Systems. *Current Opinion in Biotechnology*, **9**, 451-456.
- Rothstein, R. (1991) Targeting, Disruption, Replacement, and Allele Rescue: Integrative DNA Transformation in Yeast. *Methodes in Enzymology*, **194**, 281-301.
- Russell, S.H., Hoopes, J.L. and Odell, J.T. (1992) Directed Excision of a Transgene from the Plant Genome. *Molecular and General Genetics*, **234**, 49-59.
- Salomon, S. and Puchta, H. (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO Journal*, **17**, 6086-95.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*.
- Sauer, B. (1994) Recycling Selectable Markers in Yeast. *BioTechniques*, **16**, 1086-1088.
- Sauer, B. (1998) Inducible gene targeting in mice using the Cre/lox system. [Review] [66 refs]. *Methods*, **14**, 381-92.
- Schaefer, D.G. (1994) Molecular Genetic Approaches to the Biology of the Moss *Physcomitrella patens*. *LPC, Biologie*. University of Lausanne, Switzerland, Lausanne.
- Schaefer, D.G. (2001) Gene Targeting in *Physcomitrella patens*. *Current Opinion in Plant Biology*, **4**, 143-150.
- Schaefer, D.G., Bisztray, G. and Zryd, J.-P. (1994) Genetic transformation of the moss *Physcomitrella patens*. In Bajaj, Y.P.S. (ed.) *Plant protoplasts and genetic engineering V*. Springer Verlag, Berlin, Heidelberg, New York, Vol. 29, pp. 349-364.
- Schaefer, D.G., Vlach, S.J., Chakhparonian, M., Zryd, J.-P., Leech, M., Houba-Hérin, N., Schwartzberg, K., Pethe, C. and Laloue, M. (en préparation) Replacement of the Adenine Phosphoribosyltransferase Gene in *Physcomitrella patens*. .
- Schaefer, D.G. and Zryd, J.P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant Journal*, **11**, 1195-206.
- Schaefer, D.G., Zryd, J.-P., Knight, C.D. and Cove, D.J. (1991) Stable Transformation of the Moss *Physcomitrella patens*. *Molecular and General Genetics*, **226**, 418-424.

- Schnorr, K.M., Gaillard, C., Biget, E., Nygaard, P. and Laloue, M. (1996) A second form of adenine phosphoribosyltransferase in *Arabidopsis thaliana* with relative specificity towards cytokinins. *Plant Journal*, **9**, 891-8.
- Schulz, P., Reski, R., Maldiney, R., Laloue, M. and Schwartzenberg, K.v. (2000) Kinetics of Cytokinin Production and Bud Formation in *Physcomitrella*: Analysis of Wild Type, a Developmental Mutant and Two of Its *ipt* Transgenics. *Journal of Plant Physiology*, **156**, 768-774.
- Shah-Mahoney, N., Hampton, T., Vidaver, R. and Ratner, D. (1997) Blocking the ends of transforming DNA enhances gene targeting in *Dictyostelium*. *Gene*, **203**, 33-41.
- Shalev, G., Sitrit, Y., Avivi-Ragolski, N., Lichtenstein, C. and Levy, A.A. (1999) Stimulation of Homologous Recombination in Plants by Expression of the Bacterial Resolvase RuvC. *PNAS*, **95**, 7398-7402.
- Smih, F., Rouet, P., Romanienko, P.J. and Jasin, M. (1995) Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Research*, **23**, 5012-9.
- Somerville, C. and Somerville, S. (1999) Plant Functional Genomics. *Science*, **285**, 380-383.
- Srivastava, V., Anderson, O.D. and Ow, D.W. (1999) Single-copy transgenic wheat generated through the resolution of complex integration patterns. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 11117-11121.
- Stambuk, S. and Radman, M. (1998) Mechanism and control of interspecies recombination in *Escherichia coli*. I. Mismatch repair, methylation, recombination and replication functions. *Genetics*, **150**, 533-42.
- Stasiak, A. (1996) Getting Down to the Core of the Homologous Recombination. *Science*, **272**, 828-829.
- Strepp, R., Scholz, S., Kruse, S., Speth, V. and Reski, R. (1998) Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 4368-73.
- Szostak, J.W., Orr-Weaver, T.L., Rithstein, R.J. and Stahl, F.W. (1983) The Double-Strand-Break Repair Model for Recombination. *Cell*, **33**, 25-35.
- Takano, M., Egawa, H., Ikeda, J.E. and Wakasa, K. (1997) The structures of integration sites in transgenic rice. *Plant Journal*, **11**, 353-61.
- Thomas, K.R., Deng, C. and Capecchi, M.R. (1992) High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Molecular & Cellular Biology*, **12**, 2919-23.

- Thykjaer, T., Finnemann, J., Schauser, L., Christensen, L., Poulsen, C. and Stougaard, J. (1997) Gene targeting approaches using positive-negative selection and large flanking regions. *Plant Molecular Biology*, **35**, 523-30.
- Vergunst, A.C. and Hooykaas, P.J.J. (1999) Recombination in the plant genome and its application in biotechnology [Review]. *Critical Reviews in Plant Sciences*, **18**, 1-31.
- Vergunst, A.C., Jansen, L.E.T., Fransz, P.F., de Jong, J.H. and Hooykaas, P.J.J. (2000) Cre/lox-mediated recombination in *Arabidopsis*: evidence for transmission of a translocation and a deletion event. *Chromosoma*, **109**, 287-297.
- Villemure, J.-F., Belmaaza, A. and Chartrand, P. (1997) The Processing of DNA Ends at Double-Strand Breaks During Homologous Recombination: Different Roles for the Two Ends. *Molecular and General Genetics*, **256**, 533-538.
- von Schwartzenberg, K., Kruse, S., Reski, R., Moffatt, B. and Laloue, M. (1998) Cloning and characterization of an adenosine kinase from *Physcomitrella* involved in cytokinin metabolism. *Plant Journal*, **13**, 249-57.
- Vulic, M., Dionisio, F., Taddei, F. and Radman, M. (1997) Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 9763-7.
- Wang, T.L. and Cove, D.J. (1989) Mosses - lower plants with high potential. *Plants today*, **2**, 44-50.
- Woods, R.A., Roberts, D.G., Stein, D.S. and Filpula, D. (1984) Adenine Phosphoribosyltransferase Mutants in *Saccharomyces cerevisiae*. *Journal of General Microbiology*, **130**, 2629-2637.
- Worth, M.J., Clark, S., Radman, M. and Modrich, P. (1994) Mismatch Repair Proteins MutS and MutL inhibit RecA-catalyzed Strand Transfer Between Diverged DNAs. *PNAS*, **91**, 3238-3241.
- Xiao, Y.L. and Peterson, T. (2000) Intrachromosomal homologous recombination in *Arabidopsis* induced by a maize transposon. *Molecular & General Genetics*, **263**, 22-29.
- Yoder, J.I. and Goldsbrough, A.P. (1994) Transformation Systems for Creating Marker-Free Transgenic Plants. *Biotechnology*, **12**, 263-267.
- Zeidler, M., Gatz, C., Hartmann, E. and Hughes, J. (1996) Tetracycling-Regulated Reporter Gene Expression in the Moss *Physcomitrella patens*. *Plant Molecular Biology*, **30**, 199-205.
- Zhao, J., Hyman, L. and Moore, C. (1999) Formation of mRNA 3' Ends in Eukaryotes: Mechanisms, Regulation and Interrelationships with Other Steps in mRNA Synthesis. *Microbiology and Molecular Biology Reviews*, **63**, 405-446.

## Curriculum vitae

Mikhail CHAKHPARONIAN  
IE-PC, BB, UNILausanne  
CH-1015 Lausanne-Dorigny  
Suisse  
Mikhail.Chakhparonian@ie-pc.unil.ch  
tél. + 41 22 346 1369 ou + 41 21 692 4252

né le 15 décembre 1972 à Bakou (ex-URSS)

### Forces et aptitudes développées

Domaines de préférences et de compétences en Biologie

- biologie et génétique moléculaires, mutagenèse ciblée, recombinaison, biologie végétale, mousse *Physcomitrella patens*, culture de cellules et de tissus;
- pollution atmosphérique, bio-contrôle de pollution, biotechnologie.

Goût pour l'informatique et le multimédia

- projets d'enseignement électronique à distance ;
- création et gestion de site internet;
- gestion d'un parc informatique en matériel et logiciel.

Capacités de traduction / interprétation, bonne connaissance de plusieurs langues

- bénévolat, petits emplois en tant que traducteur ou interprète;

Faculté d'autoformation;

Travail indépendant et en équipe;

Optimisme

### Formation

Travail de doctorat «Développement d'outils de la mutagenèse ciblée par recombinaison homologue chez *Physcomitrella patens*» au Laboratoire de Phytogénétique Cellulaire de l'Université de Lausanne.

Lausanne – Suisse. Novembre 1995 - Mai 2001

Diplôme de biologiste de l'Université de Lausanne: « Développement d'un système biologique de surveillance de la pollution atmosphérique : études sur des populations d'*Arabidopsis thaliana* sous conditions contrôlées». En collaboration avec le Service de l'Environnement de l'Hygiène et du Logement de Lausanne.

Lausanne – Suisse. Août 1994 - octobre 1995

Diplôme d'ingénieur en biotechnologie de l'Université Mendéléev de Moscou (Université de Génie Chimique de Russie).

Moscou – Russie. 1989 - 1995

## Publications et participation à des congrès scientifiques

A.Hoffmann et al., "A Specific Member of the Cab Multigene Family Can Be Efficiently Targeted and Disrupted in the Moss *Physcomitrella patens*", *Mol Gen Genet* 1999, 261: 92-99;

M. Chakhparonian, D.Schaefer et J.-P. Zryd "New tools for genetic engineering of the moss *Physcomitrella patens*", dans le recueil de résumés de la conférence "Moss 2000" 30 June- 02 July 2000, Villars-s/O - Suisse

*En préparation:*

D. Schaefer et al. "Replacement of the Adenine Phosphoribosyltransferase Gene in *Physcomitrella patens*"

M. Chakhparonian et al. "Activity of Different Promoters in the Moss *Physcomitrella patens*"

M. Chakhparonian et al. "Removal of Selectable Marker Genes from the Moss *Physcomitrella patens* Using Cre/lox Recombination System"

## Expérience professionnelle

***Institut d'Ecologie – Phytogénétique Cellulaire, Université de Lausanne - Suisse.***

***Mai 1997 - Août 2000***

Fonction : assistant diplômé

Recherche sur des méthodes de génétique inverse chez la mousse *Physcomitrella patens*.

Enseignement aux travaux pratiques de génétique et de physiologie végétale ;

Encadrement des étudiants de IIe cycle aux travaux de certificat ;

Gestion de parc informatique (matériel et logiciel) ;

Gestion du site Internet du laboratoire <http://www.unil.ch/lpc>;

Réalisations :

Plusieurs publications et présentations scientifiques;

Amélioration et systématisation de la gestion du parc informatique du laboratoire ;

Elargissement du site Internet et ajout de matériel interactif.

***Projet PIAF (Projet d'Informatisation pour une Auto-évaluation Flexible) au sein du projet européen ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe). Université de Lausanne – Suisse.***

***Juillet 1998 - Novembre 1999***

Fonction : chercheur diplômé

Mise en œuvre d'un cours électronique à distance :

- Création de pages Internet interactives ;
- Création d'un logiciel d'enseignement interactif.

Réalisations :

Création et programmation d'un cours interactif multimédia en Biologie végétale ;

Publication des cours sur Internet [http://www.unil.ch/lpc/welcome\\_piaf.htm](http://www.unil.ch/lpc/welcome_piaf.htm);

<http://www.unil.ch/lpc/docs/tpgen/prep/enter.htm>

## Connaissances informatiques

Logiciels de bureautique, scientifiques et graphiques :

- Word, Excel, PowerPoint, Micrografix Designer, STATISTICA etc. ;

Logiciel de programmation des cours électroniques :

- ToolBook (Asymetrix) ;

Gestion d'un site Internet d'un laboratoire (HTML 4.0, Javascript, connaissances de Java 2.0).

## Langues parlées et écrites

*Français* : excellent (travail de doctorat, études supérieures diplômées)

*Anglais* : excellent (pratique d'enseignement, pratique courante pendant 5 ans au sein d'un laboratoire en partie anglophone, publication d'articles scientifiques)

*Allemand* : bon (cours de préparation à l'examen de Goethe Institut, large cercle d'amis germanophones)

*Russe* : langue maternelle.

## Personnes de contact :

Dr Didier Schaefer (superviseur direct du travail de thèse)

IE-PC, BB, UNIL Lausanne 1015 Suisse

+ 41 21 692 4257

Didier.Schaefer@ie-pc.unil.ch

Dr Ursula Hinz (superviseur des travaux pratiques)

+ 41 21 692 4252

Ursula.Hinz@ie-pc.unil.ch

Prof. Maia Wentland-Forte (responsable du projet PIAF)

+41 21 692 2061; +41 21 692 3429

Rectorat BRA Lausanne 1015 Suisse

mwf@unil.ch

Prof. Natalia E. Krutchinina (responsable du travail de diplôme de l'Université Mendéléev)

(*francophone*)

+7 095 978 7858 (Faculté)

Faculté de Génie Ecologique

Miousskaia sq., 9

Moscou, Russie

krutch@muctr.edu.ru

Prof. J.-P. Zryd (directeur du laboratoire)

IE-PC, BB, UNIL Lausanne 1015 Suisse

+ 41 21 692 4251

Jean-Pierre.Zryd@ie-pc.unil.ch