

La transformation génétique, un outil pour l'analyse fonctionnelle du génome du riz

Delphine Mieulet¹
 Donaldo Meynard¹
 Christelle Siré¹
 Julie Petit¹
 Jean-Christophe Breitler¹
 Christophe Périn¹
 Anne Diévert¹
 Fanchon Divol-Malgoire¹
 Brigitte Courtois¹
 Jean-Luc Verdeil¹
 Pascal Gantet²
 Emmanuel Guiderdoni¹

¹ Cirad
 UMR AGAP
 TAA108/03
 Avenue Agropolis
 34398 Montpellier cedex 5
 France
 <delphine.mieulet@cirad.fr>
 <donaldo.meynard@cirad.fr>
 <sire.christelle@gmail.com>
 <julie.petit@cirad.fr>
 <jcbreitler@gmail.com>
 <christophe.perin@cirad.fr>
 <anne.dievert@cirad.fr>
 <fanchon.divol_malgoire@cirad.fr>
 <brigitte.courtois@cirad.fr>
 <jean-luc.verdeil@cirad.fr>
 <emmanuel.guiderdoni@cirad.fr>

² Université Montpellier II
 International Joint Laboratory "Rice
 Functional Genomics and Plant
 Biotechnology" (IRD, UM2, AGI, VAAS,
 USTH)
 Agricultural Genetics Institute
 Nat. Key Lab for Plant Cell Biotechnology
 Pham Van Dong road
 Co Nhue
 Tuliem
 Hanoi
 Vietnam
 <pascal.gantet@univ-montp2.fr>

Résumé

Le riz (*Oryza sativa* L.) a le double statut de céréale la plus consommée par l'homme et d'espèce modèle pour les monocotylédones. Élucider la fonction des gènes sous-tendant les principaux caractères d'intérêt agronomique chez cette plante contribuerait donc à accélérer non seulement l'amélioration variétale du riz mais aussi celle des autres céréales. La transformation génétique est un outil permettant d'accélérer la découverte de la fonction des gènes et l'innovation variétale. La première plante transgénique de riz a été obtenue dès 1988, et, depuis, les technologies ont régulièrement évolué vers plus d'efficacité et de précision. Nous présentons ici les grandes étapes de l'évolution de ces techniques, ainsi que leurs applications pour l'analyse fonctionnelle des gènes. En particulier, les perspectives ouvertes par de nouveaux sauts technologiques en cours, qui permettent à présent une modification ciblée des génomes sans introduction d'acide désoxyribonucléique (ADN) superflu, seront illustrées et discutées.

Mots clés : amélioration génétique ; expression des gènes ; fonction physiologique ; génomique ; riz ; transgénèse.

Thèmes : amélioration génétique ; méthodes et outils ; productions végétales.

Abstract

Gene transfer: A tool for the functional analysis of the rice genome

Rice (*Oryza sativa* L.) has the dual status of being a model grass species and the main cereal for human consumption. Elucidating the function of genes underlying agronomic traits in the first model crop will contribute to hastening plant breeding not only in rice but also in other cereals. Gene transfer is a tool that has the potential to accelerate both discovery and innovation. The first transgenic rice plant was produced in 1988, and the technologies have since then improved towards more efficiency and precision. We here present the main milestones in the evolution of the technologies and their applications for the functional analysis of rice genes. In particular, we show the prospects opened by on-going technological breakthroughs that include the targeted modification of genomes without integration of superfluous DNA.

Key words: breeding; gene expression; gene transfer; genomics; physiological functions; rice.

Subjects: genetic improvement; tools and methods; vegetal productions.

En ce début de siècle, le riz se trouve à la croisée d'enjeux socio-économiques et cognitifs considérables. D'une part, on estime que la production actuelle de 600 millions de tonnes sur 150 millions

d'hectares doit augmenter de 50 % d'ici 2030 pour satisfaire la demande croissante tout en s'ajustant à un contexte de terres arables non extensibles, de réduction des intrants et d'instabilité climatique accrue

Pour citer cet article : Mieulet D, Meynard D, Siré C, Petit J, Breitler JC, Périn C, Diévert A, Divol-Malgoire F, Courtois B, Verdeil JL, Gantet P, Guiderdoni E, 2013. La transformation génétique, un outil pour l'analyse fonctionnelle du génome du riz. *Cah Agric* 22 : 484-93. doi : 10.1684/agr.2013.0650

Tirés à part : E. Guiderdoni

doi: 10.1684/agr.2013.0650

(GRiSP, 2010) ; on attend de l'amélioration variétale du riz qu'elle permette, aux côtés de l'amélioration des itinéraires techniques et de la lutte intégrée contre les pathogènes et ravageurs, de nouveaux gains de productivité (résistance accrue aux maladies et insectes, tolérance aux contraintes abiotiques de l'environnement, amélioration de la qualité culinaire et gustative du grain et augmentation du potentiel de rendement). D'autre part, cette espèce est la plante modèle de la recherche académique pour l'élucidation de la fonction de gènes d'intérêt agronomique des céréales d'ici 2020 (Zhang *et al.*, 2008). Le riz est la première plante cultivée dont le génome a été entièrement séquencé, permettant de prédire la présence d'environ 38 000 gènes (zones chromosomiques considérées comme dotées d'une fonction sur la base de leur séquence) dont il faut à présent déterminer la fonction. Outil introduit en 1988 chez le riz, la transformation génétique peut permettre de contribuer à relever ces deux défis en accélérant l'acquisition de connaissances sur le fonctionnement des gènes et du génome, et en permettant d'atteindre des objectifs de sélection difficiles à réaliser par les méthodes conventionnelles.

Historique et état des technologies

La transformation génétique chez le riz comme chez les autres espèces de plantes repose d'abord sur une méthode efficace d'introduction d'ADN exogène dans le génome du noyau ou des organites (par exemple, les chloroplastes) d'une cellule capable de se diviser jusqu'à former une plante entière. L'intégration stable de l'ADN dans le génome nucléaire étant un événement peu fréquent qui se réalise à la faveur d'une cassure de la molécule d'ADN receveuse, la discrimination de la cellule transformée parmi les cellules non transformées (généralement cultivées *in vitro*) est nécessaire *via* des étapes successives de sélection. Il est ainsi toujours nécessaire d'introduire simultanément, en plus du gène d'intérêt, un gène dont l'expression va permettre la sélection des cellules transformées (dit

gène sélectionnable). Ces deux gènes peuvent être portés par un ou deux vecteurs de transformation encore appelés plasmides (*figure 1*).

Méthodes d'introduction de l'ADN

Les méthodes permettant l'introduction d'ADN dans la cellule reposent sur l'action d'un agent chimique (polyéthylène glycol) ou physique (électroporation) appliquée à des cellules débarrassées de leur paroi pecto-cellulosique, de moyens mécaniques (bombardement de microprojectiles, aiguilles de silicone...) ou encore sur l'action biologique (transfert de l'ADN de transfert, ADN-T) d'un plasmide modifié porté par la bactérie *Agrobacterium tumefaciens* (*figure 2*). Le support végétal pour ces deux dernières méthodes est constitué par des organes excisés ou des cultures de tissus (cals) issues de ces organes. La transformation par coculture de cals issus du scutellum d'embryon zygotique avec *Agrobacterium* est aujourd'hui la méthode la plus répandue. Elle permet d'obtenir de façon régulière un grand nombre d'événements de transformation dont une majorité exprime le transgène intégré : 5 à 10 plantes transformées indépendantes sont ainsi obtenues à partir d'un seul fragment de cal cocultivé, sauf dans le cas des variétés dites « récalcitrantes ». Le cycle inoculation des cals-récolte des grains transformés dure de 4 à 6 mois.

Sélection des cellules transformées

Le produit du gène sélectionnable transféré avec le gène d'intérêt va généralement permettre de détoxifier un composé ajouté au milieu de culture mais peut également permettre d'assimiler un substrat particulier (Joersbo *et al.*, 2003 ; Lucca *et al.*, 2001). Les gènes sélectionnables ayant une fonction de détoxification utilisés chez le riz sont les gènes *hptII* (hygromycin phosphotransférase) et *nptII* (neomycin phosphotransférase) conférant respectivement une résistance à l'hygromycine et à la kanamycine. Le gène *bar* (aussi appelé *pat* – *phosphinothricine acetyl transferase*) confère une tolérance à l'herbicide

glufosinate. Cet herbicide est peu efficace pour discriminer les lignées cellulaires transformées à partir de cals alors qu'il a été utilisé avec succès pour la sélection de protoplastes transformés ou de cellules de scutellum d'embryons bombardés.

Un promoteur du gène de l'ADN ribosomal 35S du virus de la mosaïque du chou-fleur (CaMV) est souvent utilisé pour contrôler l'expression du gène sélectionnable tandis que des promoteurs plus efficaces, comme celui du gène codant pour l'ubiquitine 1 du maïs, sont généralement utilisés pour assurer l'expression constitutive du gène d'intérêt. Le gène d'intérêt et le gène sélectionnable sont en général portés par une même molécule d'ADN et vont donc s'intégrer ensemble dans le génome hôte. Il est possible cependant, par « double transformation » utilisant l'*Agrobacterium* (*figure 3*), d'introduire les deux gènes de manière indépendante et d'obtenir leur intégration dans deux sites distincts du génome hôte. Le gène sélectionnable peut ensuite être éliminé par simple ségrégation génétique dans la descendance de la plante transformée (*figure 4*).

Néoformation de plantes transformées

Il existe une grande diversité d'aptitude à la transformation entre les variétés de riz, qui dépend de l'aptitude de leurs tissus à être cultivés *in vitro*. Les variétés de riz du groupe *japonica* montrent généralement une bonne aptitude à former des cals embryogènes, mais ces cals ont une capacité variable à régénérer des plantes entières. Les variétés du groupe *indica* forment généralement des cals qui tendent à générer directement des embryons plutôt qu'à demeurer à l'état de proliférations de cellules embryogènes non différenciées. Si ce comportement est un avantage pour la néoformation finale de plantes, il s'avère désavantageux lors de la coculture avec *Agrobacterium* en conduisant à des réactions de nécrose. Au-delà de ces différences d'aptitude, la culture *in vitro* peut être source de modifications épigénétiques (par exemple, méthylation des séquences ou des histones) et génétiques (délétion, mutation ponctuelle, ou insertion d'éléments transposables) du génome,

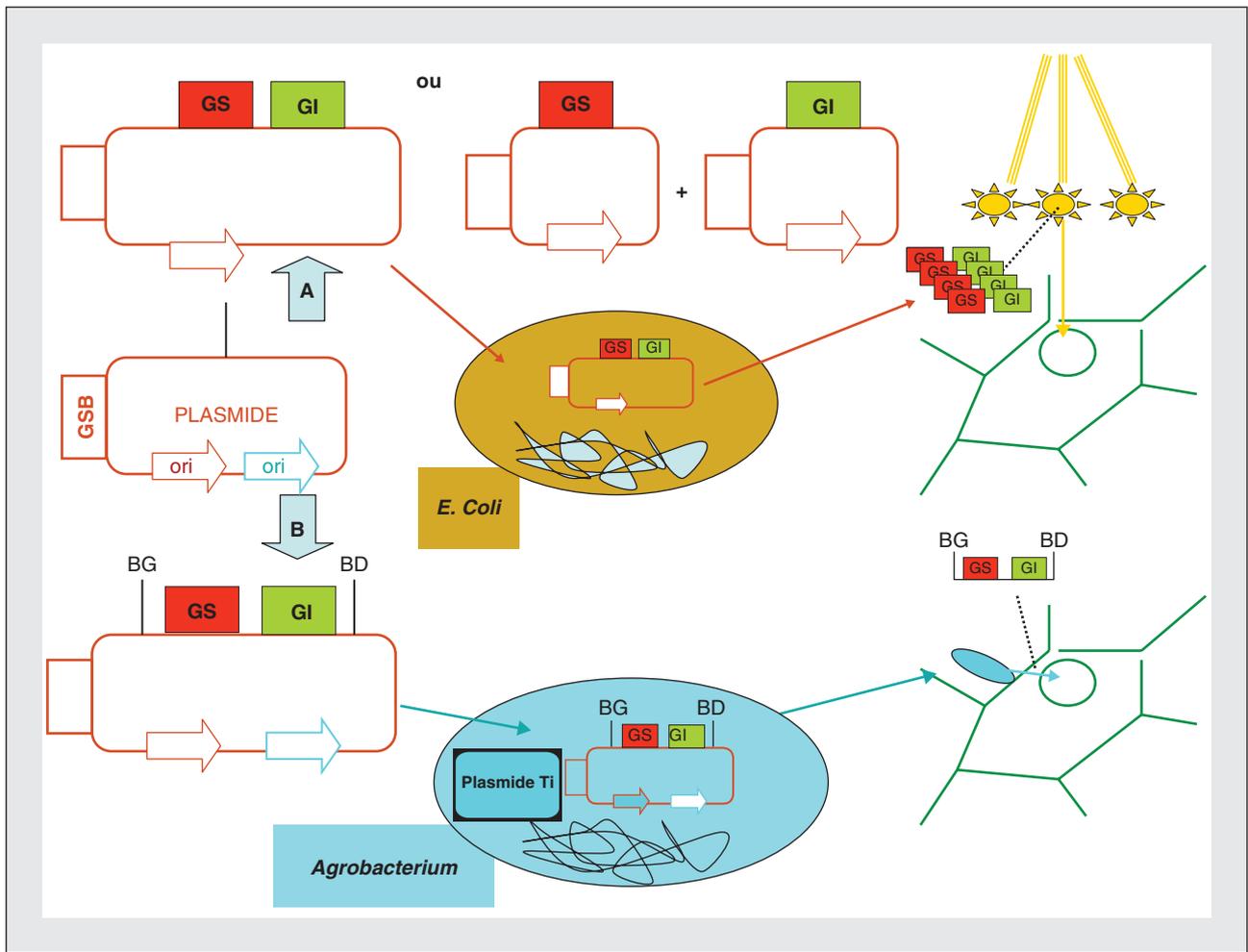


Figure 1. Vecteurs de transformation : cas de la transformation directe (flèche A) et de la transformation biologique par *Agrobacterium* (flèche B).

Figure 1. Transformation vectors for direct transformation (A arrow) and *Agrobacterium*-mediated biological transformation (B arrow).

A. Les cassettes portant les gènes d'intérêt (GI) et de sélection (GS) sont insérées ensemble ou séparément dans un ou des plasmides (molécules d'ADN circulaire) bactériens pour les multiplier dans la bactérie *E. Coli*. GSB : gène de sélection bactérien ; (ori rouge) : origine de réplication dans *E. coli*. Les plasmides sont purifiés, les cassettes isolées à l'aide d'enzymes de restriction et déposées sur les microprojectiles qui vont être introduits dans le noyau de cellules végétales en culture.

B. Origine de réplication (ori bleue) et bordures droites (BD) et gauche (BG) de l'ADN transféré (ADN-T) par *Agrobacterium*. Le plasmide est introduit dans une souche d'*Agrobacterium* portant un plasmide Ti désarmé permettant la coupure puis le transfert de l'ADN-T dans la cellule végétale. Un plasmide porteur de plusieurs ADN-T ou plusieurs plasmides porteurs d'un ADN-T différent peuvent être introduits dans *Agrobacterium*, notamment pour l'élimination du gène de sélection par ségrégation à la génération suivante (voir figure 4).

éléments conduisant à des phénotypes indésirables dans la descendance des plantes régénérées. Il est donc important de faire la différence entre les modifications liées à l'expression du transgène et celles qui sont induites par la culture *in vitro*. Le reséquençage de plantes régénérées chez le riz a montré que la majorité des modifications correspondent à l'accumulation de mutations ponctuelles (Miyao Akio *et al.*, 2012). Des efforts ont donc été faits pour raccourcir les phases de culture *in vitro* (Toki *et al.*, 2006) ou pour opérer une transformation sans culture *in vitro* (c'est-à-dire *in planta*) (Supartana *et al.*, 2005).

Intégration des transgènes

L'intégration de l'ADN-T dans le génome nucléaire des plantes à fleur se fait de manière aléatoire aussi bien en termes de nombre de copies intégrées que de site d'intégration. À ce jour, les méthodes de pilotage par homologie de séquence entre l'ADN introduit et le site receveur, sont maîtrisées seulement pour le génome des organites, d'origine procaryote. La transformation par *Agrobacterium* conduit à l'intégration d'en moyenne 2,5 copies de l'ADN-T à 1,5 site différent (figure 3). Le développement et la

caractérisation de larges collections de lignées d'insertion ADN-T chez le riz ont permis d'avoir une idée plus précise de l'intégration et de l'organisation de l'ADN-T dans le génome (An *et al.*, 2003 ; Sallaud *et al.*, 2004 ; Hsing *et al.*, 2006 ; Zhang J. *et al.*, 2006). La localisation chromosomique de ces inserts a une distribution aléatoire, avec toutefois une tendance à s'insérer préférentiellement sur les chromosomes 1, 2 et 3 qui sont également les plus longs et possèdent les plus fortes densités de gènes. Les ADN-T se retrouvent préférentiellement dans les régions subtélomériques et euchromatiques, plus riches

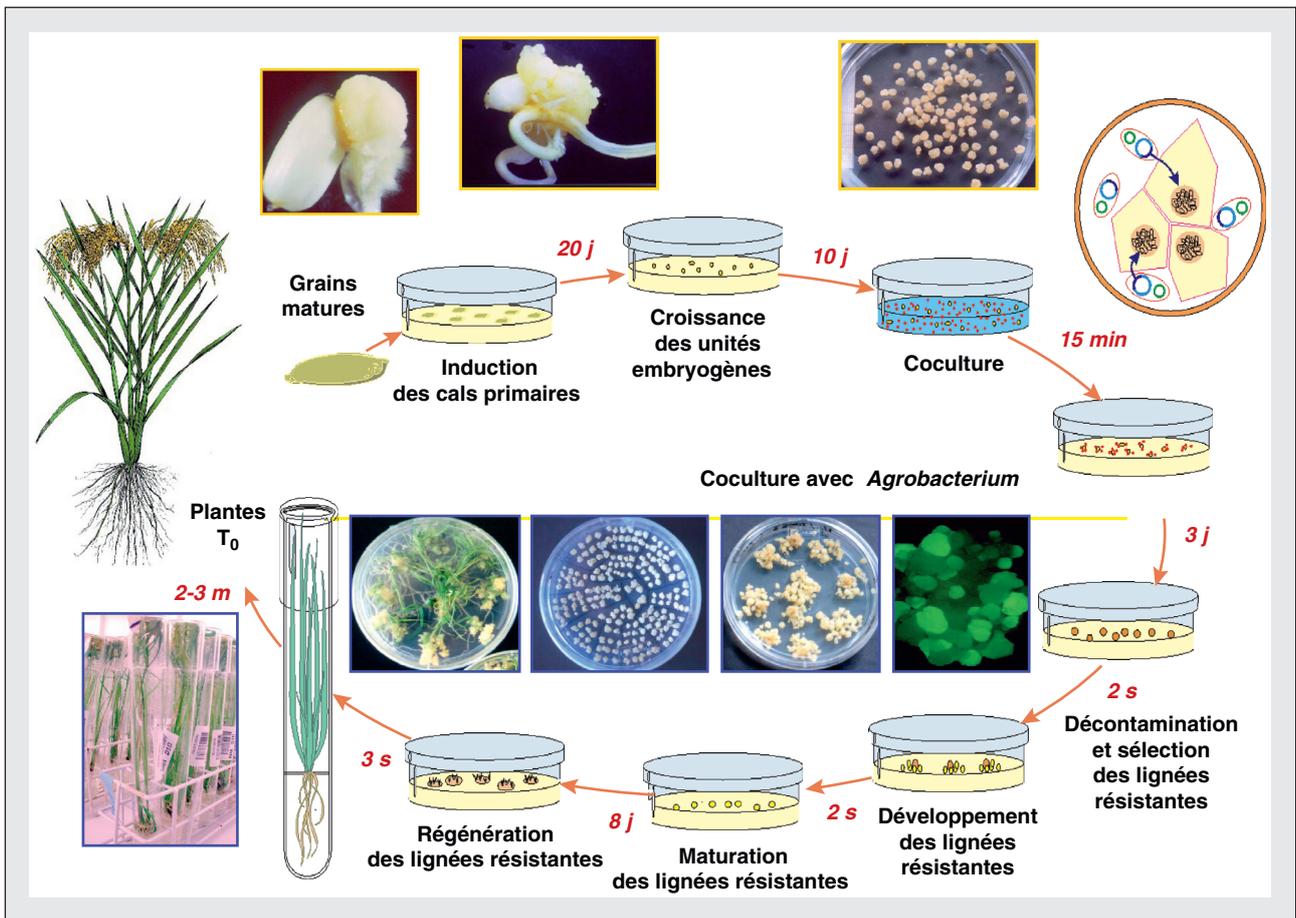


Figure 2. Transformation de culture de tissus de riz par coculture avec la bactérie *Agrobacterium tumefaciens*.

Figure 2. Transformation of rice tissue cultures through co-cultivation with a bacterial suspension of *Agrobacterium tumefaciens*.

Un cal primaire est induit à partir du scutellum de l'embryon du grain mature et va produire des embryons somatiques globulaires ou cals secondaires. Ces cals sont ensuite incubés pendant 3 jours sur un milieu de coculture solide, avant leur transfert sur un milieu permettant la décontamination d'*Agrobacterium* et la sélection des cellules transformées ayant intégré et exprimant les gènes portés par l'ADN-T. Les lignées cellulaires transformées résistantes (ici visualisées par le cotransfert du gène de la protéine de fluorescence verte (GFP) vont se développer en cal puis régénérer des plantes. Après une phase de croissance en tube, les plantes régénérées (T₀) seront transférées en serre jusqu'à la formation de grains T₁.

en gènes, par rapport aux régions centromériques et hétérochromatiques. Les fréquences d'insertion sont similaires entre régions géniques et intergéniques, plus faibles dans les régions contenant des séquences répétées. À l'échelle des gènes, une fréquence plus élevée d'intégration est observée autour du codon de début de traduction (ATG) et du codon de fin de traduction (TGA).

Amélioration des méthodes

La persistance dans le génome de la plante receveuse et de sa descendance d'un gène sélectionnable ou d'autres séquences superflues peut être évité par la mise en oeuvre de différentes

stratégies (figure 4). Un autre défi est d'obtenir l'intégration du transgène à un endroit prédéfini du génome. Cette intégration ciblée permettrait de s'affranchir des variations qualitatives et quantitatives d'expression des transgènes liées au site d'intégration comme d'éviter d'interrompre un gène utile. Le développement récent de nucléases liées à des domaines de reconnaissance spécifiques d'ADN ouvre la possibilité de catalyser l'insertion du transgène par création d'une cassure du double brin d'ADN à haute fréquence à un endroit précis du génome (D'halluin *et al.*, 2008 ; Zhang Feng *et al.*, 2010 ; Cermak *et al.*, 2012) (voir ci-après).

Un outil pour l'analyse fonctionnelle du génome

La compréhension de la fonction d'un gène repose en grande partie sur l'étude du phénotype de plantes chez lesquelles ce gène est inactivé ou au contraire surexprimé. Des informations importantes peuvent également être retirées de la localisation tissulaire ou subcellulaire du produit du gène étudié et de son interaction avec les

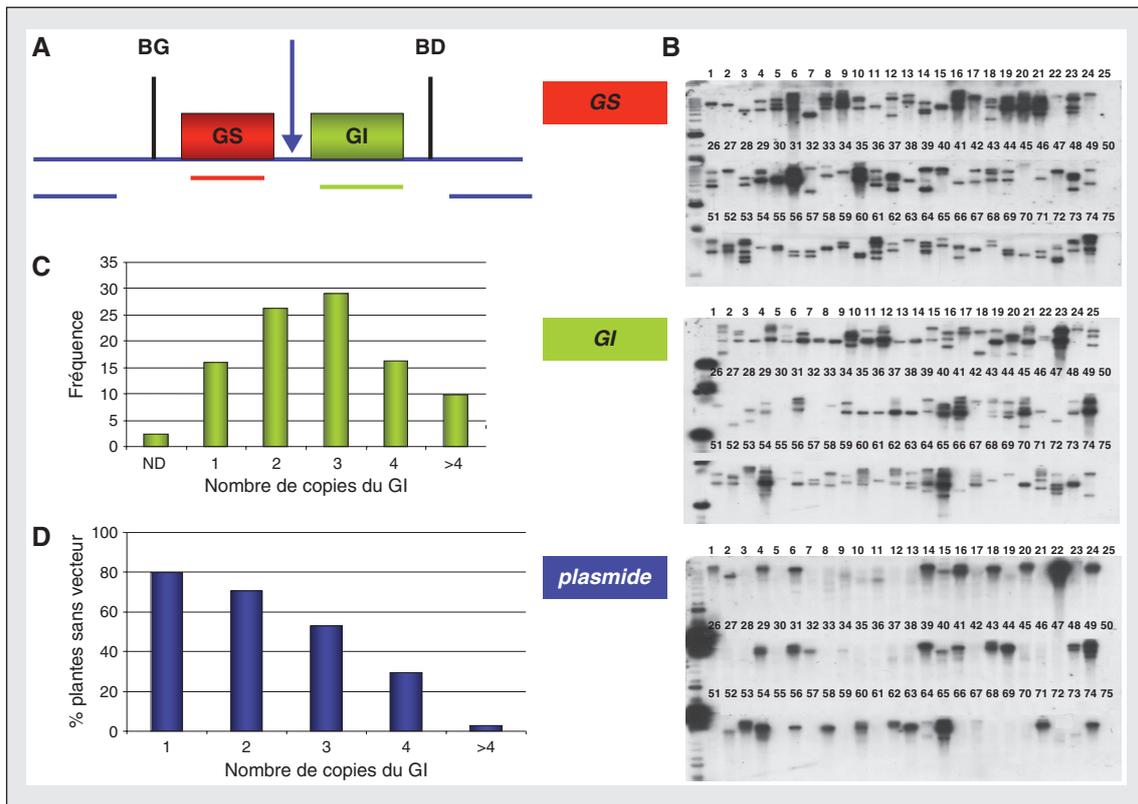


Figure 3. Analyse moléculaire de plantes transformées par *Agrobacterium*.

Figure 3. Molecular characterization of plants transformed by *Agrobacterium*.

L'ADN des plantes T0 (ou transformants primaires) est isolé à partir des feuilles puis soumis à un traitement de digestion par une enzyme de restriction qui coupe le génome de la plante en fragments, notamment au milieu de l'ADN-T intégré (flèche) (A). Les fragments d'ADN ainsi produits sont hybridés avec une sonde moléculaire marquée spécifique du gène d'intérêt (GI), du gène de sélection (GS) et du plasmide (plasmide) bactérien (B). La présence d'une bande signal correspond en général à une copie de l'ADN-T intégré. Les plantes présentant une ou deux copies intégrées et ne présentant pas de séquence plasmidique sont conservées. Le nombre de copies est en moyenne de 2,5 (C). Les plantes ayant une ou deux copies une chance plus importante de ne pas avoir intégré de séquences du plasmide bactérien (D).

produits d'autres gènes. La transformation génétique permet d'appuyer ces approches ou les rend possibles.

Inactivation aléatoire ou ciblée de gènes

L'inactivation de gènes peut être réalisée grâce à l'obtention de larges collections de mutants d'insertion. La mutagenèse insertionnelle repose sur l'intégration au hasard dans le génome d'une séquence d'ADN connue. Cette séquence peut être un élément transposable fonctionnel chez une autre espèce, par exemple les transposons Ac/Ds (Chin *et al.*, 1999) ou En/Spm (Kumar *et al.*, 2005) du maïs et introduit chez le riz par transgenèse ou, comme nous l'avons mentionné plus haut, l'ADN-T lui-même (Jeon *et al.*, 2000). Des éléments mobiles endogènes comme certains rétroélé-

ments activables par la culture *in vitro* (cas de l'élément *Tos17* chez le riz [Miyao *et al.*, 2003]) peuvent également être mis à profit pour créer des collections de mutants sans avoir besoin d'introduire un ADN étranger dans le génome. En développant des collections de plusieurs dizaines, voire centaines, de milliers de plantes contenant chacune une ou plusieurs insertions, on peut s'attendre à ce que chaque gène ait été interrompu au moins une fois dans la collection. Lorsque le mutagène s'intègre dans un gène et interromp sa séquence, il va en altérer la transcription de façon complète ou partielle. En plus de créer une interruption de la séquence, les transgènes peuvent conduire à la détection de gènes s'ils sont équipés de système rapporteur : gène de la β -glucuronidase (GUS) ou de la protéine de fluorescence verte (GFP) dont l'expression dans les cellules trans-

formées peut être détectée par colorimétrie. Dans ces cas, les transgènes sont appelés piège à gène (*gene trap*), à promoteur (*promoter trap*) ou à activateur (*enhancer trap*). Si le transgène s'insère dans un gène ou à proximité d'un gène, le gène rapporteur pourra être activé de façon dominante dans les tissus où le gène endogène est naturellement exprimé. La séquence du transgène étant connue, l'identification d'une insertion dans un gène précis au sein de la collection peut être réalisée par amplification PCR (*polymerase chain reaction*/réaction en chaîne par polymérase) des séquences bordant l'ADN inséré au sein de pools d'ADN des plantes receveuses. Cependant, dans la plupart des cas, cette recherche se fait de façon informatique après séquençage systématique des bordures des sites d'intégration du transgène (Krishnan *et al.*, 2009). Les collections

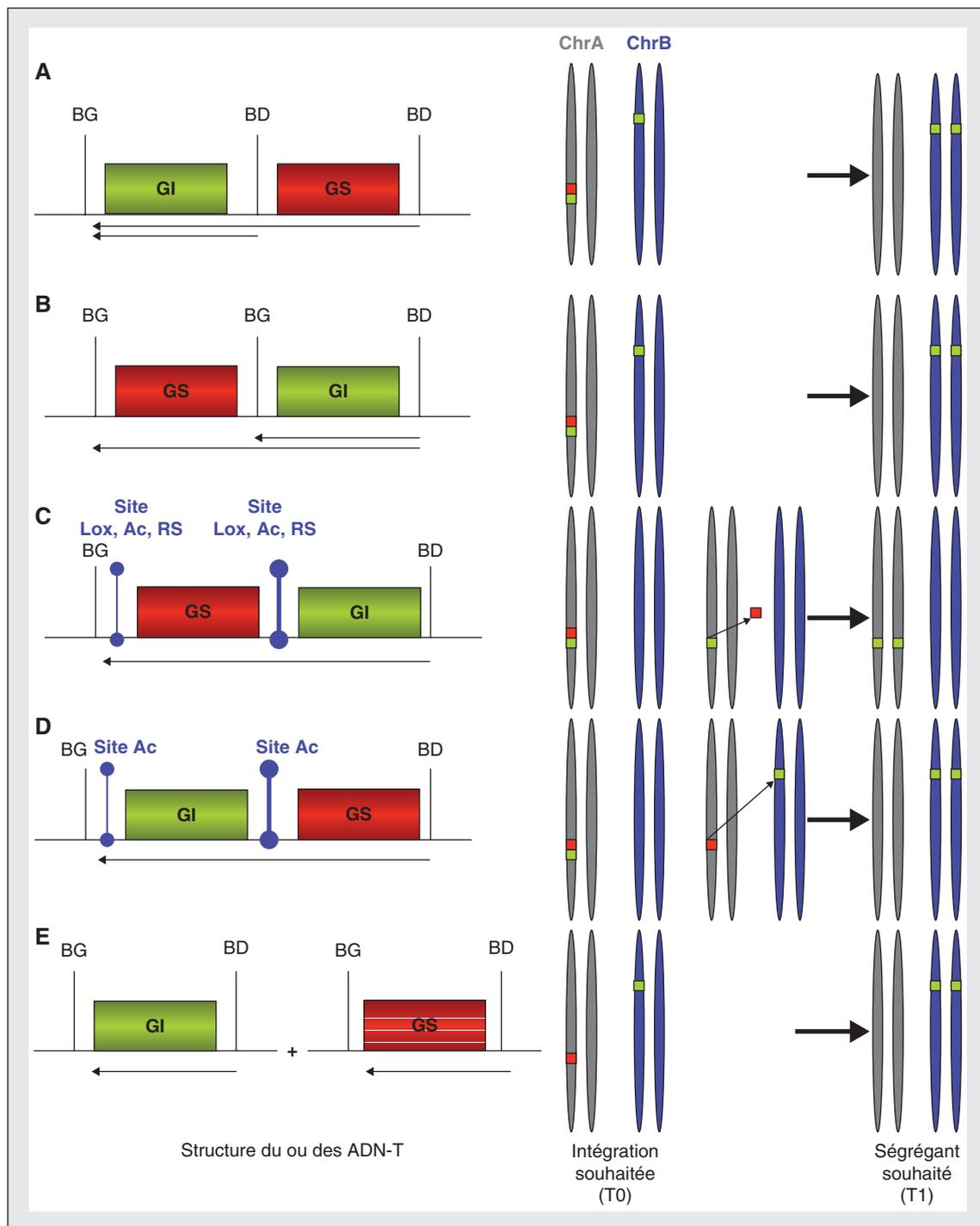


Figure 4. Stratégies pour éliminer le gène de sélection (à droite de haut en bas, les flèches représentent les molécules transférées et intégrées).

Figure 4. Strategies followed for eliminating the selectable marker gene (arrows represent the transferred molecules).

(A) ADN-T à double bordure droite (Lu *et al.*, 2001). (B) ADN-T à double bordure gauche. (C) ADN-T à gène de sélection bordé par des sites reconnus par une recombinase (Cre/lox (Sreekala *et al.*, 2005) ou R/RS (Endo *et al.*, 2002)) ou une transposase *Ac* permettant l'excision et la réintégration du gène d'intérêt à un locus non lié (Cotsaftis *et al.*, 2002). (D) ADN-T à gène d'intérêt bordé par des sites reconnus par une transposase *Ac* permettant l'excision et la réintégration du gène d'intérêt à un locus non lié (Cotsaftis *et al.*, 2002). (E) utilisation de deux ADN-T portés par un seul plasmide (Breitler *et al.*, 2004 ; Komari *et al.*, 1996), ou deux plasmides différents (systèmes pGreen/pSoup (Vain *et al.*, 2003) et leur améliorations récentes de la série pCLEAN (Thole *et al.*, 2007). À gauche, intégration souhaitée des ADN-T sur deux chromosomes différents de la plante transformée régénérée (T0) permettant l'identification dans sa descendance de plantes (T1) porteuses du gène d'intérêt à l'état homozygote dépourvues de gène de sélection.

actuellement disponibles, produites en Chine, Corée, Japon, France et États-Unis, comprennent plus de 500 000 lignées représentant 240 000 insertions dont la position génomique est connue. On estime que 70 % des 38 000 gènes prédits du riz ont été interrompus au moins une fois par les inserts caractérisés dans ces collections.

L'inactivation ciblée d'un gène repose sur une insertion conduite par une homologie de séquence entre la séquence cible et l'ADN introduit et se réalise à la faveur d'une cassure naturelle de l'ADN réparée par la voie dite de recombinaison homologue (RH). Ce type d'événement survient cependant à une fréquence très faible

(1 pour 10 000 à 1 pour 100 000 événements d'intégration) chez les plantes à fleur comme le riz (figure 5). L'augmentation de cette fréquence fait l'objet d'intenses recherches : une amélioration a été d'abord obtenue par l'utilisation de larges séquences d'homologies (plusieurs milliers de paires de bases) sur

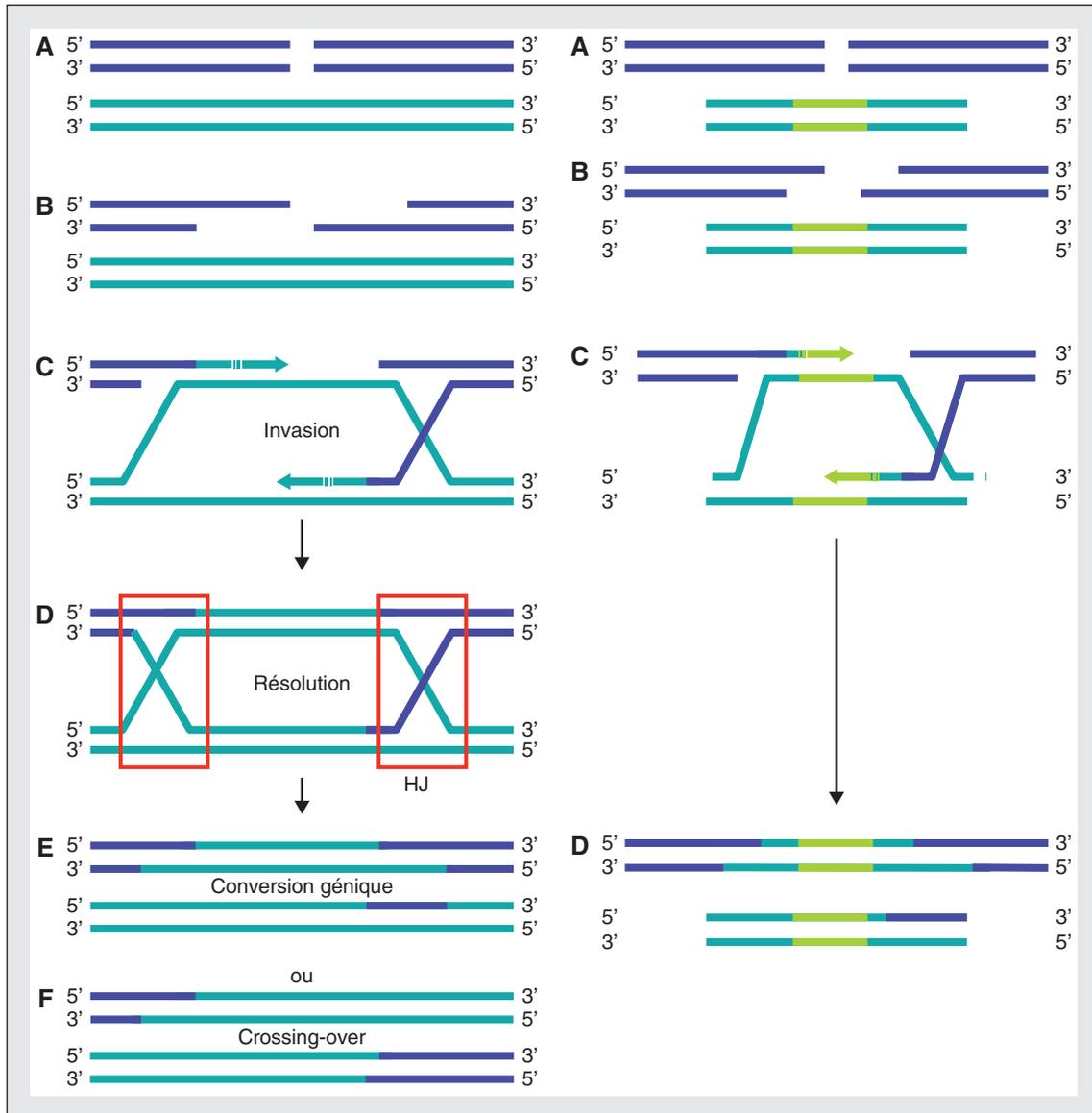


Figure 5. Réparation de l'ADN par recombinaison homologue (gauche) et application à l'intégration ciblée d'un transgène (droite).

Figure 5. Homologous recombination-mediated DNA repair (left) and application for targeted transgene integration (right).

À gauche : la molécule d'ADN cassée est représentée en bleu foncé et la matrice homologue en bleu clair. (A) Une cassure de l'ADN double brin survient. (B) Des exonucléases digèrent les extrémités 5' de la molécule. (C) Les extrémités 3' sortantes vont être utilisées comme amorces pour la synthèse d'ADN à partir d'une matrice homologue, comme un chromosome homologue ou une chromatide sœur ou encore toute séquence d'ADN disponible présentant une homologie. (D) Des jonctions de Holliday (boîtes rouges) se forment et sont résolues soit (E) par conversion génétique où les deux molécules d'ADN conservent leurs régions adjacentes ou (F) par *crossing-over* où elles procèdent à leur échange. Aucune information génétique n'est perdue durant ce processus.

À droite : légendes identiques, mis à part que la molécule d'ADN matrice est introduite par transformation génétique et contient la séquence à substituer ou à intégrer en plus (cas présenté ici en vert clair) flanquée de régions homologues (bleu clair). La molécule d'ADN résultante possède une séquence identique à la molécule de départ avec ici représentée l'intégration du transgène.

l'ADN introduit, elles-mêmes encadrées de gènes « tueurs » dont l'expression contre-sélectionne les événements d'intégration aléatoire. Des fréquences de ciblage génique de 1 à 2 % parmi les événements ayant échappé à la contre-sélection, ont été obtenues dans des expérimentations ciblant les gènes *waxy* et *adb1* (Terada *et al.*, 2002 ; Iida et Terada, 2005). Cette technique demande cependant une méthode de transformation très efficace, puisqu'il faut dimensionner l'expérimentation pour une production potentielle de plusieurs milliers d'événements de transformation, et n'est donc pas appliquée de façon routinière. Une autre voie de recherche vise donc à augmenter la fréquence de RH par manipulation de la machinerie de protéines de recombinaison, favorisant celles qui sont impliquées dans le mécanisme naturel de réparation de l'ADN géno-

mique par recherche d'homologie. Une voie complémentaire est la création avec une haute fréquence d'une cassure de l'ADN au site ciblé (utilisation de nucléases liées à des domaines spécifiques de fixation à une séquence précise de l'ADN génomique). Récemment, l'utilisation de séquences *Transcription Activator Like Effectors* (TALE) couplées à la nucléase Fok1 a permis de créer avec une forte fréquence (plusieurs dizaines de pour-cent) une mutation dans le gène *Os11N3*, impliqué dans la réponse du riz à l'infection par *Xanthomonas* (Li *et al.*, 2012). Cette dernière approche semble cette fois utilisable sans disposer de méthodes de transformation performante. De plus, les deux allèles du gène peuvent être mutés simultanément, permettant l'obtention de plantes homozygotes pour la mutation dès la génération T0.

Les méthodes d'inactivation posttranscriptionnelle sont, elles, basées sur le mécanisme d'interférence de l'ARN (ou RNAi) existant naturellement chez les eucaryotes. Il s'agit de produire un ARN dit interférent complémentaire d'une région plus ou moins importante de l'ARN messager (ARNm) cible, conduisant à la dégradation de celui-ci et de ce fait à l'inhibition de l'accumulation de la protéine correspondante. Différents types de constructions de transgène contenant la totalité ou une partie de la séquence codante ciblée peuvent être utilisés pour induire cette dégradation (Wesley *et al.*, 2001 ; Warthmann *et al.*, 2008) (figure 6). Leur efficacité dépend souvent de l'ARN messager ciblé. Avec ces méthodes, il est possible de créer une population de plantes présentant une gamme de niveaux d'inactivation de l'ARNm cible et d'altération du phénotype associé. Cette inactivation

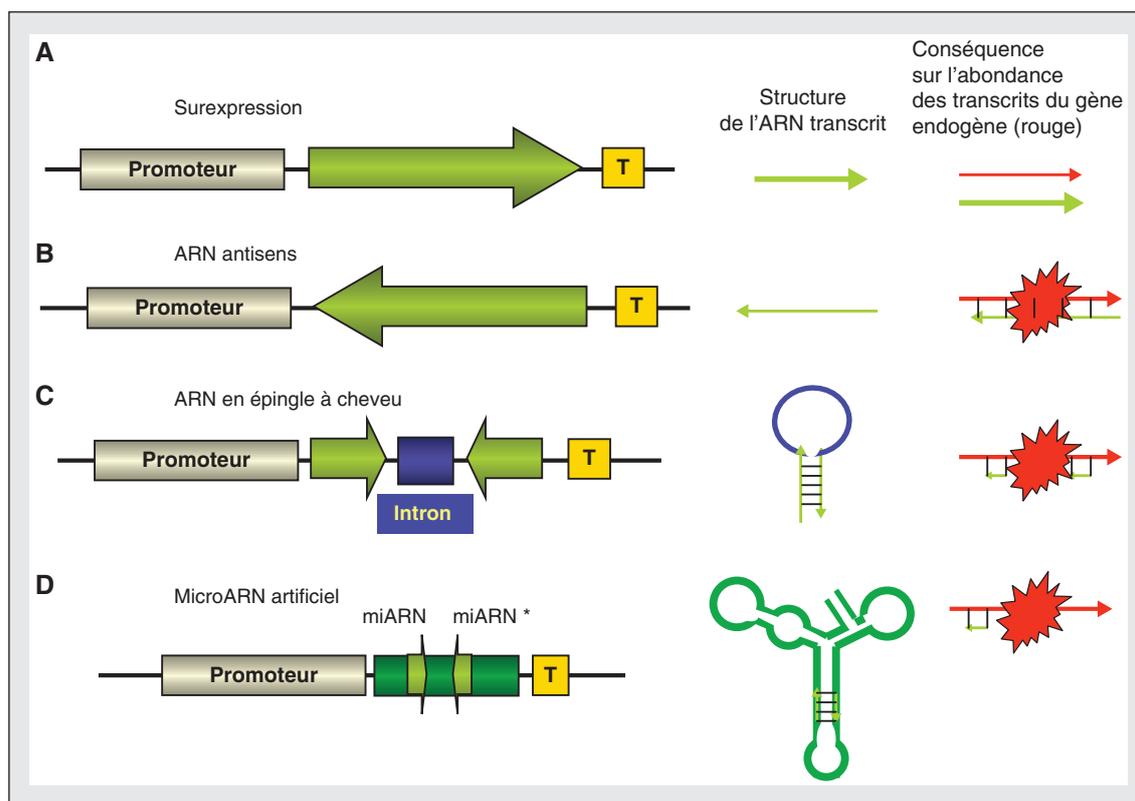


Figure 6. Construits de surexpression (A) et d'inactivation (B-D) de l'expression de gènes.

Figure 6. Transformation vectors for over expression (A) or inactivation (B-D) of gene expression.

(A) Le gène d'intérêt est placé sous le contrôle d'un promoteur constitutif, inductible ou tissu spécifique. (B) Le gène étudié est surexprimé en orientation inverse créant un ARN complémentaire à l'ARN du gène d'étude et provoquant la dégradation de l'ARN double brin formé par la cellule. (C) Des séquences spécifiques au gène étudié de 300-500 pb sont placées en orientation inverse de part et d'autre d'un intron créant un ARN en épingle à cheveu (*hairpin*), provoquant la dégradation de l'ARN double brin formé et des séquences présentant une homologie avec les petits ARN produits de cette dégradation. (D) : microARN artificiel ; deux séquences de 21 nucléotides (nt), prédites comme ciblant spécifiquement le gène étudié et complémentaires l'une avec l'autre (à un nucléotide près) sont insérées aux positions correspondantes d'un microARN naturel sur la séquence du précurseur de ce microARN, connu pour être efficacement clivé dans le génome du riz. Le précurseur de microARN à maturité va libérer un petit ARN double brin de 21 nt qui va cibler très spécifiquement la dégradation des transcrits du gène d'étude.

partielle peut être utile pour l'étude de gènes dont l'absence d'expression est létale, comme par exemple des gènes clés de la reproduction.

Surexpression constitutive, tissu spécifique ou inductible

La surexpression d'un gène est obtenue en le plaçant sous le contrôle d'une séquence régulatrice qui permet sa transcription à un niveau élevé dans tous les organes ou à tous les stades de développement (promoteur dit constitutif : ubiquitine 1 du maïs (Christensen *et al.*, 1992), CsVMV du virus de la mosaïque du manioc (Verdaguer *et al.*, 1998) et *gos2* du riz (De Pater *et al.*, 1992)) (figure 6). Des collections de lignées surexprimant constitutivement les ADNc d'*Arabidopsis* ou du riz ont été créées systématiquement par la recherche publique au Japon (Nakamura *et al.*, 2007) et par la société Crop Design (filiale de BASF). Les mutants « gain de fonction » constitutifs ainsi obtenus, présentent un phénotype correspondant à une expression exagérée du gène étudié.

L'expression constitutive de certains gènes pouvant être létale, notamment lors des divisions *in vitro* des cellules transformées, il peut s'avérer nécessaire d'utiliser un promoteur spécifique d'un organe ou d'un tissu ou inductible par un stimulus externe physique ou chimique. De même, on aura avantage à placer un gène de réponse à un stress sous le contrôle d'un promoteur dont l'activité est déclenchée par l'apparition du stress cible. Plus généralement, il peut être utile de développer des collections de mutants aléatoires portant un ADN-T ou un élément Ds activant l'expression des gènes voisins et permettant ainsi un gain plutôt qu'une perte de fonction. Ce type de construction permet de plus la surexpression du gène à proximité sans modifier le profil spatio-temporel naturel de son expression (Jeong *et al.*, 2002).

Localisation tissulaire et subcellulaire

L'étude de la région régulatrice d'un gène peut être réalisée en associant la séquence de la région promotrice du

gène (en général 1 000-2 000 paires de base en amont du site d'initiation de la traduction) à celle d'un gène rapporteur (GUS ou GFP) et en transférant cette construction dans des plantes transgéniques. De très nombreuses régions régulatrices de gènes ont été disséquées fonctionnellement de cette façon. Il est aussi possible d'étudier la compartimentation cellulaire des produits d'un gène (adressage nucléaire, réticulum, organites, membrane cytoplasmique ou tonoplasmique) en faisant s'accumuler dans la plante de façon constitutive une protéine hybride entre la protéine d'intérêt et une protéine rapportrice de type GFP dont la détection renseignera sur la localisation de la protéine d'intérêt. Ces méthodes sont susceptibles d'apporter une information supplémentaire sur la fonction des gènes.

Conclusion et perspectives

Une revue récente a indiqué que la fonction de plus de 600 gènes du riz a déjà été validée ou déterminée par les méthodes d'analyse fonctionnelle, faisant largement appel aux méthodes de transformation génétique que nous venons de présenter (Jiang *et al.*, 2011). Parmi ces gènes figurent aussi ceux issus du clonage positionnel de loci à effet quantitatif (QTL, *quantitative trait locus*) localisés par cartographie génétique ou d'association. La fonction de ces gènes, contrôlant parfois des caractères d'intérêt agronomique (productivité, résistance aux maladies et tolérance aux contraintes abiotiques) a été validée en introduisant, par transformation génétique, leur allèle favorable chez une variété qui en était dépourvue.

Le défi méthodologique majeur est aujourd'hui la maîtrise de la conversion allélique *in situ*, c'est-à-dire la modification dirigée (par substitution, addition ou délétion) dans la séquence du gène cible des bases à effet défavorables par des bases conduisant à un effet favorable, sans introduction d'autre séquence exogène et en préservant le contexte génique. Une telle méthode permettrait d'écartier les risques d'introduction de caractères défavorables et

d'accéder à une plus large gamme de variabilité du caractère cible. Les technologies basées sur des nucléases fusionnées à des domaines de reconnaissance spécifique à l'ADN couplées à une maîtrise du processus de recombinaison homologue ouvrent la voie pour réaliser une telle chirurgie génique. Les conversions à opérer pourront être inspirées par la variabilité existante du vivant qui est le fruit d'une longue adaptation évolutive. La conversion allélique pourra aussi contribuer à l'accumulation plus rapide d'allèles favorables des gènes d'intérêt agronomique, principal objectif des programmes d'amélioration génétique du riz. ■

Remerciements

Les auteurs remercient le soutien d'Agropolis Fondation qui permet *via* la plateforme internationale d'accueil REFUGE (*RicE FUnctional GENomics platform*) la mise en œuvre de ces méthodes pour les projets de recherche des étudiants et chercheurs du Sud et du Nord. Les auteurs remercient Martine Bès, Nadège Lanau, Florence Artus, Aurore Vernet, Murielle Portefaix, Christian Chaine, Rémy Michel, Eve Lorenzini pour leur participation aux projets de recherche de l'équipe. Ils remercient Nourollah Ahmadi pour sa relecture du manuscrit et Marc Lartaud pour l'infographie.

Références

- An SY, Park S, Jeong DH, Lee DY, Kang HG, Yu JH, *et al.*, 2003. Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiology* 133 : 2040-7.
- Breitler JC, Meynard D, Van Boxtel J, Royer M, Bonnot F, Cambillau L, *et al.*, 2004. A novel two T-DNA binary vector allows efficient generation of marker-free transgenic plants in three elite cultivars of rice (*Oryza sativa* L.). *Transgenic Research* 13 : 271-87.
- Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, *et al.*, 2012. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* 39: e82. doi: 10.1093/nar/gkr218.
- Chin HG, Choe MS, Lee SH, Park SH, Park SH, Koo JC, *et al.*, 1999. Molecular analysis of rice plants harboring an Ac/Ds transposable element-mediated gene trapping system. *Plant Journal* 19 : 615-23.
- Christensen AH, Sharrock RA, Quail PH, 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts

- by electroporation. *Plant Molecular Biology* 18 : 675-89.
- Cotsaftis O, Sallaud C, Breitler JC, Meynard D, Greco R, Pereira A, *et al.*, 2002. Transposon-mediated generation of T-DNA- and marker-free rice plants expressing a Bt endotoxin gene. *Molecular Breeding* 10 : 165-80.
- D'Halluin K, Vanderstraeten C, Stals E, Cornelissen M, Ruiters R, 2008. Homologous recombination: a basis for targeted genome optimization in crop species such as maize. *Plant Biotechnology Journal* 6 : 93-102.
- de Pater BS, van der Mark F, Rueb S, Katagiri F, Chua N-H, Schilperoort RA, *et al.*, 1992. The promoter of the rice gene GOS2 is active in various different monocot tissues and binds rice nuclear factor ASF-1. *Plant Journal* 2 : 837-44.
- Endo S, Sugita K, Sakai M, Tanaka H, Ebinuma H, 2002. Single-step transformation for generating marker-free transgenic rice using the ipt-type MAT vector system. *Plant Journal* 30 : 115-22.
- Hsing YI, Chern CG, Fan MJ, Lu PC, Chen KT, Lo SF, *et al.*, 2006. A rice gene activation/knockout mutant population for high throughput functional genomics analysis. *Plant Molecular Biology* 63 : 351-64.
- Iida S, Terada R, 2005. Modification of endogenous natural genes by gene targeting in rice and other higher plants. *Plant Molecular Biology* 59 : 205-19.
- Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, *et al.*, 2000. T-DNA insertional mutagenesis for functional genomics in rice. *Plant Journal* 22 : 561-70.
- Jeong DH, An SY, Kang HG, Moon S, Han JJ, Park S, *et al.*, 2002. T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiology* 130 : 1636-44.
- Jiang Y, Cai Z, Xie W, Long T, Yu H, Zhang Q, 2011. Rice functional genomics research: Progress and implications for crop genetic improvement. *Biotechnology Advances* 30 : 1059-70. doi: 10.1016/j.biotechadv.2011.08.013.
- Joersbo M, Jorgensen K, Brunstedt J, 2003. A selection system for transgenic plants based on galactose as selective agent and a UDP-glucose : galactose-1-phosphate uridylyltransferase gene as selective gene. *Molecular Breeding* 11 : 315-23.
- Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T, 1996. Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. *Plant J* 10 : 165-74.
- Krishnan A, Guiderdoni E, An G, Hsing Y-iC, Han C-d, Lee MC, *et al.*, 2009. Mutant resources in rice for functional genomics of the grasses. *Plant Physiology* 149 : 165-70. doi: 10.1104/pp.108.128918.
- Kumar CS, Wing RA, Sundaresan V, 2005. Efficient insertional mutagenesis in rice using the maize *En1* Spm elements. *Plant Journal* 44 : 879-92.
- Li T, Liu B, Spalding MH, Weeks DP, Yang B, 2012. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nature Biotechnology* 30 : 390-2.
- Lu HJ, Zhou XR, Gong ZX, Upadhyaya N, 2001. Generation of selectable marker-free transgenic rice using double right border (DRB) binary vectors. *Australian Journal of Plant Physiology* 28 : 241-8.
- Lucca P, Ye XD, Potrykus I, 2001. Effective selection and regeneration of transgenic rice plants with mannose as selective agent. *Molecular Breeding* 7 : 43-9.
- Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe KS, *et al.*, 2003. Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. *Plant Cell* 15 : 1771-80.
- Miyao A, Nakagome M, Ohnuma T, Yamagata H, Kanamori H, Katayose Y, *et al.*, 2012. Molecular spectrum of somaclonal variation in regenerated rice revealed by whole-genome sequencing. *Plant and Cell Physiology* 53 : 256-64. doi: 10.1093/pcp/pcr172.
- Nakamura H, Hakata M, Amano K, Miyao A, Toki N, Kajikawa M, *et al.*, 2007. A genome-wide gain-of-function analysis of rice genes using the FOX-hunting system. *Plant Molecular Biology* 65 : 357-71.
- Sallaud C, Gay C, Larmande P, Bes M, Piffanelli P, Piegou B, *et al.*, 2004. High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. *Plant Journal* 39 : 450-64.
- Sreekala C, Wu L, Gu K, Wang D, Tian D, Yin Z, 2005. Excision of a selectable marker in transgenic rice (*Oryza sativa* L.) using a chemically regulated *Cre/loxP* system. *Plant Cell Reports* 24 : 86-94.
- Supartana P, Shimizu T, Shioiri H, Nogawa M, Nozue M, Kojima M, 2005. Development of simple and efficient in planta transformation method for rice (*Oryza sativa* L.) using *Agrobacterium tumefaciens*. *Journal of Bioscience and Bioengineering* 100 : 391-7.
- Terada R, Urawa H, Inagaki Y, Tsugane K, Iida S, 2002. Efficient gene targeting by homologous recombination in rice. *Nature Biotechnology* 20 : 1030-4.
- Thole V, Worland B, Snape JW, Vain P, 2007. The pCLEAN dual binary vector system for *Agrobacterium*-mediated plant transformation. *Plant Physiology* 145 : 1211-9. doi: 10.1104/pp.107.108563.
- Toki S, Hara N, Ono K, Onodera H, Tagiri A, Oka S, *et al.*, 2006. Early infection of scutellum tissue with *Agrobacterium* allows high-speed transformation of rice. *The Plant Journal* 47 : 969-76.
- Vain P, Afolabi AS, Worland B, Snape JW, 2003. Transgene behaviour in populations of rice plants transformed using a new dual binary vector system: pGreen/pSoup. *Theoretical and Applied Genetics* 107 : 210-7.
- Verdaguer B, de Kochko A, Fux CI, Beachy RN, Fauquet C, 1998. Functional organization of the cassava vein mosaic virus (CsVMV) promoter. *Plant Molecular Biology* 37 : 1055-67.
- Warthmann N, Chen H, Ossowski S, Weigel D, Hervé P, 2008. Highly specific gene silencing by artificial miRNAs in rice. *PLoS ONE* 3 : e1829.
- Wesley SV, Helliwell CA, Smith NA, Wang MB, Rouse DT, Liu Q, *et al.*, 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant Journal* 27 : 581-90.
- Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, *et al.*, 2010. High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *Proceedings of the National Academy of Sciences* 107 : 12028-33. doi: 10.1073/pnas.0914991107.
- Zhang J, Li C, Wu C, Xiong L, Chen G, Zhang Q, Wang S, 2006. RMD: a rice mutant database for functional analysis of the rice genome. *Nucleic Acids Research* 34 (Database issue): D745-8.
- Zhang Q, Li J, Xue Y, Han B, Deng XW, 2008. Rice 2020: A Call For An international coordinated effort in rice functional genomics. *Molecular Plant* 1 : 715-9. doi: 10.1093/mp/ssn043.