



## Impact des nouvelles techniques de modification des génomes sur la sélection végétale.

**Josep M. CASACUBERTA**, Center for Research in Agricultural Genomics, CRAG (CSIC-IRTA-UAB-UB), Campus UAB, Cerdanyola del Vallès 08193 Barcelona, Espagne

**Fabien NOGUÉ**, Directeur de Recherche, INRA AgroParisTech, IJPB, UMR 1318, Centre INRA de Versailles, route de Saint Cyr, 78026 Versailles Cedex,

Manuscrit révisé le 7 décembre 2016 - Publié le 29 décembre 2016

**Résumé :** De nouveaux outils de modification de l'ADN sont maintenant disponibles pour l'ingénierie des génomes végétaux, y compris ceux des plantes de grande culture. Ces nouvelles techniques peuvent théoriquement accélérer les processus qui ont pour but d'augmenter la variabilité génétique. Nous expliquons ici pourquoi la diversité génétique a toujours été la matière première pour les sélectionneurs et discutons comment ceux-ci ont su tirer parti des dernières innovations scientifiques pour utiliser au mieux cette diversité génétique. Nous présentons également comment l'émergence des nouvelles techniques de réécriture des génomes donne aux sélectionneurs des outils extrêmement puissants pour l'amélioration des plantes mais aussi pourquoi leur utilisation va exiger des sélectionneurs, aidés des chercheurs, une caractérisation plus précise des gènes qui sous-tendent cette diversité génétique. Relever ces défis devrait permettre d'associer des allèles d'intérêt agronomique d'une manière inédite et contrôlée pour le développement efficace d'idéotypes, bien adaptés à des conditions environnementales diverses, et permettant des innovations pour les qualités des productions végétales au niveaux alimentaires ou industriels.

**Mots-clés :** CRISPR, diversité génétique, modification des génomes, QTL, sélection végétale, SDN.

**Abstract :** Genome engineering tools are now available in plants, including major crops, to modify in a predictable manner a given gene. These new techniques can theoretically accelerate the process the plant breeding process. Here we explain why genetic diversity has always been the raw material for breeders and discuss how breeders have taken advantage of the latest scientific innovations to make the best use of this genetic diversity. We also present how the emergence of new genome editing techniques give breeders new tools for plant breeding but also why this will require the breeders, assisted by researchers, to more precisely characterize the genes that underlie this genetic diversity. Tackling these challenges should permit the breeders to take advantage of genetic diversity in an unprecedented and controlled way for the effective development of new ideotypes, well adapted to diverse environmental conditions and allowing innovations for plant production at food or industrial levels.

**Keywords:** CRISPR, Genetic diversity, Genome engineering, Plant breeding, QTL, Site-directed nucleases.

## Glossaire

(les numéros se réfèrent aux citations dans le texte)

<sup>1</sup> **Single-Nucleotide Polymorphism (SNP)** : le polymorphisme nucléotidique simple (PNS) ou polymorphisme d'un seul nucléotide est, en génétique, la variation (polymorphisme) d'une seule paire de bases du génome, entre individus d'une même espèce.

<sup>2</sup> **Whole Genome Duplication (WGD)** : toute duplication du génome, ou polyploïdie, est le produit de non-disjonction des chromosomes lors de la méiose qui se traduit par des copies supplémentaires de l'ensemble du génome. La polyploïdie est fréquente chez les plantes, mais a également eu lieu chez les animaux. Après duplications du génome entier de nombreux gènes surnuméraires sont finalement perdus, et reviennent à l'état de singleton. Cependant, certains gènes peuvent rester en copies multiples et conduire à de l'innovation adaptative.

<sup>3</sup> **Éléments transposables** : séquences d'ADN mobiles dont l'existence a été mise en évidence par Barbara McClintock dans les années 40 chez le maïs. Ces séquences sont présentes dans toutes les branches de l'arbre du vivant et peuvent représenter une grande partie du génome (environ 40% du génome chez l'Homme et jusqu'à 90% chez le blé).

<sup>4</sup> **Syndrome de domestication** : ensemble de caractères morphologiques, physiologiques et comportementaux héréditaires présents dans un ensemble d'espèces domestiquées mais absents chez leurs variantes sauvages.

<sup>5</sup> **TILLING (Targeting Induced Local Lesions in Genomes)** : méthode de biologie moléculaire utilisée en génétique inverse. Le principe est de combiner des méthodes de mutagenèse aléatoire à des méthodes modernes d'analyse de l'ADN, permettant l'identification à haut débit de mutation ponctuelle.

<sup>6</sup> **Mutation nulle** : Mutation qui élimine complètement la fonction d'un gène, en général par délétion ou mutation stop.

<sup>7</sup> **Nucléases site dirigées, Site Directed Nucleases (SDN)** : Les nucléases site dirigées (Zinc Finger, TALEN, CRISPR-Cas9) sont des enzymes de restriction artificielles constituées d'un domaine de liaison à l'ADN, et d'un domaine ayant la capacité de cliver l'ADN. Dans le cas du CRISPR-Cas9 le domaine de liaison à l'ADN est composé d'une molécule d'ARN.

<sup>8</sup> **Cassure double brin (CDB)** : Les cassures double-brin de l'ADN sont des lésions majeures induites par divers agents mutagènes endogènes ou exogènes ; elles peuvent être source d'instabilité génétique. Deux mécanismes entrent en compétition pour la réparation des cassures double-brin : la recombinaison homologue (RH) et la jonction d'extrémités non homologues (NHEJ).

<sup>9</sup> **Quantitative Trait Loci (QTL)** : région plus ou moins grande d'ADN qui est étroitement associée à un caractère quantitatif, c'est-à-dire une région chromosomique où sont localisés un ou plusieurs gènes à l'origine du caractère en question.

<sup>10</sup> **Déséquilibre de liaison, Linkage Disequilibrium (LD)** : Il y a déséquilibre de liaison si la fréquence des gamètes porteurs des allèles de deux locus différents A et B est différente du produit des fréquences des allèles, c'est-à-dire s'il y a association préférentielle entre deux allèles.

<sup>11</sup> **Jonction d'extrémités non homologues, Non-Homologous End-Joining (NHEJ)** : La jonction d'extrémités non homologues est un terme qui recouvre des mécanismes de réparation de l'ADN qui permettent de réparer des lésions provoquant des cassures double brin. Si le NHEJ classique (C-NHEJ) est considéré comme plutôt fidèle, le NHEJ alternatif (Alt-NHEJ) est lui non conservatif et conduit à des mutations.

<sup>12</sup> **Recombinaison homologue (RH)** : La recombinaison homologue est un type de recombinaison génétique dans laquelle des séquences nucléotidiques sont échangées entre deux molécules similaires ou identiques d'ADN. Elle est utilisée par les cellules pour réparer fidèlement les cassures double brin.

<sup>13</sup> **Séquence dupliquées en tandem (TAG, Tandem Arrayed Genes)** : Les séquences répétées sont très importantes du point de vue de l'évolution des génomes. Elles sont la marque d'un processus continu de duplications et de réarrangements des chromosomes de toutes les espèces. Chez certaines espèces, comme *Arabidopsis* ou le riz, la plupart des TAG sont le résultat d'un cross-over chromosomique déséquilibré.

## 1- Les mutations à la base de la diversité génétique.

En dépit de la remarquable fidélité de la réplication de l'ADN et de la grande efficacité des systèmes de réparation du génome, des mutations apparaissent dans l'ADN à une fréquence qui rend potentiellement chaque organisme unique. Ce taux de mutation a pu être estimé chez la plante modèle *Arabidopsis thaliana* à  $7 \times 10^{-9}$  mutation par base et par génération (Ossowski et al. 2010), ce qui signifie que même pour ce génome très compact d'environ 157 millions de paires de bases, une mutation, en moyenne, est incorporée dans le génome à chaque génération. Ces mutations, qui, dans certains cas, provoquent des différences phénotypiques entre les individus, sont la matière première avec laquelle la sélection naturelle peut fonctionner, rendant l'adaptation des espèces possible.

Les mutations génétiques les plus fréquentes sont de courtes insertions/délétions ou modifications de nucléotides conduisant à des polymorphismes nucléotidiques simples<sup>1</sup> (SNP). Cependant, les génomes subissent également de nombreux changements plus drastiques, y compris des réarrangements chromosomiques, de longues délétions, des duplications et des insertions de grandes séquences comme les éléments transposables. Il est d'ailleurs considéré que les deux mécanismes majeurs à l'origine de la dynamique et de la plasticité des génomes sont, les événements de doublement du génome entier (Wendel 2015) d'une part, et d'autre part la prolifération différentielle, et spécifique à chaque individu, de différents types d'éléments transposables<sup>3</sup> (Lisch 2013). Tous ces mécanismes confèrent aux génomes végétaux une forte plasticité, par rapport aux génomes animaux (Murat et al. 2012), qui leur permet de s'adapter à de nouveaux environnements. Ces caractéristiques ont également été exploitées pour la domestication et l'amélioration des plantes. La domestication des espèces sauvages, ainsi que leur amélioration du point de vue agronomique, sont donc des versions particulières de l'évolution, dirigées principalement par la sélection humaine. Les principes et les mécanismes qui les régissent ne sont nullement différents de ceux qui sont impliqués dans l'évolution naturelle des espèces. D'ailleurs, la domestication et la sélection variétale constituent d'excellents modèles pour étudier l'évolution en général (Meyer et Purugganan 2013 ; Olsen et Wendel 2013a).

### 1.1. Des gènes majeurs impliqués dans la domestication et l'amélioration des plantes.

Au cours des 10 dernières années, de nombreux allèles, impliqués dans la domestication ou l'amélioration des plantes ont été caractérisés et nous pouvons maintenant commencer à identifier les gènes et les mutations causales de certaines différences phénotypiques clés entre les plantes cultivées et leurs parents sauvages. Ce sont essentiellement des caractères qui facilitent la culture ou la récolte des plantes considérées, et qui sont généralement rassemblés sous le terme de "syndrome de domestication"<sup>4</sup> (Olsen et Wendel 2013b). Ces différents caractères incluent la perte de l'égrenage spontané des épis, la perte de dormance des graines, l'uniformité de la germination et enfin la dominance apicale qui permet d'augmenter la densité des plantes dans un champ cultivé. Ces caractères ont été sélectionnés plusieurs fois et de manière indépendante lors du processus de domestication. Cependant, les gènes à l'origine des changements phénotypiques observés ne sont pas nécessairement identiques, laissant entrevoir une certaine flexibilité des processus d'adaptation (Gaut 2015). C'est le cas, par exemple, des deux événements indépendants de domestication du haricot commun, l'un ayant eu lieu dans ce qui est maintenant le Mexique et l'autre dans les Andes. En effet, une analyse génétique des populations a montré que, parmi les gènes présentant des signes de sélection, la proportion de

ceux qui sont communs aux deux événements n'est pas différente de celle observée sur les gènes ne présentant pas de signes de sélection (Schmutz et al. 2014).

Si certains caractères semblent contrôlés par un seul gène il faut pourtant noter que de nombreux caractères liés à la domestication ou à la sélection variétale ne sont pas monogéniques, mais impliquent de nombreux gènes avec des effets faibles cumulables (Olsen et Wendel 2013b). Même en tenant compte de la complexité des changements génomiques associés aux processus de domestication ou de sélection variétale, l'analyse des exemples relativement simples déjà caractérisés permet de déduire les tendances générales de ce processus. De plus, ces analyses nous ont fourni une collection de gènes et d'allèles spécifiques impliqués dans des changements phénotypiques importants d'un point de vue agronomique et qui peuvent être utiles pour la sélection variétale. Ces changements dans la morphologie de la plante ou dans d'autres phénotypes complexes impliquent souvent des régulateurs de la transcription et du développement, et plus rarement des protéines de structure ou des enzymes (Doebley et al. 2006 ; Olsen et Wendel 2013b).

Les mutations causales à l'origine des événements de domestication sont fréquemment des mutations pouvant aboutir au remplacement d'un acide aminé par un codon "stop", à un changement de site d'épissage, ou à un changement d'acide aminé entraînant la perte d'activité de la protéine. Par exemple, le passage d'une croissance de type rampante du riz sauvage ancestral (*Oryza rufipogon*) à une croissance de type érigée des riz cultivés (*Oryza sativa*), un des événements clés dans la domestication du riz, a été le résultat de la sélection d'une seule mutation dans le gène *PROG1* (PROSTRATE GROWTH 1). Cette mutation a conduit à une substitution d'un acide aminé dans le facteur de transcription *PROG1* entraînant une perte de fonction (Jin et al. 2008 ; Tan et al. 2008).

Des mutations induisant des changements dans les séquences cis-régulatrices des gènes ou des changements d'acides aminés conduisant à une activité modifiée sont également relativement fréquentes. Ainsi, la diminution du nombre et de la taille des branches axillaires du maïs par rapport au téosinte, qui ont été des éléments essentiels de la domestication de cette espèce, sont dus à une mutation affectant l'expression du gène *TB1* (teosinte branched 1) (Wang et al. 1999). Dans ce cas, la mutation causale est liée à l'insertion d'un rétrotransposon de type Hopscotch à une distance de 60 kb en amont de la région codante du gène *TB1*, qui augmente son niveau d'expression (Studer et al. 2011). *TB1* code un régulateur de transcription qui réprime la croissance ; sa surexpression dans les bourgeons axillaires inhibe leur développement. En ce qui concerne les mutations qui modifient la fonction d'une protéine, un exemple intéressant est celui du gène Flowering Locus T (*FT*) du tournesol qui a joué un rôle essentiel dans la domestication de cette espèce. Il a été démontré qu'une mutation par décalage de phase de lecture dans le gène *FT1* d'*Helianthus annuus* (HaFT1) génère un allèle dominant négatif qui interagit avec un autre gène paralogue *FT* et bloque son action, entraînant une floraison retardée (Blackman et al. 2010).

En plus des caractères liés au processus de domestication, les plantes d'intérêt agronomique ont également été intensivement modifiées pour d'autres caractères comme la qualité ou la diversité de leurs produits dérivés. Là encore, des mutations spontanées ont provoqué des allèles nuls (protéine non fonctionnelle) ou des modifications de l'expression de certains gènes. Un exemple intéressant qui illustre la grande plasticité des génomes végétaux est la mutation associée au gène *SUN*, cause du phénotype allongé du fruit de la tomate, qui résulte de la

duplication de ce gène suite au déplacement d'un rétrotransposon, la copie dupliquée du gène étant placée sous le contrôle d'un nouveau promoteur (Xiao et al. 2008).

En résumé, la domestication et la sélection variétale ont consisté en des modifications importantes du phénotype des plantes à travers la sélection récurrente de mutations spontanées. Ces mutations comprennent une diversité de modifications génétiques, des polymorphismes nucléotidiques simples (PNS), des petites délétions, des réarrangements chromosomiques, des duplications de gènes et des événements d'insertion/délétion de transposons, qui ont modifié un nombre important de gènes qui sont souvent des régulateurs de processus clés du développement de la plante.

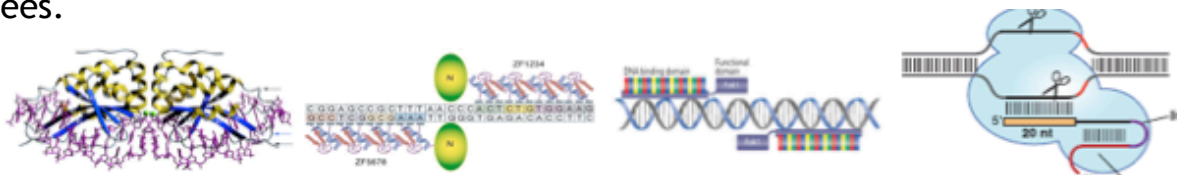
### ***1.2. L'utilisation de la mutagenèse artificielle.***

Durant la majeure partie de ses 10 000 ans d'histoire, la sélection des plantes a été faite en grande partie de manière non consciente. Cependant, au cours de son histoire récente, plusieurs avancées de la connaissance scientifique ont eu un impact immédiat sur les pratiques de sélection variétale. Ainsi, la compréhension du rôle clé de la variabilité génétique et de la sélection a profondément influencé la pratique des sélectionneurs. La découverte que des agents physiques et chimiques (rayonnements gamma, rayonnements X, méthanesulfonate d'éthyle) pouvaient induire des mutations dans l'ADN (Muller 1928 ; Stadler 1928) a entraîné le lancement de programmes de mutagenèse pour augmenter la variabilité génétique. Au cours des 70 dernières années plus de 3 200 nouvelles variétés de plantes chez plus de 200 espèces ont été obtenues par mutagenèse induite (IAEA/FAO mutant variety database, <https://mvd.iaea.org/>). Cette approche a été extrêmement utile chez certaines espèces, mais des limitations sont à noter. D'une part, il est difficile d'identifier un petit nombre d'individus présentant de nouveaux phénotypes au sein d'une grande population. D'autre part, la redondance génétique présente dans de nombreuses espèces végétales, due à la duplication de gènes et à la polyploidie, limite l'effet de nombreuses mutations sur la plante (Parry et al. 2009). Les mutations induites sont du même type que les mutations spontanées et dépendent essentiellement de l'agent mutagène utilisé. Elles consistent en des PNS, des délétions courtes ou longues, des insertions, mais également en des réarrangements chromosomiques ou des duplications (Bolon et al. 2014).

La limitation la plus importante de la mutagenèse induite, qu'elle soit chimique ou physique, est sa nature aléatoire. Ainsi, générer un mutant pour un caractère particulier va nécessiter la génération de grandes collections de mutants et le développement de cribles phénotypiques puissants. Un perfectionnement de la stratégie de mutagenèse induite est le TILLING<sup>5</sup> (Targeting Induced Local Lesions In Genomes) qui permet de sélectionner, dans la descendance de plantes mutagenisées, celles qui portent une mutation dans le gène désiré. Cependant, dans la plupart des cas, ces approches aboutissent à l'identification de mutations nulles<sup>6</sup> de sorte qu'obtenir un allèle particulier, connu pour conférer un phénotype donné, est, dans la plupart des cas, extrêmement difficile.

Au cours des dernières années, une nouvelle génération d'agents mutagènes, basée sur des nucléases site dirigées<sup>7</sup> (Site Directed Nucleases, SDN), comme les méganucléases, les « zinc finger nucleases », les TALEN ou plus récemment le système CRISPR-Cas9, a été développée (Fig.1) et est déjà utilisée par certains sélectionneurs (Shi et al. 2016). Comme on le verra dans le chapitre 3, les SDN créent des cassures double brin<sup>8</sup> de l'ADN (CDB) et permettent potentiellement de générer tous les types de mutations qui ont été sélectionnées lors de la

domestication et de la sélection variétale et ce avec une très grande précision. Dès maintenant, certains gènes montrés comme impliqués dans des caractères agronomiques simples, sont de bonnes cibles de ces SDN. De plus, de nouveaux allèles d'intérêt peuvent être découverts grâce à des cribles génétiques directs ou inverses. Pour les caractères complexes, l'utilisation de ces nouvelles techniques va nécessiter au préalable l'identification des gènes situés dans les régions génomiques impliquées.



	Meganucléase	ZFNs	TALENs	CRISPR
Nb de protéines nécessaires	1	2	2	1 + 1 ARN
Difficultés de production	difficile (modification de la protéine et sélection)	pas très facile (interactions complexes)	relativement facile (assemblage modulaire)	très facile (modification de la séquence ARN)
Prix (approximatif)	50000 €	5000 €	1000 €	10 €

**Figure 1 : Caractéristiques des Nucléases Site Dirigées.** Quatre classes de nucléases site dirigées sont utilisables pour réaliser la réécriture de gène. Elles possèdent des caractéristiques différentes.

1. Les **méganucléases** sont constituées d'une seule protéine capable de reconnaître une cible de 12 à 40 paires de bases (pb) de manière très spécifique. La modification du site de reconnaissance de ces méganucléases par ingénierie des protéines est complexe, ce qui explique leur coût élevé de production.
2. Les **Zinc Finger Nucleases (ZFN)**, ou nucléases à doigt de zinc, consistent en deux modules composés de domaines de reconnaissance de l'ADN de type Cys2-His2 (zinc finger) fusionnés chacun à une nucléase non spécifique (en général l'enzyme Fok1). Chaque "zinc finger" est capable de reconnaître spécifiquement 3 pb de l'ADN et, en les associant, on peut obtenir des modules reconnaissant 9 à 18 pb. Cependant les ZFNs sont délicates à concevoir car leur affinité vis-à-vis de la cible ADN ne dépend pas seulement de l'association des différents "zinc fingers" correspondant à des triplets de nucléotides.
3. Les **TALE nucleases (TALEN)** sont des nucléases chimériques dont le domaine de liaison à l'ADN dérive d'effecteurs de type activateur de transcription (TALE), une famille de protéines utilisée dans le processus d'infection par les agents pathogènes du genre *Xanthomonas*. Les TALE sont composés d'une série de répétitions en tandem (12-27). Chaque répétition contient 33-35 acides aminés tous identiques à l'exception de deux, en positions 12 et 13, appelé la Repeat Variable Diresidue (RVD). Chaque répétition est spécifique d'une paire de bases de l'ADN. Les TALEN consistent en deux TALE qui sont fusionnés chacun à une nucléase non spécifique (Fok1). Du fait de la difficulté d'assembler ces différents modules de reconnaissance de l'ADN, qui sont hautement homologues, différents protocoles ont été développés (GoldenGate® ou FLASH™).
4. Le système **CRISPR-Cas** est un système immunitaire adaptatif procaryote de type II que plusieurs bactéries utilisent pour se défendre contre les ADN viraux. Ce système a été adapté pour l'édition du génome en associant à la nucléase Cas9 un ARN (sgRNA) contenant une séquence de 20 pb homologue à la cible, fusionné à un ARN de type "trans-activating RNA" chargé de capter la protéine Cas9 pour son transfert sur l'ADN cible. Cette dernière technologie, qui ne demande que la modification du sgRNA pour pouvoir cibler une séquence de 20 pb bases dans le génome, est, de très loin, la moins onéreuse des quatre stratégies proposées et la plus simple à mettre en œuvre.

## 2- La diversité génétique capturée par les sélectionneurs à travers l'identification de QTL<sup>9</sup>.

### 2.1- De nouvelles techniques pour la recherche de QTL.

La cartographie des QTL a jusqu'à présent été très efficace pour identifier les régions génétiques liées à des caractères quantitatifs agronomiques. Elle a été soutenue par le développement récent de nouveaux moyens de génotypage à haut débit et de phénotypage. Cependant, il est encore difficile d'identifier des régions génomiques associées à des caractères contrôlés par différents paramètres. Ceci est particulièrement vrai dans le cas des stress abiotiques pour lesquels la confirmation de la valeur d'un QTL donné nécessite de nombreuses répétitions (Takeda et Matsuoka 2008). Par ailleurs, la cartographie des QTL est limitée à la détection de variants génétiques dans des populations issues de croisements bi-parentaux. Pour tenter de répondre à ces contraintes d'autres techniques de détection de QTL sont maintenant utilisées en routine pour la détection d'allèles intéressants, comme par exemple la génétique d'association (Takeda et Matsuoka 2008). La génétique d'association consiste à «associer» statistiquement, dans une population présentant une grande diversité génétique, les variations des caractères (phénotypes) aux variations du génome. Cela présente l'avantage, par rapport à la cartographie QTL classique, où, à un locus donné, seuls 2 allèles parentaux sont testés, de prendre en compte un grand nombre d'allèles (Takeda et Matsuoka 2008). Un facteur clé dans cette approche est le déséquilibre de liaison<sup>10</sup> (Linkage Disequilibrium, LD) qui détermine la densité de marqueurs nécessaires et les conditions de phénotypage. Pour des espèces comme *Arabidopsis* et le maïs, qui présentent des taux de recombinaison méiotiques élevés et une décroissance rapide du LD (<1Kb), cette approche peut être fructueuse (Yu et Buckler 2006). Par contre, pour les plantes cultivées autogames, comme le riz, le soja, le sorgho ou l'orge, pour lesquelles les valeurs de LD sont plus grandes, la résolution de cartographie d'association est moins bonne.

Une autre approche, visant à identifier des régions ayant été sélectionnées au cours du temps, est basée sur la théorie selon laquelle, après une forte sélection, les locus sélectionnés présentent une diminution locale de la diversité nucléotidique et une augmentation du LD (Takeda et Matsuoka 2008). Cette approche ne convient pas à toutes les plantes, et, dans le cas du sorgho par exemple, l'observation de 445 locus n'a pas permis d'identifier de régions présentant une marque de sélection (Hamblin et al. 2006). Par contre, ce type d'approche est particulièrement bien adapté aux espèces, comme le maïs, dont l'histoire démographique est bien établie.

### 2.2- L'exploitation des QTL identifiés nécessite des schémas de sélections qui peuvent être longs.

Au cours des deux dernières décennies des efforts importants ont été faits, chez de nombreuses espèces, afin d'identifier des QTL et des marqueurs liés pour une large gamme de caractères agronomiques. En plus des méthodes de sélection traditionnelles la sélection assistée par marqueurs (SAM) est l'une des techniques les plus utilisées pour sélectionner des QTL chez diverses espèces végétales. Une autre application, maintenant largement utilisée de ces techniques est connue sous le nom de "pyramidage" de QTL (Hospital et Charcosset 1997). Ici, la combinaison de plusieurs QTL responsables de différents caractères et identifiés dans des variétés ou espèces compatibles, sont introgressés dans la même lignée élite (Ashikari et Matsuoka 2006).

Cependant, des résultats contrastés ont été rapportés en ce qui concerne le taux de succès de l'introggression de QTL, surtout lorsque cinq ou plus de cinq QTL sont introgressés dans une lignée élite donnée (Semagn et al. 2006). Différents facteurs influencent l'efficacité de l'introggression multiple de QTL, notamment la faible reproductibilité de l'effet d'un QTL dans différents fonds génétiques, les interactions entre certains QTL ou la recombinaison entre des gènes présents dans le même segment chromosomique au sein du QTL (Semagn et al. 2010). Une autre limite de l'approche QTL est que leur introgression dans des lignées élites peut nécessiter des plans de croisement longs et coûteux. Ces limitations sont particulièrement marquées pour les arbres forestiers par exemple, dont les cycles de reproduction sont très longs (Lidder et Sonnino 2012). Les progrès technologiques, tels que la réécriture du génome, offrent donc de nouvelles perspectives pour tenter d'améliorer les stratégies de sélection.

### **2.3- Gènes sous-jacents aux QTL et autres gènes à "effet fort" : une source exceptionnelle de cibles pour l'ingénierie des génomes.**

Une fois qu'un QTL impliqué dans un caractère d'intérêt a été identifié, tous les allèles présents à ce locus doivent être caractérisés afin de rechercher des gènes polymorphes liés au phénotype souhaité. Lorsque le génome de l'espèce à laquelle appartient la plante chez qui le QTL a été identifié est séquencé et annoté, la fonction putative des gènes présents au locus peut orienter les analyses vers des gènes candidats. Cette approche est plus complexe pour les espèces dont les génomes ne sont pas encore disponibles ou pour les gènes dont la fonction est inconnue. La diminution continue du coût des marqueurs moléculaires et le développement de nouvelles techniques d'analyse des génomes comme le séquençage de nouvelle génération (Next Generation Sequencing, NGS), vont faciliter l'identification des gènes candidats chez de nombreuses espèces (Pérez-de-Castro et al. 2012 ; Lim et al. 2014). En plus des gènes associés aux QTL, et généralement plutôt à effets faibles, de nombreux gènes candidats à effets forts ont pu être identifiés lors de cribles génétiques par exemple. Récemment, des zones régulatrices présentes en amont des facteurs de transcription *SibZIP1* et *SibZIP2*, qui activent les gènes responsables de la synthèse et de l'accumulation des sucres dans le fruit de la tomate, ont été montrées comme impliquées dans la répression de ces mêmes gènes en présence de saccharose (Sagor et al. 2016). Leur élimination, par des techniques de transgénèse classique permet une augmentation de 50% de sucres dans les fruits. Un autre exemple concerne le gène *OR* dont la mutation conduit à l'accumulation de  $\beta$  carotène, précurseur de la vitamine A (Lu et al. 2006). Le gène *OR* est présent chez toutes les plantes : il est donc possible d'imaginer de faire produire, par mutation dirigée, du  $\beta$  carotène à des graines ou des fruits, qui en sont dépourvus.

Quelle que soit l'origine de la découverte du gène associé à un caractère donné la validation du gène candidat est l'étape clé de l'ensemble du processus. Pour cela, une stratégie possible consiste à identifier dans le gène candidat des variations de nucléotides (*e.g.* SNP, codons stop prématurés) qui sont absentes dans une autre variété de la même espèce qui ne présente pas le phénotype ou le caractère étudiés. Dans le cas où la comparaison avec d'autres variétés de la même espèce n'est pas possible, une analyse de l'incidence potentielle de ces variations nucléotidiques sur le gène candidat peut être effectuée dans une espèce modèle, *Arabidopsis* par exemple, chez laquelle l'analyse fonctionnelle de gènes (*i.e.* disponibilité de la séquence du génome, collection de mutants) est plus facile. Dans ce but, les techniques de transgénèse et de génie génétique peuvent aider à introduire des mutations ponctuelles dans un gène candidat. Une fois que le lien entre la variation nucléotidique et le phénotype est établi, le



même type de variation peut être introduit dans la variété d'intérêt agronomique ou la lignée élite. Dans un schéma de sélection conventionnel un minimum de cinq à six générations de rétrocroisements sera nécessaire pour introgresser le gène d'une variété donneuse vers la variété élite (Lidder et Sonnino 2012). Ce procédé nécessite beaucoup de temps et d'effort, en particulier lorsque plus d'un allèle doit être transféré. De ce point de vue la modification directe du gène concerné chez la lignée élite, via réécriture du génome, permettrait un gain de temps considérable.

#### ***2.4- Le rôle des plantes modèles dans la validation des gènes candidats.***

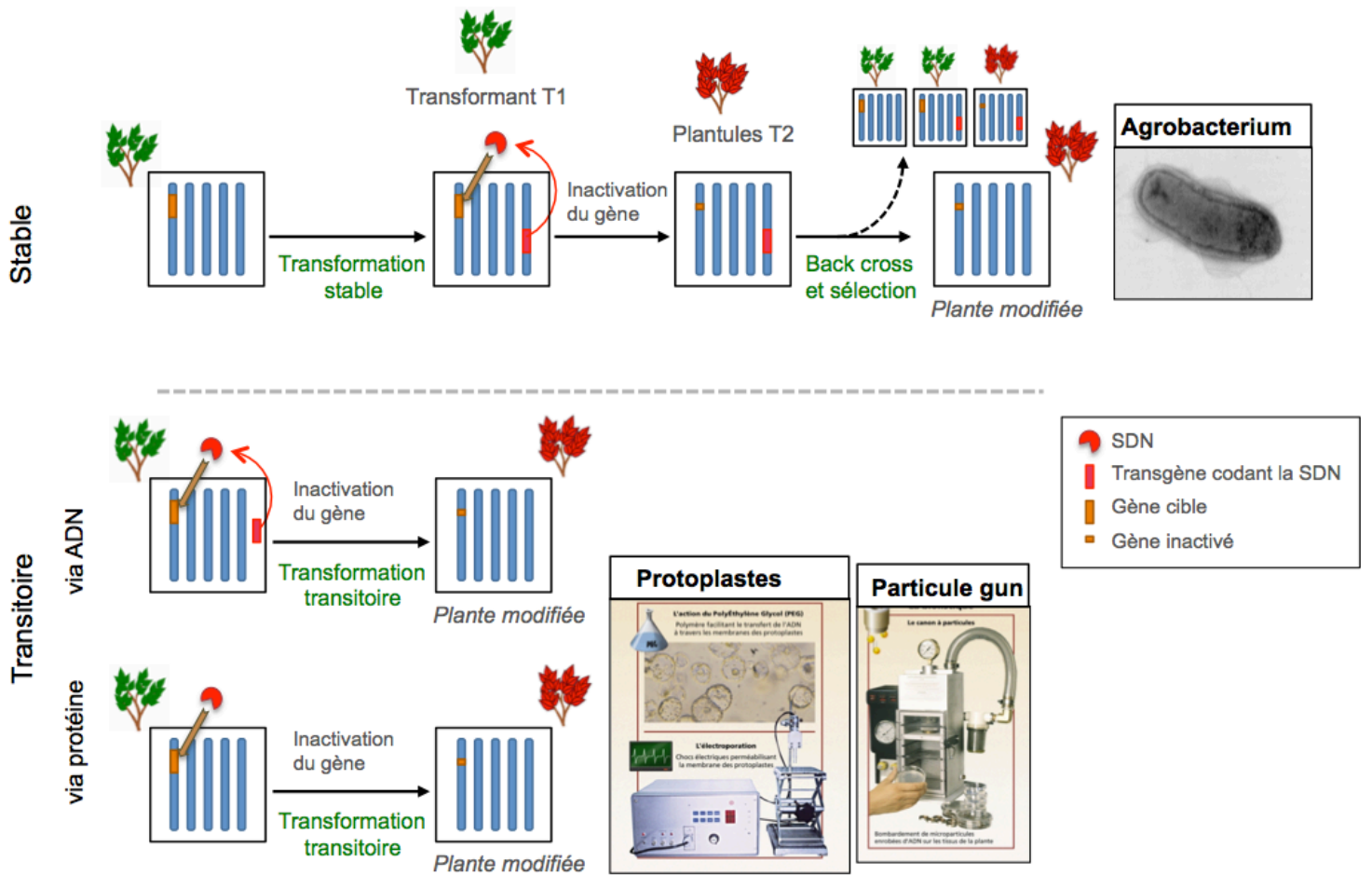
Les plantes modèles utilisées dans les laboratoires académiques présentent en général un génome bien caractérisé et de grandes ressources génétiques et moléculaires. Ces plantes sont des outils précieux pour l'analyse fonctionnelle des gènes, et la confirmation du rôle d'un gène candidat dans une espèce modèle est un bon indicateur de son implication potentielle dans une espèce cultivée apparentée. L'implication des laboratoires académiques dans ce type d'études leur permettrait d'approfondir leurs connaissances sur la génomique des plantes de grande culture et permettrait d'élargir le champ d'application de leurs découvertes. Ainsi, la coopération entre les améliorateurs, les agriculteurs et les laboratoires publics de recherche pourrait être très fructueuse pour l'innovation variétale et mériterait d'être renforcée par des programmes d'innovation appropriés.

### **3- Potentiel des nucléases site dirigées pour mettre en valeur la diversité génétique via l'ingénierie des génomes.**

#### ***3.1- Vers la modification précise des génomes de plantes d'intérêt agronomique.***

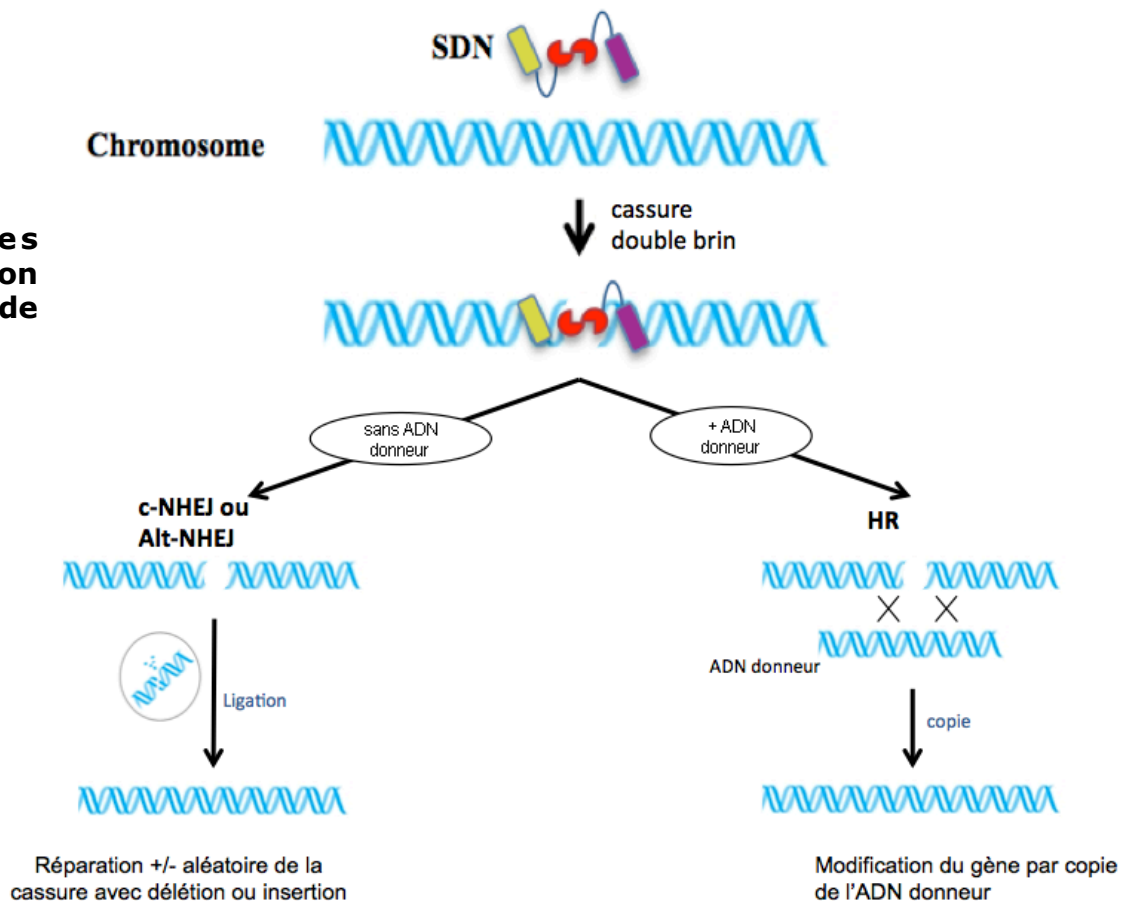
Une fois qu'un allèle responsable d'un caractère agronomique particulier a été identifié, son transfert à une variété cultivée peut parfois être complexe. De ce point de vue, l'ingénierie des génomes peut fournir des moyens efficaces pour insérer ou modifier ce gène dans les variétés cultivées. Grâce aux nucléases site dirigées (SDN) il est maintenant en effet possible de retoucher un allèle à un locus précis et ce avec des conséquences très limitées sur le reste du génome. Pour cela la SDN sera délivrée dans le noyau de la cellule végétale soit par production de transformants stables via la transformation par *Agrobacterium*, soit par expression transitoire via la transformation de protoplastes ou le bombardement de cals (Fig.2). La cassure de l'ADN induite par la nucléase va ensuite être prise en charge par un système endogène (Fig.3) de réparation des cassures de l'ADN qui procède par jonction des extrémités non homologues<sup>11</sup> (Non-Homologous End-Joining, NHEJ) et aboutir à des mutations aléatoires du gène ciblé (délétions, insertions, substitutions). De plus, en ciblant un lieu présentant des micro-homologies autour de la coupure, celle-ci peut être réparée par la voie du NHEJ alternatif (Alt-NHEJ) et donc être plus prédictible (Butler et al. 2015 ; Collonnier et al. 2016). Si, lors de la cassure de l'ADN générée par la nucléase une matrice de réparation est apportée à la cellule, cette cassure peut alors être réparée de manière fidèle (par copie de la matrice de réparation) via la voie de la recombinaison homologue<sup>12</sup> (RH) et aboutir au remplacement du gène ciblé.

En plus de l'introduction de mutations ciblées dans un gène particulier, les SDN peuvent être utilisées pour induire des mutations dans plusieurs gènes différents simultanément. Ce type d'application peut être particulièrement adapté à la modification de plusieurs gènes impliqués dans un caractère quantitatif particulier. Pour cela, plusieurs SDN ciblant des gènes différents peuvent être co-délivrés à la cellule (Ma et al. 2015 ; Shan et al. 2015 ; Lowder et al. 2015). On



**Figure 2 : Modalités d'introduction des nucléases site dirigées par transgénése ou par expression transitoire.** Pour que les nucléases site dirigées puissent induire des cassures double brin de l'ADN et générer une mutagénèse ciblée dans les plantes elles peuvent être transférées dans la cellule et son noyau par production de transformants stables (transformation via *Agrobacterium*) ou par expression transitoire (transformation de protoplastes ou bombardement de cals).

**Figure 3 : Systèmes endogènes de réparation des cassures induites de l'ADN.**



peut également envisager, à l'aide d'une même SDN, de cibler une séquence présente dans un groupe de gènes. Cette approche est très utile pour éliminer les séquences dupliquées en tandem<sup>13</sup> qui prévalent dans les génomes eucaryotes supérieurs (Christian et al. 2013 ; Qi et al. 2013 ; Zhou et al. 2014). L'obtention simultanée de mutations dans différents gènes est particulièrement intéressante pour la création de nouveaux phénotypes contrôlés par des gènes homologues (issus d'une duplication) chez les espèces diploïdes et par des gènes homéologues chez les espèces polyploïdes (Wang et al. 2014). Les SDN sont donc particulièrement intéressantes pour modifier des génomes végétaux complexes et apporter des caractères d'intérêt agronomique chez les plantes cultivées qui présentent souvent un haut niveau de ploïdie et des génomes de grande taille (Gil-Humanes et Voytas 2014).

### **3.2- Stratégies d'extinction et de remplacement de gènes et innovation variétale.**

L'extinction (knock-out) d'un gène candidat, *via* la réparation imprécise d'une cassure induite de l'ADN via une SDN, offre beaucoup de possibilités pour la création de caractères agronomiques d'intérêt. Par exemple, chez la pomme de terre, les quatre allèles (plante tétraploïde) du gène *Vlnv* (vacuolar invertase) ont pu être inactivés à l'aide de TALEN. Les plantes ainsi produites accumulent moins de sucres réducteurs dans les tubercules pendant le stockage au froid ce qui limite la production d'acrylamide pendant la cuisson (Clasen et al. 2016). D'autres caractères de qualité ont pu être obtenus en utilisant le même type de stratégie. Un riz parfumé a ainsi été créé via l'inactivation du gène *OsBADH2* (betaine aldehyde dehydrogenase homolog) (Shan et al. 2015). Des tomates présentant un mûrissement retardé ont été produites par inactivation, via l'utilisation de la stratégie CRISPR-Cas9, du gène *RIN* (ripening inhibitor) (Ito et al. 2015). Une lignée de soja avec une meilleure qualité d'huile a été obtenue en inactivant, *via* l'utilisation de TALEN, les deux gènes *FAD2* (fatty acid desaturase 2) présents chez cette espèce (Haun et al. 2014). Modifier les gènes de susceptibilité chez une plante hôte est également un bon moyen de développer de nouvelles sources de résistance aux maladies. Ainsi, dans le blé tendre, la mutation des trois gènes homéologues (le blé est hexaploïde) *MLO* (Mildew-Resistance Locus) *via* l'utilisation de TALEN, permet la résistance à l'oïdium (Wang et al. 2014). Les TALEN et les CRISPR ont également été utilisés pour générer une résistance à la bactériose vasculaire chez le riz en modifiant les sites de liaison à l'ADN de l'hôte des facteurs de transcription produits par la bactérie (*Xanthomonas*) et impliqués dans l'infection (Li et al. 2012 ; Zhou et al. 2015a).

Par conséquent, dès lors qu'un gène candidat a été identifié et que l'allèle d'intérêt agronomique correspond à une version inactive de ce gène, d'autres caractères, comme la tolérance aux stress abiotiques (e.g. sécheresse, carence en azote), la résistance aux maladies ou une meilleure qualité nutritionnelle (e.g. diminution de composés antinutritionnels, des acides gras meilleurs pour la santé), pourraient être à portée de main *via* la stratégie d'extinction de gène. Cependant, un allèle d'intérêt associé à un caractère agronomique particulier correspond bien souvent à une version modifiée du gène plutôt qu'à une version inactivée et, dans ce cadre, le remplacement précis (knock-in) d'un gène candidat *via* la recombinaison homologue, offre des possibilités encore plus grandes en terme d'amélioration variétale que la simple extinction de gène.

Cependant, un allèle associé à un caractère agronomique particulier correspond bien souvent à une version modifiée du gène plutôt qu'à une version inactivée et, dans ce cadre, le

remplacement précis (knock-in) d'un gène candidat, *via* la recombinaison homologue, offre des possibilités encore plus grandes en terme d'amélioration variétale que la simple extinction de gènes. Jusqu'à présent, très peu de caractères agronomiques ont été obtenus *via* le remplacement/modification ciblée de gènes. Ceci est dû à la faible efficacité de la stratégie de modification de gène comparée à l'efficacité d'inactivation. C'est pourquoi la majorité des exemples correspondant à des modifications de gènes correspondent à des caractères facilement sélectionnables, comme la résistance à un herbicide par exemple. C'est le cas chez le tabac pour lequel la résistance aux herbicides de type sulfonylurée a été apportée par la modification ciblée, *via* l'utilisation de SDN de type Zinc Finger Nucleases, du gène *AHAS* (Acetohydroxyacid synthase) (Townsend et al. 2009).

### ***3.3 Quel avenir pour les technologies de réécriture du génome ?***

L'association de régions particulières du génome à des traits particuliers repose sur l'utilisation de marqueurs identifiés lors de la détection de QTL. La sélection assistée par marqueurs utilise ensuite ces marqueurs pour prédire la valeur génétique des individus obtenus dans les croisements successifs du programme de sélection (Das et Rao 2015 ; Hasan et al. 2015 ; Pradhan et al. 2015). Comme proposé précédemment, la technologie de réécriture du génome pourrait permettre d'accélérer le processus d'association des gènes aux QTL qui sont normalement suivis par des méthodes de sélection traditionnelles ou par la SAM. L'utilisation de ce type de technologie serait particulièrement adaptée à l'association de gènes liés génétiquement (Flavell 2010). De plus, grâce à la possibilité de réécriture simultanée de plusieurs gènes, l'association des différents allèles pourrait se faire sur un nombre limité de générations qui dépendrait du nombre de locus à modifier (Farré et al. 2014). Il faut cependant noter que pour des caractères quantitatifs, dépendant de nombreux gènes ayant de petits effets, si les gènes candidats sont transférés dans des lignées élites par hybridation ou par l'ingénierie des génomes, la valeur du phénotype peut ne pas être exactement celle attendue, en partie en raison de la modification des fonds génétiques et de la perte d'interactions génétiques pertinentes (Takeda et Matsuoka 2008). Donc, s'il est difficilement imaginable qu'à court ou moyen terme les outils de réécriture du génome remplacent totalement les méthodes de sélection traditionnelles, de sélection génomique ou la SAM dans les schémas de sélection variétale, cette technologie devrait cependant rapidement s'imposer comme un outil supplémentaire et parfois incontournable du sélectionneur.

### ***3.4 La transformation génétique des lignées élites, point clé de l'utilisation des techniques de modification des génomes.***

Les techniques de transgénèse sont généralement employées chez des variétés ayant des capacités particulières en termes de transformation et de régénération alors que les lignées élites correspondantes sont en général récalcitrantes. L'optimisation et l'industrialisation de certaines étapes critiques de ces protocoles de transgénèse permettent dans certains cas d'accélérer les processus de sélection en transformant directement les lignées élites. Pour ces plantes, les gènes candidats peuvent être modifiés directement. Pour les autres, la modification du génome repose sur la mise au point de méthodes de transformation et de régénération fiables et si possible indépendantes du génotype utilisé. La plupart du temps, la régénération est le goulot d'étranglement et une meilleure compréhension des mécanismes sous-jacents

serait déterminante. Pour le moment, des solutions plus immédiates consistent à essayer d'augmenter l'efficacité de la transformation et la fréquence de mutation. Une voie prometteuse consiste à utiliser des vecteurs d'expression des SDN basés sur des virus végétaux (Butler et al. 2015 ; Čermák et al. 2015 ; Yin et al. 2015). Cette approche présente également l'avantage de produire des plantes modifiées sans intégration stable du transgène codant la SDN qui n'est exprimée que transitoirement dans la cellule. Le fait qu'il ne soit pas nécessaire d'éliminer le transgène par rétrocroisement et ségrégation est particulièrement intéressant pour les espèces présentant des degrés élevés d'hétérozygotie ou de longs stades juvéniles, comme les arbres fruitiers par exemple (Ilardi et Tavazza 2015). Cela est également vrai pour les espèces qui sont multipliées végétativement pour fixer leur type variétal complexe qu'on perdrait par reproduction sexuée. Des fréquences de mutation très efficaces ont également été obtenues pour certaines espèces en transfectant la cellule directement par la nucléases sous forme protéique (Luo et al. 2015 ; Woo et al. 2015). Enfin, l'efficacité de la mutagenèse peut être optimisée en utilisant des promoteurs spécifiques favorisant l'expression des gènes codants les SDN dans les cellules reproductrices (Wang et al. 2015).

### ***3.5. Les SDN, une opportunité pour les espèces n'ayant pas profité de la sélection variétale ?***

Les nouveaux outils de réécriture du génome pourraient contribuer à soutenir l'amélioration variétale des cultures orphelines qui sont d'une importance économique clé dans certains pays en voie de développement. Ces plantes cultivées, n'ayant pas ou peu fait l'objet d'un travail de sélection variétale, regroupent des espèces telles que le mil, l'arachide, le niébé, le pois chiche, pois d'Angole, le manioc, l'igname et la patate douce (Varshney et al. 2012). Elles ne bénéficient pas de grands investissements des instituts de recherche publics ni des sélectionneurs. En conséquence, très peu de données génomiques sont en général disponibles sur ces plantes, ce qui rend difficile la caractérisation de régions génomiques associées à des caractères intéressants. De plus, les approches de type SAM peuvent être rendues difficiles du fait de l'utilisation de variétés locales ou sauvages comme source d'allèles favorables. En permettant une modification précise de certaines régions génomiques et l'introgression ciblée d'allèles candidats potentiels, pouvant provenir d'autres plantes, la réécriture du génome de ces plantes orphelines via les SDN pourrait potentiellement aider à lever certaines de ces difficultés.

### ***3.6. La recherche de gènes candidats dans les espèces sauvages.***

Comme dit précédemment, la recherche de QTL et l'identification des gènes associés est une source majeure d'information pour l'élaboration de stratégies SDN. La plupart du temps, le choix des parents utilisés pour l'identification de QTL est dirigé par la capacité des lignées parentales à être croisées avec des lignées élites. L'utilisation de parents sauvages ou de variétés locales est une source originale de diversité génétique (Tanksley et McCouch 1997 ; Zhou et al. 2015b). Cependant, les croisements ne sont pas toujours faciles, ni même possibles entre les variétés domestiques (élites) et leurs équivalents sauvages. Cette barrière de compatibilité sexuelle restreint le pool génétique accessible aux sélectionneurs (Michelmore 2003). En revanche, avec la technologie SDN, la seule limitation est la présence du gène recherché dans un organisme donneur. Ce donneur peut être un parent sauvage, mais également une espèce plus éloignée, voire une espèce non végétale (une bactérie par exemple).

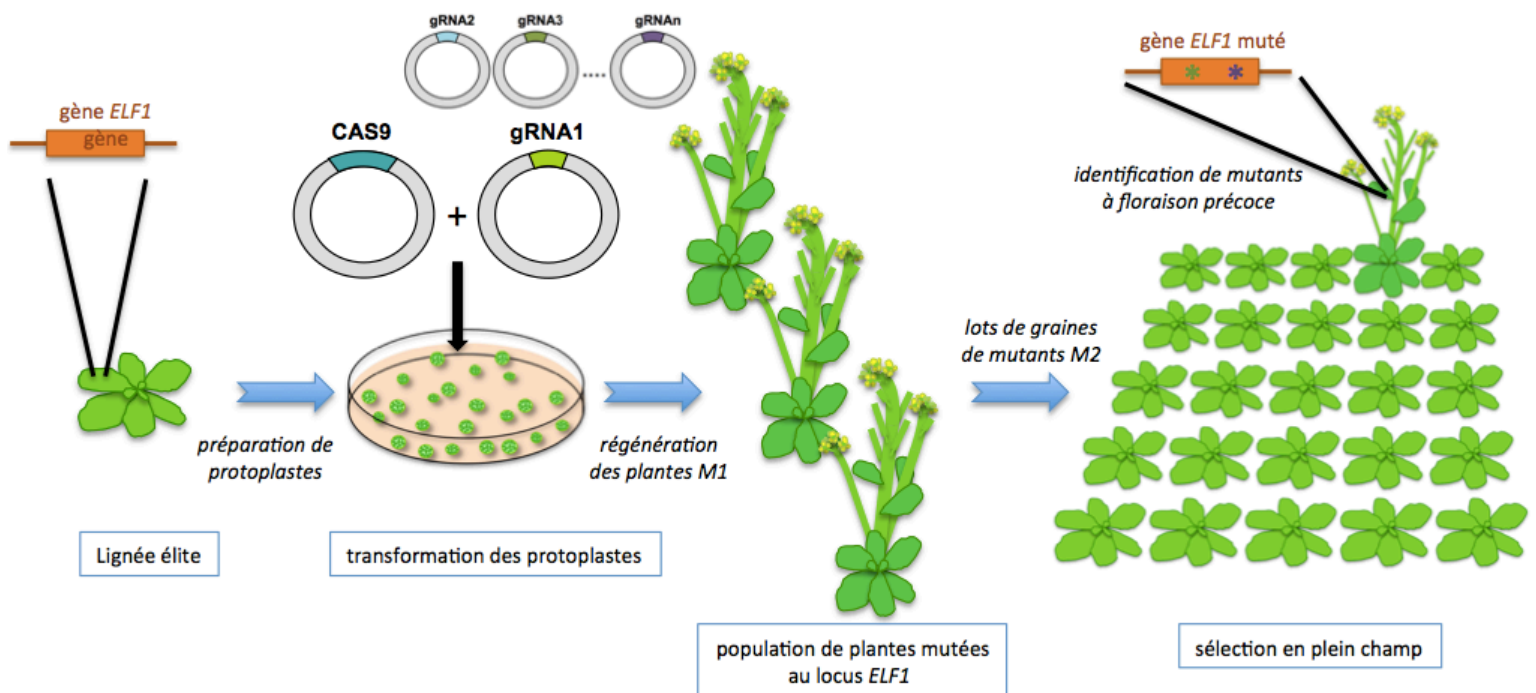
#### 4- Optimisation génomique *via* l'utilisation des nucléases site dirigées, pour aller plus loin que la sélection végétale classique.

##### 4.1. *Inventer de nouveaux gènes candidats.*

Comme dit précédemment, la recherche de nouveaux allèles et leur association est à la base de l'amélioration des plantes. Cependant, cette recherche d'allèles, à l'heure actuelle limitée à des variations spontanées, pourrait également comprendre la conception d'allèles artificiels ou issus de variations génétiques induites, conduisant à la sélection d'allèles encore non détectés dans la diversité génétique naturelle existante. En plus de la recherche classique d'allèles en utilisant la diversité génétique spontanée, plusieurs technologies peuvent être utilisées pour l'amélioration d'un gène donné, y compris la conception de nouveaux allèles basés sur l'analyse structurale de protéines (Lutz 2011), sur la mutagenèse aléatoire et enfin sur la stratégie plus récente d'évolution dirigée. L'évolution dirigée est un ensemble de technologies (PCR mutagène, mutagenèse à saturation, brassage d'ADN, métagénomique) qui exploitent les mécanismes de l'évolution pour guider une protéine vers une fonction choisie en identifiant des allèles dans des systèmes microbiens (Minshull et Stemmer 2001). Ces stratégies ont été très utiles pour améliorer ou modifier l'activité de biomolécules à usages industriels ou thérapeutiques (Packer et Liu 2015) et, bien qu'elles soient encore peu utilisées dans le domaine végétal, elles présentent un gros potentiel pour l'optimisation de certains gènes présentant un intérêt agronomique. Cependant, nous devons garder à l'esprit que ces stratégies ne sont possibles qu'en système microbien et ne seront donc valables que pour des enzymes. De plus, l'optimisation réalisée dans un système bactérien ne sera pas nécessairement confirmée dans le système végétal receveur.

##### 4.2. *La variation génétique induite ciblée (TIGV, Targeted Induced Genetic Variation), une alternative au TILLING ?*

Les mutations, en combinaison avec la recombinaison, sont les sources à partir desquelles les sélectionneurs sont capables de produire de nouvelles variétés. Comme indiqué précédemment, une façon d'augmenter la variabilité génétique disponible pour la sélection est d'induire des mutations et de les caractériser en utilisant la technique de TILLING. Cependant, l'utilisation dans les schémas de sélection de l'allèle nouvellement produit est encore un long processus et nécessite de nombreux cycles de croisements pour son introgression dans une lignée élite. Une façon d'éviter ou de limiter ces étapes pourrait être l'utilisation de nucléases site dirigées pour induire de manière ciblée de la variation génétique (TIGV) (Fig.4). Des cibles possibles de la TIGV sont les gènes de résistance de la famille *NBS-LRR*, pour lesquels des recombinaisons et des échanges de séquence permettent de diversifier les sites de reconnaissance des éliciteurs des pathogènes. La TIGV pourrait potentiellement stimuler ces échanges de séquences (Joshi et Nayak 2013). Enfin, l'utilisation de la TIGV sur des gènes différents au sein de la même plante permettrait la production d'une population de mutants qui contiendraient différentes combinaisons d'allèles à des locus différents. La production de ces populations mutagenisées de façon ciblée et leur criblage dans différentes conditions de contraintes biotiques ou abiotiques représenteraient une solution très innovante pour aborder la question importante des stress multiples et interactifs chez les plantes (Kissoudis et al. 2014). Néanmoins, la stratégie TIGV présente des limites, notamment la capacité à transformer efficacement des lignées élites.



**Figure 4 : La variation génétique induite ciblée (TIGV, Targeted induced genetic variation).** Le but de la TIGV est de produire de la variabilité génétique à un ou plusieurs locus dans le génome et de sélectionner dans la population ainsi générée des individus qui auront acquis, au niveau des locus ciblés, des allèles plus favorables pour un ou plusieurs caractères agronomiques. Dans le cas présenté ici, le caractère désiré est une lignée de colza à cycle court. Pour cela, le gène *ELF1*, qui est impliqué dans la transition florale chez *Arabidopsis*, est ciblé en utilisant la stratégie CRISPR-Cas9. Des plantes de type sauvage sont, de manière stable ou de manière transitoire (dans ce cas par fusion PEG), transformées et des plantes M1 sont générées et autofécondées. Les graines mutantes M2 sont ensuite recueillies et semées dans des conditions de culture de plein champ pour la sélection de plantes à floraison précoce.

#### 4.3. Remodelage du génome des plantes cultivées par nucléases site dirigées.

L'utilisation des nucléases site dirigées ne se limite pas à la modification de séquences codantes mais peut également être utilisée pour supprimer de grands segments chromosomiques (jusqu'à 170 kb), ou des séquences cibles répétées (Endo et al. 2014 ; Zhou et al. 2014). Comme dit précédemment, les structures génomiques comme les insertions, délétions, duplications, inversions et translocations, font partie, au même titre que le polymorphisme nucléotidique, des différences génomiques présentes au sein d'une même espèce. Ces polymorphismes de grande taille sont nommés variants structuraux (VS) (Saxena et al. 2014). Certains VS ont été clairement associés à des phénotypes spécifiques, comme par exemple la tolérance à l'aluminium chez le maïs qui est due à une augmentation du nombre de copies du gène *MATE1* (Maron et al. 2013). Cependant, les VS ne sont pas toujours associés à un gain de fonction et certains d'entre eux peuvent être préjudiciables à la plante. Par exemple, le génome du bananier est envahi par de nombreuses séquences de badnavirus parmi lesquelles certaines peuvent, suite à un stress, libérer des pararétrovirus infectieux (Chabannes et al. 2013). Ce "réveil" potentiel des virus est devenu le principal obstacle à des programmes de sélection de la banane. Une façon de se débarrasser de ce risque pourrait consister à éliminer ou rendre inactive les séquences virales dans le génome du bananier en utilisant des SDN.

#### **4.4. Imiter le système immunitaire bactérien CRISPR-Cas9 chez les plantes.**

Certains de ces nouveaux outils de l'ingénierie des génomes peuvent aussi être utilisés à d'autres fins que la modification du génome. Le système CRISPR-Cas bactérien est un système immunitaire adaptatif pour lutter contre l'invasion d'ADN étranger (phages ou plasmides) ; il pourrait être comparé à la voie des ARN interférents présente chez les eucaryotes, même si ces deux systèmes ne sont pas homologues (Marraffini et Sontheimer 2010). Récemment, différents groupes ont essayé de transposer ce système immunitaire des bactéries aux plantes. Pour cela, la nucléase Cas9 et des ARN guides ciblant différentes séquences dans le génome d'un géminivirus, ont été exprimées chez *Nicotiana benthamiana* et *Arabidopsis* (Ali et al. 2015 ; Baltés et al. 2015 ; Ji et al. 2015). Les auteurs ont pu montrer l'induction de mutations dans la séquence de l'ADN viral et une diminution du nombre de copies du virus. Cette stratégie est donc très prometteuse du fait de sa flexibilité et de sa facilité de mise en œuvre, mais son efficacité pour lutter contre les virus doit encore être démontrée dans des conditions de culture en champ. De plus, comme dans toutes les stratégies de lutte contre les virus, la stratégie CRISPR-Cas9 pourrait faire face à l'apparition de résistances spontanées, soit à la suite de mutations dans les séquences ciblées par l'ARN guide, soit à la suite de la production par le virus de suppresseurs du système CRISPR-Cas9. Cette dernière situation a déjà été observées pour certains bactériophages (van der Oost et Brouns 2015).

#### **4.5. La stratégie fondée sur les nucléases site dirigées pour contrôler la recombinaison méiotique.**

L'amélioration des espèces cultivées repose sur la possibilité de produire de nouvelles combinaisons d'allèles. Toutefois, si les sélectionneurs peuvent désormais identifier plus facilement les bons allèles, ils sont encore limités dans leur capacité à les combiner au sein de la plante élite. En effet, cela dépend du nombre et de la localisation des événements de recombinaison qui se produisent entre les chromosomes homologues au cours de la méiose. Cela est particulièrement vrai lorsque les parents du croisement appartiennent à différentes espèces. Dans ce cas, la divergence des séquences homéologues empêche la recombinaison durant la méiose (Canady et al. 2006). Par conséquent, la possibilité de maîtriser les niveaux de recombinaison lors des croisements serait un outil formidable pour contrôler la combinaison des régions chromosomiques contenant les allèles favorables. Une étape clé de la recombinaison méiotique est la formation de cassures double brin dans les chromosomes. Ces cassures vont définir les sites d'échange génétique mutuel, les crossing-overs. Dans le but de contrôler la recombinaison méiotique dans les plantes cultivées, on pourrait envisager de cibler les cassures double brin méiotiques dans des régions spécifiques afin de provoquer des crossing-overs là où l'échange génétique est nécessaire (Wijnker et de Jong 2008). La protéine SPO11 et ses partenaires sont des acteurs clés de la cassure de l'ADN programmée lors de la méiose. Dans la levure, il a été montré qu'une fusion de la protéine SPO11 au domaine de liaison à l'ADN du facteur de transcription Gal4 peut stimuler la formation de la cassure double brin, et la recombinaison génétique associée, à proximité des sites de liaison du facteur Gal4 (Peciña et al. 2002). Sur la base de ces données, il serait tentant de tester chez les plantes si les fusions de SPO11 avec le domaine de liaison à l'ADN d'une nucléase Cas9 désactivée associée à des ARN guides spécifiques, pourraient cibler les cassures double brin méiotiques dans des régions spécifiques du génome. Un tel outil serait très utile aux sélectionneurs pour accélérer l'association des allèles favorables.



## Conclusion

L'évolution des plantes repose sur des mutations spontanées du génome entraînant potentiellement de nouveaux caractères qui sont fixés par la sélection naturelle. La sélection variétale repose également sur la variabilité génétique spontanée qui a pu être artificiellement augmentée à l'aide de techniques de mutagenèse aléatoire. Les techniques de réécriture du génome fournissent maintenant des moyens pour introduire presque tout type de mutation et de réarrangement des chromosomes, et ce d'une manière très précise et rapide. Cela doit non seulement permettre aux sélectionneurs d'accélérer les programmes de sélection variétale, mais également d'offrir une gamme presque illimitée de combinaisons de nouveaux allèles tout en s'affranchissant des barrières de compatibilité sexuelles. Si la technique de remplacement de gènes reste encore difficile et doit être améliorée, la technique d'extinction de gène, elle, est dès maintenant utilisable en routine chez un grand nombre de plantes cultivées et devrait être intégrée relativement rapidement dans des programmes de sélection variétale.

Du point de vue scientifique, les points les plus limitants sont l'identification précise des gènes impliqués dans les caractères étudiés et les capacités de transformation des plantes élites cultivées. Cependant, la plus grande incertitude concernant l'avenir de cette technologie pour la sélection variétale est le statut juridique qu'auront les plantes résultant de son utilisation. Si ces plantes devaient être soumises à la même législation que celles issues de la transgénèse classique (OGM), les coûts et les délais des processus d'approbation bloqueraient probablement l'utilisation de ces techniques pour la plupart des plantes et des caractères agronomiques. Ceci rendrait impossible pour la plupart des sélectionneurs (du secteur privé ou des instituts publics), de s'engager dans le développement de nouvelles variétés utilisant la réécriture du génome, du moins en Europe.

## Bibliographie :

- Ali Z, Abulfaraj A, Idris A, et al (2015) CRISPR/Cas9-mediated viral interference in plants. *Genome Biol* 16:238. doi: 10.1186/s13059-015-0799-6
- Ashikari M, Matsuoka M (2006) Identification, isolation and pyramiding of quantitative trait loci for rice breeding. *Trends Plant Sci* 11:344-50. doi: 10.1016/j.tplants.2006.05.008
- Baltes NJ, Hummel AW, Konecna E, et al (2015) Conferring resistance to geminiviruses with the CRISPR-Cas prokaryotic immune system. *Nat Plants* 1:15145. doi: 10.1038/nplants.2015.145
- Blackman BK, Strasburg JL, Raduski AR, et al (2010) The Role of Recently Derived FT Paralogs in Sunflower Domestication. *Curr Biol* 20:629-635. doi: 10.1016/j.cub.2010.01.059
- Bolon Y-T, Stec AO, Michno J-M, et al (2014) Genome resilience and prevalence of segmental duplications following fast neutron irradiation of soybean. *Genetics* 198:967-81. doi: 10.1534/genetics.114.170340
- Butler NM, Atkins PA, Voytas DF, Douches DS (2015) Generation and Inheritance of Targeted Mutations in Potato (*Solanum tuberosum* L.) Using the CRISPR/Cas System. *PLoS One* 10:e0144591. doi: 10.1371/journal.pone.0144591
- Canady MA, Ji YF, Chetelat RT (2006) Homeologous recombination in *Solanum lycopersicoides* introgression lines of cultivated tomato. *Genetics* 174:1775-1788. doi: 10.1534/genetics.106.065144
- Čermák T, Baltes NJ, Čegan R, et al (2015) High-frequency, precise modification of the tomato genome. *Genome Biol* 16:232. doi: 10.1186/s13059-015-0796-9
- Chabannes M, Baurens F-C, Duroy P-O, et al (2013) Three infectious viral species lying in wait in the banana genome. *J Virol* 87:8624-37. doi: 10.1128/JVI.00899-13

- Christian M, Qi Y, Zhang Y, Voytas DF (2013) Targeted mutagenesis of *Arabidopsis thaliana* using engineered TAL effector nucleases. *G3 (Bethesda)* 3:1697-705. doi: 10.1534/g3.113.007104
- Clasen BM, Stoddard TJ, Luo S, et al (2016) Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnol J* 14:169-76. doi: 10.1111/pbi.12370
- Collonnier C, Epert A, Mara K, et al (2016) CRISPR-Cas9 mediated efficient directed mutagenesis and RAD51-dependent and -independent gene targeting in the moss *Physcomitrella patens*. *Plant Biotechnol J* 1-10. doi: 10.1111/pbi.12596
- Das G, Rao GJN (2015) Molecular marker assisted gene stacking for biotic and abiotic stress resistance genes in an elite rice cultivar. *Front Plant Sci* 6:1-18. doi: 10.3389/fpls.2015.00698
- Doebley JF, Gaut BS, Smith BD (2006) The Molecular Genetics of Crop Domestication. *Cell* 127:1309-1321. doi: 10.1016/j.cell.2006.12.006
- Endo M, Mikami M, Toki S (2014) Multigene Knockout Utilizing Off-Target Mutations of the CRISPR/Cas9 System in Rice. *Plant Cell Physiol* 56:41-47. doi: 10.1093/pcp/pcu154
- Farré G, Blancquaert D, Capell T, et al (2014) Engineering Complex Metabolic Pathways in Plants. *Annu Rev Plant Biol* 65:187-223. doi: 10.1146/annurev-arplant-050213-035825
- Flavell R (2010) From genomics to crop breeding. *Nat Biotechnol* 28:144-145. doi: 10.1038/nbt0210-144
- Gaut BS (2015) Evolution Is an Experiment: Assessing Parallelism in Crop Domestication and Experimental Evolution. *Mol Biol Evol* 32:1661-1671. doi: 10.1093/molbev/msv105
- Gil-Humanes J, Voytas DF (2014) Wheat rescued from fungal disease. *Nat Biotechnol* 32:886-7. doi: 10.1038/nbt.3013
- Hamblin MT, Casa AM, Sun H, et al (2006) Challenges of detecting directional selection after a bottleneck: lessons from *Sorghum bicolor*. *Genetics* 173:953-64. doi: 10.1534/genetics.105.054312
- Hasan MM, Rafii MY, Ismail MR, et al (2015) Marker-assisted backcrossing: a useful method for rice improvement. *Biotechnol Biotechnol Equip* 29:237-254. doi: 10.1080/13102818.2014.995920
- Haun W, Coffman A, Clasen BM, et al (2014) Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnol J* 12:1-7. doi: 10.1111/pbi.12201
- Hospital F, Charcosset A. 1997. Marker-Assisted Introgression of Quantitative Trait Loci. *Genetics* 147.
- Ilardi V, Tavazza M (2015) Biotechnological strategies and tools for Plum pox virus resistance: trans-, intra-, cis-genesis, and beyond. *Front Plant Sci* 6:379. doi: 10.3389/fpls.2015.00379
- Ito Y, Nishizawa-Yokoi A, Endo M, et al (2015) CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochem Biophys Res Commun* 467:76-82. doi: 10.1016/j.bbrc.2015.09.117
- Ji X, Zhang H, Zhang Y, et al (2015) Establishing a CRISPR-Cas-like immune system conferring DNA virus resistance in plants. *Nat Plants* 1:15144. doi: 10.1038/nplants.2015.144
- Jin J, Huang W, Gao J-P, et al (2008) Genetic control of rice plant architecture under domestication. *Nat Genet* 40:1365-9. doi: 10.1038/ng.247
- Joshi RK, Nayak S (2013) Perspectives of genomic diversification and molecular recombination towards R-gene evolution in plants. *Physiol Mol Biol Plants* 19:1-9. doi: 10.1007/s12298-012-0138-2
- Kissoudis C, van de Wiel C, Visser RGF, van der Linden G (2014) Enhancing crop resilience to combined abiotic and biotic stress through the dissection of physiological and molecular crosstalk. *Front Plant Sci* 5:1-20. doi: 10.3389/fpls.2014.00207
- Li T, Liu B, Spalding MH, et al (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotechnol* 30:390-2. doi: 10.1038/nbt.2199
- Lidder P, Sonnino A (2012) Biotechnologies for the management of genetic resources for food and agriculture. *Adv Genet* 78:1-167. doi: 10.1016/B978-0-12-394394-1.00001-8
- Lim J-H, Yang H-J, Jung K-H, et al (2014) Quantitative trait locus mapping and candidate gene analysis for plant architecture traits using whole genome re-sequencing in rice. *Mol Cells* 37:149-60. doi: 10.14348/molcells.2014.2336
- Lisch D (2013) How important are transposons for plant evolution? *Nat Rev Genet* 14:49-61. doi: 10.1038/nrg3374

- Lowder LG, Zhang D, Baltus NJ, et al (2015) A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiol* 169:971-85. doi: 10.1104/pp.15.00636
- Lu S, Van Eck J, Zhou X, et al (2006) The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. *Plant Cell* 18:3594-605. doi: 10.1105/tpc.106.046417
- Luo S, Li J, Stoddard TJ, et al (2015) Non-transgenic Plant Genome Editing Using Purified Sequence-Specific Nucleases. *Mol Plant* 8:1425-1427. doi: 10.1016/j.molp.2015.05.012
- Lutz S (2011) Beyond directed evolution - semi-rational protein engineering and design. *Curr Opin Biotechnol* 21:734-743. doi: 10.1016/j.copbio.2010.08.011.Beyond
- Ma X, Zhang Q, Zhu Q, et al (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol Plant* 1-11. doi: 10.1016/j.molp.2015.04.007
- Maron LG, Guimarães CT, Kirst M, et al (2013) Aluminum tolerance in maize is associated with higher MATE1 gene copy number. *Proc Natl Acad Sci U S A* 110:5241-6. doi: 10.1073/pnas.1220766110
- Marraffini LA, Sontheimer EJ (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 11:181-90. doi: 10.1038/nrg2749
- Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. *Nat Rev Genet* 14:840-852. doi: 10.1038/nrg3605
- Michelmore RW (2003) The impact zone: genomics and breeding for durable disease resistance. *Curr Opin Plant Biol* 6:397-404. doi: 10.1016/S1369-5266(03)00067-0
- Minshull J, Stemmer WPC (2001) Molecular Breeding: the Natural Approach To. *Adv Protein Chem* 55:261-292. doi: [http://dx.doi.org/10.1016/S0065-3233\(01\)55006-8](http://dx.doi.org/10.1016/S0065-3233(01)55006-8)
- Muller HJ (1928) The production of mutations by X-Rays. *Proc Natl Acad Sci* 14:714-726. doi: 10.1126/science.68.1751.59
- Murat F, Van De Peer Y, Salse J (2012) Decoding plant and animal genome plasticity from differential paleo-evolutionary patterns and processes. *Genome Biol Evol* 4:917-928. doi: 10.1093/gbe/evs066
- Olsen KM, Wendel JF (2013a) Crop plants as models for understanding plant adaptation and diversification. *Front Plant Sci*. doi: 10.3389/fpls.2013.00290
- Olsen KM, Wendel JF (2013b) A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol* 64:47-70. doi: 10.1146/annurev-arplant-050312-120048
- Ossowski S, Schneeberger K, Lucas-Lledó JI, et al (2010) The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* 327:92-4. doi: 10.1126/science.1180677
- Packer MS, Liu DR (2015) Methods for the directed evolution of proteins. *Nat Rev Genet* 16:379-394. doi: 10.1038/nrg3927
- Parry MAJ, Madgwick PJ, Bayon C, et al (2009) Mutation discovery for crop improvement. *J Exp Bot* 60:2817-2825. doi: 10.1093/jxb/erp189
- Peciña A, Smith KN, Mézard C, et al (2002) Targeted stimulation of meiotic recombination. *Cell* 111:173-84.
- Pérez-de-Castro AM, Vilanova S, Cañizares J, et al (2012) Application of genomic tools in plant breeding. *Curr Genomics* 13:179-95. doi: 10.2174/138920212800543084
- Pradhan SK, Nayak DK, Mohanty S, et al (2015) Pyramiding of three bacterial blight resistance genes for broad-spectrum resistance in deepwater rice variety, Jalmagna. *Rice (N Y)* 8:51. doi: 10.1186/s12284-015-0051-8
- Qi Y, Li X, Zhang Y, et al (2013) Targeted deletion and inversion of tandemly arrayed genes in *Arabidopsis thaliana* using zinc finger nucleases. *G3 (Bethesda)* 3:1707-15. doi: 10.1534/g3.113.006270
- Sagor GHM, Berberich T, Tanaka S, et al (2016) A novel strategy to produce sweeter tomato fruits with high sugar contents by fruit-specific expression of a single bZIP transcription factor gene. *Plant Biotechnol J* 14:1116-1126. doi: 10.1111/pbi.12480
- Saxena RK, Edwards D, Varshney RK (2014) Structural variations in plant genomes. *Brief Funct Genomics* 13:296-307. doi: 10.1093/bfgp/elu016
- Schmutz J, McClean PE, Mamidi S, et al (2014) A reference genome for common bean and genome-wide analysis of

- dual domestications. *Nat Genet* 46:707-713. doi: 10.1038/ng.3008
- Semagn K, Bjørnstad Å, Ndjiondjop MN (2006) Progress and prospects of marker assisted backcrossing as a tool in crop breeding programs. *J Biotechnol* 5:2588-2603.
- Semagn K, Bjørnstad Å, Xu Y (2010) The genetic dissection of quantitative traits in crops. *Electron J Biotechnol* 13:0-0. doi: 10.2225/vol13-issue5-fulltext-14
- Shan Q, Zhang Y, Chen K, et al (2015) Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. *Plant Biotechnol J* 13:791-800. doi: 10.1111/pbi.12312
- Shi J, Gao H, Wang H, et al (2016) ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol J*. doi: 10.1111/pbi.12603
- Stadler LJ (1928) Mutations in barley induced by X-rays and radiu. *Science* (80- ) 68:186-187. doi: 10.1126/science.68.1756.186
- Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. *Nat Genet* 43:1160-1163.
- Takeda S, Matsuoka M (2008) Genetic approaches to crop improvement: responding to environmental and population changes. *Nat Rev Genet* 9:444-457. doi: 10.1038/nrg2342
- Tan L, Li X, Liu F, et al (2008) Control of a key transition from prostrate to erect growth in rice domestication. *Nat Genet* 40:1360-1364. doi: 10.1038/ng.197
- Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277:1063-1066. doi: 10.1126/science.277.5329.1063
- Townsend J a, Wright D a, Winfrey RJ, et al (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature* 459:442-5. doi: 10.1038/nature07845
- van der Oost J, Brouns SJJ (2015) CRISPR sabotage. *Genome Biol* 16:248. doi: 10.1186/s13059-015-0820-0
- Varshney RK, Ribaut J-M, Buckler ES, et al (2012) Can genomics boost productivity of orphan crops? *Nat Biotechnol* 30:1172-1176. doi: 10.1038/nbt.2440
- Wang RL, Stec a, Hey J, et al (1999) The limits of selection during maize domestication. *Nature* 398:236-239. doi: 10.1038/18435
- Wang Y, Cheng X, Shan Q, et al (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol Advance on*:1-6. doi: 10.1038/nbt.2969
- Wang Z-P, Xing H, Dong L, et al (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. *Genome Biol* 16:144. doi: 10.1186/s13059-015-0715-0
- Wendel JF (2015) The wondrous cycles of polyploidy in plants. *Am J Bot* . doi: 10.3732/ajb.1500320
- Wijnker E, de Jong H (2008) Managing meiotic recombination in plant breeding. *Trends Plant Sci* 13:640-646. doi: 10.1016/j.tplants.2008.09.004
- Woo JW, Kim J, Kwon S Il, et al (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol* 33:1162-1164. doi: 10.1038/nbt.3389
- Xiao H, Jiang N, Schaffner E, et al (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* 319:1527-1530. doi: 10.1126/science.1153040
- Yin K, Han T, Liu G, et al (2015) A geminivirus-based guide RNA delivery system for CRISPR/Cas9 mediated plant genome editing. *Sci Rep* 5:14926. doi: 10.1038/srep14926
- Yu J, Buckler ES (2006) Genetic association mapping and genome organization of maize. *Curr Opin Biotechnol* 17:155-60. doi: 10.1016/j.copbio.2006.02.003
- Zhou H, Liu B, Weeks DP, et al (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res* 42:10903-14. doi: 10.1093/nar/gku806
- Zhou J, Peng Z, Long J, et al (2015a) Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J* 82:632-43. doi: 10.1111/tpj.12838
- Zhou Z, Jiang Y, Wang Z, et al (2015b) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat Biotechnol* 33:408-414. doi: 10.1038/nbt.3096