

L'édition génomique ou la réécriture des génomes

Depuis une dizaine d'années se sont développées les méthodes de ciblage des mutations par des nucléases et deux méthodes sont actuellement utilisées, sous les acronymes TALEN (*Transcription activator-like effector nucleases*) et CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*, ou « courtes répétitions en palindrome regroupées et régulièrement espacées »), cette dernière de plus en plus largement et bénéficiant de multiples développements techniques. Par ces méthodes, une séquence donnée du génome reconnue par les acteurs moléculaires mis en jeu, subit une coupure des deux brins d'ADN, ce qui permet alors trois types d'interventions sur ce génome (1, 2 et 3 ci-dessous). Dans la méthode TALEN, la séquence cible est reconnue par les domaines de liaison à l'ADN de protéines chimériques associées à une nucléase qui réalise la coupure. Dans la méthode CRISPR Cas, la reconnaissance se fait par liaison d'un ARN guide à l'un des deux brins de la séquence cible et la coupure est réalisée par la nucléase Cas (Cas 9 ou d'autres découvertes depuis), qui est associée à l'ARN guide.

1. On réalise la coupure des deux brins de la molécule d'ADN en un site déterminé. La réparation par la machinerie cellulaire de cette coupure, préférentiellement par les voies de réparation de type "jonction des extrémités", peut aboutir à une insertion ou une délétion de nucléotides dont la conséquence, recherchée dans ce cas, peut être l'inactivation du gène. Type d'intervention dite SDN1 (pour Site Directed Nuclease), premier cas d'édition (réécriture) génomique dans lequel il s'agit d'une inactivation ciblée.
2. On réalise la coupure en présence d'une matrice d'ADN homologue de la séquence ciblée, mais différente par une ou plusieurs mutations. Dans ce cas la coupure peut être réparée par la voie de la recombinaison homologue qui, par copie de cette matrice, introduira ces mutations dans le génome lors de la réparation. Type d'intervention dit SDN2, deuxième cas de réécriture génomique, qui peut s'extrapoler au remplacement d'un allèle par un autre présent dans l'espèce ou une espèce sexuellement compatible.
3. On réalise la coupure en présence d'une matrice d'ADN portant une séquence homologue à la séquence ciblée mais interrompue par un transgène, gène étranger au génome de l'espèce. Là encore, la coupure peut être réparée par la voie de la recombinaison homologue qui va intégrer le transgène dans le génome lors de la réparation par copie de la matrice d'ADN. Type d'intervention dit SDN3 ou transgénèse ciblée.

Les réécritures génomiques SDN1 et SDN2 consistent donc à modifier par mutation, la séquence d'un gène résident pour générer un nouvel allèle de ce gène et apporter le nouveau caractère associé. Cet allèle peut pré-exister dans l'espèce ou une espèce sexuellement compatible et la réécriture génomique dans ce cas est l'équivalent de l'introduction par croisement d'un allèle dans un génotype (introgression) ou il peut être non identifié dans l'espèce et dans ce cas la réécriture génomique est l'équivalent d'une mutation spontanée. La transgénèse SDN3 qui consiste à introduire un transgène

en un site choisi à l'avance du génome est un perfectionnement de la transgénèse qui peut présenter un intérêt en termes d'expression et d'association des transgènes.

Une autre méthode de réécriture dérivée des propriétés du système CRISPR Cas a fait son apparition récemment. Il s'agit de la production d'une protéine de fusion entre la nucléase Cas9, dont on a inactivé la fonction nucléase, avec une cytidine désaminase ou une adénine désaminase. Ces protéines hybrides, guidées comme précédemment par un ARN, peuvent modifier chimiquement respectivement une cytidine (C) de l'ADN la transformant en uridine (U), ou une adénine en hypoxanthine (H). Le résultat final est une transition ciblée d'une paire de base GC en AT ou réciproquement. A noter que la transition GC vers AT se produit majoritairement parmi les mutations induites par l'EMS, mais aléatoirement et en grand nombre dans le génome. Il est probable qu'à terme toutes les substitutions de nucléotides seront réalisables par de telles protéines chimériques.

Application à la sélection des variétés végétales

Plus que l'existence d'une variation génétique donnée conduisant à un nouveau caractère, la problématique de l'amélioration des plantes réside dans sa sélection. Cette variation génétique peut être le résultat d'une combinaison d'allèles ou le fait d'un seul gène (caractère mono-génique). Dans le premier cas la sélection génomique, à condition que les allèles ou leurs locus dans le génome soient identifiés, est une nouvelle approche pour apporter une réponse. Dans le second cas, le taux de mutation spontané (10^{-8} – 10^{-9} par paire de base de l'ADN et par génération) peut laisser penser, qu'en probabilité, toute variation mono-génique souhaitée « existe dans la nature ». Mais existence ne veut pas dire disponibilité et la question est alors de savoir comment la découvrir.

La sélection est rapide pour un caractère soumis à une forte sélection : des populations d'adventices résistantes à un herbicide apparaissent régulièrement sans qu'elles aient recours à une quelconque biotechnologie ! Il n'en va pas de même pour un caractère moins facile à observer ou quantifier, comme la présence ou l'absence d'un métabolite, la capacité d'assimilation d'un élément minéral, la composition d'une graine : ces caractères qu'ils soient mono ou polygéniques ne sont perceptibles que par l'analyse et la sélection devient alors très lourde voire matériellement impossible.

Avant la réécriture

L'idée de « *diriger la création des variétés nouvelles* » n'est pas récente et a été l'objectif récurrent de tous les perfectionnements conceptuels et techniques de l'amélioration des plantes depuis deux siècles. Parmi les techniques mises en œuvre pour des caractères mono-géniques, l'induction de mutations a pour objectif de rendre accessible la sélection par l'augmentation de la fréquence des variations. De façon très empirique dès 1907, quelques années seulement après la découverte de la radioactivité du radium (rayons gamma) et alors qu'émergeait le concept de gène, l'effet des rayonnements sur les plantes fut expérimenté. C'est en 1927 que la création de mutations héréditaires par irradiation fut démontrée chez la drosophile et l'orge ouvrant la voie à un usage systématique de cette méthode. La première variété végétale dérivée de mutagenèse par irradiation fut cultivée à partir de 1934 (tabac en Indonésie), et une deuxième en 1948 (cotonnier en Inde). La mutagenèse induite se généralisa après la seconde guerre mondiale avec le développement des applications civiles et médicales des rayonnements et la découverte de la structure chimique des gènes (de l'ADN) et des substances qui peuvent en modifier la séquence. Elle est ainsi devenue un outil banal dans la sélection d'un grand nombre d'espèces végétales. Environ 3 220 variétés de plus de 170 espèces issues directement de traitements mutagènes, essentiellement par rayonnement, ont été déclarées avant 2015 (FAO/IAEA) et un nombre indéterminé en dérivent après croisement.

La connaissance des séquences des gènes puis des génomes a conduit à des développements de méthodes et d'outils pour isoler des mutations dans des gènes connus pour leurs fonctions et par leurs séquences. Au début des années 2000 la méthode de « TILLING » a été développée, en particulier en recherche fondamentale mais également en sélection chez le riz, le blé, des graminées fourragères. Cette méthode s'appuie sur la création, par traitement de graines le plus souvent, d'un nombre élevé de mutations (chaque gène est muté au moins une fois ; on parle de mutation « à saturation ») parmi quelques milliers d'individus. La constitution de pools ordonnés et emboîtés de familles de descendants dont l'ADN est analysé permet de remonter à la famille issue de la graine qui contient la mutation désirée. Il s'agit bien d'une mutagenèse ciblée, mais a posteriori. La détection des mutations repose sur des endonucléases reconnaissant des hétéro-duplex entre les brins d'ADN du gène d'origine et ceux du gène muté. Les techniques de séquençage progressant, le TILLING par séquençage est désormais de plus en plus pratiqué.

Les prérequis techniques de la réécriture génomique pour les plantes

Sauf cas particulier, ces méthodes nécessitent la culture *in vitro* de cellules ou d'organes. Elles font appel à la transfection, transitoire ou stable (transgénèse) ou à l'expression transitoire (sans intégration dans le génome) dans la cellule végétale d'un ADN comprenant un gène codant les protéines TALEN ou, dans le cas du CRISPR-Cas, de l'endonucléase Cas (fusionnée ou non à une désaminase) et de la ou les séquences qui seront transcrites en ARN guides. Il est parfois possible de faire pénétrer dans des cellules les protéines et les ARN nécessaires à la réécriture sans avoir besoin d'exprimer les constructions génétiques correspondantes. De plus dans le cas d'un échange d'allèle (type SDN2), la matrice d'ADN nécessaire à cet échange doit être fournie également à la cellule. Les transgènes exprimés transitoirement ne se répliquant pas sont éliminés au cours des divisions cellulaires. Les transgènes intégrés dans le génome sont éliminés par ségrégation génétique dans la descendance sexuée.

Il est à noter que de la fréquence des modifications désirées dépend la faisabilité de la réécriture. Ceci est particulièrement vrai dans le cas de SDN2, où la sélection peut être problématique. Aussi le coût de mise en œuvre de ces techniques dépend essentiellement de la maîtrise des méthodes de régénération de plantes entières à partir de cellules ou tissus cultivés *in vitro* et de sélection des modifications recherchées, le volet de biologie moléculaire étant relativement accessoire. Les méthodes de réécriture génomique sont à la portée de structures de sélection qui possèdent des savoir-faire de culture *in vitro* et de biologie moléculaire que l'on peut trouver dans certaines petites ou moyennes entreprises. Elles supposent aussi certaines connaissances en génomique provenant de la littérature scientifique comme de recherches particulières, propres à l'entreprise.

En comparaison avec les précédentes méthodes

L'aléatoire qui prévaut dans l'innovation variétale quand il s'agit d'obtenir le caractère souhaité par les approches traditionnelles comme la prospection de ressources génétiques ou la mutagenèse induite, se trouve particulièrement réduit dans le cas de la réécriture génomique qui implique la connaissance des séquences et des fonctions des gènes. Pour illustrer cette précision, citons l'exemple de la production de beta-carotène dans des tissus non chlorophylliens comme des tubercules, par la simple modification d'une paire de base du gène « Or » qui substitue une histidine à l'arginine de la position 108 de la protéine correspondante.

Cette connaissance a priori se double d'un ciblage exclusif de la modification génétique produite qui n'existe pas dans les méthodes traditionnelles.

La mutagenèse aléatoire induite introduit de 5 à 10 mutations par millions de paires de bases, donc plusieurs milliers dans un génome. Elles sont à mettre en comparaison avec les quelques mutations hors cibles éventuelles en réécriture dont le nombre est proche des mutations spontanées qui apparaissent à chaque génération. Ces mutations hors cibles, plus rares chez les végétaux que chez les animaux, ne sont pas aléatoires et se produisent dans des séquences présentant certaines homologies avec la cible choisie pour la réécriture. Ces mutations hors cibles peuvent être limitées par le choix de l'ARN guide utilisé, par la maîtrise technique : expression transitoire des facteurs de réécriture plutôt que transgénèse, nouvelles générations de nucléases. Enfin, quand elles existent, elles peuvent être, contrairement aux mutations spontanées ou induites, facilement détectées et éliminées par ségrégation génétique dans la descendance sexuée.

Une source majeure de diversification des variétés réside dans des hybridations entre des types différents suivies de nombreuses générations de sélection et d'homogénéisation. L'hybridation dans l'espèce ou entre espèces sexuellement compatibles pour introduire un nouveau caractère s'accompagne de l'introduction de milliers d'allèles que les croisements en retour n'élimineront que partiellement. Par exemple le génome des tomates cultivées contient en moyenne 10% de gènes issus d'espèces sauvages, parfois non comestibles, introduits par les sélectionneurs en même temps que des caractères de qualité du fruit ou de résistance à des pathogènes.

Par ailleurs on reproduit par voie végétative certaines espèces (la vigne, la pomme de terre, des fruitiers, des plantes ornementales etc.) afin d'en conserver les caractéristiques qui se perdraient par reproduction sexuée. Aussi il n'est pas possible d'introduire un nouveau caractère par hybridation dans un cultivar d'une telle espèce, ce que permettent la réécriture génomique, la cisgénèse ou la transgénèse.

L'introduction d'un caractère (un nouvel allèle), que ce soit par voie sexuée ou par mutagenèse induite, provoque non seulement de nouvelles interactions entre le génome et cet allèle (au moins théoriquement) mais aussi un nombre incalculable de nouvelles interactions avec les milliers de mutations ou de nouveaux allèles mis en jeu par ces méthodes. La sélection a pour rôle non seulement de conserver les caractères favorables mais également de minimiser les conséquences phénotypiques qui peuvent résulter de ces interactions. De ce point de vue la réécriture génomique qui n'implique qu'une seule, ou un nombre limité de modifications connues, minimise ces effets phénotypiques dits d'épistasie et permet au sélectionneur d'en cerner les causalités. La sélection consciente d'un nouveau génotype est une intervention humaine séculaire qui soulève objectivement plus d'incertitudes que la réécriture génomique.