



Rôle de l'épigénome dans le contrôle du développement et des interactions avec le milieu. Quels apports pour l'amélioration des plantes ?

Hervé Vaucheret

Directeur de Recherche, Institut Jean-Pierre Bourgin, INRA Versailles

Manuscrit révisé le 21 septembre 2012 - Publié le 28 octobre 2013

Résumé : Pendant longtemps, l'ADN (le support de l'hérédité) et les protéines (les acteurs de la machinerie cellulaire) ont été considérés comme les plus importantes molécules de la cellule et l'ARN comme un simple intermédiaire. La découverte récente de différents types d'ARN ne codant pas des protéines mais exerçant des rôles régulateurs a changé notre vision de l'expression génique. On sait maintenant que des grands ARN non-codants ainsi que des petits ARN de 20 à 30 nucléotides jouent un rôle essentiel dans la stabilité des génomes ainsi que dans le contrôle du développement chez tous les eucaryotes. En particulier, les petits ARN régulent l'expression de nombreux gènes endogènes au cours du développement et en réponse à divers stress, contrôlent le mouvement des éléments transposables, participent à la structuration de la chromatine, et protègent les cellules contre des agents infectieux (virus, bactéries). Chez les plantes, une modification du répertoire des petits ARN peut provoquer des modifications héritables de l'expression des gènes sans que la séquence d'ADN soit altérée. On parle alors de modifications épigénétiques. Cette hérédité consécutive à des changements, même transitoires, du répertoire des petits ARNs participe peut-être à l'importante plasticité phénotypique des plantes. En effet, celles-ci sont sans cesse soumises à des modifications importantes de leur environnement et ne peuvent y échapper autrement qu'en s'adaptant. Identifier les gènes endogènes susceptibles de subir des modifications épigénétiques héritables constitue donc une étape importante dans l'établissement d'une carte des réseaux de régulations épigénétiques, utile aux agronomes et aux sélectionneurs/améliorateurs des plantes.

Introduction

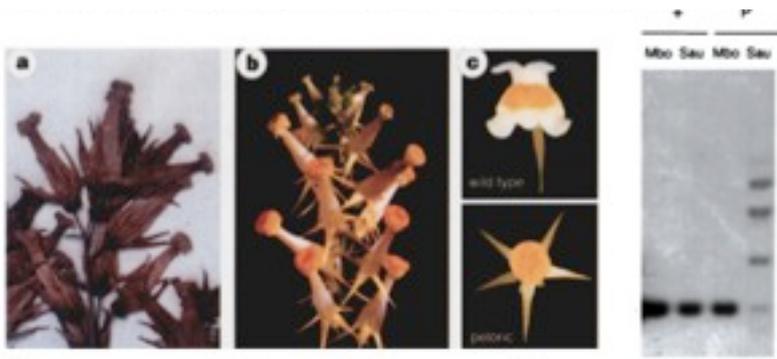
La génétique est la discipline qui étudie la transmission des caractères. On appelle allèles les différentes séquences d'ADN existant pour un même locus dans une population, la séquence la plus représentée définissant l'allèle sauvage, les autres correspondant à des mutations. Une mutation dans la séquence codante d'un gène peut affecter la séquence de la protéine, et donc potentiellement son activité, tandis qu'une mutation dans les séquences régulatrices du gène peut modifier son expression spatio-temporelle. On peut également observer des changements d'expression génique dus à la méthylation de l'ADN ou à l'acétylation, la méthylation, la phosphorylation ou l'ubiquitylation des histones associées à l'ADN sans qu'il y ait changement de la séquence de l'ADN. On parle alors d'épimutations. Ces changements d'expression sans changement de la séquence d'ADN peuvent être hérités à la méiose ; l'épigénétique est la discipline qui étudie leur transmission. Pour un gène donné, les épimutations surviennent généralement à des fréquences supérieures à celles des mutations (de l'ordre de 10^{-7} pour les mutations, jusqu'à 10% pour les épimutations). Elles révertent également à des fréquences supérieures (10^{-7} pour les mutations, jusqu'à 100% pour les épimutations). Les épimutations ont

donc très vite été considérées comme un outil particulièrement efficace utilisé par les organismes pour s'adapter rapidement aux contraintes environnementales tout en gardant la possibilité de revenir à leur forme originale. Toutefois, la démonstration d'un avantage sélectif des épimutations par rapport aux mutations n'a pas encore été faite, certainement en raison du faible nombre d'épimutations stables recensées à ce jour.

Les plantes sont sessiles et doivent donc s'adapter rapidement aux changements d'environnement puisqu'elles ne peuvent y échapper. Par ailleurs, elles développent tardivement un méristème reproductif qui donnera naissance aux gamètes. Les plantes peuvent donc transmettre à leur descendance des informations génomiques nouvelles, acquises durant la phase de développement végétatif. Enfin, de très nombreuses plantes s'autofécondent, ce qui permet en une seule génération de voir se manifester des caractères récessifs. Les plantes réunissent donc les meilleurs atouts pour étudier l'apparition, le devenir et le rôle des épimutations.

Epimutations et paramutations spontanées

Plusieurs épimutations ont été mises en évidence dans des populations naturelles de plantes (pour revue (1)). Elles correspondent toutes à des modifications morphologiques visibles à l'œil nu, et on ne peut pour l'instant pas dire si elles présentent un avantage sélectif aux plantes qui les portent. Ainsi, chez la linéaire, une plante herbacée à souche rhizomateuse, le variant « *peloric* » a été décrit pour la première fois par Linné en 1749, mais sa nature épigénétique n'a été élucidée qu'en 1999 (2). Le phénotype *peloric* est du à l'absence de la protéine codée par le gène *Cycloidea* (*LCYC*). Toutefois, le gène *LCYC* ne porte aucune mutation chez le variant *peloric*. En revanche, l'ADN de son promoteur est méthylé, ce qui empêche sa transcription.



Le promoteur du gène *Cycloidea* est hyperméthylé



Des réversions somatiques surviennent à faible fréquence !

Figure 1 : Phénotype des fleurs de plantes sauvages (wildtype) et *peloric*, ainsi que les phénotypes de réversions partielles. Les profils de méthylation de l'ADN des plantes sauvages et *peloric* ainsi que des révertants sont montrés sur la partie droite.

Cette méthylation est transmise à la méiose depuis plus de 250 ans. Toutefois, des pertes de méthylation spontanées sont observées lors de certaines mitoses, conduisant à la réexpression du gène *LCYC* et à des réversions phénotypiques (Figure 1).

L'apparition d'épimutations stables corrélées à une méthylation de l'ADN a également été observée pour le gène *CNR* de la tomate (3), le gène *WIP1* du melon (4), et les gènes *b1*, *pl1* et *r1* du maïs (pour revue (5)), sans que l'on comprenne jamais quel est l'évènement qui a induit l'apparition d'épimutations. Dans le cas des gènes *b1*, *pl1* et *r1* du maïs, les épimutations sont capables de se transmettre entre allèles à l'issue d'un croisement. On parle alors de paramutation (ou conversion épigénétique). Lors de ce processus,

l'allèle épimuté (désigné sous le terme d'allèle paramutateur) transmet son état épigénétique à un allèle exprimé (désigné sous le terme d'allèle paramutable) dans la première génération (F1) du croisement. Une fois épimuté, l'allèle paramutable devient à son tour paramutateur, c'est-à-dire capable de transmettre l'état épigénétique qu'il vient d'acquérir à un allèle exprimé introduit dans un nouveau croisement (Figure 2).

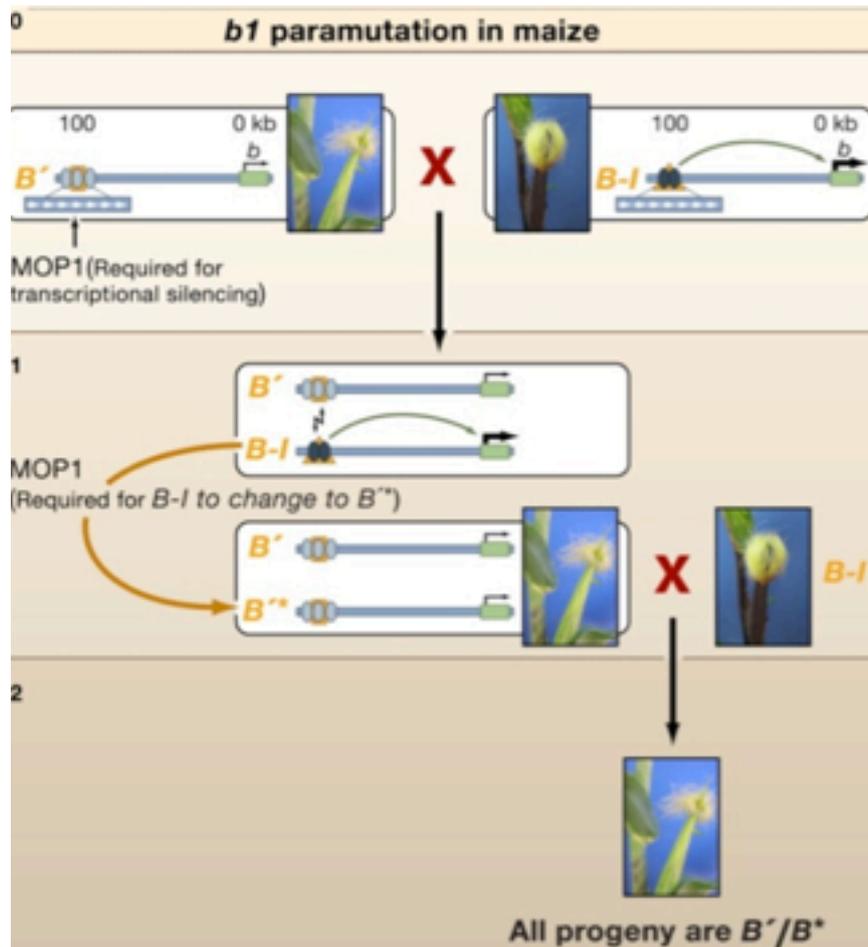


Figure 2 : Chez le maïs, le locus *b1* code un facteur de transcription qui permet la synthèse de pigments rouges (anthocyanes). Les maïs possédant l'allèle *B-I*, abondamment transcrit, accumulent des anthocyanes (à droite), alors que les anthocyanes sont absentes chez les maïs possédant l'allèle *B'* faiblement transcrit (à gauche). Une séquence d'ADN unique au génome du maïs (motif de 853 paires de bases répété 7 fois) est présente dans les deux allèles, mais elle est méthylée uniquement chez l'allèle *B'* (triangle jaune). Cette séquence contient une région activatrice ("enhancer") qui induit l'expression du gène *b1* lorsqu'elle n'est pas méthylée et qu'elle se lie à des protéines spécifiques (cercles bleu foncé ou clair). A l'issue d'un croisement entre *B-I* et *B'*, l'allèle *B-I* devient méthylé et inactif comme l'allèle *B'*, via un processus qui requiert l'ARN polymérase MOP1. L'allèle *B-I* est converti en allèle *B'* qui devient à son tour capable de convertir un allèle *B-I* en l'allèle *B'*. Cette modification épigénétique se transmet aux générations suivantes de façon invasive.

Planche extraite de Chandler, 2007 (21).

Petits ARN et modifications de la chromatine

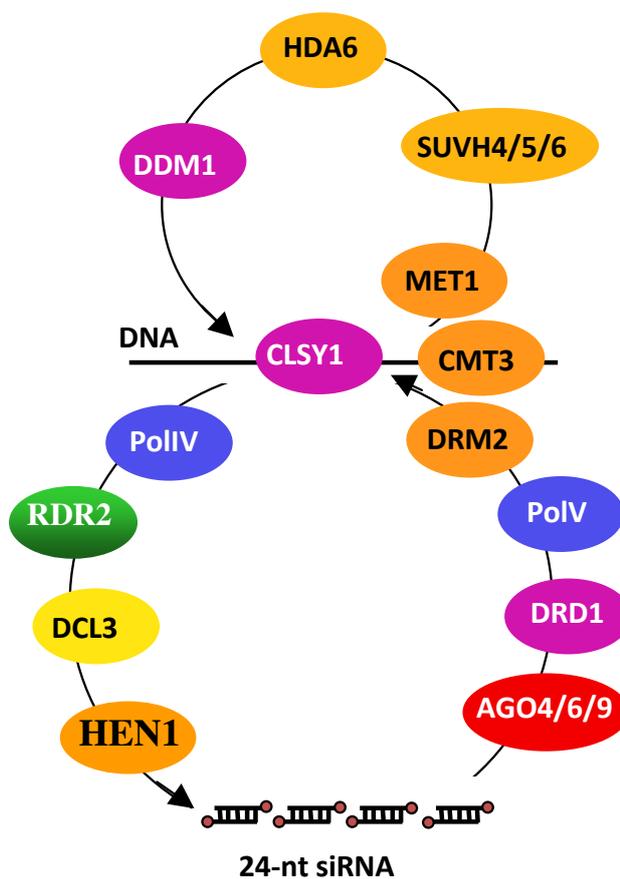
Les séquences des promoteurs des gènes capables de paramutation expriment des petits ARN interférents ("short interfering RNA" ou siRNA) de 24 nucléotides (5). Les siRNA de 24-nt constituent la catégorie la plus abondante de petits ARN chez les plantes (plus d'un million de petits ARN différents ont été identifiés chez *Arabidopsis*, dont 90% sont des siRNA de 24-nt qui couvrent un bon tiers du génome). Les siRNA de 24-nt se distinguent des microRNA (miRNA) et des siRNA de 21-nt par leur biogénèse, par les protéines avec lesquelles ils s'associent et par les mécanismes d'inactivation qu'ils engendrent (6, 7). Ainsi, les miRNA et des siRNA de 21-nt

guident le clivage des ARN messagers (mRNA) homologues ou le blocage de leur traduction. On parle alors d'inactivation post-transcriptionnelle ou PTGS (pour "post-transcriptional gene silencing"). Celle-ci est réversible et disparaît en l'absence du miRNA ou siRNA. A l'inverse, les siRNA de 24-nt guident la méthylation de l'ADN et/ou la méthylation, l'acétylation, la phosphorylation ou l'ubiquitylation des histones, ce qui a conduit à un remodelage la chromatine qui rend les séquences promotrices inaccessibles aux facteurs de transcription et/ou à la RNA Polymérase II. On parle alors d'inactivation transcriptionnelle ou TGS (pour "transcriptional gene silencing") (8, 9). Celle-ci est potentiellement réversible, mais peut parfois perdurer après disparition des siRNAs qui l'ont induite. En effet, le maintien de ces marques épigénétiques (méthylation de l'ADN, modifications des histones) est assuré par deux boucles de régulation, l'une basée sur la production de siRNA permettant la mise en place *de novo* des marques épigénétiques (mais aussi leur renforcement à chaque cycle), l'autre sur l'entretien de ces marques par des enzymes de maintenance ne nécessitant pas de siRNA pour les guider (8, 9). C'est le cas par exemple de l'ADN méthyltransférases MET1 (Figure 3).

Figure 3 : Boucles de maintenance des marques épigénétiques chez les plantes.

Une **première boucle** met en jeu la transcription des loci méthylés par les ARN polymérase Pol IV et Pol V qui sont spécifiques aux plantes. Grâce à la protéine de remodelage de la chromatine par **CLSY1**, Pol IV produit des ARN simple-brin qui sont ensuite convertis en ARN double-brin par RDR2, découpés en siRNA de 24-nt par **DCL3**, méthylés par **HEN1**, associés via **AGO4/6/9** aux ARN transcrits par **Pol V** de façon à guider la méthylation *de novo* de l'ADN par **DRM2** et le remodelage de la chromatine par **DRD1**.

La **seconde boucle** met en jeu la protéine de remodelage de la chromatine par **DDM1** et des enzymes modifcatrices des histones (**HDA6** et **SUVH4/5/6**) qui favorisent le maintien de la méthylation de l'ADN par les enzymes **MET1** et **CMT3**.



Epimutations induites artificiellement dans des fonds génétiques mutants

Des mutations dans les gènes contrôlant la boucle de TGS impliquant les siRNA conduisent à des diminutions locales de la méthylation de l'ADN (8, 9). Toutefois, ces changements ne conduisent pas à des défauts de développement importants dans des conditions normales de croissance, même après plusieurs générations, certainement en raison de la boucle de maintenance. En revanche, de telles mutations empêchent la reconnaissance des séquences étrangères et leur inactivation par TGS (8), et favorisent la réactivation des éléments transposables induits en conditions de stress (10).

Des mutations dans les gènes contrôlant la boucle d'entretien de ces marques conduisent quant à elles à des modifications importantes de la méthylation de l'ADN et du statut des histones, qui se traduisent par des changements importants de l'expression de nombreux gènes et des défauts de développement (1). De plus, certains défauts de développement peuvent être hérités après élimination des mutations qui ont conduit à leur apparition, indiquant qu'il n'est pas toujours possible de rétablir complètement les marques épigénétiques originelles. Ainsi, chez la plante modèle *Arabidopsis*, les mutants *ddm1* et *met1* accumulent des épimutations (Figure 4), et certaines d'entre elles ne sont pas corrigées après croisement avec des plantes

Epimutation dominante :
Hypométhylation et sur-expression

Epimutation récessive
Hyperméthylation et répression

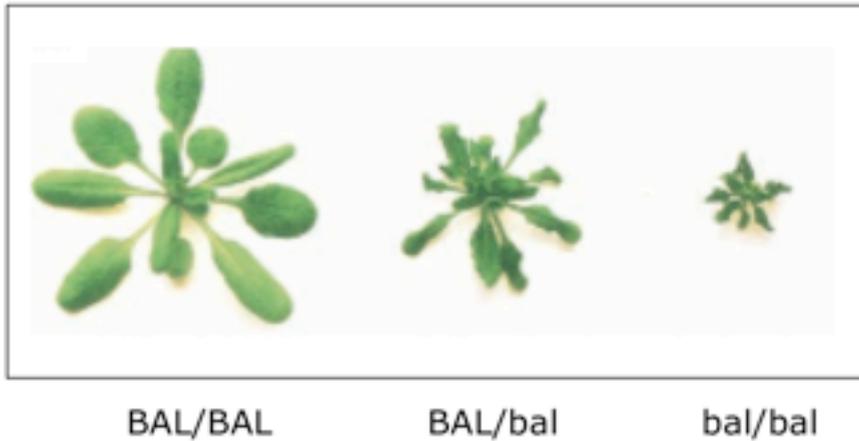


Figure 4 : Exemples d'épimutations induites chez le mutant *ddm1* : épimutation dominante (*bal*) résultant d'une hypométhylation et d'une sur-expression du gène *BAL* (à gauche), épimutation récessive (*bns*) résultant d'une hyperméthylation et d'une sous-expression du gène *BNS* (à droite).

Planches extraites de Stockes et al, 2002 (13) et Saze et al, 2007 (12).

sauvages portant des gènes *DDM1* ou *MET1* fonctionnels (11-14). En effet, on observe dans les mutants *ddm1* et *met1* qui présentent une déméthylation globale de l'ADN, une hyperméthylation locale de l'ADN comparable à ce qui est observé dans certaines cellules cancéreuses. Cette hyperméthylation est certainement due à une redistribution aberrante de la méthylation en l'absence des protéines qui instruisent son maintien. Une fois déposées, ces nouvelles marques épigénétiques sont maintenues et ne sont perdues qu'au hasard des erreurs de cette machinerie de méthylation de l'ADN qui, contrairement à la machinerie de réplication de l'ADN, ne dispose pas d'une activité de correction des erreurs (1).

Conséquences pour les sélectionneurs et améliorateurs des plantes

L'identification de millions de petits ARN différents chez les plantes a mis en évidence un pan entier d'information génétique (et épigénétique) qui avait été jusqu'à présent ignoré. Il est donc maintenant impensable de continuer à parler de séquences neutres du génome ou d'ADN poubelle (terme jusqu'alors employé pour désigner les parties du génome situés entre les gènes codants des protéines, et dont on pensait qu'elles ne portaient aucune information fonctionnelle). Des petits ARNs sont produits par la quasi-totalité de ces régions inter-géniques, indiquant donc qu'elles sont transcrites. Le terme région inter-génique est donc certainement abusif puisqu'un gène est défini comme un segment d'ADN codant une fonction biologique, ce qui est le cas de ces régions lorsqu'elles produisent un ARN régulateur. La fonction de tous ces

petits ARNs n'a pas encore été élucidée, mais il ne faut pas oublier que la fonction de tous les gènes codant des protéines n'est pas encore connue non plus. Lorsqu'il cherchera à identifier les bases moléculaires d'un caractère d'intérêt agronomique, le sélectionneur qui a recours à la sélection assistée par marqueurs devra donc prendre en considération l'information potentiellement portée par tous les segments du génome et pas seulement les régions annotées comme permettant de produire une protéine.

On sait que de nombreux stress peuvent modifier le profil des petits ARN d'une cellule et par là même son profil épigénétique. On sait par ailleurs, que les siRNA de 21-nt et 24-nt circulent au sein de la plante, soit de cellule à cellule via les plasmodesmes qui les relient entre elles, soit à longue distance via les tissus vasculaires et en particulier le phloème. La propagation des siRNA de 21-nt de cellule à cellule permet d'établir des gradients moléculaires impliqués dans la régulation dose-dépendante des gènes contrôlant la différenciation cellulaire (15). La propagation des siRNA de 21-nt à longue distance permet quant à elle d'élucider en tous points de la plante les réactions de défense contre les pathogènes avant même que ceux-ci ne se propagent au sein de la plante (16). Enfin, la propagation des siRNA de 24-nt à longue distance provoque, dans des cellules qui reçoivent ces petits ARN mais ne les produisent pas, des modifications épigénétiques potentiellement héréditaires (17). On peut donc imaginer qu'un nouveau petit ARN de 24-nt induit dans une cellule par un stress peut potentiellement causer une modification épigénétique dans une autre cellule, et que, si la modification survient dans une cellule du méristème reproductif, celle-ci sera transmise à la descendance.

On peut donc se représenter une plante comme un organisme perpétuellement soumis à des stress biotiques (agression par des pathogènes viraux, bactériens, fongiques, etc.) ou abiotiques (stress thermiques, hydriques, etc.). En réponse à certains stress, les plantes peuvent produire des myriades de petits ARNs de façon à générer de la variabilité épigénétique potentiellement transmissible à la descendance. Le maintien sur plusieurs générations d'un stress ou d'une condition environnementale particulière pourrait vraisemblablement conduire à la sélection des individus qui présentent le répertoire de petits ARNs et l'épigénome le plus adapté. Trois exemples semblent soutenir cette hypothèse : l'adaptation à des stress ionisants (18), l'adaptation à des températures élevées (19), et la sélection d'individus à haut rendement énergétique (20). Les améliorateurs des plantes et les sélectionneurs des animaux d'élevage devront donc prendre en compte le répertoire de petits ARNs et l'épigénome dans la sélection des variétés tout en gardant à l'esprit que les épimutations sont beaucoup plus labiles que les mutations. Il est d'ailleurs fort probable qu'à quelques exceptions près, les épimutations ne puissent être maintenues longtemps et que seule la transformation d'épimutations en mutations permette de fixer un nouveau caractère sur le long terme. Ceci permettrait d'expliquer qu'on ait recensé à ce jour aussi peu de caractères agronomiques causés par des épimutations.

Références

1. Paszkowski J & Grossniklaus U (2011) Selected aspects of transgenerational epigenetic inheritance and resetting in plants. (Translated from eng) *Curr Opin Plant Biol* 14(2): 195-203 (in eng).

2. Cubas P, Vincent C, & Coen E (1999) An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401(6749):157-161.
3. Manning K, *et al.* (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat Genet* 38(8):948-952.
4. Martin A, *et al.* (2009) A transposon-induced epigenetic change leads to sex determination in melon. (Translated from eng) *Nature* 461(7267):1135-1138 (in eng).
5. Chandler VL (2010) Paramutation's properties and puzzles. (Translated from eng) *Science* 330(6004):628-629 (in eng).
6. Vaucheret H (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes Dev* 20(7):759-771.
7. Vaucheret H (2008) Plant ARGONAUTES. (Translated from eng) *Trends Plant Sci* 13(7):350-358 (in eng).
8. Henderson IR & Jacobsen SE (2007) Epigenetic inheritance in plants. *Nature* 447(7143):418-424.
9. Matzke M, Kanno T, Daxinger L, Huettel B, & Matzke AJ (2009) RNA-mediated chromatin-based silencing in plants. (Translated from eng) *Curr Opin Cell Biol* 21(3):367-376 (in eng).
10. Ito H, *et al.* (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. (Translated from eng) *Nature* 472(7341):115-119 (in eng).
11. Reinders J, *et al.* (2009) Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. (Translated from eng) *Genes Dev* 23(8):939-950 (in eng).
12. Saze H & Kakutani T (2007) Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1. *Embo J* 26(15):3641-3652.
13. Stokes TL, Kunkel BN, & Richards EJ (2002) Epigenetic variation in Arabidopsis disease resistance. *Genes Dev* 16(2):171-182.
14. Teixeira FK, *et al.* (2009) A role for RNAi in the selective correction of DNA methylation defects. (Translated from eng) *Science* 323(5921):1600-1604 (in eng).
15. Chitwood DH, *et al.* (2009) Pattern formation via small RNA mobility. (Translated from eng) *Genes Dev* 23(5):549-554 (in eng).
16. Ding SW & Voinnet O (2007) Antiviral immunity directed by small RNAs. *Cell* 130(3):413-426.
17. Melnyk CW, Molnar A, & Baulcombe DC (2011) Intercellular and systemic movement of RNA silencing signals. (Translated from eng) *EMBO J* 30(17):3553-3563 (in eng).
18. Molinier J, Ries G, Zipfel C, & Hohn B (2006) Transgeneration memory of stress in plants. *Nature* 442:1046-1049.
19. Whittle CA, Otto SP, Johnston MO, & Krochko JE (2009) Adaptive epigenetic memory of ancestral temperature regime in Arabidopsis thaliana. *Botany* 87:650-657.
20. Hauben M, *et al.* (2009) Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. *Proc Natl Acad Sci U S A* 106(47):20109-20114.
21. Chandler VL (2007) Paramutation: from maize to mice. *Cell* 128(4):641-645.