



Les outils de la génomique renouvellent la biologie

Yves Chupeau

Directeur de Recherches - Institut Jean-Pierre Bourgin, INRA Versailles
Membre correspondant de l'Académie d'agriculture de France

Manuscrit révisé le 1er octobre 2012 - Publié le 27 octobre 2013

Résumé : Les concepts de la biologie moléculaire établis par la microbiologie expérimentale, ont progressivement orienté les démarches de l'ensemble des biologistes. Depuis trente ans, et de façon accélérée au cours de la dernière décennie, les progrès des techniques de la génomique renouvellent les conceptions sur les génomes et leurs fonctionnements.

L'automatisation de plus en plus poussée des outils de la génomique et leurs coûts de plus en plus bas, autoriseront prochainement une banalisation de leurs utilisations. Le recours au séquençage intégral devient routinier pour de nombreux domaines de la biologie et de la génétique. Les fonctions des gènes sont également rendues plus rapidement accessibles, grâce au développement de différents outils de génomique fonctionnelle de plus en plus précis.

Ces développements techniques produisent des avancées fondamentales majeures sur les génomes et leurs fonctionnements, qui trouvent des applications en amélioration des plantes, et donc potentiellement en agriculture. De façon générale ces connaissances révèlent l'extraordinaire malléabilité des génomes, qui sous-tend les processus d'adaptation et d'évolution.

1- Les fondements de la génomique en perspective

La résolution de la structure de l'ADN (acide désoxyribonucléique, voir <http://www.johnkyrk.com/CellIndex.fr.html>) au début des années 1950 par une équipe anglaise, formidable progrès conceptuel et technique, laissait cependant la majorité des biologistes dans l'expectative. Si cette structure permettait d'imaginer comment cette molécule pouvait se recopier à l'identique, rien ne permettait alors d'esquisser son rôle effectif dans le fonctionnement cellulaire. Ce qui permettait aux tenants du rôle vital et primordial des protéines de se gausser quelque peu...

Toutefois, les choses se précisent assez rapidement ; dès 1961 à Paris, Jacob et Monod assignent un rôle d'intermédiaire aux ARN (acide ribonucléique), molécule chimiquement caractérisée en 1956 aux USA par Volkin et Astrachan (ADN-like RNA).

Ainsi se mettait en place le schéma fonctionnel initial de la biologie moléculaire : les informations, codant les protéines contenues dans l'ADN, sont transcrites en ARN spécifiques (messagers) capables de sortir du noyau de la cellule pour être lus par les ribosomes, sièges de la traduction de ces ARN en protéines actives (enzyme, régulation...), ou structurales.

En caractérisant le mécanisme de répression dans la régulation de la β -galactosidase de *E. coli*, et donc le rôle de promoteur de transcription de certaines séquences de l'ADN, Lwoff, Jacob et Monod impriment l'élan fondateur de la biologie moléculaire, en initiant l'étude de la régulation de l'activité des gènes, qui leur vaudra le prix Nobel 1965.

En simplifiant les conceptions d'alors, le gène est constitué de deux parties essentielles. La partie codante dont la séquence comporte les informations de synthèse d'une protéine. La portion régulatrice (promoteur), généralement située en amont de la séquence codante, qui détermine dans quelle cellule, à quel moment du développement, et sous l'effet de quel signal, le gène est exprimé ou non. Ce sont ces séquences régulatrices de la transcription, et les effecteurs protéiques associés (les facteurs de transcription, activateurs, répresseurs) qui assurent les interactions de la cellule avec les autres organes, ou avec le milieu extérieur.

La caractérisation des quatre bases (éléments essentiels) de l'ADN avait permis de pronostiquer que les vingt acides aminés constitutifs des protéines devaient être codés par des triplets de bases. Le dispositif de spécification de la succession précise des acides aminés des protéines, le code génétique, fut essentiellement décrypté en 1965 au NIH (USA) par une équipe coordonnée par Nirenberg (Nobel médecine 1968).

Les enzymes de restriction, étudiés initialement dès 1960 par Arber en Suisse sur les virus d'*E.coli*, sont des enzymes qui hydrolysent des ADN exogènes. Ces activités de restriction constituent donc des mécanismes de défense utilisés par les bactéries. Selon les espèces de bactéries, il existe une grande diversité d'enzymes de restriction, chacune ayant une spécificité du site d'hydrolyse (courte séquence spécifique). Cette spécificité constitue un avantage technique irremplaçable dans les procédés de biologie moléculaire, elle permet de scinder l'ADN à étudier en fragments limités par des séquences connues. Cette propriété a permis le développement de nombreux marqueurs anonymes pour l'amélioration des plantes. Les enzymes de restriction sont surtout utilisées dans les techniques de clonage de gènes.

Le séquençage, le décryptage de l'arrangement de toutes les bases d'un organisme, constitue bien sûr une étape essentielle dans le développement des concepts et des techniques de la génomique. L'expérience du séquençage de protéine acquise par Sanger sur l'insuline (Nobel chimie 1980) à Cambridge en Angleterre, l'a probablement incité à se préoccuper du séquençage de l'ADN. Il propose une méthode de séquençage de l'ADN dès 1975, qui repose sur l'utilisation d'analogues de bases qui bloquent la synthèse d'ADN, ce procédé nécessite quatre réactions de polymérase *in vitro* avec chacun des quatre nucléotides modifiés (didésoxyribonucléotides). L'incorporation des analogues pour l'ensemble des brins en cours de synthèse, mais spécifiquement à la place de la bonne base, génère statistiquement un grand nombre de fragments de tailles différentes mais terminés par l'analogue qui bloque l'élongation. Les fragments des quatre réactions sont ensuite séparés par électrophorèse en gel. Les tailles des différents fragments synthétisés permettent de déduire la séquence (<http://www.ncbi.nlm.nih.gov/books/NBK21505/figure/A1661/>).

Cette technique assez lourde, est appliquée au séquençage complet d'un génome de virus en 1977 (Sanger, Nobel chimie 1980).

Une avancée technique décisive est publiée en 1986 par Mullis (Nobel chimie 1993). Mullis décrit l'amplification, donc la multiplication d'une séquence précise d'ADN *in vitro* par l'utilisation de deux séquences amorces complémentaires de chaque extrémité du segment d'ADN à amplifier. Ce dispositif amélioré par l'utilisation d'une ADN polymérase résistante à la chaleur, isolée d'une bactérie (*Thermus aquaticus*), permet de multiplier exponentiellement (on dit "amplifier") considérablement des fragments d'ADN *in vitro* (en alternant des cycles successifs de synthèse d'ADN, puis de séparation des deux brins complémentaires par la chaleur dans le même tube de réaction, sans avoir à remplacer l'ADN polymérase à chaque cycle).

Cette technique la **PCR (Polymerase Chain Reaction)** et ses améliorations successives connaissent une utilisation désormais routinière dans de nombreux domaines de la génomique, et tout spécialement en santé humaine, ainsi que dans les techniques de clonage de gènes. Deux améliorations sont techniquement importantes : l'amplification à partir d'ARN, grâce à l'utilisation d'une activité reverse transcriptase qui permet de synthétiser des copies ADN des RNA, et la PCR quantitative.

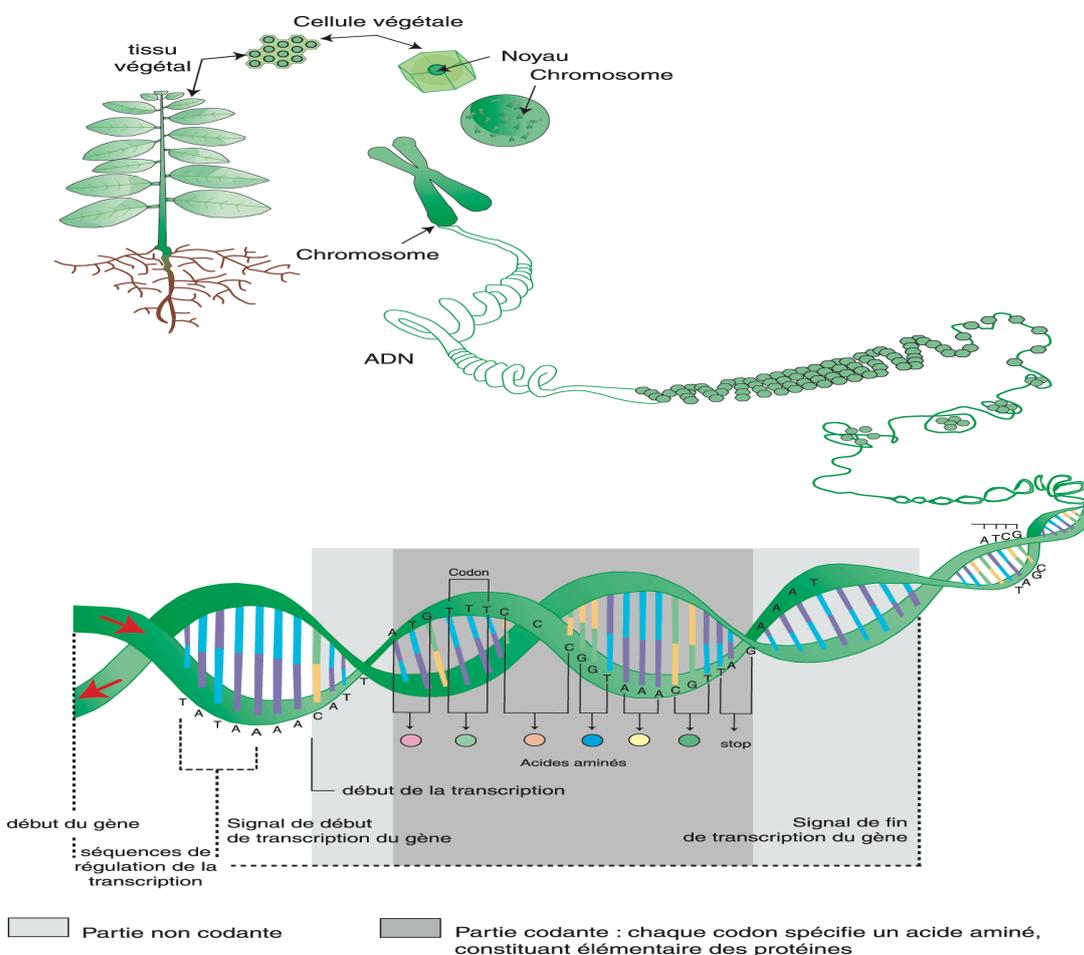
Les plasmides et le clonage de gène. À la fin des années 60, l'équipe de Cohen de Stanford University aux USA découvre les plasmides de *E.coli*, ces petites entités d'ADN circulaires se trouvent à l'extérieur des chromosomes bactériens ont des propriétés particulières. Les plasmides, « compléments » de matériels héréditaires non indispensables mais fort utiles, se répliquent normalement, en même temps que les chromosomes. Mais la multiplication inductible, jusqu'à quelques centaines de copies par

cellules, indépendante de la réplication du chromosome, constitue un moyen de réagir à une variation brutale de l'environnement plus efficace que la sexualité. Les plasmides comportent des gènes de survie, tels que des résistances à différents stress, ou de compétition par la production de toxines... Ainsi, dans la nature, en cas « d'attaque » par un antibiotique produit par une autre espèce de bactérie, l'amplification du plasmide codant une enzyme de détoxification de l'antibiotique permet à la bactérie attaquée de réagir rapidement, et donc de survivre.

Cette amplification est largement utilisée dans les procédés de clonage, car elle permet d'isoler à jusqu'à des milligrammes de plasmides, à partir de culture du clone bactérien qui héberge le plasmide étudié (pour les plasmides à grand nombre de copies).

Ces caractéristiques des plasmides associées à l'utilisation des enzymes de restriction permettent à l'équipe de Cohen de publier un premier succès de génie génétique en 1973. Après hydrolyse d'un plasmide par une enzyme de restriction qui ne coupe pas dans le gène de résistance à la tétracycline, l'un des segments de l'ADN obtenu peut s'intégrer dans un plasmide coupé par le même enzyme. Les séquences complémentaires générées par l'activité de restriction permettent l'appariement des extrémités que l'on peut ensuite lier *in vitro* de façon covalente par une enzyme de ligation. Il s'agit donc de la construction d'un nouveau plasmide, qu'il est possible de transférer dans les bactéries (*E. coli*) que l'on peut sélectionner sur les capacités de résistance aux antibiotiques ainsi transférées.

Ce schéma constitue l'un des fondements essentiel du développement et de la diversification des méthodes de clonage de gènes et donc de la génomique. Les séquences d'ADN à l'étude sont insérées dans des plasmides. Ces plasmides recombinants sont réintroduits dans des bactéries, dont la culture en milieu inducteur permet d'amplifier le plasmide recombinant. L'extraction en permet l'analyse (séquençage, par exemple) et l'utilisation. Les séquences d'ADN à l'étude peuvent être associées de différentes façons dans des constructions moléculaires en variant les séquences régulatrices, soit pour leur expression directe en masse



De la plante au gène - Schéma INRA Editions, Versailles.

dans des microorganismes, soit pour les réintroduire dans différents organismes par les procédés expérimentaux de transfert de gènes.

De nombreux plasmides types ont été progressivement développés pour des usages divers sur des organismes variés, et sont disponibles commercialement. Une amélioration récente, qui met en jeux des sites de recombinaison spécifiques sur les plasmides, simplifie considérablement les opérations de clonage (http://www.premierbiosoft.com/tech_note/Gateway_cloning.html).

2- Une révolution permanente : les progrès des techniques de séquençage

Au cours des trente dernières années, le principe du séquençage mis au point par Sanger (chain termination method) a constamment progressé, en grande partie en raison de l'automatisation de plus en plus poussée de chaque étape de la technique, de la préparation des échantillons jusqu'à la lecture et l'assemblage des séquences (bioinformatique). Pour les grands génomes, les systèmes reposent sur l'amplification de petits fragments de l'ensemble du génome. Ces fragments partiellement chevauchants, de quelques centaines de bases, sont amplifiés soit par clonage (bactéries, levures...) mais de plus en plus par PCR. L'une des premières améliorations reposait sur l'utilisation des quatre nucléotides fluorescents (à des longueurs d'ondes différentes) dans le milieu réactionnel de l'ADN polymérase *in vitro*. Les brins d'ADN synthétisés de différentes tailles sont ensuite séparés selon leur taille par électrophorèse capillaire, la fluorescence spécifique de la dernière base de chaque segment d'ADN est enregistrée et donne directement la séquence du fragment qui a servi de matrice. La séquence complète de l'organisme étudié est ensuite déduite *in silico* à partir des recouvrements des séquences des différents fragments du génome.

De nouveaux procédés ont été mis au point, souvent par amplification puis synthèse simultanée de très nombreux fragments à séquencer, et selon des principes de lecture variés.

(Sur les séquençages, voir <http://www.snv.jussieu.fr/vie/dossiers/sequencage/sequence.htm>).

L'automatisation de plus en plus sophistiquée permet de réduire progressivement les coûts de séquençage, pour des utilisations de plus en plus ciblées (Metzker, 2010).

Au début de l'année 2012, le panorama des plantes dont le génome est séquencé est impressionnant (voir http://genomevolution.org/wiki/index.php/Sequenced_plant_genomes).

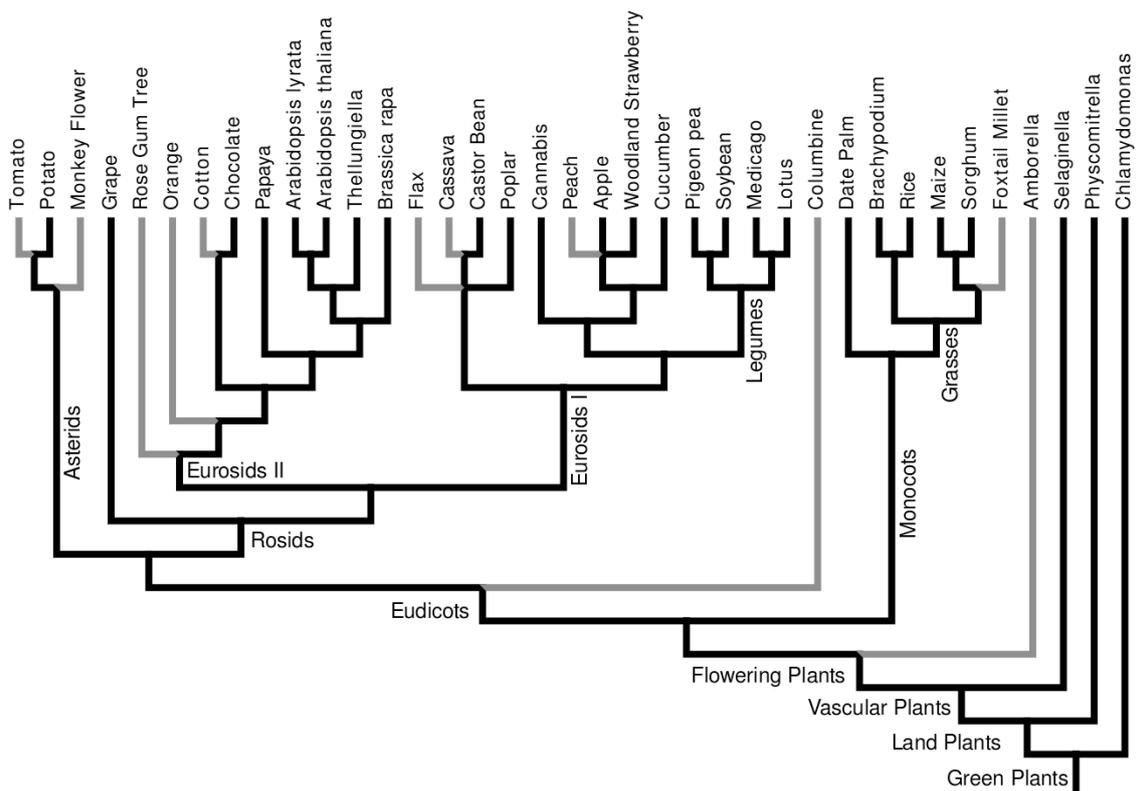


Table 1 : Panorama des plantes dont le génome est séquencé (traits noirs), et dont le séquençage est en cours (traits gris). Extrait de http://genomevolution.org/wiki/index.php/Sequenced_plant_genomes

Certains programmes en cours ne sont pas encore pris en compte dans ce panorama, c'est le cas de l'orge, par exemple, dont le séquençage devrait être publié courant 2012. De surcroît, pour certaines plantes (*Arabidopsis thaliana*, *Brachypodium distachyon*...), les programmes en cours concernent le séquençage de plusieurs centaines d'écotypes (Weigel, 2012).

Ces dispositifs de séquençage s'appliquent à tous les organismes, et donc aux organismes pathogènes et parasites, ce qui imprime une dynamique renouvelée aux études des relations hôte-pathogènes.

Ces capacités d'analyses vont d'ailleurs se renforcer bien davantage, en raison des coûts de plus en plus faibles des dispositifs de séquençage. Les formidables enjeux de santé contribuent puissamment à l'amélioration des procédés de séquençage, avant tout pour les génomes humains. Le coût du séquençage d'un génome humain est aujourd'hui de l'ordre de mille fois moins cher qu'il y a quelques années ! A la fin de l'année 2011 ce sont globalement 30 000 génomes humains qui ont été séquencés. L'automatisation et la miniaturisation de plus en plus poussée des appareils vont permettre d'en doter la majorité des hôpitaux, puis des laboratoires d'analyse...

Ainsi, dès aujourd'hui, la société Life technologies commercialise un appareil de paillasse capable de séquencer un génome humain en deux heures et pour moins de 800 euros ...<http://www.prnewswire.com/news-releases/life-technologies-introduces-the-benchtop-ion-proton-sequencer-designed-to-decode-a-human-genome-in-one-day-for-1000-136990438.html>.

De fait, la concurrence sur le séquençage est acharnée, l'année 2012 devrait voir la commercialisation de plusieurs appareils de paillasse pour moins de mille euros et très économes en réactifs, capables de séquencer directement des fragments de cent kilobases par des procédés nouveaux. C'est le cas, par exemple, de Oxford Nanopore technologies en Angleterre (<http://www.nanoporetech.com/>).

Il est assez tentant de pronostiquer le formidable impact de ces progrès technologiques sur l'accélération des démarches et des acquis de la génomique. Ces facilités expérimentales, jusque là réservées aux plantes modèles, vont s'appliquer à toutes les plantes, à tous les organismes.

Au delà de ces améliorations strictement techniques, l'une des dernières avancées des dispositifs modernes de séquençage ouvre des possibilités d'analyses biologiques encore plus larges et plus informatives sur le fonctionnement du génome : il s'agit du séquençage des ADNc obtenus par transcription inverse des ARN (RNA-seq). Il s'agit d'analyser l'ensemble des ARN exprimés par un génome dans des conditions données, en donnant accès non seulement aux ARN messagers (séquences codantes), mais aussi à l'ensemble des ARN transcrits, non codants, qui possèdent à des degrés divers des rôles importants dans la régulation de l'activité de l'ensemble des gènes. Il s'agit donc d'analyses du transcriptome bien plus riches que celles qui reposaient sur les dispositifs d'hybridation (micro-arrays). Les premières utilisations du RNA-seq révèlent effectivement de nombreux transcrits qui n'avaient pas été pris en compte dans les analyses de type micro-array, indiquant donc que la majorité du génome est exprimée. Ces démarches de « séquençage profond », et leurs raffinements en cours de perfectionnement, offrent donc l'accès à l'analyse de l'extraordinaire complexité des transcrits ainsi qu'à leur rôle biologique (Wang et al, 2009).

Au passage, ces démarches de génomique de plus en plus raffinées, et l'ensemble des connaissances qui en dérivent, renouvellent la compréhension des mécanismes d'évolution des formes de vie (Chupeau, Acad Agric France, 2012). Il faut regretter que la vision de l'extraordinaire malléabilité des génomes ne soit toujours pas correctement vulgarisée.

3- Une sophistication progressive : les outils de la génomique fonctionnelle

a- Le génie génétique expérimental appliqué aux plantes

Le développement de la biologie moléculaire en microbiologie, à partir du milieu des années 1960, avait révélé la puissance des techniques de transfert de gènes dans les démarches de génomique fonctionnelle. Il apparaissait clairement que cet outil serait nécessaire pour élucider l'organisation et l'expression des gènes, ainsi que la régulation de leur expression, puis leur rôle dans le contrôle du développement des organismes...

Pour les plantes, outre la taille et la complexité des génomes, la paroi pecto-cellulosique des cellules végétales constituait un obstacle de taille à la pénétration d'ADN vers le cytoplasme, puis le noyau des cellules cibles. De surcroît, la cohésion entre cellules des tissus végétaux constituait un autre obstacle dû à la paroi.

Dans ce contexte spécifique des cellules végétales, l'idée de faciliter la pénétration de l'ADN a également incité les équipes spécialisées à perfectionner les approches pionnières concernant l'obtention de protoplastes de cellules végétales (cellules dont la paroi a été hydrolysée), ces possibilités ont été vérifiées assez rapidement par les virologistes, qui ont utilisé des protoplastes de tabac pour l'étude de la pénétration de virus et d'acides nucléiques viraux (Aoki et Takebe, 1969).

Au cours des années 1970 des progrès décisifs dans deux domaines d'investigation vont apporter les éléments techniques véritablement déterminants pour le génie génétique :

- La capacité effective de régénérer des plantes fertiles "en masse" à partir de culture *in vitro* de protoplastes isolés de feuilles sur le tabac (Nitsch et Ohyama, 1971). Cette vérification expérimentale éclatante de la totipotence des cellules végétales sera ultérieurement étendue à de nombreuses espèces végétales (Davey et al., 2005).

- L'utilisation expérimentale des plasmides constitue effectivement l'un des tournants décisifs du génie génétique appliqué aux plantes pour deux raisons essentielles.

Techniquement, les gènes de résistance aux antibiotiques, pour la plupart directement efficaces pour les cellules végétales, ont été très utiles pour les premières mises au point des processus de transfert de gènes aux plantes (Barton et al., 1983; Herrera-Estrella et al., 1983).

Surtout, la révélation de la présence de plasmides variés chez la majorité des bactéries a renouvelé l'étude de la bactérie du sol *Agrobacterium tumefaciens*, suspectée, de longue date, de la capacité de transférer des gènes aux plantes. L'existence d'un plasmide de grande taille, le plasmide Ti (tumor inducing) spécifique d'*Agrobacterium tumefaciens*, a permis de caractériser son rôle dans la formation des "tumeurs" de plantes (Van Larebeke et al., 1975). Depuis lors, les différentes phases du processus de transfert de gènes ont été élucidées. Le processus naturel et sophistiqué déployé par ces micro-organismes consiste à réorienter le métabolisme des cellules végétales, afin de favoriser la nutrition et la sexualité des agrobactéries initiatrices. Le processus repose sur un ensemble coordonné de produits d'une centaine de gènes bactériens, qui opèrent le transfert spécifique d'une portion de plasmide vers les noyaux des cellules végétales. Cet ADN-T (ADN transféré) comporte un ensemble de gènes fonctionnels dans un environnement végétal.

En insérant des gènes marqueurs sélectionnables à la place de l'ADN-T dans le plasmide Ti, ce système naturel de transfert fournit les premiers transferts expérimentaux de gènes aux cellules de tabac, puis la régénération des premières plantes transgéniques expérimentales (Barton et al., 1983; Herrera-Estrella et al., 1983) . La connaissance progressivement affinée de ce processus naturel de transfert de gènes, a permis de le diversifier expérimentalement, et d'en améliorer l'efficacité (Chupeau, 2001), pour le rendre utilisable sur une large gamme d'espèces végétales, voire de champignons, et même de cellules humaines !

Les propriétés des agrobactéries sont le plus souvent mises à profit pour les transferts de gènes expérimentaux, mais d'autres procédés, qualifiés de « transfert direct », ont été développés. Ils font intervenir l'utilisation directe des plasmides d'*Escherichia coli*. Le principe général de ces procédés

repose sur la possibilité de délivrer des quantités importantes d'ADN à transférer dans le noyau des cellules cibles au moyen de procédures de perméation efficaces, qui ne détériorent pas irrémédiablement les cellules cibles (Chupeau et Davey, 2007).

Un progrès décisif résulte de la possibilité de s'affranchir de l'étape fastidieuse et le plus souvent limitante de la culture *in vitro*. La mise au point de l'infiltration d'agrobactéries sous vide partiel dans les inflorescences d'*Arabidopsis thaliana* a fourni la procédure pratique attendue, permettant de générer rapidement les premières collections de mutants d'insertion chez *Arabidopsis* (Bechtold et al., 1993). Les plantes traitées sont ensuite élevées en serre pour en récolter les graines, dont les semis sur des milieux sélectifs (antibiotiques, herbicides, ...) sélectionnent les événements de transfert dans la descendance. La plupart des plantes transformées sont hémizygotés pour le transgène (présent en un seul allèle), non chimériques. Ce procédé, simplifié, par l'utilisation d'un détergent au lieu de l'infiltration sous vide, est devenu routinier dans les démarches de génomique fonctionnelle d'*Arabidopsis*. Cependant aujourd'hui, cette méthode n'est applicable en routine que pour quelques membres de la famille des Brassicacées (Curtis, 2005).

Dans les démarches expérimentales, les constructions moléculaires comportent les gènes de sélection dérivés des outils de la microbiologie : les résistances aux antibiotiques. Sur le même principe d'autres marqueurs ont été développés : les résistances aux herbicides. Mais il est possible d'utiliser d'autres systèmes de sélection moins discutables en vue d'applications en amélioration des plantes, en particulier avec des gènes de plante (Mentewab et Stewart, 2005 ; Song et al., 2010). Les progrès des dispositifs expérimentaux permettent même, pour certaines plantes, de se passer de gènes de sélection (de Vetten et al., 2003).

La sélection d'évènements de transfert de gènes dans le génome chloroplastique s'effectue dans un contexte assez spécifique, dû au caractère très polyploïde du génome chloroplastique, et résulte essentiellement de recombinaisons homologues (remplacement des séquences aux mêmes sites), comme chez les bactéries. Il faut donc incorporer dans les constructions des séquences homologues choisies dans les séquences intergéniques des génomes de chloroplastes (Maliga, 2004), les fréquences d'insertion peuvent se révéler assez élevées (Langbecker et al., 2004).

Alors que les procédés de transformation expérimentale des mitochondries sont disponibles chez la levure (Johnston et al., 1988) et l'algue verte *Chlamydomonas* (Boynton et Gillham, 1996), les tentatives chez les plantes se sont révélées infructueuses jusqu'à aujourd'hui. La complexité, la polyploïdie et la plasticité du chondriome chez les plantes imposent sans doute des procédures de sélection encore plus efficaces et spécifiques que celles utilisées pour la sélection de chloroplastes transgéniques.

Ces difficultés expérimentales ne cessent de surprendre, surtout depuis que l'on connaît l'extraordinaire efficacité du transfert horizontal (d'espèce à espèce, par opposition au transfert vertical vers les descendants) de gènes aux génomes mitochondriaux de plantes dans les conditions naturelles (Bergthorsson et al., 2004).

b- Le processus d'intégration des transgènes

Le mécanisme d'intégration dans le génome nucléaire des plantes ne semble pas dépendre de la forme d'ADN utilisée (circulaire, linéaire, double brin, simple brin).

À la différence des microorganismes pour lesquels la recombinaison homologue constitue le processus le plus général, pour les plantes supérieures le mécanisme d'intégration dans le génome repose essentiellement sur la recombinaison non homologue due aux processus de réparation des cassures double brin de l'ADN receveur.

Le détail du mécanisme, assez complexe, n'est pas encore totalement élucidé. Initialement, les travaux ont surtout porté sur l'intégration de l'ADN-T (ADN simple brin, excisé du plasmide Ti). L'utilisation

d'une endonucléase (I-SceI) dont les sites sur l'ADN double brin sont très rares, permet de vérifier qu'en fait l'ADN-T est spécifiquement intégré aux sites de coupure double brin du génome receveur (Tzfira et al., 2003) et que l'ADN-T est converti en double brin avant intégration (ce qui explique d'ailleurs l'expression très rapide, et transitoire des gènes de l'ADN-T).

La vision actuelle du processus d'intégration s'oriente donc vers un mécanisme général de réparation faisant intervenir la ligation d'ADN double brin. Il reste cependant à préciser le détail de ce processus, d'autant que les composantes du complexe de réparation de l'ADN, à l'œuvre dans l'intégration des transgènes, semblent varier selon les types cellulaires (Gallego et al., 2003; Li et al., 2005).

Il importe de souligner un point important : l'intégration des transgènes expérimentaux peut s'effectuer de façon aléatoire dans l'ensemble du génome, aux sites des cassures de l'ADN receveur, par les processus naturels de réparation de l'ADN, exactement comme pour les transferts de gènes naturels.

Dans la pratique cette difficulté est surmontée par les dispositifs de sélection des évènements d'insertion, et par le tri du nombre d'insertion, par des analyses moléculaires, dans les descendances.

Pour les plantes destinées à une utilisation agronomique, ce procédé implique la caractérisation précise, tant du phénotype recherché, dénué de mutations annexes, que de la stabilité de l'expression du transgène dans des environnements de culture variés, sur plusieurs générations et saisons de culture. Un parcours absolument identique à un processus de sélection en amélioration des plantes.

Cependant, dans les utilisations scientifiques, ces processus d'insertion aléatoire posent quelques problèmes. Pour une construction donnée, même avec un promoteur de transcription efficace et constitutif, les niveaux d'expression peuvent varier assez largement selon les contrôles qui s'exercent au site d'insertion, ou au voisinage. Chaque insertion constitue de fait une mutation du génome, certaines peuvent se révéler silencieuses, mais d'autres peuvent provoquer des altérations de l'expression de gènes importants ce qui peut se révéler gênant pour l'interprétation des résultats. De plus, certains évènements d'intégration avortent, aboutissant à des intégrations tronquées ou à des réparations imparfaites des cassures double brin, engendrant des mutations dans le génome des plantes transformées.

c- La mutagenèse d'insertion

En utilisant des transgènes adaptés, qui autorisent des sélections soit sur tissus cultivés *in vitro*, soit sur semis en serre, l'effet mutagène des insertions a été rapidement exploité pour étiqueter les gènes des espèces modèles (Bouchez et Höfte, 1998). La caractérisation des phénotypes obtenus renseigne sur la fonction du gène muté, et la séquence insérée, qui fournit une « étiquette » dans le génome receveur, permet l'accès aux séquences bordures de l'insertion et facilite l'identification de la séquence du gène muté. Assez directement si le génome est séquencé, ce qui est le cas pour les génomes modèles (Arabidopsis, riz, Brachypodium...), ou par homologie de séquence si le génome n'est pas séquencé.

Des bases de données publiques très complètes, et régulièrement mises à jour, fournissent désormais l'ensemble des informations disponibles sur un génome donné (pour Arabidopsis, peuplier et vigne) voir par exemple Flag db++ (<http://urgv.evry.inra.fr/projects/FLAGdb++/HTML/index.shtml>).

De façon plus générale des bases de données plus larges (<http://www.plantgdb.org/prj/GenomeBrowser/>) permettent des analyses comparées des différents génomes séquencés.

Une autre possibilité de mutagenèse d'insertion repose sur l'utilisation de certains retro-transposons de plante dans des contextes hétérologues, par exemple chez le riz (Yamazaki et al., 2001) ou la luzerne (D'Erfuth et al., 2003) ou encore la laitue (Mazier et al., 2007).

d- L'inactivation de gènes, « Silencing », le rôle des petits ARN dans le contrôle de l'expression des gènes, l'épigénétique

Les variations surprenantes, dans certains cas, des niveaux d'expression des transgènes chez les plantes, ont permis de révéler un mécanisme de régulation des gènes totalement insoupçonné auparavant. Au cours de recherches sur le contrôle de la couleur de pétales de *Pétunia*, les tentatives de surexpression de la chalcone synthase, une enzyme clé de la synthèse des flavonoïdes et donc des anthocyanes, avaient conduit à une proportion importante de plantes à fleurs blanches ou panachées de blanc. Le phénomène avait été qualifié de co-suppression, en effet dans ces plantes à pétales blancs, la très faible expression du transgène s'accompagnait d'un niveau de la chalcone synthase résidente bien inférieur à celui des plantes témoins. Ce résultat laissait supposer une interaction à distance des deux séquences homologues, par un mécanisme qui restait à préciser (Napoli et al., 1990). Ce phénomène s'est ensuite manifesté assez souvent dans les transferts de gènes en orientation sens ou antisens (Fagard et Vaucheret, 2000), sans que l'on en comprenne totalement le mécanisme.

L'utilisation d'ARN double-brins très efficaces pour le « silencing » chez le nématode (Fire et al, 1998) a initié l'élucidation du mécanisme, qui révèle un ensemble de dispositifs très ajustés de régulation des gènes, fonctionnel chez tous les organismes, et probablement dérivé de processus de résistance contre les virus.

Le mécanisme fait intervenir des petit ARN (siRNA pour “*short interfering RNA*”, de 20 à 25 paires de bases) qui dérivent de précurseurs d'ARN double-brins non codant par l'action spécifique d'une RNase III, de type Dicer. L'un de ces brins est chargé sur des protéines Argonaute, à domaine RNaseH de type slicer, qui véhiculent le petit ARN vers la séquence complémentaire soit d'un ARN messager ou d'un ARN non codant, au sein de complexes protéiques RISC (RNA Induced Silencing Complexe) qui clivent le messenger, ou bien empêchent avec la traduction, ce qui inhibe sa fonction (PTGS Post Transcriptional Gene Silencing) (voir <http://www.youtube.com/watch?v=-9pROnSD-A>).

Les complexes RISC sont également capables de provoquer la méthylation des séquences D'ADN complémentaires aux ARN non codants ciblées par le siARN, ce qui bloque la transcription (TGS pour “*Transcriptional Gene Silencing*”) (Vaucheret, Acad Agric France, 2012).

Les protéines “Argonaute” initialement découvertes chez les plantes sont très conservées, et constituent le plus souvent des familles de protéines pour le même organisme, les différents membres ont des rôles différents, certaines sont spécifiques du contrôle par méthylation des transposons (Vaucheret, 2008). Les séquençages modernes révèlent de nombreux ARN non codants et siRNA, et l'informatique de très nombreux microARN potentiels (Voinnet, 2009).

À l'origine, ces dispositifs d'inactivation constituaient essentiellement un des moyens de contrer les infections virales. Au cours de l'évolution, ils ont dû évoluer en mécanismes plus généraux de réponse à différent stress, puis incorporés aux régulations de l'activité des gènes. C'est ainsi que sont apparus les microARN (miRNA) codés par des unités de transcription indépendantes et qui régulent en trans (à distance, sur des chromosomes différents) l'expression de nombreux gènes impliqués dans le développement ou la réponse aux stress. Il est possible de dénombrer *in silico* des milliers de séquences non codantes capables potentiellement de générer possiblement des ARN double-brins. Toutefois, chez la plupart des espèces, seuls quelques centaines de miRNA et siRNA endogènes ont une fonction biologique connue, mais ce nombre pourrait s'avérer bien plus important (Cloonan et al, 2012)

Ces processus de régulation des gènes apparaissent donc de plus en plus subtils et fondamentaux, tant dans les processus de développement que dans l'adaptation à l'environnement et, bien sûr, dans les mécanismes de réponse aux agents pathogènes. Ils sont conservés chez tous les organismes, leur étude chez l'homme renouvelle de larges secteurs de la médecine et potentiellement de la pharmacie !

Ces caractérisations ouvrent de larges perspectives fondamentales sur les contrôles génétiques, très finement régulés, des interactions avec l'environnement, eux aussi en évolution permanente (Meng et al., 2011 ; Cuperus et al., 2011), ce qui entraîne une grande diversité selon les plantes. Il existe cependant des dispositifs conservés chez de nombreuses plantes, ainsi le microARN, miR395, qui est spécifiquement induit par une carence en sulfate chez *Arabidopsis* l'est également chez le riz (Jeong et al., 2011).

Enfin, les processus de TGS éclairent de façon de plus en plus précise les mécanismes épigénétiques, qui désignent l'acquisition de caractéristiques héréditaires, sans modification de la séquence de l'ADN, à la différence des mutations. La méthylation des cytosines de certaines séquences des promoteurs de transcription est un mécanisme de répression très conservé chez tous les organismes. Les modalités des méthylations se précisent par l'étude des différentes protéines, qui avec Argonaute, contrôlent la méthylation *de novo*, puis sa maintenance ou encore la dé-méthylation. Également en cours de précision, la relation entre les modifications des histones et les méthylations des cytosines (Law et Jacobson, 2010).

Au plan pratique, ces connaissances permettent de développer des techniques complémentaires de génétique fonctionnelle. Le génie génétique permet de construire des séquences d'ADN non codant générant des portions d'ARN double-brins qui ciblent des séquences codantes, et donc de provoquer des phénomènes de PTGS (post transcriptional gene silencing) et de TGS par les techniques de transfert de gènes. Ce dispositif nécessite la connaissance de la séquence codante à cibler. Le fait que l'expression du gène ciblé n'est pas totalement abolie constitue un avantage pour l'étude des gènes dont la perte totale de fonction est létale.

e- L'inactivation de gènes induite par des virus recombinants (VIGS : virus induced gene silenced)

Les virus à ADN et à ARN régulent les mécanismes d'inactivation par la formation d'ARN double-brins puis de petits ARN double-brins (siRNA : small inducing RNA). Dans ces stratégies de défense, les ARN simple-brin peuvent également être convertis en double-brins par des ARN polymérases ARN dépendantes ce qui amplifie le processus d'inactivation.

Ces connaissances associées aux possibilités du génie génétique sur les virus utilisés de longue date et diversifiés (Marillonnet et al, 2005), ont permis de développer de nouveaux outils de génomique fonctionnelle, plus directs et plus simples que le transfert de gènes. En incorporant une séquence d'un gène de plante dans le génome d'un virus, il est possible de cibler l'extinction de ce gène dans un organe, voire une plante entière, en infectant la plante par le virus recombinant, qui diffuse le « silencing » par l'infection systémique qu'il provoque. Ces techniques dont les fondements biologiques progressent, sont déjà couramment utilisées en raison d'un certain nombre d'avantages.

Les virus recombinants permettent d'obtenir des résultats rapidement, une séquence partielle du gène à cibler est suffisante, enfin ce dispositif permet de diminuer l'expression d'une famille de gènes. La technique est limitée par la spécificité des virus. Elle s'applique cependant assez généralement en choisissant des virus capables d'infecter un grand nombre de plantes, tels que le TRV (Tobacco rattle virus) ou encore l'ASLV (Apple latent spherical virus) et le BSMV (Barley stripe mosaic virus) pour les poacées. Le méristème n'est généralement pas accessible à cette technique. La persistance de l'inactivation peut se manifester durablement chez les plantes à multiplication végétative, et peut se transmettre à la descendance sexuée pour un petit nombre de virus. (Senthil-Kumar et Mysore, 2011).

L'inactivation transcriptionnelle provoquée par VIGS est également de plus en plus utilisée, avec les mêmes avantages et limitations. Dans certains cas l'inactivation peut se transmettre à la descendance, via la méthylation de l'ADN cible, alors que le virus ne passe pas par la graine comme pour le CMV recombinant (Kanazawa et al., 2011). Cette dernière possibilité ouvre des possibilités de modifications héréditaires sans transfert de gènes. Enfin, la précision des techniques d'inactivation s'améliore encore

avec l'utilisation de virus qui comportent des microARN (Tang et al., 2010). Outre la précision obtenue par des microARN artificiels (AmiRNA) qui ciblent spécifiquement une séquence codante, cette amélioration (MIR VIGS) constitue le moyen de vérifier le rôle biologique des nombreux ARN non codants détectés par le séquençage profond, dont justement des miRNA putatifs dont la fonction n'est pas encore connue.

f- La localisation de mutations, le « Tilling » (Targeted induced local lesions in genomes)

Pour créer de la variabilité chez les plantes par mutagenèse, différentes stratégies sont disponibles aujourd'hui. L'utilisation d'agents chimiques ou physiques, l'insertion de transgènes, ou encore pour certaines plantes le recours aux transposons.

Souvent mis à contribution depuis pratiquement un siècle, les programmes de mutagenèse expérimentale ont considérablement élargi et précisé leur efficacité grâce à la génomique. Car il est possible aujourd'hui de trier spécifiquement dans l'ensemble des mutations aléatoires, une mutation dans une séquence donnée, par l'analyse des mélanges d'ADN extraits des descendances issues du traitement mutagène.

Le "Tilling" fait appel à différentes techniques, mais repose sur deux éléments essentiels : l'amplification par PCR et le clivage d'hétéroduplex d'ADN par des exonucléases spécifiques des ADN simple-brin, donc des zones de non-appariement des amplicons. La possibilité d'amplifier par PCR la séquence dans laquelle la mutation est souhaitée, suppose donc sur une connaissance *a priori* d'une séquence homéologue, qui permette la définition d'amorces à l'aide de codons, éventuellement dégénérés. L'action de l'exonucléase sur les produits de l'amplification permet de repérer les mésappariements en séparant les fragments obtenus par électrophorèse (Mac Callum et al., 2000).

Les améliorations apportées à cet outil complémentaire de génomique fonctionnelle (l'utilisation d'amorces de PCR fluorescentes, couplage avec le séquençage...), et l'utilisation de mélanges d'ADN extraits de populations pour identifier des allèles et des SNP (Single Nucleotides Polymorphisms, mutation ponctuelle d'un seul nucléotide), en font également un outil incontournable d'amélioration des plantes (Slade et Knaut, 2005).

g- La mutagenèse dirigée, recombinaison homologue et ciblage des transgènes

Dans les processus de transfert de gènes, que nous avons survolé pour les plantes supérieures, la très faible fréquence d'événements de recombinaison homologue n'avait pas permis de développer des outils réellement pratiques de ciblage des sites d'insertion (même en incorporant des séquences d'homologie dans les constructions moléculaires), contrairement aux dispositifs développés sur les cellules animales ESC (Embryo Stem Cells) (Yates et Daley, 2006).

De fait, il s'agit de trouver les moyens de surmonter la très nette efficacité des dispositifs de réparation non homologue des cassures de l'ADN des plantes supérieures, alors que la recombinaison homologue est très nettement plus efficace chez la mousse *Physcomitella patens* pour des raisons non totalement élucidées (Kamisugi et al., 2006). Les recherches en cours sont essentiellement motivées par la recherche de l'outil qui manque encore en génomique végétale, la mutagenèse dirigée, et par extension l'insertion ciblée de transgènes.

Deux dispositifs ont été majoritairement explorés jusqu'à aujourd'hui, dont le principe est le même : la génération de coupures double brin précisément ciblées de l'ADN « receveur ».

Le premier résulte des connaissances acquises sur les facteurs de transcription, protéines à doigts de zinc (Pavletich et al., 1991), dont la spécificité d'ancrage à l'ADN résulte de séquences d'acides aminés spécifiques qui reconnaissent des triplets de bases. Leur diversité, amplifiée au cours de l'évolution, permet de définir et de construire des assemblages spécifiques des codons.

(Zinc Finger Consortium <http://www.zincfingers.org/scientific-background.htm>).

L'association de plusieurs modules cible l'ancrage sur des séquences définies, et la liaison d'une nucléase non spécifique à ces assemblages permet de localiser la coupure de la double chaîne d'ADN au site précis défini par les protéines à doigt de zinc (Urnov et al., 2010).

Le second dispositif de ciblage résulte de la compréhension des mécanismes de virulence de bactéries (*Xanthomonas*) qui sont capables de flécher des protéines de type activateur de transcription : TAL pour « Transcription Activator Like ». Ces protéines de virulence comportent une séquence d'adressage au noyau et surtout, un domaine central de répétition d'acides aminés en tandem, qui contrôle l'ancrage à l'ADN de l'hôte. Dans ces répétitions un couple d'acide aminé, en position centrale, se lie spécifiquement à une base, la séquence des associations de répétitions cible une séquence d'ADN (Boch et al, 2009).

Outre les possibilités expérimentales qu'offre cette dernière découverte, elle révèle surtout que les bactéries ont développé les moyens de faire du génie génétique ciblé chez les plantes depuis la nuit des temps, nous n'avons rien inventé !

Comme pour les protéines à doigts de zinc, il est possible de construire des séquences de répétitions déterminées (par génie génétique), et d'y associer une nucléase non spécifique, afin de cibler une coupure double brin dans une séquence particulière d'ADN. La précision du ciblage se révèle supérieure à celle des protéines à doigts de zinc (Cermak et al., 2011).

Ces derniers dispositifs ont surtout été utilisés sur les cellules animales et humaines soit pour cibler des mutations, soit pour insérer des transgènes de façon parfaitement ciblée par recombinaison homologue, et même pour restaurer des fonctions vitales *in vivo* (Li et al., 2011).

Leur utilisation pour les plantes commence à peine (Li et al, 2012). Il faut, sans doute, préciser plus avant les mécanismes du ciblage, afin d'en améliorer la précision, mais il semble donc que l'on dispose désormais de l'outil expérimental adapté.

4- Conclusion et perspectives

Le rôle de l'instrumentation et de l'automatisation pour l'acquisition rapide de très nombreuses données sur l'organisation et le fonctionnement des génomes, ressort très clairement de notre survol des développements récents des outils de la génomique (Il en va de même pour les techniques d'imagerie cellulaire qui permettent de localiser et de suivre *in vivo* l'activation d'un gène puis le trajet cellulaire du produit de ce gène). Cependant, ce survol révèle également la forte imbrication entre les avancées technologiques et les progrès conceptuels et cognitifs.

Un niveau supplémentaire de régulation des gènes a été découvert, initialement chez les plantes transgéniques, qui apparaît fondamental et conservé chez tous les organismes. Il met en jeu des petits ARN qui régulent soit la transcription soit la traduction de certains gènes. Ces régulations modulent les réponses aux infections et aux variations de l'environnement, et régissent la croissance et le développement. Certains de ces petits ARN mobilisent les complexes enzymatiques qui modulent la transcription de façon spécifique, par modification des histones et de l'ADN. Les modifications de l'ADN, par méthylation des cytosines, impriment des régulations stables des gènes ciblés. Elles peuvent s'avérer héréditaires, et sont qualifiées d'épigénétiques, car sans modification de la séquence d'ADN (Vaucheret, Acad Agric France, 2012).

De fait, les régulations de l'activité des gènes se révèlent de plus en plus complexes. Il semble que les cytosines hydroxyméthylées, très récemment détectées jouent également un rôle important (Yu et al, 2012). Des mécanismes de régulation par modification des ARN messagers jouent également des rôles importants : la méthylation des adénines (Meyers et al, 2012) et le remplacement ciblé de certaines bases (Editing, Garrett et al., 2012).

En outre, les séquençages modernes révèlent que pratiquement l'ensemble du génome est transcrit. Il reste à déterminer les fonctions potentielles des très nombreux ARN non codants, qui procurent probablement des régulations additionnelles de l'activité du génome. Enfin, l'analyse plus fine des transcrits montre qu'une même séquence d'ADN peut donner lieu à plusieurs produits de fonctions différentes, ce qui modifie le concept de gène, et introduit des niveaux de régulation encore plus complexes.

L'efficacité des séquençages, et surtout des outils d'analyse des données, permet désormais d'analyser et de comparer les génomes de plusieurs centaines de génotypes par exemple de maïs (Chia et al, 2012).

Ces développements techniques et conceptuels délivrent des connaissances biologiques de plus en plus précises, qui complètent et réorientent les méthodes et les objectifs de sélection des plantes, par des démarches de plus en plus ciblées (Charcosset, Acad. Agric. France, 2012 ; Chupeau, Acad. Agric. France, 2012).

Enfin, les séquences disponibles pour de nombreux organismes, renseignent de plus en plus précisément sur les mécanismes de l'évolution des génomes (Exemple des Poacées, Salse, AAF 2012). Plus globalement, le rôle des transferts de gènes naturels apparaît comme l'un des moteurs essentiels de l'évolution (Chupeau, Acad. Agric. France, 2012).

Bibliographie

- Aoki S, Takebe I. (1969) Infection of tobacco mesophyll protoplasts by tobacco mosaic virus ribonucleic acid. *Virology* 39, 439-441.
- Barton KA, Binns AN, Matze AJM, Chilton MD. (1983) Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA and transmission of T-DNA to R1 progeny. *Cell* 32, 1033-1043.
- Bechtold N, Ellis J, Pelletier G. (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *CR Acad. Sci. III Vie* 316, 1194-1199.
- Bergthorsson U, Richardson AO, Young GJ, Goertzen LR, Palmer JD. (2004) Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*. *Proc. Natl. Acad. Sci. USA* 101, 17747-17752.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. (2009) Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* 236, 1509-1512.
- Bouchez D, et Höfte H. (1998) Fonctionnal genomics in plants. *Plant Physiol.* 118, 725-732.
- Boynton JE, Gillham NW. (1996) Genetics and transformation of mitochondria in the green alga *Chlamydomonas*. *Methods Enzymol.* 264, 279-296.
- Cermack T, Erin L. Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* 39(12) e82. Published online 2011 April 14.
- doi: 10.1093/nar/gkr218
- Charcosset, A. (2012) *Acad. Agric. France*, in Comptes rendus sur les potentiels de la science pour l'avenir de l'agriculture, de l'alimentation et de l'environnement.
- Chia et al. (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nature Genetics* doi: 10.1038/ng.2313
- Chupeau Y. (2001) Les raffinements sexuels d'une bactérie du sol au service du génie génétique... *Médecine/ sciences* 17, 856-866.
- Chupeau Y, Davey R. (2007) Gene transfer to Plants. In « Functional Plant Genomics ». Morot-Gaudry JF, Lea P, Briat JF Eds. Science Publishers, pp 123- 142.
- Chupeau, Y. (2012) La génomique renouvelle l'amélioration des plantes. *Acad. Agric. France*, in Comptes rendus sur les potentiels de la science pour l'avenir de l'agriculture, de l'alimentation et de l'environnement.
- Chupeau, Y. (2012) La génomique renouvelle la compréhension de l'évolution. *Acad. Agric. France*, in Comptes rendus sur les potentiels de la science pour l'avenir de l'agriculture, de l'alimentation et de l'environnement.

- Cloonan N, et al. (2011) MicroRNAs and their isomiRs function cooperatively to target common biological pathways. **Genome Biology** 12, R126 doi:10.1186/gb-2011-12-12-r126
- Cuperus J, Fahlgren N, Carrington JC. (2011) Evolution and Functional Diversification of MIRNA Genes. **Plant Cell** 200 23, 431-442.
- Curtis IS. (2005) Production of transgenic crops by the floral dip method. **Methods Mol. Biol.** 286, 103-110.
- Davey MR, Anthony P, Power JB, Lowe KC. (2005) Plant protoplasts : status and biotechnological perspectives. **Biotechnol. Adv.** 23, 131-171.
- D'Erfurth I, Cosson V, Eschtruth A, Lucas H, Kondorosi A, Ratet P. (2003) Efficient transposition of the Tnt1 tobacco transposon in the model legume *Medicago truncatula*. **Plant J.** 33,1-12.
- De Vetten N, Wolters AM, Raemakers K, van der Meer I, ter Stege R, Heeres E, Heeres P, Visser R. (2003) A transformation method for obtaining marker free plants of cross pollinating and vegetatively propagated crop. **Nature Biotechnol.** 21, 439-442.
- Fargard M, Vaucheret H (2000) (Trans)gene Silencing in Plants: How Many Mechanisms. **Annu. Rev. Plant Physiol. Plant Mol. Biol.** 51, 167-94.
- Fire A, Xu SQ, Montgomery MK, Kostas SA, Driver SE, Mello CC. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. **Nature** 391, 806-811.
- Gallego ME, Bleuward JY, Daoudaql-Cotterell N, White CI. (2003) Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. **Plant J.** 35, 557-565.
- Garrett, S. Rosenthal, J. J.C. (2012) RNA Editing Underlies Temperature Adaptation in K⁺ Channels from Polar Octopuses. **Science** 335, 848-851.
- Herrera-Estrella L, Depicker A, Van Montagu M, Schell J. (1983) Expression of chimeric genes transferred into plants cells using a Ti plasmid-derived-vector. **Nature** 303, 209-213.
- Jeong DH, Park S, Zhai J, Gurazada SGR, De Paoli E, Blake C. Meyers BC, J. Green PJ. (2011) Massive Analysis of Rice Small RNAs: Mechanistic Implications of Regulated MicroRNAs and Variants for Differential Target RNA Cleavage. **Plant Cell** 23, 4185-4207,
- Johnston SA, Anziano PQ, Shark K, Sanford JC, Butow RA. (1988) Mitochondrial transformation in yeast by bombardment with microprojectiles. **Science** 240, 1538-1541.
- Kanazawa A. et al. (2011) Virus-mediated efficient induction of epigenetic modifications of endogenous genes with phenotypic changes in plants. **Plant J.** 65, 156-168.
- Kamisugi Y, Schlink K, Rensing SA, Schween G, von Stackelberg M, Cuming1 AC, Reski R, Cove DJ. (2006) The mechanism of gene targeting in *Physcomitrella patens* : homologous recombination, and multiple integration. **Nucleic Acids Res.** 34, 6205-6214.
- Langbecker CL, Ye GN, Broyles DL, Duggan LL, Xu CW, Hajdukiewicz PTC, Armstrong CL, Staub JM. (2004) High frequency transformation of undeveloped plastids in tobacco suspension cells. **Plant Physiol.** 135, 39-46.
- Law JA, Jacobsen SE. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. **Nature Genetics** 11, 204-220.
- Li J, Vaidya M, White C, Vainstein A, Citovski V, Tzfira T. (2005) Involvement of ku80 in T-DNA integration in plant cells. **Proc. Natl. Acad. Sci. USA** 102, 19231-19236.
- Li H, et al. (2011) In vivo genome editing restores haemostasis in a mouse model of haemophilia. **Nature** 475, 217-222.
- Li L.T, Liu B., Spalding M.H., Weeks D.P., Yang, B. (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. **Nature Biotech.** 390-393.
- Liu Y-S, Wang Q-L, Li B-Y. (2010) Gene exchange between cell by graftings. New insights into plant graft hybridization. **Heredity** 104, 1-2; doi:10.1038/hdy.2009.115;
- Law JA, Jacobsen SE. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. **Nature Genetics** 11, 211-220.
- McCallum CM, Comai L, Greene EA, Henikoff S. (2000) Targeting Induced Local Lesions IN Genomes (TILLING) for Plant Functional Genomics. **Plant Physiol.** 123, 439-442,
- Maliga P. (2004) Plastid transformation in higher plants. **Ann. Rev. Plant Biol.** 55, 289-3313.
- Marrillonet S, Thoeringer C, Kandzia R, Klimyuk V, Gleba Y. (2005) systemic *Agrobacterium tumefaciens* mediated transfection of viral replicons for efficient transient expression in plants. **Nature Biotechnol.** 23, 718-723.

- Mazier M, Botton E, Flamain F, Bouchet JP, Courtial B, Chupeau MC, Chupeau Y, Maisonneuve B, Lucas H. (2007) Successful Gene Tagging in Lettuce Using the Tnt1 Retrotransposon from Tobacco. *Plant Physiol.* 144, 18-31,
- Meng Y, Shao C, Wang H, Chen M (2011) The Regulatory Activities of Plant MicroRNAs: A More Dynamic Perspective. *Plant Physiol.* 157, 1583-1595
- Mentewab A, Stewart CN. (2005) Overexpression of an *Arabidopsis thaliana* ABC transporter confers kanamycin resistance to transgenic plants. *Nature Biotechnol.* 23, 1177-1180.
- Metzker M. (2010) The sequencing technologies next generation. *Nature genetics* 11, 31-46.
- Meyer et al., (2012) Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 30 UTRs and near Stop Codons. *Cell* doi:10.1016/j.cell.2012.05.003
- Napoli C, Lemieux C, Jorgensen R. (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell*, 2, 279-289.
- Nitsch JP, Ohshima K. (1971) Obtention de plantes à partir de protoplastes cultivés in vitro. *CR Acad Sci. Paris*, 273, 801-804.
- Pavletich, NP, Pabo CO. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809-817.
- Salse J. (2012) La paléogénomique des plantes pour l'amélioration variétale. *Acad. Agric. France*, in Comptes rendus sur les potentiels de la science pour l'avenir de l'agriculture, de l'alimentation et de l'environnement.
- Senthil-kumar M, Mysore KS. (2011) New dimensions for VIGS in plant functional genomics. *Trends Plant Sci.* 16, 656-665.
- Slade AJ, Knauf VC. (2005) TILLING moves beyond functional genomics into crop improvement. *Transgenic Res.* 14, 109-115.
- Song SG, Sink KC, Ma Y, Herlache T, Hancock JF, Loescher WH. (2010) A novel mannose-based selection system for plant transformation using celery mannose-6-phosphate reductase gene. *Plant Cell Rep.* 29, 163-172.
- Tang T, Wang F, Zhao J, Xie K, Hong Y, Liu Y. (2010) Virus-Based MicroRNA Expression for Gene Functional Analysis in Plants. *Plant Physiol* 153, 632-641,
- Tzfira T, Frankman LR, Vaidya M, Citovsky V. (2003) Specific site integration of *Agrobacterium tumefaciens* T-DNA via double stranded intermediates. *Plant Physiol.* 133, 1011-1023.
- Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. (2010) Genome editing with engineered zinc finger nucleases. *Nature Genetics* 11, 636-646.
- Van Larebecke N, Genetello C, Schell J, Schilperoort RA, Hermans AK, Hernalsteens JP, Van Montagu M. (1975) Acquisition of tumor inducing ability of non oncogenic agrobacteria as a result of plasmid transfer. *Nature* 255, 742-743.
- Vaucheret H. (2008) Plant Argonautes. *Trends plant Sci.* 13, 350-358.
- Vaucheret H. (2012) Rôle de l'épigénome dans le contrôle du développement et des interactions avec le milieu. Quels apports pour l'amélioration des plantes ? *Acad. Agric. France*, in Comptes rendus sur les potentiels de la science pour l'avenir de l'agriculture, de l'alimentation et de l'environnement.
- Voinnet O. (2009) Origin, Biogenesis, and Activity of Plant MicroRNAs. *Cell* 136, 669-687.
- Wang X, Elling AA, Li X, Ning Li N, Peng Z, He G, Sun H, Qi Y, Liu, XS, Deng XW. (2009) Genome-Wide and Organ-Specific Landscapes of Epigenetic Modifications and Their Relationships to mRNA and Small RNA Transcriptomes in Maize. *Plant Cell* 21, 1053-1069.
- Weigel D. (2012) Natural Variation in *Arabidopsis*: From Molecular Genetics to Ecological Genomics. *Plant Physiol.* 158, 2-22.
- Yamazaki M, Tsugawa H, Miyao A, Yano M, Wu J, Yamamoto S, Matsumoto T, Sasaki T, Hirochika H (2001) The rice retrotransposon Tos17 prefers low-copy number sequences as integration targets. *Mol Genet Genomics* 265, 336-344.
- Yates F, Daley GQ. (2006) Progress and prospects: gene transfer into embryonic stem cells. *Gene Therapy* 13, 1431-1439.
- Yu et al., (2012) Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome. *Cell* doi: 10.1016/j.cell.2012.04.027