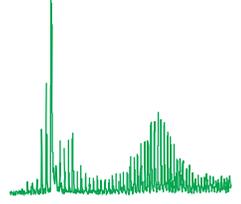
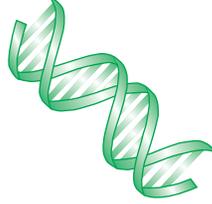
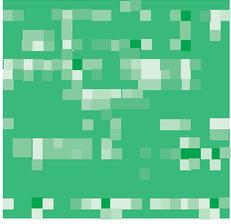
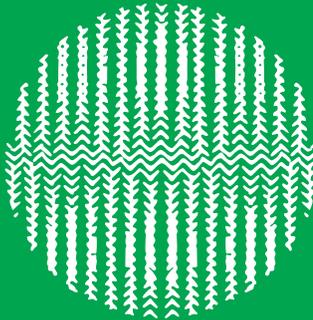


LA GÉNOMIQUE EN BIOLOGIE VÉGÉTALE



J.-F. MOROT-GAUDRY,
J.-F. BRIAT,
coord.

SCIENCE UPDATE



INRA
EDITIONS

LA GÉNOMIQUE EN BIOLOGIE VÉGÉTALE

J.-F. MOROT-GAUDRY, J.-F. BRIAT, coord.

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
147, rue de l'Université - 75338 Paris Cedex 07

Coordinateurs

Jean-François MOROT-GAUDRY
INRA
Biologie végétale
RD 10
78026 Versailles cédex

Jean-François BRIAT
INRA-ENSAM
Biologie végétale
2 place Viala
34060 Montpellier cédex 1

En vente

INRA Editions, RD 10
78026 Versailles Cedex, France
email : INRA-Editions@versailles.inra.fr

© INRA, Paris, 2004
ISBN : ~~2-7380-1167-5~~^λ
ISSN : 1159-554X

© Le code de la propriété intellectuelle du 1er juillet 1992 interdit la photocopie à usage collectif sans autorisation des ayants droit. Le non respect de cette disposition met en danger l'édition, notamment scientifique. Toute reproduction, partielle ou totale, du présent ouvrage est interdite sans autorisation de l'éditeur ou du Centre français d'exploitation du droit de copie (CFC), 20, rue des Grands-Augustins, Paris 6ème.

Sommaire

Préface – M. Caboche 7

Avant-propos – M. Lefort 9

Introduction – J.-F. Briat et J.-F. Morot-Gaudry 11

Génomique structurale et analyse *in silico*

1. Structure physique du génome nucléaire des plantes 15
M. Delseny, M. Echeverria et E. Lasserre

2. Le séquençage de génomes de plantes 33
N. Choisne, N. Demange, G. Orjeda, L. Michelet, E. Pelletier, M. Salanoubat, J. Weissenbach et F. Quétier

3. Génomique et bio-informatique : introduction 59
H. Chiapello et F. Rodolphe

4. Banques et bases de données en biologie 61
H. Chiapello

5. Prédiction de gènes 77
P. Nicolas et H. Chiapello

6. Stratégies permettant d'assigner une fonction *in silico* aux séquences protéiques issues de projets de génomique 91
J.-F. Gibrat et A. Marin

7. Les comparaisons massives de séquences protéiques..... 115
J.-L. Risler, A. Louis, S. Mohseni-Zadeh, P. Brezellec et Y. Diaz-Lazcoz

Génomique fonctionnelle : de la séquence à la fonction *in planta*

8. Les collections de mutants et la génétique inverse 133
F. Granier et D. Bouchez

9. Analyse du transcriptome végétal par les puces à ADN	155
<i>P. Dumas, C. Rothan et S. Robin</i>	
10. La protéomique et ses applications à la biologie végétale	183
<i>M. Zivy, J. Joyard et M. Rossignol</i>	
11. L'apport de la microscopie et de l'imagerie en génomique	209
<i>S. Brown, C. Poujol-Talbot, J. Kronenberger, J. Traas et B. Satiat-Jeunemaitre</i>	
12. Métabolome	237
<i>A. Krapp, L. Kerhoas, A. Hénaut, J. Einhorn et J.-F. Morot-Gaudry</i>	
13. Mesures de flux métaboliques	265
<i>M. Dieuaide-Noubhani, A.-P. Alonso, C. Roby, et P. Raymond</i>	

Les plantes modèles de la génomique végétale

14. L'espèce modèle <i>Arabidopsis thaliana</i>	289
<i>D. Bouchez</i>	
15. Le riz : une plante modèle pour la génomique végétale	303
<i>M. Delseny</i>	
16. <i>Medicago truncatula</i> , plante modèle pour les légumineuses et les interactions plantes-microorganismes	311
<i>E.-P. Journet</i>	
17. La tomate, espèce modèle pour la génomique des Solanacées et pour l'ensemble des fruits charnus	327
<i>C. Etienne, M. Bouzayen, M. Causse et C. Rothan</i>	

Génomique, variabilité génétique et amélioration des plantes

18. Les marqueurs moléculaires	349
<i>M. Falque et S. Santoni</i>	
19. Éléments transposables et analyse de la biodiversité végétale	377
<i>C. Mhiri et M.-A. Grandbastien</i>	
20. Génomique, caractérisation de gènes majeurs et de QTL	403
<i>V. Lefebvre, A. Bendahmane et M. Causse</i>	
21. Dissection génétique d'un caractère complexe chez <i>Arabidopsis</i> : lignées recombinantes et recherche de QTL	427
<i>O. Loudet et F. Daniel-Vedele</i>	

22. Évolution moléculaire et validation de gènes candidats	437
<i>D. Manicacci et A. Charcosset</i>	
23. Le maïs et le blé, céréales modèles pour la recherche en biologie intégrative et son application à la sélection	453
<i>B. Hirel, J. Le Gouis, P. Perez, M. Falque, F. Quétier, A. Murigneux, P. Rogowsky et A. Charcosset</i>	

Réflexions sur la génomique

24. Combiner génomique et modélisation pour l'analyse de la tolérance aux con- traintes environnementales	475
<i>F. Tardieu</i>	
25. De la génomique à l'auto-organisation	493
<i>P. Colonna, V. Planchot, M. Axelos, Y. Popineau</i>	
26. Protection et diffusion des résultats de génomique et biotechnologies végétales : quels enjeux pour la recherche publique ?	515
<i>B. Teyssendier de la Serve et M. Trommetter</i>	

Glossaire – <i>C. Meyer et J.-F. Morot-Gaudry</i>	535
--	-----

Index	549
--------------------	-----

Liste des auteurs	579
--------------------------------	-----

Préface

L'entrée dans le nouveau millénaire a coïncidé avec la publication de la séquence complète du génome d'*Arabidopsis*. Bien que de nombreux travaux de génomique aient précédé et permis cette avancée majeure dans le domaine végétal, cette date représente un tournant dans nos approches de la biologie végétale. En effet, il est devenu pour la première fois possible d'effectuer un inventaire exhaustif des gènes présents dans le génome nucléaire d'une plante supérieure, et avec cet inventaire de mettre en œuvre des approches systématiques pour identifier leur structure et leur fonction. Parmi ces approches de biologie à grande échelle citons la constitution d'une liste complète de gènes correctement annotés. La connaissance des séquences des gènes permet, pour chacun d'entre eux, d'identifier une séquence transcrite spécifique qui peut être utilisée pour la construction de *microarrays* après amplification PCR (projet européen CATMA) ou synthèse d'oligonucléotides spécifiques (puces Affymetrix).

Cette annotation n'est pas parfaite (66 % des gènes sont correctement prédits chez *Arabidopsis*) et sa validation définitive nécessite l'identification d'ADNc pleine longueur correspondant à chaque gène. Un programme international auquel le Génoscope en France, le RIKEN au Japon et le Salk Institute aux États-Unis contribuent, en parallèle à des initiatives privées, a pour objectif de finaliser cette annotation. Cette ressource d'ADNc complets ouvre elle-même la possibilité d'effectuer des travaux systématiques d'analyse fonctionnelle des protéines correspondantes (localisation intracellulaire de la protéine codée par un gène, exploitation des techniques de double hybride pour la recherche de protéines interactives...) et de biologie structurale (expression de la protéine étudiée en système hétérologue pour la production de cristaux permettant la détermination de la structure 3D de cette protéine).

La connaissance de la séquence du génome ouvre d'autres possibilités dans le domaine de la génétique inverse et de l'analyse fonctionnelle. Il devient possible d'inventorier les sites d'insertion de fragments d'ADN de séquence connue (transposons, rétrotransposons, ADN-T) qui constituent autant d'étiquettes pour les gènes qu'ils inactivent en s'insérant au hasard dans le génome. Ainsi plus de 200 000 sites d'insertion dans le génome d'*Arabidopsis* ont-ils été répertoriés, permettant de rechercher des mutants affectés dans un gène d'intérêt particulier grâce aux ressources bioinformatiques constituées à l'URGV dans le cadre d'un programme Génoplante, au Salk Institute (programme Signal), et au MPIZ dans le cadre du programme Gabi. Une approche complémentaire consiste à créer des lignées RNAi pour chacun des gènes identifiés dans le génome d'*Arabidopsis*, afin d'inactiver de manière constitutive ou inductible leur expression (programme européen Agrikola) en générant une série de transgéniques chez lesquels le gène d'intérêt est exprimé à divers niveaux résiduels, ce qui facilite une analyse physiologique de son rôle, même s'il est essentiel à la viabilité de la plante.

Les techniques de spectrométrie de masse viennent maintenant renforcer cet arsenal d'outils génomiques : elles permettent d'identifier des protéines purifiées par des méthodes classiques de biochimie (immunosélection, gels 2D, etc.) rendues compatibles avec le haut débit et de compa-

rer grâce à la bio-informatique la masse de fragments peptidiques issus de la protéine d'intérêt aux masses des protéines du catalogue prédites sur la base de l'annotation du génome.

La génomique apporte donc en conclusion un arsenal d'outils moléculaires qui permettent d'aborder avec efficacité l'analyse de la structure, de l'expression et de la fonction d'un gène, toutes choses qui étaient déjà possibles il y a vingt ans, mais représentaient un investissement expérimental de plusieurs années pour chaque gène étudié. La génomique a supprimé une partie importante des aspects fastidieux de ces approches, en les industrialisant au niveau des procédés. Paradoxalement cette science onéreuse est source d'économies. En effet, en effectuant une fois pour toutes la carte physique d'un génome, sa séquence et l'inventaire des transcrits, le travail des laboratoires utilisateurs est considérablement écourté et simplifié, une signature de séquence étant suffisante à confirmer la nature d'un gène ou d'une protéine. Elle a de fait permis de réaliser des économies importantes sur le coût des programmes de génétique moléculaire.

Fait plus inattendu, les travaux de génomique végétale ont progressivement renouvelé aussi les travaux sur les plantes cultivées dont les génomes sont complexes du fait de l'effet cumulé d'évènements de duplication de fragments génomiques de taille variable, allant jusqu'à celle des chromosomes entiers, et du fait de l'envahissement du génome par des éléments mobiles. La structure de génomes aussi complexes que celui du blé peut donc être analysée comme l'illustrent les travaux menés à l'URGV en ce domaine, alors que l'étude moléculaire de ces espèces à grand génome était *a priori* considérée comme trop lourde pour pouvoir être envisagée il y a encore une décennie. Le fil conducteur de la conservation de synténie entre génomes facilite ces approches, donnant à la génomique comparative une importance croissante. Néanmoins chaque espèce garde ses spécificités physiologiques. Ceci impose un nécessaire travail d'analyse fonctionnelle pour identifier les supports moléculaires de caractères agronomiques portant sur des organes et tissus particuliers (tubercule, fruit, bois, nodule...). Le seul recours aux informations extraites de l'analyse des génomes modèles est une démarche réductionniste qui a ses succès mais aussi ses limites. Ce sont des travaux sur génomes de plantes cultivées qui ont permis pour la première fois d'identifier un gène de résistance aux pathogènes, le locus S impliqué dans l'auto-incompatibilité, ou plus récemment le rôle des PPR dans les systèmes de stérilités mâles cytoplasmiques.

La génomique permet désormais d'explorer aussi la diversité biologique de chaque espèce. Le polymorphisme du génome permet de localiser le support moléculaire des caractères quantitatifs qui distinguent deux écotypes. L'analyse de ce polymorphisme chez les plantes cultivées remet en cause beaucoup d'idées reçues : en particulier cette diversité des populations, comme chez l'homme, résulte beaucoup plus de la combinatoire d'allèles que permet la recombinaison que de la variabilité allélique existant à chaque locus ; les haplotypes différents présents dans les ressources génétiques d'une espèce se comptent en général sur les doigts des deux mains. Ceci suggère que les ressources de gènes susceptibles d'améliorer une espèce cultivée sont à chercher beaucoup plus dans les espèces apparentées que dans les ressources intraspécifiques existantes. La génomique comparée a donc un champ d'application immense dans la perspective de l'exploitation des ressources génétiques présentes dans les espèces sauvages apparentées aux plantes cultivées.

Michel Caboche, URGV INRA Évry

Avant-propos

Les progrès fantastiques réalisés ces dix dernières années dans le domaine de la génomique végétale permettent aujourd'hui d'accéder à la connaissance des gènes, de leur fonction et des modalités de leur expression, comme cela est exposé dans la préface. De fait, les végétaux constituent des supports extraordinaires pour appréhender la complexité du fonctionnement du vivant et ses supports héréditaires (génomomes modèles de petite taille, descendances de très grands effectifs, temps de génération très courts...).

Les connaissances actuelles et potentielles générées par cette nouvelle branche des sciences du vivant auront des retombées spectaculaires en amélioration des plantes dans les prochaines années, quelles que soient les méthodes de sélection qui seront utilisées. En effet, à travers la compréhension du fonctionnement des plantes – grâce à l'identification et la caractérisation des gènes impliqués dans les différents processus assurant ce fonctionnement mais aussi par le biais des interactions et modes de régulation de ces derniers – l'améliorateur pourra piloter de façon très précise les innovations à construire, en exploitant la diversité disponible pour les gènes pertinents. Il pourra ainsi contribuer plus efficacement aux grands défis agricoles mondiaux en sélectionnant des variétés capables de maintenir un bon potentiel de production dans des conditions de stress importantes (sécheresse, sols salins, sols pauvres...), des variétés susceptibles de fournir des produits à haute valeur nutritionnelle permettant de surmonter des carences fortes dans certains pays en développement, ou encore contribuant à une agriculture plus respectueuse de l'environnement...

Au-delà de ces exemples en amélioration des plantes, les applications potentielles dans le domaine de la préservation et de la gestion de la biodiversité à l'échelle de l'écosystème sont aussi très prometteuses.

Cet ouvrage pédagogique illustre clairement les technologies qui forment la base des progrès de la génomique, en s'appuyant sur des travaux de recherche conduits sur des plantes modèles et des espèces cultivées. À travers de nombreux exemples, l'ouvrage fait apparaître le développement d'une nouvelle «biologie intégrative» de la plante à laquelle sont progressivement associés – outre des physiologistes, des biochimistes, des généticiens et des bio-informaticiens – des mathématiciens et des physiciens. La mobilisation des compétences de ces derniers est en effet essentielle pour modéliser le fonctionnement dynamique des génomes, à partir des données considérables générées par ces nouvelles technologies d'analyse à haut débit.

L'ouvrage s'adresse aux étudiants, enseignants et chercheurs qui souhaitent approfondir leurs connaissances de base dans ce domaine, à partir de travaux de recherche récents. Sans prétendre à l'exhaustivité, il donne un bon aperçu des applications potentielles de ce savoir pour la connaissance et la gestion des ressources génétiques ainsi que pour l'amélioration des plantes. Il vient incontestablement combler un manque pédagogique dans ce domaine.

Marianne Lefort, DGAP, INRA Versailles

Introduction

Depuis une cinquantaine d'années, de nombreuses évolutions conceptuelles et méthodologiques ont révolutionné l'étude de la biologie. Le paradigme moléculaire a poussé récemment sa logique jusqu'au séquençage complet de plusieurs génomes d'organismes eucaryotes multicellulaires. La biologie végétale tient une place de choix dans cette dynamique et la communauté scientifique française est active dans ce domaine. Tout n'est pas dit avec le séquençage, bien au contraire ! Les possibilités offertes par la génomique fonctionnelle ouvrent une ère de réconciliation entre la biologie des organismes et la biologie moléculaire pour la définition d'une biologie intégrative qui devrait permettre un regard neuf sur la construction du phénotype à partir du génotype en interaction avec son environnement.

C'est au cours de nos quelques années à la tête du département de Biologie végétale de l'INRA que l'idée de ce livre a germé. À partir de cette position privilégiée pour l'observation des changements majeurs en cours dans les approches de la biologie végétale, nous avons voulu rassembler dans cet ouvrage les différents champs relevant de la génomique qui sont en train de transformer en profondeur les recherches conduites en biologie végétale... et au-delà. Grâce aux compétences de nos collègues de l'INRA, du CNRS, du CIRAD et des universités, nous avons pu couvrir un très large éventail des méthodes et des concepts de la génomique, allant des grands programmes internationaux de séquençage aux indispensables outils de la bio-informatique, en passant par les méthodes d'analyse de l'expression des gènes incluant leurs produits métaboliques finaux et leur spécificité tissulaire et/ou cellulaire. Ce savoir naissant permet de comprendre l'intégration de grandes fonctions physiologiques et de programmes de développement. Au-delà, l'établissement de relations étroites entre la génomique et les domaines de la génétique et de l'amélioration des plantes révèle des synergies pour l'utilisation des marqueurs moléculaires, l'analyse de la variabilité génétique ou encore l'étude des caractères quantitatifs. Ce niveau d'approche permet également d'intégrer l'écophysiologie et ouvre de nouvelles perspectives pour la génomique dans le domaine de l'étude des peuplements végétaux.

C'est donc bien à la naissance d'une nouvelle biologie végétale que nous assistons, et nous espérons que ce livre contribuera à une large diffusion de ces nouvelles pratiques¹.

Ce livre s'adresse aux étudiants de fin d'études universitaires ou agronomiques, aux professeurs de l'enseignement supérieur, aux techniciens, aux ingénieurs et scientifiques qui souhaitent acquérir une connaissance rapide, relativement exhaustive et synthétique en génomique végétale.

¹ Voir aussi : D. de Vienne (éd.), *Les marqueurs moléculaires en génétique et biotechnologies végétales*, 1998 et D. Tagu, C. Moussard (éds), *Principes des techniques de biologie moléculaire*, 2^e éd. revue et augmentée, 2003. INRA Éditions, Paris.

Nous exprimons toute notre gratitude et reconnaissance aux auteurs de cet ouvrage qui ont accepté, malgré un emploi du temps très chargé, d'en écrire un chapitre. Nous remercions également H. Chiapello, A. Hénaut et D. de Vienne pour tous les conseils qu'ils nous ont prodigués au cours de la mise en place de ce livre.

Jean-François Briat et Jean-François Morot-Gaudry

**Génomique structurale
et analyse *in silico***

1. Structure physique du génome nucléaire des plantes

M. DELSENY, M. ECHEVERRIA, E. LASSERRE

Différentes méthodes d'analyse du génome existent. Elles se classent en trois grandes catégories :

– *l'analyse de la structure physique* vise à décrire l'organisation de l'ADN en termes de distance physique, en termes de succession de différents types de séquences, et en termes de distribution et d'empaquetage de l'ADN du génome au sein d'un ensemble de chromosomes. La structure physique est mesurée par des distances exprimées en nanomètres (nm) ou en paires de bases (pb).

– *l'analyse de la structure génétique* vise à comprendre comment les gènes sont distribués le long des chromosomes, comment ils ségrègent et se recombinent à chaque génération. L'analyse de la structure génétique conduit à établir des cartes génétiques des chromosomes sur lesquelles les distances sont exprimées en unités de recombinaison (les centiMorgans) qui sont en fait le reflet de l'aptitude à recombiner.

– *l'analyse de l'expression du génome* consiste à analyser les produits de transcription ou transcriptomes, ainsi que les produits de traduction ou protéomes. Le résultat de l'activité des enzymes représentées dans un protéome sur une variété de substrats produit tout un ensemble de molécules organiques qui constituent le métabolome.

Cette représentation de la dynamique du génome est relativement récente et elle est liée à l'essor, au cours des dernières années, de toute une série de technologies de plus en plus résolutive, dont l'aboutissement est le séquençage complet d'un génome.

Ce chapitre décrit les méthodes qui permettent de décrire la structure physique du génome nucléaire des plantes et les enseignements qu'elles ont apporté. La figure 1 décrit les différents niveaux de résolution de la structure physique du génome. Ces méthodes recouvrent la cytogénétique, les méthodes de fractionnement et de cartographies de restriction et l'organisation de la chromatine.

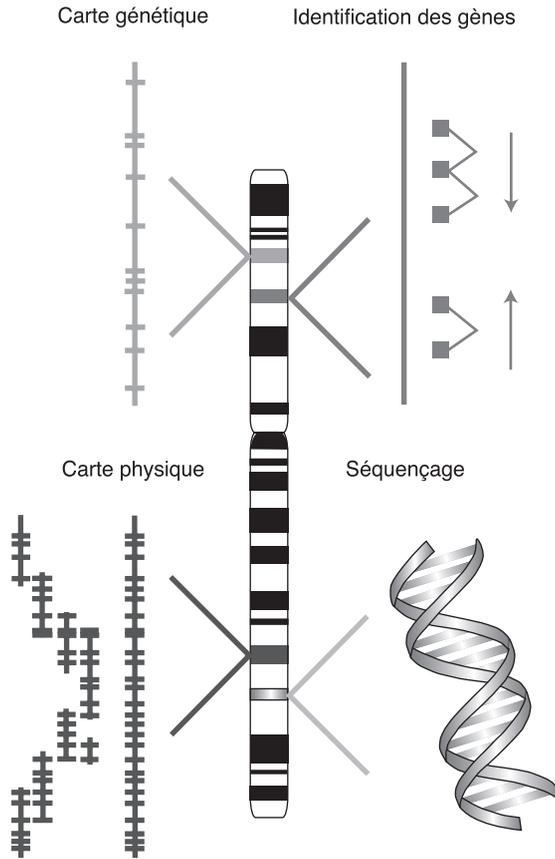


Figure 1. Différentes représentations d'un génome. Au centre représentation d'un chromosome tel qu'il apparaît après révélation des compartiments d'euchromatine et d'hétérochromatine. Sur la carte génétique sont portés des marqueurs moléculaires séparés par des distances génétiques. La carte physique est représentée par un ensemble de grands fragments chevauchants balisés par des sites de restriction. Le déchiffrement de la séquence conduit à identifier les gènes représentés par leur structure exon/intron et le sens de leur transcription.

La cytogénétique classique

Le premier niveau d'étude de la structure physique d'un génome est l'observation microscopique. Elle est réalisée à des stades du cycle cellulaire où les chromosomes sont individualisés (plaque métaphasique des noyaux en mitose) ou décondensés (stades pachytènes et diplotènes de la première division méiotique). On peut alors dénombrer les chromosomes, mesurer leur taille et observer leur morphologie, localiser la constriction centromérique ou l'organisateur nucléolaire. Les différentes régions stratégiques du chromosome peuvent alors être décrites : télomère à chaque extrémité, centromère dont la position va déterminer le type chromosomique : acrocentrique s'il est près d'une extrémité et télocentrique s'il est à peu près centré au milieu du chromosome. Les colorations de la

chromatine comme celle de Giemsa, le carmin acétique ou le DAPI, révèlent souvent des régions plus intensément colorées que d'autres qui constituent l'hétérochromatine ; il a été par la suite démontré qu'elle contient peu de gènes et correspond à des régions chromosomiques fortement condensées et peu actives. L'hétérochromatine, révélée avec le colorant de Giemsa, peut constituer des zonations caractéristiques (le *G-banding*) de chaque chromosome (lorsque ceux-ci sont suffisamment gros) qui permettent de les identifier. Ces premières observations révèlent une hétérogénéité structurale des chromosomes.

La collection de chacun des chromosomes d'une cellule constitue le caryogramme.

Le nombre de chromosomes chez les plantes est toujours pair, car elles sont des organismes diploïdes (ou polyploïdes). Il peut être extrêmement variable : 5 paires chez *Arabidopsis thaliana*, 9 paires chez le chou, 10 chez le maïs, 21 paires chez le blé et plus d'une centaine chez la canne à sucre. La taille des chromosomes est elle aussi extrêmement variable d'une espèce à l'autre, pouvant atteindre jusqu'à plusieurs microns chez des espèces comme le blé, le lis ou la fève. Ces différences suggèrent des contenus en ADN fort différents.

Les différences de taille dans le génome des plantes sont devenues évidentes le jour où l'on a su mesurer la quantité d'ADN par noyau, en utilisant soit des méthodes d'extraction suivies de dosage, soit des méthodes de cytophotométrie. Dans ce dernier type de méthodes, on procède à une évaluation de la teneur en ADN d'un échantillon sur des noyaux isolés (cytométrie en flux) ou sur des préparations microscopiques, après coloration spécifique de l'ADN (réactif de Schiff, DAPI, iodure de propidium...) et comparaison avec un standard interne. Ce type d'expérience a révélé, selon les espèces, des quantités d'ADN par noyaux très variables, allant de quelques centaines de millions de paires de bases à quelques dizaines de milliards (tabl. 1). Cette gamme de variation, sur au moins 2 ordres de grandeur, est sans rapport avec la complexité génétique (le nombre de gènes) dont il est peu probable qu'elle varie d'un facteur supérieur à 5. La quantité d'ADN d'un génome est appelée la « *C-value* » et le grand écart entre les *C-value* et la complexité génétique constitue le paradoxe de la *C-value*. On observe aussi souvent que la teneur en ADN peut varier, au sein d'une même plante, selon le tissu, ou l'organe examiné. Le génome d'une cellule végétale peut en effet subir, au

Tableau 1. Taille des génomes en millions de paires de bases (pb).

<i>Arabidopsis thaliana</i>	130
Riz	430
<i>Medicago truncatula</i>	650
<i>Brassica oleracea</i> (chou)	700
Sorgho	800
Tomate	1 000
Colza	1 200
Maïs	2 500
Orge	5 200
Pois	5 000
Blé	16 000
Tulipe	30 000

cours du développement, des phénomènes d'endoduplication. C'est le cas des cellules de l'albumen de maïs, des trichomes d'*Arabidopsis* et des cellules provasculaires qui peuvent subir jusqu'à 3 ou 4 cycles d'endoréplication. Cependant, les chromosomes polyténiques, comme ceux des glandes salivaires de la drosophile, sont tout à fait exceptionnels chez les plantes et limités aux cellules du suspenseur de l'embryon chez les Légumineuses.

L'observation des appariements chromosomiques au stade pachytène de la méiose révèle souvent des anomalies indicatrices de parentés entre chromosomes. On s'est ainsi rendu compte relativement tôt que de nombreuses espèces végétales cultivées, à comportement apparemment diploïde, étaient en fait des polyploïdes dont le génome résultait d'un doublement d'un génome ancestral (autopolyploïde) ou de l'association de deux génomes distincts (allotétraploïde). L'un des cas les mieux connus est celui du colza qui est en fait un allotétraploïde résultant de la fusion d'un génome de chou et d'un génome de navette. Le blé est également un allohexaploïde résultant de la combinaison de trois génomes A, B et D issus respectivement de trois espèces ancestrales.

Indépendamment de ces phénomènes de ploïdie, qui peuvent rendre compte de variations d'un facteur 2 à 8 dans la teneur en ADN, la cytogénétique permet aussi de révéler des altérations de la structure d'un génome portant sur des fragments de chromosomes, comme les translocations, les inversions, les délétions ou les duplications.

Les méthodes de séparation physique du génome

L'ultracentrifugation en gradient de CsCl à l'équilibre (isopycnique) repose sur la capacité d'un ADN de se positionner dans un gradient à la densité qui correspond à sa densité de flottaison. Celle-ci est strictement corrélée à la composition en paires de bases GC/AT. La détermination de la densité de flottaison à l'aide d'un réfractomètre permet de définir la composition en GC de l'ADN analysé. Les ADN bactériens produisent des pics symétriques, mais les ADN de nombreux eucaryotes présentent des profils d'éluion asymétriques, avec des pics secondaires ou des épaulements que l'on a qualifiés d'ADN satellite (fig. 2). Quelle que soit la nature de l'ADN satellite, établie ultérieurement par d'autres méthodes, dans tous les cas ce type d'observation indique une distribution hétérogène de l'ADN par rapport à une teneur moyenne en paires de bases GC. Des méthodes voisines, utilisant des drogues ou des colorants se fixant plus spécifiquement sur les régions riches en GC ou riches en AT permettent d'accentuer la séparation des différentes fractions d'ADN en fonction de leur composition nucléotidique. À ce stade de l'analyse, on sait donc que le génome des plantes se répartit entre un nombre variable de chromosomes, que sa taille peut varier sur au moins deux ordres de grandeur selon les espèces et que la distribution des différentes séquences d'ADN est hétérogène, en fonction de leur composition nucléotidique.

L'analyse de l'ADN par les méthodologies des cinétiques de dénaturation/renaturation

Cette méthodologie trouve son principe dans la nature double brin de la molécule d'ADN. Lorsque l'on chauffe une solution d'ADN dans un tampon à 0,18 M de phosphate