

POURQUOI CLONER LES GÈNES DES PROTÉINES DE RÉSERVE DU BLÉ

Philippe JOUDRIER*, Jean-Claude AUTRAN*, Marie-Françoise GAUTIER*

Pouvoir disposer d'outils simples et puissants pour maîtriser la qualité boulangère ou pastière d'un blé est à l'ordre du jour. En effet, l'évolution des technologies de transformation et la concurrence internationale exigent aujourd'hui du sélectionneur la mise rapide sur le marché de variétés non seulement productives, mais de qualité sans cesse améliorée. Les progrès réalisés dans les méthodes d'analyse biochimique ont déjà permis de mettre en évidence, parmi les protéines de réserve, des marqueurs génétiques en relation simple, directe ou indirecte, avec certaines propriétés de viscoélasticité de la pâte. Entre autres informations, la biologie moléculaire offre la possibilité de déterminer la structure de ces protéines. La mise en évidence de relations entre ces structures et les propriétés physico-chimiques de la pâte ramènerait le problème complexe qu'est la notion de qualité à un problème beaucoup plus simple.

Depuis l'après-guerre, le rendement moyen du blé en France a augmenté de 1,2 quintal par hectare et par an. De 18 quintaux à l'hectare en 1950, il dépasse 60 quintaux à l'hectare actuellement. Cette évolution, due en partie au progrès génétique (grâce à la création de nouvelles variétés) ne s'est pas faite au détriment de la qualité. La force boulangère (1), mesurée par un critère classique tel que l'indice W de l'Alvéographe Choppin, est passée, au cours de la même période, de 80-90 à 180-200 (2). Mais la tâche des généticiens et sélectionneurs français n'est

pas terminée. Elle doit au contraire être impérativement poursuivie, en raison d'une part, de l'évolution des technologies de panification (fabrication industrielle, réfrigération ou surgélation des pâtes) qui réclament des farines de plus en plus « fortes » et, d'autre part, de la concurrence internationale. Le rendement et la qualité progressent aussi à l'étranger de telle sorte que de nos jours, toute variété se trouve dépassée en peu d'années. Les agriculteurs français ne manqueraient pas de s'approvisionner en semences auprès des pays étrangers si ces derniers parvenaient à créer des variétés plus performantes que les nôtres. Dans un domaine stratégique dont dépend l'indépendance alimentaire (comme le rappellent J. Grall et B.H. Lévy dans « La guerre des semences ») (3), celui qui possèdera les semences les plus performantes détiendra certainement un immense pouvoir.

Créer et sélectionner des variétés de blé à la fois très productives, d'une qualité d'utilisation sans cesse améliorée, et adaptées aux besoins des industries utilisatrices et des consommateurs, constitue donc un objectif permanent pour la filière céréalière française.

Mais un tel objectif est difficile et coûteux

à atteindre (*voir encadré 1*). Il faut, en effet, créer des plantes toujours plus performantes ; il faut parvenir à associer sur un même génotype des caractères antagonistes tels qu'un rendement élevé et, par exemple, une haute teneur en protéines ; il faut enfin retrouver, parmi les milliers de descendants d'un croisement, la plante ou le grain qui renferme la meilleure combinaison rendement-qualité !

Il n'est guère possible de mesurer la valeur technologique par des méthodes directes classiques (fabrication de pain, de biscuits ou de pâtes alimentaires) au cours des premiers stades de la sélection, du fait du très grand nombre de plantes à analyser et des faibles quantités de grains disponibles. Le sélectionneur a donc dû s'orienter vers des micro-méthodes (modèles réduits de pains ou de pâtes alimentaires, par exemple à partir de 10 grammes de farine ou de semoule) ou des procédés indirects (gonflement de la farine dans certains solvants, quantité et viscoélasticité du gluten, résistance de la pâte à la pression ou au pétrissage). Mais les difficultés demeurent, liées à la complexité de la notion de qualité : non seulement la qualité est un caractère polygénique, mais elle est aussi fonction à la fois de la variété, des condi-

* Laboratoire de Technologies des Céréales, INRA, 9 place Viala, 34060 Montpellier Cedex.

(1) Composante de la qualité boulangère. Le W représente le travail de déformation de la pâte lorsqu'elle est soumise à une pression d'air.

(2) H. Feyt, Semences et variétés de l'an 2000. Perspectives Agricoles (1987) 111, 2, 48-56.

(3) J. Grall et B.H. Lévy, La guerre des semences : quelles moissons, quelles sociétés ? Ed. Fayard (1985).

ENCADRÉ 1 - LES CONTRAINTES DE LA SÉLECTION

Bien que les découvertes techniques et génétiques tendent actuellement à raccourcir les délais de création de nouveaux cultivars, il faut environ dix ans et parfois quinze pour mettre une nouvelle plante sur le marché (1). Pour le sélectionneur, cela signifie dépenser beaucoup d'argent sans avoir la certitude du succès. Le sélectionneur a par ailleurs la tâche extrêmement malaisée de devoir réunir sur le même génotype des caractères souvent antagonistes tels qu'une production élevée et régulière (une variété ne connaîtra de succès que si elle est au « top niveau » au plan agronomique) et une qualité technologique correspondant à la demande des utilisateurs.

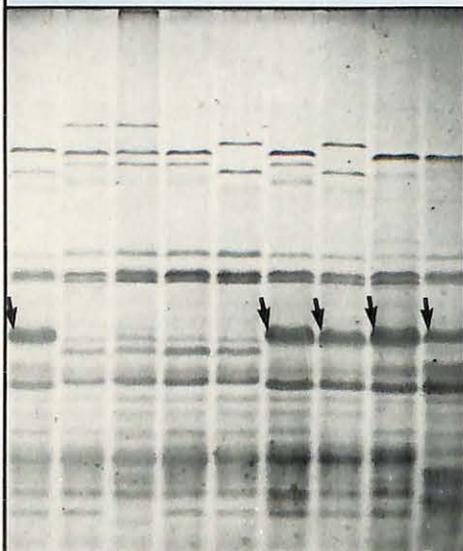
Pour le blé tendre (*Triticum aestivum* L., espèce hexaploïde), dont la principale utilisation est la transformation en farine et en pain, c'est naturellement une qualité boulangère élevée qui est recherchée, c'est-à-dire une aptitude à donner une pâte machinable et un pain léger et bien développé. Pour le blé dur (*Triticum durum* Desf., espèce tétraploïde), qui est réservé à la fabrication de semoules et de pâtes alimentaires ou de couscous, il s'agit de sélectionner des variétés de qualités pastière et culinaire élevées, c'est-à-dire donnant des pâtes alimentaires de couleur jaune ambré qui restent fermes et non collantes au cours de la cuisson.

Mais comment prévoir la demande des industries utilisatrices dix ans à l'avance sachant que les technologies évoluent ? Le sélectionneur doit, de plus, concilier cette situation avec le fait qu'il ne peut travailler que sur des objectifs pérennes ...

(1) Y. Demarly, L'impact possible des biotechnologies sur les semences de l'an 2000. Industries Agricoles et Alimentaires (1987) 104, 1, 15-21.

FIGURE 1 - DIAGRAMMES ÉLECTROPHORÉTIQUES (SDS-PAGE) DES PROTÉINES TOTALES RÉDUITES DE 10 VARIÉTÉS DE BLÉ DUR

Les flèches indiquent (chez les variétés possédant l'allèle « 45 ») la principale des gluténines LMW associées à la qualité du gluten. Les autres diagrammes correspondent aux variétés possédant l'allèle « 42 »



tions de culture et de l'interaction entre le génotype et le milieu (4). Comment alors faire la part de ce qui revient au seul génotype, unique moyen de progresser dans l'amélioration génétique ?

Les tests biochimiques : des outils puissants

La biochimie classique offre actuellement des outils très puissants d'appui à la sélection, notamment pour la prédiction, à un stade très précoce, de la valeur potentielle des génotypes. Ainsi, certaines méthodes d'étude des protéines (électrophorèse, HPLC) présentent d'énormes avantages en permettant des analyses en grandes séries sur des micro-quantités d'échantillons (et donc à un stade très précoce de la sélection) ; leurs résultats ne sont pas influencés par les facteurs agro-climatiques et permettent donc d'approcher la qualité intrinsèque des génotypes ; de plus, le déterminisme génétique des constituants protéiques étudiés est plus simple que celui des critères technologiques.

L'exemple le plus significatif qui puisse être donné est la relation très étroite, mise en évi-

dence chez le blé dur (*Triticum durum* Desf.), entre les propriétés viscoélastiques du gluten - paramètre important de la qualité culinaire des pâtes alimentaires - et l'électrophorogramme en gel de polyacrylamide des gliadines du grain (5). La présence du type allélique « gliadine 45 » permet au sélectionneur de repérer, même à partir d'un demi-grain prélevé à la génération F2, un génotype présentant un potentiel élevé de qualité culinaire. Au contraire, la présence du type « gliadine 42 » indique un faible potentiel de qualité. Une relation équivalente a été observée en électrophorèse sur gel de polyacrylamide-SDS (voir figure 1) avec certaines gluténines de faible poids moléculaire, génétiquement liées aux gliadines 45 ou 42. Or les gluténines auraient l'avantage d'être de véritables marqueurs fonctionnels de la qualité, alors que les gliadines ne seraient que des mar-

(4) G. Branlard et J.C. Autran, L'amélioration génétique de la qualité technologique du blé tendre. Culture Technique (1986) 16, 132-144.

(5) R. Damidaux, J.C. Autran and P. Feillet, Intrinsic cooking quality evaluation in durum wheats through examination of gliadin electrophoregrams and measurements of gluten viscoelasticity. Cereal Foods World (1980) 25, 754-756.

ENCADRÉ 2 - BIOSYNTÈSE ET DÉTERMINISME GÉNÉTIQUE DES PROTÉINES DE RÉSERVE DU BLÉ

L'albumen du grain de blé est le site de la synthèse et de l'accumulation d'une série de protéines de réserve qui s'identifient, pour la plupart, aux protéines du gluten. Ces protéines de réserve sont de type sécrétoire : elles sont synthétisées au niveau de polysomes liés à la membrane du réticulum endoplasmique, puis accumulées dans des organites cellulaires (corpuscules protéiques) spécialisés dans le stockage des protéines de réserve, le passage de la membrane s'accompagnant d'une modification post-traductionnelle : l'excision du peptide signal (1).

Ces protéines sont classées en deux groupes : les gliadines alcool-solubles et peu agrégatives d'une part ; d'autre part les gluténines, alcool-insolubles et très agrégatives (elles forment des complexes de poids moléculaire de l'ordre de plusieurs millions). Ces gluténines renferment deux classes de polypeptides : ceux de haut poids moléculaire ou HMW (70-120 kD) et ceux de faible poids moléculaire ou LMW (40-55 kD).

Les gliadines et les gluténines LMW sont codées par des gènes situés sur le bras court des chromosomes du groupe 1 d'homéologie (loci *Gli-A1*, *Gli-B1* et *Gli-D1*) et du groupe 6 (loci *Gli-A2*, *Gli-B2* ou *Gli-D2*). Les gluténines HMW sont codées par des gènes situés sur le bras long des chromosomes du groupe 1 (loci *Glu-A1*, *Glu-B1* et *Glu-D1*).

Ces différents loci correspondent à des familles multi-géniques considérées comme ayant divergé d'un gène ancestral par des processus de duplication, de délétion,

de mutation ponctuelle ou de crossing-over asymétrique (2). Les protéines codées au niveau de ces loci apparaissent dans les électrophorogrammes sous la forme de groupes ne donnant jamais lieu à des recombinaisons et paraissant ainsi se transmettre en bloc. Ces protéines présentent des homologies dans leur séquence N-terminale, y compris entre des espèces différentes comme les progéniteurs sauvages du blé, ou le seigle et l'orge (3).

Plusieurs allèles (de 2 à 12) ont été identifiés pour chacun des 9 loci chromosomiques des protéines de réserve du blé tendre. Après croisement entre deux variétés, l'électrophorogramme des protéines d'une lignée descendante représente une recombinaison des 9 « blocs » alléliques de chacun des parents, ce qui explique le nombre considérable des types électrophorétiques trouvés dans les blés actuels.

(1) T. Tercé-Laforgue et J.C. Pernollet, Evidence for a peptide signal in nascent wheat gliadins. *Physiol. Vég.* (1984) 22, 721-731.

(2) D.D. Kasarda, D. Lafiandra, R. Morris and P.R. Shewry, Genetic relationships of wheat gliadin proteins. *Kulturpflanze* (1984) 32, 41-60.

(3) D.D. Kasarda, J.C. Autran, E.J.L. Lew, C.C. Nimmo and P.R. Shewry, N-terminal amino acid sequences of omega-gliadins and omega-secalins. Implication for the evolution of prolamin genes, *Biochimica and Biophysica Acta* (1983) 747, 138-150.

queurs indirects (liaison génétique) (voir encadrés 2 et 3).

Alors qu'il est généralement très difficile de relier une caractéristique technologique ou agronomique à un simple marqueur biochimique, l'exemple précédent est d'une remarquable simplicité. Il constitue un modèle extrêmement précieux pour des approches à la fois physico-chimiques et moléculaires des propriétés fonctionnelles des protéines.

Les apports de la biologie moléculaire

Lorsque le sélectionneur disposera de tests pour cribler facilement plusieurs milliers de lignées par an, un progrès génétique substantiel en matière de qualité pourra être réalisé. C'est pourquoi l'utilisation de méthodes biochimiques classiques doit être poursuivie et même intensifiée, en mettant notamment à profit des outils encore plus performants et surtout automatisés de description de la variabilité génétique (HPLC, électrophorèse rapide avec informatisation de la saisie et du traitement des données). Mais il nous paraît indispensable d'aller plus loin encore.

Il s'agit en effet non seulement de constater et d'utiliser les corrélations entre marqueurs biochimiques et qualité, mais aussi de **comprendre la nature fondamentale de ces relations**. Il est clair qu'on maîtrisera d'autant mieux les problèmes de qualité qu'on saura mieux les comprendre et les expliquer en termes biochimiques et physico-chimiques.

On sait que la qualité boulangère d'un blé repose en grande partie sur les caractéristiques rhéologiques (extensibilité, ténacité, élasticité) de la pâte ou du gluten. Celles-ci dépendent essentiellement des propriétés des protéines : structure, interactions, capacité d'agrégation (voir encadré 3). On ne connaît cependant pas avec certitude les bases physico-chimiques de la qualité d'un blé, et aucun modèle de structure ne fait actuellement l'unanimité. On ne sait pas exactement quel est le rôle respectif des différentes familles protéiques (telles que les gluténines HMW et LMW) dans les propriétés d'extensibilité ou d'élasticité de la pâte. Des corrélations entre la présence de certains polypeptides et des données technologiques ont été découvertes, mais on n'a encore jamais relié

ENCADRÉ 3 – LE GLUTEN, ÉLÉMENT-CLÉ DE LA QUALITÉ DU BLÉ.

Les protéines de l'albumen du blé (et, dans une moindre mesure, celles du seigle) possèdent la propriété - unique dans le règne végétal - de pouvoir s'associer au cours du pétrissage de la pâte et de former un complexe visco-élastique : le gluten.

Le gluten correspond essentiellement aux protéines de réserve de l'albumen du grain : gliadines, gluténines LMW, gluténines HMW, dans des proportions respectives d'environ 50/35/15, ainsi que certaines protéines de structure (membranaires) et des lipides.

Ce sont vraisemblablement les gluténines (HMW et LMW) qui sont responsables des propriétés de visco-élasticité du gluten ou des pâtes obtenues à partir de farine de blé. Schématiquement, une forte tendance de ces protéines à s'associer, à s'agréger correspond à une qualité (boulangère ou pastière) élevée. Selon plusieurs modèles classiques, la structure du gluten reposerait sur l'établissement de ponts disulfures entre les résidus cystéines de polypeptides différents, d'où la formation de très longs polymères linéaires peu ou non ramifiés. La viscoélasticité serait liée à des réactions d'échanges de liaisons disulfures au cours du glissement des molécules, propriété qui disparaît d'ailleurs sous l'action d'agents réducteurs. D'autres modèles font cependant intervenir les liaisons non covalentes, les résidus glyci-

nes localisés dans le domaine central des molécules, ou encore certaines classes lipidiques se liant préférentiellement aux régions hydrophobes des protéines du gluten.

Des comparaisons intervariétales ont montré que la variation allélique considérable qui existe au niveau de chaque locus chromosomique pouvait affecter les propriétés des agrégats gluténines. En effet, toutes les protéines n'ont pas les mêmes propriétés fonctionnelles : des variétés de bonne qualité semblent renfermer des protéines ayant une forte tendance à s'agréger, contrairement aux variétés de mauvaise qualité. Chez le blé tendre, la présence ou l'absence de certains variants des gluténines HMW a pu être corrélée à des différences de qualité boulangère. Chez le blé dur, il semblerait au contraire que ce soient les gluténines LMW (génétiquement liées aux gamma-gliadines) qui jouent un rôle prépondérant dans la qualité du gluten et des pâtes alimentaires. D'où l'intérêt pour le généticien de déterminer la capacité d'agrégation des protéines (par chromatographie ou HPLC) ou d'identifier (par électrophorèse) les protéines qui sont à l'origine des propriétés fonctionnelles favorables de façon à les introduire par croisement (ou à plus long terme par manipulation génétique) dans les variétés cultivées.

avec certitude un type de structure et une propriété fonctionnelle. On ne connaît pas non plus les mécanismes de régulation de la biosynthèse des différentes protéines, ni les facteurs qui influencent favorablement ou défavorablement l'expression de la qualité au sein d'un génotype.

Pour tenter d'appréhender ces structures et ces propriétés fonctionnelles, on a utilisé jusqu'ici des méthodes descriptives, au niveau des constituants protéiques du grain de blé ou de la farine : solubilisation, électrophorèse, électrofocalisation, chromatographie... Mais pour atteindre les mécanismes intimes mis en jeu, il apparaît aujourd'hui nécessaire de remonter aux gènes eux-mêmes, à leur structure, leur localisation, leur fonctionnement. Seules les études moléculaires au niveau de l'ADN seront, en effet, en mesure de fournir l'explication complète des phénomènes observés, pour permettre un jour d'intervenir sur ces derniers. Le nouvel objectif est donc de cloner, d'isoler et de séquencer les gènes des protéines directement impliquées dans la qualité : les gènes de structure qui codent pour ces protéines, mais aussi les gènes de régulation qui modulent leur expression.

La démarche doit être analogue à celle retenue pour d'autres problèmes : résistance des plantes aux herbicides, au froid, au milieu salin ou aux agresseurs biologiques (6), modifica-

tion de la composition en acides aminés en vue de rééquilibrer la valeur nutritionnelle (7, 8).

Stratégie d'isolement d'un gène

En fonction des informations disponibles sur la protéine dont on veut isoler le gène, deux approches complémentaires peuvent être utilisées simultanément. En effet, dans la cellule l'information génétique peut être utilisée sous deux formes : l'ADN génomique qui contient la totalité de l'information, c'est-à-dire les séquences codantes, les séquences régulatrices et les séquences non codantes, et l'ARN messager (ARNm) qui correspond uniquement aux gènes exprimés dans une cellule à un moment donné. Ces deux formes peuvent être clonées à l'aide de deux stratégies différentes (voir figure 2).

(6) P. Feillet, Les biotechnologies au service de l'agriculture et des agro-industries. Futuribles (septembre 1986) 23-37.

(7) A.S. Rao and R. Singh, Improving grain protein quality by genetic engineering: some biochemical considerations. Tibtech, Elsevier Science Publishers BV. (1986) Amsterdam, pp. 108-109.

(8) C. Hedgcoth, Why clone wheat storage protein genes? Cereal Foods World (1985) 30, 11, 781-783.

Construction d'une banque d'ADN complémentaire (ADNc) de l'ARNm

Les protéines de réserve des céréales sont synthétisées uniquement dans le grain, au cours de la maturation. Pour construire une banque d'ADNc, les ARNm sont isolés de grains en cours de maturation (entre 15 et 20 jours après l'anthèse), période pendant laquelle se situe la synthèse maximale des protéines de réserve. Cette population d'ARNm correspond donc uniquement aux gènes exprimés dans le grain à un moment bien précis de son développement et est de ce fait fortement enrichie : à cette période, environ 70% des ARNm codent pour des protéines de réserve.

Après isolement et fractionnement (par exemple, selon la taille), les ARNm sont traduits dans un système acellulaire (germe de blé). Les produits de traduction sont ensuite caractérisés par électrophorèse (SDS-PAGE ou bidimensionnelle) afin de repérer les fractions protéiques auxquelles correspondent les ARNm isolés.

Grâce à une enzyme (transcriptase réverse), il est possible d'obtenir, à partir d'ARNm, des ADNc qui sont tout aussi capables de s'hybrider avec le gène correspondant que les ARNm eux-mêmes. Ces ADNc sont ensuite introduits dans des bactéries par l'intermédiaire de plasmides, ce qui permet d'une part de les produire en grande quantité et, d'autre part, de les isoler les uns des autres : la population d'ARNm se trouve en effet convertie en une population de clones. Ces clones peuvent être caractérisés selon leur taille, leur carte de restriction et, pour les plus intéressants, leur séquence. Si un ADNc se révèle coder pour une protéine que l'on recherche, il peut être marqué radioactivement. La sonde radioactive ainsi obtenue peut être utilisée pour cribler une banque génomique.

Construction d'une banque génomique

Une banque génomique est constituée d'une population de clones porteurs de fragments d'ADN représentatifs de la totalité du génome de la plante. Elle renferme donc le gène recherché sous sa forme complète (séquence codante et séquences régulatrices).

La préparation de cette banque nécessite plusieurs étapes successives :

— l'ADN génomique est isolé à partir de feuilles, racines, embryons ou tout autre organe de la plante ;

— l'ADN est découpé en gros fragments (15 à 20 kb) à l'aide d'une enzyme de restriction ;

— ces fragments d'ADN sont recombinaés, en général chez un phage lambda, ce qui permet d'avoir une population de clones. Sachant que chez le blé (*Triticum aestivum* L.) le génome haploïde correspond à $5,4 \times 10^8$ pb, il faut environ 3 à 4 millions de clones pour que la banque génomique soit représentative.

Pour isoler le gène recherché, la banque génomique est ensuite criblée avec la sonde d'ADNc marquée radioactivement. Dans le cas favorable où l'on connaît la séquence (ou une partie de la séquence) de la protéine, une autre alternative consiste à synthétiser des oligonucléotides correspondant à un fragment de la séquence protéique, ceux-ci étant ensuite utilisés pour le criblage de la banque.

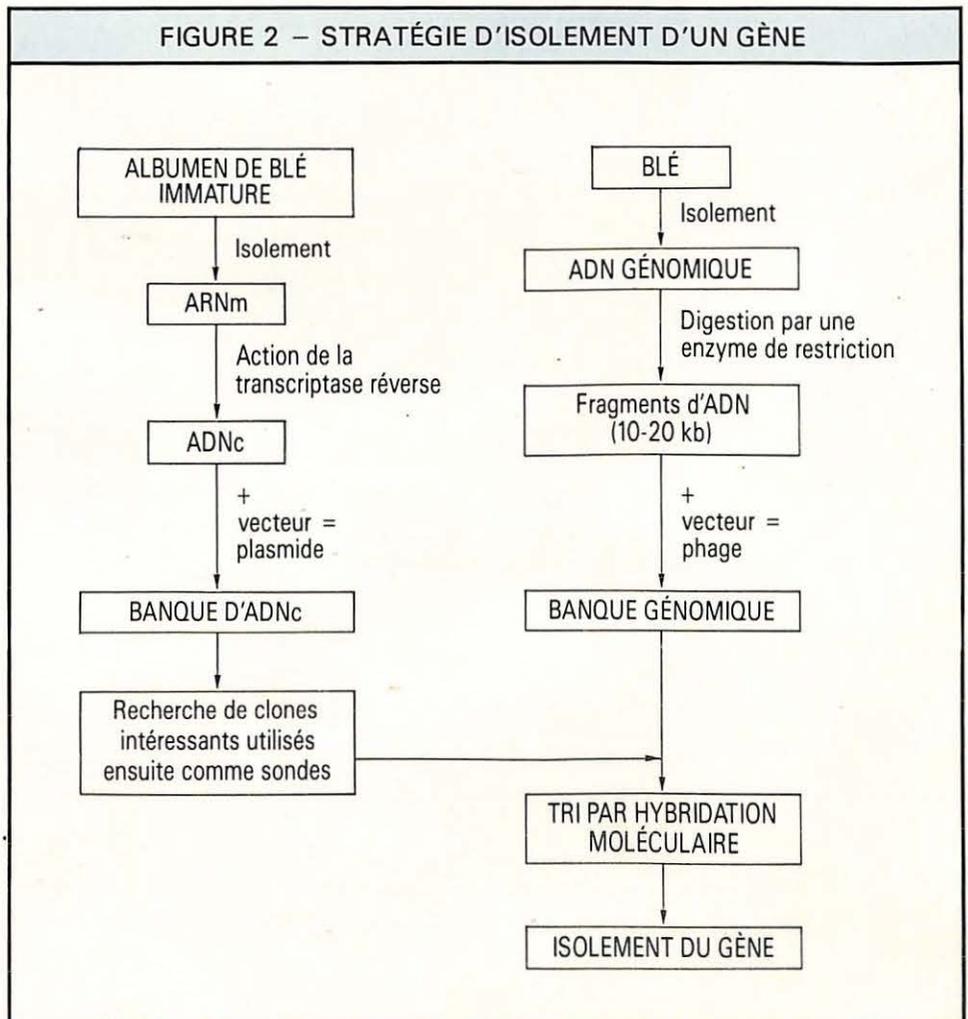
Les clones génomiques qui contiennent la séquence d'ADN complémentaire de celle de la sonde s'hybrident avec cette dernière et une autoradiographie permet ensuite de les identifier. Les clones génomiques retenus sont ensuite sous-clonés en fragments plus petits, en général dans des plasmides, ce qui permet de les étudier plus facilement (cartes de restriction, séquençage).

Principaux résultats attendus

L'isolement et la caractérisation des gènes codant pour les protéines de réserve du blé apporteront des informations à différents niveaux.

Tout d'abord, au niveau de la structure même des gènes : présence ou absence

FIGURE 2 - STRATÉGIE D'ISOLEMENT D'UN GÈNE



d'introns (9) (aucun intron n'a été mis en évidence dans les gènes des protéines de réserve du blé séquencés à ce jour), présence des séquences consensus (10) en amont et en aval de la séquence codante.

On aura également accès à la structure primaire des protéines codées par ces gènes, qui est l'information fondamentale recherchée.

On sait que la détermination de la séquence complète des acides aminés des protéines est un travail extrêmement long et difficile, parfois impossible chez les protéines de réserve des céréales, en raison soit du blocage du résidu N-terminal, soit de la présence de domaines de polyglutamine (jusqu'à 18 résidus Gln consécutifs) au coeur de la séquence. Par contre, la détermination de la séquence nucléotidique d'un gène, constitue un moyen infiniment plus rapide (et souvent le seul moyen) de connaître la séquence de la protéine.

A ce jour, plusieurs gènes codant pour des

figure 3. Sa séquence codante est organisée en 5 régions de composition différente en acides aminés.

En ce qui concerne les gènes de gluténines, les seules séquences publiées correspondent à des gluténines HMW (15, 16, 17), dont l'organisation de la séquence codante est différente de celle des gliadines. Trois régions peuvent être définies : une région N-terminale non répétitive, une région centrale répétitive représentant la majeure partie de la séquence et une région C-terminale non répétitive.

Aucune séquence n'a été déterminée sur les gluténines LMW ni, d'une façon générale, sur les protéines du blé dur (*Triticum durum* Desf.).

Ces séquences, complétées par des prédictions de structure secondaire ou tertiaire, sont susceptibles de révéler des informations précieuses pour comprendre le rôle des différentes protéines dans la valeur technologique.

la possibilité de formation de structure bêta (18) comme cela a été montré pour l'élastine (protéine du tissu conjonctif). Or les bases moléculaires des propriétés élastiques de l'élastine sont attribuées à la présence de séquences riches en glycine.

Une autre caractéristique intéressante de la structure des gluténines HMW est que les résidus cystéines sont situés aux extrémités N- et C-terminales de la chaîne polypeptidique (15, 17). Elle renforce l'hypothèse selon laquelle il y aurait formation, dans le gluten, de polymères dans lesquels les gluténines HMW - dont la partie centrale de la molécule serait sous forme de structure élastique bêta - seraient liées par des ponts disulfures via les résidus cystéines situés à leurs extrémités. Cela n'est cepen-

(9) Dans un gène eucaryote, un intron (opp : exon) représente en général les parties non codantes. Il est excisé lors de l'épissage.

(10) Séquence « idéalisée » dans laquelle chaque position représente la base rencontrée le plus fréquemment, dans un ensemble de séquences plus ou moins homologues. Elle est établie après comparaison de séquences réelles.

(11) D.D. Kasarda, T.W. Okita, J.E. Bernardin, P.A. Baecker, C.C. Nimmo, E.J.L. Lew, M.D. Dietler and F.C. Greene, Nucleic acid (cDNA) and amino acid sequences of alpha-type gliadin from wheat (*Triticum aestivum*), Proc. Natl. Acad. Sci. USA (1984) 81, 4712-4716.

(12) O.D. Anderson, J.C. Litts, M.F. Gautier and F.C. Greene, Nucleic acids sequence and chromosome assignment of a wheat storage protein gene. Nucleic Acids Research (1984) 12, 21, 8129-8144.

(13) T. Sugiyama, A. Rafalski and D. Soll, The nucleotide sequence of a wheat gamma-gliadin genomic clone. Plant Science (1986) 44, 205-209.

(14) J.A. Rafalski, Structure of wheat gamma-gliadin genes. Gene (1986) 43, 221-229.

(15) J. Forde, J.M. Malpica, N.G. Halford, P.R. Shewry, O.D. Anderson, F.C. Greene and B.J. Mifflin, The nucleotide sequence of a HMW glutenin subunit gene located on chromosome 1A of wheat (*Triticum aestivum* L.). Nucleic Acids Research (1985) 13, 19, 6817-6832.

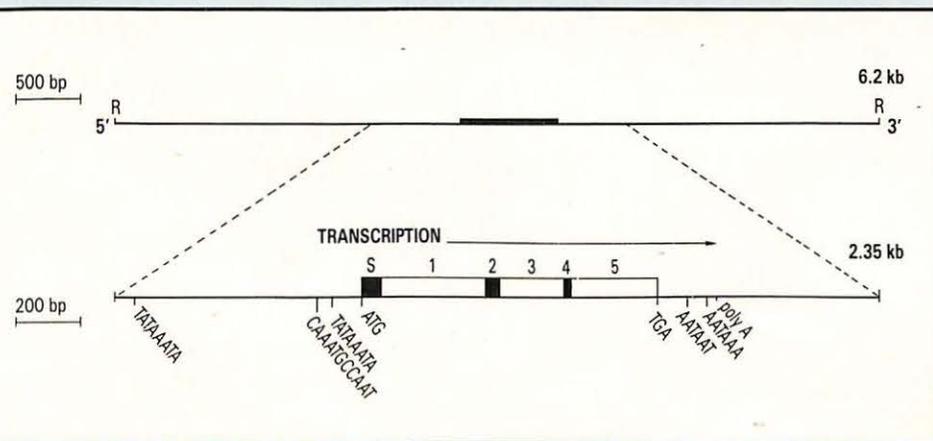
(16) R.D. Thompson, D. Bartels and N.P. Harberd, Nucleotide sequence of a gene from chromosome 1D of wheat encoding a HMW-glutenin subunit. Nucleic Acids Research (1985) 13, 9, 6833-6846.

(17) T. Sugiyama, A. Rafalski, D. Peterson and D. Soll, A wheat HMW glutenin subunit gene reveals highly repeated structure. Nucleic Acid Research (1985) 13, 24, 8729-8737.

(18) ou conformation bêta. Structure dans laquelle les chaînes polypeptidiques sont disposées parallèlement, les chaînes latérales étant situées au dessus ou en dessous du plan formé (structure en feuillet plissé).

FIGURE 3 - STRUCTURE D'UN GÈNE DE GLIADINE ALPHA/BÊTA DU BLÉ

La séquence codante de la protéine mature peut être divisée en 5 régions : 1. région N-terminale ; 2. région de polyglutamine ; 3. région centrale à séquence unique ; 4. seconde région de polyglutamine ; 5. région C-terminale à séquence unique. Cette séquence est précédée d'un peptide signal (S) de 20 résidus d'acides aminés dont la composition est caractéristique. Dans la région 5' non codante on retrouve les séquences consensus de type CAAT et TATA associées à l'initiation de la transcription. Dans la séquence 3' non codante, 2 sites de polyadénylation sont présents (d'après F.C. Greene, O.D. Anderson, J.C. Litts and M.F. Gautier, Control of wheat protein biosynthesis, Cereal Chem. 1985, 62, 5, 398-405).



protéines de réserve du blé tendre (*Triticum aestivum* L.) ont été isolés et caractérisés. En ce qui concerne les gliadines, seules des séquences de type alpha/bêta (11, 12) et gamma (13, 14) sont connues. Aucune séquence nucléotidique correspondant à une oméga gliadine n'a été publiée.

A titre d'exemple, la structure d'un gène de gliadine de type alpha/bêta est donnée sur la

En ce qui concerne plus particulièrement les gluténines, auxquelles on attribue une grande importance dans les propriétés viscoélastiques d'une pâte ou d'un gluten de blé (encadré 3), il est intéressant de noter la teneur élevée en glycine (15-20% des gluténines HMW). De plus, ces résidus glycines ne sont pas distribués au hasard dans la molécule mais sont situés dans la région centrale répétitive. Cela suggère

dant plus vrai pour une autre séquence gluténine HMW qui présente un codon cystéine dans sa partie centrale (16). Ces différences peuvent-elles être impliquées dans les différences de propriétés viscoélastiques ? Seule la comparaison d'un plus grand nombre de séquences pourra fournir des éléments de réponse.

Puisque la présence de certains allèles a été corrélée (positivement ou négativement) à des caractéristiques de qualité technologique, il est fondamental de cloner les gènes correspondant à ces allèles et de **découvrir la signification physico-chimique des corrélations**. S'il était démontré que les allèles trouvés corrélés à une qualité élevée correspondent précisément à des protéines possédant des propriétés physico-chimiques particulières, un grand pas serait

accompli dans la recherche de relations entre caractéristiques physico-chimiques et propriétés fonctionnelles des protéines, et un lien fondamental serait établi entre les résultats que les généticiens et les physico-chimistes ont obtenus chacun de leur côté.

Le modèle des gènes du locus **Gli-B1** du blé dur correspond tout à fait à ce cas. A l'un des allèles (« gliadine 45 ») correspond un potentiel élevé de qualité du gluten alors qu'à l'autre allèle (« gliadine 42 ») correspond un faible potentiel de qualité.

Une autre approche, à plus long terme, pour étudier les propriétés fonctionnelles de ces protéines consisterait à les faire produire par des microorganismes. Il est en effet très difficile d'isoler par voie biochimique classique un polypeptide pur à partir des familles multigé-

niques, fortement homologues, que sont les protéines de réserve. Grâce à l'introduction et à l'expression d'un gène dans une bactérie ou une levure permettant la production de quantités importantes de la protéine pure correspondante, des études de propriétés fonctionnelles pourraient être effectuées et il serait possible de relier, *in vitro*, structure protéique et fonction. De plus, l'utilisation de techniques de mutagenèse dirigée permettrait de modifier sélectivement des régions codantes et d'observer les effets produits au niveau des propriétés fonctionnelles des protéines ainsi modifiées.

Il serait également intéressant de pouvoir augmenter la teneur en lysine des protéines de réserve des céréales, sans modifier leurs autres propriétés. Des résultats ont déjà été obtenus en ce sens sur le maïs (19). Il faut rappeler que ces protéines de réserve présentent peu de charges électriques et que l'introduction d'acides aminés tels que la lysine pourrait entraîner des modifications importantes de structure pouvant se répercuter sur les propriétés fonctionnelles.

Ces objectifs de compréhension de la qualité sont considérés comme prioritaires. Mais les gènes isolés pourront aussi être utilisés pour des études présentant un caractère plus fondamental ou physiologique telles que :

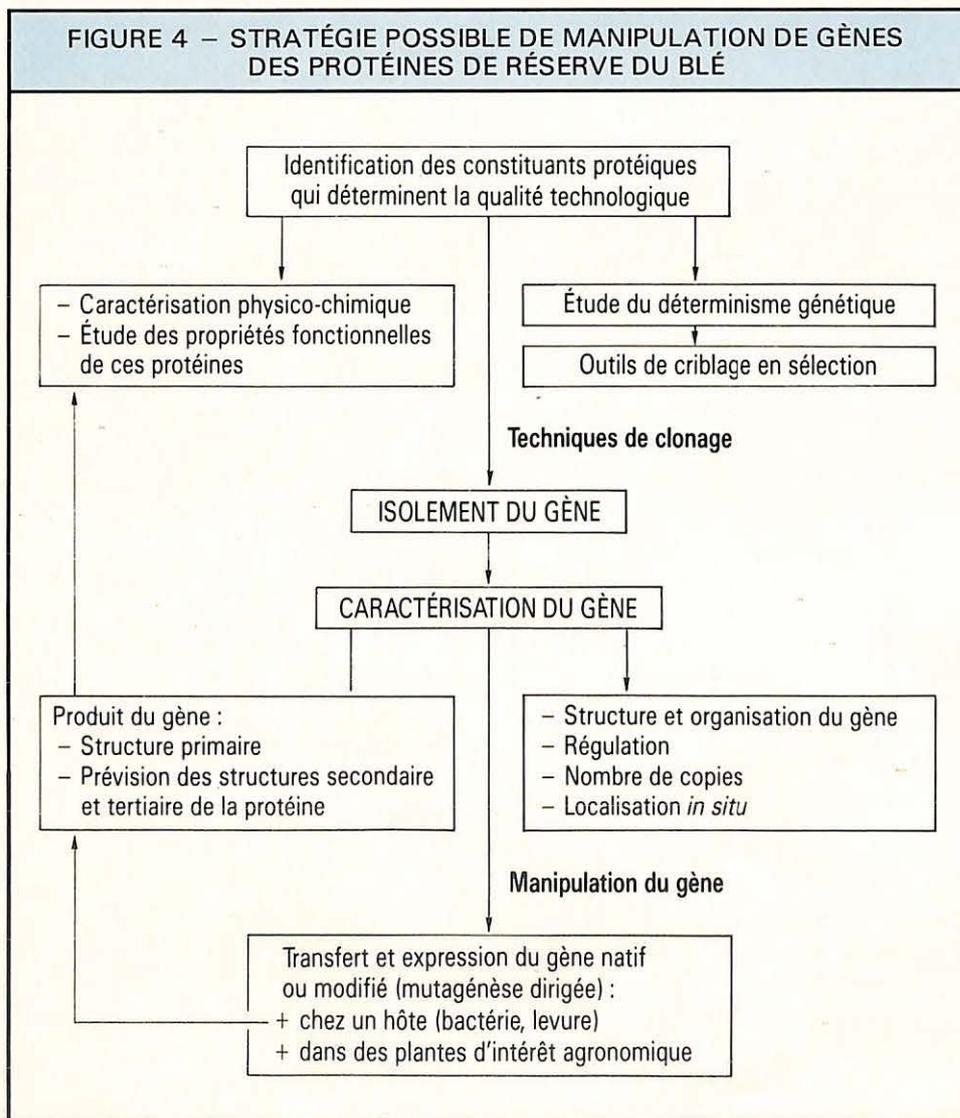
- l'estimation du nombre de copies d'un gène dans le génome, l'organisation de ces gènes sur les chromosomes, leur localisation chromosomique *in situ* ;
- l'identification, d'une part, des séquences responsables de la spécificité d'expression dans les tissus de la graine et, d'autre part, des séquences responsables d'un taux élevé d'expression pendant une courte période du développement de la plante. En d'autres termes, il s'agira de comprendre comment l'expression de ces familles multigéniques est limitée aux tissus de la graine et comment elle est coordonnée.

Transformer des plantes ?

A plus long terme, en complément des approches décrites ci-dessus, on peut envisager la création de plantes transformées par génie génétique (*voir figure 4*).

(19) B.A. Larkins and R. Cuellar, Modified Zein. European patent n° 0208418 (1986), Lubrizol Genetics Inc.

FIGURE 4 — STRATÉGIE POSSIBLE DE MANIPULATION DE GÈNES DES PROTÉINES DE RÉSERVE DU BLÉ



Les succès d'une modification par génie génétique requièrent cependant plusieurs préalables : l'identification des bases biochimiques ainsi que des produits d'expression du (ou des) gène(s) impliqué(s) dans le caractère à modifier ; l'isolement des gènes devant être modifiés et transférés ; l'aptitude à la transformation de la plante retenue ; l'expression correcte du gène transféré.

Il faut souligner que, pour beaucoup de problèmes d'intérêt agronomique, la question de savoir quel(s) gène(s) transférer demeure l'une des plus délicates. En effet, qu'il s'agisse des problèmes de rendement, de résistance aux stress (froid, sécheresse), de fixation de l'azote atmosphérique... tous ces phénomènes mettent en jeu un nombre de gènes important et pour lesquels, dans la plupart des cas, on ne connaît pas les protéines impliquées.

Il faut donc reconnaître que le modèle retenu dans ce travail, celui des gènes du locus **Gli-B1** du blé dur, constitue un cas unique par sa simplicité et pour lequel on dispose de plusieurs marqueurs protéiques parfaitement identifiés.

En ce qui concerne le transfert, le problème majeur avec les céréales se situe au niveau de la **régénération de plantes entières** à partir de protoplastes. Ce n'est plus un problème de biologie moléculaire mais de biologie cellulaire et de culture de tissus. En effet, si l'on sait transformer des protoplastes de céréales, la régénération des cals reste une étape non résolue, excepté pour le riz où un succès récent a été

obtenu (20). Des résultats très prometteurs ont été obtenus sur le seigle par injection directe d'ADN dans les inflorescences, 14 jours avant la méiose (21). Si cette technique de transformation directe se généralisait sur les céréales, toutes les étapes (protoplastes, cals, régénération de plantes entières) - qui constituent autant de verrous pour le succès des programmes visant à transférer des gènes chez les monocotylédones - seraient éliminées.

Tous ces résultats sont encourageants et l'on peut penser que, dans un proche avenir, la transformation des principales céréales cultivées sera possible.

Pour les **problèmes liés à l'expression du gène transféré**, non seulement ce gène doit s'exprimer à un taux satisfaisant mais dans un tissu spécifique et à un moment bien précis du développement de la plante. Il est évident que si l'on veut changer la composition protéique d'un blé, eu égard au fait qu'il s'agit de familles multigéniques, il faudra modifier soit un seul gène et faire en sorte qu'il soit fortement transcrit puis traduit pour observer une différence significative, soit plusieurs gènes transcrits et traduits plus faiblement. Ce n'est que dans la mesure où les phénomènes de régulation seront connus qu'il sera possible d'intervenir.

Des résultats ont déjà été obtenus chez les dicotylédones, avec des gènes qui codent pour certaines protéines de réserve des légumineuses (phaséoline du haricot, conglycine du

soja) et des céréales (zéine du maïs). Si, après transfert dans le tabac, le gène de la phaséoline est exprimé sélectivement dans la graine et à un taux d'expression comparable à celui qu'il a dans le haricot (22), on n'a pas détecté de zéine dans les tissus de tournesol transformé alors que le gène est transcrit (23). Sans aucun doute, les études d'expression des gènes transférés dans des espèces différentes fourniront des informations précieuses sur les phénomènes de régulation.

Comme on le voit, de nombreuses questions restent encore à résoudre, ce qui souligne que seuls des progrès continus dans la compréhension de la structure et du fonctionnement des gènes végétaux, sont la garantie d'un succès futur pour l'amélioration des productions végétales. □

(20) L.Y. Coulibaly and Y. Demarly, Regeneration of plantlets from protoplasts of rice, *Oryza sativa* L., *Z. Pflanzenzüchtg* (1986) 96, 79-81.

(21) A. De La Pena, H. Lorz and J. Schell, Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nature* (1987) 325, 274-276.

(22) C. Hall, N.A. Reichert, C. Sengupta-Gopalan, J.H. Cramer, K. Lea, R.F. Barker, J.L. Slightom, R. Klassy and J.D. Kemp, Regulation of bean phaseolin gene expression in yeast and tobacco seed. *Nato Asi Ser., Ser. (Mol. Form Funct. Plant Genome)* (1985) 83, 517-529.

(23) P.B. Goldsbrough, S.B. Gelvin and B.A. Larkins, Expression of maize zein genes in transformed sunflower cells. *Mol. Gen. Genet.* (1986) 202, 374-381.

Biofutur formation

Pour tout renseignement :

BIOFUTUR

Mme M.F. FOULQUIE
29 rue Buffon 75005 PARIS

Tél. (1) 43.36.96.96
47.07.11.22

Télex : 202 400

Télécopie : (1) 43.36.80.93

Les biotechnologies de A à Z

Vue d'ensemble et bases scientifiques et technologiques nécessaires à la compréhension des biotechnologies.

26-27-28 octobre
9-10-11 décembre - 5 300 F HT

La biologie aujourd'hui et demain

Des connaissances pour mieux suivre les conséquences de la "révolution biotechnologique."

14-15-16 octobre
15-16-17 décembre - 4 900 F HT

NUMÉRO 60
SEPTEMBRE 1987
PRIX : 43 F

Biofutur

LE MENSUEL EUROPÉEN DE BIOTECHNOLOGIE

INTERNATIONAL
THAÏLANDE
ET MALAISIE...

SANTÉ
THE THROMBOLYTICS WAR

AGRICULTURE
LES PROTÉINES DE
RÉSERVE DU BLÉ



DOSSIER
LA POMME DE TERRE FAIT PEAU NEUVE