

Analyse de la réponse au stress hydrique chez le maïs

Apport de la génétique quantitative moléculaire et de la protéomique

A. Leonardi¹

D. De Vienne¹

S. Pelleschi²

J.-L. Prioul²

M. Zivy¹

Introduction

La disponibilité en eau est un facteur limitant majeur pour la croissance et la productivité des plantes, particulièrement pour le maïs, qui termine son cycle à une période de faible pluviométrie. Comme la plupart des caractères qui intéressent l'améliorateur de plantes, la résistance à la sécheresse est un caractère quantitatif, c'est-à-dire à distribution continue : d'une part son déterminisme est polygénique et d'autre part l'expression de cette résistance est très influencée par les conditions environnementales. Mais tous les gènes impliqués

¹ Station de génétique végétale, Inra/UPS/INA PG, ferme du Moulon, 91190 Gif-sur-Yvette, France.

² Laboratoire de Structure et métabolisme des plantes, CNRS/UPS, IBP, 91405 Orsay, France.

dans la résistance n'ont pas la même importance : on peut supposer que le polymorphisme d'une petite partie d'entre eux est responsable d'une grande part de la variabilité génétique observée au niveau de la tolérance à la contrainte hydrique (Prioul *et al.*, 1997). L'objectif de notre travail est d'identifier ces gènes. Pour ce faire, deux approches sont menées en parallèle :

– la première consiste à décomposer la réponse au stress hydrique en mesurant des caractères à différents niveaux (biochimique, hormonal, physiologique, phénologique et morphologique) et à cartographier les régions chromosomiques qui contrôlent ces caractères (QTL, Quantitative Trait Locus) ;

– la seconde cherche à identifier les modifications d'expression du génome sous l'effet de la contrainte hydrique (transcriptome et protéome), dans différents organes, à différents stades de développement.

La combinaison de ces deux approches permet de proposer des gènes candidats dont le polymorphisme peut expliquer la variation de la réponse au stress hydrique, et ouvre la voie à l'amélioration pour une meilleure tolérance à la sécheresse.

■ Cartographie de QTL

Les QTL sont des locus dont le polymorphisme est responsable, au moins en partie, de la variation du caractère quantitatif auquel on s'intéresse. On peut détecter un QTL dès que l'on dispose d'un marqueur polymorphe en déséquilibre de liaison avec le QTL (fig. 1). Pour cartographier les QTL il faut (i) disposer d'une population en ségrégation présentant de la variabilité pour le caractère d'intérêt, (ii) déterminer le génotype de chaque individu de la population pour un grand nombre de locus marqueurs afin de construire une carte génétique saturée, (iii) mesurer la valeur du caractère quantitatif pour chaque individu de la descendance et (iv) appliquer des méthodes biométriques pour localiser les locus marqueurs liés au(x) QTL et en estimer divers paramètres génétiques. La méthode la plus simple est l'analyse de variance à un facteur, qui consiste à compa-

rer les moyennes des valeurs du caractère entre les différentes classes génotypiques au marqueur. De nombreuses méthodes plus sophistiquées existent aujourd'hui, dont la cartographie d'intervalle simple (avec le logiciel Mapmaker/QTL développé par Lander et Botstein, 1989 et plus récemment les méthodes de cartographie d'intervalle composite (Jansen, 1993 ; Zeng, 1994) qui confèrent une meilleure précision et surtout une plus grande puissance de détection.

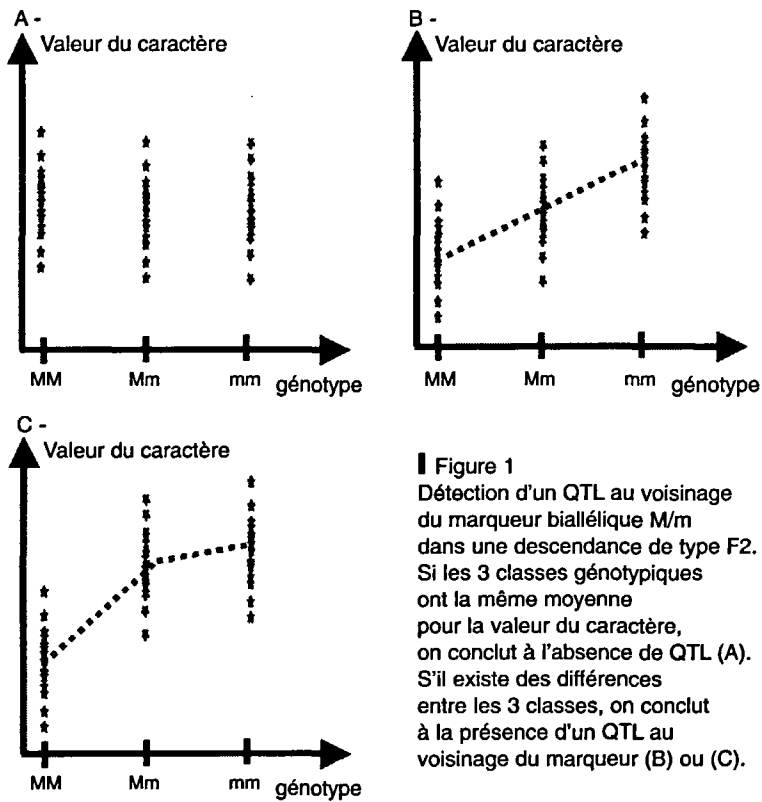


Figure 1
Détection d'un QTL au voisinage du marqueur biallélique M/m dans une descendance de type F₂. Si les 3 classes génotypiques ont la même moyenne pour la valeur du caractère, on conclut à l'absence de QTL (A). S'il existe des différences entre les 3 classes, on conclut à la présence d'un QTL au voisinage du marqueur (B) ou (C).

Les populations en ségrégation les plus couramment utilisées sont les descendance F₂, les back-cross et les populations de lignées recombinantes. Ces dernières sont obtenues par autofécondations successives à partir d'individus F₂ en descendance « monograine ».

Ce type de population présente l'avantage d'être reproductible à l'identique par autofécondation et donc de permettre la mesure d'un grand nombre de caractères dans de multiples conditions environnementales et sur un grand nombre d'individus génétiquement identiques. Un nouveau type de population est développé actuellement, les lignées hautement recombinantes : des cycles de panmixie sont intercalés avant les cycles d'autofécondations successives, ce qui augmente le taux de recombinaison entre 2 marqueurs proches, permettant ainsi de les ordonner plus précisément et donc de localiser plus finement les QTL détectés.

Le principal problème qui surgit une fois le QTL détecté est effectivement de savoir à quel(s) gène(s) il correspond et comment atteindre ce(s) gène(s). Sachant qu'avec une population de 100 lignées recombinantes les intervalles de confiance des QTL se situent au mieux autour de ± 10 cM et étant donné la grande taille des génomes végétaux (tabl. 1), un grand nombre de gènes peuvent se situer dans l'intervalle. C'est pourquoi bien que de nombreux QTL aient été détectés chez les plantes pour des caractères divers, très peu ont été caractérisés au niveau moléculaire (de Vienne *et al.*, 1999). C'est ici que se situe l'intérêt de l'approche gène/protéine candidat.

	Longueur génétique totale en cM	Longueur ADN/cM
Arabidopsis	630	140 kb
Riz	1 575	280 kb
Soja	2 700	440 kb
Tomate	1 267	750 kb
Haricot	830	780 kb
Colza	1 016	1 200 kb
Maïs	1 860	1 400 kb
Blé	3 500	4 600 kb

■ **Tableau 1**
Relation entre distances physiques et génétiques
chez les espèces pour lesquelles on dispose
de cartes de liaisons génétiques saturées
(d'après de Vienne, 1998 *Les marqueurs moléculaires
en génétique et biotechnologies végétales*, Inra, Paris).

I Approche gène/protéine candidat

À mesure que les programmes sur les génomes se développent, de plus en plus de gènes de fonction connue sont recensés dans les bases de données et sont localisés sur les cartes physiques ou génétiques. On peut ainsi sélectionner des gènes selon un premier critère qui sera un critère fonctionnel : les gènes candidats seront les gènes dont la fonction est en rapport avec le caractère d'intérêt. Le deuxième critère est génétique, il intervient quand l'expression du gène est variable dans la population :

- soit parce que le gène de structure est polymorphe ;
- soit parce que les gènes contrôlant son expression (au niveau de la transcription, traduction ou post-traduction) sont polymorphes.

On s'intéresse par exemple à la variation d'activité d'enzymes impliquées dans une voie métabolique importante dans la réalisation du caractère d'intérêt. Mais on ne peut pas mesurer les activités d'un grand nombre de ces enzymes. D'où l'intérêt de l'étude du « protéome » (The PROTEin complement expressed by a genOME, Wilkins *et al.*, 1996), qui permet de mesurer la variation des quantités de protéines. On peut en effet estimer que la variation de quantité d'une protéine est la composante majeure de la variation de son activité. Les méthodes d'identification des protéines ont fait récemment beaucoup de progrès : outre l'identification par microséquençage ou par la composition en acides aminés, l'utilisation de la spectrométrie de masse permet aujourd'hui une identification à plus grande échelle. Les protéines sont révélées sur les gels d'électrophorèse bidimensionnelle (O'Farrell, 1975) et quantifiées par analyse automatique des gels. La variation des quantités de protéines pourra être mesurée en fonction de critères physiologiques et en fonction des variations génétiques (fig. 2). Les protéines candidates pourront alors être sélectionnées, selon deux critères :

- fonctionnel : protéines dont la quantité est régulée parallèlement à la variation du caractère d'intérêt ;
- génétique : la quantité de la protéine est variable selon le génotype. On peut alors appliquer la méthodologie de détection de QTL aux quantités de protéines, en cherchant à cartographier les PQL (Protein Quantity Loci) de ces protéines (Damerval *et al.*, 1994).

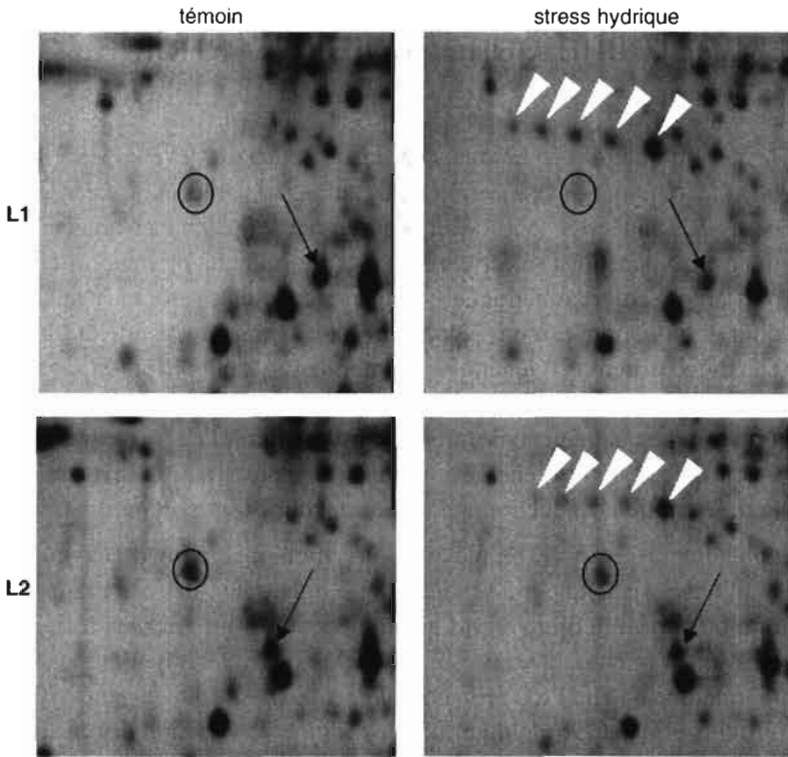
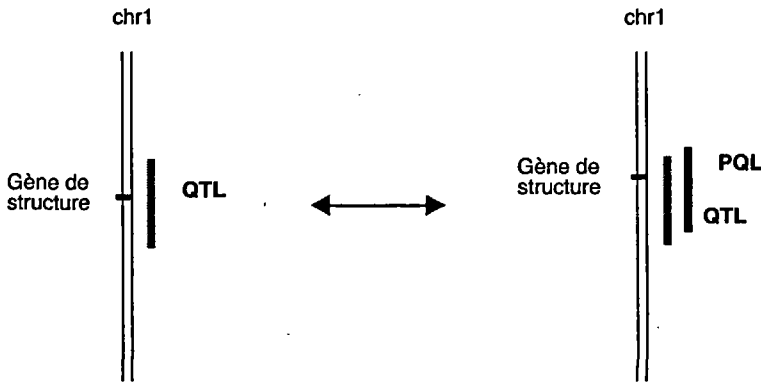


Figure 2

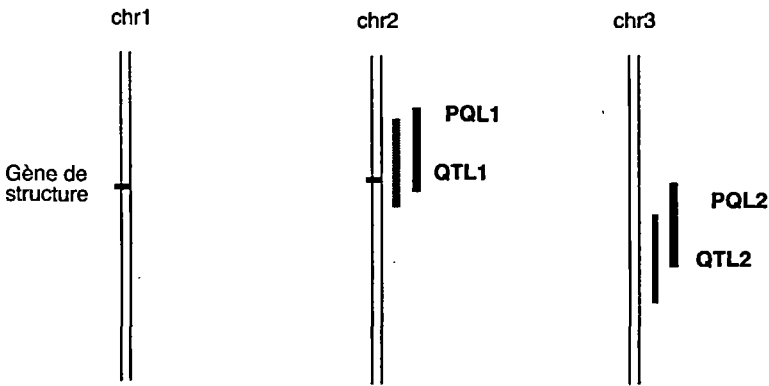
Variation des quantités de protéines révélées par électrophorèse bidimensionnelle : comparaison de 2 lignées L1 et L2, en condition témoin et en condition de stress hydrique. Le rond entoure un spot protéique qui est plus intense chez L2 que chez L1. Les flèches claires indiquent un train de spots qui apparaît avec le stress hydrique et les flèches noires une variation de position (variation allélique du gène de structure codant pour la protéine).

Les gènes et protéines candidats vont être finalement sélectionnés après la recherche de co-localisations entre les gènes de structure, les PQL de protéines et les QTL du caractère d'intérêt (fig. 3). Pour éviter les co-localisations fortuites, on s'attachera surtout aux co-localisations multiples. L'intérêt des protéines apparaît quand le gène de structure lui-même n'est pas variable mais que la quantité de protéine l'est : la cartographie des PQL devient possible et permet de relier la variation de quantité de la protéine et la variation du caractère quantitatif.

A- Polymorphisme du gène de structure



B- Polymorphisme de gènes contrôlant son expression

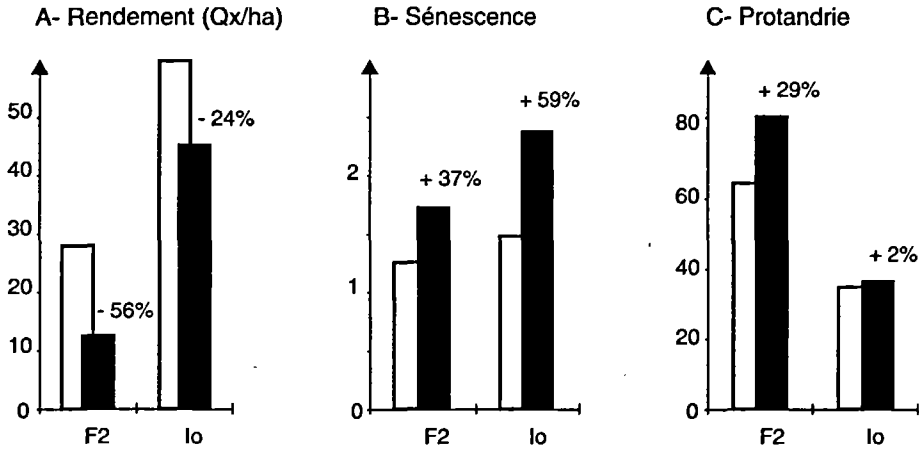


■ Figure 3

Recherche de co-localisations entre les QTL du caractère d'intérêt, le gène de structure et les PQL d'une protéine.

■ Application à l'étude de la tolérance à la sécheresse chez le maïs

Ce travail a été réalisé sur une population de 145 lignées recombinantes issues de 2 lignées de maïs, F-2 et Io. F-2 est une lignée « sensible » au stress hydrique et Io une lignée « tolérante » (fig. 4).



■ Figure 4

Variation entre les 2 lignées parentales F-2 et lo pour des caractères de réponse au stress hydrique : rendement (A), sénescence (B) (nombre de feuilles sèches/somme de températures entre floraison et notation) et protandrie (C) (somme de températures entre la floraison mâle et la floraison femelle). Blanc : condition témoin ; noir : condition de stress hydrique.

Cette population a été caractérisée pour sa réponse au stress hydrique au champ (rendement, croissance, sénescence, protandrie) et en serre (ajustement osmotique, diminution de la croissance, de la photosynthèse, synthèse d'acide abscissique (ABA) endogène). La réponse a également été mesurée au niveau biochimique : 1) par la mesure des activités d'enzymes impliquées dans le métabolisme carboné, teneurs en sucres solubles, en amidon, et 2) par les quantités de protéines révélées par électrophorèse bidimensionnelle.

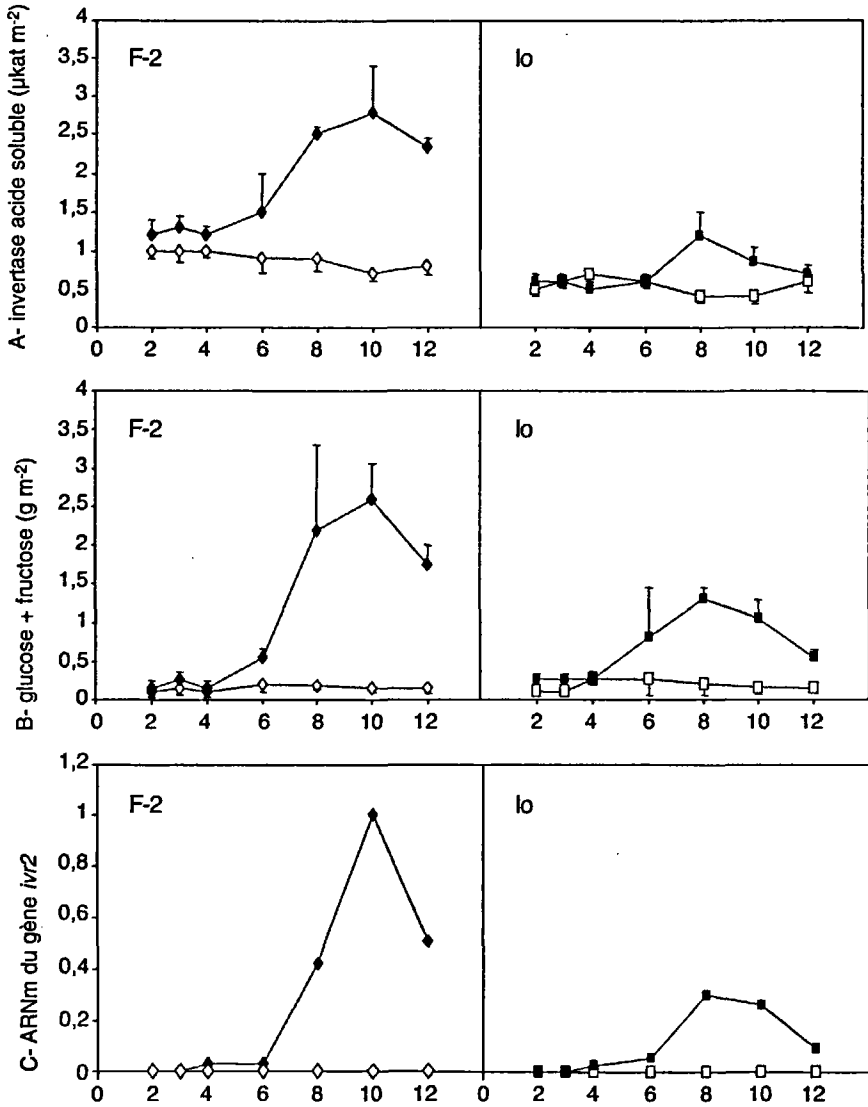
La carte de marqueurs RFLP construite avec cette population de lignées recombinantes a été utilisée pour localiser : (i) des QTL de caractères de réponse au stress hydrique, (ii) des QTL d'activités d'enzymes clés du métabolisme du carbone, (iii) des PQL de protéines affectées par la contrainte hydrique et (iv) des gènes de fonction connue, pouvant *a priori* jouer un rôle dans la tolérance à la sécheresse. Diverses co-localisations ont ensuite été recherchées. Quelques résultats marquants seront évoqués ici.

Invertase acide soluble

Les réponses physiologiques ont été analysées sur de jeunes plantes au stade 4 feuilles adultes, cultivées en serre et soumises pendant 9 jours à un arrêt d'arrosage. On observe une augmentation précoce (2^e-4^e jour) de l'activité invertase acide soluble, enzyme impliquée dans la dégradation du saccharose dans les vacuoles. Cette augmentation est très corrélée avec un accroissement des teneurs en hexoses. Cette augmentation précède la baisse de la photosynthèse, mais coïncide avec l'accumulation d'ABA endogène dans la sève du xylème. En revanche, l'état hydrique de la plante est peu modifié. L'augmentation de l'activité invertase est variable selon le génotype et corrélée avec l'accumulation du transcrite d'*Ivr2*, l'un des 2 gènes codant pour cette enzyme (fig. 5). La détection de QTL pour l'activité invertase montre qu'en condition témoin, le QTL majeur se situe au niveau du gène *Ivr2* sur le chromosome 5, alors qu'en condition de stress hydrique, bien qu'il existe un QTL au même endroit, le QTL majeur se situe sur le chromosome 2, près de *Incw2* qui code pour un gène invertase pariétale (fig. 6). Une mutation de ce gène *Incw2* produit le phénotype « grain miniature », avec les 2 activités invertase pariétales et vacuolaires réduites (Cheng *et al.*, 1996). Ce résultat montre que si *Ivr2* est un bon candidat QTL pour l'activité invertase en condition témoin. Il existe d'autres facteurs, spécifiques de la contrainte hydrique, responsables de la variation d'activité de l'invertase vacuolaire chez les plantes stressées (Pelleschi *et al.*, 1999 ; Prioul *et al.*, 1999).

Étude des protéines affectées par la contrainte hydrique

Les protéines affectées par la sécheresse ont été étudiées par électrophorèse bidimensionnelle, au niveau de la base de la feuille 6 de jeunes plantules après 10 jours de stress. La comparaison des 2 lignées parentales et de leur hybride F1 en condition de stress hydrique et en condition témoin a permis de révéler 78 protéines affectées par le stress hydrique, dont 50 étaient induites. Des varia-



■ Figure 5

Cinétique de l'activité et de la quantité d'ARNm pour l'invertase acide, et teneur en hexoses, chez les 2 lignées parentales F-2 et lo.

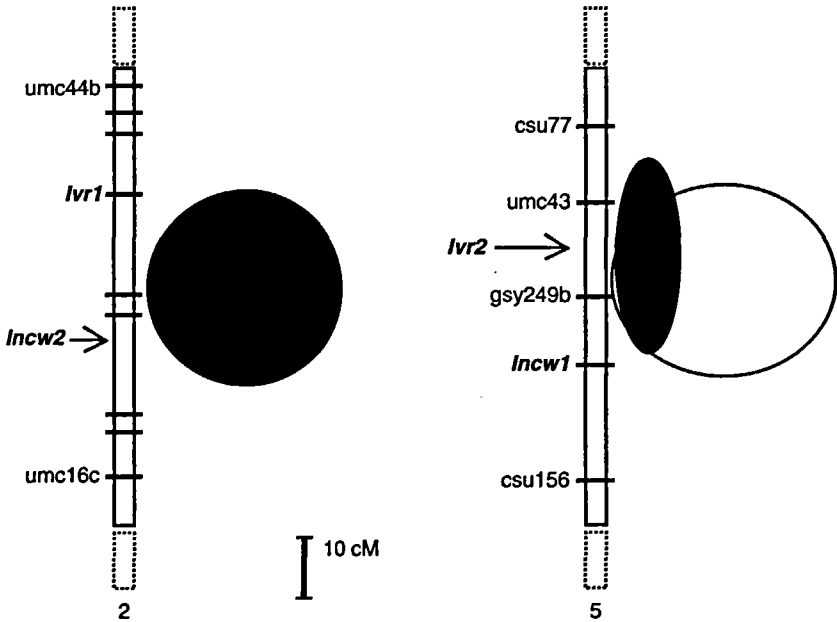
A) activité invertase acide soluble,

B) teneur en glucose + fructose,

C) quantité de l'ARNm du gène *ivr2*.

◇ condition témoin,

■ condition de stress (d'après Pelleschi *et al.*, 1999).



■ Figure 6

QTL d'activité de l'invertase acide soluble.

Les QTL sont représentés par les ellipses.

En clair : condition témoin ; en noir : condition de stress hydrique.

La hauteur de l'ellipse représente l'intervalle de confiance du QTL, sa largeur est proportionnelle à l'effet du QTL.

lvr1 et *lvr2* codent pour l'invertase acide soluble et *incw1* et *incw2* pour l'invertase pariétale (d'après Prioul *et al.*, 1999).

tions selon le génotype (effet génotype ou interaction génotype*traitement) ont été révélées pour 28 de ces protéines. Des microséquences de fragments internes de 42 de ces protéines de stress, candidates potentielles, ont permis l'identification de 36 protéines (Riccardi *et al.*, 1998) (fig. 7). L'une d'entre elles est la protéine ASR (ABA/water stress/ripening induced), dont la quantité augmente avec le stress chez Io mais qui n'est jamais détectée chez F-2. L'ADNc de l'ASR1 (isolé par M. Höfer, non publié) a été cartographié sur le chromosome 10, au niveau du PQL de présence/absence de la protéine, dans une région où se trouve un QTL

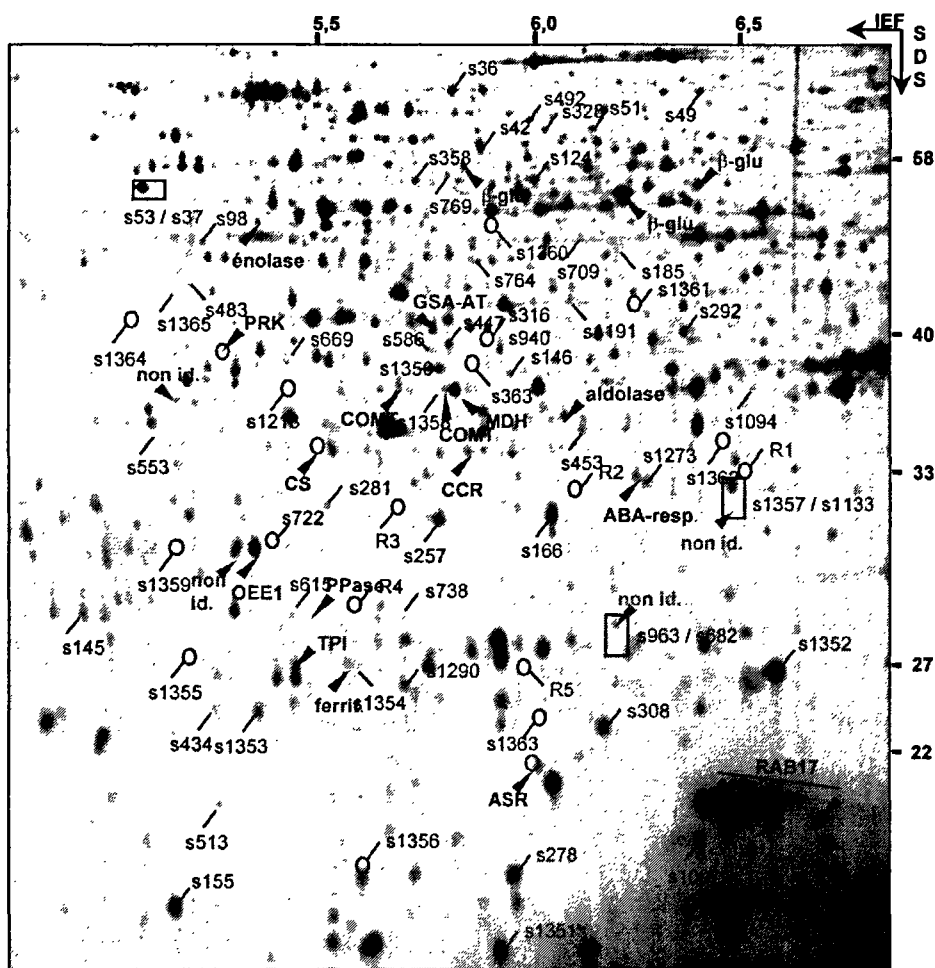


Figure 7
Protéines affectées par le stress hydrique.

pour la protandrie et un QTL pour la sénescence (fig. 8). Ce résultat suggère que le polymorphisme du gène de structure de l'ASR1 est responsable de la variation de présence/absence de la protéine, qui affecte elle-même de façon pléiotropique les autres caractères de réponse au stress. Des expériences de transformation sont en cours pour vérifier cette hypothèse. D'autres exemples sont développés dans de Vienne *et al.* (1999).

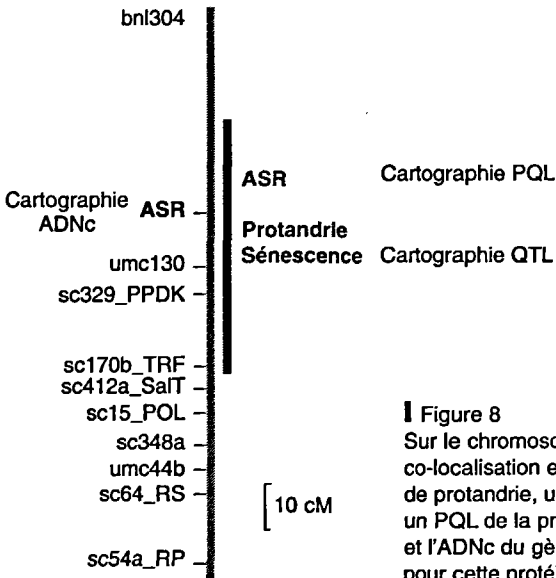


Figure 8
 Sur le chromosome 10 du maïs, co-localisation entre un QTL de protandrie, un QTL de sénescence, un PQL de la protéine ASR1 et l'ADNc du gène codant pour cette protéine (d'après de Vienne *et al.*, 1999).

Conclusion

Le protéome permet de mettre en relation les études sur la physiologie des plantes et les informations d'ordre génétique. On peut, d'une part, mettre en évidence les voies ou portion de voies métaboliques impliquées et, d'autre part, proposer un certain nombre de gènes ou protéines candidats. Ces gènes ou protéines candidats devront ensuite être validés. Différentes approches sont possibles :

- l'étude dans d'autres fonds génétiques, par exemple une population de lignées non apparentées où le déséquilibre de liaison sera minimal : si la corrélation persiste entre le polymorphisme du gène ou de la protéine et la variation du caractère quantitatif, l'hypothèse du lien de causalité est renforcée ;
- l'étude de complémentation, dans le cas où il existe un mutant correspondant au gène candidat (comme dans les travaux de Doebley chez le maïs pour un QTL d'architecture de la plante, Doebley *et al.*, 1995) ;
- les analyses de fonction ;
- la transformation.

Bibliographie

- Cheng WH, Taliencio EW, Chourey PS 1996 —
The miniature 1 seed locus of maize encodes a cell wall invertase required for normal development of endosperm and maternal cells in the pedicel. *Plant Cell* 8 : 971-983.
- Damerval C, Maurice A, Josse J-M, de Vienne D 1994 —
Quantitative trait loci underlying gene product variation - a novel perspective for analyzing regulation of genome expression. *Genetics* 137: 289-301.
- de Vienne D, Leonardi A, Damerval C, Zivy M 1999 —
Genetics of proteome variation as a tool for QTL characterization: application to drought-stress responses in maize. *J Exp Bot* 50: 303-309.
- Doebley J, Stec A, Gustus C 1995 —
Teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* 141: 333-346.
- Jansen RC 1993 —
Interval mapping of multiple quantitative trait loci. *Genetics* 135: 205-211.
- Lander ES, Botstein D 1989 —
Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185-199.
- O'Farrell PH 1975 —
High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250: 4007-4021.
- Pelleschi S, Guy S, Kim J-Y, Pointe C, Mahé A, Barthes L, Leonardi A, Prioul J-L 1999 —
Ivr2, a candidate gene for a QTL of vacuolar invertase activity in maize leaves: Gene specific expression under water stress. *Plant Mol Biol* 39: 373-380.
- Prioul J-L, Quarrie S, Causse M, de Vienne D 1997 —
Dissecting complex physiological functions into elementary components through the use of molecular quantitative genetics. *J Exp Bot* 48: 1151-1163.
- Prioul J-L, Pelleschi S, Séne M, Thévenot C, Causse M, de Vienne D, Leonardi A 1999 —
From QTLs for enzyme activity to candidate genes in maize. *J Exp Bot* 50: 1281-1288.
- Riccardi F, Gazeau P, de Vienne D, Zivy M 1998 —
Protein changes in response to progressive water deficit in maize. Quantitative variation and polypeptide identification. *Plant Physiol* 117: 1253-1263.
- Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez J-C, Yan JX, Gooley AA, Hughes G, Humperly-Smith I, Williams KL, Hochstrasser DF 1996 —
From proteins to proteomes: large-scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology* 14: 61-65.
- Zeng Z-B 1994 —
Precision mapping of quantitative trait loci. *Genetics* 136: 1457-1468.