

CHAPITRE XI

CARTOGRAPHIE GENETIQUE DU MIL

M.T. PEIGNE, J. ENJALBERT, T. ROBERT, A. RICROCH,
R. LESPINASSE, M. SANDMEIER et A. SARR

Laboratoire d'Evolution et Systématique Végétales, bâtiment 362
Université Paris Sud ; 91405 Orsay Cedex.

Le mil (*Pennisetum typhoides*) pose un problème évolutif intrigant : comment la domestication puis le maintien en sympatrie des formes cultivées et sauvages pendant 5000 ans ont-ils été possible ? La réponse à cette question passait par la connaissance précise des déterminismes génétiques des caractères différenciant les deux types de plantes. Ainsi l'observation morphologique d'une série de croisements F₂ et backcross entre trois génotypes sauvages et cinq génotypes cultivés a permis de dresser une carte des principaux caractères soumis à la domestication : une portion chromosomique unique de soixante centimorgans porte la majorité des gènes de structure de l'épillet (caducité, enveloppement de la graine, longueur du pédicelle), et est vraisemblablement dupliquée dans la plupart des croisements.

Une cartographie assistée par marqueurs RFLP a été entreprise depuis. Les premiers résultats concernant les données morphologiques, mettent l'accent sur le caractère quantitatif de la caducité et de l'enveloppement de la graine et montrent l'indispensable utilité des marqueurs génétiques pour en résoudre l'hérédité. L'établissement de la cartographie chromosomique des marqueurs du mil a été initié dernièrement. Elle permettra de visualiser directement les éventuelles duplications du syndrome de domestication. De plus ces techniques seront essentielles dans l'établissement des phylogénies grâce à la visualisation directe des remaniements chromosomiques. La cartographie précise des gènes de domestication sera un outil important pour l'utilisation de matériels exotiques dans les programmes de sélection. La compréhension de l'équilibre dynamique existant entre les deux compartiments du complexe d'espèce devrait apporter des enseignement théoriques utiles pour la gestion dynamique des ressources génétiques.

I - INTRODUCTION

L'étude du syndrome de domestication initiée par J. Pernès est abordée par notre groupe à différents niveaux : génétique, biologie florale, moléculaire, cytogénétique. Un certain nombre de problèmes demeure quand on examine son organisation entre différentes lignées : une cartographie du génome du mil a été entreprise pour répondre à ces interrogations. De plus, la cartographie est une méthode qui devrait

permettre de proposer des hypothèses quant aux sens des processus évolutifs en cause. Il a été observé dans de nombreux croisements, aussi bien cultivé x cultivé que cultivé x sauvage, des distorsions de ségrégation (Robert, 1991) qui ne s'expliquent que partiellement par des compétitions polliniques. Quelle est la part de la structuration génétique du génome dans ces phénomènes ? Une telle structuration commence à être définie dans certains croisements mais peut-elle être généralisée à l'ensemble du pool primaire du mil ?

Les formes spontanées de mil se distinguent aisément des cultivées. Leur tallage est précoce, abondant, l'émission foliaire lente, le drapeau, (dernière feuille avant la chandelle) étroit. L'involucre est sessile, l'épillet possède de nombreuses soies dont une plus longue que les autres. Toutes ces caractéristiques ont permis la définition d'un certain nombre de marqueurs discriminants qui traduisent effectivement les oppositions entre formes sauvages et cultivées. Pour répondre à ces questions nous avons entrepris l'établissement de la carte du mil par la méthode de génétique mendélienne à l'aide des marqueurs morphologiques et enzymatiques. La carte des marqueurs moléculaires est en cours d'élaboration par l'équipe de M.D. Gale au John Innes' Institute. La carte chromosomique se fera par l'intermédiaire des lignées aneuploïdes, en cours d'obtention et par marquage en hybridation *in situ*.

II - SYNDROME DE DOMESTICATION DE L'ÉPILLET

1) Marqueurs morphologiques

Dans le cadre du laboratoire GPDP à Gif, les croisements entre la forme spontanée (*Pennisetum mollissimum*) et six lignées cultivées (*typhoides*) ont été réalisés (Belliard, 1980 ; Beninga, 1981 ; Pernès, 1980 ; Sandmeier, 1981). Ils ont montré l'absence de barrière reproductive en F1 et une très grande variabilité des familles F2 qui laissait apparaître de nouvelles plantes viables, fertiles, mais également des plantes anormales incapables de fleurir et donc complètement stériles. Un cas de stérilité en croisement a été mis en évidence par Marchais et Pernès (1985).

Rey-Herme (1982) a étudié plus spécialement le croisement de la lignée Massue par la forme sauvage *mollissimum* : elle a mis en évidence des liaisons entre plusieurs marqueurs de l'épillet (tableau 1), qu'elle a reliés au marqueur enzymatique estérase. E. Joly-Ichenhauser (1984) retrouve dans d'autres confrontations sauvages x cultivées le même ordre de grandeur pour les distances entre l'enveloppe de la graine (ENV), la longueur du pédicelle de l'involucre (LPI) et la caducité (CAD) : elle met en évidence pour les lignées Drôo, Tiotandé, 23 DB (tableau 2) et pour les formes spontanées *P. mollissimum* et *P. violaceum* du Niger, une duplication chromosomique pour LPI, ENV, CAD. Un croisement Souna du Mali x *P. mollissimum* donne les mêmes résultats (tableau 3) (Enjalbert 1992).

2) Marqueurs enzymatiques

L'étude au laboratoire du GPDP à Gif par Pilate-André et Sandmeier, d'un certain nombre de F2 sauvage x cultivé a permis de situer 15 marqueurs enzymatiques, regroupés en cinq groupes de liaisons (tableau 3, groupes 1, 2, 3, 5, 7). Banuett (1979) a étudié les Adh_a et les Adh_b et a montré une liaison existant entre eux, (tableau 2, groupe 4). Tostain (1985) a relié un gène de nanisme à l'Adh_a et à l'Skdh.

De plus, certains marqueurs enzymatiques et morphologiques peuvent être reliés : c'est le cas, dans le groupe 2, pour l'estérase E avec la pubescence des feuilles (PUF) et la pilosité des feuilles (PIF), et, dans le groupe 3, pour les peroxydases anodiques avec la coloration du collet (CO.CO). Actuellement il n'est cependant pas possible de relier ces groupes de liaison à une structure chromosomique quelconque. Toutefois l'emploi des méthodes de biologie moléculaire (notamment l'hybridation *in situ*) devrait permettre la localisation de ces groupes de liaison sur les chromosomes.

III - CARTOGRAPHIE ET MARQUEURS MOLÉCULAIRES

Cette partie se voudra essentiellement prospective, les travaux de cartographie RFLP des gènes de domestication étant actuellement en cours dans notre laboratoire. Lors de l'établissement des cartes génétiques, il est nécessaire :

- de disposer de caractères dont on connaît le déterminisme génétique ;
- d'observer des coségrégations de caractères lors d'un croisement contrôlé (F2-BC).

Suivant le niveau d'observation du génotype, l'obtention de ces deux conditions est plus ou moins aisé.

1) Déterminismes génétiques

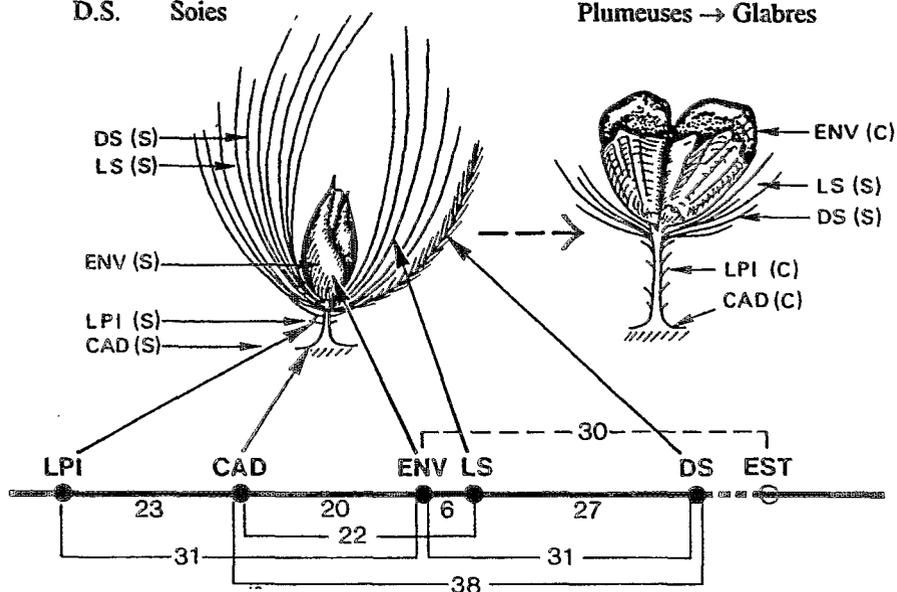
Si certains caractères propres à la domestication montrent un déterminisme génétique simple (monogénique ou digénique), la majeure partie des caractéristiques biométriques de la plante sont issus de l'expression de plusieurs gènes. Ainsi, la caducité, l'enveloppement de la graine ou la taille des graines sont des caractères quantitatifs et on observe, lors de la ségrégation, des distributions gaussiennes ; cette absence de multimodalité rend impossible la connaissance du nombre de gènes sous-jacents. Ainsi deux gènes codominants polymorphes suffisent pour obtenir une gaussienne. La discrimination des différents paramètres physiologiques, biochimiques puis génétiques responsables des variations d'un caractère permet de suivre individuellement chaque facteur et de retrouver les ségrégations mendéliennes. A l'extrême, c'est la ségrégation directe des mutations de l'ADN, responsables du caractère, qui est étudiée.

Ainsi, non seulement le polymorphisme augmente lorsque l'échelle d'observation diminue, mais les polymorphismes sont plus faciles à interpréter génétiquement.

Tableau 1
Marqueurs déterminant la structure de l'épillet

Dans ce schéma, nous avons figuré l'existence de l'estérase E1 (EST) que Béninga (1981) avait situé à une trentaine d'unité de l'aristation

Allèles		Dominants → Récessifs
L.P.I.	Pédicelle de l'involucre	Court → Long
C.A.D.	Caducité de l'épillet à maturité	Présent → Absent
E.N.V.	Enveloppement de la graine	Important → Réduit
L.S.	Soies	Longues → Courtes
D.S.	Soies	Plumeuses → Glabres



2) Estimation des distances

Le calcul de distance génétique, que ce soit par la méthode du maximum de vraisemblance ou par la méthode des produits, fait correspondre à un niveau d'association de deux caractères en ségrégation, un taux de recombinaison. Mais une corrélation entre deux caractères n'est pas forcément le seul fait des liaisons génétiques. Si un gène présente une certaine pléiotropie, s'il y a épistasie d'un locus sur un autre, alors la distance calculée (car s'est souvent possible) n'aura aucune valeur. Ainsi il est impossible de calculer une distance génétique entre le locus responsable de la couche d'abscission fonctionnelle et la note globale de caducité, car ces deux caractères sont unis par une relation de cause à effet.

Là encore, moins nombreuses seront les étapes qui séparent le polymorphisme de séquence du caractère observé, et plus faible sera la probabilité d'observer de tels phénomènes parasites.

Tableau 2
Groupes de liaison chez le mil

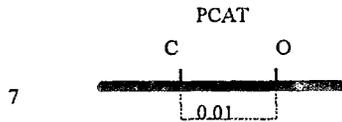
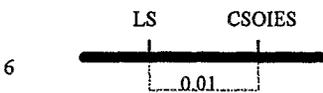
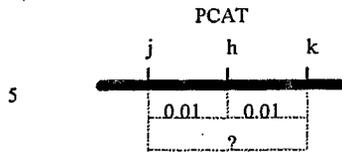
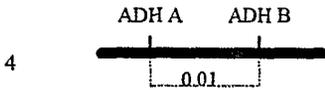
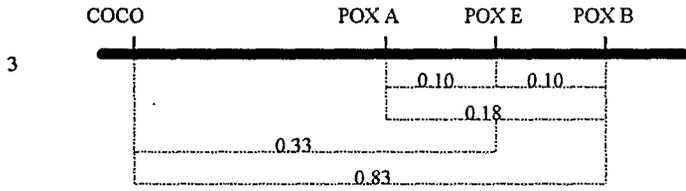
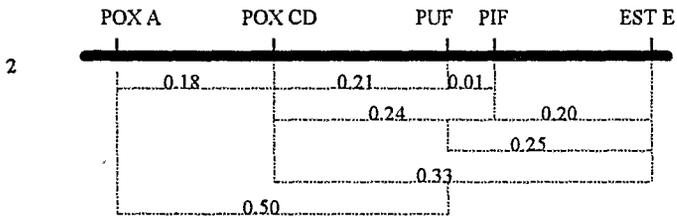
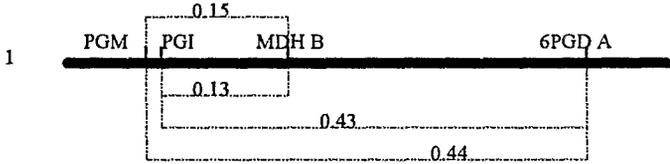
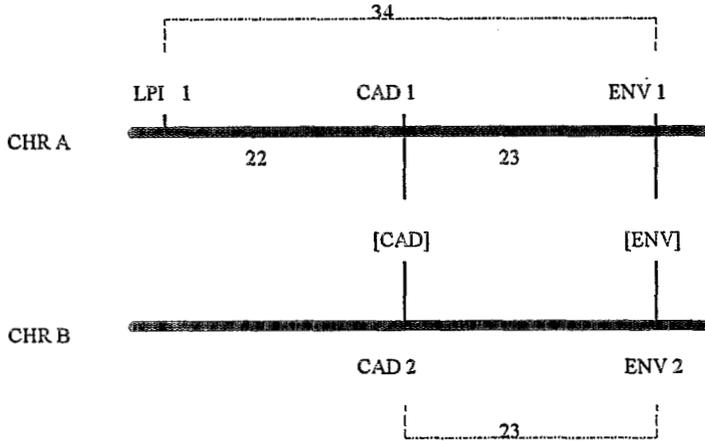
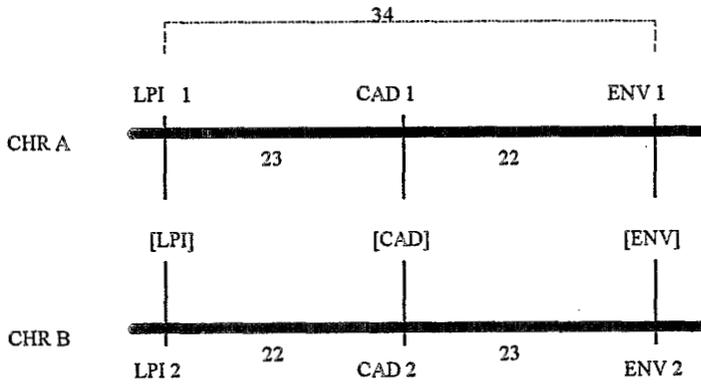


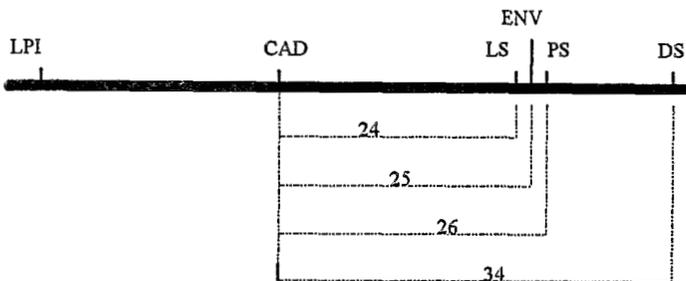
Tableau 3
Distances génétiques entre marqueurs morphologiques



Tiotandé & Drôo



23 DB



Souna x *P. mollissimum*

3) Les marqueurs moléculaires

La biologie moléculaire fournit un ensemble de caractères mendéliens (RFLP, RAPD, STS) car ils marquent directement la molécule d'ADN lors de sa ségrégation.

Toutefois, il est difficile d'obtenir un polymorphisme de caractère et de longueur de fragment de restriction provenant d'une même mutation. Par contre, du fait de leur liaison génétique, la coségrégation d'un électromorphe et d'un caractère quantitatif sont fréquents. Le déséquilibre de liaison existant entre la mutation responsable du caractère observé et les autres mutations qui lui sont liées génétiquement est ainsi exploité, et on parle alors de marqueurs (des régions chromosomiques adjacentes) neutres (car la mutation responsable du polymorphisme n'agit aucunement sur le caractère quantitatif analysé).

Les différentes sortes de marqueurs génétiques sont ainsi d'une grande utilité dans la sélection indirecte de critères coûteux ou longs à mesurer, récessifs ou encore d'expression tardive (sélection précoce). L'utilisation des marqueurs en sélection chez le mil a été illustrée par les travaux de Burton (1991), où le gène responsable d'absence de trichome a été utilisé, pour rechercher des effets hétérotiques augmentant le rendement dans des lignées exotiques, puis pour transférer le bloc chromosomique hétérotique dans du matériel amélioré. La réalisation de cartes RFLP saturées, a permis la détermination fine des gènes impliqués dans le déterminisme d'un caractère quantitatif.

4) Les "QTL" (Quantitative Trait Loci)

On dispose actuellement pour la majorité des plantes économiquement intéressantes (tomate : Benatsky, 1986; laitue : Landry, 1987 ; riz : McCouch, 1988...) de cartes moléculaires qui couvrent l'intégralité du génome par une succession de marqueurs liés entre eux de proche en proche.

Ainsi, si un QTL (Quantitative Trait Loci : locus impliqué dans un caractère quantitatif) est en ségrégation lors d'une descendance, il existera toujours un locus marqueur (polymorphe) qui lui sera lié. Entre les trois génotypes observés (MM, Mm, mm) pour un gène marqueur, une différence de moyenne apparaîtra si ce dernier est lié à un QTL. Ainsi, lorsqu'une descendance a été caractérisée biométriquement et moléculairement, une série d'analyses de variances permettra de détecter les marqueurs expliquant un certain niveau de variance d'un caractère particulier.

Les différences de moyennes observées dépendent à la fois de l'importance de l'effet du QTL sur le caractère, mais aussi de la distance QTL-Marqueur qui détermine leur niveau d'association. Ainsi, si un QTL (Quantitative Trait Loci : locus impliqué dans un caractère n dans la descendance. Si pour deux marqueurs liés les différences de moyenne d'un caractère sont significatives, et si on suppose que cet effet est dû à un seul QTL à proximité des marqueurs, alors il est possible d'estimer la distance du QTL aux deux marqueurs ainsi que la proportion de la variance expliquée par ce QTL. La méthode de "interval mapping" (Luo, 1992) est basée sur le principe du maximum de vraisemblance et fonctionne de la façon suivante :

- le programme se déplace itérativement sur les chromosomes suivant un pas désiré ;
- à chaque pas, la vraisemblance de l'existence d'un QTL est confrontée à l'hypothèse contraire ; le LOD score, logarithme du ratio des deux dernières probabilités est calculé ;
- si le LOD score dépasse la limite de signification, alors un QTL a été détecté. L'endroit du génome pour lequel le LOD score est le plus grand porte vraisemblablement le gène recherché. Cette méthode permet ainsi la résolution des caractères polygéniques en caractères mendéliens.

Nous menons actuellement l'analyse moléculaire des 253 individus F2 issus du croisement *Mollissimum* x *Souna*. Cette technique va permettre la cartographie plus précise de chaque gène majeur (CAD, LPI), et devra faire apparaître clairement la duplication, si elle existe. En effet, si cette duplication est récente et que les deux segments chromosomiques sont faiblement divergents, alors les ségrégations des RFLP dupliqués seront anormales (les homozygotes seraient en fréquence 1/16). Si par contre les deux fragments dupliqués sont très différents quant aux marqueurs RFLP, alors chaque gène dupliqué sera lié à une bande RFLP spécifique dans le meilleur des cas (4 bandes par piste).

De plus l'ensemble des caractéristiques qui différencie sauvage et cultivé (hauteur, tallage, taille des graines, des chandelles, morphologie foliaire) pourront être associés à des RFLP marquant des QTL à effet important. Il sera intéressant de regarder quelle est alors la dispersion des caractères sur le génome. Chez le couple maïs-téosinte (Doebley, 1991), les divergences morphologiques sont issues en général de plus de deux mutations. Il faut toutefois remarquer que l'ensemble des mutations se groupent en cinq linkats sur les chromosomes 1 à 5, ce qui montre une certaine concentration des gènes de domestication. De plus le segment chromosomique du bras court du chromosome 1 contient le QTL le plus fort de cinq caractères : désarticulation de l'épi, nombre de cupules/rang, longueur des entre-nœuds, épillet pédicellé, pourcentage d'épillets mâles dans la première inflorescence. L'important chemin évolutif parcouru de la forme sauvage ancestrale au maïs cultivé actuel (une sélection humaine d'environ 10000 ans les sépare) explique le grand nombre de QTL mis en évidence lors de cette étude. Chez le mil, dont la domestication est plus récente (3000 ans), il n'existe pas de fossé entre type sauvage et cultivé : on s'attend à révéler un plus petit nombre de QTL.

Les travaux effectués sur la tomate (Tanksley, 1988), ont montré qu'il fallait rester prudent sur la valeur des QTL détectés lors d'une expérimentation, car ceux-ci dépendent à la fois des conditions extérieures, mais aussi du fond génétique et le transfert de QTL est souvent décevant quant aux résultats observés. On rencontre ici la complexité des régulations physiologiques et génétiques de la cellule, bien connue des physiologistes. Mais les gènes du syndrome de domestication sont avant tout des gènes de régulation du développement et de différenciation cellulaire, et leur relative insensibilité aux conditions du milieu garantit la stabilité des QTL sur différents contextes.

Ces outils moléculaires répondront donc aux principales questions qui persistent quant à l'organisation génétique du syndrome de domestication.

5) La cytogénétique moléculaire

Nous présentons ici le programme de recherche mené par Agnès Ricroch dans l'équipe de Heslop-Harrison (John Innes'Institute, UK).

Le développement des techniques de biologie moléculaire a créé un fossé important entre la connaissance de la structure moléculaire des gènes et la connaissance chromosomique du génome. L'emploi de techniques de détection immunologique très sensibles (Fluorescent *In Situ* Hybridization: FISH) et de nouvelles colorations en microscopie électronique (Hoescht) permet maintenant une cartographie moléculaire à échelle chromosomique d'une puissance considérable.

Le programme de recherche s'articule en deux phases : l'identification des chromosomes ; cette première phase a pour but l'établissement d'un caryotype du mil, par la recherche de sondes répétées spécifiques d'une paire chromosomique, visualisés par FISH sur des étalements métaphasiques et la cartographie physique de séquences d'ADN. Dans un premier temps, une vérification de l'ordre de certains marqueurs RFLP déjà cartographiés pourrait être effectuée par hybridation sur des noyaux interphasiques de plusieurs sondes marquées par des systèmes de détection différents (coloration distincte lors de la révélation). Puis une cartographie en métaphase, par "chromosome painting" et hybridation à des RFLP montrera l'organisation chromosomique et permettra la comparaison des cartes génétiques et physique.

IV - CONCLUSION-PERSPECTIVES

Dans la perspective d'une compréhension du processus évolutif de domestication en conditions sympatriques du mil, les résultats obtenus au niveau morphologique seront avantageusement complétés à l'aide des marqueurs moléculaires :

- La visualisation des distorsions de ségrégation est directe et permettra la cartographie de QTL de compétition pollinique par exemple. Les phénomènes de régulation des flux de gènes seront ainsi analysés ;
- Il s'en suivra une modélisation des phénomènes de domestication qui donnera des résultats plus fiables ;
- Les techniques cytogénétiques permettent de visualiser directement les remaniements chromosomiques et seront pour cela indispensables dans l'étude des phénomènes évolutifs à l'échelle de la famille.

Il faut noter toutefois que l'ensemble de ces études de l'information génétique nucléaire ne pourront être que d'une utilité réduite pour affiner les hypothèses de l'origine de la domestication du mil. En effet, lorsque seront caractérisés des RFLP spécifiques de couples *typhoides-monodii*, il sera difficile de savoir si l'hypothèse d'une domestication de novo en est l'origine ou s'il s'agit après migration d'introgessions du génotype cultivé par les formes sauvages locales. Pour y répondre, l'étude des génomes cytoplasmiques est indispensable.

Dans la perspective de l'amélioration du mil, l'utilisation des marqueurs pour l'introgression de gènes majeurs (Young, 1989) ou QTL forts dans des variétés commerciales est une des applications les plus directes, alors que la sélection assistée par marqueurs pose encore de nombreux problèmes théoriques. Leur utilité dans la valorisation du matériel sauvage sera fonction des caractères. En effet, la récupération de descendances bien adaptées dans un croisement large par sélection phénotypique sera très efficace pour tous les gènes majeurs, alors qu'elle sera laborieuse pour les caractères complexes peu héréditaires. Alors les marqueurs seront d'une grande utilité dans l'accumulation des gènes d'origine cultivée.

BIBLIOGRAPHIE

- ALLARD (R.W.), 1956. Formulas and tables to facilitate the calculation of recombination values in heredity. *Hilgardia* 24,10 : 235-278.
- BANUETT-BOURILLON (F.) and HAGUE (D.R.), 1979. Genetic Analysis Alcohol Dehydrogenase Isozymes in Pearl Millet (*Pennisetum typhoides*). *Biochemical Genetics*. 17. Nos. 5/6.
- BELLIARD (J.), NGUYEN VAN (E.) et PERNES (J.), 1980. Analyse des relations génétiques entre formes spontanées et cultivées chez le mil à chandelle. I Etude des parents et des hybrides de première génération entre un écotype de *P.mollissimum* et différentes formes cultivées de *P.americanum*. *Ann. Amélior. Plantes* 30 : 229-251.
- BENATSKY (R.) and TANKSLEY (D.), 1986. Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. *Genetics* 112 : 887-898.
- BENINGA (B.M.), 1981. Structure génétique du complexe des mils penicellaires : analyse des descendances issues d'hybrides entre formes cultivées et formes spontanées. Thèse de troisième cycle, Université de PARIS XI, ORSAY.
- BURTON (G.W.) and WERNER (B.K.), 1991. Genetic markers to locate and transfer heterotic chromosome blocks for increased pearl millet yields. *Crop Science* 31 : 576-579.
- ENJALBERT (J.), 1992. Contribution à l'analyse du déterminisme génétique du syndrome de domestication du mil. DEA de "Ressources Génétiques et Amélioration des Plantes", INAPG.
- JOLY-ICHENHAUSER (H.), 1984. Hérité du syndrome de domestication chez le mil (*Pennisetum typhoides*). Thèse de troisième cycle, Université de PARIS SUD ORSAY.
- LANDRY (B.S.), 1987. A genetic map of lettuce with RFLP, Isosymes and morphological markers. *Genetics* 116 : 331-337.
- LUO (Z.W.) and KEARSEY (M.J.), 1992. Interval mapping of quantitative trait loci in an F2 population. *Heredity* 69 : 236-242

- MARCHAIS (L.) and PERNES (J.), 1985. Genetic divergence between wild and cultivated pearl millet (*P. typhoides*). I. mâle sterility. *Zeitschrift für Pflanzenzücht.* 95 : 103-112
- MARCHAIS (L.) and TOSTAIN (S.), 1985. Genetic divergence between wild and cultivated pearls millets (*P. typhoides*). II. characters of domestication. *Z. für Pflanzenzücht.* 95 : 245-261.
- MC COUCH (S.R.), KOCHERT (G.), YU (Z.H.), WANG (Z.Y.), KHUSH (G.S.), COFFMAN (W.R.) and TANKSLEY (S.D.), 1988. Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* 76 : 815-829.
- PERNES (J.), NGUYEN VAN (E.), BENINGA (M.B.) et BELLIARD (J.), 1980. Analyse des relations génétiques entre formes spontanées et cultivées chez le mil à chandelle. II étude de trois familles F2 issues d'hybrides entre une plante d'un écotype de *P.mollissimum* et trois lignées de mil cultivé *P.americanum*. *Ann. Amélior. Plantes* 30 : 253-269.
- REY-HERME (C.), 1982. Les relations génétiques entre les formes spontanées et cultivées chez le mil. Thèse de troisième cycle, Université de PARIS XI ORSAY.
- ROBERT (T.) LESPINASSE (R.), PERNES (J.) and SARR (A.), 1991. Gametophytic competition as influencing gene flow between wild and cultivated of pearl millet. *Genome* 34 : 195-200.
- SANDMEIER (M.), BENINGA (B.) et PERNES (J.), 1981. Analyse des relations entre formes spontanées et cultivées chez le mil à chandelle. Etude de l'hérédité des estérases et des peroxydases anodiques. *Agronomie I* : 486-494.
- TANKSLEY (S.D.) and HEWITT (J.), 1988. Use of molecular markers in breeding for soluble solids content in tomato - a re-examination. *Theor.Appl.Genet.* 75 : 811-823.
- TOSTAIN (S.), 1985. Mise en évidence d'une liaison génétique entre un gène de nanisme et des marqueurs enzymatiques chez le mil pennicillaire (*Pennisetum glaucum*). *J. Canad. Genet. et Cyto.* 27 : 751-758.
- YOUNG (N.D.) and TANKSLEY (S.D.), 1989. RFLP analysis of chromosomal segments retained around the Tm2 locus of tomato during backcross breeding, *Theor. Appl.Genet.* 77 : 353-539.