

## CHAPITRE V

### DIVERSITE GENETIQUE DES MILS DETECTABLE PAR RFLPs AU NIVEAU DE LA REGION DU GENE ADH1

S. PILATE-ANDRE, F. LAMY, A. SARR

Laboratoire d'Evolution et Systématique Végétales, Bâtiment 362

Université de Paris Sud ; 91405 Orsay Cedex.

L'utilisation des marqueurs RFLPs chez le mil témoigne d'une importante variabilité tant dans l'étude de génotypes très divers que dans l'analyse d'individus originaires d'une même population. La diversité détectée par RFLPs au niveau de la région du gène *Adh1* est beaucoup plus importante que celle observée par analyse enzymatique avec concordance des profils RFLPs et des classes isozymiques. Les génotypes sauvages possèdent souvent des profils et des bandes spécifiques. Des cartes de restriction de la région du gène *Adh1* ont été réalisées pour six génotypes et six enzymes de restriction. Les quatre haplotypes obtenus présentent une région centrale de 5 Kb très conservée et des régions adjacentes variables. Des résultats analogues ont été obtenus chez le maïs et la drosophile. L'unité de transcription du gène *Adh1* est située dans la région conservée et nous avons pu en préciser la localisation et l'orientation par comparaison avec des publications récentes.

Ce travail démontre l'utilité des marqueurs RFLPs et différentes perspectives se présentent. De tels marqueurs peuvent constituer des outils précieux pour suivre des introgressions du génome de plantes sauvages dans des plantes cultivées. Le marquage du syndrome de domestication, groupe de gènes qui différencient les formes sauvages des formes cultivées, semble également intéressant ainsi que celui des gènes impliqués dans les phénomènes de compétition pollinique ou de tout gène contrôlant des caractères "intéressants", comme des résistances diverses (à des maladies ou des stress environnementaux). D'une façon plus générale, l'utilisation conjointe de différents types de marqueurs, morphologiques, isozymiques, RFLPs, RAPD et autres, présentant des intérêts différents et complémentaires, semble l'approche la plus judicieuse pour un grand nombre d'études (analyse de la diversité génétique, identification de relations phylogénétiques, cartographie de gènes impliqués dans des caractères intéressants).

#### I - INTRODUCTION

M. Sandmeier (Chapitre IV) a présenté dans son exposé les isozymes qui constituent des outils de marquage très utiles dans de multiples analyses. D'autres marqueurs génétiques, les RFLPs (Polymorphismes de Longueur des Fragments de Restriction) apparaissent également très intéressants (Beckmann et Soller 1986, Tanksley *et al.* 1989). Ces marqueurs permettent de détecter la variabilité directement au niveau de

l'ADN, qu'il s'agisse de séquences codantes ou non-codantes. Ceci représente un avantage par rapport aux marqueurs enzymatiques qui ne reflètent qu'indirectement la variabilité du génome puisque la détection du polymorphisme se fait au niveau d'une classe de protéines, les enzymes et non au niveau des gènes. Ainsi, un certain nombre de mutations ne peut être décelé. De plus, la proportion du génome représenté est très faible et biaisée puisqu'il s'agit d'une fraction des séquences codantes, celle qui code pour les enzymes considérés dans l'étude. Au contraire, les marqueurs RFLPs peuvent être très nombreux et répartis sur tout le génome d'une façon quasiment uniforme. Ils offrent ainsi une bonne représentation du génome. Cependant, les marqueurs RFLPs nécessitent des technologies beaucoup plus coûteuses que celles mises en œuvre pour les analyses isozymiques et beaucoup plus complexes.

## II - RESULTATS

Nous avons analysé le polymorphisme des longueurs des fragments de restriction au niveau de la région du gène *Adh1*, chez le mil. Nous avons utilisé une sonde cDNA *ADH1* de maïs (Dennis *et al.* 1984) couplée à plusieurs enzymes de restriction. Les objectifs étaient principalement :

- d'apprécier la variabilité détectable par RFLPs chez le mil ;
- de comparer la diversité RFLP à la variabilité isozymique ;
- d'étudier finement la région chromosomique du gène *Adh1*.

Les géotypes choisis diffèrent à plusieurs titres (Tableau 1) : selon leur origine géographique (Afrique de l'ouest, Afrique de l'est ou autres origines), selon leur appartenance à différents compartiments du pool primaire du complexe des mils (cultivé, spontané ou intermédiaire), selon leur degré de consanguinité (ces géotypes ont eu de 1, pour la plupart, à 7 cycles d'autofécondation) et enfin selon leur profil enzymatique pour le système *Adh*. Chez le mil, il existe deux gènes *Adh* : *Adh1* et *Adh2*. Tous ces géotypes sont homozygotes *fast* pour le gène *Adh2* mais ils forment trois classes isozymiques pour le gène *Adh1* (*slow / fast / superfast*).

L'analyse RFLP a été réalisée avec 4 enzymes de restriction (*Bst* EII, *Eco* RI, *Bgl* II et *Hind* III). Quel que soit l'enzyme de restriction, le nombre de profils différents est élevé, de même le nombre de bandes est important (jusqu'à 7 par profil). Ceci est révélateur d'une famille multigénique. En effet, la sonde de maïs, hétérologue, hybride très certainement les deux gènes *Adh* de mil puisqu'il existe de grandes homologues de séquences entre les gènes *Adh* aussi bien au sein d'une même espèce qu'entre espèces différentes (Xie et Wu 1989).

### A - Profils RFLP

Le schéma des profils obtenus avec l'enzyme de restriction *Bst* EII est représenté figure 1. Les plantes sauvages (*Mollissimum* et 968) présentent des profils différents de ceux des plantes cultivées qui comportent souvent des bandes spécifiques. De même, le géotype intermédiaire possède également son propre profil.

Tableau 1

Profil enzymatique	Origine Géographique						
	Afrique de l'Ouest		Afrique de l'Est		Autre origine		
SLOW	Massue	(C,4)	Sénégal	990	(C,1)	Soudan	
	Ligui	(C,3)	Tchad	1041	(I,1)	Soudan	Chinois (C,1) Chine
	Tioandé	(C,3)	Sénégal			23DB	(C,3) USA
	MxT	(C,7)	Sénégal			Drôo	(C,4) Tunisie
	Molliss	(S,5)	Mali				
	inum						
FAST	781	(C,1)	Mali				
	968	(S,1)	Mali				
	1719	(C,1)	Burkina Faso				
	1746	(C,1)	Burkina Faso				
	1732	(C,1)	Burkina Faso				
	362	(C,0)	Burkina Faso				
SUPERFAST				1268	(C,1)	Tanzanie	
				1091	(C,1)	Kenya	

La lettre figurant entre parenthèses représente le type cultivé (C), sauvage (S) ou intermédiaire (I), le chiffre mentionné à la suite de cette lettre indique le nombre de cycles d'autofécondations.

## PM (Kb)

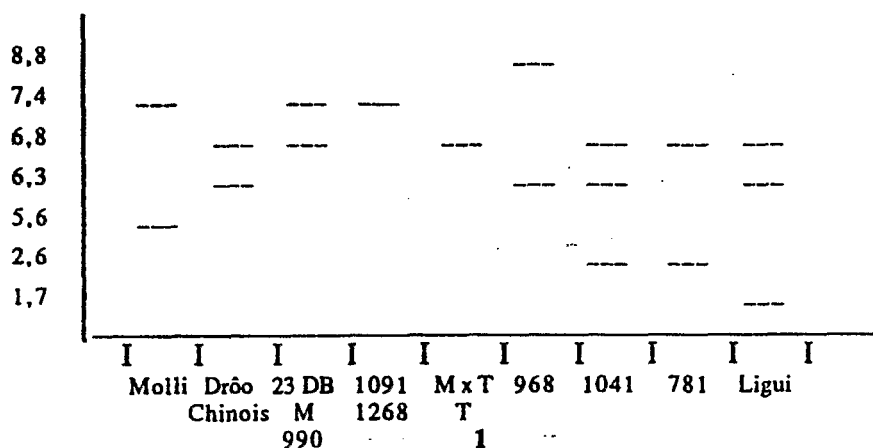


Figure 1

Schéma des profils RFLPs obtenus avec l'enzyme de restriction BstE II

BstEII : 9 profils sur 7 niveaux de bandes.

Non représenté : 4 bandes communes (poids moléculaires : 2,3 ; 3,6 ; 3,75 ; 14Kb)

Les différents profils RFLPs sont en concordance avec les trois classes isozymiques du gène *Adh1* (*slow* / *fast* / *superfast*) :

- les deux génotypes *slow* (Massue et 990) présentent toujours le même profil RFLP quel que soit l'enzyme de restriction utilisé. Il en est de même des deux génotypes *superfast* (1091 et 1268) ;
- les génotypes *fast* présentent différents profils RFLPs qui se distinguent cependant de ceux des plantes des deux autres classes isozymiques (*slow* et *superfast*). Notons que cette classe enzymatique, la plus représentée dans les populations naturelles de mil, comporte l'effectif le plus élevé de notre étude (13 génotypes sur 17) ;
- le génotype 23DB, de type *fast* au niveau enzymatique, constitue une exception puisqu'il présente, pour trois des quatre enzymes de restriction, un profil RFLP identique à celui des plantes de type *slow*.

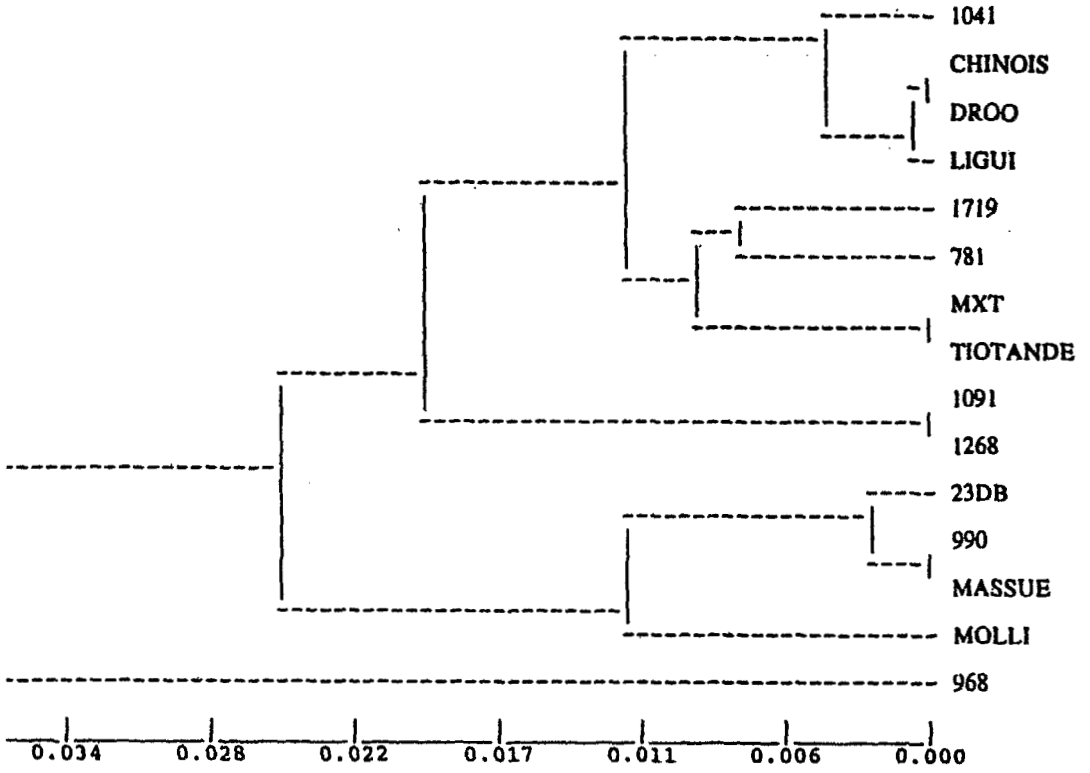
### B - Classification des profils RFLP

Les différences géographiques paraissent moins influencer sur les groupes obtenus par RFLPs que les différences de profil enzymatique. Cependant, l'Afrique de l'est n'est représentée que par quatre génotypes contre dix pour l'Afrique de l'ouest et le biais introduit ne permet pas de conclure. De même, il n'existe pas de relation apparente entre le nombre de bandes révélées par RFLPs et le niveau de consanguinité des génotypes. Ces résultats ont été analysés par le logiciel RESTSITE (Nei et Miller 1990) qui permet le calcul de distances à partir de données RFLPs. Le dendrogramme obtenu (cf figure 2) confirme les résultats discutés précédemment, à savoir notamment la concordance des profils RFLPs avec les trois classes isozymiques. L'un des deux génotypes spontané (968) est manifestement très éloigné des autres génotypes, ce qui n'est pas le cas de Mollissimum, autre génotype sauvage de notre étude. Il est également très intéressant de noter le grand rapprochement de trois génotypes d'origine géographique très éloignée (Chinois, Drôo et Ligui).

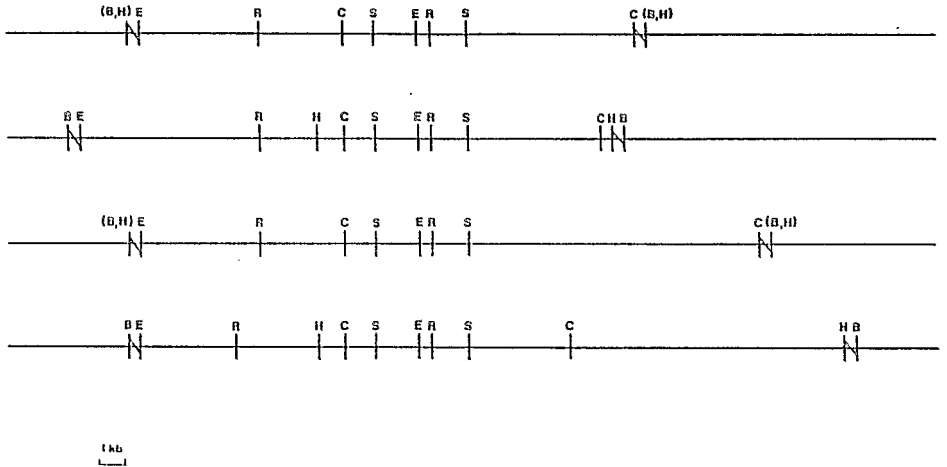
En analysant des génotypes choisis les plus différents possibles, nous avons donc pu observer un polymorphisme très important. Qu'en est-il si l'on considère des individus originaires d'une même population ? Pour le savoir, nous avons analysé 48 individus d'une population originaire du Burkina-Faso enzymatiquement monomorphe et homozygote pour Adh1 et Adh2 (allèles *fast*), avec quatre enzymes de restriction (Bst EII, Eco RI, Eco RV et Hind III). A l'exception de Eco RI qui présente seulement trois profils différents, ces enzymes de restriction génèrent de très nombreux profils différents (de 11 à 14). Ainsi nous observons un polymorphisme très important à l'intérieur d'une même population et à l'intérieur d'une même classe isozymique.

### C- Cartographie du gène ADH 1

L'analyse fine de la région chromosomique du gène Adh1 a été réalisée en prenant six génotypes et six enzymes de restriction. Comme indiqué précédemment, la sonde de maïs, hétérologue, hybride les deux gènes de mil et pour réaliser les cartes de restriction, nous avons considéré la bande d'intensité la plus forte. Dans tous les cas, il n'en existait qu'une seule par profil, à l'exception de Mollissimum qui en présentait généralement deux, qui ont été prises en compte. De cette façon, nous avons obtenu quatre haplotypes



**Figure 2**  
Dendrogramme RESTSITE

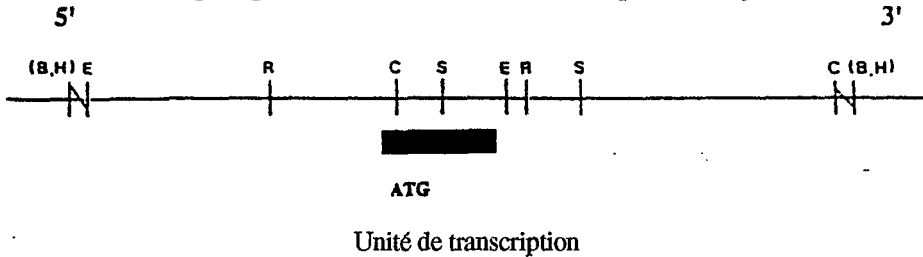


**Figure 3**  
Cartes de restriction au niveau de la région chromosomique Adh 1

- |   |  |             |              |
|---|--|-------------|--------------|
| a | haplotype 1 : Massue et Molli 1        | B = BamH I  | E = EcoR I   |
| b | haplotype 2 : M x T, Tiotandé et Ligui | C = Bcl I   | R = EcoR V   |
| c | haplotype 3 : 23 DB                    | S = BstE II | H = Hind III |
| d | haplotype 4 : Molli 2                  |             |              |

(cf. figure 3). L'observation de ces cartes montre que la région centrale est très conservée pour les six génotypes alors que les régions situées à droite et à gauche sont très variables. La région conservée consiste en 5 Kb d'ADN marqué à gauche, sur le schéma, par un site de restriction Bcl I et à droite par un site Bst E II. En dehors de cette région, le polymorphisme des sites de restriction est de règle. Sur ces cartes il est impossible de distinguer M x T, T et Ligui. Il en est de même pour Massue et Mollissimum 1, dont les différences avec 23DB ne s'observent que pour les sites Bcl I, Hind III et Bam HI situés les uns près des autres, à une distance supérieure à 10 Kb de la région conservée. Ceci pourrait être dû à un événement de type insertion/délétion puisque trois sites sont simultanément touchés ce qui pourrait résulter de l'action d'IT. Il pourrait s'agir d'un transposon. Chez le maïs la présence de transposons au niveau du gène *Adh1* a déjà été mise en évidence (Osterman and Schwartz, 1981).

Le gène *Adh1* de mil a été récemment séquencé (Bui Dang Ha *et al.* 1990, Gaut et Clegg 1991). En comparant ces différents résultats, nous pouvons localiser et orienter l'unité de transcription du gène *Adh1* sur nos cartes. Les deux sites de restriction Bcl I et Bst EII, situés à gauche, sur notre schéma, de la région conservée chez les six génotypes, se trouvent, respectivement dans le premier et le quatrième exon de l'unité de transcription du gène *Adh1*. Plus précisément, le site Bcl I est localisé à une vingtaine de nucléotides du codon d'initiation, et le site de restriction Eco RI de la région conservée est situé peu après la fin de cette unité de transcription (cf. figure 4).



**Figure 4**

Orientation de la carte *Adh 1* et localisation du codon d'initiation

### III - CONCLUSIONS ET PERSPECTIVES

L'utilisation de marqueurs RFLPs, au niveau des gènes *Adh* chez le mil a permis de détecter une variabilité très importante, que l'on considère des génotypes choisis les plus différents possibles ou des individus originaires d'une même population. Il faut noter également la concordance des profils RFLPs avec les trois classes isozymiques pour *Adh1*

Les cartes de restriction de la région du gène *Adh1*, réalisées pour six génotypes et six enzymes de restriction présentent une région centrale de 5 Kb très conservée et des régions adjacentes très variables. Des résultats analogues ont été obtenus chez le maïs et la drosophile.

Ainsi, les marqueurs RFLPs constituent des outils très précieux que ce soit pour analyser globalement la variabilité génétique, pour étudier plus finement une région chromosomique donnée ou encore pour suivre des introgressions du génome de plantes sauvages dans des plantes cultivées. Une carte génétique RFLPs du mil est en cours d'élaboration par l'équipe de M. Gale à Norwich (UK) et de nombreuses perspectives, à plus ou moins long terme se présentent. Le marquage du syndrome de domestication, groupe de gènes liés différenciant les formes sauvages des formes cultivées (Pernès 1983), a été initié par l'équipe de A. Sarr dans notre laboratoire. Le marquage d'autres gènes semble également intéressant, comme celui des gènes impliqués dans les phénomènes de compétitions polliniques (Sarr *et al.* 1988, Robert 1989) ou de tout gène contrôlant des caractères jugés "intéressants", comme des résistances à des maladies diverses ou à des stress environnementaux. A beaucoup plus long terme il peut être envisageable de développer une sélection assistée par marqueurs.

Quel que soit le type d'étude entrepris, l'approche la plus judicieuse est d'utiliser en combinaison les différents types de marqueurs (morphologiques, enzymatiques, RFLPs, et autres...) car ils sont différents et complémentaires. Dans cette optique, la création d'une carte génétique "mixte" comportant tous les marqueurs génétiques possibles paraît indispensable.

### BIBLIOGRAPHIE

- BECKMANN (J.S.) and SOLLER (M), 1986. Restriction fragment length polymorphisms and genetic improvement of agricultural species. *Euphytica* 35 : 111-124.
- BUI DANG HA (D), BUFFARD (D), BERGER (F), BRÉDA (C) and ESNAULT (R), 1990. Nucleotide sequence encoding a slow allele of *Adh1* in pearl millet. *Plant Mol. Biol.* 14 : 453-455
- DENNIS (E.S.), GERLACH (W.L.), PRYOR (A.J.), BENNETZEN (J.L.), INGLIS (A), LLEWELLYN (D), SACHS (M.M.), FERL (R.J.) and PEACOCK (W.J.). 1984. Molecular analysis of the alcohol dehydrogenase (*Adh1*) gene of maize. *Nucleic. Acids Res.* 12 : 3983-4000.
- GAUT (B.S.), CLEGG (M.T.), 1991. Molecular evolution of alcohol dehydrogenase 1 in members of the grass family. *Proc. Natl. Acad. Sci. USA* 88 : 2060-2064.
- NEI (M) and MILLER (J.C.), 1990. A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics* 125 : 873-879.
- OSTERMAN (J.C.) AND SCHWARTZ (D.), 1981. Analysis of a controlling-element mutation at the *Adh* Locus of maize. *Genetics* 99 : 267-273.
- PERNÈS (J), 1983 La génétique de la domestication des céréales. *La Recherche* 14 : 910-919.



- PILATE-ANDRÉ (S), 1992. Etude de l'organisation de la diversité génétique du complexe des mils pénicillaires (*Pennisetum* spp.) par les marqueurs enzymatiques et par l'analyse moléculaire de la région Adh. Thèse de doctorat. Université Paris XI, Orsay. 220 pp.
- ROBERT (T), 1989. Dynamique des flux de gènes entre formes sauvages et cultivées du mil (*Pennisetum typhoides* Stapf et Hubb.). Impact des sélections gamétophytiques. Thèse de doctorat. Université Paris XI, Orsay. 123 pp
- SARR (A), SANDMEIER (M) and PERNÈS (J), 1988. Gametophytic competition in pearl millet *Pennisetum typhoides* (Stapf and Hubb.). *Genome* 30 : 924-929.
- TANKSLEY (S.D.), YOUNG (N.D.), PATERSON (A.H.) and BONIERBALE (M.W.), 1989. RFLP mapping in plant breeding : new tools for an old science. *Bio/Techno* 7 : 257-264.
- XIE (Y) AND WU (R), 1989. Rice alcohol dehydrogenase genes : anaerobic induction, organ specific expression and characterisation of cDNA clones. *Plant Mol Biol* 13 : 53-68.