Theor Appl Genet (1987) 74:188–193

THEORETICAL AND APPLIED GENETICS

Enzyme diversity in pearl millet (*Pennisetum glaucum*)

1. West Africa

S. Tostain, M. F. Riandey and L. Marchais

ORSTOM, BP 11416, Niamey, Niger

Received July 9, 1986; Accepted February 18, 1987
Communicated by G. S. Khush

Summary. Polymorphism in twelve genes coding for eight enzymes in pearl millet (*Pennisetum glaucum* (L.) R. Br.): alcohol dehydrogenases (ADH), catalases (CAT), β-esterases (EST), glutamate oxaloacetate transaminases (GOT), malate dehydrogenases (MDH), 6-Phosphogluconate dehydrogenases (PGD), phosphoglucoisomerases (PGI) and phosphoglucomutases (PGM), was observed by electrophoresis on 74 cultivated samples and 8 wild samples from West Africa. Six genes: *Est A, Adh A, Pgm A, Cat A, Pgi A, Pgd A* contain 95% of the total variation. Principal component analyses and discriminant analyses of the 82 samples described by 46 allelic frequencies showed an almost complete separation into 3 groups: wilds, early maturing cultivars and late maturing cultivars. The early group has the highest enzyme diversity, with cultivated millets from Niger showing the most diversity. The high diversity of the early group and its extensive divergence from West-African wild millets suggest, firstly, the existence, elsewhere in Africa of other enzymatically different sources of wild millet, and secondly, the occurrence, prehistorically, of several different domestications. The late group of cultivars has the lowest variability and a relatively low coefficient of differentiation. This relatively homogeneous enzyme structure does not seem to be associated to ecology. A hypothesis is advanced suggesting that West African late-cultivars were derived from a common cultivated early complex. This complex must have been distributed across the Sudanian zone and must have been later sumitted to modifications by limited gene flow with local early maturing cultivars.

Key words: Isoenzymes – Evolution – Domestication – Biological structure – Genetic distances

Introduction

Since 1976, several institutes have cooperated in gathering an extensive collection of the existing wild and cultivated forms of pearl millet (*Pennisetum glaucum* (L.) R. Br.). Simultaneously, a genetic resources evaluation of the collection has been undertaken, comprising an evaluation of gene diversity in the different populations and their mutual distances. This evaluation should help reveal species biological structure, outline certain elements in pearl millet evolution (centers of origin, centers of domestication, routes of migrations) and consequently be of use to millet breeders. Bono (1973) and Marchais (1982) have demonstrated the similarity between Malian and Burkinabe millets on the one hand and Nigerien and Senegalese millets on the other. They stressed the fact that Malian millets are not intermediaries between Senegalese and Nigerien millets. On the basis of botanical observations, Portères (1962) identified 4 centers of diversity for pearl millet in Africa. Harlan (1971) and Brunken et al. (1977) hypothesized the existence of several independent domestications of pearl millet from the southern fringe of the Sahara. This study of pearl millet enzyme diversity was undertaken to try to answer some of these questions.

Materials

Seventy-four cultivated populations and eight wild populations of pearl millet collected in West Africa by ORSTOM, with the financial support of IBPGR, have been analyzed. The cultivated millets were collected in Sénégal (12), Mali (13), Niger (13), Burkina Faso (10), Togo (7), Côte d'Ivoire (5), Bénin (3), Cameroun (7) and République Centrafricaine (4). The wild millets come from western Mali (2), eastern Mali (2), Niger (3) and Burkina Faso (1).

Information on collections are reported in Clément (1985). In most of these countries, the cultivated landraces are classified by the farmers into early and late maturing types. Their a priori grouping was checked by sowing the cultivated samples on June 25, 1985 at the ICRISAT Sahelian Centre close to Niamey. A continuum of heading dates was observed from August 5 to October 20 with a rarefaction of headings between September 5 and 10. The following practical criterion was chosen: millets heading before September 6 (72 days) were classified as early; millets heading after September 5 were classified as late.

This experimental classification revealed some evident errors in the classification of samples made by the collectors. On the whole, Niger was represented solely by early landraces whereas the Côte d'Ivoire and Centrafrique were represented uniquely by late landraces. The other countries were represented by early and late landraces which in 7 cases come from the same village. The seeds analyzed come from original collected samples, each sample generally consisting of 3–4 flowering heads collected either in the field or from village granaries. For each sample and each enzyme, 24 plants were analysed.

Methods

a) Enzymes

Eight enzyme systems were observed: β-esterases (EST), alcohol dehydrogenases (ADH), phosphoglucoisomerases (PGI), phosphoglucomutases (PGM), glutamate oxaloacetate transaminases (GOT), catalases (CAT), malate dehydrogenases (MDH) and phosphogluconate dehydrogenases (PGD). Electrophoresis techniques and the genetics particular to each enzyme system have already been published: ADH (Banuett-Bourrillon and Hague 1979; Tostain and Riandey 1984), EST (Sandmeier et al. 1981; Tostain and Riandey 1984), GOT (Tostain and Lavergne 1985), MDH (Tostain and Riandey 1985) PGI and PGM (Leblanc and Pernès 1983), CAT (Tostain and Riandey 1984). The methodology with respect to PGD will be presented in detail in a subsequent paper. The PGD system includes one gene *Pgd A*, coding for a dimeric enzyme PGD A and two duplicate genes, *Pgd B* and *Pgd C*, coding for dimeric isozymes composed of intralocus dimers (PGD B, PGD C) but also of interlocus hetero-dimers (PGD B–C). PGD B and PGD C subunits do not form hetero dimers with PGD A subunits.

On the whole, the data for each millet sample provide a vector of 46 allelic frequencies associated with 12 genes.

b) Data analysis

The samples polymorphism has been measured by Nei's parameters (1975):

– Gene diversity of sample X, $Hx = \Sigma_i (1 - \Sigma_a X_{ai}^2) . X_{ai}$ denotes the frequency of allele a of gene i, for sample X.

– Coefficient of differentiation between samples of a group: $Gst = (Ht - Hs)/Ht$. Ht denotes the group gene diversity. Hs denotes the mean of sample diversities.

– The minimum distance between two samples X, Y: $Dm\ (X, Y) = \Sigma_i \Sigma_a (X_{ai} - Y_{ai})^2/2$. The intralocus variance of a group's diversity has been computed by a formula given by Brown and Weir (1983):

$$\text{Variance (Ht)} = \Sigma_i [\Sigma_a X_{ai}^3 - (\Sigma_a X_{ai}^2)^2] [4 + (K - 1)]/2\ gk$$

where g = number of samples in the group, K = number of individuals in the sample (here, K = 24) and X_{ai} = frequency of allele a of gene i in the group.

This formula is valid in the case of g individuals being randomly sampled in a panmictic population. For each individual, a constant number k of offspring is analyzed. Nei and Roychoudhury's formula (1973) is numerically very close to the previous one if the total number of alleles sampled (2 gk) is divided by the correction factor $1 + (k-1)/4$ given also by Brown and Weir (1983) to compute the variance of X_{ai}. Variance of the distance Dm has been computed by Nei and Roychoudhury's formula (1973), but with the "corrected" number of alleles.

Principal component analyses computed on the 46 unstandardized allelic frequencies were used to give a spatial representation of Nei's distances. An additional fictitious sample was included as a milestone in the scenery (marked by a circled star in the figures): it is a maximum diversity sample and has equi-frequent alleles for each locus. Spatial representations of other distances such as Arc of Cavalli-Sforza and Edwards (1967), Nei's standard distances (1975) were also obtained by the principal coordinate method (Gower 1966). The groups identified in the principal component analyses were treated by discriminant analysis to verify their cohesion.

Results

1 Single locus diversity

The 12 loci studied display 46 alleles with very variable frequencies: 24 alleles have a frequency above 0.05 in all the samples, 6 alleles are locally frequent (their frequency is above 0.05 in a small number of samples), 16 alleles are rare in all the samples (Table 1). Six loci: *Est A, Adh A, Pgm A, Cat A, Pgi A* and *Pgd A* contain 95% of the total diversity. Locus *Pgm A* has a high diversity index although it has only 2 alleles.

Table 1. Gene diversity, locus by locus, computed from the general means of frequencies

Locus	No. of alleles				Nei's diversity
	Total	Frequent[a]	Locally frequent[b]	Rare[c]	
Est A	7	6	1	0	0.827
Adh A	9	4	1	4	0.684
Pgm A	2	2	0	0	0.500
Cat A	2	2	0	0	0.354
Pgi A	4	2	1	1	0.282
Pgd A	3	2	0	1	0.157
Got A	2	1	1	0	0.054
Mdh A	5	1	1	3	0.041
Got B	4	1	0	3	0.018
Mdh C	3	1	0	2	0.016
Pgd B	2	1	1	0	0.014
Mdh B	3	1	0	2	0.000
Total	46	24	6	16	2.948

[a] Mean frequency greater than 0.05 for all 82 samples
[b] Frequency greater than 0.1 from a small number of samples
[c] Alleles rare in all samples

2 Differentiation between wild millets, early maturing cultivars, late-maturing cultivars

Principal components were computed on all samples to discern possible structure. The first two components accounted for 31.9% and 22.2% respectively of the total variance suggesting a strong association between some alleles and the existence of groups of samples. Indeed, three well-defined groups can be seen on the plane (1, 2): wild millets, cultivated early-millets, cultivated late-millets (Fig. 1). The point of maximum diversity in the middle of the early cluster suggests that early-millets are the most diverse. The clearcut divergence between early and late groups can also be seen between the 7 early/late couples originating from the same village. A discriminant analysis made on these 3 groups fully confirms their existence (Table 2). However, they are a priori groupings and they would not appear as distinct groups if they were not predetermined. The wild sample classified as "early" is a sample from western Mali enzymatically close to cultivated Malian early samples (Fig. 2). The other wild samples, especially those from Niger, are significantly distant from the cultivated group. Genetic parameters displayed in Table 3 corroborate the previous observations: the early cluster is the most diverse although its diversity is not statistically superior to that of wild millets. On the other hand, wild and early clusters possess a diversity significantly greater than that of the late one. The low diversity of the late group is associated with a relatively low coefficient of differentiation. The wild-late distance is significantly greater than wild-early distance. Measurements of distances place early millets between wild and late millets. The divergence between wild and early groups and the divergence between early and late groups are characterized by very different sets of allozymes (Table 4). They have probably been created by independant and different causes. Note also that the locally frequent allele *Got A1* is common in all the wild samples but virtually absent from cultivated samples.

Spatial representations obtained with other distances appeared very similar to the previous one and

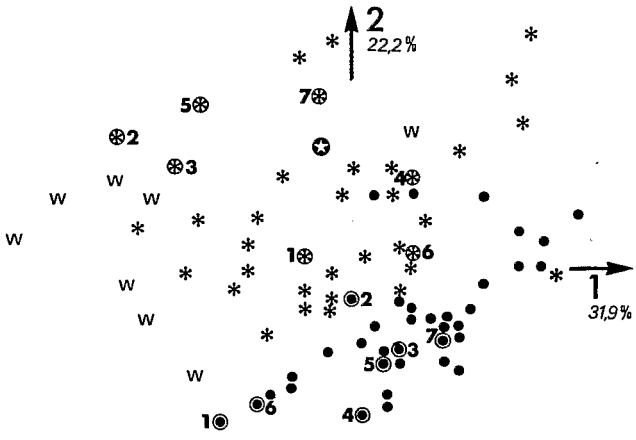

Fig. 1. Projection on the plane (1, 2) of the first two principal components of 8 wild samples (W), 37 cultivated early samples (*) and 37 cultivated late samples (●). The seven early-late pairs from the same village are numbered. 1 = Sénégal; 2 = Mali; 3 = Mali; 4 = Bénin; 5 = Burkina Faso; 6 = Burkina Faso; 7 = Togo. ✪ = the point of maximum diversity

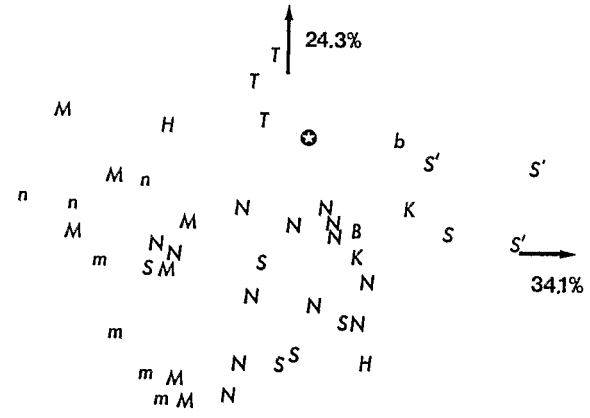

Fig. 2. Principal component analysis of the cultivated early samples with the wild samples as additional elements. Projection of wild and early samples on the plane (1, 2). M = Mali; S = Sénégal; N = Niger; T = Togo; B = Bénin; K = Cameroun; H = Burkina Faso; S′ = Tiotandé from Sénégal. m = wild samples from Mali; n = wild samples from Niger; b = wild samples from Burkina Faso. ✪ = point of maximum diversity

Table 2. Classification by discriminant analysis

Initial group	No. of cases	Final group		
		Wild	Cultivated	
			Early	Late
Wild	8	7	1	0
Cultivated				
Early	37	0	36	1
Late	37	0	1	36

Table 3. Genetic parameters of the wild, cultivated-early and cultivated-late groups

Group	Nei's gene diversity	Coefficient of differentiation	Nei's minimum distance[a]		
			Wild	Early	Late
Wild	2.870±0.147	0.202	–	0.268	0.500
Early	3.015±0.065	0.176	±0.063	–	0.197
Late	2.551±0.062	0.124	±0.083	±0.030	–

[a] Nei's minimum distances are found above the diagonal and the standard deviations below it

Table 4a. Characteristic allozymes differentiating wild and cultivated early millets

Allozyme	Mean frequency		Test T[a] (significant at 1% level)
	Wild	Early	
Got A1	0.235	0.012	5.43
Pgi A3	0.973	0.821	5.20
Adh A5	0.004	0.067	3.88
Adh A4	0.650	0.318	3.60
Adh A7	0.000	0.050	3.45

Table 4b. Characteristic allozymes differentiating cultivated early and late millets

Allozyme	Mean frequency		Test T[a] (significant at 1% level)
	Early	Late	
Cat A1	0.648	0.941	5.22
Adh A7	0.050	0.210	3.95
Est A6	0.277	0.125	4.49
Est A3	0.151	0.270	3.47
Est A7	0.196	0.091	3.92
Est A2	0.157	0.244	3.27

[a] Approximate test given in Snedecor (1985), p 98

thus only the Nei's Dm distance will be considered in the remainder of this paper.

3 Intra-group structure

A principal component analysis was carried out on the early group only with wild samples as additional elements (Fig. 2). The plane (1, 2) displays 58.4% of the total variability and distinctive regional groups.

Early Niger samples with two frontier samples from Burkina Faso and Bénin occupy the centre of the cluster. Early Mali samples form a distinct group close to Malian wild samples. Niger wild millets show greater similarity to cultivated Malian millets than to Niger cultivated millets. The Iniadi landrace samples from southern Burkina Faso and northern Togo make another homogeneous group. The Mali-Niger-Cameroun linear order suggests the existence of a west-east geographical cline but the two distinctive early Senegalese landraces (S, S') occupy positions which do not conform to geographical order. Principal component analysis carried out on the late group also displays marked regionalization in conformity with geography (Fig. 3). Late samples from Mali and Burkina Faso make a common central group surrounded on the "west" by Sénégal, on the "east" by Cameroun and République Centrafricaine, to the "south" by Côte d'Ivoire, Togo, Bénin. Both early (Sénégal excluded) and late groups display the same allelic evolution from west to east:

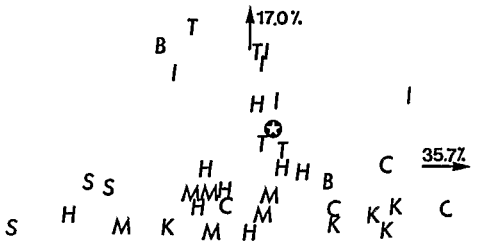

Fig. 3. Principal component analysis of the late samples. Projection on the plane (1, 2). S = Sénégal; M = Mali; H = Burkina Faso; I = Côte d'Ivoire; B = Bénin; T = Togo; K = Cameroun; C = Centrafrique. ✪ = point of maximum diversity

Pgm A1 and *Pgi A5* increase, *Adh A4* and *Est A4* decrease in frequency.

As a rule, comments made on principal component analyses are corroborated by Nei's genetic parameters (Table 5). Niger displays the highest diversity associated with a relatively low coefficient of differentiation, which means that Niger millets tend to be a same and rich large population. Countries with predominantly late varieties display a statistically lower diversity: Burkina Faso, Côte d'Ivoire, République Centrafricaine. The high Senegalese coefficient of differentiation is explained by the three very different enzymatic groups corresponding to the three landraces: Souna, Tiotandé, Sanio. The regionalization of enzyme diversity is shown in that all the distances between countries are statistically different from zero. That regional differentiation is more or less in line with geographical distances. Three exceptions can be seen: early Senegalese millets are closer to Niger than to Mali. Late Malian millets have a diversity higher than that of early Malian millets. Côte d'Ivoire is closer to Cameroun than to Mali.

Discussion

The main results of this paper are (1) the existence in West Africa of three enzymatic groups of pearl millet (wild, cultivated early and cultivated late), (2) direct correlation globally between enzymatic and geographic distances (3), the location of the maximum diversity in Niger. These new data should stimulate reflection on pearl millet domestication, evolution and breeding.

The few wild accessions from Mali and Niger that have been observed make a homogeneous group, far from the maximum diversity point and enzymatically less diverse than cultivated early accessions. This is unexpected because domestication usually reduces the genic diversity as shown by Second (1982) in rice. That anomaly can be explained if we suppose that wild millets from Mali and Niger are not the sole origin of all cultivated millets but that other and different wild

Table 5. Genetic parameters of cultivated millets from various countries

Country	Total diversity	Coefficient of differentiation (Gst)	Mean of diversities		Nei's minimum distances[a]						
			Early	Late	Sénégal	Mali	Niger	Burkina Faso	Bénin, Togo	Côte d'Ivoire	Cameroun	Centrafrique
Sénégal	2.817±0.120	0.232	2.25	1.96	–	0.140	0.075	0.090	0.152	0.228	0.088	0.16
Mali	2.824±0.099	0.133	2.41	2.49	±0.048	–	0.100	0.112	0.167	0.300	0.224	0.34
Niger	2.913±0.108	0.103	2.61	–	±0.031	±0.039	–	0.121	0.130	0.266	0.109	0.25
Burkina Faso	2.597±0.113	0.145	2.34	2.19	±0.036	±0.048	±0.041	–	0.102	0.146	0.112	0.19
Benin Togo	2.872±0.108	0.182	2.84	2.02	±0.046	±0.057	±0.043	±0.046	–	0.098	0.134	0.23
Côte d'Ivoire	2.589±0.159	0.070	–	2.41	±0.067	±0.090	±0.075	±0.062	±0.053	–	0.200	0.21
Cameroun	2.851±0.154	0.069	2.85	2.57	±0.044	±0.077	±0.055	±0.056	±0.054	±0.073	–	0.11
Centrafrique	2.168±0.187	0.081	–	1.99	±0.062	±0.102	±0.075	±0.077	±0.075	±0.090	±0.058	–

[a] Nei's minimum distances are found above the diagonal and the standard deviation below it

millets have been domesticated elsewhere, such as in Chad and Sudan, and have provided additional diversity to wild germplasm.

The enzymatic west-east cline observed among the cultivated millets is difficult to explain simply by environmental factors, given the uniformity of the Sahelian and Savannah environment. Harlan's non-centre hypothesis (Harlan 1971) can easily explain this fact: cultivated millets would issue from multiple domestications on different and independant wild sources, scattered along the southern margin of the Sahara. Gene flow between centres of domestication would thus produce the observed gradient. The maximum diversity of Niger accessions would result from their intermediate position at the centre of reciprocal gene flow between centres of domestication. The specific enzyme composition of late-millets and their low diversity does not seem to be caused by natural selection in the wet tropical climate, since these characteristics are not found in the early samples from the same ecological zone. The high divergence between wild and late millets suggests that domestication produced early millets which later on produced late millets. However, the enzymic structure common to all late types would be highly improbable in the hypothesis of multiple independant origins. The simplest explanation consists of supposing that the populations of the late-millets were generated by an old southward migration of a particular early population. That population became late-cycled and would have spread eastward and westward along the forest edge. Finally, these late millets would have established bridges of limited gene flow with local, early northern millets.

Enzymatic and morphological characters (Bono 1973) agree in demonstrating the proximity between Senegalese early-maturing varieties and Niger landraces. On the contrary, the Tiotandé landrace, which was morphologically close to early-maturing landraces from West Mali, is very dissimilar to these on the basis of enzymes. Another discordance between enzymes and morphological characters appears in the position of late-maturing samples from Sénégal and Mali, relatively to Niger samples (Marchais 1982). Discordance between enzyme and morphological distances can be foreseen if enzymes are neutral to natural and human selection whereas morphological characters are evidently submitted to human selection. The high enzymic divergence between Tiotandé and wild millets agrees well with the phenomenon of male sterility observed in a cross between Tiotandé and a wild Senegalese millet (Marchais and Pernès 1985).

The discovery of distinct groups of pearl millet by enzyme electrophoresis is encouraging. The study of enzyme diversity could be instrumental in understanding the evolutionary processes in pearl millet species.

Significant information can reasonably be expected from the enzymic study of wild and cultivated millets from central and eastern Africa and also from India. The enzymic survey of west-African pearl millets should help the management of gene banks. It should help also breeders to choose their source material among the available collections.

Acknowledgements. Our gratitude goes to Mr. John Newby (W.W.F. – Niamey) who corrected the English version of this paper. We want also to express our thanks to our colleagues Tchala M. and Billard G. for their contribution to this research.

References

Banuett-Bourrillon F, Hague DR (1979) Genetic analysis of alcohol dehydrogenase isozymes in Pearl Millet (*Pennisetum typhoïdes*). Biochem Genet 20:369–383

Bono M (1973) Contribution à la morphosystématique des *Pennisetum* annuels cultivés pour leur grain en Afrique occidentale francophone. Agron Trop (Paris) 28:229–356

Brown AHD, Weir BS (1983) Measuring genetic variability in plant populations. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, part A. Elsevier, Amsterdam Oxford New York, pp 219–239

Brunken J, de Wet MJ, Harlan JR (1977) The morphology and domestication of Pearl Millet. Econ Bot 31:163–174

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Am J Hum Gent 19:233–257

Clement JC (1985) Les mils pénicillaires de l'Afrique de l'Ouest – Prospections et collectes. IBPGR, Rome; ORSTOM, Abidjan, 231 pp

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338

Harlan JR (1971) Agricultural origins: centers and non centers. Science 174:468–474

Leblanc JM, Pernès J (1983) Enzyme polymorphism of Pennisetum americanum in the Ivory Coast. Jpn J Genet 58: 121–131

Marchais L (1982) La diversité phénotypique des mils pénicillaires cultivés au Sénégal et au Mali. Agron Trop (Paris) 37:68–80

Marchais L, Pernès J (1985) Genetic divergence between wild and cultivated Pearl Millets (*Pennisetum typhoïdes*). I. Male sterility. Z Planzenzücht 95:103–112

Nei M (1975) Molecular population genetics and evolution. North Holland, Amsterdam Oxford, 288 pp

Nei M, Roychoudhury AK (1973) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390

Portères R (1962) Berceaux agricoles primaires sur le continent africain. J Afr Hist 3:195–210

Sandmeier M, Béninga M, Pernès J (1983) Analyses des relations entre formes spontanée et cultivée chez le mil à chandelles. III. Etude de l'hérédité des estérases et des peroxydases anodiques. Agronomie 6:87–494

Second G (1982) Origin of the genic diversity of cultivated rice (*Oryza* spp.): study of the polymorphism scored at 40 isozyme loci. Jpn J Genet 57:25–57

Snedecor GW (1985) Statistical methods, 5th edn. Iowa State University Press, Ames, Iowa, p 534

Tostain S, Riandey MF (1984) Polymorphisme et déterminisme génétique des enzymes de mil (*Pennisetum glaucum* L.). Etude des alcool déshydrogénases, catalases, endopeptidases et estérases. Agron Trop (Paris) 39:335–345

Tostain S, Riandey MF (1985) Polymorphisme et déterminisme génétique des enzymes du mil pénicillaire (*Pennisetum glaucum* L.). Etude des malates déshydrogénases. Agronomie 5:227–235

Tostain S, Lavergne D (1986) Polymorphisme et déterminisme génétique des enzymes de mil (*Pennisetum glaucum* L.). Etude des glutamate oxaloacétate transaminases. CR Acad Sci, Ser III 302:181–184

Tostain S (1985) Mise en évidence d'une liaison génétique entre un gène de nanisme et des marqueurs enzymatiques chez le mil pénicillaire (*Pennisetum glaucum* L.). Can J Genet Cytol 27:751–758