

## Sommaire

	Pages
INTRODUCTION.....	1
SYNTHESE BIBLIOGRAPHIQUE .....	3
1    Taxonomie et botanique du mil .....	4
1.1    L'espèce <i>Pennisetum glaucum</i> .....	6
2    Origine et histoire évolutive de <i>Pennisetum glaucum</i> .....	8
3    Etudes de diversité de <i>Pennisetum glaucum</i> .....	10
3.1    Les études phénotypiques.....	10
3.2    Les études génétiques.....	10
4    Les différents types de marqueurs .....	11
4.1    Les marqueurs morphologiques .....	11
4.2    Les marqueurs biochimiques/enzymatiques .....	11
4.3    Les marqueurs moléculaires .....	11
MATERIEL ET METHODES.....	13
1    Matériel végétal .....	14
1.1    Extraction et quantification de l'ADN .....	14
1.2    Amplification de l'ADN et génotypage .....	16
2    Méthode .....	18
2.1    Détermination de la qualité des données .....	18
2.2    Vérification de redondances de génotypes .....	18
2.3    Etude de la diversité génétique .....	18
RESULTATS.....	20
1    Qualité des données .....	21
2    Vérification de redondances de génotypes.....	23
3    Niveau de diversité génétique .....	25
3.1    Analyse au niveau type Accession .....	25
3.2    Analyse au niveau régional des accessions locales .....	25
4    Structure de la diversité génétique .....	27
4.1    Analyse en composantes principales (ACP) .....	27
4.2    Approche Bayésienne.....	30
4.3    Relations phylogénétiques au sein de la collection .....	33
DISCUSSION.....	36
CONCLUSIONS ET PERSPECTIVES .....	39

REFERENCES BIBLIOGRAPHIQUES.....	41
ANNEXES.....	45

## Liste des acronymes

**FAO**: Food and Agriculture Organization

**ICRISAT**: International Crops Research Institute for the Semi-Arid Tropics

**IRD** : Institut de Recherche pour le Développement

**ISRA** : Institut Sénégalais pour la Recherche Agricole

**LNRPV** : Laboratoire National de Recherche sur les Productions Végétales

**UCAD** : Université Cheikh Anta Diop

## Liste des abréviations

**°C** : Degré Celsius

**ABI** : Applied Biosystems

**ACP** : Analyse en composantes principales

**ADN** : Acide désoxyribonucléique

**AFC** : Analyse factorielle des correspondances

**AFLP** : Amplified fragment-lengthpolymorphism (Polymorphisme de longueur des fragments amplifiés)

**ARN** : Acide ribonucléique

**BET** : Bromure d'éthidium

**CO<sub>2</sub>** : Dioxyde de carbone

**EBD** : Electrophorèse bidimensionnelle

**EDTA** : Acide éthylène diamine tétraacétique

**g** : Gramme

**ha** : Hectare

**ISSR** : Inter Simple SequenceRepeats

**kb** : Kilo base

**L** : Litre

**MCMC** : Markov Chaines Monte Carlo

**mM** : Millimolaire

**PCR** : Polymerasechainreaction

**pb** : Paire de bases

**RAPD** : RandomamplifiedpolymorphicDNA (Amplification aléatoire d'ADN polymorphe)

**RFLP** : Restriction fragment LengthPolymorphism (Polymorphisme de longueur des fragments de restriction)

**SSR** : Simple Sequence Repeat

**STR** : Short Tandem Repeats

**VNTR** : Variable Number Tandem Repeats

**µL** : Microlitre

**µM** : Micromolaire

## INTRODUCTION

La population mondiale ne cesse de croître de façon significative depuis la fin de la dernière guerre mondiale mettant à l'épreuve la sécurité alimentaire dans de nombreuses régions à travers le monde. Cela se manifeste par une demande accrue et une intensification de l'usage des ressources naturelles pour augmenter la production et satisfaire les besoins alimentaires (Rosenzweig and Parry, 1994). Mais il devient de plus en plus difficile d'augmenter cette production agricole à cause des changements climatiques qui peuvent occasionner la sécheresse, des inondations et l'apparition de nouveaux ravageurs mettant ainsi en péril les espèces ou variétés végétales les plus vulnérables. Cela peut conduire à long terme, à la perte de ces variétés, entraînant alors une érosion génétique, compromettant la sécurité alimentaire dans les zones où ces variétés constituent la base de l'alimentation. Il serait donc nécessaire par des programmes de sélection variétale, de produire des variétés adaptées aux contraintes environnementales locales, notamment pour les céréales qui sont la base de l'alimentation mondiale.

Les céréales présentent beaucoup d'importance dans le monde tant sur le plan économique que nutritionnel. Elles sont utilisées pour l'autoconsommation, la transformation artisanale, la transformation industrielle et la commercialisation. En Afrique de l'Ouest, elles occupent une place prépondérante dans l'alimentation des ménages où environ le tiers de la population a un régime alimentaire centré sur la consommation des céréales sèches locales ou introduites (Blein and Soulé, 2011). La production céréalière de l'Afrique de l'Ouest, multipliée par trois ces trente dernières années, a connu une croissance très forte de l'ordre de 4.6% par an et les surfaces cultivées en céréales sont ainsi passées de 19 à 44 millions d'ha sur la période 1980-2008. Elle porte actuellement sur un volume situé entre 52 et 56 millions de tonnes pour une population d'environ 300 millions d'habitants (Blein and Soulé, 2011). Parmi ces céréales, le mil occupe une place phare en Afrique. En effet elle constitue de loin la première céréale produite en Afrique de l'ouest et occupe le premier poste des transactions ; les flux d'échanges intra régionaux de produits locaux sont principalement dominés par le mil (Soule and Gansari, 2010). L'aire que couvrent ses bassins de production est plus grande que celle de toutes les autres céréales en Afrique de l'ouest. Sa production est l'apanage des pays sahéliens et sa farine utilisée dans de nombreuses préparations comme par exemple le pain le couscous ou des tartes, le mil constitue l'alimentation quotidienne de cinquante (50) millions d'habitants du Sahel (Saïdou et al., 2009). Au Sénégal il occupe une place capitale dans la filière céréalière. En raison de la crise de l'arachide au Sénégal, le mil est de plus en plus cultivé dans un but commercial et pas seulement pour l'autoconsommation (Comité national de concertation initiative mil/sorgho ministère de l'agriculture et de l'élevage Novembre 2001). Il représente 80% de la production céréalière avec une production en moyenne pour ces cinq dernières années de plus de 600 000 tonnes et de 813 295 tonnes en 2011 (Bulletin Mensuel des statistiques économiques d'Août 2012 de l'agence nationale de la statistique et de la démographie). Il constitue avec le Sorgho la première culture pratiquée par les ménages ruraux agricoles du Sénégal avec 909 % des ménages (Comité national de concertation initiative mil/sorgho ministère de l'agriculture et de l'élevage Novembre 2001).

Cependant ces céréales sont sous la menace des effets négatifs des changements climatiques. Les points majeurs sur lesquels s'accordent la quasi-totalité des modèles de simulation des changements globaux sont, le réchauffement de la terre et l'accroissement des risques de sécheresse, qui auraient des effets négatifs sur la production agricole. Ces effets seraient particulièrement prononcés en Afrique sub-saharienne où les baisses des rendements de certaines céréales comme le riz, le blé et le maïs seraient respectivement de 15 de 34 et de 10 % (Nelson et al., 2009). Par ailleurs, même si la production des céréales a connu une hausse ces trente dernières années en Afrique de l'Ouest, les rendements moyens sont restés faibles et ont peu progressé, passant de 0.9 à 1.2 tonnes par ha sur la même période (Blein and Soulé, 2011).

Un des moyens pour faire face aux effets des changements climatiques sur l'agriculture est, la production de variétés adaptées aux contraintes environnementales locales. Cela sera bénéfique à l'agriculture et à la sécurité alimentaire. Dans le but de produire des variétés qui répondent à ces changements, le sélectionneur s'appuie sur les ressources génétiques, d'où l'importance d'avoir au préalable une bonne caractérisation de ces ressources. La diversité à tous les niveaux (intraspécifique/interspécifique) représente un capital adaptatif pour les individus. Ainsi plus celle-ci est grande plus les individus sont potentiellement apte à s'adapter à des changements de conditions abiotiques (Frankel, 1970). La biodiversité peut donc être considérée comme un réservoir d'espèces susceptibles de s'adapter à certaines contraintes environnementales par les programmes de sélection variétale. Et donc la gestion durable de ces ressources génétiques, leur conservation et leur valorisation est un moyen efficace d'amoindrir les problèmes que rencontre l'agriculture d'aujourd'hui tout en faisant face à ses défis de demain. Ainsi l'objectif global de cette étude est de déterminer la diversité génétique d'accessions de mils pénicillaires du Sénégal. Les objectifs spécifiques sont : (i) Examiner la structure de cette diversité génétique à l'aide de marqueurs co-dominants et polymorphes de type microsatellites ; (ii) Identifier les (sous)groupes de mils. À terme, l'atteinte de ces objectifs permettra d'utiliser les nombreuses accessions de mil pour les programmes d'amélioration et de sélection variétale. Dans ce document nous ferons une présentation du mil et passerons en revue, les études sur sa diversité génétique. Nous présenterons également le matériel végétal et les méthodes utilisés au cours notre étude. Les résultats obtenus seront présentés, discutés et des conclusions et les perspectives de ce travail seront dégagées.

## **SYNTHESE BIBLIOGRAPHIQUE**

## SYNTHESE BIBLIOGRAPHIQUE

### 1.Taxonomie et botanique du mil

Le terme « mil » regroupe plusieurs genres de plantes appartenant à la famille des *Poaceae*, anciennement connues sous le nom de Graminées, de l'ordre des *Cyperales*, de la classe des *Liliopsida*, de la division des *Magnoliophyta*, du règne des *Plantae*. Dans cette famille nous pouvons citer les genres *Eleusine*, *Pennisetum*, *Setaria*, *Panicum* (Fig.1).

Le genre *Pennisetum* communément appelé mil pénicillaire ou mil à chandelle est caractérisé par un épis compact d'une longueur variant de 10 à 150 cm. C'est le mil le plus cultivé à travers le monde et peut pousser dans des sols sableux et pauvres. Très adapté aux conditions de sécheresse, son potentiel de rendement est le plus élevé de tous les mils cultivés en température élevée ou en condition de sécheresse (Bezançon et al., 1997). Ceci est dû au fait qu'il a la capacité de mettre en place des mécanismes physiologiques favorisant la tolérance à la sécheresse comme le ralentissement des pertes en eau au niveau des feuilles supérieures, via une régulation de l'évapotranspiration, assurant le maintien d'un niveau hydrique favorable (Winkel et al., 1997). Il est constitué par une soixantaine d'espèces qui sont distribuées dans les régions tropicales et subtropicales (Bezançon et al., 1997) dont *Pennisetum glaucum*.

Rapport GRATUIT



*Pennisetum glaucum*



*Eleusine coracana*



*Panicum miliaceum*



*Setaria italica*



*Echinochloa crusgalli*

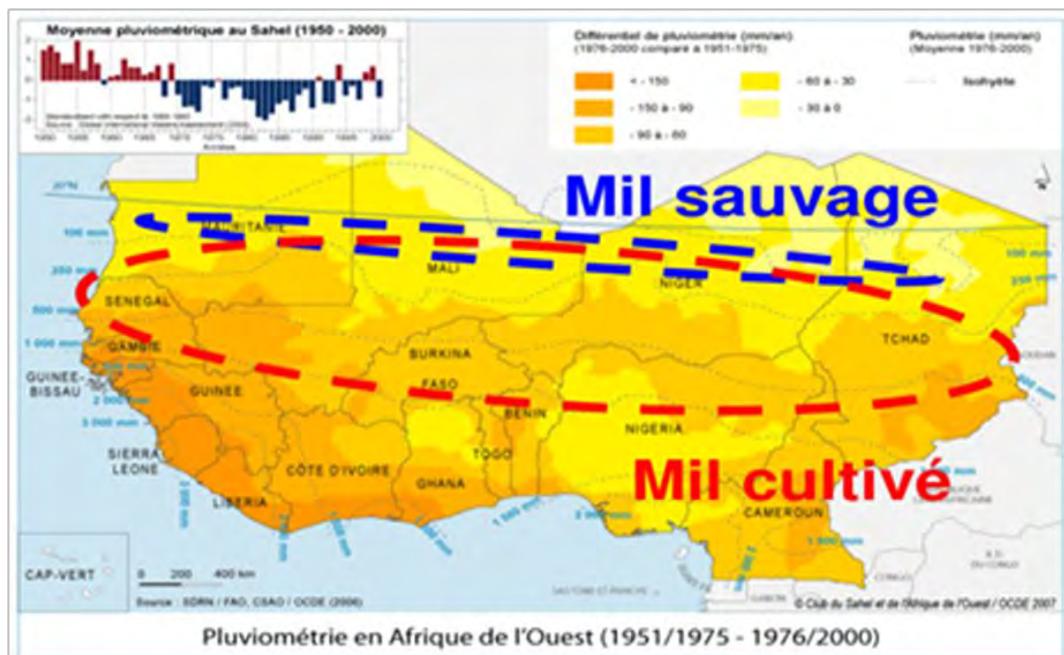
**Figure 1** : Différentes espèces de Poacées regroupées sous le terme mil (FAO et ICRISAT, 1997)

## 1.1.L'espèce *Pennisetum glaucum*

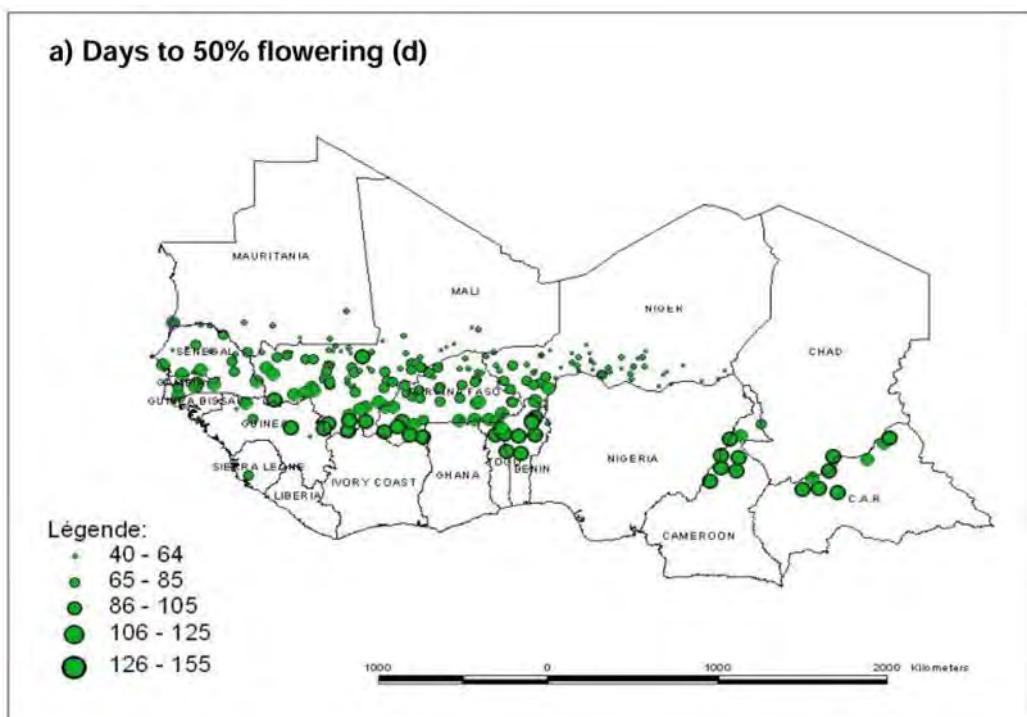
*Pennisetum glaucum* est une plante annuelle, diploïde ( $2n = 14$ ), sexuée, à fleurs hermaphrodites, à reproduction allogame favorisée par un phénomène de protogynie prononcée et à pollinisation anémophile. *Pennisetum glaucum* est constituée de trois sous-espèces :

- *Pennisetum glaucum*spp.*monodi* qui est le type sauvage, poussant à l'état naturel en dehors des champs. Elle n'existe que dans le Sahel, au sud du Sahara.
- *Pennisetum glaucum*spp.*glaucum* l'espèce cultivée. Elle se différencie du type sauvage notamment par la non-caducité des graines à maturité (Tostain, 1998), l'augmentation de la taille des graines, la perte de la dormance, la diminution de la longueur des soies de l'épi (Poncet et al., 1998) et six gènes sous sélection, intervenant dans la date de floraison (Cloutault et al., 2012).
- *Pennisetum glaucum*spp.*stenostachyum* qui regroupe l'ensemble des formes intermédiaires issues d'hybridations naturelles entre formes cultivées et formes sauvages (Marchais et al., 1993). Le nom des sous-espèces peut aussi changer en fonction des auteurs, l'espèce sauvage est aussi appelée *Pennisetum glaucum*spp.*violaceum* et la forme intermédiaire *Pennisetum glaucum*spp.*sieberianum* (Zon, 1992).

La sous-espèce cultivée peut être subdivisée en deux grandes variétés, discriminées par la longueur de leur cycle de culture, une variété précoce et l'autre tardive. Ces variétés sont réparties suivant un gradient nord-sud en fonction de la pluviosité (Fig.2 et 3). La variété tardive appelée maiwa ou Somno (au Niger et au Nigeria) et Sanio (au Sénégal et au Mali) est le plus souvent rencontrée dans les zones où les pluies sont les plus abondantes (Bezançon et al., 1997). Elle a un cycle de 110 à 150 jours tandis que la variété précoce désignée sous le nom de Guero (au Niger et au Nigeria) et Souna (au Sénégal et au Mali), a un cycle de 75 à 100 jours. Elle est cultivée dans les zones plus arides moins pluvieuses.



**Figure 2:** Carte représentant la répartition géographique des mils cultivés et sauvages en fonction de la pluviométrie mesurée au Sahel (J. CLOTAULT et al., 2011)

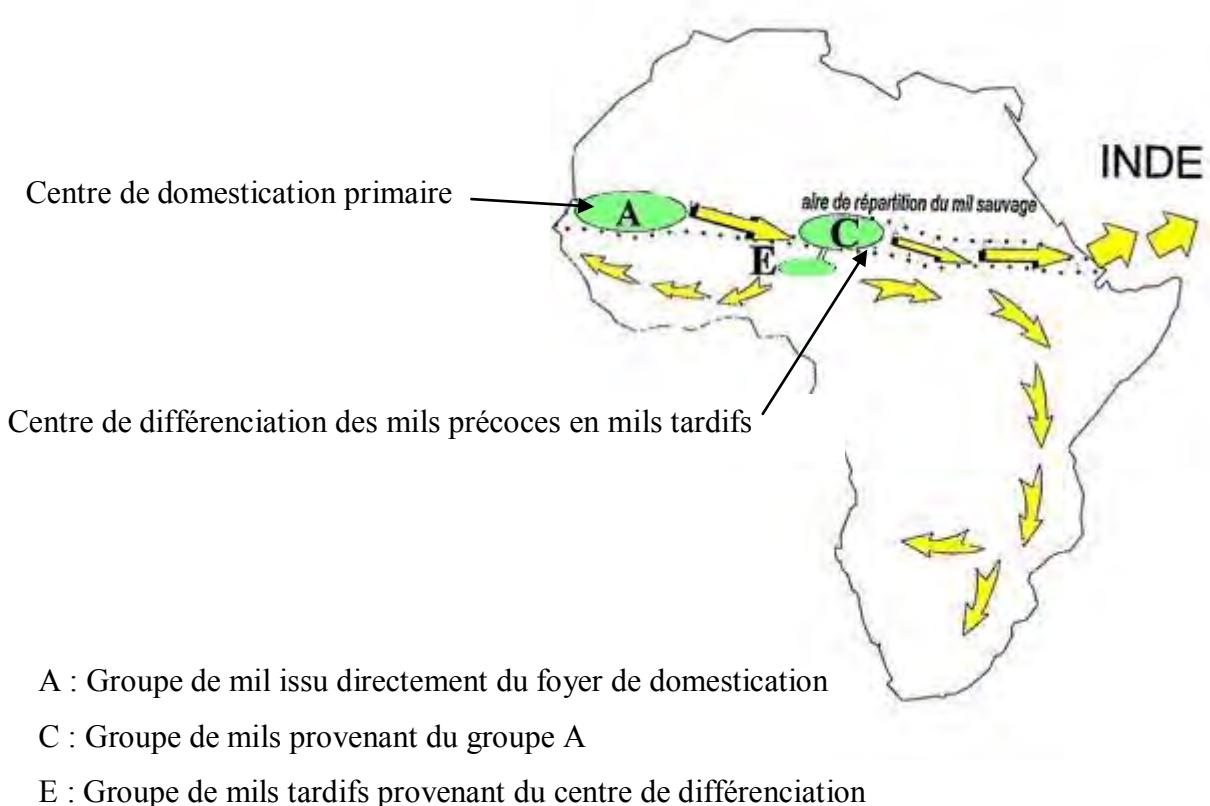


**Figure 3:** Répartition des variétés de mils cultivés (précoces et tardifs) en fonction de leurs dates de floraison (Haussmann et al 2006)

## 2. Origine et histoire évolutive de *Pennisetum glaucum*

*Pennisetum glaucum* a été domestiquée il y a plus de 4800 années (Clotault et al., 2012; Manning et al., 2011) en Afrique de l'ouest (Haussmann et al., 2006; Oumar et al., 2008) dans le Sahel Sud-Saharien, dans la zone constituée par le nord du fleuve Sénégal et le nord-ouest du Mali (Fig. 4) (Tostain, 1998). Cette hypothèse se base sur le fait que dans cette zone, on a la plus faible distance génétique entre mils sauvages et cultivés (Tostain, 1992) et que la distribution géographique du type sauvage de mil est limitée à cette zone (Bezançon et al., 1997), mais aussi sur des données archéologiques par datation(Manning et al., 2011).

C'est à partir de ce foyer de domestication (« A ») que *Pennisetum glaucum* a évolué (Fig. 4) pour engendrer un groupe de mil (« C ») qui va migrer dans trois grandes directions : Une migration ancienne vers l'Afrique de l'est et l'Inde du groupe (C), une migration de mil tardif E provenant du foyer secondaire de différenciation de mils issus du groupe C situé autour du lac Tchad vers la zone soudanienne de l'Afrique de l'ouest ; et enfin une autre migration vers le centre et le sud de l'Afrique (Tostain, 1992).



**Figure 4:** Carte représentant les bassins de domestication de *Pennisetum glaucum* et son histoire évolutive d'après S. Tostain 1998

### **3. Etudes de diversité de *Pennisetum glaucum***

#### **3.1. Les études phénotypiques**

La diversité phénotypique du mil a été étudiée au sein de différentes collections nationales et mondiales. L'étude de la diversité phénotypique des mils pénicillaires cultivés au Sénégal et au Mali a montré que les mils maliens sont plus polymorphes et très différents des mils sénégalais (Marchais, 1982). La caractérisation préliminaire de 281 accessions de mils originaires d'Afrique de l'ouest et du centre (Haussmann et al., 2006), basée sur des caractères phénotypiques comme la date de floraison, la longueur des panicules, la hauteur de la plante, le poids et la couleur des graines, a montré une large diversité pour ces caractères considérés et indique une corrélation entre date de floraison et origine géographique. La description agro-morphologique de 169 variétés de mils d'Inde (Yadav O. P., 2007) dont une partie provenant de l'ouest de Rajasthan et une autre de l'est et du centre de Rajasthan montre une variabilité de la diversité en fonction des indicateurs de diversité considérés. Par exemple, la forme et la compacité des panicules sont relativement moins variables que d'autres descripteurs considérés. En revanche, l'étude de la diversité phénotypique de 145 lignées issues de 122 variétés de mil provenant de l'ouest et du centre de l'Afrique (Stich et al., 2010) définit 5 sous-groupes et montre que certains caractères phénotypiques sont associés de manière significative à la structure génétique des populations. Ces études ont montré une diversité phénotypique relativement importante du mil, surtout à l'échelle régionale. Mais il est à noter qu'elles seules ne sont pas suffisantes pour caractériser les ressources génétiques, il est nécessaire de les compléter avec des études génétiques.

#### **3.2. Les études génétiques**

L'étude de la diversité du mil, au sein de la collection mondiale, au moyen de marqueurs enzymatiques (Tostain, 1992) a mis en évidence plusieurs groupes génétiquement distincts ; deux grands groupes de cultivars précoces, un groupe de cultivars tardifs en Afrique de l'ouest, un groupe en Afrique de l'est rassemblant des cultivars précoces et tardifs et un dernier groupe en Inde avec des cultivars précoces et tardifs. L'évaluation de la diversité des mils sauvages et cultivés au Niger, montre que les mils cultivés présentent moins de diversité génétique que les mils sauvages (Mariac et al., 2006). L'étude de la diversité génétique de 72 accessions de mils originaires du Sud du-Sahara et de l'Inde par le moyen de 34 marqueurs microsatellites (Kapila et al., 2008) a révélé une diversité importante et les classe en cinq groupes, en concordance avec leur pedigree. L'analyse de la diversité génétique des variétés de mil provenant du Zimbabwe à l'aide de marqueurs microsatellites (Chakauya and Tongoona, 2008) a révélé une opposition avec la caractérisation agro-morphologique de cette même collection. L'étude de la diversité génétique du mil a été réalisée à l'aide de marqueurs RAPD (Govindaraj et al., 2009). Ce qui a permis de déterminer la relation génétique entre 20 accessions de mil et de les classer sous huit groupes distincts. Par contre l'analyse de la diversité génétique de 24 lignées de mil par des marqueurs DArT a permis de décrire efficacement les relations génétiques entre un ensemble de lignées pures de mils (Supriya, 2010) et de les classer en deux groupes. L'étude de l'évolution des diversités phénotypiques et génétiques des mils cultivés au Niger de 1976 à 2003 (Chantereau et al., 2010) a montré une absence d'érosion génétique au cours du temps. La caractérisation génétique par des marqueurs microsatellites et enzymatiques de six variétés améliorées de mil au Brésil (Mendonça Neto et al., 2013) a permis d'observer que la plupart des semences commercialisées au Brésil appartiennent à d'autres cultivars. Toutes ces études permettent de mettre en évidence une importante diversité du mil qui est variable en fonction des zones d'étude. Mais permettent aussi de remarquer que la diversité génétique du mil n'a pas encore été étudiée dans certains pays comme le Sénégal.

## 4. Les différents types de marqueurs

Les marqueurs moléculaires sont des outils permettant de caractériser la diversité génétique à l'échelle individuelle. Un marqueur est un caractère mesurable qui permet de détecter une variabilité. Les attributs du marqueur idéal pour étudier la diversité sont : polymorphe, co-dominant, multiallelique, non épistasique (absence d'interaction avec d'autres locus et non exprimé) (de Vienne, 1998). Il existe principalement trois types de marqueurs : les marqueurs phénotypiques, les marqueurs biochimiques et les marqueurs moléculaires.

### 4.1. Les marqueurs morphologiques

Les marqueurs morphologiques sont des caractères phénotypiques (apparents) comme la couleur, la taille et la forme. Ils présentent beaucoup d'avantages et sont indispensables à la caractérisation du vivant. Cependant, ils ont aussi des caractéristiques qui répondent mal à celles d'un marqueur idéal pour l'étude de diversité car ils sont souvent dominants, influençables par le milieu, peu polymorphes et ils peuvent parfois interférer avec les conditions environnementales et engendrer de fausses appréciations.

### 4.2. Les marqueurs biochimiques/enzymatiques

Les marqueurs biochimiques sont des protéines, en général des isozymes (protéines différentes qui ont la même fonction) qui permettent d'étudier la diversité. Leur révélation est réalisée par électrophorèse (migration sur gel) ou EBD (Electrophorèse bidimensionnelle), mais ont comme inconvénients de présenter une certaine spécificité d'organe, du stade de développement ou d'état physiologique car toutes les protéines ne sont pas présentes ou actives dans tous les organes tout le temps et parfois un faible nombre de locus peuvent être révélés.

### 4.3. Les marqueurs moléculaires

Les marqueurs moléculaires sont des séquences nucléotidiques qui permettent de détecter des polymorphismes de l'ADN et sont donc qualifiés de mesures directes de la diversité. Il existe plusieurs types de marqueurs moléculaires, classés selon différents critères. Il existe une classification de ces marqueurs basée sur le type de polymorphisme détecté (polymorphisme de séquence, polymorphisme du nombre d'unité de répétition, polymorphisme d'insertion-délétion). Une autre classification dite génétique a été définie comme reconnaissant les marqueurs co-dominants révélés individuellement et les marqueurs dominants révélés en masse. Parmi les marqueurs moléculaires il y a :

Les marqueurs RFLP (Restriction fragment LengthPolymorphism). Le polymorphisme est révélé grâce à des enzymes de restriction qui coupent au niveau de leurs sites de restriction. Ils sont classés dans la catégorie des marqueurs co-dominants révélés individuellement et dans les marqueurs à polymorphisme de séquence.

Les marqueurs AFLP (amplified fragment-lengthpolymorphism). Il s'agit d'utiliser deux enzymes de restriction l'un coupant plus fréquemment que l'autre et des adaptateurs pour l'amplification. Le polymorphisme est un polymorphisme de longueurs de fragments amplifiés. Ce sont des marqueurs dominants révélés en masse et des marqueurs à polymorphisme de séquence.

Les marqueurs ISSR (Inter Simple SequenceRepeats). Consiste à utiliser des microsatellites comme amorces et d'amplifier les séquences entre eux. Il s'agit d'un polymorphisme de longueur de fragments amplifiés. Ce sont des marqueurs dominants révélés en masse et des marqueurs à polymorphisme d'unités de répétition.

Les marqueurs microsatellites ou SSR (Simple SequenceRepeat), sont des séquences répétées en tandem de 1 à 6 nucléotides. Ils sont révélés après leur amplification grâce à des amores spécifiques aux régions flanquantes. Ainsi le polymorphisme est un polymorphisme d'unités de répétition. Ce sont des marqueurs co-dominants, révélés individuellement. Ils sont répartis sur l'ensemble du génome (Selkoe and Toonen, 2006). Ce sont les marqueurs les plus fréquemment utilisés pour l'étude de diversité génétique car ils sont reproductibles, co-dominants, très polymorphes, bien repartis dans le génome. Tous ces caractères font d'eux, un excellent outil pour des études de diversité génétique.

## **MATERIEL ET METHODES**

## MATERIEL ET METHODES

### 1. Matériel végétal

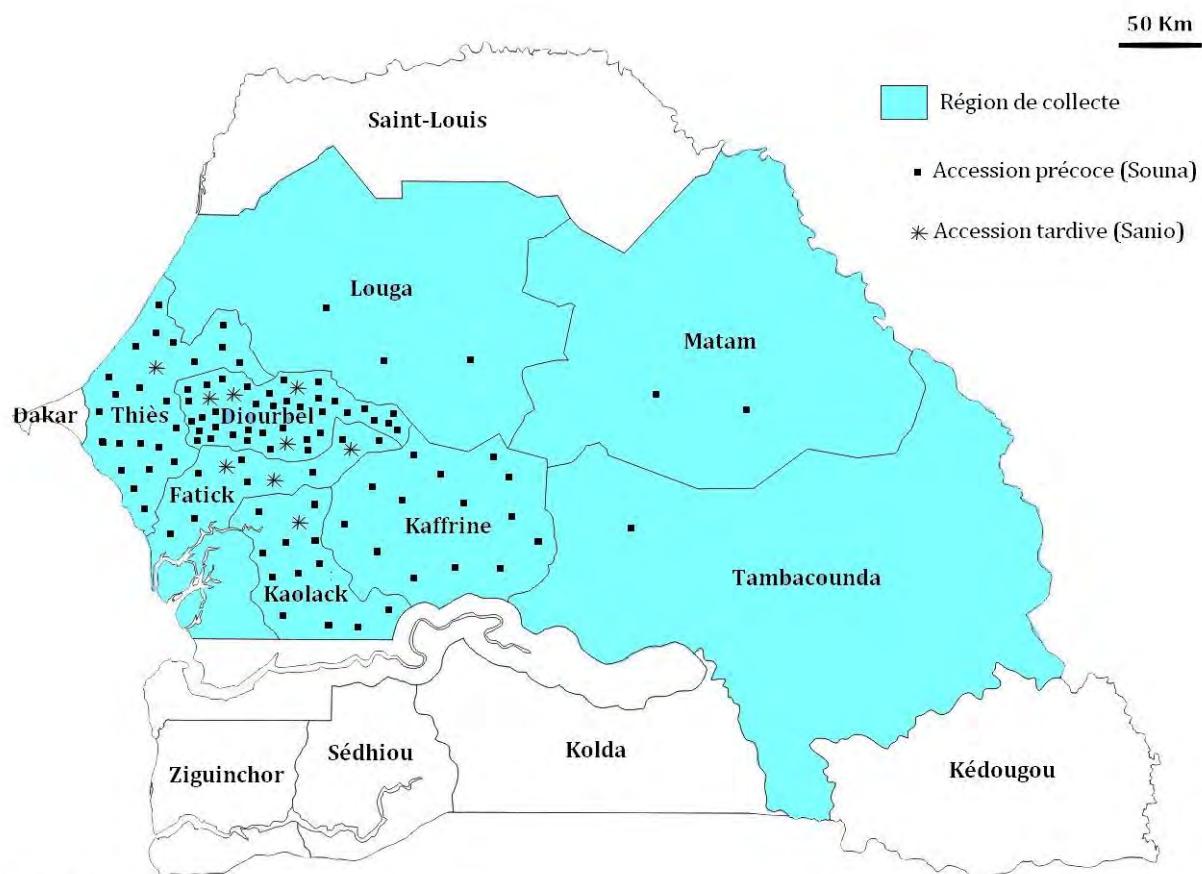
Le matériel végétal est constitué de 252 accessions de mils pénicillaires dont 10 de type Sanio et 229 Souna, collectées entre 1992, 1993 et 2010 principalement dans le bassin Arachidier du Sénégal, dans les régions de Thiès, Kaolack, Fatick, Diourbel, Tambacounda, Kaffrine et Louga (Fig.5) et 13 variétés introduites. Ces graines ont été conservées dans la collection de l'ISRA. En moyenne quatre graines par accession ont été semées dans des gaines (une gaine par accession) contenant du sol sableux léger dépourvu de nématode. Elles ont été placées en serre et quotidiennement arrosées avec l'eau de robinet pendant 15 jours, jusqu'à obtention de jeunes plantes.

#### 1.1. Extraction et quantification de l'ADN

Sur chaque jeune plante, un prélèvement foliaire d'environ 200 mg a été réalisé, afin d'en extraire l'ADN, sur la base de la méthode décrite par Doyle *et al* (1987) (Annexe 1). Les feuilles ont été broyées au pilon dans des mortiers contenant 750 µL d'une solution tampon de lyse à CTAB 2%. Les broyats ont ensuite été incubés à 74°C pendant 20 min. Après 2 min de refroidissement à la température ambiante, 750 µL d'une solution de chloroforme iso-amylique 24:1 ont été ajoutés dans chaque tube puis homogénéisés, centrifugés à 10.000 rpm à 20°C pendant 2 min. Le surnageant a été récupéré puis, ont été rajoutés, 500 µL d'isopropanol et 50 µL d'acétate de sodium 3 M à pH 5 et incubé à -20°C pendant 15 min. Les tubes ont été centrifugés à 10.000 rpm à 10°C pendant 10 min. Le culot d'ADN a été lavé avec 500 µL d'éthanol à 70° puis centrifugé à nouveau à 10.000 rpm à 10°C pendant 10 min. L'ADN précipité a été séché à la température ambiante pendant 3 h et dissout dans 100 µL d'une solution de TE 1X à pH 8 dont la solution mère contient 50mL de Tris 10mM et 250µL d'EDTA 0.1mM.

Après l'extraction, pour vérifier la quantité et la qualité de l'ADN, 15 échantillons ont été faits sur un gel d'agarose à 2% avec 6 µl d'ADN et 2 µl de bleu de bromophénol par échantillon. Le gel a ensuite été coloré au Bromure d'éthidium (BET) pendant 15 min avant de le visualiser aux rayons UV.

Pour vérifier les concentrations et la pureté des solutions d'ADN, de manière plus précise, tous les échantillons ont été passés au spectrophotomètre (un Nano Drop). Les rapports 260/230 et 260/280 indiquent respectivement les degrés de contamination en polysaccharides et en protéines. En effet, 230nm correspond à la longueur d'onde d'absorbance maximale des polyphénols, de l'EDTA et autres sels et 280nm à celle des protéines. Les valeurs des rapports de 260/230 et 260/280, pour une solution d'ADN pure doivent être comprises entre 1.8 et 2 ou 2.2. Les solutions d'ADN ont été enfin diluées à 25 ng/µl puis conservée à -20°C.



Copyright Diack et al 2013

**Figure 5:** Provenance des accessions locales de la collection de mil du Sénégal (Diack *et al* 2013)

## 1.2. Amplification de l'ADN et génotypage

### 1.2.1. Choix des marqueurs

Les 15 microsatellites utilisés (Tableau 1) pour l'étude de la diversité des accessions de mils ont préalablement été utilisés dans plusieurs études de diversité (Allouis et al., 2001; Budak et al., 2003; Mariac et al., 2006; Qi et al., 2001). Ils sont polymorphes et sont bien répartis tout au long du génome du mil.

### 1.2.2. Amplification de l'ADN

L'amplification a été réalisée à l'aide d'un Kit Qiagen contenant une solution Q 5x, un master mix 2x et du H<sub>2</sub>O (RNase-free water) où 1 µl d'ADN (25 ng) a été ajouté pour un volume total de 10 µl pour chaque puits. L'amplification a été réalisée à l'aide d'un thermocycleur TC-Plus (TECHNE) et les réactions ont été réalisées en deux multiplexes (Multiplex A et B) selon le protocole suivant : une pré-dénaturation de 95°C pendant 15 min puis 35 cycles comprenant une dénaturation à 94°C pendant 30 s, une hybridation à 55°C pendant 90 s, une élongation à 72°C pendant 60 s et une extension finale à 60°C pendant 30 min.

Après amplification, afin de vérifier que la réaction s'est bien déroulée et qu'il n'y a pas eu de contamination, deux accessions de la réaction (contrôles positifs) et deux contrôles négatifs ont été testés: A cet effet, 5 µl de produit PCR et 5 µl de bleu de bromophénol ont été mis à migrer par électrophorèse à 100 V pendant 30 min sur gel d'agarose 1%. Le gel a été coloré au bromure d'ethidium et visualisé aux UV.

Tableau 1: Liste des marqueurs microsatellites (SSR) utilisés et de leurs caractéristiques moléculaires

Nom	taille minimum	taille maximum	Dye	Multiplex	Motif	Tm°C
IRD46*	80	110	Ned	A	(CTC) <sub>6</sub>	55
Psmp2227 $\diamond$	180	215	Ned	A	(GT) <sub>7</sub>	55
Psmp2249 $\diamond$	125	160	Fam	A	(GT) <sub>7</sub>	55
Psmp2247 $\diamond$	180	231	Fam	A	(TG) <sub>6</sub>	55
Psmp2201 $\diamond$	310	375	Red	A	(GT) <sub>6</sub>	55
Psmp2237 $\diamond$	210	260	Red	A	(GT) <sub>8</sub>	55
CTM8 $\square$	225	275	Vic	A	(CT) <sub>8</sub>	55
IRD12*	110	145	Vic	A	(TA) <sub>8</sub>	55
IRD25*	165	190	Red	B	(GA) <sub>9</sub>	55
Psmp2214 $\diamond$	230	255	Red	B	(GT) <sub>9</sub>	55
Psmp2219 $\diamond$	210	295	Ned	B	(GT) <sub>7</sub>	55
Psmp2220 $\diamond$	100	145	Ned	B	(GT) <sub>11</sub>	55
Psmp2206 $\diamond$	190	210	Fam	B	(GT) <sub>11</sub>	55
Psmp2202 $\diamond$	130	165	Vic	B	(GT) <sub>8</sub>	55
Psmp2231 $\diamond$	212	262	Vic	B	(TG) <sub>12</sub> GG(TA) <sub>4</sub>	55

\*= (Mariac et al., 2006) $\diamond$ = (Qi et al., 2001) $\square$ = (Budak et al., 2003) $\diamond$ = (Allouis et al., 2001). Tm°C = température d'hybridation

### 1.2.3. Genotypage

Les fragments amplifiés par PCR ont été détectés sur un séquenceur ABI3130 (AppliedBiosystems) plateforme Génopole de Montpellier et les données de séquence ont été récupérées avec le logiciel GeneMapper 3.7 qui a aussi servi pour l'assignation des génotypes. L'assignation des pics correspondant aux allèles a été déterminée avec un standard de taille.

## 2. Méthode

### 2.1. Détermination de la qualité des données

Pour éviter d'éventuelles erreurs et s'assurer de la qualité des données et de la fiabilité des résultats, quatre échantillons Souna et quatreSanio, répétés sur les plaques PCR, ont d'abord été utilisés comme témoins positifs et 12 témoins négatifs ont aussi été utilisés pour vérifier d'éventuelles contaminations et le bon fonctionnement des réactions. Ensuite, la lecture des données sur GenMapper a été réalisée par deux personnes et au moins à deux reprises chacun. Afin d'éviter de biaiser les tests, tous les individus et marqueurs présentant plus de 10% de données manquantes ont été éliminés. Les marqueurs CTM8 et Psmp 2201 ainsi que 50 individus ont été éliminés pour le reste des analyses. Cela a ramené la totalité des individus étudiés à 202 dont 185 Souna 9 Sanio et 8 variétés améliorées et a finalement conduit à des marqueurs et des individus présentant tous 100% de données. La probabilité de présence d'allèles nuls a été testée à l'aide du logiciel Micro-Checker 2.2.3 (Van Oosterhout et al., 2004) puis un test de permutation à l'aide du logiciel Multilocus 1.3 (Agapow and Burt, 2001) a été réalisé. Ce test permet de vérifier si le nombre de marqueurs utilisés est suffisant pour mettre en évidence l'intégralité de la diversité génétique de notre jeu de données. La diversité génétique moyenne pour 1000 permutations pour un nombre n de marqueurs allant de 1 à 7 a été calculée.

### 2.2. Vérification de redondances de génotypes

Une comparaison des génotypes a été réalisée, afin de tester la présence de doublons au sein de la collection, complétée par un calcul des probabilités d'identité grâce au logiciel GenAlex (Peakall and Smouse, 2012). Ces données ont permis de réaliser un graphe expliquant la probabilité que deux accessions aient le même géotype.

### 2.3. Etude de la diversité génétique

#### 2.3.1. Niveau de diversité génétique

Après assignation des génotypes, les données ont été exportées sous forme d'une matrice rendant compte des génotypes des individus.

Un calcul des indices de diversité tels que la richesse allélique (RA), l'hétérozygotie Observée ( $H_o$ ), l'hétérozygotie attendue ( $H_{ex}$ ) a été réalisé, pour chaque locus et dans chaque population considérée grâce au logiciel GenAlex 6.5 (Peakall and Smouse, 2012).

L'hétérozygotie observée et l'hétérozygotie attendue ont permis d'obtenir l'indice de fixations (Fis) ou coefficient de consanguinité qui reflète la différence des individus à l'intérieur des populations. Sa valeur varie de -1 à 1. La valeur d'un Fis égale à 1 indique un excès d'homozygote (Cas d'autofécondations ou d'homogamie ou encore effet Whalund). La valeur d'un Fis égale à -1 traduit un excès d'hétérozygotes, tandis qu'une valeur égale à 0 signifie que la population est à l'équilibre d'Hardy et Weinberg. Tous ces paramètres permettent de pouvoir apprécier la diversité des individus d'une population étudiée. L'analyse des données a été réalisée principalement à deux niveaux distincts. D'abord à un niveau, type accession (avec d'un côté les accessions locales et de l'autre les accessions

améliorées ou introduites), puis au niveau, type régional (lieu de provenance) des accessions locales de la collection (Variété précoce ou Souna et variété tardive ou Sanio).

### **2.3.2. Structure de la diversité génétique**

#### **2.3.2.1. Analyse en composantes principales (ACP)**

Des analyses en composantes principales pour chaque niveau de considération des populations (par unité géographique et par type variétal), a été réalisée à l'aide du logiciel GenAlex 6.5 (Peakall and Smouse, 2012) avec les paramètres par défaut.

#### **2.3.2.2. Approche Bayésienne**

La structure génétique de la collection, a été mise en évidence par une approche bayésienne, suivant le modèle d'admixture (Model selon lequel les différents gènes d'un individu peuvent être hérités de différentes populations), à l'aide du logiciel STRUCTURE 2.3.3 (Falush et al., 2003; Pritchard et al., 2000) avec comme paramètres ; une période de burn de 400000 et 10 runs de 600000 répétitions de tirage et remise par la méthode Chaines de Markov Monte Carlo (MCMC) et avec  $k$  allant de 1 à 12 ( $k$  étant le nombre de groupes simulés). Les meilleures valeurs du logarithme de vraisemblance ( $\ln P(D|K)$ ) pour chaque  $K$  groupe ont été utilisées pour réaliser un graphe illustrant le nombre le plus probable de sous-groupes selon la méthode d'Evanno(Evanno et al., 2005).

#### **2.3.2.3. Relation entre accessions**

Dans le but d'avoir une représentation des individus les plus proches, entre eux au sein de la collection et de voir comment ils sont groupés, des dendrogrammes pour chaque type variétal de la collection et pour toutes les variétés de la collection prises au total, ont été construits par la méthode NeighborJoining(Saitou and Nei, 1987) avec 1000 réplications dans le logiciel Darwin 5.0.158(Perrier et al., 1999). Le but de la représentation par type variétal a été de voir dans chaque variété les individus les plus proches. Et pour la représentation de toutes les variétés prises au total a été de voir les variétés génétiquement les plus proches de la collection.

## **RESULTATS**

## RESULTATS

### 1. Qualité des données

Pour la vérification de la présence d'allèles nuls, le logiciel Micro-Checker a suggéré la possibilité qu'il y ait des allèles nuls pour les marqueurs IRD12, psmp2249,psmp2227, IRD25,psmp2231 et psmp2214 qui ont été éliminés (Tableau 2 et 3).

Pour la vérification de l'efficacité du nombre de marqueurs utilisés, pour étudier la diversité dans nos données (Fig. 6), nous observons que plus le nombre de marqueurs utilisés, est important, plus la quantité de diversité génétique mise en évidence l'est. Sur notre jeu de données, il semble que le nombre de marqueurs minimum permettant de mettre en évidence l'ensemble de la diversité génétique de nos données soit de quatre. Ainsi un nombre de marqueurs supérieur à quatre ne permet pas de mettre en évidence davantage de diversité. Or nous avons utilisé sept marqueurs sur notre jeu de données, ce qui signifie que nous avons utilisé suffisamment de marqueurs, pour mettre en évidence l'ensemble de la diversité génétique caractérisant nos données.

Tableau 2: Vérification de la présence d'allèles nuls avant élimination des locus (Micro-Checker 2.2.3)

Locus	NullPresent	Oosterhout	Chakraborty	Brookfield 1	Brookfield 2
IRD12	yes	0.2855	0.4788	0.2297	0.4884
2249	yes	0.1875	0.3021	0.1189	0.1663
2227	yes	0.0851	0.1055	0.0579	0.4812
IRD46	no	0	0	0	0.9463
2247	no	0.034	0.0363	0.0207	0.1374
2237	no	0.027	0.0408	0.0132	0.4003
IRD25	yes	0.0459	0.0488	0.0257	0.0257
2231	yes	0.1394	0.1652	0.1244	0.8269
2202	no	0	0	0	0.7633
2206	no	0.0217	0.0235	0.0198	0.0198
2220	no	0.0336	0.0413	0.0343	0.404
2219	no	0.0447	0.0508	0.0248	0.3229
2214	yes	0.1068	0.1155	0.0656	0.4303

Plusieurs loci montrent des signes d'allèles nuls.

Tableau 3: Vérification de la présence d'allèles nuls après élimination des locus ayant été détectés par le logiciel comme présentant des allèles nuls (Micro-Checker 2.2.3)

Locus	NullPresent	Oosterhout	Chakraborty	Brookfield 1	Brookfield 2
IRD46	no	0	0	0	0.947
2247	no	0.0252	0.0265	0.015	0.1367
2237	no	0.0259	0.0401	0.0137	0.2712
2202	no	0	0	0	0.7332
2206	no	0.0211	0.0235	0.0197	0.0197
2220	no	0.0464	0.0556	0.0458	0.4361
2219	no	0.0644	0.0718	0.0348	0.3306

Aucun locus ne montre des signes d'allèles nuls.

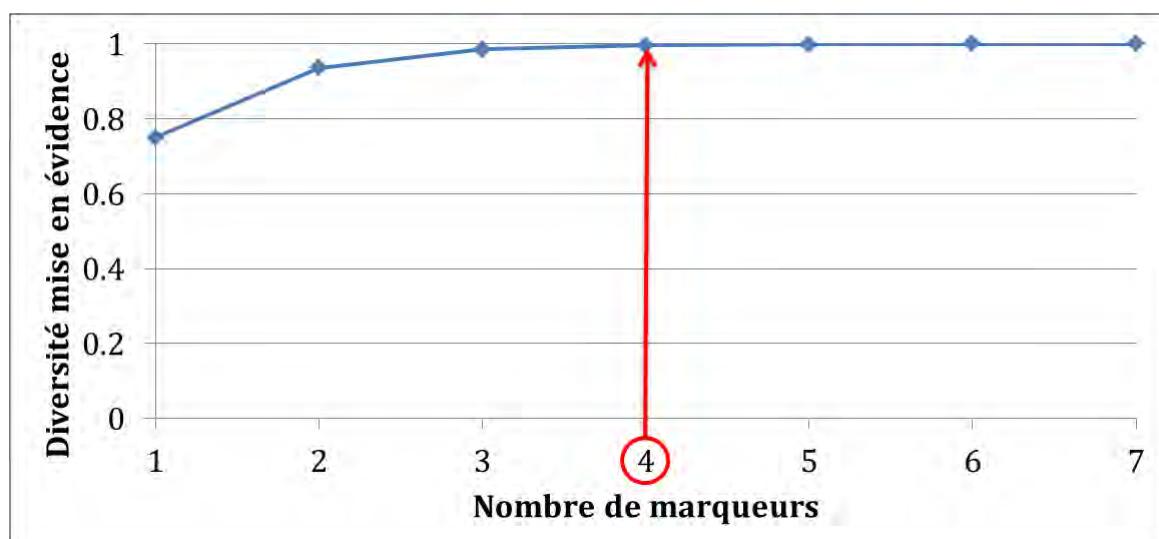
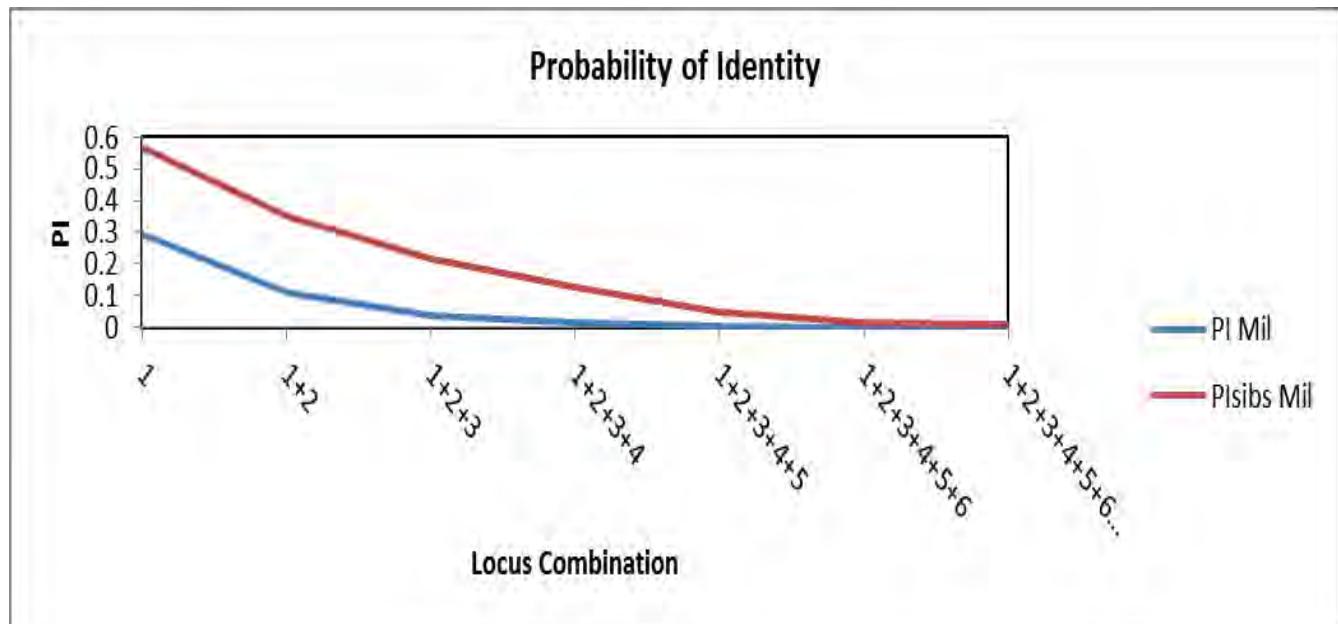


Figure 6 : Courbe Représentant la quantité de diversité génétique mise en évidence dans notre jeu de données en fonction du nombre de marqueur SSR utilisé

## 2. Vérification de redondances de génotypes

Le test de la présence de doublons s'est révélé négatif. Sur la totalité des 202 accessions testées de la collection, il n'a été détecté aucune redondance de génotypes. (Fig. 7). Nous observons que la probabilité d'identité génotypique est nulle lorsque tous les 7 locus marqueurs sont combinés. Ce qui veut dire qu'il n'y a aucun doublon dans la totalité de nos accessions.



**Figure 7 :** Courbes expliquant la probabilité que deux accessions indépendantes (en bleu) et/ou considérés comme ayant des liens de parenté (en rouge) aient le même génotype en fonction des marqueurs cumulés croissants

### **3. Niveau de diversité génétique**

Les résultats obtenus sur les marqueurs utilisés sont :d'abord leur efficacité pour cette étude mais ensuite leur polymorphisme.La richesse allélique des marqueurs varie entre trois et 14 allèles en moyenne.Les marqueurs Psmp2247et Psmp2202ont la plus faible valeur et le marqueur Psmp2220la plus grande(Tableau 4).

#### **3.1. Analyse au niveau type Accession**

La richesse allélique est de cinq et sept allèles par locus en moyenne, pour les accessions introduites et locales respectivement. Et la diversité génétique ( $H_{ex}$ ) est plus élevée chez les accessions introduites (0.601) que chez les accessions locales (0.587). Ceci fait une moyenne de 0.594 pour la diversité des accessions locales et introduites de toute la collection. L'indice de fixation a été obtenu grâce à l'hétérozygotie attendue et observée (Tableau 5). Cet indice révèle que ces deux populations de type variétal ne sont pas à l'équilibre d'Hardy et Weinberg au seuil de 0.05 signifiant un nombre d'homozygote légèrement en excès ce qui veut dire qu'au moins l'une des conditions de base à cet équilibre n'a pas été respectée (Panmixie,Sélection, mutation, migration)Ceci peut aussi être due à l'effet Whalund.

#### **3.2. Analyse au niveau régional des accessions locales**

##### **3.2.1. Accessions tardives (Sanio)**

Toutes les accessions tardives de la collection ne proviennent que de quatre régions (Diourbel, Fatick, Kaolack et Thiès).Certains paramètres pour ces accessions tardives n'ont pas été calculés dans les régions de Kaolack et de Thiès, car elles étaient représentées que par un seul individu. Par contre, la richesse allélique des accessions provenant des régions de Diourbel et Fatick est relativement semblable avec en moyenne deux et trois allèles par locus respectivement et la diversité génétique est plus grande dans la population de Fatick (0.600) que dans celle de Diourbel (0.561). La population de Diourbel a un Fis significatif autrement dit, elle n'est pas à l'équilibre d'Hardy et Weinberg au seuil de 0.05 avec des Fis de 0.265 (Tableau 6).

##### **3.2.2. Accessions précoces (Souna)**

La région de Tambacounda étant représentée par un seul individu, certains de ces paramètres n'ont pas été calculés.La richesse allélique varie en moyenne, entre deux et six allèles par locus,les accessions de Matam en ont la plus faible valeur et celles de Diourbel la plus élevée. La diversité génétique est aussi plus faible pour la population de Matam (0.547). Celle de Kaffrinea la plus grande diversité (0.631). Les individus provenant de toutes les régions, présentent des Fis indiquant l'équilibre d'Hardy et Weinberg sauf pour ceux de Thiès et Diourbelqui indiquent un Fis significatif au seuil de 0.05, pour des Fis de 0.069 et 0.068 respectivement. (Tableau 7).

Tableau 4: Récapitulatif des statistiques descriptives rendant compte de la diversité génétique mise en évidence par locus

Locus	RA	$H_o$	$H_{ex}$	Fis
IRD46	4	0.49505	0.51555	0.03986*
Psmp2247	3	0.38614	0.45035	0.14288
Psmp2237	8	0.31683	0.4334	0.26944
Psmp2202	3	0.5495	0.50572	-0.08680*
Psmp2206	10	0.72277	0.75878	0.04756*
Psmp2220	14	0.79703	0.8648	0.07855
Psmp2219	11	0.54455	0.60573	0.10122

Tableau 5: Récapitulatif des statistiques descriptives rendant compte de la diversité génétique mise en évidence par type d'accession

Accession	N	RA	$H_{ex}$	$H_o$	Fis
Locales	193	7.429 (4.276)	0.58705 (0.16593)	0.55366 (0.17039)	0.05701*
Introduites	9	5.000 (2.530)	0.60131 (0.23149)	0.40741 (0.19458)	0.33585*

Tableau 6: Diversité génétique des variétés tardives (Sanio) par région

Région	N	RA	$H_{ex}$	$H_o$	Fis
Diourbel	4	2.714 (0.756)	0.56122 (0.13318)	0.42857 (0.23780)	0.26531*
Fatick	3	3.000 (1.155)	0.60000 (0.25820)	0.52381 (0.26227)	0.15385
Kaolack	NA	NA	NA	NA	NA
Thiès	NA	NA	NA	NA	NA

Tableau 7: Récapitulatif des statistiques descriptives rendant compte de la diversité génétique mise en évidence par région des variétés précoces (Souna)

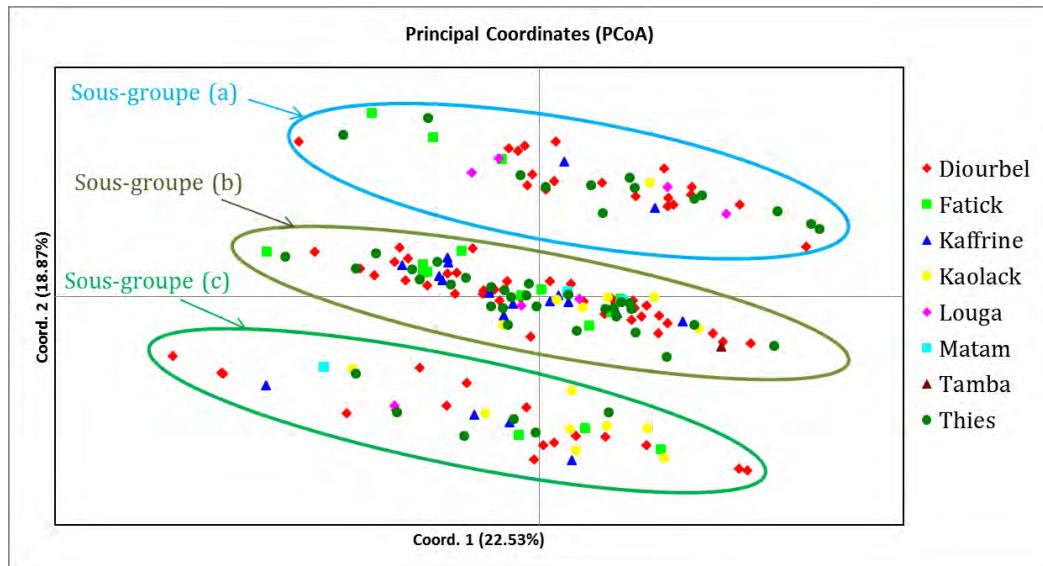
Région	N	RA	$H_{ex}$	$H_o$	Fis
Diourbel	75	6.429 (3.867)	0.58277 (0.17306)	0.54286 (0.15275)	0.06892*
Fatick	16	4.143 (1.952)	0.56740 (0.16957)	0.55357 (0.24321)	0.02516
Kaffrine	19	5.286 (2.812)	0.63178 (0.13834)	0.64662 (0.17904)	-0.02415
Kaolack	15	5.000 (2.517)	0.60690 (0.20826)	0.62857 (0.24300)	-0.03704
Louga	7	3.429 (1.512)	0.54788 (0.14736)	0.46939 (0.21372)	0.15337
Matam	3	2.667 (0.816)	0.58889 (0.22476)	0.61111 (0.25092)	-0.04762
Tambacounda	NA	NA	NA	NA	NA
Thiès	48	5.429 (2.637)	0.56864 (0.18629)	0.52976 (0.20759)	0.06905*

N = Nombre d'individus considérés, RA = Richesse Allélique,  $H_{ex}$  = Hétérozygotie attendue, NA = Valeur non calculée,  $H_o$  = Hétérozygotie observée, Fis = Indice de fixation, \* = Fis significatif à 0.05, Nb : Les valeurs entre parenthèses sont les écarts-types

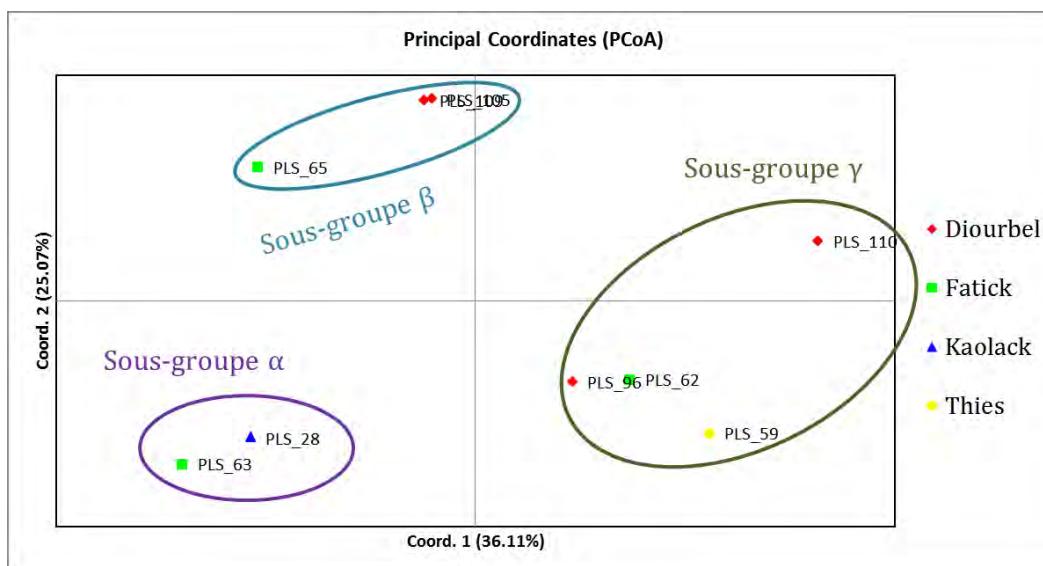
## 4. Structure de la diversité génétique

### 4.1. Analyse en composantes principales (ACP)

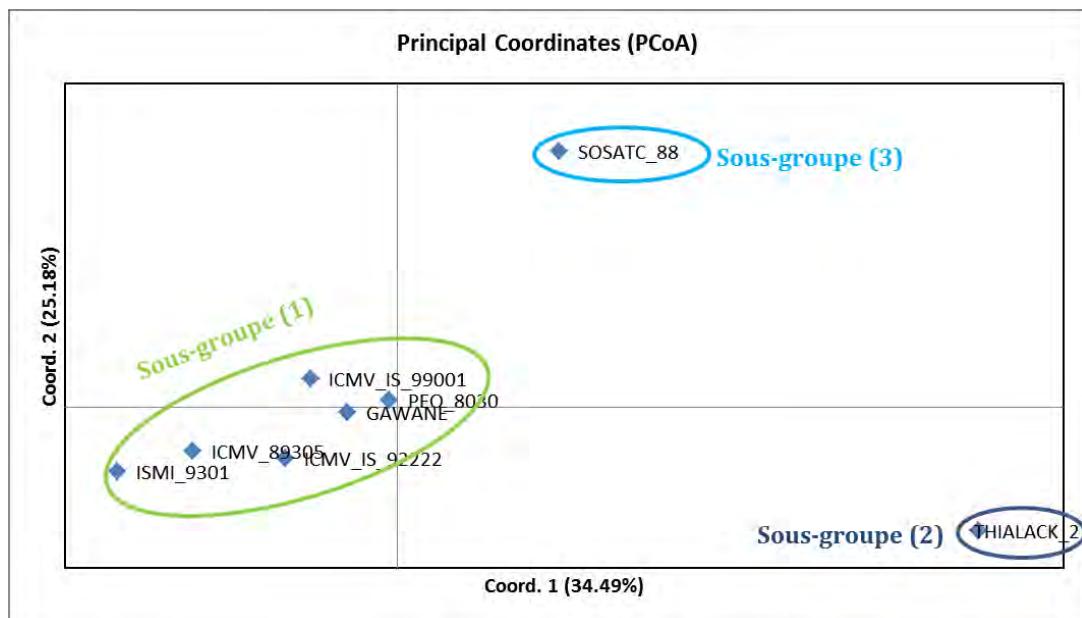
L'ACP réalisée fait apparaître une structure en trois sous-groupes pour chaque type variétal. Les variétés précoces de la collection, en sous-groupe (a), sous-groupe (b) et sous-groupe (c) de même que les variétés tardives en sous-groupes ( $\alpha$ ), ( $\beta$ ) et ( $\gamma$ ) (Fig.8 et 9). Les variétés améliorées ont une structure permettant de voir le sous-groupe (1) renfermant un certain nombre d'individus regroupés entre eux. Les variétés Thialack\_2 et SOSATC\_88 sont isolées du reste et forment chacun leurs propres sous-groupes. Le sous-groupe (2) pour Thialack\_2 et (3) pour SOSATC\_88 (Fig.10).



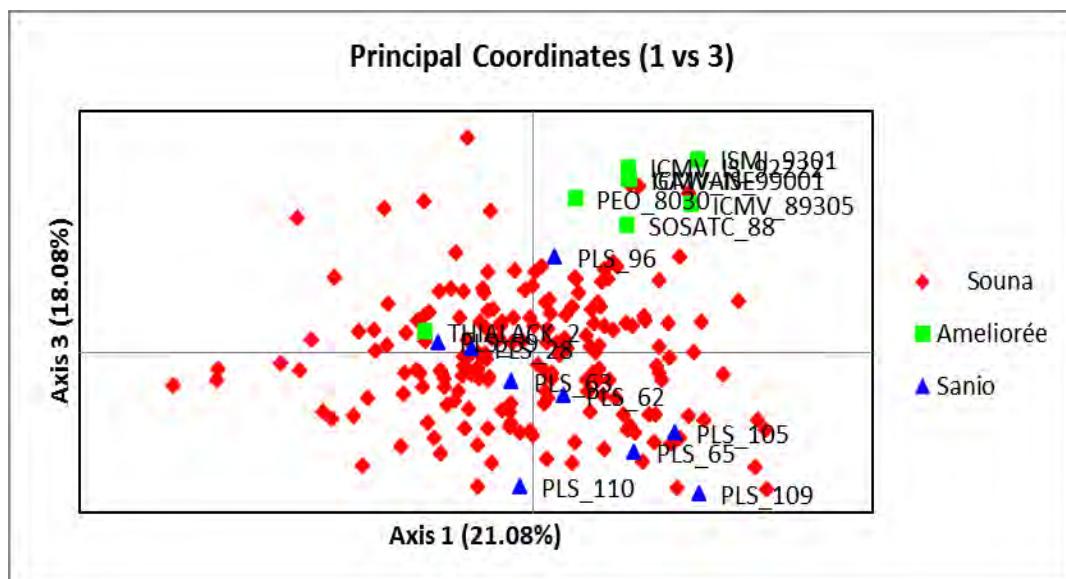
**Figure 8 :** Analyse en composantes principales des variétés précoces



**Figure 9 :** Analyse en composantes principales des variétés tardives (Sanio)



**Figure 10** : Analyse en composantes principales des variétés améliorées



**Figure 11** : Analyse en composante principale de toutes les variétés de la collection

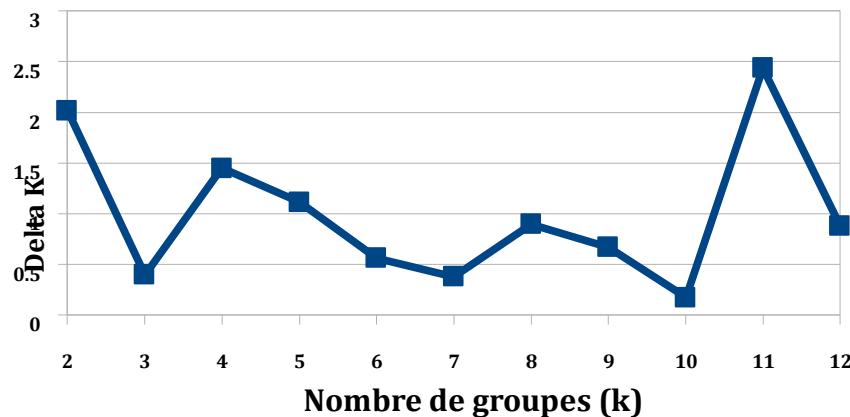
## 4.2. Approche Bayésienne

Après permutations avec le logiciel STRUCTURE, les courbes obtenues grâce à la méthode d'Evanno, pour les variétés tardives et améliorées, suggèrent plusieurs groupes. Et ne permettent pas d'être précis sur le nombre exact de groupes, à cause d'un faible nombre d'individus. En effet plus la taille des échantillons est grande, plus la puissance d'un test est fiable. Rappelons que les variétés améliorées et les variétés tardives sont constituées de huit et neuf individus respectivement, contre 185 pour les variétés précoces. Par ailleurs, la représentation graphique par STRUCTURE de ces variétés, après le test d'assignation, ne met en évidence aucun groupe.

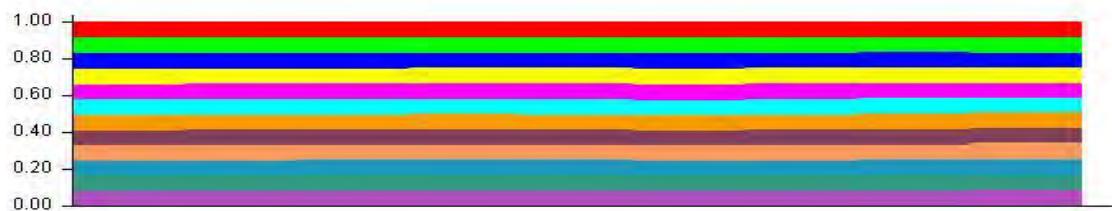
Pour les variétés précoces, la méthode d'Evanno apporte une confirmation des analyses précédentes, avec une répartition en trois sous-groupes. Cependant, il y a une autre probabilité de subdivision en six et en 12 sous-groupes, moins marquée. La représentation graphique par STRUCTURE en  $K=3$ , permet de voir qu'il n'y a aucune appartenance d'une région à un  $K$  groupe donné. Les différents groupes sont tous représentés dans toutes les régions à des proportions relativement égales.

L'analyse des accessions pour toutes les variétés confondues de la collection, indique une subdivision en deux groupes et une autre, en cinq, moins marquée. (Fig. 15-1). La représentation graphique, par région, des résultats du test d'assignation par STRUCTURE pour  $k=2$ , indique un non appartenance d'une région à un groupe.

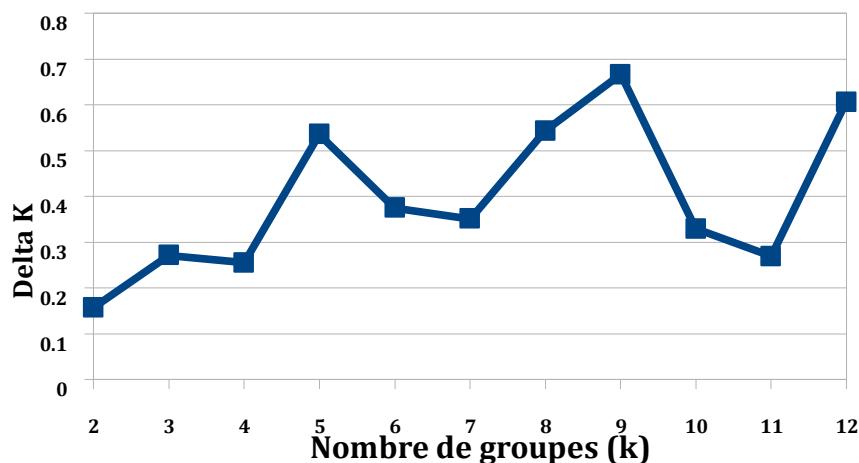
Cependant, la représentation graphique, par type variétal, des résultats du test d'assignation par STRUCTURE pour  $k=2$ , montre clairement une appartenance des variétés améliorées à un groupe (vert). Les variétés tardives appartiennent plutôt à un autre groupe (rouge). Les variétés précoces présentent une structure partagée entre ces deux groupes même si la proportion du groupe des variétés améliorées (vert) est dominante (Fig 15-3).



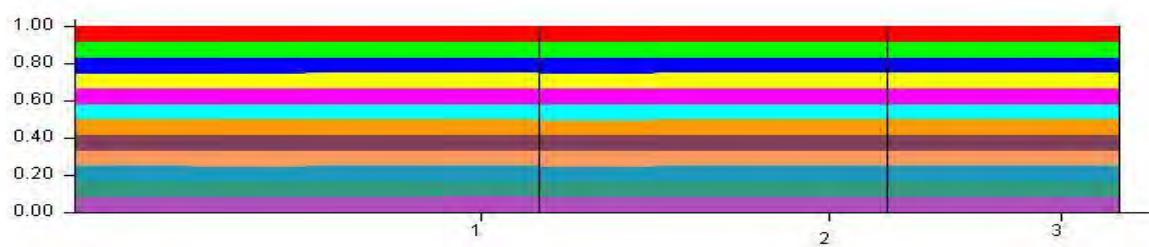
**Figure 12-1:** Représentation du nombre de groupes de variétés améliorées en fonctions de Delta K (K groupes k allant de 1 à 12 (méthode d'Evanno et al. 2005), Delta k = f(K)



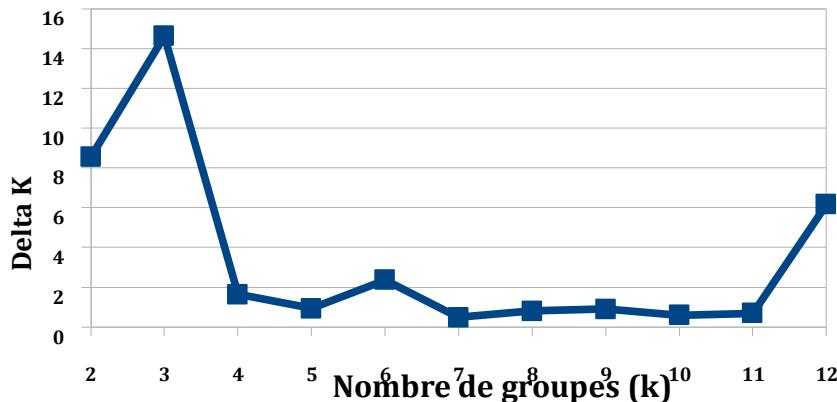
**Figure 12-2 :** Représentation graphique du test d'assignation des variétés améliorées par STRUCTURE pour K= 12. Chaque k groupe est représenté par une couleur et Chaque barre verticale représente un individu



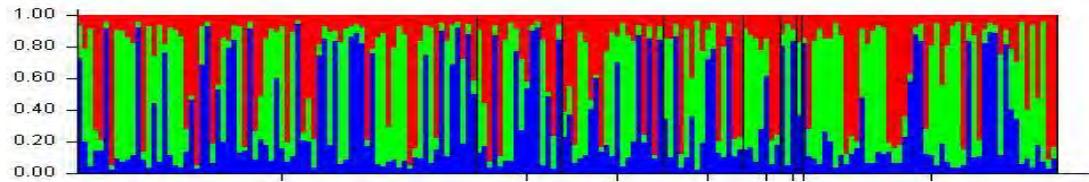
**Figure 13-1 :** Représentation du nombre de groupes des variétés tardives en fonctions de Delta K (K groupes k allant de 1 à 12 (méthode d'Evanno et al. 2005), Delta k = f(K)



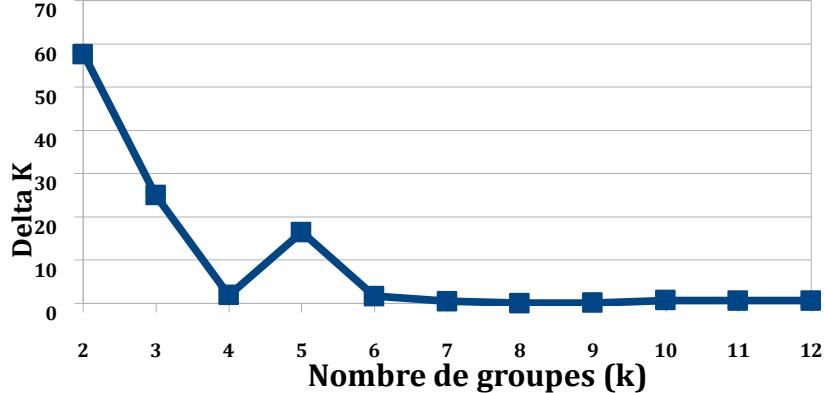
**Figure 13-2 :** Représentation graphique du test d'assignation des variétés tardives par STRUCTURE pour K= 12. Chaque k groupe est représenté par une couleur et Chaque barre verticale représente un individu. 1 =Diourbel, 2= Fatick, 3= Kaolack et Thiès



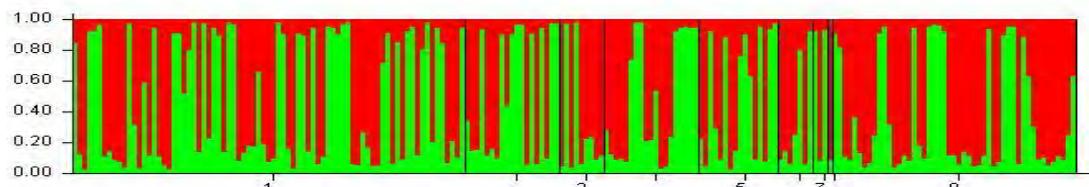
**Figure 14-1** : Représentation du nombre de groupes des variétés précoces en fonctions de Delta K (K groupes k allant de 1 à 12 (méthode d'Evanno et al. 2005), Delta k = f(K)



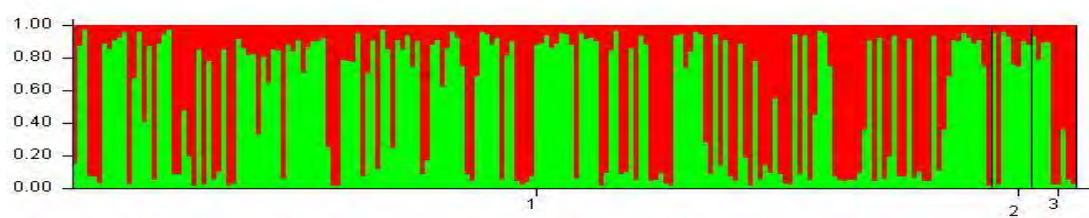
**Figure 14-2** : Représentation graphique du test d'assignation des variétés précoces par STRUCTURE pour K= 3. Chaque k groupe est représenté par une couleur et Chaque barre verticale représente un individu. 1=Diourbel, 2= Fatick, 3= Kaffrine, 4= Kaolack, 5= Louga, 6= Matam, 7= Tambacounda, 8= Thiès



**Figure 15-1** : Représentation du nombre de groupes de toutes les accessions en fonctions de Delta K (K groupes k allant de 1 à 12 (méthode d'Evanno et al. 2005), Delta k = f(K)



**Figure 15-2** : Représentation graphique du test d'assignation de toutes les accessions par STRUCTURE pour K= 2. Chaque k groupe est représenté par une couleur et Chaque barre verticale représente un individu. 1= Diourbel, 2= Fatick, 3= Introduites, 4= Kaffrine, 5= Kaolack, 6= Louga, 7= Matam, 8= Tambacounda, 9= Thiès



**Figure 15-3** : Représentation graphique du test d'assignation des types variétaux par STRUCTURE pour K= 2. Chaque k groupe est représenté par une couleur et Chaque barre verticale représente un individu. 1= Variétés précoces (Souna), 2= Variétés améliorées, 3= Variétés tardives (Sanio)

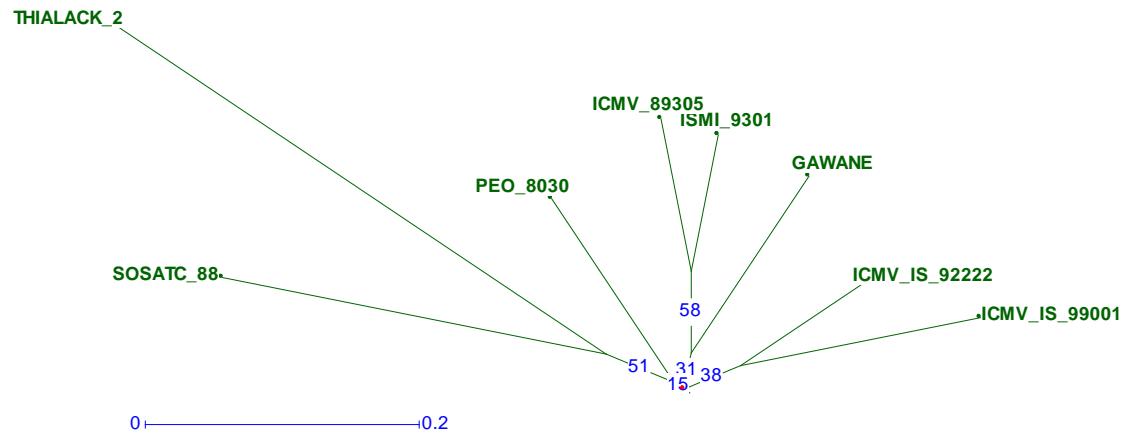
#### 4.3. Relations phylogénétiques au sein de la collection

Les dendrogrammes ont permis d'analyser les relations phylogénétiques des individus au sein de la collection. Dans le dendrogramme représentant les variétés améliorées, trois groupes peuvent être distingués. Dans le groupe 1 on peut noter SOSATC\_88 et Thialack\_2 qui s'isolent dans ce groupe tandis que PEO\_8030 forme un sous-groupe sœur. Dans le deuxième groupe on peut noter, ICMV\_89305 et ISMI\_9301 qui forment un sous-groupe, sœur à GAWANE. Le troisième groupe est formé par ICMV\_IS\_92222 et ICMV\_IS\_99001.

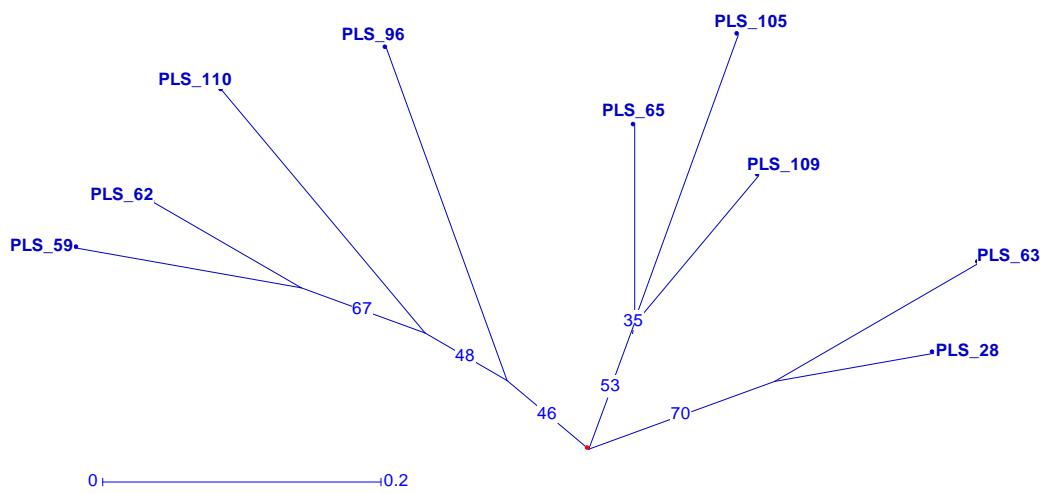
Le dendrogramme représentant les variétés tardives, peut être également subdivisé en trois grands groupes. Dans le groupe 1, il y a PLS\_59 et PLS\_62 qui forment un sous-groupe tandis que PLS\_110 et PLS\_96 s'isolent au niveau du dendrogramme. Le groupe 2 est formé par PLS\_109, PLS\_65 et PLS\_105. Quant au groupe 3, il est formé par PLS\_63 et PLS\_28.

Les relations phylogénétiques entre variétés précoces sont représentées par la Figure 18 dans laquelle on distingue trois groupes qui peuvent être subdivisés chacun en 3 sous-groupes.

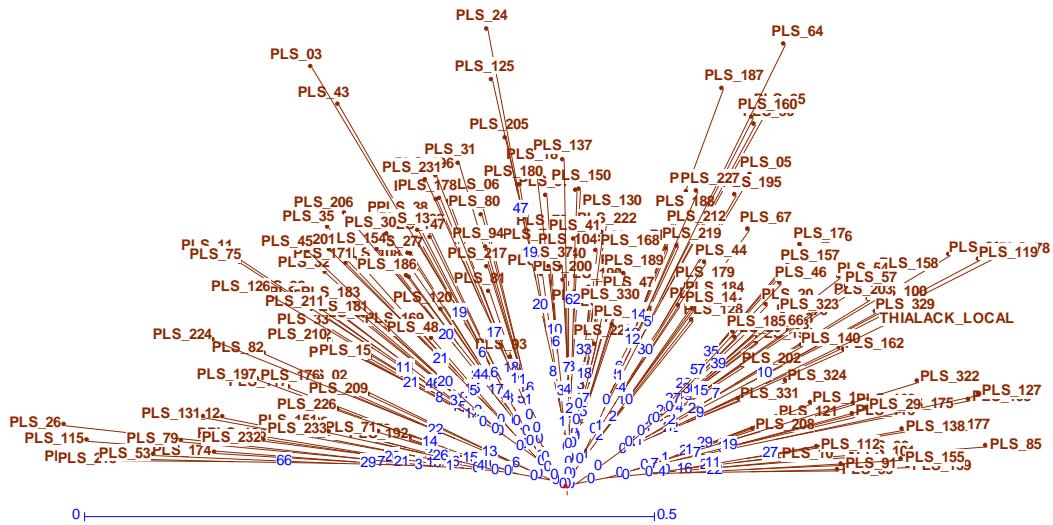
La topologie de l'arbre général regroupant tous les individus permet de les diviser en trois groupes. Dans le groupe 1 on trouve des variétés tardives et des variétés précoces. Ce groupe peut être subdivisé en 4 sous-groupes. Le groupe 2 est formé par des variétés précoces et des variétés tardives et une seule variété améliorée qui est Thialack\_2. Cette dernière se trouve dans le même groupe que son géniteur le Thialck\_local. Les variétés précoces (PLS\_194 et PLS\_189) forment un sous-groupe sœur. Le groupe 3 renferme la quasi-totalité des variétés améliorées avec seulement 2 variétés tardives (PLS\_96 et PLS\_109) et des variétés précoces.



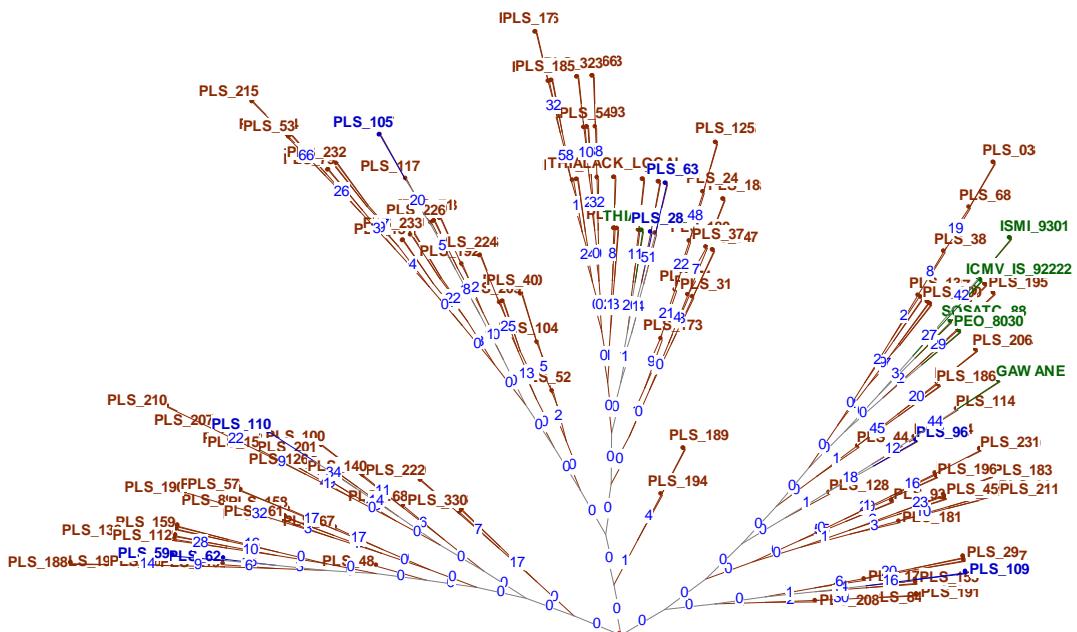
**Figure 16** : Dendrogramme montrant les relations phylogénétiques entre les variétés améliorées. Les valeurs Bootstrap sont indiquées sur les branches



**Figure 17 :** Dendrogramme montrant les relations phylogénétiques entre les variétés tardives. Les valeurs Bootstrap sont indiquées sur les branches



**Figure 18 :** Dendrogramme montrant les relations phylogénétiques entre des variétés précoces  
Les valeurs Bootstrap sont indiquées sur les branches



**Figure 19 :** Dendrogramme montrant les relations phylogénétiques entre toutes les variétés  
Les valeurs Bootstrap sont indiquées sur les branches

## **DISCUSSION**

## DISCUSSION

Pour la totalité des accessions testées dans la collection, il n'y a pas de doublons. Les 202 accessions de cette étude sont toutes génétiquement différentes les unes des autres. Ceci est normal car le mil est une plante allogame or il est rare de constater des doublons chez les plantes allogames(Noirot et al., 1993).

Les accessions de la collection présentent une diversité relativement importante avec une richesse allélique égale à celle des mils cultivés du Niger(Mariac et al., 2006) (0.62) mais avec une diversité génétique ( $H_{ex}$ ) plus élevée (0.594 contre 0.49 pour ceux du Niger). Ces valeurs de la diversité sont en accord et relativement proches de celles trouvées par(Stich et al., 2010)(richesse allélique de 0.56 et diversité génétique ( $H_{ex}$ ) de 0.69 pour 13 mils cultivés originaires du Sénégal). Ces résultats attirent aussi l'attention sur la Région de Diourbel, la seule région qui s'est révélée significative au test d'équilibre d'Hardy et Weinberg au seuil de 5% pour toutes les variétés qui en sont originaires (Souna et Sanio). En effet, elle semble présenter un excès d'homozygotes pouvant résulter d'un léger taux de consanguinité des individus issus de cette région, aussi bien pour les variétés précoces que tardives. Ces résultats suggèrent un taux faible d'échanges de semences entre la région de Diourbel et les autres régions, ou une absence de mutation, ou une sélection(Jordana et al., 2003).Ces études permettent aussi de faire des déductions sur la structure de la diversité génétique de la collection.

L'analyse comparée de la diversité génétique des Souna et des Sanio ne révèle presque aucune différence. En effet, cette diversité génétique ( $H_{ex}$ ) est de 0.584 en moyenne pour les Souna contre 0.580 pour les Sanio. Cette quasi-absence de différence dans la diversité génétique des Souna et des Sanio pourrait suggérer des fréquences similaires de recombinaisons, de mutations génétiques et des flux semblables de migration. Ce qui pourrait être expliqué par le fait que ces deux types variétaux sont presque toutes issues des mêmes localités et ont donc évolué sous les mêmes conditions de vie. Par ailleurs, la richesse allélique moyenne des Souna (quatre allèles par locus) est supérieure à celle des Sanio (deux allèles par locus). Cette différence est sans doute due à la différence de taille des populations (185 individus pour les Souna contre neuf pour les Sanio). En effet la taille de la population a une grande influence sur la richesse allélique (Foulley and OLLIVIER, 2006).

En se basant sur l'approche Bayésienne et sur les relations génétiques entre les types variétaux, l'ensemble de la collection peut être subdivisé en deux grands groupes A et B. En effet, l'analyse de la structure de l'ensemble des accessions suggère une répartition plus marquée en deux groupes. Le premier groupe (vert) est celui auquel appartient presque toutes les variétés améliorées. Le second groupe (rouge) est le groupe dominant chez les variétés tardives. Les variétés précoces sont réparties entre ces deux groupes. L'analyse des relations génétiques de toutes les variétés confondues, par le dendrogramme, permet d'avoir la même observation. Un groupe renfermant la quasi-totalité des variétés améliorées et un autre avec la majeure partie des variétés tardives. Ces deux groupes renferment tous des variétés précoces. Dans le premier groupe (A) se définissent les sous-groupes (a), (1) et (3) et deux individus de variétés tardives dont l'un provient du sous-groupe (β) et l'autre du sous-groupe (γ). Dans le groupe B nous retrouvons un sous-groupe de variétés améliorées (le sous-groupe 2), deux sous-groupes des variétés précoces (les sous-groupes (b) et (c)) et tout le reste des variétés tardives. (Fig. 8 à 10).

L'accession Thialack\_2 est la seule variété améliorée qui se regroupe avec les accessions locales du groupe B. Cela peut être expliqué par le fait qu'elle soit issue de la variété locale et traditionnelle Thialack. L'isolation de cette accession peut s'expliquer par le fait qu'elle s'est regroupée avec des accessions ayant sans doute un géniteur commun.

C'est dans le groupe A de la collection que nous rencontrons les accessions locales les plus proches génétiquement des variétés améliorées car ce groupe renferme presque toutes les accessions de

variétés améliorées avec des variétés précoces et presque pas de variétés tardives. Ce groupe pourrait probablement renfermer les géniteurs des variétés améliorées constituant à cet effet, un groupe d'intérêt pour des programmes de sélection variétale.

Dans le groupe A de la collection, il y a presque toutes les accessions améliorées avec un nombre important de variétés précoces mais il est à noter qu'il y a une quasi-absence des variétés tardives qui se trouvent presque toutes dans le groupe B où il n'y a presque pas de variétés améliorées. Cela suggère que les variétés améliorées sont génétiquement plus proches des variétés précoces. Il en ressort de cette étude que le groupe B renferme des individus plus proches des variétés tardives au sein de la collection. Cela suggère également que dans les programmes de sélection, les variétés précoces ont souvent été utilisées comme parent géniteur par contre tel ne semble pas être le cas des variétés tardives. Cela pourrait s'expliquer par l'adaptation des variétés précoces aux courtes saisons pluvieuses dans la zone sahélienne. Il est à noter que les variétés tardives ne doivent pas être négligées car elles constituent un immense réservoir de gènes potentiellement intéressants pour l'agriculture locale.

## **CONCLUSIONS ET PERSPECTIVES**

## CONCLUSIONS ET PERSPECTIVES

Les marqueurs utilisés ont été bien capables de caractériser génétiquement la collection et de montrer une absence de redondance de génotypes.

Les objectifs de cette étude ont été atteints et l'hypothèse de travail a été vérifiée et permet de dire que la collection de mil du Sénégal présente une diversité relativement importante en accord avec certaines études déjà réalisées en Afrique. Cette étude a permis de mettre en évidence la structure de la diversité génétique de la collection de mil du Sénégal d'une façon générale et particulièrement de déterminer les relations phylogénétiques entre accessions.

L'isolement de la variété SOSATC\_88 au niveau des ACP et des dendrogrammes suggèreraient une différence entre cette variété et le reste des accessions de la collection. Par ailleurs, des études biologiques portant sur sa morphologie, son écologie, son comportement vis-à-vis des pathogènes ou la composition organoleptique de ces graines pourraient être réalisées pour comprendre son isolement dans cette étude.

La région de Diourbel mérite d'être mieux étudiée pour identifier l'origine exacte du léger excès du taux de consanguinité des accessions qui en sont originaires, pour éviter la perpétuation de ce facteur, qui à la longue, diminue cette diversité et impacte négativement les capacités adaptatives de ces accessions.

Il serait aussi intéressant d'envisager un enrichissement de la collection avec de nouvelles prospections surtout vers le sud du Sénégal. Dans le but d'augmenter le nombre de variétés tardives et de variétés améliorées de la collection. Mais aussi dans le but d'avoir une bonne représentativité de toutes les régions du Sénégal dans la collection de mil, étant donné qu'aucune accession provenant du sud n'y a pas été incluse. Cela pourrait être la cause du faible nombre d'accessions tardives (Sanio) qui sont plus fréquentes dans ces régions où la pluie est plus abondante.

Il serait de même important d'utiliser les groupes d'intérêt de la collection dans les programmes de sélection variétale afin d'améliorer la tolérance aux stress abiotiques comme la sécheresse et biotiques comme le *Striga* et autres contraintes. Cela peut s'avérer très utile en termes de sécurité alimentaire dans la zone sahélienne.

## **REFERENCES BIBLIOGRAPHIQUES**

## REFERENCES BIBLIOGRAPHIQUES

- Agapow P.-M. and Burt A. (2001). Indices of multilocus linkage disequilibrium. *Mol. Ecol. Notes.* **1**: 101-102.
- Allouis S., Qi X., Lindup S., Gale M. D. and Devos K. M. (2001). Construction of a BAC library of pearl millet, *Pennisetum glaucum*. *Theor. Appl. Genet.* **102**: 1200-1205.
- Bezançon G., Renno J.-F. and Anand Kumar K. (1997). Le mil. In L'amélioration des plantes tropicales. Charrier A., Jacquot M., Hamon S. and Nicolas D. (eds) (Montpellier: CIRAD). 457-482.
- Blein R. and Soulé B. G. (2011). Les céréales au cœur de la souveraineté alimentaire en Afrique de l'Ouest. *Grain de sel.* **54-56**: 3-15.
- Budak H., Pedraza F., Cregan P. B., Baenziger P. S. and Dweikat I. (2003). Development and Utilization of SSRs to Estimate the Degree of Genetic Relationships in a Collection of Pearl Millet Germplasm. *Crop Sci.* **43**: 2284-2289.
- Chakauya E. and Tongona P. (2008). Analysis of genetic relationships of pearl millet (*Pennisetum glaucum* L.) landraces from Zimbabwe, using microsatellites. *International Journal of Plant Breeding and Genetics.* **2(1)**: 1-7.
- Chantereau J., Deu M., Pham J. L., Kapran I., Vigouroux Y. and Bezancon G. (2010). Evolution des diversités phénotypiques et génétiques des sorghos et mils cultivés au Niger de 1976 à 2003. Le sélectionneur français. **61**: 33-45.
- Cloutault J., Thuillet A.-C., Buiron M., De Mita S., Couderc M., Haussmann B. I. G., Mariac C. and Vigouroux Y. (2012). Evolutionary history of pearl millet (*Pennisetum glaucum* [L.] R. Br.) and selection on flowering genes since its domestication. *Mol. Biol. Evol.* **29**: 1199-1212.
- De Vienne D. (1998). Les marqueurs moléculaires en génétique et biotechnologies végétales. Inra-Quae (ed.) 200 p.
- Evanno G., Regnaut S. and Goudet J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14**: 2611-2620.
- Falush D., Stephens M. and Pritchard J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics.* **164**: 1567-1587.
- Foulley J.-L. and Ollivier L. (2006). Estimating allelic richness and its diversity. *Livest. Sci.* **101**: 150-158.
- Frankel O. H. (1970). Variation - the essence of life. *Proc. Linn. Soc. New South Wales.* **95**: 158-169.
- Govindaraj M., Selvi B., Prabhu D. A. and Rajarathinam S. (2009). Genetic diversity analysis of pearl millet (*Pennisetum glaucum* [L.] R. Br.) accessions using molecular markers. *Afr. J. Biotechnol.* **8**: 6046-6052.
- Haussmann B. I. G., Boubacar A., Boureima S. S. and Vigouroux Y. (2006). Multiplication and preliminary characterization of west and central African pearl millet landraces. *Int. Sorghum and Millet Newsletter.* **47**: 110-112.

Jordana J., Alexandrino P., Beja-Pereira A., Bessa I., Cañon J., Carretero Y., Dunner S., Laloë D., Moazami-Goudarzi K., Sanchez A. and Ferrand N. (2003). Genetic structure of eighteen local south European beef cattle breeds by comparative F-statistics analysis. *J. Anim. Breed. Genet.* **120**: 73-87.

Kapila R. K., Yadav R. S., Plaha P., Rai K. N., Yadav O. P., Hash C. T. and Howarth C. J. (2008). Genetic diversity among pearl millet maintainers using microsatellite markers. *Plant Breed.* **127**: 33-37.

Manning K., Pelling R., Higham T., Schwenniger J.-L. and Fuller D. Q. (2011). 4500-Year old domesticated pearl millet (*Pennisetum glaucum*) from the Tilemsi Valley, Mali: new insights into an alternative cereal domestication pathway. *J. Archaeol. Sci.* **38**: 312-322.

Marchais L. (1982). Diversité phénotypique des mils pénicillines cultivés au Sénégal et au Mali. *L'agronomie tropicale.* **XXXVII-1**: 68-80.

Marchais L., Tostain S. and Amoukou I. (1993). Signification taxonomique et évolutive de la structure génétique des mils pénicillaires. In *Le mil en Afrique. Diversité génétique et agro-physiologique : potentialités et contraintes pour l'amélioration génétique et l'agriculture. Colloques et séminaires.* Hamon S. (ed.) ORSTOM. Paris, France. 119-127.

Mariac C., Luong V., Kapran I., Mamadou A., Sagnard F., Deu M., Chantereau J., Gerard B., Ndjeunga J., Bezançon G., Pham J. L. and Vigouroux Y. (2006). Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. *Theor. Appl. Genet.* **114**: 49-58.

MendonçaNeto R. P., Von Pinho E. V. R., Carvalho B. L. and Pereira G. S. (2013). Identification of pearl millet cultivars using both microsatellites and enzymatic markers. *Genet. Mol. Res.* **12**: 1-14.

Nelson G. C., Rosegrant M. W., Koo J., Robertson R., Sulser T., Zhu T., Ringler C., Msangi S., Palazzo A., Batka M., Magalhaes M., Valmonte-Santos R., Ewing M. and Lee D. (2009). Changement climatique Impact sur l'agriculture et coûts de l'adaptation. Rapport de l'IFPRI (Institut international de recherche sur les politiques alimentaires). 30p.

Noirot M., Hamon S. and Anthony F. (1993). L'obtention d'une collection de cafiers : définition des groupes d'échantillonnage et méthodologie. XVI<sup>e</sup> Colloque scientifique international sur le café. Paris, France, ASIC. 183-190.

Oumar I., Mariac C., Pham J. L. and Vigouroux Y. (2008). Phylogeny and origin of pearl millet (*Pennisetum glaucum* [L.] R. Br) as revealed by microsatellite loci. *Theor. Appl. Genet.* **117**: 489-497.

Peakall R. and Smouse P. E. (2012). GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research an update. *Bioinformatics.* **28**: 2537-2539.

Perrier X., Flori A. and Bonnot F. (1999). Les méthodes d'analyse des données. In *Diversité Génétique des Plantes Tropicales Cultivées.* Hamon P., Seguin M., Perrier X., Glaszmann J.-C. (eds.) Montpellier, CIRAD. 43-87.

Poncet V., Lamy F., Enjalbert J., Joly H., Sarr A. and Robert T. (1998). Genetic analysis of the domestication syndrome in pearl millet (*Pennisetum glaucum* L., Poaceae): inheritance of the major characters. *Heredity.* **81**: 648-658.

Pritchard J. K., Stephens M. and Donnelly P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*. **155**: 945-959.

Qi X., Lindup S., Pittaway T. S., Allouis S., Gale M. D. and Devos K.M. (2001). Development of simple sequence repeat markers from bacterial artificial chromosomes without subcloning. *BioTechniques*. **31**: 355-362.

Rosenzweig C. and Parry M. L. (1994). Potential impact of climate change on world food supply. *Nature*. **367**: 133-138.

Saïdou A.-A., Mariac C., Luong V., Pham J. L., Bezançon G. and Vigouroux Y. (2009). Association studies identify natural variation at *PHYC* linked to flowering time and morphological variation in pearl millet. *Genetics*. **182**: 899-910.

Saitou N. and Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.

Selkoe K. A. and Toonen R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9**: 615-629.

Soule B. G. and Gansari S. (2010). La dynamique des échanges régionaux des céréales en Afrique de l'ouest. Rapport, Université de l'état de Michigan. 111p.

Stich B., Haussmann B. I., Pasam R., Bhosale S., Hash C. T., Melchinger A. E. and Parzies H. K. (2010). Patterns of molecular and phenotypic diversity in pearl millet [*Pennisetum glaucum* (L.) R. Br.] from West and Central Africa and their relation to geographical and environmental parameters. *BMC Plant Biol.* **10**: 216.

Supriya A. (2010). Genetic Diversity Analysis and QTL Mapping in Pearl Millet (*Pennisetum glaucum*) using Diversity Arrays Technology (DART). Thèse de doctorat, Université agricole Chaudhary Charan Singh Haryana. 98p.

Tostain S. (1992). Enzyme diversity in pearl millet (*Pennisetum glaucum* L.). *Theor. Appl. Genet.* **83**: 733-742.

Tostain S. (1998). Le mil, une longue histoire : hypothèses sur sa domestication et ses migrations. In *Plantes et paysages d'Afrique. Une histoire à explorer*. Chastanet K. M. (ed.) Centre de Recherches Africaines. Paris, France. Chap. **16**: 461-490.

Van Oosterhout C., Hutchinson W.F., Wills D. P. M. and Shipley P. (2004). micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes*. **4**: 535-538.

Winkel T., Renno J.-F. and Payne W. A. (1997). Effect of the timing of water deficit on growth, phenology and yield of pearl millet (*Pennisetum glaucum* (L.) R. Br.) grown in Sahelian conditions. *J. Exp. Bot.* **48**: 1001-1009.

Yadav O. P. (2007). Morphological description and yield performance of pearl millet landraces collected from Rajasthan, India. *Journal of SAT Agricultural Research*. **5(1)**. 9p.

Zon V. D. (1992). Graminées du Cameroun. Thèse de doctorat, Université agricole Wageningen : 557p.



## **ANNEXES**

## ANNEXE 1

### PROTOCOL D'EXTRACTION D'ADN SOLUTIONS

Tampon Lyse: qsp d'eau milliQ

Tris 100 mM

NaCl 1.25 mol/l

EDTA 0.02 mol//l

Ajuster à pH 8 avec de l'acide chlorhydrique HCl et stocker à température ambiante.  
(ou mettre le tris et l'EDTA à pH 8 avant).

*à rajouter extemporanément dans la quantité de tampon juste nécessaire*

2 % deMATAB (0119 mol/L) ou CTAB

2 % deβmercaptoethanol ou DTT (0.01M)

Solution de chloroforme/isoamylique 24:1 :

48 mL de chloroforme

2 mL d'alcool isoamylique

Acétate de sodium 3 M pH 5 : qsp 250 mL

20412 g d'acétate de sodium (3 mol/L)

qsp 500 mL d'eau milliQ

Ajuster à pH 5 avec de l'acide acétique glacial

Tampon TE 1X : qsp 500 mL d'eau milliQ

06055 g de Tris (001 mol/L)

01861 g d'EDTA (0001 mol/L)

Ajuster à pH 8 avec de l'acide chlorhydrique HCl 5 N et stocker à température ambiante.

## ANNEXE 2

### PCR MULTIPLEX KIT QIAGEN ET GENOTYPAGE

Le PCR multiplex kit a été mis au point pour les PCR en multiplex.

Préparer un volume de réaction avec 10% de plus que nécessaire pour les réactions.

Concentrations finales du mix :

1x de Master mix (solution 2x)	5 µl
0.2 µM de chaque primer	0.2 µl (solution à 10 µM)
1x de solution Q (solution à 5x)	2 µl
Volume de l'échantillon d'ADN	1 µl (concentration à 25 ng/µl)
Qsp eau Rnase free.	

Distribuer le mix dans les tubes. 9 µl

Distribuer ensuite les échantillons d'ADN. 1 µl

Tableau 8 : Tableau récapitulatif des caractéristiques moléculaires des marqueurs microsatellites utilisés dans notre étude et les auteurs de publication

Nom	Taille minimum	Taille maximum	AmorceForward = Amorce sens (5' 3')	Amorce Reverse = Amorce anti sens (3' 5')	Dye	Multiplex	Motif	Publication
IRD46	80	110	GCCGACCAAGAACCTCATACA	GAACAATTGCTTCTGTAATATTGCTT	Ned	A	(CTC) <sub>6</sub>	(Mariac et al.2006)
Psmp2227	180	215	ACACCAAAACACCAACCATAAAG	TCGTCAGCAATCACTAATGACC	Ned	A	(GT) <sub>7</sub>	(Allouis et al.2001)
Psmp2249	125	160	CAGTCTCTAACAAACAAACACGGC	GACAGCAACCAACTCCAAACTCCA	Fam	A	(GT) <sub>7</sub>	(Allouis et al.2001)
Psmp2247	180	231	CCAAACCGTAACCTGAAAAGCTACTG	GTGTCGGTTGCTTCGTTCCCTT	Fam	A	(TG) <sub>6</sub>	(Allouis et al.2001)
Psmp2201	310	375	CCCGACGTTATGCGTTAAGTT	TCCATCCATCCATTAAATCCACA	Red	A	(GT) <sub>6</sub>	(Qi et al.2001)
Psmp2237	210	260	TGGCCTTGGCCTTCCACGCTT	CAATCAGTCCTAGTCCACACCCCCA	Red	A	(GT) <sub>8</sub>	(Allouis et al.2001)
CTM8	225	275	GCTGCATCGGAGATAGGGAA	CTCAGCAAGCACGCTGCTCT	Vic	A	(CT) <sub>8</sub>	(Budak et al.2003)
IRD12	110	145	CGGGGAAGAGACAGGGCTACT	ACTCGTTCGGATGCACTTCT	Vic	A	(TA) <sub>8</sub>	(Mariac et al.2006)
IRD25	165	190	GCAAGCCACAAGCCTATCTC	CGGAGCTCCTATCATTCAA	Red	B	(GA) <sub>9</sub>	(Mariac et al.2006)
Psmp2214	230	255	GATTGAGCAGAAAAACCAGC	CGCACAGTACGTGTGAGTGAAG	Red	B	(GT) <sub>9</sub>	(Qi et al.2001)
Psmp2219	210	295	ACTGATGGAATCTGCTGTGGAA	GCCCCGAAGAAAAGAGAACATAGAA	Ned	B	(GT) <sub>7</sub>	(Qi et al.2001)
Psmp2220	100	145	GCATCCTTCACCATTCAAGACA	TGGGAAACAGAACATGGAGAAAAGAG	Ned	B	(GT) <sub>11</sub>	(Qi et al.2001)
Psmp2206	190	210	AGAAGAAGAGGGGTAAAGAAGGAG	AGCAACATCCGTAGAGGTAGAAG	Fam	B	(GT) <sub>11</sub>	(Qi et al.2001)
Psmp2202	130	165	CTGCCTGTTGAGAATAATGAG	GTTCCGAATATAGAGCCCAAG	Vic	B	(GT) <sub>8</sub>	(Qi et al.2001)
Psmp2231	212	262	TTGCCTGAAGACGTGCAATCGTCC	CTTAATGCGTCTAGAGAGTTAAGTTG	Vic	B	(TG) <sub>12</sub> GG(TA) <sub>4</sub>	(Allouis et al.2001)

Tableau 9 : Liste des Accessions de l'étude et leurs génotypes

Sample	Genotype	Sample	Genotype
PLS_01	8491205205260260146146203207131149287287g	PLS_57	9191199205237237146163199201118131280287g
PLS_02	9198199205260260146163199203126129287289g	PLS_59	8491199199235260146163203205118133287287g
PLS_03	8491204204235235146163203209118131280282g	PLS_62	8491199205260260155163203205129133287287g
PLS_05	9191199205235235146163199209120135262287g	PLS_63	9191205205260260163163203207118118280280g
PLS_06	8498199205260260146163201203120135283287g	GAWANE	9191205205235258146163201203129133287287g
PLS_10	8491199205260260163163199203131131287287g	PLS_64	9191199205235235163163203207118118280280g
PLS_11	8484199204260260146146199205131131287287g	PLS_65	8498205205258260163163199201118122266280g
PLS_12	8491199205235260146163203205120120264282g	PLS_66	9191199205260260146163199199118133280287g
PLS_13	8491205205235258146163203205131137287287g	PLS_67	9191199205258258146163199199131133287287g
PLS_14	9191199205258260146163203209129131287287g	PLS_68	8491199205235258146163203207118131280287g
PLS_15	8491199205260260146146199203129131282287g	PLS_69	8491205205235260163163201203129131287289g
PLS_16	8498205205235260146163205209129131287287g	PLS_71	8491205205235260146163199203120131283287g
PLS_17	9198199205260260146163199199118122266280g	PLS_72	8491205205235260163163203203135135287287g
PLS_18	8491205205260260146163199205122122266269g	PLS_73	8491205205260260146163205207122131266287g
PLS_19	9191199205260260146163199205129131287287g	PLS_74	9191199205260260146163203205122129269287g
PLS_20	9198199205235260146163203205118129280287g	PLS_75	8484199199260260146163205205131137287287g
PLS_24	9191199199235262146146203205122129269287g	PLS_76	8491199205260260146163199205118131280282g
PLS_25	919120520526026014616320511131133287287g	PLS_77	9191205205260260146163203203120120283283g
PLS_26	9198204204258260146146199203120120266283g	PLS_78	9898205205260260146146205207118118280280g
PLS_27	8491205205258260146146199203129135282287g	PLS_79	91101199199260260146163199203120120266282g
PLS_28	9191205205260260146163201203118122266280g	PLS_80	8498205205260260146163199199120135287287g
PLS_29	8491205205260260163163205207118124282287g	PLS_81	84101205205260260146163199203131135287287g
PLS_30	8491205205235260146146199209120129283287g	PLS_82	8491204204260260146163203203120129283287g
PLS_31	8484205205260260146163201205122133264287g	PLS_83	9191199205260260146163201203131133287287g
PLS_32	8491205205258260146146203203133137287287g	PLS_84	9191205205260260163163199199135135287287g
PLS_33	8498199205260260146146203205129131287287g	PLS_85	9191204204258260163163203203131131282287g
PLS_35	8491199206235260146163201205129135287287g	PLS_86	98101199199260260146163205205118124282287g
PLS_36	8491205205260260146146201209131135287287g	PLS_87	9191199205258260146163199203118120280287g
PLS_37	9191205205260260146163199205122131266287g	PLS_88	9198199205260260146163199205120129283287g
PLS_38	8491205205235264146163199203124131282287g	PLS_89	9191199205260260146163203203133137287287g
PLS_39	9198199205260260163163203205118131280287g	PLS_90	8491199205260262163163199203131131287289g
PLS_40	9191205205260260146146203203129131287287g	PLS_91	8491199205260260163163199209118131280287g
PLS_41	9191205205260260146163199203118120280283g	PLS_92	8484205205260260146146199199131133287287g
PLS_42	8491205205235260146163199207118133280287g	PLS_93	8491205205260260146163199203131135287287g
PLS_43	8498205205235258146146199203118131280282g	PLS_94	8498205205260260146163199203122129269287g
PLS_44	919120520526026146163199207120120287287g	PLS_95	9191199205235235146146203203137137287287g
PLS_45	8491199205258260146146203203129129287289g	PLS_96	9191205205258260146146199209129131287287g
PLS_46	9198199199260260146163199203118135280287g	PLS_97	8491205205260260146146199211122131266287g
PLS_47	8491205205260260146163201203118131280287g	PLS_98	9191205205260260146146199203118120280283g
PLS_48	8491199205260262146163199203129131287287g	PLS_99	9191199199260260146146199199139283283g
PLS_50	9191199205235237146146203209118120280287g	PLS_100	8498199205260262146163205205118129280287g
PLS_52	9191205205260260146163199203120129283287g	PLS_101	8491205205258260163163203203118131280287g
PLS_53	9191199199260260146163203207137137283283g	PLS_103	8491199205260262163163199199118124282282g
PLS_54	849119920523526014616319921118120280283g	PLS_104	9191205205260260146146199203126131283287g
PLS_55	91101199205260260146146199207131133287287g	PLS_105	8498205205258260163163199205120120283287g
PLS_56	9191205205235260146163199201120131287287g	PLS_106	9198199205260260146163203203118122269280g

Sample	Genotype	Sample	Genotype
PLS_107	8491205205235260146146199203135147287287g	PLS_175	9198205205260260163163205205133135282287g
PLS_108	8491205205260262146163199199118120282287g	PLS_176	9191204204260260146163199203131133287287g
PLS_109	8498205205260260163163199199122131266287g	PLS_177	8484199199260260163163203203129131287287g
PLS_110	848419919926026014616319920913113287287g	PLS_178	8484205205258260163163199203129131282287g
PLS_111	8491205205235262146163199205118129280287g	PLS_179	9191205205235260146163199205133135287287g
PLS_112	8491199205260260163163203203131144287287g	PLS_180	9191199205260260146146199201122129269287g
PLS_113	9191199205260260146163199209118133280287g	PLS_181	8491205205260260146146203203120131282287g
PLS_114	9191205205258260146163199203129133282287g	PLS_182	9191199205260260146146199207129135287287g
PLS_115	9191199199260260163163201205122122266269g	PLS_183	8491199205260260146146201203133137287287g
PLS_116	8491205205260260163163199199124133282287g	PLS_184	9191205205258260146163203207131133287287g
PLS_117	8491199205258260163163203205120129283287g	PLS_185	9198199205260260146163199203118135280287g
THIALACK_2	9191205205260260146146207207122147266266g	PLS_186	8491205205235260146163199207133133287287g
PLS_119	9191199199260260146146205205118147280280g	PLS_187	9198199199258258163163199203129133287287g
PLS_120	8491205205235260146163199205120129287287g	PLS_188	9191199205258260163163199203129133287289g
PLS_121	9191205205260260163163199205131131287287g	PLS_189	9191205205260260146163199201131131280287g
PLS_122	8491199205260260146146207213131133287287g	PLS_190	9191199205260260146163201203131135287287g
PLS_125	9191199205235235146146199199122129269287g	PLS_191	9198205205260260163163199199131131287287g
PLS_126	9191199199260266146146203205129131287287g	PLS_192	8491199205260260146163199203120131283287g
PLS_127	9191205205235260163163199201124124282282g	PLS_193	8491199205260260146163199199118133280287g
PLS_128	9198205205258260146163199203118131287287g	PLS_194	9191205205260260146163201203120131287287g
PLS_130	9191199205260260146163205205133133287287g	PLS_195	9191205205235235146163199207135135287287g
PLS_131	9191199205258260146163201205120131264286g	PLS_196	848420520525826014616320321135135287287g
PLS_137	9191205205260260146146201201129147282287g	PLS_197	8498205205260260163163203209120129283287g
PLS_138	8491199199260260163163203203133133287287g	PLS_198	8491199205260260146163199201118120280287g
PLS_140	9198199205260260146163205205129129287287g	PLS_199	9191199205258260163163201205129133287287g
PLS_147	8491205205235260146163201207131135282287g	PLS_200	9191205205260260146146199203129135287287g
PLS_150	9191205205260260163163199203118129280289g	PLS_201	8484205205260260163163205207129131287287g
PLS_151	8491199205260260146163205207120129283287g	PLS_202	9191199205260260146163199205118129280287g
PLS_153	8491205205260260146163203205120131283287g	PLS_203	9191199206219260146163199205118131280287g
PLS_154	8491205205260262146163199199122133269287g	PLS_204	8491199205235260146163199203118129280287g
PLS_155	8498205205260260163163197199118131280287g	PLS_205	9191199205260262146146199199122129269287g
PLS_157	8491199205258260146163199203118122269280g	PLS_206	8491205205260262146163199199118122264280g
PLS_158	9191199205262262146163205205118131287289g	PLS_207	8491199205260260146146199203129135287287g
PLS_159	9191199199260260163163203205118133280287g	PLS_208	9191205205260260163163199203131137287287g
PLS_160	9191199205235260163163203203120133283289g	PLS_209	9198205205260260146163203203129139287287g
PLS_161	9191199205260260146163203205129131287287g	PLS_210	8491199205260260146146199207129129287287g
PLS_162	8491205205260260146163205205118129280289g	PLS_211	8484205205260260146146199203131137287287g
PLS_163	9191205205260260163163199207122137269287g	PLS_212	9191205205235258146163199203129129287287g
PLS_164	919119920526026014616319920313113287287g	PLS_215	9191199199260260146146199207120139266283g
PLS_166	9191199205260260163163199203120131283283g	PLS_217	8491205205260260146163201203120122269287g
PLS_167	9191199205260260163163203207129133287287g	PLS_219	9191199205235260163163199203129135287287g
PLS_168	91101205205260260146163205205129131287287g	PLS_222	9191199205260260146163199205124133282287g
PLS_169	8491199205235260146163199205129131287287g	PLS_223	9191205205260260146163199203131133287287g
PLS_170	8484205205260260163163203203129133287287g	PLS_224	9198204204260260146163203207133135227287g
PLS_171	8484199205260260146163207209129131287287g	PLS_226	8491205205260260163163199203120135283287g
PLS_172	9191199205260260146163199199120131283287g	PLS_227	9191205205235235163163199203120129287287g
PLS_173	9191205205260260146163203207122131269287g	PLS_231	8484205205260260163163199199133133287287g
PLS_174	9191199199260260146163199203122129269289g	PLS_232	9191199205260260146163201203120135283283g

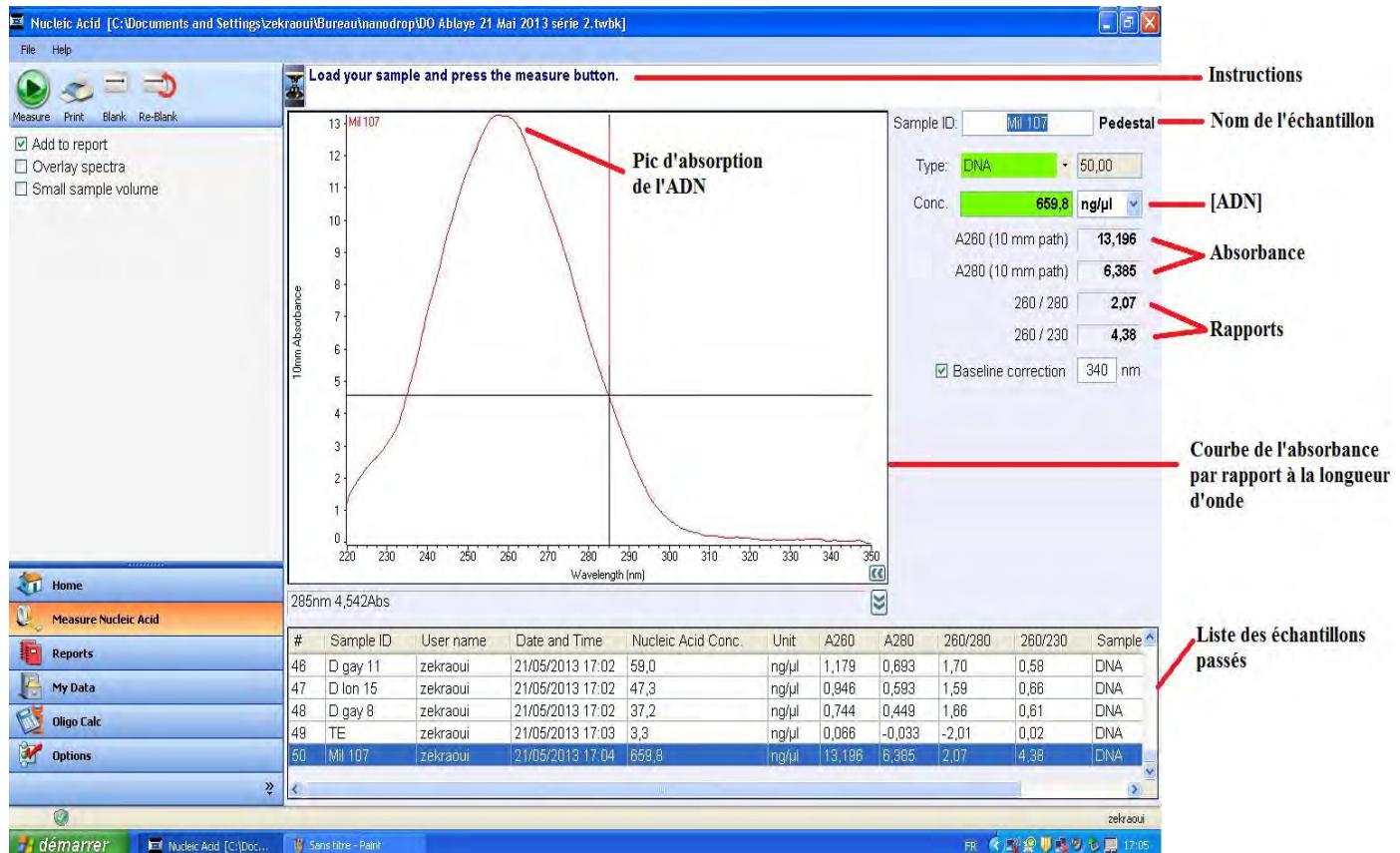
Sample	Genotype	Sample	Genotype
PLS_233	9191199205260260146163205205120131283287g	ISMI_9301	9191205205235235163163203203129129287287g
PLS_322	9198205205260260163163199199120122269283g	THIALACK_LOCAL	91101205205235260146163199205118124280282g
PLS_323	8491199205260260146146199199118131280287g	ICMV_89305	9191205205235235163163205209129131287287g
PLS_324	91101205205260260146163199199118129280287g	ICMV_IS_99001	9191205205235235155163192203126126282287g
PLS_329	9191199205260260146146203205118135280286g	SOSATC_88	9898205205235235146146201201126126287287g
PLS_330	9191205205260260146163199203133133287287g	ICMV_IS_92222	9191205205235235146163203203131131282287g
PLS_331	9198205205260260146163199203118129280287g	PEO_8030	9191205205235235146163201205120131283287g

## ANNEXE 3

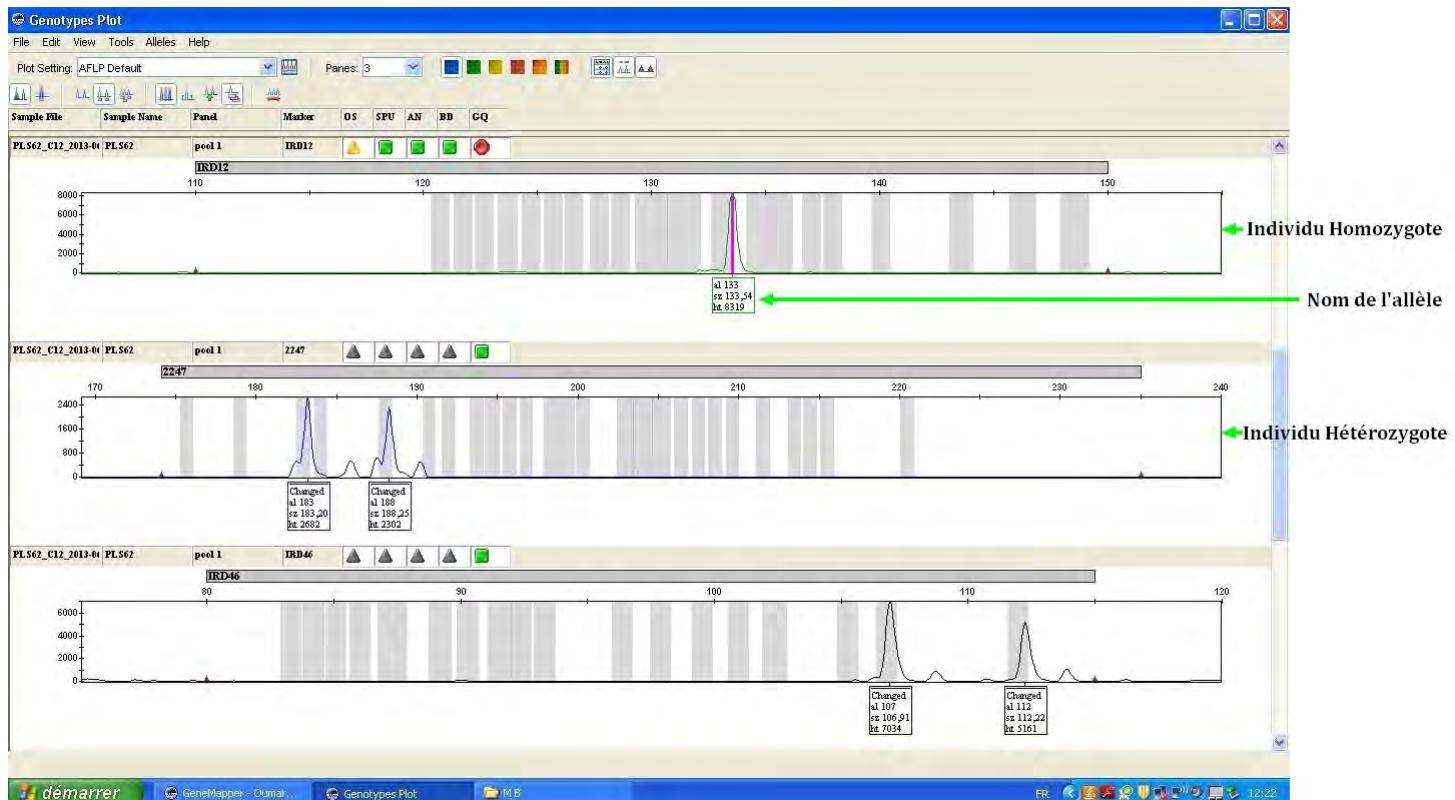
### Photos illustrant quelques étapes du stage



**Figure 20** : Photo des semis des échantillons de mil en serre (O. Diack)



**Figure 21** : Photo de la représentation à l'ordinateur de la mesure d'un échantillon par le Nano Drop (M. Texier)



**Figure 22 :** Photo du profil d'un individu homozygote et hétérozygote par rapport à un locus sur GeneMapper (O. Diack)

# ANNEXE 4 : Poster



IRD  
Institut de recherche  
pour le développement

## Etudes de la diversité génétique de la collection de mil [*Pennisetum glaucum* spp *glaucum*] (L.) R. Br.] DU Sénégal



Oumar Diack<sup>1,2</sup>, Ndjido Kane<sup>1</sup>, Adeline Barnaud<sup>1</sup>, Diaga Diouf<sup>2</sup>

1 Laboratoire National de Recherche sur les Productions Végétales (LNRPV)/ISRA, Route des hydrocarbures, Bel-air, BP 3120 Dakar Sénégal. 2 Laboratoire de Biotechnologies Végétales, Université Cheikh Anta Diop de Dakar, Faculté des Sciences et Techniques, Département de Biologie Végétale, BP 5005, Dakar, Sénégal

E-mail : omar.diack@isra.sn

### Introduction

Le mil ou *Pennisetum glaucum* spp. *glaucum* est la céréale la plus cultivée en Afrique Sub-Saharienne, où il constitue l'aliment de base de cinquante (50) millions d'individus de la population. Dans le contexte des changements globaux actuels, l'agriculture et la sécurité alimentaire sont menacées dans cette zone et doivent être renforcées par la production de variétés adaptées. Dans le but de produire des variétés qui répondent à ces changements, le sélectionneur s'appuie sur les ressources génétiques, d'où l'importance d'avoir au préalable une bonne caractérisation de ces ressources. Dans cette étude nous faisons la caractérisation génétique de la collection de mil du Sénégal à l'aide de marqueurs moléculaires de type microsatellite.

**Objectif global:** Déterminer la diversité génétique d'accessions de mils pénicillaires

**Objectifs spécifiques:** (i) Examiner la structure de cette diversité génétique à l'aide de marqueurs codominants et polymorphes de type microsatellites ; (ii) Identifier les (sous) groupes de mils.

### Matériaux et Méthodes

#### Matériaux :

252 Accessions de la collection de mil du Sénégal (13 accessions introduites et 239 accessions locales dont 229 accessions précoces et 10 accessions tardives). 15 marqueurs microsatellites (SSR)

#### Méthodes :

Analysé du niveau de diversité : Calcule des indices de diversité (Richesse allélique, Hétérozygotie observée, Hétérozygotie attendue, Indice de fixation)

Analysé de la structure de la diversité : Analyse en composantes principales, Approche Bayésienne et dendrogrammes

### Résultats

Une subdivision en trois sous-groupes pour chaque type variétal

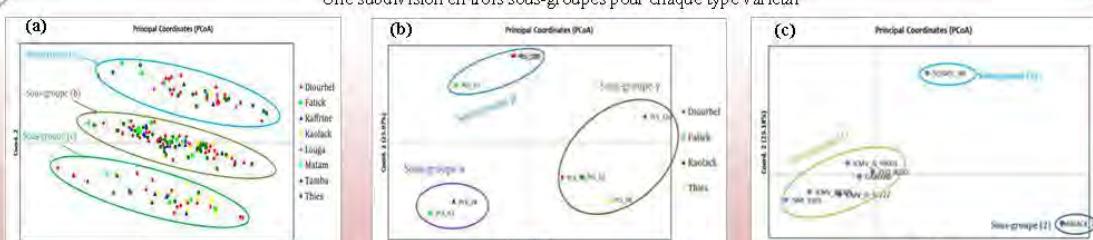


Figure 1 : ACP des variétés précoces (a), tardives (b) et améliorées (c)

Subdivision plus marquée des accessions en 2 groupes

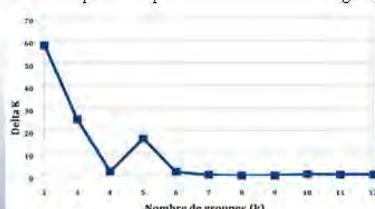


Figure 2 : Représentation du nombre de groupes de toutes les accessions en fonction de Delta (méthode d'Evanno et al. 2005)

Diversité relativement importante

	RA	Hex	Ho	Fis
Souna	4	0.58	0.56	0.02
Sanio	2	0.58	0.47	0.2
Améliorée	5	0.6	0.4	0.33

Figure 3 : Indices de diversité

Deux grands groupes (A et B)



Figure 4 : Dendrogramme montrant les relations phylogénétiques entre toutes les variétés de la collection

Appartenance des variétés améliorées et tardives de la collection à différents groupes

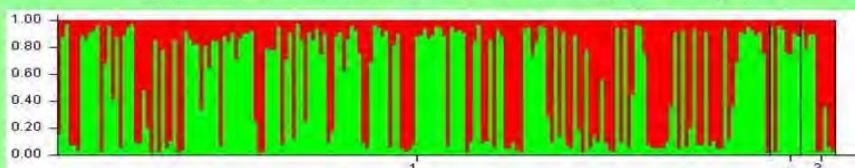


Figure 5 : Représentation graphique du test d'assignation des types variétaux de la collection par STRUCTURE pour K= 2. 1= Variétés précoces (Souna), 2= Variétés améliorées, 3= Variétés tardives (Sanio)

### Conclusion

La collection de mil du Sénégal présente une diversité relativement importante

Il existe deux grands groupes (A et B) dans la collection

Le groupe A présente de l'intérêt pour la sélection variétale

Les variétés précoces sont génétiquement plus proches des variétés améliorées que les variétés tardives

En perspective, il serait intéressant d'envisager un enrichissement de la collection avec de nouvelles prospections intégrant les régions du Sud du pays, dans le but d'augmenter les variétés tardives et variétés améliorées de la collection. Il serait aussi intéressant d'envisager l'utilisation des groupes d'intérêt de la collection pour la sélection variétale