THÈSE
Pour obtenir le grade de
Docteur

Délivré par le
Centre international d'études supérieures en sciences agronomiques
Montpellier

Préparée au sein de l'école doctorale ***Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosciences et Environnement (SIBAGHE)***
Et de l'UMR AGAP

Spécialité : **Biologie des populations**

Présentée par **Vanesse Labeyrie**

L'organisation sociale des plantes cultivées

Influence des échanges, représentations et pratiques sur la diversité du sorgho (*Sorghum bicolor* [L.] Moench) chez les peuples du mont Kenya

Soutenue le 19 décembre 2013 devant le jury composé de

Mme Victoria REYES-GARCIA, Pr, Universitat de Barcelona	Rapportrice
M. Thierry ROBERT, Pr, Université Paris-Sud XI	Rapporteur
Mme Devra JARVIS, Dr, Bioversity International	Examinatrice
M. Jacques DAVID, Pr, SupAgro Montpellier	Examinateur
M. Jean-Christophe GLASZMANN, Dr, CIRAD	Directeur de thèse
M. Christian LECLERC, Dr, CIRAD	Co-directeur de thèse

REMERCIEMENTS

Je souhaite en premier lieu remercier les agricultrices et agriculteurs qui ont participé à cette étude, et sans l'implication de qui elle n'aurait pas pu voir le jour. Au Kenya, je remercie Zachary Muthamia, directeur de la banque de gènes, Peterson Wambugu et Joseph Kamau, pour leur accueil et leur aide, ainsi que le reste de l'équipe, notamment la « caracterization-team », qui a fait preuve d'une grande rigueur dans l'acquisition de données. Sur le terrain, je souhaite remercier Catherine Wanjira, ainsi que Wanja, Kagendo et Pamela, qui ont été mes guides et mes interprètes, les professeurs et les élèves de l'école primaire de Kabururu, pour leur implication dans ce projet et pour avoir hébergé notre parcelle expérimentale. Mes remerciements vont également à Joseph Mugo, Bernard Rono et Caroline Mwongera, ainsi qu'à David Williamson.

En France, mes remerciements vont à Jean-Christophe Glaszmann, Christian Leclerc et Adeline Barnaud pour avoir encadré cette thèse, pour leurs conseils et leur soutien. Je souhaite également remercier Monique Deu, pour son aide tout au long de cette thèse, Marylène Buiron, Caroline Calatayud et Ronan Rivallan, qui ont réalisé le génotypage, Léo Valette, qui a considérablement contribué à l'acquisition de données sur le terrain, parfois à ses risques et périls. Merci aussi à Claire, Jean-François, Fabien, Marc et le reste de l'équipe EGV pour les divers coups de mains, et bien sur à Guylaine.

L'aide de Xavier Perrier, Cécile Dubois, Philippe Letourmy, Luc Baudouin et Stéphanie Manel pour les analyses statistiques a été précieuse, ainsi que celle de Jacques Chantereau et de Gilles Trouche, qui m'ont notamment aidé à préciser le protocole de phénotypage. Geo Coppens d'Eeckenbrugge a consacré un temps considérable à relire mes projets d'articles, je tiens à le remercier pour son aide tout au long de ma thèse. Vincent Freycon m'est venu en aide pour l'analyse de données pédologiques. Merci également à Anne-Marie Peatrik pour avoir partagé un peu de son expérience chez les Meru. L'équipe de Centre de Bio-Archéologie et d'Ecologie m'a accueillie pour des essais de morphométrie, merci notamment à Sarah Ivorra et Sandrine Picq pour leur aide. Enfin, je souhaite remercier Jean-Louis Noyer, qui a contribué à faire naitre ce projet de thèse, ainsi que Jean-Louis Pham et Jean-Pierre Labouisse, coordinateurs du projet ARCAD. Par ailleurs, ce projet n'aurait pu être mené à bien sans le soutien financier de la fondation Agropolis, qui m'a accordé une bourse de trois ans, celui du projet ARCAD, qui a financé les missions sur le terrain, et celui du CIRAD qui a hébergé cette thèse.

Je remercie enfin ma famille, pour son soutien constant, jusque dans la traque sans pitié des coquilles, ainsi que mes amis et Jef, qui ont été là quand il le fallait.

Ce voyage s'est achevé au Kenya, mais il avait commencé dans le Pacifique. Je remercie mes amies et amis des tribus de Wakatr à Iaai (Ouvéa) et de Mia à Canala, ainsi que celles et ceux du Sénat Coutumier de Nouvelle-Calédonie – notamment Raoul Bouacou. Entre autre choses, ils m'ont appris que le monde des hommes et celui de plantes ne font qu'un.

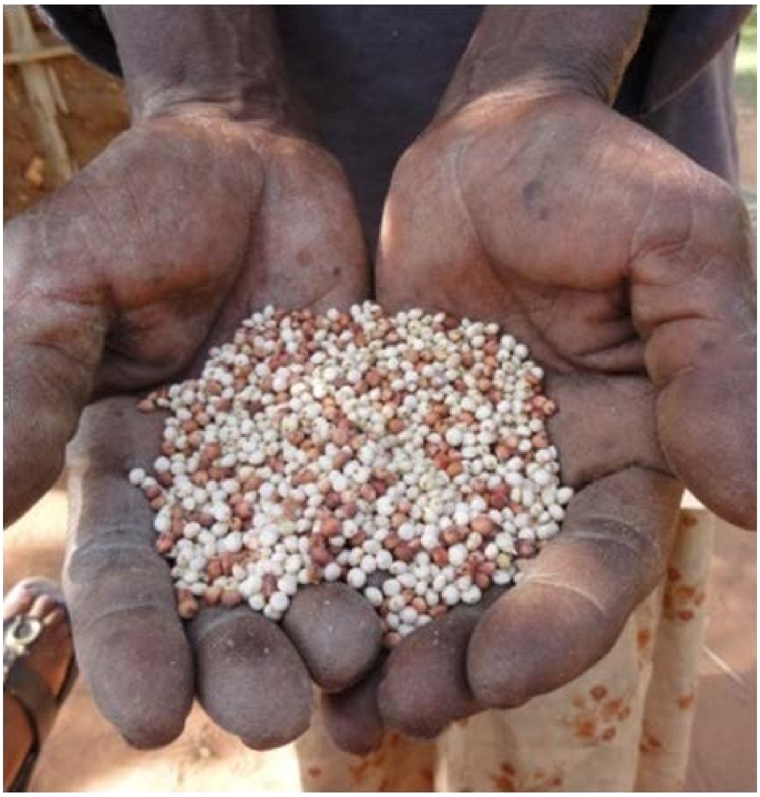
TABLE DES MATIERES

TABLE DES FIGURES ET TABLEAUX

TABLE DES ANNEXES

“Crops are artifacts made and molded by man as much as a flint arrowhead, a stone ax-head, or a clay pot”

(Harlan 1975)



INTRODUCTION GENERALE

I- Enjeu : Caractériser et conserver les ressources génétiques

Le développement des sociétés humaines, associé à l'adoption de l'agriculture lors de l'entrée dans l'ère néolithique, a été largement conditionné par la diversité des ressources génétiques végétales et animales dont elles disposaient (Diamond 2002). Ces ressource génétiques ont été domestiquées, diffusées, puis sélectionnées localement dans diverses zones du globe depuis 13 000 ans, comme l'attestent divers sites archéologiques. La diversité ainsi générée a soutenu le développement des sociétés néolithiques, permettant la production alimentaire dans diverses conditions agro-climatiques, face aux pathogènes et ravageurs des cultures, et tout en répondant aux différents usages et goûts des sociétés humaines (Gepts 2006). A l'inverse, l'uniformité des ressources génétiques sur lesquelles certaines sociétés humaines ont basé leur alimentation a parfois menacé leur stabilité. La grande famine provoquée par l'épidémie de Phytophtora dans les années 1840 en Irlande en est un des exemples les plus patents.

L'importance de la diversité des ressources génétiques a été officiellement reconnue en 1992 dans le cadre de la CBD (Convention on Biological Diversity). Cette reconnaissance faisait cependant écho aux menaces qui pesaient sur elle, et que certains scientifiques avaient fait connaitre dès les années 1936 (Harlan and Martini 1936), mais principalement depuis les années 1970 (Harlan 1972; Hawkes 1971; Frankel and Bennett 1970; Wilkes 1977). L'érosion génétique, c'est-à-dire la perte de diversité génétique des plantes cultivées, s'est accrue pendant la seconde moitié du 20 ème siècle. L'augmentation démographique considérable à laquelle les pays en voie de développement ont du faire face et la globalisation des économies, des déplacements et des moyens de communication seraient les causes profondes de cette rapide perte de diversité (Gepts 2006). Dans ce contexte, les politiques d'intensification de l'agriculture mises en place dans de nombreuses régions du globe, notamment pendant le seconde moitié du 20 ème siècle, ont largement contribué à la disparition d'une part considérable des ressources génétiques des plantes cultivées. Cette érosion génétique a d avantage touché les zones dont le potentiel agricole était le plus élevé, à commencer par l'Europe avec la Politique Agricole Commune après la seconde guerre mondiale, l'Amérique du Sud, l'Inde et l'Asie du sud-est avec la « révolution verte » à partir des années 1960.

Avec l'intensification de l'agriculture, les ressources génétiques des plantes cultivées ont commencé à représenter d'importants enjeux financiers. Alors que ces ressources avaient été considérées comme un héritage commun de l'humanité depuis leur domestication, la CBD les a placées sous la souveraineté des états et elles sont devenues l'objet de droits de propriété intellectuels (Gepts 2004b). La radicalisation du système UPOV depuis sa création en 1961 a quasiment effacé les variétés paysannes du paysage agricole dans de nombreux pays, dont la France. Le système de propriété intellectuelle concernant les plantes cultivées s'étend maintenant aux pays en voie de développement, qui y sont souvent contraints par la mise en place d'accords de libre échange avec les pays ayant déjà adopté ce système (Gepts 2004b). Dans les pays où une majeure part de la population dépend des semences de ferme, l'adoption d'un tel système apparait inadaptée alors que les échanges de semences produites à la ferme sont un pilier de la sécurité alimentaire et de l'économie locale. Ce système de propriété intellectuelle, bien que difficilement applicable et contrôlable dans les pays du sud, est néanmoins une menace de plus pour le maintien de la diversité génétique des plantes cultivées.

Dans ce contexte d'érosion accélérée des ressources génétiques, leur caractérisation et leur conservation sont devenues un enjeu central pour la recherche agronomique. Paradoxalement, le développement de l'amélioration variétale formelle, qui est une des causes d'érosion des ressources génétiques *in situ*, dépend du maintien d'une large gamme de diversité génétique (Gepts 2006). Cette diversité représente le matériel de base des sélectionneurs. Elle détermine le potentiel d'adaptation des systèmes agricoles face aux contraintes agro-climatiques, biotiques et à leur évolution rapide, ainsi qu'aux divers usages que notre société fait des produits agricoles. Face à l'urgence de conserver la diversité des ressources génétiques, des initiatives de conservation *ex situ*, au sein de banques de gènes (Plucknett et al. 1987), puis in situ, au sein des systèmes de cultures où elles ont été générées (Brush 2000), ont été mises en place. Aujourd'hui, près de 1500 banques de gènes conservent près de 5,5 millions d'accessions. Onze d'entre elles ont le statut de banques de gènes internationales, relevant du CGIAR (Consultative Group on International Agricultural Research), le reste étant des banques de gènes nationales ou régionales (FAO 1996). Les initiatives de conservation *in situ* visent à promouvoir le maintien de la diversité des plantes cultivées auprès des populations locales. Ces initiatives menées dans le cadre de programmes nationaux ou internationaux, ciblent généralement les centres de diversité des différentes espèces cultivées (Brush 2000).

Cependant, les collections conservées dans le réseau de banques de gènes ne représentent qu'imparfaitement la diversité façonnée au cours de milliers d'années d'évolution au sein des écosystèmes cultivés dans le monde entier. En effet, les stratégies de collectes sur lesquelles étaient basées les prospections dans les années 1970, et qui ont permis d'alimenter les banques de gènes, étaient largement limitées par les connaissances et des moyens disponibles (Guillaumet and Pernès 1984). Ces collectes n'ont généralement couvert que les zones les plus accessibles et, malgré les efforts faits en ce sens, leur couverture spatiale est largement incomplète. Lorsque cela était possible logistiquement et financièrement, les prospections ont cherché à considérer la diversité agro-écologique des pays concernés. En revanche, l'importance des sociétés humaines dans le façonnage des ressources génétiques a largement été négligée lors de ces prospections. Ainsi, tout en reconnaissant l'importance des sciences humaines, Guillaumet et Pernes regrettaient qu'aucune prospection en Afrique n'ait à leur connaissance impliqué un anthropologue (Guillaumet and Pernès 1984).

Malgré l'intérêt qui leur est porté depuis les années 1970, et l'augmentation considérable de nos connaissances dans ce domaine, les ressources génétiques cultivées restent encore largement méconnues. La connaissance des processus qui ont généré et façonné la distribution spatiale des ressources génétiques est un enjeu majeur pour la conservation de celles-ci au regard du coût élevé des programmes de caractérisation, de collecte et de conservation des ces ressources. En effet, la connaissance de ces facteurs permet d'optimiser les stratégies d'échantillonnage, c'est-à-dire de cibler les zones de collectes *in situ* où le maximum de variabilité génétique pourra être collecté. Ces informations sont également essentielles pour la mise en place de programmes de conservation *in situ*. La compréhension des mécanismes impliqués dans l'évolution des plantes cultivées doit donc concentrer davantage d'efforts de recherche, et sous un angle pluridisciplinaire (Gepts 2006).

II- Le rôle de l'agriculteur dans l'évolution des plantes cultivées

Les mécanismes qui ont et continuent de façonner la diversité des plantes cultivées *in situ* restent largement méconnus. Comme leurs parentes sauvages, les populations de plantes cultivées sont soumises à quatre principales forces évolutives : la sélection, la migration, la dérive et la mutation. L'adaptation aux conditions climatiques, édaphiques ou encore la pression biotique ou la compétition avec d'autres espèces sont des facteurs bien connus dans l'évolution des plantes. En revanche, les plantes cultivées sont soumises à un facteur supplémentaire : l'action humaine. Ce facteur a largement été négligé dans la caractérisation et la collecte des ressources génétiques, alors que les sociétés humaines ont et continuent d'occuper une place centrale dans la domestication, la diffusion et la sélection des plantes cultivées.

La domestication stricto sensu désigne le processus évolutif par lequel l'homme a sélectionné des populations afin de les adapter au milieu cultivé et de satisfaire ses usages (Gepts 2004a). Suite à la domestication, les plantes cultivées deviennent dépendantes des pratiques culturales humaines pour leur survie, leur dispersion et donc leur perpétuation. La domestication s'accompagne en effet de changements phénotypiques considérables, désignés sous le terme de syndrome de domestication. Ces changements portent sur les caractères d'intérêt concernant l'usage (e.g. augmentation de la taille des grains chez les céréales), l'aptitude à être cultivées (e.g. perte de dormance des graines, tallage basal) et récoltées (e.g. perte de la dispersion des graines, synchronisation de la maturité) (Glémin and Bataillon 2009). La domestication s'est faite au fil d'une longue interaction entre les populations humaines et les plantes, et elle se poursuit aujourd'hui au sein des systèmes agricoles familiaux. Dans son acceptation plus large, le terme domestication désigne en effet les processus continus de sélection humaine entrainant des changements phénotypiques dans les populations de plantes cultivées (Gepts 2004a).

De la cueillette à l'agriculture

Bien avant l'adoption de l'agriculture, l'homme pratiquait largement la cueillette. Sans qu'il s'agisse à proprement parler de domestication, cette pratique a pu, sous certaines formes, commencer à provoquer des modifications des caractéristiques phénotypiques et génétiques des populations de plantes sauvages (Gepts 2004a). Il est probable que consciemment ou non l’homme ait favorisé le développement de plantes dont les caractéristiques phénotypiques présentaient un intérêt pour ses usages. De telles pratiques ont notamment été observées au Mexique sur certaines cactées (*Stenocereus stellatus*) produisant des fruits comestibles (Casas et al. 1997), ou encore en Papouasie Nouvelle-Guinée sur le sagoutier (*Metroxylon sagu*) où les individus épineux ont été quasiment éliminés des populations situées à proximité des villages (Barrau 1962).

Il est probable que des pratiques d’ensemencement à partir de graines récoltées dans les populations sauvages soient ensuite progressivement apparues, ce qui s'apparente à un processus de pre-domestication (Willcox et al. 2007; Tanno and Willcox 2006). Entre 11 000 et 8 500 avant J.C, on observe sur certains sites du Moyen-Orient, tels que celui de Jerf el Ahmar en Syrie, une diminution de l’usage des plantes sauvages à petites graines, parallèlement à l’augmentation de la fréquence de l’orge (*Hordeum spontaneum*), l’engrain (*Triticum boeoticum*), et l’amidonnier (*Triticum dicoccoides)* dans les restes archéologiques, les capacités de stockage se développent. Ces observations suggèrent que la cueillette et la culture des céréales sauvages a eu lieu sur une longue durée avant que les premiers signes de domestication à proprement parler n'apparaissent (Willcox et al. 2007). Ainsi, bien que certains travaux suggèrent que la domestication des céréales aurait pu avoir eu lieu en moins de 200 ans (Hillman and Davies 1990), les preuves archéologiques, étayées par des études de modélisation des processus évolutifs, confortent au contraire l’hypothèse d’une domestication progressive, diffuse et ayant potentiellement impliqué des évènements multiples de domestication, ou encore des contributions réitérées des parents sauvages au génome des populations domestiquées (Tanno and Willcox 2006; Allaby et al. 2008).

De la domestication à la diffusion

La domestication des différentes espèces cultivées aurait ainsi eu lieu de façon indépendante dans diverses zones situées au Proche-Orient, en Asie, en Océanie, en Amérique et en Afrique, et sur une période s’étendant d’environ 13 000 ans à 3 000 ans avant notre époque. Dans une même zone, différentes espèces ont parfois été domestiquées à des époques différentes, comme c’est le cas pour le maïs, domestiqué il y a environ 9 000 ans (Piperno et al. 2009), et le haricot, il y a 4 000 ans (Chacón et al. 2005). Par ailleurs, la domestication d’une même espèce a eu parfois lieu indépendamment dans différentes régions. L’un des exemples les mieux documentés est celui du haricot commun (*Phaseolus vulgaris*), domestiqué indépendamment en Mésoamérique et en Amérique du Sud (Gepts et al. 1986).

Au cours de ces évènements de domestication, les caractères constituant le syndrome de domestication ont été fixés dans les populations cultivées par la succession de cycles de recombinaison-sélection. La domestication a par ailleurs entrainé un effet de dérive génétique dans les populations cultivées, se traduisant par une perte de diversité dans l'ensemble du génome. En effet, l'homme n'a échantillonné qu'une infime fraction des populations sauvages au cours de ce processus.

Diverses hypothèses ont été émises concernant les raisons de cette transition vers l'agriculture, certaines évoquant une « adaptation asymétrique » (l'homme est contraint de s'adapter aux modifications de son environnement) due a des changements climatiques (Richerson et al. 2001), ou a une augmentation démographique (Cohen 1977), d'autres supposant plutôt que l'homme aurait été réellement acteur de cet amélioration de ses ressources (Smith 2011) et qu'il s'agissait d'une initiative culturelle délibérée (Cauvin et al. 1997). Selon Diamond, enfin, cette transition aurait été favorisée dans certaines zones par rapport à d'autres car les espèces végétales et animales qui s'y trouvaient étaient plus aptes à être domestiquées (Diamond 2002).

Que l'adoption de l'agriculture en ait été la cause ou la conséquence, il est en tout cas établi que cette période a été caractérisée par une expansion des sociétés néolithiques (Diamond 2002). La diffusion de l'agriculture, et donc des plantes cultivées se serait alors réalisée de façons directe (démique) et indirecte (culturelle), mais la part des deux mécanismes reste sujet à controverse. L'hypothèse d'une diffusion démique suggère que les populations néolithiques auraient migré et diffusé ainsi directement les plantes qu'elles avaient domestiquées dans leur zone d'origine, ainsi que leurs propres gènes, langage et culture, remplaçant progressivement les sociétés de chasseurs-cueilleurs. Il s'agit de l'hypothèse la plus couramment évoquée pour l'expansion des populations Bantoues en Afrique de l'Ouest, celle des populations Austronésiennes dans le Pacifique, et celle des populations Indo-européennes en Eurasie (Cavalli-Sforza 2002). Cependant, la comparaison de données linguistiques, génétiques, et anthropologiques vient remettre partiellement en cause ce scénario jugé trop simpliste par certains, notamment pour le peuplement du Pacifique (Donohue and Denham 2010). En effet, l'agriculture se serait également propagée par diffusion culturelle, les plantes cultivées et les techniques agricoles auraient alors été diffusées de proche en proche, sans mouvements massifs de populations depuis leurs centres de domestication. Ce mode de diffusion aurait largement été basé sur les réseaux d'échanges régionaux, dépendant des relations sociales existantes entre sociétés.

Par la combinaison de ces mécanismes, les plantes cultivées ont été diffusées dans des zones éloignées de leurs aires de domestication. Ultérieurement, les mouvements humains ont encore accentué la diffusion des plantes cultivées à grande échelle. C'est ainsi que la banane, domestiquée en Asie du Sud-est, aurait atteint l'Afrique de l'Est, il y a probablement plus de 3 000 ans, vraisemblablement introduite par des navigateurs Austronésiens (Perrier et al. 2011). Des connexions préhistoriques entre l'Amérique du Sud et la Polynésie auraient conduit à l'introduction de la patate-douce en Océanie alors qu'elle avait été domestiquée sur le continent Américain (Roullier et al. 2013). En sens inverse, le cocotier qui avait été domestiqué en Asie du Sud-est aurait été introduit en Equateur par des navigateurs Austronésiens, il y a plus de 2 000 ans (Baudouin and Lebrun 2009). La distribution des plantes cultivées reflète donc l'histoire des sociétés humaines et de leurs relations.

L'adoption des plantes cultivées par des sociétés différentes a favorisé leur diversification car elles ont été sélectionnées pour répondre à différents usages, goûts et préférences, ainsi que pour leur adaptation à des environnements agro-climatiques divers. Le sorgho (*Sorghum bicolor*) illustre bien cette diversification en relation avec son adoption par des sociétés différentes. Cette espèce cultivée présente une grande diversité phénotypique, cinq races botaniques ont été distinguées par Harlan dans ce continuum de diversité (Harlan and De Wet 1972), avec des caractéristiques morphologiques distinctes. La distribution spatiale de quatre de ces races botaniques en Afrique coïncide avec celle des grandes familles de langues, la race caudatum étant associée aux familles de langues Chari-Nil, la race kafir aux populations Bantoues, la race guinea aux populations Niger-Congo et la race durra aux populations de langue Afro-asiatiques, principalement les langues Arabes (Harlan and Stemler 1976).

Les sociétés, au cours de leur histoire, ont ainsi largement contribué à modeler la diversité des plantes cultivées en modifiant profondément les forces de migration, de sélection et de dérive qui s'appliquait à ces populations. Cependant, ces mécanismes ne font pas partie du passé, et les sociétés humaines continuent aujourd'hui dans de nombreuses régions du monde à modeler la diversité des plantes cultivées.

La domestication : un processus continu et actuel

La domestication des plantes, dans sa compréhension large d'une action humaine modifiant les caractéristiques phénotypiques des populations de plantes, continue de se dérouler au sein des systèmes agricoles familiaux, peu intensifiés. Ces systèmes occupent encore une part importante des terres agricoles, notamment dans les foyers de diversité des plantes cultivées, et constituent de véritables modèles pour comprendre l'évolution des plantes cultivées (Brush 2000). En effet, contrairement aux systèmes agricoles intensifs, les plantes cultivées au sein des systèmes familiaux sont fortement soumises aux pressions évolutives de leur environnement agro-climatique et biotique en raison du faible degré d'artificialisation du milieu. De plus, elles sont souvent sympatriques avec des espèces ou sous-espèces sauvages apparentées, avec lesquelles des flux de gènes peuvent avoir lieu. Enfin, elles sont soumises aux pratiques des agriculteurs, portant principalement sur la gestion des semences, qui agissent sur les forces de sélection, de dérive et de migration appliquées à ces populations.

Migration via les flux de pollen et recombinaisons

Au sein des champs cultivés ou entre champs proches, des flux de pollen peuvent avoir lieu entre variétés de la même espèce ainsi qu'avec des apparentées sauvages ou adventices (Jarvis and Hodgkin 1999). Le type de système de culture a un impact considérable sur les flux de pollen, qui vont être déterminés principalement par la densité d'implantation, le mélange de variétés et leurs différence de phénologie, ou encore l'association de plusieurs espèces au sein de la parcelle. Les flux de pollen entre parcelles vont être influencés par leur éloignement et la présence de haies ou d'obstacles les séparant.

L'intensité des recombinaisons dépend ensuite de la biologie reproductive de la plante, qui peut varier considérablement au niveau intraspécifique. Ainsi, des espèces considérées majoritairement autogames peuvent présenter des taux d'allogamie considérables, allant jusqu'à 40 % chez le sorgho (Barnaud et al. 2008), ou encore 70 % chez le haricot commun (Wells et al. 1988). Même chez des espèces à reproduction végétative, des évènements de recombinaison peuvent avoir lieu lors de rares évènements de reproduction sexuée. De telles observations ont notamment été faites sur le manioc (*Manihot esculenta*) dans diverses régions du monde (Elias et al. 2001; Sardos et al. 2008), sur l'igname (*Dioscorea spp.*) (Scarcelli et al. 2006) et suggérées sur le taro (*Colocasia esculenta*) (Caillon et al. 2006;

Sardos et al. 2012). L'effet de ces recombinaisons, mêmes si elles sont rares, peuvent avoir des conséquences importantes sur la diversification des espèces (Gepts 1990).

Sélection

La sélection humaine intervient ensuite pour retenir les individus qui participeront à la génération suivante. Zohary, reprenant les observations de Darwin, distingue deux types de sélection: la sélection « consciente », par laquelle l'agriculteur conserve intentionnellement les individus dont les caractéristiques l'intéressent, de la sélection « inconsciente », qui est en fait la sélection naturelle ayant lieu dans les conditions particulières du champ cultivé et qui est favorisée inconsciemment par l'agriculteur en sélectionnant les individus les plus vigoureux (Zohary 2004).

Selon les pratiques des agriculteurs, les nouveaux phénotypes générés par recombinaisons seront alors fixés ou au contraire contre-sélectionnés (sélection conservative). Les descriptions disponibles dans la littérature montrent que les pratiques de sélection des agriculteurs varient selon les espèces, et entre des communautés différentes pour une même espèce. Louette et Smale décrivent la sélection conservative du maïs par les agriculteurs Mexicains malgré les flux de gènes qui génèrent des variantes morphologiques (Louette and Smale 2000). Par ailleurs, toujours sur le maïs, des observations montrent que l'introgression des populations cultivées par la téosinte (*Zea diploperennis*) est exploitée volontairement par les agriculteurs pour l'amélioration variétale (Benz et al. 1990 cité dans Jarvis and Hodgkin 1999). Au Cameroun, les individus considérés comme hors-type sont activement contre-sélectionnés (Barnaud et al. 2007), Doggett & Majisu avaient également observé que les hybrides de sorgho sauvages et cultivés étaient enlevés des champs, mais les agricultrices identifiaient malgré tout un tiers des descendants de tels croisements comme du sorgho cultivé, apte à être planté (cité dans Jarvis and Hodgkin 1999). Dans le cas des plantes à reproduction végétative, de tels évènements de recombinaison sont souvent prisés par les agriculteurs. Quiros et al. (1992) montrent par exemple que les agriculteurs andins utilisent les graines de pomme-de-terre pour créer de nouveaux cultivars. La sélection d'individus issus de reproduction sexuée a été documentée pour l'igname au Bénin (Chaïr et al. 2010), et pour le manioc en Amazone (Pujol et al. 2005) où une sélection est faite en faveur des hétérozygotes.

Boster suggère que les variétés paysannes de manioc sont sélectionnées par les agriculteurs pour la combinaison de traits morphologiques qui permet de les distinguer (Boster 1985).

Bien que ne présentant pas d'autre utilité que celle de distinguer les variétés entre elles, ces caractères morphologiques sont maintenus et sélectionnés par les agriculteurs. Ils servent de base à la classification paysanne au sein de laquelle une combinaison de traits est associée à un nom (Gibson 2009). De ce fait, les individus ambigus, présentant les attributs de plusieurs variétés, sont couramment contre-sélectionnés afin de maintenir un type conforme (Barnaud et al. 2007 ; Louette and Smale 2000). Des individus génétiquement différents mais non différenciés au niveau phénotypiques ne sont pas distingués par les agriculteurs (Gibson 2009). Cette sélection pour la « distinction perceptuelle » expliquerait notamment pourquoi les plantes cultivées exhibent une telle diversité de forme et de couleurs par rapport aux espèces sauvages, comme le relevait déjà Darwin :

"*When we look to the individuals of the same variety or sub-variety of our older cultivated plants and animals, one of the first points which strikes us, is, that they generally differ much more from each other, than do the individuals of any one species or variety in a state of nature*." (Darwin 1859)

Dérive

Les pratiques de sélection des agriculteurs peuvent également provoquer la dérive génétique des plantes qu'ils cultivent (Hodgkin et al. 2007). En effet, lorsque le nombre d'individus sélectionnés pour constituer la génération suivante est restreint ou lorsque la diversité génétique au sein de ce sous-ensemble est faible, des allèles rares peuvent être perdus et la diversité génétique sur l'ensemble du génome diminuée. Il semblerait que dans la plupart des systèmes de culture, ce phénomène de dérive n'ait pas d'effet considérable (Hodgkin et al. 2007), sauf dans certains systèmes où les populations cultivées sont de très petites tailles, comme dans la région de Cuzalpa au Mexique (Louette et al 2005 cité dans Hodgkin et al. 2007). La disposition des champs cultivés, l'arrangement des plantes au sein de ces champs ainsi que le système reproductif des espèces considérées sont également des facteurs influant sur la force de la dérive génétique.

Migration via les échanges de semences

Enfin, l'homme est également le principal acteur des forces de migration appliquées aux populations de plantes cultivées au travers des échanges de semences. La part du système semencier « formel », c'est-à-dire commercial, est généralement réduite dans les systèmes familiaux. L'autoproduction des semences et les échanges à courte distance sont les principales sources de lots de semences pour les agriculteurs dans ces systèmes, mais certaines études rapportent néanmoins de rares échanges sur de grandes distances (Leclerc and Coppens d'Eeckenbrugge 2012 ; Hodgkin et al. 2007).

Les pratiques individuelles des agriculteurs sur les populations de plantes cultivées ont été documentées pour de nombreuses espèces et dans différentes communautés. Cependant, les observations rapportées dans ces études ethnobotaniques ont généralement été faites auprès d'un nombre restreint d'agriculteurs et se sont focalisées sur leurs actions et motivations individuelles (Leclerc and Coppens d'Eeckenbrugge 2012). Bien qu'une part de variation individuelle existe, nous allons voir que l'effet des sociétés sur l'évolution des plantes ne se résume pas à une somme de pratiques individuelles indépendantes. En effet, au sein de ces sociétés, les individus sont insérés dans un réseau social complexe qui guide les échanges d'information, conditionnant ainsi l'apprentissage des pratiques, ainsi que les échanges de semences, qui influencent la diffusion des plantes.

III- De l'individu à la société

Une image généralement simpliste des sociétés rurales est parfois véhiculée sous le terme de « communauté », correspondant à celle d'un agrégat d'individus échangeant savoir et biens de façon aléatoire, ignorant l'existence d'une organisation sociale interne (Leclerc and Coppens d'Eeckenbrugge 2012). Les pratiques des agriculteurs dans ce contexte sont parfois décrites d'un point de vue purement utilitariste, envisagées sous le simple objectif de maximiser le profit face aux contraintes de l'environnement (physique, biotique, économique). La réalité est bien plus complexe, et nécessite le recours à l 'anthropologie sociale et culturelle, disciplines qui visent à comprendre comment sont organisées les relations sociales entre individus et quel est leur impact sur la façon dont ces derniers perçoivent, raisonnent et agissent sur leur environnement (Figure 1).

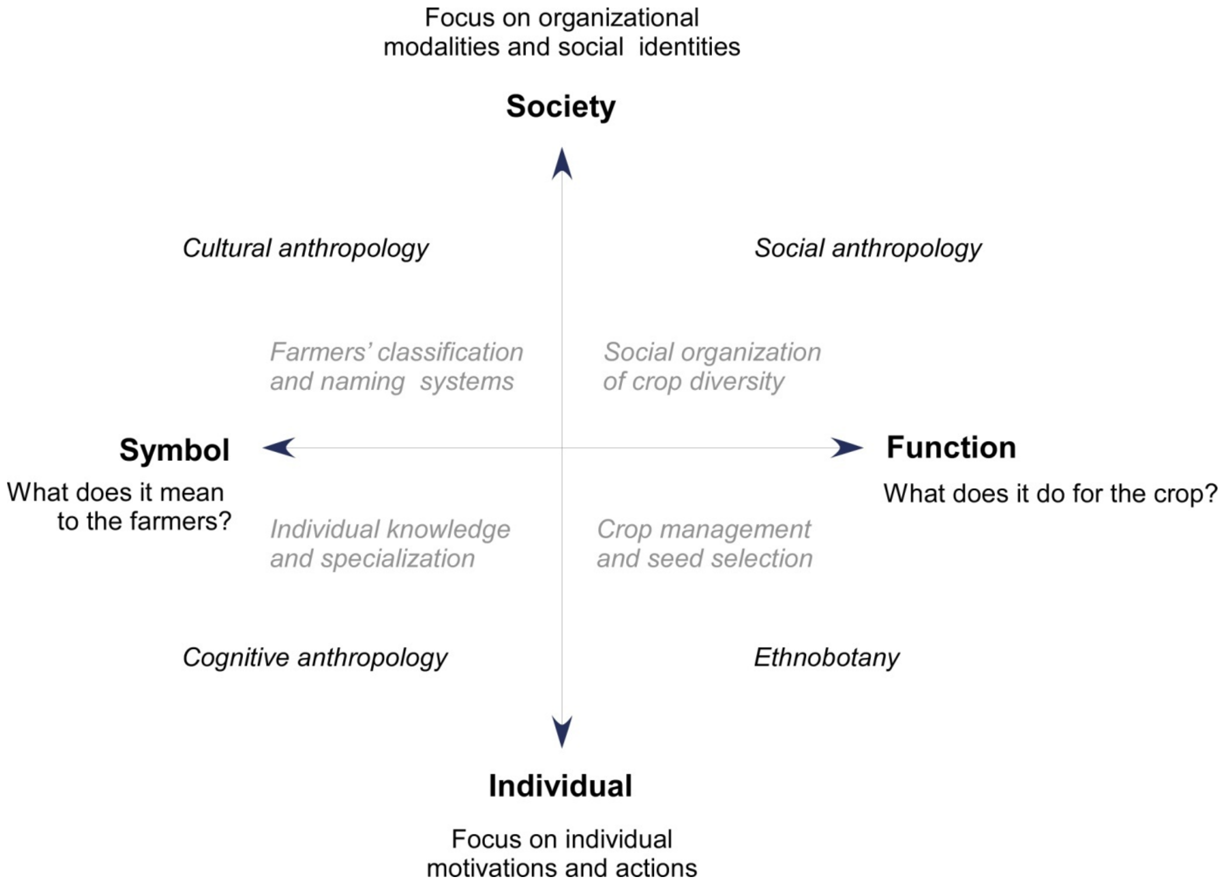


Figure 1. Diagramme représentant les principaux domaines de l'anthropologie applicables à l'étude de l'action anthropique sur la diversité des plantes cultivées. Des exemples de thèmes de recherche sont présentés en gris, selon qu'ils se focalisent sur une analyse fonctionnelle ou symbolique, et qu'ils adoptent des approches centrées sur les individus ou sur les sociétés. D'après Leclerc and Coppens d'Eeckenbrugge 2012.

Association entre l'organisation spatiale des plantes et celle des sociétés humaines

La relation entre l'organisation spatiale de la diversité génétique et phénotypique des plantes cultivées et celle des sociétés humaines a été observée dans un petit nombre d'études. La coïncidence entre la distribution des races botaniques de sorgho et celle des familles de langues en Afrique, relevée par Stemler et Harlan (Stemler et al. 1975 ; Harlan and Stemler 1976), est la première en date. Une étude ultérieure menée au Niger, creuset de diversité culturelle, a mis en évidence une relation entre la distribution géographique des groupes ethniques et la structure de la diversité génétiques du sorgho (Deu et al. 2008). Cependant, à l'échelle continentale comme à l'échelle nationale, l'interaction entre la distance géographique séparant les populations de sorgho et la distribution spatiale des groupes ethniques ou linguistiques ne permet pas de conclure définitivement à une relation causale.

Des observations similaires ont été faites pour le maïs en Amérique, malgré la différence de régime de reproduction entre les deux espèces. Une relation entre les profils de diversité génétique des populations de maïs cultivées et les groupes ethniques les cultivant est soulignée par Doebley (1990). Les études ultérieures, à l'échelle régionale, ont montré que les deux principaux groupes culturels rencontrés au Chiapas (Mexique), les Mestizos de langue espagnole et les groupes indigènes de langue Maya, présentent des différences significatives concernant les types morphologiques et raciaux de maïs qu'ils cultivent, et que ces différences sont maintenues dans toutes les zones agro-écologiques où les deux groupes coexistent (Brush and Perales 2007). Les mêmes auteurs ont également montré à plus petite échelle, toujours en Chiapas, que deux groupes ethnolinguistic distincts, mais adjacents et résidant dans la même zone agro-écologique, maintenaient des populations de maïs différenciées phénotypiquement (Perales et al. 2005). La force de la démonstration de ces dernières études réside sans doute dans le contrôle des conditions environnementales et le fait de minimiser le biais que la distance géographique entre groupes ethniques pourrait avoir dans leurs analyses. Enfin, d'autres études ont suggéré des relations similaires pour les plantes à multiplication végétatives (Sardos et al. 2012), ou encore pour les animaux (Berthouly et al. 2009). Ces diverses observations suggèrent donc une relation entre la diversité culturelle des sociétés et la diversité génétique ou phénotypique des populations de plantes, ou d'animaux, domestiques.

Réseaux d'échanges de semences et organisation sociale

La dépendance des réseaux d'échanges de semences vis-à-vis de l'organisation sociale des sociétés est un premier mécanisme pouvant expliquer la relation entre la distribution géographique de la diversité génétique des plantes cultivées et celle des sociétés. Peu d'études ont porté spécifiquement sur cette question, cependant il ressort de la revue de littérature effectuée par Leclerc et Coppens d'Eeckenbrugge (2012) que la majeure partie des semences utilisées dans les systèmes agricoles familiaux sont obtenues via les réseaux sociaux lorsqu'elles ne sont pas autoproduites. En effet, la plupart des études rapportent une forte part d'autoproduction (entre 58 et 99 % selon les études considérées), le reste des semences étant obtenues auprès de membres de la famille, de personnes du même village ou de la même communauté, et lorsque plusieurs groupes ethniques cohabitent, les semences sont préférentiellement obtenues auprès de membres du même groupe. La proportion de semences obtenues de personnes inconnues est donc très faible. En effet, la confiance est un élément essentiel dans les échanges de semences, comme l'a montré Badstue et al. (2002) car un simple examen visuel des semences ne suffit pas à déterminer leur qualité, et des semences de mauvaise qualité peuvent compromettre la récolte.

Cependant, David et Sperling (1999) rapportent un recours important aux marchés locaux en Afrique de l'Est, attribué à une évolution des réseaux semenciers vers l'acquisition de semences. En effet, bien que les agriculteurs considèrent les semences données par des proches comme de meilleures qualités que celles acquises sur le marché, les trop faibles quantités disponibles et le sentiment de dépendance pousseraient les agriculteurs à se tourner vers les marchés locaux.

La part culturelle de la sélection paysanne

Une seconde clé d'explication pour comprendre l'influence des sociétés sur l'organisation de la diversité des plantes cultivées se trouve potentiellement dans la différenciation de leurs pratiques de sélection.

De la perception aux pratiques

La façon dont les individus perçoivent et se représentent leur environnement influence profondément leurs pratiques à son endroit. Cette relation a notamment été décrite par Atran et Medin (2008) à partir de comparaisons interculturelles entre sociétés partageant le même environnement et impliquées dans les mêmes activités. Ils montrent notamment que trois groupes culturels distincts, les natifs Itza' Maya, les immigrants Q'eqchi' Maya et les immigrants Ladinos, bien que vivant au sein du même environnement présentent des différences considérables dans la façon dont ils le perçoivent et se le représentent. Ces différences se répercutent sur la façon dont ils gèrent leur environnement, les migrants Q'eqchi' et Ladinos ayant des pratiques agro-forestière bien moins durables que celles des Itza'.

Des études interculturelles menées sur le maïs au Mexique suggèrent que la façon dont les agriculteurs se représentent la variabilité phénotypique des plantes qu'ils cultivent détermine leurs pratiques de sélection, et que celles-ci peuvent diverger considérablement entre communautés adjacentes (Perales et al. 2005; Pressoir and Berthaud 2004). La sélection des semences, en tant que pratique, dépend de la façon dont les agriculteurs perçoivent et se représente la diversité des plantes cultivées. Ce système de représentation est reflété dans les systèmes de classification, ou taxonomies, auxquels ont recours les agriculteurs.

Trois mécanismes distincts bien qu'interdépendants interviennent dans le processus de classification (Friedberg 1991):

- l'identification : l'agriculteur perçoit les caractéristiques de la plante accessibles à ses sens, principalement visuels
- la nomination : il associe un nom à cet ensemble de caractéristiques
- la classification : il regroupe les plantes dans des classes en fonction de leurs caractéristiques communes de façon hiérarchique, comme le font les taxonomies scientifiques

L'agriculteur définit ainsi des variétés, distinguées par leurs caractéristiques perceptibles (principalement morphologiques), qu'il va chercher à maintenir pour conserver une association stable et sans équivoque entre une catégorie morphologique et un nom.

Cohérence des taxonomies locales

Les systèmes de taxonomie locaux ont fait l'objet de nombreuses études pour ce qui est de la classification des espèces (Atran 1985; Berlin 1973), mais très peu d'études ont porté sur le niveau intraspécifique. Cette absence de formalisation des taxonomies utilisées par les agriculteurs pour classer la variabilité de leurs plantes cultivées justifie que l'on se questionne sur la cohérence de ces systèmes (Quiros et al. 1990).

Les études menées sur des espèces à multiplication végétative (Salick et al. 1997 ; Quiros et al. 1990) ou à reproduction sexuée tels que le sorgho (Barnaud et al. 2007; Rabbi et al. 2010; Soler et al. 2013; Teshome et al. 1997), le maïs (Louette et al. 1997 ; Badstue et al. 2002) ou encore le riz (Nuijten and Almekinders 2008) ont testé la correspondance entre la taxonomie paysanne et la structure de la diversité génétique et phénotypique des populations de plantes cultivées. Les études qui ont porté sur le maïs ont montré des degrés de correspondance variables selon les études entre la taxonomie paysanne et la structure de la variabilité phénotypique (Louette et al. 1997 ; Badstue et al. 2002). Globalement, cependant, une assez bonne correspondance entre caractéristiques morphologiques et noms de variétés a été observée (Teshome et al. 1997 ; Louette et al. 1997), bien que cette correspondance semble dépendre de l'échelle considérée. Certaines études ont en effet suggéré un manque de concordance des noms utilisés par différents agriculteurs pour désigner une même catégorie morphologique (Salick et al. 1997; Nuijten and Almekinders 2008; Sadiki et al. 2007). Ces études rapportent que les agriculteurs appartenant à des villages différents, mais parfois aussi au même village, utilisent des noms différents pour désigner les mêmes catégories morphologiques.

Au niveau génétique, cette correspondance semble également variable. Quiros et al. (1990) rapportent une bonne correspondance entre les profils isozymes et les noms de variétés donnés aux cultivars de pomme-de-terre par les agriculteurs andins. Pour le sorgho, Barnaud et al. (2007) observent une correspondance variable entre les noms et la structure de la diversité génétique selon les variétés de sorgho nommées par les Duupa du Cameroun. Rabbi et al. (2010) décrivent deux situations différentes au Soudan et au Kenya, les variétés paysannes collectées au Kenya ne montrant pas de différenciation génétique tandis que celles collectées au Soudan représentaient des unités génétiques clairement distinctes. Enfin, Soler et al. (2013) observent également que les variétés paysannes cultivées dans deux villages du Nord-Cameroun sont des unités génétiques relativement distinctes.

Apprentissage et partage d'une taxonomie commune

La taxonomie est également un support pour organiser les informations relatives aux caractéristiques et usages de ces variétés, et joue ainsi un rôle central dans la transmission du savoir et l'apprentissage, ainsi que dans les échanges de semences. Comme le décrit Boster (1986), chaque agriculteur apprend des autres quelles sont les caractéristiques phénotypiques permettant d'identifier les variétés et quel est le nom approprié pour désigner chaque variété. C'est ce qui va lui permettre d'échanger des semences et de communiquer avec les membres de sa société. L'expérience propre des agriculteurs concernant la variabilité phénotypique des plantes est également essentielle, mais l'apprentissage, un processus éminemment social, a un rôle fondamental dans la façon dont l'agriculteur catégorise la variabilité, nomme et ordonne ces catégories.

L'apprentissage ne se fait pas de manière aléatoire, et selon les sociétés et les domaines considérés, les modalités de transmission du savoir ne sont pas les mêmes (Reyes-García et al. 2009 ; Cavalli-Sforza and Feldman 1981). Les principaux modes de transmission du savoir sont la transmission verticale, de parents à enfants, la transmission horizontale, entre personnes de la même cohorte d'âge, ou encore la transmission oblique, des personnes de la génération des parents, mais non apparentées à ces derniers, vers les enfants. Ces modalités vont avoir un impact sur la structuration du savoir au sein des sociétés. Une diffusion verticale étant plutôt conservative, elle aura tendance à maintenir une différenciation du savoir entre groupes familiaux tandis que la diffusion horizontale homogénéisera le savoir au sein de la société. Ces principaux modes de transmission présentent des variantes, par exemple la transmission entre affins[1] est courante des les sociétés patrilinéaires[2] et virilocales[3] (Herbich and Dietler 2008). C'est-à-dire que lorsqu'une jeune fille se marie et s'établit dans la famille de son mari, elle passe par un processus de « resocialisation » par lequel sa belle-mère et ses belles-sœurs vont lui apprendre à effectuer les activités quotidiennes telles qu'elles mêmes les pratiquent. Ce mode de transmission maintient également des différenciations entre groupes familiaux car il s'apparente en définitive à la transmission verticale.

Enfin, les réseaux de relations sociales et d'échanges de semences vont également jouer un rôle fondamental dans le partage d'une nomenclature et d'une taxonomie commune. Comme l'a montré Boster chez les Aguaruna d'Amazonie, les agricultrices ayant des liens de parenté proche échangent davantage de semences et nomment de façon similaire les différents clones de manioc. Les échanges de semences, qui se font généralement par les réseaux de relations sociales, favoriseraient le partage d'une expérience commune des caractéristiques des variétés par les agriculteurs. Les modalités d'apprentissage et de diffusion des taxonomies variétales joueraient ainsi un rôle central dans le maintien des différences de perception et de représentation entre sociétés, et parfois au sein des sociétés.

---

[1] Affins: parents par alliance (Larousse)

[2] Patrilinéaires: se dit d'un mode de filiation pour lequel seule compte la parenté paternelle

[3] Virilocales : se dit du mode de résidence de jeunes époux qui doivent résider dans le village des parents du mari

Pour une approche emic des pratiques des agriculteurs et la prise en compte de leur diversité culturelle

La synthèse des études concernant l'impact des pratiques des agriculteurs sur les populations de plantes cultivées illustre la complexité de cette relation. Il s'agit d'un domaine qui reste largement méconnu, malgré les enjeux qu'il représente pour la conservation des ressources génétiques. Deux principaux manques sont identifiables à l'issu de cette revue de littérature concernant le rôle des facteurs anthropiques dans la domestication des plantes cultivées au sens large du terme.

Premièrement, l'approche des pratiques s'est faite au niveau individuel, sans considérer la place de l'individu dans la société. Il en résulte une absence de prise en compte de la diversité culturelle des agriculteurs dans ces études, alors que le savoir n'est pas distribué de façon homogène dans les sociétés comme nous l'avons souligné précédemment. La mise en évidence de ces différences culturelles nécessite l'usage d'approches probabilistes utilisées en anthropologie culturelle, plutôt que de se baser sur des observations ponctuelles limitées à un petit nombre d'individus (Romney et al. 1986).

Deuxièmement, la grande majorité de ces études ont adopté une approche etic, se référant à un cadre conceptuel et un système de représentation externe, au lieu de chercher à mettre en évidence la logique et la cohérence interne des systèmes de représentation et leur rapport avec les systèmes de relations. F. Grenand (2002) et M.A. Martin (1975) estiment qu'une approche emic, privilégiant le point de vue de l'agriculteur et les relations qu'ils entretiennent entre eux, serait plus adaptée pour comprendre la logique sous-jacente à ces taxonomies.

Ces lacunes sont révélatrices du manque d'implication des divers champs disciplinaires des sciences sociales dans l'étude des plantes cultivées. Le champ de l'ethnobotanique a été fréquemment mis à contribution, contrairement à ceux de l'anthropologie sociale et culturelle qui pourtant recèlent des approches pertinentes pour comprendre la relation entre les sociétés et leurs plantes.

IV- Objectif de la thèse

L'objectif général de cette thèse est d'améliorer notre compréhension des mécanismes par lesquels l'organisation des sociétés humaines influence l'organisation de la diversité génétique et phénotypique des plantes cultivées *in situ*. Elle porte donc sur les deux pratiques majeures par lesquelles les sociétés humaines influent sur l'évolution des plantes cultivées : les pratiques de sélection et d'échanges des semences (Figure 2). Si les pratiques de sélection sont influencées par la façon dont les agriculteurs perçoivent et se représentent la diversité intraspécifique, nous devons par conséquent reconnaitre un rôle central de la culture, propre à chaque société, dans l'identification, la dénomination et la classification des variétés. Ces trois mécanismes seront par la suite désignés sous l'appellation unique de taxonomie paysanne.

Notre étude a été menée sur le versant est du mont Kenya, dans les groupes Chuka, Tharaka et Mbeere. Bien qu'étant adjacents spatialement, ces trois groupes sont culturellement et linguistiquement différenciés et possèdent une organisation sociale propre. On peut à ce titre considérer qu'ils constituent des sociétés distinctes[4]. Ces groupes évoluent dans un même environnement agro-écologique et économique et sont impliqués dans les mêmes activités, principalement l'agriculture. Le sorgho (*Sorghum bicolor* [L.] Moench) est la principale céréale cultivée par les trois groupes dans cet environnement semi-aride.

Ce contexte de terrain permet ainsi d'effectuer une comparaison interculturelle des représentations et des pratiques des agriculteurs concernant la diversité variétale du sorgho. En contrôlant la variabilité environnementale (E), il est ainsi possible d'étudier uniquement l'interaction entre les sociétés (S) et les plantes qu'elles cultivent (G), suivant le modèle G x E x S proposé par Leclerc et Coppens d'Eeckenbrugge (2012).

La première section de cette thèse décrit la relation entre l'organisation spatiale de la diversité du sorgho et l'organisation des sociétés Chuka, Tharaka et Mbeere. Deux niveaux

---

[4] Société : ensemble d'êtres humains vivant en groupe organisé. Milieu humain dans lequel quelqu'un vit, caractérisé par ses institutions, ses lois, ses règles (Larousse)

Culture: ensemble des phénomènes matériels et idéologiques qui caractérisent un groupe ethnique ou une nation, une civilisation, par opposition à un autre groupe ou à une autre nation. (Larousse) hiérarchiques d'organisation sociale seront considérés : le groupe de voisinage (Chapitre I), qui est le plus petit niveau d'organisation après le foyer (Lambert 1947 cité dans Middleton 1953), et le groupe ethnolinguistique (Chapitre II). Dans cette section, nous confronterons deux approches de la diversité variétale: celle des agriculteurs (point de vue emic) en décrivant la distribution des variétés telles qu'ils les nomment, et celle des biologistes (point de vue etic) en décrivant la distribution de la variabilité génétique neutre et de la variabilité phénotypique.

Dans la seconde section, nous aborderons les mécanismes impliqués dans la relation entre la distribution des variétés nommées, celle de la variabilité biologique, et l'organisation des sociétés :

Dans un premier temps, nous caractériserons le lien entre la taxonomie paysanne et la variabilité génétique et phénotypique du sorgho. Pour cela, nous testerons d'abord si les agriculteurs appartenant à une même société identifient, nomment et classent les variétés de manière similaire au regard de leurs caractéristiques biologiques (Chapitre III). Nous testerons ensuite si les taxonomies paysannes diffèrent selon les sociétés et leur culture (Chapitre IV).

Dans un second temps, nous analyserons la dépendance des échanges de semences vis-à-vis de l'organisation des sociétés (Chapitre V). Ce dernier chapitre nous permettra de faire le lien entre l'organisation sociale des agriculteurs, la façon dont ils perçoivent et se représentent la diversité du sorgho, et la façon dont ils agissent sur elle.

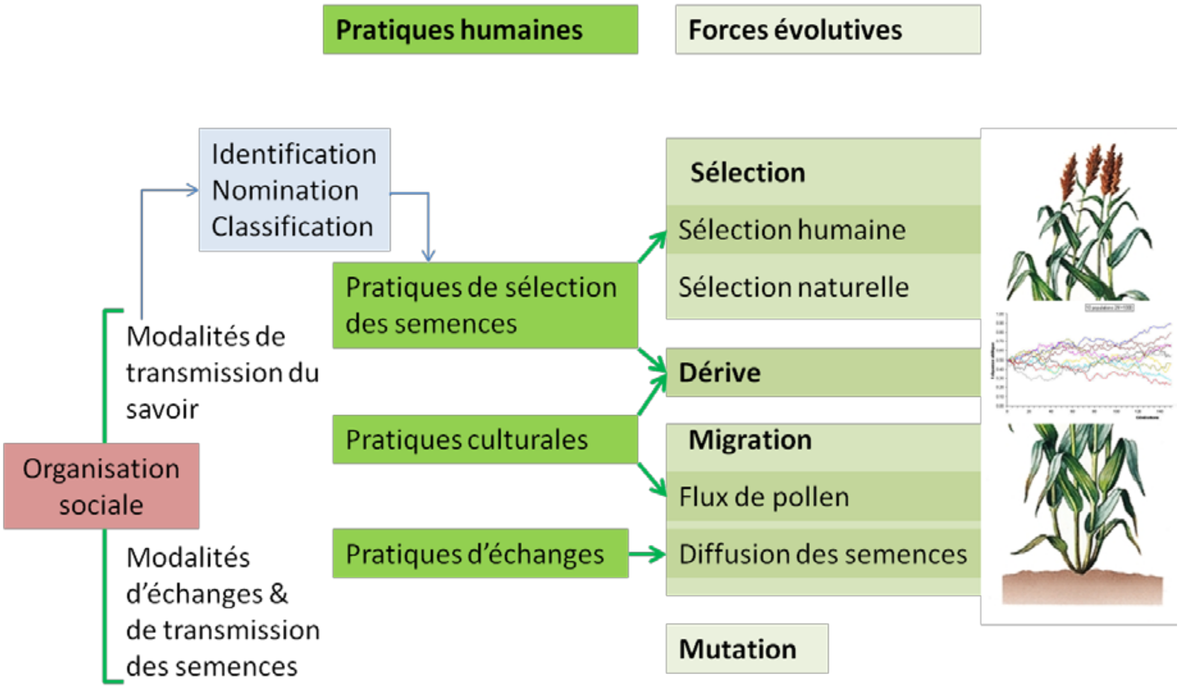


Figure 2. Schéma récapitulatif de l'influence des sociétés humaines sur les plantes cultivées. Trois des quatre principales forces évolutives qui s'appliquent sur les populations de plantes cultivées sont influencées par les pratiques humaines. Les pratiques de sélection des semences déterminent la force et la direction de la sélection, mais elles peuvent aussi provoquer la dérive génétique de ces populations. Les pratiques culturales, notamment l'arrangement spatial des différentes espèces et variétés, jouent également sur les forces de dérive, et déterminent aussi l'intensité des flux de pollen entre plantes principalement à l'échelle de la parcelle. Enfin, les pratiques d'échanges de semences jouent sur la dispersion des semences, entrainant des flux de gènes à plus grande échelle.

Les pratiques individuelles de **sélection** et les pratiques culturales seraient largement déterminées culturellement. En effet, la diversité des pratiques individuelles serait modelée par les modalités d'échanges de savoir, tendant soit à maintenir des différenciations, soit à uniformiser le savoir et les pratiques (Reyes-García et al. 2009; Cavalli-Sforza and Feldman 1981). La façon dont la diversité est perçue, c'est-à-dire identifiée, et classée par les agriculteurs serait impliquée dans le processus de sélection paysanne. Or, ces mécanismes cognitifs seraient acquis par apprentissage, et seraient donc culturellement déterminés (Chapitre III et IV de la thèse).

L'échange de semences repose sur la relation de confiance entre le donneur et le receveur (Badstue et al. 2002). Les réseaux d'échanges de semences seraient donc largement dépendants des réseaux de relations sociales (Chapitre V de la thèse).

V- Contexte institutionnel et déroulement

Cette thèse a été réalisée dans le cadre du projet ARCAD (Agropolis Resource Centre for Crop Conservation, Adaptation and Diversity, http://www.arcad-project.org), dont l'objectif général est de développer une plateforme dédiée à la caractérisation et à la valorisation de l'agrobiodiversité des plantes méditerranéennes et tropicales. Elle s'inscrit plus particulièrement dans le sous projet « Céréales en Afrique : des espèces majeures aux espèces sous-utilisées » qui vise notamment à identifier les principaux facteurs évolutifs impliqués dans la structuration de la diversité. Cette étude a été faite en partenariat avec la Banque de Gènes du Kenya Celle-ci a contribué aux collectes de matériel végétal *in situ*, aux enquêtes anthropologiques et à l'essai *in situ* en milieu contrôlé. C'est également la Banque de Gènes qui a effectué la caractérisation morphologique des panicules de sorgho. Le transfert de matériel végétal depuis le Kenya vers la France s'est réalisé sous le couvert d'un accord signé entre le CIRAD et le Kenyan Agricultural Research Institute (KARI). Les analyses génétiques ont été réalisées sur le Grand Plateau Technique Régional de Génotypage (GPTR) localisé sur le campus Lavalette du CIRAD. La figure 3 résume les différentes étapes du processus de collecte, de caractérisation et d'analyse.

Cette thèse a bénéficié d'un financement de trois ans par une bourse de la Fondation Agropolis. Elle a été accueillie par le CIRAD, au sein de l'équipe Evaluation, gestion et valorisation des ressources génétiques (EGV) qui fait partie de l'unité mixte de recherche amélioration génétique et adaptation des plantes tropicales et méditerranéennes (UMR AGAP). Au total, six mois ont été passés sur le terrain, au Kenya.

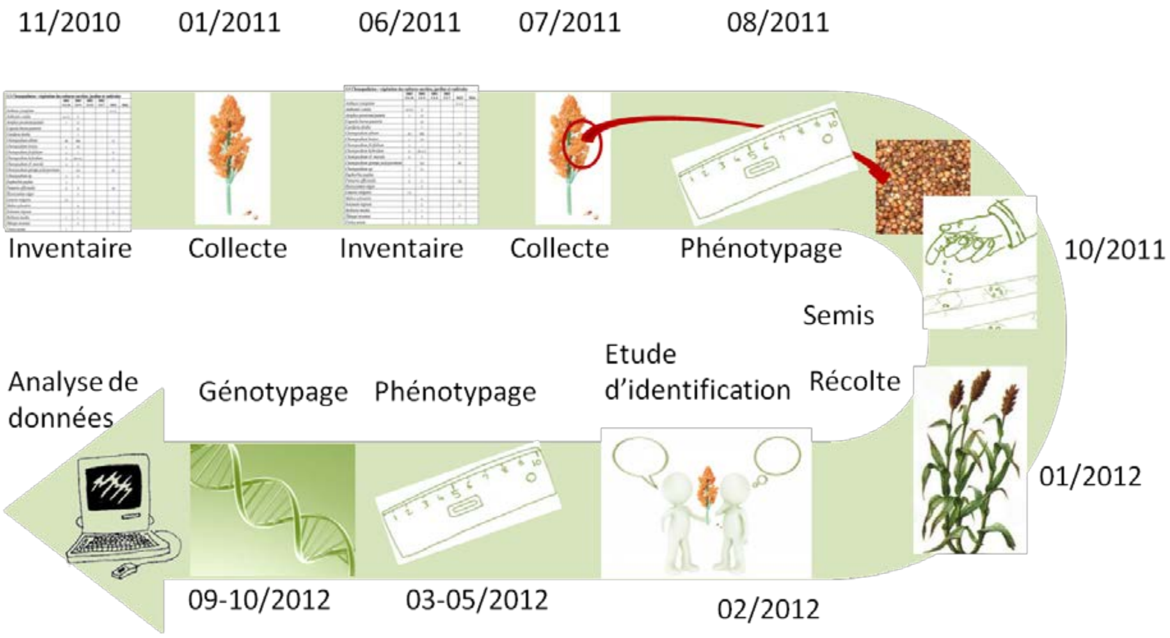


Figure 3. Planning du déroulement de la thèse

Tableau 1. Tableau récapitulatif de la thèse

Chapitres	Méthodes	Echantillon	Outils	Publications
De l’organisation des sociétés à l’organisation de la diversité du sorgho				
Chap. I : Echelle des groupes de voisinage (*ntora*)	Inventaires des variétés	95 foyers		Acceptée par *Agriculture and Human Values*
Chap. II : Echelle des groupes ethnolinguistique	Inventaires des variétés, collecte in situ, génotypage, description morphologique	130 foyers 297 plantes	18 SSR 15 descripteurs morphologiques	Soumise à *PLOS ONE*
Représentations et échanges, reflets de l’organisation des sociétés				
Chap. III : Cohérence des taxonomies locales à l’échelle de la société et différences interculturelles	Description morphologique Expérience d’identification	287 panicules 96 agricultrices	15 descripteurs morphologiques	Préparée pour *American Journal of Botany*
Chap. IV : Identification et nomination des variétés: la part de la culture	Expérience d’identification	293 panicules 96 agricultrices		Préparée pour *Philosophical Transactions of the Royal Society B-Biological Sciences*
Chap. V : Influence de l’organisation des sociétés sur les échanges	Enquêtes	218 foyers		Préparée pour *American Anthropologist*

TROIS SOCIETES, UN MEME ENVIRONNEMENT

I- Localisation et environnement agro-climatique

Notre zone d'étude se situe sur le versant est du mont Kenya, dans la Province est du Kenya. Cette étude s'est déroulée sur deux sites (Figure 4):

La principale zone d'étude (site 1: 0°24'27.88"S, 37°46'35.59"E) se situe sur la zone de contact entre trois groupes ethnolinguistiques : les Chuka, les Tharaka et les Mbeere, à la limite entre les comtés (Counties[5]) *Tharaka-Nithi* et *Embu*. Elle inclue les Locations[1] (unités administratives) de *Kawimbi*, *Kamaindi* et *Kithangani*, et couvre plus précisément les Sublocations[1] de *Kiaritha*, *Igamba Ng'ombe* and *Rianthiga* sur lesquelles 1255 foyers ont été recensés en 2009 (National Census 2009). Sur cette zone d'étude, l'altitude varie entre 810 to 946 m et le climat est semi-aride, les températures moyennes varient entre 21.7°C and 23.9°C au cours de l'année et la pluviométrie annuelle est en moyenne de 700 à 800 mm (Camberlin et al. 2012; Camberlin et al. 2009).

La seconde zone d'étude (site 2 : 0°19'16.72"S, 37°54'13.01"E) se situe dans le comté *Tharaka-Nithi,* sur les locations de *Chiakariga* et *Kamanyaki*. L'altitude varie entre 650 et 750 m sur ce site, la pluviométrie annuelle se situe aux alentours de 600 à 700 mm.

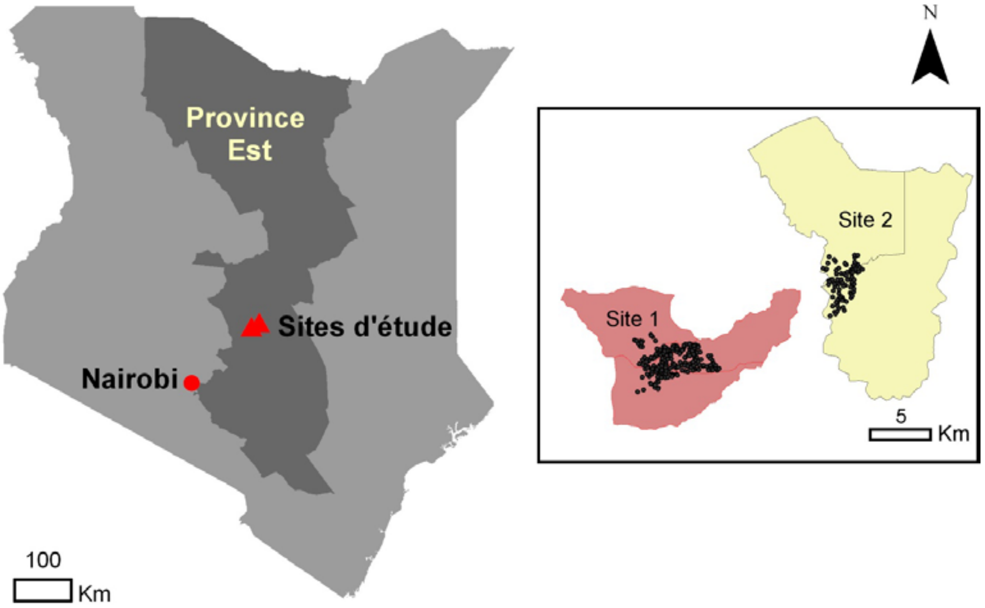


Figure 4. Localisation des sites d'étude

---

[5] Les Counties sont des unités administratives définies dans la nouvelle constitution du Kenya (2010). Ils sont subdivisés en Sub-counties, puis en Locations et Sub-locations (plus petite unité administrative)

Climat

Le régime des pluies est bimodal sur le versant est du mont Kenya, et les deux sites d'étude sont caractérisés par deux saisons humides, la première d’octobre à décembre et la seconde de mars à mai. La seconde saison des pluies est plus longue, mais moins fiable que la première (Camberlin et al. 2012; Camberlin et al. 2009). Ce régime pluviométrique permet donc aux agriculteurs d’effectuer deux saisons de culture. L'ethno-climatologie des agriculteurs diffère en distinguant non pas seulement deux, mais cinq saisons[6]:

Muratho: Mois de novembre à décembre ou janvier (selon les années). Période de forte pluies et saison de culture la plus fiable.
Kiathu: Mois de février à mars. Courte période sèche et chaude, considérée par les agriculteurs comme la plus difficile, notamment pour le bétail.
Nthanu: Mois d’avril à début juin, commence parfois dès mars. Courte période des pluies et seconde saison de culture, considérée comme la moins fiable.
Ndundu : Mois de juin et juillet, période fraiche et sèche, saison des récoltes.
Thano: Mois d’août à octobre, période sèche et ensoleillée.

Sols

Le mont Kenya est un vaste massif volcanique, culminant à 5 199 m, qui s’est formé avec l’apparition du rift est-africain il y a environ 3 millions d’années. Les sols des nos zones d’études sont donc issus de l’altération de ces roches volcaniques. Sur le site 1, il s’agit principalement de Ferrasols, de texture limono-sableuse et présentant un bon drainage ainsi qu’une fertilité modérée (Jaetzold et al. 2007). Des analyses de sols réalisées sur la zone d’étude montrent que la texture est peu variable, mais que la profondeur des sols s’accroit considérablement à proximité des rivières (données non présentées). Sur le site 2, il s’agit également de Ferrasols, mais davantage sableux et moins fertiles qu’au site 1.

---

[6] entretiens réalisés chez les Chuka sur le site 1

II- Les peuples du versant est du mont Kenya

Notre étude porte sur la comparaison interculturelle entre les groupes ethnolinguistiques Chuka, Tharaka et Mbeere[7]. La différenciation ethnique entre ces trois groupes, qui se base sur la croyance en une ascendance commune, est associée à des différences linguistiques, culturelles et historiques qui font l'objet de cette section.

Histoire

Les populations d'agriculteurs établies sur le versant est du mont Kenya sont locuteurs d'une langue bantoue. Combinant les données linguistiques et archéologiques, Holden (2002) estime que ces populations auraient atteint le Kenya il y a environ 1 800 ans, suite à leur expansion depuis une zone située entre le sud-est du Nigeria et l'ouest du Cameroun qui aurait eu lieu entre 3 000 et 5 000 ans avant notre époque (Holden 2002).

Deux principaux groupes sont distingués au sein des populations bantoues du versant est du mont Kenya : le groupe Meru et le groupe Embu. Cependant, des subdivisions ethniques basées sur des différences culturelles et linguistiques existent au sein de chacun de ces groupes. Des groupes de taille restreinte, comme les Chuka et les Tharaka ont été rattachés administrativement au groupe Meru qui était composé de sept sous-groupes seulement avant la colonisation: les Igembe, Tigania, Imenti, Miutini, Igoji, Mwimbi et Muthambi[8] (Middleton 1953). En dépit de ce regroupement d'origine coloniale, les Chuka et les Tharaka présentent encore aujourd'hui des différences linguistiques et culturelles notables au sein de l'ensemble Meru.

Les observations des administrateurs coloniaux, puis des anthropologues qui leur succédèrent sur la zone, suggèrent que les Meru n'ont jamais fonctionné comme une unité politique

---

[7] Dans ce document, on désignera ces groupes sous l'appellation « groupes ethnolinguistiques » ou « groupes ethniques ». Ethnie vient du latin *ethnos*, désignant des « gens de même origine ». Max Weber définit les groupes ethniques comme des groupes humains partageant la croyance en une ascendance commune.

[8] Selon les auteurs, la définition des sous-groupes composant initialement le groupe Meru varie. Lambert (1956) suggère que les Meru au sens strict seraient les Tigania, les Igembe et les Imenti.

cohérente. Les guerres tribales étaient fréquentes dans le passé, donnant lieu à des raids visant à dérober du bétail aux groupes adjacents. Des chaines d'alliances, basées sur les liens de parenté biologiques ou rituels, existaient entre groupes. Ce système d'alliance est désigné sous le nom de *Gishiaro* (naissance), les groupes ainsi alliés se considèrent comme frères de sang, soit parce qu'ils descendent d'un ancêtre commun, soit parce qu'ils ont effectué un rituel d'adoption. Les Chuka et les Tharaka étaient ainsi traditionnellement alliés (Fadiman 1993). Ces liens persistent encore fortement aujourd'hui : « *Chuka and Tharaka are one* » affirmaient ainsi couramment nos informateurs Chuka et Tharaka.

Les Mbeere, en revanche, seraient un sous-groupe issu des Embu (Chesaina 1997; Lambert 1956; Mwaniki 1973). Ils ont entretenu des relations conflictuelles avec les groupes Meru par le passé, en particulier avec les Chuka qui étaient leurs voisins directs (Glazier 1970; Mwaniki 1973). Une certaine méfiance persiste encore aujourd'hui, exacerbée par les conflits fonciers.

L'ensemble des populations du versant est du mont Kenya parle des langues appartenant à la famille Kamba-Kikuyu[9] (Moehlig et al. 1980). Le *gichuka* et le *kitharaka* sont des langues distinctes, tandis que le *kimbeere* est considéré comme un dialecte de la langue *kiembu* (le préfix ki-/gi-, rajouté devant le nom du groupe ethnique, désigne leur langage). La proximité entre ces trois langues est cependant forte, et elles sont inter-intelligibles.

L'histoire des peuples du mont Kenya est complexe et reste largement méconnue. Les groupes Meru auraient une origine différente des groupes Embu-Mbeere, mais l'origine des groupes Chuka et Tharaka est sujet à controverse. L'histoire commune des différents groupes *Meru*, excepté les Chuka, est décrite dans leur tradition orale par un mythe commun : le mythe de *Mbwa*. Ce mythe est l'unique information dont nous disposons à l'heure actuelle concernant l'histoire du peuplement. Il convient donc de rester prudents quant à sa fiabilité, d'autant que les données linguistiques ne semblent pas soutenir ce scénario (Bernard 1972). Selon ce mythe, les *Meru* seraient originaires de la zone côtière qu'ils auraient fuit en raison de raids effectués par des groupes d'origines Arabe (les *Nguo ntuni* – « habits rouges ») visant à capturer des esclaves. Ils auraient alors migré le long de la rivière *Tana* jusqu'aux pentes du Mont Kenya où ils se seraient divisés, chacun des groupes occupant les interfluves (Fadiman

---

[9] Classification: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Central, E, Kikuyu-Kamba

1993). Leur installation dans cette zone se serait faite aux alentours de 1700 - 1750, bien que les indices permettant d'établir cette datation soient ténus (Bernard 1972).

L'absence de ce mythe chez les Chuka suggérerait, selon certains auteurs, une origine ancienne de ce groupe dans la zone. En effet, les traditions orales mentionnent l'existence auparavant de peuples de chasseurs cueilleurs sur la zone, les *Gumba* (Middleton 1953), dont les Chuka pourraient être issus. Cependant, les preuves d'une telle origine sont une fois de plus trop faibles pour que l'on puisse réellement y faire confiance (Bernard 1972). L'origine des Tharaka est également mystérieuse. Plusieurs hypothèses ont été évoquées, Lambert suggère qu'il pourrait s'agir d'un groupe d'individus « déviants », exclus, provenant de différents sous-groupes Meru (Lambert 1950 cité dans Middleton 1953), hypothèse reprise par Fadiman plus tard (1993). Cependant peu de faits permettent de confirmer cette hypothèse.

Les Mbeere seraient issus du groupe Embu. Il n'existe pas réellement de consensus concernant leur histoire. Leur histoire orale suggère une origine commune au nord du Kenya, dans une zone nommée *Tuku* ou *Uru,* potentiellement située en Ethiopie. Ils auraient migré ensembles sous la forme d'un groupe unique avant de se séparer suite à un conflit. Ils se seraient installés dans les zones qu'ils occupent actuellement aux alentours du 16 ème ou 17 ème siècle (Mwaniki 1973). Le manque de preuves, ici encore, laisse une large part de doutes sur la fiabilité de ce scénario.

Organisation sociale: *ntora, mwiriga et nthuke*

L'organisation sociale des trois groupes présente de nombreuses similarités. Le foyer ou *mucii*, et composé généralement de l'homme, de sa ou ses épouses et de leurs enfants. Les jeunes hommes mariés et leurs familles sont également établis au sein ou a proximité de l'enclos familial. La polygamie est cependant en perte de vitesse depuis la christianisation et la plupart des hommes ont une seule épouse. Lorsqu'un homme a plusieurs épouses, chacune d'entre elle possède sa propre maison, son propre champ et son propre grenier mais l'entraide entre coépouses est la règle pour les travaux quotidiens (Middleton 1953).

Les foyers sont spatialement dispersés (Figure 5), mais ils sont organisés territorialement en *ntora*[10] ou *ituura*[11], unité territoriale nommée comptant 40 à 100 foyers et délimitée par une frontière. Le terme *ntora* est traduit généralement comme « village » ou « groupe de voisinage », et constitue l'unité politique de base (Glazier 1970; Middleton 1953). Les *ntora* possédaient auparavant une aire de battage commune, les rituels s'organisaient également à cette échelle (Peatrik 1999). Les relations entre voisins restent aujourd'hui encore privilégiées, notamment pour l'entraide concernant les travaux des champs (Linsig 2009).

Figure 5. Photo satellite du site d'étude 2. Les foyers sont délimités par des enclos circulaires ou rectangulaires, et les parcelles sont délimitées par des haies. L'habitat est dispersé.

Chaque *ntora* regroupe des foyers appartenant à plusieurs clans. Le clan, patrilinéaire, joue un rôle central dans l'organisation sociale des communautés. Le clan est désigné par le terme *mwiriga* [ch] or *muviriga* [mb], qui signifie « subdivision ». On recense un grand nombre de termes désignant les clans dans les trois groupes ethniques, cependant, ces termes semblent

---

[10] En kichuka [ch] et kitahraka [th]

[11] Kimbeere [mb]

désigner aussi bien des catégories claniques que des sous-clans, le terme *mwiriga* s'appliquant à tous les niveaux. De plus, certains clans sont désignés sous plusieurs noms synonymes. Chez les Chuka et les Tharaka, les clans sont à la base d'un système d'alliances: certains clans se réclament d'origine commune et considèrent qu'ils sont « frères de sang ». Cette ascendance commune peut être réelle, mais elle peut aussi provenir d'un rituel conclu dans le passé afin de sceller la fraternité entre les deux clans. Ce lien est désigné sous le terme de *gishiaro*, faisant référence à la naissance commune. Les membres de clans *gishiaro* ont l'interdiction de se marier et d'échanger des semences et du bétail. Chez les Mbeere, le système d'alliance ne semble pas exister (Middleton 1953). En revanche, Glazier décrit la répartition des clans Mbeere selon deux moitiés, les mariages se concluent entre personnes appartenant à des moitiés différentes, et les alliances se forgeraient également en termes de moitiés (Glazier 1970). Le mariage entre personnes appartenant au même clan est interdit dans les trois groupes.

Un troisième niveau d'organisation reposait sur la stratification horizontale de la société en classes d'âges. Les classes d'âges constituaient un élément central de l'organisation politique des Chuka, Tharaka et Mbeere, bien que des différences aient existé entre ces groupes. Cependant, l'ère coloniale a eu raison des classes d'âges, qui ont semble-t-il cessé d'être nommées dans les années 1940 en zone Mbeere (Glazier 1970), et vraisemblablement à la même époque chez les Chuka et les Tharaka. L'initiation des jeunes garçons et jeunes filles avait lieu périodiquement, cérémonie au cours de laquelle ils étaient circoncis et excisées. Les classes d'initiations successives étaient regroupées en classes d'âges ou *nthuki* chez les Meru. Le nombre de groupes d'initiation formant une *nthuki* variait entre les groupes, ils étaient recrutés sur une durée variable, se situant autour de 12 ans. Chaque classe d'âge avait un rôle distinct dans la société, les individus parcouraient ainsi des échelons au cours de leur vie, occupant des rôles différents dans la vie politique. Peu d'informations sont disponibles sur ce système dans les groupes Chuka, Tharaka et Mbeere, mais Anne-Marie Peatrik décrit spécifiquement ce systèmes dans les groupes Tigania et Igembe (Peatrik 1999).

Territoires et économies

Du point de vue territorial, les groupes Chuka, Tharaka et Mbeere sont voisins. Les Tharaka se trouvent principalement dans les plaines semi-arides, en dessous de 900 m, tandis que les Chuka occupent principalement les zones à café et à thé, plus humides et situées en altitude entre 900 et 1500 m. Les Mbeere sont principalement situés dans la zone d'altitude entre 900 et 1200 m, bien que leur territoire s'étend jusqu'à 1500 m d'altitude.

La zone principalement étudiée dans cette thèse (site 1) aurait été peuplée assez récemment d'après les témoignages de nos informateurs les plus âgés. Les premières familles Chuka se seraient installées dans la zone aux alentours des années 1900. Ces petits groupes commencèrent à défricher la forêt et à cultiver. A cette époque, les familles exploitaient le gradient altitudinal et possédaient des parcelles à différentes altitudes (Figure 6). Bananes et tubercules étaient cultivés en altitude tandis que les céréales et les légumineuses étaient cultivées à plus basse altitude. Les groupes familiaux étaient donc mobiles le long du gradient altitudinal et il était courant d'installer un parent sur les parcelles les plus éloignées (Bernard 1972). Les Tharaka seraient arrivés plus tardivement, et Bernard décrit leur migration vers les altitudes plus élevées à partir des années 1950 en raison de sécheresses répétées (Bernard 1972). Les Mbeere étaient historiquement situés de l'autre coté de la rivière *Thuchi*, à environ 5 km du cœur de notre zone d'étude, et leur migration semble plus récente mais peu d'informations sont disponibles sur ce point.

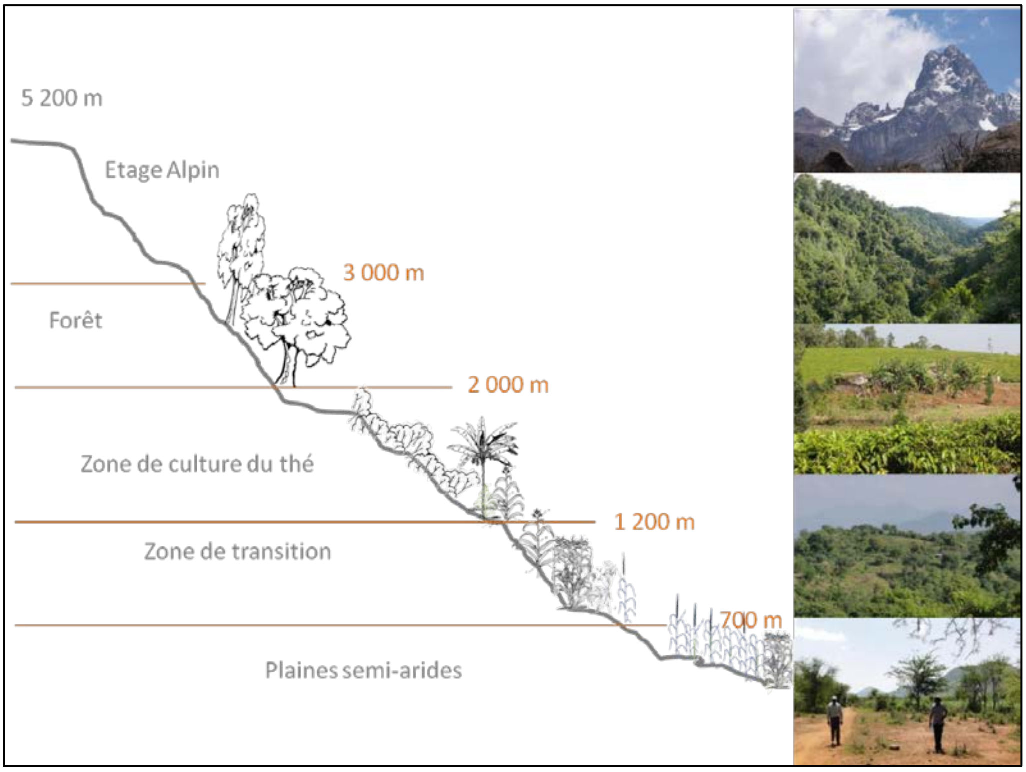


Figure 6. Gradient altitudinal et zones agro-écologiques sur le mont Kenya

Entre les années 1965 et 1970, un vaste programme de consolidation foncière a été mis en place dans le district Meru en raison d'une pression foncière élevée dans la zone d'altitude intermédiaire. L'objectif de cette réforme foncière était de désengorger la zone d'altitude intermédiaire, ou « homestead-zone » en délocalisant les foyers vers les altitudes supérieures ou inférieures. Du fait de cette réforme, les terres sont passées d'un régime de gestion communautaire à celui de propriétés individuelles avec la mise en place de titres de propriétés. Cette politique a entrainé une profonde modification du système traditionnel où la majorité des foyers étaient établis entre 1 200 et 1 600 m d'altitude et exploitaient le gradient altitudinal (Bernard 1972). Les zones de plus basse altitude ont ainsi été peuplées jusqu'à la limite du territoire Tharaka, et les familles ont été cantonnées à une seule zone agro-écologique.

Les trois groupes combinent agriculture et élevage dans des systèmes principalement vivriers à l'altitude de notre zone d'étude (900 m). La culture du tabac, et dans une bien moindre mesure du coton, est pratiquée par la plupart des foyers mais la vente de ces produits ainsi que des éventuels surplus de récolte ne constitue pas un revenu considérable. L'élevage des vaches, chèvres et moutons est généralisé, mais les troupeaux sont généralement de taille restreinte, excédant rarement une quinzaine d'animaux. Les animaux constituent une assurance en cas de sécheresse ou pour les périodes de soudures. Ils sont alors vendus pour acheter des céréales et des légumineuses. Le lait est consommé quotidiennement, mais la viande est réservée aux occasions. La production de charbon de bois s'est fortement développée dans la zone d'étude depuis quelques années et a conduit à l'éradication de la majorité des arbres. L'élevage occupe une place plus importante chez les Tharaka, situés à plus basse altitude dans les plaines semi-arides et peu fertiles (Glazier 1970; Bernard 1972). Les populations Chuka situées à plus haute altitude, dans la zone de la culture du café et du thé ont pour leur part basé leur économie sur les cultures de rentes.

Systèmes de culture

Les champs des trois groupes ethniques sont généralement de taille limitée, Bernard estime la taille moyenne des parcelles dans la zone du site 1 à moins d'1.4 Ha par foyer (Bernard 1972). Les champs abritent une grande diversité d'espèces cultivées. L'association entre céréales et légumineuses est à la base de ces systèmes. Le sorgho (Munya - *Sorghum bicolor*[12]) et le mil (mwere - *Pennisetum glaucum)* étaient les deux principales céréales cultivées sur la zone avant l'arrivée du maïs (mpempe - *Zea mays*), qui est en limite d'adaptation agro-écologique et connait de sérieux dégâts dus à la sécheresse (Figure 7). L'éleusine (ugiimbi - *Eleusine coracana*) et le petit mil (munyaki - *Setaria italica*) étaient également marginalement cultivés dans le passé, mais surtout en altitude. Une grande diversité de légumineuses est également rencontrée, parmi lesquelles le niébé (nthoroko - *Vigna unguiculata*), le dolichos (ncaabi – *Dolichos lablab*), le pois d'Angole (ncugu - *Cajanus cajun*) ainsi que le haricot commun (mung'au - *Phaseolus vulgaris*) et le haricot mungo (nkina - *Vigna radiata*), probablement introduits ultérieurement, sont les plus fréquentes. Les courges (kirenge - *Cucurbita pepo*), et la canne à sucre (kigwa – *Saccharum officinarum*) sont également cultivées à cette altitude.

Dans les plaines occupées par les Tharaka (site 2), les parcelles sont visiblement plus grandes mais nous n'avons pas d'estimation de leur taille. Le sorgho et le mil sont largement majoritaires dans les champs, ainsi que le niébé et le haricot mungo. Le maïs et les haricots sont plus rarement cultivés en raison de l'aridité de la zone.

Figure 7. Parcelle de maïs anéantie par la sécheresse

---

[12] Le nom en kimeru est indiqué en premier, suivi du nom scientifique

III- Le sorgho

Le sorgho est une céréale majeure dans le monde, cultivée sur les cinq continents notamment en raison de son adaptation aux zones arides auxquelles les autres céréales majeures ne sont pas adaptées (Wayne Smith and Frederiksen 2000). En 2012, sa production mondiale a dépassé les 58 millions de tonnes et il a été cultivé sur plus de 37 millions d'hectares, ce qui le place au sixième rang mondial en termes de surface occupée (FAOSTAT 2012). Au Kenya, le sorgho est la troisième culture alimentaire en termes de surface occupée (224 000 Ha), et la huitième en termes de production (166 627 tonnes). Il est principalement cultivé à l'ouest du pays, et dans la zone du mont Kenya.

Position taxonomique

Du fait de sa grande diversité morphologique, la classification du sorgo a représenté un réel défi pour les taxonomistes. C'est à Moench que l'on doit la création du genre *Sorghum*, en 1794. Snowden subdivisa bien plus tard, en 1936, ce genre en 52 espèces dont 31 cultivées. Ces 31 espèces cultivées furent réduites à une par de Wet and Huckabay (1967). Le sorgho est actuellement classé dans la famille des Poacée et dans le genre *sorghum*, qui contient cinq sections. L'espèce *S. bicolor* [L.] Moench à laquelle appartient le sorgho cultivé est l'une des trois espèces de la section *sorghum*. Cette espèce est subdivisée en trois sous-espèces: le sorgho cultivé appartient à ssp. *Bicolor*, la forme adventice à ssp. *Drummondii* et la forme sauvage à ssp. *Verticilliflorum* (Figure 8).

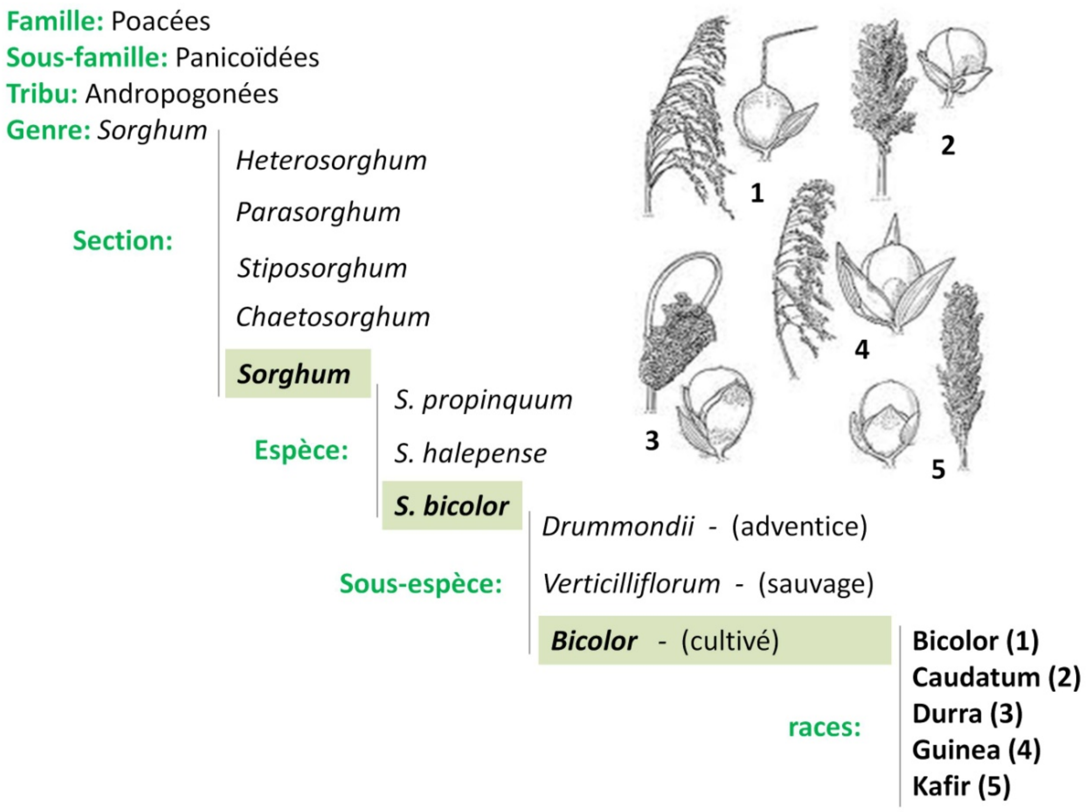


Figure 8. Classification du sorgho

Harlan et de Wet proposèrent une classification simplifiée pour le sorgho cultivé spp. *Bicolor* en 1972. Cette classification, basée sur les caractéristiques des épillets et de l'inflorescence à maturité, distingue 5 races botaniques et 15 intermédiaires. Ils proposent également de distinguer quatre races botaniques au sein de ssp. *verticilliflorum* (aethiopicum, arundinaceum, verticilliflorum et virgatum). La race arundinaceum est principalment rencontrée dans les zones humides et forestières d'Afrique centrale et de l'ouest. Virgatum est rencontrée sur les rives du Nil et dans des zones marginales au Soudan. Aethiopicum est rencontré au Soudan, en bordure du Sahara, et en Ethiopie. Verticilliflorum est largement distribué en Afrique sub-saharienne (Figure 9).

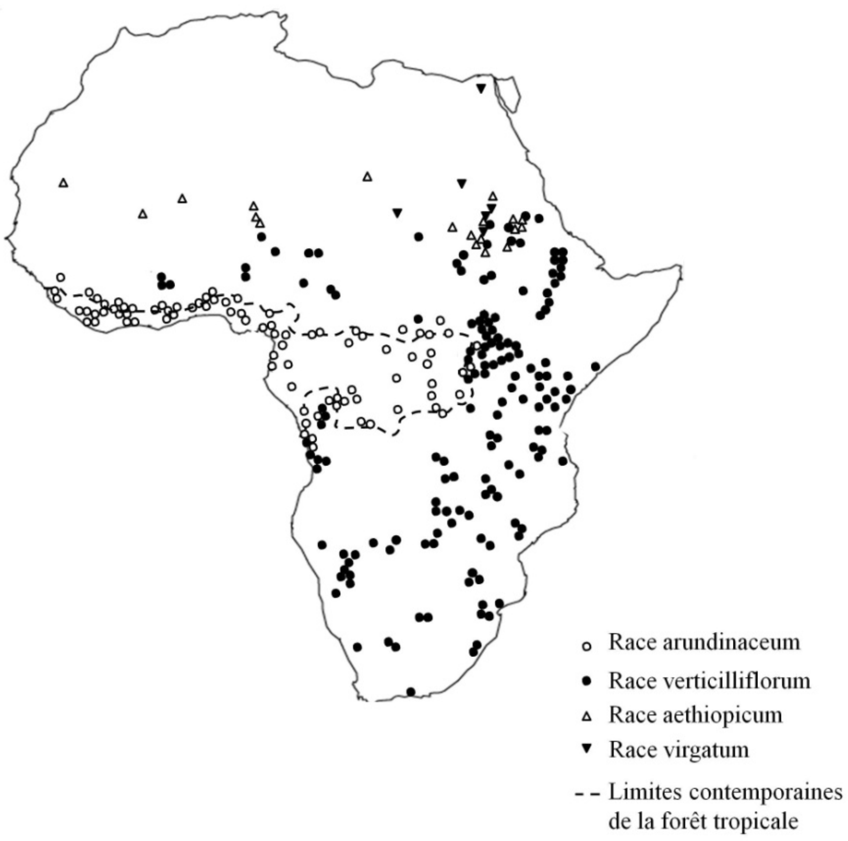


Figure 9. Distribution des différences races botaniques de la sous-espèce sauvage verticilliflorum (*S. bicolor spp.verticilliflorum*) en Afrique. (D'après Harlan 1995)

Biologie de la reproduction

Le sorgho présente des fleurs hermaphrodites et la pollinisation de cette espèce est anémophile. Le sorgho cultivé est considéré comme majoritairement autogame mais le taux d'allofécondation est fortement variable et peut atteindre jusqu'à 40 % *in situ* (Barnaud et al. 2008). En conditions contrôlées, des taux de 19 % (Ollitrault et al. 1997) à 30 % (Chantereau and Nicou 1991; Ellstrand and Foster 1983) ont été mesurés. Des flux de pollen semblent limités au-delà d'un vingtaine de mètres (Schmidt and Bothma 2006). Les différentes races de sorgho cultivé peuvent se croiser avec la sous-espèce sauvage *verticilliflorum* et cela a lieu assez fréquemment *in situ* (Deu et al. 1999), elles peuvent également s'hybrider avec l'espèce *S. propinquum*, qui est également diploïde (de Wet et al. 1976), et avec l'espèce *S. halepense,* tétraploïde (Arriola and Ellstrand 1996). Ces flux de gènes contribuent à l'enrichissement génétique du pool cultivé (Doggett 1988).

Histoire évolutive

Chez le sorgho, la principale marque de domestication est la perte du caractère déhiscent des grains, dépendant de la solidité du rachis, rendant donc sa dispersion dépendante de l'homme. La sélection pour ce caractère a vraisemblablement été possible du fait d'innovations technologiques pour la récolte (Zohary 1996). L'espèce sauvage à partir de laquelle il aurait été domestiqué fait l'objet de débats. Snowden émit l'hypothèse que le sorgho aurait été domestiqué à plusieurs endroits : la race sauvage *aethiopicum* aurait donné naissance aux races botaniques durra et bicolor, *verticilliflorum* à la race kafir et *arundinaceum* à la race guinea (Snowden 1936). Harlan suggère qu'au regard de son aire de distribution et de sa morphologie, *S. verticilliflorum* serait le parent le plus plausible, qui aurait donné naissance à la race domestiquée bicolor, jugée la plus primitive (Wayne Smith and Frederiksen 2000).

La date à laquelle le sorgho a été domestiqué ne fait pas consensus dans la communauté scientifique, principalement en raison du peu d'indices archéologiques disponibles. Trois principales hypothèses ont été proposées concernant la date de domestication du sorgho: la première est celle d'une domestication précoce. Murdock (1959 cité dans Wayne Smith and Frederiksen 2000) fait l'hypothèse d'une domestication en Afrique de l'Ouest aux alentours de 4 500 ans avant J.C. par les populations Mandé qui auraient ensuite connu une expansion et auraient introduit le sorgho au Soudan vers 4 000 av. J.C., mais cette hypothèse est peu soutenue. Plusieurs études suggèrent plutôt une domestication dans la zone du Soudan et de l'Ethiopie (Doggett 1988), Mann et al. (1983 cité dans Wayne Smith and Frederiksen 2000) proposent des dates plus précoces, aux alentours de 6 000 av. J.C. Wendorf et al. (1992) ont découvert des restes de sorgho carbonisés datant de 8 000 ans avant notre ère au sud de l'Egypte, mais le doute persiste concernant leur statut sauvage ou domestiqué. L'hypothèse d'une domestication plus tardive, autour de 1 000 ans av. J.C. est en revanche proposée par Stemler (1980) soutenant que les plus anciennes et irréfutables preuves archéologiques de culture du sorgho domestiqué datent de cette époque (Wayne Smith and Frederiksen 2000). Haaland (1998 cité dans Wayne Smith and Frederiksen 2000) propose une toute autre hypothèse, selon lui le sorgho sauvage aurait été introduit en Inde et c'est là qu'il aurait été domestiqué, mais cette hypothèse est peu soutenue.

Selon l'hypothèse d'une domestication dans la zone soudanienne, qui est la plus soutenue, la domestication du sorgho sauvage aurait donné naissance à la race bicolor. Diffusée vers l'Afrique de l'Est et du Sud par les migrations Bantoues, la race kafir se serait alors différenciée. La race guinea se serait différenciée en Afrique de l'Ouest, la race durra se serait vraisemblablement différenciée dans la zone de domestication et aurait ensuite été diffusée par les mouvements de population de langue afroasiatiques, notamment vers l'Inde. La race caudatum aurait été domestiquée plus tardivement et aurait été diffusée par les populations Chari-Nil (Stemler et al. 1975).

Le sorgho chez les peuples du mont Kenya

Place du sorgho

Le sorgho est la principale céréale cultivée dans notre zone d'étude, et sa place dans les systèmes de culture s'accroit à mesure que l'on s'approche des plaines occupées par les Tharaka. En revanche, on n'observe actuellement que rarement du sorgho à plus haute altitude, les principales cultures vivrières étant le maïs, les haricots, les bananes et les tubercules. Nos observations suggèrent que le sorgho n'est pas réellement prisé par les populations locales, qui préfèrent le maïs (surtout les jeunes générations) ou le mil. Il est principalement cultivé en raison de son adaptation au climat semi-aride de la zone.

Figure 10. Diversité des types de sorgho cultivés au sein d'un foyer

Les différents types : ratoon et cycle court

Les agriculteurs classent les variétés de sorgho en deux grandes catégories en fonction de la durée de leur cycle de culture. Les variétés à cycle court (*Mbura imwe* – « une saison ») peuvent être semées soit en octobre pour une récolte en janvier, soit en mars pour une récolte en juin. Les variétés à cycle long (*Mbura ciiri* – « deux saisons »), qualifiées de « ratoon », sont semées en octobre et leur partie végétative est coupée avant que les grains arrivent à maturité, aux alentours de janvier, et leur récolte a lieu en juillet. Cette pratique de ratooning est décrite par Plucknett et al. (1970), elle a pour objectif de stimuler la repousse depuis les bourgeons basaux. Il semblerait, selon les dire des agriculteurs, que ces variétés permettent d'assurer une récolte en juillet car l'établissement des semis des variétés à cycle court est fréquemment perturbé par les conditions climatiques aléatoires. Les variétés ratoon ayant déjà un système racinaire bien implanté, elles ne seraient pas affectées par ce problème. De plus, la partie végétative récoltée en janvier permet de constituer des réserves de fourrage sec pour nourrir le troupeau pendant la saison sèche.

Selon nos informateurs, les variétés ratoon seraient « *le sorgho des Chuka et des Mwimbi*[13] ». C'est en effet quasiment les seules variétés cultivées à plus haute altitude, dans la zone de résidence principale et originelle de ces groupes. Les Mbeere sembleraient ne cultiver qu'un seul type ratoon, nommé « Muruge Kimbeere » (« Muruge des Mbeere »). Des photos des différentes variétés de sorgho sont fournies en annexe (Annexe 3).

Itinéraires de culture

Selon les exploitations, les variétés de sorgho sont mélangées, en portions séparées d'un même champ ou encore dans des parcelles séparées, ce qui détermine l'intensité des flux de pollen entre elles. Dans les trois groupes, environ la moitié des agriculteurs sème les variétés en mélange et l'autre moitié sépare les variétés au sein d'une même parcelle. Une part négligeable des agriculteurs cultivent les différentes variétés dans des parcelles séparées. Les différentes espèces cultivées sont mélangées au sein des parcelles.

---

[13] Groupe ethnolinguistique adjacent des Chuka

Différents modes de semi sont adoptés par les agriculteurs selon le matériel dont ils disposent. Le semi en poquets était le mode de semi utilisé traditionnellement (« *Chuka style* »), puis le labour animal à la charrue a été introduit autour des années 1960. Les agriculteurs qui sèment en poquets ont davantage tendance à mélanger les graines de différentes variétés dans la parcelle et ceux qui sèment en ligne à les séparer. La majorité des agriculteurs déclarent démarier les plants de sorgho lorsqu'ils sont trop serrés au sein des poquets ou des lignes.

Les pratiques de sélection

A l'approche de la maturité, les panicules de sorgho sont fortement vulnérables vis-à-vis des attaques d'oiseaux. Les enfants ou les hommes âgés surveillent les champs, armés de frondes, et la récolte des panicules se fait généralement au fur et à mesure de leur maturité. Tous les jours, les femmes récoltent les panicules matures, qu'elles stockent dans leur *kiondo* (sac en sisal et coton). Les panicules sont ensuite mises à sécher, généralement sur le toit en tôle des habitations, avant d'être stockées dans le grenier en attendant d'être battues. L'aire de battage est préalablement préparée avec un enduit à base de bouse de vache qui, une fois sec, permet d'obtenir une surface lisse et non-poussiéreuse.

La grande majorité des agricultrices sélectionnent les panicules semencières après la récolte, séparément pour chaque variété, avant de les battre (90%). la sélection s'effectue alors uniquement sur les caractéristiques de la panicule. Chaque variété de sorgho est disposée en tas sur l'aire de battage, les panicules semencières sont alors choisies au sein de chaque tas. Les panicules présentant de nombreux grains, bien remplis et sans attaques d'insectes ou de maladies sont choisies. La quasi-totalité des agricultrices contre-sélectionnent les panicules hors-types, jugées non-conformes aux critères morphologiques définissant chaque variété, dans un souci de maintenir les variétés « pures ». Les agricultrices craignent que si elles sélectionnent ces panicules, la variété « change complètement » et finisse par « donner du sorgho sauvage ». Une minorité d'agricultrices, en revanche, ont déclaré sélectionner des panicules dont la coloration inhabituelle des grains les a attirées, leur permettant ainsi d'obtenir de nouvelles variétés.

Une large part des agricultrices ont déjà observé des panicules hors-types dans leur récolte. Selon elles, ces hors types apparaissent lorsque les semences ont été reproduites à la ferme un grand nombre de fois, surtout si la sélection n'a pas été accomplie consciencieusement. Il s'agit généralement de variations de coloration des grains, des glumes et de la forme de la panicule qui de compacte devient lâche, ou par une forte déhiscence des grains qui est un caractère sauvage. Parfois, c'est également la hauteur de certaines variétés qui était uniforme (« *wa nthuke imwe* » - « un seul âge ») et devient variable. Lorsque le changement observé sur le morphotype est mineur, le type variant est nommé comme la variété dont il est issu. En revanche, plusieurs termes sont utilisés pour les stades plus avancés de dégénérescence : le terme « *Matharara*» (« stupide »[14]) désigne les panicules dont les grains commencent à devenir déhiscents. Après avoir été encore replantée, la variété devient alors « *Kiamaguna* » (« sorgho des singes »), impropre à la consommation. Certaines agricultrices connaissent le processus de pollinisation, qu'elles attribuent aux insectes. Certaines agricultrices attribuent ces changements au fait que les abeilles ont butiné des fleurs colorées qui teignent le grain, ou encore au sol ou aux conditions climatiques.

La quantité de semences sélectionnées est très variable selon les agricultrices, elle se situe en moyenne entre 3.5 et 4.5 Kg (sur la base de leur déclaration). Elle est rarement de moins de deux kilos et peut atteindre 10 kg. En général les agriculteurs préfèrent garder de grandes quantités de semences lorsque la récolte est bonne pour pouvoir assurer plusieurs semis, et ce sur plusieurs saisons, en raison des conditions climatiques difficiles et imprévisibles.

La sélection est dans la quasi-totalité des cas réalisée par les femmes (93%), lorsqu'un homme a plusieurs épouses, chacune sélectionne ses propres semences provenant de son champ, et chacune possède son propre grenier. Belle-mère et belles-filles au sein de la même unité familiale sélectionnent également chacune leur propre récolte. Les hommes ne sélectionnent les semences que lorsqu'ils sont veufs ou célibataires, ou lorsque leur épouse est inexpérimentée (originaire d'une zone où le sorgho est peu cultivé, par exemple chez certains Kikuyu). Les jeunes hommes non-mariés font parfois sélectionner leurs semences par leurs mères.

---

[14] Se dit par exemple d'un enfant indiscipliné

Figure 11. Vannage du mil à l'aide d'une calebasse

Usages

Les usages du sorgho ont été décrits par N. Linsig (2009) dans un groupe adjacent, les Mwimbi, et sont d'après nos observations assez similaires à ceux observés dans notre zone d'étude. Le sorgho est essentiellement utilisé sous forme de farine pour préparer une sorte de porridge consommée notamment le matin ou emportée lors des travaux au champ. Une version de cette préparation se fait avec de la farine broyée manuellement à la meule de pierre additionnée de farine de mil et est fermentée. Le sorgho est également parfois consommé sous forme de grains entiers bouillis, mais cette préparation est plus rare. Enfin, le sorgho est également utilisé pour préparer de la bière, traditionnellement réservé aux occasions en raison du temps nécessaire à sa préparation. L'introduction d'une nouvelle variété de sorgho destinée au brassage industriel, fermentant plus rapidement au dire des agriculteurs, semble avoir développé une production de bière locale régulière.

Figure 12. Usages alimentaires : à gauche : mouture du mil à la meule en pierre ; en haut à droite : calebasse emplies de porridge acide ; en bas à droite : Kitheli, principal plat consommé sur la zone, composé de haricots et de maïs bouilli.

DE L’ORGANISATION DES SOCIETES A L’ORGANISATION DE LA DIVERSITE DU SORGHO

CHAPITRE I- ECHELLE DES GROUPES DE VOISINAGE

ARTICLE: HOW SOCIAL ORGANIZATION SHAPES CROP DIVERSITY: AN ECOLOGICAL ANTHROPOLOGY APPROACH AMONG THARAKA FARMERS OF MOUNT KENYA

Vanesse Labeyrie, Bernard Rono, Christian Leclerc

Publié dans *Agriculture and Human Values*

Agric Hum Values
DOI 10.1007/s10460-013-9451-9

How social organization shapes crop diversity: an ecological anthropology approach among Tharaka farmers of Mount Kenya

Vanesse Labeyrie · Bernard Rono · Christian Leclerc

Accepted: 29 April 2013

Abstract The conservation of in situ crop diversity is a key issue to ensure food security. Understanding the processes that shape it is crucial for efficiently managing such diversity. In most rural societies, crop diversity patterns are affected by farmers' practices of seed exchange, transmission, and selection, but the role of social organization in shaping those practices has been overlooked. This study proposes an ecological anthropology approach to investigate the relation between crop diversity patterns and the social organization of Tharaka farmers in Kenya. The Tharaka are organized in neighborhood-groups, clans, and age-sets. We quantified the influence of these three major social institutions on crop diversity patterns, for both crop species and sorghum landraces. General linear models were used to test the relations between crop species richness and each social factor, while the crop species and sorghum landraces compositions of cropping systems were compared separately through a between-class correspondence analysis. Crop species and sorghum landraces are not randomly distributed among farms, and neighborhood-groups constitute a significant factor organizing crop diversity at both specific and infraspecific levels. Adjacent neighborhood-groups present significantly different crop richness and composition. The results for species were consistent with those obtained for sorghum landraces, confirming that crop diversity was socially structured. The influence of social organization on seed networks and selection processes is discussed.

Keywords Agrobiodiversity · Sorghum · Social networks · Seed exchange · Farmers' selection · Crop domestication

Introduction

Subsistence farming systems, which ensure food supplies for one-third of the world's population, are mostly based on mixed cropping. Crop diversity at specific and infraspecific levels ensures the resilience of smallholder farming systems in changing environments. Conservation of crop genetic resources is therefore a major issue for food security (Thrupp 2000).

To develop efficient in and ex situ conservation strategies, it is necessary to identify the mechanisms that shape crop diversity in situ. Indeed, the distribution of crop diversity is not random. Crop evolution, like that of wild plants, is driven by genetic drift, natural selection, and migration. However, crops are highly dependent on human selection and seed exchange practices. In most subsistence farming systems, farmers are involved in a social organization that shapes relationships and thus affects seed exchanges as well as knowledge and practices diffusion. Crop diversity in situ is not only shaped by environmental barriers, but also by social barriers. Indeed, the social relationships favor the diffusion of planting material, practices, and information between farmers. Contrarily, social barriers limit both the exchanges of seed and the transmission of knowledge and practices between farmers' communities, thereby affecting the distribution of crop genetic diversity

V. Labeyrie (✉) · C. Leclerc
CIRAD, UMR AGAP, 34398 Montpellier, France
e-mail: vanesse.labeyrie@gmail.com

C. Leclerc
e-mail: christian.leclerc@cirad.fr

B. Rono
KARI, PO Box 27, Embu, Kenya
e-mail: bernard.rono@gmail.com

in situ (Leclerc and Coppens d'Eeckenbrugge 2012). Seed migration depends on social networks of exchanges (McGuire 2008; Pautasso et al. 2013), while both vertical transmission of ethnobotanical knowledge and the limitation of horizontal transmission favor cultural differentiation (McElreath and Strimling 2008; Cavalli-Sforza and Feldman 1981). This latter social mechanism can explain the divergence of agricultural knowledge and practices between farmers' communities (Boster 1986). The limitation of seed exchanges, knowledge, and practices diffusion should thus lead to the differentiation of crop populations between communities. Perales et al. (2005) notably illustrated the influence of ethnolinguistic organization on maize diversity in Mexico by showing how the Tzotzil and Tzeltal farmers' communities, living in the same environment, have divergent selection practices. In northern Benin, Baco et al. (2008) showed that the differentiation of yam landraces was related to different ethnic groups. However, the influence of farmers' social organization on crop diversity remains insufficiently studied.

Ecological anthropology, which is defined by Orlove (1980) as the study of the relations among the population dynamics, social organization, and culture of human populations and the environments in which they live, provides theoretical and methodological insights for studying the relations between social organization and crop diversity. With this context we investigate the influence of rural communities' social organization on crop diversity patterns in situ. Among the Tharaka farming community, located in the southeastern slope of Mount Kenya, relationships are embedded within three major social institutions: neighborhood-groups, clans, and age-sets. The effect of each social institution on crop composition and species richness of households' cropping systems is quantified using ecology approaches. Results are then discussed regarding ethnographic observations and literature concerning Tharaka social relationship system. These results can be applied to the study and the conservation of crop genetic resources as they contribute to the global understanding of mechanisms that shape crop diversity in situ. The design and implementation of in situ conservation programs, as well as the collection and sampling of crop genetic resources for genetic studies and efficient ex situ conservation require a better understanding of the relation between the social organization of rural communities and the distribution of crop diversity. It is also essential in designing participatory research programs and ensuring the efficient diffusion of their results.

Conceptual framework

Leclerc and Coppens d'Eeckenbrugge (2012) proposed a multidisciplinary methodological framework combining social anthropology and crop population genetics. Through many examples, they illustrated that traditional knowledge as well as seed exchange systems are embedded into social structure, favoring vertical transmission of knowledge, practices, and plant genetic resources through a centripetal system. Thus, they consider that crop diversity patterns result not only from an interaction between genetic and environmental factors, G × E, but from a three-way interaction G × E × S, where "S" stands for the social differentiation factors. Following this framework, the present study investigates how Tharaka farmers' social organization contributes to shaping the diversity of their crops.

The Tharaka are one of the nine dialectal sub-groups composing the Meru group, which is part of the Bantu linguistic family (Moehlig et al. 1980). They settled about two centuries ago in the semi-arid plains on the eastern slope of Mount Kenya (Middleton 1953; Fadiman 1993) (Fig. 1). The economy of Tharaka smallholders is based on subsistence farming systems, involving a wide diversity of crop species and landraces. They carry out two cropping seasons, the first from October to December and the second from March to May. Their farmlands are frequently hit by drought, followed by severe food shortages, of which one of the most notable was described by Ambler (1988). Moreover, this semi-arid area is facing climate change, which severely affects the production (Downing 1992). Shift in rainfall seasonality and increasing temperatures cause changes of crop species and varieties, as some crops grown in the area are at their limit of adaptation (Mati 2000).

According to previous ethnographic studies, the Tharaka social organization is based on clan ties, neighborhood-groups, and age-sets (Lambert 1947; Middleton 1953). Children belong to the clan of their father and women adopt the clan of their husband when they get married. The clan identity determines social relations such as exchanges, cooperation, and marriage opportunities. Indeed, Tharaka clans are exogamous, prohibiting marriage between people belonging to the same clan. Affiliated clans live together within a neighborhood-group, called *ntora*. After marriage, women usually settle in the native ntora of their husbands. The residence is thus patrilocal and women frequently come from a different ntora or even a different territorial subdivision from that of their husbands (Peatrik 1999). Each household is thus included within a named ntora. The ntora constitutes the lowest level of political organization (Middleton 1953).

The age-set system is the third important Tharaka social institution described by Lambert (1947) and Middleton (1953), who defined it as the group of youths circumcised at the same time, irrespective of their clan or territorial affiliation. Tharaka are thus not only organized horizontally, with

Fig. 1 Study site location

differentiated clans and ntora, but also vertically with different age-sets, however the colonial administration considerably weakened this system (Peatrik 1999). The planting, selection, and trading of seeds are done by women. Men prepare the fields before the sowing, manage tobacco and some clonally reproduced crops such as yam, banana or sugar cane (Middleton 1953; authors' observations).

Our main hypothesis is that the social organization of the Tharaka influences the exchanges between farmers, including the diffusion of planting material, practices, and information. If ntora, clans or age-sets are social barriers limiting the flows of planting material and information between farmers, we expect the specific and the infraspecific portfolios to be different between ntora, or between clans, or between age-sets. Indeed, according to the three way interaction models, G × E × S, proposed by Leclerc and Coppens d'Eeckenbrugge (2012), these three social institutions (ntora, clan, and age) should structure the exchanges within the Tharaka society. However, the respective influence of each institution on the exchanges of planting material and information is unknown. In this paper, we thus investigate which of these three social institutions have a significant influence on crop diversity.

Methods

Crop diversity was measured through richness and composition of cropping systems, using quantitative approaches developed in ecology. We tested whether a relationship existed among the number of crop species cultivated per household (i.e., "species richness"), the household's crop species portfolio, which is the assemblage of crops that are cultivated together in the same household, (i.e., "crop composition"), and the ntora, clans, and age-sets of farmers. The results obtained at the interspecific level were compared with those obtained at the infraspecific level for sorghum (*Sorghum bicolor* (L.) Moench), which is the major crop of the Tharaka. The environmental factors were controlled in our field sampling strategy to avoid confusion in our inferential procedure.

Study site

Our study site was approximately 10 km^2. It was selected for its edaphic and topographic uniformity, with a constant altitude of 700 m above sea level (±50 m). The soils are also uniform across the study site. They are infertile, drained, deep, clayey, and powdery red ferrasols resulting from volcanic rock alteration (Jaetzold et al. 2007). The mean annual temperature is 22.9 °C and the mean annual rainfall is 600–700 mm (Camberlin et al. 2012; 2009).

Ninety-five households were randomly sampled on an aerial shot in order to represent around 40 % of the total population of the study area. In each household visited, women were preferentially interviewed (83 %) because agriculture is their field of competence. The households belonged to 11 ntora (Fig. 2) and 14 patrilinear and exogamous sub-clans. The five major ntora and the three main clans were retained for the analysis. Age categories between 20 and 50 years were equally represented in the sample and the mean age of the interviewed people was 42. Four age categories were constituted a priori for ease of analysis (15–30, 31–40, 41–50, 51–92).

The interviews were conducted with each farmer individually, and consisted of three sections. The first section dealt with social information concerning the household heads: their date and location of birth, their native clan, and the ntora they belonged to. The second section listed the crop species planted in the October 2010 cropping season according to the informant, and the third section reported

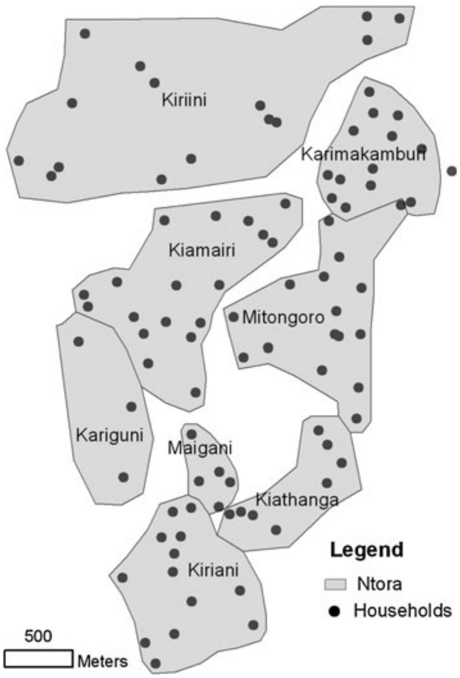


Fig. 2 Localization of the ntora

the inventory of sorghum landraces for the same season. We also surveyed the main source of each sorghum seed lot planted that season. We indicated whether they were own saved seed, purchased from market, or received from relatives. In the last case, the social link between the seed-provider and the informant was specified (belonging to the same ntora, the same clan, or the same age-set).

Data analysis

We performed statistical analyses to test the relationship among the social factors and the diversity of Tharaka cropping systems, at both specific and infraspecific levels. We tested whether the ntora, clans, and age categories (explanatory factors) explained crop diversity (response variable). The diversity of cropping systems was measured for each household by considering (1) their species and sorghum landraces richness and (2) their specific and sorghum infraspecific composition. All analyses were carried out with the R 2.13.0 software (R Development Core Team 2011).

Richness

Species richness is the number of different species (or sorghum landraces) inventoried in each household. The mean and cumulative richness were calculated for each ntora, clan or age category to compare them. The cumulative richness is the total number of distinct species (or sorghum landraces) inventoried in a population, and the mean richness is the sum of households' richness divided by the number of households in the population.

The relationship between crop species richness and each of the three social explanatory factors was assessed by a Poisson loglinear model, which is adequate for count data (Agresti 2007). For a single explanatory factor x, the Poisson model is

$$\log\mu = \alpha + \beta x,$$

where μ the mean (and the variance) of the explained variable (richness). Log-likelihood procedures determine the values of model parameters that maximize the probability of observing the empirical data. As the third and second order interactions were not significant, we studied the effect of each social factor without interaction. To assess whether the mean richness differed significantly between each social variable level (population), we used a Tukey's Honestly Significant Difference test, which compares the observed pairwise differences between populations' means to the distribution of expected pairwise differences under the null hypothesis (no differences between populations). The difference between two means is significant if it is higher than that observed in 95 % of the cases under the null hypothesis.

Composition

We established an exhaustive list of crop species and sorghum landraces cultivated in each household. The occurrence of each species and landrace was recorded in two separate matrix (one for the species and the other for the landraces), where the presence of a given species or landrace in a household was coded 1 and its absence was coded 0. Such a presence-absence matrix enables us to compute a distance between households based on their crop composition: the more similar is the crop composition of households, the shorter is the distance between them. The distance between households was measured using the Jaccard similarity index, computed for crop composition between pairs of households as follows:

$$dij = \frac{a}{a+b+c},$$

where dij is the Jaccard similarity between households i and j; a is the number of species present in both i and j households; b is the number of species present in the household i and absent in j; c is the number of species present in the household j and absent in i. A high number of common species between two households would result in a high similarity index, and vice versa. Pairwise similarity index of crop composition between households were computed separately for crop species and for sorghum landraces, and stored in separate matrix.

The composition of crop species and sorghum landraces was then compared among ntora, clans and age categories using a Between-Class correspondence analysis (Chessel et al. 2004; Doledec and Chessel 1989). The Between-Class analysis was performed on crop composition distance matrix, separately for crop species and sorghum landraces. This multivariate ordination analysis tests whether crop composition is more similar within groups (ntora, clans, or age categories) than between. Principal components are computed in order to maximize the variance between groups instead of maximizing the total variance, as correspondence analysis does. It is then possible to assess whether crop composition differs between groups using both graphical representation and Monte-Carlo test described thereafter. This constrained ordination method was developed in ecology to compare species composition between sites. An illustration of its application is given by Paillex et al. (2009) in ecology and by Baty et al. (2006) in genetics.

To test whether the crop specific and sorghum infra-specific compositions of households differed significantly between groups (ntora, clans, age categories), we used a Monte-Carlo permutation test (Manly 1997). This non-parametric test used 9999 permutations of the data to draw a random distribution of crops. It then tested whether the observed crop composition was more similar within groups than it would be under the null hypothesis, if crops were randomly distributed.

Results

Our survey confirmed previous ethnographic observations on the social organization of Tharaka, reporting that residence rule was patrilocal. Indeed, patrilocality (settlement within the ntora of the man's father) was observed in 70 % of the sampled households, and 84 % of the spouses originated from a different ntora other than their husbands' ntora. These observations are relevant for the understanding of seed, practices, and information diffusion pathways.

Patterns of richness and composition of cropping systems for crop species

Sixteen crop species were inventoried among the 95 households sampled, while each household cultivated, on average, between five and six species (mean: 5.3 ± 1.4, [1] min: 2; max: 9). Sorghum (*Sorghum bicolor*), pearl millet (*Pennisetum glaucum*), cowpea (*Vigna unguiculata*) and green grams (*Vigna radiata*) were grown by 80 % of the households. Forty percent of the households grew only those four crops, which are adapted to drought and form the basis of Tharaka cropping systems. Maize (*Zea mays)* and pigeon pea (*Cajanus cajan*) were cultivated by around 60 % of the households. Lastly, marginal species like pumpkins (*Cucurbita* sp.), cotton (*Gossypium* sp.), beans (*Phaseolus* sp.), cassava (*Manihot esculenta*), or dolichos (*Dolichos* sp.*)* were grown by less than 10 % of the households.

Table 1 p value of Tukey's HSD pairwise comparisons for the mean specific richness between ntora

	Kiamairi	Kiriani	Kiriini	Mitongoro
Karimakamburi	0.999	0.128	0.494	0.002**
Kiamairi		0.543	0.309	<0.001***
Kiriani			0.936	0.726
Kiriini				0.227

Significance codes: *** 0.001; ** 0.01; * 0.1

The species richness did not differ significantly between clans and between age categories, whereas the species richness of a household depended significantly on the ntora it belonged to. The p values of the pairwise Tukey's HSD test (Table 1) were significant for the differences of mean species richness between ntora. The specific richness of households belonging to Kiamairi (4.5 ± 1.1 species) and Karimakamburi ntora (4.6 ± 1.4 species) were significantly lower ($p < 0.01$) than that of Mitongoro (6.5 ± 1.75 species). Despite their adjacency, households belonging to Mitongoro maintained more diversity than those of Karimakamburi (Fig. 3).

The Between-Class analysis according to ntora showed that crop species were not randomly distributed as there were significant differences of species composition between ntora. The ntora factor explained 9 % of the total variation of species composition, which is a significant proportion of between-group variation. The specific composition of the cropping systems of Mitongoro differed from that of Kiamairi and Karimakamburi (Fig. 4). We observed that the p value of the Monte-Carlo test for species composition differentiation was significant for ntora ($p = 0.0171$; Fig. 5), but not for clans ($p = 0.0977$) and age categories ($p = 0.7254$), meaning that species composition differed significantly between ntora, contrary to clans and age categories which were not significant explanatory factors for species composition.

Sorghum and millet were cultivated in every household. By contrast, the proportion of households growing maize varied between 35 % in Kiamairi to 87 % in Mitongoro, and from 35 to 81 % for pigeon pea in the same groups (Table 2). Fewer households in Karimakamburi were growing pigeon pea (43 %), maize (57 %), cowpea (79 %) and green gram (79 %) as compared to Mitongoro (100 % for cowpea and 94 % for green gram).

---

[1] Means are followed by their standard deviation.

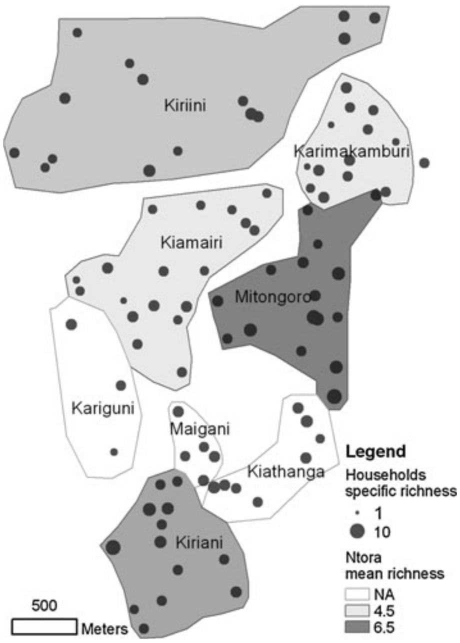


Fig. 3 Mean species richness of ntora and households' species richness

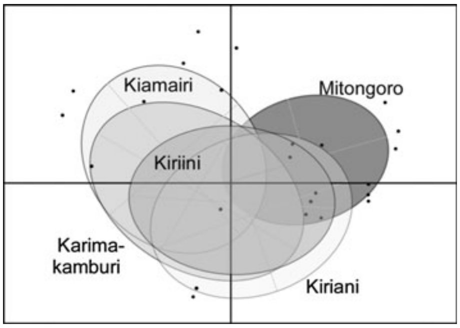


Fig. 4 Graphic display of the Between-Class analysis for crop species composition between ntora: projection of farms' portfolios similarity on axis 1 and 2 with ellipse and gravity centers of each ntora. The first and the second components of the Between-Class analysis represented respectively 66 and 14 % of the between-ntora variability

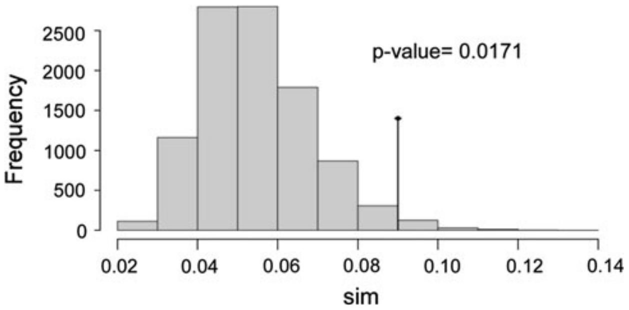


Fig. 5 Histogram of the 9999 simulated values of the Monte-Carlo test for the Between-Class analysis on crop species composition between ntora. The observed value is given by the vertical line

Patterns of richness and composition of cropping systems for sorghum landraces

While the Tharaka cultivated 21 different sorghum landraces on our study site, each household grew few of them (1.79 ± 0.74). The frequency distribution of sorghum landraces (Fig. 6) indicates that large differences in composition exist between households. *Mucarama* was the most common landrace, grown by 60 % of households; it was followed by *Kaguru* and *Mugeta*. The remaining 18 landraces were grown by less than 8 % of households, showing that each household separately cultivates only a small part of the whole infraspecific diversity of sorghum of the community.

Two-thirds of the sorghum landraces inventoried were local, which means that they have been grown within the household for at least one generation. Farmers produced a large proportion of their seed lots, thereby favoring their own local landraces. Seed source indicated that 77 % of seed lots sown in October 2010 were own seed or obtained from farmers belonging to the same ntora. On the other hand, the percentage of local landraces varied between ntora. Kiamairi had the highest percentage of local landraces, with 87 % of the October 2010 seed lots, while this percentage was 52 % in Karimakamburi, 58 % in Kiriini and 59 % in Mitongoro.

No significant relation was observed between the sorghum landraces richness and any of the three explanatory social variables (ntora, clans, age categories), but there were major differences of cumulative richness among ntora (Table 3). Table 3 shows that the cumulative richness was much lower in Karimakamburi (4 landraces) than in Kiriani (10 landraces) for a similar number of households visited.

The patterns of sorghum landraces composition were not random and the Between-Class analysis showed that there were significant differences in sorghum infraspecific composition among ntora. The ntora factor explained 11 % of the total variation in sorghum infraspecific composition. The sorghum landraces composition of Mitongoro differed from the compositions of Kiriani and Kiamairi, despite the spatial proximity between Kiamairi and Mitongoro (Fig. 7). The sorghum landraces composition of Karimakamburi also differed from that of Kiamairi. The Monte-Carlo test was significant ($p = 0.0010$; Fig. 8), confirming that sorghum landraces composition differed among ntora and that crop diversity was socially organized. By contrast, clans and age categories were not significant explanatory factors for sorghum infraspecific composition (Monte-Carlo $p = 0.6238$ between clans and $p = 0.3915$ between age categories).

Table 2 Percentage of households growing the major species in each ntora

	Number of farms	Sorghum	Millet	Cowpea	Green gram	Pigeon pea	Maize
Kiamairi	17	100	88	94	82	35	35
Mitongoro	16	100	94	100	94	87	81
Karimakamburi	14	100	100	79	79	43	57
Kiriini	14	100	93	93	100	50	71
Kiriani	13	100	100	100	100	54	77

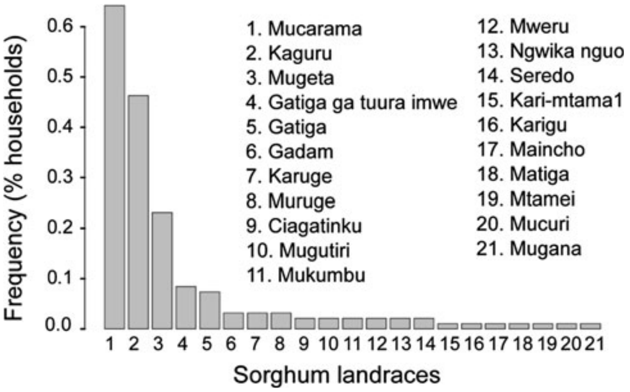


Fig. 6 Frequency of sorghum landraces (proportion of farms where each landrace is cultivated)

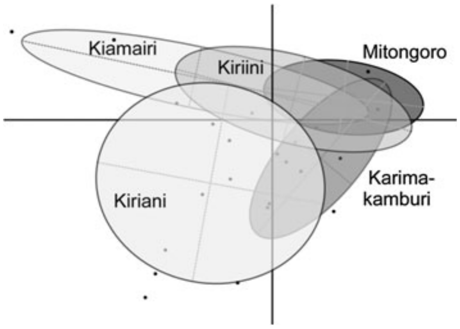


Fig. 7 Graphic display of the Between-Class analysis for sorghum infraspecific composition between ntora: projection of farms' portfolios similarity on axis 1 and 2 with ellipse and gravity centers of each ntora. The first and second projection axes represented respectively 53 and 23 % of the between-ntora variation

Table 3 Mean infraspecific richness (mean number of landraces cultivated per household within each ntora) and cumulative infraspecific richness (total number of distinct landraces cultivated in the ntora) of sorghum among the five main ntora

Neighborhood group	Number of farms	Mean richness	Cumulative richness
Karimakamburi	14	1.6	4
Kiriini	14	1.9	5
Mitongoro	16	1.7	7
Kiamairi	17	1.8	8
Kiriani	13	1.7	10
Total in the area	95	1.8	21

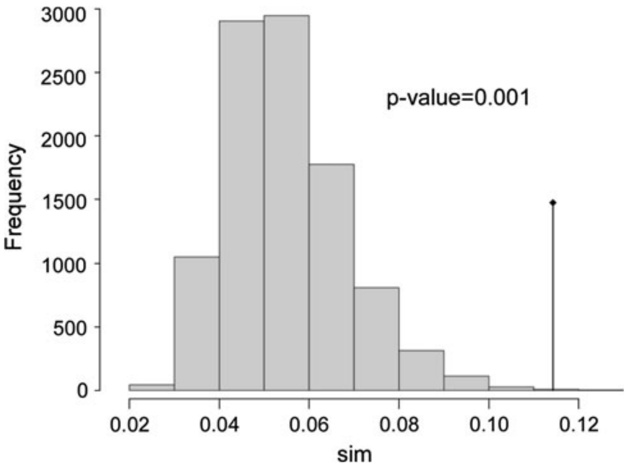


Fig. 8 Histograms of the 9999 simulated values of the Monte-Carlo test for the Between-Class analysis on sorghum infraspecific composition between ntora. The observed value is given by the vertical line

Discussion

This study highlighted linkages between Tharaka social organization and crop systems, for both crop species and crop landraces. Its aim was to quantify the effect of three major social institutions of the Tharaka, which are ntora (neighborhood-groups), clans, and age-sets, on the distribution of species and sorghum landraces diversity.

Our results showed that rural communities' social organization significantly influences crop diversity patterns in situ. The ntora organization contributed significantly to the diversity patterns of both crop species and sorghum landraces, while clan and age did not significantly influence the organization of specific and infraspecific diversity. On the contrary, no environmental effect was detected between the adjacent Mitongoro and Kiamairi ntora whose farms displayed significantly different crop richness and composition. Moreover, the proportion of local landraces differed between Mitongoro and Kiamairi. This latter group presented the highest proportion of local sorghum landraces and the lowest proportion of households growing maize, which was introduced in the 1960s.

Relation between crop diversity and the social organization

Leclerc and Coppens d'Eeckenbrugge (2012) explain that the social organization of human groups affects the differentiation of their crop germplasm. They propose that social anthropology can help to understand the differentiation mechanisms involved through the study of knowledge transmission, seed inheritance, and exchanges paths, together with marriage and residence rules.

Our results corroborate previous studies showing that farmer's behavior is strongly influenced by the behavior of others (Bandiera and Rasul 2006) and that social network plays a crucial role in the access to information (Van den Broeck and Dercon 2011) and seeds (David and Sperling 1999; Bellon 2004; McGuire 2008). The social network determines the access to seed sources as well as the opportunities to learn about crop species and varieties (Bellon 2004). The dependence of both information and seed exchanges on social organization can explain why such differences of crop diversity were found at both specific and infraspecific level between spatially close ntora. Ethnographic studies (Lambert 1947; Middleton 1953; Peatrik 1999) underlined the importance of the ntora in the social organization of the Tharaka and other Meru groups, it is thus likely that the exchanges of information and planting material take place mainly within it. Social barriers limit information and seed exchanges, and contribute to explain the differences of richness and composition we observed between ntora.

No relation between crop diversity and clan or age was established in this study, and little ethnographic information concerning these social institutions is available for the Tharaka. The complexity of clan and age-sets systems and the fact that they were considerably affected by the colonial administration (Brokensha and Glazier 1973; Peatrik 1999) could explain why their contribution to shaping seed and information exchanges appeared limited. Introduction of formal education and abandoning of indigenous cultures could also explain why their contribution is not significant.

Seed exchanges are limited by the social barriers

Among the Tharaka, the existence of social barriers is underpinned by the fact that most of the seed exchanges take place within the ntora, which corroborates the observation of Badstue et al. (2007) concerning the importance of trust in seed exchanges. It also supports the findings of McGuire (2008) who showed the importance of social networks in shaping seed networks. Centripetal seed exchanges, combined with vertical transmission customs, thus favor the differentiation of crop diversity between ntora (Leclerc and Coppens d'Eeckenbrugge 2012). At the same time, the Tharaka customary seed transmission pathway favors the maintenance of species and landraces within the ntora as ethnographic observations document that the mother-in-law provides her seeds to the newly married daughter-in-law who settles within the same ntora. On the other hand, one can argue that the proportion of seed lots coming from outside the ntora is potentially high and can lead to the homogenization of crops across ntora. The fact that this does not happen suggests that the social organization also contributes indirectly to restrict the exchanges of planting material through the limitation of information flows between farmers from different ntora.

Knowledge and practices are shaped by the social network

Several studies have shown that the social network is crucial for the adoption of new technologies and new planting material. They have suggested that it was more relevant to focus on small-scale social interaction (Conley and Udry 2001; Bandiera and Rasul 2006) and intensive exchanges of information between neighbors has been highlighted in several countries (Conley and Udry 2010; Van den Broeck and Dercon 2011). Knowledge transmission pathways affect cultural differentiation (Cavalli-Sforza and Feldman 1981), they are consequently involved in the divergence of farmers' practices. The vertical transmission of knowledge from parents or relatives to children favors cultural differentiation, contrary to the horizontal transmission between individuals belonging to the same age cohort that favors cultural uniformity. According to previous observations in another Meru group, the newly married spouses are trained to farming and household running by their mothers-in-law until the latter judge that the spouses are ready to manage their own household (Linsig personal communication). The knowledge transmission concerning crops thus follows the transmission of seeds from mothers-in-law to daughters-in-law. This vertical transmission pathway for agricultural practices favors their divergence between patrilinear families. In addition, horizontal knowledge exchanges between ntora appeared limited among the Tharaka. Indeed, we noticed that most mutual help and cooperation takes place within the neighborhood-groups, which is consistent with observations in other Meru groups (Peatrik 1999). It is therefore likely that both vertical transmission of knowledge from the mother-in-law to her daughter-in-law and horizontal transmission within the ntora favor the divergence of agricultural practices and thus the differentiation of crops between neighborhood-groups. For instance, it is striking that maize frequency is much lower and local sorghum landraces frequency is much higher in Kiamairi as compared to Mitongoro. This may result from a common reluctance to introduced crops in Kiamairi. Personal observations support this hypothesis as some Kiamairi farmers had a negative image of maize because it was not traditional. In addition, the fact that maize is not well adapted to the harsh climatic conditions of the area may explain why farmers' strategies concerning this particular species are different.

Last, the vertical knowledge transmission pathways may lead to the divergence of farmers' selection practices between neighborhood-groups, and thus contribute to infraspecific genetic diversity patterns. Indeed, the selection practices are culturally determined, as shown by Pressoir and Berthaud (2004) and Perales et al. (2005), who highlighted the divergence of maize selection practices among villages and ethnolinguistic groups in Mexico. Similar divergence between Aguaruna families was described by Boster (1986) for the identification of cassava cultivars, which is an activity closely related to selection. The divergence of selection practices can contribute to the observed infraspecific diversity pattern of sorghum, but a precise characterization of the agro-morphological diversity is necessary to assess the importance of this effect.

Perspectives

Our sampling strategy focused on ntora and enabled us to characterize the effect of this factor with sufficient statistical power. However, we lacked information concerning clans and age-sets. This strategy was therefore not optimal for studying the latter factors. Due to sample limitation, it was not possible to test interactions between ntora, clans, and age. Nonetheless, such interactions may exist. Each ntora is composed of major clans because of the patrilocality settlement custom, so it might be difficult to figure out which of the two factors is involved in diversity patterns. Moreover, clan identity plays a central role in alliances and exchanges. For instance, the prohibition of seed exchanges between clans related by the Gishiaro link suggests that this level of social organization may have an impact on crop diversity. Larger-scale studies are needed to address this issue. Due to the lack of information concerning age-set systems among Tharaka women, age categories were created a priori and without ethnographic bases. This choice may have affected our analysis. However, our own observations complemented by literature suggest that, since women initiation was abandoned at least 40 years ago, the age-set system has lost most of its significance (Peatrik 1999).

Uncontrolled environmental factors, such as the local edaphic variations (Bazile et al. 2008), the proximity to the local market or to an NGO, as well as the economic status of households (Rana et al. 2007), could have interfered with the ntora organization. Against this possible criticism, we would like to stress the consistency of our results for crop species and for sorghum landraces and the fact that adjacent ntora have different crop compositions. Moreover, these observations disagree with the hypothesis that the structure of information and seed exchange could result from geographical proximity.

Our study of sorghum diversity was based on local Tharaka nomenclature and we were not able to identify synonyms or homonyms as neither morphological nor genetic characterization were carried out. However, local names of sorghum landraces are indicators of cultural variations. This is illustrated by the work of Boster (1986), who showed that agreement between Aguaruna informants concerning the names of cassava cultivars is correlated to social distance between them. Common naming systems are indicators of cultural proximity and frequent exchanges between farmers. Studying rice cultivars names in Gambia, Nuijten and Almekinders (2008) observed that their uniformity reflects the intensity of seed exchange. When frequent seed exchanges of the same variety between two villages occur, that variety may obtain the same name in both villages. Farmers' variety names are exchanged as other language elements and thus can be used as a social differentiation index.

Applications for crop diversity study and conservation

As most rural populations retain their social organization, our demonstration is not limited to the Tharaka. The works of Perales et al. (2005), Pressoir and Berthaud (2004), and Baco et al. (2008) confirm that our findings among the Tharaka refer to more general mechanisms. In addition, their findings indicate that our approach could be generalized at different levels of social organization, from family units to ethnic or linguistic groups, as proposed recently (Leclerc and Coppens d'Eeckenbrugge 2012). Most studies dealing with crop diversity focus on the individual choices of farmers, emphasizing their dependence on environmental constraints (Lacy et al. 2006). They usually overlook the importance of social organization, neglecting that crops are reproduced, selected, and preferentially exchanged within socially defined groups. Bypassing the study of social organization may thus hamper the investigation of evolutionary processes involved in crop diversity patterns. We therefore recommend that research and conservation initiatives take more account of the social organization impact on crop diversity.

Our finding can be applied to conserve and sample genetic resource in situ more efficiently. Indeed, genetic resources collections generally neglect the sampling strategy at the local scale, targeting a limited number of accessible households. Furthermore, up to now, most crop collections have been based on geographic distance and agro-ecologic zonation, even though the link between crop diversity and climate is not always clear (Deu et al. 2008). While the cultural diversity and social structure of farmers has rarely been considered, our results show that it can have a strong impact on the spatial structure of diversity even at the local scale. They demonstrate that social anthropology surveys prior to inventories and sampling would help to capture crop diversity more efficiently.

Acknowledgments We thank J.I. Kamau (KARI) for his help in fieldwork, Geo Coppens d'Eeckenbrugge (CIRAD) for its useful comments on the manuscript, as well as Jean-Louis Noyer (CIRAD) and Adeline Barnaud (IRD), and the Tharaka farmers who agreed to devote time to our survey. The authors also thank two anonymous reviewers for their thoughtful comments. This study was supported by Agropolis Fondation in the framework of the ARCAD project No. 0900-001. V. Labeyrie was supported by Agropolis Fondation through a Ph.D. grant and received financial support from CIRAD for this study.

References

Agresti, A. 2007. *An introduction to categorical data analysis*. Hoboken: Wiley.

Ambler, C., ed. 1988. The Great Famine, 1897–1901. In *Kenyan communities in the age of imperialism. The central region in the late nineteenth century*. New Haven and London: Yale University Press.

Baco, M.N., G. Biaou, J.L. Pham, and J.P. Lescure. 2008. Geographical and social factors of cultivated yam diversity in northern Benin. *Cahiers Agricultures* 17(2): 172–177.

Badstue, L.B., M.R. Bellon, J. Berthaud, A. Ramírez, D. Flores, and X. Juárez. 2007. The dynamics of farmers' maize seed supply practices in the central valleys of Oaxaca, Mexico. *World Development* 35: 1579–1593.

Bandiera, O., and I. Rasul. 2006. Social networks and technology adoption in northern Mozambique. *The Economic Journal* 116(514): 869–902.

Baty, F., M. Facompré, J. Wiegand, J. Schwager, and M. Brutsche. 2006. Analysis with respect to instrumental variables for the exploration of microarray data structures. *BMC Bioinformatics* 7: 422.

Bazile, D., S. Dembélé, M. Soumaré, and D. Dembélé. 2008. Utilisation de la diversité variétale du sorgho pour valoriser la diversité des sols au Mali. *Cahiers Agricultures* 17(2): 86–94.

Bellon, M.R. 2004. Conceptualizing interventions to support on-farm genetic resource conservation. *World Development* 32(1): 152–172.

Boster, J.S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88(2): 428–436.

Brokensha, D., and J. Glazier. 1973. Land reform among the Mbeere of Central Kenya. *Africa* 43: 182–206.

Camberlin, P., J. Boyard-Micheau, N. Philippon, C. Baron, C. Leclerc, and C. Mwongera. 2012. Climatic gradients along the windward slopes of Mount Kenya and their implication for crop risks. Part 1: Climate variability. *International Journal of Climatology*. doi:10.1002/joc.3427.

Camberlin, P., V. Moron, R. Okoola, N. Philippon, and W. Gitau. 2009. Components of rainy seasons' variability in Equatorial East Africa: Onset, cessation, rainfall frequency and intensity. *Theoretical and Applied Climatology* 98(3): 237–249.

Cavalli-Sforza, L., and M. Feldman. 1981. Cultural transmission and evolution: A quantitative approach. *Monographs in population biology* 16: 1–388.

Chessel, D., A. Dufour, and J. Thioulouse. 2004. The ade4 package I. One-table methods. *R News* 4: 5–10.

Conley, T., and C. Udry. 2010. Learning about a new technology: Pineapple in Ghana. *The American Economic Review* 100(1): 35–69.

Conley, T., and C. Udry. 2001. Social learning through networks: The adoption of new agricultural technologies in Ghana. *American Journal of Agricultural Economics* 83(3): 668–673.

David, S., and L. Sperling. 1999. Improving technology delivery mechanisms: Lessons from bean seed systems research in Eastern and Central Africa. *Agriculture and Human Values* 16(4): 381–388.

Deu, M., F. Sagnard, J. Chantereau, C. Calatayud, D. Hérault, C. Mariac, J.L. Pham, et al. 2008. Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theoretical and Applied Genetics* 116(7): 903–913.

Doledec, S., and D. Chessel. 1989. Seasonal successions and spatial variables in fresh-water environments. Part 2: Distinction between seasonal and environmental-effects within floral and faunistical data. *Acta Oecologica-Oecologia Generalis* 10: 207–232.

Downing, T.E. 1992. *Climate change and vulnerable places: Global food security and country studies in Zimbabwe, Kenya, Senegal, and Chile*. Research Report–Environmental Change Unit, University of Oxford.

Fadiman, J.A. 1993. *When we began there were Witchmen: An oral history from Mount Kenya*. Los Angeles, CA: University of California Press.

Jaetzold, R., H. Schmidt, B. Hornetz, and C. Shisanya. 2007. *Farm management handbook of Kenya Vol II, Part C, East Kenya. Subpart C1, Eastern Province*. Nairobi, Kenya: Ministry of Agriculture/GTZ.

Lacy, S.M., D.A. Cleveland, and D. Soleri. 2006. Farmer choice of sorghum varieties in southern Mali. *Human Ecology* 34(3): 331–353.

Lambert, H. 1947. *The use of indigenous authorities in tribal administration: Studies of the Meru in Kenya Colony*. Cape Town, South Africa: School of African Studies, University of Cape Town.

Leclerc, C., and G. Coppens d'Eeckenbrugge. 2012. Social organization of crop genetic diversity. The G x E x S interaction model. *Diversity* 4(1): 1–32.

Manly, B.F.J. 1997. *Randomization, bootstrap, and Monte-Carlo methods in biology*. London, UK: Chapman and Hall.

Mati, B.M. 2000. The influence of climate change on maize production in the semi-humid–semi-arid areas of Kenya. *Journal of Arid Environments* 46(4): 333–344.

McElreath, R., and P. Strimling. 2008. When natural selection favors imitation of parents. *Current Anthropology* 49(2): 307–316.

McGuire, S. 2008. Securing access to seed: Social relations and sorghum seed exchange in eastern Ethiopia. *Human Ecology* 36(3): 217–229.

Middleton, J. 1953. *The central tribes of the north-eastern Bantu. Ethnographic survey of Africa*. London, UK: International African Institute.

Moehlig, W.J.G., G. Guarisma, and S. Platiel. 1980. La dialectométrie: une méthode de classification synchronique en Afrique. In *Dialectologie et comparatisme en Afrique Noire. Actes des journées de dialectologie Africaine, Ivry, 2-5 juin 1980*, ed. G. Guarisma, and S. Platiel. Paris, France: SELAF.

Nuijten, E., and C.J.M. Almekinders. 2008. Mechanisms explaining variety naming by farmers and name consistency of rice varieties in the Gambia. *Economic Botany* 62(2): 148–160.

Orlove, B.S. 1980. Ecological anthropology. *Annual Review of Anthropology* 9: 235–273.

Paillex, A., E. Castella, and G. Carron. 2009. Aquatic macroinvertebrate response along a gradient of lateral connectivity in river floodplain channels. *Journal of the North American Benthological Society* 26(4): 779–796.

Pautasso, M., G. Aistara, A. Barnaud, S. Caillon, P. Clouvel, O. Coomes, M. Delêtre, et al. 2013. Seed exchange networks for agrobiodiversity conservation. *A review. Agronomy for Sustainable Development* 33(1): 151–175.

Peatrik, A.M. 1999. *La vie à pas contés. Génération, âge et société dans les Hautes Terres du Kenya (Meru Tigania-Igembe)*. Nanterre, France: Société d'ethnologie.

Perales, H., B. Benz, and S. Brush. 2005. Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* 102(3): 949–954.

Pressoir, G., and J. Berthaud. 2004. Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity* 92(2): 95–101.

Rana, R., C. Garforth, B. Sthapit, and D. Jarvis. 2007. Influence of socio-economic and cultural factors in rice varietal diversity management on-farm in Nepal. *Agriculture and Human Values* 24(4): 461–472.

R Development Core Team. 2011. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Thrupp, L. 2000. Linking agricultural biodiversity and food security: the valuable role of agrobiodiversity for sustainable agriculture. *International Affairs* 76(2): 265–281.

Van den Broeck, K., and S. Dercon. 2011. Information flows and social externalities in a Tanzanian banana growing village. *The Journal of Development Studies* 47(2): 231–252.

Author Biographies

Vanesse Labeyrie is a PhD student at the International Center for Higher Education in Agriculture Sciences (SupAgro) in Montpellier, France. She holds an MSc in Sustainable Crop Production from SupAgro. Her research focuses on the influence of farmers' social organization on crop diversity, using a multidisciplinary approach combining population biology and anthropology.

Bernard Rono is a Rural Sociologist at the Kenya Agricultural Research Institute (KARI) in Embu, Kenya. He holds a Master of Arts Degree in Sociology from University of Nairobi, Kenya. His main research interest is the influence of farmers' social organization on seed systems.

Christian Leclerc PhD is an ecological anthropologist at the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) in Montpellier, France, and research coordinator of the ARCAD program in Kenya. He holds a PhD in Anthropology from Paris University. His main research interest is to develop multidisciplinary approaches combining crop population genetic and social anthropology to understand how farmers' cultural and crop genetic differentiation can be linked over time.

CHAPITRE II- ECHELLE DES GROUPES ETHNOLINGUISTIQUES

ARTICLE: INFLUENCE OF ETHNOLINGUISTIC DIVERSITY ON SORGHUM EVOLUTION IN SUBSISTENCE FARMING SYSTEMS IN EASTERN KENYA

V. Labeyrie, M. Deu, A. Barnaud, C. Calatayud, M. Buiron, P.Wambugu, S. Manel, J-C. Glaszmann, C. Leclerc

Accepté sous révisions mineures par *PLOS one*

INFLUENCE OF ETHNOLINGUISTIC DIVERSITY ON SORGHUM EVOLUTION IN SUBSISTENCE FARMING SYSTEMS IN EASTERN KENYA

V. Labeyrie[1], M. Deu[1], A. Barnaud[3], C. Calatayud[1], M. Buiron[1], P.Wambugu[2], S. Manel[4,5], J-C. Glaszmann[1], C. Leclerc[1]

[1]CIRAD, UMR AGAP, F-34398 Montpellier, France

[2] KARI, National Genebank of Kenya, Nairobi, Kenya

[3] IRD, UMR DIADE, Montpellier, France

[4] Université Aix-Marseille / IRD, UMR LPED, Marseille, France

[5]CIRAD, UMR AMAP, F-34398 Montpellier, France

Corresponding author: Vanesse Labeyrie, vanesse.labeyrie@gmail.com

ABSTRACT

Understanding the effects of actions undertaken by human societies on crop evolution processes is a major challenge for the conservation of genetic resources. This study investigated the mechanisms whereby social boundaries associated with patterns of ethnolinguistic diversity have influenced the on-farm evolution of sorghum. Social boundaries limit the diffusion of planting material, practices and knowledge, thus shaping crop diversity *in situ*.

To assess the effect of social boundaries, this study was conducted in the contact zone between the Chuka, Mbeere and Tharaka ethnolinguistic groups in eastern Kenya. Sorghum varieties were inventoried and samples collected in 130 households. In all, 297 individual plants derived from seeds collected under sixteen variety names were characterized using a set of 18 SSR molecular markers and 15 morphological descriptors. The structure of genetic diversity was characterized using both a Bayesian assignment method and distance-based clustering. Principal Coordinates Analysis was used to describe the structure of the morphological diversity of the panicles. The distribution of the varieties and the main genetic clusters across ethnolinguistic groups was described using a non-parametric MANOVA and pairwise Fisher tests.

The spatial distribution of landrace names and the overall genetic spatial patterns were significantly correlated with ethnolinguistic partition. However, the structure based on molecular makers did not discriminate the short-cycle landraces despite their morphological distinctness. The cases of two improved varieties highlighted possible fates of improved materials. The most recent one was often given the name of local landraces. The second one, that was introduced a dozen years ago, displays traces of admixture with local landraces with differential intensity between ethnic groups. The patterns of congruence or discordance between the nomenclature of farmers' varieties and the structure of both genetic and morphological diversity highlight the effects of the social organization of communities on the diffusion of seed, practices, and variety nomenclature.

INTRODUCTION

Identifying factors involved in crop evolution is of great importance for genetic resource conservation and crop improvement. Crop genetic diversity patterns result from selection, migration and genetic drift processes which are strongly influenced by human action. Recent studies combining linguistic, archeological and genetic data have unraveled the past domestication and diversification processes of crops such as banana [1] and sweet-potatoes [2], on a large time-space scale, by linking global diversity patterns to human migrations. However, the evolution of crops is still ongoing in smallholder farming systems under the pressure of agro-ecological conditions and farmers' management practices [3]. The study of these processes at the community scale is complementary to large time-space approaches and contributes to the general understanding of the *in situ* genesis of crop genetic patterns.

Social boundaries contribute to the evolution of crop populations both directly, by determining seed flows, and indirectly, by inducing the divergence of seed selection practices [4]. Previous studies notably showed that the ethnic organization of farming communities plays an important role in differentiating the domesticated populations of allogamous crops [5], vegetatively-propagated crops [6] and animals [7].

Sorghum (*Sorghum bicolor* L. Moench) is an annual cereal extensively cultivated in smallholder farming systems because of its ability to grow under harsh climatic conditions. De Wet et al. [8] and Harlan et al. [9] suggested that the spatial distribution of sorghum botanical races in Africa was related to that of the ethnic groups, but this hypothesis was not further tested. In a study undertaken in Niger, Deu et al. [10] suggested that human ethnic diversity has probably a greater impact on sorghum diversity than recent environmental constraints. However, the authors were not able to assess this hypothesis as the spatial localization of the different ethnic groups in Niger corresponded to different agro-ecological regions. Thus, deciphering how the social organization of farmers affects the structure of sorghum diversity remains a challenge.

This article addresses the role of social boundaries in sorghum evolution and diversification processes. It set out to identify the mechanisms whereby social boundaries, associated with ethnolinguistic diversity patterns, shape sorghum genetic diversity on-farm. To study only the main effect of social boundaries, this study focused on an ethnolinguistic contact zone where both geographical distance between ethnic groups and agro-ecological variability were limited. If social boundaries do not limit seed-mediated gene flows and the diffusion of selection practices, then no relation should be observed between ethnic diversity patterns and both the genetic and morphological structure of sorghum diversity. Otherwise, it would reflect the impact of social boundaries on the evolutionary mechanisms that shape sorghum diversity *in situ*.

Farmers' varieties are relevant units for studying on-farm crop diversity as they are consciously defined and named by farmers for management, selection, seed exchanges and knowledge transmission purposes [11]. Farmer's nomenclature and taxonomy of crop varieties is a marker of knowledge diffusion and exchanges across communities [12], while the structure of the genetic and morphological diversity of crop populations reflects gene flows and selection forces [5]. This study thus used molecular markers to estimate genetic diversity and compared the spatial distribution of varieties with genetic spatial patterns according to ethnic groups. These patterns were then discussed regarding the congruence between farmer's varieties and the structure of their genetic and morphological diversity. Combining these three approaches enabled us to investigate the influence of social boundaries on the evolutionary mechanisms that shape sorghum diversity *in situ*.

Clarifying the effect of social boundaries on crop evolutionary mechanisms has important applications for crop genetic resource collection, characterization and conservation. It contributes to increasing the overall understanding of on-farm crop diversification processes. By highlighting the overall role of societies in shaping crop diversity, it stresses the relevance of multidisciplinary approaches for crop genetic diversity studies.

MATERIAL AND METHODS

Ethics Statement

This was a collaborative study between CIRAD and KARI-National Genebank of Kenya, who were the Kenyan national partners. KARI has the national mandate for the collection and conservation of all plant genetic resources and documentation of all accompanying information. It is under this framework and mandate that the study was mounted and hence no specific permission was required. This notwithstanding, the mandate as well as the importance of the study, both nationally and globally, was explained to the farmers and concurrence was sought before undertaking the study activities. We confirm that sorghum, the study crop, is neither endangered nor protected.

Study site: agro-ecological conditions and ethnic organization

This study was conducted on the eastern slope of Mount Kenya (0°24'27.88"S, 37°46'35.59"E), in an ethnolinguistic contact zone between Chuka, Tharaka and Mbeere groups (Figure II-1). The three ethnolinguistic groups (hereafter ethnic) live within the same agro-ecological zone, as defined by Jaetzold et al. [13]. The study site was 15 km-square, and the elevation ranged from 810 to 946 m above sea level, so rainfall and temperature variability was limited. The area is semi-arid with a mean temperature ranging between 21.7°C and 23.9°C. The mean rainfall is about 700-800 mm per year, distributed across two rainy seasons with the Long Rains occurring from March to May and the Short Rains from October to December [14]. Soil characteristics are homogeneous in the area occupied by the three ethnic groups, corresponding to well drained Ferralsols, with a loamy-sand texture and moderate fertility [13].

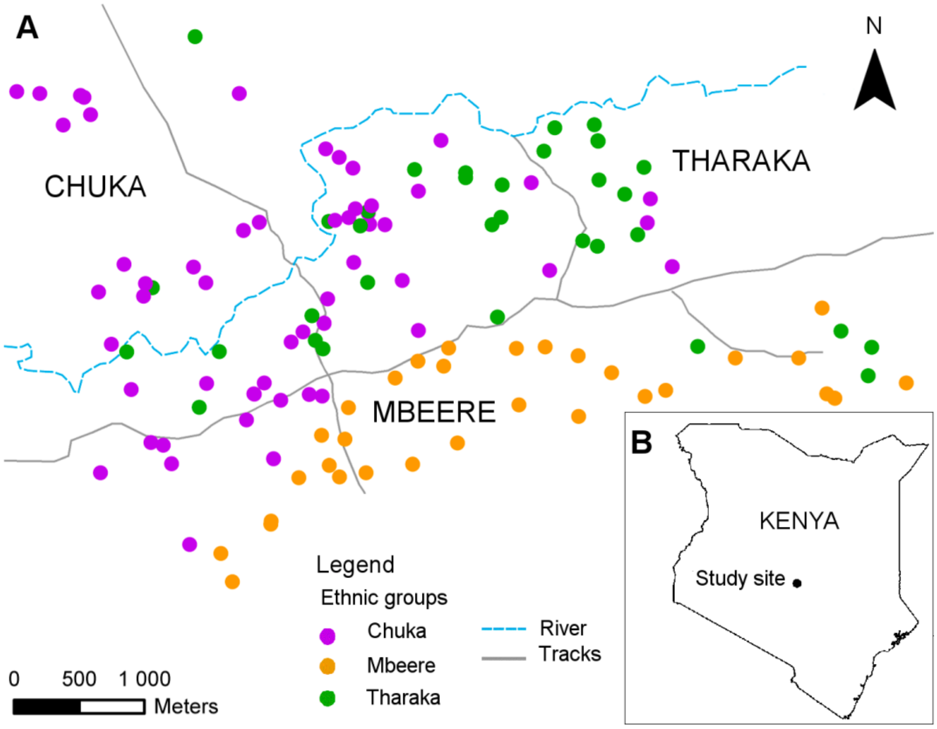


Figure II-1. Study site location. Location of the farms where sorghum samples were collected. Colors correspond to the ethnic identity of the male house-head.

The three ethnic groups, Chuka, Tharaka and Mbeere, migrated to the study area by the end of the 19th century, either because of a population increase or because of recurrent drought [15]. Social boundaries exist between Chuka, Tharaka and Mbeere groups as revealed by their distinct ethnic identity, and their current cultural and linguistic differences [16,17]. The Mbeere are closely related to the Embu group [18], while the Chuka and Tharaka are related to the Meru group. The Mbeere and Chuka had conflictual mutual relationships in the past [19], while the Chuka and Tharaka maintain strong social ties and consider they are kin [20]. Intermarriage is usual between the Chuka and Tharaka, while it is very uncommon between the Mbeere and Chuka or Tharaka (unpublished data). Men usually settle near their father's compound once they get married. The residence is thus patrilocal [16]. The three ethnic groups present a non-random spatial distribution. The Mbeere households are located in the southern part of the study area, the Tharaka mostly on the north-eastern side, and the Chuka on the north-western side (Figure II-1). Consistently with the social relationships between groups depicted above, a clear spatial boundary was found between the Mbeere and both the Chuka and Tharaka, while the Chuka and Tharaka appeared to be spatially more mixed.

The three ethnic groups manage low-input cropping systems that harbor high specific and infra-specific crop diversity. Cropping systems are based on cereals and legumes that are usually intercropped. Sorghum (*Sorghum bicolor*), cowpeas (*Vigna unguiculata*), maize (*Zea mays*), mungo bean (*Vigna radiata*) and pearl millet (*Pennisetum glaucum)* are the main crop species grown in the area. Sowing is done either by hand-dibbling or by drilling, while plowing is done with animals. The different sorghum varieties are either grown in separate plots or mixed together within farmers' fields. Improved varieties, mainly disseminated by the extension services of the Kenyan Ministry of Agriculture, have also been adopted by the farmers. They are cultivated together in the same field with the local varieties (or landraces). Farmers distinguish between short-cycle varieties that can be grown either from October to January or from March to June, and long-cycle varieties that are subjected to the ratooning practice [21]. These long-cycle varieties are sown in October, the vegetative part being cut before the grains are mature to stimulate regrowth from basal buds, and panicles are finally harvested in July.

Data collection

Sorghum inventory and germplasm collection

The field work consisted of two stages. A preliminary survey was carried out to estimate the frequency of varieties in the three ethnic groups. The strategy for on-farm germplasm collection was then based on that estimation of diversity, as it aimed to represent the diversity and frequencies of each variety in each ethnic group.

The preliminary inventory survey was conducted in both January (Short Rains cropping season) and June 2011 (Long Rains cropping season), just before harvesting and prior to germplasm collection. The inventory of sorghum varieties was based on the local names as reported by women farmers who were in charge of sorghum selection in each of the 124 households surveyed. Indeed, grain crop farming comprising seed sowing, harvesting, selection and trading is ensured by women ([16], personal observation). The ethnolinguistic identity of male house-heads or single women was also recorded.

Sorghum panicles were then collected from 130 households in January (56 households) and in July (96 households), with 22 households providing panicles in both seasons. Half of these 130 households were visited during the preliminary surveyed described above. About half of the total number of households in the area was thus sampled, insuring a good representativeness. Sixty households belonged to the Chuka ethnic group, 35 to the Mbeere and 35 to the Tharaka. In order to be representative of the sorghum population of each ethnic group, all the varieties that were previously inventoried were collected. Depending on its occurrence frequency previously estimated over the 124 households, from one to twenty panicles of each variety were sampled from each ethnic group, without exceeding two individual panicles per variety for a given household. Sixteen different varieties were collected out of the seventeen inventoried because one was not found at the time of harvest. The mean number of varieties collected per household was 1.5 (min: 1, max: 6). It was similar across ethnic groups, as well as the mean number of samples of each variety collected per household (Table II-S1). The fraction of households where each variety was collected for the study of genetic diversity was correlated to the fraction of households where each variety was previously inventoried (Linear regression R^2: 0.77, Figure II-S1). Only the three improved varieties (*Kaguru, Gadam* and *Serendo*) were sampled in fewer households as compared to their occurrence frequency estimated from the inventory. In all, 290 samples were collected on-farm after harvest, each consisting of a single panicle. About 47 % of the individual plants were sampled from the Chuka ethnic group, 30 % from the Tharaka and 23 % from the Mbeere. Information concerning the names, the origin (local or improved) and the cycle length of each sampled panicle was recorded from women house-heads.

DNA extraction and SSR genotyping

Seeds from the 290 panicles collected on-farm were sown in an experimental field *in situ*, and the leaves of one sibling randomly chosen for each mother plant were collected and stored on silicagel. Leaves from seven individuals grown from certified seeds of the improved varieties *Serendo* and *Gadam* were also collected as controls. In total 297 individual plants were thus used for the genetic diversity study.

Twenty-two pairs of primers were selected for their high polymorphism in central Kenya (unpublished data) and West Africa [10]; twenty of them were part of a set of reference microsatellite markers proposed by Billot and colleagues ([22], http://sat.cirad.fr/sat/sorghum_SSR_kit/). Loci were distributed over the 10 chromosomes. DNA was extracted from dried leaves and the polymerase chain reaction amplifications were done following the procedure described previously [10,22]. The fluorescent dye–labeled PCR products from differentially labeled primers and with non-overlapping size were pooled and subjected together to capillary electrophoresis using a 24-capillary 3500xL System (Applied Biosystems®). GeneMapper v 4.1 (Applied Biosystems®) was used for genotype scoring. GeneScan™ 600 LIZ® Size Standard v2.0 was added to each well, and three control samples were used to facilitate allele scoring [22]. Genotyping was done at the Montpellier Languedoc-Roussillon Genopole platform located on the CIRAD campus in Montpellier (France).

Four markers presenting either a high number of missing data, or low polymorphism (at a 99 % threshold) were discarded from the analysis, so eighteen markers were kept, covering 9 chromosomes out of 10. The percentage of missing data for the 18 markers kept was 1 %.Table II-S2 provides a list of these18 markers and their description.

Panicle morphological characterization

Fifteen qualitative morphological traits were measured on the panicles of the 297 individuals that were genotyped (Table II-S3). Eight morphological descriptors were selected from the IPGRI descriptors [23] and were completed by seven additional descriptors for seeds and glumes characteristics that showed variability on the sorghum collected in our study area. Descriptors covered the characteristics of the whole panicle (panicle shape), seeds (color, presence of sub-coat, pericarp thickness, shape, endosperm texture and shattering) and glumes (color, adherence, covering, opening, texture, hairiness, awning and transversal wrinkle). Only qualitative traits were kept for these analyses because they are stable characteristics on which farmers base their nomenclature and classification [24]. Multiple characterizations of randomly sampled individuals enabled to check for morphological trait scoring consistency.

Data analysis

Comparing sorghum assemblages between ethnic groups

We characterized each household by its sorghum assemblage, which is the panel of co-occurring sorghum varieties that are cultivated by the household. The differentiation of sorghum assemblages across ethnic groups was tested using a non-parametric Multivariate Analysis of Variance (perMANOVA, [25]). The PerMANOVA was implemented under the *adonis* function in the R package *vegan* [26]. The presence/absence matrix for sorghum varieties in each household was transformed into a distance matrix using the Bray-Curtis index [27]. The *adonis* function partitions the distance matrix according to grouping factors (ethnic groups) and compares the sum of squared distances within groups (which is the sum of squared distances from individual replicates to their group centroid) and between groups (which is the sum of squared distances from group centroids to the overall centroid). A pseudo F-ratio is then computed and compared to its distribution under the null hypothesis simulated using 4000 random permutations of the raw data. Pairwise Fisher exact tests implemented in the R package *fmsb* [28] were then used to compare the occurrence frequencies of the most frequent varieties across the Chuka, Mbeere and Tharaka ethnic groups. The calculation of p-values was corrected for multiple comparisons using the False Discovery Rate (FDR) procedure [29] implemented in the *p.adjust* function.

Genetic structure of sorghum populations

The genetic diversity of sorghum populations sampled in each ethnic group was assessed using several indexes. The observed number of alleles and the observed heterozygosity were calculated using GENETIX 4.05.2 software [30]. The allelic richness corrected for sample size [31], the unbiased gene diversity (expected heterozygosity) corrected for small sample size [32], and the F_{IS} [33] of multi-locus genotypes were estimated using the procedures implemented in FSTAT 2.9.3.2 software [34].

Two complementary approaches, Bayesian clustering and Neighbor-Joining tree, were used to assess the global structure of genetic diversity without defining a-priori populations. First, the genetic structure of sorghum populations was characterized using the Bayesian clustering algorithm implemented in STRUCTURE 2.3.3 software [35] and run on the Bioportal server (http://www.bioportal.uio.no). The admixture model with correlated allele frequencies was used, assuming that the genome of each individual resulted from the mixture of *K* ancestral populations. The estimated proportions of each individual's genotype originating from each of the *K* ancestral populations (*q*) was calculated for *K* ranging from 2 to 10 ancestral populations (or clusters), with twenty runs for each *K* value. The burn-in period was set at 500 000 and 1 000 000 iterations were performed. The criterion suggested by Evanno et al. [36], based on the rate of change in the log probability of data between successive *K* values, was used to determine the most likely number of clusters (*K*). Second, a Neighbor-Joining tree [37] was built from a simple matching genetic dissimilarity matrix [38] using Darwin V5 software [39]. The results of both the Bayesian clustering and Neighbor-Joining methods were then compared to check for the consistency of the clusters. This led to what we refer to as an MMb (molecular-marker-based) classification scheme.

For further analysis, individuals whose estimated proportion of genome originating from one population (*q*, hereafter admixture coefficient) was below a 0.8 threshold were considered as resulting from admixture between the populations. Individuals whose *q* value was equal to or above 0.8 for a population were assigned to that population (hereafter cluster). To explain the MMb genetic structure, the assignment of individuals to clusters thus defined was crossed with information concerning their origin and cycle length as reported by farmers during the collection of samples *in situ*. The occurrence frequencies of each MMb genetic cluster were then compared across ethnic groups using Pearson's Chi-squared test, and pairwise Fisher exact tests with False Discovery Rate (FDR) correction for multiple comparisons.

The mean unbiased allelic richness, gene diversity and F_{IS} of sorghum populations were compared between ethnic groups using paired Pairwise Wilcoxon tests with False Discovery Rate (FDR) correction implemented in R (package *stats, pairwise.wilcox.test* function). Pairwise F_{ST} [33] were computed between the sorghum populations collected in the three ethnic groups, and the significance of differences was assessed using a permutation procedure (3 000 permutations) and corrected using a sequential Bonferroni procedure [40]. A multilocus G-test of differentiation, known to be accurate for measuring the genetic differentiation between populations with unbalanced sizes [41], was used to test the genetic differentiation between the populations sampled in each ethnic group (10 000 permutations). Calculations were carried out using FSTAT 2.9.3.2. Pairwise G-tests implemented in GENEPOP 4.2 [42] were used to estimate the genotypic differentiation between pairs of populations and p-values were corrected for multiple tests using FDR correction (*p.adjust* function in the R package *stats*).

Morphological structure of sorghum populations

To describe the structure of individual panicle morphological diversity, a dissimilarity matrix was computed on the basis of the 15 morphological traits coded through a total of 43 modalities using the simple matching index [38]. The morphological similarity between individuals was then assessed using a Principal Coordinates Analysis (PCoA) using the R package *ade4*.

RESULTS

Differences in variety assemblages across ethnic groups

On the basis of their local names, seventeen different varieties were inventoried among the 124 households visited during both the January and June surveys (Chuka: 14, Mbeere: 10, Tharaka: 14) out of which 9 were shared by the three ethnic groups. The mean number of varieties inventoried in both cropping seasons per household was similar across ethnic groups (2.77, SE: 0.17 for the Chuka, 2.65, SE: 0.17 for the Mbeere, 3.02, SE: 0.21 for the Tharaka). The most frequent variety was *Kaguru,* (76% of the households), followed by *Gadam* (48% of the households), both of which are improved varieties. *Ngirigacha, Mugeta*, *Mbura imwe, Muruge mbura ciiri*, and *Muruge mbura imwe* were the most frequent local varieties (landraces) (Figure II-2).

The non-parametric perMANOVA showed that sorghum variety assemblages differed significantly between ethnic groups (Table II-S4), even though the ethnic partition explained a limited part of variability (pseudo-$F_{2,121}$= 4.971, p-value = 0.0002, R^2 = 0.076). Pairwise Fisher exact tests confirmed that the frequency of three out of the five most frequent landraces differed significantly between ethnic groups, while the frequency of improved varieties (*Gadam*, *Kaguru* and *Serendo*) did not differ significantly between ethnic groups. *Muruge mbura imwe* and *Mugeta* were significantly less frequent in the Mbeere group than in the Chuka and Tharaka groups while *Ngirigacha* was significantly more frequent.

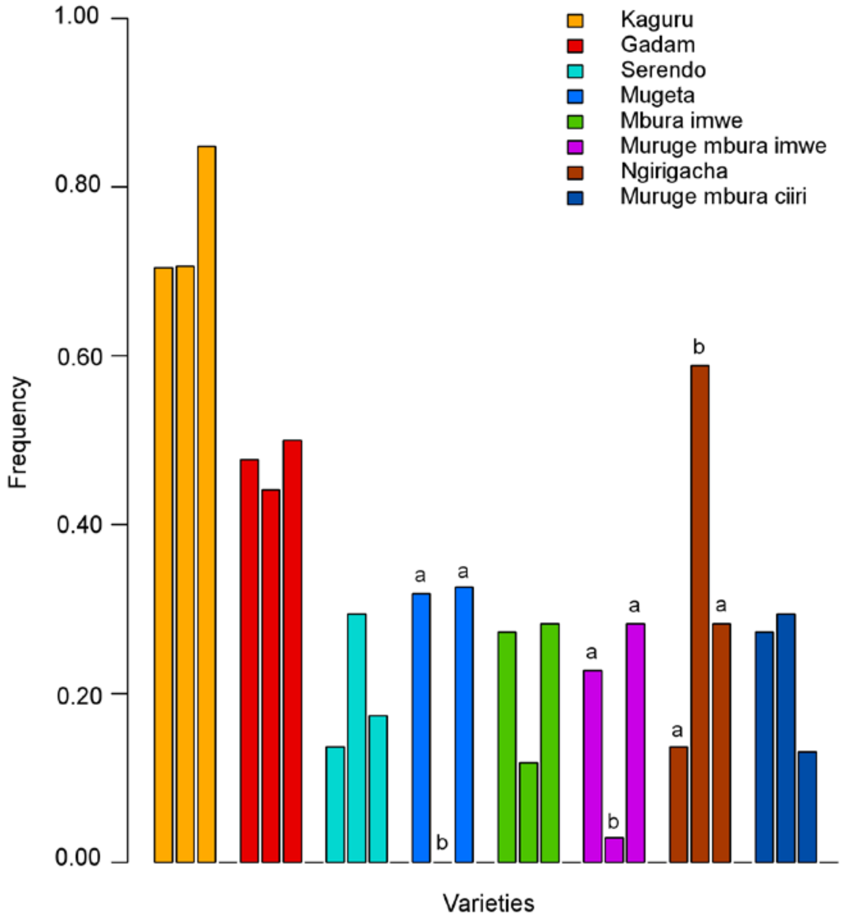


Figure II-2. Frequency of the eight major varieties in each ethnic group. The vertical axis displays the percentage of farms where the variety was cultivated. Ethnic groups are present in the following order for each variety: Chuka, Mbeere, Tharaka. The letters (a,b) on top of the bars indicate the statistical significance of differences (Fisher test) at a 5% level after correction for multiple testing (FDR). For a given variety, ethnic groups with the same letter did not present significantly different frequencies.

Global structure of the diversity among sorghum varieties

The most likely number of populations (*K*) identified by STRUCTURE was *K* = 4. Indeed, the log-probability of data increased up to *K* = 4, where it reached a plateau. This was congruent with Evanno's ΔK curve which presented a clear peak for *K* = 4. The populations (clusters) inferred by STRUCTURE for *K* = 4 (Figure II-3.A) corresponded to distinct groups on the Neighbor-Joining tree (Figure II-4.A). Cluster A and C were distinct and showed higher genetic uniformity than cluster B and D. Most of the individuals sampled (88 %)

showed an admixture coefficient (*q*) above or equal to $q = 0.8$, and they were thus assigned to the corresponding cluster. The remaining 12 % of the individuals were considered to result from admixture between clusters.

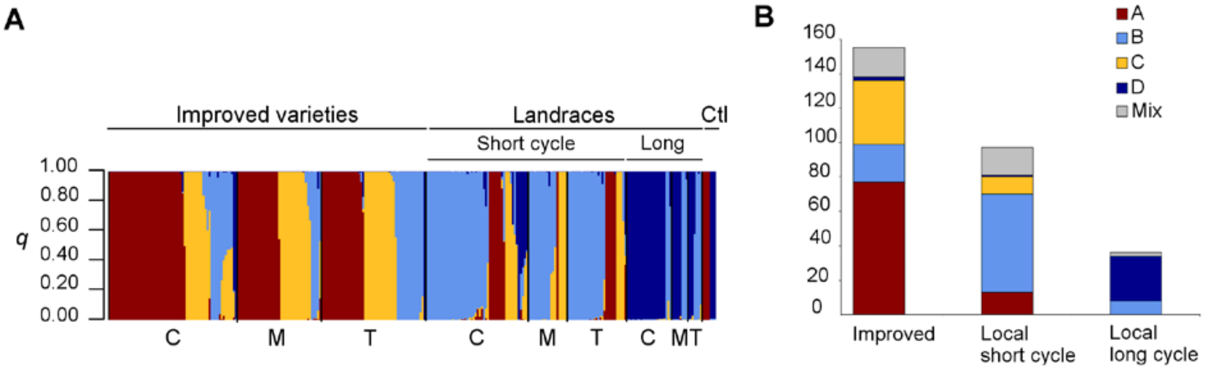


Figure II-3.

A. Cluster assignment of 297 sorghum individuals estimated using STRUCTURE for K=4. The genome of each individual is represented by a vertical line, which is partitioned into K colored segments that represent the admixture coefficient (*q*), i.e the estimated proportion of membership of its genome in each of the K clusters. Thick black lines separate the individuals identified by farmers as improved varieties, short-cycle landraces or long-cycle landraces, and control individuals (Ctrl), as labeled above the figure. Thin black lines separate individuals sampled in the different ethnic groups (Chuka: C, Mbeere: M, Tharaka: T, as labeled below the figure. The figure shown is based on the highest probability run at K=4. Red: cluster A, light blue: cluster B, yellow: cluster C, dark blue: cluster D.

B. Percentage of individuals classified according to their origin and cycle length (farmer information) assigned to each MMb genetic cluster. The vertical axis indicates the number of individuals assigned to each cluster. Individuals were assigned to a cluster when their estimated admixture coefficient (*q*) for this cluster was equal to or over 0.8. Admixed individuals are represented in gray.

The diversity in terms of morphological characters could be summarized by the PCoA (Figure II-4.B). The first axes accounted for 29 and 13 % of the variation, respectively. Axis 1 isolated a clear group on its positive side (II), corresponding to the major share of individuals assigned to MMb cluster D, while the rest of the individuals were broadly distributed along axes 1 and 2. Individuals assigned to MMb clusters A and C displayed narrow distributions indicating uniform morphological types, which is consistent with their improved origin and recent introduction. Individuals assigned to cluster C formed a distinct morphological group (I), discriminated on the third axis of the PCoA (expressing 10.7 % of the total variation, data not shown). Individuals assigned to MMb cluster B displayed a broad distribution, reflecting high variability and continuous distribution between diverse morphotypes, some being similar to individuals from cluster A. Nevertheless, part of the individuals assigned to the MMb cluster B clustered in a separate morphological groups (III).

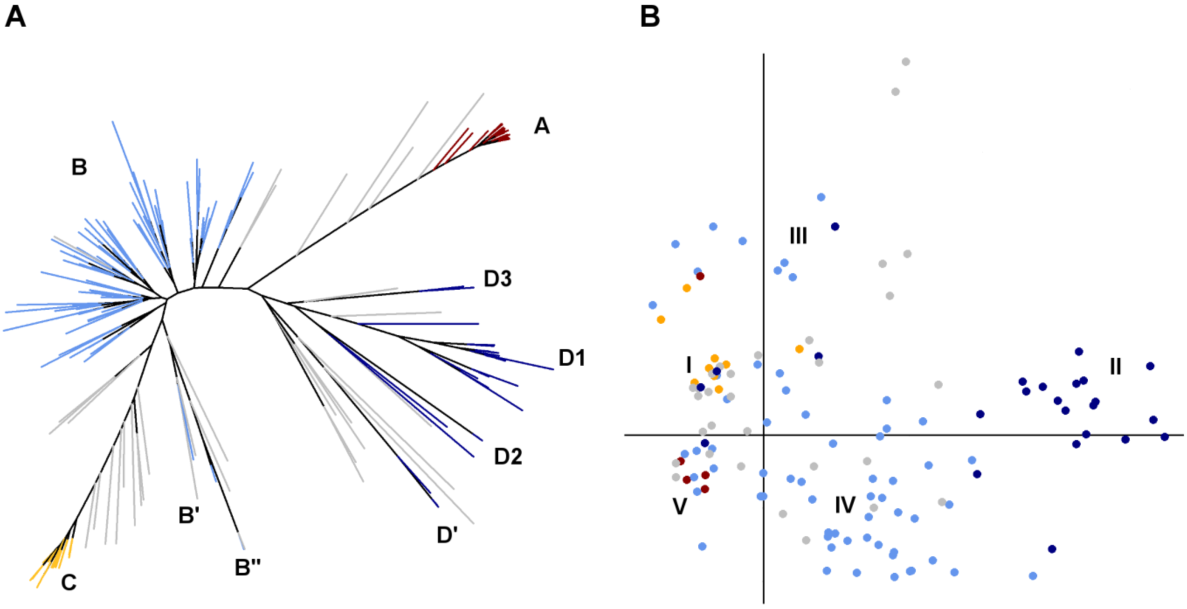


Figure II-4.

A. Neighbor-Joining tree based on 18 SSRs among sorghum plants using the simple matching index. Genetic clusters inferred by STRUCTURE are displayed using different colors: Cluster A: red, B: light blue; C: yellow, D: dark blue. Sub-clusters are identified by letters followed by a number.

B. Plot of the two first axes of the Principal Coordinates Analysis (PCoA) based on 15 panicle morphological traits using the simple matching index. The first axis (x) expresses 29.3 % of the total variation and the second axis 13.2 %. The main morphological groups are indicated by roman numerals and the MMb genetic assignment of individuals for K= 4 is displayed with the same colors as in figure II-4.a.

The MMb genetic structure was found to be strongly related to the improvement status of the germplasm – improved varieties or local landraces, and by differences in growth-cycle length (Figure II-3.B). Individuals assigned to the uniform clusters A and C were mostly improved varieties introduced by the extension services, while individuals assigned to the broader clusters B and D were mainly classified by farmers as local landraces. Moreover, almost all individuals assigned to cluster D were identified by farmers as long-cycle varieties (ratoon) while those individuals assigned to clusters A, B and C were mainly identified as short-cycle varieties.

Despite this global coherence, the variety names used by farmers showed some divergence from the MMb genetic classification. Twenty-two percent (22 %) of the individual plants that were identified as long-cycle landraces by farmers during the collection were assigned to cluster B by STRUCTURE (Figure II-3.B). A substantial proportion of individuals identified by farmers as short-cycle landraces were assigned to clusters A (13 %) or C (10 %). Conversely, 14 % of the individuals identified by farmers as improved varieties were assigned to cluster B. Indeed, young farmers may consider as local the varieties that were introduced a long time ago, perhaps before they began farming. In terms of morphological diversity, it is noteworthy that some of the individuals assigned to MMb genetic clusters A and B displayed morphological similarity (Figure II-4.B.), which may induce possible confusion in naming the recent improved variety and the local landraces (homonymy).

Genetic differentiation of sorghum populations across ethnic groups

Various indexes were used to characterize the diversity displayed within each ethnic group (Table II-1). The unbiased gene diversity estimates (H_e) of Chuka and Tharaka sorghum populations were significantly higher than that of the Mbeere (Wilcoxon test: p-value < 0.01). Similar results were found for the unbiased allelic richness. F_{IS} was very high in the three groups, yet it was significantly lower in the Chuka population as compared to those of both the Tharaka and the Mbeere (Wilcoxon test: p-value < 0.05 for both pairwise comparisons), in relation with the higher heterozygosity found within the Chuka sorghum population (0.033) compared to the other two populations (0.022 for the Mbeere and 0.023 for the Tharaka).

Table II-1. Summary of the genetic polymorphism indexes of sorghum individuals sampled in each ethnic group. N_i: number of samples, N_{hh}: number of households, $N_{Al.}$: Mean number of observed alleles over the 18 loci, $R_{Al.}$: unbiased allelic richness corrected for sample size, H_e: unbiased gene diversity, H_o: observed heterozygosity, F_{IS}: fixation index. The letters (a, b) next to the R_{Al}, H_e and F_{IS} values indicate the statistical significance of their differences between ethnic groups (Wilcoxon test) at a 5 % level after correction for multiple testing (FDR). For a given index, ethnic groups with the same letter did not present significant differences.

Ethnic group	N_i	N_{hh}	$N_{Al.}$	$R_{Al.}$	H_e	H_o	F_{IS}
Chuka	135	60	6.8	6.1^a	0.590^a	0.033	0.943^a
Mbeere	68	35	4.7	4.6^b	0.544^b	0.022	0.959^b
Tharaka	87	35	6.1	5.9^a	0.569^a	0.023	0.961^b
Total	290	130	7.7	7.7	0.574	0.028	0.952

An exact G-test of genetic differentiation of sorghum across ethnic groups was significant (p-value = 0.0205). The differentiation was clearer (G-test p-value = 0.0026) when removing from the analysis the individuals assigned to cluster A, derived from the recent introduction of the *Gadam* improved variety. The Pairwise G-tests showed that genetic differentiation was highly significant between the sorghum populations of the three groups, being highest between the Chuka and both the Tharaka (p-value < 0.0001) and Mbeere (p-value = 0.0002) populations and lowest between the Tharaka and Mbeere populations (p-value = 0.0083). The F_{ST} values between the sorghum populations of the three ethnic groups were low: 0.027 between the Chuka and Mbeere sorghum populations and 0.019 between the Chuka and Tharaka populations, both significant at a 5 % confidence level; 0.010 and non significant between the Mbeere and Tharaka populations.

The spatial distribution of the four MMb genetic clusters was not uniform (Figure II-5.B) and they were not evenly distributed across the three ethnic groups (Table II-2). Pearson's Chi-squared test led to rejecting independence between the genetic clusters and the ethnic groups (p-value = 0.003).

Table II-2. Number of individuals sampled in each ethnic group and assigned to each MMb genetic cluster. Individuals with a *q* value equal to or above the threshold of 0.8 for a cluster were assigned to that cluster. The Chi-Square statistics and p-value compare, for each MMb cluster, the observed and the expected frequencies under the null hypothesis of independence. For each cluster, the letters indicate the statistical significance of the differences in its frequency between ethnic groups (Fisher test) at a 5 % level after correction for multiple testing (FDR). For a given cluster, ethnic groups with the same letter did not present significant differences.

MMb	Chuka	Mbeere	Tharaka	Total	Chi^2	P-value
A	44 (33 %) ^{a}	21 (31 %) ^{a}	25 (29 %) ^{a}	90 (31 %)	0.37	0.832
B	36 (27 %) ^{a}	18 (27 %) ^{a}	34 (39 %) ^{a}	88 (30 %)	4.49	0.106
C	12 (9 %) ^{a}	17 (25 %) ^{b}	18 (21 %) ^{b}	47 (16 %)	10.5	0.005
D	21 (15 %) ^{a}	5 (7 %) ^{ab}	4 (4 %) ^{b}	30 (11 %)	7.7	0.021
Mix	22 (16 %) ^{a}	7 (10 %) ^{a}	6 (7 %) ^{a}	35 (12 %)	4.7	0.097
Total	135 (100 %)	68 (100 %)	87 (100 %)	290 (100 %)		

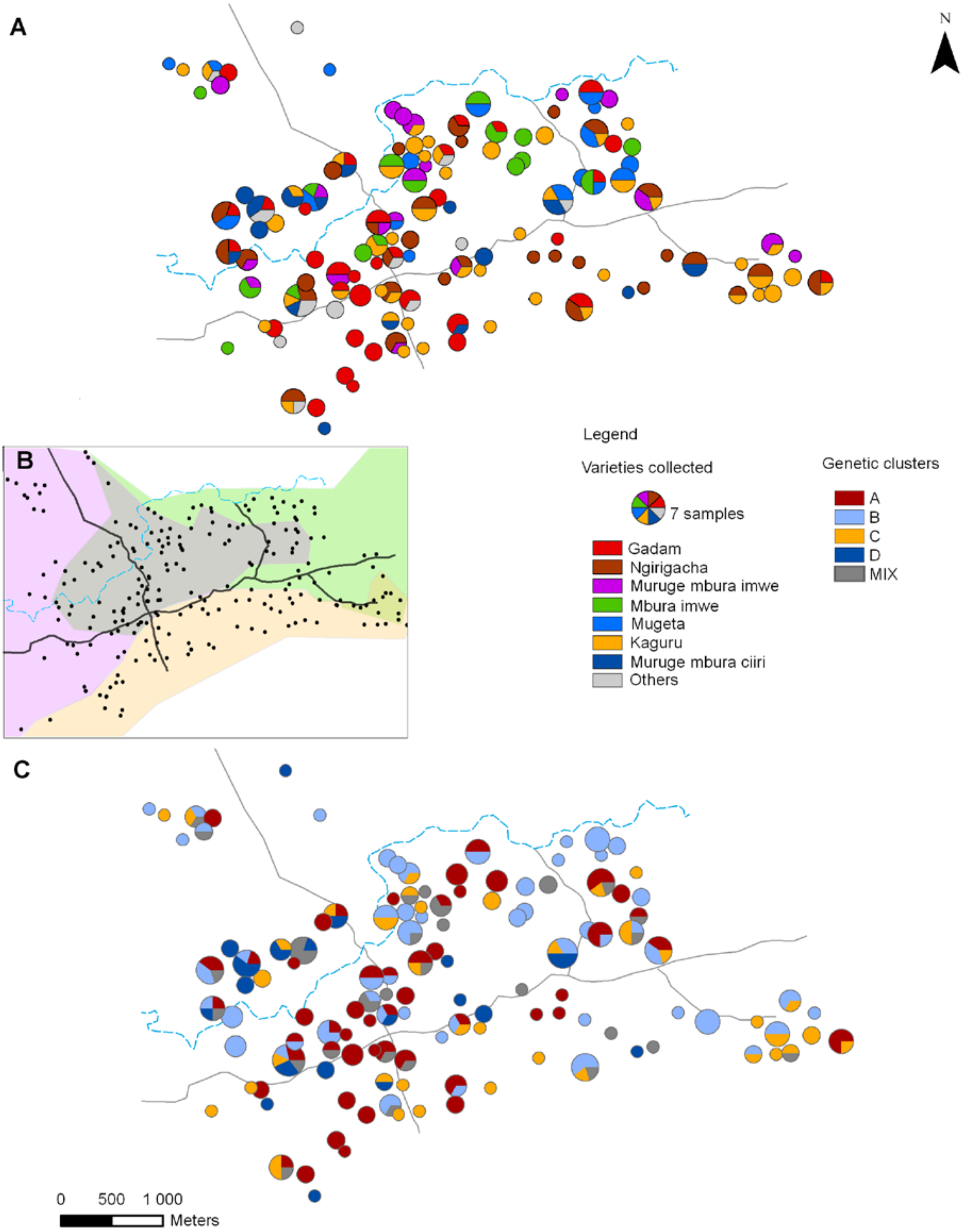


Figure II-5. A. Map of the named varieties collected in each ethnic group. Pie charts represent the number of samples of each variety collected in each household. The size of each circle is proportional to the number of individuals sampled.
B. Location of the ethnic groups. Purple: Chuka, Green: Tharaka, Orange: Mbeere. The grayish area is inhabited by both the Chuka and Tharaka groups.

C. Map of the number of sorghum individuals in each household assigned to each of the four MMb genetic clusters. Individuals were assigned to a cluster if their estimated genome fraction to that cluster, i.e. admixture coefficient (*q*), was higher than 0.8.

Correspondence between global diversity patterns and farmers’ variety names

The MMb cluster A was clearly separated from the others, as illustrated by the Neighbor-Joining tree. It included the four control individuals stemming from certified seeds of the *Gadam* improved variety, which has been disseminated in the area since 2009. Most of the other individuals assigned to cluster A were identified by farmers as *Gadam* (Chuka: 50 %, Mbeere: 71 %, Tharaka: 48 %), confirming the cluster A – *Gadam* correspondence. Yet cluster A also included 46 % of varieties collected under other names, mainly *Ngirigacha* (Chuka: 29 %, Mbeere: 24 %, Tharaka: 28 %) and *Mbura-imwe* (Chuka: 11 %, Tharaka: 16 %). As a result, cluster A was distributed throughout the study area and its spatial distribution appeared more uniform than that of the individuals designated by farmers as *Gadam* (Figure II-5).

The MMb cluster C was also clearly separated from the others, yet with an array of individuals that appeared as intermediates (along the branch of the Neighbor-Joining tree). The major share of the individuals assigned to cluster C was identified by farmers as an improved variety called *Kaguru*, which was introduced in the area about ten years ago (Chuka: 83 %, Mbeere: 94 %, Tharaka: 100 %). *Kaguru* individuals originated uniformly from the study area (Figure II-5.A) and in the three ethnic groups (Figure II-S1), but were less frequent in the Chuka area (Figure II-5.C). This share of the Chuka sorghum population assigned to cluster C was significantly smaller (9 %) as compared to the Tharaka (21 %) population (Fisher test: p-value = 0.023), and the Mbeere (25 %) population (Fisher test: p-value = 0.009). Half of the individuals (52 %) collected under the name *Kaguru* among the Chuka farmers were admixed, while this proportion was significantly lower among the Mbeere farmers (15 %, Pairwise Fisher test p-value: 0.0300) and among the Tharaka farmers (10 %, p-value = 0 .0190). Accordingly, the genetic diversity parameter estimates calculated for the *Kaguru* individuals collected in the Chuka farms were significantly higher than those for the Mbeere and Tharaka farms (Table II-3). Altogether, these observations suggest that more admixture occurred between the *Kaguru* population and local landraces within the Chuka cropping systems than within the Tharaka and Mbeere systems.

Table II-3. Summary of the genetic polymorphism indexes of the *Kaguru* variety sampled in the three ethnic groups. N_i: number of samples, N_{hh}: number of households, R_{Al}: unbiased allelic richness. H_e: unbiased gene diversity, H_o: observed heterozygosity, F_{IS}: fixation index. The letters (a, b, c) next to the R_{Al} and H_e values indicate the statistical significance of their differences between ethnic groups (Wilcoxon test) at a 5 % level after correction for multiple testing (FDR). For a given index, ethnic groups with the same letter did not present significant differences.

Ethnic group	N_{hh}	N_i	$R_{Al.}$	He	Ho	F_{IS}
Chuka	19	22	3.29 ^{a}	0.339 ^{a}	0.049	0.857
Mbeere	17	20	2.33 ^{b}	0.184 ^{b}	0.003	0.985
Tharaka	13	20	1.83 ^{b}	0.091 ^{c}	0.006	0.939

The MMb cluster D appears clearly separated but rather heterogeneous on the Neighbor-Joining tree. On a morphological basis, these varieties mostly fall in a clearly distinct group. Most individuals assigned to cluster D were identified as long-cycle landraces by farmers (*Muruge mbura ciiri*, *Mugana*, *Muthigo*, *Mucuri*, *Kathirigwa)* and a few as short-cycle improved varieties (*Serendo* and *Musalama)*. The latter individuals identified as improved varieties, both collected on-farm and stemming from certified seeds, formed a distinct genetic sub-group D' on the Neighbor-Joining tree and STRUCTURE confirmed these results for K = 5. The rest of the individuals assigned to cluster D were distributed across three major sub-clusters (Figure II-4.A). Most *Muruge mbura ciiri* individuals clustered together in a separate branch on the Neighbor-Joining tree (D1). *Mugana* and *Kathirigwa* formed another branch (D2), and *Mucuri* a third one (D3). Hence, there was a clear correspondence between the farmers' nomenclature and the genetic structure of individuals assigned to MMb cluster D, as well as with the structure of panicle morphological diversity (Figure II-S3). Cluster D was mainly observed in the Chuka area, as seen on Figure II-5.C and confirmed with pairwise Fisher tests (p-value < 0.05). Interestingly, the few Tharaka households where we collected individuals assigned to cluster D were located in the Chuka area. Moreover, one household located on the eastern side presented several individuals assigned to cluster D, but it was a Chuka household settled in the Tharaka area (Figure II-5).

The MMb cluster B is both central and diverse on the basis of molecular markers as well as morphological traits. The individuals assigned to cluster B were mainly identified as local landraces bearing various local names, whose occurrence frequency differed across ethnic groups (Table II-4). Most of those collected in the Chuka and Tharaka farms were named

Muruge mbura imwe, *Mugeta* and *Mbura imwe* while no or very few individuals collected in the Mbeere farms were named as such. Moreover, most of those collected in the Mbeere farms were named *Ngirigacha* (61 %), while fewer individuals bore that name in the Chuka (8 %) and Tharaka (18 %) populations. Cluster B accounted for a uniformly large share among the farmers of the Tharaka (39 %), the Chuka (27 %) and the Mbeere (26 %) ethnic groups. It showed little internal sub-structure with no clear correspondence to farmers' varieties, and a morphological differentiation between *Muruge mbura imwe* and *Mugeta* (Figure II-S2). As the only individuals with peculiar features, four individuals assigned to MMb cluster B for $K = 4$ formed a separate branch on the genetic Neighbor-Joining tree (B') and their difference was confirmed by STRUCTURE for $K = 5$. It could be explained by their foreign origin, as farmers reported purchasing these seeds at a lowland market. A fifth individual assigned to cluster B for $K=4$ formed a long branch (B") indicative of a marked genetic differentiation. It was identified as *Muthigo wa mwimbi* which means that it was introduced from another ethnic group (*Mwimbi*).

Table II-4. Proportion of individuals of each variety assigned to MMb cluster B regarding their collection ethnic group. Percentages in brackets.

Variety	Chuka	Mbeere	Tharaka
Muruge mbura imwe	14 (39 %)	1 (5 %)	10 (29 %)
Mugeta	10 (28 %)	-	8 (24 %)
Ngirigacha	3 (8 %)	11 (61 %)	6 (18 %)
Mbura imwe	7 (19 %)	-	8 (23 %)
Muthigo wa mwimbi	1 (3 %)	-	-
Others :			
Gadam	1 (3 %)	2 (12 %)	2 (6 %)
Kaguru	-	1 (5 %)	-
Muruge mbura ciiri	-	3 (17 %)	-
Total	36 (100 %)	18 (100 %)	34 (100 %)

DISCUSSION

Our study showed that in a uniform agro-ecological environment, social boundaries associated with ethnolinguistic diversity patterns have impacted the distribution of sorghum varieties and their genetic spatial patterns. If seeds, knowledge and practices were freely exchanged across the three ethnic groups, we would expect their sorghum varieties to be similar and, because of their geographical proximity and similar environmental conditions, to display no genetic differentiation. Quite the contrary, we showed that ethnic groups maintained different sorghum landraces, whereas improved varieties were uniformly distributed across groups. The structure of their diversity, as assessed with molecular markers, reflected the influence of improved variety dissemination and a differentiation in terms of cycle duration and phenology. The overall distribution patterns were clearly associated with the farmers' ethnic partition. Long-cycle landraces formed a genetically distinct cluster which was more frequently encountered in the Chuka sorghum population than in the Tharaka sorghum population. The improved *Kaguru* variety showed more admixture with the local landraces in the Chuka sorghum population than in the Mbeere and Tharaka ones. As a result of the unbalanced frequency of the different genetic clusters across ethnic groups, the genetic differentiation of their sorghum populations was significant. However, the structure based on molecular makers did not discriminate the three major short-cycle landraces whose frequency varied markedly across ethnic groups and which display different morphological characteristics.

The uneven distribution of named landraces across the Chuka, Tharaka and Mbeere ethnic groups is consistent with the results of Baco et al. [43], who reported that different ethnic groups in Benin cultivated different varieties of yam. A similar relationship between the structure of the genetic diversity of domesticated populations and farmers' social organization was found in taro populations across linguistic groups in Vanuatu [6], and in goat populations across ethnic groups in Vietnam [7]. However, a common caveat to such crop diversity studies conducted on large spatial scales is the difficulty involved in assessing whether the spatial patterns of crop diversity are related to variations in agro-ecological conditions, geographical distance, or to socio-cultural differences between human societies [10]. The community-scale approach we used in this study revealed that social boundaries have contributed to the differentiation of sorghum populations across spatially-close ethnic groups living in the same agro-ecological environment. Such an approach is thus complementary to country or regional-scale studies. In addition, such an approach makes it possible to investigate the mechanisms behind the relationship by jointly analyzing the distribution of varieties and the structure of genetic and morphological diversity in relation to the social organization of the communities concerned.

The ethnic identity of human groups is maintained by social boundaries that impede their cultural homogenization [44]. Our results suggest that these social boundaries also maintain differences between crop populations across ethnic groups. Indeed, gene flows in crop populations greatly depend on the exchange of seed, which is facilitated by social relationships and limited by social boundaries [4]. In addition, farmers' seed selection practices have a strong impact on crop populations and can differ considerably across communities [5,45]. The comparison of the structure of the genetic and morphological diversity of sorghum populations provides information concerning gene flows and selection forces, while the study of the nomenclature given to farmers' varieties tracks the diffusion of knowledge across farming communities. Thus, by combining the two approaches it is possible to investigate the respective influence of seed exchanges and the diffusion of selection practices across ethnic groups on sorghum genetic diversity patterns.

Limited diffusion of long-cycle landraces across ethnic groups

Long-cycle landraces formed a distinct MMb genetic cluster, whose frequency differed across ethnic groups. It was more frequent among the Chuka than among the Tharaka, and, interestingly, these results confirmed farmers' reports stating that long-cycle landraces were "*Muvia wa Chuka*", the sorghum of Chuka people. Moreover, certain sub-types within this cluster (sub-clusters D2 and D3) were not present in the Mbeere population and corresponding landraces were not inventoried in that ethnic group. The relation between the spatial distribution of the MMb genetic clusters and that of ethnic groups suggests that social boundaries limit the diffusion of planting material. Indeed, in most rural societies, seed exchanges depend on social networks as trust is required for seed transactions [46]. On the one hand, social relationships directly shape the seed exchanges because they facilitate access to seeds [47,48,49]. On the other hand, the social network is the major pathway for information exchange [50] and indirectly helps shape seed exchanges, as farmers tend to imitate relatives [51]. The joint action of these two mechanisms can thus explain the uneven distribution of long-cycle landraces across ethnic groups. In addition, the small grains and the bitter taste could explain the low economic value of these landraces, which probably helps limit their diffusion.

Management practices of improved varieties differ across ethnic groups

In contrast to the case of some landraces, improved varieties were uniformly distributed and their frequencies did not differ between ethnic groups. The recently introduced *Gadam* variety was genetically distinct from the landraces and showed limited introgression from the other genetic clusters. It was genetically uniform and complied with certified control. However, farmers also gave the names of local and already known variety to individuals that have the same genetic profile as *Gadam*, an improved variety. This can be explained by a morphological similarity. Yet it raises the question of the consequences this will have for the on-farm evolution of the improved variety. *Kaguru*, for instance, which was introduced in the area 10-15 years ago, seems to have evolved differently across ethnic groups. High admixture was detected between *Kaguru* and the local landraces in the Chuka population, resulting in a range of genetically diverse materials still called *Kaguru*, while this variety was found to be genetically more uniform in the Mbeere and Tharaka populations. As a result, the genetic diversity of *Kaguru,* as identified by farmers, was greater in the Chuka population than in the Mbeere and Tharaka populations. According to the farmers, the variety was introduced simultaneously in the three ethnic groups but little information is available concerning the origin of the seed lots. The divergence of the *Kaguru* variety across ethnic groups within a few decades could thus be the result of differences in their management practices, be it in planting (spatial arrangement of the varieties) or in seed selection. The higher admixture rate between *Kaguru* and the local landraces among the Chuka could be due to more intense gene flows within fields or to less stringent selection practices. As our observations suggest that the cropping systems used by the three ethnic groups were similar, the hypothesis of different selection practices is more likely. Cases of divergent selection practices between geographically close communities were observed by Pressoir and Berthaud [45], and by Perales et al. [5], who hypothesized that social boundaries impede the homogenization of selection practices. However, the hypothesis of the introduction of seed lots with different genetic characteristics in the three ethnic groups cannot be excluded.

Divergence in the nomenclature of the landraces between ethnic groups

Comparing the genetic structure of short-cycle landraces, their morphological characteristics and the farmers' nomenclature, raises interesting questions concerning the relation between farmers' nomenclature and the diffusion of planting material. Indeed, the frequencies of the majority of named short-cycle landraces differed significantly between ethnic groups even though they were assigned to the same genetic pool and no clear correlation was detected between named landraces and the MMb genetic sub-structure. The short-cycle landraces grown by the different ethnic groups thus appeared to belong to the same genetic pool. Yet, the analysis of the morphological characteristics of the panicles suggests that the landraces presented morphological differences that were not detected with neutral genetic markers. *Mugeta* and *Muruge mbura imwe*, mainly grown by the Chuka and Tharaka, corresponded to two distinct morphological groups while *Ngirigacha*, which is mainly grown by the Mbeere, was distributed over the entire PCoA plot. These results suggest that ethnic groups use different names for landraces with similar morphotypes: the Chuka and Tharaka identify and name two main short-cycle landraces corresponding to distinct morphotypes while the Mbeere mainly use the name *Ngirigacha* for all the morphotypes corresponding to the short-cycle landrace group.

This difference in folk-nomenclature and classification between the Mbeere and both the Chuka and Tharaka groups may result from limited knowledge diffusion. This is consistent with a number of observations concerning the conflictual relationship between the Chuka and Mbeere groups [19]. The impact of social relationships on the diffusion of folk-taxonomy and nomenclature among farming communities was demonstrated by Boster [52], who showed that the cassava nomenclature used by kin-related women was more similar than that used by non-kin in the Aguaruna community in Peru. Nuijten and Almekinders [12] also reported that the naming of rice varieties was more consistent within villages than between villages in Gambia. They pointed out that information concerning varieties, such as names, is not necessarily passed on with the seed lots. Hence, seed exchanges between communities can be more intense than knowledge diffusion, leading to the use of different names for similar morphotypes and genotypes. Further comparison of farmers' nomenclature and taxonomy between ethnic groups is required to confirm this hypothesis.

Previous studies showed that different sorghum varieties may display no genetic differentiation despite being morphologically distinct. Notably, in Cameroon, Barnaud et al. [3] showed that considerable gene flows existed between Guinea sorghum landraces while farmers kept on selecting them for their morphological distinctiveness. Rabbi et al. [53] reported similar results in western Kenya, while varieties collected in eastern Sudan were clearly genetically distinct. He explained these results by the varietal isolation practiced in Sudan, while Kenyan farmers mixed varieties within their fields. Soler et al. [54] found that landraces were distinct genetic units, but in that study each landrace was sampled in a single field belonging to one farmer, which considerably limits the variability. As farmers' taxonomy and nomenclature is based on morphological traits with a simple genetic determinism, morphological differences can be maintained even though gene flows occur within farmers' fields. The 18 SSRs used in our study were selected because they revealed high polymorphism in previous diversity studies, and they proved to be adequate for characterizing the genetic sub-structure of long-cycle landraces. However, their resolution power may not be sufficient to reveal a finer-scale genetic sub-structure in short-cycle landraces.

Effect of community social organization on the diffusion of seeds, knowledge and farmers' practices

According to local elders, the three ethnic groups migrated to the study area about a century ago. Our results suggest that even though they have lived in proximity since then, the way knowledge, practices and seeds are diffused has maintained differences between sorghum populations across ethnic groups. Ethnographic observations of community social organization provide explanations for such limited exchanges across geographically close communities. Indeed, information transmission and diffusion appear to be confined within the residential groups (parents and married sons) first, which is common in patrilineal and patrilocal societies [55], and next within the neighborhood group, which is a major social institution among eastern Kenyan Bantu communities [16,56], (Linsig pers.com). The way knowledge is transmitted and diffused is very conservative and favors cultural differentiation between communities [57,58]. It thus probably plays a major role in maintaining differences in nomenclature and practices between ethnic groups, and maybe also in limiting seed exchanges.

Conclusion

Our study highlights the importance of local short-scale studies to investigate farm crop evolution processes. To date, emphasis has been placed on the effect of agro-ecological conditions on crop evolution processes, as in the study of the evolution of wild plants. The influence of the cultural diversity and social organization of farming communities has consequently been neglected, although the major role of smallholders in the management of crop diversity has been acknowledged [59]. Crop evolution is still ongoing in smallholder farming systems and such systems occupy a substantial proportion of croplands in developing countries, especially in Africa [60]. Most of these rural communities have retained pre-colonial social institutions that continue to shape the relationships between people. Sixty-eight living language groups were inventoried in Kenya and about 2 146 linguistic groups in Africa [61], so the situation of ethnic co-existence described in this paper is not an isolated case. This study confirms the influence of the ethnolinguistic patterns of rural communities on gene flows and on farmers' selection practices that shape crop diversity *in situ*. Crop diversity patterns, thus result not only from an interaction between genetic and environmental factors, G × E, but from a three-way interaction G × E × S, where ''S'' stands for effects of the social boundaries [4]. Investigating this relation in other communities, with different social organizations and rules for the transmission of knowledge, would thus help gain a clearer picture of crop evolution dynamics in subsistence farming systems.

A further study is now needed to probe the mechanisms involved. Notably, the link between seed exchange networks and social organization deserves more investigation to confirm whether seed exchanges are confined within ethnic groups. This would explain why the diffusion of long-cycle landraces is more limited than that of short-cycle landraces. Moreover, further comparison of the local sorghum nomenclature and classification systems (folk taxonomy) across ethnic groups would make it possible to test whether their definition of landraces differs, and whether it influences their seed selection practices.

The uneven distribution of the genetic clusters across ethnic groups within a restricted geographic area highlights the need to take the social relationship and exchanges into account in the characterization, collection, and conservation of crop diversity. Accounting for the impact of human practices on crop populations would help capture their diversity more efficiently and, to this end, ethnic contact zones are of major interest for their potentially high genetic diversity. This study paves the way for participatory plant breeding as it shows that farmers' individual choices concerning planting material are not only determined by agro-ecological conditions or economic interest, but also by their cultural background.

ACKNOWLEDGEMENTS

We wish to thank Joseph I. Kamau and Catherine Wanjira for their contribution to field work, the staff of the National Genebank of Kenya and its director Zachary Muthamia, as well as Philippe Letourmy and Xavier Perrier for their help with statistical analysis. Thanks are also due to the participating farmers for their collaboration in our study. We feel indebted to the headmaster, teachers and students of Kabururu primary school for hosting our experimental trials on their school farm. We are also deeply grateful to the extension and local administrative staff in the study area. The support of the chiefs and assistant chiefs is particularly acknowledged.

REFERENCES

1. Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, et al. (2011) Multidisciplinary perspectives on banana (Musa spp.) domestication. Proc Nat Acad Sci USA 108: 11311-11318.

2. Roullier C, Benoit L, McKey DB, Lebot V (2013) Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. Proc Nat Acad Sci USA 110: 2205-2210.

3. Barnaud A, Deu M, Garine E, McKey D, Joly H (2007) Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. Theor Appl Genet 114: 237-248.

4. Leclerc C, Coppens d’Eeckenbrugge G (2012) Social organization of crop genetic diversity. The G × E × S interaction model. Diversity 4: 1-32.

5. Perales H, Benz B, Brush S (2005) Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. Proc Nat Acad Sci USA 102: 949-954.

6. Sardos J, Noyer JL, Malapa R, Bouchet S, Lebot V (2012) Genetic diversity of taro (Colocasia esculenta (L.) Schott) in Vanuatu (Oceania): an appraisal of the distribution of allelic diversity (DAD) with SSR markers. Genet Resour Crop Evol 59: 805-820.

7. Berthouly C, Do Ngoc D, Thévenon S, Bouchel D, Van T, et al. (2009) How does farmer connectivity influence livestock genetic structure? A case-study in a Vietnamese goat population. Mol Ecol 18: 3980-3991.

8. De Wet JMJ, Huckabay JP (1967) The origin of Sorghum bicolor. II. Distribution and domestication. Evolution 21: 787-802.

9. Harlan JR, De Wet JMJ, Stemler ABL (1976) Origins of African plant domestication. The Hague: De Gruyter Mouton. 498 p.

10. Deu M, Sagnard F, Chantereau J, Calatayud C, Hérault D, et al. (2008) Niger-wide assessment of *in situ* sorghum genetic diversity with microsatellite markers. Theor Appl Genet 116: 903-913.

11. Bellon MR, Brush SB (1994) Keepers of maize in Chiapas, Mexico. Econ Bot 48: 196-209.

12. Nuijten E, Almekinders CJ (2008) Mechanisms explaining variety naming by farmers and name consistency of rice varieties in the Gambia. Econ Bot 62: 148-160.

13. Jaetzold R, Schmidt H, Hornetz B, Shisanya C (2007) Farm management handbook. Vol II, Part C, East Kenya. Subpart C1, Eastern Province. Nairobi Ministry of Agriculture / GTZ.

14. Camberlin P, Boyard-Micheau J, Philippon N, Baron C, Leclerc C, et al. (2012) Climatic gradients along the windward slopes of Mount Kenya and their implication for crop risks. Part

1: climate variability. Int J Climatol. Available: http://onlinelibrary.wiley.com/doi/10.1002/joc.3427/full. Accessed 2 July 2013.

15. Ambler C (1988) The Great Famine, 1897-1901. In: Ambler C, editor. Kenyan Communities in the Age of Imperialism: The central region in the late Nineteenth Century. New Haven, London: Yale University Press. pp.123-149.

16. Middleton J (1953) The central tribes of the north-eastern Bantu; Forde D, editor. London: International African Institute. 105 p.

17. Moehlig WJG, Guarisma G, Platiel S (1980) La dialectométrie: une méthode de classification synchronique en Afrique. In: Guarisma G, Platiel S editors. Dialectologie et comparatisme en Afrique Noire. Paris: SELAF. pp. 27-45.

18. Chesaina C (1997) Chapter One: Historical and cultural background. In: Chesaina C editor. Oral literature of the Embu and Mbeere. Nairobi: East African Educational Publishers. pp. 3-10.

19. Glazier J (1970) Ritual and social conflict: circumcision and oath-taking in Mbeere. Nairobi: Institute of Development Studies, Nairobi University. 17 p.

20. Fadiman JA (1993) When we began there were witchmen: an oral history from Mount Kenya. Los Angeles: University of California Press. 395 p.

21. Plucknett DL, Evenson JP, Sanford WG (1970) Ratoon cropping. Adv Agron 22: 285-330.

22. Billot C, Rivallan R, Sall M, Fonceka D, Deu M, et al. (2012) A reference microsatellite kit to assess for genetic diversity of Sorghum bicolor (Poaceae). Am J Bot 99: e245-e250.

23. International Plant Genetic Resources Institute (IPGRI) (1993) Descriptors for Sorghum [Sorghum bicolor (L.) Moench]. Rome: IBPGR / ICRISAT. 38 p.

24. Gibson RW (2009) A review of perceptual distinctiveness in landraces including an analysis of how its roles have been overlooked in plant breeding for low-input farming systems. Econ Bot 63: 242-255.

25. Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26: 32-46.

26. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. (2012) Vegan: Community Ecology Package. 2.0-5 ed.

27. Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27: 325-349.

28. Nakazawa M (2013) fmsb: Functions for medical statistics book with some demographic data. 0.3.8 ed.

29. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1: 289-300.

30. Belkhir K, Borsa P, Chikhi L, Raufaste N, Catch F (2004) GENETIX, software under Windows™ for the genetics of populations. 4.05 ed. Montpellier: University of Montpellier.

31. Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12: 844-855.

32. Nei M (1987) Molecular Evolutionary Genetics. New York: Columbia University Press.

33. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-1370.

34. Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices 2.9.3 ed. Lausanne: Institute of Ecology.

35. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

36. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620.

37. Saitou N, Nei M (1987) The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.

38. Sokal R, Michener CD (1958) A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin 38: 1409-1438.

39. Perrier X, Jacquemoud-Collet JP (2006) DARwin software 5.0.156 ed. Montpellier: CIRAD.

40. Rice WR (1989) Analyzing tables of statistical tests. Evolution 43: 223-225.

41. Goudet J, Raymond M, de Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. Genetics 144: 1933-1940.

42. Raymond M, Rousset F (1995) GENEPOP version 1.2: population genetics software for exact tests and ecumenicism. J Hered 86: 248-249.

43. Baco MN, Biaou G, Pham JL, Lescure JP (2008) Geographical and social factors of cultivated yam diversity in northern Benin. Cah Agric 17: 172-177.

44. Barth F (1969) Introduction In: Barth F, editor. Ethnic groups and boundaries: The social organization of culture difference. Boston: Little, Brown. pp. 9-38

45. Pressoir G, Berthaud J (2004) Population structure and strong divergent selection shape phenotypic diversification in maize landraces. Heredity 92: 95-101.

46. Badstue LB, Bellon MR, Berthaud J, Ramírez A, Flores D, et al. (2007) The dynamics of farmers' maize seed supply practices in the central valleys of Oaxaca, Mexico. World Dev 35: 1579-1593.

47. McGuire S (2008) Securing access to seed: Social relations and sorghum seed exchange in eastern Ethiopia. Hum Ecol 36 217-229.

48. David S, Sperling L (1999) Improving technology delivery mechanisms: lessons from bean seed systems research in Eastern and Central Africa. Agric Human Values 16: 381-388.

49. Bellon MR (2004) Conceptualizing interventions to support on-farm genetic resource conservation. World Dev 32: 159-172.

50. Van den Broeck K, Dercon S (2011) Information flows and social externalities in a Tanzanian banana growing village. J Dev Stud 47: 231-252.

51. Bandiera O, Rasul I (2006) Social networks and technology adoption in northern Mozambique. Econ J 116: 869-902.

52. Boster JS (1986) Exchange of varieties and information between Aguaruna manioc cultivators. Am Anthropol 88: 428-436.

53. Rabbi IY, Geiger HH, Haussmann BI, Kiambi D, Folkertsma R, et al. (2010) Impact of farmers' practices and seed systems on the genetic structure of common sorghum varieties in Kenya and Sudan. Plant Genet Resour 8: 116-126.

54. Soler C, Saidou AA, Hamadou TVC, Pautasso M, Wencelius J, et al. (2013) Correspondence between genetic structure and farmers' taxonomy - a case study from dry-season sorghum landraces in northern Cameroon. Plant Genet Resour 11: 36-49.

55. Herbich I, Dietler. M (2008) The long arm of the mother-in-law: Learning, postmarital resocialization of women, and material culture style. In: Stark MT, Bowser BJ, Horne L, editors. Breaking down boundaries: Anthropological approaches to cultural transmission and material culture. Tucson: University of Arizona Press. pp. 223-244.

56. Labeyrie V, Rono B, Leclerc C (2013) How social organization shapes crop diversity: an ecological anthropology approach among Tharaka farmers of Mount Kenya. Agric Human Values. In press.

57. Reyes-García V, Broesch J, Calvet-Mir L, Fuentes-Peláez N, McDade TW, et al. (2009) Cultural transmission of ethnobotanical knowledge and skills: an empirical analysis from an Amerindian society. Evol Hum Behav 30: 274-285.

58. Cavalli-Sforza LL, Feldman MW (1981) Cultural transmission and evolution: a quantitative approach. Princeton: Princeton University Press. 388 p.

59. Brush SB (2000) Genes in the field: on-farm conservation of crop diversity; Brush SB, editor. Rome: Copublished by International Plant Genetic Resources Institute. 288 p.

60. International Fund for Agricultural Development (IFAD) (2001) Rural Poverty Report 2001: The Challenge of Ending Rural Poverty. Oxford: Oxford University Press for IFAD.

61. Lewis MP, Gary FS, Fennig CD (2013) Ethnologue: Languages of the World, Seventeenth edition. Dallas: SIL International. Available: http://www.ethnologue.com. Accessed 25 August 2013.

SUPPORTING INFORMATION

Table II-S1. Summary of the sampling of planting material. Mean number of varieties collected per household (Mean no. varieties / household) and mean number of samples of each variety collected per household (Mean no. samples / variety / household) in each ethnic group, followed by their standard error (SE).

	Chuka	Mbeere	Tharaka
Mean No. varieties / household	1.67 [SE: 0.14]	1.46 [SE: 0.13]	1.71 [SE: 0.20]
Mean No. samples / variety / household	1.51 [SE: 0.06]	1.32 [SE: 0.07]	1.56 [SE: 0.08]

Table II-S2. Summary of information and genetic diversity estimates per locus. Minimum and maximum size of alleles (Size), chromosome where the locus is located (Ch), percentage of missing data per locus (Miss), number of sampled alleles (N_{Al}), *He*: unbiased gene diversity, F_{IS}: Fixation index.

SSR Marker	Size (pb)		Ch	Miss	$N_{Al.}$	He	F_{IS}
	min	max					
mSbCIR223	120	140	2	0%	3	0.125	1
mSbCIR248	100	120	5	1%	4	0.051	0.932
sb4-72	196	235	6	1%	3	0.401	0.913
sb5-206	121	174	9	1%	13	0.774	0.951
sb6-84	192	240	2	0%	6	0.58	0.94
sbAgb02	111	179	7	1%	5	0.519	0.967
xcup02	199	229	9	0%	4	0.699	0.941
xcup14	224	256	3	0%	4	0.472	0.971
xcup53	200	230	1	0%	3	0.503	0.945
xcup61	214	223	3	0%	2	0.493	0.937
xtxp10	140	174	9	0%	8	0.73	0.962
xtxp12	180	230	4	4%	13	0.679	0.968
xtxp21	180	220	4	4%	13	0.76	0.976
xtxp289	270	346	5	2%	11	0.667	0.947
xtxp295	150	210	7	0%	13	0.753	0.949
xtxp320	271	304	1	1%	10	0.723	0.908
xtxp321	160	180	8	1%	11	0.685	0.975
xtxp57	240	276	6	2%	13	0.742	0.957

Table II-S3. Morphological descriptors used for panicle description.

Morphological descriptors	No. of modalities	Modalities
Panicle shape	8	Broom; Very loose; Loose; Semi-loose; Semi-compact long; Semi-compact ; Compact elliptic; Very compact
Grain shattering	2	Mid; High
Seed color	5	White; Cream; Grey; Brown; Red
Seed shape	2	Asymetric; Non-asymetric
Pericarp thickness	2	Thin; Thick
Subcoat	2	Present; Absent
Endosperm texture	2	Mainly vitreous, Mainly floury
Glume adherence	3	High; Mid; Low
Glume opening	4	Half-open; Highly open; Mid; Tight
Glume covering	2	Full; Mid
Awn	2	Present; Absent
Glume transversal wrinkle	2	Present; Absent
Glume texture	2	Hard; Papery
Glume color	3	Black; Red; Tan
Glume hairiness	2	Mid; High

Table II-S4. Results from the perMANOVA comparing the effect of ethnic groups on sorghum variety assemblages. Df: degrees of freedom, Ssq: sequential sum of squared distance between individuals and their group's centroïd, Mean Ssq = Ssq/Df, F.Model: pseudo F ratio, R^2: coefficient of determination [Ssq Etnic group / Ssq Total]

	Df	SSq	Mean SSq	F.Model	R^2	Pr(>F)
Ethnic group	2	2.047	1.023	4.971	0.076	0.0002*
Residuals	121	24.912	0.206	0.924		
Total	123	26.959	1			

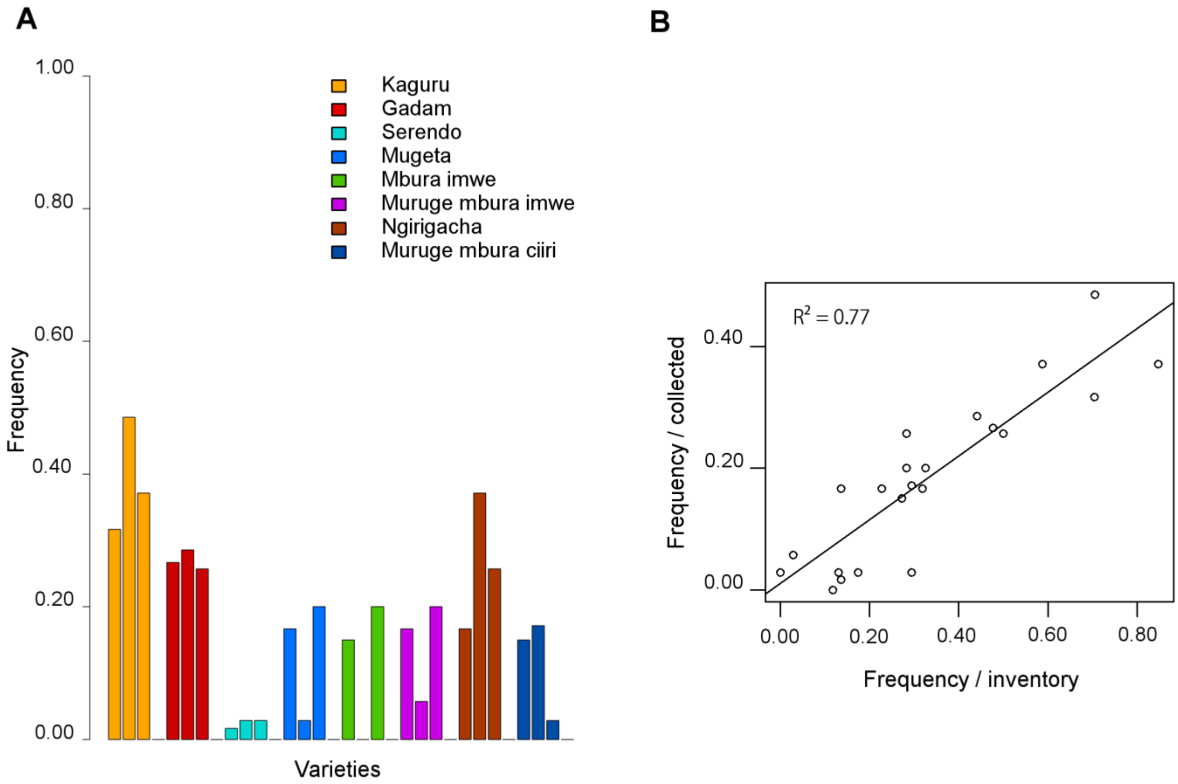


Figure II-S1. Comparison between the inventory of varieties and their sampling.

A. Number of households where each variety was sampled for the genetic diversity study (130 households)

B. Linear correlation between the proportions of households where each variety was inventoried (vertical axis, 124 households) and where it was collected (horizontal axis, 130 households) in each ethnic group.

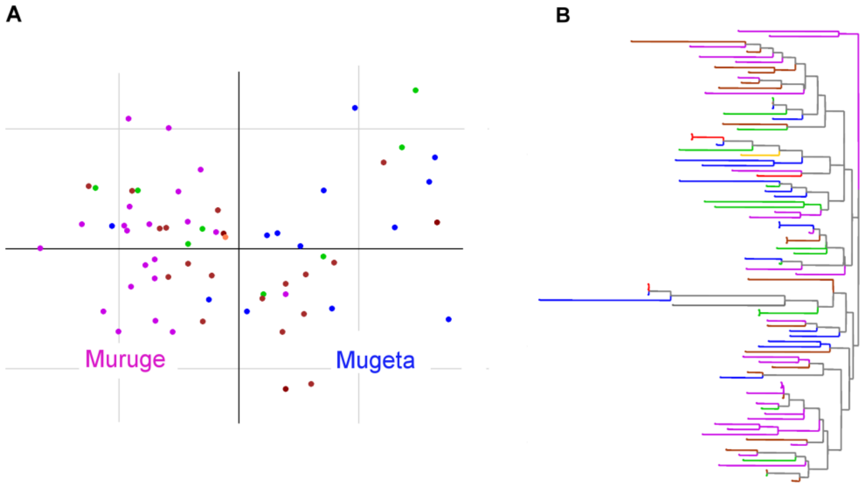


Figure II-S2. Structure of the morphological and genetic diversity within the MMb cluster B.

A. Plot of the two first axes of the Principal Coordinates Analysis (PCoA) based on 15 panicle morphological traits. PCoA was done on the sorghum plants assigned to the MMb cluster B. The first axis (x) expresses 35.1 % of the total variation and the second axis 13.1 %. Varieties are displayed using the following color code: Blue: *Mugeta*, purple: *Muruge mbura imwe*, green: *Mbura imwe*, brown: *Ngirigacha*, Red: *Gadam*, yellow: *Kaguru*, salmon: *Muthigo wa mwimbi*.

B. Neighbor-Joining tree based on 18 SSRs among the sorghum plants assigned to the MMb cluster B using the simple matching index.

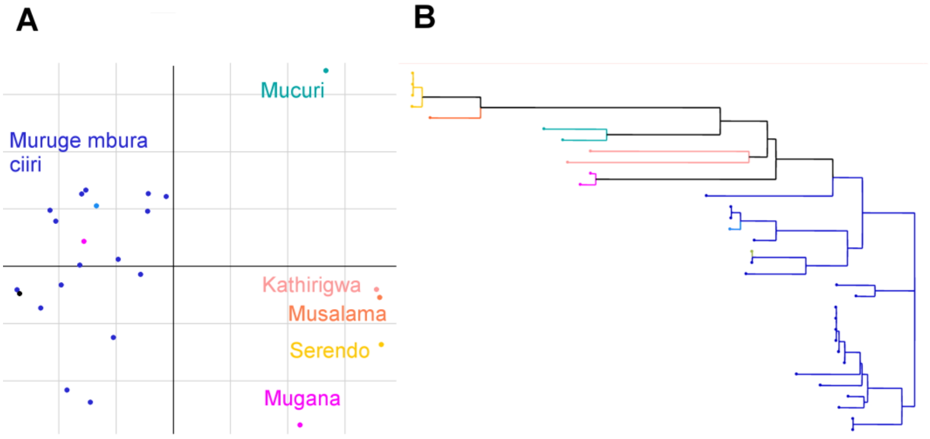


Figure II-S3. Structure of the morphological and genetic diversity within the MMb cluster D.

A. Plot of the two first axes of the Principal Coordinates Analysis (PCoA) based on 15 panicle morphological traits. PCoA was done on the sorghum plants assigned to the MMb cluster D. The first axis (x) expresses 51.5 % of the total variation and the second axis 18.0 %. Varieties are displayed using the following color code: Yellow: *Serendo*, orange: *Musalama*, light-pink: *Kathirigwa*, Fushia: *Mugana*, Greenish blue: *Mucuri*, dark-blue: *Muruge mbura ciiri*, black: *Muthigo*, blue: *Mugeta*.

B. Neighbor-Joining tree based on 18 SSRs among the sorghum plants assigned to the MMb cluster D using the simple matching index.

REPRESENTATIONS ET ECHANGES, REFLETS DE L'ORGANISATION DES SOCIETES

CHAPITRE III- COHERENCE DES TAXONOMIES LOCALES A L'ECHELLE DE LA SOCIETE ET DIFFERENCES INTERCULTURELLES

ARTICLE: WHAT IS A LANDRACE? FORMALIZING FARMERS' TAXONOMY OF SORGHUM THROUGH AN EMIC APPROACH

V. Labeyrie, J.I. Kamau, C. Dubois, C. Catalayud, R. Rivalan, A. Barnaud, C. Leclerc

En préparation pour soumission au journal *American Journal of Botany*

WHAT IS A LANDRACE? FORMALIZING FARMERS' TAXONOMY OF SORGHUM THROUGH AN EMIC APPROACH

V. Labeyrie, J.I. Kamau, C. Dubois, C. Catalayud, R. Rivalan, A. Barnaud, C. Leclerc

En préparation pour soumission au journal *American Journal of Botany*

ABSTRACT

Crop species present a large phenotypic variability at the intraspecific level as compared to wild plants. Farmers largely contributed in the generation of this diversity of shapes and colors through their selection practices. For management and communication purposes, farmers identify and name categories within this continuum of phenotypic variability. The so-called landraces, of farmers' varieties were used in number of studies for the characterization and sampling of crop diversity on-farm. However, the lack of formalization of farmers' taxonomy raised questions from biologists concerning their consistency in identifying, naming and classifying varieties.

This paper proposes an emic approach to describe how farmers perceive and categorize the variability of sorghum in eastern Kenya, instead of referring to an outsider scientific frame as it has been done up to now. Furthermore, it investigates whether three adjacent ethnolinguistic groups differ in the way they identify, name and classify sorghum diversity using an intercultural comparison approach.

A set of 287 panicles, representative of the sorghum varieties grown in the area of study was presented to 96 female farmers randomly selected in each ethnic group. They were independently asked to name each panicle. The morphological characteristics of these same panicles were further scored using 15 descriptors, and the neutral genetic variability of plants on which panicles were collected was described using 18 SSR genetic markers. The originality of our approach is that we considered the names used by 96 farmers to identify each panicle instead of considering that used by a single farmer, as it was done in previous studies. It hence provides a picture of the taxonomy used at the society level, among inter-individual variability. A comparison between farmers' taxonomy and the structure of sorghum genetic and morphological variability was first achieved using clustering methods without *a* *priori*. Then, a priori clustering methods were used to test specifically whether farmers' varieties presented distinct morphological characteristics. The morphological variability covered by each variety name, in each ethnic group was measured using diversity index.

A poor match was observed between farmers' taxonomy and the structure of sorghum genetic diversity. On the contrary, farmers were found to identify, name and classify different varieties accurately and consistently in the wide and continuous range of morphological variability covered by the sorghum population they grow. Farmers' taxonomy was based on a small number of salient morphological traits, and farmers' varieties displayed an overall distinction and uniformity regarding these criteria. Nevertheless, the accuracy and consistency of farmers in naming and classifying varieties regarding their morphological characteristics differed among ethnolinguistic groups, suggesting that exchanges of planting material and information may be limited among them despite their spatial adjacency.

INTRODUCTION

Varieties within crop species present a remarkable morphological variability, as it was notably stressed by Darwin in the *Origin of species* (Darwin 1859). This diversity is generated and modeled by various evolutionary mechanisms occurring within subsistence farming systems (Hodgkin et al. 2007): natural and farmers' selection (Boster 1985; Louette and Smale 2000), gene flows, both pollen-mediated at the field-scale (Barnaud et al. 2007) and seed-mediated a longer distance (McGuire 2008; Vom Brocke et al. 2003), mutation and drift. The varieties named by farmers are assumed to be distinct units organizing this amazing continuum of phenotypic and genetic diversity (Harlan et al. 1976; Bellon and Brush 1994; Badstue et al. 2007). However, the impressive number of varieties inventoried on-farm in a number of studies raises the question of farmers' consistency and accuracy in naming varieties regarding their phenotypic characteristics (Jarvis et al. 2008; Sadiki et al. 2007).

The lack of formalization of farmers' variety nomenclature and taxonomy has raised skepticism from geneticists concerning the correspondence between farmers' classification and the patterns of genetic and phenotypic variability (Quiros et al. 1990). This question has been central for ethnobotanical research at the species level, for which folk taxonomies were found to be stable across cultures and highly consistent with the Linnaean taxonomy (Berlin

1973). Fewer studies, however, addressed this issue at the infraspecific level for crop varieties, and they draw different conclusions.

The correspondence between the structure of the phenotypic diversity of crop populations and farmers' taxonomy was found to vary depending on the crop species and the farming community studied. Farmers' variety names corresponded to distinct phenotypes of sorghum in Ethiopia (Teshome et al. 1997) and of maize in Mexico (Louette et al. 1997). However, Badstue (2002) observed that folk taxonomy poorly represented the morphological variability of maize, still in Mexico. Salick et al. (1997) found that one cassava phenotype can bear different names across families in Peruvian Amazonia.

Quiros et al. (1990) found an overall correspondence between the structure of potato genetic diversity revealed by isozymes markers and the variety names used by farmers in the Andes. More recent studies compared farmers' classification of sorghum varieties with the structure of the genetic diversity revealed by SSR markers (Barnaud et al. 2007). They observed among the sorghum populations grown in a single Duupa village in northern Cameroon that some varieties correspond to a distinct genetic cluster, while others do not correspond at all. Rabbi et al. (2010) described two contrasted situation in Kenya and Sudan. Landraces collected in Western Kenya showed weak genetic differentiation while they were morphologically distinct. Contrarily, East-Sudanese varieties were clearly genetically distinct. Soler et al. (2013) found that sorghum landraces corresponded to distinct genetic clusters in two villages of northern Cameroon, each variety being sampled in a single farm.

Most of these studies, however, were based on the biological characteristics of plants as viewed by the researchers. The dominant use of such etic approaches, referring to an outsider's scientific frame (Pike 1954), for the study of farmers' variety taxonomy may explain the widespread skepticism concerning the consistency of such folk classification. Such approach may fail to reveal the logic of folk classification systems and conclude hastily to their lack of consistence. Indeed, the absence of consistence of farmers in identifying, naming and classifying varieties could hardly be justified when considering the function of taxonomy. Folk taxonomy and nomenclature constitute "*a frame for storing and conveying experience and information*", as stated by Bulmer (Bulmer 1974 in Friedberg 1991). How learning and information exchanges as well as the exchange of planting material could exist without a common reference frame shared by individuals? To communicate, farmers must show some consistence in naming varieties regarding their phenotypic characteristics (Boster 1986; Nuijten and Almekinders 2008).

Another shortfall of most studies on farmers' variety taxonomy lies in their lack of regard concerning the cultural diversity associated with the social organization of the communities studied. The ways people think of and classify the natural world differ across cultures. Atran and Medin (2008) notably showed that different cultural groups living in the same agro-ecological environment can present noteworthy differences in their knowledge and modeling of their environment. A remarkable example of cultural differences in classification is provided by cross-cultural comparison of colors categories, different cultural groups presenting different number of basic colors categories (Saunders 2007). These examples clearly illustrate that culture contributes largely to determine the way human classify the continuum of biological variability which surrounds them.

There is thus a need for further formalization of the taxonomy used by farmers to classify crop varieties through an emic approach, by describing how people perceive and categorize the world instead of referring to an outsider scientific frame (Pike 1954, Grenand 2002). In addition, such formalization also requires the understanding of the impact of cultural differences on the way farmers identify, name and classify crop varieties.

In this study, we adopted an emic approach to study the taxonomy of cultivated sorghum (*Sorghum bicolor* [L.] Moench) used by farmers in three ethnolinguistic groups in eastern Kenya. This paper deals with two major questions: Is farmers' variety taxonomy consistent with the patterns of genetic and morphological variability of sorghum? Does farmers' variety taxonomy differ among cultural groups? To address these questions, we used a probabilistic approach which enabled us first to depict the common taxonomy, shared by the majority of farmers in each ethnolinguistic group, and secondly to investigate cross-cultural differences beyond inter-individual variability.

MATERIAL AND METHODS

Study site

The study site was located in the Eastern Province of Kenya at the limit boundary between Tharaka-Nithi and Embu counties (0°24'27.88"S, 37°46'35.59"E). We focused on a 15 km-square contact zone between Chuka, Tharaka and Mbeere ethnolinguistic groups presenting uniform agro-ecological conditions (Jaetzold et al. 2007). Chuka, Tharaka and Mbeere groups present cultural and linguistic differences, and members of each group believe in their common ascendance, which founds their distinct ethnic identity (Middleton 1953; Moehlig et al. 1980). The relationship among these three groups was described by ethnographers and colonial administrators. Mbeere had conflictual relationship with the Meru groups and notably the Chuka who were their direct neighbor (Glazier 1970; Mwaniki 1973) and current land-pressure maintains this rivalry. On the contrary, Tharaka and Chuka were allied in the past and consider they are blood brothers, or *gishiaro* in meru language (Fadiman 1993). Intermarriage is frequent between the Tharaka and the Chuka while it is very uncommon between both groups and the Mbeere (personal observations). This ethnic organization and relationship system is traduced on the territorial organization of the three ethnic groups, Tharaka and Chuka being spatially mixed and settled in the northern part of the study site while the Mbeere are located separately from the two other groups in the southern part of the area. The maintenance of this spatial partition among ethnic groups results mainly from the patrilocality residence rules, implying that most married men settle near the compound of their father (Middleton 1953).

Collection of sorghum varieties on-farm

Sorghum panicles were sampled in January and July 2011 in the three ethnic groups following the strategy described previously (Labeyrie et al., submitted), which aimed at representing the diversity of sorghum varieties named by farmers. Seeds from the panicles collected on-farm were sown in October 2011 in an experimental field under controlled and uniform growing conditions. We harvested panicles of 287 descendants randomly sampled to maximize the range of morphological variability.

Varieties identification experiment

The 287 panicles harvested in the experimental field were presented to a panel of informants from the three ethnic groups. Thirty-two female informants were randomly chosen in each ethnic group. The native and husband's ethnic group of each informant were recorded. Following the procedure used by Boster (1986), each informant was independently asked to identify each of the 287 panicles that were successively presented to her. A field assistant recorded the variety name used by each informant to identify each panicle. Spelling standardization was later done with the field assistants to ensure the consistency in scoring names by eliminating the differences due to variation in pronunciation across informants. We established for each variety name whether it was motivated or not (Grenand 2002). When the variety name was motivated, it was translated and we recorded to which characteristic of the variety it referred.

Genotyping

The DNA extraction, amplification, migration and alleles' size scoring for 18 microsatellite SSR loci was done in a previous study for 268 individuals out of the 287 considered in the identification experiment. The full methodological details are provided in Labeyrie et al. (submitted).

Description of panicles' morphological traits

The 287 panicles used for the variety identification experiment were scored for 15 qualitative morphological descriptors at the KARI National Gene Bank of Kenya. The study was limited to the characteristics of the panicles because the selection of seeds by farmers is done at home, before threshing and thus without considering the characteristics of the whole plant. Only qualitative descriptors were scored because they are the main criteria on which farmers base their perceptual distinctiveness (Gibson 2009). The 15 descriptors were selected for their polymorphism in the sorghum population studied and their ease of scoring. They included the main criteria that farmers reported using for identifying sorghum varieties, according to information collected during semi-directive interviews.

The traits scored concerned the whole panicle shape (7 modalities), seeds characteristics (color: 5 modalities; lateral shape, shattering, endosperm texture, sub-coat presence and pericarp thickness) and glumes characteristics (color, opening, adherence, covering, awning, hairiness, texture and presence of a transversal wrinkle). Part of these descriptors was selected among those recommended by the IPGRI (International Plant Genetic Resources Institute 1993), and more precise descriptors of seed and spikelets were added among some of those used by Snowden (1936). Procedures of double characterization of panicles randomly sampled enabled to ensure the consistency of operators in scoring morphological traits, and double data entry was performed to limit typing errors.

Statistical analyses

Classification based on sorghum genetic and morphological characteristics without a priori

The pairwise genetic distances between the 18 SSR genotypes scored on the 268 individuals were calculated with the Simple Matching index (Sokal and Michener 1958) using DARwin 5.0.156 software (Perrier and Jacquemoud-Collet 2006). Neighbor-joining tree (Saitou and Nei 1987) was then built on this distance matrix for the unique genotypes scored for the 18 SSR using the algorithm implemented in the R package ape (Paradis et al. 2004; R Development Core Team 2011). To further investigate the structure of the genetic diversity, we ran STRUCTURE software on the unique 18 SSR genotypes using the admixture model with correlated alleles' frequencies (Falush et al. 2003; Pritchard et al. 2000). Ten runs were performed for each K value comprised between 1 and 20, using a burning period of 500 000 and 1 000 000 iterations. For further analysis, we selected the run maximizing the log-likelihood of data for each K value. To identify the optimal K value supported by our data, we used the criterion based on the rate of change in the log probability of data between successive K values proposed by Evanno (2005).

A dissimilarity matrix was computed between the 287 individuals for the 15 morphological descriptors described above. The Simple Matching index was used to calculate the dissimilarity between pairs of panicles, simply by calculating the number of unmatching modalities between two individuals divided by the total number of descriptors (matching + unmatching). The computation of the dissimilarity matrix was implemented in DARwin 5.0.156. A Principal Coordinates Analysis (PCoA) was implemented on this matrix of morphological dissimilarity to characterize the overall structure of panicles morphological diversity using the R package ade4 (Dray and Dufour 2007).

A priori classification of panicles morphological characteristics based on variety names

Decision trees were built to identify the morphological traits on which farmers based their classification of sorghum varieties in each ethnic group (Breiman et al. 1984). The explanatory variable for this analysis was the variety name used by each farmer (32 in each ethnic group) to identify each of the 287 panicles, and the dependent variables were the 15 morphological traits measured on these panicles. The recursive algorithm evaluates all the possible dichotomic splits within the data set and chooses the split criterion that maximizes the information gain (or entropy reduction) within categories; it then recursively split each category using the same procedure. The algorithm used was implemented in the R package *rpart* (Therneau et al. 2012) and employed the Gini index to measure the information gain. The tree was then pruned, and its optimum complexity was determined to minimize the cost-complexity parameter based on the cross-validation error. The optimum complexity level hence determined, however, was influenced by imbalanced variety names frequency in the identification experiment data. We thus further ran a decision tree within the terminal categories which showed considerable entropy. To measure the accuracy of the decision tree in traducing the classification system of farmers, the sensitivity and specificity index were calculated. For a class C*i*, the sensitivity represents the number of panicles identified under the name C*i* by both the informants and the tree model (true positive), based on the total number of panicles identified under the name C*i* by informants (true positive + false negative).The specificity represents the number of panicles that were not identified under the name C*i* by both the informants and the tree model (true negative), based on the total number of panicles that were not identified under the name C*i* by informants (true negative + false positive).

To explain the discrepancies between the classification produced by the decision tree and that of farmers, a correspondence analysis (Benzécri 1973) was performed to identify the combination of morphological characteristics associated to each variety names. A contingency matrix was first built to count the number of times where each of the 96 informants associated each variety names to each modality of the 15 morphological traits for the 287 panicles they identified during the experiment. The correspondence analysis was then done only for the major names identified as terminal categories on the decision tree. Moreover, five morphological traits displaying rare modalities (Glume hairiness, wrinkle, glumes texture, awning, and glumes covering) were not considered in the analysis. The correspondence analysis displayed hereafter was thus done on 10 morphological traits. Analyses were done using the R package *ade4* (Dray and Dufour 2007).

Diversity index for measuring morphological variability covered by variety names

We calculated an index of morphological variability for each variety name in each ethnic group, using the formula proposed by Jost (2006) to measure diversity:

$D = exp(H)$ where H is the Shannon index : H= $-sum\ p_i\ log(2)\ p_i$, where p_i is the proportional abundance of the morphological trait i. Jost's D was calculated on the varieties x morphological traits contingency matrix, where the value in each cell corresponded to the number of times where an informant associated each morphological trait to each variety name. This index measures the morphological variability covered by each variety name, in each ethnic group.

RESULTS

Lack of correspondence between biological variability patterns and farmers' taxonomy

Identification experiment: a high diversity of variety names

76 different variety names were cited by the 96 informants during the identification experiment, after basic spelling homogenization. 68 % were unique lexemes and 32% were composed names (basic root and a determinant). All these names were based on 54 basic roots, 20 being unmotivated (the words have no known meaning except naming variety) and 34 being motivated, and 12 determinants were used. The majority of the 76 names recorded

(66%) had a motivated root, and all the determinants were motivated (See table in supplementary material I).

The basic roots lexemes could be classified in 12 main categories of meaning. One fifth of the basic roots were verbs or adjectives referring to the appearance of the crop. They referred to the color of the panicle, to the shape (*Mukumbu* – "compact"), to the hairiness (*Kiamanguo* – "furry"), or to the glumes characteristics (*Mwitumiria* – "closed"). Adjective and verbs used for human attitudes were also found (*Gatengu* – "provoke", *Mubenania* – "liar", *Msalama* – "peaceful", *Ciankoma* – "mad", *Kitharara* –"stupid"). The rest of the roots referred to animals (*Maguna* – "monkeys", *Magwika* – "baboons", *Kambumbu* – "bats"), to the human anatomy (*Macuiri* -"hairs", *Macara* - "hands"), to the use of the variety (*Gachobi* – "for beer", *Ndume* – "traditional beer"), or its processing (*Muthigo* – "dehulling"), and two names described their taste (*Mugana* and *Karuru* - "bitter"). Four roots were names of commercial varieties (*Gadam*, *Serendo*, *Mtama one*) or research stations (*Kaguru*). Other names were found to refer to crops varieties (*Muvia* – "sorghum", *Kiminchi* – "green peas"), to food preparations (*Kiamiga te* – "bread"), to materials (*Naironi* – "nylon"), or to the origin of the plant (*Mwitika* – "feral"). Last, the very commonly used root *Ngirigacha* means simply "agriculture".

Determinants referred to the appearance of the crop, such as the color (*mutune* – "red", *mbilu* – "black", *muceru* – "white", *muruge* – "looks cooked" which refers to its grey color) or its height (*ka nthuke imwe* – "one age", referring to the uniform size of the plants), to the length of its growth-cycle (*mbura imwe* – "one season", *mbura ciiri* – "two seasons"). Interestingly, reference to the ethnic groups was also found: *kichuka* and *kimbeere*. Four varieties names showed the prefix *cia-*, which is widely used for women names.

Out of type panicles, considered in course of degenerating because of crossing events with wild sorghum were identified by different names depending on their stage of degeneration. A panicle presenting highly dehiscent seeds will be called *Kitharara* ("stupid"), if sown severally, it will become *Kiamaguna* ("for the monkeys") and considered unsuitable for human consumption.

The frequency of names varied considerably (Figure III-1). Seven names were clearly more frequently cited than the others (more than 1 000 citations, or 5% of the total number of citations): *Gadam, Muruge, Kaguru, Mugeta, Serendo, Mbura imwe, Ngirigacha.*

The names whose citation frequency was low showed two different situations:

Most names were cited by few informants. 54 names were considered rare as they were cited by ten or less informants (i.e. less than 10 % of the 96 informants).

Some names were not frequently cited, but they were cited by a large number of informants. Such situation was observed for *Mukumbu*, *Mbunge*, *Mugana*, *Kathirigwa* and *Mucuri* which were cited less than 100 times by the informants over the 287 panicles, but which were cited by more than one third of the 96 informants (Figure III-1).

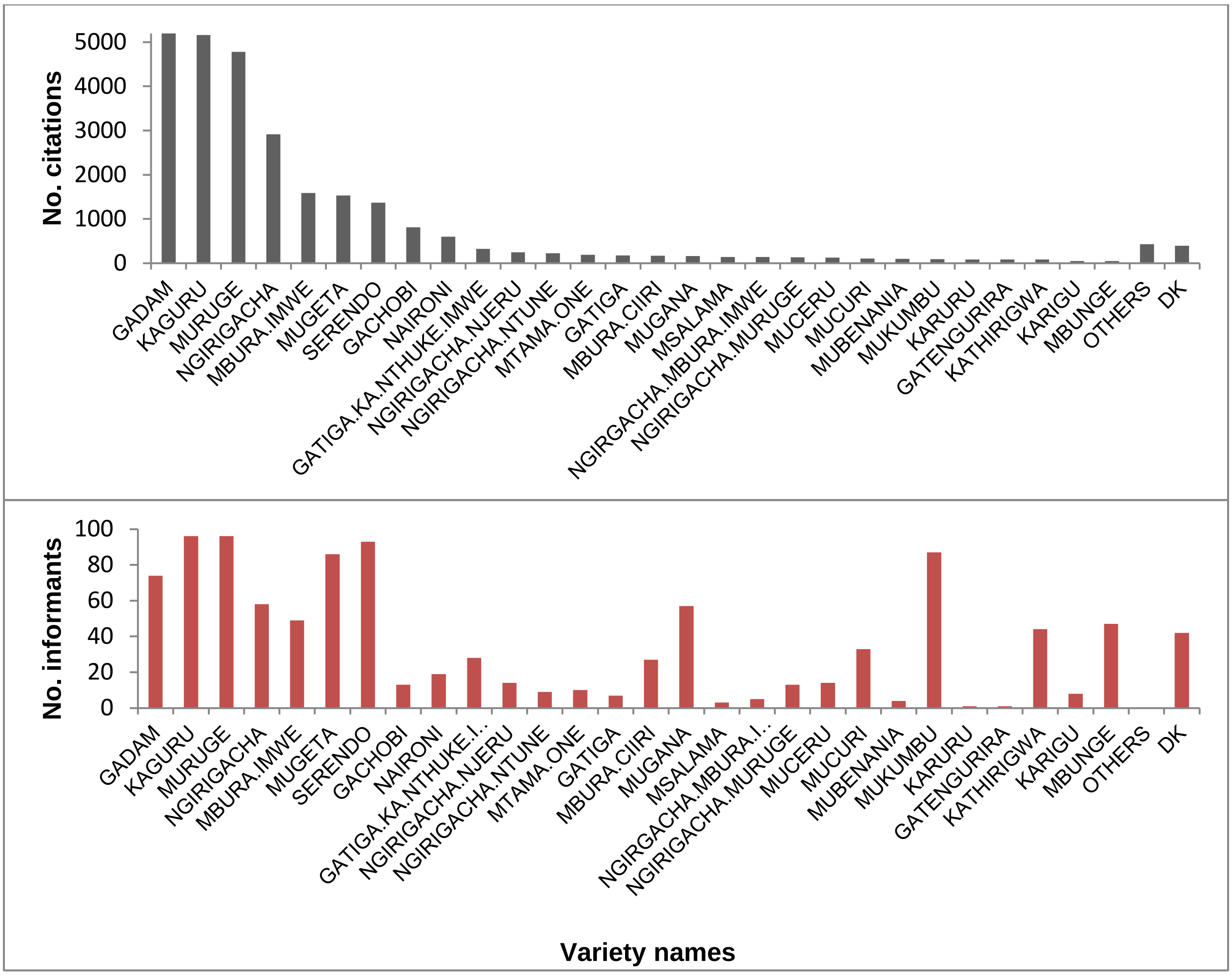


Figure III-1. Top: Overall citation frequency of the 76 variety names (96 informants, 287 panicles); Bottom: Number of informants who cited each variety names.

Lack of correspondence between sorghum genetic and morphological variability patterns and farmers' taxonomy

The Neighbor-Joining tree showed that the sorghum population studied was clearly structured (Figure III-2). Evanno's criteria suggested that an adequate number of populations for describing our data was $K = 9$, values for which the between-run variance of the log-likelihood of data was also minimal and congruence between runs was high. An overall good congruence was observed between the genetic clusters displayed on the Neighbor-Joining tree and those inferred by STRUCTURE for $K = 9$.

An overall discrepancy was observed between farmers' taxonomy and the structure of sorghum genetic diversity (Figure III-2). Four main genetic clusters showed some correspondence with the variety names used by farmers. Individuals assigned to the cluster D were mainly identified as *Kaguru*, those assigned to the cluster F as *Serendo*, those assigned to the cluster A as *Gadam* and those assigned to the cluster E as *Muruge*. However, the genetic clusters B, C, G and H did not correspond to distinctly named varieties. In addition the most frequent variety names, notably *Mugeta*, *Muruge*, *Mbura Imwe*, *Ngirigacha* and *Gadam* (depending on ethnic groups) were scattered over the Neighbor-Joining tree. *Kaguru* was the only variety name found in only one genetic cluster. Hence, most variety names did not correspond to distinct and uniform genetic units.

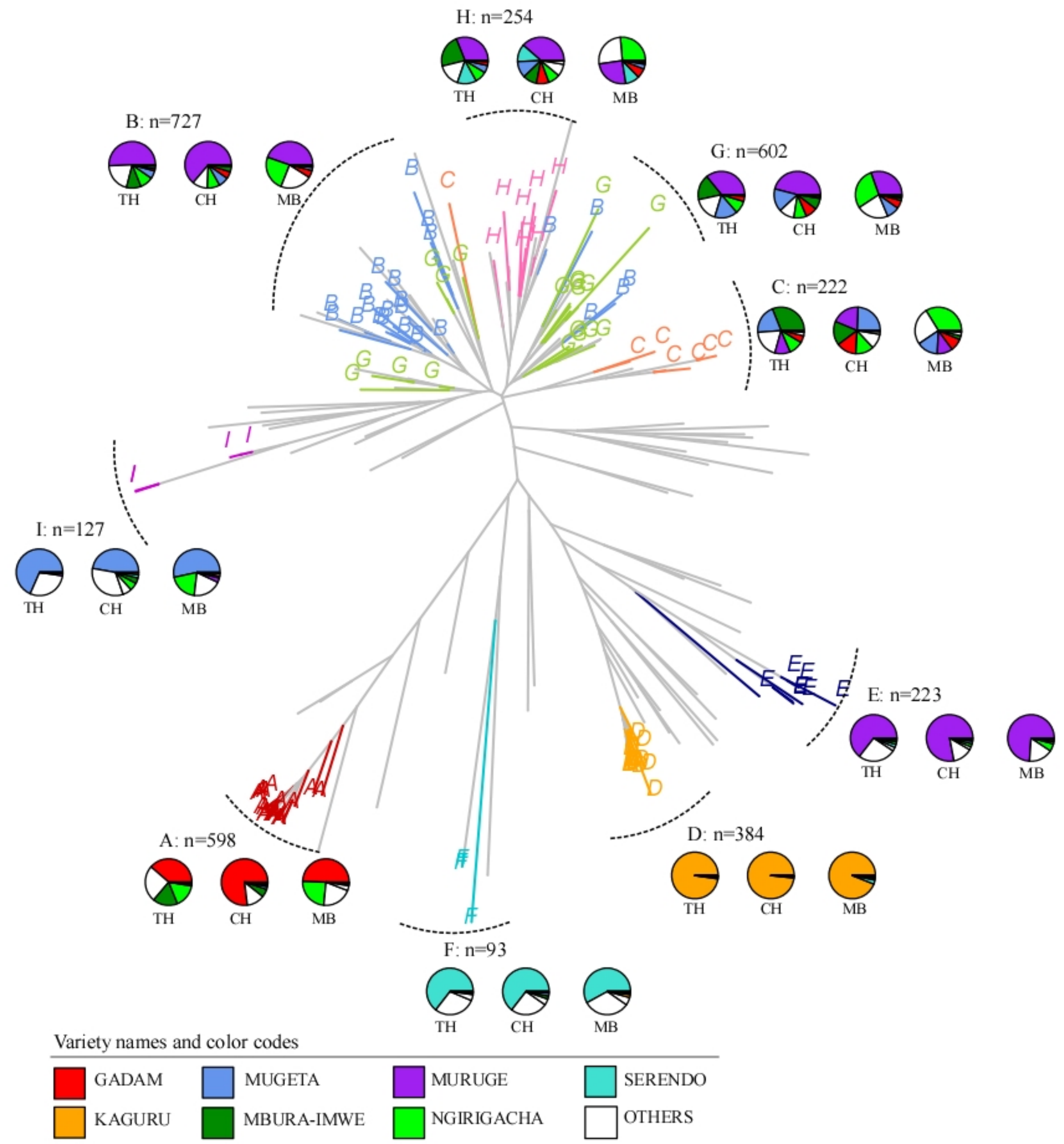


Figure III-2. Neighbor-Joining tree of the genetic distance between individuals (Simple Matching index) calculated using 18 SSR. Colors on the tree branches and letters identify the 9 genetic clusters inferred by STRUCTURE ($q > 0.8$) and admixed individuals are displayed in gray. Pie-charts display the fraction of variety names used by informants in each ethnic group (Chuka: CH, Tharaka: TH, Mbeere: MB) to identify the individuals assigned to each of the 9 genetic clusters inferred by STRUCTURE ($q > 0.8$). n: total number of variety names cited by the 96 informants to identify all the individuals belonging to each genetic cluster (n = No. of individuals in the cluster x No. of informants).

The range of morphological variability covered by the sorghum panicles displayed on the PCoA was wide and continuous (Figure III-3). No clearly distinct morphological clusters could be identified in such a continuum of variability. Obviously, no correspondence can be found between variety naming and the morphological variability.

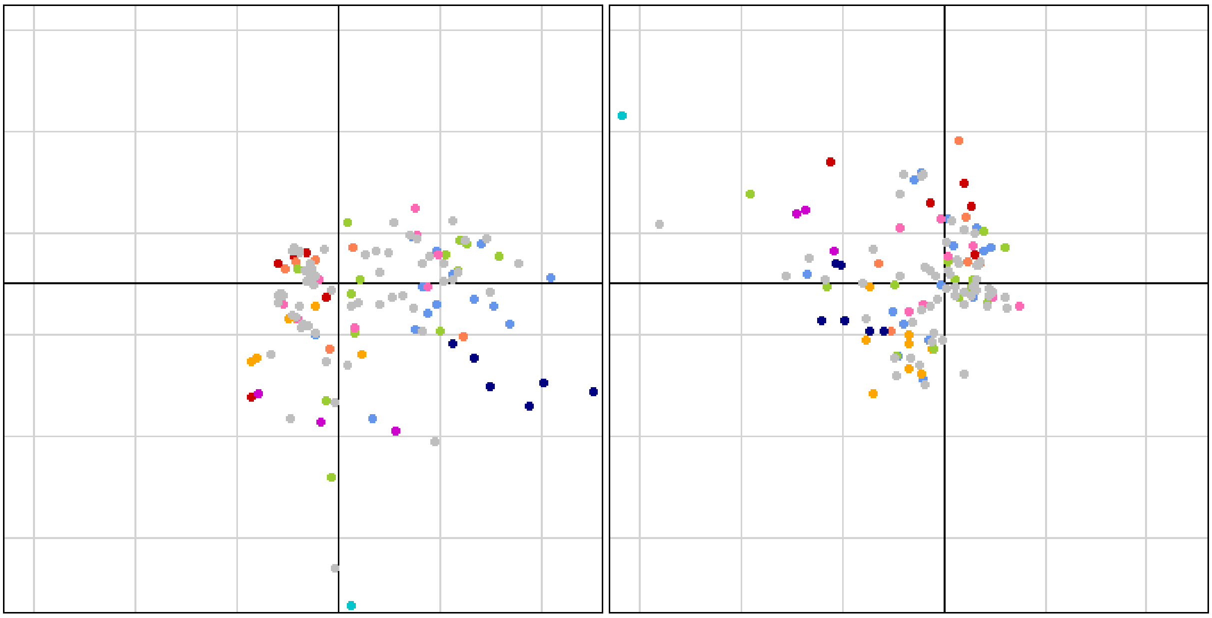


Figure III-3. PCoA on the 15 morphological traits (Sokal and Michener index). Colors correspond to the genetic clusters defined on figure III-2.left: Axis 1 (22.8 %) & 2 (14.8%), right: Axis 2 & 3 (10.6%).

Emic approach of farmers' variety taxonomy: farmers' varieties are distinct morphological units

Farmers' criteria for the classification of varieties

The decision tree schematizes the classification system of farmers in the three ethnic groups (Figure III-4). This model showed a good adequacy with the data, the mean sensitivity[15] being 0.81, and the mean specificity 0.94. This shows that farmers' panicle names mostly correspond to the named morphological categories as defined by the decision tree (high

---

[15] Mean over categories and ethnic groups

sensitivity), and that names used by farmers are usually associated to a single morphological category (high specificity). This regression tree thus gives an accurate picture of farmers' variety taxonomy, and it showed an overall good correspondence with the traits associated to each variety by the correspondence analysis (Figure III-5).

Mugeta, Gadam, *Mbura imwe* (Tharaka), *Ngirigacha* (Mbeere), *Mbunge* and *Mukumbu* presented white or cream seed color. Further distinction between these varieties was based on the glumes color (*Mugeta*), the panicles' shape, or glume covering (*Mbunge*) (Figure III-4). The classification of panicles presenting white or cream seed color differed among ethnic groups. All groups seemed to distinguish the compact-elliptic panicles, mainly named *Gadam,* from the looser panicles which displayed more diversity in the names that informants attached to them. Panicles assigned to this last category were mainly named *Mbura-imwe* by the Tharaka or *Ngirigacha* by the Mbeere, while the Chuka used several names (Figure III-4).

Mukumbu was classified with a high specificity in the three ethnic groups (Table III-1). *Mbunge* was observed in the Mbeere and Tharaka group, where it showed a high sensibility and specificity. It was cited by a restricted fraction of informants in the Chuka group. The correspondence analysis clearly discriminated *Mbunge* and *Mukumbu* which presented rare morphological characteristics: *Mbunge* had fully covering glumes and *Mukumbu* a bundled panicle shape and both displayed awns. These individuals were then removed for further correspondence analysis.

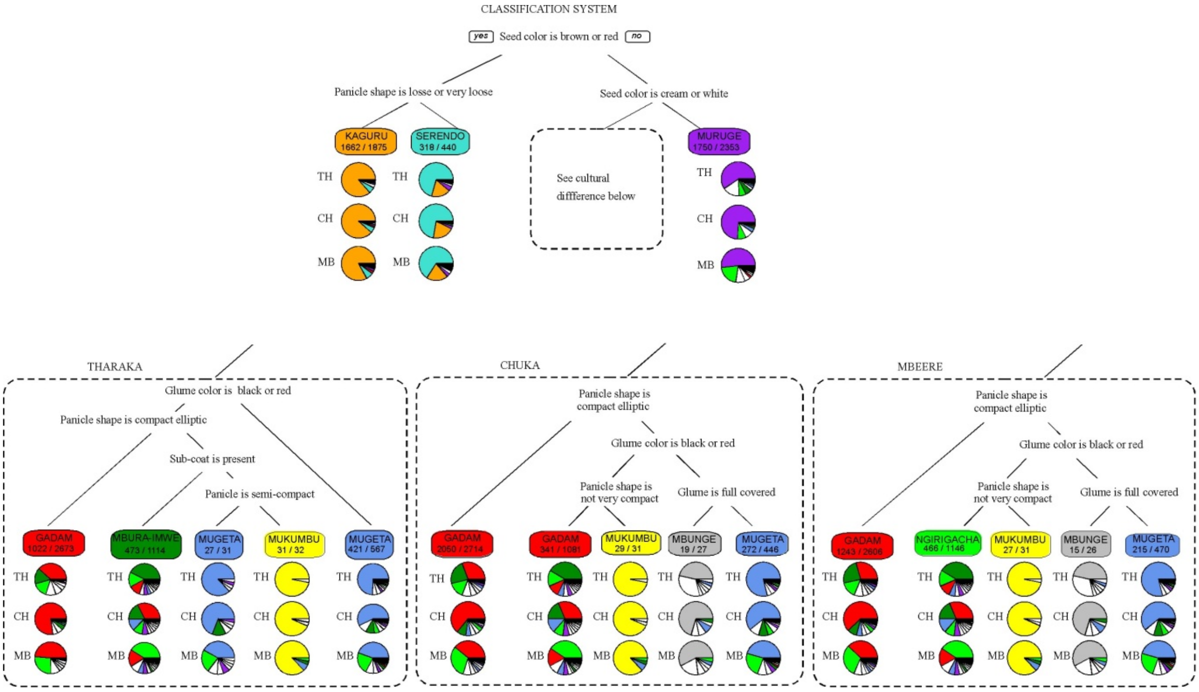


Figure III-4. Decision tree depicting farmers' taxonomy in the three ethnic groups. It creates morphological categories, defined by a classification key, in order to minimize the diversity of variety names within each category. Pie charts display the fraction represented by each variety name in each morphological category, and for each ethnic group (TH: Tharaka, CH: Chuka, MB: Mbeere).

Table III-1. Sensitivity (Ss) and specificity (Sp) of the classification inferred by the decision tree regarding farmers' classification. For a class Ci, the sensitivity represents the portion of panicles identified by informants as Ci which were classified by the tree in the class Ci. The specificity represents the portion of panicles not named Ci by informants and which were not classed as Ci by the tree.

	Tharaka		Chuka		Mbeere	
Variety names	Ss	Sp	Ss	Sp	Ss	Sp
KAGURU	0.94	0.97	0.94	0.97	0.93	0.96
SERENDO	0.67	0.99	0.68	0.99	0.67	0.98
MURUGE	0.91	0.87	0.93	0.92	0.89	0.85
GADAM	0.86	0.79	0.98	0.78	0.80	0.82
MUGETA	0.80	0.98	0.45	0.98	0.59	0.97
MBURA-IMWE	0.45	0.92	-	-	-	-
NGIRIGACHA	-	-	-	-	0.26	0.91
MUKUMBU	1.00	1.00	1.00	1.00	1.00	1.00
MBUNGE	-	-	1.00	1.00	1.00	1.00

Misclassified panicles tells us more about farmers' taxonomy keys

Some confusion was observed between the *Kaguru* and *Serendo* varieties, as a substantial share of panicles classified as *Serendo* by the tree were named *Kaguru* by informants, and reversely (Figure III-4). Indeed, the correspondence analysis showed that the black glumes and the red seed colors were also associated to the *Kaguru* variety on the first axis (40 % inertia), while these criteria were not used by the classification tree. A closer look at the morphological characteristics of the panicles classified as *Kaguru* by the decision tree while named differently by farmers confirms that the glumes color was also involved in the classification process. Two panicles classified as *Kaguru* were named differently by the majority of informants (> 80 %) because they presented tan glumes color. Some panicles were identified as *Kaguru* by part of the informants and as *Serendo* by the others because they combined characteristics from both varieties: red glumes and brown seed colors associated to *Serendo,* but loose panicles associated to *Kaguru*. Reversely, panicles showing semi-compact shape, but red seed or black glumes colors caused the same identification dilemma to informants.

The sensitivity of *Mugeta* category was 0.61 in average, suggesting that the decision tree did not perfectly represent the classification criteria of farmers for this variety. Indeed, *Mugeta* was also associated to the absence of sub-coat, vitreous endosperm and cream seed color on the third axis of the correspondence analysis (24.6 % inertia). The absence of sub-coat was detected as an additional criterion by the decision tree in the Tharaka group. The specificity and sensitivity of *Mugeta* varied strongly among ethnic groups.

The specificity of *Gadam* was the lowest (0.80), meaning that 20 % of the panicles which were not identified as *Gadam* by the informants were classified as such by the tree. Indeed, substantial confusion between *Gadam* and *Mbura-imwe* or *Ngirigacha* was observed as they only differ in the compactness of their panicle. The sensitivity of *Gadam* (0.88 in average) displayed large differences among ethnic groups.

The three ethnic groups displayed a morphological category defined by its white/cream seed color, red/black glumes color and non compact-elliptic shape which presented a large diversity of names. The name *Mbura imwe* was preeminent in this morphological category for the Tharaka, the name *Ngirigacha* for the Mbeere, but this category bore multiple names in the Chuka group. *Ngirigacha* and *Mbura imwe* were not clearly differentiated from the other varieties or associated to particular traits on the five first axis of the correspondence analysis (Figure III-5). The Jost index was high for these two variety names (Table III-2) and they displayed the highest diversity for panicle shape. Indeed, a large part of the Mbeere informants used *Ngirigacha* name to identify panicles corresponding to *Gadam*, *Muruge* or *Mugeta* morphological categories and the Tharaka used *Mbura-imwe* for panicles classified in *Gadam* morphological category (Figure III-4). This was indicated by the Jost index, which was higher for *Ngirigacha* in the Mbeere group (22.89) than for *Mbura-imwe* in the Tharaka group (20.86).

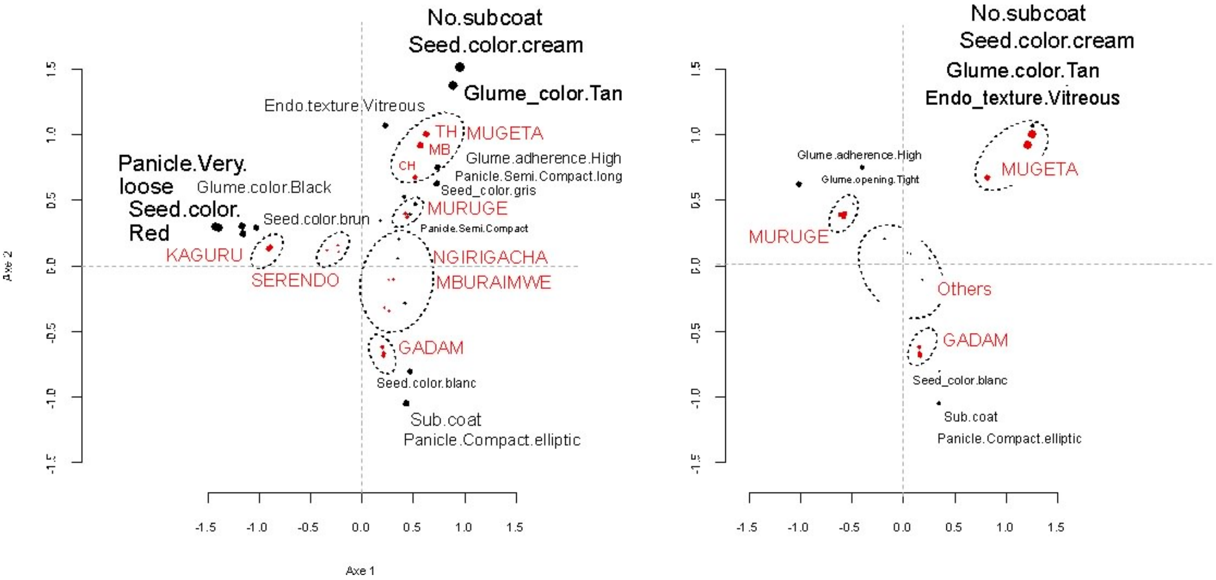


Figure III-5. Correspondence analysis of the seven most frequent varieties regarding 10 morphological traits. Left: axis 1 (40 %) and 2 (28.5 %), right: axis 2 and 3 (24.6 %). Letter emphasis is proportional to its distance from the barycenter

Morphological variability covered by variety names

The Jost index, which reflects the morphological variability covered by a given variety name, varied among variety names (Table III-2). It was the lowest for the improved varieties *Gadam* (17.05), *Kaguru* (19.36) and *Serendo* (21.47). The names corresponding to local landraces, notably *Mugeta* and *Muruge*, covered a wider range of morphological variability. *Mugeta* had the highest morphological variability (24.72), followed by *Mbura imwe* (22.6) and *Ngirigacha* (22.54). The morphological variability varied also strongly depending on the morphological characteristics considered: the higher variability index for *Mugeta* was due to the high variability of most traits, excepted for the color of the glumes. *Ngirigacha* and *Mbura imwe* showed the highest variability in the shape of the panicles.

Table III-2. Jost's D index of morphological variability (15 qualitative traits) calculated for each variety. Overall Mean: mean of D calculated on the fifteen traits across the three ethnic groups, mean of D across ethnic groups for the five major morphological criteria. Standard deviation across the three ethnic groups is indicated between brackets.

	Overall	Glume	Panicle	Seed color	Endosperm	Sub-coat
Gadam	17.05 (0.5)	1.22 (0.07)	1.82 (0.19)	1.19 (0.15)	1.02 (0.02)	1.02 (0.02)
Kaguru	19.36	1.46 (0.02)	2.63 (0.08)	1.66 (0.02)	1.26 (0.01)	1.01 (0.01)
Serendo	21.47	2.06 (0.04)	3.25 (0.03)	2.28 (0.19)	1.29 (0.03)	1.03 (0.02)
Muruge	21.95	1.69 (0.08)	2.94 (0.12)	1.5 (0.1)	1.1 (0.02)	1.04 (0.02)
Ngirigacha	22.54	1.76 (0.29)	3.87 (0.36)	2.23 (0.39)	1.19 (0.11)	1.18 (0.12)
Mbura	22.60	1.88 (0.27)	4.13 (0.12)	2.05 (0.47)	1.28 (0.19)	1.3 (0.25)
Mugeta	24.72	2.04 (0.30)	2.94 (0.46)	2.53 (0.22)	1.97 (0.06)	1.96 (0.03)

Cultural differences in naming and classifying sorghum varieties

The overall structure of farmers’ variety taxonomy did not show strong differences among ethnic groups. However, the accuracy and consistency of ethnic groups in classifying varieties regarding the morphological characteristics of panicles was found to differ clearly. Such differences among ethnic groups were observed by comparing on the one hand the diversity of variety names within morphological categories, and on the other hand the morphological diversity covered by each variety name.

Comparing the diversity of names within morphological categories

The diversity of variety names observed in each morphological category varied clearly across ethnic groups. Referring at the pie charts, the diversity of variety names in the morphological category labeled *Muruge* was notably lower in the Chuka group than in the other ethnic groups (Figure III-4). Similarly, the diversity of variety names in the morphological category labeled *Mugeta* was lower in the Tharaka group than in the other ethnic groups. Striking differences were observed for the morphological category labeled *Gadam* in the three ethnic groups. It showed an especially high diversity of variety names in the Tharaka and Mbeere groups while the name *Gadam* was dominant for this category in the Chuka group. On the other hand, the morphological categories labeled *Mbura-imwe* or *Ngirigacha* showed a higher diversity of variety names in the Chuka group than in the other groups.

Comparing the morphological diversity covered by variety names

The morphological variability covered by each variety name differed across ethnic groups. The value of Jost index notably varied among ethnic groups for *Mugeta*, *Mbura-imwe* and *Ngirigacha* varieties that showed clearly less morphological variability within the Tharaka group than in the other ethnic groups (Table III-3). Indeed, the *Mugeta* name was confined in one major morphological category in the Tharaka decision tree, while it was found in various morphological categories in the Chuka and Mbeere decision trees. Similarly, *Mbura-imwe* was found mainly in two morphological categories in the Tharaka decision tree while *Ngirigacha* was scattered across the different morphological categories. The name *Gadam* was clearly more frequently used by the Chuka than by the other groups to identify panicles with semi-compact shape. Last, *Muruge* name covered a higher morphological diversity in the Mbeere group (D = 22.31) than in the Chuka (D = 21.65) and Tharaka group (D = 21.90). Indeed, it was scattered across several morphological classes in the decision tree while it was bounded to one morphological category in the Chuka and Tharaka groups.

Table III-3. Jost index D of morphological variability (15 qualitative traits) calculated for each variety in each ethnic group. D: Jost index, N: Number of citation for that variety name on which the index was calculated.

	Tharaka		Chuka		Mbeere	
	D	N	D	N	D	N
Gadam	16.67	15767	16.87	32968	17.61	20739
Kaguru	19.2	24917	19.35	25592	19.55	24110
Serendo	21.97	5466	21.53	5379	20.9	5096
Muruge	21.9	20253	21.65	24800	22.31	17323
Ngirigacha	20.66	10513	24.07	5270	22.89	24074
Mbura	20.86	14650	22.73	6104	24.22	1182
Mugeta	23.95	7824	25.42	8400	24.78	5061

The differences observed between the Tharaka decision tree, and the Chuka and the Mbeere ones further indicate differences of informants' accuracy for sorghum classification across ethnic groups. For instance, the decision tree discriminated a morphological category labeled *Mbunge* in the Chuka and Mbeere groups but it was not discriminated in the Tharaka group, where a minority of informants used that name to identify this morphotype (Figure III-4). Furthermore, an additional morphological category labeled *Mugeta* was discriminated by the decision tree in the Tharaka group but not in the other groups where a lower share of informants used that name to identify this morphotype.

DISCUSSION

The etic approach fails to unravel the logic underlying farmers' taxonomy

An overall discrepancy between farmers' taxonomy and the genetic structure of sorghum was found as most variety names did not corresponded to distinct and uniform genetic units. Similar results were observed to some extent by Barnaud et al. (2007), and by Rabbi et al. (2010) while Soler et al. (2013) found an overall good correspondence between farmers' taxonomy and the genetic structure of dry season sorghum.

The three genetic clusters showing a higher uniformity for variety names corresponded to improved varieties: *Kaguru*, *Serendo* and *Gadam,* thus explaining this relatively good correspondence between names and genetic patterns. In addition, individuals in the cluster E were mainly named *Muruge.* Indeed, these individuals were previously identified as long-cycle local varieties (Labeyrie et al., submitted). The genetic distinction of this set of individuals is probably maintained because of their later flowering date, which must limit considerably the pollen flows with the rest of the varieties (data not shown). Individuals in the clusters B, C, G and H were identified as short-cycle local varieties and were found to be part of the same global genetic pool in a previous study (Labeyrie et al., submitted). The simple genetic determinism of the morphological traits used by farmers to identify the varieties probably explains the discrepancy between the genetic sub-structure in this pool and farmers' taxonomy.

The neutral genetic sub-structure within short-cycle landraces shows that different sorghum populations coexist in the area. The spatial distribution of these populations is random, and several populations are often present within a single household (data not shown). A possible explanation for this pattern would be the introduction of seed lots from different populations, potentially through the market. Indeed, this semi-arid area is frequently hit by severe drought and farmers regularly purchase seed at the local markets or borrow small quantities from other farmers. The practice of "completing seed lots" was frequently reported by farmers, a single variety within one household can thus originate from various seed origins. The genetic differentiation between these populations shall be maintained by limited outcrossing (data not shown) and by the dilution of pollen flows within fields, sorghum being planted at a low density and intercropped with various other species.

Sorghum is known to present a challenge for taxonomists because of its large and continuous diversity range (Teshome et al. 1997). Indeed, several classifications have been proposed, from that of Snowden (1936) which distinguished 52 species (31 being cultivated) to that of Harlan and de Wet (1972) who defined five races and fifteen intermediate races. The lack of structure in the morphological variability observed for the sorghum population studied illustrates this classification challenge. Our results thus raise questions concerning the ability of farmers to identify, name and classify different varieties in this continuum of variability. This suggests that the morphological criteria have different weights in the identification key used by farmers. Clustering methods without a priori thus give a poor picture of the logic underlying farmers' variety taxonomy. Similar results were obtained by Teshome in Ethiopia (1997), who showed that a priori clustering methods were adapted to depict farmers' classification system of sorghum varieties in continuum of morphological variation.

Farmers' varieties are distinct morphological units

This study showed that farmers' variety taxonomy is based on a small number of salient morphological traits, as suggested by Gibson (2009).The *a priori* classification approach was clearly more appropriate than without *a priori* analysis to describe the taxonomy of farmers in which the different morphological traits do not have the same weight. It is noteworthy that farmers are able to classify and identify consistently varieties in such a large continuum of morphological variability. These results confirm that farmers' varieties are "recognizable morphologically”, as it was already noticed by Harlan et al. (1976), and observed in previous studies (Teshome et al. 1997; Louette et al. 1997). Most of these studies, however, dealt with sorghum populations displaying a more clear structure of their morphological variability notably because they collected panicles in a large range of agro-ecological conditions.

Some variety names, however, covered several morphological categories. Our observations notably suggest that *Ngirigacha*, in the Mbeere group, and *Mbura-imwe*, in the Tharaka group may be inclusive categories (i.e higher order taxa) and not basic categories. Indeed, these names are used by part of the informants to identify various morphological categories that the major part of informants names *Gadam*, *Muruge* or *Mugeta* in the case of *Ngirigacha*, or *Gadam* in the case of *Mbura-imwe*. This is supported by the wide meaning of these names, *Ngirigacha* meaning “Agriculture” and *Mbura-imwe* meaning “one season”. Farmers may use these names in different situations. They can be used to label panicles presenting ambiguous characteristics regarding farmers’ taxonomy, either because of admixture between varieties or because they correspond to an exogenous variety. These inclusive category names may also be used by inexperienced farmers with limited knowledge in this domain, or by farmers having little experience of this particular variety. Last, farmers sometimes use inclusive categories when they judge that their interlocutor is not competent in the domain considered. *Mbura-imwe* seems to include only panicles showing white/cream seeds and red/black glumes, while *Ngirigacha* appears to include any variety except those with red/brown seeds.

Despite the clear classification scheme of farmers, some panicles were ambiguous regarding folk taxonomy because the combined attributes of two varieties or presented an unusual combination of characters. In the first case, panicles showed a partition between two variety names across informants, while in the second case, panicles were labeled using various names or inclusive categories names. Indeed, the strategy we used for selecting this set of panicles aimed at maximizing the continuum of morphological variability, “atypical” panicles were thus included

We considered a large number of morphological descriptors and they proved to be accurate for depicting farmers’ taxonomy. However, we missed some important criteria for farmers’ classification. Notably, a large share of informants used the name *Kathirigwa* to label a single panicle while it was not discriminated by the decision tree because it showed no particular morphological characteristics for the 15 traits scored. Indeed, farmers reported that the main criterion enabling them to identify *Kathirigwa* was the small size of its seeds, which was not measured in this study.

The formalization of farmers' varieties taxonomy we provide here argues against the belief that taxonomy would be primarily defined by the use and function of the varieties (Hunn 1982). The functional criteria basing variety names' motivation are distinct from the identification criteria. Friedberg notably stressed that the lower level of folk classifications are determined by perceptual characteristics, mainly visual (Friedberg 1991).The fact that the name of an object is motivated by its function or use does not mean that it is as an identification criterion. For instance, the fact that a plant presents a short growth-cycle ("*Mbura-imwe*") is not a perceptual trait; it is through perceiving a set of morphological traits that a farmer is able to infer this characteristic of the plant.

Intercultural differences of farmers' accuracy and consistency in varieties' classification

Considerable intercultural differences thus exist in the accuracy and consistence of ethnic groups in classifying varieties regarding the morphological characteristics of panicles. Such differences among ethnic groups were observed by comparing on the one hand the diversity of variety names within morphological categories, and on the other hand the morphological diversity covered by each variety name. Loosely speaking, this represents how well defined are varieties named by farmers, or how vague they are regarding the morphological characteristics. Interestingly, the Mbeere, as a group, was found to be less accurate and consistent than the other groups in classifying panicles regarding their morphological characteristics. Indeed, the morphological diversity covered by variety names and the diversity of names within the morphological categories were the highest in the Mbeere group. This was notably due to the extensive use of the inclusive category *Ngirigacha* in this ethnic group.

The Chuka taxonomy was found to be especially accurate in classifying *Muruge* regarding morphological characteristics, and similar results were found for *Mugeta* in the Tharaka group. In both cases, these variety names covered the lowest morphological diversity as they were bounded within few morphological categories in which they were largely dominant. It is noteworthy that the three ethnic groups showed such clear differences in the classification of the recently introduced *Gadam*. On the one hand, the Chuka appeared to use extensively and abusively the name *Gadam*. On the other hand, the Mbeere and the Tharaka seemed to have a restrictive use of this name. Indeed, the Chuka largely named *Gadam* panicles showing semi-compact shape, and assigned to the genetic pool of local landraces (data not shown). Reversely, the Mbeere and Tharaka largely identified as *Mbura-imwe* or *Ngirigacha* panicles showing compact-elliptic shape, and showing the genetic signature of *Gadam*. *Gadam* variety was introduced in the area about two years ago at the time of our study, and was still in course of diffusion. Interestingly, *Gadam* presented morphological similarity with the germplasm present locally (Labeyrie et al., submitted), which seems to cause some disturbance in the classification system of farmers. The Tharaka and the Mbeere did not appear to be fully experienced in its identification and naming as they mostly used the inclusive category names *Mbura-imwe* and *Ngirigacha* for this purpose. On the contrary, the Chuka mostly adopted this name, but they also used it abusively to identify local germplasm.

Conclusion

Our probabilistic analysis of variety names evidenced the lack of correspondence between the structure of sorghum genetic diversity and the taxonomy used by farmers as a group. Moreover, the set of panicles sampled displayed a wide and continuous range of morphological variability, but farmers proved to be able to identify, name and classify different varieties accurately and consistently in such a continuum. Farmers' taxonomy was based on a small number of salient morphological traits, and farmers' varieties displayed an overall distinction and uniformity regarding these criteria. However, intercultural differences were found in the morphological distinction and uniformity of named varieties because the accuracy and consistency of farmers in naming and classifying varieties regarding their morphological characteristics differed among ethnic groups.

A major methodological contribution of our study was the use of probabilistic approaches to describe farmers' taxonomy at the community scale, while previous studies were affected by the high inter-informant variability because they relied on a single name attached to each plant by one informant. We showed that the taxonomy used by the majority of farmers' in each ethnic group was based on eight named categories and two inclusive categories for the

Tharaka and the Mbeere, while 76 different names were inventoried in the identification task. Individually, farmers cited various names while collectively they use a small number of names to classify the diversity of morphotypes. The inter-individual variability in naming varieties is thus considerable, and justifies the approach used in this study.

By depicting the overall correspondence between variety names used by communities and plants morphological characteristics, this study raised the question of inter-individual consistency in naming varieties. Indeed, we observed that the accuracy and consistency of farmers in naming varieties regarding their morphological characteristics was limited in some cases. Hence, the study of inter-informants agreement patterns would give further insight concerning the reasons of such discrepancies.

REFERENCES

Atran, S., and D. Medin. 2008. *The native mind and the cultural construction of nature*. Cambridge, Massachusetts: MIT press.

Badstue, L.B., M.R. Bellon, X. Juárez, I. Manuel, and A.M. Solano. 2002. Social Relations and Seed Transactions among Small-Scale Maize Farmers in the Central Valleys of Oaxaca, Mexico. *CIMMYT Economics Working Paper* No. 02–02.

Badstue, L.B., M.R. Bellon, J. Berthaud, A. Ramírez, D. Flores, and X. Juárez. 2007. The dynamics of farmers' maize seed supply practices in the central valleys of Oaxaca, Mexico. *World Development* 35 (9):1579-1593.

Barnaud, A., M. Deu, E. Garine, D. McKey, and H. Joly. 2007. Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theoretical and Applied Genetics* 114 (2):237-248.

Barnaud, A., G. Trigueros, D. McKey, and H. Joly. 2008. High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity* 101 (5):445-452.

Bellon, M.R., and S. B. Brush. 1994. Keepers of maize in Chiapas, Mexico. *Economic Botany* 48 (2):196-209.

Benzécri, J.P. 1973. *L'analyse des données. II L'analyse des correspondances*. Paris: Bordas.

Berlin, B. 1973. Folk systematics in relation to biological classification and nomenclature. *Annual Review of Ecology and Systematics* 4:259-271.

Boster, J.S. 1985. Selection for perceptual distinctiveness: evidence from Aguaruna cultivars of Manihot esculenta. *Economic Botany* 39 (3):310-325.

Boster, J.S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88 (2):428-436.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks.

Darwin, C. 1859. *On the origins of species by means of natural selection*. London: Murray.

Dray, S., and A.B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22 (4):1-20.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14 (8):2611-2620.

Fadiman, J.A. 1993. *When we began there were witchmen: an oral history from Mount Kenya*. Los Angeles: University of California Press.

Falush, D., M. Stephens, and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164 (4):1567-1587.

Friedberg, C. 1991. Operative aspects of Folk Classification. In *Man and a half: essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer*, ed. A. Pawley, 102-109. Auckland: The Polynesian Society.

Gibson, R.W. 2009. A review of perceptual distinctiveness in landraces including an analysis of how its roles have been overlooked in plant breeding for low-input farming systems. *Economic Botany* 63 (3):242-255.

Glazier, J. 1970. Ritual and social conflict: circumcision and oath-taking in Mbeere. Nairobi: Institute of Development Studies, Nairobi University.

Grenand, F. (2002). 2002. Stratégies de nomination des plantes cultivées dans une société tupi-guarani, les Wayãpi. *Amerindia* 26/27:209-247.

Harlan, J. R., and J.M.J. De Wet. 1972. A simplified classification of cultivated sorghum. *Crop Science* 12 (2):172-176.

Harlan, J.R., J.M.J. De Wet, and A.B.L. Stemler. 1976 *Origins of African plant domestication*. The Hague: De Gruyter Mouton.

Hodgkin, T., R. Rana, J. Tuxill, D. Balma, A. Subedi, I. Mar, R. Karamura et al. 2007. Seed Systems and Crop Genetic Diversity in Agroecosystems. In *Managing biodiversity in agricultural ecosystems*, eds. D. Jarvis, C. Padoch, and D. Cooper, 77-116. New York: Colombia University Press.

Hunn, E. 1982. The utilitarian factor in folk biological classification. *American Anthropologist* 84 (4):830-847.

International Plant Genetic Resources Institute (IPGRI). 1993. *Descriptors for Sorghum [Sorghum bicolor (L.) Moench]*. Rome: IBPGR / ICRISAT.

Jaetzold, R., H. Schmidt, B. Hornetz, and C. Shisanya. 2007. *Farm management handbook. Vol II, Part C, East Kenya. Subpart C1, Eastern Province*. Nairobi Ministry of Agriculture / GTZ.

Jarvis, D., A. Brown, P. Cuong, L. Collado-Panduro, L. Latournerie-Moreno, S. Gyawali, T.Tanto et al. 2008. A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. *Proceedings of the National Academy of Sciences of the United States of America* 105 (14):5326-5331.

Jost, L. 2006. Entropy and diversity. *Oikos* 113 (2):363-375.

Labeyrie, V., M. Deu, A. Barnaud, C. Catalayud, M. Buiron, P. Wambugu, S. Manel, J.-C. Glaszmann, and C. Leclerc. 2013. Influence of ethnolinguistic diversity on sorghum evolution in subsistence farming systems in eastern Kenya. *Submitted for publication.*

Louette, D., A. Charrier, and J. Berthaud. 1997. In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Economic Botany* 51 (1):20-38.

Louette, D., and M. Smale. 2000. Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. *Euphytica* 113 (1):25-41.

McGuire, S. 2008. Securing access to seed: Social relations and sorghum seed exchange in eastern Ethiopia. *Human Ecology* 36 (3):217-229.

Middleton, J. 1953. *The central tribes of the north-eastern Bantu*. Ethnographic Survey of Africa. London: International African Institute.

Moehlig, W.J.G., G. Guarisma, and S. Platiel. 1980. La dialectométrie: une méthode de classification synchronique en Afrique. In *Dialectologie et comparatisme en Afrique Noire*, 27-45.

Mwaniki, H.S.K.1973. *The living history of Embu and Mbeere to 1906*. East African Literature Bureau.

Nuijten, E., and C.J. Almekinders. 2008. Mechanisms explaining variety naming by farmers and name consistency of rice varieties in the Gambia. *Economic Botany* 62 (2):148-160.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289-290.

Perrier, X., and J.P. Jacquemoud-Collet. 2006. DARwin software. Montpellier: CIRAD.

Pike, K.L.1954. *Language in Relation to a Unified Theory of the Structure of Human Behavior*. Glendale: Summer Institute of Linguistics.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155 (2):945-959.

Quiros, C.F., S. Brush, D.S. Douches, K.S. Zimmerer, and G. Huestis. 1990. Biochemical and Folk Assessment of Variability of Andean Cultivated Potatoes. *Economic Botany* 44 (2):254–266.

R Development Core Team. 2011. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/.

Rabbi, I.Y., H.H. Geiger, B.I. Haussmann, D. Kiambi, R. Folkertsma, and H.K. Parzies. 2010. Impact of farmers' practices and seed systems on the genetic structure of common sorghum varieties in Kenya and Sudan. *Plant Genetic Resources: Characterization and Utilization* 8 (2):116-126.

Sadiki, M., D. Jarvis, D. Rijal, J. Bajracharya, N. Hue, T. Camacho-Villa, L. Burgos-May et al. 2007. Variety names: An entry point to crop genetic diversity and distribution in agroecosystems? In *Managing Biodiversity in Agricultural Ecosystems*, eds. D. Jarvis, C. Padoch, and D. Cooper, 34–76. New York: Colombia University Press.

Saitou, N., and M. Nei. 1987 The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Salick, J., N. Cellinese, and S. Knapp. 1997. Indigenous Diversity of Cassava: Generation, Maintenance, Use and Loss among the Amuesha, Peruvian Upper Amazon. *Economic Botany* 51 (1):6–19.

Saunders, B.A.C. 2007. Towards a new topology of colour. In A*nthropology of color: Interdisciplinary multilevel modeling*, eds. R. E. MacLaury, G.V. Paramei, and D. Dedrick, 467–479. Amsterdam: John Benjamins.

Snowden, J.D. 1936. *The cultivated races of sorghum*. London: Adlard and Son.

Sokal, R., and C.D. Michener. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38:1409-1438.

Soler, C., A.A. Saidou, T.V.C. Hamadou, M. Pautasso, J. Wencelius, and H.H. Joly. 2013. Correspondence between genetic structure and farmers' taxonomy - a case study from dry-season sorghum landraces in northern Cameroon. *Plant Genetic Resources: Characterization and Utilization* 11 (1):36-49.

Teshome, A., B.R. Baum, L. Fahrig, J.K. Torrance, T.J. Arnason, and J.D. Lambert. 1997. Sorghum [Sorghum bicolor (L.) Moench] landrace variation and classification in north Shewa and south Welo, Ethiopia. *Euphytica* 97 (3):255–263.

Therneau, T., B. Atkinson, B. Ripley, and 2012. rpart: Recursive Partitioning: R package version

Vom Brocke, K., A. Christinck, R.E. Weltzien, T. Presterl, and H.H. Geiger. 2003. Farmers' seed systems and management practices determine pearl millet genetic diversity patterns in semiarid regions of India. *Crop Science* 43 (5):1680-1689.

SUPPLEMENTARY MATERIALS

Table III-S1. Variety nomenclature. List of the basic roots (A.) and determinants (B.), their meaning and translation

A.

Roots	Meaning	Translation
Motivated		
GATUMI	appearance	"Tight"
KIAMANGUO	appearance	"Furry"
MWITUMIRIA	appearance	"Closed"
MUKUMBU	appearance	"Compact"
MBUNGE	appearance	"Fully covered"
MSALAMA	attitude	"Peaceful"
CIANKOMA	attitude	"Mad"
GATENGU	attitude	"Provoke"
GATENGURIRA	attitude	"Provoke"
MUBENANIA	attitude	"Liar"
KITHARARA	attitude	"Stupid""
KARUIRA	attitude	"Fight"
MACUIRI	anatomy	"Hair"
MACARA	anatomy	"Hand"
KAMBUMBU	animal	"Bat"
MAGUNA	animal	"Monkey"
MAGWIKA	animal	"Baboon"
MUCERU	color	"White"
MURUGE	color	"as cooked"
NTUNE	color	"Red"
GADAM	commercial variety	
KAGURU	commercial variety	Research station
MTAMA.ONE	commercial variety	
SERENDO	commercial variety	
KIMINCHI	crop species	"Green peas"

MUVIA	crop species	"Sorghum"
KARURU	taste	"Bitter"
MUGANA	taste	"Bitter"
GACHOBI	use	"for beer"
NDUME	use	"Beer" (traditional)
MUTHIGO	processing	"dehulling"
KIAMIGA.TE	food	"Bread"
NAIRONI	material	"Nylon"
MWITIKA	origin	"feral"
NGIRIGACHA	agriculture	"Agriculture"
GAKUNDI	undetermined	"First"

Unmotivated
CIAMABUA
CIAMAGURU
CIANJOMI
DIGA
GACHUKU
GATIGA
GETIA
KARIGU
KARUGU
KATHIRIGWA
KIBOFI
KITHUKE
MATHIGUKA
MBARIE
MUCURA
MUCURI
MUGAMBI
MUGETA
MUTHUNGA

B.

Determinant	Meaning	Translation
MURUGE	color	"as cooked"
MUTUNE (NTUNE)*	color	"Red"
MBILU	color	"Black"
MUCERU (NJERU)*	color	"White"
KIMBEERE	ethny	Mbeere
KICHUKA	ethny	Chuka
MBURA CIIRI	growth-cycle length	"Two
MBURA IMWE	growth-cycle length	"One
IGUURU	plant height	"Top"
KA NTHUKE IMWE (IRIA NTHUKE)*	plant height	"same age"
NAIRONI	material	"Nylon"
ITHETU	material	"Soil"

* names into brackets are the kimbeere spelling

CHAPITRE IV- IDENTIFICATION ET NOMINATION DES VARIETES : LA PART DE LA CULTURE

ARTICLE: FARMERS' CONSISTENCY IN NAMING SORGHUM VARIETIES: THE SHARE OF CULTURE

V. Labeyrie, J. I. Kamau, C. Leclerc

En préparation pour soumission au journal *Philosophical Transactions of the Royal Society B-Biological Sciences*

FARMERS' CONSISTENCY IN NAMING SORGHUM VARIETIES: THE SHARE OF CULTURE

V. Labeyrie, J. I. Kamau, C. Leclerc

En préparation pour soumission au journal *Philosophical Transactions of the Royal Society B-Biological Sciences*

ABSTRACT

To manage the phenotypic variability of crops, farmers use taxonomies as a frame for reasoning, communicating and learning about it. The identification, naming and classification of varieties regarding their perceptual characteristics is socially learnt and further constructed by individuals through their direct experience of the plants. This study aims at characterizing the share of farmers' culture in their identification and naming of sorghum varieties among three ethnolinguistic groups in eastern Kenya. It further investigates whether the history of the different varieties in the area influences the consistency of farmers in naming them.

A probabilistic approach was used to investigate the agreement patterns of farmers in identifying and naming sorghum varieties. A set of 293 panicles representative of the phenotypic diversity of sorghum in the area of study was presented independently to 32 female informants in each ethnic group for identification. The name used by each informant to identify each panicle was recorded, and an agreement index was computed between pairs of informants using the simple-matching index. The mean agreement within ethnolinguistic groups was compared. It was further computed separately for subsets of panicles according to their history and improvement status, traced with genetic molecular markers. The mean agreement within and between ethnolinguistic groups was computed for: i. an anciently released improved variety, ii. a recently released improved variety, iii. local short-cycle varieties and iv. local long-cycle varieties.

Our results suggest that the diffusion of experience concerning varieties' identification and naming was limited among ethnic group. Furthermore, agreement patterns for naming varieties reflected the social relationship among the three ethnic groups. The history of varieties in the area determined the consistency of farmers in naming them. Anciently

released improved varieties were named highly consistently, suggesting a uniform distribution of knowledge on their identification and naming among ethnolinguistic groups, which may result from their common learning from an external source of information: the extension services or the NGOs. On the contrary, farmers in the three groups displayed large differences of consistence in naming the various landraces.

INTRODUCTION

Humans have been in close relation to crops for about 13 000 years, and a large share of the world population continues to entertain this inter-dependence relation in subsistence farming systems (Diamond 2002; Gepts 2004). In these systems, the detailed appreciation of crop varieties' characteristics is of upmost importance for the survival of societies as part of their adaptative strategy. Societies hence developed classification systems to categorize crop diversity just like they did for the rest of their environment. Classification systems are used as a frame for reasoning and managing it (Atran and Medin 2008). These so called "folk taxonomies" involve identification, naming and classification processes, which are distinct even though related (Friedberg 1991). The identification is a perceptual process through which farmers assign a plant to a category on the basis of its perceived characteristics. Naming is the process through which these categories are labeled, mainly for communication purpose.

The identification, naming and classification of varieties regarding their morphological characteristics is socially learnt and further constructed by individuals through their direct experience of the plants (Boster 1986). As stated by Boster: "*For informants to agree in the identification of something, they must recognize it as a member of the same distinct category and must label the category with the same name*". On the one hand, he argues that the identification process alone largely relies on farmers' experience in distinguishing the characteristics of the varieties. Agreement among farmers in identifying a variety would hence rely on their shared experience of its morphological characteristics. On the other hand, the sharing of a variety name to label a given morphological category would depend on the path for learning and knowledge transmission. Indeed, names are used for communicating, it is hence essential to learn what the culturally appropriate name to label a given category is.

The naming of a plant by an individual hence reflects the experience that he had from the variety, and the culture in which his learning took place.

Few studies dealt with the issue of crop variety identification and naming by farmers since Boster's work. Nuijten and Almekinders (2008) observed in Gambia that overall, inter-informant agreement in naming rice varieties was higher within villages than among, and that some villages presented lower inter-informant agreement than others. Low inter-informant agreement was also reported by Salick et al. (1997) for naming cassava clones in Peru, and in few other studies (Sadiki et al. 2007). None of these studies, however, investigated the role of farmers' culture in identifying and naming of crop varieties consistently.

In this study, we addressed the issue of cultural differences in the identification and naming of sorghum varieties on an ethnolinguistic contact zone in eastern Kenya. In a previous study (Labeyrie et al., in prep.) the Chuka, Tharaka and Mbeere adjacent ethnolinguistic groups were found to differ in their accuracy and consistency in naming sorghum varieties regarding their morphological characteristics even though they shared an overall common taxonomy. This suggested that farmers among the three ethnic groups did not fully shared experience concerning the identification and naming of sorghum varieties, despite their close spatial adjacency. This question is hence specifically investigated here. We tested whether farmers within and among adjacent ethnic groups shared experience in identifying and naming the various sorghum varieties grown in the area of study. This was done through the characterization of their agreement pattern in naming a set of panicles representing the diversity of sorghum in the area. Indeed, farmers are expected to name panicles in the same way if they share experience of varieties' morphological characteristics as well as a common nomenclature (Boster 1986).

Furthermore, this study gives insights concerning the exchange of knowledge on varieties identification and naming among and within ethnolinguistic groups. If information exchanges are not limited, high consistency in naming varieties is expected among and within ethnic groups. Departures from this hypothesis are discussed on the one hand regarding the characteristics of the farming communities: their social organization and the modalities of knowledge transmission (Reyes-García et al. 2009; Cavalli-Sforza and Feldman 1981), and on the other hand, regarding the characteristics of the sorghum varieties. Indeed, as agreement among farmers increases with their shared experience of varieties, it is expected to depend also on the history of varieties' introduction and on their commonness in the area. Some varieties were introduced in the area of study by national breeding programs a long time ago. Their diffusion and adoption in the three groups was a common experience, and the names used to identify these varieties are expected to be used consistently by farmers. On the contrary, we expect that rare landraces would not be consistently named by farmers, because, following Boster (1986), they cannot agree in naming a variety from which they have no experience.

This study provides a further formalization of folk taxonomy for crop varieties, which has been largely neglected as compared to the species level. It is relevant for the characterization, the collection and the conservation of crop diversity *on-farm* as most studies based their sampling strategies and estimations of intraspecific diversity on the variety names used by farmers.

MATERIALS AND METHODS

This study was conducted in an ethnolinguistic contact zone between Chuka, Tharaka, and Mbeere ethnic groups. The site was selected for conducting a cross-cultural comparison given the absence of spatial isolation between ethnic groups. Furthermore, the study area presented uniform agro-ecological conditions (Jaetzold et al. 2007), making it possible to control for spatial and environmental biases.

Chuka, Tharaka and Mbeere farmers have maintained different ethnic, linguistic and cultural characteristics despite their spatial proximity on Mount Kenya (Middleton 1953; Heine and Moehlig 1980; Lambert 1956). Linguistic and cultural differences suggest that the relationships between these ethnic groups are socially restricted. The relationship between the three groups was described by ethnographers and colonial administrators. Mbeere had tensed relationship with Chuka who were their direct neighbor (Glazier 1970; Mwaniki 1973) and the current land-pressure reinforces this rivalry. On the contrary, Tharaka and Chuka were allied in the past and consider they are kin, or *gishiaro* (Fadiman 1993). Intermarriage is frequent between Tharaka and Chuka while it is very uncommon between the two latter groups and the Mbeere, according to observations by one of us (VL). This ethnic organization and relationship system relates to the spatial organization of the three ethnic groups. Tharaka and Chuka were spatially mixed and settled in the northern part of the study site while the Mbeere were located separately from the two other groups in the southern part of the area.

Varieties identification survey

293 sorghum panicles were presented to a panel of women representative from the three ethnic groups. The sorghum panicles used for the identification survey were selected to represent the morphological diversity of the varieties collected *on-farm* in the three ethnic groups. They were selected among a larger set of descendants from panicles collected on-farm in January and July 2011 that were grown in an experimental field under controlled conditions, as described in detail in Labeyrie et al. (Labeyrie et al., submitted).

Thirty-two female informants were randomly chosen in each ethnic group and invited to participate to the identification survey. The native and husband's ethnic group and age of each informant were recorded. Following the procedure used by Boster (1986) for the study of cassava clones naming among the Aguaruna Indians, each informant was independently asked to identify each of the 293 panicles that were individually presented to her. A field assistant recorded the varieties names used by each informant to identify each panicle.

Analyses

The methodology used for this study follows the probabilistic approach used by Boster (1986), and developed later by Romney et al. (1986) to measure the knowledge of individuals by reference to that of their cultural group. However, we did not apply this conceptual frame to measure individual knowledge, but rather to measure the shared experience of farmers within and among ethnic groups in identifying and naming sorghum varieties. A low concordance between the names used by farmers to identify the same panicles would indicate either that they do not share experience concerning the morphological characteristics of the varieties considered because they did not experienced it physically, or that farmers belong to different cultural groups and that they do not share a common pool of information (D'Andrade 1981).

To measure the concordance of the names used by informants to identify the set of panicle, an index of agreement between each pair of informants was calculated based on the Simple-Matching index (Sokal and Michener 1958). It was calculated as the ratio of the number of times the pair of informants used the same names to identify panicles divided by the overall number of panicles they identified:

$$Ag = {No.agree}\Big/{No.agree + No.disagree}$$

To compare the strength of inter-informant agreement within and among ethnic groups, we computed the mean agreement over all pairs of informants within each ethnic group (intra-group), and over pairs of informants belonging to different ethnic groups (inter-groups). To test the significance of the differences of intra-group agreement among ethnic groups, a permutational multivariate analysis of variance (perMANOVA) was performed (Anderson 2001) using the *adonis* function in the R package *vegan* (Oksanen et al. 2012). This analysis is performed on a squared matrix of pairwise disagreement (1-*Ag*) between informants, which is partitioned according to ethnic groups. Following the ANOVA principle, the sum of squared disagreement within groups (which is the sum of squared distances from individual replicates to their group centroid) and between groups (which is the sum of squared distances from group centroids to the overall centroid) are compared. A pseudo F-ratio is then computed and compared to its distribution under the null hypothesis simulated using 4000 random permutations of the raw data.

To test whether the inter-informant agreement was influenced by the origin and improvement status of varieties, the mean intra-group agreement was computed separately for i) the recently introduced variety *Gadam* (86 panicles)*, ii)* the anciently introduced variety *Kaguru* (40 panicles), iii) local short-cycle landraces (110 panicles), and iv) local long-cycle landraces (15 panicles). The assignment of panicles to one of these four genetic groups was done on the basis of previous genetic studies (Labeyrie et al., submitted).

RESULTS

Overall consistency of farmers in naming varieties differs among ethnic groups

The overall consistency of farmers in naming the whole set of panicles differed among ethnic groups (Figure IV-1.A). The mean inter-informant agreement was especially high within the Chuka group. Non-parametric perMANOVA confirmed that these differences were significant at a 1 % threshold (Supplementary materials 1). In addition, the mean agreement within the Chuka group was clearly higher than between the Chuka and members of the two other ethnic groups. The lowest mean agreement was observed between the Mbeere and the Tharaka (Figure IV-1.B).

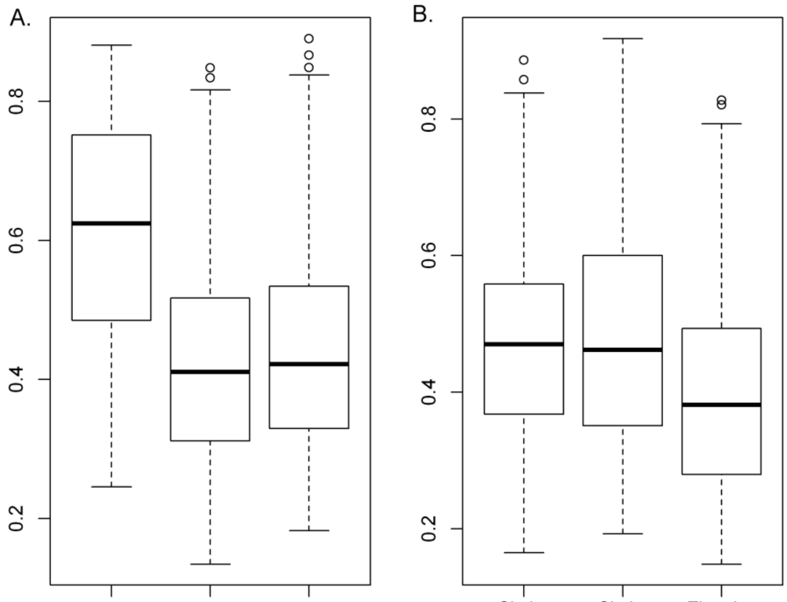


Figure IV-1. Inter-informants agreement within (left) and between (right) ethnic groups. Mean, quartiles and standard deviation.

Consistency of farmers in naming varieties regarding their history

The consistency of farmers in naming sorghum varieties varied clearly depending on the history of the variety considered, notably regarding its improvement status and date of introduction (Figure IV-2). The three ethnic groups displayed a high agreement in naming anciently introduced varieties (Figure IV-2.A) while their agreement was lower for local landraces (Figure IV-2.C) and recent improved varieties (Figure IV-2.B). Furthermore, the strength of intra-group agreement differed among ethnic groups for these latter categories (perMANOVA p-value < 0.001). Differences of intra-group agreement were observed for the various local landraces among ethnic groups (Figure IV-2).

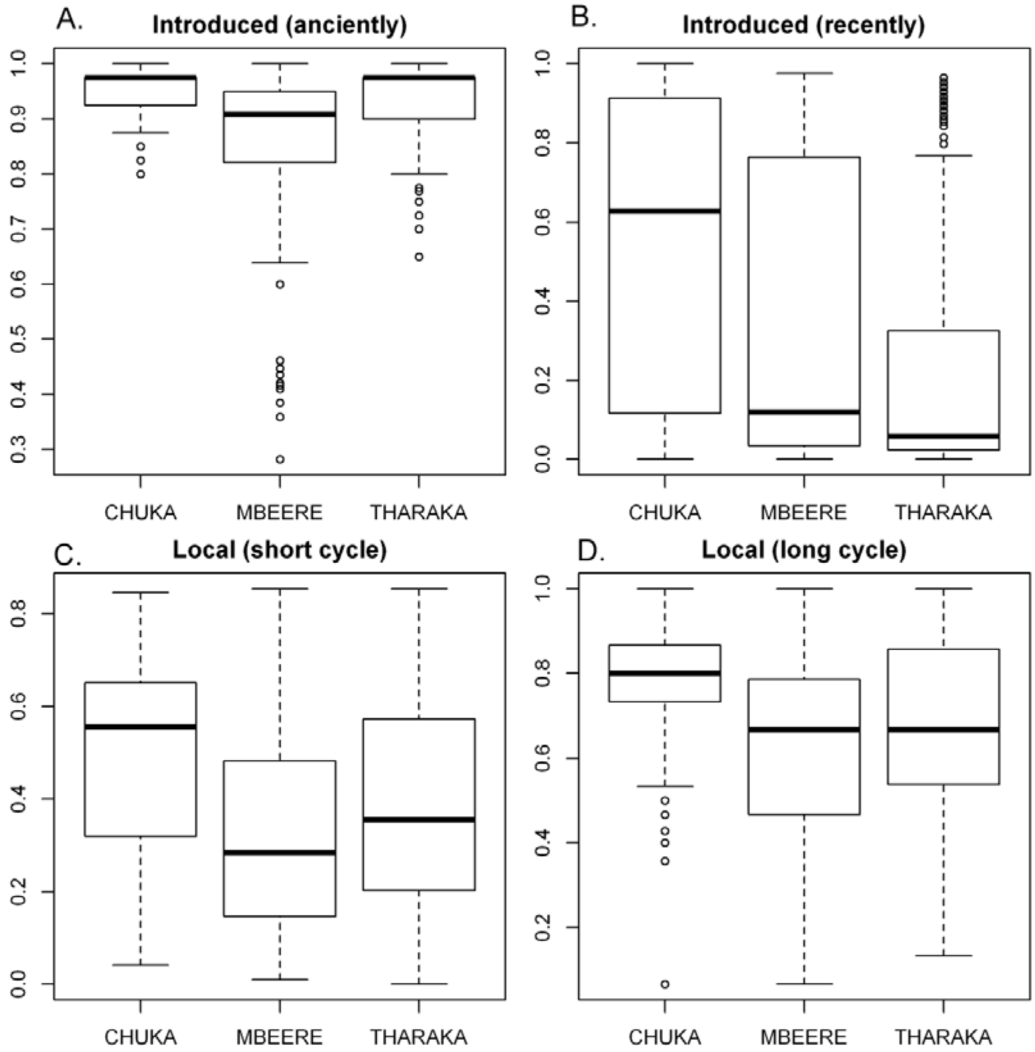


Figure IV-2. Intra-group agreement in naming improved varieties and local landraces. Mean, quartiles and standard deviation.

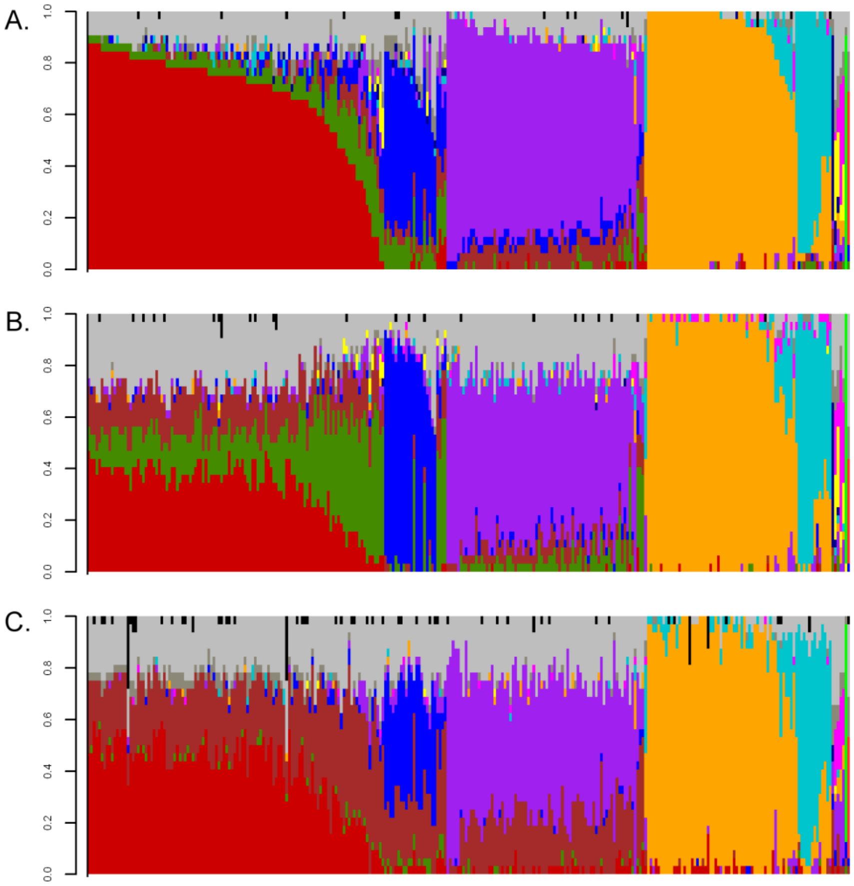


Figure IV-3. Fraction of informants who cited each variety name to identify each panicle. Each vertical bar represents a panicle and is divided in 10 colored sections representing the fraction of informants who cited each variety name. From top to bottom: Chuka, Tharaka, Mbeere. Red: *Gadam*, green: *Mbura Imwe*, brown: *Ngirigacha*, blue: *Mugeta*, purple: *Muruge*, orange: *Kaguru*, turquoise: *Serendo*, dark blue: *Kathirigwa*, yellow: *Mucuri*, magenta: *Mugana*, light green: *Mukumbu*, brick red: *Mbunge*, light grey: others, dark grey: unknown, black: missing data.

Improved varieties

The mean intra-group agreement in naming the improved variety *Kaguru* that was anciently introduced in the area (around the 1990's) was exceptionally high in the three ethnic group (Figure IV-2.A), and did not differ among groups (perMANOVA p-value = 0.624). Similar results were observed for another improved variety which was introduced anciently: *Serendo* (in turquoise, Figure IV-3). By contrast, the consistency of farmers in naming the recently introduced variety *Gadam* differed strongly among ethnic groups (Figure IV-2.B; perMANOVA p-value < 0.001). The Chuka informants were remarkably consistent in using *Gadam* name, contrary to the informants in the two other ethnic groups. The Chuka consistently identified a large set of panicles as *Gadam* (in red, Figure IV-3.A) while a substantial share of the Tharaka and Mbeere informants identified this same set of panicles as *Ngirigacha* (in brown, Figure IV-3.C) or *Mbura imwe* (in green, Figure IV-3.B).

Local varieties

The three ethnic groups were found to differ in their degree of agreement in naming the local varieties (perMANOVA p-value < 0.001 for both short-cycle and long-cycle varieties). The Chuka notably displayed the highest intra-group agreement in naming local varieties, and the Mbeere the lowest while the Tharaka were in-between (Figure IV-2.C and IV-2.D). In addition, members from the Chuka and the Tharaka groups agreed more strongly together in naming local varieties than they agreed with the Mbeere for both short and long cycle local varieties (Figure IV-4.A and IV-4.B).

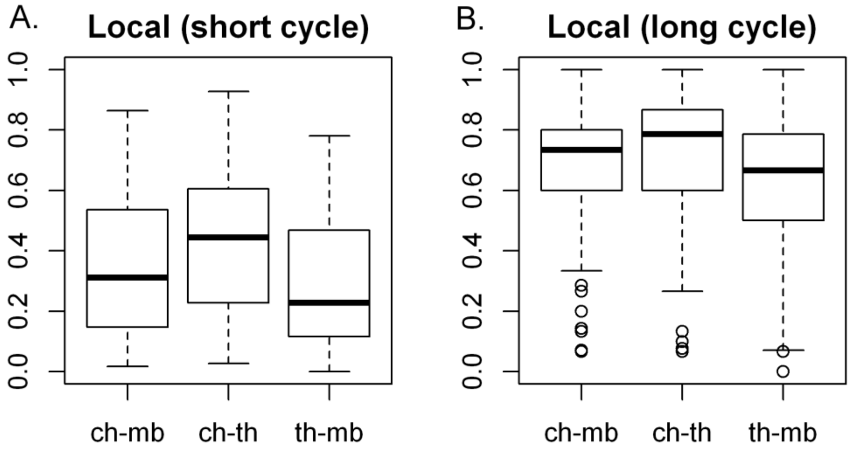


Figure IV-4. Inter-group agreement of farmers in naming the local varieties. ch: Chuka, mb: Mbeere, th: Tharaka

The high consistency of the Chuka in naming local varieties was mainly due to their noteworthy consistency in naming the *Muruge* variety (in purple, Figure IV-3.A), which represented about one fourth of the overall panicle set. By contrast, the Mbeere showed the lowest consistency in using the name *Muruge*. The set of panicles that was highly consistently named *Muruge* by the Chuka was named *Ngirigacha* (in brown, Figure IV-3.C) by a large proportion of the Mbeere. The Tharaka showed a degree of consistency in naming *Muruge* in-between that of the two other groups (Figure IV-3.B).

Nevertheless, the overall comparison of inter-informant agreement in naming the local varieties did not describe accurately the differences of consistency in using the name *Mugeta* among ethnic groups. Indeed, the Tharaka were found highly consistent in naming *Mugeta* variety (in blue, Figure IV-3.B) as compared to the other ethnic groups, especially the Mbeere (Figure IV-3.C). Last, farmers in the three ethnic groups were quite consistent in naming long-cycle varieties. Important intra-group variation was nonetheless observed in the Tharaka and especially in the Mbeere groups, while the Chuka were the most consistent in naming these varieties and displayed little variation in their intra-group degree of agreement as compared to the two other groups (Figure IV-2).

Uncommon variety names

The variety names *Mukumbu*, *Kathirigwa*, *Mugana*, *Mucuri*, and *Mbunge* were rarely cited, but they were used with varying level of consistency by informants to identify a small set of panicles. The name *Mukumbu* was used highly consistently by informants in the three groups to identify a single panicle (Figure IV-5, in light green). Indeed, 84 % (Mbeere), 91 % (Chuka) and 97 % (Tharaka) of the informants identified this panicle as *Mukumbu* variety. The consistency of informants in using the variety name *Kathirigwa* (in dark blue) differed strongly among ethnic groups. This name was used by a majority of Chuka informants (72 %) to identify a single panicle while the Mbeere were clearly inconsistent in naming this same panicle, and the Tharaka were in-between (56 %). The Chuka informants were also more consistent in using the name *Mbunge* (59 %, in brick red) than the Tharaka (41 %) and the Mbeere (48 %). The use of *Mugana* (in magenta) and *Mucuri* (in yellow) names was inconsistent in the three ethnic groups for the panicles studied, and especially in the Mbeere group.

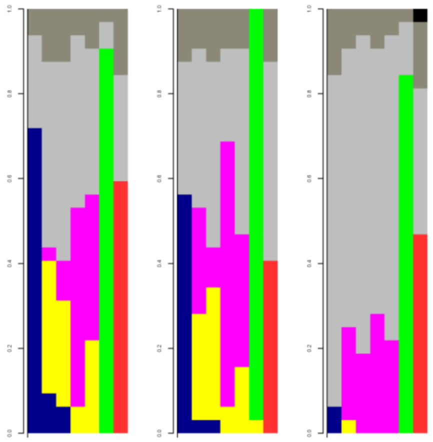


Figure IV-5. Zoom on the fraction of informants who cited the uncommon variety names. Each vertical bar represents a panicle and is divided in 10 colored sections representing the fraction of informants who cited each variety name. From left to right: Chuka, Tharaka, Mbeere. Dark blue: *Kathrigwa*, yellow: *Mucuri*, magenta: *Mugana*, light green: *Mukumbu*, brick red: *Mbunge*, light grey: others, dark grey: unknown, black: missing data

DISCUSSION

In this study, we showed that intra-group agreement varied strongly among ethnic groups. The comparison of inter-informants agreement patterns gives insight concerning their shared experience of the morphological characteristics of varieties on the one hand, and the diffusion of nomenclature on the other hand (Boster 1986). Our results first suggest that the three ethnic groups did not fully share experience concerning varieties’ identification or that their nomenclature differed in some extent, because the overall agreement between them was low. Furthermore, farmers within the Chuka group clearly shared experience concerning varieties’ identification and referred to a common nomenclature, while it was not as clear in the Mbeere and Tharaka groups.

Such differences of farmers’ consistency in naming varieties were previously observed among villages in Gambia, farmers in some villages showing less consistency in naming rice varieties than in others (Nuijten and Almekinders 2008). Authors suggested that these patterns were linked to a differential intensity of seed exchanges within villages. This argument was also developed by Boster, the degree of agreement between informants in naming cassava clones being linked to the intensity of planting material exchanges between them (Boster 1986). Boster notably argued that seed exchanges increased the shared experience of farmers concerning the morphological characteristics of varieties. Nevertheless, the overall agreement patterns hide noteworthy differences of intra and inter-group agreement among varieties, notably regarding their improvement status and date of introduction.

Farmers' consistence in naming improved varieties depends on their date of introduction

Kaguru and *Serendo* improved varieties, which were anciently introduced in the area, were identified and named with a noteworthy consistence by farmers in the three ethnic groups. This relates to the uniform distribution of knowledge concerning the identification and naming of these varieties among ethnic groups. Indeed, farmers in the three groups probably learnt uniformly about these varieties from an external source of information: the extension services or the NGOs. In addition, *Kaguru* was extensively grown in the area mainly because of its high price on the market, which probably contributed to increase the shared experience of farmers concerning this variety and the diffusion of its name.

By contrast, farmers in the three ethnic groups displayed striking differences of consistency in identifying and naming *Gadam* variety (commercial name), which was recently introduced at the time of the survey as its diffusion started two years ago. The large majority of Chuka farmers appeared to be familiar with that variety at the time of the survey and used consistently its commercial name. In the Tharaka and Mbeere groups, however, farmers did not shared experience concerning its identification and naming. Some of them used the commercial name *Gadam* but a large share of informants used the names *Mbura-imwe* and *Ngirigacha.* Previous study suggested that *Mbura-imwe* and *Ngirigacha* were used to label inclusive categories, i.e. higher order taxa (Labeyrie et al. , in prep.). The assignment of this new variety to these categories hence denotes that farmers were not familiar with its commercial name and probably did not learn how to identify it yet. Various other names were also used by the Mbeere and the Tharaka to identify *Gadam*. The reasons for such differences are unknown, but they suggest that the extension services or NGO may have distributed seeds mainly in the Chuka area and information did not diffuse in the other ethnic groups at the moment of the survey.

Farmers' consistence in naming local varieties reflects their social relationship

This study showed that despite their spatial proximity, farmers in the three ethnic groups displayed different degrees of consistency in naming local sorghum varieties. The Chuka farmers were the most consistent in naming local varieties while the Mbeere farmers were the less consistent, and the Tharaka were in-between. The strength of inter-informant agreement reflects on the one hand their experience of the varieties, and on the other hand the social paths of knowledge transmission and learning (Romney et al. 1986; Boster 1986). Hence, our observations suggest first that the Mbeere as a group had a limited experience of the local sorghum varieties as compared to the Chuka, and secondly that knowledge transmission from the Chuka to the Mbeere was limited despite their spatial proximity. Indeed, if some Mbeere households settled in the area of study long time ago, most Mbeere farmers migrated during the last decades and their tensed relationship with the Chuka (Mwaniki 1973; Glazier 1970) could explain why experience and nomenclature was not shared between the two groups.

By contrast, the inter-group agreement in naming local varieties was the highest between the Chuka and the Tharaka group. This is congruent with the descriptions of the relationship between the two groups found in the literature and with our own field observations. The Chuka and Tharaka consider they are blood-brother (*gishiaro*), resulting from their common ascendance (Fadiman 1993). Intermarriage is frequent between the two groups. It is hence likely that the close social relationship between the two groups favored the sharing of experience and nomenclature, homogenizing their identification and naming processes. Nevertheless, the homogenization of identification process and nomenclature between the Chuka and the Tharaka was no full. Notably, the Chuka were quite inconsistent in using the name *Mugeta* while the Tharaka used it with a noteworthy consistence. This suggests that the shared experience of the Tharaka informants concerning the identification and naming of *Mugeta* was much higher than that of the other ethnic groups. These observations are in line with the fact that *Mugeta* was considered as a Tharaka variety by other famers in the area (Personal observation). *Mugeta* is the third most frequently grown variety in the Tharaka lowlands (Labeyrie et al. 2013).

Last, a large share of the Mbeere and Tharaka informants used the names *Ngirigacha* and *Mbura-imwe* quite inconsistently. These names have a wide meaning, *Ngirigacha* (meaning "Agriculture") and *Mbura - imwe* ("One season"). As suggested previously, these names may in fact label inclusive categories, or higher order taxa, that farmers used when they could not identify clearly a panicle either because they had limited experience of the characteristics of this variety, because they did not know the appropriate name, or because the panicle presented an ambiguous morphotype regarding their local taxonomy. The use of *Ngirigacha* and *Mbura-imwe* was marginal in the Chuka group, emphasizing the limited diffusion of nomenclature or the lack of shared experience among groups.

Farmers' consistence in using uncommon names

Some variety names which were uncommonly cited (Labeyrie et al. , in prep.) were nonetheless found to be used with some consistency by farmers to label panicles. *Mukumbu* name was notably cited with a noteworthy consistency by informants to label one panicle. This observation somehow contradicts the fact that agreement is linked to shared experience, as farmers had potentially limited opportunity to become familiar with the characteristics of this variety. One explanation may be that its particular and easily distinguishable morphotype, reflected by its motivated name, helps in the process of memorization (Grenand 2002; Boster 1985). Another explanation could be that it might have been largely grown in the past and later abandoned by most farmers.

Informants in the three ethnic groups varied in their consistency in using the name *Kathirigwa*; the Chuka being highly consistent while the Mbeere were clearly inconsistent and the Tharaka intermediate. This case is especially interesting because it traduces that the Mbeere had little opportunity to learn how to identify and name this uncommon variety, likely because of their lack of social relationship with the Chuka, while the Tharaka who have close relationship with this latter group appeared to be more familiar with the identification and naming of this variety even though all the informants were not familiar with it.

Mbunge, *Mucuri* and *Mugana* names were not used very consistently. However, we probably failed to include representative morphotypes of these varieties in the set of panicles used for the identification task. Indeed, part of the long-cycle varieties was not mature at the time of the identification task and it was not possible to wait for them to finish maturing. An identification experiment specifically dedicated to long-cycle landraces would be necessary to reach a better understanding of their taxonomy and the consistence of farmers in naming them.

Agreement patterns reflects the modalities of knowledge transmission

We underlined above the influence of learning modalities on the consistency of farmers in naming varieties through the demonstration that learning from improved varieties by an external source of information enhanced the naming consistency. Furthermore, the intra-group and inter-group agreement patterns give insights concerning the farmer-to-farmer transmission of knowledge and sharing of experience. In addition, our results suggest that the diffusion of information within groups is slow. Indeed, the identification and naming of supposedly exogenous varieties remained incompletely homogenized within groups. For instance, part of the Chuka identified consistently *Mugeta*, but a substantial share of informants in this group was not able to do so. Similar observations were done for *Muruge* and *Mugeta* in the Mbeere group. These observations suggest that modalities of knowledge diffusion may be quite conservative within these groups and limit the diffusion of knowledge (Cavalli-Sforza and Feldman 1981; Reyes-García et al. 2009).

The examination of the path for knowledge transmission hint some ways of explanation. The three ethnic groups studied are patrilinear and virilocal societies, which rules are known to favor the maintenance of cultural differentiation. Indeed, the knowledge transmission from the mother in-law to its daughters-in-law was found to be of major importance in several studies done on this kind of societies, notably for craft-making (Herbich and Dietler 2008). A previous study in closely related groups suggested that young brides pass through a resocialization process during which their mother-in-law completes their education concerning cultivation and household tasks (Linsig pers.com.).This path for knowledge diffusion is conservative, as knowledge diffusion is mainly bounded within the residence group (Cavalli-Sforza and Feldman 1981; Reyes-García et al. 2009; Herbich and Dietler 2008). However, further study would be necessary to assess whether this path of knowledge and experience transmission is really dominant in these societies, and to measure the horizontal (between members of the same age cohort), vertical (parents to children), and oblique (non-parental elders to children) components of knowledge transmission.

Implications for the estimation of crop variety diversity on-farm

It is worth noting that the differences in variety naming consistency among ethnic groups coincided with the differences in the frequency of varieties among the three ethnic groups as described by Labeyrie et al. (submitted). Indeed, *Mugeta* was not inventoried in the Mbeere households during on-farm inventories while it was frequent in the Tharaka and Chuka households. Inventories also showed that *Muruge* was more frequently grown in the Chuka and Tharaka households than in the Mbeere households, and *Ngirigacha* was more frequently grown in the Mbeere area while uncommon in the other ethnic groups. These results illustrate the impact of farmers' consistency in naming varieties on the measure of crop diversity, even at such a small scale. Most studies dealing with the diversity of crop genetic resources on-farm based their estimation or sampling strategies on variety names used by farmers (Jarvis et al. 2008; Deu et al. 2008) because it is supposed to be farmers' management unit, notably for seed selection (Bellon and Brush 1994).

We showed that differences of identification and naming process exist, even between adjacent communities living in a uniform environment. The definition of what is a "community" shall hence be considered carefully in field surveys because spatial proximity between households does not reflect necessary their social proximity. In reality, in most rural areas, the relationships between farmers form a complex network imbedded in different social institutions and governed by rules (Leclerc and Coppens d'Eeckenbrugge 2012). In our case, it is clear that the social organization of farmers, notably at the ethnic level, influences the way in which they identify and name sorghum varieties. This study thus demonstrates the necessity of considering the cultural diversity of farmers for the study and management of crop diversity on-farm because it clearly determines the identification and naming of varieties. Social and cultural anthropology approaches are hence worth receiving further interest from researchers and conservationists for the study and management of crop diversity on-farm.

REFERENCES

Anderson, M.J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26 (1):32-46.

Atran, S., and D. Medin. 2008. *The native mind and the cultural construction of nature*. Cambridge, Massachusetts: MIT press.

Bellon, M.R., and S.B. Brush. 1994. Keepers of maize in Chiapas, Mexico. *Economic Botany* 48 (2):196-209.

Boster, J. S. 1985. Selection for perceptual distinctiveness: evidence from Aguaruna cultivars of Manihot esculenta. *Economic Botany* 39 (3):310-325.

Boster, J.S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88 (2):428-436.

Cavalli-Sforza, L.L., and M.W. Feldman. 1981. Cultural transmission and evolution: a quantitative approach. *Monographs in population biology*. Princeton: Princeton University Press.

D'Andrade, R.G. 1981. The cultural part of cognition. *Cognitive science* 5 (3):179-195.

Deu, M., F. Sagnard, J. Chantereau, C. Calatayud, D. Hérault, C. Mariac, J.L. Pham et al. 2008. Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theoretical and Applied Genetics* 116 (7):903-913.

Diamond, J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* 418 (6898):700-707.

Fadiman, J.A. 1993. *When we began there were witchmen: an oral history from Mount Kenya*. Los Angeles: University of California Press.

Friedberg, C. 1991. Operative aspects of Folk Classification. In *Man and a half: essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer*, ed. A. Pawley, 102-109. Auckland: The Polynesian Society.

Gepts, P. 2004. Crop domestication as a long-term selection experiment. *Plant breeding reviews* 24 (2):1-44.

Glazier, J. 1970. *Ritual and social conflict: circumcision and oath-taking in Mbeere*. Nairobi: Institute of Development Studies, Nairobi University.

Grenand, F. 2002. Stratégies de nomination des plantes cultivées dans une société tupi-guarani, les Wayãpi. *Amerindia* 26/27:209-247.

Heine, B., and W.J.G. Moehlig. 1980. Geographical and historical introduction : language and society. In *Language and dialect atlas of Kenya*. Volume I, eds. B. Heine, and W.J.G. Moehlig. Berlin: Dietrich Reimer Verlag.

Herbich, I., and M. Dietler. 2008. The long arm of the mother-in-law: Learning, postmarital resocialization of women, and material culture style. In *Breaking down boundaries*:

Anthropological approaches to cultural transmission and material culture, eds. M.T. Stark, B.J. Bowser, and L. Horne. Tucson: University of Arizona Press.

Jaetzold, R., H. Schmidt, B. Hornetz, and C. Shisanya. 2007. *Farm management handbook. Vol II, Part C, East Kenya*. Subpart C1, Eastern Province. Nairobi Ministry of Agriculture / GTZ.

Jarvis, Devra, Anthony Brown, Pham Cuong, Luis Collado-Panduro, Luis Latournerie-Moreno, Sanjaya Gyawali, Tesema Tanto et al. 2008. A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. *Proceedings of the National Academy of Sciences of the United States of America* 105 (14):5326-5331.

Labeyrie, V., B. Rono, and C. Leclerc. 2013. How social organization shapes crop diversity: an ecological anthropology approach among Tharaka farmers of Mount Kenya. *Agriculture and Human Values*:1-11.

Labeyrie, V., M. Deu, A. Barnaud, C. Catalayud, M. Buiron, P. Wambugu, S. Manel, J.-C. Glaszmann, and C. Leclerc. 2013. Influence of ethnolinguistic diversity on sorghum evolution in subsistence farming systems in eastern Kenya. *Submitted for publication*.

Labeyrie, V., J.I. Kamau, C. Dubois, C. Catalayud, R. Rivalan , A. Barnaud, and C. Leclerc. 2013. What is a landrace ? Formalizing farmers' taxonomy of sorghum through an emic approach. *In preparation for publication*.

Lambert, H.E. 1956. *Kikuyu social and political institutions*. London: Oxford University Press.

Leclerc, C., and G. Coppens d'Eeckenbrugge. 2012. Social organization of crop genetic diversity. The G x E x S interaction model. *Diversity* 4 (1):1-32.

Middleton, J. 1953. *The central tribes of the north-eastern Bantu*. Ethnographic Survey of Africa. London: International African Institute.

Mwaniki, H.S.K. 1973. *The living history of Embu and Mbeere to 1906*. East African Literature Bureau.

Nuijten, E., and C.J. Almekinders. 2008. Mechanisms explaining variety naming by farmers and name consistency of rice varieties in the Gambia. *Economic Botany* 62 (2):148-160.

Oksanen, J., F.G. Blanchet, R. Kindt, P. Legendre, P.R. Minchin, R. O'Hara, G.L. Simpson, P. Solymos, M. Stevens, and H. Wagner. 2012. Vegan: Community Ecology Package.

Reyes-García, V., J. Broesch, L. Calvet-Mir, N. Fuentes-Peláez, T.W. McDade, S. Parsa, S. Tanner, T. Huanca, W.R Leonard, and M.R. Martínez-Rodríguez. 2009. Cultural transmission of ethnobotanical knowledge and skills: an empirical analysis from an Amerindian society. *Evolution and Human Behavior* 30 (4):274-285.

Romney, A.K., S.C. Weller, and W.H. Batchelder. 1986. Culture as Consensus: A Theory of Culture and Informant Accuracy. *American Anthropologist* 88(2):313-338.

Sadiki, M., D. Jarvis, D. Rijal, J. Bajracharya, N. Hue, T. Camacho-Villa, L. Burgos-May et al. 2007. Variety names: An entry point to crop genetic diversity and distribution in agroecosystems? In *Managing Biodiversity in Agricultural Ecosystems*, eds. D. Jarvis, C. Padoch, and D. Cooper, 34–76. New York: Colombia University Press.

Salick, J., N. Cellinese, and S. Knapp. 1997. Indigenous Diversity of Cassava: Generation, Maintenance, Use and Loss among the Amuesha, Peruvian Upper Amazon. *Economic Botany* 51 (1):6–19.

Sokal, R., and C.D. Michener. 1958. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin 38:1409-1438.

SUPPLEMENTARY MATERIAL

Table IV-S1: Results from the perMANOVA comparing the effect of ethnic groups on farmers' agreement in naming varieties. Df: degrees of freedom, Ssq: sequential sum of squared distance between individuals and their group's centroid, Mean Ssq = Ssq/Df, F.Model: pseudo F ratio, R2: coefficient of determination [Ssq Etnic group / Ssq Total], Pr: p-value (4 000 permutations)

	Df	Ssq	Mean Ssq	F.Model	R^2	Pr(>F)
Ethnic group	2	1.366	0.683	4.7	0.098	2e-04 ***
Residuals	87	12.552	0.144		0.902	
Total	89	13.918			1	

CHAPITRE V- INFLUENCE DE L'ORGANISATION DES SOCIETES SUR LES ECHANGES

FARMER-TO-FARMER SEED EXCHANGES AS A MIRROR OF FARMERS' SOCIAL ORGANIZATION

V. Labeyrie, J. I. Kamau, C. Leclerc

En préparation pour soumission au journal *American Anthropologist*

FARMER-TO-FARMER SEED EXCHANGES AS A MIRROR OF FARMERS' SOCIAL ORGANIZATION

V. Labeyrie, J. I. Kamau, C. Leclerc

En préparation pour soumission au journal *American Anthropologist*

ABSTRACT

The seed-mediated gene flow is a major mechanism of crop evolution on-farm, enhancing the genetic homogenization of their populations. In smallholder farming systems, farmer-to-farmer seed exchanges continue to occupy a major role in seed systems. However, little information concerning the dependence of seed exchanges on farmers' social relationship systems are available. In this study, we depicted the seed system of sorghum in three ethnolinguistic groups on the Mount Kenya region. The patterns of ethnolinguistic and residential endogamy were further compared with those of seed exchanges. Finally, the impact of seed exchange networks on the spatial distribution of varieties and on the diffusion of varietal taxonomy among ethnolinguistic groups was discussed.

The major seed sources of farmers in the area of study were identified through interviews. The farmer-to-farmer seed exchange was then investigated through interviews in 218 households belonging to the three ethnolinguistic groups. Incoming and outgoing seed exchanges were recorded and the ethnolinguistic group, neighborhood-group, gender and age of seed donors and receivers was mentioned as well as their kinship relation. Endogamy rate was computed for 193 households in which the native ethnolinguistic group and neighborhood-group of husband and spouse were recorded.

Seed exchanges were found to be embedded in the social organization of farmers. Indeed, seed exchanges and marriage were mainly bounded within ethnolinguistic groups. In addition, ancestral blood-brotherhood between two groups was witnessed by their important rate of inter-marriage as well as substantial seed exchanges, while limited marriage and seed exchanges existed with the third group. Moreover, a large share of seed exchanges occurred within the neighborhood-group. Kinship, and particularly the relation between women and the family of their husbands, was found to be of central importance for seed exchanges.

INTRODUCTION

Among the major evolutionary processes that shape crop genetic diversity *in situ*, the seed-mediated migration contributes to the homogenization of crop populations. The global patterns of crop genetic diversity observed nowadays are largely due to seed migration which followed that of humans. For instance, the diffusion of Banana to Africa followed human migrations from the New Guinea area, and its further diffusion in Africa is supposed to be linked to the Bantu expansion (Perrier et al. 2011). Sweet potato was probably brought from South-America to Oceania by human travels (Roullier et al. 2013). In Africa, the spatial distribution of the sorghum caudatum botanical race was likely shaped by the migration of Chari-Nil people (Stemler et al. 1975), while the distribution of the kafir race appeared to be linked to that of Bantu populations, the guinea race to that of Niger-Congo, and the durra to Afroasiatic populations (Harlan and Stemler 1976).

At a shorter time-scale, however, seed-mediated migration mainly occurs within short distance through seed exchanges (Hodgkin et al. 2007). Although numerous studies dealt with seed networks in small-scale farming systems, little attention has been paid on their dependence regarding farmers' social relationship. Some rare studies, nevertheless, suggested that farmers' social organization plays a major role in shaping seed exchanges (Leclerc and Coppens d'Eeckenbrugge 2012). On the one hand, the upmost importance of trust in the seed provider (Badstue et al. 2002) favors exchanges with relatives, neighbors, friends, or *compadres* (in South-America), but also with people of the same ethnic group. On the other hand, institutionalized rules of filiation, residence, marriage and inheritance also play a major role in seed diffusion (Longley 2000).

These farmer-to-farmer seed exchanges are also associated to the diffusion of information and knowledge concerning the crop varieties exchanged. Boster (1986) notably showed that kin women exchanged more planting material and also showed a higher agreement in naming cassava cultivars in Aguaruna society from the Amazonia basin. Nuijten and Almekinders (2008) also observed a link between seed exchanges and the consistency of farmers in naming rice varieties in Gambia. The exchanges of seed are thus a potentially good tracer of knowledge diffusion between farmers, and potentially homogenize crop variety identification, naming and classification processes among farmers.

Previous studies on seed networks rarely had the opportunity to assess at the same time their direct impact on the patterns of genetic and phenotypic diversity of crop on-farm, and to address their indirect impact on these diversity patterns through their role in the homogenization of variety naming and classification. On an ethnolinguistic contact zone located East of Mount Kenya, we previously found a relationship between farmers' ethnic organization and sorghum varieties distribution on the one hand, and the spatial patterns of sorghum genetic diversity on the other hand (Labeyrie et al., submitted). Furthermore, we found that the consistency of farmers in naming sorghum varieties regarding their morphological characteristics varied across ethnic groups, suggesting limited knowledge diffusion (Labeyrie et al., In prep.-a).

In this paper, we depict the seed system of sorghum in these three ethnic groups to test the dependence of famer-to-farmer seed exchanges regarding their social relationship. It enables us to further discuss the uneven distribution of varieties across ethnic groups that was observed previously (Labeyrie et al., submitted; Labeyrie et al. 2013). In addition, the present study gives insights concerning the link between the seed exchanges and the diffusion of knowledge concerning varietal taxonomy and contributes to explain the differences observed between the three ethnic groups for this field of knowledge.

MATERIAL AND METHOD

Site of study

The site of study, located on the eastern side of the Mount Kenya, in the Tharaka-Nithi County, covers three major administrative sublocations, Kiaritha, Igamba Ng'ombe and Rianthiga, in which a total of 1255 households were censed in 2009 (National Census 2009). The area of study was 15 km-square and a total of 233 scattered households were surveyed. The study site ranges from 810 to 946 in altitude. At this altitude, climate is semi-arid and displays a bimodal rain distribution with long rains from March to May and short rains from October to December (Camberlin et al. 2009). Farmers thus cultivate crops on two growing seasons, the long rains being poorly predictable (Camberlin et al. 2012). This site was selected for our study because its particular setting enabled us to compare the seed systems of different ethnolinguistic groups while minimizing the spatial distance between them. This field setting enabled us to test the effect of social relationship on seed exchanges while ensuring that the distance did not bias our results.

Farmers claimed to belong to three different ethnolinguistic groups: the Chuka, the Tharaka and the Mbeere, which present historical, linguistic and cultural differences (Middleton 1953; Lambert 1947). The Chuka and Tharaka groups consider themselves as blood-brothers (*Gishiaro*) and believe in their common ascendance (Fadiman 1993). On the contrary, observers reported that the relationship between the Mbeere and the Chuka were tensed in the past (Mwaniki 1973). 15 main neighborhood-groups, or *ntora* in kimeru, were delimited in the area of study. It is a group of usually less than 100 households who share the feeling of unity despite their spatial scattering. The *ntora* is a central unit of social organization for the Bantu populations settled on the eastern side of the Mount Kenya, it was notably described as the smallest level of political organization (Middleton 1953). The *mucii* is the group of people living in the same compound, typically the house head, his wife or wives (polygamy is now uncommon), their unmarried children and the married sons with their families.

SURVEY

Inter-marriage

193 households composed of a man and his wife were surveyed over the three ethnolinguistic groups. In each household, the native ethnolinguistic group and neighborhood-group (*ntora*) of both male and female house heads were recorded. Data concerning the native *ntora* of the wife and the husband were available for 135 couples. The frequency of marriages within and between ethnic groups and *ntora* was calculated.

Origin of seed lots

A first set of interviews aimed at investigating the origin of seed lots that were planted for each growing season. A seed lot is the bulk of seed of a given variety, selected, borrowed or purchased by a farmer for sowing during a given cropping season (Badstue et al. 2002).

Interviews were conducted in 150 households in January 2011 and in 182 households in June 2011. For each variety grown in October 2010 - January 2011 and March - July 2011 growing seasons, farmers were asked whether seed was self-produced in the farm, bought at market, given by NGO or government, or obtained from another farmer. To identify the major sources of seed for the farmers in the area, we calculated the share of the different seed lots origin over all the seed lots sown in October 2010 and March 2011.

Market survey

This survey conducted in the major local market, located in the small town of *Ishiara,* aimed at inventorying the sorghum varieties sold. 156 seed sellers were randomly interviewed over 13 market days between February and April 2011, during the sowing period of the short-rains season (average 12 seed sellers interviewed per market day). The sellers were mainly Mbeere women (70 %), Chuka and Tharaka represented 22 % of the people interviewed (11 % each), and the remaining were mainly of Kikuyu or Kamba origin. Seed sellers were asked to list all the varieties of sorghum they were selling. We calculated the occurrence frequency of each variety over all the market-days surveyed.

Farmer-to-farmer seed exchanges

Further interviews concerning farmer-to-farmer seed exchanges were later conducted in 218 households (43% Chuka, 29% Tharaka, 23% Mbeere), mainly with the female house heads (84%) who are the main stakeholder in seed management (Middleton 1953). During these interviews, we asked to each informant to cite all the people from whom they received and to whom they gave seed. The kinship relation between the informants and the people they cited as seed provider or receiver was recorded, as well as their ethnolinguistic group, neighborhood group and sublocation of residence, age and name. The variety and the quantity of seed exchanged, the reason and the year of the exchange were also recorded.

The frequencies of seed exchanges within (intra) and between (inter) ethnic groups were computed. For the most frequently exchanged varieties, we tested whether their frequency in the intra-group exchanges differed among ethnic groups using a pairwise fisher tests with fdr correction for multiple comparisons. The same test was used to assess the significance of differences of inter-groups seed exchanges, used to test whether the frequency of a given variety differed among the Chuka-Tharaka, the Chuka-Mbeere and the Tharaka-Mbeere exchanges. The share of the seed exchanges that occurred within *ntora* was computed.

To further characterize the paths of seed exchanges, we computed the share of the incoming and outgoing seed exchanges which originated or were addressed to kin, neighbors, friends or unrelated farmers. Further characterization of the kinship relation between the donors and the receivers was done. Moreover, we compared the share of incoming and outgoing exchanges which originated or were addressed to kin from the same generation (g), grandparental generation (g-2), parental generation (g-1), the children generation (g+1), and the grandchildren generation (g+2). For the seed exchanges that occurred between friends and neighbors, we computed a contingency table of seed exchanges between the age classes, defined as following (less than 35 years, 35-47, 48-62, more than 63).

RESULTS

Social organization

Intermarriage

In the majority of the 193 couples censed, men and women shared the same ethnolinguistic identity (67 % of the couples). The Mbeere showed the higher rate of ethnolinguistic endogamy as 83 % of the Mbeere men married Mbeere women. Marriage between *gishiaro* groups (Chuka and Tharaka) was common: 27 % of the Chuka men married Tharaka women, and reversely 24 % of the Tharaka men married Chuka women (Table V-1). On the contrary, inter-marriage between the Mbeere and the Chuka or Tharaka groups was uncommon. By contrast, marriages were exogamous at the *ntora*-scale, the rate of intra-*ntora* marriage being 14 %.

Table V-1. Percentage of men in each ethnic group who married women in the same or in another ethnic group. N: Number of men surveyed in each ethnic group

		WOMEN					
		Chuka	Mbeere	Tharaka	Others	N	Total
MEN	Chuka	58%	6%	27%	9%	79	100%
	Mbeere	4%	83%	10%	2%	48	100%
	Tharaka	24%	5%	65%	6%	66	100%

Sources of farmers' seed lots

A large share of the seed lots were self-produced by farmers (39 - 41 %) or purchased at the market (34 - 29 %) in October 2010 and March 2011. Only 10 % of the seed lots were received from other farmers (Table V-2). NGO, government and industry constituted about one-fifth of the sources of seed. No differences were observed among ethnic groups concerning the share represented by the various seed sources (data not shown), and few differences were observed between the two cropping seasons.

Table V-2. Origin of the seed lots declared during January inventory (150 households, 254 seed lots) and July inventory (182 household, 320 seed lots)

	January		June	
Origine of seed lots	No.	Percentage	No.	Percentage
Own	99	39%	131	41%
Market	86	34%	94	29%
Other farmers	25	10%	31	10%
Governement	28	11%	33	10%
Brewery	4	2%	-	-
Church	3	1%	-	-
NGO	9	4%	31	10%
Total	254	100%	320	100%

Varieties given by the NGOs or by the government were mostly *Gadam*, disseminated since 2009, *Kaguru* disseminated in 1994, and again in few households in 2010. *Serendo* was disseminated since 1996 and up to 2011. According to the January survey, the share of seed lots purchased at the market was similar for the improved varieties and for the local ones, but the local varieties were more self produced (56 % of the seed lots) than the improved ones (20 % of the seed lots). The most common varieties sold at the market were *Kaguru*, followed by *Ngirigacha*, *Muruge* and *Gadam. Mbura ciiri*, *Muceru*, *Muruge mbura ciiri*, *Serendo*, *Gatinga* and *Mugeta* varieties were uncommon, as well as several other varieties which were observed only once (Figure V-1).

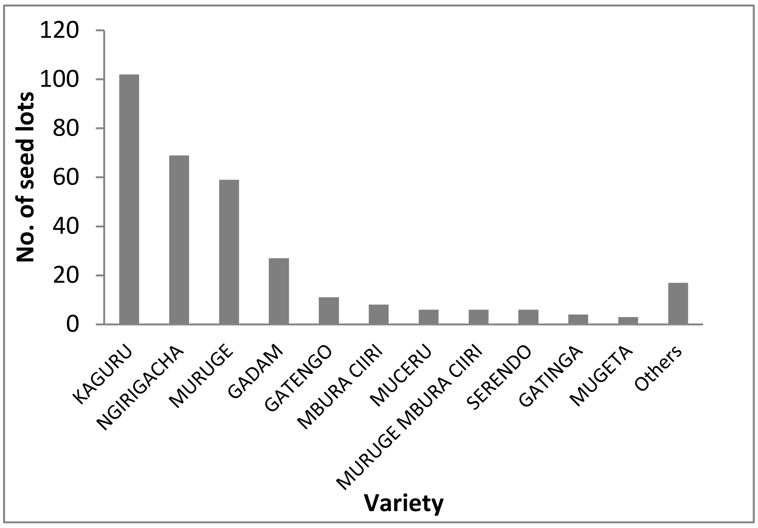


Figure V-1. Number of seed lots inventoried for each sorghum variety over 13 market days

Farmer-to-farmer seed exchanges

During the interviews, farmers stated that the large majority of farmer-to-farmer seed exchanges were due to lack of seed to sow (93%). Only 5% of the exchanges were motivated by the desire to test a new variety. The large majority of exchanges reported by women were concluded with other women.

Seed exchanges between ethnic groups

The majority of seed exchanges occurred between farmers belonging to the same ethnic group (63 %). The portion of within-group seed exchanges was the highest in the Mbeere group

(Outgoing: 77 %; incoming: 81 %; Table V-3.A and B). Seed exchanges between the Chuka and the Tharaka were noteworthy as they represented 29 to 38 % of the incoming and outgoing seed exchanges in both groups. On the contrary, the Chuka and the Tharaka rarely exchanged seeds with the Mbeere (Table V-3).

Table V-3. Fraction of the total seed exchanges outgoing (A) or incoming (B) from/to each ethnic group to/from farmers from the same (bold) or from other ethnic groups. Chuka: CH, Mbeere: MB, Tharaka: TH, Others: Oth.

A.

		To :				
		CH	MB	TH	Oth.	Total % outgoing
From : (% outgoing)	CH	0.64	0.04	0.29	0.03	1.00
	MB	0.12	0.77	0.09	0.02	1.00
	TH	0.38	0.06	0.54	0.02	1.00
	Oth.	0.72	0.06	0.22	0	1.00

B.

		To : (% incoming)			
		CH	MB	TH	Oth.
From:	CH	0.62	0.09	0.37	0.55
	MB	0.06	0.81	0.06	0.17
	TH	0.30	0.10	0.57	0.28
	Oth.	0.02	0.00	0.01	-
	Total % incoming	1.00	1.00	1.00	1.00

The principal varieties concerned by the farmer-to-farmer exchanges were *Kaguru* (27 % of the total number of exchanges), *Muruge* (21%) and *Gadam* (15%). *Ngirigacha*, *Mbura-imwe* and *Mugeta* represented equivalent fractions (8-9 %). The frequency of exchange of the different varieties varied considerably between ethnic groups (Figure V-2). *Mugeta* was mainly exchanged within (in blue, Figure V-2.A) and between the Chuka and Tharaka groups (Figure V-2.B). No exchange of *Mugeta* was recorded within the Mbeere group, and exchanges of this variety were very rare with the two other groups. Most of the overall exchanges of *Mugeta* took place within and between the Tharaka and Chuka groups (93%).

Mugeta represented 14 % of the intra-group exchanges for the Tharaka, 8% for the Chuka, and only 1 % for the Mbeere (Figure V-2.A) and all the pairwise comparisons between ethnic groups were significant (Fisher test: p-value < 0.05; Table V-4).

The exchanges of *Mbura-imwe* also occurred mainly within and between Chuka and Tharaka. *Mbura-imwe* accounted for 11 % of the exchanges within the Tharaka group, 8 % within the Chuka and 3 % within the Mbeere group (in green, Figure V-2.A) and pairwise differences between the Mbeere and the Chuka or Tharaka were significant (Table V-4). Reversely, the exchanges of *Ngirigacha* mostly took place within the Mbeere group (in brown, Figure IV-2.A) and were uncommon among ethnic groups (Figure IV-2.B). *Ngirigacha* accounted for 28 % of the seed exchanges within the Mbeere group, which was significantly more than for the Tharaka (5%) and the Chuka (2%). The varieties exchanged among ethnic groups did not differ significantly, except for *Ngirigacha* which was more exchanged between the Mbeere and the Tharaka than between this latter group and the Chuka (in brown, Figure VI-2.B).

The frequency of exchanges of *Kaguru*, *Gadam* and *Muruge* within groups did not differ significantly (Table V-4). The varieties *Mugana*, *Mucuri* and *Mukumbu* (not represented on the graphs) were not commonly exchanged (less than 15 e xchanges) and they were exclusively exchanged within and between the Chuka and the Tharaka groups, but never with the Mbeere.

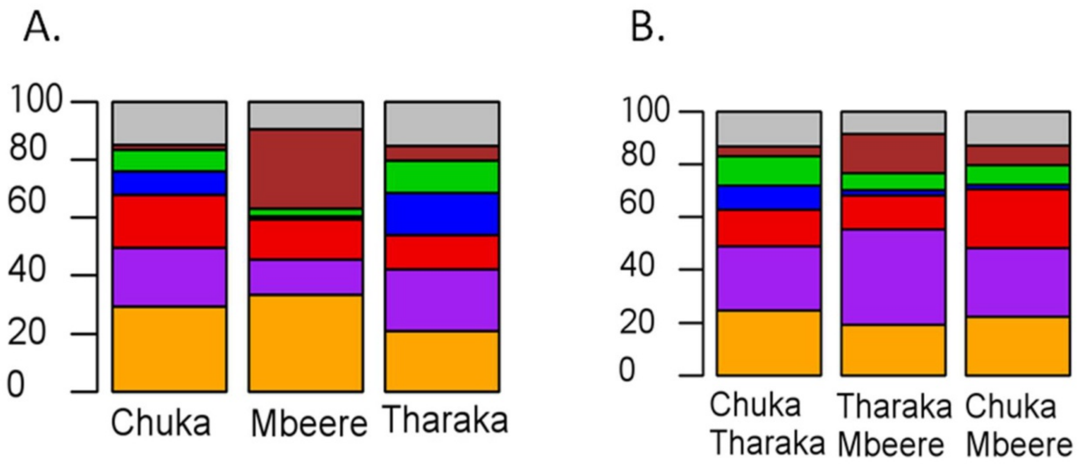


Figure V-2. Share of the main varieties in the exchanges within (A.) and between (B.) ethnic groups. Y-axis: number of intra-group exchanges. Colors represent varieties: Kaguru: orange, Gadam: red, Muruge: purple, Mugeta: Blue, Mbura-imwe: green, Ngirigacha: brown, others: grey.

	Within-groups			Between groups		
	CH-MB	TH-MB	CH-TH	TH/CH TH/MB	TH/CH CH/MB	TH/MB CH/MB
Kaguru	0.465	0.084	0.136	0.870	0.870	0.870
Gadam	0.430	0.570	0.250	1.000	0.660	0.660
Muruge	0.053	0.053	0.839	0.720	0.870	0.810
Mugeta	0.001***	<0.001*	0.029*	0.230	0.230	1.000
Mbura-	0.030*	0.004**	0.239	0.950	0.950	1.000
Ngirigacha	<0.001***	<0.001***	0.049*	0.029*	0.349	0.349

Table V-4. Significance of the differences of exchange frequencies for each variety within-group (left) and between-group (right). P-value of the pairwise Fisher tests corrected for multiple comparisons.

Seed exchanges between neighborhood-groups

More than half of the seed exchanges (53 %) occurred between farmers belonging to the same neighborhood-group. This percentage was the highest for *Nduti* neighborhood-group (outgoing: 57 %, incoming: 61 %) and *Makomora* (outgoing: 62%, incoming: 58 %) where exchanges were especially intense with farmers from the same neighborhood-group. *Kabururu* (outgoing: 53%, incoming: 62 %) and *Mbaraga* (outgoing: 57%, incoming: 43 %) also showed similar trends.

Kinship and seed exchanges

The large majority of farmer-to-farmer seed exchanges occurred among kin (outgoing: 74 %, incoming: 79 %). A large part of the exchanges arise between women and the family of their husband (outgoing: 33 %, incoming: 35 %, Table V-5). On the one hand, sisters-in-law (wives of the husband's brothers) were the second recipient of seed gifts (12%) after the friends. On the other hand, the mother-in-law (14%), the sisters-in-laws (5%) and husbands' sisters (5%) were the main seed providers in the husband's family.

A substantial share of the outgoing seed exchanges were addressed to friends (15%), parents (13%), sons' and their family including their in-laws (11%), and neighbors (9%). The women own maternal (15%) and paternal (8%) families were also a non-negligible source of seed, as well as friends (12%) and neighbors (7%).

Table V-5. Kinship relation of farmers to which (outgoing) and from which (incoming) the female informants declared to give and receive seeds. No.: Number of exchanges, Perc.: Percentage of the total outgoing or incoming exchanges.

	Outgoing (%)		Incoming (%)	
Relationship	No.	Perc.	No.	Perc.
Husband's familly	177	33%	139	35%
Friends	80	15%	47	12%
Wife's familly	71	13%	94	23%
Sons, in-laws and grandchildren	59	11%	21	5%
Neighbors	49	9%	29	7%
Sisters, in-laws and nephews	36	7%	21	5%
Brothers, in-laws and nephews	30	6%	19	5%
Daughters, in-laws and grandchildren	27	5%	21	5%
Others	8	1%	6	2%
Total	537	100%	397	100%

The majority of seed exchanges with the kin occurred between farmers from the same generation, or were addressed to the children's generation (g+1; Table V-6). Exchanges with the parental generation were less frequent than the reverse. Exchanges between grandchildren and farmers belonging to their grandparents' generation were very uncommon. The proportions of outgoing and incoming seed exchanges between generations are symmetrical: the parental generation gives seed to children, and children receive from the parental generation.

Table V-6. Number and percentage of outgoing and incoming seed exchanges reported by women farmers to/from kin according to their generation: g: same generation as the informant, g+1: parental, g+2: grandparental, g-1: children, g-2: grandchildren. No.: Number of outgoing or incoming exchanges.

	Outgoing		Incoming	
Generation	No.	Perc.	No.	Perc.
g	131	33%	114	36%
g+1	176	44%	43	14%
g+2	5	1%	3	1%
g-1	81	20%	143	45%
g-2	6	2%	12	4%
N	399	100%	315	100%

DISCUSSION

Farmer-to-farmer seed exchanges reflect the social organization

This study shows that seed exchanges are linked to the social organization of farmers. Indeed, seed exchanges and marriage occur mainly within ethnic groups. In addition, ancestral blood-brotherhood (*gishiaro*) between ethnic groups is witnessed by the large share of inter-marriage between the Chuka and the Tharaka, confirming observations of Fadiman (Fadiman 1993), and clearly favors seed exchanges between these groups. The limited relationship between these latter groups and the Mbeere is traduced by the scarcity of marriage between them, and limit their exchanges of seed.

Our results confirmed residential exogamy at the *ntora* (neighborhood-group) scale as most women were native from a different *ntora* group from that of their husband. Nevertheless, a large share of seed exchange occurred within *ntora*. The high rate of within-*ntora* exchanges is partially explained by the intensity of exchanges between women and their husband's family members, mainly their mother-in-law and sisters-in-law. During interviews, informants always emphasized the importance of trust in seed exchanges, which is congruent with the observations of Badstue in Mexico (Badstue et al. 2002). According to farmers, asking for seed is a delicate matter, first because farmers feel shameful to ask and fear rejection, but also because the seed may be “cursed” or simply of bad quality. Getting seeds from relatives is thus a guarantee. In addition, part of our female informants reported that asking for seed from their own family was regarded as shameful because it means that the family of the husband was not able to fulfill its needs. It would be only after several years of marriage that women could get seed from their own family. Similar observations were done in an adjacent group, the Mwimbi (Linsig, pers.com.). Such preferential interaction with the in-laws appears to be a common pattern in patrilinear and virilocal societies (Herbich and Dietler 2008).

Seed exchanges explain the patterns of diversity

The fact that seed exchanges are mainly within intra-ethnic groups and intra-neighborhood groups explains the uneven distribution of varieties observed across ethnic groups, and across *ntora*. Indeed, the exchange frequencies of the different varieties among ethnic groups were congruent with their distribution patterns depicted previously in the area (Labeyrie et al., submitted). *Mugeta* and *Mbura imwe* were almost absent from the Mbeere seed network, and they were not inventoried in the Mbeere farms previously. Reversely, the *Ngirigacha* variety which was frequent in the Mbeere area also represented a large share of the Mbeere seed network. The long-cycle varieties *Mugana*, *Mucuri* and *Mukumbu* were not inventoried previously in the Mbeere farms, and their exchanges never involved the Mbeere. The correspondence observed previously between the spatial patterns of distribution of sorghum varieties and those of ethnic groups is thus explained by the limited exchanges of seed between groups despite their spatial proximity and their common reliance on the same market as a main seed source.

The survey on the origin of seed lots emphasized the importance of the local market. However, only a small set of sorghum varieties are frequently sold on the market as compared to the diversity of varieties inventoried on-farm. The widespread recourse to the market hence appeared to concern only few varieties, mainly *Kaguru*, *Ngirigacha*, *Muruge* and *Gadam* while other varieties were mainly maintained through farmer-to-farmer exchanges. *Mugeta* in the Tharaka group and *Mugana*, *Mucuri* and *Mukumbu* in the Chuka group notably appeared to be maintained through this network.

Observations done previously in another site located in the Tharaka lowlands showed that the assemblages of crop species and sorghum varieties cultivated within farms differed significantly among *ntora* (Labeyrie et al. 2013). The importance of intra-*ntora* exchanges probably explains these patterns. Even though they had the opportunity to get seed of any variety, notably through the market, farmers appeared to imitate their neighbors. Moreover, these lowlands Tharaka farmers appeared to rarely buy seed at the market, suggesting a predominance of farmer-to-farmer seed exchanges.

The link between seed exchanges and knowledge diffusion

The hypothesis of limited diffusion of knowledge from the Chuka and Tharaka groups to the Mbeere is comforted by the scarcity of seed exchanges between them. The limited communication between these groups explains why they maintain different denominations for sorghum varieties. Notably, the Mbeere extensively used the inclusive category *Ngirigacha* where the majority of the Chuka and Tharaka distinguished and named different varieties (Labeyrie et al., In prep.-b). The *Ngirigacha* variety name was also widely used by the Mbeere seed sellers in the market but the Chuka did not use this denomination, suggesting that little exchange of information occurs within the market among Mbeere and Chuka sellers. The social organization that we observed on-farm, with limited knowledge exchanges between Mbeere and Chuka, is thus reproduced into the markets.

It is noteworthy that the Tharaka group, in which the exchanges of *Mugeta* variety were more intense, showed also a higher consistency in naming this variety during the identification task (Labeyrie et al., In prep.-a). This corroborates the hypothesis of Boster (1986), and later Nuijten and Almekinders (2008) according to which farmers' consistency is linked to seed exchanges. Nevertheless, the path for seed and knowledge exchanges appeared to differ in some extent as rarely exchanged varieties such as *Mukumbu* or *Kathirigwa* were still named consistently.

The seed exchanges network depicted here reflected a preferential relationship between members of the same *mucii* (household) and neighborhood-group, which probably also facilitated the diffusion of knowledge. Knowledge transmission through the affinal link is a common pattern in patrilinear and virilocal societies, where the interaction between the bride and its in-laws favors her "resocialization" in her new social group (Herbich and Dietler 2008). On an adjacent ethnolinguistic group, the Mwimbi, observations suggest that the bride passes through a learning process of several months before the mother-in-law allows her daughter-in-law to manage her own garden, granary and kitchen (Linsig, pers.com.). In addition, we observed that many recently-married girls had a limited knowledge concerning sorghum varieties and their naming, suggesting that they may complete their learning after they get married , as observed by one of us (VL) However, the fact that seed exchanges mainly take place within the *mucii* may be a recent evolution of the societies studied. It seems that settlement of married sons in their father's compound, which is the most frequent situation nowadays, may be a recent evolution due to land pressure. Observations suggest that in the past, families displayed a strategy aiming at maximizing the range of agro-ecological conditions covered by their lands (Bernard 1972). House heads thus commonly installed their sons in the different plot of land they possessed.

Conclusion

This study showed that farmer-to-farmer seed exchanges are clearly influenced by their social relationships, emphasizing the influence of ethnic partition, neighborhood and kinship in the flow of seed. Furthermore, their common reliance on the same local market does not appear to homogenize their varietal portfolio and the way they identify, name and classify varieties. This small-scale case study depicted mechanisms which were probably involved in the shaping of crop morphological and genetic diversity at larger time and spatial scale. Indeed, it shows that gene flows between crop populations cultivated by different cultural groups are restricted despite their spatial proximity. Such isolation from gene flows, combined to the divergence of farmers' selection practices in the absence of knowledge diffusion may thus be a major driver in crop evolution (Leclerc and Coppens d'Eeckenbrugge 2012). Such mechanisms may have contributed to shape the crop phenotypic and genetic diversity at larger scale, as it was for instance suggested by Harlan and Stemler (1976) to explain the distribution of sorghum diversity in Africa.

REFERENCES

Badstue, L.B., M.R. Bellon, X. Juárez, I. Manuel, and A.M. Solano. 2002. Social Relations and Seed Transactions among Small-Scale Maize Farmers in the Central Valleys of Oaxaca, Mexico.*CIMMYT Economics Working Paper* No. 02–02.

Bernard, F.E. 1972. *East of Mount Kenya: agriculture in transition*. Afrika-Studien. München: Weltforum verlag.

Boster, J.S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88 (2):428-436.

Camberlin, P., J. Boyard-Micheau, N. Philippon, C. Baron, C. Leclerc, and C. Mwongera. 2012. Climatic gradients along the windward slopes of Mount Kenya and their implication for crop risks. Part 1: climate variability. *International Journal of Climatology* (online).

Camberlin, P., V. Moron, R. Okoola, N. Philippon, and W. Gitau. 2009. Components of rainy seasons' variability in Equatorial East Africa: onset, cessation, rainfall frequency and intensity. *Theoretical and Applied Climatology* 98 (3):237-249.

Fadiman, J.A. 1993. *When we began there were witchmen: an oral history from Mount Kenya*. Los Angeles: University of California Press.

Harlan, J.R., and A. Stemler. 1976. The races of sorghum in Africa. In *Origins of African plant domestication*, ed. J. R. Harlan, 465-478. The Hague: Mouton.

Herbich, I., and M. Dietler. 2008. The long arm of the mother-in-law: Learning, postmarital resocialization of women, and material culture style. In *Breaking down boundaries: Anthropological approaches to cultural transmission and material culture*, eds. M.T. Stark, B.J. Bowser, and L. Horne. Tucson: University of Arizona Press.

Hodgkin, T., R. Rana, J. Tuxill, D. Balma, A. Subedi, I. Mar, R. Karamura et al. 2007. Seed Systems and Crop Genetic Diversity in Agroecosystems. In *Managing biodiversity in agricultural ecosystems*, eds. D. Jarvis, C. Padoch, and D. Cooper, 77-116. New York: Colombia University Press.

Labeyrie, V., M. Deu, A. Barnaud, C. Catalayud, M. Buiron, P. Wambugu, S. Manel, J.-C. Glaszmann, and C. Leclerc. 2013. Influence of ethnolinguistic diversity on sorghum evolution in subsistence farming systems in eastern Kenya. *Submitted for publication.*

Labeyrie, V., J.I. Kamau, and C. Leclerc. In prep.-a. Farmers consistency in naming sorghum varieties: the share of culture. *In preparation for publication.*

Labeyrie, V., J.I. Kamau, C. Dubois, C. Catalayud, R. Rivalan , A. Barnaud, and C. Leclerc. In prep.-b. What is a landrace ? Formalizing farmers' taxonomy of sorghum through an emic approach. *In preparation for publication.*

Labeyrie, V., B. Rono, and C. Leclerc. 2013. How social organization shapes crop diversity: an ecological anthropology approach among Tharaka farmers of Mount Kenya. *Agriculture and Human Values*:1-11.

Lambert, H. 1947. *The use of indigenous authorities in tribal administration: studies of the Meru in Kenya Colony*. School of African Studies, University of Cape Town.

Leclerc, C., and G. Coppens d'Eeckenbrugge. 2012. Social organization of crop genetic diversity. The G x E x S interaction model. *Diversity* 4 (1):1-32.

Longley, C. 2000. *A social life of seeds: Local management of crop variability in North-Western Sierra Leone*. University of London, London, UK.

Middleton, J. 1953. *The central tribes of the north-eastern Ban*tu. Ethnographic Survey of Africa. London: International African Institute.

Mwaniki, H.S.K. 1973. *The living history of Embu and Mbeere to 1906*. East African Literature Bureau.

National Census, 2009. Kenya National Bureau of Statistics census. https://www.opendata.go.ke/vision. Accessed October 2013.

Nuijten, E., and C.J. Almekinders. 2008. Mechanisms explaining variety naming by farmers and name consistency of rice varieties in the Gambia. *Economic Botany* 62 (2):148-160.

Perrier, X., E. De Langhe, M. Donohue, C. Lentfer, L. Vrydaghs, F. Bakry, F. Carreel et al. 2011. Multidisciplinary perspectives on banana (Musa spp.) domestication. *Proceedings of the National Academy of Sciences of the United States of America* 108 (28):11311-11318.

Roullier, C., L. Benoit, D.B. McKey, and V. Lebot. 2013. Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. *Proceedings of the National Academy of Sciences of the United States of America* 110 (6):2205-2210.

Stemler, A.B., J. R. Harlan, and J.M. de Wet. 1975. Caudatum sorghums and speakers of Chari-Nile languages in Africa. *Journal of African History* 16 (2):161-183.

DISCUSSION GENERALE

Cette thèse a mis en évidence une relation entre l'organisation des sociétés Chuka, Tharaka et Mbeere et l'organisation de la diversité du sorgho *in situ*. Cette relation se manifeste d'une part sur la distribution de la diversité telle qu'elle est perçue par les agriculteurs (les variétés nommées), et d'autre part sur celle de la diversité telle que nous l'avons caractérisée à l'aide de marqueurs génétiques et phénotypiques. Cependant, ces deux approches de la diversité ne coïncident pas parfaitement. Différentes variétés locales à cycle court sont notamment distinguées par les agriculteurs alors qu'elles ne correspondent pas à des unités génétiques distinctes et homogènes.

Cette relation s'explique par le fait que l'organisation des sociétés structure d'une part la diffusion des systèmes locaux de taxonomie et de nomenclature variétale, et d'autre part les échanges de semences. Les agriculteurs appartenant à un même groupe ethnique identifient, nomment et classent globalement de la même manière les variétés de sorgho. Des différences de taxonomie ont en revanche été mises en évidence entre groupes. Certains groupes ethniques affichent notamment un plus fort consensus que d'autres pour l'identification des variétés de sorgho. De plus, nous avons montré qu'il existe une relation entre l'organisation des échanges de semences et l'organisation des sociétés. En effet, les mariages et les échanges de semences ont lieu principalement au sein des groupes ethniques ou entre groupes alliés (*gishiaro*).

Afin de clarifier la relation entre la distribution des variétés nommées, celle de la variabilité biologique, et l'organisation des sociétés, nous analyserons d'abord l'impact de l'organisation des sociétés sur les échanges de semences et de savoir, et donc sur la diffusion des systèmes de taxonomie et de nomenclature locaux. Nous caractériserons ensuite l'effet de ces différences de représentations (traduites par les taxonomies) sur les pratiques de sélection des semences. Enfin, nous discuterons de l'effet combiné des pratiques de sélection et des échanges de semences sur l'organisation de la diversité génétique et phénotypique du sorgho. Les applications de ces résultats pour la caractérisation, la conservation et l'amélioration des ressources génétiques seront ensuite abordées.

I- De l'organisation des sociétés aux représentations qu'elles se font de la diversité

Cette thèse montre qu'il existe une cohérence globale entre les relations sociales qu'entretiennent les agriculteurs, leurs pratiques d'échanges de semences, et la façon dont ils perçoivent, se représentent et nomment la variabilité phénotypique de leurs plantes. En effet, lorsqu'une variété est plus fréquente dans un groupe ethnique que dans un autre (Chapitre II), elle est également identifiée de façon plus consensuelle (Chapitre IV) et elle est davantage échangée (Chapitre V) par les agriculteurs de ce groupe.

Dans cette première partie de discussion, nous considérerons d'abord l'influence de l'organisation des sociétés sur les échanges de semences et de savoir, puis nous examinerons l'influence des échanges sur la diffusion des systèmes de taxonomie et de nomenclature locaux.

Des relations sociales aux échanges

L'organisation des sociétés se reflète de façon frappante dans l'organisation des échanges de semences comme nous l'avons vu dans le chapitre V. La majeure partie des échanges a lieu à l'intérieur des groupes ethniques, ou entre les Chuka et les Tharaka qui sont *gishiaro*. En revanche, les échanges entre les Mbeere et les deux autres groupes sont rares. Cette structuration des échanges de semences correspond à celle observée concernant les mariages. Les descriptions ethnographiques disponibles dans la littérature ainsi que le discours de nos informateurs corroborent nos observations à propos des relations entre groupes ethnolinguistiques. Les Chuka et les Tharaka se considèrent comme « frères de sang », ou *gishiaro* tandis que les Chuka et les Mbeere entretiennent historiquement des relations difficiles qui, d'après nos observations sur le terrain, sont aujourd'hui exacerbées par la pression foncière.

La relation entre les échanges de semences et l'organisation des sociétés a déjà été mise en évidence dans des études menées notamment sur le maïs en Mesoamérique (Badstue et al. 2002), sur les tubercules dans les Andes (Zimmerer 2003), ou encore sur le sorgho en Ethiopie (McGuire 2008), (voir revue dans Leclerc and Coppens d'Eeckenbrugge 2012). Cette relation s'explique par la nécessaire confiance entre le donneur et le receveur de semences, comme l'a souligné Badstue (2007). Réclamer des semences est souvent vu comme une chose honteuse par les agriculteurs. Ils craignent de se voir refuser leur requête ou encore que la quantité de semences soit trop réduite et qu'elles soient de mauvaise qualité. Le receveur se retrouve alors dans une situation de dépendance vis-à-vis du donneur. Il est donc plus aisé de demander des semences à des personnes avec qui l'on est proche. Cela assure également une certaine garantie concernant la qualité des semences. Par exemple, certains de nos informateurs nous ont rapporté se procurer des semences uniquement auprès de leurs proches de crainte de se voir procurer des semences maudites. Le semi de telles semences, auxquelles une personne mal intentionnée aurait jeté un sort, risquerait de provoquer la perte de l'ensemble de la récolte. Il est donc compréhensible que les échanges ne se fassent pas avec n'importe qui dans ces conditions.

L'importance des échanges entre affins, c'est-à-dire essentiellement entre belles-mères et belles-filles ainsi qu'entre belles-sœurs, est en cohérence avec le discours des agricultrices interrogées sur notre zone d'étude ainsi qu'avec les observations faites chez les Mwimbi, un groupe ethnolinguistique adjacent des Chuka et des Tharaka (Linsig 2009). La relation privilégiée entre les femmes et la famille de leur époux est une caractéristique répandue dans les sociétés patrilinéaires et virilocales (Herbich and Dietler 2008). Les échanges de semences, et donc potentiellement les échanges de savoir, semblent se faire suivant deux principales modalités dans les sociétés étudiées, suivant la typologie introduite par Cavalli-Sforza and Feldman (Reyes-García et al. 2009; Cavalli-Sforza and Feldman 1981) : i. Entre affins, ce qui dans une société virilocale a les mêmes conséquences qu'une transmission verticale de parents à enfants, ii. Horizontalement, entre personnes de la même cohorte d'âge. De plus, les échanges horizontaux sont largement circonscrits au groupe de voisinage et au groupe ethnolinguistique. Ces modalités de diffusion semblent être assez conservatives, c'est-à-dire que la diffusion du savoir mais aussi des variétés par ce réseau est relativement lente. En effet, l'identification et la nomination des variétés exogènes (*Gadam*, *Mugeta* chez les Chuka et les Mbeere, *Muruge* chez les Mbeere et les Tharaka) ne sont pas encore homogénéisées entre groupes malgré leur proximité spatiale.

Des échanges aux représentations

L'organisation de ces sociétés se reflète également dans la façon dont elles identifient, nomment et classent les variétés de sorgho, notamment pour les variétés considérées comme locales, comme cela a été montré dans le chapitre IV. Nos résultats suggèrent que la diffusion de la nomenclature et le partage de l'expérience concernant l'identification des variétés est limitée entre groupes, particulièrement entre les Mbeere et les groupes Chuka et Tharaka. En effet, les agriculteurs Chuka montrent un consensus assez élevé concernant l'identification des variétés locales tandis que les Mbeere sont clairement moins consensuels. Par ailleurs, le consensus entre agriculteurs Chuka et Tharaka est plus élevé qu'entre les agriculteurs de ces deux groupes et les Mbeere. Le savoir n'est donc pas uniformément distribué entre les trois groupes et reflète leurs relations ainsi que l'intensité des échanges de semences entre eux.

La relation que nous avons mise en évidence entre le partage d'une taxonomie et d'une nomenclature variétale commune et les échanges de semences rejoint les observations faites par Boster chez les Aguaruna d'Amazonie (Boster 1986). Il montre en effet que les échanges de savoir et de semences sont corrélés et dépendants des relations sociales entre agricultrices, notamment des relations de parenté. Nuijten et Almekinders (2008) suggèrent également que le manque de consensus concernant l'identification des variétés de riz entre villages Gambiens est attribuable à la rareté de leurs échanges. Cependant, cette dernière étude n'a pas examiné le lien entre les échanges et l'organisation sociale des agriculteurs, qui pourrait expliquer l'absence de consensus concernant la dénomination des variétés au sein de certains villages. Les échanges de semences permettent de renforcer l'expérience commune des agriculteurs concernant la perception des caractéristiques phénotypiques des variétés, leur assignement à des classes, et leur dénomination (Boster 1986). Le lien entre les échanges et le partage d'une taxonomie et d'une nomenclature commune s'explique donc aisément, celles-ci constituant un cadre de référence pour stocker et échanger l'expérience et les informations[16] (Bulmer 1974 in Friedberg 1991).

Le marché : découplage entre échanges de semences et de savoir

Le marché a une importance considérable pour les agriculteurs de la zone de contact ethnolinguistique (site 1). Tous l'ont cité comme leur source privilégiée de semences pour plusieurs raisons. Les marchés ne semblaient pas exister chez les groupes Meru par le passé, comme le rapporte Bernard (1972), bien que des échanges commerciaux existaient entre

---

[16] "a frame for storing and conveying experience and information"

groupes sous d’autres formes. Il s’agit donc d’une pratique relativement récente, initiée sous l’aire coloniale. Nos informateurs considèrent que le recours fréquent au marché pour l’achat de semences est une conséquence de l’évolution de leurs sociétés vers d’avantage d’individualisme. Tout d’abord, l’achat de semences affranchit les agriculteurs de la dépendance vis-à-vis des donneurs, ils peuvent obtenir les quantités qu’ils souhaitent et choisir les variétés qu’ils désirent semer. Des raisons similaires sont rapportées par David et Sperling (1999) pour expliquer la recrudescence du recours au marché par les agriculteurs d’Ouganda, du Rwanda, du Burundi et du Congo.

Une raison supplémentaire citée par de nombreux agriculteurs pour justifier leur recours au marché est originale et tout à fait spécifique à l’organisation de ces sociétés. Les agriculteurs Chuka et Tharaka rapportent préférer acheter des semences au marché car ils n’ont pas à se préoccuper de l’identité clanique de leur fournisseur de semence. En effet, la grande majorité des vendeuses de semences au marché local sont d’origine Mbeere. Or, les clans Mbeere n’ont pas de relations d’alliance (*gishiaro*) avec les clans Chuka et Tharaka, relations qui prohibent tout échange de semences. Au regard du flou existant concernant la dénomination des clans, il est possible que les jeunes générations maitrisent moins bien que les anciennes les complexes relations d’alliances qui lient les clans entre eux. Dans le doute, ils semblent préférer avoir recours au marché. N. Linsig (2009) décrit une situation similaire chez un groupe voisin, les Mwimbi, qui évitent de se rendre à un marché où la majorité des vendeuses appartiennent à un clan interdit.

Chez les Tharaka situés à plus basse altitude, éloignés de la zone de contact ethnique (site 2), le recours au marché est nettement moins marqué que sur la zone de contact ethnolinguistique (site 1). Ils sont pourtant situés dans une zone agro-écologique moins favorable et seraient potentiellement plus souvent touchés par des pertes de récolte nécessitant de renouveler leur stock de semences. Les raisons de ces différences sont inconnues, une explication partielle pourrait venir du fait que les Tharaka sont réputés pour être le groupe le plus traditionnaliste, et le moins affecté par les changements sociétaux. Leur société est implantée depuis longtemps dans les zones de plaines tandis que la zone de contact ethnolinguistique est de peuplement assez récent, ce qui a probablement bouleversé quelque peu l’organisation sociale.

Le marché induit la diffusion uniforme d'une partie des variétés dans les trois groupes ethniques, en revanche il ne semble pas contribuer à l'uniformisation du savoir relatif à l'identification et à la nomination des variétés. Seul un nombre restreint de dénominations ont été relevées auprès des vendeuses en comparaison à celles recensées sur la zone d'étude, en particulier chez les Chuka et les Tharaka. Le nom *Ngirigacha* est par exemple largement utilisé par les vendeuses Mbeere alors qu'il est anecdotique chez les Chuka. Ces résultats sont importants car ils vont à l'encontre des idées reçues selon lesquelles le recours au marché entraine forcément une uniformisation des variétés et des savoirs. Le recours à cette nouvelle pratique se fait en continuité avec les pratiques préexistantes, et non en rupture.

II- De l'organisation des sociétés à leurs pratiques : quel impact sur la diversité des plantes ?

Nous avons montré précédemment que l'organisation des sociétés contribue à déterminer la représentation que les agriculteurs se font de la diversité des plantes. Il est donc nécessaire d'aborder maintenant la question de la relation entre représentations et pratiques. Une variété est-elle maintenue par sélection uniquement si elle est identifiée et distinguée des autres par les agriculteurs? Les différences de représentations que se font les trois groupes ethniques de la diversité phénotypique du sorgho ont-elles un impact sur celle-ci ?

D'autre part, nous avons montré que les échanges d'agriculteurs à agriculteurs sont fortement dépendants de leurs relations sociales, mais que le recours au marché est fréquent. Quelles sont les conséquences de ces pratiques d'échanges de semences sur la distribution de la diversité génétique et phénotypique du sorgho?

Dans cette seconde partie, nous discuterons de l'influence des différences de représentations, traduites par les différences de taxonomies, sur les pratiques de sélection des semences. Nous aborderons ensuite l'effet des pratiques de sélection et des échanges de semences sur l'organisation de la diversité génétique et phénotypique du sorgho.

Les pratiques de sélection

Dans les chapitres III et IV, nous avons montré l'influence de l'organisation des sociétés sur leur façon de percevoir et de se représenter la variabilité phénotypique du sorgho. Atran et

Medin ont montré que les représentations que se font les sociétés de leur environnement influencent considérablement la façon dont elles agissent sur celui-ci. De telles différences ont-elles un impact sur les pratiques de sélection des semences ?

La variété *Mugeta* est identifiée de façon fortement consensuelle au sein du groupe Tharaka, contrairement au groupe Mbeere (chapitre IV). *Mugeta* correspond à une catégorie morphologique distincte, caractérisée notamment par ses glumes claires (Chapitre III). Cette variété n'est pourtant pas homogène, ni distincte génétiquement des autres variétés locales. Ces résultats suggèrent que les Tharaka maintiennent activement les caractéristiques phénotypiques de cette variété par sélection, essentiellement pour la couleur de la glume qui est fortement héritable. Cela semble être aussi le cas chez les Chuka, mais pas chez les Mbeere où nos observations suggèrent que ce phénotype est anecdotique au champ. En effet, la quasi-totalité des panicules présentant les caractéristiques de *Mugeta* ont été collectées dans des foyers Chuka et Tharaka (données non-présentées), confirmant que cette variété n'est quasiment pas cultivée par les Mbeere. La distribution hétérogène de ce phénotype entre les groupes ethniques souligne de surcroit l'absence d'échanges des Chuka et Tharaka vers les Mbeere.

Par ailleurs, nous avons vu qu' une part conséquente des Chuka nomment *Gadam* des individus pourtant assignés génétiquement au pool local (Chapitre III). Potentiellement, cela pourrait entrainer la diversification de *Gadam* par introgression du pool génétique local, comme cela a été observé par exemple pour la variété *Kaguru* dans ce même groupe ethnique. A l'inverse, une large part des Mbeere et des Tharaka identifient comme *Mbura-Imwe* et *Ngirigacha* des individus portant la signature génétique de *Gadam*. La confusion dans l'identification de *Gadam*, due à sa similarité phénotypique avec une partie des variétés présentes localement, pourrait donc favoriser un mélange considérable entre les deux pools génétiques locaux et introduits.

Cela nous amène à nous interroger sur les pratiques de sélection concernant *Ngirigacha* et *Mbura-imwe*, qui semblent être des catégories inclusives (taxons d'ordre supérieur) regroupant plusieurs variétés présentant des caractéristiques communes : grains blancs et glumes foncés pour *Mbura-imwe*, panicule semi-compacte à compacte et grains blancs, crème ou gris pour *Ngirigacha*. Si *Ngirigacha* est une catégorie inclusive, la situation est moins claire pour *Mbura-imwe*, qui est aujourd'hui confondue avec *Gadam* mais qui était utilisée pour désigner toutes les variétés locales ne présentant ni les grains gris de *Muruge*, ni les glumes beiges de *Mugeta*. Il n'a pas été possible d'observer systématiquement si la large gamme de phénotype identifiés comme *Ngirigacha* et *Mbura-imwe* par les agricultrices sont sélectionnés ensembles comme une seule et même catégorie, où si les agricultrices distinguent et sélectionnent séparément des sous-catégories phénotypiques mieux définies au sein de ces catégories inclusives. Une unique observation chez une agricultrice montre qu'elle sélectionne *Mbura-imwe* comme s'il s'agissait d'une catégorie unique, en choisissant des panicules présentant des phénotypes divers tant par la couleur de leurs glumes que par la forme de leur panicule. Des observations systématiques seraient nécessaires pour conclure. Enfin, les Chuka ont rarement fait usage de ces catégories inclusives durant l'expérience d'identification. Ces différences de taxonomie entre groupes ethniques peuvent avoir un impact sur les populations de sorgho qu'ils cultivent puisqu'elles traduisent des différences de perception et de représentation, qui influencent les pratiques de sélection.

La situation observée pour *Kaguru* est assez singulière, puisqu'il est davantage diversifié chez les Chuka que dans les deux autres groupes, bien qu'aucune différence d'identification n'ait été observée entre groupes ethniques. Il ne faut cependant pas négliger que l'identification ne reflète pas complètement les pratiques : une panicule identifiée comme appartenant à une catégorie n'est pas forcément jugée sélectionnable et représentative de cette catégorie par la majorité des agricultrices. Il est donc possible que les Tharaka et les Mbeere aient une sélection plus conservatrice que les Chuka, mais il faut néanmoins rester prudent sur ce point car d'autres phénomènes peuvent entrer en jeu comme la distribution récente de semences par des ONG dans certains groupes, la différence de qualité des semences diffusées initialement, celle de la taille des populations cultivées, ou encore celle des mélanges variétaux et arrangement des espèces dans les systèmes de culture. Ces facteurs, donc, peuvent également avoir généré ces différences de diversité dans les populations de *Kaguru* échantillonnées dans les trois groupes

Les échanges de semences

Nous avons vu dans le chapitre V que les échanges de semences directs, d'agriculteur à agriculteur, étaient fortement cloisonnés. Ce cloisonnement n'a pas seulement un impact sur la nomenclature des variétés, il a manifestement aussi un impact direct sur la distribution de certaines variétés paysannes qui ne sont quasiment pas commercialisées au marché, comme les variétés à cycle long *Mugana*, *Mucuri*, *Kathirigwa, Mukumbu* ou *Mbunge*. Ces variétés sont peu fréquentes, mais des échanges ont été relevés chez les Chuka et les Tharaka où elles ont été inventoriées alors qu'elles semblent absentes du réseau d'échanges des Mbeere, chez qui elles n'ont pas été inventoriées. Les populations de sorgho cultivées par les Tharaka et surtout les Chuka sont davantage diversifiées génétiquement que celles des Mbeere. De plus ces populations sont différenciées génétiquement, principalement en raison de la plus grande fréquence des variétés à cycle long, porteuses d'allèles rares, chez les Chuka par rapport aux Mbeere (Chap. II). Le cloisonnement ethnique des échanges de semences d'agriculteur à agriculteur limite donc l'uniformisation génétique des populations de sorgho cultivées par les trois groupes ethniques, et ce malgré un recours commun et fréquent au marché.

Au sein du pool génétique des variétés locales à cycle court, plusieurs sous-groupes génétiques ont été identifiés. Ils ne correspondent cependant pas à des unités phénotypiques distinctes et uniformes, et ne présentent pas de structuration spatiale. Cela suggère premièrement que les variétés appartenant au pool local cultivées sur la zone descendent de plusieurs sous-populations différenciées génétiquement, et que peu de flux de gènes existent entre elles au sein des champs. En effet, les agriculteurs ont accès à des semences venant probablement de populations différenciées génétiquement par le biais du marché. Les flux de gènes sont ensuite vraisemblablement limités à l'échelle des parcelles du fait du régime principalement autogame du sorgho et de l'arrangement des parcelles. Le maintien de caractéristiques phénotypiques différentes au sein de ces sous-populations est attribuable aux pratiques de sélection des agriculteurs.

Des pratiques à la distribution de la diversité

Cette thèse montre donc qu'en l'absence de variabilité environnementale, des sociétés adjacentes maintiennent des différences génétiques et phénotypiques entre les populations de sorgho qu'elles cultivent. Le maintien de ces différences malgré un recours commun au marché local s'explique par l'organisation des échanges de semences et dans une certaine mesure par des différences dans leurs pratiques de sélection. Le marché n'uniformise pas ces différences car seul un petit nombre de variétés y sont vendues et que les différents groupes continuent malgré tout de maintenir leurs propres variétés, comme le font les Tharaka avec *Mugeta*. Ces mécanismes, observés à l'échelle locale, ne sont pas anecdotiques. Le Kenya à lui seul compte plus de 68 langues, traduisant autant de différences culturelles, le continent

Africain en compte lui plus de 2 000 (Lewis et al. 2013). Cette grande diversité culturelle traduit également des modalités d'organisation sociale diverse qui ont un impact sur la façon dont les semences et le savoir sont transmis (Leclerc and Coppens d'Eeckenbrugge 2012). Par exemple, les différences concernant les modalités de mariage entre sociétés matrilinéaire ou patrilinéaire impliquent des différences de modalités de diffusion des semences qui se répercutent sur la structuration spatiale de la diversité génétique (Delêtre et al. 2011). Il est donc vraisemblable que l'hypothèse émise par Harlan et Stemler (1976) pour expliquer la distribution des races de sorgho en Afrique devienne vérifiée. L'impact des sociétés sur la distribution spatiale de la variabilité génétique et phénotypique des plantes cultivées est sans doute plus important que n'ont pu le laisser croire les études uniquement centrées sur l'interaction génotype × environnement.

III-Implications pour la caractérisation, la conservation et l'amélioration des ressources génétiques

Dans la caractérisation des facteurs d'organisation de la diversité génétique et phénotypique des plantes cultivées *in situ*, la diversité culturelle des sociétés a largement été négligée. Leur statut de ressource domestiquée, modelée par l'homme pour reprendre les termes de Harlan (1975), n'a que trop peu été pris en considération. En effet, la majorité des prospections et collectes des ressources génétiques ont été conçues pour maximiser la couverture des zones agro-climatiques. Cela a été notamment le cas pour le sorgho, bien que l'effet du climat sur la distribution de ces ressources génétiques *in situ* ne soit pas toujours vérifié (Deu et al. 2008). L'effet de l'homme sur les ressources génétiques est parfois présenté comme une somme de pratiques individuelles sans réelle cohérence globale en dehors d'un objectif commun aux individus qui serait de maximiser la production au regard des contraintes de l'environnement. Cette thèse montre qu'au contraire, une cohérence est mise en évidence lorsque l'on examine les représentations et les pratiques des agriculteurs au regard de l'organisation des sociétés dans lesquelles ils sont insérés. La prise en compte de la diversité culturelle des agriculteurs et de l'organisation interne des sociétés ainsi que des relations qu'elles ont entre elles permettraient de mieux caractériser les ressources génétiques qui continuent d'être maintenue dans les systèmes familiaux.

La correspondance entre la taxonomie paysanne et la structure de la variabilité génétique et phénotypique est notamment une problématique centrale pour la collecte et la conservation des ressources génétiques *in situ.* L'étude présentée dans le chapitre IV montre que l'usage de méthode de classification sans *a priori* pour caractériser la diversité phénotypique ne permet pas de mettre en évidence les critères qui fondent la taxonomie des agriculteurs. Cela a des conséquences pour les programmes de conservation, pour lesquels l'identification des critères qui distinguent les variétés paysannes, unités de gestion des agriculteurs (Bellon and Brush 1994), est un enjeu majeur. Or, l'identification de ces critères de classification est difficilement abordable par le biais de simples entretiens. Les approches probabilistes avec le recours à des arbres de régression sur un large échantillon d'agriculteurs en considérant le nom de variété comme un cluster *a priori* se montrent bien adaptées pour décrire le système de taxonomie local. Ce système sert de cadre de référence aux agriculteurs, même s'il existe des disparités de savoir et de compétences dans l'utilisation de la taxonomie locale au sein des sociétés.

Dans le cas des populations étudiées, une variété locale est bien une unité phénotypique mais n'est pas une unité génétique distincte, du moins dans le cas des variétés à cycle court. De plus, au sein même des exploitations agricoles, des individus appartenant à une même variété nommée peuvent être distribués dans différents sous-groupes génétiques (données non présentées dans cette thèse). Nous sommes donc dans un cas bien différent de celui décrit par Soler et al. (2013) au nord-Cameroun, ou Rabbi et al. (2010) au Soudan où les variétés paysannes correspondent à des clusters génétiques distincts. Ces différences doivent être prises en compte pour la collecte et la conservation des ressources génétiques. Dans notre cas d'étude, l'usage de marqueurs génétiques neutres s'est montré insuffisant pour caractériser la diversité phénotypique des variétés, et inadapté pour mettre en évidence la logique des taxonomies locales.

Les observations décrites dans cette thèse peuvent également être utiles pour l'amélioration variétale, notamment dans le cadre de programmes de sélection participative. Nos observations soulignent notamment l'importance de la distinction perceptuelle des nouvelles variétés introduites (perceptual distinctiveness), comme cela avait été déjà argumenté par Boster (1985) et Gibbson (2009). Comme le montre le cas de *Gadam*, une variété qui ne se différencie pas clairement des variétés présentes localement sera confondue et potentiellement moins bien maintenue par les agriculteurs. D'autre part, les caractéristiques phénotypiques des variétés peuvent faire appel au système symbolique des agriculteurs. Les entretiens avec les agricultrices suggèrent que parmi les raisons de la large adoption de *Kaguru* figurent des qualités nutritionnelles symbolisées par sa couleur rouge. En effet, les agricultrices considèrent que cette variété est « bonne pour le sang » et notamment recommandé pour les femmes venant d'accoucher. Il ne s'agit que d'une part d'explication pour son succès durable, mais qui nous laisse entrevoir l'interconnexion entre systèmes de représentation et systèmes symboliques, et l'impact que cela peut avoir sur l'adoption d'une variété.

Malgré l'introduction massive et fréquente de variétés améliorées par les services d'extension du ministère de l'agriculture et des ONG, la part de variétés locales reste importante dans les systèmes de culture étudiés. Les agriculteurs continuent à maintenir diverses variétés locales bien que le sorgho ne soit pas une espèce spécialement valorisée et appréciée au niveau alimentaire. D'après nos observations, les agricultrices considèrent que les variétés locales sont mieux adaptées que les variétés introduites pour la préparation de porridge, mais en dehors de cela, les différentes variétés ne semblent pas correspondre à des usages particuliers. Il semble donc que des raisons utilitaires ne suffisent pas à expliquer le maintien d'une telle diversité, notamment le maintien de certaines variétés à faible fréquence. La diversité peut cependant être valorisée pour différentes raisons comme par exemple des raisons esthétiques, ou de prestige (Caillon 2005). Ces questions demanderaient des enquêtes systématiques, qui n'ont pu être réalisées dans cette étude.

IV- Pour une reconnaissance des variétés paysannes

Les pays du Sud sont incités par les pays du Nord à adopter des systèmes de propriété intellectuelle sur le vivant depuis l'entrée en vigueur du Trade-Related Aspects of Intellectual Property Right (TRIPs) en 1995, à l'instar de l'Organisation Mondiale du Commerce. Les pays du Sud doivent de ce fait s'engager à définir les droits de propriété intellectuelle concernant les variétés végétales soit par un système de brevet, soit par un système *sui generis* de type Protection des Obtentions Végétales mais laissant aux états la possibilité d'effectuer des modification notamment relatives au droits des agriculteurs (Gepts 2004). Cependant, les droits des agriculteurs évoqués dans l'article 8 (j) de la CBD ne sont pas définis clairement.

Le Kenya a adhéré au système UPOV en 1999, espérant ainsi dynamiser le secteur de l'amélioration variétale. L'adoption d'un tel système est clairement destinée à protéger les firmes semencières et s'avère totalement inadaptée à l'agriculture familiale au Kenya. L'argument selon lequel l'adoption de ce système dynamise la recherche en favorisant les innovations destinées aux zones marginales est difficilement convainquant quand on examine la diversité de ces systèmes agricoles. Comment définir un « idéotype de sélection» adapté à cette diversité de situations agro-écologiques, économiques et culturelles ? Une alternative pour cela résiderait dans les démarches de sélection participatives (vom Brocke et al. 2010; Morris and Bellon 2004), mais les systèmes familiaux ne représentent pas un intérêt économique suffisant pour que ce type de démarche soit développée à large échelle par les firmes industrielles.

L'adoption de ce système de marchandisation du vivant par les pays du Sud éveille des inquiétudes concernant l'effet qu'il peut avoir sur l'accélération de l'érosion génétique. La première inquiétude tient au remplacement des variétés paysannes par des variétés améliorées. Cependant, dans des zones marginales comme celle où cette thèse a été menée, cette situation ne semble pas devoir être redoutée trop fortement dans l'immédiat. D'une part, même si certaines variétés issues du secteur formel ont un certain succès, les agriculteurs conservent leurs variétés locales malgré tout. D'autre part, le coût de ces semences et des intrants qu'elles nécessitent excède les bénéfices que les agriculteurs peuvent espérer en tirer. Dans les zones présentant des conditions plus favorables, en revanche, le remplacement des variétés paysannes par les variétés améliorées est largement amorcé. Un autre sujet d'inquiétude serait l'interdiction de commercialisation de semences non inscrites au catalogue, mais une telle interdiction semble inapplicable actuellement dans des contextes comme ceux du Kenya.

Dans ce contexte de marchandisation des ressources génétiques, certains réfléchissent à la possibilité de protéger les variétés paysannes et les savoirs traditionnels associés. En effet, ce système a entrainé des dérives notoires, comme le cas du brevet déposé sur un cultivar de riz aromatique issu de croisement entre riz Basmati et riz long grain qui avait provoqué de vives polémiques en Inde. Certaines communautés revendiquent des droits de propriété intellectuelle sur leurs variétés, c'est par exemple le cas en Nouvelle-Zélande où des communautés Maori revendiquent de tels droits sur des variétés de patate-douce (Roullier 2012). La définition de ces droits parait difficilement envisageable au regard de la complexité des pratiques et des réseaux d'échanges ayant généré les variétés paysannes. Cette étude pointe donc, une fois de plus, la nécessité de repenser les systèmes de droits de propriétés sur le vivant pour une meilleure protection des droits des agriculteurs à échanger et cultiver librement leurs variétés paysannes.

V- Forces et limites de l'étude

Cette étude présente l'originalité de retracer l'influence des sociétés sur l'organisation de la diversité des plantes au travers de l'analyse de leurs représentations et de leurs pratiques. Les études sur cette problématique ont généralement traité les deux aspects séparément : soit elles ont constaté la correspondance entre la diversité culturelle des agriculteurs et la distribution de la diversité génétique (Sardos et al. 2012; Berthouly et al. 2009; Deu et al. 2008) ou phénotypique (Perales et al. 2005), soit elles ont traité de la relation entre organisation sociales et pratiques d'échanges (Badstue et al. 2007; Badstue et al. 2002; McGuire 2008; Zimmerer 2003), rarement en lien avec l'analyse de la concordance des dénominations entre agriculteurs (Boster 1986; Nuijten and Almekinders 2008). Cette thèse propose donc de combiner ces deux approches, la première appartenant plutôt au domaine de la génétique des populations, et la seconde au domaine de l'anthropologie sociale et culturelle, pour expliquer l'influence de l'organisation des sociétés sur l'organisation de la variabilité génétique et phénotypique du sorgho.

Cependant, une limitation majeure de notre étude est l'absence d'observations directes concernant les pratiques de sélection, permettant leur comparaison entre les trois groupes. En effet, le rôle des pratiques de sélection dans la distribution hétérogène de la diversité phénotypique et génétique reste une question ouverte à l'issue de cette thèse. Nous n'avons pu que formuler des hypothèses concernant leur rôle dans la diversification considérable de *Kaguru* chez les Chuka par rapport aux deux autres groupes, ou encore concernant l'absence de phénotypes correspondant à la variété *Mugeta* chez les Mbeere. Il serait nécessaire de confirmer ces hypothèses par des observations directes des pratiques de sélection. Ce type d'enquête est néanmoins difficile à mettre en œuvre, raison pour laquelle peu d'informations et d'études systématiques sont disponibles à ce sujet dans la littérature. Il serait particulièrement difficile de mettre en œuvre ce type d'observation dans les sociétés étudiées en raison des distance considérables à parcourir entre foyers, de la difficulté de connaitre les jours où le battage a lieu dans chaque foyer, et du nombre élevé d'observations qui seraient nécessaires pour pouvoir distinguer les tendances culturelles au delà de la variabilité individuelle. De telles observations permettraient pourtant d'éclaircir la relation entre identification, nomination et sélection, et notamment de tester si une panicule pour lesquels les agriculteurs présentent un agrément fort est forcément sélectionnée, et à l'inverse, si une panicule recueillant un faible agrément est contre-sélectionnée. Il serait également possible d'établir si les catégories *Ngirigacha* et *Mbura-imwe* sont sélectionnées et maintenues sciemment par les agriculteurs ou bien s'ils distinguent et sélectionnent des sous-catégories phénotypiques au sein de ces catégories inclusives.

L'approche probabiliste des taxonomies locales utilisée dans cette thèse est inspirée des travaux de Boster (1986) et de Romney (1986), également repris par Atran et Medin (2008). Ces travaux ont marqué une étape méthodologique en anthropologie culturelle, notamment suite aux controverses qui ont agité ce domaine concernant les travaux de Margaret Mead au Samoa, à laquelle il était reproché d'avoir tiré des interprétations biaisés par le choix de ses informateurs (Freeman 1983). Ces approches probabilistes sont relativement méconnue dans le domaine de l'ethnobotanique, et rarement mises en œuvre dans les études traitant des savoirs paysans. Bien que complexes à mettre en œuvre auprès d'un grand nombre de participants, en particulier dans des zones où l'habitat est dispersé, ces approches permettent une estimation plus rigoureuse de la distribution des connaissances au sein des sociétés. Leur usage pour l'étude des taxonomies variétales et des pratiques de sélection gagnerait à être répandu et conférerait davantage de rigueur dans l'utilisation de concepts tels que celui de « savoirs paysans », qui sont fréquemment manipulés mais rarement étudiés rigoureusement.

L'utilisation de ces méthodes issues du domaine de l'anthropologie culturelle s'inscrit dans une démarche emic, visant à mettre en évidence la logique et la cohérence interne des taxonomies paysannes en privilégiant le point de vue de l'agriculteur. Elle s'est montrée adéquate pour comprendre les taxonomies locales du sorgho. Ce type d'approche devrait clairement être privilégié pour caractériser les variétés paysannes, qui ont jusqu'à présent été considérées sous un angle etic, se référant à un cadre conceptuel et un système de représentation externe.

La comparaison interculturelle, suivant le modèle G×E×S (Leclerc and Coppens d'Eeckenbrugge 2012), a prouvé son intérêt pour l'étude de l'influence de l'organisation des sociétés sur celle des plantes cultivées via le contrôle des conditions environnementales, notamment la distance spatiale, les conditions agro-écologiques, et le contexte économique. Ce type d'approche n'est possible qu'à l'échelle locale et démontre la complémentarité de cette échelle d'étude vis-à-vis des échelles régionales ou continentales qui sont largement privilégiées aujourd'hui. Les études locales permettent une compréhension fine des mécanismes à l'œuvre, qu'il n'est pas possible d'obtenir à d'autres échelles. Loin d'être de simples études de cas de portée restreinte, elles permettent de mettre en évidence des mécanismes qu'il serait impossible de montrer à plus grande échelle en raison de l'interaction des facteurs. Elles permettent également de dégager des principes généralisables, applicables à d'autres situations.

Dans le cadre de cette thèse, nous avons décrit des sociétés présentant des modalités spécifiques d'organisation sociale, et qui plus est, dans une situation d'ajustement face à des changements récents de leur mode de vie. Cependant, l'objet de notre démonstration n'est pas de généraliser cette situation particulière à l'ensemble des systèmes agricoles familiaux, mais bien de démontrer que chaque société doit être considéré au regard de son organisation propre. Cette thèse illustre la complexité des facteurs entrant en jeu dans les pratiques des agriculteurs, et montre qu'il est illusoire de chercher à expliquer leurs pratiques par le simple objectif de maximiser leur production face aux contraintes de leur environnement agro-écologique ou économique.

REFERENCES

Agresti, A. 2007. *An introduction to categorical data analysis*. Hoboken: John Wiley & Sons.

Allaby, R.G., D.Q. Fuller, and T.A. Brown. 2008. The genetic expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences of the United States of America* 105 (37):13982-13986.

Ambler, C. 1988. The Great Famine, 1897–1901. In *Kenyan communities in the age of imperialism. The central region in the late nineteenth century*. New Haven and London: Yale University Press.

Anderson, M.J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26 (1):32-46.

Arriola, P.E., and N.C. Ellstrand. 1996. Crop-to-weed gene flow in the genus *Sorghum* (Poaceae): Spontaneous interspecific hybridization between johnsongrass, *Sorghum halepense*, and crop sorghum, *S. bicolor*. *American Journal of Botany* 83 (9):1153-1159.

Atran, S. 1985. The Nature of Folk-Botanical Life Forms. *American Anthropologist* 87 (2):298-315.

Atran, S., and D. Medin. 2008. *The native mind and the cultural construction of nature*. Cambridge, Massachusetts: MIT press.

Baco, M.N., G. Biaou, J.L. Pham, and J.P. Lescure. 2008. Geographical and social factors of cultivated yam diversity in northern Benin. *Cahiers Agricultures* 17(2): 172–177.

Badstue, L.B., M.R. Bellon, J. Berthaud, A. Ramírez, D. Flores, and X. Juárez. 2007. The dynamics of farmers' maize seed supply practices in the central valleys of Oaxaca, Mexico. *World Development* 35 (9):1579-1593.

Badstue, L.B., M.R. Bellon, X. Juárez, I. Manuel, and A.M. Solano. 2002. Social Relations and Seed Transactions among Small-Scale Maize Farmers in the Central Valleys of Oaxaca, Mexico. *CIMMYT Economics Working Paper* No. 02–02.

Bandiera, O., and I. Rasul. 2006. Social networks and technology adoption in northern Mozambique. *The Economic Journal* 116(514): 869–902.

Barnaud, A., G. Trigueros, D. McKey, and HI. Joly. 2008. High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity* 101 (5):445-452.

Barnaud, A., M. Deu, E. Garine, D. McKey, and H. Joly. 2007. Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theoretical and Applied Genetics* 114 (2):237-248.

Barrau, J. 1962. *Les plantes alimentaires de l'Océanie : origines, distribution et usages*. Faculté des Sciences de Marseille, Marseille.

Barth F. 1969. Introduction In: Barth F, editor. *Ethnic groups and boundaries: The social organization of culture difference*. Boston: Little, Brown. pp. 9-38

Baty, F., M. Facompré, J. Wiegand, J. Schwager, and M. Brutsche. 2006. Analysis with respect to instrumental variables for the exploration of microarray data structures. *BMC bioinformatics* 7: 422.

Baudouin, L., and P. Lebrun. 2009. Coconut (*Cocos nucifera* L.) DNA studies support the hypothesis of an ancient Austronesian migration from Southeast Asia to America. *Genetic Resources and Crop Evolution* 56 (2):257-262.

Bazile, D., S. Dembélé, M. Soumaré, and D. Dembélé. 2008. Utilisation de la diversité variétale du sorgho pour valoriser la diversité des sols au Mali. *Cahiers Agricultures* 17(2): 86–94.

Belkhir K., P. Borsa, L. Chikhi, N. Raufaste, F. Catch (2004) GENETIX, software under Windows™ for the genetics of populations. 4.05 ed. Montpellier: University of Montpellier.

Bellon, M.R., and S.B. Brush. 1994. Keepers of maize in Chiapas, Mexico. *Economic Botany* 48 (2):196-209.

Bellon, M.R. 2004. Conceptualizing interventions to support on-farm genetic resource conservation. *World Development* 32(1): 152–172.

Benjamini Y, Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1: 289-300.

Benz, B.F., L.R. Sanchez-Velasquez, and F.J. Santana Michel. 1990. Ecology and ethnobotany of Zea diploperennis: preliminary investigations. *Maydica* 35 (2):85-98.

Benzécri, J.P. 1973. *L'analyse des données. II L'analyse des correspondances*. Paris: Bordas.

Berlin, B. 1973. Folk systematics in relation to biological classification and nomenclature. *Annual Review of Ecology and Systematics* 4:259-271.

Bernard, F.E. 1972. *East of Mount Kenya: Mem agriculture in transition*. Afrika-Studien. München: Weltforum verlag.

Berthouly, C., D. Do Ngoc, S. Thévenon, D. Bouchel, T. Van, et al. 2009. How does farmer connectivity influence livestock genetic structure? A case-study in a Vietnamese goat population. *Molecular Ecology* 18: 3980-3991.

Billot C., R. Rivallan, M. Sall, D. Fonceka, M. Deu, et al. 2012. A reference microsatellite kit to assess for genetic diversity of Sorghum bicolor (Poaceae). *American Journal of Botany* 99: e245-e250.

Boster, J.S. 1985. Selection for perceptual distinctiveness: evidence from Aguaruna cultivars of Manihot esculenta. *Economic Botany* 39 (3):310-325.

Boster, J.S. 1986. Exchange of varieties and information between Aguaruna manioc cultivators. *American Anthropologist* 88(2): 428–436.

Bray, J.R., and J.T Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* 27: 325-349.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks.

Brokensha, D., and J. Glazier. 1973. Land reform among the Mbeere of Central Kenya. *Africa* 43: 182–206.

Brush, S.B. 2000. *Genes in the field: on-farm conservation of crop diversity*. Rome: Copublished by International Plant Genetic Resources Institute.

Brush, S.B. 2000. The issues of *in situ* conservation of crop genetic resources. In *Genes in the field: on-farm conservation of crop diversity*, ed. S.B. Brush, 3-28. Rome: Copublished by International Plant Genetic Resources Institute.

Brush, S.B., and H.R. Perales. 2007. A maize landscape: Ethnicity and agro-biodiversity in Chiapas Mexico. *Agriculture, ecosystems & environment* 121 (3):211-221.

Caillon, S. 2005. *Pour une conservation dynamique de l'agrobiodiversité: Gestion locale de la diversité variétale d'un arbre «des Blancs» (cocotier, Cocos nucifera L.) et d'une plante «des ancêtres» (taro, Colocasia esculenta (L.) Schott) au Vanuatu*. Université d'Orléans, Orléans.

Caillon, S., J. Quero-Garcîa, J. P. Lescure, and V. Lebot. 2006. Nature of taro (Colocasia esculenta (L.) Schott) genetic diversity prevalent in a Pacific Ocean island, Vanua Lava, Vanuatu. *Genetic Resources and Crop Evolution,* 53 (6):1273-1289.

Camberlin, P., J. Boyard-Micheau, N. Philippon, C. Baron, C. Leclerc, and C. Mwongera. 2012. Climatic gradients along the windward slopes of Mount Kenya and their implication for crop risks. Part 1: climate variability. *International Journal of Climatology* wileyonlinelibrary.com.

Camberlin, P., V. Moron, R. Okoola, N. Philippon, and W. Gitau. 2009. Components of rainy seasons' variability in Equatorial East Africa: onset, cessation, rainfall frequency and intensity. *Theoretical and Applied Climatology* 98 (3):237-249.

Casas, A., B. Pickersgill, J. Caballero, and A. Valiente-Banuet. 1997. Ethnobotany and domestication in xoconochtli, Stenocereus stellatus (Cactaceae), in the Tehuacán Valley and la Mixteca Baja, Mexico. *Economic Botany* 51 (3):279-292.

Cauvin, J., M.C. Cauvin, D. Helmer, and G. Willcox. 1997. L'homme et son environnement au Levant nord entre 30 000 et 7 500 BP. *Paléorient* 23 (2):51-69.

Cavalli-Sforza, L. L. 2002. Demic diffusion as the basic process of human expansions. In *Examining the farming/language dispersal hypothesis*, eds. P. Bellwood, and C. Renfrew, 79-88. Cambridge: McDonald Institute for Archaeological Research.

Cavalli-Sforza, L., and M. Feldman. 1981. Cultural transmission and evolution: a quantitative approach. *Monographs in population biology* 16:1–388.

CBD. 1992. The convention on biological diversity. Secretariat of the Convention on Biological Diversity, United Nations Environment Programme, Montreal.

Chacón, M. I., B. Pickersgill, and D. G. Debouck. 2005. Domestication patterns in common bean (Phaseolus vulgaris L.) and the origin of the Mesoamerican and Andean cultivated races. *Theoretical and Applied Genetics* 110 (3):432-444.

Chaïr, H., D. Cornet, M. Deu, M.N. Baco, A. Agbangla, M.F. Duval, and J. L. Noyer. 2010. Impact of farmer selection on yam genetic diversity. *Conservation Genetics* 11 (6):2255-2265.

Chantereau, J., and R. Nicou. 1991. *Le Sorgho*. Paris: Maisonneuve & Larose.

Chesaina, C. 1997. *Oral literature of the Embu and Mbeere*. Nairobi: East African Educational Publishers.

Chessel, D., A. Dufour, and J. Thioulouse. 2004. The ade4 package. I. One-table methods. *R News* 4: 5–10.

Cohen, M.N. 1977. *The food crisis in prehistory: overpopulation and the origins of agriculture*. New Haven: Yale University Press.

Conley, T., and C. Udry. 2001. Social learning through networks: the adoption of new agricultural technologies in Ghana. *American Journal of Agricultural Economics* 83(3): 668–673.

Conley, T., and C. Udry. 2010. Learning about a new technology: pineapple in Ghana. *The American Economic Review* 100(1): 35–69.

D'Andrade, R.G. 1981. The cultural part of cognition. *Cognitive science* 5 (3):179-195.

Darwin, C. 1859. *On the origins of species by means of natural selection*. London: Murray.

David, S., and L. Sperling. 1999. Improving technology delivery mechanisms: lessons from bean seed systems research in Eastern and Central Africa. *Agriculture and Human Values* 16(4): 381–388.

De Wet, J.M.J, J.R. Harlan, and E.G. Price. 1976. Variability in *Sorghum bicolor*. In *Origins of African plant domestication* 453-463. La Haye, Pays Bas: Harlan, J.R. et al.

De Wet, J.M.J., and J.P. Huckabay. 1967. The origin of Sorghum bicolor. II. Distribution and domestication. *Evolution* 21 (4):787-802.

Delêtre, M., D.B. McKey, and T.R. Hodkinson. 2011. Marriage exchanges, seed exchanges, and the dynamics of manioc diversity. *Proceedings of the National Academy of Sciences of the United States of America* 108 (45):18249-18254.

Deu, M., F. Sagnard, J. Chantereau, C. Calatayud, D. Hérault, C. Mariac, J.L. Pham et al. 2008. Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theoretical and Applied Genetics* 116(7): 903–913.

Deu, M., P. Hamon, F. Bonnot, and J. Chantereau. 1999. Le sorgho. In *Diversité Génétique des Plantes Tropicales Cultivées*, eds. P. Hamon, M. Seguin, X. Perrier, and J.C. Glaszmann, 351-387. Paris, France: Cirad.

Diamond, J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* 418 (6898):700-707.

Doebley, J. 1990. Molecular evidence and the evolution of maize. *Economic Botany* 44 (3):6-27.

Doggett, H. 1988. *Sorghum*. Second edn. Aufl. New York: John Wiley and Sons Inc.

Doledec, S., and D. Chessel. 1989. Seasonal successions and spatial variables in freshwater environments. Part 2: Distinction between seasonal and environmental-effects within floral and faunistical data. *Acta Oecologica-Oecologia Generalis* 10: 207–232.

Donohue, M., and T. Denham. 2010. Farming and Language in Island Southeast Asia Reframing Austronesian History. *Current Anthropology* 51 (2):223-256.

Downing, T.E. 1992. *Climate change and vulnerable places: global food security and country studies in Zimbabwe, Kenya, Senegal, and Chile*. Research Report–Environmental Change Unit, University of Oxford.

Dray, S., and A.B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22 (4):1-20.

Elias, M., L. Penet, P. Vindry, D. McKey, O. Panaud, and T. Robert. 2001. Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (Manihot esculenta Crantz), in a traditional farming system. *Molecular Ecology* 10 (8):1895-1907.

Ellstrand, N.C., and K.W. Foster. 1983. Impact of population structure on the apparent outcrossing rate of grain sorghum (*Sorghum bicolor*). *Theoretical and Applied Genetics* 66 (3-4):323-327.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14 (8):2611-2620.

Fadiman, J.A. 1993. *When we began there were witchmen: an oral history from Mount Kenya*. Los Angeles: University of California Press.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164 (4):1567-1587.

FAO. 1996. Report on the state of the world's plant genetic resources. In *International Technical Conference on Plant Genetic Resources*. Leipzig, Germany: FAO, Rome.

FAOSTAT. 2012. Agriculture Organization of the United Nations. Accessed October 2013.

Frankel, O., and E. Bennett. 1970. *Genetic resources in plants: their exploration and conservation*. Oxford: Blackwell.

Freeman, D. 1983. *Margaret Mead and Samoa: The making and unmaking of an anthropological myth*. Cambridge, MA: Harvard University Press.

Friedberg, C. 1991. Operative aspects of Folk Classification. In *Man and a half: essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer*, ed. A. Pawley, 102-109. Auckland: The Polynesian Society.

Gepts, P. 1990. Biochemical evidence bearing on the domestication of Phaseolus (Fabaceae) beans. *Economic Botany* 44 (3):28-38.

Gepts, P. 2004a. Crop domestication as a long-term selection experiment. *Plant breeding reviews* 24 (2):1-44.

Gepts, P. 2004b. Who owns biodiversity, and how should the owners be compensated? *Plant physiology* 134 (4):1295-1307.

Gepts, P. 2006. Plant genetic resources conservation and utilization. *Crop Science* 46 (5):2278-2292.

Gepts, P., T.C. Osborn, K. Rashka, and F.A. Bliss. 1986. Phaseolin-protein variability in wild forms and landraces of the common bean (Phaseolus vulgaris): evidence for multiple centers of domestication. *Economic Botany* 40 (4):451-468.

Gibson, R.W. 2009. A review of perceptual distinctiveness in landraces including an analysis of how its roles have been overlooked in plant breeding for low-input farming systems. *Economic Botany* 63 (3):242-255.

Glazier, J. 1970. *Ritual and social conflict: circumcision and oath-taking in Mbeere.* Nairobi: Institute of Development Studies, Nairobi University.

Glémin, S., and T. Bataillon. 2009. A comparative view of the evolution of grasses under domestication. *New phytologist* 183 (2):273-290.

Goudet, J. 2001 FSTAT, a program to estimate and test gene diversities and fixation indices 2.9.3 ed. Lausanne: Institute of Ecology.

Goudet J., M. Raymond, T. de Meeüs, F. Rousset. 1996. Testing differentiation in diploid populations. *Genetics* 144: 1933-1940.

Grenand, F. 2002. Stratégies de nomination des plantes cultivées dans une société tupi-guarani, les Wayãpi. *Amerindia* 26/27:209-247.

Guillaumet, J.L., and J. Pernès. 1984. Stratégies de prospection. In *Gestion des ressources génétiques des plantes. Tome 2: Manuel*, ed. J. Pernès, 109-136. Paris: Agence de Coopération Culturelle et Technique.

Haaland, R. 1998. Theory and evidence of the archaelogical interpretation of the transition from gathering to domestication. In *Changes in agrarian systems*, eds. J. Hather, and C. Gosden. London: Routledge.

Harlan, H.V., and M.L. Martini. 1936. Problems and results in barley breeding. In *Yearbook of agriculture*. Washington, DC: USDA.

Harlan, J. R. 1972. Genetics of disaster. *Journal of Environmental Quality* 1 (3):212-215.

Harlan, J. R., and J. M. J. De Wet. 1972. A simplified classification of cultivated sorghum. *Crop Science* 12 (2):172-176.

Harlan, J. R. 1975. *Crop and man*. Madison, Wisconsin: American Society of Agronomy.

Harlan, J. R., and A. Stemler. 1976. The races of sorghum in Africa. In *Origins of African plant domestication*, ed. J. R. Harlan, 465-478. The Hague: Mouton.

Harlan, J.R., J.M.J. De Wet, and A.B.L. Stemler. 1976 *Origins of African plant domestication*. The Hague: De Gruyter Mouton.

Harlan, J.R. 1995. *The living fields: our agricultural heritage*. Cambridge, UK: Cambridge University Press.

Hawkes, J. G. 1971. Conservation of plant genetic resources. *Outlook on Agriculture* 6 (6):248-253.

Heine, B., and W. J. G. Moehlig. 1980. Geographical and historical introduction: language and society. In *Language and dialect atlas of Kenya*. Volume I, eds. B. Heine, and W. J. G. Moehlig. Berlin: Dietrich Reimer Verlag.

Herbich, I., and M. Dietler. 2008. The long arm of the mother-in-law: Learning, postmarital resocialization of women, and material culture style. In *Breaking down boundaries: Anthropological approaches to cultural transmission and material culture*, eds. M.T. Stark, B.J. Bowser, and L. Horne. Tucson: University of Arizona Press.

Hillman, G. C., and M. S. Davies. 1990. Domestication rates in wild type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society,* 39 (1):39-78.

Hodgkin, T., R. Rana, J. Tuxill, D. Balma, A. Subedi, I. Mar, R. Karamura et al. 2007. Seed Systems and Crop Genetic Diversity in Agroecosystems. In *Managing biodiversity in agricultural ecosystems*, eds. D. Jarvis, C. Padoch, and D. Cooper, 77-116. New York: Colombia University Press.

Holden, C. J. 2002. Bantu language trees reflect the spread of farming across sub-Saharan Africa: a maximum-parsimony analysis. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269 (1493):793-799.

Hunn, E. 1982. The utilitarian factor in folk biological classification. *American Anthropologist* 84 (4):830-847.

International Fund for Agricultural Development (IFAD), 2001. *Rural Poverty Report 2001: The Challenge of Ending Rural Poverty*. Oxford: Oxford University Press for IFAD.

International Plant Genetic Resources Institute (IPGRI). 1993. *Descriptors for Sorghum* [*Sorghum bicolor (L.) Moench*]. Rome: IBPGR / ICRISAT.

Jaetzold, R., H. Schmidt, B. Hornetz, and C. Shisanya. 2007. *Farm management handbook of Kenya Vol II, Part C, East Kenya. Subpart C1, Eastern Province*. Nairobi, Kenya: Ministry of Agriculture/GTZ.

Jarvis, D. I., and T. Hodgkin. 1999. Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular Ecology* 8 (s1):S159-S173.

Jarvis, D., A. Brown, P. Cuong, L. Collado-Panduro, L. Latournerie-Moreno, S. Gyawali, T. Tanto et al. 2008. A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. *Proceedings of the National Academy of Sciences of the United States of America* 105 (14):5326-5331.

Jost, L. 2006. Entropy and diversity. *Oikos* 113 (2):363-375.

Labeyrie, V., B. Rono, and C. Leclerc. 2013. How social organization shapes crop diversity: an ecological anthropology approach among Tharaka farmers of Mount Kenya. *Agriculture and Human Values*:1-11.

Labeyrie, V., J.I. Kamau, and C. Leclerc. In prep.-a. Farmers' consistency in naming sorghum varieties: the share of culture. *In preparation for publication*.

Labeyrie, V., J.I. Kamau, C. Dubois, C. Catalayud, R. Rivalan, A. Barnaud, and C. Leclerc. In prep.-b. What is a landrace? Formalizing farmers' taxonomy of sorghum through an emic approach. *In preparation for publication*.

Labeyrie, V., M. Deu, A. Barnaud, C. Catalayud, M. Buiron, P. Wambugu, S. Manel, J.-C. Glaszmann, and C. Leclerc. 2013. Influence of ethnolinguistic diversity on sorghum evolution in subsistence farming systems in eastern Kenya. *Submitted for publication*.

Lacy, S.M., D.A. Cleveland, and D. Soleri. 2006. Farmer choice of sorghum varieties in southern Mali. *Human Ecology* 34(3): 331–353.

Lambert, H. 1947. *The use of indigenous authorities in tribal administration: studies of the Meru in Kenya Colony*. School of African Studies, University of Cape Town.

Lambert, H.E. 1950. History of the tribal occupation of the land. In *The system of land tenure in the Kikuyu land unit.*, ed. H.E. Lambert. Cape Town: The School of African Studies, Cape Town University.

Lambert, H.E. 1956. *Kikuyu social and political institutions*. London: Oxford University Press.

Leclerc, C., and G. Coppens d'Eeckenbrugge. 2012. Social organization of crop genetic diversity. The G × E ×S interaction model. *Diversity* 4(1): 1–32.

Lewis, M.P., F.S. Gary, and C.D. Fennig. 2013. *Ethnologue: Languages of the World, Seventeenth edition.* Dallas: SIL International.

Linsig, N. 2009. *"We were not created the same way": Influence du genre sur l'agriculture chez les Meru du Mont Kenya*. Université de Neuchâtel, Neuchâtel.

Longley, C. 2000. *A social life of seeds: Local management of crop variability in North-Western Sierra Leone*. University of London, London, UK.

Louette, D., A. Charrier, and J. Berthaud. 1997. In situ conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Economic Botany* 51 (1):20-38.

Louette, D., and M. Smale. 2000. Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. *Euphytica* 113 (1):25-41.

Manly, B.F.J. 1997. *Randomization, bootstrap, and Monte-Carlo methods in biology*. London, UK: Chapman and Hall.

Mann, J.A., C.T. Kimber, and F.R. Miller. 1983. *The origin an early cultivation of sorghums in Africa*. Texas Agricultural Experiment Station.

Martin, M.A. 1975. L'ethnobotanique, science per se? À propos d'un livre de B. Berlin «The principles of Tzeltal plant classification». *Journal d'Agriculture tropicale et de Botanique appliquée* 22:237-276.

Mati, B.M. 2000. The influence of climate change on maize production in the semi-humid–semi-arid areas of Kenya. *Journal of Arid Environments* 46(4): 333–344.

McElreath, R., and P. Strimling. 2008. When natural selection favors imitation of parents. *Current Anthropology* 49(2): 307–316.

McGuire, S. 2008. Securing access to seed: social relations and sorghum seed exchange in eastern Ethiopia. *Human Ecology* 36(3): 217–229.

Middleton, J. 1953. *The central tribes of the north-eastern Bantu*. Ethnographic Survey of Africa. London: International African Institute.

Moehlig, W.J.G., G. Guarisma, and S. Platiel. 1980. La dialectométrie: une méthode de classification synchronique en Afrique. In G. Guarisma, and S. Platiel. *Dialectologie et comparatisme en Afrique Noire.* Actes des journées de dialectologie Africaine, Ivry, 2-5 juin 1980. Paris, France: SELAF.

Morris, M.L., and M.R. Bellon. 2004. Participatory plant breeding research: opportunities and challenges for the international crop improvement system. *Euphytica* 136 (1):21-35.

Murdock, G.F. 1959. *Africa: Its people and their cultural history*. London: Mc Graw-Hill.

Mwaniki, H.S.K.1973. *The living history of Embu and Mbeere to 1906*. East African Literature Bureau.

Nakazawa, M. 2013. fmsb: Functions for medical statistics book with some demographic data. 0.3.8 ed.

National Census, 2009. Kenya National Bureau of Statistics census. https://www.opendata.go.ke/vision. Accessed October 2013.

Nei, M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nuijten, E., and C.J.M Almekinders. 2008. Mechanisms explaining variety naming by farmers and name consistency of rice varieties in the Gambia. *Economic Botany* 62(2): 148–160.

Oksanen, J., F.G. Blanchet, R. Kindt, P. Legendre, P.R. Minchin, R. O'Hara, G.L. Simpson, P. Solymos, M. Stevens, and H. Wagner. 2012. Vegan: Community Ecology Package.

Ollitrault, O., J.L. Noyer, J. Chantereau, and J.C. Glaszmann. 1997. Structure génétique et dynamique des variétés traditionnelles de sorgho au Burkina Faso. In *Gestion des ressources génétiques des plantes en Afrique des savanes*. Bamako, Mali.

Orlove, B. S. 1980. Ecological anthropology. *Annual Review of Anthropology* 9: 235–273.

Paillex, A., E. Castella, and G. Carron. 2009. Aquatic macroinvertebrate response along a gradient of lateral connectivity in river floodplain channels. *Journal of the North American Benthological Society* 26(4): 779–796.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289-290.

Pautasso, M., G. Aistara, A. Barnaud, S. Caillon, P. Clouvel, O. Coomes, M. Delêtre et al. 2013. Seed exchange networks for agrobiodiversity conservation. A review. *Agronomy for Sustainable Development* 33(1): 151–175.

Peatrik, A.M. 1999. *La vie à pas contés. Génération, âge et société dans les Hautes Terres du Kenya (Meru Tigania-Igembe).* Nanterre, France: Société d'ethnologie.

Perales, H., B. Benz, and S. Brush. 2005. Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* 102(3): 949–954.

Perrier, X., and J. P. Jacquemoud-Collet. 2006. DARwin software. Montpellier: CIRAD.

Perrier, X., E. De Langhe, M. Donohue, C. Lentfer, L. Vrydaghs, F. Bakry, F. Carreel et al. 2011. Multidisciplinary perspectives on banana (Musa spp.) domestication. *Proceedings of the National Academy of Sciences of the United States of America* 108 (28):11311-11318.

Petit, R.J, A. El Mousadik, O. Pons .1998. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* 12: 844-855.

Pike, K.L. 1954. *Language in Relation to a Unified Theory of the Structure of Human Behavior*. Glendale: Summer Institute of Linguistics.

Piperno, D. R., A. J. Ranere, I Holst, J. Iriarte, and R. Dickau. 2009. Starch grain and phytolith evidence for early ninth millennium BP maize from the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* 106 (13):5019-5024.

Plucknett, D. L., J. P. Evenson, and W. G. Sanford. 1970. Ratoon cropping. *Advances in Agronomy* 22:285-330.

Plucknett, D.L., N.J.H. Smith, J.T. Williams, and N. Murthi Anishetty. 1987. *Gene banks and the world's food*. Princeton, NJ: Princeton University Press.

Pressoir, G., and J. Berthaud. 2004. Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity* 92(2): 95–101.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155 (2):945-959.

Pujol, B., P. David, and D. McKey. 2005. Microevolution in agricultural environments: how a traditional Amerindian farming practice favours heterozygosity in cassava (Manihot esculenta Crantz, Euphorbiaceae). *Ecology Letters* 8 (2):138-147.

Quiros, C.F., R. Ortega, L. Van Raamsdonk, M. Herrera-Montoya, P. Cisneros, E. Schmidt, and S.B. Brush. 1992. Increase of potato genetic resources in their center of diversity: the role of natural outcrossing and selection by the Andean farmer. *Genetic Resources and Crop Evolution* 39 (2):107-113.

Quiros, C. F., S. Brush, D. S. Douches, K. S. Zimmerer, and G. Huestis. 1990. Biochemical and Folk Assessment of Variability of Andean Cultivated Potatoes. *Economic Botany* 44 (2):254–266.

R Development Core Team. 2011. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/.

Rabbi, I.Y., H.H. Geiger, B.I. Haussmann, D. Kiambi, R. Folkertsma, and H.K. Parzies. 2010. Impact of farmers' practices and seed systems on the genetic structure of common sorghum varieties in Kenya and Sudan. *Plant Genetic Resources: Characterization and Utilization* 8 (2):116-126.

Rana, R., C. Garforth, B. Sthapit, and D. Jarvis. 2007. Influence of socio-economic and cultural factors in rice varietal diversity management on-farm in Nepal. *Agriculture and Human Values* 24(4): 461–472.

Raymond, M., F. Rousset. 1995. GENEPOP version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248-249.

Reyes-García, V., J. Broesch, L. Calvet-Mir, N. Fuentes-Peláez, T.W. McDade, S. Parsa, S. Tanner, T. Huanca, W.R Leonard, and M.R. Martínez-Rodríguez. 2009. Cultural transmission of ethnobotanical knowledge and skills: an empirical analysis from an Amerindian society. *Evolution and Human Behavior* 30 (4):274-285.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43: 223-225.

Richerson, P.J., R. Boyd, and R.L. Bettinger. 2001. Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. *American Antiquity* 66 (3):387-411.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. *Heredity* 88 (4):221-228.

Romney, A.K., S.C. Weller, and W.H. Batchelder. 1986. Culture as Consensus: A Theory of Culture and Informant Accuracy. *American Anthropologist* 88.

Roullier, C. 2012. *Aux origines de la diversité de la patate douce (Ipomoea batatas): Une enquête phylogéographique en Amérique tropicale (aire d'origine) et en Océanie (aire d'introduction)*. Université Montpellier II, Montpellier.

Roullier, C., L. Benoit, D.B. McKey, and V. Lebot. 2013. Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. *Proceedings of the National Academy of Sciences of the United States of America* 110 (6):2205-2210.

Sadiki, M., D. Jarvis, D. Rijal, J. Bajracharya, N. Hue, T. Camacho-Villa, L. Burgos-May et al. 2007. Variety names: An entry point to crop genetic diversity and distribution in agroecosystems? In *Managing Biodiversity in Agricultural Ecosystems*, eds. D. Jarvis, C. Padoch, and D. Cooper, 34–76. New York: Colombia University Press.

Saitou, N., and M. Nei. 1987 The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Salick, J., N. Cellinese, and S. Knapp. 1997. Indigenous Diversity of Cassava: Generation, Maintenance, Use and Loss among the Amuesha, Peruvian Upper Amazon. *Economic Botany* 51 (1):6–19.

Sardos, J., D. McKey, M.F. Duval, R. Malapa, J.L. Noyer, and V. Lebot. 2008. Evolution of cassava (Manihot esculenta Crantz) after recent introduction into a South Pacific Island system: the contribution of sex to the diversification of a clonally propagated crop. *Genome* 51 (11):912-921.

Sardos, J., J.L. Noyer, R. Malapa, S. Bouchet, and V. Lebot. 2012. Genetic diversity of taro (*Colocasia esculenta* (L.) Schott) in Vanuatu (Oceania): an appraisal of the distribution of allelic diversity (DAD) with SSR markers. *Genetic Resources and Crop Evolution* 59 (5):805-820.

Saunders, B.A.C. 2007. Towards a new topology of colour. In A*nthropology of color: Interdisciplinary multilevel modeling*, eds. R. E. MacLaury, G.V. Paramei, and D. Dedrick, 467–479. Amsterdam: John Benjamins.

Scarcelli, N., S. Tostain, Y. Vigouroux, C. Agbangla, O. Daïnou, and J. L. Pham. 2006. Farmers’ use of wild relative and sexual reproduction in a vegetatively propagated crop. The case of yam in Benin. *Molecular Ecology* 15 (9):2421-2431.

Schmidt, M., and G. Bothma. 2006. Risk assessment for transgenic sorghum in Africa: Crop-to-crop gene flow in *Sorghum bicolor* (L.) Moench. *Crop Science* 46 (2):790-798.

Smith, B.D. 2011. A cultural niche construction theory of initial domestication. *Biological Theory* 6 (3):260-271.

Snowden, J.D. 1936. *The cultivated races of sorghum*. London: Adlard and Son.

Sokal, R., and C.D. Michener. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38:1409-1438.

Soler, C., A.A. Saidou, T.V.C. Hamadou, M. Pautasso, J. Wencelius, and H.H. Joly. 2013. Correspondence between genetic structure and farmers' taxonomy - a case study from dry-season sorghum landraces in northern Cameroon. *Plant Genetic Resources: Characterization and Utilization* 11 (1):36-49.

Stemler, A.B., J.R. Harlan, and J.M. de Wet. 1975. Caudatum sorghums and speakers of Chari-Nile languages in Africa. *Journal of African History* 16 (2):161-183.

Stemler, A.B.L. 1980. Origins of plant domestication in the Sahara and the Nile Valley. In *The Sahara and the Nile*, eds. M.A.J Williams, and F. Faure, 503-526. Balkema.

Tanno, K.I., and G. Willcox. 2006. How fast was wild wheat domesticated? *Science* 311 (5769):1886-1886.

Teshome, A., B.R. Baum, L. Fahrig, J.K. Torrance, T.J. Arnason, and J.D. Lambert. 1997. Sorghum [*Sorghum bicolor* (L.) Moench] landrace variation and classification in north Shewa and south Welo, Ethiopia. *Euphytica* 97 (3):255–263.

Therneau, T., B. Atkinson, B. Ripley, and 2012. rpart: Recursive Partitioning: R package version

Thrupp, L. 2000. Linking agricultural biodiversity and food security: the valuable role of agrobiodiversity for sustainable agriculture. *International Affairs* 76(2): 265–281.

Van den Broeck, K., and S. Dercon. 2011. Information flows and social externalities in a Tanzanian banana growing village. *The Journal of Development Studies* 47(2): 231–252.

vom Brocke, K., A. Christinck, R.E. Weltzien, T. Presterl, and H.H. Geiger. 2003. Farmers' seed systems and management practices determine pearl millet genetic diversity patterns in semiarid regions of India. *Crop Science* 43 (5):1680-1689.

vom Brocke, K., G. Trouche, E. Weltzien, C.P. Barro-Kondombo, E. Gozé, and J. Chantereau. 2010. Participatory variety development for sorghum in Burkina Faso: Farmers’ selection and farmers’ criteria. *Field Crops Research* 119 (1):183-194.

Wayne Smith, C., and R.A. Frederiksen. 2000. *Sorghum: origin, history, technology, and production*. New-York: Wiley.

Weir, B.S., Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Wells, W.C., W.H. Isom, and J.G. Waines. 1988. Outcrossing rates of six common bean lines. *Crop science* 28 (1):177-178.

Wendorf, F., A.E. Close, R. Schild, K. Wasylikowa, R.A. Housley, J.R. Harlan, and H. Królik. 1992. Saharan exploitation of plants 8,000 years BP. *Nature* 359 (6397):721-724.

Wilkes, G. 1977. The world's crop plant germplasm: An endangered resource. *Bulletin of the Atomic Scientists* 33 (2):8-16.

Willcox, G., S. Fornite, and L. Herveux. 2007. Early Holocene cultivation before domestication in northern Syria. *Vegetation History and Archaeobotany* 17 (3):313-325.

Zimmerer, K.S. 2003. Geographies of seed networks for food plants (potato, ulluco) and approaches to agrobiodiversity conservation in the Andean countries. *Society &Natural Resources* 16 (7):583-601.

Zohary, D. 1996. The mode of domestication of the founder crops of Southwest Asian agriculture. In *The origins and spread of agriculture and pastoralism in Eurasia*, ed. D.R. Harris, 142–158. London: University College London Press.

Zohary, D. 2004. Unconscious selection and the evolution of domesticated plants. *Economic Botany* 58 (1):5-10.

ANNEXES

ANNEXE 1: ESTIMATION DU TAUX D’ALLOGAMIE

Material and Methods

One panicle was collected on six plants that were previously genotyped. These six plants were selected to represent the four main genetic groups previously identified in that population, corresponding to: i. a recently introduced improved variety (*Gadam*), ii. an anciently introduced improved variety (*Kaguru*), iii. a local variety presenting short growing cycle, iv. a local variety presenting a long growing cycle (Labeyrie et al., submitted). To maximize the probability of detecting gene flows, we selected individuals that were sown in the experimental field in-between half-sib lines belonging to a different genetic group from their own. The seeds of these six panicles where then sown in a greenhouse and the leaves of 22 descendants per panicles were collected, lyophilized and DNA was extracted as described previously. The 16 most polymorphic SSR loci were selected for this study out of the 18 used in previous study (Labeyrie et al., submitted).

As we had doubts concerning the performance of the calculations implemented in MLTR software (Ritland 2002) for such small samples presenting few heterozygous loci, we used a simpler and intuitive measure of outcrossing. For each loci which was found heterozygous in the maternal genotype each, we thus simply measured the significance of the departure from the ½, ¼, ¼ proportions for heterozygous and homozygous individuals in the progenies using a conformity Chi-squared test. We also scored at each locus for each progeny the number of descendants presenting alleles that were not found in the maternal genotype.

Results & discussion

Out of the six progenies, five showed no out-crossing events as the observed frequencies of homozygous and heterozygous individuals for each loci were not significantly different from that expected in case of full selfing. For only one progeny, two descendants out of 22 (9%) were found to result from out-crossing events with neighboring plants. The mother of this progeny was genetically assigned to the gene-pool of short-cycle local varieties. We detected pollen flows from two other plants, one of which was the direct neighbor of the mother-plant in the experimental field, and the other located at about 3 meters from it. Both plants belonged to the local short-cycle genetic pool. No pollen flows were detected with the direct neighboring plant which was assigned to the long-cycle local genetic pool.

The individual belonging to the long-cycle landraces genetic pool, and to improved varieties groups showed a null outcrossing rate. However, results for only one individual cannot help to conclude on the mating system of these varieties. Some gene flows appeared to occur between local short-cycle varieties, but they seemed limited. The observations of farmers concerning “out of type” panicles, and the important part of admixture between *Kaguru* and the local gene pool comfort the existence of gene flows in farmers’ fields. In the case of short-cycle landraces, the existence of gene flows can explain the loose genetic structure within that genetic pool, and the absence of structure regarding variety names. This suggests that the selection of farmers for perceptual distinctiveness is central for the maintenance of varieties phenotypic characteristics.

The estimation of outcrossing done in this survey did not aim at measuring precisely the variability of mating system within and between the sorghum varieties. The objective was rather to get a rough estimate of gene flows within farmers’ fields for the main genetic groups identified in the population of study. These values of outcrossing rates we estimated are in the range of outcrossing rates estimated in previous studies, notably in that of Barnaud et al. (Barnaud et al. 2008) who found outcrossing rates comprised between 0 % and 73 %.

The strength of our approach as compared to previous studies is that we scored the maternal genotypes, which gives more power to the analysis. Method used in most studies (MLTR) were found to be inadequate for our data and raised concern on their adequacy for small samples size. We thus preferred to compare the allelic frequencies observed in the progenies with those expected under full selfing. The main caveat of this approach is that no correction was applied for biparental inbreeding, which is potentially high in our experimental design because we arranged the half-sib in lines instead of randomly. Hence, the maternal individuals were surrounded by 9 half-sibs, which have probably contributed largely to the pollen cloud. Our results were insufficient to compare the different genetic groups because we only considered a limited number of progenies. In addition, strong variations of the outcrossing rates can be observed between cropping seasons (Abdel-Ghani et al. 2004), measures based on a unique cropping season are thus insufficient to build a good estimation.

ANNEXE 2: CHEVAUCHEMENT DES DATES DE FLORAISON

Material and Methods

The dates of seedling emergence and 50 % flowering were observed for 66 half-sib progenies sown in the experimental field. Observations were done on half-sib lines with 10 individuals per progeny per bloc, each progeny being replicated in three blocs. Progenies lines were randomly arranged in each bloc. The duration between the 50 % flowering date and the emergence date was then calculated.

To estimate the overlay of flowering period of the different varieties, the mean and standard deviation of the duration between emergence and 50 % flowering was calculated over the three blocs for each half-sib lines. Mean and standard deviation were then calculated over the half-sib lines belonging to the same genetic group defined in Labeyrie et al. (submitted).

Results & discussion

The number of day between emergence and 50 % flowering varied between 48.3 and 70.4 days between the half-sib lines. Most of the half-sib lines flowered before 55 days, and 9 flowered after 56 days. Individuals assigned to the genetic groups A, B and C showed concomitant flowering period, while the 14 individuals assigned to the cluster D, corresponding to long-cycle landraces, flowered much later (68.5 days in average). Later flowering dates were notably observed for the varieties *Mukumbu* (56), *Mucuri* (59) *and Mugana* (65), *Kathirigwa* being much later-flowering than all the other (70 days).

Considering a mean flowering duration of seven days (Chantereau and Nicou 1991), the flowering period of some long-cycle landraces potentially overlays that of some short-cycle landraces. However, these differences in flowering period potentially limit the gene flows between the long-cycle varieties and the short-cycle ones, and also within part of the long-cycle varieties which showed some lag. This can contribute to the maintenance of a clear morphological and genetic distinction of the different long-cycle landraces. The short-cycle varieties from both improved and local origins, on the other hand, had concomitant flowering periods.

Growing conditions on-farm can considerably influence the flowering period, and the estimation done in this survey was just a baseline which shall be considered with care.

Moreover, the estimation of 50 % flowering can be difficult to score and the date of flag-leaf emergence is often preferred for a better estimate.

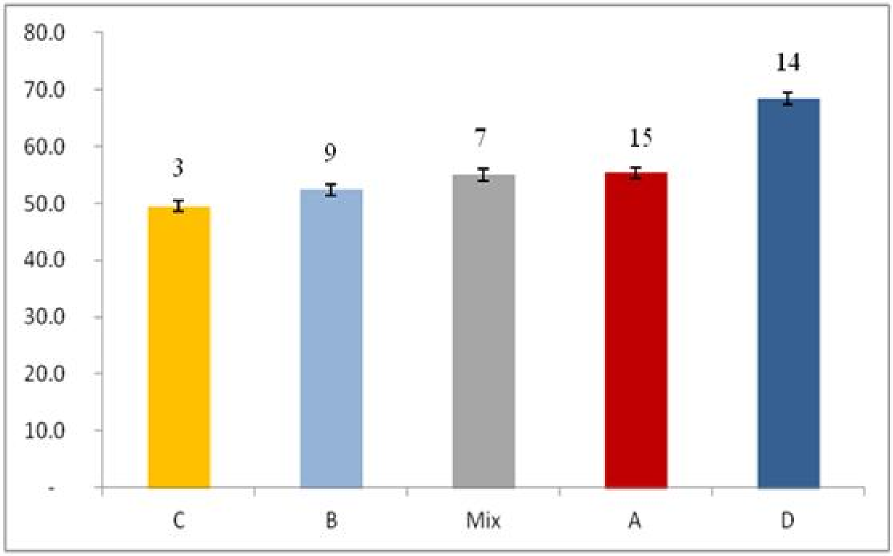


Figure A2-1. Mean duration (number of days) of the emergence - 50 % flowering spell for individuals assigned to each genetic group (K = 4, q > 0.8). Error bars represent the standard deviation, number above the bars represent the number of individuals

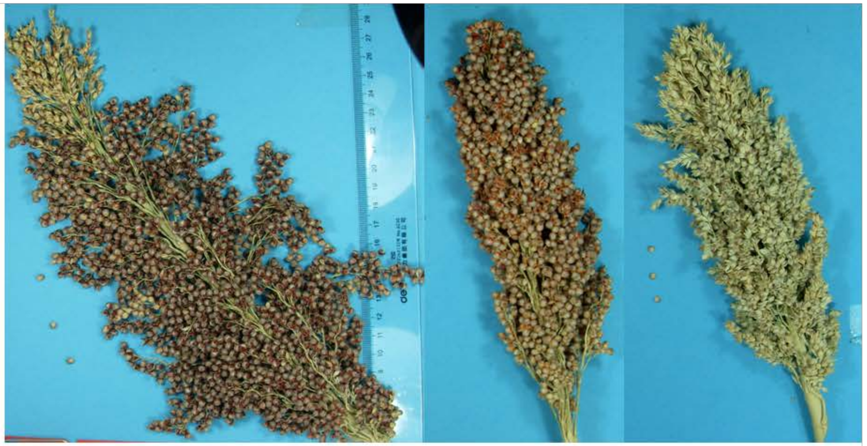
ANNEXE 3: PHOTOS DES VARIETES

Planche 1: Variétés à cycle court

Planche 2 : Variétés à cycle long

Planche 3 : Variabilité morphologique des variétés Mbura-imwe et Ngirigacha

Mugeta



Muruge

Gadam

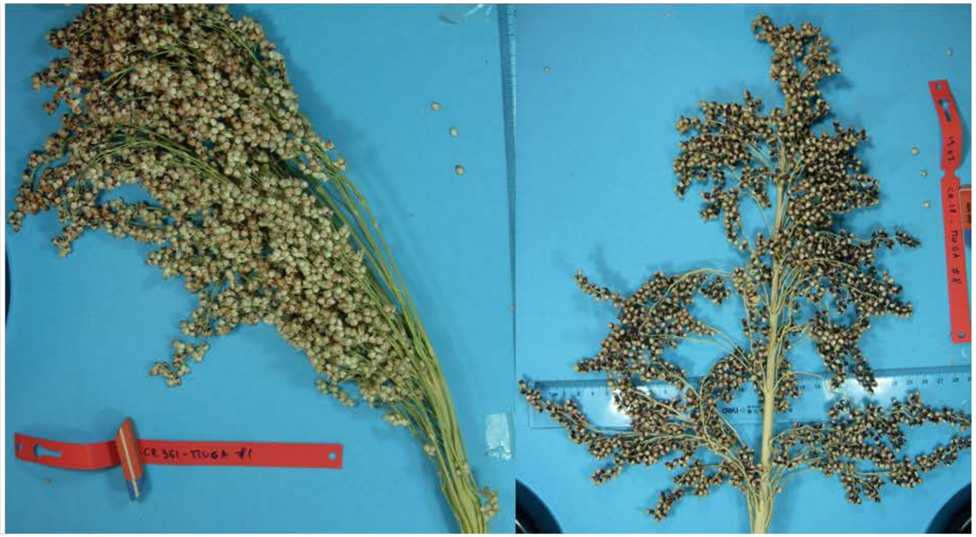
Serendo

Kaguru

Mbunge

Mukumbu

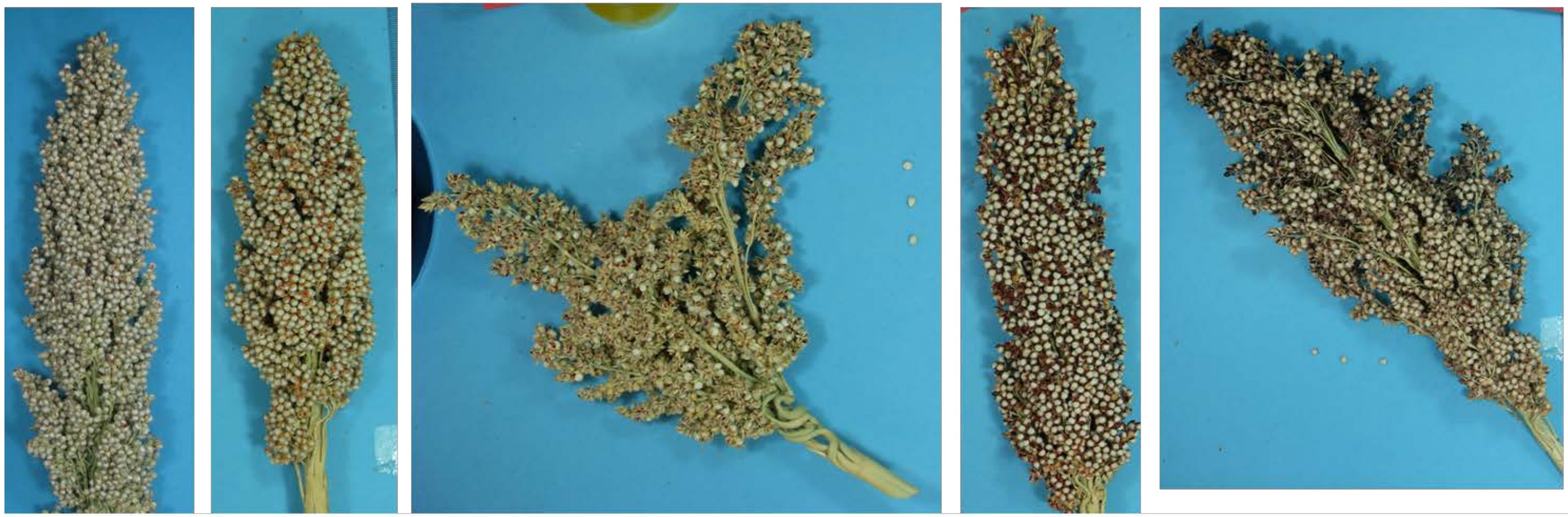
Kathirigwa



Mugana

Muruge mbura ciiri

Mucuri



Mbura-imwe

Ngirigacha

Abstract

Crop genetic resources are elaborated by multiple environmental factors *in situ*, among which the human action plays a major role. Seed diffusion and selection are the main human practices that influence crop evolution on farm. In addition, in most small-scale farming systems, farmers' practices of seed exchanges and selection are not independent. Indeed, the individual practices are considerably influenced by the organization of societies that shapes seed and knowledge diffusion.

This thesis addresses the mechanisms through which the organization of societies shapes crop diversity *in situ*. We combined populations' biology with social and cultural anthropology approaches to analyze the relationship between the organization of Chuka, Tharaka and Mbeere societies and the patterns of sorghum diversity (*Sorghum bicolor* [L.] Moench) on the eastern slope of Mont Kenya.

This study first describes the patterns of sorghum diversity with regards to the organization of societies. We show that ethnolinguistic organization shapes the distribution of sorghum diversity as perceived by farmers (the varieties they name), and as characterized by us using neutral genetic markers and phenotypic descriptors. Nevertheless, these approaches of diversity have evidenced some divergence, notably because the local varieties were not distinct and uniform genetic units. This observation raised questions concerning the correspondence between farmers' taxonomy and the structure of genetic and phenotypic diversity.

This led us to test whether farmers belonging to a same ethnolinguistic group identify, name and classify varieties consistently regarding their phenotypic characteristics, and whether these local taxonomies differ among groups. These analyses showed that the organization of societies impacts the diffusion of local taxonomy.

Examination of seed exchange networks finally helped to understand this latter relationship by showing that seed, and the knowledge probably associated, are exchanged mainly between people from the same ethnolinguistic group.

We further discuss the effect of differences in the way farmers represent sorghum diversity to themselves, reflected by their taxonomies, on their seed selection practices.

This study showed that the organization of societies shapes seed and knowledge exchanges, hence influencing farmers' individual practices of diversity management. This pluridisciplinary approach hence gives new perspectives for the characterization, the conservation and the improvement of crop genetic resources

Keywords: **farmers' varieties, genetic resources, social organization, folk taxonomy, seed exchanges, social and cultural anthropology, cultural diversity, sorghum, Kenya.**

Résumé

La diversité des ressources génétiques est façonnée par de multiples facteurs *in situ*, parmi lesquels l'action anthropique joue un rôle majeur. La diffusion et la sélection des semences sont les principales pratiques humaines qui influencent l'évolution des plantes cultivées. Au sein des systèmes agricoles familiaux, les pratiques individuelles d'échange et de sélection ne sont pas indépendantes entre elles. En effet, l'organisation des sociétés a un impact considérable sur les pratiques individuelles des agriculteurs car elle influence la diffusion des semences et du savoir qui y est associé.

L'objectif de cette thèse est d'améliorer notre compréhension des mécanismes par lesquels l'organisation des sociétés humaines influence l'organisation de la diversité des plantes cultivées *in situ*. Cette étude combine pour cela des approches de biologie des populations et d'anthropologie sociale et culturelle pour analyser la relation entre l'organisation des sociétés Chuka, Tharaka et Mbeere et l'organisation de la diversité du sorgho (*Sorghum bicolor* [L.] Moench) sur le versant est du mont Kenya.

Dans un premier temps, cette étude décrit la distribution de la diversité du sorgho au regard de l'organisation des sociétés. Nous montrons que l'organisation ethnolinguistique structure la distribution de la diversité telle qu'elle est perçue par les agriculteurs (les variétés nommées), et telle que nous l'avons caractérisée à l'aide de marqueurs génétiques et phénotypiques. Cependant, ces deux approches de la diversité ne coïncident pas parfaitement, notamment car certaines variétés locales distinguées par les agriculteurs ne correspondent pas à des unités génétiques distinctes et homogènes. Ces résultats nous amènent à nous interroger sur la cohérence des taxonomies locales vis-à-vis de la structure de la diversité génétique et phénotypique.

Dans un second temps, nous testons donc si les agriculteurs appartenant à un même groupe ethnolinguistique identifient, nomment et classent les variétés de manière similaire au regard de leurs caractéristiques phénotypiques, et si les taxonomies paysannes diffèrent entre les groupes. Ces analyses montrent que l'organisation des sociétés a un impact sur la diffusion des systèmes de taxonomies locaux. L'examen des réseaux d'échanges de semences permet de comprendre cette relation en montrant que les échanges de semences, et vraisemblablement du savoir qui y est associé, se font principalement entre personnes de même groupe ethnique.

L'effet des différences de représentations, traduites par les taxonomies, sur les pratiques de sélection des semences est enfin discuté.

Cette thèse montre donc que l'organisation des sociétés structure la diffusion des semences et du savoir, et qu'elle influence ainsi les pratiques individuelles de gestion de la diversité des agriculteurs. Cette approche pluridisciplinaire ouvre de nouvelles perspectives pour la caractérisation, la conservation et l'amélioration des ressources génétiques.

Mots clés: **variétés paysannes, ressources génétiques, organisation sociale, taxonomies locales, échanges de semences, anthropologie sociale et culturelle, diversité culturelle, sorgho, Kenya.**