

**Le berbère et les Berbères :
Diversité linguistique et génétique**

DUGOUJON Jean-Michel

Action " Origine de l'Homme, du Langage et des Langues "

A. FICHE ADMINISTRATIVE

Titre du projet :

Le berbère et les Berbères : Diversité linguistique et génétique.....

Mots-clés :

Berbère - Marqueurs génétiques – Afrique du Nord – Langues- Diversité humaine

Résumé du projet (10 lignes maximum) :

Notre projet a pour but d'étudier la diversité génétique des populations Berbères en relation avec les différentes langues parlées par ces populations réparties dans toute l'Afrique du Nord, de la Mauritanie à l'oasis de Siwa (Egypte) en passant par le Sahara (Touareg) et l'Atlas (Maroc). De plus, nous proposons un ancrage dans le temps avec l'étude d'ADN fossile. Notre travail collectif s'inscrit ainsi dans le contexte plus large du peuplement de l'Afrique du Nord, de la préhistoire à l'époque moderne.

Avec l'étude de marqueurs génétiques (allotypes Gm des Ig, ADN mitochondrial, chromosomes X et Y, autosomes permettant la construction d'haplotypes), nous mesurerons quels sont les degrés de parenté linguistique et génétique entre les différentes populations Berbères, puis avec les populations voisines appartenant à des familles linguistiques différentes.

1. Responsable scientifique du projet

Nom : ..DUGOUJON.....

Prénom : ..Jean-Michel.....

Grade : ..CR1

Discipline du responsable scientifique: ..Anthropologie génétique - Génétique des populations

Établissement de rattachement : ..CNRS.....

Adresse professionnelle : ..39, allées Jules Guesde.....

Code postal : 31000 Commune : ..TOULOUSE.....

Tél : 05 61 14 59 80

Fax : 05 61 14 59 79

E-Mail : dugoujon@cict.fr.....

2. Laboratoire ou organisme de rattachement de l'équipe de recherche

IntituléCentre d'Anthropologie.....

Type de formation (*cocher la case utile*)

- Unités CNRS : unité propre du CNRS • unité associée ou mixte du CNRS

Préciser le code unité : UMR 8555

Préciser la délégation régionale : .Midi-Pyrénées.....

- Unités hors CNRS : unité universitaire (*Préciser l'université*)

Autre (*Préciser : EPST, Association, équipe étrangère...*)

Nom du directeur de l'organisme : ALBERT Jean-Pierre.....

Adresse : 39, allées Jules Guesde.....

Code postal : 31000 Commune : TOULOUSE

Tél 05 61 55 80 70

Fax 05 61 14 59 79

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : Dynamique du Langage, UMR 5596.....

Discipline(s) couverte(s) par l'équipe : .Linguistique.....

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : LOUALI-RAYNALPrénom : Naïma.....

Grade : CR1E-Mail : Naima.Louali@ish-lyon.cnrs.fr

Adresse : ISH, 14 avenue Berthelot.....

Code postal 69363 Commune .Lyon cedex 07.....

Tél 33 (0)4 72 72 64 93

Fax 33 (0)4 72 72 65 90

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : Dept de Biologie Animale / Anthropologie; Faculté de Biologie; Université de Barcelone

Discipline(s) couverte(s) par l'équipe : .Anthropologie biologique

Nom du directeur de l'équipe : .MORAL Pedro.....

Membre(s) partenaire(s) :

Nom : MORAL Prénom : Pedro.....
Grade : Professeur.....E-Mail : pmoral@porthos.bio.ub.es

Nom : JUNYENT Prénom : Carmen.....
Grade : Professeur (linguiste).....E-Mail :

Nom : HARICH Prénom : Nourdin.....
Grade : Professeur à l'Université d'El Jadida (Maroc) E-Mail :

Adresse : Av. Diagonal, 645.....
.....

Code postal : 08028 Commune :BARCELONE Espagne.....

Tél : 34 93 402 14 59

Fax : 34 93 403 57 40

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : ADN Mitochondrial et Populations

Discipline(s) couverte(s) par l'équipe : ADN mitochondrial, ADN ancien, Génétique des populations humaines anciennes et actuelles

Nom du directeur de l'équipe :.BERAUD-COLOMB Eliane

Membre(s) partenaire(s) :

Nom : BERAUD-COLOMBPrénom : Eliane
Grade : CR1 INSERM.....E-Mail : beraud@luminy.univ-mrs.fr

Nom : SPADONIPrénom : Jean-Louis
Grade : Doctorant..... E-Mail : Spadoni@luminy.univ-mrs.fr

Nom : STEVANOVITCHPrénom : Alain
Grade : DoctorantE-Mail : Steva@luminy.univ-mrs.fr

Nom : PARIS *Prénom : François
Grade : DR IRDE-Mail : paris@ird.fr

Adresse : Faculté des Sciences de Luminy, Case 901.....

163, Av de Luminy.....

Code postal : 13288 Commune :MARSEILLE cedex 09.....

Tél : 04 91 82 97 46

Fax : 04 91 82 93 09

*Adresse : Institut français d'archéologie orientale, c/o Service de la Valise Diplomatique.....
128 bis, rue de l'Université

Code postal : 75351 Commune :PARIS cedex 07.....

Et : Représentation de l'IRD ; PO Box 26 12211 GIZA République Arabe d'Egypte

Tel : (202) 362 05 30 Fax : (202) 262 24 49

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : INSERM U535 Génétique épidémiologique et Structure des populations humaines
Discipline(s) couverte(s) par l'équipe : Phylogénétique
Nom du directeur de l'équipe : DARLU Pierre

Membre(s) partenaire(s) :

Nom : DARLUPrénom : Pierre
Grade : DR2 CNRSE-Mail : darlu@kb.inserm.fr

Nom : HAZOUT *Prénom : Serge
Grade : ProfesseurE-Mail : hazout@urb.jussieu.fr

Nom : DUPANLOUP**Prénom : Isabelle
Grade : Post-DoctoranteE-Mail : isabelle.dupanloup@anthro.unige.ch

* Adresse : Equipe de Bioinformatique Moléculaire et Génomique ; INSERM U436 ; Université de Paris 7 ; 2, place Jussieu, 75251 PARIS cedex 05

** Adresse : Université de Genève ; Département d'anthropologie et d'Ecologie ; 12, rue Gustave Revilliod, 1227 CAROUGE – GENEVE, Suisse

Adresse : Batiment Grégory Pincus.....
80, rue du Général Leclerc
Code postal : 94276 Commune : KREMBLIN BICETRE cedex.....

Tél : 01 49 59 53 40

Fax : 01 49 59 53 31

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : Centre d'Anthropologie, UMR 8555
Nom du directeur de l'équipe : CRUBEZY Eric

Membre(s) partenaire(s) :

Nom : CRUBEZYPrénom : Eric
Grade : ProfesseurE-Mail : crubezyeric@wanadoo.fr

Nom : LUDESPrénom : Bertrand
Grade : ProfesseurE-Mail : Bertrand.Ludes@iml-ulp.u-strasbg.fr

Nom : SEVINPrénom : André
Grade : IR1E-Mail : sevin@cict.fr

Nom : CLISSONPrénom : Isabelle
Grade : DoctoranteE-Mail :

Nom : BEN NCER*Prénom : Abdel
Grade : ConsevateurE-Mail : abdelbenncer@caramail.com

Nom : LACOMBE**Prénom : Jean-Paul

Grade : ChercheurE-Mail :

* Adresse : Musée de l'Archéologie, 23, rue Al-Brihi, BP 503 RABAT-CHELLAH, Maroc

** Adresse : UMR 5809, Université de Bordeaux 1 ; Mission littorale franco-marocaine

Adresse : 39, allées Jules Guesde.....

Code postal : 31000 Commune : TOULOUSE.....

Tél : 05 61 55 80 65

Fax : 05 61 14 59 79

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : Laboratoire d'Ecologie Humaine, Université Cadi Ayyad

Nom du directeur de l'équipe : BAALI Abddelatif

Membre(s) partenaire(s) :

Nom : BAALIPrénom : Abdellatif

Grade : ProfesseurE-Mail : baali@ucam.ac.ma

Nom : CHERKAOUIPrénom : Mohamed

Grade : Enseignant-Chercheur.....E-Mail :

Nom : HILALIPrénom : Mohamed Kamal

Grade : Enseignant-Chercheur.....E-Mail :

Nom : AMORPrénom : Akima

Grade : Enseignant-ChercheurE-Mail :

Nom : LOUKIDPrénom : Mohamed

Grade : Enseignant-ChercheurE-Mail :

Adresse : Bd My Abdellah, BP2390.....

Code postal : 40000 Commune : MARRAKECH (Maroc).....

Tél : (212) 4 43 46 49 p 522

Fax : (212) 4 43 74 12

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : Institute of Biological Anthropology

Nom du directeur de l'équipe : WARD Ryk

Membre(s) partenaire(s) :

Nom : BOUZEKRIPrénom : Nourdine

Grade : Post-Doctorant.....E-Mail : nourdine.bouzekri@bioanth.ox.ac.uk

Nom : WARDPrénom : Ryk

Grade : ProfesseurE-Mail : ryk.ward@bioanth.ox.ac.uk

Adresse : Zoology Building, South Parks Road.....
Code postal : OX1 3PS Commune : OXFORD (Angleterre).....

Tél : 44 (0) 1865281905 Fax : 44 (0) 1865274699

Date : Signature du Responsable Scientifique :

3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé : Centre National de Génotypage
Nom du directeur de l'équipe : LATHROP M. (Directeur Général)

Membre(s) partenaire(s) :

Nom : LathropPrénom : Marc
Grade : ProfesseurE-Mail : christine.lestringuez@cng.fr

Adresse : Bat G2, 2, rue Gaston Crémieux, CP 5721
Code postal : 91057 cédex Commune : EVRY

Tél 01 60 87 25 00 Fax : 01 60 87 84 85
Contact : C. Lestringuez

B. PROJET SCIENTIFIQUE

I - Les objectifs du projet

Notre projet a pour but d'étudier la diversité génétique des populations berbères en relation avec les différentes langues parlées par ces populations réparties dans toute l'Afrique du Nord, de la Mauritanie à l'oasis de Siwa (Egypte) en passant par le Sahara (Touareg) et l'Atlas où elles sont particulièrement nombreuses (Chamla, 1971 ; Camps, 1974). Les degrés de diversité mis en évidence n'auront de signification que si nous arrivons à proposer un ancrage dans le temps. Notre travail collectif s'inscrit ainsi dans le contexte plus large du peuplement de l'Afrique du Nord, de la préhistoire à l'époque moderne.

En s'appuyant tout d'abord sur les subdivisions linguistiques et anthropologiques les plus marquées (les deux groupes : Berbère du nord / Touareg et un troisième en marge du domaine géographique linguistique : le Siwi (Siwa, en Égypte) et le Zénaga (en Mauritanie), nous aborderons à travers l'alliance des trois disciplines (la linguistique, l'anthropologie et la génétique) la question de la proximité linguistique et génétique et de la justification d'une classification dialectale dans la langue Berbère. De même, compte tenu du nombre des populations étudiées, nous aborderons dans une même aire géographique ou linguistique des études de diversité intra et inter populations. C'est à dire quels sont les degrés de parenté linguistique et génétique entre les différentes populations Berbères ? Quelle est la variabilité génétique et linguistique au sein d'une même population ? Quelles relations génétiques avec les populations voisines ?

II - La langue berbère

Le berbère appartient à la famille dite Chamito-sémitique ou Afro-asiatique. La zone d'origine de cette famille de langues est très controversée avec un berceau primitif au Moyen Orient (Syrie-Palestine ou péninsule arabe) ou en Afrique (Afrique de l'Est) ou Sahara central) La

langue berbère est actuellement en usage dans neuf pays sur le continent Africain : Mauritanie, Maroc, Algérie, Tunisie, Libye, Égypte, Niger, Mali, Burkina-Fasso. Le Maroc et l'Algérie abritent le pourcentage le plus élevé de berbérophones, respectivement: 30 à 40% (28 millions) et 20 à 25% (30 millions). Les populations touarègues représentent le troisième groupe important avec ± 900000 individus que se partagent les pays du Sahel (Niger et Mali) et les pays du Sahara (Algérie et Libye). Pour le Burkina-Fasso, il s'agit d'une implantation récente. Elle remonte aux années 70 et elle est liée à la sécheresse et aux conflits avec les militaires du Mali et du Niger. Il y a des isolats qui se situent sur les marges du domaine de répartition où le berbère est voué à l'extinction : (i) à très court terme comme en Tunisie où l'on évalue les berbérophones à 60 000 individus répartis entre l'île de Djerba et environ cinq villages au sud du pays ; (ii) à moyen terme comme en Mauritanie (± 10 000 individus) et à l'oasis de Siwa en Égypte (± 10 000 individus) : sur ces deux points extrêmes la transmission de la langue se fait encore (le berbère y est préservé pour quelques années grâce à l'éloignement d'un pouvoir central qui prône l'unité arabe et islamique) et (iii) à long terme comme au nord Libye (Zouara et Djebel Nefoussa) où le sentiment d'appartenance berbère est vécu de manière positive et l'on manifeste de la résistance à l'égard de l'arabe.

La classification interne du berbère fait apparaître un ensemble de dialectes : le kabyle, le zénaga, le touareg, etc... Prasse (1972) considère toutefois le touareg comme une langue plutôt que comme un dialecte. Il fonde cette distinction sur la base du critère d'intercompréhension entre le touareg et le berbère du nord. Pour s'identifier, un berbérophone a recours d'abord à un micro-repère qui est sa tribu, ou sa région, et ensuite à un macro-repère qui varie selon qu'il parle le berbère du nord : amazigh, ou le berbère du sud (le touareg) : amhegh (au sud de l' Algérie et au sud de la Libye), amajegh (au Niger) et amchigh (au Mali). La langue berbère est constituée d'une multitude d'idiomes, dont un nombre important demeure inconnu. R. Basset (1952) propose une classification à partir de "critères phonétiques". C'est ainsi qu'il distingue trois groupes : les dialectes forts, faibles et intermédiaires. Il est évident que les notions de "fort/faible" ne correspondent pas à une réalité linguistique. Faut-il les disloquer en plusieurs groupes ? En effet, la question reste posée puisque la classification proposée par Willms (1980) est à certains égards une classification géographique malgré son caractère novateur. Il propose une classification en quatre groupes : (i) Chleuh et Beraber du sud, Zayan et Zemmour, Aït Ndir ; (ii) Beni Iznassen, Senhaja, Beni Snouss et Rif ; (iii) Mzab, Ouargla, Nefoussa, Chaouïa, Siwa, Touareg et (iv) Kabyle. Cependant, le classement ne couvre pas tout le domaine berbère (omission des parlers zenaga et des parlers du sud tunisien). Le traitement du groupe touareg comme un bloc relève d'une classification géographique : les parlers touaregs sont très diversifiés et ne se limitent pas aux parlers du Hoggar (Algérie) (Louali, 1992). La considération du kabyle comme une "langue" n'est pas justifiée. Ameur (1990) discute d'une classification fondée sur des bases ethniques et conclut par une énumération des critères à retenir dans une classification (socio-linguistique, phonologique, syntaxique, morphologique et lexicale. Enfin, Aikhenvald (1988) propose de classer les parlers berbères en quatre groupes sur la base de critères morphologiques (les pronoms personnels, le déterminant et le mode) et syntaxiques (l'ordre des constituants de la phrase et les marques morphologiques du sujet et de l'objet direct) : (i) le berbère ouest (Zénaga) ; (ii) le berbère sud (Touareg) ; (iii) le berbère est (Libye, Égypte) et (iv) le berbère nord (les parlers du Maroc, d'Algérie, de Tunisie et de Libye (Zouara). Au terme de ce bref rappel, on s'aperçoit que le problème est loin d'être résolu et qu'aucune de ces classifications n'est complète (le critère phonologique n'est pas pris en compte ou elles ne couvrent pas tout le domaine).

C'est pour cela que dans le cadre de ce projet, nous avons prévu l'organisation à Lyon d'un colloque international : *Internal & External Classification of Berber languages*. L'objectif est de faire le point sur la classification linguistique interne et externe. Ce sera aussi un forum de

discussion ouvert sur les autres disciplines et l'occasion pour tisser des liens entre chercheurs linguistes, anthropologues ou généticiens. La participation de ces derniers aux débats des linguistes permettra de coordonner les approches et de dégager quelques "notions" qui feront consensus et contribueront à mettre au point un "langage commun".

III - Les populations et les méthodes

A - Les populations

Les populations Berbères

Région du nord

I- *Les populations de l'Atlas du Maroc :*

Populations Shilha (langage Tamazight) de l'Atlas central.

Populations Tachelhit du sud Atlas.

Nous disposons dans notre laboratoire à Toulouse de 130 échantillons prélevés à Amizmiz (Maroc). A Barcelone, l'équipe de P. Moral propose aussi 140 Berbères du moyen Atlas répartis en Zaiane (n = 86), Ichkern (n = 15), Ait Sgougou (n = 10), Beni M'Guild (n = 9) et divers (n = 20) (Cela est très intéressant pour notre projet, car nous pouvons ainsi commencer les analyses génétiques sur 270 échantillons).

II- *Les Zenati*

Ils se répartissent entre la Tunisie, la Lybie, l'Algérie et le Maroc. Nous nous intéresserons principalement à 3 groupes : un groupe à l'est (à l'île de Djerba en Tunisie et Zuara dans le nord de la Lybie) ; un 2^{ème} groupe de Ghomora (nord de la région de Tamazight) et le 3^{ème} avec des sujets de la région de Ghardaia (Mzab) d'Algérie. Nous avons pour ces derniers plus de 100 échantillons qui pourraient être inclus dans l'étude (après accord).

Région de l'est

I- *Les Sokna* (région de Tripoli en Lybie)

II- *Les Awjila* (est de la Lybie)

III- *Siwa* (dans l'oasis de Siwa en Egypte et des villages où l'on parle le Siwa)

Région de l'ouest

Les Zenaga (côte atlantique du sud de la Mauritanie)

Les Touaregs

Ce sont des nomades qui vivent dans plusieurs pays du Sahara (Algérie, Burkina-Faso, Lybie, Mali, Niger et Nigeria)

Les populations Tchadique

Nous souhaitons faire des prélèvements sur des populations parlant un langage d'origine Afro-Asiatique (à proximité des populations berbères), en particulier la sous-famille Tchadique. Nous avons choisi les *Hausa*. Ils se répartissent entre le Nigéria, le Niger, le Benin, le Burkina-Faso et le Cameroun.

Les populations des phylums Nilo-Saharien et Niger-Kordofanien

L'accès à ces populations nous permettra d'étudier les relations entre les populations Berbères et sub-Sahariennes. Trois groupes seront analysés :

I - Nilo-Saharienne

Les *Sara* : ils vivent principalement en RCA, mais aussi au Tchad, Cameroun et Nigeria

II – Niger-Kordofanien

Les *Peuls* : Mali, Benin, Burkina-Faso, Tchad, RCA et Niger

Les *Mande* (Bobo Madare): nord du Mali, Sénégal, Burkina-Fasso, Cameroun et Côte d'Ivoire

Les populations Sémitiques

Quelles sont les affinités entre les populations Berbères et Sémitiques ? Est-ce que les Bédouins et les Berbères, qui partagent la même région géographique ont aussi la même structure génétique ? Quel est le poids de la proximité géographique et de l'affinité linguistique ? Les Bedouins vivent dans plusieurs régions du Sahara. Nous souhaitons étudier différentes populations géographiquement éloignées (de l'est : Egypte et de l'ouest : Maroc)

Environ 100 échantillons seront prélevés par population (voir tableau en annexe). Nous prioriserons les hommes pour avoir la meilleure information apportée par les données de l'ADN mt, du chromosome Y, des haplotypes sur le X et les autosomes.

B - L'analyse des données

A partir des données génétiques et linguistiques, on peut reconstruire un scénario de l'histoire génétique des différentes populations berbères et des populations voisines apparentées, ainsi que l'histoire évolutive des langues parlées dans ces régions. Les questions auxquelles on se propose de répondre sont les suivantes :

i) La représentation en arbre des données génétiques et linguistiques est-elle significativement plus informative qu'une représentation euclidienne (ou de degré supérieur à 2) (Guénoche et Garreta., 2000) ? Si ce n'est pas le cas, la ressemblance génétique ou linguistique entre populations serait simplement dépendante de la distance géographique qui les séparent.

ii) Dans cette hypothèse, existe-t-il des zones de discontinuités ? C'est-à-dire existe-t-il des zones où la relation entre ressemblance génétique ou linguistique et distance géographique ne serait pas régulière : application de la méthode du Wombling (adaptation à des données sous forme de matrice de distances) (Bocquet-Appel et Baccro, 1994).

iii) Si l'hypothèse de l'arbre est statistiquement confortée, la congruence entre l'histoire évolutive des gènes et des langues est-elle significative pour la totalité de l'arbre, ou seulement pour des parties ou encore pour un ou plusieurs sous-ensemble de caractères : extension des méthodes de resampling (Farris et al., 1995), et de comparaisons d'arbres (Robinson et Foulds, 1981).

Plusieurs méthodes classiques illustrant la diversité génétique (et linguistique) seront aussi utilisées, avec en particulier une analyse en composantes principales et une analyse factorielle de correspondance. Nous compléterons par une méthode originale appelée "méthode des sites mobiles", combinant des distances géographiques et des distances génétiques qui aboutira à la construction de cartes de similarités génétiques (qui peuvent être représentées sous forme de cartes géographiques déformées), ainsi que de zones de discontinuités géographie/génétique. Les cartes des zones de discontinuité géographie-génétique feront apparaître les zones de divergence (populations géographiquement proches mais génétiquement distantes) et celles de convergence (populations génétiquement similaires). Ces différentes représentations rendent compte des

différences observées dans la structure génétique des migrations successives et font aussi apparaître le poids de certains haplotypes ou marqueurs génétiques dans la différenciation génétique (sous la forme de points-guides) (Hazout et al., 1993 ; Dugoujon et al., 1995 ; Mourrieras et al., 1997). Outre l'intérêt de rechercher les relations génétiques (ou linguistiques) entre les populations dans une même région, nos résultats seront plus globalement interprétés sous une forme dynamique (pattern génétique des vagues successives de migration, présence de gradients génétiques). Les voies possibles de migrations seront visualisées sous la forme d'un "spanning tree" superposé sur la carte de similarité génétique.

Une autre méthode originale sera aussi réalisée avec la recherche de la perméabilité de frontières linguistiques au flux génique: Cette procédure (Dupanloup de Ceuninck 1999; Dupanloup de Ceuninck et al. in press), permet de quantifier la perméabilité de frontières linguistiques, et de différentes sections de ces frontières, à l'échange génétique entre populations. L'application de cette procédure permet de tester si la corrélation généralement observée entre génétique et linguistique peut avoir été induite par l'action des frontières linguistiques sur le flux de gènes entre populations, et non pas forcément (ou uniquement) par une synchronisation des processus de différenciation génétique et linguistique dans l'histoire des populations humaines. La méthode permet également de localiser géographiquement des frontières génétiques, c'est-à-dire des zones où le taux de variation génétique est particulièrement important. Une fois localisées, ces frontières peuvent alors être comparées avec les frontières linguistiques et/ou écologiques.

IV - Les populations modernes

Les données archéologiques et génétiques apportent des arguments convainquants sur l'origine africaine des populations modernes. Cependant, la date et la (les) localisation(s) géographique(s) de leur expansion sont sujettes à débat (Stringer, 2000). A cet égard, les populations nord-africaines ont dû faire partie de ce scénario et elles ont probablement joué un rôle important dès les premières migrations et dans l'évolution des populations modernes. Des liens ont été démontrés entre les populations sub-sahariennes et celles de l'Europe de l'Ouest jusqu'à un passé récent. Dans ce contexte, les Berbères (une des principales cultures dans cette région) ont influencé les différenciations culturelles (démontré par les affinités linguistiques) et biologiques (affinités génétiques) entre l'Europe et les populations sub-Sahariennes (Excoffier et al., 1987). Linguistiquement, le berbère est une des cinq familles majeures du phylum Afro-Asiatique (Rulhen, 1987). Les affinités linguistiques entre le berbère et le tchadique suggèrent des affinités culturelles jusqu'environ - 15000 ans.; cela contraste avec l'apparente discontinuité morphologique entre les Hausa d'origine sub-Saharienne et les populations Shilha du Maroc (plus "Caucasoides"). Cela pourrait être expliqué par une longue période de contact culturel et un possible "remplacement" de la langue, avec un relativement faible flux génique.

Tout ceci démontre en Afrique du Nord et de l'Ouest l'interaction des gènes et de la culture durant ces dernières 40000 années. Le contraste entre les affinités génétiques des Berbères avec des populations européennes, et linguistiques avec les populations africaines plus au sud, montre qu'une analyse telle que nous la proposons dans ce projet éclairera une des questions les plus passionnantes de l'évolution humaine.

Les outils que nous avons à notre disposition ont montré tout leur intérêt en génétique des populations (voir annexes). Les marqueurs Gm par exemple se sont particulièrement révélés informatifs pour l'étude des populations d'Afrique (Excoffier et al., 1991 ; Dugoujon et al., 1997) et des Berbères en particulier (Chaabani et al, 1984 ; Yacoubi et al., soumis). Les marqueurs moléculaires (ADN) portent sur des polymorphismes ayant des taux de mutation variables, allant de l'ADN mt aux SNPs (Single Nucléotide Polymorphisms) jusqu'aux microsattellites, minisatellites hypervariables et les séquences ALU. Ainsi, l'évolution sera mesurée avec plusieurs échelles de temps. Des haplotypes très résolutifs seront construits sur des distances d'environ 100 kb. Des recherches ont déjà été entreprises pour caractériser de tels haplotypes sur le chromosome X, de même que dans certaines régions candidates à des maladies génétiques dans des populations d'Afrique de l'Ouest. Ils apporteront des données originales et une vue inédite du panorama génétique des populations d'Afrique du Nord.

V - Les populations anciennes

Le peuplement du nord de l'Afrique de la préhistoire récente à l'époque moderne est

relativement mal connu, et les données de la morphologie osseuse ont fait l'objet, pour les populations les plus récentes, notamment depuis 14 000 ans environ, d'interprétations radicalement opposées (Crubézy 1992 ; Crubézy et al., 1999). Aux populations de chasseurs-cueilleurs, appelés anciennement mechtoïdes ou cromagnoïdes, de morphologie très robuste, retrouvées dans le Sahara, la vallée du Nil (Wadi-Halfa) et le Maroc (Taforalt, Afalou Bou Rhumel) ont fait suite, à partir du néolithique, des populations plus graciles. Si pour certains, il y a là une rupture dans le peuplement (Camps, 1974), pour d'autres (Carlson et Van Gerven, 1977) il y aurait eu évolution *in situ* avec des changements morphologiques, notamment une gracilisation due à un changement de mode de vie. A l'époque historique, l'expansion arabe sous les Omayyades jusqu'en 750 après J.-C. a fait l'objet de quelques hypothèses avec en particulier l'arrivée plus ou moins massive de populations.

Dans ce contexte, les influences culturelles avec le sud de l'Europe longtemps négligées, sont de mieux en mieux appréciées, particulièrement celles intéressant le néolithique. En effet, elles sont contemporaines des principaux changements morphologiques observés, mais surtout de changements démographiques majeurs dans l'histoire de l'humanité. Sur les zones où les populations de langue berbère sont aujourd'hui rencontrées, pour le paléolithique supérieur (entre - 35000 et - 10000), les influences européennes semblent pouvoir être écartées. La situation change avec la mise en place des premières civilisations connaissant l'usage de la terre cuite (Guilaine, 1994). Après le néolithique, les influences et les contacts du Maghreb furent constants avec l'Europe, certes variables en intensité en raison des prises et des déprises historiques. Les relations du Nord de l'Afrique avec la zone sub-saharienne et l'Afrique noire n'ont véritablement été abordées, tant sur le plan culturel qu'anthropologique, que dans la vallée du Nil, véritable corridor rejoignant ces deux régions du monde (Crubézy et al., 1999). Dans le Maghreb, les caravanes et les Touaregs ont joué un rôle important.

En fait, les phases de peuplement se passent dans un contexte écologique et climatologique très particulier : celui de l'assèchement du Sahara. Cela va aboutir, entre 3000 et 4000 avant J.-C. à une rupture entre le nord et le sud du continent et à une augmentation de la segmentation en différentes entités culturelles et géographiques du Nord de l'Afrique. Rappelons qu'entre 10000 et 5500 avant J.C., le Sahara connut une phase (le Grand Humide) qui vit le développement des savanes, la mise en place de plusieurs grands systèmes lacustres, le développement des oasis et celui du développement de la végétation méditerranéenne sur les hautes terres. Des groupes de chasseurs-collecteurs nomades sillonnaient alors toute cette zone et atteignaient même ce qui sera plus tard le désert de la Lybie méridionale ou de l'Ouest égyptien. Quelques uns de ces groupes se sont parfois sédentarisés près des points d'eau, d'autres sont restés nomades. Par la suite le Sahara va s'assécher, atteignant vers 3000 avant J.C., en Egypte par exemple, un aspect proche de la période actuelle. Quel a été le devenir des populations sahariennes lors de l'assèchement de cette région ? Sédentaires ou mobiles, elles ont été forcées de partir vers d'autres lieux. On peut suggérer un éclatement vers le bassin du lac Tchad, la vallée du Nil, les côtes méditerranéennes et atlantiques, voir certains hauts plateaux de l'Atlas qui devaient être plus humides. Dans ce contexte, l'origine des populations Berbères a toujours questionné les archéologues et les anthropologues.

Les périodes les plus intéressantes à étudier sont essentiellement l'épipaléolithique (mais les squelettes sont rares). C'est pour cela que l'apport du site de Taforalt est très précieux. Le néolithique de toute l'Afrique du Nord, dont la nécropole de Skhirat (Maroc) a fourni le meilleur échantillon actuellement disponible, mais aussi la période médiévale, jalon intéressant entre le néolithique et l'époque actuelle et dont de nombreux squelettes sont disponibles, tant au Maroc (collaboration d'Abdel Ben Ncer) que dans le sud de la péninsule ibérique (nécropoles sarrazines

de l'Algarve, collaboration Pr. E. Cunha, Université de Coïmbra).

La difficulté principale dans l'étude génétique des populations modernes est d'établir la chronologie selon lesquelles se sont produites les différentes migrations. En conduisant une étude coordonnée des mêmes marqueurs génétiques dans des populations modernes et préhistoriques on pourra préciser la chronologie des migrations. Notre programme impliquant l'ADN ancien sera le suivant :

(i) haplotypes " population-spécifique " des individus de Taforalt, population épipaléolithique. Dans un travail récent, l'équipe d'E. Béraud-Colomb a montré qu'il était possible d'extraire, d'amplifier et de séquencer des fragments d'ADN provenant d'ossements humains de cette population de Cromagnoïdes Africains ayant vécu au Maroc il y a 12 000 ans (Béraud-Colomb et al ., 1995 ; Stevanovitch et al., soumis). Il a été amplifié et séquencé un fragment de la D-Loop de l'ADN mitochondrial. Cette séquence s'est révélée être particulièrement polymorphe. Le séquençage de cette région d'ADN doit être complété et élargi à d'autres régions afin de pouvoir définir l'haplotype. Nous utiliserons la même méthodologie que celle que nous utilisons dans l'étude de populations humaines actuelles : séquençage des régions HV1 ET HV2 de la D-Loop et détermination de la présence (ou absence) de 14 sites polymorphiques répartis sur le génome permettant la classification des haplotypes dans les 10 haplogroupes définis par A. Torroni (Mogentale-Profizi et al., soumis) . D'autres individus de Taforalt ont été étudiés (5 individus). Ces premiers résultats montrent que ce ne sont pas des individus apparentés aux sub-Soudanais actuels et soulèvent un grand nombre de questions. Avant tout, qui sont les descendants de Taforalt? Ou sont-ils? En Europe ou en Afrique du Nord? Etaient-ils Berbères? Leurs descendants sont-ils parmi les populations Berbères actuelles? Cette étude sera complétée par l'analyse de nouveaux spécimens de Taforalt, mais aussi par l'analyse d'une autre population ayant vécu au Maghreb il y a 10 000 ans: la population Afalou. Enfin, de nouveaux spécimens Berbères seront aussi étudiés (en collaboration avec F. Paris).

(ii) La même méthodologie, augmentée de marqueurs nucléaires pour l'étude de locus microsatellites (D3S1358, vWA, FGA, D8S1179, D2S11, D18S51, D5S818, D13S317, D7S820) sera utilisée pour des populations plus récentes afin d'évaluer l'évolution du pool génique dans ces parties du monde. Seront ainsi étudiées : des populations néolithiques du Maroc (dans celle de Skhirat notamment, et sur la côte atlantique à 30 kms de Rabat : fouilles de la "mission du littoral" de la Commission des fouilles Franco-Marocaines dont le Dr. Lacombe finit actuellement l'étude anthropologique classique ; (ii) dans des populations post-néolithiques d'Afrique du Nord (Maroc notamment : celles sous la responsabilité actuelle d'Abdel Ben Ncer) et "sarrazines" d'Andalousie (iii) des recherches semblables "de contrôle" seront menées sur les populations d'époque pro et proto-historique de la vallée du Nil. Ces dernières recherches seront effectuées dans le cadre du programme sur "*Le peuplement de la vallée du Nil, apport de l'ADN fossile*" dirigé par le Pr. E. Crubézy dans le cadre de l'Institut Français d'Archéologie Orientale.

Bibliographie :

- Aikhenvald A. Y. (1988), "A Structural and Typological Classification of Berber Languages", *Progress Tradition in African and Oriental Studies Special Issue*, n° 21, p. 37-43.
- Ameur M. (1990), "À propos de la classification des dialectes berbères", *Études et Documents Berbères*, 7, p. 15-27.
- Barbujani, G., 1991, What do languages tell us about human microevolution? *Trends Ecol. Evol.*, 6,151-156.
- Basset A. (1952), *La langue berbère*, Oxford University Press, International African Institute.
- Béraud-Colomb E., Roubin R., Martin J., Maroc N., Gardeisen A., Trabuchet G., and Goossens M.(1995). Human B-globin gene Polymorphisms characterized in DNA Extracted from Ancient Bones 12,000 years old. *Am. J. Hum. Genet.*, 57 : 1267-1274.
- Bocquet-Appel J.P.,Baccro J.N., 1994. Generalized Wombling. *Systematic Biology* 43(3)442-444
- Camps G. 1974. *Les civilisations préhistoriques de l'Afrique du Nord et du Sahara*. Doin Edts, Paris.
- CarlsonD.S., Van Gervend.P., 1977.- Masticatory function and post-Pleistocene evolution in Nubia. *American*

- Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton University Press, Princeton, New Jersey
- Chaabani, H., Helal, A.N., van Loghem, E., Langaney, A., Benammar Elgaaiied, A., Rivat Peran, L., and Lefranc, G., 1984, Genetic study of Tunisians Berbers. I. Gm, Am and Km immunoglobulin allotypes and AB0 blood groups. *J. Immunogenet.*, 11, 107-113.
- Chamla MC (1971) L'anthropologie biologique. PUF, Collection " Que sais-je "
- Sous la Direction de E.Crubézy (1992). "*Paléo-ethnologie Funéraire et Paléo-biologie*". Ouvrage de 12 contributions et 135 pages. Archéo-Nil, Bulletin N°2.
- Crubézy E, Telmon N, Sevin A, Picard J, Rougé D, Larrouy G, Braga J, Ludes B, Murail P. (1999) *Microévolution d'une population historique. Etude des caractères discrets de la population de Missiminia (Soudan, III^e-VI^e siècle)*. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*. T11,n°1-2. 213 pages
- Dugoujon J.M., B.Mourrieras, M.T. Senegas, E. Guitard, A. Sevin, E. Bois, S. Hazout.1995. Human genetic diversity (immunoglobulin GM allotypes), linguistic data and migrations of Amerindian tribes. *Human Biology* 67:231-249
- Dugoujon JM, Guitard E, Sénégal MT, Loirat F, Hazout S, Sevin A. Anthropologie génétique en Afrique: L'apport des marqueurs génétiques des immunoglobulines (systèmes GM des IGHG). Paper presented at Colloque des Anthropologues de langue française. Strasbourg, France, 1997, June 22-24
- Dupanloup de Ceuninck, I. (1999) *Evaluation et synthèse des contributions de la linguistique et de la génétique à l'étude de la différenciation des populations humaines pendant la préhistoire récente*. Thèse de Doctorat: Laboratoire de Génétique et Biométrie, Université de Genève.
- Dupanloup de Ceuninck, I., Schneider, S., Langaney, A. & Excoffier, L. (sous presse). Inferring the impact of linguistic boundaries on population differentiation: application to the Afro-Asiatic/Indo-European case". *European Journal of Human Genetics*.
- Excoffier L, Harding R, Sokal RR, Pellegrini B, Sanchez-Mazas A.1991. Spatial differentiation of RH and GM haplotype frequencies in Sub-Saharan Africa and its relation to linguistic affinities. *Hum Bio* 63:273-297
- Excoffier L, Pellegrini B, Sanchez-Mazas A, Simon C, Langaney A Genetics and History of Sub-Saharan Africa.1987. *Yearbook of Phys Anthropol*, 30:151-194
- Farris J.S., Källersjö M., Kluge A.G., But C., 1995. Testing significance of incongruence. *Cladistics* 10:315-319.
- Guilaine(J, 1994.- *La mer partagée. La Méditerranée avant l'écriture, 7000-2000 BC*. Paris, Hachette.
- Guénoche A., Garreta H., 2000. Quelle confiance accordée à une représentation arborée. *Communication JOBIM*, 3-5 Mai, 2000, Montpellier.
- Hazout S., G. Guasp, F. Loirat, P. Maurieras, G. Larrouy, J.M. Dugoujon. 1993. A new approach for interpreting the genetic diversity in space : "Mobile Site Method". Application to Gm haplotype distribution of twenty seven populations from North and Central America. *Annals of Human Genetics* 57:221-237
- Louali N. (1992). "Le système vocalique touareg." *Pholia* 7 (CRLS-Université Lumière Lyon 2 : 83-115.
- Mogentale-Profizi N, Chollet L, Dubut V, Stévanovitch A, Poggi C, Pradié MP, Spadoni JL and Béraud-Colomb E. Mitochondrial DNA sequence diversity in two groups of italian, Veneto speaker from Veneto (soumis)
- Mourrieras B., J.M. Dugoujon, L. Buffat, S. Hazout. 1997. Assessment of genetic diversity in space by superimposition of a distorted geographic map with a spatial population clustering. Application to Gm haplotypes of Native Amerindians tribes. *Annals of Human Genetics*. 61: 45-55
- Prasse K. G. (1972), Manuel de grammaire touarègue (Tahaggart) : Phonétique, Ecriture, Pronoms, Copenhagen, Akademisk Forlag.
- Robinson D.F., Foulds L.R., 1981. Comparison of phylogenetic trees. *Mathematical Biosciences* 53:131-147.
- Ruhlen. M. (1987). *A guide to the World's languages*. Stanford University press
- Stevanovitch A, Garcin G, Taillé A, Spadoni JL, Frackowiak S,Béraud-Colomb E. Direct evidence of ancient DNA from human bones up to twelve thousand years old by probe hybridisation. (soumis)
- Stringer C. (2000), Coasting out of Africa. *Nature* 405:24-27.
- Willms (1980), Die dialektale Differenzierung des Berberischen, Berlin, Verlag von Dietrich Reimer.
- Yacoubi L.B., Ennafaa H., Marrakchi R., Dugoujon J.M., Lefranc M., Sanchez-Mazas A., Ben Ammar Elgaaiied A. Gm allotypes and RFLP of gamma-IGCH genes in three different groups of Jerba Island. (soumis)

Annexes

I - Les marqueurs génétiques

Les allotypes GM et KM (marqueurs génétiques des IGHG)

Les gènes qui codent pour les IgG sont situés sur la partie distale du chromosome 14 (14q32.3). On décrit à ce jour 18 allotypes GM portés par les parties constantes des chaînes lourdes des IgG (4 sur les IgG1, 1 sur les IgG2 et 13 sur les IgG3). Ils définissent 15 'allèles' (4 sur les IgG1, 2 sur les IgG2 et 9 sur les IgG3) qui se combinent pour former 15 haplotypes communs. Il existe dans les populations des différences à la fois qualitatives et quantitatives. Le système KM est défini par 3 allotypes. Ils sont portés par la partie constante des chaînes légères kappa. Les gènes qui codent pour ces chaînes sont situés sur le chromosome 2 (2p12) (voir revue générale de Lefranc et Lefranc, 1990).

Tous les sérums seront typés pour les allotypes G1M(1,2,3,17), G2M(23), G3M(5,6,10,11,13,14,15,16,21,24,28) et KM(1) (Field et Dugoujon, 1989). Nous estimerons les fréquences des haplotypes par une méthode du maximum de vraisemblance.

ADN mitochondrial (ADNmt)

L'ADN mt est composé de plusieurs gènes (cytochrome oxydase, cytochrome b, etc...), de régions de régulation, mais aussi de régions non codantes (Anderson et al., 1981). Les cellules ont des centaines de copies d'ADN mt dans leur cytoplasme. Ceci est un grand avantage par rapport à l'ADN nucléaire, ce qui en fait une molécule de choix dans l'étude des populations anciennes et modernes. L'ADN mt est haploïde: durant l'évolution les mutations se sont accumulées sans recombinaison. Il évolue 5 à 10 fois plus vite que l'ADN nucléaire. Par sa transmission uniparentale, nous définissons en fait des phylogénies maternelles.

Nous effectuerons la séquence complète de l'ADN mt de 10 échantillons par population. Pour les autres échantillons, les régions contrôle I et II (aussi appelées régions hypervariables en raison du taux de mutations très élevé) seront séquencées. De même, des PCR-RFLP seront effectuées. Tout cela permettra de définir des haplotypes très informatifs.

Le chromosome Y

Le chromosome Y apporte des informations complémentaires de l'ADNmt. Il est représenté le plus grand segment non recombinant du génome (environ 60 Mb) et permet la construction de phylogénies de lignées paternelles. Il y a plusieurs types de polymorphismes, et des taux de mutations différents:

- les marqueurs bi-alléliques: ils sont basés sur des substitutions de bases ou de petites délétions. Nous étudierons en particulier: les YAP, DYS257, DYS271, PN2 et PN3 (Hammer et al, 1998)
- les microsatellites: ce sont des répétitions de 2 à 5 nucléotides et leur riche polymorphisme est basé sur leur nombre. Ils sont détectés par PCR. Plusieurs microsatellites seront analysés sur tous les échantillons
- les minisatellites: les séquences répétées sont plus longues (6 à 100 nucléotides). A ce jour, un seul a été décrit: MSY1 (voir Bouzekri et al, 1998 pour revue). C'est un marqueur très discriminant. Les MVR-PCR seront effectuées sur une partie des échantillons (sélectionnés sur la base des haplotypes définis à partir des marqueurs bi-alléliques et des microsatellites).

Les autosomes

Ces marqueurs ont été étudiés les premiers pour reconstruire l'histoire de l'évolution de l'Homme (voir Cavalli-Sforza et al, 1994). De nombreux haplotypes ont été identifiés. Ils seront définis dans l'équipe de M. Lathrop. De nouveaux SNPs seront détectés.

Nous utiliserons également les haplotypes informatifs du chromosome X qui sont très intéressants pour les études de populations sur une période allant de -10000 à -50000 ans. Quatre régions haplotypiques sur le bras long du chromosome s'étalant sur 50 Mb environ seront étudiées: 1) la région Xq26 définie par 4 SNPs; 2) Xq137.977-138.062 (Xq27) qui s'étend sur 80 kb (12 SNPs et deux microsatellites); 3) la région IDS (Xq27) sur 30 kb (10 SNPs) et 4) la région G6PD (Xq28) sur 40 kb (10 SNPs). Par la suite, nous définirons des régions similaires dans un segment de 10 à 15 kb pour construire des haplotypes.

De même, sur les autosomes, nous étudierons les polymorphismes de 15 insertions ALU: HS2.43, HS4.14, HS4.65, HS4.75, Sb19.3, Sb19.12, APO, B65, CO13A1, D1, PV92, TPA25, CD4, ACE y HS4.32 (Arcot et al. 1998; Milewicz et al. 1996). Comparé à d'autres marqueurs, ceux-ci ont l'avantage de représenter des événements uniques dans l'évolution (probabilité très faible de transposition indépendante). Leur utilité sera démontrée pour une approche de l'histoire démographique des populations (Harpending et al, 1998; Lopez-Alomar et al, sous presse).

Certains polymorphismes se révéleront informatifs, selon que nous nous intéresserons à des événements récents (il y a environ 1000 ans), ou plus anciens (de 10000 à 40000 ans). Notre riche base de données faisant appel à tous les types de marqueurs, nous l'utiliserons au mieux pour appréhender la co-évolution du langage, de la culture et des gènes.

Bibliographie

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA,

- Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. (1981) Protein, Nucleotide, Genome Sequence and organization of the human mitochondrial genome. *Nature*. 290:457-465.
- Arcot SS, Adamson AW, Risch G, LaFleur J, Lameerdin JE, Carrano AV, Batzer MA (1998) High-resolution cartography of recently integrated chromosome 19-specific Alu fossils. *J Mol Biol* 281:843-855
- Bouzekri N, Taylor PG, Hammer MF, Jobling MA. (1998) Novel mutation processes in the evolution of a haploid minisatellite, MSY1: array homogenization without homogenization.. *Hum Mol Genet*. 7:655-659.
- Field, L.L., and Dugoujon, J.M., 1989, Immunoglobulin allotyping (Gm, Km) of GAW 5 families. *Genet. Epidemiol.*, 6, 31-34.
- Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL.(1998). Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Mol Biol Evol*. 15:427-441.
- Harpending HC, Batzer MA, Gurven M, Jorde L, Rogers AR, Sherry ST (1998) Genetic traces of ancient demography. *Proc Nat Acad Sci USA* 95:1961-1967
- Lefranc MP, Lefranc G. *Molecular genetics of immunoglobulin allotype expression*; in: Shakib F. (eds): The human IgG subclasses: molecular analysis of structure, function and regulation. Pergamon Press, 1990, Oxford pp 43-78
- Lopez-Alomar A, Harich N, Kandil M, Valveny N, Esteban E, Gonzalez E, Bao M, Moral P (2000) Microgeographical variation of four polymorphic Alu insertions (CD4, TPA25, APo, ACE) in Western Mediterranean: frequencies on several groups from the Iberian Peninsula and Morocco. (in press).
- Milewicz DM, Byers PH, Reveille J, Hughes AL, Duvic M (1996) A dimorphic Alu Sb-like insertion in COL3A1 is ethnic-specific. *J Mol Evol* 42:117-123.

II - Les populations

Population		Famille linguistique	Sous-famille	Pays
Atlas Berbere	Tamazight	Afro-Asiatique	Berbere	Maroc
Atlas Berbere	Tachlihit	Afro-Asiatique	Berbere	Maroc
Zenati	Djerba	Afro-Asiatique	Berbere	Tunisie
Zenati	Zuari	Afro-Asiatique	Berbere	Lybie
Zenati	Ghomara	Afro-Asiatique	Berbere	Maroc (Rif)
Zenati	Mzab	Afro-Asiatique	Berbere	Algerie
Est	Sokna	Afro-Asiatique	Berbere	Libye
Est	Awjila	Afro-Asiatique	Berbere	Libye
Est	Siwa	Afro-Asiatique	Berbere	Egypte
Ouest	Zenaga	Afro-Asiatique	Berbere	Mauritanie
Touareg	Ajjer	Afro-Asiatique	Berbere	Libye
Touareg	Iforas	Afro-Asiatique	Berbere	Mali
Touareg	Ioullemeden	Afro-Asiatique	Berbere	Niger
Hausa		Afro-Asiatique	Tchadique	Nigeria
Hausa		Afro-Asiatique	Tchadique	Niger
Bedouins		Afro-Asiatique	Semitique	Maroc ?
Bedouins		Afro-Asiatique	Semitique	Egypte ?
Sara	Kaba	Nilo-Saharien	Sara	RCA, Tchad
Peul	Fulfulde	Niger-Kordofanien	Peul	Tchad, RCA, Niger
Mande	Bobo Madare	Niger-Kordofanien	Mande	Mali, Senegal

