

Université de Montréal

Variabilité Génétique des Populations Ouest-Africaines

par

Elias GBEHA

Programmes de biologie moléculaire

Faculté de Médecine

Mémoire présenté à la Faculté de Médecine

en vue de l'obtention du grade de

Maître ès science (M.Sc.)

en Biologie Moléculaire

juillet, 2008

© Elias GBEHA, 2008

Université de Montréal
Faculté de Médecine

Ce mémoire intitulé :

Variabilité Génétique des Populations Ouest-Africaines

présenté par :
Elias GBEHA

a été évalué par un jury composé des personnes suivantes :

Dr Edward BRADLEY,
président-rapporteur

Dr Damian LABUDA,
directeur de recherche

Dr Marc TREMBLAY,
membre du jury

RÉSUMÉ

Notre patrimoine génétique dévoile, de plus en plus, les passerelles démogénétiques d'une susceptibilité plus accrue de certains individus à des maladies infectieuses complexes. En vue d'une caractérisation de la variabilité génétique des populations ouest-africaines, nous avons analysé 659 chromosomes X au locus *dys44* qui comprend, 35 SNPs et un microsatellite distribués sur 2853 pb en amont et 5034 pb en aval de l'exon 44 du gène de la dystrophine en Xp21.3. Les génotypes obtenus, par ASO dynamique et électrophorèse sur gel d'acrylamide, ont servi à la détermination des haplotypes. Des paramètres comme la diversité haplotypique (G) et l'indice de fixation (F_{st}) ont été calculés. Des analyses en composantes principales ainsi que multidimensionnelles ont été réalisées. Sur 68 haplotypes détectés, 26 sont nouveaux, et cette région, avec une diversité haplotypique moyenne (G_{moy}) de $0,91 \pm 0,03$, se révèle beaucoup plus hétérogène que le reste du continent ($G_{moy} = 0,85 \pm 0,04$). Toutefois, malgré l'existence de disparités sous régionales dans la distribution des variants du marqueur *dys44*, l'AMOVA montre d'une manière générale, une faible érosion de l'éloignement génétique entre les populations subsahariennes ($F_{st} = 1,5\%$; $p < 10^{-5}$). Certains variants tel que l'haplotype eurasiens B006 paraissent indiquer des flux transsahariens de gènes entre les populations nord-africaines et celles subsahariennes, comme l'exemplifie le pool génétique de l'une des populations ubiquitaires de la famille linguistique Nigéro-congolaise : Les Fulani. Nos résultats vont aussi dans le sens d'un héritage phylétique commun entre les Biaka, les Afro-américains et les populations de la sous-famille de langues *Volta-Congo*.

Mots-clés : variabilité, SNP, haplotype, *dys44*, populations, Ouest-Afrique.

ABSTRACT

The unravelling of our genetic heritage has revealed a demogenetic segueway leading to an increased susceptibility of certain individuals to complex infectious diseases. In order to characterize genetic variability among the West African populations, we analyzed 659 X chromosomes at the *dys44* locus which comprises 35 SNPs and a microsatellite spanning a region 2853 bp upstream and 5034 bp downstream of exon 44 of the dystrophine gene in Xp21.3. The resulting genotypes, obtained by dynamic allele specific oligonucleotide hybridization and acrylamide gel electrophoresis, were used for haplotype construction. Gene diversity parameters such as the haplotypic diversity (G) and fixation indexes (F_{st}) were estimated. Multidimensional analysis of the data, including principal component analysis was also performed. Of the 68 distinct haplotypes detected in our data set, 26 were novel. The mean haplotypic diversity (G_{moy}) was 0.91 ± 0.03 for this West African region which was shown to be more heterogeneous than the rest of the continent ($G_{moy} = 0.85 \pm 0.04$). However, despite certain sub-regional differences in the distribution of *dys44* variants, the analysis of molecular variance showed an overall decline in the genetic distance between Sub-Saharan populations ($F_{st} = 1.5\%$; $p < 10^{-5}$). Certain variants, such as the Eurasian-specific haplotype B006, appear to suggest a Trans-Saharan gene flux between North African and Sub-Saharan populations as exemplified by the observed genetic pool of one of the ubiquitous populations of the Nigerian-Congolese linguistic family: The Fulani. Our results are also in agreement with a phyletic heritage between the Biaka, the Afro-Americans and the populations of the *Volta-Congo* language subfamilies.

Key words: variability, SNP, haplotype, *dys44*, populations, West-Africa.

TABLE DES MATIÈRES

RÉSUMÉ	<i>iii</i>
ABSTRACT	<i>iv</i>
LISTE DES TABLEAUX	<i>viii</i>
LISTE DES FIGURES	<i>ix</i>
LISTE DES ABRÉVIATIONS	<i>x</i>
REMERCIEMENTS	<i>xii</i>
Chapitre I :	<i>1</i>
INTRODUCTION	<i>1</i>
1. Le patrimoine génétique des hommes modernes : un héritage de plusieurs milliers d'années d'évolution	<i>2</i>
1.1. L'Afrique : berceau de l'humanité	<i>3</i>
1.1.1. L'Afrique occidentale subsaharienne au regard de l'Archéologie, de l'Anthropologie, de la Paléontologie et de la Génétique des populations humaines	<i>5</i>
1.1.2. L'Afrique de l'Ouest : une préhistoire récente	<i>7</i>
1.1.3. Données ethnohistoriques et linguistiques de quelques populations contemporaines de l'Afrique de l'Ouest	<i>9</i>
1.1.3.1. Les Peulh	<i>9</i>
1.1.3.2. Les Mossi	<i>10</i>
1.1.3.3. Les Rimaibe	<i>10</i>
1.1.3.4. Les Tamasheq	<i>11</i>
1.1.3.5. Les Ewe	<i>11</i>
1.1.3.6. Les Akebou	<i>12</i>
1.1.3.7. Les Kabiye	<i>12</i>
1.1.3.8. Les Nawdam (ou Nawdm)	<i>12</i>
1.1.3.9. Les Ditammari	<i>13</i>
1.1.3.10. Les Adja, Fon, Goun, Kotafon, Mahi, Nago, Ahizo et Weme	<i>13</i>
1.2. Refonte de l'histoire évolutive des humains à travers l'étude de leur patrimoine génétique	<i>14</i>
1.3. Objectifs	<i>19</i>
Chapitre II	<i>20</i>
MATERIEL ET METHODES	<i>20</i>
1. Populations ouest-africaines étudiées	<i>21</i>
2. Extraction d'ADN	<i>25</i>
3. Système <i>dys44</i> et amplification <i>in vitro</i> par PCR	<i>26</i>

3.1.	Description de <i>dys44</i>	26
3.2.	Amplification par PCR	27
4.	Hybridation ASO	29
4.1.	Transfert sur membranes	29
4.2.	Marquage des sondes oligonucléotidiques.....	29
4.3.	Hybridation et exposition.....	31
4.4.	Décrochage des sondes	32
4.5.	Le microsatellite Tn.....	33
5.	Base de données et outils statistiques	34
Chapitre III.....		38
RESULTATS		38
1.	Les haplotypes du marqueur <i>dys44</i> au sein des populations africaines et afro-américaines.....	39
1.1.	Cohorte C1	39
1.1.1.	Au Bénin.....	50
1.1.2.	Au Burkina Faso	52
1.1.3.	Au Togo	52
1.2.	Cohortes C1 à C6 dans un contexte global.....	54
1.2.1.	L'haplotype B001	54
1.2.2.	L'haplotype B002	55
1.2.3.	L'haplotype B003	56
1.2.4.	L'haplotype B005	57
1.2.5.	L'haplotype B007	58
1.2.6.	L'haplotype B010	59
1.2.7.	L'haplotype B012	60
1.2.8.	L'haplotype B016	61
1.2.9.	L'haplotype B022	62
1.2.10.	L'haplotype B040	63
1.2.11.	L'haplotype B160	64
2.	Le poly T et les haplotypes étendus B..._Tn en Afrique de l'Ouest.....	65
Chapitre IV.....		70
DISCUSSION.....		70
1.	L'haplotype B003 et ses probables recombinants.....	75
2.	Haplotypes spécifiquement subsahariens en général présents hors du continent sous des formes recombinées	78

3. Les haplotypes majoritairement partagés entre l’Afrique subsaharienne et les autres régions du globe.....	85
4. La “migration retour” de l’Eurasie vers l’Afrique subsaharienne	87
5. Diversité de <i>dys44</i> en Afrique de l’Ouest : Un reflet du brassage génétique entre les populations des trois grandes familles linguistiques africaines	90
<i>CONCLUSION</i>	94
<i>BIBLIOGRAPHIE</i>	98
<i>ANNEXES</i>	101
<i>Annexe A</i>	xiv
<i>Annexe B</i>	xiv
<i>Annexe C</i>	xv
<i>Annexe D</i>	xxiii

LISTE DES TABLEAUX

Tableau I : Classification linguistique et modes de vie des populations ouest-africaines étudiées	22
Tableau II : Origine et nombre de chromosomes X analysés	24
Tableau III : Liste des courts fragments <i>dys44</i> et séquences des amorces correspondantes utilisées pour la réaction de PCR	28
Tableau IV : Liste des Oligonucléotides utilisés pour l'hybridation spécifique des allèles du locus <i>dys44</i> par ASO	30
Tableau V : Distribution des haplotypes de <i>dys44</i> et leurs poly T associés dans notre cohorte ouest africaine	43
Tableau VI : Paramètres de diversité et variance S^2 du microsatellite Tn en fonctions des modes T ₁₅ et T ₂₂ au sein des populations africaines, eurasiennes et océaniques	xiv
Tableau VII : Distribution des haplotypes (Hap) de <i>dys44</i> et leurs poly T associés en Afrique subsaharienne et chez les Africains-Américains	xv
Tableau VIII : Distribution des haplotypes de <i>dys44</i> en Afrique subsaharienne et dans quelques populations non-africaines	xxii

LISTE DES FIGURES

Figure 1 : Carte de la répartition géographique des populations ouest-africaines étudiées	21
Figure 2 : Le système <i>dys44</i> et ses sites polymorphes	26
Figure 3 : Structure allélique des haplotypes de <i>dys44</i> dans la cohorte C1	40
Figure 4 : Distribution des haplotypes par pays dans la cohorte C1	47
Figure 5 : Filiation des haplotypes de <i>dys44</i> et leurs fréquences relatives en % par population dans la cohorte C1	53
Figure 6 : Distribution spatiale de l'haplotype B001 en Afrique	54
Figure 7 : Distribution spatiale de l'haplotype B002 en Afrique	55
Figure 8 : Distribution spatiale de l'haplotype B003 en Afrique	56
Figure 9 : Distribution spatiale de l'haplotype B005 en Afrique	57
Figure 10 : Distribution spatiale de l'haplotype B007 en Afrique	58
Figure 11 : Distribution spatiale de l'haplotype B010 en Afrique	59
Figure 12 : Distribution spatiale de l'haplotype B012 en Afrique	60
Figure 13 : Distribution spatiale de l'haplotype B016 en Afrique	61
Figure 14 : Distribution spatiale de l'haplotype B022 en Afrique	62
Figure 15 : Distribution spatiale de l'haplotype B040 en Afrique	63
Figure 16 : Distribution spatiale de l'haplotype B160 en Afrique	64
Figure 17 : Distribution du microsatellite Tn au niveau sous régional ouest-africain	66
Figure 18 : MDS 1 basée sur la matrice des pairwise Fst des populations de la cohorte C1	67
Figure 19 : Analyse en composantes principales (ACP) des populations de la cohorte C1	68
Figure 20 : MDS 2 basée sur la matrice des pairwise Fst des populations subsahariennes et afro-américaines	69
Figure 21 : Diversité haplotypique de <i>dys44</i> en Afrique	xiii

LISTE DES ABRÉVIATIONS

ADN: acide désoxyribonucléique
ASO: *allelic specific oligonucléotide*
ATP : adénosine triphosphate
°C : degré Celsius
cm : centimètre
EDTA : *ethylene diamine tetraacetic acid*
HCl : acide chlorhydrique
ka : mille ans
kb : kilobase
KCl : chlorure de potassium
M : molaire
MDS: *multidimensional scaling*
MgCl₂ : chlorure de magnésium
min : minute
mJ : millijoule
mL : millilitre
mM : millimolaire
mm : millimètre
NaCl : chlorure de sodium
NaH₂PO₄ : orthophosphate de sodium
NaOH : hydroxide de sodium
pb: paire de base
PBS : *phosphate buffered saline*
PCR: *polymerase chain reaction*
pmol: picomole
RFLP: *restricted fragment length polymorphism*
s: seconde
SDS: *sodium dodecyl sulfate*
SNP: *single nucleotide polymorphisms*
SSPE: *saline-sodium phosphate-EDTA*
Tris-HCl : *trisaminomethane hydrochloride*
UV: *ultraviolet*
µL : microlitre
µmol : micromole
µCi : microcurie
µg : microgramme

A. Eusèbe Abléfonlin

REMERCIEMENTS

J'aimerais témoigner ma gratitude à toutes les personnes qui ont contribué à mon travail de maîtrise et, particulièrement

au Dr Damian Labuda, mon directeur de recherche, auprès de qui j'ai eu le privilège d'ébaucher cette captivante odyssée de la génétique des populations humaines,

au Dr Ambaliou Sanni qui m'a initié à la pratique de la recherche en sciences biomédicales appliquées,

à Vania, Claudia, Stéphane, Jean-François, Claude, Pascale et tous les membres de mon laboratoire d'accueil dont j'ai gracieusement pu bénéficier de l'expertise pluridisciplinaire,

au Dr Daniel Sinnett pour « les balises »,

au Dr Maja Krajinovic pour ses conseils,

à Dominika, Robert, Jasmine, Mathieu, Marie-Ève et tous les membres de la dream team au pavillon videotron pour leur soutien,

à Nina et Jean-Yves Ndiaye pour leur assistance inconditionnelle,

à Pearl et Jack Rothenberg que la providence m'a fait croiser dans l'aboutissement de mon projet académique à l'Université de Montréal,

à ma mère Marie, mon père Jean, mon frère et mes sœurs pour leurs prières.

Chapitre I :
INTRODUCTION

1. Le patrimoine génétique des hommes modernes : un héritage de plusieurs milliers d'années d'évolution

Dans la parution de la revue *Pour la Science* de Janvier 1997, le zoologiste Richard Dawkins reprenant sa théorie du "*gène égoïste*" (Dawkins 1976), déclarait que « la véritable fonction d'utilité de la vie, ce vers quoi tout tend dans la nature, c'est la survie de l'ADN... peu importe que la transmission d'un gène se fasse au détriment de quelqu'un ou de quelque chose, parce que les gènes ne se préoccupent de rien ». Si des modifications du génome, suite à une mutation, une recombinaison ou une conversion génique, se révèlent favorables ou délétères pour un groupe d'organismes vivants, ce serait parce qu'à travers le temps, des mécanismes démographiques, écologiques, combinés à des processus de sélection naturelle, de dérive génétique, de flux génique, etc. ont, de manière significative, avantage leur accumulation au sein du patrimoine génétique de ce groupe d'organismes.

Au milieu du XIXe siècle, suite à la théorie darwiniste de *L'Origine des espèces* en 1859, plusieurs modèles théoriques ont été développés, afin de débayer les scories qui font encore écran à la compréhension de la dynamique d'introduction et de dispersion de diversité génétique au sein des organismes vivants. N'étant pas soumis aux mêmes contraintes évolutives, et ce dépendamment des facteurs biotiques et abiotiques en jeu, des différences morphogénétiques s'accumulent au sein des groupes d'organismes vivants, aboutissant après quelques générations, à la formation – ou à la disparition – de taxons nouveaux. Bien évidemment, au faîte de la pyramide des formes de vie existant sur terre, l'homme n'échappe pas à ces forces dont les effets – parfois

dommageables – au niveau individuel, intra-populationnel et inter-populationnel, trouvent leur fondement à travers plusieurs milliers d’années d’évolution.

Depuis quelques décennies, la génétique des populations humaines, au confluent de plusieurs disciplines scientifiques, rend effective l’analyse et la prédiction des processus de distribution de diversité génétique au sein des populations humaines et des individus qui les composent. Comment, à l’aune de son histoire, le patrimoine génétique du genre *Homo* illustre-t-il l’héritage phylétique transmis depuis la première espèce humaine jusqu’à la nôtre en ce début du XXI^e siècle ?

1.1. L’Afrique : berceau de l’humanité

Ces deux dernières décennies, les données paléanthropologiques ont centré l’Afrique au cœur des débats sur l’origine de l’homme moderne, *Homo sapiens* (Stringer 2002). Il y a environ 5 millions d’années (5000 ka), au Pliocène, seraient apparus en Afrique les premiers Hominines. Parmi ceux-ci, nous pouvons citer le genre *Australopithecus* et le genre *Homo*. Concernant le genre *Homo*, les conclusions divergent sur le nombre d’espèces qui auraient existé depuis le Pléistocène (1800 – 12 ka). Certaines études reconnaissent l’existence en cette période d’une seule espèce, l’espèce *Homo sapiens* (Hawks, Hunley *et al.* 2000). Selon d’autres études, il aurait existé au Pléistocène au moins quatre espèces du genre *Homo* : *Homo erectus*, son probable descendant *Homo heidelbergensis*, lui-même ancêtre potentiel de *Homo neanderthalensis* et de *Homo sapiens* (Stringer 2002).

Par ailleurs, de l’analyse des données génétiques et des fossiles, deux modèles prédominants s’opposent sur l’origine des hommes modernes :

(1) Le premier modèle, *Recent African Origin model* ou RAO (aussi nommé *Recent out-of-Africa* ou encore *Replacement model*), soutient l'existence, il y a 100 - 200 ka, d'une seule espèce. Cette dernière se serait dispersée dans le monde entier et aurait remplacé les espèces primitives comme *Homo neanderthalensis* en Europe, *Homo erectus* en Asie et d'autres espèces préexistantes avec lesquelles les hommes modernes n'eurent pas (ou très peu si ce fut le cas) d'unions fécondes (Cavalli-Sforza et Feldman 2003).

(2) Le second modèle, *Multiregional Evolution*, rejette le premier en mettant l'accent sur la continuité génétique qui existe entre les populations des divers continents, suite à l'exode des tous premiers migrants hors d'Afrique, il y a 100 -1800 ka (Wolpoff et Caspari 1996).

À ces modèles s'ajoutent deux autres secondaires :

a- Le *Hybridization and Replacement model*, soutient l'approche d'une hybridation à grande échelle entre les migrants africains et les populations indigènes ;

b- L'*Assimilation model*, situe l'origine de *Homo sapiens* en Afrique, et souligne l'importance du flux génique, du métissage et des variations de pression de sélection dans la dispersion et les changements morphologiques de notre espèce à travers la planète (Stringer 2002).

Il importe de noter que la conformité de ces modèles à l'origine de notre espèce dépend de l'interprétation des données génétiques et/ou paléanthropologiques.

1.1.1. L’Afrique occidentale subsaharienne au regard de l’Archéologie, de l’Anthropologie, de la Paléontologie et de la Génétique des populations humaines

Bien que ces deux dernières décennies, d’importants progrès aient été réalisés dans l’étude de la diversité génétique des populations humaines aux fins du traitement des maladies génétiques complexes et de la compréhension de l’histoire évolutive de l’espèce *Homo sapiens* depuis ses origines africaines jusqu’à nos jours, beaucoup d’interrogations demeurent quant aux processus démographiques et génétiques qui ont modulé – et qui modulent – le patrimoine génétique des populations humaines en général, et celui des populations africaines au sud du Sahara en particulier. Dans un grand nombre de ces études, sur un minimum de 1800 communautés ethniques que compte l’Afrique, l’ensemble du continent noir est soit représenté par les Afro-américains, soit par les Yoruba et/ou quelques populations, comme par exemple : les Mbuti, Biaka, Khoisan, Kung, Fulani, Mandenka, Bamiléké (Olson 1996). D’une manière générale, en dehors de cette récurrence de certaines populations africaines, celles subsahariennes, lorsqu’elles sont incluses, sont souvent sous représentées dans les études de génétique des populations (Reed et Tishkoff 2006).

La découverte des fossiles de primates du Pliocène, tels que *Australopithecus africanus* en 1924 à Taung en Afrique du sud, *Australopithecus afarensis* à Hadar en Ethiopie en 1974, celle des empreintes d’Hominidés des sites de Laetoli en Tanzanie et celle de beaucoup d’autres primates anciens dans la région de la *Rift Valley* (système de faille qui s’étend de la Jordanie au Mozambique, en passant par la mer Morte, la mer Rouge et des pays comme l’Ethiopie, le Kenya, etc.), ont supporté, entre autres, la théorie de l’*East side Story* sur la séparation de la branche des Hominidés du reste des Hominoïdes (Johanson et Taieb 1976; Leakey et Hay 1979; Coppens 1981).

Toutefois, ces données archéologiques et paléanthropologiques sont sujettes à controverse sur l'origine des premiers hommes anatomiquement modernes (HAM) et les voies possibles de migrations qui ont servi à leur expansion dans la conquête du globe. Jusqu'au début des années 1980, l'idée communément admise était que l'espèce *Homo sapiens* aurait évolué depuis environ 100 ka (Brauer 1984). Cependant, des restes fossiles d'HAM datant de 160 - 195 ka furent retrouvés en Ethiopie sur les sites de Herto et Omo Kibish ainsi que sur les sites de Qena en Egypte, Skhul et Qafzeh en Israël (sépultures datant de 90 - 120 ka) (Clark, Beyene *et al.* 2003). Qu'en était-il des fouilles archéologiques au centre et à l'ouest du vieux continent ? Par ailleurs, dans la seconde moitié des années 1980, l'analyse de séquences d'ADN mitochondrial issues d'une cohorte mondiale de 147 individus suggéra que l'ancêtre commun aux HAM pouvait être retracé jusqu'à une femme – l'Eve mitochondriale – de l'est africain dont l'existence remonterait entre 150 - 200 ka ans (Cann, Stoneking *et al.* 1987). Au milieu des années 1990, l'étude d'une portion de 2,6 kb du chromosome Y de seize humains et quatre chimpanzés, révéla que l'âge du chromosome Y ancestral remonterait à environ 188 ka (Hammer 1995). Deux ans plus tard, les analyses d'une région de 3 kb comprenant le gène de la β -globine soutinrent l'origine africaine de l'humanité avec l'âge de notre plus récent ancêtre commun estimé à 800 ka environ (Harding, Fullerton *et al.* 1997).

Bien qu'il demeure dans la communauté scientifique quelques poches de résistance contre l'origine africaine des HAM, on s'accorde, quels que soient les scénarios démohistoriques et/ou démogénétiques sur une chose : le genre *Homo* – anatomiquement moderne ou non – a conquis le reste du globe à partir de l'est de la *Rift Valley* en Afrique. Qu'en a-t-il alors été de l'odyssée de notre espèce à l'intérieur du continent noir ? Plus spécifiquement, que se passait-il à l'ouest de la région des grands

lacs de l'Afrique ? Quelles étaient les voies de migrations possibles qui ont servi au peuplement de l'Afrique occidentale en particulier, et de l'Afrique subsaharienne en général par l'homme ? À quand remontent les premières présences de vie humaine dans l'ouest subsaharien ? Quelle est l'amplitude de la variabilité génétique des populations dans cette région et, à travers les individus qui les composent, quelle histoire évolutive sous-tend l'introduction et la distribution de diversité dans leur pool génétique depuis le plus récent ancêtre commun jusqu'à l'heure actuelle ?

1.1.2. L'Afrique de l'Ouest : une préhistoire récente

S'étendant du Sahel à l'océan Atlantique (Figure 1 page 21), et des côtes de la Mauritanie aux limites occidentales du Tchad, l'Afrique de l'Ouest compte une constellation de populations qui se différencient les unes des autres, aussi bien par leurs pratiques socio-économiques et culturelles que leurs spécificités ethnolinguistiques. Jusqu'à quel point les vestiges du passé reflètent-ils cette richesse du pool anthropologique actuel de l'Afrique de l'Ouest ?

Depuis les années 1900 jusqu'à présent, le tableau de chasse des archéologues, anthropologues, paléontologues, préhistoriens etc. est demeuré vierge de tous restes ostéologiques du paléolithique spécifiant les premiers Hominidés en Afrique de l'Ouest. Pour la plupart, sur la base des restes osseux, l'occupation de cette région daterait seulement du début de l'Holocène (environ 12 ka), témoignant d'une préhistoire encore plus récente que celle des populations amérindiennes (environ 20 ka).

Cependant, l'analyse des archives sédimentaires des sites d'Ounjougou (au Mali), révéla l'existence d'industries lithiques (hachereaux, pointes de flèche en pierre taillée, éclats de grès etc.), constituant des indices d'occupation humaine de la région depuis plus de 500 ka (Soriano et Huysecom 2007). Contrairement aux autres régions

nord, centre, est et sud du continent, où la présence de restes organiques – d'origine humaine et/ou animale – permet une reconstitution à une échelle temporelle au carbone 14 du mode de vie des chasseurs cueilleurs paléolithiques, l'acidité des sols ferrugineux en Afrique de l'Ouest semble l'avoir dépouillée des témoignages osseux d'une occupation très ancienne. Ainsi, l'une des preuves d'un peuplement de l'ouest africain par les premiers HAM demeure enfouie dans la diversité des faciès (type de culture matérielle selon la terminologie des préhistoriens) des industries lithiques des strates du paléolithique moyen d'Ounjougou au Mali (150 ka environ). D'autres indicateurs lithiques d'occupation humaine au paléolithique inférieur de la région de Mauritanie ont été aussi décelés. Par le biais de ces vestiges lithiques, il est possible, sur des bases paléoclimatologiques et paléoanthropologiques, d'inférer l'évolution climatique et les événements majeurs qui ont modulé la dynamique culturelle des populations humaines de l'ouest africain. Ainsi, l'apparition de conditions climatiques arides défavorables (ou l'absence d'humidité et d'une faune abondante) dans la région aurait contraint jusqu'à la fin de la période ogolienne (13 - 25 ka), ces populations à des navettes cycliques nord-sud/sud-nord depuis les latitudes du Sahara jusqu'au littoral du golfe de Guinée. Il y a 11 ka avant notre ère, le retour d'une végétation luxuriante de graminées favorisa une sédentarisation de ces hommes qui, progressivement, (notamment il y a 5 ka) développèrent à la place d'une cueillette intensive et de la chasse, l'agriculture, la poterie, l'élevage domestique (Huysecom 2007).

1.1.3. Données ethnohistoriques et linguistiques de quelques populations contemporaines de l’Afrique de l’Ouest

1.1.3.1. Les Peulh

La population Peulh (ou Fulani, Foulah, Peul, Peuhl, Pulaar, Fulbe etc.), qui comporte environ 30 millions d’individus, répartis dans 17 états africains, constitue l’un des plus grands groupes ethniques de l’Afrique subsaharienne. Les données ethnohistoriques situent le foyer peulh dans le royaume du Tekrur à l’est de la Guinée dans la vallée du fleuve Sénégal, d’où ils auraient migré à partir du XI^e siècle pour occuper une aire géographique qui s’étend du Cameroun jusqu’à la ceinture du Sahel (du Sénégal au Tchad). Au XV^e siècle, ils fondèrent l’Empire du Macina dans la région centrale de la République du Niger, puis au XIX^e siècle, l’Empire du Sokoto dans la moitié nord de la République du Nigeria. Selon des études archéologiques, les Peulh dont les origines remonteraient jusqu’en Egypte, auraient introduit l’élevage bovin en Afrique de l’Ouest (Tauxier 1937; Dupuy 1999). On distingue dans cette ethnie deux grands groupes : celui des nomades – éleveurs – (parfois semi-nomades) qui, à la quête de pâturage et d’eau, effectuent des mouvements migratoires saisonniers de quelques centaines à plusieurs milliers de kilomètres, et le groupe des sédentaires qui pratiquent généralement l’agriculture et l’élevage domestique. D’un point de vue anthropologique, les Peulh constituent un groupe hétérogène que les traits du visage rapprochent des peuples nord-africains mais dont la couleur de la peau – plus foncée – les en éloigne. Les Peulh parlent le fulfulde (Tableau I page 22) qui est un des dialectes de la branche atlantique de la famille des langues nigéro-congolaises (Ruhlen 1987).

1.1.3.2. Les Mossi

Du XIe au XIVe siècle et principalement au XIIIe siècle, suite à la chute de l'empire du Ghana, les populations Mossi se séparèrent du groupe Ashanti (localisé dans la région centrale de l'actuelle République du Ghana) et migrèrent dans le bassin de la Volta au Burkina Faso pour fonder le royaume Mossi (Roy 1998). Les Mossi qui constituent au moins 40% de la population actuelle du Burkina Faso vivent principalement d'agriculture et parlent le Mòoré qui est une langue Oti-volta du groupe Gur (Tableau I page 22). Ce dialecte présente jusqu'à 35% de similitude avec le Nawdam. Ce qui laisse supposer une origine linguistique commune à ces deux ethnies.

1.1.3.3. Les Rimaibe

Le terme Rimaibe (Tableau I) regroupe des populations Bambara, Soninké, Bwa et Dogon qui auraient été au XIXe siècle sous la domination de l'empire peulh du Macina. Ces anciens esclaves, actuellement localisés au Burkina Faso, adoptèrent les habitudes socioculturelles de leurs maîtres – auxquels ils sont parfois confondus – et parlent un dialecte Samo du groupe des langues mandées de la famille linguistique Nigéro-congolaise (Gordon 2005).

1.1.3.4. Les Tamasheq

Encore connus sous les noms de Touareg, Tuareg, Tamajak, Tamahak, les Tamasheq sont un groupe de Berbères nomades dont les origines remonteraient à l'Asie mineure, mais aussi à l'Europe du nord, la péninsule ibérique, les îles de la péninsule italienne (Camps 1981). De nos jours, on retrouve cette population au Sahara central (Mali, Burkina Faso, Niger, Algérie, Libye et Tchad). Les données anthropologiques placent à la tête de la lignée berbère, les protoméditerranéens capsians datant d'il y a 9 ka. Quant aux données linguistiques, elles militent en faveur d'une lointaine origine orientale de ce peuple dont la langue, le Tamasheq (Tableau I page 22), appartient à la branche berbère des familles de langues afro-asiatiques ou chamito-sémitiques (Greenberg 1963; Cohen 2007).

1.1.3.5. Les Ewe

Les Ewe sont principalement localisés au sud des Républiques du Ghana et du Togo. Ils seraient issus des mouvements migratoires, au début du deuxième millénaire (vers le XIIIe siècle), de populations de la cité yoruba d'Ifè au Nigeria. Ces déplacements les conduisirent dans un premier temps dans la cité de Tado au Togo. Au début du XVIIe siècle (vers 1610), l'expansion démographique et l'épuisement des ressources entraîna une partie du groupe dans une deuxième vague de migration au sud-ouest de Tado vers la rive gauche de la Volta au Ghana, où s'établit le foyer ethnique Ewe. Les Ewe (Tableau I), qui sont pour la plupart des agriculteurs ou des pêcheurs, parlent Ewe qui est un des dialectes des langues Gbe du groupe linguistique Kwa (Westermann 1930; Gordon 2005).

1.1.3.6. Les Akebou

Il s'agit d'un peuple des montagnes atacoriennes de la région frontalière centre-sud du Togo et du Ghana (Morrisson, Mitchell *et al.* 1989). Estimés au début du XXI^e siècle à environ 60000 habitants et vivant principalement de l'agriculture, les locuteurs du dialecte Akebou (Tableau I) qui appartient au groupe des langues Kwa, parlent aussi – pour la plupart – le Ewe. (Gordon 2005).

1.1.3.7. Les Kabiye

Les Kabiye, dont les "tous premiers ancêtres seraient descendus du ciel", auraient occupé la région du Lama Dessi au nord-ouest du Togo dès le Xe siècle de notre ère (Tcham 1990). Cette ethnie qui est constituée surtout d'agriculteurs, parle le Kabiye (Tableau I page 22), un dialecte des langues Grusi du groupe linguistique Gur (Bendor-Samuel 1965).

1.1.3.8. Les Nawdam (ou Nawdm)

Cette population sympatrique des Kabiye fait partie des migrants pré-adjas qui ont établi leur foyer dans la région frontalière de Lama Kara (au nord-est du Togo) vers la fin du premier millénaire. Vers le XV^e siècle, ces excellents cultivateurs du palmier à huile, à la recherche de meilleures terres, effectuèrent par vagues successives des mouvements migratoires nord-sud dans la région de Notsé puis jusqu'aux régions côtières de Lomé et d'Accra (Cornevin 1962; Gayibor 1997). Ce groupe ethnique parle le dialecte Nawdam (Tableau I) qui est une des langues Oti-Volta du groupe Gur.

1.1.3.9. Les Ditammari

Localisés au nord-est du Bénin dans la région montagneuse de Natitingou, les Ditammari (aussi appelés Somba) auraient leurs origines au Burkina Faso. Ce peuple vit principalement des activités fermières et parle un dialecte des langues Oti-volta du groupe Gur, telle que la précédente ethnie (Gordon 2005).

1.1.3.10. Les Adja, Fon, Goun, Kotafon, Mahi, Nago, Ahizo et Weme

Les Adja comme les Fon et les Goun font partie des migrants adjas dont les origines remonteraient vers le XIIe siècle au pays Ashanti (actuel Ghana) (Couao-Zotti 2007). Entre le XIIe et le XVe siècle, ces populations de l'ère Adja-Ewe se seraient déplacées vers la région de Tado (au Togo). Vers le XVIe siècle, ces groupes de population auraient migré de Tado, d'abord sur les rives du fleuve Mono où s'installa le groupe ethnique Adja, puis au sud du Bénin où le groupe des Alladanou fonda vers 1575 le royaume d'Allada (www.beninensis.net). Des querelles intestines conduisirent à la séparation des Alladanou en trois groupes. L'un des groupes demeura à Allada, tandis que les deux autres furent entraînés dans une troisième vague de migration. Le groupe ethnique Fon s'installa au nord d'Allada pour fonder vers les années 1620 le royaume d'Abomey. Celui des Goun fonda au sud-est d'Allada, vers les années 1730, le royaume de Porto-Novo. Il faudrait noter que sur le plateau d'Allada, vivaient déjà les Ayizo de provenance indéterminée – même si les ethnologues situent leurs origines au Ghana vers le XVIIe siècle (Morrisson, Mitchell *et al.* 1989; Decalo 1994). Par ailleurs, sur le plateau d'Abomey vivaient avant la domination des Fon, les Nago issues du noyau de peuplement yoruba du royaume de Ketou fondé au XIIIe siècle. Les Kotafon constitueraient le fruit du métissage entre populations Fon et Adja. Les Mahi, seraient

issues du métissage entre le XIII^e et le XVIII^e siècle des souches de peuples adjas et yorubas (de Lespinay 1994). Ces populations auraient par la suite fuit au XVII^e siècle la dynastie d'Abomey vers la région des collines de Savalou au centre-ouest du pays. Les Weme pour échapper à la domination des Fon, auraient aussi connu la migration forcée vers leur emplacement actuel au XVIII^e siècle (Pellissier 1962). Toutes ces populations parlent des langues Gbe du groupe Kwa des familles de langues nigéro-congolaises (Tableau I).

1.2. Refonte de l'histoire évolutive des humains à travers l'étude de leur patrimoine génétique

Au début des années 2000, avec l'achèvement du séquençage du génome humain, la caractérisation de la majorité des variations génétiques à l'échelle de tout le génome humain s'est avérée réalisable. Non seulement cela facilite la découverte des bases génétiques de susceptibilité de certains individus aux maladies complexes mais aussi, nous renseigne sur l'évolution de notre espèce. Ainsi par des analyses de données paléanthropologiques, archéologiques et génétiques entre humains d'origines ethnolinguistiques – voire socioculturelles – différentes et celles de primates non humains, il est de plus en plus possible de comprendre, au bénéfice des progrès de la génétique des populations humaines, les raisons d'un risque relativement plus élevé de certains individus ou de certaines populations humaines à certaines maladies – voire d'une réponse différentielle à certains médicaments.

L'existence de variation génétique à l'échelle moléculaire entre les populations humaines fut démontrée pour la première fois au début des années 1900 sur le système des groupes sanguins ABO (Hirszfeld et Hirszfeld 1919). L'introduction de la technique

d'électrophorèse pour l'étude des variants d'hémoglobine, trente années plus tard, contribua fortement à l'étude des marqueurs protéiques (Pauling, Itano *et al.* 1949). La technique de RFLP favorisa la détection de mutations affectant des sites de restriction (Botstein, White *et al.* 1980), jusqu'à la découverte en 1986 de la technique de PCR par Kary Mullis pour l'étude d'un plus grand nombre d'allèles d'ADN. Avec le développement du séquençage automatisé de l'ADN au début des années 1990, le champ d'investigation des variations au niveau de l'ADN s'est très rapidement élargi.

Au cours de ces deux dernières décennies, l'histoire génétique de notre espèce a été beaucoup étudiée à l'aide d'ADN mitochondrial (ADNmt) (Cann, Stoneking *et al.* 1987; Vigilant, Stoneking *et al.* 1991), et du chromosome Y (ChrY), notamment sa portion non-recombinante (Hammer 1995; Cruciani, Santolamazza *et al.* 2002).

L'étude de la variation à l'échelle globale de 1544 ChrY montra qu'une large proportion de la diversité paternelle africaine aurait ses racines en Asie (Hammer, Karafet *et al.* 1998). En Afrique, les séquences YAP⁺ portant la mutation SRY₄₀₆₄ (datant de 31 ka environ) marqueraient la lignée qui aurait migré de l'Asie. Cette vague de migration proviendrait du sud-est asiatique (Templeton 1998).

Au début des années 2000, l'analyse du ChrY de 608 individus issus de 22 populations africaines révéla au sein de populations du nord Cameroun (Daba, Ouldeme, Fali, Fulani et d'autres populations mixtes), la preuve – représentée par l'haplotype 117 – d'une récente “migration retour” de populations humaines d'Asie vers l'Afrique (Cruciani, Santolamazza *et al.* 2002). L'âge de coalescence de cet haplotype a été estimé à 4100 ans en Afrique, ceci refléterait la limite minimale de son entrée sur ce continent depuis l'Asie.

En 2004, les études du marqueur ADNmt de 77 Fulani de la République de Guinée-Bissau montrèrent en plus de la présence à haute fréquence d'haplogroupes de la famille L (L1b, L2a et L3b), la présence de quelques haplogroupes eurasiens (U5 et U6) (Rosa, Brehm *et al.* 2004).

En dépit de leur utilité indéniable pour l'histoire des populations humaines, ces deux marqueurs présentent quelques faiblesses dans l'inférence des mécanismes d'évolution. L'ADNmt, transmis de la mère à ses enfants, reflète les lignées maternelles, tandis que le ChrY, transmis de père en fils, reflète les lignées paternelles. Par ailleurs, à ce "biais de genre" s'ajoute un paramètre important dans l'analyse des données génétiques : le facteur temps.

En effet à l'échelle moléculaire, la majeure partie de la variabilité entre espèces est transmise de façon neutre et maintenue grâce à l'équilibre entre les mutations et la dérive génétique. Ceci implique, sous un modèle neutre, que la durée moyenne de fixation d'un polymorphisme est de $4N_e$ générations, où N_e correspond à l'effectif efficace de la population étudiée. Ce terme de $4N_e$ traduit, dans un modèle probabiliste comme celui de la coalescence, le nombre moyen de générations séparant les individus d'une population donnée de leur plus récent ancêtre commun (Kingman 1982; Hudson 1990). L'âge du plus récent ancêtre commun est appelé TMRCA. Par ailleurs, de l'union féconde d'un homme et d'une femme, le nombre de copies respectives transmissibles d'ADN autosomal, de ChrX, de ChrY et d'ADNmt à la génération suivante est dans un rapport 4 : 3 : 1 : 1. Ainsi en fonction des marqueurs ADN autosomal, ChrX, ChrY, et ADNmt analysés, la limite informative du polymorphisme à l'étude peut correspondre à $4N_e$, $3N_e$, N_e et N_e générations respectivement. Sur cette base, certaines études de l'ADNmt datent l'apparition des hommes anatomiquement modernes à 160 ka (Kivisild,

Shen *et al.* 2006), tandis que la valeur de TMRCA obtenue pour le ChrY est de 60 - 104 ka (Tang, Siegmund *et al.* 2002). Cette valeur de TMRCA du ChrY reflète un effectif efficace plus faible comparé à celui de l'ADNmt (Harding et McVean 2004; Garrigan et Hammer 2006). Une telle hétérogénéité des TMRCA pour ces deux marqueurs nécessiterait la diversification des marqueurs utilisés, permettant notamment une reconstruction temporelle plus profonde, sans biais de genre comme c'est le cas pour le chromosome sexuel X.

Le ChrX est un marqueur intéressant de l'histoire génétique de notre espèce. En effet, au niveau des hommes, il facilite la lecture directe des haplotypes (combinaison d'allèles sur un chromosome) et l'estimation du déséquilibre de liaison (Zietkiewicz, Yotova *et al.* 2003).

Les premières études faites avec le ChrX, pour expliquer l'histoire génétique de notre espèce, ont été effectuées sur un système expérimental proposé il y a une décennie (Zietkiewicz, Yotova *et al.* 1997). Il s'agit du système *dys44* (Figure 2 page 26) qui correspond à une portion intronique entourant l'exon 44 du gène de la dystrophine et localisée en Xp21.3. L'analyse de *dys44* à l'échelle mondiale a révélé l'existence d'haplotypes assez fréquents partagés par tous les continents tels que B001, B002, B003 et B005, un excès d'haplotypes spécifiques à l'Afrique, ainsi que des haplotypes spécifiques au reste du monde à l'exception de l'Afrique, comme l'haplotype B006 (Labuda, Zietkiewicz *et al.* 2000; Zietkiewicz, Yotova *et al.* 2003). Ces études soutiennent l'existence en Afrique d'au moins deux lignées fondatrices qui auraient évolué séparément pendant une longue période. La première serait demeurée sur le continent africain tandis que la seconde aurait mené au peuplement du reste du monde il y a 27 - 56 ka. Cette dernière se serait aussi métissée à la première en Afrique. Ceci

expliquerait la grande diversité génétique observée parmi certaines populations africaines au sud du Sahara par l'étude de plusieurs marqueurs génétiques.

D'autre part, pour le système *dys44*, une troisième lignée est représentée par un haplotype ancien, l'haplotype B006. Cet haplotype, qui est virtuellement absent de l'Afrique subsaharienne, est retrouvé en Eurasie et en Amérique, et aurait contribué à la colonisation des autres continents lors d'une première vague de migration il y a 160 ka. L'haplotype B006 (qui ne diffère de l'haplotype ancestral que par quatre mutations) se serait par la suite métissé avec les haplotypes de la deuxième lignée (Zietkiewicz, Yotova *et al.* 2003). Or parmi les populations subsahariennes étudiées, un individu Rimaibe fut détecté porteur de B006. La présence de cet haplotype constitue-t-elle une contribution non africaine au pool génétique des populations subsahariennes ? Le profil du système *dys44* dans la population Rimaibe constitue-t-il la seule preuve ténue d'une ancienne migration de populations eurasiennes en Afrique subsaharienne ou cela représente-t-il la pointe de l'iceberg dans cette région ouest-africaine ? Quel est le degré de différenciation génétique des populations dans cette région et à quel point les processus démogénétiques ont-ils modulé la distribution de ce marqueur au sein des populations ouest-africaines en particulier et subsaharienne en général ?

1.3. Objectifs

Notre projet de recherche vise principalement :

- la description de la diversité génétique des populations ouest-africaines à travers l'étude du système *dys44*.

Pour cela nous allons :

- étudier la distribution des haplotypes de *dys44* au sein de cinq-cent-cinquante individus issus de dix-sept ethnies ouest-africaines,
- déduire du calcul des paramètres de diversité, l'amplitude d'homogénéité génétique entre les populations de cette région,
- inférer de ces données et de celles d'études antérieures, les processus génétiques et démographiques qui, dans la région, sur le continent africain et dans le reste du globe, ont contribué à façonner le pool génique des populations subsahariennes.

Chapitre II
MATERIEL ET METHODES

1. Populations ouest-africaines étudiées

Nous avons analysé des échantillons d'ADN (441 hommes et 109 femmes) provenant du sang périphérique de 550 individus anonymes de degré d'apparentement inconnu. La collecte a été effectuée pendant les années 2003-2004 par le Dr Ambaliou SANNI et ses collaborateurs au Togo, au Burkina Faso et au Bénin (Figure 1 ; Tableau I page 22).

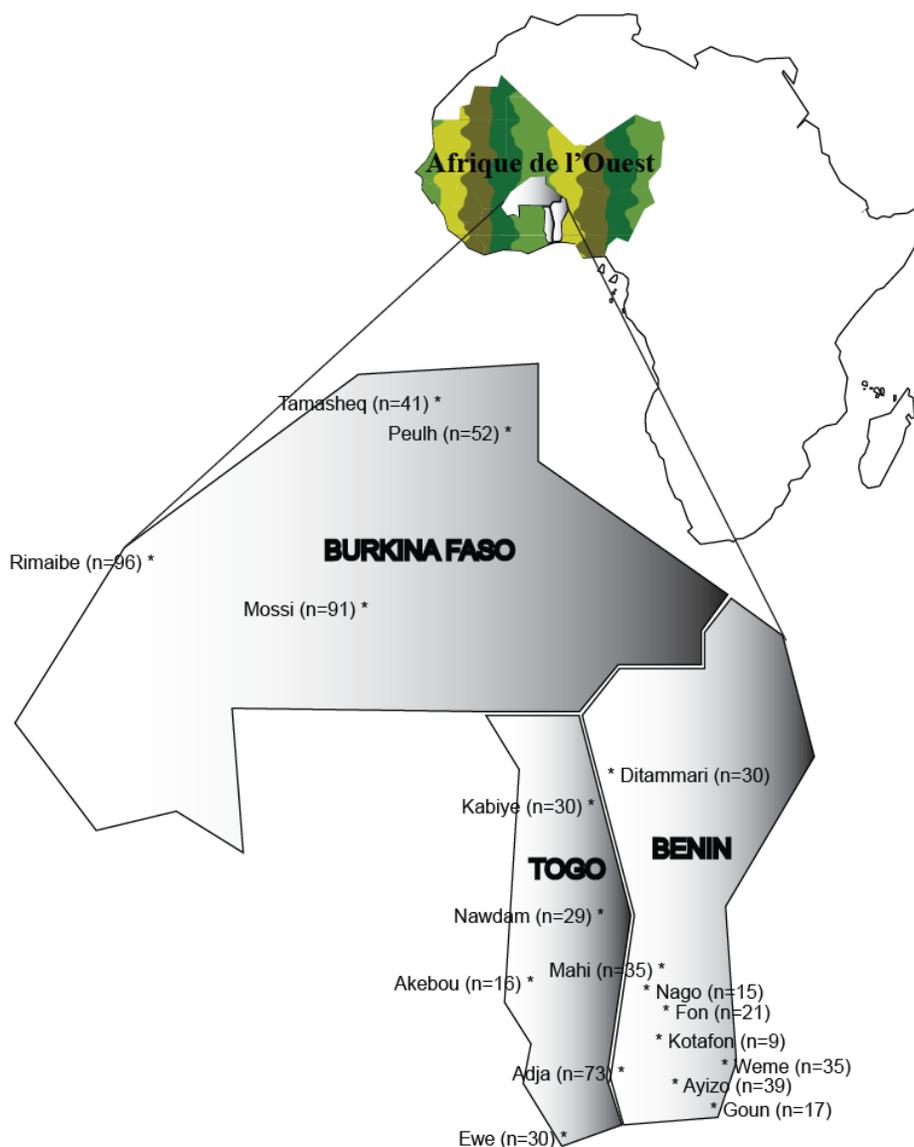


Figure 1 : Carte de la répartition géographique des populations ouest-africaines étudiées **n** correspond au nombre de chromosomes sexuels X analysés au sein de chaque ethnie

Populations	ChrX	Pays	Familles de langues	Branches	Groupes	Langues	Dialectes	Modes de vie	
Ditammari	30	Benin	Nigéro-congolaise	Volta-Congo	Kwa	Oti-Volta	Ditammari	Sédentaire, agricole	
Adja	73						Adja		
Ayizo	39						Ayizo		
Fon	21						Fon		
Goun	17						Goun		
Kotafon	9						Kotafon		
Mahi	35						Mahi		
Weme	35						Weme		
Nago	15	Burkina Faso	Afro-asiatique	Bénoué-Congo	Defoïde	Yoruboïde	Nago	Nomade, semi-nomade, agricole	
Peulh	52			Atlantique	Fulani-Wolof	Fula	Fulfulde		
Rimaibe	96			Mande	Eastern	Samo	Rimaibe		
Mossi	91			Volta-Congo	Gur	Oti-Volta	Mòoré		Sédentaire, agricole
Tamasheq	41			Berbère	Tamasheq	Southern	Tamasheq		Nomade, semi-nomade
Kabiye	30			Togo	Nigéro-congolaise	Volta-Congo	Gur		Grusi
Nawdam	29	Oti-Volta	Nawdam						
Akebou	16	Kwa	Kebu-Animere				Akebou		
Ewe	30		Gbe				Ewe	Sédentaire, agricole, halieutique	

Tableau I : Classification linguistique et modes de vie des populations ouest-africaines étudiées

Suite à l'agrément éclairé et signé dans un formulaire de consentement de chacun des participants, des gouttes de sang ont été prélevées, à l'extrémité d'un doigt, sur papier Whatman (*FTA Classique card*) à l'aide d'autopiqueur muni d'une lancette à usage unique (*Safety lancet, Sarstedt*). Le tableau II à la page 24 donne un aperçu du nombre de chromosomes analysés par population dans notre cohorte ouest-africaine C1. D'autres populations africaines et eurasiennes précédemment étudiées, dans notre laboratoire d'accueil (Zietkiewicz, Yotova *et al.* 2003; Xiao, Yotova *et al.* 2004; Lovell, Moreau *et al.* 2005), ont été ajoutées à notre cohorte C1, et analysées dans le cadre de notre étude (Tableau II page 24).

n ChrX	Localisations	Populations	Cohortes
73	Bénin	Adja	C1
39		Ayizo	
30		Ditamari	
21		Fon	
17		Goun	
9		Kotafon	
35		Mahi	
15		Nago	
35		Weme	
30		Kabiye	
16	Togo	Akebou	C2
30		Ewe	
29		Nawdam	
52		Peulh	
96		Rimaibe	
91		Mossi	
41		Tamasheq	
31		Rimaibe 2	
30		Mossi 2	
39		Ga	
65	AA	C1'	
85	Centrafrique	Biaka	C1''
58	Congo Démocratique	Mbuti	C3
30		Amhara	
19	Ethiopie	Oromo	C4
16		Juif Ethiopien	
18	Afrique du Sud	Khoisan	C5
29	Maroc	Juif Marocain	
19	Sahara Occidental	Saharawi	C6
31	Yémen	Juif Yéménite	C7
110	Pol, Rou, Ukraine	Juif Ashkénaze	
23	Palestine	Palestinien	C8
20	Espagne	Basque	
26	Italie	Italien	C9
78	Chine	Chinois	
65	Japon	Japonais	C10
24	PNG	Papou c	
41			Papou h

Tableau II : Origine et nombre de chromosomes X analysés

n ChrX = nombre de chromosomes X

E.U. = Etats-Unis d'Amérique ; Pol = Pologne ; Rou = Roumanie ; PNG = Papouasie-Nouvelle-Guinée

2. Extraction d'ADN

Une ponction d'environ 6 mm de diamètre de papier Whatman imprégné de sang est découpée en quatre et mise – sans être tassée – dans un tube eppendorf de 0.6 ml. Au contenu de chaque tube sont ajoutés 400 μ l d'une solution fraîchement préparée de PBS 1x. Le tout délicatement inversé environ toutes les 5 min, est laissé pendant 20 min à température ambiante. Le surnageant chargé d'hémoglobine et de déchets est pipetté, laissant un culot de papier Whatman lavé, réhydraté et riche en cellules nucléées au fond du tube. Au culot sont ajoutés 40 μ l de solution de lyse fraîchement préparée, composée comme suit : 10 mM de NaOH, 200 mM de NaCl et 0,05% de SDS. L'ensemble est modérément vortexé pendant 20 s, chauffé pendant 6 min à 95°C, puis centrifugé pendant 3 min à 16000 g, et le surnageant transvidé dans une plaque PCR de 96 puits (*ultraAmp PCR Plates, BioScience Inc.*). Ces isolements d'ADN conservés à -20°C, constituent notre stock d'ADN dont les aliquots serviront à la réaction de PCR. Il importe de notifier que des échantillons contrôles, génotypés auparavant, ont été ajoutés sur toutes nos plaques PCR.

3. Système *dys44* et amplification *in vitro* par PCR

3.1. Description de *dys44*

Le système *dys44* (Figure 2) précédemment décrit par Zietkiewicz *et al.* (1997) comprend un segment d'environ 8 kb, soit 2853 paires de bases en amont et 5034 paires de bases en aval de l'exon 44 (148 pb) du gène de la dystrophine au locus Xp21.3 (*Genbank* U94396). On y trouve trente-cinq sites polymorphes dont trente-et-un constituent des substitutions nucléotidiques simples, un site triallélique, deux sites correspondant chacun à une délétion de triplet nucléotidique, et un site qui consiste en une duplication de huit nucléotides. À ces trente-cinq polymorphismes, s'ajoute un microsatellite résultant d'une répétition en tandem de nucléotide T (poly T) dont la longueur varie entre quatorze et vingt-trois répétitions.

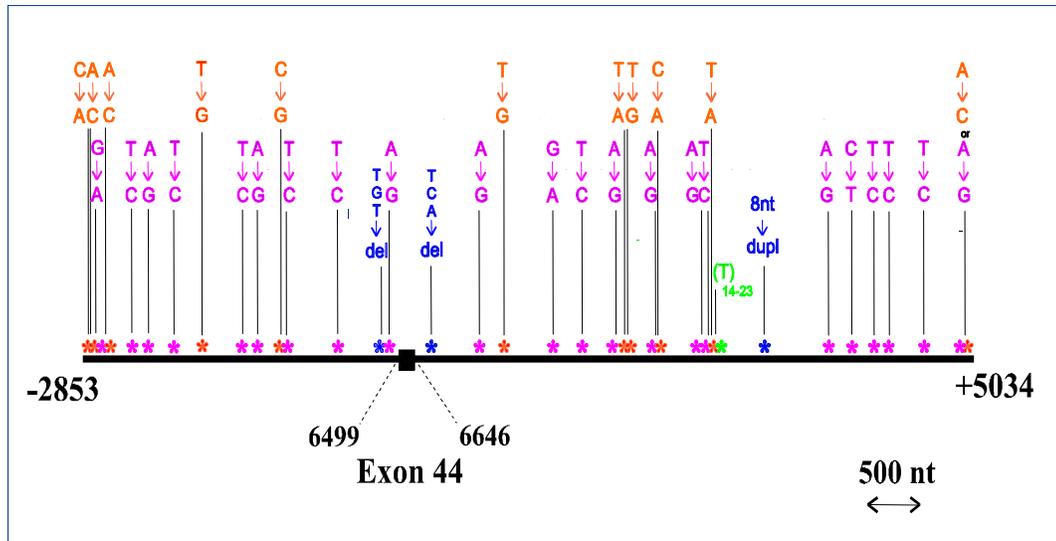


Figure 2 : Le système *dys44* et ses sites polymorphes (Zietkiewicz, Yotova *et al.* 1998).

3.2. Amplification par PCR

À partir de nos isollements d'ADN, nous avons réalisé des dilutions 12,5x qui ont ensuite été utilisées pour le génotypage. La polymérisation en chaîne du segment de *dys44* a été effectuée par amplification en duplex de courts fragments (Tableau III page 28) d'environ 100 à 300 pb chacun contenant les sites polymorphes. Le milieu réactionnel est composé comme suit : 20 μ l de mélange PCR contenant en concentration finale 1x de *Tampon* (20 mM Tris-HCl (pH 8,4) ; 50 mM KCl), 0,2 mM de dNTP, 1,5 mM de MgCl₂, 0,25 μ M de chaque amorce (*forward* et *reverse*), 0,05 unité (U) de Taq Platinum (*Invitrogen*) et 2 μ l d'ADN. Ce mélange est soumis dans un thermocycleur (GeneAmp PCR System 9700, *Applied Biosystems*) à une succession de cycles comprenant une dénaturation initiale de l'ADN à 94°C pendant 5 min, suivie de cycles (45x) de dénaturation à 94°C pendant 30 s, d'appariement à 55°C pendant 30 s, d'élongation à 72°C pendant 30 s, puis d'un cycle final de terminaison à 72°C pendant 7 min. La qualité de l'amplification a ensuite été vérifiée par électrophorèse de 3 μ l de produit PCR sur gel d'agarose (*UltraPure Agarose, Invitrogen*) à 1%.

N° Frag	Long (pb)	Amorces forward 5' - 3'	Amorces reverse 5' - 3'	Duplex
1	216	TGTTGAATTTGATAGAGGATGG	GTTTGCTGTTTTTCCTACCC	1-40
2	274	GAGGAGGGTTGGAGTTAGC	TTCCCTTTAAAAATGAATTGC	2-93
3	263	GATCATTTTGTGTGGAAGG	TCCCCTTATTTTATTAACCTACCC	3-95
4	124	GGAAAAATAATTCACAAGAGAGC	CCAAGTAATTTGCCAAAAGG	4-45
5	144	AATTCCTGAATGAAGGTAGAGG	AGCCATCCTACAGATTACCG	5-32
20	194	AATCTGTAGGATGGCTATTGG	TCAGTATTCTGGAGGTTGG	20-33
25	206	AAATAACAGCGTCAGTGTCC	AGCATTTGGTCACCTTCC	25♦
32	203	TGATCAAATCAGGTTTCTTGG	AATATTATTCTTGATCCAAAAGTGC	
33	231	TAAAGGGAAAAATTGCAACC	TATGGATCAAGAAAAATAGATGG	
38	205	GGTTCGAAGGGTGATGG	TGCTGTTACGATGCTTCC	38-86
40	256	CAATGCAGAGTTAATTGTTGG	GCAACATGAACAATTTGAGC	
45	162	GCTCAAATTTTCATGTTGC	TGATACTGCTTTTGTGAATCC	
48	214	CAAACCTGCATGCACACC	CATTGAAAATGGCAAGTGG	48-55
50	182	TGAATTTGCTACATATTTTCAAGG	AAATGCACAGATTCAAGAGC	50-64
55	286	ATCTTTTCAAAGTCTAATCTATGG	TACCAGTCATCCAAAAGAAGC	
64	231	TGCTTCTTTGGATGACTGG	GTAGTGGGGAAAAAGTAGCC	
70	215	TGATCACTAGTCAGCCTTGG	TGGACTTGTTTTCAAGTTGG	70-87
11Pol	186	TGATCACTAGTCAGCCTTGG	GCTGAACTTAATCTCCT	11 Pol♦
85	200	AAGCAATTCCTGCCTTAGC	TCTTTTTAAAAATTATCCAGTTGGCC	85♦
86	242	CGATAAGGCAAATCAAAGC	AGGCTTCAGACTCCAACG	
87	281	CTTTTGAAGATATTCCAGAGC	TCTGAGTCCAGCATTTAAC	
93	226	CACTTGGAATAGCATTTTGC	TAATCCAAGAGCCCTTATGC	
95	216	AAGCTGAACACCATGAAGG	TGGCAACTACAAACATTTTCC	

Tableau III : Liste des courts fragments *dys44* et séquences des amorces correspondantes utilisées pour la réaction de PCR

«♦» correspond à un fragment amplifié seul.

N° Frag = numéro de fragments *dys44* ; Long = longueurs

4. Hybridation ASO

Le géotypage a été réalisé par la technique d'hybridation ASO décrite dans Zietkiewicz *et al.* (1997), (Bourgeois et Labuda 2004).

4.1. Transfert sur membranes

Les produits PCR sont dénaturés par ajout de 165 μ l de solution de dénaturation composée comme suit : NaOH 0,4 M, EDTA 20 mM, 12 ml (Orange/TE 1x) d'Orange II à 0,0005%. Ces fragments d'ADN ainsi dénaturés sont transférés grâce à un dispositif de pompe à vide sur deux membranes jumelles de nylon 13/10 cm (Hybond-N⁺, Amersham), à raison de 90 μ l de produit PCR dénaturé par membrane. Chaque spot d'ADN sur une membrane donnée possède donc sa réplique exacte sur une seconde membrane. Ces membranes sont ensuite soumises dans un four (*Stratagene, UV Stratalinker 1800*) à un rayonnement ultraviolet de 120 mJ/cm² pour la fixation covalente des fragments d'ADN sur elles, puis rincées à l'eau distillée.

4.2. Marquage des sondes oligonucléotidiques

Longues de 15-mer (en moyenne) et, synthétisées par la compagnie *Invitrogen*, ces sondes spécifiques des allèles de *dys44* (Tableau IV page 30) sont marquées en 5' au radioisotope ³²P comme ci-après : 25 pmol de sonde, 1x de *forward Buffer* (70 mM Tris-HCl (pH 7,6) ; 10 mM MgCl₂ ; 100 mM KCl ; 1 mM 2-mercaptoethanol) (*GIBCOBRL*), 0,67 U de T₄ kinase, 10 μ Ci de ³²P- γ ATP (*Invitrogen*), sont complétés après 1h d'incubation à 37°C avec une concentration finale de 5 mM d'EDTA.

Frag	Positions	Oligonucléotides 1	Tm	Oligonucléotides 2	Tm
1	-2782	TCT GTG AAA CAG GTT	56	TCT GTG CCA CAG GTT	61
	-2781				
	-2777	TCT GTG CAA CAG GTT	59	TCT GTG CAA CAA GTT	56
	-2678	AGT AGC TAA AAG TGT	53	AGT AGC TAA CAG TGT	56
2	-2441	TTC TGA CTC TCA ATA	53	TTC CGA CTC TCA ATA	56
	-2291	TGA TGC AAA TGC ATG	56	TGG TGC AAA TGC ATG	59
3	-2033	CTT CGG CTA CCT TCG	64	CTT TGG CTA CCT TCG	61
4	-1760	GAA ACA AGA GAT ATT	50	GAA ACA AGA GAG ATT	53
5	-1430	ATC CAT TGC TTC TTA	53	ATC CAC TGC TTC TTA	56
20	-1313	ATG CTT GCT TGG ACA	59	ATG CTT GCT TGG GCA	61
25	-1098	ATG CCC CAG TTG ATT	59	ATG CGC CAG TTG ATT	59
	-1058	AGG TAA ACA TAC AGC	56	AGG TAA ACA TAT AGC	53
32	-599	ATA ATT TCT CTG TGG	53	ATA ACT TCT CTG TGG	56
33	-182	AAT GTT GTG TGT ACA	53	AAT GT...G TGT ACA	43
	-115	AGT TCC CTG CAT TTG	59	AGT TCC TTG CAT TTG	46
38	116	GTT GTC ATC ATT ATA	50	GTT GTC ... ATT ATA	40
40	556	ATA GAG AAA CAG CAT	53	ATA GGG AAA CAG CAT	56
45	800	GGC CTT AAA AAT TGG	56	GGC ATT AAA AAT TGG	53
48	1241	TCT ATA TCT ACC CCA	56	TCT ATA TCT ATC CCA	53
50	1485	AGG TTA ACT AGG GAG	59	AGG TTA ATT AGG GAG	56
55	1813	CCA TGT AGT ATA TAT	50	CCA TGT AGT ATG TAT	53
	1851	CCA ACT TTT ACT TCT	53	CCA AAT TTT ACT TCT	50
	1854	CCA AAT TAT ACT TCT	50	CCA ACT TAT ACT TCT	53
64	2149	AAT ATT CTC AAC CCT	53	AAT ATT CTC AAA CCT	50
	2150	AGG GCT GAG AAT ATT	56	AGG TCT GAG AAT ATT	53
70	2590	TTT AAC AGA AGC CTC	56	TTT AAC AAA AGC CTC	53
	2650	AAC CGT TGT AGC ATA	56	AAC CGC TGT AGC ATA	59
	2663	AGA CAA AAT ATA TGC	50	AGA CAA TAT ATA TGC	50
85	3144	CTG GTC TAG ATC TC.....C	61	CTG GTC TAG ATC TCT AGA TCT CC	73
86	3720	GTG AAA AAA ATC AAC	50	GTG AAA AAA GTC AAC	53
87	4099	TGG AGA ACC TCT CTA	59	TGG AGA ATC TCT CTA	56
	4128	CAT TTT TTT ACC TGA	50	CAT TTT TTT ACC CGA	53
	4264	ATC CAC CAA AAC AGT	56	ATC CAC CAA AAT AGT	53
93	4586	ACT GAA TTA TCG TCT	53	ACT GAA TTA TTG TCT	50
95	4987	CTA TGA TAG AAT GAG	53	CTA TAA TAG AAT GAG	50
		CTA TGA TAG AAT GAG	53	CTA TCA TAG AAT GAG	53

Tableau IV : Liste des Oligonucléotides utilisés pour l'hybridation spécifique des allèles du locus *dys44* par ASO

Les nucléotides en caractères gras correspondent aux mutations (ou SNP) à l'étude.
Frag = fragments *dys44* ; Positions = positions polymorphes

4.3. Hybridation et exposition

Cette étape se déroule dans un milieu réactionnel maintenu en incubation dans un système de four rotatif (*Isotemp dual – chamber incubator, Fisher*), pour l'hybridation spécifique d'une sonde à sa cible désoxyribonucléique fixée sur la membrane de nylon. Pour ce faire, chaque réplique d'une paire de membrane de nylon portant les amplicons d'ADN est mise dans un tube d'hybridation (*Fisher Scientific, 35x300 mm*) de manière à favoriser la plus grande surface d'échange entre la face portant les amplicons et la lumière du tube. Ensuite, 20 ml de solution d'hybridation composée de SSPE 1x (NaCl 150 mM, NaH₂PO₄ 10 mM, EDTA 1 mM, pH 7,4) NaCl 0,75 mM, 1% SDS, Héparine 0,8 µg/ml, Tris-HCl 70 mM, sont ajoutés au contenu de chaque tube. L'ensemble est laissé pendant 30 min en rotation modérée dans le four pour une préhybridation des membranes à 70°C. Suite à la préhybridation, un mélange sonde/compétiteur dans un rapport de 1:30 (soit 2,5 pmol de sonde et 75 pmol de compétiteur) est aussi vite que possible ajouté au contenu de chacun des tubes et la température du four baissée à 37°C. Quinze min après que la température d'hybridation ait atteint 37°C dans le four, les tubes sont vidés de leur solution d'hybridation qui est remplacée par un même volume de solution de lavage (SSPE 2x, SDS 0,1%). Après trois rinçages – à la température de la pièce – des membranes, celles-ci sont disposées côte à côte suivant la même orientation dans un sac plastique qui est scellé et placé au contact d'un film radiographique médical (*Fuji Medical X-Ray Film*), dans une cassette (Autoradiography Cassette FBXC 1417, *FisherBiotech*), pour exposition pendant six à douze heures à -80°C. Après exposition, le film est développé.

4.4. Décrochage des sondes

Après lecture du film, les membranes sont ôtées du sac plastique et immédiatement plongées pendant cinq à dix minutes dans une solution de dissociation (SDS 1%) chauffée à 95°C, ce qui permet le décrochage des sondes de leurs cibles. Ainsi ces membranes de nylon peuvent – si nécessaire – servir de nouveau à d'autres ASO.

4.5. Le microsatellite Tn

L'analyse s'effectue sur gel de polyacrylamide, suite à l'amplification par PCR du fragment 11 Pol (Tableau III page 28). 10 μ l de mélange PCR contenant en concentration finale, 1x de *Tampon* (20 mM Tris-HCl (pH 8,4) ; 50 mM KCl), 1,5mM de MgCl₂, 0,5 μ M d'amorce *forward*, 0,5 μ M d'amorce *reverse* fluorescente, 0,2 mM de dNTP, et 0,075U Taq Platinum sont soumis dans un thermocycleur à une succession de cycles. Une dénaturation initiale de l'ADN à 94°C pendant 5 min, suivie de cycles (30x) de dénaturation à 94°C pendant 30 s, d'appariement à 55°C pendant 30 s, d'élongation à 72°C pendant 30 s, puis d'un cycle final de terminaison à 72°C pendant 7 min. Au produit PCR, on ajoute 3 μ l de tampon de charge et l'échantillon est dénaturé à 95°C pendant 5 min. La dénaturation effectuée, la plaque PCR est immédiatement refroidie dans un bac à glace, et 0,4 μ l de l'échantillon déposé pour l'électrophorèse sur gel de polyacrylamide 6% pendant trois heures dans un analyseur de séquence d'acides nucléiques (NEN Model 4300 DNA analyser LI-COR, *Bioscience*).

5. Base de données et outils statistiques

L'information sur les 35 sites polymorphes (plus le poly T) de *dys44* de chaque individu participant est enregistrée dans une base de données appelée PANDORA, créée pour la vérification et l'utilisation efficace des résultats. Les allèles détectés par ASO et les variations de longueur du poly T analysées sur gel de polyacrylamide sont dans un premier temps enregistrés dans la base de données puis, définitivement archivés après contre-vérification par un autre utilisateur de la base de données. Ensuite nous procédons à la lecture directe des haplotypes des hommes. La reconstitution des haplotypes féminins est effectuée à l'aide du logiciel *Phase 2.1* (Stephens, Smith *et al.* 2001; Stephens et Donnelly 2003) disponible sur internet (<http://www.stat.washington.edu/stephens/software.html>). Les anciens haplotypes de la base de données, ainsi que les nouveaux, sont identifiés suivant une nomenclature codifiée avec une lettre (B ou b) et trois chiffres. La lettre B est associée à un haplotype retrouvé au moins une fois chez un homme ou une femme homozygote ou simple hétérozygote – à un site seulement – tandis que la lettre b renseigne sur des haplotypes jamais retrouvés chez les hommes et inférés à partir des femmes hétérozygotes pour deux sites ou plus. On étend le nom de ces haplotypes par un chiffre qui correspond à la longueur de l'allèle T_n (ou poly T_n) associé à chacun d'eux au niveau des ChrX génotypés (Tableau V page 43). Ainsi l'haplotype B001 associé par exemple aux T_n 14, 15, 16, ..., etc. sera désigné suivant cette nomenclature par les haplotypes étendus B001_14, B001_15, B001_16 etc. Les fréquences de tous les haplotypes directement inventoriés dans la base de données sont utilisées pour le calcul de certains paramètres de diversité à l'aide du logiciel Arlequin 3.11 (Excoffier, Laval *et al.* 2005) (<http://cmpg.unibe.ch/software/arlequin3/>).

Comme paramètres nous avons :

- le nombre d'haplotypes différents (k) présents dans une population,
- le nombre de sites polymorphes (S) qui, au niveau d'une population, correspond au nombre de tous les sites haplotypiques différents de ceux de l'haplotype ancestral (ici $S_{\max} = 35$ car nous ne recensons pas de nouveaux sites),
- la diversité haplotypique (G) ou l'hétérozygotie attendue au sein d'une population, qui traduit la probabilité que deux haplotypes choisis aléatoirement au sein du pool génique de cette population soient différents,
- la diversité nucléotidique H_n qui correspond, pour une sous-population donnée, au rapport entre la somme des hétérozygoties (H_i) de tous les sites polymorphes (excepté le poly T) et la longueur totale ($L = 7622$) du segment d'ADN analysé.
- l'indice de fixation ($F_{st} = 1 - H_S/H_T$) qui mesure la réduction d'hétérozygotie en fonction des différences de fréquences alléliques entre populations (H_S correspond à l'hétérozygotie moyenne des hétérozygoties observées au niveau des sous-populations analysées; H_T correspond à l'hétérozygotie totale observée au niveau de toutes les populations prises ensemble comme un groupe) (Wright 1978). Ces différences découlent généralement de l'effet de dérive génétique. Toutefois, à la dérive peuvent s'ajouter d'autres facteurs comme les recombinaisons et les migrations dont les effets modulent le degré de séparation génétique entre les populations (Jobling, Hurles *et al.* 2004).

En fonction du nombre de différences – excepté le poly T – entre toutes les paires possibles d'haplotypes, nous avons réalisé une matrice de distance. Chaque substitution mononucléotidique, délétion de triplet nucléotidique, ou duplication octanucléotidique

est comptée comme une différence, donc correspond à un site mutationnel. À une différence par site polymorphe entre deux haplotypes, nous attribuons la valeur 1 et, à une ressemblance la valeur 0. Toutes les différences polymorphiques identifiées parmi chaque paire d'haplotypes sont sommées en une valeur correspondant à la distance mutationnelle séparant ces haplotypes. L'exemple ci-dessous explique brièvement l'algorithme suivi :

Anc	CAGATATTTACTTtgtAtcaATGTATTACATTT8nACTTTA	$d_{Anc-B001} = 12$
B001	CCGATACTTACCTtgtGtcaAGGCAAGACGTTT8nACTCCG	$d_{B001-B006} = 12$
B002	$d_{B002-B002} = 0$
	
	
B006	CAGACATTACCTtgtAtcaATGTATTACATTT16nACTTIG	$d_{Anc-B006} = 4$
	
	

etc.

Les différences polymorphiques entre l'haplotype ancestral (Anc) et les autres haplotypes (B001 ou B006) sont en couleur rouge ; celles entre B001 et B006 sont doublement soulignées ; $d_{Anc-B001}$ correspond à la distance mutationnelle entre l'haplotype ancestral et B001 et ainsi de suite pour $d_{B001-B006}$, etc.

La matrice triangulaire ou carrée – au choix – obtenue à partir du décompte des différences polymorphiques au sein de chaque paire d'haplotypes est traitée avec le programme Neighbor-Joining (NJ) disponible dans Phylip 3.67 (Felsenstein 2005) (<http://evolution.genetics.washington.edu/phylip.html>). Cet ensemble de programmes nous permet de réaliser un arbre de distance (Figure 5 page 53) qui regroupe les haplotypes les plus apparentés et facilite l'identification des haplotypes parentaux pour les nouveaux recombinants.

Par ailleurs, à partir des fréquences des haplotypes et F_{st} , les analyses en composantes principales (ACP) et *multidimensional scaling* (MDS) sont réalisées à l'aide du logiciel Statistica 6.0 (*Statsoft Inc.*) pour une représentation multidimensionnelle du regroupement des sous-populations analysées.

Chapitre III
RESULTATS

1. Les haplotypes du marqueur *dys44* au sein des populations africaines et afro-américaines

1.1. Cohorte C1

Nous avons génotypé six cent cinquante-neuf chromosomes X provenant de dix-sept populations subsahariennes réparties dans trois pays co-frontaliers de l'Afrique de l'Ouest (Bénin, Burkina-Faso et Togo). Cette région a révélé un pool de soixante-huit haplotypes dont vingt-six nouveaux, pour la plupart retrouvés en une copie (B219, B220, B221, B222, B223, B224, b225, B226, B228, b230, b231, b233, B234, b235, b238, B239, B240, B241, b242, b244, b245, B246), à l'exception des haplotypes B243 et b229 (en deux copies), B227 et B232 (en quatre copies) (Figure 3 page 40). Des quarante-deux haplotypes précédemment identifiés (Labuda, Zietkiewicz *et al.* 2000; Zietkiewicz, Yotova *et al.* 2003; Xiao, Yotova *et al.* 2004; Lovell, Moreau *et al.* 2005), quinze (B001, B002, B003, B005, B008, B016, B019, B023, B039, B048, B050, B067, B143, B145 et B146) sont partagés avec les autres continents, vingt-et-un (B007, B009, B010, B012, B017, B022, B027, B032, B040, B058, B074, B088, B117, b121, B136, B151, B153, B159, B160, B164, b207) sont spécifiques au continent africain et six (B006, B011, B013, B029, B030, B166) sont très peu fréquents en Afrique subsaharienne. La distribution des haplotypes de *dys44* et leurs poly T associés au sein de notre cohorte C1 se trouve dans la figure 3 à la page suivante (voir aussi Tableau V page 43).

		Position	-2782	-2781	-2777	-2678	-2441	-2291	-2033	-1760	-1430	-1313	-1098	-1058	-699	-182	-115	116	556	800	1241	1485	1813	1851	1854	2149	2150	2590	2650	2663	2684	3144	3720	4099	4128	4264	4586	4987
		Ancestral	C	A	G	A	T	A	T	T	T	A	C	T	T	tgt	A	tca	A	T	G	T	A	T	T	A	C	A	T	T	na	8N	A	C	T	T	T	A
		Dérivé	A	C	A	C	C	G	C	G	C	G	G	C	C	D	G	D	G	G	A	C	G	A	G	G	A	G	C	A	13-24	16N	G	T	C	C	C	C/G
		Pos ID	2	3	5	8	10	12	14	15	18	20	25	30	32	33	35	38	40	45	48	50	55	58	59	64	65	70	71	72	Oligo-T	85	86	87	88	90	93	95
Famille	HaploID	Géo																																				
B006	B006					C								C															16	I							G	
B002	B002		A		C			C						C									A					C	14, 15, 16, 17, 18, 19, 20								G	
	B023		A		C			C						C									A					C	15								G	
	B048		A		C			C						C								C		A				C	16									
	B224	Af	A		C			C						C			D					C		A				C	15									
	B223	Af	A		C			C						C								C		A				C	15									
	B166		A		C			C			G			C									A					C	17									
	B145		A		C			C						C									A					C	15, 16									
	b221	Af	A		C			C						C									A					C	16								G	
	B219	Af	A		C			C						C		G							A					C	17									
	B151	Af	A		C			C						C									A					C	14				C		C	G		
	B143		A		C			C						C									A				G	C	15				C	C	C	G		
	B220	Af	A		C			C						C									A					C	22				C	C	C	G		
B050	Af	A		C			C						C									A					C	17							C	G		
B222	Af	A		C			C						C									A					C	19				C		C	G			
B010	B010	Af		C				C			G		C				D					A	G					C	15, 20, 21, 22, 24				C	C	C	G		
	b245	Af		C				C			G		C			G	D					A	G					C	17			C	C	C	G			
	b231	Af						C			G		C									A	G					C	19							G		
	B017	Af						C			G		C									A	G					C	15, 19									
	B232	Af						C			G		C									A	G				C	16										
B007	B007	Af						G	C			G					G					A	G				A	19, 20, 21										
	B058	Af						G	C			G					G					A	G			G		C	15, 17, 21				C	C	G			
	b238	Af						G	C			G									C		A	G		G		C	15			C	C	C	G			
	B234	Af						C			G	G										A	G				A	19										
	B088	Af						C		C	G		C				G					C		A	G			C	20, 21			C	C	C	G			
	b235	Af						C		C	G		C	C			G					C		A	G			C	21			C	C	C	G			
B008	B008							C			G		C				G					C	G	A	G		G	13, 14, 15, 21					C	C	C	G		
	B027	Af						C			G		C				G					C	G	A	G	G		C	14, 15				C	C	C	G		
	b207	Af						C			G		C				G		G			C	G	A	G	G		C	14				C	C	G			

(début Figure 3)

		Position	-2782	-2781	-2777	-2678	-2441	-2291	-2033	-1760	-1430	-1313	-1098	-1058	-599	-182	-115	116	556	800	1241	1485	1813	1851	1854	2149	2150	2590	2650	2663	2694	3144	3720	4099	4128	4264	4586	4987	
		Ancestral	C	A	G	A	T	A	T	T	T	A	C	T	T	tgt	A	tca	A	T	G	T	A	T	T	A	C	A	T	T	na	8N	A	C	T	T	T	A	
		Dérivé	A	C	A	C	C	G	C	G	C	G	G	C	C	D	G	D	G	G	A	C	G	A	G	G	A	G	C	A	13-24	16N	G	T	C	C	C	C/G	
		Pos ID	2	3	5	8	10	12	14	15	18	20	25	30	32	33	35	38	40	45	48	50	55	58	59	64	65	70	71	72	Oligo-T	85	86	87	88	90	93	95	
Famille	HaploID	Géo																																					
B005	B005							C			G		C			G		G			C			A	G	A			15, 21, 22					C	C	G			
	B039			C				C			G		C			G		G				C		A	G	A			22					C	C	G			
	B011			C				C			G		C			G		G				C		A	G	A			22					C	C	G			
	B067	Af		C				C			G		C			G		G				C		A	G	A			21					C	C	G			
	B013							C			G		C			G		G				C		A	G	A			21					C	C	G			
	B146							C			G		C			G		G				C		A	G	A			20					C	C	G			
	B040	Af						C			G		C			G		G				C		A	G	A		A	18, 19					C	G				
	B136	Af						C			G		C			G		G				C		A	G	A		A	19										
	B153	Af		C							G		C			G		G				C		A	G	A			20, 22								G		
	b229	Af						C			G		C			G						C		A	G	A		C	A	19									
	b230	Af						C			G		C			G						C		A	G	A			17								G		
	B228	Af				C		C			G		C			G		G				C		A	G	A			19						C	G			
	b244	Af		C				C			G		C			G		G				C		A	G	A			20					C	C	G			
	B241	Af		C				C			G		C			G		G				C		A	G	A			20					C	C	G			
B164	Af		C				C			G		C			G						C		A	G	A			16, 20					C	C	G				
B243	Af		C				C			G		C			G						C		A	G	A			20					C	C	G				
B003	B003			C				C					C			G		G			C		A	G	A	G			14, 15					C	C	G			
	b242	Af		C				C					C			G		G				C		A	G	A	G		14					C	C	G			
	B016			C				C			G		C									C		A	G	A	G		15					C	C	G			
	B009	Af		C				C			G		C									C		A	G	A	G		15					C	C	G			
	B160	Af						C			G		C									C		A	G	A	G		15, 16					C	C	G			
	B159	Af				C		C			G		C									C		A	G	A	G		15					C	C	G			
b233	Af						C			G		C		C							C		A	G	A	G		15					C	C	G				
B001	B001			C				C					C			G		G			C		A	G	A				14, 15, 16, 17, 19					C	C	G			
	B019			C				C					C			G		G			C		A	G	A			G	C	15, 17				C	C	G			
	B029	In	A			C		C					C			G		G			C		A	G	A			G	C	17				C	C	G			
	B030		A			C		C					C			G		G			C		A	G	A			G		15				C	C	G			
	B032	Af		C				C			G		C			G		G				C		A	G	A			15					C	C	G			
	B240	Af		C				C					C			G		G				C		A	G	A			15					C	C	G			
B246	Af		C				C			G		C									C		A	G	A			15					C	C	G				

(suite Figure 3)

		Position	-2782	-2781	-2777	-2678	-2441	-2291	-2033	-1760	-1430	-1313	-1098	-1058	-999	-182	-115	116	556	800	1241	1485	1813	1851	1854	2149	2150	2590	2650	2663	2684	3144	3720	4099	4128	4254	4586	4987	
		Ancestral	C	A	G	A	T	A	T	T	T	A	C	T	T	tgt	A	tca	A	T	G	T	A	T	T	A	C	A	T	T	na	8N	A	C	T	T	T	A	
		Dérivé	A	C	A	C	C	G	C	G	C	G	G	C	C	D	G	D	G	G	A	C	G	A	G	G	A	G	C	A	13-24	16N	G	T	C	C	C	C/G	
		Pos ID	2	3	5	8	10	12	14	15	18	20	25	30	32	33	35	38	40	45	48	50	55	58	59	64	65	70	71	72	Oligo-T	85	86	87	88	90	93	95	
B012	B012	Af							C		C		G			G								A	G				A	17, 20						C	G		
	B239	Af			A				C		C		G			G								A	G					21						C	C	G	
	b121	Af			A				C		C		G									C		A	G					21, 22							C	C	G
	B022	Af	A			C			C					C	C		G							A	G					19, 21, 22							C	C	G
	B227	Af	A			C			C	G				C	C		G							A	G					22							C	C	G
	b225	Af	A			C			C					C	C		G							A	G			C		22							C	C	G
	B117	Af	A			C			C	G	C			C			G							A	G					21							C	C	G
	B074	Af	A			C				G	C			C			G							A	G				A	18, 19					C		C	G	
	B226	Af	A			C			C				G	C				D						A	G					22							C	C	G

Figure 3 : Structure allélique des haplotypes de *dys44* dans la cohorte C1 (n=659)

Les haplotypes sont groupés par un code de couleur en fonction de leur similitude, la spécificité géographique (Géo) de certains haplotypes étendus Bxxx_Tn au continent africain « Af » est indiquée par les poly T en rouge

Haplotypes	Benin										Burkina Faso				Togo				Total	Fréquence %
	Tn	Adja	Avizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamashaq	Kabiyè	Akebou	Ewe	Nawdam		
B001	14									1	1							2	9.56	
	15	5	3	10		3	1		4	4	7	8	2	1	3	1	3	55		
	16						1			1	1					1		4		
	17	1																1		
	19*									1								1		
B002	14												1					1	16.2	
	15	2	1	1	1	1	2	1	1	2	2	4	12	1	3	1		1	36	
	16	4	1	2		1		1	1	2	1	1	2	3		1	2	22		
	17	3	1	3	1		1	2	3	1	5	6	7	3	2		2	4	44	
	18*											1					1	2		
	19	1																1		
	20*												1					1		
B003	14	1	2	1	2					1						1	2	10	2.88	
	15						1		2	4				2				9		
B005	15	1																1	12.8	
	21									1		3						4		
	22	7	10	5	2	2		3	2	4	5	6	13	4	8	2	1	5	79	
B006	16										3		1					4	0.61	
B007	19*									1		1						2	4.55	
	20	6	1	2	1	3	1	5	3	1		1	2	1				27		
	21*		1															1		
B008	13*									2								2	3.79	
	14	4				4	1		1	1		2	2	1		1		17		
	15	3	2															5		
	21											1						1		
B009	15	2							1	3	1	1		3			1	12	1.82	
B010	15*												1					1	7.13	
	20*											1	2					3		
	21*			1		1						5	2	1	1		2	13		
	22*		1	1		1	1		3	1	3	2	6	3	3	2		27		
	24*				1				1		1							3		
B011	22												2					2	0.3	
B012	17								1	10	2				1			14	3.64	
	20*			1	1		1				3	1	1	1			1	10		

(début du Tableau V ; voir les deux pages suivantes pour le reste du Tableau V)

Haplotypes	Benin										Burkina Faso				Togo			Total	Fréquence %
	Tn	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamashaq	Kabiye	Akebou	Ewe		
B013	21								1									1	0.15
B016	15	9	1		1		2	2	1		2	5		1		1	2	27	4.09
B017	15*									1	1							2	3.03
	19*	2	4			1		2			6	2				1		18	
B019	15	1																1	0.91
	17*		2					1					1			1		5	
B022	19*	1																1	3.18
	21*											1						1	
	22*	2		1	1		2				5	4	2		1		1	19	
B023	15											1						1	0.15
B027	14	1							1									2	0.76
	15*			1		1										1		3	
B029	17*										1							1	0.15
B030	15											1						1	0.15
B032	15		1			1		1										3	0.45
B039	22										2							2	0.3
B040	18*										9							9	2.58
	19		1		1				1		4						1	8	
B048	16*															1		1	0.15
B050	17*												1					1	0.15
B058	15	1	2			1		3	3	3	3	1	1			2		17	3.18
	17*		1								2							3	
	21*											1						1	
B067	21*			3												1	1	5	0.76
B074	18*										1							1	1.21
	19										3	1			1		2	7	
B088	20*													1			1	2	2.58
	21		1		3	1				3		3	1	1	1	1	1	15	
B117	21*		1				3	1				1			1	1		8	1.21
b121	21*					1						1						2	0.45
	22*											1						1	
B136	19	4									1	1						6	0.91
B143	15					1												1	0.15
B145	15	1																1	0.3
	16	1																1	

(suite Tableau V)

Haplotypes	T _n	Benin								Burkina Faso				Togo			Total	Fréquence %		
		Adja	Ayizo	Ditammaré	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamashéq	Kabiye	Akebou			Ewe	Nawdam
B146	20*				1				1										2	0.3
B151	14*											1							1	0.15
B153	20	1																	1	0.3
	22	1																	1	
B159	15																1		1	0.15
B160	15*	3		1	2	1			1	1		1	1	1		2	1	2	17	2.73
	16*								1										1	
B164	16*											1							1	0.61
	20*	1											2						3	
B166	17*	1																	1	0.15
b207	14*											1							1	0.15
B219	17*											1							1	0.15
B220	22*								1										1	0.15
b221	16*	1																	1	0.15
B222	19*								1										1	0.15
B223	15*																1		1	0.15
B224	15*																1		1	0.15
b225	22*											1							1	0.15
B226	22*		1																1	0.15
B227	22*								1			1	1				1		4	0.61
B228	19*											1							1	0.15
b229	19*											2							2	0.3
b230	17*											1							1	0.15
b231	19*											1							1	0.15
B232	16*		1			2				1									4	0.61
b233	15*											1							1	0.15
B234	19*	1																	1	0.15
b235	21*											1							1	0.15
b238	15*	1																	1	0.15
B239	21*																1		1	0.15

(suite Tableau V)

Haplotypes	T _n	Benin								Burkina Faso				Togo				Fréquence %		
		Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamashaq	Kabiye	Akebou	Ewe		Nawdam	Total
B240	15*															1			1	0.15
B241	20*														1				1	0.15
b242	14*									1									1	0.15
B243	20*							2											2	0.3
b244	20*											1							1	0.15
b245	17*												1						1	0.15
B246	15*															1			1	0.15
Total		73	39	30	21	17	9	35	15	35	52	96	91	41	30	16	30	29	659	100%

Tableau V : Distribution des haplotypes de *dys44* et leurs poly T associés dans notre cohorte ouest africaine

* indique des haplotypes étendus Bxxx_Tn (combinaison d'haplotypes Bxxx et d'allèles du microsatellite Tn) spécifiques à l'Afrique subsaharienne.

Ici, ce sont les valeurs des fréquences relatives de chaque haplotype Bxxx retrouvé dans notre cohorte ouest-africaine (n=659) qui sont indiquées.

À l'échelle sous-régionale ouest-africaine (cohorte C1), les haplotypes B002, B005 et B001 sont les haplotypes les plus fréquents (Figure 4 page 47). Ces variants en plus des haplotypes B003, B008 et B016 font partie du groupe des variants du marqueur *dys44* principalement partagés avec les populations non-africaines. Au Burkina Faso et au Togo, l'haplotype B010 est aussi fréquent que B001 (25 ChrX contre 26 ChrX pour le Burkina Faso et 11 ChrX contre 9 ChrX pour le Togo). Par ailleurs, l'haplotype B007, absent du Togo, est cinq fois plus fréquent (25 ChrX) au Bénin qu'au Burkina Faso. C'est la tendance inverse qu'on observe pour les haplotypes B012 et B040 qui sont au moins quatre fois plus fréquents au Burkina Faso qu'au Bénin ou au Togo. Quant à l'haplotype B003, sa fréquence pratiquement semblable à celles de B040, B088 et B160, est plus faible que celle d'une dizaine d'haplotypes (B002, B005, B001, B010, B007, B016, B008, B012, B022, B058, B017) dans la région ouest africaine.

Hormis quelques rares haplotypes partagés entre ces pays pris deux à deux, nous pouvons noter la présence de deux nouveaux haplotypes : B227 représenté par quatre chromosomes X portés par un Béninois, un Togolais et deux Burkinabé, et l'haplotype B232 portés exclusivement par quatre hommes du Bénin.

Concernant les haplotypes spécifiques au continent africain, hormis B010 à la fréquence de 7,13% et dans une moindre mesure B007 à 4,55%, aucune discontinuité franche ne semble caractériser la huitaine d'haplotypes (B012, B022, B058, B017, B160, B040, B088 et B009) qui suivent les deux premiers par ordre de fréquence (Tableau V page 43)

1.1.1. Au Bénin

Au sein du groupe ethnique Adja, la tendance générale des haplotypes communs au reste du globe, et celle des haplotypes spécifiques à l'Afrique subsaharienne semble maintenue avec cinq haplotypes majeurs (B001, B002, B005, B008 et B016) qui cumulent 55% des soixante-treize chromosomes analysés dans cette ethnie. Le tiers des chromosomes, soit 34% caractérise chez les Adja, des haplotypes (B007, B009, B017, B022, B027, B058, B136, B153, B160, B164) spécifiques à l'Afrique. Six haplotypes dont trois nouveaux (B145, B153, B166, b221, B234, et b238), se révèlent exclusifs à ce groupe en Afrique de l'Ouest. Par ailleurs, l'haplotype B003 à 1,4% n'est représenté dans cette sous-population que par un seul chromosome (Figure 5 page 53).

Au niveau de l'ethnie Ayizo, 51% des trente-neuf chromosomes sont représentés par cinq haplotypes (B001, B002, B003, B005 et B019), avec une prédominance de l'haplotype étendu B005_22 porté par dix chromosomes, soit une fréquence de 26%. La seconde moitié des chromosomes dans ce groupe, soit dix-neuf chromosomes X, est distribuée entre douze haplotypes (dont deux nouveaux B226, B232) spécifiques à l'Afrique (Figure 5).

Les trente chromosomes ditammaris de notre cohorte sont principalement représentés par trois haplotypes majeurs B001, B002, B005 à une fréquence cumulée de 70%. Le reste de chromosomes, soit neuf ChrX, est partagé entre sept haplotypes spécifiques au continent africain (Figure 5).

Treize haplotypes sont partagés entre vingt-et-un chromosomes dans la population Fon d'où l'haplotype B001 est absent. Dans ce groupe, deux haplotypes majeurs, B067 et B088, sont chacun représentés par trois ChrX (Figure 5).

On distingue dans la population Goun deux haplotypes majeurs, B001 et B007, à raison de trois chromosomes par haplotype. Deux hommes de ce groupe ethnique sont porteurs du nouvel haplotype B232 retrouvé aussi chez un homme Ayizo (Figure 5). Sur les neuf chromosomes kotafons, quatre sont B008 et trois sont B002.

L'ethnie Mahi affiche 65% de ses chromosomes qui spécifient quinze haplotypes exclusivement africains. B007 à 14% est l'haplotype majeur suivi de B002 à 11%, puis B005 et B117 à 9%. Hormis quatre haplotypes (B001, B016, B022, B243) représentés chacun par des doubletons, le reste des haplotypes – y compris B003 – de ce groupe, n'existe qu'en une copie. Par ailleurs, en plus de B243, trois autres haplotypes nouveaux, B220, B222 et B227 ont été retrouvés au sein de cette population (Figure 5).

Les chromosomes nagos analysés se répartissent entre sept haplotypes dont B002 à 33% et B007 à 20%. Dans ce groupe aucun individu n'est porteur de l'haplotype B001 ni B003.

Dans l'ethnie Weme, bien que la tendance générale par rapport aux haplotypes majeurs B001, B002, B005, B007 et B010 soit observée, aucun chromosome B003 n'est présent. L'haplotype nouveau B232 retrouvé au sein des groupes Ayizo et Goun est aussi porté chez les Weme par un homme. On note dans ce groupe la présence d'un chromosome B013 (Figure 5).

1.1.2. Au Burkina Faso

19% des chromosomes peulhs analysés sont B012, suivi par ordre de fréquence de B002, B001, B005, B008 et B003. Par ailleurs, trois individus (un homme et deux femmes) se sont révélés porteurs de l'haplotype B006 à une fréquence de 6% égale à celle de B003 dans cette population. De plus un nouvel haplotype b242 a été retrouvé au sein du groupe ethnique Peulh (Figure 5 page 53).

Au sein de la population Rimaibe, treize ChrX spécifient l'haplotype majeur du groupe, l'haplotype B040 dont la fréquence de 14% dépasse légèrement celle de B002 à 13%. Quatre individus sont porteurs d'un chromosome B003 dans ce pool génique rimaibe d'où l'haplotype B006 est absent. Quant aux haplotypes B219, b225, B228, b229 et b231, ils sont nouveaux et exclusifs à ce groupe. Un chromosome B227, retrouvé chez un Mahi au Bénin, est aussi présent dans la population Rimaibe (Figure 5).

Chez les Mossi, 23% des chromosomes sont B002, suivi de B005 à 18% et de B001 à 9%. Aucun chromosome B003 n'est présent dans cette population au sein de laquelle quatre nouveaux haplotypes (B227, b230, b233, b235) sont détectés (Figure 5).

Environ le quart des chromosomes analysés chez les Tamasheq sont B010. Ensuite viennent par ordre de fréquence les haplotypes B002 à 22%, B005 à 10%, B001 et B008 à 5% chacun. Un chromosome B006 et un B011 ainsi que deux nouveaux haplotypes b244 et b245 sont présents dans la population Tamasheq. Quant à l'haplotype B003, aucun individu n'en est porteur dans ce groupe (Figure 5).

1.1.3. Au Togo

Avec une fréquence de 27%, l'haplotype B005 est le plus prépondérant dans le groupe ethnique Kabiye. On y retrouve deux chromosomes B003 et un seul B001. Un nouvel haplotype B241 y est aussi présent (Figure 5 page 53).

Trois copies des haplotypes B001 et B010 ainsi que deux copies de B002 et B005 sont présentes chez les Akebou dont le pool génique ne comporte aucun chromosome B003 (Figure 5).

Chez les Ewe, on note presque autant d'haplotypes différents que de chromosomes analysés. Cependant, dans ce groupe, B002 est présent en cinq copies, et six nouveaux haplotypes (B223, B224, B227, B239, B240, B246) ont été détectés (Figure 5).

Le groupe ethnique Nawdam avec ses haplotypes majeurs B002 et B005 à 17% chacun, suivi de B001 à 10% et B003 à 7% ne présente aucune spécificité particulière.

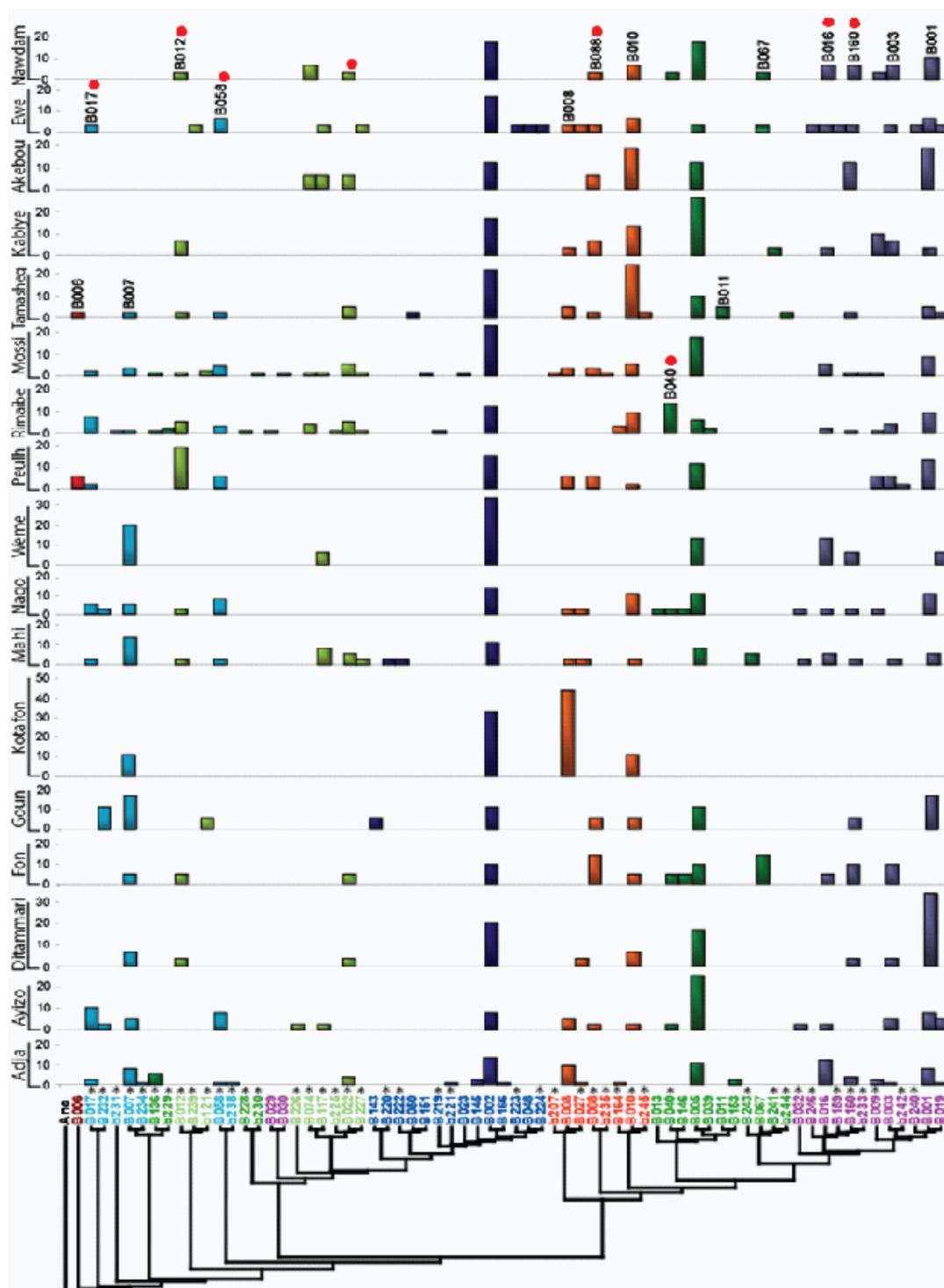


Figure 5 : Filiation des haplotypes de *dys44* et leurs fréquences relatives en % par population dans la cohorte C1

* haplotypes afro-spécifiques ; ● haplotypes ouest et centre-ouest africains virtuellement absents du reste de l'Afrique subsaharienne

Les codes de couleurs reflètent la similitude existant entre certaines séquences d'haplotype

1.2. Cohortes C1 à C6 dans un contexte global

Les figures 6 à 16, réalisées à l'aide du logiciel *Surfer 8*, donnent une vue générale (extrapolée pour certaines sous-régions) de la distribution spatiale des haplotypes majeurs du locus *dys44* sur le continent africain. Les fréquences utilisées ici correspondent aux fréquences de ces haplotypes calculées à partir de nos données et celles de Zietkiewicz *et al.* 2003 ainsi que celles de Lovell *et al.* (2005) (Tableau VIII page xxii en annexes)

1.2.1. L'haplotype B001

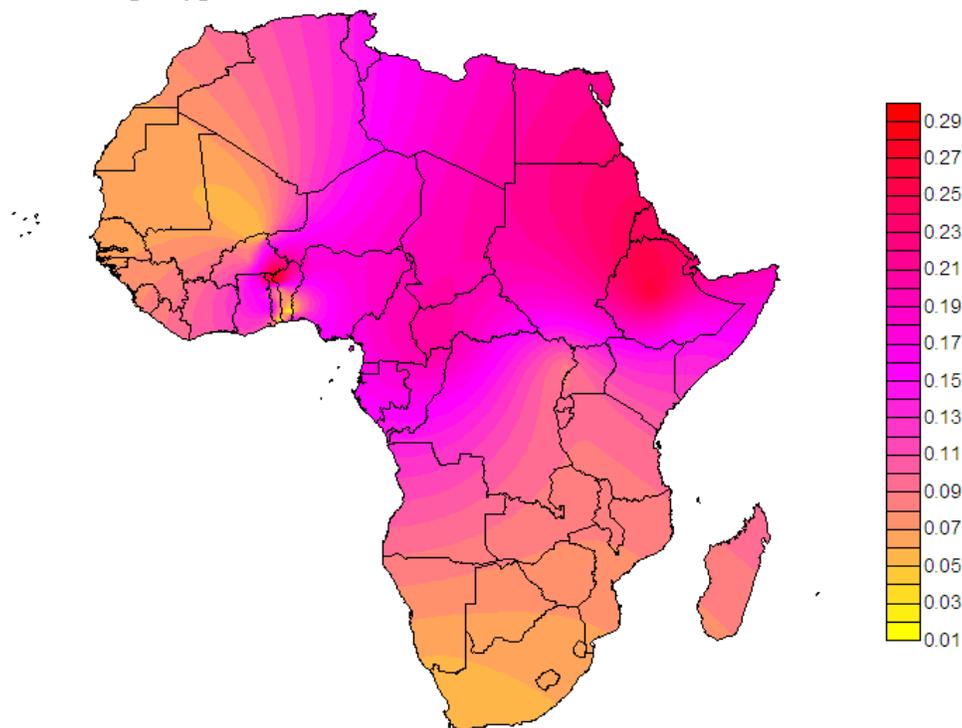


Figure 6 : Distribution spatiale de l'haplotype B001 en Afrique
Les valeurs associées au gradient de couleur dans la barre de légende correspondent aux fréquences de B001

L'haplotype B001, fréquent dans le monde, se révèle commun aux populations de la ceinture intertropicale en Afrique. Dans la majorité des cas, on le retrouve sous sa forme étendue B001_15.

1.2.2. L'haplotype B002

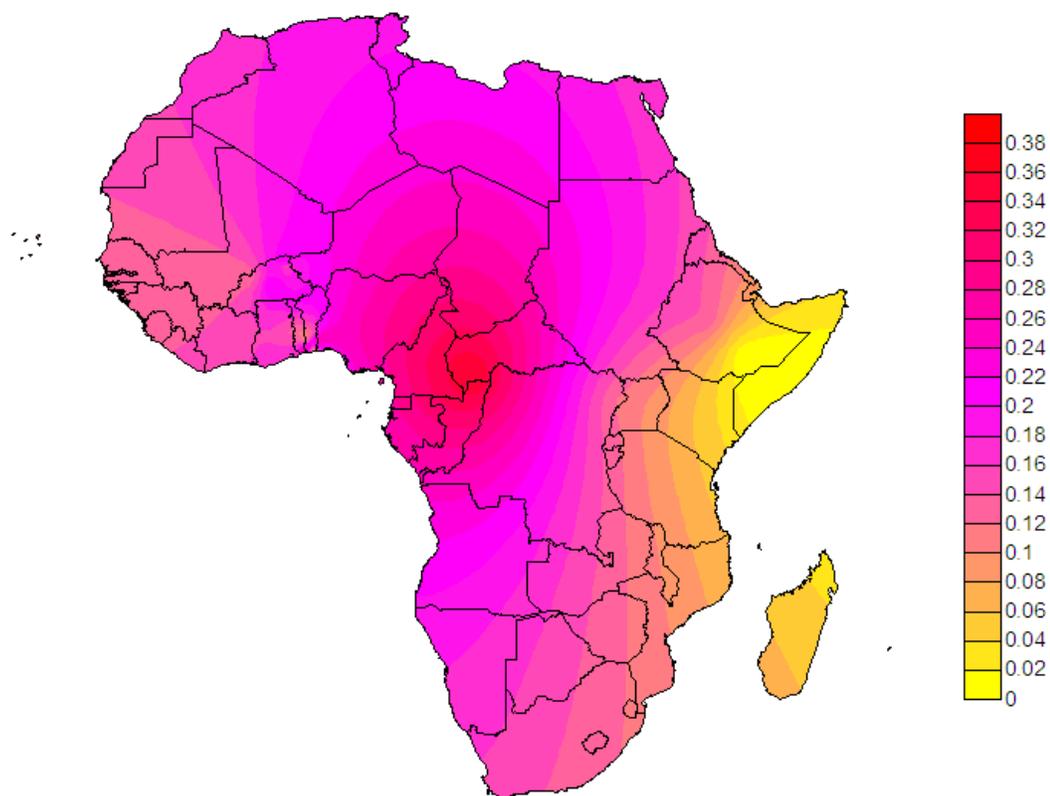


Figure 7 : Distribution spatiale de l'haplotype B002 en Afrique

Cet haplotype constitue l'haplotype majeur en Afrique d'une manière générale. B002 présente sa plus grande diversité au niveau des populations de la famille Nigéro-congolaise.

1.2.3. L'haplotype B003

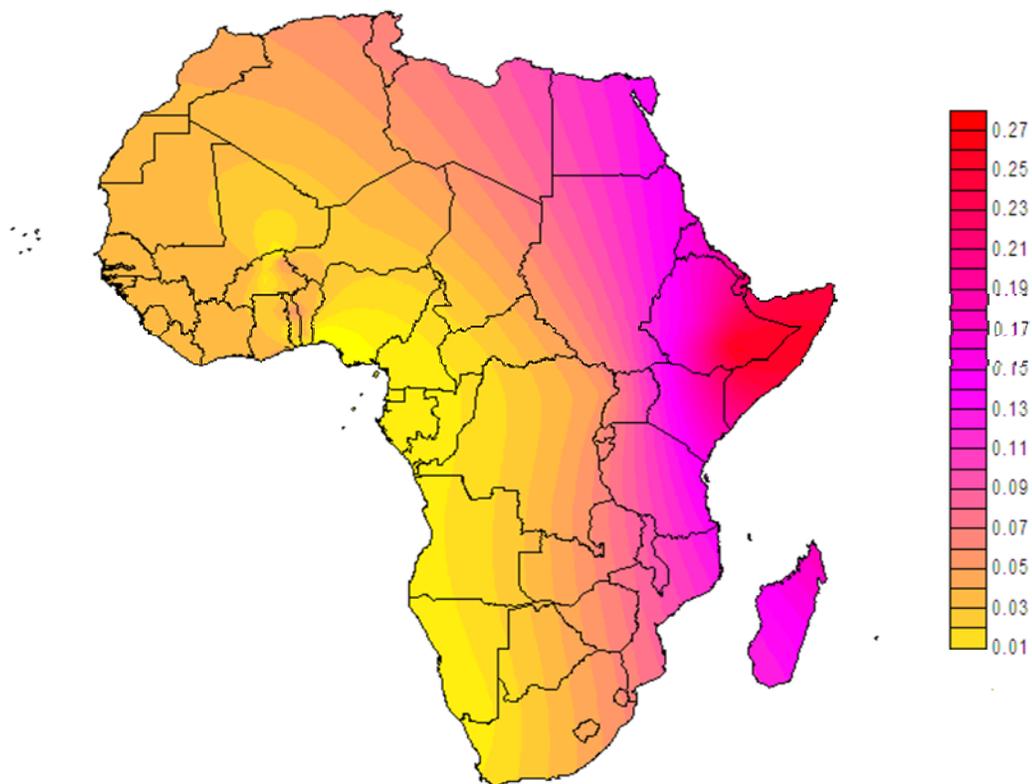


Figure 8 : Distribution spatiale de l'haplotype B003 en Afrique

À l'échelle du globe, un chromosome sur vingt environ (38/794) est B003_14 contre B003_15. La diversité de cet haplotype en Europe s'accroît de cinq chromosomes B003_22, et chez un canadien français (donnée non encore publiée) d'un ChrX B003_16. Contrairement au B003_15 qui est assez bien représenté dans la plupart des régions du globe, l'haplotype B003_14 semble inféodé aux Européens et caucasiens d'origine européenne. En Afrique au sud du Sahara par contre, nous avons retrouvé les ChrX B003_14 et B003_15 dans un rapport 14 : 23. Sur les quatorze chromosomes T₁₄, un seul est sud-africain. Les treize autres sont spécifiques aux populations ouest africaines principalement du bassin guinéen (10/13). Quant aux vingt-trois ChrX B003_15, ils

semblent exclusifs aux populations (Peulh, Rimaibe, Biaka, Mbuti, Amhara) de la ceinture tropico-subéquatorienne.

1.2.4. L'haplotype B005

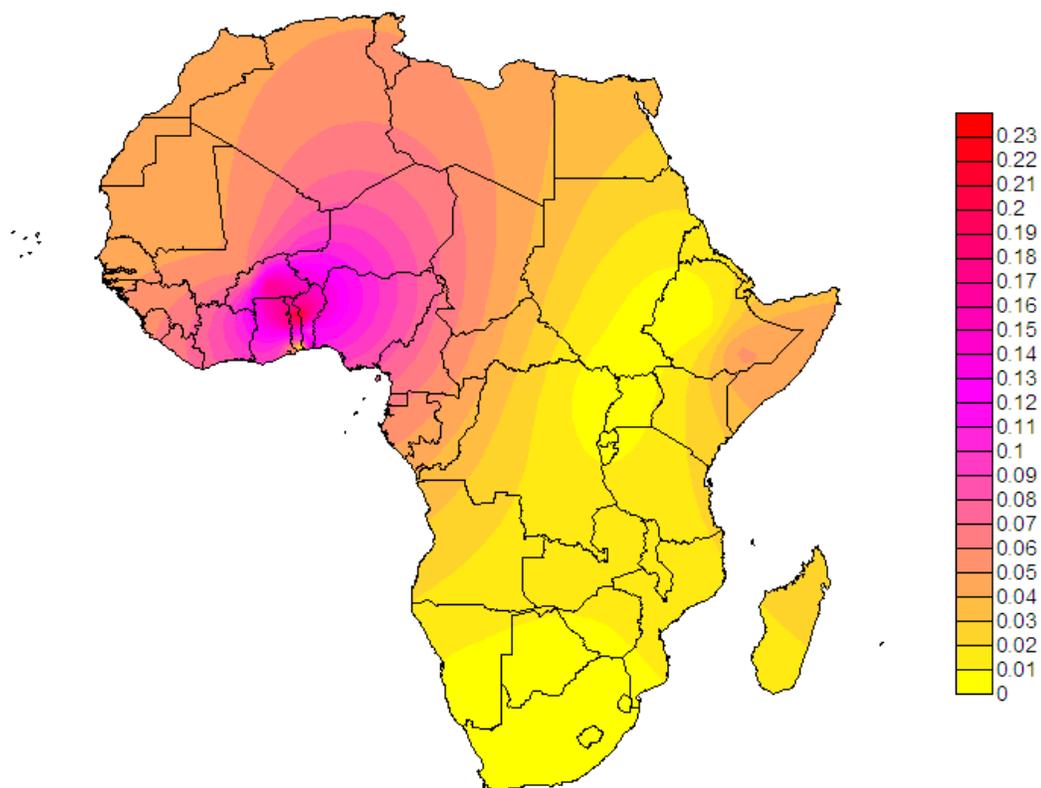


Figure 9 : Distribution spatiale de l'haplotype B005 en Afrique

L'haplotype B005 est le troisième par ordre de fréquence en Afrique subsaharienne (9,4%). Dans la cohorte C1, cette fréquence augmente à 13%. B005 principalement associé au T₂₂ en Afrique noire a été aussi retrouvé associé au T₁₅ (un exemplaire chez les Adja et Oromo), au T₂₁ (1 ChrX peulh, 1 ChrX rimaibe et 3 ChrX mossi). À l'est et au sud du continent, une seule copie de cet haplotype a été détectée chez un Oromo et un Ovambo. Hors du continent africain en Eurasie, l'haplotype B005 vient en cinquième position avec une fréquence de 6,5%.

1.2.5. L'haplotype B007

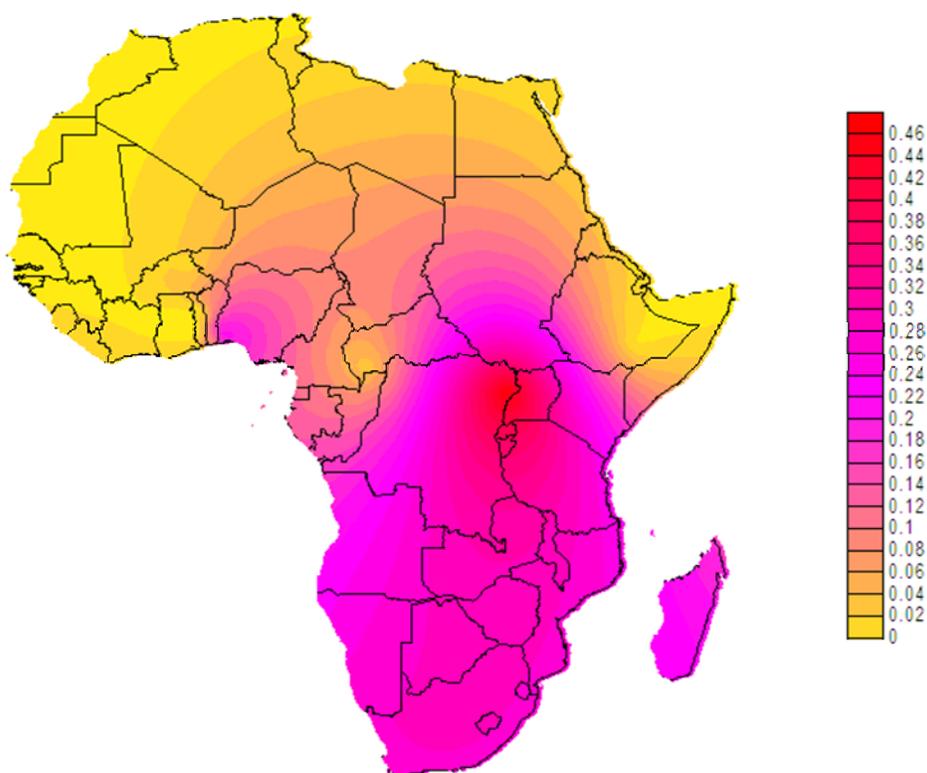


Figure 10 : Distribution spatiale de l'haplotype B007 en Afrique

Hormis un Français des Midi-Pyrénées, tous les ChrX B007 analysés à l'échelle du globe sont africains. Bien que quelques chromosomes épars aient été trouvés chez trois Mossi, un Rimaibe, un Tamasheq de notre cohorte (plus deux Rimaibe et un Mossi d'études précédentes), l'haplotype B007_20 semble spécifier les Alladanou du Bénin et les populations Pygmées de la région centrale du continent. Par ailleurs, l'association B007_19 chez un Rimaibe, un Mossi et un Weme ainsi que cinq Khoisan, semble inexistante chez les Pygmées. Au Bénin, un Ayizo s'est révélé porteur de l'haplotype B007_21. Aucun Peulh n'est porteur de ChrX B007.

1.2.6. L'haplotype B010

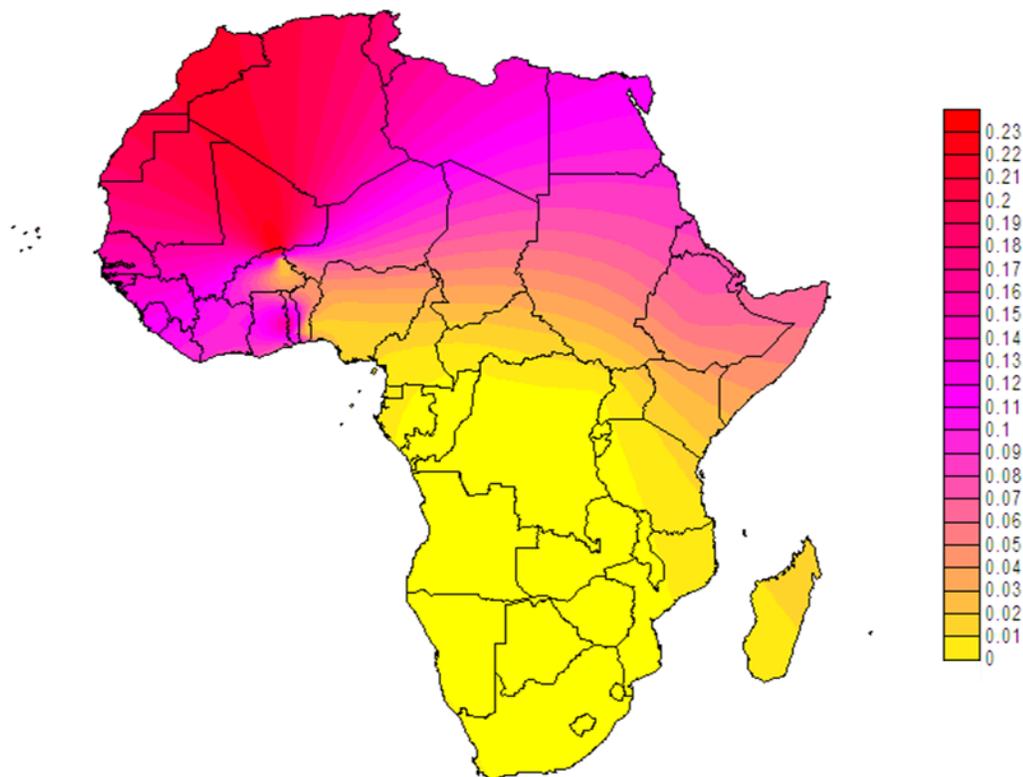


Figure 11 : Distribution spatiale de l'haplotype B010 en Afrique

Absent des populations Khoisan, Biaka et Mbuti – bien que retrouvés chez trois individus du Gabon, l'haplotype B010 est assez présent en Afrique de l'Ouest au sud du Sahara. Cet haplotype principalement associé au T₂₂ dans la sous-région (quarante-et-un ChrX) se trouve aussi associé au T₂₁, T₂₄ et T₁₅ avec respectivement dix-huit, cinq, et deux chromosomes. Deux individus Amhara (un T₂₂, un T₁₅) et une femme Peulh (B010_ T₂₂/B012_T₂₂) sont porteurs de cet haplotype

1.2.7. L'haplotype B012

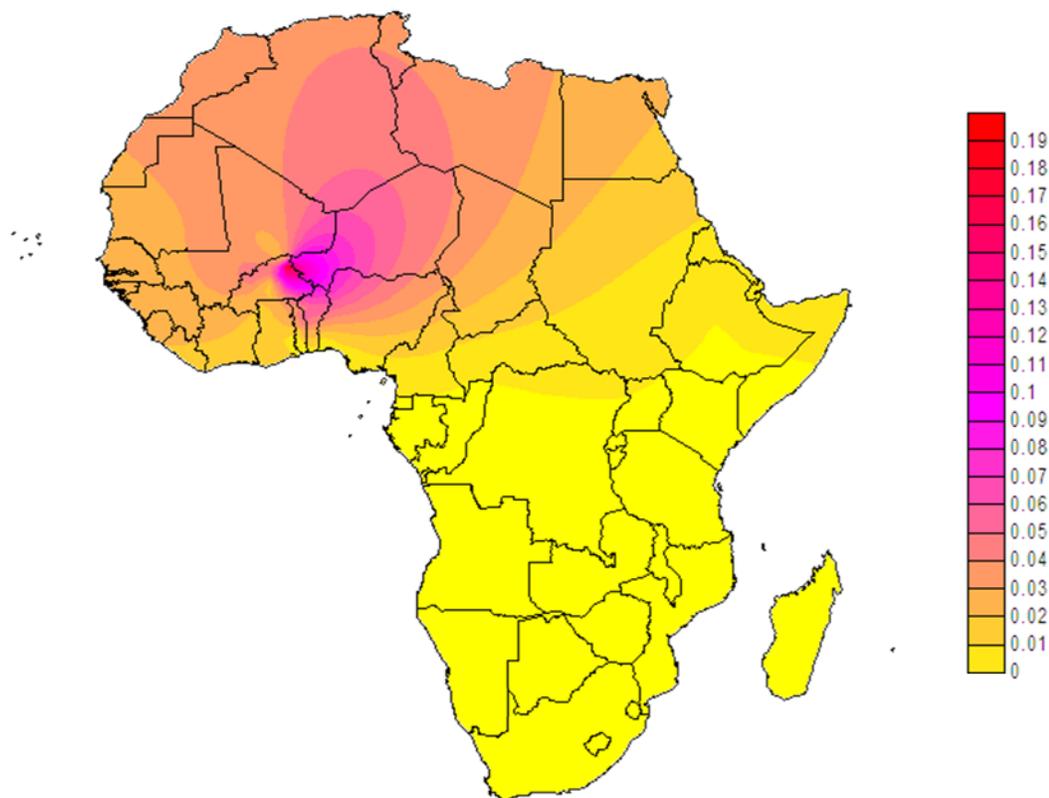


Figure 12 : Distribution spatiale de l'haplotype B012 en Afrique

La dispersion de cet haplotype, associé au T₁₇ et au T₂₀, paraît limitée à toute la région africaine à l'ouest du lac Tchad entre les latitudes de Cotonou au Sud du Bénin jusqu'à celle du détroit de Gibraltar. Quant au reste du continent, aucun chromosome B012 n'a été détecté. Les haplotypes dérivés de cette famille B012 (B022, B074, B117 etc.) présentent pratiquement la même distribution. Cependant, un chromosome gabonais B022_22, un congolais B074_19 et un chromosome khoïsan B074_19 ont été détectés sur le continent. D'autres haplotypes comme B017, B058 et B088 présentent sensiblement la même dispersion que l'haplotype B012 en raison de leur absence du pool génique des populations africaines à l'est et au sud du Sahara.

1.2.8. L'haplotype B016

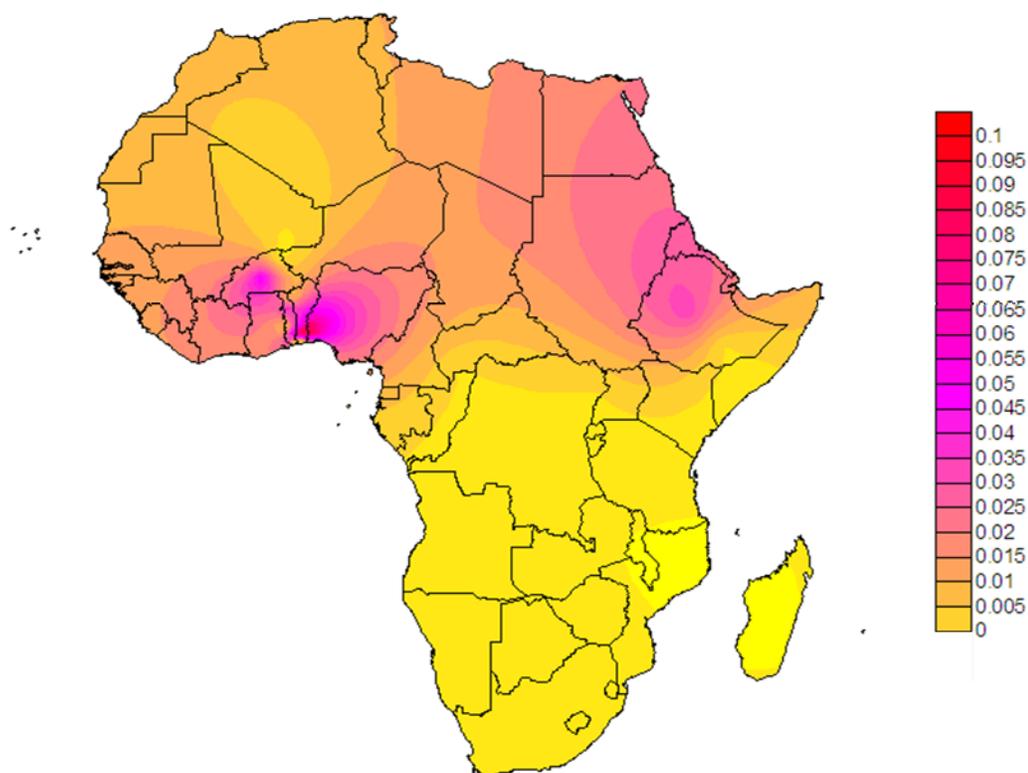


Figure 13 : Distribution spatiale de l'haplotype B016 en Afrique

Les quatre-vingt-quatorze haplotypes B016 génotypés à travers toutes les populations du globe (données de Zietkiewicz *et al.* 2003, celles de Lovell *et al.* 2005, nos données et celles non encore publiées) sont associés au microsatellite T₁₅. En Europe la dispersion de cet haplotype semble limitée à une aire géographique s'étendant du Portugal au Moyen-Orient. On le retrouve aussi chez les caucasiens d'origine européenne. En Afrique, excepté un chromosome amhara, les trente chromosomes B016 sont exclusivement retrouvés dans les populations (de notre cohorte C1) à l'ouest du bassin tchadien à l'exception des Peulh.

1.2.9. L'haplotype B022

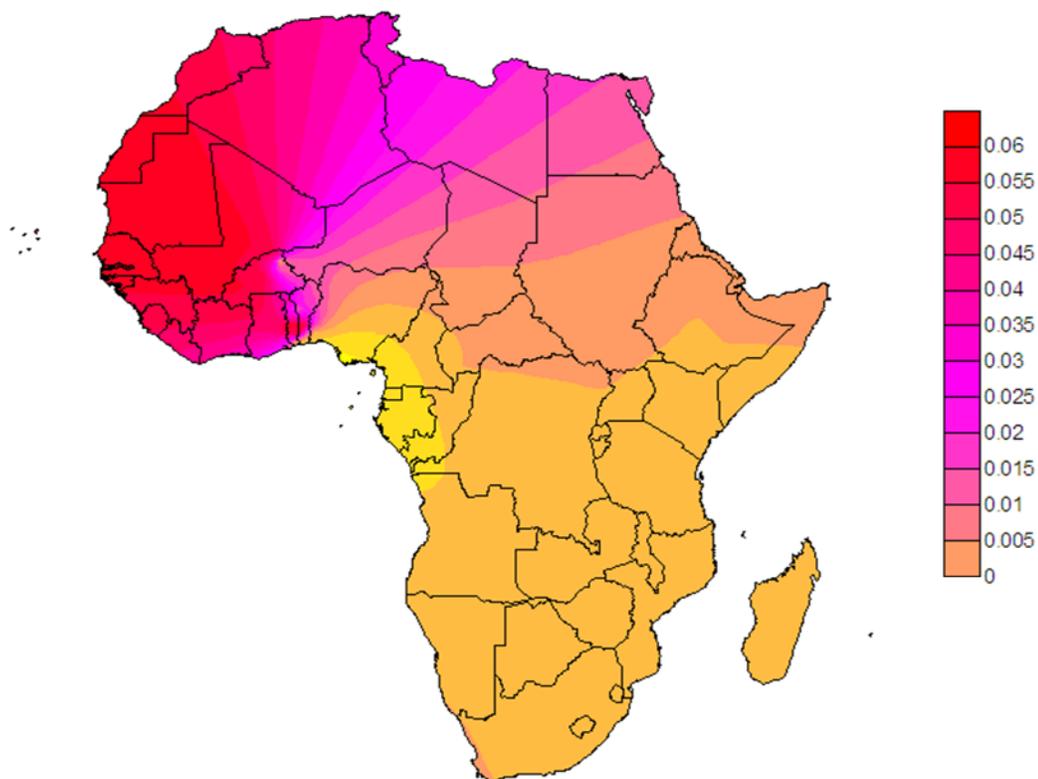


Figure 14 : Distribution spatiale de l'haplotype B022 en Afrique

L'haplotype B022 spécifique à l'Afrique subsaharienne semble caractériser uniquement les populations ouest-africaines au nord de l'équateur entre le cinquième parallèle et le tropique du cancer. Au nombre des haplotypes spécifiquement africains, cet haplotype, principalement associé au poly T₂₂, vient au quatrième rang par ordre de fréquence

1.2.10. L'haplotype B040

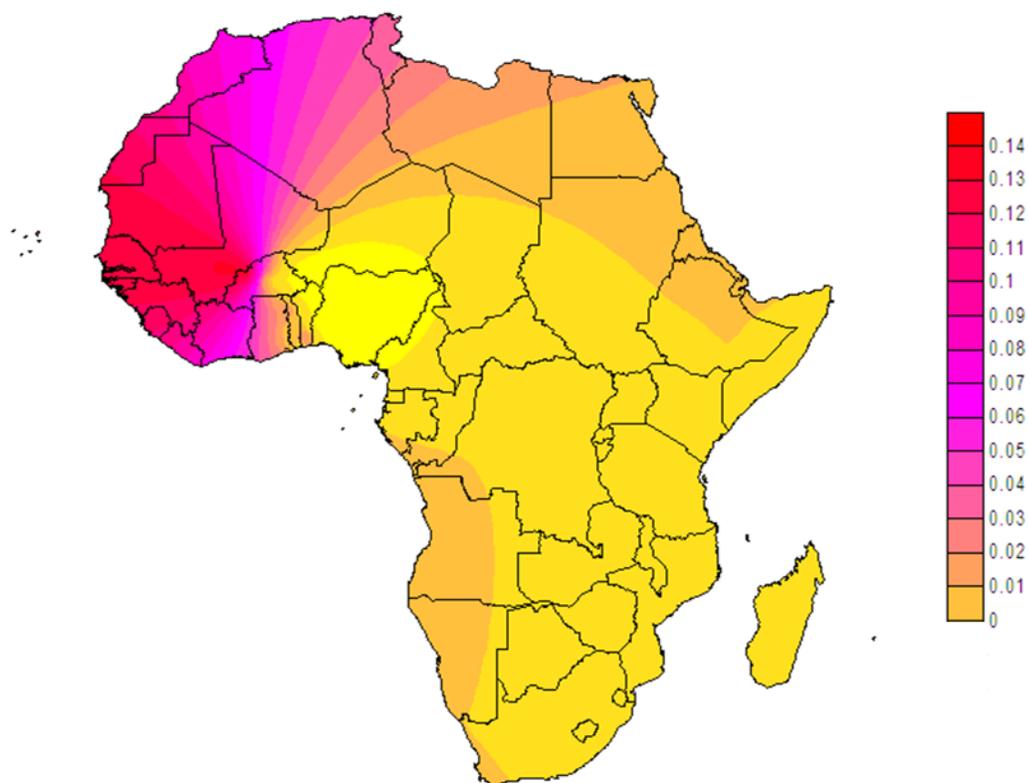


Figure 15 : Distribution spatiale de l'haplotype B040 en Afrique

Des vingt-deux chromosomes B040 répertoriés (données de Zietkiewicz *et al.* 2003, nos données et celles non encore publiées), treize ont été retrouvés dans la population Rimaibe de notre cohorte C1, dont neuf sont associés au T₁₈ contre quatre associés au T₁₉. Quant aux neuf ChrX B040 restant, cinq associés au T₁₉ ont été retrouvés dans cinq populations du golfe de Guinée ; un Mossi, un Africain Américain et un Français de Picardie en sont aussi porteurs. Par ailleurs, l'association B040_20 a été détectée chez un Africain Américain. Il faut aussi noter qu'un individu Yoruba du Nigéria (Poly T inconnu) est aussi porteur de l'haplotype B040. Aucun Peulh n'en est porteur.

1.2.11. L'haplotype B160

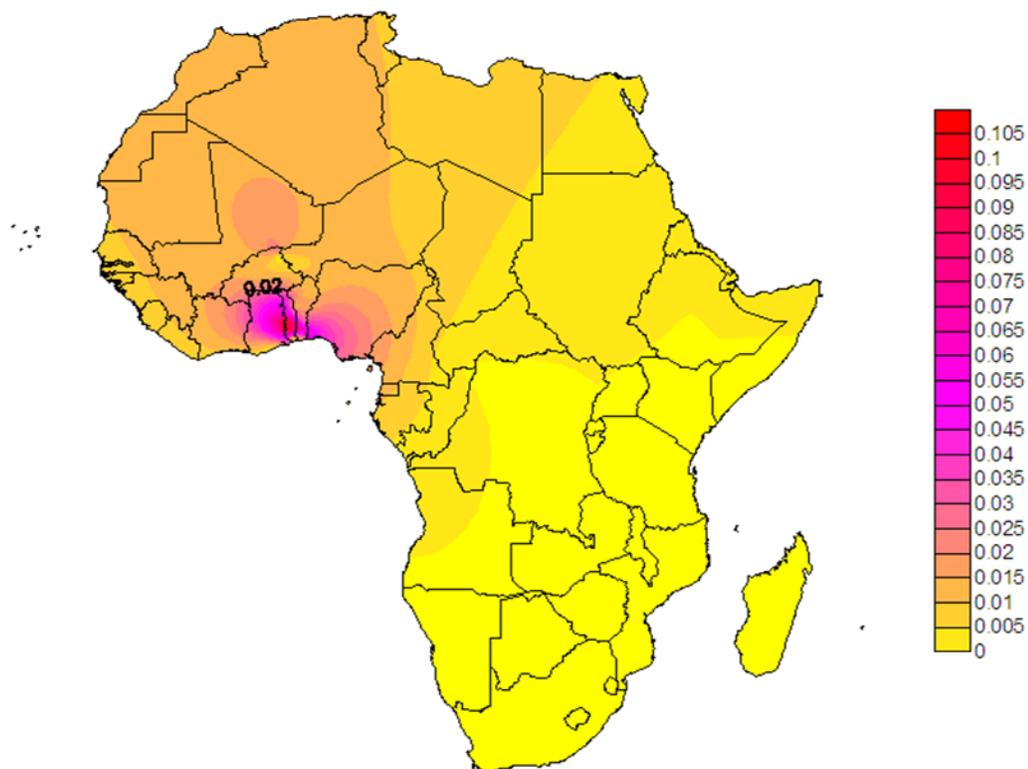


Figure 16 : Distribution spatiale de l'haplotype B160 en Afrique

Les dix-neuf ChrX B160 tous T_{15} – sauf un Mahi T_{16} – présentent en Afrique une dispersion analogue à celle de B016 à l'exception d'un X sud africain. Ici aussi aucun Peulh n'en est porteur. En Afrique du nord et au Proche-Orient, des traces de l'haplotype B160 ont aussi été détectées.

2. Le poly T et les haplotypes étendus B..._Tn en Afrique de l'Ouest

L'analyse du polymorphisme de longueur Tn au sein des populations ouest-africaines montre une diffusion assez élevée d'haplotypes associés principalement aux poly T₁₅ et poly T₂₂. Cependant cette distribution en « M » (bimodale) des Tn semble déplacée en second mode sur le poly T₂₁ chez les Fon, en T₂₀ chez les Goun (voire Mahi) et en T₁₇ chez les Peulh (voire Nago). Au niveau des Rimaibe du Burkina Faso, la distribution du Tn est trimodale centrée en T₁₉. Aussi cette population Rimaibe se singularise-t-elle – exception faite d'un chromosome Ewe – par la présence à une fréquence assez élevée (11 chromosomes) du poly T₁₈ absent de toutes les autres populations ouest africaines à l'étude.

Au niveau sous-régional, cette distribution à deux modes (T₁₅ et T₂₂) des fréquences alléliques du poly T est nettement prononcée au Bénin et au Togo. Au Burkina Faso, la fréquence relativement élevée des haplotypes associés au T₁₇ par rapport à ceux associés au T₂₂ semble séparer ces populations du Sahel de celles de la région tropicale humide et subéquatoriale que constituent les populations béninoises et togolaises. La figure 17 à la page suivante montre la distribution des allèles Tn au Togo, au Burkina Faso et au Bénin.

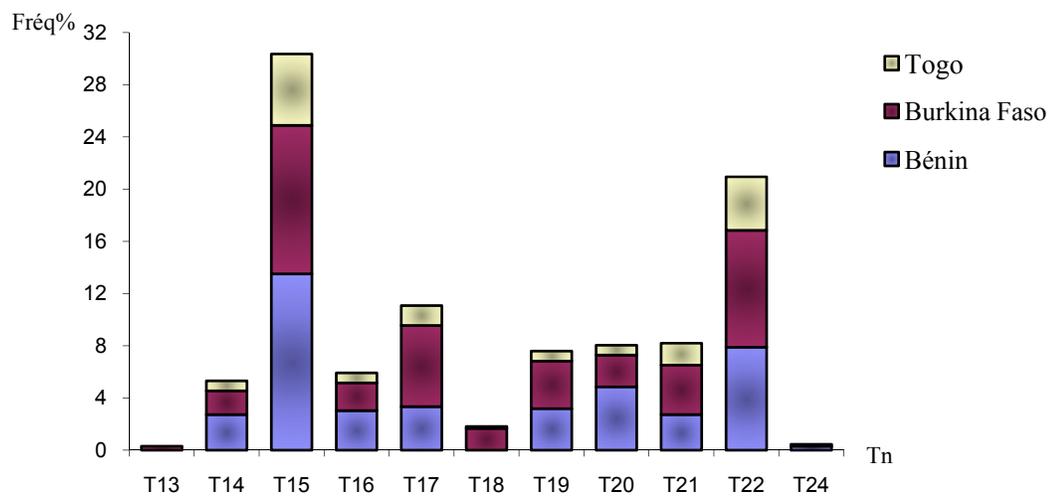


Figure 17 : Distribution du microsatellite Tn au niveau sous-régional ouest-africain (n=659)

La répétition de dix-huit nucléotides T est décrite pour une première fois dans l'étude du marqueur dys44.

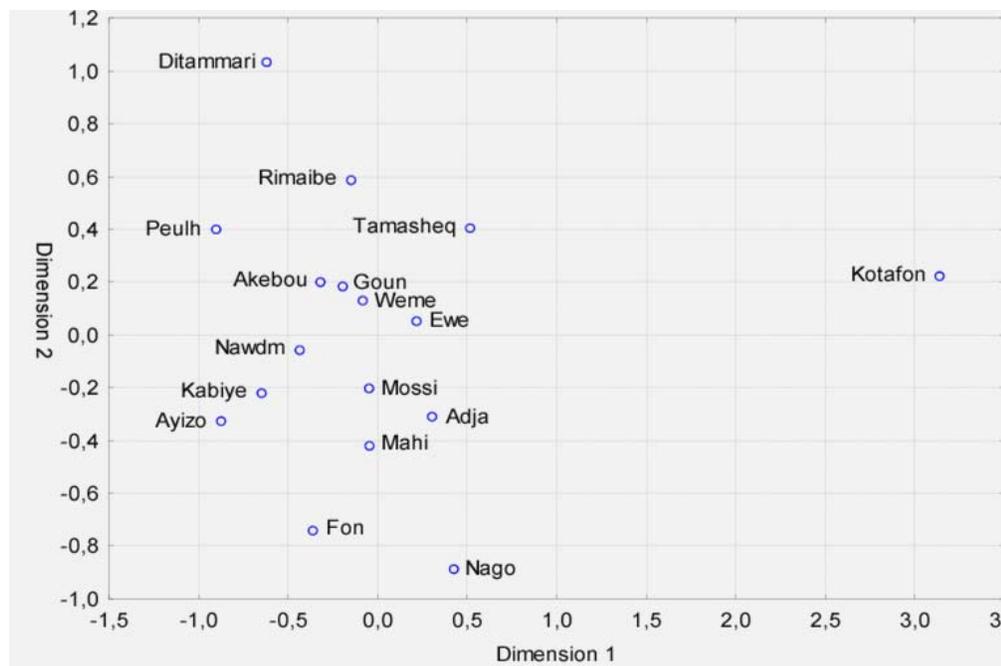


Figure 18 : MDS 1 basée sur la matrice des pairwise F_{st} des populations de la cohorte C1 (F_{st} totale = 1,86% ; $p < 10^{-5}$)

Le profil bi-dimensionnel du degré de différenciation génétique existant entre les populations de la cohorte C1 montre une séparation en première dimension des Kotafon du reste de la cohorte (Figure 18). Ceci s'explique par le faible nombre de ChrX analysés chez les Kotafon ($n = 9$), ainsi que par la faible diversité haplotypique caractérisant ce groupe ethnique. En deuxième dimension, la dispersion des groupes ethniques semble séparer les Ditammari du reste des populations béninoises. La même tendance s'observe avec l'ethnie yoruboïde Nago qui semble génétiquement moins distante des Fon que du reste des populations du Bénin. Toutefois, ce faible degré de différenciation génétique ressort significatif ($p < 0,05$) entre les Peulh et chacune des ethnies de la cohorte C1 (exceptée l'ethnie Kabiye du Togo), ce qui n'est pas le cas des Ditammari, ni des Nago par rapport au reste de la cohorte C1. Les mêmes relations entre ces différentes ethnies ressortent de l'analyse en composantes principales (Figure 19

page 68), où l'éloignement des Nago du reste de la cohorte C1, résulte principalement de la composante B016 dans cette ethnie comparativement aux autres ethnies de la région.

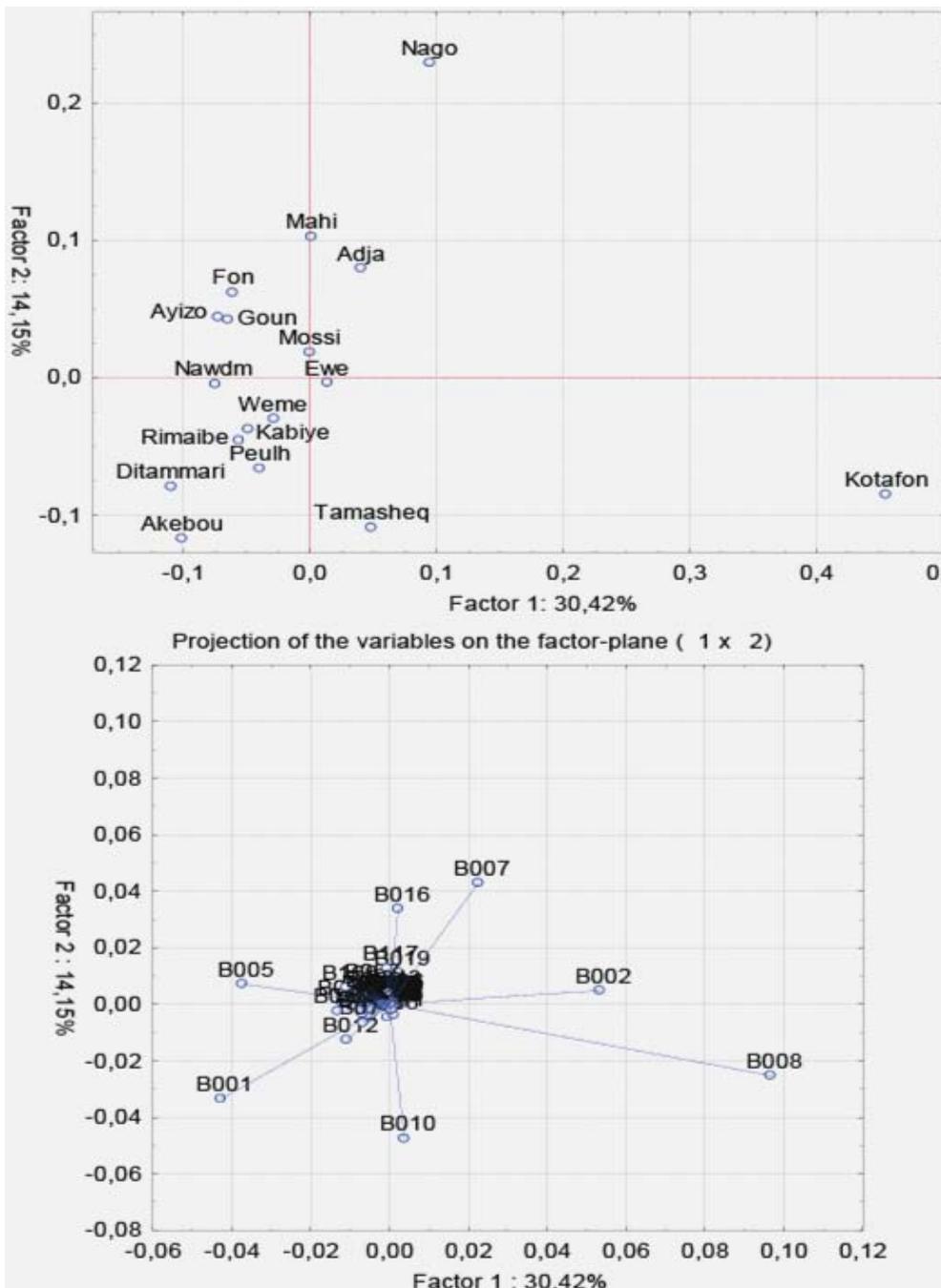


Figure 19 : Analyse en composantes principales (ACP) des populations de la cohorte C1

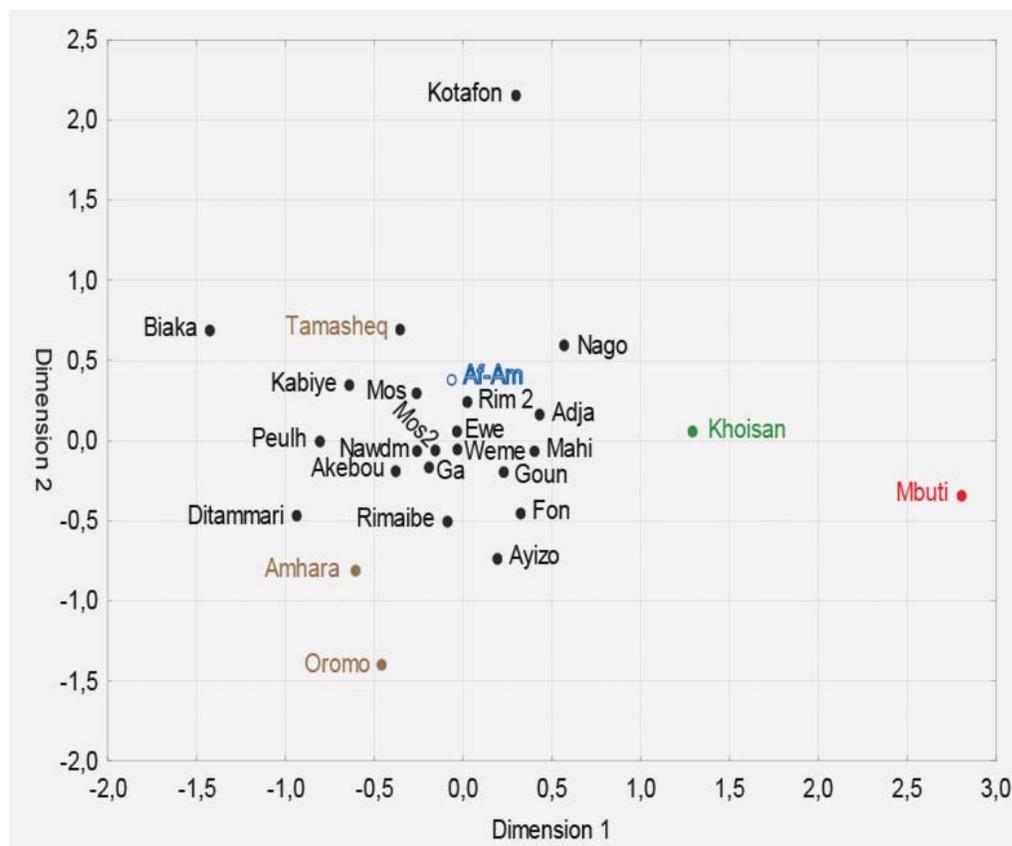


Figure 20 : MDS 2 basée sur la matrice des pairwise Fst des populations subsahariennes et afro-américaines (Fst totale = 1,5% ; $p < 10^{-5}$)

Mos = Mossi ; Rim = Rimaibe ; Af-Am = Afro-américains

- population de la famille linguistique nigéro-congolaise
- population de la famille linguistique Nilo-saharienne
- population de la famille linguistique Afro-asiatique
- population de la famille linguistique Khoisane

La figure 20 révèle une proximité génétique entre les populations de la famille linguistique Nigéro-congolaise et les Afro-américains. Les Mbuti de la République démocratique du Congo ressortent plus distants des populations ouest-africaines que les Khoisan de l'Afrique du Sud (ou les Oromo de l'Éthiopie).

Chapitre IV
DISCUSSION

1. L'haplotype B003 et ses probables recombinants

Les haplotypes du marqueur *dys44* présentent une distribution différentielle suivant la localisation géographique et/ou l'origine ethnique des populations analysées. En Eurasie où cinq haplotypes majeurs (B001, B003, B002, B006, B005 par ordre de fréquence) se partagent 82,5% de la diversité génétique au locus *dys44*, dix-huit chromosomes représentent 82,2% de cette diversité de *dys44* au sein des populations subsahariennes (Tableau VII page xv et Tableau VIII pages xxii en annexes). Sur ce continent, hormis B002, B001, B005, B007 et B010 qui varient de 18,3% à 5,6%, treize haplotypes – y compris B003 – sont retrouvés à des fréquences allant de 3,7 à 1,2%. Par sa distribution en Afrique et en Eurasie, notamment sa faible fréquence en Afrique subsaharienne (3,7%), l'haplotype B003 semble assez particulier.

En effet en Eurasie, cet haplotype est le deuxième plus fréquent après B001 (16,7%). Si nous considérons la variabilité du poly T associé, B003 est retrouvé en Eurasie sous ses formes étendues B003_14, B003_15, B003_22 avec une prédominance de l'haplotype étendu B003_15 (95% du total des B003 étendus). En Afrique subsaharienne, seuls B003_14 et B003_15 ont été détectés dans un rapport 1:2. À l'exception d'un chromosome sud-africain B003_14, la distribution de l'haplotype B003 semble restreinte aux populations sahéliennes et ouest africaines. Hormis les Peulh et Rimaibe qui présentent dans leur pool génique les deux variants B003_14 et B003_15, toutes les populations subsahariennes étudiées pour le marqueur *dys44* ne présentent que l'un d'eux (soit B003_14, ou soit B003_15). L'haplotype étendu B003_15 est retrouvé chez les Biaka, Mbuti,

Amhara et Oromo tandis qu'au sein des populations subsahariennes à l'ouest du bassin tchadien, B003_15 est retrouvé chez les Adja, Ayizo, Ditammari, Fon, Ewe, Ga et Nawdam, et B003_14 chez les Mahi et Kabiye. Dans la cohorte C2 (Tableau II page 24) B003_14 est plus fréquent que son variant allélique B003_15 (treize ChrX B003_14 contre neuf ChrX B003_15). Par ailleurs, aucun des cent vingt-et-un et quarante-et-un ChrX (données de Zietkiewicz *et al.* 2003 et les nôtres) analysés respectivement dans la population Mossi et Tamasheq n'est B003. Compte tenu de la taille de ces deux échantillons (Mossi et Tamasheq) comparativement au nombre relativement plus faible de chromosomes analysés pour les autres populations à l'étude, on s'attendrait à retrouver des chromosomes B003 au sein de ces populations. Ce n'est pas le cas. L'absence d'un haplotype d'un échantillon ne signifie pas nécessairement que cet haplotype n'existe pas dans ladite population. La probabilité p de ne pas retrouver à une fréquence f un haplotype donné dans une population de taille N est : $(1 - f)^N$ (Coia, Destro-Bisol *et al.* 2005). Pour $p = 0,05$ et $N = 121$ (respectivement 41), f serait de 2,4% (respectivement 7%). Hors pour toute l'Afrique de l'Ouest, soit 659 ChrX, la fréquence de l'haplotype B003 est de 3% (Tableau VII page xv en annexes). Il est donc possible que nous ne décelions aucun haplotype B003 chez les Mossi, et fort probable que nous ne retrouvions aucun chromosome B003 chez les Tamasheq.

Par ailleurs, compte tenu de la distribution des formes étendues B003_14 et B003_15 (Tableau VII) en Afrique subsaharienne, l'haplotype B003 semble indiquer deux courants migratoires de populations de la famille de langues nigéro-congolaises. Le variant allélique B003_15 suggérerait l'hypothèse d'un courant

migratoire sahélien, tandis que B003_14 (virtuellement absent du reste du continent excepté l'Afrique de l'Ouest) suggèrerait plutôt la migration de populations humaines le long des côtes du golfe de Guinée. Au cas où l'haplotype B003 aurait été créé au sein de populations à l'ouest (ou au sud-ouest) du bassin tchadien, cela impliquerait – l'effet de dérive aidant – que seuls les porteurs du variant B003_15 auraient pu migrer vers l'est au-delà du lac Tchad pour coloniser l'Éthiopie. Il faudrait noter que le long du golfe de Guinée, les porteurs du variant B003_14 appartiennent principalement aux populations de la sous-famille linguistique *Volta-Congo* (Tableau I page 22). Ce qui placerait l'origine possible de l'haplotype étendu B003_14 dans cette région du continent. Un peu à l'image de l'haplogroupe mitochondrial L1, qui suite à la ré-expansion des populations africaines il y a 60 – 80 ka, devint minoritaire au profit des haplogroupes L2 et L3, l'haplotype B003 semble avoir subi une dilution au bénéfice des haplotypes B001, B002 et B005 suite à ladite vague de ré-expansion (Watson, Forster *et al.* 1997; Forster 2004). Un autre évènement majeur est que la restriction virtuelle de B003 à la ceinture sahélo-subéquatorienne indique qu'en Afrique au sud du Sahara, il y a eu des flux migratoires impliquant des groupes de populations qui probablement n'ont jamais franchi en direction du sud les limites du Bénoué (entre le sud du Nigéria et le Nord du Cameroun). Cette distribution de B003 supporterait aussi la séparation linguistique faite entre les populations de la famille de langues nigéro-congolaises A et les Bantu correspondant à la famille Nigéro-congolaise B. De plus, la distribution des haplotypes étendus de B003 chez les Afro-américains (Tableau VII page xv en annexes) confirme la contribution majeure de ces populations du golfe de Guinée à la traite négrière.

Absent des populations Peulh et Tamasheq, l'haplotype B016 exclusivement associé au poly T₁₅ – donc qui n'existe que sous la forme étendue B016_15 – constitue un autre haplotype dont la distribution en Afrique au sud du Sahara reflète une spécificité aux populations de l'ouest tchadien. Ici aussi, les traces de la contribution de ces populations à la traite des noirs se reflètent par la fréquence de B016 (5,5%) chez les Afro-américains.

B016_15 qui ne diffère de B003_15 qu'aux positions -1313 et -115 (Figure 3 page 40), n'a été détecté dans le reste du continent au sud du Sahara que chez un homme Amhara de l'Ethiopie. Cet haplotype B016_15 est ensuite retrouvé sur une aire géographique couvrant les longitudes du Portugal jusqu'au Moyen-Orient où sa diversification semble s'être arrêtée. B016_15 est aussi retrouvé dans le nouveau monde.

L'analogie de séquence entre les haplotypes B016_15 et B003_15 laisse supposer que B016_15 a été créé suite à une conversion génique qui aurait affecté B003_15. Si tel a été le cas, il est probable – compte tenu de la fréquence de B003_15 en Europe – que l'haplotype B016_15 ait ses origines hors d'Afrique, et que sa présence en Afrique de l'Ouest reflète la “migration retour” de populations humaines d'Europe vers le continent noir. B016_15 a peut être bien pu provenir de B003_15 en Afrique subsaharienne ; dans ce cas l'origine de B016_15 se situerait plus probablement au sein des populations porteuses de l'haplotype B003_15 (Rimaibe, Mbuti, Amhara, Oromo). Cependant B016_15 est principalement retrouvé – comme dans le cas de B003_14 – au sein des populations de la sous-famille *Volta-Congo*. La création de B016_15, sur la base de B003_14, nécessiterait non seulement des événements de recombinaison (ou de conversion

génique), mais aussi l'insertion d'un nucléotide T au niveau du microsatellite Tn. Il est aussi envisageable que B016 ait été engendré en Ethiopie, et qu'il ait divergé vers l'ouest du continent et vers le nord hors du continent par la péninsule du Sinaï vers le Moyen-Orient, puis vers l'ouest en Europe. On s'attendrait à en trouver les traces au sein des populations de la ceinture sahélo-subéquatorienne, et non uniquement en Afrique de l'Ouest. L'absence de B016_15 du pool des Peulh dont les origines remonteraient en Égypte (Dupuy 1999), et Tamasheq de la sous-famille des Berbères montre que l'introduction – ou la naissance – de cet haplotype ne s'est probablement pas effectuée via ces populations en Afrique subsaharienne. Bien que nos données ne couvrent pas toute l'Afrique de l'Ouest depuis les limites de la Mauritanie, il est possible que la distribution actuelle de l'haplotype B016_15 reflète, entre l'Europe et l'Afrique de l'Ouest, des mouvements de porteurs de ce variant du locus *dys44*, via les Canaries et/ou le détroit de Gibraltar. Notons qu'il subsistait des points d'ombre sur l'origine africaine de l'haplotype B003 (Zietkiewicz, Yotova *et al.* 2003). Cette ambiguïté existe aussi quant à l'origine de B016 en Afrique au sud du Sahara. Chose curieuse, un troisième variant du marqueur *dys44*, l'haplotype B160, monomorphe lui aussi au niveau du microsatellite Tn, (dans le cas d'espèce, le T₁₅), est virtuellement absent du reste du continent – et du globe – hormis l'Afrique de l'Ouest. L'analyse de la séquence de cet haplotype suggère sa création à partir de l'haplotype B016_15 dont il ne diffère qu'à l'extrémité 5' par la position -2781 (Figure 3 page 40). Cet haplotype B160 a pu naître suite à un croisement entre porteurs de ChrX B016 (ou B003) et porteurs de chromosomes B005 (le deuxième plus fréquent haplotype en Afrique de l'Ouest). Si en Afrique de l'Ouest, le

brassage entre les populations de cette sous-région subsaharienne a pu servir à engendrer l'haplotype B160_15, en Europe où B003_15, B016_15 et B005 sont relativement assez fréquents, quel(s) serai(en)t le(s) processus à la base du faible nombre de copies B160 retrouvées en Afrique du nord et au Proche-Orient ? Si nous admettons que l'expansion hors d'Afrique des hommes anatomiquement modernes s'est effectuée de l'Ethiopie vers l'Eurasie, via des porteurs de chromosomes B003 et B005 (entre autres), c'est qu'il y a probablement eu peu de croisement entre ces deux groupes qui aurait pu donner naissance aux chromosomes B160, pour plusieurs raisons: l'une d'elles seraient qu'en Afrique les porteurs de B003 et B005 constituaient des pools distincts de populations dont les migrations hors du continent, via l'Ethiopie, ont eu lieu suivant différentes vagues et/ou différentes voies ; B003 ayant probablement emprunté la voie du nord. Plus on s'éloigne de la vallée du Rift en direction du sud vers les populations de la Papouasie-Nouvelle-Guinée, plus la fréquence de l'haplotype B003 diminue. Par contre, du Proche-Orient en direction de la mer du Nord et du détroit de Behring, la distribution de B003 semble en faveur de flux migratoire(s) de populations humaines via la voie du nord. La présence relativement récente de l'haplotype B003 en Europe comme le suggèrent d'autres études antérieures (Zietkiewicz, Yotova *et al.* 2003) n'a pas favorisé sur ce continent la création de B160 au sein des populations Eurasiennes. L'alternative serait que l'haplotype B160 provienne de croisement entre B016 et B005. Compte tenu de la plus faible fréquence et aussi de la présence probablement récente en Europe de B016 (comparativement à B003), il est peu probable que la rencontre entre porteurs B005 et porteurs B016 ait suffisamment eu d'impact pour que l'effet de dérive

génétiq ue en garde des traces au Proche-Orient et/ou en Afrique au nord du Sahara. Les cinq copies B160 retrouvées dans la zone géographique couvrant l'Afrique du Nord et le Proche-Orient résulteraient donc de métissage de ces populations aux populations subsahariennes (autrement ouest-africaines).

Sur la base de la similitude entre les séquences, les nouveaux haplotypes b242_14 et b233_15 – retrouvés chacun en une copie – respectivement génotypés chez une femme Peulh et une femme Mossi, ont été regroupés avec l'haplotype B003 (figures 3 et 5). L'haplotype b242_14 qui ne diffère de B003_14 que par la transition A → G en position 556 (mutation 40) proviendrait d'une double recombinaison entre les haplotypes B003 et B005. Quant à l'haplotype b233_15, il diffère de l'haplotype B160_15 d'une mutation : la transition T → C en position -599 (mutation 32) spécifique à l'Afrique subsaharienne. Cette substitution de l'allèle T au niveau de B160_15 par l'allèle C sur b233_15 résulte probablement d'une double recombinaison entre les haplotypes B160 et B022 retrouvés pratiquement aux mêmes fréquences en Afrique de l'Ouest.

2. Haplotypes spécifiquement subsahariens en général présents hors du continent sous des formes recombinaées

La dispersion ouest-africaine de l'haplotype B022 – virtuellement absent du reste du continent – et sa variabilité allélique au niveau du poly T laissent envisager une histoire évolutive relativement ancienne de cet haplotype. B022 principalement associé au poly T₂₀ existe aussi sous d'autres combinaisons alléliques chez un Ga (B022_15), un Adja (B022_19), un Rimaibe (B022_21) et un Gabonais (B022_22). Les haplotypes B022_19 et B022_21 ont pu être créés par délétion ou addition d'un seul nucléotide T de l'allèle majeur B022_20, reflétant le modèle de saut mutationnel d'une unité nucléotidique à la fois (Ohta et Kimura 1973). Quant à l'haplotype B022_15, il résulterait plutôt d'un phénomène de recombinaison avec un haplotype lié au poly T₁₅, que d'un saut mutationnel de cinq nucléotides T. En faveur de l'ancienneté de B022, les nouveaux haplotypes b225_22 et B227_22 – respectivement retrouvés en une et quatre copies – présentent la même séquence que B022_22 à l'exception de la transversion T → G en position -1760 (mutation 15) pour B227_22, et de la transition T → C en position 2650 (mutation 71) pour b225_22. La distribution de B227_22 du Sahel à la région côtière du golfe de Guinée semble conforter l'idée d'une présence assez ancienne de l'allèle majeur B022_20 dans la région ouest-africaine.

Toujours en Afrique de l'Ouest, un autre haplotype spécifique au continent, l'haplotype B012 – absent du reste de l'Afrique subsaharienne – se singularise par la variabilité des poly T associés, à savoir T₁₇ et T₂₀. Au Sahel chez les Saharawi,

Peulh, Mossi et Rimaibe, l'haplotype étendu B012_20 prédomine sur son variant allélique B012_17 avec un effet fondateur marqué chez les Peulh où 20% des ChrX sont B012_20. À mesure qu'on s'éloigne du Sahel vers les côtes du golfe de Guinée, B012_20 devient absent – sauf chez un Weme du Bénin – remplacé par quelques ChrX B012_17 épars au sein des populations du bas Togo et du bas Bénin, présumant ainsi d'un goulot d'étranglement de flux migratoire nord-sud de populations humaines depuis le Sahel vers les côtes atlantiques du golfe de Guinée. Par ailleurs, une recombinaison entre un chromosome B022 et un des chromosomes de l'haplogroupe B012, l'haplotype b121, semble avoir produit le chromosome B239 retrouvé chez un Ewe du Togo. Cependant, ni l'haplotype B012, ni B022 n'ont été détectés dans la population Ewe. La population Adja qui vers le XVI^e siècle constituait avec les Ewe la deuxième composante de l'ère Adja-Ewe ne présente que des individus porteurs de l'haplotype B022. Toutefois, les Fon et Goun qui, comme les Adja, proviennent du bassin originel Adja-Ewe (vers la fin du XVI^e siècle) présentent dans leur pool génique les haplotypes B012 et B022 pour les Fon, et b121 pour les Goun. L'haplotype B239 pourrait donc avoir été créé au sein des populations de la composante Adja, et sa présence chez l'homme Ewe résulterait d'un métissage entre les deux composantes de l'ère Adja-Ewe.

En Afrique, B007 constitue l'un des haplotypes qui paraissent n'avoir franchi la grande dépression de la vallée du Rift que sous une forme recombinée, comme c'est le cas avec l'haplotype B058 dont il diffère à l'extrémité 3' de deux sites polymorphes : les sites 2590 (mutation 70) et 2663 (mutation 72). En effet, la

dispersion de cet haplotype semble exclusive aux populations ouest, centre et sud-africaines. La plus grande variabilité de B007 à travers ses formes étendues n'est visible qu'au sein des populations ouest-africaines où, en plus du variant majeur B007_20, sont présents les variants mineurs B007_19 et B007_21. Dans la population Biaka où seulement 2,3% des quatre-vingt-cinq ChrX analysés sont B007_20, la fréquence de 46% de cet haplotype étendu au sein des populations Mbuti, sympatriques des Biaka, indique un effet fondateur. Chez les Khoisan de l'Afrique du sud, même si la fréquence de B007 à 28% indique aussi un effet fondateur, cet effet est exclusivement marqué par le variant B007_19. L'absence de variabilité du poly T associé à cet haplotype B007 chez les populations du centre et sud de l'Afrique – conjuguée aux effets fondateurs séparés chez les Mbuti (B007_20) et les Khoisan (B007_19) – semble indiquer une histoire relativement ancienne de B007 au sein des populations africaines au sud du Sahara. Suite à la création des variants B007_19 et B007_20, deux groupes de populations auraient quitté le pool originel de B007 et auraient chacun évolué suffisamment longtemps pour que l'effet de dérive génétique doublé d'effet fondateur crée, chez ces populations subsahariennes, la distribution actuelle observée de l'haplotype B007. Mieux, l'haplotype B058 probablement issu d'une recombinaison B007 x B008 est – comme B008 – absent du pool des populations Biaka, Mbuti et Khoisan. Cet haplotype B058 assez fréquent sous sa forme majeure B058_15 en Afrique de l'Ouest, y est aussi retrouvé associé aux poly T₁₇ et T₂₁. Malgré son absence du pool Biaka, Mbuti et Khoisan, B058 existe en une copie chez un Oromo et un Palestinien. À l'image de B058, l'haplotype B008 est

présent en Afrique de l'Ouest et après la vallée du Rift jusqu'à l'Eurasie. En plus de B007, la grande variabilité des poly T associés aux haplotypes B008 et B058 en Afrique de l'Ouest par rapport au reste du continent – voire le reste du globe – suppose dans l'ouest africain une coévolution de B007 et B008 pour la création relativement ancienne de B058. Ensuite le mouvement de groupes humains, porteurs principalement de B008, le long de la ceinture du Sahel jusqu'à l'Eurasie, aurait servi au peuplement de ces régions.

La diversité du pool de gènes ouest-africains se trouve augmentée des haplotypes B017 et B088 avec leurs variants majeurs B017_19, B088_21, et mineurs B017_15 et B088_20. Ces haplotypes aussi, virtuellement absents des populations Biaka, Mbuti, Amhara, Oromo et Khoisan semblent au nombre des chromosomes à l'origine des nouveaux haplotypes b235_21, b238_15 et B234_19 détectés en Afrique de l'Ouest. B235_21 résulterait d'une substitution T → C en position -599 (mutation 32) de chromosomes B088, b238_15 d'une transition T → C en position 1485 (mutation 50) sur des chromosomes B058, et B234_19 d'une conversion génique dans la partie proximale de l'haplotype B007 (Figure 3 page 40).

Deux autres haplotypes spécifiques à l'Afrique subsaharienne, dont la distribution dépeint une histoire relativement ancienne des populations à l'ouest du bassin tchadien, constituent les haplotypes B010 et B040. L'haplotype B010 par sa fréquence élevée et sa grande variabilité au sein des Rimaibe, Mossi et Tamasheq, semble s'être diversifié de cette région du continent vers les autres populations subsahariennes où il est beaucoup moins fréquent et peu variable au niveau du

microsatellite Tn (Tableau VII page xv en annexes). Chez les Tamasheq du Burkina Faso, environ un chromosome sur quatre est B010. De plus dans cette population, tous les variants de B010 – excepté B010_24 – sont présents, ce qui suppose chez les Tamasheq – voire les populations berbères – une histoire ancienne de cet haplotype par ailleurs quasi absent des chromosomes Peulh (une copie) et inexistant chez les Biaka, Mbuti et Khoisan. Le seul chromosome B010 retrouvé chez une femme peulh – porteuse aussi du chromosome peulh B012 – proviendrait d'un métissage avec les populations Tamasheq ou Rimaibe (ou Mossi peut-être) avec lesquels ils partagent le nord du Burkina Faso. Il est probable que cette copie de B010 chez la femme peulh provienne d'une femme Rimaibe, ces derniers ayant été au XIXe siècle sous la domination des Peulh. Que ce fût en Afrique ou hors de ce continent, suite à l'annexion d'un territoire, les vainqueurs faisaient non seulement main basse sur les richesses des autochtones, mais aussi sur les « belles filles » avec lesquelles ils pouvaient avoir des unions fécondes, dans le cas d'espèce, « la fille » Peulh porteuse des chromosomes X B012 et B010. Chose intéressante, la fréquence du nouvel allèle caractérisant l'haplotype B010 (soit la délétion du triplet TCA en position 116) est de 6% en Afrique (n = 1118 chromosomes analysés au total), ce qui place cet allèle dans la gamme de fréquence de 5% à 15% caractérisant les allèles nouveaux d'âges estimés de 100 - 200 ka (Zietkiewicz, Yotova *et al.* 1998). Rappelons qu'au nombre des haplotypes spécifiques au continent noir (n = 1118), l'haplotype B010 est le deuxième – après B007 – par ordre de fréquence. L'origine probable de cet haplotype se situant en Afrique de l'Ouest, l'occupation de cette sous-région du continent au sud du Sahara remonterait à 100 ka environ. Bien que cet âge soit discutable, notamment

pour les Tamasheq qui sont de la famille de langues afro-asiatiques, l'occupation de l'Afrique de l'Ouest par les populations humaines se révèle plus ancienne que le néolithique suggéré par les données paléanthropologiques. Selon Huysecom (2007), les traces d'occupation humaine de cette sous-région dateraient de 150 ka environ.

Deux copies du variant majeur B010_22 ont été génotypées chez les Amhara tandis qu'une copie du variant B010_15 – probablement recombiné – a été détectée chez un Oromo. D'autres recombinaisons de l'haplotype B010 semblent avoir servi à la création de nouveaux haplotypes en Afrique de l'Ouest. C'est le cas de B226_22 (combinaison B002 proximale et B010), b231_19 (combinaison B006 et B017), B232_16 (combinaison B017 et B002 distal), b245_17 (transition A → G en position -115 sur l'haplotype B010).

En ce qui concerne l'haplotype B040, près de 70% des chromosomes de ce type retrouvés en Afrique sont rimaibes. Dans cette population, un ChrX B040 sur trois est associé au T19, tandis que les autres sont associés à un poly T, pour la première fois décrit pour le marqueur *dys44* : la répétition en tandem de 18 nucléotides T. La singularité de B040_18 chez les Rimaibe supporte l'idée d'une identité par l'ascendance (IBD) de tous les porteurs de cet haplotype étendu. L'haplotype B040, qui probablement constitue le produit de croisement entre porteurs B005 et porteurs de B012, fait partie de ces haplotypes dont la présence actuelle, dans le pool génique des populations ouest-africaines, témoigne de l'intensité particulière du flux de gènes entre les populations subsahariennes à l'ouest du continent. Au Burkina Faso, chez les Peulh, la fréquence des haplotypes B012 et B005 est particulièrement élevée pour qu'on s'attende à retrouver au sein de cette population, des porteurs du ChrX B040.

Or dans le bassin de gènes peulhs, aucun haplotype B040 n'a été détecté. Dans ce pays, seuls les Rimaibe présentent le variant B040 du locus *dys44*, à une fréquence de 13,5%. Ceci pourrait indiquer chez les Rimaibe, une cohabitation relativement plus ancienne des haplotypes B005 et B012 qui sont présents quasiment aux mêmes fréquences dans cette population (soit 6,2% et 5,2% respectivement). Contrairement aux dix ChrX B012 sans doute IBD en raison de leur monomorphisme au niveau du microsatellite T₁₉, les neuf haplotypes B012 Rimaibe (Tableau VII page xv en annexes) sont partagés entre les répétitions de 19 et 20 nucléotides T, avec cinq B012_19 et quatre B012_20. Non seulement le bassin originel des B012 et B040 – voire B005 – semble se situer en Afrique de l'Ouest, mais aussi les Rimaibe du Burkina Faso paraissent constituer l'une des plus anciennes populations de cette sous-région du continent. Rappelons qu'en plus de B007, B012 et B040 caractérisent (en plus des rares recombinants) les séquences afro-spécifiques porteuses du nouvel allèle 2663A dont l'âge estimé est de 100 – 200 ka environ (Zietkiewicz, Yotova *et al.* 2003).

3. Les haplotypes majoritairement partagés entre l'Afrique subsaharienne et les autres régions du globe

Au nombre des haplotypes communs à l'Afrique et au reste du globe, B005 représente l'un des haplotypes qui se sont peut-être diversifiés à partir du bassin ouest-tchadien. En effet, contrairement au reste du continent où cet haplotype n'est présent que sous sa forme B005_22 chez les Biaka (quatre copies sur quatre-vingt-cinq ChrX), le deuxième variant (B005_21) n'a été retrouvé que chez les Mossi, Peulh, Rimaibe et un Basque. La combinaison B005_15 détectée chez un Adja, un Oromo et un Ovambo, présume plutôt d'une recombinaison que d'une mutation du poly T, comme précédemment observé chez trois Khalkha – sans doute IBD – (Zietkiewicz, Yotova *et al.* 2003). La seule copie de B005 sous sa forme recombinée B005_15 au sud du continent chez l'individu Ovambo, proviendrait d'un métissage récent de porteur – peut-être caucasien – de cet haplotype avec la population autochtone Ovambo. Toutefois, le faible nombre de ChrX ovambo (ou sud-africain) dans la cohorte africaine, recommande une certaine réserve quant à cette interprétation sur l'origine plausible du B005_15 ovambo. En Eurasie, trois copies B005_23 – sans doute IBD – ont été détectées chez trois kazakh. De fait, l'haplotype B005, sous ses formes étendues, paraît moins variable en Afrique qu'en Eurasie. Ce qui laisse supposer une origine eurasiennne de B005. Parmi les nouveaux recombinants, les haplotypes B228_19, b229_19, b230_17, B241_20, B243_20 et B244_20 reflètent le résultat de recombinaisons multiples entre les chromosomes de type B005 et ceux du type B002. Ce dernier et l'haplotype B001 – les deux haplotypes les plus fréquents en Afrique subsaharienne et en Eurasie – présentent un gradient opposé suivant le contexte continental de leur distribution.

En Eurasie – et même chez les Américains – l’haplotype B001 occupe la tête du peloton suivi de B002 (par ordre de fréquence). Par contre chez les Africains au sud du Sahara, les Afro-américains, les Papous de la Nouvelle-Guinée, l’haplotype B002 est le plus fréquent, suivi de B001. À l’exception des Rimaibe et Peulh qui présentent trois à quatre variants étendus de B001, les autres populations analysées présentent principalement le variant majeur B001_15 et parfois une seule copie de l’un des variants mineurs B001_14, B001_16 ou B001_17. L’haplotype B002 absent du pool des ChrX Oromo, mais présent sous deux à quatre variants au sein des autres populations subsahariennes, semble par sa distribution chez les Biaka et Mbuti, s’être diversifié à partir de ces populations. Pour B002 en Afrique, l’effet de dérive semble en faveur d’une meilleure diffusion du variant B002_17, tandis que hors du continent, B002_15 paraît avoir principalement contribué à la conquête de l’Eurasie et de l’Océanie par les premiers HAM. Dans ce contexte de l’haplotype B002, six nouveaux recombinants ont été détectés en Afrique de l’Ouest : les haplotypes B223_15 et B224_15 qui diffèrent de l’haplotype B048_16 asiatique respectivement par les mutations 65 (C → A en position 2150) et 38 (TCA → delTCA en position 116), l’haplotype b221_16 qui résulterait d’une recombinaison simple B002 x B017, l’haplotype B219_17 qui serait créé à partir d’une recombinaison multiple B002 x B007, l’haplotype B220 créé par recombinaison B002 x B005/B010, et l’haplotype B222 qui constituerait le résultat d’un croisement B002 x B074.

4. La “migration retour” de l’Eurasie vers l’Afrique subsaharienne

L’hypothèse, sur la base de l’haplotype B006 rimaibe (Zietkiewicz, Yotova *et al.* 2003), d’une “migration retour” de populations humaines depuis l’Eurasie vers l’Afrique subsaharienne, semble plus découler d’un métissage des Rimaibe avec les populations Berbères (dans notre étude les Tamasheq) et Peulh. En effet, aucun des quatre-vingt-seize chromosomes analysés dans notre cohorte C1 (Tableau II page 24) ne correspond à une séquence B006. Les quatre haplotypes B006 (tous associés au T₁₆) présents dans cette sous-région ont été détectés au sein des populations Peulh et Tamasheq. Ceci va dans le sens d’autres études qui proposent que, suite à l’introduction depuis le Proche-Orient des haplogroupes mitochondriaux U6 et H en Afrique du nord il y a 20 - 50 ka au sein des populations Berbères, ces séquences mitochondriales auraient été distribuées en Afrique subsaharienne via les populations pastorales Peulh dont les origines remonteraient à l’Egypte (Tauxier 1937; Macaulay, Richards *et al.* 1999; Salas, Richards *et al.* 2002). Une autre preuve du métissage des populations Peulh aux Berbères est la présence, dans leur pool de gènes, de séquences mitochondriales U5 analogues à celles caractérisant les populations Berbères du Maroc, du Sahara occidental et de la Tunisie (Brakez, Bosch *et al.* 2001; Plaza, Calafell *et al.* 2003). Bien que classés dans la famille linguistique Nigéro-congolaise, les Peulh constituent actuellement l’une des rares populations dont le pool de génique actuel reflète le mieux l’influence génétique des autres populations des familles Afro-asiatiques, Nilo-sahariennes et Nigéro-congolaises dont ils ont adopté les modes de vie – en plus des leurs. Ce n’est donc pas surprenant que, quel que soit le

marqueur génétique, les Peulh se distinguent des autres populations subsahariennes par la présence dans leur génome de variants génétiques spécifiant les populations eurasiennes, comme c'est le cas de l'haplotype B006 dans notre cohorte C1. Si la présence de l'haplotype B006 résulte d'une "migration retour" de populations eurasiennes vers cette région, sa distribution au sein des populations de cette région serait détectée non seulement au niveau des Peulh mais aussi au sein d'autres populations de la famille Nigéro-congolaise, compte tenu des évidences de brassages historiques entre ces Peulh et leurs voisins sympatriques. Au nord du Cameroun, l'étude du chromosome Y a révélé à des fréquences variables – quoique très élevées parfois – son variant R1-173 (notamment l'haplotype 117) appartenant à l'haplogroupe M9. Ce sous-groupe R1-173 non africain à l'origine, compte tenu de sa distribution (fréquence de 6,7% à 95,2%) au sein des populations Fali, Tali, Uldeme, Fulbe, Daba et Ewondo, permit à Cruciani *et al.* (2002) de suggérer le scénario d'une "migration retour" de populations humaines de l'Asie vers l'Afrique au sud du Sahara. Cependant l'analyse du marqueur mitochondrial au sein de ces mêmes populations ne révéla que les haplogroupes L1, L2, L3, U5 et U6, contrairement aux données du R1-173 (Coia, Destro-Bisol *et al.* 2005). Ce qui n'est pas surprenant dans la mesure où, si la "migration retour" de ces populations asiatiques s'est opérée principalement par des hommes, il est peu probable que la transmission foëto-maternelle de l'ADNmt révèle des lignées asiatiques. Notons qu'au niveau de l'ADNmt, cette "migration retour" est plus supportée par l'haplogroupe M1, notamment par le sous-groupe M1c dont l'introduction via la péninsule du Sinaï en Afrique du nord-ouest

remonterait vers 30 ka environ (Gonzalez, Larruga *et al.* 2007). Cette différence dans la phyléogéographie du ADNmt et du ChrY au sein de la même population a été remarquée au sein d'autres populations comme les populations nord-africaines où la contribution maternelle subsaharienne est plus importante que la contribution paternelle subsaharienne (Richards, Rengo *et al.* 2003). Notre modèle moléculaire utilisé, le locus *dys44*, ne présente pas ce biais de sexe. Nos résultats reflètent donc aussi bien l'histoire démogénétique des hommes que des femmes. Par ailleurs, en analysant les populations du nord Cameroun pour le marqueur *dys44*, nous pourrions ne déceler aucune trace de cette lignée eurasiennne B006 au sein de ces populations, si ce processus a eu lieu principalement via les porteurs de ChrX autres que B006 comme les lignées ubiquitaires B001 et B002.

5. Diversité de *dys44* en Afrique de l'Ouest : Un reflet du brassage génétique entre les populations des trois grandes familles linguistiques africaines

La très grande diversité de l'Afrique de l'Ouest (principalement des populations du Bénin, Burkina-Faso et Togo) comparativement au reste du continent (Tableau VI page xiv) suggère, pour le locus *dys44*, une histoire évolutive relativement ancienne de ces populations. Dans l'ouest africain, la diversité moyenne des haplotypes du marqueur *dys44* de $0,91 \pm 0,03$ (comparativement à $0,85 \pm 0,04$ dans le reste du continent), témoigne aussi de l'intensité et de la profondeur temporelle du brassage génétique entre les populations de l'Afrique de l'Ouest. Non seulement la séparation de ces dernières des autres populations subsahariennes paraît très ancienne, mais aussi l'Afrique de l'Ouest semble avoir bénéficié de la vague de ré-expansion des populations humaines à l'intérieur du continent il y a 60 – 80 ka (Watson, Forster *et al.* 1997) dans l'enrichissement de son pool génétique. Cependant, cette subdivision des populations n'empêcha pas une certaine quantité de flux de gènes le long de la ceinture sahélo-subéquatorienne entre l'ouest et l'est, de l'ouest – voire centre-ouest – vers le sud, et entre l'est et le sud du continent. Ce qui expliquerait la faible valeur de F_{st} globale de 1,5% calculée pour l'Afrique subsaharienne – y compris les Afro-américains (Figure 20 page 69). Selon les analyses en composantes principales, les Biaka qui constitueraient l'une des plus anciennes populations subsahariennes (Chen, Olckers *et al.* 2000) paraissent génétiquement moins distants des populations ouest-africaines, que les Mbuti. Ceci, de prime abord, peut sembler ambigu quand on sait que le terme « pygmée » est généralement utilisé pour indiquer les populations Mbuti et Biaka. Cependant, il importe de noter que

contrairement aux Mbuti qui sont de la famille de langues nilo-sahariennes, les Biaka sont classés dans la famille Nigéro-congolaise de laquelle provient la quasi-totalité des chromosomes ouest-africains. La proximité anthropologique des Biaka et Mbuti n'implique donc pas une identité génétique de ces deux populations subsahariennes, comme l'éloignement géographique ne certifie pas l'isolement génétique de ces populations d'autres populations plus à l'ouest ou plus au sud du continent.

L'une des particularités du bassin de gènes des populations ouest-africaines au niveau du locus *dys44* constitue la diversité du microsatellite. En effet, à l'exception du polyT₂₃ (associé à B005) kazakh, cette région présente tous les variants de la répétition de nucléotides T retrouvés dans le monde (Figure 17 page 66). De surcroît, la présence du polyT₁₈ détecté pour la toute première fois dans cette région au Burkina-Faso au sein des populations humaines (Tableau VII page xv en annexes), ainsi que la spécificité de l'haplotype B040 associé, abondent dans le mosaïcisme de la contribution des lignées qui ont contribué à la diversité actuelle du pool génétique des populations subsahariennes en particulier, et des populations du monde en général (Zietkiewicz, Yotova *et al.* 2003). Au sud du Sahara, la variance observée pour le poly T en Afrique de l'Ouest (exception faite des Goun voire Ditammari et Kabiye) suggère une histoire évolutive relativement ancienne des populations humaines de cette région. Cependant, dans un contexte géographique, une grande diversité génétique n'implique pas nécessairement que les populations concernées sont plus anciennes. Autrement, les Afro-américains dont les paramètres de diversité se rapprochent plus des Africains de l'Ouest, et

dont le pool génétique est constitué d'haplotypes africains, eurasiens et spécifiquement afro-américains (B044, B045 etc.), seraient considérés – au détriment de leur contexte historique – comme étant le pool originel des humains actuels. Heureusement, dans le cas de la population noire américaine, aucune évidence ostéologique n'est indispensable pour expliquer (ou conforter) les processus démogénétiques qui ont sous-tendu l'introduction de diversité dans leur bassin de gènes.

L'Afrique de l'Ouest constitue une région dont les populations bénéficient, sans doute depuis plusieurs millénaires, des flux de gènes afro-asiatiques via le Maghreb, des flux de gènes nilo-sahariens (principalement des populations de langues songhaï) et des flux de gènes nigéro-congolais. Cette singularité géo-linguistique non seulement concourt probablement à la riche diversité caractérisée au niveau du marqueur *dys44* pour les populations ouest-africaines mais aussi, confirme l'importance de cette région dans la description à l'échelle moléculaire des variations communes au niveau du génome humain, comme c'est le cas dans le projet HapMap avec l'inclusion judicieuse de la population ouest-africaine Yoruba. Toutefois, quel que soit le groupe ethnique étudié, son pool de gènes ne saurait couvrir toute la variabilité génétique qui existe en Afrique de l'Ouest, encore moins en Afrique subsaharienne. Il aurait été peu probable que la persistance de la lactase soit dévoilée au sein des chasseurs-cueilleurs Hadza de la Tanzanie (Tishkoff, Reed *et al.* 2007), si l'étude de ce phénotype n'avait seulement été réduite qu'à la cohorte Yoruba du Consortium HapMap. Il est souhaitable qu'on tienne compte, à l'exemple du marqueur *dys44*, de la distribution différentielle des variants génétiques au niveau des populations

humaines, pour la cartographie des polymorphismes ayant un impact fonctionnel au sein des groupes d'individus qui les portent.

CONCLUSION

La caractérisation de la diversité génétique du marqueur *dys44* au sein des populations ouest-africaines est la première du genre à impliquer à cette échelle les populations de la sous-famille de langues *Volta-Congo*. Ces populations présentent au niveau de notre séquence d'intérêt, la plus grande variabilité jamais observée pour le locus *dys44* au sein des populations humaines. Bien que ces populations se soient révélées peu distantes génétiquement des autres populations subsahariennes, la spécificité de certains haplotypes dans cette sous-région, ainsi que la variabilité des microsatellites associés, sont en faveur d'une histoire évolutive relativement ancienne du marqueur *dys44* au sein de ces populations de l'ouest africain. Cette sous-région ouest-africaine présente la particularité d'être au confluent des populations des familles Afro-asiatique, Nilo-saharienne et Nigéro-congolaise, ce qui explique aussi la faiblesse de la différenciation génétique entre ces populations africaines au sud du Sahara. Les Afro-américains ressortent plus proches des populations ouest-africaines, ce qui confirme la contribution majeure de l'Afrique de l'Ouest à la traite négrière. Quant à la "migration retour" de populations eurasiennes vers l'Afrique subsaharienne, nos résultats indiquent qu'en dehors du métissage récent des populations nord-africaines aux populations Peulh, les mouvements des porteurs de certains variants du marqueur *dys44* entre l'Eurasie et l'Afrique subsaharienne ont pu s'effectuer via la péninsule ibérique par le détroit de Gibraltar et/ou les îles Canaries. Toutefois les mécanismes génétiques et les processus démographiques à la base de la création et de la diffusion exclusivement ouest-africaine de certains haplotypes restent à être élucidés. L'étude du marqueur DXS1238 (Yotova, Lefebvre *et al.* 2007) en déséquilibre de liaison avec le locus

dys44 contribuerait substantiellement à comprendre ces paramètres démogénétiques, qui ont servi à l'introduction différentielle de ces haplotypes dans le pool génique ouest-africain.

BIBLIOGRAPHIE

- Bendor-Samuel J. T. (1965). "*The Grusi Sub-Group of the Gur Languages*"
Journal of West African Languages 2: 47-55.
- Botstein D., R. L. White, M. Skolnick and R. W. Davis (1980). "*Construction of a genetic linkage map in man using restriction fragment length polymorphisms.*" Am J Hum Genet 32(3): 314-31.
- Bourgeois S. and D. Labuda (2004). "*Dynamic allele-specific oligonucleotide hybridization on solid support.*" Anal Biochem 324(2): 309-11.
- Brakez Z., E. Bosch, H. Izaabel, O. Akhayat, D. Comas, J. Bertranpetit and F. Calafell (2001). "*Human mitochondrial DNA sequence variation in the Moroccan population of the Souss area.*" Ann Hum Biol 28(3): 295-307.
- Brauer G. (1984). "*[Presapiens hypothesis or Afro-European sapiens hypothesis?]*." Z Morphol Anthropol 75(1): 1-25.
- Camps G. (1981). "*L'Origine des Berbères*". CNRS. Paris, Les Cahiers C.R.E.S.M.
- Cann R. L., M. Stoneking and A. C. Wilson (1987). "*Mitochondrial DNA and human evolution.*" Nature 325(6099): 31-6.
- Cavalli-Sforza L. L. and M. W. Feldman (2003). "*The application of molecular genetic approaches to the study of human evolution.*" Nat Genet 33 Suppl: 266-75.
- Chen Y. S., A. Olckers, T. G. Schurr, A. M. Kogelnik, K. Huoponen and D. C. Wallace (2000). "*mtDNA variation in the South African Kung and Khwe and their genetic relationships to other African populations.*" Am J Hum Genet 66(4): 1362-83.
- Clark J. D., Y. Beyene, G. WoldeGabriel, W. K. Hart, P. R. Renne, H. Gilbert, A. Defleur, G. Suwa, S. Katoh, K. R. Ludwig, J. R. Boissarie, B. Asfaw and T. D. White (2003). "*Stratigraphic, chronological and behavioural contexts of Pleistocene Homo sapiens from Middle Awash, Ethiopia.*" Nature 423(6941): 747-52.
- Cohen D. (2007). "*Chamito-Sémitiques (Langues)*", Encyclopaedia Universalis.
- Coia V., G. Destro-Bisol, F. Verginelli, C. Battaglia, I. Boschi, F. Cruciani, G. Spedini, D. Comas and F. Calafell (2005). "*Brief communication: mtDNA variation in North Cameroon: lack of Asian lineages and implications for back migration from Asia to sub-Saharan Africa.*" Am J Phys Anthropol 128(3): 678-81.
- Coppens Y. (1981). "*Le Cerveau des Hommes Fossiles.*" Comptes rendus de l'Académie des Sciences Paris IIA(292): 3-24.
- Cornevin R. (1962). "*Histoire du Togo*". Paris, Editions Berger Levrault.
- Couao-Zotti E. (2007). "*Centenaire de la mort du roi Gbêhanzin : L'histoire d'un royaume et les racines d'un souverain.*" from www.gouv.bj/legislatives2007/affiche.php?numero_article=240.
- Cruciani F., P. Santolamazza, P. Shen, V. Macaulay, P. Moral, A. Olckers, D. Modiano, S. Holmes, G. Destro-Bisol, V. Coia, D. C. Wallace, P. J. Oefner, A. Torroni, L. L. Cavalli-Sforza, R. Scozzari and P. A. Underhill (2002). "*A back migration from Asia to sub-Saharan Africa is supported by*

- high-resolution analysis of human Y-chromosome haplotypes.*" Am J Hum Genet 70(5): 1197-214.
- Dawkins R. (1976). "*The Selfish Gene*", Oxford University Press.
- de Lespinay C. (1994). "*Le Sud Bénin : Kétu et la question des origines*" Cahier du CRA 8: 21-147.
- Decalo S. (1994). "*Historical Dictionary of Benin*", The Scarecrow Press.
- Dupuy C. (1999). "*Les Apports de l'Archéologie et de l'Ethnologie à la Connaissance de l'Histoire Ancienne des Peuls*". Figures Peules. Paris, Karthala: 53-72.
- Excoffier L. G., I. Laval and S. Schneider (2005). "*Arlequin ver. 3.0: An Integrated Software Package for Population Genetics Data Analysis.*" I: 47-50.
- Felsenstein J. (2005). "*Using the quantitative genetic threshold model for inferences between and within species.*" Philosophical Transactions of the Royal Society of London series B 360: 1427-1434.
- Forster P. (2004). "*Ice Ages and the mitochondrial DNA chronology of human dispersals: a review.*" Philos Trans R Soc Lond B Biol Sci 359(1442): 255-64; discussion 264.
- Garrigan D. and M. F. Hammer (2006). "*Reconstructing human origins in the genomic era.*" Nat Rev Genet 7(9): 669-80.
- Gayibor N. L. (1997). "*Histoire des Togolais, Volume I: Des Origines à 1884.*" Presses de l'U.B. I: 135-139.
- Gonzalez A. M., J. M. Larruga, K. K. Abu-Amero, Y. Shi, J. Pestano and V. M. Cabrera (2007). "*Mitochondrial lineage M1 traces an early human backflow to Africa.*" BMC Genomics 8: 223.
- Gordon R. G. (2005). "*Ethnologue: Languages of the World*", Dallas, Tex: Sil International.
- Greenberg J. (1963). "*The Languages of Africa*". Bloomington, Indiana University.
- Hammer M. F. (1995). "*A recent common ancestry for human Y chromosomes.*" Nature 378(6555): 376-8.
- Hammer M. F., T. Karafet, A. Rasanayagam, E. T. Wood, T. K. Altheide, T. Jenkins, R. C. Griffiths, A. R. Templeton and S. L. Zegura (1998). "*Out of Africa and back again: nested cladistic analysis of human Y chromosome variation.*" Mol Biol Evol 15(4): 427-41.
- Harding R. M., S. M. Fullerton, R. C. Griffiths, J. Bond, M. J. Cox, J. A. Schneider, D. S. Moulin and J. B. Clegg (1997). "*Archaic African and Asian lineages in the genetic ancestry of modern humans.*" Am J Hum Genet 60(4): 772-89.
- Harding R. M. and G. McVean (2004). "*A structured ancestral population for the evolution of modern humans.*" Curr Opin Genet Dev 14(6): 667-74.
- Hawks J., K. Hunley, S. H. Lee and M. Wolpoff (2000). "*Population bottlenecks and Pleistocene human evolution.*" Mol Biol Evol 17(1): 2-22.
- Hirszfeld L. and H. Hirszfeld (1919). "*Serological differences between the blood of different races.*" Lancet 2: 675-679.
- Hudson R. R. (1990). "*Gene genealogies and the coalescence process.*" Oxf Sur Evol Biol 7: 1-44.

- Huysecom E. (2007). "*Un Néolithique ancien en Afrique de l'Ouest?*" Pour La Science 358: 44-49.
- Jobling M. A., M. E. Hurles and C. Tyler-Smith (2004). "*Human Evolutionary Genetics: Origins, Peoples and Disease*". New York, Garland Science.
- Johanson D. C. and M. Taieb (1976). "*Plio--Pleistocene hominid discoveries in Hadar, Ethiopia.*" Nature 260(5549): 293-7.
- Kingman J. F. C. (1982). "*On the genealogy of large populations.*" J Appl Prob 19A: 27-43.
- Kivisild T., P. Shen, D. P. Wall, B. Do, R. Sung, K. Davis, G. Passarino, P. A. Underhill, C. Scharfe, A. Torroni, R. Scozzari, D. Modiano, A. Coppa, P. de Knijff, M. Feldman, L. L. Cavalli-Sforza and P. J. Oefner (2006). "*The role of selection in the evolution of human mitochondrial genomes.*" Genetics 172(1): 373-87.
- Labuda D., E. Zietkiewicz and V. Yotova (2000). "*Archaic lineages in the history of modern humans.*" Genetics 156(2): 799-808.
- Leakey M. D. and R. L. Hay (1979). "*Pliocene footprints in the Laetoli Beds at Laetoli, northern Tanzania.*" Nature 278: 317-323.
- Lovell A., C. Moreau, V. Yotova, F. Xiao, S. Bourgeois, D. Gehl, J. Bertranpetit, E. Schurr and D. Labuda (2005). "*Ethiopia: between Sub-Saharan Africa and western Eurasia.*" Ann Hum Genet 69(Pt 3): 275-87.
- Macaulay V., M. Richards, E. Hickey, E. Vega, F. Cruciani, V. Guida, R. Scozzari, B. Bonne-Tamir, B. Sykes and A. Torroni (1999). "*The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs.*" Am J Hum Genet 64(1): 232-49.
- Morrisson D. G., C. R. Mitchell, J. P. Paden and H. M. Stevenson (1989). "*Black Africa: A Comparative Handbook*". New York, Free Press.
- Ohta T. and M. Kimura (1973). "*A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population.*" Genetic Research 22: 201-204.
- Olson J. S. (1996). "*The Peoples of Africa: An Ethnohistorical Dictionary*". Westport, CT 06881, Greenwood Press.
- Pauling L., H. A. Itano, S. J. Singer and I. C. Wells (1949). "*Sickle cell anemia a molecular disease.*" Science 110(2865): 543-8.
- Pellissier P. (1962). "*Les pays du Bas-Ouémé, une région témoin du Dahomey méridional*", Les Cahiers d'Outre Mer.
- Plaza S., F. Calafell, A. Helal, N. Bouzerna, G. Lefranc, J. Bertranpetit and D. Comas (2003). "*Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean.*" Ann Hum Genet 67(Pt 4): 312-28.
- Reed F. A. and S. A. Tishkoff (2006). "*African human diversity, origins and migrations.*" Curr Opin Genet Dev 16(6): 597-605.
- Richards M., C. Rengo, F. Cruciani, F. Gratrix, J. F. Wilson, R. Scozzari, V. Macaulay and A. Torroni (2003). "*Extensive female-mediated gene flow from sub-Saharan Africa into near eastern Arab populations.*" Am J Hum Genet 72(4): 1058-64.

- Rosa A., A. Brehm, T. Kivisild, E. Metspalu and R. Villems (2004). "*MtDNA profile of West Africa Guineans: towards a better understanding of the Senegambia region.*" *Ann Hum Genet* 68(Pt 4): 340-52.
- Roy C. D. (1998). "*The Art of Burkina Faso*".
- Ruhlen M. (1987). "*A Guide to the World's Languages*". Stanford, CA: Stanford University Press.
- Salas A., M. Richards, T. De la Fe, M. V. Lareu, B. Sobrino, P. Sanchez-Diz, V. Macaulay and A. Carracedo (2002). "*The making of the African mtDNA landscape.*" *Am J Hum Genet* 71(5): 1082-111.
- Soriano S. and E. Huysecom (2007). "*Un paléolithique ignoré.*" *Pour La Science* 358: 36-43.
- Stephens M. and P. Donnelly (2003). "*A comparison of bayesian methods for haplotype reconstruction from population genotype data.*" *Am J Hum Genet* 73(5): 1162-9.
- Stephens M., N. J. Smith and P. Donnelly (2001). "*A new statistical method for haplotype reconstruction from population data.*" *Am J Hum Genet* 68(4): 978-89.
- Stringer C. (2002). "*Modern human origins: progress and prospects.*" *Philos Trans R Soc Lond B Biol Sci* 357(1420): 563-79.
- Tang H., D. O. Siegmund, P. Shen, P. J. Oefner and M. W. Feldman (2002). "*Frequentist estimation of coalescence times from nucleotide sequence data using a tree-based partition.*" *Genetics* 161(1): 447-59.
- Tauxier L. (1937). "*Mœurs et Histoires des Peuls*". Payot. Paris: 422.
- Tcham B. (1990). "*Les Peuples du Togo.*" Presses de l'U.B. Vol I.
- Templeton A. R. (1998). "*Human race: a genetic and evolutionary perspective.*" *Am Anthropol* 100(3): 632-650.
- Tishkoff S. A., F. A. Reed, A. Ranciaro, B. F. Voight, C. C. Babbitt, J. S. Silverman, K. Powell, H. M. Mortensen, J. B. Hirbo, M. Osman, M. Ibrahim, S. A. Omar, G. Lema, T. B. Nyambo, J. Gori, S. Bumpstead, J. K. Pritchard, G. A. Wray and P. Deloukas (2007). "*Convergent adaptation of human lactase persistence in Africa and Europe.*" *Nat Genet* 39(1): 31-40.
- Vigilant L., M. Stoneking, H. Harpending, K. Hawkes and A. C. Wilson (1991). "*African populations and the evolution of human mitochondrial DNA.*" *Science* 253(5027): 1503-7.
- Watson E., P. Forster, M. Richards and H. J. Bandelt (1997). "*Mitochondrial footprints of human expansions in Africa.*" *Am J Hum Genet* 61(3): 691-704.
- Westermann D. H. (1930). "*A Study of the Ewe Languages*". London, Oxford University Press.
- Wolpoff M. and R. Caspari (1996). "*The modernity mess.*" *Hum Evo* 30(2): 167-171.
- Wright S. (1978). "*Variability Within and Among Natural Populations*". Chicago, University of Chicago Press.
- Xiao F. X., V. Yotova, E. Zietkiewicz, A. Lovell, D. Gehl, S. Bourgeois, C. Moreau, C. Spanaki, A. Plaitakis, J. P. Moisan and D. Labuda (2004).

- "*Human X-chromosomal lineages in Europe reveal Middle Eastern and Asiatic contacts.*" *Eur J Hum Genet* 12(4): 301-11.
- Yotova V., J. F. Lefebvre, O. Kohany, J. Jurka, R. Michalski, D. Modiano, G. Utermann, S. M. Williams and D. Labuda (2007). "*Tracing genetic history of modern humans using X-chromosome lineages.*" *Hum Genet* 122(5): 431-43.
- Zietkiewicz E., V. Yotova, D. Gehl, T. Wambach, I. Arrieta, M. Batzer, D. E. Cole, P. Hechtman, F. Kaplan, D. Modiano, J. P. Moisan, R. Michalski and D. Labuda (2003). "*Haplotypes in the dystrophin DNA segment point to a mosaic origin of modern human diversity.*" *Am J Hum Genet* 73(5): 994-1015.
- Zietkiewicz E., V. Yotova, M. Jarnik, M. Korab-Laskowska, K. K. Kidd, D. Modiano, R. Scozzari, M. Stoneking, S. Tishkoff, M. Batzer and D. Labuda (1997). "*Nuclear DNA diversity in worldwide distributed human populations.*" *Gene* 205(1-2): 161-71.
- Zietkiewicz E., V. Yotova, M. Jarnik, M. Korab-Laskowska, K. K. Kidd, D. Modiano, R. Scozzari, M. Stoneking, S. Tishkoff, M. Batzer and D. Labuda (1998). "*Genetic structure of the ancestral population of modern humans.*" *J Mol Evol* 47(2): 146-55.

ANNEXES

Annexe A

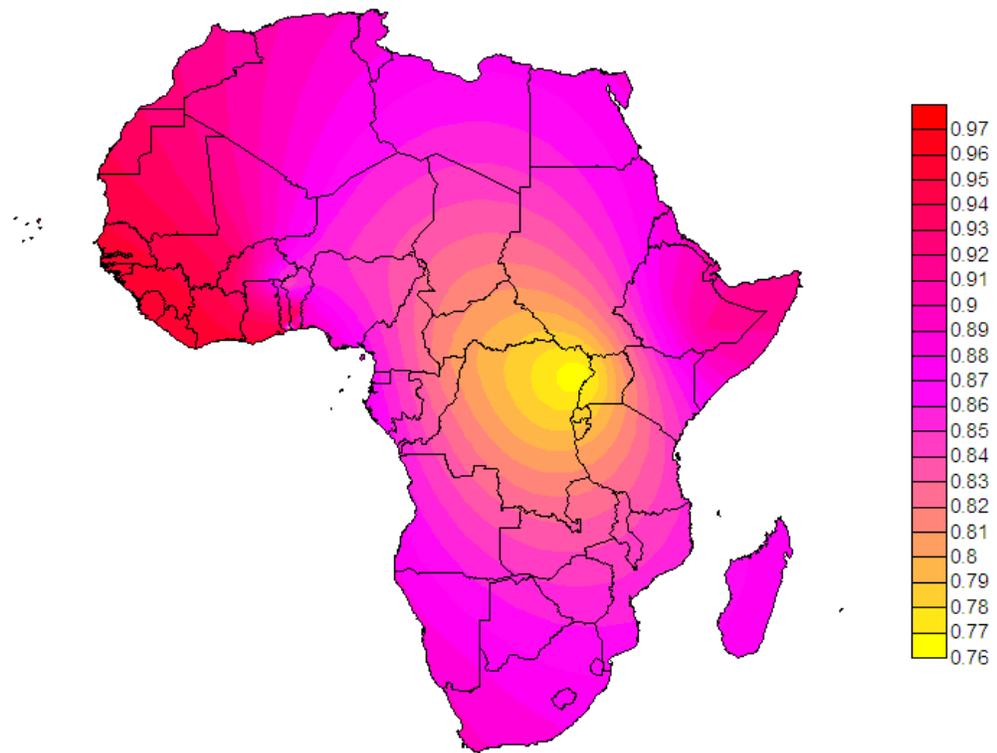


Figure 21: Diversité haplotypique de *dys44* en Afrique

Annexe B

Popul	n	S	k	G (\pm S.E.)	Hn $\times 10^3$ (\pm S.E.)	T ₁₅		T ₂₂	
						n	S ²	n	S ²
Adja	73	24	22	0,93 (\pm 0,01)	0,95 (\pm 0,22)	46	0,663	27	1,712
Ayizo	39	25	17	0,91 (\pm 0,03)	0,83 (\pm 0,19)	18	0,967	21	1,648
Ditam	30	26	10	0,83 (\pm 0,04)	0,91 (\pm 0,21)	19	0,690	11	0,855
Fon	21	25	13	0,95 (\pm 0,03)	0,84 (\pm 0,19)	7	1,000	14	1,456
Goun	17	25	10	0,93 (\pm 0,04)	0,97 (\pm 0,22)	9	0,250	8	0,696
Kot	9	19	4	0,75 (\pm 0,11)	0,87 (\pm 0,22)	7	1,238	2	2,000
Mahi	35	28	20	0,96 (\pm 0,02)	0,99 (\pm 0,22)	14	0,725	20	1,145
Nago	15	22	7	0,86 (\pm 0,06)	1,03 (\pm 0,24)	9	1,000	6	0,967
Weme	35	25	18	0,94 (\pm 0,02)	0,9 (\pm 0,21)	20	0,618	15	2,429
Kab	30	24	11	0,88 (\pm 0,03)	0,87 (\pm 0,2)	14	0,863	16	0,667
Ake	16	24	9	0,92 (\pm 0,04)	0,9 (\pm 0,21)	7	0,143	9	1,028
Ewe	30	28	23	0,97 (\pm 0,02)	0,99 (\pm 0,22)	20	0,779	10	1,878
Naw	29	25	14	0,93 (\pm 0,02)	0,96 (\pm 0,22)	15	1,114	14	1,516
Peulh	52	27	13	0,90 (\pm 0,02)	1 (\pm 0,21)	41	1,480	11	0,855
Rim	96	26	25	0,93 (\pm 0,01)	0,96 (\pm 0,22)	37	0,859	58	2,438
Mos	91	28	26	0,90 (\pm 0,02)	0,99 (\pm 0,22)	47	0,850	44	1,021
Tama	41	28	17	0,89 (\pm 0,03)	0,94 (\pm 0,21)	19	1,228	22	0,814
Rim 2	31	29	16	0,94 (\pm 0,02)	1,09 (\pm 0,23)	20	1,200	11	0,891
Mos 2	30	26	19	0,95 (\pm 0,02)	0,94 (\pm 0,21)	14	0,901	16	1,629
Ga	45	26	26	0,95 (\pm 0,02)	1,02 (\pm 0,23)	22	1,100	15	1,571
Biaka	85	23	16	0,81 (\pm 0,03)	0,98 (\pm 0,24)	69	0,390	16	1,563
Mbu	58	20	11	0,76 (\pm 0,05)	0,79 (\pm 0,2)	21	0,662	37	0,027
Amh	30	24	12	0,89 (\pm 0,03)	0,85 (\pm 0,2)	22	0,476	8	0,786
Oro	19	19	11	0,91 (\pm 0,05)	0,49 (\pm 0,15)	17	0,375	2	0,000
Khoi	18	23	9	0,89 (\pm 0,05)	0,88 (\pm 0,21)	4	0,333	14	1,846
AA	65	32	26	0,93 (\pm 0,02)	1,06 (\pm 0,23)	44	0,877	21	1,900
Sah	23	23	7	0,87 (\pm 0,04)	0,78 (\pm 0,18)	11	0,473	8	0,000
J Ethi	16	23	9	0,85 (\pm 0,08)	0,8 (\pm 0,19)	15	0,552	1	Na
J Yém	31	19	9	0,77 (\pm 0,07)	0,68 (\pm 0,17)	29	0,429	2	0,000
J Mar	29	22	9	0,68 (\pm 0,09)	0,48 (\pm 0,12)	27	0,071	2	4,500
J Ash	110	23	12	0,73 (\pm 0,03)	0,59 (\pm 0,14)	100	0,120	10	0,400
Palest	23	22	9	0,81 (\pm 0,06)	0,87 (\pm 0,2)	23	0,300	0	Na
Basq	20	21	9	0,89 (\pm 0,04)	0,91 (\pm 0,22)	17	0,485	3	0,333
Ital	26	20	7	0,76 (\pm 0,06)	0,56 (\pm 0,14)	24	0,087	2	0,000
Chin	78	20	8	0,68 (\pm 0,04)	0,83 (\pm 0,22)	72	0,000	6	0,000
Jap	65	19	7	0,57 (\pm 0,05)	0,7 (\pm 0,19)	63	0,000	2	0,000
PNGc	24	18	8	0,77 (\pm 0,08)	0,85 (\pm 0,22)	18	0,000	6	0,000
PNGh	41	18	6	0,55 (\pm 0,09)	0,55 (\pm 0,15)	39	0,000	2	0,000

Tableau VI : Paramètres de diversité et variance S² du microsatellite Tn en fonction des modes T₁₅ et T₂₂ au sein des populations africaines, eurasiennes et océaniques

Na = non applicable ; Jap = Japonais ; Chin = Chinois ; Ital = Italien ; Basq = Basque ; Palest = palestinien ; J = Juif ; Ash = Ashkénaze ; Mar = Marocain ; Yém = Yéménite ; Ethi = Ethiopien ; Sah = Saharawi ; AA = Afro-américain ; Khoi = Khoisan ; Oro = Oromo ; Amh = Amhara ; Mbu = Mbuti ; Mos = Mossi ; Rim = Rimaibe ; Tam = Tamasheq ; Naw = Nawdam ; Ake = Akebou ; Kab = Kabiye ; Kot = Kotafon ; Ditam = Ditammari ; Popul = Population

Annexe C

Populations	Bénin								Burkina Faso					Togo				RCA		Eth		RSA	Gh	USA	Fréquence %			
	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Rimaibe2	Mossi	Mossi2	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoi San		Gra	AF-Am	
Hap	Tn																											
B001	14								1	1																	0.19	
	15	5	3	10		3		1	4	4	7	3	8	4	2	1	3	1	3	18	5	7	3	1	5	3	10.00	
	16						1			1	1						1										0.38	
	17	1																									0.10	
	19*									1																	0.10	
	21																					1					0.10	
B002	14													1						7	2						0.96	
	15	2	1	1	1	1	2	1	1	2	2	4	3	12	1	1	3	1	1	12	2	1		1	1	6	6.06	
	16	4	1	2		1		1	1	2	1	1		2	2	3		1	2	11		1		1	1	4	4.04	
	17	3	1	3	1		1	2	3	1	5	6	3	7	2	3	2		2	4	1	3	1		5	4	6.06	
	18*										1							1									0.19	
	19	1																									0.10	
	20*														1												0.10	
	22																					1		1			0.19	
	n.d																								1		0.10	
B003	14	1	2	1	2					1		1						1	2						2	3	1.54	
	15						1			2	4					2				1	3	5	5			4	2.60	
B004	15											1														1	0.19	
B005	15	1																					1				0.19	
	21									1		1	3														0.48	
	22	7	10	5	2	2		3	2	4	5	6		13	2	4	8	2	1	5	4				3	3	8.75	
	n.d																								1			0.10
B006	15											1																0.10
	16									3					1							1				3		0.77

(début Tableau VII)

Populations	Bénin									Burkina Faso					Togo				RCA		Eth		RSA	Gh	USA	Fréquence %	
	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Wene	Peulh	Rimaibe	Rimaibe2	Mossi	Mossi2	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoi San	Ga		AF-Am
B007	19*								1		1	1											5				0.77
	20	6	1	2	1	3	1	5	3	1		1	1	2	1					2	27	2			1	2	6.06
	21*		1																								0.10
B008	13*													2													0.19
	14	4				4	1		1	1			2		2	1		1				3	2			2	2.31
	15	3	2																								0.48
	21												1														0.10
B009	15	2							1	3	1	2	1			3			1	7	4					2	2.60
B010	15*														1								1				0.19
	20*												1		2												0.29
	21*			1		1					5	2	2		1	1			2						1	1	1.63
	22*		1	1		1	1		3	1	3		2	2	6	3	3	2				2			1	4	3.46
	24*				1				1		1		1													1	0.48
	n.d																							1			0.10
B011	15																				1	1					0.19
	22													2												1	0.29
B012	17								1	10	2	3		2		1										1	1.92
	20*			1	1		1				3	1	1		1	1		1							1	1	1.25
B013	21													1													0.10
B016	15	9	1		1		2	2	1		2		5	1		1		1	2			1			1	3	3.17
B017	15*										1	1															0.19
	19*	2	4				1		2		6	1	2				1								1	2	2.12
B018	19										2									1						2	0.48
B019	15	1																									0.10
	17*		2					1						1	1			1									0.58

(suite Tableau VII)

Populations	Bénin								Burkina Faso					Togo				RCA	Eth	RSA	Gh	USA	Fréquence %					
	Adja	Ayizo	Ditammani	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Rimaibe2	Mossi	Mossi2	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara		Oromo	Khoi San	Ga	AF-Am	
B021	20																			3								0.29
B022	15*																								1			0.10
	19*	1																										0.10
	21*											1																0.10
	22*	2		1	1		2			5		4	1	2		1		1								1	2.02	
B023	15											1																0.10
B023	n.d																								3			0.29
B024	15												1								1				1	1	0.38	
B027	14	1						1			1															2	0.48	
	15*			1			1										1											0.29
B029	17*										1																	0.10
B030	15											1																0.10
B032	15		1				1	1												6	1		2		1		1.25	
B036	17																					1						0.10
B037	16																			4								0.38
B038	20												1															0.10
B039	22										2		1							1								0.38
B040	18*										9																	0.87
	19		1		1			1			4		1					1							1			0.96
B044	15																									1	0.10	
B045	14																									1	0.10	
B046	19																			1								0.10
	20																					4						0.38
B047	22												1															0.10

(suite Tableau VII)

Populations	Bénin								Burkina Faso					Togo				RCA	Eth	RSA	Gh	USA	Fréquence %						
	Adja	Ayizo	Ditammani	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimabe	Rimabe2	Mossi	Mossi2	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara		Oromo	Khoi San	Ga	AF-Am		
B048 16*																	1											0.10	
B050 17*											1			1															0.19
B056 20																					2								0.19
b057 14																										1		0.10	
B058 15	1	2				1		3	3	3		1	1				2					1				1		1.83	
17*		1										2																0.29	
21*												1																0.10	
b059 20																										1		0.10	
b060 21																										1		0.10	
b061 19																				1								0.10	
20																				4								0.38	
b062 19																				1								0.10	
b063 14																				1		1						0.10	
16																				1								0.10	
B067 21*				3													1	1										0.48	
22																						1			1			0.19	
B072 20												1																0.10	
B074 18*										1																		0.10	
19										3		1				1		2				1		1				0.87	
B088 20*															1			1										0.19	
21	1			3	1				3			3	1	1	1	1	1											1.54	
B106 22																						1						0.10	
B107 20																							1					0.10	
21																						1						0.10	
B108 22																							1	2					0.29

(suite Tableau VII)

Populations	Bénin								Burkina Faso					Togo				RCA		Eth		RSA	Gh	USA	Fréquence %		
	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimabe	Rimabe2	Mossi	Mossi2	Tamasheq	Kabye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoi San		Ga	Al-Am
b110	14																		1								0.10
	20												1														0.10
b115	20												1														0.10
b116	21												1														0.10
B117	21*	1					3	1				2	1			1	1							1			1.06
b118	15											1															0.10
B119	19																		1								0.10
b120	19																				1						0.10
b121	21*				1								1														0.19
	22*												1														0.10
B124	21																								1		0.10
B125	17																								1		0.10
B136	19	4									1		1											2			0.77
B143	15				1																						0.10
B145	15	1																									0.10
	16	1																						1			0.19
	22																							1			0.10
B146	20*			1				1																			0.19
B151	14*												1														0.10
	15																								1		0.10
B152	17																								1		0.10
B153	20	1																									0.10
	22	1																							1		0.19
B154	22																								1		0.10
B155	17																								1		0.10
B156	22																								1		0.10
B158	20																								2		0.19

(suite Tableau VII)

Populations	Bénin								Burkina Faso					Togo				RCA	Eth	RSA	Gh	USA	Fréquence %					
	Adja	Ayizo	Ditammani	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Rimaibe2	Mossi	Mossi2	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara		Oromo	Khoi San	Ga	Al-Am	
B159	15																1											0.10
	16																							1				0.10
B160	15*	3	1	2	1		1	1		1		1		1		2	1	2										1.63
	16*						1																					0.10
B163	20																						1					0.10
B164	16*									1																		0.10
	20*	1								2																		0.29
B166	17*	1																										0.10
b207	14*											1																0.10
B219	17*										1																	0.10
B220	22*						1																					0.10
b221	16*	1																										0.10
B222	19*						1																					0.10
B223	15*																1											0.10
B224	15*																1											0.10
b225	22*									1																		0.10
B226	22*	1																										0.10
B227	22*						1				1	1					1											0.38
B228	19*									1																		0.10
b229	19*									2																		0.19
b230	17*											1																0.10
b231	19*									1																		0.10
B232	16*	1			2			1																				0.38
b233	15*											1																0.10

(suite Tableau VII)

Populations	Bénin								Burkina Faso					Togo				RCA		Eth		RSA	Gh	USA	Fréquence %								
	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Rimaibe2	Mossi	Mossi2	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoi San		Ga	AF-Am						
B234 19*	1																																0.10
b235 21*												1																				0.10	
b238 15*	1																															0.10	
B239 21*																	1															0.10	
B240 15*																	1															0.10	
B241 20*															1																	0.10	
b242 14*									1																							0.10	
B243 20*							2																									0.19	
b244 20*														1																		0.10	
b245 17*														1																		0.10	
B246 15*																	1															0.10	
Total Chr X	73	39	30	21	17	9	35	15	35	52	96	31	91	30	41	30	16	30	29	85	58	30	19	18	45	65					100.00		

Tableau VII : Distribution des haplotypes (Hap) de *dys44* et leurs poly T associés en Afrique subsaharienne et chez les Africains-Américains Africains-Américains = Af-Am

RCA=République Centrafricaine; RSA=République d'Afrique du Sud; Eth=Ethiopie; Gh=Ghana; Hap=Haplotypes; Fréq. %=Fréquence des haplotypes en % ; Tn=Répétition en tandem de nucléotide T; * correspond à un haplotype étendu Bxxx_Tn spécifique à l'Afrique

Annexe D

Haplotypes	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoisan	Ga	Mossi 2	Rimaibe 2	Saharawi	J Eth	J Yemenite	J Mor	J Ashk	Palestinien	Basque	Italien	Chinois	Japanese	PNGe	PNGh	AF-Am	Fréquence %
B001	6	3	10		3		2		4	7	9	8	2	1	3	2	3	18	5	8	3	1	5	4	3	3	6	14	16	44	9	3	11	37	39	2	3	3	19.66
B002	10	3	6	2	2	3	4	5	5	8	12	21	9	5	2	5	5	31	7	4		3	8	5	6	2	3	5	1	10	1	1	2	24	18	11	27	14	19.00
B003	1	2	1	2			1			3	4			2		1	2	1	3	5	5		2		1	6	1	2	5	35	2	5	7	5	1	1		7	7.40
B004																									1							1		1		1		0.33	
B005	8	10	5	2	2		3	2	4	6	6	16	4	8	2	1	5	4			1		4	2	1	5	1			4		2		6	2	4	3	8.06	
B006										3			1							1					1	2	1		2	7	5	4	2	2			3	2.23	
B007	6	2	2	1	3	1	5	3	2		1	3	1					2	27	2		5	1	1	2												2	4.72	
B008	7	2				4	1		1	1	3	1	1	1		1				3	2						4	1	2	2	2						2	2.82	
B009	2								1	3	1	1		3			1	7	4						2												2	1.77	
B010		1	2	1	1	1	1		4	1	9	5	10	4	3	2	2			2	1	3	3	3	2												6	4.19	
B011													2							1	1				3		2	1	2			2					1	0.98	
B012			1	1			1		1	10	5	1	1	2			1						1	2	4	1											2	2.23	
B013									1																					2		1						0.26	
B014																																	2					0.13	
B016	9	1		1			2	2	1		2	5		1		1	2			1		1	1				1	1									3	2.29	
B017	2	4					1		2	1	7	2				1							1		1													2	1.57
B018																		1								2												2	0.33
B019	1	2						1								1								1															0.46
B020																																				3			0.20
B021																			3																				0.20
B022	3		1	1			2				5	5	2		1		1					1	1															1	1.57
B023												1																						1				0.33	
B024																			1				1	1													1	0.26	
B026																																			2				0.13
B027	1		1				1		1							1									1													2	0.52
B028																																							0.13
B029											1																									2		0.39	
B030												1																								1	4		0.20
B031																																1					2		0.07

(début Tableau VIII)

Haplotypes	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoisan	Ga	Mossi 2	Rimaibe 2	Saharawi	J Eth	J Yemenite	J Mor	J Ashk	Palestinien	Basque	Italien	Chinois	Japanese	PNGc	PNGh	AF-Am	Fréquence %
B032		1					1		1									6	1		2		1															0.85	
B033																													1										0.07
B035																																							0.07
B036																					1													2				0.20	
B037																		4																				0.26	
B038																									1													0.07	
B039											2							1							1													0.26	
B040		1		1					1		13						1						1		1													1.25	
B042																																			2	2		0.26	
B044																																				1		0.07	
B045																																				1		0.07	
B046																	1		4																			0.33	
B047																								1														0.07	
B048															1																							0.07	
B050													1												1													0.13	
b054																																		1				0.07	
B056																			2																			0.13	
b057																																				1		0.07	
B058	1	3					1		3	3	3	4	1			2					1								1								1	1.57	
b059																																					1	0.07	
b060																																				1		0.07	
b061																		5																				0.33	
b062																		1																				0.07	
b063																		1		1																		0.13	
B067				3											1	1					1				1													0.52	
b068																																		1				0.07	
B072																								1														0.07	
B074											4	1			1		2						1	1														0.66	
B082																																1					0.07		
b084																													1									0.07	

(suite Tableau VIII)

Haplotypes	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoisan	Ga	Mossi 2	Rimaibe 2	Saharawi	J Eth	J Yemenite	J Mor	J Ashk	Palestinien	Basque	Italien	Chinois	Japanese	PNGc	PNGh	AF-Am	Fréquence %		
B088	1			3	1					3	3	1	2		1	1	1							1																1.18	
b089																													1											0.07	
B106																				1																				0.07	
B107																				1		1																		0.13	
B108																						1	2																	0.20	
b110																		1						1																0.13	
b114																													1											0.07	
b115																								1																	0.07
b116																							1																		0.07
B117	1						3	1				1			1	1							1		2															0.72	
b118																								1																	0.07
B119																		1																						0.07	
b120																			1																					0.07	
b121					1							2																												0.20	
B124																																					1		0.07		
B125																																					1		0.07		
B126																											1													0.07	
B127																											1													0.07	
b128																											1													0.07	
B129																											1													0.07	
B130																																	1							0.07	
b131																																								0.07	
b132																																								0.07	
B135																												1												0.07	
B138																												1												0.07	
B136	4										1	1																	1											0.59	
B137																														1										0.07	
B143					1																																			0.07	
B145	2																																							0.26	
B146				1					1																															0.13	

(suite Tableau VIII)

Haplotypes	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoisan	Ga	Mossi 2	Rimaibe 2	Saharawi	J Eth	J Yemenite	J Mor	J Ashk	Palestinien	Basque	Italien	Chinois	Japanese	PNGc	PNGh	AF-Am	Fréquence %		
B151												1											1																0.13		
B152																							1																	0.07	
B153	2																						1																	0.20	
B154																							1																	0.07	
B155																							1																	0.07	
B156																							1																	0.07	
B158																							2																	0.13	
B159																1							2																	0.13	
B160	3	1	2	1	1	1	1	1	1	1	1	1	1	2	1	2							1																1.18		
B163																						1																		0.07	
B164	1										3																													0.26	
B166	1																																							0.07	
b207												1																												0.07	
B219											1																													0.07	
B220							1																																	0.07	
b221	1																																							0.07	
B222							1																																	0.07	
B223																1																								0.07	
B224																1																								0.07	
b225											1																													0.07	
B226	1																																							0.07	
B227							1				1	1			1																										0.26
B228											1																														0.07
b229											2																														0.13
b230												1																													0.07
b231											1																														0.07
B232	1				2				1																															0.26	
b233												1																													0.07
B234	1																																							0.07	
b235												1																													0.07

(suite Tableau VIII)

Haplotypes	Adja	Ayizo	Ditamari	Fon	Goun	Kotafon	Mahi	Nago	Weme	Peulh	Rimaibe	Mossi	Tamasheq	Kabiye	Akebou	Ewe	Nawdm	Biaka	Mbuti	Amhara	Oromo	Khoisan	Ga	Mossi 2	Rimaibe 2	Saharawi	J Eth	J Yemenite	J Mor	J Ashk	Palestinien	Basque	Italien	Chinois	Japanese	PNGc	PNGh	Af-Am	Fréquence %		
b238	1																																								0.07
B239																1																									0.07
B240																1																									0.07
B241														1																											0.07
b242										1																															0.07
B243							2																																		0.13
b244													1																												0.07
b245													1																												0.07
B246																1																									0.07
Total	73	39	30	21	17	9	35	15	35	52	96	91	41	30	16	30	29	85	58	30	19	18	45	30	31	23	16	31	29	110	23	20	26	78	65	24	41	65	100.00		

Tableau VIII : Distribution des haplotypes de *dys44* en Afrique subsaharienne et dans quelques populations non-africaines