

2m11.3035.2

UNIVERSITE DE MONTREAL

**STRUCTURE GENETIQUE DES POPULATIONS :
MODELE AMERINDIEN ET LES HAPLOTYPES DU CHROMOSOME X**

Par
Stéphane Bourgeois

Programme de biologie moléculaire
Faculté des études supérieures

Mémoire présenté à la Faculté des études supérieures
En vue de l'obtention du grade de
Maître ès sciences (M. Sc.)
En biologie moléculaire

Décembre, 2002

©, Stéphane Bourgeois, 2002à

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

UNIVERSITE DE MONTREAL

Faculté des études supérieures

Ce mémoire intitulé :

**STRUCTURE GENETIQUE DES POPULATIONS :
MODELE AMERINDIEN ET LES HAPLOTYPES DU CHROMOSOME X**

Présenté par :

Stéphane Bourgeois

A été évalué par un jury composé des personnes suivantes :

Dr Bernard Brais

président-rapporteur

Dr Damian Labuda

directeur de recherche

Dr Hervé Philippe

membre du jury



RESUMÉ

Les trois groupes linguistiques amérindiens sont censés refléter trois vagues indépendantes de colonisation du Nouveau Monde à partir du détroit de Béring. Les études génétiques de l'ADN mitochondrial et du chromosome Y, transmis maternellement et paternellement respectivement, ne permettent pas une distinction claire entre les individus de ces trois groupes linguistiques. Nous présentons ici les données d'un marqueur génétique du chromosome X, étudié pour des populations représentant les groupes linguistiques *NaDene* et *Amerind*. Les haplotypes du système *Dys-44* consistent en 35 sites polymorphiques dans une portion intronique de 8kb entourant l'exon 44 du gène de la dystrophine, en Xp21. Des six principaux haplotypes non-africains, identifiés dans un échantillon de 1315 chromosomes X distribués à travers le monde, quatre sont d'un intérêt particulier. Les haplotypes B002 et B005, avec une fréquence globale de 21,3% et 7%, respectivement, sont virtuellement absents en Amérique du Nord (*NaDene* et *Amerind*), en Amérique Centrale (*Amerind*) ainsi qu'en Amérique du Sud (*Amerind*). Cette découverte procure un argument en faveur d'un seul effet fondateur réunissant les populations appartenant aux groupes linguistiques *Amerind* et le *NaDene*. D'autre part, la fréquence de l'haplotype B004 est dramatiquement augmentée chez les Mayas (31,3%), et encore plus chez les Karitianas (63,9%), en comparaison des Ojibwas (3,7%) et des Chipewyans (4%); cela suggère des effets fondateurs relatifs à la colonisation de l'Amérique Centrale et du Sud, qui ne corrèlent toutefois pas avec les principales subdivisions linguistiques. Les évidences pointent actuellement vers un unique événement majeur de colonisation et une subséquente diversification des populations du Nouveau Monde.

Mots clefs : Amérindiens, chromosome X, Dystrophine, haplotypes, génétique des populations

SUMMARY

Genetic studies of maternally inherited mitochondrial DNA and paternally inherited Y-chromosome do not provide clear distinction between the three linguistic groups of the American populations believed to represent three independent waves of colonization of the New World through Beringia. Here, we present the data on a nuclear X-linked marker studied in NaDene and Paleo-Indian populations. *Dys-44* haplotypes consist of 35 segregating sites found within 8 kb of the intronic DNA flanking exon 44 of the dystrophin gene on Xp21. Out of the six major non-African haplotype identified in a worldwide sample (n=1315) of X-chromosomes four are of particular interest. The haplotype B002 and B005 with overall worldwide frequencies of 21.3% and 7%, respectively, appear to be virtually absent in Northern America (NaDene and Paleo-Indians) as well as Central (Paleo-Indians) and Southern America (Paleo-Indians). This finding provides an argument for a single founder effect including both Paleo-Indians and NaDene speakers. On the other hand, the frequency of haplotype B004 is dramatically increased in Maya (31.3%) and even more in Karitiana (63.9%) as compared to 3.7% in Ojibwa and 4% in NaDene. This local founder effect, which however does not correlate with major linguistic subdivisions, could reflect population movements from North to South, or South to North, during the colonization of the Americas. Present evidence point to a single major colonization event and subsequent diversification of the population of the New World.

Keywords: Amerindians, population genetics, X chromosome, Dystrophin, haplotypes.

TABLE DES MATIERES

	Page
Identification du jury	i
Résumé	ii
<i>Summary</i>	iii
Table des matières	iv
Liste des tableaux	vii
Liste des figures	ix
Liste des sigles et abréviations	xi
Dédicace	xii
Remerciements	xiii
Chapitre I : INTRODUCTION	1
A) Données archéologiques, anthropologiques, paléoclimatiques et linguistiques	5
1. Paléoclimatologie	5
2. Archéologie, anthropologie et linguistique	10
a. Données des sites archéologiques, anthropologie	10
b. Voyages maritimes	17
3. Linguistique	22
B) Etude de l'ADN mitochondrial	27
1. Définitions et nomenclature	27
2. Voies de migration	28
3. Lignées fondatrices	31
4. Datation des migrations	33
5. Origine des Amérindiens	36

C) Etude du chromosome Y	37
1. Nomenclature	37
2. Lignées fondatrices, origine des Amérindiens et vagues de migration	40
3. Datation des migrations	41
D) Etude des autosomes, co-évolution de micro-organismes	42
1. Système HLA	42
2. Groupes sanguins	43
3. Co-évolution de micro-organismes	45
E) Projet de recherche	46
Chapitre II : MATERIEL ET METHODES	48
A) Populations étudiées	49
B) Extraction de l'ADN	53
C) Détermination de la concentration en ADN	54
1) Spectrophotomètre	54
2) Fluorimètre	55
D) Système <i>Dys 44</i> et amplification par PCR	57
1) Description du système	57
2) Amplification par PCR	57
E) Transfert sur membranes	60
F) Hybridation ASO	61
1) Marquage radioactif des sondes	63
2) Protocole	63
G) Base de données, étude des haplotypes	67

Chapitre III : RESULTATS	72
A) Structure des populations amérindiennes	73
B) Populations amérindiennes et Eurasie	85
C) Influence récente des Européens	104
Chapitre IV : DISCUSSION	114
BIBLIOGRAPHIE	126

LISTE DES TABLEAUX

	Page
Tableau I : Classification des ères géologiques	6
Tableau II : Diversité et densité des stocks pour les zones continentales majeures	22
Tableau III : Définition des haplogroupes du chromosome Y	38
Tableau IV : Composition ethnique de deux populations brésiliennes	44
Tableau V : Populations étudiées et nombre de chromosomes X analysés pour chacune d'elles.	50
Tableau VI : Composition de la solution CAS	56
Tableau VII : Polymorphismes de <i>Dys44</i>	58
Tableau VIII : Conditions d'amplification	59
Tableau IX : Sondes oligonucléotidiques pour hybridation ASO, avec leur T _m	64
Tableau X : Proportions des ingrédients nécessaires pour le marquage radioactif des sondes oligonucléotidiques	64
Tableau XI : composition des solutions utilisées lors de l'hybridation ASO	67
Tableau XII : description des haplotypes de <i>Dys44</i>	75
Tableau XIII : fréquence des haplotypes dans les différentes populations amérindiennes	76
Tableau XIV : Comparaison des populations amérindiennes d'Amérique du Nord et du Sud par le test du χ^2	83
Tableau XV : Comparaison des populations amérindiennes d'Amérique du Nord et Centrale par le test du χ^2	83
Tableau XVI : Comparaison des populations amérindiennes d'Amérique Centrale et du Sud par le test du χ^2	83
Tableau XVII : Matrices utilisées pour le test de Mantel	84
Tableau XVIII : fréquence des haplotypes dans les différentes populations amérindiennes, Européennes et Asiatiques	88
Tableau XIX : Comparaison des populations amérindiennes et européennes par le test du χ^2	96

Tableau XX : Comparaison des populations amérindiennes et asiatiques par le test du χ^2	96
Tableau XXI : Diversité génétique des différentes populations amérindiennes, asiatiques et européennes, telle que calculée par Arlequin	97
Tableau XXII : Valeurs des F_{ST} telles que calculées par Arlequin, à partir de la fréquence des haplotypes	100
Tableau XXIII : fréquence des haplotypes dans les différentes populations amérindiennes, métisses amérindiennes, Européennes et Asiatiques	106
Tableau XXIV : Comparaison des populations métisses amérindiennes et amérindiennes par le test du χ^2	109
Tableau XXV : Comparaison des populations métisses amérindiennes et européennes par le test du χ^2	109
Tableau XXVI : Comparaison des populations métisses amérindiennes et asiatiques par le test du χ^2	109
Tableau XXVII : Diversité génétique des différentes populations amérindiennes, métisses amérindiennes, asiatiques et européennes	110
Tableau XXVIII : Valeurs des F_{ST} telles que calculées par Arlequin, à partir de la fréquence des haplotypes	112

LISTE DES FIGURES

	page
Figure 1: Carte des présumés sites paléindiens sur le continent américain, et leur(s) datation(s).	13
Figure 2 : Fréquence des haplotypes fondateurs en Amérique	30
Figure 2b : Classification en haplogroupes des haplotypes du chromosome Y dans un arbre de parcimonie maximale	39
Figure 3 : Localisation des populations Amérindiennes étudiées	51
Figure 4 : Localisation des populations non-amérindiennes étudiées	52
Figure 5 : Courbe standard de lecture de la concentration en ADN	56
Figure 6 : Gel d'électrophorèse	60
Figure 7 : Hybridation ASO	62
Figure 8 : Ecran d'accueil de la base de données Pandora	68
Figure 9 : Recherche de données spécifiques à certains polymorphismes de <i>Dys44</i> dans Pandora	69
Figure 10 : Grilles de saisie des résultats d'hybridation	69
Figure 11 : Résultats présentés sous forme de génotypes	70
Figure 12 : Différentes possibilités d'analyse des fréquences dans Pandora	71
Figure 13 : Répartition de l'haplotype B001 chez les populations amérindiennes	77
Figure 14 : Répartition de l'haplotype B002 chez les populations amérindiennes	78
Figure 15 : Répartition de l'haplotype B003 chez les populations amérindiennes	79
Figure 16 : Répartition de l'haplotype B004 chez les populations amérindiennes	80
Figure 17 : Répartition de l'haplotype B006 chez les populations amérindiennes	81
Figure 18 : Répartition des autres haplotypes chez les populations amérindiennes	82
Figure 19 : Répartition de l'haplotype B001 chez les populations amérindiennes, européennes et asiatiques	89
Figure 20 : Répartition de l'haplotype B002 chez les populations amérindiennes, européennes et asiatiques	90
Figure 21 : Répartition de l'haplotype B003 chez les populations amérindiennes, européennes et asiatiques	91

Figure 22 : Répartition de l'haplotype B004 chez les populations amérindiennes, européennes et asiatiques	92
Figure 23 : Répartition de l'haplotype B005 chez les populations amérindiennes, européennes et asiatiques	93
Figure 24 : Répartition de l'haplotype B006 chez les populations amérindiennes, européennes et asiatiques	94
Figure 25 : Répartition des autres haplotypes chez les populations amérindiennes, européennes et asiatiques	95
Figure 26 : Analyse en composantes principales (PCA) des populations amérindiennes, européennes et asiatiques	98
Figure 27 : Analyse en composantes principales (PCA) des populations amérindiennes, européennes et asiatiques	99
Figure 28 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Asiatiques et les Européens de l'Ouest	101
Figure 29 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Asiatiques et les Européens de l'Ouest	102
Figure 30 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Asiatiques et les Européens de l'Ouest	103
Figure 31 : Répartition des différents haplotypes dans les populations amérindiennes et métisses amérindiennes	107
Figure 32 : Répartition des différents haplotypes dans les populations amérindiennes, métisses amérindiennes, asiatiques et européennes	108
Figure 33 : Analyse en composantes principales (PCA) des populations amérindiennes, métisses amérindiennes, européennes et asiatiques	111
Figure 34 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Métisses Amérindiens, les Asiatiques et les Européens de l'Ouest	113
Figure 35 : Voies de migration et colonisation des Amériques	123
Figure 36 : modèle expliquant l'origine des liens entre les Européens de l'ouest, les Sibériens et les Amérindiens.	124

LISTE DES SIGLES ET ABREVIATIONS

- ADN:** Acide désoxyribonucléique
- ASO:** *Allele Specific Oligonucleotide*
- BSA:** *Bovine Serum Albumin*
- CAS :** *Capillary Assay Solution*
- CM:** Centimorgan
- DHPLC:** *Denaturing High Performance Liquid Chromatography*
- D.O.:** Densité optique
- DMD:** *Duchenne Muscular Dystrophy*
- dNTP:** Deoxynucleotidine-5'-triphosphate
- EDTA:** *Ethylenediamine-tetraacetic acid*
- Kb:** Kilobase
- M:** Molaire, soit 1 mole par litre
- MDS:** *Multi-Dimensional Scaling*
- PCA:** *Principal Component Analysis*
- PCR:** *Polymerization Chain Reaction*
- RFLP:** *Restriction Fragment Length Polymorphism*
- SDS:** Sodium Dodecyl Sulfate, détergent
- SIL :** *Sociedade Internacional de Lingüística*
- SSC:** Saline-Sodium Citrate, solution tampon
- SSPE:** Saline-Sodium Phosphate-EDTA, solution tampon
- STR :** *Short Tandem Repeat*
- TE:** Tris EDTA, solution tampon
- Tm:** *Melting temperature*
- TNE :** Tris NaCl EDTA, solution tampon
- Tris:** Tris (hydroxymethyl) aminomethane
- UV:** Ultra-violet

*A mes parents,
Éliane et Jean-Claude.*

REMERCIEMENTS

Je tiens tout particulièrement à remercier mon directeur de recherche, le Dr Damian Labuda, pour m'avoir donné la chance de travailler dans son laboratoire alors que je venais d'un milieu scientifique éloigné de la biologie, ainsi que pour sa patience, ses enseignements et sa générosité.

Il me faut également remercier le Dr Daniel Sinnett, qui s'est avéré d'une aide comparable à celle d'un co-directeur de recherche et m'a toujours accordé du temps pour me conseiller, tant sur le plan scientifique que celui de ma carrière.

Je remercie grandement les collègues de mon laboratoire, le Dr Vania Yotova, le Dr Ewa Zietkiewicz, le Dr Alan Lovell, Mlle Claudia Moreau, M. Dominik Gehl et le Dr Fengxia Xiao, pour leur soutien scientifique et technique.

Je remercie toutes les personnes précédemment citées, ainsi que mes nombreux collègues des laboratoires voisins, pour la formidable ambiance de travail dans laquelle il m'a été possible de travailler, et leur amitié.

Enfin, je tiens à remercier nos collaborateurs, le Dr Michalski, le Dr Ruiz-Linares, le Dr Bedoya, ainsi que les populations Amérindiennes, sans lesquels cette étude n'aurait pas pu être effectuée.

Chapitre I

INTRODUCTION

La génétique des populations, qui s'intéresse à la structure des populations, aux migrations et à l'histoire démographique, peut sembler d'un intérêt purement académique au prime abord, mais s'avère aujourd'hui d'une importance grandissante en génétique médicale. Avec l'énorme expansion des données génétiques humaines ces dix dernières années, principalement grâce au projet de séquençage du génome humain, les outils statistiques développés pour la génétique des populations s'avèrent de grande utilité pour l'identification des fonctions des gènes ainsi que pour la découverte de mutations responsables de maladies.

Par exemple, les bases génétiques de nombreuses maladies monogéniques humaines ont été découvertes en analysant le co-héritage de l'allèle marqueur de la condition étudiée et d'autres allèles marqueurs dans des familles atteintes (analyse de liaison). La plupart des maladies courantes, comme le diabète, l'asthme ou de nombreux cancers sont complexes et ne peuvent être étudiés par une étude familiale classique, bien que de nouvelles approches dans des populations choisies semblent partiellement contourner le problème [1]. Les études d'associations, qui prennent principalement la forme d'études cas-contrôles, sont dans ce cas considérées comme la meilleure alternative pour étudier la base génétique de ces désordres [2-5] et l'étude du déséquilibre de liaison se montre dans ce cas d'une importance capitale [2, 6-8]. Comme l'intérêt pour ces méthodes statistiques augmente, l'analyse de ces méthodes est poussée et de nouvelles méthodes voient le jour [9-11].

Même si les études montrent que les différences entre les groupes ethniques comptent pour moins de 15% de la variation génétique de notre espèce [12-15], ces informations ethno-géographiques peuvent être très importantes pour les études d'épidémiologie génétiques. La *Food and Drug Administration* (FDA) a dernièrement approuvé la mise en place d'une étude clinique sur la combinaison de médicaments dans le traitement de déficiences cardiaques chez les Afro-Américains, après qu'une étude ait montré que l'enalapril, un inhibiteur de l'enzyme de conversion de l'angiotensine, avait une plus faible efficacité dans la population afro-américaine que dans la population américaine de descendance européenne [16, 17].

Bien entendu, les considérations médicales ne sont pas les seules, et ne furent d'ailleurs pas les premières, à pousser les chercheurs à étudier la génétique des populations. Le développement de notre intelligence et notre soif de savoir nous ont toujours poussé à nous poser des questions philosophiques sur notre place dans l'univers et l'origine de la vie. Que les convictions aient poussé les Hommes vers des réponses religieuses ou scientifiques, l'origine de notre espèce a toujours été au cœur d'un fervent débat. Aujourd'hui encore, même si un consensus qui place le berceau de notre espèce en Afrique a émergé, la question passionne toujours et les recherches continuent, tentant de trouver notre plus vieil ancêtre, la période à laquelle l'Homme a quitté l'Afrique, combien de vagues de migration ont eu lieu et par quelles voies.

Si les différentes ethnies que l'on peut rencontrer sur les « vieux » continents donnent naissance à de multiples débats concernant leur origine, peu ont suscité un intérêt aussi vif que les populations du Nouveau Monde. Depuis la récente découverte de ce continent, qui semblait inaccessible, et du fait que des humains y vivaient, la question de l'origine des Amérindiens a passionné les archéologues, anthropologues, linguistes et plus récemment les généticiens. Bien qu'il existe un intérêt médical à l'étude de ces populations, dans le cadre de l'épidémiologie génétique [18-20], de la pharmacogénétique [21-24] ou de l'étude de maladies génétiques rares et spécifiques [25, 26], une partie importante de la littérature consacrée à ces populations traite de leurs origines et du peuplement du continent américain [27-36].

Les études génétiques et les déductions concernant l'origine des différentes populations et ethnies, dont les Amérindiens, sont basées sur l'observation des variations génétiques existant actuellement dans ces populations. Les profils de variation observés peuvent avoir été générés de nombreuses façons par la démographie, les mutations et/ou les processus de sélection [37], en plus de l'histoire. Par conséquent les affirmations sur les populations anciennes à partir des populations actuelles sont issues d'un processus de déduction, souvent soumis à des modèles et hypothèses simplistes.

Cela souligne la nécessité d'intégrer aux études génétiques les informations provenant de l'archéologie, l'anthropologie et la paléontologie, qui se basent sur l'observation directe des populations anciennes à travers de ce qu'il en reste aujourd'hui. Les analyses d'ADN ancien pourraient en principe changer cette situation, mais leur impact a été bien moindre que prévu [38-43], principalement à cause de limitations techniques [44-47] et des contraintes sur la détermination de la contamination des échantillons [48]. C'est sans doute parce que l'Homme de Neandertal avait des allèles mitochondriaux notablement différents de ceux des populations modernes qu'il a été possible d'exclure les risques de contamination, permettant l'analyse directe de la génétique de cette ancienne population [49-52].

Par conséquent, la première partie de cette introduction sera consacrée aux données archéologiques, anthropologiques, linguistiques et paléoclimatiques disponibles pour les populations amérindiennes, ainsi que les débats qui les entourent. Les trois parties suivantes concerneront les études portant sur l'ADN mitochondrial, le chromosome Y et enfin les autosomes et protéines. La dernière partie introduira notre étude dans le contexte précédemment défini.

A) Données archéologiques, anthropologiques, paléoclimatiques et linguistiques

1. Paléoclimatologie

La nature épisodique des migrations humaines pourrait être due en partie au climat. Il est connu que les changements climatiques au cours des deux derniers millions d'années ont été un facteur déterminant des changements dans la distribution de beaucoup d'espèces animales et végétales [53-55], et cela pourrait également être le cas pour les différentes espèces *Homo* [56]. La connaissance des conditions climatiques sur le continent américain et le nord-est du continent asiatique, durant le Pléistocène et le début de l'Holocène, est indispensable à l'établissement de théories concernant le peuplement des Amériques.

La plupart des publications en paléoclimatologie, archéologie et anthropologie utilisent une classification temporelle géologique. On découpe les temps géologiques en ères géologiques, les ères en périodes et les périodes en époques. Le tableau I page 6 récapitule cette classification. Les dates, dans le tableau I ainsi que dans le reste de ce chapitre, sont exprimées en années radiocarbone avant le présent (YBP). Les années radiocarbone avant le présent sont en fait les années radiocarbone avant 1950, période à partir de laquelle les explosions atomiques ont modifié le taux de carbone 14 dans l'atmosphère. De plus, il a été montré que l'activité du carbone 14 dans l'atmosphère et dans les autres réservoirs (océans et biosphère), et donc également celle de l'activité initiale des échantillons datés, a varié dans le temps [57], et une correction est nécessaire pour passer des années radiocarbone aux années réelles. Afin de simplifier la lecture, toutes les dates seront exprimées en années, sans rappeler qu'il s'agit d'années radiocarbone, sauf mention spéciale.

Millions d'années	Eres géologiques	Période	Epoque	
4000	Précambrien	Archéen		
2500		Protérozoïque		
570	Primaire ou Paléozoïque	Cambrien		
510		Ordovicien		
438		Silurien		
410		Dévonien		
355		Carbonifère		
290		Permien		
250		Secondaire ou Mésozoïque	Trias	
205	Jurassique			
135	Crétacé			
65	Tertiaire ou Cénozoïque	Paléogène	Paléocène	
53			Eocène	
37		Néogène	Oligocène	
23			Miocène	
5			Pliocène	
1,65				Pléistocène
0,010		Quaternaire		Holocène

Tableau I : Classification des ères géologiques

L'histoire du Pléistocène est dominée par les grandes glaciations. Entre les périodes froides, des périodes de réchauffement, ou interglaciaires, permettaient à la flore et à la faune de reconquérir les territoires abandonnés par la glace. Dans les régions tropicales, ces fluctuations climatiques se traduisirent par des alternances de périodes pluviales et interpluviales, enregistrées par les niveaux des grands lacs fonctionnant comme de vastes pluviomètres.

L'Alaska, le Territoire du Yukon et le nord-est de la Sibérie partagent une histoire unique lors de la dernière ère glaciaire. A peu près toutes les autres régions de hautes latitudes de l'Hémisphère Nord étaient recouvertes d'une calotte glaciaire lors des glaciations du Pléistocène, contrairement aux plaines des régions mentionnées plus haut, qui restèrent non gelées pendant une grande partie du Pléistocène.

Le niveau de la mer subit aussi des variations de l'ordre de 100 à 200 m, baissant pendant les glaciations alors que des masses importantes d'eau étaient emprisonnées sous forme de glace continentale. Ainsi, le niveau global de la mer était environ 120m plus bas que le niveau actuel pendant la dernière glaciation.

Les plates-formes continentales entre l'Alaska et la Sibérie sont pour la plupart moins de 100m sous le niveau actuel de la mer, et étaient par conséquent exposées lors de la dernière glaciation, formant un pont entre les terres non gelées de Sibérie, d'Alaska et du Yukon. Cette grande région est connue sous le nom de région de Béring.

Les études paléoclimatiques ont montré que la région de Béring avait un climat assez froid pour le développement et la persistance d'une calotte glaciaire, mais était trop aride pour un tel développement, comme les dépôts et les dunes de sable de la fin du Pléistocène en attestent.

Les estimations des paléotempératures du dernier maximum glaciaire ont été réalisées à partir d'assemblages d'insectes (scarabées) fossilisés en Alaska et au Yukon [58] et au nord-est de la Sibérie [59]. Ces estimations sont basées sur la méthode de l'amplitude climatique mutuelle (*Mutual climatic Range* ou MCR). Cette méthode prône la détermination d'un ensemble de paramètres climatiques compatibles avec toutes les espèces se trouvant dans l'assemblage. La première étape est donc la construction des paramètres climatiques de chaque espèce. Certaines espèces de scarabées, appartenant aux espèces sténothermiques, ont des tolérances climatiques très précises. Certaines de ces espèces ne se trouvent plus aujourd'hui que dans une région restreinte et unique du monde, et d'autres dans de multiples régions différentes, mais précises, aux conditions climatiques similaires (sommets de montagnes ou parties isolées de déserts). Superposer les paramètres climatiques de chacune des espèces sténothermiques d'un assemblage permet la détermination d'un ensemble climatique précis de la région à l'époque considérée.

Les assemblages de fossiles de la région de Béring, dans une période de temps de 20 000 à 18 000 ans avant le présent, permirent d'estimer que la température annuelle maximale il y a 20 000 ans était de 5°C inférieure à la température annuelle maximale actuelle, et 0,9°C plus chaude il y a 18 000 ans qu'actuellement, soit environ 12-13°C. Les températures minimales ne semblent pas différer de plus de quelques degrés Celsius des températures minimales actuelles. Par conséquent, la différence majeure entre le dernier maximum glaciaire et le climat actuel dans cette région réside dans une continentalité accrue. Cela montre que les conditions climatiques dans la région de Béring étaient relativement douces en comparaison des refroidissements dramatiques des étés et des hivers dans les régions au sud de la calotte glaciaire continentale.

La fin de la période glaciaire, 14 000 à 10 000 ans auparavant, amena une série de changements environnementaux de grande échelle dans la région de Béring. Les fluctuations climatiques ont provoqué d'importants changements dans la distribution des plantes et des animaux de la région, et pourraient avoir joué un rôle primordial dans l'extinction de nombreuses espèces de mammifères [60]. D'après les données tirées des analyses des scarabées de la région [61], les températures annuelles maximales ont commencé à augmenter il y a au moins 12 000 ans, atteignant des niveaux supérieurs aux niveaux actuels il y a 11 000 ans. En Alaska Arctique, les températures maximales annuelles dépassaient de plus de 7°C les températures maximales annuelles actuelles. Ce changement climatique important était probablement dû à la combinaison d'un plus grand insolemment de la région, et de l'inondation du détroit de Béring, qui amena des eaux relativement chaudes du Pacifique vers la côte Nord-Ouest de l'Alaska.

Durant la dernière glaciation, la région de Béring fut le refuge majeur des grands mammifères adaptés à la vie dans les hautes latitudes de l'Hémisphère Nord. Il y avait une abondante diversité de grands mammifères, comme il n'en a jamais plus existé dans les régions de hautes latitudes. Le seul écosystème moderne qui démontre une variété similaire d'animaux herbivores est la Savane est-africaine. On pouvait trouver des mammoths, des rhinocéros laineux, des antilopes saïga, des chevaux du Pléistocène, des bisons à larges cornes et deux espèces de musk-ox. On pouvait également trouver des caribous, dont l'alimentation comporte, en plus des herbes, des lichens et arbustes.

Les mammoths semblent avoir dominé les régions arctiques, alors que les chevaux du Pléistocène et les bisons dominaient les régions subarctiques [62]. Ces herbivores disparurent à la fin de la dernière glaciation. Pour les autres herbivores, seul le musk-ox vit encore en Alaska, et l'antilope saïga disparût de l'Alaska et du Yukon il y a environ 13 000, mais se retrouve encore aujourd'hui en Asie Centrale où un habitat aride et herbeux persiste. Les autres grands animaux vivant encore dans ces régions ne sont pas des herbivores stricts, et sont les élans, les caribous et les chèvres des montagnes.

Mais là où il y a un grand nombre d'herbivores, il y a des prédateurs, et la région de Béring n'échappe pas à cette règle. On pouvait trouver des lions du Pléistocène, des tigres aux dents de sabre, et l'ours géant à petite tête, pour ce qui est des espèces qui se sont éteintes, et des ours polaires, des ours grizzly, des loups et des lynx, qui subsistent encore dans ces régions. De tels prédateurs pourraient avoir rendu la région assez dangereuse pour les populations humaines.

Une des hypothèses concernant le peuplement des Amériques est que les humains arrivèrent sur le continent américain en traversant la région de Béring, pourchassant du gros gibier, puis migrèrent vers le sud par un couloir ouvert dans les glaciers dans l'Ouest canadien. Outre le fait que la région ait pu être assez inhospitalière à cause de la présence de nombreux prédateurs, et l'absence de forêts (donc des moyens limités pour faire du feu), la découverte du site archéologique de Monte Verde dans le sud du Chili [63] datant d'il y a 12 500 ans place les humains sur le continent américain bien avant la fin de la dernière glaciation (affirmation discutée dans la partie 2). L'ouverture d'un couloir dans les glaciers est aujourd'hui estimée avoir eu lieu il y a seulement 12 400 ans environ [64]. Il apparaît donc que les humains ont quitté la région de Béring avant la dernière période glaciaire, soit il y a 60 000 à 30 000 ans [60] (période interglaciaire mid-Wisconsin, pendant laquelle le passage terrestre était dégagé), ou bien qu'ils sont entrés en Amérique par une autre voie.

2. Archéologie, anthropologie et linguistique

a. Données des sites archéologiques, anthropologie

Que la région de Béring ait joué un rôle déterminant dans le peuplement du continent américain ou non, les évidences archéologiques des vallées des rivières Tanana et Nenana [65-69] montrent que l'est de la région était peuplée il y a environ 11 800 ans (d'après les critères récents de validité, comme discuté ci-après). Trois sites possiblement plus anciens ont été trouvés, Bluefish Caves, Lime Hills Caves et Trail Creek Caves, mais un consensus n'a pas encore été atteint concernant la validité de ces sites.

La principale théorie concernant le peuplement du continent américain par des chasseurs de gros gibiers d'origine mongoloïde il y a environ 12 000 ans [70-77] était basée sur de nombreux sites archéologiques nord-américains, semblant mettre en évidence le développement en Amérique du Nord d'une même culture, la culture Clovis. La culture fut ainsi nommée en référence aux pointes taillées particulières qui la caractérisent, qui furent originellement découvertes dans un site archéologique près de la ville de Clovis, au Nouveau Mexique, dans les années trente. Ce mode de vie, dans un environnement ouvert et tempéré était typique dans cet hémisphère, et ce n'est seulement qu'après l'extinction du gros gibier, le réchauffement du climat et la montée du niveau de la mer que ce mode de vie a changé, poussant les populations vers d'autres régions, dont les côtes [71, 77, 78].

Cette élégante théorie était basée sur des données archéologiques anciennes, et les techniques d'excavation et de datation de l'époque, ainsi que les critères d'acceptation, étaient assez rudimentaires et imprécises par rapport à celles présentement disponibles. Depuis lors, de grandes avancées dans différents domaines scientifiques comme la taphonomie, l'étude des processus de dépôt des sédiments et des objets, et de nouvelles méthodes de datation au carbone 14, ont permis de poser un regard nouveau sur les découvertes archéologiques en Amérique. Ces progrès ont mené à l'établissement de critères précis pour l'évaluation des sites archéologiques paléindiens anciens [79-91]. Par exemple, des séries de datations statistiquement précises et ayant une erreur standard de 300 ans au plus sont actuellement exigées pour la validation incontestable d'un site. Les objets datés doivent être bien identifiés, propres et stratigraphiquement associés à des objets manufacturés. A cet égard, les fruits d'espèces botaniques bien identifiées et les acides aminés extraits d'os ou de dents d'animaux sont particulièrement fiables. De plus, la validation de cultures régionales doit reposer sur l'étude de différents sites répondant aux critères précédemment cités, et présentant des datations, des objets manufacturés et des restes biologiques de même nature.

Dans ce nouveau contexte, les données concernant la civilisation Clovis ont beaucoup évolué. La plupart des sites considérés comme appartenant à cette culture se trouvent en fait dans le sud-est de l'Amérique du Nord, le cœur de la culture Clovis, et non dans la zone que cette civilisation était sensée avoir colonisée. La grande partie des sites des supposés descendants de la culture Clovis, dans la zone de colonisation, étaient pauvres en restes biologiques et culturels et manquaient de séries de datations fiables plaçant les sites au Pléistocène. Les restes biologiques utilisés pour les datations étaient hautement sensibles à la contamination, et certains ont maintenant été expérimentalement trouvés contaminés par de grandes quantités de carbone géologique [85, 92-96].

Seuls cinq sites situés dans les hautes plaines (Lehner, Murray Springs, Domebo, Dent et Anzick) ont des séries de datations fiables. Les séries de datation n'ont donné aucune date fiable antérieure à 11 200 ans, le spectre s'étendant jusqu'à 10 800 ans, avec une moyenne de 10 900 ans [97]. La civilisation Clovis est donc plus jeune que l'on supposait, et n'est plus la civilisation la plus ancienne en Amérique du Nord, les sites de Tanana et Nenana étant plus anciens. Cela signifie également que de nombreuses autres cultures sont contemporaines à la culture Clovis, et non ses descendantes. Tous les sites archéologiques américains importants sont présentés dans la figure 1 page 13, ainsi que leur(s) datation(s).

Les évidences archéologiques et biologiques actuelles ne supportent pas la théorie de la migration Clovis. Par exemple, il n'y a aucune trace de pointes caractéristiques de cette civilisation lors de la fin du Pléistocène au point d'entrée du continent américain, toutes celles découvertes datant du début de l'Holocène [68, 98].

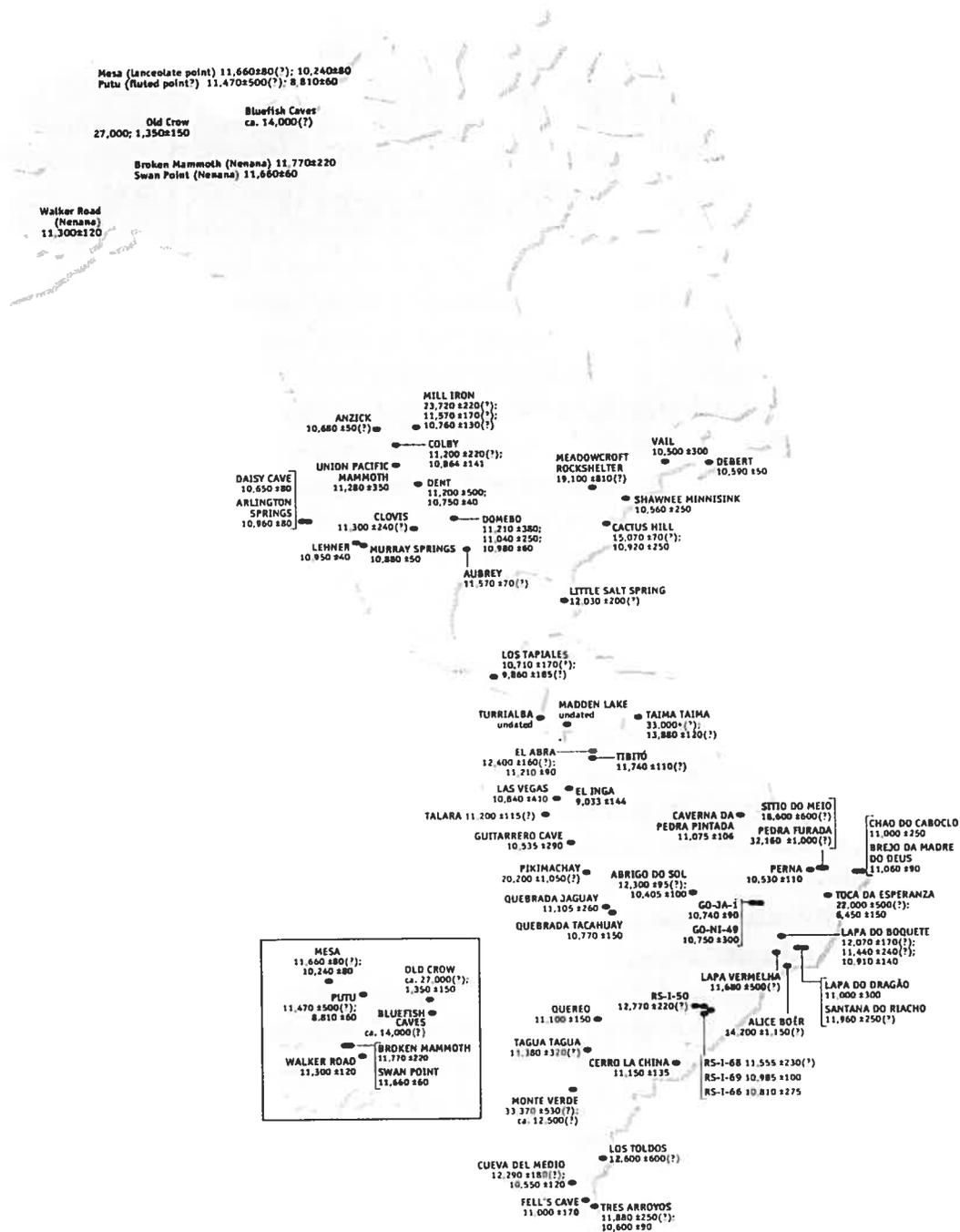


Figure 1 : Carte des pr sum s sites pal oindiens sur le continent am ricain, et leur(s) datation(s). Les points d'interrogation indiquent que la datation est discutable. [97]

Malgré des efforts intensifs, aucune recherche n'a abouti à la découverte de cultures similaires à la culture Clovis à l'intérieur du nord-est de l'Asie. Les cultures sibériennes de l'époque utilisaient des pointes très différentes de celles utilisées par les Clovis et avaient des moyens de subsistance diversifiés [99-101]. Les cultures de la fin du Pléistocène le long des côtes au nord de l'Asie ont également des moyens de subsistance diversifiés et d'autres types de pointes [102, 103], dont certaines étaient taillées avec des techniques caractéristiques des Clovis et d'autres Paléindiens. La technique de taille bifaciale des Clovis et certaines formes d'outils, parfois utilisées comme preuve de l'origine des Paléindiens dans la culture solutréenne de la fin du Pléistocène en Europe [90, 104], étaient également communes à la fin du Pléistocène dans l'est de l'Asie [99, 101-103]. D'autres formes d'outils sont particulières à la civilisation Clovis et ne semblent avoir aucun lien avec des civilisations européennes ou asiatiques [74].

L'interprétation de la biologie du squelette des Paléindiens, en terme d'origines géographiques, a également beaucoup changé au cours des dernières décennies. Les caractéristiques métriques et non-métriques ont été revues pour environ 25 squelettes directement datés de la fin du Pléistocène et du début de l'Holocène [105-109]. Cela révéla que la morphologie crâniofaciale des Paléindiens différait de celle des Amérindiens nord-américains modernes et des peuples mongoloïdes comme les Japonais ou le Chinois Han. Bien qu'il ait été suggéré par quelques archéologues que les Paléindiens pourraient alors descendre d'Européens de type caucasoïde [110-114], les traits crâniens étudiés ne sont pas considérés comme étant sous un contrôle génétique poussé [115] et étaient de moins en moins communs en Europe à la fin du Pléistocène [115].

De multiples caractères des dents humaines connus comme étant sous contrôle génétique sont très communs chez les Paléindiens et les populations du nord-est de l'Asie. La forme des incisives, les motifs des bords des molaires et le nombre de racines chez les Paléindiens ont un fort taux de similarité avec ce que l'on retrouve chez les populations de l'est de l'Asie, mais un faible taux de similarité avec les dents

provenant d'Européens de la fin du Pléistocène et de l'Holocène [115]. Comme il sera discuté dans la partie concernant l'ADN mitochondrial, même l'haplogroupe mitochondrial X, qui semblait relier les Paléoindiens aux Européens, est maintenant trouvé en Asie. Il n'y a donc aucune preuve empirique statistiquement significative qui relie les Paléoindiens plus fortement aux Européens qu'aux Asiatiques, mais de nombreuses données les relient fortement à l'est de l'Asie.

Quant aux sites Clovis et pré-Clovis en Amérique Centrale et du Sud, aucun ne s'est avéré indiscutable. Les pointes originellement identifiées comme étant de culture Clovis soit datent de l'Holocène, soit s'avèrent ne pas correspondre aux critères définissant la culture Clovis. De plus il n'y a toujours aucune preuve de la chasse de gros gibiers par les chasseurs de la fin du Pléistocène dans ces régions. Toutes les cultures dont la nature a été adéquatement établie et publiée ne dépassent pas les 11 200 ans, et toutes comportent quelques objets manufacturés bifaciaux et pas exclusivement monofaciaux.

Par contre, diverses cultures paléoindiennes régionales contemporaines à la culture Clovis, mais chassant une vaste gamme de gibiers et vivant également de la cueillette, utilisant des pointes triangulaires, ont été trouvées dans divers contextes environnementaux. Certaines théories plaçaient ces caractéristiques post-Pléistocène, et considéraient que les pointes triangulaires étaient de type archaïque, en analogie avec des transitions dans certaines régions d'Amérique du Nord (passage de pointes de type Clovis au type archaïque triangulaire).

La datation radiométrique de différentes pointes triangulaires, en Amérique du Nord et du Sud, établit clairement qu'elles datent de la fin du Pléistocène et précèdent donc les formes archaïques de pointes triangulaires. De plus les pointes triangulaires trouvées en Amérique du Sud sont morphologiquement distinctes des pointes triangulaires archaïques. Les datations et la diversité des sites montrent que ces caractéristiques étaient communes à la fin du Pléistocène, et que ces formes de

pointes étaient probablement connues des premiers immigrants plutôt que le fruit d'un développement local.

Le plus célèbre site pré-Clovis d'Amérique du Sud est sans doute le site de Monte Verde, à la pointe Sud du Chili vers la côte du Pacifique. Ce site a été décrit comme datant possiblement de plus de 33 000 ans [116, 117], mais beaucoup de questions à propos de ce site restent en suspens et le site a été critiqué pour différentes raisons [118, 119]. Tout d'abord la stratigraphie de la zone est très difficile à interpréter sans le repérage spécifique des objets trouvés, y compris leur datation, dans des coupes avec une échelle de mesure précise. La nature de l'association des quatre objets indubitablement manufacturés, des spécimens fauniques et botaniques avec les dates du radiocarbone est floue. De plus aucun débris des tailles de pointes n'a été trouvé, alors qu'ils sont en général bien plus nombreux que les pointes elles-mêmes.

La datation présente également un autre problème : deux datations de 33 000 ans, bien que traitées par les archéologues du site comme étant des dates culturelles, sont caractérisées dans la notation des laboratoires de datation sous la forme de dates infinies, ce qui provient de carbone trop vieux pour encore contenir du carbone 14. Il existe de nombreuses sources possibles de vieux carbone pour ce site, depuis la présence de bitume sur les outils, substance présente dans les dépôts marins de pétrole proches du site, à la présence de détergents et d'essence dans les eaux polluées ruisselantes. La datation des pollens est également en conflit avec les insectes trouvés, suggérant que du matériel de différents âges ont été mélangés.

Une revue récente des données par une équipe d'archéologues nord-américains résulta dans une déclaration de support moral pour le site de Monte Verde comme étant pré-Clovis et datant de 12 500 ans [120]. Toutefois, ce consensus a récemment disparu après la circulation du second rapport publié sur le site, et des archéologues se sont prononcés contre la validité de l'occupation pré-Clovis du site de Monte Verde [118, 121].

b. Voyages maritimes

Les développements récents en paléoclimatologie, qui placent l'ouverture d'un couloir dans les glaciers à 12 400 ans, et laissent entendre que la survie des humains dans les conditions environnementales de l'époque étaient probablement très précaires [60, 122-125], l'inondation de la région de Béring il y a 11 000 ans, le manque de sites archéologiques datant du Paléocène dans la région du couloir entre les glaciers, ainsi des datations de sites archéologiques d'Amérique du Sud de plus de 12 500 ans (bien que cela soit grandement discuté) poussent vers l'établissement d'une nouvelle théorie du peuplement des Amériques, alternative à la théorie « terrestre ». Même si les sites archéologiques d'Amérique du Sud dont la datation est présentement confirmée datent au plus de 11 200 ans, la période de temps entre l'ouverture du couloir et l'arrivée des humains en Amérique du Sud semble relativement courte.

Les recherches ont également montré que les côtes de l'Alaska et de la Colombie Britannique n'étaient pas couvertes par la calotte glaciaire pendant une grande partie du Pléistocène et étaient plus propices à la vie humaine qu'on le supposait précédemment [126-133].

C'est sans doute la sous-estimation des capacités d'exploitation des ressources marines par les humains de la fin du Pléistocène qui a poussé à la longue marginalisation de l'idée que le peuplement des Amériques puisse avoir été maritime, en longeant les côtes.

Un consensus concernant le peuplement de la planète par l'Homme Anatomiquement Moderne est lentement apparu depuis les quinze dernières années. Il aurait quitté l'Afrique il y a environ 150 000 à 200 000 ans et se serait rapidement répandu dans le monde, remplaçant les autres hominidés déjà présents. Comme prédit par cette théorie, les restes fossiles les plus anciens de l'Homme Moderne sont trouvés en Afrique [134-136] et datent d'entre 120 000 et 60 000 ans.

Des recherches récentes dans la région de la rivière Semliki, au Zaïre, ont permis la découverte de harpons (en os) et d'une pêche intensive dans cette région il y a 90 000 à 80 000 ans [137-139].

C'est également à cette époque que les restes humains les plus anciens sont trouvés hors d'Afrique. Curieusement, ils n'ont pas colonisé la plus grande partie de l'Europe avant encore 45 000 à 55 000 ans [140], peut-être à cause de la présence de l'Homme de Neandertal. Par contre l'Homme Moderne s'était répandu dans le sud de l'Asie il y a au moins 70 000 à 60 000 ans. De là, en quelques millénaires, ils ont probablement fait les multiples voyages maritimes, d'île en île d'Asie du Sud-Est, nécessaires au peuplement de Sahul, le grand continent qui unissait l'Australie, la Nouvelle Guinée et la Tasmanie pendant les périodes de bas niveau de la mer (périodes glacières).

Une théorie anthropologique avançait que les adaptations maritimes et la navigation étaient apparues tardivement dans notre espèce, et n'existaient pas lors du Pléistocène, principalement basée sur le manque de preuves archéologiques de telles activités. D'un point de vue géologique, il y a toutes les raisons de penser que les preuves archéologiques d'adaptation côtière sont sérieusement sous-estimées, principalement à cause des changements dramatiques du niveau global la mer lors des périodes glacières, causant d'importants, et très variables, changements dans la géographie côtière. De nombreux sites ont donc probablement été inondés lors de la montée du niveau de la mer, détruits par l'érosion côtière associée ou encore les deux.

Lors de la montée et descente du niveau de la mer, les paysages côtiers se déplacent généralement de façon latérale. Les mouvements latéraux maximaux des lignes côtières pendant les 20 000 dernières années, par exemple, ont varié jusqu'à 1000 km dans certaines zones (nord de l'Australie), et de moins d'un kilomètre dans d'autres. Mais les régions dans lesquelles les lignes côtières se sont déplacées de 10 km ou moins sont rares et semblent corrélées avec des sites d'occupation côtière relativement anciens [141].

Par conséquent les sites côtiers actuels étaient pour la plupart éloignés de plus de 10 km des côtes lors du Pléistocène. L'étude des chasseurs-cueilleurs côtiers actuels suggère qu'ils voyagent rarement plus de 5 ou 10 km pour la recherche de nourriture et ne transportent pas les restes de leur chasse si loin [142, 143].

Si certains sites suggèrent une certaine utilisation des ressources marines par *Homo erectus* et l'archaïque *Homo sapiens*, il semble que l'utilisation intensive des ressources marines soit liée à l'Homme Moderne. Les preuves de voyage maritime systématique et sophistiqué lors du Pléistocène viennent de plusieurs régions du monde, mais la plupart proviennent d'Australie, de la Mélanésie et de l'est de l'Asie. Les preuves de voyages maritimes de 20 à 200 km dans ces régions sont maintenant très bien documentées.

L'une des raisons majeures du changement des théories sur le voyage maritime est probablement la démonstration non équivoque que les humains avaient atteint l'Australie il y a au moins 20 000 ans [144], une migration qui a nécessité de multiples traversées maritimes entre les îles de l'Asie du Sud-Est, même lorsque le niveau de la mer baissait de plus de 125 mètres lors des glaciations. Au milieu des années 70, le registre archéologique australien a été étendu à environ 34 000 ans au lac Mungo [145], et des recherches subséquentes poussèrent cette datation à 40 000 ans [146], 50 000 ans [147] et maintenant possiblement 60 000 ans ou plus [148].

Les données australiennes étaient déconcertantes pour deux raisons. Tout d'abord, lors des temps historiques, les aborigènes australiens ont été décrits comme n'ayant pas d'embarcations sophistiquées capables de traversées maritimes substantielles [149]. De plus, comme la plupart des autres sites dans le monde, il n'y avait pas de trace en Australie de débris de coquillages ou d'autres preuves d'utilisation des ressources marines datant du Pléistocène. En fait, la grande majorité des débris de coquillages trouvés en Australie étaient plus récents que 5000 ou 6000 ans.

Des données confirmant le scénario de colonisation de Sahul parvinrent lors des années 80 par la découverte de débris de coquillages datant du Pléistocène (la datation actuelle donne au moins 35 000 ans) dans l'archipel de Bismarck et les îles Salomon, à l'est de la Nouvelle Guinée [150-152]. En plus de la confirmation du voyage maritime dans cette région, les sites découverts présentaient entre autre des restes de coquillages et de poissons, comme cela peut être attendu pour des civilisations maritimes. Contrairement à la situation en Australie, les côtes proches des sites plongent rapidement et profondément dans la mer, et les fluctuations du niveau de la mer lors des glaciations ont eu peu d'effet sur la géographie côtière locale.

Les sites mélanésiens semblent également montrer que les capacités de navigation ont grandement progressé entre 35 000 ans et 15 000 ans. Alors que la migration initiale vers Sahul, la Nouvelle Angleterre et la Nouvelle Irlande nécessita des voyages atteignant près de 100 km, le peuplement de Buka, dans les îles Salomon, il y a au moins 28 000 ans, nécessita une traversée maritime d'au moins 140 km et peut-être même autant que 175 km [153]. Il y a 15 000 ans, l'une des dates possibles de la première colonisation du continent Américain, les navigateurs mélanésiens avaient atteint l'île de Manus, qui imposa un voyage ininterrompu de 200 à 220 km, alors que l'île était donc hors de vue depuis la terre la plus proche [153].

Il est clair que les eaux relativement chaudes de l'Asie du Sud-Est et de la Mélanésie opposèrent aux navigateurs du Pléistocène un tout autre défi que celui des eaux plus froides du Pacifique Nord et de la région de Béring [154]. Des preuves de navigation dans cette région du monde viennent des îles du Japon. Alors que le Japon lui-même était relié au reste de l'Asie pendant les périodes de très bas niveau de la mer, et par conséquent que sa colonisation ne nécessita pas obligatoirement de navigation, la présence humaine sur l'île d'Okinawa il y a environ 32 000 ans [155], et sur de plus petites îles de la chaîne de Ryukyu (avec des datations de 26 000 et 15 000 ans), démontrent l'utilisation de la navigation dans cette région, une des distances de traversée nécessaires pour atteindre Okinawa depuis le Japon étant

d'environ 75 km. De plus longs voyages, jusqu'à 150 km, ont même été nécessaires pour atteindre d'autres îles, comme l'île de Miyako (datation d'un site à 26 000 ans) [155]. D'autres données archéologiques montrent qu'au Japon même des traversées d'environ 50 km étaient effectuées il y a 21 000 ans [156].

La présence de navigateurs dans l'archipel japonais lors du Pléistocène est importante et démontre que la navigation dans les eaux du Pacifique Nord existait à une époque suffisamment reculée pour être compatible avec la théorie de la colonisation initiale du continent américain par voie maritime.

3. Linguistique

Les linguistes identifient les groupes de langues apparentées à différents niveaux taxonomiques. Le niveau appelé stock est le plus haut niveau qui puisse être reconstruit par les méthodes comparatives standards de la linguistique historique. Le terme « famille » est un terme générique pour des groupes généalogiques de toutes sortes, incluant les stocks qui sont un type particulier de famille. Il existe environ 6000 langues dans le monde, tombant dans environ 250 groupes généalogiques différents (stocks) et un grand nombre de types structuraux différents, voir tableau II ci-après. Il est important de noter que certains stocks ne sont définis que par une seule langue dans des populations précises (par exemple le Basque), alors que d'autres regroupent plus de 1000 langues (par exemple l'Austronésien).

Zone	Stocks	Densité en stocks	Langages	Langages par stock	Date de colonisation
Afrique	20	4,4	2614	130,7	>100
Asie du Sud et du Sud-Est	10	3,8	1998	199,8	60
Nouvelle Guinée	27	227,3	1109	41,1	50
Australie	15	13,0	234	15,6	50
Eurasie du Nord	18	3,3	732	40,7	40
Amériques	157	27,2	1219	7,8	12

Tableau II : Diversité et densité des stocks pour les zones continentales majeures [157-161] . Le nombre des langages a été obtenu en additionnant les totaux des pays, et certains sont comptés deux fois [159]. Toutefois les valeurs restent représentatives des zones considérées [157, 159].

En se basant sur des données linguistiques, génétiques et dentaires, Greenberg classa les langues amérindiennes en trois groupes distincts : l'Eskimo-Aleut, le NaDene et l'Amerind, ces groupes étant eux-même décomposés en sous-groupes [162, 163]. L'Eskimo-Aleut a été classé comme appartenant au groupe Eurasiatique, y rejoignant le Chukotian, le Gilyak (Nivkhi), le Japonais, le Coréen, l'Ainu, l'Altaïque, l'Uralique-Yukaghir et l'Indo-Européen. C'est sur cette base qu'il énonça la théorie qui porte maintenant son nom et propose que les Amériques aient été peuplées en trois vagues distinctes de migration, chacune donnant naissance à l'un des groupes linguistiques, le groupe le plus récent étant l'Eskimo-Aleut, suivi du NaDene et l'Amerind étant le plus ancien.

D'autres travaux se basent sur les stocks et leur densité [157, 158, 164-169]. A cet égard les Amériques sont exceptionnellement riches, puisque plus de la moitié des stocks sont amérindiens (tableau II page 22). Les stocks les plus anciens sont approximativement datés à 6000 ans, comme par exemple l'Indo-Européen, l'Uralique et l'Austronésien. En considérant que pour un stock, par période de 6000 ans, le nombre moyen de ramifications qui survivent est de 1,5 [164] et que chacune de ces ramifications va donner naissance à un stock qui lui-même donnera naissance à 1,5 stocks et ainsi de suite, permet de proposer des scénarios concernant la date d'entrée des humains sur le continent américain et le nombre de migrations.

Sur cette base, si l'on ne considère que les stocks appartenant à l'Amerind, censé représenter la première vague de migration au Nouveau Monde, et une seule vague de migration, il faudrait que cette lignée ait commencé à se ramifier il y a au moins 50 000 ans pour avoir donné naissance aux stocks actuels [164]. Cette date étant hautement improbable, il semble plus sage de considérer de multiples migrations. Deux scénarios semblent alors compatibles avec les données linguistiques actuelles. Il aurait fallu 10 à 12 vagues de migration, ayant débuté il y a 20 000 à 30 000 ans [164], ou encore peu d'entrées avant le dernier maximum glaciaire, plusieurs par la suite et un plus grand taux de prolifération pour les premiers migrants suivant leur entrée sur le continent (par rapports aux migrants arrivés plus tardivement) [169].

La comparaison des stocks amérindiens avec ceux des autres continents indique que les liens linguistiques les plus forts sont trouvés avec des langues asiatiques de Mélanésie et d'Asie du Sud-Est, et non avec les langues d'Asie du Nord ou Centrale. Après leur entrée en Amérique, la radiation linguistique semble avoir eu lieu des côtes vers l'intérieur des terres, et le plus souvent du Sud vers le Nord.

De nombreux doutes subsistent concernant les travaux précédemment cités, pour deux raisons principales : tout d'abord, de petits changements dans les paramètres utilisés peuvent générer des dates variant entre 12 000 et 91 000 ans, entraînant une très grande incertitude, et les hypothèses utilisées dans la méthode de datation précédemment mentionnée semblent non fondées (voir paragraphe suivant).

L'unité de temps utilisée, soit 6000 ans, est hautement hasardeuse. Tout d'abord, il n'y a aucun point fixe contre lequel calibrer la date, en notant au passage qu'un stock est défini par un degré de similarité linguistique et non par une datation connue, et il est assez peu probable que le taux de changement linguistique soit constant [170]. La logique sous-jacente est que la diversité croît linéairement avec le temps jusqu'à l'établissement éventuel d'un équilibre. Théoriquement il n'y a aucune raison de penser que la diversité soit une fonction linéaire du temps. La divergence des langues commence quand des événements démographiques ou sociaux conduisent à l'éclatement en groupe de ce qui constituait avant une communauté homogène. Rien ne permet de penser que de tels événements se produisent à un taux constant.

La théorie du taux de ramification constant des langues peut être testée en combinant les données archéologiques avec le nombre et la densité des stocks sur chaque continent. Les données montrent que le nombre de stocks n'a pas tendance à augmenter avec le temps, voir tableau II page 22. En fait la tendance serait plutôt à l'inverse, bien que non significative.

Il ressort des données que les Amériques présentent une diversité exceptionnelle qui n'est retrouvée sur aucun autre continent. Le nombre absolu des stocks en Amérique apparaît toutefois beaucoup moins important lorsque la diversité est exprimée relativement à la taille du continent (densité des stocks) et c'est en fait la Nouvelle Guinée qui présente une densité extrême des stocks. Par contre la corrélation entre la densité des stocks et le temps n'est pas significative. De plus, il est assez artificiel d'utiliser la surface brute d'un continent pour évaluer sa densité en stocks, les densités de population variant grandement d'un continent à un autre. Par exemple, la Nouvelle Guinée est luxuriante et offre une niche écologique à plus de 1000 communautés [159].

Une mesure plus appropriée de la diversité linguistique relative pourrait être le nombre de stocks rapporté au nombre de groupes sociaux habitant la zone considérée. Le nombre de groupes sociaux peut être approximé par le nombre de langues, en assumant qu'en général les groupements linguistiques coïncident avec les groupements socio-économiques [159]. Le nombre de langues par stock est fortement, et positivement, relié à la durée d'habitation, c'est-à-dire que plus une population continentale est « âgée » (il s'agit ici du temps écoulé depuis la colonisation du continent en question, d'après les données archéologiques) moins elle contient de stocks, et plus ces stocks comportent de langues. L'Afrique et l'Eurasie ont peu de stocks, regroupant chacun un grand nombre de langues, les Amériques ayant de nombreux stocks rassemblant chacun peu de langues.

Cela suggère donc que le continent américain est le plus récent à avoir été colonisé, mais ne permet en aucun cas de définir une date pour cette colonisation. La tendance observée est compatible avec un modèle proposé par Dixon [171]. Dixon émet l'hypothèse que la ramification linguistique intervient lors de courtes et infréquentes périodes de temps de bouleversements rapides de populations, qu'il nomme ponctuations.

L'un des exemples avancés est celui du développement de l'agriculture en Eurasie qui provoqua, en augmentant grandement le taux de croissance de la population, le dispersement sur le continent de nombreuses populations en croissance rapide [172]. Comme les populations se sont séparées et fragmentées, les langues de ces populations ont commencé à diverger et se sont séparées, donnant l'arborescence observée clairement dans la famille linguistique de l'Indo-Européen [173, 174]. Entre les ponctuations se trouvent de longues périodes pendant lesquelles toutes les niches écologiques du continent sont saturées, et aucune communauté n'a d'avantage démographique ou technologique suffisant sur ses voisins pour les déplacer. Pendant de telles périodes d'équilibre l'apparition de nouvelles lignées est rare.

La ponctuation la plus significative qui ait pu arriver sur le continent américain pourrait bien être l'arrivée des humains dans un habitat entièrement nouveau. Dans un environnement si peu habité, la croissance de la population serait rapide et les groupes de chasseurs-cueilleurs se répandraient et se diviseraient avec une grande fréquence alors qu'ils progresseraient à travers le continent [175]. Chacune de ces divisions serait associée avec le fondement d'une nouvelle lignée linguistique. La croissance de la population diminuerait comme la pression de population et la compétition entre les groupes pousseraient l'évolution culturelle vers une extraction plus intense des ressources naturelles [176] et l'établissement de sociétés plus grosses et plus complexes [177]. Dixon insiste sur le fait que l'extinction est une force d'évolution aussi significative que la ramification. Les lignées et les langues peuvent disparaître si les communautés les parlant sont victimes de catastrophes naturelles, se dispersent vers d'autres groupes ou sont absorbées par le développement de leurs voisins. De telles extinctions ne sont pas rares dans les registres ethnographiques.

B) Etude de l'ADN mitochondrial

1. Définitions et nomenclature

Tout d'abord, un haplotype consiste en la liaison de polymorphismes. Comme l'ADN mitochondrial ne subit aucune recombinaison et qu'il n'est transmis que par la mère [178-180], tous les polymorphismes trouvés dans l'ADN mitochondrial d'un individu peuvent être groupés comme faisant partie d'un (même) haplotype. Cela peut être une portion de séquence continue de l'ADN mitochondrial ou une collection de polymorphismes de restriction (RFLPs) provenant de la séquence entière, comportant 16569 nucléotides, de l'ADN circulaire de la mitochondrie [181].

Un haplogroupe est une collection d'haplotypes fortement apparentés qui partagent tous un ou plusieurs polymorphismes, en comparaison d'autres haplogroupes. C'est en 1985 que fut montré que les Amérindiens avaient un sous-échantillonnage de polymorphismes trouvés en Asie, mais à des fréquences drastiquement différentes, suggérant un important effet fondateur entre l'Asie et l'Amérique [182]. Par la suite, il fut montré que 97% de la variation trouvée dans trois populations d'Amérique du Nord, Centrale et du Sud était expliquée par quatre haplogroupes [33] qui furent nommés A, B, C et D [183]. Ces haplogroupes étaient initialement définis par des RFLPs et une délétion de neuf paires de bases.

Plus spécifiquement, l'haplogroupe A était défini par le gain d'un site de restriction *Hae III* en position 663, qui n'est présent chez aucun autre groupe amérindien. L'haplogroupe C se définissait par la perte d'un site de restriction *Hinc II* en position 13259, et l'haplogroupe D par la perte d'un site de restriction *Alu I* en position 5176. L'haplogroupe B était défini par une délétion de neuf paires de bases dans la région V du génome mitochondrial, une courte région non-codante entre les gènes codant pour la cytochrome oxydase II et l'ARNt pour la lysine. La plupart des humains ont deux copies de cette séquence de neuf paires de bases, contrairement aux individus de l'haplogroupe B qui n'en ont qu'une.

Les haplogroupes A et B, ainsi que l'haplogroupe X de Brown, perdent un site de restriction *Alu I* en position 10397 et *Dde I* en position 10394, alors que les haplogroupes C et D, ainsi que les haplogroupes X6 et X7, gardent ces sites. La perte du site *Alu I* en position 10397 est très caractéristique de l'Asie, et n'est trouvée que chez des populations ou des individus de descendance asiatique, y compris les Insulaires du Pacifique et les Amérindiens. La perte du site de restriction *Hinc II* en position 13259 (haplogroupe C) est toujours accompagnée par le gain des sites *Dde I* en position 10394 et *Alu I* en position 10397 dans le Nouveau Monde et en Asie. Alors que certaines populations non-asiatiques se présentent avec une perte du site de restriction *Hinc II* en position 13259, elles n'ont pas ces gains en position 10394 et 10397.

La délétion de neuf paires de bases semble avoir eu lieu plusieurs fois, y compris une fois, et peut-être plus, en Afrique [184-188], et au moins deux fois en Asie [188-191]. Quoi qu'il en soit, dans les populations descendant d'Asie, 99% des individus avec la délétion dérivent d'un événement de délétion unique, il y a probablement 30 à 60 000 ans [192]. En plus des marqueurs RFLPs, des mutations ponctuelles dans la région contrôle de l'ADN mitochondrial, la région non-codante située entre les nucléotides 15975 et 430, peuvent également être utilisées pour définir les haplogroupes A, B, C, D et X [190, 193-196].

2. Voies de migration

La détermination des voies de migration en utilisant les données génétiques est probablement ce qu'il y a de plus difficile à faire, à moins que de l'ADN ancien recouvré le long des voies supposément utilisées puisse fournir des informations suffisamment solides pour pouvoir argumenter contre ou en faveur d'un scénario particulier. Par contre, on peut regarder si les données sont incompatibles avec certaines théories

De façon assez certaine, 99% de tous les haplotypes amérindiens prennent leur origine en Asie. Cinq à sept haplotypes, qui appartiennent à autant d'haplogroupes (A, B, C, D, X6 et 7, X de Brown), ont été plus ou moins reconnus comme étant partagés par l'Asie et l'Amérique. Si l'acceptation de ce fait pour les haplogroupes A, B, C, D et X de Brown semble maintenant être générale, ce n'est pas le cas pour les haplogroupes X6 et X7 qui sont fortement apparentés aux haplogroupes C et D [197, 198].

Environ 98% des échantillons amérindiens génotypés sont A, B, C ou D, comme précédemment définis [183]. Les 2% restant semblent être X6, X7 [199, 200], ou X de Brown [201]. Donc, la très grande majorité des Amérindiens actuels sont une variante de A, B, C ou D, et cela indépendamment de l'origine tribale, de la langue parlée et de la localisation géographique, voir figure 2 page 30.

Les données de l'ADN ancien sont bien plus rares, mais presque tous les échantillons complètement génotypés pour les marqueurs des lignées mitochondriales fondatrices se sont avérés être A, B, C, D ou un des X. Jusqu'à aujourd'hui, aucun reste humain ancien associé à une route maritime, qu'elle soit côtière ou non, n'est disponible. Il est fort possible que de tels sites, s'ils existent, soient maintenant situés sous les eaux à cause de la montée importante du niveau de la mer depuis la dernière ère glaciaire. Cela est particulièrement vrai pour la route transatlantique proposée par Stanford et Bradley [104].

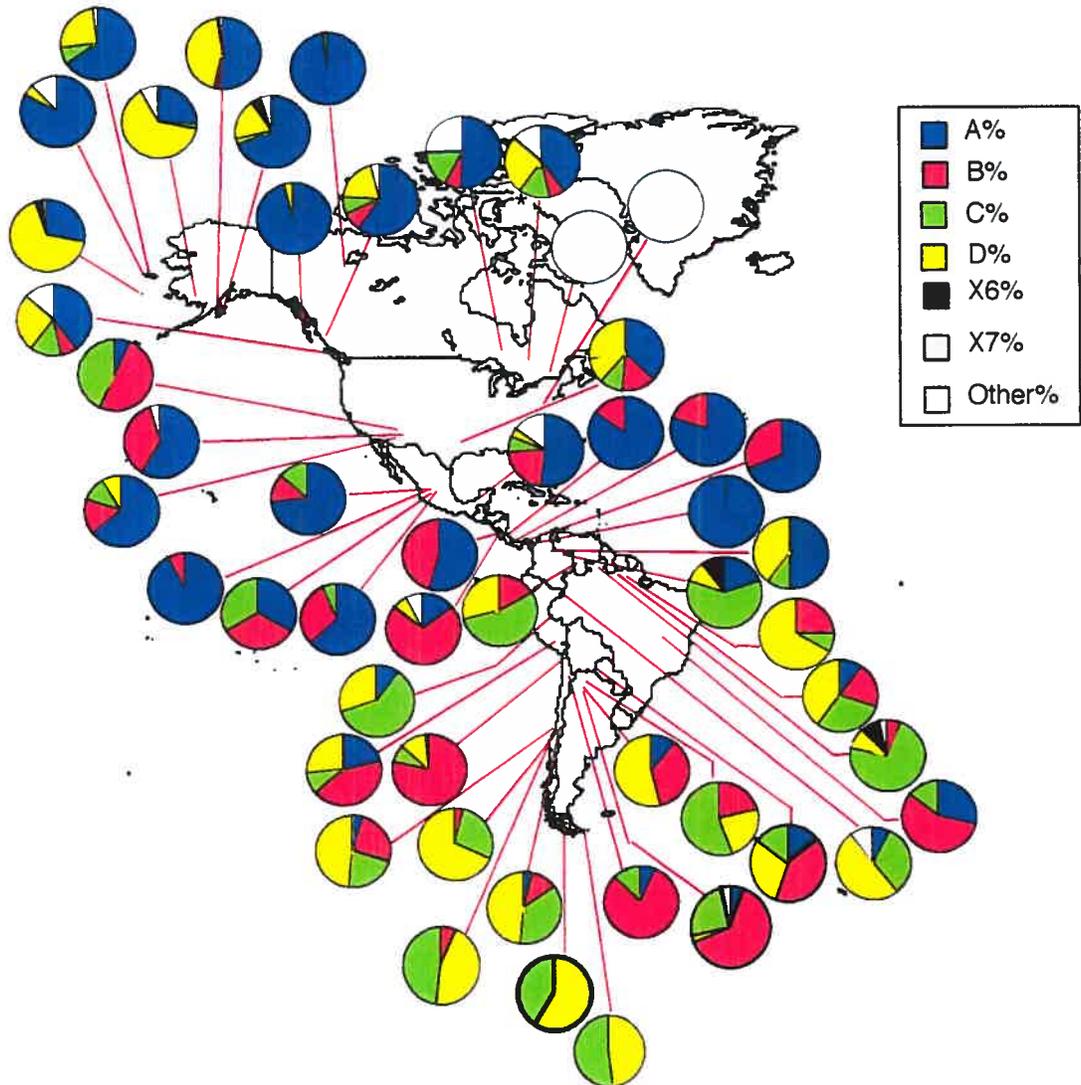


Figure 2 : Fréquence des haplotypes fondateurs en Amérique [201].

3. Lignées fondatrices

En 1985 la plupart des lignées fondatrices des haplogroupes du Nouveau Monde avaient été découvertes [182]. Elles furent identifiées plus précisément par la suite et 95 à 98% des Amérindiens présentaient l'un de ces quatre types [33, 183]. Les autres 2 à 5% étaient supposés être dus au métissage avec les Européens (dans les populations génotypées à l'époque cela devait être en grande partie vrai). Quoi qu'il en soit, de nouveaux haplotypes, et possiblement un nouvel haplogroupe, communs au Nouveau Monde et à l'Asie furent identifiés en 1994 [202] et confirmés en 1996 [199, 200]. Ils furent nommés A1, A2, B1, C1, C2, D1, D2, X6 et X7 [200]. En 1998 un nouvel haplogroupe fondateur a été identifié par Brown [201], malencontreusement également appelé X (d'où la référence à cet haplogroupe comme X de Brown), non-apparenté aux haplotypes X6 et X7 précédemment décrits [199].

A l'exception du X de Brown, tous ces haplotypes sont également retrouvés en Asie, parfois même tous dans une même population (la Mongolie [203], le Tibet [204]), mais la plupart sont absents d'Europe, d'Afrique et d'Australie. Cet argument a été utilisé en faveur d'une seule vague de migration vers le Nouveau Monde [36, 193, 198, 203, 205-207].

Il faut noter que le site *Hae III* 16517 est hypervariable, ce qui diminue son utilité comme marqueur phylogénétique. Toutefois, pour de courtes échelles de temps, cette hypervariabilité peut être un atout intéressant lorsque comparé au « fond » des variations de l'ADN mitochondrial. Cela signifie que les sous-types, qui sont uniquement définis par ce marqueur (soit A1 par rapport à A2, X6 par rapport à X7 etc.), diffèrent par une mutation qui aurait très bien pu apparaître dans le Nouveau Monde plutôt qu'avant l'entrée sur ce continent, malgré le fait que cette mutation soit retrouvée des deux côtés du détroit de Béring. Cependant, il y a des définitions des haplotypes A1 et A2, et D1 et D2, par la région contrôle, et ces haplotypes sont également trouvés sur les deux continents [197], ce qui supporte également fortement

la notion que plus d'un variant des haplogroupes A et D sont parvenus sur le continent américain avec la vague initiale de migration [192, 197].

Une lignée fondatrice doit satisfaire quatre critères pour être effectivement considérée comme fondatrice. Tout d'abord la lignée doit être partagée entre le Nouveau Monde et la population parentale, l'haplotype devrait être central (nodal) dans la reconstruction phylogénétique amérindienne et devrait être répandu sur le continent. De plus, il devrait être trouvé dans l'ADN ancien datant d'avant le premier contact avec les Européens.

Si l'on applique ces critères, les haplotypes A, B, C, D, X6, X7 et X de Brown satisfont certaines ou toutes les conditions requises. Tous sont trouvés dans de l'ADN datant d'avant le premier contact avec les Européens [205, 208, 209]. Les variants de A, B, C et D sont clairement au centre des reconstructions phylogénétiques et sont partagés entre l'Amérique et l'Asie.

Les lignées X6 et X7 ont été décrites comme dérivant des haplogroupes C et D. X6 et X7 se retrouvent sur de grandes étendues en Asie et en Amérique [199, 200], et à travers d'un grand nombre de cultures linguistiques différentes, ce qui laisse penser qu'elles sont entrées sur le continent américain comme lignées fondatrices. Leur faible fréquence indique que les événements mutationnels menant de C ou D à X6 ou X7 sont assez peu fréquents. Le séquençage complet des lignées X6 et X7 amérindiennes est en cours afin de déterminer si elles sont identiques à celles retrouvées en l'Asie [192, 197].

L'haplogroupe X de Brown est trouvé à travers l'Amérique du Nord (et de nombreux groupements linguistiques) [201, 210], mais n'a pas été encore confirmé en Amérique Centrale et du Sud. Il semble également qu'il soit présent dans l'ADN ancien [209]. Il n'a pas été positivement identifié en Asie (une différence de six mutations existe entre les haplogroupes asiatiques classifiés comme X de Brown et

ceux du Nouveau Monde), mais partage quatre mutations avec un ancien haplogroupe européen.

Les haplogroupes X de Brown européens et amérindiens se seraient donc séparés il y a une grande période de temps. L'absence d'identification sûre de cet haplogroupe dans les populations situées entre l'Europe et l'Amérique, ainsi que des ressemblances archéologiques entre la civilisation Clovis et une ancienne civilisation ibérique, la civilisation des Solutréens, mena vers la théorie d'une connexion européenne entre l'Europe et le Nouveau Monde, à travers l'océan Atlantique. Cette théorie est grandement spéculative, et il est tout à fait possible que l'haplogroupe X de Brown se soit éteint en Asie dans la période de temps considérée. Il existe deux haplotypes partiels trouvés en Asie montrant possiblement une relation avec l'haplogroupe X de Brown, en Corée [211, 212] et en Mongolie [193], identiques à des haplotypes Nootka dans le Nouveau Monde [196]. Cela montre en fait clairement que des études plus approfondies des populations asiatiques sont nécessaires.

4. Datation des migrations

L'idée d'utiliser une horloge moléculaire est problématique pour une période temporelle aussi courte, parce que les hypothèses sur lesquelles cette horloge repose demandent que les mutations se produisent au hasard, sans pression de sélection (sauf celles sélectionnant contre les mutations très nuisibles pour la santé), et s'accumulent à la façon dont le temps s'écoule pour une horloge, et cela dans toutes les lignées.

L'horloge moléculaire est moins précise sur les courtes périodes de temps, parce que les mutations n'arrivent pas réellement à des intervalles de temps constants. Les changements biochimiques qui créent ces mutations sont causés par des erreurs de réplication, des dommages causés par l'environnement et/ou une mauvaise réparation de ces dommages. Sur de longues périodes de temps, le nombre moyen de mutations par nucléotide et par an tend à se stabiliser, si tant est que les autres facteurs demeurent constants.

Une population en croissance rapide dans une nouvelle niche écologique pourrait permettre plus facilement l'introduction de nouvelles mutations dans la population. Les nouvelles populations étant petites, toute nouvelle mutation s'y trouve automatiquement à plus haute fréquence que si elle était advenue dans une plus grande population (dans une population de 10 individus, une nouvelle mutation a une fréquence de 10%, alors qu'elle n'aurait une fréquence que de 0.1% dans une population comptant 1000 individus). Si tous les individus ont un grand nombre de descendants, la place disponible n'étant pas un facteur limitant, au moins dans les premiers temps de la colonisation, une nouvelle mutation a des chances de voir sa fréquence augmenter rapidement.

Comme pour l'archéologie, il est toujours possible de trouver de nouvelles séquences, différentes de celles précédemment découvertes, qui n'avaient pas été détectées à cause d'un échantillonnage inadéquate. La population amérindienne est aujourd'hui estimée à plus de 2,5 millions d'individus en Amérique du Nord [213] et probablement plus de 10 millions en combinant l'Amérique du Nord, Centrale et du Sud. Au moment du premier contact entre les Amérindiens et les Européens, la population du Nouveau Monde est estimée se situer entre 43 et 65 millions d'individus [214]. Au 16^e siècle il semble que la population amérindienne ait diminué de 89% [215]. Les échantillons d'ADN mitochondrial génotypés aujourd'hui représentent donc seulement quelques dixièmes de pourcents de la population amérindienne totale actuelle, et moins d'un centième de pourcent de la population amérindienne disparue. De plus, les échantillons ne correspondent qu'à un nombre très restreint des langues parlées sur ce continent, même si environ un tiers des groupements linguistiques majeurs sont couverts. Les données manquantes pourraient complètement changer l'état actuel des connaissances, en modifiant la fréquences des haplotypes et la diversité génétique, et par conséquent en modifiant les relations entre la population amérindienne et les autres populations, y compris la datation de la divergence entre ces populations.

Une fois les mutations identifiées et le taux de mutation que l'on veut appliquer aux calculs, il faut encore déterminer quelles mutations sont intervenues dans la population parentale avant la migration, et lesquelles se sont produites dans la population dérivée après la migration. Dans le meilleur des cas il s'agit des mutations trouvées en Asie (Sibérie, Mongolie) et dans le Nouveau Monde.

La plupart des estimations publiées sur l'âge des lignées mitochondriales dans le Nouveau Monde ne sont pas basées sur des mutations partagées par le continent Asiatique et Américain, mais plutôt sur des mutations particulières aux Amérindiens, menant à une surestimation de l'époque d'un ancêtre commun pour les lignées du Nouveau Monde.

En tenant compte de toutes ces considérations, la moyenne des âges d'un ancêtre commun, pour chaque haplogroupe du Nouveau Monde, a été estimé à 25 000 ans, bien qu'il fut suggéré qu'il y ait eu plusieurs vagues de migration pour ce scénario, et que l'haplogroupe B apparaissait plus jeunes que les haplogroupes A, C et D [189, 190, 204, 216-218]. Il a été montré, par la suite que les temps de divergence pour les haplogroupes A, B, C et D n'étaient pas statistiquement différents les uns des autres [219]. La date de 25 000 ans est celle d'un ancêtre commun en Asie, mais pas obligatoirement celle de l'entrée dans le Nouveau Monde, qui pourrait être considérablement plus récente.

5. Origine des Amérindiens

Il y a un grand nombre d'hypothèses concernant le berceau des civilisations amérindiennes, dont la plus communément acceptée, et la moins précise, est que ce berceau est situé en Asie. L'étude des peuples de Sibérie et d'une partie de ceux de l'Asie du Nord et du Sud-Est ont mis à jour certaines possibilités quant à l'origine des fondateurs des Amérindiens. La Mongolie, la Manchourie et le sud de la Sibérie furent suggérées en 1994 [220], la Mongolie ayant été également suggérée par des études antérieures sur l'ADN mitochondrial [203]. Des études sur le chromosome Y indiquèrent la région du lac Baïkal dans le sud de la Sibérie [221, 222], comme certaines analyses de l'ADN mitochondrial [223, 224]. Il fut aussi suggéré que l'haplogroupe B arriva sur le continent américain plus récemment que A, C ou D et que sa répartition plutôt côtière (côte Ouest des Amériques et côte Est de l'Asie) indiquait une voie de migration maritime le long des côtes [189, 190, 204, 216, 217], en accord avec certaines observations linguistiques [165, 225].

Selon l'approche utilisée pour l'analyse des données, l'interprétation des résultats obtenus peut grandement varier, et mener à différentes conclusions. Une approche parcimonieuse serait de considérer le nombre minimal de migrations nécessaires à la réconciliation avec les données observées. Par contre, une approche basée sur la fréquence considèrera que chaque lignée provient de l'emplacement auquel elle a actuellement la plus haute fréquence. Ces deux approches reposent aussi sur la proximité de la source et du Nouveau Monde afin de départager les différentes populations qui pourraient se trouver en compétition, la population idéale étant située le plus près possible du détroit de Béring.

C) Etude du chromosome Y

1. Nomenclature

Malheureusement, aucune nomenclature n'a été encore universellement acceptée et utilisée pour la classification des haplotypes du chromosome Y et deux systèmes de notation semblent coexister, compliquant les comparaisons entre les différentes études.

Les premières recherches systématiques de polymorphismes de restriction (RFLPs) sur le chromosome Y suggèrent une grande pauvreté de ce chromosome en terme de polymorphismes [226, 227]. L'utilisation du chromosome Y dans les études des populations gagna de l'intérêt après la découverte de plusieurs polymorphismes informatifs. Ces polymorphismes sont de plusieurs natures : des substitutions de bases [228, 229], des duplications [230], l'insertion d'une séquence Alu [231], des changements dans le nombre de répétitions dans des microsatellites [232] ou dans la région centromérique du bloc alphas [231].

Le premier système de classification employé pour les Amérindiens est celui utilisé en 1995 par Pena et al [27], tel que précédemment décrit par Santos et al. [232]. Ce système se base sur l'analyse de deux polymorphismes : le microsatellite DYS19 (Y-27H39) et la variation de séquence dans les répétitions alpha de la région centromérique DYZ3 ($\alpha(h)$). DYS19 permet d'isoler cinq allèles, classés de A à E d'après le nombre de répétitions [232], et $\alpha(h)$ 23 allèles numérotés de I à XXIII. La combinaison de ces deux polymorphismes résulta en 37 haplotypes différents dans le monde, répertoriés en combinant la classification de DYS19 et $\alpha(h)$, soit IA, IIA et ainsi de suite. Par la suite, neuf allèles ont été identifiés pour DYS19, donnant plus de 45 haplotypes lorsque combinés avec $\alpha(h)$ [32, 233].

Le second système utilisé est celui décrit par Underhill et al. en 1997 [234], continuant ses investigations à l'aide de la technique DHPLC [235], alors nouvelle, pour trouver de nouveaux sites polymorphiques bi-alléliques sur le chromosome Y. Aux trois sites précédemment décrits (DYS287, DYS271 et DYS199) vinrent s'ajouter 19 nouveaux sites, et les sites furent classés de M1 à M22 et trois haplogroupes furent définis et classés A (défini par la séquence ancestrale, trouvée pour 150 chromosomes humains sur les 718 analysés), B (défini par M1, une insertion *Alu* [236] parfois décrite par l'abréviation YAP) et C (défini par le marqueur M9, une transversion non-africaine C vers G). Ce système gagna encore grandement en puissance en 2000 après l'ajout par Underhill [237] de 86 marqueurs bi-alléliques aux 80 déjà identifiés à cette époque, et un marqueur tri-allélique [234, 238]. Cela définit 116 haplotypes, classés de M1 à M116, et les trois haplogroupes précédemment définis firent place à dix nouveaux haplogroupes (notés de I à X) d'après un arbre basé sur le principe de parcimonie maximale, dont la racine a été déterminée à partir de la séquence trouvée chez 6 primates appartenant à trois espèces différentes [235, 237], voir tableau III page 38 et figure 2b page 39.

Haplogroupe	Mutation caractéristique la plus récente	Nombre d'haplotypes par haplogroupe
I	M91	8
II	M60	10
III	M96	21
IV	M174	4
V	M130	5
VI	M89 et absence de M9	23
VII	M175	15
VIII	M9 et absence de M175 et M45	11
IX	M173	13
X	M74 et absence de M173	6
Total	NA	116

Tableau III : Définition des haplogroupes du chromosome Y [237].

2. Lignées fondatrices, origine des Amérindiens et vagues de migration

Le système basé sur DYS19 et $\alpha(h)$ permet d'identifier 37 haplotypes répartis dans 12 populations amérindiennes à travers différents groupes linguistiques reliés à l'Amerind [27]. L'haplotype IIA (combinaison de $\alpha(h)II$ et DYS19 allèle A) fut retrouvé chez 74% des individus testés, ou 91% en excluant les Mapuches, qui présentent le plus haut niveau estimé de métissage avec les Européens (12%). En se basant sur la très haute fréquence trouvée, cet haplotype a été défini comme étant l'haplotype ancestral majeur chez les Amérindiens, même si l'haplotype le plus fréquent n'est pas obligatoirement l'haplotype ancestral. La plupart des études qui ont suivi, et qui ont utilisé ces deux marqueurs, entre autres, pour la reconstruction des haplotypes, sont en accord avec la conclusion concernant une seule vague de migration [32, 239, 240].

Le marqueur DYS199 semble être polymorphique presque exclusivement chez les Amérindiens (allèle T), et trouvé à très haute fréquence (entre 18 et 90%) [221, 234, 235, 239, 241-243], en accord avec la théorie d'une seule vague de migration. Des études plus récentes sur un différent nombre de marqueurs, dont DYS199, ont permis d'identifier des populations asiatiques polymorphiques pour ce marqueur [221, 222, 240, 241, 243]. Les populations présentant l'allèle T pour le marqueur DYS199 sont donc potentiellement la source de l'allèle T amérindien. Cet allèle est retrouvé chez les Evens, les Chukchi et les Eskimos de Sibérie, ces trois populations se trouvant en Sibérie [221, 240, 241, 243]. Toutefois, la co-étude de DYS199 et d'autres marqueurs amena certains auteurs à identifier deux lignées fondatrices des haplotypes amérindiens [222, 240, 242, 243], ajoutant la région générale du lac Baïkal et la Mongolie comme autre source possible des haplotypes du chromosome Y retrouvés chez les amérindiens.

Cette observation amena Lell à conclure à deux vagues de migration différentes de l'Asie vers l'Amérique, en provenance de Sibérie [243]. Toutefois cette interprétation a été critiquée, sur la base d'un manque de parcimonie [35], une seule vague de migration suffisant à expliquer les données observées : l'un des deux haplotypes utilisés par Lell pour conclure à deux vagues de migration est l'un des plus fréquents en Europe [244], et pourrait donc grandement dû au métissage des Amérindiens avec les Européens, et que l'autre haplotype est trouvé dans des populations avec l'allèle T de DYS199, en Sibérie.

La grande différence de diversité du chromosome Y par rapport à celle reportée par les études sur l'ADN mitochondrial a été expliquée par une différence de dynamique entre les lignées maternelles et paternelles. Les activités spécifiques aux hommes, principalement la chasse et la guerre, étant sans doute une cause de plus grande mortalité par rapport à la mortalité des femmes, conduisirent probablement à un nombre effectif de chromosomes inférieur chez les hommes (par rapport aux femmes). De plus, les chefs de tribus aurait pu participer de façon excessive au nombre de chromosomes Y transmis, la polygynie étant souvent pratiquée.

3. Datation des migrations

Les migrations ont été datées par les différents auteurs, principalement à partir du marqueur DYS199, par des études de coalescence, en utilisant un taux de mutation arbitraire et une durée moyenne de génération de 27 ans [235], et que les allèles DYS199 C et T étaient tous deux arrivés sur le continent américain lors de la vague de migration. Des dates très différentes ont été suggérées : entre 2147 et 30 000 ans selon le taux de mutation [235], 22 770 ans (minimum 13 500 ans, maximum 58 700 ans) [240], 7600 ans plus ou moins 5000 ans [222] et entre 9000 et 11 000 ans [242]. Bien entendu il s'agit ici de l'âge de la mutation DYS199 T, qui n'est pas obligatoirement celui de la migration ayant fondé les populations amérindiennes.

D) Etude des autosomes, co-évolution de micro-organismes

1- Système HLA

Le complexe génétique HLA est la région codante la plus polymorphique du génome humain [245], et sa distribution reflète des affiliations à des régions géographiques ou des groupes ethniques [246].

Une étude détermina la distribution de la fréquence allélique des gènes polymorphiques HLA de classe I (A et B) et de classe II (DRB, DQA1 et DBQ1) par analyse génétique et sérologique chez 89 Amérindiens de la tribu des Mazatecs [247]. Les données furent comparées à celles de 6050 individus répartis sur 29 groupes ethniques différents, sur les cinq continents. Les analyses phylogénétiques supportent le fait que les Amérindiens d'Amérique Centrale et d'Amérique du Sud ont des ancêtres communs originaires du Nord de la Sibérie. En outre, à part ces contributions, les Amérindiens d'Amérique Centrale ne démontrent pas d'affinité avec les autres groupes, contrairement aux Africains, Européens, Australiens, Polynésiens Eskimos et autres.

Dans une autre étude, huit populations sibériennes furent étudiées pour trois des loci HLA de classe II les plus polymorphiques, soit DRB, DQA1 et DQB1 [248]. Le nombre d'allèles par population variait de 16 à 25, sept à huit et neuf à 14 pour les loci DRB1, DQA1 et DQB1 respectivement, définissant 66 haplotypes différents. La plupart de ces haplotypes apparaissaient d'origine ancienne, bien que quelques-uns aient pu apparaître dans la population sibérienne. Les analyses phylogénétiques ont montré que les haplotypes asiatiques et amérindiens se trouvaient regroupés et isolés des haplotypes européens et africains, suggérant un ancêtre commun. Les asiatiques se divisèrent en deux groupes, le premier réunissant la presque totalité des haplotypes sibériens et amérindiens, le second réunissant les populations d'Asie Centrale, de l'Est et du Sud-Est. Le temps de divergence de ces deux groupes fut situé entre 21 000 et 24 000 ans. Trois branches sibériennes différentes semblent avoir contribué

au fondement des populations amérindiennes. Les données supportent la théorie de trois vagues de migrations, la première en provenance de l'est de la Sibérie il y a moins de 21 000 ans.

Une étude des séquences de l'exon 2 de HLA-DRB1 (et de DQB1, mais sans obtention de matériel à analyser) de l'ADN ancien provenant de l'Homme des Glaces du Tyrol détermina que l'allèle porté par cet homme était d'un type très rare, HLA-DRB1*1402. Aucune personne impliquée dans l'étude de ces échantillons ne présentait cet haplotype, rare chez les Européens, excluant ainsi la possibilité de contamination des échantillons par du matériel génétique extérieur (la contamination de matériel génétique ancien provenant généralement des personnes manipulant les échantillons).

2- Groupes sanguins

Trois études récentes sur différents marqueurs du groupe sanguin, de 11 à 16 marqueurs (dont ABO, RH et MNSs), ont été rapportées [249-251] pour différentes populations amérindiennes.

La première étude s'intéressait à une population du Centre des Andes, au Pérou, vivant à 3800 mètres d'altitudes, les Tayacaja [250]. Seize systèmes génétiques furent utilisés pour génotyper 141 individus. Aucune déviation de l'équilibre d'Hardy-Weinberg ne fut constatée. Les fréquences des allèles trouvés dans cette population étant comparables à celles trouvées dans des populations de basses altitudes, aucune preuve d'adaptation à la haute altitude n'a pu être mise en évidence. Les résultats pour cette population pour 12 des 16 marqueurs furent utilisés afin de déterminer la variabilité génétique en Amérique du Sud, en ajoutant les données des Tayacajas à d'autres précédemment obtenues pour ces 12 marqueurs dans des populations amérindiennes. Les tests d'autocorrelation spatiale n'ont indiqué aucun motif typique en génétique des populations [252-255]. Un marqueur, AcP1*a, a montré un net gradient géographique. Des tests subséquents suggèrent un peuplement indépendant

de l'Est et de l'Ouest de l'Amérique du Sud ou un faible niveau d'échange génétique entre ces deux régions.

L'étude des Ayoreos et des Lenguas de la région de Chaco, au Paraguay, et leur comparaison avec d'autres tribus de la région et une tribu externe montra une homogénéité des tribus de cette région (Lenguas, Tobas, Mataco, Choroti) mais avec une séparation claire des Ayoreos [249] des autres groupes de la région. La tribu externe à la région de Chaco, les Aches, se montrent très différents de toutes les autres tribus.

La dernière étude s'intéressait au métissage des Amérindiens avec les Européens et les Africains dans deux communautés brésiliennes, à Monte Negro et Portuchuelo [249-251], en génotypant 11 marqueurs différents. Les résultats de la composition ethnique sont présentés dans le tableau IV ci-après.

Population	Africain	Amérindien	Caucasoïde
Portuchuelo	0.21 ± 0.046	0.44 ± 0.064	0.35 ± 0.069
Monte Negro	0.25 ± 0.032	0.12 ± 0.046	0.63 ± 0.054

Tableau IV : Composition ethnique de deux populations brésiliennes

La participation presque quatre fois supérieure des Amérindiens dans le pool génétique de Portuchuelo en comparaison de celui de Monte Negro est en accord avec l'histoire de ces régions. On remarque que la contribution africaine est équivalente dans les deux régions et assez significative. Les populations brésiliennes rurales et urbaine d'Amazonie résultent d'une composition ethnique variée [251, 256].

Il existe d'autres études sur les populations amérindiennes, utilisant des marqueurs autosomaux. Cependant, ces études focalisent généralement sur une tribu ou ethnie très précise, dans un contexte médical, et sont peu informatives concernant la génétique des populations et le peuplement des Amériques [18, 257-265].

3- Co-évolution de micro-organismes

Lorsque un pathogène est introduit dans un nouvel environnement, et que ce pathogène connaît un succès face à ses hôtes, son développement est très rapide, réduisant sa diversité. Cela peut faire des pathogènes humains une source intéressante d'informations concernant les mouvements des populations. Deux études récentes de co-évolution de micro-organismes avec les humains ont été faites dans le contexte des populations amérindiennes.

La première étude s'intéresse à champignon pathogène, *Coccidioides immitis*, et à son expansion en Amérique du Sud [266]. Les analyses de diversité des marqueurs microsattellites montre une partition de la diversité génétique en Amérique du Nord, mais une grande homogénéité en Amérique du Sud. Les souches trouvées en Amérique du Sud sont très proches, et parfois identiques, à des souches trouvées au Texas. La variance en Amérique du Sud montre qu'il y a eu une augmentation rapide de la population de ce micro-organisme. Les auteurs concluent que le micro-organisme a été apporté en Amérique du Sud lors de la colonisation de cette partie du continent, par un mouvement de population en provenance d'Amérique du Nord. La très faible variabilité est en accord avec une colonisation assez récente de l'Amérique du Sud, mais la datation effectuée donne un intervalle de temps compris entre 9000 ans et 140 000 ans, ne permettant pas d'affiner les estimations données par d'autres systèmes génétiques.

Une étude sur le polyomavirus humain JC (JCV) a classé les types de JCV trouvés chez les Eskimos de Sibérie et les Inuits Canadiens [267]. Ce virus avait déjà été classé en dix types géographiquement distincts. Tout d'abord, le JCV ancestral se divisa en trois superfamilles notées A, B et C. Le type A a généré deux sous-types, Eu-a et Eu-b, regroupant principalement des groupes européens et méditerranéens. Le type B se sépara en Af2 (le type africain majeur), Bl-c (un sous-type européen mineur) et différents sous-types asiatiques. Enfin, le type C généra un unique sous-type, Af1, regroupant des communautés d'Afrique de l'Ouest.

Les JCVs trouvés dans les populations sibériennes et arctiques furent classés comme appartenant principalement au type A. En refaisant l'analyse du type A pour les populations mondiales, après l'ajout des résultats sur les populations arctiques et sibériennes, le type A se divisa en trois sous-types, Eu-a, Eu-b et Eu-c. Le sous-type Eu-c était exclusivement composé d'isolats du Nord-Est de la Sibérie, principalement des Nanais, vivant dans la basse Région de la Rivière Amur. La plupart des sibériens et des peuples de la région arctique étaient des Chukchis, des Koryaks et des Inuits du Canada, et se retrouvaient groupés dans un sous-type de Eu-a.

La séparation du type A en Eu-c et non Eu-c, puis de non Eu-c en Eu-a et Eu-b a été estimée s'être produite il y a entre 30 000 et 40 000 ans [267, 268]. Pour la séparation du sous-type Eu-a la date des différentes séparations a été évaluée entre 10 000 et 20 000 ans. Comme JCV est d'habitude transmis de parents à enfants à cause d'une longue cohabitation et est rarement transmis entre les populations humaines [269, 270], il est probable que cela ait produit une forte association entre les génotypes de JCV et les populations humaines. Cela suggère que les Européens ont contribué à la population ancestrale des Sibériens du Nord-Est, et par conséquent à celle des Eskimos de l'Amérique du Nord.

E) Projet de recherche

Comme Francisco M. Salzano le faisait remarquer dans sa récente revue de littérature [271], bien que beaucoup d'informations soient disponibles sur la variabilité dans les populations amérindiennes, ces informations sont hautement hétérogènes par rapport aux populations et aux types de systèmes génétiques étudiés. Seulement neuf populations amérindiennes ont été étudiées au niveau du chromosome X [257, 272-276], et neuf systèmes différents ont été utilisés (en comparaison, 50 populations ont été étudiées au niveau du chromosome Y, et 66 systèmes différents ont été utilisés).

S'il est vrai que les recombinaisons peuvent compliquer l'interprétation des données obtenues, principalement au niveau temporel, le chromosome X présente des caractéristiques très intéressantes pour l'étude des populations, et peut grandement compléter les systèmes génétiques localisés sur d'autres chromosomes. De plus, les recombinaisons peuvent présenter certains avantages quant à la détection des métissages.

Tout d'abord, l'utilisation d'individus de sexe masculin permet une lecture directe des haplotypes, limitant ainsi le débat sur les problèmes de reconstruction haplotypiques. Le chromosome X étant transmis par la mère et le père, le biais sexuel rencontré pour l'ADN mitochondrial et le chromosome Y, respectivement seulement transmis par la mère et par le père, disparaît. Ce mode de transmission permet aussi de limiter les effets de la dérive génétique en portant la population effective relative du chromosome X à $\frac{3}{4}$, par rapport aux autosomes, contre $\frac{1}{4}$ pour l'ADN mitochondrial et le chromosome Y. Ainsi, il est possible d'espérer une reconstruction temporelle plus profonde, ainsi qu'une meilleure étude des populations isolées.

Le système *Dys44*, composé de polymorphismes dans une région intronique du gène de la dystrophine, en Xp21, a déjà été étudié de manière exhaustive dans de nombreuses populations à travers le monde [273, 274, 277-279], mais relativement peu dans les populations amérindiennes. En ajoutant des données sur les populations amérindiennes on inscrit donc ces populations dans le cadre d'une étude beaucoup plus large. Cela nous permettra de nombreuses comparaisons entre les populations amérindiennes, mais également entre ces dernières et les populations asiatiques et européennes. En réduisant fortement l'hétérogénéité des données pour les populations amérindiennes, au moins concernant le chromosome X, le système *Dys 44* devrait s'avérer d'une grande utilité et compléter de manière significative les autres systèmes génétiques et les données d'autres domaines de recherche.

Chapitre II

MATERIEL ET METHODES

A) Populations étudiées

La liste complète des populations considérées dans le cadre de ce travail, ainsi que le nombre de chromosomes X étudiés pour chacune d'elles, peut être trouvée dans le tableau V page 50. L'analyse des populations citées dans la partie supérieure du tableau sont le fruit de mon propre travail, les autres ayant précédemment fait l'objet d'études dans notre laboratoire [273]. Il est à noter que pour certaines populations (Chipewyan, Ojibwa, Mexicaine, Ibérienne et Chinoise), les échantillons que j'ai génotypés se sont ajoutés à d'autres des mêmes populations, dont le génotypage avait été effectué par le Dr Yotova dans le cadre des études citées plus haut.

Les échantillons provenant de diverses populations d'Amérique du Sud (collection Coriell, Venezuela, Guyane, Brésil, HD18), d'Amérique du Nord (collection Coriell, Mexique, HD08, et autres) et d'Antioquia, en Colombie, ont été analysés dans le cadre d'une étude en collaboration avec le Dr Ruiz-Linares. Les résultats seront discutés dans le contexte de la récente colonisation du continent américain par les Européens, puisque ces populations sont connues pour leur important métissage avec les Européens.

Chacune des populations est approximativement située sur une carte, figure 3 page 51 pour les populations Amérindiennes et métisses, et figure 4 page 52 pour les autres populations.

POPULATIONS	LOCALISATION	NOMBRE DE CHROMOSOMES ETUDIÉS
Chipewyans	Amérique du Nord	75
Crees	Amérique du Nord	36
Ojibwas	Amérique du Nord	52
Navajos	Amérique du Nord	24
Ticunas	Amérique du Sud	18
Antioquiens	Amérique du Sud	56
Mexicains	Amérique du Nord	61
Coriell HD18	Amérique du Sud	13
Aches	Amérique du Sud	40
Ibériens	Sud-Ouest de l'Europe	53
Chinois	Est de l'Asie	96
Algonquins	Amérique du Nord	33
Mayas	Amérique Centrale	80
Karitianas	Amérique du Sud	83
Japonais	Asie	65
Mongols	Asie	296
Sibériens	Asie	53
Guinéens (PNG)	Océanie	67
Africains	Afrique Sub-Saharienne	204
Asiatiques du Sud-Est	Asie	91
Européens de l'Ouest	Europe	410

Tableau V : Populations étudiées et nombre de chromosomes X analysés pour chacune d'elles. Les Européens de l'Ouest comprennent des Français, des Italiens et des Allemands. Les Espagnols et Basques composent les Ibériens et sont compris dans le groupe des Européens de l'Ouest lors des analyses statistiques, à moins que cela ne soit spécifié. Les Africains sont composés de Mossi, Rimaibe, Biaka et Mbuti.



Figure 3 : Localisation des populations Amérindiennes étudiées. En **blanc**, les populations représentant le groupe linguistique *Amerind*, en **jaune** celles représentant le groupe linguistique *Na Dene*, et en **rose** les populations à fort métissage.

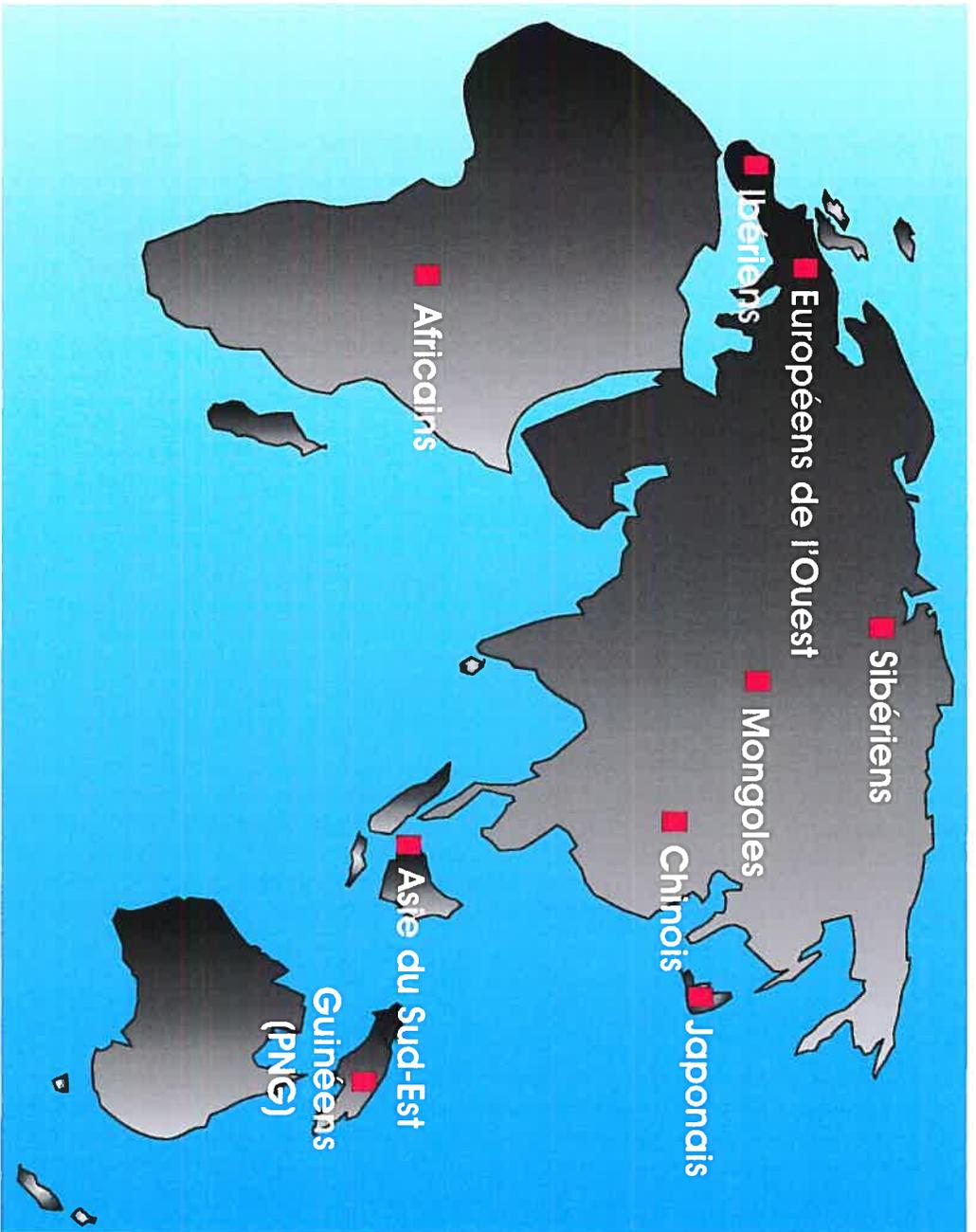


Figure 4 : Localisation des populations non-américaines étudiées.

B) Extraction de l'ADN

Lorsque les échantillons consistaient en extraits sanguins, l'ADN était extrait en utilisant le kit d'extraction *DNA Purification Kit* de PUREGENE, à partir de trois ou dix millilitres de sang. Le protocole d'extraction pour trois millilitres de sang est décrit ci-après. Il est à noter que la composition des solutions employées n'est pas fournie par le fabricant.

Dans un tube de 15ml, 3ml de sang sont ajoutés à 9ml de solution de lyse des globules rouges (*RBC Lysis Solution*). Après quelques inversions du tube, le tout est laissé 10 minutes à température ambiante pour incubation. Le tube est ensuite centrifugé à 2000g pendant 10 minutes.

Le surnageant est jeté, en prenant soin de garder au fond du tube le culot de globules blancs et environ 100µl du liquide résiduel. Après une vigoureuse agitation du tube pour resuspendre le culot de cellules dans le liquide résiduel, 3ml de solution de lyse (*Cell Lysis Solution*) sont ajoutés et le tout est mélangé en pipetant. D'ordinaire, aucune incubation n'est nécessaire; toutefois, lorsque des amas de cellules étaient présents, le tube était placé à 37°C avec agitation jusqu'à homogénéisation de la solution.

Sont alors ajoutés au lysat 15µl de solution RNase (*RNase A Solution*), et après 25 inversions du tube, le tout est laissé 15 à 60 minutes à 37°C. Après avoir laissé la solution refroidir jusqu'à température de la pièce, 1ml de solution de précipitation des protéines (*Protein Precipitation Solution*) est ajouté au lysat et le tout est ensuite agité à haute vitesse pendant 20 secondes. Une centrifugation à 2000g pendant 10 minutes est alors pratiquée, créant un culot protéique.

Le surnageant est transféré dans un nouveau tube de 15ml contenant 3ml d'isopropanol 100%, et le tout est doucement inversé environ 50 fois, jusqu'à formation d'un amas visible d'ADN. Le tube est ensuite centrifugé 3 minutes à 2000g, à la suite de quoi l'ADN est visible sous forme d'un petit culot blanc. Le surnageant est jeté et 3 ml d'éthanol 70% sont alors ajoutés, puis le tube est inversé quelques fois afin de laver le culot d'ADN.

Une centrifugation à 2000g pendant 1 minute est pratiquée, et l'éthanol est doucement vidé du tube. Par la suite, le tube est inversé sur un papier absorbant et laissé sécher à l'air pendant environ 10 à 15 minutes.

L'ADN est alors réhydraté par l'ajout de 250 μ l de TE et laissé pour incubation à 65°C pendant 1 heure ou bien à température ambiante pendant 8 à 10 heures. Lorsque la viscosité de la solution trahissait une très importante concentration en ADN, du TE supplémentaire, jusqu'à 250 μ l parfois, était ajouté.

C) Détermination de la concentration en ADN

Lorsque isolé par nos soins, étant alors disponible en volume important, l'ADN fut quantifié en utilisant un spectrophotomètre (Ultrospec 2000 de *Pharmacia Biotech*). Certains échantillons furent reçus de collaborateurs externes, directement sous forme d'ADN en solution, mais en faible volume et avec une concentration inconnue. Pour ces derniers, il fut décidé d'employer un fluorimètre (*Hoefler TKO 120*) afin de quantifier l'ADN, cette méthode permettant la quantification de faibles concentrations d'ADN, et donc de travailler avec de plus faibles volumes.

1) Spectrophotomètre

Ce système de mesure est très simple et donne une mesure directe de la concentration d'ADN et de protéines, en mesurant le taux d'absorption de la solution à 260 et 280nm, respectivement. Un volume total d'au moins 100 μ l est nécessaire afin que toute la surface de mesure soit utilisée. Dans le cas contraire, la mesure est faussée à cause d'une rupture du niveau d'absorbance.

Après chaque mesure, la solution est récupérée et la cuvette lavée. Afin de limiter les risques de contamination des solutions stock d'ADN, des dilutions furent utilisées pour les mesures de concentration : 10 μ l de solution d'ADN à mesurer furent dilués dans 90 μ l de TE. 100 μ l de TE furent placés dans une seconde cuvette comme référence. Il suffisait donc de multiplier par 10 la concentration donnée par la mesure, exprimée en ng/ μ l, pour connaître la concentration de la solution stock.

2) Fluorimètre

J'ai effectué les mesures en utilisant 2 μ l de solution d'ADN auxquels furent ajoutés 43 μ l de TNE1x et 5 μ l de solution CAS. La composition de la solution CAS se trouve dans le tableau VI page 56. La première étape fut d'établir une courbe standard de concentration d'ADN ; la valeur mesurée par le fluorimètre pour différentes concentrations connues d'ADN d'extrait de thymus de veau est notée et reportée sur un graphe, voir figure 5 page 56. La courbe ainsi obtenue sert à faire la correspondance entre la valeur lue sur le fluorimètre et la mesure de la concentration en ADN de la solution.

Il est à noter que la solution CAS doit être protégée de la lumière, et par conséquent être préparée juste avant la mesure. La mesure s'effectue dans de fins capillaires en verre, qui sont bouchés par une pâte une fois la solution en place. La rapidité d'exécution peut être déterminante pour la précision des mesures, car il faut minimiser l'exposition des solutions à la lumière et éviter de trop faire chauffer la lampe du fluorimètre. Afin de s'assurer de la précision des mesures, les dilutions standards utilisées pour la courbe furent remesurées à la fin des manipulations.

CAS	
TNE 10x	100 μ l
Hoechst 33258 1mg/ml	20 μ l
Eau distillée	882 μ l

Tableau VI : Composition de la solution CAS pour la mesure de la concentration en ADN d'une solution à l'aide d'un fluorimètre à capillaires.

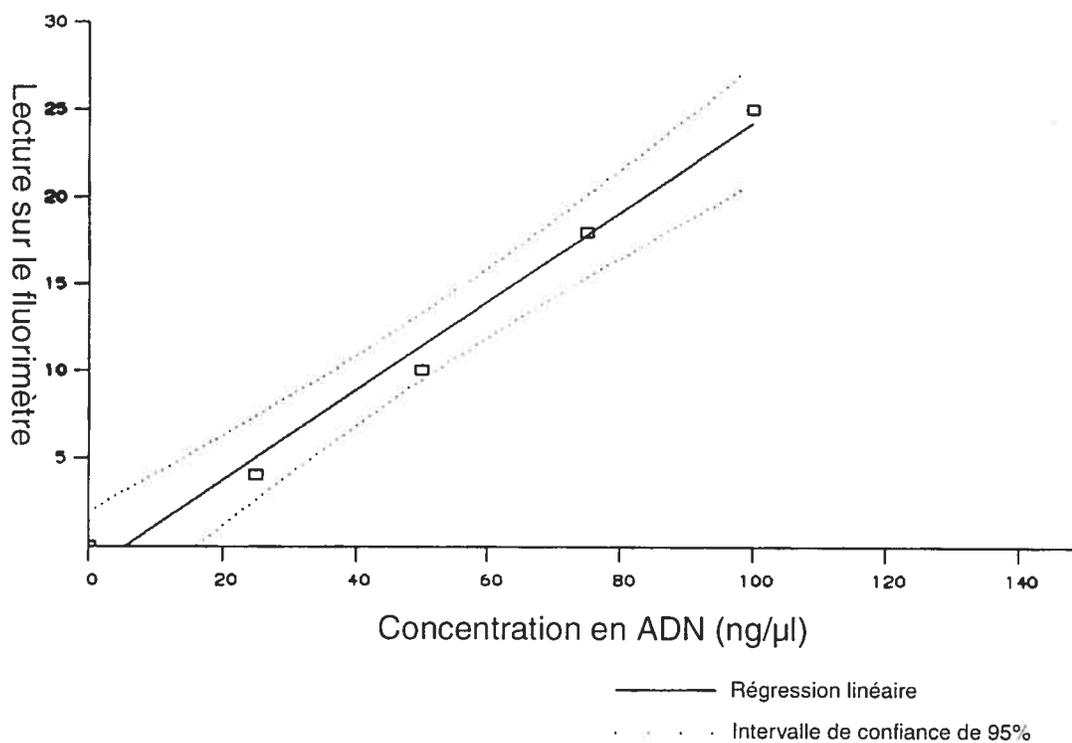


Figure 5 : Courbe standard de lecture de la concentration en ADN, établie à partir de la mesure de concentrations connues en ADN de thymus de veau.

D) Système *Dys 44* et amplification par PCR

1) Description du système

Le système *Dys44* consiste en 35 polymorphismes se trouvant dans la portion intronique entourant l'exon 44 du gène de la dystrophine, protéine impliquée dans la dystrophie musculaire de Duchenne (DMD), sur le chromosome X en Xp21.3, tel que décrit précédemment par Zietkiewicz, Yotova et al [273], et référencé dans *Genbank* (<http://www.ncbi.nlm.nih.gov>) par le code U94396. Des 35 polymorphismes trouvés dans cette région, 31 sont des polymorphismes simples, les autres consistant en deux délétions de trois nucléotides, une duplication de huit nucléotides et un microsatellite (répétition de T, dont le nombre varie de 14 à 24).

Le détail des polymorphismes, y compris leur localisation et le nom du fragment de PCR sur lequel ils se trouvent, est présenté dans le tableau VII page 58. Il est à noter que la localisation des polymorphismes, comme celle des amorces de PCR ou des sondes ASO, s'effectue par rapport au premier nucléotide de l'exon 44 du gène de la dystrophine, ce dernier prenant la valeur +1. Par conséquent, une valeur négative indique que l'on se trouve en amont (3') de l'exon 44, et une valeur positive que l'on se trouve en aval (5'). L'ancienne nomenclature, utilisée dans les anciennes publications, est également mentionnée afin de faciliter la comparaison avec ces dernières[273, 274, 278, 279].

2) Amplification par PCR

Il fut décidé, pour des raisons pratiques et économiques, d'amplifier la région d'intérêt en la décomposant en quatre fragments de tailles similaires. Non seulement cela permet de limiter à quatre le nombre d'amplifications qui sont nécessaires pour l'étude de l'ensemble des polymorphismes, d'où l'intérêt économique, mais cela permet également de faciliter la lecture des résultats d'hybridation, en effet, lorsqu'une amplification fonctionne, elle atteint un même niveau d'amplification pour un groupe de polymorphismes, limitant ce type de disparité.

Frag. PCR	Position	Ancien nom du polymorphisme	Ancien numéro du polymorphisme	Afr	Allèle ancestral	Nouvel Allèle	Brin
DMD1	-2782	H1/H2	2	N	C	A	F
	-2781		3		A	C	
	-2777	H3/H5	5	N	G	A	F
	-2678	1A-176	8	N	A	C	F
	-2441	1B-187	10	N	T	C	F
	-2291	1B-337	12	O	A	G	F
	-2033	1B-595/2-100	14	N	T	C	F
	-1760	2up-279	15	O	T	G	F
	-1430	3med-161	18	O	T	C	F
	-1313	3up-278	20	N	A	G	F
	-1098	3up-493	25	O	C	G	F
	-1058	3low-533	30	O	T	C	F
DMD2	-599	4B-149/173	32	N	T	C	F
	-182	5low-356-358	33	N	TGT	del	F
	-115	5low-423	35	N	A	G	R
	+116	7A-107-109	38	O	TCA	del	F
	+556	7B-291	40	N	A	G	F
	+800	8-140	45	N	T	G	R
DMD3	+1241	9low-138	48	O	G	A	R
	+1485	9up-382	50	N	T	C	F
	+1813	10med-151	55	N	A	G	F
	+1851	10up-189-192	58	N	T	A	R
	+1854		59	N	T	G	R
	+2149	10low-487-488	64	N	A	G	F
	+2150		65	N	C	A	R
	+2590	11Aup-158	70	N	A	G	F
	+2650	11Aup-218	71	N	T	C	F
	+2663	11Aup-231	72	O	T	A	R
	+2684	11POL	73	N	Microsatellite Tn		11Pol
	+3144	12up-299-306	85	N	8N	Dupl	F
DMD4	+3720	13A-271	86	O	A	G	F
	+4099	14-161	87	N	C	T	F
	+4128	14-190	88	O	T	C	F
	+4264	14-326	90	N	T	C	F
	+4586	15low-240	93	N	T	C	F
	+4987	15up-641-1	95	O	A	C	F
					A	G	F

Tableau VII : Polymorphismes de *Dys44* avec leur localisation par rapport au premier nucléotide de l'exon 44, ainsi que l'ancienne nomenclature. « O » dans la colonne Afr. signifie que le polymorphisme est caractéristique des populations africaines. F code pour le brin sens et R pour le brin anti-sens.

Après de nombreux essais, il fut possible d'obtenir un bon niveau d'amplification pour les quatre fragments, dans les mêmes conditions d'amplification, à partir de 20ng d'ADN génomique. Les séquences des amorces utilisées, leur localisation, l'identification des fragments amplifiés et les conditions précises des amplifications sont présentées dans le tableau VIII ci-après. Après chaque PCR, 5µl du produit d'amplification étaient prélevés et déposés sur un gel d'agarose ultra-pur à 2%, contenant 20 microgrammes, afin de vérification de la qualité de l'amplification par électrophorèse, voir la figure 6 page 60. La variation observée du taux d'amplification est principalement fonction du niveau de conservation de l'ADN utilisé.

Frag.	Début	Fin	Primer Forward 5' – 3'	Primer Reverse 5' – 3'
DMD 1	-2824	-865	GGCAATTACCAATATGATGAGTGATTG	CTAAACTATGCGATGCCTTCTTTAATCC
DMD2	-892	1188	GGATTAAAGAAGGCATCGCATAGTTTAG	CCATTATGCTACATTAGCCCCAAAAT
DMD 3	1164	3375	ATTTTGGGCTAATGTAGCATAATGG	GGAAAGAAAATACCTTAACTGTTTGG
DMD 4	3351	5024	CCAAACAGTTAAGGTATTTCTTTCC	CTACTGGCAACTACAAACATTTTCC

Taq	Buffer	Magnésium	dNTPs	Primer Forward	Primer Reverse	ADN	Volume Final
2U	1x	1.5 mM	212.5 µM	0.25 µM	0.25 µM	20 ng	40 µl

1 cycle		45 cycles						1 cycle	
Dénaturation		Dénaturation		Hybridation		Elongation		Terminaison	
°C	temps	°C	temps	°C	temps	°C	temps	°C	temps
95	240s	95	30s	55	30s	68	120s	68	600s

Tableau VIII : Conditions d'amplification des différents fragments utilisés pour le génotypage des polymorphismes du système *Dys44*. Dans la première partie du tableau sont présentées les séquences des oligonucléotides utilisés, ainsi que leur localisation. La seconde partie présente les conditions biochimiques de l'amplification, la Taq utilisée étant la *Taq Platinum* d'*Invitrogen* et la concentration en dNTPs étant donnée pour chacun d'eux. Les réglages de l'appareil à PCR sont présentés dans la dernière partie du tableau.

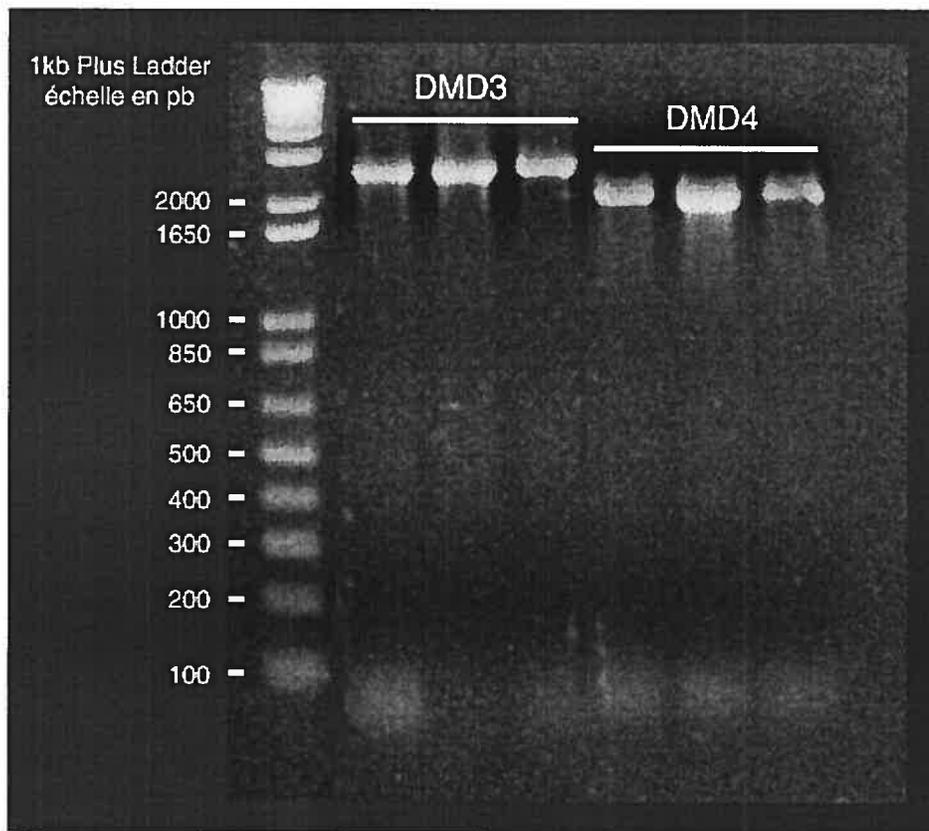


Figure 6 : Gel d'électrophorèse à 2% d'agarose ultra-pur après amplifications des fragments DMD3 et DMD4, avec le *1kb Plus Ladder* d'*Invitrogen*.

E) Transfert sur membranes

Les fragments de PCR sont dénaturés avant d'être transférés sur une membrane de nylon pour y être fixés. Pour environ 3000 échantillons, la solution de dénaturation est préparée comme suit : 200ml d'hydroxyde de sodium (NaOH) à 1M sont ajoutés à 20ml d'EDTA et 280ml d'eau distillée. A la solution ainsi obtenue 12.1ml de colorant orange, 0,02% d'Orange II dans du TE 1x, sont ajoutés afin de permettre la visualisation du transfert sur la membrane. Pour chaque échantillon, 165 μ l de cette solution de dénaturation sont ajoutés directement dans le puits de la plaque de PCR.

Deux membranes de nylon Hybond-N⁺ (*Amersham Pharmacia Biotech*), préalablement identifiées selon un code arbitraire, sont disposées dans des appareils avec aspiration pour permettre le transfert. La moitié du contenu de chaque puits est alors transféré sur la première membrane, et l'autre moitié sur la seconde membrane. Les deux membranes sont alors brièvement rincées dans un peu d'eau distillée avant d'être disposées dans un four à UV (*Stratagene UV stratalinker 1800*). Les membranes sont alors exposées à 120 mJ de rayonnement UV afin de lier de façon covalente l'ADN à la membrane.

F) Hybridation ASO

L'hybridation spécifique d'un allèle, ou hybridation ASO, est une technique d'hybridation permettant le génotypage de mutations préalablement identifiées. La stabilité de l'association d'une sonde oligonucléotidique avec un brin d'ADN dépend non seulement de la température, mais également du degré d'appariement. Un mésappariement, spécialement lorsqu'il est situé proche du milieu de l'oligonucléotide, induit une instabilité de l'association entre la sonde et l'ADN, en comparaison à un appariement parfait. Par conséquent, on peut génotyper un échantillon d'ADN pour des variations connues en effectuant une série d'hybridations avec des oligonucléotides rendant compte de ces variations dans leur séquence.

Il est sans doute plus aisé de comprendre le fonctionnement de cette technique en observant un exemple d'hybridation, accompagné de la méthode d'interprétation, comme cela est présenté dans la figure 7 en page 62. La méthode de détection ici présentée est celle de l'autoradiographie d'oligonucléotides radioactivement marqués, toutefois, d'autres techniques de marquage et de détection, comme la fluorescence par exemple, sont envisageables.

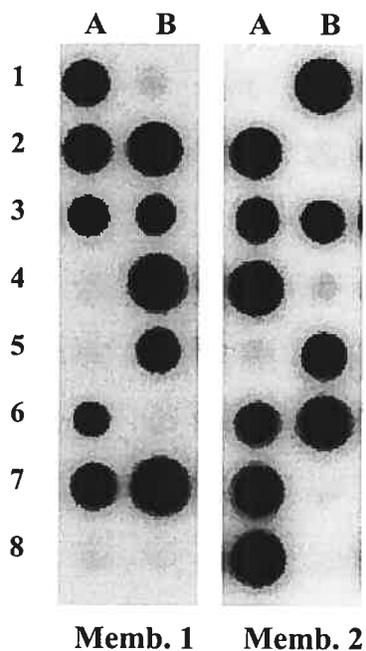


Figure 7 : Hybridation ASO. La membrane 2 est une réplique parfaite de la membrane 1 ; ceci est obtenu en fixant la première moitié des produits d'amplification sur la membrane 1 et la seconde moitié sur la membrane 2. Ainsi, l'échantillon fixé en position A1 de la membrane 1 est le même que celui fixé en position A1 de la membrane 2, et il en va de même pour les autres positions. Les sondes oligonucléotidiques ayant servies à l'hybridation pour les membranes 1 et 2 sont identiques à l'exception d'une position, celle du polymorphisme. La première sonde est spécifique du premier allèle, alors que la seconde sonde est spécifique du deuxième allèle. En position A1, l'hybridation n'est positive que pour la membrane 1 (et donc la sonde 1) et par conséquent l'individu est homozygote pour l'allèle 1. Le contraire est observé pour l'individu en B1, qui est homozygote pour l'allèle 2. L'hybridation en position A2 est positive pour les deux membranes, l'individu considéré est donc hétérozygote. Les polymorphismes étudiés se trouvant sur le chromosome X, tous les individus de sexe mâle apparaissent donc homozygotes, et leur haplotype est résolu en même temps que leur génotype.

1) Marquage radioactif des sondes

La liste complète des sondes oligonucléotidiques utilisées dans cette étude est présentée dans le tableau IX, page 64. Les sondes furent commandées chez *Invitrogen* et furent reçues sous forme de poudre, dans une quantité avoisinant les 50nmol pour chaque oligonucléotide. Les oligonucléotides furent réhydratés de manière à ce que la solution obtenue ait une concentration de 200 μ M. Pour chacune de ces solutions stock, une solution diluée à 25 μ M fut faite, afin de limiter les risques de contamination.

Le ^{32}P utilisé pour le marquage des sondes fut commandé chez *Amersham* sous forme de γ ATP (le phosphore radioactif est en position γ de l'ATP) en solution, d'activité totale de 5mCi pour un volume de 30 μ l. Les quatre premières semaines d'utilisation, cette solution était diluée à 1/10, pour être utilisée pure par la suite (la demi-vie du ^{32}P est 14,5 jours).

Le protocole de marquage est en fait très simple, il suffit de laisser incuber pendant une heure une solution contenant l'oligonucléotide à marquer, un tampon d'échange, de la kinase T4, de l'eau et du γ ATP radioactif. Les proportions exactes de chacun de ces éléments sont présentées dans le tableau X, page 64. Après l'incubation, 15 μ l d'EDTA à 10mM sont ajoutés à la solution afin d'arrêter la réaction.

2) Protocole

Les deux membranes de nylon sur lesquelles furent fixés les échantillons d'ADN dénaturés sont placées dans deux tubes de verre différents. Lorsque plusieurs membranes sont placées dans un même tube à hybridation, celles-ci sont disposées sur une toile synthétique qui est par la suite enroulée, évitant ainsi que les membranes ne se collent les unes aux autres, ce qui pourrait entraver l'hybridation en ne permettant pas la bonne circulation de la solution d'hybridation. La composition des solutions employées lors de cette procédure est disponible dans le tableau XI page 67.

	Position	Séquence oligonucléotide 1	Tm 1	Séquence oligonucléotide 2	Tm 2
DMD1	-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
	-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
	-2033	CTT CGG CTA CCT TCG	64	CTT TGG CTA CCT TCG	61
	-1760	GAA ACA AGA GAT ATT	50	GAA ACA AGA GAG ATT	53
	-1430	ATC CAT TGC TTC TTA	53	ATC CAC TGC TTC TTA	56
	-1313	ATG CTT GCT TGG ACA	59	ATG CTT GCT TGG GCA	61
	-1098	ATG CCC CAG TTG ATT	59	ATG CSC CAG TTG ATT	59
	-1058	AGG TAA ACA TAC AGC	56	AGG TAA ACA TAT AGC	53
DMD2	-599	ATA ATT TCT CTG TGG	53	ATA ACT TCT CTG TGG	56
	-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	56
	116	GTT GTC ATC ATT ATA	50	GTT GTC --- ATT ATA	40
	556	ATA GAG AAA CAG CAT	53	ATA GGG AAA CAG CAT	56
	800	GGC CTT AAA AAT TGG	56	GGC ATT AAA AAT TGG	53
DMD3	1241	TCT ATA TCT ACC CCA	56	TCT ATA TCT ATC CCA	53
	1485	AGG TTA ACT AGG GAG	59	AGG TTA ATT AGG GAG	56
	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
	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
	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
3144	CTG GTC TAG ATC TC- --- -C	61	CTG GTC TAG ATC TCT AGA TCT CC	73	
DMD4	3720	GTG AAA AAA ATC AAC	50	GTG AAA AAA GTC AAC	53
	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
	4586	ACT GAA TTA TCG TCT	53	ACT GAA TTA TTG TCT	50
	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 IX : Sondes oligonucléotidiques pour hybridation ASO, avec leur Tm.

Oligonucleotide (25µM)	5x Exchange Buffer	T4 kinase	Eau nanopure	32P
1µl	3µl	1µl	9µl	1µl

Tableau X : Proportions des ingrédients nécessaires pour le marquage radioactif des sondes oligonucléotidiques.

Un volume suffisant de solution d'hybridation, préalablement chauffée, est ajouté dans le tube, et ce dernier est fermé et mis pour pré-hybridation dans un four. Un volume suffisant est tel qu'une partie des membranes baigne constamment dans la solution d'hybridation pendant les rotations du tube à hybridation, placé horizontalement, dans le four de type rôtissoire. Pour quatre membranes dans un tube *Fisher Scientific* de 35x300mm, 50ml suffisaient pour répondre à ces exigences. La solution d'hybridation est préalablement placée dans un bain-marie à la température souhaitée, non seulement pour diminuer le temps de pré-hybridation, mais également pour permettre la bonne dissolution des cristaux qui se forment à température ambiante lorsque du SDS est en présence de chlorure de sodium.

La pré-hybridation, d'une durée approximative de 30 minutes, est effectuée à la température d'hybridation, et permet à la solution d'hybridation d'atteindre la température voulue avant l'ajout des sondes radioactives ; c'est la température d'hybridation qui va déterminer la stringence de l'hybridation. En générale la température d'hybridation est choisie proche de la température de *melting*, T_m , de la sonde employée. Cette température est celle à laquelle 50% des sondes sont hybridées, bien entendu, il s'agit là d'un équilibre dynamique. Elle peut être approximativement déterminée par un calcul prenant en compte la salinité de la solution d'hybridation et la composition de la sonde en G et C. La T_m des sondes employées, est fournie dans le tableau IX, page 64.

Les tubes sont ensuite retirés du four, et les sondes sont ajoutées, à raison de 3 μ l de sondes marquées, comme décrit précédemment, pour 50ml de solution d'hybridation. Afin de s'assurer d'une plus grande spécificité des hybridations, il fut décidé de pratiquer la technique du blocage, qui consiste en l'ajout d'un excès de sonde froide spécifique pour un allèle différent de celui auquel la sonde radioactive va s'hybrider, ici, 6 μ l de sonde froide à 25 μ M furent ajoutés. Ces sondes froides, qui ne seront pas détectées par autoradiographie, permettent de saturer les sites auxquels les sondes marquées pourraient se lier de façon non spécifique.

Les tubes sont ensuite remis dans le four pour une durée d'environ 30 minutes, temps amplement suffisant pour permettre l'hybridation. A la fin de cette période, les tubes sont retirés du four et les solutions sont récupérées dans des contenants spéciaux pour être traitées adéquatement par une entreprise spécialisée. Deux lavages successifs, de 5 minutes chacun, sont alors effectués à l'aide de 50ml de solution de lavage, à température ambiante. Les membranes sont alors transférées dans des sacs en plastique qui seront scellés et placés dans une cassette d'exposition. En chambre noire, un film radiographique (Fuji Medical X-Ray 100NIF) est placé dans la cassette au contact du sac contenant les membranes. Des écrans réflecteurs peuvent également être placés dans la cassette afin d'amplifier le signal. Le temps d'exposition peut varier de 6 à 48 heures, selon la « fraîcheur » de la radioactivité employée, la présence ou non d'écrans réflecteurs et la température à laquelle l'exposition s'effectue (les cassettes peuvent être placées à -80°C pour augmenter la sensibilité du film).

Malgré les précautions prises, il peut arriver que l'hybridation s'avère peu spécifique. Au lieu de recommencer l'hybridation à une plus haute température, il est possible de relaver les membranes à une température plus élevée que celle utilisée pour l'hybridation, ce qui permet la dissociation des sondes hybridées de façon non-spécifique. Ces lavages peuvent être réitérés, à des températures croissantes, jusqu'à obtention d'un résultat clair et spécifique.

Enfin, lorsque les résultats sont satisfaisants, toutes les sondes sont dissociées de l'ADN en laissant baigner les membranes dans une solution de dissociation préalablement chauffée à 100°C . Après vérification à l'aide d'un compteur Geiger qu'il ne reste pas trace de radioactivité sur les membranes, ces dernières sont alors disponibles pour une autre expérience d'hybridation.

Solution d'hybridation (400ml)		Solution de lavage (300ml)	
Produit	Volume (ml)	Produit	Volume (ml)
SSPE 20x	20	SSPE 20x	30
NaCl 5M	60	SDS 10%	3
SDS 10%	40	Eau distillée	267
Héparine à 200 ng/ μ l	1,6	Solution de dissociation (525ml)	
Tris 1M	28	Produit	Volume (ml)
Eau distillée	250,4	SDS 10%	25
		Eau distillée	500

Tableau XI : composition des solutions utilisées lors de l'hybridation ASO.

G) Base de données, étude des haplotypes

Une base de données, nommée Pandora, a en grande partie été conçue pour ce projet, par Dominik Gehl, secondé par Hérin Vasquez, Patrick Beaulieu et Richard Côté. Développée en MySQL avec une interface HTML, la base de données a pour but de faciliter le stockage et l'analyse des résultats, tout en limitant les risques d'erreur humaine, voir figure 8 page 68. Cela permet, entre autres, de centraliser l'information disponible sur les différents polymorphismes ainsi que la séquence des fragments étudiés, de garder un inventaire précis des échantillons disponibles dans la banque de tissus, d'archiver les informations disponibles pour chaque échantillon (ethnie, langues parlées, localisation géographique) et d'effectuer des recherches basées sur ces informations, voir figure 9 page 69.

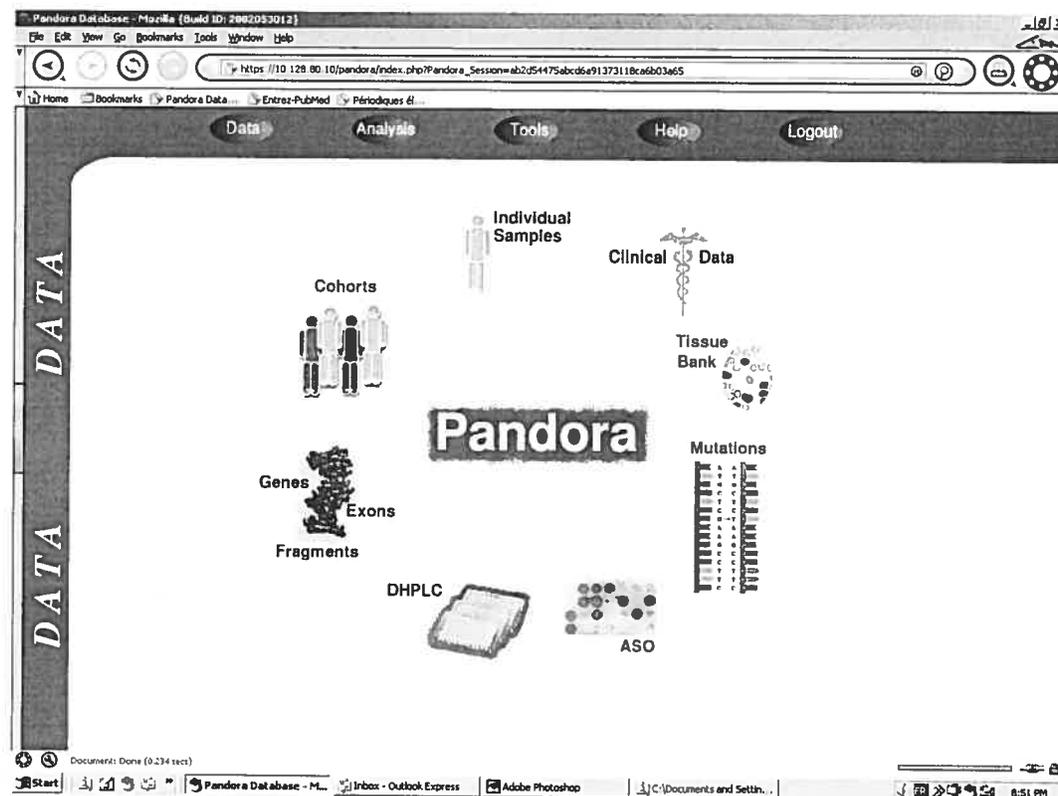


Figure 8 : Ecran d'accueil de la base de données Pandora avec les différentes options disponibles.

Une collaboration étroite entre l'équipe de recherche et l'équipe de programmation a permis l'établissement d'outils spécifiques pour la saisie des résultats d'hybridation, voir figure 10 page 69. Le fait que l'accès à la base de données soit géré par un nom d'utilisateur et un mot de passe propre à chaque utilisateur ne permet pas seulement d'assurer la confidentialité des informations disponibles, mais également de mettre en place un système de contrôle de la qualité des résultats. En effet, avant d'être disponible pour analyse, chaque résultat doit faire l'objet d'une seconde lecture par un autre membre du laboratoire, ceci étant assuré par l'association du nom de l'utilisateur à la grille de résultats. Seule une lecture identique de la part des deux chercheurs permet l'enregistrement définitif des informations, dans le cas d'un désaccord les résultats sont relus en groupe, et les hybridations recommencées si besoin est. Une fois les résultats enregistrés, ils sont directement disponibles sous forme de génotypes, voir figure 11 page 70.

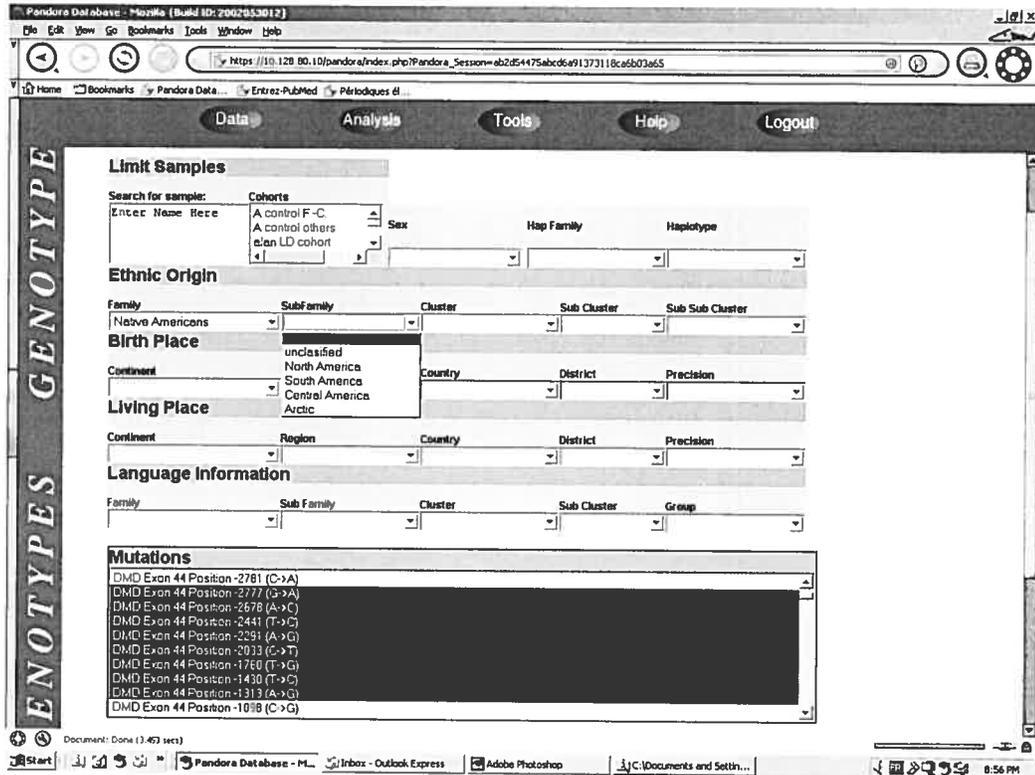


Figure 9 : Recherche de données spécifiques à certains polymorphismes de *Dys44* dans Pandora, en fonction de la classification ethnique.

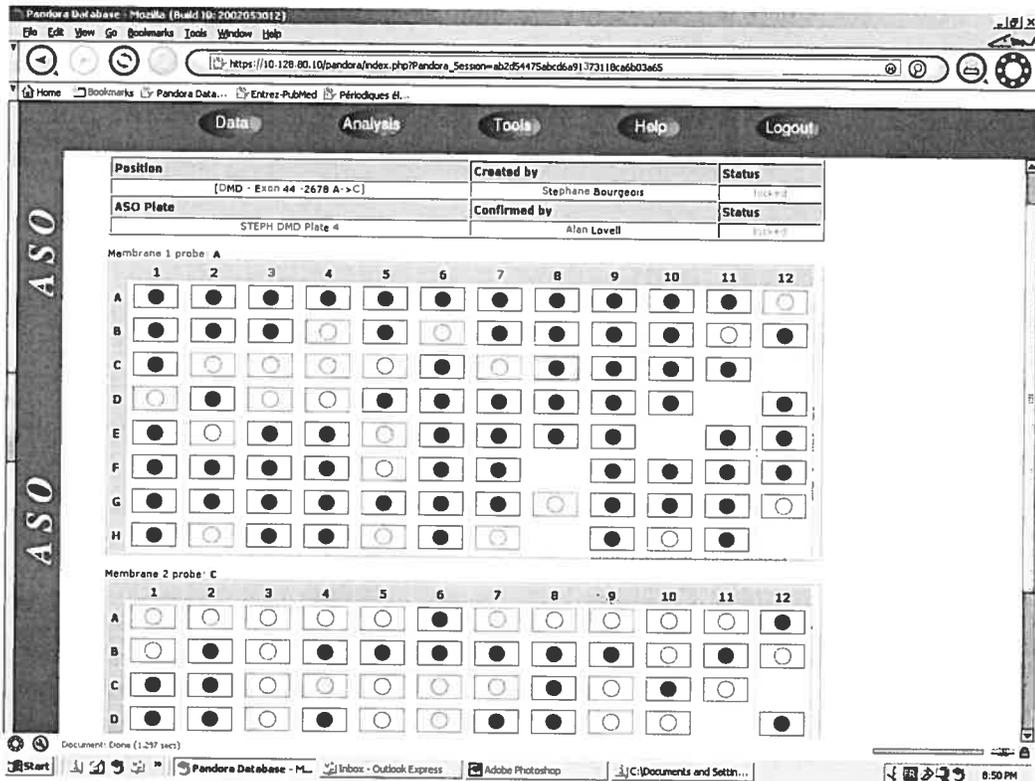


Figure 10 : Grilles de saisie des résultats d'hybridation

Click to download as Arlequin file

Sample	Ethnicity	DMD Exon 44 Position -2777 (G->A)	DMD Exon 44 Position -2678 (A->C)	DMD Exon 44 Position -2441 (T->C)	DMD Exon 44 Position -2291 (A->G)	DMD Exon 44 Position -2033 (C->T)	DMD Exon 44 Position -1760 (T->G)	DMD Exon 44 Position -1430 (T->C)	DMD Exon 44 Position -1313 (A->G)
11820	Navajo	G	A	C	A	T	T	T	A
11821	Navajo	G	C	T	A	C	T	T	A
11822	Navajo	G	A	C	A	T	T	T	A
11823	Navajo	G	A	T	A	C	T	T	A
11824	Navajo	G	A	C	A	T	T	T	A
11825	Navajo	G	A	T	A	C	T	T	A
11826	Navajo	G	A	T	A	C	T	T	A
11827	Navajo	G	C	T	A	C	T	T	A
11828	Navajo	G	C	T	A	C	T	T	A
11829	Navajo	G	C	T	A	C	T	T	A
11830	Navajo	G	A	T	A	C	T	T	A
11831	Navajo	G	A	C	A	T	T	T	A
11832	Navajo	G	A	C	A	T	T	T	A
11833	Navajo	G	A	T	A	C	T	T	A
11834	Navajo	G	A	T	A	C	T	T	A
11835	Navajo	G	A	T	A	C	T	T	A
11836	Navajo	G	A	T	A	C	T	T	A
11837	Navajo	G	A	T	A	C	T	T	A
11838	Navajo	G	A	T	A	C	T	T	A
11839	Navajo	G	A	T	A	C	T	T	A
11840	Navajo	G	A	T	A	C	T	T	A
11841	Navajo	G	A	T	A	C	T	T	A
11842	Navajo	G	A	T	A	C	T	T	A
11843	Navajo	G	A	T	A	C	T	T	A
11844	Navajo	G	A	T	A	C	T	T	A
11845	Navajo	G	A	T	A	C	T	T	A
11846	Navajo	G	A	T	A	C	T	T	A
11847	Navajo	G	A	T	A	C	T	T	A
11848	Navajo	G	A	T	A	C	T	T	A
11849	Navajo	G	A	T	A	C	T	T	A
11850	Navajo	G	A	T	A	C	T	T	A
11851	Navajo	G	A	T	A	C	T	T	A
11852	Navajo	G	A	T	A	C	T	T	A
11853	Navajo	G	A	T	A	C	T	T	A
11854	Navajo	G	A	T	A	C	T	T	A
11855	Navajo	G	A	T	A	C	T	T	A
11856	Navajo	G	A	T	A	C	T	T	A
11857	Navajo	G	A	T	A	C	T	T	A
11858	Navajo	G	A	T	A	C	T	T	A
11859	Navajo	G	A	T	A	C	T	T	A
11860	Navajo	G	A	T	A	C	T	T	A
11861	Navajo	G	A	T	A	C	T	T	A
11862	Navajo	G	A	T	A	C	T	T	A
11863	Navajo	G	A	T	A	C	T	T	A
11864	Navajo	G	A	T	A	C	T	T	A
11865	Navajo	G	A	T	A	C	T	T	A
11866	Navajo	G	A	T	A	C	T	T	A
11867	Navajo	G	A	T	A	C	T	T	A
11868	Navajo	G	A	T	A	C	T	T	A
11869	Navajo	G	A	T	A	C	T	T	A
11870	Navajo	G	A	T	A	C	T	T	A
11871	Navajo	G	A	T	A	C	T	T	A
11872	Navajo	G	A	T	A	C	T	T	A
11873	Navajo	G	A	T	A	C	T	T	A
11874	Navajo	G	A	T	A	C	T	T	A
11875	Navajo	G	A	T	A	C	T	T	A
11876	Navajo	G	A	T	A	C	T	T	A
11877	Navajo	G	A	T	A	C	T	T	A
11878	Navajo	G	A	T	A	C	T	T	A
11879	Navajo	G	A	T	A	C	T	T	A
11880	Navajo	G	A	T	A	C	T	T	A
11881	Navajo	G	A	T	A	C	T	T	A
11882	Navajo	G	A	T	A	C	T	T	A
11883	Navajo	G	A	T	A	C	T	T	A
11884	Navajo	G	A	T	A	C	T	T	A
11885	Navajo	G	A	T	A	C	T	T	A
11886	Navajo	G	A	T	A	C	T	T	A
11887	Navajo	G	A	T	A	C	T	T	A
11888	Navajo	G	A	T	A	C	T	T	A
11889	Navajo	G	A	T	A	C	T	T	A
11890	Navajo	G	A	T	A	C	T	T	A
11891	Navajo	G	A	T	A	C	T	T	A
11892	Navajo	G	A	T	A	C	T	T	A
11893	Navajo	G	A	T	A	C	T	T	A
11894	Navajo	G	A	T	A	C	T	T	A
11895	Navajo	G	A	T	A	C	T	T	A
11896	Navajo	G	A	T	A	C	T	T	A
11897	Navajo	G	A	T	A	C	T	T	A
11898	Navajo	G	A	T	A	C	T	T	A
11899	Navajo	G	A	T	A	C	T	T	A
11900	Navajo	G	A	T	A	C	T	T	A
11901	Navajo	G	A	T	A	C	T	T	A
11902	Navajo	G	A	T	A	C	T	T	A
11903	Navajo	G	A	T	A	C	T	T	A
11904	Navajo	G	A	T	A	C	T	T	A
11905	Navajo	G	A	T	A	C	T	T	A
11906	Navajo	G	A	T	A	C	T	T	A
11907	Navajo	G	A	T	A	C	T	T	A
11908	Navajo	G	A	T	A	C	T	T	A
11909	Navajo	G	A	T	A	C	T	T	A
11910	Navajo	G	A	T	A	C	T	T	A
11911	Navajo	G	A	T	A	C	T	T	A
11912	Navajo	G	A	T	A	C	T	T	A
11913	Navajo	G	A	T	A	C	T	T	A
11914	Navajo	G	A	T	A	C	T	T	A
11915	Navajo	G	A	T	A	C	T	T	A
11916	Navajo	G	A	T	A	C	T	T	A
11917	Navajo	G	A	T	A	C	T	T	A
11918	Navajo	G	A	T	A	C	T	T	A
11919	Navajo	G	A	T	A	C	T	T	A
11920	Navajo	G	A	T	A	C	T	T	A
11921	Navajo	G	A	T	A	C	T	T	A
11922	Navajo	G	A	T	A	C	T	T	A
11923	Navajo	G	A	T	A	C	T	T	A
11924	Navajo	G	A	T	A	C	T	T	A
11925	Navajo	G	A	T	A	C	T	T	A
11926	Navajo	G	A	T	A	C	T	T	A
11927	Navajo	G	A	T	A	C	T	T	A
11928	Navajo	G	A	T	A	C	T	T	A
11929	Navajo	G	A	T	A	C	T	T	A
11930	Navajo	G	A	T	A	C	T	T	A
11931	Navajo	G	A	T	A	C	T	T	A
11932	Navajo	G	A	T	A	C	T	T	A
11933	Navajo	G	A	T	A	C	T	T	A
11934	Navajo	G	A	T	A	C	T	T	A
11935	Navajo	G	A	T	A	C	T	T	A
11936	Navajo	G	A	T	A	C	T	T	A
11937	Navajo	G	A	T	A	C	T	T	A
11938	Navajo	G	A	T	A	C	T	T	A
11939	Navajo	G	A	T	A	C	T	T	A
11940	Navajo	G	A	T	A	C	T	T	A
11941	Navajo	G	A	T	A	C	T	T	A
11942	Navajo	G	A	T	A	C	T	T	A
11943	Navajo	G	A	T	A	C	T	T	A
11944	Navajo	G	A	T	A	C	T	T	A
11945	Navajo	G	A	T	A	C	T	T	A
11946	Navajo	G	A	T	A	C	T	T	A
11947	Navajo	G	A	T	A	C	T	T	A
11948	Navajo	G	A	T	A	C	T	T	A
11949	Navajo	G	A	T	A	C	T	T	A
11950	Navajo	G	A	T	A	C	T	T	A
11951	Navajo	G	A	T	A	C	T	T	A
11952	Navajo	G	A	T	A	C	T	T	A
11953	Navajo	G	A	T	A	C	T	T	A
11954	Navajo	G	A	T	A	C	T	T	A
11955	Navajo	G	A	T	A	C	T	T	A
11956	Navajo	G	A	T	A	C	T	T	A
11957	Navajo	G	A	T	A	C	T	T	A
11958	Navajo	G	A	T	A	C	T	T	A
11959	Navajo	G	A	T	A	C	T	T	A
11960	Navajo	G	A	T	A	C	T	T	A
11961	Navajo	G	A	T	A	C	T	T	A
11962	Navajo	G	A	T	A	C	T	T	A
11963	Navajo	G	A	T	A	C	T	T	A
11964	Navajo	G	A	T	A	C	T	T	A
11965	Navajo	G	A	T	A	C	T	T	A
11966	Navajo	G	A	T	A	C	T	T	A
11967	Navajo	G	A	T	A	C	T	T	A
11968	Navajo	G	A	T	A	C	T	T	A
11969	Navajo	G	A	T	A	C	T	T	A
11970	Navajo	G	A	T	A	C	T	T	A
11971	Navajo	G	A	T	A	C	T	T	A
11972	Navajo	G	A	T	A	C	T	T	A
11973	Navajo	G	A	T	A	C	T	T	A
11974	Navajo	G	A	T	A	C	T	T	A
11975	Navajo	G	A	T	A	C	T	T	A
11976	Navajo	G	A	T	A	C	T	T	A
11977	Navajo	G	A	T	A	C	T	T	A
11978	Navajo	G	A	T	A	C	T	T	A
11979	Navajo	G	A	T	A	C	T	T	A
11980	Navajo	G	A	T	A	C	T	T	A
11981	Navajo	G	A	T	A	C	T	T	A
11982	Navajo	G	A	T	A	C	T	T	A
11983	Navajo	G	A	T	A	C	T	T	A
11984	Navajo	G	A	T	A	C	T	T	A
11985	Navajo	G	A	T	A	C	T	T	A
11986	Navajo	G	A	T	A	C	T	T	A
11987	Navajo	G	A	T	A	C	T	T	A
11988	Navajo	G	A	T	A	C	T	T	A
11989	Navajo	G	A	T	A	C	T	T	A
11990	Navajo	G	A	T	A	C	T	T	A
11991	Navajo	G	A	T	A	C	T	T	A
11992	Navajo	G	A	T	A	C	T	T	A
11993	Navajo	G	A	T	A	C	T	T	A
11994	Navajo	G	A	T	A	C	T	T	A
11995	Navajo	G	A	T	A	C	T	T	A
11996	Navajo	G	A	T	A	C	T	T	A
11997	Navajo	G	A	T	A	C	T	T	A
11998	Navajo	G	A	T	A	C	T	T	A
11999	Navajo	G	A	T	A	C	T	T	A
12000	Navajo	G	A	T	A	C	T	T	A

Figure 11 : Résultats présentés sous forme de génotypes.

Les résultats peuvent être exportés dans des formats compatibles avec des logiciels externes de résolution d'haplotypes (PHASE), puis réimportés sous forme d'haplotypes. Finalement, différentes analyses de fréquences (haplotypes, polymorphismes...) sont disponibles directement dans Pandora, voir figure 12 page 71, et la combinaison des données haplotypiques et des fréquences peut être exportée vers le logiciel Arlequin pour l'analyse statistique propre à la génétique des populations.

Comme mentionné précédemment, les haplotypes ont été reconstruits en utilisant le logiciel PHASE [280], dont l'algorithme de reconstruction des haplotypes semble être l'un des meilleurs actuellement disponibles [281]. Bien qu'une reconstruction des haplotypes ne soit pas nécessaire pour les individus de sexe masculin, tous les génotypes ont été analysés dans PHASE. De plus, les haplotypes déjà identifiés dans un échantillonnage mondial furent ajoutés dans des proportions respectant celles trouvées dans la population mondiale, obligeant ainsi le logiciel à

Chapitre III

RESULTATS

Toutes populations confondues, plus de 2900 chromosomes X ont été génotypés, et 125 haplotypes différents sont actuellement disponibles dans la base de données, résultat de différents projets de recherche comparables à celui présenté ici. Les haplotypes sont arbitrairement numérotés, en commençant par B001, voir le tableau XII page 75. L'emploi d'un B majuscule signifie que l'haplotype a été trouvé chez un mâle, alors que l'emploi d'un b minuscule signifie que l'haplotype n'a jamais été trouvé chez un mâle et par conséquent est le fruit d'une reconstruction.

amérigai
323ici

Six haplotypes seuls, de B001 à B006, constituent plus de 80% de la totalité des haplotypes, B001 étant le plus fréquent et représentant à lui seul plus de 33% des haplotypes. Pour cette raison, dans les figures et les discussions, l'accent sera mis sur ces six haplotypes, les autres étant regroupés dans une catégorie « autres ». Toutefois, dans les analyses statistiques tous les haplotypes ont été utilisés sans aucune ségrégation artificielle.

A) Structure des populations amérindiennes

Afin de comparer la structure des différentes populations amérindiennes, la fréquence de chaque haplotype pour les différentes populations (voir tableau XIII page 76) est représentée sous forme d'un camembert placé sur la carte des Amériques selon la localisation de chaque ethnie (figures 13 à 18, page 77 à 82). La comparaison des populations d'Amérique du Nord, Centrale et du Sud, deux à deux, par le test du Chi carré (χ^2) est disponible dans les tableaux IV à XVI page 83.

Dans le contexte des populations amérindiennes, l'haplotype B001 n'est pas informatif, sa répartition étant à peu près égale en Amérique du Nord, Centrale et du Sud (figure 13 page 77), comme cela est confirmé par le test du Chi carré (voir tableaux XIV à XVI page 83).

L'haplotype B002 est presque complètement absent du continent américain, seulement un Ojibwa et un Navajo possèdent cet haplotype (figure 14 page 78). Quant à l'haplotype B005, il est totalement absent du continent. Aucune information ne peut être tirée de ces haplotypes dans un contexte purement amérindien, toutefois ces résultats seront explorés dans d'autres contextes dans les deux prochains chapitres.

Pour les haplotypes B003, B004 et B006 (voir figures 15 à 17 pages 79 à 81), la différence entre les Amérindiens du Nord et du Sud est très prononcée et statistiquement significative (tableau XIV page 83). Les Amérindiens du Centre semblent présenter des fréquences haplotypiques intermédiaires, entre celles des Amérindiens du Nord et du Sud, donnant ainsi une impression de gradient Nord/Sud de la répartition de ces haplotypes. Les tests de Chi carré comparant l'Amérique Centrale avec l'Amérique du Nord et du Sud (tableaux XV et XVI page 83), confirment cette impression, en notant toutefois un excès de B004 par rapport au Nord et un déficit en B003 par rapport au Sud. Si un gradient semble apparaître pour les fréquences moyennes au Nord, au Centre et au Sud, il existe des disparités importantes au niveau des populations composant ces groupes.

La fréquence moyenne des autres haplotypes dans ces trois groupes apparaît stable et non significative, voir tableaux XIII à XVI pages 76 et 83. Là encore, il existe des différences marquées entre les populations.

La structure géographique des haplotypes a été comparée à leur structure spatiale et linguistique lors d'un test d'auto-corrélation de Mantel, effectué dans Arlequin. Le résultat confirme une corrélation entre la localisation géographique et la fréquence haplotypique, mais n'indique aucune corrélation entre cette dernière et la langue parlée. Les matrices utilisées pour ce test, ainsi que les résultats, sont présentés dans le tableau XVII page 84. Cela confirme ce que l'on pouvait l'observer sur les figures, les disparités entre les ethnies ne sont pas reliées à l'appartenance à un groupe linguistique.

Ancestral	C	A	G	A	T	A	T	T	A	C	T	T	G	A	T	C	A	T	G	T	A	T	A	C	A	T	T	8N	A	C	T	T	A							
B001	-	C	-	-	-	-	C	-	-	C	-	-	-	G	-	-	-	G	-	C	-	A	G	-	-	G	-	-	-	-	-	-	-	C	C	G				
B002	A	-	-	C	-	-	C	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
B003	-	C	-	-	-	-	C	-	-	C	-	-	-	G	-	-	-	-	-	-	C	-	A	G	-	A	G	-	-	-	-	-	-	-	-	C	C	G		
B004	A	-	-	C	-	-	C	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	G	
B005	-	-	-	-	-	-	C	-	-	C	-	-	-	G	-	-	-	-	-	-	C	-	A	G	-	A	-	-	-	-	-	-	-	-	-	-	-	C	C	G
B006	-	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	
...																																								
...																																								
...																																								
b125	A	-	-	C	-	-	C	-	-	C	-	-	-	-	-	-	-	-	-	-	A	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G

Tableau XII : description des haplotypes de *Dys44*. La première ligne donne l'haplotype ancestral, tel que trouvé chez d'autres primates (chimpanzés et Orang-Outang). Les haplotypes sont présentés en comparaison à l'haplotype ancestral : seuls les changements par rapport à ce dernier sont indiqués.
8N signifie 8 nucléotides et « DUP » signifie duplication.

Echantillon	Population	B001	B002	B003	B004	B005	B006	Autres
36	Crees	0.33333	0	0.13889	0.27778	0	0.13889	0.11111
52	Ojibwas	0.38462	0.01923	0.21154	0.05769	0	0.19231	0.13462
24	Navajos	0.41667	0.04167	0.04167	0.20833	0	0.29167	0
75	Chipewyans	0.44	0	0.26667	0.04	0	0.22667	0.02667
33	Algonquins	0.24242	0	0.0303	0.18182	0	0.51515	0.0303
220	Amérindiens du Nord	0.37727	0.00909	0.17273	0.12273	0	0.25455	0.06364
80	Mayas	0.275	0	0.1125	0.3125	0	0.2	0.1
40	Aches	0.825	0	0	0.175	0	0	0
18	Ticunas	0.27778	0	0.11111	0.38889	0	0.05556	0.16667
83	Karitanas	0.13253	0	0	0.66265	0	0.18072	0.0241
141	Amérindiens du Sud	0.34752	0	0.01418	0.48936	0	0.11348	0.03546
441	Amérindiens (total)	0.34921	0.00454	0.11111	0.27438	0	0.19955	0.06122

Tableau XIII : fréquence des haplotypes dans les différentes populations amérindiennes.

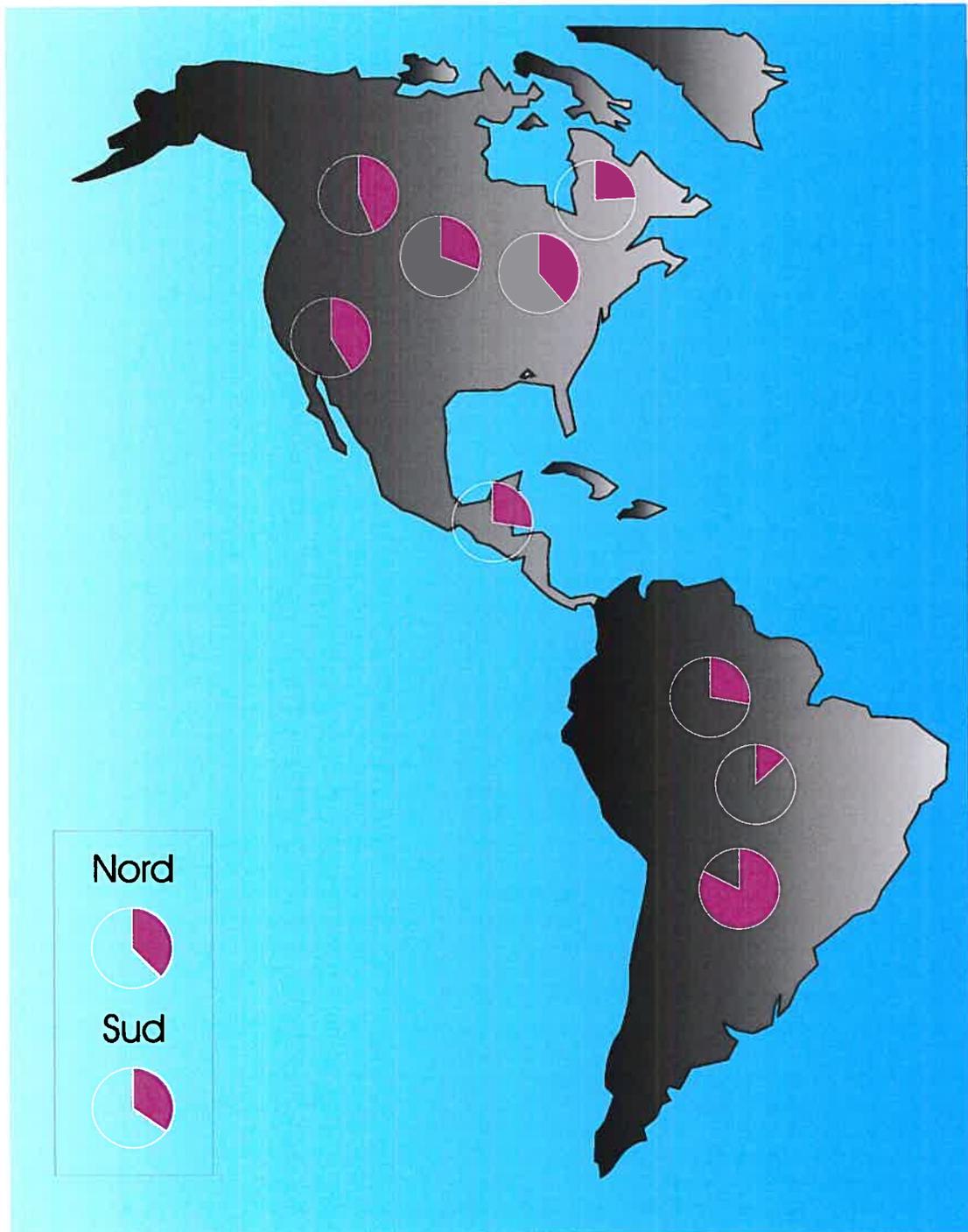


Figure 13 : Répartition de l'haplotype B001 chez les populations amérindiennes.

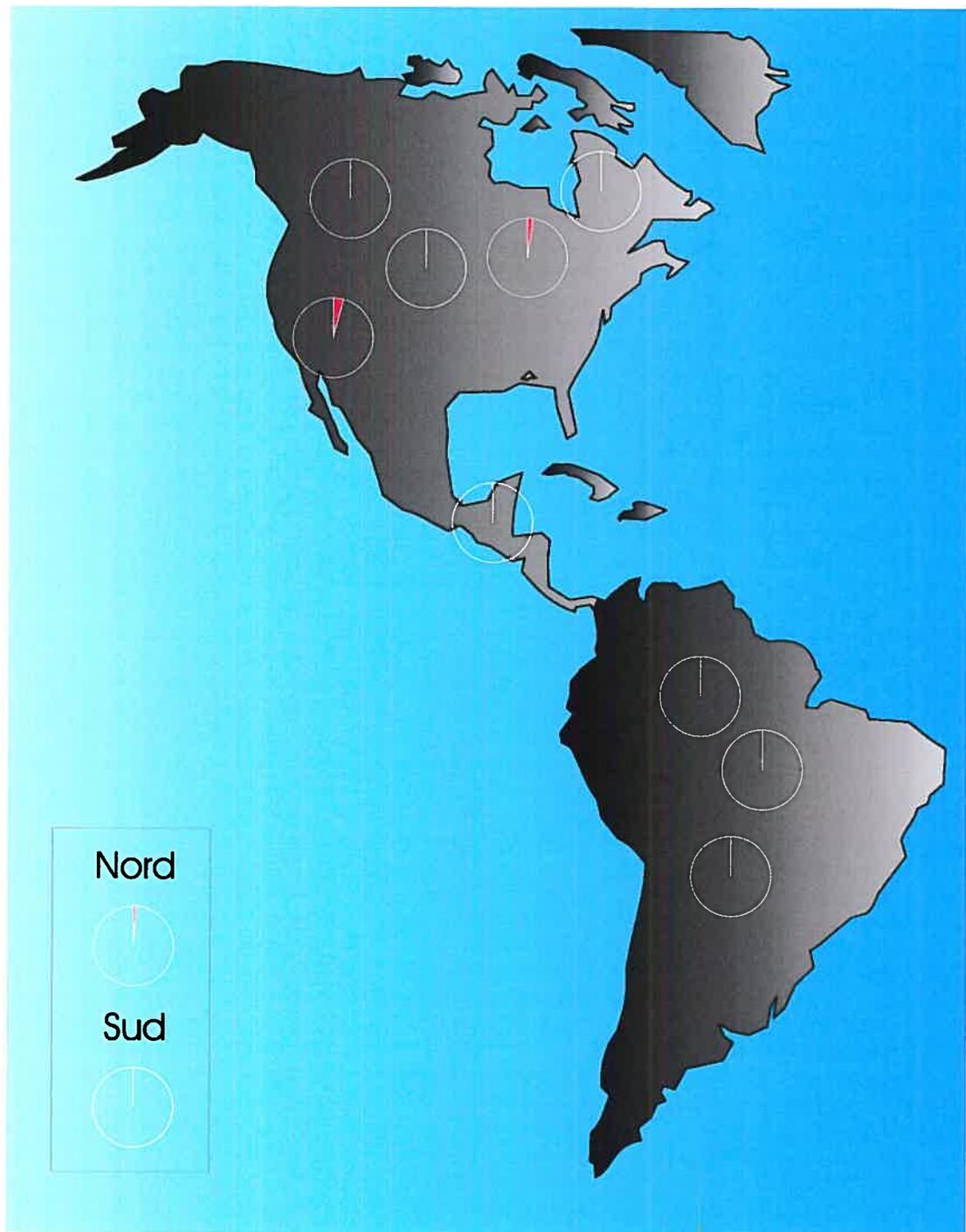


Figure 14 : Répartition de l'haplotype B002 chez les populations amérindiennes.

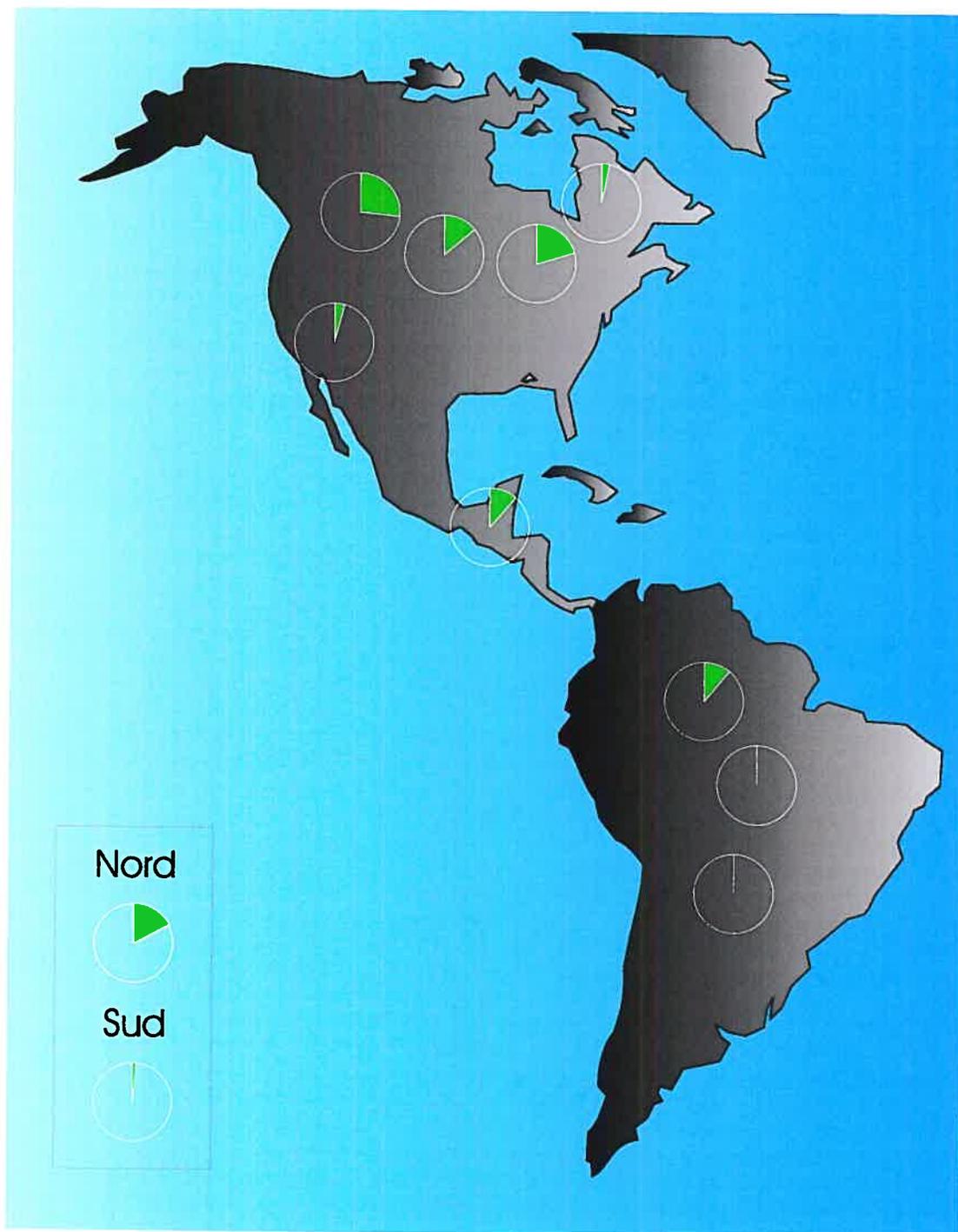


Figure 15 : Répartition de l'haplotype B003 chez les populations amérindiennes.

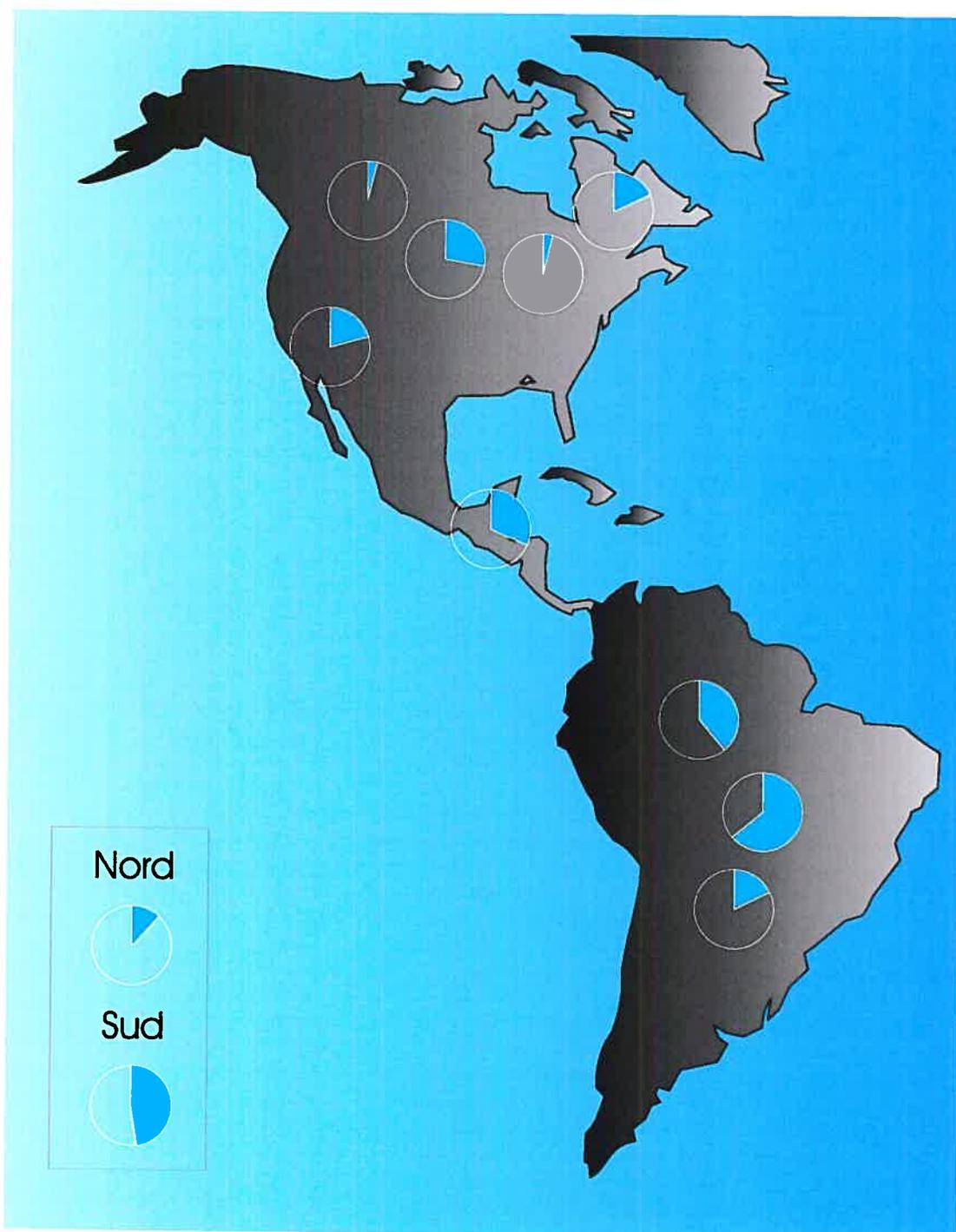


Figure 16 : Répartition de l'haplotype B004 chez les populations amérindiennes.

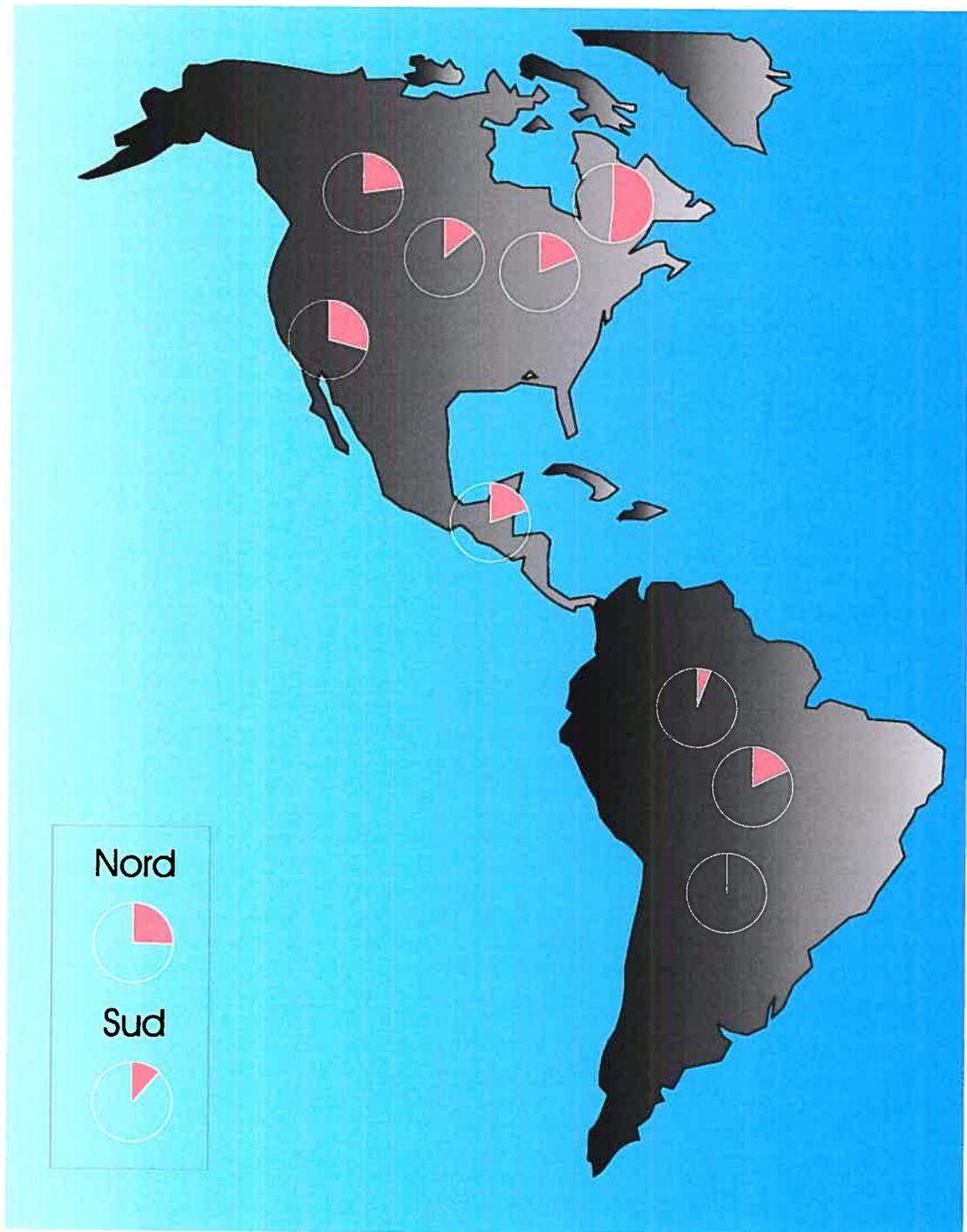


Figure 17 : Répartition de l'haplotype B006 chez les populations amérindiennes.

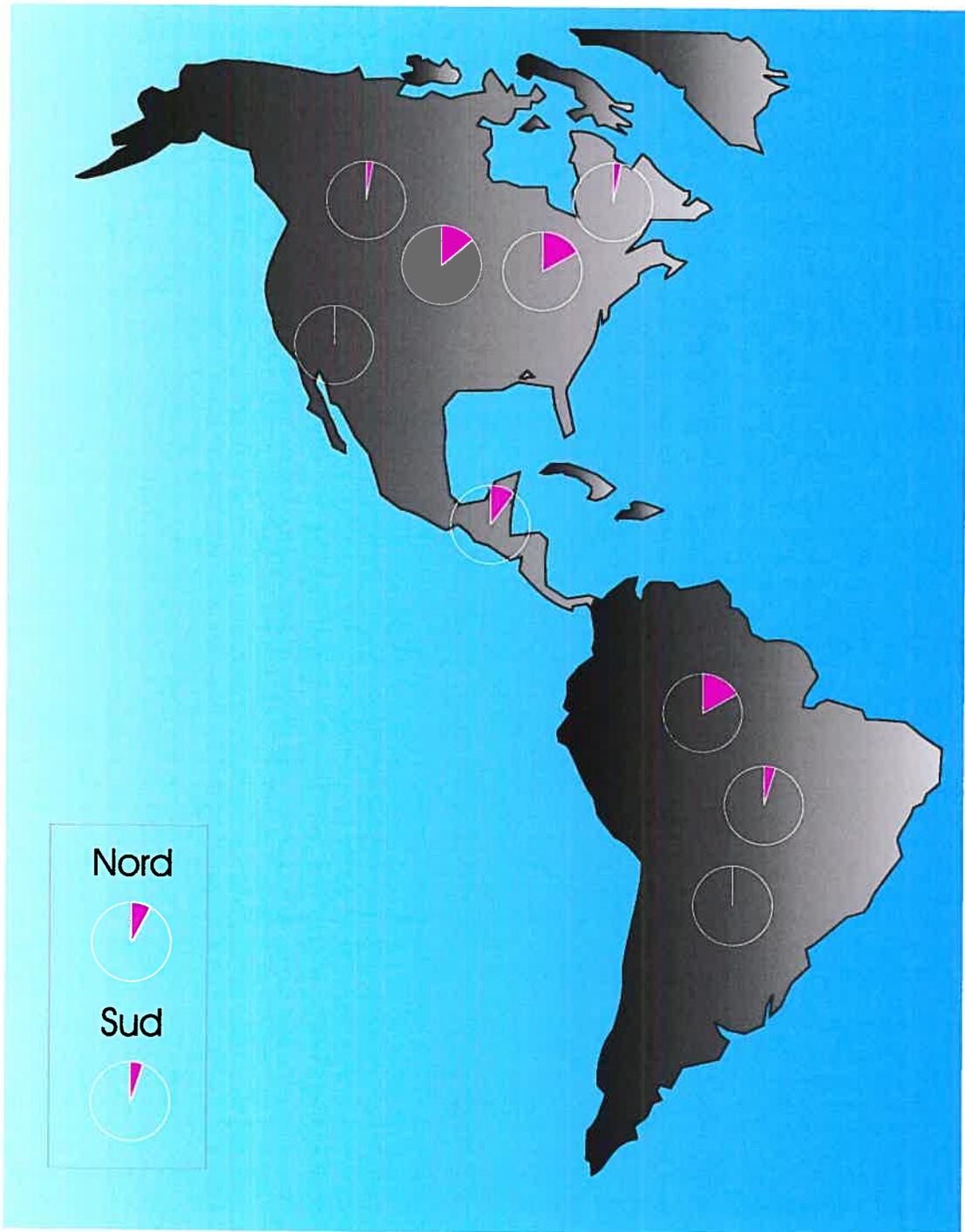


Figure 18 : Répartition des autres haplotypes chez les populations amérindiennes.

	B001	B002	B003	B004	B005	B006	Autres
Amérindiens du Nord	83	2	38	27	0	56	14
Amérindiens du Sud	49	0	2	69	0	16	5
<i>p value</i>	0.64829	0.25756	1.01E-05	4.38E-11	NA	0.00341	0.25493

Tableau XIV : Comparaison des populations amérindiennes d'Amérique du Nord et du Sud par le test du χ^2

	B001	B002	B003	B004	B005	B006	Autres
Amérindiens du Nord	83	2	38	27	0	56	14
Amérindiens du Centre	22	0	9	25	0	16	8
<i>p value</i>	0.18547	0.39377	0.24383	0.00048	NA	0.39377	0.30371

Tableau XV : Comparaison des populations amérindiennes d'Amérique du Nord et Centrale par le test du χ^2

	B001	B002	B003	B004	B005	B006	Autres
Amérindiens du Centre	22	0	9	25	0	16	8
Amérindiens du Sud	49	0	2	69	0	16	5
<i>p value</i>	0.36069	NA	0.00164	0.05269	NA	0.10427	0.05729

Tableau XVI : Comparaison des populations amérindiennes d'Amérique Centrale et du Sud par le test du χ^2

Matrice 1	Algonquins	Crees	Aches	Karitianas	Navajos	Ojibwas	Ticunas	Chipewyans	Mayas
Algonquins	0	2270	8528	6650	4243	773	5825	2613	4175
Crees	2270	0	9766	7858	2179	1643	7189	1035	4073
Aches	8528	9766	0	1918	9761	8563	2703	10696	6058
Karitianas	6650	7858	1918	0	8041	6649	888	8778	4330
Navajos	4243	2179	9761	8041	0	3490	7623	2792	3718
Ojibwas	773	1643	8563	6649	3490	0	5882	2236	3652
Ticunas	5825	7189	2703	888	7623	5882	0	8061	3998
Chipewyans	2613	1035	10696	8778	2792	2236	8061	0	5107
Mayas	4175	4073	6058	4330	3718	3652	3998	5107	0

Matrice 2	Algonquins	Crees	Aches	Karitianas	Navajos	Ojibwas	Ticunas	Chipewyans	Mayas
Algonquins	0	1	4	4	4	1	4	4	4
Crees	1	0	4	4	4	1	4	4	4
Aches	4	4	0	3	4	4	4	4	4
Karitianas	4	4	3	0	4	4	4	4	4
Navajos	4	4	4	4	0	4	4	2	4
Ojibwas	1	1	4	4	4	0	4	4	4
Ticunas	4	4	4	4	4	4	0	4	4
Chipewyans	4	4	4	4	2	4	4	0	4
Mayas	4	4	4	4	4	4	4	4	0

Matrice 3	Algonquins	Crees	Aches	Karitianas	Navajos	Ojibwas	Ticunas	Chipewyans	Mayas
Algonquins	0	0.0853	0.3869	0.2304	0.0243	0.0887	0.134	0.1109	0.06
Crees	0.0853	0	0.2027	0.1384	-0.0004	0.0199	-0.0134	0.0457	-0.0091
Aches	0.3869	0.2027	0	0.4516	0.1989	0.1916	0.2769	0.1948	0.2212
Karitianas	0.2304	0.1384	0.4516	0	0.1954	0.272	0.0829	0.3111	0.0993
Navajos	0.0243	-0.0004	0.1989	0.1954	0	0.0139	0.0364	0.033	0.0043
Ojibwas	0.0887	0.0199	0.1916	0.272	0.0139	0	0.0672	-0.007	0.0426
Ticunas	0.134	-0.0134	0.2769	0.0829	0.0364	0.0672	0	0.1065	-0.0057
Chipewyans	0.1109	0.0457	0.1948	0.3111	0.033	-0.007	0.1065	0	0.0693
Mayas	0.06	-0.0091	0.2212	0.0993	0.0043	0.0426	-0.0057	0.0693	0

Coefficient de corrélation géographique/F_{ST}	0.399750
<i>p value</i>	0.0097
Coefficient de corrélation linguistique/F_{ST}	0.096500
<i>p value</i>	0.3546

Tableau XVII : Matrices utilisées pour le test de Mantel, et résultat du test. La matrice 1 donne une distance géographique en kilomètres, la matrice 2 une distance linguistique en fonction de la classification linguistique de la base de données et la matrice 3 est la matrice de F_{ST} telle que calculée par Arlequin à partir des fréquences haplotypiques.

B) Populations amérindiennes et Eurasie

La fréquence de chaque haplotype pour les différentes populations amérindiennes, asiatiques et européennes (tableau XVIII page 88), est représentée sous forme d'un camembert placé sur la carte du monde selon la localisation de chaque ethnie (figures 19 à 25, page 89 à 95). La comparaison des populations amérindiennes, avec les populations européennes et asiatiques, deux à deux, par le test du Chi carré (χ^2) est disponible dans les tableaux XIX et XX page 96.

L'haplotype B001 a une fréquence moyenne similaire en Amérique, en Europe et en Asie. Les variations notables entre les populations semblent limitées à quelques cas particuliers, comme les Aches, les Karitianas et les Guinéens (voir tableau XVIII page 88 et figure 19 page 89). Hormis le fait que les variations inter-populations permettent de distinguer trois populations particulières, la répartition de l'haplotype B001 n'est pas informative sur les relations existantes entre les Amérindiens, les Européens et les Asiatiques.

L'haplotype B002 est présent à très haute fréquence en Asie (32,4%), dans toutes les populations asiatiques (figure 20 page 90). La fréquence de cet haplotype est également importante en Europe (5,4%) en étant toutefois bien moindre qu'en Asie. Par contre, la fréquence très faible de cet haplotype en Amérique (0,5%) indique une forte dissension entre les populations amérindiennes et les populations eurasiatiques. La même observation peut être faite en observant la répartition de l'haplotype B005 (figure 23 page 93), dont la fréquence en Amérique est nulle alors qu'elle est assez élevée en Eurasie, à un niveau comparable en Europe de l'Ouest (6,9%) et en Asie (7,1%).

L'haplotype B003 est présent à très haute fréquence en Europe de l'Ouest (24%), à haute fréquence en Amérique (11%) et en Asie (5%) (figure 21 page 91). Les Amérindiens se situent entre les Européens de l'Ouest et les Asiatiques, mais restent significativement différents de ces deux groupes (tableaux XIX et XX page 96).

L'haplotype B004 marque également une grande différence entre les populations amérindiennes et eurasiatiques, mais cette fois avec une forte prévalence de cet haplotype en Amérique (figure 22 page 92). En effet, alors que cet haplotype est presque absent d'Europe de l'Ouest (0,2%) et présent à faible fréquence en Asie (2,4%) la fréquence en Amérique atteint 26,6% et l'haplotype est présent dans toutes les populations amérindiennes. En Asie, ce sont les Sibériens qui présentent la plus haute fréquence de l'haplotype B004 (7,2%), se distinguant nettement des autres populations asiatiques, mais demeurant à une fréquence bien plus faible que dans les populations amérindiennes.

L'haplotype B006 est également présent à très haute fréquence en Amérique (19,9%), alors qu'il est présent à haute fréquence en Europe de l'Ouest (11%) et à fréquence moyenne en Asie (3,8%) (figure 24 page 94). Une fois encore, avec une fréquence en B006 de 12,2%, les Sibériens se distinguent des autres populations asiatiques en arborant une fréquence en B006 comparable à celle des Européens de l'Ouest, se rapprochant ainsi des Amérindiens.

C'est chez les amérindiens que la fréquence des autres haplotypes, de 6,1%, est la plus faible, contre 10,2% chez les asiatiques (~~10,2%~~) et 18,8% chez les européens de l'ouest (figure 25 page 95 et tableau XX page 96). Cela nous amène à regarder la diversité génique, telle que définie par Nei en 1987, c'est à dire la probabilité, lorsque sont choisis au hasard deux chromosomes dans une population, d'obtenir deux chromosomes (haplotypes) différents (tableau XXI page 97).

Bien que la fréquence des autres haplotypes soit significativement supérieure en Asie par rapport à l'Amérique, l'Asie montre une diversité génétique légèrement inférieure à celle trouvée en Amérique (0,73 contre 0,75 respectivement). C'est en Europe de l'Ouest que la diversité génétique est la plus élevée (0,81). Mais une fois encore, la Sibérie se montre plus proche de l'Europe de l'Ouest que de l'Asie, avec une diversité génétique comparable à celle trouvée en Europe (0,81). Deux populations d'Amérique du Sud, les Aches et les Karitianas, montrent une diversité génétique particulièrement basse (0.30 et 0.55 respectivement).

Afin d'approfondir les relations existant entre les différentes populations, plusieurs analyses en composantes principales ont été effectuées à partir de la fréquence des haplotypes (figures 26 et 27 pages 98 et 99). Les valeurs de F_{ST} ont également été calculées (tableau XXII page 100), et plusieurs représentations graphiques de ces valeurs ont été réalisées par *Multi Dimensional Scaling* (MDS) (figures 28 à 30 pages 101 à 103).

Les PCA et les MDS concordent et placent les Européens de l'Ouest plus proches des Amérindiens que les Asiatiques. La décomposition de ces groupes ne montre pas de différence notable entre les Ibériens et le reste des Européens de l'Ouest. Cela met cependant en évidence d'importantes différences pour les populations Asiatiques. C'est une fois encore les Sibériens qui sont les plus proches des Européens de l'Ouest et des Amérindiens. Lorsque les Amérindiens sont divisés selon le Nord, le Centre et le Sud, les Amérindiens du Nord apparaissent comme étant les plus proches des Européens de l'Ouest et des Sibériens, les Amérindiens du Centre et du Sud étant plus en retrait.

Echantillon	Population	B001	B002	B003	B004	B005	B006	Autres
30	Mossi	0.13333	0.16667	0	0	0.06667	0	0.63333
31	Rimaibe	0.09677	0.19355	0.03226	0.03226	0.03226	0.03226	0.58065
85	Biaka	0.21176	0.36471	0.01176	0	0.04706	0	0.36471
58	MButi	0.08621	0.12069	0.05172	0	0	0	0.74138
204	Africains	0.14706	0.2402	0.02451	0.0049	0.03431	0.0049	0.54412
53	Ibériens	0.30189	0.09434	0.16981	0.01887	0.09434	0.13208	0.18868
410	Européens de l'Ouest	0.34146	0.04878	0.24878	0	0.06585	0.10732	0.1878
463	Européens de l'Ouest (total)	0.33693	0.054	0.23974	0.00216	0.06911	0.11015	0.1879
67	Guinéens	0.07463	0.56716	0.01493	0.01493	0.0597	0	0.26866
65	Japonais	0.6	0.27692	0.01538	0	0.03077	0	0.07692
96	Chinois	0.45833	0.35417	0.05208	0.01042	0.0625	0.02083	0.04167
91	Asiatiques du Sud-Est	0.36264	0.42857	0.04396	0.02198	0.01099	0	0.13187
53	Sibériens	0.35849	0.20755	0.11321	0.03774	0.09434	0.11321	0.07547
296	Mongols	0.40878	0.26351	0.05743	0.03041	0.09797	0.05743	0.08446
668	Asiatiques	0.39072	0.32635	0.0509	0.02246	0.07036	0.03743	0.1018
36	Crees	0.33333	0	0.13889	0.27778	0	0.13889	0.11111
52	Ojibwas	0.38462	0.01923	0.21154	0.05769	0	0.19231	0.13462
24	Navajos	0.41667	0.04167	0.04167	0.20833	0	0.29167	0
75	Chipewyans	0.44	0	0.26667	0.04	0	0.22667	0.02667
33	Algonquins	0.24242	0	0.0303	0.18182	0	0.51515	0.0303
220	Amérindiens du Nord	0.37727	0.00909	0.17273	0.12273	0	0.25455	0.06364
80	Mayas	0.275	0	0.1125	0.3125	0	0.2	0.1
40	Aches	0.825	0	0	0.175	0	0	0
18	Ticunas	0.27778	0	0.11111	0.38889	0	0.05556	0.16667
83	Karitianas	0.13253	0	0	0.66265	0	0.18072	0.0241
141	Amérindiens du Sud	0.34752	0	0.01418	0.48936	0	0.11348	0.03546
441	Amérindiens (total)	0.34921	0.00454	0.11111	0.27438	0	0.19955	0.06122

Tableau XVIII : fréquence des haplotypes dans les différentes populations amérindiennes, Européennes, Africaines et Asiatiques.

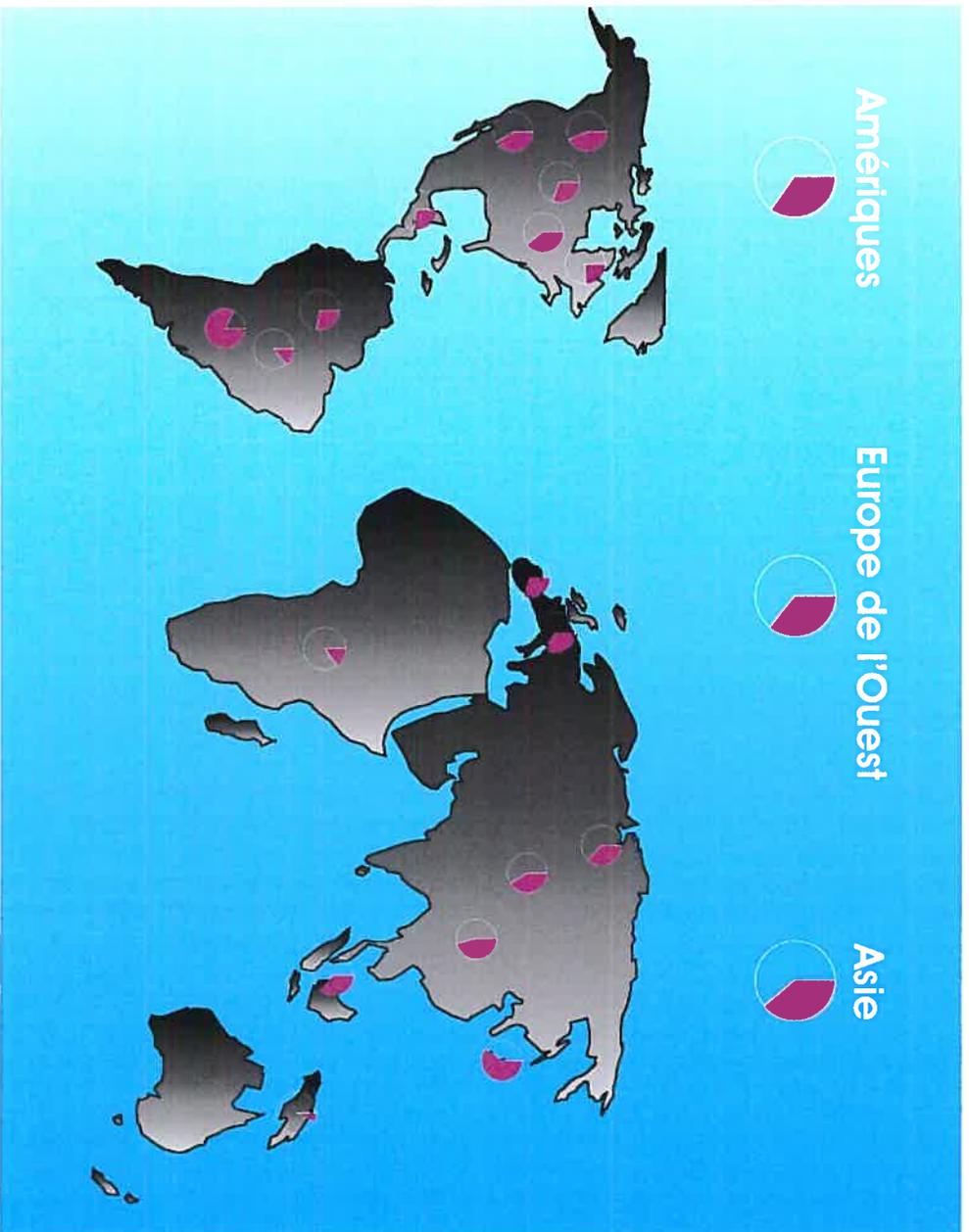


Figure 19 : Répartition de l'haplotype B001 chez les populations amérindiennes, européennes, africaines et asiatiques.

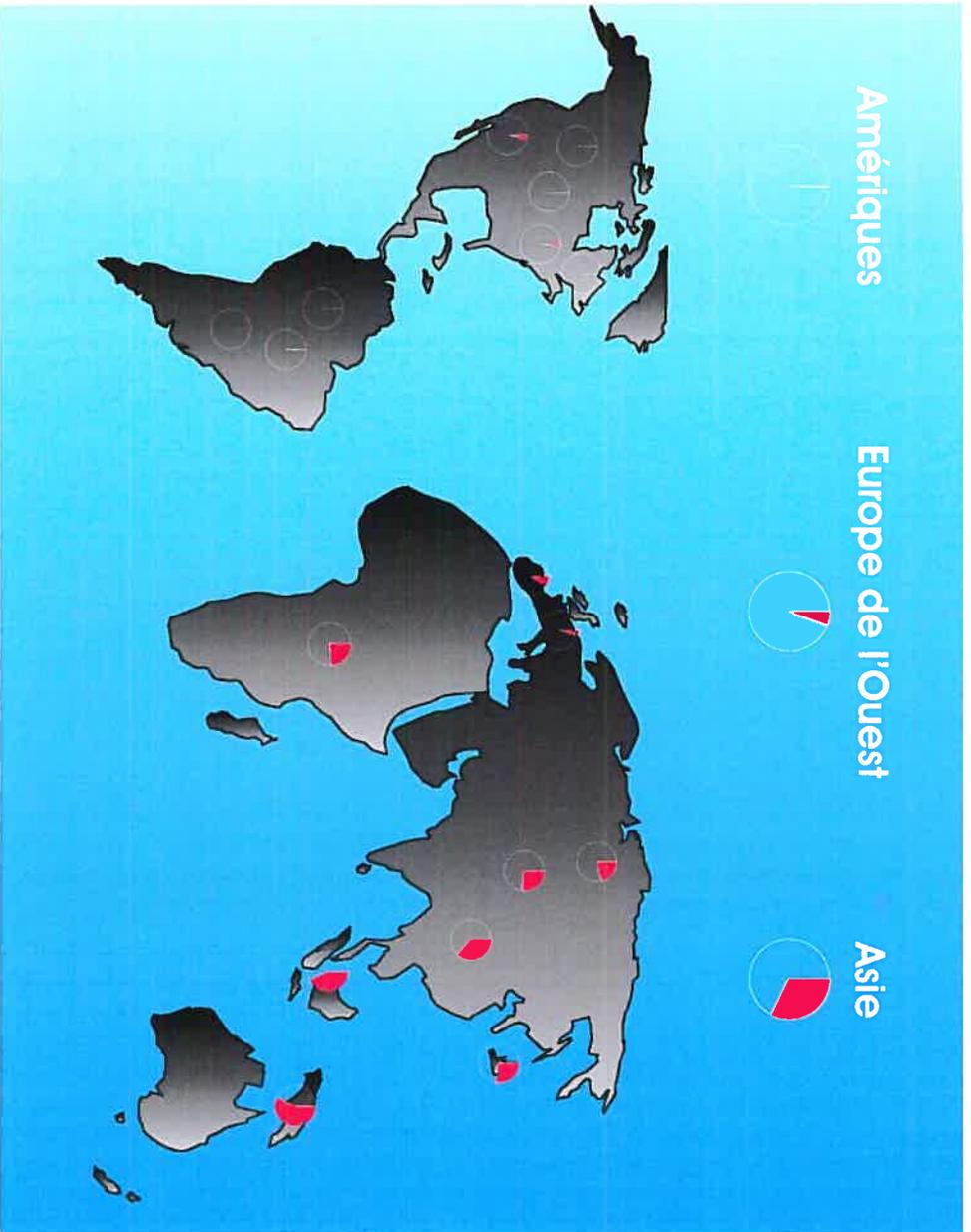


Figure 20 : Répartition de l'haplotype B002 chez les populations amérindiennes, européennes, africaines et asiatiques.

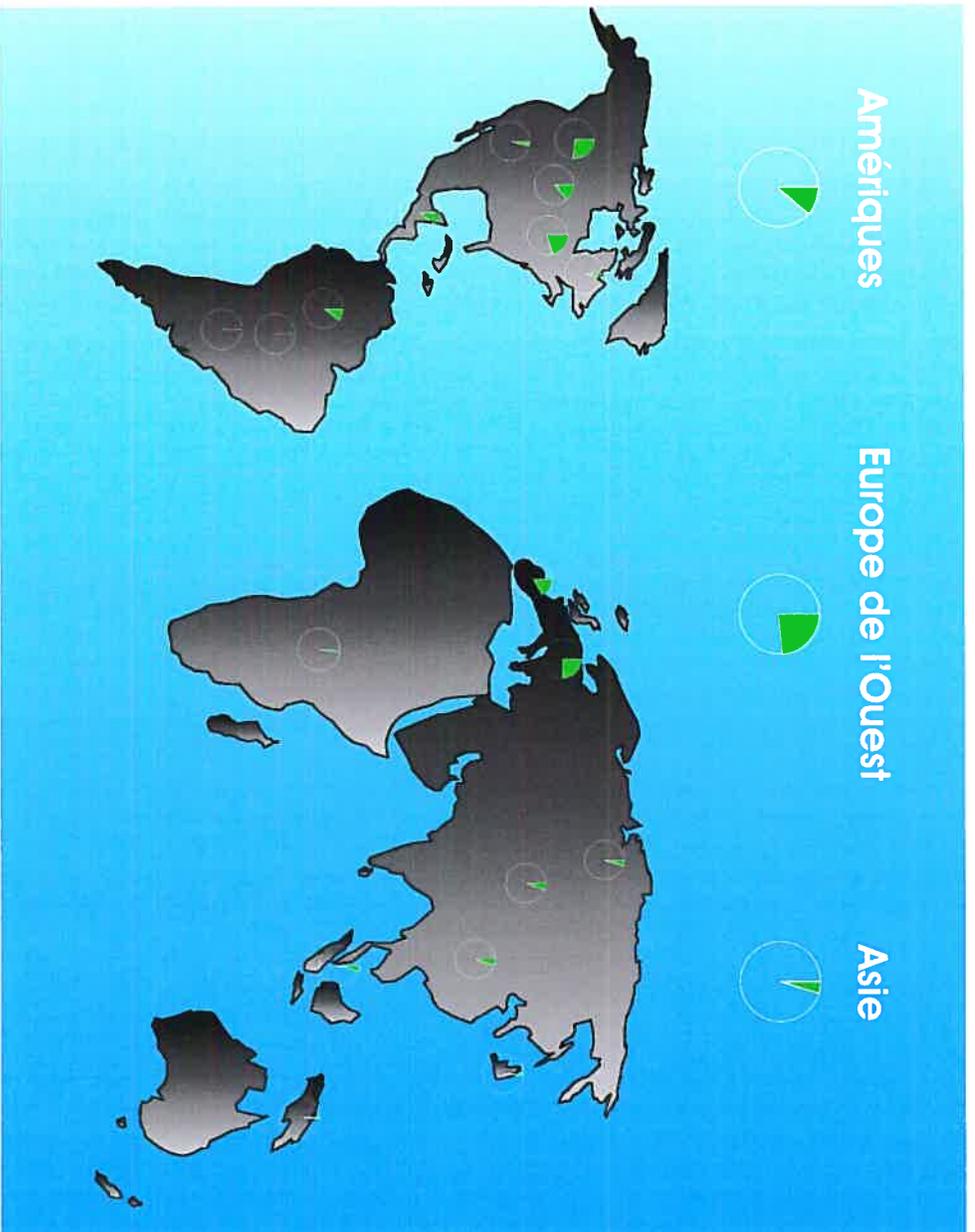


Figure 21 : Répartition de l'haplotype B003 chez les populations amérindiennes, européennes, africaines et asiatiques

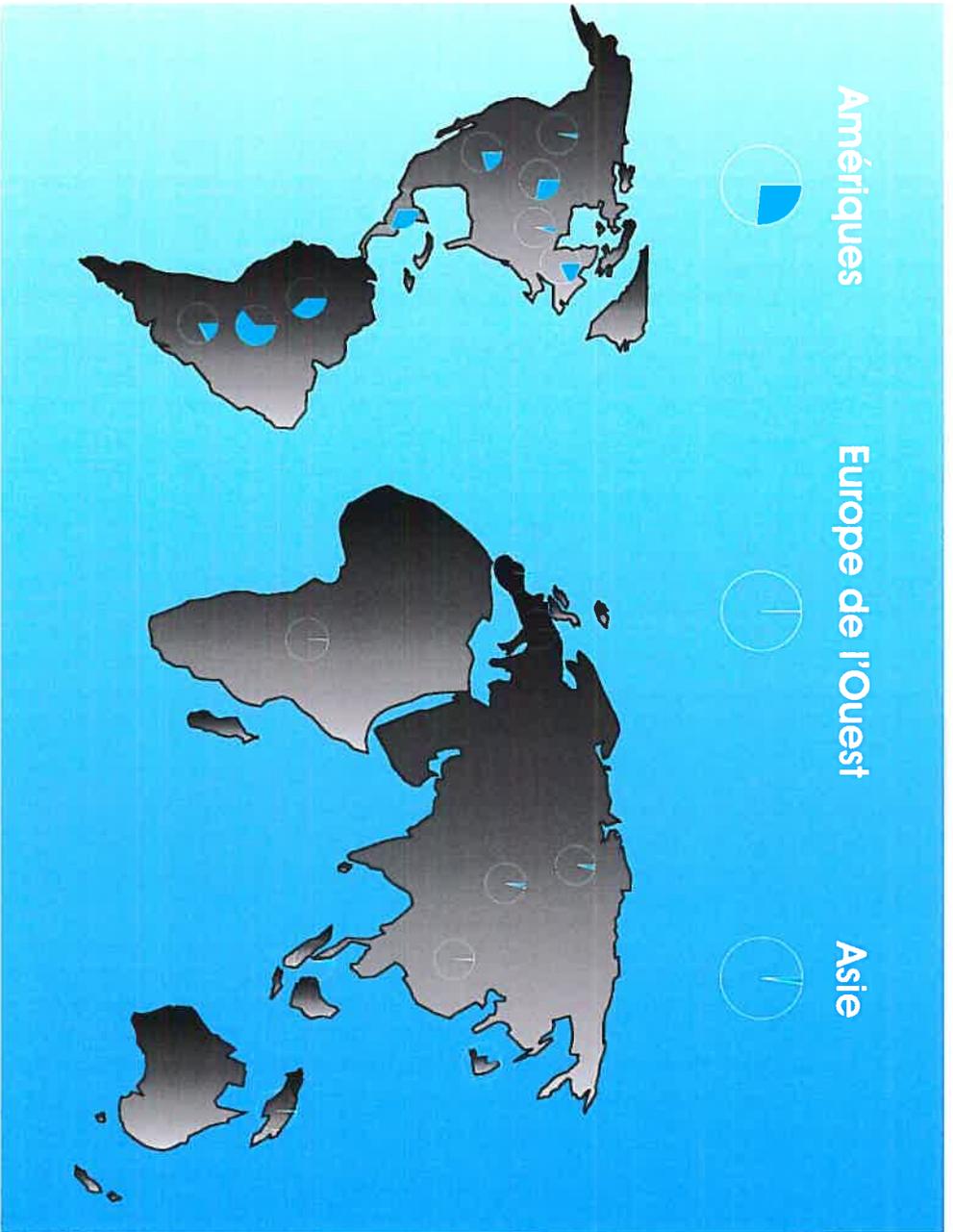


Figure 22 : Répartition de l'haplotype B004 chez les populations amérindiennes, européennes, africaines et asiatiques.

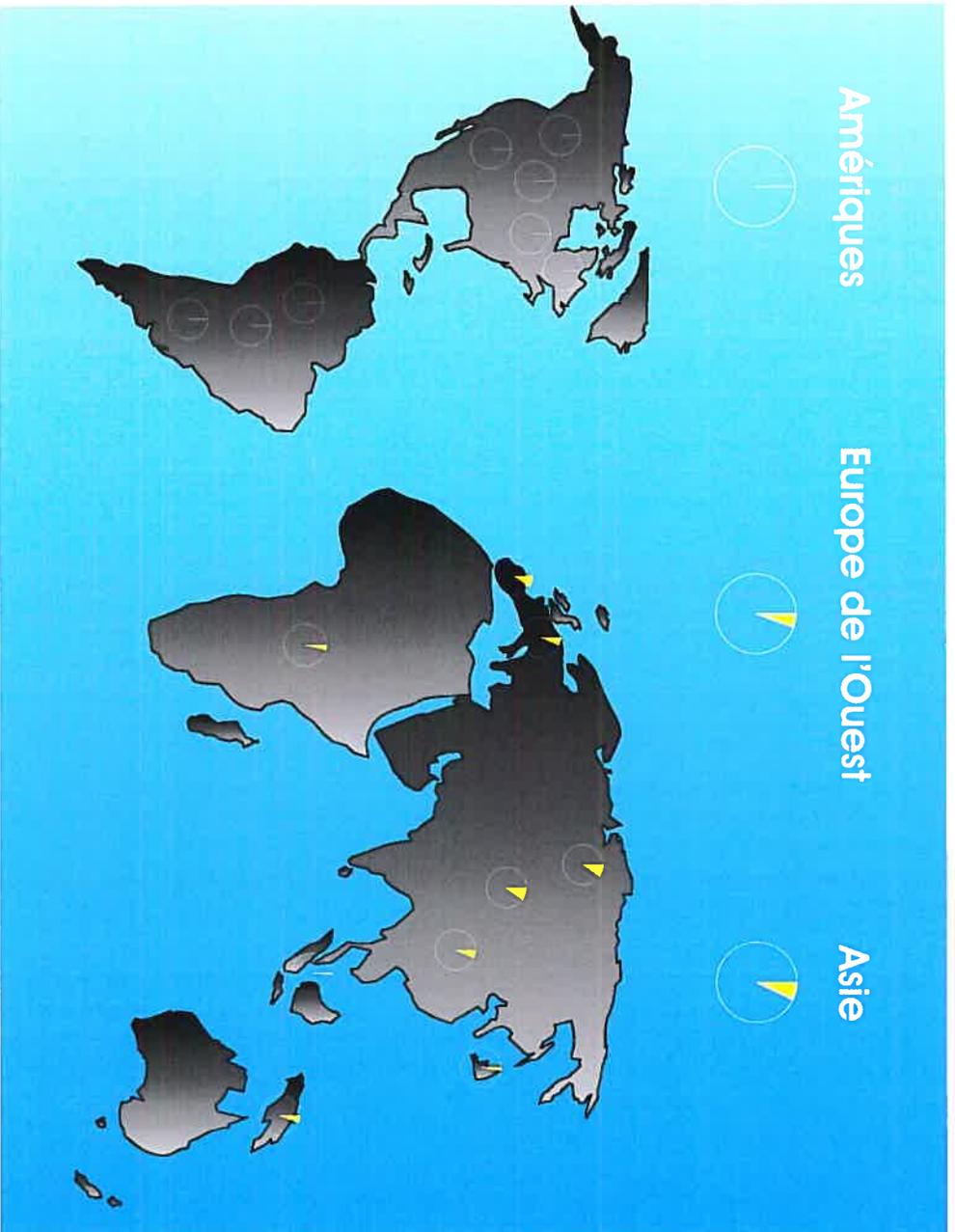


Figure 23 : Répartition de l'haplotype B005 chez les populations amérindiennes, européennes, africaines et asiatiques.

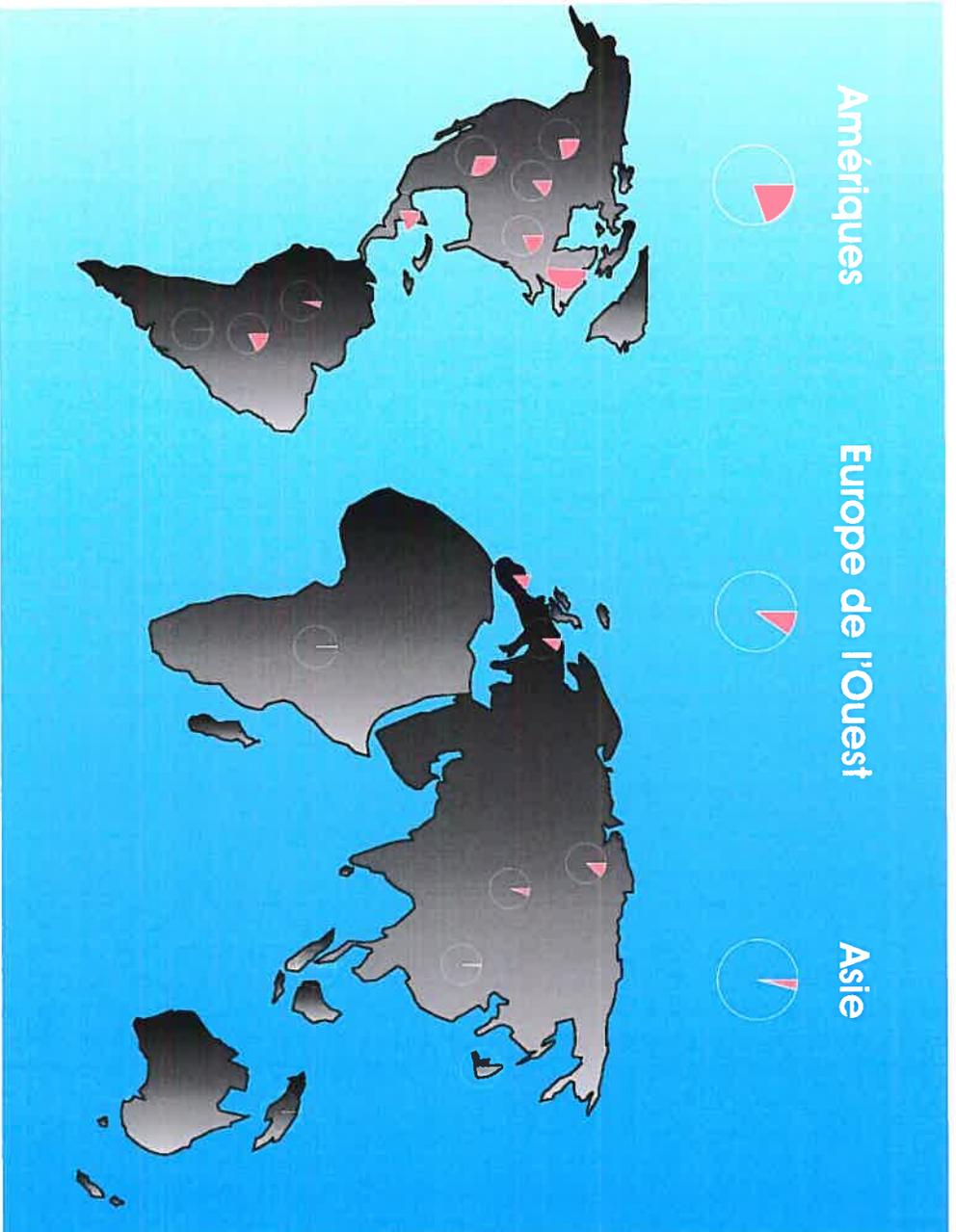


Figure 24 : Répartition de l'haplotype B006 chez les populations amérindiennes, européennes, africaines et asiatiques.

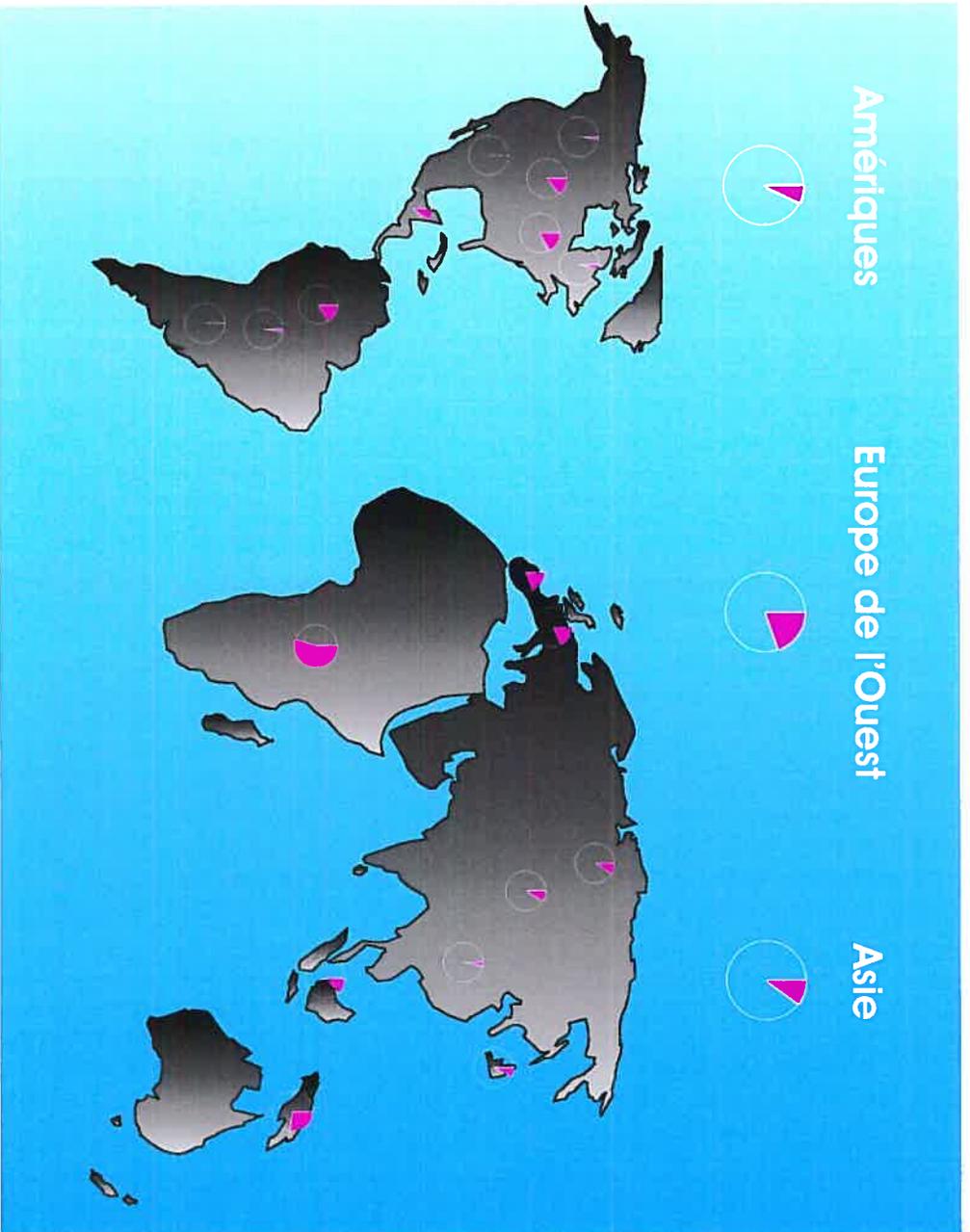


Figure 25 : Répartition des autres haplotypes chez les populations amériindiennes, européennes, africaines et asiatiques.

	B001	B002	B003	B004	B005	B006	Autres
Amérindiens	154	2	49	121	0	88	27
Européens de l'Ouest	156	25	111	1	32	51	87
<i>p value</i>	0.75277	0.00002	4.33E-06	8.34E-29	3.37E-08	0.00061	8.26E-08

Tableau XIX : Comparaison des populations amérindiennes et européennes par le test du χ^2

	B001	B002	B003	B004	B005	B006	Autres
Amérindiens	154	2	49	121	0	88	27
Asiatiques	261	218	34	15	47	25	68
<i>p value</i>	0.26872	5.19E-32	0.00033	9.52E-32	2.54E-08	1.26E-16	0.02387

Tableau XX : Comparaison des populations amérindiennes et asiatiques par le test du χ^2

Population	Diversité génétique
Africains	0.8890 +/- 0.0127
Ibériens	0.8527 +/- 0.0298
Européens de l'Ouest*	0.8010 +/- 0.0128
Européens de l'Ouest	0.8065 +/- 0.0118
Chinois	0.6636 +/- 0.0317
Japonais	0.5688 +/- 0.0526
Guinéens	0.6653 +/- 0.0627
Asiatiques du Sud-Est	0.6867 +/- 0.0329
Sibériens	0.8055 +/- 0.0350
Mongols	0.7476 +/- 0.0177
Asiatiques	0.7319 +/- 0.0116
Crees	0.7889 +/- 0.0367
Ojibwas	0.7768 +/- 0.0378
Navajos	0.7246 +/- 0.0533
Chipewyans	0.6908 +/- 0.0290
Algonquins	0.6610 +/- 0.0611
Amérindiens du Nord	0.7507 +/- 0.0158
Aches	0.2962 +/- 0.0788
Ticunas	0.7843 +/- 0.0680
Karitianas	0.5166 +/- 0.0539
Amérindiens du Sud	0.6202 +/- 0.0244
Mayas/Amérindiens du Centre	0.7816 +/- 0.0221
Amérindiens	0.7520 +/- 0.0097

Tableau XXI : Diversité génétique des différentes populations africaines, amérindiennes, asiatiques et européennes, telle que calculée par Arlequin.

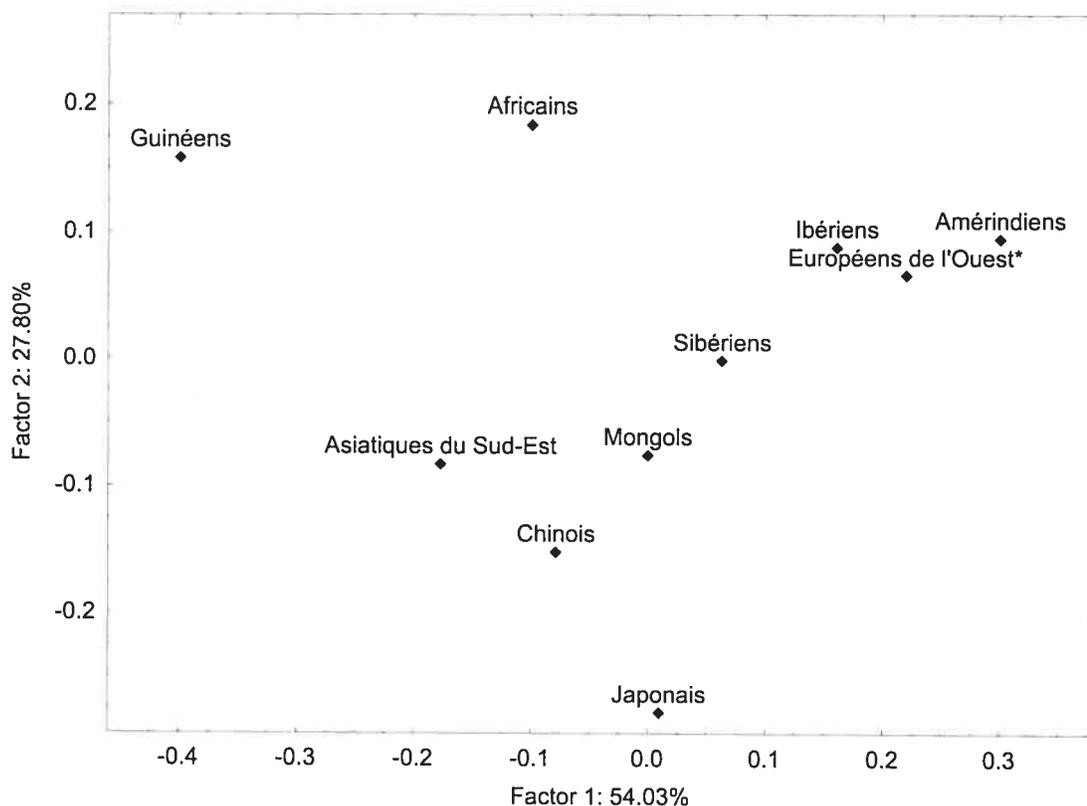


Figure 26 : Analyse en composantes principales (PCA) des populations amérindiennes, européennes et asiatiques. Les populations asiatiques sont présentées individuellement, et les Ibériens sont présentés à part du reste des Européens de l'Ouest (signifié par l'astérisque). Il est à noter que l'importance des deux axes est inégale, et est exprimée en fonction (pourcentage) de la variance totale reconstituée sur un axe. L'addition de la variance reconstituée par chaque axe donne la variance totale maximale qu'il a été possible de représenter en deux dimensions.

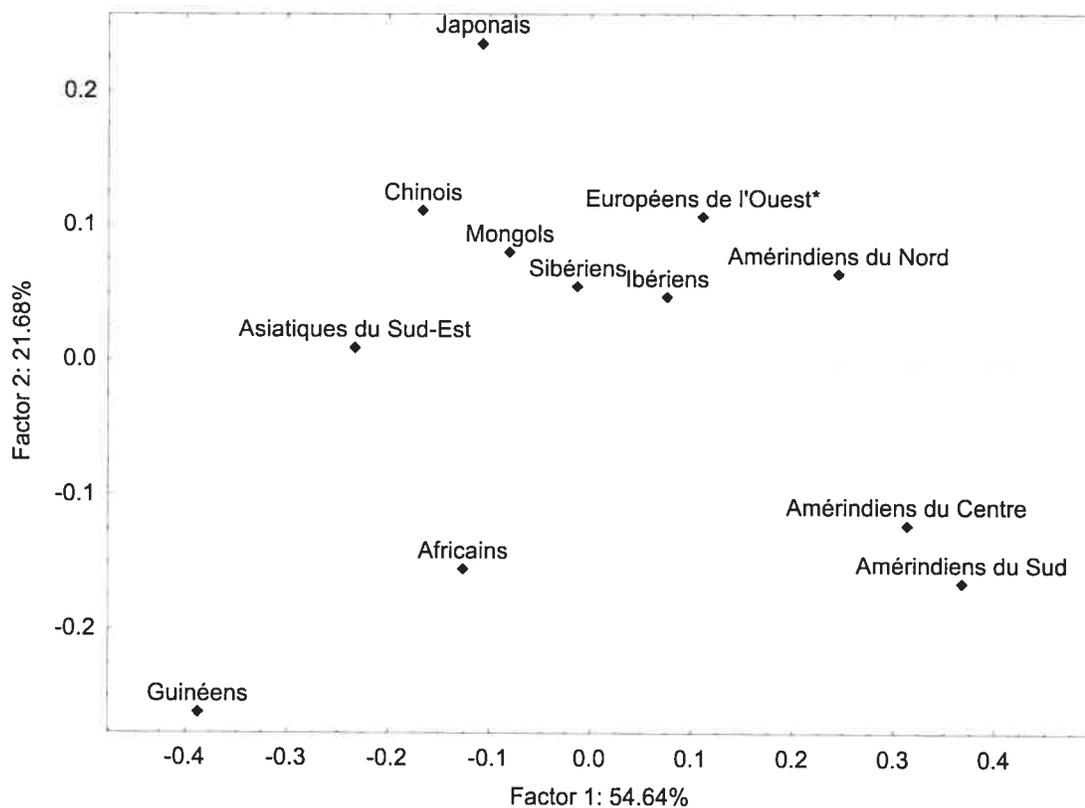


Figure 27: Analyse en composantes principales (PCA) des populations amérindiennes, européennes et asiatiques. Les populations asiatiques sont présentées individuellement, et les Ibériens sont présentés à part du reste des Européens de l'Ouest (signifié par l'astérisque). Les populations amérindiennes sont présentées en trois groupes : les Amérindiens du Nord, du Centre et du Sud.

	Européens de l'Ouest	Chinois	Japonais	Guinéens	Asiatiques du Sud-Est	Sibériens	Mongoles	Asiatiques	Amérindiens	Africains
Européens de l'Ouest	0	0.08469	0.10216	0.20276	0.10944	0.01866	0.05202	0.07316	0.06358	0.08866
Chinois	0.08469	0	0.0123	0.12605	0.0049	0.01749	0.00341	-0.00152	0.13748	0.07583
Japonais	0.10216	0.0123	0	0.2247	0.05167	0.05073	0.02531	0.02791	0.14733	0.12239
Guinéens	0.20276	0.12605	0.2247	0	0.07025	0.13401	0.12633	0.09922	0.25704	0.07943
Asiatiques du Sud-Est	0.10944	0.0049	0.05167	0.07025	0	0.03687	0.02281	0.00658	0.15958	0.06089
Sibériens	0.01866	0.01749	0.05073	0.13401	0.03687	0	0.02281	0.00814	0.15958	0.06089
Mongoles	0.05202	0.00341	0.02531	0.12633	0.02281	-0.00168	0	0.00173	0.09622	0.06193
Asiatiques	0.07316	-0.00152	0.02791	0.09922	0.00658	0.00814	0.00173	0	0.11986	0.06111
Amérindiens	0.06358	0.13748	0.14733	0.25704	0.15958	0.06193	0.09622	0.11986	0	0.13311
Africains	0.08866	0.07583	0.12239	0.07943	0.06089	0.04684	0.06193	0.06111	0.13311	0

Tableau XXII : Valeurs des Fst telles que calculées par Arlequin, à partir de la fréquence des haplotypes.

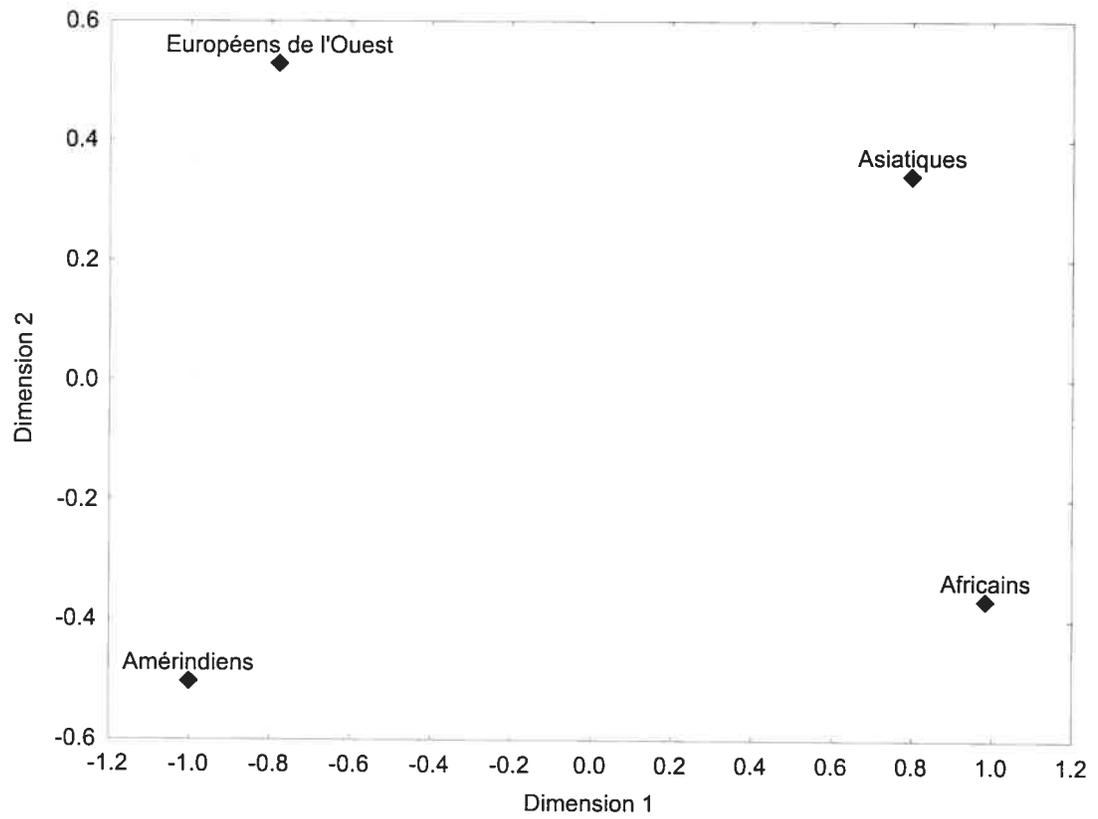


Figure 28 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Asiatiques et les Européens de l'Ouest.

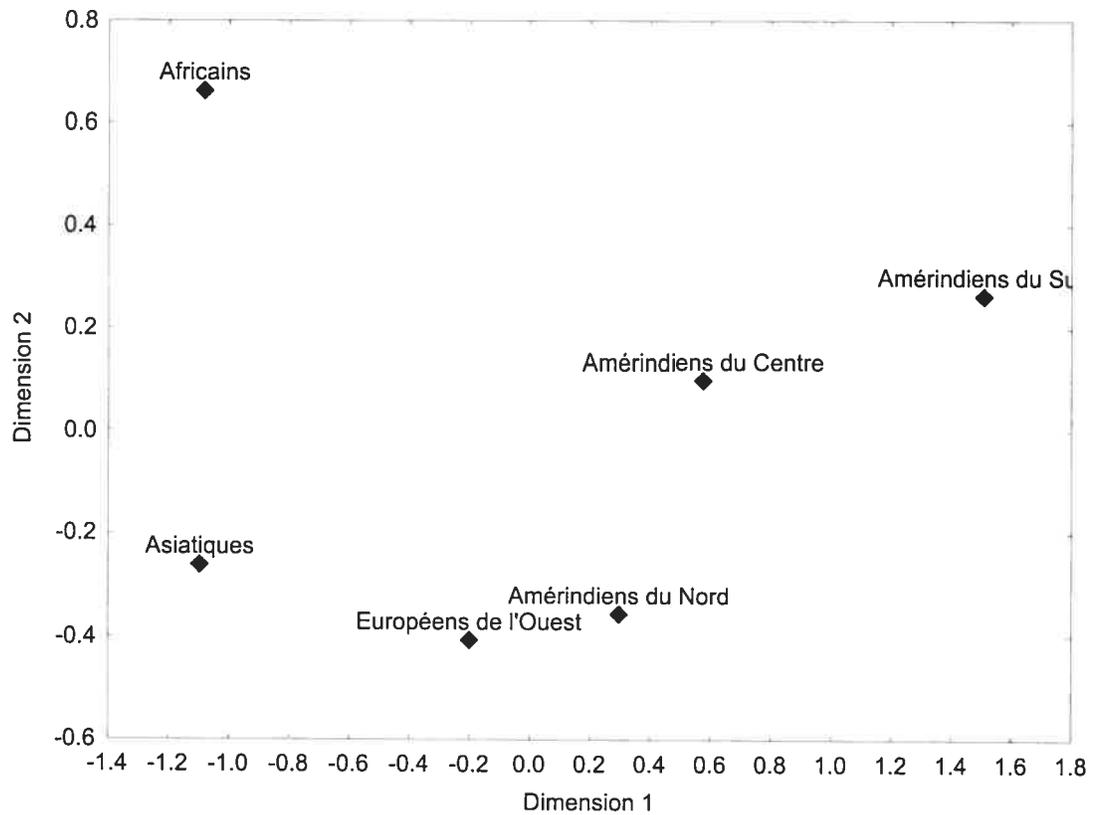


Figure 29 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Asiatiques et les Européens de l'Ouest. Les Amérindiens ont été décomposés en Amérindiens du Nord, du Centre et du Sud.

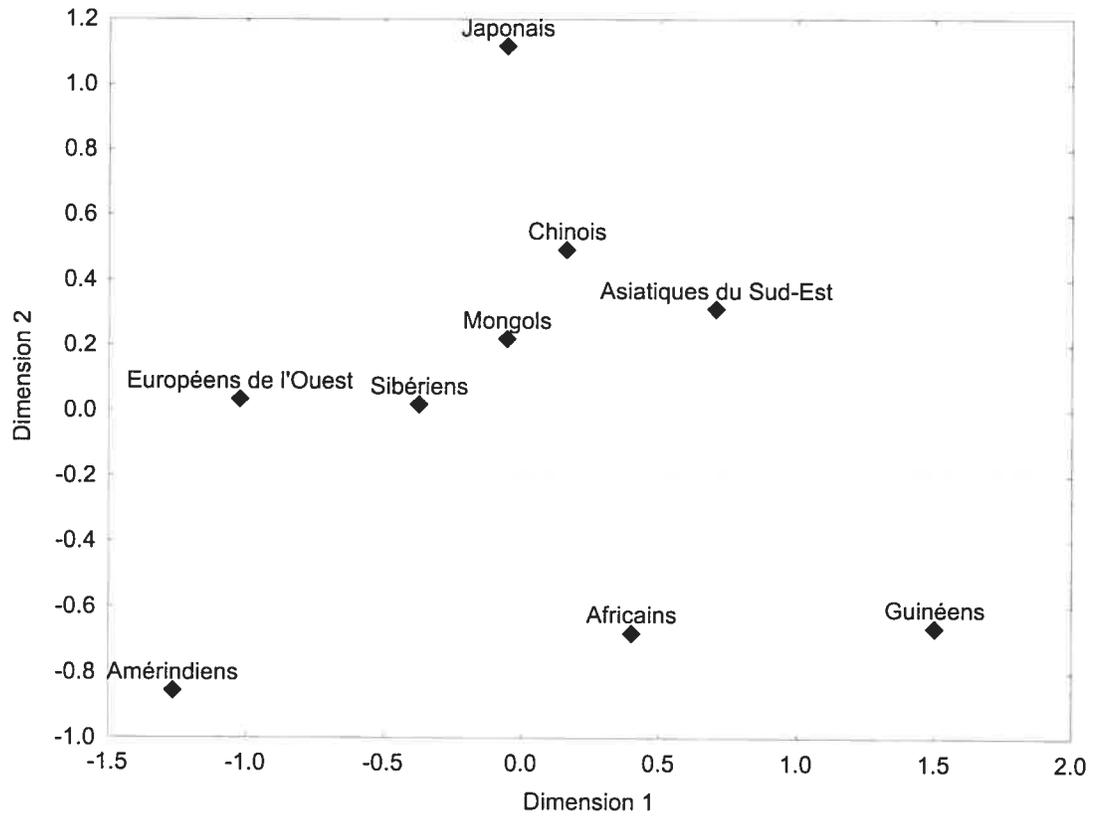


Figure 30 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Asiatiques et les Européens de l'Ouest.

C) Influence récente des Européens

La fréquence des haplotypes pour les différentes populations amérindiennes, métisses amérindiennes, asiatiques et européennes (tableau XXIII page 106), est représentée sous forme d'un camembert placé sur la carte de l'Amérique et du monde selon la localisation de chaque ethnie (figures 31 et 32, pages 107 et 108). La comparaison des populations métisses amérindiennes, avec les populations amérindiennes, européennes et asiatiques, deux à deux, par le test du Chi carré (χ^2) est disponible dans les tableaux XXIV à XXVI page 109.

Il y a de nombreuses différences concernant les fréquences haplotypiques entre les populations métisses amérindiennes et amérindiennes. B002, qui était presque absent chez les Amérindiens (0,5%), est retrouvé avec une fréquence assez élevée chez les Métisses Amérindiens (7%). De même, l'haplotype B005 est retrouvé à une fréquence de 2,6% chez les Métisses Amérindiens, alors qu'il était complètement absent chez les Amérindiens. Alors que la fréquence des haplotypes B004 et B006 chute chez les Métisses Amérindiens (10,5% et 9,6% respectivement, contre 26,6% et 19,9% chez les Amérindiens), la fréquence des autres haplotypes (14,9%) monte presque au niveau de celle trouvée chez les Européens de l'Ouest (18,8%).

Les tests du khi carré (tableaux XXIV à XXVI page 109), ainsi que les valeurs de F_{ST} (tableau XXVIII page 112), et leur représentation en MDS (figure 34 page 113), ainsi que les analyses en composantes principales (figure 33 page 111), concordent et montrent que les populations métisses amérindiennes sont plus proches des populations européennes de l'Ouest que des populations amérindiennes. Toutefois ces populations métisses restent plus proches des populations amérindiennes que des populations asiatiques. De même, la diversité génétique est plus importante chez les Métisses Amérindiens que chez les Amérindiens (tableau XXVII page 110), se rapprochant de celle trouvée chez les Européens de l'Ouest.

Curieusement, chez les Antioquiens, la diversité génétique est faible par rapport à celle trouvée chez les autres populations métisses amérindiennes. La population d'Antioquia diffère également des autres populations métisses amérindiennes par l'absence de l'haplotype B002. Dans les autres haplotypes trouvés dans cette population, un haplotype très particulier se distingue des autres, présentant un polymorphisme typique des populations africaines de la région sub-saharienne. Plusieurs de ces polymorphismes typiquement africains ont également été trouvés dans deux échantillons mexicains. Il est également important de noter qu'un haplotype complet et typiquement africain a été trouvé chez un individu espagnol de sexe masculin. Dans toutes les populations de cette étude, seuls les haplotypes mentionnés ci-avant présentaient des polymorphismes typiquement africains.

Echantillon	Population	B001	B002	B003	B004	B005	B006	Autres
30	Mossi	0.13333	0.16667	0	0	0.06667	0	0.63333
31	Rimaibe	0.09677	0.19355	0.03226	0.03226	0.03226	0.03226	0.58065
85	Biaka	0.21176	0.36471	0.01176	0	0.04706	0	0.36471
58	MButi	0.08621	0.12069	0.05172	0	0	0	0.74138
204	Africains	0.14706	0.2402	0.02451	0.0049	0.03431	0.0049	0.54412
53	Ibériens	0.30189	0.09434	0.16981	0.01887	0.09434	0.13208	0.18868
410	Européens de l'Ouest	0.34146	0.04878	0.24878	0	0.06585	0.10732	0.1878
463	Européens de l'Ouest (total)	0.33693	0.054	0.23974	0.00216	0.06911	0.11015	0.1879
67	Guinéens	0.07463	0.56716	0.01493	0.01493	0.0597	0	0.26866
65	Japonais	0.6	0.27692	0.01538	0	0.03077	0	0.07692
96	Chinois	0.45833	0.35417	0.05208	0.01042	0.0625	0.02083	0.04167
91	Asiatiques du Sud-Est	0.36264	0.42857	0.04396	0.02198	0.01099	0	0.13187
53	Sibériens	0.35849	0.20755	0.11321	0.03774	0.09434	0.11321	0.07547
296	Mongols	0.40878	0.26351	0.05743	0.03041	0.09797	0.05743	0.08446
668	Asiatiques	0.39072	0.32635	0.0509	0.02246	0.07036	0.03743	0.1018
36	Crees	0.33333	0	0.13889	0.27778	0	0.13889	0.11111
52	Ojibwas	0.38462	0.01923	0.21154	0.05769	0	0.19231	0.13462
24	Navajos	0.41667	0.04167	0.04167	0.20833	0	0.29167	0
75	Chipewyans	0.44	0	0.26667	0.04	0	0.22667	0.02667
33	Algonquins	0.24242	0	0.0303	0.18182	0	0.51515	0.0303
220	Amérindiens du Nord	0.37727	0.00909	0.17273	0.12273	0	0.25455	0.06364
80	Mayas	0.275	0	0.1125	0.3125	0	0.2	0.1
40	Aches	0.825	0	0	0.175	0	0	0
18	Ticunas	0.27778	0	0.11111	0.38889	0	0.05556	0.16667
83	Karitanas	0.13253	0	0	0.66265	0	0.18072	0.0241
141	Amérindiens du Sud	0.34752	0	0.01418	0.48936	0	0.11348	0.03546
441	Amérindiens (total)	0.34921	0.00454	0.11111	0.27438	0	0.19955	0.06122
61	Méxicains	0.34426	0.11475	0.13115	0.14754	0.01639	0.09836	0.14754
56	Antioquians	0.46429	0.01786	0.14286	0.08929	0.03571	0.14286	0.10714
13	Coriell HD18	0.38462	0.15385	0.07692	0.07692	0.07692	0.07692	0.15385
130	Métisses Amérindiens (total)	0.4	0.07692	0.13077	0.11538	0.03077	0.11538	0.13077

Tableau XXIII : fréquence des haplotypes dans les différentes populations amérindiennes, métisses amérindiennes, européennes, africaines et asiatiques.

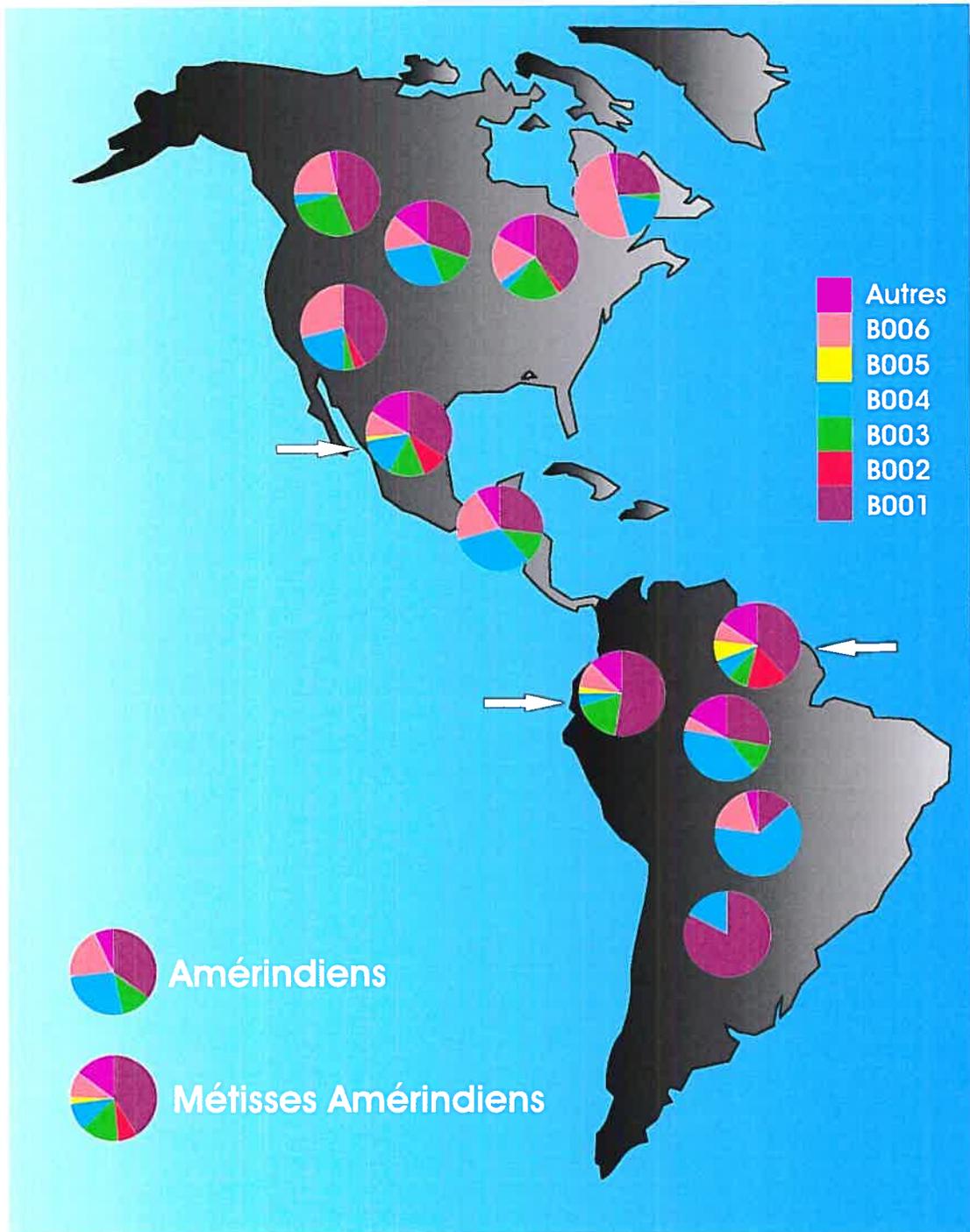


Figure 31 : Répartition des différents haplotypes dans les populations amérindiennes et métisses amérindiennes. En bas à gauche du graphique, se trouve la répartition moyenne des haplotypes chez les Amérindiens et les Métisses Amérindiens. Les populations métisses amérindiennes sont indiquées par une flèche. Pour une liste complète des populations, voir la figure 3 page 51.

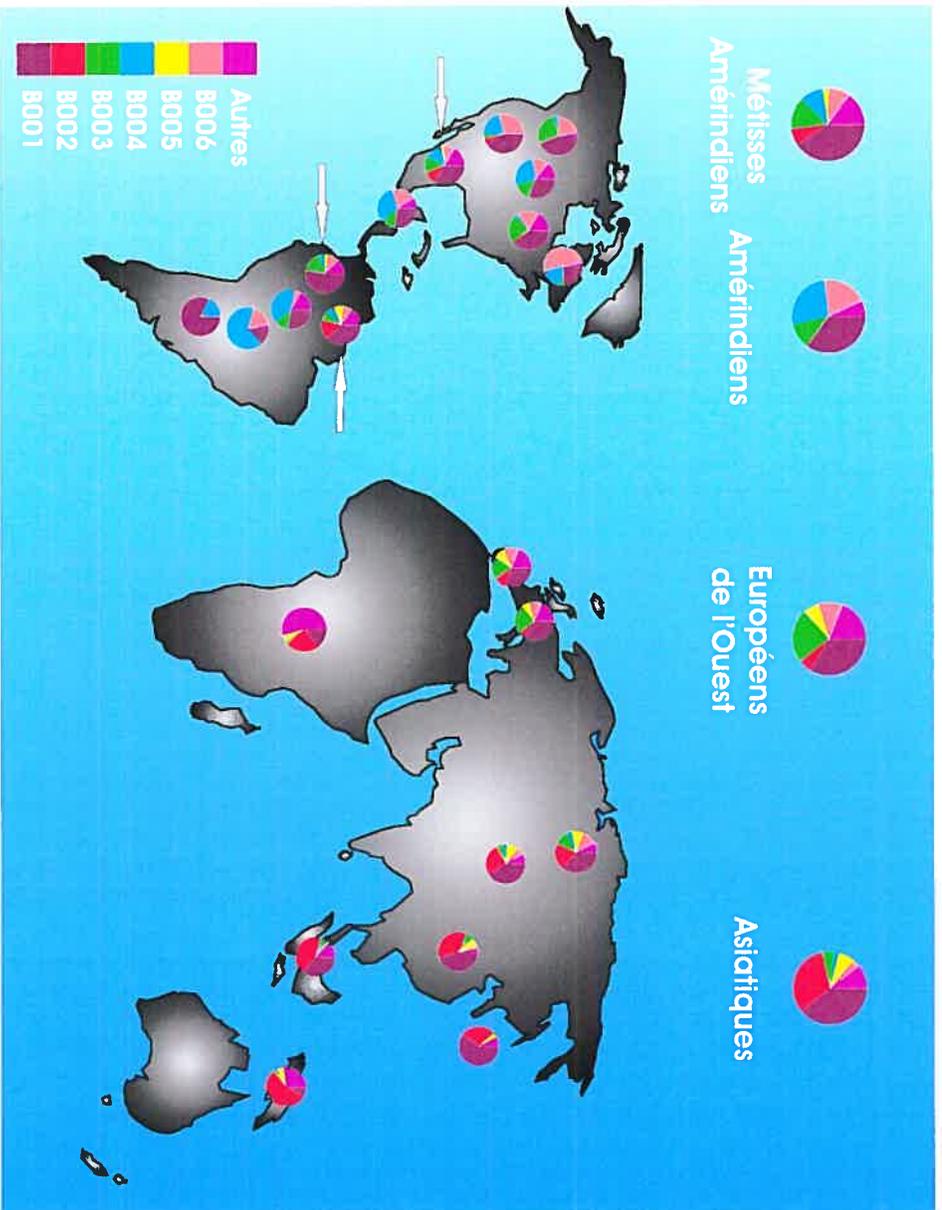


Figure 32 : Répartition des différents haplotypes dans les populations amérindiennes, métisses amérindiennes, asiatiques, africaines et européennes. En haut du graphique, se trouve la répartition moyenne des haplotypes chez les Métisses Amérindiens, les Amérindiens, les Européens de l'Ouest et les Asiatiques. Les populations métisses amérindiennes sont indiquées par une flèche. Pour une liste complète des populations, voir la figure 3 page 51.

	B001	B002	B003	B004	B005	B006	Autres
Métisses Amérindiens	52	10	17	15	4	15	17
Amérindiens	154	2	49	121	0	88	27
<i>p value</i>	0.39680	5.63E-07	0.56234	0.00110	0.00023	0.04708	0.01206

Tableau XXIV : Comparaison des populations métisses amérindiennes et amérindiennes par le test du χ^2

	B001	B002	B003	B004	B005	B006	Autres
Métisses Amérindiens	52	10	17	15	4	15	17
Européens de l'Ouest	156	25	111	1	32	51	87
<i>p value</i>	0.28335	0.34171	0.01813	3.80E-12	0.11690	0.87442	0.16928

Tableau XXV : Comparaison des populations métisses amérindiennes et européennes par le test du χ^2

	B001	B002	B003	B004	B005	B006	Autres
Métisses Amérindiens	52	10	17	15	4	15	17
Asiatiques	261	218	34	15	47	25	68
<i>p value</i>	0.87714	1.13E-06	0.00098	5.74E-07	0.10233	0.00028	0.35441

Tableau XXVI : Comparaison des populations métisses amérindiennes et asiatiques par le test du χ^2

Population	Diversité génétique
Africains	0.8890 +/- 0.0127
Ibériens	0.8527 +/- 0.0298
Européens de l'Ouest*	0.8010 +/- 0.0128
Européens de l'Ouest	0.8065 +/- 0.0118
Chinois	0.6636 +/- 0.0317
Japonais	0.5688 +/- 0.0526
Guinéens	0.6653 +/- 0.0627
Asiatiques du Sud-Est	0.6867 +/- 0.0329
Sibériens	0.8055 +/- 0.0350
Mongols	0.7476 +/- 0.0177
Asiatiques	0.7319 +/- 0.0116
Antioquiens	0.7448 +/- 0.0506
Coriell HD18	0.8590 +/- 0.0886
Mexicains	0.8301 +/- 0.0325
Métisses Amérindiens	0.7940 +/- 0.0279
Crees	0.7889 +/- 0.0367
Ojibwas	0.7768 +/- 0.0378
Navajos	0.7246 +/- 0.0533
Chipewyans	0.6908 +/- 0.0290
Algonquins	0.6610 +/- 0.0611
Amérindiens du Nord	0.7507 +/- 0.0158
Aches	0.2962 +/- 0.0788
Ticunas	0.7843 +/- 0.0680
Karitianas	0.5166 +/- 0.0539
Amérindiens du Sud	0.6202 +/- 0.0244
Mayas/Amérindiens du Centre	0.7816 +/- 0.0221
Amérindiens	0.7520 +/- 0.0097

Tableau XXVII : Diversité génétique des différentes populations africaines, amérindiennes, métisses amérindiennes, asiatiques et européennes.

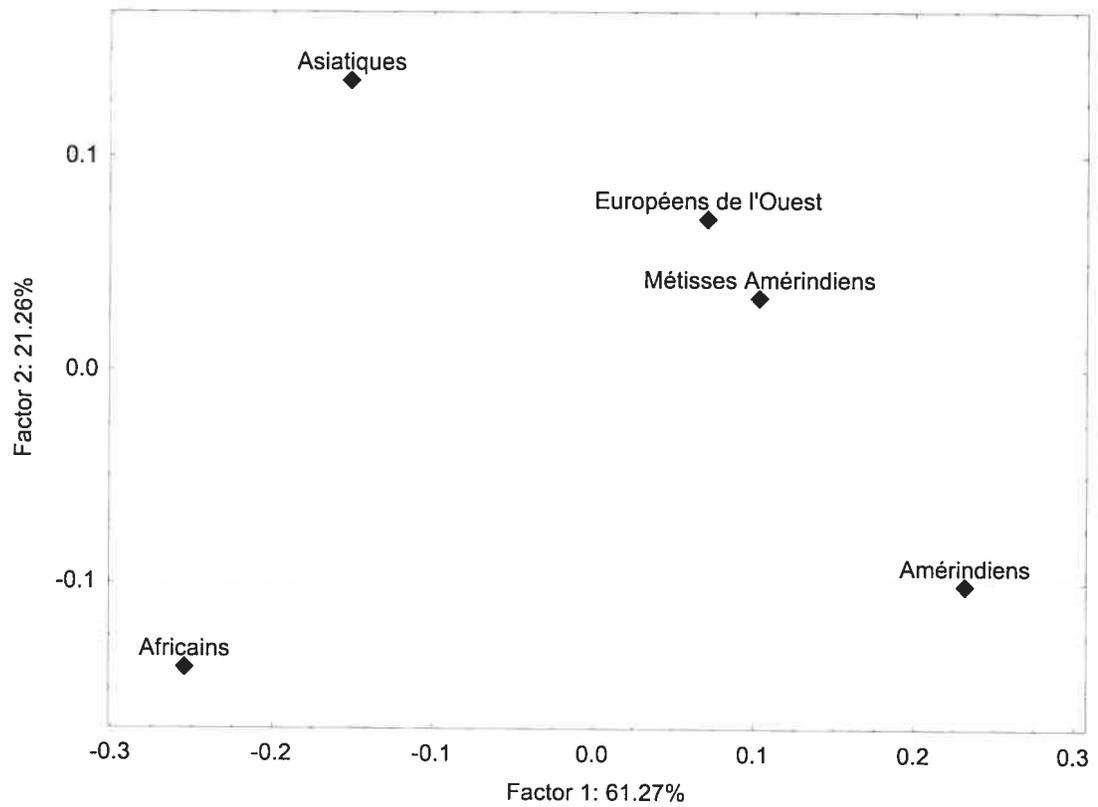


Figure 33 : Analyse en composantes principales (PCA) des populations amérindiennes, métisses amérindiennes, européennes et asiatiques.

	Européens de l'Ouest	Chinois	Japonais	Guinéens	Asiatiques du Sud-Est	Sibériens	Mongoles	Asiatiques	Métisses Amérindiens	Amérindiens	Africains
Européens de l'Ouest	0	0.08469	0.10216	0.20276	0.10944	0.01866	0.05202	0.07316	0.01621	0.06358	0.08866
Chinois	0.08469	0	0.0123	0.12605	0.0049	0.01749	0.00341	-0.00152	0.06112	0.13748	0.07583
Japonais	0.10216	0.0123	0	0.2247	0.05167	0.05073	0.02531	0.02791	0.0692	0.14733	0.12239
Guinéens	0.20276	0.12605	0.2247	0	0.07025	0.13401	0.12633	0.09922	0.20276	0.25704	0.07943
Asiatiques du Sud-Est	0.10944	0.0049	0.05167	0.07025	0	0.03687	0.02281	0.00658	0.08822	0.15958	0.06089
Sibériens	0.01866	0.01749	0.05073	0.13401	0.03687	0	-0.00168	0.00814	0.00687	0.06193	0.04684
Mongoles	0.05202	0.00341	0.02531	0.12633	-0.00168	0	0	0.00173	0.03114	0.09622	0.06193
Asiatiques	0.07316	-0.00152	0.02791	0.09922	0.00814	0.00173	0	0	0.05106	0.11986	0.06111
Métisses Amérindiens	0.01621	0.06112	0.0692	0.20276	0.08822	0.00687	0.03114	0.05106	0	0.02264	0.08134
Amérindiens	0.06358	0.13748	0.14733	0.25704	0.15958	0.06193	0.09622	0.11986	0.02264	0	0.13311
Africains	0.08866	0.07583	0.12239	0.07943	0.06089	0.04684	0.06193	0.06111	0.08134	0.13311	0

Tableau XXVIII : Valeurs des Fst telles que calculées par Arlequin, à partir de la fréquence des haplotypes.

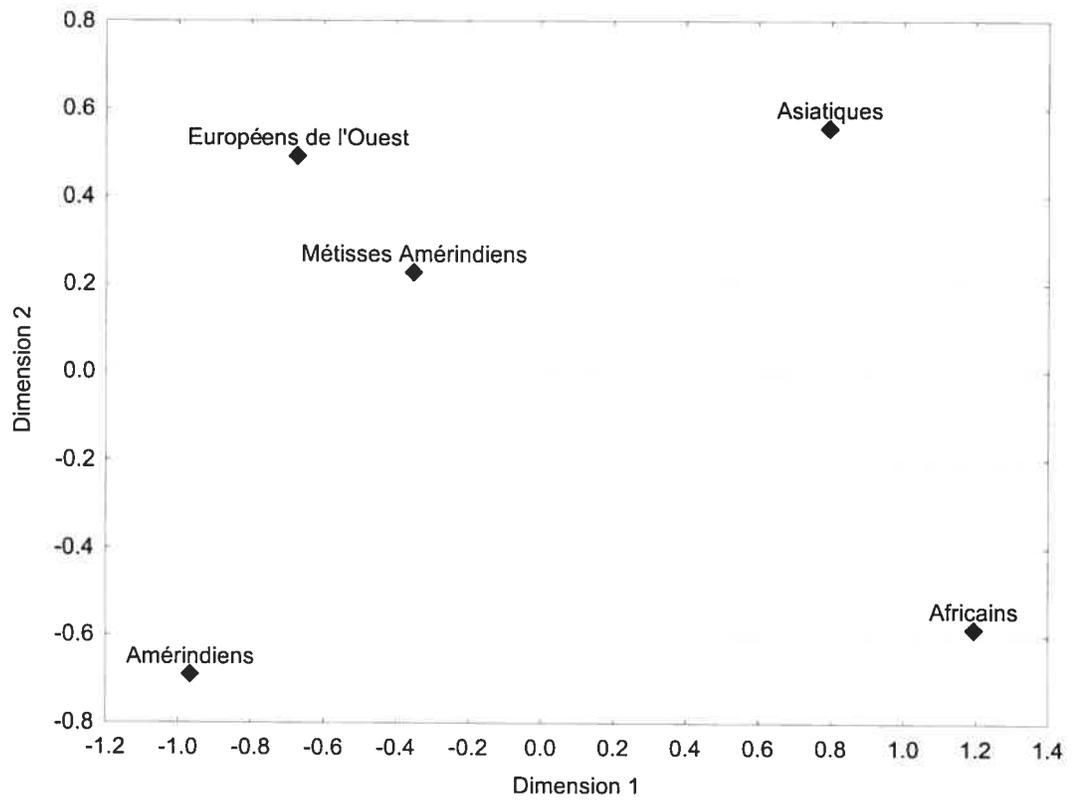


Figure 34 : MDS à partir des valeurs de F_{ST} pour les Amérindiens, les Métisses Amérindiens, les Asiatiques et les Européens de l'Ouest.

Chapitre IV

DISCUSSION

Les haplotypes du système *Dys44* sont composés de 30 polymorphismes bi-alléliques de substitution de nucléotide (SNPs), un SNP tri-allélique, deux délétions de trois nucléotides et une duplication de huit nucléotides. Les 2900 chromosomes X indépendants actuellement disponibles dans notre base de données, fruit de différentes études [273, 274, 279], dont 323 chromosomes analysés lors du travail présenté ici, se présentent avec 125 haplotypes différents qui furent arbitrairement classés de B001 à B125.

On retrouve quatre haplotypes majeurs dans les populations amérindiennes, les haplotypes B001, B003, B004 et B006, totalisant près de 94% de tous les haplotypes.

L'haplotype B001 est le plus fréquent dans le monde, avec une fréquence globale d'environ 34,9%. Il se retrouve à une fréquence comparable en Amérique du Nord, Centrale et du Sud et ne présente pas de variation remarquable d'une population à l'autre, cela à travers les deux groupes linguistiques représentés dans l'échantillon étudié, soit le NaDene et l'Amerind. Les haplotypes B003, B004 et B006 sont respectivement trouvés avec une fréquence de 11,1%, 27,4% et 20%. Ils sont retrouvés dans toutes les populations amérindiennes quelle que soit la localisation géographique ou la langue parlée, à l'exception d'une tribu d'Amérique du Sud, les Aches (figures 13 à 18 pages 77 à 82 et le tableau XIII page 76). Le nombre d'haplotypes majeurs et leur répartition à travers les groupes linguistiques est en parfait accord avec les données de l'ADN mitochondrial [192]. Leur fréquence et leur répartition les classent donc comme des haplotypes fondateurs au Nouveau Monde.

A l'exception de l'haplotype B001, la fréquence des trois autres haplotypes fondateurs varie grandement d'une population à l'autre, et varie géographiquement du nord au sud. Lorsque la fréquence moyenne au Nord, au Centre et au Sud est considérée, il semble y avoir un gradient de fréquences du nord au sud. Les haplotypes B003 et B006 voient leur fréquence diminuer en allant vers le sud

(respectivement 17,3% et 25,5% au Nord, 11,3% et 20% au Centre et 1,4% et 11,3% au Sud) alors que le contraire est observé pour l'haplotype B004 (48,9% au Sud, 31,3% au Centre et 12,3% au Nord) (figures 15 à 17 pages 79 à 81). Cela suggère un effet fondateur interne, c'est-à-dire un effet fondateur lors de la colonisation de l'Amérique du Nord, Centrale et du Sud. Dans les études sur le chromosome Y, le marqueur DYS199 montra également des différences dans sa répartition entre le Nord et le Sud [221, 234, 235, 237, 239, 241-243, 282]. Cela semble également partiellement en accord avec certaines données linguistiques [168, 225], observant des gradients Sud-Nord et Ouest-Est. Le nombre de populations ici étudiées, ainsi que leur localisation géographique, ne permet pas de tester l'établissement de gradients Est-Ouest.

Il faut souligner que l'Amérique Centrale n'est ici représentée que par une unique population, les Mayas, et pourrait par conséquent ne pas être représentative de cette région. De plus, sur les trois populations sud-américaines, deux démontrent une diversité génétique particulièrement faible, les Karitianas et les Aches (tableau XXI page 97). Ce manque de diversité est particulièrement dramatique chez les Aches pour lesquels les seuls deux haplogroupes fondateurs retrouvés, B001 et B004, forment 100% des haplotypes de cette population. Les Aches sont d'ailleurs la seule population présentant une variation significative de la fréquence de l'haplotype B001 par rapport aux autres populations, et à ne pas posséder les quatre haplogroupes fondateurs. Une étude sur des marqueurs du groupe sanguin avait déjà observé une grande dissimilarité entre les Aches et d'autres tribus paraguayennes [249].

Les Karitianas ne comptent que 150 individus (SIL 1995), et les Aches 639 (recensement de 1992 au Paraguay), ce qui laisse penser que la dérive génétique joue un rôle important dans la structure génétique de ces populations. Pour les Aches, nos données sont également en accord avec une diminution récente de leur population, en effet, il semblerait que les Aches aient été victimes d'un génocide sous le régime du Général Stroessner, de 1954 à 1989 [283-287].

Les tests de corrélation de Mantel des valeurs de F_{ST} avec les coordonnées géographiques et les différences linguistiques (tableau XVII page 84), confirment que les différences observées ne corrélaient pas avec les principales classifications linguistiques, mais avec la distance géographique séparant les tribus, renforçant l'idée que les haplotypes majeurs observés sont des haplotypes fondateurs. Il faut noter que seulement quatre classes différentes ont été utilisées dans la matrice linguistique, reflétant la classification actuelle des langues amérindiennes. Lorsque le test de Mantel fut effectué en utilisant également seulement quatre classes pour les distances géographiques, les résultats ne changèrent que très peu, la corrélation géographique restant la seule corrélation significative.

Lorsque comparées à l'Eurasie, les Amériques présentent de nombreuses et importantes différences. L'haplotype B001, avec une fréquence de 34,8% en Amérique, 33,7% en Europe et 39% en Asie, ne présente pas de différence significative entre les trois continents. Les haplotypes B002, dont la fréquence n'est que de 0,5% chez les Amérindiens, et B005, complètement absent dans les populations amérindiennes, se retrouvent à haute fréquence en Europe et en Asie (figures 20 et 23 pages 90 et 93 et tableau XVIII page 88). La fréquence de B002 est de 5,4% en Europe et atteint 32,4% en Asie, la fréquence de B005 est d'environ 7% en Europe et en Asie. De plus, ces deux haplotypes se retrouvent dans toutes les populations eurasiatiques étudiées.

L'haplotype B004 est assez caractéristique des populations amérindiennes, présent à 26,6% en Amérique contre 2,4% en Asie et seulement 0,2% en Europe, et sa fréquence est plus élevée en Amérique du Sud qu'en Amérique du Nord (figure 22 page 92 et tableau XVIII page 88). Il faut noter que parmi les populations asiatiques, ce sont les Sibériens qui présentent la plus haute fréquence de cet haplotype. Ces caractéristiques sont très proches de celles de l'allèle T du marqueur DYS1999 sur le chromosome Y [221, 234, 235, 237, 239, 241-243, 282].

L'haplotype B006 voit également sa fréquence augmenter en Amérique (19,9%), mais d'une manière moins marquée que B004, puisque sa fréquence en Europe est de 11% et de 3,8% en Asie. Une fois encore, c'est en Sibérie que la fréquence asiatique de cet haplotype est la plus élevée (figure 24 page 94 et tableau XVIII page 88). L'haplotype B003 se retrouve à haute fréquence en Europe (24%), à une fréquence intermédiaire en Amérique (11,1%) et à plus faible fréquence, quoiqu'encore assez élevée, en Asie (5,5%)

La fréquence des autres haplotypes est plus faible en Amérique (6,1%) que sur les autres continents (18,8% en Europe et 10,2% en Asie). Malgré une plus grande fréquence des autres haplotypes en Asie qu'en Amérique, la diversité génétique des Amérindiens apparaît plus élevée que celle des Asiatiques (tableau XXI page 97). Il apparaît donc que l'ensemble des autres haplotypes trouvés en Asie regroupe moins d'haplotypes très rares que ce qui est trouvé chez les populations amérindiennes. Ce sont les Européens qui présentent la plus grande diversité génétique, en accord avec la très haute fréquence de leurs autres haplotypes. Une fois encore, la Sibérie se distingue des autres populations asiatiques en présentant une diversité génétique comparable à celle trouvée chez les Européens, et donc supérieure à celle trouvée chez les Amérindiens.

Le fait que deux haplotypes très fréquents en Asie, et aussi en Europe, soient presque totalement absents chez les Amérindiens suggère que les quatre haplotypes fondateurs retrouvés en Amérique (B001, B003, B004 et B006) sont arrivés sur le continent américain lors d'une migration unique. Cette migration a été accompagnée d'un important effet fondateur, B002 et B005 n'étant pas représentés chez les Amérindiens, B006 et surtout B004 étant sur-représentés en Amérique et B003 ayant une fréquence significativement différente de celle trouvée en Europe et en Asie. L'explication la plus parcimonieuse est que la migration unique ayant donné naissance aux populations amérindiennes provienne d'une population asiatique présentant tous les haplotypes fondateurs, ce qui semble, d'après nos données, indiquer la Sibérie comme berceau des populations amérindiennes.

En cela, nos données semblent en parfait accord avec les données de l'ADN mitochondrial [36, 192, 200, 203] et du chromosome Y [221, 240] qui suggèrent une seule vague de migration en provenance d'Asie, plus particulièrement en provenance de la Sibérie ou de la Mongolie.

Afin d'appuyer nos conclusions, nous avons observé les relations existant entre les populations que nous avons étudiées par analyse en composantes principales (ACP) (figures 26 et 27 et 33 pages 98 et 99 et 111), et en représentant graphiquement les données de F_{ST} par MDS (figures 28 à 30 et 34 pages 101 à 103 et 113). De façon assez surprenante, c'est la population européenne qui apparaît comme étant la plus proche des Amérindiens. La décomposition des Amérindiens selon la géographie (Nord, Centre, Sud) montre que les trois groupes sont plus proches des Européens que des Asiatiques, les Amérindiens du Nord étant toutefois les plus proches des Européens (figure 29 page 102).

En décomposant alors la population asiatique générale en ses sous-populations, d'après les groupes ethniques, la MDS des F_{ST} (figure 30 page 103), et la PCA des fréquences confirment que les Européens sont plus près des Amérindiens que n'importe quelle population asiatique. Les Sibériens se retrouvent comme population asiatique la plus proche des Amérindiens, mais ils se retrouvent également plus proches des Européens que des Amérindiens.

Afin de vérifier si nos résultats étaient le fruit de métissages des Européens avec les Amérindiens et les Sibériens, trois populations amérindiennes connues pour leur fort métissage européen ont été étudiées : les Mexicains, les Antioquiens et des Métisses d'Amérique du Sud (Coriell HD18) (figures 31 et 32 pages 107 et 108 et tableau XXIII page 106).

Les haplotypes B002 et B005 font leur apparition dans ces populations à des fréquences de 7% et 2,6%, respectivement, alors qu'ils étaient virtuellement absents chez les Amérindiens. La fréquence de B004 et de B006 se trouve également diminuée, alors que la fréquence des autres haplotypes augmente et que la diversité génétique devient comparable à celle trouvée chez les Européens, voir tableau XXVII page 110. Les analyses PCA (figure 33 page 108), et MDS (figure 34 page 111), montrent que les populations métisses se placent plus près des populations européennes que des populations amérindiennes ou asiatiques.

De plus, des mutations africaines ont été trouvées chez deux individus Mexicains et un Antioquian, semblant indiquer un métissage entre les Amérindiens et Africains, principalement des esclaves qui fuyaient les Etats-Unis, en plus de celui avec les Européens, comme cela avait déjà été montré avec d'autres systèmes génétiques dans différentes populations métisses [30, 288-298]. Toutefois, la présence d'un haplotype africain complet chez un individu espagnol montre que des mutations typiquement africaines peuvent aussi avoir été véhiculées par des Européens du Sud, en accord avec l'histoire de la péninsule ibérique, qui fut longtemps sous domination des Maures. Il serait important de pousser l'étude des populations ibériques afin de déterminer si une partie importante du métissage attribué aux Africains, dans les populations amérindiennes, ne serait pas d'origine espagnole ou portugaise.

Quant aux populations sibériennes, il est très peu probable que la fréquence des haplotypes B004 et B002 puisse être expliquée par un métissage européen. Toutefois, aucune conclusion ne peut être faite à ce propos pour l'haplotype B006 et les autres haplotypes.

Conclusions.

Afin de complètement exclure des migrations différentielles selon la classification linguistique, comme proposé par Greenberg [162], il sera nécessaire de compléter la collection d'échantillons avec des populations provenant de la troisième classification linguistique, l'Eskimo-Aleut. Les populations sibériennes devront être investiguées plus finement en augmentant le nombre des ethnies sibériennes représentées dans l'échantillonnage. Enfin, les populations d'Amérique Centrale et du Sud devront être diversifiées afin de palier aux caractères particuliers de deux des trois populations sud-amérindiennes étudiées et à la représentation de l'Amérique Centrale par une unique population.

Des fréquences de B002, B005, B004 et B006 dans les populations amérindiennes, asiatiques, européennes et métisses nous pouvons conclure que nos observations ne sont pas dues à un métissage européen. L'explication considérant les populations européennes comme faisant partie des populations ancestrales des Amérindiens [104], ne sont pas compatibles avec les données que nous observons. L'explication la plus parcimonieuse est que les populations amérindiennes ont été fondées par une seule vague de migration provenant de Sibérie, en accord avec les données mitochondriales et du chromosome Y, et se sont ensuite réparties sur le continent américain entraînant ainsi des effets fondateurs internes (gradients Nord-Sud de la fréquence des haplotypes B003, B004 et B006).

Si nos données sont en accord avec une seule vague de migration, il est difficile de se prononcer sur la voie de migration empruntée (terrestre ou côtière) (figure 35 page 123). Dans le cas de la voie terrestre, le mouvement migratoire se serait effectué du nord vers le sud, laissant possiblement apparaître des effets fondateurs internes au continent américain, comme ceux que nous observons. Mais un tel mouvement migratoire Nord-Sud, combiné à un mouvement migratoire Ouest-Est, est également attendu dans le cas du scénario de la voie côtière.

Si l'Amérique du Sud était la première région d'Amérique à avoir été largement colonisée, comme certains archéologues le prétendent, alors la seule voie possible de migration serait la voie côtière. Dans ce cas, la colonisation de l'Amérique du Nord pourrait être le résultat de migrations en provenance d'Amérique Centrale et du Sud, possiblement en combinaison avec une expansion de l'ouest vers l'est de populations qui avaient colonisé les côtes de l'Amérique du Nord lors de la première vague migratoire.

Nos données vont à l'encontre de cette dernière théorie, plaçant les Amérindiens d'Amérique Centrale et du Sud comme étant les plus éloignés des populations Asiatiques et Européennes (en comparaison aux Amérindiens d'Amérique du Nord), et s'opposant donc à la théorie d'une importante migration du sud vers le nord comme origine des populations amérindiennes d'Amérique du Nord.

Le mouvement migratoire se serait donc effectué du nord vers le sud, et possiblement de l'ouest vers l'est. Par contre, le nombre de populations étudiées et leur localisation ne permettant pas de distinguer d'éventuels effets fondateurs internes lors d'un mouvement de migration Ouest-Est, il est ici impossible de départager le scénario terrestre du scénario côtier.

La parenté semblant exister entre les populations amérindiennes, sibériennes et européennes, en accord avec la présence de l'haplogroupe mitochondrial X de Brown [201, 299] et de l'allèle HLA-*DRB1**1402 [300], pourrait très bien prendre ses racines dans la colonisation de l'Europe et de l'Asie du Nord. En effet, plusieurs études semblent indiquer qu'il y aurait eu deux voies de migration hors d'Afrique, une voie Nord et une voie Sud [301-303]. La voie Nord aurait mené au peuplement de l'Europe, du Caucase et du Nord de l'Asie alors que la voie Sud serait à l'origine du peuplement de l'Asie du Sud, du Sud-Est et de l'Est [302, 304] (figure36 page 124).

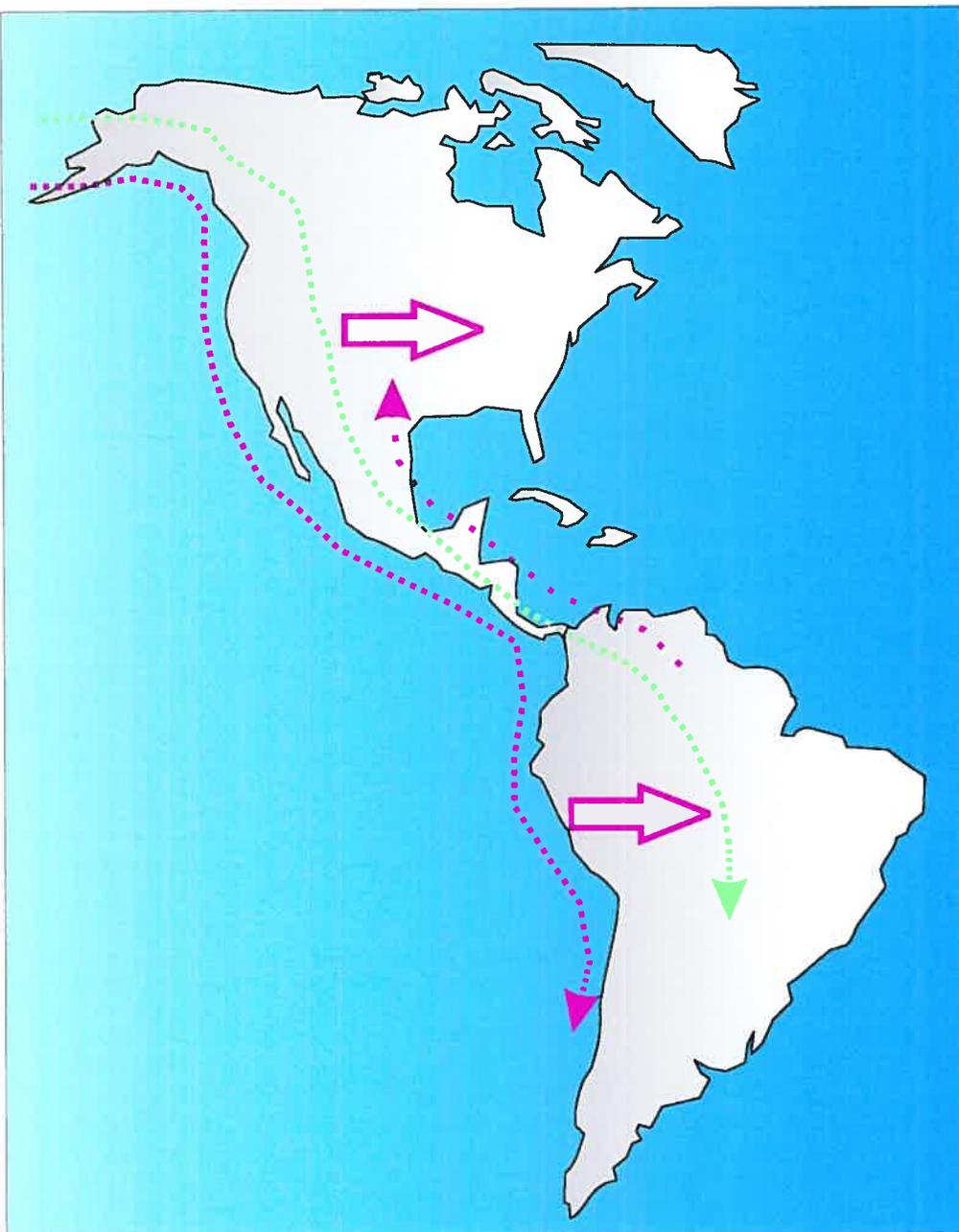


Figure 35 : Voies de migration et colonisation des Amériques

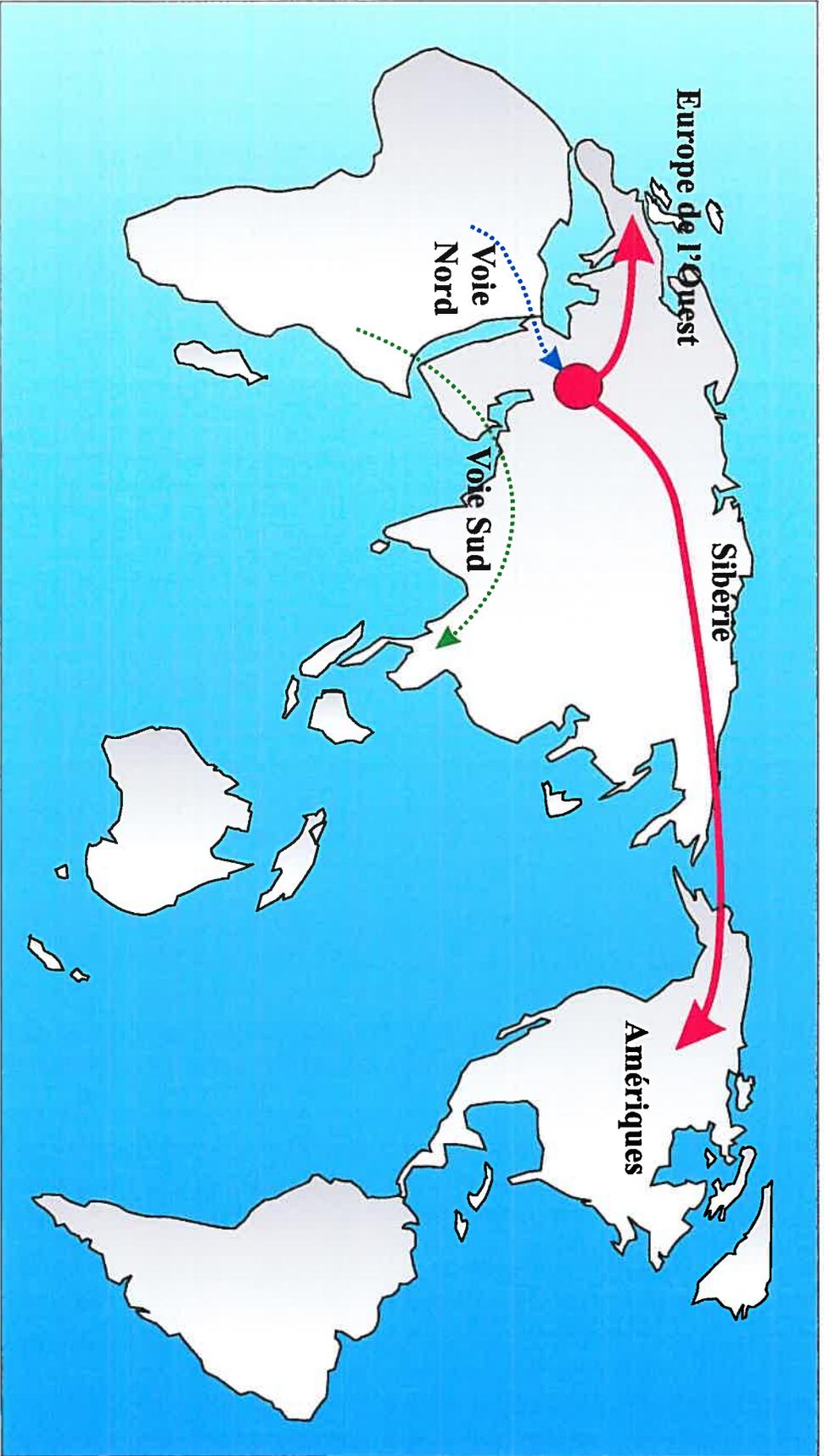


Figure 36 : modèle expliquant l'origine des liens entre les Européens de l'ouest, les Sibériens et les Amérindiens.

Des membres de populations d'Asie du Nord, probablement des Sibériens, ont ultérieurement migré vers l'ouest jusqu'en Amérique, donnant ainsi naissance aux populations amérindiennes. Les liens entre les populations amérindiennes, sibériennes et d'Europe de l'Ouest pourraient donc être le résultat de mouvements de migration hors d'Afrique vers le Caucase par la voie Nord, et de subséquentes migrations du Caucase en Europe de l'Ouest et en Asie du Nord, et finalement de l'Asie du Nord en Amérique (figure 36 page124).

Des populations d'Amérique du Sud, du groupe Eskimo-Aleut et de Sibérie vont être prochainement analysées par notre laboratoire pour le système *Dys44*, afin de lever certaines incertitudes concernant la représentation de ces régions dans l'échantillonnage actuel. Afin d'obtenir une datation relative des mouvements migratoires en Amérique, un microsatellite se trouvant dans l'intron 44 du gène de la Dystrophine est en cours d'analyse pour l'ensemble des échantillons. A ce système va venir s'ajouter l'analyse d'une insertion Alu et d'un autre microsatellite dans la même région, ce qui devrait nous permettre d'observer la structure des populations amérindiennes avec une plus grande finesse et de déterminer un âge approximatif pour la colonisation du continent américain.

BIBLIOGRAPHIE

1. Gulcher, J.R., A. Kong, and K. Stefansson, *The role of linkage studies for common diseases*. *Curr Opin Genet Dev*, 2001. **11**(3): p. 264-7.
2. Weiss, K.M. and J.D. Terwilliger, *How many diseases does it take to map a gene with SNPs?* *Nat Genet*, 2000. **26**(2): p. 151-7.
3. Risch, N. and K. Merikangas, *The future of genetic studies of complex human diseases*. *Science*, 1996. **273**(5281): p. 1516-7.
4. Risch, N.J., *Searching for genetic determinants in the new millennium*. *Nature*, 2000. **405**(6788): p. 847-56.
5. Menzel, S., *Genetic and molecular analyses of complex metabolic disorders: genetic linkage*. *Ann N Y Acad Sci*, 2002. **967**: p. 249-57.
6. Goldstein, D.B. and M.E. Weale, *Population genomics: linkage disequilibrium holds the key*. *Curr Biol*, 2001. **11**(14): p. R576-9.
7. Pritchard, J.K. and M. Przeworski, *Linkage disequilibrium in humans: models and data*. *Am J Hum Genet*, 2001. **69**(1): p. 1-14.
8. Pritchard, J.K. and N.A. Rosenberg, *Use of unlinked genetic markers to detect population stratification in association studies*. *Am J Hum Genet*, 1999. **65**(1): p. 220-8.
9. Pesch, K., et al., *Case populations must match the respective disease model: Genotype diversity causes linkage disequilibrium mapping failure in monogenic disorders*. *Int J Mol Med*, 2001. **8**(1): p. 53-8.
10. Gulcher, J., A. Kong, and K. Stefansson, *The genealogic approach to human genetics of disease*. *Cancer J*, 2001. **7**(1): p. 61-8.
11. Tregouet, D.A. and L. Tiret, *Applications of the estimating equations theory to genetic epidemiology: a review*. *Ann Hum Genet*, 2000. **64**(Pt 1): p. 1-14.
12. Barbujani, G., et al., *An apportionment of human DNA diversity*. *Proc Natl Acad Sci U S A*, 1997. **94**(9): p. 4516-9.
13. Jorde, L.B., et al., *The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data*. *Am J Hum Genet*, 2000. **66**(3): p. 979-88.
14. Nei, M. and A.K. Roychoudhury, *Genic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids*. *Am J Hum Genet*, 1974. **26**(4): p. 421-43.
15. Nei, M. and A.K. Roychoudhury, *Evolutionary relationships of human populations on a global scale*. *Mol Biol Evol*, 1993. **10**(5): p. 927-43.
16. Dries, D.L., et al., *Racial differences in the outcome of left ventricular dysfunction*. *N Engl J Med*, 1999. **340**(8): p. 609-16.
17. Exner, D.V., et al., *Lesser response to angiotensin-converting-enzyme inhibitor therapy in black as compared with white patients with left ventricular dysfunction*. *N Engl J Med*, 2001. **344**(18): p. 1351-7.
18. Grimaldi, R., et al., *Prevalence of the CCR5Delta32 mutation in Brazilian populations and cell susceptibility to HIV-1 infection*. *Hum Genet*, 2002. **111**(1): p. 102-4.
19. Palatnik, M., et al., *Ethnicity and type 2 diabetes in Rio de Janeiro, Brazil, with a review of the prevalence of the disease in Amerindians*. *Hum Biol*, 2002. **74**(4): p. 533-44.

20. Loureiro, C.L., et al., *High prevalence of GB virus C/hepatitis G virus genotype 3 among autochthonous Venezuelan populations*. J Med Virol, 2002. **68**(3): p. 357-62.
21. Bradford, L.D., *CYP2D6 allele frequency in European Caucasians, Asians, Africans and their descendants*. Pharmacogenomics, 2002. **3**(2): p. 229-43.
22. Jorge, L.F. and T.D. Arias, *[Human biodiversity and its effects on the pharmacological variability: CYP2D6 and NAT2 enzymes in Amerind populations of Panama, Colombia and Costa Rica]*. Rev Med Panama, 1995. **20**(3): p. 98-107.
23. Jorge, L.F., et al., *Comparative evolutionary pharmacogenetics of CYP2D6 in Ngawbe and Embera Amerindians of Panama and Colombia: role of selection versus drift in world populations*. Pharmacogenetics, 1999. **9**(2): p. 217-28.
24. Hall, D., et al., *Variability at the uridine diphosphate glucuronosyltransferase 1A1 promoter in human populations and primates*. Pharmacogenetics, 1999. **9**(5): p. 591-9.
25. Betard, C., et al., *Localization of a recessive gene for North American Indian childhood cirrhosis to chromosome region 16q22-and identification of a shared haplotype*. Am J Hum Genet, 2000. **67**(1): p. 222-8.
26. Chagnon, P., et al., *A Missense Mutation (R565W) in Cirhin (FLJ14728) in North American Indian Childhood Cirrhosis*. Am J Hum Genet, 2002. **71**(6): p. 1443-9.
27. Pena, S.D., et al., *A major founder Y-chromosome haplotype in Amerindians*. Nat Genet, 1995. **11**(1): p. 15-6.
28. Pena, S.D., *Origin of the Amerindians*. Science, 1999. **283**(5410): p. 2017.
29. Salzano, F.M., *The peopling of the Americans as viewed from South America*. Acta Anthropogenet, 1984. **8**(1-2): p. 111-23.
30. Salzano, F.M., et al., *Reconstructing history: the Amazonian Mura Indians*. Hum Biol, 1990. **62**(5): p. 619-35.
31. Santos, S.E., et al., *Multiple founder haplotypes of mitochondrial DNA in Amerindians revealed by RFLP and sequencing*. Ann Hum Genet, 1996. **60** (Pt 4): p. 305-19.
32. Santos, F.R., et al., *North and South Amerindians may have the same major founder Y chromosome haplotype*. Am J Hum Genet, 1996. **58**(6): p. 1369-70.
33. Schurr, T.G., et al., *Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages*. Am J Hum Genet, 1990. **46**(3): p. 613-23.
34. Carvalho-Silva, D.R., et al., *The phylogeography of Brazilian Y-chromosome lineages*. Am J Hum Genet, 2001. **68**(1): p. 281-6.
35. Tarazona-Santos, E. and F.R. Santos, *The peopling of the Americas: a second major migration?* Am J Hum Genet, 2002. **70**(5): p. 1377-80; discussion 1380-1.
36. Merriwether, D.A., F. Rothhammer, and R.E. Ferrell, *Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World*. Am J Phys Anthropol, 1995. **98**(4): p. 411-30.
37. Pannell J.R., C.B., *Effects of metapopulation processes on measures of genetic diversity*. Philosophical Transactions: Biological Sciences, 2000. **355**(1404): p. 1851-1864.

38. Cano, R.J., *Analysing ancient DNA*. Endeavour, 1996. **20**(4): p. 162-7.
39. Cano, R.J., et al., *Sequence analysis of bacterial DNA in the colon and stomach of the Tyrolean Iceman*. Am J Phys Anthropol, 2000. **112**(3): p. 297-309.
40. Paabo, S., J.A. Gifford, and A.C. Wilson, *Mitochondrial DNA sequences from a 7000-year old brain*. Nucleic Acids Res, 1988. **16**(20): p. 9775-87.
41. Paabo, S., *Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification*. Proc Natl Acad Sci U S A, 1989. **86**(6): p. 1939-43.
42. Paabo, S., R.G. Higuchi, and A.C. Wilson, *Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology*. J Biol Chem, 1989. **264**(17): p. 9709-12.
43. Paabo, S., *Ancient DNA*. Sci Am, 1993. **269**(5): p. 86-92.
44. Handt, O., et al., *Ancient DNA: methodological challenges*. Experientia, 1994. **50**(6): p. 524-9.
45. Handt, O., et al., *The retrieval of ancient human DNA sequences*. Am J Hum Genet, 1996. **59**(2): p. 368-76.
46. Lindahl, T., *Facts and artifacts of ancient DNA*. Cell, 1997. **90**(1): p. 1-3.
47. Lindahl, T., *Recovery of antediluvian DNA*. Nature, 1993. **365**(6448): p. 700.
48. Cooper, A. and H.N. Poinar, *Ancient DNA: do it right or not at all*. Science, 2000. **289**(5482): p. 1139.
49. Cooper, A., et al., *Neandertal genetics*. Science, 1997. **277**(5329): p. 1021-4.
50. Krings, M., et al., *DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen*. Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5581-5.
51. Krings, M., et al., *A view of Neandertal genetic diversity*. Nat Genet, 2000. **26**(2): p. 144-6.
52. Ovchinnikov, I.V., et al., *Molecular analysis of Neanderthal DNA from the northern Caucasus*. Nature, 2000. **404**(6777): p. 490-3.
53. Hewitt, G., *The genetic legacy of the Quaternary ice ages*. Nature, 2000. **405**(6789): p. 907-13.
54. Taberlet, P., et al., *Comparative phylogeography and postglacial colonization routes in Europe*. Mol Ecol, 1998. **7**(4): p. 453-64.
55. Webb T III, B.P., *Global changes during the last 3 million years: climatic controls and biotic responses*. Annu. Rev. Ecol. Syst., 1992. **23**: p. 141-73.
56. Lahr MM, a.F.R., *Genes, fossils and behaviour: when and where do they fit?*, in *Genes, Fossils and Behaviour: an integrated approach to modern human origins.*, P.D.a.R. Foley, Editor. 2001, NATO: Brussels. p. 13-48.
57. Stuiver, M., and Braziunas T. F., *Modeling atmospheric 14C influences and 14C ages of marine samples back to 10,000 BC*. Radiocarbon, 1993. **35**(1): p. 137-189.
58. Elias, S.A., *Late Pleistocene climates of Beringia, based on fossil beetle analysis*. Quat. Res., 2000. **53**: p. 229-235.
59. Alfimov, A.V. and D.I. Berman, *Beringian climate during the Late Pleistocene and Holocene*. Quat. Sci. Rev., 2001. **20**: p. 127-134.
60. Elias, S.A., *Setting the Stage: Environmental Conditions in Beringia as People Entered the New World*, in *The First Americans*, N.G. Jablonski, Editor. 2002, California Academy of Sciences: sAN FRANCISCO.

61. Elias, S.A., *Mutual climatic range reconstruction of seasonal temperatures based on Late Pleistocene fossil beetle assemblages in Eastern Beringia*. *Quat. Sci. Rev.*, 2001. **20**: p. 77-91.
62. Kunz, M.L., et al., *The life and times of Paleoindians in Arctic Alaska*. Arctic Research of the United States, 1999. **13**: p. 33-39.
63. Dillehay, T.D., *The cultural relationships of Monte Verde: A Late Pleistocene settlement site in the sub-antarctic forest of south-central Chile*, in *New Evidence for the Pleistocene Peopling of the Americas*, A.L. Bryan, Editor. 1986, Center for the Study of Early Man: Orono. p. 319-337.
64. Catto, N.R., *Richardson Mountains, Ykon-Northwest Territories: the northern portal of the postulated "ice-free corridor."* *Quat. Int.*, 1996. **32**: p. 3-19.
65. Goebel, T., R. Powers, and N. Bigelow, *The Nenana complex of Alaska and Clovis origins*, in *Clovis: Origins and Adaptations*, R. Bonnicksen and K.L. Turnmire, Editors. 1991, Center for the Study of the First Americans, Department of Anthropology, Oregon State University: Corvallis. p. 49-80.
66. Hoffecker, J.F., W.R. Powers, and T. Goebel, *The colonization of Beringia and the peopling of the New World*. *Science*, 1993. **259**: p. 46-53.
67. Hoffecker, J.F., W.R. Powers, and N.H. Bigelow, *Dry Creek*, in *American Beginings: The Prehistory and Paleoecology of Beringia*, F.H. West, Editor. 1996, University of Chicago Press: Chicago. p. 343-354.
68. Yesner, D.R., *Human adaptation at the Pleistocene-Holocene border*, in *Humans at the End of Ice Age*, L.G. Strauss, et al., Editors. 1996, Plenum: New York. p. 255-276.
69. West, F.H., *American Beginings: The Prehistory and Paleoecology of Beringia*. 1996, Chicago: University of Chicago Press. 576.
70. Fagan, B., *The Great Journey: The Peopling of Ancient America*. 1987, London: Thames and Hudson. 288.
71. Fiedel, S.J., *Prehistory of the Americas*. 1987, Cambridge: Cambridge University Press. 386.
72. Fiedel, S.J., *Prehistory of the Americas*. Second Edition ed. 1992, Cambridge: Cambridge University Press. 400.
73. Fiedel, S.J., *The peopling of the New World: Present evidence, new theories, and future directions*. *J. Archaeol. Res.*, 2000. **8**: p. 39-103.
74. Haynes, C.V., Jr., *Fluted projectile points: Their age and dispersion*. *Science*, 1964. **145**: p. 1408-1413.
75. Jennings, J.D., *Origins*, in *Ancient North Americans*, J.D. Jennings, Editor. 1983, W. H. Freeman and Company: New York. p. 25-69.
76. Martin, P.S., *Prehistoric overkill*, in *Pleistocene Extinctions: The Search for a Cause*, P.S. Martin and H.E. Wright, Editors. 1967, Yale University Press: New Haven. p. 75-120.
77. Martin, P.S., *Prehistoric extinctions: The global model*, in *Quaternary Extinction: A Prehistoric Revolution*, P.S. Martin and R.G. Klein, Editors. 1984, University of Arizona: Tucson. p. 354-403.
78. Lynch, T.F., *The Paleo-Indians*, in *Ancient South Americans*, J.D. Jennings, Editor. 1983, Freeman, W. H.: San Francisco. p. 87-138.
79. Bowman, S., *Radiocarbon dating*. 1990: University of California Press, Berkeley, and British Museum, London. 62.

80. Aitken, M.J., *Science-Based Dating in Archaeology*. 1990, London: Longman. 274.
81. Taylor, R.E., *Radiocarbon Dating, An Archaeological Perspective*. 1987, New York: Academic Press. 212.
82. Taylor, R.E., *Radiocarbon dating of bone: To collagen and beyond*, in *Radiocarbon Dating after Four Decades: An Interdisciplinary Perspective*, R.E. Taylor, A. Long, and R.S. Kra, Editors. 1992, Springer-Verlag: New York. p. 375-402.
83. Taylor, R.E., *Radiocarbon dating*, in *Chronometric Dating in Archaeology*, R.E. Taylor and M.J. Aitken, Editors. 1997, Plenum Press: New York. p. 65-96.
84. Taylor, R.E. and M.J. Aitken, *Chronometric Dating in Archaeology*. Advance in Archaeological and Museum Science. Vol. 1. 1997, New York: Plenum Press. 395.
85. Haynes, C.V., Jr., *Contributions of radiocarbon dating to the geochronology of the peopling of the New World*, in *14 C Dating and the Peopling of the New World*, R.E. Taylor, A. Long, and R.S. Kra, Editors. 1992, Springer-Verlag: New York. p. 355-374.
86. Mead, J.I. and D.J. Meltzer, *North American Quaternary extinctions and the radiocarbon record*, in *Quaternary Extinctions: A Prehistoric Revolution*, P.S. Martin and R.G. Klein, Editors. 1984, University of Arizona: Tucson. p. 440-450.
87. Meltzer, D.J. and J.I. Mead, *Dated Late Pleistocene extinctions: Theoretical issues, analytical bias, and substantive results*, in *Environments and Extinctions: Man in the Late Glacial North America*, J.I. Mead and D.J. Meltzer, Editors. 1985, Center for the Study of Early Man, University of Maine: Orono. p. 145-173.
88. Roosevelt, A.C., *Paleoindian and Archaic occupations in the Lower Amazon: A summary and comparison*, in *Explorations in American Archaeology: Essays in Honor of Wesley R. Hurt*, M.G. Plew, Editor. 1998, University Press of America: Lanham. p. 165-191.
89. Roosevelt, A.C., *Clovis classification: A follow-up*. Mammoth Trumpet, 1998. **13**: p. 14-17.
90. Stanford, D., *Clovis origins and adaptations: An introductory perspective*, in *Clovis: Origins and Adaptations*, R. Bonnicksen and K.L. Turnmire, Editors. 1991, Center for the Study of the First Americans, Department of Anthropology, Oregon State University: Corvallis. p. 1-4.
91. Toth, N., *The material record*, in *The First Americans: Search and Research*, T.D. Dillehay and D.J. Meltzer, Editors. 1991, CRC Press: Boca Raton. p. 53-76.
92. Haffner, S., et al., *Prevalence of hypertension in Mexico City and San Antonio, Texas*. *Circulation*, 1994. **90**(3): p. 1542-9.
93. Holliday, V.T., *Paleoindian Georarchaeology of the Southern High Plains*. 1997, Austin: University of Texas. 297.
94. Haynes, C.V., Jr., et al., *Keresey-Kuner Terrace investigations at the Dent and Bernhardt sites, Colorado*. *Geoarchaeology*, 1998. **13**: p. 201-218.

95. Dixon, E.J., *Bones, Boats & Bison*. 1999, Albuquerque: University of New Mexico Press. 322.
96. Stanford, D.J., *Pre-Clovis occupation south of the ice sheets*, in *Early Man in the New World*, R. Shutler, Jr., Editor. 1983, Sage Publications: Beverly Hills. p. 65-72.
97. Roosevelt A.C., D.J.a.B.L., *The Migrations and Adaptations of the First Americans Clovis and Pre-Clovis Viewed from South America*, in *The First Americans*, N.G. Jablonski, Editor. 2002, California Academy of Sciences: San Francisco. p. 224-235.
98. Clark, D.W., *The northwestern (Alaska-Ykon) fluted points*, in *Clovis: Origins and Adaptations*, R. Bonnichsen and K.L. Turnmire, Editors. 1991, Center for the Study of the First Americans, Department of Anthropology, Oregon State University: Corvallis. p. 35-48.
99. Derev'anko, A.P., *The Paleolithic of Siberia: New Discoveries and Interpretations*. 1998, Urbana: University of Illinois. 406.
100. Dikov, N.N., *The Ushki Sites: Kamchatka Peninsula*, in *American Beginnings: Prehistory and Paleoecology of Beringia*, F.H. West, Editor. 1996, University of Chicago: Chicago. p. 244-250.
101. Powers, W.R., *Siberia in the late glacial and early postglacial*, in *Humans at the End of the Ice Age: The Archaeology of the Pleistocene-Holocene Transition*, L.G. Strauss, et al., Editors. 1996, Plenum: New York. p. 229-242.
102. Aikens, C.M. and T. Higuchi, *The Prehistory of Japan*, ed. A. Press. 1982, New York. 354.
103. Aikens, C.M. and T. Akazawa, *The Pleistocene-Holocene transition in Japan and adjacent northeast Asia: Climate and biotic change, broad-spectrum diet, pottery, and sedentism*, in *Humans at the End of the Ice Age: The archaeology of the Pleistocene-Holocene Transition*, L.G. Strauss, et al., Editors. 1996, Plenum: New York. p. 216-228.
104. Stanford, D. and B. Bradley, *Ocean Trails and Prairie Paths? Thoughts about Clovis Origins*, in *The First Americans*, N.G. Jablonski, Editor. 2002, California Academy of Sciences: San Francisco. p. 255-271.
105. Lahr, M.M., *Patterns of modern human variation: Implications for Amerindian origins*. Yrbk. Phys. Anthropol., 1995. **38**: p. 163-198.
106. Powell, J.F., *Dental evidence for the peopling of the New World: some methodological considerations*. Hum Biol, 1993. **65**(5): p. 799-819.
107. Powell, J.F. and W.A. Neves, *Dental diversity of early New World populations: Taking a bite out of the tripartite model*. Am. J. Phys. Anthropol., 1998. **Suppl.26**: p. 169.
108. Powell, J.F. and J. Rose, *Report on the Osteological Assessment of the "Kennewick Man" Skeleton (CENWW.97.Kennewick)*. Discovering Archaeology, Special Report, 1999: p. 1-20.
109. Scott, G.R. and C.J. Turner II, *The Anthropology of Modern Human Teeth: Dental Morphology and its Variation in Recent Human Populations*. 1997, Cambridge: Cambridge University Press. 382.
110. Begley, S.A. and A. Murr, *The first Americans*. Newsweek, 1999: p. 50-57.
111. Booth, W., *Early migrants may have come by land and sea*. The Los Angeles Times, 1999: p. A13.

112. Egan, T., *Expert panel recasts origin of fossil man in northwest*. The New York Times, 1999: p. A11.
113. Preston, J., *The lost man*. The New Yorker, 1997: p. 69-81.
114. Wilford, J.N., *New answers to old questions: Who got here first?* The New York Times, 1999: p. D1, D4.
115. Turner, C.G.I., *Teeth, Needles, Dogs, and Siberia: Bioarchaeological Evidence for the Colonization of the New World*, in *The First Americans*, N.G. Jablonski, Editor. 2002, California Academy of Sciences: San Francisco. p. 123-158.
116. Dillehay, T.D., *Monte Verde: A Late Pleistocene Settlement in Southern Chile*. Vol. 1. 1989, Washington: Smithsonian Institution.
117. Dillehay, T.D., *Monte Verde*. Vol. 2. 1997, Washington: The Smithsonian Institution. 1071.
118. Fiedel, S.J., *Artifact provenience at Monte Verde: Confusion and contradictions*. Discovery Archaeology, Special Report, 1999: p. 1-12.
119. Lynch, T.F., *Glacial-age man in South America? A critical review*. Am. Antiquity, 1990. **55**: p. 199-228.
120. Meltzer, D.J., et al., *On the Pleistocene antiquity of Monte Verde, Southern Chile*. Am. Antiquity, 1997. **62**: p. 659-663.
121. Haynes, C.V., Jr., *Monte Verde and the pre-Clovis situation in America*. Discovering Archaeology, Special Report: Monte Verde Revisited, 1999: p. 17-19.
122. MacDonald, G.M., *Postglacial vegetation history of the MacKenzie River basin*. Quat. Res., 1987. **28**: p. 245-262.
123. Mandryk, C.S., *Paleoecology as Contextual Archaeology: Human Variability of the Late Quaternary Ice-Free Corridor, Alberta, Canada*. 1991, PhD Thesis, University of Alberta: Edmonton, Canada.
124. Clague, J.J., et al., *Quaternary geology of the Canadian cordillera*, in *Quaternary Geology of Canada and Greenland*, R.J. Fulton, Editor. 1989, Geological Society of Canada: Ottawa. p. 17-96.
125. Wright, H.E.J., *Environmental conditions for Paleoindian immigration*, in *The First Americans: Search and Research*, T.D. Dillehay and D.J. Meltzer, Editors. 1991, CRC Press: Boca Raton. p. 113-135.
126. Barrie, V.K., et al., *Submerged late Quaternary terrestrial deposits and paleoenvironment of northern Hecate Strait British Columbia continental shelf, Canada*. Quat. Int., 1993. **20**: p. 123-129.
127. Mann, D.H., *Wisconsin and Holocene glaciation of southeast Alaska*, in *Glaciation in Alaska: The Geological Record*, K.M. Reed and R.M. Thorson, Editors. 1986, Alaska Geological Society: Anchorage. p. 237-265.
128. Molnia, B.F., *Glacial history of the northeastern Gulf of Alaska-a synthesis*, in *Glaciation in Alaska: The Geological Record*, T.D. Hamilton, K.M. Reed, and R.M. Thorson, Editors. 1986, Alaska Geological Society: Anchorage. p. 219-236.
129. Mann, D.H. and D.M. Peteet, *Extent and timing of the Last Glacial Maximum in southwestern Alaska*. Quat. Res., 1994. **42**: p. 136-148.

130. Heaton, T.H. and F. Grady, *Fossil grizzly bears (Ursus arctos) from Prince of Whales Island, Alaska, offer new insights into animal dispersal, interspecific competition, and age of deglaciation*. *Curr. Res. Pleistocene*, 1993. **10**: p. 98-100.
131. Mann, D.H. and T.D. Hamilton, *Late Pleistocene and Holocene paleoenvironments of the North Pacific Coast*. *Quat. Sci. Rev.*, 1995. **14**: p. 449-471.
132. Dixon, E.J., et al., *Late Quaternary regional geoarchaeology of southern Alaska karst: A progress report*. *Geoarchaeology*, 1997. **12**: p. 689-712.
133. Josenhans, H.W., et al., *Early humans and rapidly changing Holocene sea levels in the Queen Charlotte Islands-Hecate Strait British Columbia, Canada*. *Science*, 1997. **277**: p. 71-74.
134. Klein, R.G., *Anatomy, behavior, and modern human origins*. *J. World Prehist.*, 1995. **9**: p. 167-241.
135. Singer, R. and J. Wymer, *The Middle Age at Klasies River Mouth in South Africa*. 1982, Chicago: University of Chicago Press. 234.
136. Rightmire, G.P. and H.J. Deacon, *Comparative studies of Late Pleistocene human remains from Klasies River Mouth, South Africa*. *J. Hum. Evol.*, 1991. **20**: p. 131-156.
137. Brooks, A.S., et al., *Dating and context of three Middle Stone Age sites with bone points in the Upper Semliki Valley, Zaire*. *Science*, 1995. **268**: p. 548-552.
138. Yellen, J.E., et al., *A Middle Stone Age worked bone industry from Katanda, Upper Semliki Valley, Zaire*. *Science*, 1995. **268**: p. 553-556.
139. Yellen, J.E., *Barbed bone points: Tradition and continuity in Saharan and sub-Saharan Africa*. *African Archaeol. Rev.*, 1998. **15**: p. 173-198.
140. Klein, R.G., *Why anatomically modern people did not disperse from Africa 100,000 years ago*, in *Neandertals and Modern Humans in Western Asia*, T. Akazawa, K. Aoki, and O. Bar-Yosef, Editors. 1998, Plenum: New York. p. 509-521.
141. Erlandson, J.M., *The archaeology of aquatic adaptations: Paradigm for a new millennium*. *J. Archaeol. Res.*, 2001. **9**: p. 287-350.
142. Bigalke, E.H., *The exploitation of shellfish by the coastal tribesmen of the Transkei*. *Annual of the Cape Province Museum*, 1973. **9**: p. 159-175.
143. Meehan, B., *Shell Bed to Shell Midden*. 1982, Canberra: Australian Institute of Aboriginal Studies. 189.
144. Lampert, R.J., *Burrill Lake and Currarong*. 1971, Canberra: Australian National University. 86.
145. Bowler, J.M. and A.G. Thorne, *Human remains from Lake Mungo*, in *The Origin of the Australians*, R.I. Kirk and A.G. Thorne, Editors. 1976, Humanities Press, NJ. p. 127-138.
146. Groube, L., et al., *A 40,000 year old human occupation site at Huon Peninsula, Papua New Guinea*. *Nature*, 1986. **324**: p. 453-435.
147. Roberts, R.G., R.H. Jones, and M.A. Smith, *Thermoluminescence dating of a 50,000 year old human occupation site in northern Australia*. *Nature*, 1990. **345**: p. 153-156.

148. Thorne, A., et al., *Australia's oldest human remains: Age of the Lake Mungo 3 skeleton*. *J. Hum. Evol.*, 1999. **36**: p. 591-612.
149. Flood, J., *Archaeology of a Dreamtime*. 1990, New Haven: Yale University Press. 302.
150. Allen, J., C. Gosden, and J.P. White, *Pleistocene dates for the human occupation of New Ireland, northern Melanesia*. *Nature*, 1988. **331**: p. 707-709.
151. Wickler, S. and M.J.T. Spriggs, *Pleistocene human occupation of the Solomon Islands, Melanesia*. *Antiquity*, 1988. **62**: p. 703-707.
152. Allen, J., C. Gosden, and J.P. White, *Human Pleistocene adaptations to the tropical-island Pacific: Recent evidence from New Ireland, a Greater Australian outlier*. *Antiquity*, 1989. **63**: p. 548-561.
153. Irwin, G., *The Prehistoric Exploration and Colonisation of the Pacific*. 1992, Cambridge: Cambridge University Press. 240.
154. Erlandson, J.M., *California's coastal prehistory: A circum-Pacific perspective*. *Proceedings of the Society for California Archaeology*, 1993. **6**: p. 23-36.
155. Matsu'ura, S., *A chronological review of Pleistocene human remains from the Japanese archipelago*, in *Interdisciplinary Perspectives on the Origins of the Japanese*, K. Omoto, Editor. 1996, International Research Center for Japanese Studies. p. 181-197.
156. Oda, S., *A review of archaeological research in the Izu and Ogasawara islands*. *Man and Cultures in Oceania*, 1990. **6**: p. 53-79.
157. Nettle, D., *Linguistic diversity of the Americas can be reconciled with a recent colonization*. *Proc Natl Acad Sci U S A*, 1999. **96**(6): p. 3325-9.
158. Nichols, J., *Linguistic Diversity in Space and Time*. 1992, Chicago: University of Chicago Press.
159. Nettle, D., *Linguistic Diversity*. 1999, Oxford: Oxford University Press.
160. Diamond, J., *Guns, Germs and Steele: The Fates of Human Societies*. 1997, London: Jonathan Cape.
161. Roberts, R.G., R.H. Jones, and M.A. Smith, *Nature*, 1990. **345**: p. 153-156.
162. Greenberg, J.H., C.J. Turner II, and S.L. Zegura, *The settlement of the Americas: A comparison of linguistic, dental, and genetic evidence*. *Curr. Anthropol.*, 1986. **27**: p. 477-497.
163. Greenberg, J.H., *Language in the Americas*. 1987, Stanford: Stanford University Press.
164. Nichols, J., *Linguistic diversity and the first settlement of the New World*. *Language*, 1990. **63**(3): p. 475-521.
165. Nichols, J., *The spread of language around the Pacific Rim*. *Evolutionary Anthropology*, 1995. **3**(6): p. 206-15.
166. Nichols, J. and D.A. Peterson, *The Amerind personal pronouns*. *Language*, 1996. **72**(2): p. 336-71.
167. Nichols, J., *Modeling ancient population structures and population movement in linguistics and archeology*. *Annual Review of Anthropology*, 1997. **26**: p. 359-84.

168. Nichols, J., *The origin and dispersal of languages: Linguistic evidence.*, in *The Origin and Diversification of Language*, N.G. Jablonski and L.C. Aiello, Editors. 1998, California Academy of Sciences: San Francisco. p. 127-70.
169. Nichols, J., *Estimating dates of early American colonization events.*, in *Time Depth in Historical Linguistics*, C. Renfrew, A. McMahon, and L. Trask, Editors. 2000, McDonald Institute for Archaeological Research: Cambridge. p. 643-664.
170. Nettle, D., *Using Social Impact Theory to simulate linguistic change.* *Lingua*, 1999. **108**: p. 95-117.
171. Dixon, R.M.W., *The Rise and Falls of Languages.* 1997, Cambridge: Cambridge Univ. Press.
172. Ammerman, A. and L.L. Cavalli-Sforza, in *The Explanation of Culture Change: Models in Prehistory*, C. Renfrew, Editor. 1973: London. p. 335-358.
173. Renfrew, C., *Archaeology and Language: The Puzzle of Indo-European Origins.* 1987, London: Jonathan Cape.
174. Renfrew, C., *Cambridge Archaeol. J.*, 1991. **1**: p. 3-23.
175. Steele, J., J.P. Adams, and T. Sluckin, *World Arch.*, 1998. **30**: p. 286-305.
176. Cohen, M.N., *The Food Crisis in Prehistory.* 1977, New Haven: Yale Univ. Press.
177. Jonhson, A. and T. Earle, *The Evolution of Human Societies: From Foraging Group to Agrarian State.* 1987, Stanford: Stanford Univ. Press.
178. Case, J.T. and D.C. Wallace, *Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblasts.* *Somatic Cell Genet*, 1981. **7**(1): p. 103-8.
179. Giles, R.E., et al., *Maternal inheritance of human mitochondrial DNA.* *Proc Natl Acad Sci U S A*, 1980. **77**(11): p. 6715-9.
180. Merriwether, D.A., et al., *The structure of human mitochondrial DNA variation.* *J Mol Evol*, 1991. **33**(6): p. 543-55.
181. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome.* *Nature*, 1981. **290**(5806): p. 457-65.
182. Wallace, D.C., K. Garrison, and W.C. Knowler, *Dramatic founder effects in Amerindian mitochondrial DNAs.* *Am J Phys Anthropol*, 1985. **68**(2): p. 149-55.
183. Wallace, D.C. and A. Torroni, *American Indian prehistory as written in the mitochondrial DNA: a review.* *Hum Biol*, 1992. **64**(3): p. 403-16.
184. Soodyall, H., et al., *mtDNA control-region sequence variation suggests multiple independent origins of an "Asian-specific" 9-bp deletion in sub-Saharan Africans.* *Am J Hum Genet*, 1996. **58**(3): p. 595-608.
185. Chen, Y.S., et al., *Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups.* *Am J Hum Genet*, 1995. **57**(1): p. 133-49.
186. Vigilant, L., et al., *Mitochondrial DNA sequences in single hairs from a southern African population.* *Proc Natl Acad Sci U S A*, 1989. **86**(23): p. 9350-4.
187. Vigilant, L., et al., *African populations and the evolution of human mitochondrial DNA.* *Science*, 1991. **253**(5027): p. 1503-7.

188. Redd, A.J., et al., *Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific*. Mol Biol Evol, 1995. **12**(4): p. 604-15.
189. Torroni, A., et al., *Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations*. Genetics, 1992. **130**(1): p. 153-62.
190. Torroni, A., et al., *Asian affinities and continental radiation of the four founding Native American mtDNAs*. Am J Hum Genet, 1993. **53**(3): p. 563-90.
191. Ballinger, S.W., et al., *Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations*. Genetics, 1992. **130**(1): p. 139-52.
192. Merriwether, D.A., *A mitochondrial perspective on the peopling of the New World*, in *The First Americans*, N.G. Jablonski, Editor. 2002, California Academy of Sciences: San Francisco.
193. Kolman, C.J., N. Sambuughin, and E. Bermingham, *Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders*. Genetics, 1996. **142**(4): p. 1321-34.
194. Lorenz, J.G. and D.G. Smith, *Distribution of sequence variation in the mtDNA control region of Native North Americans*. Hum Biol, 1997. **69**(6): p. 749-76.
195. Ginther, C., et al., *Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes*. Exs, 1993. **67**: p. 211-9.
196. Ward, R.H., et al., *Extensive mitochondrial diversity within a single Amerindian tribe*. Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8720-4.
197. Merriwether, D.A., et al., *Gene flow and genetic variation in the Yanomama as revealed by mitochondrial DNA*, in *America Past, America Present: Genes and Languages in the Americas and Beyond*, C. Renfrew, Editor. 2000, The McDonald Institute for Archaeological Research: Cambridge.
198. Merriwether, D.A. and K.F. A., *Mitochondrial Recombination? (Continued)*. Science, 1999. **285**: p. 837.
199. Easton, R.D., et al., *mtDNA variation in the Yanomami: evidence for additional New World founding lineages*. Am J Hum Genet, 1996. **59**(1): p. 213-25.
200. Merriwether, D.A. and R.E. Ferrell, *The four founding lineage hypothesis for the New World: a critical reevaluation*. Mol Phylogenet Evol, 1996. **5**(1): p. 241-6.
201. Brown, M.D., et al., *mtDNA haplogroup X: An ancient link between Europe/Western Asia and North America?* Am J Hum Genet, 1998. **63**(6): p. 1852-61.
202. Bailliet, G., et al., *Founder mitochondrial haplotypes in Amerindian populations*. Am J Hum Genet, 1994. **55**(1): p. 27-33.
203. Merriwether, D.A., et al., *mtDNA variation indicates Mongolia may have been the source for the founding population for the New World*. Am J Hum Genet, 1996. **59**(1): p. 204-12.

204. Torroni, A., et al., *Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude*. *Am J Phys Anthropol*, 1994. **93**(2): p. 189-99.
205. Merriwether, D.A., F. Rothhammer, and R.E. Ferrell, *Genetic variation in the New World: ancient teeth, bone, and tissue as sources of DNA*. *Experientia*, 1994. **50**(6): p. 592-601.
206. Bonatto, S.L. and F.M. Salzano, *A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data*. *Proc Natl Acad Sci U S A*, 1997. **94**(5): p. 1866-71.
207. Forster, P., et al., *Origin and evolution of Native American mtDNA variation: a reappraisal*. *Am J Hum Genet*, 1996. **59**(4): p. 935-45.
208. Stone, A.C. and M. Stoneking, *mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World*. *Am J Hum Genet*, 1998. **62**(5): p. 1153-70.
209. Stone, A.C. and M. Stoneking, *Analysis of ancient DNA from a prehistoric Amerindian cemetery*. *Philos Trans R Soc Lond B Biol Sci*, 1999. **354**(1379): p. 153-9.
210. Smith, D.G., et al., *Distribution of mtDNA haplogroup X among Native North Americans*. *Am J Phys Anthropol*, 1999. **110**(3): p. 271-84.
211. Lee, S.D., et al., *Sequence variation of mitochondrial DNA control region in Koreans*. *Forensic Sci Int*, 1997. **87**(2): p. 99-116.
212. Burkhardt, F., A. von Haeseler, and S. Meyer, *HvrBase: Compilation of mtDNA control region sequences from primates*. *Nucleic Acids Res.*, 1999. **27**: p. 138-142.
213. Thornton, R., *Population history of native North Americans*, in *A population History of North America*, M.R. Haines and R.H. Steckel, Editors. 2000, Cambridge University Press: Cambridge. p. 9-50.
214. Butzer, K.W., *The Americas Before and After 1492: An Introduction to Current Geographical Research*. *Annals of the Association of American Geographers*. **82**(3): p. 345-368.
215. Denevan, W.M., *The pristine myth: The landscape of the Americas in 1492*. *Annals of the Association of American Geographers*, 1992. **82**: p. 345-369.
216. Torroni, A., et al., *mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans*. *Am J Hum Genet*, 1993. **53**(3): p. 591-608.
217. Torroni, A., et al., *mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico*. *Am J Hum Genet*, 1994. **54**(2): p. 303-18.
218. Torroni, A., et al., *Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America*. *Proc Natl Acad Sci U S A*, 1994. **91**(3): p. 1158-62.
219. Bonatto, S.L. and F.M. Salzano, *Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World*. *Am J Hum Genet*, 1997. **61**(6): p. 1413-23.

220. Neel, J.V., R.J. Biggar, and R.I. Sukernik, *Virologic and genetic studies relate Amerind origins to the indigenous people of the Mongolia/Manchuria/southeastern Siberia region*. Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10737-41.
221. Lell, J.T., et al., *Y chromosome polymorphisms in native American and Siberian populations: identification of native American Y chromosome haplotypes*. Hum Genet, 1997. **100**(5-6): p. 536-43.
222. Karafet, T.M., et al., *Ancestral Asian source(s) of new world Y-chromosome founder haplotypes*. Am J Hum Genet, 1999. **64**(3): p. 817-31.
223. Schurr, T.G., et al., *Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic*. Am J Phys Anthropol, 1999. **108**(1): p. 1-39.
224. Starikovskaya, Y.B., et al., *mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of Ancient Beringia and the peopling of the New World*. Am J Hum Genet, 1998. **63**(5): p. 1473-91.
225. Nichols, J., *The First American Languages*, in *The First Americans*, N.G. Jablonski, Editor. 2002, California Academy of Sciences: Sans Francisco. p. 273-293.
226. Jakubiczka, S., et al., *A search for restriction fragment length polymorphism on the human Y chromosome*. Hum Genet, 1989. **84**(1): p. 86-8.
227. Malaspina, P., et al., *The human Y chromosome shows a low level of DNA polymorphism*. Ann Hum Genet, 1990. **54** (Pt 4): p. 297-305.
228. Ngo, K.Y., et al., *A DNA probe detecting multiple haplotypes of the human Y chromosome*. Am J Hum Genet, 1986. **38**(4): p. 407-18.
229. Seielstad, M.T., et al., *Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition*. Hum Mol Genet, 1994. **3**(12): p. 2159-61.
230. Spurdle, A. and T. Jenkins, *Y chromosome probe p49a detects complex PvuII haplotypes and many new TaqI haplotypes in southern African populations*. Am J Hum Genet, 1992. **50**(1): p. 107-25.
231. Persichetti, F., et al., *Disequilibrium of multiple DNA markers on the human Y chromosome*. Ann Hum Genet, 1992. **56** (Pt 4): p. 303-10.
232. Santos, F.R., S.D. Pena, and J.T. Epplen, *Genetic and population study of a Y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique*. Hum Genet, 1993. **90**(6): p. 655-6.
233. Santos, F.R., et al., *Geographic differences in the allele frequencies of the human Y-linked tetranucleotide polymorphism DYS19*. Hum Genet, 1996. **97**(3): p. 309-13.
234. Underhill, P.A., et al., *Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography*. Genome Res, 1997. **7**(10): p. 996-1005.
235. Underhill, P.A., et al., *A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history*. Proc Natl Acad Sci U S A, 1996. **93**(1): p. 196-200.
236. Hammer, M.F. and S. Horai, *Y chromosomal DNA variation and the peopling of Japan*. Am J Hum Genet, 1995. **56**(4): p. 951-62.

237. Underhill, P.A., et al., *Y chromosome sequence variation and the history of human populations*. Nat Genet, 2000. **26**(3): p. 358-61.
238. Shen, P., et al., *Population genetic implications from sequence variation in four Y chromosome genes*. Proc Natl Acad Sci U S A, 2000. **97**(13): p. 7354-9.
239. Bianchi, N.O., et al., *Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers*. Am J Phys Anthropol, 1997. **102**(1): p. 79-89.
240. Bianchi, N.O., et al., *Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations*. Am J Hum Genet, 1998. **63**(6): p. 1862-71.
241. Karafet, T., et al., *Y chromosome markers and Trans-Bering Strait dispersals*. Am J Phys Anthropol, 1997. **102**(3): p. 301-14.
242. Ruiz-Linares, A., et al., *Microsatellites provide evidence for Y chromosome diversity among the founders of the New World*. Proc Natl Acad Sci U S A, 1999. **96**(11): p. 6312-7.
243. Lell, J.T., et al., *The dual origin and Siberian affinities of Native American Y chromosomes*. Am J Hum Genet, 2002. **70**(1): p. 192-206.
244. Semino, O., et al., *The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective*. Science, 2000. **290**(5494): p. 1155-9.
245. Kappes, D. and J.L. Strominger, *Human class II major histocompatibility complex genes and proteins*. Annu Rev Biochem, 1988. **57**: p. 991-1028.
246. Terasaki, P.I., et al., *Gene frequencies and maps of their distributions.*, in *HLA 1997*, P.I. Terasaki and D.W. Gjertson, Editors. 1997, UCLA Tissue Typing Laboratory: Los Angeles. p. 427-460.
247. Arnaiz-Villena, A., et al., *HLA genes in Mexican Mazatecans, the peopling of the Americas and the uniqueness of Amerindians*. Tissue Antigens, 2000. **56**(5): p. 405-16.
248. Uinuk-Ool, T.S., et al., *Origin and affinities of indigenous Siberian populations as revealed by HLA class II gene frequencies*. Hum Genet, 2002. **110**(3): p. 209-26.
249. Goicoechea, A.S., et al., *New genetic data on Amerindians from the Paraguayan Chaco*. Am J Human Biol, 2001. **13**(5): p. 660-7.
250. Luiselli, D., et al., *Genetic structure of Quechua-speakers of the Central Andes and geographic patterns of gene frequencies in South Amerindian populations*. Am J Phys Anthropol, 2000. **113**(1): p. 5-17.
251. Ferreira, R.G., et al., *Ethnic admixture composition of two western Amazonian populations*. Hum Biol, 2002. **74**(4): p. 607-14.
252. Sokal, R.R. and N.L. Oden, *Spatial autocorrelation in biology - 2. Some biological implications and four applications of evolutionary and ecological interest*. Biol J Linn Soc, 1978. **10**: p. 228-249.
253. Sokal, R.R. and N.L. Oden, *Spatial autocorrelation in biology - 1. Methodology*. Biol J Linn Soc, 1978. **10**: p. 199-228.
254. Barbujani, G., *Autocorrelation of gene frequencies under isolation by distance*. Genetics, 1987. **117**(4): p. 777-82.

255. Crawford, M.H., *The origin of Native Americans. Evidence from archaeological genetics*. 1998, Cambridge: Cambridge University Press.
256. Batista dos Santos, S.E., et al., *Differential contribution of indigenous men and women to the formation of an urban population in the Amazon region as revealed by mtDNA and Y-DNA*. Am J Phys Anthropol, 1999. **109**(2): p. 175-80.
257. Mingroni-Netto, R.C., et al., *Distribution of CGG repeats and FRAXAC1/DXS548 alleles in South American populations*. Am J Med Genet, 2002. **111**(3): p. 243-52.
258. Duran, R. and M. Ruiz-Garcia, *Genetic population history relationships of the population of Bogota, Colombia, by using the DIS80, VWA, and TH01 molecular markers*. Am J Human Biol, 2001. **13**(3): p. 374-83.
259. Rupert, J.L., et al., *Beta-fibrinogen allele frequencies in Peruvian Quechua, a high-altitude native population*. Am J Phys Anthropol, 1999. **109**(2): p. 181-6.
260. Gaspar, P.A., et al., *Polymorphisms of CYP1A1, CYP2E1, GSTM1, GSTT1, and TP53 genes in Amerindians*. Am J Phys Anthropol, 2002. **119**(3): p. 249-56.
261. Gaspar, P.A., et al., *TP53 polymorphisms and haplotypes in South Amerindians and neo-Brazilians*. Ann Hum Biol, 2001. **28**(2): p. 184-94.
262. Gaspar, P.A., et al., *High frequency of CYP1A1*2C allele in Brazilian populations*. Hum Biol, 2002. **74**(2): p. 235-42.
263. de Andrade, F.M., et al., *High heterogeneity of apolipoprotein E gene frequencies in South American Indians*. Ann Hum Biol, 2000. **27**(1): p. 29-34.
264. De Andrade, F.M., et al., *Lipoprotein lipase and APOE/APOC-I/APOC-II gene cluster diversity in native Brazilian populations*. Am J Human Biol, 2002. **14**(4): p. 511-8.
265. Calafell, F., et al., *Genealogy reconstruction from short tandem repeat genotypes in an Amazonian population*. Am J Phys Anthropol, 1999. **108**(2): p. 137-46.
266. Fisher, M.C., et al., *Biogeographic range expansion into South America by Coccidioides immitis mirrors New World patterns of human migration*. Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4558-62.
267. Sugimoto, C., et al., *JC virus strains indigenous to northeastern Siberians and Canadian Inuits are unique but evolutionally related to those distributed throughout Europe and Mediterranean areas*. J Mol Evol, 2002. **55**(3): p. 322-35.
268. Hatwell, J.N. and P.M. Sharp, *Evolution of human polyomavirus JC*. J Gen Virol, 2000. **81**(Pt 5): p. 1191-200.
269. Kunitake, T., et al., *Parent-to-child transmission is relatively common in the spread of the human polyomavirus JC virus*. J Clin Microbiol, 1995. **33**(6): p. 1448-51.
270. Kato, A., et al., *Lack of evidence for the transmission of JC polyomavirus between human populations*. Arch Virol, 1997. **142**(5): p. 875-82.
271. Salzano, F.M., *Molecular variability in Amerindians: widespread but uneven information*. An Acad Bras Cienc, 2002. **74**(2): p. 223-63.

272. Zago, M.A., et al., *Interpopulational and intrapopulational genetic diversity of Amerindians as revealed by six variable number of tandem repeats*. Hum Hered, 1996. **46**(5): p. 274-89.
273. Zietkiewicz, E., et al., *Nuclear DNA diversity in worldwide distributed human populations*. Gene, 1997. **205**(1-2): p. 161-71.
274. Zietkiewicz, E., et al., *Genetic structure of the ancestral population of modern humans*. J Mol Evol, 1998. **47**(2): p. 146-55.
275. Kolman, C.J., et al., *Reduced mtDNA diversity in the Ngobe Amerinds of Panama*. Genetics, 1995. **140**(1): p. 275-83.
276. Scozzari, R., et al., *Differential structuring of human populations for homologous X and Y microsatellite loci*. Am J Hum Genet, 1997. **61**(3): p. 719-33.
277. Zietkiewicz, E., et al., *Single-strand conformational polymorphisms (SSCP): detection of useful polymorphisms at the dystrophin locus*. Hum Genet, 1992. **89**(4): p. 453-6.
278. Zietkiewicz, E., N. Akalin, and D. Labuda, *Neutral polymorphisms in the deletion-prone regions of the dystrophin gene*. Hum Hered, 1995. **45**(2): p. 80-3.
279. Labuda, D., E. Zietkiewicz, and V. Yotova, *Archaic lineages in the history of modern humans*. Genetics, 2000. **156**(2): p. 799-808.
280. Stephens, M., N.J. Smith, and P. Donnelly, *A new statistical method for haplotype reconstruction from population data*. Am J Hum Genet, 2001. **68**(4): p. 978-89.
281. Lin, S., et al., *Haplotype inference in random population samples*. Am J Hum Genet, 2002. **71**(5): p. 1129-37.
282. Underhill, P.A., et al., *The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations*. Ann Hum Genet, 2001. **65**(Pt 1): p. 43-62.
283. Arens, R., *"Manhunt." Genocide in Paraguay*. 1976, Philadelphia: Temple University Press.
284. Davis, S., *Land rights and indigenous peoples*. 1988, Cambridge: Cultural Survival Inc.
285. Hill, K. and M. Hurtado, *Ache Life History*. 1996, New York: Aldine De Gruyter.
286. Münzel, M., *The Aché Indians: Genocide in Paraguay*. 1973, Copenhagen: Iwgia.
287. Münzel, M., *The Ache: Genocide Continues in Paraguay*. 1974, Copenhagen: Iwgia.
288. Long, J.C., et al., *Genetic variation in Arizona Mexican Americans: estimation and interpretation of admixture proportions*. Am J Phys Anthropol, 1991. **84**(2): p. 141-57.
289. Cerda-Flores, R.M., et al., *Gene diversity and estimation of genetic admixture among Mexican-Americans of Starr County, Texas*. Ann Hum Biol, 1992. **19**(4): p. 347-60.
290. Morera, B., R. Barrantes, and R. Marin-Rojas, *Gene admixture in the Costa Rican population*. Ann Hum Genet, 2003. **67**(Pt 1): p. 71-80.

291. Cerda-Flores, R.M., et al., *Maximum likelihood estimates of admixture in Northeastern Mexico using 13 short tandem repeat loci*. Am J Human Biol, 2002. **14**(4): p. 429-39.
292. Cerda-Flores, R.M., et al., *Genetic admixture in three Mexican Mestizo populations based on DIS80 and HLA-DQA1 loci*. Am J Human Biol, 2002. **14**(2): p. 257-63.
293. Yunis, J.J., E.J. Yunis, and E. Yunis, *Genetic relationship of the Guambino, Paez, and Ingano Amerindians of southwest Colombia using major histocompatibility complex class II haplotypes and blood groups*. Hum Immunol, 2001. **62**(9): p. 970-8.
294. Probst, C.M., et al., *HLA polymorphism and evaluation of European, African, and Amerindian contribution to the white and mulatto populations from Parana, Brazil*. Hum Biol, 2000. **72**(4): p. 597-617.
295. Sans, M., *Admixture studies in Latin America: from the 20th to the 21st century*. Hum Biol, 2000. **72**(1): p. 155-77.
296. Sans, M., F.M. Salzano, and R. Chakraborty, *Historical genetics in Uruguay: estimates of biological origins and their problems*. Hum Biol, 1997. **69**(2): p. 161-70.
297. Castro de Guerra, D., et al., *Genetic study in Panaquire, a Venezuelan population*. Hum Hered, 1996. **46**(6): p. 323-8.
298. Yunis, J.J., et al., *Major histocompatibility complex class II alleles and haplotypes and blood groups of four Amerindian tribes of northern Colombia*. Hum Immunol, 1994. **41**(4): p. 248-58.
299. Malhi, R.S. and D.G. Smith, *Brief communication: Haplogroup X confirmed in prehistoric North America*. Am J Phys Anthropol, 2002. **119**(1): p. 84-6.
300. Fischer, G.F., et al., *An HLA class-II allele frequent in Eskimos and Amerindians is found in the Tyrolean Ice Man*. Ann Hum Genet, 2001. **65**(Pt 4): p. 363-9.
301. Cavalli-Sforza, L.L., P. Menozzi, and A. Piazza, *The History and Geography of Human Genes*. 1994, Princeton: Princeton University Press.
302. Kivisild, T., et al., *Deep common ancestry of indian and western-Eurasian mitochondrial DNA lineages*. Curr Biol, 1999. **9**(22): p. 1331-4.
303. Lahr MM, a.F.R., *Multiple dispersals and modern human origins*. Evol Anthropol, 1994. **3**: p. 48-60.
304. Chu, J.Y., et al., *Genetic relationship of populations in China*. Proc Natl Acad Sci U S A, 1998. **95**(20): p. 11763-8.