



UNIVERSITÉ MOHAMMED V – AGDAL  
FACULTÉ DES SCIENCES  
RABAT



N° d'ordre : 2678

**THÈSE de DOCTORAT**

Présentée par :

**Rachid ABOUKHALID**

Discipline : Biologie

Spécialité : Génétique médico-légale, des populations et biologie moléculaire

**Caractérisation anthropo-génétique d'un échantillon de la population marocaine en utilisant les marqueurs uni-parentaux (Chromosome Y et ADN mitochondrial)**

Soutenue le **9 Novembre 2013** devant le jury composé de :

**Président :**

**Pr. Saaïd AMZAZI**

**Professeur de l'enseignement Supérieur (PES)  
Faculté des Sciences Rabat**

**Examineurs :**

**Pr. Hassan JAZIRI**

**Professeur de l'enseignement supérieur (PES)  
Ecole Normale Supérieur Rabat**

**Pr. Touria BENZAOU**

**Professeur de l'enseignement Supérieur (PES)  
Faculté des Sciences Rabat**

**Pr. Abdeljalil BOUZOUGGAR**

**Professeur Habilité (PH)  
Institut Nationale des Sciences de l'Archéologie  
et du Patrimoine de Rabat**

**Pr. Abdelaziz CHAFIK**

**Professeur de l'enseignement Supérieur (PES)  
Faculté Polydisciplinaire d'El Jadida**

**Invité :**

**Dr. Jala TALBI**

**Laboratoire de la Police Scientifique (LPS) de  
Casablanca**

**À mes parents**

**À mes sœurs**

**À ma Famille**

**À mes amis (es)**

## Avant-propos

*Le présent travail est le fruit d'une collaboration étroite entre l'Unité de génétique médico-légale du laboratoire de Biochimie et d'immunologie de la Faculté des Sciences de Rabat et divers laboratoires et organisme (Maroc et Etats-Unis):*

- *Unité de Biologie et de Recherche Médicale du Centre Nationale de l'Énergie des Sciences et des Techniques Nucléaires (CNESTEN, Rabat).*
- *Unité de recherche de l'ADN mitochondriale du Laboratoire d'Identification génétique des Forces Armées américaines (AFDIL, Dover, Delaware, USA).*
- *Laboratoire des Sciences Médico-légales de l'université George Washington (Washington, District of Columbia, USA).*
- *GENOME Biotechnologies.*

*Cette thèse fait partie intégrante du projet d'établissement de la base de données pour les empreintes génétiques. Ce travail concerne les résultats des marqueurs uni-parentaux (ADN mitochondrial et chromosome Y) chez un échantillon représentatif de la population marocaine qui renferme trois groupes (Arabophones, Berbérophones, et Sahraouis).*

*Ce travail n'aurait vu le jour sans la confiance et la patience de mon directeur de thèse **Monsieur Saïd AMZAZI**, Professeur et doyen à la faculté des sciences de Rabat, pour m'avoir donné la chance de mener une thèse dans les meilleures conditions. Pour avoir dirigé de près les travaux présentés dans cette thèse, pour son soutien en me laissant travailler très librement, ses conseils et ses encouragements. Les qualités humaines dont vous bénéficiez me touchant à jamais, je vous suis redevant. Veuillez trouver ici le fruit de votre confiance.*

*J'adresse mes sincères remerciements à **Madame Touria BENAZZOU**, Professeur à la Faculté des Sciences de Rabat, d'avoir accepté d'être rapporteur de cette thèse et de critiquer ce travail. Je suis particulièrement honorée de votre présence dans ce jury de soutenance. Je vous exprime ici ma plus haute considération.*

*Je remercie vivement **Monsieur Hassane JAZIRI**, Professeur et Directeur de l'Ecole Normale Supérieur de Rabat (ENS) pour l'attention que vous m'avez portée au tout début de ma recherche et de l'honneur que vous me faites en acceptant la charge de rapporteur de ma thèse.*

Mes plus sincères remerciements vont pour **Monsieur Abdeljalil BOUZOUGAR**, Professeur à l'Institut National des Sciences de l'Archéologie et du Patrimoine (INSAP), Chef du département d'archéologie de la préhistoire et chercheur associé au département de l'évolution humaine à l'Institut Max Planck (Lepzig, Allemagne) qui m'a fait l'honneur de juger mon travail de thèse. La grande estime et admiration que je lui porte m'ont naturellement conduit à solliciter son jugement.

Je remercie infiniment **Monsieur Abdelaziz CHAFIK**, Professeur et doyen de la Faculté Polydisciplinaire d'El Jadida, pour avoir aimablement accepté de lire mon manuscrit et aussi de se déplacer volontierement pour juger ce travail.

Je tiens à remercier **Docteur Jalal TALBI**, chercheur au laboratoire de la Police scientifique de Casablanca, d'avoir accepté d'honorer ce jury, pour sa sympathie constante et de l'intérêt qu'il a manifesté pour cette thèse.

Je souhaite exprimer ma profonde reconnaissance à **Monsieur le Colonel Louis FINELLI**, responsable du laboratoire d'identification génétique des forces armées américaines (AFDIL), de m'avoir donné l'opportunité de réaliser mes travaux de recherche sous la direction du responsable de l'unité de recherche de l'ADN mitochondrial **Dr. Michael D. COBLE**, qui m'a accueilli chaleureusement avec les membres de son équipe : **Dr. Jodi A. Irwin**, **Mme Kimberly ANDREAGI STURK**, **Dr. Odile LOREILLE**, **Dr. Rebeca JUST**, **Mme Jessica SAUNIER**, **Mme Toni Marie DIEGOLI** et **Mme Melissa SCHEIBLÉ**, que je remercie très sincèrement et qui m'ont donné les moyens de mener à bien ma recherche. J'ai apprécié aussi la qualité des passionnantes discussions scientifiques et vos orientations qui ont contribué à mon épanouissement scientifique.

Je remercie **Monsieur Daniele PODINI**, Professeur et responsable du laboratoire des Sciences médico-légale à l'université **GEORGE WASHINGTON** pour m'avoir accueilli au sein de son laboratoire et pour m'avoir accordé sa confiance en m'offrant la possibilité de poursuivre ce travail.

J'aimerais également remercier **Docteur Driss SQUALLI HOUSSAINI**, directeur de la société **GENOME Biotechnologies**, pour le soutien et la qualité de ses relations humaines et discussions constructives qu'on a échangées.

Je tiens à exprimer ma reconnaissance et mes sincères remerciements à mon ami et frère **Monsieur Mehdi BOUABDELLAH** qui m'a encouragé à contribuer dans le projet et soutenu dans mes moments d'enthousiasme étincelant comme dans les moments de doute inconsolables.

Je remercie le Docteur **Mohammed MZIBRI** pour m'avoir chaleureusement accueilli au sein de son laboratoire de l'Unité de biologie et de recherche médicale du Centre Nationale de l'Énergie des Sciences et des Techniques Nucléaires (CNESTEN, Rabat) et pour m'avoir accordé sa confiance en m'offrant la possibilité d'entamer mon sujet de thèse et son soutien tout au long de ce travail. **Un clin d'œil à ce cher ABI 310...**

Un grand merci à ceux qui étaient mes collègues mais sont aujourd'hui des amis (es) précieux, **Dr. Hicham EL OSSMANI, Monsieur Hafid NATEQ, Dr. Hassan BENHESSOU, Mlle Falak AZZAM, Pr. Laila SBABOU, Mlle Imane BENNIS, Mlle Hanane BELLAÏOU, Mme Badia FOUJIL, Dr. Rachid CHATER, Monsieur Mohammed AL GHAMARTI, Monsieur Abdellah BERRADA et Mlle Noâma BERRADA**. Ces personnes m'ont inlassablement encouragé, supporté et soutenu en participant aux moments joyeux et aux épreuves dures de ma vie.

Je ne puis citer tous ceux qui de près ou de loin ont contribué à mon travail mais je ne les oublie pas. Je ne peux pas ignorer que la réalisation de cette thèse de doctorat n'aurait jamais été possible sans leur contribution. Merci à toutes et à tous!

Mes remerciements les plus affectueux reviennent à mes parents et à toute ma famille. Merci pour vos soutiens, moral et financier, de tous les instants. Merci pour votre écoute, votre patience, vos conseils, votre réconfort et pour tous ces moments de vie.

Ces études ont bénéficié de l'aide financière et scientifique des différents collaborateurs cités ci-dessus ; Une partie de ce travail a été financée par **US National Institute of Justice** «Interagency Agreement 2005-DN-R-086». Je tiens à exprimer ma sincère gratitude à la commission marocaine-américaine pour l'éducation et l'échange culturel (MACECE) qui, en collaboration avec le programme Fulbright, m'a permis de faire une partie de ma thèse aux USA.

# Table des matières

*Liste des publications*

*Liste des communications*

*Résumé*

*Abstract*

*Liste des abréviations*

*Liste des figures*

*Liste des tableaux*

## **Introduction générale**

### **Partie I : Données de la littérature**

#### **Chapitre A**

*Le Maroc : son origine à nos jours*

I. Histoire du Maroc -----	1
I.1. Les estampilles de peuplement au Maroc du Paléolithique inférieur au Néolithique-----	1
I.2. Le Maroc antique-----	4
I.3. Conquête arabo-musulmane et islamisation du Maroc -----	6
I.4. Soubassement du Maroc -----	7

#### **Chapitre B**

*Concepts et génétique des populations*

I. La génétique des populations -----	9
I.1. Transmission des caractères-----	9
I.2. Définition et objectifs -----	9
I.3. Équilibre de Hardy-Weinberg et les forces évolutives -----	10
II. Le polymorphisme génétique -----	12
II.1. Les modifications à la source du polymorphisme génétique-----	13
II.2. Les différents polymorphismes -----	15

III. Les marqueurs haplotypiques -----	16
--	----

## ***Chapitre C***

### *Les marqueurs haplotypiques*

## **L'ADN MITOCHONDRIAL**

I. Généralités sur l'ADN mitochondrial -----	18
I.1. Définition -----	18
I.2. Propriétés et avantages de l'ADNmt -----	20
II. L'hétéroplasmie: cas particulier rencontré au cours de l'analyse de l'ADNmt -----	21
II.1. Hétéroplasmie ponctuelle -----	21
II.2. Hétéroplasmie de taille -----	22
III. Exemples d'utilisation de l'ADNmt -----	22
III.1. Les origines de l'homme moderne -----	23
III.1. Les études paléanthropologiques -----	24
III.1. L'ADN mitochondrial de l'intérêt scientifique à la pratique judiciaire -----	25
IV. Les haplogroupes mitochondriaux -----	25

## **LE CHROMOSOME Y**

I. Généralités sur le chromosome Y -----	29
I.1. Évolution du chromosome Y -----	30
I.2. Structure actuelle du chromosome Y -----	31
I.3. Diversité du chromosome Y -----	33
I.3.1. Les haplotypes du chromosome Y -----	34
I.3.2. Les haplogroupes du chromosome Y -----	34
I.3.3. Intérêt du chromosome Y en généalogie et en génétique médico-légale -----	35

## Partie II: Matériel & Méthodes

I. Objectif du travail-----	38
II. Echantillonnage-----	38
III. De l'ADN dans l'échantillon à l'ADN dans le Tube-----	38
III.1. Extraction des acides nucléiques-----	38
III.2. Quantification de l'ADN par PCR en temps Réel-----	39
IV. Analyse du polymorphisme des marqueurs génétique du chromosome Y-----	40
IV.1. Technologie du génotypage des STRs du chromosome Y-----	41
IV.1.1. La PCR multiplex du kit AmpFISTR®Yfiler™ (Applied Biosystems®)-----	42
IV.1.2. L'électrophorèse capillaire-----	42
IV.1.3. L'obtention des profils Y-STRs-----	44
IV.1.4. Les outils informatiques-----	47
<i>i. La structure génétique</i> -----	47
<i>ii. Les distances entre les populations</i> -----	48
IV.1.5. Contrôle de qualité par YHRD-----	47
IV.2. Technologie du génotypage des marqueurs bi-alléliques du chromosome Y-----	50
IV.2.1. Sélection des marqueurs bi-allélique du chromosome Y-----	52
IV.2.2. La conception des amorces-----	52
<i>i. Pour la PCR</i> -----	52
<i>ii. Pour l'extension d'amorce (Primer Extension, PEX)</i> -----	53
IV.2.3. L'amplification PCR multiplex et purification-----	53
IV.2.4. La réaction du miniséquençage multiPEX et détection de la fluorescence-----	54
V. Mise en évidence du polymorphisme de l'ADN Mitochondrial-----	55
V.1. Le séquençage de l'ADN mitochondrial-----	55
V.2. Stratégie d'analyse de la région contrôle de l'ADNmt de notre population étudié-----	55
V.2.1. Amplification et séquençage de la région contrôle de l'ADNmt-----	55
V.2.2. Analyse des données et contrôle qualité-----	58
V.2.3. L'analyse statistique des données-----	58

## **Partie III : Résultats et Discussion**

<b>I. Analyse du polymorphisme des marqueurs du chromosome Y</b> -----	59
I.1. Les marqueurs STR-----	59
I.2. Les marqueurs bi-allélique (Y-SNPs)-----	72
I.2.1. Analyse de l'haplogroupe E-----	73
i. L'haplogroupe E1b1b1 (E-M35)-----	74
ii. L'haplogroupe E1 (E-P147)-----	78
I.2.2. L'analyse de l'haplogroupe J-----	78
<b>II. Analyse du polymorphisme de l'ADN mitochondriale</b> -----	80
II.1. L'Hétéroplasmie ponctuelle et de taille-----	82
II.2. Duplication et nouvelles mutations-----	82
II.3. Haplogroupes présents dans notre population-----	83
II.4. L'haplogroupe détectés principalement dans les populations nord-africaines : U6-----	85
II.5. Les haplogroupes fréquents en Eurasie-----	86
II.6. Les haplogroupes fréquents dans les populations sub-sahariennes-----	86

*Conclusions et Perspectives*

*Références Bibliographiques*

*Articles*

*Annexe*

# LISTE DES PUBLICATIONS

- **R. Aboukhalid**, K. Sturk-Andreaggi, M. Bouabdellah, D. Squalli, J. A. Irwin, S. Amzazi. Mitochondrial DNA control region variation from samples of the Moroccan population. *International Journal of Legal Medicine*. 2013 Jan 3.
- Kaoutar Bentayebi, Antònia Picornell, Mehdi Bouabdellah, José Aurelio Castro, **Rachid Aboukhalid**, Driss Squalli, Maria Misericòrdia, Saaïd Amzazi. Genetic diversity of 12 X-chromosomal short tandem repeats in the Moroccan population. *Forensic Science International: Genetics* 6 (2012) e48–e49.
- Picornell, K. Bentayebi, M. Bouabdellah, **R. Aboukhalid**, D. Squalli, J.A. Castro, M.M. Ramon, S. Amzazi. X-chromosomal haplotype frequencies of four linkage groups in a North African population. *Forensic Science International: Genetics Supplement Series* 3 (2011) e19–e20.
- **Rachid Aboukhalid**, Mehdi Bouabdellah, Ronald Lai, Nicholas Bradford, Hicham EL Ossmani, Jalal Talbi, Kaoutar Bentayebi, Driss Squalli, Daniele Podini and Saaïd Amzazi. Y Chromosomal SNP Analysis Using the Minisequencing Strategy in a Moroccan Population Samples. *Journal of Forensic Research* (2010) doi:10.4172/2157-7145.1000113.
- **Rachid Aboukhalid**, Mehdi Bouabdellah, Meriame Abbassi, Kaoutar Bentayebi, Mohammed Elmzibri, Driss Squalli, Saaïd Amzazi. Haplotype frequencies for 17 Y-STR loci (AmpFISTR®Y-filer™) in a Moroccan population sample. *Forensic Science International: Genetics* 4 (2010) e73–e74.
- Hicham El Ossmani, Brahim Bouchrif, **Rachid Aboukhalid**, Mehdi Bouabdellah, Bouchaib Gazzaz, Driss Zaoui, Abdelaziz Chafik, Jalal Talbi. Assessment of phylogenetic structure of Berber-speaking population of Azrou using 15 STRs of Identifiler kit. *Journal of Legal medicine* (2009) 52-56, doi: 10.1016.
- M. Bouabdellah, F. Ouenzar, **R. Aboukhalid**, M. Elmzibri, D. Squalli, S. Amzazi. STR data for the 15 AmpFISTR Identifiler loci in the Moroccan population *Forensic Science International: Genetics Supplement Series* 1 (2008) 306–308.

## **Liste des communications**

1. ***Communication affichée*** : **STR data for 15 AmpFLSTR Identifier loci in the Moroccan population** au 22<sup>ème</sup> congrès de la société internationale de génétique forensique. 2007. Copenhagen ; Danmark.
2. ***Communication affichée*** : **Y Chromosomal SNP Analysis Using the Minisequencing Strategy in a Moroccan Population Samples** au congress annuelle de l'American Academy of Forensic Sciences. 2009. Dover; USA.
3. ***Communication affichée*** : **Haplotype frequencies for 17 Y-STR loci (AmpFISTR1Y-filerTM) in a Moroccan population sample** au 23<sup>ème</sup> congrès de la société internationale de génétique forensique. 2009. Buenos–Aires ; Argentine.
4. ***Communication affichée*** : **Mitochondrial DNA control region variation from samples of the Moroccan population** au 24<sup>ème</sup> congrès de la société internationale de génétique forensique. 2011. Vienne ; Autriche.
5. ***Communication Orale*** : **Uni-parental markers in human identity testing including forensic DNA analysis and evolutionary genetic for Moroccan population samples** au 2ème congrès Internationale des sciences medico-légale. 2013. Las-Vegas ; USA.

## **Résumé :**

Les marqueurs génétiques STR (Short Tandem Repeat) à transmission parentale ont été utilisés en identité humaine il y'a environ 25 ans. Plus récemment, les marqueurs uniparentaux comme l'ADN mitochondrial (hérité de la mère) et le chromosome Y (hérité du père) ont été introduits dans les tests d'identification humaine en vue qu'ils échappent à la recombinaison méiotique. Ainsi, contrairement aux autosomes ou au chromosome X, tous les ADN mitochondriaux présents dans les populations humaines peuvent provenir d'un ancêtre maternel commun et les chromosomes Y, d'un ancêtre paternel commun. Ces derniers ont donc été utilisés pour préciser géographiquement et historiquement l'origine de nos ancêtres communs les plus récents et les interactions entre les différentes populations humaines. Ce présent travail décrit la diversité génétique d'un échantillonnage de la population marocaine refermant les trois groupes ethniques (Arabophones, Berbérophones et Sahraouis) selon deux types de polymorphismes : certains marqueurs microsatellites et bi-alléliques du chromosome Y et la région contrôle de l'ADN mitochondrial. Notre étude a pour objectif d'établir une base de données pour les marqueurs uniparentaux (YHRD et EMPOP) et retracer l'origine et l'histoire génétique de la population du Nord-Ouest d'Afrique (Maroc) en s'appuyant sur des données archéologiques et historiques. Pour l'ensemble des marqueurs, nos résultats montrent une composition génétique mosaïque issue d'une population très ancienne provient peut-être du Paléolithique et/ou du Néolithique. Une très nette différenciation génétique est observée entre notre population marocaine étudiée et une population d'Egypte; une distinction moins importante est rencontrée entre les différents groupes de notre population.

**Mots-clés :** Identité humaine, génétique des populations, ADN mitochondrial, Chromosome Y, histoire génétique du Maroc.

**Abstract:**

Bi-parentally inherited short tandem repeat polymorphisms have been routinely used in human identity testing for about 25 years. More recently, the analysis of uni-parentally inherited markers from the human Y chromosome and mitochondrial DNA was introduced for special applications of DNA-based human identification. Thus, unlike the autosomes and the X chromosome, all mitochondrial DNA present in human populations may come from a common maternal ancestor and Y chromosomes, a common paternal ancestor. These have been used to identify geographically and historically the origin of our most recent common ancestors and the interactions between different human populations. The present work describes the genetic variation of the Moroccan population samples belonging to the three ethnic groups (Arabic-speaking, Berber speaking and Sahrawi) using two types of polymorphisms: Y-STR and bi-allelic markers and the control region of the mitochondrial DNA. Our study aims to establish a database for uni-parental markers (YHRD and EMPOP) and trace the origin and genetic history of the population of Northwest Africa (Morocco) based on archaeological and historical data. For all markers, our results show a mosaic genetic composition from a very old population may stem from the Paleolithic and/or Neolithic. A clear genetic differentiation was observed between our study populations and a population from Egypt, a less important distinction is found between the different Moroccan ethnic groups.

**Key words:** Human identity, genetic populations, mitochondrial DNA, Y Chromosome, Morocco genetic history.

## **Liste des abréviations**

<b>ADN</b>	<b>Acide Désoxyribo Nucléique</b>
<b>pb</b>	<b>paire de base</b>
<b>SNP</b>	<b>Single Nucleotide Polymorphism</b>
<b>A</b>	<b>Adénine</b>
<b>T</b>	<b>Thymine</b>
<b>C</b>	<b>Cytosine</b>
<b>G</b>	<b>Guanine</b>
<b>UV</b>	<b>Ultra-Violet</b>
<b>HLA</b>	<b>Human Leucocyte Antigen</b>
<b>ADNmt</b>	<b>ADN mitochondrial</b>
<b>ARNt</b>	<b>Acide RiboNucléique de transfert</b>
<b>ARNr</b>	<b>Acide RiboNucléique ribosomal</b>
<b>D-Loop</b>	<b>Displacement-Loop</b>
<b>CRS</b>	<b>Cambridge Reference Sequence</b>
<b>rCRS</b>	<b>revised Cambridge Reference Sequence</b>
<b>IUPAC</b>	<b>International Union of Pure and Applied Chemistry</b>
<b>HVI</b>	<b>HyperVariable Region I</b>
<b>HVII</b>	<b>HyperVariable Region II</b>
<b>HVIII</b>	<b>HyperVariable Region III</b>
<b>BP</b>	<b>Before Present</b>
<b>SXD</b>	<b>SeX Determining</b>
<b>SST</b>	<b>Sexually Selected Trait</b>
<b>SOX3</b>	<b>Sry-related HMG box</b>
<b>SRX</b>	<b>Sex determining Region Y</b>
<b>HMG</b>	<b>High Mobility Group</b>
<b>PAR</b>	<b>Pseudo-Autosomal Region</b>
<b>MSY</b>	<b>Male Specific Y chromosome</b>
<b>NRY</b>	<b>Non Recombining Y</b>
<b>DYS</b>	<b>DNA Y-chromosome Segment number</b>
<b>STR</b>	<b>Short Tandem Repeat</b>
<b>Hg</b>	<b>Haplogroupe</b>
<b>YCC</b>	<b>Y Chromosome Consortium</b>
<b>ISOGG</b>	<b>International Society of Genetic Genealogy</b>
<b>HGD</b>	<b>Human Genome Diversity</b>
<b>Ct</b>	<b>Cycle threshold</b>
<b>PCR</b>	<b>Polymerase Chain Reaction</b>
<b>ISFG</b>	<b>International Society of Forensic Genetics</b>
<b>GD</b>	<b>Gene Diversity</b>

## *Liste des abréviations*

---

<b>HD</b>	<b>H</b> aplotype <b>D</b> iversity
<b>DC</b>	<b>D</b> iscrimination <b>C</b> apacity
<b>NUH</b>	<b>N</b> umber of <b>U</b> nique <b>H</b> aplotypes
<b>Fst</b>	<b>F</b> -statistics
<b>AMOVA</b>	<i><b>A</b>nalysis of <b>M</b>Olecular <b>V</b>ariance</i>
<b>ddNTPs</b>	<b>d</b> idéoxy <b>N</b> ucléotide <b>T</b> ri <b>P</b> hosphate
<b>BET</b>	<b>B</b> romure d' <b>E</b> thidium
<b>TBE</b>	<b>T</b> ris <b>B</b> orate <b>E</b> DTA
<b>EDTA</b>	<b>E</b> thylene <b>D</b> iamine <b>T</b> etra <b>A</b> cétique
<b>PEX</b>	<b>P</b> rimier <b>E</b> xtension

## Liste des Figures

- Figure 1** : Carte du nord Afrique montrant les provinces de la Maurétanie : Maurétanie Tingitane à l'ouest (Maroc actuel), Maurétanie Césarienne au centre-ouest, Numidie au centre-est et Africa à l'est----5
- Figure 2** : (a) Transmission partiel de l'information génétique au niveau des chromosomes autosomiques à travers les générations. (b) Transmission totale de l'information génétique contenue au niveau du chromosome Y et de l'ADN mitochondrial. ----- 17
- Figure 3** : Structure du génome mitochondrial humain ----- 19
- Figure 4** : Hétéroplasmie ponctuelle, superposition de deux bases, ici, la Thymine et la Cytosine, elle est noté Y d'après le code IUPAC (International Union of Pure and Applied Chemistry) ----- 22
- Figure 5** : Hétéroplasmie de taille HVI, Transition de la Thymine en Cytosine avec un décalage de la lecture due à une insertion d'une cytosine à la position 16193 rendant le chromatogramme illisible par la suite----- 23
- Figure 6** : Les haplogroupes mitochondriaux présents dans les populations humaines mondiales----- 26
- Figure 7** : Arbre phylogénétique simplifié des haplogroupes mitochondriaux rencontrés dans les populations humaines ----- 27
- Figure 8** : Carte géographique retraçant les principales migrations des haplogroupes mitochondriaux ----- 28
- Figure 9** : La paire de chromosomes sexuels humains. Vue prise au microscope électronique à balayage ---- 29
- Figure 10** : Schéma illustrant l'évolution du chromosome Y ----- 31
- Figure 11** : Schéma de la structure du chromosome Y ----- 32
- Figure 12** : Arbre simplifié pour les haplogroupes du chromosome Y ----- 34
- Figure 13** : Carte géographique retraçant les principales migrations des haplogroupes du chromosome Y ----- 36
- Figure 14** : Marqueurs de fluorescence utilisés dans le kit AmpFISTR®Yfiler™ (AppliedBiosystems, Foster City,CA, USA) et taille des loci amplifiés ----- 41
- Figure 15** : Echelles alléliques (ladder) utilisées dans la PCR multiplex Yfiler™ (Applied Biosystems) ----- 44
- Figure 16** : Exemple d'un profile Y-STR d'un individu obtenu après la PCR multiplex Yfiler™ (Applied Biosystems) ----- 45
- Figure 17** : Principe du génotypage par SNaPshot™ Multiplex par extension allèle spécifique d'amorce, selon le SNP----- 51
- Figure 18** : Stratégie d'amplification et du séquençage de la région contrôle de l'ADN mitochondrial Humain suivi à AFDIL----- 57

<b>Figure 19</b> : Illustration du profile Y-STR pour le locus DYS458 issu du logiciel GeneMapper® v 3.2 de l'individu qui présente l'allèle variant (allèle 13) -----	60
<b>Figure 20</b> : Distribution des fréquences allélique pour le locus DYS458 des métapopulations suivantes : Eurasienne, Asie de l'EST, Aborigène d'Australie, Africain, Eskimo-Aleut, Afro-Asiatique et les américains autochtones-----	60
<b>Figure 21</b> : Illustration graphique des individus ayant l'allèle 13 pour le locus DYS458 dans 5 métapopulations : Africain Américain, Asiatique, Caucasien, Hispanique et Autochtone Américain -----	61
<b>Figure 22</b> : Illustration schématique du locus DYS385 a/b dupliqué sur le chromosome Y -----	68
<b>Figure 23</b> : Illustration graphique de la distance génétique entre la population marocaine étudiée et 9 populations soumis sur la base de données YHRD sur une échelle multidimensionnelle -----	69
<b>Figure 24</b> : Fréquences de l'ensemble des haplogroupes rencontrés dans la population marocaine étudiée (n =159) regroupant les trois groupes ethnies [Arabophones, n = 42 ; Berbèrophones n = 67 ; Sahraouis, n = 50)-----	73
<b>Figure 25</b> : Carte illustrant la distribution géographique de l'haplogroupe E1b1b1 (E-M35) -----	74
<b>Figure 26</b> : Carte illustrant la distribution géographique de l'haplogroupe E1b1b1a (E-M78)-----	76
<b>Figure 27</b> : Carte illustrant la distribution géographique de l'haplogroupe E1b1b1b1a (E-M81) -----	77
<b>Figure 28</b> : Carte illustrant la distribution géographique de l'haplogroupe E1b1b1b2a (E-M123)-----	77
<b>Figure 29</b> : Carte illustrant la distribution géographique de l'haplogroupe J2 (J-M172) -----	78
<b>Figure 30</b> : Carte illustrant la distribution géographique de l'haplogroupe J1 (J-M267) -----	79
<b>Figure 31</b> : Illustration des individus ayant l'hétéroplasmie de taille (HVI, HVII, HVIII et AC) et ponctuelle dans notre population étudiée -----	83
<b>Figure 32</b> : Ascendant maternelle de la population marocaine étudiée telle que définie par les lignées d'ADN mitochondrial humain -----	84
<b>Figure 33</b> : Ascendant maternelle de la population marocaine étudiée telle que définie par les lignées d'ADN mitochondrial humain -----	85

## Liste des Tableaux

<b>Tableau 1</b>	: Résultat de la lecture du profil électrophorétique Y-STR d'un individu par ordre de taille de séparation -----	46
<b>Tableau 2</b>	: les Fluorochromes des ddNTP du kit ABI Prism® SNaPshot™ -----	52
<b>Tableau 3</b>	: Les séquences des amorces utilisées pour l'amplification et le séquençage de la région contrôle de l'ADNmt -----	56
<b>Tableau 4</b>	: Fréquences des haplotypes pour 17 Loci Y-STR de notre population étudiée (n = 267) -----	62
<b>Tableau 5</b>	: Distribution des fréquences alléliques ainsi la diversité génétique des 17 loci Y-STR de la population étudiée-----	67
<b>Tableau 6</b>	: Calcul de la distance génétique entre la population marocaine étudiée (n = 267) et 9 populations citées précédemment par le test AMOVA. Les valeurs $\Phi_{st}$ sont représentées au-dessus de la diagonale et les valeurs p sont représentées au-dessous de la diagonale -----	70
<b>Tableau 7</b>	: Les haplogroupes rencontrés dans notre population marocaines étudiées (n = 159) constitué de trois groupes ethnique (Arabes, Berbères et Sahraouis)-----	72
<b>Tableau 8</b>	: L'analyse statistique des haplotypes de la région contrôle obtenu chez la population marocaine étudiée. Les statistiques sont basées sur les séquences entières de la région de contrôle (16024-576). Insertions au 16193, 309 et 573 ont été ignorées de l'analyse-----	80
<b>Tableau 9</b>	: Le Tableau illustre les haplotypes partagés ainsi les haplogroupes de notre population analysée. Pour les échantillons possédant l'hétéroplasmie de taille, les haplotypes énumérés ci-dessous représentent les composantes majeures -----	80

## **Introduction générale**

## Introduction générale

---

Le peuplement du monde par les premiers humains modernes (*Homo sapiens*) place son origine au niveau de l'Afrique de l'Est (théorie de l'origine mono centrique dite « Out of Africa »). C'est le phénomène global le plus ancien qu'ait connu notre espèce, conduisant cette dernière jusqu'aux derniers recoins inhabités de la planète, en dépit d'environnements parfois hostiles. A l'heure actuelle, de nombreux chercheurs étudient cette globalisation, depuis son origine la plus reculée, il y a plus 100.000 ans, jusqu'aux dernières conquêtes s'achevant dans le courant de notre ère: paléontologues, archéologues, linguistes et généticiens des populations sont à l'œuvre pour reconstituer notre passé.

De manière classique, la génétique des populations a pour objet l'étude de l'impact des différentes forces évolutives sur la répartition de la diversité génétique entre les populations et dans les populations. Connaissant la diversité actuelle de notre espèce, les outils de la génétique des populations permettent de retracer l'histoire des populations, de détecter si certains gènes ont été soumis à la sélection dans notre espèce dans son ensemble ou seulement dans certaines populations. C'est donc un des outils pertinents pour retracer le passé évolutif d'une espèce et comprendre sa diversité actuelle. Parmi les modèles d'études disponibles, les généticiens des populations étudient en particulier le chromosome Y et l'ADN mitochondrial qui, échappant au phénomène de recombinaison lors de la méiose, permettent d'explorer avec un certain degré de fiabilité respectivement les lignées paternelles et maternelles.

C'est dans ce contexte multidisciplinaire que s'inscrit le sujet de cette thèse. En effet, si les vestiges archéologiques et paléanthropologiques mis au jour témoignent de l'ancienneté (Paléolithique) de l'occupation du Maroc par l'Homme, la question de l'évolution de ces groupes humains jusqu'à nos jours reste posée en raison des divers évènements préhistoriques et historiques.

Le nom même de « **Maroc** » apparaît, semble-t-il, sous les Saadiens, dynastie ayant régné de 1554 à 1659. Il serait le résultat de la contraction du nom de la ville de Marrakech, la principale de leurs capitales. Auparavant, pour désigner le Maroc, on parlait de *Maghreb el-Aqça* ou Maghreb extrême.

Le Maroc situé dans le nord-ouest de l'Afrique appartient au monde de la Méditerranée. Séparé de l'Europe par un détroit large de quelques kilomètres à peine, le pays est bordé par l'océan Atlantique à l'ouest, par l'Espagne, le détroit de Gibraltar et la mer Méditerranée au nord, par l'Algérie à l'est, et de

## Introduction générale

---

*facto* au sud par la Mauritanie au-delà du Sahara atlantique. Le Maroc fait partie des États du *Maghreb*, il est considéré comme étant le pays le plus occidental.

Certaines populations avaient une place de premier ordre : **les populations Amazighes**. En effet, leurs ancêtres sont considérés comme les plus anciens autochtones Nord-Africains, sans doute depuis le Paléolithique (ou le Capsien), mais dont la présence serait incontestablement marquée au Néolithique (vers 7 000 ans B.P dans les régions sahariennes et 4 300 ans B.P au Maghreb). Ces hommes ont ensuite connu un passé très riche en invasions, conquêtes et tentatives d'assimilation notamment les phéniciens, carthaginois, romains, vandales, byzantins. Ces migrations et expansions de populations ont certainement laissé leurs empreintes dans la composition génétique des groupes Amazighes. Ce travail de thèse s'appuie sur cette remarque d'ordre biologique et décrit, à travers l'étude des populations actuelles, le patrimoine génétique de la population marocaine qui renferme à nos jours trois groupes ethniques (Arabophones, Berbérophones et Sahraouis). Il tente également de mesurer si celui-ci a pu être modifié lors des divers événements historiques qui se sont succédés au Nord de l'Afrique. La période la plus importante et la plus marquante pour l'histoire du nord de l'Afrique a été la présence des arabes (amorcée au VII<sup>ème</sup> siècle). Celle-ci a induit de profonds changements culturels dans les populations Amazighes en imposant tout d'abord la religion, l'Islam, puis, dans un deuxième temps, la langue de ces conquérants, l'arabe. Notre objectif est de mieux appréhender les impacts génétiques des migrants arabes sur les Berbères du Nord-Ouest d'Afrique et de voir si une éventuelle différenciation génétique entre ces deux groupes est à mettre en parallèle avec leur langue.

Dans ce travail, la description de la structure génétique d'un échantillonnage de la population marocaine est basée sur la recherche de divers polymorphismes situés sur les marqueurs uni-parentaux: deux polymorphismes au niveau du chromosome Y, les marqueurs bi-alléliques (SNP) et les microsatellites Y-STRs, et le polymorphisme extranucléaire de l'ADN mitochondrial (ADNmt). Ces deux types de marqueurs ont un grand intérêt aussi en génétique médico-légale. Le chromosome Y ne pouvant provenir que du père, tous les hommes de la même lignée paternelle présente un profil identique. Ce type de polymorphisme est principalement utilisé dans le cadre des dossiers de meurs, lorsqu'une faible quantité d'ADN masculin est mélangé à une grosse quantité d'ADN féminin issu de la victime. L'analyse de l'ADNmt peut donc être une alternative intéressante lorsque l'analyse de l'ADN

## Introduction générale

---

nucléaire a peu de chance de réussite, principalement pour l'analyse des cheveux, d'échantillons très anciens ou encore d'échantillons dégradés. D'autant plus que l'avantage d'utiliser l'ADN mitochondrial pour l'analyse de la diversité génétique de nos ancêtres, réside dans le fait que les mitochondries sont transmises uniquement par la mère. Cela permet donc de suivre des populations en comparant le degré de similarité de leur ADN mitochondrial.

Ainsi, notre analyse de ces différents types de polymorphismes des marqueurs microsatellites et bi-alléliques du chromosome Y, et l'ADN mitochondrial est menée selon quatre principaux objectifs :

1. Situer la diversité génétique de notre échantillonnage de la population marocaine parmi celle de populations géographiquement proches, notamment les populations du Nord d'Afrique, Sud-Ouest Européennes, d'Asie et de l'Afrique.
2. Décrire la composition génétique de notre échantillonnage pour les marqueurs bi-alléliques.
3. Elargir l'échantillonnage et étaler l'analyse sur toute la région contrôle pour le polymorphisme de l'ADN mitochondriale (précédemment limité selon la littérature que sur la région HVI et HVII de la région contrôle).
4. Etablir la base de donnée pour les marqueurs STR du chromosome Y sur YHRD ([www.yhrd.org](http://www.yhrd.org)) et sur EMPOP ([www.empop.org](http://www.empop.org)) pour l'ADN mitochondrial.

# *Chapitre A*

**Le Maroc :**

**Son origine à nos jours**

## **I. Histoire du Maroc**

Le **Maroc** est une région du nord-ouest de l'Afrique dont le nom même, dérivé de Marrakech, ville impériale et capitale de 1062 à 1273, remonte au XVI<sup>e</sup> siècle. Le pays a connu des peuplements phéniciens, carthaginois, romains, vandales, byzantins avant d'être islamisé par les arabes. C'est lors de son exil qu'Idriss Premier, fuyant les persécutions du califat des Abbassides, en 788, a donné naissance à un État dans ce Maghreb *el-Aqça* ou Maghreb extrême (Bernard Lugan, 2000). Le Maroc a toujours gardé, si ce n'est son indépendance absolue, une très forte autonomie.

Selon l'historien Bernard Lugan l'attrait des richesses provenant du commerce du Sud (Sahara) vers le Nord (l'occident) va attirer les convoitises de diverses tribus avec pour ville carrefour Marrakech (la porte du désert) qui deviendra naturellement la capitale de diverses dynasties en particulier celles venant du Sud (Almoravides, Almohades, Saadiens); c'est la raison pour laquelle, toute l'histoire du Maroc (des Idrissides aux Alaouites) fut marquée par le commerce des richesses du Sud vers le Nord. L'histoire et l'origine du Maroc furent, sont et seront marquées par le lien avec le Sahara.

### **I.1. Les estampilles de peuplement au Maroc du Paléolithique au Néolithique**

L'Homme a laissé de nombreuses traces au cours de toute la période préhistorique, marque d'un peuplement très ancien.

#### ***Paléolithique inférieur***

Des indices datant d'au moins 1.2 Millions d'années (Raynal et *al.*, 2010) traduisent une première activité humaine. Ces hommes vivaient principalement de la cueillette et de la chasse. Les outils de cette époque sont les galets aménagés, le biface, les hachereaux découverts notamment dans les régions de Casablanca et de Salé.

#### ***Paléolithique moyen***

D'après les travaux de plusieurs chercheurs (Roche, 1953, 1967 et 1969 ; Hublin et *al.*, 1988), l'Atérien est précédé par le Moustérien. Mais peu d'indications sont disponibles sur sa date (environ

30 000 ans B.P.) et sur sa localisation dans le gisement de Taforalt par exemple. Néanmoins, cette relation est plus claire dans la grotte de Rhafas au niveau de la Région d'Oujda (Wengler, 1993).

Les nouvelles fouilles dans plusieurs sites au Maroc ont plus mis l'accent sur la précision du cadre stratigraphique, chronologique et paléo-environnemental dans cette période. Parmi les sites les plus importants on peut citer la grotte des pigeons à Taforalt compte tenu de la complexité de la stratigraphie du remplissage, la cavité a été subdivisée dans un premier lieu en plusieurs secteurs puis dans un second lieu une interprétation de l'ensemble du site a été proposée (Bouzouggar et al., 2007). Le Paléolithique moyen est principalement identifié dans les secteurs 1 et 2. La plus importante caractéristique de cette période correspond à plusieurs niveaux atériens dont la base date au-delà de 100 000 ans B.P. (Humphrey et al., 2012) et le sommet de la séquence est encore difficile à interpréter malgré l'obtention de plusieurs dates, car les objets archéologiques trouvés dans des niveaux sûrs, présentent des caractères à la fois du Paléolithique moyen et du Paléolithique supérieur.

D'après les auteurs des nouvelles recherches dans la grotte des Pigeons à Taforalt, elle a révélé le plus ancien niveau atérien au Maroc contenant des objets de parure (*Nassarius gibbosulus*) ocrés considérés comme parmi les plus anciens au monde et qui ont été datés par quatre méthodes différentes qui ont fourni un âge au-delà 100 000 ans, (Bouzouggar et al., 2007 ; Humphrey et al., 2012).

### ***Paléolithique supérieur***

D'après l'abbé Jean Roche, le Paléolithique supérieur représenté dans la grotte de Tafouralt par l'Ibéromaurusien date de 21 900 ans B.P., ce qui fait de ce site avec le gisement de Tamar Hat en Algérie (Saxon et al., 1974) comme les plus anciens témoins du Paléolithique supérieur en Afrique du Nord et qui a perduré jusqu'à environ 10 800 ans B.P.

La grotte des Pigeons a également fonctionné au cours d'une période du Paléolithique supérieur, comme une nécropole. Les fouilles de l'abbé Jean Roche ont mis au jour plus de 180 squelettes (Roche, 1963) et qui ont été étudiés en détail par Denise Ferembach (Ferembach et al., 1963). Dans le niveau de la nécropole a été également exhumé un squelette dont le crâne présente les traces d'une trépanation considérée comme la plus ancienne au monde. Les radiographies (Ferembach,

1962) ont démontré la présence d'un processus de cicatrisation ce qui impliquerait que l'individu a survécu à cette opération.

Si les phases récentes du Paléolithique supérieur dans la grotte sont claires et ne diffèrent guère de ce qui a été décrit par l'abbé Jean Roche (Roche, 1963), les phases anciennes sont très difficiles à interpréter et correspondraient à l'un des plus anciens faciès de l'Ibéromaurusien en Afrique du Nord (Barton et *al.*, 2007 ; Humphrey et *al.*, 2012 ; Bouzouggar et *al.*, 2006).

L'identification des charbons de bois (anthracologie), montre la présence du cèdre à la base de l'Ibéromaurusien et au sommet de la séquence dans cette grotte (Ward, 2007).

### **Le Néolithique**

Le mot **Néolithique** désigne littéralement le *nouvel âge de la pierre*. Ce terme a été proposé en 1865 par le préhistorien John Lubbock. Le Néolithique est une période de la Préhistoire qui a occupé une durée de temps relativement courte s'échelonnant du VII<sup>e</sup> au III<sup>e</sup> millénaire av. J.C. et marquée par de profondes mutations techniques, économiques et sociales, liées à l'adoption par les groupes humains d'un modèle de subsistance fondé sur l'agriculture et la domestication des animaux, et impliquant le plus souvent une sédentarisation. Les principales innovations techniques sont la généralisation de l'outillage en pierre polie et de la poterie. Ces productions et modes de vies sont communément appelés « néolithisation » et correspondent, en réalité, à plusieurs Néolithiques résultant des adaptations aux changements climatiques dans différentes régions. Ces transformations se sont mises en place progressivement dans le contexte marocain, ainsi les cultures possédant au moins un critère de néolithisation est considéré comme néolithique.

C'est dans le *contexte climatique* que s'est effectuée la transition de l'Ibéromaurusien vers le Néolithique. La rareté des données archéologiques ne permet cependant pas une reconstitution satisfaisante. Il semble toutefois que le *contexte culturel* ait toujours été fortement marqué par des productions et des systèmes de subsistance connus au cours de l'Ibéromaurusien, comme la chasse, la cueillette et la fréquence d'un outillage composé essentiellement d'outils taillés. Malgré cette continuité apparente, certains auteurs soutiennent que la diffusion du Néolithique s'est effectuée du nord vers le sud, à partir de la Péninsule ibérique. Cette hypothèse s'appuie sur l'étude d'un type de céramique décorée à la coquille de cardium « *cardium edule* », d'où l'appellation de céramique cardiale (Mohamed Kably et *al.*, 2011).

La navigation dans le bassin méditerranéen ayant été vraisemblablement pratiqués, la céramique cardiale s'inscrirait dans un courant d'échanges entre les deux rives de la méditerranée par le détroit de Gibraltar, et non dans un contexte de déplacements massifs ou de remplacement de la population. Les données de l'anthropologie biologique ne vont pas à l'encontre de ce constat, car, à cette époque, le peuplement humain au Maroc n'a pas été bouleversé par de nouveaux **arrivants** puisqu'il était marqué par la présence de **Mechtoïdes**. Mais, progressivement, d'autres types humains ont fait leur apparition. Le Néolithique montrerait donc une diversité biologique et culturelle qui serait due à une ouverture sur l'espace régional, à des adaptations aux fluctuations climatiques et l'adoption de régimes alimentaires différents selon les aires géographiques.

Sur le plan répartition spatiale, le Néolithique est connu dans toutes les régions du Maroc, avec des spécificités locales. Cependant, l'état actuel des connaissances ne couvre que les zones du Sud Saharien, du Haut-Atlas, du Nord, de l'Oriental et de la Meseta marocaine.

## I.2. Le Maroc antique

Les Phéniciens, commerçants entreprenants, installent leurs premiers établissements sur les côtes marocaines dès le XII<sup>e</sup> siècle av. J.-C. selon les sources et fondent des port-comptoirs comme *Tingi* (Tanger) ou *Lixus* (Larache). C'est à partir de la fondation de Carthage (en Tunisie, Maghreb de l'est) que la région commence à être réellement mise en valeur. L'influence punique (carthaginois) se fera sentir près de mille ans au Maroc, dans ses relations avec les chefs de tribus berbères locales : en effet à partir du VI<sup>e</sup> siècle, les carthaginois en quête d'or (tiré de l'Atlas), de pourpre (coquillage que l'on trouve à Mogador par exemple, et qui donne la teinture du même nom), vont commercer avec les habitants du Maroc.

C'est à partir du IV<sup>e</sup> siècle av. J.-C. que, dans le nord du Maroc, apparaît la première organisation politique du pays : le royaume de Maurétanie apparaît, résultat de la fédération de différentes tribus berbères qui avaient profité de l'influence punique (Jean Carcopino, 1943) (Figure 1).

Lorsque les Romains arrivent vers le II<sup>e</sup> siècle av. J.-C., après la destruction de Carthage, ils sont d'abord alliés à ce royaume de Maurétanie qui leur permet de lutter et de prendre à revers le chef numide Jugurtha. La Maurétanie devient un royaume ami, un « état-client », qui, s'il dépend étroitement de Rome et prendra part à toutes les querelles internes de l'Empire, reste de fait indépendant. En année 40 (Calendrier Julien), le royaume des Maures perd son roi. Caligula, qui l'a

fait assassiner, fait face à la guerre d'Aedemon : Il faudra quatre ans pour mater cette révolte et en 46, l'empereur Claude annexe le royaume qui devient la Maurétanie tingitane (chef-lieu *Tingi*, devenu Tanger). La domination romaine se limite aux plaines du nord (jusqu'à la région de Volubilis près de Meknès) et l'Empire ne cherche pas à contrôler la région très fermement : il semble que les tribus berbères autonomes et pacifiques étaient imbriquées dans les possessions romaines. Pour autant Rome doit lutter sans cesse contre les Berbères montagnards.



**Figure 1:** Carte du nord Afrique montrant les provinces de la Maurétanie : Maurétanie Tingitane à l'ouest (Maroc actuel), Maurétanie Césarienne au centre-ouest, Numidie au centre-est et Africa à l'est [H. Kiepert, Antique Atlas, Berlin].

Au même titre que le reste de l'Afrique du nord, la Maurétanie Tingitane va connaître la christianisation. Des dizaines d'évêchés couvrent la région, s'adressant d'abord aux populations romaines puis aux romanisés. C'est en 298, à Tanger, sous Dioclétien que saint Marcel, centurion romain, est décapité (Anne Bernet, 2006). Les berbères du Maroc ne seront, à la différence des berbères d'Algérie et de Tunisie, que très peu christianisés. Deux évêchés ont été identifiés en Tingitane (à Tanger et Larache).

Au III<sup>e</sup> siècle, l'Empire recule. C'est aussi le cas en Afrique du Nord et en particulier au Maroc : la Maurétanie Tingitane se retrouve réduite à la seule ville de Tingi et à la côte nord. Les villes du sud sont toutes abandonnées, y compris la grande cité *Volubilis*. Au sud seul le port de *Sala* est conservé à l'Empire. Les raisons de ce repli sont mal connues : pression des berbères montagnards et du sud?

Crise économique plus violente dans cette région ? Affaiblissement dû aux conflits dynastiques de l'Empire avec l'épisode des Gordiens? (Bernard Lugan, 1992).

Profitant de l'affaiblissement de l'Empire romain d'occident, une troupe de barbares de langue teutonne, formées de Suèves, de Vandales et d'Alains traverse le Rhin en 406. Les Vandales descendent alors en Espagne et passent en Afrique en 429. Ils atteignent Hippone (Algérie) en 430. Le gouvernement de Constantinople engage en vain une expédition navale contre cette invasion. Les Vandales s'installent dans l'Afrique du nord-ouest pour plus d'un siècle.

La Maurétanie Tingitane, quant à elle, n'est d'abord pas touchée par la conquête et la domination vandale. Ceux-ci ne contrôleront que les côtes méditerranéennes. La région passe sous contrôle byzantin en 534. Mais les berbères, habitués à une large autonomie depuis plus d'un siècle, s'ils sont encore « romanisés », ne sont plus « romains », et ils vont résister farouchement autour du prince Garmel (Jérôme Carcopino, 1947). Après la victoire byzantine, la province connaît un certain renouveau économique et démographique.

### **I.3. Conquête arabo-musulmane et islamisation du Maroc**

En 638, les Arabes prennent Alexandrie. En 649, ils atteignent le Maghreb. Mais ce n'est qu'à la cinquième campagne (681) qu'ils entrent au Maroc. Ils font alors face à une farouche résistance berbère, à la suite de certaines erreurs diplomatiques. Les Berbères, qu'ils soient montagnards, ou des plaines aujourd'hui marocaines ou algériennes, vont permettre à l'empire byzantin de se maintenir jusqu'en 698. L'empire byzantin est alors vaincu et ne subsiste que la résistance berbère. Cette résistance tient encore quinze ans. En 708, le Maroc berbère se convertit massivement à l'islam. Cette conversion, qui touchait des populations qui n'avaient pas été christianisées, ne semble pas être remise en cause par les Berbères. Très vite, Les musulmans utilisent les capacités guerrières des nouveaux convertis : l'Espagne wisigothique est conquise en trois ans, les troupes arabes et berbères arrivent en Navarre (Nord d'Espagne) en 715. Ils seront vaincus à Poitiers en 732.

L'ensemble du Maroc côtier est sous domination de l'empire Omeyyade. Dans la région du Rif s'établit un petit émirat berbère autonome : l'émirat de Nekor ou Nokour (Henri Terrasse, 1949).

En 740 a lieu la première révolte berbère face au pouvoir arabe : le *kharijisme* sert de prétexte pour remettre en cause le califat d'orient. C'est, pour ses fidèles, la volonté de choisir « le meilleur »

pour gouverner, et non pas forcément un descendant du prophète (ce que veut le chiisme), ou un candidat choisi par les sages (ce que veut le sunnisme). Le *kharijisme* est la thèse la plus appréciée par les peuples berbères, qui ont des sentiments relativement démocratiques : le chef se doit d'être choisi par tous, et non pas imposé (Bernard Lugan, 2000). Le califat omeyyade ne peut l'accepter, et un conflit éclate. En 750, à Damas, les Omeyyades sont renversés par les Abbassides. Le Maroc se retrouve dans une quasi-anarchie.

#### **I.4. Soubassement du Maroc**

Le soubassement du Maroc, pays se considérant arabo-berbère, africain et musulman, se fait avec les Idrissides qui allèrent à leur cause diverses tribus arabo-berbère-afro-musulmanes contrôlant des petits royaumes ou territoires indépendants de tout pouvoir central (Abdellah ben Mlih, 1990). Au fur et à mesure des alliances, les Idrissides vont étendre leur influence territoriale avec des populations autochtones et lancer les bases de l'organisation d'un état constitué (Makhzen) reprises par les dynasties suivantes. Si les Idrissides vont commencer à dessiner les bases de l'état et des frontières de l'actuel Maroc ce sont les Almoravides qui en créant leur capitale Marrakech donneront au pays son nom ; ils consolideront et élargiront l'œuvre débutante et fragile des Idrissides ; les dynasties suivantes hériteront de l'expérience étatique précédente.

Même si d'autres civilisations du bassin méditerranéen (Rome, Carthage etc.) ont enrichi l'histoire du pays et même, si des populations de l'actuel Maroc vont participer à l'essor de ces civilisations, les historiens du Maroc considèrent que l'impact de ces civilisations a été limité à l'extrême nord du Maroc et à certains comptoirs commerciaux.

A propos du Maroc, le terme *Empire* est parfois utilisé car par définition, un empire est un ensemble d'états ou de royaumes. Ceci explique l'appellation « villes impériales » utilisée encore de nos jours pour qualifier les villes de Fès, Marrakech, Meknès et Rabat (Samuel Pickens, Françoise Peuriot, Philippe Ploquin, 1995).

Lorsque le Maroc se fonde, le reste du Maghreb est éclaté sous forme de royaumes ou territoires indépendants, parfois concurrents ou en guerre, sans pouvoir central c'est-à-dire non organisés en état dirigé par des populations autochtones.

## *Partie I : Données de la littérature*

### *Chapitre A: Le Maroc son origine à nos jours*

---

L'organisation en Etat structuré permit aux Saadiens et aux Alaouites de s'opposer à l'avancée Othmane qui s'arrêta à la rivière Melwiya et qui s'étendait sur une grande partie des autres pays arabes actuels.

Des désaccords apparus au début du XX<sup>e</sup> siècle dans la famille Alaouite suite à des problèmes de gestion du pays, créèrent une période d'instabilité dont vont profité plusieurs puissances coloniales (Allemagne, Angleterre, Espagne, France) pour essayer de s'emparer du pays qui possède entre autres une position géostratégique intéressante, à la veille de la Première Guerre mondiale. Après bien des tractations houleuses et secrètes qui faillirent déclencher, dès 1912, la première guerre mondiale (coup d'Agadir), le Maroc fut partagé entre la France et l'Espagne.

# *Chapitre B*

**Concepts et génétique des populations**

## I. La génétique des populations

### I.1. Transmission des caractères

La génétique initiée par Gregor Mendel, appelée classiquement **génétique mendélienne**, a pour objectif de comprendre le déterminisme et la transmission des caractères par l'analyse de la descendance d'un croisement contrôlé entre individus de génotypes différents (proportions des diverses catégories de descendants). Après la découverte du support de l'information génétique (ADN), la **génétique moléculaire** continue à rechercher les mécanismes fins du déterminisme, de l'expression et de la transmission des caractères. Elle trouve aujourd'hui de nombreuses extensions avec les programmes de génomique (séquençage des génomes et identification des gènes) et de protéomique (inventaire et fonction des protéines d'un organisme).

La compréhension du déterminisme et de la transmission des caractères nécessite l'étude des individus dans les conditions naturelles où ils sont génétiquement uniques et libres de se reproduire avec n'importe quel autre individu de la même espèce. Cette partie de la génétique, qui considère les individus en interactions avec leur environnement, est intitulée « **la génétique des populations** ».

### I.2. Définition et objectifs

La génétique des populations s'efforce d'évaluer la diversité génétique et d'établir des lois décrivant dans le temps et dans l'espace le maintien ou la modulation de la variation génétique au sein d'une population. Elle vise donc à comprendre pourquoi et comment l'information génétique évolue au cours du temps au sein des espèces et des populations. Les fondements de la génétique des populations reposent sur deux concepts: la *diversité génétique* et la *population*. Le point de départ de la discipline est la recherche de la variabilité génétique ou **polymorphisme** observable au niveau de divers loci géniques (allèles) dans une population donnée. L'étape suivante consiste à exprimer ce polymorphisme génétique par le calcul des fréquences (proportions) relatives des différents allèles dans la population. La comparaison des résultats obtenus pour la même population, à différentes échelles de temps, donne alors accès à des informations sur son évolution génétique, qui se traduit par une variation (augmentation, diminution ou maintien) des fréquences alléliques. La comparaison génétique de

différentes populations permet ensuite de mesurer les différences (ou les similitudes) observées et ainsi d'accéder à l'histoire génétique de chaque population, d'un ensemble de populations ou encore de l'espèce. Cette histoire génétique s'attache, dans la mesure du possible, à déterminer les origines génétiques des populations mais aussi à retracer leurs interactions au fil du temps (brassage génétique, flux géniques). Les modifications des fréquences alléliques sont non seulement liées au comportement même des individus constituant les populations (choix culturel, religieux, démographique, sanitaire, etc., du conjoint ; guerre ; épidémies) mais également à d'autres processus connus sous le terme de **forces évolutives** : la mutation, la migration, la sélection naturelle et la dérive génétique. Les trois principaux objectifs de la génétique des populations sont :

- a. Mesurer la variabilité génétique, appelée aussi diversité génétique, par la fréquence des différents allèles d'un même gène ;
- b. Comprendre comment la variabilité génétique se transmet d'une génération à l'autre ;
- c. Comprendre comment et pourquoi la variabilité génétique évolue au fil des générations.

### I.3. Équilibre de Hardy-Weinberg et les forces évolutives

Les modèles de génétique des populations retracent l'évolution de paramètres décrivant la population prise dans son ensemble. Ces paramètres sont abordés dans les études empiriques par le calcul de quantités (estimateurs) mesurées sur des échantillons représentatifs de la population. Le modèle théorique central de la génétique des populations a été décrit en 1908 par le mathématicien anglais G. Hardy et le médecin allemand W. Weinberg. Il est connu sous le nom «**d'équilibre de Hardy-Weinberg**». Ce modèle correspond, sous certaines conditions, à un équilibre des fréquences génotypiques attendues dans une descendance, en fonction des fréquences alléliques parentales (Hardy 1908, Weinberg 1908).

Bien que ses hypothèses soient des simplifications évidentes, l'utilité du modèle de Hardy-Weinberg est grande pour plusieurs raisons. En effet, au niveau théorique il a permis de préciser les facteurs pouvant modifier les constitutions génétiques des populations. En levant telle ou telle hypothèse on peut rechercher le rôle propre de chacun des facteurs, puis, en levant simultanément plusieurs hypothèses, on

## Partie I : Données de la littérature

### Chapitre B : Concepts et génétique des populations

---

étudie les effets cumulés des différents facteurs. Quatre forces évolutives peuvent influencer sur l'état d'équilibre de la population : la mutation, la sélection, la migration et la dérive génétique.

☞ La **mutation** est la principale source de la variabilité génétique. Elle désigne n'importe quel changement héréditaire, d'origine biologique, physique ou chimique, intervenu dans la séquence de l'ADN. Il peut s'agir d'une modification ponctuelle d'un ou plusieurs nucléotides sans modification de la longueur de la séquence, ou d'une modification de celle-ci par insertion, délétion ou transposition de segments plus importants du génome. De ce fait, en produisant de nouveaux allèles, les mutations se considèrent les moteurs de la diversité génétique.

☞ La **sélection** se traduit par une valeur sélective ou adaptative différente selon les génotypes. Ce principe de sélection stipulant que certains individus sont plus aptes à survivre et à se reproduire dans un environnement donné, a été l'un des trois principes à la base de la théorie de l'évolution des espèces par sélection naturelle, présentée par Darwin en 1859 dans son célèbre ouvrage « De l'origine des espèces » (Darwin 1859). Les deux autres principes sont : le principe de variation (morphologique, physiologique...) des individus d'une même population et le principe de l'hérédité précisant que les jeunes ressemblent plus à leurs géniteurs qu'à des individus auxquels ils ne sont pas apparentés. Ainsi, les populations naturelles se composent d'individus génétiquement différents pour un grand nombre de gènes, et c'est ce polymorphisme génétique qui constitue un véritable capital adaptatif permettant la survie des individus et des populations. Cependant, comme un allèle peut, selon les conditions, se trouver défavorable dans une population et favorable dans une autre, la sélection ne réduit pas forcément la diversité génétique intra spécifique : elle la limite à la diversité inter populationnelle.

☞ La **migration** est l'occasion de tout échange génétique (transmission d'allèles) entre les populations. Elle modifie, bien évidemment, la fréquence des allèles dans les populations concernées et peut même conduire à une homogénéisation des fréquences alléliques entre les divers groupes (brassage). Son effet est d'autant plus fort que l'effectif immigrant est grand et que la différence de fréquences alléliques entre les deux populations est significative.

☞ La **dérive génétique (Drift)** se produit par échantillonnage des gamètes à chaque génération. Ces derniers ne participent pas tous à la reproduction. La génération qui remplace celle qui l'a générée présente alors, du fait des hasards de la méiose et de la fécondation, des fréquences alléliques différentes. Cette fluctuation des fréquences est d'autant plus grande que la taille de la population est petite. Ainsi, dans les populations à effectif réduit, ces changements aléatoires peuvent entraîner une perte d'allèles. Au contraire, dans une population de grande taille, les fréquences alléliques varieront peu d'une génération à l'autre par l'effet du grand nombre de géniteurs potentiels. On peut même considérer qu'elles sont stables sur un temps assez court. Un cas particulier de la dérive génétique est l'**effet fondateur** qui se produit lorsqu'un petit groupe de personnes constitue le noyau fondateur d'une nouvelle population. Ce groupe migrant, porteur d'une petite fraction de la variation génétique totale de la population d'origine, peut induire une concentration de certains traits génétiques, favorables ou délétères, à l'intérieur d'une population. Les fréquences alléliques de cette nouvelle population sont alors différentes de celles de la population d'origine et peuvent même correspondre à la surreprésentation d'allèles initialement rares ou peu fréquents et responsables de certaines maladies génétiques.

## II. Le polymorphisme génétique

D'après Gouyon et Henry (Pierre-Henri Gouyon, Jean-Pierre Henry, 1998), le polymorphisme génétique a été défini par la coexistence de plusieurs allèles pour un gène ou locus donnés, dans une population. La *diversité* de l'**ADN** (acide désoxyribonucléique) provient de l'accumulation de mutations transmises par nos ancêtres. Cette diversité, connue sous la forme de polymorphisme de l'**ADN** ou bien polymorphisme **génétique** (du grec "*poly*" plusieurs et "*morphos*" formes), peut être analysée et utilisée pour rechercher notre histoire génétique et notre parenté, complétant par conséquent les approches plus traditionnelles de l'archéologie, de la paléontologie et de la linguistique.

L'**ADN** appartient à la famille chimique des acides nucléiques. C'est une molécule formée de l'assemblage linéaire de **nucléotides** composés chacun d'une base azotée, d'un sucre à cinq carbones (le désoxyribose) et d'un acide phosphorique (Harry 2001). Il existe quatre bases dont deux appartiennent au groupe des purines, l'**adénine** (A) et la **guanine** (G), et deux au groupe des pyrimidines, la **cytosine** (C)

et la **thymine** (T). Par simplification, le terme de « base » est couramment employé pour désigner le nucléotide lui-même.

*In vivo*, l'ADN se présente sous la forme d'une double hélice dont les deux brins sont complémentaires et d'orientations opposées. Cette structure hélicoïdale de l'ADN (structure secondaire) en deux chaînes antiparallèles a été décrite en 1953 par J.D. Watson et F.H. Crick (Watson et Crick 1953). Pour cette découverte fondamentale, ils ont reçu en 1962 le prix Nobel de physiologie et médecine. L'ADN est la forme de stockage de l'information génétique : Il constitue le **patrimoine génétique** de l'organisme, son **génom**e et son support de l'hérédité vu qu'il est transmis de génération en génération. Chaque cellule eucaryote renferme deux types de génomes : l'ADN nucléaire (présent dans le noyau) et l'ADN mitochondrial (présent dans les mitochondries cytoplasmiques).

Le génome humain, constitué de 6 milliards de paires de bases (pb) comporte environ 35000 gènes (International Human Genome Mapping Consortium, 2001 ; Venter et *al.*, 2001). Cependant, plus de 90% du génome est non-codant, c'est à dire ne comporte pas d'information en rapport direct avec la synthèse de protéines. Contrairement à l'ADN codant, la partie non codante du génome n'est pas soumise à la forte pression de sélection ce qui permet aux mutations qui y surviennent d'être conservées et transmises à la descendance. Ces régions où l'on rencontre une forte variabilité sont très informatives pour la génétique criminalistique tout en étant phénotypiquement neutres. Plusieurs types de polymorphismes de l'ADN ont été utilisés comme marqueurs génétiques, chacun révélant un niveau différent de variabilité. La génétique des populations étudie les variabilités d'origine génétique des populations, c'est-à-dire les différences existant, en divers loci géniques, entre les individus d'une même espèce. Ces différences sont traduites par la présence de plusieurs formes alléliques, ou **polymorphisme**, à chaque locus.

## II.1. Les modifications à la source du polymorphisme génétique

Le polymorphisme génétique est la conséquence directe de changements survenus dans la séquence de l'ADN. Ces modifications peuvent être causées par divers mécanismes tels que les mutations ponctuelles, les insertions, les délétions ou les transpositions d'un segment de l'ADN.

## Partie I : Données de la littérature

### Chapitre B : Concepts et génétique des populations

---

Les **mutations ponctuelles** sont la source prépondérante du polymorphisme. Elles n'intéressent qu'une région très limitée du génome (quelques nucléotides, au maximum). Les modifications les plus simples correspondent au remplacement d'une base de l'ADN par une autre (SNP, *Single Nucleotide Polymorphism*). Ces remplacements sont soit des **transitions** (substitution d'une base purique par une base purique ( $A \leftrightarrow G$ ) ou d'une base pyrimidique par une base pyrimidique ( $C \leftrightarrow T$ )), soit des **transversions** (substitution d'une base purique par une base pyrimidique et réciproquement ( $A, G \leftrightarrow C, T$ )). Il existe donc quatre transitions possibles ( $A \rightarrow G, G \rightarrow A, C \rightarrow T$  et  $T \rightarrow C$ ) et huit transversions ( $A \rightarrow C, A \rightarrow T, G \rightarrow C, G \rightarrow T, C \rightarrow A, C \rightarrow G, T \rightarrow A$  et  $T \rightarrow G$ ). Les autres mutations ponctuelles correspondent à l'**insertion** ou à la **délétion** d'une base. Dans tous les cas, les mutations modifient la séquence de l'ADN.

Lorsqu'elles se produisent à proximité ou dans les gènes eux-mêmes, les mutations modifient le cadre de lecture de ces gènes, ce qui peut altérer (ou non) la transcription et la traduction de l'information génétique. Tout changement de la séquence de l'ADN peut conduire à la nonexpression des gènes (mutations délétères) ou également à la modification des acides aminés et donc des propriétés des protéines exprimées. Par contre, si les mutations ont peu d'effet sur les structures ou le fonctionnement de l'organisme, on dit qu'elles sont « neutres » (à la base de la théorie neutraliste proposée par Kimura en 1983). Lorsque ces dernières surviennent dans l'ADN non codant, elles ne sont pas soumises à une quelconque pression de sélection et peuvent s'accumuler dans le génome de génération en génération. Ces caractéristiques permettent d'expliquer que les régions non codantes de l'ADN présentent des variations importantes d'un individu à l'autre.

Ces événements générant du polymorphisme peuvent être liés à divers mécanismes. Les mutations ponctuelles pourraient s'expliquer par des erreurs du système de réplication (duplication de l'ADN avant chaque division cellulaire) ou de réparation de l'ADN (erreurs de l'ADN polymérase), par des altérations chimiques spontanées (hydrolyse des bases) ou par exposition à des agents mutagènes endogènes (radicaux libres) ou exogènes (rayonnement UV). Les délétions, les insertions, les duplications et les transpositions peuvent résulter de mauvais appariements entre certains chromosomes homologues (crossing-over inégaux) aboutissant à des régions génomiques de structures et/ou de tailles différentes.

## II.2. Les différents polymorphismes

Les polymorphismes génétiques sont généralement classés en deux catégories : les polymorphismes classiques, détectables directement (ou indirectement dans le cas des groupes sanguins érythrocytaires) à partir des produits de l'expression génique ; et les polymorphismes moléculaires qui analysent directement la variation de l'ADN.

☞ Les **polymorphismes classiques** : Avant le développement de la génétique moléculaire, l'analyse des gènes (et des génotypes) était indirecte et limitée à un petit nombre de marqueurs. Ces derniers représentaient les produits des gènes, les protéines, dont une centaine était connue. Aujourd'hui, de nombreux polymorphismes classiques sont étudiés. On les répertorie en quatre ensembles :

- a. les groupes sanguins (ABO, MNS, Rh, Duffy, Kidd...);
- b. les enzymes érythrocytaires (phosphoglucomutases, estérases...);
- c. les protéines sériques (haptoglobines, immunoglobulines...);
- d. le système HLA (*Human Leucocyte Antigen*) mis en évidence au niveau des leucocytes et des plaquettes sanguines (Dausset, 1985).

☞ Les **polymorphismes moléculaires** : Ils rassemblent tous les polymorphismes de la séquence d'ADN. Les avantages d'analyser les polymorphismes génétiques de l'ADN plutôt que ceux des produits de gènes sont divers :

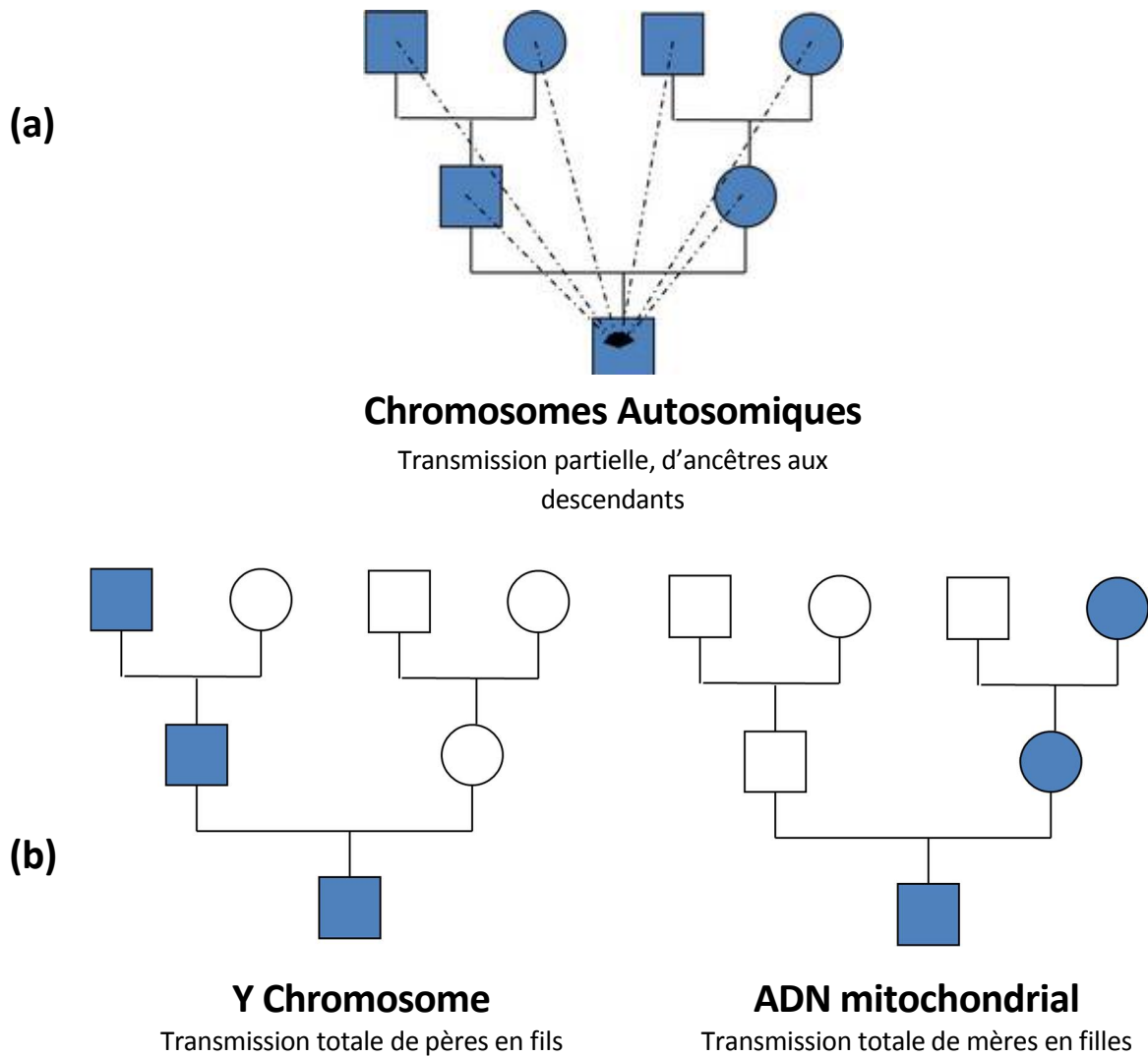
- a. Il y a plus d'informations génétiques dans la séquence d'ADN que dans la séquence protéique ;
- b. Il existe un plus grand nombre de polymorphismes génétiques ;
- c. Les techniques d'analyse de biologie moléculaire sont souvent les mêmes quel que soit le segment d'ADN considéré ;
- d. L'automatisation de ces techniques est plus facile, ce qui est d'un grand secours lorsque l'on doit analyser un grand nombre de marqueurs.

En effet, Les polymorphismes moléculaires concernent aussi bien le génome nucléaire que les génomes extranucléaires et touchent de même la fraction codante et non codante de l'ADN. Dans le génome nucléaire humain, ils se situent aussi bien sur les **autosomes** que sur les **chromosomes sexuels**. On rencontre donc des variations au niveau de l'ADN mitochondrial, de gènes nucléaires, de séquences non codantes répétées ou dispersées, d'autosomes, du chromosome X, du chromosome Y, etc.

Dans nos travaux, trois types de polymorphismes moléculaires ont été analysés dans la population marocaine étudiée: le polymorphisme des microsatellites ou **STR** (abréviation du terme anglais **Short Tandem Repeat**), le polymorphisme ponctuel de séquence **SNP** (**S**ingle **N**ucleotide **P**olymorphism) du **chromosome Y**, et celui de l'**ADN mitochondrial**. Le polymorphisme moléculaire au niveau du chromosome Y et de l'ADN mitochondrial sont considérés comme des **marqueurs haplotypiques** ou bien **marqueurs uni-parentaux**.

### III. Les marqueurs haplotypiques

Les marqueurs autosomaux subissent un brassage génétique à chaque génération vu qu'ils sont transmis de manière biparentale. Ainsi, la moitié de l'information génétique d'un individu lui provient de son père et l'autre moitié de sa mère (Figure 2, a). Les marqueurs uni-parentaux, c'est-à-dire ceux situés sur le chromosome Y et sur l'ADN mitochondrial, sont transmis d'une génération à l'autre sans changement sauf dans le cas de mutations. Les marqueurs de l'ADN mitochondrial, transmis de mère à enfant, permettent de retracer les lignées maternelles (Oota et *al.*, 1995) alors que ceux du chromosome Y, transmis de père en fils, permettent de reconstituer les lignées paternelles (Schultes et *al.*, 1999)(Figure 2, b). L'hérédité uni-parentale et l'absence de recombinaison du chromosome Y et la transmission cytoplasmique de l'ADN mitochondrial sont des avantages inhérents (comme la construction précise des haplotypes des individus, un seul ancêtre paternel/maternel, une estimation facile de la chronologie des événements des arbres phylogénétiques utilisant des taux de mutation calibrés,...) qui rendent l'interprétation des histoires génétiques des hommes moins complexe que celles déduites de l'analyse des gènes présents sur les autosomes. Cette caractéristique rend les marqueurs moins informatifs pour l'identification individuelle. L'information génétique de chaque marqueur uni-parental est appelée haplotype au lieu de génotype puisqu'un seul allèle est détecté par individu.



**Figure 2:** (a) Transmission partielle de l'information génétique au niveau des chromosomes autosomiques à travers les générations.

(b) Transmission totale de l'information génétique contenue au niveau du chromosome Y et de l'ADN mitochondrial.

**Source:** J.M. Butler (2005) *Forensic DNA Typing*, 2<sup>nd</sup> Edition © 2005 Elsevier Academic Press.

# *Chapitre C*

**Les marqueurs haplotypiques**

# L'ADN MITOCHONDRIAL

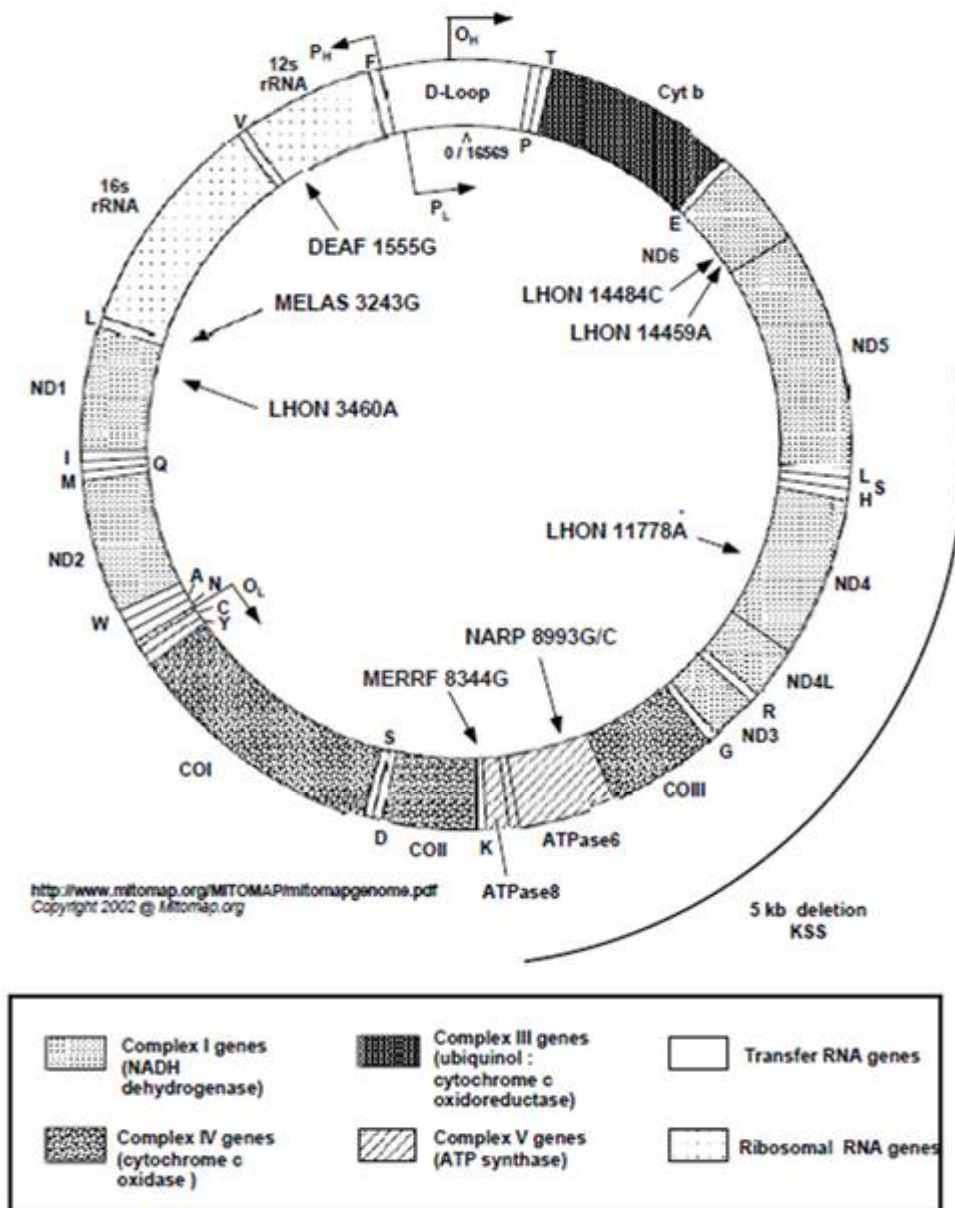
## I. Généralités sur l'ADN mitochondrial

### I.1. Définition

Chez l'Homme, le génome mitochondrial consiste en une molécule d'ADN double brin circulaire fermée, longue de **16 569** paires de bases (pb), soit une taille d'environ 5mm. Il représente moins de 1% de l'ADN du noyau cellulaire. On le rencontre en de nombreuses copies (2 à 10) dans les mitochondries, organites cellulaires cytoplasmiques d'origine endosymbiotique dont la fonction primaire est de fournir de l'énergie à la cellule par phosphorylation oxydative. Chaque cellule contient des centaines de mitochondries et par conséquent des milliers de molécules d'ADN mitochondrial (ADNmt).

La mitochondrie possède un système génétique complet, partiellement autonome par rapport à l'ADN du noyau. La réplication et la transcription de l'ADNmt lui sont propres. Le **génome** mitochondrial est très compact (c'est à dire présentant très peu d'introns) et se compose de deux principales régions : une *région codante* et une *région non codante*. La première contient les 37 gènes codant 13 protéines de la chaîne respiratoire, 22 ARN de transfert (ARNt) et 2 ARN ribosomiaux (ARNr). La région non codante, dont la partie la plus étendue est la D-Loop (*Displacement-loop*), comporte les origines de réplication des deux filaments (léger et lourd) de la molécule d'ADN, divers éléments de régulation de l'expression des gènes et des régions hypervariables (Figure 3).

Le génome mitochondrial humain a été intégralement séquencé en 1981 (Anderson *et al.*, 1981) à partir d'un échantillon de placenta. La séquence obtenue a été révisée en 1999 (Andrews *et al.*, 1999) et est utilisée depuis comme séquence de référence dite séquence **CRS** (*Cambridge Reference Sequence*) ou **rCRS** (*revised Cambridge Reference Sequence*).



**Figure 3 :** Structure du génome mitochondrial humain.

**Source :** <http://www.mitomap.org/MITOMAP/mitomapgenome.pdf>

Le génome mitochondrial comporte différentes régions fonctionnelles : des gènes codant pour 2 ARN ribosomiaux (12S et 16S), des gènes codant pour 22 ARN de transfert indiqués par des lettres individuelles (représentant les acides aminés correspondant) et des gènes codant pour 13 protéines (ex : ND1 à 6, ATPase 6 et 8...). Le génome se compose de 2 filaments : léger (L) et lourd (H), ce dernier étant plus riche en base G et T. La région non codante contient les promoteurs pour la transcription des 2 filaments ainsi que leurs origines de réplication (OH et OL). La région non codante la plus étendue est la D-loop (*Displacement loop*).

## I.2. Propriétés et avantages de l'ADNmt

L'ADNmt a une **hérédité** cytoplasmique (non mendélienne) *uniparentale* et plus précisément **maternelle** (Giles *et al.*, 1980 ; Case et Wallace 1981). En effet, lors de la fécondation, les mitochondries du spermatozoïde (de l'ordre d'une cinquantaine) ne passent quasiment pas dans l'ovocyte ; le cytoplasme de la cellule-oeuf provient uniquement du gamète femelle (contenant de 400 à 200 000 mitochondries). La rare présence de l'ADNmt paternel équivaut à 1% de l'ADNmt total du zygote et il n'a pas été mis en évidence d'ADNmt mixé par les deux parents.

Ce caractère haploïde de l'ADNmt fait qu'il n'est pas soumis à la recombinaison génétique. Ainsi, les diverses mutations qui ont pu se produire au cours de l'évolution se sont accumulées et ont été transmises telles quelles de mère en fille, de génération en génération. L'ADNmt a un **taux de mutation** très élevé, de 10 à 17 fois supérieur à celui de l'ADN nucléaire (Brown *et al.*, 1980 ; Wallace *et al.*, 1987), conduisant à un haut niveau de polymorphisme. Ainsi, cette petite fraction d'ADN extranucléaire est particulièrement adaptée aux études de filiation génétique des lignées maternelles.

Le **polymorphisme** de l'ADNmt est donc la conséquence de l'accumulation successive de mutations dans la séquence nucléotidique aussi bien au niveau de la région contrôle (segments hypervariables I II et III) que dans la région codante. Les combinaisons de ces mutations ont permis de décrire différents haplotypes que l'on peut regrouper, par certaines positions nucléotidiques spécifiques, en **haplogroupes**. Chaque haplogroupe mitochondrial représente un ensemble de séquences qui dérivent toutes d'une séquence initiale ancestrale. Mais, selon le niveau de description des mutations (typées dans la région contrôle et/ou dans la région codante), cette définition des haplogroupes peut poser quelques problèmes lors de la comparaison de différents travaux. L'uniformisation des données devant être respectée, on doit parfois tenir compte seulement des grandes catégories d'haplogroupes (H, L, T, X, etc.), perdant alors certaines informations sur la diversité mitochondriale totale des populations considérées. Néanmoins, on a pu constater que les fréquences des haplogroupes varient suivant les populations et les régions géographiques. Quelques-uns de ces marqueurs ont même été détectés seulement dans certaines populations (Mishmar *et al.*, 2003).

La majorité des mutations dans la séquence de l'ADNmt sont sélectivement neutres et n'ont donc pas de conséquence sur la fonction du génome. Néanmoins, certaines mutations peuvent être responsables de **maladies** (Wallace *et al.*, 1999).

Une autre caractéristique des mutations génétiques mitochondriales est le fait qu'elles ne sont pas nécessairement représentées sur toutes les molécules d'ADNmt. Il y a en général coexistence de gènes mutés et de gènes normaux, dans le même tissu, dans la même cellule voire dans la même mitochondrie. Cette propriété est connue sous le nom d'**hétéroplasmie**. Cette hétéroplasmie joue un rôle important dans l'expression des mutations mitochondriales lors des divisions cellulaires et son taux serait lié au vieillissement de l'organisme (Michikawa *et al.*, 1999 ; Thèves *et al.*, 2006).

## II. L'hétéroplasmie: cas particulier rencontré au cours de l'analyse de l'ADNmt

Au moment de la scission des mitochondries lors du recopiage de l'ADN mitochondrial, il se peut produire une erreur de réplication, plus fréquente pour l'ADN mitochondrial que pour l'ADN génomique. Une même cellule peut posséder dans son cytoplasme des mitochondries avec un ADN mutée et des mitochondries avec un ADN provenant de la mitochondrie souche. Un même sujet peut avoir deux types d'ADN mitochondriaux.

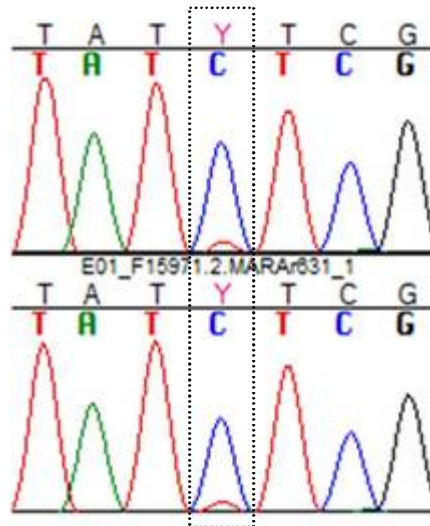
Les méthodes d'analyse ne permettent pas de séparer les deux types d'ADN mitochondriaux (mutes et non mutes) présents dans la cellule. Lors de la révélation, les deux séquences des deux ADN mitochondriaux vont se superposer, le chromatogramme va présenter une hétéroplasmie.

Il existe deux types d'hétéroplasmies, celles induisant une variation de séquence (**Hétéroplasmie ponctuelle**) et celles induisant une variation de longueur (**Hétéroplasmie de taille**).

### II.1. Hétéroplasmie ponctuelle

Une transversion ou une transition sur l'un des ADN, superposition de deux bases sur le chromatogramme à une même position (Figure 4). La présence d'une hétéroplasmie ponctuelle va augmenter le pouvoir discriminant de l'analyse de l'ADNmt. Un exemple remarquable d'un individu dont l'hétéroplasmie a été découverte par hasard **Nicolas II de Russie**, dont l'hétéroplasmie (et

celle de son frère) ont servi à convaincre les autorités russes de l'authenticité de ses restes (Pavel et al., 1996).



**Figure 4 :** Hétéroplasmie ponctuelle, superposition de deux bases, ici, la Thymine et la Cytosine, elle est noté Y d'après le code IUPAC (International Union of Pure and Applied Chemistry).

**Source :** Donnée brute d'une séquence d'un individu de la population Marocaine analysée dans le cadre de cette thèse.

## II.2. Hétéroplasmie de taille

Ce type d'hétéroplasmie est fréquemment observé au niveau des régions hypervariables de la région contrôle de l'ADNmt HVI, HVII et HVIII qui contiennent des segments de séquences qui sont homopolymériques (c'est-à-dire base C). Dans HVI, une région homopolymérique commence au nucléotide numéro 16184. Quand une transition en position 16189 (T-C) est observée, une hétéroplasmie souvent apparaît (Figure 5).



populations étaient les plus anciennes (forte accumulation de mutations) et donc que la source du pool génique mitochondrial actuel se situait en Afrique. Ils ont présenté leurs résultats sous la forme d'un arbre généalogique. Selon le principe de l'horloge moléculaire, ils ont daté le type mitochondrial ancestral autour de 140 000 à 190 000 ans. Cette hypothétique femme à l'origine de toutes les lignées maternelles actuelles a été baptisée « l'Ève mitochondriale ». Cette conclusion apparaît toutefois comme abusive à de nombreux chercheurs (Pellegrini 1995). De nombreuses critiques ont été faites sur le choix des individus contemporains analysés, sur l'effectif insuffisant de ces groupes pour estimer correctement les fréquences géniques, sur l'arbre phylogénétique retenu ou bien encore sur le fait que les auteurs se basent sur l'hypothèse peu concevable que la fréquence des mutations soient constantes dans le temps (concept de l'horloge moléculaire).

### III.2. Les études paléanthropologiques

L'ADN mitochondrial est un marqueur de choix dans les études paléanthropologiques lorsqu'il existe une importante dégradation des échantillons, une grande ancienneté du matériel ou de mauvaises conditions de conservation. En effet, dans ces situations, l'ADN mitochondrial reste en bon état par rapport à l'ADN nucléaire fragmenté et, de par sa présence en de nombreuses copies dans les cellules, on peut l'extraire facilement. Sa recherche dans les populations anciennes a fait l'objet de diverses publications. Par exemple, l'étude de l'ADNmt de « l'Homme des glaces » (Spindler 1995) a révélé qu'il était porteur d'un type mitochondrial encore commun en Europe centro-septentrionale et présent en Italie du Nord, l'haplogroupe K1 (Handt *et al.*, 1994 ; Rollo *et al.*, 2006). L'analyse de petits fragments d'ADN extrait de squelettes de Néandertaliens a montré que leurs séquences étaient assez éloignées des séquences humaines actuelles (Krings *et al.*, 1997 ; Ovchinnikov *et al.*, 2000 ; Serre *et al.*, 2004 ; Orlando *et al.*, 2006). Les auteurs suggèrent une nette discontinuité génétique entre les Hommes anatomiquement modernes et les Néandertaliens tandis qu'une continuité expliquerait la généalogie des hommes modernes depuis le Paléolithique supérieur jusqu'à nos jours. L'équipe du Laboratoire d'Anthropologie Moléculaire de Strasbourg (dirigée par le Pr B. Ludes) s'est spécialisée sur l'étude de l'ADN ancien et travaille particulièrement sur des restes osseux ou momifiés provenant de Mongolie (Keyser-Tracqui *et al.*, 2003), Sibérie (Ricaud *et al.*, 2005), et Yakoutie (Ricaud *et al.*, 2006). Enfin, les polymorphismes de l'ADNmt de sujets anciens d'Espagne (Casas *et al.*, 2006) et du Maroc (Kéfi *et al.*, 2005) ont récemment pu être analysés.

### III.3. L'ADN mitochondrial de l'intérêt scientifique à la pratique judiciaire

L'ADN mitochondrial est particulièrement intéressant du fait du nombre important de copies par cellule et de sa protection à l'intérieur des mitochondries, le rendant ainsi moins sensible aux dégradations, contrairement à l'ADN nucléaire.

Du fait de sa plus grande sensibilité, l'ADN mitochondrial est utile lorsque les prélèvements sont dégradés, voire putréfiés ou quand un cheveu est dépourvu de bulbe et que l'étude de l'ADN génomique se révèle impossible.

Cependant, cette distinction maternelle s'avère problématique lorsqu'il s'agit d'inclure un individu puisque toutes les personnes de la même lignée maternelle (frère, sœur, mère, grand-mère, cousin par la lignée maternelle...) possèdent le même ADN mitochondrial.

L'ADN mitochondrial ne permet pas d'inclure un individu comme l'ADN génomique peut faire. Cette analyse doit être pratiquée en complément de l'analyse génomique ou dans les cas d'absence de résultats sur l'ADN génomique. Toutefois, l'ADN mitochondrial permet d'exclure un individu.

L'analyse de l'ADNmt a permis de résoudre quelques spectaculaires cas historiques comme, par exemple, l'identification de restes de la famille Romanov (Gill et *al.*, 1994; Ivanov et *al.*, 1996), de Jesse James (Stone et *al.*, 2001) et de Martin Bormann (Anslinger et *al.*, 2001).

Il existe des bases de données tel que **EMPOP** (<http://empop.org/>), **mtDNA manager** (<http://mtmanager.yonsei.ac.kr/>), **GeneBank** (<http://www.ncbi.nlm.nih.gov/genbank/>), recensant les mitotypes de plusieurs populations mondiales. La comparaison du mitotype obtenu lors de l'analyse d'un extrait au cours d'une enquête criminelle avec les bases de données internationales nous renseigne sur la fréquence d'apparition de ce mitotype. Cela permet de déterminer sa rareté relative. Plus des bases de données seront de grandes de tailles, plus la fréquence d'apparition de ce mitotype sera précise.

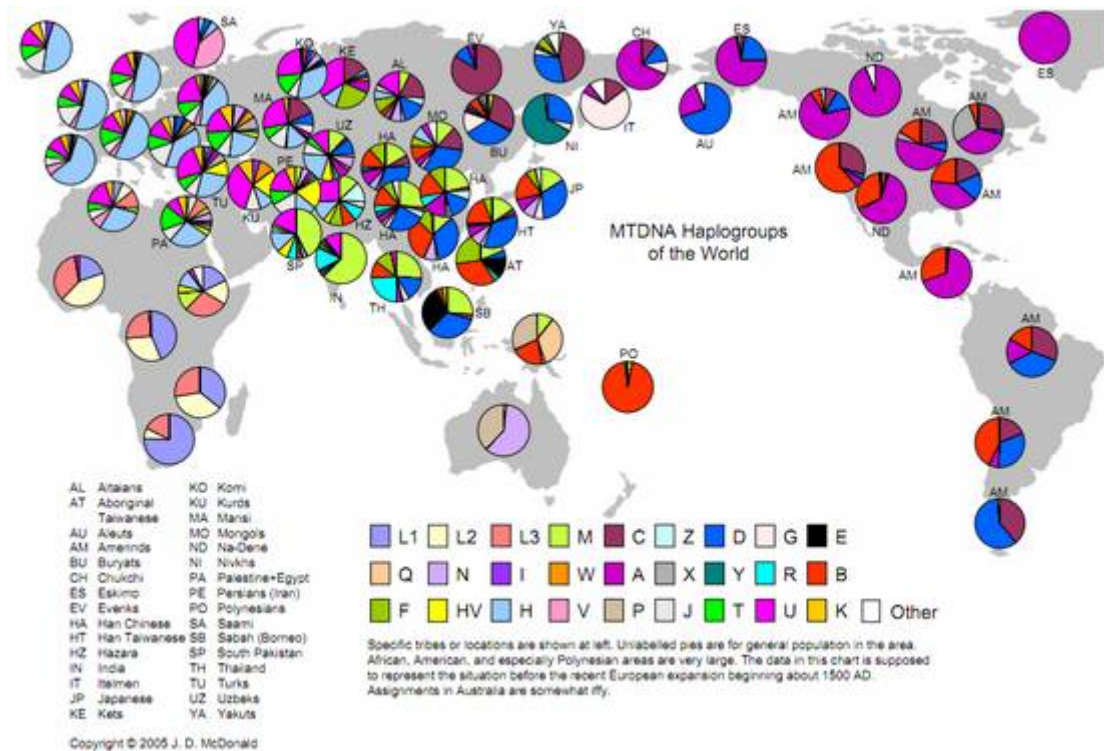
## IV. Les haplogroupes mitochondriaux

Les études des mutations de l'ADN mitochondrial de diverses populations humaines mondiales ont permis de mettre en évidence différents haplotypes qui ont été regroupés, d'après leurs

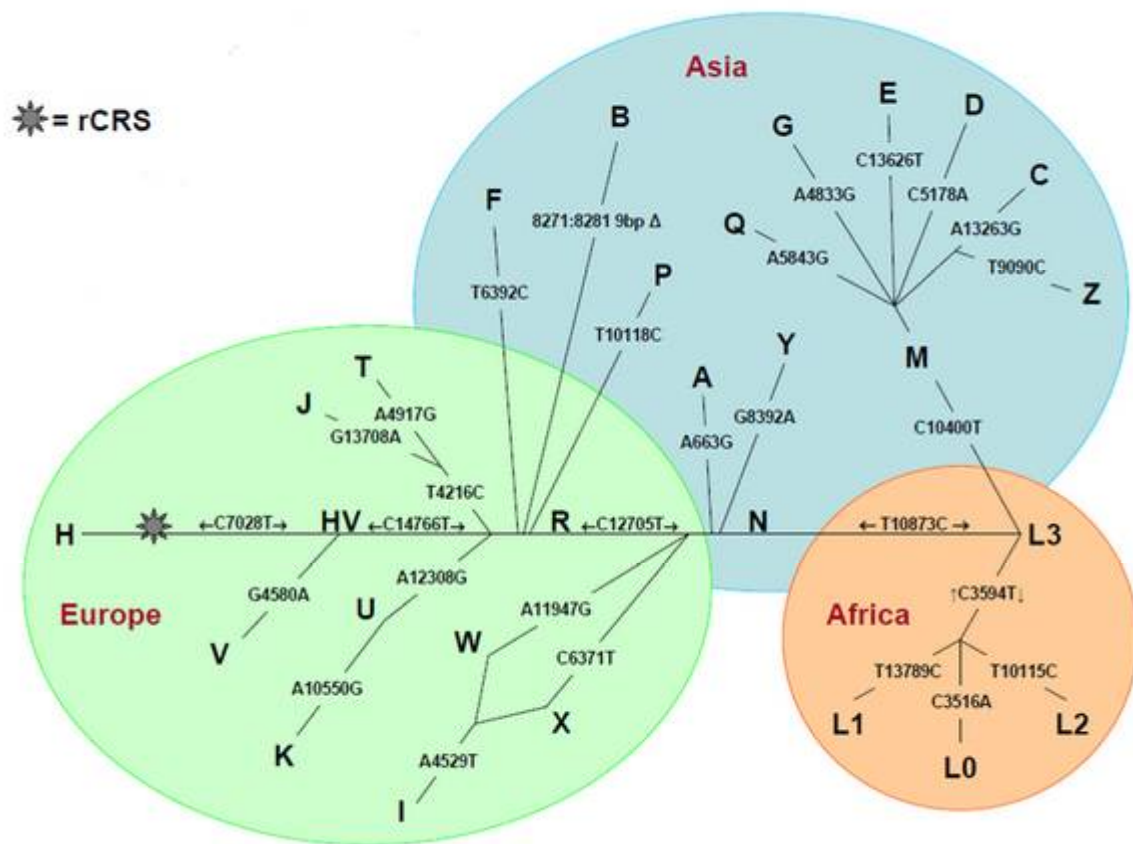
relations phylogénétiques, en haplogroupes. Ces haplogroupes mitochondriaux sont répartis en différentes familles très fréquentes dans certaines régions géographiques voire même dans certaines populations. On distingue aujourd’hui divers ensembles d’haplogroupes rencontrés à des fréquences élevées en Afrique, en Europe, en Asie, dans certaines îles du Pacifique et en Amérique.

Naturellement, des haplogroupes communément observés dans une région donnée sont aussi retrouvés ailleurs (mais à des fréquences plus basses), en raison des mouvements de populations (figure 6). Les lignées A, B, C et D sont fréquentes en Amérique et à l’Est de l’Asie. Les populations européennes ont des fréquences élevées d’haplogroupes H, HV, I, J, K, T, U, V, W et X. Les lignées fréquentes au Proche-Orient, au sud de l’Asie et en Océanie sont M et N (lignées ancestrales), E, F, G, P, Q, R, Y et Z. Une lignée est uniquement détectée dans les populations nord-africaines : U6.

Enfin, les régions sub-sahariennes présentent un pourcentage élevé d’haplogroupes de la famille L. Un arbre phylogénétique simplifié des relations de parenté entre les divers haplogroupes mitochondriaux est présenté dans la figure 7. Une carte géographique des chemins de migrations humaines ayant conduit à la répartition actuelle des haplogroupes de l’ADNmt est présentée dans la figure 8.

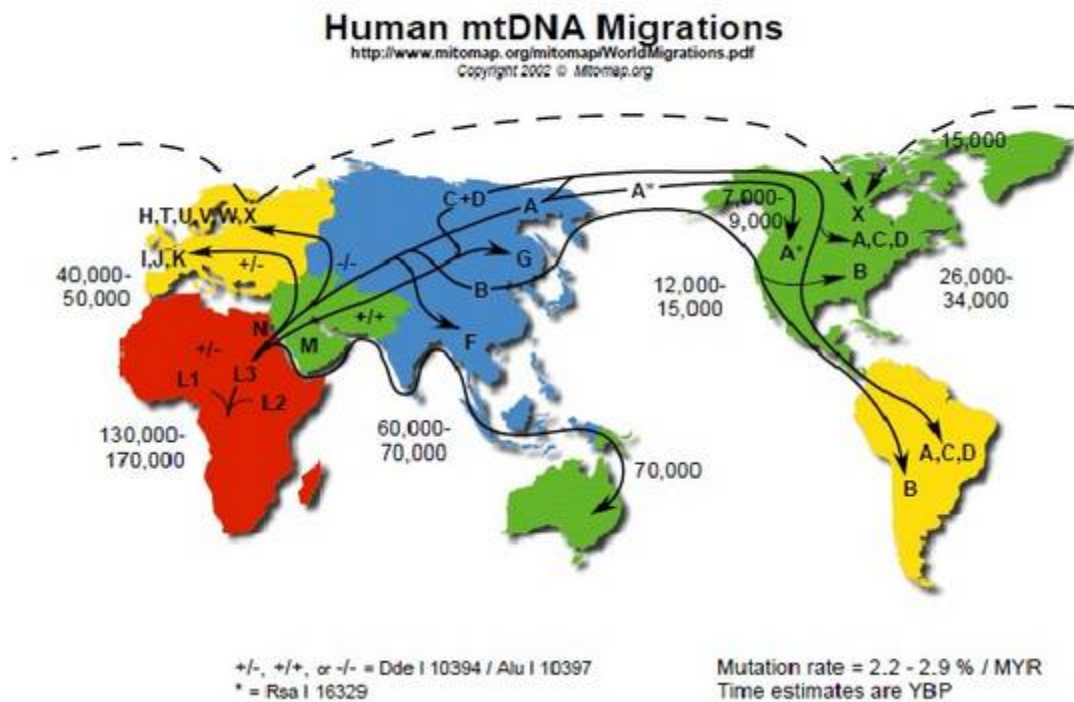


**Figure 6 :** Les haplogroupes mitochondriaux présents dans les populations humaines mondiales.  
**Source :** <http://www.scs.uiuc.edu/mcdonald/WorldHaplogroupsMaps.pdf>.



**Figure 7 :** Arbre phylogénétique simplifié des haplogroupes mitochondriaux rencontrés dans les populations humaines.

**Source:** <http://mitomap.org/pub/MITOMAP/MitomapFigures/simple-treemitomap 2009.pdf>



**Figure 8 :** Carte géographique retraçant les principales migrations des haplogroupes mitochondriaux.

**Source :** <http://www.mitomap.org/mitomap/WorldMigrations.pdf>.

Copyright 2002 © Mitomap.org

L0 serait apparu en Afrique entre 156.000 et 170.000 ans BP. Ceci correspondrait avec les dates d'apparition de l'Homme moderne en Afrique (environ 150 000 ans, Stoneking 1993). Une première expansion africaine, il y a entre 122.000 et 132.000 ans BP, aurait donné les lignées L1 et une deuxième expansion (entre 85.000 et 95.000 ans BP) aurait conduit à l'émergence des lignées L2. Une nouvelle expansion vers 59.000-69.000 ans BP aurait permis l'apparition des lignées L3 dont une partie est à l'origine des haplogroupes mitochondriaux M et N, précurseurs de toutes les autres lignées répertoriées, marquant ainsi la première migration hors d'Afrique (Maca-Meyer *et al.*, 2001). Cependant, il est probable que le pool génétique des hommes modernes africains était déjà composite mais que seule la fraction de la population portant l'haplogroupe ancestral L3 aurait quitté l'Afrique pour coloniser le reste du Monde il y a moins de 60 000 ans. D'autres migrations ont ensuite permis l'émergence de nouveaux haplogroupes en Europe (40.000-50.000 ans BP), en Asie et Océanie (60.000-70.000 ans BP) puis en Amérique (12.000-15.000 ans BP).

# LE CHROMOSOME Y

*Jusqu'à présent, le chromosome Y semble jouer un rôle d'un délinquant juvénile entre les autres chromosomes humain ; pauvre en attributs utiles, retissant avec ses voisins et il a une tendance inéluctable à dégénérer...*

*(Mark Jobling et Chris Tyler-Smith 2003)*

## I. Généralités sur le chromosome Y

### I.1. Évolution du chromosome Y

Au cours de l'évolution des vertébrés, la différenciation sexuelle, aboutissant le plus souvent à des phénotypes différents chez les mâles et les femelles, s'est accompagnée de la spécialisation d'une paire de chromosomes, portant alors le nom de chromosomes sexuels ou gonosomes. Auparavant, il n'existait que des autosomes et le déterminisme sexuel était sous la dépendance de facteurs environnementaux, comme la température ou la nutrition, ainsi qu'on peut l'observer encore aujourd'hui pour certains reptiles ou poissons. Cette spécialisation d'une paire de chromosomes vers des fonctions de déterminisme du sexe et de reproduction s'est accompagnée d'une divergence des gonosomes, à la fois en terme de taille et de contenu en gènes, pour aboutir actuellement à un hétéromorphisme marqué entre les deux chromosomes sexuels (Figure 9).



**Figure 9 :** La paire de chromosomes sexuels humains. Vue prise au microscope électronique à balayage.

**Source :** Nature (2003) 423 : 810-812.

Chez les mammifères, ce phénomène de spécialisation et de divergence a porté principalement sur le chromosome Y puisque le chromosome X n'est pas particulièrement enrichi en gènes impliqués dans le déterminisme sexuel ou les fonctions reproductrices (Ravel et *al.*, 2006). Il y a environ 300 millions d'années, sur une paire d'autosomes, aurait existé un gène **SXD** (**SeX Determining**) impliqué dans la différenciation du sexe et existant sous deux formes alléliques, **SXD** et **sxd**. La combinaison **SXD/sxd** aurait été caractéristique des mâles et la combinaison **sxd/sxd** des femelles. Sur la même paire chromosomique, un autre gène, **SST** pour **Sexually Selected Trait**, aurait muté pour donner un allèle **SSTM** qui aurait avantage les mâles **SXD/sxd** la portant. Ainsi aurait été couplée une combinaison de gènes **SXD** et **SSTM** favorisant, d'une part un gène de déterminisme mâle, et d'autre part, un gène de fertilité, assurant ainsi une sélection positive à cette association. Les gènes **SOX3**<sup>1</sup> (**Sry-related HMG box**) et **SRY** (**Sex determining Region Y**) actuels, portés respectivement par les chromosomes X et Y, pourraient donc être les descendants du couple d'allèles **SXD** et **sxd**. Le gène SRY est le gène majeur de déterminisme testiculaire et il présente encore des analogies importantes avec le gène SOX3 comme l'existence d'une boîte **HMG** (**High Mobility Group**) et l'absence d'intron. La divergence entre les allèles **SXD-sxd** ou **SRY-SOX3** aurait ensuite été rendue possible grâce à la suppression des recombinaisons génétiques entre les deux proto gonosomes en raison de la survenue de quatre inversions sur le proto chromosome Y au cours des 300 derniers millions d'années (Lahn et *al.*, 1999). En effet, en cas d'inversion, l'appariement de deux chromosomes homologues au cours de la méiose ne peut se faire que par la formation d'une boucle au niveau du segment inversé. Si un crossing over survient à l'intérieur de la boucle, les chromosomes recombinés sont anormaux, car ils portent en plus et en moins les segments extérieurs à la boucle. La transmission de tels chromosomes déséquilibrés à la descendance donne des individus porteurs de tare<sup>2</sup> génétique grave ce qui, d'un point de vue évolutif, revient à ce que les recombinaisons soient interdites à l'intérieur de l'inversion. Ne pouvant plus mélanger leurs allèles sur des régions de plus en plus grandes, les deux proto gonosomes ont donc amorcé une évolution divergente qui les a amenés aux chromosomes X et Y actuels. Le reliquat de ces inversions est représenté par les régions distales des deux gonosomes qui sont restées strictement homologues, puisque les inversions sont des phénomènes interstitiels, et qui portent le nom de régions pseudo-autosomiques, ou **PARs 1 et 2** (**Pseudo-Autosomal Region**).

---

<sup>1</sup> **Gènes SOX** : Famille de gènes qui code pour les facteurs de transcriptome, qui se lient sur des sillons d'ADN mineur et appartiennent à une sous-famille de gène caractérisé par une homologie de séquence connu sous le nom : Groupe de Haute Mobilité (HMG).

<sup>2</sup> **Tare** : Défaut physique ou psychique, généralement héréditaire.

La séparation progressive des chromosomes X et Y a ensuite été accentuée par d'autres événements chromosomiques majeurs comme de nouvelles inversions, des translocations ou des délétions, qui à chaque fois ont empêché encore un peu plus les recombinaisons (Figure 10). Le fait que l'Y ne pouvait plus recombiner avec un homologue, alors que les chromosomes X pouvaient toujours le faire chez les femelles XX, a eu des conséquences multiples sur la structure et la composition en gènes du chromosome Y. La première est que toutes les mutations délétères, ou toutes les variations neutres, ne pouvaient plus être éliminées par recombinaison avec un allèle homologue normal et qu'elles se sont accumulées dans l'Y, transformant ce dernier en une sorte de poubelle de l'évolution. Le chromosome Y est transmis uniquement de père en fils. De nouveau, c'est cette particularité qui a été utilisée pour étudier l'origine des populations humaines, cette fois-ci, du côté des hommes (Mitchell et al., 1996).

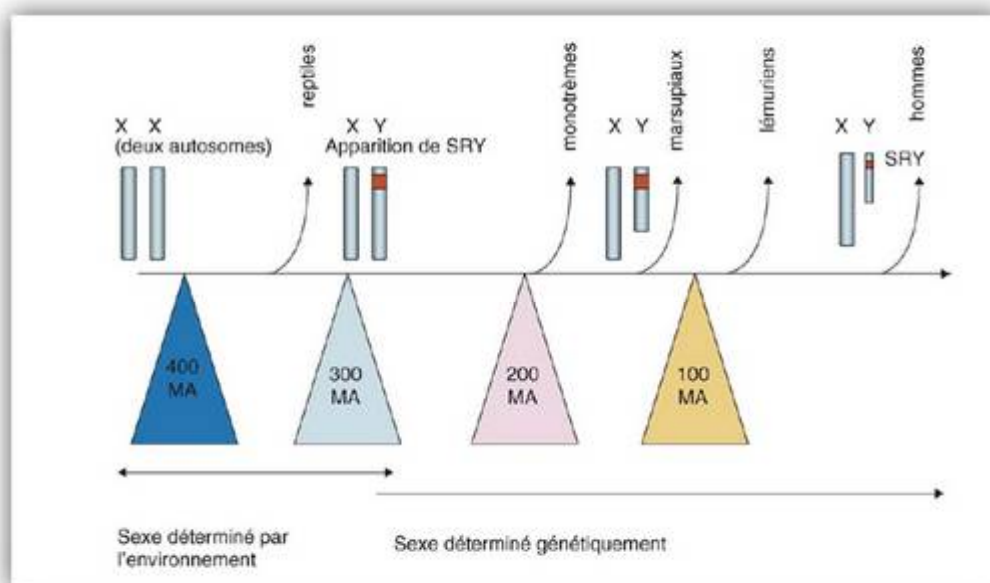


Figure 10 : Schéma illustrant l'évolution du chromosome Y.

Source : Ravel et al., 2006.

## I.2. Structure actuelle du chromosome Y

La région spécifiquement mâle, c'est-à-dire située entre les régions **PARs** 1 et 2, est appelée **MSY** (**M**ale **S**pecific **Y** chromosome) et représente la partie de l'Y qui normalement ne recombine jamais avec le chromosome X (précédemment appelée **NR**Y pour **N**on **R**ecombining **Y**

chromosome). La région **MSY** constitue près de 95 % de la longueur du chromosome et montre, dans la partie distale du bras long, une zone d'hétérochromatine non fonctionnelle qui présente des variations de taille importantes selon les hommes. La partie fonctionnelle ou euchromatique de **MSY** s'étend sur environ 23 méga bases, à la fois sur le bras court et la partie proximale du bras long, ce qui représente un peu moins de 1 % du génome humain (Figure 11).

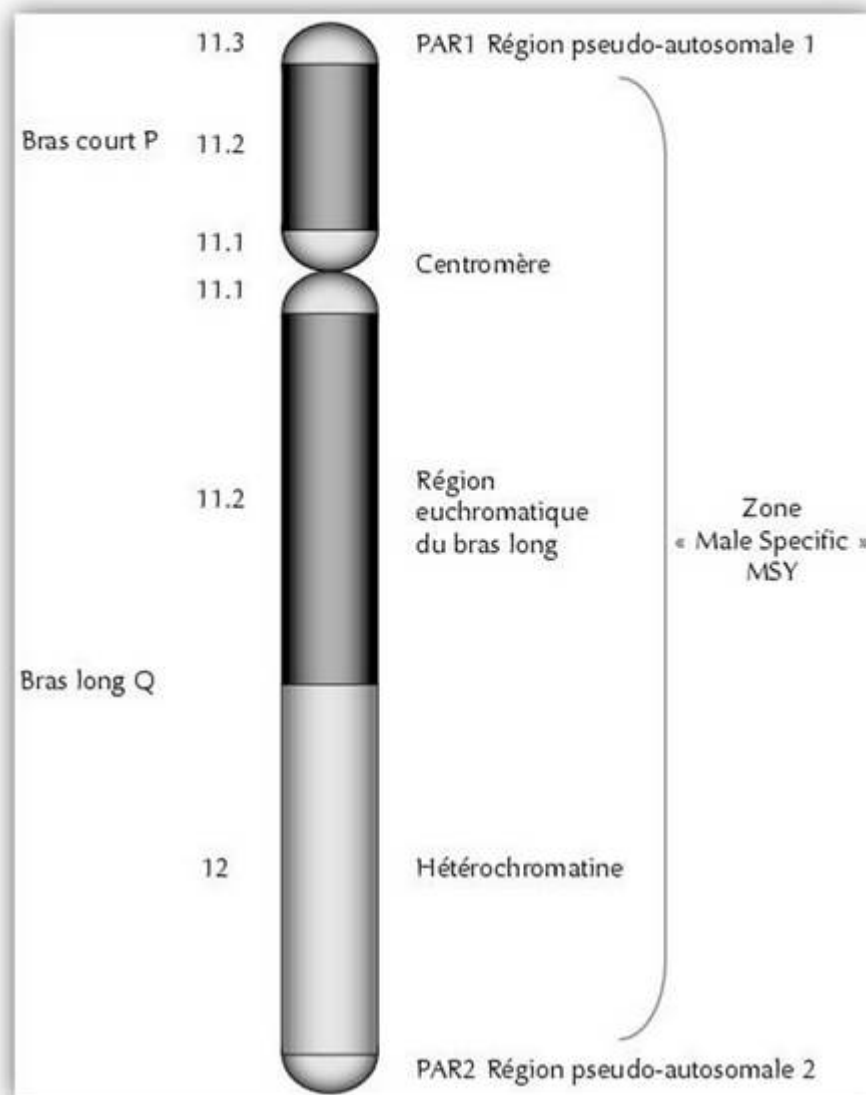


Figure 11: Schéma de la structure du chromosome Y.

Source : <http://www.jle.com/>

La structure chromosomique de **MSY** diffère de celle des autres chromosomes sur deux points particuliers :

- ✓ Il s'agit d'une région qui n'existe que chez les mâles ;
- ✓ Cette région ne présente pas de recombinaison méiotique inter chromosomique (crossing-over).

De part et d'autre se trouvent deux régions pseudo autosomales, nommées **PAR (Pseudo Autosomal Region)** où la recombinaison inter chromosomique est possible avec le chromosome X, événement survenant normalement et fréquemment pendant la méiose masculine.

### **1.3. Diversité du chromosome Y**

La conséquence directe de l'absence de recombinaison du chromosome Y avec le chromosome X, en dehors des zones PAR, est que toute modification de la structure de ce chromosome (mutation, délétion, duplication, insertion ou inversion) va se trouver fixée et être transmise de génération en génération.

L'utilisation de marqueurs génétiques (marqueurs binaires d'évolution lente type **SNP** pour **Single Nucleotide Polymorphism** ou encore marqueurs microsatellites d'évolution rapide) permet de reconnaître des variétés différentes de chromosomes Y. Toutes ces variétés dressent un véritable arbre phylogénétique du chromosome Y humain (Jobling et *al.*, 2003). Ces différentes sortes de chromosomes Y traduisent la diversité et l'extension de l'espèce humaine sur l'ensemble de la planète. Une mutation binaire survenant sur un chromosome Y particulier pourra donc donner naissance à un nouvel haplogroupe. Ces modifications ponctuelles sont à l'origine de plus de 150 haplogroupes différents, dont les fréquences varient selon les origines ethniques et géographiques (Chromosome Y Consortium, 2002). Ces différents haplogroupes ont donc une transmission paternelle et ne changent que par mutation. L'étude de ces mutations dans la région mâle spécifique du chromosome Y permet d'étudier des phénomènes intra-alléliques, comme le taux de mutations spontanées dans une population donnée, sans l'interférence des processus inter-alléliques qui surviennent dans les autres chromosomes. Il est ainsi possible d'obtenir des informations spécifiques sur les taux de mutation dans une population donnée (Jobling et *al.*, 2003).

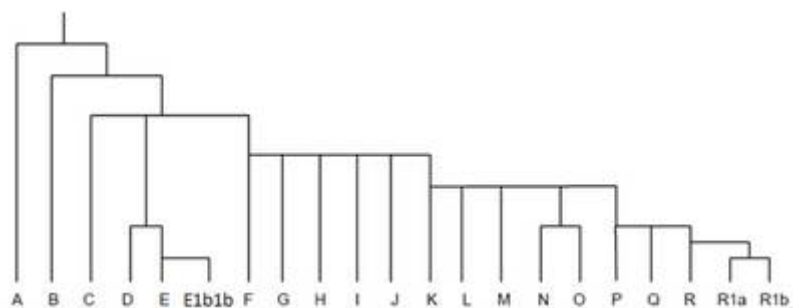
### I.3.1. Les haplotypes du chromosome Y

Un haplotype est un *groupe d'allèles* de différents **loci** situés sur un même chromosome et habituellement transmis ensemble. Haplotype est un mot-valise formé par la contraction de la locution anglaise ***haploid genotype***, ou génotype haploïde. L'ensemble des gènes situés sur un même chromosome et dont les allèles ségrégent ensemble lors de la méiose constituent un haplotype.

La meilleure façon pour caractériser un haplotype et qui est largement répandu est l'utilisation des marqueurs microsatellites ou **STR (Short Tandem Repeat)** du chromosome Y. Ce dernier contient des séquences répétées de nucléotides (de paires de bases). Le nombre de répétitions varie d'une personne à l'autre. Un STR du chromosome Y est désigné par un nombre **DYS (DNA Y-chromosome Segment number)**. Lorsqu'on « test » une personne, on associe au marqueur **DYS** le nombre de répétitions de la séquence STR du chromosome la personne concernée. Ils sont utilisés pour définir les haplotypes, la résolution de l'haplotype croît avec le nombre de marqueurs STR. à l'heure actuel, plus de 77 marqueurs STR du chromosome Y sont répertoriés (Butler et *al.*, 2008).

### I.3.2. Les haplogroupes du chromosome Y

En génétique des populations, chaque branche majeure s'appelle **haplogroupe** (Hg) et chaque sous-branche majeure **sous-haplogroupe**. Le terme *haplogroupe* ou *sous-haplogroupe* n'est pas absolu, il est relatif à l'endroit étudié de l'arbre (Figure 12). La définition de cet arbre est loin d'être achevée si bien que la dénomination des haplogroupes change régulièrement (Jobling et *al.*, 2003).



**Figure 12:** Arbre simplifié pour les haplogroupes du chromosome Y.

**Source :** McDonald, 2005 ; [www.scs.illinois.edu/mcdonald/WorldHaplogroupsMaps](http://www.scs.illinois.edu/mcdonald/WorldHaplogroupsMaps)

La majorité des études utilisent à l'heure actuelle cet arbre généalogique de l'ADN-Y avec sa nomenclature associée. Cette nomenclature a été définie une première fois en 2002 par le **Y Chromosome Consortium (YCC)**.

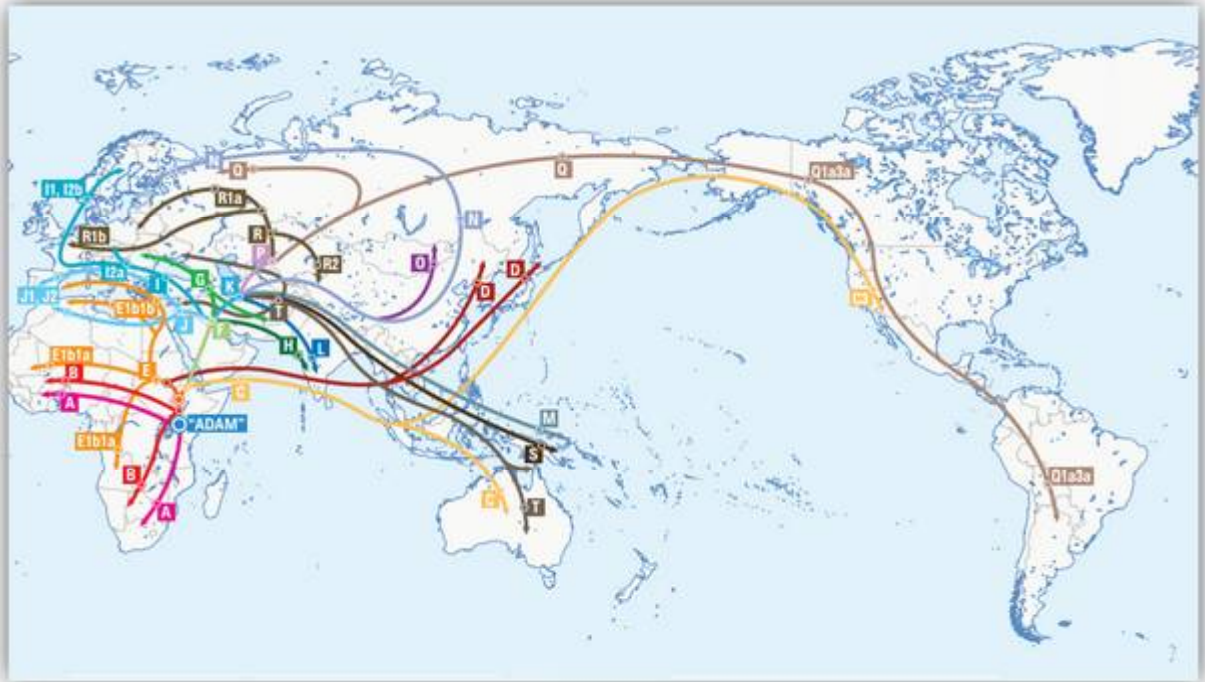
La classification des haplogroupes du chromosome Y est actuellement utilisée pour estimer le groupe patri-linéale d'une population donnée. Les haplogroupes sont identifiés par des lettres allant de A jusqu'à T. La cartographie des haplogroupes pour les marqueurs génétiques *SNP* du chromosome Y sont mise à jour presque chaque année par **ISOGG (International Society of Genetic Genealogy)**.

### **I.3.3. Intérêt du chromosome Y en généalogie et en génétique médico-légale**

L'étude du chromosome Y présente deux avantages principaux : la spécificité de l'ADN masculin lors de l'analyse de mélanges d'ADN et la possibilité de suivre les lignées paternelles. Outre dans les recherches médicales comme celle de l'origine de l'infertilité masculine (Carvalho et al., 2003), les marqueurs du chromosome Y peuvent être utilisés pour de nombreuses applications telles que :

- ✓ L'analyse criminalistique d'indices dans le cadre d'agressions sexuelles (Sibille et al., 2002 ; Parson et al., 2003) en permettant l'amplification spécifique de l'ADN masculin. Ceci peut éviter les extractions différentielles pour séparer les spermatozoïdes des cellules épithéliales ainsi que le masquage du profil masculin par le profil féminin ;
- ✓ La recherche de personnes disparues pour laquelle tout parent en lignée paternelle peut être utilisé comme échantillon de référence (Dettlaff-Kakol et al., 2002 ; Koyama et al., 2002);
- ✓ Les tests de paternité déficiente pour relier les enfants masculins à une lignée paternelle (Santos et al., 1993 ; Jobling et al., 1997 ; Rolf et al., 2001);
- ✓ Les études d'évolution et de migrations humaines (Figure 13), car l'absence de recombinaisons (région non recombinante du chromosome Y, *Non recombining region of the Y chromosome*, NRY) permet la comparaison d'individus masculins séparés par de longues périodes de temps (Underhill et al., 2001 ; Ke et al., 2001) ;
- ✓ Les recherches historiques ou généalogiques puisque dans la majorité des sociétés le nom de famille se transmet de père en fils (Foster et al., 1998 ; Jobling, 2001, Sykes et Irven, 2000 ; Trumme et al., 2004).

Pour tous les marqueurs uni parentaux, l'interprétation statistique est plus compliquée et des corrections appropriées tenant compte des sous-populations et des erreurs d'échantillonnage sont nécessaires (Roewer et al., 2000 ; Roewer et al., 2001).



**Figure 13:** Carte géographique retraçant les principales migrations des haplogroupes du chromosome Y.

**Source :** <http://www.familytreedna.com/pdf-docs/2010-ydna-migration-map.pdf>.

## **Matériel et méthodes**

## I. Objectif du travail

Au début de cette étude, l'objectif est de construire un échantillon de la population marocaine, renfermant trois groupes ethniques (Arabophones, Berbérophones et Sahraouis) afin d'analyser le polymorphisme et établir une base de données (YHRD et EMPOP) en identité humaine des marqueurs uni-parentaux, c'est-à-dire ceux situés sur la région contrôle de l'ADN mitochondriale et ceux du chromosome Y (Marqueurs STR et bi-allélique ou bien SNP).

## II. Echantillonnage

L'étude a été réalisée sur 509 échantillons non apparentés et issus de différentes régions du Maroc. Dans notre collecte, nous nous sommes référés aux recommandations du comité international **Human Genome Diversity (HGD)**. Le consentement des personnes permettant la réalisation de ces travaux a été précédée d'une information appropriée sur la nature et la finalité de l'étude. Nous avons respecté dans l'établissement du questionnaire certaines règles d'ordre éthique comme le respect d'autonomie de la personne, l'anonymisation la plus protectrice vis-à-vis de la confidentialité. Aussi, nous avons tenu dans notre questionnaire à remonter jusqu'à la 3<sup>ème</sup> génération pour vérifier l'appartenance ethnique de chaque individu. (Annexe A).

## III. De l'ADN dans l'échantillon à l'ADN dans le Tube

### III.1. Extraction des acides nucléiques

Quel que soit le type de matrice, le principe des techniques d'extraction d'ADN est sensiblement le même et comprend quatre grandes étapes : la lyse des cellules, la dégradation des protéines libérées, la séparation des acides nucléiques des autres composants et la purification de l'ADN.

Le protocole ci-dessous permet l'extraction des acides nucléiques en utilisant le kit DNA IQ™ System-Database (Promega, Madison, USA). Un volume de 25 µl de sang est mélangé avec 80 µl d'une solution préalablement préparée du *Tampon d'incubation/Protéinase K*. Après incubation d'une heure à 56°C dans un bain marie à sec, un volume de 232 µl d'une solution composé de 225 µl du *Tampon de lyse* et 7 µl de la *Résine* (Billes magnétiques) est ajouté au mélange. Après, le mélange est vortexer pendant 3 secondes au bout de chaque minute pendant 5 minutes. Les tubes sont placés dans un portoir magnétique. Une nette pelote bien distincte est formée sur la paroi du tube en

juxtaposition avec le portoir. Tout le liquide contenu dans le tube est ensuite éliminé sans atteindre la pelote de résine. Par la suite, 100 µl du *tampon de lyse* est ajouté aux échantillons, qui sont vortexés pendant 2 secondes à haute vitesse et à nouveau remis au portoir magnétique. Le *tampon de lyse* est éliminé de chaque tube sans toucher la pelote de résine puis 100 µl du *tampon de lavage* est ajouté par la suite. Les tubes sont vortexés pendant 2 secondes à grande vitesse puis remis dans le portoir magnétique afin d'éliminer la solution du *tampon de lavage*. Cette étape est répétée deux fois pour un total de trois lavages. Sur le portoir magnétique les tubes sont incubés à température ambiante en gardant les bouchons ouverts pour permettre l'évaporation des résidus du *tampon de lavage*. Après lavage, 100 µl du *tampon d'élution* est ajouté aux tubes. L'ensemble est vortexé pendant 2 secondes à grande vitesse puis incubé pendant 5 minutes à une température de 65°C. Après incubations, les tubes sont vortexés pendant 2 secondes et placés immédiatement dans le portoir magnétique. Enfin, l'ADN est transféré soigneusement dans les tubes appropriés.

### III.2. Quantification de l'ADN par PCR en temps Réel

Au vu de la disparité qualitative et quantitative des échantillons à partir desquels l'ADN peut être extrait, il apparaît indispensable de quantifier les extraits avant leur analyse. La détermination quantitative d'ADN se base sur le nombre de cycles minimum (Ct) nécessaire à l'atteinte du seuil correspondant à une fluorescence donnée dans le tube réactionnel. Le seuil est déterminé pour correspondre à la zone d'amplification exponentielle de la réaction de PCR. On distingue deux types de quantifications, l'une absolue se référant à une gamme de concentration connue et la seconde relative se référant à une autre cible interne à l'échantillon tel un gène domestique.

Dans cette étude, les quantifications absolues ont été effectuées grâce au kit Quantifiler™ Human DNA Quantification Kit (Applied Biosystems®, Foster City, CA, USA) basé sur la méthode nucléase fluorogénique 5' Nucléase (TaqMan®). La gamme de concentration pouvant être détectée est comprise entre 23 pg et 50 ng/µl pour 2 µL d'extrait. Les amplifications et lectures ont été faites au moyen du système PCR en temps réel 7500 suivant le protocole décrit par le fabricant (Applied Biosystems®, Foster City, CA, USA).

## IV. Analyse du polymorphisme des marqueurs génétique du chromosome Y

Deux catégories de marqueurs moléculaires ont été utilisés afin d'analyser la diversité génétique du chromosome Y dans la population marocaine :

- a) Les *marqueurs multi-allélique* ou *STR* (Short Tandem Repeat) renferment plus de 200 marqueurs STR au niveau du chromosome Y (Y-STR) (Buttler 2003, Kayser et *al.*, 2004). Les loci multi-alléliques peuvent être utilisés pour différencier les haplotypes du chromosome Y avec une grande résolution qui est due à leurs taux de mutation très élevés.
- b) Les *marqueurs bi-alléliques* tels que les polymorphismes d'un seul nucléotide (*Single Nucleotide Polymorphisms, SNP*) ou les petites insertions ou délétions (indels) représentent une autre classe importante des marqueurs du chromosome Y. On se réfère parfois à ces marqueurs en les nommant les polymorphismes d'événement unique car leur taux de mutation est beaucoup plus faible que celui des STR ( $\approx 10^{-8}$  contre  $\approx 10^{-3}$  mutations par génération) (de Knijff, 2000 ; Kayser et *al.*, 2000 ; Kayser et Sajantila, 2001). Cependant, la majorité des marqueurs bi-alléliques sont localisés dans la région non recombinante du chromosome Y et de ce fait n'a pas d'impact détectable sur le phénotype d'un individu. Ces SNP neutres sont très utiles en tant que marqueurs en génétique des populations (Nielsen, 2004 ; Marth et *al.*, 2004), pour les études de l'évolution moléculaire (Zhao et *al.*, 2003 ; Y chromosome consortium, 2002) et les analyses génétiques appliquées à la criminalistique (Gill, 2001 ; Gill et *al.*, 2004). Le premier marqueur bi-allélique du chromosome Y a été décrit en 1994 (Hammer, 1994).

### IV.1. Technologie du génotypage des STRs du chromosome Y

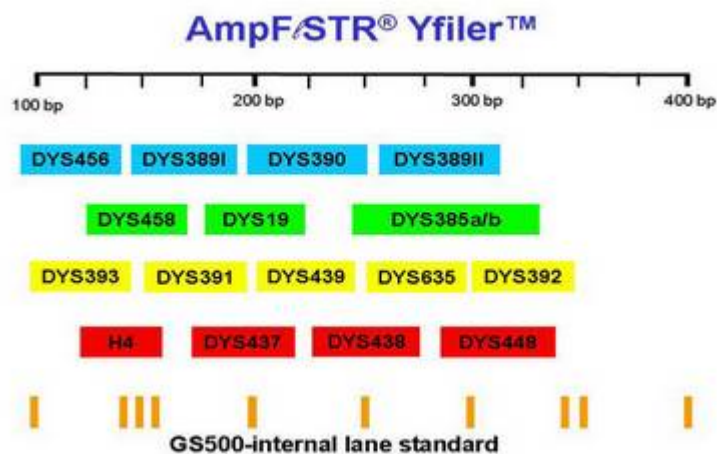
Grâce à l'utilisation de molécules fluorescentes, une méthode rapide, fiable et hautement précise a pu être développée afin de déterminer le profil génétique STR du chromosome Y d'un individu : la *PCR multiplex*. Basée sur le principe de base de la PCR. Les PCR multiplex sont réalisées à l'aide du kit AmpFSTR®Yfiler™ (Applied Biosystems®, Foster City, CA, USA). Le kit permet l'amplification simultanée (dans un seul tube d'analyse) de 17 loci STR du chromosome Y. Les loci STR amplifiés par le kit Yfiler™ peuvent être subdivisés en trois catégories :

- ◇ **Haplotype minimal Européen** (European minimal Haplotype) est déterminé par les loci suivants: DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393.
- ◇ **Scientific Working Group-DNA Analysis Methods** (SWGDM) recommande d'ajouter sur le panel Haplotype minimale Européen les loci suivants : DYS438 et DYS439.
- ◇ D'autant plus, d'autre marqueur hautement polymorphe ont été additionné par le fabricant (Applied Biosystems®, Foster City, CA, USA) qui sont: DYS437, DYS448, DYS456, DYS458, DYS635 (Y GATA C4) et Y GATA H4.

Les procédés chimiques de la PCR Yfiler™ utilisent quatre couleurs (de différentes longueurs d'ondes) pour marquer les amorces qui serviront de point de départ à l'amplification d'un Y-STR spécifique (Figure 14) :

- ◇ **6-FAM** (bleu) pour les loci DYS456, DYS389I, DYS390, et DYS389II
- ◇ **VIC** (vert) pour les loci DYS458, DYS19 et DYS385a/b
- ◇ **NED** (jaune) pour les loci DYS393, DYS391, DYS439, DYS635 et DYS392
- ◇ **PET** (rouge) pour les loci Y GATA H4, DYS437, DYS438 et DYS448

Une 5ème couleur (LIZ, orange) est utilisée lors de l'électrophorèse pour surligner le standard de taille.



**Figure 14:** Marqueurs de fluorescence utilisés dans le kit AmpFISTR®Yfiler™ (AppliedBiosystems, Foster City, CA, USA) et taille des loci amplifiés.

**Source :** <http://www.cstl.nist.gov/strbase/kits/Yfiler.htm>

Ces colorations sont ensuite détectées par un LASER (dans un Analyseur Génétique) lors de la séparation électrophorétique des produits de PCR. Les informations sont ensuite transmises à un ordinateur qui va les collecter et les traiter. Des logiciels spécifiques attribuent alors à chaque allèle

STR une taille (en paires de bases) à partir de laquelle le nombre de motifs répétés et le nom de cet allèle sont déduits. Un profil génétique est ainsi généré pour chaque individu analysé.

#### **IV.1.1. La PCR multiplex du kit AmpFISTR®Yfiler™ (Applied Biosystems®)**

La réaction de polymérisation en chaîne de 17 loci Y-STRs est effectuée en multiplex en utilisant le kit AmpFISTR®Yfiler™ (Applied Biosystems, Foster City, CA, USA). Pour chaque échantillon, 10 µl d'ADN (0.5 à 1 ng/µl) est amplifié dans un volume réactionnel final de 25 µl dont 9.2 µl du Mix de la réaction, 5.0 µl des amorces, 0.8 µl de l'ADN polymérase AmpliTaq Gold. En parallèle, deux «témoins» sont utilisés, l'un dit «positif» pour valider le bon déroulement de la réaction (ADN contrôle fourni avec le kit) et l'autre dit «négatif» pour s'assurer de la non contamination du milieu réactionnel (eau milliQ).

Le déroulement de la PCR Yfiler™ est fait dans un thermocycleur GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA, USA) en trois étapes: après une incubation initiale à 95°C pendant 11 minute, succèdent 30 cycles composés chacun de trois paliers de température (dénaturation : 1 min à 94 °C, hybridation 1 min à 61 °C, élongation 1 min à 72 °C). A la fin une extension finale est réalisée à 60°C pendant 80 minutes.

#### **IV.1.2. L'électrophorèse capillaire**

L'électrophorèse capillaire permet de détecter et quantifier les différents produits de PCR mais aussi de les séparer en fonction de leur taille et de leur fluorescence. Dans une électrophorèse capillaire, la séparation résulte de la migration des fragments d'ADN dans un champ électrique au sein d'un capillaire de quelques microns de diamètre et rempli de gel de polyacrylamide. Les molécules d'ADN sont séparées en fonction de leur taille : les plus petites traversent plus facilement le maillage du polymère et migrent donc plus rapidement au travers de ce réseau. La détection et la quantification de la fluorescence des fragments (sous forme de pics colorés) sont effectuées par un détecteur LASER placé en fin du parcours migratoire. Ainsi, les fragments les plus petits sont chronologiquement détectés en premier et si deux (ou plusieurs) fragments ont la même taille, ils sont distingués par la couleur de leur fluorescence. Ce procédé d'utilisation de colorations de différentes longueurs d'onde inclus dans la PCR multiplex permet ainsi de distinguer des fragments chevauchants (même taille).

Dans cette étude, les électrophorèses capillaires ont été effectuées dans un Analyseur Génétique à détection LASER de la fluorescence (modèle ABI Prism 3130xl, Applied Biosystems, Foster City, CA, USA). Les séparations ont été réalisées dans un capillaire de 47 cm de longueur (36 cm de la zone d'injection jusqu'au détecteur laser), de 50  $\mu\text{m}$  de diamètre interne et rempli de gel de polyacrylamide à 4% ou POP4. Les échantillons sont électro cinétiquement injectés dans le capillaire en 5 secondes. La migration électrophorétique de chaque échantillon dure 30 minutes, à 15 kV et à une température de 60°C.

Avant la séparation électrophorétique des fragments d'ADN, 1,5  $\mu\text{l}$  du produit PCR multiplex est mélangé avec 24.5  $\mu\text{l}$  du Hi-Di™ formamide et 0,5  $\mu\text{l}$  d'un standard de taille de fluorescence orange (GeneScan 500 LIZ, Applied Biosystems®, Foster City, CA, USA). Comme dans une électrophorèse classique, ce marqueur de taille «interne», préalablement calibré et dont la taille de chacun des fragments-standards est connue, doit migrer en même temps que l'échantillon afin de corrélérer l'ensemble des paramètres pouvant influencer la migration (vitesse, temps de détection par le LASER, pureté du gel, température...).

Dans chaque plaque d'échantillons soumis à l'électrophorèse capillaire, un autre marqueur de taille (« externe ») est ajouté : le ladder STR du chromosome Y. Il s'agit d'une « échelle allélique » constituée de brins d'ADN synthétique dont la composition et la taille (le nombre de répétitions) sont connus. Le ladder contient, en fait, tous les allèles STR du kit Yfiler™ pouvant être mis en évidence avec une PCR multiplex donnée et il y a une échelle allélique pour chaque fluorochrome utilisé. Comme pour un échantillon, le standard de taille interne est ajouté au ladder, les fragments d'ADN (allèles STR) sont séparés lors de l'électrophorèse et la fluorescence est détectée par le LASER. La Figure 15 présente le ladder des 4 couleurs utilisées dans la PCR multiplex Yfiler™ (Applied Biosystems®, Foster City, CA, USA).



Figure 15: Echelles alléliques (ladder) utilisées dans la PCR multiplex Yfiler™ (Applied Biosystems).

Les échelles alléliques définissent, pour chaque locus STR, tous les allèles pouvant être mis en évidence dans la technologie du kit Yfiler™ (Applied Biosystems). Chaque pic coloré est un allèle et correspond à une taille spécifique de fragment d'ADN et donc à un nombre précis de répétitions.

#### IV.1.3. L'obtention des profils Y-STRs

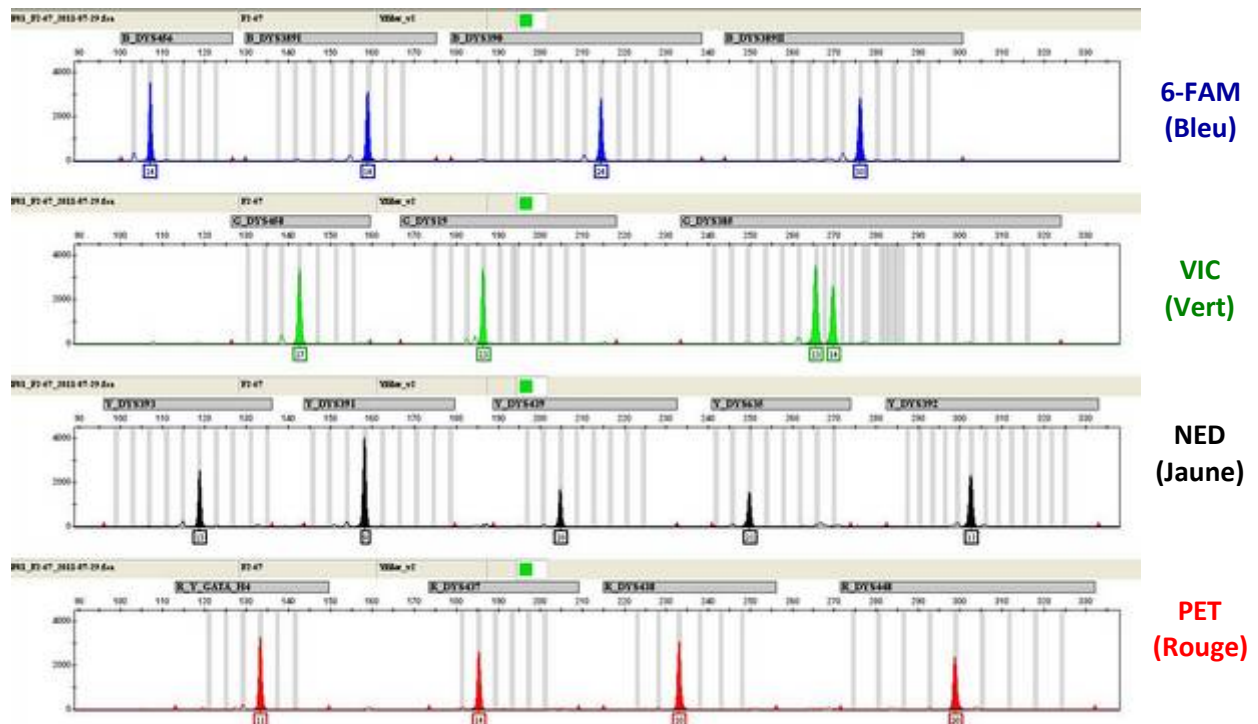
L'obtention du profil Y-STR d'un individu consiste à traiter informatiquement les informations collectées par le LASER lors de l'électrophorèse. La détection des produits de PCR d'un échantillon par le LASER est transmise à un ordinateur (couplé au séquenceur automatique) et les Y-STR sont représentés sous forme de pics colorés (allèles Y-STR). Le Traitement des données se fait en deux temps :

1. Assignment d'une taille (en paires de bases, pb) à chaque pic ;
2. Détermination du nom de l'allèle Y-STR.

En premier temps, le logiciel GeneMapper® ID (version 3.2, Applied Biosystems) trie les différents fragments selon leurs fluorescences et leurs longueurs. Cette longueur est proportionnelle au temps

de migration devant le LASER et se mesure en SCAN (Southern 1979). GeneScan détermine alors le *profil de séparation* du standard de taille interne (de couleur orange) en corrélant les longueurs (en SCAN) et les tailles (en pb) de chaque pic. Des profils de séparation sont ensuite tracés pour chaque échantillon (ainsi que pour le ladder) et, par comparaison avec le standard de taille, le logiciel estime la taille (en pb) de chaque fragment STR. L'analyseur génétique a un haut pouvoir de résolution puisqu'il permet de distinguer des fragments d'ADN variant de seulement 1 pb.

En deuxième temps, le logiciel commence par calibrer le standard de taille externe (ladder). A partir de la taille (en pb) de chaque allèle, le logiciel en déduit le nombre de motifs répétés et attribue cette valeur chiffrée (équivalente au nom de l'allèle) à chacun des pics du ladder. Les profils de séparation de chaque échantillon sont ensuite comparés aux échelles alléliques de référence et les noms des allèles STR sont déterminés pour tous les loci. On obtient ainsi le « profil électrophorétique » de chaque individu, c'est-à-dire son « profil Y-STR » (Figure 16).

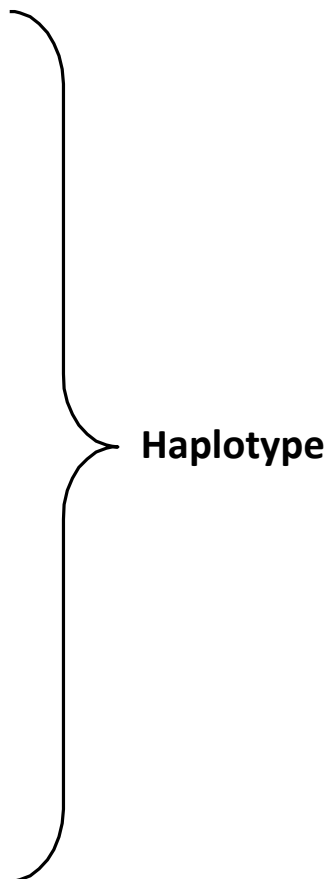


**Figure 16:** Exemple d'un profil Y-STR d'un individu obtenu après la PCR multiplex Yfiler™ (Applied Biosystems).

Le profil Y-STR d'un individu décrit les allèles (nombre de répétitions du motif nucléotidique) de 17 loci du chromosome Y. Chaque allèle est représenté sous forme d'un pic coloré correspondant à une taille de fragment d'ADN et donc à un nombre précis de répétitions. L'attribution des allèles pour les 17 loci du chromosome Y suit les lignes directives de la société Internationale de Génétique médico-légale (ISFG) (Gusmao et *al.*, 2006). Le groupe d'allèle obtenu après la lecture du profile représente un **haplotype** pour l'individu analysé. Un haplotype est une combinaison d'allèles de différents loci situés sur un même chromosome, étroitement liés (faible distance entre les loci) et transmis en bloc des parents aux enfants. Divers haplotypes peuvent être regroupés en haplogroupes (Tableau 1).

**Tableau 1 :** Résultat de la lecture du profil électrophorétique Y-STR d'un individu par ordre de taille de séparation.

Locus	Allèle
DYS456	14
DYS393	13
GATA H4	11
DYS458	17
DYS391	9
DYS389 I	14
DYS437	14
DYS19	13
DYS439	10
DYS390	24
DYS438	10
DYS635	21
DYS385 a/b	13 / 14
DYS389 II	30
DYS448	20
DYS392	11



Ce tableau récapitule les allèles présents à chacun des 17 loci STR du chromosome Y pour l'individu de la Figure 16.

#### IV.1.4. Les outils informatiques

Divers outils et méthodes statistiques sont employés pour décrire la structure génétique de notre population marocaine étudiée ( $n = 267$ ) subdivisée en trois groupes ethniques [Arabe ( $n = 130$ ), Berbère ( $n = 69$ ), Sahraoui ( $n = 68$ )]. Les premiers calculs effectués permettent de quantifier les 17 marqueurs STR du chromosome Y étudiés (fréquences alléliques, fréquences haplotypiques, diversité génétique,...). Ensuite notre population est comparée à d'autres populations et ordonnées selon des procédures de classification automatique ou de regroupement hiérarchique afin de retracer leurs affinités génétiques.

##### *i. La structure génétique*

L'étude des microsatellites du chromosome Y de notre population analysée a été réalisée sur la base des fréquences alléliques de chaque locus. Différentes analyses statistiques sont réalisées par le logiciel **Arlequin** v.3.11 (Excoffier et al., 2005) (URL : <http://cmpg.unibe.ch/software/arlequin3/>). L'analyse par Arlequin permet d'estimer :

- ☞ La **Diversité Génétique** (*Gene Diversity*, GD) : La diversité génétique définit par Nei en 1973 mesure comme l'indique son nom le degré du polymorphisme génétique dans une population ou le pouvoir de discrimination de chaque locus. Elle est calculée par la formule suivante où  $n$  est le nombre des échantillons étudiés et  $p_i$  est la fréquence allélique :

$$GD = (N/n - 1)(1 - \sum p_i^2)$$

- ☞ La **Diversité Haplotypique** (*Haplotype Diversity*, HD) : Elle présente la probabilité que deux haplotypes choisis au hasard soient différents. **DH** est calculée par la même formule que **GD** où  $p_i$  est la fréquence haplotypique au lieu de allélique.
- ☞ La **Capacité de Discrimination** (*Discrimination Capacity*, DC) : est par définition, la fréquence par laquelle deux individus non apparentés peuvent être discriminés. Elle est calculée en divisant le nombre des différents haplotypes par le nombre total des haplotypes.
- ☞ Le **Nombre des Haplotypes Uniques** (*No. Of Unique Haplotypes*, **NUH**) : NUH est le nombre des haplotypes observés une seule fois dans la population.

## ii. Les distances entre les populations

La comparaison des populations passe par la quantification de leurs différences (génétiques, géographiques, linguistiques, culturelles...), qui se traduit par le calcul de distances entre les groupes analysés. Ces calculs et l'établissement de matrices de distances sont à la base des analyses phénétiques réalisées en Phylogénie.

Par définition, une distance est une mesure de l'éloignement entre objets : plus les objets sont différents, plus la distance est grande ou, inversement, plus la similarité globale entre deux objets est grande, plus la distance qui les sépare est petite. Une distance entre deux objets **A** et **B** vérifie **trois conditions** mathématiques :

- ☞ La distance est positive si A est différent de B, mais nulle si A est égal à B (il s'agit lors de la distance d'un objet à lui-même) : *propriété de séparation*.
- ☞ La distance de A à B est égale à celle de B à A (la distance est indifférente du sens du parcours) : *propriété de symétrie*.
- ☞ La distance entre A et B est la plus directe (c'est-à-dire plus courte que celle qui passe par un troisième objet C) : inégalité triangulaire. Une distance vérifiant cette inégalité triangulaire est une distance métrique.

Les distances utilisées dans les méthodes de reconstruction phylogénétique doivent avoir ces propriétés des distances définies mathématiquement. Cependant, la distance génétique sera aussi définie par d'autres types de critères et devra donc répondre à un modèle génétique expliquant le processus ayant conduit à la différenciation de deux populations (par exemple la mutation, la mutation et la dérive, la sélection...). La distance de Nei, utilisée dans la méthode du Neighbor-Joining, est linéairement proportionnelle au temps de divergence entre deux populations et suppose que les vitesses d'évolution sont constantes d'un locus à l'autre et d'une génération à l'autre (Nei 1972 et 1978). Les distances **Fst** mesurent la divergence des populations due à la dérive et sont donc appropriées à des temps courts où les mutations sont négligeables (Reynolds 1983). La valeur d'une distance **Fst** varie entre **0** (pas de différenciation génétique entre les populations) et **1** (différenciation la plus extrême, isolation complète des populations). Les distances Fst dérivent des indices Fst. Ces derniers ne sont pas des distances à proprement parlé car ils ne respectent pas la

règle de l'inégalité triangulaire. Ils renseignent néanmoins sur les différenciations génétiques observables entre des populations et sur leur significativité. Les *indices Fst* sont calculés entre chaque paire de populations et leur significativité est testée par une procédure non-paramétrique « **test AMOVA** ».

L'**AMOVA** (*Analysis of MOlecular VAriance*) ou Analyse de la Variance Moléculaire permet d'étudier la diversité génétique des populations d'une même espèce selon différents critères: géographiques, culturels, linguistiques, etc. (Excoffier *et al.*, 1992). Elle est basée à la fois sur l'analyse de la variance des fréquences géniques et sur le nombre de mutations entre les haplotypes moléculaires. Sur la base d'une division hiérarchique des populations selon des critères non génétiques, l'approche estime les composants de la variance totale observée à trois niveaux : (i) dans une population, (ii) entre les populations et (iii) entre des groupes de populations.

Dans notre étude nous avons comparé notre population marocaine en utilisant l'outil AMOVA ( $\Phi$ st statistics) disponible sur la base de donnée **YHRD** (URL : [www.yhrd.org](http://www.yhrd.org)) avec neuf populations : **Algérie** (Robino *et al.*, 2008), **Tunisie** (Onofri *et al.*, 2008), **Nord d'Egypte** (Omran *et al.*, 2008), **Iran** (Roewer *et al.*, 2009), **Grèce** (Kovatski *et al.*, 2009), **Italie** (Ferri *et al.*, 2009), **Portugal** (Carvalho *et al.*, 2003), **Namibie** (Fujihara *et al.*, 2009), **Chine** (Zhu *et al.*, 2008)]. Les résultats issus du test AMOVA permet de visualiser en deux graduations multidimensionnelles la distance génétique entre les populations analysées.

#### IV.1.5. Contrôle de qualité par YHRD

La publication des résultats du typage des STR du chromosome Y de la population marocaine étudiées (n = 267) à la base de donnée de référence YHRD ([www.yhrd.org](http://www.yhrd.org)) a été précédé par la réalisation d'un test de contrôle qualité dans le but d'évaluer la nomenclature correcte des allèles selon les recommandations de l'ISFG « International Society of Forensic Genetics » pour l'analyse des STR du chromosome Y (Gusmao *et al.*, 2006). Cinq échantillons d'ADN inconnus reçus de la base de données YHRD ont été analysés pour les 17 locus en utilisant le kit Yfiler™ (Applied Biosystems®, Foster City, CA, USA). Suite à l'évaluation des résultats et la certification (annexe B), des numéros d'accessions ont été attribués pour les trois groupes ethniques: Rabat, Morocco [Arab] YA003437, Rabat, Morocco [Berber] YA003438, et Rabat, Morocco [Sahrawi] YA003439. Les données du polymorphisme des 17 marqueurs STR du chromosome Y des populations citées précédemment

peuvent être retrouvés sur la base de données YHRD (URL : [www.yhrd.org](http://www.yhrd.org)) soit par leurs numéros d'accès ou bien par le nom de la population.

## IV.2. Technologie du génotypage des marqueurs bi-alléliques du chromosome Y

Les méthodes de typage des marqueurs bi-alléliques ou bien SNP (*Single Nucleotide Polymorphisms*) font intervenir deux étapes :

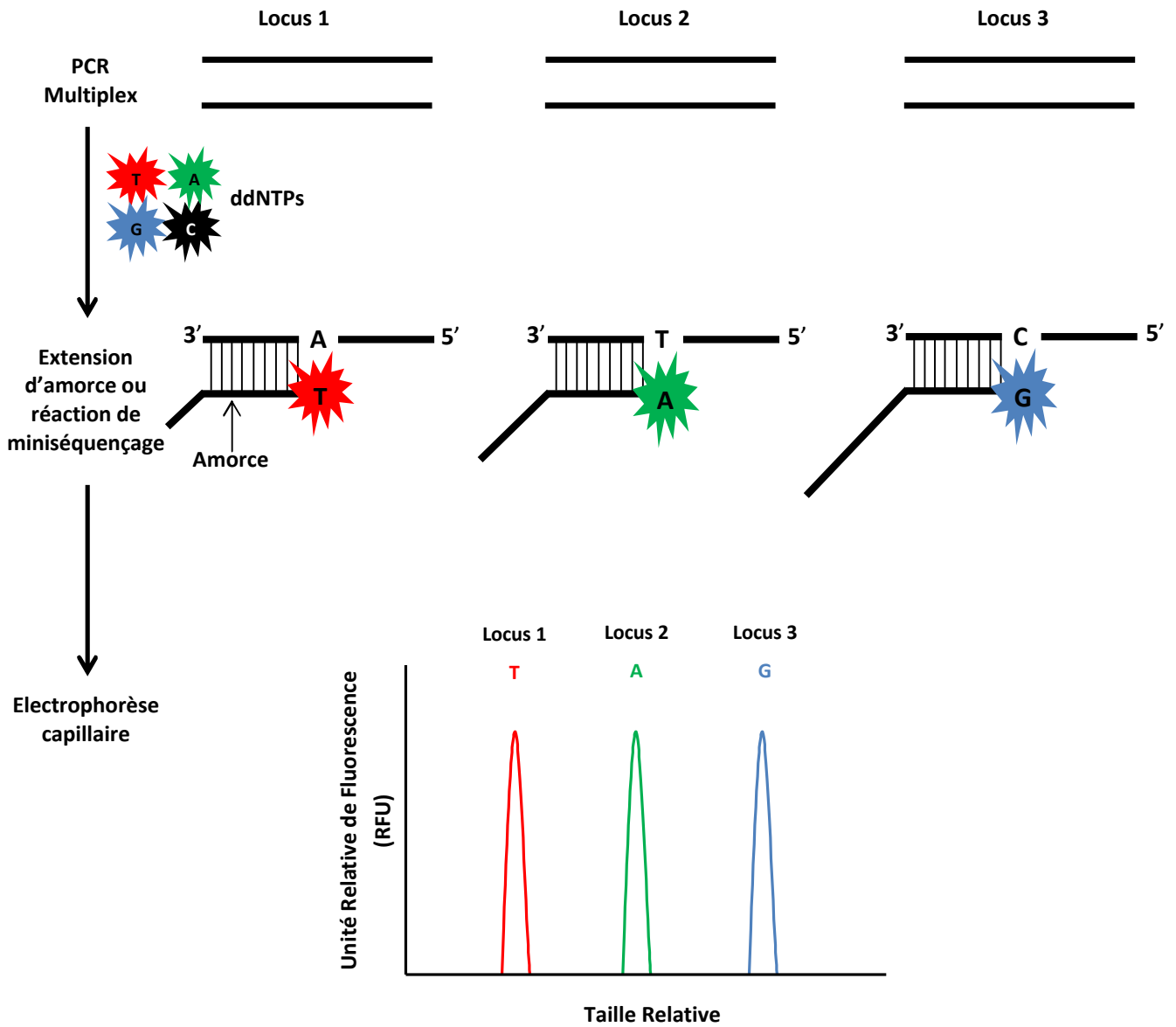
1. Une étape d'amplification par PCR de la région contenant le SNP;
2. Une étape d'analyse post-PCR pour déterminer le variant allélique au site polymorphe considéré.

Les analyses **post-PCR** sont fondées soit sur l'hybridation d'une sonde sur le produit amplifié, soit sur une ligation des amorces, soit sur un clivage de sonde, soit encore sur une extension des amorces. Il existe également différents systèmes de détection des produits de la discrimination allélique parmi lesquels les puces à ADN, l'électrophorèse, la PCR en temps réel, le pyroséquençage et la spectrométrie de masse. Par la présence d'un instrument d'électrophorèse capillaire ou bien l'analyseur génétique 3130 (Applied Biosystems, Foster City, CA, USA) au laboratoire des sciences médico-légales à l'université George Washington (Washington, DC, US), la combinaison technologique employée est une **extension d'amorce**, qui répond aux besoins qualitatifs d'une analyse d'identification génétique de notre population marocaine étudiée (**n = 159**), subdivisée en trois groupes ethniques [Arabe (**n = 42**), Berbère (**n = 67**), Sahraoui (**n = 50**)].

Cette étape, également appelée **miniséquencage**, est précédée par une étape d'amplification de la région d'ADN contenant le SNP d'intérêt au moyen d'un couple d'amorces. Les produits PCR obtenus sont purifiés puis soumis à la réaction du **miniséquencage** ou **d'extension d'amorce**. Cette réaction repose sur l'hybridation de l'amorce à une base en amont du polymorphisme et son extension de manière complémentaire à l'allèle présent sur l'ADN cible. Cette extension se fait grâce à une ADN polymérase qui incorpore de manière indifférente les didéoxyNucléotide TriPhosphate (ddNTPs) (Figure 17).

Nous avons choisi pour cette étude de discriminer les allèles par extension d'amorce en utilisant la technologie du kit ABI Prism® **SNaPshot™** Multiplex (Applied Biosystems, Foster city, CA, US)

constituée de l'AmpliTaq® DNA Polymerase et d'un mélange de ddNTP marquée par des fluorochromes (Tableau 2).



**Figure 17:** Principe du génotypage par SNaPshot™ Multiplex par extension allèle spécifique d'amorce, selon le SNP.

**Tableau 2:** les Fluorochromes des ddNTP du kit ABI Prism® SNaShot™

ddNTP	Fluorochrome	Couleur des données analysées
A	dR6G	Vert
C	dTAMRA™	Noir
G	dR110	Bleu
T (U)	dROX™	Rouge

#### IV.2.1. Sélection des marqueurs bi-allélique du chromosome Y

Dans l'objectif de déterminer l'haplogroupe le plus fréquent dans la population Marocaine, nous avons ciblé dans notre étude 22 marqueurs bi-alléliques du chromosome Y subdivisés en quatre multiplex :

- ☞ **MY01:** A-M91 (A), B-M60 (B), C-RPS4Y711 (C), D-M174 (D), J-M267 (J1), I-M170 (I), R-M207 (R);
- ☞ **MY02:** E-P147 (E1), E-M132 (E1a), E-P189 (E1b1a), E-M215 (E1b1b), E-M81 (E1b1b1b), E-M54 (E2b), J-M172 (J2);
- ☞ **MY03:** J-L24 (J2a4h), J-M221 (J2b), R-L63 (R1a), R-M343 (R1b), R-M269 (R1b1b2), T-M70 (T) ;
- ☞ **MY04:** E-M35 (E1b1b1), E-M78 (E1b1b1a).

Ces derniers, ont été sélectionnés selon une stratégie hiérarchique basée sur l'arbre phylogénétique reconnue par le Consortium du Chromosome Y (CCY) (Y Chromosome Consortium, 2002). La nomenclature des loci suit les recommandations de Karafet et collaborateurs (2002). Les séquences de références des loci choisis ont été prises depuis : International Society of Genetic Genealogy, Y-DNA Haplogroup Tree 2010 (<http://isogg.org/tree/>).

#### IV.2.2. La conception des amorces

##### *i. Pour la PCR*

La mise au point d'une réaction PCR spécifique et efficace dépend fortement du choix des amorces. La conception des amorces a été faite avec l'utilitaire « Integrated DNA Technologies » (URL: <http://eu.idtdna.com/Scitools/Applications/Primerquest/>). La température d'hybridation a été fixée à 55°C pour tous les couples d'amorces afin de développer des multiplex, le pourcentage de GC varie entre 45 et 50%, la taille des fragments à amplifier a été limitée à 295 pb (annexe C).

Les dimères et les structures en épingle à cheveux des amorces ont été vérifiés en utilisant l'utilitaire AutoDimer (Vallone et *al.*, 2004).

Dans un premier temps, les fragments contenant les SNP d'intérêt ont été amplifiés individuellement (PCR single) dans le but de valider la spécificité de l'amplification. Les réactions d'amplifications ont été développées dans des conditions identiques pour tous les fragments d'intérêt. Les produits d'amplification ont été observés sous lumière UV après électrophorèse sur gel à 2% d'agarose standard contenant du BET, dans du tampon TBE 1X (89 mM Tris base, 89 mM acide borique, 2 mM EDTA, pH 8,3).

#### **ii. Pour l'extension d'amorce (Primer Extension, PEX)**

Le choix des amorces pour l'extension allèle spécifique d'amorce est limité par rapport à celui des amorces PCR. En effet, une amorce PEX doit s'hybrider immédiatement en amont du polymorphisme. La longueur de l'amorce étant limitée par la température de fusion imposée, le choix entre les orientations sens et anti-sens sera guidé par la "qualité" de la séquence uniquement.

La conception des amorces PEX a été réalisée par le logiciel Primer Express™ v. 2.0 (Applied Biosystems, Foster city, CA, US). Une queue homopolymérique en Thymine a été ajoutée d'une manière croissante afin d'éviter le chevauchement entre les différents produits de réaction de miniséquençage au cours de l'électrophorèse en multiplex (annexe D).

Dans un premier temps, les SNP d'intérêt ont été analysés individuellement (singlePEX) dans le but de valider la spécificité d'extension. Les PEX simultanées (multiPEX) n'ont été optimisées qu'après validation de la faisabilité technique de l'analyse des SNP et de leur informativité.

#### **IV.2.3. L'amplification PCR multiplex et purification**

Pour Chaque échantillon, 2 µl d'ADN génomique (3 à 5 ng/µl) ont été amplifiées par les 4 multiplex MY01 (*Heptaplex*), MY02 (*Heptaplex*), MY03 (*Hexaplex*) et MY04 (*Duplex*) cités précédemment. Le mix réactionnel du volume final de 25 µl est composé par 1X GeneAmp® PCR Gold (Applied Biosystems, Foster city, CA, US), 400 µM de dNTPs, 3 mM du MgCl<sub>2</sub>, 0.65 mg/µl de L'albumine de sérum bovin ou (ASB) et 1.5 Unité de l'AmpliTaq Gold® DNA polymérase (Applied Biosystems). La concentration finale des amorces est comprise entre 0.08 et 0.8 pmol. Les

amplifications ont été conduites sur un thermocycleur GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA, USA). Après une activation pendant 10 min à 95°C, succèdent 35 cycles d'amplification composé de trois paliers de température (Dénaturation : 30 sec à 94°C, Hybridation : 30 sec à 55°C, élongation : 30 sec à 72°C). À la fin du dernier cycle, une élongation finale pendant 10 min à 72°C.

Après amplification, un volume de 5 µl de produits PCR a été purifié des amorces non fixées et les dNTPs non incorporées par l'ajout d'une unité de l'Exonucléase I (USB Corporation) et une unité de Shrimp Alkaline Phosphatase (SAP). L'incubation des produits PCR a été conduite sur un thermocycleur GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA, USA) selon le programme suivant : 70 min à 37°C et 20 min à 65°C.

#### **IV.2.4. La réaction du miniséquençage multiPEX et détection de la fluorescence**

La réaction du miniséquençage a été conduite dans un volume réactionnel de 15 µl contenant 1.5 µl d'ABI Prism® SNaPshot™ multiplex kit (Applied Biosystems), 2 µl de produit PCR purifiée et 1.5 pmol de chaque amorces PEX. Les réactions du miniséquençage multiPEX ont été conduites sur un thermocycleur GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA, USA) avec un programme comprenant 25 cycles de Dénaturation (10 sec à 95°C), Hybridation (5 sec à 50°C) et une élongation (30 sec à 60°C).

Les électrophorèses capillaires des produits non purifiés de la réaction du miniséquençage multiPEX ont été réalisées dans un analyseur génétique (ABI Prism® 3130, Applied Biosystems, Foster City, CA, USA). Les séparations ont été effectuées dans un capillaire de 47 cm et rempli de gel de polyacrylamide POP7.

Après électrophorèse, les résultats ont été analysés en utilisant le logiciel GeneMapper v 4.0 (Applied Biosystems, Foster City, CA, USA).

## V. Mise en évidence du polymorphisme de l'ADN Mitochondrial

### V.1. Le séquençage de l'ADN mitochondrial

Les premiers séquençages de l'ADN mitochondrial (ADNmt) humain ont été réalisés à partir des années 1990 et concernaient uniquement la partie non codante du génome. La première étude des segments hypervariables (HVI et HVII) de la région contrôle d'individus africains (121 sub-Sahariens) a été faite en 1991 (Vigilant *et al.*). Les auteurs ont alors montré que la région non codante de l'ADNmt a une diversité 3 à 4 fois plus élevée que celle de la région codante. L'analyse de populations amérindiennes, effectuée en 1993, a défini la nomenclature à utiliser pour désigner les haplogroupes mitochondriaux (Torroni *et al.*, 1993). A partir de là, les haplogroupes des autres régions du Monde ainsi que la phylogénie des divers marqueurs ont été décrits (Torroni *et al.*, 1994b et c ; Richards *et al.*, 1998). Aujourd'hui, le séquençage de l'ADN mitochondrial concerne l'ensemble des fragments hypervariables de la D-loop. Depuis peu, le séquençage complet de la molécule d'ADNmt (régions codante et non codante) se développe, augmentant ainsi la précision des définitions des haplogroupes mitochondriaux (Torroni *et al.*, 2006).

### V.2. Stratégie d'analyse de la région contrôle de l'ADNmt de notre population étudié

L'amplification et le séquençage de la région contrôle (*D-loop*) de l'ADN mitochondrial de notre échantillonnage (**n = 509**) suit la stratégie du laboratoire d'identification génétique des forces armée **AFDIL (Armed Forces DNA Laboratory, Rockville, Maryland, USA)** selon la procédure décrite par Brandstätter *et al.*, (2004) avec des modifications mineures au niveau des amorces pour le séquençage (Irwin *et al.*, 2007).

#### V.2.1. Amplification et séquençage de la région contrôle de l'ADNmt

Dans notre étude, l'amplification de la région contrôle de l'ADN mitochondrial a été subdivisé en 3 PCR (Figure 18). Les séquences des amorces pour la PCR et les séquences des amorces pour le séquençage sont illustrées dans le tableau 3.

**Tableau 3:** Les séquences des amorces utilisées pour l'amplification et le séquençage de la région contrôlée de l'ADNmt.

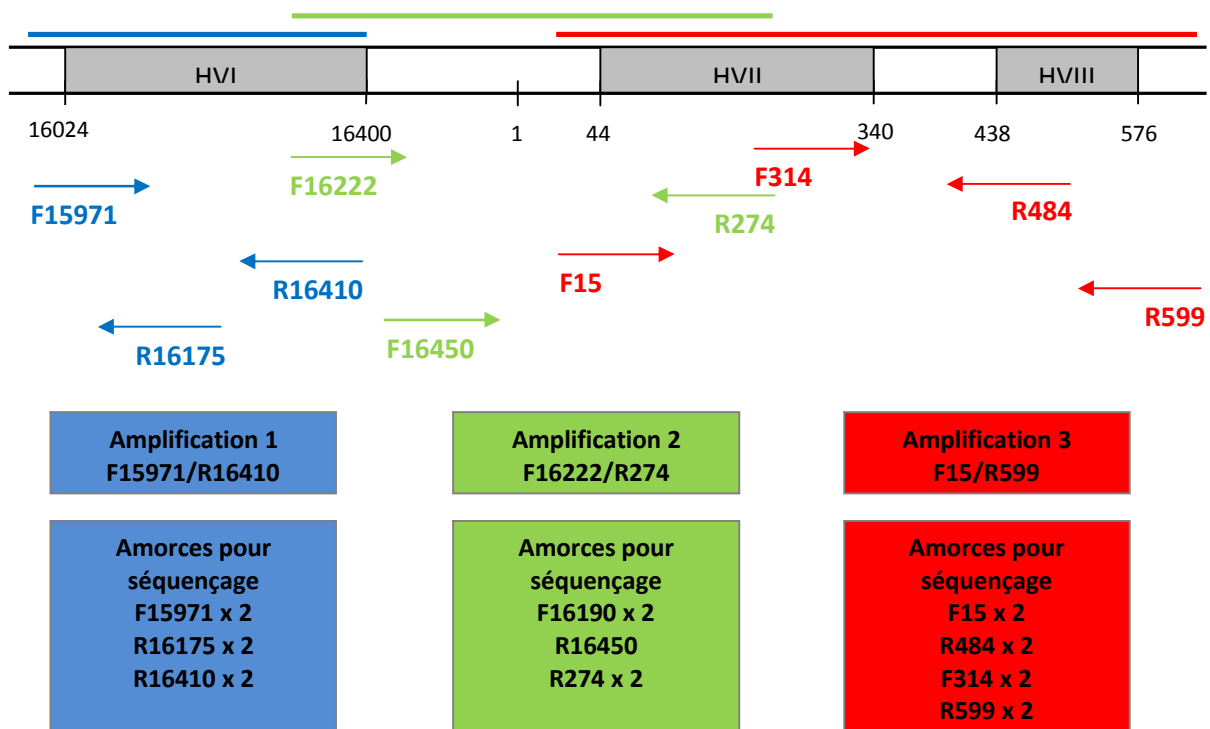
Région d'amplification	Position et séquences des couples d'amorces PCR	Position et séquences des amorces pour le séquençage
Amp1	<b>F15971</b> : 5'-TTA ACT CCA CCA TTA GCA CC-3'	<b>F15971</b> : 5'-TTA ACT CCA CCA TTA GCA CC-3'
	<b>R16410</b> : 5'-GAG GAT GGT GGT CAA GGG A-3'	<b>R16175</b> :5'-TGG ATT GGG TTT TTA TGT A-3'
		<b>R16410</b> : 5'-GAG GAT GGT GGT CAA GGG A-3'
Amp2	<b>F16222</b> : 5'-CCT CAA CTA TCA CAC ATC-3'	<b>F16222</b> : 5'-CCT CAA CTA TCA CAC ATC-3'
	<b>R274</b> : 5'-TGT GTG GAA AGT GGC TGT GC-3'	<b>R16450</b> :5'-GCT CCG GGC CCA TAA CAC TTG-3'
		<b>R274</b> : 5'-TGT GTG GAA AGT GGC TGT GC-3'
Amp3	<b>F15</b> : 5'-CAC CCT ATT AAC CAC TCA CG-3'	<b>F15</b> : 5'-CAC CCT ATT AAC CAC TCA CG-3'
		<b>R484</b> :5'-TGA GAT TAG TAG TAT GGG AG-3'
	<b>R599</b> : 5'-TTG AGG AGG TAA GCT ACA TA-3'	<b>F314</b> : 5'-CCG CTT CTG GCC ACA GCA CT-3'
		<b>R599</b> : 5'-TTG AGG AGG TAA GCT ACA TA-3'

Pour chaque échantillon, 3 µl d'ADN (5 à 10 ng/µl) a été amplifié par les trois couples d'amorces PCR dans un volume final de 50 µl. Le mélange réactionnel comprends 5 µl du tampon 10X GeneAmp® PCR Gold (Applied Biosystems), 4 µl dNTP (10 mM), 2 µl de chaque amorce (200 nM) et 0,5 µl d'AmpliTaq Gold® DNA polymérase.

Les amplifications sont conduites sur un thermocycleur GeneAmp™ PCR System 9700 (Applied Biosystems®, Foster City, CA, USA). Les conditions de la PCR sont identiques pour tous les paires d'amorces, avec une activation initiale de la Taq Gold polymérase à 96°C pendant 10 min et suivie de 36 cycles d'amplification (dénaturation : 30 sec à 94°C, hybridation 30 sec à 56°C, élongation 1 min à 72°C), puis d'une élongation finale (7 min à 72°C). Après la PCR, la présence du produit PCR a été confirmée sur un gel d'agarose 1 %. Les produits de PCR sont par la suite été purifiés en ajoutant dans chaque tube 1,5 µl d'ExoSAP-IT (USB Corporation) et 18,5 µl du tampon de dilution SAP (shrimp alkaline phosphatase). La purification a été conduite dans un thermocycleur GeneAmp™ PCR System 9700 (Applied Biosystems®, Foster City, CA, USA) selon le programme suivant : une étape d'activation de l'enzyme 80 °C pendant 15 minutes qui est suivie par une incubation à 37 °C pendant 20 minutes.

Dans un volume final de 20 µl, les réactions de séquences ont été réalisées en double à l'exception de l'amorce *R16450* pour l'amplicon 2. Le milieu réactionnel contient 1,5 µl du BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems®, Foster City, CA, USA), 6 µl du tampon de

séquençage (Applied Biosystems®), 0.5 µl du dGTP, 2 µl soit 10 µM d’amorce et 2 µl du produit PCR purifié. La réaction de séquence est conduite sur un thermocycleur GeneAmp™ PCR System 9700 (Applied Biosystems®, Foster City, CA, USA) selon les conditions suivantes: une dénaturation initiale a 96°C pendant une minute suivie de 25 cycles (dénaturation : 10 sec a 95°C, hybridation 5 sec a 50°C, élongation 4 min a 60°C). La réaction de séquences est suivie par une purification à l’aide des colonnes AGTC (Edge Biosystems, Gaithersburg, MD, USA) selon le protocole décrit par le fabricant. Après la purification, l’électrophorèse capillaire est réalisée sur l’analyseur génétique 3730xl (Applied Biosystems®, Foster City, CA, USA).



**Figure 18:** Stratégie d’amplification et du séquençage de la région contrôle de l’ADN mitochondrial Humain suivi à AFDIL.

### V.2.2. Analyse des données et contrôle qualité

Dans la mesure d'assurer une très grande qualité de donnée, pour chaque échantillon une double analyse a été effectuée. Les séquences consensus générées sont alignées et comparées à la séquence de référence rCRS par deux scientifiques indépendants à l'aide du programme Sequencher v4.7 (GeneCodes, Ann Arbor, MI). S'il n'y a pas de divergence entre les deux résultats et la validation de la nomenclature des différences par rapport à la séquence de référence rCRS, l'haplotype ainsi est déterminé. La séquence consensus finale est exportée du programme Sequencher ensuite importer vers la base de données interne du laboratoire.

Une vérification finale a été effectuée par l'envoi des séquences brutes de chaque individu ainsi les fichiers export de la séquence consensus généré par Sequencher à l'institut de médecine légale d'Innsbruck (Autriche), où ils seront réalignés une autre fois par deux scientifiques indépendants. Dans le cas d'absence de divergence, les données ont été éditées et incorporés dans la base de données EMPOP (<http://www.empop.org>). Les haplotypes sont enregistrés sous les numéros d'accessions suivants : *Moroccan Arab* EMP00280 (n = 42), *Moroccan Berber* EMP00281 (n = 69), *Moroccan Sahrawi* EMP00282 (n = 53) et *Moroccan Mixed* EMP00283 (n = 345). Les séquences consensus finaux de notre échantillonnage peuvent aussi être aussi trouvées sur GenBank sous les numéros d'accessions suivants : JF749296 - JF749804.

Après la confirmation des séquences par EMPOP, l'étape suivante consiste de classifier chaque haplotype de notre échantillonnage en haplogroupe et sous haplogroupes en fonction des mutations caractéristiques par l'utilisation de l'arbre phylogénique de l'ADN mitochondriale ([www.phylotree.org](http://www.phylotree.org)), (Van oven et *al.*, 2009) (annexe E).

### V.2.3. L'analyse statistique des données

Différents paramètres statistiques tel que la diversité génétique, La probabilité de rencontrer le même profil génétique chez deux individus inapparentés (Random Match Probability) et la comparaison par paire entre les individus (Pairwise) ont été calculé par le programme LISA du laboratoire AFDIL (Laboratory Information Systems Applications, Future Technologies Inc., Fairfax, VA) pour l'analyse de **509** haplotypes de notre population étudié.

## **Résultats et Discussion**

## I. Analyse du polymorphisme des marqueurs du chromosome Y

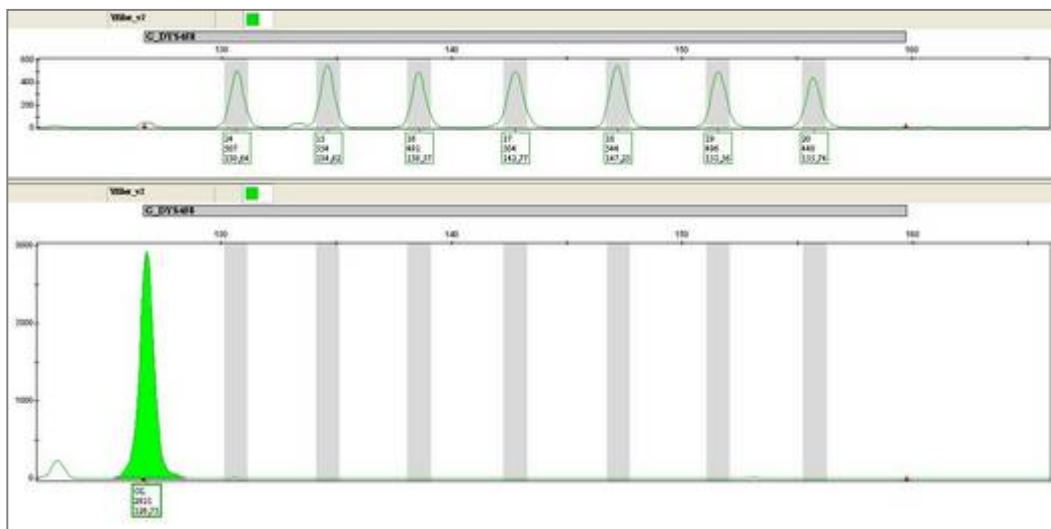
### I.1. Les marqueurs STR

Selon les résultats issus de l'analyse du polymorphisme des 17 loci STR du chromosome Y (DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (Y GATA C4) et Y GATA H4) en utilisant le kit AmpF/STR®Yfiler™ (Applied Biosystems, Foster city, CA, USA), le **tableau 4** montre que sur un totale de 267 individus masculins marocain non apparentés et appartenant au trois groupes ethnique [*Arabophones* (n = 130), *Berbérophones* (n = 69), *Sahraouis* (n = 68)], deux cent cinquante-sept individus, soit 96,25% de notre population, présentent différents haplotypes, dans lesquels 237 sont des haplotypes uniques (observés une seule fois dans la population étudiée) et 10 haplotypes dont chacun observé chez deux individus. Par ailleurs, quatre individus appartenant au groupe ethnique *Arabe* de notre population présentent des duplications au niveau du chromosome Y pour les loci **DYS448** (n = 2), **DYS456** (n = 1) et **DYS635** (n = 1).

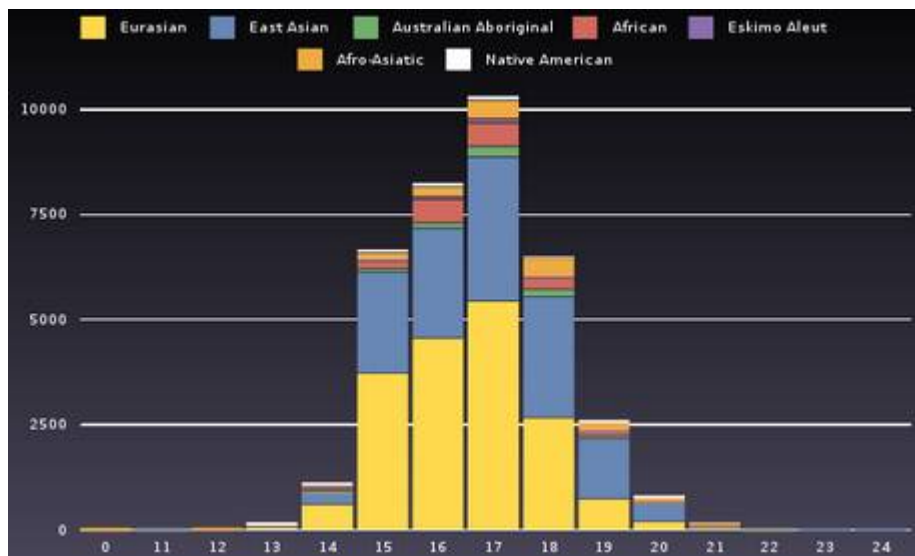
L'amplification des fragments ADN renfermant des régions répétées ou STR peuvent présenter un phénomène rare connu sous le nom **allèle nul**. Ce type de phénomène peut avoir lieu dans le cas d'une mutation au niveau de la région 3' (drop out) du site d'hybridation des amorces sens. Notre étude a montré que sur les 267 individus analysés, 19 entre eux présentent des allèles nuls, respectivement 17 au niveau du locus **DYS456** [*Arabophones*, n = 12 ; *Berbérophones*, n = 5] et chez deux individus *arabophones* pour le locus **DYS458**. Afin de confirmer la présence de ces allèles nuls, une deuxième amplification a été réalisée par le kit AmpF/STR®Yfiler™ (Applied Biosystems, Foster city, CA, USA). Nous étions limités d'utiliser le kit PowerPlex® Y (Promega corporation) dont le panel est de 11 marqueurs (DYS391, DYS389 I, DYS439, DYS389 II, DYS438, DYS437, DYS19, DYS392, DYS393, DYS390 et DYS385 a/b), en raison de l'absence des marqueurs génétiques **DYS456** et **DYS458**.

Un autre cas a été aussi observé dans notre population étudiée, où on rencontre des allèles qui ne tombent pas dans les 0,5 pb du lieu correspondant à l'échelle allélique d'un locus donné. Ce type de variation allélique est désigné par les *allèles variants* du terme anglais *off-ladder* ou *microvariant*. Deux types de cas d'allèles variants ont été rencontrés à l'échelle de notre population étudiée. D'une part un allèle variant qui est plus petit par rapport à l'échelle allélique du locus **DYS458**, et ceci a été

observé chez un seul individu appartenant au groupe ethnique Sahraoui qui présente l'allèle 13 avec une fréquence de 0,004 (Figure 19). D'après la base de données européenne YHRD ([www.yhrd.org](http://www.yhrd.org)), la distribution de l'allèle 13 pour le locus **DYS458** est très rare chez 7 métapopulations (Figure 20). En outre, sur la base de données américaine pour le chromosome Y « USYSTR » ([www.usystrdatabase.com](http://www.usystrdatabase.com)), l'allèle 13 est trouvé chez seulement 26 individus sur un total de 18193 individus (données Mai 2013) (Figure 21).

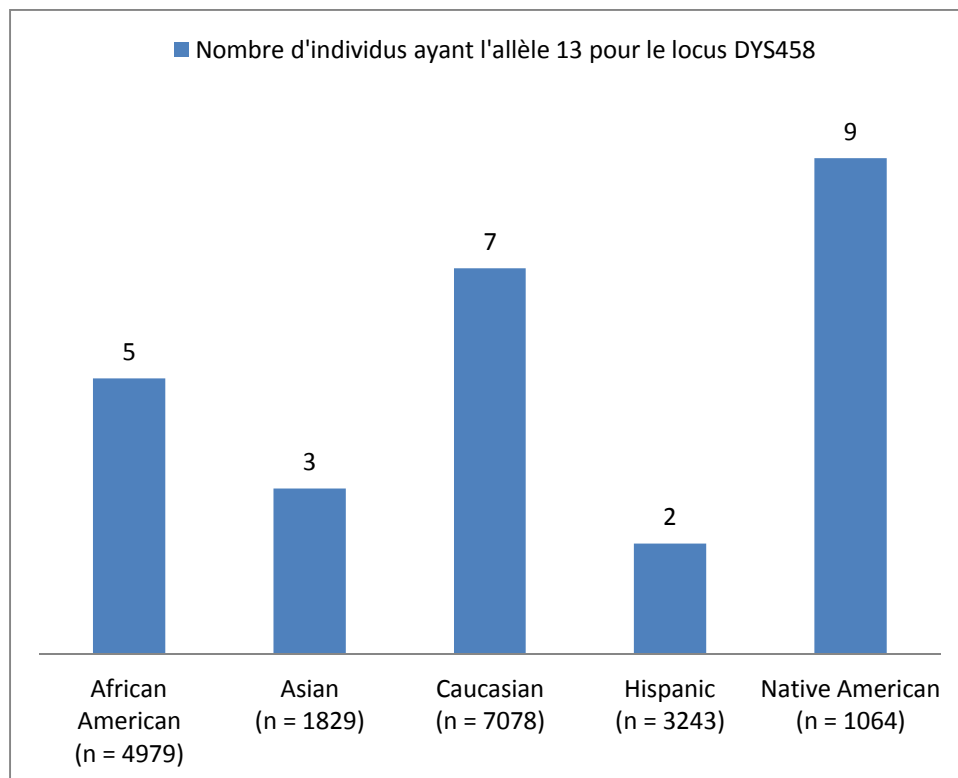


**Figure 19:** Illustration du profil Y-STR pour le locus DYS458 issu du logiciel GeneMapper® v 3.2 de l'individu qui présente l'allèle variant (allèle 13).



**Figure 20:** Distribution des fréquences allélique pour le locus DYS458 des métapopulations suivantes : Eurasienne, Asie de l'EST, Aborigène d'Australie, Africain, Eskimo-Aleut, Afro-Asiatique et les américains autochtones.

**Source :** [www.yhrd.org](http://www.yhrd.org)



**Figure 21:** Illustration graphique des individus ayant l'allèle 13 pour le locus DYS458 dans 5 métapopulations : Africain Américain, Asiatique, Caucasien, Hispanique et Autochtone Américain.

**Source :** [www.usystrdatabase.org](http://www.usystrdatabase.org).

D'autre part et pour le même locus (**DYS458**), un allèle tombe entre les barreaux de l'échelle allélique. Au niveau de notre population étudiée, ce type d'allèle variant ou *microvariant* a été rencontré chez 14 individus qui présentent l'allèle 18.2 [*Arabophones*, n = 8 ; *Sahraouis*, n = 6] et un individu appartenant au groupe ethnique *Sahraoui* qui présente l'allèle 19.2 (Tableau 4). Dans ce cas particulier, le microvariant possède une répétition partielle par rapport à un allèle normale pour le locus DYS458 avec un ordre de répétition complète, dont la structure est  $[GAAA]_n$  qui tombe directement sur l'un des barreaux de l'intervalle de l'échelle allélique. A cet égard, la fréquence des deux *microvariant* 18.2 et 19.2 observés dans notre population est petite par rapport à ce qui été observé chez une population situé au sud de l'Égypte (Omran et al., 2008).

**Tableau 4:** Fréquences des haplotypes pour 17 Loci Y-STR de notre population étudiée (n = 267)

Hp	Fréq	GE	DYS 19	DYS 385a/b	DYS 389 I	DYS 389 II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 437	DYS 438	DYS 439	DYS 448	DYS 456	DYS 458	DYS 635	Y GATA H4
Hp1	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	16	19	21	12
Hp2	1	Ar	14	13,18	13	30	23	11	11	12	14	10	10	20	14	18.2	20	11
Hp3	1	Ar	13	13,14	13	29	24	9	11	13	14	10	10	20	15	19	21	12
Hp4	1	Ar	14	13,19	13	29	23	11	11	12	14	10	11	20	14	18.2	21	11
Hp5	1	Ar	14	12,12	13	30	22	8	11	13	14	10	13	22	15	16	21	12
Hp6	1	Ar	14	12,14	13	29	24	9	11	13	14	10	10	20	16	18	21	12
Hp7	1	Ar	14	13,14	14	30	23	9	11	13	14	10	10	20	16	18	21	13
Hp8	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	14	18	21	11
Hp9	1	Ar	18	15,16	13	30	21	13	11	17	14	11	10	21	15	17	21	12
Hp10	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	15	17	21	12
Hp11	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	16	17	21	12
Hp12	1	Ar	14	13,18	13	31	23	11	11	12	14	10	12	20	14	18.2	21	11
Hp13	1	Ar	14	13,15	14	29	24	9	11	13	14	10	10	20	15	18	21	12
Hp14	1	Ar	13	13,14	14	30	24	9	11	12	14	10	11	20	16	19	22	12
Hp15	1	Ar	13	13,14	14	31	23	9	11	13	14	10	10	20	15	19	22	12
Hp16	1	Ar	15	11,12	13	30	21	10	11	13	14	10	12	21	14	16	20	11
Hp17	1	Ar	15	11,12	13	29	21	9	11	13	14	11	12	21	14	16	20	11
Hp18	1	Ar	13	13,14	14	30	24	9	11	13	14	10	11	20	16	18	22	12
Hp19	1	Ar	14	13,19	14	31	23	11	11	12	14	10	12	20	14	18.2	21	11
Hp20	1	Ar	13	13,14	15	31	24	9	11	13	14	10	9	20	15	18	21	12
Hp21	2	Ar\Sw	13	13,14	14	30	24	9	11	13	14	10	10	20	15	18	21	12
Hp22	1	Ar	14	12,14	13	29	24	10	13	13	15	12	12	19	17	17	23	12
Hp23	1	Ar	17	14,15	13	29	23	11	13	13	14	12	13	19	15	16	23	11
Hp24	1	Ar	14	13,17	13	30	23	11	11	12	14	10	11	20	14	18.2	20	12
Hp25	1	Ar	15	13,14	14	30	24	9	11	13	14	10	10	20	15	17	20	12
Hp26	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	21	16	19	22	12
Hp27	1	Ar	15	14,15	12	29	21	10	11	13	15	10	11	22	15	17	21	12
Hp28	1	Ar	16	13,14	12	29	21	10	11	15	16	10	11	22	15	18	22	11
Hp29	1	Ar	13	17,18	12	29	22	10	11	13	14	10	11	20	15	14	21	11
Hp30	1	Ar	13	13,14	13	29	24	9	11	13	14	10	10	20	15	20	19	12
Hp31	1	Ar	13	13,15	14	30	24	9	11	13	14	10	11	20	15	18	22	12
Hp32	1	Ar	13	13,14	13	30	24	9	11	13	14	10	11	20	16	18	22	12
Hp33	1	Ar	14	13,15	14	30	24	9	11	13	14	10	10	20	15	18	21	12
Hp34	1	Ar	15	16,17	13	32	21	11	11	14	15	11	12	21	15	19	21	12
Hp35	1	Ar	15	16,18	13	29	21	10	11	14	14	11	12	20	14	16	21	11
Hp36	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	16	18	21	12
Hp37	1	Ar	13	17,18	13	31	25	10	11	13	14	11	12	20	16	15	19	12
Hp38	1	Ar	14	13,13	13	30	25	10	11	13	14	11	13	19	15	18	23	11
Hp39	1	Ar	14	13,19	13	30	23	11	11	12	14	10	11	20	14	18.2	21	11
Hp40	1	Ar	13	13,16	14	30	24	9	11	13	14	10	10	20	16	18	21	12
Hp41	1	Ar	13	13,14	13	29	24	9	11	13	14	10	10	20	16	17	21	12
Hp42	1	Ar	13	14,14	13	29	23	9	11	13	14	10	10	20	17	17	21	12
Hp43	1	Ar	13	13,14	13	29	24	9	11	13	14	10	10	19	16	17	22	12
Hp44	1	Ar	15	17,2	13	30	21	10	11	14	14	11	11	21	16	17	21	11
Hp45	1	Ar	15	17,18	12	29	21	10	11	14	14	11	12	21	15	19	21	11
Hp46	1	Ar	14	13,14	14	30	24	9	11	13	14	10	10	20	16	18	22	12
Hp47	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	15	18	20	12
Hp48	1	Ar	13	13,14	13	29	23	9	11	13	14	10	10	20	15	18	21	12
Hp49	1	Ar	13	12,16	14	33	22	9	11	13	14	10	14	22	15	14	22	11

## Résultats et Discussion

### Suite

Hp50	1	Ar	14	13,19	14	30	23	10	11	12	14	10	11	20	14	18.2	22	11
Hp51	1	Ar	14	11,14	14	30	23	11	13	12	15	12	11	19	15	16	23	12
Hp52	1	Ar	14	11,14	14	31	23	11	13	12	15	12	11	19	15	16	23	12
Hp53	1	Ar	13	16,17	12	29	23	10	11	13	14	10	10	20	15	14	21	11
Hp54	1	Ar	13	13,13	14	30	24	9	11	13	14	10	10	20	15	17	21	12
Hp55	1	Ar	13	13,15	14	30	24	9	11	13	14	10	10	20	16	19	22	12
Hp56	1	Ar	14	13,17	13	30	23	12	11	12	14	10	11	20	14	18.2	21	11
Hp57	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	15	19	22	12
Hp58	1	Ar	14	11,14	13	29	24	11	13	13	15	11	12	19	17	17	24	12
Hp59	1	Ar	15	14,15	12	28	22	10	11	13	15	10	11	21	15	17	21	12
Hp60	1	Ar	13	13,14	14	32	23	9	11	13	14	10	10	20	15	17	21	11
Hp61	1	Ar	13	13,14	13	29	24	8	11	13	14	10	10	20	15	18	21	12
Hp62	1	Ar	14	16,16	14	32	24	10	11	12	15	9	12	19	15	16	20	12
Hp63	1	Ar	13	13,14	14	30	23	9	11	13	14	10	10	20	16	17	21	12
Hp64	1	Ar	13	13,14	14	30	23	9	11	13	14	10	10	20	15	17	21	12
Hp65	1	Ar	13	13,14	13	29	23	9	11	13	14	10	10	20	15	17	21	12
Hp66	1	Ar	13	13,16	14	30	24	9	11	13	14	10	10	20	16	17	21	12
Hp67	1	Ar	15	11,11	13	31	25	11	11	13	14	10	11	20	14	18	21	11
Hp68	1	Ar	13	17,18	12	29	22	10	11	14	14	10	10	20	14	14	21	11
Hp69	1	Ar	14	11,14	13	29	24	11	13	13	15	12	11	19	15	17	23	12
Hp70	1	Ar	14	14,17	14	30	24	10	11	12	15	9	11	21	15	20	21	11
Hp71	1	Ar	13	17,19	12	29	23	10	11	10	14	10	10	20	15	14	21	11
Hp72	1	Ar	16	16,17	13	30	22	10	11	14	15	13	12	21	15	16	21	10
Hp73	1	Ar	12	13,14	14	30	24	9	11	13	14	10	10	20	15	17	21	12
Hp74	1	Ar	18	14,14	12	29	22	10	11	13	17	10	12	20	13	15	20	11
Hp75	1	Ar	13	17,18	12	29	23	10	11	10	14	10	10	20	15	14	21	11
Hp76	1	Ar	13	17,14	13	30	23	9	11	13	14	10	10	20	14	17	21	12
Hp77	1	Ar	14	13,19	13	30	23	10	11	12	14	10	11	20	14	20	21	11
Hp78	1	Ar	14	13,19	13	30	23	12	11	12	14	10	11	20	13	19	21	11
Hp79	1	Ar	15	13,13	13	30	23	10	11	12	14	10	11	20	14	0	21	11
Hp80	1	Ar	14	13,2	13	30	23	11	11	12	14	10	11	20	13	19	21	11
Hp81	1	Ar	14	13,19	13	30	22	11	11	8	14	10	12	20	0	19	21	10
Hp82	1	Ar	13	14,14	14	30	24	9	11	8	14	10	10	20	0	16	21	10
Hp83	1	Ar	14	13,13	14	30	24	9	11	13	14	10	10	20	15	15	23	12
Hp84	1	Ar	13	14,14	14	30	23	9	11	13	14	10	10	20	15	17	21	12
Hp85	1	Ar	14	13,19	14	30	23	11	11	12	14	10	11	20	14	19	21	11
Hp86	1	Ar	13	13,14	14	30	25	9	11	13	14	10	10	20	16	18	21	12
Hp87	2	Ar\Sw	13	13,14	12	28	24	9	11	13	14	10	11	20	14	18	21	12
Hp88	1	Ar	14	13,19	13	30	23	11	11	8	14	10	11	20	0	19	21	10
Hp89	1	Ar	13	13,14	14	30	24	9	11	8	14	10	10	21	0	16	22	10
Hp90	1	Ar	13	17,18	10	30	23	10	11	13	14	10	10	20	16	0	21	8
Hp91	1	Ar	14	13,2	13	30	23	10	11	12	14	10	11	20	14	19	21	11
Hp92	1	Ar	13	13,14	14	30	23	9	11	8	14	10	10	21	0	17	21	10
Hp93	1	Ar	16	15,15	12	31	21	10	11	14	14	11	12	21	15	18	21	11
Hp94	1	Ar	13	18,18	12	29	22	10	11	8	14	10	11	20	18	19	21	10
Hp95	1	Ar	14	13,2	13	29	23	11	11	8	14	10	11	21	0	20	21	10
Hp96	1	Ar	11	13,14	9	28	19	9	11	12	14	8	9	20,21	17	16	19	8
Hp97	1	Ar	14	12,19	14	31	23	12	11	8	14	10	11	20	0	19	20	10
Hp98	1	Ar	14	13,18	13	30	24	11	11	12	14	10	11	20	14	19	21	11
Hp99	1	Ar	15	12,14	12	29	23	10	14	8	14	10	12	21	17	19	19	10
Hp100	1	Ar	14	13,2	13	29	23	10	11	8	14	10	11	20	0	19	21	10
Hp101	1	Ar	14	11,13	14	31	24	11	13	14	15	12	13	19	15	14	23	12
Hp102	1	Ar	14	13,19	13	29	23	11	11	8	14	10	11	20	0	19	21	10

## Résultats et Discussion

### Suite

Hp103	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	19	16	18	21	11
Hp104	1	Ar	13	17,18	12	29	23	10	11	10	14	10	10	20	15	15	21	11
Hp105	1	Ar	13	13,14	14	30	25	9	11	8	14	10	10	20	0	18	21	12
Hp106	1	Ar	14	13,2	13	30	23	11	11	8	14	10	11	20	0	18	21	10
Hp107	1	Ar	13	13,14	13	29	24	9	11	13	14	10	11	20	15	19	20	12
Hp108	1	Ar	15	16,18	13	29	24	10	10	8	14	9	12	19	0	17	21	10
Hp109	1	Ar	14	13,18	13	30	23	11	11	13	14	10	11	20	14	19	22	11
Hp110	1	Ar	14	13,15	13	32	22	9	11	8	14	10	15	23	0	14	22	10
Hp111	1	Ar	13	13,14	14	30	25	9	11	14	14	10	11	20	16	19	21	12
Hp112	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	15,16	18	21	12
Hp113	1	Ar	13	13,14	14	31	23	9	11	13	14	10	10	20	15	16	21	11
Hp114	1	Ar	14	12,16	13	30	22	9	11	13	14	10	12	22	15	14	21,22	13
Hp115	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	20	15	16	21	12
Hp116	1	Ar	14	13,19	13	29	23	11	11	12	14	10	12	20	14	18	21	11
Hp117	1	Ar	14	13,19	13	31	23	11	11	12	14	10	11	20	14	19	21	11
Hp118	1	Ar	14	11,15	13	31	22	9	11	13	14	10	13	23	15	19	21	11
Hp119	1	Ar	14	13,18	10	29	23	9	11	8	14	10	11	20	0	16	21	8
Hp120	1	Ar	14	13,14	13	29	23	11	11	13	14	10	11	20	14	18	21	11
Hp122	1	Ar	13	13,14	14	30	24	9	11	13	14	10	10	19	15	17	21	12
Hp123	1	Ar	17	14,15	12	30	22	10	11	8	17	10	12	20	0	14	20	10
Hp124	1	Ar	13	12,14	13	29	24	9	10	13	14	10	10	20	16	17	20	12
Hp125	1	Ar	13	13,19	14	30	24	11	10	8	14	10	12	20	0	18	21	10
Hp126	1	Ar	13	13,14	13	31	23	9	11	13	14	9	10	20	16	18	21	12
Hp127	1	Ar	13	17,14	13	30	22	9	10	8	14	10	10	20	14	17	21	12
Hp128	1	Ar	14	13,19	14	30	23	11	10	8	14	10	12	20	0	19	21	10
Hp129	1	Ar	13	13,14	14	30	23	9	11	13	14	10	11	19	15	18	21	12
Hp130	1	Br	13	13,14	14	30	24	9	12	14	14	11	10	19	16	18	22	10
Hp131	1	Br	17	17,19	14	29	21	10	10	8	14	10	11	20	14	18	21	13
Hp132	1	Br	13	13,13	14	30	24	9	11	13	14	10	10	20	16	17	22	12
Hp133	1	Br	14	13,14	13	29	23	11	10	12	14	10	11	19	14	18	21	11
Hp134	1	Br	14	14,15	13	29	24	11	11	12	14	10	11	19	15	18	21	12
Hp135	1	Br	13	13,14	13	30	23	11	10	12	14	10	11	19	14	18	21	12
Hp136	1	Br	13	17,18	12	29	23	10	11	8	14	10	10	20	18	19	21	12
Hp137	1	Br	14	14,14	14	29	23	9	12	8	14	10	10	20	14	16	22	10
Hp138	1	Br	14	13,14	14	30	23	9	11	13	14	10	10	20	16	15	22	12
Hp139	1	Br	14	14,17	14	30	24	10	13	13	14	9	12	19	16	15	20	11
Hp140	1	Br	14	14,16	14	30	24	9	12	13	14	10	10	20	15	17	22	12
Hp141	1	Br	13	13,14	13	29	24	9	11	13	14	10	10	20	16	19	21	12
Hp142	1	Br	15	14,15	12	29	22	10	11	13	17	10	13	20	16	15	21	11
Hp143	1	Br	14	13,19	13	30	23	12	11	12	14	10	11	20	14	18	21	11
Hp144	1	Br	15	14,19	13	29	23	10	11	12	15	9	12	20	15	15	23	10
Hp145	1	Br	16	17,18	13	30	21	10	11	15	13	11	9	21	15	15	21	11
Hp146	1	Br	15	14,16	13	31	21	10	11	14	13	11	12	19	15	19	24	11
Hp147	1	Br	14	13,14	14	30	23	9	11	13	13	10	10	20	16	18	21	13
Hp148	1	Br	15	13,13	13	29	23	10	11	13	14	12	11	20	14	17	21	12
Hp149	1	Br	13	17,18	12	29	22	10	11	10	14	10	10	20	15	14	21	11
Hp150	1	Br	13	13,14	15	31	24	9	11	13	14	10	10	20	15	18	21	12
Hp151	1	Br	14	17,19	14	31	24	10	11	13	14	10	12	19	15	14	22	13
Hp152	1	Br	14	13,19	13	30	23	11	11	12	14	10	12	20	14	18	21	11
Hp153	1	Br	15	12,16	12	29	23	10	11	13	16	11	12	21	17	17	21	12
Hp154	1	Br	14	13,18	13	30	23	11	11	12	14	10	11	20	15	18	21	11
Hp155	1	Br	16	14,17	11	28	23	10	11	13	17	10	9	19,20	15	15	21	11
Hp156	1	Br	14	14,19	12	28	26	11	11	12	14	11	11	20	16	17	24	11

## Résultats et Discussion

### Suite

Hp157	1	Br	13	14,14	14	30	23	9	11	13	14	10	10	20	16	19	21	12
Hp158	1	Br	15	14,15	12	29	21	10	11	13	16	10	11	22	15	17	21	12
Hp159	1	Br	13	13,15	14	30	24	9	11	13	14	10	11	20	14	18	22	12
Hp160	1	Br	14	13,16	14	30	24	9	11	13	14	10	10	20	16	18	21	12
Hp161	2	Br\Br	15	12,16	12	29	22	10	11	14	16	10	11	21	17	17	21	12
Hp162	2	Br\Sw	15	13,14	13	29	24	11	13	13	14	12	12	19	15	17	23	13
Hp163	1	Br	11	14,14	14	30	24	9	11	13	14	10	10	20	15	17	22	12
Hp164	1	Br	13	16,17	12	29	22	10	11	13	14	10	10	20	15	14	21	11
Hp165	1	Br	13	13,15	14	30	24	9	11	13	14	10	10	21	16	18	21	12
Hp166	1	Br	13	13,14	14	30	24	9	11	13	14	10	10	20	16	18	23	12
Hp167	1	Br	13	13,15	14	30	23	9	11	13	14	10	9	20	16	18	21	12
Hp168	1	Br	17	14,15	12	29	22	10	11	13	17	10	8	19	14	15	20	11
Hp169	1	Br	15	16,16	14	29	24	10	11	12	14	11	13	21	14	17	22	11
Hp170	1	Br	13	13,14	14	30	24	10	11	13	14	10	11	20	16	17	20	12
Hp171	1	Br	14	13,19	13	30	23	11	11	12	14	10	12	20	14	19	21	11
Hp172	1	Br	13	13,14	12	28	24	9	11	13	14	11	11	19	16	17	21	12
Hp173	1	Br	14	16,19	12	30	23	10	11	12	14	10	10	20	15	15	22	11
Hp174	1	Br	13	14,14	14	30	23	9	11	12	15	10	10	21	16	18	22	12
Hp175	1	Br	13	13,14	13	30	24	9	11	13	14	10	10	20	15	18	22	12
Hp176	1	Br	14	14,16	12	32	24	10	11	12	15	9	13	19	15	16	20	12
Hp177	1	Br	14	13,19	13	29	23	11	11	12	14	10	10	20	14	18	21	11
Hp178	1	Br	13	13,15	14	30	24	9	11	13	14	9	10	20	15	18	21	12
Hp179	1	Br	14	13,19	13	30	23	11	11	12	14	10	11	20	14	18	21	11
Hp180	1	Br	14	12,15	13	30	24	10	13	13	15	12	11	19	16	17	23	13
Hp181	2	Br\Br	14	18,19	13	30	23	11	11	12	14	10	11	20	14	18	21	11
Hp182	1	Br	11	14,15	14	30	24	9	11	13	14	10	11	21	15	17	22	12
Hp183	1	Br	17	16,16	12	31	21	11	11	14	14	11	11	21	16	17	21	11
Hp184	1	Br	14	15,16	12	28	23	11	11	13	15	8	12	20	15	17	21	12
Hp185	2	Br\Sw	17	16,16	13	32	21	11	11	14	14	11	11	21	16	17	21	11
Hp186	2	Br\Sw	15	15,16	13	29	21	10	11	14	14	11	12	20	16	19	20	11
Hp187	1	Br	13	13,14	13	29	24	9	11	13	14	10	10	19	17	18	21	12
Hp188	2	Br\Sw	13	17,17	12	29	23	10	11	13	14	10	10	20	15	14	22	11
Hp189	2	Br\Sw	15	14,14	12	30	22	10	11	14	13	10	12	21	15	16	20	11
Hp190	1	Br	13	17,18	12	29	22	10	11	13	14	10	11	21	15	14	21	11
Hp191	1	Br	11	14,15	14	30	24	9	11	13	14	10	11	20	15	17	22	12
Hp192	1	Br	16	16,16	13	31	23	11	11	13	15	10	11	21	16	17	22	11
Hp193	1	Br	14	15,16	12	28	24	11	11	12	14	8	12	20	15	17	21	13
Hp194	1	Br	17	16,16	13	32	21	11	11	14	14	11	11	19	16	17	21	11
Hp195	1	Br	15	15,16	12	31	22	11	11	13	15	8	11	21	15	16	21	12
Hp196	1	Br	14	12,15	13	30	24	11	13	13	15	12	11	19	16	17	23	13
Hp197	1	Br	11	14,14	14	30	24	9	11	13	14	10	11	20	15	17	22	12
Hp198	1	Br	17	16,16	13	31	21	11	11	14	14	11	11	21	16	17	21	11
Hp199	1	Br	14	15,16	12	28	23	11	11	13	14	8	12	20	15	17	21	12
Hp200	1	Sw	14	14,15	12	28	22	10	11	13	15	10	11	21	15	17	21	12
Hp201	1	Sw	13	13,14	13	29	24	9	11	13	14	10	10	20	17	18	21	12
Hp202	1	Sw	15	14,15	12	28	22	10	11	13	15	10	11	21	15	17	21	12
Hp203	1	Sw	13	17,18	12	29	22	10	11	13	14	10	10	21	15	14	21	11
Hp204	1	Sw	13	17,18	12	29	22	10	11	14	14	10	10	20	15	14	21	11
Hp205	1	Sw	13	13,14	14	30	23	9	11	13	14	10	10	21	16	17	21	12
Hp206	1	Sw	14	14,16	14	32	24	10	11	12	15	9	12	19	15	16	20	12
Hp207	1	Sw	13	13,14	14	30	24	9	11	13	14	10	10	20	15	18	20	11
Hp208	1	Sw	13	14,16	14	29	23	10	11	12	14	10	12	20	14	18	20	12
Hp209	1	Sw	13	17,18	12	29	23	10	11	13	14	10	10	20	15	14	21	11

## Résultats et Discussion

### Suite

Hp210	1	Sw	13	13,15	14	30	24	9	11	13	14	10	11	20	16	18	21	12
Hp211	1	Sw	14	13,18	13	30	23	11	11	12	14	10	11	20	15	19	21	11
Hp212	1	Sw	16	15,19	14	34	21	11	11	14	14	11	13	20	15	16	21	11
Hp213	1	Sw	13	15,16	13	31	24	10	11	13	14	10	12	20	15	16	22	12
Hp214	1	Sw	13	13,14	14	30	24	10	11	13	14	10	11	20	15	19	21	12
Hp215	2	Sw\Sw	14	13,19	13	29	23	10	11	12	14	10	11	20	14	20	21	11
Hp216	1	Sw	13	16,18	12	29	24	10	11	13	14	10	11	20	15	14	22	11
Hp217	1	Sw	13	13,15	14	30	24	9	11	13	14	10	10	21	16	18	22	12
Hp218	1	Sw	13	13,14	14	30	24	9	11	13	14	10	10	20	17	19	21	12
Hp219	1	Sw	13	13,14	14	31	24	9	11	13	14	10	10	20	15	18	21	12
Hp220	1	Sw	13	13,15	14	30	24	9	11	13	14	10	10	20	15	18	21	12
Hp221	1	Sw	15	16,18	13	30	21	10	11	14	14	11	12	20	14	16	21	11
Hp222	1	Sw	13	13,14	13	30	24	9	11	13	14	10	10	20	16	17	21	11
Hp223	1	Sw	13	14,15	14	30	24	10	11	13	14	10	11	20	16	18	21	12
Hp224	1	Sw	13	13,14	14	30	24	9	11	12	14	10	10	20	14	18	20	11
Hp225	1	Sw	13	13,14	13	29	24	9	11	13	14	10	10	20	15	18	21	12
Hp226	1	Sw	15	16,16	12	30	21	9	11	13	14	10	10	20	15	16	21	12
Hp227	1	Sw	13	13,14	14	30	24	9	11	13	14	10	10	20	15	19	21	12
Hp228	1	Sw	13	14,14	14	30	23	9	11	13	14	10	10	20	16	17	21	12
Hp229	1	Sw	13	13,13	14	30	24	9	11	13	14	10	10	20	16	18	21	12
Hp230	1	Sw	13	13,15	14	30	24	9	11	13	14	10	10	20	16	17	21	13
Hp231	1	Sw	13	16,17	12	29	23	9	11	13	14	9	10	19	15	14	21	11
Hp232	1	Sw	14	14,17	14	32	24	10	11	12	15	9	12	20	15	16	20	12
Hp233	1	Sw	13	13,14	14	30	23	9	11	13	14	10	10	20	15	17	20	12
Hp234	1	Sw	14	13,14	14	28	22	9	11	14	14	9	10	20	15	17	20	12
Hp235	1	Sw	15	11,14	13	30	25	11	11	13	14	11	10	20	15	18	23	13
Hp236	1	Sw	14	13,14	14	30	24	9	11	13	14	10	10	20	15	17	22	12
Hp237	1	Sw	15	13,15	14	31	24	11	13	13	14	12	12	19	15	17	23	12
Hp238	1	Sw	13	17,18	12	29	23	10	11	13	14	10	10	21	15	14	21	11
Hp239	1	Sw	13	13,14	14	30	24	9	11	13	14	10	10	20	17	18	21	11
Hp240	1	Sw	15	14,14	12	28	22	10	11	13	15	10	11	20	15	17	21	12
Hp241	1	Sw	13	14,14	14	30	23	9	11	13	14	10	10	20	16	18	22	12
Hp242	1	Sw	15	13,13	13	29	24	10	11	12	14	11	10	19	17	15	23	11
Hp243	1	Sw	16	15,15	12	30	21	10	11	14	14	11	12	21	15	18	21	11
Hp244	1	Sw	13	13,14	14	28	24	9	11	13	14	10	12	19	14	17	21	12
Hp245	1	Sw	15	12,14	12	29	23	10	14	12	14	10	12	21	15	19	20	11
Hp246	1	Sw	15	16,17	13	30	21	10	11	14	14	11	11	21	15	18	23	11
Hp247	1	Sw	13	13,14	13	29	24	9	11	13	14	9	11	20	14	19	21	12
Hp248	1	Sw	14	13,15	13	32	22	9	11	13	14	10	15	23	15	15	22	11
Hp249	1	Sw	13	13,14	14	30	23	9	11	12	14	11	10	20	16	18	21	12
Hp250	1	Sw	14	13,19	13	30	23	11	11	12	14	10	11	20	14	19.2	21	11
Hp251	1	Sw	15	12,14	12	29	23	11	14	12	14	10	12	20	15	13	20	11
Hp252	1	Sw	14	13,19	13	30	23	12	11	12	14	10	12	20	14	18.2	21	11
Hp253	1	Sw	14	13,19	13	29	22	9	11	13	14	10	11	20	14	18.2	21	11
Hp254	1	Sw	14	13,19	14	31	23	10	11	12	14	10	11	20	14	18.2	21	11
Hp255	1	Sw	13	17,19	13	30	23	11	13	12	14	10	11	23	14	18.2	22	11
Hp256	1	Sw	16	13,19	12	30	21	9	11	12	14	10	10	20	14	18.2	22	11
Hp257	1	Sw	15	17,18	13	29	23	11	11	12	14	10	11	20	14	18.2	21	12

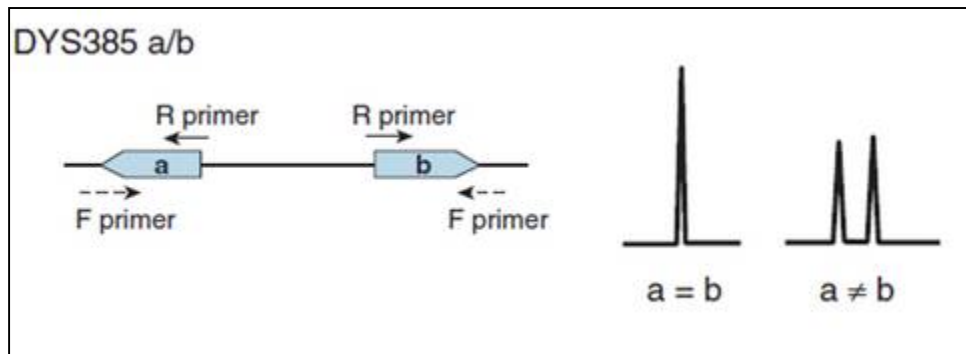
Le **tableau 5** illustre la distribution des fréquences alléliques ainsi que la diversité génétique (équivalent au pouvoir de discrimination) pour les 17 loci STR du chromosome Y analysés dans notre population étudiée. Le locus **DYS385 a/b** indique un degré de polymorphisme élevé avec une valeur de 0,887 suivi par le locus **DYS458** (0,820), comme étant un locus unique. La particularité du locus **DYS385 a/b**, dont la structure répétitive est un tétra-nucléotidique [GAAA]<sub>n</sub>, c'est qu'il se présente en duplicat sur le chromosome Y (**a/b**) avec une distance entre la région **a** et **b** de 40 775 pb et ces deux régions sont amplifiées par les mêmes amorces sens et anti-sens (Figure 22).

**Tableau 5** : Distribution des fréquences alléliques ainsi la diversité génétique des 17 loci Y-STR de la population étudiée.

Allèles	DYS 19	DYS 389 I	DYS 389 II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 437	DYS 438	DYS 439	DYS 448	DYS 456	DYS 458	DYS 635	Y GATA H4	Génotypes	DYS 385 a/b	
8					0.007		0.086		0.019	0.004					0.011	11,11	0.004	
9		0.004			0.434				0.049	0.019						11,12	0.007	
10		0.007			0.307	0.030	0.015		0.764	0.416						0.079	11,13	0.004
11	0.019	0.004			0.228	0.891			0.116	0.330					0.386	11,14	0.022	
12	0.004	0.206			0.004	0.011	0.213		0.049	0.184					0.479	11,15	0.004	
13	0.442	0.382			0.004	0.056	0.566	0.019	0.004	0.037		0.011	0.004		0.045	12,12	0.004	
14	0.315	0.390				0.011	0.109	0.846		0.004		0.206	0.090			12,14	0.022	
15	0.146	0.007					0.007	0.097		0.007		0.434	0.052			12,15	0.007	
16	0.034							0.019				0.225	0.097			12,16	0.019	
17	0.034						0.004	0.019				0.049	0.251			12,19	0.004	
18	0.007											0.007	0.273			13,13	0.030	
18.2													0.052			13,14	0.292	
19				0.004							0.135		0.150	0.015		13,15	0.056	
19.2													0.004			13,16	0.011	
20											0.667		0.019	0.109		13,17	0.007	
21				0.101							0.154			0.633		13,18	0.026	
22				0.120							0.022			0.157		13,19	0.094	
23				0.356							0.015			0.071		13, 20	0.019	
24				0.390										0.011		14,14	0.052	
25				0.026												14,15	0.049	
26				0.004												14,16	0.019	
27																14,17	0.015	
28			0.056													14,19	0.007	
29			0.303													15,15	0.007	
30			0.494													15,16	0.030	
31			0.097													15,19	0.004	
32			0.041													16,16	0.034	
33			0.004													16,17	0.022	
34			0.004													16,18	0.015	
15,16												0.004				16,19	0.004	
19, 20											0.004					17,14	0.007	
20,21											0.004					17,17	0.007	
21,22														0.004		17,18	0.064	
AN												0.064	0.007			17,19	0.015	
																17, 20	0.004	
																18,18	0.004	
																18,19	0.007	
DG	0.684	0.662	0.652	0.698	0.668	0.203	0.617	0.275	0.399	0.685	0.514	0.700	0.820	0.560	0.615		0.889	

DG: Diversité génétique

AN: Allèle nul



**Figure 22:** Illustration schématique du locus DYS385 a/b dupliqué sur le chromosome Y.

Si  $a = b$  un seul pic est obtenu et si  $a$  est différent de  $b$  (Divergence) deux pics différents sont obtenus

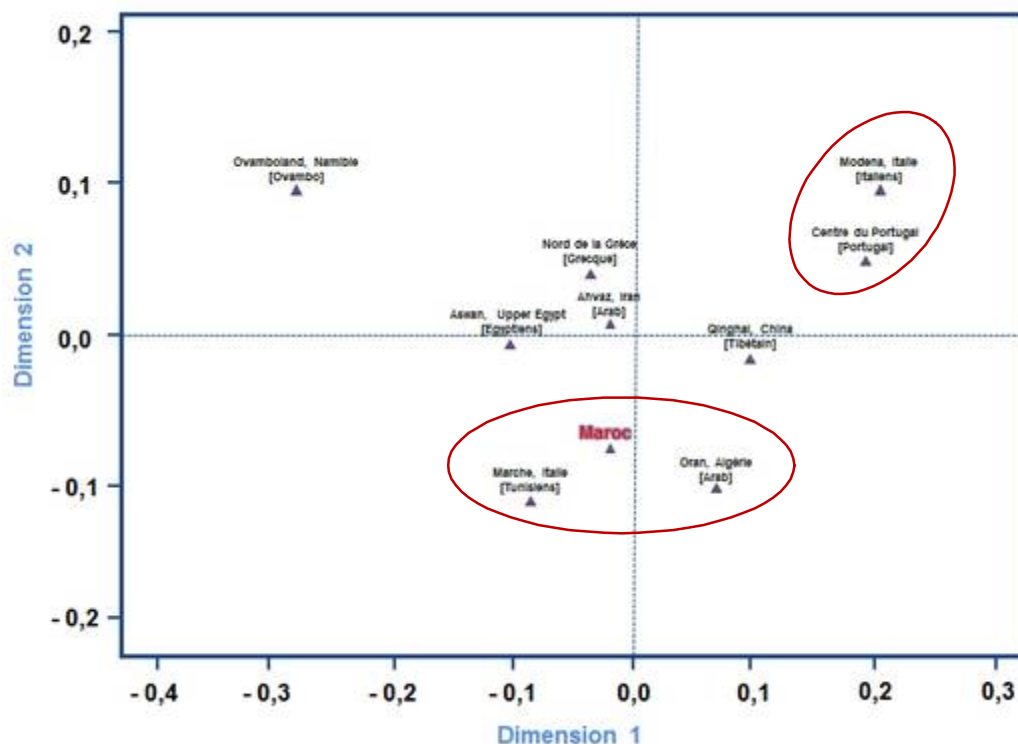
**Source:** Butler, J.M. (2005) Forensic DNA Typing, 2nd Edition, Figure 9.1, ©Elsevier Science/Academic Press.

L'analyse du locus **DYS385 a/b** dans notre population étudiée a révélé 37 génotypes avec une prédominance du génotype (13,14) dont la fréquence est 0.292.

La description de la structure génétique de notre population étudiée a fait pour objet le calcul de la diversité haplotypique des individus (**HD**), la capacité de discrimination (**DC**) et les haplotypes uniques (**UH**). Les résultats de ces différents paramètres précédemment décrits sont respectivement **0.9991**, **0.963** et **247**. Ces différentes valeurs nous permettent de dire que notre population marocaine étudiée présente un degré de polymorphisme très élevé. Contrairement aux marqueurs bi-alléliques du chromosome Y (Y-SNPs), les marqueurs STRs du chromosome Y présentent un taux de mutation plus rapide (0.0012 par locus par génération) (Heyer et *al.*, 1997 ; Kayser et *al.*, 1997 ; Bianchi et *al.*, 1998), qui leurs permettent d'être utilisés pour distinguer une généalogie génétique récente (patronymique) (Roewer et *al.*, 1996 ; Pérez-Lezaun et *al.*, 1997). Ce qui revient à dire que les haplotypes des STRs du chromosome Y forment un groupe « *cluster* » de résultats peu semblables. Typiquement ce groupe de résultat aura une origine ancestrale la plus probable connu sous le terme de *l'haplotype modèle* (probablement proche de l'haplotype fondant l'événement) et aussi une diversité haplotypique auquel il est devenu répandu. Par conséquent, plus l'évènement est ancien plus la diversité haplotypique est grande. D'autre part, si la diversité haplotypique est petite, ceci indique que la population a un ancêtre commun récent, ou bien une récente expansion démographique de la population.

Il est important de noter que même si deux individus partagent le même haplotype pour les marqueurs STRs du chromosome Y, ceci ne veut pas dire qu'ils partagent le même ancêtre du moment que les événements des Y-STR ne sont pas unique. Au lieu de cela, les groupes « cluster » des haplotypes STR du chromosome Y résultent d'un héritage d'événements différents et des histoires différentes qui ont tendance à se chevaucher.

Inversement, l'analyse multidimensionnelle (Figure 23) et les mesures  $\Phi_{st}$  (Tableau 6) montrent une faible différenciation génétique entre notre population marocaine étudiée avec deux populations du Nord-Ouest d'Afrique (Algérie et Tunisie) et la population Ahvaz d'Iran par rapport à certaines populations localisées sur le pourtour de la mer méditerranée tel que la Grèce et l'Égypte. Une nette différenciation génétique a été observée en comparant notre population étudiée avec une population du Sud-ouest de l'Afrique (la Namibie), deux populations du sud de l'Europe (l'Italie et le Portugal) et une population de la région Qinghai en Chine.



**Figure 23:** Illustration graphique de la distance génétique entre la population marocaine étudiée et 9 populations soumises sur la base de données YHRD sur une échelle multidimensionnelle.

**Tableau 6 :** Calcule de la distance génétique entre la population marocaine étudiée (n = 267) et 9 populations citées précédemment par le test AMOVA. Les valeurs  $\Phi_{st}$  sont représentées au-dessus de la diagonale et les valeurs p sont représentées au-dessous de la diagonale.

Populations	Ahvaz, Iran [Arabe]	Aswan, Upper Egypt [Égyptienne]	Central Portugal [Portugal]	Marche, Italy [Tunisiens]	Modena, Italy [Italiens]	Northern Greece [Grèce]	Oran, Algeria [Arabe]	Ovamboland, Namibia [Ovambo]	Qinghai, China [Tibétain]	Marocaine
Ahvaz, Iran [Arabe]	-	0.0783	0.0000	0.0049	0.0000	0.0825	0.0000	0.0000	0.0000	0.0000
Aswan, Upper Egypt [Égyptienne]	0.0507	-	0.0000	0.1192	0.0002	0.0338	0.0023	0.0010	0.0000	0.0049
Central Portugal [Portugale]	0.1866	0.2837	-	0.0000	0.0121	0.0000	0.0000	0.0000	0.0000	0.0000
Marche, Italy [Tunisienne]	0.0682	0.0461	0.3121	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0049
Modena, Italy [Italiens]	0.1985	0.2815	0.0112	0.3388	-	0.0000	0.0000	0.0000	0.0000	0.0000
Northern Greece [Grèce]	0.0140	0.0705	0.1865	0.1027	0.1881	-	0.0000	0.0000	0.0000	0.0000
Oran, Algeria [Arabe]	0.1138	0.1925	0.1704	0.1062	0.2207	0.1359	-	0.0000	0.0000	0.0017
Ovamboland, Namibia [Ovambo]	0.2610	0.2429	0.4794	0.3002	0.4974	0.2448	0.4202	-	0.0000	0.0000
Qinghai, China [Tibétains]	0.1481	0.2572	0.1839	0.2178	0.1991	0.1855	0.1286	0.3890	-	0.0000
Marocaine	0.0724	0.1170	0.2327	0.0393	0.2700	0.1143	0.0287	0.2977	0.1777	-

Le chromosome Y est un important marqueur patri-linéale pour étudier les relations génétiques entre populations. En effet, outre son rôle crucial en génétique médico-légale et en phylogénie, ce système est très informatif pour comparer les profils génétiques des populations humaines en relation avec des événements historiques et des migrations passées. Le Nord d’Afrique a été le théâtre d’importants mouvements de populations à travers l’histoire, mais également d’événements plus localisés qui ont pu avoir une influence importante sur les différenciations génétiques de ces populations. En effet, On constate, qu’il existe une nette différenciation génétique entre la population marocaine étudiée et la population d’Égypte. Cette dernière, située dans la partie la plus orientale de l’aire de répartition de la population autochtone d’Égypte (Berbère), auraient-ils une histoire génétique différente de celle des groupes du Nord-Ouest Africains (Maroc, Algérie, Tunisie, Mauritanie et la Libye)? Les découvertes archéologiques attestent l’occupation de l’oasis de Siouah (Ouest de l’Égypte) au Paléolithique supérieur/Néolithique autour de 6700 - 8800 ans av J.-C. (Hassan, 1978 ; Fakhry, 1973) avec une industrie lithique particulière dite « *libyco-capsienne* ». Celle-

ci se distingue nettement des industries lithiques contemporaines présentes au Maghreb et tendrait plutôt à se rapprocher de la culture Natoufienne<sup>1</sup> du Proche-Orient. D'un point de vue génétique, différentes analyses : système allotypie GM (Coudray et *al.*, 2006), chromosome Y , ADN mitochondrial (Arredi et *al.*, 2004 ; Dugoujon et *al.*, 2009) s'accordent sur le fait que la population de l'Ouest d'Égypte a une structure génétique particulière avec un fort pourcentage de marqueurs « sub-sahariens ». Deux hypothèses sont actuellement émises sur l'origine du peuplement du Maghreb : soit les premiers Berbères seraient venus de l'Est avec la migration des Proto-Méditerranéens depuis l'Égypte vers le Maghreb ; soit ils seraient arrivés depuis l'Ouest avec la migration secondaire de groupes berbères depuis le Maghreb jusqu'en Égypte.

Au Maroc, selon les résultats issus de l'analyse du polymorphisme des marqueurs STR du chromosome Y de notre population étudiée, on constate qu'il n'y a pas de distinction génétique entre les trois groupes ethniques arabophones, berbérophones et Sahraouis. Ceci laisserait alors supposer que l'arrivée des arabes a été, sur les populations marocaines, un phénomène principalement culturel avec l'Islamisation de certains groupes. Qu'en est-il alors de l'impact génétique ? En fait, les marqueurs bi-allélique du chromosome Y (Y-SNPs) ont été choisis comme outil pour mesurer et comprendre l'impact génétique des migrants Arabes sur les populations Nord-Africaines (Partie suivante). On pourrait donc penser que cette similarité génétique entre les différents groupes du Maroc était déjà présente avant l'arrivée des arabes. On peut aussi imaginer qu'elle résulte, pour partie, d'une homogénéisation du pool génique par échanges réciproques entre les diverses populations.

---

<sup>1</sup> La Natoufien est une culture préhistorique du Proche-Orient (Palestine) dont les sites ont été découverts dans les régions bordant la côte méditerranéenne de l'Asie (notamment près du Mont Carmel et dans le Néguev). Le natoufien marque la transition entre la fin de l'Épipaléolithique et le début du Néolithique. La culture natoufienne y succède au Kébarien et précède le Khiamien. (François Raymond Valla, 1976).

## 1.2. Les marqueurs bi-allélique (Y-SNPs)

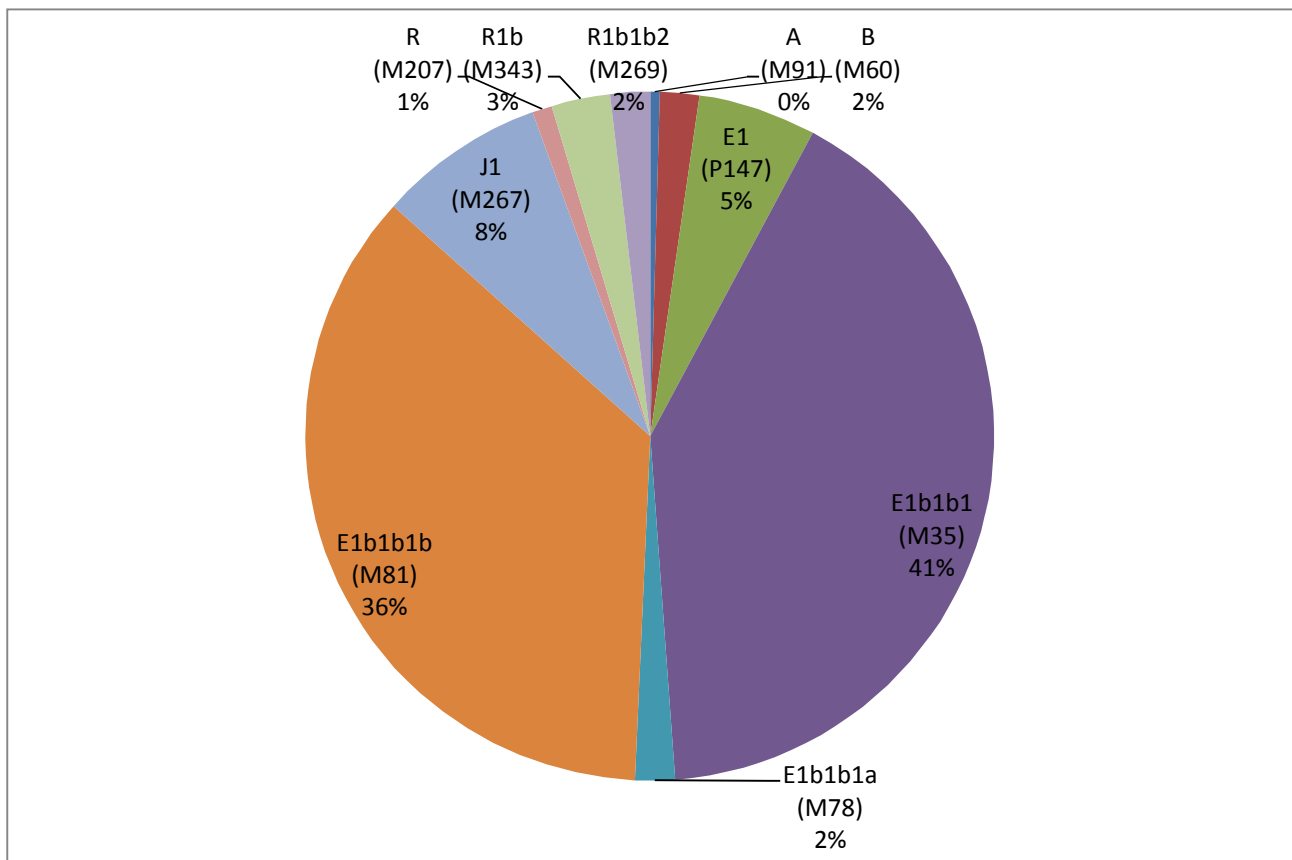
Plusieurs études sur l'analyse de la diversité génétique des populations d'Afrique ont été menées sur des populations sub-sahariens et du Nord Est de l'Afrique dont la localisation géographique de ces différentes populations est situé sur le couloir de migration vers le reste du monde (Tishkoff et Williams, 2002). Pourtant, l'Afrique du Nord-ouest a été l'endroit d'une direction évolutive distincte et nécessite plus d'investigation.

Les résultats d'analyses de 22 marqueurs bi-alléliques du chromosome Y pour les haplogroupes suivants : A-M91 (**A**), B-M60 (**B**), C-RPS4Y711 (**C**), D-M174 (**D**), J-M267 (**J1**), I-M170 (**I**), R-M207 (**R**), E-P147 (**E1**), E-M132 (**E1a**), E-P189 (**E1b1a**), E-M215 (**E1b1b**), E-M81 (**E1b1b1b**), E-M54 (**E2b**), J-M172 (**J2**), J-L24 (**J2a4h**), J-M221 (**J2b**), R-L63 (**R1a**), R-M343 (**R1b**), R-M269 (**R1b1b2**), T-M70 (**T**), E-M35 (**E1b1b1**), E-M78 (**E1b1b1a**), ont permis d'identifier **10 Haplogroupes** dont les plus fréquents sont E-M35 (**E1b1b1**), E-M81 (**E1b1b1b**) et J-M267 (**J1**), respectivement 56%, 49% et 10% (Tableau 7). La figure 24 illustre la fréquence de l'ensemble des haplogroupes rencontré dans notre population marocaine étudiée.

**Tableau 7 :** Les haplogroupes rencontrés dans notre population marocaines étudiées (n = 159) constitué de trois groupes ethnique (Arabophones, Berbèrophones et Sahraouis).

Haplogroupes	Mutations	Groupes ethnique du Maroc (n = 159)			Totale	Fréquence de chaque haplogroupe* %
		Arabophones (n = 42)	Berbèrophones (n = 67)	Sahraouis (n = 50)		
A	M91	0	1	0	1	0.6
B	M60	2	1	1	4	2.5
E1	P147	2	5	5	12	7.5
<b>E1b1b1</b>	M35	20	49	20	89	<b>56</b>
E1b1b1a	M78	1	3	0	4	2.5
<b>E1b1b1b</b>	M81	17	18	43	78	<b>49</b>
<b>J1</b>	M267	5	6	6	17	<b>10.6</b>
R	M207	1	1	0	2	1.2
R1b	M343	2	1	3	6	3.7
R1b1b2	M269	2	1	1	4	2.5

\* Fréquence de l'haplogroupe pour l'ensemble des individus (n = 159).



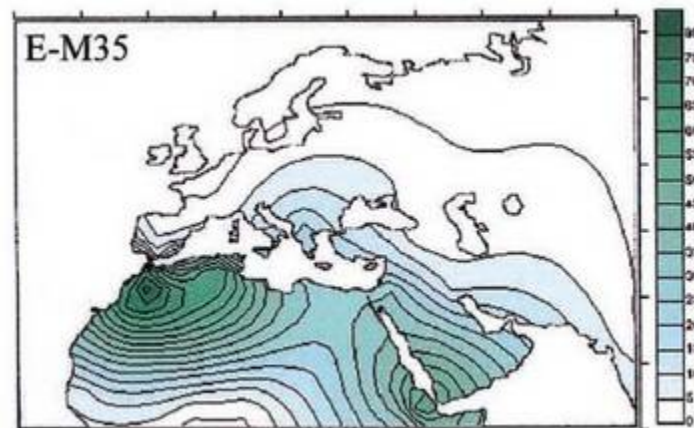
**Figure 24:** Fréquences de l'ensemble des haplogroupes rencontrés dans la population marocaine étudiée (n =159) regroupant les trois groupes ethnies [Arabes, n = 42 ; Berbères n = 67 ; Sahraouis, n = 50).

### I.2.1. Analyse de l'haplogroupe E

L'haplogroupe E du chromosome Y est le troisième haplogroupe prédominant en Afrique, probablement issu d'une population d'Afrique de l'Est qui a été l'origine d'une expansion vers l'extérieur de l'Afrique « Out of Africa ». La distribution géographique des grands embranchements de ce haplogroupe, suggère que la majorité des colonisations en dehors de l'Afrique par des membres ayant l'haplogroupe E est impliquée par une mutation récente de l'embranchement E-M35 vers les sous-embranchements E-M78, E-M81 et E-M123 qui ont migré vers la péninsule d'Arabie et la côte nord de la méditerranée (Charoni et *al.*, 2009).

i. L'haplogroupe E1b1b1 (E-M35)

L'haplogroupe **E1b1b** est caractérisé par la mutation M215. Son sous-groupe dominant est l'haplogroupe **E1b1b1** (ou **E-M35**), caractérisé par la mutation M35. L'haplogroupe **E1b1b1** représente la plus récente migration majeure de l'Afrique vers l'Europe. Les grands sous-embanchement de E-M35 sont originaires de l'Afrique orientale voisine du Proche Orient et du Nord d'Afrique il y a approximativement 22 400 ans (Cruciana et *al.*, 2007), et s'est dispersé à travers le Proche-Orient pendant le Paléolithique supérieur et le Mésolithique (Figure 25).



**Figure 25:** Carte illustrant la distribution géographique de l'haplogroupe E1b1b1 (E-M35).

Source : Semino et *al.*, 2004.

Certains sous-embanchement de l'haplogroupe E-M35 ont quitté l'Afrique il y a approximativement 1000 ans. Par exemple Battaglia et ses collaborateurs en 2008 ont estimé que E-M78 (E1b1b1a) a été en Europe il y a plus de 10 000 ans. Très récemment, des analyses ADN menées par Lacan et ses collaborateurs (2011) sur des restes humains trouvés dans une sépulture Espagnole qui date de 7000 ans ont montré la présence de ce haplogroupe. Néanmoins, E-M35 présente une migration récente vers l'extérieur de l'Afrique par rapport à l'haplogroupe C (M-130) qui représente une migration costale des premiers Homo sapiens d'Afrique vers l'Asie. Par exemple, Underhill (2002) croit que la structure et le modèle régional des sous-embanchement d'E-M35 procure potentiellement des événements avec lesquels on peut déduire des épisodes d'histoires des populations qui s'associent avec l'expansion agriculture au Néolithique. Concernant la présence d'E-M35 en Europe, Underhill et Kivisild (2007) ont remarqué qu'E-M215 semble représenter une

ancienne migration durant la période du Pléistocène<sup>2</sup> du Nord d'Afrique à travers la péninsule Égyptienne *Sinaï* vers l'Europe. Tandis que cette proposition reste incontestée, parce que récemment Trombetta et ses collaborateurs (2011) propose qu'il y'a une migration supplémentaire de E-M215 de l'Afrique vers le Sud-Ouest d'Europe par la voie maritime.

Les populations européennes qui possèdent l'haplogroupe **E1b1b1** (E-M35) sont pour la plupart originaires du pourtour méditerranéen. Elles auraient reçu l'influence génétique de l'Afrique du Nord ou du Proche-Orient. Ce haplogroupe aurait été amené par des populations interagissant directement avec l'Afrique du Nord (**M81**) ou bien avec le Proche- Orient (**M78-V13**). Ces dernières populations sont venues d'Afrique, via l'Asie Mineure, et ont apporté l'agriculture en Europe au Néolithique. L'haplogroupe E1b1b1 est divisé en 3 sous-groupes principaux : **E1b1b1a1**<sup>\*3</sup> (E-M78), **E1b1b1a1**<sup>\*</sup> (E-M81) et **E1b1b1b2a**<sup>\*</sup> (E-M123).

#### L'haplogroupe E1b1b1a1 (E-M78)

L'haplogroupe **E1b1b1a1** (ou E-M78, Précédemment E1b1b1a) la mutation M78 a été estimé qu'elle soit apparu en Afrique de l'Est il y a environ 18 600 ans (17 000 à 20 000 ans) (Cruciana et al., 2007). E-M78 aurait migré hors de l'Afrique via l'Égypte à la fin de la période glaciaire il y a environ 11 000 ans. E-M78 est largement distribué au Nord et Est de l'Afrique (Égypte, Lybie, Ethiopie, Somalie, Kenya), à l'Asie de l'Ouest (Turquie), en Europe (Albanie, Grèce, Sud d'Italie, Sardaigne) (Figure 26). L'haplogroupe E1b1b1a1 (E-M78) a été retrouvé dans notre population étudié à une faible fréquence ≈ **2,5 %** [*Arabophone*, n = 1 ; *Berbérophones*, n = 3 ; *Sahraouis*, n = 0] contre 6,9 % chez des individus *Berbérophones* de Marrakech (n = 29) analysés pour cette mutation par Cruciana et collaborateurs (2004). Cependant, la faible fréquence de la mutation M78 au Maroc nous laisse supposés que cette dernière a été soit originaire ou réintroduite au Nord-Ouest de l'Afrique « Back to Africa » avec l'une des conquêtes que la Maroc a eues avant son islamisation.

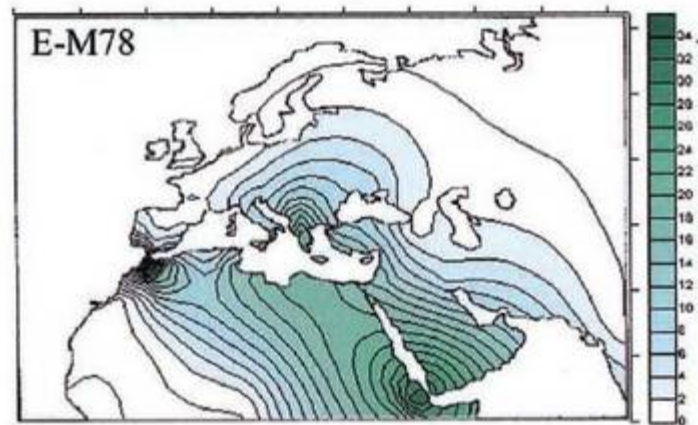
L'haplogroupe E1b1b1a1 (E-M78) est divisé en 4 branches principales : E1b1b1a1a (E-V12), E1b1b1a1b (E-V13), E1b1b1a1c (E-V22) et E1b1b1a1d (E-V65).

---

<sup>2</sup> Le Pléistocène (du grec ancien pleistos, nombreux, et kainos, récent) est la première époque géologique du Quaternaire. Sa fin correspond à celle du Paléolithique et il couvre la plupart des glaciations récentes.

\* La nomenclature des haplogroupes ont été mise à jour sur le texte selon ISOGG (International Society of Genetic Genealogy), Y-DNA Haplogroup Tree 14 Janvier 2013 ([www.isogg.org/tree/](http://www.isogg.org/tree/)).

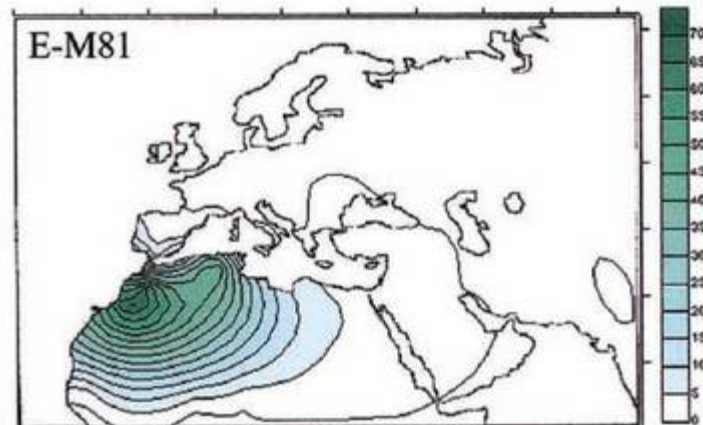
<sup>3</sup> La nomenclature des haplogroupes ont été mise à jour sur le texte selon ISOGG (International Society of Genetic Genealogy), Y-DNA Haplogroup Tree 14 Janvier 2013 ([www.isogg.org/tree/](http://www.isogg.org/tree/)).



**Figure 26:** Carte illustrant la distribution géographique de l'haplogroupe E1b1b1a (E-M78).  
**Source :** Semino et *al.*, 2004.

#### L'haplogroupe E1b1b1a (E-M81)

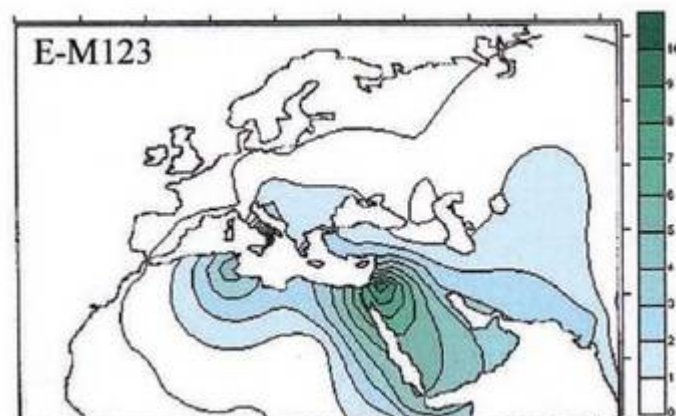
L'haplogroupe **E1b1b1a** (ou E-M81, Précédemment E1b1b1b1) du chromosome Y, le plus répandu au Maghreb est dominé par son sous-groupe E1b1b1a2 (E-M183) (Figure 27). La mutation E-M81 serait apparue dans cette région il y a environ 5 600 ans (Cruciana et *al.*, 2004, Barbara et *al.*, 2004). La fréquence de E1b1b1a se situe généralement autour de 50 % au Maghreb mais atteint parfois 100 % dans certaines populations isolées (Bosch et *al.*, 2001). Cependant, dans notre population marocaine étudiée, la fréquence de l'haplogroupe E1b1b1a est de **49 %** [*Arabophones*, n = 17 ; *Berbérophones*, n = 18 ; *Sahraouis*, n = 43]. Les résultats montrent que la mutation E-M81 est très fréquente chez nos individus *Sahraouis* (86 %) par rapport aux individus étudiée appartenant aux groupes *Arabophones* (34 %) et *Berbérophones* (36 %). Ces résultats sont en accord avec ceux trouvés par Arredi et *al.*, (2004) chez des individus issus du groupe ethnique Sahraouis du Maroc ( $\approx 76$  %). Arredi et *al.*, (2004) a reporté que la fréquence de l'haplogroupe E1b1b1a diminue du Nord-Ouest vers le Nord-Est de l'Afrique ( $\approx 10$  % en Égypte)(Figure 27). En Europe, l'haplogroupe E-M81 se retrouve surtout dans la péninsule Ibérique, où sa fréquence dépasse parfois les 10 % avec la plus forte concentration en Cantabrie, dans le nord de l'Espagne, où sa fréquence atteint 40 % chez les Pasiegos. Il est aussi présent à des fréquences plus faibles en Italie et dans le sud de la France.



**Figure 27:** Carte illustrant la distribution géographique de l'haplogroupe E1b1b1b1a (E-M81).  
**Source :** Semino et al., 2004.

L'haplogroupe E1b1b1b2a (E-M123)

L'haplogroupe **E1b1b1b2a** (ou E-M123, précédemment E1b1b1c) caractérisé par la mutation M123, rare en Europe, est surtout présent en Éthiopie et au Moyen-Orient (Jordanie, Yémen etc) (Figure 28). Il est aussi assez fréquent chez les Juifs et les Berbères Kabyles (10 %). Sa fréquence est également très faible dans le Caucase (< 1 %). La mutation M123 n'a pas été analysée sur notre échantillonnage vu que sa distribution géographique au Nord-Ouest de l'Afrique est presque nulle (Semino et al., 2004).



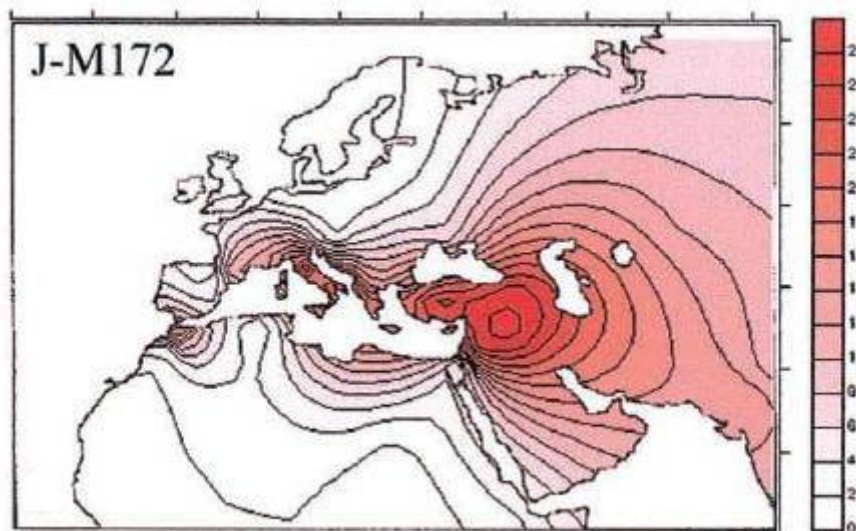
**Figure 28:** Carte illustrant la distribution géographique de l'haplogroupe E1b1b1b2a (E-M123).  
**Source :** Semino et al., 2004.

ii. L'haplogroupe E1 (E-P147)

L'haplogroupe **E1** (ou E-P147) descendant de son ancien ancêtre E-M96 est apparu en Afrique il y'a environ 50 000 ans. La mutation P147 de l'haplogroupe E-P147 a été découverte en 2008 par Karafat et ses collaborateurs. La fréquence de l'haplogroupe E-P147 mutation P147 dans notre population étudiée est de 5 % [Arabophones, n = 2 ; Berbérophones, n = 5 ; Sahraouis, n = 5].

**I.2.2. L'analyse de l'haplogroupe J**

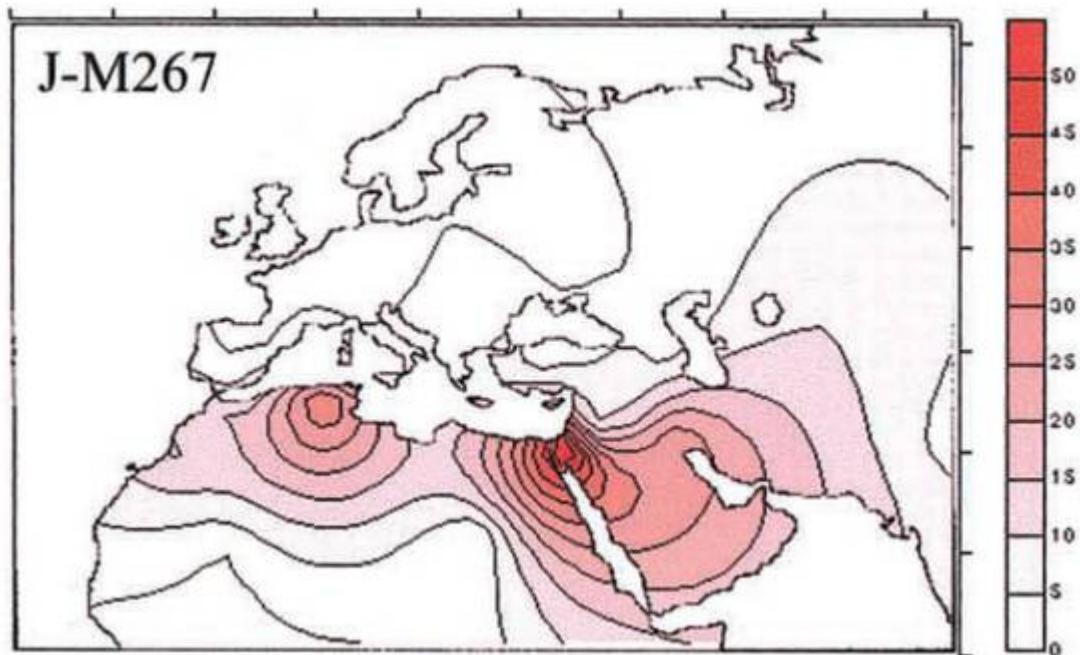
L'haplogroupe J originaire du Moyen Orient est caractérisé par la mutation 12f2a, divisé entre la branche **J2** (J-M172) au Nord (Mésopotamie et Anatolie) qui est la plus répandue en Europe (Figure 29) et la branche **J1** (J-M267) au Sud (Péninsule arabe) (Cinnioglu et al., 2004).



**Figure 29:** Carte illustrant la distribution géographique de l'haplogroupe J2 (J-M172).

Source : Semino et al., 2004.

L'haplogroupe J2 est probablement apparu dans le Nord de la Mésopotamie (pays entre deux fleuves) (Di Giacomo et al., 2004), et s'est répandu vers l'Anatolie le sud de l'Europe, et vers l'Est en Perse et en Inde. J2 est associé au ancien Etrusques, Grecs (minoens), Anatoliens méridionaux, Phéniciens, Assyriens et Babyloniens (Semino et al., 2004). J2 est arrivé en Grèce au début du Néolithique, ou peut-être même avant. Cependant, l'analyse de l'haplogroupe J2 (J-M172) avec ses deux sous embranchements **J2a4h** (J-L24) (récemment **J2a1h**, ISOGG 15 Mars 2013) et **J2b** (J-M221) dans notre population étudiée n'a révélé aucun individu appartenant à ces différents haplogroupes.



**Figure 30:** Carte illustrant la distribution géographique de l'haplogroupe J1 (J-M267).

**Source :** Semino et al., 2004.

L'haplogroupe J1 typiquement sémitique défini par la mutation M267, est apparu il y'a 34 000 ans (Di Giacomo et al., 2004) et constituant la plus grosse partie de la population de la péninsule arabe. Les hommes qui présentent la mutation M267 indique qu'ils ont un ancêtre paternel commun (Cinnioglu et al., 2004). Sa plus forte densité a été constatée au Yemen (72 %), qui doit être sa terre d'origine (Figure 30). Suite à l'analyse de La mutation M267 dans notre population étudiée, la fréquence de l'haplogroupe J1 est de **10,6 %** avec une proportion presque égales entre les trois groupes ethniques [*Arabophones*, n = 5 ; *Berbérophones*, n = 6 ; *Sahraouis*, n = 6]. Les résultats trouvés dans notre échantillonnage sont en accord avec ceux de Semino et collaborateurs (2004).

La conquête musulmane au 7ème siècle du Moyen Orient vers l'Afrique du nord, et dans une moindre mesure de l'Espagne et de la Sicile, propagèrent J1 bien au-delà de l'Arabie, créant un nouveau monde arabe (Arredi, 2004 et Semino, 2004). Il est aussi trouvé avec des fréquences significatives chez les Juifs d'Afrique du Nord particulièrement ceux avec le nom de famille Cohen.

## II. Analyse du polymorphisme de l'ADN mitochondriale

Le tableau 8 résume les analyses statistiques des différents haplotypes de la région contrôle de l'ADN mitochondriale obtenu chez **509** individus marocains non apparentés et appartenant aux trois groupes ethniques (Arabophone, Berbérophone et Sahraoui) de notre population étudiée.

**Tableau 8 :** L'analyse statistique des haplotypes de la région contrôle obtenu chez la population marocaine étudiée. Les statistiques sont basées sur les séquences entières de la région de contrôle (16024-576). Insertions au 16193, 309 et 573 ont été ignorées de l'analyse.

Statistique de la population étudiée		Maroc (n = 509)
Probabilité de concordance aléatoire (RMP)	Empirique	0,52 %
	La somme des carrés	0,64 %
Haplotypes		417 (46 partagés)
Sites polymorphe		236
Les différences moyennes par paire		13,24
La diversité génétique		0,9916

L'analyse du polymorphisme de la région contrôle de l'ADN mitochondriale de notre population étudiée, nous a permis de révéler 417 différents haplotypes, dans lesquels 371 sont des haplotypes uniques et 46 haplotypes partagés (Tableau 9).

**Tableau 9 :** Le Tableau illustre les haplotypes partagés ainsi les haplogroupes de notre population analysée. Pour les échantillons possédant l'hétéroplasmie de taille, les haplotypes énumérés ci-dessous représentent les composantes majeures.

Individus	Arabe (n=42)	Berbère (n=69)	Sahraouis (n=53)	Mixe (n=345)	Haplogroupes	Haplotypes
13	1	1	0	11	R0	16519C 263G 315.1C
11	1	1	0	9	R0	16519C 263G 309.1C 315.1C
8	0	3	1	4	HV0	16153A 16298C 72C 195C 263G 315.1C
5	0	0	1	4	L1b	16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C 73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524-
5	0	3	0	2	L3e5	16041G 16223T 16519C 73G 150T 263G 315.1C 398C 523-524-
4	0	0	2	2	R0	16519C 152C 263G 315.1C
4	0	2	0	2	U6a	16172C 16219G 16278T 73G 152C 263G 315.1C
4	0	0	0	4	W	16223T 16292T 16519C 73G 189G 194T 195C 204C 207A 263G 309.1C 315.1C
3	0	0	0	3	HV0	16153A 16298C 72C 195C 263G 309.1C 315.1C
3	0	0	0	3	HV0	16189C 16298C 72C 263G 315.1C

Suite

3	0	0	0	3	<b>L3b1b1</b>	16124C 16223T 16278T 16294T 16362C 16519C 73G 152C 217C 263G 315.1C 523- 524-
3	0	0	0	3	<b>R0</b>	16362C 16519C 263G 309.1C 315.1C
3	0	0	2	1	<b>R0</b>	16519C 152C 263G 309.1C 315.1C
3	0	0	0	3	<b>R0</b>	16519C 263G 309.1C 309.2C 315.1C
3	0	1	0	2	<b>T1a</b>	16126C 16163G 16186T 16189C 16193- 16294T 16519C 73G 263G 309.1C 315.1C
3	0	0	1	2	<b>T2c</b>	16126C 16213A 16292T 16294T 16296T 16519C 73G 146C 195C 263G 309.1C 315.1C 390G
2	0	1	0	1	<b>H2a1</b>	16354T 263G 309.1C 315.1C
2	0	0	0	2	<b>HV0</b>	16093Y 16298C 16519C 72C 207A 263G 309.1C 315.1C
2	0	0	0	2	<b>HV0</b>	16298C 16519C 72C 263G 309.1C 315.1C
2	0	1	0	1	<b>HV0</b>	16298C 72C 195C 263G 309.1C 315.1C
2	0	0	0	2	<b>HV0</b>	16298C 72C 195C 263G 309.1C 315.1C 489C
2	0	0	0	2	<b>HV0</b>	16298C 72C 195C 263G 315.1C
2	0	0	0	2	<b>HV1</b>	16067T 16274A 16519C 263G 309.1C 309.2C 315.1C
2	0	0	0	2	<b>J2b</b>	16069T 16126C 16193T 16278T 73G 150T 152C 196.1T 263G 295T 309.1C 315.1C 489C
2	0	0	0	2	<b>K1a</b>	16086C 16224C 16256T 16519C 73G 150T 263G 315.1C 497T 523- 524-
2	0	0	1	1	<b>L1b</b>	16126C 16145A 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C 73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524-
2	0	2	0	0	<b>L1b</b>	16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16400T 16519C 73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524-
2	0	0	0	2	<b>L1b1</b>	16126C 16187T 16189C 16223T 16264T 16270T 16278T 16311C 16519C 73G 152C 182T 185T 189G 195C 247A 263G 315.1C 357G 523- 524-
2	0	2	0	0	<b>L2a1l</b>	16189C 16192T 16223T 16278T 16294T 16309G 16390A 73G 143A 146C 152C 195C 263G 315.1C 534T
2	0	0	0	2	<b>L2c</b>	16086C 16223T 16278T 16311C 16318G 16390A 73G 146C 150T 152C 182T 195C 198T 263G 315.1C 325T 523- 524-
2	0	1	0	1	<b>L2c</b>	16223T 16224C 16278T 16311C 16318G 16390A 73G 146C 150T 152C 182T 195C 198T 263G 315.1C 325T 523- 524-
2	0	0	0	2	<b>L3b1</b>	16124C 16223T 16278T 16311C 16362C 16519C 73G 263G 315.1C 523- 524-
2	0	0	2	0	<b>L3b1b1</b>	16124C 16223T 16278T 16291T 16311C 16362C 16519C 73G 152C 263G 315.1C 523- 524-
2	0	0	0	2	<b>L3e5</b>	16041G 16223T 16325C 16519C 73G 150T 263G 315.1C 398C 523- 524-
2	0	0	1	1	<b>L3e5</b>	16041G 16223T 16519C 73G 150T 152C 263G 309.1C 315.1C 398C 523- 524-
2	0	0	1	1	<b>M1</b>	16129A 16182C 16183C 16189C 16249C 16311C 16519C 73G 195C 263G 309.1C 309.2C 315.1C 489C
2	0	0	0	2	<b>M1</b>	16129A 16183C 16189C 16193.1C 16223T 16249C 16311C 16399G 16519C 73G 195C 263G 309.1C 315.1C 489C
2	0	0	0	2	<b>M1</b>	16183C 16189C 16193.1C 16249C 16311C 16519C 73G 195C 263G 315.1C 489C

Suite

2	0	1	1	0	<b>T2c</b>	16126C 16292T 16294T 16519C 73G 146C 152C 263G 279C 315.1C 374G
2	1	1	0	0	<b>U3a</b>	16260T 16343G 16390A 16519C 73G 150T 263G 309.1C 315.1C
2	0	1	0	1	<b>U4</b>	16356C 16362C 16519C 73G 195C 263G 309.1C 315.1C 499A 524.1A 524.2C 524.3A 524.4C
2	0	0	0	2	<b>U5a1</b>	16192T 16256T 16270T 16311C 16327T 16362C 16399G 73G 263G 315.1C
2	0	0	1	1	<b>U6a</b>	16172C 16183C 16189C 16193.1C 16219G 16278T 16295T 73G 103A 263G 315.1C
2	1	0	0	1	<b>U6a</b>	16172C 16219G 16278T 73G 150T 152C 263G 315.1C
2	0	0	0	2	<b>W</b>	16223T 16292T 16519C 73G 189G 194T 195C 199C 204C 207A 263G 315.1C
2	0	1	0	1	<b>X1b1</b>	16183C 16189C 16193.1C 16223T 16278T 16311C 16519C 73G 146C 153G 256T 263G 309.1C 309.2C 315.1C

## II.1. L'Hétéroplasmie ponctuelle et de taille

L'hétéroplasmie ponctuelle a été identifiée chez 43 individus (8,45 % de notre base de données). Deux positions d'hétéroplasmie ponctuelle ont été trouvées chez un individu de notre population marocaine analysée (16182 et 16256).

L'hétéroplasmie de taille a été observée chez 291 individus de notre population marocaine étudiée (57,17 % de notre base de données) situé au niveau des zones où il existe une succession de cytosine, HVI (19,06 %), HVII (47,35 %), et HVIII (1,77 %).

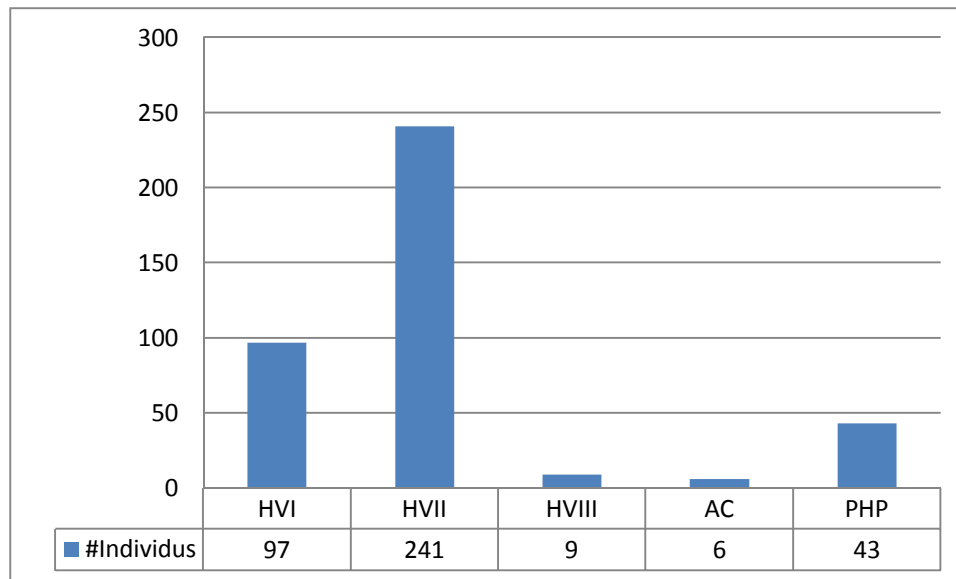
L'hétéroplasmie de taille au niveau de la zone AC de la région contrôle a été observé chez 6 individus de notre population marocaine étudiée (1,18 % de notre base de données) (Figure 31).

## II.2. Duplication et nouvelles mutations

Nous avons identifié dans notre population étudiée une rare duplication de 204 pb chez un individu (Haplogroupe **L2a1**) dans la région contrôle de l'ADN mitochondriale de la position nucléotidique 360 à la position 563. Ce type de duplication en tandem hétéroplasmique de taille ≈260 pb et de faible abondance a été observé pour la première fois par Brockington et *al.*, (1993) chez des patients atteints de myopathies mitochondriales. Torroni et *al.*, (1994) a aussi trouvé une duplication similaire ≈260 pb chez des individus sains appartenant à l'haplogroupe I. Chez ces sujets, le

pourcentage de molécules d'ADNmt dupliqués est toujours faible, et représente moins de 2 % de l'ADNmt totale (Bouzidi et *al.*, 1997).

Des nouveaux polymorphismes jamais décrits ont été également observés dans notre base de données, avec un échantillon qui présente des insertions 524,1 G 524.2C.



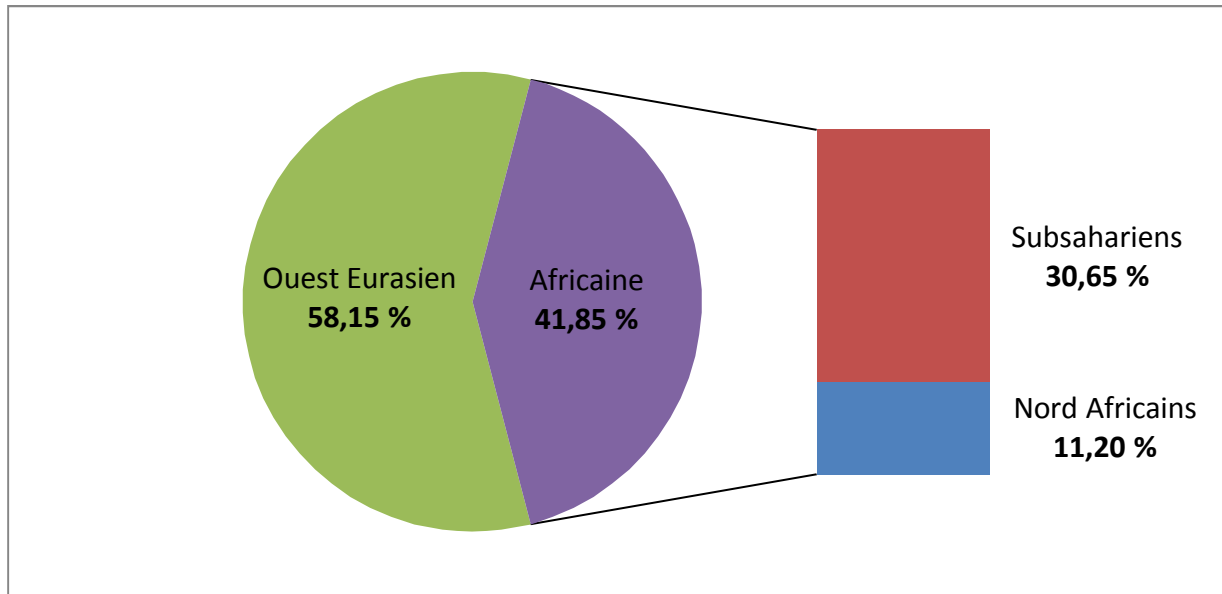
**Figure 31:** Illustration des individus ayant l'hétéroplasmie de taille (HVI, HVII, HVIII et AC) et ponctuelle dans notre population étudiée.

Dans l'ensemble de nos individus marocains analysés (n = 509) pour l'ADN mitochondriale, l'haplogroupe le plus souvent rencontré est **16519C 263G 315.1C** (Haplogroupe **R0**), en d'autres termes l'haplogroupe a été observé chez 28 individus (5,50 % des échantillons). Le résultat obtenu par rapport à cet haplogroupe est en accord avec le résultat rapporté par Turchi et *al.*, (2009) sur la lignée la plus répandue au Nord-Ouest d'Afrique. De plus, un autre haplogroupe vient en deuxième lieu et qui a été observé chez 11 individus (2,16 %) **16153A 16298C 72C 195C 263G 315.1C** (Haplogroupe **HV0**).

### II.3. Haplogroupes présents dans notre population

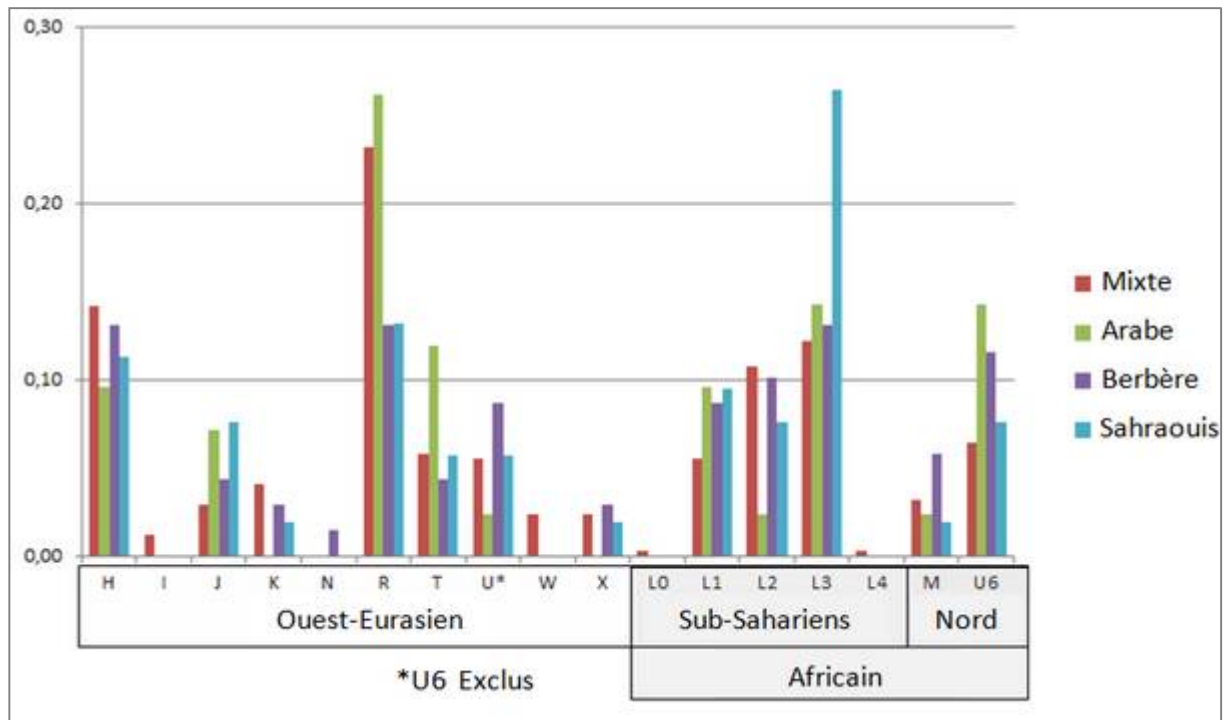
Les données d'analyse de l'ADN mitochondrial des différents groupes de notre population étudiée présentent une composition mosaïque des lignées Ouest-eurasien, des lignées de l'Afrique Subsahariens et des lignées d'Afrique du Nord avec des fréquences de 58,15 %, 30,65% et 11,20%, respectivement (Figure 32). D'une manière générale, les résultats trouvés sont en accord avec ce qui

a été rapporté par la littérature (Brakez et al., 2001 ; Plaza et al., 2003 ; Coudray et al., 2008 et Turchi et al., 2009).



**Figure 32:** Ascendant maternelle de la population marocaine étudiée telle que définie par les lignées d'ADN mitochondrial humain.

Toutefois, en comparant les fréquences des lignées Ouest Eurasien et Subsahariens trouvées dans notre population marocaine étudiée par rapport à ce qui a été rapporté dans la littérature (Turchi et al., 2009), on observe que la fréquence de la lignée Ouest Eurasien obtenue est légèrement inférieure (58,15 vs 63 %), en revanche la fréquence de la lignée Subsaharienne est légèrement supérieure (30,65 vs 25 %). D'une manière plus précise, les haplogroupes les répandus dans notre population marocaine étudiée sont : **R0** (17,49%) et **HV0** (8,25%) qui appartiennent à la lignée Ouest Eurasien, suivie par les lignées du Nord d'Afrique **U6a** (7,27%) et **M1** (3,34%), et puis les haplogroupes de la lignée Subsaharienne **L1b** (5,11%), **L2A1** (3,73%), et **L3e5** (3,34%) (Figure 33).



**Figure 33:** Ascendant maternelle de la population marocaine étudiée telle que définie par les lignées d'ADN mitochondrial humain.

#### II.4. L'haplogroupe détectés principalement dans les populations nord-africaines : U6

L'origine de la lignée **U6** se situe le plus probablement au Nord-Ouest de l'Afrique il y a environ 30.000 ans BP. **U6** est aujourd'hui considéré comme le marqueur de la continuité régionale des populations nord-africaines. Le sous-haplogroupe **U6a** (+/- 27 000 ans BP) marquerait la première expansion humaine paléolithique depuis le Maghreb vers le Nord-Est de l'Afrique. Le sous-groupe **U6a1** semble être le signal d'un mouvement postérieur depuis l'Est de l'Afrique vers le Maghreb et vers le Proche-Orient. Cette migration coïnciderait avec l'expansion des langues afro-asiatiques. Les sous-haplogroupes **U6b** (+/- 10 000 ans BP) et **U6c** (+/- 7000 ans BP) sont uniquement présents à l'Ouest de l'Afrique et ont connu une expansion plus locale : **U6b** aurait même atteint la Péninsule Ibérique par une migration depuis le nord de l'Afrique et **U6b1** et **U6c1** seraient les indicateurs de l'arrivée des colons nord-africains dans l'archipel canarien, probablement à la suite du dessèchement du Sahara (Maca-Meyer et al., 2003).

## II.5. Les haplogroupes fréquents en Eurasie

Le macro-haplogroupe **M** serait issu des populations Est-africaines (Descendant de son ancêtre l'haplogroupe L3) et aurait été répandu en Eurasie il y a environ 60.000 ans BP. Après leur arrivée en Asie, les groupes portant l'haplogroupe fondateur **M** auraient connu une nette expansion démographique et géographique. De nouvelles mutations ont alors conduit à la division de **M** en différents sous-clades: M (2, 5, 7, 8, 9), C, D, G et Q.

Deux explications ont été envisagées comme causes possibles de la présence de **M1** en Afrique :

- Seuls les porteurs de l'haplogroupe **M1** seraient restés confinés en Afrique de l'Est jusqu'à il y a 10.000 à 20.000 ans BP. A partir de là, ils se seraient à nouveau répandus (Quintana-Murci *et al.*, 1999).
- La présence de l'haplogroupe M1 en Afrique est le résultat d'une migration secondaire de l'Asie vers l'Afrique « retour en Afrique » (Gonzalez *et al.*, 2007).

Les fréquences les plus élevées de **M1** sont en Afrique de l'Est (Kivisild *et al.*, 2004) et sont plutôt rares voire absentes au nord de l'Afrique (Rando *et al.*, 1998 ; Brakez *et al.*, 2001 ; Plaza *et al.*, 2003).

## II.6. Les haplogroupes fréquents dans les populations sub-sahariennes

Ces haplogroupes appartiennent à la famille **L**. Ils sont subdivisés en huit lignées : de **L0** à **L7** (Torroni *et al.*, 2006). Ils sont rencontrés à des fréquences très élevées dans les populations de la moitié sud du continent africain (fréquence supérieure à 90%), à des fréquences élevées en Afrique de l'Est et dans la Péninsule Arabe (de 30 à 50%), à des fréquences moyennes au nord de l'Afrique (15 à 30%) et enfin à des fréquences faibles au Proche-Orient (inférieures à 10%) ou quasi nulles en Europe et en Asie.

Le paragroupe **L1** inclurait l'ancêtre commun le plus récent de tous les ADNmt humains et qui serait daté entre 150.000 et 170.000 ans BP (Horai *et al.*, 1995 ; Ingman *et al.*, 2000). L'haplogroupe L1a serait originaire d'une migration depuis l'Afrique de l'Est et aurait environ 40.350 ans BP. Il est commun (≈20 à 25%) à l'est, au centre et au sud-est de l'Afrique. Il est absent au nord, à l'ouest et au sud du continent (Salas *et al.*, 2002). L'haplogroupe L1b a une distribution géographique complètement différente de celle de L1a. Il est concentré à l'ouest de l'Afrique mais est aussi

retrouvé au centre et au nord du continent, notamment dans les régions connectées par la côte Atlantique. Il est très peu présent à l'est, ou sud-est et au sud de l'Afrique. L1b aurait environ 30.550 ans BP (Salas *et al.*, 2002).

Le paragroupe **L2** aurait environ 70.100 ans BP (Salas *et al.*, 2002). L'haplogroupe L2a est le plus fréquent et le plus répandu en Afrique. Il aurait environ 55.150 ans BP. Son origine géographique est difficile à établir mais se situerait probablement dans une région entre l'est et l'ouest de l'Afrique. Les mouvements de populations auraient ensuite répandu L2a dans les deux directions le long du corridor sahélien (Salas *et al.*, 2002). L'haplogroupe L2b est confiné à l'ouest et au centre-ouest de l'Afrique. Il est aussi isolément représenté en Afrique de l'Est et dans la Péninsule Ibérique. Il serait daté autour de 31.600 ans BP (Salas *et al.*, 2002).

Le paragroupe **L3** est distribué à travers tout le continent africain mais atteint les fréquences les plus élevées à l'Est de l'Afrique où il compte pour plus de la moitié de la structure mitochondriale des populations de ces régions. Ce profil de fréquence suggère que L3 serait issu d'une migration depuis l'Afrique de l'Est (Watson *et al.*, 1997). L3 est divisé en différents sous-clades (b, c, d, e, f, h, i, x et w) (Torroni *et al.*, 2006) et serait daté autour de 61.300 ans BP (Salas *et al.*, 2002). Cette date correspond à l'expansion des hommes modernes hors d'Afrique il y a 60.000 à 80.000 ans BP. Au cours de cette migration, une branche de L3 se serait répandue hors d'Afrique et serait à l'origine des haplogroupes ancestraux M et N, eux-mêmes définis comme les précurseurs de tous les lignages mitochondriaux rencontrés en Asie et en Europe. Dans les régions sub-sahariennes, l'haplogroupe L3e est le répandu, le plus fréquent et le plus ancien des haplogroupes L3 (49.250 ans BP, Salas *et al.*, 2002). Il serait originaire de la région du Soudan (Bandelt *et al.*, 2001). L3e1 est distribué dans toutes les populations sub-sahariennes mais il est particulièrement commun au sud-est de l'Afrique. Il est absent à l'ouest et proviendrait de mouvements de populations depuis l'Afrique centrale (Salas *et al.*, 2002). L'haplogroupe L4g est présent uniquement dans les populations est-africaines (Ethiopie). Il aurait 68.800 ans BP (Kivisild *et al.*, 2004).

## **Conclusion générale et perspectives**

## **Conclusion générale et perspectives**

Les résultats obtenus ont permis la description de l'histoire génétique et de l'évolution de la population Nord-Ouest de l'Afrique (Maroc), tout en prenant en considération l'origine, l'emplacement géographique et l'histoire de cette population afin de bien interpréter nos résultats génétiques.

Nos principales conclusions sont les suivantes :

- La population marocaine étudiée a une structure génétique composée de polymorphismes également observés dans d'autres régions du monde mais aussi de marqueurs plus spécifiques notamment les haplogroupes mitochondriaux U6a, M1 et l'haplogroupe E-M81 du chromosome Y.
- Les haplogroupes trouvés dans notre échantillonnage ont une parenté génétique avec les populations européennes tandis que d'autres polymorphismes seraient issus de flux géniques depuis les régions sahariennes. Une différenciation génétique très nette est observée entre notre population marocaine étudiée et une population du Sud d'Égypte ; une faible distinction génétique entre les différents groupes de notre population ; La diversité génétique est très ancienne prouvée par les haplogroupes mitochondriaux et par les haplogroupes du chromosome Y, provient peut-être du Paléolithique et/ou du Néolithique.
- Des mutations nouvellement trouvées dans la population marocaine analysée.
- Les événements de l'époque antique (Phéniciens, Carthaginois, les Romains, les Vandales ou les Byzantins) n'ont pas tous eu d'impacts génétiques et démographiques considérables sur la population Nord-Ouest d'Afrique. La période qui a induit les changements les plus importants est l'arrivée des arabes.
- Les impacts furent majoritairement culturels mais également génétiques selon les polymorphismes des marqueurs bi-alléliques du chromosome Y et de l'ADN mitochondrial « Haplogroupe M1 ». Les populations migrantes ont permis une augmentation et/ou un maintien de la diversité déjà en place. Il se confirmerait donc qu'on se trouve en présence, au Maroc, d'un mélange fortement brassé issu d'une ancienne population méditerranéenne, à caractère européen, et des immigrants venus d'Asie, de l'Afrique subsaharienne et de l'Afrique de l'Est.

## *Conclusion générale et perspectives*

---

- De plus, en ciblant les SNP adéquats, la méthode développée dans ce travail intéresse à la fois la criminalistique et la paléoanthropologie, où la quantité et la qualité de l'échantillon sont limitantes.

Ce travail sera le point de départ pour d'autres perspectives qui consisteront à compléter les données obtenues en:

- Augmentant considérablement la taille de l'échantillon étudié pour les marqueurs bi-alléliques du chromosome Y puis en élargissant les investigations avec d'autres marqueurs SNP (notamment l'haplogroupe E et J);
- Augmentant le nombre des marqueurs STR du chromosome Y ;
- Affiner les résultats pour les haplogroupes obtenus de la région contrôle de l'ADN mitochondrial par le séquençage complet de toute la molécule ;
- Réaliser des missions futures pour un échantillonnage considérable des différents groupe Berbérophones du Maroc dans l'objectif d'affiner le degré de parenté par les mêmes outils génétiques pour mieux corrélér l'histoire de ces populations. De même, le typage d'un plus grand nombre de polymorphismes donnerait une image beaucoup plus précise sur l'état actuelle de la structure de la population moderne et des facteurs intra-populationnelle qui ont permis son évolution.
- Introduire la nouvelle technologie du séquençage à haut débit dans la criminalistique pour l'analyse des différents marqueurs moléculaire.
- Analyser des fossiles humains.

## Références Bibliographiques

## **Références Bibliographiques**

- Abdellah BM. Structures politiques du Maroc colonial - Histoire et perspectives méditerranéennes. 1990. p135-137.
- Anne Bernet. *Les chrétientés d'Afrique*. 2006. éditions de Paris ; p174-176.
- Anslinger K, Weichhold G, Keil W, Bayer B, Eisenmenger W. Identification of the skeletal remains of Martin Bormann by mtDNA analysis. *Int J Legal Med*. 2001. 114: 194 – 196.
- Arredi B, Estella SP, Silvia P, Tatiana Z, Dahmani MF, Mohamed M, Vincenzo LP, Andrea N, Chris TS. A Predominantly Neolithic Origin for Y-Chromosomal DNA Variation in North Africa. *American Journal of Human Genetics*. 2004. 75: 338–345, doi: 10.1086/423147.
- Bandelt HJ, Alves SJ, Guimaraes PE, Santos MS, Brehm A, Pereira L, Coppa A, Larruga JM, Rengo C, Scozzari R, Torroni A, Prata MJ, Amorim A, Prado VF, Pena SD. Phylogeography of the human mitochondrial haplogroup L3e: a snapshot of African prehistory and Atlantic slave trade. *Ann Hum Genet*. 2001. 65 (Pt 6): 549-563.
- Barton RNE, Bouzouggar A, Bronk RC, Collcutt SN, Higham TFG, Humphrey LT, Parfitt S, Rhodes EJ, Schwenninger JL, Stringer CB, Turner E, Ward S. Abrupt climatic change and chronology of the Upper Palaeolithic in northern and eastern Morocco. In Bar-Yosef, O., Mellars, P., Stringer, C., & Boyle, K (eds). *Rethinking the Human Revolution: New Behavioural & Biological Perspectives on the Origins and Dispersal of Modern Humans*. Research Monographs of the Macdonald Institute, Cambridge. 2007. p177-186.
- Battaglia FS, Al-Zahery N, Olivieri A, Pala M, Myres NM, King RJ, Rootsi S. Y-chromosomal evidence of the cultural diffusion of agriculture in southeast Europe. *European Journal of Human Genetics*. 2008. 17 (6): 820–830, doi:10.1038/ejhg.2008.249, PMC 2947100, PMID 19107149.
- Bernard lugan. *Histoire du Maroc des origines à nos jours*. 2000. p12-15.
- Bouzidi M, Poyau A, Godinot C. Co-existence of high levels of a cytochrome b mutation and of a tandem 200 bp duplication in the D-loop of muscle human mitochondrial DNA. *Hum Mol Genet*. 1998. 7:385–91.
- Bouzouggar A, Barton N, Vanhaeren M, d'Errico F, Collcutt S, Higham T, Hodge E, Parfitt S, Rhodes E, Schwenninger JL, Stringer C, Turner C, Ward S, Moutmir A, Stambouli A. 82,000 year-old shell beads from North Africa and implications for the origins of modern human behavior. *Proceedings of the National Academy of Sciences (USA)*. 2007. 104 (24): 9964-9969.
- Bouzouggar A, Barton RNE, Collcutt SN, Parfitt S, Higham TFG, Rhodes E, Gale R. Le Paléolithique Supérieur au Maroc: apport des sites du Nord-Ouest et de l'Oriental. 2006. pp. 138-150.

## Références Bibliographiques

---

- Brakez Z, Bosch E, Izaabel H, Akhayat O, Comas D, Bertranpetit J, Calafell F. Human mitochondrial DNA sequence variation in the Moroccan population of the Souss area. *Ann Hum Biol.* 2001. 28(3): 295-307.
- Brandstätter A., Christine T. Peterson CT., Irwin JA., Mpoke S., Davy K., Koech DK., Parson W., Parsons TJ. Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database. *Int J Legal Med.* 2004. 118: 294–306 DOI 10.1007/s00414-004-0466-z.
- Brockington M, Sweeney MG, Hammans SR, Morgan-Hughes JA, Harding AE. A tandem duplication in the D-loop of human mitochondrial DNA is associated with deletions in mitochondrial myopathies. *Nature Genet.* 1993. 4, 67–71.
- Brookes AJ. The essence of SNPs. *Gene* 1999. 234(2):177-86.
- Brown WM. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc Natl Acad Sci USA.* 1980. 77(6): 3605-3609.
- Butler JM, Kline MC, Decker AE. Addressing Y-chromosome short tandem repeat (Y-STR) allele nomenclature. *Journal of Genetic Genealogy.* 2008. 4(2): 125-148.
- Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature.* 1987.325(6099): 31-36.
- Carcopino J. La Maurétanie tingitane : le Maroc des Romains. 1944. p286-289.
- Carvalho CM, Fujisawa M, Shirakawa T, Gotoh A, Kamidono S, Freitas Paulo T. Lack of association between Y chromosome haplogroups and male infertility in Japanese men. *Am. J. Med. Genet. A.* 2003.116(2):152-8.
- Carvalho M, Anjos MJ, Andrade L, Lopes V, Santos MV, Gamero JJ, Corte Real F, Vide MC. Y-chromosome STR haplotypes in two population samples: Azores Islands and Central Portugal, *Forensic Sci. Int.* 2003. 134 (1) 29–35.
- Casas MJ, Hagelberg E, Fregel R, Larruga JM, Gonzalez AM. Human mitochondrial DNA diversity in an archaeological site in al-Andalus: Genetic impact of migrations from North Africa in medieval Spain. *Am J Phys Anthropol.* 2006. 131(4): 539-551.
- Case JT, Wallace DC. Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblasts. *Somatic Cell Genet.* 1981. 7(1): 103-108.
- Charoni J, Underhill P, Cavalli-Sforza LL. Y Chromosome Diversity, Human Expansion, Drift, and Cultural Evolution. *PNAS.* 2009. 106 (48); 20174-20179.

- Chromosome Consortium Y. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* 2002. 12:339–48.
- Cinnioğlu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC. Excavating Y-chromosome haplotype strata in Anatolia. *Human Genetics.* 2004. 114 (2): 127–48. doi:10.1007/s00439-003-1031-4. PMID 14586639.
- Coudray C, Evelyne G, Morgane G, André S, Georges L, Jean-Michel D. Genetic diversity (GM allotypes and STRs) of the Berber populations and the peopling of Northern Africa. *Antropo.* 2006. 11, 75–84.
- Cruciani F, La Fratta R, Trombetta B, Santolamazza P, Sellitto D, Colomb EB, Dugoujon JM, Crivellaro F. Tracing Past Human Male Movements in Northern/Eastern Africa and Western Eurasia: New Clues from Y-Chromosomal Haplogroups E-M78 and J-M12. *Molecular Biology and Evolution.* (2007). 24 (6): 1300–1311, doi:10.1093/molbev/msm049, PMID 17351267.
- Cruciani F, Roberta LF, Piero S, Daniele S, Roberto P, Pedro M, Elizabeth W, Valentina G, Eliane BC, Boriani Z, João L, Giuseppe V, Rashid A, Francesco C, Nejat A, Martin R, Antonio T, Andrea N, Rosaria S. Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. *Am J Hum Genet.* 2004. 74(5): 1014–1022.
- Darwin C. De l'origine des espèces. 1862. D'après l'édition de 1896 Schleicher Frères, Éditeurs. Traduit de l'édition anglaise définitive par Ed. Barbier.
- De Knijff P. Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. *Am. J. Hum. Genet.* 2000. 67(5):1055–61.
- Dettlaff-Kakol A, Pawlowski R. First Polish DNA "manhunt"-an application of Y chromosome STRs. *Int. J. Legal Med.* 2002. 116(5):289–91.
- Di Giacomo F, Luca F, Anagnou N, Ciavarella G, Corbo RM, Cresta M, Cucci F, Di Stasi L. Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. *Molecular Phylogenetics and Evolution.* 2003. 28 (3): 387–95. doi:10.1016/S1055-7903(03)00016-2. PMID 12927125.
- Dugoujon JM, Coudray C, Torrioni A, Cruciani F, Scozzari F, Moral P, Louali N, Kossmann M. The Berber and the Berbers: Genetic and linguistic diversities. In: *Become Eloquent*. Edited by J.M. Hombert and F. d'Errico. Ed. John Benjamins. 2009. pp 123–146.
- Elena B, Francisc C, David C, Peter JO, Peter AU, Jaume B. High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between north-western Africa and the Iberian Peninsula. *Am J Hum Genet.* 2001. 68: 1019–1029, doi:10.1086/319521.

## Références Bibliographiques

---

- Excoffier LLG, Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*. 2005. **1**:47-50.
- Fakhry A. Siwa Oasis (The American University in Cairo Press). 1973. p214.
- Ferembach D, Dastugue J, Poitrat-Targowla MJ. La Nécropole Epipaléolithique de Taforalt (Maroc Oriental): Étude des squelettes humains. Editions Casablanca, Rabat. 1962. P???
- Ferri G, Alu M, Corradini B, Radheshi E, Beduschi G. Slow and fast evolving markers typing in Modena males (North Italy), *Forensic Sci. Int. Genetic*. 2009. **3** - e31-e33.
- Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R et al. Jefferson fathered slave's last child. *Nature*. 1998. **396**(6706):27-8.
- François RV. Le Natoufien, une culture préhistorique en Palestine. *Cahiers de la Revue Biblique*. 1976 n° 15, Volume 53 ; p350-352.
- Fujihara J, Yuasa I, Muro T, Iida R, Tsubota E, Nakamura H, Imamura S, Yasuda T, Takeshita H. Allele frequencies and haplotypes for 28 Y-STRs in Ovambo population, *Leg Med (Tokyo)*. 2009. **11**(4), 205-208.
- Giles RE, Blanc H, Cann HM, Wallace DC. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA*. 1980. **77**(11): 6715-6719.
- Gill P, Ivanov PL, Kimptom C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K. Identification of the remains of the Romanov family by DNA analysis. *Nat Genet*. 1994. **6**: 130-135.
- Gill P, Werrett DJ, Budowle B, Guerrieri R. An assessment of whether SNPs will replace STRs in national DNA data-bases-joint considerations of the DNA working group of the ENFSI and the SWGDAM. *Sci. Justice*. 2004. **44**(1): 51-3.
- Gill P. An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. *Int. J. Legal Med*. 2001. **114**(4-5): 204-10.
- González Ana M, Larruga José M, Abu-Amero Khaled K, Yufei Shi, Pestano J, Cabrera Vicente M. Mitochondrial lineage M1 traces an early human backflow to Africa. *BMC Genomics*. 2007. doi:10.1186/1471-2164-8-223.
- Gusmao L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Tyler-Smith C, Schneider PM. DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, *Forensic Sci. Int*. 2006. **157**; 187-197.

## Références Bibliographiques

---

- Hammer MF. A recent insertion of an alu element on the Y chromosome is a useful marker for human population studies. *Mol. Biol. Evol.* 1994. 11(5): 749-61.
- Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, Georgiev O, Bauer K, Stone A, Hedges R, Schaffner W, Utermann G, Sykes B, Paabo B. Molecular genetic analyses of the Tyrolean Ice Man. *Science.* 1994. 264(5166): 1775-1778.
- Hardy GH. Mendelian proportions in a mixed population. *Science.* 1908. 28: 49-50.
- Harry M. Génétique moléculaire et évolutive. 2001. Editions Maloine, Paris, p326.
- Hassan FA. Archeological explorations of the Siwa Oasis Region, Egypt. *Curr Anthropol.* 1978. 19, 146-148.
- Henri T. Histoire du Maroc. 1949. Tome 1, p104.
- Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc Natl Acad Sci USA.* 1995. 92(2): 532-536.
- Humphrey L, Bello SM, Turner E, Bouzouggar A, Barton N. Iberomaurusian funerary behaviour: Evidence from Grotte des Pigeons, Taforalt, Morocco. *Journal of Human Evolution.* 2012. 62,261-273.
- Ingman M, Kaessmann H, Paabo S, Gyllensten U. Mitochondrial genome variation and the origin of modern humans. *Nature.* 2000. 408(6813): 708-713.
- International Society of Genetic Genealogy. Y-DNA Haplogroup Tree 2010, Version: Version: 4.28 Date: 11 December 2009 , <http://www.isogg.org/tree/>.
- Irwin JA, Jessica LS, Strouss KM, Sturk KA, Toni MD, Just RS, Coble MD, Parson W, Parsons TJ. Development and expansion of high quality control region databases to improve forensic mtDNA evidence interpretation. *Forensic Science International: Genetics.* 2007. 154–157.
- Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parson TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat Genet.* 1996. 12: 417 -420.
- Jobling MA, Chris TS. The Human Y chromosome: An Evolutionary Marker Comes of Age. *Nature.* 2003. Reviews/genetics; Volume 4; doi:10.1038/nrg1124.
- Jobling MA, Pandya A, Tyler-Smith C. The Y chromosome in forensic analysis and paternity testing. *Int. J. Legal Med.* 1997. 110(3):118-24.

## Références Bibliographiques

---

- Jobling MA. Y-chromosomal SNP haplotype diversity in forensic analysis. *Forensic Sci. Int.* 2001. 118(2-3):158-62.
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 2008. 18: 830-838.
- Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am. J. Hum. Genet.* 2000. 66(5):1580-8.
- Kayser M, Sajantila A. Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Sci. Int.* 2001. 118(2-3):116-21.
- Ke Y, Su B, Song X, Lu D, Chen L, Li H et al. African origin of modern humans in East Asia: a tale of 12,000 Y chromosomes. *Science.* 2001. 292(5519):1151-3.
- Kéfi R, Stevanovitch A, Bouzaid E, Béraud-Colomb E. Diversité mitochondriale de la population de Taforalt (12.000 ans BP – Maroc): une approche génétique à l'étude du peuplement de l'Afrique du Nord. *Anthropologie.* 2005. 43: 1-11.
- Keyser-Tracqui C, Crubezy E, Ludes B. Nuclear and mitochondrial DNA analysis of a 2,000- year-old necropolis in the Egyin Gol Valley of Mongolia. *Am J Hum Genet.* 2003. 73(2): 247-260.
- Kivisild T, Reidla M, Metspalu E, Rosa A, Brehm A, Pennarun E, Parik J, Geberhiwot T, Usanga E, Villems R. Ethiopian mitochondrial DNA heritage: tracing gene flow across and around the Gate of Tears. *Am J Hum Genet.* 2004. 75(5): 752-770.
- Kovatsi L, Saunier JL, Irwin JA. Population genetics of Y-chromosome STRs in a population of Northern Greeks. *Forensic Sci. Int. Genet.* 2009. doi: 10.1016/j.fsigen.2009.01.001.
- Koyama H, Iwasa M, Tsuchimochi T, Maeno Y, Isobe I, Matsumoto T, Nagao M. Utility of Y-STR haplotype and mtDNA sequence in personal identification of human remains. *Am. J. Forensic Med. Pathol.* 2002. 23(2):181-5.
- Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Paabo S. Neandertal DNA sequences and the origin of modern humans. *Cell.* 1997. 90(1): 19-30.
- Lacan M, Keyser C, Ricaut FX, Brucato N. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *PNAS.* 2011. doi:10.1073/pnas.1113061108, PMID 22042855.

- Lahn BT, Page DC. Four evolutionary strata on the human X chromosome. *Science*. 1999. 286:964–7.
- Lincoln P, Carracedo A. Publication of population data of human polymorphism. *Forensic Sci. Int.* 2000. 110; 3–5.
- Lubbock J. *Prehistoric Times*. 1865. Londres, Williams and Norgate.
- Maca-Meyer N, Gonzalez AM, Larruga JM, Flores C, Cabrera VM. Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet.* 2001. 2: 13.
- Maca-Meyer N, González AM, Pestano J, Flores C, Larruga JM, Cabrera VM. Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. *BMC Genetics*. 2003. 4: 15. doi:10.1186/1471-2156-4-15. PMC 270091. PMID 1456321.
- Marth GT, Czubarka E, Murvai J, Sherry ST. The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. *Genetics*. 2004.166(1):351-72.
- Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA*. 1977. 74: 560-564.
- Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G, Attardi G. Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. *Science*. 1999. 286(5440): 774-779.
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD, Sukernik RI, Olckers A, Wallace DC. Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA*. 2003. 100(1): 171-176.
- Mitchell RJ, Hammer MF. Human evolution and the Y chromosome. *Curr Opin Genet Dev*. 1996. 6 : 737-42.
- Mohamed Kably et collaborateurs. *Histoire du Maroc – Réactualisation et synthèse*. Edition de l'institut Royal pour la recherche sur l'histoire du Maroc, Rabat. 2011. p 54-60.
- Nielsen R. Population genetic analysis of ascertained SNP data. *Hum. Genomic*. 2004. 1(3): 218-24.
- Omran GA, Rutty GN, Jobling MA Diversity of 17-locus Y-STR haplotypes in Upper (Southern) Egyptians, *Forensic Sci. Int. Gene. Suppl.* 2008. 230–232.
- Onofri V, Alessandrini F, Turchi C, Pesaresi M, Tagliabracci A. Y-chromosome markers distribution in Northern Africa: high-resolution SNP and STR analysis in Tunisia and Morocco populations, *Forensic Sci. Int. Gene. Suppl.* 2008.235–236.

## Références Bibliographiques

---

- Orlando L, Darlu P, Toussaint M, Bonjean D, Otte M, Hanni C. Revisiting Neandertal diversity with a 100,000 year old mtDNA sequence. *Curr Biol*. 2006. 16(11): R400-402.
- Ovchinnikov IV, Gotherstrom A, Romanova GP, Kharitonov VM, Liden K, Goodwin W. Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature*. 2000. 404(6777): 490-493.
- Parson W, Niederstätter H, Brandstätter A, Berger B. Improved specificity of Y-STR typing in DNA mixture samples. *Int. J. Legal Med*. 2003. 117:109–114.
- Pavel LI, Mark JW, Rhonda KR, Mitchell MH, Victor W, Thomas J. Parsons. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nature Genetics*. 1996. 12, 417 – 420 doi: 10.1038/ng0496-417.
- Pellegrini B. L'Ève imaginaire. Les origines de l'homme, de la biologie à la paléontologie. 1995. Payot-rivages (Ed), Paris, p228.
- Pierre-Henri G, Henry PJ, Masson. Précis de Génétique des populations, 1998.
- Plaza S, Salas A, Calafell F, Corte-Real F, Bertranpetit J, Carracedo A, Comas D. Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. *Hum Genet*. 2004. 115(5): 439-447.
- Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nat Genet*. 1999. 23(4): 437-441.
- Rando JC, Pinto F, Gonzalez AM, Hernandez M, Larruga JM, Cabrera VM, Bandelt HJ. Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European near-eastern and sub-Saharan populations. *Ann Hum Genet*. 1998. 62(Pt 6): 531-550.
- Ravela C, Chantot-Bastaroud S, McElreavey K, Siffroib JP. Molecular anomalies of the Y chromosome: consequences on male fertility. *Gynécologie Obstétrique & Fertilité*. 2006. 34 885–893.
- Raynal JP, Sbihi Alaoui FZ, Mohib A. Bilan des recherches récentes sur le Paléolithique de Casablanca (Maroc), *Les Nouvelles de l'archéologie*. 2010. n° 120-121, p. 102-109.
- Raynal JP. Taforalt. Mission Préhistorique et paléontologique française au Maroc: rapport d'activité pour l'année. *Bulletin d'Archéologie Marocaine*. 1980. 12: 69-71.
- Ricaud FX, Fedoseeva A, Keyser-Tracqui C, Crubezy E, Ludes B. Ancient DNA analysis of human neolithic remains found in northeastern Siberia. *Am J Phys Anthropol*. 2005. 126(4): 458-462.

## Références Bibliographiques

---

- Robino C, Crobu F, Di Gaetano C, Bekada A, Benhamamouch S, Cerutti N, Piazza A, Inturri S, Torre C. Analysis of Y-chromosomal SNP haplogroups and STR haplotypes in an Algerian population sample. *Int. J. Legal Med.* 2008. 122 (3); 251–255.
- Roche J. Cadre chronologique de l'Épipaléolithique marocain. Actes du IX<sup>e</sup> Congrès de l'UISPP : Chronologie et synchronisme dans la préhistoire circum-méditerranéenne. 1976. pp. 153-67.
- Roche J. L'Épipaléolithique Marocaine. Fondation Calouste Gulbenkian, Lisbon. 1963.
- Roche J. La grotte de Taforalt. *Anthropologie.* 1953. 57: 375-380.
- Roche J. L'Aterien de la grotte de Taforalt (Maroc oriental). *Bulletin d'Archéologie Marocaine.* 1967. 7: 11-56.
- Roche J. Les industries paléolithiques de la grotte de Taforalt (Maroc oriental). *Quaternaria.* 1969. 11: 89-100.
- Roewer L, Kayser M, de Knijff P, Anslinger K, Betz A, Caglia A et al. A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Sci. Int.* 2000. 114(1):31-43.
- Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A et al. Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. *Forensic Sci. Int.* 2001. 118(2-3):106-13.
- Roewer L, Willuweit S, Stoneking M, Nasidze I. A Y-STR database of Iranian and Azerbaijani minority populations. *Forensic Science International: Genetics.* 2009. 4(1), e53-e55.
- Rolf B, Keil W, Brinkmann B, Roewer L, Fimmers R. Paternity testing using Y-STR haplotypes: assigning a probability for paternity in cases of mutations. *Int. J. Legal Med.* 2001. 115(1):12-5.
- Rollo F, Ermini L, Luciani S, Marota I, Olivieri C, Luiselli D. Fine characterization of the Iceman's mtDNA haplogroup. *Am J Phys Anthropol.* 2006. 130(4): 557-564.
- Salas A, Richards M, De la Fe T, Lareu MV, Sobrino B, Sanchez-Diz P, Macaulay V, Carracedo A. The making of the African mtDNA landscape. *Am J Hum Genet.* 2002. 71(5): 1082-1111.
- Samuel P, Françoise P, Philippe P. Maroc : les Cités Impériales. 1995. P ???
- Santos FR, Epplen JT, Pena SD. Testing deficiency paternity cases with a Y-linked tetranucleotide repeat polymorphism. *EXS.* 1993. 67:261-5.

## Références Bibliographiques

---

- Saxon EC, Close A, Cluzel C, Morse V, Shackleton NJ. Results of recent excavations at Tamar Hat. Libyca . 1974. 22: 49-91.
- Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery Nadia, Battaglia V, Mac-Cioni L, Triantaphyllidis C. Origin, Diffusion, and Differentiation of Y-Chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area. The American Journal of human genetics. 2004. 74 (5): 1023-34.DOI:101086. PMC 1181965. PMID 15069642.
- Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Menecier P, Hofreiter M, Possnert G, Pääbo S. No evidence of Neandertal mtDNA contribution to early modern humans. PLoS Biol. 2004. 2(3): 313-317.
- Sibille I, Duverneuil C, Lorin de la Grandmaison G, Guerrouache K, Teissiere F, Durigon M, de Mazancourt P. Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. Forensic Sci. Int. 2002. 125(2-3):212-6.
- Southern EM. Measurement of DNA length by gel electrophoresis. Anal Biochem. 1979. 100: 319-323.
- Spindler K. The man in the ice. 1995. London: Phoenix.
- Stone AC, Starrs JE, Stoneking M. Mitochondrial DNA analysis of the presumptive remains of Jess James. J Forensic Sci. 2001. 46: 173 -176.
- Stoneking M. DNA and recent human evolution. Evol Anthropol. 1993. 2: 60-73.
- Sykes B, Irven C. Surnames and the Y chromosome. Am. J. Hum. Genet. 2000. 66(4):1417-9.
- Thèves C, Keyser-Tracqui C, Crubezy E, Salles JP, Ludes B, Telmon N. Detection and quantification of the age-related point mutation A189G in the human mitochondrial DNA. J Forensic Sci. 2006. 51(4): 865-873.
- Tishkoff SA, Williams SM. Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet. 2002. 3:611–621.
- Torrioni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. Harvesting the fruit of the human mtDNA tree. Trends Genet. 2006. 22(6): 339-345.
- Torrioni A, Lott MT, Cabell MF, Chen YC, Lavergne L, Wallace DC. Mt DNA and the origin of the Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. Am. J. Hum. Genet. 1994. 55, 760–776.

- Torrioni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC. mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am J Hum Genet.* 1994b. 55(4): 760-776.
- Torrioni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, Wallace DC. Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. *Am J Phys Anthropol.* 1994c. 93(2): 189-199.
- Torrioni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet.* 1993. 53(3): 563-590.
- Trombetta B, Cruciani F, Sellitto D, Scozzari R, MacAulay V. A New Topology of the Human Y Chromosome Haplogroup E1b1 (E-P2) Revealed through the Use of Newly Characterized Binary Polymorphisms. *PLoS.* 2011. (1): e16073, doi:10.1371/journal.pone.0016073, PMC 3017091, PMID 21253605.
- Trumme T, Herrmann B, Hummel S. Genetics in genealogical research--reconstruction of a family tree by means of Y-haplotyping. *Anthropol. Anz.* 2004. 62(4):379-86.
- Underhill PA, Bellwood, Renfrew. *Inference of Neolithic Population Histories using Y-chromosome Haplotypes*, Cambridge: McDonald Institute for Archaeological Research. 2002. ISBN 1-902937-20-1.
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA et al. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.* 2001. 65(Pt 1):43-62.
- Underhill, Kivisild T. Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. *Annu.Rev.Genet.* 2007. 41:539-64, DOI:10.1146/annurev.genet., PMID 18076332.
- Vallone PM, Butler JM. AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques.* 2004. 37(2): 226-231.
- Van OM, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30(2):E386-E394. <http://www.phylotree.org>.doi10.1002.
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. *Science.* 1991. 253(5027): 1503-1507.
- Wallace DC, Brown MD, Lott MT. Mitochondrial DNA variation in human evolution and disease. *Gene.* 1999. 238(1): 211-230.

## Références Bibliographiques

---

- Wallace DC, Ye JH, Neckelmann SN, Singh G, Webster KA, Greenberg BD. Sequence analysis of cDNAs for the human and bovine ATP synthase beta subunit: mitochondrial DNA genes sustain seventeen times more mutations. *Curr Genet.* 1987. 12(2): 81-90.
- Ward S. Reconstructing Late Pleistocene Vegetation Dynamics from Archaeological Cave Sites in the Western Mediterranean: Links with Climate and Cultural Changes. DPhil Doctoral Thesis. 2007. University of Oxford.
- Watson E, Forster P, Richards M, Bandelt HJ. Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet.* 1997. 61(3): 691-704.
- Watson JD, Crick FH. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature.* 1953. 171(4356): 737-738.
- Weinberg W. Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg.* 1908. 64: 368-382.
- Wengler L. Formations quaternaires et cultures préhistoriques au Maroc oriental. Thèse de Doctorat d'État à l'Université Bordeaux I. 1993.
- Willard HF. Tales of the Y-chromosome. *Nature.* 2003. 423: 810-812.
- Y chromosome consortium. A nomenclature system for the three of human Y-chromosomal binary haplogroups. 2002. 339-348.
- Y Chromosome, Haplotype Reference Database, [www.yhrd.org](http://www.yhrd.org).
- Zhao Z, Fu YX, Hewett-Emmett D, Boerwinkle E. Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution. *Gene.* 2003. 312:207-13.
- Zhu B, Wu Y, Shen C, Yang T, Deng Y, Xun X, Tian Y, Yan J, Li T. Genetic analysis of 17 Y-chromosomal STRs haplotypes of Chinese Tibetan ethnic group residing in Qinghai province of China. *Forensic Sci. Int.* 2008. 175 (2-3); 238-243.

## Articles

Research article

# STR data for the 15 AmpFISTR Identifiler loci in the Moroccan population

M. Bouabdellah<sup>a,b,c,\*</sup>, F. Ouenzar<sup>a,b</sup>, R. Aboukhalid<sup>a,b</sup>, M. Elmzibri<sup>b</sup>,  
D. Squalli<sup>c</sup>, S. Amzazi<sup>a</sup>

<sup>a</sup> *Biochemistry and Immunology laboratory, Science Faculty, Mohamed V-Agdal University, Rabat, Morocco*

<sup>b</sup> *Biology and Medical Research laboratory, CNESTEN, Rabat, Morocco*

<sup>c</sup> *GENOME Biotechnologies, Casablanca, Morocco*

Received 16 August 2007; accepted 7 October 2007

## Abstract

Allele frequencies for the 15 STR loci included in the AmpFISTR<sup>®</sup> Identifiler<sup>™</sup> PCR Amplification Kit panel from Applied Biosystems (D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338, D16S539) and several statistical parameters were estimated from a sample of 425 unrelated individuals representing all ethnic groups in Morocco (Berbers, Arabs and Sahraouis). DNA was extracted from blood samples, by the DNA IQ<sup>™</sup> System and the PCR products were typed using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The results were analyzed using GeneMapper ID v3.2 software (Applied Biosystems) and the values of heterozygosity, polymorphic information content (PIC), power of discrimination (PD), power of exclusion (PE), Typical Paternity Index (TPI) and matching probability (MP) were calculated.

© 2008 Elsevier Ireland Ltd. All rights reserved.

**Keywords:** Moroccan population; Ethnic groups; STR DNA typing; Population data; Allele frequencies

## 1. Population

Blood samples were collected from unrelated healthy donors from the Regional Center of Blood Transfusion in Morocco. The donors were subjected to an interview in order to make sure of their ethnic origin and to obtain their informed consent.

## 2. DNA extraction

DNA from blood samples was extracted using DNA IQ<sup>™</sup> System (Promega, Madison, USA) according to manufacturer's instructions.

## 3. PCR

Simultaneous amplifications of 16 STR loci (multiplexed PCR) were performed by using the AmpFISTR<sup>®</sup> Identifiler<sup>™</sup> PCR Amplification Kit according to the user's manual recommendations [1]. The 16 loci amplified in this study are

D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, D19S433, D2S1338, D16S539 and the gender determination marker, Amelogenin.

## 4. Typing

The separation and detection of amplified products were conducted with the ABI Prism<sup>®</sup> 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) following manufacturer's protocols. Data collection was performed with Data Collection v. 2.0 software (Applied Biosystems, Foster City, CA, USA) and samples were analyzed by GeneMapper ID v. 3.2 software (Applied Biosystems, Foster City, CA, USA).

## 5. Quality control

Laboratory internal control standards and kit controls.

## 6. Statistical analyses

The frequency of each allele for each locus, the observed heterozygosity (Ho) and the expected heterozygosity (He)

\* Corresponding author at: GENOME Biotechnologies, Parc d'activité Oukacha, Imm B N35, AS, Casablanca, Morocco. Tel. +212 61 09 20 60.

E-mail address: [bou.mehdi@gmail.com](mailto:bou.mehdi@gmail.com) (M. Bouabdellah).

Table 1  
Allele frequencies and statistical parameters of AmpFISTR Identifier PCR Amplification Kit loci in Moroccan population

	D3S1358	vWA	FGA	TH01	TPOX	CSF1PO	D5S818	D13S317	D7S820	D8S1179	D21S11	D18S51	D16S539	D2S1338	D19S433
5								0.001							
5.3				0.004											
6		0.001		0.185	0.015										
7		0.001		0.233	0.032	0.015		0.001	0.005						
8		0.019		0.192	0.409	0.029	0.048	0.076	0.15	0.001		0.001	0.0033		
9		0.003		0.25	0.194	0.024	0.037	0.049	0.093	0.001			0.105		
9.3		0.003		0.136	0.077										
10	0.001	0.003			0.274	0.307	0.063	0.049	0.338	0.097	0.001	0.008	0.064		0.001
11		0.009				0.34	0.275	0.326	0.268	0.141		0.019	0.295	0.004	0.01
12	0.001	0.001				0.284	0.391	0.363	0.146	0.122	0.003	0.152	0.301		0.137
13	0.001	0.009					0.185	0.135		0.206	0.001	0.122	0.202	0.004	0.249
13.2											0.003				0.043
14	0.071	0.128								0.235		0.137			0.271
14.2															0.071
15	0.284	0.193								0.196		0.15		0.003	0.148
16	0.257	0.23									0.001	0.171		0.048	0.048
17	0.246	0.23	0.001									0.134		0.282	0.015
18	0.139	0.15	0.003									0.075		0.099	0.003
18.2			0.001												
19		0.02	0.046								0.001			0.138	
20			0.14								0.003	0.031		0.17	0.003
21			0.161								0.001			0.06	
22			0.2												
23			0.211											0.09	
23.2											0.001				
24			0.14											0.09	
25			0.097												
26														0.012	
27											0.027				
28											0.122				
28.2															
29											0.246				
29.2															
30											0.264				
30.2															
31											0.085				
31.2											0.103				
32															
32.2											0.121				
33															
33.2															
34															
34.2											0.004				
35											0.011				
MP	0.095	0.065	0.046	0.08	0.133	0.147	0.114	0.113	0.092	0.058	0.058	0.033	0.1	0.049	0.065
TPI	2.07	2.25	2.05	2.36	1.55	1.68	1.82	1.8	1.75	2.35	2.53	2.51	2.08	2.82	2.91
Ho	0.766	0.8352	0.7862	0.7896	0.7284	0.7226	0.7382	0.7441	0.7429	0.8042	0.8184	0.8175	0.7759	0.8412	0.8113
He	0.7777	0.7995	0.8595	0.824	0.7046	0.7346	0.7413	0.765	0.7869	0.833	0.8465	0.8338	0.7806	0.8602	0.8096
PD	0.905	0.935	0.954	0.92	0.867	0.853	0.886	0.887	0.908	0.942	0.942	0.967	0.9	0.952	0.935
PIC	0.73	0.79	0.82	0.76	0.67	0.65	0.69	0.69	0.72	0.79	0.8	0.85	0.73	0.82	0.79
PE	0.523	0.559	0.519	0.578	0.395	0.432	0.469	0.464	0.449	0.575	0.603	0.6	0.526	0.642	0.652

Ho, observed heterozygosity; He, expected heterozygosity; TPI, paternity index; MP, matching probability; PD, power of discrimination; PIC, polymorphic information content.

were assessed by GENETIX (4.05) software [2]. Forensic parameters (polymorphism information content, PIC; power of discrimination, PD; and power of exclusion, PE) were calculated using Microsoft Excel-PowerStats program [3] downloaded from Promega corporation website at <http://www.promega.com/geneticidtools/powerstats/>.

## 7. Analysis of results

The observed heterozygosity (Ho) ranges from 0.7226 (CSF1PO) to 0.8142 (D2S1338). The power of discrimination varies between 0.853 (CSF1PO) and 0.967 (D18S51), and the combined power of discrimination for the 15 STR loci is

0.99999999. The probability of excluding paternity (PE) varies between 0.395 (TPOX) and 0.652 (D19S433), and the combined probability of excluding paternity for the 15 loci is 0.9999.

Present work of the Moroccan population was conducted in order to complete the STR profiling of the different ethnic groups of the country. Previous studies were published on the Arab and Berber speaking groups [4–10] but this is the first large work that encloses both groups plus a small Sahraoui pool. All statistical values reported in this work showed no significant difference from those reported precedently (Table 1).

### Acknowledgements

This work was funded by GENOME Biotechnologies, the exclusive distributor of Applied Biosystems products in Morocco. The authors acknowledge the blood donors, the technicians of the Regional Center of Blood Transfusion of Rabat and anyone who contribute to this work.

### Funding source

This work was sponsored by GENOME Biotechnologies, Applied Biosystems products distributor in Morocco. The sponsor had no involvement in the development of the paper or decisions related to the paper.

### Conflict of interest

None.

### References

- [1] AmpFISTR Identifier™ User's Manual, Applied Biosystems, USA, 2001.
- [2] K. Belkhir, P. Borsa, L. Chikhi, N. Raufaste, F. Bonhomme, 1996–2004 GENETIX 4.05, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
- [3] A. Tereba, Tools for analysis of population statistics, Profiles DNA 2 (1999) 14–16.
- [4] A. Perez-Lezaun, F. Calafell, J. Clarimon, E. Bosch, E. Mateu, L. Gusmao, A. Amorim, N. Benchemsi, J. Bertranpetit, Allele frequencies of 13 short tandem repeats in population samples from the Iberian Peninsula and northern Africa, Int. J. Legal Med. 113 (2000) 208–214.
- [5] E. Bosch, J. Clarimon, A. Perez-Lezaun, F. Calafell, STR data for 21 loci in northwestern Africa, Forensic Sci. Int. 116 (2001) 41–51.
- [6] S. Dios, J.R. Luis, J.C. Carril, B. Caeiro, Sub-Saharan genetic contribution in Morocco: microsatellite DNA analysis, Hum. Biol. 73 (2001) 675–688.
- [7] L. Abdin, I. Shimada, B. Brinkmann, C. Hohoff, Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria, Legal Med. 5 (2003) S150–S155.
- [8] A. Jauffrit, H. El Amri, F. Airaud, M.T. Andre, O. Herbert, G. Landeau-Trottier, S. Giraudet, C. Richard, A. Chaventre, J.P. Moisan, DNA short tandem repeat profiling of Morocco, J. Forensic Sci. 48 (2003), pp. 458–159.
- [9] F. Chbel, S. Nadifi, C. Martinez-Bouzas, S. Louahlia, H. Azeddoug, M. Martinez De Pancorbo, Population genetic data of eight tetrameric short tandem repeats (STRs) in Casablanca resident population to use in forensic casework, Forensic Sci. Int. 132 (2003) 82–83.
- [10] C. Coudray, E. Guitard, C. Keyser-Tracqui, M. Melhaoui, M. Cherkaoui, G. Larrouy, J.-M. Dugoujon, Population genetic data of 15 tetrameric short tandem repeats (STRs) in Berbers from Morocco, Forensic Sci. Int. 167 (2007) 81–86.



Contents lists available at ScienceDirect

Legal Medicine

journal homepage: [www.elsevier.com/locate/legalmed](http://www.elsevier.com/locate/legalmed)

## Announcement of Population Data

## Assessment of phylogenetic structure of Berber-speaking population of Azrou using 15 STRs of Identifiler kit

Hicham El Ossmani<sup>a,\*</sup>, Brahim Bouchrif<sup>b</sup>, Rachid Aboukhalid<sup>c</sup>, Mehdi Bouabdillah<sup>c</sup>, Bouchaïb Gazzaz<sup>d</sup>, Driss Zaoui<sup>a</sup>, Abdelaziz Chafik<sup>a</sup>, Jalal Talbi<sup>a</sup><sup>a</sup>Laboratory of Anthropogenetics and Physiopathology, Department of Biology, University of Chouaïb Doukkali, Faculty of Sciences, El Jadida, Morocco<sup>b</sup>Immunology and Biochemistry Laboratory (Forensic Genetic Unit), Mohammed V Agdal University, Faculty of Sciences, Rabat, Morocco<sup>c</sup>Laboratory of Molecular Biology, Institute of Pasteur, Luis Pasteur Place 1, Casablanca, Morocco<sup>d</sup>Laboratory of Genetics and Molecular Pathology, Faculty of Medicine, Casablanca, Morocco

## ARTICLE INFO

## Article history:

Received 29 April 2009

Received in revised form 12 August 2009

Accepted 16 October 2009

Available online xxxx

## Keywords:

Berber-speaking

Azrou

STRs

Allele frequencies

Phylogenetic tree

## ABSTRACT

Allele frequencies for 15 STR autosomal loci of Identifiler kit (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818 and FGA) in the Moroccan population of Berber-speaking of Azrou, were assessed from a sample of 201 unrelated individuals. Markers D18S51, D2S1338, FGA and D21S11 present the highest power of discrimination (PD) values while D21S11 was the most polymorphic locus in the studied population. The phylogenetic tree established among worldwide populations, shows that Berber-speaking population of Azrou was so close to the Berber-speaking population of Asni but also to the Arab-speaking population of southern Morocco. Nevertheless, a significant distance was observed between populations of Azrou and Bouhria even they share the same dialect (Amazigh) and belong to the same geographical area (Morocco). The 15 STR loci studied appear to be highly discriminating, thus providing a powerful tool for forensic applications, paternity investigation, individual identification and anthropological studies.

© 2009 Elsevier Ireland Ltd. All rights reserved.

**Introduction:** Some Moroccan populations have been characterized in term of allele frequencies of 15 autosomal short tandem repeats (STRs) [1–3]. Their forensic parameters and their phylogenetic position among worldwide populations have been assessed. Nevertheless, many other Moroccan populations are not yet characterized, so their forensic parameters and phylogenetic situation still ignored. Berber-speaking of Azrou is one of these populations. In fact, data about Berber-speaking populations may help in forensic and anthropological investigations. On the other hand, only language and some cultural traits permit to distinguish Berbers from Arabs. Then, genetic characterization comes to give great answers about biological identity. In Morocco, about 40% acknowledge a Berber identity, though many more have Berber ancestry.

Historically, Berbers lived in North Africa long before the arrival of the Arabs, and their culture probably dates back more than 4000 years. Berber states known as Mauritania and Numidia ex-

isted in classical times. They comprise a clear majority of the population of North Africa in terms of ethnicity, but in terms of identity, a considerable minority. It is essential to understand this difference between ethnicity and identity in order to grasp the meaning of being Berber. The influx of Arabs in North Africa has been too insignificant throughout history to justify those large numbers of people now claiming to be Arabs. And the influx of other peoples in North Africa has not been of any significance since the Vandals in the 5th century.

In this study, we used 15 STR of Identifiler kit to provide their allelic distribution and their forensic parameters in the Berber-speaking population of Azrou (Morocco). These markers will help to situate genetically the studied population among the local and worldwide populations.

**Linguistics parameters:** Berber belongs to the Afro-Asiatic language family, and has many dialects. The three main dialects used in Morocco are Tachelhit, Tamazight and Tarifit. Collectively, they are known as “shelha” in Arabic. The populations of Asni, Azrou and Bouhria speak, respectively, these three dialects and belong respectively to Berbers of Atlas, Souss and Rif.

\* Corresponding author. Mobile: +212 679827887.

E-mail address: [helossmani@yahoo.fr](mailto:helossmani@yahoo.fr) (H. El Ossmani).

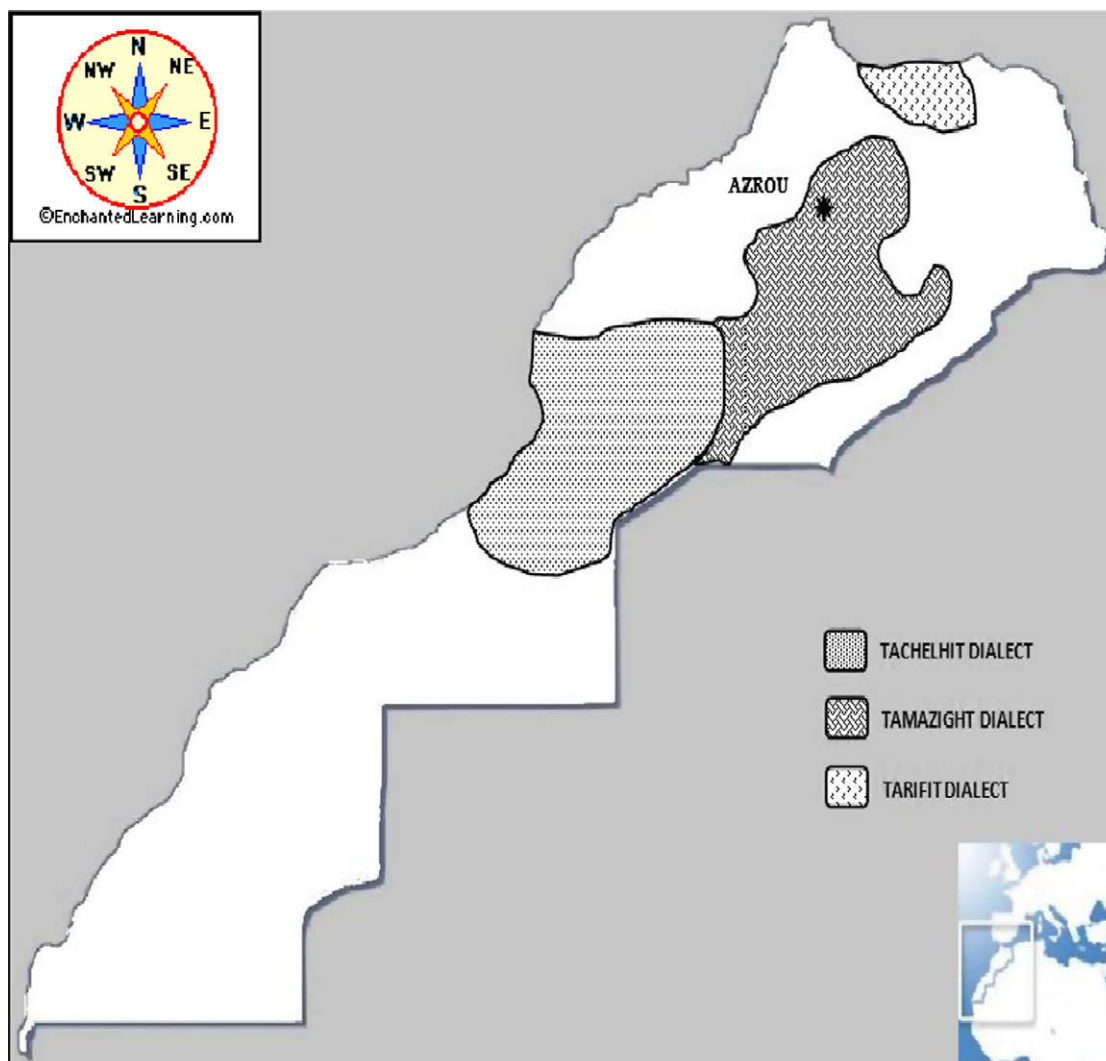


Fig. 1. Repartition of ethnolinguistic groups in Morocco.

Fig. 1 shows the geographical distribution of the different ethnolinguistic groups in Morocco.

- *Tachelhit* (sometimes known as “soussi” or “chleuh”) is spoken in south-west Morocco, in an area between Ifni in the south, Agadir in the north and Marrakech and the Draa/Sous valleys in the east.
- *Tamazight* (or Berber, Amazigh, Zaian) is spoken in the Middle Atlas Azrou, between Taza, Khemisset, Azilal and Errachidia.
- *Tarifit* (or Rifia) is spoken in the Rif area of northern Morocco. This dialect continues along the Algerian border in eastern Morocco all the way to Figuig.

**Population information:** Blood samples were obtained from 201 unrelated healthy individuals of Berber-speaking population of Azrou in Morocco. Their ancestry was traced back at least two generations.

**DNA extraction:** Genomic DNA was extracted using the standard Phenol–Chloroform method [4] and purified by ethanol precipitation.

**PCR:** PCR amplification was performed for 15 STR loci using AmpFI STR® Identifiler™ kit (Applied Biosystems, Foster City, CA, USA) according to manufacturer’s specifications [5].

**Typing:** The PCR products were detected with the ABI PRISM® 3130xl Genetic Analyzer (Applied Biosystems) and sized with GeneScan-500 LIZ internal lane size standard. Allelic calls and genotyping were carried out by comparison to the reference allelic ladder included in the kit, using GeneMapper ID v3.2 (Applied Biosystems).

**Quality control:** Identifiler™ kit included allelic ladder and positive control.

**Statistical and phylogenetic analysis:** For data analysis the Arlequin software package version 3.1 [6] was used to assess Hardy–Weinberg equilibrium (HWE) using Fisher’s exact test [7]. Several forensic parameters were also examined including power of discrimination (PD), polymorphic information content (PIC) and power of exclusion (PE) using the PowerStats program v1.2 [8,9]. STR allele frequencies of 14 worldwide populations previously published [2,3], were introduced in the analysis to build phylogenetic tree using Neighbor-Joining (NJ) program within PHYLIP 3.67 software [10] (Table 1). Phylogenetic tree was built using MEGA v4 [22].

**Access to data:** Complete data can be acquired upon request to helossmani@yahoo.fr.

**Results:** Refer to Table 1 and Fig. 2.

**Other remarks:** Table 2 shows that all markers have high PD values (>0.865). High values of PD and PIC were observed for D18S51, D2S1338, FGA and D21S11 with the highest ones in mar-

**Table 1**

The worldwide populations introduced in the study.

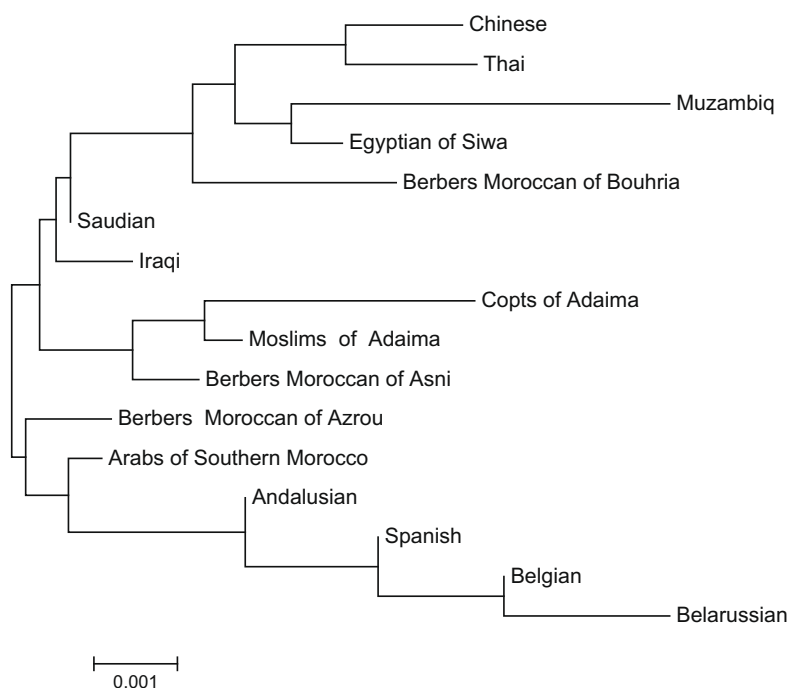
Population		Number
Northern Africa	Arabs of Southern Morocco [2]	204
	Berbers of Bouhria «Morocco» [3]	104
	Berbers of Asni «Morocco» [3]	105
	Berbers of Siwa «Egypt» [11]	98
	Muslims of Adaima «Egypt» [11]	99
	Copts of Adaima «Egypt» [11]	100
Sub-Saharan Africa	Mozambiq [12]	108–126
Middle East	Saoudian [13]	94
	Iraqi [14]	103
Eastern Asia	Chinese [16]	200
	Thai [17]	210
Europe	Belgian [21]	100
	Andalusian [18]	114
	Natives of Spain [19]	342
	Belarussian [20]	176

ker D18S51. The lowest value of PD was observed for CSF1PO while both of markers D5S818 and TPOX have the lowest value of PIC. D16S539 and vWA were the most powerful markers for paternity testing with the highest values of PE and TPI but, also, the highest value of heterozygosity. The most polymorphic marker was D21S11 with 18 alleles. Deviation from Hardy–Weinberg equilibrium was observed after Benferroni's correction in, only, one marker (TH01).

**Discussion:** Berber-speaking population of Azrou is approximately in equilibrium. Only one of the 15 studied markers (TH01) doesn't present a panmictic distribution. Molecular definition of this marker could explain its deviation. In fact, as we showed in our previous works, an unusual rate of mutations as well as possible genetic drift could be likely to change the allelic distribution in some loci but not in others [1,2]. Matrimonial behavior in this population could be, also, responsible of the observed disequilibrium.

As shown in phylogenetic tree and confirmed by genetic distance in Table 3, population of Azrou from Morocco seems to be nearest to the Berber-speaking population of Asni and also to the Arab-speaking population of southern Morocco. This traduces some continuity between the Berber's genetic pool and the Arabs one. Indeed even they may have different origins, Berbers and Arabs lived in the same geographical area for long centuries, which may be sufficient to make them closer. In a previous work Shepard and Herrera [15] found that Andalusia and Arabs from Morocco bifurcate together away from the Berber population with a bootstrap value equal to 76% which corroborates our result.

By comparison to Berber-speaking of Bouhria (Morocco) and those of Egypt, Berber-speaking of Azrou and those of Asni were distant. Actually, while Berber-speaking of Azrou and those of Asni were in the same cluster with Arabs and somewhat nearest to Europeans, Berber-speaking of Bouhria (Morocco) and those of Egypt were closer to Asians and populations of Middle East. This

**Fig. 2.** NJ-phylogenetic tree of 15 STRs used in Identifiler kit.

**Table 2**  
Allele frequencies and medico-legal parameters of 15 STRs in the Berber-speaking population of Azrou ( $n = 201$ ).

Allele	D5S818	FGA	D8S1179	D21S11	D7S820	CSF1PO	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D3S1358
5.3	–	–	–	–	–	–	0.010	–	–	–	–	–	–	–	–
6	–	–	–	–	–	–	0.152	–	–	–	–	–	0.015	–	–
6.3	–	–	–	–	–	–	0.010	–	–	–	–	–	–	–	–
7	–	–	–	–	–	0.015	0.225	–	–	–	–	–	0.020	–	–
8	0.034	–	0.015	–	0.167	0.025	0.181	0.069	0.015	–	–	0.005	0.387	0.005	–
8.3	–	–	–	–	–	–	0.010	–	–	–	–	–	–	–	–
9	0.025	–	–	–	0.078	0.015	0.245	0.054	0.113	–	–	0.005	0.196	–	–
9.2	–	–	–	–	0.010	–	–	–	–	–	–	–	–	–	–
9.3	–	–	–	–	–	–	0.113	–	–	–	–	–	–	–	–
10	0.034	–	0.098	0.005	0.328	0.294	0.043	0.039	0.059	–	0.005	–	0.083	0.005	–
11	0.328	–	0.132	–	0.250	0.299	–	0.294	0.275	–	0.010	0.010	0.289	0.005	–
12	0.368	–	0.093	0.005	0.137	0.309	–	0.348	0.348	–	0.132	–	0.010	0.176	–
12.2	–	–	–	–	–	–	–	–	–	–	0.005	–	–	–	–
13	0.191	–	0.211	–	0.029	0.039	–	0.123	0.186	–	0.206	0.010	–	0.078	–
13.2	–	–	–	–	–	–	–	–	–	–	0.054	–	–	–	–
14	–	–	0.191	–	0.003	0.005	–	0.068	–	–	0.348	0.113	–	0.142	0.064
14.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
15	0.005	–	0.211	–	–	–	–	0.005	0.005	–	0.132	0.162	–	0.152	0.265
15.2	–	–	–	–	–	–	–	–	–	–	0.044	–	–	–	–
16	–	–	0.093	–	–	–	0.015	–	–	0.025	0.039	0.260	–	0.123	0.289
16.2	–	–	–	–	–	–	–	–	–	–	0.005	–	–	–	–
17	–	0.005	0.010	–	–	–	–	–	–	0.270	0.020	0.206	–	0.127	0.230
17.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
18	0.005	–	–	–	–	–	0.005	0.005	–	0.118	–	0.132	–	0.078	0.137
18.2	–	0.005	–	–	–	–	–	–	–	–	–	–	–	–	–
19	–	0.015	–	0.010	–	–	–	–	–	0.103	–	0.083	–	0.044	0.015
19.2	–	–	–	–	–	–	–	–	–	–	–	–	–	0.005	–
20	–	0.118	–	0.005	–	–	–	–	–	0.162	–	0.015	–	0.025	0.005
20.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
21	–	0.176	–	0.005	–	–	–	–	–	0.093	–	–	–	0.020	–
21.2	–	0.005	–	–	–	–	–	–	–	–	–	–	–	–	–
22	–	0.206	–	–	–	–	–	–	–	0.029	–	–	–	–	–
23	–	0.211	–	–	–	–	–	–	–	0.054	–	–	–	0.015	–
23.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
24	–	0.108	–	–	–	–	–	–	–	0.098	–	–	–	–	–
25	–	0.083	–	–	–	–	–	–	–	0.025	–	–	–	–	–
26	–	0.054	–	–	–	–	–	–	–	0.020	–	–	–	–	–
27	–	0.015	–	0.010	–	–	–	–	–	0.005	–	–	–	–	–
28	–	–	–	0.152	–	–	–	–	–	–	–	–	–	–	–
29	–	–	–	0.211	–	–	–	–	–	–	–	–	–	–	–
30	–	–	–	0.245	–	–	–	–	–	–	–	–	–	–	–
30.2	–	–	–	0.015	–	–	–	–	–	–	–	–	–	–	–
30.4	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
31	–	–	–	0.064	–	–	–	–	–	–	–	–	–	–	–
31.2	–	–	–	0.069	–	–	–	–	–	–	–	–	–	–	–
32	–	–	–	0.010	–	–	–	–	–	–	–	–	–	–	–
32.2	–	–	–	0.123	–	–	–	–	–	–	–	–	–	–	–
33	–	–	–	0.010	–	–	–	–	–	–	–	–	–	–	–
33.2	–	–	–	0.044	–	–	–	–	–	–	–	–	–	–	–
34.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
35	–	–	–	0.010	–	–	–	–	–	–	–	–	–	–	–
35.2	–	–	–	0.010	–	–	–	–	–	–	–	–	–	–	–
$H_0$	0.676	0.784	0.804	0.833	0.735	0.725	0.775	0.794	0.853	0.814	0.794	0.853	0.696	0.804	0.745
$H_e$	0.721	0.850	0.841	0.780	0.730	0.774	0.823	0.767	0.756	0.857	0.798	0.830	0.724	0.883	0.823
$P$	0.370	0.581	0.332	0.474	0.513	0.172	0.000 <sup>c</sup>	0.487	0.093	0.643	0.492	0.287	0.232	0.144	0.350
PD	0.876	0.955	0.948	0.954	0.916	0.865	0.926	0.893	0.876	0.958	0.924	0.939	0.869	0.967	0.906
PIC	0.670	0.830	0.820	0.830	0.740	0.680	0.780	0.730	0.710	0.820	0.770	0.800	0.670	0.870	0.730
PE	0.393	0.570	0.606	0.662	0.485	0.469	0.553	0.588	0.701	0.625	0.588	0.701	0.422	0.606	0.501
TPI	1.550	2.320	2.550	3.000	1.890	1.820	2.220	2.430	3.400	0.680	2.430	3.400	1.650	2.550	1.960

$H_0$ : observed heterozygosity;  $H_e$ : expected heterozygosity;  $P$ : exact test of Hardy–Weinberg equilibrium; PD: power of discrimination; PE: power of exclusion; PIC: polymorphic information content; TPI: typical pattern index; c: Bonferroni's correction.

agrees with the result of our previous work [2]. This corroborates, also, the hypothesis of Myles and colleagues [23] that suggest that contemporary Berber populations possess the genetic signature of a past migration of pastoralists from the Middle East and that they share a dairying origin with Europeans and Asians.

#### Conflict of interest statement

The authors declare to have no financial and personal relationships with other people or organizations that could inappropriately influence this work.

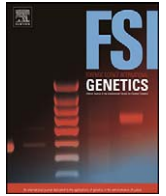
**Table 3**

Nei's genetic distances among the populations.

	Arabs of Southern Morocco (10 <sup>-3</sup> )	Morocco of Asni (10 <sup>-3</sup> )	Morocco of Bouhria (10 <sup>-3</sup> )	Egyptian of Siwa (10 <sup>-3</sup> )	Muslims of Adaima (10 <sup>-3</sup> )	Copts of Adaima (10 <sup>-3</sup> )	Mozambiq (10 <sup>-3</sup> )	Thai (10 <sup>-3</sup> )	Chinese (10 <sup>-3</sup> )	Saudian (10 <sup>-3</sup> )	Iraqi (10 <sup>-3</sup> )	Andalusian (10 <sup>-3</sup> )	Spanish (10 <sup>-3</sup> )	Belarussian (10 <sup>-3</sup> )	Belgian (10 <sup>-3</sup> )
Berber-speaking of Azrou		1.872	6.608	4.375	4.077	8.676	6.248	6.374	7.005	2.388	4.214	2.936	4.505	8.651	6.065
	Arabs of Southern Morocco	2.609	6.387	3.848	4.697	6.990	8.224	5.776	5.962	2.235	3.594	3.049	3.355	6.545	4.988
	Morocco of Asni		5.318	5.937	1.912	5.092	7.967	8.155	6.728	3.205	4.295	4.400	5.916	10.820	7.930
	Morocco of Bouhria			5.340	6.676	9.984	6.802	5.521	6.252	4.904	3.944	7.071	8.219	11.736	9.133
	Egyptian of Siwa				5.838	8.812	5.148	2.819	4.559	3.474	4.782	6.380	7.786	12.477	8.876
	Muslims of Adaima					3.688	8.181	7.967	4.500	2.918	4.767	4.736	7.274	12.353	8.119
	Muslims of Adaima						16.097	10.890	8.938	5.566	5.081	6.612	8.348	12.441	9.565
	Mozambiq							8.476	8.406	5.849	8.676	9.796	13.642	18.147	13.802
	Thai								2.973	4.587	5.821	8.323	8.385	14.087	11.593
	Chinese									3.724	5.514	7.562	9.564	15.732	11.887
Saudian										1.405	1.592	3.653	7.194	4.430	
Iraqi											2.765	5.061	7.579	5.581	
Andalusian												1.808	3.103	1.463	
Spanish													1.626	1.684	
Belarussian														1.303	
Belgian															1.303

**References**

- [1] El Ossmani H, Bouchrif B, Talbi J, El Amri H, Chafik A. La diversité génétique de 15 STR chez la population arabophone de Rabat-Salé-Zemmour-Zaer. *Antropo* 2007;15:55–62. [www.didac.edu.es/antropo](http://www.didac.edu.es/antropo).
- [2] El Ossmani H, Talbi J, Bouchrif B, Chafik A. Allele frequencies of 15 autosomal STR loci in the Southern Morocco population with phylogenetic structure among worldwide populations. *J Leg Med* 2009;11:155–8.
- [3] Coudray C, Guitard E, Keyser-Tracqui C, Melhaoui M, Cherkaoui M, Larrouy G, et al. Population genetic data of 15 tetrameric short tandem repeats (STRs) in Berbers from Morocco. *Forensic Sci Int* 2007;167:81–6.
- [4] Sambrook J, Russell DW. *Molecular cloning: a laboratory manual*. 3rd ed. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press; 2001.
- [5] Applied Biosystems. *Applied biosystems AmpFISTR<sup>®</sup> identifier<sup>™</sup> PCR amplification kit user's manual, instruction for use of products*. Foster City, CA: Applied Biosystems; 2001.
- [6] Excoffier L, Laval G, Schneider S. Arlequin ver. 3.1: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005;1:47–50.
- [7] Guo S, Thompson E. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* 1992;48:361–72.
- [8] Tereba A. Tools for analysis of population statistics, profiles in DNA. *Promega Corporation* 1999;2:14–6.
- [9] Brenner C, Morris J. Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. In: *Proceedings for the international symposium on human identification 1989*. Madison, WI: Promega Corporation; 1990. p. 21–53.
- [10] Felsenstein J. *Phylogeny inference package (PHYLIP) version 3.67*. Seattle, USA: Department of Genetics, University of Washington; 2007.
- [11] Coudray C, Guitard E, El-Chennawi F, Larrouy G, Dugoujon JM. Allele frequencies of 15 short tandem repeats (STRs) in three Egyptian populations of different ethnic groups. *Forensic Sci Int* 2007;169:260–5.
- [12] Alves C, Gusmao L, Damasceno A, Soares B, Amorim A. Contribution for an African autosomic STR database (AmpF/STR Identifier and Powerplex 16 System) and a report on genotypic variations. *Forensic Sci Int* 2004;139:201–5.
- [13] Alshamali F, Alkhatay AQ, Budowle B, Watson ND. STR population diversity in nine ethnic populations living in Dubai. *Forensic Sci Int* 2005;152(2–3):267–79.
- [14] Barni F, Berti A, Pianese A, Boccellino A, Miller MP, Caperna A, et al. Allele frequencies of 15 autosomal STR loci in the Iraq population with comparisons to other populations from the middle-eastern region. *Forensic Sci Int* 2007;167:87–92.
- [15] Shepard EM, Herrera RJ. Iranian STR variation at the fringes of biogeographical demarcation. *Forensic Sci Int* 2006;158:140–8.
- [16] Yang B, Wang G, Liu Y, Yang W. Population data for the AmpF/STR Identifier PCR Amplification Kit in China Han in Jilin Province China. *Forensic Sci Int* 2005;151(2–3):293–7.
- [17] Rerkamnuaychoke B, Rinthachai T, Shotivaranon J, Jomsawat U, Siriboonpiputtana T, Chaiatchanarat K, et al. Thai population data on 15 tetrameric STR loci-D8S1179 D21S11 D7S820 CSF1PO D3S1358 TH01 D13S317 D16S539 D2S1338 D19S433 vWA TPOX D18S51 D5S818 and FGA. *Forensic Sci Int* 2006;158(2–3):234–7.
- [18] Coudray C, Calderon R, Guitard E, Ambrosio B, Gonzalez-Martín A, Dugoujon JM. Allele frequencies of 15 tetrameric short tandem repeats (STRs) in Andalusians from Huelva (Spain). *Forensic Sci Int* 2007;168:21–4.
- [19] Camacho MV, Benito C, Figueiras AM. Allelic frequencies of the 15 STR loci included in the AmpF/STR Identifier<sup>™</sup> PCR Amplification Kit in an autochthonous sample from Spain. *Forensic Sci Int* 2007;173(2–3):241–5.
- [20] Rebała K, Wysocka J, Kapinska E, Cybulska L, Mikulich AI, Tsybovsky IS, et al. Belarusian population genetic database for 15 autosomal STR loci. *Forensic Sci Int* 2007;173(2–3):235–7.
- [21] Decorte R, Engelen M, Larno L, Nelissen K, Gilissen A, Cassiman JJ. Belgian population data for 15 STR loci (AmpF/STR SGM Plus and AmpF/STR profiler PCR amplification kit). *Forensic Sci Int* 2004;139(2–3):211–3.
- [22] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596–9. Available from: <http://www.kumarlab.net/publications>.
- [23] Myles S, Bouzekri N, Haverfield E, Cherkaoui M, Dugoujon JM, Ward R. Genetic evidence in support of a shared Eurasian–North African dairying origin. *Hum Genet* 2005;117:34–42.



## Announcement of Population Data

Haplotype frequencies for 17 Y-STR loci (AmpFISTR<sup>®</sup>Y-filer<sup>™</sup>) in a Moroccan population sample

Rachid Aboukhalid<sup>a,b,\*</sup>, Mehdi Bouabdellah<sup>a,b,c,1</sup>, Meriame Abbassi<sup>a,b,2</sup>, Kaoutar Bentayebi<sup>a,b,3</sup>, Mohammed Elmzibri<sup>b,4</sup>, Driss Squalli<sup>c,5</sup>, Saaïd Amzazi<sup>a,6</sup>

<sup>a</sup> Biochemistry and Immunology Laboratory (Forensic Genetic Unit), Mohammed V Agdal University, Faculty of Sciences, BP 1014, Avenue Ibn Batouta Agdal, Rabat, Morocco

<sup>b</sup> Biology and Medical Research Unit, Centre National de l'Energie des Sciences et Techniques Nucléaires (CNESTEN), BP 1382, R.P. 10001, Rabat, Morocco

<sup>c</sup> GENOME Biotechnologies Parc d'activités Oukacha imm. B, N° 35 Ain Sebaâ, Casablanca, Morocco

## ARTICLE INFO

## Article history:

Received 15 August 2008

Received in revised form 1 June 2009

Accepted 8 June 2009

## Keywords:

Y-Chromosomal STR

Haplotype

Morocco

Population data

## ABSTRACT

A sample of 267 unrelated Moroccan males from different ethnic groups (Arabs, Berbers and Sahrawi), was typed for 17 Y-STR loci (DYS19, DYS385, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, Y GATA H4). Discrimination capacity (96.3%) and haplotype diversity (99.91%) were calculated. A total of 257 haplotypes were identified, of which 237 were unique and 10 were found in two individuals each. DYS385 showed the highest diversity (0.887) followed by DYS458 (0.820) as a single locus marker.

© 2009 Elsevier Ireland Ltd. All rights reserved.

## 1. Population

Morocco is a country located in northwest Africa on the Mediterranean Sea and the Atlantic Ocean. Inhabited from ancient times by Berbers, the region became a Roman province in the 1st century and was conquered by Arabs in the 7th century. The country was later united (11th–13th century) under Berber-Muslim dynasties. Morocco is the third most populous Arab country, after Egypt and Sudan. Morocco's official language is classical Arabic. The country's distinctive Arabic dialect is called Moroccan Arabic (DARIJA). In rural areas, Berbers speak three different dialects (Tarfitt, Tashelhit, and Tamazight) either as a first language or bilingually with the spoken Arabic dialect [1].

## 2. Sampling

Blood samples were obtained from 267 unrelated, healthy male individuals from the three ethnic groups (Arab-speaking  $n = 130$ , Berber-speaking  $n = 69$  and Sahrawi  $n = 68$ ) belonging to the whole area of Morocco. All donors signed the informed consent and provided information about their ethnolinguistic origin as well as the one of their parents and grandparents.

## 3. DNA extraction

Genomic DNA was extracted using DNA IQ<sup>™</sup> System (Promega, Madison, USA) according to manufacturer's instructions.

## 4. PCR amplification

PCR for 17 Y-STR loci was performed in multiplex reaction using AmpFISTR<sup>®</sup>Y-filer<sup>™</sup> kit (Applied Biosystems, Foster City, CA, USA). The total amplification reaction volume was 25  $\mu$ l, which contained 9.2  $\mu$ l of reaction mix, 5.0  $\mu$ l of primer set, 0.8  $\mu$ l of AmpliTaq Gold DNA polymerase and 10  $\mu$ l of genomic DNA (approximately 0.5–1 ng/ $\mu$ l). Amplifications were performed as described in the manufacturer's protocols of the AmpFISTR<sup>®</sup>Y-filer<sup>™</sup> kit using the GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) [2].

\* Corresponding author at: Biochemistry and Immunology Laboratory (Forensic Genetic Unit), Mohammed V Agdal University, Faculty of Sciences, BP 1014, Avenue Ibn Batouta Agdal, Rabat, Morocco. Tel.: +212 61 24 32 20; fax: +212 37 77 42 61.

E-mail addresses: [a.rachid22@gmail.com](mailto:a.rachid22@gmail.com) (R. Aboukhalid), [Bou.mehdi@gmail.com](mailto:Bou.mehdi@gmail.com) (M. Bouabdellah), [abbassi.meriame@gmail.com](mailto:abbassi.meriame@gmail.com) (M. Abbassi), [kaoutarbely55@hotmail.com](mailto:kaoutarbely55@hotmail.com) (K. Bentayebi), [mzibri@yahoo.com](mailto:mzibri@yahoo.com) (M. Elmzibri), [dr.squalli@gmail.com](mailto:dr.squalli@gmail.com) (D. Squalli), [amzazi@gmail.com](mailto:amzazi@gmail.com) (S. Amzazi).

<sup>1</sup> Tel.: +212 61 09 20 60; fax: +212 22 34 41 35.

<sup>2</sup> Tel.: +212 67 41 10 72; fax: +212 37 77 42 61.

<sup>3</sup> Tel.: +212 11 97 10 59; fax: +212 37 77 42 61.

<sup>4</sup> Tel.: +212 66 83 59 11; fax: +212 37 71 18 46.

<sup>5</sup> Tel.: +212 61 40 92 95; fax: +212 22 34 41 35.

<sup>6</sup> Tel.: +212 61 90 93 78; fax: +212 37 77 42 61.

## 5. Electrophoresis and typing

Separation and detection of the 17 Y-STR loci were performed using the ABI Prism 3130xl DNA Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Each sample was prepared by adding 1.5  $\mu$ l PCR product to 24.5  $\mu$ l of Hi-Di™ formamide and 0.5  $\mu$ l GeneScan-500 LIZ Internal Size Standard (Applied Biosystems). Samples were analyzed using GeneMapper ID v3.2 (Applied Biosystems). Alleles were assigned according to the International Society of Forensic Genetics (ISFG) guidelines for forensic Y-STR analysis [3].

## 6. Data analysis

Haplotype frequencies were calculated using the Arlequin software v.3.1 [4]. Haplotype diversity (HD) was determined as  $HD = N/N - 1(1 - \sum pi^2)$ , where  $N$  is the population size and  $pi$  is the frequency of the  $i$ th haplotype [5]. Gene diversity (GD) (equivalent to the power of discrimination) was calculated with the same equation using allele frequencies rather than haplotype frequencies. The discrimination capacity (DC) and unique haplotype (UH) were also determined. To compare our data to other populations, analysis of molecular variance (AMOVA) was assessed with  $\Phi_{st}$  values, and the distance was visualized in two multi-dimensional scaling (MDS) using YHRD online tools ([www.yhrd.org](http://www.yhrd.org)) [6].

## 7. Quality control

Laboratory internal control standards and kit controls. The data have been submitted to the YHRD for quality checks in advance of publication and received the following accession numbers: Rabat, Morocco [Arab] YA003437, Rabat, Morocco [Berber] YA003438, and Rabat, Morocco [Sahrawi] YA003439. The populations above-cited can be searched at [www.yhrd.org](http://www.yhrd.org) [6] by contributor, accession number or by population name.

## 8. Results

Table 1 presents the haplotype distribution of the Y-chromosome obtained among 267 individuals analyzed.

## 9. Other remarks

A total of 257 (96.25%) different haplotypes were found from 267 unrelated individuals, of which 237 were unique and 10 were found in two individuals each. In four individuals, two alleles per locus were observed at the loci DYS448 (2), DYS456 (1) and DYS635 (1) (See Table 1). Null alleles were observed in seventeen individuals at DYS456 and in two individuals at DYS458, a second retyping for these individuals was done revealing the same profile.

The gene diversity values of 17 loci showed the highest diversity at DYS385 (0.887) followed by DYS458 (0.820) as a single locus marker. The haplotype diversity (HD), discrimination capacity (DC) and unique haplotype (UH) were, respectively, 0.9991, 0.963 and 247. Y-STR polymorphisms in Moroccan population, using Y-filer, could be then a powerful discrimination tool for routine forensic applications.

Two microvariant alleles (18.2 and 19.2) were observed at the locus DYS458 with a lower frequency (5.99%) than that observed in the Upper (Southern) Egyptian population [7]. We observed also one sample with the smallest allele 13 at the same locus.

The haplotype data were compared with the populations submitted to the YHRD database, namely Algeria [8], Tunisia [9], Upper Egypt [7], Iran [6], Greek [10], Italy [11], Central Portugal [12], Namibia [6] and China [13]. As shown in the multidimensional scaling (MDS) (Fig. 1) and confirmed by genetic distances in Table 2, our studied sample seems to be nearest to Tunisian, Algerian and Iranian population than other Mediterranean populations like Upper Egyptian and Greek. A high genetic distance was observed when comparing our population with Namibian, Italian, Portuguese and Chinese. Cultural, geographical and historic background may be responsible of this distribution.

This paper follows the guidelines for publication of population data requested by the journal [14].

## Acknowledgement

This work was supported by GENOME Biotechnologies, the exclusive distributor of Applied Biosystems in Morocco. The authors acknowledge Dr. Jalal Talbi and Dr. Hicham El Ossmani for their contribution in the results interpretation and discussion, the blood donors and the technicians of the Regional Center of Blood Transfusion of Rabat.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.fsigen.2009.06.004](https://doi.org/10.1016/j.fsigen.2009.06.004).

## References

- [1] Muqadima, Ibn Khaldoun, d'après les Prolégomènes d'Ibn Khaldoun (1332-1406 de J.C) traduits en français et commentés par W. Mac Guckin de Slane, 1801–1878.
- [2] Applied Biosystems, AmpFISTR® Yfiler® PCR Amplification Kit User's Manual, Foster City, CA, 2004.
- [3] L. Gusmão, J.M. Butler, A. Carracedo, P. Gill, M. Kayser, W.R. Mayr, N. Morling, M. Prinz, L. Roewer, C. Tyler-Smith, P.M. Schneider, DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, *Forensic Sci. Int.* 157 (2006) 187–197.
- [4] L. Excoffier, G. Laval, S. Schneider, Arlequin version 3.11: an integrated software package for population genetics data analysis, *Evol. Bioinform. Online* 1 (2005) 47–50.
- [5] M. Nei, *Molecular Evolutionary Genetics*, Columbia University Press, New York, 1987, p. 178.
- [6] Y Chromosome, Haplotype Reference Database, [www.yhrd.org](http://www.yhrd.org).
- [7] G.A. Omran, G.N. Ruttly, M.A. Jobling, Diversity of 17-locus Y-STR haplotypes in Upper (Southern) Egyptians, *Forensic Sci. Int. Gene. Suppl.* (2008) 230–232.
- [8] C. Robino, F. Crobu, C. Di Gaetano, A. Bekada, S. Benhamamouch, N. Cerutti, A. Piazza, S. Inturri, C. Torre, Analysis of Y-chromosomal SNP haplogroups and STR haplotypes in an Algerian population sample, *Int. J. Legal Med.* 122 (3) (2008) 251–255.
- [9] V. Onofri, F. Alessandrini, C. Turchi, M. Pesaresi, A. Tagliabracci, Y-chromosome markers distribution in Northern Africa: high-resolution SNP and STR analysis in Tunisia and Morocco populations, *Forensic Sci. Int. Gene. Suppl.* 1 (2008) 235–236.
- [10] L. Kovatsi, J.L. Saunier, J.A. Irwin, Population genetics of Y-chromosome STRs in a population of Northern Greeks, *Forensic Sci. Int. Genet.* (2009), [doi:10.1016/j.fsigen.2009.01.001](https://doi.org/10.1016/j.fsigen.2009.01.001).
- [11] G. Ferri, M. Ald, B. Corradini, E. Radhesi, G. Beduschi, Slow and fast evolving markers typing in Modena males (North Italy), *Forensic Sci. Int. Genetic*, 3 (2009) e31–e33.
- [12] M. Carvalho, M.J. Anjos, L. Andrade, V. Lopes, M.V. Santos, J.J. Gamero, F. Corte Real, M.C. Vide, Y-chromosome STR haplotypes in two population samples: Azores Islands and Central Portugal, *Forensic Sci. Int.* 134 (1) (2003) 29–35.
- [13] B. Zhu, Y. Wu, C. Shen, T. Yang, Y. Deng, X. Xun, Y. Tian, J. Yan, T. Li, Genetic analysis of 17 Y-chromosomal STRs haplotypes of Chinese Tibetan ethnic group residing in Qinghai province of China, *Forensic Sci. Int.* 175 (2–3) (2008) 238–243.
- [14] P. Lincoln, A. Carracedo, Publication of population data of human polymorphism, *Forensic Sci. Int.* 110 (2000) 3–5.

# Y Chromosomal SNP Analysis Using the Minisequencing Strategy in a Moroccan Population Samples

Rachid Aboukhalid<sup>1,4\*</sup>, Mehdi Bouabdellah<sup>1,4</sup>, Ronald Lai<sup>2</sup>, Nicholas Bradford<sup>2</sup>, Hicham EL Ossmani<sup>3</sup>, Jalal Talbi<sup>3</sup>, Kaoutar Bentayebi<sup>1</sup>, Driss Squalli<sup>4</sup>, Daniele Podini<sup>2</sup> and Saaid Amzazi<sup>1</sup>

<sup>1</sup>Forensic Genetic Unit, Biochemistry and Immunology Laboratory, Mohammed V University, Faculty of Science, BP 1014, Avenue Ibn Batouta Agdal, Rabat, Morocco

<sup>2</sup>Department of Forensic Science, The George Washington University, 2100 Foxhall Road NW Washington, DC 20007, United States

<sup>3</sup>Laboratory of Anthropogenetics and Physiopathology, Department of Biology, University of Chouaib Doukkali, Faculty of Science, El Jadida, Morocco

<sup>4</sup>GENOME Biotechnologies, Parc d'activité Oukacha, Imm E N°16, Ain Sebaâ, Casablanca, Morocco

## Abstract

The Y chromosome contains the largest non-recombining portion in the human genome. Y-Binary polymorphisms, also known as a single nucleotide polymorphisms (SNPs), are a series of biallelic polymorphisms occurring on the non-recombining region of the Y chromosome (NRY), which represent a precious tool for human evolutionary studies and, potentially, for forensic applications. Low mutation rate, paternal inheritance, and absence of recombination make Y-SNPs particularly suitable for the identification of stable paternal lineages and the reconstruction of an ancestral state from which to explore the evolution of humans. Also, these markers would allow inference of the paternal ancestry of unknown samples which could be useful in forensic applications. In the present study we analyzed 22 biallelic Y-SNPs in 159 males belonging to three ethnic groups (Arab n = 42, Berber n = 67 and Sahrawi n = 50) from Morocco. A total of 10 different haplogroups were identified in this sample representative of Moroccan population. The most common Y chromosome haplogroups is E1b1b1, E1b1b1b and J1 with frequencies of 56%, 49% and 10% respectively.

**Keywords:** Y-chromosomal SNP; Mini-sequencing; SNaPshot; Morocco

## Introduction

Biallelic markers, such as single nucleotide polymorphisms (SNPs) and insertion/deletions (indels), represent an important class of markers on the Y-chromosome [1]. Y-chromosome SNPs (Y-SNPs) are mostly used in molecular anthropology for evolutionary research. Moreover, typing a sample's Y chromosome haplogroup allows paternal ancestry inference. This may be useful, in forensic applications, when a conventional Short Tandem Repeat (STR) profile, generated from DNA collected at a crime scene, does not match any of the identified suspects and doesn't "hit" any profile on the available databases.

The kingdom of Morocco is a country located in the northwestern corner of the African continent with coasts on the Atlantic Ocean and the Mediterranean Sea; it is bordered by Algeria to the east and Mauritania to the south. Modern-day Morocco is inhabited by three major ethnic groups (Arab, Berber and Sahrawi). Several dialects such as Arabic (Moroccan dialect or Darija), Berber (Tarifit, Tachelhit and Tamazight) and Sahrawi (El Hassania) are spoken in the country. The aim of the present study was to develop an assay to genotype a selected panel of Y-SNPs using single base extension assay (SBE), or minisequencing [2], in order to determine the most frequent Y-chromosomal haplogroup in Moroccan population. The analysis of single nucleotide polymorphisms (SNPs) is a promising application in forensic casework; since the forensic scientists is often faced with degraded and/or very low amounts of DNA.

## Materials and Methods

### Population

Buccal swabs were collected from 159 unrelated healthy adult men belonging to the three ethnic groups (Arab n = 42, Berber n = 67 and Sahrawi n = 50). Informed consent was obtained from all participants in this study, and information about the geographical origin of their grand-parents and about their first language was recorded.

### DNA Isolation

Genomic DNA was extracted from buccal swab punches, using DNA IQ™ System (Promega; Madison, Wisconsin) on the Biomek® 2000 (Beckman Coulter, Brea, CA) robotic platform according to manufacturer's instructions.

### Y-SNP selection and multiplex design

Assay design and development together with sample testing were conducted in the Forensic Molecular Biology Laboratory of the Forensic Sciences Department of The George Washington University. A total of 22 Y chromosome biallelic markers were selected for this study following a hierarchical strategy based on the phylogenetic tree of Y chromosome recognized by the Y Chromosome Consortium (YCC) [3]. Loci Nomenclature is based on Karafet et al. [4]. The reference sequence for each Y-SNP was taken from the International Society of Genetic Genealogy, Y-DNA Haplogroup Tree 2010 (<http://www.isogg.org/tree>). The 22 markers were genotyped in two heptaplex, one hexaplex, and one duplex PCR reactions. The Multiplex MY01 allows the detection of major clades (A-R), A-M91 (A), B-M60 (B), C-RPS4Y<sub>711</sub> (C), D-M174 (D), J-M267 (J1), I-M170 (I) and R-M207 (R). The Multiplexes MY02 E-P147 (E1), E-M132 (E1a), E-P189 (E1b1a), E-M215 (E1b1b), E-M81 (E1b1b1b), E-M54 (E2b), J-M172 (J2), MY03 J-L24 (J2a4h), J-M221 (J2b), R-L63 (R1a), R-M343 (R1b), R-M269 (R1b1b2),

\*Corresponding author: Rachid Aboukhalid, Genetic Unit, Biochemistry and Immunology Laboratory, Mohammed V University, Faculty of Science, BP 1014, Avenue Ibn Batouta Agdal, Rabat, Morocco, GENOME Biotechnologies, Parc d'activité Oukacha, Imm E N° 16, Ain Sebaâ, Casablanca, Morocco Tel: +212 61 24 32 20; E-mail: [a.rachid22@gmail.com](mailto:a.rachid22@gmail.com)

Received December 16, 2010; Accepted December 29, 2010; Published December 30, 2010

Citation: Aboukhalid R, Bouabdellah M, Lai R, Bradford N, El Ossmani H, et al. (2010) Y Chromosomal SNP Analysis Using the Minisequencing Strategy in a Moroccan Population Samples. J Forensic Res 1:113. doi:10.4172/2157-7145.1000113

Copyright: © 2010 Aboukhalid R, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



T-M70 (T) and **MY04** E-M35 (E1b1b1), E-M78 (E1b1b1a) subdivides haplogroups E, J, R and T to define the most frequent haplogroups in NorthWest Africa. Multiplex minisequencing primers were designed for each multiplex PCR. PCR primers used in the primer extension assays are listed in Table 1. The extension primers designed in four multiplex sets labeled msMY01, msMY02, msMY03 and msMY04 are listed in Table 2.

### PCR amplification

Amplifications were performed in a final volume of 25 $\mu$ L using a master mix containing 1X GeneAmp<sup>®</sup> PCR Gold buffer (Applied Biosystems, Foster City, CA), 400 $\mu$ M of dNTPs, 3 mM of MgCl<sub>2</sub>, 0.65 mg/ $\mu$ L bovine serum albumin (BSA), and 1.5 units of AmpliTaq Gold<sup>®</sup> DNA polymerase (Applied Biosystems). The thermal cycling program was carried out on a GenAmp 9700 (Applied Biosystems) using the following conditions: 95°C for 10 minutes, 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 30 seconds, and a final extension at 72°C for 10 minutes.

Following PCR amplification, unincorporated primers and dNTPs were removed by adding 1 unit of Exonuclease I (USB Corporation) and 1 unit of Shrimp Alkaline Phosphatase (SAP) to 5 $\mu$ L of PCR products and incubating a 37°C for 70 minutes followed by 20 minutes at 65°C.

### Primer extension assay using fluorescence detection

Multiplex primer extension reactions were carried out in a total volume of 15 $\mu$ L using 1.5 $\mu$ L of ABI Prism<sup>®</sup> SNaPshot<sup>™</sup> multiplex kit mix (Applied Biosystems), 2 $\mu$ L of purified PCR product and 1.5 pmols of each extension primer. Extension reaction was as follows: 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 30 seconds.

A 1.5 $\mu$ L of non-purified primer extension product was diluted in 9.8 $\mu$ L Hi-Di<sup>™</sup> formamide and 0.2 $\mu$ L of GS120-LIZ internal lane standard (Applied Biosystems) and analyzed on ABI Prism<sup>®</sup> 3130 Genetic Analyzer (Applied Biosystems). Separation was performed on a 36 cm array using POP<sup>™</sup> - 7 (Applied Biosystems). Results were analyzed using GeneMapper v 4.0 (Applied Biosystems).

### Results and Discussion

A total of 10 haplogroups were observed across the 159 male samples that were tested (Table 3). The most common Y chromosome haplogroups in the combined Moroccan population resulted E1b1b1, E1b1b1b and J1 (56%, 49% and 10% respectively).

The E-M215 derivative E1b1b1 is observed in significant frequencies in Moroccan samples. It's defined by the E-M35 SNP which appears to have originated in East Africa then migrating to the Near East and then on to North Africa and Europe.

Previous studies have shown that E1b1b1b (E-M81), formerly E3b1b or E3b2, is the predominant haplogroup in northwestern Africa [5 - 8]. Bosh et al. [9] obtained a higher frequency (64%) of E-M81 (E1b1b1b) in Northwest African populations than the present study. This disparity could be explained by the difference of geographic area and circumstances of sample collection. However, according to our results and those of Bosh et al. [9], it seems that the E1b1b1b haplogroup is more characteristic of Sahrawi than Berbers. It is also seen, although at low frequency, in the Iberian Peninsula (4%) and Sicily (3%) due to recent gene flow from Northwest Africa through Gibraltar strait [9 - 10]. Nevertheless, this haplogroup (E1b1b1b) is

not found in sub-Saharan Africa and its frequency sharply declines through the continent towards the east.

The J1 haplogroup, defined by the single nucleotide polymorphism (SNP) M267, is most frequent in the Arabian Peninsula especially in Yemen (76%) [11]. This could be attributed to the early medieval period during which the Semitic expansion spread J1 out of Arabia into North Africa [12].

The multiplex assays described herein were designed to explore the shallowest branches of Y chromosome haplogroups in Moroccan population. They could also be applied to human evolution and human geneticists studies as well as to forensic casework for ancestry inferences.

### Acknowledgement

We would like to express our sincere gratitude to the Moroccan-American Commission for Educational and Cultural Exchange (MACECE) and the Fulbright Program for the Fulbright scholarship, for supporting a part of our research in the United States. We would like, also, to thank the George Washington University and the Armed Forces DNA Identification Laboratory (AFDIL). We are greatly indebted to Kimberly Andreaggi Sturk, Jodi A. Irwin, Joni Johnson and Katherine Butler for discussion and technical assistance. Finally, an honorable mention goes to the support offered by GENOME Biotechnologies.

### Reference

1. Butler JM (2003) Recent developments in Y-single tandem repeat and Y-single nucleotide polymorphism analysis. *Forensic Sci Rev* 15: 91.
2. Podini D, Vallone PM (2009) SNP genotyping using multiplex single base primer extension assays. *Methods Mol Biol* 578: 379-391.
3. Y chromosome consortium (2002) A nomenclature system for the three of human Y-chromosomal binary haplogroups. 339-348.
4. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, et al. (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18: 830-838.
5. Onofri V, Alessandrini F, Turchi C, Pesaresi M, Tagliabracci A (2008) Y-chromosome markers distribution in Northern Africa: High-resolution SNP and STR analysis in Tunisia and Morocco populations. *Forensic Sci Int: Genetics* 235-236.
6. Robino C, Crobu F, Gaetano D, Bekada A, Benhamamouch S, et al. (2008) Analysis of Y-Chromosomal SNP haplogroups and STR haplotypes in Algerian population sample. *Int J Legal Med* 122: 251-255.
7. Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, et al. (2004) A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. *Am J Hum Genet* 75: 338-345.
8. Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, et al. (2004) Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. *Am J Hum Genet* 74: 1014-1022.
9. Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, et al. (2001) High-Resolution Analysis of Human Y-Chromosome Variation Shows a Sharp Discontinuity and Limited Gene Flow between Northwestern Africa and the Iberian Peninsula. *Am J Hum Genet* 68: 1019-1029.
10. Adams SM, Bosch E, Balaesque PL, Ballereau SJ, Lee AC, et al. (2008) The Genetic Legacy of Religious Diversity and Intolerance: Paternal Lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet* 83: 725-736.
11. Alshamali F, Pereira L, Budowle B, Poloni ES, Currat M (2009) Local population structure in Arabian Peninsula revealed by Y-STR diversity. *Hum Hered* 68: 45-54.
12. Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, et al. (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74: 1023-1034.



Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

## Forensic Science International: Genetics

journal homepage: [www.elsevier.com/locate/fsig](http://www.elsevier.com/locate/fsig)

## Letter to the Editor

**Genetic diversity of 12 X-chromosomal short tandem repeats in the Moroccan population***Dear Editor,*

Morocco is a country located in northwest Africa. It has a long coastline on the Atlantic Ocean that reaches past the Strait of Gibraltar into the Mediterranean Sea. Morocco has been the home of the Berbers since the second millennium BC. It was annexed by Rome in the 1st century and invaded by Arabs in the 7th century. The country was later united (11th–13th century) under Berber-Muslim dynasties. Since several invasions have successively interspersed with the original population throughout history (e.g. Phoenicians, Romans, Arabs, Spanish, French and Sub-Saharan Africans) [1], actually, the Moroccan population is an admixture of Mediterranean, African and Oriental populations, as evidenced by the heterogeneous gene pool [2].

Since Paleolithic time North-West Africa have been inhabited by several human groups that are different by their ethnic identity, language and culture. In Morocco, the country's official language is classical Arabic, although different languages or dialects are spoken by the different ethnic groups. Thus, Arabic people dialect is called Darija; Sahrawis, who live mostly in Western Sahara, speak another Arabic dialect, Hassania. The latest group is Berbers, who forms approximately 8 million of the Moroccan population; they speak three different Berber dialects: Tarifit in the north, Tamazight in the central region and Tashelhit in the south area of Morocco, either as a first language or bilingually with the spoken Arabic dialect [1].

The present study aimed to build an X-STR database of the Moroccan population for anthropological and forensic purposes, as well as to provide reference data for comparison with other studies based on the analysis of the 4 X-STR trios included in the new kit Investigator Argus X-STR 12 kit (Qiagen, Hilden, Germany).

Blood samples were collected from 145 unrelated, healthy individuals (97 male and 48 female) from the three ethnic groups (Arab, Berber and Sahrawi) living in the whole area of Morocco already typed for 15 autosomal short tandem repeats [3] and 17 Y-STR markers [4]. All participants signed the informant consent and provided information about their ethno-linguistic as well as their parents and grandparents origins. Genomic DNA was extracted using DNA IQ™ System (Promega, Madison, USA) and PCR amplification for 12 X-STR loci was performed using Investigator Argus X-12 Kit (Qiagen) according to the manufacturer's instructions [5]. For genetic typing, an ABI Prism 3130 DNA Genetic Analyzer along with GeneMapper ID 3.2 software (Applied Biosystems, Foster City, CA) was used.

Forensic parameters were calculated at the Chromosome X homepage (<http://www.chrx-str.org>). Allele frequencies, exact test of the Hardy–Weinberg equilibrium for the female samples (HWE), pairwise exact test of linkage disequilibrium (LD) and haplotype diversity for the male samples, and population pairwise genetic distance (*Fst*) were carried out using the Arlequin v.3.0 software [6].

Allele frequencies and statistic parameters for forensic evaluation for each of the 12 X-STR markers are displayed in **Supplementary Table 1**. No significant differences were found between the three ethnic groups. DXS10146 was the most polymorphic X-STR, with 24 alleles, and DXS8378 was the least polymorphic one, with 5 alleles. Overall values obtained for the power of discrimination were high ( $>0.999999$ ) in both females and males. Combined mean exclusion chance for trios and duos were 0.99999999 and 0.9999998, respectively. No deviations from Hardy–Weinberg equilibrium were observed after Bonferroni correction in the Moroccan population. The present study shows that there is a high genetic diversity in the Moroccan population for these 12 X-STRs.

Considering the *p*-values from locus by locus comparisons between Morocco and other published population data [7–20], our population seems to be nearest to Algerian and some European populations than to sub-Saharan Africans (**Supplementary Table 2**). Nevertheless, the results showed some differences even with Algerian, in accordance with the X chromosomal SNPs study, where a high overall genetic homogeneity was observed in Mediterranean populations except in Morocco [21]. These results are consistent with the influence of the Romans and the history of Moroccan colonization, since the largest migratory contingent in this population was European, composed mainly of French, Spanish and Portuguese between the fourteenth and nineteenth centuries [22]. However, the Moroccan population showed significant differences from the Asian populations in almost all loci implying genetic differences according to the geographic location.

The exact test for linkage disequilibrium was performed for all pairs of loci in all the sampled population. After Bonferroni correction, only a significant *p* value was obtained for DXS10101–DXS10146 pair of loci ( $p \leq 0.0000$ ), which are relatively quite distant on the chromosome X (over 22 Mb). The LD does not only depend on the distance between marker pairs, but may be associated with a random genetic drift, founder effect, mutations, selection and population admixture or stratification [23]. Thus, studies with larger sample size would be necessary to obtain true LD information.

Since it has been recently suggested that the use of haplotypes of closely linked X-STR trios instead of a single or duos of STRs might be a powerful tool in forensic identification and kinship testing [24], three markers of each linked group have been handled as haplotypes for genotyping in a sample of 97 Moroccan man. The 4 X-STR trios of linkage group 1–4 revealed 88, 64, 72 and 76 haplotypes, respectively (**Supplementary Table 3**). Of all the observed haplotypes, 95% showed frequencies  $<0.021$ , and only four haplotypes were found in four or five individuals each, displaying a frequency higher than 0.040. Haplotype diversity values (**Supplementary Table 4**) in the Moroccan population were all higher than 0.9897. The linkage group 2 (DXS7132–DXS10079–DXS10074) had the lowest haplotype diversity value. These values, as expected, are higher than those found in the previous study of

four X-STR duos in the Korean population [17] and nearest to those found in Koreans involving the same X-STR trios set [18].

In conclusion, the forensic efficiency parameters of the twelve X-STRs investigated in this work demonstrate the usefulness of these markers in population genetic studies, human identification and complex kinship cases. Therefore, these results will contribute to establish this system as standard for the X-STR analysis in the Moroccan population.

Moreover, few X-STR multiplex kits have become commercially available and validated for forensic casework. As a consequence, there are few data for comparison because there is no a set of X-STR markers widely used by the forensic community. Here we present the first population study performed using the new Argus X-12 kit (Qiagen), to our knowledge, which it will serve as a reference for comparison with other studies based on these 4 X-STR trios.

This work was supported by an AVERROES grant (Erasmus MUNDUS, European Commission).

This study follows the ISFG recommendations [25] and the guidelines for publication of population data proposed by the journal [26].

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2011.03.008.

## References

- [1] Muqadima, Ibn Khaldoun, d'après les Prolégomènes d'Ibn Khaldoun (1332-1406 de J.C) traduits en français et commentés par W. Mac Guckin de Slane, 1801–1878.
- [2] S. Plaza, F. Calafell, A. Helal, N. Bouzerna, G. Lefranc, J. Bertranpetit, D. Comas, Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean, *Ann. Hum. Genet.* 67 (2003) 312–328.
- [3] M. Bouabdeallah, F. Ouenzar, R. Aboukhalid, M. Elmzibri, D. Squalli, S. Amzazi, STR data for 15 AmpFISTR Identifier loci in the Moroccan population, *Forensic Sci. Int. Genet. Suppl.* 1 (2008) 306–308.
- [4] R. Aboukhalid, M. Bouabdeallah, M. Abassi, K. Bentayebi, M. Elmzibri, D. Squalli, S. Amzazi, Haplotype frequencies for 17 Y-STR loci (AmpFISTR1 Y-filer™) in a Moroccan population sample, *Forensic Sci. Int. Genet.* 4 (2010) e73–e74.
- [5] Qiagen, Investigator Argus 12 X STR Kit Handbook, Madrid, S.A, Spain, 2010.
- [6] L. Excoffier, G. Laval, S. Schneider, Arlequin ver. 3.0: an integrated software package for population genetics data analysis, *Evol. Bioinform. Online* 1 (2005) 47–50.
- [7] A. Bekada, S. Benhamamouch, A. Boudjema, F. Mustapha, S. Menegon, C. Torre, C. Robino, Analysis of 21 X-chromosomal STRs in an Algerian population sample, *Int. J. Legal Med.* 124 (2010) 287–294.
- [8] M. Poetsch, A. Knop, D. El-Mostaqim, N. Rakotomavo, N. Wurmb-Schwark, Allele frequencies of 11 X-chromosomal loci of two population samples from Africa, *Int. J. Legal Med.* (2010) (Epub ahead of print).
- [9] M. Hedman, J.U. Palo, A. Sajantila, X-STR diversity in the Finnish and the Somali population, *Forensic Sci. Int. Genet.* 3 (2009) 173–178.
- [10] I. Gomes, C. Alves, K. Maxzud, R. Pereira, M. Prata, P. Sánchez-Diz, A. Carracedo, A. Amorim, L. Gusmão, Analysis of 10 X-STRs in three African populations, *Forensic Sci. Int. Genet.* 1 (2007) 208–211.
- [11] M. Aler, P. Sánchez-Diz, I. Gomes, M. Gisbert, A. Carracedo, A. Amorim, L. Gusmão, Genetic data of 10 X-STR in a Spanish population sample, *Forensic Sci. Int.* 173 (2007) 193–196.
- [12] C. Robino, A. Giolitti, S. Gino, C. Torre, Analysis of 12 X-chromosomal short tandem repeats in the North-West Italian population by means of two multiplex PCRs, *Int. Congr. 1288* (2006) 316–318.
- [13] J. Edelmann, S. Hering, M. Michael, R. Lessig, D. Deichsel, G. Meier-Sundhausen, L. Roewer, I. Plate, R. Szibor, 16 X-chromosome STR loci frequency data from a German population, *Forensic Sci. Int.* 124 (2001) 215–218.
- [14] R. Pereira, I. Gomes, A. Amorim, L. Gusmão, Genetic diversity of 10 X chromosome STRs in northern Portugal, *Int. J. Legal Med.* 121 (2007) 192–197.
- [15] I. Gomes, M. Prinz, R. Pereira, C. Meyers, R. Mikulasovich, A. Amorim, A. Carracedo, L. Gusmão, Genetic analysis of three US population groups using an X-chromosomal STR decaplex, *Int. J. Legal Med.* 121 (2007) 198–203.
- [16] J.A. Martins, R.H.A. Silva, A. Freschi, G.G. Paneto, R.N. Oliveira, R.M.B. Cicarelli, X-chromosome genetic variation in São Paulo State (Brazil) population, *Ann. Hum. Biol.* 37 (2010) 598–603.
- [17] E.J. Lim, H.Y. Lee, J.E. Sim, W.I. Yang, K.J. Shin, Genetic Polymorphism and Haplotype Analysis of 4 Tightly Linked X-STR Duos in Koreans, *Croat. Med. J.* 50 (2009) 305–312.
- [18] J.E. Sim, H.Y. Lee, W.I. Yang, K.J. Shin, Population genetic study of four closely-linked X-STR trios in Koreans, *Mol. Biol. Rep.* 37 (2010) 333–337.
- [19] L. Kang, S. Li, X-chromosome STR polymorphism of Luoba Ethnic Group in Tibet (SW China), *Forensic Sci. Int.* 156 (2006) 88–90.
- [20] M.Y. Chen, C.E. Pu, Population data on the X chromosome short tandem repeat loci DXS10011, DXS101, DXS6789, DXS7132, DXS8377, and DXS9895 in Taiwan, *Forensic Sci. Int.* 146 (2004) 65–67.
- [21] C. Tomas, J.J. Sanchez, A. Barbaro, C. Brandt-Casadevall, A. Hernandez, M. Ben Dhiab, M. Ramon, N. Morling, X-chromosome SNP analyses in 11 human Mediterranean populations show a high overall genetic homogeneity except in North-west Africans (Moroccans), *BMC Evol. Biol.* 8 (2008) 75.
- [22] B. Luga, Histoire du Maroc—Des origines à nos jours, Édition Perrin, Paris, 2000.
- [23] A. Chakravarti, Population genetics—making sense out of sequence, *Nat. Genet.* 21 (1999) 56–60.
- [24] R. Szibor, X-chromosomal markers: past, present and future, *Forensic Sci. Int. Genet.* 1 (2007) 93–99.
- [25] B. Olaisen, W. Bär, B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, P. Lincoln, W.R. Mayr, S. Rand, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, DNA recommendations 1997 of the International Society for Forensic Genetics, *Vox Sang.* 74 (1998) 61–63.
- [26] A. Carracedo, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, Publication of population data for forensic purposes, *Forensic Sci. Int. Genet.* 4 (2010) 145–147.

Kaoutar Bentayebi<sup>a,b,\*</sup>

<sup>a</sup>Biochemistry and Immunology Laboratory-Forensic Unit,  
Mohammed V Agdal University, Faculty of Sciences,  
Rabat, Morocco

<sup>b</sup>Institut Universitari d'Investigacions en Ciències de la Salut (IUNICS)  
i Laboratori de Genètica, Departament de Biologia,  
Universitat de les Illes Balears, Palma de Mallorca, Spain

Antònia Picornell<sup>b</sup>

<sup>b</sup>Institut Universitari d'Investigacions en Ciències de la Salut (IUNICS)  
i Laboratori de Genètica, Departament de Biologia,  
Universitat de les Illes Balears, Palma de Mallorca, Spain

Mehdi Bouabdeallah<sup>a,c</sup>

<sup>a</sup>Biochemistry and Immunology Laboratory-Forensic Unit,  
Mohammed V Agdal University, Faculty of Sciences,  
Rabat, Morocco

<sup>c</sup>GENOME Biotechnologies Parc d'activitats Oukacha imm. B,  
N8 35 Ain Sebaâ, Casablanca, Morocco

José Aurelio Castro<sup>b</sup>

<sup>b</sup>Institut Universitari d'Investigacions en Ciències de la Salut (IUNICS)  
i Laboratori de Genètica, Departament de Biologia,  
Universitat de les Illes Balears, Palma de Mallorca, Spain

Rachid Aboukhalid<sup>a,c</sup>

<sup>a</sup>Biochemistry and Immunology Laboratory-Forensic Unit,  
Mohammed V Agdal University, Faculty of Sciences,  
Rabat, Morocco

<sup>c</sup>GENOME Biotechnologies Parc d'activitats Oukacha imm. B,  
N8 35 Ain Sebaâ, Casablanca, Morocco

Driss Squalli<sup>c</sup>

<sup>c</sup>GENOME Biotechnologies Parc d'activitats Oukacha imm. B,  
N8 35 Ain Sebaâ, Casablanca, Morocco

Maria Misericòrdia<sup>b</sup>

<sup>b</sup>Institut Universitari d'Investigacions en Ciències de la Salut (IUNICS)  
i Laboratori de Genètica, Departament de Biologia,  
Universitat de les Illes Balears, Palma de Mallorca, Spain

Saaïd Amzazi<sup>a</sup>

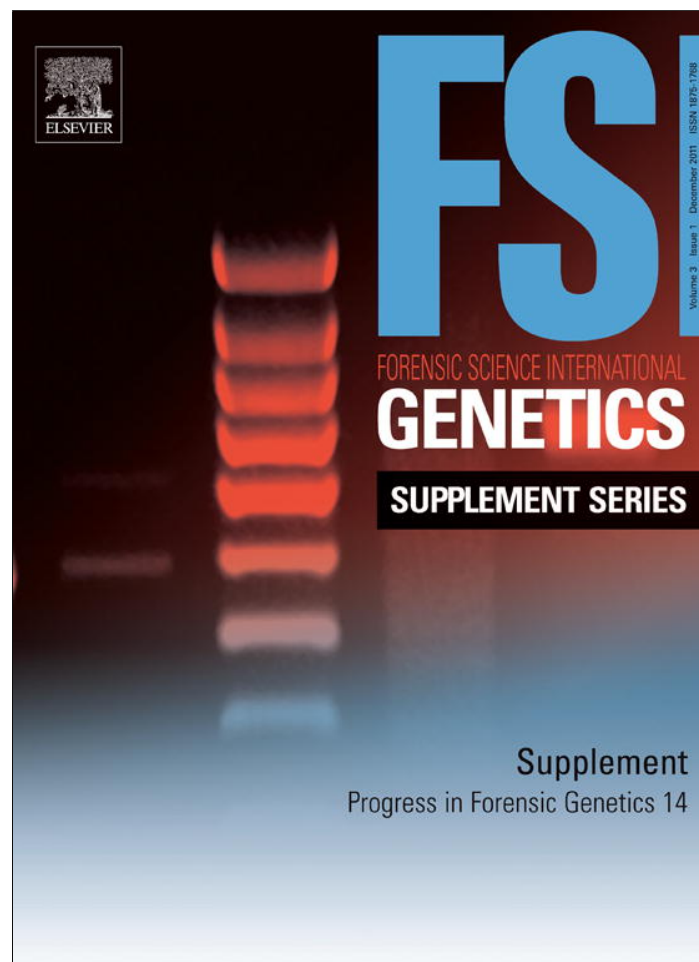
<sup>a</sup>Biochemistry and Immunology Laboratory-Forensic Unit,  
Mohammed V Agdal University, Faculty of Sciences,  
Rabat, Morocco

\*Corresponding author. Tel.: +34971173143

E-mail address: kaoutar.bentayebi@gmail.com (K. Bentayebi).

2 December 2010

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: [www.elsevier.com/locate/FSIGSS](http://www.elsevier.com/locate/FSIGSS)

## X-chromosomal haplotype frequencies of four linkage groups in a North African population

A. Picornell<sup>a,1,\*</sup>, K. Bentayebi<sup>a,b,1</sup>, M. Bouabdeallah<sup>b</sup>, R. Aboukhalid<sup>b</sup>, D. Squalli<sup>b</sup>, J.A. Castro<sup>a</sup>, M.M. Ramon<sup>a</sup>, S. Amzazi<sup>b</sup>

<sup>a</sup> Institut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Departament de Biologia, Universitat de les Illes Balears, Palma de Mallorca, Spain

<sup>b</sup> Biochemistry and Immunology Laboratory-Forensic Unit, Mohammed V Agdal University, Faculty of Sciences, Rabat, Morocco

### ARTICLE INFO

#### Article history:

Received 5 August 2011

Accepted 13 August 2011

#### Keywords:

X-STR

Haplotype frequency

Linkage groups

Moroccan population

### ABSTRACT

Twelve X-STR loci in four clusters of linked markers were typed in a sample of 97 Moroccan men from different ethnic groups (Arabs, Berbers and Sahrawi) using the Investigator Argus X-12 Kit. Forensic statistical parameters for these 12 X-chromosome STR, particularly the haplotypes of the four closely linked X-STR trios, indicated that they are highly informative for anthropological and forensic purposes in Moroccan population.

© 2011 Elsevier Ireland Ltd. All rights reserved.

### 1. Introduction

X-STR markers have a series of special characteristics that justify its increasing interest in forensic practice, population genetics and anthropology. Compared with autosomes, the X-chromosome has lower recombination rate, lower mutation rate and smaller effective population size resulting in a faster genetic drift. In consequence, both linkage disequilibrium (LD) and population structure in the X chromosome are expected to be stronger than those in autosomes. The X-STRs are particularly useful in paternity testing and kinship analysis, especially in deficiency cases, such as grandmother–granddaughter, aunt–niece and cousins. Owing to the complexity of kinship investigation, a new approach of substituting single STRs by stable haplotypes of closely linked loci has been recently suggested [1].

Morocco is a country located in northwest Africa on the Mediterranean Sea and the Atlantic Ocean. Inhabited from ancient times by Berbers, the region became a Roman province in the 1st century and was conquered by Arabs in the 7th century. The country was later united (11th–13th century) under Berber-Muslim dynasties. Since Paleolithic time North-West Africa has been inhabited by several human groups that are different by their ethnic identity, language and culture. In Morocco, the country's

official language is classical Arabic, although different languages or dialects are spoken by the different ethnic groups. Thus, Arabic people dialect is called Darija; Sahrawis, who live mostly in Western Sahara, speak another Arabic dialect, Hassania. The latest group is Berbers, who speak three different Berber dialects: Tarifit, Tamazight and Tashelhit, either as first language or bilingually with the spoken Arabic dialect [2].

The present study aimed to build an X-STR database of the Moroccan population for anthropological and forensic purposes, as well as to provide reference data for comparison with other studies based on the analysis of the 4 X-STR trios included in the new Investigator Argus X-STR 12 Kit (Qiagen, Hilden, Germany).

### 2. Materials and methods

#### 2.1. DNA sample

Blood samples were collected from 97 unrelated males, living in the whole area of Morocco; with known ancestors until at least the third generations and belonging to the three ethnic groups (Arab, Berber and Sahrawi). All participants signed the informant consent and provided information about their ethno-linguistic as well as their parents and grandparents origins.

#### 2.2. PCR and genotyping

Genomic DNA was extracted using DNA IQ™ System (Promega, Madison, USA) and PCR amplification for the 12 X-STR (DXS10148,

\* Corresponding author. Tel.: +34 971 172050; fax: +34 971 173184.

E-mail address: [apicornell@uib.es](mailto:apicornell@uib.es) (A. Picornell).

<sup>1</sup> Both authors contributed equally to this work.

**Table 1**  
Forensic parameters of 4 X-chromosomal STR trios in 97 Moroccan men.

	Linkage group 1 (lg 1)	Linkage group 2 (lg 2)	Linkage group 3 (lg 3)	Linkage group 4 (lg 4)
Number of haplotypes	88	64	72	76
Unique haplotypes	80 (90.9%)	41 (64.0%)	53 (73.6%)	61 (80.3%)
Haplotype Diversity $\pm$ S.D.	0.9979 $\pm$ 0.0019	0.9897 $\pm$ 0.0032	0.9927 $\pm$ 0.0028	0.9942 $\pm$ 0.0025
Discrimination capacity	90.72%	65.98%	74.23%	78.35%
Match probability	0.21%	1.03%	0.73%	0.58%
Frequency of the most common haplotype	0.0309	0.0515	0.0515	0.0309

lg 1: DXS10148–DXS10135–DXS8378; lg 2: DXS7132–DXS10079–DXS10074; lg 3: DXS10103–HPRTB–DXS10101; and lg 4: DXS10146–DXS10134–DXS7423.

DXS10135, DXS8378, DXS7132, DXS10079, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS10134 and DXS7423) was performed using Investigator Argus X-12 Kit (Qiagen) according to the manufacturer's instructions [3]. For genetic typing, an ABI Prism 3130 DNA Genetic Analyzer along with GeneMapper ID 3.2 software was used.

### 2.3. Statistical analysis

Allele and haplotype frequencies, pairwise exact test of LD and haplotype diversity were carried out using the Arlequin v.3.0 software. Discrimination capacity was calculated as the percentage of different haplotypes and haplotype match probability as 1-haplotype diversity. Other forensic parameters were calculated at the Chromosome X homepage (<http://www.chrx-str.org>).

## 3. Results and discussion

The present study shows that there is a high genetic diversity in the Moroccan population for the 12 X-STRs, with polymorphic information content (PIC) values ranging between 0.626 and 0.931. DXS10135 and DXS10146 were the most polymorphic X-STRs, with 23 alleles, and DXS8378 was the least polymorphic one, with 5 alleles. Pairwise linkage disequilibrium between all pairs of the twelve loci was tested. After Bonferroni correction, only a significant *P* value was obtained for DXS10101–DXS10146 pair of loci ( $P < 0.0001$ ). Overall value obtained for the power of discrimination was high ( $>0.999999$ ). Combined mean exclusion chance for trios and duos were 0.999999999 and 0.9999998, respectively.

The statistic parameters for forensic evaluation for each of the 4 X-STR linkage groups are displayed in Table 1. Among the 97 males analyzed, the 4 X-STR trios of linkage groups 1–4 revealed 88, 64, 72 and 76 haplotypes, respectively. Of all the observed haplotypes, 94% showed frequencies  $<0.021$ . Most of them (78%) were only

observed once, and the other haplotypes were shared by two to five men. Haplotype diversity values in the Moroccan population were all higher than 0.9897. The linkage group 2 (DXS7132–DXS10079–DXS10074) had the lowest haplotype diversity value. The discrimination capacity and the match probability ranged from 65.98% (lg 2) to 90.72% (lg 1) and between 1.03% (lg 2) and 0.21% (lg 1), respectively.

The present work demonstrates that haplotypes for the four X-STR trios included in the Investigator Argus X-12 Kit are highly discriminating and, therefore, provide a powerful tool for solving complex kinship cases in Moroccan population.

This study follows the ISFG recommendations [4] and the guidelines for publication of population data proposed by the journal.

### Role of funding

This work was partially supported by grants AAEE133/09 from the *Direcció General de R+D+I (Comunitat Autònoma de les Illes Balears)* and an AVERROES grant (*Erasmus MUNDUS, European Commission*) to K.B.

### Conflict of interest

None.

### References

- [1] R. Szibor, X-chromosomal markers: past, present and future, *Forensic Sci. Int.: Genet.* 1 (2007) 93–99.
- [2] Muqadima, Ibn Khaldoun. d'après les *Prolégomènes d'Ibn Khaldoun (1332-1406 de J.C)* traduits en français et commentés par W. Mac Guckin de Slane, 1801–1878.
- [3] Qiagen, Investigator Argus 12 X STR Kit Handbook, <http://www.qiagen.com/products/investigatorargusx-12kit.aspx#Tabs=t2>.
- [4] A. Carracedo, J.M. Butler, L. Gusmão, et al., Publication of population data for forensic purposes, *Forensic Sci. Int.: Genet.* 4 (2010) 145–147.

*Mitochondrial DNA control region  
variation from samples of the Moroccan  
population*

**Rachid Aboukhalid, Kimberly Sturk-  
Andreaggi, Mehdi Bouabdellah, Driss  
Squalli, Jodi A. Irwin & Saaid Amzazi**

**International Journal of Legal  
Medicine**

ISSN 0937-9827

Int J Legal Med

DOI 10.1007/s00414-012-0813-4



**Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

# Mitochondrial DNA control region variation from samples of the Moroccan population

Rachid Aboukhalid · Kimberly Sturk-Andreaggi · Mehdi Bouabdellah · Driss Squalli · Jodi A. Irwin · Saaïd Amzazi

Received: 12 September 2012 / Accepted: 17 December 2012  
© Springer-Verlag Berlin Heidelberg 2012

**Abstract** In an effort to facilitate forensic mitochondrial DNA (mtDNA) testing in Morocco, high-quality control region sequences from 509 individuals were generated using a comprehensive processing and data review system. This large dataset of random samples from various Moroccan population groups (Arab speaking, Berber speaking, and Sahrawi speaking) exhibited a low random match

probability (0.52 %) and a mean of pairwise comparisons of 13.24. The Moroccan mtDNA gene pool studied here was defined entirely by West Eurasian (58.15 %) and African haplogroups (41.85 %).

**Keywords** mtDNA · Control region · Morocco

**Electronic supplementary material** The online version of this article (doi:10.1007/s00414-012-0813-4) contains supplementary material, which is available to authorized users.

R. Aboukhalid (✉) · M. Bouabdellah · S. Amzazi  
Forensic Genetic Unit, Mohammed V Agdal University,  
Faculty of Sciences Rabat, BP 1014 Avenue Ibn Batouta Agdal,  
Rabat, Morocco  
e-mail: a.rachid22@gmail.com

M. Bouabdellah  
e-mail: bou.mehdi@gmail.com

S. Amzazi  
e-mail: amzazi@gmail.com

R. Aboukhalid · D. Squalli  
GENOME Biotechnologies Parc d'activités Oukacha imm.,  
E N° 13 Ain Sebaâ,  
Casablanca, Morocco

D. Squalli  
e-mail: dr.squalli@gmail.com

K. Sturk-Andreaggi · J. A. Irwin  
American Registry of Pathology, Camden, DE, USA

K. Sturk-Andreaggi  
e-mail: kimberly.andreaggi@us.army.mil

J. A. Irwin  
e-mail: irwinjodi@gmail.com

K. Sturk-Andreaggi · J. A. Irwin  
Armed Forces DNA Identification Laboratory, Dover, DE, USA

## Introduction

The current day population of Morocco reflects both the country's long and rich population histories on the African continent, as well as more recent demographic features resulting from its geographic proximity to Europe. Inhabited since ancient times by Berbers, the region became a Roman province in the first century and was conquered by Arabs in the seventh century. The country was later united (11th–13th century) under Berber-Muslim dynasties; yet, the most recent influences come from Portugal, Spain, and France. Morocco is the third most populous Arab country after Egypt and Sudan, with a population described as over 99 % Arab-Berber. Previous studies of Moroccan mitochondrial DNA (mtDNA) have detected a highly admixed gene pool, with the majority of lineages belonging to west Eurasian haplogroups [1–4]. Here, we expand upon these mtDNA studies by increasing both the sample size and sequence range available for the Moroccan population.

## Population samples

Blood and saliva samples were collected from 509 unrelated individuals living in Morocco. In addition to 345 individuals representing a number of ethnic groups (Arab, Berber, and Sahrawi), 42 individuals from the Arab ethnic group, 69

individuals from the Berber ethnic group, and 53 individuals from the Sahrawi ethnic group were sampled. Informed consent was obtained from all participants in this study, and information about the geographical origin of their grandparents and about their first language was recorded.

### Sample processing

DNA from whole blood samples was manually extracted using the DNA IQ™ System (Promega, Madison, WI) according to manufacturer's instructions. The DNA IQ™ extraction of the saliva samples, collected with Buccal DNA Collectors™ (Bode Technology Group, Inc., Lorton, VA), was automated on the Biomek® 2000 Laboratory Automation Workstation (Beckman, Coulter, Brea, CA). PCR amplification and amplicon sequencing reactions were performed as described in the study of Irwin et al. [5] with minor modifications to the primers. Electrophoresis was performed on a 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA).

### Analysis of data

Electropherograms were evaluated in Sequencher version 4.7 (Gene Codes, Ann Arbor, MI) by two independent scientists. The nucleotide positions considered for analysis were 16,024–16,569 and 1–576. They were aligned and compared to the revised Cambridge Reference Sequence following nomenclature guidelines for mtDNA typing [6]. Sequences were assigned to mitochondrial DNA haplogroups on the basis of control region polymorphisms reported to be associated with particular mtDNA lineages [7]. In some cases, this assignment is conservative and reflects the broadest level of classification that can be determined with control region information alone. Haplotypes likely belong to one of the many sub-lineages within the assigned haplogroup, but coding region data would be necessary for this level of classification.

Genetic diversity indices, random match probability, and pairwise comparisons were generated by the custom software package (Laboratory Information Systems Applications, Future Technologies, Inc., Fairfax, VA). Hypervariable C-stretch length variants at sites 16193, 309, and 573 were ignored in the analyses.

### Access to data

Sequence data in electronic form are available from the authors upon request and have been submitted to the EMPOP database (<http://www.empop.org>). Haplotypes are available under the following accession numbers: Moroccan Arab EMP00280, Moroccan Berber EMP00281, Moroccan

Sahrawi EMP00282, and Moroccan Mixed EMP00283. Data are also available under GenBank accession numbers JF749296–JF749804.

### Results

Summary statistics and entire control region haplotypes for all individuals are presented in supplementary data (Tables S1 and S2, respectively). Direct pairwise comparisons among the sequences in this dataset resulted in an empirically determined random match probability of 0.52 %. On average, sequences differed at 13.24 sites, and in total, 88.97 % (371 out of 417) were unique haplotypes.

Point heteroplasmies were identified in 43 individuals (8.45 % of database), and 1 sample exhibited two positions of point heteroplasmy (16182 and 16256). Length heteroplasmy was observed in 291 individuals (57.17 % of the database) and occurred primarily in regions of polycytosine stretches: hypervariable (HV)I (19.06 % of the database), HVII (47.35 %), and HVIII (1.77 %). Length heteroplasmy in the AC repeat region was observed in six individuals. One sample exhibited a rare 204-bp duplication of nucleotide positions 360–563. This duplication has been previously reported [8] and, when present as the major component of the heteroplasmic mixture (as it is in the Moroccan sample), has been associated with a specific non-synonymous change in the cytochrome *b* gene. However, upon further sequencing of the cytochrome *b* gene, the variant in question was not found in this particular sample. A number of novel polymorphisms were also observed in the dataset, with one sample showing the previously undescribed insertion 524.1 G 524.2C.

In the complete Moroccan dataset, the most frequently observed haplotype was 16519C 263 G 315.1C (a member of Hg R0 or its subhaplogroups). This haplotype, along with its C-stretch variants, was seen altogether in 28 samples (5.50 % of samples) and is consistent with the most common Moroccan lineage reported by Turchi et al. [4]. The second most common haplotype 16153A 16298C 72C 195C 263 G 315.1C (Hg HV0), and its C-stretch variants, was found in 11 samples (2.16 % of samples), of which 7 belonged to the Moroccan mixed population, 3 belonged to the Moroccan Berber, and 1 belonged to the Moroccan Sahrawi.

### Discussion

The mtDNA haplotypes from the four groups show a mosaic composition of West Eurasian, sub-Saharan, and North African lineages with frequencies of 58.15, 30.65, and 11.20 %, respectively. Generally speaking, this pattern is consistent with previous reports of Moroccan mtDNA [1–4],

though the data presented here exhibit slightly fewer West Eurasian lineages (58.15 vs ~63 %) and slightly more sub-Saharan lineages (30.65 vs ~25 %). More specifically, the most common haplogroups in the Moroccan mtDNA pool studied here were members of R0 (17.49 %) and HV0 (8.25 %), followed by African lineages U6a (7.27 %), L1b (5.11 %), L2a1 (3.73 %), M1 (3.34 %), and L3e5 (3.34 %). The occurrence of clades M1 and U6 in our dataset corresponds well with previous reports of their distributions in the east and north of Africa. These two lineages are believed to have originated in southwestern Asia and then moved collectively to Africa approximately 45,000 years ago [9]. This migration, followed by periods of stability and expansion, has led to relatively high frequencies of U6 and M1 in northern and eastern Africa, with limited distributions elsewhere.

The entire mtDNA control region sequences from the Moroccan population groups studied here are available for forensic DNA purposes. This paper follows the recommendations of the International Society for Forensic Genetics on the use of mtDNA in forensic analysis and the guidelines for publication of population data requested by the journal [10].

**Acknowledgments** We would like to express our sincere gratitude to the Moroccan-American Commission for Educational and Cultural Exchange (MACECE) who, together with the Fulbright Program, offered R. A. the opportunity to perform this research. We are also indebted to Jessica Saunier, Toni Marie Diegoli, Melissa Scheible, and numerous others for data review, as well as Dr. Michael Coble, COL Louis Finelli, James Canik, Dr. Brion Smith, Rebecca Just, the Armed Forces Medical Examiner System, and the American Registry of Pathology for logistical and administrative support. I would like to express my deepest gratitude to Dr. Mohammed Mzibri who offered his continuous advice and encouragement. Finally, thanks to

GENOME Biotechnologies for their ongoing support. Portions of this work were funded by the US National Institute of Justice through Interagency Agreement 2005-DN-R-086.

## References

1. Brakez Z, Bosch E, Izaabel H et al (2001) Human Mitochondrial DNA sequence variation in the Moroccan population of the Souss area. *Ann Hum Bio* 28:295–307
2. Plaza S, Calafell F, Helal A et al (2003) Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. *Ann Hum Genet* 67:312–28
3. Coudray C, Olivieri A, Achilli A et al (2008) The complex and diversified mitochondrial gene pool of Berber populations. *Ann Hum Genet* 1469–809
4. Turchi C, Buscemi L, Giacchino E et al (2009) Polymorphisms of mtDNA control region in Tunisian and Moroccan populations: an enrichment of forensic mtDNA databases with Northern Africa data. *For Sci Int Genet* 3:166–72
5. Irwin J, Saunier J, Strouss K et al (2007) Development and expansion of high quality control region databases to improve forensic mtDNA evidence interpretation. *For Sci Int Genet* 1:154–7
6. Bär W, Brinkmann B, Budwole B et al (2000) DNA commission of the International Society for Forensic Genetics: guidelines for mitochondrial DNA typing. *Int J Legal Med* 113:193–6
7. Van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30(2):E386–94. <http://www.phylotree.org>. doi:10.1002/humu.20921
8. Bouzidi M, Poyau A, Godinot C et al (1998) Co-existence of high levels of a cytochrome b mutation and of a tandem 200 bp duplication in the D-loop of muscle human mitochondrial DNA. *Hum Mol Genet* 7:385–91
9. González A, Larruga J, Abu-Amero K et al (2007) Mitochondrial lineage M1 traces an early human backflow to Africa. *BMC Genomics* 8:223
10. Parson W, Lutz R (2010) Publication of population data of linearity inherited DNA markers in the International Journal of Legal Medicine. *Int J Legal Med* 124:505–9

## **Annexes**

## Annexe A

### Questionnaire

Université Mohamed V-Agdal  
Faculté des Sciences Rabat  
Département de Biologie  
Laboratoire de Biochimie Immunologie

N° du prélèvement :

Date de prélèvement :

**PROJET D'ETUDE GENETIQUE DE LA POPULATION MAROCAINE  
QUESTIONNAIRE DE PRELEVEMENT SANGUIN OU BUCCAL POUR  
EXTRACTION D'ADN**

Le progrès qu'a connu le projet du génome humain a permis de localiser des milliers de marqueurs polymorphes, qui sont aujourd'hui utilisés pour l'étude de la structure génétique des populations humaines. Parmi les plus importants, les séquences microsatellites sont devenues un outil très puissant dans plusieurs domaines et notamment la génétique des populations. Le grand polymorphisme de ces marqueurs sera mis à profit dans cette étude pour la caractérisation génétique de la population marocaine et l'investigation de son histoire évolutive

#### I – VOTRE IDENTIFICATION

N° Individu : \_\_\_\_\_ Date de prélèvement: \_\_\_\_\_ N° ADN : \_\_\_\_\_

Initiales : \_\_\_\_\_

Sexe : M :  F :

Prénom : \_\_\_\_\_

Age : \_\_\_\_\_

Lieu de naissance : \_\_\_\_\_ Origine ethnique : \_\_\_\_\_

Téléphone : \_\_\_\_\_

Email : \_\_\_\_\_

#### II – L'IDENTIFICATION DES PARENTS

Initiales du nom de la mère : \_\_\_\_\_

ORIGINE ETHNIQUE (Arabe, berbère, Sahraoui) \_\_\_\_\_

LANGUE D'ORIGINE:  Dialecte marocain  Sahraoui  
 Berbère Tachelhit  Berbère Tarifit  Berbère tamazight

Initiales du nom du père: \_\_\_\_\_

ORIGINE ETHNIQUE (Arabe, berbère, Sahraoui) \_\_\_\_\_

LANGUE D'ORIGINE:     Dialecte marocain     Sahraoui  
 Berbère Tachelhit     Berbère ~~Tarifit~~     Berbère tamazight

<b>III – L'IDENTIFICATION DES GRANDS PARENTS</b>
--

Initiales du nom de la grand mère: \_\_\_\_\_

ORIGINE ETHNIQUE (Arabe, berbère, Sahraoui) \_\_\_\_\_

LANGUE D'ORIGINE:     Dialecte marocain     Sahraoui  
 Berbère Tachelhit     Berbère ~~Tarifit~~     Berbère tamazight

Initiales du nom du grand père: \_\_\_\_\_

ORIGINE ETHNIQUE (Arabe, berbère, Sahraoui) \_\_\_\_\_

LANGUE D'ORIGINE:     Dialecte marocain     Sahraoui  
 Berbère Tachelhit     Berbère ~~Tarifit~~     Berbère tamazight

Les données recueillies durant cette étude serviront strictement pour celle-ci et pour l'interprétation des résultats obtenus, elles resteront confidentielles et ne seront absolument pas utilisées pour quelconque autre objectif.

**Signature :**

---

## Annexe B

### Certificat contrôle qualité de YHRD pour le typage des 17 locus STR du chromosome Y



Rachid ABOUKHALID  
Unité de Génétique Forensique  
Laboratoire de Biochimie Immunologie  
Département de Biologie  
Faculté des Sciences de Rabat  
4, Avenue Ibn Battouta, B.P 1014 RP  
Rabat, Maroc

Institute of Legal Medicine and Forensic Sciences  
Charité – University Medicine Berlin  
Hannoversche Strasse 6  
D-10115 Berlin, Germany

Dept. Forensic Genetics  
Chair: Lutz Roewer, PhD  
Tel: ++49 (030) 450 525032  
Fax: ++49 (030) 450 525912  
e-mail: lutz.roewer@charite.de

Internet Y-STR database: [www.yhrd.org](http://www.yhrd.org)

Berlin, August 1, 2008

#### Certificate of participation

The

**Unité de Génétique Forensique, Laboratoire de Biochimie Immunologie  
Département de Biologie, Faculté des Sciences de Rabat, Maroc**

has participated in the Y-STR Haplotyping Quality Assurance Exercise 2008.

The alleles at the YHRD core loci

DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385ab,  
DYS438, DYS439

as well as the additional loci

DYS437, DYS448, DYS456, DYS458, DYS635, GATA-H4

have been typed correctly according to the published nomenclature and the updated ISFG recommendations for Y-STR analysis (Forensic Sci Int. (2008) 157(2-3):187-97).

Lutz Roewer  
YHRD curator

## Annexe C

**Séquences des amorces sens et anti-sens des quatre multiplex utilisés pour l'analyse des marqueurs SNP du chromosome Y de notre population étudiée (n = 159)**

Marqueurs	Amorces PCR Sens	Amorces PCR Anti-Sens	Taille Amplicon (bp)	Concentration finale (pmol)	
MY01	M91	AGAAACACACTGCTAATGATCCG	GGCAGTGCCCTTCCAAATAA	295	0.2
	M60	AAGACTAGGAGTAGGGTTGTGGAC	AAGACTAGGAGTAGGGTTGTGGAC	180	0.4
	RPS4Y711	GAAGGAAGGTATGAGGCATGTGTG	CAGTAAGTCGAATGCCCTTTCC	98	0.2
	M174	CGTCTCTCTGAATACCTTCTGGAG	TTACCTGACTAGAAAGGTCCTGG	145	0.6
	M267	GGATATACCAAGTCTGGATAGCGG	CTCGTAGCTTCAAGCAATCCAC	147	0.6
	M170	CATGTTTGTTCAAATAATTGCAGCTC	TACAGTGAGACACAACCCACAC	129	0.4
	M207	TCAGAAGTATCCCTGAAGAAGGA	GGTCACTTCAACCTCTTGTGG	125	0.8
MY02	P147	CTCTCCCTACTACTGCATTCTAGC	CTAATTTCCAGCAGACAGCTC	248	0.12
	M132	GTCTCAGTCACAGCAATTCTATCC	GTTCTGTACTTTCAATTAATCACCTCT	285	0.4
	P189	GGTTATGAAGAGCTGAACACTG	CAGGTAGTATAGCTAGATTCTGGTCC	142	0.08
	M215	ACATACTTGCTGCATTAAGACAAAC	CTGTAGTCCCAGCTACTCAGAA	274	0.8
	M81	AGCACTATCATACTCAGCTACAC	CGCACTAATTTGTAGCCTAGGAGC	161	0.2
	M54	CAGAGCAGTCACATCTTCAGAG	CAAAGCCCAGTCTCACTTCATC	245	0.12
	M172	CTCATTACCTGCCTCTCAGTATC	GCCAGGTACAGAGAAAGTTTGG	144	0.08
MY03	L24	CTGGACTGCATACATTAGTGGG	CATCTGAACTGCTCTTGGAGTC	86	0.08
	M221	GGAAAGGTCAAAC TGGGTACTG	ACACTGACAGAACTCACTGAT	260	0.6
	L63	CACTATGTCACCATCAACATTGCC	GGGATAAGAAATAGTTTGTAGCTGAGG	74	0.2
	M343	CTTGTAGAGAGAGAGAGGTAGGAG	GCGTTAGAACTGGGTTCACTC	324	0.6
	M269	CATGCCTAGCCTCATTCTCTA	GGGTTTCACCATGTTAGCCT	267	0.8
	M70	CAGCCTCCAGAATTCCTTCTCTAC	CATGTATGTGATCTGAGGGCTG	160	0.08
MY04	M35	TACTCAGGAGGAACTGAGAGG	CACTAAGGAGCTGTGGTGAATG	146	0.8
	M78	CAGTGGTTTCTGCATTAATCCG	GCTATAGTGTTCCTTCACCTTTCC	180	0.2

## Annexe D

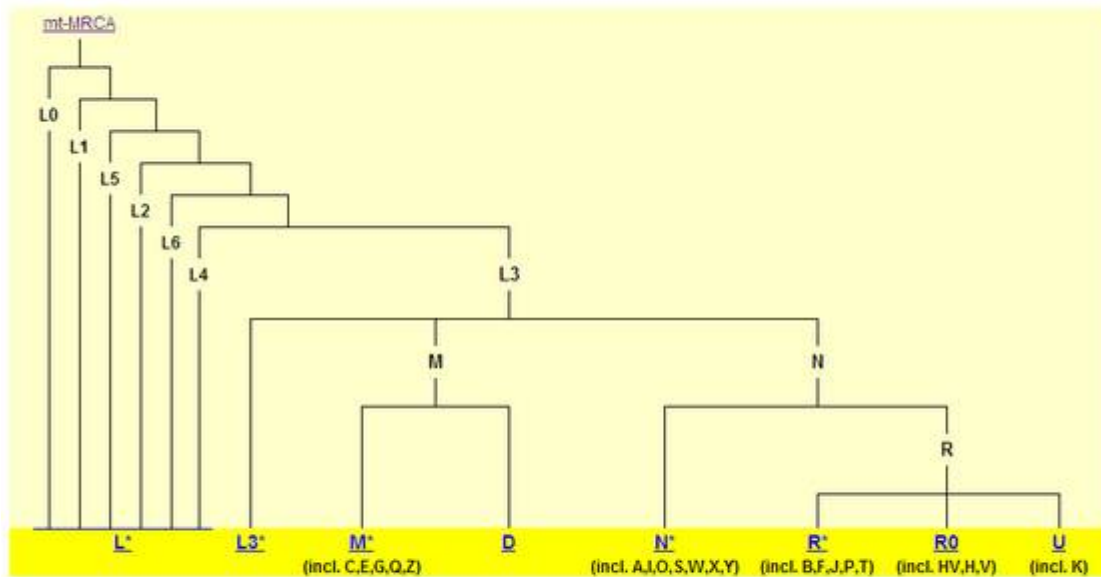
### Séquences des amorces d'extension des quatre multiplex pour les 21 marqueurs SNP du chromosome Y incluant les queues homopolymère T

	Marqueurs	Séquence d'extension d'amorces	Brin	Longueur (pb)	SNP (Ancestral/Muté)
msMY01	M174	GAATACCTTCTGGAGTGCCC	F	20	C à T
	RPS4Y711	(T) <sub>5</sub> GGGCAATAAACCTTGGATTTC	F	26	C à T
	M60	(T) <sub>5</sub> GAGTAACCACTGTGTGCCTGAT	R	31	1 bp (-/+)
	M267	(T) <sub>13</sub> CGATGGAAGCATTTTTGTAAATA	F	36	G à T
	M91	(T) <sub>16</sub> CCTACATGCTATTCTGTTTTTTTT	F	41	1 bp (-/+)
	M207	(T) <sub>20</sub> CAAATGTAAGTCAAGCAAGAAATTTA	F	46	A à G
	M170	(T) <sub>25</sub> TTACTATTTTATTTACTTAAAAATCATTGTTC	F	57	A à C
msMY02	M54	CTCAGGCAGCCGCAC	F	15	A à T
	P147	(T) <sub>1</sub> CCAAAAGTTCTCTTCCTG	R	20	A à G
	M172	(T) <sub>5</sub> CCAAACCCATTTTGATGCTT	F	25	A à G
	P189	(T) <sub>7</sub> AAATTGCTCATAAATCAAAGAC	F	30	G à T
	M215	(T) <sub>11</sub> TACTTGCTGCATTAAGACAAACTG	R	35	G à T
	M132	(T) <sub>13</sub> GGTAAAACAGATAGAAAAAAGCTATG	F	40	A à G
	M81	(T) <sub>18</sub> TTGGTTTGTGTGAGTATACTCTATGAC	R	45	C à T
msMY03	M343	AGTGCCCTCGTGTCCA	F	17	A à G
	L63	(T) <sub>1</sub> CAACATTGCCATTGCTTTAGT	F	22	A à G
	M70	(T) <sub>6</sub> GGGATTCTGTTGTGGTAGTCTTAG	R	27	C à T
	M269	(T) <sub>10</sub> GAATGATCAGGGTTTGGTTAAT	F	32	A à C
	L24	(T) <sub>13</sub> GATGTCAATTGTTAGAAGGGTAG	F	37	C à T
	M221	(T) <sub>16</sub> ATATTATCACTTGCATCAGTGCTAAA	F	42	G à T
msMY04	M35	CAATTTTCCTTTGGGACACTG	F	21	G à C
	M78	(T) <sub>7</sub> TATTTTGAAATATTTGGAAGGGC	R	30	C à T

## Annexe E

### PhyloTree.org - mtDNA tree Build 15 (30 Sep 2012)

For your convenience, the online tree is divided into eight subtrees accessible through the links in the scheme below. Alternatively, the entire tree can be downloaded as single file: [mtDNA tree Build 15 \(30 Sep 2012\)](#)



PhyloTree collaborates with [EMPOP](#) and [mtDNACommunity](#) to enrich our knowledge of the human mtDNA phylogeny in its evolution and present-day variation.

Maintained by Mannis van Oven (m.vanoven@erasmusmc.nl)  
Department of Forensic Molecular Biology  
Erasmus MC, University Medical Center Rotterdam  
The Netherlands