Université Mohammed V
AGDAL - RABAT

UIB
Universitat de les Illes Balears

GENETIC PROFILE OF WESTERN MEDITERRANEAN POPULATIONS: CONTRIBUTION OF ARAB AND JEWISH GROUPS

Kaoutar Bentayebi

Ph D Thesis, 2012

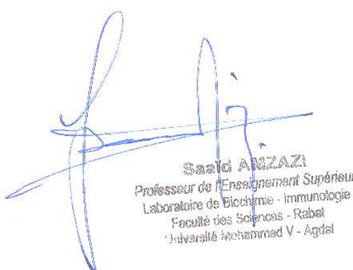
Université Mohammed V
AGDAL - RABAT

UIB
Universitat de les Illes Balears

GENETIC PROFILE OF WESTERN MEDITERRANEAN POPULATIONS: CONTRIBUTION OF ARAB AND JEWISH GROUPS

Thesis submitted by Miss **Kaoutar Bentayebi** to aspire to the degree of Doctor in Biotechnology, Genetics and Cellular Biology (at the University of Balearic Islands) and in Biology (at the Mohamed V-Agdal University)

Thesis supervisors: Dr. Antònia Picornell Rigo and Dr. Saaid Amzazi



Dr. Antònia Picornell Rigo | Dr. Saaid Amzazi | Kaoutar Bentayebi

Rabat-Palma de Mallorca 2012

RESUME

Ce travail décrit la diversité génétique des populations ouest-méditerranéennes actuelles à travers deux polymorphismes : neufs marqueurs Alus et douze microsatellites (STR) présents au niveau du chromosome X. Dans ce travail, des résultats originaux de populations Marocaines, Espagnoles, sud Italiennes et de populations Juives sont présentés. Notre étude multidisciplinaire s'appuie sur des données biologiques, archéologiques, historiques, géographiques et linguistiques pour retracer les origines et l'histoire génétique des populations ouest-Méditerranéennes. Pour l'ensemble des marqueurs, nos résultats montrent une proximité génétique entre les populations Nord Africaines et les populations du sud-ouest de l'Europe mais une différenciation entre les groupes nord-africains et sub-sahariens. Aussi, nous constatons qu'au nord-ouest de l'Afrique, aucune importante différenciation génétique entre les Berbères et les Arabes n'apparait. L'analyse de cinq populations juives les a groupé dans un même cluster, témoignant de leur origine ancestrale commune qu'il ont conservé à travers le temps malgré la Diaspora, avec une nette distinction entre eux et leurs voisins non-juifs.

MOTS-CLEFS

Ouest Méditerranéenne, Anthropologie, génétique des populations, Forensique Génétique, Arabes, nord de l'Afrique, Diaspora juive, polymorphismes génétiques, STR, Alu, linguistique, histoire génétique, peuplement, Iles Baléares.

RESUMEN

Este trabajo describe la diversidad genética de las poblaciones actuales del Oeste del Mediterráneo según dos polimorfismos: nueve polimorfismos Alus y doce microsatelites (STR) presente en el cromosoma X. En este trabajo, se presentan resultados originales de poblaciones de Marruecos, España, sur de Italia y de poblaciones judías. Nuestro estudio multidisciplinario se basa sobre datos biológicos, arqueológicos, históricos, geográficos y lingüísticos para reconstruir los origines y la historia genética del oeste del Mediterráneo. Para la totalidad de los marcadores, nuestros resultados muestran una proximidad genética entre las poblaciones del norte de África y del sur de Europa con una diferencia entre los grupos norte africanos y sub-

saharianos. También se puede concluir que en el noroeste de África, no hay una fuerte diferenciación genética entre los beréberes y los árabes. .El análisis de cinco poblaciones judías muestra que se agrupan en un mismo cluster, indicando su ancestral común origen, que se ha conservado durante el tiempo a pesar de la diáspora, con una clara distinción entre ellos y sus vecinos no-judíos.

PALABRAS CLAVES

Oeste del Mediterráneo, Antropología, genética de población, genética forense, Árabes, Norte de África, Diáspora judía, polimorfismos genéticos, STR, Alu, lingüístico, historia genética, población, Islas Baleares.

ABSTRACT

This research describes the genetic diversity of current Western Mediterranean populations through two polymorphisms: nine Alus polymorphisms and twelve Short Tandem Repeats (STR), located in the X chromosome. Original results are presented for populations from Morocco, Spain, South of Italy and Jewish Diaspora. Our multidisciplinary study is based on biological, archaeological, historical, geographical and linguistic data in order to track the roots of the Western Mediterranean origins and genetic history. For all markers, our results show that the North African are genetically close to European populations but differentiated from sub-Saharan groups. Furthermore, we notice that in North-West Africa, there isn’t any striking genetic differentiation between Berber and Arabic. The analysis of five Jewish populations plotted them in the same cluster, arguing for their shared ancestral origins that they conserved over time despite the Diaspora, with a clear distinction found between them and their non-Jewish neighbors

KEY WORDS

Western Mediterranean, Anthropology, Population Genetics, Forensic Genetics, Arabic, North Africa, Jewish Diaspora, genetic polymorphisms, STR, Alu, linguistics, genetic history, settlement, Balearic Islands.

A mes chers parents….

A la mémoire de Papy….

A ma famille et ami(e)s…

PREFACE

The present work was supported by an Averroes fellowship, which was awarded to me in my last two years. This has led to the collaboration and the co-supervision of my PhD work through Université Mohammed V - Agdal and Universidad de las Islas Baleares. In this context, I would like here first to convey my heartfelt thanks to both Universities for offering me an ideal environment in which I felt free and could concentrate on my research.

I am grateful for all the institutions and co-authors for their help, availability and orientation that made my PhD Work and publication possible: My supervisors “Dr. Saaïd AMZAZI and Dr. Antònia PICORNELL”, Dr. Misericòrdia Ramon, Dr. José Aurelio, Dr. Driss Squalli, Dr. Anna Barbaro, Mehdi Bouabdellah and Rachid Aboukhalid.

Indeed, this work leads to the following publications:

Published papers

Bentayebi K, Picornell A, Amzazi S. Genetic Diversity: New research: Chapter 5: Genetic History of Morocco. Novapublishers. USA. 2012.

Bentayebi K, PicornellA ,Bouabdeallah M, Castro J A, Aboukhalid R , Squalli D, Misericordia MM , Amzazi S: Genetic Diversity of 12 X-chromosomal short tandem repeats in the Moroccan population, Forensic Sci. Int. Gent. 2012. 6: e48-e49.

Bentayebi K, Ramon MM, Castro JA, Barbaro A, Aboukhalid R, Amzazi S and Picornell A. Inferring ethnicity from the X-chromosome ALU insertions: data from Western Mediterranean human groups, Forensic Sci. Int. Gent. Supp Series. 2011: e27-e28.

Picornell A, **Bentayebi K**, Bouabdeallah M, Castro JA, Bouabdellah M, Squalli D, Misericordia M, Amzazi S. X-chromosomal haplotype frequencies of four linkage groups in a North african population, *Forensic Sci. Int. Gent*. Supp Series. 2011, e19-e20.

Aboukhalid R, Bouabdellah M, Abbassi M, **Bentayebi K**, Elmzibri M, Squalli D, Amzazi D, Amzazi S. Haplotype frequencies for 17 Y-STR loci (AmpFISTR Y-filer) in a Moroccan population sample. Forensic Sci. Int. Gent. 2009. 4: e73-e74.

Aboukhalid R, Bouabdellah M, Lai R, Bradford N, EL Ossmani H, Talbi J, **Bentayebi K**, Squalli D, Podini D and Amzazi S. Y chromosomal SNP analysis using the minisequencing strategy in a Moroccan population samples. *Journal of forensic research*.2010.1 (2). 1000113.

Submitted papers

Aboukhalid R, Andreaggi KS, Bouabdellah M, **Bentayebi K**, El Mzibri M, Squalli D, Irwin J.A, and Amzazi S. Mitochondrial DNA control region variation from samples of the Moroccan population.ForensicSciInt Gent. 2012. Submitted paper.

Current papers

Bentayebi K, Ramon M.M, Castro J.A, Barbaro A, Aboukhalid R, Amzazi S and Picornell A. Populationgenetic inference from X chromosome *Alu*insertions in Western-Mediterranean region (*current*).

Bentayebi K*, Ferragut JM*, Castro JA, Ramon MM, Amzazi S, Picornell A (* both authors contributed equally to this work).Gene Pool Structure of Jewish people as Inferred from the X-Chromosome (*current*).

Ferragut JM*, **Bentayebi K***, Castro JA, Ramon MM, Amzazi S, Picornell A (* both authors contributed equally to this work).Gene Diversity in Balearic Islands Genetic Isolates and Valencia using Argus X-12 STR (*current*).

Bentayebi K, Picornell A, Bouabdeallah M, Castro JA, Aboukhalid R, Squalli D, Ramon MM, Amzazi S : Genetic analysis of 18 X chromosomal SNP in the Moroccan population (*in preparation*).

Bouabdellah M, **Bentayebi K** et al. 48 autosomal Single nucleotide Polymorphisms for Moroccan Population studies and Forensic applications(*in peparetion*).

International Communications

Participation by two posters entitled « **Inferring ethnicity from the X-chromosome ALU insertions: data from Western Mediterranean human groups** » and "**X-chromosomal haplotype frequencies of four linkage groups in a North african population**" at24th Congress of the International Society for Forensic Genetics in Vienna, Autria: 29-03 Sep (2011).

Participation by a poster entitled "**Haplotype frequenies for 17 Y-STR loci (AmpFlSTRYfilerTM) in a moroccan population sample**" at the 23th Congress of the International Society for Forensic Genetics in Puerto Madero, Buenos Aires, Argentina: 14-18 Sep (2009).

Participation by a poster entitled « **Haplotypefrequencies for 17 Y-STR loci (AmpFlSTR®Y-filerTM) in a Moroccanpopulation sample**» at the 4th international congress of genetics and molecular biology and the 4th International Congress of Biotechnology, Ouarzazat, Morocco: 05-08 Nov (2008).

Patrimony

Morocco is like a tree, whose roots lie in Africa, but whose leaves breathe in European air.
(Le Maroc est comme un arbre, dont les racines se situent en Afrique, mais dont les feuilles en air europeen.)
—King Hassan II of Morocco (ruled 1961-1999)

Acknowledgement

Pursuing a Ph.D. project is both painful and enjoyable experience. It's just like climbing a high peak, step by step, accompanied with bitterness, hardships, frustration, encouragement and trust and with so many people's kind help. When I found myself at the top enjoying the beautiful scenery, I realized that it was, in fact, teamwork that gone me there. Though it will not be enough to express my gratitude in words to all those people who helped me, I would still like to give my many, many thanks to all these people.

First of all, I'd like to give my sincere thanks to my supervisor, prof. ***Saaïd AMZAZI****. Professor and dean of the Faculty of Sciences at the University of Mohammed V-Agdal. It has been an honor to be his PhD student. He has taught me, both consciously and unconsciously, how good biologist is done, I appreciate all his contribution of time, ideas, and funding to make my PhD experience productive and stimulating. The joy and enthusiasm he has for his work was contagious and motivational for me, even during tough times in the PhD pursuit. I am also thankful for the excellent example he has provided as a successful man biologist and professor. Since my first day as a graduate student at the university he has been successively my professor, my idol, my master's degree supervisor and my PhD supervisor as well as a very good friend.*

Special thanks are also given to prof. ***Antònia Picornell****. Professor and vice-dean of the Faculty of Sciences at the Universidad de las Islas Baleares. She is the one who accepted me as a PhD scholar, at the Laboratory of Geneticsof the UIB, for my research in 2010. Thereafter, she accepted me as her PhD student without any hesitation. She offered me so much advice, patiently supervising me, and always guiding me in the right direction. I've learned a lot from her at both, the professional and personal level, without her help I could not finished my dissertation successfully.*

I would like also to thank Prof. ***Misericòrdia Ramon****, the head of the Laboratory of Genetics at the Universidad de las Islas Baleares, that welcomed me in her laboratory. Her encouragement and help made me confident to fulfill my desire and to overcome every difficulty I encountered.*

I also appreciate and thank prof. ***José Aurelio Castro****. The head of the Department of Biology at the UIB, for his welcoming, advices and help. Your*

permanent smile, French songs and the cookies that you used to give me at the end of every workday rob out my stress and tiredness and offered me an ideal environment.

I would like to present my gratitude to prof. ***Youssef Bakri****. Professor at the faculty of Sciences – UMV Agdal. He was the first one who introduced me to the molecular biology, I learned a lot from him and I am very grateful. Thank you for being part of my reading and oral defense committee.*

Many thanks to prof. ***Touria Benazzouz****. Professor at the Faculty of Sciences – UMV Agdal. She was my professor of genetics and taught me the basics of DNA and genes. You accepted to be the president of my oral defense, thank you for this honor.*

A special thanks to prof. ***Abdelaziz Chafik****. Professor and dean of the "Faculté polydisciplinaire of El Jadida-Université Chouaib Doukkali". You agreed to take time out of your busy schedule to be participate in my reading and oral defense committee. Your truly honor me with your presence.*

I am presenting also my acknowledgement to my friends and colleagues in Morocco: Mehdi and Rachid as well as my colleagues at the UIB: Virginia, Marina, Joana M and Joana F for their collaboration and the good moment we spent together.

I gratefully acknowledge the so many friends, both in Rabat and Mallorca, who offered me so much kindly help. I give my sincere thanks to all these people.

To everyone that contributed to this work

Table of contents

Abstract
Preface
Acknowledgements
Table of contents
List of Tables
List of Figures
List of Abbreviations

GENERAL INTRODUCTION.. 1

Part I: BIBLIOGRAPHICAL REVIEW

A. The linguistic, historic and cultural context of Mediterranean populations

I. Geographical and cultural context of the Mediterranean region…………… 7
II. Historical and demo-linguistic data of Morocco, Balearic Islands and south of Italy………………………………………………………………………. 15

B. Religious diversity between populations of interest

I. Insight of the religious diversity between South of Europe, North Africa and Middle East……………………………………………………………………… 21
II. The historical and geographical movement of the Jewish communities….. 24

C. Concepts

I. Anthropology: Concept and History……………………………………… 29
II. Population Genetics: Concept and History……………………………….. 30
1. Definition………………………………………………………… 30
2. Hardy Weinberg equilibrium and evolutionary forces…………... 32

D. Usefulness of Forensic Genetics in Human Identification

I. DNA and Human Identification …………………………...……..……… 35
II. Usefulness of the different genetic markers in human identification…….. 37
1. Short Tandem Repeats…………………………………………… 37
2. Nomenclature of STR markers…………………………………... 39
3. Alu insertion polymorphism……………………………………... 40
4. Alu family classification………………………………………… 42
III. X Chromosome: properties and relevance in human identification and population genetics………………………………………………………………… 45
1. Use of Chromosome X in Forensic Genetics…………………… 49

Part II: MATERIAL & METHODS

A. Chromosome X: Genetic Analyses of 9 Alu polymorphisms

I. Material………………………………………………………… 53
II. Methods………………………………………………………… 54
1. DNA extraction…………………………………………………. 54
2. Quantification of isolated DNA……………………………… 54

3. Amplification of DNA with Polymerase Chain Reaction (PCR)… 55
4. Analysis of the PCR products……………………………………. 56
5. Statistical Analysis of Data………………………………………. 57

B. Chromosome X: Genetic Analyses of 12 STR polymorphism

I. Material……………………………………………………………… 63
II. Methods……………………………………………………………… 63
1. DNA extraction……………………………………………………… 63
2. Quantification of isolated DNA…………………………………… 64
3. Amplification of DNA with Polymerase Chain Reaction (PCR)…. 64
4. Capillary Electrophoresis…………………………………………. 68
5. Statistical Analysis of Data……………………………………….. 70

Part III: RESULTS & DISCUSSION

A. Genetic characterization of the Western Mediterranean population

I. Genetic diversity of 12 X-chromosomal short tandem repeats in the Moroccan population………………………………………………………………. 73

II. Gene Diversity in Balearic Islands Genetic Isolates and Valencia using Argus X-12 STR…………………..…………………………………………………… 83

III. Population genetic inference from X chromosome *Alu* insertions in Western-Mediterranean region……………………………………...…………… 97

B. Gene Pool Structure of Jewish people as Inferred from the X-Chromosome… 105

Part IV: GENERAL DISCUSSION

A. The ethno-religious origins of the Western Mediterranean region: North African and South Europeans populations……………………………………………………… 139

B. Jewish Diaspora: Origins and relationship with the Western Mediterranean populations……………………………………………..………………….… 145

Conclusion & Prospects……………………………………………………………….. 149

References.. 151

Appendix.. 165

List of Tables

Part I

Table 1. Comparison of population-genetic markers………………………………..….. 47

Part II

Table2.PCR Protocols: X chromosome Alu insertions………………………..…… 55

Table 3. Oligonucleotid primer sequences and sizes of the PCR products in the presence of each Alu insertion………………………………………….................. 55

Table 4. Thermocycling conditions for the 9 Alu polymorphism…………………... 56

Table 5. Chromosomal mapping of the Investigator Argus X-12 Kit……………… 65

Table 6. Locus-specific information of the Investigator Argus X-12 Kit………….. 67

Part III

Table 7.Allele frequencies and statistical parameters of 12 X-chromosomal short tandem repeat loci in 145 Moroccan samples (Arabic: 36 men and 18 women; Berber: 32 men and 16 women and Sahrawi: 29 men and 14 women)……………………………………………………………… 75

Table 8. Genetic distances between the Moroccan and other African, European, American and Asian populations (*F*st values)…………………….………….. 78

Table 9.X-chromosome haplotypes of four linked STR trios in 36 Arabic, 32 Berber and 29 Sahrawi males living in Morocco………………………………... 79

Table 10.Haplotype frequencies (F) of 4 linked X-chromosomal short tandem repeat trios in 97 Moroccan men…………………….………………………… 81

Table 11.Allele frequencies and statistical parameters of 12 X-chromosomal short tandem repeat loci in 250 Spanish samples (Majorca: 39 men and 22 women; Minorca: 39 men and 25 women; Ibiza: 39 men and 20 women and Valencia: 39 men and 27 women)…………………………………………….………… 84

Table 12. Genetic distances between the Moroccan and other African, European, American and Asian populations populations (*F*st values)………………………… 87

Table 13. Haplotype diversity of X-STR clusters in Ivorians (Tie et al. 2010), Algerians and Italians (Inturri et al. 2010 and Robino et al. 2006)………………….. 89

Table.14. X-chromosome haplotypes of four linked STR trios in 39 men from each Mojorca, Minorca, Ibiza and Valencia……………………………….. 91

Table 15. Allele frequencies and heterozygosity per locus and population for the X-chromosomal ALU insertions Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4, Ya5DP13 in 11 Western Mediterranean populations…………..……………………………. 98

Table 16. Reynods' distances for the eleven Western-Mediterranean populations 99

Table 17. Estimated natural logarithm of the probability of the data (Ln P(X/K)) for each value of K and proportion of membership of each pre-defined population in each of the five clusters……………………………...………………. 102

Table 18. Allele frequencies and heterozygosity per locus and population for the X-chromosomal ALU insertions Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4, Ya5DP13 in 5 worldwide Jewish populations………………………………………………… 108

Table 19. Reynods' distances for the five Jewish studied populations…………… 108

Table 20. Estimated natural logarithm of the probability of the data (Ln P(X/K)) for each value of K and proportion of membership of each pre-defined population in each of the five clusters…………………………………………… 111

Table 21.Allele frequencies and statistical parameters of 12 X-chromosomal short tandem repeat loci in 302 Jewish samples (Ashkenazis: 26 men and 33 women; Sephardic (Bulgaria & Turkey): 19 men and 25 women; North African (Morocco, Tunisia and Libya): 29 men and 25 women , Oriental (Iraq & Iran): 29 men and 26 women) and Chuetas: 49 men and 41 women)……………………………………………..……………….. 113

Table 22.X-chromosome haplotypes of four linked STR trios in males belonging to five different Jewish communities………………………..……………….. 120

Table 23. Genetic distances between the Jewish communities and other North African and European populations (*F*st values)……………………………………. 138

List of Figures

Part I

Figure 1. Map of the Western Mediterranean region……………………………... 2
Figure 2. The Jewish Diaspora between 70 to 1497……………………………… 24
Figure 3. The Jewish migration pattern and worldwide distribution…..…………. 27
Figure 4. Short Tandem Repeats structure……………………………………... 37
Figure 5.STR allele’s structure……………………………………………... 39
Figure 6.A typical human Alu element and its retroposition…………………….. 41
Figure 7. Expansion of recently integrated human Alu subfamilies…………….... 44
Figure 8. Chromosome X ideogram…………………………………………… 46
Figure 9. Variation between loci…………………..………………………….... 48

Part II

Figure 10. Photograph of a 2 % agarose gel containing the PCR products of Ya5DP62……………………………………………………………... 57
Figure 11. Photograph of a 2 % agarose gel containing the PCR products of Yb8NBC578, Yd3JX437, Ya5NBC491……………………………...…………… 57
Figure 12. The ideogram of the X-chroosome. Distances from the p-telomere are shown in Mb………………………………………………………………….. 66
Figure 13. Electropherogram of the DNA Size Standard 550 (BTO), fragments with lengths in bp…………………………………………………….…………. 69
Figure 14. Electropherogram of the allelic ladder Argus X-12 analyzed on an ABI PRISM 310 Genetic Analyzer…………………………………..……… 70

Part III

Figure 15. Multidimensional scaling analysis (MDS) based on Reynolds genetic distances calculated between Western Mediterranean populations for 9 ALU insertions…………………………………………………………….. 90

Figure 16. Multidimensional scaling analysis (MDS) based on Reynolds genetic distances calculated between populations………..…………………..… ……... 100

Figure 17. Multidimensional scaling analysis (MDS) based on Reynolds genetic distances calculated between five Jewish groups for 9 ALU insertions……… 109

Figure 18. Multidimensional scaling analysis (MDS) based on Reynolds genetic distances calculated between 24 populations for 9 ALU insertions…..……… 109

Figure 19. Multidimensional scaling analysis (MDS) based on Reynolds genetic distances calculated between five 21 populations for 9 ALU insertions……... 110

Figure 20. The triangle plot of Q. Each individual is represented by a colored point. The colors correspond to the prior population labels. The estimated ancestry vector for an individual consists of K components which add up to 1. For K=5, the ancestry vectors are plotted onto a triangle. For a given point, each of the five components is given by distance to one edge of the triangle………………………………………..………………. 111

Figure 21. Population structure inferred by STRUCTURE analysis. Each individual is represented by a vertical (100%) stacked column of genetic components proportions shown in color for K=3…... 112

Figure 22. Multidimensional scaling analysis (MDS) based on pairwise genetic distances calculated between 13populations for 12 X-STR……….............. 117

List of Abbreviations

AMOVA	Analysis of Molecular Variance
ARB	Arabs
AS	Autosomal
ASH	Ashkenazi
BC	Before Christ
BP	Before Present
BRB	Berbers
CE	Christian Era
Chr	Chromosome
CHU	Chuetas
CODIS	Combined DNA Index System
DNA	DeoxyriboNucleic Acids
GD	Gene Diversity
H	Haplotype Diversity
Het_{obs}	Observed Heterozygosity
Het_{exp}	Expected Heterozygosity
HWE	Hardy-Weinberg Equilibrium
IBD	Identical By Descent
IB	Ibiza
LD	Linkage Disequilibrium
LIA	Little Ice Age
MA	Majorca
Ma	Mega Annum (one million years, geological chronology)
MDS	Multidimensional Scaling
MEC_{duo}	Mean Exclusion Chance in trios involving daughter
MEC_{trio}	Mean Exclusion Chance in father/daughter or mother/son duos
MI	Minorca
MO	Pooled Moroccan Population
MWP	Medieval Warm Period
MENA	Middle East and North Africa
MtDNA	Mitochondrial DNA
N	Size

NAFR	North African Jews
ND	Not Done
ORT	Oriental
PCR.	Polymerase Chain Reaction
PD	Power of discrimination
PE	Power of Exclusion
PIC	Polymorphism Information Content
RSQ	Squared Correlation
SHR	Sahrawi
SNP	Single Nucleotide Polymorphism
STR	Short Tandem Repeat
VL	Valencia
X-Alu	X chromosome Alu
X-STR	X chromosome STR

GENERAL INTRODUCTION

Western Mediterranean region is the junction point between Africa on the South, Europe on the North, the Atlantic Ocean on the west and Euro-Asia on the East. Complex interplay between geographical and ongoing climatic processes has been pivotal and influenced the patterns of geographical diversification that are currently observed in these human populations. Yet, the Western Mediterranean region is settled now by many populations, different by their origins, culture and history. Track the population's root and understand their evolutionary processes have been always fascinating subjects for anthropologists since the XXth century. The recent progress of biology allowed a clear description to the genetic diversity of human populations by a genetic reconsideration of their history. These information, associated to the archeological, paleontological and linguistic data, give an inevitably possibility to rebuild the human evolutionary history. This is the main challenge of the anthropologist.

Our thesis is carried out in this context within a multidisciplinary approach. In fact, the evolutionary history of these human groups up today, remains related to the prehistorically and historical events evidenced by the archeological and paleontological shadows.

Within the same issue of Western Mediterranean population, certain human groups have a special interest: Romans, Arabic, Berbers and Jews moved, settled and had been expended for different reasons, including for a commercial and political account.

The Mediterranean was one of the world's greatest trading seas. At its eastern end, sometime in the second millennium BC, Minoans, Greeks and Phoenicians set forth. After the 8th century BC, Phoenicians from Tyre in what is now southern Lebanon moved into the Western Mediterranean.

Figure 1.
Map of the Western Mediterranean region.

The Iberian Peninsula had silver and tin and better watering places than North Africa. Since it was a long way from Tyre, they soon established a line of settlements on the shores and islands that lay between. The greatest of these settlements was Carthage, in what is now Tunisia. It was founded around the end of the ninth century BC, according to tradition, in 814 BC. Other settlements followed, including Rusaddir (now Melilla) on the Mediterranean coast, and Lixus, near what is now Larache, and near the modern town Essaouira, on the Atlantic. The Phoenicians had braved the Strait of Gibraltar and pushed southwards down the coast of Africa. The earliest traces of occupation at Lixus go back to the seventh century BC, but it is uncertain how much further the Phoenicians went. In the fifth century BC, Persian armies overran the eastern coast of the Mediterranean, and cut off the western settlements from the old metropolis at Tyre. Carthage became the pre-eminent Phoenician city and began to expand its influence westwards. A literary account, now known as the *Periplus* of Hanno, describes a trip between 475 and 450 BC which, it is sometimes claimed, reached the Gulf of Guinea. It may only have reached Essaouira or perhaps Dakhla on the modern Mauritanian coast. What really spread Carthaginian influence into the African interior was war with the Greek city states, particularly in Sicily. The war lasted, almost continuously, for over a century and the booty and the Greek prisoners who were taken to Carthage as slaves made the city extremely wealthy. This took Carthage into the mainstream of Mediterranean civilization, which was largely Greek.

Also, the Arabic conquest to North Africa (VIIth century) and then to the Iberian Peninsula (in the year 711) and Balearic Islands, was a crucial period in the Western Mediterranean history. This movement generated a deep change in culture, religion and partially in language.

The aim of this project is to describe the genetic diversity of the Western Mediterranean population and to assess their kinship degree with neighbor populations with a different culture and language. We tried to answer another challenging question in this study, concerning the different Jewish populations and their kinship level, after the long scattering they undergone. In the other hand we tried to compare them with the Western Mediterranean population where some of these groups settled for many centuries after their exile from Judaea by Titus, or even by Nebuchadnezzar.

In this work, the description of the Western Mediterranean genetic structure is based on the investigation of two polymorphisms: Alu insertion polymorphisms and Short Tandem repeats on the X chromosome. The STR as well as Alu markers showed widely their great interest in population genetics in order to compare human populations in different geographic regions of the world and track the root of their evolutionary history. Our main objective is to demonstrate their high power of discrimination at the individual and population level. In the other hand, their location in the X chromosome that have a special characteristics, giving it many futures and justifying its increasing use in population genetics and population genetics fields; the global evolutionary tendency of the Western Mediterranean population could be pulled. In the global discussion part of this thesis we complemented our results, supported some hypothesis or denied others by adding and exhibiting results obtained from other systems, which had as a concern the Western Mediterranean genetic structure or the Jewish origins: The mitochondrial DNA, autosomal SNP, GM Immunoglobulin allotype polymorphism, Y STR and SNP and autosomal STR.

Thus, our analysis of two polymorphisms X-Alu and X-STR have three aims:

1) Assess the genetic diversity of the Western Mediterranean populations comparing the genetic pool of our studied populations each other and with their neighbors, including sub-Saharan and North European.
2) Describe the genetic pool of the Western Mediterranean populations and plot their kinship degree.
3) Show that the 9 X-Alu polymorphisms and 12 X-STR are reliable in the forensic science field. Validate their use in our populations and generate an allelic and haplotype frequencies database that the international and local forensic science community can use to resolve complex kinship cases.

This study in organized in IV parts. The first one is bibliographic, where we try firstly, to explain the geographical, climatic, anthropological, cultural, linguistic, religious and historic events that characterized these populations. This part exhibits the different hypothesis of their origins. This information is provided for the three main regions that constitute the Western Mediterranean: North Africa, South Europe and the

Jewish populations from Middle East. Secondly, other chapters define the two fields handling this study: Population Genetics and Forensic Genetics. We also describe the STRs and Alu polymorphisms as well as the X chromosome and its interest in both fields. Part II is devoted to describe the material and methods, biologic, informatics and statistical we used to get and analyze our genetic and comparative results. In Part III, we presented our results and their discussion guise scientific six scientific papers where I am the first author, as well as a book chapter that describe the genetic history of Morocco. A general discussion, in the same part gathered all these results and faced them to a deep discussion based on other scientific results and synchronized them with the archeological and historical events.

We consider the archeological, paleanthropological, historical, linguistic and biological data quoted from different books, reviews and investigations in order to understand how the Western Mediterranean as well as Jewish Diaspora bore in this region and how each ethnical group generated and kept its identity and culture and to assess their kinship and relationship.

Could we describe one or many evolutionary history for the Western Mediterranean populations? Is the Mediterranean Sea acted as barrier to the genetic flux? And what was the role of the Sub-Sahara desert in the cultural and genetic exchange between these populations? At a biologic level, what were the historical impact and especially the Arabic conquest on the human Western Mediterranean groups? Is the linguistic distinction between these groups of populations only cultural? Did the Jewish communities conserved their ancestral genetic pool from their earliest beginning to our own day or they lost it? How important is the Arabic contribution in the actual Western Mediterranean genetic structure? All these questions are challenging issues that we tried to answer through the present thesis.

A. The linguistic, historic and cultural context of Mediterranean populations

I. Geographical and cultural context of the Mediterranean region

Complex interplays between historical, cultural and ongoing geographical diversification of the Mediterranean, generated the genetical structure that is currently observed in the human populations of this region.

Geographically, the Mediterranean is the westernmost part of the global scale Alpine-Himalayan orogenic belt which stretches from Spain to New Zealand. The landscapes of the region have a long and complex history that includes both horizontal and vertical crustal movements and the creation and destruction of oceans. This began with the break up of the supercontinent Pangea around 250 Ma, which generated the Tethys Ocean—the forerunner to the present-day Mediterranean Sea. It is the collision of Africa and Eurasia, and the associated tectonics that have been largely responsible for generating the Mediterranean Sea, its subsequent history, and the landscapes that surround it. This collisional history progressively reduced the connectivity of the Mediterranean Sea with surrounding marine bodies by closing and restricting marine gateways. During the Miocene, for example, the Mediterranean basin became completely isolated from surrounding marine bodies in what is known as the 'Messinian Salinity Crisis'. This period saw major changes to the regional water balance leading to evaporation and draw-down of the Mediterranean Sea (Woodward, 2009). This had profound impacts on all aspects of the physical geography of the region including the climatology, biogeography, and geomorphology and its legacy can be seen across the region today.

Furthermore, the Mediterranean region has a highly distinctive climate due to its position between 30 and 45ºN to the west of the Euro-Asian landmass. With respect to

the global atmospheric system, it lies between subtropical high pressure systems to the south, and westerly wind belts to the north. The major climatic episodes of the last millennium in Europe are the Medieval Warm Period (MWP), the Little Ice Age (LIA), and Current Warm Period. Whereas the MWP and the LIA were due to natural variability, in the case of the LIA solar variability and volcanic activity, the Current Warm Period is widely considered to be due to human activities causing global warming. The MWP lasted from the tenth to the fourteenth century. The LIA lasted from around the fourteenth to the end of the nineteenth century. Both the MWP and LIA are visible in the records of past climates from the Mediterranean (Bolle et al, 2003).

In north-west Spain, the MWP was around 1.5ºC warmer than the present day (Martinez-Cortizas et al, 1999). In the eastern Mediterranean, it was a period of wetter conditions with, for example, high water levels in the Dead Sea and Sea of Galilee (Schilman et al, 2001). The LIA was a period of glacier advance in the Apennines and Pyrenees, and wintertime temperatures sometimes as much as 3ºC colder than at present (Giraudi, 2005).

The Mediterranean or sea ‘between the lands’ as he is known in English and romance languages, goes and has gone by many names: ‘Our Sea’ for the Romans, the ‘White Sea’ (Akdeniz) for the Turks, the ‘Great Sea’ (Yam gadol) for the Jews, the ‘Middle Sea’ (Mittelmeer) for the Germans, and more doubtfully the ‘Great Green’ of the ancient Egyptians. Modern writers have added to the vocabulary, coining epithets such as the ‘Inner Sea’, the ‘Encircled Sea’, the ‘Friendly Sea’, the ‘Faithful Sea’ of several religions, the ‘Bitter Sea’ of the Second World War, the ‘Corrupting Sea’ of dozens of micro-ecologies transformed by their relationship with neighbors who supply what they lack, and to which they can offer their own surpluses; the ‘Liquid Continent’ that, like a real continent, embraces many peoples, cultures and economies within a space with precise edges. It is important, then, to begin by defining its limits. The Black Sea washes shores from which grain, slaves, furs and fruit were exported into the Mediterranean since antiquity, but it was a sea penetrated by Mediterranean merchants rather than a sea whose inhabitants participated in the political, economic and religious changes taking place in the Mediterranean itself – its links across land, towards the Balkans, the Steppes and the Caucasus, gave the civilizations along its shores a different outlook and character to those of the Mediterranean. This is not true of the Adriatic,

which has participated strongly in the commercial, political and religious life of the Mediterranean, thanks to the Etruscans and Greeks of Spina, the Venetians and Ragusans in the medieval and early modern period, and the businessmen of Trieste in more modern times. The boundaries of the Mediterranean have been set where first nature and then man set them: at the Straits of Gibraltar; at the Dardanelles, with occasional forays towards Constantinople since it functioned as a bridge between the Black Sea and the White Sea; and at the littoral running from Alexandria to Gaza and Jaffa (Abulafia, 2011).

At the human level, the ethnic, linguistic, religious and political diversity of the Mediterranean, was constantly subject to external influences from across the sea, and therefore in a constant state of flux. From the earliest history of this region, the edges of the Mediterranean Sea have provided meeting-points for peoples of the most varied backgrounds who have exploited its resources and learned, in some cases, to make a living from transferring its products from better-endowed to ill-endowed regions. Even more than fish, which keeps well only after salting or drying, grain has long been the major product carried across the sea, originally grown around its shores or brought down from the Black Sea, but, by the seventeenth century, increasingly of north European origin. Access to supplies of vital foodstuffs and other primary materials enabled cities to grow, whether Corinth, Athens or Rome in antiquity, or Genoa, Venice and Barcelona in the middle Ages. For these cities and many others, denial of access to basic supplies by one’s enemies meant strangulation. Less glamorous than the famous and better documented spice trade, the trade in wheat, wood and wool provided a sure foundation on which it was then possible to build commerce in silk, gold and pepper, items often produced far from the shores of the Mediterranean itself (Horden et al, 2000). The struggle for access to all these commodities set off bloody conflicts between rivals, while the more the Mediterranean was criss-crossed by ships full of rich cargoes, the more these vessels were likely to be preyed upon by pirates, whether ancient Etruscans or early modern Barbary corsairs and Uskoks (Bigelow et al. 2005).

Keeping the sea safe was thus an important function of governments. It could be achieved the Roman way, by actively suppressing pirates in a series of vigorous campaigns, and then policing the sea; or, in times when no one was master of great tracts of the sea, merchant fleets could demand the protection of armed convoys, such as

the Venetian muda. Pirate states in Barbary and elsewhere could be the object of eager negotiation, in the hope of securing guarantees for the safety of those with whom the ruler had treaties, or they could be confronted aggressively, as the Americans successfully chose to do at the start of the nineteenth century. There were bigger dangers to shipping as well, when great land empires reached the shores of the Mediterranean and began to interfere with movement across its surface: the Persians in antiquity, the Ottoman Turks from the late fourteenth century onwards, and (though attempts to acquire permanent bases failed) the Russians in the eighteenth century (Pryor, 1988). Perhaps the most extraordinary case of imperial expansion within the Mediterranean is that of Great Britain, a kingdom with no Mediterranean shores, which thanks to its acquisitions stretching from Gibraltar to Suez, managed to exercise a degree of control that aroused the ire and envy of powers whose lands actually bordered the Mediterranean, notably France.

The history of the Mediterranean is also the story of the port cities of very varied political loyalties in which merchants and settlers from all over the sea and far beyond gathered and interacted. These port cities acted as vectors for the transmission of ideas, including religious beliefs, bringing Greek gods to Etruscan Tarquinia, and much later acting as focal points for the spread of proselytizing Judaism, Christianity and Islam, each of which left an extraordinarily powerful imprint on the societies of the lands around the Mediterranean (Orvietani, 2001).

Those individuals who transformed the Mediterranean world were sometimes visionaries, such as Alexander the Great or St Paul, to cite two very different cases. It is noticeable that they always seem to be men. At a time when gender has become the focus of so much historical debate, one might ask: how male is the Mediterranean? Sedentary merchants might be women, as among the Jews of eleventh-century Egypt and the Christians of twelfth-century Genoa. In that era, at least, wives did not accompany their husbands on trading expeditions, let alone travel for trade in their own right, though attitudes to participation in business varied between Jews, Christians and Muslims. A few European women could be found in the Genoese trading colony in late thirteenth-century Tunis, mainly offering sexual services to the Christian business community. Female participation in naval warfare, a twenty-first-century phenomenon, has not been tested within the Mediterranean. But among migrants, whether the Alans

and Vandals invading Africa at the time of St Augustine, or the Sephardim expelled from Spain in 1492, there was often, though not invariably, a large female component – even the armies of the early crusades were accompanied by both noblewomen and bands of prostitutes. Female pilgrims appear in the record as early as the first decades of the Christian Roman Empire: a fragment from the late fourth-century records the travels of the intrepid Egeria (or Aetheria) from either Gaul or northern Spain to the Holy Land. It is less clear whether the Bronze Age raiders known as the Sea Peoples came accompanied by women to the lands in Syria, Palestine and elsewhere that they settled; indeed, a likely explanation for the rapid abandonment of their Aegean culture by the early Philistines is that they intermarried with the Canaanites, adopted their gods and learned their language. Yet one group of women has a particular importance for the history of the Mediterranean: female slaves, whose fortune varied enormously, from the extraordinary power it might be possible to exercise within an Ottoman harem to the sad exploitation and debasement of those used for sexual purposes or assigned lowly work in the villas of prosperous Romans. During the Middle Ages, many of these slaves, both male and female, were brought out of the Black Sea, but those who inhabited the shores of the Mediterranean in the age of the Barbary corsairs (and at many other periods) also knew the horror of raiding parties that picked people off the shore – Christians off the coasts of Italy, France and Spain, Muslims off the coasts of Morocco, Algeria and Tunisia (Husain, 2007). When King Francis I of France permitted the Turks to visit Marseilles and occupy Toulon in 1543, they kidnapped the nuns of Antibes, among other victims (Abulafia, 2011).

Among all those who traversed the Mediterranean, merchants generally reveal most, for several reasons. One is simply that ever since Phoenician merchants spread the art of alphabetic writing across the Mediterranean, traders have been anxious to record their transactions; we therefore know a great deal about them, whether in Roman Puteoli, near Naples, in medieval Genoa and Venice, or modern Smyrna and Livorno (Gilmour, 2011). But the merchant pioneer is almost by definition an outsider, someone who crosses cultural and physical boundaries, encountering new gods, hearing different languages, and finding himself (much more rarely, herself) exposed to the sharp criticisms of the inhabitants of the places he visits in search of goods unavailable at home. This ambiguous image of the merchant as a desirable outsider is there in our earliest sources. It has been seen that Homer was uneasy about merchants, showing

contempt for mere traders of Phoenicia, and suggesting that they were deceitful and unheroic, despite glorying, paradoxically, in the trickery of Odysseus; the somewhat hypocritical sense that trade dirtied one's hands remained strong among patrician readers of Homer in ancient Rome. It was the Phoenicians, however, who ventured as far as southern Spain, establishing colonies side-by-side with but often apart from the native populations of the western Mediterranean – typically on offshore islands that were easy to guard, for one never knew how long relations with neighboring peoples would remain warm. Then, as the Phoenician colony at Carthage became an economic and political power in its own right, this booming city became the hub of new networks of communication, a cosmopolitan meeting-point between Levantine and North African cultures, a place where divergent cultures fused and a new identity may be said to have emerged, even if the city elite continued to describe themselves as 'people of Tyre' (Horden, 2000). Greek culture too gained a purchase in Carthage, whose citizens identified the Phoenician god Melqart with Herakles Gods and goddesses as well as merchants criss-crossed the ancient Mediterranean. Additionally, the presence on the shores of Italy of Phoenicians and Greeks, individuals with a distinct cultural identity, acted as yeast that transformed the villages of rural Etruria into cities whose richer inhabitants possessed an insatiable hunger for the foreign: for Greek vases, Phoenician silver bowls, Sardinian bronze figurines. Alongside merchants who came for the metals of Italy, we can soon detect artisans who travelled west to settle in the lands of the barbaroi, knowing that their skills would probably earn them greater esteem than at home, where each was one of many (Blake et al. 2005).

There are striking parallels in later centuries. Alien traders are an obvious feature of the medieval Mediterranean, where we have the intriguing phenomenon of the ghettoized merchant visiting Islamic or Byzantine territory, enclosed in an inn or fonduk that also functioned as a warehouse, chapel, bake-house and bath-house, with one inn for each major 'nation': Genoese, Venetian, Catalan and so on. The sense that the merchant might be a source of religious contamination and political subversion led the rulers of Egypt to lock the doors of these inns at night-time (the keys being held by Muslims on the outside). This only enhanced the solidarity and sense of community that held these merchants together, while underlining the differences between the various groups of Italians and Catalans, who coexisted in rivalry Muslim emirs proved adept at exploiting. The Byzantines too set the Italian merchants apart in a walled compound

during the twelfth century, feeding xenophobia in their capital city, with the ugly consequences of anti-Latin pogroms. The idea of enclosing distinct communities behind walls was not, then, particularly novel when the king of Aragon first segregated the Majorcan Jews around 1300, and was quite venerable by the time the government of Venice enclosed the Jews in the ghetto nuovo in 1516; these merchant communities provided a useful model for the ghetto. The enclosed areas, whether of Jews or of European merchants, were places where a certain amount of privilege – self-government, freedom to practice one's religion, tax exemptions – was counter-balanced by constraint – limitations on free movement and reliance on often capricious public authorities for protection (Abulafia, 2011).

To speak of the Jews is to speak of traders who had an unusual ability to cross the boundaries between cultures, whether in the early days of Islam, during the period of ascendancy of the Genizah Jews from Cairo, with their trans- and ultra-Mediterranean connections, or in the period of Catalan commercial expansion, when they could exploit their family and business ties to their co-religionists and penetrate deep into the Sahara in search of gold, ostrich feathers and other African products that were beyond the reach of their Christian compatriots still stuck within their trading compounds. These Jewish merchants were able to bring back information about the world beyond the Mediterranean ports that was recorded and disseminated across Mediterranean Europe and further afield in the remarkable portolan charts and world maps produced in late medieval Majorca. As merchants moved around, so did information about the physical world (Montville, 2011).

The concept of the Mediterranean as a 'faithful sea', to cite the title of a recent collection of essays, needs to take into account its role as a surface across which moved not merely poor and anonymous pilgrims but also charismatic missionaries such as Ramon Llull, who died in 1316 after writing hundreds of books and pamphlets on how to convert Muslims, Jews and Greeks to the true faith, without, it must be said, ever converting anyone (Pawer, 1988). Yet Llull's career is a reminder that religious friction and confrontation are only part of the picture. He imitated Sufi verses and hobnobbed with kabbalists; he was at once a keen missionary and an exponent of old-fashioned Iberian convivencia, recognizing the God of the three Abrahamic religions as the same single God. A different sort of convivencia existed in the minds of members of the

religious communities that were expelled or forced to convert as Spain asserted its Catholic identity in 1492 and afterwards: the Marranos and Moriscos, Jews and Muslims who might or might not adhere to their ancestral religion in private, while being expected to practice the Catholic faith in public. The ascendancy of the Sephardic merchants in the early modern Mediterranean is astonishing in any number of ways: their ability to acquire and shed different identities, as 'Portuguese' able to enter Iberia and as Jews resident in Livorno or Ancona – an ability to cross cultural, religious and political boundaries reminiscent of their forebears in the Cairo Genizah six centuries earlier. These multiple identities are an extreme case of a wider Mediterranean phenomenon: there were places where cultures met and mixed, but here were individuals within whom identities met and mixed, often uneasily (Abulafia, 2011).

There is an understandable tendency to romanticize the Mediterranean meeting-places, and the darker reality of trans-Mediterranean contact in the early modern period also needs to be born in mind: the ascendancy, between the fifteenth and the early nineteenth centuries, of the Barbary corsairs, and the close intersection between piracy and trade. Before the final suppression of the Barbary corsairs, the Mediterranean had only ever really been free of a serious threat from piracy under Roman imperial rule, as a result of Rome's political control of more or less all its shores and islands. But piracy reveals some of the most extraordinary cases of mixed identity: corsairs from as far away as Scotland and England who, outwardly at least, accepted Islam and preyed on the shipping of the nation from which they came. This darker side of Mediterranean history also encompasses the history of those already mentioned whom the pirates carried back and forth: male and female slaves and captives, though they too, like the historian Polybios, could play a notable role in cultural contact between the opposing shores of the Mediterranean (Ruiz-Domenec, 2004).

The unity of Mediterranean history thus lies, paradoxically, in its swirling changeability, in the Diasporas of merchants and exiles, in the people hurrying to cross its surface as quickly as possible, not seeking to linger at sea, especially in winter, when travel became dangerous, like the long-suffering pilgrims Ibn Jubayr and Felix Fabri. Its opposing shores are close enough to permit easy contact, but far enough apart to allow societies to develop distinctively under the influence of their hinterland as well as of one another. Those who cross its surface are often hardly typical of the societies from

which they come. If they are not outsiders when they set out, they are likely to become so when they enter different societies across the water, whether as traders, slaves or pilgrims. But their presence can have a transforming effect on these different societies, introducing something of the culture of one continent into the outer edges, at least, of another (Abulafia, 2011). The Mediterranean thus became probably the most vigorous place of interaction between different societies on the face of this planet, and it has played a role in the history of human civilization that has far surpassed any other expanse of sea.

II. Historical and demo-linguistic data of Morocco, Balearic Islands and South of Italy

Morocco

Situated in the northwest corner of Africa and, on a clear day, visible from the Spanish coast, Morocco has resisted outside invasion while serving as a meeting point for European, Eastern, and African civilizations throughout history. Its early inhabitants were Tamazight-speaking nomads; many of these became followers of Christianity and Judaism, which were introduced during a brief period of Roman rule. In the late 7th century, Arab invaders from the East brought Islam, which local Imazighen gradually assimilated. Sunni Islam triumphed over various sectarian tendencies in the 12th and 13thcenturies under the doctrinally rigorous Almohad dynasty. The Christian reconquest of Spain in the later middle Ages brought waves of Muslim and Jewish exiles from Spain to Morocco, injecting a Hispanic flavor into Moroccan urban life. Apart from some isolated coastal enclaves, however, Europeans failed to establish a permanent foothold in the area. In the 16th century, Ottoman invaders from Algeria attempted to add Morocco to their empire, thus threatening the country's independence (Douglas, 2005). They, too, were thwarted, leaving Morocco virtually the only Arab country never to experience Ottoman rule. In 1578, three kings fought and died near Ksar el-Kebir (Alcazarquivir), including the Portuguese monarch Sebastian. This decisive battle, known as the Battle of the Three Kings, was claimed as a Moroccan victory and put an end to European incursions onto Moroccan soil for three centuries. The 17th century saw the rise of the "Alawite dynasty" of sharifs, who still rule Morocco today. This dynasty fostered trade and cultural relations with sub-Saharan Africa, Europe, and the Arab lands, though religious tensions between Islam and Christendom often threatened the

peace. By the late 17th century, Morocco's cultural and political identity as an Islamic monarchy was firmly established. The figure of the strong sultan was personified by Mawlay Ismail (1672–1727), who used a slave army, known as the 'Abīd al-Bukhari, to subdue all parts of the country and establish centralized rule. Subsequent monarchs often used their prestige as religious leaders to contain internal conflicts caused by competition among tribes. In the late 18th and early 19th centuries, when Europe was preoccupied with revolution and continental war, Morocco withdrew into a period of isolation. On the eve of the modern era, despite their geographic proximity, Moroccans and Europeans knew little about each other (Penell, 2003 and Mckenna, 2011).

Balearic Islands

The Balearic Islands (Majorca, Minorca, and Ibiza) were settled by different people throughout their history (Casasnovas, 1998), which has contributed to the genetic pool of the actual population, in particular, it is important to emphasize the contribution of the Romans in the 3rd century BC and the Catalans in the early 13th century, although Ibiza differs from the other two islands, especially in the origins of founding settlements. Whereas Majorca and Minorca were inhabited since 5,500 years ago there is no archaeological evidence for permanent human habitation on Ibiza before the arrival of the Carthaginians in 654 BC. If the island was occupied during the prehistoric period, its residents failed to leave evidence of the "talaiotic" (stone builder) culture that is prominent in both Majorca and Minorca. The Carthaginians remained in Ibiza for at least five centuries and colonized the coastal areas and interior of the island. Ibiza was annexed by the Roman Empire as part of a political pact, but apparently the Romans failed to occupy it. During the last seven centuries, the Ibiza population has been reproductively isolated, and has thus received little gene flow from outside. The small size of the autochthonous population and its reproductive isolation resulted in a moderate incidence of consanguineous marriages (about 6%; Valls, 1969).

It is important to mention the Arab/Muslims who, after their taking full control of the Islands, left their stamp in Balearic population. The Arabs first came to the Balearic in the early the 8th century when in 707, a Muslim fleet, under the command of Abd Allah Ibn Musa, son of the governor of Ifriqiya, Musa Ibn Nusayr, reached the islands. However, they did not fully occupy the Islands until 902 when the Islands were annexed to the Emirate of Cordoba. In that year the affluent Moorish man of commerce, Al

Hawlani convinced the ruler of Al-Andalusia to conquer the Balearics. Al Hawlani himself took the command over the fleet, conquered the Balearics and was made the first Governor of Mallorca. During their rule, the Arabs gave the islands a long period of cultural blooming and prosperity, which saw Medina Mayurka, the Arabic name for Palma, the Balearic Island's capital, become one of the major trading ports in the Mediterranean and a city of culture and wealth. In the 12th century, Ramon Berenguer III troops plundered Mallorca and Ibiza. In 1203, the Almohades dominated the Balearic territory and the Islands stayed under Islamic domination until 1229 when the King James I (Aragon Crown) began to occupy them.

Another deme present in the Balearic Archipelago is represented by the Chuetas, a small and inbred community of descendants of Sephardic Jews. Although the presence of Jews on Majorca Island goes back to ancient Rome, the Jewish communities that have a historical continuity with the so-called Chuetas only go back to the Moslem period (10th–13thcenturies). The Christian occupation of Majorca in 1229 guaranteed the survival of the Jewish population mainly for their technical and commercial superiority, but it compelled them to isolation. In spite of their official conversion to Christianity (1391–1435), they kept their traditions and beliefs. For this reason, the Inquisition persecuted them until the17th century. Their descendants, the Chuetas, were excluded by their neighbors, and consequently, intermarriage with non-Jews did not take place until the middle of the 20th century. This isolation almost exclusively affected the bearers of the 15 surnames of the accused people in the last Inquisition's "autos-de-fe" (Braunstein, 1936).

The Balearic populations have not received any remarkable contribution of foreign genes for the last seven centuries, until very recently when the tourist boom promoted active immigration, principally from mainland Spain, but also from other European countries. In 1980, the Balearic Islands got the access to the autonomy under the article 143 of the constitution.

The language of the Balearic Islands is Catalan, although the population can be considered bilingual, being Spanish and Catalan both official languages. The areas inhabited historically by Catalan peoples almost all contain a significant percentage of Catalan speakers today. These are, in order of size and population: Catalonia, which

accounts for over half of the population of the Catalan-speaking areas; Valencia, with over a third; the Balearic Islands, with just under 8 percent; and then Catalunya Nord in France, the Principality of Andorra, the border areas known as the Franja in Aragon, Alguero in Sardinia, and some population centers in Carxe in Murcia, which together make under 4 percent of the total population of nearly 14 million people (Micula, 2004).

South of Italy

The region of Calabria is located in southern Italy, south of Naples. It covers 15,080 km^2 and has a population of just over 2 million. Calabria has an interesting history. From the VIIIth to the Vth centuries BC, Calabria was colonized by Greeks who established, in a special way along the Ionian coast, some important and populous urban settlements, such as Sybaris (Sibari), Kroton (Crotone), and Locrys (Locri). Subsequently, the territory was occupied by other groups: Romans, Longobards, Byzantines, Normans, Arabs, Hispanics, and the French (Lenormant, 1976).All of these groups probably contributed genetic influences to the indigenous population. Moreover, the economic status of Calabria at the beginning of this century and up until today was based on gathering and sheep-herding, but it was never like the economic conditions in northern Italian regions. These conditions contributed to the emigrations at the beginning of this century to transoceanic regions, and then during the1960s to northern Europe and northern Italy. Furthermore, the precarious economic status and the geographical peculiarities of the territory did not allow communication between the three Calabrian provinces until the beginning of this century. Until this period, there was only one road, the state road"S.S.19" which follows the ancient "Via Popilia," built for military reasons following the occupation by the Romans. Improvement of the roads during the period following World War II allowed better communication between Cosenza, Catanzaro, and Reggio Calabria. But, when the towns along the Tyrrhenian and Ionian coasts improved their communications with other towns, those in the internal region were again isolated. The populations living within the three Calabrian provinces do not have a different style of life, but geographically the three provinces are different. Within Cosenza, a "passive" barrier, represented by a coastal mountain chain, separates the Tyrrhenian coast from the internal area. Within Catanzaro, both coasts are separated by a relatively short (45 km) area that allows good communication between both the coasts and the internal area. Within Reggio Calabria, the Aspromonte mountain chain

does not represent a barrier to thee ventual fusion between the population living on the coasts and the internal area (Ancestry.com, 2007).

The capital city of Calabria is Catanzaro and the official national language of Calabria has been Standard Italian since before unification in 1861, as a consequence of its deep and colorful history, Calabrian dialects have been spoken in the region for centuries. Most linguists divide the various dialects into two different language groups. In the northern one-third of the region, the Calabrian dialects are considered part of the Neapolitan language (or Southern Italian) and are grouped as Northern Calabrian or *Cosentino*. In the southern two-thirds of the region, the Calabrian dialects are considered part of the Sicilian language and are often grouped as Central and Southern Calabrian (Tagarelli et al. 2000).

Other historical languages have left an imprint on the region. In isolated pockets, as well as some quarters of Reggio Calabria (historical stronghold of the Greek language in Italy), a hybrid language that dates back to the 9th century, called Griko, is spoken. A variety of Occitan can also be found in certain communities and French has had an influence on many Calabrian words and phrases. In several villages, the Arbëresh dialect of the Albanian language has been spoken since a wave of refugees settled there in the 15th century. In addition, since Calabria (as well as other parts of southern Italy and Sicily) was once ruled by the Spanish, some Calabrian dialects exhibit Spanish derivatives. In the south-west of Calabria is located the largest island in the Mediterranean Sea southern Italian region “Sicily”. It is the closest to the mainland and shares considerable history with the Italian peninsula and the populations of Southern Italy.

During the historic period, Sicily was populated by different peoples of both Indo-European (Sicels) and non-Indo-European (Sicans, Elymes) origin. Subsequently, significant influences were imparted by the Phoenicians in the West and Greeks in the East. After the Roman period, Sicily was invaded by German tribes such as the Visigoths, and later ruled by the Arabs and Normans. During these centuries, an appreciable immigration from both Central-North Europe and North Africa is historically documented (Leighton, 1999).

B. Religious diversity between populations of interest

I. Insight of the religious diversity between South of Europe, North Africa and Middle East

Religion has left its striking place in the worldwide regions, leading to an important influence in different issues. The examination of the current relationship between selected religious actors and the state in Europe, Middle East and North Africa (MENA) is important to well understand the genetic mapping of the related populations andhow the religion influences the genetic life and future of people.

Islam is the second-largest religion in the world (after Christianity) and will soon be the second-largest religion in America (The List, 2010). Since we are more familiar with Christianity, we know without thinking that there is great diversity in this religion. Christianity expresses itself in many forms and contexts. There are different Christian churches or sects (from Baptists to Unitarians, Roman Catholics to Greek and Russian Orthodox), existing in different cultures (North American, Middle Eastern, European, Asian, and African). The result is a diversity of beliefs and practices within what we call Christianity. So too in Islam, although Muslims maintain that there is one divinely revealed and mandated Islam, there are many Muslim interpretations of Islam. There are two major branches, Sunni (85 percent of the world's Muslims) and Shiite (15 percent). Within them are diverse schools of theology and law. In addition, Islam has a rich mystical tradition that includes many Sufi orders or brotherhoods (Ramadan, 2011).

Islam represents a basic unity of belief within a rich cultural diversity. Islamic practice expresses itself in different ways within a vast array of cultures that extend from North Africa to Southeast Asia as well as Europe and America. In recent decades, Islam has gone from being invisible in America and Europe to being a prominent feature in the religious landscape. Muslims represent a broad spectrum of racial and ethnic groups. The racial and ethnic diversity of Islam is represented in Europe by two broad Muslim groups: Europeans converted and immigrant.

Judaism, Christianity, and Islam, in contrast to Hinduism and Buddhism, are all monotheistic faiths that worship the God of Adam, Abraham, and Moses—creator, sustainer, and lord of the universe. They share a common belief in the oneness of God (monotheism), sacred history (history as the theater of God's activity and the encounter of God and humankind), prophets and divine revelation, angels, and Satan. All stress moral responsibility and accountability, Judgment Day, and eternal reward and punishment. All three faiths emphasize their special covenant with God, for Judaism through Moses, Christianity through Jesus, and Islam through Muhammad. Christianity accepts God's covenant with and revelation to the Jews but traditionally has seen itself as superseding Judaism with the coming of Jesus. Thus Christianity speaks of its new covenant and New Testament. So, too, Islam and Muslims recognize Judaism and Christianity: their biblical prophets (among them Adam, Abraham, Moses, and Jesus) and their revelations (the Torah and the Gospels). Muslim respect for all the biblical prophets is reflected in the custom of saying "Peace and blessings be upon him" after naming any of the prophets and in the common usage of the names Ibrahim (Abraham), Musa (Moses), Daoud (David), Sulayman (Solomon), and Issa (Jesus) for Muslims. In addition, Islam makes frequent reference to Jesus and to the Virgin Mary, who is mentioned more times in the Quran than in the New Testament (Ramadan, 2008).

History teaches us that religion is a powerful force that has been used for good and for ill. From Egypt, Sudan, and Nigeria to Pakistan, Indonesia, and the southern Philippines, Muslims have clashed with Christians. Moreover, despite an impressive record of religious pluralism in the past, the situation in the Middle East has gotten worse rather than better. It is often difficult to identify specific conflicts as primarily motivated by religion as opposed to politics or economics (Bowker, 2002).

It is useful to recall that historically Islam's attitude toward other religions, especially Judaism and Christianity, was more tolerant than that of Christianity. However, Muslim-Christian relations have deteriorated over time under the influence of conflicts and grievances, from the Crusades and European colonialism to contemporary politics. Part of the legacy of colonialism is a deep-seated Muslim belief, nurtured by militant religious leaders, that indigenous Christians were favored by and benefited from colonial rule or that they are the product of the European missionaries and their schools

that converted local Muslims, and somehow retain a connection to a Christian West. The situation is compounded in areas where Christians proved more affluent or successful. The creation of the state of Israel and subsequent Arab-Israeli wars and conflicts has contributed to a deterioration of relations between Palestinian Muslims and Christians and IsraeliJews (Davies, 2003).At the heart of the conflict in Israel-Palestine is the creation of the state of Israel and the Palestinian demand for a Palestinian state and the right to return to their lands. At the same time, for a significant minority of Muslims and Jews, the struggle is at its heart based upon conflicting religious claims to the land.

The birth of Islam in North Africa began in the 7th and 8thcentury with the Umayyads, who brought the religion from the Middle East to North Africa. Along the coast of Africa, Islam spread among Berbers, who joined the Muslim community and drove north among the Mediterranean into Europe. The Muslim conquest of Iberia was a great military success. In seven short years (711 CE-718 CE), Islamic inhabitants from North Africa became the authoritative rule in the Iberian Peninsula. Working their way in through Gibraltar, the Muslims were successful all the way to the Pyrenees Mountains until they were defeated by the Franks in 732 CE at the Battle of Tours. While the amount of Muslim power varied from the 8th to 15th centuries in the Peninsula, the effect that Islamic culture had on Spanish lifestyle can still be seen in the architecture and art in modern day. In Morocco, Muslims founded the city of Fez (808), wish soon thereafter gave refuge to Andalusian Muslims fleeing an uprising in Cordoba (Menocal, 2009).

A large mixture resulted due to the successive movements of colonization and expansions even by Christians, Muslims or Jewish especially between Mediterranean countries. Therefore, genetic studies arestrongly required to assess the mixture degree of the three religious affiliations.

II. The historical and geographical movement of the Jewish communities

The national and religious origins of Jews have been traced, archeologically, to the Middle East in the second millennium BC, what happened next has been more opaque. Ever since, they kept genetic, cultural and religious traditions coherence despite, migrations from the Middle East into Europe, North Africa, and beyond over the centuries. Nowadays, within the Jewish community, various groups are differentiated depending to their history and settlement. The three main groups are Orientals, Sephardic and Ashkenazi Jews (LivresGroupe, 2010) (Figure 2).

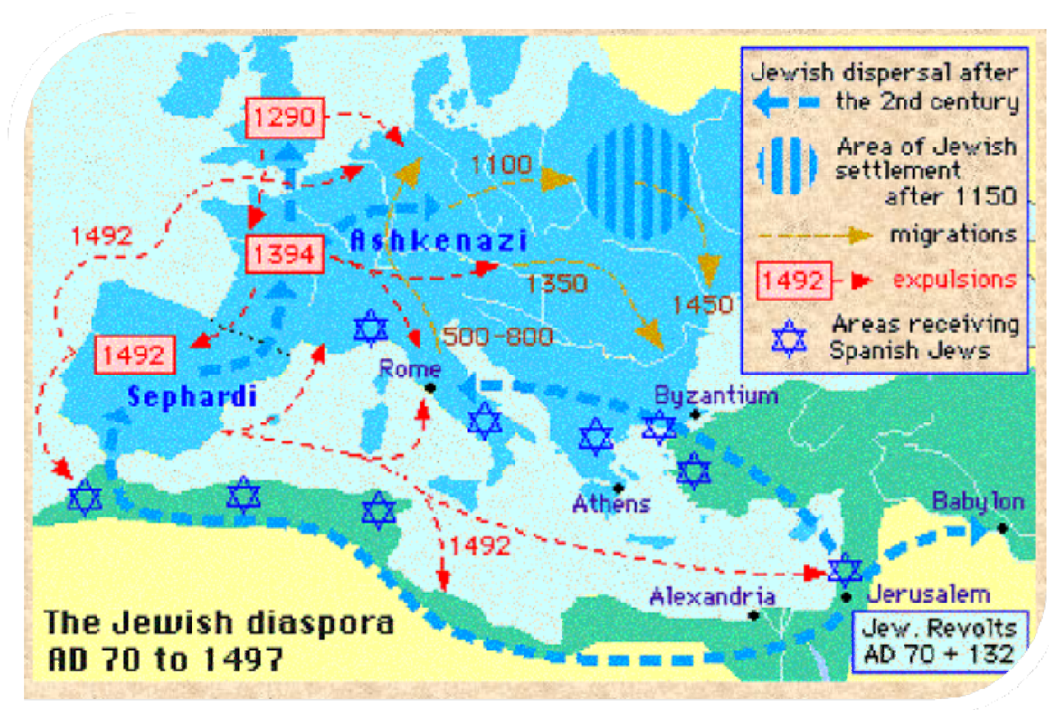


Figure 2: The Jewish Diaspora between 70 to 1497 AD.
(http://www.worldreligions.psu.edu/world_religions18.htm)

The Oriental Jews settlement is Israel, Palestine, Iran, Iraq, Asia Central and Arabic peninsula, come from communities originated from Persian and Babylonian empires (IIIV centuries BC).

The Sephardic Jews (from Hebrew Sepharad: Spain) lived in Spain and Portugal. The current Jewish communities of Balkans, Italy, North Africa and Syria were formed during the classical antiquity and they have been mixed with Sephardic Jews, wish migrated after their expulsion from the Iberian Peninsula in the XV century. According to old traditions, without documentary evidence, the first Jewish reached Spain together with the Phoenician shipment members; even if it's possible in some cases their arrival as refugees, due the destruction of the holy temple (temple in Jerusalem), in the year 30 of our age. In those earlier centuries, Spain was identified as the "Sepharad Biblica". Thus, the Spanish Jews were called "Sephardies/Sepharditas" (Baer, 2001). The Jews settlement was mainly in the Mediterranean cost, in cities like Gerona, Ampurias, Tarragona, Malaga and Cadiz and left tracks of their presence since the II century, in the same time of the Roman dominance, coexisting as workers or even slaves. But progressively they lost their prestige, as in the IV century; the council of Elvira (Granada) drew up some laws against Sephardic, because of the religious danger that they could present, wish led to the forced conversion in some places like in Mahón (Menorca). It is widely believed that the first persecution of Jews in Spain occurred under the Christians. Rather the first persecution of Jews in Spain occurred under the Visigoths in the 7th Century, not under Christian rule. Jews gained rapidly in importance by the beginning of the Muslim's invasion of the Iberian Peninsula (the year 711) (Jane et al, 1992). Under the Almoravides, the Jewish community prospered, but they were persecuted when the Almohade began fighting for the power, especially in Cordoba. In 1478, a war outbroke to put an end to the Arabic colonization of Spain. After that, in 1492, an expulsion decree was signed by the Catholic Kings. In the other hand, Jewish were expulsed from most of European countries between the XI and XV centuries. The North African Jews, even though they are considered as Sephardic, in some studies they were analyzed as a group apart. Nevertheless, they are a blend of pre-existing Jews in North Africa and the populations that settled their before their expulsion from European countries (Balta et al, 2003).

The Ashkenazis Jews (from Hebrew Ashkenaz: Germany) settled in the valley of Rin, originally from the Oriental Jews of Palestine, during the VI-IX centuries and the immigrated to the East of Europe between the XI-XV centuries, especially to Polony, Lituany, Byelorussia, Ukraine and Russia, being the most large Jewish community. Some theories rises about their Sorbian (with a Slavic language in Germany) and Khazarian origins. Ashkenazi Jews developed their own language, the "*Yiddish*" (blend of and old German and Hebrew, with some words stemmed from Slavic language, wrote with a Hebrew alphabet) (Straten, 2011).

Geography is linked to genetic variation, and people who have the same geographic ancestry are more likely, on average, to be genetically similar than people who do not. Thus, it is important to analyze the Jewish community looking for a possible evidence of genetic indicators for geographic and ethnical ancestry and to verify if this group stayed close genetically or if he became mixed with other surrounding populations. Some historians situated the Jewish mixture by the beginning of the Jewish world, applying for the possible limitation of mixture in the last 2000 years because of their religious principles (Bonne-Tamir, 1992; Frederic et al. 2011 and Goldstein, 2008).

Several studies focused on the Jewish community since the XX century, trying to best characterize the interrelation between this community and the populations they coexisted with and between themselves during the Diaspora. The first genetic studies on the Jewish populations were carried out with classical markers like blood type polymorphism and in different blood markers as serum protein and erythrocytes' enzymes. Some published results, assumed the Middle Eastern origins of the Jewish people, with some genetic similarities between most of them but the mixture degree with the local populations remain important. The development and use of some techniques based on DNA analyses, helped widely in the population genetics field, to characterize the Jewish population. The analyses of mitochondrial DNA, as well as autosomes, Y chromosomal and recently X chromosomal markers especially STR and SNP were highly used (i.e. Hammer et al. 2000; Ostrer 2001; Thomas et al. 2002; Atzmon et al. 2010; Behar et al. 2010).

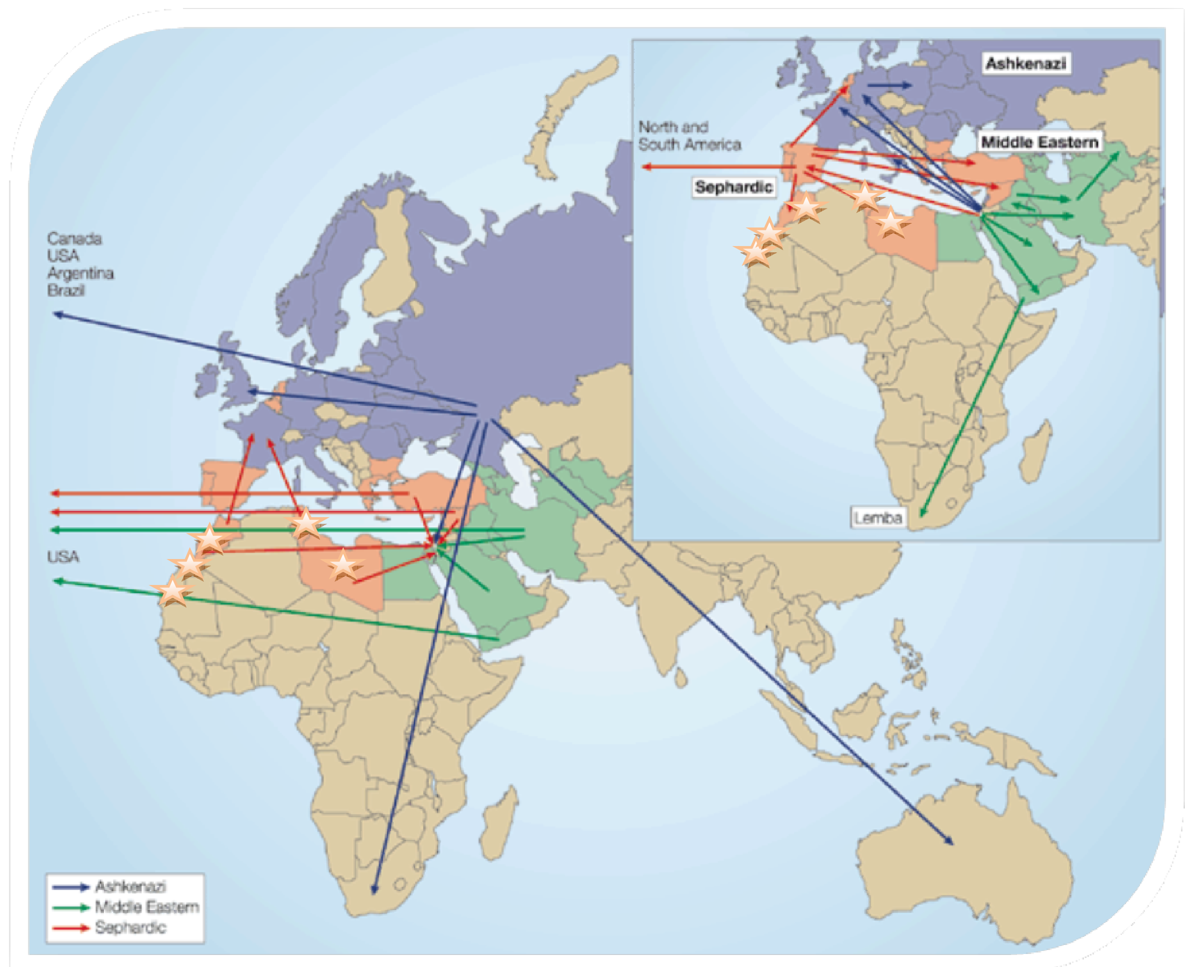


Figure 3: The Jewish migration pattern and worldwide distribution.
(www.jewishgene.org)

C. *Concepts*

I. Anthropology: Concept and History

Anthropology, from Greek ***anthrôpos***, man, and ***logos***, science, is the scientific study of the origin and behavior of man, including the development of societies and culture. It began in the XIXth century and was progressively institutionalized in Europe and USA in the XXth century to be part of human and social sciences. Anthropology studies the human being and their ancestors through time and space and in relation to physical character, environmental and social relations and culture. It includes different fields of research as the social and cultural anthropology, the linguistic anthropology, the historic anthropology, the physical anthropology, etc…, and of course the Anthropology we deal with in this thesis: the biological anthropology or anthropobiology which studies the specificity and diversity of human populations, from the ancestry origins to this day, analyzing the interactions between biology and human populations culture(Crebezy et al. 2002) past and future along with the reconstitution of their scalable history(Susanne et al. 2003).

Two main sciences are identified with anthropological biology: human sciences and life sciences. The first one merges all the disciplines useful to redraw the cultural and behavioral particularities of human and ethnical groups as well as their relation throughout time: sociology, ethnology, linguistic, history, etc. The second one is based on all the sciences allowing the physical and biological characterization of human: anatomy, physiology, molecular biology, genetic, medicine, etc.

The way in which the complex constitution of man is described and analyzed depends in part on genetic, behavioral, sanitary and environmental purposes, explaining the interconnection of the different disciplines described below. Furthermore, the human studies can't be separated from what happen both inside (genetic pool) and outside

(environment and life conditions) the body. Some discoveries studying the interaction between man and environment in earth's sediment, leads to a third concept: science of earth including: geology, paleontology, paleoanthropology, archeology, geography, etc.

Biologic Anthropology cannot be distinguished by its own techniques and methods. It is for some such reason that, as most writers on the subject freely admit, biological anthropology-unlike other disciplines, might appear as a "soft" and less exact science than it purports to be. However, its originality lies without doubt in its capacity to gather several approaches used in different fields of research in order to answer the different issues it deals with: understand and redraw the biologic diversity of ancestral and present man. Indeed, the multidisciplinary of biological anthropology allows widely studying the human origins, its physical and cultural characteristics as well as the current and past relations between contemporary subjects. The main purpose of this thesis leans basically upon results using multidimensional scaling as an anthropological and population genetics or anthropogenetic analysis.

The anthropogenetic operational practice was promptly evolved from the methodological principals of molecular biology. The arrival of PCR method and DNA-based markers gave the real boost to this field of research. Since 1989, the number of investigations increased including some different animal and vegetal groups: species recently eradicated by man (Aepyornis), glacial periods representatives (mammoth), or even domestic species (pork) as preferred target. The scopes of the investigation increased, in order to well understand the species, populations and genomes evolution: population genetics, species phylogeny, domestication, migration of populations, paleopathology, paleogenomic and molecular evolution, which joined a field in booming.

II. Population genetics: concept and history

1. Definition

Any species of organism in nature lives in a form of **population**. A population of organisms is characterized by some sort of cooperative or inhibitory interaction between

members of the population. Thus, the rate of growth of a population depends on the population size or density in addition to the physical environment in which the population is placed. When population density is below a certain level, the members of the population often interact cooperatively, while in a high density they interact inhibitorily. In organisms with separate sexes, mating between males and females is essential for the survival of a population. Interactions between individuals are not confined within a single species but also occur between different species. The survival of a species generally depends on the existence of many other species which serve as food, mediator of mating, shelter from physical and biological hazards, etc.

A population of organisms has properties or characteristics that transcend the characteristics of an individual. The growth of a population is certainly different from that of an individual. The differences between ethnic groups of man can be described only by distributions of certain identifiable genes. All these measurements are characteristics of populations rather than of individuals.

Population genetics is aimed to study the genetic structure of populations and the laws by which the genetic structure changes. By genetic structure we mean the types and frequencies of genes or genotypes present in the population. Natural populations are often composed of many subpopulations or of individuals which are distributed more or less uniformly in an area. In this case the genetic structure of populations must be described by taking into account the geographical distribution of gene or genotype frequencies. The genetic structure of a population is determined by a large number of loci. At the present time, however, only a proportion of the genes present in higher organisms have been identified. Therefore, our knowledge of the genetic structure of a population is far from complete. Nevertheless, it is important and meaningful to know the frequencies of genes or genotypes with respect to a certain biologically important locus or a group of loci.

The population genetics is considered as a dialog between predictions based on principles of Mendelian inheritance and results obtained from empirical measurement of genotype and allele frequencies that form the basis of hypothesis tests. Idealized predictions stemming from general principles form the basis of hypotheses that can be tested. At the same time, empirical patterns observed within and among populations

require explanation through the comparison of various processes that might have caused a pattern. Population genetics is based on two concepts: Genetic diversity and population. Chiefly, it tries to find the genetic variability or observable polymorphism in different genetic loci (alleles), in a given population and then, to assess this genetic polymorphism, calculating the frequencies (proportions) of all the alleles present in this population. The results obtained, when compared in different scales of time in the same population, give information about the genetic evolution of this population, through a variation (increase, decrease or constancy) of allelic frequencies. The genetic comparison of different populations allow to assess the observant differences (or similarities) and then to get the key for the genetic history of each population, a group of populations or even a species. This genetic history aimed to characterize the genetic origins of populations but also to redraw their interaction over time (genetic admixture, gene flow).Furthermore to the population behavior (cultural choice, religious, demographic, sanitary, etc), the allelic frequencies variations are associated to other process “evolutionary forces”: mutation, migration, natural selection and genetic drift. The main objectives of the population genetics are:

1º) To assess the genetic variation of natural populations and to describe the organization basic grounding for this variation.
2º) To explain the origins, the evolution and stability of the genetic variation, by the evolutionary forces effect.

2. Hardy-Weinberg equilibrium and evolutionary forces

The population genetics models redraw the evolution of parameters describing the entire population. Those parameters are set up by calculating some quantities measured out from representative samples of population. The central theoretical model of population genetics was described in 1908 by the British mathematician G. Hardy and the German doctor W. Weinberg. Known as, Hardy Weinberg equilibrium, this model match under a given conditions, the genotypic frequencies equilibrium expected in an offspring, contingent on the parental allelic frequencies (Hardy, 1908 and Weinberg, 1908).

In spite of the Hardy-Weinberg hypothesis simplicity and evidence, this model remains useful for many reasons. Theoretically, it allowed stating the different factors that can modify the population genetics constitutions. If we make a hypothesis, we can assess the role of each factor, than assuming simultaneously several hypotheses, we can study the cumulative effects of different factors.

Four evolutionary forces can influence de population equilibrium state: the mutation, the migration and the genetic drift.

The **mutation** is the principal source of genetic variability. It decides of any hereditary variation in the DNA sequence despite its biologic, physical or chemical origin. It can be a pinpoint modification of one or many nucleotides without modification in the length of the sequence, or a modification by insertion, deletion or transposition of more important part of the genome. Concluding, mutations are responsible for genetic diversity by generating new alleles.

The **selection** is expressed by a different selective or adaptative value according to genotypes. This concept assuming that some individuals can survive and mate in a given environment more than others, was one of the three principles of the evolution theory of species by natural selection, presented by Darwin in 1859, in his famous book "***On the Origin of Species***"(Darwin, 1859). The two other principles are: the variation concept of individuals from the same population (morphologic, physiologic…) and the heredity concept assuming that individuals are closer to their progenitors than to other individuals without family ties. Then, natural populations are formed by individuals genetically different by many genes. This genetic polymorphism allows the individuals and populations survival. The darwinien selection leads to the intrapopulation genetic diversity disappearance. However, as an allele can be, according to the conditions, unfavourable in a population and positive in other, selection does not necessarily reduce the intraspecific genetic diversity: its limits are the interpopulation diversity.

The **migration** is a potential reason for the genetic exchange (allele transmission) between populations. It modifies, evidently, the allelic frequencies in a given population and can even leads to a homogenization of the allelic frequencies between different

groups. It is as strong as the number of immigrants is important and the allelic frequencies difference between the two populations is meaningful.

The **genetic drift** occurs by sampling the gametes in each generation, just someone participate to the breeding. The generation that stand in for its genitors, presents due to the random effects of meiosis and fecundation, different allelic frequencies. This frequencies' variation is as big as the population size is small. So, in the small populations, those random variations can cause loss of alleles. In opposite, in some bigger populations, the allelic frequencies do not change a lot from a generation to another, due to the big number of the potential genitors. We can even consider them constant over a short time. The **founder effect** is a particular case of the genetic drift. The founder effect is generated when a small group of persons is the founder of a new population. This migrant group, holder of a small fraction of total genetic variation of the original population, can induce a concentration of some genetic futures, positive or deleterious, inside a population. The allelic frequencies of this new population are different from those of the original population and can even match the overrepresentation of alleles initially seldom or infrequent and responsible of some genetic diseases. For example, due to a fonder effect, there is a very high prevalence of the familial hypercholesterolemia in Tunisians (Slimane et al. 1993 and Slimane et al. 2002), or also a higher prevalence of many hereditary diseases (myotonic dystrophy, dystrophy oculopharyngeal,..) in a north eastern population of Quebec(Heyer et al. 1995 and Yotova et al. 2005).

Despite the random mating of the human populations, some observed violation of this concept can be due to some preferential mating between related individuals, as it is the case for the homogamy, population subdivisions, **endogamy** or **inbreeding**. The endogamy in populations can be the consequence of the geographic confinement of some breeding mechanism or behavioral characteristics (cultural, social). The inbreeding generates a fall of heterozygosity and then a loss of the population genetic diversity. Furthermore, in the medical field, the increasing of the homozygosity probability can unfortunately increase some risks of recessive genetic pathology (albinism, mucoviscidosis, phenylketonuria …). This risk is higher in the populations where preferential mating between related individuals is common.

D. Usefulness of Genetic Forensic in Human Identification

I. DNA and Human Identification

DNA fingerprinting, or DNA typing (profiling) as it is now known, was first described in 1985 by an English geneticist named Alec Jeffreys, who found that certain regions of DNA contained DNA sequences that were repeated over and over again next to each other. He also discovered that the number of repeated sections present in a sample could differ from individual to individual. By developing a technique to examine the length variation of these DNA repeat sequences, Dr. Jeffreys created the ability to perform human identity tests.

These DNA repeat regions became known as VNTRs, which stands for *variable number of tandem repeats*. The technique used by Dr. Jeffreys to examine the VNTRs was called restriction fragment length polymorphism (RFLP) because it involved the use of a restriction enzyme to cut the regions of DNA surrounding the VNTRs. This RFLP method was first used to help in an English immigration case and shortly thereafter to solve a double homicide case. Since that time, human identity testing using DNA typing methods has been widespread.

The past two-and-a-half decades have seen tremendous growth in the use of DNA evidence in crime scene investigations as well as paternity and genetic genealogy testing. Today around 200 public and private forensic laboratories and several dozen private paternity testing laboratories conduct hundreds of thousands of DNA tests annually in North America. In addition, most countries in Europe, South America, and Asia as well as Australia, New Zealand, and some countries in Africa have forensic DNA programs. The number of laboratories around the world conducting DNA testing

will continue to grow as the technique gains in popularity within the law enforcement community (Li, 2008).

Since the mid-1990s, computer databases containing DNA profiles from crime scene samples, convicted offenders, and in some cases persons simply arrested for a crime, have provided law enforcement with the ability to link offenders to their crimes. Application of this technology has enabled tens of thousands of crimes — particularly horrible serial crimes by repeat offenders, to be solved around the world (Butler, 2010).

Due to the growth and effectiveness of national DNA databases of convicted offenders, which now contain in some cases millions of DNA profiles with a specific set of core STR loci, it is unlikely that other classes of genetic markers will have a major impact in the forensic community for the foreseeable future. Rather, single nucleotide polymorphisms (SNPs) and other possible forensic DNA typing systems will probably see use in a supplemental rather than a supplanting role over the core STR loci that provide the common currency of data exchange in today's national DNA databases.

There are a number of applications for human identity testing involving DNA analysis using STR markers including (1) parentage analysis to help identify a child's father, (2) disaster victim identification to literally help put the pieces back together after a major natural or man-made disaster, (3) genetic genealogy and ancestry tests to attempt to gain a better understanding of one's heritage, and (4) historical and missing persons investigations to help link recovered bones from unknown individuals back to their family members.

In DNA typing, multiple markers or loci are examined. The more DNA markers examined and compared, the greater the chance that two unrelated individuals will have different genotypes. Alternatively, each piece of matching information adds to the confidence in connecting two matching DNA profiles from the same individual. If each locus is inherited independent of the other loci, then a calculation of a DNA profile frequency can be made by multiplying each individual genotype frequency together. This is known as the *product rule* (National Research council, 1996).

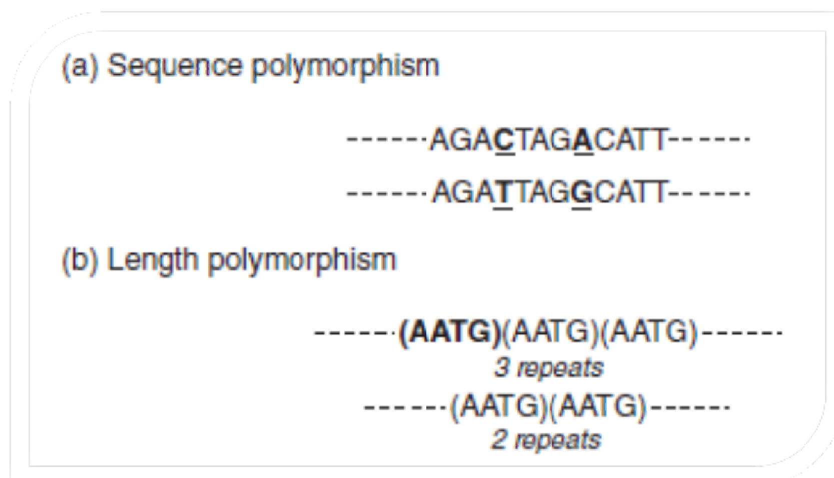


Figure 4

Short Tandem Repeats Structure. *Two primary forms of variation exist in DNA: (a) sequence polymorphisms; (b) length polymorphisms. The short tandem repeat DNA markers discussed in this book are length polymorphisms(Butler, 2010).*

II. Usefulness of the different genetic markers in human identification

1. Short Tandem Repeats (STR)

A short tandem repeat (STR) is a DNA locus containing short segments of DNA (usually four nucleotides long for forensic DNA analysis) that are repeated one after another (in tandem) in differing numbers throughout the population. There are thousands of STRs that can potentially be used for forensic analysis. STR loci are spread throughout the genome, including the 22 autosomal chromosomes and the X and Y sex chromosomes. They have a core unit of between 1 bp[2] and 6 bp and the alleles typically range from 50 bp to 300 bp. The majority of the loci that are used in forensic genetics are tetranucleotide repeats, which have a 4 bp repeat motif (Figure 4).STR DNA segments do not code for proteins or other phenotype characteristics, and therefore no dominance exists for these loci. Both, the maternal and paternal alleles for each homologous STR locus are present in an individual; a person is heterozygous for a particular STR if the maternal and paternal alleles differ; a person is homozygous for that locus if the alleles are identical. The detectable differences in STR alleles are

caused by size variations created by differences in the number of times the four-base units are repeated (Goodwin, 2011).

The STRs that are used for forensic DNA analysis are all highly ***polymorphic*****,** which means they are present in many forms in the population. The different STRs chosen for DNA identification all assort independently of each other because they are either located on different chromosomes or are far apart on the same chromosome. Thus, the ***probability*** of a person having any particular STR allele can be multiplied by the probability of them having any other, leading to the great ***power of discrimination*** achieved when many different STRs are examined together (National Research Council, 1996).

STRs are currently the most commonly analyzed genetic polymorphism in forensic genetics. They were introduced into casework in the mid-1990s and are now the main tool for just about every forensic laboratory in the world – the vast majority of forensic genetic casework involves the analysis of STR polymorphisms (Kobilinsky et al, 2006).

STRs satisfy all the requirements for a forensic marker: they are robust, leading to successful analysis of a wide range of biological material; the results generated indifferent laboratories are easily compared; they are highly discriminatory, especially when analyzing a large number of loci simultaneously (multiplexing); they are very sensitive, requiring only a few cells for a successful analysis; it is relatively cheap and easy to generate STR profiles; and there is a large number of STRs throughout the genome that do not appear to be under any selective pressure.

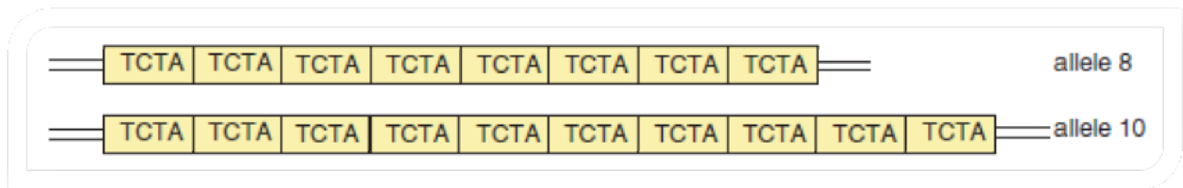


Figure 5

STR's allele structure*. This example shows the structure of two alleles from the locus D8S1179.1 The DNA either side of the core repeats is called flanking DNA. The alleles are named according to the number of repeats that they contain – hence alleles 8 and 10(Butler, 2010).*

2. Nomenclature of STR markers

The nomenclature for DNA markers is fairly straightforward. If a marker is part of a gene or falls within a gene, the gene name is used in the designation. For example, the short tandem repeat (STR) marker TH01 is from the human *t*yrosine *h*ydroxylase gene located on chromosome 11. The '01' portion of TH01 comes from the fact that the repeat region in question is located within intron1 of the tyrosine hydroxylase gene. Sometimes the prefix HUM- is included at the beginning of a locus name to indicate that it is from the human genome. Thus, the STR locus TH01 would be correctly listed as HUMTH01.DNA markers that fall outside of gene regions may be designated by their chromosomal position. The STR loci D5S818 and DYS19 are examples of markers that are not found within gene regions. In these cases, the 'D' stands for DNA. The next character refers to the chromosome number, 5 for chromosome5 and Y for the Y chromosome. The 'S' refers to the fact that the DNA marker is a single copy sequence. The final number indicates the order in which the marker was discovered and categorized for a particular chromosome. Sequential numbers are used to give uniqueness to each identified DNA marker. Thus, for the DNA marker D16S539:

D16S539

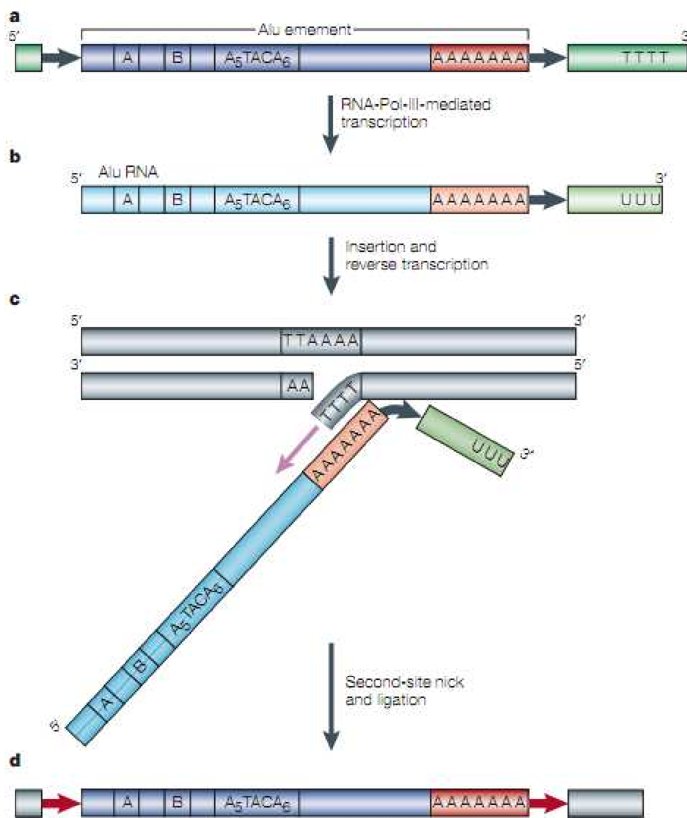
D**:** DNA; ***16*****:** chromosome 16; ***S*****:** single copy sequence; ***539*****:** 539th locus described on chromosome 16

3. ALU insertions polymorphism

Alu insertion polymorphisms and other SINE elements are robust markers for evolutionary and phylogenetic studies, because, they have a unique mutational mechanism, an absence of back mutation, and a lack of recurrent forward mutation. A specific Alu insertion and nearby flanking sequences are identical by descent in all individuals in whom they occur. Thus, sets of related chromosome regions marked by an Alu insertion event can be distinguished from a pool of ancestral chromosomes that lack the element. These features give each locus genetic polarity that allows the independent assignment of an ancestral state and a root for phylogenetic analyses. One should take care to include only those markers which show high hetrozygosity.

Alu elements are the most successful and largest class of repetitive mobile sequences that are dispersed ubiquitously throughout the primate genome. They were named after the AluI restriction enzyme site within the consensus Alu sequence (280 bp in length). The full length Alu elements are 300 bp long and commonly found in introns, 3'UTR and intergenic genomic regions. They are the most abundant SINES comprising 10% of the genome reaching a high copy number of 1.1 million during the last 65 Myr.

Alu elements have amplified by duplicating via an RNA intermediate that is reverse transcribed by target primed reverse transcription and integrated into the genome, hence, Alu are postulated to be the retro transposons. Alu elements are unable to retropose autonomously, so they are thought to appropriate the necessary mobilization machinery from the long interspersed elements retro transposons family, which encodes a protein possessing endonuclease and reverse transcriptase activity (Deiniger et al. 1992 and Donaldson et al.2002).



The structure of each Alu element is bi-partite, with the 3' half containing an additional 31-bp insertion (not shown) relative to the 5' half. The total length of each Alu sequences ~300 bp, depending on the length of the 3' oligo(dA)-rich tail. The elements also contain a central A-rich region and are flanked by short intact direct repeats that are derived from the site of insertion (black arrows). The 5' half of each sequence contains an RNA-polymerase-III promoter (A and B boxes). The 3' terminus of the Alu elemental most always consists of a run of As that is only occasionally interspersed with othe rbases (a).

Alu elements increase in number by retrotransposition — a process that involves reverse transcription of an Alu-derived RNA polymerase III transcript. As the Alu element does not code for an RNA-polymerase-III termination signal, its transcript will therefore extend into the flanking unique sequence (b). The typical RNA-polymerase-III terminator signal is a run of four or more Ts on the sense strand, which results in three Us at the 3' terminus of most transcripts. It has been proposed that the run of As at the 3' end of the Alu might anneal directly at the site of integration in the genome for target-primed reverse transcription (mauve arrow indicates reverse transcription) (c). It seems likely that the first nick at the site of insertion is often made by the L1 endonuclease at the TTAAAA consensus site. The mechanism for making the second-site nick on the other strand and integrating the other end of the Alu element remains unclear. A new set of direct repeats (red arrows) is created during the insertion of the new Alu element (d).

Figure 6:
A typical human Alu element and its retroposition (Batzer et al. 2002).

The largely human specific Alu subfamilies represent only ~ 0.5% of all of the Alu repeats in the human genome in the past 4 - 6 million years after the divergence of humans and African Apes, but most of them integrated before the African radiation of humans, so these Alu repeats are monomorphic for their insertion sites among diverse human genome. However, approximately 25% of the young Alu repeats have inserted so recently that they are dimorphic for the presence or absence of the insertion, which makes them useful source of genomic polymorphisms.

Alu elements that are present in the genomes of some individuals and absent from others are referred as ***Alu insertion polymorphisms***.

These are the autosomal markers that thus reflects both the maternal and paternal history of a population, moreover, they are the stable markers that unique evolutionary events namely the insertion of an Alu element into a new chromosomal location. These polymorphisms have several characteristics that make them unique markers for the study of human population genetics (Batzer et al. 1990; Caroll et al. 2001 and Roy-Angel et al. 2001).

Individuals that share Alu insertion polymorphism have inherited the Alu elements from a common ancestor, which makes the Alu insertion alleles identical by descent. Also, because the likelihood of two Alu elements independently inserting into the same location of the genome is extremely small, and as there is no known biological mechanism for the specific excision of Alu elements from the genome, so Alu insertions can be considered homoplasy free characteristic. Furthermore, the ancestral states of this polymorphism are known to be the absence of Alu elements at a particular genomic location. Thus, this facilitates to construct the population phylogeny.

4. ALU family classification

The human specific Alu are further classified as sub families according to diagnostic nucleotide substitution along their sequence. Due to this characteristic, they can serve as unique evolutionary milestones. Phylogenetic studies of Alu elements suggest that only a small number of Alu elements deemed "master" or source genes are retropositionally competent, overtime the eventual accumulation of new mutations within the master or source genes created the hierarchy of Alu subfamilies. Diagnostic mutation sites can be used to classify each individual element according to subfamily. Alu classification is based upon age which is:

1. The oldest (J)
2. Intermediate (S)
3. Youngest (Y).

Some young Alu subfamilies have amplified so recently that they are virtually absent from the genome of nonhuman primates, as the result of the recent integration of young

Alu subfamily members with in the human genome, individual human can be polymorphic for the presence of Alu element at particular loci. Almost all of the recently integrated human Alu elements belong to one of the several small and closely related young Alu subfamilies known as Y, Yel, Yc2, Ya5,Ya5a2, Ya8, Yb8, Yb9 and Yd (Roy-Angel et al. 2001 and Batzer et al. 1995).Although some newly integrated Alu elements result in detrimental mutation events in the human genome, the vast majority of recently integrated Alu elements have had no apparent negative impact on the genome and represent new, essentially neutral, mutation events. After a new, neutral Alu insertion integrates into the genome, it is subjected to **genetic drift**. So, the probability that it will be lost from the population is initially quite high, depending on the size of the population (the greater the population size, the more likely it is to be lost). But, over a short period of time, the Alu element will increase in frequency in the population. Because the amplification of Alu repeats is a continuing process, a series of Alu elements must have integrated into the Alu-insertion polymorphisms are essentially **homoplasy** free characteristics that can be used to study human population genetics. In addition, there is no evidence for any type of process that specifically removes Alu elements from the genome; even when a rare deletion occurs; it leaves behind a molecular signature. By contrast, other types of genetic polymorphism, such as variable numbers of tandem repeats, RFLPs and single-nucleotide polymorphisms (SNPs), are merely identical by state; that is, they have arisen as the result of several independent parallel mutations at different times and have not been inherited from a common ancestor. Alleles that are identical by descent have been directly inherited from a common ancestor. Alleles that are identical by state have the same character state, but have not been inherited from a common ancestor.

Alu insertion polymorphisms have several attractive features that make them unique elements for the study of human population genetics. First of all, the genotypes of Alu insertion polymorphisms are easy to determine by typing with rapid, nonradioactive, simple PCR based assays. They are biallelic polymorphisms with three possible genotypes: homozygous for the presence of the Alu element, heterozygous with one chromosome having the Alu element and the other lacking it and homozygous for the absence of the Alu element (Batzer et al. 1995).

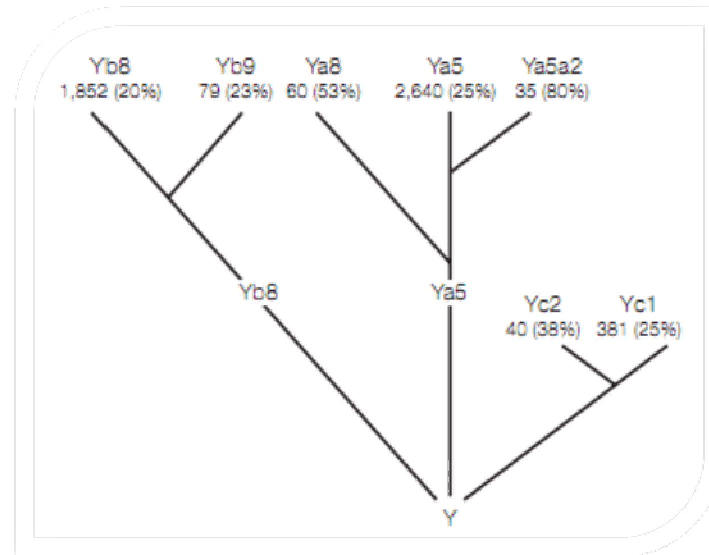


Figure 7
Expansion of recently integrated human Alu subfamilies.
Several subfamilies of Alu elements have expanded simultaneously in the human genome primarily from three Y-subfamily lineages, termed 'Ya', 'Yb' and 'Yc' in accordance with standard Alu nomenclature on the basis of commonly shared mutations. The approximate copy number of each subfamily is given as estimated from computational analysis of the draft sequence of the human genome. The percentage of insertion polymorphisms in each family is given In brackets. Alu subfamilies with smaller copy numbers and higher levels of insertion polymorphism are generally thought to be more recent in origin in the human genome (Batzer, 2002)

Secondly, once inserted into a new location, an Alu element is rarely subject to deletion. Even if deleted, it would not be an exact excision, but instead it would leave behind a molecular signature of the original insertion event by either retaining a part of the Alu element and/or deleting some of the flanking region (Novick et al. 1992).Therefore, Alu insertion polymorphisms are stable markers that reflect a unique evolutionary event, which is the insertion of an Alu element into a new chromosomal location.

Thirdly, Alu insertion polymorphisms display unique events that occurred during human evolution. Since there are 3 billion nucleotides in the haploid human genome, the probability that an Alu element would insert between the exact same two nucleotides at two different times during evolution is insignificant. Therefore, there is no parallel gain or loss of Alu elements at a particular chromosomal location, so all chromosomes that carry a polymorphic Alu element must be identical by descent. Hence, polymorphic Alu insertions reflect population relationships more accurately than

other genetic markers, such as RFLP, SNPs, STRs, microsatellites, mtDNA markers, *etc*. The disadvantage of these latter genetic markers lies in the fact that they have arisen as the result of several independent parallel mutations at different times. Therefore, they are identical by state rather than descent and thus may have not been inherited from a common ancestor. This is because the same allele could arise independently at different times during human evolution (Edwards et al. 1992).

Lastly, the ancestral state of Alu insertion polymorphisms is known to be the absence of the Alu element at a particular chromosomal location and the derived state is the presence of the Alu element. The precise knowledge of the ancestral state of a genomic polymorphism, which is very important in phylogenetic analyses, permits the construction of phylogenetic trees without making too many assumptions (Batzer et al. 2002; Batzer et al. 1994; Batzer et al. 1996a; Stoneking et al. 1997).

III- X Chromosome: properties and relevance in human identification and population genetics

Both the intrinsic characteristics of a marker and of the population of interest are responsible for the choice of marker in population-genetic studies.

The X chromosome has several characteristics that distinguish it from the autosomes despite their physical similarity. X chromosome is the most stable nuclear chromosome spanning approximately150 million base pairs (Mb) in humans (~5% ofthe genome among mammals). The X and Y chromosomes are thought to have evolved from a single autosome ~300 million years ago. Both chromosomes are homologous and recombine with each other near their ends, in the two pseudo-autosomal regions. Elsewhere, however, they have taken different evolutionary paths. The Y chromosome has lost the bulk of both its sequence and its genes, and has developed a unique pattern of repeated sequence. By contrast, the X chromosome has not lost its autosomal character. In terms of age, autosomes record slightly older time periods than the X chromosome, but both record substantially older histories than either the Y chromosome or mtDNA (Harris et al. 1999).

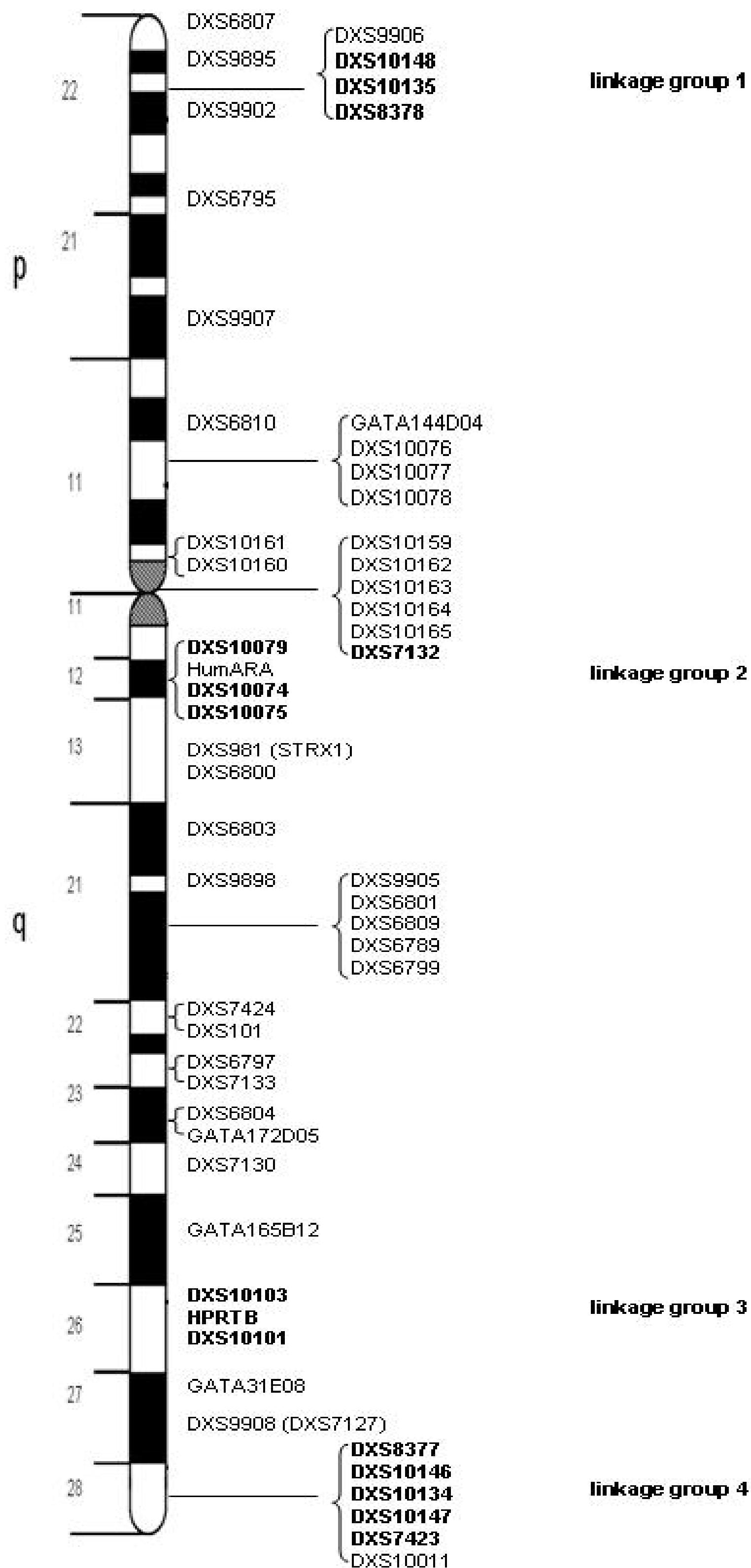


Figure 8.
Chromosome X ideogram. Localisation of ChrX STRs used in forensic practise. The order and approximate position of STRs on the ChrX ideogram is based upon publicised map data (Marshfield, NCBI). Pair-wise genetic distances (in cM) were calculated from maximum likelihood estimates of pairwise recombination fractions using the Kosambi mapping function(www.chrx-str.org).

The X chromosome has several distinguishing characteristics. First, it spends two third of its lifetime in females since males have only one copy of this chromosome. This fact, explains its lower genetic diversity because the nucleotide mutation rate in females is several-fold lower than in males. Diversity is further reduced by the effective population size (Ne) of the X chromosome, is three-quarters of that of an autosome when there are equal numbers of breeding males and females and random variation in offspring.

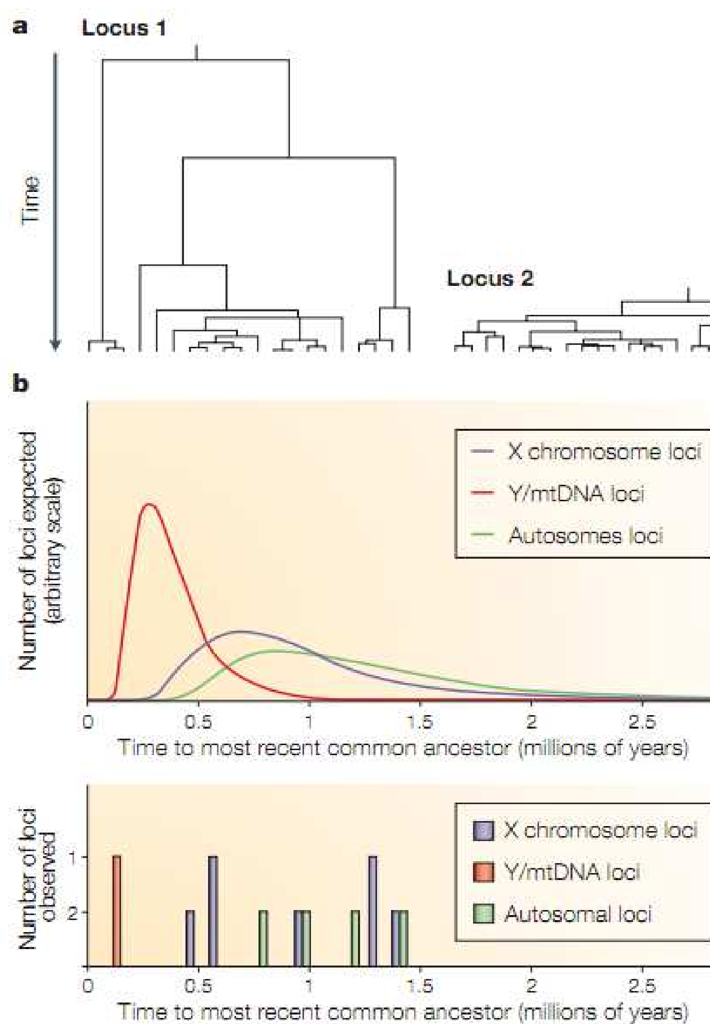
Table 1: Comparison of population-genetic markers

	Marker type				References
	mtDNA	Y chromosome	X chromosome	Autosomes	
Size (Mb)	0.017	60	150	3,000	15,16,59
Number of usable loci	1	1	Hundreds	Thousands	–
Mutation rate (mutations per Mb per generation)	Very high (1–300)	High (0.033)	Low (0.015)	Moderate (0.020)	19,60,61
Recombination rate (cM/Mb)	0	0	0.8	1.1	21
Diversity (fraction of discordant base pairs)	Very high (0.4%)	Low (0.02%)	Moderate (0.04%)	High (0.08%)	42,62
Accessible haplotypes*	Yes	Yes	Yes	No	–
Genetic drift‡§	High	High	Moderate	Low	–
Age§	100,000 years	100,000 years	750,000 years	1,000,000 years	–
Effective population size (relative to autosomes)	1/4	1/4	3/4	1	–

**A haplotype is a set of genetic markers that is present on one chromosome; ‡genetic drift describes the random changes in allele frequency that occur because genes that appear in offspring are not a perfectly representative sample of the parental genes (for example, as occurs in small populations); § these entries are approximate population genetics inferences, based on the consensus estimate for the effective population size in humans. cM, centiMorgan; Mb, megabase; mt, mitochondrial (Schaffner, 2004).*

In the other hand, recombination on the X chromosome can occur only in females; therefore, only two of the three potentially transmitted X chromosomes can be product of a recombination event as males have a single copy. The measured recombination rate for the X chromosome is, in fact, almost exactly two-thirds of the genome average. As a result, we can expect linkage disequilibrium (LD) to be greater on the X chromosome and the size of regions with a single genetic history to be larger. This effect is reinforced by the younger age of the X chromosome, as younger loci have had less time for recombination to break down LD (Yu et al. 2002).Other property of X chromosome is its lower mutation rate argued by the high mutation rate in males than females,

presumably because of the long generation time and unusually large number of male mitoses (*Estimate of the Mutation Rate per Nucleotide in Humans*). There are several reasons why the X chromosome might exhibit a lower substitution rate than the autosomes. First is male-driven molecular evolution. If most mutations arise in the male germ line, then the X chromosome is expected to have a lower substitution rate than the autosomes because the X chromosome spends only one-third of its time in males (Haldan, 1947). Second is a lower mutation rate on the X chromosome, independent of sex-specific effects to protect itself against deleterious mutations (Miyata, 1987).



Unlike the Y chromosome and mtDNA, the X chromosome contains many independent loci, each with its own phylogenetic tree. It is a characteristic of genealogies, whatever chromosome they occur on, that they vary randomly; that is, under identical circumstances, the phylogenetic trees for two loci can be very different, both in shape and in depth. The two trees shown in ***panel a*** *are the result of simulations of a constant-sized population for two loci, and are typical of the amount of variation observed. Although the two simulated loci share an identical population history, the age (and therefore the diversity) of locus 1 is many times that of locus 2; inferring the characteristics of the population from either tree alone will therefore give a badly skewed result.* ***Panel b*** *shows the full range of ages expected for the three types of chromosome, on the basis of an OUT OF AFRICA MODEL of human origins. As the X chromosome has three times the effective population size of the Y chromosome or mtDNA, loci on the X chromosome can be expected to be much older; the same is true for autosomes, which have four times the effective population of the Y chromosome. Note the broad age range expected for different loci from the same type of chromosome. The histogram shows published estimates of the age of various loci; all have large uncertainties (not shown). Similar variation from locus to locus occurs in other inferences, such as those concerning BOTTLENECKS in population size, or about the source of migrations into a region.*

Figure 9. Variation between loci (Schaffner, 2004).

X-Y chromosome comparison was carried out to compare their divergence from the inferred ancestral sequence; a higher male mutation rate will be reflected in a higher rate of substitutions on the Y chromosome copy. Measurements of male and female mutation rate estimated about a fivefold higher rate in males, but two studies have yielded ratios close to two. In the other hand, the much lower level of variation occurs at the low-recombination loci indicating that **natural selection** has a strong effect on levels of X chromosome. Both a selective sweep model and a background selection model are responsible of reduced variation in regions of low recombination and the strong differentiation between populations at low recombination genes. The selective sweep model assumes differentiation in regions of low recombination is due to the locally favored substitution. In contrast the background selective model assumes that the differentiation is caused by the continual removal of deleterious alleles in regions of low recombination which result in lower effective population sizes and thus in a lower migration rate for the low recombination locus (McVeanet al.1997).

The use of the X chromosome in population genetics is still in its infancy (Schaffner SF 2004). It has already proved its worth in studies of the early history of modern *Homo sapiens*, but in most research areas its potential remains largely untapped. That potential is needed — the Y chromosome and mtDNA, despite their enormously fruitful contributions, are not very informative about some questions (such as the size of ancestral populations), and the information that they can provide about others (such as population history before the Out of Africa migration) has largely already been mined. The X chromosome is therefore the logical place to turn for more information. Many of the same questions can be addressed by either the X chromosome or the autosomes, but the X chromosome has a clear advantage in allowing easy access to haplotypes; the cost of extracting haplotypes from autosomes remains high, even as sequencing and genotyping become much faster and cheaper.

1. Use of chromosome X markers in Forensic Genetics

Following the ISFH recommendations for the forensic application of microsatellite markers, trimeric, tetrameric and pentameric microsatellites can be used in practice if they have suitable population genetic properties (Hardy-Weinberg equilibrium, sufficiently high degree of polymorphism, known linkage disequilibrium etc.). The Genome Database (http://www.gdb.org) lists a total of26 trinucleotide and 90

tetranucleotide repeat polymorphisms on ChrX, but only 18 tetranucleotide and 3 trinucleotide STRs, plus the VNTR locus DXS52 appear to be in common forensic use. The X chromosome has many futures in forensic genetics especially in complex kinship cases:

☑ Paternity testing in trios and duos

Paternity cases involving the common trio constellation of mother, offspring and alleged father can usually be solved with autosomal (AS) STRs alone, and do not seem to require any additional or alternative markers. However, when father/daughter relationships are to be tested it may be worthwhile including ChrX markers, too. This is especially the case when difficult to analyze template materials are involved, such as DNA from exhumed skeletons or historical or prehistorical samples. In such instances sufficient statistical power has to come from a small number of low sizes STRs. Fortunately, ChrX STRs are usually characterized by relatively high Mean Exclusion Chances (MECs), even at a low to medium degree of polymorphism.

☑ Paternity cases involving blood relatives

In paternity cases involving close blood relatives as alternative putative fathers, the exclusion power of STRs is substantially decreased and ChrX STRs may be superior to AS markers. For example, if two alleged fathers are father and son, they would not share any X-chromosomal alleles identical by descent (ibd) so that ChrX markers would be more efficient than AS markers. Brothers, in contrast, share a given maternal ChrX allele with probability0.5, which equals the probability that two alleles are shared ibd at an AS locus. For three unlinked ChrX loci, the chance of ibd sharing would be 0.53=0.125. However, when the markers are closely linked, they do not segregate independently. As with AS markers, they would instead represent a single haplotype, that is again shared with a probability approaching 0.5. The ChrX contains three linkage groups which can provide nearly independent genotype information. At present, it is proposed to use clusters DXS6807-DXS9895-DXS9902-DXS8378 (Xp22–21), DXS7132-ARA-DXS6800 (Xq11-Xq13), and DXS7423-DXS8377-DXS10011 (Xp27–28) to define haplotypes for forensic practice.

☑ Paternity testing in rape and incest cases

After criminal sexual assault or incest, pregnancies may be terminated by suction abortion. An aborted 6–8 week product of conception consists of small amounts of non-identifiable fetal organs as well as maternal blood and other tissues. In such cases, the microscopically dissection of chorionic villi is not generally successful, and samples most often contain a mixture of fetal and maternal DNA. Efficient paternity testing of such material is still possible for male fetuses, using ChrY markers. Paternity testing of female fetuses, in contrast, can only include AS and ChrX markers, the latter of which represent a more efficient means of paternity exclusion under all circumstances. A positive proof of paternity, however, relies mainly upon fetal alleles not shared with the respective mother. In incest cases in which a father is rightfully charged with abusing his daughter, ChrX testing of an abortion can therefore contribute only very limited information towards a positive proof of paternity. This is because all fetal alleles would necessarily coincide with alleles of the daughter.

Maternity testing. There are situations in which mother/child testing may be required. For example, due to the high rate of illegitimate paternity in modern societies, the identification of skeletons or corpses by mother/child testing is more reliable than through the assessment of father/child relationships. Although maternity can be demonstrated by sequencing mitochondrial DNA, this technique is nevertheless expensive and does not always yield the level of certainty required in forensic science. This is especially the case when individuals are involved for whom appropriate population genetic data are not available. Typing of ChrX STRs may thus represent a sensible alternative option when assessing maternity. For testing mother-daughter relationships, ChrX markers are equivalent to AS markers and do not provide any specific advantage. Testing mother-son kinship, however, is more efficiently performed using ChrX markers. The exclusion chance in such cases is identical to that of ChrX STRs in father/ daughter tests.

☑ Deficiency paternity cases

The major advantage of ChrX markers arises in deficiency paternity cases, i.e. when a biological sample from a putative father is not available and DNA from paternal relatives has to be analyzed instead. When female individuals have the same father, they also share the same paternal ChrX. An investigation of ChrX markers of two sisters or half-sisters can thus exclude paternity, namely through the presence of four different alleles or haplotypes, even when none of the parents is available for testing. AS markers cannot provide such information. A positive proof of paternity is also possible without parental genotype information, but is generally less reliable. This is due to the fact that sisters usually inherit only partially matching haplotypes from their mother (Szibor et al, 2003).

A. X Chromosome: Alu insertion polymorphism analysis

I. Material

A total of **892** blood samples from healthy unrelated individuals male and female were collected from sixteen Mediterranean populations. Samples from the three ethnical groups of Morocco were obtained: 49 (Arabic), 46 (Berber) and 34 (Sahrawi). In the Balearic Islands, 53 individuals were sampled from Mallorca, 45 from Minorca and 70 from Ibiza; 60 individuals from the Iberian Peninsula (Valencia) and 288 from South of Italy (79 samples from Sicily, 30 from Cosenza, 61 from Catanzaro and 118 from Calabria). Five Jewish communities were also sampled: 43 Sephardic, 45 North African, 41 Oriental, 93 Chuetas and 25 Ashkenazi. The regions of sampling for all of the individuals studied are represented in Figure 2. Samples were collected from healthy individuals and special care was taken to avoid sampling from related individuals. The sampling was anonymous in order to prevent linkage to the original donor and every person investigated gave informed consent. Individuals whose families had been from a same particular area for at least three generations were selected and this was recorded with the questionnaires filled by the donors. As an example, the questionnaire used in Moroccan samples can be seen in the Appendix. The permission of the DNA donors was also taken and a form to obtain the consent of the sampled individuals is shown in the Appendix.

The DNA samples of Balearic Islands, Valencia and Chueta populations belonged to the collection of the Genetics Laboratory, University of Balearic Islands. While, the DNA samples of Calabria Region and Sicily populations belonged to the collection of the Studio Indagini Mediche e Forensi (SIMEF) (Reggio Calabria, Italy), and the DNA samples of Morocco were provided from the Laboratory of Biochemistry and Immunology-Forensic Genetics Unit DNA database. DNA samples of the four Jewish populations (Ashkenazi, Sephardic, Oriental and North-African) were obtained from the

collection of The National Laboratory for the Genetics of Israeli Populations at Tel-Aviv University.

II. Methods

1. DNA extraction

Blood samples (~ 10 ml) were collected into EDTA containing tubes to prevent coagulation and stored at +4 °C until use. 10 ml blood was completed to 50 ml with 2X lysis buffer to lyse the red blood cells. Tubes were mixed for 10 minutes by inversion and then centrifuged at 3000 rpm for 10 minutes at +4 °C to precipitate nuclei. The pellet was resuspended in 3 ml of salt-EDTA buffer by vortexing. Then, 0.3 ml of 10% SDS and 150 ml proteinase-K (10 mg/ml) were added and the tubes were incubated at +55 °C for 3 hours. After the incubation, 3 ml phenol was added and this suspension was centrifuged at 3000 rpm for 10 minutes at +4 °C. The supernatant was mixed with 3 ml phenol-chloroform-isoamylalcohol solution (25:24:1) and centrifuged at 3000 rpm for 10 minutes at +4 °C. Then, the upper phase was collected carefully and transferred into a new glass tube by a transfer pipette. Then 1 ml 3 M Sodium Acetate (NaAc) and about 2 volumes of 99% ethanol were added to precipitate and collect the DNA. Glass tubes were mixed gently by inversion and then the DNA was transferred to eppendorf tubes containing 500 ml Tris-EDTA (TE) buffer (pH: 7.5) and stored at -20 °C.

2. Quantification of isolated DNA

Each DNA sample was quantified using the NanoDrop® ND-8000 spectrophotometer for the Moroccan samples and the NanoDrop® ND-1000 spectrophotometer for European and Jewish samples. The resulting concentrations and purity (A_{260}/A_{230} values) demonstrated that neither yield nor sample purity were affected by the blood collection tube type.

3. Amplification of DNA with Polymerase Chain Reaction (PCR)

The PCR amplification of the DNA samples for each Alu insertion Polymorphism (Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4 and Ya5DP13) was performed in 20µl amplification reactions using the components presented in table 2.

Table 2.
PCR Protocols: X chromosome Alu insertions (All numbers are in µl/well; T_a: annealing temperature; DNA template should be around 100 ng/ml)

ElémentAlu	DNA (any concentration between 20-100 ng/µl)	Taq. (5U/µl)	Buffer (10X)	MgCl (50mM)	dNTP (2,5mM foreachdNTP)	Primer 5' (100 pmol/µl)	Primer3' (100 pmol/µl)	H2O	Volume total (µl)
Ya5DP62	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Yb8DP49	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Yd3JX437	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Yb8NBC634	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Ya5DP77	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Ya5NBC491	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Yb8NBC578	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Ya5DP4	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20
Ya5DP13	1	0.2	2	1.2	1.5	0.1	0.1	13.9	20

Table 3.
Oligonucleotidprimer sequencesand sizes of the PCR products in the presence of each Alu insertion

ALU	Size (bp) of the PCR product with the insertion and without it	Forward primer	Reverse primer
Ya5DP62	894/ ~600	CTGCCTTCCATGAACGTTTT	AAATGAAGCCTTGCTGCTGT
Yb8DP49	700/380	GACTAGGGGTTTGTGCCAGA	TCCCCCATTTCTGTTGTTGT
Yd3JX437	547/235	TGGTGTACCTTAGTCCAAAGACC	TTTGCATCTCAGAACTTTTTCCT
Yb8NBC634	420/95	AACAGAAAGGCATCATTTGC	GGGGGCATTTATTACTGCTT
Ya5DP77	620/298	GAAGGATGATCTCTCCTTAC	TGCAAGGAGAGTTGGCATAA
Ya5NBC491	435/96	ACATGAATGTGCCATTGGTT	CAAGAAGGCAGCTGTCCTAGA
Yb8NBC578	380/72	TTTTTGCAGATGCTTCCCTA	CCCTTGATCCAGATGTGATG
Ya5DP4	649/334	AACACCTCTGATGTAGCTTATG	CTAGGCCACCATTAAGCCAA
Ya5DP13	494/141	CTAGTGGAAGCTCCGTTTGG	AGGCTGACCACTGGATATGG

Hot start at 94 °C for 5 minutes is only applied to genomic DNA just before the addition of the reaction mixture to improve the accuracy of primer annealing. The

specific oligonucleotide primer sequences together with the annealing temperatures used for each Alu insertion are given in Table 3 and 4.

Table 4: Thermocycling conditions for the 9 Alu polymorphism

Alu	Dén. initial		Dénaturation		Hybridation		Extension		Finalisation		∞	Cicles (D-H-E)
Ya5DP62	94°	5’	94°	1’	57.5	1’	72°	1’	72°	5’	4°	32
Yb8DP49	94°	5’	94°	1’	57.5	1’	72°	1’	72°	5’	4°	32
Yd3JX437	94°	5’	94°	1’	58	1’	72°	1’	72°	5’	4°	32
Yb8NBC634	94°	5’	94°	1’	57.5	1’	72°	1’	72°	5’	4°	32
Ya5DP77	94°	5’	94°	1’	58.9	1’	72°	1’	72°	5’	4°	32
Ya5NBC491	94°	5’	94°	1’	58	1’	72°	1’	72°	5’	4°	32
Yb8NBC578	94°	5’	94°	1’	56.5	1’	72°	1’	72°	5’	4°	32
Ya5DP4	94°	5’	94°	1’	59	1’	72°	1’	72°	5’	4°	32
Ya5DP13	94°	5’	94°	1’	58	1’	72°	1’	72°	5’	4°	32

4. Analysis of the PCR products

Analysis of the PCR products was done by agarose gel electrophoresis. 2 % agarose gel was prepared by boiling agarose in 0.5X TBE buffer, pouring it into an electrophoresis plate and leaving it at RT for 30 minutes for polymerization. 9 ml of each PCR product was mixed with 7 ml of 6X bromophenol blue dye and loaded into the wells of the gel. The gel was run in 0.5X TBE buffer at 100 V until the bands reached the end of the gel. Then, gel was stained in 0.5 ml/ml ethidium bromide (Et-Br) solution and the amplification products were directly visualized by UV fluorescence. The photograph of the gel was obtained by a gel image system. Some examples of gel photographs showing the result of the amplification of Alu insertions with different genotypes are shown in Figures 10-11. The length of the PCR products is given in Table 3.

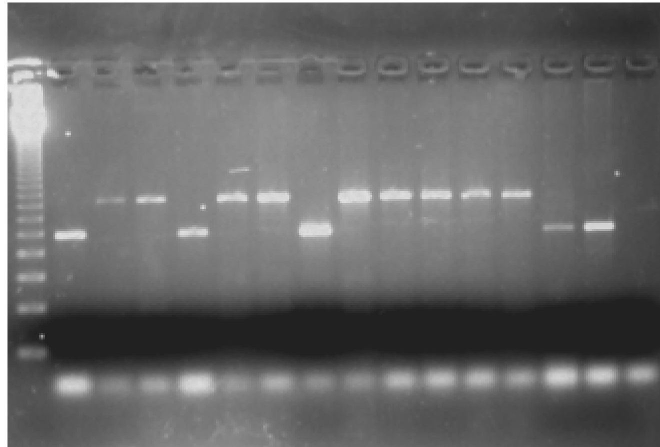


Figure 10. Photograph of a 2 % agarose gel containing the PCR products of Ya5DP62.

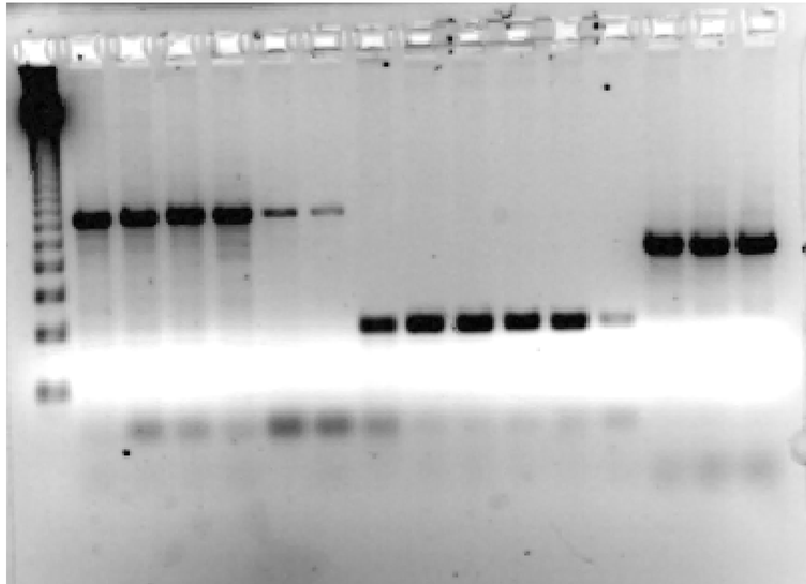


Figure 11. Photograph of a 2 % agarose gel containing the PCR products of Yb8NBC578, Yd3JX437, Ya5NBC491.

5. Statistical Analysis of Data

☑ *Allele Frequencies, Heterozygosities and Evaluation of Hardy-Weinberg Equilibrium*

The statistical analysis of the data started with the calculation of relative frequencies of each Alu insertion polymorphism in each gender apart, and then used to calculate the total allelic frequency in the studied populations. This calculation gives the allele

frequencies of each Alu insertion, which are used frequently in the subsequent steps of the data analysis. Allele frequencies of the insertions were calculated according to the formula given below;

$$\text{Total allelic frequency} = \frac{(2 \text{ x women allelic frequency}) + (1 \text{ x men allelic frequency})}{3}$$

The men's allelic frequency is established by direct gene counting and the women's allelic frequency according to the following formula:

$$\text{Women's allelic frequency} = \frac{(2 \text{ x homozygotes}) + (1 \text{ x heterozygotes})}{2N}$$

N is the total number of women

Heterozygosity as well as HWE were calculated only for females.

Heterozygosity which helps to measure the genetic variation in a population was calculated with the formula below;

$$\text{Observed Heterozygosity} = \frac{\text{Number of heterozygotes}}{\text{Total number of individuals}}$$

Calculation of the the observed and expected heterozygositiesas well as *p*-value were performed with GENEPOP(Rousset, 2010) and verified with ARLEQUIN 3.0(Excoffier et al. 2005). The imput file for the two programs was generated using GENETIX 4.0 software(Belkhir et al. 2001).

In statistical significance testing, the *p*-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. One often "rejects the null hypothesis" when the *p*-value is less than the significance level α, which is in our case 0.05. When the null hypothesis is rejected, the result is said to bestatistically significant.

Furthermore to the *p* value, the Hardy-Weinberg (HW) equilibrium in our populationswas performed by calculating the expected genotype frequencies and comparing them with the observed ones. HW equilibrium is based on the following assumptions: (i) mating is random, (ii) allelic frequencies are conserved from generation

to generation, (iii) no significant migrations occur, (iv) mutation, selection, genetic drift and gene flow are negligible (Hedrick, 2000). All of these requirements were assumed to be true while calculating the expected genotype frequencies, which were calculated as follows;

Expected frequency of homozygotes = p^2 and q^2

Expected frequency of heterozygotes = 2pq, where p + q = 1

The agreement between observed and expected values was tested by the chisquare (χ^2) test statistic (Daniel, 1999). The general formula for $\chi^2$is;

$\chi^2 = \sum$ (Observed values - Expected values)^2/ Expected values

Chi-square values for each Alu insertion were calculated and these values were evaluated in chi-square distribution. The decision on the null hypothesis (Ho), which states that observed and expected frequencies are not different from each other, was made.

☑ *Reynolds Genetic Distance (DA) and Neighbor Joining (NJ) Tree*

Genetic distance analysis, which focuses on average genetic distance between populations, is quite efficient while constructing an evolutionary tree from allele frequency data. Reynold’s genetic distances, between the pairs of the Western Mediterranean population and other populations (Stoneking et al 1997; Nasidze et al. 2001 and Romuladi et al. 2002)were calculated with the GENDIST program in the PHYLIP program package (Felsenstein, 1993). Neighbor-joining trees were produced by using the NEIGHBOR program, 1000 bootstrap replicates were generated by the SEQBOOT program and a consensus tree was built with the CONSENSE program as implemented in PHYLIP 3.6 which is available at http://evolution.genetics.washington.edu/phylip.html.

☑ *Multidimensional Scaling (MDS)*

To analyze population relationships and to determine the relative positions of populations in 3 dimensional space, the Multi-Dimensional Scaling analysis (MDS) was performed using a computer program called SPSS 15.0: Statistical Package for the Social Sciences (Inc, 2007).

In MDS from the input matrix giving dissimilarities between pairs of populations, a set of independent compound axes are configured and relative positions of the populations are visualized in the space generated by these axes. The first axis will explain the highest variation of the all data that can be accounted by the compound axes; the second will explain the next highest variation, and so on. Inspection of the weightings of the first few axes will show which variables contribute most to the differences between individuals.

☑ *The AMOVA analysis*

A series of hierarchical measures of heterozygosity are defined:

H_I = mean observed heterozygosity per individual within subpopulations.
H_S = mean expected heterozygosity within random mating subpopulations = $2p_iq_i$
H_T = expected heterozygosity in random mating total population = $2pq$

Using these three different hierarchical measures of H, we can define three hierarchical F-statistics, defined below:

- ✓ ***Inbreeding coefficient***, that measures the extent of genetic inbreeding within subpopulations. It can range from -1.0 (all individuals are heterozygous) to +1.0 (no observed heterozygotes)

$$F_{IS} = (H_S - H_I)/H_S$$

- ✓ ***Fixation Index***, that measures the extent of genetic differentiation among subpopulations. It ranges from 0.0 (no differentiation) to 1.0 (complete differentiation – subpopulations fixed for different alleles).

$$F_{ST} = (H_T - H_S)/H_T$$

- ✓ ***Overall Fixation Index***, that means the reduction in H of an individual relative to the total population.

$$F_{IT} = (H_T - H_I)/H_T$$

The relationship between the three F-statistics is:

$$(1\text{-}F_{IT}) = (1\text{-}F_{IS})\ (1\text{-}F_{ST})$$

Polymorphism and population structure within and between groups was tested by analysis of molecular variance (AMOVA) using the ARLEQUIN v3.0 program. Global F_{ST} values were estimated by averaging partial values, and the resultant probability was calculated by combining probabilities from each individual test.

☑ *Structure analysis*

Apart from the usual distance-based clustering methods described above, a model-based method was also used to infer population structure by means of the STRUCTURE 2.1 program(Pritchard et al. 2000).A model of K population groups (where K might be unknown) was assumed. This model was tested for several values of K using a specific Markov Chain Monte Carlo algorithm (the Gibbs sampler). STRUCTURE estimates the 'natural logarithm of the probability of the data' for each value of K, briefly referred to as 'Ln P (X\K)'. Among the estimated K values, that yielding the lowest absolute value of the Ln P (X\K) is the one that best describes the data. In our data sets, we ran the Gibbs sampler under the admixture model (INFERALPHA¼1.0), using prior population information and assuming correlated allele frequencies. All runs included a burn-in period of 50 000 iterations followed by 106 iterations, and they were repeated three times each in order to test the consistency of the results.

B. X CHROMOSOME: Genetic Analysis of 12 STR polymorphism

I. Material

Blood samples were collected from **697**unrelated, healthy individuals. **145** individual (97 men and 48 women) from the three ethnic groups (Arab, Berber and Sahrawi) living in the whole area of Morocco; **250** Spanish samples from the Balearic Islands and Valencia: (Majorca: 39 men and 22 women; Minorca: 39 men and 25 women; Ibiza: 39 men and 20 women and Valencia: 39 men and 27 women); and **302** Jewish samples from different worldwide areas (Ashkenazis: 26 men and 33 women; Sephardic (Bulgaria & Turkey): 19 men and 25 women; North African (Morocco, Tunisia and Libya): 29 men and 25 women, Oriental (Iraq & Iran): 29 men and 26 women and Chuetas (Majorca): 49 men and 41 women). All participants signed the annexed informant consent and provided information about their ethno-linguistic as well as their parents and grandparents origins.

II. Methods

1. DNA extraction

The DNA extraction was performed from blood samples using DNA IQ™ System (www.promega.com). It is a DNA isolation and quantification system designed specifically for the forensic and paternity communities. This system employs a novel technology with magnetic particles to prepare clean samples for short tandem repeat (STR) analysis easily and efficiently. The DNA IQ™ System can be used to extract DNA from stains or liquid samples such as blood or resulting DNA solutions.

The DNA IQ™ resin has defined DNA-binding capacity in the presence of excess DNA. Yields will be consistent within a single sample type but will differ with different

sample types. The typical yield for FTA®blood card punches is 50-100ng, for liquid blood, 50-200ng; and for buckle swab, 100-500ng.

The DNA IQ™ System avoid the use of some harmful organic products like phenol and consume less time than other classical methods as it reduce many centrifugation and purification steps. Since recovery depends on the sampling method, solid support and sample type, laboratories will need to determine the average yield for a single sample type. Once this average yield has been determined, the researcher can bypass the quantification step typically necessary with other purification procedures.

2. Quantification of isolated DNA

Each DNA sample was quantified using the Termo Scientific NanoDrop® ND-8000 spectrophotometer. The resulting concentrations and purity (A_{260}/A_{230} values) demonstrated that neither yield nor sample purity were affected by the blood collection tube type.

3. Amplification of DNA with Polymerase Chain Reaction (PCR)

A number of 12 X-chromosome STR markers (DXS10148, DXS10135, DXS8378, DXS7132, DXS10079, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS10134 and DXS7423) were amplified using the Investigator Argus X-12 Kit (Qiagen GmbH, Hilden, Germany) (Tables 5 and 6). The manufacturer's recommendations were followed.

The Investigator Argus X-12 Kit (www.Qiagen.com) enables simultaneous amplification of 12 X-chromosomal STR loci, which are highly informative for kinship and paternity testing, as well as population genetics and anthropological studies. The heightened discriminatory power of the kit meets the demands of complicated deficiency cases involving at least one female. In addition, this kit is highly suited for analysis of forensic stains, such as female traces on a male background, and is an informative supplement to the Investigator Argus Y-12 QS Kit for kinship and paternity testing. The markers of the Investigator Argus X-12 are clustered into 4 linkage groups

(3 markers per group), and thus each set of 3 markers is handled as a haplotype for genotyping (Figure 12).

The polymerase chain reaction (PCR) of 12 X chromosomal STR markers was performed in a multiplex reaction. For each sample, 2 µl of DNA were amplified in a final reaction volume of 26 µl. The final reaction mixture contains: 15.9 µl of Nuclease-free water, 5 µl of reaction mix A, 2.5 µl of primer mix and 0.6 µl of Multi Taq2 DNA Polymerase.Positive and negative controls were included in each PCR run.

The PCR amplification is held in a the Veriti® Thermal Cycler (Applied Biosystems, Foster City, CA, USA) in five steps: a “hot start” PCR at 94ºC for 4 min, in order to activate the Multi Taq2 DNA Polymerase and to prevent the formation of non-specific amplification products followed by 5 cycles, each one include (denaturation: 30s at 96ºC, hybridation: 120s at 63ºC and extension: 75s at 72ºC) and 25 other cycles with (denaturation: 30s at 94ºC, hybridation: 120s at 60ºC and extension: 75s at 72ºC), a final extension at 68ºC for 60 min and lastly 10ºC for de PCR product conservation.

Table 5: Chromosomal mapping of the Investigator Argus X-12 Kit(www.Qiagen.com).

Locus	Chromosomal mapping	Label dye
Amelogenin	Xp22.1–22.3, Yp11.2	6-FAM
DXS7132	Xq11.2	6-FAM
DXS7423	Xq28	BTY
DXS8378	Xp22.31	6-FAM
DXS10074	Xq12	BTG
DXS10079	Xq12	BTY
DXS10101	Xq26.2	BTG
DXS10103	Xq26.2	6-FAM
DXS10134	Xq28	6-FAM
DXS10135	Xp22.31	BTG
DXS10146	Xq28	BTY
DXS10148	Xp22.31	BTR
HPRTB	Xq26.2	BTR

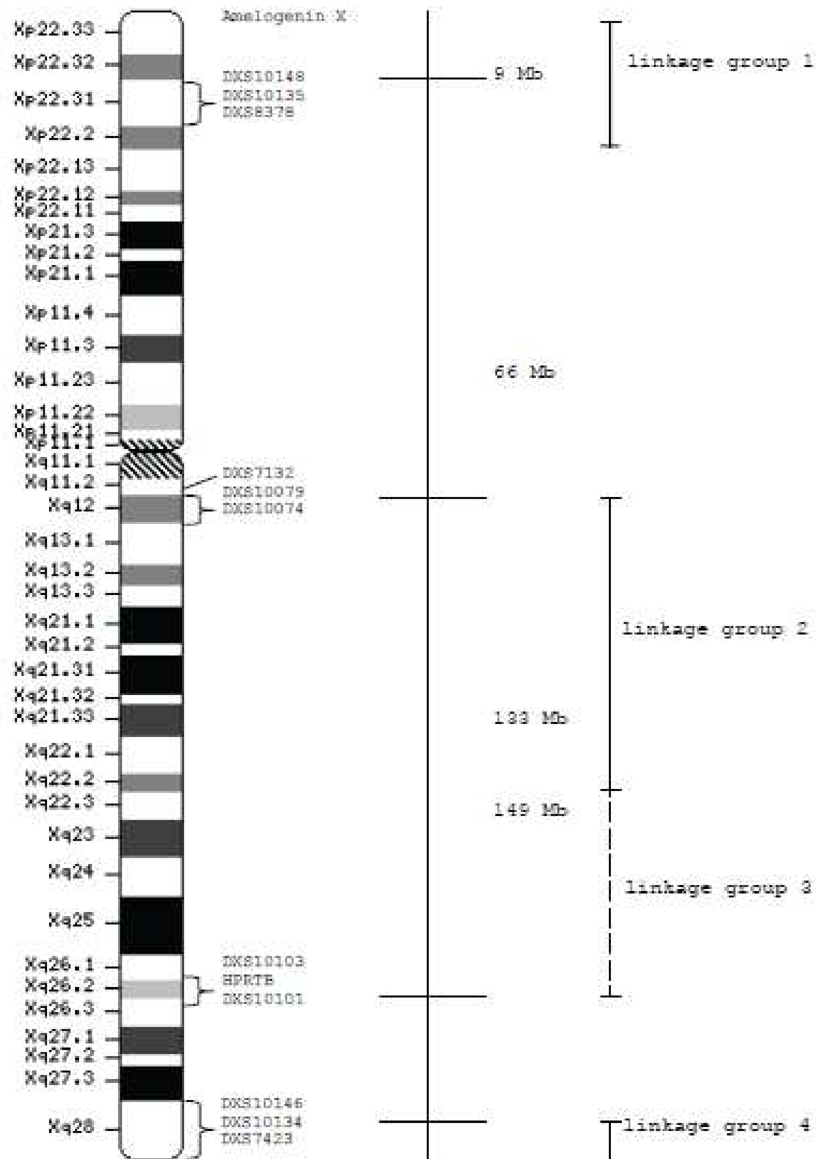


Figure 12. The ideogram of the X-chromosome. Distances from the p-telomere are shown in Mb. (http://www.ncbi.nlm.nih.gov/genome/guide/human as at 10/2009)

Table 6: Locus-specific information of the Investigator Argus X-12 Kit (www.qiagen.com).

Locus	GenBank® accession number	Repeat motif of the reference allele	Reference allele	Allele range
Amelogenin X	M55418			
Amelogenin Y	M55419			
DXS7132	G08111	$[TCTA]_{13}$	13	8-20
DXS7423	AC109994	$[TCCA]_{3}$ TCTGTCCT $[TCCA]_{12}$	15	8-19
DXS8378	G08098	$[CTAT]_{12}$	12	7-15
DXS10074	AL356358	$[AAGA]_{14}$	14	4-21
DXS10079	AL049564	$[AGAG]_{3}$ TGAAAGAG $[AGAA]_{17}$ AGAG $[AGAA]_{3}$	21	14-25
DXS10101	AC004383	$[AAAG]_{3}$ GAAAGAAG $[GAAA]_{3}$ A $[GAAA]_{4}$ AAGA $[AAAG]_{5}$ AAAAAGAA $[AAAG]_{13}$ AA	28.2	24-38
DXS10103	BV680555	$[TAGA]_{2}$ CTGA $[CAGA][TAGA]_{11}[CAGA]_{4}[TAGA]$	19	15-21
DXS10134	AL034384	$[GAAA]_{3}$ GAGA $[GAAA]_{4}$ AA [GAAA] GAGA $[GAAA]_{4}$ GAGA $[GACAGA]_{3}$ [GAAA] GTAA $[GAAA]_{3}$ AAA $[GAAA]_{4}$ AAA $[GAAA]_{15}$	35	28-46.1
DXS10135	AC003684	$[AAGA]_{3}$ GAAAG $[GAAA]_{20}$	23	13-39.2
DXS10146	AL034384	$[TTCC]_{3}$ T $[TTCC]_{3}$ TTTC CTCCCTTCC [TTCC] [TCCC] TTCTTCTTTC $[TTCC]_{2}$ TTTCTT $[CTTT]_{2}$ CTTC $[CTTT]_{10}$ T $[CTTT]_{2}$	26	24-46.2
DXS10148	AC003684	$[GGAA]_{4}[AAGA]_{12}[AAAG]_{4}$ N_{8} $[AAGG]_{2}$	22	13.3-38.1
HPRTB*	M26434	$[AGAT]_{12}$	12	6-19

4. Capillary Electrophoresis

Capillary electrophoresis can be used to detect and separate the different PCR product according to their size and fluorescence.

The principle of this method substitutes the other conventional electrophoresis methods. In a capillary electrophoresis, the separation result from the DNA fragments migration in an electric field inside a capillary, with a diameter of few microns full of polyacrylamide gel. The DNA molecules are separated according to their respective size: the smaller cross easily the polymer meshing and thus migrate faster through this network. The fluorescence detection and quantification (guise a colored picks) is performed via a LASER detector located at the end of the migratory route. So, the smallest fragments are chronologically detected firstly and if two (or more) fragments have the same size, they can be distinguished by the fluorescence color. The use of colors with different diffusivity included in the PCR multiplex, allow distinguishing the overlapping fragments (same size).

In our laboratory, the capillary electrophoresis is carried out in a genetic analyzer with a fluorescence Laser detection system (ABI Prism 3130, Applied Biosystems). Separations are performed in a capillary with a length of 47 cm (36 cm from the injection zone until the laser detector), 50 µm in diameter and full of polyacrylamide gel 4% or POP4 (*Performance Optimized Polymer 4,* Applied Biosystems[1]). Samples are electrokineticly injected in a capillary in 5 s. The electrophoretic migration of every sample last 30 min, at 15 kV and a temperature of 60ºC.

Before the electrophoretic separation of DNA fragments, samples were prepared by the addition of 12µl of Hi-Di™ formamide and 0.5 µl of DNA Size Standard BTO to 1 µl of each PCR product. As in all the classical electrophoresis, this internal standard size (Figure 13) previously calibrated with a known size of each standard fragment, must migrate in parallel with samples to correlate with all the parameters that can influence the migration (speed, detection's time with laser, gel purity, temperature…).

[1]POP4 polymer composition : 4% of poly-diméthylacrylamide, 8 M or urea, 5% of 2-pyrrolidinone and 100mM of N-tris (hydroxymethyl)-methyl-3-aminopropanesulfonique acid with pH 8.

In each 96 well reaction plate prepared for capillary electrophoresis, 1 µl of another size marker ("external") is added to 12 of Hi-Di™ formamide: the STR ladder or allelic ladder. It's an "allelic scale" formed by a synthetic strands of DNA with a predefined composition and size (number of repetitions). The ladder contains all the alleles than can be analyzed and identified by a multiplex PCR with an allelic scale for each used dye (Figure 14).As for a sample, the internal size standard is added to ladder, DNA fragments are separated during the electrophoresis and fluorescence is detected by laser system. The figure presents the allelic ladder for the four used colors in the PCR multiplex 12 X-STR (Applied Biosystems).

Samples are than centrifuged and in a thermo-cycler denaturized at 95ºC for 1 min and transferred immediately in ice for 3 min.

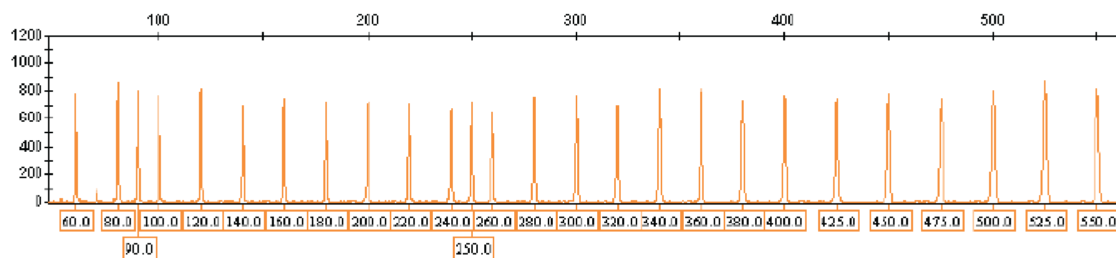


Figure 13:*Electropherogram of the DNA Size Standard 550 (BTO), fragments with lengths in bp (www.quiagen.com)*

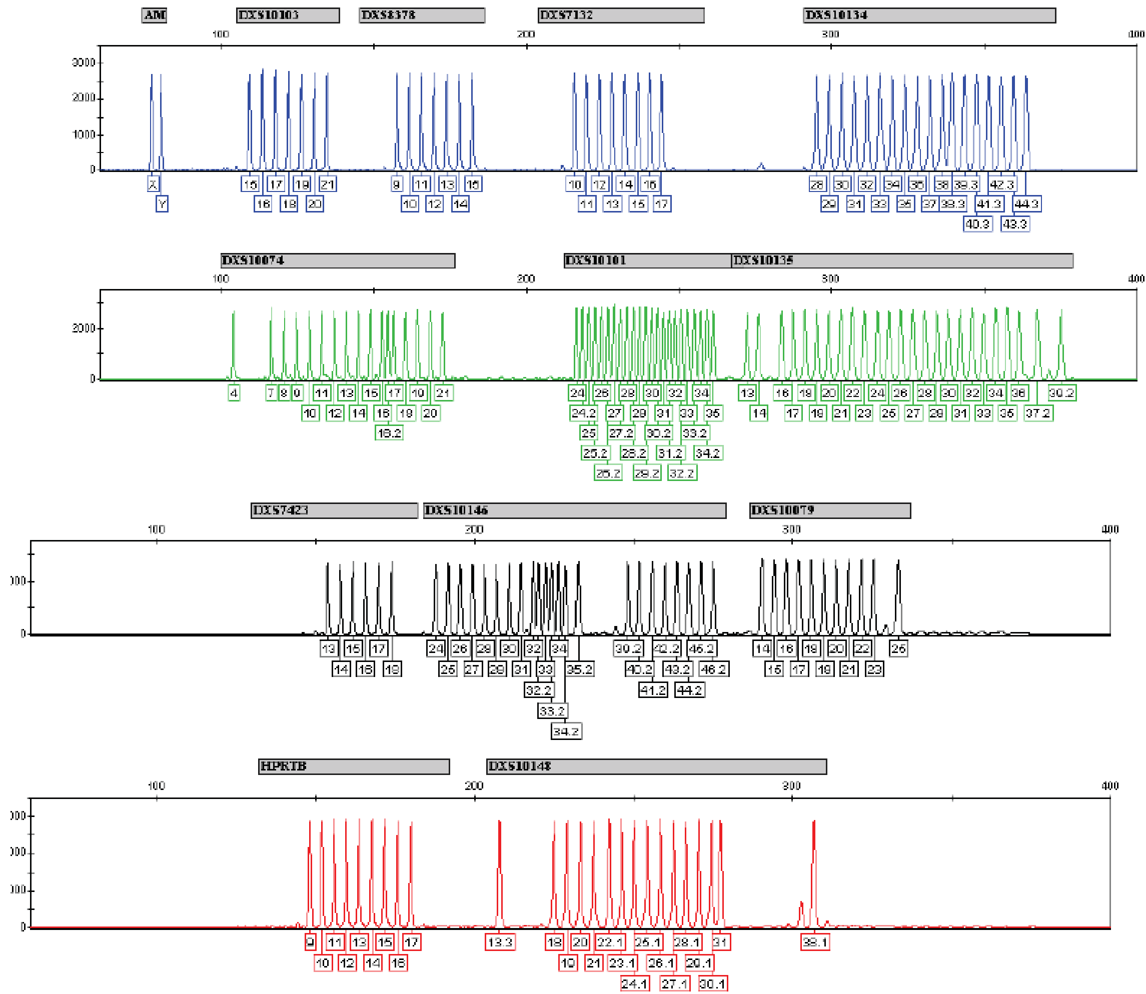


Figure 14: Electropherogram of the allelic ladder Argus X-12 analyzed on an ABI PRISM 310 Genetic Analyzer. Allele assignment was performed using the GeneMapper ID Software and the Investigator Argus X-12 Template File (www.quiagen.com).

5. Statistical Analysis of Data

Allele frequencies, Heterozigosities and evaluation of Hardy-Weinberg equilibrium have been calculated as explained in Alu insertion polymorphism analysis.

The following parameters provide information on the power of forensic analysis using the respective markers. Formulae are listed below. Two of the parameters listed, namely the polymorphism information content (PIC) (Bostein et al. 1980) and the expected heterozygosity (Het) (Nei et al. 1974), have been devised for more general purposes and are valid for both AS and ChrX markers.

$$1 - (\sum_{i=1}^{n} p_i^2) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2p_i^2 p_j^2$$

The mean exclusion chance (MECKRÜ) was introduced by (Krüger et al. 1968)for AS markers typed in trios involving mother, child and putative father (formula I). This parameter is not suitable for ChrX markers except for deficiency cases in which the paternal grandmother is investigated instead of the alleged father.

I $$\sum_i f_i^3 (1-f_i)^2 + \sum_i f_i (1-f_i)^3 + \sum_{i<j} f_i f_j (f_i + f_j)(1 - f_i - f_j)^2$$

(Kishida et al. 1997), devised a MECKIS for ChrX markers which covers trios including a daughter (formula II). If MECKRÜ is compared to MECKIS, the latter is considerably larger. This highlights the fact that in trios involving a daughter, ChrX markers are more efficient than AS markers.

II $$\sum_i f_i^3 (1-f_i) + \sum_i f_i (1-f_i)^2 + \sum_{i<j} f_i f_j (f_i + f_j)(1 - f_i - f_j)$$

Finally, Desmarais et al. (1998) introduced formulae for the mean exclusion chance of ChrX markers in trios involving daughters (formula III) and in father/daughter duos lacking maternal genotype information (formula IV). MEC(III) is equivalent to MEC(II) whilst MEC(IV) is also appropriate for maternity testing of mother/son duos. PD_M and PD_F are parameters suitable to evaluate the power of the markers for forensic identification purposes in males and females, respectively.

III $$1 - \sum_i f_i^2 + \sum_i f_i^4 - \left(\sum_{i<j} f_i^2\right)^2$$

IV $$1 - 2\sum_i f_i^2 + \sum_i f_i^3$$

V $$1 - 2\left(\sum_i f_i^2\right)^2 + \sum_i f_i^4$$

VI $$1 - \sum_i f_i^2$$

Genetic characterization of the Western Mediterranean population using the X-STR and X-Alu polymorphisms

In this chapter we focus on the genetic characterization of the Western Mediterranean population analyzing the X chromosome through two polymorphisms:

- ☑Four X-STR triplets in the linkage group 1-4 (DXS10148-DXS10135-DXS8378, DXS7132-DXS10079-DXS10074, DXS10103-HPRTB-DXS10101 and DXS10146-DXS10134-DXS7423) performed in a multiplex reaction.

- ☑Nine human specific X-Alu polymorphisms: Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4 and Ya5DP13 performed in monoplex reactions.

Our main aims are: 1) Set the genetic diversity of the Western Mediterranean population to one another and with other populations (European, American, Asiatic and Subsaharan), 2) to describe the genetic structure of the Western Mediterranean population and evaluate their kinship, 3) To verify the hypothesis about the gene flow boundary that could constitute the Mediterranean Sea and 4) to highlight the relevance and validate the use of the 12 X-STR and 9 X-Alu polymorphisms in the population genetic investigations and forensic genetics fields.

I. Genetic diversity of 12 X-chromosomal short tandem repeats in a Moroccan population sample

Blood samples were collected from 145 unrelated, healthy individuals (97 male and 48 female) from the three ethnic groups (Arab, Berber and Sahrawi) belonging to the whole area of Morocco.

No shared haplotypes were found when studying the 12 X-STRs simultaneously, consequently the haplotype diversity is 1. Allele frequencies observed for each marker are displayed in Table 7. Statistic parameters obtained for forensic evaluation are shown in the same table. DXS10146 and DYS10153 were the most polymorphic X-STR, with 23 alleles, and DXS8378 was the least

polymorphic one, with 5 alleles. Overall values obtained for the power of discrimination were high in females (PD_{femal}> 0.999999) and males (PD_{male}> 0.999999). Combined mean exclusion chance for trios and duos were 0.999999996 and 0.9999998, respectively. No deviations from Hardy-Weinberg equilibrium were observed ($p \geq 0.0042$; significant level after Bonferroni correction). The present study shows that there is a high genetic diversity in the Moroccan population for the 12 X-STRs. Forensic efficiency parameters also demonstrated that the DXS8378 is the least informative marker, in concordance with the study already done on 100 individual predominately born and living in Marrakech (Poetsch et al. 2010).

The haplotype frequencies of the 4 linkage groups in 97 Moroccan men are shown in table 9. The 4 linkage trios of DXS10148-DXS10135-DXS8378, DXS7132-DXS10079-DXS10074, DXS10103-HPRTB-DXS10101, and DXS10146-DXS10134-DXS7423 revealed 88, 64, 72 and 77 haplotypes, respectively. Of all showed haplotypes, 94% showed frequencies < 0.021, and some haplotypes displayed high frequency > 0.040.

The exact test of population differentiation was used to compare our data with the other African, European, American and Asian samples previously described in the literature (Bekada et al. 2010; Poetsch et al. 2010; Hedman et al. 2009; Gomez et al. 2007; Aler et al. 2007; Robino et al. 2006; Edelmann et al. 2001; Pereira et al. 2007; Gomes et al. 2007; Martins et al. 2010; Lim et al. 2009; Kang et al. 2006 and Chen et al. 2004). As shown in Table 8, our population seems to be nearest to Algerian, Spanish, and Portuguese populations than the other one like Brazil and Germany. These results are consistent with 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 (Lugan, 1999). However, The Moroccan population showed significant differences from the Korean, Chinese and Taiwanese populations in almost all loci implying genetic differences according to the geographic location.

Forensic efficiency parameters proved that the twelve X-STR used in this work are highly discriminating and therefore useful for forensic purposes. Overall values of the power of discrimination were high, supporting the potential of this multiplex in forensic identification tests. The high values obtained for combined MEC_{trio} and MEC_{duo} support the potential of this kit in a specific kinship analysis context when the offspring is female or when the father/daughter relationships are being investigated.

Table 7.
Allele frequencies and statistical parameters of 12 X-chromosomal short tandem repeat loci in 145 Moroccan samples (Arabic: 36 men and 18 women; Berber: 32 men and 16 women and Sahrawi: 29 men and 14 women)

DXS10103	ARB	BRB	SHR	MO
15	0.009	-	0.012	**0.007**
16	0.130	0.031	0.128	**0.096**
17	0.102	0.146	0.211	**0.149**
18	0.213	0.135	0.153	**0.170**
19	0.416	0.479	0.283	**0.398**
20	0.130	0.188	0.189	**0.166**
21	-	0.021	0.024	**0.014**
PIC	0.701	0.655	0.769	0.720
Het$_{obs}$ †	0.833	0.813	0.786	0.813
Het$_{exp}$	0.761	0.692	0.810	0.757
MEC$_{trio}$	0.701	0.655	0.769	0.720
MEC$_{duo}$	0.562	0.511	0.643	0.584
PD$_{female}$	0.895	0.868	0.930	0.906
PD$_{male}$	0.737	0.694	0.799	0.753
PE	0.488	0.419	0.597	0.516
HWE†	0.461	0.862	0.523	0.068

DXS8378	ARB	BRB	SHR	MO
9	0.028	-	-	**0.010**
10	0.241	0.240	0.270	**0.249**
11	0.425	0.427	0.341	**0.401**
12	0.250	0.312	0.341	**0.298**
13	0.056	0.021	0.047	**0.042**
PIC	0.642	0.594	0.631	0.626
Het$_{obs}$ †	0.833	0.750	0.857	0.813
Het$_{exp}$	0.716	0.674	0.707	0.693
MEC$_{trio}$	0.642	0.594	0.631	0.626
MEC$_{duo}$	0.496	0.447	0.485	0.480
PD$_{female}$	0.854	0.818	0.844	0.841
PD$_{male}$	0.695	0.662	0.693	0.687
PE	0.430	0.372	0.418	0.408
HWE†	0.713	0.877	0.817	0.565

DXS7132	ARB	BRB	SHR	MO
10	0.009	-	-	**0.003**
11	0.092	0.073	0.024	**0.066**
12	0.130	0.084	0.141	**0.118**
13	0.269	0.218	0.259	**0.249**
14	0.351	0.427	0.375	**0.384**
15	0.130	0.135	0.177	**0.145**
16	-	0.063	0.024	**0.028**
17	0.019	-	-	**0.007**
PIC	0.726	0.702	0.631	0.714
Het$_{obs}$ †	0.889	0.813	0.786	0.833
Het$_{exp}$	0.758	0.745	0.754	0.833
MEC$_{trio}$	0.726	0.702	0.631	0.714
MEC$_{duo}$	0.591	0.563	0.558	0.578
PD$_{female}$	0.908	0.896	0.890	0.902
PD$_{male}$	0.762	0.736	0.740	0.750
PE	0.530	0.485	0.493	0.510
HWE†	0.955	0.420	0.689	0.694

DXS10134	ARB	BRB	SHR	MO
29	0.019	0.010	0.024	**0.017**
30	0.056	-	0.071	**0.042**
31	0.019	0.031	-	**0.017**
33	0.074	0.083	0.023	**0.062**
33.2	-	0.010	-	**0.003**
34	0.092	0.042	0.093	**0.076**
35	0.222	0.125	0.260	**0.201**
36	0.204	0.156	0.189	**0.183**
37	0.046	0.250	0.186	**0.155**
38	0.065	0.166	0.047	**0.093**
38.3	0.009	-	-	**0.003**
39	0.037	-	0.024	**0.021**
39.3	-	0.010	-	**0.003**
40	0.009	-	-	**0.003**
40.3	0.055	0.073	0.059	**0.062**
41	0.009	-	0.024	**0.010**
41.1	0.019	-	-	**0.007**
41.3	0.046	0.021	-	**0.024**
42.3	0.019	0.021	-	**0.014**
PIC	0.867	0.838	0.822	0.864
Het$_{obs}$ †	1.000	0.875	0.786	0.896
Het$_{exp}$	0.889	0.862	0.841	0.896
MEC$_{trio}$	0.867	0.838	0.822	0.881
MEC$_{duo}$	0.776	0.735	0.713	0.772
PD$_{female}$	0.974	0.963	0.956	0.973
PD$_{male}$	0.877	0.854	0.840	0.876
PE	0.750	0.702	0.676	0.747
HWE†	0.991	0.876	0.670	0.964

DXS10074	ARB	BRB	SHR	MO
7	0.073	0.021	0.036	**0.045**
8	0.250	0.208	0.152	**0.208**
9	0.019	0.042	0.047	**0.035**
10	-	-	0.024	**0.007**
11	0.009	-	0.024	**0.010**
12	0.047	-	0.024	**0.024**
13	0.037	0.031	0.107	**0.056**
14	0.055	0.021	0.012	**0.031**
15	0.065	0.052	0.118	**0.076**
16	0.092	0.136	0.070	**0.100**
16.2	0.019	-	-	**0.007**
17	0.167	0.240	0.222	**0.208**
18	0.093	0.198	0.117	**0.135**
19	0.028	0.041	0.047	**0.038**
20	0.037	0.010	-	**0.017**
21	0.009	-	-	**0.003**
PIC	0.862	0.813	0.864	0.857
Het$_{obs}$ †	0.889	0.875	1.000	0.917
Het$_{exp}$	0.892	0.841	0.874	0.845
MEC$_{trio}$	0.862	0.813	0.864	0.857
MEC$_{duo}$	0.770	0.700	0.771	0.762
PD$_{female}$	0.973	0.952	0.973	0.970
PD$_{male}$	0.873	0.833	0.876	0.870
PE	0.741	0.662	0.747	0.734
HWE†	0.955	1.000	0.938	0.949

DXS10101	ARB	BRB	SHR	MO
24.2	-	0.010	-	**0.003**
25	0.009	-	0.011	**0.007**
25.2	-	-	0.024	**0.007**
26	-	0.010	-	**0.003**
27	0.009	-	0.024	**0.010**
27.2	0.046	0.010	0.012	**0.024**
28	0.028	0.052	0.071	**0.049**
28.2	0.092	0.104	0.106	**0.100**
29	0.046	0.042	0.058	**0.048**
29.2	0.140	0.084	0.082	**0.104**
30	0.083	0.042	0.106	**0.076**
30.2	0.102	0.125	0.059	**0.097**
31	0.111	0.104	0.106	**0.107**
31.2	0.148	0.125	0.047	**0.111**
32	0.093	0.125	0.176	**0.120**
32.1	-	0.021	-	**0.007**
32.2	0.028	0.063	0.011	**0.034**
33	0.065	0.073	0.083	**0.073**
33.2	-	0.010	-	**0.010**
34	-	-	0.024	**0.007**
PIC	0.893	0.901	0.898	0.907
Het$_{obs}$ †	0.889	1.000	1.000	0.958
Het$_{exp}$	0.911	0.934	0.924	0.917
MEC$_{trio}$	0.893	0.901	0.898	0.907
MEC$_{duo}$	0.814	0.826	0.822	0.836
PD$_{female}$	0.982	0.984	0.984	0.986
PD$_{male}$	0.902	0.908	0.905	0.914
PE	0.799	0.812	0.807	0.823
HWE†	0.782	1.000	1.000	0.825

DXS10135	ARB	BRB	SHR	MO
17	0.009	0.052	0.012	**0.024**
17.1	-	-	0.012	**0.003**
18	0.065	0.052	-	**0.041**
18.1	0.019	-	0.094	**0.035**
19	0.102	0.073	0.071	**0.083**
19.1	-	0.042	-	**0.014**
20	0.047	-	0.118	**0.052**
20.1	0.009	-	0.012	**0.007**
21	0.092	0.094	0.129	**0.104**
21.1	0.009	0.042	0.024	**0.024**
22	0.065	0.042	0.059	**0.055**
22.1	-	0.010	0.012	**0.007**
23	0.092	0.063	0.057	**0.073**
23.1	0.065	-	-	**0.024**

DXS7423	ARB	BRB	SHR	MO
13	0.147	0.146	0.036	**0.115**
14	0.380	0.302	0.329	**0.339**
15	0.213	0.375	0.318	**0.297**
16	0.195	0.167	0.317	**0.221**
17	0.065	-	-	**0.024**
18	-	0.010	-	**0.003**

DXS10146	ARB	BRB	SHR	MO
24	0.018	0.010	-	**0.010**
25	0.018	0.073	0.036	**0.041**
26	0.148	0.063	0.166	**0.124**
27	0.130	0.135	0.116	**0.128**
28	0.130	0.156	0.083	**0.124**
29	0.176	0.073	0.104	**0.121**
29.2	-	0.021	-	**0.007**
30	0.074	0.104	0.081	**0.086**
31	0.074	0.052	0.024	**0.052**
32	-	-	0.023	**0.007**
33	0.018	0.042	0.024	**0.028**
34	0.037	0.042	-	**0.028**
34.1	-	-	0.024	**0.007**
34.2	0.019	-	0.047	**0.021**

Table 7 (Continued)

DXS10135	**ARB**	**BRB**	**SHR**	**MO**
24	0.093	0.083	0.059	**0.079**
24.1	-	0.010	-	**0.003**
25	0.046	0.104	0.024	**0.059**
26	0.065	0.063	0.070	**0.066**
27	0.083	0.031	0.023	**0.049**
28	0.084	0.051	0.106	**0.080**
29	0.055	0.146	0.070	**0.090**
30	-	0.042	0.024	**0.021**
31	-	-	0.024	**0.007**
PIC	0.920	0.917	0.915	0.931
Het_{obs} †	1.000	0.938	0.857	0.938
Het_{exp}	0.941	0.938	0.928	0.941
MEC_{trio}	0.920	0.917	0.915	0.931
MEC_{duo}	0.857	0.853	0.849	0.875
PD_{female}	0.989	0.989	0.838	0.992
PD_{male}	0.925	0.923	0.689	0.935
PE	0.847	0.842	0.411	0.871
HWE†	1.000	0.357	0.520	0.178

DXS7423	**ARB**	**BRB**	**SHR**	**MO**
PIC	0.707	0.669	0.624	**0.687**
Het_{obs} †	0.778	0.813	0.643	**0.750**
Het_{exp}	0.724	0.730	0.707	**0.736**
MEC_{trio}	0.707	0.669	0.624	**0.687**
MEC_{duo}	0.568	0.526	0.477	**0.546**
PD_{female}	0.896	0.871	0.838	**0.882**
PD_{male}	0.746	0.719	0.689	**0.734**
PE	0.504	0.458	0.411	**0.483**
HWE†	0.795	0.024	0.748	**0.219**

DXS10146	**ARB**	**BRB**	**SHR**	**MO**
35.2	-	0.010	-	**0.003**
39.2	0.018	0.042	-	**0.020**
40.2	0.009	-	0.059	**0.021**
41.2	0.009	-	0.023	**0.010**
42.2	0.037	0.010	0.095	**0.045**
43.2	0.055	0.125	0.070	**0.083**
44.2	0.009	0.010	-	**0.007**
45.2	-	0.031	0.024	**0.017**
46.2	0.018	-	-	**0.007**
PIC	0.885	0.899	0.904	0.907
Het_{obs} †	0.944	0.625	0.929	0.833
Het_{exp}	0.905	0.925	0.901	0.915
MEC_{trio}	0.885	0.899	0.904	0.907
MEC_{duo}	0.803	0.824	0.831	0.837
PD_{female}	0.980	0.984	0.985	0.986
PD_{male}	0.894	0.907	0.910	0.914
PE	0.784	0.809	0.817	0.824
HWE†	0.595	0.001	0.901	0.003

DXS10079	**ARB**	**BRB**	**SHR**	**MO**
14	0.009	-	-	**0.003**
15	0.019	0.021	0.012	**0.017**
16	0.019	0.031	0.047	**0.031**
17	0.064	0.052	0.094	**0.069**
18	0.213	0.188	0.106	**0.173**
19	0.269	0.250	0.247	**0.256**
20	0.185	0.271	0.224	**0.225**
21	0.139	0.125	0.211	**0.155**
22	0.074	0.062	0.058	**0.065**
23	0.009	-	-	**0.003**
PIC	0.794	0.778	0.794	0.795
Het_{obs} †	0.833	0.875	0.786	0.833
Het_{exp}	0.835	0.824	0.841	0.825
MEC_{trio}	0.794	0.778	0.794	0.795
MEC_{duo}	0.675	0.655	0.675	0.677
PD_{female}	0.943	0.935	0.943	0.943
PD_{male}	0.818	0.805	0.819	0.820
PE	0.633	0.609	0.634	0.636
HWE†	0.960	0.535	0.944	0.895

HPRTP	**ARB**	**BRB**	**SHR**	**MO**
9	0.019	-	-	**0.007**
10	-	-	0.024	**0.007**
11	0.092	0.177	0.106	**0.125**
12	0.361	0.198	0.306	**0.290**
13	0.323	0.406	0.199	**0.315**
14	0.158	0.156	0.283	**0.194**
15	0.047	0.063	0.082	**0.062**
PIC	0.685	0.697	0.731	0.720
Het_{obs} †	0.667	0.813	0.714	0.729
Het_{exp}	0.750	0.744	0.782	0.764
MEC_{trio}	0.685	0.697	0.731	0.720
MEC_{duo}	0.544	0.557	0.597	0.584
PD_{female}	0.882	0.891	0.909	0.903
PD_{male}	0.729	0.736	0.768	0.759
PE	0.475	0.487	0.541	0.526
HWE†	0.778	0.928	0.319	0.744

DXS10148	**ARB**	**BRB**	**SHR**	**MO**
13.1	-	-	0.024	**0.007**
13.3	0.009	0.063	0.083	**0.048**
16	0.009	-	-	**0.003**
18	0.176	0.114	0.188	**0.159**
19	0.073	0.020	-	**0.035**
20	0.019	0.010	0.024	**0.017**
20.1	-	-	0.024	**0.007**
22	-	0.020	-	**0.007**
22.1	0.019	-	-	**0.007**
23	0.028	0.020	0.047	**0.031**
23.1	-	0.063	0.140	**0.062**
24	0.019	0.063	0.047	**0.042**
24.1	0.092	0.094	0.059	**0.083**
25.1	0.222	0.209	0.140	**0.194**
26.1	0.102	0.094	0.024	**0.076**
27.1	0.056	0.063	0.046	**0.056**
28.1	0.111	0.084	0.059	**0.087**
29.1	0.046	0.020	0.071	**0.045**
30.1	-	0.020	0.012	**0.010**
31	-	-	0.012	**0.003**
37.1	-	0.020	-	**0.007**
38.1	0.019	0.020	-	**0.014**
PIC	0.864	0.892	0.889	0.893
Het_{obs} †	0.778	0.750	0.714	0.750
Het_{exp}	0.889	0.912	0.903	0.904
MEC_{trio}	0.864	0.892	0.889	0.893
MEC_{duo}	0.771	0.814	0.809	0.815
PD_{female}	0.973	0.982	0.981	0.983
PD_{male}	0.876	0.900	0.897	0.901
PE	0.746	0.795	0.790	0.797
HWE†	0.846	0.032	0.011	0.021

*Abbreviations: ARB - Arabic; BRB – Berber; SHR – Sahraoui; MO – Pooled Moroccan Population ; PIC – polymorphic information content; Het_{obs}– observed heterozygosity;Het_{exp}– expected heterozygosity; MEC_{trio}– mean exclusion chance in trios involving daughter; MEC_{duo} – mean exclusion chance in father/daughter or mother/son duos; PD_{female} – power of discrimination in women; PD_{male} – power of discrimination in men; HWE – Hardy Weinberg equilibrium. †P value calculating using the female data.

The exact test for linkage disequilibrium was performed for all pairs of loci in all the sampled population. For a significance level of 0.00076 (after Bonferroni correction for 66 comparisons in our population) only a significant *p* value was obtained for DXS10101-DXS10146 pair of loci ($p \leq 0.0000$), which are quite distance 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 (Chakravarti, 1999). Thus, studies with larger sample size are necessary to obtain true LD information.

Since it has been recently suggested that the use of triplets in STR cluster instead of a single or duos of STRs might be a powerful tool in forensic identification and kinship testing, three markers of each linked group have been handled as haplotypes for genotyping in a sample of 97 Moroccan men. A total of 64 to 88 haplotypes for each linkage triplets were observed. The linkage group 2 (DXS7132-DXS10079-DXS10074) had the lowest haplotype diversity values. Haplotype frequencies for the Moroccan population can be considered as a reference to compare other haplotype population data using Investigator 12 X STR amplification kit as a definite set of closely linked X-chromosomal markers.

In conclusion, this work demonstrates the usefulness of those 4 X-STR triplets in both population genetic related studies and in human identification. A specific database for this multiplex should be used in forensic casework and kinship analysis in the Moroccan population. Therefore, the result of this work, especially the frequencies of the 4 closely linked haplotypes will contribute to establish this system as standard for the X-STR analysis in the Moroccan population and will serve as a data for comparison with other studies consisting in the analysis of those 4 X-STR triplets.

Table 8
Genetic distances between the Moroccan and other African, European, American and Asian populations populations (*F*st values)

		Algeria	Madagascar	Somali	Uganda	Spain	Italy	Germany	Portugal	Afro-American	Brasil	Korea	China	Taiwan
DXS10103	*F*st	0.00737	ND	ND	ND	ND	ND	ND	ND	ND	ND	**0.03406**	ND	ND
	*p*value	0.09009										**0**		
DXS8378	*F*st	0.00226	**0.01429**	0.00143	-0.01158	0.00297	- 0.00261	0.00311	0.00584	- 0.00521	ND	**0.11134**	**0.34529**	ND
	p value	0.20721	**0.00901**	0.29730	0.91892	0.57658	0.54054	0.20721	0.09910	0.80180		**0**	**0**	
DXS7132	*F*st	- 0.00299	**0.06423**	0.01616	**0.04135**	- 0.00080	- 0.00293	0.00139	0.00099	- 0.00113	0.00259	**0.02008**	0.00694	**0.39717**
	*p*value	0.70270	**0**	0.06306	**0**	0.50450	0.64865	0.25225	0.27027	0.40541	0.16216	**0**	0.12613	**0**
DXS10134	*F*st	**0.052380**	ND	0.00450	ND	ND	ND	ND	ND	ND	ND	**0.00659**	ND	ND
	p value			0.05405								**0.01802**		
DXS10074	*F*st	- 0.00368	ND	**0.05079**	ND	ND	ND	ND	ND	ND	ND	**0.00886**	ND	ND
	*p*value	0.97297		**0**								**0.01802**		
DXS10101	*F*st	**0.008770**	ND	**0.01825**	ND	ND	ND	ND	ND	ND	ND	**0.01549**	ND	ND
	*p*value			**0**								**0**		
DXS10135	*F*st	0.00163	ND	**0.00620**	ND	ND	ND	ND	ND	ND	ND	**0.00657**	ND	ND
	p value	0.14414		**0**								**0.01802**		
DXS7423	*F*st	0.00955	0.00392	**0.2269**	0.00077	**0.01347**	ND	**0.23708**	**0.01201**	**0.03480**	ND	**0.13538**	**0.23708**	ND
	*p*value	0.05405	0.15315	**0**	0.36937	**0.00901**		**0**	**0.01802**	**0**		**0**	**0**	
DXS10146	*F*st	- 0.00235	ND	ND	ND	ND	ND	ND	ND	ND	ND	**0.00735**	ND	ND
	*p*value	0.81982										**0.00901**		
DXS10079	*F*st	- 0.00226	ND	ND	ND	ND	ND	ND	ND	ND	ND	0.00487	ND	ND
	p value	0.73874										0.05405		
HPRTP	*F*st	0.00169	**0.08690**	0.00260	**0.07572**	**0.18726**	- 0.00289	- 0.00416	**0.03143**	**0.06001**	ND	**0.00757**	0.00494	ND
	p value	0.22523	**0**	0.22523	**0**	**0**	0.54955	0.96396	**0**	**0**		**0.02703**	0.19820	
DXS10148	*F*st	**0.00795**	ND	ND	ND	ND	ND	ND	ND	ND	ND	**0.01164**	ND	ND
	*p*value	**0**										**0**		

p values less than 0.05 were considered statistically significant (in bold)
ND not done

Table 9.
X-chromosome haplotypes of four linked STR trios in 36 Arabic, 32 Berber and 29 Sahrawi males living in Morocco

Nº	DXS10148	DXS10135	DXS8378	Arabic	Berber	Sahrawi	DXS7132	DXS10079	DXS10074	Arabic	Berber	Sahrawi
1	12	29	29.1	1			13	16	20	2		
2	10	21	27.1	1			11	8	19	2		
3	11	18	24.1	1			13	8	18	1	1	
4	11	23	19	1			15	15	23	1		
5	11	25	18	1			15	7	20	1		
6	11	23	27.1	1		1	12	21	22	1		
7	11	28	13.3	1	1		14	18	18	1		2
8	10	25	24.1	1			13	17	21	1		1
9	11	24	19	1			14	18	14	1		
10	12	19	18	1			14	18	19	1	1	
11	10	22	27.1	1			13	8	21	1		1
12	11	21	16	1			14	18	21	1	2	2
13	12	18	19	1			13	12	22	1		
14	10	23	18	1			14	17	19	2		
15	11	28	24	1			12	28	19	1		
16	10	23	27.1	1			13	18	20	1	2	
17	11	21	24	1			21	26	21	1		
18	13	20.1	28.1	1			13	7	18	1		
19	11	19	26.1	1			14	17	21	1		1
20	12	18	25.1	1			14	17	20	2		
21	12	20	19	1			14	8	17	1		
22	11	21	23	1	1		14	9	20	1		
23	11	17	24.1	1			14	20	21	1		
24	12	25	28.1	1			14	20	15	1		
25	10	21.1	27.1	1			13	8	19	1	1	2
26	11	19	25.1	1			13	16	19	1		
27	12	27	22.1	1			14	15	20	1	1	2
28	10	29	25.1	1			12	8	19	1		
29	12	28	18	1			13	9	21	1		
30	11	28	25.1	1			10	19	16	1		
31	10	19	23	2	1		14	15	15	1		
32	9	24	18	1			14	11	16	1		
33	10	26	22.1	1			13	19	16		1	
34	12	28	24.1	1			13	18	19		1	
35	13	23.1	27.1	1			12	17	21		1	
36	10	18	24.1		1		13	17	19		1	1
37	12	17	26.1		1		11	19	19		1	
38	11	22	19		1		14	14	18		2	
39	12	21.1	28.1		1		16	18	18		1	
40	10	18	25.1		1		11	17	19		2	
41	11	22	29.1		1		12	16	20		1	
42	12	19	25.1		2		13	13	21		1	
43	11	21	29.1		1		13	20	19		1	
44	10	26	26.1		1		14	8	18		1	
45	12	24	20		1		14	18	20		1	
46	11	18	28.1		1		13	17	20		1	1
47	11	25	25.1		1		13	17	22		1	1

Table 9 (Continued)

Nº	DXS10148	DXS10135	DXS8378	Arabic	Berber	Sahrawi	DXS7132	DXS10079	DXS10074	Arabic	Berber	Sahrawi
48	11	22	26.1		1		12	17	19		1	
49	11	27	25.1		1		12	17	17		1	1
50	11	29	25.1		1		14	15	18		1	
51	11	27	27.1		1		14	8	20		1	1
52	11	29	23.1		1		15	18	22		1	
53	10	24.1	19		1		16	17	20		1	
54	12	25	27.1		1		14	15	19		1	
55	10	24	18		1	1	15	16	19			3
56	10	26	22		1		12	17	22			2
57	10	21.1	25.1		1		15	13	21			1
58	10	22.1	26.1		1		13	16	18			1
59	11	21	26.1		1		14	8	19			1
60	12	24	23.1		1		14	8	21			1
61	11	22	13.3		1		14	7	17			1
62	12	27	22		1		14	14	21			1
63	11	24	25.1		1		15	17	15			1
64	10	17.1	23.1			1	12	8	20			1
65	11	26	25.1			1						
66	10	26	18			1						
67	12	26	18			1						
68	12	17	27.1			1						
69	11	28	23.1			1						
70	11	29	27.1			2						
71	11	23	30.1			1						
72	10	18.1	18			1						
73	12	21	23			2						
74	10	18.1	31			1						
75	12	21	23.1			1						
76	11	20.1	23.1			1						
77	12	22.1	18			1						
78	10	23	23.1			1						
79	10	21	25.1			1						
80	12	23	25.1			1						
81	12	22	24.1			1						
82	12	26	25.1			1						
83	12	27	25.1			1						
84	12	28	25.1			1						
85	10	21	13.3			1						
86	11	23	18			1						
87	11	28	28.1			1						
88	10	27	23.1			1						
89												
h*				0.9984 ± 0.0070	0.9980 ± 0.0085	0.9951 ± 0.0106				0.9937 ± 0.0082	0.9919 ± 0.0099	0.9803 ± 0.0140

Table 10
Haplotype frequencies (F) of 4 linked X-chromosomal short tandem repeat trios in 97 Moroccan men

Nº	DXS10103	HPRTB	DXS10101	Arabic	Berber	Sahrawi	DXS10146	DXS10134	DXS7423	Arabic	Berber	Sahrawi
1	18	28.2	12	1			41	16	40.2	1		
2	19	31.2	13	1			36	13	29	1		
3	19	31	9	1			36	14	30	1		
4	20	27	15	1			36	14	41.2	1		
5	19	30	13	1			33	14	34.2	1		
6	18	31	13	3			36	14	29	2		
7	18	30.2	12	1			39	16	27	1		
8	18	29	12	1		1	35	16	27	1		
9	19	31	12	1			41.3	17	29	1		
10	20	31	13	1			30	14	31	2		
11	19	29.2	13	1	1		38.3	15	44.2	1		
12	20	29.2	11	1		1	33	16	29	1		
13	16	29	14	1			36	16	27	1	1	1
14	19	30	12	1			38	14	42.2	1		
15	19	31.2	12	1			37	15	27	1		1
16	19	32.2	13	1	1		34	16	28	1		
17	19	31	13	1			38	14	26	1		
18	16	32	13	1		4	36	16	26	1		
19	16	33	14	1		1	36	16	30	1		
20	16	27.2	12	1			35	13	26	1		
21	18	30	11	1			40	15	28	1		
22	17	31	13	1	1		35	15	34.2	1		
23	16	28.2	12	2			41.3	14	26	2	1	
24	19	29	12	1			34	14	29	1		2
25	19	30	9	1			35	14	29	1		2
26	18	25	12	1		1	37	14	28	1	2	
27	19	28.2	11	1		1	34	15	29	1		
28	19	31.2	11	1			37	14	29	1		
29	18	28	14	1			38	16	31	1		
30	15	31.2	14	1			34	14	42.2	1		
31	17	30	14	1	1	1	36	14	26	1		
32	17	32	13	1			36	14	31	1		
33	20	29.2	13	1			39	14	28	1		
34	19	26	14		1		40.3	14	28		1	
35	20	30.2	11		1		38	15	31		1	
36	20	30.2	14		1		35	15	29		2	
37	19	28.2	12		2		35	15	26		1	
38	20	31.2	13		2		37	15	30		1	
39	19	29.2	12		1	2	36	13	25		1	
40	16	32	11		1		33.2	14	45.2		1	
41	19	32.2	14		1		38	15	43.2		2	

Table 10 (Continued)

Nº	DXS10103	HPRTB	DXS10101	Arabic	Berber	Sahrawi	DXS10146	DXS10134	DXS7423	Arabic	Berber	Sahrawi
42	19	33	13		2		35	14	44.2		1	
43	19	24.2	13		1		31	16	30		1	
44	16	28.2	11		1		35	16	45.2		1	
45	19	30.2	13		1	1	36	15	28		1	
46	18	29	14		1		39.3	15	27		1	
47	16	33.2	14		1		36	15	25		1	
48	19	31	14		2		37	16	27		1	
49	19	30.2	12		1		35	16	25		1	
50	19	28	12		1		35	15	28		1	
51	19	27.2	11		1		37	18	35.2		1	
52	18	30	13		1		37	15	29		2	1
53	19	33	11		1		41.3	14	29		1	
54	18	32	13		1		35	14	45.2		1	
55	20	28.2	13		1		37	15	31		1	
56	20	31	14		1		29	13	31		1	
57	19	29	13		1		38	14	24		1	
58	15	32	14			1	36	16	42.2		1	
59	17	31.2	12			1	34	16	29			1
60	16	30	14			1	37	14	27			2
61	18	31.2	12			1	34	16	30			1
62	16	28.2	13			1	35	14	43.2			2
63	17	31	15			1	36	16	41.2			2
64	10	27.2	13			1	35	15	25			1
65	17	31	14			2	33	16	27			1
66	18	28.2	13			1	37	15	32			2
67	17	29	14			1	38	15	26			1
68	19	29	11			1	37	16	28			1
69	17	32	14			1	40.3	16	40.2			1
70	18	30	23.1			1	35	13	30			1
71	19	32.2	12			1	37	15	26			1
72	17	32	12			1	37	14	30			1
73							33	16	30			1
74							38	14	27			1
75							36	16	29			1
76							37	16	30			1
h*				0.9937 ± 0.0090	0.9919 ± 0.0099	0.9803 ± 0.0172				0.9952 ± 0.0078	0.9919 ± 0.0099	0.9852 ± 0.0124

*Haplotype diversity ± standard deviation .

II. Gene Diversity in Balearic Islands Genetic Isolates and Valencia using Argus X-12 STR

Blood samples from 250 unrelated Spanish samples from Balearic Islands (Majorca: 39 men and 22 women; Minorca: 39 men and 25 women; Ibiza: 39 men and 20 women) and Valencia (39 men and 27 women) already typed for 9 X-Alus were analyzed.

Allele frequencies of 12 X-STR loci (DXS7132, DXS7423, DXS8378, HPRTB, DXS10074, DXS10101, DXS10134, and DXS10135) as well as the values for polymorphism information content, heterozygosity, mean exclusion chance, and power of discrimination are shown in Table 11.

DXS10135 was the most polymorphic X-STR, with 29 alleles, and DXS7423 was the least polymorphic one, with 6 alleles. In particular, the most frequently observed alleles of the DXS10103 and DXS7423 loci had 50.4% and 43.5% occurrences, respectively. Three new alleles not reported in the Investigator Argus 12 X STR were recovered: DXS10146-36.2, DXS8378-11.2 and DXS10148-20.3. Confirmation of these alleles by sequencing is planned. Combined Forensic efficiency parameters were calculated for all loci, as follows: combined power of discrimination for males 0.999999999 (Balearic Islands) and 0.999999998 (Valencia), combined power of discrimination for females is greater than 0.999999999 in both Balearic Islands and Valencia, and combined mean exclusion chance for trios 0.9999999 (Balearic Islands) and 0.9999999 (Valencia). No significant deviations from the Hardy-Weinberg equilibrium were observed in these 12 markers. No significant linkage disequilibrium was observed in Minorca. The system DXS7423 shows some significant linkage disequilibrium with other loci in Majorca and Ibiza.

Table 11
Allele frequencies and statistical parameters of 12 X-chromosomal short tandem repeat loci in 250 Spanish samples (Majorca: 39 men and 22 women; Minorca: 39 men and 25 women; Ibiza: 39 men and 20 women and Valencia: 39 men and 27 women)

DXS10103	MA	MI	IB	BI	VL
10	-	0.008	-	**0.003**	-
15	0.015	-	0.025	**0.013**	0.021
16	0.093	0.093	0.059	**0.082**	0.064
17	0.063	0.109	0.025	**0.066**	0.102
18	0.151	0.184	0.236	**0.190**	0.261
19	0.504	0.409	0.430	**0.448**	0.402
20	0.174	0.153	0.225	**0.184**	0.120
21	-	0.044	-	**0.015**	0.030
PIC	0.722	0.722	0.656	0.682	0.704
Het$_{obs}$ †	0.681	0.800	0.650	0.710	0.615
Het$_{exp}$	0.690	0.776	0.709	0.725	0.727
MEC$_{trio}$	0.722	0.722	0.656	0.682	0.704
MEC$_{duo}$	0.586	0.587	0.512	0.540	0.566
PD$_{female}$	0.908	0.908	0.865	0.884	0.896
PD$_{male}$	0.753	0.753	0.704	0.718	0.740
PE	0.515	0.515	0.435	0.457	0.493
HWE†	0.374	0.234	0.177	0.262	0.235

DXS8378	MA	MI	IB	BI	VL
9	0.047	-	-	**0.016**	0.010
10	0.261	0.375	0.394	**0.343**	0.249
11	0.354	0.232	0.235	**0.274**	0.401
11.2	-	-	0.008	**0.003**	-
12	0.330	0.336	0.286	**0.317**	0.298
13	0.008	0.044	0.059	**0.037**	0.042
14	-	-	0.017	**0.006**	-
20	-	0.013	-	**0.004**	-
PIC	0.630	0.630	0.650	0.648	0.626
Het$_{obs}$ †	0.636	0.640	0.650	0.642	0.615
Het$_{exp}$	0.717	0.677	0.730	0.708	0.726
MEC$_{trio}$	0.630	0.630	0.650	0.648	0.626
MEC$_{duo}$	0.484	0.484	0.506	0.503	0.480
PD$_{female}$	0.844	0.844	0.865	0.856	0.841
PD$_{male}$	0.691	0.691	0.704	0.705	0.687
PE	0.414	0.414	0.435	0.436	0.408
HWE†	0.667	0.302	0.750	0.573	0.748

DXS7132	MA	MI	IB	BI	VL
11	0.015	0.008	-	**0.008**	0.026
12	0.132	0.126	0.251	**0.170**	0.090
13	0.252	0.244	0.269	**0.255**	0.291
14	0.366	0.341	0.336	**0.348**	0.393
15	0.218	0.241	0.143	**0.201**	0.184
16	0.017	0.039	-	**0.019**	0.008
17	-	-	-	**-**	0.008
PIC	0.706	0.702	0.681	0.701	0.671
Het$_{obs}$ †	0.591	0.813	0.700	0.701	0.846
Het$_{exp}$	0.736	0.745	0.743	0.741	0.735
MEC$_{trio}$	0.706	0.702	0.681	0.701	0.671
MEC$_{duo}$	0.567	0.563	0.539	0.561	0.529
PD$_{female}$	0.894	0.896	0.878	0.891	0.873
PD$_{male}$	0.749	0.736	0.740	0.745	0.718
PE	0.507	0.485	0.478	0.500	0.456
HWE†	0.525	0.420	0.779	0.575	0.355

DXS10134	MA	MI	IB	BI	VL
32	0.015	0.026	0.008	**0.016**	0.021
33	0.032	0.013	0.017	**0.021**	0.034
34	0.084	0.080	0.068	**0.077**	0.094
35	0.235	0.260	0.209	**0.235**	0.179
35.2	0.015	-	-	**0.005**	-
35.3	-	-	-	**-**	0.021
36	0.136	0.212	0.287	**0.212**	0.278
36.2	-	0.022	-	**0.007**	0.013
37	0.188	0.160	0.135	**0.161**	0.150
37.2	0.015	0.027	-	**0.014**	-
37.3	-	0.008	-	**0.003**	0.013
38	0.043	0.062	0.117	**0.074**	0.103
38.2	0.015	-	-	**0.005**	-
38.3	0.015	-	0.017	**0.011**	-
39	0.015	0.013	0.008	**0.012**	0.013
39.3	0.039	0.013	0.008	**0.020**	0.038
40	0.024	0.008	-	**0.011**	-
40.3	0.041	0.008	0.058	**0.036**	-
41.2	-	0.013	-	**0.004**	-
41.3	0.056	0.066	-	**0.041**	0.026
42.2	-	-	-	**-**	0.008
42.3	0.024	0.008	0.017	**0.016**	-
43.2	-	-	-	**-**	0.008
43.3	0.008	-	0.050	**0.019**	-
PIC	0.830	0.828	0.811	0.843	0.827
Het$_{obs}$ †	0.682	0.920	0.950	0.851	0.846
Het$_{exp}$	0.886	0.853	0.825	0.855	0.858
MEC$_{trio}$	0.830	0.828	0.811	0.843	0.827
MEC$_{duo}$	0.724	0.721	0.698	0.742	0.721
PD$_{female}$	0.960	0.959	0.952	0.965	0.959
PD$_{male}$	0.846	0.844	0.830	0.857	0.844
PE	0.670	0.684	0.657	0.709	0.683
HWE†	0.010	0.484	0.310	0.268	0.550

DXS10074	MA	MI	IB	BI	VL
7	0.015	0.044	0.042	**0.034**	0.073
8	0.138	0.174	0.235	**0.182**	0.162
9	0.030	0.013	0.025	**0.023**	0.026
13	0.008	-	-	**0.003**	0.008
14	0.008	0.022	0.042	**0.024**	0.013
15	0.112	0.110	0.033	**0.085**	0.162
16	0.194	0.136	0.093	**0.141**	0.150
17	0.207	0.233	0.202	**0.214**	0.145
17.2	-	-	-	**-**	0.017
18	0.194	0.202	0.235	**0.210**	0.201
19	0.093	0.066	0.093	**0.084**	0.043
PIC	0.817	0.816	0.804	0.821	0.838
Het$_{obs}$ †	0.955	0.840	0.800	0.865	0.846
Het$_{exp}$	0.847	0.845	0.843	0.845	0.862
MEC$_{trio}$	0.817	0.816	0.804	0.821	0.838
MEC$_{duo}$	0.705	0.704	0.688	0.710	0.734
PD$_{female}$	0.953	0.953	0.948	0.955	0.962
PD$_{male}$	0.837	0.837	0.826	0.841	0.855
PE	0.670	0.670	0.649	0.676	0.705
HWE†	0.022	0.977	0.742	0.580	0.698

Table 11 (continued)

DXS10101	MA	MI	IB	BI	VL
25.2	-	-	-	**-**	0.008
26.2	-	-	-	**-**	0.026
27	0.015	0.008	-	**0.008**	-
27.2	0.017	0.062	0.059	**0.046**	0.085
28	0.071	0.067	0.076	**0.071**	0.038
28.2	0.127	0.110	0.126	**0.121**	0.030
29	0.039	0.035	-	**0.025**	0.034
29.1	-	0.008	-	**0.003**	-
29.2	0.213	0.192	0.110	**0.172**	0.162
30	0.017	0.022	0.050	**0.030**	0.068
30.2	0.157	0.139	0.135	**0.144**	0.158
31	0.080	0.079	0.167	**0.109**	0.073
31.1	-	0.008	-	**0.003**	-
31.2	0.125	0.146	0.117	**0.129**	0.167
32	0.039	0.052	0.025	**0.039**	0.056
32.2	0.024	0.022	0.059	**0.035**	0.013
33	0.030	0.035	0.059	**0.041**	0.073
33.2	-	-	-	**-**	0.008
34	0.030	-	0.017	**0.016**	-
35	0.015	0.013	-	**0.009**	-
PIC	0.891	0.880	0.883	0.884	0.881
Het_{obs} †	0.864	0.960	0.850	0.891	0.808
Het_{exp}	0.896	0.885	0.901	0.894	0.891
MEC_{trio}	0.891	0.880	0.883	0.884	0.881
MEC_{duo}	0.812	0.794	0.798	0.800	0.796
PD_{female}	0.982	0.978	0.779	0.979	0.978
PD_{male}	0.899	0.889	0.893	0.893	0.891
PE	0.794	0.774	0.780	0.781	0.777
HWE†	0.713	0.724	0.134	0.524	0.089

DXS10135	MA	MI	IB	BI	VL
16	-	0.022	-	**0.007**	0.008
16.1	-	0.013	-	**0.004**	-
17	-	0.075	-	**0.025**	0.034
17.1	-	0.008	-	**0.003**	-
18	0.071	0.013	0.043	**0.042**	0.051
18.1	-	-	0.008	**0.003**	0.008
19	0.080	0.088	0.083	**0.084**	0.098
19.1	0.041	0.017	-	**0.019**	-
20	0.063	0.008	0.050	**0.040**	0.068
20.1	0.008	-	-	**0.003**	0.013
21	0.117	0.084	0.100	**0.100**	0.120
21.1	-	0.008	-	**0.003**	0.034
22	0.110	0.057	0.059	**0.075**	0.128
22.1	0.030	0.022	0.008	**0.020**	-
23	-	0.149	0.034	**0.061**	0.056
23.1	-	0.017	-	**0.006**	-
24	0.063	0.088	0.042	**0.064**	0.021
25	0.110	0.096	0.168	**0.125**	0.077
26	0.015	0.013	0.059	**0.029**	0.056
27	0.080	0.030	0.092	**0.067**	0.068
28	0.039	0.079	0.134	**0.084**	0.021
28.1	0.008	-	-	**0.003**	-
29	0.054	0.049	0.025	**0.043**	0.056
29.2	-	0.013	-	**0.004**	-
30	0.063	0.013	0.034	**0.037**	0.021
31	0.024	-	0.034	**0.019**	0.008
32	0.024	0.035	0.017	**0.025**	0.026
33	-	-	-	**-**	0.026
33.1	-	-	0.008	**0.003**	-
PIC	0.919	0.919	0.904	0.928	0.922
Het_{obs} †	0.864	0.920	0.950	0.911	0.923
Het_{exp}	0.943	0.941	0.917	0.934	0.925
MEC_{trio}	0.919	0.919	0.904	0.928	0.922
MEC_{duo}	0.855	0.855	0.832	0.870	0.861
PD_{female}	0.989	0.989	0.985	0.991	0.990
PD_{male}	0.924	0.924	0.911	0.932	0.927
PE	0.844	0.844	0.818	0.862	0.850
HWE†	0.125	0.564	0.446	0.378	0.874

DXS7423	MA	MI	IB	BI	VL
11	-	-	0.016	**0.005**	-
13	0.084	0.084	0.067	**0.078**	0.094
14	0.315	0.330	0.336	**0.327**	0.376
15	0.435	0.430	0.395	**0.420**	0.359
16	0.120	0.126	0.126	**0.124**	0.162
17	0.045	0.030	0.059	**0.045**	0.008
PIC	0.628	0.628	0.658	0.642	0.638
Het_{obs} †	0.727	0.440	0.850	0.672	0.654
Het_{exp}	0.687	0.689	0.719	0.698	0.691
MEC_{trio}	0.628	0.628	0.658	0.642	0.638
MEC_{duo}	0.483	0.483	0.516	0.498	0.494
PD_{female}	0.845	0.845	0.865	0.855	0.851
PD_{male}	0.682	0.682	0.707	0.693	0.695
PE	0.402	0.402	0.439	0.418	0.420
HWE†	0.647	0.012	0.114	0.258	0.171

DXS10146	MA	MI	IB	BI	VL
24	0.015	0.022	0.050	**0.029**	0.021
25	0.039	0.049	0.092	**0.060**	0.073
26	0.082	0.070	0.126	**0.093**	0.120
27	0.162	0.175	0.101	**0.146**	0.154
28	0.228	0.171	0.144	**0.181**	0.192
29	0.134	0.214	0.168	**0.172**	0.141
30	0.078	0.066	0.142	**0.095**	0.094
31	0.073	0.027	0.017	**0.039**	0.008
32	-	0.008	-	**0.003**	0.013
36.2	-	0.022	-	**0.007**	-
39.2	0.024	0.013	0.050	**0.029**	0.030
40.2	0.039	0.040	-	**0.026**	0.030
41.2	0.032	0.044	0.026	**0.034**	0.026
42.2	0.008	0.013	0.008	**0.010**	0.021
43.2	0.008	0.040	0.008	**0.019**	0.013
44.2	0.032	-	-	**0.011**	0.026
45.2	0.045	0.008	0.059	**0.037**	0.013
46.2	-	0.017	-	**0.006**	-
47.2	-	-	0.008	**0.003**	0.025
PIC	0.866	0.430	0.876	0.878	0.875
Het_{obs} †	0.864	1.000	0.950	0.938	1.000
Het_{exp}	0.891	0.881	0.894	0.889	0.894
MEC_{trio}	0.866	0.430	0.876	0.878	0.875
MEC_{duo}	0.775	0.771	0.788	0.792	0.788
PD_{female}	0.974	0.973	0.976	0.977	0.977
PD_{male}	0.877	0.875	0.887	0.888	0.886
PE	0.749	0.745	0.768	0.771	0.766
HWE†	0.065	0.978	0.424	0.489	0.697

Table 11 (continued)

DXS10079	MA	MI	IB	BI	VL
11	-	-	0.017	**0.006**	-
12	-	-	0.017	**0.006**	-
14	-	-	0.017	**0.006**	-
15	0.024	0.013	0.008	**0.015**	-
16	0.047	0.040	0.026	**0.038**	0.013
17	0.071	0.080	0.076	**0.076**	0.107
18	0.196	0.234	0.236	**0.222**	0.162
19	0.276	0.279	0.201	**0.252**	0.171
20	0.228	0.235	0.176	**0.213**	0.303
21	0.086	0.084	0.218	**0.129**	0.205
22	0.054	0.013	0.008	**0.025**	0.030
23	0.017	-	-	**0.006**	0.008
24	-	0.008	-	**0.003**	-
30	-	0.013	-	**0.004**	-
PIC	0.791	0.767	0.793	0.792	0.769
Het_{obs} †	0.864	0.840	1.000	0.901	0.731
Het_{exp}	0.834	0.815	0.832	0.827	0.815
MEC_{trio}	0.791	0.767	0.793	0.792	0.769
MEC_{duo}	0.672	0.642	0.673	0.673	0.643
PD_{female}	0.942	0.929	0.942	0.942	0.930
PD_{male}	0.815	0.797	0.818	0.817	0.798
PE	0.627	0.593	0.633	0.631	0.596
HWE†	0.780	0.544	0.359	0.561	0.810

HPRTP	MA	MI	IB	BI	VL
9	-	0.008	-	**0.003**	-
10	0.008	0.013	-	**0.007**	0.021
11	0.166	0.092	0.117	**0.125**	0.158
12	0.332	0.320	0.320	**0.324**	0.363
13	0.304	0.395	0.387	**0.362**	0.256
14	0.149	0.118	0.126	**0.131**	0.167
15	0.041	0.044	0.050	**0.045**	0.034
16	-	0.008	-	**0.003**	
PIC	0.672	0.697	0.668	0.684	0.708
Het_{obs} †	0.840	0.813	0.750	0.801	0.846
Het_{exp}	0.738	0.744	0.724	0.735	0.744
MEC_{trio}	0.672	0.697	0.668	0.684	0.708
MEC_{duo}	0.531	0.557	0.526	0.544	0.570
PD_{female}	0.875	0.891	0.872	0.882	0.896
PD_{male}	0.717	0.736	0.716	0.749	0.748
PE	0.455	0.487	0.453	0.475	0.507
HWE†	0.689	0.928	0.712	0.776	0.705

DXS10148	MA	MI	IB	BI	VL
18	0.096	0.140	0.092	**0.109**	0.111
19	0.081	0.044	0.017	**0.047**	0.008
20	-	0.013	-	**0.004**	-
20.3	-	-	0.017	**0.006**	-
21	-	-	0.008	**0.003**	-
22	0.016	0.008	-	**0.008**	0.008
22.1	0.017	0.034	0.034	**0.028**	0.026
23	0.034	0.017	0.034	**0.028**	0.051
23.1	0.082	0.030	0.050	**0.054**	0.051
24	-	0.035	0.008	**0.014**	0.021
24.1	0.129	0.189	0.202	**0.173**	0.235
25	0.017	-	-	**0.006**	-
25.1	0.181	0.123	0.244	**0.183**	0.230
26.1	0.161	0.163	0.109	**0.144**	0.103
27.1	0.032	0.088	0.134	**0.085**	0.060
27.2	-	-	0.008	**0.003**	-
28.1	0.129	0.075	0.034	**0.079**	0.056
29.1	0.008	0.027	-	**0.012**	0.026
30.1	-	-	0.008	**0.003**	0.013
31.1	0.016	0.013	-	**0.010**	-
PIC	0.871	0.872	0.839	0.872	0.840
Het_{obs} †	0.905	0.880	0.900	0.895	0.885
Het_{exp}	0.893	0.899	0.871	0.888	0.854
MEC_{trio}	0.871	0.872	0.839	0.872	0.840
MEC_{duo}	0.781	0.783	0.737	0.783	0.738
PD_{female}	0.975	0.975	0.963	0.975	0.964
PD_{male}	0.882	0.883	0.855	0.883	0.855
PE	0.759	0.761	0.704	0.761	0.705
HWE†	0.887	0.814	0.647	0.783	0.327

*Abbreviations: MA - Mallorca; MI – Menorca; IB– Ibiza; BI – Pooled Balearec Island Population; VL- Valencia ; PIC – polymorphic information content; Het_{obs}– observed heterozygosity;Het_{exp}– expected heterozygosity; MEC_{trio}– mean exclusion chance in trios involving daughter; MEC_{duo} – mean exclusion chance in father/daughter or mother/son duos; PD_{female} – power of discrimination in women; PD_{male} – power of discrimination in men; HWE – Hardy Weinberg equilibrium. †P value calculating using the female data.

Table 12.
Genetic distances between Balearic Islands and other African, European, American and Asian populations populations (*F*st values)

		Morocco	Algeria	Valencia	Italy	Germany	Portugal	Afro-American	Spain	Danemark	Somaly	Hungary	Ghana	Korea
DXS10103	*F*stpval ue	0.00108 0.29730	0.00173 0.31532	-0.01054 0.99099	ND	ND	ND	ND	ND	-0.00272 0.74775	-0.00700 0.84685	ND	ND	ND
DXS8378	*F*stp value	**0.01231** **0.01802**	-0.00015 0.39640	0.00813 0.16216	0.00841 0.12613	0.00120 0.26126	0.00174 0.27027	**0.00000** **0.00901**	0.00259 0.16216	0.00104 0.29730	-0.00332 0.72072	0.00557 0.05405	**0.02436** **0.00901**	**0.07222** **0.00000**
DXS7132	*F*stp value	0.00157 0.25225	0.00331 0.21622	- 0.00413 0.59459	- 0.00339 0.72072	0.00150 0.19820	0.00005 0.44144	0.00341 0.18919	- 0.00550 0.96396	0.00018 0.40541	**0.00996** **0.03604**	ND	ND	ND
DXS10134	*F*stp value	-0.00149 0.63063	**0.05625** **0.00000**	-0.00255 0.57658	ND	ND	ND	ND	ND	-0.00007 0.51351	0.01184 0.25225	ND	**0.01195** **0.01802**	**0.00685** **0.01804**
DXS10074	*F*stp value	0.001510.2 7928	0.00046 0.40541	-0.00181 0.47748	ND	ND	ND	ND	ND	**0.00982** **0.01802**	ND	**0.00486** **0.03604**	**0.08404** **0.00000**	**0.02797** **0.00000**
DXS10101	*F*stp value	0.00459 0.06306	**0.01128** **0.00000**	-0.00026 0.40541	ND	ND	ND	ND	ND	-0.00130 0.63063	**0.03935** **0.00000**	**0.00531** **0.00901**	**0.04871** **0.00000**	**0.03060** **0.00000**
DXS10135	*F*stp value	0.000300.3 3333	0.001170.2 4324	- 0.001100.6 0360	ND	ND	ND	ND	ND	0.00020 0.43243	**0.00632** **0.00000**	-0.00030 0.53153	**0.01022** **0.00000**	**0.00494** **0.00000**
DXS7423	*F*stp value	0.012330.0 3604	0.00017 0.31532	-0.00506 0.64865	ND	**0.25960** **0.00000**	-0.00282 0.80180	**0.04747** **0.00000**	- 0.00412 0.72072	-0.00003 0.32432	0.00839 0.05405	0.00105 0.29730	**0.03577** **0.00000**	**0.05808** **0.00000**
DXS10146	*F*stp value	0.00304 0.13514	0.003410.0 5405	- 0.007600.9 8198	ND	ND	ND	ND	ND	-0.00193 0.45661	**0.02641** **0.00000**	ND	ND	ND
DXS10079	*F*stp value	- 0.003230.8 1081	- 0.000680.5 4955	0.004750.1 9820	ND	ND	ND	ND	ND	0.00257 0.26126	0.00360 0.18018	ND	ND	ND
HPRTP	*F*stp value	- 0.001050.4 5045	**0.00329** **0.13514**	**0.00076** **0.33333**	0.00317 0.19820	0.00031 0.32432	ND	**0.08219** **0.00000**	**0.22399** **0.00000**	0.00669 0.07207	0.01449 0.00000	-0.00212 0.82883	0.00055 0.36937	**0.00556** **0.03604**
DXS10148	*F*stp value	**0.00795** **0.00000**	ND	ND	ND	ND	ND	ND	ND	-0.00112 0.63063	**0.02795** **0.00000**	ND	ND	ND

p values less than 0.05 were considered statistically significant (in bold)
ND not done

Significant values of LD of other pairs of loci were also observed in Valencia (Kruglyak et al. 1999) concluded that linkage disequilibrium would not extend much beyond about 3 kb, but also noted that this could be significantly greater in populations that have undergone a severe bottleneck and admixture. The comparison of our data with (Bentayebi et al. 2012) on the Moroccan population and (Bekada et al. 2010) on the Algerian population, where the LD between loci was reached, agreed the hypothesis that a genetic bottleneck has affected the European but not African populations. In the other hand, the limited extent of LD found in the African populations, compared to Europeans (Tomas et al. 2012), and reflects the longer evolutionary history and relatively constant population size that shaped the genetic background of African populations. A more exhaustive sampling of these populations is suitable to confirm this observation.

Locus by locus population pairwise genetic distances (F_{ST}) between the Balearic population and relevant population samples (Morocco, Algeria, Italy, Germany, Denmark, Hungary, Somalia, Ghana, Japan and Korea) are shown in Table 12. Significant genetic distances, mainly from the Asiatic and sub-Saharan populations are distributed rather heterogeneous between different STRs. The North African (mainly Moroccan) and European populations seem to be close genetically. It could be explained by the history of the Balearic Islands that were settled by different people, which has contributed to the genetic pool of the actual population. In particular, it is important to emphasize the contribution of the Romans in the 3rd century BC and the Catalans in the early 13thcentury, although Ibiza differs from the other two islands, especially in the origins of founding settlements. Whereas Majorca and Minorca were inhabited since 5,500 years ago, there is no archaeological evidence for permanent human habitation on Ibiza beforethe arrival of the Carthaginians in 654 BC. The genetic influence of North African (especially Moroccan) on the Balearic Islands populations could be evidenced by the long settlement of Arabic in the islands that lasted for around 3 centuries.

Pairwise F_{ST} values calculated between 14 populations for eight of the 12 X-STRs were averaged and represented in an MDS diagram (Figure 15). As shown in this Figure, the populations tended to group according to their continent of origin. The clear differentiation of the Mediterranean and North European populations could be explained by a probable sex-biased contribution of North European genes (males) to the Mediterranean population (females)that has probably been occurred.

Table 13:
Haplotype diversity of X-STR clusters in Ivorians (Tie et al. 2010), Algerians and Italians (Inturri et al. 2010 and Robino et al. 2006) and Balearic Islands

	Balearic Islands	Morocco	Algeria	Italy	IvoryCoast
DXS10148-DXS10135-DXS8378,	0.858	0.997	0.998	0.998	0.998
DXS7132-DXS10074-DXS10074	0.8025	0.989	0.994	0.984	0.994
HPRTB-DXS10101-DXS10103	0.7616	0.989	0.989	0.985	0.986
DXS10146-DXS10134-DXS7423	0.7949	0.991	0.999	1	1

Making recently use of closely linked X-STR loci, to detect kinship between alleged relatives in large and incomplete pedigrees kinship testing instead of a single STR, we have expanded the investigation to the haplotype analysis of 4 linked X-STR duos in 300 men from our populations. Haplotype frequencies were counted for each linkage group. Haplotyping results are summarized in Table 14.

A total of 38 to 59 haplotypes for each linkage trio and 76.2 to 85.8% of haplotype diversities were observed, which are lower values than those found in Germans (96.0 – 97.6%) and in other populations. In Table 13, haplotype diversity calculated for the six clusters of linked loci in the Balearic Islands population sample is compared to that previously observed in Moroccan, Algerians Italians and Ivorians. X-chromosomal variability in the sample from Balearic Islands and Valencia is lower than that found in the Mediterranean basin and sub-Saharan populations, as expected according to the geographical isolation, small population size, high endogamy, and inbreeding, an increased genetic differentiation among subpopulations is depicted as a consequence of founder effect and genetic drift (Angius et al. 2001 and Frammene et al. 2003).

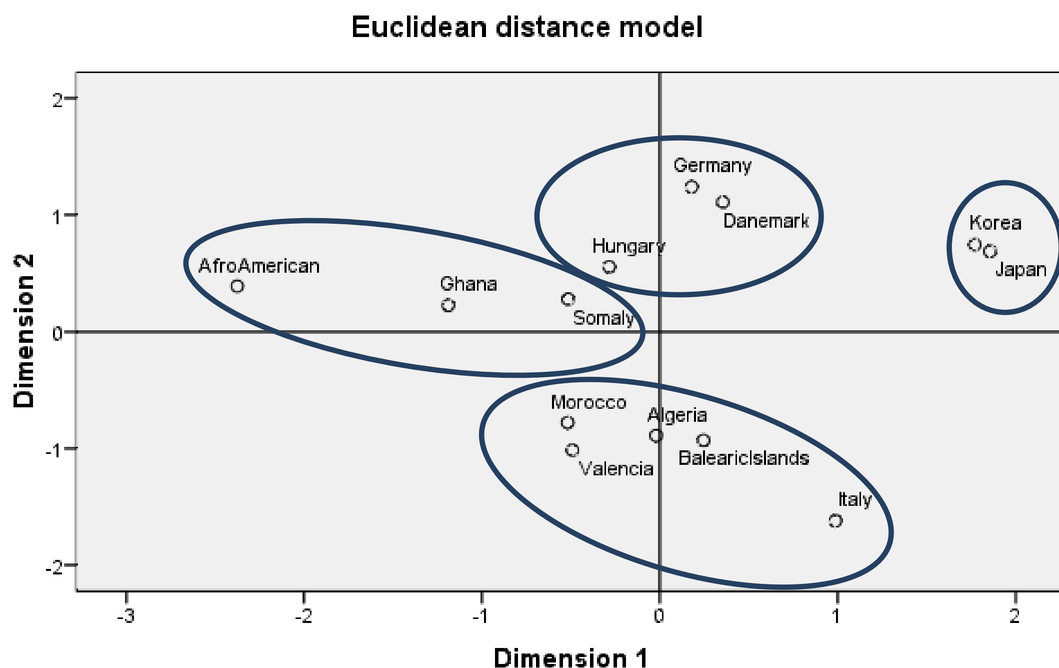


Figure 15: Multidiensional scaling analysis (MDS) based on Reynolds genetic distances calculated between 14 populations for 8 X-STRs.
Stress: 0,17467 RSQ= 0,83551

As described for other populations, a high level of polymorphism was observed for the 12 genetic markers included in Argus X-12 kit in the three populations analyzed in this study. A significant lower variability together with a high genetic differentiation from the other investigated populations was shown in the Balearic Islands population. Significant linkage disequilibrium was observed between markers inside linkage groups. As previously discussed by others, haplotype frequencies of each linkage group instead of allele frequencies need to be used in forensic cases. In order to increase the forensic utility of the combined X-STR markers, larger databases are then required.

Table. 14
X-chromosome haplotypes of four linked STR trios in 39 men from each Majorca, Minorca, Ibiza and Valencia

Nª	DXS10148	DXS10135	DXS8378	MA	MI	IB	VL	DXS7132	DXS10079	DXS10074	MA	MI	IB	VL
1	11	25	22.1	1	2			14	8	18	2			
2	12	32	28.1	1				15	15	21	1			1
3	9	21	23.1	1				13	15	15	1			
4	10	19	25.1	3				13	19	18	1	1	1	
5	11	28.1	26.1	1				14	15	21	1			
6	11	27	19	1				14	8	20	2		1	
7	12	27	19	1				12	17	20	1			
8	10	22	23	2				14	15	20	1	1		
9	11	27	23	1				16	15	20	1			
10	12	28	22.1	1				13	16	19	1	1		1
11	12	19.1	25.1	1				12	8	19	1			
12	10	21	23	1				14	18	21	1	3		1
13	12	18	25	1				13	17	19	1			1
14	11	18	18	1				13	8	19	1	2		1
15	12	25	24.1	1				13	13	23	1			
16	10	20	18	1				14	17	19	2		2	
17	9	24	24.1	1				14	17	18	2			
18	11	30	23.1	1				14	17	20	1			2
19	12	24	28.1	1				14	16	19	1			
20	11	29	25.1	1				13	16	16	1			
21	10	21	26.1	1				14	15	17	1			1
22	19	20	25.1	1				15	18	23	1			
23	12	27	28.1	1				13	18	20	1			1
24	13	30	26.1	1				14	8	19	1		1	
25	11	31	25.1	1				15	16	17	1	1		
26	11	18	24.1	1				15	16	18	1			
27	11	22	24.1	1				14	17	22	1	1		
28	10	19.1	26.1	1				16	19	19	1			
29	11	25	28.1	1				12	17	19	1	2	1	
30	11	22	25	1				13	8	20	1			
31	11	19.1	25.1	2				15	17	19	1			
32	12	20.1	23.1	1				14	8	17	1	2		1
33	12	19	29.1	1				15	14	20	1			
34	10	25	25.1	1				14	18	16	1			
35	12	20	23.1	1	1			14	18	18	1	4	2	1
36	13	25	21			1		13	19	20			1	
37	10	21	27.1			2		13	18	15			1	
38	11	23	22.1			1		12	8	18			2	
39	12	18	27.2			1		13	9	18			1	
40	13	22.1	25.1			1		15	17	16			3	
41	11.2	25	23			1		14	19	20			1	
42	13	22	25.1			3		13	17	21			1	1
43	12	25	26.1			1		13	19	21		1	2	
44	12	25	23			1		14	7	18			1	
45	11	23	25.1		1	1		13	16	17			1	
46	14	30	25.1			2		14	18	19			2	
47	12	18.1	24			1		13	8	18		1	2	
48	12	18	27.1			1		15	8	18		1	1	
49	12	22	25.1			1		15	16	20			1	
50	12	32	26.1			1		14	16	20			1	
51	10	25	24.1			3	1	12	14	18			1	
52	12	26	24.1			1		12	16	19		1	1	
53	11	27	24.1			1	1	14	17	17			1	
54	12	26	22.1			1		15	17	17			1	
55	11	25	25.1			1	1	12	18	21			1	
56	12	18	23.1			2		12	8	21			1	
57	10	32	28.1			1		14	18	18			1	
58	11	33.1	18			1		15	8	19		2		
59	10	18	27.1			1		13	17	20		1		1
60	12	27	24.1			1		16	8	17		1		
61	10	27	28.1			1		15	14	18		1		
62	10	22	24.1			1		16	17	19		1		
63	11	24	25.1			1		15	17	20		1		
64	11	28	25.1			1		12	16	18		1		
65	11	28	30.1			1		16	8	18		1		

Table 14. (Continued)

N[a]	DXS10148	DXS10135	DXS8378	MA	MI	IB	VL	DXS7132	DXS10079	DXS10074	MA	MI	IB	VL
66	10	29	26.1			1		12	19	17		1		
67	10	26	24.1			1		12	18	20		1		
68	12	25	19		2			13	7	20		1		1
69	11	19	24.1		1			15	18	21		1		
70	12	24	18		1			11	16	24		1		
71	10	21.1	23		1			15	8	16		1		
72	10	23	18		1			15	15	20		1		
73	12	32	27.1		1			14	17	16		1		
74	13	27	26.1		1			15	17	21		1		
75	12	21	27.1		1			12	7	20		1		
76	12	19.1	18		1			13	18	21		1		
77	12	17.1	24.1		1			14	13	20				1
78	12	27	22.1		1			13	16	18				1
79	12	28	25.1		1			12	15	19				1
80	12	22	26.1		1			13	15	18				1
81	10	23.1	18		1			15	16	21				1
82	10	22.1	24.1		1			14	9	17				2
83	10	29	22		1			15	17.2	20				1
84	10	19	26.1		1			13	18	17				1
85	13	24	26.1		1			17	8	21				1
86	11	28	18		1			14	15	18				1
87	12	24	24		1			12	19	21				1
89	11	23	28.1		1		1	14	9	21				1
88	11	23	18		1			16	8	19				1
89	11	22	23.1		1			14	17.2	22				1
90	12	25	27.1		1			13	15	19				1
100	12	23	18		1			13	16	21				1
101	11	24	23		1			14	19	19				1
102	12	16	25.1		1			15	15	20				1
103	12	19	22.1		1			15	15	19				1
104	12	17	24.1		1			12	16	21				1
105	10	23.1	25.1		1			15	15	23				1
106	12	21	25.1		1									
107	11	28	27.1		1									
108	12	19	24.1		1									
109	12	19	26.1				1							
110	10	20	25.1				1							
111	9	20	28.1				1							
112	12	19	25.2				1							
113	10	21.1	23.1				1							
114	11	29	19				1							
115	12	27	25.1				1							
116	11	18	23.1				1							
117	12	18	25.1				1							
118	12	26	25.1				1							
119	10	31	24.1				1							
120	11	25	24.1				1							
121	11	30	25.1				1							
122	10	25	18				1							
123	12	18.1	24.1				1							
124	10	18	24.1				1							
125	13	21	24.1				1							
126	10	27	18				1							
127	11	29	24.1				1							
128	13	24	24.1				1							
129	12	23	25.1				1							
130	9	16	25.1				1							
131	12	28	24.1				1							
132	11	19	22				1							
133	10	18	26.1				1							
134	13	26	25.1				1							
135	11	18	25.1				1							
136	11	21	18				1							
137	10	19	27.1				1							
138	12	17	26.1				1							

Table 14. (Continued)

N[a]	DXS10148	DXS10135	DXS8378	MA	MI	IB	VL	DXS7132	DXS10079	DXS10074	MA	MI	IB	VL
139	11	27	25.1				1							
140	12	18	18				1							
141	12	22	23.1				1							
142	12	22	24				1							

	MA	MI	IB	VL	MA	MI	IB	VL
h*	**0.8606 +/- 0.5229**	**0.8393 +/- 0.5126**	**0.8583 +/- 0.5217**	**0.8380 +/- 0.5117**	**0.7957 +/- 0.4907**	**0.8170 +/- 0.5016**	**0.8025 +/- 0.4940**	**0.8174 +/- 0.5014**

Table 14. (Continued)

Nª	DXS10103	HPRTB	DXS10101	MA	MN	IB	VL	DXS10146	DXS10134	DXS7423	MA	MN	IB	VL
1	20	31.2	12	1				33	14	31	1			1
2	18	31	13	1		2		37	14	27	1		1	1
3	18	29.2	12	1				37	15	28	1			
4	19	28.2	14	1	1			35	16	29	1			
5	18	29.2	11	1				38	13	27	1			
6	18	28.2	11	1				35	16	28	1			
7	19	27.2	13	1				35	15	28	1	2	1	
8	19	27.2	11	1		1		40.3	15	28	1			
9	19	31.2	12	1	1			41.3	15	41.2	1			
10	20	28.2	13	2				36	15	26	1			
11	19	29	12	1				37	15	31	1			
12	19	30.2	12	3		1	1	38	14	26	1			
13	19	31	14	1				36	15	31	1			
14	20	29.2	12	2	2	1	1	40.3	15	41.2	1			
15	17	31	13	2				37	16	29	1			
16	18	31.2	14	1				35	14	29	1			
17	19	31.2	11	1				40	14	26	1			
18	20	30.2	14	1				38	16	30	2			
19	19	29.2	12	2	1	1	1	36	14	25	1			
20	19	32.2	13	1				35	15	44.2	1			
21	16	28	10	1				37	15	26	1		1	
22	19	28	15	1				36	15	29	1		1	
23	18	30	15	1		1		41.3	15	28	1			
24	19	30.2	11	2	3	1	2	34	16	31	1			
25	16	30	12	1				41.3	15	44.2	1			
26	20	29.2	13	1	2		1	35	15	31	1			
27	19	29.2	15	1				36	16	28	1			
28	19	28.2	11	1		1		36	15	28	1	1		
29	19	28.2	12	1				35	14	26	2			1
30	18	32	13	1	1			40.3	14	39.2	1			
31	19	28	12	1		2		43.3	15	28	1		1	1
32	20	29.2	14	1				37	14	29	1			1
33	18	29.2	14			2		42.3	15	40.2	1			
34	19	30.2	12			2		36	14	42.2	1		1	
35	19	34	13			2		39.3	14	43.2	1			
36	18	32.2	11			1		33	15	28	1			
37	18	28.2	12			1		38	15	27	1			1
38	20	31.2	13			2		38	14	29		1	2	
39	16	33	14			1		39	14	26			1	
40	18	31	15			1		34	14	26			1	1
41	19	27.2	12			2		38.3	15	41.2			2	
42	17	30	13			1		36	15	28			2	
43	19	32	13			1		36	14	28			2	1
44	18	29.2	13			1		34	14	27			2	1
45	20	28.2	12			1		35	15	29			2	
46	19	29.2	13		1	3		37	14	45.2			1	
47	20	30.2	13		1	2		36	16	26			2	
48	19	31	13			1	1	37	14	30		1	1	
49	19	33	13			1		36	13	25			2	1
50	20	28	12			1		36	16	27		1	2	
51	18	30.2	12		1	1		40.3	14	28			1	
52	19	29.2	11			1	2	36	16	30			1	
53	15	33	12			1		39.3	15	41.2			1	
54	19	30.2	13		1			35	14	27			1	
55	18	32.2	13		1			36	16	43.2			1	
56	18	29	13		1			37	17	29			1	
57	18	32	15		1			34	15	28			1	
58	17	31	14		1		1	32	15	47.2			1	
59	18	31	14		1			37	16	25			1	
60	19	27.2	12		1			35	14	28			1	
61	20	29.1	16		1			36	15	27		2	1	
62	17	30	14		1			36	15	30		1	3	
63	20	31.2	13		1			41.3	16	41.2		1		
64	17	33	12		1			36	15	29		2		
65	18	30.2	13		1			38	16	29		2		

Nª	DXS10103	HPRTB	DXS10101	MA	MN	IB	VL	DXS10146	DXS10134	DXS7423	MA	MN	IB	VL
65	16	29.2	13		1			40	16	29		1		
66	18	29.2	12		1			35	15	25		1		
67	16	28.2	12		1			37	13	28		1		
68	21	31.1	12		1			35	14	26		1		
69	19	29.2	14		2			35	17	29		1		
70	18	30.2	9		1			41.3	14	28		1		
71	10	31.2	13		1			35	15	30		1		1
72	17	27	13		1			36	17	26		1		
73	16	32	15		1			37	14	24		1		
74	21	29.2	12		1			41.3	14	45.2		1		
75	17	31	12				1	36.2	14	29		1		
76	15	33	13				1	38	14	27		1		1
77	18	30	13				1	35	15	27		4		1
78	18	27.2	12				1	40.3	15	28		1		
79	18	28.2	13				1	32	15	46.2		1		
80	19	29	13				1	32	15	29		1		
81	18	33	13				1	36	15	41.2		1		
82	16	25.2	12				1	35	13	29		1		
83	19	33	14				1	42.3	14	36.2		1		
84	17	31	13				1	37.3	16	29		1		
85	16	29.2	12				1	39.3	14	26				1
86	19	32	15				1	36	14	29				1
87	18	30	12				1	32	14	29				1
88	20	30.2	12				2	39.3	15	39.2				2
89	16	27.2	12				1	37	16	28				1
90	18	33.2	12				1	34	14	28				1
91	21	29.2	10				1	34	15	29				1
92	19	31.2	14				1	35	15	26				1
93	16	29	13				1	37	13	26				1
94	20	28.2	11				1	34	15	25				1
95	19	31.2	13				1	35	15	42.2				1
96	16	33	13				1	42.3	15	27				1
97	21	29	13				1	36	15	25				2
98	19	29	12				1	33	16	40.2				1
99	16	32	12				1	35	15	28				1
100								35	17	30				1
101								36	16	40.2				1
102								38	14	24				1
103								35	15	31				1
104								35.3	14	28				1

	MA	MI	IB	VL	MA	MI	IB	VL
h*	**0.7818 +/- 0.4837**	**0.8129 +/- 0.4991**	**0.7616 +/- 0.4737**	**0.8092 +/- 0.4974**	**0.8210 +/- 0.5032**	**0.7947 +/- 0.4905**	**0.7949 +/- 0.4902**	**0.8025 +/- 0.4940**

III. Population genetic inference from X chromosome Alu insertions in Western-Mediterranean region

A total of 654 blood samples from healthy unrelated individuals males and females were collected from eleven Mediterranean populations from Balearic Island (Mallorca, Minorca, Ibiza), Iberian Peninsula (Valencia), South of Italy (Sicily, Cosenza, Catanzaro, Reggio Calabria) and Morocco (Arabs, Berber and Sahrawi).

Allele frequencies, Hardy-Weinberg equilibrium and heterozygosity

Genotyping of eleven Western Mediterranean population samples for the 9 X-chromosomal *Alu* insertions resulted in allele frequencies as shown in Table 15. Most of these loci were found to be highly polymorphic in the studied populations; however some of them are monomorphic and fixed for the absence or the presence of one allele across four populations mainly in the isolated areas like Ibiza. The Arabic Moroccan population showed a small level of variation at Yd3JX437 system. Our results clearly distinguished Reggio Calabria, Sicily and Catanzaro from the rest of Mediterranean populations at Ya5DP4, due to the remarkably higher allelic frequencies. Most of the studied populations were in Hardy-Weinberg equilibrium; nevertheless some of them showed a significant departure from equilibrium, considering Bonferroni correction (Yb8DP49 in Moroccan Arabic, Sahrawi, Majorca and Reggio Calabria; Yd3JX437 in Sahrawi; Yb8NBC634 in Reggio Calabria; Ya5DP77 in Minorca and Yb8NBC578 in Reggio Calabria). Since none of the deviations are assigned to a particular locus or populations, they probably represent random statistical fluctuations. Considering the mean heterozygosity for each of population, there was a general trend for moderate to high diversity in Western Mediterranean group as shown in Table 16. The average heterozygosity per locus ranged from 0.068 in Ya5DP77 to 0.251 in Ya5DP62. The average heterozygosity per population was lower in Ibiza (0.080) that in the rest of populations, ranging between 0.102 in Reggio Calabria and 0.185 in Valencia.

Reynolds' distances and Multidimensional scaling plot

The estimates of various genetic distance measures between each of the eleven Western Mediterranean populations were calculated on the basis of Alu insertion frequency data Table 16. Reynolds genetic estimate was the lowest (0.002749) between Moroccan Berbers and Sahrawi indicating the lowest genetic distance but highest genetic identity between these two populations for X chromosomal Alu polymorphism. On the other hand, the highest pairwise value (0.110513)

was found between Catanzaro and Moroccan Arabic suggesting greatest genetic distance but lowest genetic identity between these two populations.

Table 15:
Allele frequencies and heterozygosity per locus and population for the X-chromosomal ALU insertions Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4, Ya5DP13 in 11 Western Mediterranean populations.

	Nº of ChX	Ya5DP62	Yb8DP49	Yd3JX437	Yb8NBC634	Ya5DP77	Ya5NBC491	Yb8NBC578	Ya5DP4	Ya5DP13	*H* per population
MorArbc	72	0.750	0.722	0.042	0.944	1.000	0.917	0.931	0.014	0.889	0.163
MorBerb	67	0.791	0.821	0.239	0.955	0.896	0.955	0.881	0.015	0.955	0.167
MorShrw	49	0.796	0.816	0.204	0.959	0.939	0.959	0.898	0.000	0.980	0.165
Mallorca	68	0.721	0.824	0.103	0.971	0.985	0.971	0.985	0.015	0.812	0.154
Minorca	71	0.732	0.887	0.099	0.972	0.887	0.986	0.958	0.000	0.845	0.155
Ibiza	70	0.857	0.914	0.100	1.000	0.986	1.000	1.000	0.000	0.943	0.080
Valencia	60	0.800	0.715	0.133	0.967	1.000	0.900	0.917	0.017	0.833	0.185
R.Calabria	164	0.921	0.927	0.152	0.957	0.970	1.000	1.000	0.140	1.000	0.102
Sicilia	84	0.905	0.893	0.095	0.964	0.988	0.988	0.952	0.107	0.989	0.106
Catanzaro	61	0.918	0.918	0.152	0.951	0.967	1.000	1.000	0.198	1.000	0.112
Cosenza	30	0.933	0.933	0.167	0.967	0.967	1.000	1.000	0.000	1.000	0.116
H per locus		0.251	0.242	0.233	0.074	0.068	0.055	0.079	0.104	0.118	

Abreviation; MorArb: Moroccan Arabic, MorBerb: Moroccan Berber, MorShrwi: Sahrawi, *H*: Heterozygosity

Population distance relationships were represented through a Multidimensional scaling graph. This analysis highlights the similarity between the 11 Western Mediterranean groups of the present study, grouping them according to the Reynolds genetic distances calculated between pairs of population. At the end of the analysis an RSQ estimate is reported. This value is defined as the proportion of variance of the scaled data (disparities) in the partition which is accounted for by their corresponding distances. The RSQ (0.99481) indicates an average goodness-of-fit of the 2-dimensional projection to the data. The multidimensional scaling (MDS) plot places the populations into three discrete clusters (North African, Spain and South Italy) and one dispersed cluster of Ibiza and Crete Island. Ibiza is located in the left bottom of the plot. The South Italy (SI) populations cluster to the top left of the plot, well separated from the other populations. Within the SI cluster, Cosenza is on the far top of the cluster, while the rest of SI populations are on the far bottom of the cluster, and Sicily is the most differentiated, close to the center of the diagram. The North African populations (NA) are found close to the center of the plot, although the Moroccan Arabic and Tunisia populations are clearly separated to the right of the other NA populations. The Spanish (S) populations form a tight cluster in the lower right of the plot adjacent to the NA cluster. Within the S cluster, Balearic Islands are slightly removed from the other ES populations.

The analysis was repeated with further populations from the Mediterranean Sea, North of Africa, Sub-Sahara and America in order to evaluate the strength of the X chromosomal Alu insertion in population genetics. Three clear clusters resulted, all the Western Mediterranean group is grouped in a same cluster and cross the plot in its right side from the top to the bottom. The sub-Saharan (SS) population is displayed on the far top left of the plot, whereas the Amerindian cluster is located in the opposite side from the SS cluster, in the lower left of the plot, which improve the genetic difference between the three clusters.

Table 16: Reynods' distances for the eleven Western-Mediterranean populations

	MorArbc	MorBerb	MorShrw	Majorca	Minorca	Ibiza	Valencia	R.Calabria	Sicily	Catanzaro	Cosenza
MorArbc	0.000000										
MorBerb	0.044561	0.000000									
MorShrw	0.036477	0.002749	0.000000								
Majorca	0.019198	0.039666	0.036992	0.000000							
Minorca	0.035767	0.029742	0.031104	0.011366	0.000000						
Ibiza	0.059162	0.047069	0.036967	0.041860	0.035154	0.000000					
Valencia	0.009670	0.030497	0.027863	0.018756	0.036079	0.058521	0.000000				
R.Calabria	0.100355	0.054725	0.051136	0.083882	0.074211	0.037530	0.085176	0.000000			
Sicily	0.065810	0.045900	0.038000	0.063937	0.058043	0.023439	0.061105	0.009181	0.000000		
Catanzaro	0.110513	0.065979	0.065056	0.094387	0.086641	0.058810	0.094643	0.004368	0.017116	0.000000	
Cosenza	0.110391	0.051263	0.050696	0.089637	0.081974	0.051446	0.087655	0.004001	0.020015	0.005654	0.000000

Analysis of molecular variance

Most population pairwise F_{ST} tests were significant. Non-significant pairwise F_{ST} tests occurred between the populations in locus Yd3JX437, Yb8NBC634 and Ya5DP77. A global AMOVA estimated that 96.74% of the variation was within populations and only 3.36% was among populations (data not shown). When each locus was analyzed separately the AMOVA analysis showed values between 0% for locus Yb8NBC634 and 7.34% for locus Ya5DP13.

All Western Mediterranean populations were grouped into two geographical groups "South European" and "North African". The frequency variance between the two groups (F_{CT} = 1.20%) resulted lower than the diversity among populations within groups (F_{SC}= 2.99%).

Model-based inference of population structure

The ADMIXTURE model implemented in STRUCTURE program was used in order to assign individuals to genetically homogeneous clusters and thus infer the underlying population structure. The values of K we used for this purpose were ranged from K=1 to K=11. The lowest K likelihood value was for K=5 as shown in Table 17. However there was no significant difference

between K=4 and K=5 for inferring clusters to our samples. The same table displays the proportion of membership of each pre-defined population related to the 4 clusters. In this way, the pattern of membership for almost populations in the 4 clusters was 1:1:1:1 excepted for the Moroccan Arabic population, who had a 2:2:2:1 as pattern. If we consider a K=5, the inferring clusters change considerably.

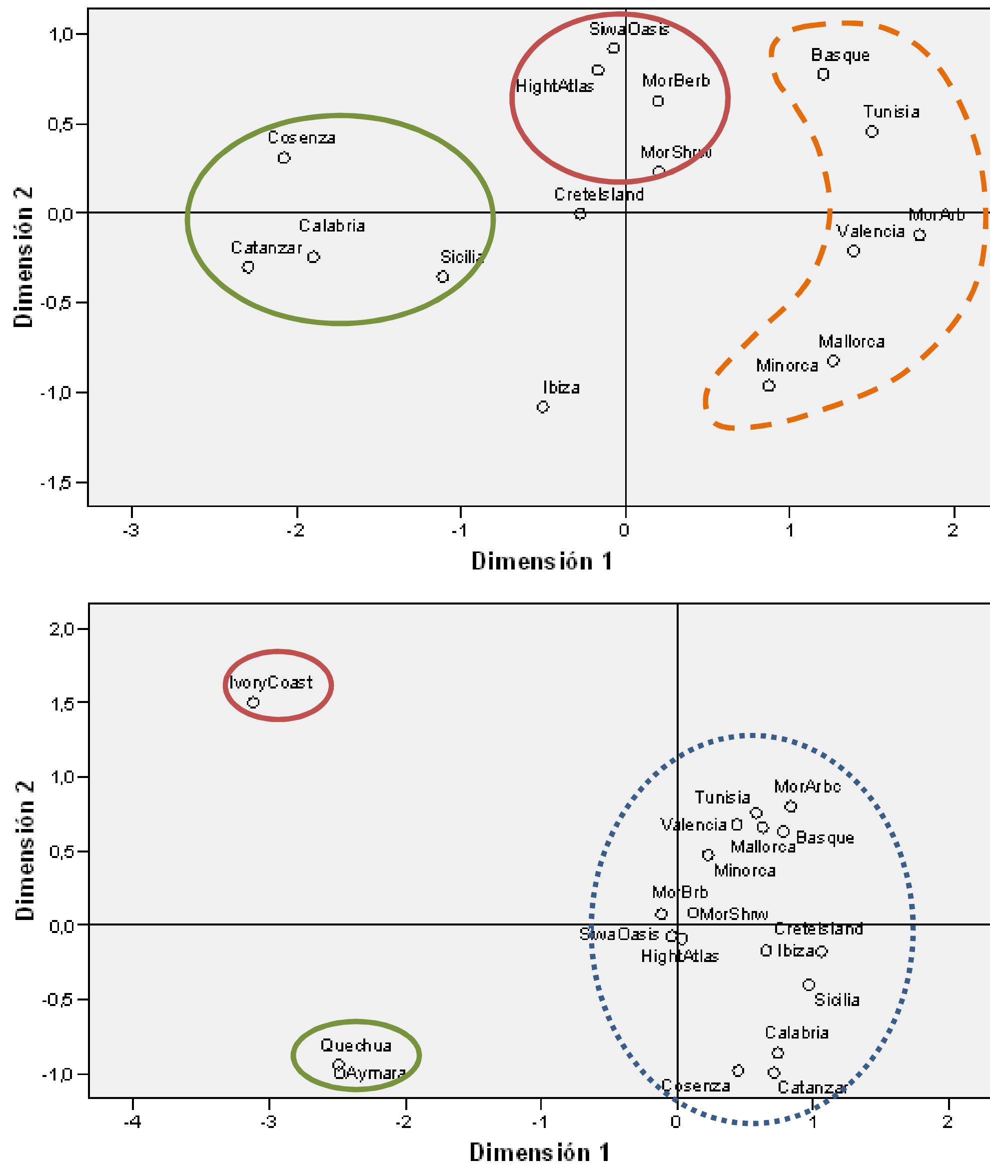


Figure 16: Multidimensional scaling analysis (MDS) based on Reynolds genetic distances calculated between populations.

That what the pattern of membership was 2:2:2:1:3 for the Moroccan Arabic population and 2:2:2:2:1 for the four Italian populations. Ibiza resulted different from the rest of the Spanish

populations with 2:2:2:1:1 against 2:2:2:1:2 for Majorca, Minorca and Valencia. Moroccan Berbers and Sahrawi belong to the same clusters 1:1:1:1:1. In fact the forth cluster seems to be characteristic to the South Italian population and the fifth one have been shown to be exclusive for describing the Moroccan Arabic population.

Discussion

The following study aimed to describe, as clear as the reliability of the nine X chromosomal *Alu* insertions used, the genetic structure of the Western Mediterranean region and to extent and complete other studies done with the same interest. The current study provides additional insight into Western Mediterranean substructure and differences among different ethnic groups that may impact our understanding of the patterns of human diversity, within and among the two neighboring continents. The allelic frequencies and average heterozygosity lead to an important conclusion about the genetic distribution of some very specific populations such the isolated Mediterranean Islands. Ibiza shows as well as the south Italian populations different patterns to other Mediterranean populations. The rest of populations remain in accordance with the general patterns described previously (Bentayebi et al. 2012). Overall, the general pattern of diversity is consistent with the major south/north increasing gradient as the difference within North African and European groups (mean heterozygosity for the North African samples: 0.165, against 0.126 for the European ones) confirming earlier studies (Pasino et al. 2011). The lowest gene diversity was found respectively in Ibiza (0.080) and south of Italy (0.109), supporting the hypothesis that the genetic differentiation of the two populations is a result of their particular demographic histories (Tomas et al. 2012; Edelmann et al. 2011 and Inturri et al. 2011). It is interesting to note the extreme frequency values found in seven loci near the fixation in some populations, for both absence (Ya5DP4) and presence (Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578 and Ya5DP13). The absence of an Alu insertion at a locus is the known ancestral condition. Fixed Alu are essential for use as shared derived characters to diagnose common ancestry among populations and species, whereas unfixed polymorphic Alu are identical by descent; that is, they are inherited from a common ancestor. These attributes also enable the six described Alu to be used to establish genealogies below the species level with minimal assumptions compared with other standard markers.

Table 17.
Estimated natural logarithm of the probability of the data (Ln P(X/K)) for each value of K and proportion of membership of each pre-defined population in each of the five clusters.

K	Ln P(X/K) All populations
1	-1918.6
2	-1664.7
3	-1496.9
4	-1460.7
5	-1457.7
6	-1474.9
7	-1525.8
8	-1582.6
9	-1725.9
10	-1795.6
11	-2275.3

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Morrocan Arabic	0.204	0.203	0.204	0.082	0.307
Morrocan Berber	0.184	0.184	0.185	0.217	0.231
Morrocan Sahrawi	0.204	0.204	0.204	0.201	0.187
Majorca	0.205	0.204	0.204	0.151	0.236
Minorca	0.200	0.201	0.200	0.157	0.243
Ibiza	0.230	0.230	0.230	0.168	0.143
Valencia	0.175	0.175	0.174	0.184	0.292
R.Calabria	0.219	0.219	0.219	0.234	0.109
Sicily	0.223	0.222	0.223	0.200	0.132
Catanzaro	0.219	0.218	0.219	0.233	0.112
Cosenza	0.226	0.225	0.225	0.215	0.109

This close genetic similarity of the Western Mediterranean group was also confirmed by the hierarchical AMOVA analysis. The genetic variation between the two sides of the Mediterranean Sea was low and insignificant, as revealed by the variation between the South-European and North-African groups of the Mediterranean region (F_{CT} = 1.20%). Some previous studies focusing on the analysis of the genetic variation between Western-Mediterranean populations supported and somehow confirmed the hypothesis of the North versus South differentiation gradient. Previous studies carried out by (Georgious et al. 2007) on the same set of Alu insertions used in this analysis in six Mediterranean populations and on polymorphic autosomal *Alu* insertions in the Western-Mediterranean populations (Comas et al. 2000 and Gonzalez-Perez et al. 2003), showed a lightly higher values of variations than ours (respectively F_{ct} = 1.24%; F_{ct} = 1.80% and F_{ct} = 1.96%). However, the population variation within groups in Comas et al. and Gonzalez-Perez et al. surveys was widely lower (respectively F_{sc} = 2.30% and F_{sc} = 0.47%) while in Georgious et al. 2007 was closely related (F_{sc} = 3.06%). The present results give evidence of the relevance of the X chromosome Alu insertions compared with the autosomes one and therefore, correlate the higher

population variation of those markers to the effect of the reduced population size of the X chromosome on population differentiation.

The Reynolds distances as well as the MDS diagram displayed an important genetic repartition of the different studied groups. When grouping our results with Georgious et al. ones, we revealed three clusters of homogeneous groups. The first one comprise south Italian population, the second one has two subgroups, on one hand brings Valencia and Basque Country to Moroccan Arabic and Tunisia, and on the other the Balearic Island group excepted Ibiza, which remain isolated. The third one, groups together three African Berbers population and Sahrawi population. The results seems support other studies concerning the Berbers origins who clustered separately together as showed further below, that suggest that the ancestors of the Berbers must come from the Europe and the Near East since the Neolithic.

The Mediterranean is formed in the process of continental drift, when Africa crashes against Eurasia. The resulting sea was of a size and a shape almost perfect for the development of civilization. More than 10,000 km of coastline, around a relatively calm sea, with plentiful harbours and numerous islands, provided an ideal setting for intricate patterns of trade, migration and warfare - all of which stimulate a mood of creative energy in human communities. Much of the rugged coastline on the northern shore is difficult terrain, hindering gene flow between the two continents. However, the geographical distance between the two shores at its narrowest part is short enough to allow ample migration (15 Km). Our result shows that gene flow between populations either within South-Western Europe or within North-Western Africa is not particularly insignificant and evidence to the presence of a genetic boundary through the Gibraltar Strait. Certain futures may have impacted on the level of gene flow between populations across this region: A wave of immigration occurred in the Neolithic synchronously along the two Mediterranean shores (Bosch et al. 1997 and Simoni et al. 1999). This fact may also have generated a cultural difference by bringing Indoeuropean languages to the Northern Mediterranean shore and Afroasiatic languages to the Southern shore (Renfrew et al. 1999 and Barbujani et al. 1994). Such cultural factors showing the same geographical discontinuity may have acted as enhancers of the genetic separation creating a positive feed-back mechanism of differentiation and producing the demographic scenario whose genetic consequences have been detected in the present study.

The results seems support other studies concerning the Berbers origins as well as Sahrawi who clustered separately together as showed further below, which imply that their arabisation has

certainly a cultural impact but did not replace the ancestry population. Several studies suggest that the ancestors of the Berbers must come from the Europe and the Near East since the Neolithic. In the other hand the Iberian populations included in our study clustered together with the North African populations witch can be explained by the genetic and cultural influence of the Arabic during their colonization to the Spain. As isolated area geographically, Balearic Island seems to be genetically differentiated but homogeneous even have been the cradle of various civilization. Excepted Ibiza, one of the three major islands of the Balearic archipelago, who was reproductively isolated for centuries, with a reduced population effective size due to the infectious diseases and a high number of consanguineous marriages reported in the 15th-17th centuries (Tomas et al. 2006). Our result supports the finding of Tagarelli et al. (2000) regarding Calabria region as a collection of many "human genetic isolates", due its population heterogeneity, geographic variation, and variability of ancient historical events. From the VIIIth to the Vth centuries BC, Calabria was colonized by Greeks; the territory was occupied by other groups: Romans, Longobards, Byzantines, Normans, Arabs, Hispanics, and the French (Lenormant et al. 1976). All of these groups probably contributed genetic influences to the indigenous population.

When more disparate populations are included, Amerindian (Quechua, Aymara) (Gaya-Vidal et al. 2010) and Sub-Saharan (Ivory Coast), a wide differentiation was detected and three clusters are formed, grouping the South European and the North African populations in a same clusters and defining a Sub-Saharan cluster apart as well as an Amerindian one. Those results argue about the reliability of the nine Alu markers used in inferring ethnicity and in studding the population differentiation and gene flow movement and the complex relationships reflect different population origins that include migration, admixture and isolation.

The Reynolds distances classified our samples as four distinct clusters as shown in the table below and evidenced by the MDS plot. The use of the admixture model yields the same clustering pattern and a similar trend for the likelihood assigned to each K. Thus, we are confident that our results represent the real structuring of the data.

In summary, this method of analysis allowed as when comparing our results with others to answer the questions mentioned below. It permits an appropriate comparison of results with those from the literature and completes other investigations focused on the Mediterranean populations because to our knowledge, it is the first time that these specific markers have been used in studying such Mediterranean populations. It also relates the results to natural, historical, and social events that characterize any territory and the people living there.

Gene Pool Structure of Jewish people as Inferred from the X-Chromosome

Demographic changes are among the major factors in the overall transformation of world Jewry during the last one hundred years. Changing patterns of family formation and childbearing, morbility, as well as socioeconomic and cultural changes have deeply affected the size, geographic distribution, and socio-demographic structure of Jewish populations worldwide. Demographic factors affected the inner structure of Jewish society, its variability and chances for continuity, and its relationship to the broader world (Coffman et al. 2005).

Jewish historians, linguists, anthropologists, and population genetists have not pooled their knowledge to produce a unified, acceptable classification of Jews. The Major Jewish communities are: Oriental Jewry who represents the original "gene pool" of the Jewish people; Sephardic Jewry who are an outgrowth of Oriental Jewry. The name Sephardi (in Hebrew meaning "Spanish") was acquired after expulsion from Spain, many subgroups evolved from Sephardic Jewry with communities along the northern (Turkey, Bulgaria and the Majorcan converted Jews: Chuetas) and southern (North African Jews: Morocco, Tunisia and Libya) shores of the Mediterranean and also in parts of Western Europe and North and South America; and Ashkenazi Jewry, who are an outgrowth of Oriental Jewry, mainly in Palestinian segment. The name Ashkenazi was acquired from their early roots in Germany (in Hebrew, Ashkenazi means "Germans"), and after a period of time they moved into Central and Eastern Europe. The main problems in classifying the Jewish communities deals with the many historic events that have molded these groups, and their various migration with exposure to forces, such as religious conversion, assimilation and intermarriage. Nevertheless, the amazing fact is the tenacity of the bonds that have kept most of these groups together throughout their difficult periods in history, resulting in the maintenance of distinct ethnic identities. In addition to Judaism serving as a binding force, the various Jewish languages that developed in the Diaspora also aided in uniting the people. Common to all these languages was the use of Hebrew alphabet. Perhaps the

two best known of these Jewish languages are Yiddish (Judeo-German) and Ladino (Judeo-Spanish) (Ben-Sasson et al. 1984; De-Lange et al.1984 and Stillman et al. 1979).

Several methodological problems are posed in the scientific study of Jewish populations. A primary difficulty concerns the availability and quality of data. High geographical dispersion of the Jewish population implies exposure of the Jews to a wide variety of political regimes and regulations concerning the organization of their own communities. Consequently, relevant data are very scattered, often not very reliable, and in many instances lacking altogether. In particular, after the Shoah (Holocaust) the vast majority of diaspora Jews live in countries with a separation between state and church, which results in the absence of official governmental data on religious groups. Thus the study of Jewish demography has become to depend predominantly on a variety of private research initiative; a few of wish are of excellent quality, but which often lack a common basis of concepts definition, and working techniques (Bonné-Tamir et al. 1992).

The very definition of the field of investigation, the “Who is a Jew?” question , has tended to become more complicated in the course of time because of the increasing frequency of intermarriage and other subtle processes of identificational change among modern and contemporary Jewries (Dever et al. 2003).

At a more analytical level, there is a need for comprehensive interpretative framework of Jewish population changes expanding beyond mere description of the statistical facts. A major issue is whether Jewish demography can be shown to have responded over time to a distinctive set of factor peculiar to the Jewish group regardless of time and space, or whether, on the contrary, Jews simply reflected the demographic patterns of the non-Jewish populations among which they lived. A related question concerns the development of Israel of a Jewish population majority, in contrast with the minority condition that was and still is typical of Diaspora Jewries. To subsume these two distinct situations in the one comprehensive description and meaningful interpretation is a challenging task.

Recent research makes it possible to present a brief overview of some of the leading patterns of change in Jewish populations over the last several generations. The

purpose of this chapter is the genetic characterization of five Jewish groups: Oriental, North African, Sephardic, Chuetas and Ashkenazi by means of STRs and Alus X chromosome analysis.

X-Chromosome Alu study

A large DNA sample collection of 247 individuals from five different geographic locations, Chuetas from the UIB collection and the rest from whom The National Laboratory for the Genetics of Israeli Populations at Tel-Aviv University, collected detailed ethnological and geolinguistic data. Inclusion criteria comprised an autochthonous origin confirmed by surnames and birthplaces of each individual, their parents and grandparents. Furthermore, the inclusion of participants of more than 60 years old was held to increase the probability of a deep local ancestry. On the basis of geographic location, linguistic affiliation, and surnames, individuals were attributed to five major zone categories: Ashkenazi, Sephardic (Turkey and Bulgaria), North African (Morocco, Tunisia and Libya), Oriental (Iraq and Iran) and Chueta Jews (Majorca).

Results

Genotype distributions and probability values concerning allele frequencies and heterozygosity levels per group and per locus for the 9 loci in the five Jewish communities are reported in Table 18. No deviation from Hardy–Weinberg equilibrium tests was detected considering the Bonferroni correction. Most of the Alu insertions were polymorphic in all populations, the exceptions being Ya5DP77 in Sephardic, Ya5NBC491, Yb8NBC578 and Ya5DP13 in which the Alu+ allele was fixed in Ashkenazi and Ya5DP4 in which Alu– was fixed in Oriental Jews. Considering the mean heterozygosity for each population, a low values were noticed without any striking difference between the most of Jewish groups (0.169 for Chuetas against 0.167 for Oriental and 0.165 for North African). Ashkenazi had the lowest heterozigozity values. Of all the loci analyzed, the one with the highest heterozygosity is Ya5DP62 (0.367) and the lowest one is Ya5DP4 with (0.029).

Table18:
Allele frequencies and heterozygosity per locus and population for the X-chromosomal ALU insertions Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4, Ya5DP13 in 5 worldwide Jewish populations.

	Nº of ChX	Ya5DP62	Yb8DP49	Yd3JX437	Yb8NBC634	Ya5DP77	Ya5NBC491	Yb8NBC578	Ya5DP4	Ya5DP13	*H* per population
Sephardic	111	0.793	0.721	0.135	0.964	1.000	0.946	0.928	0.018	0.883	0.145
North Africa	78	0.769	0.731	0.038	0.962	0.974	0.910	0.923	0.013	0.897	0.165
Oriental	75	0.693	0.733	0.013	0.933	0.840	0.987	0.933	0.000	0.947	0.167
Chuetas	141	0.716	0.816	0.092	0.972	0.901	0.979	0.965	0.014	0.830	0.169
Ashkenazis	35	0.800	0.829	0.143	0.943	0.971	1.000	1.000	0.029	1.000	0.119
H per locus		0.367	0.146	0.153	0.087	0.114	0.067	0.092	0.029	0.115	

In order to assess the relationship between the populations analyzed, Reynolds genetic distances were calculated and displayed in Table 19. The results indicated that around 76% of the distance values were significant. The highest distance was observed between Ashkenazi and Oriental groups (0.047071). It is worth noting that the distance between the Sephardic and North African samples of this study was the lowest values found (0.009090).

When we compared with the host populations whom with the Jewish groups lived during the Diaspora, the average distance value between the three ethnic groups of Morocco, Tunisia and the North African Jews was 0.055; the mean distance between the Balearic Islands (Majorca, Minorca and Ibiza) and Chuetas was 0.018, while the average distance between the Iberian peninsula (Basque country and Valencia) and Sephardic Jews was 0.011.

Table 19: Reynods' distances for the five Jewish studied populations.

	Sephardic	North African	Orientals	Chuetas	Ashkenazi
Sephardic	0.000000				
North African	0.009090	0.000000			
Orientals	0.037065	0.022497	0.000000		
Chuetas	0.020610	0.019916	0.021992	0.000000	
Ashkenazi	0.026391	0.037013	0.047071	0.034508	0.000000

Population distance relationships were represented through an MDS analysis (Figure 17) that illustrates the scattered position of the five Jewish groups, mainly between Orientals and Ashkenazi. In the second MDS we compared our samples with other populations including some new and unreported studies (Athanasiadis et al. 2007;

Bentayebi et al. 2011 and Bentayebi et al. 2012) (Figure 19). The Amerindian populations as well as the sub-Saharan one lies clearly in the other extremity of the graphic axe, while in the third MDS the studied Jewish Diaspora overlap with the countries where they settled during many centuries but remain plotted close to each other.

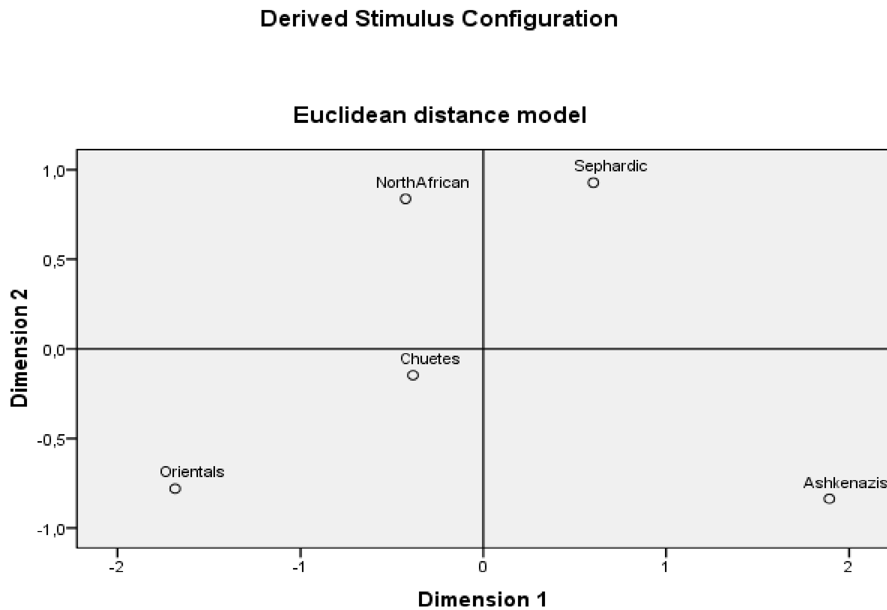


Figure 17: Multidiensional scaling analysis (MDS) based on Reynolds genetic distances calculated between five Jewish groups for 9 ALU insertions.
Stress: 0.00489 RSQ= 0.99982

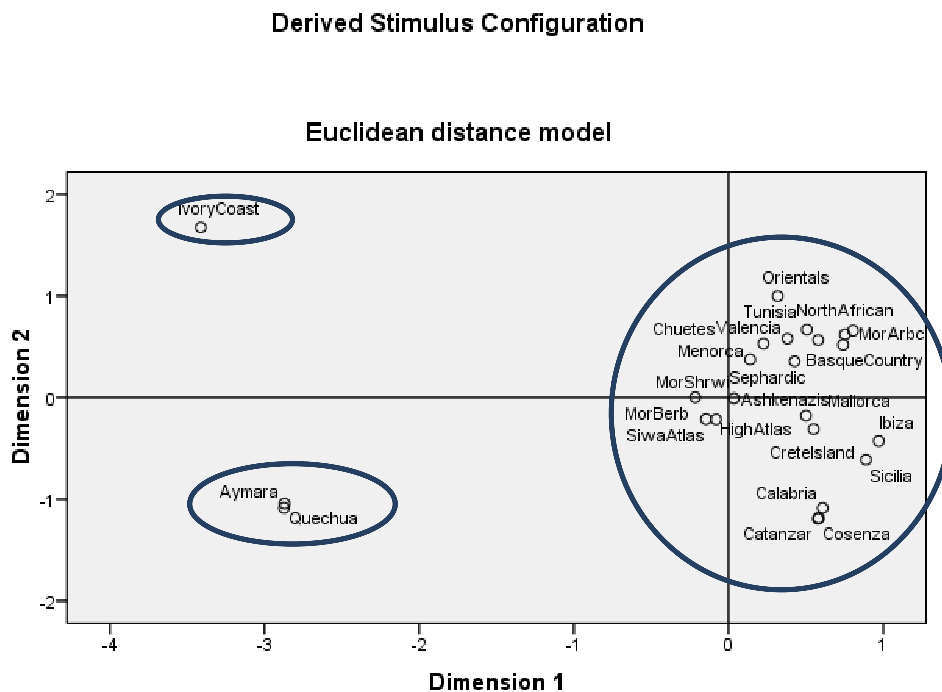


Figure 18: Multidiensional scaling analysis (MDS) based on Reynolds genetic distances calculated between 24 populations for 9 ALU insertions.
Stress: 0.07584 RSQ= 0.98607

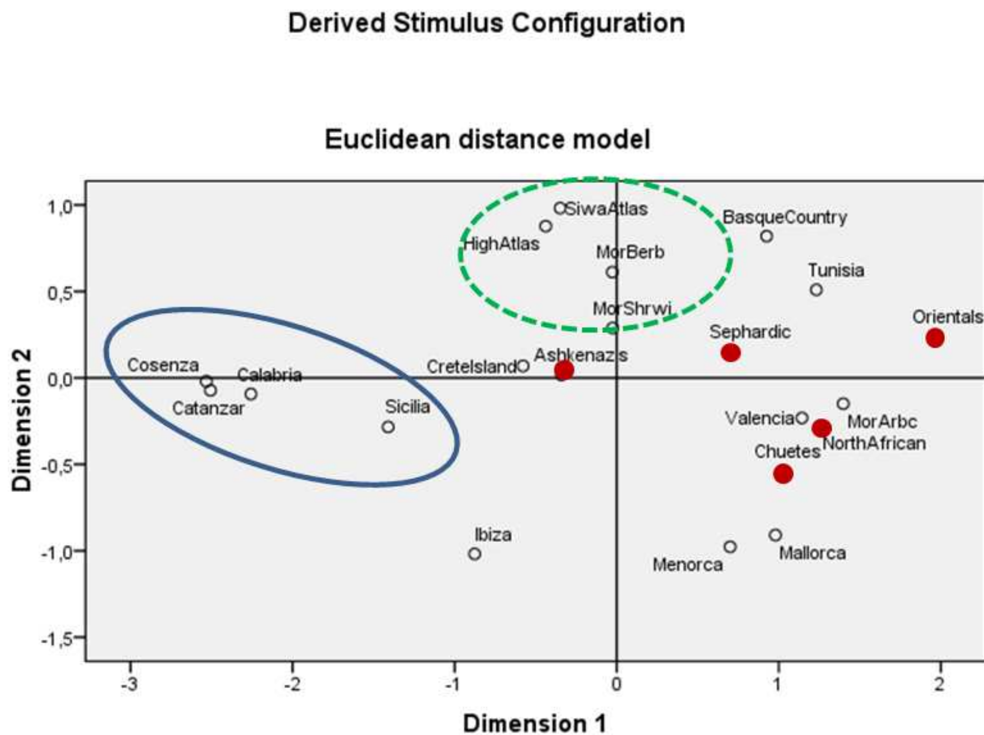


Figure 19: Multidiensional scaling analysis (MDS) based on Reynolds genetic distances calculated between five 21 populations for 9 ALU insertions.
Stress: 0.10506 RSQ=95393

Genetic differentiation among populations studied was estimated by calculating F_{ST} values of each Alu repeat. F_{ST} values are a measure of the amount of differentiation among subpopulations. The global F_{st} for the five Jewish communities was 3.77% and the F_{st} estimations by locus range from 0%for Ya5DP13 to4% Yb8NBC634. A hierarchical AMOVA analysis was performed on the Jewish populations. In this context, the total between-population diversity (F_{CT}= 0.90%) appears clearly lower than the diversity within groups (*Fsc* 2.87%).

Phylogenetic relationships evident in the MDS graph (Figures 19) is mirrored by the Structure bar plots assuming five ancestral populations (K=5). Individual ancestry proportions inferred by the Structure algorithms (Table 20) are consistent with the clustering patterns observed in the MDS plot discriminating among the five Jewish

populations. The membership of individuals within the graphical representation inferred by K=5 revealed that the pattern of membership of the Ashekenazi is the most differentiated with 2.4% of membership in cluster 4 and 30% in cluster 5 but generally there is no striking population substructure.

Table 20.
Estimated natural logarithm of the probability of the data (Ln P(X/K)) for each value of K and proportion of membership of each pre-defined population in each of the five clusters.

K	Ln P(X/K) All populations
1	-1243.9
2	-928.9
3	-785.2
4	-699.1
5	-651.6

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Ashkenazi	0.280	0.277	0.119	0.024	0.300
Sephardic	0.238	0.230	0.230	0.066	0.236
North African	0.207	0.206	0.237	0.107	0.244
Oriental	0.195	0.195	0.273	0.082	0.255
Chuetas	0.225	0.205	0.107	0.107	0.231

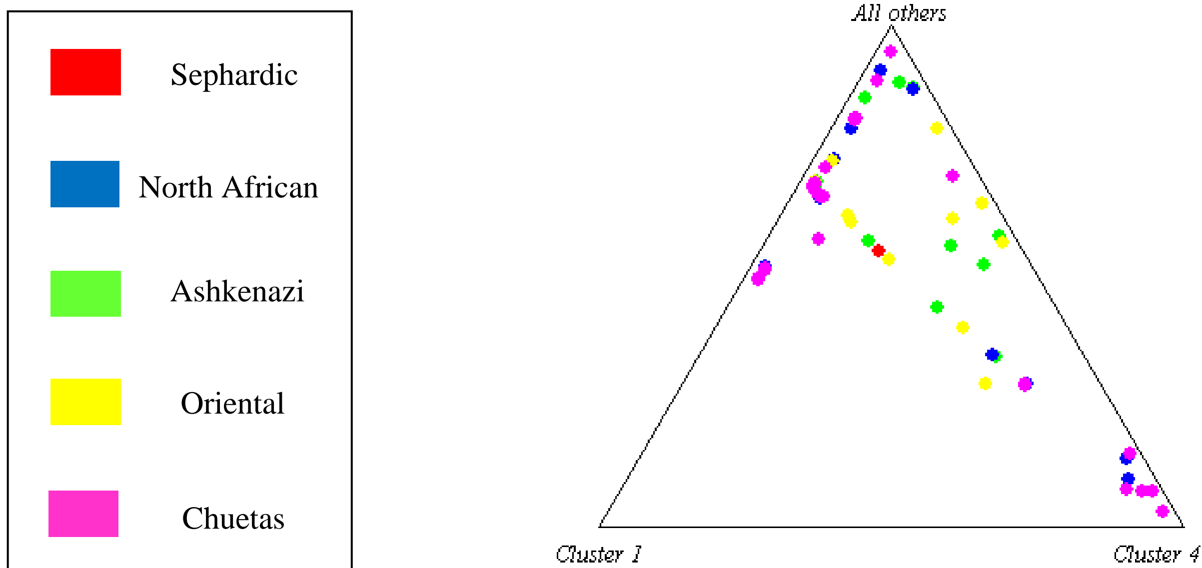


Figure 20: The triangle plot of Q. Each individual is represented by a colored point. The colors correspond to the prior population labels. The estimated ancestry vector for an individual consists of K components wish adds up to 1. For K=5, the ancestry vectors are plotted onto a triangle. For a given point, each of the five components is given by distance to one edge of the triangle.

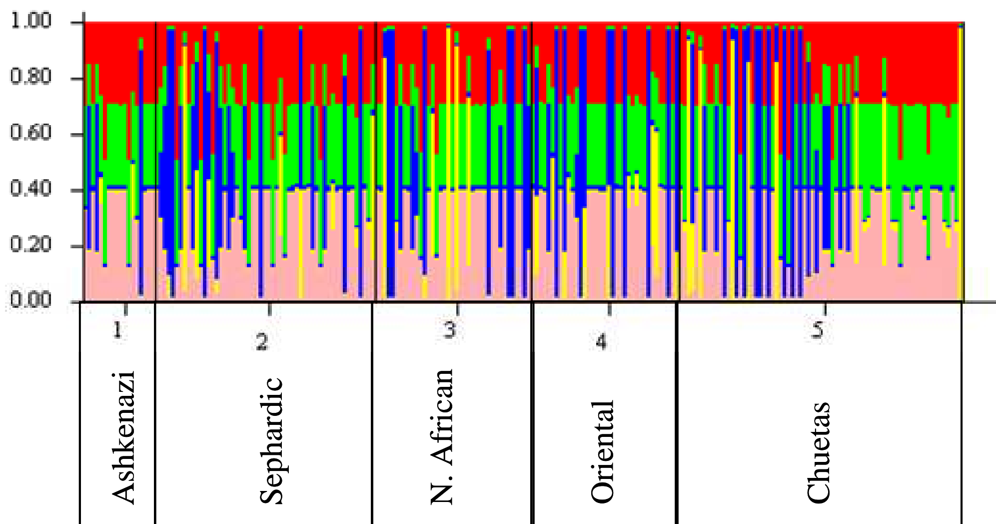


Figure 21: Population structure inferred by STRUCTURE analysis. Each individual is represented by a vertical (100%) stacked column of genetic components proportions shown in colour for K=3. The Jewish communities are labelled in colour and bold.

X-Chromosome STR study

The twelve X-linked STR markers were successfully amplified in one single-PCR multiplex reaction, following the manufacture's conditions. Allele frequencies obtained for the twelve X-STR loci studied (DXS10148, DXS10135, DXS8378, DXS7132, DXS10079, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS10134 and DXS7423) in the five Jewish communities (Ashkenazi, Sephardic, North African, Oriental and Chuetas) are presented in Table 21. Forensic statistical evaluation parameters were calculated in all five groups and are shown in the same table. Some alleles not reported in the Investigator Argus 12 X-STR manual of Qiagen were detected at seven of the loci studied: allele 34 at the DXS7132 locus (Chuetas); alleles 8 and 11 (Chuetas) and 25 (Ashkenazis) at the DXS10103 locus; alleles 8 and 18 (Chuetas) at DXS10134 locus;alleles 16 (Ashkenazis, Sephardic), 19 (Chuetas) and 21 and 21.2 (Sephardic) at DXS10101; alleles 26, 27 and 28 (Chuetas) at DXS10079; allele 12 (Chuetas) at DXS7423; and the alleles 16, 18 and 19 (Chuetas) and 47.2 (Ashkenazi) at DXS10146 locus. Confirmation of these alleles by sequencing is planned.

Table 21
Allele frequencies and statistical parameters of 12 X-chromosomal short tandem repeat loci in 265 Jewish samples (Ashkenazis: 26 men and 33 women; Sephardic (Bulgaria and Turkey): 19 men and 25 women; North African (Morocco, Tunisia and Libya): 29 men and 25 women , Oriental (Iraq and Iran): 29 men and 26 women) and Chuetas: (Majorca): 12 men and 41 women).

DXS10103

	ASH	SEPH	NAF	ORT	CHU
8	-	-	-	-	0.008
11	-	-	-	-	0.008
15	0.036	-	-	-	0.007
16	0.093	-	0.107	0.075	0.068
17	0.073	0.151	0.103	0.073	0.063
18	0.160	0.168	0.137	0.172	0.264
19	0.453	0.465	0.512	0.517	0.476
20	0.150	0.204	0.116	0.162	0.099
21	0.023	0.013	0.025	-	0.007
25	0.010	-	-	-	-
PIC	0.702	0.646	0.652	0.627	0.642
Het_{obs} †	0.697	0.760	0.680	0.654	0.805
Het_{exp}	0.736	0.702	0.648	0.672	0.695
MEC_{trio}	0.702	0.646	0.652	0.627	0.642
MEC_{duo}	0.563	0.500	0.507	0.480	0.498
PD_{female}	0.898	0.859	0.868	0.850	0.858
PD_{male}	0.731	0.691	0.683	0.666	0.685
PE	0.477	0.415	0.402	0.378	0.406
HWE†	0.941	0.022	0.437	0.122	0.044

DXS8378

	ASH	SEPH	NAF	ORT	CHU
8	-	0.013	-	-	-
9	0.071	0.013	0.013	-	0.070
10	0.270	0.265	0.342	0.372	0.273
10.2	0.010	-	-	-	-
11	0.408	0.319	0.328	0.350	0.329
11.2	-	-	-	-	0.007
11.3	-	-	0.011	-	-
12	0.208	0.323	0.294	0.233	0.288
12.1	0.010	-	-	-	-
13	0.023	0.040	0.011	0.011	0.025
14	-	0.027	-	0.034	-
15	-	-	-	-	0.008
PIC	0.663	0.669	0.624	0.621	0.679
Het_{obs} †	0.697	0.880	0.880	0.654	0.829
Het_{exp}	0.713	0.731	0.710	0.709	0.733
MEC_{trio}	0.663	0.669	0.624	0.621	0.679
MEC_{duo}	0.520	0.527	0.478	0.474	0.538
PD_{female}	0.869	0.871	0.839	0.837	0.877
PD_{male}	0.712	0.721	0.689	0.684	0.729
PE	0.446	0.462	0.411	0.403	0.474
HWE†	0.253	0.345	0.248	0.401	0.045

DXS7132

	ASH	SEPH	NAF	ORT	CHU
11	-	0.013	-	0.013	0.015
12	0.158	0.124	0.027	0.096	0.084
13	0.220	0.190	0.263	0.309	0.352
14	0.337	0.412	0.420	0.373	0.434
15	0.204	0.199	0.187	0.160	0.068
16	0.051	0.062	0.090	0.050	0.031
17	0.030	-	0.013	-	-
34	-	-	-	-	0.016
PIC	0.732	0.697	0.663	0.684	0.620
Het_{obs} †	0.788	0.680	0.680	0.615	0.537
Het_{exp}	0.765	0.753	0.724	0.749	0.672
MEC_{trio}	0.732	0.697	0.663	0.684	0.620
MEC_{duo}	0.597	0.557	0.520	0.543	0.476
PD_{female}	0.910	0.892	0.870	0.882	0.839
PD_{male}	0.768	0.735	0.710	0.729	0.675
PE	0.541	0.485	0.445	0.474	0.390
HWE†	0.265	0.761	0.278	0.359	0.101

DXS10074

	ASH	SEPH	NAF	ORT	CHU
7	0.010	0.018	-	0.024	0.014
8	0.089	0.159	0.212	0.144	0.100
9	0.026	0.035	0.076	0.037	0.044
10	0.010	-	-	0.013	-
11	-	0.027	-	-	0.008
13	-	0.031	-	-	-
14	0.013	0.027	0.025	-	0.067
15	0.103	0.044	0.126	0.034	0.130
16	0.193	0.155	0.168	0.185	0.112
16.2	0.010	-	-	-	-
17	0.188	0.226	0.158	0.288	0.216
17.2	-	-	-	-	-
18	0.183	0.155	0.151	0.161	0.233
19	0.165	0.124	0.059	0.103	0.022
19.3	0.010	-	0.011	-	-
20	-	-	-	0.013	0.039
21	-	-	0.013	-	-
28.2	-	-	-	-	0.008
30.2	-	-	-	-	0.008
PIC	0.869	0.838	0.836	0.801	0.834
Het_{obs} †	0.727	0.760	0.800	0.846	0.878
Het_{exp}	0.864	0.880	0.861	0.825	0.858
MEC_{trio}	0.869	0.838	0.836	0.801	0.834
MEC_{duo}	0.780	0.734	0.730	0.684	0.730
PD_{female}	0.975	0.962	0.961	0.946	0.961
PD_{male}	0.880	0.855	0.853	0.823	0.851
PE	0.755	0.704	0.701	0.642	0.697
HWE†	0.054	0.231	0.838	0.860	0.774

DXS10134

	ASH	SEPH	NAF	ORT	CHU
8	-	-	-	-	0.008
18	-	-	-	-	0.008
30	-	-	0.013	-	-
31	-	-	-	-	0.023
32	0.010	-	-	0.048	0.007
33	0.063	0.040	0.086	0.228	0.031
34	0.103	0.190	0.162	0.109	0.157
35	0.163	0.212	0.217	0.206	0.157
35.2	-	-	-	-	-
35.3	-	-	-	-	0.007
36	0.247	0.142	0.231	0.113	0.288
36.3	0.010	-	-	-	-
37	0.147	0.124	0.141	0.123	0.050
37.1	-	-	-	-	0.008
37.2	0.010	0.013	-	-	-
37.3	-	0.018	-	-	0.055
38	0.071	0.088	0.011	0.050	0.054
38.2	-	-	-	-	-
38.3	0.043	0.013	-	-	0.023
39	0.023	0.035	-	-	0.023
39.3	0.033	0.013	-	-	0.052
40	-	-	-	0.011	0.007
40.3	0.023	0.013	0.050	0.063	-
41.2	-	-	-	-	-
41.3	-	0.027	0.027	0.024	0.025
42.2	-	-	0.013	-	-
42.3	0.053	0.040	0.023	-	0.016
43.2	-	-	-	-	-
43.3	-	0.013	0.025	-	-
44.2		0.018	-	-	-
44.3	-	-	-	0.024	-
PIC	0.850	0.856	0.822	0.840	0.839
Het_{obs} †	0.848	0.920	0.920	0.807	0.927
Het_{exp}	0.845	0.876	0.871	0.857	0.853
MEC_{trio}	0.850	0.856	0.822	0.840	0.839
MEC_{duo}	0.753	0.760	0.713	0.737	0.738
PD_{female}	0.968	0.970	0.956	0.963	0.965
PD_{male}	0.864	0.869	0.841	0.855	0.853
PE	0.723	0.733	0.676	0.706	0.700
HWE†	0.006	0.028	0.232	0.336	0.764

DXS10101	ASH	SEPH	NAF	ORT	CHU
16	0.010	0.013	-	-	-
19	-	-	-	-	0.016
21	-	0.013	-	-	-
21.2	-	0.018	-	-	-
24.2	-	-	0.011	-	-
25	-	0.018	-	-	-
26	-	0.013	-	-	-
26.2	-	0.013	0.013	0.050	-
27	0.030	0.018	-	-	-
27.1	-	-	0.013	0.013	-
27.2	0.010	0.062	0.025	0.023	0.038
28	0.020	0.035	0.025	0.025	0.120
28.2	0.145	0.102	0.065	0.108	0.131
29	0.043	0.031	-	-	0.025
29.1	0.010	-	-	-	-
29.2	0.178	0.120	0.176	0.184	0.122
30	0.058	0.102	0.076	0.048	0.008
30.1	0.010	-	-	-	-
30.2	0.142	0.098	0.099	0.201	0.164
31	0.079	0.098	0.084	0.024	0.062
31.1	-	-	-	-	-
31.2	0.073	0.115	0.128	0.086	0.220
32	0.020	0.018	0.194	0.137	0.045
32.2	0.089	0.088	0.051	0.011	0.014
33	0.051	0.013	0.027	0.088	0.020
33.2	0.020	-	-	-	-
34	-	0.013	0.011	-	0.015
34.2	0.010	-	-	-	-
PIC	0.889	0.910	0.872	0.861	0.856
Het_{obs} †	0.909	0.920	0.800	0.692	0.854
Het_{exp}	0.901	0.939	0.897	0.885	0.873
MEC_{trio}	0.889	0.910	0.872	0.861	0.856
MEC_{duo}	0.808	0.840	0.783	0.766	0.760
PD_{female}	0.981	0.987	0.975	0.971	0.970
PD_{male}	0.897	0.916	0.883	0.873	0.869
PE	0.790	0.829	0.761	0.741	0.733
HWE†	0.143	0.603	0.111	0.098	0.079

DXS10135	ASH	SEPH	NAF	ORT	CHU
14	-	-	-	-	0.016
15	-	-	-	-	0.007
16	-	-	-	-	0.023
17	-	0.013	0.027	-	0.023
17.1	-	0.013	-	0.050	-
18	-	0.040	0.025	-	-
18.1	-	-	0.011	0.087	-
19	0.124	0.053	0.090	0.039	0.060
19.1	0.023	-	0.013	0.109	0.014
20	0.033	0.075	0.074	0.024	0.068
20.1	0.010	0.027	0.013	0.048	0.055
21	0.073	0.035	0.074	0.013	0.007
21.1	0.033	-	0.053	-	-
21.2	-	0.013	-	0.087	-
22	0.046	0.075	0.023	-	0.115
22.1	0.010	-	0.038	0.062	0.014
23	0.102	0.107	0.088	-	0.049
23.1	-	0.013	0.013	0.048	-
23.3	-	-	-	-	0.007
24	0.084	0.088	0.038	0.086	0.103
25	0.106	0.115	0.086	0.039	0.125
26	0.048	0.133	0.040	0.096	0.101
27	0.061	0.048	0.097	0.090	0.031
28	0.089	0.079	0.053	-	0.053
28.1	-	-	-	0.047	-
29	0.111	0.040	0.090	-	0.060
29.2	-	-	-	0.013	-
30	0.046	-	0.011	-	0.039
30.2	-	-	0.013	0.050	-
31	-	-	0.027	0.013	0.008
32	-	0.031	-	-	-
33.1	-	-	-	-	0.022
PIC	0.911	0.915	0.931	0.927	0.921
Het_{obs} †	0.818	0.920	0.920	0.962	0.878
Het_{exp}	0.917	0.931	0.942	0.935	0.938
MEC_{trio}	0.911	0.915	0.931	0.927	0.921
MEC_{duo}	0.842	0.849	0.874	0.868	0.858
PD_{female}	0.987	0.988	0.992	0.991	0.990
PD_{male}	0.917	0.920	0.934	0.931	0.925
PE	0.831	0.837	0.867	0.859	0.848
HWE†	0.030	0.363	0.188	0.645	0.001

DXS7423	ASH	SEPH	NAF	ORT	CHU
13	0.023	0.027	0.078	-	0.070
14	0.320	0.314	0.240	0.441	0.260
15	0.471	0.381	0.315	0.397	0.423
15.1	-	-	-	-	0.036
16	0.134	0.252	0.254	0.124	0.171
17	0.033	0.027	0.113	0.039	0.014
18	0.020	-	-	-	-
26	-	-	-	-	0.008
27	-	-	-	-	0.008
28	-	-	-	-	0.008
PIC	0.597	0.631	0.720	0.558	0.683
Het_{obs} †	0.697	0.760	0.760	0.538	0.585
Het_{exp}	0.618	0.666	0.776	0.625	0.730
MEC_{trio}	0.597	0.631	0.720	0.558	0.683
MEC_{duo}	0.451	0.485	0.583	0.412	0.542
PD_{female}	0.823	0.844	0.902	0.791	0.882
PD_{male}	0.656	0.691	0.760	0.631	0.723
PE	0.363	0.415	0.527	0.330	0.464
HWE†	0.879	0.338	0.775	0.379	0.023

DXS10079

	ASH	SEPH	NAF	ORT	CHU
12	-	-	-	-	0.008
13	-	-	-	-	0.008
14	-	-	-	-	0.008
15	0.010	0.040	0.023	0.011	0.031
16	0.033	0.031	0.011	0.062	-
17	0.056	0.075	0.036	0.271	0.081
18	0.180	0.181	0.151	0.213	0.225
19	0.280	0.283	0.314	0.366	0.273
20	0.256	0.235	0.279	-	0.114
20.1	-	0.013	-	0.051	-
21	0.122	0.102	0.111	0.013	0.221
22	0.043	0.040	0.048	0.013	0.022
23	0.020	-	0.027	-	-
24	-	-	-	-	0.008
30	-	-	-		-
PIC	0.775	0.787	0.753	0.822	0.777
Het$_{obs}$ †	0.818	0.920	0.680	0.731	0.805
Het$_{exp}$	0.808	0.815	0.797	0.733	0.812
MEC$_{trio}$	0.775	0.787	0.753	0.698	0.777
MEC$_{duo}$	0.651	0.666	0.625	0.560	0.654
PD$_{female}$	0.933	0.940	0.923	0.891	0.634
PD$_{male}$	0.802	0.812	0.783	0.740	0.805
PE	0.603	0.621	0.569	0.493	0.608
HWE†	0.504	0.550	0.927	0.587	0.164

DXS10148

	ASH	SEPH	NAF	ORT	CHU
13.3	-	-	0.013	0.023	0.046
14	-	0.013	-	-	-
17	0.010	-	-	-	-
18	0.154	0.137	0.090	0.050	0.063
19	0.033	0.058	0.025	0.024	0.050
20	-	-	0.027	-	0.007
20.3	-	-	-	-	-
21	0.010	-	-	-	0.014
22	-	-	-	-	0.007
22.1	-	0.027	-	0.037	0.069
23	0.020	0.040	0.011	-	0.007
23.1	0.097	0.031	0.086	0.097	0.092
23.3	-	0.018	-	-	-
24	0.010	0.013	0.025	0.013	-
24.1	0.103	0.013	0.053	0.025	0.041
25	-	-	-	0.011	0.029
25.1	0.312	0.349	0.240	0.202	0.094
26.1	0.118	0.155	0.154	0.257	0.288
26.2	-	0.013	-	-	-
27.1	0.063	0.062	0.114	0.185	0.163
27.2	-	0.013	-	-	-
28.1	0.036	0.027	0.124	0.050	0.023
28.2	-	0.013	0.011	-	-
29.1	0.020	0.018	0.025	0.025	-
30.1	-	-	-	-	0.007
31.1	-	-	-	-	-
36.1	0.013	-	-	-	-
PIC	0.821	0.807	0.856	0.822	0.844
Het$_{obs}$ †	0.727	0.920	0.880	0.654	0.756
Het$_{exp}$	0.651	0.524	0.886	0.849	0.874
MEC$_{trio}$	0.821	0.807	0.856	0.822	0.844
MEC$_{duo}$	0.713	0.694	0.760	0.713	0.744
PD$_{female}$	0.958	0.953	0.970	0.956	0.967
PD$_{male}$	0.837	0.823	0.869	0.840	0.856
PE	0.670	0.641	0.733	0.676	0.708
HWE†	0.462	0.218	0.987	0.056	0.075

DXS10146

	ASH	SEPH	NAF	ORT	CHU
16	-	-	-	-	0.008
18	-	-	-	-	0.016
19	-	-	-	-	0.008
23	-	-	0.013	-	-
24	0.020	-	0.013	-	0.022
25	0.033	0.084	0.091	0.013	0.030
26	0.051	0.186	0.063	0.124	0.181
27	0.135	0.120	0.185	0.123	0.155
28	0.140	0.133	0.151	0.146	0.068
29	0.101	0.092	0.113	0.160	0.184
29.1	-	0.026	-	-	-
30	0.102	0.026	0.097	0.113	0.027
31	0.053	0.013	0.011	0.113	0.060
32	0.010	-	0.013	-	-
32.2	-	-	-	0.013	-
33	-	0.017	-	-	0.046
35.2	-	0.017	-	-	-
36.2	-	0.017	-	-	-
37	0.023	-	-	-	-
38.2	-	0.013	-	-	-
39.2	0.033	0.043	0.064	0.011	0.023
40.2	0.051	0.040	0.013	0.039	0.007
41.2	0.043	0.030	0.037	-	0.016
42.2	0.041	0.040	0.013	0.013	0.093
42.4	-	-	0.013	-	-
43.1	-	-	0.013	-	-
43.2	0.043	0.047	0.013	0.024	0.008
44.1	0.020	-	-	-	-
44.2	0.045	0.043	0.073	0.071	0.049
45.2	0.030	0.017	-	0.025	-
46.2	0.013	-	0.011	0.013	-
47.2	0.013	-	-	-	-
PIC	0.916	0.898	0.887	0.878	0.874
Het$_{obs}$ †	0.818	0.840	0.840	0.846	0.878
Het$_{exp}$	0.924	0.923	0.903	0.902	0.888
MEC$_{trio}$	0.916	0.898	0.887	0.878	0.874
MEC$_{duo}$	0.851	0.822	0.806	0.791	0.787
PD$_{female}$	0.989	0.984	0.981	0.977	0.976
PD$_{male}$	0.921	0.905	0.896	0.889	0.885
PE	0.839	0.806	0.788	0.772	0.764
HWE†	0.004	0.057	0.507	0.192	0.000

HPRTP

	ASH	SEPH	NAF	ORT	CHU
10	-	0.053	-	0.176	0.007
11	0.122	0.177	0.074	0.123	0.118
12	0.342	0.332	0.258	0.305	0.243
13	0.378	0.349	0.399	0.184	0.484
14	0.125	0.075	0.156	0.162	0.094
15	0.033	0.013	0.099	0.039	0.030
16	-	-	0.013	0.011	0.008
18	-	-	-	-	0.016
PIC	0.658	0.682	0.694	0.770	0.641
Het$_{obs}$ †	0.667	0.760	0.800	0.538	0.659
Het$_{exp}$	0.703	0.749	0.732	0.798	0.676
MEC$_{trio}$	0.658	0.682	0.694	0.770	0.641
MEC$_{duo}$	0.515	0.541	0.555	0.644	0.496
PD$_{female}$	0.864	0.880	0.889	0.931	0.857
PD$_{male}$	0.709	0.728	0.734	0.799	0.683
PE	0.442	0.473	0.484	0.597	0.402
HWE†	0.172	0.514	0.655	0.001	0.025

*Abbreviations: ASH - Ashkenazi; SEPH– Sephardic; **NAF** – North African; ORT – Oriental; CHU- Chuetas ; PIC – polymorphic information content; Het$_{obs}$– observed heterozygosity;Het$_{exp}$– expected heterozygosity; MEC$_{trio}$– mean exclusion chance in trios involving daughter; MEC$_{duo}$ – mean exclusion chance in father/daughter or mother/son duos; PD$_{female}$ – power of discrimination in women; PD$_{male}$ – power of discrimination in men; HWE – Hardy Weinberg equilibrium. †P value calculating using the female data.

All loci selected for this multiplex study revealed to be highly polymorphic and as a result confirm their potential use for forensic purposes in these populations. DXS10135 was revealed to be the most polymorphic for Chueta, Oriental and North African Jewish communities while DXS10146 was the most polymorphic for Sephardic and Ashkenazi Jews followed by DXS10134 as the second most polymorphic marker for the five Jewish groups. The least discriminating locus in Ashkenazis was HPRTP, in Chuetas DXS7132 and in Sephardic, North African and oriental DXS742, as they have the lower number of alleles. The high values obtained for combined MEC_T and MEC_D in all five Jewish communities support the potential of this multiplex system in a specific kinship analysis context when the offspring is female or when father/daughter relationships are being investigated (Szibor et al. 2003). The same was observed for the overall values of PD_F and PD_Min all five populations. These high values of power of discrimination obtained both in females and in males support the value of this X-STR multiplex in forensic identity testing.

The exact test for linkage disequilibrium was performed for all pairs of loci in the five population groups. The Chueta population showed the highest level of LD. For this population, significant LD was observed in all loci pairs inside each linkage group. Indeed a higher LD have been observed in the European (Ashkenazi, Chuetas and Sephardic) compared to the North African group, supporting Arlden et al. hypothesis that assume that shows that LD for European is higher compared to that in human populations (Ardlie et al. 2002).

Haplotype frequencies studied for each of the four linkage groups and for each population (Ashkenazi, Sephardic, North African, Oriental and Chuetas) are shown in Table 22.All Jewish populations show very high gene diversity (*GD*). A low percentage of unique haplotypes was observed among Chuetas compared to the percentage observed in the other Jewish populations.

Population differentiations of the five Jewish population groups studied in the present work have been evaluated by genetic distance analysis. Data for the 12 studied X-STR systems were collected in order to calculate pairwise F_{ST} values between populations. The highest F_{ST} value was observed for DXS10079 between North African and Oriental (F_{st}=0.073). The X-STR systems DXS10135, DXS7423, HPRTP and

DXS10148 reached F_{ST} values higher than 0.020, which mainly involved the Oriental population (data no shown). Pairwise F_{ST} values calculated between 13 populations (Bentayebi et al. 2012; Bentayebi et al (in course); Tomas et al. 2012; Edelmann et al. 2011; Inturri et al. 2011 and Bekada et al. 2010) for the 12 X-STRs were averaged and represented in an MDS plot (Table 23 and Figure 22). As shown in this diagram, all the Jewish communities clustered together in the same side of the diagram while the western Mediterranean populations lie in the opposite side of the MDS according to their geographical origins and genetic affinities.

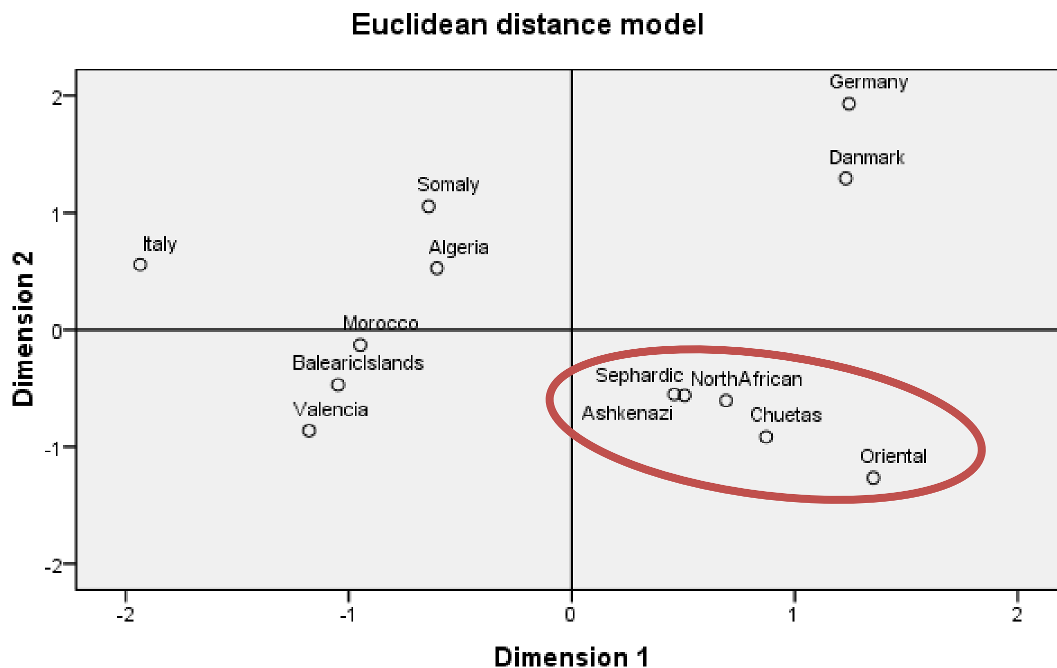


Figure 22: Multidiensional scaling analysis (MDS) based on pairwise genetic distances calculated between 13 populations for 12 X-STRs.
Stress: 0.18645 RSQ= 0.81648

X-Chromosome STR plus Alu polymorphism study

Samples form five Jewish communities were analyzed for nine X chromosomal Alu insertions as well as for 12 X-STR chromosome and few discrepancies were observed between the results obtained from both genetic markers in the studied Jewish communities. As described for other populations, a high level of polymorphism was observed for the 12 genetic markers included in the Investigator Argus X-12 kit in the

five populations included in this investigation. Significant linkage disequilibrium was observed between markers inside linkage groups, especially in Chuetas. As previously discussed by others, haplotype frequencies of each linkage group instead of allele frequencies need to be used in forensic cases, the present results could serve as the first database of the combined X-STR markers for five main the Jewish communities that should be completed by other investigations in order to increase their forensic utility.

The hierarchical AMOVA analysis (in both kind of markers) revealed that there is only a small genetic variance between North African-Oriental Jews and the other European Jewish communities.

Reynold distance for the Alu polymorphisms as well as pairwise F_{st} for X-STR markers indicate that the genetic distances between the five Jewish groups are slightly different but remains tightly related when compared with other populations. In the MDS built from the 9 Alu Reynold distances most Jewish samples, overlies non-Jewish samples. As shown in the Diagram of Jewish sub-clusters: the first group is located close to the center and consists of Ashkenazi and Sephardic Jews. The second group, comprising the Chuetas and North African Jewish communities, is positioned within the large conglomerate of non-Jewish Spanish and Moroccan Arabic populations. According to historical data there were migrations between Jews settled in mainland Spain and Balearic Islands to North Africa, explaining the similarity of these three populations. Also it is interesting to point out that the Chuetas are in the middle of Sephardic Jews and Balearic Islands populations, indicating the admixture with their host population described in the literature. The third group contains only a tight cluster of the Oriental Jews. The Jewish communities could be plotted in a wide-ranged cluster that is centrally located on the Multidimensional Scaling analysis (MDS) plot when compared with the other non-Jewish populations with a strike tendency to the right part of the diagram axe's where the Oriental Jews community plots (Figure 19). This Jewish cluster consists of samples from most Jewish communities studied here, which together cover more than 90% of the current world Jewish population (Pergola et al. 1997); this is consistent with an ancestral Levantine contribution to much of contemporary Jewry.

To glean further details of Jewish genetic structure, we repeated MDS based on the F_{st} pairwise genetic distances of 12 X-STR markers(Figure 22). This analysis

confirm the common genetic pool of the studied Jewish groups as they are depicted in only one tight cluster comprising the Ashkenazi, Oriental (Iranian and Iraqi), North

African (Moroccan, Tunisian and Libyan), Sephardi (Bulgarian and Turkish) and Chuetas (Majorca) Jewish communities plotted in the other extremity of the diagram distant from the European and Western Mediterranean populations included in this analysis.

After elucidation of these groupings by MDS, we turned to the algorithm ADMIXTURE (INTERALPHA=1.0) included in the program STRUCTURE, using prior population information and assuming correlated allele frequencies to assign individuals proportionally to hypothetical ancestral populations (Supplementary Note 3). All run included a burn-in period of 50 000 iterations followed by 10^6 iterations. Initially, all Jewish samples were analyzed jointly. This analysis significantly refines and reinforces the previously proposed partitioning of Old World population samples into continental groupings. We note that membership of a sample in a component that is predominant in, but not restricted to, a specific geographic region is not sufficient to infer its genetic origins. Membership in several genetic components can imply either a shared genetic ancestry or a recent admixture of sampled individuals. An illustrative example at K=5 is the pattern of membership of Ashkenazi, Chuetas (Majorca), Oriental (Iranian and Iraqi), north African (Moroccan) and Sephardic (Bulgarian and Turkish) Jewish communities in the light-green, blue, pink, red and yellow genetic components, which is very similar in all of theme suggesting a shared regional origin of these Jewish communities despite their actual geographic origins. This inference is consistent with historical records describing the dispersion of the people of ancient Israel throughout the Old World. Our conclusion favoring common ancestry over recent admixture is further supported by the fact that our sample contains individuals that are known not to be admixed in the most recent one or two generations.

Table 22.
X-chromosome haplotypes offour linked STR trios in males belonging to five different Jewish communities

N	DXS10148	DXS10135	DXS8378	ASH	SEPH	NAFR	ORT	CHU
1	9	24	23.1	2				
2	10	23	36.1	1				
3	11	19	25.1	1				
4	12	30	23.1	1				
5	10	26	19	1				1
6	11	24	23.1	1				
7	11	20	26.1	1				
8	10	19.1	23.1	1				
9	10	28	27.1	1				
10	11	30	26.1	1				
11	11	21	28.1	1				
12	9	24	28.1	1				
13	10	26	25.1	1				
14	10	21.1	25.1	1				
1	11	22	25.1	1		1		
5	10	22	25.1	1			1	
16	11	26	26.1	2				
17	12	23	25.1	1				
18	11	25	26.1	1			1	
19	10	28	26.1	1				
20	10	25	18	1	1			
21	10	23	26.1	1				
22	11	23	25.1	1				
23	13	28	24.1	1				
24	12	21	29.1		1			

Table 22 (continued)

Nº	DXS10148	DXS10135	DXS8378	ASH	SEPH	NAFR	ORT	CHU
25	10	21	26.1		1			
26	12	26	25.1		1			1
27	12	27	27.1		1			
28	11	26	19		1			1
29	12	28	26.1		1			
30	11	22	27.1		1			
31	11	20	18		1			
32	10	28	25.1		1			
33	12	28	25.1		1			
34	12	22	18		1			
35	11	24	25.1		1			
36	10	32	18		1			
37	10	27	23.1		1			
38	12	24	25.1		1			
39	11	26	25.1		1			
40	11	20	25.1		1			
41	10	25	23.1		1			
42	12	30	25.1			1		
43	11.3	18.1	27.1			1		
44	13	20	26.1			1	1	
45	12	25	25.1			1		
46	10	23	27.1			1	1	
47	10	22.1	23			1		
48	12	23	28.1			1		
49	11	27	28.2			1		
50	12	19	23.1			1		

Table 22 (continued)

Nº	DXS10148	DXS10135	DXS8378	ASH	SEPH	NAFR	ORT	CHU
51	12	18	23.1			1		
51	10	25	23.1			1		
52	11	19	26.1			1	1	
53	11	20	28.1			1		
54	10	24	23.1			1		
55	10	23	28.1			1		
56	11	29	25.1			1		
57	11	25	18			1		
58	10	27	27.1			1		
59	10	22	28.1			1		
60	11	27	28.1			1		
61	12	21	25.1			1		
62	12	27	26.1			1		
63	10	21	19			1		
64	11	27	25.1			1		
65	11	21	24			1		
66	10	25	25.1			1		1
67	10	20	18			1		
68	11	29	29.1			1		
69	11	18	26.1				1	
70	10	20	26.1				1	
71	11	25	27.1				1	
72	10	22	27.1				1	
73	10	19	22.1				1	
74	10	29	28.1				1	
75	12	21	27.1				1	
76	12	27	23.1				3	

Table 22 (continued)

N	DXS10148	DXS10135	DXS8378	ASH	SEPH	NAFR	ORT	CHU
77	12	29	27.1				1	
78	12	29	26.1				1	1
79	11	20	19				1	
80	10	24	25.1				1	
81	12	23	26.1				1	
82	14	20	13.3				2	
83	10	24	26.1				1	
84	11	27	22.1				1	
85	14	21	22.1				1	
86	10	31	26.1				1	
87	10	20.1	18				1	
89	11	25	25				1	
90	12	19	24.1					1
91	9	22	24.1					1
92	12	17	27.1					1
93	11	15	30.1					1
94	10	19.1	25.1					1
95	10	28	22					1
96	11.2	20	25					1
97	12	18	25.1					1
98	12	19	21					1
99	11	26	27.1					1
100	10	29	25.1					1
101	10	20	27.1					1
102	11	22	26.1					3
103	11	33.1	26.1					2
104	12	24	24.1					2

Table 22 (continued)

N	DXS10148	DXS10135	DXS8378	ASH	SEPH	NAFR	ORT	CHU
105	12	19	25.1					1
106	12	20	18					1
107	10	28	18					1
108	10	23.3	18					1
109	12	17	21					1
110	11	19	27.1					1
111	10	22.1	26.1					2
112	12	16	26.1					1
113	11	20.1	19					1
114	12	21	26.1					1
115	11	29	27.1					1
116	10	25	26.1					1
117	10	25	20					1
118	11	24	26.1					1
119	12	19	18					1
120	11	25	24.1					1
121	10	30	27.1					1
122	10	29	18					1
123	11	30	27.1					1
124	12	20	24.1					1
125	11	19.1	19					1
126	10	27	23					1
128	12	28	18					1
129	9	22	25.1					1
				0.830769 +/- 0.514607	0.820663 +/- 0.516923	0.849754 +/- 0.522060	0.848112 +/- 0.521239	0.849490 +/- 0.514715

_N[a]	DXS7132	DXS10079	DXS10074	ASH	SEPH	NAFR	ORT	CHU
1	12	18	20	2				1
2	15	18	20	1		1		
3	14	8	17	1		1		
4	15	16	18	1				1
5	12	19	19	1	1			
6	14	18	16	1				
7	15	9	20	1				
8	15	9	19	1	1			
9	15	17	20	1				
10	14	19	19	1	1			1
11	14	8	18	1				
12	15	19	18	1	1	1		
13	12	19	21	1				
14	13	14	19	1				
15	14	16	19	1				
16	14	17	19	2				
17	12	16	21	1				
18	14	17	21	1				3
19	14	17	22	1				
20	15	17	19	1	2			
21	13	15	21	1				
22	13	16	19	1				1
23	15	8	17	1				
24	14	8	20		1	1	1	
25	12	18	18		1			1
26	15	8	21		1	1		
27	14	8	21		1	2		

Table 22 (continued)

_N[a]	DXS7132	DXS10079	DXS10074	ASH	SEPH	NAFR	ORT	CHU
28	16	16	19		2			
29	14	9	18			1	1	
30	14	13	20		1			
31	14	18	18		1			
32	13	15	19		1			
33	13	19	17		2			
34	15	7	16		1			
35	13	17	19		1			
36	16	16	18		1			
37	13	9	19			1		
38	13	19.3	22			1		
39	13	15	16			1		
40	13	17	18			1	1	
41	14	8	22			1		
42	14	16	20			1		
43	14	19	15			1		
44	15	16	21			1		
45	13	18	18			1	1	6
46	16	15	20			1		
47	15	19	22			1		
48	13	19	19			1		
49	14	16	17			1		
50	13	14	20			1		
51	14	18	20			1	1	
52	13	18	21			1		2
53	14	9	19			1		2
54	15	15	20			2		

Table 22 (continued)

_N[a]	DXS7132	DXS10079	DXS10074	ASH	SEPH	NAFR	ORT	CHU
55	13	8	18			1	1	1
56	16	8	18			1		
57	13	16	20				1	
58	13	17	20				3	
59	15	8	18				1	
60	15	16	16				1	
61	12	18	19				1	
62	12	16	18				1	
63	14	17	20				1	
64	13	15	17				1	
65	16	18	18				1	
66	15	18	19				2	
67	12	8	20				3	
68	14	15	20				1	
69	14	7	19				1	
70	15	18	17				1	
71	13	15	20				3	
72	13	17	21				2	
73	13	16	15				1	
74	14	20	21				1	
75	14	17	17				1	
76	13	16	17				1	
77	13	7	18				1	
78	16	8	19				1	
79	14	15	17				1	
80	13	18	20				1	
81	12	18	17				1	

Table 22 (continued)

_N[a]	DXS7132	DXS10079	DXS10074	ASH	SEPH	NAFR	ORT	CHU
82	14	16	18				2	
83	13	18	19				1	
84	15	9	17				1	
85	14	18	19				1	
86	14	14	20				1	
87	13	14	21				1	
88	15	8	19				1	
89	13	7	22				1	
90	11	8	18				1	
91	15	16	19				1	
92	13	19	20				1	
93	14	15	22				1	
94	12	17	21				1	
95	13	14	18				1	
				0.807179 +/- 0.502737	0.832359 +/- 0.522891	0.816913 +/- 0.505614	0.733169 +/- 0.463537	0.784014 +/- 0.482348

Table 22 (continued)

_N[a]	DXS10103	HPRTB	DXS10101	ASH	SEPH	NAFR	ORT	CHU
1	18	30	14	2				
2	16	31	12	1				
3	19	32.2	14	1				
4	18	29.2	12	1				
5	19	28.2	11	1	1		1	1
6	19	32.2	13	1	2			
7	20	30	13		1			
8	20	28.2	13	1				
9	19	31	15	1				
10	15	30.2	11	2				
11	19	29.2	13	2			1	1
12	20	31.2	12	1			2	
13	17	29	14	2		1	4	1
14	19	29.2	12					1
15	20	30.2	12	1				1
16	18	30.2	13	1			2	1
17	20	29.2	13	1				
18	18	31	14	1				
19	21	29.2	11	1				
20	18	28.2	12	1				
21	18	32.2	12	1				1
22	19	28	12	1		1		
23	17	29	13		1			
24	19	31.2	11		1			
25	18	31.2	14		1			
26	20	30.2	11		1			
27	18	31	12		1			

Table 22 (continued)

_N[a]	DXS10103	HPRTB	DXS10101	ASH	SEPH	NAFR	ORT	CHU
28	18	30	12		1		1	
29	19	27	12		1			
30	19	27.2	13		1			1
31	18	25	12		1			
32	20	28	11		1			
33	18	32	13		1			2
34	20	27.2	12		1			
35	19	28.2	14		1			
36	19	21.2	13		1			
37	19	28.2	12		1	1	3	
38	19	27.2	12			1		1
39	19	34	13			1		
40	19	30	15			2		
41	18	32	14			1		
42	19	31	12			1		
43	20	32	12			1		
44	21	24.2	13			1		
45	19	30.2	11			1		
46	19	32	14			1		
47	18	30.2	12			1		1
48	19	31	13			1		
49	18	29.2	12			1		
50	20	31.2	15			1		
51	19	29.2	11			2	1	1
52	19	30.2	12			1		2
53	19	30.2	13			1		5
54	17	31	12			1	2	

Table 22 (continued)

_N[a]	DXS10103	HPRTB	DXS10101	ASH	SEPH	NAFR	ORT	CHU
55	19	31.2	13			1		3
56	18	31	13			1		
57	17	31	13			1		1
58	18	32.2	15			1		
59	19	31.2	12			1		
60	19	27.2	10				1	
61	19	32	10				1	
62	19	32	11				1	
63	19	29.2	14				1	
64	20	32.2	16				1	
65	16	32	13				1	
66	17	28.2	12				1	1
67	19	31.2	14				1	
68	18	30	13				1	
69	18	29.2	11				1	
70	19	26.2	12				1	
71	18	27.2	13				1	
72	19	30.2	10				1	
73	19	33	14				1	
74	16	32.2	14					1
75	15	32	14					1
76	19	28	15					1
77	19	33	12					1
78	19	28.2	13					2
79	18	31.2	12					9
80	20	28.2	12					1
81	16	34	11					1

Table 22 (continued)

_N[a]	DXS10103	HPRTB	DXS10101	ASH	SEPH	NAFR	ORT	CHU
82	18	28.2	13					1
83	16	33	13					1
84	21	29.2	10					1
85	18	33	13					1
86	19	28	13					1
87	20	30.2	13					1
88	20	29.2	12					1
89	16	31	14					1
90	20	29.2	15					1
91	19	28	15					1
92	19	33	12					1
93	19	28.2	13					2
94	18	31.2	12					9
95	20	28.2	12					1
96	16	34	11					1
97	18	28.2	13					1
98	16	33	13					1
99	21	29.2	10					1
100	18	33	13					1
101	19	28	13					1
102	20	30.2	13					1
103	20	29.2	12					1
104	16	31	14					1
105	20	29.2	15					1
				0.818462 +/- 0.508416	0.803119 +/- 0.507963	0.733169 +/- 0.463537	0.774220 +/- 0.484190	0.747732 +/- 0.464359

Table 22 (continued)

_N[a]	DXS10146	DXS10134	DXS7423	ASH	SEPH	NAFR	ORT	CHU
1	36	15	44.2	2			1	
2	39.3	15	39.2	1				
3	36	15	26	1				
4	35	14	26	2		1	1	
5	34	13	30	1				
6	36	15	30	1				
7	36	15	27	1				3
8	35	15	27	4				
9	40.3	15	30	1				
10	37	15	28	1				
11	37	15	26	1		1		
12	36	14	46.2	1				
13	36	14	28	1			1	
14	42.3	17	37	1				
15	35	15	41.2	1				
16	38.3	15	30	1				
17	39	14	28	1				
18	36	15	47.2	1				
19	36	16	31	1				
20	35	15	43.2	1	1			
21	33	14	25	1				
22	34	15	26		1			
23	35	16	26		1			
24	36	14	25		1	1		
25	35	14	33		1			

Table 22 (continued)

_N[a]	DXS10146	DXS10134	DXS7423	ASH	SEPH	NAFR	ORT	CHU
26	37	16	41.2		1			
27	39	16	39.2		1			
28	35	14	44.2		1			
29	34	14	45.2		1		1	
30	44.2	14	36.2		1			
31	35	14	28		1			
32	37.3	14	28		1			
33	39	15	29		1			
34	36	14	29		1			1
35	35	15	29		1	1	1	
36	34	16	35.2		1			
37	38	16	26		1			
38	38	16	43.2		1			
39	34	15	28		1	1		
40	43.3	16	39.2			1		
41	35	16	30			1		
42	38	13	31			1		
43	33	15	44.2			2		
44	34	14	30			1		
45	42.3	17	29			1		
46	40.3	14	46.2			1		
47	37	14	29			1	2	
48	34	17	28			1		
49	36	15	43.1			1		
50	34	14	28			1		
51	35	14	27			1	1	
52	36	15	28			1		1

Table 22 (continued)

_N[a]	DXS10146	DXS10134	DXS7423	ASH	SEPH	NAFR	ORT	CHU
53	33	16	30			2		
54	36	17	44.2			1		
55	34	15	29			1		
56	37	15	43.2			1		
57	40.3	16	40.2			1		
58	35	16	41.2			1		
59	36	16	28			1		1
60	42.3	17	44.2			1		
61	34	15	30			1		
62	34	14	26				1	1
63	35	15	28				2	
64	41.3	14	29				1	
65	35	14	30				1	
66	35	16	29				1	
67	40	14	43.2				1	
68	40.3	14	28				1	
69	37	15	44.2				1	
70	38	14	26				1	
71	33	15	31				1	
72	35	14	31				1	
73	32	15	27				1	
74	34	14	27				2	1
75	34	16	44.2				2	
76	33	15	39.2				1	
77	37	14	28				1	1
78	44.3	14	28				1	
79	35	15	30				1	

Table 22 (continued)

_N[a]	DXS10146	DXS10134	DXS7423	ASH	SEPH	NAFR	ORT	CHU
80	35	16	28					1
81	31	15	42.2					1
82	39.3	14	26					2
83	33	16	29					1
84	37	16	39.2					1
85	37	17	29					1
86	39.3	14	42.2					1
87	36	16	30					2
88	38.3	16	27					1
89	36	13	25					1
90	39.3	15	42.2					1
91	34	15.1	27					1
92	34	14	31					3
93	36	15.1	24					1
94	34	15	27					2
95	36	15	25					1
96	35.3	13	33					1
97	38	15	27					1
98	36	15	29					2
99	34	16	29					1
100	36	15.1	27					1
101	37	14	30					1
102	35	16	40.2					1
103	40	14	26					1
104	35	16	33					1
105	34	16	31					1
106	39	15	27					1

Table 22 (continued)

_N[a]	DXS10146	DXS10134	DXS7423	ASH	SEPH	NAFR	ORT	CHU
107	37.3	15	26					1
108	37	16	30					1
109	35	15	42.2					1
110	38	17	29					1
111	36	15	24					1
112	34	14	31					1
113	36	15.1	24					1
114	34	15	27					2
115	36	15	25					1
116	35.3	13	33					1
117	38	15	27					1
118	36	15	29					3
119	34	16	29					1
120	36	15.1	27					1
121	37	14	30					1
122	35	16	40.2					1
123	40	14	26					1
124	35	16	33					1
125	34	16	31					1
126	39	15	27					1
127	37.3	15	26					1
128	37	16	30					1
129	35	15	42.2					1
130	38	17	29					1
131	36	15	24					1
				0.751795 +/- 0.474806	0.830409 +/- 0.521897	0.858785 +/- 0.526578	0.782430 +/- 0.488315	0.827381 +/- 0.503799

Table 23
Genetic distances between the Jewish communities and other North African and European populations (*F*st values) based on X-STRs

	Morocco	Algeria	Valencia	Balearic I.	Italy	Germany	Denmark	Somaly	Ashkenazi	Sephardic	N African	Oriental	Chuetas
Morocco	0.000												
Algeria	0.008	0.000											
Valencia	0.000	0.006	0.000										
Balearic I.	0.003	0.006	0.000	0.000									
Italy	0.007	0.004	0.006	0.004	0.000								
Germany	0.023	0.027	0.000	0.025	0.048	0.000							
Denmark	0.027	0.004	0.071	0.054	0.030	0.014	0.000						
Somaly	0.016	0.016	0.010	0.012	0.038	0.033	0.018	0.000					
Ashkenazi	0.009	0.014	0.005	0.007	0.038	0.031	0.009	0.019	0.000				
Sephardic	0.006	0.010	0.016	0.008	0.037	0.029	0.009	0.019	0.000	0.000			
N. African	0.007	0.015	0.009	0.009	0.041	0.030	0.013	0.020	0.000	0.000	0.000		
Oriental	0.024	0.024	0.024	0.023	0.050	0.039	0.027	0.028	0.010	0.006	0.011	0.000	
Chuetas	0.017	0.019	0.016	0.015	0.045	0.038	0.017	0.023	0.006	0.003	0.004	0.014	0.000

GENERAL DISCUSSION

This thesis demonstrates the complex nature of the factors that determine the population genetic structure and the genetic history of the Western Mediterranean population since their origins until now.

By examining two kind of genetic polymorphisms in the X chromosome (Alus and STRs) and comparing our biologic results with the geographic, linguistic, archeological, paleoanthropological, prehistoric and historic data of this region, we are suggesting a multidisciplinary discussion on the origins of the Western Mediterranean populations and answering a challenging question on the genetic contribution of the Arabic and Jews in this region as well as describing the evolutionary, cultural and genetic events that may explain the genetic and linguistic diversity of the actual worldwide Jewish communities and Western Mediterranean populations.

The ethno-religious origins of the Western Mediterranean region: North African and South Europeans populations

The ethnical appurtenance of an individual is related to two characteristics allowing inferring to a person (or population) a name or designation. The first is anthropological, related to the ancestors and the second is the parameter we used in our biological investigation to cluster and classify the samples: linguistic. The population origins have two aspects: (paleo)anthropological and cultural. For each one, let’s discuss the mentioned hypothesis in part I and let’s compare them and try to find a link between the emergence of the proto-Mediterranean and the introduction of the three predominant religions present within the region.

Many investigations focused on the genetic characterization of the North African and the South European population (Bosch et al. 1997; Aboukhalid et al. 2010; El Ossmani et al. 2008; El Ossmani et al. 2009; Chafik et al. 2003; Chbel et al. 2003;Robino et al. 2006 and Tomas et al. 2012; Gonzales et al. 2003 and Picornell et al. 1996). Most of

analysis underlined the proximity between the North-African and South-European populations (especially Spain), in spite of their different genetic structures. This kinship is characterized by the presence in the genetic pool of Berbers and North-African Arabic, of high frequencies of some markers that we find commonly in the European populations. Thus, the analysis of the human immunoglobulin allotypes reveals that the populations of Maghreb enclose from 66 (in the Algerian Touareg) to 92% (in Douiret-Chenini and Kesra, Tunisia) of Gm haplotypes characteristic to the European populations. The autosomal microsatellite investigations show the global genetic bridge between the North-African and South-European populations. Some mitochondrial haplogroups found in the European and Euroasian (H, HV, U*, J, T, X…) are observed between 54% (in Mozabites of Ghardaia) and 87% (in Douiret-Chenini, Tunisia). In the Moroccan population are present in a 55.64% and in Siouah in a 60%.

In our study we gave a new insight investigating two set of a reliable markers Alu and STRs in the X chromosome, that spends two-thirds of its lifetime in female subjects where the nucleotide mutation rate in much lower than in male subjects. These values explain, complete and support the observed genetic closeness between both groups as displayed in our MDS graphs. Indeed, when the frequency values of 9 X-Alu were examined in a Western Mediterranean population, it became evident that these values are in the frequency range of North African neighbors Europeans.

Alu insertion loci are biallelic; hence they have a maximum possible heterozygosity of 0.5. In the present study, the average heterozygosity for nine Alu insertions in our Western Mediterranean population was calculated as 0.137. Since heterozygosity is a measure of genetic diversity in a population, it can be concluded that Alu insertion polymorphism is considerable in our sample. Moreover, the average observed heterozygosity values catch up with the expected one; which was also proven by the presence of the Hardy-Weinberg equilibrium in most analyzed systems. Both values of heterozygosities of North African and Europeans were calculated, which are 0.165 and 0.126, respectively. These results respect to the geographic location and since in peopling of Europe there were movements through North Africa (especially Morocco) and Europe and the opposite. In order to further assess population relationships in a three dimensional space, the Multidimensional Component analysis was applied to our data set based on Reynolds genetic distance calculated. The diagram showed a clear

clustering of human populations according to their geographical locations. In both of the plots of Multidimensional Scaling Analysis, the North African population was obviously placed in the European cluster as illustrated in the respective figure (Figure 15 and 16). Another very important property of principal component analysis is that it manifests the most effective variables in differentiating the populations and explaining their variation. The analysis implied that Ya5DP62, Yb8DP49, Yd3JX437insertion polymorphisms were the most successful Alu insertion in differentiating the populations. Therefore, this result will help in the choice of Alu insertions for further evolutionary studies. When these results are compared with those of Athanasiadis et al. (2007), which ranked 13X-Alu insertions according to their power of differentiation, it is observed that the importance of Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4 and Ya5DP13 should not be underestimated.

In addition to Reynolds genetic distances, Fst analysis was performed to determine the genetic distance between North African and European. Non-significant Fst values were observed in pairs of markers, which obviously implies that North African is genetically very close to Europeans.

In the Multidimensional Scaling Analysis (MDS), the North African population was obviously placed in the European cluster as illustrated in the respective figure (Figure 18 b). Although if we focus only in Mediterranean populations, we can observed a cluster consisting of Sahrawi and Berbers populations, separated of the cluster with Arabic and Spanish population (Figure 18a). Also it is interesting to point out that, surprisingly, Ibiza and Calabria region were plotted separately in the diagram. Therefore these genetic differences can be attributed to the geographical isolation of populations and also to their history and fits with the studies of Tomas et al. (2008) and Tragelli et al. (2000).

To glean further details of the genetic structure of our populations, we added the STRUCTURE analysis. Five clusters were inferred from whole examined genotypes. The genetic patterns observed could possibly explain why the software has failed to detect population substructure in the Western Mediterranean populations, implying that there is no population-specific genetic patterns representative enough to allow us to

assign, with certainty, individuals to populations. Apparently differences do exist, but they are not striking enough to allow the definition of different clusters within the Mediterranean region. Clustering appears only when a quite distinct human group is added, such as the sub-Sahara and the Ivory Coast.

Our investigation on the X chromosomal STR leads almost to the same conclusion of a close genetic affinity and mixture occurring between South of Europe and North Africa. The 12 STRs markers included in our study were highly polymorphic as described below in the papers with no significant differences in allele distribution between the male and female subsamples. For twelve loci, we report, to our knowledge, the first population data for Western European populations including some genetic isolates dwelling on islands with a huge history like Ibiza, Minorca and Majorca. Among them, three new alleles DXS10146-36.2, DXS8378-11.2 and DXS10148-20.3 previously unobserved in worldwide populations were detected. Most of the alleles characterizing the investigated locus in the Moroccan and Algerian populations ranged within the values described for the European populations. Thus, they possibly represent a trace for the North African contribution in the genetic pool of European populations, already evidenced by autosomal chromosomes and mitochondrial DNA studies (Khodjet et al. 2005; Bosch et al. 2000; Comas et al. 2000; Harich et al. 2002; Arredi et al. 2004 and Cruciani et al. 2004). Based on the observed and expected distribution of genotypes in the female subsample, and after Bonferroni correction for multiple testing, no significant deviations from HWE were observed in these 12 X-STR systems. However, some pair of loci shows a significant LD in the European populations while the linkage equilibrium was reached in the North African populations which supports the hypothesis of the genetic bottleneck that affected only the European populations and explains that the limited extent of LD found in Africans reflects the longer evolutionary history and constant population size that shaped the genetic background of African populations.

It can be noticed from locus by locus pairwise genetic distances (F_{st}) between our samples and relevant population samples, that most of the statically significant F_{st} values were found for sub-saharan African, Asian and American populations. As can be seen from the MDS plot summarizing the relationships between our samples and other populations (Figure 16), the North African groups tightly cluster with the South

European. Also, the most shared haplotypes have been observed predominantly between the Western Mediterranean populations, whereas the other populations seem to be more diversified. The observed pattern is coherent with the results of a previous chromosome X investigation performed on a macro-geographic scale, which confirm the recent out-of-Africa expansion for human populations (Pasino et al. 2011).

Many investigations on other polymorphisms lead to the same observations. For example, Khodjet El Khil et al. (2005) suggested a relative « Mixture » between Arabic Tunisian of Jerba and the European populations. Gonzales-Peres et al. (2003) noticed a North-African and South-west European influence since the prehistoric and historic area and added that the North-African Berbers are closer to Andalusia populations than the other groups of the Iberic Peninsula. Rando et al. (1998) suggested that the limitation of genetic influx between Europe and North-Western Africa, could be due to the recent human migration. However, some studies (Bosch et al. 2000; Comas et al. 2000; Harich et al. 2002; Arredi et al. 2004 and Cruciani et al. 2004) revealed a difference between the Northern and Southern shores of the Mediterranean Sea, suggesting that the Strait of Gibraltar acted as barrier for the genetic influx.

How to explain this genetic proximity between North-Africa and South-Western Europe? The must implicit explanation implies the common origin of all these populations. Two hypotheses are discussed (Barbuhani et al. 1994 and Myles et al. 2005), the first suggest that this origins has a superior paleolithic provenance with the expansion of humans anatomically modern from the Middle East to the two shores of the Mediterranean Sea (Ferembach, 1985). The second confirms this Middle Eastern origin, but give it an earliest age: It could be occurred during the Neolithic scattering, 10.000 years BC (Ammerman et al. 1984). In our point of view we also consider these hypotheses that contribute to explain the origins of modern humans. Nowadays, most of anthropologists (population genetists) that *Homo sapiens sapiens* appears in Africa (about 150.000 to 160.000 years ago) and from there he spread in the other parts of the world, but also in Africa. This is imply that the population of North Africa and the migration of the modern human to the Middle East occurred in the same period, by individuals from the same population, who have the same genetic markers. As these frequent polymorphisms in the European and North African populations are few observed (or absent) in the Eastern and central Africa, it seems that the common genetic

origins of the North African, Middle Eastern and European populations is go back to a specific part of the ancestral population (African), which evidenced the genetic diversity in Africa since the antiquity. Thus, the polymorphism near to them found in the European population and observed in the genetic pool of the actual North African populations would be present in the ancestral population. However, their presence and/or their maintaining can be related to different human migration over time, especially the Neolithic one. The genetic exchanges could be occurred together with the spread of agriculture (model of demic diffusion). Lastly, as lately evoked for Europe (Barbujani et al. 2001), the genetic polymorphism emergence could be also done in the North African populations during the Paleolithic, Mesolithic and/or Neolithic.

Some authors suggested a genetic difference between the North African and South European populations because of the geographic barrier of Strait of Gibraltar that was also a barrier for the genetic influx. We do not support this hypothesis, because even the Strait of Gibraltar can appears as an uncrossable geographic barrier, it does not prevent the population movement. Its important to underline that the street of Gibraltar, as it is now, is due to the global warming that succeeded to the last glacial period (20.000 BC). At the glacial period, the sea's level was 135m lower than the present. This geography of the street of Gibraltar is related to the most important population's movement. Some groups spread quickly on the African and European shores between 18.000 and 9.000 years old BC before undergoing the global warming and the sea level rise on islands and shores. Our study argued for the permeability of the street of Gibraltar according to all the results discussed below. Furthermore, other studies (Coudray, 2006), showed according to the mitochondrial haplotypic frequencies (predominantly found in European populations) in the Berbers populations: Haplogroups H and V, are considered as the markers of the European populations spreading from the franco-cantabrian region after the last glacial maximum and are present today around 40% in Maghreb. In the other hand, the street of Gibraltar could not be a barrier against human migration, notably South versus North, since the history report it to the invasion and occupation during seven centuries of the Iberian Peninsula and Balearic Islands by Berbers and Arabic army leaded by the Arabic leader Tariq Ibn Zyad (Gibraltar was adapted from "Djbel Tariq"). All these events argue, justify and support the genetic mixture between the Western Mediterranean populations over both shores of the Mediterranean Sea and disjoining two great continents.

Jewish Diaspora: Origins and relationship with the Western Mediterranean populations

The Jewish state comes to an end in 70 AD, when the Romans begin to actively drive Jews from the home they had lived in for over a millennium. But the Jewish Diaspora ("diaspora" ="dispersion, scattering") had begun long before the Romans had even dreamed of Judaea. When the Assyrians conquered Israel in 722, the Hebrew inhabitants were scattered all over the Middle East; these early victims of the dispersion disappeared utterly from the pages of history. However, when Nebuchadnezzar deported the Judeans in 597 and 586 BC, he allowed them to remain in a unified community in Babylon. Another group of Judeans fled to Egypt, where they settled in the Nile delta. So from 597 onwards, there were three distinct groups of Hebrews: a group in Babylon and other parts of the Middle East, a group in Judaea, and another group in Egypt. Thus, 597 is considered the beginning date of the Jewish Diaspora. While Cyrus the Persian allowed the Judeans to return to their homeland in 538 BC, most chose to remain in Babylon. A large number of Jews in Egypt became mercenaries in Upper Egypt on an island called the Elephantine. All of these Jews retained their religion, identity, and social customs; both under the Persians and the Greeks, they were allowed to run their lives under their own laws. Some converted to other religions; still others combined the Yahweh cult with local cults; but the majority clung to the Hebraic religion and its new-found core document, the Torah.

In 63 BC, Judaea became a protectorate of Rome. Coming under the administration of a governor, Judaea was allowed a king; the governor's business was to regulate trade and maximize tax revenue. While the Jews despised the Greeks, the Romans were a nightmare. Governorships were bought at high prices; the governors would attempt to squeeze as much revenue as possible from their regions and pocket as much as they could. Even with a Jewish king, the Judeans revolted in 70 AD, a desperate revolt that ended tragically. In 73 AD, the last of the revolutionaries were holed up in a mountain fort called Masada; the Romans had besieged the fort for two years, and the 1,000 men, women, and children inside were beginning to starve. In desperation, the Jewish revolutionaries killed themselves rather than surrender to the Romans. The Romans then destroyed Jerusalem, annexed Judaea as a Roman province, and systematically drove

the Jews from Palestine. After 73 AD, Hebrew history would only be the history of the Diaspora as the Jews and their world view spread over Africa, Asia, and Europe.

Three main Jews communities branched: Ashkenazi, Sephardic (including North African), and Oriental. The Ashkenazi survived in North and Eastern Europe until their decimation by the Hitler regime, and now lives mostly in the United States and Israel. The Sephardic were exiled from Spain in 1492 and Portugal in 1497, moving to the Ottoman Empire, North Africa (joining to the Jewish populations that were settled previously in that region) and the Netherlands and the Middle Eastern or Oriental stayed in Middle East mostly in Iraq, Iran and Yemen.

The researchers proposed to answer the question whether the scattered groups of modern Jews can be identified as the descendants of the ancient Hebrews of the Bible, or whether their common ancestry has been diluted through influx of converts and through intermarriage so that little remains of their “Jewish genes”.

The complex recorded history of dispersal from the Land of Israel and subsequent residence in and movements between various countries in Europe, North Africa and the Middle East is expected to produce a complex pattern of genetic relationships among Jewish populations as well as between them and the non-Jewish peoples among whom they lived.

Many studies focused on the genetic structure and ancestry of the Jewish populations. I make mention here of the most recent.

Jews in Europe and the Middle East share inherited genes going back 3,000 years, two studies indicate (Behar et al. 2010 and Atzmon et al. 2010), refuting a hypothesis put forward by a historian in 2010 (Sand, 2010).

The genetic surveys, the first to use genome-wide scanning devices comparing Jewish communities around the world, contradict the hypothesis posed by Schlomo Sand in his book, "The Invention of the Jewish People". He had suggested Jews have no common origin but are a mix of people from Europe and Central Asia who converted to Judaism at various times.

One of the surveys, published in the current American Journal of Human Genetics, was conducted by Gil Atzmon of the Albert Einstein College of Medicine and Harry Ostrer of New York University. The other, led by Doron M. Behar of the Rambam Health Care Campus in Haifa and Richard Villems of the University of Tartu in Estonia, is published in Nature review. One of the major revelations from both surveys is the apparent genetic closeness of the two Jewish communities of Europe, the Ashkenazim and the Sephardic.

Our studies on different Jewish communities came to shed light on this concern by evaluating for the first time the genetic pool of the Jews through the X chromosome markers. Indeed our results based on the Alu insertion and X-chromosomal STRs support the first hypothesis of the Jewish common origins.

Most of markers used markers either STR or Alu resulted highly polymorphic and discriminant in these populations. Some new and unreported alleles were characteristic of these communities especially in the group of Chuetas that remains isolated in the Island of Majorca. We think that this group conserved an part of unmixed genetic pool of Jewish people, because of the inbreeding and intermarriage that occurs for almost the totality of their settlement there, as it was definitely prohibited from the indigenous people of the Island to get any kind of relationship with Majorca Jews and strictly to get married with one of them. Even when they were forced to convert to Christianity, some of them were persecuted by the Inquisition by their crypto-judaic traditions. The descendants of the Inquisition's convicts are the called Chuetas. They were discriminated by the Majorcans and mixed marriages did not take place until very present times. All these facts help them to keep a considerable proportion of their original genetic pool that track their origins compared with their neighbor Jews, although they also present some similarities with Balearic Islands populations due to the admixture. Even though it would be expected that the Sephardic and Ashkenazi had lost their genetic identity by the countless historical events they undergone, most recent studies showed the opposite. As shown in our Multidimensional Scaling analysis, the Jewish group lies clearly apart in the same cluster. Furthermore, the Admixture analyses based on the X-Alu results identified an unrecognized genetic substructure and traced the shared origins of Jewish Diaspora communities.

CONCLUSIONS & FUTURE PROSPECTS

The present thesis was a collaboration research between the Laboratory of Biochemistry and Immunology-Forensic Science Unit (Mohammed V University-Agdal) and the Laboratory of Genetics (Universidad de las Islas Baleares). This program is based on two thematic: Forensic Genetics and Population Genetics. The purpose of this thesis was to describe the genetic history of Western Mediterranean populations and the contribution of Arabs and Jews. As this biologic study can't be achieved without considering the hypothesis on their origins, history, culture and geographical situation, the interpretation of our genetic results is multidisciplinary.

Our main conclusions are the following:

- ✓ Anthological origins: The Mediterranean was a "melting pot" of many populations but was conserved compared to other regions since it was isolated at its western side, the roads connecting it to the Middle East were narrow and in the North, Europe was glacial. Indeed, the proto-Mediterranean capsiens of the Mesolithic are the most probably ancestors. They could be originated from Africa with a potential regional continuity since the Iberomaurisian.
- ✓ Cultural origins: The Western Mediterranean has been a mixing bowl for the ingredients of western civilization from Roman times and eastern civilization from the Arabic culture.
- ✓ Biological origins: Some polymorphisms that we can find in North African populations genetic structure has some polymorphisms present within certain populations in a given regions of the world but have other specific polymorphisms commonly found in South Europe. They have some kinship with Europeans while other polymorphisms would come from a Middle Eastern genetic influx that reaches also the Iberian Peninsula and Balearic Islands. The genetic diversity of the Western Mediterranean population especially the North African is very old (argued also by the Mitochondrial DNA, autosomal and Y STR and SNP) and could be originated since the Paleolithic, Mesolithic and Neolithic. Not all of the historic events let a strong genetic and demographic influence on the Western Mediterranean population. The period that mostly induced striking differentiations is the Arabic conquest. The impacts were mainly cultural but also genetic (obviously support by the Y chromosome

analysis). These populations increased and maintained a pre-established diversity.

Our investigations underlined the difficulty to obtain a clear vision on the origins of the Western Mediterranean populations and to assess their genetic diversity. This difficulty comes up from the discrepancies of some results rather than the lack of arguments. Nevertheless, we had a global view on the evolution history of the Western Mediterranean population; however the history of each local population or ethnical group should be more complicated. In the other hand we aimed in this thesis to track the genetic and historical roots of the Jewish Diaspora. While some hypothesis and earlier studies argued for the genetic admixture of the Jewish communities with their host populations, other most recent and reliable studies especially (Behar M et al. 2010)analysis on 14 Jewish groups as well as ours investigation, revealed that most Jewish populations were "genetically closer" to each other than to their non-Jewish neighbors. It also revealed genetic ties between globally dispersed Jews and non-Jewish populations in the Western Mediterranean populations. This fits with the idea that most contemporary Jews descended from ancient Hebrew and Israelite residents in the Middle Eastern region known as the Levant. It provides a trace of the Jewish Diaspora.

Our investigation has another scientific dimension: “Forensic Genetics”. The results obtained from this work, especially the STR haplotype frequencies, and gene diversity values constitute a valuable tool and a database for many interesting populations that the international scientific community can use in order to resolve some cases implying their use like the complex kinship cases.

As perspective to our work we project to increase our database of samples to have more significant results and to homogenize the sampling, including other parts of the Western Mediterranean region. In this term, we expect to collaborate with other Arabic and Middle Eastern laboratories to compare our results with a challenging pole that is the “Middle East”. That how we can shed light and get a wider and clear overview of the Western Mediterranean population genetic structure’s and to assess the contribution of the Arabic and Jews in this region. In the other hand, we are working in another part of this project that consists on the analysis of X chromosomal SNP of our DNA database samples in order to have a deeper and clearer insight of this Western Mediterranean population structure and history.

References

1. Aboukhalid R, Bouabdeallah M, Abassi M, Bentayebi K, Elmzibri M, Squalli D, Amzazi S (2010). Haplotype frequencies for 17 Y-STR loci (AmpFISTR® Y-filer™) in a Moroccan population sample. Forensic Sci. Int. Gent (4): e73-e74.
2. Abulafia D, Rackham O. The Mediterranean in history,(Thames & Hudson Ltd, 2003).
3. Abulafia D. The Mediterranean in History, (J. Paul Getty Museum, 2011).
4. Abulafia D. The Great Sea: A Human History of the Mediterranean, (Penguin Group, 2011).
5. Abulafia D (2011). Mediterranean History as Global History. History and Theory (2): 220-228
6. Aler M, Sánchez-Diz P, Gomes I, Gisbert M, Carracedo A, Amorim A, Gusmão I (2007), Genetic data of 10 X-STR in a Spanish population sample. Forensic Sci. Int. (1): 193-196.
7. Ammerman A, Cavalli-Sforza LL .The Neolithic transition and the Genetics of populations in Europe,(Princeton University Press, Princeton, 1984).
8. Ancestry.com. The Calabria Name in History, (Ancestry.com, 2007)
9. Angius A, Melis PM, Morelli L, Petretto E, Casu G, Maestrale GB, Fraumene C, Bebbere D, Forabosco P, Pirastu M. (2001). Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits. Hum Genet(109):198–209.
10. Ardlie KG, Kruglyak L, Seielstad M (2002). Patterns of linkage disequilibrium in the human genome. NatureReviewsGenetics, (3), 299–309.
11. Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, Tyler-Smith C (2004). A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. Am J Hum Genet. 75(2): 338-345.
12. Athanasiadis G, Esteban E, Via M, Dugoujon JM, Moschonas N, Chaabani H, Moral P (2007). The X chromosome Alu insertions as a tool for human population genetics: data from European and African human groups. Eur J Hum Genet, (15):578-583.

13. Atzmon G, Hao L, Pe'er I, Velez C, Pearlman A, Palamara PF, Morrow B, Friedman E, Oddoux C, Burns E, Ostrer H (2010) Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry. Am J Hum Genet. 86(6): 850-859.
14. Baer Y. A History of the Jews in Christian Spain. (Varda Books, 2001).
15. Balta P, Dana C, Dhoguois-Cohen R. La Méditerranée des Juifs: Exodes et enracinements. (Editions L'Harmattan, 2003).
16. Barbujani G, Bertorelle G (2001). Genetics and the population history of Europe. ProcNatlAcadSci U S A. 98(1): 22-25.
17. Barbujani G, Pilastro A, De Domenico S, Renfrew C (1994). Genetic variation in North Africa and Eurasia: Neolithic demic diffusion vs. Palaeolithiccolonisation. Am J PhysAnthropol. . 95(2): 137-154.
18. Batzer MA, Deninger PL (2002). *Alu* repeats and human genomic diversity. Nature Reviews, (3): 370-380.
19. Batzer MA, Kilroy GE, Richard PE, Shaikh TH, Desselle TD, Hoppens CL, Deininger PL (1990). Structure and variability of recently inserted *Alu* family members. Nucleic Acids Research, 18(23): 6793-6798.
20. Batzer MA, Deininger PL (1991). A human-specific subfamily of *Alu* sequences. Genomics, (9): 481-487.
21. Batzer MA, Rubin CM, Hellmann-Blumberg U, Alegria-Hartman M, Leeflang EP, Stern JD, Bazan HA, Shaikh TH, Deininger PL, Schmid CW (1995). Dispersion and insertion polymorphism in two small subfamilies of recently amplified human *Alu* repeats. Journal of Molecular Biology, (247): 418-427.
22. Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh, TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, Deininger PL (1994). African origin of human-specific polymorphic *Alu* insertions. Proceedings of the National Academy of Sciences, (91): 12288-12292.
23. Batzer MA, Arcot SS, Phinney JW, Alegria-Hartman M, Kass DH, Milligan SM, Kimpton C, Gill P, Hochmeister M, Ioannou PA, Herrera RJ, Boudreau DA, Scheer WD, Keats BJB, Deininger PL, Stoneking M (1996a). Genetic variation of recent *Alu* insertions in human populations. Journal of Molecular Evolution, (42): 22-29.

24. Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, Rootsi S, Chaubey G, Kutuev I, Yudkovsky G, Khusnutdinova EK, Balanovsky O, Semino O, Pereira L, Comas D, Gurwitz D, Bonne-Tamir B, Parfitt T, Hammer MF, Skorecki K, Villems R. (2010) The genome-wide structure of the Jewish people.Nature, (466): 238-242.
25. Bekada A, Benhamamouch S, Boudjema A, Fodil M, Menegon S, Torre C, Robino C (2010) Analysis of 21 Xchromosomal STRs in an Algerian population sample. Int J Legal Med (124):287–294.
26. Belkhir K, Borse P, Chikhi L, Goudet J, Bonchomme F, (2001). Genetix 4.0 windows™ Software for population genetics. Laboratoire Genome et Populations, University of Montpellier 2, Montpellier, France. http://www.univmontp2. fr/~genetix/gnetix.htm.
27. Ben-Sasson HH. A history of the Jewish People, (Harvard University Press, 1976)
28. Bentayebi K, Picornell A , Bouabdeallah M, Castro JA, Aboukhalid R, Squalli D, Misericordia MM, Amzazi S (2012). Genetic Diversity of 12 X-chromosomal short tandem repeats in the Moroccan population, Forensic Sci. Int. Gent 6(1): e48-49
29. Bigelow PA, Fogarty T. The Lure of the Mediterranean: The Ship Dweller a Story of a Happy Cruise. (Kessinger Publishing, LLC. 2005).
30. Blake E, Knapp A B. The archeology of Mediterranean prehistory. (Blackwell Publishing Ltd, 2005)
31. Bolle HJ et al. Mediterranean Land-surface Process Assessed from Space. (Springer, 2003).
32. Bonne-Tamir B, Adam A, Ticher A. Genetic Diversity among Jews. Diseases and Markers at the DNA Level. (Oxford University Press, 1992).
33. Bosch E, Calafell F, Pérez-Lezaun A, Clarimon J, Comas D, Mateu E, Martinez-Arias R, Morera B, Brakez Z, Akhayat O, Sefiani A, Hariti G, Cambon-Thomsen A, Bertranpetit J (2000). Genetic structure of north-west Africa revealed by STR analysis. Eur J Hum Genet..8(5): 360-366.

34. Bosch E, Calafell F, Pérez-Lezaun A. *et al*. (1997). A population history of Northern Africa: evidence from classical genetic markers. Hum. Biol.1997. (69): 295-311.
35. Botstein D, White RI, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet (32):314–331
36. Bouabdallah M, Ouenzar F, Aboukhalid R, D Elmzibri, .Squalli D, Amzazi S,(2008).STR data for 15 AmpFISTR Identifiler loci in the Moroccan population, Forensic Sci.Int. Genet. Suppl.(1): 306–308.
37. Bowker J. The Cambridge Illustrated History of religions. (Cambredge University Press, 2002).
38. Braunstein B. The Chuetas of Majorca. Conversos and the Inquisition of Majorca. (Ktav Publishing House, Inc., 1936).
39. Butler MJ. Fundamentals of forensic DNA typing. (Elsevier: Science&technologybooks. 2010).
40. Calliman PA, Hedges DJ, Salem AH *et al.* (2003). Comprehensive analysis of Alu-associated diversity on the human sex chromosomes. Gene. (317): 103-110.
41. Carroll ML, Roy-Engel AM, Nguyen SV, Salem AH, Vogel E, Vincent B, Myers J, Ahmad Z, Nguyen L, Sammarco M, Watkins WS, Henke J, Makalowski W, Jorde LB, Deininger PL, Batzer M (2001). Large-Scale Analysis of the *Alu*Ya5 and Yb8 Subfamilies and their contribution to human genomic diversity. Journal of Molecular Biology, (311): 17-40.
42. Casasnovas MA. Història de les Illes Balears. Palma de Mallorca. (Ed. Moll, 1998)
43. Chbel F, Pancorbo MM, Martinez-Bouzas C, Azeddoug H, Alvarez-Alvarez M, Rodriguez-Tojo M J, Nadifi S (2003). Polymorphism of six Alu insertions in Morocco: Comparative study between Arabs, berbers and Casablanca Residents. Human Genetics. 39 (10). 1398-1405.
44. Chafik A, Moundib N, Barakat A *et al*. (2003). Etude du polymorphisme moléculaire du marqueur SRY-8299 sur le chromosome Y chez des populations marocaines, arabes berbères et sahraouies. 3ème Congrès National de Génétique et Biologie Moléculaire, Tanger, pp. 120.
45. Chafik A, Moundib N, Barakat A, Rouba H. 2003. Etude du polymorphisme moléculaire du marqueur Yap sur le chromosome Y chez des populations

marocaines arabes, berbères et sahraouies. Proceeding du 26 colloque du GALF, pp. 93.

46. Chen MY, Pu CE (2004). Population data on the X chromosome short tandem repeat loci DXS10011, DXS101, DXS6789, DXS7132, DXS8377, and DXS9895 in Taiwan. Forensic Sci. Int. (146): 65-67.
47. Coffman E-L, (2005). A mosaic of people: the Jewish story and a reassessment of the DNA evidence. Cambridge. J of Gen Genealogy. (1): 193–223.
48. Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, Batzer MA, Bertranpetit J, Sajantila A (2000). Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. Hum Genet. 107(4): 312-319.
49. Comas D, Calafell F, Benchemsi N et al (2000). Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. Hum Genet. (107): 312–319.
50. Coudray C (2006). Histoire génétique et evolution des populations berbérophones nord-africaines. Université Toulouse III-Paul Sabatier. Ecole doctorale Biologie, santé, biotechnologies.
51. Crebezy E, Janin B, Midant.Reynes B. Adaima la nécropode prédynastique. Institut Français d'Archéologie Orientale, (Le Caire, 2002).
52. Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, Lavinha J, Vona G, Aman R, Cali F, Akar N, Richards M, Torroni A, Novelletto A, Scozzari R (2004). Phylogeographic analysis of haplogroup E3b (EM215) y chromosomes reveals multiple migratory events within and out of Africa. Am J HumGenet. 74(5): 1014-1022.
53. Davies B. An Introduction to the Philosophy of Religion. (OUP Oxford, 2003).
54. Darwin C. The Origin of Species by Means of Natural Selection. (John Murray, 1859)
55. Deiniger PL, Batzer MA, Hutchison CA, Edgell MH (1992). Master genes in mammalian repetitive DNA amplification. Trends in Genetics, 8(9): 307-311.
56. Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L (1998) Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene(HUMARA). J Forensic Sci (43):1046–1049.

57. Dever WG. Who were the early Israelites and where did they come from? (Eerdmans Publishing Co, Gand Rapids, 2003).
58. Donaldson CJ, Crapanzano JP, Watson JC, Levine EA, Batzer MA (2002). PROGINS *Alu*i nsertion and human genomic diversity. Mutation Research, (501): 137-141.
59. Douglas P. The Conquest of Morocco, (Frarrar, Straus and Giroux, 2005).
60. Edelmann J, Lutz-Bonengel S, Hering S (2011) X-chromosomal haplotype frequencies of four linkage groups using the Investigator Argus X-12 kit. FSI Genetics. doi:10.1016/j.fsigen.2011.01.001.
61. Edelmann J, Hering S, .Michael M, .Lessig R, Deichsel D, Meier-Sundhausen G, Roewer L, Plate I, Szibor R (2001). 16 X-chromosome STR loci frequency data from a German population. Forensic Sci. Int. (124): 215-218.
62. Edwards M. and Gibbs R (1992). A human dimorphism resulting from loss of an Alu. Genomics, (14): 590-597.
63. El Ossmani H, Talbi J, Bouchrif, B, Chafik, A (2009). Allele frequencies of 15 autosomal STR loci in the southern Morocco population with phylogenetic structure among worldwide populations, Legal Medicine. (11): 155-158.
64. El Ossmani H, Bouchrif B, Glouib K., Zaoui D., El AmriH , Chafik, A, Etude du polymorphisme des groupes sanguins, (ABO, Ss, Rhésus et Duffy) chez la population arabophone du plateau de Beni Mellal. 2008, Lebanes Sciences Journal. (9):17-28.
65. Excoffer L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances amang DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. (131): 479-491.
66. Excoffier L, Laval G, Schneider S (2005). Arlequin ver.3.0: an integrated software package for population genetics data analysis, Evol. Bioinform. Online (1): 47–50.
67. Felsenstein, J., (1993). PHYLIP (Phylogeny Inference Package), version 3.6c. Department of Genetics, University of Washington, Seattle. http://evolution.genetics.washington.edu/phylip.html
68. Ferembach D (1985). On the origin of the Iberomaurusians. A new hypothesis. J Hum Evol. (14): 393-397.
69. Frederic PM, Agnes FV, McBrewster J. Genetic Studies on Jews. (Betascript Publishing, 2011).

70. Gilmour D. The Pursuit of Italy: A History of a Land, Its Regions, and Teir Peoples, (Farrar, Straus and Giroux, 2011).
71. Goldstein D. Jacob's Legacy: A Genetic View of Jewish History. (Yale University, 2008).
72. Gomes I, Alves C, Maxzud K, Pereira R, Prata M, Sánchez-Diz P, .Carracedo A, .Amorim A, Gusmão L (2007), Analysis of 10 X-STRs in three African populations. Forensic Sci. Int. Gent. (1): 208-211.
73. Gomes I, prinz M, .Pereira R, Meyers C, Mikulasovich R, .Amorim A, .Carracedo A, L.Gusmão (2007). Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int. J. Legal Med. 121, 198-203.
74. Gonzalez-Perez E, Via M, Esteban E et al (2003). Alu insertions in the Iberian Peninsula and north west Africa – genetic boundaries or melting pot? CollAntropol. (27): 491– 500.
75. Goodwin W, Linacre A, Hadi S. An introduction to Forensic Genetics, (Willy-Blackwell, 2011).
76. Giraudi C. (2005). Middle to Late Holocene glacial variations, periglacial processes and alluvial sedimentation on the higher Apennine massifs (Italy). Quat. Res., 64(2) : 176–184.
77. Grove A, Rackham O. The Nature of Mediterranean Europe: an Ecological History. (New Haven, CT, 2001).
78. Haldane J B S (1947) The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. Ann. Eugen. 13, 262–271.
79. Hammer MF, Redd AJ, Wood ET, Bonner MR, Jarjanazi H, Karafet T, Santachiara-Benerecetti S, Oppenheim A, Jobling MA, Jenkins T, Ostrer H, Bonné-Tamir B (2000) Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. ProcNatlAcadSci USA(97): 6769-6774.
80. Hardy HG (1908). Mendelian proportions in a mixed population. Science (28):49–50.
81. Harich N, Esteban E, Chafik A, Lopez-Alomar A, Vona G, Moral P (2002). Classical polymorphisms in Berbers from Moyen Atlas (Morocco): genetics, geography and historical evidence in the Mediterranean peoples. Ann Hum Biol. 29(5): 473-487.

82. Hedman M, Palo JU, Sajantila A (2009).X-STR diversity in the Finnish and the Somali population. Forensic Sci. Int. Gent. (3): 173-178.
83. Harris EE, Hey J (1999). X chromosome evidence for ancient human histories. ProcNatlAcadSci U S A ;(96): 3320–3324.
84. Heyer E, Tremblay M (1995). Variability of the genetic contribution of Quebec population founders associated to some deleterious genes. Am J Hum Genet. 56(4): 970-978.
85. Horden, Purcell N. The corrupting Sea: a study of Mediterranean History. (Oxford, 2000).
86. Husain A, Fleming K. A Faithful Sea: The Religious Cultures of the Mediterranean. (Oxford, 2007).
87. IncSPSS (2007). Statistical package for the social sciences (Windows), version 15.0. SPSS. Inc., Chicago, Illinois.
88. Inturri S, Menegon S, Amoroso A *et al.* (2011) Linkage and linkage disequilibrium analysis of X-STRs in Italian families. FSI Genetics 5:152–154.
89. Jane S. Gerber, The Jews of Spain. A History of the Sephardic Experience*,* New (The Free Press, 1992).
90. Kang L, Li S (2006), X-chromosome STR polymorphism of Luoba Ethnic Group in Tibet (SW China), L.Kang, S.Li. Forensic Sci. Int. 156, 88-90.
91. Khodjet el Khil H, Marrakchi RT, Loueslati BY, Langaney A, Fellous M, BenAmmarElgaaied A (2005). Distribution of Y chromosome lineages in Jerba island population. Forensic Sci Int. 148(2-3): 211-218.
92. Kishida T, Wang W, Fukuda M, Tamaki Y (1997) Duplex PCR of the Y-27H39 and HPRT loci with reference to Japanese population data on the HPRT locus. NipponHoigakuZasshi .(51): 67-69.
93. Kobilinsky L, Levine L, Margolis-Nunno H. Inside Forensic Science: Forensic DNA Analysis. (Chelsea House Publishers, 2006).
94. Klug WS, Cummings MR. Concepts of Genetics (5^{th}edn Prentice Hall, 1997). Rousset F (2010). GENEPOP 4.0 for Windows and Linux.
95. Kruglyak L (1999). Prospects for whole-genome linkage disequilibrium mapping common disease genes. NatGenet, 22:139-144.
96. Krüger J, Fuhrmann W, Lichte KH, Steffens C (1968) ZurVerwendung der saurenErythrocytenphosphatasebei der Vaterschaftsbegutachtung.Dtsch Z Gerichtl Med 64:127–146.

97. Leighton R. Sicily before History: an Archeological Surveys from the Paleolithic to the Iron Age (London, 1999).
98. Li R. Forensic Biology: Identification and DNA Analysis of Biological Evidence. (CRC Press, 2008).
99. Lim EJ, Lee HY, Sim JE, Yang WI, Shin KJ (2009). Genetic Polymorphism and Haplotype Analysis of 4 Tightly Linked X-STR Duos in Koreans. CMJ. (50): 305-12.
100. Livres Groupe. Diaspora Juive: Histoire Des Juifs En Espagne, Histoire Des Juifs En Inde, Falashas, Histoire Des Juifs En Alg Rie. (Books LLC, WiliSeries, 2010).
101. Martins JA, Silva RHA, .Freschi A, .Paneto GG, Oliveira RN, .Cicarelli RMB (2010). X-chromosome genetic variation in São Paulo State (Brazil) population. Ann. of. HumanBiology. 37,598-603.
102. Martínez-Cortizas A, Pontevedra-Pombal X, García-Rodeja E, Nóvoa-Muñoz JC, Shotyk W (1999) Mercury in a Spanish peat bog: archive of climate change and atmospheric metal deposition. Science, 284, 939–942.
103. Mckenna A. The history of Northern Africa. (Britannica Educational Publishing, 2011).
104. McVean GT. Hurst LD (1997). Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. Nature 386:388–392.
105. Menocal MR. The Ornament of the World: How Muslims, Jews, and Christians Created a Culture of Tolerance in Medieval Spain. (Back Bay Books, 2009).
106. Micula G. Mallorca, Minorca and Ibiza. (Eyewitness Travel, 2004).
107. Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T (1987). Male-driven molecular evolution: a model and nucleotide sequence analysis. Cold Spring HarbSympQuantBiol; 52:863–867.
108. Montville J, Cohen MR. History as Predule: Muslims and Jews in the Medieval Mediterranean, (Lexington Books, 2011).
109. Myles S, Bouzekri N, Haverfield E, Cherkaoui M, Dugoujon JM, Ward R (2005). Genetic evidence in support of a shared European-North African dairying origin. Hum Genet. 117(1): 34-42.
110. Nasidze I, Risch GM, Robichaux M, Sherry ST, Batzer MA, Stoneking M (2001). *Alu* insertion polymorphisms and the genetic structure of human

populations from the Caucasus. European Journal of Human Genetics, (9): 267-272.

111. National Research Council. The Evaluation of Forensic DNA Evidence. (National Academy Press, 1996).

112. Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics (76):379–390.

113. Novick GE, Batzer MA, Deininger PL, Herrera RJ (1996). The mobile genetic element *Alu* in the human genome. BioScience (46): 32-40.

114. Orvietani B.S. Medieval Mediterranean Ports: the Catalan and Tuscan Coasts. (Leiden, 2001).

115. Ostrer H (2001) A genetic profile of contemporary Jewish populations. Nat Rev Genet(2): 891-898.

116. Pasino S, Caratti S, Massimiliano DP, Santovito A, Torre C, Robino C (2011). Allele and haplotype diversity of X-chromosomal STRs in Ivory Coast. Int J Legal Med. 125(7): 49–752.

117. Pawer J. The History of Jews in the Latin Kingdom of Jerusalem (Oxford, 1988).Penell C.R. Morocco From Empire to Independence. (OneWord Publications, Oxford, 2003).

118. Pereira R, Gomes I, Amorim A, Gusmão L (2007). Genetic diversity of 10 X chromosome STRs in northern Prtugal. Int. J. Legal Med. (121): 192-197.

119. Poetsch JM, Knop A, El-Mostaqim D, .Rakotomavo N, Wurmb-Schwark N (2010). Allele frequencies of 11 X-chromosomal loci of two population samples from Africa, Int. J. Legal Med. Epub ahead of print.

120. Picornell A, Gómez-Barbeito L, Tomàs C, Castro JA, Ramon MM (2005)Mitochondrial DNA HVRI variation in Balearic populations. Am J Phys Anthropol,(128):119-130.

121. Pryor J. Geography, Technology, and War: Studies in the Maritime History of the Mediterranean. (Cambridge, 1988).

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

123. Ramadan T. L'islam et le réveil arabe. (Presses du châtelet, 2011).

124. Ramadan T. The Messenger : The Meanings of the life of Mohammad. (Penguin, 2008).

125. Rando JC, Pinto F, Gonzalez AM, Hernandez M, Larruga JM, Cabrera VM, Bandelt HJ (1998).Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European near-eastern and sub-Saharan populations. Ann Hum Genet.(62): 531-550.

126. Robino C, Giolitti A, Gino S, Torre C (2006). Analysis of 12 X-chromosomal short tandem repeats in the North-West Italian population by means of two multiplex PCRs. International Congress Series. (1288): 316-318.

127. Romualdi C, Balding D, Nasidze IS, Risch G, Robichaux M, Sherry ST, Stoneking M, Batzer MA, Barbujani G (2002). Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. Genome Research, (2): 602-612.

128. Roy-Engel AM, Carrol ML, Vogel EAH, Garber RK, Nguyen SV, Salem AH, Batzer MA, Deininger PL (2001). *Alu*insertion polymorphisms for the study of human genomic diversity. Genetics Society of America, (159): 279-290.

129. Ruiz-Domenec JE. El Mediterraneo: Historia y cultura, (Peninsular Publishing Company, 2004)

130. Schaffner SF (2004). The X chromosome in population genetics. Nat Rev Genet; (5): 43– 51.

131. Sand S. The invention of the Jewish people, (Verso, First Edition edition, 2010).

132. Schilman PE, Lazarri C.R, Manrique G, (2001), Comparison of disturbance stridulations in five species of triatominae bugs. ActaTropica, 79 (2): 171-178.

133. Simoni L, Gueresi P, Pettener D *et al*. Patterns of gene flow inferred from genetic distances in the Mediterranean region. *Hum. Biol.*1999. 71: 399-415.

134. Slimane MN, Pousse H, Maatoug F, Hammami M, Ben Farhat MH (1993). Phenotypic expression of familial hypercholesterolaemia in central and southern Tunisia. Atherosclerosis. 104(1-2): 153-158.

135. Slimane MN, Lestavel S, Clavey V, Maatouk F, Ben Fahrat MH, Fruchart JC, Hammami M, Benlian P (2002). CYS127S (FH-Kairouan) and D245N (FH-Tozeur) mutations in the LDL receptor gene in Tunisian families with familial hypercholesterolaemia. J Med Genet. 39(11): e74.

136. Stoneking M, Fontius JJ, Clifford SL, Soodyall H, Arcot SS, Saha N, Jenkins T, Tahir MA, Deininger PL, Batzer MA (1997). Alu insertion polymorphisms

and human evolution: evidence for a larger population size in Africa. Genome Research, 7: 1061-1071.

137. Straten JV. The origin of Ashkenazi Jewry : The controversy Unraveled. Walter de Gruyter& Co, 2011).
138. Straus LG (1989). Age of the modern Europeans. Nature. 342: 476-477.
139. Susanne C.; Rebatoe.; Chiarelli B. Anthropologie biologique. De Boeck Université, Bruxelles, 2003).
140. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D (2003). Use of X-linked markers for forensic purposes. It J Legal Med. 117:67-74.
141. Tagarelli A, Piro A, Tagarelli G, Zinno F (2000). Color-blindness in Calabria (Southern Italy): a North-South decreasing trend. Am J Hum Biol,12:17-24.
142. "The List: The World's Fastest-Growing Religions". May 14, 2007. Retrieved 2010-05-16.
143. Thiele K, Löffler S, Löffler J *et al.* (2008) Population data of eight X-chromosomal STR markers in Ewe individuals from Ghana. FSI Genetics Supplement Series (1):167–169.
144. Thomas MG, Weale ME, Jones AL, Richards M, Smith A, Redhead N, Torroni A, Scozzari R, Gratix F, Tarekegn A, Wilson J, Bradman N, Goldstein DB (2002) Founding mothers of Jewish communitites: geographycally separated Jewish groups were independently founded by very few female ancestors. Am J Hum Genet(70): 1411-1420.
145. Tomas C, Sanchez JJ, Barbaro A, Brandt-Casadevall C, Hernandez A, Ben Dhiab M, Ramon M, Morling N, X-chromosome SNP analyses in 11 human Medi- terranean populations show a high overall genetic homogeneity except in North- west Africans (Moroccans), BMC Evol. Biol. 8 (2008) 75.
146. Tomas C, Preira V, Morling N, (2012). Analysis of 12 X-STRs in Greenlanders, Danes and Somalis using Argus X-12. Int J Legal Med. (126): 121–128.
147. Tie J, Uchigasaki S, Oshida S (2010) Genetic polymorphisms of eight X-chromosomal STR loci in the population of Japanese. FSI Genetics (4):e105–e108.
148. Valls A. (1969). Inbreeding frequencies in the Balearic Islands (Spain). Z Morphol Anthropol: (61) 343–351.

149. Weinberg W (1908) .On the demonstration of heredity in man. In: Boyer SH,trans (1963) Papers on human genetics. Prentice Hall, Englewood Cliffs, NJ.

150. Woodward J. The physical geography of the Mediterranean. (Oxford Uneversity Press, 2009).

151. www.promega.com. DNA IQ™ SystemDatabaseProtocol, Madison, USA.

152. www.qiagen.com, Investigator Argus 12 X STR Kit Handbook (2010), Madrid, S.A, Spain

153. Yotova V, Labuda D, Zietkiewicz E, Gehl D, Lovell A, Lefebvre JF, Bourgeois S, Lemieux- Blanchard E, Labuda M, Vezina H, Houde L, Tremblay M, Toupance B, Heyer E, Hudson TJ, Laberge C (2005). Anatomy of a founder effect: myotonic dystrophy in Northeastern Quebec. Hum Genet. 117(2-3): 177-187.

154. Yu N, Fu YX, Li WH (2002). DNA polymorphism in a worldwide sample of human X chromosomes. MolBiolEvol; (19): 2131– 2141.

155. Zalán A, Völgyi A, Brabetz W et al. (2008) Hungarian population data of eight X-linked markers in four linkage groups. ForensicSciInt(175):73–78.

156. Zalán A, Völgyi A, Jung M, Peterman O et al. (2007) Hungarian population data of four X-linked markers: DXS8378, DXS7132, HPRTB, and DXS7423. Int J Legal Med(121):74–77.

Published papers

Bentayebi K, Picornell A, Amzazi S. Genetic Diversity: New research: Chapter 5: Genetic History of Morocco. Novapublishers. USA. 2012.

Bentayebi K, PicornellA ,Bouabdeallah M, Castro J A, Aboukhalid R , Squalli D, Misericordia MM , Amzazi S: Genetic Diversity of 12 X-chromosomal short tandem repeats in the Moroccan population, Forensic Sci. Int. Gent. 2012. 6: e48-e49.

Bentayebi K, Ramon MM, Castro JA, Barbaro A, Aboukhalid R, Amzazi S and Picornell A. Inferring ethnicity from the X-chromosome ALU insertions: data from Western Mediterranean human groups, Forensic Sci. Int. Gent. Supp Series. 2011: e27-e28.

Picornell A, **Bentayebi K**, Bouabdeallah M, Castro JA, Bouabdellah M, Squalli D, Misericordia M, Amzazi S. X-chromosomal haplotype frequencies of four linkage groups in a North african population, *Forensic Sci. Int. Gent*. Supp Series. 2011, e19-e20.

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 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 important, les séquences microsatellites sont devenus 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

Ne pas remplir ce cadre SVP

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 ☐

III – L'IDENTIFICATION DES GRANDS PARENTS

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 :