

Final

História Concisa das Linhagens de DNA Mitocondrial de Uma População de Judeus Curdos Através da Análise da Sua Região Controlo

DISSERTAÇÃO DE MESTRADO

Catarina Jeppesen Cruz

MESTRADO EM BIOQUÍMICA APLICADA



UNIVERSIDADE da MADEIRA

A Nossa Universidade

www.uma.pt

abril | 2016

M
U His
D-R

T/14
577
CRU His
+ ED-R

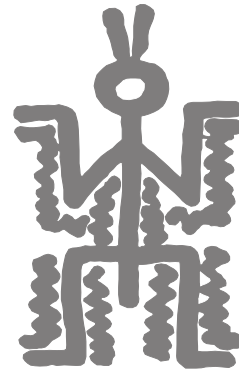
UNIVERSIDADE DA MADEIRA
BIBLIOTECA

**História Concisa das Linhagens de DNA Mitocondrial
de Uma População de Judeus Curdos Através
da Análise da Sua Região Controlo**

DISSERTAÇÃO DE MESTRADO

Catarina Jeppesen Cruz
MESTRADO EM BIOQUÍMICA APLICADA

ORIENTADOR
António Manuel Dias Brehm



Laboratório de Genética Humana

**História concisa das linhagens de DNA mitocondrial de uma
população de Judeus Curdos através da análise da sua Região
Controlo**

**Tese submetida à UNIVERSIDADE DA MADEIRA
com vista para obtenção do grau de Mestre em Bioquímica Aplicada**

Catarina Jeppesen Cruz

**Trabalho efetuado sob orientação de:
Professor Doutor António Manuel Dias Brehm**

FUNCHAL - PORTUGAL

“If you can’t fly then run, if you can’t run then walk, if you can’t walk then crawl,
but whatever you do you have to keep moving forward.”

Martin Luther King, Jr.

Agradecimentos

A elaboração desta tese não seria possível sem o apoio e orientação de várias pessoas, que ainda contribuíram para o meu desenvolvimento pessoal e acadêmico.

Em primeiro lugar, gostaria de especialmente agradecer ao meu orientador, Professor Doutor António Brehm, pela oportunidade de efetuar o trabalho de investigação no Laboratório de Genética Humana. Agradeço também pelo seu apoio, experiência e orientação durante todo este processo e para a conclusão do projeto.

Quero agradecer também a todos os membros do Laboratório pela ajuda, apoio e carinho que demonstraram durante todo este processo. Quero agradecer em especial à Sara Gomes por todo o apoio e orientação que me dedicou durante todo o trabalho de investigação e para a conclusão da dissertação.

Quero ainda agradecer à Doutora Alexandra Rosa por todo o apoio que me deu para a compreensão da parte do desenho e tratamento das árvores filogenéticas, e à Professora Manuela Gouveia pela disponibilidade do seu aparelho de PCR.

Por fim, quero agradecer à Ana Freitas e à Priscilla Figueira pela amizade e apoio durante todo o curso, e à minha família por todo o apoio e amor incondicional.

Trabalhos científicos

No decorrer do 2º ciclo de Bioquímica Aplicada foi possível desenvolver alguns trabalhos científicos paralelos a esta dissertação, dos quais resultaram em publicações e comunicações por poster. Em relação ao trabalho desenvolvido durante esta dissertação, este resultou numa comunicação oral e, posteriormente, irá servir de base para uma publicação.

Trabalho publicado:

Porto-Figueira, P., Freitas, A., **Cruz C. J.**, Figueira, J., Câmara J. S. Profiling of passion fruit volatiles: An effective tool to discriminate between species and varieties. *Food Res Int.* 2015; 77(3):408-418.

Comunicações:

➤ *EM POSTER EM CONFERÊNCIAS NACIONAIS E INTERNACIONAIS:*

Porto-Figueira, P., Freitas, A., **Cruz, C. J.**, Figueira, J. A., Câmara, J. S. GC-MS on Profiling Passion Fruit Volatiles. An Effective Tool to Discriminate Between Species and Varieties. 9th National Meeting of Chromatography and XVI Colacro, Lisboa, Portugal, Janeiro 2016.

Porto-Figueira, P., Freitas, A., **Cruz, C. J.**, Figueira, J. A., Câmara, J. S. Discrimination of different *Passiflora* L. varieties based on Volatomic Profile and Multivariate Analysis. 18th EuroAnalysis, Bordéus, França, Setembro 2015.

Porto-Figueira, P., Freitas, A., **Cruz, C. J.**, Figueira, J. A., Câmara, J. S. Differentiation of the volatomic pattern of several *Passiflora* L. 12th Meeting of Food Chemistry, Lisboa, Portugal, Setembro 2014. (ISBN: 978-989-98541-5-4).

Porto-Figueira, P., Freitas, A., **Cruz, C. J.**, Figueira, J. A., Câmara, J. S. Screening of the volatomic profile of different *Passiflora* L. species through headspace solid phase microextraction tandem with gas chromatography-mass spectrometry analysis. Riva 2014, 40th International Symposium of Capillary Chromatography and 13th GC x GC Symposium, Riva del Garda, Itália, Maio 2014.

➤ *ORAIS EM CONFERÊNCIAS NACIONAIS:*

Cruz, C. J., Gomes, S. C., Rosa, A., Brehm, A. A concise history of Kurdish Jewish mtDNA lineages through the assemblage of HVS-I, -II and -III gene sequences. 3rd CQM Annual Meeting, Funchal, Portugal, Abril 2016, página 31.

Resumo

O povo curdo, oriundo do Médio Oriente, é constituído por cerca de 36 milhões de indivíduos, dispersos por quatro países diferentes: a Turquia, o Irão, o Iraque e a Síria. Não obstante as suas diversas origens étnicas, a linguagem e a cultura estão intimamente relacionadas com a população persa (correspondente à região do atual Irão). Neste estudo foi traçado o perfil genético da linhagem materna desta população de que professa a religião judaica através da análise da região controlo do DNA mitocondrial (DNAMt) e, posteriormente, comparado com outras populações da Europa, de África e do Médio Oriente, bem como com demais populações judaicas. Identificou-se uma elevada diversidade genética com a prevalência dos haplogrupos mitocondriais H, J1 e N1, comuns do Médio Oriente. A pesquisa por uma possível relação genética entre esta população de judeus curdos com outras populações relevantes discriminadas na literatura apontou para uma relação mais próxima com as populações da Bulgária, do Irão e do Azerbaijão e com os judeus da mesma região do que com as demais populações do Médio Oriente. Os presentes resultados sugerem que o povo curdo judeu conseguiu manter ao longo do tempo um certo isolamento genético relativamente às influências de populações circundantes.

Palavras-chave: DNA mitocondrial; Região controlo; Haplogrupos mitocondriais; Judeus curdos; História materna.

Abstract

The Kurdish people, originating from the Middle East, is composed of 36 million people scattered across four different countries: Turkey, Iran, Iraq and Syria. Despite their diverse ethnic origins, language and culture are closely related to the Persian population (corresponding to present-day Iran). In this study, the genetic profile of the maternal lineage of Kurdish Jews was investigated through the analysis of the control region of mitochondrial DNA (mtDNA), and then compared with other populations from Europe, Africa and Middle East, as well as other Jewish populations. We identified a high genetic diversity with the prevalence of mitochondrial haplogroups H, J1 and N1, all common in the Middle East. The search for a possible genetic relationship between this population of Kurdish Jews and other relevant populations specified in the literature pointed to a closer relationship to populations from Bulgaria, Iran and Azerbaijan and with the Jews in the same region than with other peoples from the Middle East. The present results suggest that the Kurdish Jewish people, while subject of dispute by other populations, remained a certain genetic isolation from the influences of surrounding populations.

Keywords: Mitochondrial DNA; Control region; Mitochondrial haplogroups; Kurdish Jews; Maternal history.

Índice

AGRADECIMENTOS	5
TRABALHOS CIENTÍFICOS	7
RESUMO	9
ABSTRACT	10
LISTA DE FIGURAS	13
LISTA DE TABELAS	14
LISTA DO MATERIAL SUPLEMENTAR	14
LISTA DE ABREVIATURAS	15
CAPÍTULO 1 – INTRODUÇÃO	17
1.1. DNA MITOCONDRIAL (DNAMT).....	18
1.1.1. Particularidades do DNAMt	20
1.1.1.1. Elevado número de cópias.....	20
1.1.1.2. Hereditariedade materna.....	22
1.1.1.3. Homoplasmia vs Heteroplasmia.....	23
1.1.1.4. Ausência de recombinação	23
1.1.1.5. Elevada taxa de mutação	24
1.1.2. DNAMt: potencial como marcador genético	25
1.1.3. Haplogrupos mitocondriais	25
1.2. OS CURDOS	32
1.2.1. A história do seu nome.....	32
1.2.2. A origem dos curdos	33
1.2.2.1. As conquistas muçulmanas	34
1.2.2.2. As dinastias no Curdistão.....	34
1.2.2.3. A invasão dos mongóis	35
1.2.2.4. Os curdos a partir do século XVI	35
1.2.2.5. Os curdos após a 1ª Guerra Mundial	36
1.2.3. Religião no Curdistão.....	37
1.3. OBJETIVOS DO ESTUDO	38
CAPÍTULO 2 – FUNDAMENTOS TEÓRICOS DAS TÉCNICAS EXPERIMENTAIS	41
2.1. EXTRAÇÃO POR CHELEX.....	42
2.2. REAÇÃO EM CADEIA DE POLIMERASE (PCR).....	43
2.3. ELETROFORESE EM GEL	45
2.4. SEQUENCIACÃO PELO MÉTODO DE SANGER	46
CAPÍTULO 3 – MATERIAL E METODOLOGIA	49
3.1. COLEÇÃO DA AMOSTRA POPULACIONAL	50
3.2. EXTRAÇÃO DE DNA	50

3.3.	AMPLIFICAÇÃO POR PCR DA REGIÃO DE CONTROLO DO DNAMT	50
3.4.	PURIFICAÇÃO DO DNAMT AMPLIFICADO.....	51
3.5.	SEQUENCIAÇÃO AUTOMÁTICA DA REGIÃO DE CONTROLO DO DNAMT	51
3.6.	ANÁLISE DA REGIÃO DE CONTROLO	52
3.7.	CONSTRUÇÃO DAS ÁRVORES FILOGENÉTICAS	53
3.8.	ANÁLISE ESTATÍSTICA	54
CAPÍTULO 4 – RESULTADOS E DISCUSSÃO		55
4.1.	ANÁLISE DO DNAMT NUMA POPULAÇÃO DE JUDEUS DO CURDISTÃO.....	56
4.1.1.	Árvores filogenéticas dos judeus curdos	58
4.1.1.1.	Macrohaplogrupo H	59
4.1.1.2.	Macrohaplogrupo J.....	62
4.1.1.3.	Macrohaplogrupo K	65
4.1.1.4.	Macrohaplogrupo M.....	67
4.1.1.5.	Macrohaplogrupo N	68
4.1.1.6.	Macrohaplogrupo R.....	69
4.1.1.7.	Macrohaplogrupo T.....	70
4.1.1.8.	Macrohaplogrupo U	72
4.1.1.9.	Macrohaplogrupo X	74
4.2.	ANÁLISE DA VARIÂNCIA MOLECULAR E DISTÂNCIAS GENÉTICAS	75
4.3.	ANÁLISE DAS COMPONENTES PRINCIPAIS.....	77
CAPÍTULO 5 – CONCLUSÕES.....		83
REFERÊNCIAS		87
MATERIAL SUPLEMENTAR		97

Lista de Figuras

Figura 1.1: Mapa do DNA mitocondrial humano.....	19
Figura 1.2: Esquema da estrutura do D-loop.....	19
Figura 1.3: Segregação do DNAMt durante a oogénese e a embriogénese.....	21
Figura 1.4: Árvore Filogenética simplificada dos haplogrupos do DNAMt.....	26
Figura 1.5: Esquema do padrão das migrações humanas mundiais.....	28
Figura 1.6: Árvore dos haplogrupos mitocondriais e respetivas distribuições.....	31
Figura 1.7: Mapa da região do Curdistão.....	32
Figura 2.1: Representação esquemática do processo de amplificação do DNA.....	45
Figura 2.2: Representação esquemática do princípio eletroforético.....	46
Figura 2.3: Representação esquemática do processo de Sequenciação de Sanger...	47
Figura 4.1: Árvore Filogenética do macrohaplogrupo H.....	62
Figura 4.2: Árvore Filogenética do macrohaplogrupo J.....	63
Figura 4.3: Árvore Filogenética do macrohaplogrupo K.....	66
Figura 4.4: Árvore Filogenética do haplótipo M1a1.....	67
Figura 4.5: Árvore Filogenética do macrohaplogrupo N.....	69
Figura 4.6: Árvore Filogenética do haplótipo R0a1a.....	70
Figura 4.7: Árvore Filogenética do macrohaplogrupo T.....	71
Figura 4.8: Árvore Filogenética do macrohaplogrupo U.....	73
Figura 4.9: Árvore Filogenética do macrohaplogrupo X.....	75
Figura 4.10: PCA obtida para as relações dos Judeus Curdos com outras populações que possam ter tido alguma influência sobre o Curdistão.....	78
Figura 4.11: PCA obtida para as relações dos Judeus Curdos com outras populações judaicas.....	78

Lista de Tabelas

Tabela 3.1: Fragmentos amplificados da região controlo do DNAm e respetivos primers utilizados.....	51
Tabela 3.2: Posições utilizadas na confirmação dos haplogrupos.....	53
Tabela 4.1: Frequências dos macrohaplogrupos classificados na população de Judeus Curdos.....	56
Tabela 4.2: Parâmetros estatísticos obtidos a partir da análise da região de controlo...	57
Tabela 4.3: AMOVA das populações judaicas, não judaicas e de todas as populações	76

Lista do Material Suplementar

Tabela S.1: Populações incluídas no estudo.....	97
Tabela S.2: Haplogrupos classificados para cada uma das amostras	98
Tabela S.3: Polimorfismos identificados na região codificante	100
Tabela S.4: Número de correspondências exatas distribuídos por país	102
Tabela S.5: Número de haplótipos parecidos distribuídos por país	104
Tabela S.6: Índices F_{ST} obtidos para as populações que tiveram alguma influência sobre o Curdistão em estudo	106
Tabela S.7: Índices F_{ST} obtidos para as populações judaicas em estudo.....	107

Lista de Abreviaturas

°C	Graus Celsius
μM	Micromolar (μmol/L)
A	Adenina
a.C.	Antes de Cristo
DNA	Ácido Desoxirribonucleico
DNAmt/mtDNA	DNA Mitocondrial (<i>Mitochondrial DNA</i>)
AMOVA	Análise de Variância Molecular
RNA	Ácido Ribonucleico
RNAr	Ácido Ribonucleico Ribossomal
RNA _t	Ácido Ribonucleico de Transferência
ATP	Adenosina Trifosfato
BSA	Albumina do Soro Bovino (<i>Bovine Serum Albumin</i>)
C	Citosina
cm	Centímetro
CRS	Sequência de Referência de Cambridge (<i>Cambridge Reference Sequence</i>)
d.C.	Depois de Cristo
ddNTP	Didesoxinucleótido Trifosfato (<i>Dideoxynucleotide Triphosphate</i>)
D-loop	Displacement Loop
dNTP	Desoxinucleótido Trifosfato (<i>Deoxynucleotide Triphosphate</i>)
EDTA	Ácido Etilenodiamino Tetra-acético
Fe ²⁺	Ião Ferro
F _{ST}	Estatística F (F-statistics)
FW	<i>Forward</i> (o que inicia a cadeia leve)
G	Guanina
H	Hidrogênio
H ₂ O	Água
HAM	Homem Anatomicamente Moderno
HCl	Ácido Clorídrico
HS	Cadeia pesada (<i>Heavy strand</i>)
HSP	Promotor da Cadeia Pesada
HVS	Segmento Hipervariável
kb	Quilobase (1 kb = 1000 pb)
km	Quilômetros
LHON	Neuropatia Ótica Hereditária de Leber (<i>Leber's Hereditary Optic Neuropathy</i>)
LPE	Extração em Fase Líquida
LS	Cadeia leve (<i>Light strand</i>)
LSP	Promotor da Cadeia Leve
M	Molar (mol/L)
mg	Miligrama
Mg ²⁺	Ião Magnésio
MgCl ₂	Cloreto de Magnésio
MJ	Algoritmo <i>Median-Joining</i>
mL	Mililitro

mM	Milimolar (mmol/L)
MRCA	Ancestral Comum Mais Recente (<i>Most Recent Common Ancestor</i>)
mtSSB	Proteína de ligação da cadeia simples de DNAm (<i>Mitochondrial Single-Stranded DNA binding protein</i>)
NaAc	Acetato de Sódio
NaCl	Cloreto de Sódio
OH	Ião Hidroxilo
OXPPOS	Fosforilação Oxidativa
pb	Pares de base
PCA	Análise de Componentes Principais
PCR	Reação em Cadeia de Polimerase
pH	Potencial de Hidrogénio (Acidez de uma solução)
pK	Constante de Dissociação
PKK	<i>Kurdistan Worker's Party</i>
pmol	Picomol
RCLB	Tampão de Lise das Células Vermelhas (<i>Red Cell Lysis Buffer</i>)
rCRS	Sequência de Referência de Cambridge revista (<i>Revised Cambridge Reference Sequence</i>)
RM	Algoritmo <i>Reduced-Median</i>
rpm	Rotações por minuto
RV	<i>Reverse</i> (o que inicia a cadeia pesada)
SDS	Dodecilsulfato de Sódio (<i>Sodium Dodecyl Sulfate</i>)
SNP	Polimorfismo de Nucleotídeo Único (<i>Single-Nucleotide Polymorphism</i>)
SPE	Extração em Fase Sólida
STR	Microssatélite (<i>Short Tandem Repeat</i>)
T	Timina
T _a	Temperatura de <i>annealing</i>
U	Unidade
μL	Microlitro

Capítulo 1 – Introdução



1.1. DNA mitocondrial (DNAMt)

A mitocôndria é um organelo com dupla membrana localizado no citosol das células eucarióticas tendo, provavelmente, resultado de uma relação endossimbiótica entre uma célula eucariótica primitiva sem funções respiratórias e uma α -proteobactéria há cerca de 1500 milhões de anos [1, 2]. A mitocôndria é um organelo altamente especializado, responsável principalmente pela produção de adenosina trifosfato (ATP) utilizada pela célula, através da fosforilação oxidativa (OXPHOS) – biossíntese de ATP [2, 3]. Para além desta principal função, participa, igualmente, na homeostasia celular do cálcio, na regulação do metabolismo celular, na biossíntese do grupo prostético heme (cofator que consiste num ião Fe^{2+} contido no centro de um anel orgânico que existe na hemoglobina, por exemplo) e desempenha um papel fundamental na morte programada celular [2-6].

As mitocôndrias possuem um genoma próprio, o DNA mitocondrial (DNAMt – Figura 1.1), que nos mamíferos surge sob a forma circular de dupla cadeia, com cadeias assimétricas a nível da sua constituição, uma cadeia pesada (*heavy strand* – *HS*) e cadeia leve (*light strand* – *LS*). A cadeia pesada possui uma percentagem de guaninas superior à outra cadeia. A molécula de DNAMt apresenta dimensões reduzidas, 16569 pares de base (pb), o que a torna compacta e, praticamente, sem regiões intergénicas.

O DNAMt encontra-se dividido em duas partes principais consoante a sua função: a região codificante com cerca de 15.5 kb e a região não codificante ou região de controlo com, aproximadamente, 1.1 kb.

Na região codificante estão presentes 37 genes codificados pelo DNAMt, incluindo 13 polipéptidos necessários na OXPHOS, e 2 RNAt e 2 RNAr essenciais à síntese proteica intramitocondrial [7-9].

A região não codificante, ou de controlo, contém os promotores de transcrição das cadeias *HS* e *LS* (*HSP* e *LSP* respetivamente) e a origem da replicação da cadeia pesada (O_H) [10]. Já a origem de replicação da cadeia leve surge na região codificante num agrupamento de 5 RNAt. A região controlo por ser uma zona hipervariável, com polimorfismos que se diferenciam entre indivíduos, tornou-se alvo de variados estudos populacionais. Esta região está subdividida em três segmentos hipervariáveis, HVS-I (16024-16365), HVS-II (57-372) e HVS-III (438-576) [11-14].

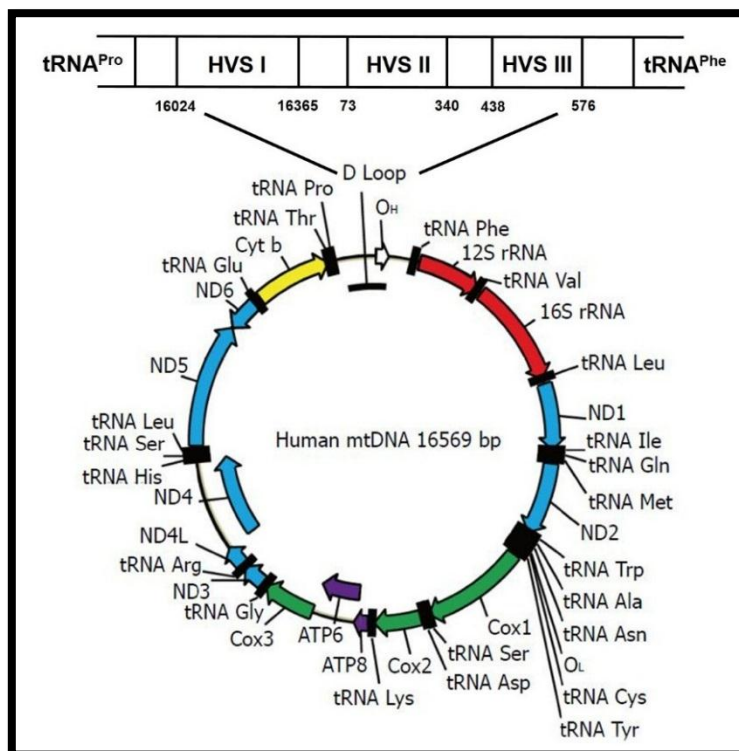


Figura 1.1: Mapa do DNA mitocondrial humano. (HVS I-III: Segmentos hipervariáveis I a III; O_H e O_L: origens de replicação da cadeia pesada e leve, respectivamente; ND1-ND6: Subunidades da NADH desidrogenase; COX1-COX3: Subunidades da citocromo oxidase; ATP6 e ATP8: Subunidades 6 e 8 da ATPase mitocondrial; Cyt b: Citocromo.) (Adaptado de Shokolenko [7].)

Uma grande parte da região de controlo incorpora, ainda, uma terceira cadeia de DNA com cerca de 650 pb, denominada de *displacement loop* (D-loop – Figura 1.2). Esta terceira cadeia intitula-se por 7S DNA e não está presente em todas as moléculas de DNAm_t [15].

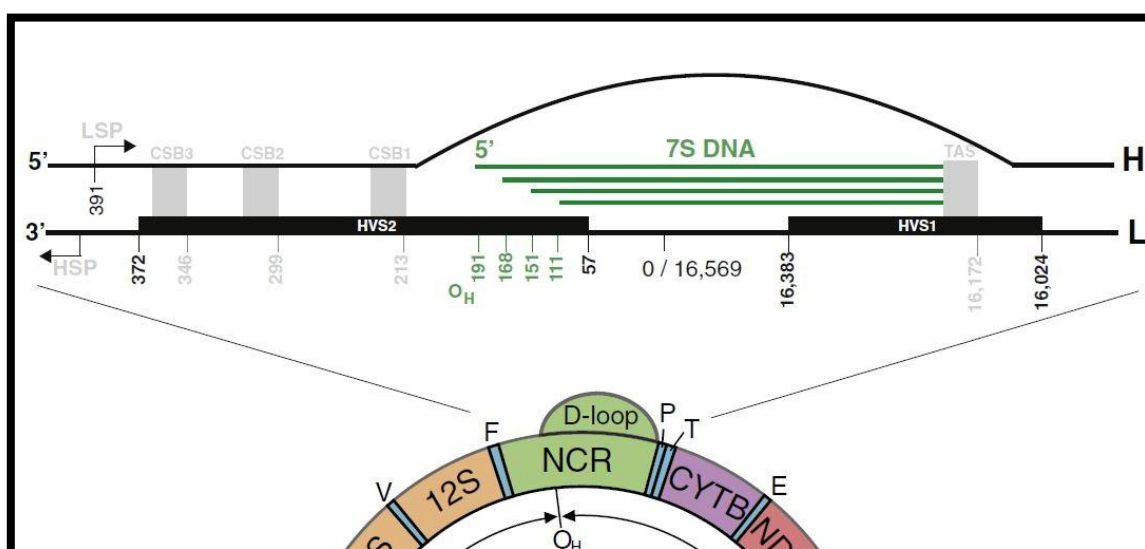


Figura 1.2: Esquema da estrutura do D-loop, com indicação das principais extremidades 5' de 7S DNA. (LSP e HSP: Promotores das cadeias leve e pesada, respectivamente; H e L: cadeias pesada e leve, respectivamente; CSB: Bloco de sequência conservada; TAS: Sequências de terminação associadas.) (Adaptado de Nicholls [15].)

O DNAm foi sequenciado, pela primeira vez, na sua totalidade e publicado em 1981, por Anderson e seus colegas [9], surgindo assim a Sequência de Referência de Cambridge (*Cambridge Reference Sequence, CRS*). Em 1999, Andrews e seus colaboradores [16] reviram a sequência original, apresentando uma nova versão, referenciada como Sequência de Referência de Cambridge revista (*revised Cambridge Reference Sequence, rCRS*), utilizada, atualmente, como base universal de comparação.

1.1.1. Particularidades do DNAm

O DNAm possui características particulares que o tornaram numa molécula de destaque em estudos da história e evolução da população humana. Destacam-se a herança matrilinear, a ausência de recombinação, uma taxa de mutação considerável, um elevado número de cópias e a homoplasia. Existem ainda associações deste genoma a patologias humanas (como a doença de *Alzheimer*, doença de *Parkinson* e diabetes, entre outros), envelhecimento e cancro [17, 18].

1.1.1.1. Elevado número de cópias

As células animais são características por possuírem um elevado número de cópias de moléculas de DNAm, variando o seu número consoante o tipo de célula. Nas células somáticas, cada mitocôndria possui entre uma a 15 cópias da molécula de DNAm e cada célula pode conter entre 1.6 a 4.1×10^3 cópias, dependendo da quantidade de mitocôndrias presente em cada célula. Quanto às células germinativas, um ócito maturo pode conter entre 100 a 200×10^3 cópias com uma ou duas cópias por mitocôndria, ao passo que um espermatozoide maturo pode conter 100 cópias, aproximadamente, com 1 cópia por mitocôndria [19-21]. As moléculas de DNAm dentro das mitocôndrias estão agrupadas em nucleoides, estruturas estáveis e dinâmicas que dividem e acompanham a dinâmica da rede mitocondrial, provavelmente, com o objetivo de assegurar a transmissão do DNAm para a descendência durante a divisão e crescimento da mitocôndria [22]. Cada nucleóide é então constituído por proteínas associadas ao DNAm (por exemplo, a proteína de ligação da cadeia simples de DNAm – *Mitochondrial Single-Stranded DNA binding protein, mtSSB*), *Twinkle* (presumível helicase do DNAm), fatores de replicação,

transcrição e processamento de RNA e por duas a dez cópias de moléculas de DNAm [22-24].

Atualmente, sabe-se que existe um controlo rigoroso da quantidade de DNAm durante a oogenese e a embriogénese. A este controlo denomina-se de efeito “*bottleneck*”, em que o número de moléculas de DNAm é reduzido a cerca de 100. É então seguido por uma sobre-replicação que retorna a quantidade de DNAm por célula. Deste modo, é também prevenida a acumulação de mutações (modificações nucleotídicas presentes numa sequência genética) patogénicas [21, 25]. Os mecanismos envolvidos neste processo não estão ainda bem estudados, mas sabe-se que ocorrem dois efeitos “*bottlenecks*”, um durante a oogenese e outro durante a embriogénese (ver Figura 1.3). Na oogenese existe um decréscimo na quantidade de DNAm e durante a maturação um aumento extraordinário, que se finaliza durante a metáfase II, provavelmente para a preparação das necessidades metabólicas da fertilização e implantação. O segundo efeito acontece no início da embriogénese, onde não ocorre replicação do DNAm entre o zigoto e o blastocisto e ao mesmo tempo sucede uma rápida divisão celular. A replicação é retomada apenas mais tarde no desenvolvimento embrionário, atingindo a quantidade já referida, anteriormente, para cada célula [21, 26].

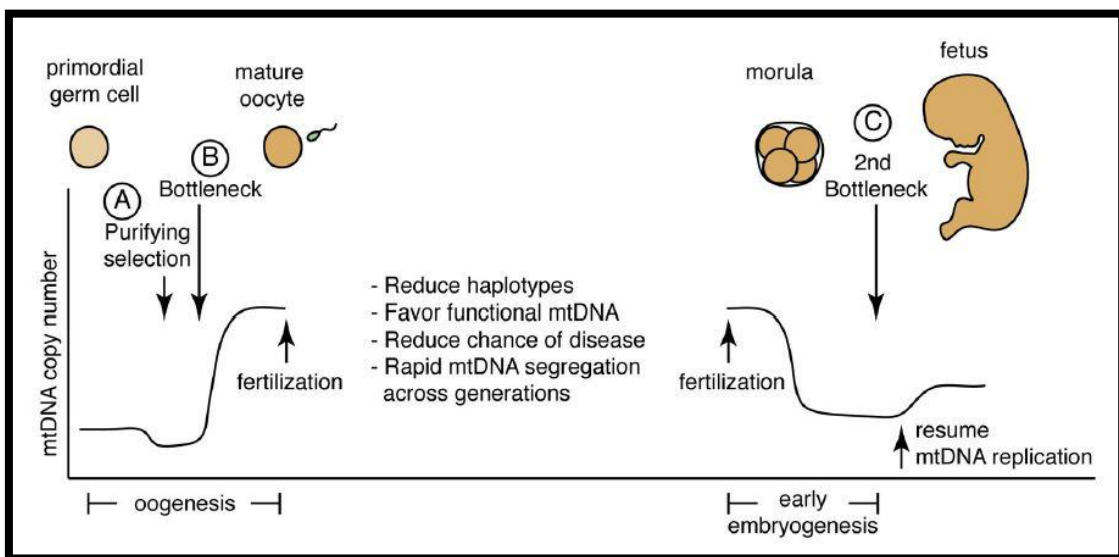


Figura 1.3: Segregação do DNAm durante a oogenese e a embriogénese, com os respetivos efeitos “*bottleneck*” e respetiva quantidade relativa de cópias de DNAm. (A – Seleção por purificação (contribuição na diminuição da transmissão de mutações prejudiciais à descendência); B – Primeiro efeito “*bottleneck*”; C – Segundo efeito “*bottleneck*”). (Retirado de Mishra [26])

1.1.1.2. Hereditariedade materna

O DNAm_t é herdado de maneira não-Mendeliana sendo aceite que, na maioria dos animais, este é herdado das mitocôndrias existentes no citoplasma do oócito em que o embrião se desenvolve [27]. A herança é considerada uniparental, ou seja, provém da linha germinativa materna [28-30], sendo pela primeira vez descoberta em humanos em 1980 [28].

Foram propostas duas hipóteses que tentam explicar esta herança uniparental nos animais [27, 31]. A primeira afirma que este sistema desenvolveu-se com o intuito de excluir o DNAm_t proveniente dos espermatozoides. Uma vez que o DNAm_t nos espermatozoides é vulnerável a danos oxidativos, tornando-os propensos a carregar mutações prejudiciais à descendência, a herança uniparental via linhagem materna vem, deste modo, impedir a sua transmissão [32, 33]. A segunda hipótese afirma que a presença de haplótipos geneticamente diferentes de DNAm_t pode ser desfavorável, já que pode permitir a seleção natural tanto à custa de indivíduos isolados, como, posteriormente, ao nível populacional [31].

Alguns estudos expõem a presença do DNAm_t paterno durante a fertilização até a fase da mórula, vindo a desaparecer após a mesma [27, 34]. Alguns mecanismos, ainda que mal entendidos, foram expostos com o intuito de explicar o que sucede ao DNAm_t da linhagem paterna [27, 31, 35, 36]:

- Diluição do DNAm_t paterno pelo materno devido à diferença de quantidade de cada um;
- Marcação seletiva através da ubiquitinação (ubiquitina é uma proteína que se liga a outras proteínas marcando-as para posterior degradação pela 26S proteassoma) e posterior destruição das mitocôndrias presentes nos espermatozoides;
- Ocorrência do efeito “*bottleneck*” genético capaz de excluir pequenas contribuições alélicas, como por exemplo o DNAm_t paterno.

Em Agosto de 2002, foi reportado um caso, em humanos, de hereditariedade paterna [37]. Um indivíduo de 28 anos possuía um severo distúrbio metabólico que o tornava intolerante a qualquer exercício físico. O seu DNAm_t foi sequenciado e foram encontrados dois haplótipos distintos apenas nas células musculares. Após a sequenciação do DNAm_t da mãe, do pai, da irmã e do tio, concluiu-se que o indivíduo em questão

possuía 90% do DNAm de origem paterna nas células musculares. Possuía, ainda, uma deleção de dois pb no gene ND2 responsável pela codificação de uma enzima essencial à respiração da mitocôndria [37, 38].

1.1.1.3. Homoplasmia vs Heteroplasmia

Como previamente referido, as células contêm milhares de moléculas de DNAm e acreditava-se que todos os tecidos possuíssem moléculas idênticas, estado que se denomina de homoplasmia [39]. No entanto, já em 1996, Bendall [40] levantou a possibilidade de o estado de heteroplasmia, ocorrência de mais do que um tipo de molécula de DNAm num indivíduo ou tecido, ser mais comum que o previsto [21, 40]. Este estado tem sido, frequentemente, associado apenas a doenças, como a Neuropatia Óptica Hereditária de Leber (*Leber's Hereditary Optic Neuropathy – LHON*) ou a intolerância ao exercício (exemplo referido no ponto 1.2.1.2) [38, 41, 42].

No entanto, estudos recentes têm vindo a demonstrar que indivíduos saudáveis também possuem estados de heteroplasmia. No estudo de Naue [43], e devido ao avanço da tecnologia que permitiu o aumento da sensibilidade dos métodos através da sequenciação, demonstrou que 88% dos indivíduos possuíssem, no mínimo, uma posição heteroplásmica [43, 44].

No estado de homoplasmia podemos encontrar dois tipos de polimorfismos (mutações não prejudiciais com ocorrem com alguma frequência, cerca de 1%, numa determinada população): de tamanho e de sequência. Polimorfismos de tamanho referem-se ao facto das sequências diferirem no tamanho, através de duplicações, adições ou deleções. Polimorfismos de sequência referem-se a alterações pontuais por transições ou transversões, designadas por polimorfismos de nucleótido único (*Single-Nucleotide Polymorphism – SNP*) que serão abordados mais detalhadamente mais abaixo [44].

1.1.1.4. Ausência de recombinação

O DNAm não recombina com os restantes fragmentos genómicos como ocorre com os cromossomas nucleares [45, 46]. Porém, esta ausência de recombinação nos mamíferos já foi questionada inúmeras vezes nos últimos anos [47]. Numa revisão realizada em 2004 por Piganeau e Eyre-Walker[48], estudos anteriores foram reunidos

cujos dados foram reanalisados e tratados sob novos testes de homoplasmia, verificando a relação entre desequilíbrio de ligação e a distância entre sítios. No entanto, concluíram que não havia suficiente evidência de recombinação e, caso houvesse de facto recombinação, seria em níveis muito baixos, cujos testes utilizados não seriam suficientemente sensíveis para os detetar [48, 49].

Ainda assim, a mitocôndria possui toda a maquinaria necessária à recombinação [50]. Estudos recentes confirmam que este organelo possui um mecanismo de reparação do seu genoma através de recombinação homóloga, que muitas vezes procede a quebra da dupla cadeia (para uma leitura mais pormenorizada [51, 52]). Na possibilidade de se provar que a recombinação é real, sentir-se-á o impacto não só na compreensão da etiologia e transmissão de doenças mitocondriais, bem como na história da evolução do homem [50, 53].

1.1.1.5. Elevada taxa de mutação

O DNAm_t apresenta uma taxa de mutação mais elevada que a calculada para o genoma nuclear: cerca de cinco a 100 vezes maior [54-56]. Esta discrepância de taxas de mutação podem ser atribuídas a alguns fatores, de entre os quais a falta de mecanismos de reparação eficientes e de um sistema de proteínas protetoras como as histonas presentes no genoma nuclear; a grande exposição aos radicais livres de oxigénio, provenientes da via metabólica de OXPHOS que aumentam o dano oxidativo no DNAm_t; e a presença de um sistema de replicação rápido propenso a erros [54, 57, 58].

Dentro da própria molécula, as taxas de mutação variam ao longo das regiões do DNAm_t [55, 59, 60]. A região controlo possui uma taxa de mutação maior que a região codificante, por não estar sujeita à reparação genómica uma vez que esta não codifica proteínas. Simultaneamente, as transições são favorecidas às transversões [54, 61]. A taxa das mutações sinónimas é relativamente uniforme ao longo do genoma mitocondrial, enquanto que a das mutações não-sinónimas varia consideravelmente entre os genes, dependendo das restrições funcionais [55]. Mutações sinónimas e RNAr evoluem cerca de 20 vezes mais que os seus equivalentes no genoma nuclear e os RNAt evoluem cerca de 100 vezes mais [54, 55].

1.1.2. DNAMt: potencial como marcador genético

Devido às suas características individuais, o DNAMt é usado desde há muito para investigar a origem, evolução e migração de populações humanas (filogenética e filogeografia) [62, 63], sendo também utilizado em medicina forense [64, 65] e medicina clínica [3, 17].

A ocorrência de mutações é a base na variação de cada indivíduo. Assim, o conjunto destas modificações irá determinar o perfil genético individual. Dado que o DNAMt acumula alterações nucleotídicas de forma mais rápida que o DNA nuclear, a sequência de um indivíduo irá se diferenciar de outro, em média, em 25 pb [63]. As mutações podem ocorrer de forma espontânea, através de erros na replicação, ou de forma induzida, com agentes físicos ou químicos, e apenas estas são transmitidas à descendência [66, 67].

A maioria dos estudos filogenéticos utiliza os polimorfismos como marcadores ancestrais: polimorfismos de nucleótido único (*SNP – Single Nucleotide Polymorphism*) e microssatélites (*STR – Short Tandem Repeat*) [68].

Os *STRs* são pequenas repetições de 3-7 nucleótidos, que se distinguem entre si pelo número de repetições presentes em cada *locus*. Estas repetições estão presentes ao longo de todo o genoma humano, com exceção do DNAMt [69].

Os *SNPs* são as diferenças mais simples entre dois genomas diferentes e consistem em substituições simples de pares de bases (transversões ou transições) [70]. No DNAMt, os *SNPs* são responsáveis por cerca de 90% das variantes genómicas e ocorrem principalmente na região de controlo [68, 70].

Para além dos *SNPs* e dos *STRs*, outros polimorfismos identificados, como inserções ou deleções de um ou mais pares de base, podem ser identificados tanto no DNAMt como no restante genoma humano [71].

1.1.3. Haplogrupos mitocondriais

A partir dos estudos iniciais sobre a evolução das populações humanas foi possível uma classificação de todas as sequências de DNAMt numa árvore de natureza filogenética (Figura 1.4), na qual grupos específicos de haplótipos (combinação de estados alélicos de um conjunto de marcadores polimórficos que estão na mesma

molécula de DNA, como por exemplo um cromossoma ou uma região de um cromossoma [72]) estão assim filogeneticamente relacionados. No presente estudo, estes haplótipos são então definidos pelos polimorfismos que podem ser específicos de grupos populacionais distintos à escala global, como por exemplo os africanos [73], europeus [74, 75], americanos nativos [76, 77] e asiáticos [78]. As populações que possuem uma origem comum ou geograficamente próximas mostram frequências similares de haplogrupos (clado monofilético de haplótipos que partilham polimorfismos característicos e derivam de um haplótipo ancestral fundador [72]). Assim, estes haplogrupos uniformizados permitem a possibilidade de obter detalhes específicos sobre as migrações humanas ao longo da história e sobre a sua demografia consoante o sexo, mesmo que enviesada devido às particularidades desta molécula ser transmitida apenas pela linha materna.

Convencionou-se que os haplogrupos e sub-haplogrupos são designados por letras do alfabeto e por números, respetivamente, seguindo a regra letra-número-letra (por exemplo, J1b). Este sistema foi primeiramente introduzido por Torroni [77], que diferenciou quatro grupos fundadores da população americana (A, B, C e D) através de uma amostra de americanos nativos.

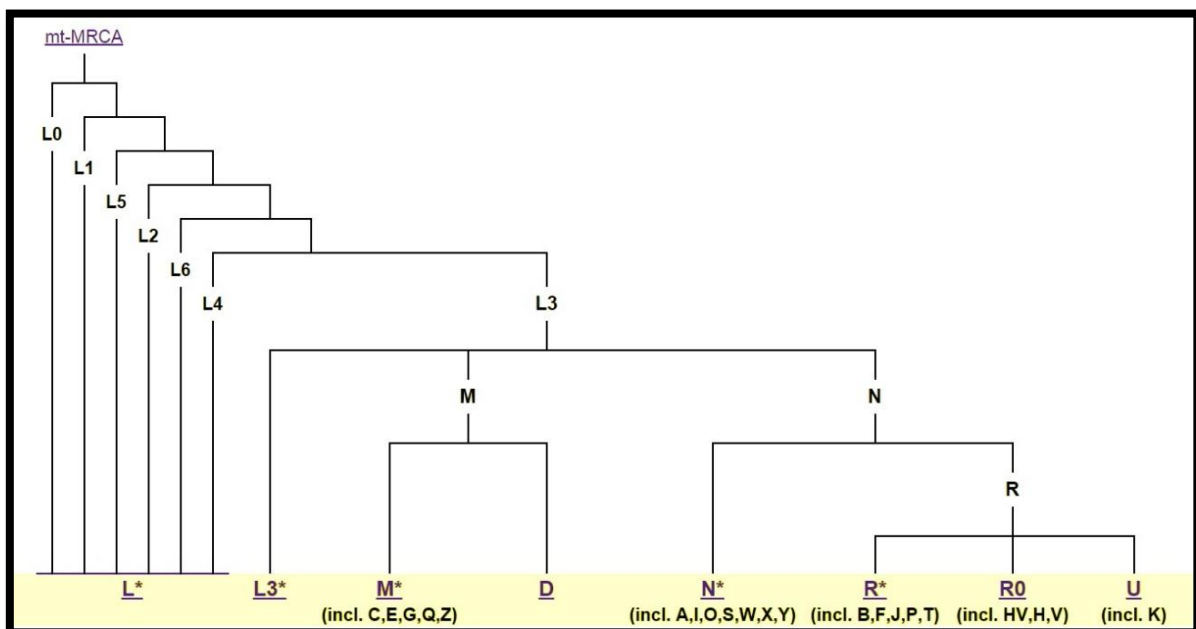


Figura 1.4: Árvore Filogenética simplificada dos haplogrupos do DNAm (adaptado de www.phylotree.com, a 03-10-2015).

Através de evidências genéticas e paleontológicas, chegou-se a um consenso onde se considera África o local de origem dos humanos modernos descendentes do mais

recente ancestral comum (*MRCA – Most Recent Common Ancestor*) que lá viveu há cerca de 200 mil anos [79, 80]. Existem algumas teorias para o aparecimento e consequente dispersão do homem anatomicamente moderno (HAM) [80-82]. Numa ponta do espectro, o modelo *Out-of-Africa* pressupõe que um grupo de humanos modernos, com origem em África há cerca de 100-200 mil anos, se propagou pelo mundo fora, substituindo populações humanas mais “arcaicas” (como os Neandertais) sem se intermisturarem. Na outra ponta do espectro surge o modelo do multirregionalismo onde se destaca o papel da continuidade genética ao longo do tempo e do fluxo de genes entre as populações “arcaicas” sustentando que os humanos modernos surgiram não só em África, mas também na Europa e Ásia. Entre estes dois modelos, assumem-se dois cenários alternativos: o modelo *Out-of-Africa* conjugado com hibridação, que pressupõe a mistura dos humanos modernos com as populações arcaicas locais; e o modelo da Assimilação que, aceitando a origem africana, pressupõe que os humanos modernos se propagaram e, com um processo de cruzamentos mais complexo e longo, os genes e a morfologia poderiam modificar-se localmente, ao invés de uma rápida substituição [80-82]. Mais recentemente, Harris [83] provou que cerca de 2-4% do material genético das populações fora de África derivam de cruzamentos entre os Neandertais e os homens anatomicamente modernos.

Na raiz desta árvore evolutiva encontra-se um conjunto de sequências, agrupadas em haplogrupos mais antigos e de onde se originaram todos os outros a que se denomina haplogrupo L. A raiz deste haplogrupo L terá tido a sua origem na África central há cerca de 180 mil anos [72, 84]. A primeira divisão (e a mais antiga) ocorre entre L0 e L1-L6. L0 possui provavelmente origem no sul do continente africano há cerca de 130 mil anos, sendo frequente nesta zona em grupos como os Khoe, San e falantes Bantu (exceção dos subgrupos L0a, L0b e L0f que possuem origem a este de África). Por outro lado, L1-L6 possui origem na África central/este. Este ramo é mais frequente em África do que L0, com L5, L6 e L4 limitados à zona este de África e L1 (125 mil anos) encontrado na África central/oeste. L2-L6 é datado há cerca de 130-140 mil anos, com divisões mais recentes que posteriormente deram origem às linhagens do oeste de África [72, 79, 84]. Fora de África, estas linhagens são raras e limitadas às zonas que receberam um fluxo genético histórico, através da escravatura, como, por exemplo, oeste da Ásia e a América [79]. A Figura 1.5 traduz as migrações da linhagem de DNAm_t, e as restantes pelo resto do mundo que serão descritas abaixo.

o Norte de África. As linhas mais antigas da Europa são U5 (cerca de 37 mil anos) e U8 (cerca de 50 mil anos). Como grupos basais da Europa mais recentes temos, ainda, os haplogrupos R0 (designado inicialmente de pré-HV) e JT, derivados do haplogrupo R, e N1, N2 e X, derivados diretamente do haplogrupo N [88-91]. Posteriormente, foram surgindo outros haplogrupos na Europa, como V, H1, H3 (os três com origem há apenas 11 mil anos) e H5 (cerca de 14 mil anos), todos com origem aparente no sudoeste da Europa. Mais tarde, migrações do Médio Oriente trouxeram para a Europa outros haplogrupos como K, W, T* e X há cerca de 15 mil anos e J1a, J2a e K2a há cerca de 8-9 mil anos, juntamente com o advento da agricultura [88-90]. Atualmente, o haplogrupo H é considerado o mais frequente nestas duas regiões [88, 90].

Na Ásia, M e N contribuíram igualmente para a expansão mitocondrial. As duas rotas da saída de África contribuíram para o povoamento da Ásia: pela rota norte o HAM dirigiu-se para a Ásia central; e pela rota sul povoou toda a linha costeira da Ásia [92]. O continente Indiano foi povoado há pelo menos 50 mil anos e aqui desenvolveram-se haplogrupos indo-específicos: U2a-c, R5-8, R30, R31, N1d e N5 (derivados do haplogrupo N); M2-6, M30-47 (derivados do haplogrupo M, fazendo parte de cerca de 60% da população da Índia) [92-94]. A zona sul da Ásia serviu também como um corredor para as dispersões do HAM após a saída de África, povoando até a zona este da Ásia. Estudos recentes mostram ainda que existiram dispersões pelo interior, de Myanmar até o sudoeste da China, auxiliando na povoação da zona este da Ásia [95]. Em toda esta zona, é comum encontrar os haplogrupos A-G, Z e M7-M9 [96, 97]. Quanto à região sudoeste da Ásia, os haplogrupos desta zona são comuns com os do Médio Oriente, referidos acima [88]. O HAM chegou à Sibéria do lado do Pacífico há cerca de 40 mil anos pela linha sul costeira da Ásia e esta zona possui fundadores tanto da Europa como da zona este da Ásia [98, 99]. Podemos, então, encontrar por toda a Sibéria haplogrupos da linhagem U diferentes dos já descritos para a Europa, como o U2a, U4a-b e U5a1, e das linhagens H, HV, N, X2 (nomeadamente X2e2), W e Z1 [98, 99]. Também se encontram algumas linhagens que posteriormente deram origem às populações que se dirigiram em direção ao continente americano, como o A2, B2, C4c, D1, D3 [98].

A Oceânia terá sido povoada por duas ondas migratórias principais: uma primeira há cerca de 50 mil anos, derivada da saída de África pela rota sul e através da Índia, que chegavam à Nova Guiné e Austrália; e uma segunda onda que ocorreu muito mais tarde, há cerca de 5-6 mil anos, que chegava à Oceânia Remota (parte este da

Oceânia em que as distâncias entre as ilhas são maiores, que inclui as ilhas Fiji e Polinésia, entre outras) [100, 101]. Os haplogrupos nativos Q, P, S (este último apenas na Austrália) e M27-29 são associados à primeira onda migratória [100, 102, 103]. A segunda onda migratória é associada à disseminação do haplogrupo B, principalmente o B4a1 e suas linhas descendentes, restritos a esta zona [102].

Por fim, estudos apontam que o continente americano foi o último a ser povoado pelo HAM, muito provavelmente através da Ásia pela zona da *Beringia* (zona que conectava a Ásia à América nos períodos em que o nível do mar era baixo, correspondendo ao atual Estreito de Bering) [104-106]. Pode-se referir que o início da colonização da América foi feito num modelo de três etapas: 1) expansão da Ásia central para a *Beringia*, após divergência dos próprios asiáticos (há cerca de 43-36 mil anos); 2) isolamento da população na *Beringia* (durante cerca de 20 mil anos); e 3) expansão da *Beringia* para a América (há cerca de 16 mil anos) [104, 106]. Atualmente existem 16 linhagens descobertas que deram origem aos nativos americanos: A2*, B2*, C1b, C1c, C1d*, C1d1, C4c, D1, D4h3a, D4e1c, X2a e X2g, sendo estes parte da principal vaga que colonizou a América de norte a sul [79, 105, 106]; D2a (expansão dos paleo-esquimós ao longo do Ártico pelo norte do Canadá e Gronelândia) [79, 107]; A2a, A2b e D3 que substituíram a linhagem D2a através da expansão dos neo-esquimós [79, 107].

O haplogrupo X é uma exceção ao padrão que temos visto anteriormente, de limitação geográfica dos haplogrupos. É encontrado na Europa, apresentando apenas frequências apreciáveis no Sul de Portugal, e Médio Oriente, e também nos grupos nativos norte americanos. Contudo, está ausente na Sibéria e nas populações da zona este da Ásia. Este haplogrupo é dividido em dois subhaplogrupos, X1 e X2. X1 é comum e restrito ao norte e este de África, enquanto o X2 encontra-se espalhado pela Europa e Médio Oriente. O haplogrupo X presente nos nativos americanos é uma derivação do X2, por uma combinação de cinco mutações características, muito provavelmente originado do Médio Oriente, sendo este o ponto inicial para a posterior dispersão [108].

Estas distribuições dos haplogrupos pelos continentes podem ser verificadas na Figura 1.6.

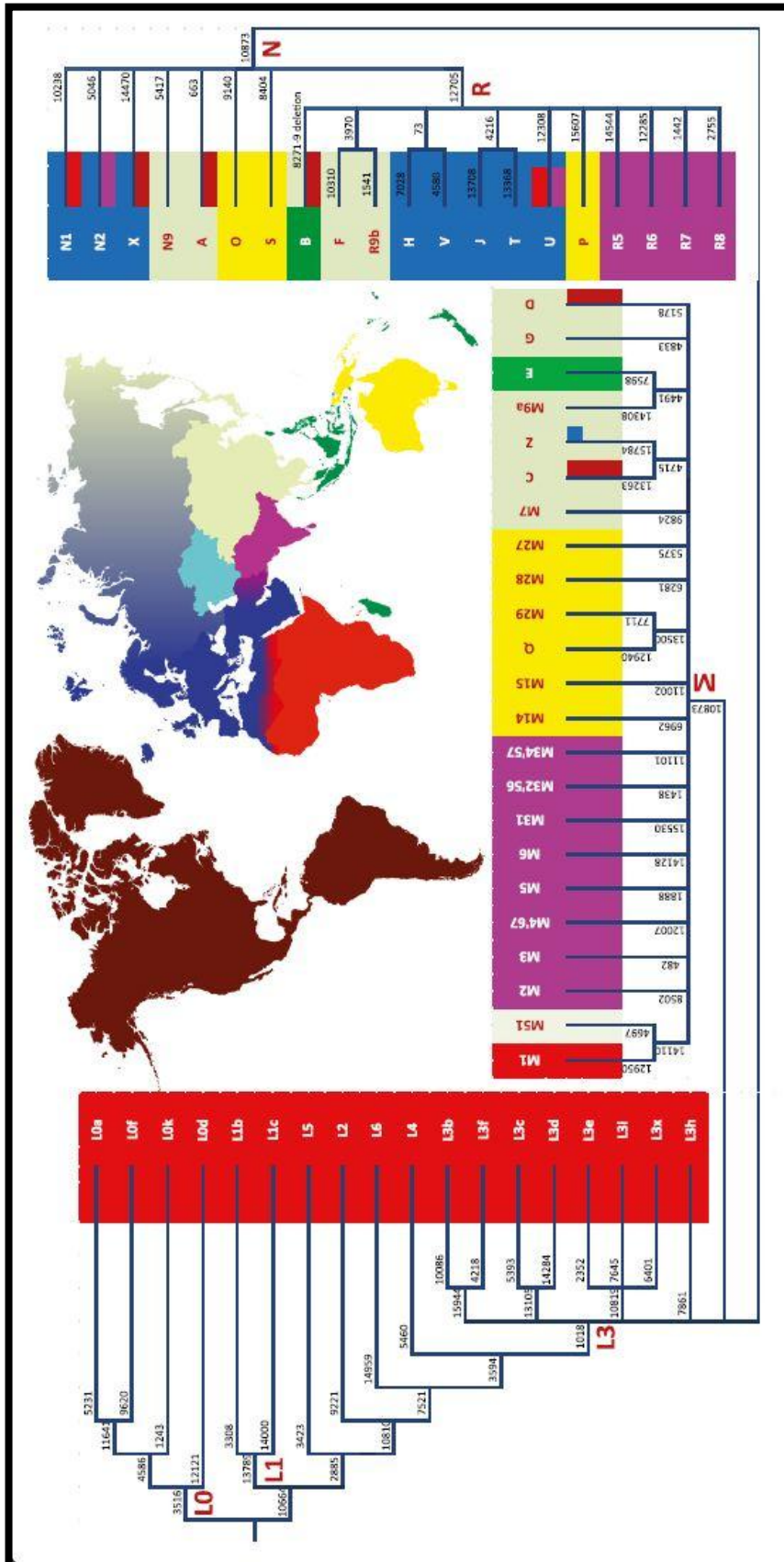


Figura 1.6: Árvore dos haplogrupos mitocondriais e respectivas distribuições pelas diversas regiões do planeta. (Retirado de Kivisild [80])

1.2. Os curdos

Os curdos são considerados um grupo Indo-Europeu originário do Médio Oriente e habitam uma região conhecida como Curdistão. Trata-se de uma região que compreende áreas de quatro países atuais: sudeste da Turquia, oeste do Irão, nordeste do Iraque e norte da Síria, e ainda em pequenas regiões da Arménia (Figura 1.7) [109, 110]. Este grupo possui várias origens étnicas, não obstante, as suas linguagens e sua cultura estão intimamente relacionadas com a população persa. Assim, os curdos são considerados maioritariamente de origem persa [109, 111].



Figura 1.7: Mapa da região do Curdistão com as respetivas delimitações compreendidas entre os quatro países que o consistem (Turquia, Irão, Iraque e Síria).

1.2.1. A história do seu nome

Os árabes foram os primeiros a utilizar a palavra “*Kurd*” no século VII d.C. Nesta altura, a palavra “curdo” significava “nómada”. A partir do século XI em diante, muitos viajantes e historiadores começaram a tratar a palavra como sinónimo de “bandido” (no sentido da pessoa que vive num *gang* e que rouba), significado também partilhado pelos viajantes europeus do século XIX. A meio do século XIX, a palavra

“curdo” era também utilizada como “*tribespeople*” (pessoas de tribo) que falavam a língua dos curdos [112].

Entretanto, esta palavra aparece ainda em diferentes formas com escritores romanos e gregos. A palavra “*kardakes*” é aplicado a uma classe de mercenários asiáticos. Estes são chamados deste modo por viverem do roubo, pois “*karda*” significa “ másculo” e “guerreiro”. Ao contrário, na Assíria (atualmente norte do Iraque) “*qardu*” significa “forte” ou “herói”, e “*qaradu*” significa “ser forte” [109, 110].

1.2.2. A origem dos curdos

De acordo com os historiadores [110, 112-114], os curdos não possuem apenas uma origem. A área que hoje ocupam tem sido desde os seus primórdios uma zona de constante fluxo e passagem de diversos povos [110].

Acredita-se que os curdos são descendentes de diversas hordas de tribos Indo-Europeias que atravessavam a região em direção a oeste através do Irão, possivelmente no segundo milénio a.C. [112]. No entanto, muito pouco se conhece sobre essas tribos.

Os primeiros habitantes do Curdistão terão sido os povos antigos nativos das montanhas de Zagros (maior montanha do Irão, Curdistão e Turquia) e Este de Taurus (Turquia). Entre os das montanhas de Zagros distinguem-se três grupos: Guti, Lullubi (Lullumi) e Kashshu (Kassites), estando estes relacionados com os Hallapi (Elamites). Os habitantes de Taurus eram do grupo Hurrian, que também se pensa serem relacionados com os Elamites [110].

No século IX a.C., ouviu-se falar pela primeira vez sobre as invasões persas. Estas moviam-se em direção a sul, possivelmente através do Caucasus (fronteira entre a Europa e a Ásia, entre os mares Negro e Cáspio), com destino a Fars (uma das 31 províncias do Irão). No final deste século, o reino de Mannai existia na maior parte do sul do Curdistão e oeste do lago Urmiya (representado na Figura 1.7), e consistia num atenuador entre a Assíria, o seu inimigo Urartu a norte e os Medes que se estabeleceram entre o Tehran e Hamadan (ambos noroeste do Irão, perto do lago Urmyia). No século VI a.C., a Pérsia constitui-se, então, num império consistente [110, 112, 114].

1.2.2.1. As conquistas muçulmanas

Durante o período de *Sassanian* (último império persa antes da ascensão do Islamismo), os curdos parecem ter ganho poder e influência de forma constante. As tropas muçulmanas encontraram-nos na posse das fronteiras montanhosas do oeste da Pérsia (império que irá dar origem ao presente Irão), onde estes resistiram à entrada dos muçulmanos, dando um forte suporte ao império de *Sassanian* (639-644 d.C.) [110, 112].

Com a conquista muçulmana, os curdos emergiram de uma obscuridade histórica. O primeiro contacto com as tropas muçulmanas deu-se com a conquista destes da Mesopotâmia (atual Iraque, Síria e Kuwait) em 637 d.C., com a ocupação de Takrit (norte do Iraque) e de Hulwan (ponta oeste do Irão). Quando se tornou claro que o império estava condenado, os chefes curdos renderam-se às tropas muçulmanas. Os curdos ainda participaram nalgumas revoltas nesta altura tendo sido a mais eminente a revolta de al-Khirit em Fars, onde se juntaram aos persas e cristãos e fizeram comuns as suas causas. Outras revoltas se deram em diversos momentos da história e contra diferentes impérios: 645, 659, e 666 d.C. contra os muçulmanos; 685, 702 e 708 d.C. contra o califado do Umayyads; e 840, 846 e 866 d.C. contra o império de Abbasid [110, 112].

1.2.2.2. As dinastias no Curdistão

As dinastias no Curdistão surgiram nos séculos X e XI d.C. à medida que aproveitavam todo o território que conquistavam, o poder do império de Abbasid decaía e era eliminado pelas dinastias turcas. Este período foi então marcado pela transição entre dois períodos importantes: o colapso do domínio muçulmano e a ascensão do poder turco [110, 112].

Neste período, surgiram, então, algumas importantes dinastias: Shaddadis (951-1075) no este da Trascaucasia, entre os rios Kur e Araxes; Marwanids (984-1083) em Diyarbakir (representado na Figura 1.7) até o norte de Jazira (Síria); Hasanwayhids (959-1095) em Khuzistan (província do Irão); e Ayyubids (1171-1250) nas zonas montanhosas da Síria, e também no Egito [112].

1.2.2.3. A invasão dos mongóis

A primeira metade do século XIII mostrou ser desastrosa para o Curdistão. Em 1217 e até 1230, os *Khwarazmians* (povo nómada provenientes da Ásia central) iniciaram ataques continuados ao Curdistão apenas abandonando a região com o surgimento de uma força ainda mais temerosa, os mongóis (também provenientes da Ásia central) [110, 112].

A invasão dos mongóis criou uma grande destruição e deslocamento da população no Curdistão. Muitas regiões foram despovoadas pelos massacres ocorridos e migrações constantes. Alguns dos fugitivos deslocaram-se para o Egito e para a Arménia [110, 112].

Já mais de um século depois dos mongóis, os curdos voltam a sofrer nova devastação. Em 1393, Tamerlane (conquistador turco-mongol) capturou Bagdad e dirigiu-se a norte para Mosul (representado na Figura 1.7). Em 1401, após uma revolta dos curdos, Tamerlane volta a devastar Arbil, Mosul e Jazira. Acredita-se que apenas uma vila cristã foi poupada em toda a área [110, 112].

1.2.2.4. Os curdos a partir do século XVI

Após a ascensão dos impérios Otomano e Safávida na Turquia e na Pérsia, respetivamente, o Curdistão e os próprios curdos tornaram-se um ponto de disputa entre estes dois impérios. O império Otomano, com origens tribais, optou pelo sedentarismo da população e, conscientemente, criou uma forma de governo centralizado com uma cultura cívica e formal. O império Safávida conseguiu obter a maior parte do norte do Curdistão, mas foi obrigado a desistir destas aquisições para o império Otomano [110, 112].

A batalha de Chaldiran em 1514 foi marcante para a história do Curdistão. Apesar do império Otomano ter controlado o poder exercido pelo império Safávida, não conseguiu destruí-lo totalmente, resultando na perda do Curdistão por parte do império Safávida. Esta guerra entre a Turquia e a Pérsia continuou por mais três séculos com evidentes repercussões respetivas para os curdos [110, 112].

Em 1639, foi oficializada a linha da fronteira entre estes dois impérios pelo Tratado de Zuhab, dividindo o Curdistão a meio. Esta linha persistiu até 1914, apesar das disputas e invasões constantes que tiveram um grande impacto sobre o Curdistão. Cada império teve que ponderar, então, o quão longe poderia controlar para além das margens

da fronteira estabelecida. Entretanto, os chefes curdos tinham que escolher que império seria melhor reconhecer tendo em conta o desejo de máxima liberdade da interferência do governo contra os locais [110, 112].

Em 1736, Nadir Shah trouxe novas mudanças e transformações aos curdos, após ter derrubado o império Safávida e usurpado o império Otomano. Este conquistador utilizou os curdos em larga escala nas suas expedições militares provocando destruição e perdas de vida nas suas inúmeras guerras com o império de Ottoman. Nadir Shah morreu em 1747, possibilitando o emergir de um notável líder curdo, que, por quase 30 anos, governou a Pérsia. Durante o seu reinado (1750-1779), o Curdistão usufruiu de uma relativa acalmia [110, 112].

A primeira parte do século XIX foi marcada por guerras para a unificação e independência do Curdistão. Foi em 1847 que o último principado curdo independente colapsou [112, 115]. De 1847 até 1881, foram observadas novas revoltas, sob a liderança de chefes tradicionais, para a criação do estado do Curdistão. Até a 1ª Guerra Mundial (1914-1918), estas revoltas foram seguidas por outras esporádicas e regionais contra o governo central, as quais foram duramente reprimidas pelo império Otomano [112, 115].

1.2.2.5. Os curdos após a 1ª Guerra Mundial

O povo curdo entrou na 1ª Guerra Mundial dividido e sem um plano coletivo para o seu futuro. A região afundou-se numa ainda maior desordem que aquela verificada desde o Chaldiran, com uma divisão acentuada seguindo a derrota do império Otomano, em 1918, pelos poderes da Aliança (França, Grã Bretanha e Estados Unidos). Incentivados por esta derrota, um grupo de independentes formou uma delegação presente na Conferência de Versailles para apresentar as reivindicações de uma nação curda [112, 115].

O Tratado Internacional de Sèvres, assinado entre os aliados e o império Otomano, foi concluído em 1920 e recomendava a criação de um estado curdo em parte do território do Curdistão. Tal tratado jamais seria levado à prática devido à fragilidade dos poderes existentes na região [112, 115].

O Tratado de Lausanne foi assinado em 1923 com o renascimento militar da Turquia. Este invalidava o Tratado de Sèvres e, sem garantias a respeito dos direitos dos curdos, anexou a maior parte do Curdistão ao novo estado da Turquia [112, 115].

Durante as décadas de 1920 e 1930, muitas revoltas ocorreram no Curdistão, sendo, constantemente, reprimidas pela Turquia. Foi alvo de bombardeamentos aéreos nos últimos anos, resultando no abate de milhares de curdos turcos. Outras revoltas também foram reprimidas no Irão e no Iraque durante estes anos [112, 113, 115].

A queda da monarquia do Iraque em 1958 prometia uma era mais esperançosa para os curdos. O novo regime revolucionário no Iraque tomou algumas medidas calculadas, como o retorno de alguns exilados, entre eles Mulla Mustafa, para reconciliar-se com os curdos e conquistá-los de novo. No entanto, na década de 1960, uma agitação por parte dos curdos da zona do Iraque para manter um Curdistão unido e autónomo, levou a uma guerra prolongada entre estes, liderados por Mulla Mustafa, e o Iraque [112, 113, 115].

Em 1974, o governo do Iraque procurou impor um plano para uma autonomia limitada no Curdistão. Com o estabelecimento da República Islâmica no Irão em 1979, foi lançada uma campanha de genocídio contra os curdos que habitavam essa zona. Os ataques iraquianos continuaram durante a guerra entre o Irão e o Iraque (1980-1988), finalizando em 1988 com ataques com gases mortíferos em vilas curdas para reprimir a resistência [113, 115]. Desde então, uma organização nacionalista, PKK ou *Kurdistan Worker's Party* (grupo fundado em 1978 e ligado a ações terroristas), tem travado lutas constantes contra o estado turco a favor dos direitos político-culturais e autonomia para os curdos na Turquia [113, 115].

Os curdos podem ser considerados o maior grupo étnico sem estado/país [110, 112, 115]. Apesar de terem sido postos sob sucessivos domínios de vários conquistadores, como árabes, turcos, persas, entre outros, os curdos mantiveram uma certa integridade sem muitas misturas com os invasores, grande parte devido à sua montanhosa e inóspita terra [113].

1.2.3. Religião no Curdistão

O Curdistão é uma zona com grande diversidade de crenças e práticas religiosas. Cerca de 75% dos curdos seguem o Sunismo (ramo do Islamismo que acredita que o sucessor adequado do profeta islâmico Maomé como califa – chefe muçulmano – foi o seu sogro). No Curdistão, existem ainda seguidores do outro ramo do Islamismo, Xiismo (*Shia* – acredita que o sucessor adequado do profeta islâmico Maomé como califa foi o

seu genro) [110, 112]. No entanto, antes da introdução do Islamismo, a maioria dos curdos seguiam o Zoroastrianismo (primeira manifestação do monoteísmo – um deus –, cujas características, como a ressurreição, o paraíso e o inferno, e a vinda de um messias, influenciaram outras religiões como o Cristianismo, o Judaísmo e o Islamismo), que ainda é praticado por uma pequena quantidade de curdos [110, 112]. Além destes, ainda se pode encontrar aqueles que seguem o Iazidismo, uma comunidade religiosa do próprio Curdistão (antiga religião sincrética que inclui elementos do Zoroastrianismo, Islamismo, Cristianismo, Judaísmo e Shamanismo). Até o início dos anos de 1800, os iazidis foram bastante poderosos e numerosos [110, 112]. Outras comunidades religiosas existem ainda no Curdistão, fazendo parte de uma comunidade cultural curda ainda mais ampla. Aqui podemos referir os cristãos, que sempre foram uma comunidade considerável no Curdistão, e os judeus, que têm vivido nesta região há mais de 2000 anos [110, 112].

O Talmude (livro sagrado dos judeus) conta que alguns judeus foram deportados para o Curdistão há cerca de 2800 anos e, eventualmente, foi dada a permissão para que estes convertessem, com sucesso, os curdos locais. Mais tarde, já como comunidades constituídas, os judeus curdos, parte cultural dos curdos que será estudada na presente tese, eram sobretudo comerciantes e artesãos. No início do século XIX, já existiam comunidades substanciais com sinagogas em diversas cidades (como Sulaymaniya e Irbil – representados na Figura 1.7). Apesar do êxodo Sionista de 1948-1952, alguns judeus permaneceram no Curdistão, enquanto os restantes migraram para Israel [112, 116].

1.3. Objetivos do estudo

Este projeto teve como objetivos:

- i. A análise das linhagens de DNAMt de uma população judaica originária da região do Curdistão, recorrendo a 72 indivíduos, através do sequenciação da região de controlo do DNAMt – HVS-I, HVS-II, HVS-III –, com posterior classificação em haplogrupos por via dos polimorfismos identificados.
- ii. Maior conhecimento sobre a variação haplotípica do DNAMt na atualidade da dos judeus curdos, já que existe muito pouca informação sobre esta população do ponto de vista filogenético.

- iii. A análise da variação genética inter-populacional por comparação dos resultados desta população em estudo com os resultados de outras populações descritas na literatura tendo em conta as suas crenças religiosas, e as populações que tiveram alguma influência sobre os curdos ao longo da sua história, de modo a desvendar possíveis relações genéticas entre as populações do Médio Oriente.

- iv. Identificar qualquer influência de outras populações sobre os judeus curdos, ao longo da sua extensa história e existência.

Capítulo 2 – Fundamentos teóricos das técnicas experimentais



2.1. Extração por *Chelex*

A extração de DNA de amostras sanguíneas ou tecidos permite a purificação dos ácidos nucleicos pela combinação de métodos físicos e químicos, ou seja, permite a separação do DNA dos restantes constituintes para que seja utilizado nos subsequentes testes com maior eficácia, nomeadamente na sua amplificação em Reação em Cadeia de Polimerase (*Polymerase Chain Reaction – PCR*), sequenciação, entre outros. Este passo é de especial importância pois a qualidade e integridade dos ácidos nucleicos isolados vão determinar o sucesso nos resultados dos testes realizados.

As técnicas de extração em geral compreendem quatro fases principais: (1) lise celular, destruindo a estrutura celular e libertando o seu conteúdo, (2) inativação de endonucleases, particularmente DNase e RNase, (3) extração física dos ácidos nucleicos dos restantes componentes celulares, e (4) precipitação e ressuspensão dos ácidos nucleicos em soluções-tampão apropriados ou água livre de RNases e DNases [117].

A maioria das técnicas de extração estão compreendidas entre duas categorias: extração em fase líquida (*Liquid Phase Extraction – LPE*) e extração em fase sólida (*Solid Phase Extraction – SPE*). A escolha do tipo de extração a utilizar depende da matriz inicial que contém o DNA ou RNA.

No presente trabalho, foi utilizado o método de extração por *Chelex*, sendo este um dos métodos amplamente utilizados em biologia molecular no isolamento de ácidos nucleicos. Este método é do tipo líquido-líquido (LPE), onde se joga com a solubilidade do que se pretende extrair ou purificar.

Chelex é uma resina quelante utilizada para purificar outros compostos por troca iónica, com grande afinidade para iões metálicos polivalentes. Esta resina é composta por copolímeros de estireno divinilbenzeno, contendo iões emparelhados de iminodiacetato que atuam como grupos quelantes na ligação a iões metálicos [118]. Assim, o método de extração por *Chelex* baseia-se na prevenção da degradação do DNA, durante a ebulição e armazenamento, através da seletiva absorção dos iões metálicos (especialmente o ião magnésio) que atuam como catalisadores (cofatores das DNases) na degradação do DNA a temperaturas elevadas em soluções de baixa força iónica [118, 119].

2.2. Reação em Cadeia de Polimerase (PCR)

A Reação em Cadeia de Polimerase (PCR) é uma técnica muito utilizada em biologia molecular, desenvolvida nos anos de 1980 por Kary Mullis. Esta técnica permite copiar e amplificar pequenos segmentos de DNA ou RNA a baixo custo. Com este método simples e prático, é possível criar em laboratório milhões de cópias (amplificados) de um segmento de DNA ou RNA em apenas poucas horas. O desenvolvimento desta técnica possibilitou então o descarte da utilização de bactérias para amplificar os ácidos nucleicos [120, 121].

A PCR consiste numa reação enzimática *in vitro*, baseada nas reações que ocorrem *in vivo*. Para a reação se dar é necessário preparar a amostra de DNA. A solução de reação de PCR contém vários componentes [120-122]:

- **Cadeia de DNA** – esta cadeia irá servir de molde e contém a sequência que se pretende amplificar.
- **Primers** – pequenas sequências de DNA fabricadas em laboratório com, aproximadamente, 20 nucleótidos, podendo ter qualquer sequência que se pretenda; são desenhados para corresponder, exatamente, aos limites da cadeia de DNA que se deseja amplificar (um *primer* para limitar o início e outro para o fim da sequência). Os *primers* vão permitir também à polimerase iniciar a replicação.
- **Polimerase de DNA** – complexo de proteínas cuja função é copiar o DNA presente nas células antes desta se dividir em duas. A polimerase liga-se ao final do *primer* (que se encontra hibridizado com a cadeia molde) e começa a adicionar nucleótidos à cadeia complementar. A polimerase regularmente utilizada nesta reação é a Taq polimerase, que provém da bactéria *Thermus aquaticus*, é então termoestável, mantendo-se ativa a temperaturas elevadas, e a sua temperatura ótima é aproximadamente 72°C.
- **Nucleótidos trifosfatos (Deoxinucleotide Triphosphate – dNTPs)** – bases de adenina (A), timina (T), guanina (G) e citosina (C), livres em solução, que são usadas pela polimerase para sintetizar a nova cadeia de DNA.
- **Solução-tampão** – solução que fornece um ambiente ótimo e estável para a atividade da polimerase.

→ **Catiões bivalentes** – são cofatores da polimerase do DNA. Normalmente é utilizado o íon Mg^{2+} .

Esta reação desenrola-se num termociclador, aparelho automatizado que controla e varia as temperaturas com períodos programados e em determinado número de ciclos (geralmente entre 30 e 40 ciclos). Cada ciclo é dividido em três passos: (1) desnaturação, (2) hibridação ou *annealing*, e (3) extensão [122-124]:

- 1) No passo da desnaturação, o tubo contendo a amostra a amplificar é elevado a uma temperatura superior a $90^{\circ}C$, geralmente, entre 94 e $96^{\circ}C$, para que as ligações mais fracas, as ligações por pontes de hidrogénio, se quebrem e a dupla cadeia se separe em duas cadeias simples. Neste passo, as ligações mais fortes, as ligações covalentes entre o fosfato e a desoxirribose, continuam intactas.
- 2) No passo da hibridização ou *annealing*, os dois *primers* hibridizam com as cadeias de DNA, um para cada cadeia desnaturada, marcando o início da sequência que se pretende amplificar. Este passo dá-se a uma temperatura mais baixa, a chamada temperatura de *annealing*, geralmente entre 40 e $60^{\circ}C$. A escolha da temperatura varia com a sequência e o comprimento dos *primers* utilizados e irá permitir uma hibridização com maior especificidade.
- 3) Por fim, no passo da extensão, a temperatura é novamente elevada até $72^{\circ}C$ e a Taq polimerase começa a replicar a cadeia, iniciando na região marcada pelos *primers* e seguidamente adicionando os nucleótidos complementares às cadeias presentes em solução. A replicação dá-se no sentido $5'$ para $3'$, já que a polimerase sintetiza as cadeias de DNA exclusivamente nesse sentido.

As primeiras cópias obtidas irão ter tamanhos diferentes dado que a polimerase não reconhece o final da sequência. No entanto, estas irão servir de molde no próximo ciclo, limitando então o outro *primer* à outra extremidade da sequência. Após alguns ciclos, obtêm-se, maioritariamente, cópias apenas da sequência pretendida, de comprimento definido [122].

Este processo pode ser verificado sequencialmente na Figura 2.1.

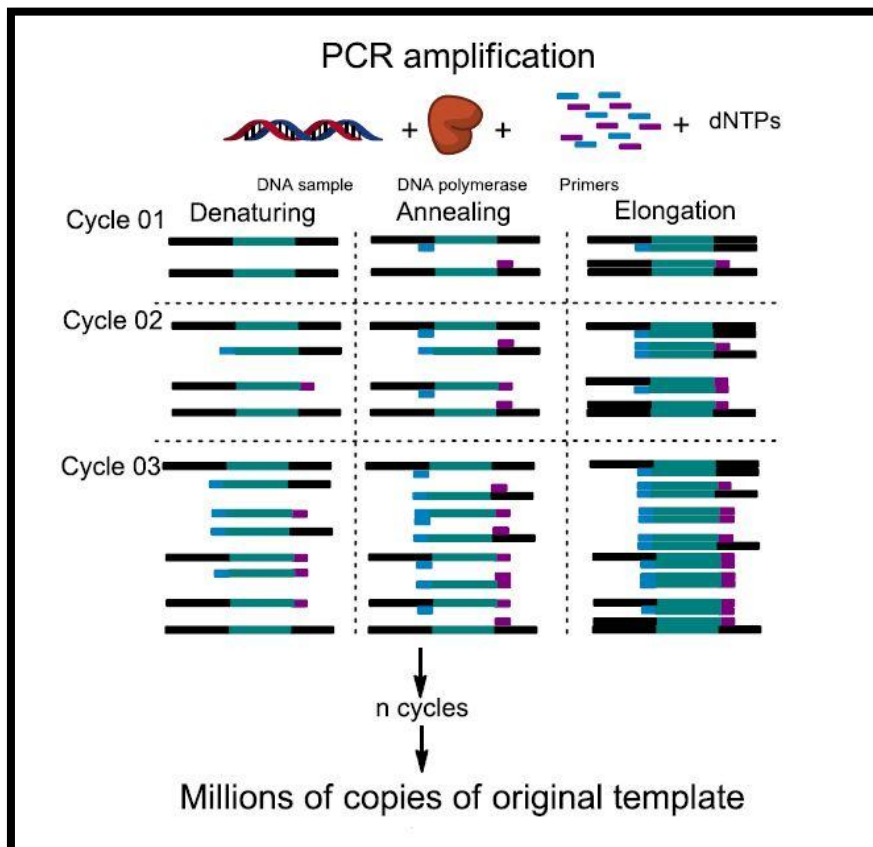


Figura 2.1: Representação esquemática do processo de amplificação do DNA através da técnica do PCR. (Adaptado de Morey [131])

2.3. Eletroforese em gel

A eletroforese é considerada uma técnica de separação que pode ser atingida com grande eficiência com, relativamente, poucos equipamentos. Esta técnica é, principalmente, utilizada com intuito analítico (caracterização qualitativa da substância ou mistura, controlo de pureza) [125, 126].

A eletroforese consiste na migração de moléculas e partículas carregadas na direção do eletrodo com a carga oposta, sob contínua influência de um campo elétrico (Figura 2.2). A separação da mistura das diferentes moléculas e partículas em simples frações dá-se pela diferença de velocidades de migração respetivas, devido às diferentes massas e cargas de cada fração [125, 126]. A mobilidade eletroforética, medida da velocidade de migração, é um parâmetro característico de cada molécula ou partícula, e é dependente dos valores de pK de cada grupo carregado e do tamanho de cada molécula ou partícula. A mobilidade é influenciada pelo tipo, pH e concentração do tampão utilizado, pela temperatura, pela força do campo utilizado e pela natureza do material de

suporte. As separações podem ser realizadas em soluções livres, como capilares e sistemas de fluxo livre, ou em meios de estabilização, como placas de camada fina, películas ou géis [127].

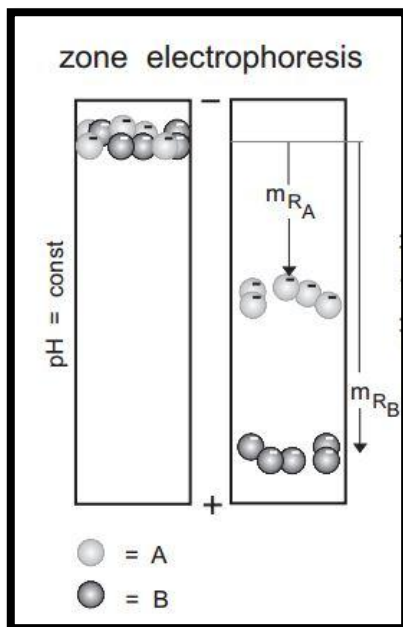


Figura 2.2: Representação esquemática dos três princípios eletroforéticos. A e B são componentes da amostra. (Retirado de Westermeyer [127]).

2.4. Sequenciação pelo método de Sanger

A sequenciação de DNA é uma técnica que nos permite analisar de maneira detalhada um dado genoma providenciando a informação mais elementar de um ser vivo: a sequência genética. Com este conhecimento todo é possível, por exemplo, localizar sequências regulatórias e de genes específicos ou comparar genes homólogos entre espécies e identificar mutações, aplicável nas mais diversas áreas, como forense, genômica comparativa e evolução, medicina aplicada em diagnóstico e terapia, entre outros [128].

Em 1975, Sanger [129] desenhou um método que simula o processo natural da replicação do DNA. Dois anos mais tarde, Sanger [130] publicou uma variação do método mais eficiente. Este método torna-se então o método *standard* pela sua facilidade e praticabilidade, que ainda hoje em dia se utiliza [130, 131].

O método de Sanger (ou também chamado de método de terminação de cadeia) baseia-se na utilização de didesoxinucleótidos (*Dideoxynucleotide Triphosphate* – *ddNTPs*) em adição aos *dNTPs* utilizados, habitualmente, numa reação de amplificação

(referidos no ponto 2.2) [130]. Os *ddNTPs* são, essencialmente, idênticos aos *dNTPs* com a exceção da presença de apenas um átomo de hidrogênio (H) ao invés de um grupo hidroxilo (OH) no carbono da extremidade 3' [132]. Assim, quando integrados numa sequência, estes nucleótidos modificados evitam a adição de mais nucleótidos na sequência por não se formar uma ligação fosfodiéster o nucleótido modificado e o próximo nucleótido, terminando a cadeia [132].

Com o equilíbrio apropriado entre os dois tipos de nucleótidos, a competição entre os dois irá resultar em diversos fragmentos, diferindo no comprimento das cadeias terminadas. Após a limpeza de nucleótidos livres, *primers*, e polimerase do DNA e subsequente desnaturação, as moléculas resultantes são ordenados por massa molecular e lidas, sequencialmente, por eletroforese capilar [128, 132]. Este processo está representado na Figura 2.3.

Para a leitura ser possível, os nucleótidos modificados são marcados com marcadores e emitem fluorescência com diferentes comprimentos de onda (Figura 2.3), associando, assim, uma cor diferente a cada nucleótido [133].

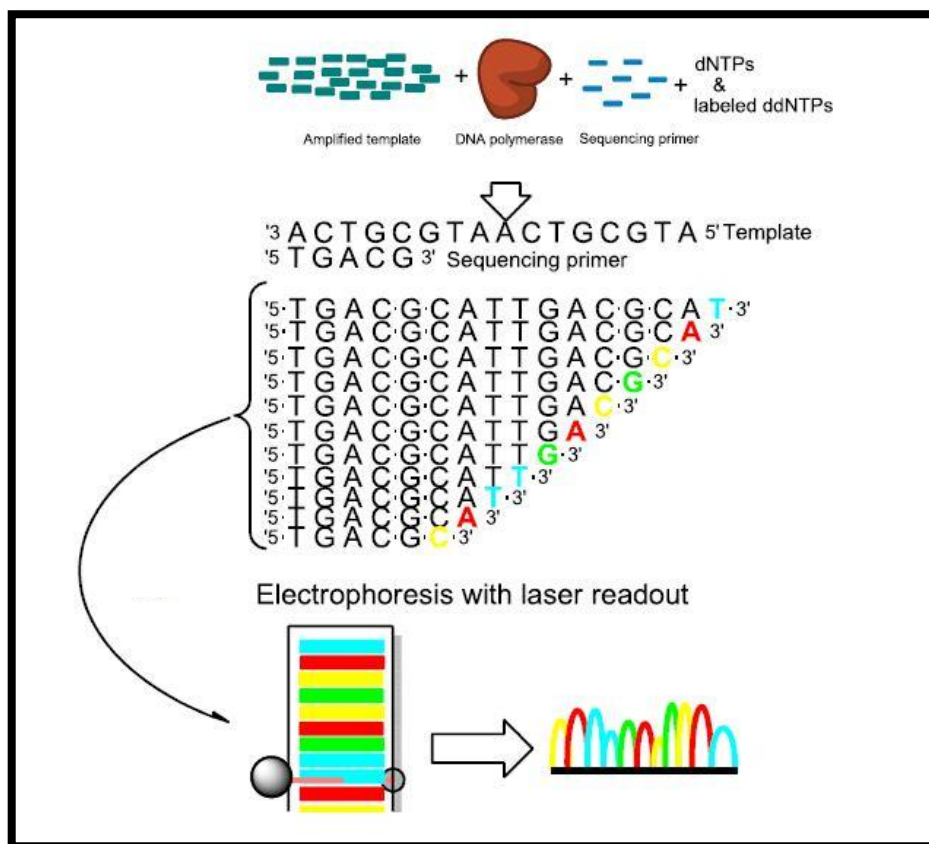


Figura 2.3: Representação esquemática do processo de Sequenciamento de Sanger, com o método de terminação de cadeia devido aos *ddNTPs* (devidamente marcados com fluorescência) e posterior separação por eletroforese dos fragmentos por tamanho. A detecção do sinal é automatizado com auxílio de um laser e um computador. (Adaptado de Morey [131]).

Capítulo 3 – Material e metodologia



3.1. Coleção da amostra populacional

Para este estudo, foram recolhidas amostras de saliva e de sangue de 72 indivíduos voluntários e anónimos, sem grau de parentesco entre os mesmos, da região do Curdistão, refugiados em Israel (Médio Oriente). A maioria dos antecessores destes indivíduos eram do Norte do Iraque. Estes curdos pertenciam à coleção de DNA estabelecida no laboratório em Israel (Laboratório de Bioantropologia e DNA Antigo, Divisão Dentária de Anatomia e Biologia Celular, Faculdade de Medicina Dentária – Universidade Hebraica de Jerusalém) para o estudo de doenças hematológicas ou de voluntários que se autoidentificaram de judeus curdos. Estas amostras foram colhidas com o consentimento informado de todos os indivíduos, seguindo as orientações éticas.

3.2. Extração de DNA

A extração do DNA das amostras obtidas foi feita através do método por *Chelex* [119]. Todo o processo de extração destas amostras utilizadas no presente estudo foi realizado no mesmo Laboratório de Bioantropologia e DNA Antigo, Divisão Dental de Anatomia e Biologia Celular, Faculdade de Medicina Dentária da Universidade Hebraica de Jerusalém. As amostras foram posteriormente cedidas pela Professora Marina Faerman ao Laboratório de Genética Humana da Universidade da Madeira (LGH) num contexto de colaboração.

3.3. Amplificação por PCR da região de controlo do DNAMt

A amplificação da região de controlo do DNAMt foi baseada no método descrito por Hamoy [134]. A região de controlo do DNAMt da população em questão foi, então, amplificada no termociclador 2720 Thermal Cycler (Applied Biosystems) com os *primers* descritos na Tabela 3.1. A solução de amplificação para cada amostra era constituída por 10 pmol de cada *primer* (Eurofins Genomics), 160 μ M de cada dNTP (Promega), 0.8 mM de $MgCl_2$ (Solis Biodyne), 10x de tampão de reação (Solis Biodyne), 1 U de FIREPol Taq polimerase do DNA (Solis Biodyne) e 1 μ L de DNA extraído, num volume total de 25 μ L. O perfil de amplificação otimizado utilizado foi: desnaturação inicial a 94°C durante 5 minutos, seguido de 35 ciclos com desnaturação a 94°C durante 1 minuto, *annealing* a T_a °C (consoante o *primer* utilizado, descrito na Tabela 3.1) durante 40 segundos e extensão a 72°C durante 50 segundos, e terminando com uma extensão final

a 72°C durante 1 minuto. Os amplificadores foram posteriormente verificados por eletroforese, em gel de agarose (Lonza SeaKem® LE) a 1.5%, aplicando voltagem constante a 90 volts, durante cerca de 30 minutos.

Tabela 3.1: Fragmentos amplificados da região controlo do DNAm_t e respetivos primers utilizados. (Legenda: pb – pares de base; T_a – temperatura de annealing).

Fragmento	Tamanho do fragmento (pb)	-	Nome	Sequência (5'-3')	T _a
MTT1	1822	FW	14898forb	TAGCCATGCACTACTCACCAGA (22)	61
		RV	151revb	GGATGAGGCAGGAATCAAAGAC (22)	
MTT2	1758	FW	16488fora	CTGTATCCGACATCTGGTTCCT (22)	60
		RV	1677reva	GTTTAGCTCAGAGCGGTCAAGT (22)	

3.4. Purificação do DNAm_t amplificado

Para a purificação dos fragmentos de DNAm_t amplificados, colocou-se 500 µL de Sephadex G-50 (GE Healthcare) em cada coluna de purificação GFX (Pharmacia), e centrifugou-se durante 3 minutos a 4000 rpm (Kubota 1300). De seguida, colocou-se a solução de reação de amplificação em cada coluna sobre o Sephadex G-50 e centrifugou-se sob as mesmas condições descritas anteriormente. Neste passo, a coluna foi colocada dentro de um tubo eppendorf para recuperar o DNAm_t amplificado purificado.

3.5. Sequenciação automática da região de controlo do DNAm_t

A região de controlo do DNAm_t foi posteriormente sequenciada pelo método de Sanger, num equipamento 3130 xl Genetic Analyzer (Applied Biosystems) de 16 capilares (36 cm, polímero POP 7) do Serviço Especializado de Epidemiologia e Biologia Molecular (SEEBMO), do Hospital de Santo Espírito da ilha Terceira, Açores, ao abrigo de um acordo de cooperação entre este e o Laboratório de Genética Humana da Universidade da Madeira.

Para cada amostra, foi realizada uma reação de amplificação com o primer específico para posterior leitura de sequência. Os *primers* utilizados foram L15997 e H017 (quando necessário) para o primeiro fragmento e L16555 e H599 (quando necessário) para o segundo fragmento. A solução de reação era constituída por 2 µL de diluente, 1 µL de pré-mix de sequenciação (Big Dye v3.1 premix diluído com tampão recomendado pelo fabricante), 1.6 pmol de *primer* FW e 1 µL de DNA *template*, num

volume total de 10 μL . O perfil de amplificação utilizado foi: desnaturação inicial a 96°C durante 1 minuto, seguido de 25 ciclos com desnaturação a 96°C durante 10 segundos, *annealing* a 50°C durante 5 segundos e extensão a 60°C durante 4 minutos, e terminando com uma extensão final a 60°C durante 4 minutos.

Procedeu-se à limpeza da mistura pelo método de purificação por Etanol/NaAc/EDTA. Primeiramente, centrifugou-se brevemente os tubos contendo a reação de sequenciação, de modo a recolher qualquer condensação que se tenha formado. Adicionou-se 1 μL de EDTA 125 mM (pH=4.6) a cada tubo, e de seguida, 1 μL de NaAc 3 M (pH=4.8) e 25 μL de etanol a 100% (-20°C). Incubou-se a temperatura ambiente no escuro durante 15 minutos. Seguiu-se uma centrifugação na velocidade máxima a 4°C durante 30 minutos e descartou-se o sobrenadante. Adicionou-se, então, 35 μL de etanol a 70% (-20°C) a cada tubo e centrifugou-se na velocidade máxima a 4°C durante 15 minutos. Removeu-se todo o sobrenadante cuidadosamente e ressuspendeu-se em 15 μL de formamida (Applied Biosystems). Por fim, desnaturou-se as amostras a 95°C durante 2 minutos e foi feita a leitura da sequência no sequenciador automático.

3.6. Análise da região de controlo

As sequências obtidas foram alinhadas com a Sequência de Referência de Cambridge revista (*rCRS*) [9, 16], recorrendo ao auxílio do programa Sequencher® (versão 5.3), por forma a aferir da existência de eventuais polimorfismos (*SNPs*) em cada amostra, sendo estes marcados como mutações. O conjunto destes polimorfismos permitiu a atribuição de haplogrupos específicos a cada indivíduo, recorrendo ao programa HaploGrep 2.0 (<https://haplogrep.uibk.ac.at/>) e com o auxílio da Árvore Filogenética (<http://www.phyloree.org/>). Para todas as amostras, recorreu-se à região codificante para uma fidelidade maior na determinação do haplogrupo de cada indivíduo, utilizando a restrição enzimática para verificar as posições características de cada haplogrupo (as enzimas de restrição utilizadas estão descritas na Tabela 3.2). A mistura enzimática (constituída por 20 unidades de enzima, 1x tampão da enzima específica, 2.5 μL de BSA quando necessário) foi misturada com 3 μL de DNA amplificado, obtendo uma mistura de reação total de 25 μL . A digestão foi feita a 37°C, em banho-maria, durante a noite. Os fragmentos foram posteriormente visualizados em gel de agarose a 2.5%. Para algumas das amostras, foi necessário recorrer à técnica da sequenciação automática, devido a não ser possível o recurso às enzimas de restrição para uma

confirmação viável desse mesmo haplogrupo (posições descritas na Tabela 3.2 com ausência da descrição da enzima de restrição).

Os índices de diversidade haplotípicas e nucleotídicas das amostras foram calculados recorrendo ao programa Arlequin v. 3.5.2.2 [135]. Os polimorfismos de comprimento (inserções de C), nomeadamente nas posições 309 e 315, foram ignoradas para efeitos de cálculos dos índices por serem posições *hotspot*.

Tabela 3.2: Posições utilizadas na confirmação dos haplogrupos e respetivas enzimas de restrição utilizadas. Foi utilizada a sequenciação automática nas posições em que a enzima não é descrita.

Haplogrupo	Mutação	Enzima	Zona amplificada
H	T7028C	AluI	6511/7315
H1	G3010A	–	2499/3346
H2	G1438A	–	1245/2007
H2a	G4769A	–	3967/4812
H13	C14872T	MboI	13809/14998
J1	G3010A	NcoI	2646/3346
J1b2	C1733T	–	1245/2007
J1b3	A8460G	AluI	7937/8797
J1b6	A5501G	–	3734/5571
J2	G15257A	AccI	14898/15349
K1	T1189C	–	16488/1677
K2	T9716C	BsaI	9230/10130
M1	C12403T	MnII	12104/12818
M1a1	C12346T	HpyCH4V	12104/12818
N	C12705T	MboII	12541/12818
N1a	A13780G	HpyCH4III	13477/13830
N1b1	A12822G	BsrBI	11977/13830
N2a	C739T	–	16488/1677
R	T12705C	MboII	12541/12818
R0a1	A827G	–	16488/1677
T	G13368A	BamHI	13338/13507
T1	C12633A	AvaII	12541/12818
U1	A13104G	MboI	12571/13507
U6	A3348G	–	2646/3947
U7	C5360T	TasI	5255/6031
X2	G1719A	–	1404/2007

3.7. Construção das Árvores Filogenéticas

As Árvores Filogenéticas das linhagens de DNAmT obtidas neste estudo foram construídas manualmente, verificadas com o auxílio do programa HaploGrep 2.0 (<https://haplogrep.uibk.ac.at/>) e do programa Network 5.0.0.0 (Fluxus Technology Ltd.),

combinando os algoritmos *Reduced-Median* (RM) e *Median-Joining* (MJ) [136, 137] e, posteriormente desenhadas recorrendo às ferramentas de desenho existentes no sistema operativo *Windows*. É de notar que as árvores foram apenas construídas por macrohaplogrupo separadamente. As relações filogenéticas foram baseadas na variação nucleotídica da região controlo, conferindo maior prioridade à informação da região codificante. A resolução das árvores foi melhorada através da atribuição de informação sobre a mutabilidade nas diferentes posições [138], alterando o seu peso respetivo na contabilização do polimorfismo consoante a sua importância na classificação do haplogrupo, tendo por base os pesos utilizados por Kloss-Branstätter [139] com o objetivo de resolver as redes que se formavam.

3.8. Análise estatística

Para a análise estatística, os índices F_{ST} e respetivas significâncias entre as populações foram estimadas pelas frequências haplotípicas através da análise de permutações, utilizando 2000 permutações, recorrendo ao programa Arlequin v. 3.5.2.2 [135]. Como a resolução dos haplogrupos mitocondriais não era uniforme entre os estudos utilizados, os respetivos haplogrupos foram reduzidos aos marcadores mais informativos partilhados. Para ser possível o estudo da estrutura populacional dos judeus curdos, a variação genética foi repartida dentro e entre populações, agrupando-as em grupos segundo a sua localização geográfica com o teste AMOVA (Análise da Variância Molecular), por meio do programa Arlequin v. 3.5.2.2 [135]. Os dados utilizados foram retirados da literatura em comparação com a nossa população em estudo. As populações e respetivas referências utilizadas estão descritas na Tabela S.1 do Material Suplementar.

As Análises de Componentes Principais (PCA) foram realizadas recorrendo ao programa IBM SPSS Statistic versão 23 (SPSS Inc., Chicago, Illinois) (IBM SPSS Statistics) [140], com duas componentes principais, para comparar os judeus curdos tanto com outras populações descritos na literatura que influenciaram a população em estudo durante a sua história (populações de 20 países), como com outras populações judaicas (14 populações judaicas). Os dados utilizados foram retirados da literatura em comparação com a nossa população em estudo. As populações e respetivas referências utilizadas estão descritas na Tabela S.1 do Material Suplementar.

Capítulo 4 – Resultados e discussão



4.1. Análise do DNAm numa população de judeus do Curdistão

A análise da região de controlo do DNAm dos 72 indivíduos judeus curdos, em comparação com a *rCRS*, permitiu a classificação desta população em haplogrupos mitocondriais característicos europeus (descritos na Tabela 4.1 e nas Tabelas S.2 e S.3 do Material Suplementar). H, J1 e N1 foram os macrohaplogrupos maioritários encontrados, correspondendo a 66.66% do total da amostra populacional. As frequências das classes presentes nesta população estão descritas na Tabela 4.1.

Tabela 4.1: Frequências dos macrohaplogrupos classificados na população de Judeus Curdos.

Macrohaplogrupo	Judeus Curdos (n=72) %
H	33.33
J1	19.44
J2	1.39
K1	5.56
K2	1.39
M1	1.39
N1	13.89
N2	1.39
R0	1.39
T1	2.78
T2	4.17
U1	4.17
U7	2.78
X2	6.94

Dos haplogrupos identificados, cerca de 98.6% correspondem a haplogrupos característicos da região do Médio Oriente, enquanto que os restantes (cerca de 1.4%), nomeadamente M1a1 (1 indivíduo) se enquadram em haplogrupos característicos do Norte de África (Figura 1.6). Estes resultados estão de acordo com os dados descritos na literatura. Sendo esta região uma zona de passagem entre África e os continentes asiático e europeu onde há cerca de 30 mil anos ocorreu a migração de volta ao Norte de África a partir do Médio Oriente, é provável que os haplogrupos M1 e U6 tenham entrado no *pool* genético desta população durante este fenómeno [88], explicando a sua presença, ainda que mínima, nesta população.

No geral os resultados alcançados estão em linha com os apresentados no estudo efetuado por Comas [141], que utilizou uma amostra populacional de 29 curdos. O macrohaplogrupo H surgiu com maior frequência em ambos os estudos (33.3% no presente estudo e 44.8% no estudo efetuado por Comas [141]) e os macrohaplogrupos T, U1 e X figuram em percentagens semelhantes. Curiosamente, a frequência do macrohaplogrupo K surge reduzida no nosso estudo (6.9% da amostra populacional), contrariando os 17.2% apresentados por Comas [141]. Este último identificou ainda três haplogrupos que não foram encontrados na nossa amostra, U3, W e I. Já os macrohaplogrupos J, M, N e R caracterizados neste estudo não foram encontrados na amostra populacional de curdos utilizada por Comas. É importante referir que apenas 29 indivíduos, uma diminuta amostra populacional de curdos revela ser, notoriamente, insuficiente para uma extrapolação para toda a população. Por outro lado, a amostra populacional, não totalmente ideal, no presente estudo, consegue espelhar de forma mais representativa a diversidade populacional existente na região.

Dos resultados obtidos através do programa Arlequin v. 3.5.2.2 [135], observa-se um nível considerado de diversidade para a população em estudo. Com base, unicamente, na análise da região controlo, encontraram-se 46 haplótipos distintos (Tabela 4.2), o que é desde logo um número relativamente elevado, considerando que se trata apenas de 72 amostras diferentes. Apresenta ainda uma diversidade haplotípica de 0.9789 ± 0.0071 (Tabela 4.2), ou seja, escolhendo ao acaso dois haplótipos há cerca 98% de probabilidade de serem diferentes. Tendo em conta que o estudo cingiu-se a 72 indivíduos, considera-se este número de haplótipos, substancialmente, elevado, podendo ser consequência direta do baixo número de efetivos da amostra populacional. Caso esta fosse maior, poderia ser possível verificar frequências superiores de determinados haplótipos.

Tabela 4.2: Parâmetros estatísticos obtidos a partir da análise da região de controlo do DNAmT nos indivíduos Judeus Curdos. (Legenda: KJ – Judeus Curdos)

Parâmetros estatísticos	Região controlo KJ
Número de haplótipos	46
Diversidade haplotípica	0.9789 ± 0.0071
Diversidade nucleotídica	0.0136 ± 0.0069
Transições observadas	66
Transversões observadas	14
Inserções e deleções observadas	10
Sítios polimórficos	89

Verifica-se também uma diversidade nucleotídica minorada de 0.0136 ± 0.0069 na região controlo, ou seja, escolhendo dois nucleótidos ao acaso há apenas cerca de 1.3% de probabilidade de serem diferentes. Apesar de apresentar uma diversidade haplotípica relevante, os polimorfismos ocorrem relativamente nos mesmos *loci* (apenas 89 sítios polimórficos). Confrontando os resultados deste estudo com os valores obtidos por Comas [141], verifica-se que são ligeiramente mais elevados (diversidade haplotípica: 0.958; diversidade nucleotídica: 0.0123), no entanto, continuam a ser mais baixos do que os definidos na região, tal como descrito por Comas [141]. Foram, ainda, calculados alguns parâmetros estatísticos básicos para toda a região controlo do DNAm, que podem ser consultados na Tabela 4.2.

4.1.1. Árvores filogenéticas dos judeus curdos

A abordagem filogenética (do grego *phylon* = raça; *genetic* = nascimento) é a classificação de *taxa*, baseando-se na sua relação de proximidade em termos evolutivos. A variação e o padrão existentes das relações das linhagens são expressos pela construção de árvores filogenéticas com intuito de organizar e ordenar as relações evolutivas entre diferentes variantes de forma relevante e significativa [142].

Para cada macrohaplogrupo, construiu-se uma árvore filogenética com o propósito de exibir estas relações de proximidade na população de judeus curdos. Os macrohaplogrupos estão representados nas Figuras 4.1-4.9. Para esta população não foi possível representar apenas uma árvore filogenética representando a totalidade das linhagens identificadas nos judeus curdos, já que não foi estudada a molécula de DNAm toda. Não foi possível calcular os respetivos tempos de coalescência pois seria necessário que as respetivas árvores possuíssem uma forma do tipo estrela, um fundador claro no grupo em questão e/ou a existência de pelo menos cinco haplótipos diferentes por grupo para obter uma estimativa fiável [143].

Adicionalmente, com o auxílio das ferramentas existentes na base de dados EMPOP v3/R11 (<http://empop.online/>), desenvolvida pelo Instituto de Medicina Legal na Universidade de Medicina de Innsbruck e pelo Instituto de Matemática da Universidade de Innsbruck (consiste em dados forenses, literatura publicada e sequências não publicadas de laboratórios participantes), efetuou-se uma comparação dos haplótipos identificados com os de demais estudos. Os resultados desta compilação encontram-se

descritos nas Tabelas S.4 e S.5, a primeira descrevendo as correspondências exatas e a segunda descrevendo outros haplótipos com uma ou duas diferenças dos haplótipos identificados nos judeus curdos.

4.1.1.1. Macrohaplogrupo H

O macrohaplogrupo H predomina nesta população dos judeus curdos. Este macrohaplogrupo é de difícil caracterização uma vez que possui menos polimorfismos em relação à *rCRS* ao longo da região controle do DNAm. Nenhuma das linhagens identificadas nos judeus curdos corresponde à linhagem H2a2a1 pois todas possuem uma transição na posição A263G (Figura 4.1). Todas as linhagens foram testadas na posição 7028 para determinar se eram HV ou H e constatou-se que todas elas não possuíam a transição na posição 7028C, evidenciando que pertencem efetivamente ao macrohaplogrupo H. Para uma melhor compreensão, cada linhagem será discutida em separado.

As linhagens identificadas nos indivíduos 48, 61, 72, 504 e 516 são caracterizadas por uma única transição na posição T16189C. Segundo a base de dados EMPOP, apenas existe uma correspondência exata em Portugal (Tabela S.4). Quanto aos haplótipos com uma ou duas diferenças, foram encontrados distribuídos por todos os continentes, com 422 na América, principalmente nos Estados Unidos, cuja metapopulação era maioritariamente do oeste da Eurásia, e 344 na Europa, principalmente na Alemanha, Espanha e Portugal (Tabela S.5). A partir desta linhagem desenvolveram-se outras cinco, com a transição na posição T16189C em comum. A linhagem identificada nos indivíduos 34, 38 e 521 é caracterizada pelas transições nas posições C16270T e T16519C. A linhagem identificada no indivíduo 30 é caracterizada pela transição na posição A234G. A linhagem identificada no indivíduo 45 é caracterizada pela transição na posição T16183C. Desta última, desenvolveu-se a linhagem identificada no indivíduo 514 caracterizada pela inserção de dois C e desta desenvolveu-se a linhagem identificada no indivíduo 502 caracterizada pela inserção de TAA na posição 244. Segundo a base de dados EMPOP, nenhuma destas linhagens possuía correspondências exatas. Quanto aos haplótipos com uma ou duas diferenças, para todas estas linhagens foram encontrados haplótipos parecidos, maioritariamente, nos Estados Unidos, cuja

metapopulação da maior parte era igualmente do oeste da Eurásia, na Alemanha e em Espanha.

A linhagem identificada no indivíduo 71 é definida pelas transições nas posições T16311C e T16519C. Segundo a base de dados EMPOP, oito correspondências exatas foram encontradas na Suécia, oito na Alemanha e um no Bahrain. Quanto aos haplótipos com uma ou duas diferenças, foram encontrados distribuídos por todos os continentes de forma semelhante ao descrito em cima.

A linhagem identificada no indivíduo 524 é definida pelas transições nas posições T152C, T204C e G15152A e pela transversão na posição G499C. Segundo a base de dados EMPOP, não foram encontrados nenhuma correspondência exata para esta linhagem. Quanto aos haplótipos com uma ou duas diferenças, foram encontrados nove haplótipos parecidos nos Estados Unidos, cinco em Portugal e quatro na Alemanha.

A linhagem identificada no indivíduo 65 é definida pelas transições nas posições T146C e T16519C e pela transversão na posição A16524C. Segundo a base de dados EMPOP, nenhuma correspondência exata desta linhagem foi encontrada. Quanto aos haplótipos com uma ou duas diferenças, possuem correspondentes idênticos um pouco por todo o mundo, com maior representação nos Estados Unidos, na Alemanha, em Espanha, em Portugal e Marrocos.

A linhagem identificada no indivíduo 32 é definida pelas transições nas posições T146C, T16311C e T16519C e possui uma correspondência exata nos Estados Unidos. Para esta linhagem, não foram encontrados haplótipos com uma ou duas diferenças.

A linhagem identificada nos indivíduos 3, 15, 28, 66 e 513 é definida pelas transições nas posições C16278T e T16519C. Estas linhagens possuem correspondências exatas na Grécia, na Alemanha, em Marrocos e nos Estados Unidos. Quanto aos haplótipos com uma ou duas diferenças, foram encontrados haplótipos parecidos por todo o mundo, com maior representação nos Estados Unidos, na Alemanha, em Espanha, no Uzbequistão, em Portugal e Marrocos.

A linhagem identificada no indivíduo 501 é definida pelas transições nas posições A200G, C16261T e T16519C e possui três correspondências exatas nos Estados Unidos. Quanto aos haplótipos com uma ou duas diferenças, foram encontrados haplótipos parecidos em todos os continentes, com maior representação nas mesmas regiões geográficas descritas para a linhagem identificada nos indivíduos 3, 15, 28, 66 e 513.

A linhagem identificada no indivíduo 8 é caracterizada pela transição na posição A16300G. Esta linhagem não possui nenhuma correspondência exata. Quanto aos haplótipos com uma ou duas diferenças, as mesmas representações maioritárias podem ser referidas que para as linhagens dos indivíduos 3, 15, 28, 66, 513 e 501.

Por fim, a linhagem identificada no indivíduo 57 é definida pelas transições nas posições C150T e T16519C. Segundo a EMPOP, foram encontradas três correspondências exatas em Portugal, cinco em Espanha e 11 nos Estados Unidos. Quanto aos haplótipos com uma ou duas diferenças, segue a mesma linha descrita para as linhagens identificadas nos indivíduos 3, 15, 28, 66, 513, 501 e 8.

A árvore filogenética do macrohaplogrupo H (Figura 4.1) é apenas uma mera representação baseada apenas nas sequências da região controlo do DNAMt.

O macrohaplogrupo H é encontrado, predominantemente, ao longo de toda a Europa, representando mais de 40% da variabilidade genética do Oeste da Eurásia, decrescendo a sua frequência a Sul e Este até cerca de 10-30% no Médio Oriente e no Cáucaso [144]. Segundo Roostalu [145], os haplogrupos mais frequentes do macrohaplogrupo H no Médio Oriente e no Cáucaso são H1, H2, H5 e H13.

Para este macrohaplogrupo, não existem certezas sobre quais as linhagens mais recentes ou mais antigas, podendo no entanto tentar inferir da sua antiguidade por via secundária. A linhagem representada no indivíduo 524 poderá ser uma das primeiras linhagens H introduzidas na população de judeus curdos devido à acumulação de mutações privadas (significa que não foram encontradas noutros haplogrupos) e por não possuir nenhuma correspondência exata e muito poucos haplótipos parecidos. Outra linhagem que pode ter sido introduzida na mesma altura é a representada nos indivíduos 48, 61, 72, 504 e 516. É considerada um ancestral comum de outras seis demonstradas na Figura 4.1 e possui uma correspondência exata em Portugal, o que possibilita inferir da existência de uma migração pontual possivelmente desta população para o Oeste da Europa. Poderá ainda ter existido um ancestral comum, não identificado nesta amostra populacional, de onde derivaram as linhagens identificadas nos indivíduos 32 e 65. Para qualquer uma destas linhagens não existe nenhuma correspondência exata o que sugere que um ancestral comum foi introduzido na população num passado mais longínquo tendo posteriormente divergido nestas duas linhagens. As restantes poderão ter sido introduzidas mais recentemente e ao longo do tempo, não sendo possível datar a sua entrada na população de judeus curdos, objeto do nosso estudo.

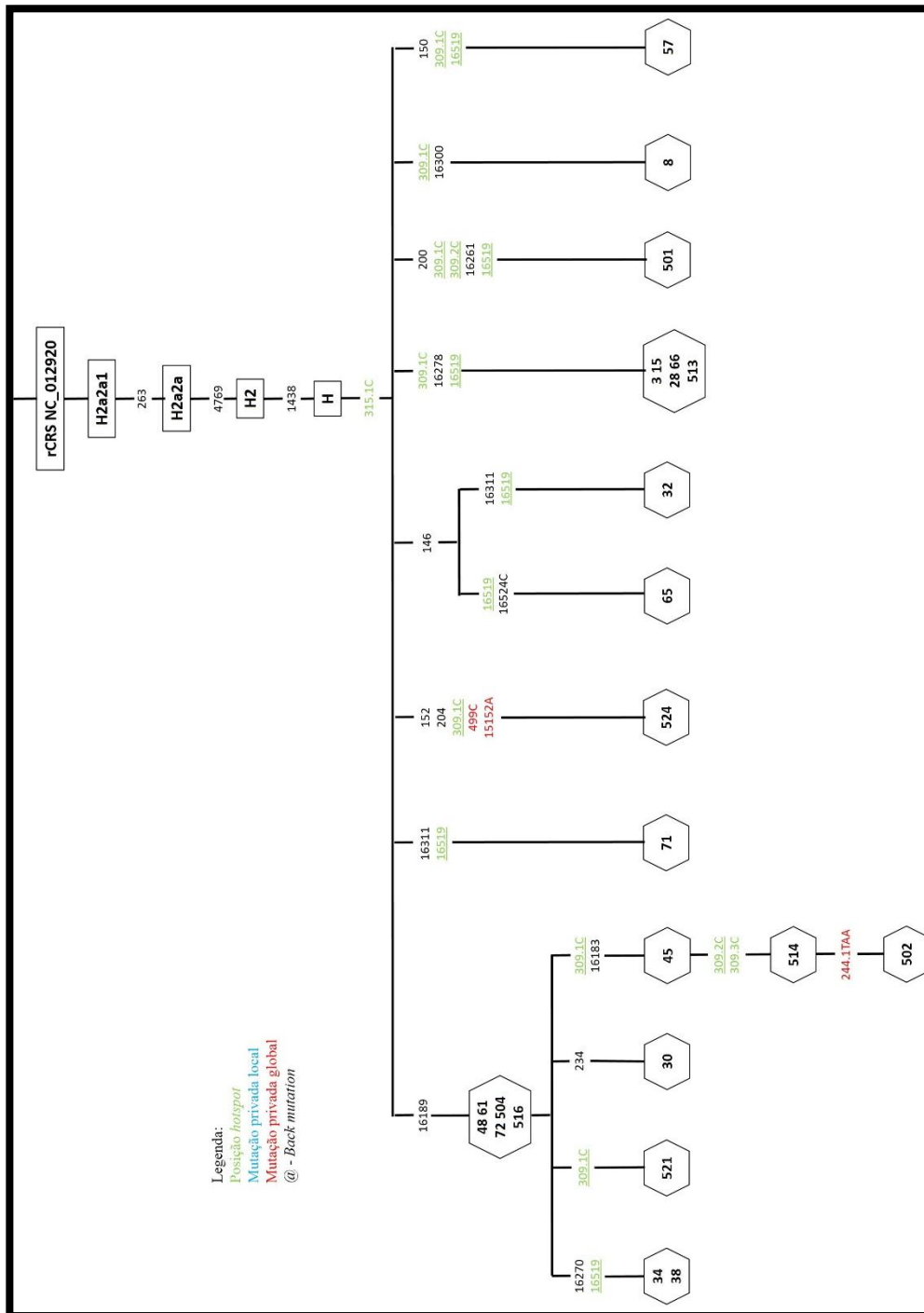


Figura 4.1: Árvore Filogenética do macrohaplogrupo H com os respectivos haplótipos presentes nos judeus curdos. Os números inscritos nos hexágonos correspondem à identificação das amostras classificadas.

4.1.1.2. Macrohaplogrupo J

Este macrohaplogrupo representa o segundo grupo mais predominante nos judeus curdos. O macrohaplogrupo J é, então, caracterizado pelas transições nas posições C295T, T489C e C16069T (Figura 4.2). Na nossa população foram encontradas três

linhagens distintas do subhaplogrupo J1b: J1b1b, J1b2 e J1b6. A linhagem J1b1b caracteriza-se por uma pela transição na posição C271T. A linhagem J1b2 é definida pela transição na posição C1733T e a linhagem J1b6 pela transição na posição A5501G. No subhaplogrupo J1c foi encontrada a linhagem J1c1, definida pela transição na posição T482C. O haplogrupo J2 está representado na nossa população por uma única linhagem, J2a2b2, definida pela transição na posição C456T.

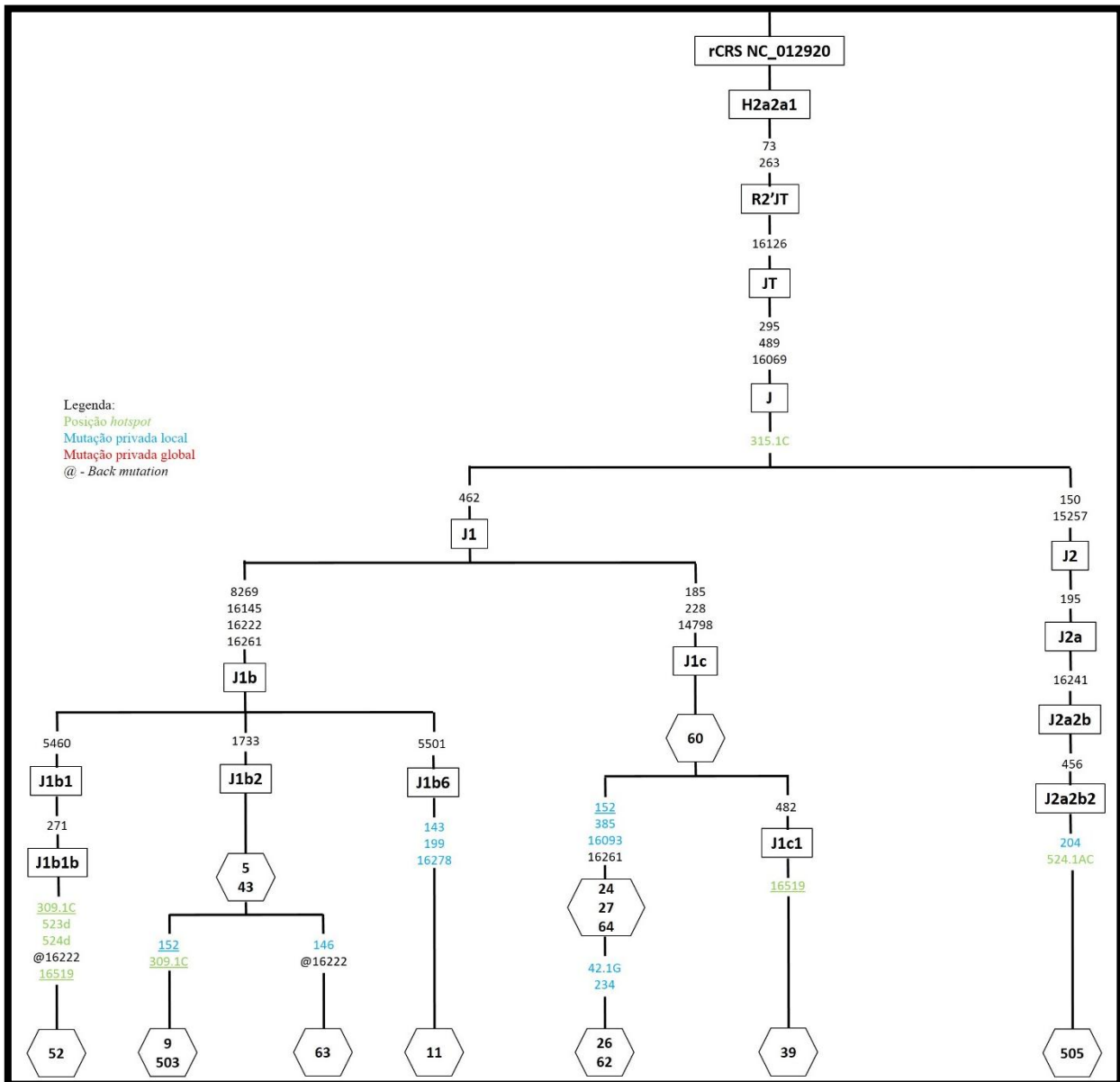


Figura 4.2: Árvore Filogenética do macrohaplogrupo J com os respectivos haplótipos presentes nos judeus curdos. Os números inscritos nos hexágonos correspondem à identificação das amostras classificadas.

O macrohaplogrupo J é encontrado na Europa e no Médio Oriente, compreendendo a 9% e 13% da totalidade dos haplogrupos, respetivamente. Quanto ao

haplogrupo J1, que teve origem há cerca de 33 mil anos [146], este representa cerca de 80% de todas as linhagens J e é encontrado nas mesmas regiões já descritas, no entanto com maior incidência no Médio Oriente, atingindo cerca de 16% no Yémen e 12% na Arábia Saudita [146-148]. Quanto ao haplogrupo J2, que teve origem há cerca de 37 mil anos [146], é representado no Médio Oriente e parte do Norte de África (apesar de ser muito mais raro que o anterior), atingindo maior incidência na Líbia com cerca de 7% [146, 149].

Nos judeus curdos, podemos supor que as linhagens do macrohaplogrupo J tiveram diferentes introduções em diferentes escalas temporais. Algumas delas possuem acumulação de mutações privadas e por isso podem constituir parte das linhagens fundadoras da população, enquanto outras tiveram uma evolução paralela às já existentes e outras ainda possam ter sido introduzidas mais recentemente.

Segundo a EMPOP, não foram encontradas correspondências exatas para as linhagens J1b1b, J1b2 (identificada no indivíduo 63), J1b6, J1c (identificada nos indivíduos 24, 26, 27, 62 e 64) e J2a2b2. Para as restantes linhagens do haplogrupo J1b2 foram encontradas sete correspondências exatas em Portugal, cinco no Kuwait, uma no Iraque e uma no Dubai. Para as restantes linhagens J1c foram encontradas seis correspondências exatas na Alemanha, seis na Suécia, quatro em Portugal e duas na Hungria (Tabela S.4). Quanto aos haplótipos com uma ou duas diferenças, para a linhagem J1c identificada nos indivíduos 26 e 62 não foram encontrados outros relacionados. Para as restantes linhagens J1c foram encontrados três haplótipos no Médio Oriente, 11 no Sul da Ásia (nomeadamente no Uzbequistão) e 156 na Europa (com maior incidência na Alemanha e na Roménia). Para a linhagem J1b1b foram encontrados três no Kuwait, dois no Iraque, dois no Dubai e dois no Paquistão. Para todas as linhagens J1b2 foram encontrados 43 haplótipos relacionados no Médio Oriente, seis no Sul da Ásia e 29 na Europa. Para a linhagem J1b6 foram encontrados seis em Portugal, três no Kuwait e um no Iraque. Por fim, para a linhagem J2a2b2 foram apenas encontrados dois haplótipos parecidos nos Estados Unidos (Tabela S.5).

Assim, é natural assumir que as linhagens J1b1b, J1b2 (identificadas nos indivíduos 5 e 43) e J1c (identificadas no indivíduo 60) possam ter feito parte das linhagens fundadoras da nossa população. Estas podem ser os fundadores ancestrais tanto na própria população dos judeus curdos, como na região. A partir destas ter-se-ão desenvolvido linhagens “secundárias” posteriores com a acumulação de mutações

privadas. Quanto a J1b6, teve uma introdução na população mais recente a partir do Kuwait ou do Iraque, apesar da sua evolução a ter afastado das restantes linhagens da região. J1c1 pode representar uma introdução muito mais recente a partir da Europa, de onde foi originária há cerca de 11 mil anos [146]. Para terminar, a linhagem J2a2b2 pode também ter tido uma introdução relativamente recente, já que sua origem é também recente e está localizada no Reino Unido, apesar de pouco mais se saber sobre a mesma [99].

4.1.1.3. Macrohaplogrupo K

As linhagens do macrohaplogrupo K encontradas nos judeus curdos são caracterizadas pelas transições nas posições T16224C e T16311C (Figura 4.3). Nos judeus curdos foi possível identificar duas linhagens distintas do haplogrupo K1. A linhagem K1a5b é definida pela transversão na posição T408A e a linhagem K1a9 pela transição na posição A16524G. Dentro do haplogrupo K2 foi possível identificar apenas uma linhagem, K2a2a1, definida pela transversão na posição A512C.

O haplogrupo K encontra-se bem representado ao longo de toda a Europa e um pouco pelo Norte de África e Sul da Ásia. No Médio Oriente, atinge percentagens médias de 6% e na Europa chega aos 10% [90]. Para além desta distribuição, cerca de 32% da ascendência dos judeus Ashkenazi são do macrohaplogrupo K. Segundo Fernández [150] e Costa [151], existem três subhaplogrupos fundadores destes judeus dentro do haplogrupo K: K1a1b1a, K1a9 e K2a2, sendo que estes dois últimos foram identificados na nossa população de judeus curdos.

É possível alegar que a linhagem K1a5b poderá ter tido uma introdução mais remota, em parte devido à acumulação de outras mutações às definidas para a classificação deste haplogrupo, ao passo que as linhagens K1a9 e K2a2a1 podem ser remetidas para uma introdução mais recente. Estas duas linhagens são as linhagens fundadoras dos judeus Ashkenazi que estão presentes na nossa população de judeus curdos. Resta ainda a hipótese de que estas últimas duas linhagens terem sofrido introgressão de mulheres de outros grupos étnicos ou de outras regiões, tendo sido “incorporadas” na população de judeus curdos.

Segundo a EMPOP, a linhagem K1a5b não possui nenhuma correspondência exata dos haplótipos identificados nos judeus curdos. Para a linhagem K1a9, a base de

dados encontrou oito correspondências exatas na Hungria e uma na Áustria. Por sua vez, a linhagem K2a2a1 possui 14 correspondências exatas também na Hungria (Tabela S.4). Quanto aos haplótipos com uma ou duas diferenças, para as linhagens derivadas do K1 foram encontrados similares na Macedónia, na Alemanha, no Norte de África entre outras regiões. Para a linhagem derivada do K2, foram encontrados haplótipos com uma ou duas diferenças na Macedónia, Uzbequistão, Alemanha, Dubai, entre outros (Tabela S.5).

É possível que a linhagem K1a5b tenha sido introduzida na população primeiro que as restantes linhagens K, devido à acumulação de mutações ao longo da tempo e ao facto de não existirem correspondência exata na base de dados.

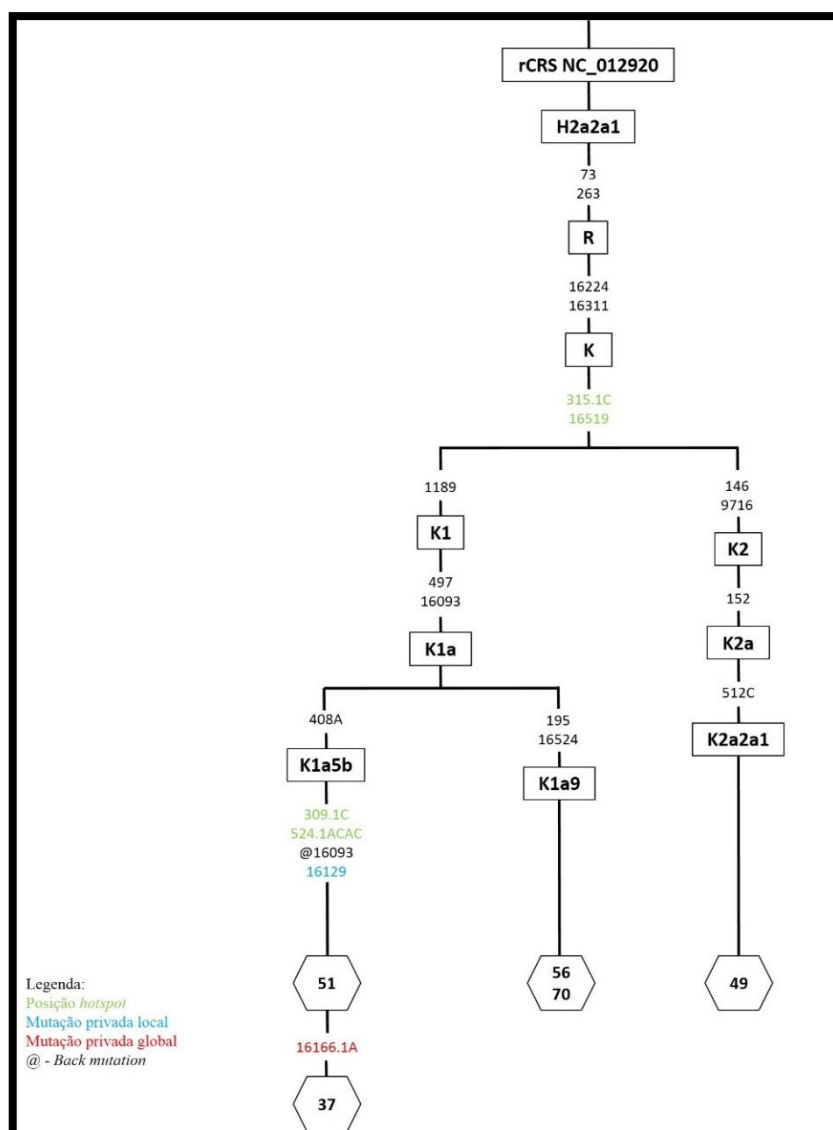


Figura 4.3: Árvore Filogenética do macrohaplogrupo K com os respetivos haplótipos presentes nos judeus curdos. Os números inscritos nos hexágonos correspondem à identificação das amostras classificadas.

4.1.1.4. Macrohaplogrupo M

O macrohaplogrupo M é caracterizado pela transição na posição T489C. O subhaplogrupo M1a1 foi a única linhagem identificada nos judeus curdos, e é definida pela transição na posição T16359C (Figura 4.4).

Segundo o estudo de González [152], o haplogrupo M1 é encontrado, maioritariamente, entre o Norte de África e o Médio Oriente. Na zona do Norte de África, a sua frequência diminui de Este para Oeste e de Norte para Sul, enquanto no Médio Oriente está representado apenas com baixas frequências em toda esta região.

A presença de apenas uma amostra da linhagem M1a1 revela uma entrada recente ou um vestígio da altura da migração dos haplogrupos M1 e U6 para o Norte de África ocorrido provavelmente há cerca de 30 mil anos, como já foi referido anteriormente.

Segundo a EMPOP, não foi encontrada nenhuma correspondência exata deste haplótipo específico (Tabela S.4). No entanto, no Egito e no Kuwait, a base de dados encontrou um haplótipo em cada uma destas populações com uma ou duas diferenças do haplótipo encontrado nos judeus curdos (Tabela S.5). Estas diferenças podem, possivelmente, suportar a ideia de que tenha havido uma migração de volta ao Norte de África em que esta teria sido derivado de um dos fundadores do haplogrupo M1. O haplótipo encontrado nos judeus curdos mostra uma acumulação de mutações privadas que indiciam um desenvolvimento paralelo relativamente às restantes linhagens encontradas tanto no Médio Oriente como no Norte de África.

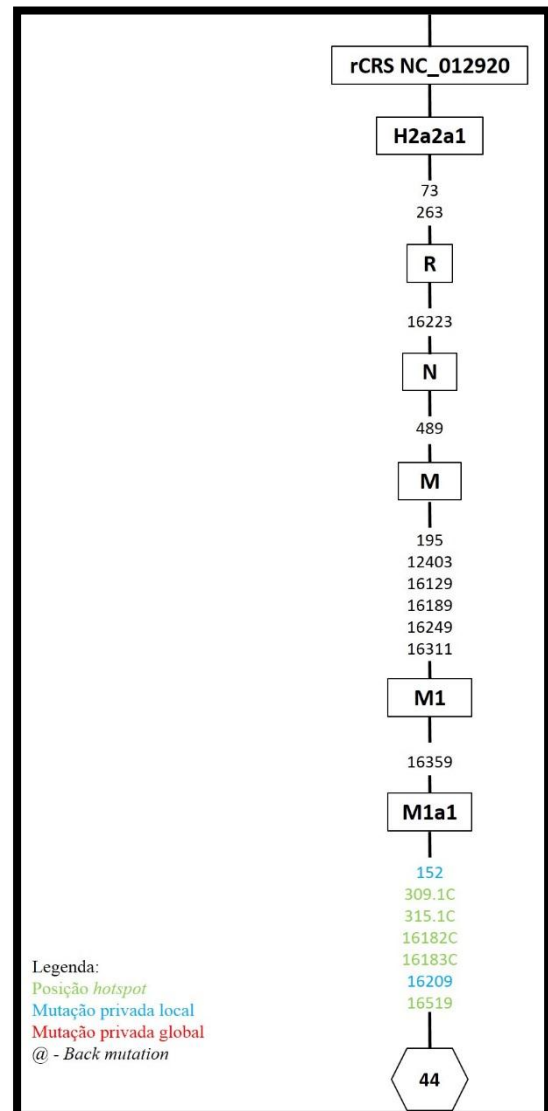


Figura 4.4: Árvore Filogenética do haplótipo M1a1, presente nos judeus curdos. O número inscrito no hexágono corresponde à identificação da amostra classificada.

4.1.1.5. Macrohaplogrupo N

O macrohaplogrupo N foi confirmado pelas transições nas posições C12705T e C16223T (Figura 4.5). No subhaplogrupo N1a foi apenas encontrada uma linhagem N1a3a, definida pela transição na posição C16201T. No subhaplogrupo N1b foi identificada igualmente uma só linhagem, N1b1, definida pela transição na posição G16145A. Do haplogrupo N2, apenas foi identificado um dos seus subhaplogrupos, N2a, caracterizado por transições nas posições T199C, C739T, G16153A e G16319A.

O macrohaplogrupo N, por ser um dos responsáveis pelo *Out-of-Africa*, está presente em toda a Eurásia, Norte de África e Oceânia. Segundo Fregel [85], o seu haplogrupo N1 divergiu há cerca de 51.9 mil anos e encontra-se presente ao longo do Oeste da Eurásia e Norte de África, enquanto que o haplogrupo N2 teve origem há cerca de 48.3 mil anos encontrando-se presente no Oeste e Sul da Eurásia. Adicionalmente, a linhagem N1b representa ainda 9.2% dos ascendentes dos judeus Ashkenazi. Segundo Costa [151], esta faz ainda parte de uma das linhagens fundadoras destes judeus, juntamente com as já descritas do macrohaplogrupo K.

Nos judeus curdos, cada uma das três linhagens pode ter tido introduções diferentes e relativamente recentes na população.

Segundo a base de dados EMPOP, não foi encontrada nenhuma correspondência exata para qualquer das linhagens do macrohaplogrupo N presentes na população de judeus curdos (Tabela S.4). Quanto aos haplótipos com uma ou duas diferenças, para a linhagem N1a3a foram encontradas 14 sequências semelhantes no Kuwait. Para a linhagem N1b1, foram encontrados um na Macedónia, um em Marrocos e um no Egito. Quanto à linhagem N2a, foi encontrado um no Paquistão (Tabela S.5). Mais uma vez se constata que estas linhagens podem ter tido introduções independentes ao longo do tempo, no entanto, relativamente recentes. A linhagem N1a3a pode ter proveniência do Kuwait e ter seguido uma evolução paralela, facto que é suportado pela presença na nossa população de curdos judeus de dois indivíduos distintos com o haplótipo idêntico. O mesmo se pode aferir para a linhagem N1b1, com oito indivíduos com o mesmo haplótipo. É igualmente possível que neste caso possa ter havido um ancestral comum, já que N1b terá sido linhagem fundadora dos judeus Ashkenazi, a partir da qual um segmento evoluiu na população dos judeus curdos e outro se dirigiu para o Norte de

África. A linhagem N2a pode ter tido uma introdução independente, não sendo possível distinguir claramente a sua região de origem.

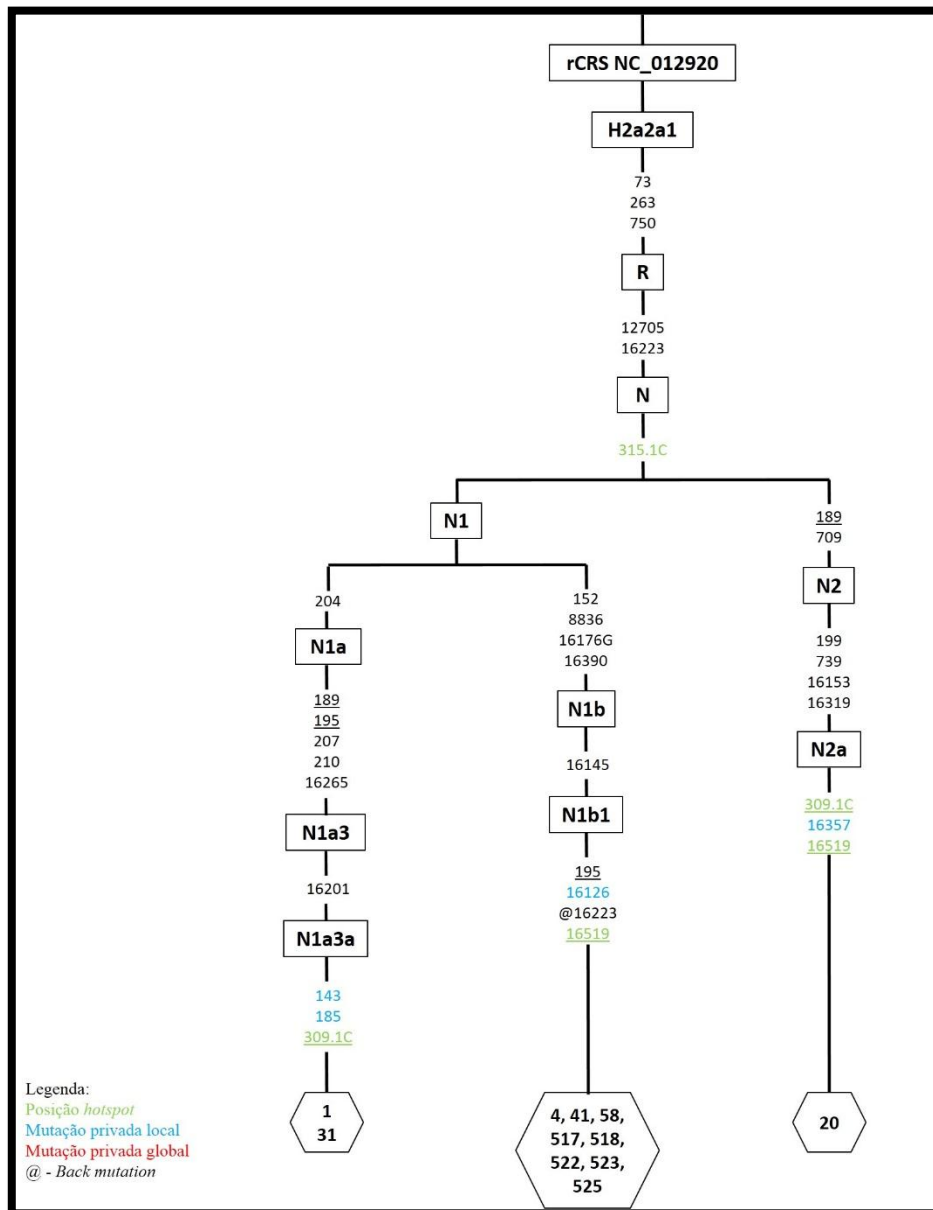


Figura 4.5: Árvore Filogenética do macrohaplogrupo N com os respectivos haplótipos presentes nos judeus curdos. Os números inscritos nos hexágonos correspondem à identificação das amostras classificadas.

4.1.1.6. Macrohaplogrupo R

Nos judeus curdos apenas foi identificada uma linhagem do macrohaplogrupo R, R0a1a. A linhagem R0a1a é, então, definida pelas transições nas posições T146C e C16355T (Figura 4.6).

O macrohaplogrupo R é encontrado em vastas áreas, desde a Europa até o sul da Ásia, incluindo Norte de África e Médio Oriente. Quanto ao haplogrupo R0, este teve uma origem há cerca de 26.4-35 mil anos, possivelmente, na Península Arábica [153]. Encontra-se bem representado, principalmente, no Médio Oriente e Sul da Europa, apresentando uma das frequências mais altas em Soqatra, Iémen (cerca de 38%) [154].

De modo semelhante à situação do macrohaplogrupo M, esta linhagem é representada em apenas um indivíduo, e remete tanto para uma introdução recente como um vestígio das primeiras linhagens deste haplogrupo.

Segundo a EMPOP, este haplótipo tem seis correspondências exatas no Kuwait, uma na Alemanha e uma na Argentina (Tabela S.4). Foram ainda encontrados haplótipos com um ou dois polimorfismos diferentes no Médio Oriente (Kuwait, Bahrain e Dubai), Norte de África (Egito) e Sul da Europa (Grécia, Roménia e Hungria) (Tabela S.5). Deste modo, com as comparações das correspondências exatas e assumindo que não possui mutações privadas acumuladas, pode-se afirmar que, possivelmente, esta linhagem é remetida para uma introdução mais recente na população dos judeus curdos, com possível origem no Kuwait.

4.1.1.7. Macrohaplogrupo T

O macrohaplogrupo T é caracterizado pelas transições nas posições C16294T da região controlo e nas posições T10463C e G13368A da região codificante (Figura 4.7). O haplogrupo T1, única linhagem identificada nos judeus curdos, é caracterizado pela transição na posição A16163G e pela transversão na posição C12633A. Com origem no

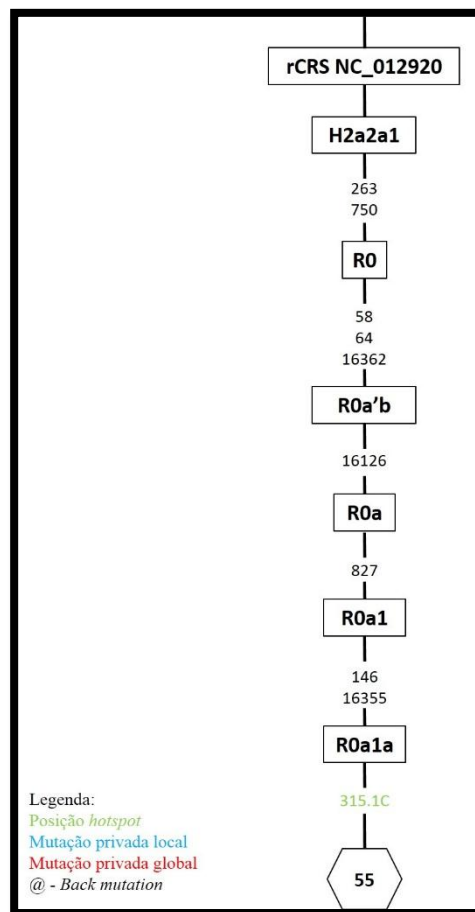


Figura 4.6:  rvore Filogen tica do hapl tipo R0a1a, presente nos judeus curdos. O n mero inscrito no hex gono corresponde   identificaç o da amostra classificada.

haplogrupo T2, foram identificados dois subhaplogrupos nos judeus curdos: T2b e T2c. No primeiro destes foi identificada a linhagem T2b3 definida pela transição na posição A10750G. Dentro do subhaplogrupo T2c foi identificada a linhagem T2c1 definida pela transição na posição C16292T.

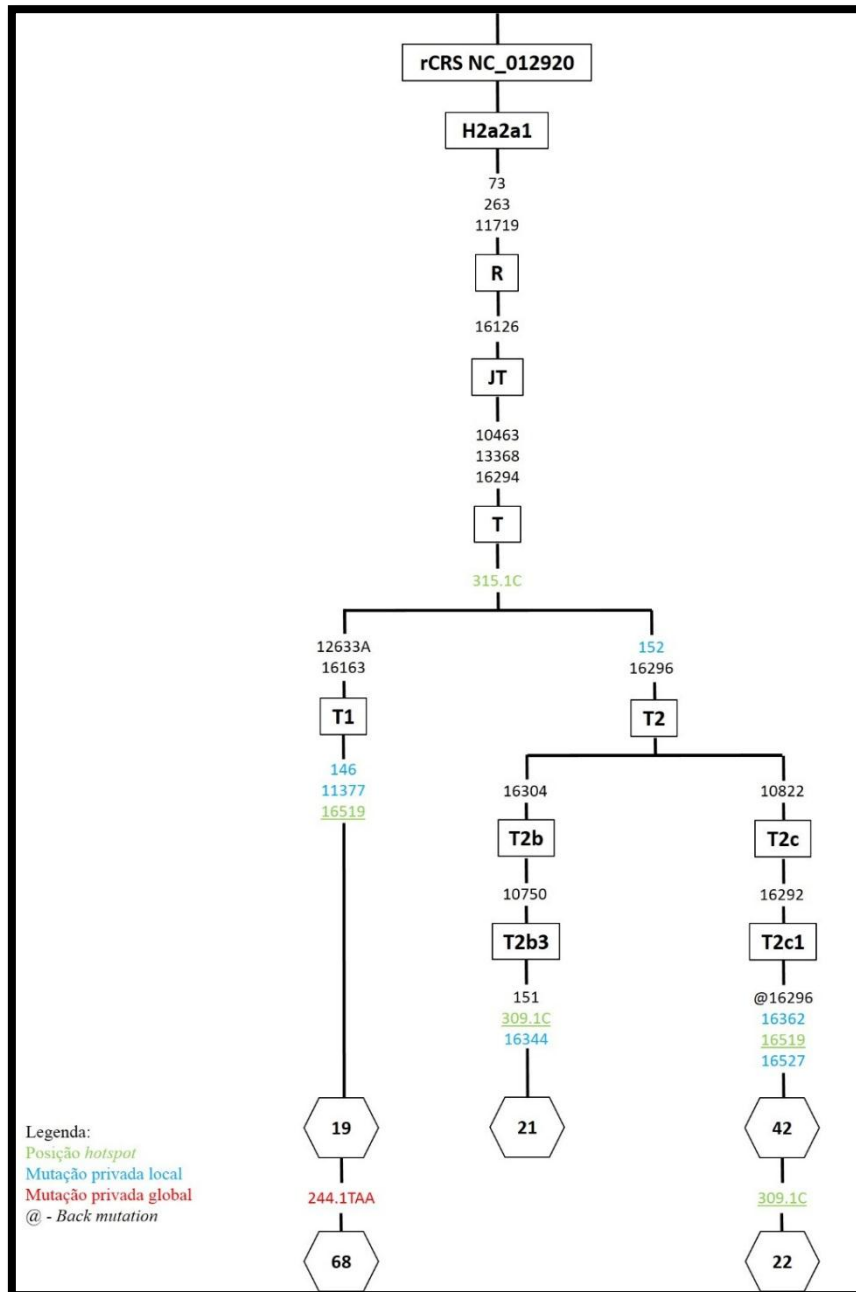


Figura 4.7: Árvore Filogenética do macrohaplogrupo T com os respectivos haplótipos presentes nos judeus curdos. Os números inscritos nos hexágonos correspondem à identificação das amostras classificadas.

O macrohaplogrupo T encontra-se representado ao longo do Oeste e Centro da Eurásia, constituindo cerca de 10% dos haplogrupos na Europa e cerca de 8% no Médio Oriente [146]. Segundo Pala [146], os haplogrupos T1 e T2 tiveram origem há cerca de

21 mil anos, sendo que o haplogrupo T2 possa ser mais antigo com presença na Europa. O haplogrupo T1 está representado em cerca de 2% do Oeste da Europa e cerca de 3% no Este da Europa e Médio Oriente. Quanto ao haplogrupo T2, que representa cerca de 80% de todas as linhagens T, este está representado em apenas 8% do Oeste da Europa e 5% do Médio Oriente [146].

A acumulação de mutações em todas as linhagens T encontradas nos judeus curdos permite afirmar que a sua introdução nesta população tenha sido mais remota. É de notar ainda que nesta população já houve derivações dos haplótipos mais ancestrais, tanto na linhagem T1, com a inserção de TAA na posição 241, como na linhagem T2c1, com a inserção de C na posição 309. Podemos ainda afirmar que algumas destas linhagens possam ter sido fundadores desta população.

De acordo com a EMPOP, não foi encontrada nenhuma correspondência exata de quaisquer dos haplótipos identificados nesta população (Tabela S.4). Contudo, foram encontrados haplótipos com uma ou duas diferenças para todas as linhagens. Para a T1, foram encontrados três haplótipos parecidos no Paquistão, três na Alemanha, um na Macedónia e um na Grécia. Para a T2b3, foram encontrados cinco nos Estados Unidos e um na Argentina. Para a T2c1, foram encontrados um na Hungria, um na Roménia e um em Marrocos (Tabela S.5). Uma vez que não se encontraram correspondências exatas, podemos talvez inferir que a sua introdução é mais remota na população, tendo sido paralela a evolução destas linhagens em relação às da região. Será possível que a linhagem mais antiga na população tivesse sido a T2c1, seguindo-se a introdução de T1 e, por fim, a T2b3?

4.1.1.8. Macrohaplogrupo U

O macrohaplogrupo U é caracterizado pelas transições nas posições A11467G, A12308G e G12372A (Figura 4.8). Nos judeus curdos foram encontradas apenas duas linhagens. Dentro do subhaplogrupo U1a foi identificada uma única linhagem, o U1a1a, definida pelas transições A385G e G3591A, e pela inserção de uma base T na posição 3158. Quanto ao haplogrupo U7, foi igualmente identificada uma linhagem apenas, U7a, definida pela transição na posição C151T.

O macrohaplogrupo U está representado em toda a Eurásia e no Norte de África. Mais especificamente, o seu derivado U1 é encontrado a baixas frequências na Europa,

sendo mais frequente na zona Este da Europa, Médio Oriente e Sul da Ásia [155]. Por outro lado, o haplogrupo U7 encontra-se mais restringido ao Oeste da Eurásia, incluindo o Médio Oriente, atingindo as frequências mais altas neste último, com cerca de 6% no Irão [156].

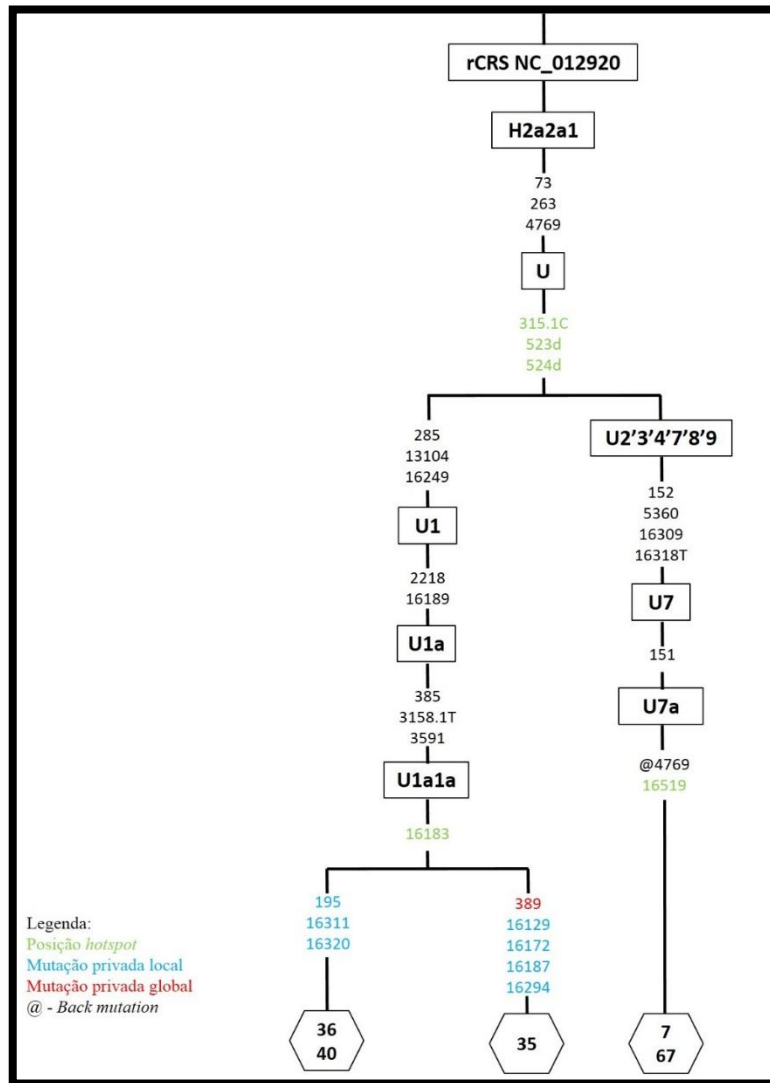


Figura 4.8:  rvore Filogen tica do macrohaplogrupo U com os respetivos hapl tipos presentes nos judeus curdos. Os n meros inscritos nos hex gonos correspondem   identificaç o das amostras classificadas.

Pela observa o das linhagens identificadas nos judeus curdos, podemos supor que as linhagens U1a1a tenham sido introduzidas, atrav s de um ancestral comum, primeiramente nos judeus curdos. A linhagem U7a pode, portanto, representar uma introdu o mais recente atrav s da incorpora o de mulheres de outras regi es.

Na EMPOP n o foram encontradas correspond ncias exatas para qualquer uma das linhagens U1a1a. Quanto  s linhagens U7a, foram encontradas duas correspond ncias

no Kuwait, uma no Bahrain e uma no Iraque (Tabela S.4). Quanto aos haplótipos com uma ou duas diferenças, as linhagens observadas nos indivíduos 36 e 40 possuem um relacionado no Kuwait e um nos Estados Unidos e a linhagem do indivíduo 35 possui apenas um em Marrocos. As linhagens observadas nos indivíduos 7 e 67 possuem 30 haplótipos relacionados no Paquistão, 13 no Uzbequistão, três no Kuwait e dois no Bahrain (Tabela S.5). Deste modo, podemos afirmar que a linhagem U1a1a foi introduzida na população antes da linhagem U7a, devido a não existirem correspondências exatas e devido à acumulação de mutações privadas. Devido à quantidade de mutações privadas e à existência de duas linhagens paralelas provenientes de um ancestral comum, não detetado nesta amostra populacional, podemos supor que este ancestral comum poderá até ser uma possível linhagem fundadora nos judeus curdos. Quanto à linhagem U7a, é possível que esta tenha tido origem no Kuwait e, posteriormente, ter sido introduzida nesta população.

4.1.1.9. Macrohaplogrupo X

O macrohaplogrupo X é caracterizado pelas transições nas posições T16189C e C16278T. Dentro deste, foi apenas identificado um único haplogrupo, o X2, definido pelas transições nas posições T195C e G1719A (Figura 4.9).

Como já foi descrito, o haplogrupo X2 está representado pela Europa e Médio Oriente, atingindo uma das percentagens mais altas na população Druze, da Geórgia e das ilhas Orkney, com 11.8%, 8% e 7%, respetivamente [108]. Em Portugal reside uma das percentagens mais altas do haplogrupo X da Europa, atingindo 5.5% [157].

Assim como para o haplogrupo U1, todas as linhagens identificadas nos judeus curdos são derivadas do haplogrupo X2. Com a acumulação de mutações privadas, é exetável que a sua introdução na população tenha sido mais remota, podendo até fazer parte das linhagens fundadoras desta população.

Na EMPOP não existem correspondências exatas dos haplótipos identificados nos judeus curdos (Tabela S.4). Quanto aos haplótipos com uma ou duas diferenças, para a linhagem identificada no indivíduo 69 foi encontrado um haplótipo na Suécia e para as restantes foi encontrado um haplótipo no Kuwait (Tabela S.5). Deste modo, podemos de facto afirmar que a sua introdução foi mais primordial, podendo ser uma das linhagens fundadoras da população. Já que não existe nenhuma correspondência exata e muito

poucos haplótipos parecidos, é notável que de um ancestral comum da região se tenham formado várias ramificações em que duas delas evoluíram de forma paralela nesta população.

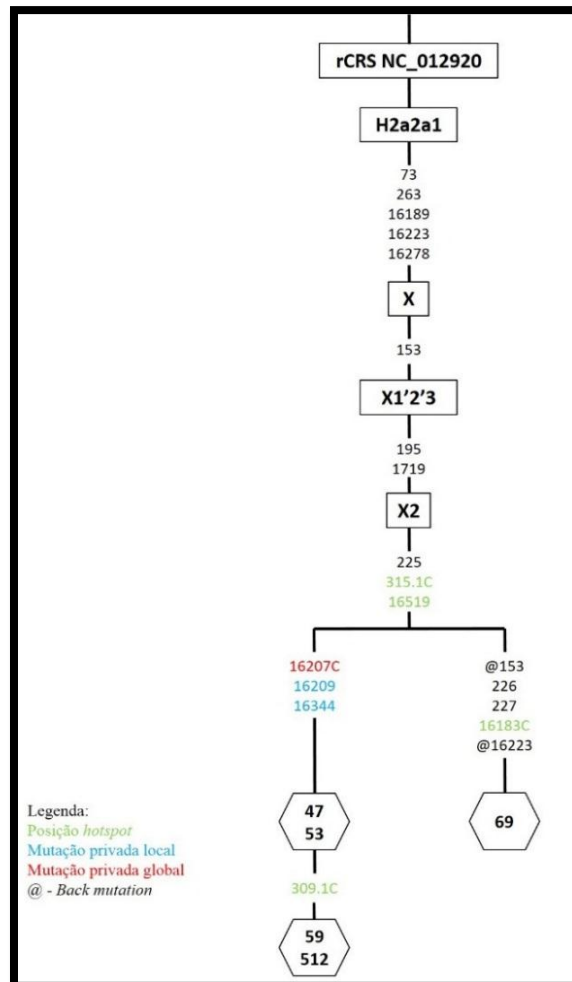


Figura 4.9: Árvore Filogenética do macrohaplogrupo X com os respectivos haplótipos presentes nos judeus curdos. Os números inscritos nos hexágonos correspondem à identificação das amostras classificadas.

4.2. Análise da variância molecular e distâncias genéticas

Através da Análise da Variância Molecular (AMOVA) investigou-se a estrutura genética dos judeus curdos refugiados em Israel, comparando-a com populações da Ásia, Norte de África e Europa, e com outras populações judaicas, agrupando as respectivas populações segundo a sua localização geográfica (Tabela 4.3).

Tabela 4.3: AMOVA das populações judaicas, não judaicas e de todas as populações, baseada nas frequências dos haplogrupos de cada população.

Grupos	Intrapopulações			Entre populações dentro dos grupos			Entre grupos		
	%	F _{ST}	P	%	F _{SC}	P	%	F _{CT}	P
Judeus ¹	84.82	0.15180	0.00000	12.28	0.12648	0.00000	2.9	0.02899	0.03211
Não-Judeus ²	95.48	0.04515	0.00000	2.94	0.02989	0.00000	1.57	0.01573	0.00099
Judeus + Não-Judeus ³	93.57	0.06433	0.00000	5.47	0.05519	0.00000	0.96	0.00968	0.00000

¹Grupos: (KJ, IrJ, IrqJ); (AzbJ, GeoJ); (EthJ); (BJ, CJ); (MarJ, TnJ, LbJ); (TurJ, BgJ, BelJ); (YeJ)

²Grupos: (KJ, Ir, Irq, Sir, Tur); (ArS, Ye, Jor); (Quig, Uzb, Tqm, Afg, Czq); (Alg, Lb, Tn, Egt); (Mng); (Geo, Azb); (Bg)

³Grupos: (KJ, Ir, Irq, Sir, Tur, IrJ, IrJ, TurJ); (ArS, Ye, Jor, YeJ); (Quig, Uzb, Tqm, Afg, Czq); (Alg, Lb, Tn, Egt, EthJ, LbJ, MarJ, TnJ); (Mng, BJ, CJ); (Geo, Azb, GeoJ, AzbJ); (Bg, BelJ, BgJ)

Usualmente, recorre-se à Análise de Variância Molecular para a determinação da estrutura genética de determinadas populações. Utilizam-se marcadores moleculares, neste caso, frequências dos respetivos haplogrupos mitocondriais, sendo o próprio utilizador que define uma determinada estrutura genética ao definir grupos de populações. Uma análise hierarquizada da variância divide a variância total em componente de covariância devido a diferenças intra-indivíduos, diferenças inter-indivíduos e/ou diferenças interpopulações [158]. Assim, a estrutura genética das populações judaicas exibiu que 84.82% da variação encontrada deve-se a diferenças no seio das próprias populações, ou seja, variação intrapopulacional. Verificou-se uma variação elevada dentro dos grupos formados (12.28%) decaindo, consideravelmente, quando se investiga a variação entre os mesmos (2.9%). A ocorrência de uma variação genética intrapopulacional leva-nos a aprofundar o facto com o auxílio das PCAs, no ponto 4.4.

No que concerne às populações não-judaicas que tiveram determinada influência sobre o Curdistão ao longo da história, observou-se que 95.48% da variação surge de diferenças intrapopulacionais. Como, identicamente, se verificou com as populações judaicas mas com menor diferença, também as populações não-judaicas apresentaram uma maior variação dentro dos grupos do que entre si (2.94% e 1.57%, respetivamente). Tais resultados serão, posteriormente, comparados e discutidos com o auxílio das PCAs, no ponto 4.4. Quando se englobam todas as populações acima mencionadas num estudo apenas, a variação genética observada dentro das populações é de 93.57%, dentro dos grupos de 5.47% e entre grupos de 0.96%. Todos os valores apresentados na Tabela 4.3 são significativos ($P < 0.05$).

4.3. Análise das Componentes Principais

Toda a informação genética obtida e posteriores comparações podem ser apresentadas por meio de gráficos, tornando assim mais fácil a sua compreensão. Neste trabalho, estas informações serão apresentadas sob a forma de Análise de Componentes Principais (PCA). A PCA é um método que utiliza a redução de dimensões de modo a explicar a variação dos dados de determinadas variáveis (as componentes principais). Após redução das variáveis a componentes principais, as coordenadas são inseridas num gráfico dimensional de modo a representar um panorama genético. Para n populações, são necessárias $n-1$ dimensões para representar totalmente as divergências genéticas com reduzida perda de informação. Comumente é utilizado um gráfico com duas dimensões, que expõe uma boa distribuição das semelhanças relativas genéticas, sendo satisfatória quando a variação retida ronda os 60-75%. Não obstante demonstrarem resultados semelhantes às árvores de “clusters”, as PCAs não devem ser utilizados como método único de análise [142].

A base molecular dos polimorfismos possibilita a determinação de distâncias evolutivas baseadas no número de diferenças dos mesmos. A mais simples distância genética entre duas populações para um gene bialélico seria a diferença de frequências alélicas p_i e p_j [142, 159]. Deste modo, a relação evolutiva das populações, em termos de estrutura genética, pode ser medida através de índices F_{ST} . A distância genética avalia, assim, a média da frequência génica e da variação que lhe está associada, podendo incorporar as distâncias moleculares entre os haplogrupos. O índice F_{ST} varia entre os valores zero e um, sendo que os mais altos indicam uma maior distância evolutiva, ou seja, um índice de 0.3 significa que 30% da frequência alélica é atribuída a diferenças interpopulacionais, enquanto que a restante variação é encontrada dentro da própria população. A hipótese nula reflete ausência de diferença entre as populações, cuja significância é testada pelo teste de significância P (rejeição da hipótese nula quando $P < 0.05$).

Com o objetivo de analisar as relações genéticas entre os judeus curdos do presente estudo com populações europeias, norte africanas e asiáticas descritas na literatura, efetuou-se uma Análise de Componentes Principais (PCA) utilizando as respetivas frequências haplotípicas. Foram obtidas duas PCAs refletindo as relações entre os judeus curdos e as populações que tiveram influência sobre o Curdistão, e entre os judeus curdos e outras populações judaicas (Figuras 4.10 e 4.11, respetivamente).

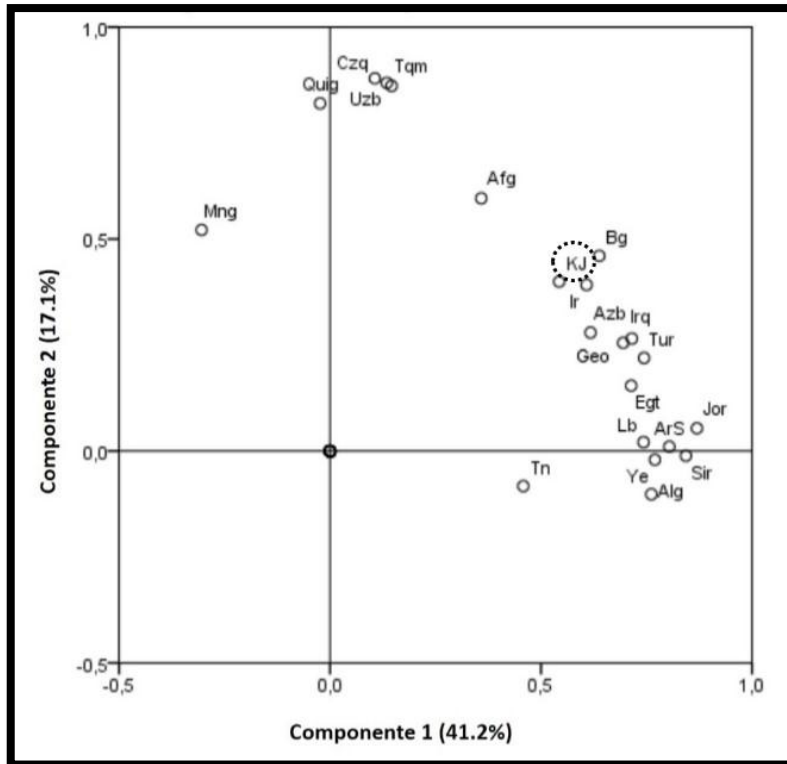


Figura 4.10: Mapa bidimensional da PCA obtida para as relações dos Judeus Curdos com outras populações que possam ter tido alguma influência sobre o Curdistão ao longo da sua história, baseada nos dados de DNAm. As duas primeiras componentes representam 58.3% da variância total: 41.2% na primeira componente e 17.1% na segunda componente. Abreviaturas estão como descritas na Tabela S.1.

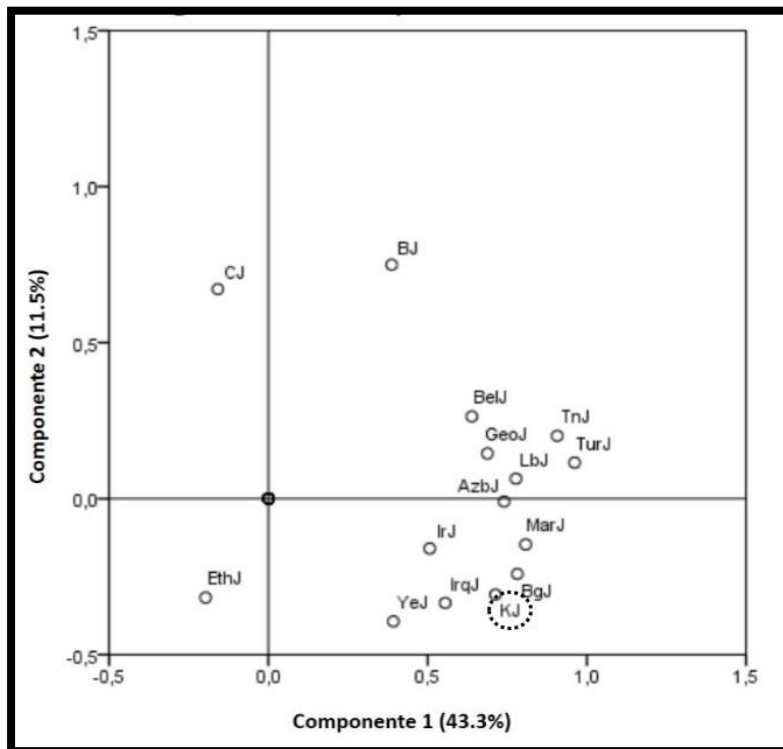


Figura 4.11: Mapa bidimensional da PCA obtida para as relações dos Judeus Curdos com outras populações judaicas, baseada nos dados de DNAm. As duas primeiras componentes representam 54.8% da variância total: 43.3% na primeira componente e 11.5% na segunda componente. Abreviaturas estão como descritas na Tabela S.1.

Da análise das PCAs, podemos inferir a relação genética das populações em estudo através da posição topográfica que ocupam no espaço. Na primeira PCA (Figura 4.10), que inclui a variação genética das populações que tiveram influência no Curdistão, as populações agrupam-se de acordo com a sua localização geográfica, emergindo claramente dois grupos principais: Norte de África/Médio Oriente e Ásia. Apesar desta distribuição, subsiste ainda uma certa uniformidade com todas as populações agrupadas num só quadrante. Esta proximidade topográfica vem espelhar uma certa homogeneidade genética entre elas. Como descrito anteriormente, por ter sido um ponto de passagem entre África e a Europa/Ásia, o Médio Oriente, assim como a população dos curdos, terá, certamente, algumas características africanas e outras asiáticas, o que realmente é patente na topografia das PCAs. O agrupamento das populações do Médio Oriente encontra-se entre os outros dois agrupamentos, indicando que não tem apenas características de um ou do outro grupo. A maior heterogeneidade verificada nesta PCA dentro dos agrupamentos deve-se, principalmente, às populações asiáticas do Afeganistão e da Mongólia. Curiosamente, a população da Síria aproxima-se mais das populações do Norte de África do que do Médio Oriente. O facto de não ser significativamente separada pela primeira componente poderá explicar esta situação (dados não descritos). A comparação dos judeus curdos com as restantes populações mostra semelhanças na estrutura genética, principalmente com as populações do Irão, da Bulgária e do Azerbaijão. Confrontando com os valores obtidos da AMOVA, podemos identificar uma variação genética baixa entre as populações dentro dos grupos, com as populações asiáticas apresentando a maior variação, como referido anteriormente. Assim, o agrupamento das populações segundo a sua localização geográfica verificada na PCA apoia a estrutura genética para a AMOVA. Entre os grupos verificou-se, igualmente, um valor baixo para a variação genética provocada pelo posicionamento próximo das populações agrupadas, maioritariamente, num único quadrante.

Quando a variabilidade se fundamenta em haplogrupos, os judeus curdos exibem uma variabilidade genética de 0.8294 ± 0.0293 . De modo geral, os índices da variação genética são constantes em todas as populações, variando apenas entre 0.88 e 0.97 (Tabela S.1). No Norte de África, os índices da variação genética tendem a decrescer do Este para Oeste. Nas restantes populações podemos identificar uma variação superior no Irão decrescendo, radialmente, nas populações circundantes. Os índices F_{ST} são todos significativos exceto para a relação entre o Azerbaijão e a Jordânia. Segundo estes índices,

os judeus curdos aproximam-se do Azerbaijão e Bulgária, e afastam-se da Geórgia, Tunísia e Mongólia. Tais relações estão representadas na respetiva PCA.

Na segunda PCA (Figura 4.11) averiguou-se a variação genética de várias populações judaicas. Verifica-se somente uma ligeira separação segundo a sua localização geográfica, menos evidente que na primeira PCA: populações judaicas da Índia num *cluster* e da Etiópia noutra *cluster*. Aqui, a separação entre judeus norte-africanos e os do Médio Oriente não é tão evidente. No estudo de Behar [160], do qual foram retirados os resultados genéticos dos judeus, as amostras foram recolhidas de vários grupos étnicos em Israel, cada indivíduo com indicação do local de nascimento da mãe, avó materna e da bisavó. Neste estudo, os judeus do norte de África pertenciam, no geral, ao grupo genético do oeste da Eurásia. De facto, na PCA observa-se uma mistura entre as populações destas duas regiões. Ostrer e Skorecki [161] afirmaram, igualmente, que as populações judaicas, à exceção das da Etiópia e da Índia, possuíam genomas mitocondriais com origem no Médio Oriente. Este facto é aparente na PCA, com uma boa separação das populações judaicas da Etiópia e da Índia em relação às restantes populações. Nas duas exceções descritas, parece ser mais provável que a culturização tenha aqui um papel primordial, muito maior que o atribuído a um possível fluxo genético, acompanhando assim a evolução populacional do local onde se encontravam [160, 161]. A análise desta PCA vem explicar os valores elevados dentro dos grupos obtidos pela AMOVA. Uma separação das populações por localização geográfica, utilizada como estrutura genética para o estudo em questão, verificou-se pouco esclarecedora. Os valores obtidos entre os grupos são, também, ligeiramente elevados quando comparados com os restantes resultados obtidos para este parâmetro, observando-se a razão, igualmente, na PCA. Assim, aquando da comparação com as populações judaicas, os judeus curdos aproximam-se, geneticamente, das populações judaicas da Bulgária e do Iraque.

Quando a variabilidade é sustentada pelos haplogrupos, agora em relação a outras populações judaicas, os judeus curdos apresentam, como foi descrito anteriormente, uma variabilidade genética de 0.8294 ± 0.0293 . De modo geral, os índices da variação genética são constantes em todas as populações variando entre 0.75 e 0.91, à exceção dos judeus de Belmonte (Portugal) com cerca de 0.1, os do Azerbaijão e da Geórgia com cerca de 0.6 e os de Bombay com cerca de 0.5 (Tabela S.1). No Norte de África, os índices da variação genética tendem a crescer do Este para o Oeste. Nas restantes populações podemos identificar a maior variação no Iémen e no Irão e vai

decrecendo radialmente nas populações circundantes. Os índices F_{ST} são todos significativos exceto para a relação entre os judeus da Turquia, da Tunísia, da Bulgária e de Marrocos e para a relação entre os judeus da Bulgária e do Curdistão. Segundo estes índices, os judeus curdos aproximam-se dos judeus da Turquia e da Bulgária e afastam-se dos judeus da Geórgia, do Bombay e de Belmonte, relações estas também representadas na respetiva PCA.

Capítulo 5 – Conclusões

Da análise efetuada à região controlo do DNAm de 72 indivíduos judeus curdos foi possível considerar esta população como uma população do Médio Oriente com algumas influências europeias. Esta população de judeus curdos apresenta uma diversidade genética considerável, característica da zona do Médio Oriente. Cerca de 99% dos haplogrupos identificados são característicos do Médio Oriente e da Europa, sendo os haplogrupos maioritários H, J1 e N1. Os judeus curdos possuíam ainda linhagens fundadoras determinadas para os judeus Ashkenazi. Supõe-se que das linhagens identificadas neste estudo, J1b1b, J1b2, J1c, T2c1, U1a1a e X2 possam ter feito parte das linhagens fundadoras dos judeus curdos.

Esta população é, geneticamente, mais próxima das populações judaicas da Bulgária e da Turquia, seguindo-se das populações da Bulgária e do Azerbaijão. Os judeus curdos acabam por ser geneticamente mais distintos das populações judaicas de Belmonte, Bombay e Geórgia e das populações da Mongólia e da Tunísia.

Com os dados obtidos da classificação dos haplogrupos da população em análise, podemos inferir que a mesma manteve-se isolada, pese embora todo o historial de invasões e anexações sofridas com o passar do tempo por intermédio da pressão exercida pelos grandes impérios reinantes na região: persas, turcos e mongóis.

Alguns historiadores referem populações judaicas exiladas no Curdistão há cerca de 2800 anos, algumas delas convertendo, passado pouco tempo, os curdos locais. Até o grande êxodo de 1950-1951, os judeus curdos viviam, maioritariamente, na região do Iraque. Assim, aqui é demonstrado o fluxo da ideia do Judaísmo em adição à recruta de mulheres por parte de outros homens judeus nesta região.

Posto isto, o estabelecimento como comunidade judaica poderá ter ocorrido há menos de 2800 anos. Esta era constituída por curdos locais e judeus exilados pelo rei assírio Shalmaneser III (reinado de 858-824 a.C.). Apenas pela análise feita neste estudo não é possível identificar com exatidão quais as linhagens que tiveram na fundação desta população. Há possibilidade ainda que algumas das linhagens fundadoras possam ter sido perdidas ao longo do tempo, dando origem a outras linhagens diferentes com polimorfismos desenvolvidos apenas no ceio dos judeus curdos. Os judeus curdos mantiveram-se, relativamente, isolados durante a sua história devido a fatores económicos e políticos que determinavam, basicamente, a sua vida, até 1950-1951. Uma propaganda contra atividades sionistas (movimento nacionalista e política de judeus e cultura judaica) que teve início em 1925 no Iraque afetou, adversamente, a posição dos

judeus no Curdistão. Esta propaganda atingiu seu máximo em 1941 com a revolta de Rashid Ali, levando à invasão da Grã-Bretanha. Em 1950-1951 dois bombardeamentos contra os judeus no Iraque levaram ao êxodo dos judeus. Após este incidente, alguns judeus curdos mantiveram-se nas zonas do Irão, Turquia e Síria.

Em sùmula, os judeus curdos são uma população com características típicas do Médio Oriente, com diversidade genética elevada e que permaneceu, relativamente, isolada durante a sua história.

Referências

1. Gray MW, Burger G, Lang BF. The origin and early evolution of mitochondria. *Genome Biology*. 2001;2(6):1018.1-5.
2. van der Giezen M, Tovar J. Degenerate mitochondria. *EMBO Reports*. 2005;6(6):525-30.
3. Chinnery PF, Hudson G. Mitochondrial genetics. *Br Med Bull*. 2013;106:135-59.
4. Giorgi C, Romagnoli A, Pinton P, Rizzuto R. Ca²⁺ signaling, mitochondria and cell death. *Curr Mol Med*. 2008;8:119-30.
5. Lopez J, Tait SW. Mitochondrial apoptosis: killing cancer using the enemy within. *Br J Cancer*. 2015;112(6):957-62.
6. Pizzo P, Drago I, Filadi R, Pozzan T. Mitochondrial Ca²⁺ homeostasis: mechanism, role, and tissue specificities. *Pfluegers Arch/Eur J Physiol*. 2012;464(1):3-17.
7. Shokolenko IN, Wilson GL, Alexeyev MF. Aging: a mitochondrial DNA perspective, critical analysis and an update. *World J Exp Med*. 2004;4(4):46-57.
8. Kukat C, Wurm CA, Spahr H, Falkenberg M, Larsson N-G. Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc Natl Acad Sci USA*. 2011;108(33):13534-9.
9. Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290:457-65.
10. Holt IJ, Reyes A. Human mitochondrial DNA replication. *Cold Spring Harb Perspect Biol*. 2012;4:a012971.
11. Horai S, Hayasaka K. Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Amer J Hum Genet*. 1990;46:828-42.
12. Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, et al. A high observed substitution rate in the human mitochondrial DNA control region. *Nat Genet*. 1997;15:363-8.
13. Lutz S, Witting H, Weisser H-J, Heizmann J, Junge A, Dimo-Simonin N, et al. Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? *Forensic Sci Int*. 2000;113:97-101.
14. Phylotree mt. Annotated mtDNA reference sequences: revised Cambridge Reference Sequence (rCRS) [02-03-2016].
15. Nicholls TJ, Minczuk M. In D-loop: 40 years of mitochondrial 7S DNA. *Exp Gerontol*. 2014;56:175-81.
16. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 1999;23:147.
17. Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. *Nat Rev Genet*. 2005;6(5):389-402.
18. Lee H-C, Wei Y-H. Mitochondria and aging. *Adv Exp Med Biol*. 2012;942:311-27.
19. Satoh M, Kuroiwa T. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp Cell Res*. 1991;196:137-40.

20. Malka F, Lombès A, Rojo M. Organization, dynamics and transmission of mitochondrial DNA: focus on vertebrate nucleoids. *Biochim Biophys Acta*. 2006;1763:463-72.
21. White DJ, Wolff JN, Pierson M, Gemmell NJ. Revealing the hidden complexities of mtDNA inheritance. *Mol Ecol*. 2008;17(23):4925-42.
22. Garrido N, Griparic L, Jokitalo E, Wartiovaara J, van der Blik AM, Spelbrink JN. Composition and dynamics of human mitochondrial nucleoids. *Mol Biol Cell*. 2003;14:1583-96.
23. Ashley N, Harris D, Poulton J. Detection of mitochondrial DNA depletion in living human cells using PicoGreen staining. *Exp Cell Res*. 2005;2005:432-46.
24. Bogenhagen DF. Mitochondrial DNA nucleoid structure. *Biochim Biophys Acta*. 2012;1819:914-20.
25. Howell N, Halvorson S, Kubacka I, McCullough DA, Bindoff LA, Turnbull DM. Mitochondrial gene segregation in mammals: is the bottleneck always narrow? *Hum Genet*. 1992;90:117-20.
26. Mishra P, Chan DC. Mitochondrial dynamics and inheritance during cell division, development and disease. *Nat Rev Mol Cell Biol*. 2014;15(10):634-46.
27. Sato M, Sato K. Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *Biochim Biophys Acta*. 2013;1833(8):1979-84.
28. Giles RE, Blanc H, Cann HM, Wallace DC. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA*. 1980;77(11):6715-9.
29. Morris AA, Lightowers RN. Can paternal mtDNA be inherited? *Lancet*. 2000;355(9212):1290-1.
30. Wolff JN, Gemmell NJ. Mitochondria, maternal inheritance, and asymmetric fitness: why males die younger. *Bioessays*. 2012;35:93-9.
31. Wolff JN, Nafisinia M, Sutovsky P, Ballard JWO. Paternal transmission of mitochondrial DNA as an integral part of mitochondrial inheritance in metapopulations of *Drosophila simulans*. *Heredity*. 2013;110:57-62.
32. Kumar DP, Sangeetha N. Mitochondrial DNA mutations and male infertility. *Indian Journal of Human Genetics*. 2009;15(3):93-7.
33. Sampson MJ, Decker WK, Beaudet AL, Ruitenbeek W, Armstrong D, Craigen WJ. Immobile sperm and infertility in mice lacking mitochondrial voltage-dependent anion channel type 3. *J Biol Chem*. 2001;276(42):39206-12.
34. Kaneda H, Hayashi J-I, Takahama S, Taya C, Lindahl KF, Yonekawa H. Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. *Proc Natl Acad Sci USA*. 1995;92:4542-6.
35. Schwartz M, Vissing J. New patterns of inheritance in mitochondrial disease. *Biochem Biophys Res Commun*. 2003;310:247-51.
36. Birky CWJ. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci USA*. 1995;92:11331-8.
37. Schwartz M, Vissing J. Paternal inheritance of mitochondrial DNA. *New Engl J Med*. 2002;347(8):576-80.
38. Bromham L, Eyre-Walker A, Smith NH, Smith JM. Mitochondrial Steve: paternal inheritance of mitochondrial in humans. *Trends Ecol Evol*. 2003;18(1):2-4.
39. Ashley MV, Laipis PJ, Hauswirth WW. Rapid segregation of heteroplasmic bovine mitochondria. *Nucleic Acids Res*. 1989;17(18):7325-31.
40. Bendall KE, Macaulay VA, Baker JR, Sykes BC. Heteroplasmic point mutations in the human mtDNA control region. *Amer J Hum Genet*. 1996;59:1276-87.

41. Boursot P, Yonekawa H, Bonhomme F. Heteroplasmy in mice with deletion of large coding region of mitochondrial DNA. *Mol Biol Evol.* 1987;4(1):46-55.
42. Lightowlers RN, Chinnery PF, Turnbull DM, Howell N. Mammalian mitochondrial genetics: heredity, heteroplasmy and disease. *Trends Genet.* 1997;13(11):450-5.
43. Naue J, Hörer S, Sängler T, Strobl C, Hatzler-Grubwieser P, Parson W, et al. Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. *Mitochondrion.* 2014;20:82-94.
44. Just RS, Irwin JA, Parson W. Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. *Forensic Sci Int Genet.* 2015.
45. Stoneking M. Mitochondrial DNA and human evolution. *J Bioenerg Biomembr.* 1994;26(3):251-9.
46. Stoneking M, Soodyall H. Human evolution and the mitochondrial genome. *Curr Opin Genet Dev.* 1996;6(6):731-6.
47. Wiuf C. Recombination in human mitochondrial DNA? *Genetics.* 2001;159(2):749-56.
48. Piganeau G, Eyre-Walker A. A reanalysis of the indirect evidence for recombination in human mitochondrial DNA. *Heredity.* 2004;92:282-8.
49. Rokas A, Laddoukakis E, Zouros E. Animal mitochondrial DNA recombination revisited. *Trends Ecol Evol.* 2003;18(8):411-7.
50. White DJ, Bryant D, Gemmell NJ. How good are indirect tests at detecting recombination in human mtDNA? *G3.* 2013;3(7):1095-104.
51. Chen XJ. Mechanism of homologous recombination and implications for aging-related deletions in mitochondrial DNA. *Microbiol Mol Biol Rev.* 2013;77(3):476-96.
52. Thyagarajan B, Padua RA, Campbell C. Mammalian mitochondria possess homologous DNA recombination activity. *J Biol Chem.* 1996;271(44):27536-43.
53. Hey J. Human mitochondrial DNA recombination: can it be true? *Trends Ecol Evol.* 2000;15(5):181-2.
54. Brown WM, Prager EM, Wang A, Wilson AC. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol.* 1982;18:225-39.
55. Pesole G, Gissi C, Chirico AD, Saccone C. Nucleotide substitution rate of mammalian mitochondrial genomes. *J Mol Evol.* 1999;48:427-34.
56. Brown WM, George MJ, Wilson AC. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA.* 1979;76(4):1967-71.
57. Shadel GS, Clayton DA. Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem.* 1997;66:409-35.
58. Bandy B, Davison AJ. Mitochondrial mutations may increase oxidative stress: implications for carcinogenesis and aging? *Free Radical Biol Med.* 1990;8:523-39.
59. Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, et al. The role of selection in the evolution of human mitochondrial genomes. *Genetics.* 2006;172(1):373-87.
60. Howell N, Kubacka I, Mackey DA. How rapidly does the human mitochondrial genome evolve? *Amer J Hum Genet.* 1996;59:501-9.
61. Wallace DC, Ye J, Neckelmann SN, Singh G, Webster KA, Greenberg B. Sequence analysis of cDNAs for the human and bovine ATP synthase B subunit: mitochondrial DNA genes sustain seventeen times more mutations. *Curr Genet.* 1987;12:81-90.
62. Huang S, Wang C, Li H. Natural selection on human mitochondrial DNA. *Biotechnology.* 2014;3(1):1-7.
63. Chinnery PF, Howell N, Andrews RM, Turnbull DM. Mitochondrial DNA analysis: polymorphisms and pathogenicity. *J Med Genet.* 1999;36(7):505-10.

64. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, et al. Identification of the remains of the Romanov family by DNA analysis. *Nat Genet.* 1994;6(2):130-5.
65. Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat Genet.* 1996;12(4):417-20.
66. Wallace DC. Mitochondrial DNA sequence variation in human evolution and disease. *Proc Natl Acad Sci USA.* 1994;91:8739-46.
67. Galtier N, Nabholz B, Glémin S, Hurst GDD. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol.* 2009;18:4541-50.
68. Kundu S, Ghosh SK. Trend of different molecular markers in the last decades for studying human migrations. *Gene.* 2015;556:81-90.
69. De Kniff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, et al. Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med.* 1997;110:134-40.
70. Collins FS, Guyer MS, Chakravarti A. Variations on a theme: cataloging human DNA sequence variation. *Science.* 1997;278:1580-1.
71. Budowle B, Allard MW, Wilson MR, Chakraborty R. Forensics and mitochondrial DNA: applications, debates, and foundations. *Annu Rev Genom Hum Genet.* 2003;4:119-41.
72. Behar DM, Villemans R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, et al. The dawn of human matrilineal diversity. *Amer J Hum Genet.* 2008;82:1130-40.
73. Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Amer J Hum Genet.* 1995;57:133-49.
74. Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, et al. Classification of European mtDNAs from an analysis of three European populations. *Genetics.* 1996;144:1835-50.
75. Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, et al. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Amer J Hum Genet.* 2002;70:1152-71.
76. Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, et al. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Amer J Hum Genet.* 1993;53:591-608.
77. Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, et al. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Amer J Hum Genet.* 1993;53:563-90.
78. Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, et al. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Amer J Hum Genet.* 1990;46:613-23.
79. Kivisild T. Maternal ancestry and population history from whole mitochondrial genomes. *Investig Genet.* 2015;6:10.
80. Reyes-Centeno H, Hubbe M, Hanihara T, Stringer C, Harvati K. Testing modern human out-of-Africa dispersal models and implications for modern human origins. *J Hum Evol.* 2015;87:95-106.
81. Blum MGB, Jakobsson M. Deep divergences of human gene trees and models of human origins. *Mol Biol Evol.* 2011;28(2):889-98.

82. Stringer C. Why we are not all multiregionalists now. *Trends Ecol Evol.* 2014;29(5):248-51.
83. Harris K, Nielsen R. The genetic cost of Neanderthal introgression. *Genetics.* 2016.
84. Rito T, Richards M, Fernandes V, Alshamali F, Cerny V, Pereira L, et al. The first modern human dispersals across Africa. *PLoS One.* 2013;8(11):16.
85. Fregel R, Cabrera VM, Larruga JM, Abu-Amero KK, González AM. Carriers of mitochondrial DNA macrohaplogroup N lineages reached Australia around 50,000 years ago following a northern Asian route. *PLoS One.* 2015;10(6):e0129839.
86. Campbell MC, Tishkoff SA. The evolution of human genetic and phenotypic variation in Africa. *Curr Biol.* 2010;20:R166-R73.
87. Soares P, Alshamali F, Pereira JB, Fernandes V, Silva NM, Afonso C, et al. The expansion of mtDNA haplogroup L3 within and out of Africa. *Mol Biol Evol.* 2012;29(3):915-27.
88. Soares P, Achilli A, Semino O, Davies W, Macaulay VA, Bandelt H-J, et al. The archaeogenetics of Europe. *Curr Biol.* 2010;20:R174-R83.
89. Torroni A, Achilli A, Macaulay VA, Richards M, Bandelt H-J. Harvesting the fruit of the human mtDNA tree. *Trends Genet.* 2006;22(6):339-45.
90. Richards M, Macaulay VA, Hickey E, Vega E, Sykes BC, Guida V, et al. Tracing european founder lineages in the Near Eastern mtDNA pool. *Amer J Hum Genet.* 2000;67:1251-76.
91. Mishmar D, Ruiz-Pesini E, Golik P, Macaulay VA, Clark AG, Hosseini S, et al. Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA.* 2003;100(1):171-6.
92. Palanichamy MG, Sun C, Agrawal S, Bandelt H-J, Kong QP, Khan F, et al. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Amer J Hum Genet.* 2004;75(6):966-78.
93. Chaubey G, Metspalu M, Kivisild T, Villems R. Peopling of South Asia: investigating the caste-tribe continuum in India. *Bioessays.* 2007;29(1):91-100.
94. Majumder PP. The human genetic history of South Asia. *Curr Biol.* 2010;20(4):R184-R7.
95. Li YC, Wang HW, Tian JY, Liu LN, Yang LQ, Zhu CL, et al. Ancient inland human dispersals from Myanmar into interior East Asia since the Late Pleistocene. *Scientific Reports.* 2015;5:9473.
96. Kivisild T, Tolk H-V, Parik J, Wang Y, Papiha S, Bandelt H-J, et al. The emerging limbs and twigs of the East Asian mtDNA tree. *Mol Biol Evol.* 2002;19(10):1737-51.
97. Stoneking M, Delfin F. The human genetic history of East Asia: weaving a complex tapestry. *Curr Biol.* 2010;20(4):R188-R93.
98. Sukernik RI, Volodko NV, Mazunin IO, Eltsov NP, Drymov SV, Starikovskaya EB. Mitochondrial genome diversity in the Tubalar, Even, and Ulchi: contribution to prehistory of native Siberians and their affinities to Native Americans. *Amer J Phys Anthropol.* 2012;148(1):123-38.
99. Derenko M, Malyarchuk B, Denisova G, Perkova M, Litvinov A, Grzybowski T, et al. Western Eurasian ancestry in modern Siberians based on mitogenomic data. *BMC Evol Biol.* 2014;14:217.
100. Kayser M. The human genetic history of Oceania: near and remote views of dispersal. *Curr Biol.* 2010;20:R194-R201.

101. Hudjashov G, Kivisild T, Underhill PA, Endicott P, Sanchez JJ, Lin AA, et al. Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Natl Acad Sci USA*. 2007;104(21):8726-30.
102. Duggan AT, Evans B, Friedlaender FR, Friedlaender JS, Koki G, Merriwether DA, et al. Maternal history of Oceania from complete mtDNA genomes: contrasting ancient diversity with recent homogenization due to the Austronesian expansion. *Amer J Hum Genet*. 2014;94:721-33.
103. Gomes SM, Bodner M, Souto L, Zimmermann B, Huber G, Strobl C, et al. Human settlement history between Sunda and Sahul: a focus on East Timor (Timor-Leste) and the Pleistocenic mtDNA diversity. *BMC Genomics*. 2015;16:70.
104. Fagundes NJR, Kanitz R, Bonatto SL. A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. *PLoS One*. 2008;3(9):e3157.
105. Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, Kashani BH, et al. The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. *Genome Res*. 2010;20(9):1174-9.
106. Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, et al. Beringian standstill and spread of Native American founders. *PLoS One*. 2007;2(9):e829.
107. Gilbert MT, Kivisild T, Grønnow B, Andersen PK, Metspalu E, Reidla M, et al. Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. *Science*. 2008;320(5884):1787-9.
108. Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk H-V, et al. Origin and diffusion of mtDNA haplogroup X. *Amer J Hum Genet*. 2003;73:1178-90.
109. Shoup JA. *Ethnic groups of Africa and Middle East: an encyclopedia: ABC-CLIO*; 2011.
110. Jwaideh W. *Kurdish national movement: its origins and development: Syracuse University Press*; 2006.
111. Nasidze I, Quinque D, Ozturk M, Bendukidze N, Stoneking M. MtDNA and Y-chromosome variation in Kurdish groups. *Ann Hum Genet*. 2005;69:401-12.
112. McDowall D. *A modern history of the Kurds*. 3^a ed: I. B. Tauris; 2003. 504 p.
113. Kurds The Columbia Encyclopedia, 6^a ed.2015 [12-10-2015]. Available from: <http://www.encyclopedia.com/topic/Kurds.aspx#4>.
114. Hennerbichler F. The origins of Kurds. *Advances in Anthropology*. 2012;2(2):64-79.
115. Nezan K. A brief survey of the history of the Kurds Kurdish Institute of Paris [12-10-2015]. Available from: http://www.institutkurde.org/en/institute/who_are_the_kurds.php.
116. Encyclopaedia of Kurdistan. Religion in Kurdistan: Judaism 2008 [19-10-2015]. Available from: <http://kurdistanica.com/?q=node/105>.
117. Rogacs A, Marshall LA, Santiago JG. Purification of nucleic acids using isotachopheresis. *J Chromatogr A*. 2014;1335:105-20.
118. Bio-Rad. Chelex® 100 and Chelex 20 Chelating Ion Exchange Resin: Instruction Manual. California: Bio-Rad Laboratories; 2000.
119. Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*. 1991;10(4):506-13.
120. Baumforth KR, Nelson PN, Digby JE, O'Neil JD, Murray PG. Demystified... the polymerase chain reaction. *Mol Pathol*. 1999;52(1):1-10.
121. Genetic Science Learning Center. PCR University of Utah. 2008 [08-12-2014]. Available from: <http://learn.genetics.utah.edu/content/labs/pcr/>.

122. NCBI. Polymerase Chain Reaction (PCR) 2014 [10-12-2014]. Available from: <http://www.ncbi.nlm.nih.gov/probe/docs/techpcr/>.
123. Gibbs RA. DNA amplification by the polymerase chain reaction. *Anal Chem.* 1990;62(13):1202-14.
124. Kim Y, Flynn TR, Donoff RB, Wong DT, Todd R. The gene: the polymerase chain reaction and its clinical application. *J Oral Maxillofac Surg.* 2002;60(7):808-15.
125. Franks F. Protein biotechnology: isolation, characterization, and stabilization. New Jersey: The Humana Press; 1993. 592 p.
126. Gel electrophoresis of nucleic acids: a practical approach. Second ed. Rickwood D, Hames BD, editors. Oxford University Press: IRL Press; 1990. 311 p.
127. Westermeier R. Electrophoresis in practice. 3^a ed: Wiley-VCH; 2001.
128. Metzker ML. Emerging technologies in DNA sequencing. *Genome Res.* 2005;15(12):1767-76.
129. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol.* 1975;94(3):441-8.
130. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA.* 1977;74(12):5463-7.
131. Morey M, Fernández-Marmiesse A, Castiñeiras D, Fraga JM, Couce ML, Cocho JA. A glimpse into past, present, and future DNA sequencing. *Mol Genet Metab.* 2013;110:3-24.
132. Kircher M, Kelso J. High-throughput DNA sequencing - concepts and limitations. *Bioessays.* 2010;32(6):524-36.
133. França LT, Carrilho E, Kist TB. A review of DNA sequencing techniques. *Q Rev Biophys.* 2002;35(2):169-200.
134. Hamoy IG, Ribeiro-dos-Santos AM, Alvarez L, Barbosa S, Silva A, Santos S, et al. A protocol for mtGenome analysis on large sample numbers. *Bioinform Biol Insights.* 2014;8:127-34.
135. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010;10(3):564-7.
136. Bandelt H-J, Forster P, Sykes BC, Richards M. Mitochondrial portraits of human populations using median networks. *Genetics.* 1995;141(2):743-53.
137. Bandelt H-J, Macaulay VA, Richards M. Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. *Mol Phylogen Evol.* 2000;16(8-28).
138. Hasegawa M, Cao Y, Yang Z. Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous rate ratio is much higher within species than between species. *Mol Biol Evol.* 1998;15(11):1499-505.
139. Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, et al. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum Mutat.* 2011;32(1):25-32.
140. IBM Corp. IBM SPSS Statistics for Windows. 22.0 ed. Armonk, NY: IBM Corp; 2013.
141. Comas D, Calafell F, Bendukidze N, Fañanás L, Bertranpetit J. Georgian and Kurd mtDNA sequence analysis shows a lack of correlation between languages and female genetic lineages. *Amer J Phys Anthropol.* 2000;112:5-16.
142. Rosa A. Phylogenetic structure of Guinea-Bissau ethnic groups for mitochondrial DNA and Y chromosome genetic systems: Tese de Doutorado; Universidade da Madeira; 2007.

143. Serk P. Human mitochondrial DNA haplogroup J in Europe and Near East: Tese de Mestrado; Universidade de Tartu; 2004.
144. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, et al. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun.* 2013;4:1764.
145. Roostalu U, Kutuev I, Loogväli E-L, Metspalu E, Tambets K, Reidla M, et al. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. *Mol Biol Evol.* 2007;24(2):436-48.
146. Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, Reidla M, et al. Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Amer J Hum Genet.* 2012;90(5):915-24.
147. Abu-Amero KK, Larruga JM, Cabrera V, González AM. Mitochondrial DNA structure in the Arabian Peninsula. *BMC Evol Biol.* 2008;8:45.
148. Cerny V, Mulligan CJ, Rídl J, Zaloudková M, Edens CM, Hájek M, et al. Regional differences in the distribution of the sub-Saharan, West Eurasian, and South Asian mtDNA lineages in Yemen. *Amer J Phys Anthropol.* 2008;136:128-37.
149. Fadhaoui-Zid K, Rodríguez-Botigué L, Naoui N, Benammar-Elgaaïed A, Calafell F, Comas D. Mitochondrial DNA structure in North Africa reveals a genetic discontinuity in the Nile Valley. *Amer J Phys Anthropol.* 2011;145:107-17.
150. Fernández E, Pérez-Pérez A, Gamba C, Prats E, Cuesta P, Anfruns J, et al. Ancient DNA Analysis of 8000 B.C. Near Eastern Farmers Supports an Early Neolithic Pioneer Maritime Colonization of Mainland Europe through Cyprus and the Aegean Islands. *PLoS Genet.* 2014;10(6):e1004401.
151. Costa MD, Pereira JB, Pala M, Fernandes V, Olivieri A, Achilli A, et al. A substantial prehistoric European ancestry amongst Ashkenazi maternal lineages. *Nat Commun.* 2013;4:2543.
152. González AM, Larruga JM, Abu-Amero KK, Shi Y, Pestano J, Cabrera V. Mitochondrial lineage M1 traces an early human backflow to Africa. *BMC Genomics.* 2007;8:223.
153. Brandstätter A, Zimmermann B, Wagner J, Göbel T, Röck A, Salas A, et al. Timing and deciphering mitochondrial DNA macro-haplogroup R0 variability in Central Europe and Middle East. *BMC Evol Biol.* 2008;8:191.
154. Cerny V, Pereira L, Kujanová M, Vasíková A, Hájek M, Morris M, et al. Out of Arabia-the settlement of island Soqatra as revealed by mitochondrial and Y chromosome genetic diversity. *Amer J Phys Anthropol.* 2009;138(4):439-47.
155. Rai N, Taher N, Singh M, Chaubey G, Jha AN, Singh L, et al. Relic excavated in western India is probably of Georgian Queen Ketevan. *Mitochondrion.* 2014;14(1):1-6.
156. Derenko M, Malyarchuk B, Bahmanimehr A, Denisova G, Perkova M, Farjadian S, et al. Complete mitochondrial DNA diversity in Iranians. *PLoS One.* 2013;2(11):e80673.
157. González AM, Brehm A, Pérez JA, Maca-Meyer N, Flores C, Cabrera V. Mitochondrial affinities at the Atlantic fringe of Europe. *Amer J Phys Anthropol.* 2003;120(4):391-404.
158. Excoffier L, Lischer HE. User Manual: Arlequin ver 3.5: Swiss Institute of Bioinformatics; 2015.
159. Nei M. *Molecular Evolutionary Genetics.* Nova Iorque: Columbia University Press; 1987.

160. Behar DM, Metspalu E, Kivisild T, Rosset S, Tzur S, Hadid Y, et al. Counting the founders: the matrilineal genetic ancestry of the Jewish diaspora. *PLoS One*. 2008;3(4):e2062.
161. Ostrer H, Skorecki K. The population genetics of the Jewish people. *Hum Genet*. 2013;132(2):119-27.
162. Irwin JA, Ikramov A, Saunier J, Bodner M, Sylvain A, Röck A, et al. The mtDNA composition of Uzbekistan: a microcosm of Central Asian patterns. *Int J Legal Med*. 2010;124:195-204.
163. Bekada A, Arauna L, Deba T, Calafell F, Benhamamouch S, Comas D. Genetic heterogeneity in Algerian human populations. *PLoS One*. 2015;10(9):e0138453.
164. Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, et al. Where west meets East: the complex mtDNA landscape of the Southwest and Central Asian corridor. *Amer J Hum Genet*. 2004;74:827-45.
165. Karachanak S, Carossa V, Nesheva D, Olivieri A, Pala M, Kashani BH, et al. Bulgarians vs the other European populations: a mitochondrial DNA perspective. *Int J Legal Med*. 2012;126(497-503).
166. Elmadawy MA, Nagai A, Gomaa GM, Hegazy HMR, Shaaban FE, Bunai Y. Investigation of mtDNA control region sequences in an Egyptian population sample. *Legal Med*. 2013;15:338-41.
167. Al-Zahery N, Saunier J, Ellingson K, Parson W, Parsons TJ, Irwin JA. Characterization of mitochondrial DNA control region lineages in Iraq. *Int J Legal Med*. 2013;127(2):373-5.
168. González AM, Karadsheh N, Maca-Meyer N, Flores C, Cabrera V, Larruga JM. Mitochondrial DNA variation in Jordanians and their genetic relationship to other Middle East populations. *Ann Hum Biol*. 2008;35(2):212-31.
169. Cheng B, Tang W, He L, Dong Y, Lu J, Lei Y, et al. Genetic imprint of the Mongol: signal from phylogeographic analysis of mitochondrial DNA. *J Hum Genet*. 2008;53:905-13.
170. Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, et al. The genome-wide structure of the Jewish people. *Nature*. 2010;466(7303):238-42.
171. Ennafaa H, Fregel R, Khodjet-et-khil H, González AM, Mahmoudi HAE, Cabrera V, et al. Mitochondrial DNA and Y-chromosome microstructure in Tunisia. *J Hum Genet*. 2011;56:734-41.

Material Suplementar

Tabela S.1: Populações incluídas no estudo, com respectivas abreviaturas, número da amostra populacional, referência e diversidade genética. (Legenda: N – Indivíduos na população utilizados no estudo; sd – Desvio padrão)

População	Abreviatura	N	Referência	Diversidade genética	sd
Afganistão	Afg	98	[162]	0.9102	0.0194
Argélia	Alg	336	[163]	0.8798	0.0102
Arábia Saudita	ArS	553	[147]	0.9334	0.0052
Azerbaijão	Azb	40	[164]	0.8910	0.0377
Bulgária	Bg	855	[165]	0.9007	0.0066
Cazaquistão	Czq	256	[162]	0.9525	0.0063
Egito	Egt	99	[166]	0.9586	0.0072
Geórgia	Geo	58	[164]	0.9153	0.0199
Iémen	Ye	185	[148]	0.9172	0.0089
Irão	Ir	352	[156]	0.9687	0.0030
Iraque	Irq	182	[167]	0.9370	0.0099
Jordânia	Jor	145	[168]	0.8917	0.0177
Líbia	Lb	269	[149]	0.9454	0.0046
Mongólia	Mng	201	[169]	0.9271	0.0088
Quirguistão	Quig	249	[162]	0.9456	0.0073
Síria	Sir	215	[170]	0.9245	0.0083
Tunísia	Tn	80	[171]	0.9073	0.0171
Turquemenistão	Tqm	249	[162]	0.9426	0.0078
Turquia	Tur	383	[170]	0.8994	0.0094
Uzbequistão	Uzb	328	[162]	0.9645	0.0039
Judeu Azerbaijão	AzbJ	58	[160]	0.6334	0.0662
Judeu Belmonte	BelJ	30	[160]	0.1287	0.0792
Judeu Bombay (Índia)	BJ	34	[160]	0.4902	0.0975
Judeu Bulgária	BgJ	71	[160]	0.8861	0.0252
Judeu Cochim (Índia)	CJ	45	[160]	0.8404	0.0255
Judeu Etiópia (Beta Israel)	EthJ	29	[160]	0.9187	0.0253
Judeu Geórgia	GeoJ	74	[160]	0.6275	0.0607
Judeu Iémen	YeJ	119	[160]	0.9181	0.0094
Judeu Irão	IrJ	82	[160]	0.9021	0.0133
Judeu Iraque	IrqJ	135	[160]	0.8786	0.0133
Judeu Líbia	LbJ	83	[160]	0.7376	0.0346
Judeu Marrocos	MarJ	149	[160]	0.8299	0.0251
Judeu Tunísia	TnJ	37	[160]	0.8468	0.0463
Judeu Turquia	TurJ	123	[160]	0.8535	0.0265

Continuação da Tabela S.2

KJ52	J1b1b	95.75	16069T	16126C	16145A	16261T	16519C	73G	263G	271T	295T	309.1C	315.1C	462T	489C	523-524d
KJ53	X2	90.5	16189C	16207C	16209C	16223T	16278T	16344T	16519C	73G	153G	195C	225A	263G	315.1C	
KJ55	R0a1a	100	16126C	16355T	16362C	58C	64T	146C	263G	315.1C						
KJ56	K1a9	100	16093C	16224C	16311C	16519C	16524G	73G	195C	263G	315.1C	497T				
KJ57	H	100	16519C	150T	263G	309.1C	315.1C									
KJ58	N1b1	90.49	16126C	16145A	16176G	16390A	16519C	73G	152C	195C	263G	315.1C				
KJ59	X2	90.5	16189C	16207C	16209C	16223T	16278T	16344T	16519C	73G	153G	195C	225A	263G	309.1C	315.1C
KJ60	J1c	100	16069T	16126C	73G	185A	228A	263G	295T	315.1C	462T	489C				
KJ61	H	100	16189C	263G	315.1C											
KJ62	J1c	87.41	16069T	16093C	16126C	16261T	42.1G	73G	152C	185A	228A	234G	263G	295T	315.1C	385G 462T 489C
KJ63	J1b2	94.64	16069T	16126C	16145A	16261T	73G	146C	263G	295T	315.1C	462T	489C			
KJ64	J1c	93.69	16069T	16093C	16126C	16261T	73G	152C	185A	228A	263G	295T	315.1C	385G	462T	489C
KJ65	H	100	16519C	16524C	146C	263G	315.1C									
KJ66	H	100	16278T	16519C	263G	309.1C	315.1C									
KJ67	U7a	91.39	16309G	16318T	16519C	73G	151T	152C	263G	315.1C	523-524d					
KJ68	T1	90.77	16126C	16163G	16294T	16519C	73G	146C	244.1TAA	263G	315.1C					
KJ69	X2	82.03	16183C	16189C	16278T	16519C	73G	195C	225A	226C	227G	263G	315.1C			
KJ70	K1a9	100	16093C	16224C	16311C	16519C	16524G	73G	195C	263G	315.1C	497T				
KJ71	H	100	16311C	16519C	263G	315.1C										
KJ72	H	100	16189C	263G	315.1C											
KJ501	H	100	16261T	16519C	200G	263G	309.1CC	315.1C								
KJ502	H	100	16183C	16189C	241.1TAA	263G	309.1CCC	315.1C								
KJ503	J1b2	100	16069T	16126C	16145A	16222T	16261T	73G	152C	263G	295T	309.1C	315.1C	462T	489C	
KJ504	H	100	16189C	263G	315.1C											
KJ505	J2a2b2	97.53	16069T	16126C	16241G	73G	150T	195C	204C	263G	295T	315.1C	456T	489C	524.1AC	
KJ512	X2	90.5	16189C	16207C	16209C	16223T	16278T	16344T	16519C	73G	153G	195C	225A	263G	309.1C	315.1C
KJ513	H	100	16278T	16519C	263G	309.1C	315.1C									
KJ514	H	100	16183C	16189C	263G	309.1CCC	315.1C									
KJ516	H	100	16189C	263G	315.1C											
KJ517	N1b1	88.65	16126C	16145A	16176G	16390A	16519C	73G	152C	195C	263G	315.1C				
KJ518	N1b1	88.65	16126C	16145A	16176G	16390A	16519C	73G	152C	195C	263G	315.1C				
KJ521	H	100	16189C	263G	309.1C	315.1C										
KJ522	N1b1	88.65	16126C	16145A	16176G	16390A	16519C	73G	152C	195C	263G	315.1C				
KJ523	N1b1	88.65	16126C	16145A	16176G	16390A	16519C	73G	152C	195C	263G	315.1C				
KJ524	H	100	152C	204C	263G	309.1C	315.1C	499C								
KJ525	N1b1	88.65	16126C	16145A	16176G	16390A	16519C	73G	152C	195C	263G	315.1C				

Tabela S.3: Polimorfismos identificados na região codificante para a confirmação dos respectivos haplogrupos, em comparação com a rCRS.

Amostra	Haplogrupo	Região codificante			
KJ01	N1a3a	12705T			
KJ03	H				
KJ04	N1b1	8836G	12705T		
KJ05	J1b2	1733T	13708A		
KJ07	U7a	5360T			
KJ08	H	1438G	4769G		
KJ09	J1b2	1733T			
KJ11	J1b6	5501G	8269A	13708A	
KJ15	H				
KJ19	T1	11377A	11719A	12633A	
KJ20	N2a	709A	739T	750G	12705T
KJ21	T2b3	10463C	10750G	13368A	
KJ22	T2c1	10822T	13368A		
KJ24	J1c	14798C			
KJ26	J1c	14798C			
KJ27	J1c	14798C			
KJ28	H				
KJ30	H				
KJ31	N1a3a	12705T			
KJ32	H	15326G			
KJ34	H				
KJ35	U1a1a	3158.1T	3591A	13104G	
KJ36	U1a1a	3591A	13104G	13368G	
KJ37	K1a5b	1189C	9716T		
KJ38	H				
KJ39	J1c1				
KJ40	U1a1a	2218T	13104G		
KJ41	N1b1	8836G	12705T		
KJ42	T2c1	10822T	13368A		
KJ43	J1b2	1733T	13708A		
KJ44	M1a1	12403T			
KJ45	H				
KJ47	X2	1719A			
KJ48	H				
KJ49	K2a2a1	9716C			
KJ51	K1a5b	1189C	9716T		
KJ52	J1b1b	5460A			
KJ53	X2	1719A			
KJ55	R0a1a	750G	827G	12705C	
KJ56	K1a9	1189C	9716T		
KJ57	H				
KJ58	N1b1	8836G	12705T		

Continuação da Tabela S.3.

KJ59	X2	1719A			
KJ60	J1c	4216C	14798C		
KJ61	H				
KJ62	J1c	14798C			
KJ63	J1b2	1733T	13708A		
KJ64	J1c	14798C			
KJ65	H	15326G			
KJ66	H				
KJ67	U7a	5360T			
KJ68	T1	11377A	11719A	12633A	
KJ69	X2	1719A			
KJ70	K1a9	1189C	9716T		
KJ71	H				
KJ72	H				
KJ501	H				
KJ502	H				
KJ503	J1b2	1733T			
KJ504	H				
KJ505	J2a2b2	15257A			
KJ512	X2	1719A			
KJ513	H				
KJ514	H				
KJ516	H				
KJ517	N1b1	12705T			
KJ518	N1b1	12705T			
KJ521	H				
KJ522	N1b1	12705T			
KJ523	N1b1	12705T			
KJ524	H	1438G	4769G	15152A	15326G
KJ525	N1b1	12705T			

Tabela S.4: Número de correspondências exatas distribuídos por país, em relação aos haplótipos identificados nos judeus curdos.

	Médio Oriente				Sul da Ásia			Europa							África	América			
	Kuwait	Bahrain	Iraque	Dubai	Paquistão	Uzbequistão	Afganistão	Turquia	Macedónia	Hungria	Grécia	Suécia	Portugal	Espanha	Alemanha	Áustria	Marrocos	Estados Unidos	Argentina
KJ01																			
KJ03										1				2			1	2	
KJ04																			
KJ05	3		1	1								6	1					15	
KJ07	2	1	1		1	1					1								
KJ08																			
KJ09	2											1							
KJ11																			
KJ15										1				2		1	2		
KJ19																			
KJ20																			
KJ21																			
KJ22																			
KJ24																			
KJ26																			
KJ27																			
KJ28										1				2		1	2		
KJ30																			
KJ31																			
KJ32																		1	
KJ34																			
KJ35																			
KJ36																			
KJ37																			
KJ38																			
KJ39											1								
KJ40																			
KJ41																			
KJ42																			
KJ43	3		1	1								6	1					15	
KJ44																			
KJ45																			
KJ47																			
KJ48												1							
KJ49									14									3	
KJ51																			
KJ52																			
KJ53																			
KJ55	6													1					1

Continuação da Tabela S.4.

KJ56									8				1		7		
KJ57											3	5			11		
KJ58																	
KJ59																	
KJ60				1			2	2	1	5	4		6	2	1	18	
KJ61											1						
KJ62																	
KJ63																	
KJ64																	
KJ65																	
KJ66									1				2		1	2	
KJ67	2	1	1		1	1				1							
KJ68																	
KJ69																	
KJ70									8					1		7	
KJ71		1			1			1	1	1	8		1	8	1	3	16
KJ72											1						
KJ501																3	
KJ502																	
KJ503	2										1						
KJ504											1						
KJ505																	
KJ512																	
KJ513									1				2		1	2	
KJ514																	
KJ516											1						
KJ517																	
KJ518																	
KJ521																	
KJ522																	
KJ523																	
KJ524																	
KJ525																	

Tabela S.5: Número de haplótipos parecidos distribuídos por país com uma ou duas diferenças, em relação aos haplótipos obtidos nos judeus curdos. Para os haplótipos com mais de 150 exemplos, foram agrupados por zonas geográficas (valores a negrito).

	Médio Oriente				Sul ásia		Europa										África			América				
	Kuwait	Bahrain	Iraque	Dubai	Paquistão	Uzbequistão	Chipre	Macedónia	Hungria	Roménia	Grécia	Suécia	Portugal	Espanha	Alemanha	Áustria	Fed. da Rússia	Marrocos	Egito	Somália	Estados Unidos	México	Argentina	
KJ01	14																							
KJ03	124						587										79			732				
KJ04								1										1	1		1			
KJ05	1	2	1	1	1			2	1									1	4		2			
KJ07	3	2	1	3	30	13	1		1	1	1	1							2		3			
KJ08	93						298										52			345				
KJ09	16	3	3	6	3	2		1	1		2		6	1					1		16	1		
KJ11	3		1	1								6	1								15			
KJ15	124						587										79			732				
KJ19					3	3		1			1			1	3				1		2			
KJ20					1																			
KJ21																					5		1	
KJ22									1	1								1			6			
KJ24																					4			
KJ26																								
KJ27																					4			
KJ28	124						587										79			732				
KJ30	3			2	1	5		3	4	1	6	3	6	7	10	3			10		53		1	
KJ31	14																							
KJ32																								
KJ34	47						247										38			281				
KJ35																		1						
KJ36	1																				1			
KJ37	1							1						1					1		6			
KJ38	47						247										38			281				
KJ39						2		2	8	1	2	9	7	6	8	7					70			
KJ40	1																				1			
KJ41								1										1	1		1			
KJ42									1	1								1			6			
KJ43	1	2	1	1	1			2	1									1	4		2			
KJ44	1																		1		1			
KJ45	3		1	3	1	6		4	5	1	6	4	6	10	15	8		5	11		67		5	
KJ47	1																							
KJ48	108						344										61			422				
KJ49				1		4		1		38	3	3	1	1	13	7		1			46		3	

Continuação da Tabela S.5.

KJ51	1							1							1			1		6		
KJ52	3	1	2	2	2					1							1					
KJ53	1																					
KJ55	3	3		4	1	1			1	2	3						1	1	1			
KJ56						3		1				1	2		3		1	1		19		3
KJ57																						
KJ58								1									1	1		1		
KJ59	1																					
KJ60	2		1				9	1	12	23	7	10	6	8	27	9	3	4	1		118	3
KJ61																						
KJ62																						
KJ63	8		1	1				1		2	5		6	1							17	1
KJ64																					4	
KJ65																						
KJ66																						
KJ67	3	2	1	3	30	13	1		1	1	1	1							2		3	
KJ68					3	3		1			1			1	3				1		2	
KJ69												1										
KJ70						3		1				1	2		3			1	1		19	3
KJ71																						
KJ72																						
KJ501																						
KJ502				1		1							1		2	2		1			7	1
KJ503	16	3	3	6	3	2		1	1		2		6	1							16	
KJ504																						
KJ505																					2	
KJ512	1																					
KJ513																						
KJ514	3		1	3	1	6		4	5	1	6	4	6	10	15	8		5	11		67	5
KJ516																						
KJ517								1										1	1		1	
KJ518								1										1	1		1	
KJ521	3			2	1	5		3	4	1	6	3	6	7	10	3			10		53	1
KJ522								1										1	1		1	
KJ523								1										1	1		1	
KJ524	1	1				1	1			2	1		5		4						9	
KJ525								1										1	1		1	

Tabela S.6: Índices F_{ST} obtidos para as populações que tiveram alguma influência sobre o Curdistão em estudo. Abreviaturas consoante a Tabela S.1.

	KJ	Irq	Ir	Sir	Tur	ArS	Ye	Jor	Quig	Uzb	Tqm	Afg	Czq	Alg	Lb	Tn	Egt	Mng	Geo	Azb	Bg	
KJ	0																					
Irq	0.0824	0																				
Ir	0.0599	0.0199	0																			
Sir	0.0541	0.0481	0.0308	0																		
Tur	0.0439	0.0665	0.0416	0.0045*	0																	
ArS	0.0650	0.0095	0.0260	0.0394	0.0547	0																
Ye	0.0745	0.0178	0.0313	0.0517	0.0699	0.0161	0															
Jor	0.0425	0.0624	0.0444	0.0136	0.0165	0.0524	0.0565	0														
Quig	0.084	0.0515	0.0313	0.0509	0.0577	0.0533	0.0609	0.0614	0													
Uzb	0.0517	0.0352	0.0145	0.0281	0.0312	0.0373	0.0449	0.0361	0.0092	0												
Tqm	0.0369	0.0458	0.0232	0.0311	0.0301	0.0422	0.0488	0.0309	0.0142	0.0044*	0											
Afg	0.0838	0.0558	0.0446	0.0591	0.0668	0.0684	0.0790	0.0686	0.0552	0.0374	0.0424	0										
Czq	0.0745	0.0442	0.0241	0.0444	0.0496	0.0470	0.0553	0.0555	0.0036*	0.0034*	0.0101	0.0472	0									
Alg	0.0608	0.0730	0.0615	0.0317	0.0349	0.0609	0.0520	0.0251	0.0708	0.0473	0.0438	0.0855	0.0644	0								
Lb	0.0867	0.0279	0.0266	0.0382	0.0544	0.0283	0.0247	0.0468	0.0489	0.0325	0.0434	0.0641	0.0408	0.0329	0							
Tn	0.1133	0.0606	0.0507	0.0657	0.0855	0.0634	0.0361	0.0711	0.0694	0.0553	0.0654	0.0864	0.0646	0.0570	0.0282	0						
Egt	0.0630	0.0204	0.0162	0.0308	0.0463	0.0243	0.0185	0.0382	0.0385	0.0222	0.0295	0.0477	0.0329	0.0311	0.0094	0.0299	0					
Mng	0.1169	0.0661	0.0470	0.0715	0.0843	0.0668	0.0757	0.0879	0.0142	0.0263	0.0358	0.0717	0.0187	0.0964	0.0627	0.0821	0.0561	0				
Geo	0.1006	0.0430	0.0279	0.0606	0.0709	0.0434	0.0537	0.0777	0.0579	0.0431	0.0513	0.0766	0.0523	0.0924	0.0562	0.0850	0.0442	0.0771	0			
Azb	0.0237*	0.0693	0.0416	0.0241	0.0163*	0.0566	0.0625	0.0126**	0.0575	0.0310	0.0201	0.0676	0.0506	0.0358	0.0609	0.0865	0.0451	0.0864	0.0676	0		
Bg	0.0251	0.0613	0.0361	0.0293	0.0204	0.0575	0.0655	0.0247	0.0539	0.0246	0.0182	0.0526	0.0429	0.0397	0.0513	0.0757	0.0405	0.0846	0.0672	0.0136*	0	

Níveis de significância:

$P=0.0000$

* $0.000 < P < 0.05$

** $P > 0.05$

Tabela S.7: Índices F_{ST} obtidos para as populações judaicas em estudo. Abreviaturas consoante a Tabela S.1.

	KJ	IrJ	IrqJ	AzbJ	GeoJ	EthJ	BJ	CJ	MarJ	TnJ	LbJ	TurJ	BgJ	BelJ	YeJ
KJ	0														
IrJ	0.0690	0													
IrqJ	0.0582	0.0254*	0												
AzbJ	0.2169	0.1989	0.1946	0											
GeoJ	0.2341	0.2082	0.2085	0.3431	0										
EthJ	0.1202	0.0862	0.0954	0.2374	0.2436	0									
BJ	0.3057	0.2660	0.2752	0.4249	0.4258	0.3016	0								
CJ	0.1606	0.1161	0.1369	0.2670	0.2698	0.1171	0.1884	0							
MarJ	0.0236	0.0933	0.0872	0.2006	0.1952	0.1286	0.2904	0.1646	0						
TnJ	0.0345*	0.0872	0.0929	0.2245	0.2282	0.1062	0.3145	0.1564	0.0090**	0					
LbJ	0.0846	0.1483	0.1511	0.2670	0.2884	0.1726	0.3537	0.2134	0.0772	0.0289*	0				
TurJ	0.0179*	0.0681	0.0684	0.1781	0.2052	0.1155	0.2807	0.1486	0.0070**	0.0034**	0.0615	0			
BgJ	0.0139**	0.0567	0.0500	0.1890	0.2079	0.0919	0.2779	0.1358	0.0085*	0.0108**	0.0754	0.0030**	0		
BelJ	0.4398	0.4031	0.3913	0.5666	0.5565	0.4798	0.6811	0.4783	0.3923	0.4802	0.4821	0.3923	0.4125	0	
YeJ	0.0942	0.0691	0.0717	0.1705	0.1391	0.0479	0.2552	0.1093	0.0925	0.0853	0.1466	0.0906	0.0720	0.3785	0

Níveis de significância:

$P=0.0000$

* $0.000 < P < 0.05$

** $P > 0.05$