

IV. Discussão e Conclusão.

A análise de polimorfismos do ADN é bastante utilizada para o estudo genético de populações, a sua evolução e a sua proximidade filogenética. O ADN mitocondrial é uma ferramenta importante neste estudo filogenético e populacional, uma vez que apresenta características específicas que o diferenciam dos marcadores autossómicos e do cromossoma Y. As análises destes marcadores (autossómicos, cromossoma Y e ADN mitocondrial) complementam-se e permitem, quando comparadas entre si, fazer uma análise mais completa sobre determinada população. De facto, quando analisamos marcadores autossómicos estamos a visualizar a população como um todo. São marcadores comuns às mulheres e homens de uma população e que sofrem recombinação; os marcadores do cromossoma Y são de linhagem exclusivamente paterna, sendo herdados em bloco pelos filhos do sexo masculino. Assim sendo, a análise destes marcadores permite traçar a evolução dos homens de determinada população. O ADN mt aparece como o “complementar” do cromossoma Y, uma vez que apresenta herança exclusivamente materna, não sofre recombinação, permitindo a análise da evolução e migração das mulheres da população em questão. O ADN mt possui ainda uma taxa de mutação elevada, uma grande resistência a fenómenos degradativos, por ser circular, e um elevado número de cópias por célula, características que lhe conferem ainda mais relevância e robustez para estudos relacionados com evolução humana, migrações e o povoamento do planeta, genética das populações e genética forense.

A obtenção de uma base de dados genéticos de sequências das regiões hipervariáveis do ADN mt é importante, mas não é suficiente por si só. Numa perspectiva de evolução populacional é importante comparar os dados obtidos com os disponíveis para outras populações para podermos avaliar a proximidade existente entre as populações, a sua (co)ancestralidade e as migrações populacionais inerentes à origem e evolução das populações analisadas. Numa perspectiva de identificação forense é importante que a população de referência adoptada no processo de identificação seja a mais adequada, para que a frequência de cada sequência seja correctamente determinada. Adicionalmente, as fronteiras físicas actuais podem não corresponder às barreiras sociais e reais de determinada população ou grupo populacional, que coabitando fisicamente no mesmo espaço geográfico que outras populações

ou grupos populacionais (etnias, por exemplo) pode não se relacionar e permanecer geneticamente isolado.

A análise de uma população é sempre mais completa quanto maior for a informação genética disponível. Assim, e como já foi referido, a existência de informação relativa a marcadores autossómicos e / ou do cromossoma Y (linhagem paterna) em complemento à informação obtida pela análise do ADN mt (linhagem materna) é sempre uma mais-valia e deve ser considerada numa análise populacional mais global.

No presente trabalho foram obtidas as sequências das duas regiões hipervariáveis do ADN mt de uma amostragem representativa de dois países geograficamente próximos, Cabo Verde e Guiné-Bissau, as bases de dados de sequências obtidas foram analisadas por si só numa primeira fase e depois comparadas com bases de dados de populações dos vários Continentes, com particular atenção para as populações africanas, europeias, nomeadamente portuguesas e outros países de língua portuguesa em África e na América. Aqui fazemos ainda uma análise conjunta tendo em consideração a informação disponível sobre marcadores do cromossoma Y e autossómicos, considerando estudos que versam sobre a mesma amostragem das populações alvo do presente trabalho e outras publicações sobre as mesmas populações.

1. ANÁLISE DAS REGIÕES HIPERVARIÁVEIS I E II DO ADN MITOCONDRIAL

1.1. ARQUIPÉLAGO DE CABO VERDE

No arquipélago de Cabo Verde foram obtidas as sequências das duas regiões hipervariáveis, HVI e HVII, de 77 indivíduos cuja ascendência materna foi declarada como sendo proveniente das diferentes ilhas do arquipélago e representativa dos diferentes grupos coexistentes no mesmo (Tabela II.1). A análise por ilha das sequências das regiões hipervariáveis HVI e HVII demonstra que o grupo composto pelas ilhas do Barlavento (São Nicolau, São Vicente e Santo Antão) se encontra significativamente distante da maioria do grupo das ilhas do Sotavento, sendo esta distância mais significativa para as ilhas do Fogo e de Santiago, do que relativamente ao grupo composto pelas ilhas Brava e Maio. Esta situação repete-se quando analisamos as regiões hipervariáveis isoladas e em conjunto (Figura III.3).

Este facto poderia ser explicado por uma diferencial povoação destas ilhas, aquando dos descobrimentos portugueses, podendo estar na sua origem um povoamento por escravas africanas provenientes de etnias diferentes das restantes ilhas. Contudo a amostragem disponível e considerada (8 amostras do Barlavento e 6 amostras das ilhas Brava e Maio) é reduzida, não nos permitindo retirar conclusões definitivas sobre a(s) origem(ens) da(s) linhagem(ens) materna(s) destas ilhas e das restantes, isoladamente. Pode exceptuar-se a ilha de Santiago onde possuímos uma amostragem mais significativa, 45 indivíduos.

A amostragem referente ao grupo do Barlavento Caboverdiano apresenta a menor diversidade de sequência (**J**) do que os restantes, em particular quando consideramos apenas a região HVI (Tabela III.3). Mais uma vez este facto pode dever-se a uma baixa amostragem e não ser representativo de todo o Barlavento do arquipélago de Cabo Verde, ou, pode dever-se a uma maior homogeneidade de sequências do ADN mt derivada de uma ascendência da população destas ilhas comum, relativamente limitada quanto ao número de indivíduos e não muito distante temporal e evolutivamente.

No que diz respeito à classificação em haplogrupos, verificamos que 4 amostras do Barlavento pertencem ao haplogrupo **L1c1**, 1 amostra ao **L2a1** e 3 amostras ao **L2c**. A amostragem do Sotavento apresenta uma maior heterogeneidade a nível de haplogrupos, encontrando-se os haplogrupos observados no Barlavento em minoria. A ilha Brava apresenta uma amostragem composta por 4 amostras do haplogrupo **L2**, nomeadamente **L2a1**, **L2b**, **L2c** e **L2d1**. A ilha Maio está representada por 2 amostras dos haplogrupos **L1b1** e **L3e4**. A amostragem das ilhas de Santiago e Fogo apresenta uma maior heterogeneidade de haplogrupos, com amostras representativas de vários sub-haplogrupos dos haplogrupos **L1**, **L2** e **L3**. Quatro amostras provenientes da ilha de Santiago apresentam sequências características de haplogrupos não – africanos, Oeste Euro-Asiático: **X** e **X2d** e Este Asiático: **D4k** e **N9b**. No geral, a grande maioria das amostras de Cabo Verde (94,8 %) pertencem aos haplogrupos Africanos **L0**, **L1**, **L2**, **L3** e **U6**, embora 1 e 2 amostras, respectivamente (Tabela III.2), pertençam aos haplogrupos **L0** e **U6**, reflectindo uma origem Africana da linhagem feminina Caboverdiana.

Os valores encontrados neste estudo estão em concordância com os já reportados para este país por Brehm A, *et al*, em 2002, que referem uma percentagem de haplogrupos Africanos de cerca de 97 %, ligeiramente superior à encontrada neste estudo, sendo a maioria pertencente aos Haplogrupos L1, L2 e L3, caracteristicamente Sub-Saharianos, mas detectando a presença do haplogrupo Norte-Africano, U6, em 3 % das amostras analisadas. Estes autores reportam a presença de haplogrupos não africanos em percentagem diminuta, cerca de 3 %, alguns também detectados neste estudo (X e N*) e outros que não detectámos, H, U5, T, pre-HV e pre*V.

Brehm A, *et al* (2002), analisaram a região HVI de 289 amostras de Cabo Verde e também verificaram a existência de diferenças significativas entre o Sotavento e o Barlavento e uma menor diversidade de sequências (haplótipos) e haplogrupos nas amostras das ilhas do Barlavento caboverdiano. Este fenómeno pode ter origem no facto de estas ilhas terem, em parte, sido povoadas por escravos fugidos das ilhas do Sotavento ou mesmo por negros livres que terão ido para estas ilhas como comerciantes, missionários ou capitães de navios.

No geral, a base de dados de sequências obtida para a população de Cabo Verde apresenta uma amostragem (**N** = 77), diversidades de sequência (**J**) e nucleotídica (**π**) significativas (Tabela III.9, Tabela III.12), constituindo uma ferramenta forense importante para uma correcta identificação de vestígios, recorrendo à análise das regiões hipervariáveis do ADN mt, e uma ferramenta filogenética também relevante para a análise da origem e evolução matrilinear da população deste arquipélago.

1.2. REPÚBLICA DA GUINÉ-BISSAU

Na República da Guiné-Bissau foram obtidas as sequências das duas regiões hipervariáveis, HVI e HVII, de 79 indivíduos de diferentes etnias e cuja ascendência materna é proveniente deste país (Tabela II.2). A análise das sequências das regiões hipervariáveis HVI e HVII da amostragem agrupada por etnia demonstra que, no geral, existe proximidade entre as diferentes etnias. Apenas a etnia Papel apresenta uma distância significativa das etnias Beafada e Manjaco quando consideramos a região HVI e o seu conjunto com a região HVII. A análise isolada da HVII não confirma a distância entre a etnia Papel e Manjaco, posicionando esta última a uma distância significativa das etnias Beafada e Fula (Figura III.6). Contudo, a análise isolada desta região não possui muita relevância quer filogenética, quer forense.

As diversidades de sequência (**J**) calculadas para cada grupo étnico são semelhantes e elevadas, mesmo considerando a baixa amostragem de alguns grupos. O que aparenta suceder é que esta pode ser um pouco exacerbada devido à amostragem ser diminuta aproximando-se bastante de 1. Este valor reflecte uma situação em que número de sequências (**K**) é igual número de amostras (**N**). O facto de terem sido analisadas poucas amostras poderá ser a razão de não haver nessa amostragem sequências comuns a mais do que uma amostra (Tabela III.6).

A amostragem de cada etnia é reduzida, variando entre 1 e 18 amostras por etnia. Assim será precipitado retirar conclusões definitivas com base na análise de grupos étnicos, podemos apenas inferir que, naturalmente, haverá grupos étnicos mais próximos filogeneticamente e outros que possuirão um ancestral comum mais antigo e que por isso apresentarão uma maior distância genética com base na análise realizada.

No que diz respeito à classificação em haplogrupos, não verificamos uma correlação directa entre um determinado haplogrupo e uma ou outra etnia. Contudo, verificamos que as amostras das etnias Beafada e Fula pertencem todas aos haplogrupos **L1** (**L1b1** e **L1c1**) e **L2** (**L2a**, **L2a1**, **L2c** e **L2d1**). Exclui-se uma amostra de etnia Beafada cujo haplogrupo não foi determinado, embora apresente um perfil que se aproxima do característico do haplogrupo **L2c**. As restantes etnias, com amostragem superior a 5 indivíduos, possuem amostras com sequências características de diferentes haplogrupos: as etnias Balanta e Mancanha possuem amostras pertencentes aos haplogrupos **L0**, **L1**, **L2** e **L3**; a etnia Manjaco possui amostras dos haplogrupos **L0**, **L2** e **L3**; e a etnia Papel amostras dos haplogrupos **L1**, **L2** e **L3**. À semelhança

do sucedido com Cabo Verde, também na amostragem Guineense encontramos amostras que apresentam sequências características de haplogrupos não – africanos, Oeste Euro-Asiático: **X2**, **U3** e **R1**, correspondentes a 3 amostras, 2 de etnia Papel e 1 de etnia Balanta, respectivamente. No geral a grande maioria das amostras da Guiné-Bissau (92,4 %) apresentam sequências de ADN mt características dos haplogrupos Africanos **L0**, **L1**, **L2**, **L3** e **U6**, embora com uma prevalência do haplogrupo **U6** residual (Tabela III.5).

Em 2004, Rosa A, *et al*, estudaram a região HVI de uma amostragem alargada desta população (N=372) tendo encontrado uma percentagem de haplogrupos africanos de cerca de 97,3 %, um pouco superior à encontrada neste estudo, sendo maioria (94 % do total) pertencente aos Haplogrupos L0, L1, L2 e L3, tipicamente Sub-Saharianos, mas detectando a presença do haplogrupo Norte-Africano, U6, em 2,2 % das amostras analisadas. Dos haplogrupos africanos detectados na nossa amostragem a sua maioria são comuns ao detectados por Rosa A, *et al*. Exclui-se apenas o haplogrupo L0a, encontrado em duas amostras de etnia Manjaco no presente estudo e que não foi encontrado na amostragem analisada por Rosa A, *et al*, e o haplogrupo L3f1 que não foi detectado na nossa amostragem e cuja presença foi reportada na amostragem Guineense estudada por Rosa A, *et al* (2004).

Estes autores reportam a presença de haplogrupos não africanos em percentagem diminuta, cerca de 2,7 % das amostras pertencem ao haplogrupo U5, que é tipicamente europeu. No presente estudo não encontramos amostras deste haplogrupo, mas verificámos que 3,8 % das amostras possuem um perfil de ADN mt típico de haplogrupos do Oeste Euro-Asiático (X2, U3, R1). Os autores referidos também reportam a existência destes haplogrupos europeus, detectados neste estudo (X e N*) e de outros haplogrupos europeus que não foram detectados no presente trabalho: H, U5, T, pre-HV e pre*V. Apesar das semelhanças entre o presente estudo e o de Rosa A, *et al*, considerando as duas populações Guineenses estas apresentam algumas discrepâncias e distância entre elas, o que poderá ser explicado pelo facto da amostragem ser diferente entre os estudos no que diz respeito às etnias representadas.

A base de dados de sequências obtida para a população da Guiné-Bissau no seu geral apresenta uma amostragem (N = 79), diversidades de sequência (J) e nucleotídica (π) significativas (Tabela III.9, Tabela III.12), constituindo uma ferramenta forense importante para uma correcta identificação de vestígios, recorrendo à análise das regiões hipervariáveis do ADN mt, e uma ferramenta filogenética também relevante para a análise da origem e evolução matrilinear da população desta república do oeste africano.

1.3. COMPARAÇÃO ENTRE CABO VERDE E GUINÉ-BISSAU

As sequências nucleotídicas do ADN mt das amostras Caboverdianas e Guineenses são caracteristicamente Africanas, dado confirmado pela sua classificação em haplogrupos. As populações possuem várias sequências em comum, 13, 12 e 6 considerando as regiões HVI, HVII e HVI+HVII, respectivamente, e partilham a maioria dos haplogrupos e sub-haplogrupos encontrados, sendo os haplogrupos L1b1; L2a1; L2c; L3b; L3d e L3e4 os mais significativos em ambas as populações, respectivamente 65% e 59% das amostras das populações de Cabo Verde e Guiné-Bissau. Contudo, nem todos estes haplogrupos se encontram na mesma proporção nas duas populações, o haplogrupo L2c é muito mais significativo em Cabo Verde do que na Guiné-Bissau, 21% *versus* 8%. O haplogrupo é ligeiramente mais frequente em Cabo Verde, 10% *versus* 8%. Os haplogrupos L1b1, L2a1 e L3d encontram-se com uma frequência ligeiramente maior (2 a 4%) na população da Guiné-Bissau, 11%, 13% e 10% *versus* 8%, 9% e 8% em Cabo Verde, respectivamente. O haplogrupo L3e4 possui uma frequência significativa em ambas as populações e está representado numa proporção equivalente, cerca de 9% das amostras pertencem a este haplogrupo. Os haplogrupos não africanos presentes nas populações estudadas não são comuns e apresentam uma prevalência muito diminuta em ambas (Figura III.1, Figura III.3 e Tabela III.8).

Estas diferenças da composição em haplogrupos entre Cabo Verde e Guiné-Bissau, que embora sejam, na sua grande maioria, comuns às duas populações, se encontram com uma frequência diferente, suportam a hipótese de que o povoamento das diferentes ilhas de Cabo Verde foi diferenciado.

As populações possuem parâmetros de diversidade muito semelhantes, e posições mais polimórficas muito coincidentes e com uma percentagem muito semelhante, o que aparenta estar de acordo com a sua proximidade geográfica. Ambas apresentam taxas de Transversão:Transição elevadas, apresentando uma razão superior na população Caboverdiana para a região HVI e na população Guineense para a região HVII.

O fenómeno de Heteroplasmia de comprimento ocorre em ambas as populações nas duas regiões hipervariáveis. Sempre que uma amostra apresenta uma heteroplasmia de comprimento na HVI está, em todos os casos, associada a uma transição T→C na posição 16189 (Bendall & Sykes, 1995).

A análise da matriz de distâncias obtida pelo método de *Pairwise difference*, programa Arlequin 3.11 (Excoffier L, *et al*, 2005), não revela a existência de uma distância significativa entre a população de Cabo Verde e da Guiné-Bissau na sua globalidade. Uma análise sub-populacional revela que existe distância significativa ($P > 0,05$) entre algumas ilhas do arquipélago Caboverdiano e algumas etnias Guineenses. O grupo das ilhas do Barlavento representadas no presente estudo apresenta-se distante da quase totalidade de etnias Guineenses, excluindo-se as etnias Beafada e Fula. Por sua vez a etnia Manjaco apresenta proximidade apenas ao grupo composto pelas ilhas Brava e Maio. Adicionalmente, a ilha Fogo apresenta-se significativamente distante das etnias Beafada e Manjaco. Esta análise é comum ao conjunto das duas regiões hipervariáveis HVI+HVII e à região HVI isoladamente, com exceção do distanciamento entre a ilha Fogo e a etnia Manjaco, o que não se verifica na análise individual da região hipervariável I (Anexo V).

1.4. COMPARAÇÃO GLOBAL

As sequências da região hipervariável I (16090 – 16365) e as sequências resultantes da conjugação desta com a região HVII (073 – 340) das amostras de Cabo Verde ($N = 77$) e da Guiné-Bissau ($N = 79$) foram comparadas com as sequências publicadas para populações dos vários continentes. A análise da região HVI isoladamente permitiu uma comparação com um maior número de populações, 120 no total, sendo algumas referentes ao mesmo país, mas com uma amostragem diferente. Quando analisamos as duas regiões em conjunto o número de populações disponível para comparação reduz drasticamente: no actual estudo comparamos as sequências conjuntas HVI + HVII de Cabo Verde e Guiné-Bissau com um total de 55 populações dos cinco continentes (Tabela III.10).

A existência de sequências comuns entre duas ou um grupo de populações pode indicar alguma proximidade histórica e / ou geográfica entre estas. Contudo, quando comparamos as sequências das regiões hipervariáveis temos de ter em consideração a elevada taxa de mutação do ADN mt, o facto de duas populações não possuírem sequências comuns não é por si só indicativo de uma grande distância filogenética entre elas. Outro factor importante a ter em conta nesta análise comparativa é o tamanho da base de dados em questão; quando comparamos uma população com outra o número de sequências comuns pode ser mais ou menos significativo consoante o número de sequências diferentes (K) encontradas nessa mesma população, e até a frequência dessas sequências.

A análise comparativa simples das sequências da região HVI e do conjunto das regiões HVI + HVII das populações de Cabo Verde e da Guiné-Bissau revelam claramente uma maior partilha de sequências com populações de África, seguindo-se a Europa (Oeste), América (EUA, México, Caraíbas e Brasil) e a Ásia (Médio Oriente e Índia). A análise ponderada dada pela taxa de sequências comuns, calculada pela percentagem de sequências comuns entre as populações estudadas e o número de sequências diferentes (K) encontradas para cada uma das populações consideradas para esta comparação, confirma que existem mais sequências comuns com populações de África, especialmente com populações do Oeste Africano. A taxa de sequências comuns é superior entre as populações estudadas e as populações Americanas, com sequências comuns, do que com as populações Europeias que partilham sequências com as populações de Cabo Verde e da Guiné-Bissau. Entre as populações estudadas e as

populações asiáticas da Arábia Saudita e Índia, esta taxa não é superior a 2 % na HVI e é residual ou inexistente para o conjunto da HVI e da HVII (Tabela IV.1, Tabela IV.2).

Tabela IV.1 – Populações que partilham sequências com Cabo Verde e Guiné-Bissau, para a região HVI (16090 – 16365), Número e Taxa de sequências comuns com base no número de sequências diferentes (**K**) de cada população.

Código	População	HVI		Cabo Verde		Guiné-Bissau	
		Nº amostras (N)	Nº sequências diferentes (K)	Nº sequências comuns	Taxa sequências comuns (%)	Nº sequências comuns	Taxa sequências comuns (%)
1	África do Sul	74	19	1	5,26 %	1	5,26 %
4	Moçambique 1	307	115	8	6,96 %	8	6,96 %
5	Moçambique 2	109	50	4	8,00 %	2	4,00 %
9	Angola 1	44	34	4	11,76 %	3	8,82 %
10	Angola 2	59	43	5	11,63 %	3	6,98 %
11	Angola - Cabinda	110	73	4	5,48 %	2	2,74 %
12	São Tomé	103	61	11	18,03 %	9	14,75 %
13	São Tomé e Príncipe	50	32	5	15,63 %	5	15,63 %
14	Guiné Equatorial 1	11	10	2	20,00 %	1	10,00 %
15	Guiné Equatorial 2	45	18	1	5,56 %	2	11,11 %
16	Camarões	104	66	6	9,09 %	7	10,61 %
17	Oeste África	160	108	6	5,56 %	9	8,33 %
18	Serra Leoa	277	147	14	9,52 %	14	9,52 %
19	Guiné-Bissau 1	372	180	26	14,44 %	27	15,00 %
20	Guiné-Bissau 2	79	59	14	23,73 %	59	100,00 %
21	Cabo Verde 1	289	114	29	25,44 %	20	17,54 %
22	Cabo Verde 2	77	52	52	100,00 %	14	26,92 %
23	Senegal 1	111	47	8	17,02 %	8	17,02 %
24	Senegal 2	121	85	14	16,47 %	13	15,29 %
25	Este África	88	75	2	2,67 %	3	4,00 %
26	Quênia	100	81	3	3,70 %	5	6,17 %
28	Mauritânia	30	22	4	18,18 %	5	22,73 %
29	Saara Ocidental	25	20	3	15,00 %	4	20,00 %
30	Marrocos 1	92	62	3	4,84 %	2	3,23 %
32	Argélia	85	15	1	6,67 %	2	13,33 %
33	Egipto	58	47	1	2,13 %	1	2,13 %
34	Espanha - Canárias	54	42	2	4,76 %	1	2,38 %
35	Portugal - Arq. Madeira	155	87	5	5,75 %	5	5,75 %
36	Portugal - Açores	64	45	1	2,22 %	1	2,22 %
37	Portugal - Açores 2	94	39	2	5,13 %	1	2,56 %
38	Portugal 1	106	68	1	1,47 %	2	2,94 %
39	Portugal 2	241	141	3	2,13 %	1	0,71 %

Tabela IV.1 – Populações que partilham sequências com Cabo Verde e Guiné-Bissau, para a região HVI (16090 – 16365), Número e Taxa de sequências comuns com base no número de sequências diferentes (**K**) de cada população.

Código	População	HVI		Cabo Verde		Guiné-Bissau	
		Nº amostras (N)	Nº sequências diferentes (K)	Nº sequências comuns	Taxa sequências comuns (%)	Nº sequências comuns	Taxa sequências comuns (%)
51	Itália 2	49	32	0	0,00 %	1	3,13 %
56	Polónia	435	218	2	0,92 %	2	0,92 %
72	Arábia Saudita	556	266	5	1,88 %	4	1,50 %
82	Índia 1	96	51	1	1,96 %	1	1,96 %
94	Canadá + EUA	63	27	1	3,70 %	1	3,70 %
95	EUA 1	326	74	7	9,46 %	9	12,16 %
99	México	85	64	2	3,13 %	2	3,13 %
100	Caraíbas	93	54	7	12,96 %	8	14,81 %
108	Brasil 1	247	164	4	2,44 %	8	4,88 %
109	Brasil 2	167	121	5	4,13 %	9	7,44 %

Entre as populações estudadas existe uma taxa de sequências comuns elevada apresentando valores entre os 25 % e 27 % quando consideramos apenas a região HVI e valores entre os 11 % e 13 % quando consideramos o conjunto das duas regiões hipervariáveis estudadas, HVI + HVII (Tabela IV.1, Tabela IV.2). Esta taxa é sempre superior para as populações do Oeste Africano com sequências comuns e Angola, São Tomé e Príncipe e Moçambique, que apresentam uma taxa de sequências comuns significativa, a qual é ligeiramente superior para Cabo Verde. A partilha de sequências com populações do Norte de África também existe, sendo mais significativa para a população da Guiné-Bissau, com destaque para a população da Mauritânia e do Saara Ocidental (Tabela IV.1).

O número de sequências partilhadas com as populações Europeias, Asiáticas e Americanas é residual, excluindo-se as populações dos EUA 1 (Índios Apaches e Navajos), Caraíbas e Brasil com taxas de sequências comuns significativas, as quais apresentam valores sempre superiores quando comparadas com a população Guineense em detrimento da população de Cabo Verde (Tabela IV.1)

A análise conjunta das duas regiões hipervariáveis, HVI + HVII, apresenta um maior poder discriminatório, pois aumenta significativamente a variabilidade de sequências dentro de uma população. Esta análise confirma a análise da região HVI isoladamente na sua generalidade. Contudo, o conjunto de populações disponível para comparação é significativamente menor e a análise populacional comparativa da conjugação das duas regiões hipervariáveis apresenta taxas de sequências comuns muito inferiores. Considerando o conjunto das regiões HVI e HVII, verificamos que o número de sequências comuns entre Cabo Verde e a Guiné-Bissau com populações Europeias é quase nula. A população dos EUA 1 (Índios Apaches e Navajos) não apresenta nenhuma sequência comum com qualquer uma das populações estudadas. Adicionalmente, a quantidade de sequências partilhadas com o Brasil e as Caraíbas é bastante reduzida, sendo praticamente inexistente para Cabo Verde (Tabela IV.2).

Tabela IV.2 – Populações que partilham sequências com Cabo Verde e Guiné-Bissau, para o conjunto das regiões HVI (16090 – 16365) + HVII (073 – 340), Número e Taxa de sequências comuns com base no número de sequências diferentes (**K**) de cada população.

Código	População	HVI + HVII		Cabo Verde		Guiné-Bissau	
		Nº amostras (N)	Nº sequências diferentes (K)	Nº sequências comuns	Taxa sequências comuns (%)	Nº sequências comuns	Taxa sequências comuns (%)
1	África do Sul	74	21	0	0,00 %	1	4,76 %
5	Moçambique 2	109	63	2	3,17 %	2	3,17 %
10	Angola 2	59	50	2	4,00 %	1	2,00 %
11	Angola - Cabinda	110	87	1	1,15 %	1	1,15 %
12	São Tomé	103	70	7	10,00 %	5	7,14 %
20	Guiné-Bissau 2	79	67	8	11,94 %	67	100,00 %
22	Cabo Verde 2	77	66	66	100,00 %	8	12,12 %
23	Senegal 1	111	52	5	9,62 %	4	7,69 %
26	Quênia	100	89	0	0,00 %	4	4,49 %
36	Portugal - Açores 1	64	53	1	1,89 %	0	0,00 %
37	Portugal - Açores 2	94	49	0	0,00 %	0	0,00 %
38	Portugal 1	106	81	0	0,00 %	1	1,23 %
39	Portugal 2	241	179	0	0,00 %	0	0,00 %
72	Arábia Saudita	556	330	1	0,30 %	1	0,30 %
82	Índia 1	96	55	0	0,00 %	0	0,00 %
95	EUA 1	326	92	0	0,00 %	0	0,00 %
100	Caraíbas	93	62	1	1,61 %	5	8,06 %
108	Brasil 1	247	84	0	0,00 %	1	1,19 %
109	Brasil 2	167	133	1	0,75 %	3	2,26 %

Como foi referido, a análise única de sequências comuns, mesmo que ponderando a sua significância dentro das bases de dados existentes, não é suficiente para concluirmos sobre a proximidade filogenética entre populações. Uma análise mais cuidada com base na Matriz de distâncias F_{ST} , os valores de P respectivos e a sua significância permite inferir com maior precisão sobre a proximidade ou distanciamento filogenético das populações e desenhar árvores filogenéticas que o ilustrem (Figura III.8, Anexo V e VI).

Para a população de Cabo Verde, a análise da significância dos valores P da matriz de distâncias F_{ST} obtidos, para a região HVI, revela uma proximidade genética entre a sua linhagem materna da Guiné-Bissau (020), $F_{st} = 0,00052$ com $P = 0,40541+0,0339$, e também com as populações do Senegal (023 e 024), $F_{st} = 0,00947$ com $P = 0,06306+0,0273$ e $F_{st} = 0,00488$ com $P = 0,09009+0,0303$, respectivamente. Para a HVI+HVII, confirma-se a proximidade entre Cabo Verde e a Guiné-Bissau (020), $F_{st} = 0,00290$ com $P = 0,20721+-0,0273$ (Anexo V, Rando J C, *et al*, 1998; Graven L, *et al*, 1995).

Estes achados corroboram a hipótese da colonização de Cabo Verde ter sido efectuada recorrendo essencialmente a escravos oriundos da Costa Oeste Africana Senegalesa, ou seja, do actual Senegal e Guiné-Bissau, em detrimento de outros locais de África que à data também estariam sob o domínio português. Não podemos distinguir se houve um maior contributo da região da Guiné-Bissau do que do Senegal, como seria esperado, uma vez que estas duas populações também se encontram filogeneticamente próximas.

A comparação entre a população de Cabo Verde estudada e a já publicada por Brehm A, *et al*, em 2002 (021), revela proximidade entre estes dois estudos, $F_{st} = -0,00070$ com $P = 0,47748+0,0471$, o que é concordante com a análise dos dados provenientes da classificação em haplogrupos (Anexo V).

A linhagem materna Guineense apresenta proximidade com as linhagens maternas de várias populações africanas, continentais e insulares e de uma população do Continente Americano que foi uma colónia Espanhola. Para a HVI verifica-se a proximidade entre a Guiné-Bissau e São Tomé e Príncipe (013), $F_{st} = 0,00196$ com $P = 0,32432+0,0492$, Serra Leoa (018),

Fst = 0,00452 com P = 0,07207+0,0264, Cabo Verde (022), Fst = 0,00052 P = 0,40541+0,0339, Caraíbas (100), Fst = 0,00818 com P = 0,08108+0,0252, e uma população representativa de vários países do Oeste de África (017), Fst = 0,00104 com P = 0,29730+0,0305. A proximidade com Cabo Verde (022), Fst = 0,00290 com P = 0,20721+0,0273, e com as Caraíbas (100), Fst = 0,00742 com P = 0,05405+0,0201, confirma-se na análise da HVI+HVII (Anexo V, Jackson B A, *et al*, 2005; Salas A, *et al*, 2005; Mateu E, *et al*, 1997; Watson E, *et al*, 1997; Watson E, *et al*, 1996).

A proximidade existente entre a população Guineense e outras populações do Continente Africano será reflexo das migrações dos povos de África e apoia a teoria que os povos do Oeste Africano terão tido uma origem comum e que terão ocorrido vários movimentos migratórios dentro da região. Da análise da matriz de distâncias verifica-se também proximidade a mais duas populações insulares, além de Cabo Verde, nomeadamente à população de São Tomé e Príncipe, situada a sul da Guiné-Bissau ao largo da costa da Guiné Equatorial e do Gabão, e à população das Caraíbas, conjunto de ilhas ao largo da América Central, outrora colonizadas pelos espanhóis. Esta proximidade evidencia uma contribuição significativa de povos Guineenses, escravos, para a colonização destes arquipélagos.

A população da República Guineense estudada no presente trabalho apresenta-se filogeneticamente próxima da população Guineense estudada em 2004 por Rosa A, *et al*, Fst = 0,00015 com P = 0,35135+0,0684 (Anexo V).

Ao contrário do esperado, estas duas populações da Guiné-Bissau (019 e 020), a estudada por Rosa A, *et al*, em 2004, e a que se refere o presente trabalho, apresentam uma distância filogeneticamente significativa da População Caboverdiana (021) descrita por Brehm A, *et al*, em 2002, embora esta se apresente próxima da população de Cabo Verde aqui estudada (Anexo V). Esta situação pode suceder pelo facto de estes dois estudos não possuírem uma amostragem da população Caboverdiana equivalente. Uma justificação provável é a existência de diferenças significativas entre as populações das diferentes ilhas que depois se reflectem na análise global do arquipélago, de acordo com a amostragem de cada ilha que foi incluída no estudo. Mais uma vez, podemos estar a visualizar um fenómeno associado à diferente colonização das ilhas do arquipélago por grupos étnicos africanos que se encontram histórica e filogeneticamente distantes, embora partilhem o espaço geográfico.

Da análise da matriz de distâncias F_{st} é também evidente a proximidade entre as populações de Angola, São Tomé e Príncipe e Guiné Equatorial, reflectindo um povoamento do arquipélago de São Tomé com povos continentais provenientes, também, de Angola e da Guiné Equatorial e uma proximidade filogenética entre estes dois países do continente Africano (Anexo V, Plaza S, *et al*, 2004; Beleza S, *et al*, 2005; Trovoada M J, *et al*, 2003; Mateu E, *et al*, 1997; Pinto F, *et al*, 1996).

As populações Europeias consideradas nesta análise comparativa apresentam, na sua generalidade, pouca divergência filogenética entre si e com as populações do Norte de África, confirmando a proximidade entre elas. Embora, algumas populações se destaquem por se mostrarem filogeneticamente distantes de praticamente todas as restantes, nomeadamente, a população dos Açores descrita por Santos C, *et al*, em 2005, as populações Espanholas do País Basco e da Galiza, a população da Croácia, da Noruega, da Finlândia e da Suécia (Anexo V, Carvalho M, *et al*, 2006; Poetsch M, *et al*, 2004; Poetsch M, *et al*, 2003; Brehm A, *et al*, 2003; Carvalho M, *et al*, 2003; Malyarchuk B A, *et al*, 2002; Tolk H-V, *et al*, 2001; Vona G, *et al*, 2001; Malyarchuk B A & Derenko M V, 2001; Kouvatsi A, *et al*, 2001; Salas A, *et al*, 2000; Belledi M, *et al*, 2000; Saillard J, *et al*, 2000; Kittles R A, *et al*, 1999; Rando J C, *et al*, 1998; Salas A, *et al*, 1998; Rousselet F & Mangin, 1998; Delghandi M, *et al*, 1998; Lutz S, *et al*, 1998; Baasner A, *et al*, 1998; Parson W, *et al*, 1998; Crespiello M, *et al*, 1997; Calafell F, *et al*, 1996; Francalacci P, *et al*, 1996; Torroni A, *et al*, 1996; Pinto F, *et al*, 1996; Côrte-Real H B S M, *et al*, 1996; Bertranpetit J, *et al*, 1995; Piercy R, *et al*, 1993).

Na Oceânia encontramos também um conjunto de populações que se apresentam filogeneticamente próximas, a população de Samoa, da Papuásia Nova Guiné e a população de Ateoroa – Nova Zelândia (Anexo V, Redd A J, *et al*, 1995; Sykes B, *et al*, 1995).

As árvores filogenéticas ‘unrooted’ obtidas para a análise da região HVI e a análise conjunta da HVI e HVII, após selecção de algumas populações representativas, apresentam uma organização quase continental. Verificamos a presença de um ramo quase exclusivo das populações africanas, em que as populações do Sul de África (003) e Botswana (002) são as que mais se distanciam das restantes populações. Seguem-se as populações Europeias, Asiáticas, Americanas e da Oceânia, embora estas apresentem alguma miscigenação a nível de localização na árvore. Esta disposição parece concordar com a Hipótese ‘*out of Africa*’ e

representar o que terão sido as migrações do Homem aquando do povoamento do planeta, sendo a aparente mistura de populações de diferentes países reveladora de vários movimentos migracionais sucessivos e não só no sentido África → Europa → Ásia → Oceânia / América, mas, por vezes, no sentido contrário também (Figura III.8, Shriver M D & Kitles R A, 2004, Grande Atlas Universal, 1997, O Novo Atlas do Mundo, 1992, A Aventura da Vida, 1989).

As árvores obtidas para a região HVI e para o conjunto HVI+HVII são semelhantes, contudo, a população da Guiné-Bissau apresenta-se mais distante de Cabo Verde na análise em que consideramos apenas a região HVI. O mesmo sucede com a localização das populações portuguesas consideradas ao desenhar estas árvores: apresentam-se mais próximas quando consideramos as duas regiões hipervariáveis. De salientar é também a proximidade que as populações de Moçambique (005) e Angola (010) apresentam relativamente às populações de Cabo Verde e Guiné-Bissau estudadas (Figura III.8, Pereira L, *et al*, 2001).

A análise da árvore filogenética 'em bruto', referente à região HVI, permite distinguir um grupo composto essencialmente pelas populações oriundas de África (001 a 033), onde aparecem algumas populações Americanas que estarão filogeneticamente próximas deste continente, nomeadamente, duas referentes a dois estudos sobre o Brasil (108 e 109), a população das Caraíbas (100) e uma população dos Estados Unidos da América (097), referente a Afro-Americanos aí residentes, e em que três populações de ilhas da Oceânia também aparecem relativamente próximas: Papuásia Nova Guine 1 (113), Tonga (117) e Tahiti + Ilhas Austrais (120+121). Apenas as populações de Marrocos (030 e 031) e Argélia (032) não se encontram neste grupo, estando localizadas entre populações Europeias. Esta análise é também válida quando consideramos a região HVII, embora visando muito menos populações (Figura III.8, Anexo VI, Barbosa A B G, *et al*, 2008; Berniell-Lee G, *et al*, 2008; Ely B, *et al*, 2006; Salas A, *et al*, 2005; Beleza S, *et al*, 2005; Jackson B A, *et al*, 2005; Brandstätter A, *et al*, 2004; Rosa A, *et al*, 2004; Plaza S, *et al*, 2004; Cerný V, *et al*, 2004; Trovoada M J, *et al*, 2003; Stevanovitch A, *et al*, 2003; Brehm A, *et al*, 2002; Salas A, *et al*, 2002; Pereira L, *et al*, 2001; Alves-Silva J, *et al*, 2000; Chen Y-S, *et al*, 2000; Quintana-Murci L, *et al*, 1999; Rando J C, *et al*, 1998; Vigilant L, *et al*, 1990 in Watson E, *et al*, 1997; Mateu E, *et al*, 1997; Watson E, *et al*, 1997; Côrte-Real H B

S M, et al, 1996; Watson E, et al, 1996; Pinto F, et al, 1996; Soodyall H, et al, 1996; Sykes B, et al, 1995; Graven L, et al, 1995; Vigilant L, et al, 1989).

Dentro do grupo composto por populações Africanas podemos identificar as populações alvo deste estudo num subgrupo constituído essencialmente por populações do Oeste Africano. A proximidade das linhagens maternas destas populações e a sua ligação ao continente africano, nomeadamente à costa Oeste Africana, verificada neste estudo e outros, é indicativa de que a povoação feminina do arquipélago de Cabo Verde terá sido realizada recorrendo a mulheres, escravas, vindas de África, nomeadamente de tribos que habitavam na costa ocidental africana próxima das ilhas desabitadas que hoje constituem o arquipélago Caboverdiano, e com um significativo contributo da Guiné-Bissau na altura também colonizada por Portugal (Jackson B A, et al, 2005; Rosa A, et al, 2004; Brehm A, et al, 2002).

A classificação em Haplogrupos das sequências das regiões hipervariáveis do ADN mt das populações estudadas confirma a sua proximidade à região oeste do Continente Africano e revela uma composição característica desta região. Os Haplogrupos encontrados nestas populações são maioritariamente Sub-Saharianos: L1a-L1k; L2 e L3A (todos os haplogrupos L3 com excepção dos haplogrupos M e N); e incluem alguns relatados como sendo específicos do antigo povo Bantu: L0a, L1a, L1e, L2, L3b e L3e (Salas A, et al, 2002; Bandelt H-J, et al, 2001; Pereira L, et al, 2001; Alves-Silva J, et al, 2000; Chen Y-S, et al, 2000; Soodyall H, et al, 1996).

No geral a composição em Haplogrupos e/ou sub-haplogrupos verificada neste estudo para a Guiné-Bissau e Cabo Verde apresenta um perfil semelhante ao reportado para as populações da costa Ocidental Africana e para o conjunto destas (Salas A, et al, 2002).

Ambas as populações, Cabo Verde e Guiné-Bissau, apresentam uma presença residual, inferior a 3 %, do haplogrupo U6, nomeadamente U6a, característico do Norte de África e de haplogrupos característicos do Oeste Euro-asiático. Este fenómeno poderá ser explicado pela ocorrência de um pequeno fluxo populacional da região do Noroeste Africano, migrações Berberes, ou da Península Ibérica, Portugal e Espanha, durante o período dos descobrimentos (Coia V, et al, 2005; Rosa A, et al, 2004; Cerný V, et al, 2004; Maca-Meyer N, et al, 2003; Brehm

A, *et al*, 2002; Salas A, *et al*, 2002; Quintana-Murci L, *et al*, 1999; Rando J C, *et al*, 1998; Watson E, *et al*, 1997; Pinto F, *et al*, 1996; Graven L, *et al*, 1995).

A presença do Haplogrupo L0a e L0a1 na Guiné-Bissau, aproximadamente 2,5 % e 5 % da amostragem, já havia sido reportada: Foi sugerida a sua origem num fluxo migratório, no fim da última era glacial (Holoceno), vindo do Este e Centro de África e é reportado como uma evidência da origem sudanesa da etnia Balanta. Contudo, no presente trabalho, estes haplogrupos aparecem em 4 amostras de etnia Manjaco, uma amostra de etnia Mancanha e 1 amostra de etnia Balanta (Rosa A, *et al*, 2004). O haplogrupo L0a1 está presente em apenas uma amostra de Cabo Verde e não foi encontrado no estudo realizado por Brehm A, *et al*, em 2002.

O haplogrupo L1b é característico do Oeste Africano, sendo o sub-haplogrupo L1b1 o mais comum quer nas populações do presente estudo quer noutras populações do Oeste de África (Rosa A, *et al*, 2004; Brehm A, *et al*, 2002; Salas A, *et al*, 2002; Rando J C, *et al*, 1998; Graven L, *et al*, 1995).

O haplogrupo L2: L2a, L2b, L2c e L2d, é originário do Oeste Africano, sendo os três últimos sub-haplogrupos comuns à região Centro-Oeste de África. Há relatos da presença destes haplogrupos em diversas populações da costa Ocidental Africana (Jackson B A, *et al*, 2005; Rosa A, *et al*, 2004; Brehm A, *et al*, 2002; Salas A, *et al*, 2002; Rando J C, *et al*, 1998; Watson E, *et al*, 1997; Graven L, *et al*, 1995). O haplogrupo L2b estende-se a Este e Norte de África sendo também detectado na Península Ibérica (Salas A, *et al*, 2002). Os sub-haplogrupos L2c2 e L2d2 encontrados em apenas 1 amostra Guineense e 4 amostras Caboverdianas foram reportados apenas na população de Serra Leoa no Oeste Africano. A sua presença em Cabo-Verde apoia os registos históricos que reportam a povoação deste arquipélago recorrendo a escravos oriundos da região entre o Senegal e a Serra Leoa (Jackson B A, *et al*, 2005). O sub-haplogrupo L2c2 foi também encontrado em 2 amostras oriundas da região de Cabinda – Angola; 3 amostras de São Tomé e Príncipe e o sub-haplogrupo L2d2 numa amostra de São Tomé e Príncipe (Beleza S, *et al*, 2005; Trovoadá M J, *et al*, 2003)

O Paragrupa L3 terá tido a sua origem no Este de África, tendo depois migrado para praticamente todo o Continente Africano e através do Globo, derivando em muitos haplogrupos distintos (Chen Y-S, *et al*, 2000; Watson E, *et al*, 1997). Os haplogrupos L3b e L3d são comuns nas populações do Oeste Africano, verificando-se por vezes a presença do sub-haplogrupa do L3f, característico do Este Africano, L3f1 e uma presença até significativa nesta região do sub-haplogrupa do L3e, bastante disperso em África, L3e4. A amostragem do presente reflecte claramente esta situação (Salas A, *et al*, 2002; Bandelt H-J, *et al*, 2001; Watson E, *et al*, 1997).

A amostragem de Cabo Verde e da Guiné-Bissau analisada no presente estudo apresenta alguma proximidade, sequências e haplogrupos comuns, com outras populações do Continente africano, essencialmente ex-províncias ultramarinas portuguesas (Moçambique, Angola e São Tomé e Príncipe), a razão poderá ser o tráfico de escravos entre colónias ou rotas comuns às várias colónias. De facto, verifica-se em Cabo Verde a existência de um maior número de sequências homólogas a estas antigas colónias portuguesas relativamente à Guiné-Bissau, embora esta diferença seja baixa apoia a hipótese de este arquipélago também ter sido povoado com recurso a escravos de distintas regiões de África (Tabela IV.1, Tabela IV.2). Esta proximidade populacional entre a Costa Oeste e o Este de África poderá ter também origem, em parte, num período histórico muito anterior, as migrações de povos africanos autóctones, e em migrações mais recentes entre colónias (Rosa A, *et al*, 2004; Plaza S, *et al*, 2004; Brehm A, *et al*, 2002; Salas A, *et al*, 2002; Bandelt H-J, *et al*, 2001; Pereira L, *et al*, 2001; Watson E, *et al*, 1997; Mateu, *et al*, 1997).

Alguns dos haplogrupos encontrados nas populações estudadas e considerados tipicamente Africanos, L1b; L2a; L3b e L3e, aparecem também em populações do continente americano com alguma frequência, nomeadamente em comunidades Afro-Americanas e populações do Centro e Sul da América. A análise de sequências revela também algumas homólogas com estas populações Americanas, mais uma vez com uma maior frequência para a população de Cabo Verde, à semelhança do que sucede com outras populações africanas de língua Portuguesa (Barbosa A B G, *et al*, 2008; Ely B, *et al*, 2006; Salas A, *et al*, 2005; Plaza S, *et al*, 2004; Salas A, *et al*, 2002; Bandelt H-J, *et al*, 2001; Pereira L, *et al*, 2001).

A presença de haplogrupos ou sequências tipicamente africanas é também verificada na Península Ibérica, incluindo Portugal, explicada pelo facto de alguns escravos oriundos de África terem também vindo para Portugal, essencialmente para o sul e para as ilhas com o objectivo de as povoar, e ter ocorrido o cruzamento entre homens portugueses caucasóides e escravas negróides (Tabela IV.1, Tabela IV.2; Brehm A, *et al*, 2003; Santos C, *et al*, 2003; Salas A, *et al*, 2002; Pereira L, *et al*, 2000).

No geral, os resultados obtidos no presente trabalho, quer sob a forma de árvores filogenéticas, partilha de sequências entre populações distintas e parâmetros estatísticos (diversidade de sequência - J - e diversidade nucleotídica - π -), são concordantes com a hipótese do surgimento do Homem em África seguida de expansão para outros continentes, ocorrendo casualmente alguma miscigenação de povos em várias fases da evolução e história do Homem, apoiando a teoria de uma origem única do homem moderno '*out of África*' e não a teoria multirregional do aparecimento do homem (Purves, *et al*, 2001; Jorde LB, *et al*, 1998; Jorde LB, *et al*, 1995).

2. COMPARAÇÃO COM MARCADORES DO CROMOSSOMA Y

As populações estudadas no presente trabalho já foram alvo de estudo, em conjunto com outras populações de expressão Portuguesa, de marcadores autossómicos e do cromossoma Y por Corte Real F em 1999, tendo-se partido da amostragem recolhida para esses estudos para a realização do presente trabalho.

As amostras de Cabo Verde ($N = 47$) estudadas por Corte Real F (1999) para a determinação dos microssatélites do cromossoma Y apresentam a seguinte proveniência da linhagem paterna:

- Negros: 26 pessoas
- Mestiços: 16 pessoas
- Brancos: 5 pessoas.

As amostras da Guiné-Bissau ($N = 33$) estudadas por Corte Real F (1999) para a determinação dos microssatélites do cromossoma Y apresentam a seguinte proveniência da linhagem paterna:

- etnia Balanta: 7 pessoas
- etnia Papel: 6 pessoas
- etnia Manjaco: 5 pessoas
- etnia Fula: 3 pessoas
- etnia Mancanha: 3 pessoas
- etnia Beafada: 2 pessoas
- etnia Mandinga: 1 pessoa
- etnia Bijagós: 1 pessoa
- etnia Cassanga: 1 pessoa
- etnia Banhune: 1 pessoa
- não sabe ou não responde: 3 pessoas.

Foi estudada ainda amostragem proveniente de Angola (Norte), Moçambique (Sul), São Tomé e Príncipe, Portugal (Centro) e duas populações do Brasil (Amazónia e São Paulo) e estudadas as relações de proximidade (coancestralidade) e semelhança genética entre elas para vários marcadores do cromossoma Y.

Da análise e comparação entre populações efectuada por Corte-Real F (1999), baseada nas diferenças estatisticamente significativas, análise de haplótipos comuns, distâncias de Nei e valores de F_{st} , foi verificado que podia agrupar estas populações em três grupos: um grupo Ibero-Americano, e dois grupos Africanos, o grupo das populações Africanas Continentais e o grupo das populações Africanas Insulares.

O grupo Ibero-Americano é composto pela população de Portugal e as duas populações do Brasil, Amazónia e São Paulo, cujas linhagens paternas se encontram filogeneticamente próximas. Este grupo é caracterizado por apresentarem linhagens e haplótipos de cromossoma Y caracteristicamente e maioritariamente caucasianos.

O grupo de populações Africanas Continentais é composto pela população da Guiné-Bissau, Angola (Norte) e Moçambique (Sul). Estas populações apresentam uma grande diferença e distância filogenética das populações Ibero-Americanas analisadas e apresentam linhagens e haplótipos de cromossoma Y caracteristicamente e maioritariamente africanos.

O grupo das populações Africanas Insulares é composto pela população de Cabo Verde e de São Tomé e Príncipe, estas aparecem numa posição intermédia entre as populações Africanas Continentais e as populações Ibero-Americanas que se encontram nitidamente distantes, verificando-se contudo uma maior proximidade entre as linhagens do cromossoma Y das populações Africanas Insulares e as das populações Ibero-Americanas do que com as das populações Africanas Continentais, pois apresentam haplótipos de cromossoma Y maioritariamente caucasianos.

De realçar o facto de entre as populações Africanas insulares, Cabo-Verde aparecer sempre mais próximo das populações Ibero-Americanas do que a população de São Tomé e Príncipe, sugerindo uma maior influência genética caucasiana neste arquipélago.

Relativamente à análise de microssatélites do cromossoma Y da população de Cabo Verde apresenta-se próximo do arquipélago de São Tomé e Príncipe, não se tendo contabilizado nenhuma diferença estatisticamente significativa. Embora apresente alguma distância de Portugal e do Brasil (São Paulo e Amazónia), esta distância é, no geral, inferior à que se verifica entre Cabo Verde e as populações do continente Africano, nomeadamente Moçambique, a Guiné-Bissau e Angola, as quais se encontram já a uma distância significativa.

A linhagem paterna da população da Guiné-Bissau apresenta-se próxima das linhagens paternas das outras populações do continente Africano, Angola e Moçambique, embora com algumas diferenças estatisticamente significativas. Esta população apresenta-se algo distante de Cabo Verde e de São Tomé e Príncipe e apresenta-se com uma grande distância e diferenciação relativamente às populações do grupo Ibero-Americano, Portugal e Brasil.

Os achados de Corte-Real F (1999) são concordantes com os relatos históricos que descrevem a povoação de Cabo-Verde recorrendo a escravos Africanos oriundos da costa Ocidental de África e de outros territórios Africanos que estariam na altura sob domínio Português (Guiné-Bissau, Angola e Moçambique) e homens Portugueses e outros Europeus (Espanhóis e Italianos, por exemplo). Estes seriam essencialmente exilados e criminosos que eram enviados para as ilhas no atlântico com o objectivo de ajudar à sua povoação e exploradores que acabavam por ficar nas ilhas descobertas. Estes dados são apoiados por estudo posteriores em que foram analisados os haplogrupos de cromossoma Y de Cabo Verde e da Guiné-Bissau (Rosa A, *et al*, 2007; Gonçalves R, *et al*, 2003).

Os haplogrupos do cromossoma Y presentes em Cabo Verde apresentam proveniência da Euroásia, nomeadamente da Península Ibérica e da região mediterrânica, e de diferentes regiões de África, Oeste, Norte e também do Este. Estes dados confirmam uma origem mista da linhagem paterna caboverdiana, essencialmente de populações Ibéricas, mediterrâneas e da costa Oeste Africana como a Guiné-Bissau, com uma pequena influência de populações do Este e Norte de África e Médio Oriente, verificando-se especificamente influência do povo Judeu (Gonçalves R, *et al*, 2003).

À semelhança do indiciado pela análise do ADNmt, verifica-se uma diferença entre as frequências dos haplogrupos do cromossoma Y presentes nas ilhas do Sotavento e do Barlavento Caboverdiano, na frequência e no tipo de haplogrupos encontrados (Gonçalves R, *et al*, 2003; Brehm A, *et al*, 2002).

A República da Guiné-Bissau apresenta haplogrupos do cromossoma Y com uma diversidade de muito reduzida, uma característica comum a várias populações sub-Saharianas, e caracteristicamente africanos, embora tenham detectado uma pequena presença de haplogrupos do cromossoma Y tipicamente europeus (Rosa A, *et al*, 2007; Wood E T, *et al*, 2005; Destro-Bisol G, *et al*, 2004).

A diversidade étnica presente na Guiné-Bissau é justificada por migrações de povos Africanos para a região que terão ocorrido durante vários séculos e em alturas distintas, esta diversidade é evidente nas linhagens maternas e paternas desta população, existindo grupos étnicos que distam significativamente uns dos outros (análise de F_{st} , $P < 0,05$) no que respeita à análise das linhagens do cromossoma Y (haplogrupos) e análise das sequências das regiões hipervariáveis do ADNmt (Rosa A, *et al*, 2007; Wood E T, *et al*, 2005; Destro-Bisol G, *et al*, 2004).

3. COMPARAÇÃO COM MARCADORES AUTOSSÓMICOS

Como já foi referido, Corte Real F, em 1999, estudou vários marcadores autossómicos das populações de Cabo Verde e Guiné-Bissau, entre outras populações de expressão Portuguesa, e partiu-se dessa mesma amostragem para a realização do presente trabalho.

As amostras da Guiné-Bissau (N = 92) estudadas por Corte Real F (1999) para a determinação dos microssatélites autossómicos apresentam a seguinte proveniência:

- etnia Manjaco: 15 pessoas
- etnia Papel: 15 pessoas
- etnia Balanta: 11 pessoas
- etnia Fula: 9 pessoas
- etnia Mancanha: 8 pessoas
- etnia Beafada: 3 pessoas
- etnia Mandinga: 2 pessoas
- etnia Bijagós: 1 pessoa
- etnia Cassanga: 1 pessoa
- etnia Nalu: 1 pessoa
- etnias Manjaco/Papel: 4 pessoas
- etnias Beafada/Papel: 3 pessoas
- etnias Fula/Papel: 2 pessoas
- etnias Balanta/Mancanha: 2 pessoas
- etnias Balanta/Banhune: 1 pessoa
- etnias Balanta/Mandinga: 1 pessoa
- etnias Balanta/Papel: 1 pessoa
- etnias Beafada/Nalu: 1 pessoa
- etnias Felupe/Manjaco: 1 pessoa
- etnias Fula/Manjaco: 1 pessoa
- etnias Mandinga/Manjaco: 1 pessoa
- etnias Mandinga/Papel: 1 pessoa
- etnias Mansoanca/Papel: 1 pessoa
- etnias Nalu/Papel: 1 pessoa
- não sabe ou não responde: 5 pessoas.

As amostras de Cabo Verde ($N = 107$) estudadas por Corte Real F (1999) para a determinação dos microssatélites autossômicos apresentam a seguinte proveniência:

- Mestiços: 86 pessoas
- Negros: 19 pessoas
- Brancos: 2 pessoas.

A análise de marcadores autossômicos destas populações, da população de Angola (Norte), Moçambique (Sul), São Tomé e Príncipe, Portugal (Centro) e duas populações do Brasil (Amazônia e São Paulo) e a sua análise comparativa, baseada em diferenças estatisticamente significativas, análise de haplótipos comuns, distâncias de Nei e valores de F_{st} , permitiu a Corte-Real F (1999) agrupar estas populações em dois grupos: um grupo Ibero-Americano, e um grupo Africano, e não os três encontrados para a análise do cromossoma Y.

O grupo Ibero-Americano é composto pela população de Portugal e as duas populações do Brasil, Amazônia e São Paulo, cujas linhagens paternas se encontram filogeneticamente próximas, este grupo é caracterizado por apresentarem linhagens e haplótipos autossômicos caracteristicamente e maioritariamente caucasianos.

O grupo de populações Africanas é composto pela população da Guiné-Bissau, Angola (Norte), Moçambique (Sul), Cabo Verde e de São Tomé e Príncipe. Estas populações apresentam uma grande diferença e distância filogenética das populações Ibero-Americanas analisadas e apresentam resultados autossômicos caracteristicamente e maioritariamente africanos. Contudo, as populações Africanas insulares, Cabo Verde e de São Tomé e Príncipe, aparecem numa posição intermédia entre as populações Africanas Continentais e as populações Ibero-Americanas, sugerindo uma maior influência genética caucasiana.

A análise de marcadores autossômicos realizada por Corte-Real F (1999) revela que população de Cabo Verde não apresenta um número de diferenças estatisticamente significativas que indiquem distância filogenética significativa relativamente a São Tomé e Príncipe, Guiné-Bissau, Moçambique e Angola. Relativamente às populações de Portugal e Brasil, estas apresentam-se a uma distância significativa da população Caboverdiana.

Por seu lado, a população da Guiné-Bissau encontra-se próxima das populações africanas estudadas: Moçambique, Angola, Cabo Verde e de São Tomé e Príncipe; e distante da população de Portugal e das populações Brasileiras.

Embora esta análise pareça entrar em contradição com a análise de marcadores do cromossoma Y, a realidade é que vem apoiar algumas das conclusões da mesma e os relatos históricos que descrevem que a povoação de Cabo-Verde foi realizada recorrendo maioritariamente a escravos Africanos oriundos das colónias Portuguesas em África, nomeadamente da Guiné-Bissau, e a homens Portugueses e Europeus, essencialmente exilados, criminosos e exploradores, que se miscigenaram com mulheres oriundas de África originando uma nova 'raça', os 'mulatos' ou 'mestiços'. Os territórios do Continente Africano agora pertencentes à Guiné-Bissau, Angola e Moçambique não se encontravam despovoadas aquando da sua descoberta, não tendo ocorrido miscigenação entre de Portugueses e/ou Europeus e Africanos autóctones significativa.

Esta relativa proximidade de Cabo-Verde a populações Europeias, nomeadamente Portugal, e a diferenciação observada relativamente às populações do Oeste Africano, como a Guiné-Bissau, é também reportada em outros estudos sobre esta população. Esta particularidade será devida à influência Europeia, caucasiana, e a fenómenos de deriva genética, acumulação de diferenças no pool genético da população devidas essencialmente a drásticas reduções no número de habitantes deste arquipélago em diferentes períodos da sua história (Spínola H, *et al*, 2005; Fernandes AT, *et al*, 2003; Spínola H, *et al*, 2002; Parra EJ, *et al*, 1995).

Embora ainda muito haja para estudar relativamente a polimorfismos do ADN das populações Africanas, existem já vários estudos sobre Cabo Verde que versam sobre vários tipos de polimorfismos, além dos polimorfismos do cromossoma Y e das sequências das regiões hipervariáveis do ADN mt. Já foram estudados para esta população polimorfismos enzimáticos eritrocitários, polimorfismos de proteínas plasmáticas, de antígenos humanos leucocitários (HLA) e de "short tandem repeat's" (STR's) autossómicos. Os estudos autossómicos desta população relatam alguma distância entre Cabo Verde e as populações do Oeste Africano e colocam esta população numa posição intermédia entre a Europa e as populações do Oeste de África (Spínola

H, *et al*, 2005; Brehm A, *et al*, 2003; Fernandes AT, *et al*, 2003; Spínola H, *et al*, 2002; Parra EJ, *et al*, 1995).

A análise da população de Cabo Verde para polimorfismos HLA e STR's revela alguma diferenciação entre a população do Barlavento e do Sotavento Caboverdiano, como indiciado na análise das sequências das regiões hipervariáveis do ADN mt apresentada no presente trabalho. A população do grupo de ilhas do Sotavento apresenta um perfil genético típico de populações sub-Saharianas, enquanto que a população do grupo de ilhas do Barlavento apresenta uma maior influência caucasiana. Este facto pode ser devido a sua diferencial povoação, recorrendo a povos e/ou etnias diferentes. Adicionalmente, o Barlavento foi povoado significativamente depois das ilhas do Sotavento e em grande parte com escravos fugidos de outras ilhas (Spínola H, *et al*, 2005; Brehm A, *et al*, 2003; Fernandes AT, *et al*, 2003; Spínola H, *et al*, 2002).

Ao contrário do Cabo Verde, não existem muitos estudos sobre as populações do Oeste de África, como a república da Guiné-Bissau. Em 2005, Spínola H, *et al*, estudaram os polimorfismos de HLA desta população e de Cabo-Verde, tendo verificado a forte ligação da Guiné-Bissau à África sub-Sahariana, embora apresentando maiores semelhanças com populações do Norte de África do que outras populações do Oeste Africano, evidenciando uma influência de povos oriundos desta região, do Saara e do Este Africano, o estudo e análise das regiões hipervariáveis do ADN mt indiciam também a existência dessa influência (Spínola H, *et al*, 2005; Rosa A, *et al*, 2004).

Embora este país tenha estado sob o domínio Português durante aproximadamente 500 anos, não terá ocorrido miscigenação significativa entre portugueses e africanos, uma vez que geneticamente esta população é nitidamente africana, quer no que concerne a análise de polimorfismos autossômicos (HLA e STR's), quer polimorfismos de linhagem, cromossoma Y e sequências das regiões hipervariáveis do ADN mt, a presença de haplótipos tipicamente caucasianos, quando detectada, é extremamente diminuta (Spínola H, *et al*, 2005; Rosa A, *et al*, 2004; Spínola H, *et al*, 2002; Corte Real F, 1999).

4. CONCLUSÕES GERAIS

O presente trabalho permite retirar algumas conclusões de ordem forense e filogenética sobre estudos populacionais em geral, as populações estudadas, as populações consideradas para o estudo comparativo e a evolução e migrações do homem em geral:

- ✓ A análise das duas regiões hipervariáveis do ADN mt em conjunto, HVI + HVII, reflecte-se num aumento da diversidade genética de sequência de uma população (J), devido ao facto de as sequências do ADN mt de duas amostras distintas pode ser comum para a região HVI e ser distinta para a região HVII. Logo, ao considerarmos as duas regiões hipervariáveis I e II aumentamos o número de sequências diferentes dentro de determinado grupo de amostras (população).

- ✓ Verifica-se sempre uma diminuição da diversidade nucleotídica (π) de uma população quando consideramos a conjugação das sequências nucleotídicas das regiões HVI e HVII. Esta diminuição deve-se ao facto de a região HVII possuir um número mais restrito de posições que apresentam 'mutação', apresentam uma base diferente da reportada por Anderson S, *et al* (1981), e outras que praticamente nunca se apresentam nesse estado de 'mutação', ao contrário da HVI onde as mutações ocorrem de forma mais homogênea ao longo das diferentes posições nucleotídicas (Pereira L, *et al*, 2000).

- ✓ A criação de bases de dados populacionais de sequências das duas regiões hipervariáveis do ADN mt, HVI e HVII, é de extrema relevância a nível forense, uma vez que aumenta o poder discriminatório destas bases de dados, permitindo que numa investigação forense, criminalista ou de identificação humana, seja efectuada

uma análise e um cálculo da frequência da sequência na população mais correcto e fiável, desde que a população considerada para esta análise seja a mais adequada.

- ✓ A análise filogenética populacional realizada com base nas duas regiões hipervariáveis do ADN mt, HVI e HVII, não se revelou muito diferente da análise realizada apenas com base na primeira região hipervariável, HVI, embora a incorporação dos dados da região HVII na análise populacional comparativa tenha, por vezes, levando a que populações se apresentem mais próximas ou distantes uma da outra.

- ✓ A análise da região HVII isoladamente possui pouca informação quando comparada com a análise baseada na região HVI, quer a nível forense, quer a nível filogenético, uma vez que esta região possui uma baixa diversidade de sequência (J) e nucleotídica (π). Adicionalmente, esta região apresenta posições específicas que estão classificadas como posições de mutação rápida (*'fast mutation'*), as quais rapidamente apresentam uma base diferente da reportada por Anderson S *et al*, em 1981, ou revertem para a base considerada como referência. Qualquer mutação encontrada nestas posições têm de ser considerada com o devido cuidado para não ser sobrevalorizada numa análise forense ou estudo filogenético. Contudo, as variações encontradas na zona dos poli-C desta região, na forma de inserções ou deleções que se podem traduzir em heteroplasmias de comprimento, possuem alguma relevância filogenética e forense (Pereira L, *et al*, 2000; Salas A, *et al*, 2000; Baasner A, *et al*, 1998).

- ✓ As bases de dados, de sequências das regiões hipervariáveis I e II do ADNmt das populações de Cabo Verde e da Guiné-Bissau, resultantes do presente trabalho, foram obtidas a partir de uma amostragem independente, a nível de linhagem, com número significativo e representativa das populações estudadas, apresentam valores de diversidade, de sequência (J) e nucleotídica (π), significativas, o que lhes confere utilidade e relevância a nível forense e que permitiu retirar algumas conclusões sobre a origem e evolução destas populações.

- ✓ A análise comparativa global, realizada com recurso a populações de todos os continentes, apoia a teoria do surgimento do homem em África, '*out of África*', e a posterior povoação do planeta por movimento migratórios, por vezes bidireccionais. De facto, as populações Africanas apresentam valores de diversidade de sequência (J) e nucleotídica (π), regra geral, superiores às populações dos restantes Continentes, sendo as populações da Oceânia as com valores mais baixos para estes parâmetros, revelando uma evolução temporal menor.

- ✓ A população de Cabo Verde apresenta uma linhagem materna, dada pela análise das sequências do ADNmt, maioritariamente e tipicamente Africana. Estes dados, confirmam que este arquipélago foi povoado com recurso a escravos provenientes, na sua grande maioria, da costa Ocidental Africana, nomeadamente da região entre o Senegal e Serra Leoa. É detectada alguma influência de populações do Este de África, ligeiramente superior à detectada em populações do Oeste Africano, indicando que terá ocorrido algum fluxo directamente para este arquipélago, provavelmente devido ao tráfico de escravos. A presença de sequências do ADNmt características do Norte de África e Sul da Europa (Península Ibérica e Mediterrânica) revela algum fluxo populacional para estas ilhas, em parte durante a sua povoação, mas também oriunda de migrações muito anteriores de povos do Norte para o Oeste de África.

A análise do ADNmt de Cabo Verde, linhagem materna, e a sua comparação com estudos de polimorfismos do cromossoma Y, linhagem paterna, revela uma linhagem materna essencialmente Africana em contraste com uma linhagem paterna com grande influência Europeia. Tal facto está de acordo com os relatos históricos que afirmam que os homens europeus que colonizaram estas ilhas se terão cruzado com mulheres negras, oriundas da costa Ocidental Africana, originando um novo grupo, os 'mestiços'.

- ✓ A linhagem materna da Guiné-Bissau é característica das populações do Oeste Africano. Identifica-se a presença de algumas sequências de ADNmt características do Centro e Este Africano, derivada de migrações de povos Africanos e de esta

região ter sido um entreposto de tráfico de escravos na época dos descobrimentos. A influência europeia detectada na população da Guiné-Bissau é explicada pela ocorrência de pequenas migrações de homens e mulheres, essencialmente da Península Ibérica e da região do mediterrâneo, aquando dos descobrimentos. O mesmo sucede com a presença de sequências mitocondriais oriundas de povos do Norte de África, geograficamente próximos (Rosa A, *et al*, 2007).

- ✓ A partilha de sequências entre as populações estudadas e outros Países Africanos de Língua Portuguesa, com Portugal e outras Populações Europeias e do Norte de África, e a sua posição pouco distante na representação em formato de árvore filogenética, pode também ser explicada pela existência de descendentes de mulheres oriundas da África Sub-Sahariana nestas populações, e não só pelo fluxo de mulheres Europeias para as colónias Africanas, uma vez que são detectadas linhagens mitocondriais tipicamente Sub-Saharianas nestas populações (Pereira L, *et al*, 2000). Esta hipótese é apoiada por alguns relatos históricos que descrevem pequenos fluxos de mulheres e homens das colónias Africanas para o sul da Europa e ilhas (Sul de Portugal, Arquipélago dos Açores e Madeira), com o objectivo de ajudar a povoar estas regiões. A proximidade geográfica entre o Sul da Europa e o Norte de África permite fluxos populacionais frequentes que levam a miscigenação dos povos destas regiões.

- ✓ A relativa proximidade e partilha de sequências das regiões hipervariáveis do ADNmt entre a Guiné-Bissau e populações Africanas insulares ao largo da costa Oeste Africana, algumas populações do Continente Americano e ilhas atlânticas Europeias, como a Madeira e as Canárias, comprovam que a região conhecida hoje como Guiné-Bissau terá sido uma fonte importante de escravos que terão sido levados para estes locais para ajudar na sua colonização e / ou povoação e como mão-de-obra.

- ✓ A análise intra-populacional de Cabo-Verde revela alguma diferenciação entre a população das ilhas do Sotavento e do Barlavento do arquipélago, embora a amostragem por ilha do presente estudo seja baixa, excepto para a Ilha de Santiago; este facto já havia sido detectado em outros estudos (Brehm A, *et al*,

2003; Fernandes AT, *et al*, 2003; Spínola H, *et al*, 2002). Verificando-se uma maior proximidade das ilhas do Sotavento ao Oeste Africano, com alguma presença de haplogrupos Europeus e Asiáticos, nomeadamente na ilha de Santiago.

A análise do Barlavento apresenta uma menor diversidade: as 8 amostras provenientes destas ilhas pertencem exclusivamente aos haplogrupos L1 e L2. Esta baixa diversidade pode ser devida à baixa amostragem, ou pela povoação destas ilhas ter sido comum e pouco diversificada ou com recurso a poucas pessoas. Os períodos de fome sentidos em Cabo Verde que levaram a uma grande redução da sua população podem ainda estar na origem desta baixa diversidade. Embora a amostragem das ilhas do Barlavento Caboverdiano seja diminuta, alguns dos achados são apoiados por outros estudos, que referem ainda uma maior aproximação das ilhas do Barlavento a populações caucasianas.

A diferenciação entre Sotavento e Barlavento pode dever-se ao facto de, uma vez que estas ilhas foram povoadas com um grande desfasamento temporal, a sua povoação ter sido realizada recorrendo a etnias e / ou povos diferentes. Pode ainda ser devido a fenómenos de deriva genética, ocorrendo um distanciamento das ilhas do Sotavento e do Barlavento por acumulação de diferenças no seu pool genético. Também aqui os períodos prolongados de fome que Cabo Verde sofreu, referidos em documentos históricos, podem ter contribuído para esta diferenciação.

- ✓ A população da Guiné-Bissau apresenta uma significativa diversidade étnica, presente nas amostras aqui analisadas, embora havendo etnias que não estão aqui representadas ou cuja representação é muito diminuta. No geral as etnias analisadas no presente trabalho não apresentam grande diferenciação entre elas. A análise filogenética com base na matriz de distâncias *Fst* e na significância dos seus valores de P, para $P = 0,05$, apresenta diferenciação entre a etnia de Beafada e Papel para as regiões HVI e HVII isoladas, e na análise conjunta destas regiões, HVI+HVII, ambas estas etnias apresentam ainda distância significativa da etnia Manjaco, para região HVI e HVI+HVII. Adicionalmente, só nas etnias Papel e Balanta são encontradas amostras cujas sequências correspondem a haplogrupos Europeus.

- ✓ Da comparação entre Cabo Verde e Guiné-Bissau, que revela uma grande proximidade entre estas populações, conseguimos detectar alguma distância significativa (considerando $P = 0,05$) entre etnias Guineenses e algumas ilhas ou grupos de ilhas de Cabo Verde. A etnia Manjaco apresenta-se significativamente distante de todas as ilhas Caboverdianas representadas no presente estudo, com excepção do grupo das ilhas Brava e Maio. Considerando apenas a região HVI a amostragem da ilha do Fogo apresenta-se significativamente distante das etnias Beafada e Manjaco, sendo que a etnia Beafada se apresenta distante do grupo composto pelas ilhas do Sotavento. O grupo composto pelas ilhas do Barlavento apresentam uma distância significativa de todas as etnias Guineenses analisadas, com excepção das etnias Beafada e Fula. De facto os únicos haplogrupos presentes na amostragem do Barlavento Caboverdiano, os haplogrupos Sub-Saharianos L1 e L2, são comuns às etnias Beafada e Fula da Guiné-Bissau. Estas particularidades podem ser reflexo de uma colonização das ilhas de Cabo Verde etnicamente diferenciada e da redução da população devido aos subsequentes períodos de fome, que podem ter reduzido ou mesmo eliminado a presença de algumas etnias de determinadas ilhas, levando à sua diferenciação.

5. LIMITAÇÕES DO ESTUDO E PROJECTOS FUTUROS.

Como refere Corte-Real F, 1999, em genética forense e genética populacional um estudo de investigação nunca se conclui, é sempre possível completá-lo e melhorá-lo. No que toca ao presente estudo, fica a vontade de uma análise mais aprofundada das populações de Cabo Verde e Guiné-Bissau a vários níveis, desde a amostragem à utilização de novas técnicas de análise do ADN mt.

Um maior número de amostras de cada população alargaria a base de dados e permitiria uma análise mais correcta das suas linhagens maternas. Em 2004 Pereira L, *et al*, tentaram determinar o número de amostras para o qual uma base de dados populacional de sequências do ADN mt atinge o seu ponto de saturação. Para tal simularam análises das duas regiões do ADN mt, HVI e HVII, de vários subgrupos de uma população Portuguesa que compreendia um total de 549 amostras. Estes subgrupos possuíam um número de amostras total de 50, 100, 200, 300 e 400 amostras, misturadas ao acaso. Com base nesta análise, calcularam os parâmetros de diversidade para os vários subgrupos e verificaram que embora não houvesse alterações significativas na diversidade nucleotídica (π) e de sequência (**J**), o número de sequências diferentes (**K**) e de posições nucleotídicas com polimorfismo (**A**) era continuamente crescente. Por extrapolação calcularam o ponto de saturação de uma base de dados das regiões HVI e HVII do ADN mt, o ponto de saturação corresponde ao número de amostras numa base de dados a partir do qual o aumento de 100 amostras reflecte-se num aumento de diversidade populacional inferior a 5%. Para a região HVI determinaram que o ponto de saturação de uma população seria de 1000 amostras, para a HVII seria de 900 e para o conjunto das duas regiões hipervariáveis a saturação da base de dados seria atingida para as 1300 amostras (Pereira L, *et al*, 2004).

Além do aumento do número de amostras da população na sua globalidade, será também relevante, tentar que a amostragem recolhida e analisada seja estatisticamente

representativa da população, ou seja, que o número de amostras consideradas para a análise global de uma população seja proporcional à sua composição étnica e distribuição geográfica. Para a população de Cabo Verde seria desejável que a amostragem estudada espelhasse a distribuição geográfica da população pelas diferentes ilhas do arquipélago, tendo em consideração a frequência das diferentes etnias em cada ilha. Para a população Guineense uma análise semelhante que espelhasse a sua diversidade étnica faz todo o sentido, uma vez que este país é rico em etnias que se diferenciam a vários níveis, não só genéticos mas também culturais, as quais estão presentes em proporções muito diferentes e possuem uma distribuição territorial específica e muito distinta.

No presente trabalho foi realizada uma análise intra-populacional que evidencia as diferenças existentes entre algumas etnias Guineenses e entre algumas ilhas caboverdianas. Um estudo individualizado e aprofundado das diferentes etnias da Guiné-Bissau e das populações de cada ilha de Cabo Verde é relevante a nível forense e também a nível antropológico, uma vez que as diferenças, nalguns casos bastante significativas, indiciam uma origem e evolução distinta de cada grupo étnico e uma colonização e evolução também distinta das ilhas de Cabo Verde. Este tipo de estudos permite obter bases de dados de sub-populações, que podem ser mais adequadas como população de comparação numa investigação forense.

Actualmente os estudos forenses e populacionais com base no ADN mt não se resumem à análise das suas regiões hipervariáveis. A sequenciação e análise de toda a região controlo do ADN mt constitui uma ferramenta mais poderosa, aumentando o poder discriminatório do ADN mt em investigações forenses. Esta análise baseia-se na amplificação e sequenciação de toda a 'D-loop', desde a posição 16024 até à posição 576 (Sala A, et al, 2009; Nakamura S, et al, 2008).

Contudo, ainda existem poucas bases de dados das sequências completas do ADN mt, logo esta análise completa da região controlo do ADN mt ainda é um pouco limitativa, quer a nível forense, quer filogenético, pois não existem muitas outras populações para as quais já se possua esta informação e com as quais se possam fazer estudos comparativos.

Pela mesma razão, é bastante pertinente a criação destas bases de dados populacionais que permitirão uma análise mais profunda das populações em questão e

aumentar significativamente o poder discriminatório do ADN mt numa investigação forense, permitindo o cálculo mais correcto da frequência de determinada sequência do ADN mt na população considerada como origem dessa sequência.

Outro tipo estudo de interesse é a análise de 'Single Nucleotide Polimorfims', SNP's, do ADN mt, aumentando a informação populacional com o objectivo final de termos um maior poder discriminatório, o que permitiria conclusões mais assertivas numa investigação forense. Esta análise baseia-se na amplificação por PCR de pequenas regiões do ADN mt e posterior detecção de polimorfismos pontuais existentes ao longo de toda a molécula de ADN mt. A amplificação pode ser realizada em multiplex, a detecção dos polimorfismos pode ser realizada por electroforese em gel ou recorrendo a uma mini-sequenciação, SNaPshot, dependendo do protocolo de amplificação (Bento A M, *et al*, 2009a; Mikkelsen M, *et al*, 2008; Medeiros S, *et al*, 2008).

Esta análise é muito útil na determinação do haplogrupo de ADN mt de determinada amostra, uma vez que a informação das duas regiões hipervariáveis, HVI e HVII, nem sempre é suficiente (Bento A M, *et al*, 2009a; Mikkelsen M, *et al*, 2008; Medeiros S, *et al*, 2008).

Adicionalmente, esta análise permite avaliar posições nucleotídicas dispersas pela molécula do ADNmt, nomeadamente, fora da região controlo. Logo, aumenta o poder discriminatório da análise do ADN mt. Na tipagem de amostras degradadas a análise de SNPs do ADN mt também se revela bastante vantajosa, uma vez que não é necessário que o ADN esteja íntegro, pois os fragmentos amplificados para esta análise são bastante pequenos (Bento A M, *et al*, 2009b; Sun X, *et al*, 2009; Dario P, *et al*, 2009; Sturk K A, *et al*, 2008; Lima G & Pinheiro M F, 2008; Álvarez-Iglesias V, *et al*, 2008)

Este método pode ser visto como um método de rastreio das amostras para os polimorfismos do ADN mt, no âmbito de uma investigação forense. É uma metodologia rápida, fácil e económica, que pode reduzir significativamente os custos e o tempo associados à análise do ADN mt. Posteriormente, se necessário, pode-se recorrer à sequenciação da região controlo completa ou das regiões hipervariáveis I e II para confirmação e complemento dos dados obtidos pela análise de SNPs (Dario P, *et al*, 2009).

Em resumo, dentro do muito que ainda há a estudar relativamente a Cabo Verde e Guiné-Bissau, as acções descritas em baixo serão as mais pertinentes:

- ✓ Aumento do número de amostras de ambas as populações em concordância com a composição étnica e geográfica das mesmas;
- ✓ Estudo individual de uma amostragem significativa de cada etnia;
- ✓ Estudo aprofundado por ilha do arquipélago de Cabo-Verde;
- ✓ Análise da sequência completa de toda a região controlo do ADN mt;
- ✓ Análise de 'Single Nucleotide Polimorfims', SNP's, do ADN mt.