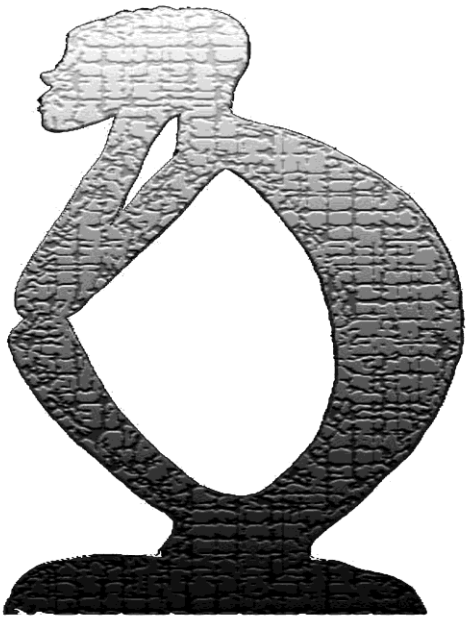
U.PORTO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

DIVERSIDADE GENÉTICA NOS PRINCIPAIS GRUPOS POPULACIONAIS EM ANGOLA - *APLICAÇÃO FORENSE*



MIGUEL MANUEL MELO

Tese de doutoramento em Ciências Médicas

2010

MIGUEL MANUEL MELO

DIVERSIDADE GENÉTICA NOS PRINCIPAIS GRUPOS POPULACIONAIS EM ANGOLA - *APLICAÇÃO FORENSE*

Tese de Candidatura ao grau de Doutor em Ciências Médicas submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador Prof. Doutor Francisco Côrte-Real Gonçalves

Afiliação – Instituto Nacional de Medicina Legal I.P. e Faculdade de Medicina da Universidade de Coimbra.

Co-orientador – Prof. Doutor Jorge Sequeiros

Afiliação – IBMC e Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Agradecimentos

Ao **Prof. Doutor Duarte Nuno P. Vieira,** Presidente do Instituto Nacional de Medicina Legal, pela amizade, confiança depositada, oportunidade concedida e pela vontade expressa em apadrinhar Angola na área da Medicina Legal.

Ao **Prof. Doutor Francisco Côrte-Real Gonçalves,** Director da Delegação do Centro do Instituto Nacional de Medicina Legal e orientador deste trabalho pelos brilhantes ensinamentos, amizade, dedicação e estímulo constante à busca do conhecimento, tão fundamental para o desenvolvimento e conclusão deste trabalho; pela vontade demonstrada em ajudar Angola na organização do Serviço de Genética Forense.

Ao **Prof. Doutor Jorge Sequeiros** co-orientador deste trabalho pela amizade, apoio na tomada de decisões, na orientação científica e pela sugestão do tema.

À **Prof. Doutora Maria de Fátima Pinheiro** pela amizade, apoio e sugestões com relação ao trabalho.

À **Prof. Doutora Helena Geada** pela amizade, apoio com relação ao trabalho.

Ao **Prof. Doutor Nuno Grande** pela amizade, oportunidade e esperança no desenvolvimento de Angola em diversas áreas do saber.

À **Prof. Doutora Maria Helena de Fátima Figueiredo** pela amizade, dedicação e atenção durante a minha estadia em Portugal.

À **Drª Virgínia Lopes** pela amizade, disponibilidade, pela ajuda no tratamento e ensinamentos dos dados estatísticos deste trabalho.

À **Dr.ª Mónica Carvalho** pela sua preciosa colaboração, inesgotável paciência e apoio científico, fundamentais para a tomada de algumas decisões.

À Mestre Heloísa Costa, Mestre Filipa Balsa, Mestre Lisa Sampaio, Dr Armando Serra, Dr Pedro Brito pela amizade e auxílio na rotina laboratorial.

À Drª M*aria* João na qualidade de Directora do Serviço de Genética e Biologia Forense da Delegação do Centro do INML pelo apoio concedido.

A todos outros funcionários dos Serviços de Genética e Biologia Forense e de outros Serviços da Delegação do Centro pela amizade e camaradagem ao longo destes anos de formação.

Ao **Governo Angolano, Ministério do Interior** assim como ao Departamento dos Recursos Humanos do **Comando Geral da Polícia Nacional de Angola** pela oportunidade de formação.

Ao Digníssimo **Comissário "Geral" Ambrósio de Lemos Freire dos Santos** Comandante Geral da Polícia Nacional de Angola pela amizade e carinho, incentivo, oportunidade de formação e pela vontade expressa na criação de um laboratório de Genética Forense.

A minha querida esposa **Edite Neves Melo** pelo amor, compreensão, dedicação e aos filhos **Wélvia**, **Túlio** e **Nickson Miguel Melo** pela minha ausência.

A **Susana Naleca, Mohamed, Celina,** pelo amor, carinho, incentivo, pela minha ausência e apoio na recolha de algumas das amostras e o transporte destas até Luanda.

Ao Mestre Azi Sajo Manuel pela camaradagem.

Aos meus pais e irmãos pelo carinho, amor e apoio apesar das grandes dificuldades e ….

A Deus……...eternamente grato.

Índice

Índice das figuras IX
Índice das tabelas XI
Abreviaturas das siglas XV
Citações XVII

1 – Justificação do tema e objectivos 18
1.1 – Justificação do tema 19
1.2 – Objectivos gerais e específicos 21
2 – Introdução 22
2.1 – Origem do homem moderno 23
2.2 – História e caracterização da população Angolana 27
2.2.1 – Primeiros habitantes de Angola 28
2.2.2 – Principais grupos étnico-linguísticos actuais em Angola 31
2.3 – Genoma humano 37
2.3.1– Genes e cromossomas 37
2.3.2 – Características dos cromossomas autossómicos 41
2.3.3 – Características do cromossoma Y 42
2.3.4 – Características do ADN mitocondrial 45
2.4 – Polimorfismos de ADN 48
2.4.1 – Polimorfismos de STRs autossómicos 50
2.4.2 – Polimorfismos de STRs do cromossoma Y 53
2.4.3 – Polimorfismos do ADN mitocondrial 57
2.4.4 – Aplicação forense dos polimorfismos do ADN 60
2.4.5– Marcadores de linhagens em genética populacional 61
3 – Material e Métodos 64
3.1 – Amostragem populacional 65
3.2 – Colheitas de amostras sanguíneas 65
3.3 – Extracção do ADN 66
3.4 – Técnicas utilizadas no estudo do ADN autossómico 66
3.4.1 – Amplificação de ADN autossómico 66
3.4.2 – Detecção dos polimorfismos de STRs autossómicos 68
3.4.3 – Designação alélica dos polimorfismos de STRs autossómicos 70
3.5 – Técnicas utilizadas no estudo de STRs do cromossoma Y 71
3.5.1 – Amplificação de STRs do cromossoma Y 71
3.5.2 – Detecção dos polimorfismos de STRs do cromossoma Y 72

3.5.3 – Designação alélica dos polimorfismos de STRs do cromossoma Y....73
3.6 – Técnicas utilizadas no estudo do ADN mitocondrial....73
3.6.1 – Amplificação da região controlo....73
3.6.2 – Purificação dos fragmentos amplificados por ExoSap-IT....74
3.6.3 – Sequenciação cíclica (directos/reversos) dos produtos amplificados....74
3.6.4 – Purificação dos fragmentos sequenciados por X Terminator....74
3.6.5 – Detecção do produto sequenciado....74
3.7 – Análise estatística....75
4 – Resultados....76
4.1– Polimorfismos de STRs autossómicos....77
4.1.1 – Polimorfismos de STRs autossómicos da população angolana....77
4.1.2 – Comparação dos polimorfismos de STRs autossómicos da população angolana com outras populações....78
4.1.3 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Bakongo....81
4.1.4 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Kimbundo....83
4.1.5 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Kwanhama....85
4.1.6 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Lunda-Tchokwe....87
4.1.7 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Nganguela....89
4.1.8 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Nhaneca-Humbe....91
4.1.9 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Ovimbundo....93
4.1.10 – Comparação dos polimorfismos de STRs autossómicos entre os sete grupos étnico-linguísticos....94
4.2 – Polimorfismos de STRs do cromossoma Y....96
4.2.1 – Polimorfismos de STRs do cromossoma Y na população angolana....96
4.2.2 – Comparações genético-populacionais de STRs do cromossoma Y....102

4.2.3 – Polimorfismos de STRs do cromossoma Y no grupo étnico-linguístico Bakongo 103
4.2.4 – Polimorfismos de STRs do cromossoma Y no grupo étnico-linguístico Kimbundo 108
4.2.5 – Polimorfismos de STRs do cromossoma Y no grupo étnico-linguístico Ovimbundo 113
4.2.6 – Análise comparativa dos polimorfismos do cromossoma Y entre os três grupos étnico-linguísticos 117
4.3 – Polimorfismos do ADN mitocondrial 119
4.3.1 – Polimorfismos do ADN mitocondrial na população de Angola 120
4.3.2 – Polimorfismos do ADNmt do grupo étnico-linguístico Bakongo 121
4.3.3 – Polimorfismos do ADNmt do grupo étnico-linguístico Kimbundo 121
4.3.4 – Polimorfismos do ADNmt do grupo étnico-linguístico Ovimbundo 122
4.3.5 – Análise comparativa dos polimorfismos de ADN mitocondrial entre três grupos étnico-linguísticos 122
5 – Discussão 124
5.1 – Microssatélites autossómicos 125
5.1.2 – Equilíbrio de Hardy-Weinberg nos grupos étnico-linguísticos estudados 127
5.1.3 – Parâmetros de eficácia *a .priori* 128
5.1.4 – Comparações genético-populacionais entre os grupos étnico-linguisticos 130
5.2 – Microssatélites do cromossoma Y na população de Angola e análise comparativa com outras populações africanas 131
5.3 – Análise comparativa dos polimorfismos do cromossoma Y entre os três grupos étnico-linguísticos 133
5.4 – Análise comparativa dos polimorfismos do ADNmt entre os três principais grupos étnico-linguísticos 137
6 – Conclusões e perspectivas de estudos futuros 140
7 – Resumo 144
8 – Summary 147
9 – Referências bibliográficas 150
10 – Anexos 165
Anexo I – Ficha preenchida no acto da colheita das amostras 166
Anexo II – Protocolo de extracção do ADN 167

Anexo III – Procedimentos de amplificação do ADN pelo método da PCR, de STRs autossómicos e do cromossoma Y 167
Anexo IV – Detecção do produto amplificado 169
Anexo V – Procedimentos de amplificação da região controlo 170
Anexo VI – Procedimentos de purificação dos fragmentos amplificados por ExoSap-IT 170
Anexo VII – Sequenciaçao cíclica (direita/reversa) dos produtos amplificados 171
Anexo VIII – Procedimentos de purificação dos fragmentos sequenciados por X Terminator /SAM solution 171
Anexo IX – Artigo, poster e comunicações escrita 172

Índice das figuras

Figura 2.1 – Origem do homem moderno e processos migratórios em África…… 25
Figura 2.2 – Mapa da divisão administrativa de Angola…….. 27
Figura 2.3 – Processos migratórios Bantu……. 29
Figura 2.4 – Caçador Bosquímano no sul de Angola……..... 30
Figura 2.5 – Mapa étnico-linguístico de Angola ……... 30
Figura 2.6 – Povos Bantu de Angola……..... 31
Figura 2.7 – Representação da relação entre a dupla hélice de ADN e os cromossomas e sua localização no interior das células…….. 37
Figura 2.8 – Classificação das sequências de ADN repetitivo no genoma humano (Fowler *et al.* 1988)……. 39
Figura 2.9 – Ideograma do cromossoma Y (adaptado de Quintana-Murci e Fellous, 2001; Skaletsky et al., 2003)……..... 43
Figura 2.10 – Organização do ADNmt humano (adaptado de Butler, 2005)…… 46
Figura 2.11 – Polimorfismos de STRs autossómicos…….. 51
Figura 2.12 – Esquema da distribuição dos marcadores do STR ao longo do cromossoma Y……. 54
Figura 2.13 – Divisão da região controlo mitocondrial……. 57
Figura 3.1 – Representação esquemática dos *loci STRs* amplificados no sistema AmpFℓSTR® Identifiler ™ na posição vertical a separação por cores e na posição horizontal a separação por tamanho em pb (adaptado de Ruitberg, 2001)…… 69
Figura 3.2 – Representação esquemática dos *loci* do cromossoma Y amplificados no sistema PowerPlex…… 72
Figura 4.1 – Árvore filogenética da comparação entre 13 diferentes populações, obtida com base nas distâncias genéticas de Nei a partir do método Neighbor-Joining e visualizada no programa Tree View…….. 80
Figura 4.2 – Árvore filogenética dos diferentes grupos étnico-linguísticos de Angola, obtida com base nas distâncias genéticas de Nei a partir do método Neighbor-Joining e visualizada no programa Tree View……. 95
Figura 4.3 – Distribuição haplotipica do *locus* DYS385 na população angolana (N=166)…… 96
Figura 4.4 – Distribuição dos alelos de cada um dos *loci* de STRs do cromossoma Y no grupo étnico-linguístico Bakongo…… 103

Figura 4.5 – Distribuição dos haplótipos do DYS385 no grupo étnico-linguístico Bakongo……. 104
Figura 4.6 – Distribuição dos alelos de cada um dos *loci* de STRs do cromossoma Y no grupo étnico-linguístico Kimbundo……. 108
Figura 4.7 – Distribuição dos haplótipos do DYS385 no grupo étnico-linguístico Kimbundo…… 109
Figura 4.8 – Distribuição dos alelos de cada um dos *loci* de STRs do cromossoma Y no grupo étnico-linguístico Ovimbundo…….. 113
Figura 4.9 – Distribuição dos haplótipos DYS385 no grupo étnico-linguístico Ovimbundo……... 114
Figura 4.10 – Árvore filogenética relativa aos três principais grupos étnico-linguísticos de Angola…… 118
Figura 4.11 – Árvore filogenética do ADNmt nos três principais grupos étnico-linguísticos de Angola…….. 123

Índice das tabelas

Tabela 2.1 – Comparação entre o ADN mitocondrial e o ADN nuclear……..……. 47
Tabela 2.2 – Comparação entre os sistemas *multiplex* utilizados em identificação humana……. 52
Tabela 3.1 – Caracterização dos *loci* de STRs presentes no sistema *AmpFLSTR® Identifiler®*…… 68
Tabela 3.2 – Principais características dos *loci* analisados pelo sistema *AmpFLSTR® Identifiler®*……. 70
Tabela 3.3 – Principais características dos *loci* de STRs do cromossoma Y analisados……..... 71
Tabela 4.1 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para a população global de Angola……... 77
Tabela 4.2 – Matriz de distâncias relativa a diferentes populações mundiais, obtida a partir do programa Arlequim 2.000, pelo cálculo de FST……..... 79
Tabela 4.3 – Comparação entre população de Angola e outras populações através da diferenciação do valor de P testada para cada *locus* ……... 79
Tabela 4.4 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de o equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori* , para o grupo étnico-linguístico Bakongo…… 81
Tabela 4.5 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Kimbundo……..... 83
Tabela 4.6 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e,

parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Kwanhama……..... 85

Tabela 4.7 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Lunda-Tchokwe…… 87

Tabela 4.8 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Nganguela…… 89

Tabela 4.9 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori* , para o grupo étnico-linguístico Nhaneca-Humbe…… 91

Tabela 4.10 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Nhaneca-Humbe…….. 93

Tabela 4.11 – Matriz de distâncias (FST) e valores de P relativa aos diferentes grupos étnico-linguísticos de Angola……... 94

Tabela 4.12 – Número de alelos e alelo mais frequente, nos marcadores estudados na população angolana……..... 96

Tabela 4.13 – Frequência alélica, diversidade génica e haplotípica dos 11 *loci* do cromossoma Y na população angolana…….. 97

Tabela 4.14 – Haplótipos do cromossoma Y na população Angolana (N=166) ……. 98

Tabela 4.15 – Matriz de distâncias génicas e respectivos valores de P entre populações Angola e outras populações Africanas…… 102

Tabela 4.16 – Análise da variância molecular entre a população Angolana e outras populações Africanas…….. 102
Tabela 4.17 – Frequência alélica, diversidade génica e haplotípica dos *loci* do cromossoma Y no grupo Bakongo (N=57)…...…… 105
Tabela 4.18 – Haplótipos de STRs do cromossoma Y do grupo étnico-linguístico Bakongo……..... 106
Tabela 4.19 – Frequência alélica, diversidade génica e haplotipica dos *loci do* cromossoma Y no grupo Kimbundo (N=56)..…….. 110
Tabela 4.20 – Haplótipos de STRs do cromossoma Y do grupo étnico-linguístico Kimbundo……... 111
Tabela 4.21 – Frequência alélica, diversidade génica e haplotipica dos *loci* do cromossoma Y no grupo Ovimbundo (N=53)..……..... 115
Tabela 4.22 – Haplótipos de STRs do cromossoma Y do grupo étnico-linguístico Ovimbundo……..….. 116
Tabela 4.23 – Matriz de distância génicas e respectivos valores de P relativa aos três principais grupos étnico-linguísticos de Angola……... 117
Tabela 4.24 – Análise da variância molecular nos três grupos étnico-linguísticos de Angola…… 118
Tabela 4.25 – Sequência de ADNmt, haplogrupo estimado e haplótipos da população angolana (N=30)……. 120
Tabela 4.26 – Sequência de ADNmt, haplogrupo estimado e haplótipos no grupo étnico-linguístico Bakongo (N=10)..……..… 121
Tabela 4.27 – Sequência de ADNmt, haplogrupo estimado e haplótipos no grupo étnico-linguístico Kimbundo (N=10)..……..... 121
Tabela 4.28 – Sequência de ADNmt, haplogrupo estimado e haplótipos no grupo étnico-linguístico Ovimbundo (N=10)...……. 122
Tabela 4.29 – Parâmetros de diversidade de ADNmt para os três grupos étnico-linguístico……... 122
Tabela 4.30 – Matriz de distância e respectivos valores de P para os três principais grupos de Angola……..…… 123
Tabela 5.1 – Parâmetros estatísticos de eficácia *a priori* e número de alelos por *locus*, para a população global angolana……. 126
Tabela 5.2 – Heterozigosidade observada nos STRs autossómicos nos grupos étnico-linguísticos……. 128
Tabela 5.3 – Poder de discriminação dos STRs autossómicos nos grupos étnico-linguísticos……. 129

Tabela 5.4 – Poder de exclusão *a priori* dos STRs autossómicos nos grupos étnico-linguísticos……. 129
Tabela 5.5 – Caracterização genética populacional por *locus* entre os três grupos étnico-linguísticos……..... 133
Tabela 5.6 – Valores da diversidade genica dos microssatélites do cromossoma Y nos três grupos……. 135
Tabela 5.7 – Diversidade de parâmetros dos11 STRs nos três grupos étnico-linguísticos de Angola…….. 136
Tabela 5.8 – Número e tipo de alterações polimórficas da região controlo total do ADNmt para os três grupos étnico-linguístico de Angola…….. 138

Abreviaturas das siglas

A – Adenina
ABI – Applied Biosystems, Inc
ADN – Ácido Desoxirribonucleico
ADNmt – Ácido Desoxirribonucleico mitocondrial
ARNs – Ácido Ribonucleico mensageiro
ATF – Adenosina trifosfato
C – Citosina
CODIS – *Combined DNA Index System*
CRS – Cambridge Reference Sequence
D-loop – Displacement *loop*, região controlo
EDNAP – European DNA profiling group
ENFSI – *European Network of Forensic Science Institutes*
FISH – *Fluorescente in situ Hibridization*
G – Guanina
G.F. – Genética Forense
H – Haplótipos
He_{esp} – heterozigotos esperados
He_{obs} – heterozigotos observados
HVI – Region hipervariável
ISFG – International Society for Forensic Genetics
LINES – Long Interspersed Repeated Sequences
MB – Mega Bytes
µl – microlitros
MSY – specific Male Y, região masculina específica
NRY – Non Recombining region of the Y chromosome
PAR1 – Região pseudoautossómica 1
pb – pares de bases
PCR – Polymerase Chain Reaction
PD – Poder de Discriminação
P.Ex – Poder de exclusão
rpm – rotação por minutos
SINES – *Short Interspersed Repeated Sequences*
STRs – Short Tandem Repeats, microssatélites
SRY – Sex Determining Region of the Y Chromosome
SWGDAM – *Scientific Working Group-DNA Analysis Methods*

T – Timina

Taq – DNA polymerase derivado da thermos aquaticus

TDF – Testis Determining Factor

YHDH – Y chromosome Haplotype Reference Database

UR – Unidades de Repetições

UV – ultra violeta

VNTR – *Variable Number Tandem Repeats*

“Costumávamos pensar que nosso destino estava escrito nas estrelas. Hoje sabemos que, em grande parte, ele está em nossos genes.”

James Watson, Time,
20 de março de 1989

“É muito importante que o homem tenha ideais. Sem eles, não se vai a parte alguma. No entanto, é irrelevante alcançá-los ou não. É apenas necessário mantê-los vivos e procurar atingi-los.”

Dalai - Lama

“Queria que vistes o que é realmente coragem, em vez de teres a ideia de que a coragem é um homem com uma arma na mão. A coragem é saberes que não tem hipótese antes de começares e, apesar de começares, conseguires acabar, sejam quais forem as dificuldades”.

Harper Lee

“ Época triste é a nossa em que é mais difícil quebrar um preconceito do que um átomo “

Albert Einstein

1. Justificação do tema

e objectivos gerais

1.1 – Justificação do tema

A Genética Forense constitui um dos ramos da Medicina Legal que se reveste de considerável importância na actualidade, devido ao seu contributo ao serviço da Justiça, na resolução de problemas do foro civil e criminal. Trata-se da aplicação da análise genética da diversidade humana à resolução de processos judiciais.

O estudo dos polimorfismos de ADN e o seu manuseamento, com recurso à técnica de *polymerase chain reaction (*PCR*),* possibilitou a superação das dificuldades apresentadas pelos marcadores clássicos de tipo sanguíneo, proteínas séricas, que apresentavam uma baixa heterozigosidade. Os avanços tecnológicos associados ao conhecimento molecular da transmissão genética permitiram o desenvolvimento da Genética Forense, referente à identificação genética individual, criminalística biológica e investigação biológica de parentesco. O estudo genético dos polimorfismos de ADN permite também a obtenção de dados relativos à origem e dispersão das populações.

O avanço da ciência e da tecnologia a nível forense teve seu ponto culminante em meados dos anos 80, quando as técnicas de ADN se tornaram uma das mais poderosas ferramentas para a investigação humana. A análise dos polimorfismos de ADN constitui um dos maiores progressos técnicos da investigação criminal desde a descoberta das impressões digitais.

Actualmente, a prova de genética forense é aceite universalmente em tribunais como prova pericial em casos criminais, para a identificação de desconhecidos, desastres em massa, e problemas de imigração, etc, assim como em investigação antropológica. A identificação de pessoas desaparecidas, de cadáveres em avançado estado de putrefacção ou resultantes de desastres em massa, reconhecendo-se as grandes dificuldades que tal tarefa implica, representa um grande desafio sobretudo em países em vias de desenvolvimento como Angola, que não beneficia ainda de serviços especializados como a Genética e/ou Antropologia Forenses.

Pretende-se com este trabalho evidenciar a importância da identificação humana, através do perfil genético de criminosos que tenham deixado vestígios biológicos, evitando a impunidade e preservando assim os direitos humanos dos cidadãos. Esta necessidade é tanto maior quando, num país como Angola, muitos cidadãos não estão sequer registados nos Arquivos Centrais de Identificação.

Por outro lado, a não identificação de vítimas resultantes de um acontecimento nefasto ou de situações de desastres em massa, acarreta situações desagradáveis com implicações de ordem familiar (por falta de acto fúnebre condigno e humano), social e económica para o Estado.

Este trabalho inclui um estudo de marcadores polimórficos de STRs, autossómicos, do cromossoma Y e do ADN mitocondrial, nos principais grupos étnico-linguísticos de Angola. Assim, julga-se que os seus resultados poderão prestar um contributo importante para a resolução de problemas do foro judicial, que muitas vezes são irresolúveis devido à escassez de estruturas e técnicos capacitados.

Poderá também marcar o início embrionário dos Serviços de Genética Forense em Angola, em prol da Justiça e dos direitos humanos, e contribuir para a criação de uma base de dados de aplicação forense, uma vez que as amostras são oriundas da população em que se pretende realizar a perícia.

1.2 – Objectivos gerais e específicos:

Objectivos gerais:

1. Estudo e caracterização da diversidade genética nos principais grupos populacionais em Angola.

2. Formação e desenvolvimento científico e tecnológico em Genética para aplicação forense em Angola.

Em termos de objectivos específicos pretendeu-se determinar e descrever:

- a diversidade genética nos principais grupos étnico-linguísticos (Bakongo, Kimbundo, Kwanhama, Nganguela, Nhaneca-Humbe, Lunda-Tchokwe, Ovimbundo) de Angola.

- a variabilidade e grau de polimorfismo dos STRs (autossómicos, do cromossoma Y e do ADN mitocondrial) por grupos étnicos, na população de Angola.

- as frequências alélicas dos STRs autossómicos e STRs do cromossoma Y na população Angolana.

- os parâmetros estatísticos de interesse forense.

- os diferentes haplótipos e respectivas frequências do cromossoma Y, nos três principais grupos étnico-linguísticos de Angola, e adicionar os dados referentes à população de Angola no banco de dados do cromossoma Y (Y-STR Haplotype Reference Database).

- a diversidade nucleotídica e haplotípica para o ADN mitocondrial e caracterização dos haplótipos dos três grupos étnico-linguísticos de Angola.

- a distância genética entre as populações angolanas e elaboração das respectivas árvores filogenéticas.

- a comparação entre os principais grupos étnico-linguísticos e a população Angolana em geral com populações de outras origens.

2. Introdução

Origem do homem moderno

História e caracterização da população angolana

Genoma humano

Polimorfismos do ADN

- Polimorfismos de STRs autossómicos
- Polimorfismos de STRs do cromossoma Y
- Polimorfismos do ADN mitocondrial

Aplicação forense dos polimorfismos de ADN

Marcadores de linhagens em Genética

2.1 – Origem do homem moderno

O ser humano é um mamífero, fazendo parte da ordem dos primatas da família *hominidae*, pertencente ao género *homo*. A precariedade de informações limita o conhecimento da origem dos humanos. As primeiras pesquisas sobre este assunto datam do final do século XIX e culminaram com descobertas de vestígios de antepassados nossos em África e noutras zonas do globo.

A origem do ser humano começou a ser estudada através de fósseis encontrados ao longo dos anos, considerando-se que o período da Pré-história se estende a partir do ano 4.000 antes de Cristo (A.C.). A evolução humana divide-se em três etapas: Paleolítico, Neolítico e Idade dos Metais.

Os primatas podem dividir-se em diversos grupos que descendem de um antepassado comum. Cada um evoluiu em determinada direcção, tendendo para uma especiação cada vez maior. Morfologicamente, o processo de evolução destes primatas foi caracterizado pelo aumento do volume cerebral e o desenvolvimento das respectivas faculdades.

Os primatas surgiram na terra há cerca de 80 milhões de anos e a separação entre os vários géneros Antropóides terá ocorrido entre 20 a 5 milhões de anos atrás. Os fósseis dos *Australopithecus* com idade entre 4 a 1.5 milhões de anos encontrados na África meridional (região dos Grandes Lagos) e oriental são os mais antigos e conhecidos até à presente data. Os *Australopithecus* tinham em média 115 centímetros de altura, andavam na posição vertical e já manuseavam instrumentos como paus e pedras no seu dia-a-dia, tendo uma dieta baseada no consumo de frutos e sementes. O *Australopithecus* é o mais antigo hominídeo que se conhece. No entanto, apesar do crânio pequeno, apresentava traços característicos dos hominídeos.

O *Homo habilis* viveu há cerca de 2.5 milhões de anos e foi contemporâneo do *australopithecus*, mas apresentava uma capacidade craniana maior quando comparada com a destes últimos. O seu cérebro era um pouco maior, o que lhe conferia mais inteligência e habilidade do que às outras espécies, facto que fez com que eles se tivessem destacado em termos de evolução. Esta espécie incluiu carne na sua alimentação, o que provocou mudanças na sua arcada dentária. O termo *Homo habilis* significa “homem habilidoso”, pelo facto de esta espécie ter manifestado destreza com as mãos e com instrumentos de trabalho.

O *Homo erectus* possuía maxilares maciços e dentes grandes, um cérebro maior do que o tipo anterior e membros melhor adaptados à postura erecta. Esta espécie apresentava altura e inteligência quase equivalentes às dos humanos actuais.

De modo geral, dizemos que existe um tronco comum do qual se originaram os grandes macacos (*pongidae*) e os humanos (*hominidae*). Em determinado momento da evolução, os dois grupos separaram-se, tendo cada qual apresentado a sua evolução própria. Os *pongidae* assumiram a forma do gorila, chimpanzé e orangotango; os hominídeos evoluíram para a forma do actual *Homo sapiens*, que terá surgido há 250 mil anos, no decurso da evolução do *Homo erectus*, tendo-se tornado a espécie dominante.

Resumindo, o Homem não evolui directamente dos macacos, mas sim descende do *Homo sapiens sapiens*, tendo passado por vários períodos de organização (Paleolítico e Neolítico) e por um longo processo de humanização ao longo de milhões de anos. Actualmente, o ser humano tenta explicar a sua origem fundamentando-se em duas teorias, que são unânimes em afirmar que o *Homo erectus* evoluiu em África e daí se expandiu para os restantes continentes, há cerca de 1-2 milhões de anos.

A teoria multirregional – ou da continuidade evolutiva – sugere que o humano moderno evoluiu a partir do *Homo erectus*, por meio dum processo de evolução. Assim sendo, o *Homo erectus* terá migrado de África há 1.5 a 2 milhões de anos e expandido pelos diversos continentes, mantendo um fluxo de genes suficiente para que surgisse o *Homo sapiens*, de forma lenta e progressiva, fruto das várias migrações entre as populações de todo o mundo habitado; logo, o *Homo neanderthalensis* não era uma espécie, mas uma população que contribuiu geneticamente para a formação do *Homo sapiens*, dando origem ao homem moderno (Green *et al.,* 2010). Além disso, fósseis humanos encontrados na Europa e na Ásia apontam para o facto de que hominídeos já habitavam o planeta antes da migração africana do *Homo erectus*. Achados de crânios com características associadas de *Homo erectus* e *Homo sapiens* em África, no sudeste Asiático e na Europa com mais de 100 mil anos, corroboram a teoria da coexistência, que sugere a hipótese de que o homem moderno conviveu com homens de Neandertal (Green *et al.*, 2010) e o *homo erectus*. O cruzamento com outras populações deu origem a formas intermédias como os Neandertais na Europa e na Ásia. A continuidade genética entre as populações seria mantida pela dispersão genética através da selecção natural.

Esta teoria sugere ainda que, há 300 mil anos, em várias regiões do planeta, o *homo erectus* evoluiu no sentido do *homo sapiens* e, posteriormente, por um processo de diferenciação regional, evoluiu para a sub-espécie do *Homo sapiens sapiens*. A

ocorrência de fluxo genético entre populações levou a intercruzamentos e não a substituição (Cann *et al.*, 1987).

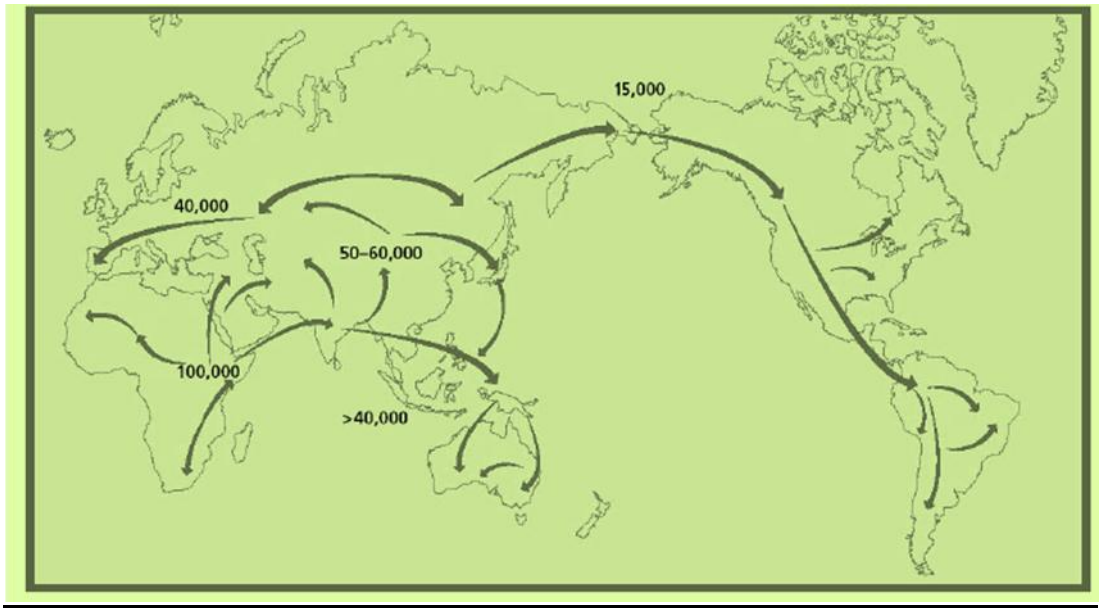


Figura 2.1 – Origem do homem moderno e processos migratórios em África (retirado da www.laboratoriogene.info/Genealogia_por_DNA/Materna.pdf).

A teoria "out of Africa", ou modelo de substituição, defende que as populações de humanos modernos terão surgido no leste de África há mais de 200 mil anos, substituindo progressivamente as outras espécies como o *Homo erectus*, na Ásia e os Neandertais na Europa e na Ásia ocidental. De acordo com este modelo, ter-se-iam realizado as primeiras migrações a partir de África há 100.000 anos, em direcção à Ásia ocidental, através da região do Suez, e em direcção à região da actual República do Iémen, tendo posteriormente rumado em direcção ao sudeste Asiático (Fig. 13). Segundo esta teoria, depois da migração inicial, as populações de *homo erectus* dispersas deram origem, por radiação adaptativa, a várias espécies; o *homo sapiens*, surgido em África há 100.000 ou 200.000 anos, substituiu as populações existentes em todo o mundo, sem trocas significativas de genes; as variações regionais são um fenómeno muito recente.

Há cerca de 40.000 anos, a partir da região ocidental do continente asiático, deu-se uma onda migratória do *Homo sapiens sapiens* para a região da Europa; a partir da região do sudeste da Ásia, terá migrado e povoado a Austrália. Outras vagas migratórias, há cerca de 35.000 a 15.000 anos, foram registadas do nordeste Asiático através da região do actual estreito de Bering até à América do Norte.

Como estas populações não eram isoladas umas das outras, o fluxo de genes ocorria, dando lugar a uma miscigenação das espécies. O processo evolutivo de hominização conduziu, a partir de um primata ainda desconhecido, à forma actual do

homem, quer física quer intelectualmente. Actualmente, é aceite que os primeiros hominídeos - todos os *australopithecus* e *Homo habilis* - evoluíram em África durante os últimos 3 milhões de anos.

A teoria do "out of Africa", - as ininterruptas migrações do *homo habilis* e a ocupação progressiva de todos os recantos do globo são confirmadas por estudos paleontológicos e genéticos, nomeadamente do ADNmt e do cromossoma Y, que apontam para a hipótese da origem do ser humano moderno com um antecessor paterno e materno africanos.

A análise do ADNmt representa uma ferramenta importante no esclarecimento sobre a origem humana (Cann *et al.*, 1987; Vigilant *et al.*, 1991; Krings *et al.*, 1997; Gutierrez *et al.*, 2002), de processos de evolução humana (Torroni *et al.*, 1996; Cann *et al.*, 1987; Mateu *et al.*, 1997), estudos da dispersão e migração de populações pelos continentes (Wallace e Torroni, 1992; Bonato e Salzano, 1997; Mesa *et al.*, 2000; Pereira *et al.*, 2001), assim como a diferenciação de grupos humanos (Chen *et al.*, 1995; Torroni *et al.*, 1996; Salas *et al.*, 2002).

Por outro lado, o continente africano apresenta maior variabilidade genética entre os indivíduos, quando comparado com os restantes continentes, podendo este ser também um bom indicativo da origem mais antiga de toda a linhagem humana moderna.

2.2 – História e caracterização da população angolana

A população angolana ronda os 15 milhões de habitantes. A República de Angola ocupa uma superfície de 1.246.700 km^2. Situada na África austral, faz fronteira a norte com a República do Congo-Brazaville e com a República Democrática do Congo (ex-Zaíre), a leste com a República Democrática do Congo e a República da Zâmbia, a sul com a República da Namíbia, e a oeste é banhada pelo Oceano Atlântico

A fig. 2.2 ilustra a divisão administrativa e política da república de Angola e a sua localização no continente africano.

Figura 2.2 – Mapa da divisão administrativa de Angola.
(retirado: http://www.google.co.uk/imgres? imgurl=http://opatifundio.com/site/wp).

Geograficamente, 60% do território é constituído por planaltos de 1.000 a 2.000 m., com uma densa e extensa rede hidrográfica. Apresenta um clima tropical com duas estações: cacimbo (seca) e chuvas (mais quente). As temperaturas variam entre 17º C (mínima) e 27º C (máxima). Angola divide-se administrativamente em 18 províncias, subdivididas em 163 municípios, cada uma delas possuindo as respectivas autoridades locais.

Muito antes da chegada dos navegadores portugueses, liderados por Diogo Cão, ao reino do Congo, em Angola, no ano de 1482, este território era constituído por diversos reinos e habitado por muitos povos. O reino do Congo localizava-se no sudoeste de África, nos territórios que correspondem hoje ao noroeste de Angola, Congo-Brazaville, parte ocidental do Congo Democrático e parte sul do Gabão.

2.2.1 – Primeiros habitantes de Angola

Os primeiros habitantes deste imenso território eram os povos Bosquímanos, Vátuas, Bantu, e os principais reinos eram os do Congo, Ngola, Matamba e Bailundo. A formação e organização daqueles reinos estavam directamente relacionadas com a sucessiva vaga de migração dos povos Bantu, que eram comuns na África central, meridional e oriental.

Bantu é uma palavra composta, na qual o radical "*ntu*" significa ser humano ou pessoa; ao acrescentar-se-lhe o prefixo "*ba*", obtém-se este termo para referir todos os povos que o utilizam para designar o ser humano, sendo "*mutu*" o seu plural. Os povos Bantu são indivíduos negróides que terão vindo do noroeste da floresta equatorial (vale do Benué, a leste da Nigéria) ou da região dos Grandes Lagos, no Centro de África, e migraram para o Sul. Estes povos seguiram o percurso dos rios Luapala, Lualaba e, posteriormente, do rio Zaire, assim como os seus afluentes; durante milhares de anos foram-se fixando em diversos pontos da África ao sul do Sahara, ao passo que outros grupos, no seu percurso para o sul, se desviaram para o deserto do Kalahari.

As causas das migrações destes povos para o sul de África poderão ser múltiplas, salientando-se a defesa e luta pela sobrevivência, as catástrofes naturais que os obrigaram a procurar terras mais férteis, ou mesmo ainda, o desentendimento entre grupos mais próximos (Redinha, 1974).

Os povos Vátuas ou pré-Bantu (formados por Cuepes e Cuissis) ou Curocas (devido ao rio Curoca da região que eles habitaram), foram habitantes anteriores aos povos Bantu naquela região e são descendentes dos pigmóides bosquímanos. Os Hotontote-Bosquímanos (homens dos bosques), que até então viviam na parte norte e sul do rio Zaire, foram expulsos de suas terras pelos Bantu e obrigados a refugiarem-se mais para o interior, nas zonas do sudoeste africano ou em zonas desérticas da faixa ocidental, onde procuravam melhores condições de vida (Redinha, 1974; Coelho *et al*., 2009).

Na figura 2.3 está ilustrada as fases I, II e III dos processos migratórios dos grupos Bantu a partir das regiões dos grandes lagos em direcção, sudeste, sudoeste e ao sul de África. A III fase do processo migratório foi mais ampla, chegou a transcender o continente africano.

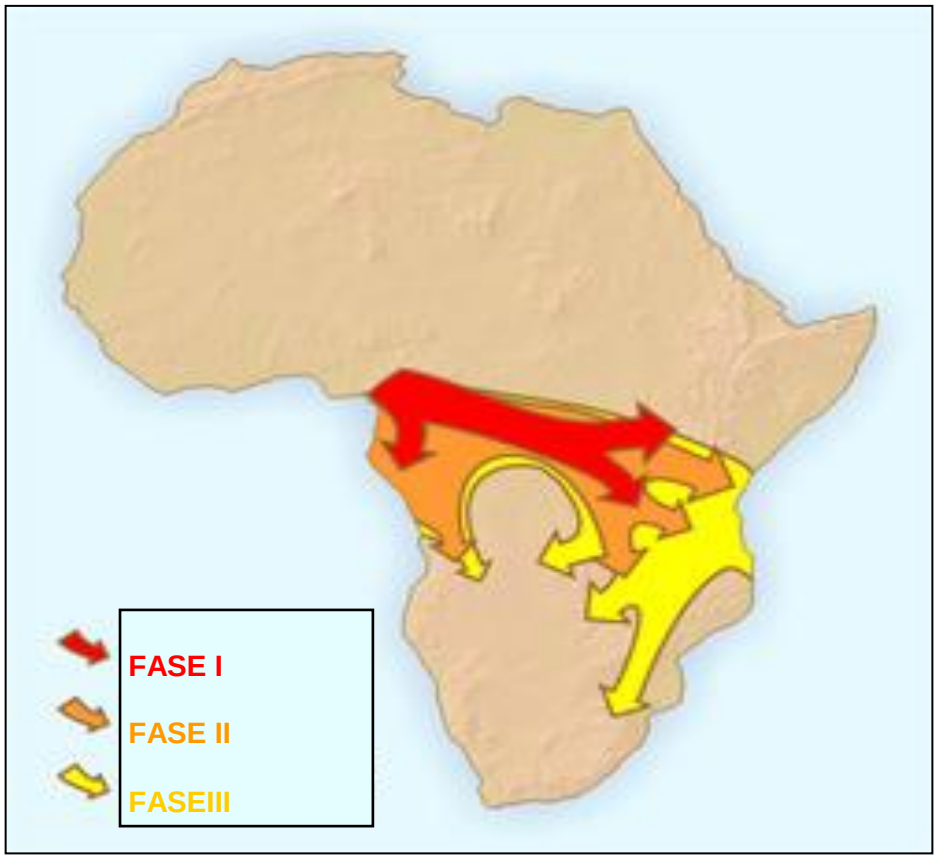


Figura 2.3 – Processos migratórios Bantu (retirado da http://ptwikipédia.org/wik/expans_bantu).

Os Bosquímanos Khoisan constituem a designação de uma família de grupos étnicos existentes no sudoeste de África que partilham algumas características físicas e linguísticas. Apresentam uma estatura média mais baixa, que os diferencia dos restantes povos africanos, possuem uma coloração da pele amarela e prega epicântica nos olhos. Outra característica física deste grupo consiste na esteatopigia (grande desenvolvimento posterior das nádegas) nas mulheres.

Os Vátuas e os Khoisan vivem ainda actualmente da caça e da colheita de frutos silvestres (Fig. 2.4). A comunicação entre eles processa-se através de estalidos linguísticos. A poligamia não é frequente nesta comunidade, a qual dificilmente aceita cruzamentos com outros grupos.

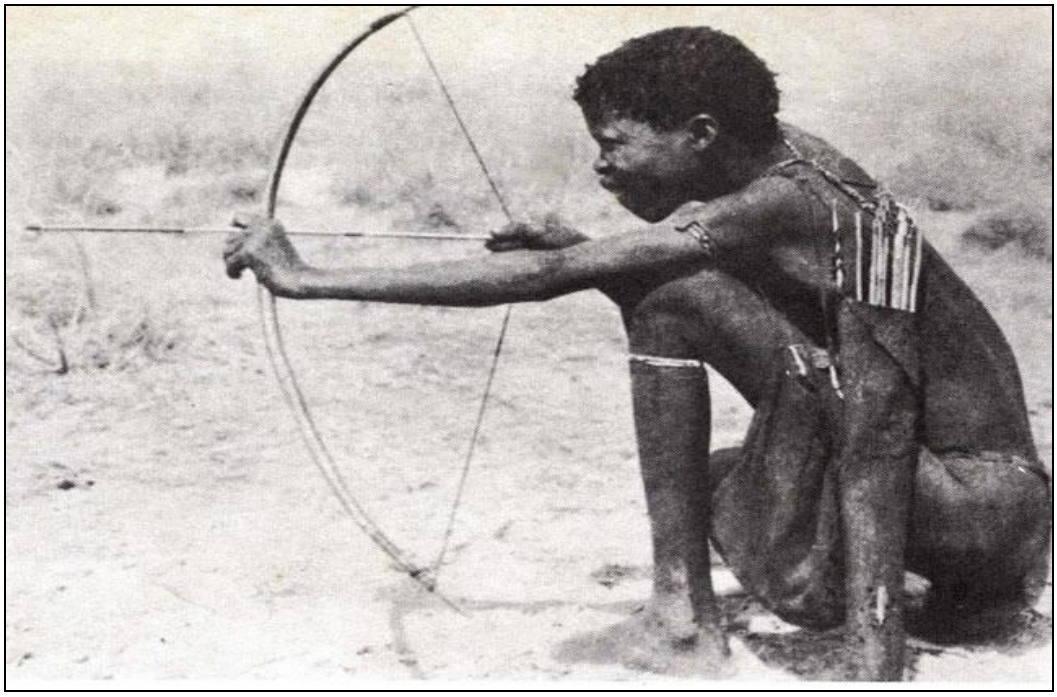


Figura 2.4 – Caçador Bosquímano no sul de Angola
(retirado de http://www.nossokimbos.net/Etnografia/Povos/index.htm).

A palavra Angola deriva do nome de um dos maiores reis, de nome Ngola, da região do Ndongo. Aquando dos primeiros contactos entre os portugueses e os nativos, houve deturpação do título hereditário do Rei Ngola, o que levou os primeiros Portugueses a designar aquela região como Angola (Oliver, 1980).

Na actualidade, a maior parte dos povos que povoam o território de Angola descende dos Bantu ocidentais e meridionais, principalmente no sudoeste, e da miscigenação dos diversos grupos, quer entre si, quer com a população caucasiana, distribuindo-se pelos seguintes grupos étnico-linguísticos (Fig. 2.5):

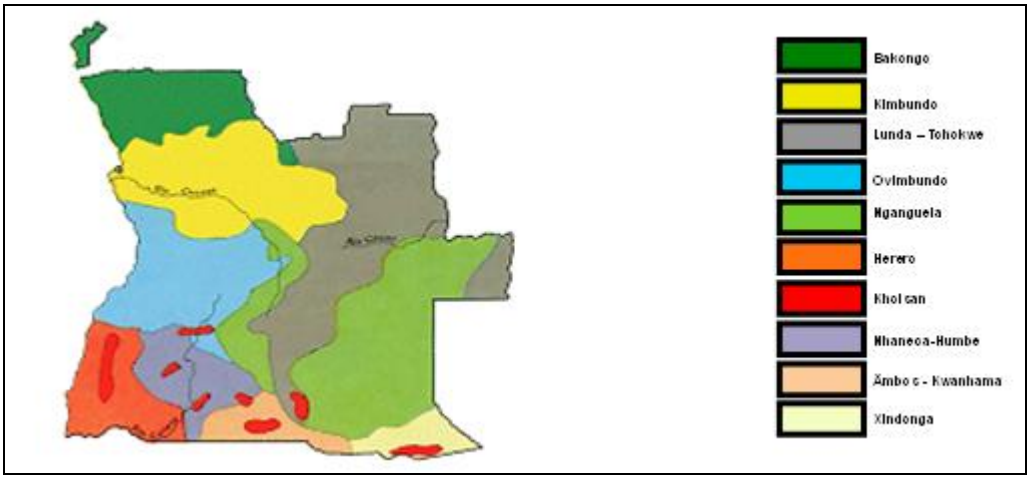


Figura 2.5 – Mapa étnico-linguístico de Angola (adaptado)
(retirado de: http://www.triplov.com/letras/americo_correia_oliveira/literatura_angolana/anexo3.htm).

2.2.2 – Principais grupos étnico-linguísticos actuais em Angola

O grupo étnico-linguístico Bakongo ocupa o norte e o nordeste de Angola, localizando-se nas províncias de Cabinda, Uíge, Zaire, na parte noroeste da província do Bengo e no noroeste da província da Lunda-Norte, e numa faixa próxima do rio Kwango (Fig. 2.2). Constatou-se aqui a centralização do poder entre os reis Bakongos. A sucessão do poder no reino ocorria de irmão para irmão ou para o sobrinho mais velho, filho da irmã uterina, e nunca para os próprios filhos. O pai era um mero progenitor, não possuía direitos nem deveres em relação aos filhos, cabendo a responsabilidade dos filhos ao irmão mais velho da esposa (Redinha, 1974).

Figura 2.6 – Povos Bantu de Angola
(retirado de http://www.nossoskimbos.net/Etnografia/Povos/index.htm).

Os Bakongos apresentam as seguintes características antropológicas: cabeça alongada, prognatismo acentuado, nariz largo e achatado, lábios espessos, estatura média ou superior, apresentando variações individuais ou agrupais devido ao cruzamento com outros grupos (Redinha, 1974).

A tradição ancestral refere que foi um chefe chamado Nimi a Lukeni que reuniu no passado todos os clãs que falavam Kikongo, fundando o Reino do Congo, com capital em Mbanza Kongo, situada na província Angolana do Zaire.

O grupo étnico-linguístico Bakongo era constituído por dezena e meia de sub-grupos que se dedicavam à agricultura, à caça e à pesca e ao comércio, e adoptou a moeda "njimbu" que consistia de conchas marinhas, o que assegurou o apogeu comercial do grande Reino do Congo. Dada a sua situação geográfica, foi um dos primeiros Reinos a ser contactado pelos portugueses, em 1482.

O grupo étnico-linguístico Kimbundo é vizinho dos Bakongos, ocupando uma área que se estende do norte a sul do médio Kwanza e do mar ao rio Kwango, correspondente à zona centro norte. Espalha-se de Luanda até aos Lunda-Tchokwe e confina a sul com os Ovimbundos, localizando-se nas províncias de Luanda, Kwanza-Norte, Kwanza-Sul e na província de Malange (Fig. 2.2). Este grupo, tal como outros, originou admiráveis organizadores de estados, dos quais se destacaram grandes sobas guerreiros (reis), tal como o rei Ngola, Ginga e os Quinguris dos Bangalas, na luta contra a ocupação portuguesa.

Os historiadores deram ao reino dos Kimbundos o nome de Ndongo-Ngola e Ndongo-Matamba, na região de Malange, onde viviam as tribos Gingas. O reino das Gingas tornou-se famoso pela crueldade das suas rainhas, que estenderam a sua supremacia sobre outros povos, sendo que metade deste grupo era dominado pelos Bakongos (Redinha, 1974).

Os Kimbundos foram bons agricultores de subsistência e criadores de gado bovino e caprino. No grupo étnico linguístico Kimbundo, é notável a existência de traços culturais do grupo étnico linguístico Bakongo, principalmente da zona Norte, talvez devido ao facto de estarem espacialmente muito próximos (Redinha, 1974). Entretanto, foi o primeiro grupo a ser invadido militarmente pelos portugueses, a partir dos meados do século XVI; na sua resistência contra a dominação estrangeira, destacaram-se várias figuras, entre elas a Rainha Nzinga, que liderou a resistência no século XVII. Finalmente, em nome da paz, os Kimbundos foram obrigados a conviver com os portugueses e a acostumarem-se à sua cultura.

Sempre se considerou que os Ovimbundos estavam inseridos nos processos migratórios Bantu, mas várias hipóteses têm sido levantadas sobre a origem deste grupo étnico-linguístico: uma delas é que seriam descendentes dos Bantu que se fixaram no Planalto Central, deslocando-se pela faixa litoral Atlântica até à região Central de Angola. Estes dados são sustentados pela linguística, visto que os Ovimbundos utilizam alguns termos próximos daqueles dos povos Igbo da Nigéria (Childs, 1949; MCculloch, 1952).

Outra hipótese, a mais aceitável para alguns historiadores, baseia-se na descoberta de pinturas rupestres da estação arqueológica de Kaniniguri, na região do

Mungo e do Bailundo, que remontam há mais de 9600 anos atrás. Com base nas evidências arqueológicas e na tradição oral, presume-se que os Ovimbundos seriam descendentes dos autores destas pinturas rupestres Hotentotes e que, ao longo dos anos, através de um processo de miscigenação e aculturação, foram adquirindo traços dos outros grupos Bantu e sendo influenciados pelos mesmos.

A origem dos Ovimbundos permanece um mistério para muitos, devido à falta de informação fidedigna sobre esta questão. Alguns autores não descartam a possibilidade de que os Ovimbundos sejam descendentes dos Bakongos, visto que eles falam uma língua – o Ombundo - que constitui uma síntese do Bantu-Congo e do Bantu-Lunda, factos que não deixam de ter evidência científica devido à posição geográfica que os mesmos ocupam.

Esta miscigenação não se limitou apenas aos aspectos somato-culturais e linguísticos, abarcando também a sua capacidade de adaptação (Neto, 1997; Childs, 1949; MCculloch, 1952). Estes povos localizam-se nas províncias do Huambo, Bié, Benguela, numa parte das províncias da Huíla e Moxico, e numa secção da província do Kwanza-Sul (Fig. 2.2).

O mais provável é que os Ovimbundos sejam descendentes dos Bantu ocidentais que, ao longo do seu percurso, se fixaram no planalto central ou, através de um processo de aculturação e miscigenação, foram adquirindo traços de outros grupos. Existe ainda a possibilidade de serem resultado de uma simbiose entre os autores das pinturas rupestres de Caninguiri e os Bantu (Melo *et al*., 2006).

É de salientar que a língua Ombundo, para além de não apresentar demarcações geográficas no interior de Angola, é falada em algumas áreas dos países vizinhos, tais como a República do Zaíre, Zâmbia e República da Namíbia. Os Ovimbundos dedicaram-se ao comércio e à agricultura, promovendo um forte intercâmbio de experiências e valores culturais, para além de serem hospitaleiros. A influência cristã foi muito forte neste grupo, o que originou um maior grau de alfabetização. Antes da chegada dos portugueses, este grupo estava organizado em pequenos estados, tais como Mbalundu, Wambu, Viye, Ndulu, Ngalangui e Chiaka, que constituíam uma divisão administrativa daquela vasta área.

O grupo étnico linguístico Lunda-Tchokwe localiza-se nas províncias da Lunda-Norte e Sul, numa parcela da província do Moxico e está também disseminado nas províncias do Kuando-Kubango, Huíla e leste do Bié, compreendendo os Lunda-Lua, Chimbe, Lunda-Demba e Tchokwe (Figs. 2.2 e 2.5).

Reza a tradição oral que os Tchokwe estavam unidos ao império Lunda e se dedicavam à agricultura, à caça e à apicultura. A causa da sua disseminação e expansão

em vários estados deveu-se à recusa de serem tributários do rei Lunda. Este reino estabelecia trocas comerciais com outros povos. A região das Lundas foi povoada pelos Bantu, Lunda-Tchokwe da região do Katanga, do reino conguês, que atravessaram o rio Kassai e se albergaram naquelas terras (Redinha, 1974).

O povo Nganguela é uma das civilizações da idade do ferro que se instalou nos grandes lagos. A sua entrada em Angola teve início em meados do século XVII, quando passaram pelo Norte da Zâmbia e se instalaram no sudeste Angolano, entre os rios Kuando e Kubango, numa área que ocupa a província do Kuando Kubango e uma parte das províncias do Cunene, Bié, Moxico e Huíla (Fig. 2.2).

Este povo era constituído por uma dezena de sub-grupos, dos quais os mais proeminentes são os Mbwelas, os Luchazi, os Nganguelas do Kuvango e do Kuchi, localizados em diversas áreas do sudeste Angolano (Redinha, 1974).

No seu percurso, alguns destes subgrupos foram ficando pelo caminho, pelo que se podem encontrar Nganguelas no interior e exterior de Angola. Actualmente, existem grupos étnico-linguísticos Nganguela na República da Zâmbia e na Namíbia. O termo Nganga, em língua Tchingangela, significa conhecedores do segredo da natureza. A sua instalação em Angola deveu-se às terras férteis para a agricultura e às boas condições para a caça e para o trabalho do ferro.

Tendo em conta a distância ao Oceano Atlântico, a dominação portuguesa apenas foi possível a partir de 1920, devido à resistência dos Bunda, chefiados pelo seu dirigente Muene Bandu, assim como dos Nhemba, encabeçados pelo Rei Chilhauku. Estavam divididos em subgrupos, facto que afectou a autoridade Central. Contudo formaram importantes Reinos, dentre os quais se destacaram os Reinos do Kubango, de Massaka, ou da Rainha Lussinga, de Senge e dos Luvale, sob a prestigiosa dinastia Nhakatolo.

Os Nhaneca-Humbe, vizinhos dos Ovimbundo e dos Ovambo, espalharam-se pelas províncias da Huíla e do Cunene, compreendendo, entre outros, os Muíla, Humbe e Gambo. Dedicavam-se à agricultura e à criação de gado, sendo alguns grupos semi-nómadas. Os Âmbos (Ovambo) localizam-se na Província de Cunene (Fig. 2.2), entre o paralelo 16º e a fronteira com a República da Namíbia, sendo constituídos por Kwanhamas, Kuamatuis, Evales e Kafimas. Praticam uma agricultura de subsistência, a sua base económica assenta na criação de gado e, devido a factores climáticos e a um baixo nível de desenvolvimento, não se libertaram da vida semi-nómada. O célebre Rei Mandume conseguiu unir o povo e organizar uma forte resistência aos portugueses até Setembro de 1917.

Os Kwanhamas representam um dos grupos com maior expressão entre os povos Ambós. Eles apresentam características antropológicas que denunciam o tipo europóide somático, de estatura alta e corpulenta. Os Kwanhamas revelam na sua linha somática alguns traços Camítoedes (Redinha, 1974; Ervedosa, 1980).

Os grupos étnico-linguísticos Hereros, Âmbós e Nhaneca-Humbe são descendentes dos Bantu meridionais e apresentam características culturais influenciadas pela cultura Camítica oriental, tal como a instituição do boi sagrado, introduzida pelos pastores Camíticos do nordeste Africano. Estes grupos terão migrado para a Angola a partir do Sul vindo da vizinha república da Namíbia.

Estas características são pronunciadas nos subgrupos Ximbas e Cuvais. Eles dedicam-se à criação de gado bovino e ao cultivo do massango. Uma das particularidades destes grupos é o facto de apresentarem a pele mais escura em comparação com indivíduos de outros grupos da zona norte, centro e leste. Os Bantu do sul de Angola apresentam normalmente uma maior estatura e uma acentuada diminuição do prognatismo.

Estes povos são fortes e saudáveis, de porte altivo, e costumavam considerar-se superiores a todos os outros. Para preservar a pureza da sua estirpe, não admitiam qualquer ligação com povos de outras tribos. Se uma mulher estabelecesse um relacionamento com um homem que não pertencesse à tribo, seria severamente castigada, e o filho que resultasse desta união marginalizado ou morto. Estes grupos dificilmente aceitavam cruzamentos com outras tribos, chegando mesmo a menosprezar os outros grupos e a considerá-los desprezíveis devido às diferenças sócio-culturais existentes entre eles.

A sua base alimentar consiste no leite e nos seus derivados, e num cereal denominado massambala. A carne para consumo provém da caça, porque eles evitam abater o seu gado. Tal como já se referiu, o ritual mais importante deste povo consiste na festa do boi sagrado, originária da cultura Camítica oriental, levada até à província do Cunene, no sul de Angola, pela migração dos pastores Camíticos do nordeste Africano (Redinha, 1974).

Em todas as sociedades de Angola realiza-se a patrilocalidade, ou seja, após a cerimónia de casamento, a mulher vai viver junto da família do marido, mantendo-se assim a unidade do grupo de referência.

Com a chegada dos primeiros portugueses, o rei do Congo converteu-se ao Cristianismo, assim como também uma parte da população. Posteriormente, deu-se o inicio do comércio escravo, que contribuiu para a expansão do domínio português em

direcção ao interior do país. Paulo Dias de Novais fundou a cidade de Luanda (capital actual de Angola), em 1575.

O comércio de escravos dominou a economia angolana, para além da exportação de cera, marfim e cobre. Os escravos eram levados em caravelas, como mão-de-obra para as plantações da cana-de-açúcar em colónias portuguesas, como as de São Tomé e Príncipe e Brasil. O número de escravos era tão grande que os portos de Angola e do Congo chegaram a totalizar um terço da exportação de escravos africanos para a América e as ilhas do Atlântico. Como o país não era densamente povoado, a procura por escravos expandiu-se até ao centro da África no século XVIII.

O comércio de escravos foi proibido em Angola em 1836, mas apenas decaiu realmente a partir da primeira metade da década de 1850, quando o Brasil fechou o seu mercado para importação de escravos. A escravatura foi oficialmente abolida no Império Português, apenas em 1875.

Depois de vários séculos de presença portuguesa, Angola acabou por se tornar independente, em 1975. Muito antes da independência, o país envolveu-se numa guerra civil que durou 27 anos e causou grandes danos às instituições políticas e sociais do país, fazendo com que muitos angolanos se tornassem refugiados na sua própria nação. O Português é a língua oficial da República de Angola, mas este país possui mais de vinte línguas nacionais. A língua com mais falantes em Angola, depois do português, é o Ombundo, falado na região Centro-Sul de Angola e em muitos meios urbanos, seguida do Kimbundo.

A população de Angola é assim constituída por quatro espaços sócio-culturais, sendo que os grupos étnico-linguísticos Bakongo, Kimbundo e Ovimbundo representam três quartos dos habitantes deste país. A maior parte da população é descendente dos povos Bantu, que não constituem uma etnia específica, mas sim um conjunto de grupos que representam uma comunidade cultural, com uma civilização e uma linguagem similar assente nas mesmas raízes.

Os povos que habitam Angola formaram-se, pois, a partir de migrações de diferentes grupos. A população que habita actualmente o território de Angola é descendente dos Bantu e, por outro lado, resulta da miscigenação e do cruzamento no tempo e no espaço entre os diferentes grupos socioculturais. O padrão étnico de Angola caracteriza-se por uma ampla miscigenação dos povos Bosquímanos, Vátuas, Bantu e dos Portugueses.

2.3 – Genoma humano

2.3.1 – Genes e cromossomas

A célula constitui a unidade básica dos seres vivos, sendo composta por uma membrana citoplasmática, um núcleo e outros organelos (Fig. 2.7). Além disso, ela dispõe de todos os elementos necessários para executar as funções de um ser vivo, entre as quais a nutrição, a produção de energia e a reprodução. No núcleo celular estão contidos os cromossomas, compostos por sequências de ácido desoxirribonucleico (ADN) e proteínas (histonas e não-histonas).

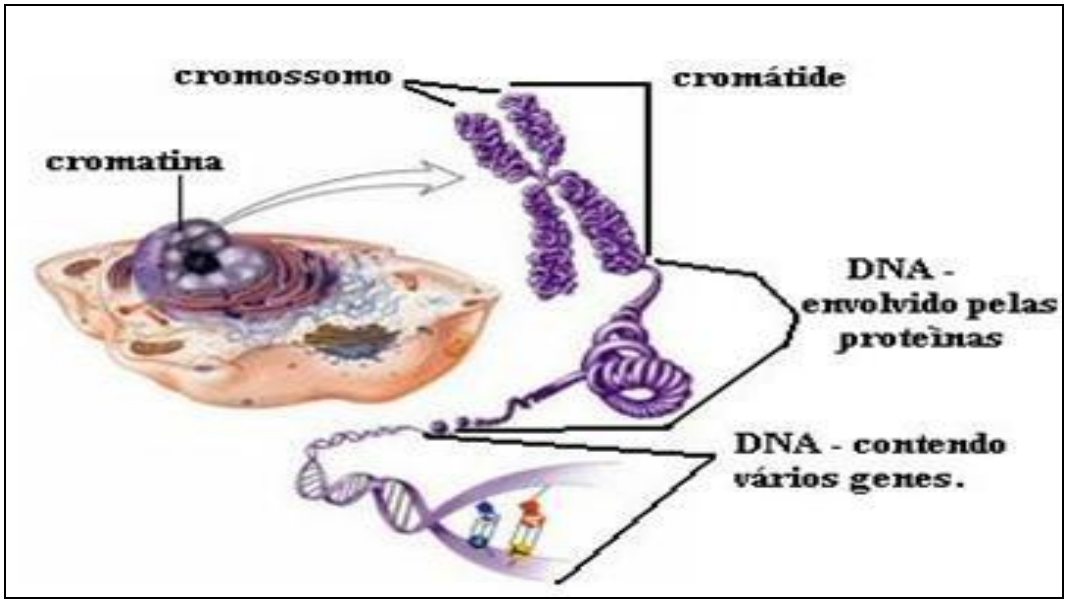


Figura 2.7 – Representação da relação entre a dupla hélice de ADN e os cromossomas e sua localização no interior das células eucariotas (retirado de ttp://www.crv.educacao.mg.gov.br/sistema_crv/banco_obje).

O ADN é a molécula da vida e a sua estrutura básica é composta por um longo polímero de unidades simples (monómeros) de nucleotídeos, cujo cerne é composto por uma base azotada, por moléculas de açúcar e um grupo de fosfato intercalados, unidos por ligações fosfodiéster. O ADN é constituído por duas cadeias helicoidais com orientação antiparalela, em que os pares de bases são complementares ou hibridam entre si. Neste caso, a base nucleotídica adenina emparelha com a timina, e a citosina emparelha sempre com a guanina.

A estrutura tridimensional do ADN foi proposta por Watson e Crick em 1953 marcando assim, o início da genética molecular. O esqueleto de cada hélice é formado pelas ligações longitudinais entre o açúcar de um nucleótido e o fosfato do nucleótido

seguinte, através dos átomos 5´ e 3´ da molécula desoxirribose. As bases de uma hélice emparelham de modo complementar com as bases da outra hélice, ligadas por pontes de hidrogénio, sendo duas ligações entre adenina (A) e timina (T) e três ligações entre a citosina (C) e guanina (G).

As pontes de hidrogénio constituem as principais forças de ligação das duas cadeias de ADN. A exposição a determinadas temperaturas e um ambiente alcalino provoca ruptura das pontes de hidrogénio entre as bases e a consequente separação das duas cadeias que formam o ADN. Este processo é reversível; quando cessam estas condições, volta-se a estabelecer a associação, segundo o princípio da complementaridade entre as suas bases (Darnell 2000; Alberts *et al.*, 1994).

O ADN tem duas funções básicas: a replicação, que é responsável pela hereditariedade, e a transcrição de genes, que produz mensagens para outros locais da célula e que dão origem à síntese de proteínas. As sequências de ADN formam os cromossomas, que são compostos por um complexo de proteínas básicas designadas histonas e um grupo heterogéneo de proteínas ácidas não-histonas.

No cromossoma está inserida a porção funcional do ADN, o gene, que constitui a unidade básica do material genético e determina e controla uma (ou mais) característica específica – a qual se transmite hereditariamente de geração em geração – e ocupa um determinado *locus* (região específica) num cromossoma. Os genes influenciam o funcionamento e o desenvolvimento dos órgãos e determinam a produção de proteínas no organismo.

O genoma humano, na sua forma halóide, é constituído por aproximadamente 3.3×10^9 pares de bases (pb), contidas nos 46 cromossomas (Jones, 2004). Entretanto, os genes são transmitidos, com variações individuais, de geração em geração e determinam a espécie do ser vivo. O genoma representa toda informação hereditária de um organismo codificada no seu ADN, incluindo os genes e as sequências não codificadoras, desempenhando um importante papel na regulação génica, entre outras funções.

As informações contidas num gene são transcritas em ARNs por enzimas, no núcleo da célula, e posteriormente convertidas em proteínas, no citoplasma. Geneticamente, a individualidade biológica de cada indivíduo baseia-se na exclusividade do seu ADN, na igualdade de conteúdo genético entre todas as células nucleadas do organismo, e na invariabilidade deste ao longo da vida.

Na célula, encontram-se dois tipos de ADN: um que se encontra no núcleo celular (nos cromossomas), designado de ADN nuclear, e outro nas mitocôndrias, designado de ADN mitocondrial ou extra-nuclear.

De acordo com a função biológica que desempenha, o ADN é classificado em ADN codificante e não codificante (Jones, 2004). O ADN codificante corresponde a cerca de 30% do genoma, sendo composto por genes, ADN de cópia única e sequências relacionadas com genes (como pseudogenes, fragmentos de genes), regiões promotoras e intrões com função estrutural e reguladora.

As regiões que codificam proteínas constituem uma pequena fracção, cerca de 3% deste tipo de ADN. O ADN codificante contém informação referente aos menos de mais de 100.000 genes humanos (que codificam o ARNs e sintetizam uma proteína génica funcional do genoma (Strachan e Read, 2002).

Portanto, o ADN codificante sofre grande pressão selectiva conservadora e, sendo pouco polimórfico, apresenta escassa variedade entre os indivíduos. Com efeito, qualquer alteração na sua estrutura influencia negativamente a produção proteica gerando consequentemente modificações funcionais e patológicas (Strachan e Read, 2002). O ADN codificante assume grande relevo na prática de medicina clínica.

O ADN não codificante corresponde a 70% do genoma, relacionando-se com sequências de ADN de cópia única ou múltiplas cópias, que correspondem à sua maior porção e se manifestam em sequências repetitivas, de carácter moderado a altamente repetitivo, com funções transcripcionalmente inactivas ou desconhecidas. Deste modo, o ADN de cópia única contém as sequências codificantes para as principais proteínas da célula e manifesta-se em extensões curtas, intercaladas com outras famílias de ADN. O ADN não codificante, nomeadamente o ADN repetitivo, apresenta alto grau de polimorfismo, surgindo representado em todo o genoma humano. Como tal, constitui uma grande fonte de marcadores genéticos, assumindo assim grande importância na genética forense e na genética populacional. O ADN repetitivo do genoma nuclear compreende diversas classes, conforme representado na Fig. 2.8.

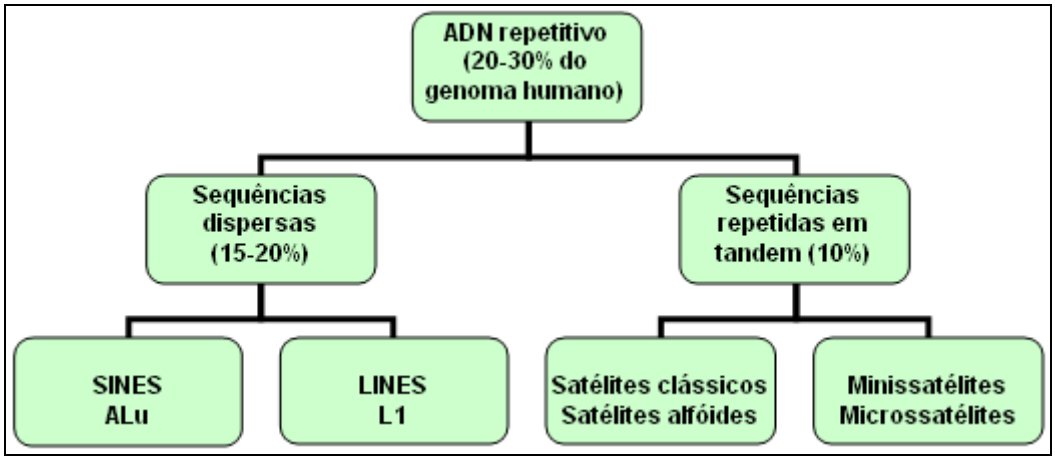


Figura 2.8 – Classificação das sequências de ADN repetitivo no genoma humano (Fowler *et al.*, 1988).

No genoma humano, entre as sequências repetitivas, existe uma classe de sequências repetitivas dispersas, compostas por cerca de 500 pb e que estão representadas por *short interspersed repeated sequences* (SINES), ou elementos nucleares dispersos curtos (sendo as sequências Alu uma representante deste tipo de sequências) e por sequências mais longas designadas por *long interspersd repeated sequences* (LINES) ou elementos nucleares dispersos longos, compostos por unidades com mais de 500 pb. Devido à sua complexidade, este tipo de ADN repetitivo não é aplicado na genética forense (Prak e Kazazian, 2000).

Existem dois tipos de repetições em *tandem* com aplicações em genética forense importantes: os *variable number of tandem repeats* (VNTRs), também designados por minissatélites – apresentam unidades de repetição que variam entre 10 a 70 pb de tamanho, podendo abranger um tamanho total de 500 pb (Bennett, 2000); e os *short tandem repeats* (STRs) ou microssatélites. A estrutura das repetições é basicamente a mesma, variando apenas no tamanho de cada unidade de repetição e na sua composição ou motivo repetido – comprimento.

Os STRs (ou microssatélites) apresentam pequenas dimensões, até 400 pb, e unidades de repetição em *tandem* compostas por 2 a 7 pb, sem localização preferencial no genoma. As suas dimensões facilitam a sua amplificação por meio das técnicas de PCR (Edwards *et al.*, 1991).

2.3.2 – Características dos cromossomas autossómicos

O ser humano possui 46 cromossomas (são diplóides), organizados em 22 pares de cromossomas autossómicos (que são semelhantes no homem e na mulher) e um par de cromossomas sexuais (XX ou XY). Cada progenitor contribui com um de cada par dos 46 cromossomas dos seus filhos.

Durante a meiose, verifica-se a troca do material genético entre os cromossomas autossómicos em cada progenitor. Este processo acontece nos ovários e nos testículos e culmina com a separação dos pares de cromossomas, nos ovócitos ou espermatozóides. Após a fecundação, os gâmetas (células haplóides) fundem-se para formar um zigoto, que dará origem ao embrião (genoma diplóide, 46 cromossomas). Portanto, é o espermatozóide quem determina o sexo da criança, pelo facto de os cromossomas sexuais do homem serem distintos.

O ADN organiza-se em cromossomas e estes transmitem a informação genética através dos genes. O genoma nuclear diplóide é autónomo e apresenta duas cópias, sendo uma de cada progenitor. Durante a meiose, ocorre a recombinação dos cromossomas autossómicos. Na mitose, cada cromossoma replica-se para formar um par de cromatídios, que se mantêm ligados pelo centrómero. Nesta fase, pode ocorrer troca de material genético entre os cromatídios, por *crossing-over*, o que aumenta a variabilidade individual. No final da divisão celular, cada célula-filha tem um número idêntico de 46 cromossomas.

Os cromossomas homólogos são morfológica e estruturalmente semelhantes, contendo genes responsáveis pelas mesmas características. Assim sendo, diz-se que um indivíduo é homozigótico quanto a uma determinada característica, sempre que os alelos de um *locus* forem idênticos. Se dois alelos, para o mesmo *locus* num par de cromossomas homólogos, forem diferentes, o indivíduo é designado por heterogozigotico. O *locus* constitui a posição exacta que cada gene ocupa num dado cromossoma.

Um marcador genético é um *locus* polimórfico, isto é, para o qual existem várias formas possíveis do mesmo gene (polialelismo). Deste modo, o genótipo constitui a composição alélica de um *locus*, e o fenótipo consiste nas características manifestadas por um indivíduo, ou na expressão externa do genótipo. O fenótipo (ou característica) pode ser herdado de modo autossómico dominante, autossómico recessivo ou ligado ao sexo. Em determinadas circunstâncias, o fenótipo resulta da interacção entre o genótipo e o meio ambiente. Neste caso, a variabilidade genética reside na possibilidade de existirem múltiplos alelos num determinado *locus*, quando se considera a globalidade da população.

Os cromossomas caracterizam-se por possuírem um braço curto (p) e um braço longo (q), separados pelo centrómero, sendo a extremidade de cada braço designada de telómero. Consoante a posição dos centrómeros nos cromossomas, estes são designados por telocêntricos, acrocêntricos, sub-metacêntricos ou metacêntricos. Os cromossomas humanos são agrupados em 7 grupos (A a G), que no cariótipo são categorizados de acordo com as suas dimensões e a posição do centrómero (classificação de Denver). Neste caso, o grupo A é constituído pelos cromossomas 1, 2 e 3; o grupo B, pelos cromossomas 4 e 5; o grupo C, pelos cromossomas 6-12 e o cromossoma X; o grupo D, pelos cromossomas 13-15; o grupo E, pelos cromossomas 16-18; o grupo F, pelos cromossomas 19 e 20; e o grupo G, pelos cromossomas 21, 22 e o cromossoma Y (Verma e Babu, 1995).

O cariótipo constitui a representação, de forma organizada, dos cromossomas de uma célula mitótica (em metafase), tendo em consideração o número, a forma, o tamanho e outras características morfológicas dos cromossomas metafásicos nas células de um indivíduo. Neste caso, o cariótipo de um indivíduo do sexo masculino é representado por 46,XY e do sexo feminino por 46,XX. O estudo das regiões cromossómicas permite diagnosticar rapidamente a existência de um número anormal de cromossomas ou alterações da sua morfologia.

Os cromossomas homólogos possuem uma estrutura idêntica para a mesma função genética, à excepção dos cromossomas sexuais, que são diferentes entre si no homem e iguais na mulher (Bailey, 1995). Pressupõe-se que os cromossomas sexuais tenham sido originados pelos cromossomas autossómicos por meio de um processo de recombinação génica. O processo evolutivo, ao longo dos anos, originou a inversão e a recombinação das regiões de ADN do cromossoma Y, que contiveram o seu alinhamento, com regiões análogas ao cromossoma X; este mecanismo de recombinação génica ao longo dos anos gerou mutações na sua estrutura, o que facultou a diferenciação de algumas porções dos dois cromossomas (Brion, 2002).

2.3.3 – Características do cromossoma Y

A amelogenina permite identificar o sexo do indivíduo, visto que este gene está presente, tanto no cromossoma Y (presente só no homem) como no X (uma só cópia no homem e duas na mulher).

O cromossoma Y humano é estruturalmente acrocêntrico, sendo um dos cromossomas mais pequenos. Possui cerca de 60 Mb de comprimento e representa

cerca de 2% do genoma humano (Quintana-Murci *et al.*, 2001). Nas zonas teloméricas de ambos os braços estão situadas as regiões pseudoautossómica 1 (PAR1), e a pseudo-autossómica 2 (PAR2), que correspondem aproximadamente a 5% da sequência total do cromossoma Y (Vogt *et al.*, 1997). As sequências destas duas regiões não se encontram relacionadas com o sexo e são homólogas às sequências do cromossoma X. Por este motivo, durante a meiose masculina, estas regiões participam no processo de emparelhamento e recombinação entre os cromossomas sexuais (Ellis e Goodfellow, 1989; Jobling *et al.*, 1997). Porém, Hassold *et al.* (1991) referem que a falta de recombinação a nível do PAR1 pode produzir alterações fenotípicas, tal como o síndrome de Klinefelter (cujo cariótipo é 47, XXY na maior parte dos casos).

A maior parte deste cromossoma, cerca de 95%, não recombina com nenhum outro, sendo designada de non recombining *region of the Y chromosome* (NRY) ou região não recombinante do cromossoma Y (Lahn *et al.* 2001).

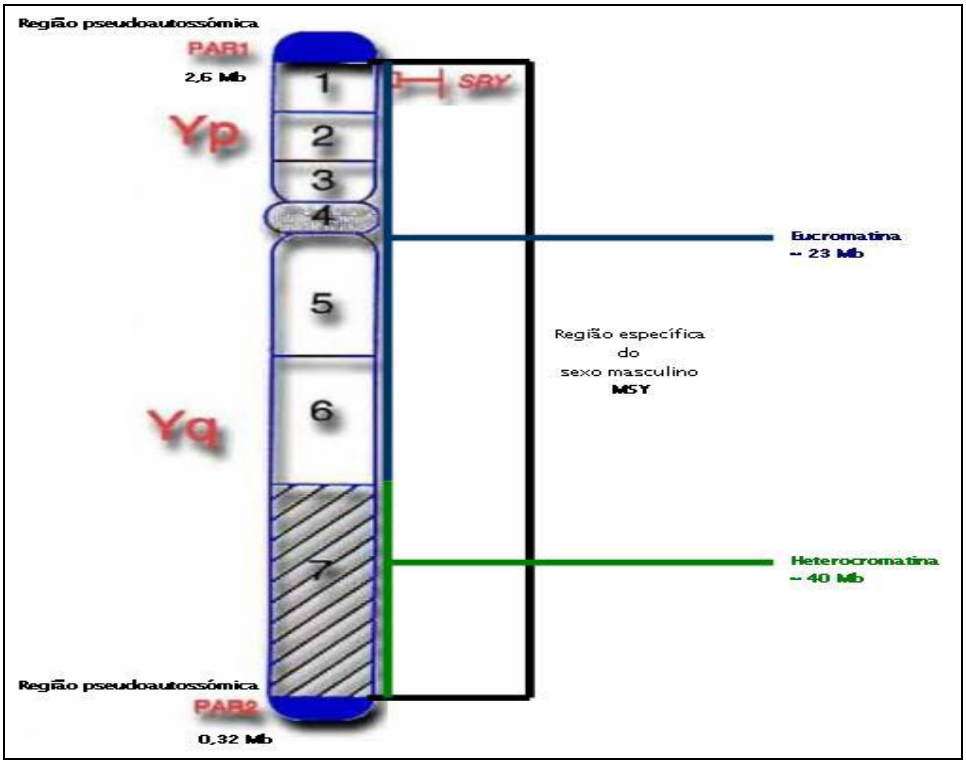


Figura 2.9 – Ideograma do cromossoma Y (adaptação de Quintana-Murci e Fellous, 2001; Skaletsky *et al.*, 2003).

O cromossoma Y tem a porção NRY do cromossoma Y, constituída por sequências polimórficas altamente repetitivas, composta pela região de heterocromatina polimórfica –, com uma dimensão de 40 Mb, localizada na zona distal do braço longo (Yq) (Roewer *et al.*, 1996), variável em indivíduos fenotipicamente normais –, e pela região de

eucromatina – que apresenta um tamanho aproximado de 23 Mb, localizada no braço curto do cromossoma (Yp), no centrómero e na zona distal do braço longo (Skaletsky *et al.*, 2003), designada por male *-specific region of the Y* (MSY). Ambas as regiões são constituídas por ADN não recombinante, com sequências repetitivas específicas do cromossoma Y (Roewer et al., 1992; Hammer e Zegura, 1996).

O cromossoma Y apresenta alguns genes funcionais, que codificam proteínas com funções biológicas específicas relacionadas com a determinação do sexo masculino, tais como o gene *sex determining region of the Y chromosome* (SRY), que determina o sexo; e um gene responsável ou envolvido na espermatogénese (Graves, 2005), responsável pelo desenvolvimento testicular, o *testis determining factor* (TDF). Este gene faz com que as gónadas indiferenciadas se transformem em testículos durante a embriogénese; mutações neste gene afectam o desenvolvimento dos testículos (Su e Lau, 1993). O estudo do cromossoma Y, com base em alguns marcadores genéticos, tem permitido investigar algumas das causas da infertilidade masculina.

O cromossoma Y é transmitido de modo uniparental (paterno) por meio de blocos de forma haplóide, sendo que o pai o transmite apenas aos indivíduos do sexo masculino. A fórmula haplóide do cromossoma Y só possui uma cópia em cada célula, o que difere dos cromossomas autossómicos. A região heterocromatínica do cromossoma Y apresenta algumas semelhanças com o ADN mitocondrial, devido à ausência de recombinação durante a meiose, e ao modo de herança haplóide, em bloco.

A porção não-recombinante do cromossoma Y reveste-se de grande importância para os estudos de patrilinhagens (Vogt *et al.,* 1997; Lahn *et al.*, 2001). A região não recombinante é transmitida em blocos (haplótipos) de pai para filho. A transmissão efectua-se de forma imutável, à excepção das mutações gradualmente acumuladas. As mutações constituem a fonte de variação e, na sua maioria, ocorrem nas regiões intrónicas e extragénicas, sendo transmitidas às gerações seguintes por meio dos haplótipos (Pena *et al.*, 2000). Neste caso, todos os indivíduos masculinos da mesma linhagem paterna apresentarão haplótipos idênticos.

Tendo em conta a sua natureza e especificidade masculina, no cromossoma Y encontram-se registados todos os eventos mutacionais ocorridos ao longo da História humana, permitindo a reconstrução e análise de linhagens paternas. As mutações que foram ocorrendo durante o processo de evolução humana geraram polimorfismos dos haplótipos do cromossoma Y, que podem assim ser utilizados como marcadores de linhagem.

2.3.4 – Características do ADN mitocondrial

As mitocôndrias são organelos presentes em todas as células eucarióticas, localizadas no citoplasma celular, na zona extra-nuclear, podendo apresentar formas e tamanhos variados, diferindo o seu número de acordo com a actividade fisiológica das células. São constituídas por duas membranas: a externa (lisa) e a interna (pregueada, formando as cristas mitocondriais, septos que delimitam a matriz mitocondrial, onde ficam dispersas as estruturas ribossomais, enzimas e diversas cópias de um filamento de ADN circular.

O ADN mitocondrial (ADNmt) é, assim, composto por uma molécula de ADN de dupla fita, de forma circular, aparelhada na matriz mitocondrial como pequenos anéis de filamento duplo, em quantidade que varia de 2 a 10 cópias por cada organiza. Uma célula possui cerca de 500 a 10.000 moléculas de ADN dispostas na matriz das mitocôndriais (Strachan e Read, 2002; Turner *et al*., 2003). A sua estrutura concede-lhe uma grande estabilidade e resistência à degradação.

Existem duas teorias que tentam explicar a origem das mitocôndrias: a teoria autogénica afirma que a célula teria surgido através da especialização de membranas internas, derivadas de invaginações da membrana plasmática; a teoria endossimbiótica sugere que a célula eucariota seria o resultado da associação de células procariotas simbióticas que envolveriam outras complementares que ficaram intactas no interior do hospedeiro.

A análise comparativa revelou que as mitocôndrias surgiram nas células eucariotas, durante a evolução, sendo este facto reforçado por evidências como: a dupla membrana, sendo a interna semelhante aos mesossomos (dobras membranosas de bactérias, ricas em enzimas respiratórias); o pequeno tamanho dos ribossomas, semelhantes aos de procariotas, e diferenciados dos encontrados no hialoplasma da mesma célula eucariota; e a presença do ADN circular.

Portanto, supõe-se que por volta de 2.5 mil milhões de anos atrás, células procarióticas terão fagocitado, sem digestão, arqueobactérias capazes de realizar respiração aeróbia, as quais disponibilizaram assim energia para a célula hospedeira, garantindo alimento e protecção (uma relação harmónica de dependência). Esta teoria é suportada pelo facto de não haver diferenças entre material genético das células eucariotas e procariotas e pelo facto de o processo da simbiose ser muito comum no mundo vivo.

O genoma extranuclear ou mitocondrial representa 1-2% do ADN nuclear. O ADNmt é uma molécula circular pequena, composta por 16.569 pb de comprimento; foi

primeiramente sequenciado e descrito em 1981, por Anderson *et al.* (Anderson *et al.,* 1981). Posteriormente, a sequência foi reanalisada por Andrews et al., (1999), sendo denominada *Cambridge Reference Sequence* (CRS); a região codificante do ADNmt corresponde a 90% e codifica 37 genes, dos quais 22 constituem genes para ARN de transferência, 2 para ARN ribossómico e 13 para a síntese de proteínas, pois participam no processo da respiração celular, sendo responsáveis pela produção de energia celular – que será armazenada em moléculas de ATP (adenosina trifosfato) – e em processos metabólicos, tais como a fosforilação oxidativa.

No entanto, existe uma interacção entre as mitocôndrias e o núcleo celular, assegurando assim o bom funcionamento das células (Anderson *et al.* 1981; Ballard e Whitlock, 2004).

Os restantes 10% do genoma mitocondrial são representados pelo ADN não codificante, que engloba a região controlo total, comportando cerca de 1.112 pb (região hansa D), e se subdivide nas regiões hipervariáveis I, II e III (Fig. 2.10) (Lutz *et al.*, 1997; Lutz *et al.*, 2000; Bini *et al.*, 2003).

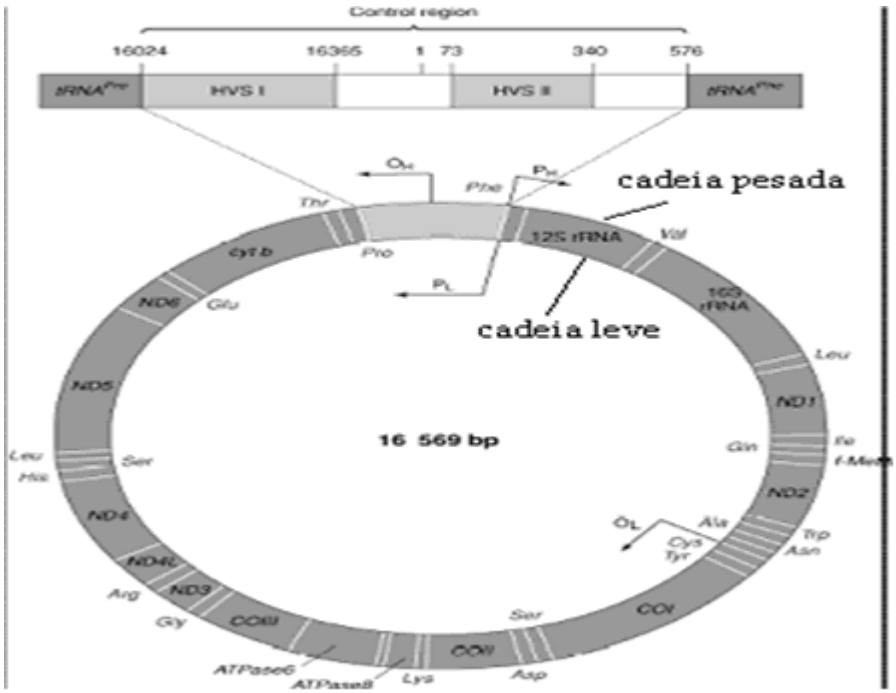


Figura 2.10 – Organização do ADNmt humano (adaptado de Butler, 2005).

A região controlo é responsável pela regulação da replicação e da transcrição de todo o ADNmt (Fig. 9). A replicação do ADNmt é bidireccional e assincrónica, é efectuada por deslocamento de uma cadeia em relação a outra, começando com a replicação da

cadeia pesada e no ponto em que esta termina dá-se a replicação da cadeia leve, mas no sentido contrário (Upholt e Dawid, 1977).

No ADNmt, as duas cadeias apresentam diferenças quanto à distribuição do número de bases (guanina e timina) em cada uma delas. Assim, uma das moléculas de bases púricas (adenina e guanina) é denominada cadeia pesada (*heavy*, H), ao passo que a molécula complementar rica de bases pirimídicas (timina e citosina) é designada de cadeia leve (*light*, L). Cada cadeia é transcrita a partir de um promotor PL e PH1, localizados na região controlo (Ballard e Whitlock, 2004).

Na Tabela 2.1 estão descritas as características que diferenciam os dois tipos de ADN (nuclear e mitocondrial) no genoma humano.

Tabela 2.1 – Comparação entre o ADN mitocondrial e o ADN nuclear (adaptado de Butler, 2005).

Característica	**ADNmt**	**ADN nuclear**
Localização	Mitocôndria	Núcleo da célula
Estrutura	Cadeia dupla circular	Dupla hélice
Nº de genes	37 genes	100.000 genes *
Tamanho do genoma	~ 16.569 pb	~ 3.2 mil milhões pb
Cópias por célula	Podem ser >1000	2 (1 alelo de cada progenitor)
Funcionamento	Necessita da cooperação do ADN nuclear	Autónomo
Característica do genoma	Haplóide	Diplóide
Herança	Materna	Materna e Paterna
Recombinação geracional	Não	Sim
Actividade da polimerase	Fraca	Elevada
Sistema de reparação	Ausente	Presente
Taxa de mutação	5 a 10 vezes > que a nuclear	Pequena
Poder de individualização	Não individualizante (partilhado pela linhamaterna)	Único para cada indivíduo (excepto gémeos monozigóticos)
Sequência de referência	Descrita em 1981, por Anderson e et.	Descrita em 2001, pelo Projecto do Genoma Humano

- * O número de genes varia entre 20.000 - 25.000, (http://ghr.nlm.nih.gov/handbook, 2010)

O ADNmt apresenta características peculiares, devido a (1) elevado número de cópias que apresenta em cada célula, (2) elevada taxa de mutação (5 a 10 vezes superior quando comparada àquela do ADN nuclear) e (3) ausência de recombinação durante a meiose.

A sua elevada taxa de mutação deve-se à ausência de mecanismos de reparação do ADNmt, e também aos danos causados pelos radicais livres de oxigénio resultante do processo de fosforilação oxidativa.

O ADNmt é transmitido de forma uniparental (materna), ou seja, todos os indivíduos (homens e mulheres) da mesma linhagem materna apresentam o mesmo haplótipo mitocondrial.

Cada célula humana possui, pois, um segundo genoma localizado nas mitocôndrias. A análise desta molécula é assim fundamental para o estudo da evolução humana.

2.4 – Polimorfismos do ADN

Os polimorfismos constituem as diversas formas de um *marcador genético* e apresentam grande variabilidade entre os indivíduos, como resultado de mutações no genoma humano, sem que, contudo, das mesmas decorram consequências patológicas ou fenotípicas (Strachan e Read, 2002). Os marcadores genéticos, polimórficos, são facilmente detectáveis na população, e podem referir-se a um gene, um sítio de restrição ou qualquer outra sequência do ADN que apresente diferentes alterações alélicas para um determinado *locus*.

Antes dos anos 80, como meio de identificação individual para o estudo da variabilidade genética, analisavam-se polimorfismos genéticos do sistema sanguíneo ABO (antigénios eritrocitários), proteínas séricas, enzimas eritrocitárias ou leucocitárias do complexo HLA (histocompatibility leucocyte antigen). Todavia, alguns destes marcadores apresentavam limitações, devido ao seu baixo grau de polimorfismo.

Os avanços tecnológicos na área da Biologia permitiram a análise dos polimorfismos de ADN, constituindo um dos maiores progressos técnicos da investigação criminal desde a descoberta das impressões digitais. Jeffreys *et al.*, (1985ª) descreveram uma nova metodologia que permite a análise de ADN, com aplicação na identificação individual e permite superar as dificuldades na análise da diversidade humana (Jeffreys *et al.* 1985ª).

O termo polimorfismo foi empregue por Ford, em 1940, para designar a aparição conjunta, de duas ou mais formas alternativas de um gene ou sequência de ADN não codificante**.** Consequentemente, uma determinada característica genética é considerada polimórfica quando é transmitida de forma independente, relativamente a outros marcadores, e as variações ou alelos mais comuns para esse *locus* devem manifestar uma frequência populacional inferior a 99% (Pestoni et al., 1995). O grau de polimorfismo genético está relacionado com o número de alelos apresentado por um marcador.

Por outro lado, o ADN não codificante apresenta grande variabilidade entre indivíduos, ou seja, é altamente polimórfico, pelo que se reveste de grande utilidade para a Medicina Legal. Para além disso, o ADN não codificante representa a maior parte do

genoma nuclear, constituindo assim uma grande fonte de marcadores genéticos (Jarreta, 1999). Geneticamente, cerca de 99.9% do genoma é semelhante entre os indivíduos e apenas o restante 0.1% gera o polimorfismo que é responsável pela diversidade humana - regiões que variam com certa frequência entre os indivíduos, permitindo assim a identificação humana (Brown, 2003).

Como tal, para o estudo dos marcadores convencionais utilizava-se a técnica de detecção do polimorfismo a nível do fenótipo, ao passo que nos polimorfismos de ADN se recorre à análise ao nível do genótipo. Os polimorfismos de ADN não sofrem modificações devido a factores externos e são invariáveis durante toda a vida do indivíduo, podendo ser aplicados a estudos populacionais, de genética clínica ou forense.

Os polimorfismos presentes nas regiões do ADN podem ser agrupados em dois tipos (Jarreta, 1999; Strachan e Read, 2002):

- Polimorfismos de comprimento - incluem variações quanto ao tamanho dum motivo repetitivo (VNTRs e STRs); caracterizam-se por sequências de nucleotídeos que se repetem em múltiplas cópias, divergindo o número de repetições entre os indivíduos para cada *locus*
- Polimorfismos de sequência - são compostos por diferentes nucleotídeos com variações numa sequência específica de bases de uma determinada região do ADN.

O ADN nuclear, tal como o mitocondrial, apresentam uma grande variedade de polimorfismos que têm sido empregues no estudo de variações genéticas entre populações, assim como em estudos evolutivos.

A introdução de novas metodologias, como a PCR, revolucionou a genética molecular, por permitir uma rápida clonagem e análise do ADN. Esta técnica foi descrita por Kary Mullis em 1986, e consiste na amplificação enzimática *in vitro* de uma determinada sequência de ADN, possuindo vastas aplicações nas áreas da Biologia e da Genética Molecular. As vantagens da PCR consistem na facilidade da sua utilização, sensibilidade, elevada capacidade de discriminação e na possibilidade de se analisarem amostras degradadas (Schuller et. Al., 2001). As suas limitações prendem-se com a contaminação das amostras, facto que pode introduzir erros na análise; para maior controlo na prática laboratorial, se utilizar sempre um controlo negativo (branco), que ajuda a detectar ADN contaminante. Actualmente, a técnica de PCR é a mais utilizada para amplificação no estudo de amostras de ADN nuclear e ADNmt.

2.4.1 – Polimorfismo de STRs autossómicos

Os microssatélites ou *short tandem repeats* (STRs) consistem em regiões de ADN repetitivo em *tandem*, multialélicas, altamente polimórficas, e encontram-se distribuídas por todo o genoma (Wyman e White, 1980; Strachan e Read, 2002). Os STRs são constituídos por 2-7 pb de comprimento, de tamanho variável que se repete em *tandem.* Apresentam um grau de heterozigosidade superior a 70% na maior parte das populações (Edwards *et al.* 1991; Lee *et al.*, 1994; Brinkmann et al., 1996). Apresentam elevada variabilidade e uma frequência de um STRs em 300-500 Kb de número de unidades repetitivas.

Os STRs apresentam uma herança mendeliana, codominante, com dois alelos presentes em cada *locus*, e estão sujeitos a recombinação e mutação.

Os STRs estão dispersos por todo o genoma humano. Estima-se que existam cerca de 200.000 STRs dos quais 6.000 a 10.000 são di, tri ou tetranucleótideos (Fig. 2.11) (Weber *et al.*, 1989; Litt e Luty, 1989; Edwards *et al.*, 1991; Kimpton *et al.*, 1993). As sequências dos STRs são denominadas de acordo com a longitude das unidades de repetições (UR). Os tetranucleótideos (AGAT ou GATA) são os mais comuns para os *loci* STRs utilizados em genética forense.

De acordo com o padrão das UR, os STRs são classificados como tendo baixa microvariação ou simples, que compreendem uma unidade de repetição constante, de 3, 4 ou 5 bp, num número variável de vezes (Brinkmann, 1996); média microvariação ou compostos (VWA, TH01), que compreendem duas ou mais unidades de repetições simples e adjacentes, com um determinado grau de variação estrutural em relação aos alelos consensuais (Brinkmann e Meyer, 1996); e os STRs de alta microvariação, que consistem em variações, quer estruturais, quer relacionadas com uma repetição em grau extensivo. Entre os STRs de microvariação intermédia ou compostos com maior interesse forense, temos o *locus* VWA. Estes sistemas apresentam duas ou mais unidades de repetições contínuas diferentes, que variam tanto na sequência, como na longitude

- **STR – repetições de sequência nucleotídicas**
  - AAAAAA – mononucleotídeo
  - ATATAT – dinucleotídeo
  - TAGTAGTAG – trinucleotídeo
  - TAGTTAGTTAGT – tetranucleotídeo
  - TAGGCTAGGCTAGGC – pentanucleotídeo
- **Alelos diferentes contêm diferentes números de repetições**
  - TTCTTCTTCTTC – quatro unidades de repetição
  - TTCTTCTTCTTCTTC – cinco unidades de repetição

Figura 2.11 - Polimorfismos de STRs autossómicos.

O número de alelos nos STRs tetranucleotídeos pode variar de 5 a 20, de acordo com o tipo de *locus*, para além de possuir entre 100 a 300 pb (Inman e Rudin, 1997). Os STRs utilizados em genética forense apresentam um elevado poder de discriminação e podem apresentar taxas de mutação estimadas entre 10^{-2} e 10^{-5} eventos por *locus* e por geração (Weber *et al.*, 1993; Brinkmann *et al.*, 1998) devido à ausência da pressão selectiva relacionada com a sua posição no genoma. Os STRs aplicados neste estudo são constituídos por sequências tetranucleotídicas com elevado grau de polimorfismo.

A comissão de ADN da *International Society for Forensic Genetics* recomenda que a designação dos alelos dos STRs se deva realizar de acordo com o número de unidades de repetição completas (Brinkmann, 1996; Jacewi R. *et al.*, 2004). Nos casos de UR incompletas, os alelos serão designados pelo número das UR completas separado por um ponto do número de pb da UR incompleta (Gill *et al.* 2001).

A descoberta de um grande número de STRs no genoma humano teve início nos anos 90, momento a partir do qual passaram a constituir uma ferramenta poderosa para a identificação individual e o estabelecimento das relações de parentesco (Weber *et al.*, 1989; Edwards *et al.*, 1991; Laréu *et al.*, 1998). O polimorfismo dos STRs reside na variabilidade de número das UR (Jarreta, 1999), que definem a diversidade entre os indivíduos (Brown, 1995).

Na Tabela 2.2 estão representados diversos sistemas *multiplex* utilizados em muitos laboratórios para o estudo de STRs autossómicos.

Tabela 2.2 – Comparação entre os sistemas *multiplex* utilizados em identificação humana.

Loci	**AmpF/STR Profiler Plus**	**AmpF/STR COfiler**	**Powerplex 1.1**	**Powerplex 2.1**	**Powerplex 16**	**SGM**	**SGM Plus**	**AmpF/STR Identifiler**
D16S539		●	●		●			●
D7S820	●	●	●		●			●
D13S317	●		●		●			●
D5S818	●		●		●			●
CSF1PO		●	●		●			●
TPOX		●	●	●	●			●
TH01		●	●	●	●	●	●	●
VWA	●		●	●	●	●	●	●
FGA	●			●	●	●	●	●
D21S11	●			●	●	●	●	●
D8S1179	●			●	●	●	●	●
D18S51	●			●	●	●	●	●
D3S1358	●	●		●	●		●	●
Amelogenina	●	●	●		●	●	●	●
Penta D					●			
Penta E				●	●			
D19S433							●	●
D2S1338							●	●

Nota: 13 STRs do CODIS (*Combined DNA Index System*), em verde e laranja; 8 STRs do ENFSI (*European Network of Forensic Science Institutes*), em laranja, em azul gene da amelogenina, em preto *loci* não padronizados (Ruitberg, Reeder e Butler, 2001).

Na escolha de *loci* de STR para estudos forenses ou populacionais, deve-se ter em consideração os seguintes factores: que apresentem um alto nível de variabilidade dentro de cada *locus*, de forma que se verifique uma baixa probabilidade de coincidência; e o comprimento dos alelos deve estar entre 90 e 500 pb (alelos com maior peso molecular tendem a apresentar menor precisão de sua medida); os *loci* podem ser seleccionados com base na localização cromossómica, para garantir que *loci* próximos não sejam seleccionados; os STR devem apresentar robustez e reprodutibilidade dos seus resultados, que são essenciais.

Inicialmente, o *European DNA Profiling Group* (EDNAP) e *European Network of Forensic Science Institutes* (ENFSI) estabeleceram quatro sistemas como padrão, designadamente: HUMTH01, HUMVWFA, HUMD21S11 e HUMFIBRA/FGA. Posteriormente, foram adicionados HUMD3S1358, HUMD8S1179, HUMD18S51 e amelogenina. Os Estados Unidos da América estabeleceram um sistema de padronização, o *Combined DNA Index System* (CODIS), composto de 13 *loci* (Tabela 2.2).

Os *loci* de STRs autossómicos apresentam um elevado grau de polimorfismo genético, cuja base molecular consiste na variação do número de unidades de repetição

(Ferreira *et al.*, 1998; Edwards *et al.*, 1991). O reduzido tamanho dos seus alelos, inferior a 500 pb, permite a sua análise a partir de amostras degradadas ou ínfimas de ADN (Edwards *et al.*, 1991; Alford *et al.*, 1994; Brinkmann, 1996). Estas características fazem-nos marcadores de eleição no campo da genética forense e permitiram superar as limitações dos sistemas de *variable number tandem* repeats repeats (VNTRs).

Entretanto, os STRs tetranucleótideos são os mais utilizados em genética forense devido ao facto de os resultados da amplificação serem facilmente reproduzíveis, fiáveis e de fácil interpretação (Edwards *et al.* 1991). Os STRs dinucleótideos, apesar da fácil amplificação, apresentam múltiplas bandas-artefacto, o que os inviabiliza para fins forenses (Kimpton *et al.*, 1993; Urquhart *et al.*, 1995; Walsh *et al.*, 1996).

A PCR permite a amplificação simultânea de vários *loci* de STRs numa única reacção *multiplex*. Esta técnica é bastante sensível e rápida, pois diminui a quantidade de reagentes e de ADN da amostra necessários para se obter um perfil genético. A utilidade dos STRs deve-se à facilidade de classificação dos alelos em função do número de repetições que aqueles contêm, sendo os seus produtos de amplificação altamente reprodutíveis e rapidamente interpretáveis (Butler, 2001).

Para fins de identificação humana, é indispensável recorrer-se a marcadores de ADN polimórficos que apresentem grande variabilidade e que, no seu conjunto, manifestem a capacidade de discriminação dos indivíduos (Butler, 2001).

2.4.2 – Polimorfismos de STRs do cromossoma Y

Em 1985 Casanova e colaboradores haviam descrito o primeiro marcador polimórfico do cromossoma Y, denominado p122f2, tendo sido posteriormente identificados outros marcadores com diferentes utilizações.

Na década de 90, iniciaram-se os primeiros estudos relativos aos STRs do cromossoma Y, devido ao elevado número de polimorfismos presentes na região não recombinante deste cromossoma (Roewer *et al.*, 1992). Mais tarde foram descritos três STRs do cromossoma Y, um dinucleotídeo (YCAI, YCAII, YCAIII), um trinucleotídeo (27H39 ou DYS19) um pentanucleotídeo (DXYS156Y). Naquela época, apenas o DYS19 tinha aplicação na resolução de casos forenses (Kayser *et al.*, 1997) e em estudos antropológicos (Roewer *et al.*, 1993). A sua aplicabilidade era limitada, devido à falta de outros marcadores.

Os marcadores do cromossoma Y foram descobertos pela seguinte ordem cronológica: DYS19 (Roewer *et al.*, 1992); YCAI, YCAII e DXYS156 (Mathias *et al.*, 1994); DYS389 I/II, DYS390, DYS391, DYS392 e DYS393 (Roewer *et al.*, 1996); DYS288,

DYS388 (Kayser *et al.*, 1997ª); DYS385 (Schneider *et al.*, 1998); A7.1, A7.2, A10, C4, H4 (White *et al.*, 1999); DYS434, DYS435, DYS436, DYS437, DYS438, DYS439 (Ayub *et al.*, 2000); G09411, G10123 (De Kniff, 2000); DYS441e DYS442 (Lida *et al.*, 2001); DYS446, DYS447, DYS448, DYS449, DYS450, DYS452, DYS453, DYS454, DYS455, DYS456, DYS458, DYS459, DYS463, DYS464 (Redd *et al.*, 2002). O conjunto da informação dos STRs de um mesmo sistema no cromossoma Y é designado haplótipo.

Na Fig. 2.12 está representada esquematicamente a distribuição ao longo do cromossoma Y de vários marcadores.

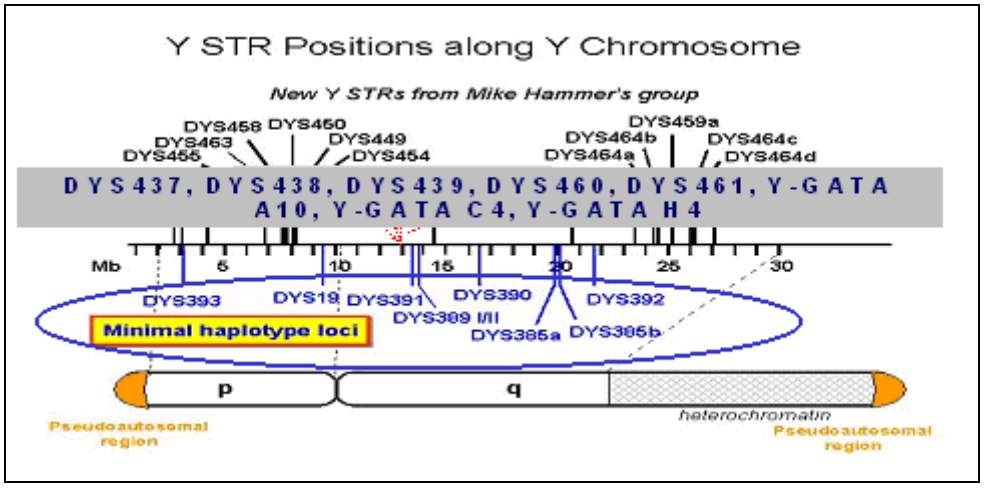


Figura 2.12 – Esquema da distribuição dos marcadores do STR ao longo do cromossoma Y (retirado http://www.cstl.nist.gov/biotech/strbase/images/Y%20STR%20Positions.jpg adaptado).

O haplotipo mínimo foi estabelecido pela Y *Chromosome Haplotype Reference Database* (YHRD), é bastante discriminativo e tem aplicação em investigações forenses (Pascali *et al.*, 1999; Roewer *et al.*, 2001). Autores como Kayser *et al.* (1997) são unânimes em afirmar que o estudo do haplótipo mínimo (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 e DYS385), permite discriminar a maior parte dos indivíduos de uma população (Kayser et al., 1997ª e 2002).

Posteriormente, o *Scientific Working Group – DNA Analysis Methods* (SWGDAM) recomendou a adição dos restantes *loci* (DYS437, DYS438 e DYS439) com a finalidade de aumentar consideravelmente o poder de discriminação do haplotipo mínimo.

O haplotipo extendido é mais informativo e discriminativo, devido ao maior número de STRs do cromossoma Y utilizados para caracterizar um indivíduo ou uma população. Como tal, esta é uma ferramenta poderosa na resolução de casos forenses, tais como nos testes de paternidade e nos crimes de natureza sexual com mistura de materiais biológicos de ambos os sexos (Jobling *et al.*, 1997); Sibille *et al.*, 2002; Cerri *et al.*, 2003).

Os STRs do cromossoma Y consistem em repetições em *tandem*, cujas unidades de repetição variam em sequência e em comprimento (Roewer *et al.*, 1992; Kayser *et al.*, 1997). A seguir descrevem-se os STRs do cromossoma Y que foram usados neste estudo:

DYS19 – é um tetranucleotídeo complexo, localizado no braço curto do cromossoma Y, e apresenta unidade repetitiva $[TAGA]_n$ (Roewer *et al.*, 1992). O seu tamanho varia entre 232 a 268 pb, em função do número de repetições em *tandem*, entre 10 a 19 vezes.

DYS389 I/II – são tetranucleotídeos complexos e apresentam unidades repetitivas $[TCTG]_n$ e $[TCTA]_m$ (Rower *et al.*, 1996). A sequência do sistema DYS389 I está incluída na sequência do DYS389 II. Por este motivo, utiliza-se apenas um par de *primers* para amplificar os dois *loci*, visto que um dos *primers* hibrida em dois locais diferentes na mesma cadeia. Devido a isso, aparecem dois produtos de amplificação, contendo repetições em número variável. Neste caso, os alelos podem ser atribuídos a qualquer um dos *loci*. Os seus alelos apresentam um tamanho que varia de 148-168 pb para DYS389I e 256-296 pb para DYS389 II, respectivamente (Edwards *et al.*, 1991).

DYS390 – trata-se de um tetranucleotídeo complexo que apresenta unidades repetitivas $[TCTG]_n$ $[TCTA]_m$ (Roewer *et al.*, 1996). Os seus alelos apresentam um tamanho que varia de 191-227 pb, em função do número de repetições em *tandem*, entre 18-27 vezes.

DYS391 – trata-se de um tetranucleotídeo que apresenta uma unidade repetitiva [TCTA], os seus alelos apresentam um tamanho que varia de 90-118 pb, em função do número de repetições em *tandem*, entre 8-13 (Edwards *et al.*, 1991; Roewer *et al.* 1996).

DYS392 – trata-se de um trinucleotídeo que apresenta uma unidade repetitiva [TAT], com tamanho que varia de 294-327 pb, em função do número de repetições em *tandem*, entre 7-18 (Edwards *et al.*, 1991).

DYS393 – trata-se de um tetranucleotídeo que apresenta unidade repetitiva $[AGAT]_n$ com tamanho que varia de 104-136 pb, em função do número de repetições em *tandem*, entre 8-16 (Edwards *et al.*, 1991).

DYS385 – trata-se de um tetranucleotídeo que apresenta unidade repetitiva [GAAA], com tamanho que varia de 243-315 pb, em função do número de repetições em *tandem*, com 7-25 (Edwards *et al.*, 1991); o sistema apresenta dois alelos de dois *loci* diferentes que, ao contrário dos outros marcadores, não podem ser atribuídos a nenhum dos *loci*, devido à sobreposição de tamanhos, sendo mais informativo e discriminativo, com uma diversidade haplotípica elevada. Por esta razão, estes alelos são analisados

como haplótipos compostos por dois alelos que podem estar em homozigotia (Edwards *et al.*, 1991; Gill *et al.,* 2001).

DYS437 – trata-se de um tetranucleotídeo complexo com um alto grau de polimorfismo e elevado poder de discriminação que apresenta unidades repetitivas $[TCTA]_n$ $[TCTG]_m$, com tamanho que varia de 183-199 pb em função de número de repetições em *tandem* , com 13-17 (Edwards *et al.*, 1991).

DYS438 – trata-se de um pentanucleotídeo altamente variável, que apresenta unidades repetitivas $[TTTTC]_n$ com tamanho que varia de 101-121 pb, em função de número de repetições em *tandem*, com 8-12 (Edwards *et al.*, 1991).

DYS439 – trata-se de um tetranucleotídeo altamente variável que apresenta unidades repetitivas $[GATA]_n$ com tamanho que varia de 203-231 pb, em função de número de repetições em *tandem*, com 8-15 (Ayub *et al.*, 2000).

A existência de STRs tetraméricos na região não recombinante do cromossoma Y, com grande variabilidade entre os indivíduos, representa uma ferramenta importante de marcadores genéticos (Mathias *et al.*, 1994; Kayser *et al.*, 1997), com aplicação em estudos forenses, genética populacional e estudos evolutivos.

Os STRs do cromossoma Y encontram-se disseminados por todo o cromossoma. Uma das suas características consiste na elevada taxa de mutação ($2.1x10^{-3}$), relacionado ao maior número de divisões apresentadas na formação dos gâmetas (Gusmão *et al.*, 2005).

Roewer *et al.,* (2005) observaram que o elevado nível da diversidade haplotípica nas populações humanas do cromossoma Y está relacionado com a elevada taxa de mutação. Os STRs do cromossoma Y que constituem o haplótipo mínimo apresentam uma taxa de mutação que varia entre os diferentes *loci*, podendo-se estabelecer uma relação entre o tamanho da unidade de repetição e a diversidade dos *loci* (Carvalho-Silva *et al.*, 1999).

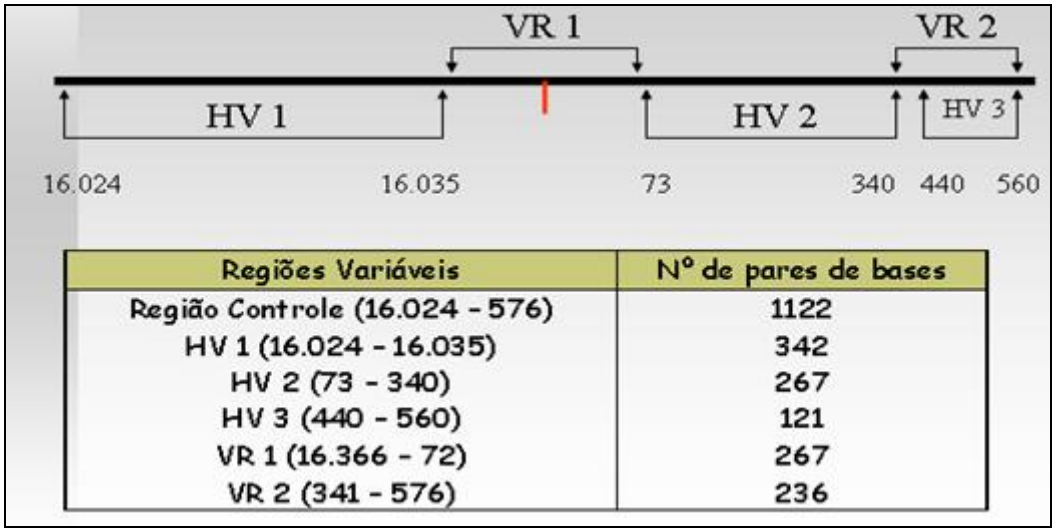
O estudo de vários polimorfismos do cromossoma Y num indivíduo permite construir o seu haplótipo, tendo revelado grande importância tanto na genética de populações, como na genética forense, para além de permitir a reconstrução da história de migrações (Hammond *et al.*, 1992; Gill e Evett, 1995; Tishkoff *et al.*, 2009).

Os haplótipos são transmitidos em bloco de *loci*, inalterados de pai para filhos, de geração em geração. Na ausência de mutações, logo, todos os indivíduos relacionados pela linhagem paterna apresentam o mesmo haplótipo (Budowle *et al.*, 2001). As mutações ocorridas durante a evolução humana geraram variações dos haplótipos que

servem como marcadores de linhagem. Neste caso, o estudo dos polimorfismos do cromossoma Y permite identificar os diferentes haplótipos, o que possibilita reconstruir uma parte da história genética de uma determinada população.

2.4.3 – Polimorfismos do ADN mitocondrial

A região " *displacement loop*" (D-loop) ou região controlo do ADNmt, com aproximadamente 1.112 pb, apresenta grande variabilidade individual. Nela estão localizados os segmentos hipervariáveis HVI, HVII e a região hipervariável menor denominada HVIII (Fig. 2.13) com interesse em aplicação forense e em estudos de Antropologia e Genética das Populações (Bini *et al.*, 2003; Lee *et al.*, 2006).



Regiões Variáveis	Nº de pares de bases
Região Controle (16.024 - 576)	1122
HV 1 (16.024 - 16.035)	342
HV 2 (73 - 340)	267
HV 3 (440 - 560)	121
VR 1 (16.366 - 72)	267
VR 2 (341 - 576)	236

Figura 2.13 – Divisão da região controlo mitocondrial
(retirado:www.seguranca.mt.gov.br/politec/3c/artigos/dna_mitocondrial.doc).

No estudo da região controlo, convencionou-se iniciar a numeração nucleotídica do ADNmt desta região, de maneira que as posições finais sejam 16.024 a 16.569 e as iniciais 1 a 576 do genoma mitocondrial, de acordo com a referência original (Anderson *et al.*, 1981). O segmento HVIII, que se estende da posição 440 à 560), foi caracterizado na década de 90 e, apesar do seu menor polimorfismo, aumenta o poder de discriminação em casos de análise forense (Lutz *et al.*, 1997). As regiões hipervariáveis apresentam grande variabilidade individual e são utilizadas para investigar a diversidade genética numa população e entre diversas populações. Neste caso, assumem grande relevo para os estudos das linhagens maternas idênticas (Lutz *et al.*, 1997; Lutz *et al.*, 2000; Bini *et al.*, 2003).

Na maior parte da sua sequência, o genoma mitocondrial é idêntico entre diferentes indivíduos. Diferenças de sequências são encontradas na região D-*loop* do ADNmt, que apresenta uma alta taxa de mutação, superior que no ADN nuclear (Jarreta, 1999), devido à falta de histonas, que actuam como um "isolador" no ADN nuclear, e de um sistema reparador do ADN eficiente. A taxa de mutação verificada no genoma mitocondrial é mais alta na região D-loop, quando comparada com a região codificante na maioria dos casos (Peric *et al.*, 2005).

As alterações permanentes na sequência de pares de bases do ADNmt podem ocorrer num único nucleótido ou serem mais extensas. Podem dever-se a erros de replicação, despurinação, desaminação e oxidação de bases do ADN, ou serem induzidas por agentes mutagénicos gerados no ambiente intracelular. As mutações pontuais que não originam alterações na qualidade de material genético são as mais frequentes. Dentro das mutações pontuais, contam-se: a mutação por substituição de uma única base, que pode ser de transição quando uma pirimidina (C ou T) é substituída por outra ou uma purina (A ou G) é substituída por outra purina (a mais frequente), ou ainda por transversão quando uma purina é substituída por uma pirimidina ou vice-versa (Belle *et al.*, 2005; Tully *et al.*, 2001). As mutações por deleção ou inserção são as menos frequentes (Tully *et al.*, 2001). As mutações pontuais revelam alterações na molécula do ADNmt, proporcionando polimorfismos.

Durante a divisão celular, as mitocôndrias são distribuídas nas células filhas, que são compostas por um único tipo de ADN, situação designada por homoplasmia. Porém, devido à sua alta taxa de mutação, podem co-existir dentro da mesma célula moléculas normais e moléculas mutantes, o que dá lugar a duas ou mais populações de moléculas de ADNmt num mesmo indivíduo, fenómeno denominado heteroplasmia. A heteroplasmia verifica-se nos segmentos da região controlo, ocorrendo por inserção e/ou deleção em zonas repetitivas (Butler e Levin, 1998; Calloway, 2000). A heteroplasmia celular pode ocorrer por linhagem germinativa ou ser devido a causas somáticas (Calloway, 2000). Um dos factores a ter em consideração, no estudo da frequência da heteroplasmia, é a relação com o factor idade e a associação a um processo patológico.

As regiões hipervariáveis apresentam uma zona rica em citosina designada por homopoliméricas poli-C – processo de replicação que leva à inserção de um número de citosinas superior na zona de poli-citosinas das regiões HVI e HVII (Calloway, 2000; Longley *et al.*, 2001).

A taxa de mutação verificada no ADNmt apresenta uma frequência de 5-10 vezes superior à registada no ADN nuclear (Brown *et al.*, 1982 e Wallace *et al.*, 1987), ocorrendo aleatoriamente e podendo ser influenciada por diversos factores. O ADNmt

não surge associado a histonas pois, encontrase sob influência dos radicais livres resultantes do processo respiratório que são potencialmente mutagénicos (Richter, 1995). A baixa eficácia de reparação da polimerase do ADNmt constitui um dos factores atribuídos ao aumento da frequência elevada de incorporação errada de nucleótidos (Richter, 1995; Jonhson e Johnson, 2001). A invariabilidade de uma mutação numa determinada população está dependente dos efeitos fenotípicos e da ineficácia da sua eliminação. Em Genética das Populações, as mutações neutras são as mais interessantes devido à acção da selecção natural. Portanto, na região controlo, a mutação apresenta ainda *hot spots* mutacionais devido à hipermutabilidade desta zona.

O polimorfismo destas regiões apresenta grande utilidade para a identificação genética individual (Seo *et al.*, 1998). O estudo da região codificante do ADNmt revelou a existência de outros polimorfismos associados a determinadas patologias ligadas aos genes. Apesar da sua grande utilidade em estudos de identificação individual e em populações, a análise dos polimorfismos do ADNmt apresenta algumas desvantagens, pois, quando comparados aos do marcador nuclear, revelam ser menos informativos, devido ao menor poder de discriminação entre os indivíduos.

O ADNmt é herdado como um único bloco, herança uniparental materna, pois os indivíduos relacionados pela linhagem materna partilham a mesma sequência de ADNmt – haplotipo que, ao serem agrupados, definem haplogrupos (Maca-Meyer *et al.*, 2001). As mutações pontuais acumuladas ao longo das várias gerações representam a base dos diferentes polimorfismos no ADNmt e as suas frequências permitem caracterizar populações em diferentes regiões geográficas (Cavalli-Sforza *et al*., 1994).

A sequenciação do ADNmt é uma técnica com múltiplas aplicações na área da genética forense, nomeadamente na identificação genética individual, em estudos de evolução humana, antropológicos ou arqueológicos. A PCR constitui a técnica de eleição para a amplificação das moléculas de ADNmt, cujo sucesso se deve à sensibilidade da mesma na amplificação de amostras com material degradado e pouco ADN nuclear, tais como dentes e ossos, ou com ínfimas quantidades de ADN (Budowle *et al.*, 2003; Parson e Bandelt, 2007). O maior número de moléculas de ADNmt numa célula diminui o risco de degradação.

Uma das desvantagens na aplicação do ADNmt reside no risco de obtenção de resultados errados, devido a uma maior susceptibilidade de contaminação das amostras biológicas. A ocorrência de heteroplasmia tem dificultado a interpretação dos dados e pareceres consensuais quanto à valorização estatística dos resultados.

A análise dos resultados é feita por comparação da sequência obtida com a sequência de referência, sendo que cada sequência de ADNmt de um indivíduo

corresponde a um haplótipo. As sequências da região controlo podem ser facilmente analisadas e utilizadas para serem comparadas com o perfil das amostras de diferentes indivíduos. As regiões hipervariáveis têm sido amplamente utilizadas para investigar a diversidade dentro de e entre populações.

2.4.4 – Aplicação forense dos polimorfismos do ADN

O estudo dos polimorfismos do ADN converteu-se numa ferramenta imprescindível, de grande importância em áreas como a Genética e Biologia Forenses, a Genética Populacional, Evolução Humana, Antropologia e outras.

Os STRs autossómicos são herdados de ambos os progenitores. A sua aplicação em genética forense consiste principalmente na resolução de casos de identificação individual, estabelecimento de parentescos (vínculos genéticos) e na investigação biológica criminal.

No entanto, os STRs do cromossoma Y têm grande interesse também no estabelecimento do vínculo de filiação, na ausência ou falecimento do progenitor (Rolf *et al.*, 2001), assim como na investigação de identidade e determinação do sexo em desastres em massa e na investigação de pessoas desaparecidas (Schults e Herrmann, 1999; Butler *et al.*, 2003; Jobling, 2003).

Na sua maioria, os crimes violentos são perpetrados por indivíduos do sexo masculino e as amostras biológicas encontradas no local do crime podem fornecer informações relevantes, se forem analisadas com os STRs do cromossoma Y (Jobling *et al.*, 1997). As características peculiares do cromossoma Y fazem dele uma ferramenta indispensável e relevante na investigação dos crimes de natureza sexual, tanto hetero como homossexual (Prinz e Sansone, 2001). Contudo, o estudo do cromossoma Y apresenta algumas limitações, pelo facto de não identificar um indivíduo, visto que todos os indivíduos da mesma linhagem paterna partilham o mesmo haplotipo, o que, em muitas ocasiões, dificulta as investigações. Portanto, apesar da sua enorme utilidade em algumas situações complexas, o estudo dos marcadores do cromossoma Y deve ser efectuado em conjunto com os marcadores STRs autossómicos.

Em Genética Forense, o estudo do ADNmt aplica-se naqueles casos em que existe reduzida quantidade de material genético para identificação de restos humanos degradados ou em casos de acidentes de aviação e outros desastres em massa.

Tendo em consideração as suas características específicas, o ADNmt permite a utilização de vários tipos de tecidos, como cabelos, ossos e dentes que, dependendo das circunstâncias, contêm pouco ADN nuclear (Vigilant *et al.*, 1989; Wilson *et al*., 1995; Budowle *et al.*, 2003; Parson e Bandelt, 2007).

Em regra, o estudo do ADNmt aplica-se naqueles casos em que a análise dos polimorfismos do ADN nuclear não é possível para a resolução de um determinado caso ou quando se pretende uma informação adicional.

2.4.5 – Marcadores de linhagens em Genética populacional

O processo de recombinação dos genes nos cromossomas autossómicos verificado a cada geração, gera uma mistura genética de todos os antepassados, o que dificulta a sua utilização em estudos evolutivos.

Os marcadores genéticos uniparentais são legados de um dos pais para os filhos e estão localizados no cromossoma Y e no ADNmt, nomeadamente. Estes marcadores representam uma ferramenta relevante em estudos da evolução humana, estudos antropológicos e populacionais (Butler *et al.*, 2003; Jobling e Tyler-Smith, 2003; Schults *et al*., 1999), e fornecem informações que permitem traçar o perfil de patrilinhagens, que revela a genealogia paterna e as relações evolutivas entre diferentes grupos de indivíduos (Bradman e Thomas, 1998; Hammer, 1995).

Os indivíduos que compartilham o mesmo haplótipo fazem-no por ancestralidade comum; assim sendo, o estudo do cromossoma Y possibilita compreender a história da linhagem patrilinear e eventos mutacionais que aconteceram ao longo da evolução em toda a linhagem (Mitchell e Hammer, 1996). A análise dos haplótipos permite compreender a evolução do cromossoma Y no passado e revelar padrões geográficos na sua distribuição, permitindo a detecção de conjuntos característicos de haplótipos específicos de determinadas populações (Mitchell e Hammer, 1996). São principalmente relevantes em estudos das populações, pois auxiliam na compreensão da história evolutiva e na identificação humana (Hammer e Zegura, 1996; Jobling e Tyler-Smith, 1995; Jobling *et al.*, 1997; Underhill *et al.*, 2001; Hammer *et al.*, 2001) de várias gerações no passado, o que permite reconstruir a história genética de uma população ou de um povo (Jobling e Tyler-Smith 2003).

Os haplótipos do ADNmt são legados de mãe para todos os filhos, sem modificação (Budowle e Brown, 2001; Alvarez *et al.*, 2001). Os seus haplogrupos caracterizam grupos humanos ou grupos étnicos com características específicas, povos

ou mesmo continentes. Os haplogrupos africanos apresentam maior variação, bem como uma raiz mais profunda na árvore filogenética. A árvore filogenética da espécie humana consiste nos seguintes haplogrupos: na África subsahariana, na sua totalidade, predominam haplogrupos do ADNmt pertencentes aos haplogrupos L0, L1, L2, L3, L4, L5 e M1 (Chen *et al.*, 1995; Chen *et al.*, 2000; Salas *et al.*, 2002; Kivisild *et al.*, 2002). O haplogrupo L0 abarca subhaplogrupos L0a, L0d, L0f e L0k, que foram classificados como subgrupos de L1 (Watson *et al.*, 1997; Salas *et al.*, 2002), sendo o subgrupo L0a o mais frequente. Pereira *et al.* (2001) e Rosa *et al.* (2004) referiram que o haplogrupo L0 não apresentava uma distribuição uniforme entre as populações africanas, sendo predominante em Moçambique e estando ausente noutras regiões de África. Os subhaplogrupos L0d e L0k são específicos das populações africanas Khoisan (Chen *et al.*, 2000; Salas *et al.*, 2002). O haplogrupo L1 encontra-se subdividido nos macrohaplogrupos L1b e L1c (Salas *et al.*, 2002). Enquanto o L1b é predominante no oeste de África, o L1c é mais frequente na África central e, com pequenas variações, desde o oeste ao sudoeste de África.

O haplogrupo L2 encontra-se subdividido nos subhaplogrupos L2a, L2b, L2c e L2d, sendo o L2a o mais representativo em quase todas as regiões de África. O haplogrupo L3 tem a sua origem no este de África e encontra-se subdividido nos subhaplogrupos L3b, L3d, L3e, L3f, L3h, L3i, L3w e L3x. O subhaplogrupo L3e é o mais frequente e com maior dispersão geográfica, representando um terço do haplogrupo L3 na África subsahariana (Kivisild *et al.*, 2002).

Os haplogrupos M e N da Euroásia derivaram do haplogrupo L3 (Quintana-Murci *et al.*, 1999). O ADNmt europeu apresenta uma variabilidade de haplogrupos caracterizados pelas linhagens H, I, J, K, M, T, U, V, W e X. Estes haplogrupos foram derivados inicialmente do macrohaplogrupo N (Torroni *et al.*, 1996; Mishmar *et al.,* 2003). Dos haplogrupos característicos das populações europeias, o haplogrupo H é o mais representativo.

Os haplogrupos C, D, E, G, Z, derivados do macrohaplogrupo M, assim como os haplogrupos A, B, F e Y, derivados do macrohaplogrupo N, são mais comuns no leste e sul asiático (Kivisild *et al.*, 2002). Na América ocorrem com maior frequência os haplogrupos A, B, C, D e X (Bolnick e Smith, 2003).

Os haplogrupos reflectem a partilha de um ancestral comum de ADNmt e podem ser utilizados para estimar a proporção de mistura em indivíduos que habitam em conhecidas rotas de migração, sendo importante calcular as distâncias genéticas entre as populações e criar árvores filogenéticas que as relacionem, para compreendermos melhor os processos migratórios do passado. Também são úteis para estabelecer os

níveis de diversidade das populações não africanas, que dependem do nível de amplitude do efeito fundador, e os efeitos *bottleneck* que ocorreram durante a colonização do Novo Mundo. Quando uma população migra, leva todos os seus haplogrupos; a idade do haplogrupo nas migrações populacionais indica quando a mutação pontual que define esse haplogrupo ocorreu e não quando ocorreu a migração (Pakendorf e Stoneking, 2005).

3. Material e Métodos

Amostragem populacional

Extracção e amplificação

Detecção automática de fragmentos e polimorfismos

Purificação e sequenciação

Detecção automática de fragmentos de sequenciação

Verificação dos polimorfismos de sequência

Análise estatística

3.1 – Amostragem populacional

Na análise dos STRs autossómicos, foram consideradas 479 amostras biológicas, sendo 212 de indivíduos do sexo masculino e os restantes do sexo feminino, pertencentes aos 7 principais grupos étnico-linguísticos de Angola (102 amostras do grupo Bakongo, 100 amostras do grupo Kimbundo, 50 amostras do grupo Kwanhama, 50 amostras do grupo Lunda-Tchokwe, 50 amostras do grupo Nganguela, 25 amostras do grupo Nhaneca-Humbe e 102 amostras no grupo Ovimbundo).

No estudo dos STRs do cromossoma Y, foram seleccionadas e analisadas 166 amostras dos três principais grupos étnico-linguísticos (57 do grupo Bakongo -ABK, 56 do grupo Kimbundo-AKI e 53 do grupo Ovimbundo-AOV).

Para o estudo do ADN mitocondrial, foram seleccionadas e analisadas 30 amostras biológicas, sendo 10 de cada um dos 3 grupos étnico-linguísticos (Bakongo - ABK, Kimbundo - AKI e Ovimbundo - OVI).

3.2 – Colheitas de amostras sanguíneas

A origem e o grupo étnico-linguístico dos indivíduos intervenientes, bem como dos seus pais e avós, foram confirmados por inquérito, no momento da colheita das amostras. Estas foram colhidas após o consentimento informado dos intervenientes e o preenchimento de uma ficha de identificação das amostras. Os indivíduos foram considerados como pertencentes a um determinado grupo étnico-linguístico, todos aqueles nascidos na região dos seus progenitores ou descendentes destes.

Foi feita a colheita das amostras sanguíneas de indivíduos saudáveis, nos 7 principais grupos étnico-linguísticos (Bakongo, Kimbundo, Kwanhamas, Lunda-Tchokwe, Nganguela, Nhaneca-Humbe e Ovimbundo) da população de Angola, sob a forma de mancha de sangue em papel.

As amostras biológicas (sangue periférico) foram colhidas por punção venosa ou por picada no dedo, com lanceta, em 479 indivíduos saudáveis e voluntários.

As amostras foram recolhidas aleatoriamente nos indivíduos intervenientes, excluindo-se os familiares consanguíneos.

As manchas sanguíneas, com cerca de um centímetro de diâmetro, foram efectuadas em papel absorvente, sem adição de qualquer substância química; posteriormente, procedeu-se à secagem das manchas de sangue, por exposição ao ar livre, evitando a irradiação solar.

Foram preenchidas fichas de identificação anónima dos indivíduos intervenientes, nas quais foram solicitados dados como local de nascimento, grupo étnico linguístico a que pertence, assim como dados dos pais e dos seus avós paternos e maternos (anexo I).

Todas as amostras foram colhidas no território angolano durante o ano de 2004, onde foram identificadas, agrupadas de acordo com o grupo de origem e depois transportadas para o Serviço de Biologia e Genética Forense, no Instituto Nacional de Medicina Legal – Delegação do Centro, para o Serviço de Genética e Biologia Forense, onde foram arquivadas e conservadas à temperatura ambiente.

3.3 – Extracção do ADN

O ADN foi extraído das manchas sanguíneas segundo o protocolo de extracção (Anexo II), utilizando o Chelex 100 (Walsh *et al*., 1991). O Chelex é um composto de co-polímeros de estireno-divinilbenzeno, com grande afinidade para iões polivalentes, prevenindo assim a degradação do ADN em presença de iões metálicos a alta temperaturas e em condições de baixa força iónica. Os procedimentos de extracção do ADN pelo método de Chelex® encontra-*se* em anexo II.

3.4 – Técnicas utilizadas no estudo do ADN autossómico

3.4.1 – Amplificação de ADN autossómico

Em 1986, Kary Mullis apresentou uma nova tecnologia de amplificação de sequências específicas de ADN, designada por *polymerase chain reaction*, *PCR*. A PCR é um processo enzimático, através do qual se obtêm múltiplas cópias de uma sequência específica do ADN, em apenas algumas horas (Mullis *et al*., 1986). Os procedimentos de amplificação encontra-se no anexo III.

O processo de amplificação de pequenos segmentos do ADN começa com uma reacção de amplificação, que inclui a amostra de ADN, uma enzima termoestável - Taq polimerase (Lodish *et al*., 1995), dois *primers* (oligonucleotídeos), nucleótidos (dNTPs), tampão de reacção e cloreto de magnésio. A reacção tem lugar no amplificador de PCR. Neste, as amostras passam por ciclos de diferentes temperaturas, em períodos diferentes de tempo, formando assim o ciclo de amplificação que multiplica a sequência - alvo de ADN da amostra.

Para a reacção de PCR e utilizada uma região específica do ADN e as sequências das extremidades desta região devem ser conhecidas. Os *primers* são complementares destas sequências que, ao hibridarem com o ADN, fazem com que a região a amplificar seja delimitada.

A PCR é um processo cíclico, composto de três etapas, que são a desnaturação do ADN, o acoplamento dos *primers* e a extensão com os dNTPs. Todas elas requerendo uma série de parâmetros e optimização. O ciclo de temperaturas é o seguinte:

● *Desnaturação:* nesta etapa ocorre a desnaturação da dupla cadeia de ADN, a altas temperaturas (90-95º C), dando assim origem a duas cadeias simples de ADN;

● *Acoplamento ou emparelhamento:* após a desnaturação da cadeia, a temperatura baixa até aos 40-60º C, durante aproximadamente 30-60 segundos, permitindo que os dois *primers* complementares da sequência - alvo se liguem à cadeia simples de ADN na terminação 3´;

● *Extensão:* nesta etapa começa a síntese do novo ADN, quando a temperatura da reacção aumenta até aos 72º C, permitindo à enzima *Taq* polimerase adicionar os (dNTPs) à cadeia de ADN e copiar a região - alvo.

Para a reacção de polimerização em cadeia, para o sistema *multiplex AmpFLSTR® Identifiler®*, as amostras forma colocadas no termociclador 2700 e submetidas ao seguinte ciclo de temperaturas: desnaturação inicial de 95º C durante 11 minutos; 28 ciclos de: 94º C durante 1 minuto (desnaturação dos iniciadores), 59º C durante 1 minuto (emparelhamento dos iniciadores), 72º C durante 1 minuto (extensão dos iniciadores e actuação da enzima taq); seguido de extensão final durante 60 minutos a 60º C após o qual a reacção foi interrompida por esfriamento a 4º C.

Em cada ciclo, no final destas três etapas, obtêm-se duas moléculas de ADN de cadeia dupla, que servirão de moldes para os próximos ciclos. Como é óbvio, para o manuseamento desta técnica é indispensável ter conhecimento da sequência de base do ADN da região alvo.

Neste trabalho, o procedimento realizado esteve de acordo com o protocolo referenciado no manual *AmpℓSTR® Identifiler™ PCR Amplification* Kit User´s Manual (*PE, Applied Biosystems*, 1997).

O *kit AmpFLSTR® Identifiler® PCR Amplification* amplifica simultaneamente os 15 *loci* STRs, para além da amelogenina para determinar o género (Tabela 3.1). Os *primers* estão marcados com fluorocromos que marcam os STRs com cores, facilitando assim a sua distinção e a sua análise. Existem *primers* marcados com 6-FAM *dye* de cor azul

(D8S1179, D21S11, D7S820, CSF1PO), com VIC dye-*primers* de cor verde (D3S1358, TH01, D3S317, D16S539, D2S1338), com NED™ dye-*primers* de cor amarela (D19S433, VWA, TPOX, D18S51) e PET™ dye-*primers* de cor vermelha (amelogenina, D5S818, FGA) (http://www.appliedbiosystems.com/).

Na Tabela 3.1, podem observar-se as características dos STRs autossómicos, presentes no sistema *Identifiler* (Applied Biosystems), utilizados neste estudo. (ABI, *AmpFlSTR® Identifiler™*):

Tabela 3.1 – Caracterização dos *loci* de STRs presentes no sistema *AmpFLSTR® Identifiler®.*

DESIGNAÇÃO DO LOCUS	LOCALIZAÇÃO CROMOSSÓMICA	UNIDADE DE REPETIÇÃO NO FORMATO ISFH	CLASSIFICAÇÃO	REPETIÇÕES	GENBANK
D8S1179	8q24.1-24.2	[TCTA] [TCTG]	COMPOSTO	13	AF216671
D21S11	21q21	[TCTA] [TCTG]	COMPLEXO	29	AP000433
D7S820	7q11.21-22	[GATA]	SIMPLES	13	AC004848
CSF1PO	5q33.3-34	[AGAT]	SIMPLES	12	X14720
D3S1358	3p	[TCTG] [TCTA]	COMPOSTO	18	Não disponível
THO1	11p15.5	[TCAT]	SIMPLES	9	D00269
D13S317	13q22-31	[TATC]	SIMPLES	11	AL353628
D16S539	16q24-qter	[GATA]	SIMPLES	11	AC024591
D2S1338	2q35-37.1	[TGCC] [TTCC]		20	AC010136
D19S433	19q12-13.1	[AAGG]		16	AC0080507
VWA	12p12-pter	[TCTG], [TCTA] [TCCA]	COMPOSTO	18	M25858
TPOX	2p23-pter	[AATG]	SIMPLES	11	M68651
D18S51	18q21.3	[AGAA]	SIMPLES	18	AP001534
Amelogenina	X, Y				
D5S818	5q21-31	[AGAT]	SIMPLES	11	AC008512
FGA	4q28	[CTTT]	COMPOSTO	21	M64982

3.4.2 – Detecção dos polimorfismos de STRs autossómicos

O equipamento de electroforese capilar utilizado para a análise dos fragmentos foi o sequenciador automático *ABI PRISM 310 Genetic Analyser*. O sequenciador é composto por um único capilar e está optimizado para suportar várias corridas de sequenciação e análise de fragmentos de 48 a 96 tubos de amostras. O *ABI Prism® 310 Genetic Analyser* está equipado com o software *ABI Prism® 310 Genetic Analyser Data Collection*, que controla as condições e todas as operações relacionadas com a corrida e transformação das emissões de fluorescência detectadas na câmara CCD em

electroforegramas (http://www.appliedbiosystems.com/). Os procedimentos de detecção dos polimorfismos de STRs autossómicos se encontram em anexo IV.

Às amostras tem que ser adicionada uma solução desnaturante, que facilita a desunião das ligações de hidrogénio entre as cadeias complementares dos produtos da PCR, e devem ser aquecidas e desnaturadas com o objectivo de se separarem as duas cadeias de cada produto de PCR e depois introduzidas no sistema para análise (anexo III).

Os produtos de amplificação, marcados com fluorocromos, são separados por electroforese capilar através do instrumento *ABI PRISM® 310 GENETIC ANALYSER utilizando Polimero POP4 e detectados por* (emissões) de fluorescência, com o *software data collection* apresentando-se sob a forma de "picos" que representam os vários alelos amplificados a partir da amostra de ADN.

Com a aplicação do software *GeneScan® Analysis*, os fragmentos de ADN são identificados, através da atribuição do tamanho de acordo com o padrão interno LIZ-500.

A representação esquemática dos *loci de STRs* amplificados no sistema *AmpFℓSTR® Identifiler* ™ assim como a distinção por cores e tamanho em pb estão representados na Fig. 3.1.

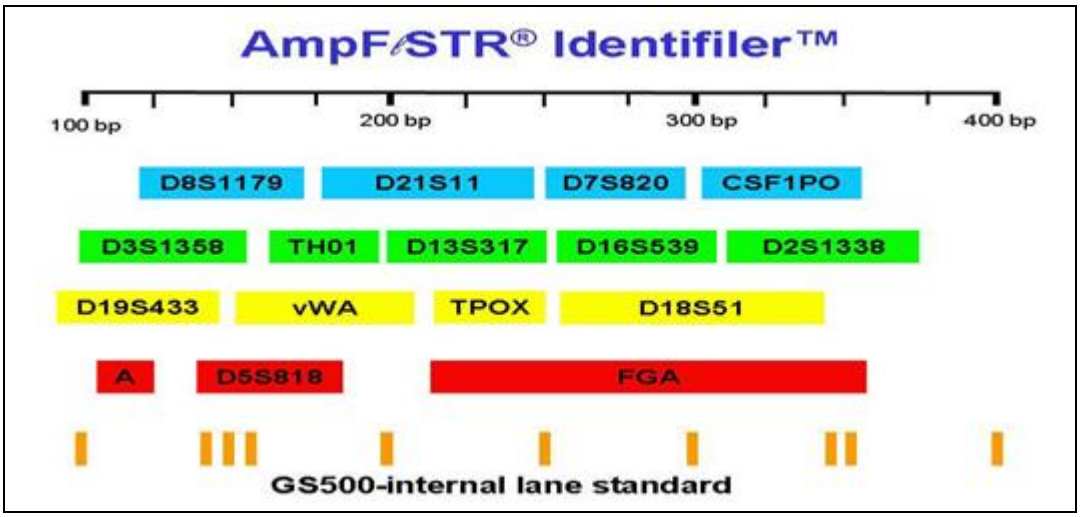


Figura 3.1 – Representação esquemática dos *loci STRs* amplificados no sistema *AmpFℓSTR® Identifiler* ™; na posição vertical, a separação por cores; na posição horizontal, a separação por tamanho em pb (adaptado de Ruitberg, 2001).
(retirado: http://www.cstl.nist.gov/strbase/kits/Identifiler.htm).

Após a determinação do tamanho dos fragmentos dos produtos de PCR da amostra, estes são comparados com os fragmentos de um *ladder* alélico, para atribuição do alelo que lhes corresponde.

3.4.3 – Designação alélica dos polimorfismos de STRs autossómicos

A designação alélica foi efectuada por comparação com o *ladder* alélico. Um *ladder* é o conjunto dos alelos mais comuns para os marcadores em estudo, que vão correr em conjunto com as amostras. Após a obtenção dos electroforegramas das amostras com tamanhos dos "picos" de fluorescência, em termos de números de pares de bases, atribuídos pelo computador, procede-se à atribuição alélica por comparação com o tamanho dos "picos" (em pares de bases) presentes no *ladder* (Tabela 3.2). O valor do "pico" (número de pares de bases) tem de ser encontrado num intervalo de confiança de 5 décimas, em relação ao valor do *ladder*.

Tabela 3.2 – Principais características dos *loci* analisados pelo sistema *AmpFLSTR® Identifiler®*.

Designação do *locus*	**Localização no cromossoma**	**Alelos incluídos no *ladder* AmpFLSTR® *IDENTIFILER®***
D8S1179	8q24.1-24.2	8,9,10,11,12,13,14,15,16,17, 18,19
D21S11	21q11.2-q21	24,24.2,25,26,27,28,28.2,29,29.2,30,30.2,31,31.2,32,32.2,33,33.2, 34,34.2,35,35.2,36,37,38
D7S820	7q11.2-22	6,7,8,9,10,11,12,13,14,15
CSF1PO	5q33.3-34	6,7,8,9,10,11,12,13,14,15
D3S1358	3p	12,13,14,15,16,17,18,19
THO1	11p15.5	4,5,6,7,8,9,9.3,10,11,13.3
D13S317	13q22-31	8,9,10,11,12,13,14,15
D16S539	16q24-qter	5,8,9,10,11,12,13,14,15
D2S1338	2q35-37.1	15,16,17,18,19,20,21,21,23,24,25,26,27,28
D19S433	19q2-13.1	9,10,11,12,12.2,13,13.2,14, 14.2,15,15.2,16,16.2,17,17.2
vWA	12p2-pter	11,12,13,14,15,16,17,18,19, 20,21,22,23,24
TPOX	2p23-2per	6,7,8,9,10,11,12,13
D18S51	18q21.3	7,8,9,10,10.2,11,12,13,13.2,14,14.2,15,16,17,18,19,20,21, 22,23,24,25,26,27
Amelogenina	X: P22.1-22.3 Y: p11.2	X Y
D5S818	5q21-31	7,8,9,10,11,12,13,14,15,16
FGA	4q28	17,18,19,20,21,22,23,24,25,26,26.2,27,28,29,30,30.2,31.2,32.2, 33.2,42.2,43.2,44.2,45.2,46.2,47.2,48.2,50.2,51.2

3.5 – Técnicas utilizadas no estudo dos STRs do cromossoma Y

Na Tabela 3.3 estão referenciadas as características do STRs do cromossoma Y do *kit PowerPlex® Y System*.

No presente estudo, foi utilizado o haplótipo extendido composto pelo haplotipo mínimo (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 e DYS385), mais os marcadores DYS437, DYS438 e DYS439 do kit *PowerPlex® Y System*.

Tabela 3.3 – principais características dos *loci* de STRs do cromossoma Y analisados

Locus STRs	**Número de acesso ao GenBank**	**Estrutura da unidade repetida**	**Número de unidades repetidas**	**Tamanho dos alelos em pb**	**Número de alelos observados**
DYS19	X77751	[TAGA] n	10-19	232-268	14
DYS389I	AF140635	[TCTG] n [TCTA] n	10-15	148-168	13
DYS389II		[TCTG] n [TCTA] n	24-34	256-296	31
DYS390	AC011289	[TCTG] n [TCTA] n	18-27	191-227	24
DYS391	G09613	[TCTA] n	6, 8-13	90-118	10
DYS392	G09867	[TAT n	7-18	294-327	13
DYS393	G09601	[AGAT] n	8-16	104-136	13
DYS385	Z93950	[GAAA n	7-25	243-315	11-14
DYS437	AC002992	[TCTA] [TCTG] n	13-17	183-199	15
DYS438	AC002531	[TTTTC] n	8-12	101-121	11
DYS439	AC002992	[GATA] n	8-15	203-231	12

3.5.1 – Amplificação de STRs do cromossoma Y

A PCR permite que várias regiões do ADN sejam copiadas simultaneamente pela simples adição de mais de um par de *primers* à mistura de reacção, que assim é denominada PCR em *multiplex*. O *kit PowerPlex® Y PCR* amplifica simultaneamente os seguintes marcadores STRs do cromossoma Y: DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS385, DYS437, DYS438 e DYS439 (procedimentos de amplificação no anexo III).

Em cada reacção de amplificação, é fundamental a utilização de controlo positivo (PCR+) e controlo negativo (PCR-). O controlo consiste na adição de apenas água, em vez de ADN.

Os *primers* estão marcados com flurocromos que marcam os diferentes STRs do cromossoma Y, facilitando assim a sua distinção e análise (Fig. 3.2).

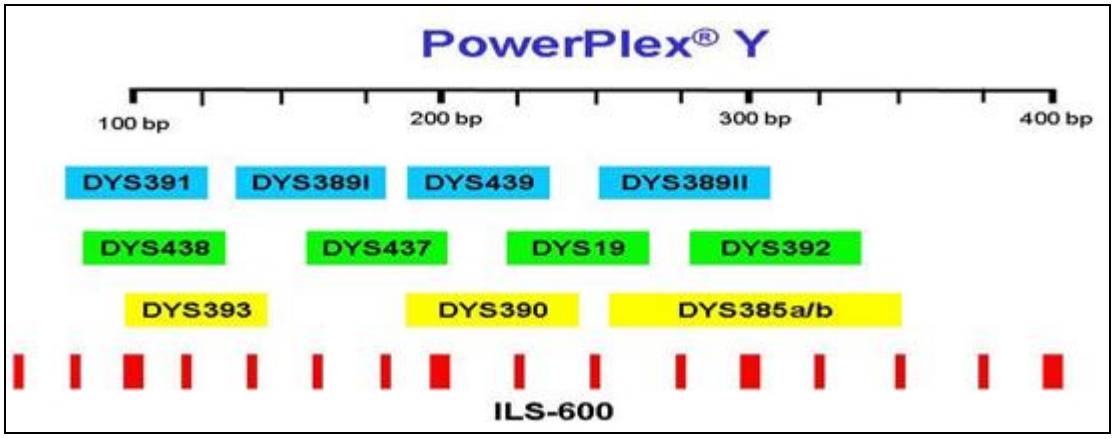


Figura 3.2 – Representação esquemática dos *loci* do cromossoma Y amplificados no sistema PowerPlex. (retirado: http://www.cstl.nist.gov/strbase/kits/PowerPlexY.htm).

Para a amplificação, as amostras forma colocadas no termociclador 9600 e submetidas ao seguinte ciclo de temperaturas: desnaturação inicial de 95º C, durante 11 minutos, e 96º C, durante 1 minuto; 10 ciclos de 94º C durante 30 segundos (desnaturação dos iniciadores), 60º C durante 30 segundos (emparelhamento dos iniciadores) e 70º C durante 45 segundos (extensão); seguem-se 20 ciclos de: 90º C durante 30 segundos (desnaturação), 58º C durante 30 segundos (hibridação) e 70º C durante 45 segundos (extensão), seguido de extensão final durante 10 minutos a 60º C, após o qual a reacção foi interrompida por esfriamento a 4ºC.

3.5.2 – Detecção dos polimorfismos de STRs do cromossoma Y

O equipamento de electroforese capilar utilizado para a análise dos fragmentos foi o sequenciador automático *ABI PRISM 310 Genetic Analyser*. Os procedimentos de aplicação e detecção dos polimorfismos de STRs do cromossoma Y são similares aos utilizados na detecção do STRs autossómicos. Os procedimentos de detecção dos polimorfismos de STRs do cromossoma Y se encontram em anexo IV.

Os fragmentos de ADN do cromossoma Y são identificados através da atribuição do tamanho, de acordo com o padrão interno ILS-600 recorrendo ao software *GeneScan® Analysis*.

Após a determinação do tamanho dos fragmentos dos produtos de PCR da amostra, estes são comparados com os fragmentos dum *ladder* alélico, para atribuição do alelo que lhes corresponde.

3.5.3 – Designação alélica dos polimorfismos de STRs do cromossoma Y

A designação alélica foi efectuada por comparação com o *ladder* alélico fornecido com o respectivo kit. A metodologia da designação alélica dos polimorfismos do cromossoma Y é similar à aplicada na designação alélica dos STRs autossómicos.

3.6 – Técnicas utilizadas no estudo do ADN mitocondrial

Neste trabalho, foi analisada a região controlo total do ADNmt, para o estudo de diversidade génica e haplotípica na população de Angola. Foram analisadas as posições 16.024 a 16.569 e as posições 1 a 576 da região controlo total do ADNmt.

3.6.1 – Amplificação da região controlo

A região controlo, com cerca de 1112 pb, é constituída pelas regiões hipervariáveis I, II e III; foi amplificada por PCR, utilizando dois pares de *primers* específicos da região controlo que se ligam à extremidade 5´ da cadeia molde, conforme o anexo V.

Nas 30 amostras referentes aos três grupos étnico-linguísticos, foi analisada a região controlo do ADNmt. Para a sua amplificação, foram utilizados os *primers com a seguinte constituição*: o L15971 (5´-TTA ACT CCA CCA TTA GCA CC-3´), o L16450 (5´-GCT CCG GGC CCA TAA CAC TTG-3´); o H017 (5´-CCC GTG AGT GGT TAA TAG GGT-3´) e o H599 (5´-TTG AGG AGG TAA GCT ACA TA-3´). Os *primers* foram produzidos de maneira a obter 2 fragmentos de cerca de 600 pb, para permitir a sequenciação de todas as posições.

Os procedimentos de amplificação do ADNmt encontram-se detalhados no anexo V.

Para a reacção de polimerização em cadeia, as amostras foram colocadas no termociclador 9700 (AB Applied Biosystems) e submetidas ao seguinte ciclo de temperaturas: desnaturação inicial de 95º C, durante 15 minutos; 35 ciclos de 94º C, durante 30 segundos (desnaturação), 58º C durante 90 segundos (emparelhamento dos iniciadores) e 72º C durante 60 segundos (extensão); seguido de extensão final durante 10 minutos a 72º C, após o qual a reacção foi interrompida por esfriamento a 4º C.

Como regra, foi sempre usado, em todas as reacções de amplificação, um controlo negativo, em que a amostra foi substituída por água desionizada esterilizada.

A estratégia adoptada de amplificação do ADNmt pares de primers mix associado a BigDey/Better/X Terminator (Lee *et al.*, 2008) permite reduzir o tempo de

procedimentos, melhorar a qualidade, a quantidade de dados e reduzir significativamente o custo por reacção.

3.6.2 - Purificação dos fragmentos amplificados por ExoSap- IT

Os produtos da PCR foram purificados com ExoSap-IT, para remover os primers e dNTPs não incorporados durante o processo de amplificação. Os procedimentos de purificação dos fragmentos amplificados estão descritos no anexo VI.

3.6.3 – Sequenciação cíclica (directos/reversos) dos produtos amplificados

A sequenciação da região controlo directa/reversa foi feita segundo o protocolo da *ABI PRISM BigDye Terminator Cycle Sequence V3.1 (Applied Biosystems*). Quatro marcadores fluorescentes são unidos a cada ddNTP diferente: o ddTTP é marcado com dicloro [ROX], o ddCTP é marcado com dicloro [TAMRA], o ddATP é marcado com dicloro [R6G], e o ddGTP é marcado com dicloro [R110].

Os procedimentos desta metodologia encontram-se descritos no anexo VII.

O programa de sequenciação da região controlo directos/reversos foi efectuado no termociclador 9700, de acordo com as seguintes condições: desnaturação inicial a 96º C durante 2 minutos; 35 ciclos, para as cadeias directas, e 38 ciclos, para as cadeias reversas, de desnaturação a 96º C durante 15 segundos, emparelhamento a 50º C, durante 9 segundos, e extensão a 60º C, durante 2 minutos; extensão final a 60º C, durante 10 minutos, após o que a reacção foi interrompida por esfriamento a 4º C.

3.6.4 – Purificação dos fragmentos sequenciados por *X Terminator*

Após a conclusão da reacção de sequenciação, as amostras foram submetidas a uma nova etapa de purificação, com a finalidade de remover terminadores não incorporados do BigDye e sais desnecessários. Os procedimentos realizados encontram-se descritos em detalhe no anexo VIII.

3.6.5 – Detecção do produto sequenciado

As amostras preparadas foram colocadas no aparelho 3130 *Applied Biosystems*, que processa 4 amostras simultaneamente, tendo sido definidos todos os parâmetros necessários para se efectuar a corrida electroforética, assim como a detecção automática dos polimorfismos de sequência e, posteriormente, a análise dos resultados, por comparação com a sequência de referência de Cambridge (CRS) (Anderson *et al*., 1981).

3.7 – Análise estatística

Foram calculadas as frequências alélicas de cada um dos 15 *loci* STRs autossómicos (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, VWA, TPOX, D18S51, D5S818, FGA e amelogenina), assim como o número de homozigotos e heterozigotos observados e esperados. A heterozigosidade foi calculada como descrito por Nei (1987).

O equilíbrio de Hardy-Weinberg, para cada uma das etnias e para a população global, relativamente a cada marcador, foi calculado com recurso ao *software Genepop,* (Guo, 1992), sendo os correspondentes valores de P obtidos pelo método das cadeias de Markov (Guo e Thompson, 1992). Com o mesmo *software* foram determinados os valores obtidos e esperados da heterozigosidade.

A utilização do *software* Arlequin permitiu obter a matriz de distâncias genéticas e correspondentes valores de P entre os diversos grupos étnico-linguísticos, com recurso ao método Neighbor-Joining (Saitou and Nei, 1987). As árvores filogenéticas foram determinadas por aplicação do *software TreeView* (TreeView versão 1.5.2, 1998). Nos testes realizados foram considerados diferenças estatisticamente significativas para valores em que P era <0.05 (Nei *et al*., 1987).

As frequências génicas e a diversidade haplotípica dos marcadores do PowerPlex Y (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS385, DYS437, DYS438 e DYS439) do cromossoma Y foram determinadas de acordo com Nei (1987).

Relativamente ao ADN mitocondrial foram calculadas a diversidade nucleotídica e diversidade de sequência para a região controlo total, nos três principais grupo étnico-linguísticos de Angola, segundo Nei e Tajima (1981). O programa Arlequin 2000 permitiu efectuar a análise molecular da variância AMOVA (Excoffier *et al*., 1992) e obter medidas de distância genética inter-populacional para estabelecer as relações filogenéticas.

4. Resultados

Polimorfismos de STRs autossómicos

Polimorfismos de STRs do cromossoma Y

Polimorfismos do ADN mitocondrial

4.1 – Polimorfismos de STRs autossómicos

4.1.1 – Polimorfismos de STRs autossómicos da população angolana

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses da população global, estão representados na Tabela 4.1.

Tabela 4.1 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para a população global de Angola (N= 479).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
5								0.0031							
6			0.0010	0.0010		0.0656		0.0010				0.0865			
7			0.0052	0.0635		**0.4219**						0.0146			
8			0.2531	0.0750		0.3229	0.0125	0.0167				**0.2927**		0.0573	
9			0.1385	0.0542		0.1396	0.0073	0.2604		0.0031		0.2104		0.0333	
9.3						0.0417									
10	0.0021		**0.3063**	**0.2750**	0.0021	0.0083	0.0198	0.1031		0.0135		0.0958	0.0042	0.0490	
10.2													0.0031		
11	0.0490		0.1792	0.2250			0.2938	**0.3573**		0.0771	0.0135	0.2781	0.0052	0.1885	
11.2										0.0021					
12	0.0990		0.0896	0.2531	0.0052		**0.4594**	0.1260	0.0010	0.1260		0.0188	0.0260	**0.3479**	
12.2										0.0521					
13	0.1948		0.0250	0.0458	0.0052		0.1615	0.1219	0.0010	**0.2958**	0.0104	0.0010	0.0365	0.2990	
13.2										0.0646			0.0010		
14	**0.3979**		0.0021	0.0073	0.0646		0.0427	0.0094	0.0010	0.1813	0.0698		0.0625	0.0219	0.0010
14.2										0.0594			0.0031		
15	0.1927				0.2958		0.0031	0.0010	0.0063	0.0615	0.2021	0.0021	0.1885	0.0031	
15.2										0.0417			0.0021		
16	0.0531				**0.3219**				0.0667	0.0063	**0.2740**		**0.2094**		0.0021
16.1															0.0010
16.2										0.0146			0.0010		
17	0.0094				0.2490				0.1073		0.2063		0.1802		0.0031
17.2										0.0010					
18	0.0021				0.0542				0.0458		0.1396		0.1198		0.0115
18.2															0.0146
19					0.0021				0.1490		0.0615		0.0750		0.0844
19.2															0.0010
20									0.0802		0.0177		0.0458		0.0375
20.2													0.0010		
21									0.1469		0.0042		0.0219		0.0990
21.2													0.0010		
22									**0.1635**		0.0010		0.0094		0.1698
23									0.0583				0.0031		0.1552
24		0.0010							0.0635						**0.1813**
24.2		0.0052													0.0010
24.3		0.0010													
25		0.0010							0.0688						0.1188
26		0.0010							0.0344						0.0344
27		0.0771							0.0052						0.0458
28		**0.2406**							0.0010						0.0083
29		0.1656													0.0073
30		0.1833													0.0021
30.2		0.0115													0.0052
31		0.0865													
31.2		0.0490													0.0083
32		0.0177													
32.2		0.0469													0.0031
33		0.0094													
33.1		0.0010													
33.2		0.0208													0.0010
34		0.0198													
34.2		0.0010													
35		0.0469													
36		0.0104													
37		0.0021													
38		0.0010													
42.2															0.0021
43.2															0.0010
He_{obs}	0.7604	0.8854	0.7625	0.8083	0.7396	0.6792	0.6563	0.7729	0.8813	0.8500	0.8167	0.7646	0.8542	0.7375	0.8542
He_{esp}	0.7521	0.8604	0.7833	0.7958	0.7396	0.6938	0.6750	0.7625	0.8917	0.8417	0.8125	0.7771	0.8604	0.7479	0.8792
P	0.5797 ± 0.0168	0.4591 ± 0.0312	0.2999 ± 0.0157	0.0184* ± 0.0026	0.0543 ± 0.0043	0.4701 ± 0.0087	0.0117* ± 0.0012	0.1633 ± 0.0142	0.0150* ± 0.0013	0.5113 ± 0.0189	0.4963 ± 0.0146	0.0630 ± 0.0042	0.8461 ± 0.0187	0.8049 ± 0.0084	0.6654 ± 0.0302
PD	0.7515	0.8593	0.7822	0.7949	0.7397	0.6922	0.6742	0.7628	0.8916	0.8414	0.8128	0.7755	0.8595	0.7468	0.8759
PEx	0.5414	0.7235	0.5757	0.6008	0.5022	0.4450	0.4234	0.5514	0.7810	0.6952	0.6315	0.5644	0.7205	0.5235	0.7468

*** correcção de Bonferroni (0.05/15=0.0033)**

Na população global de Angola, verificou-se que os alelos mais frequentes em cada *locus* foram: 14 para D8S1779, 28 para D21S11, 10 para D7S820, 10 para CSF1PO, 16 para D3S1358, 7 para TH01, 12 para D13S317, 11 para D16S539, 22 para D2S1338, 13 para D19S433, 16 para VWA, 8 para TPOX, 16 para D18S51, 12 para D5S818 e 24 para FGA.

A aplicação do teste exacto de Guo e Thompson a cada um dos 15 *loci* mostrou que a população de Angola se encontra em equilíbrio de Hardy-Weinberg para a maioria dos *loci* genéticos estudados, à excepção do *loci* CSF1PO (P=0.0184), D13S31 (P=0.0117) e D2S1338 (P=0.0150); após correcção de Bonferroni, não havia diferenças estatisticamente significativas.

Quanto aos parâmetros estatísticos de eficácia *a priori*, verificou-se que os valores de heterozigosidade média observada eram superiores a 0.7 para todos os 15 *loci* STRs estudados.

Os valores referentes ao poder de discriminação *a priori* variam de 0.6742 (D13S317) a 0.8759 (FGA) e o poder de discriminação acumulada nos 15 *loci* foi de 0.999999999968716.

A probabilidade de exclusão *a priori* apresentou valores que variam de 0.4234 (D13S317) a 0.7810 (D2S1338) e apresentaram um valor médio de 0.6017, sendo que o poder de exclusão *a priori* acumulado nos 15 *loci* foi de 0.999999028.

4.1.2 – Comparação dos polimorfismos de STRs autossómicos da população angolana com outras populações

As comparações genético-populacionais realizadas entre a população angolana e cada uma das 12 populações Africana, latino americanas e europeias por meio da matriz de distância diferenciação do valor de P testada para cada *locus*, estão representadas nas (Tabelas 4.2 e 4.3).

Tabela 4.2 – Matriz de distâncias relativa a diferentes populações mundiais, obtida a partir do programa Arlequim 2.000, pelo cálculo de FST.

Populações	Moçambi-que	Angola	Bahia-Basil	Venezuela	México	Argentina	Somália	Uganda	Guiné Equatorial	Afro-americanos	Espanha	Portugal	Namíbia
Moçambique	---	0.0261	0.0837	0.1139	0.1400	0.1109	0.0783	0.0622	0.0266	0.0263	0.1352	0.1218	0.0455
Angola	0.0261	---	0.0881	0.1269	0.1547	0.1253	0.0756	0.0484	0.0101	0.0183	0.1417	0.1327	0.0123
Bahia-Brasil	0.0837	0.0881	---	0.0255	0.0456	0.0257	0.0731	0.0847	0.0758	0.0499	0.0279	0.0236	0.1068
Venezuela	0.1139	0.1269	0.0255	---	0.0226	0.0190	0.0791	0.0974	0.1049	0.0731	0.0206	0.0196	0.1504
México	0.1400	0.1547	0.0456	0.0226	---	0.0256	0.0953	0.1185	0.1339	0.0951	0.0443	0.0432	0.1735
Argentina	0.1109	0.1253	0.0257	0.0190	0.0256	---	0.0890	0.1113	0.1086	0.0741	0.0235	0.0210	0.1471
Somália	0.0783	0.0756	0.0731	0.0791	0.0953	0.0890	---	0.0381	0.0659	0.0510	0.0979	0.1017	0.0938
Uganda	0.0622	0.0484	0.0847	0.0974	0.1185	0.1113	0.0381	---	0.0467	0.0317	0.1236	0.1217	0.0659
Guiné Equatorial	0.0266	0.0101	0.0758	0.1049	0.1339	0.1086	0.0659	0.0467	---	0.0163	0.1165	0.1121	0.0232
Afro-americanos	0.0263	0.0183	0.0499	0.0731	0.0951	0.0741	0.0510	0.0317	0.0163	---	0.0869	0.0827	0.0344
Espanha	0.1352	0.1417	0.0279	0.0206	0.0443	0.0235	0.0979	0.1236	0.1165	0.0869	---	0.0055	0.1584
Portugal	0.1218	0.1327	0.0236	0.0196	0.0432	0.0210	0.1017	0.1217	0.1121	0.0827	0.0055	---	0.1500
Namíbia	0.0455	0.0123	0.1068	0.1504	0.1735	0.1471	0.0938	0.0659	0.0232	0.0344	0.1584	0.1500	---

Tabela 4.3 – Comparação entre população de Angola e outras populações através da diferenciação do valor de P testada para cada *locus*.

Loci	Angola vs Moçambique	Angola vs Bahía Brasil	Angola vs Venezuela	Angola vs México	Angola vs Argentina	Angola vs Somália	Angola vs Uganda	Angola vs Guiné Equatorial	Angola vs Afro-Americanos	Angola vs Espanha	Angola vs Portugal	Angola vs Namíbia
D8S1179	0.2520	0.0001	0.0000	0.0000	0.0000	0.0008	0.0809	0.5965	0.0296	0.0000	0.0000	0.6419
D21S11	0.0128	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.9370	0.7869	0.0000	0.0000	0.5335
D7S820	0.2386	0.0055	0.0000	0.0000	0.0000	0.0000	0.0691	0.9169	0.3525	0.0000	0.0000	0.6232
CSF1PO	0.6188	0.0019	0.0000	0.0000	0.0000	0.0000	0.0697	0.5565	0.8852	0.0000	0.0000	0.1212
D3S1358	0.3988	0.2409	0.2527	0.0000	0.0037	0.6855	0.3867	0.9210	0.8982	0.0000	0.0000	0.6728
TH01	0.0980	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4021	0.0010	0.0000	0.0000	0.7990
D13S317	0.8070	0.0000	0.0000	0.0000	0.0000	0.0000	0.3663	0.9666	0.5307	0.0000	0.0000	0.8808
D16S539	0.0193	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0131	0.1018	0.0000	0.0000	0.2979
D2S1338	0.6287	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.2198	0.5369	0.0000	0.0000	0.3447
D19S433	0.0236	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5443	0.5439	0.0000	0.0000	0.5072
VWA	0.6053	0.4262	0.0013	0.0000	0.0000	0.0000	0.0002	0.8235	0.5962	0.0895	0.0000	0.2357
TPOX	0.4572	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.2316	0.2508	0.0000	0.0000	0.4482
D18S51	0.2406	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017	0.9998	0.0469	0.0000	0.0000	0.3105
D5S818	0.0457	0.0004	0.0000	0.0000	0.0000	0.1826	0.7093	0.8890	0.7609	0.0000	0.0000	0.7761
FGA	0.0634	0.1011	0.0000	0.0149	0.0000	0.0000	0.0000	0.9404	0.0756	0.0000	0.0000	0.9143

Moçambique (Alves, *et al*., 2004), Bahia (Brasil) (Santos *et al*., 2004), Venezuela (Bernal, *et al*., 2006), México (Hernández, *et al*., 2005), Argentina (Bozzo, *et al*., 2007), Somália (Tilmar, *et al*., 2009), Uganda (Gomes, *et al*., 2009), Guiné Equatorial (Alves, et *al*., 2005), Afro-Americanos (EUA) (Butler, et *al*., 2003), Espanha (Camacho, *et al*., 2007), Portugal (Lopes, *et al*., 2009), Namíbia (Muro, *et al*., 2008).

A árvore filogenética relativa às 13 populações estudadas está representada na Fig. 4.1.

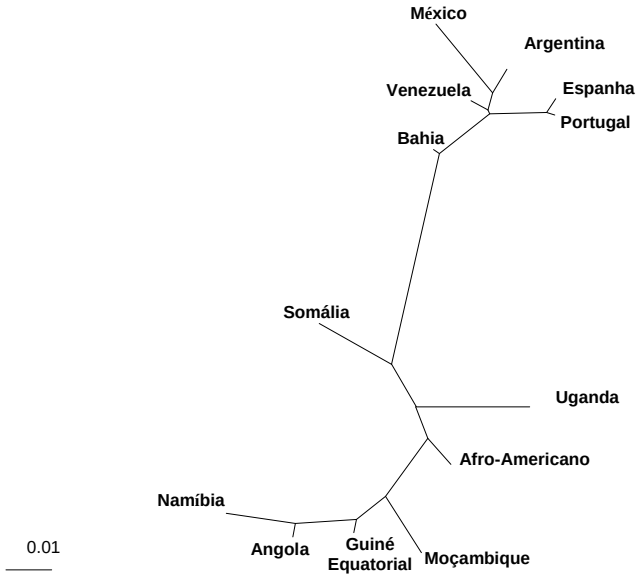


Figura 4.1 – Árvore filogenética da comparação entre 13 diferentes populações, obtida com base nas distâncias genéticas de Nei a partir do método Neighbor-Joining e visualizada no programa Tree View.

4.1.3 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Bakongo

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicações forenses do grupo étnico-linguístico Bakongo, estão representados na Tabela 4.4.

Tabela 4.4 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de o equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori* , para o grupo étnico-linguístico Bakongo (N = 102).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
5								0.0098							
6						0.0392						0.0686			
7			0.0098	0.0539		0.4657						0.0245			
8			0.1765	0.0686		0.3138	0.0098	0.0049				0.2549		0.0686	
9			0.1519	0.0637		0.1372		0.2501		0.0049		0.2794		0.0441	
9.3						0.0392									
10	0.0049		0.3187	0.2255		0.0049	0.0196	0.1470		0.0147		0.0784	0.0049	0.0392	
11	0.0245		0.1765	0.2255			0.3089	0.3285		0.0686	0.0049	0.2647	0.0049	0.1764	
11.2										0.0049					
12	0.0833		0.1176	0.3089	0.0049		0.4020	0.1127		0.1177		0.0196	0.0196	0.3627	
12.2										0.0392					
13	0.2157		0.0441	0.0490	0.0049		0.2009	0.1372		0.2893	0.0098		0.0245	0.2892	
13.2										0.0588			0.0049		
14	0.3873		0.0049	0.0049	0.0784		0.0490	0.0098	0.0049	0.1765	0.0735		0.0539	0.0147	
14.2										0.0588			0.0098		
15	0.2304				0.3138		0.0098		0.0098	0.0735	0.2205	0.0098	0.1715	0.0049	
15.2										0.0686					
16	0.0441				0.3138				0.0686	0.0049	0.2402		0.2500		0.0049
16.2										0.0196			0.0049		
17	0.0098				0.2254				0.1078		0.2058		0.1617		
18					0.0539				0.0343		0.1421		0.1421		0.0196
18.2															0.0147
19					0.0049				0.1471		0.0735		0.0686		0.1029
20									0.1226		0.0196		0.0588		0.0686
21									0.1471		0.0049		0.0098		0.0882
22									0.1520		0.0049		0.0098		0.1666
23									0.0588						0.1568
24									0.0539						0.1715
25									0.0637						0.1029
26									0.0245						0.0245
27		0.1029							0.0049						0.0294
28		0.2059													0.0098
29		0.1667													0.0196
30		0.2060													
30.2		0.0343													
31		0.0833													
31.2		0.0490													0.0098
32		0.0147													
32.2		0.0343													
33.2		0.0245													0.0049
34		0.0196													
35		0.0245													
36		0.0245													
37		0.0098													
42.2															0.0049
He_{obs}	0.7647	0.9118	0.7843	0.8431	0.6862	0.6470	0.7549	0.7843	0.8529	0.8921	0.8529	0.7745	0.9019	0.7353	0.8333
He_{exp}	0.7451	0.8627	0.8039	0.7941	0.7451	0.6666	0.7059	0.7843	0.8922	0.8529	0.8235	0.7745	0.8529	0.7451	0.8823
P	0.5308 ± 0.0091	0.3397 ± 0.0137	0.8038 ± 0.0059	0.2281 ± 0.0080	0.1556 ± 0.0088	0.6177 ± 0.0086	0.0326* ± 0.0027	0.2582 ± 0.0081	0.9193 ± 0.0064	0.4983 ± 0.0167	0.9270 ± 0.0050	0.0169* ± 0.0017	0.8184 ± 0.0154	0.4302 ± 0.0080	0.1387 ± 0.0137
PD	0.7408	0.8627	0.7972	0.7888	0.7432	0.6628	0.6997	0.7762	0.8900	0.8474	0.8197	0.7749	0.8492	0.7452	0.8818
PEx	0.5169	0.7279	0.6050	0.5922	0.5092	0.4080	0.4469	0.5680	0.7772	0.7060	0.6420	0.5637	0.7023	0.5232	0.7622

*** correcção de Bonferroni (0.05/15=0.0033)**

Nos Bakongos, da região Norte de Angola, verificou-se que os alelos mais frequentes em cada *locus* eram: 14 para D8S1779, 30 para D21S11, 10 para D7S820, 12 para CSF1PO, 15 e 16 para D3S1358, 7 para TH01, 12 para D13S317, 11 para

D16S539, 22 para D2S1338, 13 para D19S433, 16 para VWA, 9 para TPOX, 16 para D18S51, 12 para D5S818 e 24 para FGA.

O valor de heterozigosidade média observada foi de 0.80. Entretanto, a aplicação do teste exacto de Guo e Thompson a cada um dos 15 *loci*, mostrou que o grupo Bakongo se encontra em equilíbrio de Hardy-Weinberg para a maioria dos sistemas estudados, à excepção dos *loci* D13S317 (P=0.0326) e TPOX (P=0.0169); após correcção de Bonferroni nenhum dos valores obtidos de P apresentou diferenças significativas.

Os valores referentes ao poder de discriminação *a priori* variaram entre 0.6628 (TH01) a 0.8900 (D2S1338), com um valor médio de 0.7920, e o poder de discriminação acumulada nos 15 *loci* foi de 0,999999999972540.

A probabilidade de exclusão *a priori* apresentou valores que variaram de 0.4080 (TH01) a 0.7772 (D2S1338) e um valor médio de 0.60338, sendo o poder de exclusão *a priori* acumulado nos 15 *loci* de 0.999999492.

4.1.4 – Polimorfismo de STRs autossómicos no grupo étnico-linguístico Kimbundo

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses do grupo étnico-linguístico Kimbundo, estão representados na Tabela 4.5.

Tabela 4.5 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Kimbundo (N = 100).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
5								0.0050							
6				0.0050		0.0950						0.0900			
7				0.0750		0.4150						0.0150			
8			0.2650	0.0600		0.2450	0.0150	0.0150				0.2600		0.0850	
9			0.1400	0.0650		0.1700	0.0100	0.2450				0.2150		0.0150	
9.3						0.0550									
10			0.3250	0.3650	0.0050	0.0200	0.0200	0.1050		0.0100		0.1400	0.0150	0.0550	
10.2													0.0050		
11	0.0650		0.1400	0.1750			0.2850	0.3850		0.1100	0.0150	0.2600	0.0050	0.1800	
12	0.0750		0.1200	0.2100	0.0050		0.4550	0.1200		0.1350		0.0200	0.0350	0.3750	
12.2										0.0450					
13	0.2050		0.0100	0.0400	0.0050		0.1550	0.1100		0.2750	0.0100		0.0250	0.2600	
13.2										0.0600					
14	0.4250			0.0050	0.0850		0.0600	0.0150		0.1750	0.0600		0.0700	0.0300	
14.2										0.0600					
15	0.1550				0.2150				0.0100	0.0650	0.1800		0.1800		
15.2										0.0400			0.0050		
16	0.0550				0.3750				0.0600	0.0100	0.2850		0.2000		
16.2										0.0150					
17	0.0150				0.2600				0.1100		0.2250		0.1700		0.0050
17.2															
18	0.0050				0.0500				0.0450		0.1300		0.1000		
19									0.1250		0.0700		0.0750		0.1150
20									0.0600		0.0200		0.0550		0.0150
21									0.1850		0.0050		0.0250		0.0800
22									0.1600				0.0250		0.1650
23									0.0600				0.0100		0.1600
24		0.0050							0.0600						0.1700
24.2		0.0050													
25									0.0750						0.1500
26									0.0400						0.0450
27		0.0800							0.0050						0.0300
28		0.2600							0.0050						0.0150
29		0.1850													0.0050
30		0.1850													0.0100
30.2		0.0050													0.0100
31		0.0700													
31.2		0.0250													0.0200
32		0.0200													
32.1		0.0550													0.0050
32.2		0.0150													
33		0.0100													
33.2		0.0250													
34.2		0.0500													
36		0.0050													
He_{obs}	0.7700	0.9000	0.7600	0.7900	0.7700	0.6700	0.6700	0.7900	0.8700	0.8900	0.8300	0.6700	0.8600	0.7500	0.8400
He_{exp}	0.7400	0.8500	0.7700	0.7800	0.7400	0.7300	0.6900	0.7600	0.8900	0.8500	0.8100	0.7300	0.8800	0.7500	0.8800
P	0.3787 ± 0.0084	0.7512 ± 0.0168	0.4475 ± 0.0055	0.6355 ± 0.0106	0.3805 ± 0.0103	0.5975 ± 0.0055	0.5836 ± 0.0075	0.8993 ± 0.0047	0.1837 ± 0.0109	0.9933 ± 0.0012	0.9255 ± 0.0040	0.5976 ± 0.0060	0.6241 ± 0.0171	0.5385 ± 0.0079	0.6432 ± 0.0179
PD	0.7402	0.8451	0.7705	0.7770	0.7358	0.7264	0.6834	0.7538	0.8887	0.8480	0.8096	0.7903	0.8717	0.7480	0.8721
PEx	0.5326	0.6972	0.5560	0.5818	0.5029	0.4992	0.4385	0.5423	0.7762	0.7035	0.6275	0.5864	0.7444	0.5286	0.7230

Nos Kimbundos, da região Centro-Norte de Angola, verificou-se que os alelos mais frequentes em cada *locus* foram: 14 para D8S1779, 28 para D21S11, 10 para D7S820, 10 para CSF1PO, 16 para D3S1358, 7 para TH01, 12 para D13S317, 11 para

D16S539, 21 para D2S1338, 13 para D19S433, 16 para VWA, 8 e 11 para TPOX, 16 para D18S51, 12 para D5S818 e 24 para FGA.

O valor da heterozigosidade média observada foi 0.7887. A aplicação do teste exacto de Guo e Thompson a cada um dos 15 *loci* mostrou que a população Kimbundo se encontra em equilíbrio de Hardy-Weinberg para todos os *loci* estudados (P>0.05).

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicações forenses do grupo étnico-linguístico Kimbundo, estão representados na Tabela 4.5.

Os valores referentes ao poder de discriminação *a priori* variam de 0.6834 (D13S317) a 0.8887 (D2S1338); o poder de discriminação acumulada nos 15 *loci* foi de 0.999999999967503. A probabilidade de exclusão *a priori* apresentou valores que variam de 0.4385 (D13S317) a 0.7762 (D2S1338), com um valor médio de 0.5941, sendo o poder de exclusão *a priori* acumulado nos 15 *loci* de 0.999999427.

4.1.5 – Polimorfismo de STRs autossómicos no grupo étnico-linguístico Kwanhama

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses do grupo étnico-linguístico Kwanhama, estão representados na Tabela 4.6.

Tabela 4.6 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Kwanhama (N = 50).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
6			0.0100			0.0200						0.0900			
7			0.0100	0.0400		0.4600						0.0300			
8			0.3700	0.0900		0.3600	0.0200	0.0200				0.3300		0.0600	
9			0.1100	0.0700		0.1200	0.0100	0.2600				0.1300		0.0200	
9.3						0.0300									
10			0.2900	0.2300		0.0100	0.0100	0.0900				0.0700		0.0300	
11	0.0300		0.1900	0.2700			0.2600	0.3200		0.0300	0.0100	0.3400		0.2100	
12	0.0900		0.0200	0.2600			0.4800	0.1500		0.0700		0.0100	0.0200	0.2800	
12.2										0.1200					
13	0.1300			0.0400			0.1600	0.1600		0.3100	0.0100		0.0800	0.3800	
13.2										0.0700					
14	0.4300				0.0500		0.0500			0.2400	0.0900		0.0500	0.0200	
14.2										0.0700					
15	0.2600				0.3300		0.0100			0.0600	0.2000		0.1900		
15.2										0.0200					
16	0.0600				0.2600				0.0900	0.0100	0.3100		0.1500		
17					0.2900				0.1000		0.1600		0.2500		0.0100
18					0.0700				0.0800		0.1700		0.1200		
18.2															0.0300
19									0.1800		0.0500		0.0500		0.0800
19.2															0.0100
20									0.0500				0.0700		0.0200
21									0.1600				0.0200		0.1300
22									0.1000						0.0700
23									0.0400						0.1800
24									0.0700						0.1900
24.2															0.0100
25									0.0900						0.1300
26									0.0400						0.0300
27		0.0500													0.0900
28		0.2600													
29		0.1300													
30		0.2000													
31		0.0800													
31.2		0.0300													0.0100
32		0.0100													
32.2		0.0600													
33		0.0200													
33.2		0.0200													
34		0.0300													
34.2		0.0100													
35		0.0900													
38		0.0100													
43.2															0.0100
He_{obs}	0.7000	0.8600	0.7600	0.8400	0.8000	0.7000	0.6400	0.8600	0.9200	0.7400	0.7800	0.7800	0.8200	0.7200	0.8200
He_{exp}	0.7200	0.8600	0.7400	0.8000	0.7400	0.6400	0.6800	0.7800	0.9000	0.8200	0.8000	0.7600	0.8600	0.7400	0.8800
P	0.4311 ± 0.0051	0.2941 ± 0.0162	0.1356 ± 0.0058	0.8640 ± 0.0035	0.1748 ± 0.0054	0.0462* ± 0.0033	0.0711 ± 0.0051	0.3310 ± 0.0052	0.4138 ± 0.0087	0.2277 ± 0.0099	0.3484 ± 0.0076	0.0421* ± 0.0032	0.1335 ± 0.0060	0.8010 ± 0.0047	0.0268* ± 0.0069
PD	0.7180	0.8520	0.7302	0.7904	0.7320	0.6430	0.6732	0.7734	0.8888	0.8122	0.7986	0.7446	0.8474	0.7278	0.8756
PEx	0.4925	0.7112	0.4922	0.5911	0.4873	0.3765	0.4313	0.5608	0.7754	0.6414	0.6070	0.5232	0.6981	0.4920	0.7505

*** correcção de Bonferroni (0.05/15=0.0033)**

Nos, da região mais ao Sul de Angola, verificou-se que os alelos mais frequentes em cada *locus* foram: 14 para D8S1779, 28 para D21S11, 8 para D7S820, 11 para CSF1PO, 17 para D3S1358, 7 para TH01, 12 para D13S317, 11 para D16S539, 19 para

D2S1338, 13 para D19S433, 16 para VWA, 11 para TPOX, 17 para D18S51, 13 para D5S818 e 24 para FGA.

O valor de heterozigosidade média observada foi 0.7827. O grupo Kwanhama encontra-se em equilíbrio de Hardy-Weinberg para a maioria dos 15 *loci* estudados, à excepção dos sistemas TH01 (P=0.0462), TPOX (P=0.0421) e FGA (P= 0.0268); após correcção de Bonferroni, não havia diferenças significativas também para estes 3 *loci*.

Os valores referentes ao poder de discriminação *a priori* variaram entre 0.6430 (TH01) e 0.8888 (D2S1338) e o poder de discriminação acumulado nos 15 *loci* era de 0.999999999907182.

A probabilidade de exclusão *a priori* apresentou valores que variaram de 0.3765 (TH01) a 0.7754 (D2S1338), com um valor médio de 0.57753, sendo que o poder de exclusão *a priori* acumulado nos 15 *loci* foi de 0.999998597.

4.1.6 – Polimorfismos de STRs autossómicos no grupo étnico-linguístico Lunda-Tchokwe

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses do grupo étnico-linguístico Lunda-Tchokwe, estão representados na Tabela 4.7.

Tabela 4.7 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Lunda-Tchokwe (N = 50).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVIWA	TPOX	D18S51	D5S818	FIBRA/FGA
6						0.0600		0.0100				0.0500			
7				0.0400		0.3900									
8			0.2300	0.1100		0.3200	0.0100	0.0200				0.3600		0.0100	
9			0.1800	0.0600		0.2000	0.0100	0.2400		0.0200		0.1500		0.0300	
9.3						0.0300									
10			0.3800	0.2400	0.0100		0.0200	0.1600				0.1100		0.0300	
10.2													0.0100		
11	0.0600		0.1200	0.2500			0.3400	0.3100		0.0900	0.0200	0.3000	0.0100	0.1800	
11.2										0.0100					
12	0.1100		0.0600	0.2800			0.4300	0.1500		0.0900		0.0300	0.0100	0.4100	
12.2										0.0600					
13	0.2000		0.0200	0.0100			0.1800	0.1000		0.2300			0.0200	0.3300	
13.2										0.0900					
14	0.4100		0.0100	0.0100	0.0600		0.0100			0.2400	0.0500		0.0800	0.0100	
14.2										0.0400					
15	0.1300				0.2400			0.0100		0.0700	0.2200		0.1900		
15.2										0.0300					
16	0.0600				0.3900				0.0900	0.0100	0.2500		0.1900		0.0100
16.1															0.0100
16.2										0.0200					
17	0.0300				0.2500				0.0900		0.2200		0.1700		
18					0.0500				0.0400		0.1400		0.1500		0.0100
18.2															0.0400
19									0.1900		0.0700		0.0900		0.0700
20									0.0700		0.0200		0.0100		0.0200
21									0.1400		0.0100		0.0400		0.1000
21.2													0.0100		
22									0.1300				0.0200		0.2200
23									0.0500						0.1400
24									0.0800						0.2500
24.3		0.0100													
25		0.0100							0.0900						0.0500
26									0.0300						0.0300
27		0.0900													0.0300
28		0.2500													0.0100
29		0.1900													
30		0.2100													
30.2															0.0100
31		0.0700													
31.2		0.0400													
32		0.0100													
32.2		0.0400													
33		0.0100													
34		0.0200													
35		0.0400													
36		0.0100													
Heobs	0.7800	0.8200	0.6600	0.8400	0.7600	0.6800	0.7400	0.7000	0.8400	0.9200	0.8200	0.6400	0.8800	0.7000	0.8800
He_{exp}	0.7600	0.8400	0.7600	0.8000	0.7200	0.7000	0.6800	0.8000	0.9000	0.8600	0.8200	0.7400	0.8600	0.7000	0.8400
P	0.9709 ± 0.0012	0.4121 ± 0.0169	0.1035 ± 0.0039	0.0251* ± 0.0023	0.9132 ± 0.0025	0.4617 ± 0.0044	0.8451 ± 0.0064	0.2926 ± 0.0082	0.3906 ± 0.0077	0.5671 ± 0.0126	0.7429 ± 0.0079	0.2312 ± 0.0046	0.4444 ± 0.0161	0.9878 ± 0.0011	0.8438 ± 0.0134
PD	0.7548	0.8386	0.7518	0.7840	0.7216	0.7010	0.6664	0.7876	0.8868	0.8532	0.8128	0.7424	0.8590	0.6886	0.8280
PEx	0.5545	0.6837	0.5344	0.5786	0.4801	0.4459	0.3985	0.5863	0.7716	0.7132	0.6300	0.5167	0.7174	0.4327	0.6833

*** correcção de Bonferroni (0.05/15=0.0033)**

No grupo étnico-linguístico Lunda-Tchokwe, da região Nordeste de Angola, verificou-se que os alelos mais frequentes em cada *locus* foram: 14 para D8S1779, 28 para D21S11, 10 para D7S820, 12 para CSF1PO, 16 para D3S1358, 7 para TH01, 12 para D13S317, 11 para D16S539, 19 para D2S1338, 14 para D19S433, 16 para VWA, 8 para TPOX, 15 e 16 para D18S51, 12 para D5S818 e 24 para FGA.

O valor de heterozigosidade média observada foi de 0.7773. A aplicação do teste exacto de Guo e Thompson a cada um dos 15 *loci* mostrou que o grupo Lunda-Tchokwe se encontra em equilíbrio de Hardy-Weinberg para a maioria dos sistemas estudados, à excepção dos sistemas CSF1PO (P=0.0251); após correcção de Bonferroni, não havia diferenças significativas.

Os valores referentes ao poder de discriminação *a priori* variaram entre 0.6664 (D13S317) a 0.8868 (D2S1338) e o poder de discriminação acumulada nos 15 *loci* foi de 0.999999999924266.

A probabilidade de exclusão *a priori* apresentou valores que variaram de 0.3985 (D13S317) a 0.7716 (D2S1338), com um valor médio de 0.58179, sendo que o poder de exclusão *a priori* acumulado nos 15 *loci* foi de 0.999998821.

4.1.7 – Polimorfismo de STRs autossómicos no grupo étnico-linguístico Nganguela

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses do grupo étnico-linguístico Nganguela, estão representados na Tabela 4.8.

Tabela 4.8 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Nganguela (N = 50).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
6						0.1100						0.1100			
7				0.0600		0.4300									
8			0.2500	0.0600		0.3200						0.2700		0.0500	
9			0.1000	0.0400		0.1100	0.0300	0.3000				0.1900		0.0600	
9.3						0.0300									
10			0.2500	0.3000			0.0100	0.0800		0.0400		0.1200		0.0600	
11	0.0400		0.2900	0.2400			0.2800	0.4200		0.0600		0.3000		0.2200	
12	0.1200		0.0900	0.2100			0.5300	0.1000	0.0100	0.1300		0.0100	0.0100	0.3400	
12.2										0.0500					
13	0.2500		0.0200	0.0900			0.1200	0.0900	0.0100	0.3300	0.0200		0.0700	0.2700	
13.2										0.0500					
14	0.4000				0.0400		0.0300	0.0100		0.1700	0.0700		0.0800		0.0100
14.2										0.0900					
15	0.1500				0.3700				0.0100	0.0300	0.2200		0.1800		
15.2										0.0300			0.0100		
16	0.0400				0.2400				0.0700		0.2400		0.2500		
16.2										0.0200					
17					0.2900				0.1300		0.2700		0.1900		0.0100
18					0.0600				0.0300		0.0900		0.1100		0.0300
18.2															0.0100
19									0.1200		0.0400		0.0500		0.0700
20									0.0600		0.0400		0.0300		0.0400
20.2													0.0100		
21									0.1100		0.0100		0.0100		0.0900
22									0.2100						0.1500
23									0.1000						0.1500
24									0.0600						0.2400
25									0.0300						0.1100
26									0.0400						0.0100
27		0.0600							0.0100						0.0800
28		0.3100													
29		0.1600													
30		0.1400													
31		0.0900													
31.2		0.0400													
32		0.0400													
32.2		0.0600													
33.2		0.0200													
34		0.0300													
35		0.0500													
He_{obs}	0.6800	0.8600	0.7400	0.7400	0.7200	0.7200	0.5600	0.8200	0.8800	0.8400	0.8200	0.7600	0.7600	0.8600	0.9000
He_{exp}	0.7400	0.8400	0.7800	0.8000	0.7200	0.7000	0.6200	0.7200	0.9000	0.8400	0.8200	0.7800	0.8400	0.7600	0.8800
P	0.5508 ± 0.0048	0.8645 ± 0.0069	0.2533 ± 0.0053	0.2696 ± 0.0057	0.2816 ± 0.0039	0.2334 ± 0.0041	0.3041 ± 0.0068	0.3798 ± 0.0068	0.0319* ± 0.0048	0.1090 ± 0.0074	0.4711 ± 0.0071	0.9672 ± 0.0012	0.3333 ± 0.0130	0.0992 ± 0.0029	0.5874 ± 0.0123
PD	0.7374	0.8364	0.7724	0.7914	0.7162	0.6876	0.6244	0.7090	0.8866	0.8248	0.8044	0.7744	0.8418	0.7534	0.8630
PEx	0.5151	0.6877	0.5558	0.5952	0.4653	0.4384	0.3736	0.4745	0.7735	0.6698	0.6181	0.5596	0.6874	0.5299	0.7280

*** correcção de Bonferroni (0.05/15=0.0033)**

Nos Nganguela, da região Leste de Angola, verificou-se que os alelos mais frequentes em cada *locus* foram: 14 para D8S1779, 28 para D21S11, 11 para D7S820, 10 para CSF1PO, 15 para D3S1358, 7 para TH01, 12 para D13S317, 11 para D16S539, 22 para D2S1338, 13 para D19S433, 17 para VWA, 11 para TPOX, 16 para D18S51, 12 para D5S818 e 24 para FGA.

O valor de heterozigosidade média observada foi de 0.7773. A aplicação do teste exacto de Guo e Thompson a cada um dos 15 *loci* mostrou que o grupo Nganguela se encontra em equilíbrio de Hardy-Weinberg para a maioria dos sistemas estudados, à excepção do *loci* D2S1338 (P= 0.0319); após correcção de Bonferroni, esta diferença não era significativa.

Os valores referentes ao poder de discriminação *a priori* variaram de 0.6244 (D13S317) a 0.8866 (D2S1338) e o poder de discriminação acumulada nos 15 *loci* foi de 0.999999999907385.

A probabilidade de exclusão *a priori* apresentou valores que variavam entre 0.3736 (D13S317) e 0.7735 (D2S1338), com um valor médio de 0.57813, sendo que o poder de exclusão *a priori* acumulado nos 15 *loci* foi de 0.999998646.

4.1.8 – Polimorfismo de STRs autossómicos no grupo étnico-linguístico Nhaneca-Humbe

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses do grupo étnico-linguístico Nhaneca-Humbe, são apresentados na Tabela 4.9.

Tabela 4.9 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori* , para o grupo étnico-linguístico Nhaneca-Humbe (N = 25).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
6						0.0200						0.1800			
7				0.0600		0.4000						0.0200			
8			0.2600	0.0200		0.3800		0.0200				0.2400		0.0200	
9			0.1600	0.0600		0.1600		0.2400				0.3000			
9.3						0.0200									
10			0.3600	0.2600		0.0200		0.0400		0.0200		0.0200		0.0800	
11	0.1200		0.1200	0.2600			0.2400	0.4400		0.1200	0.0400	0.2400		0.2200	
12	0.2000		0.0600	0.2800			0.5200	0.1200		0.1200			0.0400	0.2400	
12.2										0.0600					
13	0.1800		0.0400	0.0200			0.2000	0.1400		0.2000	0.0200		0.0600	0.4200	
13.2										0.0400					
14	0.3600			0.0400	0.0200		0.0400			0.2200	0.1000		0.0600	0.0200	
14.2										0.1000			0.0200		
15	0.1400				0.4400					0.1000	0.1800		0.2000		
16					0.2600				0.0600		0.2200		0.1600		
17					0.2000				0.1400		0.2000		0.0800		
17.2										0.0200					
18					0.0800				0.0400		0.1800		0.1600		
18.2															0.0200
19									0.0600		0.0400		0.1400		0.0400
20									0.0600		0.0200		0.0600		0.0400
21									0.1400				0.0200		0.1400
22									0.2000						0.1400
23									0.0600						0.1400
24									0.0800						0.1600
24.2		0.0200													
25									0.0800						0.1800
26									0.0800						0.0600
27		0.0600													0.0800
28		0.2800													
29		0.1000													
30		0.1600													
31		0.0600													
31.2		0.1800													
32		0.0200													
33		0.0200													
33.2		0.0600													
35		0.0400													
He_{obs}	0.8800	0.8800	0.7200	0.8400	0.7200	0.7200	0.6000	0.7600	0.8800	0.9200	0.7600	0.9200	0.8400	0.8000	0.8800
He_{exp}	0.8000	0.8400	0.7600	0.8000	0.7200	0.6800	0.6400	0.7200	0.9200	0.8800	0.8400	0.7600	0.8800	0.7200	0.8800
P	0.1758 ± 0.0028	0.8264 ± 0.0083	0.2617 ± 0.0047	0.0562 ± 0.0039	0.7334 ± 0.0034	0.9246 ± 0.0027	0.3910 ± 0.0054	0.5165 ± 0.0031	0.4428 ± 0.0102	0.3751 ± 0.0097	0.1571 ± 0.0066	0.4998 ± 0.0057	0.5535 ± 0.0106	0.0681 ± 0.0031	0.5720 ± 0.0089
PD	0.7640	0.8400	0.7576	0.7768	0.6920	0.6688	0.6304	0.7128	0.8856	0.8568	0.8328	0.7616	0.8696	0.7104	0.8696
PEx	0.5477	0.6885	0.5416	0.5685	0.4418	0.4028	0.3689	0.4841	0.7702	0.7139	0.6656	0.5322	0.7383	0.4687	0.7361

Nos, da região Sul de Angola, verificou-se que os alelo mais frequente em cada *locus* foram: 14 para D8S1779, 28 para D21S11, 10 para D7S820, 12 para CSF1PO, 15 para D3S1358, 7 para TH01, 12 para D13S317, 11 para D16S539, 22 para D2S1338, 14

para D19S433, 16 para VWA, 9 para TPOX, 15 para D18S51, 13 para D5S818 e 24 para FGA.

A heterozigosidade média observada era de 0.8080. O grupo Nhaneca-Humbe se encontra em equilíbrio de Hardy-Weinberg para todos os *loci* estudados (P>0.05).

Os valores referentes ao poder de discriminação *a priori* variaram de 0.6304 (D13S317) a 0.8856 (D2S1338) e o poder de discriminação acumulada nos 15 *loci* foi de 0.999999999928865.

A probabilidade de exclusão *a priori* apresentou valores que variaram entre 0.3689 (D13S317) e 0.7702 (D2S1338), com um valor médio de 0.5779, sendo que o poder de exclusão *a priori* acumulado nos 15 *loci* era de 0.999998858.

4.1.9 – Polimorfismo de STRs autossómicos no grupo étnico-linguístico Ovimbundo.

Os resultados relativos ao estudo das frequências alélicas de STRs autossómicos e parâmetros estatísticos com aplicação forenses do grupo étnico-linguístico Ovimbundo, são apresentados na Tabela 4.10.

Tabela 4.10 – Frequências alélicas para cada um dos 15 marcadores, número de heterozigotos observados (He_{obs}) e esperados (He_{esp}), valor de P e desvio padrão nos testes de equilíbrio de Hardy-Weinberg e, parâmetros estatísticos de eficácia *a priori:* poder de discriminação (PD) e probabilidade de exclusão (PEx.) *a priori*, para o grupo étnico-linguístico Ovimbundo (N = 102).

Alelos	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
6								0.0098							
7						0.0392						0.0686			
8			0.0098	0.0539		0.4656						0.0245			
9			0.1764	0.0686		0.3137	0.0098	0.0049				0.2549		0.0686	
9.3			0.1519	0.0637		0.1372		0.2500		0.0049		0.2794		0.0441	
10						0.0392									
10.2	0.0049		0.3186	0.2254		0.0049	0.0196	0.1470		0.0147		0.0784	0.0049	0.0392	
11	0.0245		0.1764	0.2254			0.3088	0.3284		0.0686	0.0049	0.2647	0.0049	0.1764	
12										0.0049					
12.2	0.0833		0.1176	0.3088	0.0049		0.4019	0.1127		0.1176		0.0196	0.0196	0.3627	
13										0.0392					
13.2	0.2156		0.0441	0.0490	0.0049		0.2009	0.1372		0.2892	0.0098		0.0245	0.2892	
14										0.0588			0.0049		
14.2	0.3872		0.0049	0.0049	0.0784		0.0490	0.0098	0.0049	0.1764	0.0735		0.0539	0.0147	
15										0.0588			0.0098		
15.2	0.2303				0.3137		0.0098		0.0098	0.0735	0.2205	0.0098	0.1715	0.0049	
16										0.0686					
16.2	0.0441				0.3137				0.0686	0.0049	0.2402		0.2500		0.0049
17										0.0196			0.0049		
18	0.0098				0.2254				0.1078		0.2058		0.1617		
18.2					0.0539				0.0343		0.1421		0.1421		0.0196
19															0.0147
20					0.0049				0.1470		0.0735		0.0686		0.1029
21									0.1225		0.0196		0.0588		0.0686
22									0.1470		0.0049		0.0098		0.0882
23									0.1519		0.0049		0.0098		0.1666
24									0.0588						0.1568
24.2									0.0539						**0.1715**
25									0.0637						0.1029
26									0.0245						0.0245
27		0.1029							0.0049						0.0294
28		0.2058													0.0098
29		0.1666													0.0196
30		0.2059													
30.2		0.0343													
31		0.0833													
31.2		0.0490													0.0098
32		0.0147													
32.2		0.0343													
33		0.0245													0.0049
33.1		0.0196													
33.2		0.0245													
34		0.0245													
35		0.0098													
36															0.0049
42.2								0.0098							
He_{obs}	0.7745	0.9020	0.8137	0.7843	0.7353	0.6863	0.5686	0.7157	0.9216	0.7745	0.7941	0.7941	0.8627	0.6765	0.8725
He_{exp}	0.7647	0.8823	0.7941	0.8040	0.7451	0.6961	0.6666	0.7647	0.8823	0.8137	0.8040	0.7647	0.8530	0.7647	0.8823
P	0.7155 ± 0.0081	0.5101 ± 0.0191	0.4945 ± 0.0101	0.4116 ± 0.0084	0.0943 ± 0.0058	0.7628 ± 0.0051	0.1573 ± 0.0044	0.1838 ± 0.0056	0.8205 ± 0.0076	0.0397* ± 0.0050	0.4933 ± 0.0096	0.0526 ± 0.0040	0.9210 ± 0.0059	0.1855 ± 0.0071	0.9728 ± 0.0043
PD	0.7661	0.8748	0.7885	0.8017	0.7428	0.6917	0.6639	0.7615	0.8819	0.8107	0.7999	0.7636	0.8490	0.7635	0.8740
PEx	0.5595	0.7505	0.5851	0.6114	0.5108	0.4437	0.4149	0.5491	0.7634	0.6484	0.6129	0.5492	0.7009	0.5536	0.7396

*** correcção de Bonferroni (0.05/15=0.0033)**

Nos Ovimbundos, localizados na região do Planalto Central de Angola, verificou-se que os alelos mais frequentes em cada *locus* foram: 14.2 para D8S1779, 30 para D21S11, 10.2 para D7S820, 12.2 para CSF1PO, 15.2 e 16.2 para D3S1358, 8 para TH01, 12.2 para D13S317, 11 para D16S539, 23 para D2S1338, 13.2 para D19S433, 16.2 para VWA, 9.3 para TPOX, 16.2 para D18S51, 12.2 para D5S818 e 24.2 para FGA.

A heterozigosidade média observada foi de 0.7784. O grupo Ovimbundo se encontra em equilíbrio de Hardy-Weinberg para a maioria dos *loci* estudados, à excepção do *loci* D19S433 (P=0.0397); após correcção de Bonferroni, a diferença não se apresentou significativa.

Os valores referentes ao poder de discriminação *a priori* variaram de 0.6639 (D13S317) a 0.8819 (D2S1338) e o poder de discriminação acumulada nos 15 *loci* foi de 0.999999999962836.

A probabilidade de exclusão *a priori* apresentou valores que variavam entre 0.4149 (D13S317) e 0.7634 (D2S1338), com um valor médio de 0.57792, sendo que o poder de exclusão *a priori* acumulado nos 15 *loci* foi de 0.999999358.

4.1.10 – Comparação dos polimorfismos de STRs autossómicos entre os 7 grupos étnico-linguísticos.

Na Tabela 4.11 estão representados os resultados relativos a matriz de distância e respectivos valores de P dos 7 grupos étnico-linguísticos.

Tabela 4.11 - Matriz de distâncias (FST) e valores de P relativa aos diferentes grupos étnico-linguísticos de Angola.

FST / P	Bakongo	Kimbundo	Kwanhama	Lunda-Tchokwe	Nganguela	Nhaneca-Humbe	Ovimbundo
Bakongo		0.0008	0.0029	-0.0009	0.0012	0.0010	0.0006
Kimbundo	0.1699		0.0039	-0.0008	0.0006	0.0030	0.0009
Kwanhama	0.0268	0.0053		0.0010	0.0010	0.0003	0.0009
Lunda-Tchokwe	0.7412	0.7104	0.2760		0.0025	0.0023	0.0007
Nganguela	0.1696	0.3316 ± 0.0095	0.2661	0.1104		-0.0003	-0.0005
Nhaneca-Humbe	0.2800	0.0810	0.3914	0.1673	0.5180		0.0022
Ovimbundo	0.2519	0.1623	0.2493	0.3081	0.6622	0.1663	

A árvore filogenética destes principais grupos étnico-linguísticos de Angola é mostrada na Fig. 4.2.

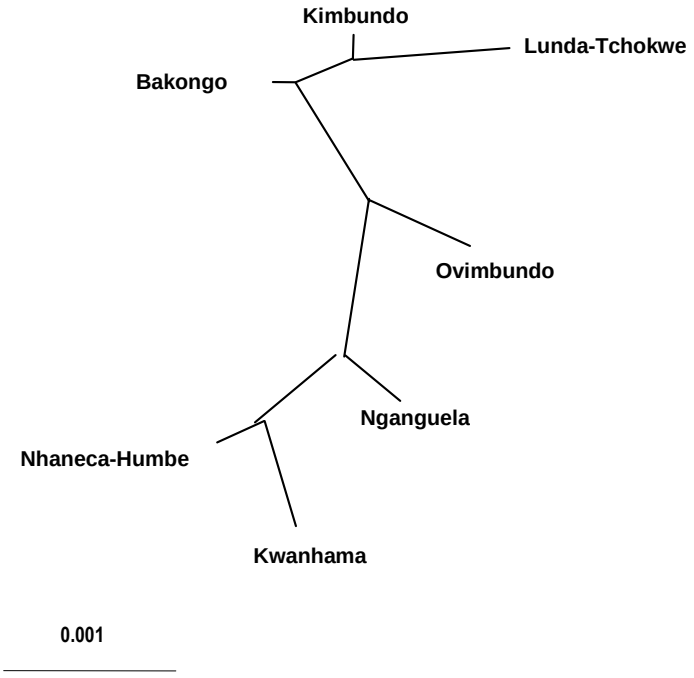


Figura 4.2 - Árvore filogenética dos diferentes grupos étnico-linguísticos de Angola, obtida com base nas distâncias genéticas de Nei, a partir do método Neighbor-Joining, e visualizada no programa Tree View.

4.2. – Polimorfismos de STRs do cromossoma Y

4.2.1 – Polimorfismos de STRs do cromossoma Y na população angolana

A população angolana, na sua maioria, é de origem Bantu e apresenta uma grande diversidade génica entre os diversos grupos étnico-linguísticos. O estudo de frequências alélicas e haplotípicas do cromossoma Y na população de Angola representa grande interesse em estudos de identificação genética individual, investigações forenses e para o estabelecimento de relações filogenéticas entre as populações.

Na Tabela 4.12 estão representados os resultados relativos à distribuição dos alelos por cada *locus*, número de alelos e o alelo mais frequente por *locus* da população Angolana

Tabela 4.12 – Número de alelos, alelo mais frequente, nos marcadores estudados na população angolana (N = 166).

Loci	**Nº de alelos**	**Alelo mais frequente**
DYS19	5	15
DYS389 I	5	13
DYS389 II	7	31
DYS390	7	21
DYS391	4	10
DYS392	5	11
DYS393	6	13
DYS437	5	14
DYS438	4	11
DYS439	4	12

A distribuição haplotípica do *locus* DYS385 da população Angolana está representada na Fig. 4.3.

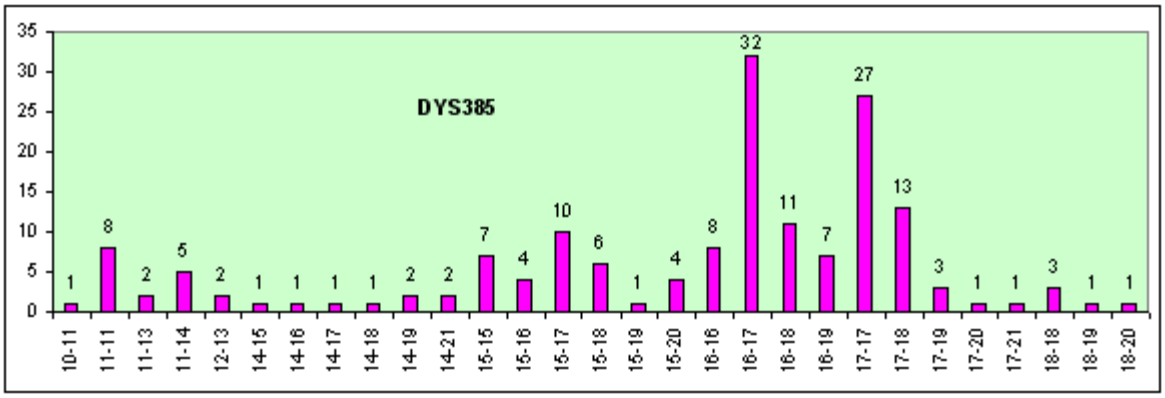


Figura 4.3 – Distribuição dos haplótipos do *locus* DYS385 na população Angolana (N=166).

No marcador DYS385 obtiveram-se 29 haplótipos diferentes em 166 indivíduos do sexo masculino da população angolana, sendo 10 únicos, 4 haplótipos comuns em 2 indivíduos, 2 haplótipos em 3 indivíduos, 2 haplótipos em 4 indivíduos, 2 haplótipos em 8 indivíduos, 1 haplótipo em 5, 6, 7, 10, 11 e em 13 indivíduos. Os haplótipos 16-17 e 17-17

foram os mais frequentes e comuns em 32 e 27 indivíduos, respectivamente (Fig. 4.3). A diversidade haplotípica na população global obteve o valor de 0.996933187295, o que traduz o elevado polimorfismo do haplótipo identificado com este grupo de marcadores do cromossoma Y.

Na Tabela 4.13 estão apresentadas as frequências alélicas (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) e haplotípicas do *locus* DYS385 referente a população Angolana e a respectiva diversidade génica.

Tabela 4.13 – Frequência alélica, diversidade génica e haplotípica dos 11 *loci* do cromossoma Y na população angolana.

Alelos	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplótipos	DYS385
9					0.006				0.024		10-11	0.006
10					**0.723**	0.012			0.048	0.030	11-11	0.048
11		0.006			0.259	**0.922**	0.012		**0.801**	0.386	11-13	0.012
12		0.060			0.012	0.012	0.012	0.006	0.127	**0.452**	11-14	0.030
13	0.012	**0.771**				0.048	**0.614**	0.030		0.133	12-13	0.012
14	0.084	0.157					0.205	**0.873**			14-15	0.006
15	**0.584**	0.006				0.006	0.145	0.078			14-16	0.006
16	0.181						0.012	0.012			14-17	0.006
17	0.139										14-18	0.006
20				0.018							14-19	0.012
21				**0.813**							14-21	0.012
22				0.036							15-15	0.042
23				0.036							15-16	0.024
24				0.066							15-17	0.060
25				0.024							15-18	0.036
26			0.006	0.006							15-19	0.006
27											15-20	0.024
28			0.024								16-16	0.048
29			0.145								16-17	**0.193**
30			0.343								16-18	0.066
31			**0.392**								16-19	0.042
32			0.078								17-17	0.163
33			0.012								17-18	0.078
											17-19	0.018
											17-20	0.006
											17-21	0.006
											18-18	0.018
											18-19	0.006
											18-20	0.006
DG	0.5994	0.3772	0.7017	0.3317	0.4102	0.1484	0.5592	0.2387	0.3392	0.6376	-	-

DG: Diversidade génica

Os *loci* DYS19 e DYS389 II, DYS393 e DYS439 apresentaram valor de diversidade génica superior à 0.5, os restantes *loci* apresentaram valores mais baixos.

Na Tabela 4.14 estão representados a distribuição haplotípicas do cromossoma Y da população Angolana.

Tabela 4.14 – Haplótipos do cromossoma Y na população Angolana (N=166).

Haplótipos	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H001 ANG	15	12	29	21	10	11	14	14	12	10	17-17
H002 ANG	14	12	28	26	10	11	13	14	11	13	14-21
H003 ANG	15	13	30	21	10	11	13	14	11	10	17-17
H004 ANG	16	13	30	22	10	11	14	14	11	10	17-17
H005 ANG	15	13	29	21	10	11	13	14	11	11	16-17
H006 ANG	17	13	29	21	10	11	15	14	12	11	18-18
H007 ANG	15	13	30	21	10	11	13	14	11	11	16-17
H008 ANG	17	13	30	21	10	11	14	14	11	11	17-17
H009 ANG	15	13	31	21	10	11	13	14	11	11	15-18
H010 ANG	15	13	31	21	10	11	14	14	11	11	17-18
H011 ANG	15	13	31	21	10	11	15	14	11	11	15-15
H012 ANG	16	13	29	21	10	11	15	14	11	12	16-19
H 013ANG	16	13	29	22	10	11	16	14	11	12	17-21
H 014ANG	17	13	29	21	10	11	14	14	11	12	15-17
H015 ANG	14	13	29	23	10	11	13	14	12	12	17-19
H016 ANG	14	13	29	24	10	15	13	14	12	12	11-11
H017 ANG	17	13	30	21	10	11	13	12	11	12	17-17
H018 ANG	15	13	30	21	10	11	15	13	11	12	17-17
H019 ANG	15	13	30	21	10	11	13	14	11	12	16-19
H020 ANG	15	13	30	21	10	11	14	14	11	12	15-16
H021 ANG	15	13	30	21	10	11	15	14	11	12	16-18
H022 ANG	15	13	30	21	10	11	15	14	11	12	17-17
H023 ANG	16	13	30	21	10	11	15	14	12	12	17-18
H024 ANG	15	13	31	21	10	11	13	14	11	12	15-15
H025 ANG	15	13	31	21	10	11	13	14	11	12	15-15
H026 ANG	15	13	31	21	10	11	13	14	11	12	15-20
H027ANG	16	13	31	21	10	11	13	14	11	12	15-15
H028 ANG	15	13	31	21	10	11	13	14	12	12	16-16
H029 ANG	15	13	32	21	10	12	13	14	11	12	16-17
H030 ABK	15	13	30	20	10	11	13	14	11	13	17-17
H031 ANG	15	13	30	21	10	11	14	14	11	13	17-17
H032 ANG	16	13	30	21	10	11	15	14	11	13	17-18
H033 ANG	16	13	30	21	10	11	15	14	11	13	17-19
H 034ANG	17	13	30	21	10	11	15	14	11	13	18-18
H035 ANG	16	13	31	21	10	11	15	14	11	13	17-18
H036 ANG	15	13	31	21	10	11	13	14	12	13	16-17

(continua)

(continuação)

H037 ANG	15	14	31	21	10	11	14	14	11	11	16-18
H038 ANG	17	14	31	21	10	11	15	14	11	11	17-18
H039 ANG	17	14	31	21	10	11	15	14	11	11	17-19
H 040 ANG	17	14	32	21	10	11	13	14	11	11	14-16
H041 ANG	16	14	31	21	10	11	13	14	11	12	17-18
H042 ANG	16	14	31	21	10	11	15	14	11	12	17-18
H043 ANG	17	14	31	21	10	11	14	14	11	12	16-16
H044 ANG	15	14	32	24	10	11	13	15	10	12	11-11
H045 ANG	16	14	32	24	10	11	13	15	10	12	11-11
H046 ANG	17	14	33	23	10	11	13	15	10	12	10-11
H047 ANG	15	14	32	21	10	11	13	14	11	13	16-17
H048 ANG	15	14	32	21	10	11	13	14	11	13	16-17
H049 ANG	14	12	28	25	11	13	13	14	12	12	11-14
H050 ANG	15	13	30	21	11	11	13	14	11	11	17-17
H051 ANG	15	13	30	21	11	11	13	14	11	11	17-17
H052 ANG	15	13	31	21	11	11	13	14	11	11	15-17
H053 ANG	15	13	31	21	11	11	13	14	11	11	16-17
H054 ANG	15	13	31	21	11	11	13	14	11	11	16-17
H055 ANG	14	13	29	25	11	13	13	14	12	12	11-14
H056 ANG	14	13	29	24	11	11	13	15	12	13	11-14
H057 ANG	16	13	30	21	12	11	14	13	11	13	16-17
H058 ANG	15	13	30	21	9	11	13	14	11	11	16-18
H059 ANG	15	12	29	21	10	11	14	14	12	10	17-17
H060 ANG	15	12	29	21	10	11	14	14	12	10	17-17
H061 ANG	15	13	29	21	10	11	13	14	11	11	16-17
H062 ANG	16	13	30	21	10	11	14	13	11	11	17-18
H063 ANG	14	13	30	21	10	11	14	14	11	11	15-16
H064 ANG	16	13	30	21	10	11	14	14	11	11	15-20
H065 ANG	17	13	30	21	10	11	15	14	11	11	17-18
H066 ANG	15	13	31	21	10	11	13	14	11	11	15-17
H067 ANG	15	13	31	21	10	11	13	14	11	11	15-18
H068 ANG	15	13	31	21	10	11	13	14	11	11	16-19
H069 ANG	15	13	32	21	10	11	13	14	11	11	16-17
H070 ANG	15	13	29	21	10	11	15	14	11	12	18-20
H071 ANG	16	13	29	21	10	11	16	14	11	12	16-19
H072 ANG	17	13	30	21	10	11	13	13	11	12	17-17
H073 ANG	15	13	30	21	10	11	13	14	11	12	16-17
H074 ANG	15	13	30	21	10	11	14	14	11	12	15-17
H075 ANG	15	13	30	21	10	11	14	14	11	12	16-16
H076 ANG	15	13	30	21	10	11	14	14	11	12	18-19
H077 ANG	16	13	30	21	10	11	14	14	11	12	17-18
H078 ANG	16	13	30	21	10	13	14	14	11	12	16-17
H079 ANG	17	13	30	21	10	11	13	14	11	12	17-17
H080 ANG	17	13	30	21	10	11	14	14	11	12	17-20
H081 ANG	17	13	30	21	10	11	15	14	11	12	15-20

(continua)

(continuação)

H082 ANG	17	13	30	21	10	11	15	14	11	12	17-17
H083 ANG	15	13	31	22	10	10	11	15	9	12	12-13
H084 ANG	15	13	31	24	10	11	13	15	10	12	11-11
H085 ANG	16	13	31	24	10	11	13	15	10	12	11-11
H086 ANG	15	13	31	21	10	11	13	14	11	12	15-15
H087 ANG	16	13	31	21	10	11	13	14	11	12	15-15
H088 ANG	15	13	31	21	10	11	13	14	12	12	15-18
H089 ANG	15	13	33	21	10	12	13	14	11	12	15-17
H090 ANG	15	13	30	21	10	11	14	14	11	13	16-18
H091 ANG	16	13	31	22	10	10	11	16	9	13	12-13
H092 ANG	15	13	31	21	10	11	13	14	11	13	14-19
H093 ANG	15	13	31	21	10	11	13	14	11	13	15-18
H094 ANG	15	13	31	21	10	11	13	14	11	13	15-20
H095 ANG	15	14	31	21	10	11	13	14	11	11	16-17
H096 ANG	16	14	30	21	10	11	15	14	11	12	16-19
H097 ANG	15	14	31	20	10	11	13	14	12	12	16-16
H098 ANG	15	14	32	21	10	11	13	14	12	12	16-17
H099 ANG	17	15	32	21	10	11	14	14	11	12	16-17
H100 ANG	15	12	30	21	11	11	13	14	11	11	17-17
H101 ANG	15	13	29	21	11	11	13	14	11	11	17-17
H102 ANG	15	13	30	21	11	11	13	14	11	11	15-17
H103 ANG	15	13	30	21	11	11	13	14	11	11	16-17
H104 ANG	15	13	30	21	11	11	13	14	11	11	16-17
H105 ANG	15	13	30	21	11	11	13	14	11	11	17-17
H106 ANG	15	13	31	21	11	11	13	14	11	11	16-17
H107 ANG	15	13	31	21	11	11	13	14	11	11	17-17
H108 ANG	15	13	31	21	11	11	13	14	11	11	17-17
H109 ANG	15	13	31	21	11	11	14	14	11	11	15-17
H110 ANG	15	13	31	21	11	11	13	14	11	12	16-17
H111 ANG	14	13	29	23	11	13	13	15	12	13	11-14
H112 ANG	15	14	32	21	11	11	13	14	11	11	16-17
H113 ANG	14	14	28	24	11	13	13	15	12	13	11-13
H114 ANG	13	10	26	22	10	13	12	14	9	12	15-15
H115 ANG	14	12	29	23	10	11	13	14	11	11	14-19
H116 ANG	16	12	29	21	10	11	13	14	11	12	15-16
H117 AOV	16	13	29	21	10	11	14	14	11	11	16-18
H118 ANG	15	13	30	21	10	11	13	14	11	11	16-17
H119 ANG	15	13	30	21	10	11	13	14	11	11	16-17
H120 ANG	15	13	30	21	10	11	13	14	11	11	16-17
H121 ANG	15	13	31	21	10	11	13	14	11	11	16-16
H122 AOV	15	13	31	21	10	11	13	14	11	11	16-17
H123 ANG	16	13	30	21	10	11	15	13	11	12	16-18
H124 ANG	15	13	30	21	10	11	13	14	11	12	14-15
H125 ANG	15	13	30	21	10	11	13	14	11	12	15-16
H126 ANG	16	13	30	21	10	11	14	14	11	12	17-18

(continua)

(continuação)

H127 ANG	17	13	30	21	10	11	14	14	11	12	17-17
H128 ANG	17	13	30	21	10	11	15	14	11	12	17-17
H129 ANG	17	13	30	21	10	11	15	14	11	12	18-18
H130 ANG	16	13	31	24	10	11	13	15	10	12	11-11
H131 ANG	15	13	31	21	10	11	13	14	11	12	16-17
H132 ANG	15	13	31	24	10	11	13	14	11	12	11-11
H133 ANG	16	13	31	21	10	11	13	14	11	12	16-18
H134 ANG	16	13	31	21	10	11	13	14	11	12	16-18
H135 ANG	17	13	31	21	10	11	14	14	11	12	17-17
H136 ANG	17	13	31	21	10	11	14	14	11	12	17-18
H137 ANG	13	13	32	24	10	11	13	14	10	13	16-16
H138 ANG	15	14	31	22	10	11	13	14	11	11	16-18
H139 ANG	16	14	31	21	10	11	15	14	11	11	16-18
H140 ANG	17	14	31	21	10	11	14	14	12	11	16-19
H141 ANG	16	14	31	23	10	11	13	15	10	12	11-11
H142 ANG	15	14	31	20	10	11	13	14	11	12	15-18
H143 ANG	15	14	31	21	10	11	14	14	11	12	16-16
H144 ANG	15	14	31	21	10	11	14	14	12	12	14-17
H145 ANG	14	12	28	25	11	11	13	14	11	11	14-21
H146 ANG	15	12	29	21	11	11	13	14	11	11	16-17
H147 ANG	14	13	29	25	11	11	13	14	11	11	14-18
H148 ANG	14	13	29	23	11	13	13	15	12	11	11-14
H149 ANG	15	13	30	21	11	11	13	14	11	11	16-17
H150 ANG	15	13	30	21	11	11	13	14	11	11	16-17
H151 ANG	15	13	30	21	11	11	13	14	11	11	17-17
H152 ANG	15	13	31	21	11	11	13	14	11	11	15-17
H153 ANG	15	13	31	21	11	11	13	14	11	11	16-16
H154 ANG	15	13	31	21	11	11	13	14	11	11	16-17
H155 ANG	15	13	31	21	11	11	13	14	11	11	16-19
H156 ANG	15	13	31	21	11	11	14	14	11	11	15-17
H157 ANG	15	13	31	21	11	11	13	15	11	11	17-17
H158 ANG	15	13	32	21	11	11	13	14	11	11	15-18
H159 ANG	14	13	29	24	11	13	13	16	12	12	11-13
H160 ANG	15	13	31	21	11	11	13	14	9	12	15-17
H161 ANG	15	13	31	21	11	11	13	14	11	12	16-17
H162 ANG	15	13	31	21	11	11	13	14	11	12	16-18
H163 ANG	15	13	30	21	11	11	15	14	11	13	15-19
H164 ANG	15	13	31	21	11	11	13	14	11	13	16-17
H165 ANG	15	14	32	21	11	11	12	14	11	11	17-18
H166 ANG	15	13	30	21	12	11	14	14	11	11	17-17

Nos marcadores DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 e DYS385 observou-se, a partir dos 166 indivíduos do sexo masculino da população angolana, que 138 haplótipos eram diferentes, sendo 120 únicos, e 12 haplótipos diferentes comuns em 2 indivíduos cada, 4 haplótipos

diferentes comuns em 4 indivíduos cada, e 2 haplótipos diferentes comuns em 3 indivíduos cada. A diversidade haplotípica dos 11 *loci,* equivalendo ao PD (Poder de Discriminação) nos STRs autossómicos, apresentou um valor notável de 0.996933187295. Em caso de perícias médico-legais, os estudos do cromossoma Y devem ser associados aos estudos dos STRs autossómicos, para informação adicional.

4.2.2 – Comparações genético-populacionais de STRs do cromossoma Y

A análise comparativa entre a população global com outras populações encontra-se nas Tabelas 4.15 e 4.16 incluindo, matriz de distâncias, valor de P e análise da variância molecular.

Tabela 4.15 – Matriz de distâncias génicas e respectivos valores de P entre populações Angola e outras populações Africanas.

FST / FST P	Angola	Namíbia	Guiné. Bissau	Uganda	Moçambique	Somália	Guiné Equatorial
Angola	*	-0.0064	0.0345	0.2054	0.0231	0.63459	0.0303
Namíbia	0.8291 ± 0.0036	*	0.0289	0.1755	0.0157	0.62371	0.0348
Guine. Bissau	0.0000 ± 0.0000	0.0040 ± 0.0006	*	0.1431	0.0254	0.61900	0.0506
Uganda	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	*	0.1329	0.4719	0.1367
Moçambique	0.0025 ± 0.0005	0.0618 ± 0.0025	0.0010 ± 0.0003	0.0000 ± 0.0000	*	0.6009	0.0186
Somália	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.00000 ± 0.0000	*	0.5692
Guine Equatorial	0.0004 ± 0.0002	0.0062 ± 0.0008	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0114 ± 0.0011	0.0000 ± 0.0000	*

Namíbia: (Fujihara *et al*., 2009); Guine. Bissau: (Rosa *et al*., 2006); Uganda: (Gomes *et al*., 2009); Moçambique: (Alves *et al*., 2003); Somália: (Hallenberg *et al*., 2005); Guine Equatorial: (Arroyo-Pardo *et al*., 2005).

Tabela 4.16 – Análise da variância molecular entre a população Angolana e outras populações Africanas.

Fonte de variação	Soma dos quadrados	Componentes da variância	Percentagem da variância
Entre população	3087.496	3.98113 Va	38.33
Dentro da população	5796.878	6.40539 Vb	61.67
Total	841.247	5.11556	

4.2.3 – Polimorfismos de STRs do cromossoma Y no grupo étnico-linguístico Bakongo

O grupo étnico-linguístico Bakongo representa ¼ da população. Foram analisados os resultados relativos à distribuição dos alelos por cada *locus* (Fig. 4.4), distribuição haplotípica do *locus* DYS385 (Fig. 4.5).

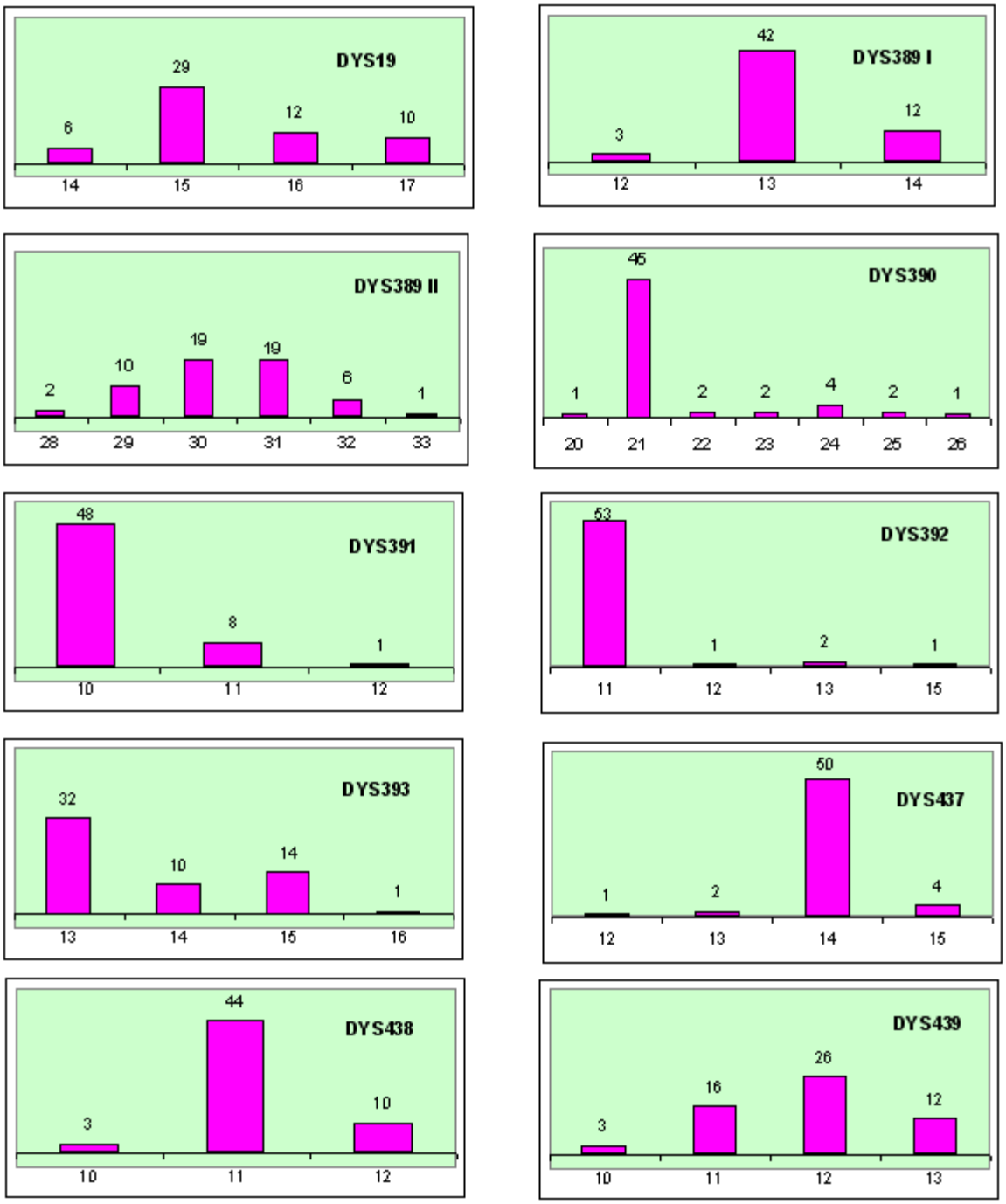


Figura 4.4 – Distribuição dos alelos de cada um dos *loci* de STRs do cromossoma Y no grupo étnico-linguístico Bakongo.

Pela análise da Fig. 4.4, verificou-se que os *loci* de STRs do cromossoma Y com maior número de alelos foram: DYS390 com 7 alelos e o DY389 II com 6 alelos respectivamente. Os *loci* DYS389 I, DYS391 e o DYS438 apresentaram 3 alelos cada, e os restantes *loci* apresentaram 4 alelos cada.

Na Fig. 4.5 está representado a distribuição haplotípica do *locus* DYS385 no grupo étnico-linguístico Bakongo.

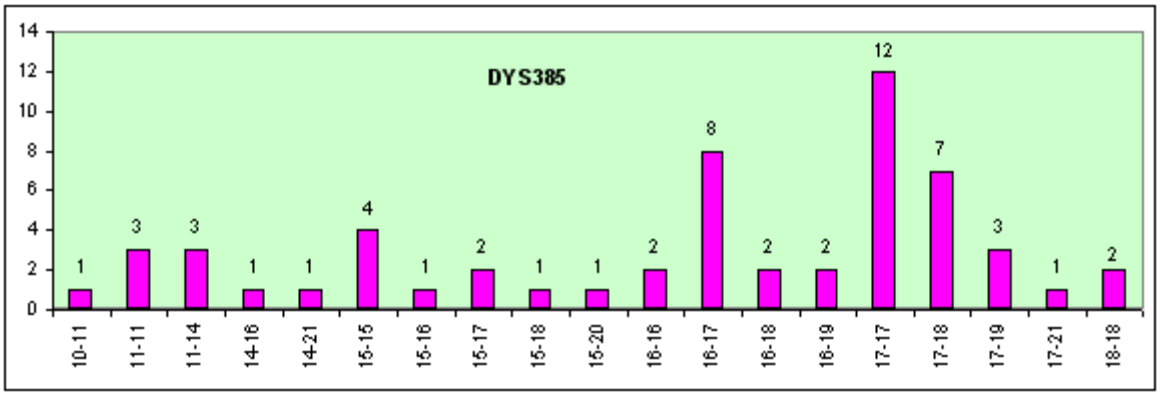


Figura 4.5 – Distribuição dos haplótipos do *locus* DYS385 no grupo étnico-linguístico Bakongo

No marcador DYS385 obtiveram-se 19 haplótipos diferentes em 57 indivíduos do sexo masculino pertencentes ao grupo étnico-linguístico Bakongo, sendo 7 de ocorrência num único indivíduo, 5 haplótipos diferentes comuns a 2 indivíduos, 3 haplótipos diferentes comuns a 3 indivíduos, 2 haplótipos diferentes comuns a 4 e 7 indivíduos. Os haplótipos 17-17 (0.21%) e 16-17 (0.14%) foram os mais frequentes e comuns a 12 e 8 indivíduos, respectivamente (Fig. 4.5).

Na Tabela 4.17 estão apresentadas as frequências alélicas (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) e haplotípicas do *locus* DYS385 referente ao grupo étnico-linguístico Bakongo.

Tabela 4.17 – Frequência alélica, diversidade génica e haplotípica dos *loci* do cromossoma Y no grupo Bakongo (N = 57).

Alelo	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplótipos	DYS385
10					0.842				0.053	0.053	10-11	0.017
11					0.140	0.930			0.772	0.281	11-11	0.053
12		0.053			0.017	0.017		0.017	0.175	0.456	11-14	0.053
13		0.737				0.035	0.561	0.035		0.210	14-16	0.017
14	0.105	0.211					0.175	0.877			14-21	0.017
15	0.509					0.017	0.246	0.070			15-15	0.070
16	0.211						0.017				15-16	0.017
17	0.175										15-17	0.035
20				0.017							15-18	0.017
21				0.789							15-20	0.017
22				0.035							16-16	0.035
23				0.035							16-17	0.140
24				0.070							16-18	0.035
25				0.035							16-19	0.035
26				0.017							17-17	0.210
27											17-18	0.123
28			0.035								17-19	0.053
29			0.175								17-21	0.017
30			0.333								18-18	0.035
31			0.333									
32			0.105									
33			0.017									
DG	0.6557	0.4100	0.7344	0.3675	0.2708	0.1336	0.5934	0.2241	0.3706	0.6661	-	-

DG: Diversidade génica

Os alelos mais frequentes por marcador foram: DYS19 (15), DYS389 I (13), DYS389 II (30 e 31), DYS390 (21), DYS391 (10), DYS392 (11), DYS393 (13), DYS437 (14), DYS438 (11), DYS439 (12) e no marcador DYS385 o haplotipo mais frequente foi 17-17 (Tabela 4.17).

Este grupo étnico-linguístico apresentou uma diversidade génica que varia de 0.1336 no (DYS392) a 0.7344 no *locus* (DYS389 II). A diversidade haplotípica, neste grupo étnico Bakongo, da região Norte de Angola, obteve o valor de 0.9974937343, o que traduz elevado polimorfismo do haplótipo identificado com este grupo de marcadores do cromossoma Y.

Na Tabela 4.18 estão descritos os haplótipos do STRs do cromossoma Y (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS385) do grupo Bakongo e a respectiva diversidade génica.

Tabela 4.18 – Haplótipos de STRs do cromossoma Y do grupo étnico-linguístico Bakongo.

Haplótipo	N	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H01 ABK	1	15	12	29	21	10	11	14	14	12	10	17-17
H02 ABK	1	14	12	28	26	10	11	13	14	11	13	14-21
H03 ABK	1	15	13	30	21	10	11	13	14	11	10	17-17
H04 ABK	1	16	13	30	22	10	11	14	14	11	10	17-17
H05 ABK	1	15	13	29	21	10	11	13	14	11	11	16-17
H06 ABK	1	17	13	29	21	10	11	15	14	12	11	18-18
H07 ABK	1	15	13	30	21	10	11	13	14	11	11	16-17
H08 ABK	1	17	13	30	21	10	11	14	14	11	11	17-17
H09 ABK	1	15	13	31	21	10	11	13	14	11	11	15-18
H10 ABK	1	15	13	31	21	10	11	14	14	11	11	17-18
H11 ABK	1	15	13	31	21	10	11	15	14	11	11	15-15
H12 ABK	1	16	13	29	21	10	11	15	14	11	12	16-19
H13 ABK	1	16	13	29	22	10	11	16	14	11	12	17-21
H14 ABK	1	17	13	29	21	10	11	14	14	11	12	15-17
H15 ABK	1	14	13	29	23	10	11	13	14	12	12	17-19
H16 ABK	1	14	13	29	24	10	15	13	14	12	12	11-11
H17 ABK	1	17	13	30	21	10	11	13	12	11	12	17-17
H18 ABK	1	15	13	30	21	10	11	15	13	11	12	17-17
H19 ABK	1	15	13	30	21	10	11	13	14	11	12	16-19
H20 ABK	1	15	13	30	21	10	11	14	14	11	12	15-16
H21 ABK	1	15	13	30	21	10	11	15	14	11	12	16-18
H22 ABK	1	15	13	30	21	10	11	15	14	11	12	17-17
H23 ABK	1	16	13	30	21	10	11	15	14	12	12	17-18
H24 ABK	2	15	13	31	21	10	11	13	14	11	12	15-15
H25 ABK	1	15	13	31	21	10	11	13	14	11	12	15-20
H26 ABK	1	16	13	31	21	10	11	13	14	11	12	15-15
H27 ABK	1	15	13	31	21	10	11	13	14	12	12	16-16
H28 ABK	1	15	13	32	21	10	12	13	14	11	12	16-17
H29 ABK	1	15	13	30	20	10	11	13	14	11	13	17-17
H30 ABK	1	15	13	30	21	10	11	14	14	11	13	17-17
H31 ABK	1	16	13	30	21	10	11	15	14	11	13	17-18
H32 ABK	1	16	13	30	21	10	11	15	14	11	13	17-19
H33 ABK	1	17	13	30	21	10	11	15	14	11	13	18-18
H34 ABK	1	16	13	31	21	10	11	15	14	11	13	17-18
H35 ABK	1	15	13	31	21	10	11	13	14	12	13	16-17
H36 ABK	1	15	14	31	21	10	11	14	14	11	11	16-18
H37 ABK	1	17	14	31	21	10	11	15	14	11	11	17-18

(continua)

(continuação)

H38 ABK	1	17	14	31	21	10	11	15	14	11	11	17-19
H39 ABK	1	17	14	32	21	10	11	13	14	11	11	14-16
H40 ABK	1	16	14	31	21	10	11	13	14	11	12	17-18
H41 ABK	1	16	14	31	21	10	11	15	14	11	12	17-18
H42 ABK	1	17	14	31	21	10	11	14	14	11	12	16-16
H43 ABK	1	15	14	32	24	10	11	13	15	10	12	11-11
H44 ABK	1	16	14	32	24	10	11	13	15	10	12	11-11
H45 ABK	1	17	14	33	23	10	11	13	15	10	12	10-11
H46ABK	2	15	14	32	21	10	11	13	14	11	13	16-17
H47 ABK	1	14	12	28	25	11	13	13	14	12	12	11-14
H48 ABK	2	15	13	30	21	11	11	13	14	11	11	17-17
H49 ABK	1	15	13	31	21	11	11	13	14	11	11	15-17
H50ABK	2	15	13	31	21	11	11	13	14	11	11	16-17
H51 ABK	1	14	13	29	25	11	13	13	14	12	12	11-14
H52 ABK	1	14	13	29	24	11	11	13	15	12	13	11-14
H53 ABK	1	16	13	30	21	12	11	14	13	11	13	16-17

Na Tabela 4.18 estão descritos os marcadores DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 e no marcador DYS385, observou-se, a partir dos 57 indivíduos do sexo masculino pertencentes ao grupo étnico-linguístico Bakongo, que 53 haplótipos eram diferentes, sendo 49 únicos e 4 haplótipos diferentes comuns a 2 indivíduos cada (H24 ABK, H46 ABK, H48 ABK e H50 ABK).

4.2.4 – Polimorfismos de STRs do cromossoma Y no grupo étnico-linguístico Kimbundo

O grupo étnico-linguístico Kimbundo representa ¼ da população e foram analisadas os resultados relativos à distribuição dos alelos por cada *locus* do cromossoma Y (Fig. 4.6) distribuição haplotípica do *locus* DYS385 (Fig. 4.7).

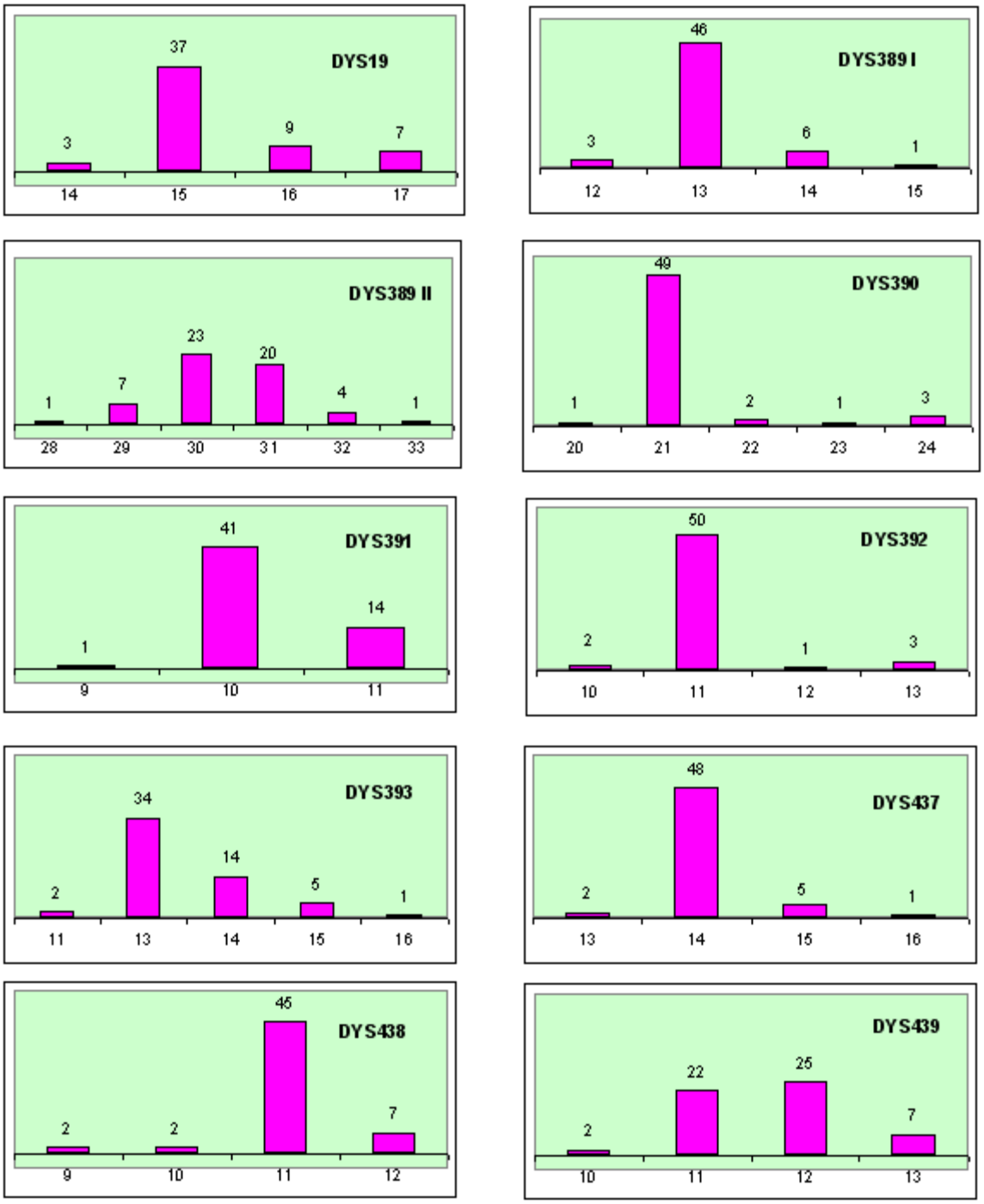


Figura 4.6 – Distribuição dos alelos de cada um dos *loci* de STRs do cromossoma Y no grupo étnico-linguístico Kimbundo.

Pela análise da Fig. 4.6 verificou-se que os *loci* de STRs do cromossoma Y com maior número de alelos foram: o DY389 II com 6 alelos, seguido dos *loci* DYS390 e DYS393 com 5 alelos cada e os restantes apresentam em média 4 alelos.

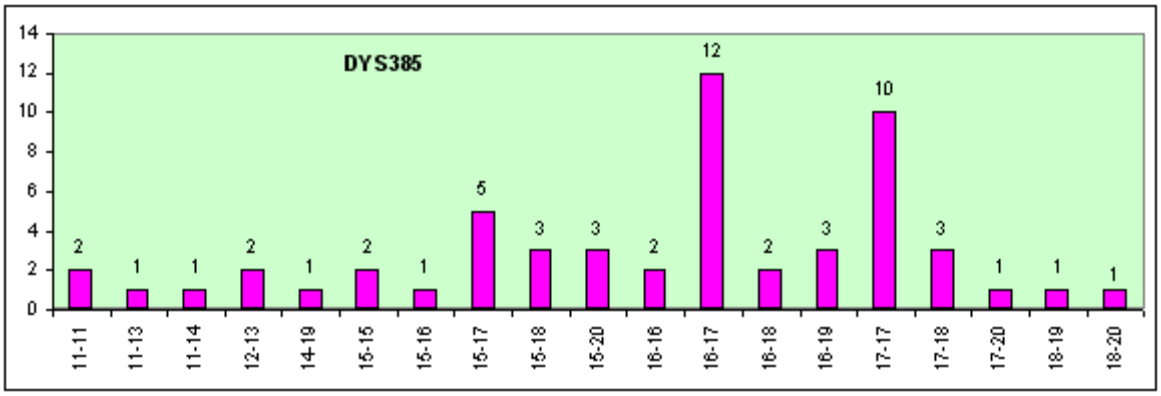


Figura 4.7 - Distribuição dos haplótipos do *locus* DYS385 no grupo étnico-linguístico Kimbundo

No marcador DYS385 obtiveram-se 19 haplótipos diferentes em 57 indivíduos do sexo masculino pertencentes ao grupo étnico-linguístico Kimbundo, sendo 7 de ocorrência num único indivíduo, 5 haplótipos diferentes comuns a 2 indivíduos, 4 haplótipos diferentes comuns a 3 indivíduos, 1 haplótipos de ocorrência num único indivíduo. Os haplótipos 16-17 (21.4%) e 17-17 (17.9%) foram os mais frequentes e comuns a 12 e 10 indivíduos, respectivamente (Fig. 4.7).

Na Tabela 4.19 estão apresentadas as frequências alélicas (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) e haplotípicas do *locus* DYS385 referente ao grupo étnico-linguístico Kimbundo e a respectiva diversidade génica.

Tabela 4.19 – Frequência alélica, diversidade génica e haplotipica dos *loci do* cromossoma Y no grupo Kimbundo (N = 56).

Alelos	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplótipos	DYS385
9					0.018				0.036		11-11	0.036
10					0.732	0.036			0.036	0.036	11-13	0.018
11					0.250	0.893	0.036		0.804	0.393	11-14	0.018
12		0.054				0.018			0.125	0.446	12-13	0.036
13		0.821				0.054	0.607	0.036		0.125	14-19	0.018
14	0.054	0.107					0.250	0.857			15-15	0.036
15	0.661	0.018					0.089	0.089			15-16	0.018
16	0.161						0.018	0.018			15-17	0.089
17	0.125										15-18	0.054
20				0.018							15-20	0.054
21				0.875							16-16	0.036
22				0.036							16-17	0.214
23				0.018							16-18	0.036
24				0.054							16-19	0.054
28			0.018								17-17	0.179
29			0.518								17-18	0.054
30			0.411								17-20	0.018
31			0.357								18-19	0.018
32			0.071								18-20	0.018
33			0.018									
DG	0.5191	0.3106	0.6824	0.2296	0.4012	0.1983	0.5593	0.2557	0.3361	0.6295	-	-

DG: Diversidade génica

Os alelos mais frequentes por marcador foram os seguintes: DYS19 (15), DYS389 I (13), DYS389 II (30), DYS390 (21), DYS391 (10), DYS392 (11), DYS393 (13), DYS437 (14), DYS438 (11), DYS439 (12) (Fig. 20) e no marcador DYS385 o haplotipo mais frequente foi 16-17 (Tabela 4.19).

Este grupo étnico-linguístico Kimbundo apresentou uma diversidade génica que varia de 0.1984 no DYS392 a 0.6824 no *locus* DYS389 II (Tabela 4.19). A diversidade haplotipica, neste grupo étnico-linguístico Kimbundo, da região Centro Norte de Angola, obteve o valor de 0.9980519481, o que traduz o elevado polimorfismo do haplótipo identificado com este grupo de marcadores do cromossoma Y.

Na Tabela 4.20 estão descritos os haplótipos do STRs do cromossoma Y (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS385) do grupo Kimbundo.

Tabela 4.20 – Haplótipos de STRs do cromossoma Y do grupo étnico-linguístico Kimbundo.

Haplótipos	N	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H01 AKI	1	15	13	30	21	9	11	13	14	11	11	16-18
H02 AKI	2	15	12	29	21	10	11	14	14	12	10	17-17
H03 AKI	1	15	13	29	21	10	11	13	14	11	11	16-17
H04 AKI	1	16	13	30	21	10	11	14	13	11	11	17-18
H05 AKI	1	14	13	30	21	10	11	14	14	11	11	15-16
H06 AKI	1	16	13	30	21	10	11	14	14	11	11	15-20
H07AKI	1	17	13	30	21	10	11	15	14	11	11	17-18
H08 AKI	1	15	13	31	21	10	11	13	14	11	11	15-17
H09 AKI	1	15	13	31	21	10	11	13	14	11	11	15-18
H10 AKI	1	15	13	31	21	10	11	13	14	11	11	16-19
H11 AKI	1	15	13	32	21	10	11	13	14	11	11	16-17
H12 AKI	1	15	13	29	21	10	11	15	14	11	12	18-20
H13 AKI	1	16	13	29	21	10	11	16	14	11	12	16-19
H14 AKI	1	17	13	30	21	10	11	13	13	11	12	17-17
H15 AKI	1	15	13	30	21	10	11	13	14	11	12	16-17
H16 AKI	1	15	13	30	21	10	11	14	14	11	12	15-17
H17 AKI	1	15	13	30	21	10	11	14	14	11	12	16-16
H18 AKI	1	15	13	30	21	10	11	14	14	11	12	18-19
H19 AKI	1	16	13	30	21	10	11	14	14	11	12	17-18
H20 AKI	1	16	13	30	21	10	13	14	14	11	12	16-17
H21 AKI	1	17	13	30	21	10	11	13	14	11	12	17-17
H22 AKI	1	17	13	30	21	10	11	14	14	11	12	17-20
H23 AKI	1	17	13	30	21	10	11	15	14	11	12	15-20
H24 AKI	1	17	13	30	21	10	11	15	14	11	12	17-17
H25 AKI	1	15	13	31	22	10	10	11	15	9	12	12-13
H26 AKI	1	15	13	31	24	10	11	13	15	10	12	11-11
H27 AKI	1	16	13	31	24	10	11	13	15	10	12	11-11
H28 AKI	1	15	13	31	21	10	11	13	14	11	12	15-15
H29 AKI	1	16	13	31	21	10	11	13	14	11	12	15-15
H30 AKI	1	15	13	31	21	10	11	13	14	12	12	15-18
H31 AKI	1	15	13	33	21	10	12	13	14	11	12	15-17
H32 AKI	1	15	13	30	21	10	11	14	14	11	13	16-18
H33 AKI	1	16	13	31	22	10	10	11	16	9	13	12-13
H34 AKI	1	15	13	31	21	10	11	13	14	11	13	14-19
H35 AKI	1	15	13	31	21	10	11	13	14	11	13	15-18
H36 AKI	1	15	13	31	21	10	11	13	14	11	13	15-20
H37 AKI	1	15	14	31	21	10	11	13	14	11	11	16-17
H38 AKI	1	16	14	30	21	10	11	15	14	11	12	16-19

(continua)

(continuação)

H39 AKI	1	15	14	31	20	10	11	13	14	12	12	16-16
H40 AKI	1	15	14	32	21	10	11	13	14	12	12	16-17
H41 AKI	1	17	15	32	21	10	11	14	14	11	12	16-17
H42 AKI	1	15	12	30	21	11	11	13	14	11	11	17-17
H43 AKI	1	15	13	29	21	11	11	13	14	11	11	17-17
H44 AKI	1	15	13	30	21	11	11	13	14	11	11	15-17
H45 AKI	2	15	13	30	21	11	11	13	14	11	11	16-17
H46 AKI	1	15	13	30	21	11	11	13	14	11	11	17-17
H47 AKI	1	15	13	31	21	11	11	13	14	11	11	16-17
H48 AKI	2	15	13	31	21	11	11	13	14	11	11	17-17
H49 AKI	1	15	13	31	21	11	11	14	14	11	11	15-17
H50 AKI	1	15	13	31	21	11	11	13	14	11	12	16-17
H51 AKI	1	14	13	29	23	11	13	13	15	12	13	11-14
H52 AKI	1	15	14	32	21	11	11	13	14	11	11	16-17
H53 AKI	1	14	14	28	24	11	13	13	15	12	13	11-13

Nos marcadores DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 e no marcador DYS385, observou-se a partir dos 56 indivíduos do sexo masculino pertencentes ao grupo étnico-linguístico Kimbundo, que 53 haplótipos eram diferentes, dos quais 50 haplótipos eram únicos e 3 haplótipos diferentes, comuns a 2 indivíduos cada. Os Haplótipos mais frequentes foram (H02 AKI, H45 AKI e o H48 AKI).

4.2.5 – Polimorfismos de STRs do cromossoma Y no grupo étnico-linguístico Ovimbundo

O grupo étnico-linguístico Ovimbundo representa ¼ da população e foram analisadas os resultados relativos à distribuição dos alelos por cada *locus* do cromossoma Y (Fig. 4.8), distribuição haplotipica do *locus* DYS385 (Fig. 4.9).

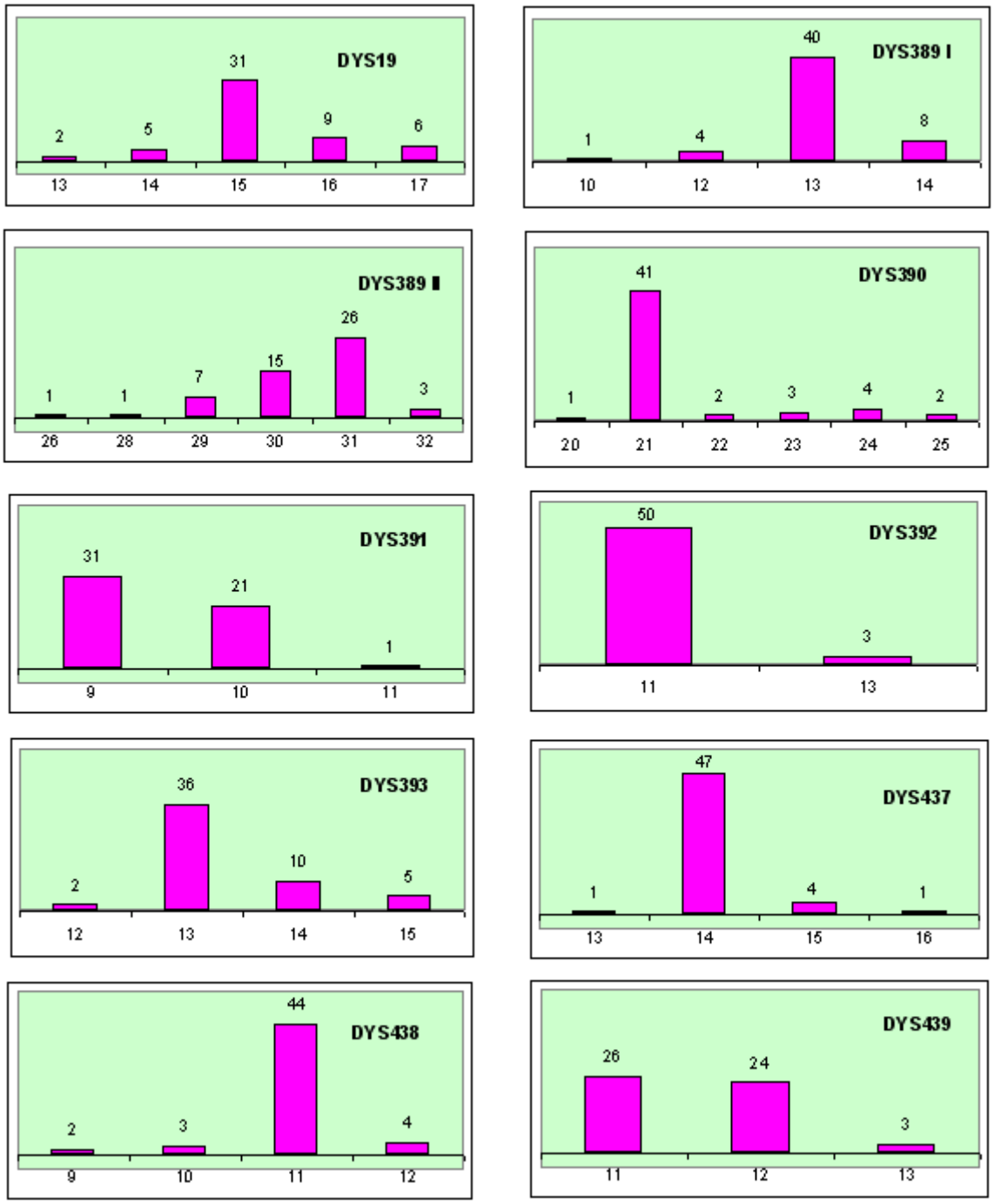


Figura 4.8 - Distribuição dos alelos de cada um dos *loci* de STRs do cromossoma Y no grupo étnico-linguístico Ovimbundo.

Pela análise da Fig. 4.8 constatou-se que os *loci* de STRs do cromossoma Y com maior número de alelos foram: o DY389 II e DYS390 com 6 alelos cada, DYS19 com 5 alelos e os restantes apresentaram em média 3 alelos cada.

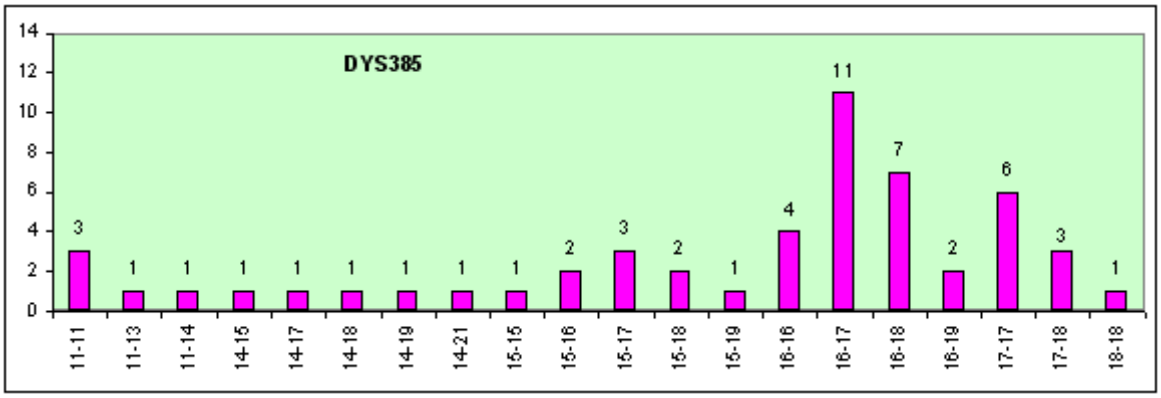


Figura 4.9 – Distribuição dos haplótipos do *locus* DYS385 no grupo étnico-linguístico Ovimbundo.

No marcador DYS385 obtiveram-se 20 haplótipos diferentes em 53 indivíduos do sexo masculino pertencentes ao grupo étnico-linguístico Ovimbundo, sendo 10 de ocorrência em um único indivíduo, 3 haplótipos diferentes comuns em 2 indivíduos, 3 haplótipos diferentes comuns em 3 indivíduos, 1 haplótipo comum em 4 indivíduos e 1 haplótipo comum em 6 indivíduos. Os haplótipos 16-17 (0.206%) e 16-18 (0.132%) foram os mais frequentes e comuns em 11 e 7 indivíduos, respectivamente (Fig.4.9).

Na Tabela 4.21 estão apresentadas as frequências alélicas (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439) e haplotípicas do *locus* DYS385 referente ao grupo étnico-linguístico Ovimbundo e a respectiva diversidade génica.

Tabela 4.21 – Frequência alélica, diversidade génica e haplotipica dos *loci* do cromossoma Y no grupo Ovimbundo (N = 53).

alelos	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplótipos	DYS385
9					0.585				0.038		11-11	0.057
10		0.019			0.396				0.057		11-13	0.019
11					0.019	0.943			0.830	0.491	11-14	0.019
12		0.075							0.075	0.453	14-15	0.019
13	0.038	0.755				0.057	0.679	0.019		0.057	14-17	0.019
14	0.094	0.151					0.189	0.887			14-18	0.019
15	0.585						0.094	0.075			14-19	0.019
16	0.161							0.019			14-21	0.019
17	0.113										15-15	0.019
20				0.019							15-16	0.038
21				0.774							15-17	0.057
22				0.038							15-18	0.038
23				0.057							15-19	0.019
24				0.075							16-16	0.075
25				0.038							16-17	0.207
26			0.019								16-18	0.132
28			0.019								16-19	0.038
29			0.132								17-17	0.113
30			0.283								17-18	0.057
31			0.491								18-18	0.019
32			0.057									
DG	0.6059	0.4016	0.6588	0.3895	0.5005	0.1068	0.49**30**	0.,2072	0.3005	0.5511	-	-

DG: Diversidade génica

Os alelos mais frequentes por marcador foram os seguintes: DYS19 (15), DYS389 I (13), DYS389 II (31), DYS390 (21), DYS391 (15), DYS392 (11), DYS393 (13), DYS437 (14), DYS438 (11), DYS439 (11) (Fig. 4.20).

Este grupo étnico-linguístico Ovimbundo apresentou uma diversidade génica que varia de 0.1068 no (DYS392) a 0.6059 nos *loci* (DYS19 e DYS391). A diversidade haplotipica, neste grupo étnico Ovimbundo obteve o valor de 0.996371552975, o que traduz o elevado polimorfismo do haplótipo identificado com este grupo de marcadores do cromossoma Y.

Na Tabela 4.22 estão descritos os haplótipos do STRs do cromossoma Y (DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS385) do grupo Ovimbundo.

Tabela 4.22 – Haplótipos de STRs do cromossoma Y do grupo étnico-linguístico Ovimbundo.

Haplótipos	N	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H01 AOV	1	13	10	26	22	10	13	12	14	9	12	15-15
H02 AOV	1	14	12	29	23	10	11	13	14	11	11	14-19
H03 AOV	1	16	12	29	21	10	11	13	14	11	12	15-16
H04 AOV	1	16	13	29	21	10	11	14	14	11	11	16-18
H05 AOV	3	15	13	30	21	10	11	13	14	11	11	16-17
H06 AOV	1	15	13	31	21	10	11	13	14	11	11	16-16
H07 AOV	1	15	13	31	21	10	11	13	14	11	11	16-17
H08 AOV	1	16	13	30	21	10	11	15	13	11	12	16-18
H09 AOV	1	15	13	30	21	10	11	13	14	11	12	14-15
H10 AOV	1	15	13	30	21	10	11	13	14	11	12	15-16
H11 AOV	1	16	13	30	21	10	11	14	14	11	12	17-18
H12 AOV	1	17	13	30	21	10	11	14	14	11	12	17-17
H13 AOV	1	17	13	30	21	10	11	15	14	11	12	17-17
H14 AOV	1	17	13	30	21	10	11	15	14	11	12	18-18
H15 AOV	1	16	13	31	24	10	11	13	15	10	12	11-11
H16 AOV	1	15	13	31	21	10	11	13	14	11	12	16-17
H17 AOV	1	15	13	31	24	10	11	13	14	11	12	11-11
H18 AOV	2	16	13	31	21	10	11	13	14	11	12	16-18
H19 AOV	1	17	13	31	21	10	11	14	14	11	12	17-17
H20 AOV	1	17	13	31	21	10	11	14	14	11	12	17-18
H21 AOV	1	13	13	32	24	10	11	13	14	10	13	16-16
H22 AOV	1	15	14	31	22	10	11	13	14	11	11	16-18
H23 AOV	1	16	14	31	21	10	11	15	14	11	11	16-18
H24 AOV	1	17	14	31	21	10	11	14	14	12	11	16-19
H25 AOV	1	16	14	31	23	10	11	13	15	10	12	11-11
H26 AOV	1	15	14	31	20	10	11	13	14	11	12	15-18
H27 AOV	1	15	14	31	21	10	11	14	14	11	12	16-16
H28 AOV	1	15	14	31	21	10	11	14	14	12	12	14-17
H29 AOV	1	14	12	28	25	11	11	13	14	11	11	14-21
H30 AOV	1	15	12	29	21	11	11	13	14	11	11	16-17
H31 AOV	1	14	13	29	25	11	11	13	14	11	11	14-18
H32 AOV	1	14	13	29	23	11	13	13	15	12	11	11-14
H33 AOV	2	15	13	30	21	11	11	13	14	11	11	16-17
H34 AOV	1	15	13	30	21	11	11	13	14	11	11	17-17
H35 AOV	1	15	13	31	21	11	11	13	14	11	11	15-17
H36 AOV	1	15	13	31	21	11	11	13	14	11	11	16-16
H37 AOV	1	15	13	31	21	11	11	13	14	11	11	16-17

(continua)

(continuação)

H38 AOV	1	15	13	31	21	11	11	13	14	11	11	16-19
H39 AOV	1	15	13	31	21	11	11	14	14	11	11	15-17
H40 AOV	1	15	13	31	21	11	11	13	15	11	11	17-17
H41 AOV	1	15	13	32	21	11	11	13	14	11	11	15-18
H42 AOV	1	14	13	29	24	11	13	13	16	12	12	11-13
H43 AOV	1	15	13	31	21	11	11	13	14	9	12	15-17
H44 AOV	1	15	13	31	21	11	11	13	14	11	12	16-17
H45 AOV	1	15	13	31	21	11	11	13	14	11	12	16-18
H46 AOV	1	15	13	30	21	11	11	15	14	11	13	15-19
H47 AOV	1	15	13	31	21	11	11	13	14	11	13	16-17
H48 AOV	1	15	14	32	21	11	11	12	14	11	11	17-18
H49 AOV	1	15	13	30	21	12	11	14	14	11	11	17-17

Nos marcadores DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 e no marcador DYS385, observou-se, a partir dos 53 indivíduos do sexo masculino pertencentes ao grupo étnico-linguístico Ovimbundo, que 49 haplótipos eram diferentes, dos quais 46 haplótipos eram únicos e 1 haplótipo diferente comum a 3 indivíduos e dois haplótipos diferentes comuns a 2 indivíduos cada. O haplótipo mais frequente foi (H05 AOV) e apresentou uma frequência de 5.7%.

4.2.6 – Análise comparativa dos polimorfismos do cromossoma Y entre os três grupos étnico-linguísticos.

Tabela 4.23 – Matriz de distância génicas e respectivos valores de P relativa aos três principais grupos étnico-linguísticos de Angola.

FST / FST P	Bakongo	Kimbundo	Ovimbundo
Bakongo	---	0.0088	0.0201
Kimbundo	0.1658 ± 0.0035	---	-0.0001
Ovimbundo	0.0576 ± 0.0027	0.4112 ± 0.0049	---

Embora os grupos Kimbundo e Ovimbundo estejam mais próximos entre si (FST = -0.00013; P = 0.16583) relativamente aos Bakongo, não há diferenças estatísticas significativas entre eles (valor de P> 0.05).

Na Fig. 4.9 está representada a árvore filogenética do estudo do cromossoma Y relativa aos três grupos étnico-linguísticos

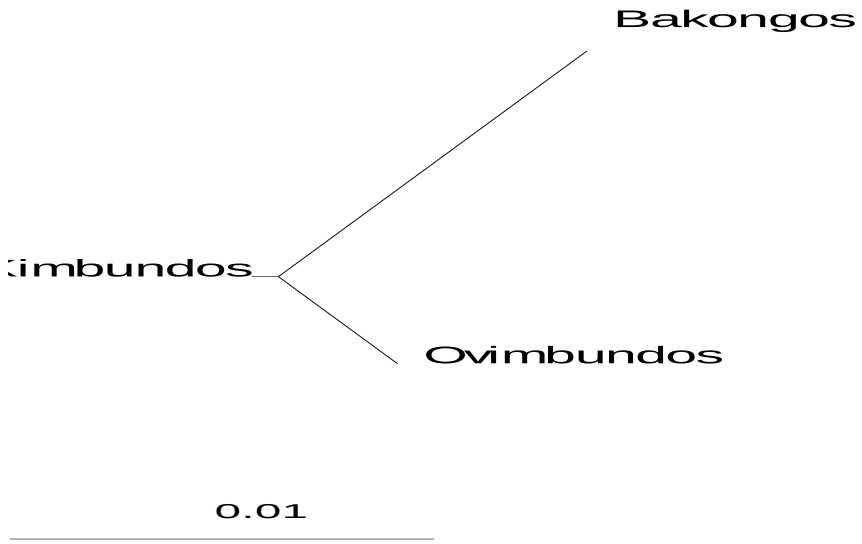


Figura 4.10 - Árvore filogenética relativa aos três principais grupos étnico-linguísticos de Angola.

Os resultados relativos a análise da variância molecular relativo ao estudo do cromossoma Y nos três grupos étnico-linguísticos de Angola estão representados na Tabela 4.24.

Tabela 4.24 - Análise da variância molecular nos três grupos étnico-linguísticos de Angola.

Fonte de variação	**Soma dos quadrados**	**Componentes da variância**	**Percentagem da variância**
Entre população	15.864	0.05186	1.01
Dentro da população	825.383	5.06370	98.99
Total	841.247	5.11556	

A alta diversidade haplotípica observada nos três grupos étnico-linguísticos analisados consolida a alta variabilidade intra-populacional dos haplótipos do STRs do cromossoma Y. O teste de Amova mostra que a população apresenta uma variação maior (98.99%) dentro de cada população do que a variação inter-população.

4.3 – Polimorfismos do ADN mitocondrial

A região controlo do ADNmt comporta cerca de 1122 pb, sendo altamente polimórfica, devido à alta taxa de mutação em comparação com a região codificante. Esta região apresenta maior variabilidade entre os indivíduos e análises dos polimorfismos de sequência têm grande aplicabilidade em estudos populacionais e forenses.

O poder de discriminação apresentado pelo ADNmt deve-se ao polimorfismo natural das regiões hipervariáveis da região controlo. Os haplogrupos do ADNmt são relevantes para o esclarecimento da história e o passado demográfico das populações, que podem reflectir a relação filogenética entre os diversos grupos étnico-linguísticos.

O estudo dos três principais grupos étnico-linguísticos tem como finalidade a caracterização genética da população angolana e a criação de uma base de dados e, posteriormente, a sua aplicação em casos forenses. Procedeu-se à análise de 30 amostras referentes aos grupos Bakongo, Kimbundo e Ovimbundo da população de Angola e, posteriormente, à análise comparativa entre si. Os resultados referentes aos polimorfismos de ADN mitocondrial da população angolana e dos grupos étnico-linguísticos Bakongo, Kimbundo e Ovimbundo encontram-se nas Tabelas 4.25, 4.26, 4.27 e 4.28, respectivamente.

4.3.1 – Polimorfismos do ADN mitocondrial na população global

Tabela 4.25 - Sequência de ADNmt, haplogrupo estimado e haplótipos da população angolana (N=30).

População	Haplogrupo estimado	N	Haplótipos
ANG 1	L3b1a	1	16129A 16213A 16223T 16278T 16355T 16362C 16390A 73G 150T 152C 182T 204C 263G 315.1C 418T
ANG 2	L3e1	1	16223T 16327T 73G 150T 189G 199C 200G 263G 309.1C 315.1C
ANG 3*	L1c3b	1	16129A 16189C 16193.1C 16223T 16274A 16278T 16293G 16294T 16311C 16360T 16519C73G 89C 93G 95C 152C 182T 186A 189C 236C 247A 263G 297G 315.1C 316A
ANG 4	L1c3b	1	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T 16519C73G 151T 152C 186A 189C 247A 263G 315.1C 316A
ANG 5	L3d1a	1	16124C 16223T 16319A 73G 150T 152C 263G 315.1C
ANG 6	L2b	1	16114A 16129A 16213A 16223T 16278T 16355T 16362C 16390A 73G 150T 152C 182T 195C 198T 204C 263G 315.1C 418T
ANG 7*	L1c1	1	16129A 16189C 16193.1C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 93G 151T 152C 182T 186A 189C 195C 247A 248G 263G 315.1C 316A
ANG 8	L3f1b1	1	16129A 16209C 16223T 16292T 16295T 16311C 16519C 73G 152C 189G 263G 272G 315.1C
ANG 9	L1c2	1	16129A 16187T 16189C 16214T 16223T 16265C 16278T 16286A 16291T 16294T 16311C 16360T 16519C 16527T 73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A
ANG 10*	L3e2b*	1	16172C 16183C 16189C 16193.1C 16223T 16320T 16519C 73G 150T 152C 195C 263G 315.1C
ANG 11	L0a1b	1	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16311C 16320T 93G 95C 185A 189G 236C 247A 263G 309.1C 315.1C
ANG 12*	L2a2	1	16169T 16189C 16193.1C 16223T 16229C 16278T 16294T 16311C 16389A 16519C 73G 152C 182T 263G 315.1C
ANG 13	L1c1	1	16093C 16129A 16187T 16189C 16223T 16263C 16278T 16293G 16294T 16311C 16360T 16399G 16519C 73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A
ANG 14	L3f1b	1	16129A 16209C 16223T 16292T 16311C 16519C 73G 189G 200G 263G 272G 309.1C 315.1C
ANG 15*	L2a1	1	16189C 16193.1C 16223T 16278T 16294T 16309G 16390A 16519C 73G 146C 152C 195C 263G 315.1C
ANG 16*	L3e2b	1	16126C 16172C 16182C 16183C 16189C 16193.1C 16223T 16320T 16519C 73G 150T 195C 263G 315.1C
ANG 17	L1c3b	1	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A
ANG 18	L2a1	1	16086C 16155G 16223T 16278T 16294T 16309G 16390A 73G 143A 146C 152C 195C 198T 263G 315.1C
ANG 19	L0a1b	1	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G 16311C 16320T 93G 95C 185A 189G 236C 247A 263G 309.1C 315.1C
ANG 20*	L1b1*	1	16126C 16189C 16193.1C 16223T 16264T 16270T 16278T 16311C 16519C 73G 151T 152C 182T 185T 195C 247A 263G 315.1C 357G
ANG 21	L1b1	1	16126C 16166G 16187T 16189C 16193T 16223T 16264T 16270T 16278T 16293G 16311C 16519C 73G 152C 182T 185T 189G 195C 247A 263G 315.1C 357G
ANG 22	L3e1	1	16223T 16327T 73G 150T 189G 200G 263G 309.1C 315.1C
ANG 23	L1c3b	1	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A
ANG 24	L2b1	1	16114A 16129A 16213A 16223T 16278T 16355T 16362C 16390A 73G 150T 152C 182T 195C 198T 204C 263G 309.1C 315.1C 418T
ANG 25	L0a1b	1	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G 16311C 16320T 93G 95C 185A 189G 236C 247A 263G 309.1C 315.1C
ANG 26	L3b	1	16093C 16124C 16192T 16223T 16278T 16362C 16519C 73G 146C 263G 309.1C 315.1C
ANG 27	L0a2	2	16148T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T 16519C 64T 93G 152C 189G 204C 207A 236C 247A 263G 315.1C
ANG 28	L2	1	16223T 16264T 16278T 16390A 73G 93G 146C 150T 152C 182T 195C 198T 263G 315.1C
ANG 29	L3e2a	1	16223T 16320T 16519C 73G 150T 195C 198T 263G 315.1C

* Heteroplasmia de comprimento na zona poly (C) da região hipervariável I
** Heteroplasmia de comprimento na zona poly (C) da região hipervariável II

4.3.2 – Polimorfismos do ADNmt do grupo étnico-linguístico Bakongo

Tabela 4.26 - Sequência de ADNmt, haplogrupo estimado e haplótipos do grupo étnico linguístico Bakongo (N=10).

População	Haplogrupo estimado	N	Haplótipos
ABK1	L3b1a	1	16129A 16213A 16223T 16278T 16355T 16362C 16390A 73G 150T 152C 182T 204C 263G 315.1C 418T
ABK2	L3e1	1	16223T 16327T 73G 150T 189G 199C 200G 263G 309.1C 315.1C
ABK3*	L1c3b	1	16129A 16189C 16193.1C 16223T 16274A 16278T 16293G 16294T 16311C 16360T 16519C 73G 89C 93G 95C 152C 182T 186A 189C 236C 247A 263G 297G 315.1C 316A
ABK4	L1c3b	1	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 151T 152C 186A 189C 247A 263G 315.1C 316A
ABK5	L3d1a	1	16124C 16223T 16319A 73G 150T 152C 263G 315.1C
ABK6	L2b	1	16114A 16129A 16213A 16223T 16278T 16355T 16362C 16390A 73G 150T 152C 182T 195C 198T 204C 263G 315.1C 418T
ABK7*	L1c1	1	16129A 16189C 16193.1C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 93G 151T 152C 182T 186A 189C 195C 247A 248G 263G 315.1C 316A
ABK8	L3f1b1	1	16129A 16209C 16223T 16292T 16295T 16311C 16519C 73G 152C 189G 263G 272G 315.1C
ABK9	L1c2	1	16129A 16187T 16189C 16214T 16223T 16265C 16278T 16286A 16291T 16294T 16311C 16360T 16519C 16527T 73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A
ABK10*	L3e2b*	1	16172C 16183C 16189C 16193.1C 16223T 16320T 16519C 73G 150T 152C 195C 263G 315.1C

* - Heteroplasmia de comprimento na zona da poly (C) da região hipervariável I
** - Heteroplasmia de comprimento na zona da poly (C) da região hipervariável II.

4.3.3 – Polimorfismos do ADNmt do grupo étnico-linguístico Kimbundo

Tabela 4.27 - Sequência de ADNmt, haplogrupo estimado e haplotipos do grupo étnico-linguístico Kimbundo (N=10).

População	Haplogrupo estimado	N	Haplótipos
AKI 1	L0a1b	1	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16311C 16320T 93G 95C 185A 189G 236C 247A 263G 309.1C 315.1C
AKI 2*	L2a2	1	16169T 16189C 16193.1C 16223T 16229C 16278T 16294T 16311C 16389A 16519C 73G 152C 182T 263G 315.1C
AKI 3	L1c1	1	16093C 16129A 16187T 16189C 16223T 16263C 16278T 16293G 16294T 16311C 16360T 16399G 16519C 73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A
AKI 4	L3f1b	1	16129A 16209C 16223T 16292T 16311C 16519C 73G 189G 200G 263G 272G 309.1C 315.1C
AKI 5*	L2a1	1	16189C 16193.1C 16223T 16278T 16294T 16309G 16390A 16519C 73G 146C 152C 195C 263G 315.1C
AKI 6*	L3e2b	1	16126C 16172C 16182C 16183C 16189C 16193.1C 16223T 16320T 16519C 73G 150T 195C 263G 315.1C
AKI 7	L1c3b	1	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A
AKI 8	L2a1	1	16086C 16155G 16223T 16278T 16294T 16309G 16390A 73G 143A 146C 152C 195C 198T 263G 315.1C
AKI 9	L0a1b	1	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G 16311C 16320T 93G 95C 185A 189G 236C 247A 263G 309.1C 315.1C
AKI 10*	L1b1*	1	16126C 16189C 16193.1C 16223T 16264T 16270T 16278T 16311C 16519C 73G 151T 152C 182T 185T 195C 247A 263G 315.1C 357G

* - Heteroplasmia de comprimento na zona da poly (C) da região hipervariável I
** - Heteroplasmia de comprimento na zona da poly (C) da região hipervariável II.

4.3.4 – Polimorfismos do ADNmt do grupo étnico-linguístico Ovimbundo

Tabela 4.28 - Sequência de ADNmt, haplogrupo estimado e haplótipos do grupo étnico-linguístico Ovimbundo (N=10).

População	Haplogrupo estimado	N	Haplótipos
AOV 1	L1b1	1	16126C 16166G 16187T 16189C 16193T 16223T 16264T 16270T 16278T 16293G 16311C 16519C 73G 152C 182T 185T 189G 195C 247A 263G 315.1C 357G
AOV 2	L3e1	1	16223T 16327T 73G 150T 189G 200G 263G 309.1C 315.1C
AOV 3	L1c3b	1	16129A 16163G 16187T 16189C 16209C 16223T 16278T 16293G 16294T 16311C 16360T 16519C 73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A
AOV 4	L2b1	1	16114A 16129A 16213A 16223T 16278T 16355T 16362C 16390A 73G 150T 152C 182T 195C 198T 204C 263G 309.1C 315.1C 418T
AOV 5	L0a1b	1	16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16293G 16311C 16320T 93G 95C 185A 189G 236C 247A 263G 309.1C 315.1C
AOV 6	L3b	1	16093C 16124C 16192T 16223T 16278T 16362C 16519C 73G 146C 263G 309.1C 315.1C
AOV 7	L0a2	2	16148T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T 16519C 64T 93G 152C 189G 204C 207A 236C 247A 263G 315.1C
AOV 8	L2	1	16223T 16264T 16278T 16390A 73G 93G 146C 150T 152C 182T 195C 198T 263G 315.1C
AOV 9	L3e2a	1	16223T 16320T 16519C 73G 150T 195C 198T 263G 315.1C

* - Heteroplasmia de comprimento na zona da poly (C) da região hipervariável I
** - Heteroplasmia de comprimento na zona da poly (C) da região hipervariável II.

4.3.5 – Análise comparativa dos polimorfismos de ADNmt entre os três grupos étnico-linguísticos.

Os parâmetros de diversidade de ADNmt relativo aos três grupos étnico-linguístico estão apresentados na Tabela 4.29.

Tabela 4.29 – Parâmetros de diversidade de ADNmt para os três grupos étnico-linguísticos

Populações	Diversidade de sequência	Diversidade nucleotídica
Bakongo	1.0000 ± 0.0447	0.0157 ± 0.0086
Kimbundo	1.0000 ± 0.0447	0.0168 ± 0.0092
Ovimbundo	1.0000 ± 0.0447	0.0171 ± 0.0094

Os resultados relativos a matriz de distância e respectivos valores de P dos três grupos étnico-linguísticos estão representados na Tabela 4.29.

Tabela 4.30 – Matriz de distância e respectivos valores de P para os três principais grupos de Angola.

FST / FST P	Bakongo	Kimbundo	Ovimbundo
Bakongo	*	0.5080 ± -0.0053	0.3686 ± -0.0048
Kimbundo	-0.0070	*	0.7495 ± -0.0043
Ovimbundo	0.0071	-0.0356	*

Na Fig. 4.10 está representado a árvore filogenética do estudo do ADNmt relativo aos três grupos étnico-linguísticos.

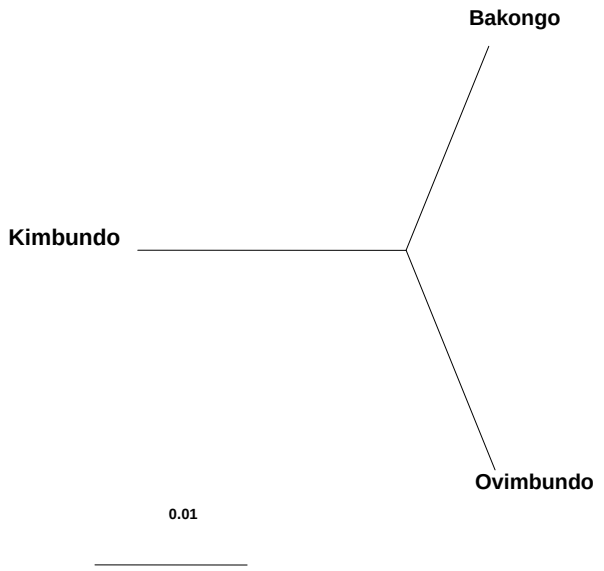


Figura 4.11 - Árvore filogenética do ADNmt nos três principais grupos étnico-linguísticos de Angola.

5. Discussão

5.1 – Microssatélites autossómicos

A população de Angola é constituída por dezenas de grupo étnico-linguísticos com características e comportamentos socioculturais que os distinguem dos demais. Na sua maioria, os povos que habitam Angola são descendentes dos Bantu que migraram da região dos grandes Lagos. A maior parte da população angolana está concentrada junto da faixa litoral. Para além da língua portuguesa, o Ovimbundo é a língua mais falada entre as populações.

Os diversos povos angolanos apresentam uma grande diversidade étnico-linguística, o que tem grande interesse para o estudo genético populacional. Neste trabalho, foram estudados marcadores polimórficos autossómicos, do cromossoma Y e mitocôndriais, nos diversos grupos étnico-linguísticos de Angola.

Para se efectuar a avaliação estatística de STRs autossómicos com aplicação médico-legal, é preciso verificar a existência de equilíbrio de Hardy-Weinberg. Os *loci* do Identifiler® da Promega aplicados neste estudo apresentam independência alélica entre os cromossomas. Nestas circunstâncias, torna-se possível o uso das frequências alélicas obtidas no cálculo de investigação biológica de parentesco e criminalística biológica. Para a análise informativa *a priori* de conjuntos de marcadores para a aplicação forense é importante que seja determinado os parâmetros de eficácia *a priori.*

A aplicação do teste exacto de Guo e Thompson, nos 15 *loci* do Identifiler na população angolana apresentou valores de P superiores a 0.05 (nível de significância escolhido) na maioria dos *loci*, à excepção dos *loci* CSF1P0, D13S317 e D2S1338. Após a correcção de Bonferroni não apresentaram diferenças estatisticamente significativas, o que permitiu concluir que a população angolana se encontra em equilíbrio de Hardy-Weinberg (Tabela.4.1).

Na população angolana, o valor de heterozigosidade variou de 0.6563 (D13S317) a 0.8854 (D21S11) e o valor da heterozigosidade média foi de 0.7882, superior a 0.7, valor proposto por Gill *et al.* (1996) na escolha de *loci* para aplicação forense.

O poder de discriminação nos *loci* analisados apresentou um valor próximo ou superior a 0.8 (Tabela 4.1). Este é o valor mínimo proposto por Urquhart *et al*. (1996) na escolha de *loci* para aplicação forense. O poder de discriminação acumulado dos 15 *loci* foi de 0.999999999968716.

A probabilidade de exclusão *a priori* apresentou um valor médio de 0.6017 e o poder de exclusão acumulado apresentou um valor de 0.999999028. Este valor é

relativamente superior do mínimo proposto (99.9%) pelo grupo Luso-Espanhol de Hemogenética Forense (GEP, 1992)

Tabela 5.1 – Parâmetros estatísticos de eficácia *a priori* e número de alelos por *locus*, para a população angolana.

Loci	**H**	**PD**	**PEx**	**Número de alelos**
D8S1779	0.7604	0.7515	0.5414	9
D21S11	0.8854	0.8593	0.7235	23
D7S820	0.7625	0.7822	0.5757	9
CSF1PO	0.8083	0.7949	0.6008	9
D3S1358	0.7396	0.7397	0.5022	9
HUMTH01	0.6792	0.6922	0.4450	6
D13S317	0.6563	0.6742	0.4234	8
D16S539	0.7729	0.7628	0.5514	10
D2S1338	0.8813	0.8916	0.7810	17
D19S433	0.8500	0.8414	0.6952	15
HUMVWA	0.8167	0.8128	0.6315	11
TPOX	0.7646	0.7755	0.5644	9
D18S51	0.8542	0.8595	0.7205	21
D5S818	0.7375	0.7468	0.5235	8
FIBRA/FGA	0.8542	0.8759	0.7468	26

A análise dos resultados dos parâmetros de eficácia *a priori* dos 15 *loci*, observou-se que os *loci* D21S11, D2S1338, D18S51 e FGA (Tabela 5.1) apresentaram valores mais elevados e maior número de alelos, confirmando assim uma maior capacidade informativa destes *loci* em relação aos restantes. Os valores mais baixos foram observados nos *loci* HUMTH01 e D13S317 para os parâmetros analisados

Na análise da matriz de distância (Tabela 4.2) com base nas frequências constatou-se que a população de Angola está mais próxima das populações Bantu do que de outras populações não Bantu. Entre as populações africanas a Guiné Equatorial é a mais próxima da população de Angola apesar de não terem fronteiras geográficas em comum. Este facto pode ser explicado com base em processos migratórios Bantu (Fig. 2.3), visto que a população angolana é descendente dos Bantu que vieram da região dos grandes Lagos.

Na Tabela 4.3, comparou-se a estrutura genética da população de Angola e de outras populações, em termos de frequência alélicas. As diferenças significativas quanto à distância genética são com Portugal (Lopes *et al*., 2009), Argentina (Bozzo *et al*., 2007) e o México (Hernández *et al*., 2005), para todos os 15 *loci* analisados; com a Espanha (Camacho, *et al*., 2007) e a Venezuela (Bernal *et al*., 2006) em 14 *loci*; com a Somália (Tilmar *et al*., 2009) em 13 *loci*; com a Bahía-Brasil, (Alves, *et al*., 2004) em 12 *loci*; com o Uganda (Gomes, *et al*., 2009) em 9 *loci*; com Moçambique (Alves, *et al*., 2004) em 4 *loci*; com os afro-americanos (Butler *et al*., 2003) em 3 *loci*; com a Guiné Equatorial (Alves, *et al*., 2005) em 1 *locus* apenas; com a Namíbia (Muro, *et al*., 2008), não são demonstráveis diferenças estatísticas em nenhum dos *loci* estudados.

5.1.2 – Equilíbrio de Hardy-Weinberg nos grupos étnico-linguísticos estudados

A árvore filogenética das 13 populações comprova os resultados obtidos nas comparações genético-populacionais com a matriz de distâncias génicas e os valores de teste exacto de P (Tabela 4.2 e 4.3). Podemos considerar três grupos distintos: um composto por países europeus e da América latina, outro extremo da árvore composto pelo grupo de países africanos e por um grupo intermédio composto pela, Somália, Uganda e afro-americanos (Fig. 4.1).

A aplicação do teste exacto de Guo e Thompson no grupo Bakongo no *locus* D13S317 (P = 0.0326) e TPOX (P = 0.0169) na Tabela 4.4, no grupo Kwanhama no *locus* HUMTH01 (P=0.0462), TPOX (P = 0.0421) e FGA (P = 0.0268) na Tabela 4.6, no grupo Lunda-Tchokwe nos *locus* CSF1PO (P = 0.0251) na Tabela 4.7, no grupo Nganguela no *locus* D2S1338 (P=0.0319) na Tabela 4.8, no grupo Ovimbundo no *locus* D19S433 (P = 0.0397) na Tabela 4.10 mostraram valores de P inferiores a 0.05 (nível de significância escolhido). Após a correcção de Bonferroni, nos *loci* com valores de P inferiores a 0.05, verificou-se que os grupos étnico-linguísticos se encontram em equilíbrio de Hardy-Weinberg. Assim sendo, as frequências alélicas obtidas podem ser utilizadas em estudo genético-populacionais e forenses.

5.1.3 - Parâmetros de eficácia *a priori*

Os valores de heterozigosidade observada média foram superiores a 0.7 para a maioria dos sistemas estudados, à excepção dos *loci* HUMTH01 e D13S317 (Tabela 5.2) que apresentaram valores relativamente próximos, valor que deve estabelecer um dos requisitos fundamental na escolha de *loci* para aplicação em casos forenses (Gill *et al* 1996; Urquhart *et al*., 1996). Contudo os *loci* que apresentaram valores mais elevados mostraram maior poder de discriminação e poder de exclusão *a priori*.

Tabela 5.2 – Heterozigosidade observada nos STRs autossómicos nos grupos étnico-linguísticos

População	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
Bakongo	0.7647	0.9118	0.7843	0.8431	0.6862	0.6470	0.7549	0.7843	0.8529	0.8921	0.8529	0.7745	0.9019	0.7353	0.8333
Kimbundo	0.7700	0.9000	0.7600	0.7900	0.7700	0.6700	0.6700	0.7900	0.8700	0.8900	0.8300	0.6700	0.8600	0.7500	0.8400
Kwanhama	0.7000	0.8600	0.7600	0.8400	0.8000	0.7000	0.6400	0.8600	0.9200	0.7400	0.7800	0.7800	0.8200	0.7200	0.8200
Lunda-Tchokwe	0.7800	0.8200	0.6600	0.8400	0.7600	0.6800	0.7400	0.7000	0.8400	0.9200	0.8200	0.6400	0.8800	0.7000	0.8800
Nganguela	0.6800	0.8600	0.7400	0.7400	0.7200	0.7200	0.5600	0.8200	0.8800	0.8400	0.8200	0.7600	0.7600	0.8600	0.9000
Nhaneca-Humbe	0.8800	0.8800	0.7200	0.8400	0.7200	0.7200	0.6000	0.7600	0.8800	0.9200	0.7600	0.9200	0.8400	0.8000	0.8800
Ovimbundo	0.7745	0.9020	0.8137	0.7843	0.7353	0.6863	0.5686	0.7157	0.9216	0.7745	0.7941	0.7941	0.8627	0.6765	0.8725
Média	0.7642	0.8763	0.7483	0.8111	0.7416	0.6890	0.6476	0.7757	0.8806	0.8538	0.8081	0.7627	0.8464	0.7488	0.8608

O poder de discriminação dos *loci* D21S11, D2S1338, D19S433, D18S51 e FGA transpôs os requisitos de 0.8 propostos por Urquhart *et al.* (1996), (Tabela 5.3). Os restantes *loci* apresentaram valores próximos ou inferiores ao proposto por Urquhart *et al.* (1996), o que não constitui obstáculo para a sua aplicação em Genética Forense, uma vez que eles são analisados em conjunto. O poder de discriminação acumulado dos 15 *loci* foi superior a 0.9999 em todos os grupos étnico-linguísticos estudados.

Tabela 5.3 – Poder de discriminação dos STRs autossómicos nos grupos étnico-linguísticos

População	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
Bakongo	0.7408	0.8627	0.7972	0.7888	0.7432	0.6628	0.6997	0.7762	0.8900	0.8474	0.8197	0.7749	0.8492	0.7452	0.8818
Kimbundo	0.7402	0.8451	0.7705	0.7770	0.7358	0.7264	0.6834	0.7538	0.8887	0.8480	0.8096	0.7903	0.8717	0.7480	0.8721
Kwanhama	0.7180	0.8520	0.7302	0.7904	0.7320	0.6430	0.6732	0.7734	0.8888	0.8122	0.7986	0.7446	0.8474	0.7278	0.8756
Lunda-Tchokwe	0.7548	0.8386	0.7518	0.7840	0.7216	0.7010	0.6664	0.7876	0.8868	0.8532	0.8128	0.7424	0.8590	0.6886	0.8280
Nganguela	0.7374	0.8364	0.7724	0.7914	0.7162	0.6876	0.6244	0.7090	0.8866	0.8248	0.8044	0.7744	0.8418	0.7534	0.8630
Nhaneca-Humbe	0.7640	0.8400	0.7576	0.7768	0.6920	0.6688	0.6304	0.7128	0.8856	0.8568	0.8328	0.7616	0.8696	0.7104	0.8696
Ovimbundo	0.7661	0.8748	0.7885	0.8017	0.7428	0.6917	0.6639	0.7615	0.8819	0.8107	0.7999	0.7636	0.8490	0.7635	0.8740
Média	0.7467	0.8499	0.7669	0.7872	0.7262	0.6830	0.6631	0.7535	0.8869	0.8362	0.8111	0.7645	0.8554	0.7338	0.8663

A probabilidade de exclusão *a priori* apresentou valores médios superiores a 0.7 nos *loci* D21S11, D2S1338, D19S433, D18S51 e FGA (Tabela 5.4) e a maioria dos *loci* apresentou valores médios ligeiramente superiores a 0.5. O poder de exclusão *a priori* acumulado nos 15 *loci* dos grupos étnico-linguísticos foi superior a 0.999999, o que confirma a importância da utilização conjunta dos 15 loci, em estudos de genética forense.

Tabela 5.4 – Poder de exclusão *a priori* dos STRs autossómicos nos grupos étnico-linguísticos

População	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
Bakongo	0.5169	0.7279	0.6050	0.5922	0.5092	0.4080	0.4469	0.5680	0.7772	0.7060	0.6420	0.5637	0.7023	0.5232	0.7622
Kimbundo	0.5326	0.6972	0.5560	0.5818	0.5029	0.4992	0.4385	0.5423	0.7762	0.7035	0.6275	0.5864	0.7444	0.5286	0.7230
Kwanhama	0.4925	0.7112	0.4922	0.5911	0.4873	0.3765	0.4313	0.5608	0.7754	0.6414	0.6070	0.5232	0.6981	0.4920	0.7505
Lunda-Tchokwe	0.5545	0.6837	0.5344	0.5786	0.4801	0.4459	0.3985	0.5863	0.7716	0.7132	0.6300	0.5167	0.7174	0.4327	0.6833
Nganguela	0.5151	0.6877	0.5558	0.5952	0.4653	0.4384	0.3736	0.4745	0.7735	0.6698	0.6181	0.5596	0.6874	0.5299	0.7280
Nhaneca-Humbe	0.5477	0.6885	0.5416	0.5685	0.4418	0.4028	0.3689	0.4841	0.7702	0.7139	0.6656	0.5322	0.7383	0.4687	0.7361
Ovimbundo	0.5595	0.7505	0.5851	0.6114	0.5108	0.4437	0.4149	0.5491	0.7634	0.6484	0.6129	0.5492	0.7009	0.5536	0.7396
Média	0.5265	0.7067	0.5475	0.5884	0.4853	0.4306	0.4104	0.5379	0.7725	0.6852	0.6290	0.5473	0.7127	0.5041	0.7318

5.1.4 - Comparações genético-populacionais entre os grupos étnico-linguísticos

Na análise comparativa entre os diversos grupos étnico-linguísticos estudados com base na matriz de distância observaram-se diferenças estatisticamente significativas entre o grupo Bakongo, da região mais ao norte, e o grupo Kwanhama (FST=0.0029, P=0.0268) da região mais ao sul de Angola (Tabela 4.11). Esta diferença com base nas frequências alélicas pode ser devido ao reduzido número da amostra ou devido ao facto do grupo Kwanhama serem descendentes dos Bantu meridionais que terão entrado em Angola pela fronteira sul. Os hábitos e costumes do povo Kwanhama são similares aos do povo Ovambo localizado no Norte da república da Namíbia.

As comparações genético-populacionais entre os grupos étnico-linguísticos demonstram que o grupo Bakongo, quando comparado com outros grupos, está mais próximo dos seus vizinhos Lunda-Tchokwe, seguido pelos Ovimbundo e Kimbundo (Tabela 4.11).

A árvore filogenética (Fig. 4.2) representa os resultados da matriz de distâncias entre os 7 grupos étnico-linguísticos e por outro lado, reflecte a posição geográfica que cada grupo étnico ocupa no contexto geográfico angolano (Fig. 2.5).

5.2. – Microssatélites do cromossoma Y na população angolana e análise comparativa com outras populações africanas

Na população angolana verificou-se que os *loci* de STRs do cromossoma Y com maior número de alelos foram DYS389 II e DYS390 com 7 alelos, DYS393 com 6 alelos, DYS19, DYS389 I, DYS392 e DYS437 com 5 alelos cada.

Os alelos mais frequentes por *locus* DYS19, DYS389 I, DYS389 II DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 e DYS439 (Tabela 4.12) são similares aos da população da Guiné Bissau (Rosa *et al*., 2006).

O *locus* DYS385 apresentou 29 haplótipos diferentes, prevendo-se uma maior capacidade informativa quando comparado com os demais (Fig. 4.3).

Nos *loci* DYS19, DYS389 II, DYS393 e DYS439, os alelos mais frequentes apresentaram frequências superiores a 0.5 e mostraram também maior diversidade génica em relação aos restantes. O *locus* DYS389 II apresentou elevado valor de diversidade génica. O *locus* DYS390 apesar de apresentar maior número de alelos apresentou diversidade génica inferior a 0.5 (Tabela 4.12). Alguns *loci* que apresentaram diversidades génicas mais baixas, inferiores a 0.5, devido às elevadas frequências dos alelos mais comum, o que se traduz numa variabilidade muito baixa destes marcadores (Tabela 4.13).

A frequência do haplótipo mais frequente no *locus* DYS385 foi de 0.193 (Tabela 4.13). A diversidade haplotípica nos 11 STRs do cromossoma Y do sistema PowerPlex, na população angolana foi de 0.996933187295 e a capacidade de discriminação foi de 0.8313, o que traduz o seu elevado polimorfismo.

Na população angolana foram detectados 120 (72.3%) haplótipos únicos, sendo os mais frequentes H007, H050, H053 e H103 e apresentaram uma frequência de 0.024 (Tabela 4.14). Realizou-se uma pesquisa destes haplótipos na base de dados mundial do cromossoma Y (YHRD) e constatou-se o seguinte:

O haplótipo H007 da população de Angola foi encontrado nas seguintes populações: em 3 de 1018 Afro-americanos, em 3 de 53 Angola (Ovimbundo), em 1 de 1100 Norte americanos, em 1 de 57 Angola (Bakongo), em 1 de 39 Benin (Fon), em 1 de 57 Angola (Cabinda).

O haplótipo H050 da população de Angola foi encontrado nas seguintes populações: em 2 de 57 Angola (Bakongo), em 2 de 75 Angola (Cabinda), em 1 de 53 Angola (Ovimbundo), em 1 de 53 Angola (Kimbundo).

O haplótipo H053 da população de Angola foi encontrado nas seguintes populações: em 3 de 54 Namíbia (Ovamboland), em 2 de 327 Argentina (Córdoba), em 2 de 1018 Afro-americanos, em 2 de 57 Angola (Bakongo), em 2 de 75 Angola (Cabinda),

em 1 de 53 Angola (Ovimbundo), em 1 de 53 Angola (Kimbundo), em 1 de 224 Argentina (Mendonza), em 1 de 205 Brasil (Bahia), em 1 de 39 Benin (Fon), em 1 de 139 Brasil, (Porto Velho), em 1 de 173 Venezuela (Maracay), em 1 de 50 Brasil (Mato Grosso do Sul.)

O haplótipo H103 da população de Angola foi encontrado nas seguintes populações: em 3 de 1018 afro-americano, em 2 de 53 Angola (Ovimbundo), em 1 de 255 Brasil (Rio Grande do Sul), em 1 de 53 Angola (Kimbundo), em 1 de 205 Brasil (Bahia), em 1 de 54 Namíbia (Ovambo).

Contudo, o haplótipo H053 é o mais frequente fora das fronteiras angolanas e o haplotipo H103 o menos frequente. Este resultado pode ser explicado pela escassez de dados referentes a outras populações na base de dados da YHRD.

As comparações genético-populacionais realizadas entre a população angolana masculina e cada uma das 6 populações, localizadas em distintas áreas geográficas de África Bantu publicadas referentes aos 11 STRs do cromossoma Y do PowerPlex®, revelaram diferenças estatisticamente significativas com a maioria das populações: valor de $P<0.05$, com a Guiné-Bissau (Rosa *et al*., 2006); com Uganda (Gomes *et al*., 2009); com Moçambique (Alves *et al*., 2003); com a Somália (Hallenberg *et al*., 2005); com a Guine Equatorial (Arroyo-Pardo *et al*., 2005), à excepção da população da Namíbia (Fujihara *et al*., 2009) que se encontra mais próxima ($P = 0.8291$, $FST =-0.0064$) (Tabela 4.15).

A análise da matriz de distância, neste estudo, revela que a população masculina angolana quando comparada com outras populações está mais próxima dos seus vizinhos da África austral (Namíbia $FST = -0.0064$ e Moçambique $FST = 0.0231$) e a mais distante da Somália ($FST = 0.63459$) localizada no leste de África (Tabela 4.15). A aproximação génica entre povos da África austral pode dever-se ao facto de pertencerem ao grupo Bantu que se deslocou das regiões dos grandes lagos e se fixou no Sul de África.

A análise da variância molecular entre as populações demonstrou que a variação dentro de cada grupo é maior (61.67%) que a variação entre os grupos (Tabela 4.16).

Foi constituído um banco de dados de frequências alélicas e haplotípicas do STRs do cromossoma Y do PowerPlex®, com base na tipagem dos 166 indivíduos da população angolana escolhidos ao acaso na população.

5.3 – Análise comparativa dos polimorfismos do cromossoma Y entre os três principais grupos étnico-linguísticos

A caracterização genética populacional étnico-linguístico por *locus* do cromossoma Y através do sistema Y-STR PowerPlex®, encontra-se na Tabela 5.5. Os três grupos analisados apresentam polimorfismo para todos os marcadores analisados.

Tabela 5.5 - Caracterização genética populacional por *locus* entre os três grupos étnico-linguísticos.

Loci	**Bakongo**	**Kimbundo**	**Ovimbundo**
DYS19	14-15-16-17	14-15-16-17	****13***-14-15-16-17
DYS389 I	12-13-14	12-13-14-***15****	****10***-12-13-14
DYS389 II	28-29-30-31-32-***33****	28-29-30-31-32-***33***	****26***-28-29-30-31-32
DYS390	20-21-22-23-24-25-***26****	20-21-22-23-24	20-21-22-23-24-*25*
DYS391	10-11-***12****	9-10-11	9-10-11
DYS392	11-12-13-***15****	****10***-11-12-13	11-13
DYS393	13-14-15-16	****11***-13-14-15-16	****12***-13-14-15
DYS437	****12***-13-14-15	13-14-15-***16****	13-14-15-***16****
DYS438	10-11-12	****9***-10-11-12	****9***-10-11-12
DYS439	****10***-11-12-13	****10***-11-12-13	*11*-12-13

***** Alelos específicos de cada grupo

O sistema estendido do PowerPlex® do cromossoma Y estudado nos grupos étnico-linguísticos apresenta variabilidades em diferentes populações, quanto ao número de alelos por *locus.*

No sistema DYS19 os grupos Bakongo e Kimbundo apresentam quatro alelos (DYS19*14, *15**,** *16 e *17), enquanto o grupo Ovimbundo apresenta um alelo adicional (DYS19 *13), sendo o alelo DYS19*15 o mais representativo nos três grupos. O alelo (DYS19 *13) surge apenas no grupo Ovimbundo.

No sistema DYS389 I o grupo Bakongo apresenta 3 alelos (DYS389 I *12, *13 e*14), enquanto o grupo Kimbundo apresenta um adicional (DYS389 I *15) e o Ovimbundo apresenta um outro adicional (DYS389 I *10), sendo o alelo DYS389 I *13 o mais representativo nos três grupos.

No sistema DYS389 II, os grupos Bakongo e Kimbundo apresentam 6 alelos similares (DYS389 II *28, *29, *30, *31 *32 e *33) e o grupo Ovimbundo também 6 alelos (DYS389 II *26, *28, *29, *30, *31 e*32), sendo o alelo DYS389 II *30 o mais

representativo nos grupos Bakongo e Kimbundo e o alelo DYS389 II *31 nos grupos Bakongo e Ovimbundo respectivamente.

No sistema DYS390, o grupo Kimbundo apresenta 5 alelos (DYS390 *20, *21, *22, *23 e*24), o grupo Ovimbundo apresenta um adicional (DYS390 *25) e o grupo Kimbundo 2 adicionais (DYS390 *25 e *26), sendo o alelo DYS390 *21 o mais representativo nos três grupos.

No sistema DYS391, os grupos Kimbundo e Ovimbundo apresentaram 3 alelos cada (DYS391 *9, *10 e*11) e o grupo Bakongo apresentou 3 alelos (DYS391 *10, *11 e *12), sendo o alelo DYS391 *10 o mais representativo no grupo Bakongo e Kimbundo ao passo que no grupo Ovimbundo foi o alelo 11 (Tabela 5.5).

No sistema DYS392 o grupo Ovimbundo apresentou 2 alelos (DYS392*11 e*13), o grupo Kimbundo apresentou 2 alelos adicionais (DYS392 *10 e *12) e o grupo Bakongo apresentou também 2 alelos adicionais (DYS392 *12 e *15), sendo o alelo DYS392 *11 o mais representativo nos três grupos.

No sistema DYS393, o grupo Bakongo apresenta 4 alelos (DYS393 *13, *14, *15 e *16), o Kimbundo apresenta um adicional (DYS393 *11**)** e o grupo Ovimbundo apresenta 4 alelos com a seguinte sequência (DYS393 *12 *13 *14 e *15), sendo o alelo DYS393 *13 o mais representativo nos três grupos.

No sistema DYS437, os grupos Kimbundo e Ovimbundo apresentam 4 alelos similares (DYS437 *13 *14 *15 *16) e o grupo Bakongo apresenta 4 alelos distintos (DYS437 *12 *13 *14 *15), sendo o alelo DYS437 *14 o mais representativo nos três grupos.

No sistema DYS438 o grupo Bakongo apresentou 3 alelos (DYS438 *10, *11 e *12) e os grupos Kimbundo e Ovimbundo apresentaram um alelo adicional cada (DYS438 *9), sendo o alelo DYS438 *11 o mais representativo nos três grupos.

No sistema DYS439, o grupo Ovimbundo apresenta três alelos (DYS439 *11, *12 e *13) e o grupo Kimbundo e Bakongo apresentam um alelo adicional cada (DYS439 *10), sendo o alelo DYS439 *11 o mais representativo no grupo Bakongo e o alelo DYS439 *12 o mais representativo no grupo Kimbundo e Ovimbundo, respectivamente.

No sistema DYS437, os alelos DYS437 *12 e DYS390 *26 são característicos apenas dos Bakongos, na região Norte. Os alelos DYS389 I *15 e DYS393 *11 são característicos dos Kimbundos, na região Centro Norte de Angola. Os alelos DYS19 *13 DYS389I *10, DYS389II *26 e o DYS393 *12 estão presentes apenas entre os Ovimbundos da região Centro-Sul de Angola.

Os *loci* DYS389 II e DYS390 foram os que maior variação apresentaram quanto ao número de alelos, nos três grupos étnico-linguísticos, apresentando maior diversidade alélica.

Constatou-se que alguns grupos étnico-linguísticos apresentaram alelos não verificados noutros. No entanto, a falta de dados adicionais e de outros estudos com diferentes marcadores nos três grupos étnicos linguísticos não permite afirmar que os mesmos alelos sejam caracterizados como marcadores populacionais destes grupos, visto que os mesmos foram observados em outras populações.

O *locus* DYS385 nos três grupos étnico-linguísticos apresentou maior número de alelos. No grupo Bakongo o haplótipo mais frequente foi o 17-17 (12), no Kimbundo foi o 16-17 (12) e no grupo Ovimbundo 16-17 (11), sendo o *locus* DYS385 mais discriminativo, nos três grupos étnico-linguísticos.

Tabela 5.6 – Valores da diversidade génica dos STRs do cromossoma Y nos três grupos

População	**DYS19**	**DYS389 I**	**DYS389 II**	**DYS390**	**DYS391**	**DYS392**	**DYS393**	**DYS437**	**DYS438**	**DYS439**
Bakongo	0.6557	0.4100	0.7344	0.3675	0.2708	0.1336	0.5934	0.2241	0.3706	0.6661
Kimbundo	0.5191	0.3106	0.6824	0.2296	0.4012	0.1983	0.5593	0.2557	0.3361	0.6295
Ovimbundo	0.6059	0.4016	0.6588	0.3895	0.5005	0.1068	0.4930	0.2072	0.3005	0.5511

Os elevados valores da diversidade génica nos sistemas DYS19, DYS389 II, DYS393 e DYS439 (Tabela 5.6) devem-se a uma distribuição mais uniforme dos seus alelos. Os baixos valores da diversidade génica nos restantes sistemas podem ser justificados com a elevada frequência de um alelo nesses *loci*.

Os três grupos étnico-linguísticos são similares quanto ao alelos mais frequentes na maioria dos *loci* do PowerPlex do cromossoma Y analisados à excepção dos *loci* DYS389 II e DYS391.

A diversidade haplotípica nos três grupos varia de 0.996371552975 a 0.9980519481, o que traduz elevado polimorfismo no haplótipo identificado com este grupo de marcadores do cromossoma Y.

A análise da matriz de distâncias génicas dos três grupos étnico-linguísticos revelou que o grupo étnico-linguístico Kimbundo ocupa uma posição intermediária entre outros grupos (Fig. 5.2). Estes resultados corroboram a posição dos grupos no contexto geográfico e divisão político-administrativa de Angola. Não foram encontradas diferenças estatisticamente significativas entre as populações, revelando a sua grande homogeneidade.

Os três grupos mostraram uma grande variabilidade haplotípica, sendo que maior frequência de haplótipos distintos e diversidade haplotípica foi verificada no grupo Kimbundo (Tabela 5.7). A alta capacidade de discriminação do conjunto dos 11 *loci* do sistema PowerPlex® do cromossoma Y mostra valores relevantes que permitem a sua utilização em Genética Forense, nomeadamente na identificação do sexo masculino, assim como a sua aplicação em genética populacional.

Tabela 5.7 – Diversidade de parâmetros dos 11 STRs nos três grupos étnico-linguísticos de Angola.

População	**N**	**Ht Dif**	**D**	**h**
Bakongo	57	53	0.9298	0.9975
Kimbundo	56	53	0.9464	0.9981
Ovimbundo	53	49	0.9245	0.9964
Total	166	138	0.8313	0.9969

Ht. Dif. Haplótipos diferentes; D: capacidade de discriminação; h: diversidade haplotipica

O poder de discriminação nos três grupos étnico-linguísticos apresentou valores elevados (acima de 0.9245). Entretanto, considerando os três grupos em conjunto, o poder de discriminação na população angolana apresentou o valor de 0.8313, devido ao reduzido número de haplótipos diferentes observados (Tabela 5.7).

A diversidade haplotípica na população angolana e nos três grupos étnico-linguísticos analisados separadamente foram elevados, apresentando valores acima de 0.99. Este parâmetro indica a probabilidade de dois indivíduos escolhidos ao acaso apresentarem haplótipos diferentes.

A alta diversidade haplotípica nos grupos analisados reflecte a alta variabilidade intra-populacional e permite melhor diferenciação dos grupos étnico-linguísticos. A análise da variância molecular, AMOVA, utilizando os *loci* de STR do cromossoma Y não revelou diferenças estatisticamente significativas entre os diferentes grupos. Os resultados mostraram uma alta variabilidade dentro das populações (98.99%) e baixa variabilidade entre as populações (Tabela 4.16).

Os resultados referentes aos três grupos étnico-linguísticos foram adicionados a base de dados mundial do cromossoma Y (Y-STR Haplotype Refence Database), com o seguinte número de acesso: Nº YA003640 AKI- Kimbundo, YA003641 ABK- Bakongo e YA003642 AOV - Ovimbundo. Esta base de dados constitui uma importante ferramenta com aplicação em estudos populacionais e casos forenses.

5.4 – Análise comparativa dos polimorfismos do ADNmt entre os três principais grupos étnico-linguísticos

Considerando o conjunto da população angolana, a análise do ADNmt da região controlo, o tipo de polimorfismo de sequência predominante foi a substituição por transição pirimidínica, localizados na zona hipervariável I. No entanto foram identificadas 27 (90%) haplótipos diferentes dos quais 24 (80%) eram únicos tendo em consideração a região controlo total, na população angolana. A taxa de heteroplasmia foi de 7 (23.3%) na população e foram localizados na zona poli C da região hipervariável I.

Na população de Angola, o haplogrupo representativo foi o L1c3b (13.3%), seguido pelo haplogrupo L0a1b (10%) e os haplogrupos L0a2, L1b1, L1c1, L3e1, L3e2b (6.7%) (Tabela 4.25). As populações com elevadas frequências de haplótipos únicos são consistentes com um estabelecimento antigo no local.

Na análise comparativa entre os grupos étnico-linguísticos observou-se no grupo Bakongo, o haplogrupo L1c3b como o mais frequente. O haplogrupo L1c é representativo na África central, oeste e sudeste africano; no grupo Kimbundo o haplogrupo L2a1 foi o mais representativo e apresentou uma frequência de 0.2. O haplogrupo L2 está localizado no oeste e centro oeste do continente africano; no grupo Ovimbundo o haplogrupo predominante foi L0a2 com uma frequência de 0.2. O macrohaplogrupo L0 teve a sua origem no leste de África.

A diversidade de sequência ou haplotípica para os três grupos étnico-linguísticos foi de 1.0000 ± 0.0447. A região controlo apresenta uma diversidade elevada que permite um maior poder de distinção entre os haplótipos. A alta diversidade haplotípica na população angolana evidencia a variabilidade populacional dos haplótipos do ADNmt.

Tabela 5.8 – Número e tipo de alterações polimórficas da região controlo total do ADNmt para os três grupos étnico-linguísticos de Angola.

Populações	Transições	Transversões	Substituições	Nº de posições polimórficas
Bakongo	45	7	52	51
Kimbundo	49	7	56	54
Ovimbundo	59	6	55	53

Entre os 3 grupos étnico-linguísticos foram identificados 29 haplótipos, dos quais 28 eram únicos, expressos por 51 posições polimórficas no grupo Bakongo, 54 posições polimórficas no Kimbundo e 53 posições polimórficas no grupo Ovimbundo (Tabela 5.8). O grupo Kimbundo apresentou maior número de posições polimórficas e consequentemente maior número de substituições. O grupo Ovimbundo apresentou maior número de substituições por transições que os restantes grupos.

As inserções foram detectadas na zona poli-C, situadas na posição 309 da região hipervariável II no ABK 2, AKI 1, AKI 4, AKI 9, AOV 2, AOV 4, AOV 5 e AOV 6 (Tabelas 4.26, 4.27 e 4.28) respectivamente. Entretanto, foram excluindo a inserção 315.1C, que é comum, apresentando uma taxa de 100% em todos os grupos. A ocorrência de heteroplasmia na zona de poli-C foi observada na região hipervariável II e apresentou uma taxa 30% no grupo Bakongo e 40% no grupo Kimbundo, podendo estar relacionada com a alta taxa de mutação verificada no ADNmt.

A substituição por transição (T/C e G/A) no grupo Bakongo correspondeu a 88.2%, no grupo Kimbundo a 91% e no Ovimbundo a 111%, sendo as substituições por transversões menos frequentes nos três grupos (Tabela 5.8). O segmento HVI da região controlo apresentou maior variabilidade polimórfica.

A diversidade nucleotídica foi de 0.0157 ± 0.0086 no grupo Bakongo, 0.0168 ± 0.0092 no grupo Kimbundo e de 0.0171 ± 0.0094 no grupo Ovimbundo (Tabela 4.29). Podemos considerar que a diversidade de sequência foi elevada nos três grupos étnico-linguísticos.

Na análise dos haplótipos, há semelhança entre as amostras AKI 1 com AOV 2; entre AKI 9 com AOV 5; e entre AKI 7 com AOV 3, presumindo-se não haver uma

diferenciação por grupo populacional. Verifica-se mais semelhanças na globalidade, entre Bakongo e Ovimbundo do que entre Bakongo e Kimbundo.

A análise da árvore filogenética, baseada na distância molecular das populações (Fig. 4.10) revelou que os três grupos étnico-linguísticos, estão muito próximos uns dos outros, não apresentando diferenças estatisticamente significativas (Tabela 4.30). Pode-se afirmar que a população apresentou alta homogeneidade em relação ao conjunto de marcadores. O estudo dos grupos étnicos mostra uma grande variabilidade genética intra-populacional com altos valores de diversidade nucleotídica e de sequência.

Tendo em consideração que se trata de uma população maioritariamente Bantu localizada na África austral, era de esperar que a população angolana apresentasse haplótipos pertencentes ao macrohaplogrupo L. Os haplogrupos L são específicos das populações Africanas que evidenciam a origem do homem moderno. O haplogrupo L3 deu origem a haplogrupos M, N e R, que englobam todas as variações observadas fora da África, de acordo com a teoria de *out-of-Africa* (Van Oven *et al*., 2008). Os haplogrupos do ADNmt são importantes para compreender a história e o passado demográfico de uma população, uma vez que reflectem a relação filogenética entre as populações.

6. Conclusões e Perspectivas

de

Estudos Futuros

O estudo dos principais grupos étnico-linguísticos de Angola permitiu concluir o seguinte:

- Obtiveram-se frequências alélicas de 15 *loci* microssatélites (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, VWA, TPOX, D18S51, D5S818 e FGA) nos principais grupos étnico-linguísticos e na população global de Angola, as quais se encontravam em equilíbrio de Hardy-Weinberg;

- Os loci D21S11, D2S1338, D18S51 e FGA apresentaram o maior grau de polimorfismo e, consequentemente, uma maior capacidade informativa em todos os grupos étnico-linguísticos;

- Constatou-se a presença de grupos étnico-linguísticos (Bakongo/Kwanhamas e Kimbundo/Kwanhamas) distintos entre si (diferentes frequências génicas, estatisticamente significativas);

- A análise de AMOVA com os marcadores STRs autossómicos mostrou uma grande variabilidade dentro das populações e diferenças relativas estatisticamente significativas;

- A análise das distâncias genéticas permitiu afirmar que a população de Angola, no seu todo, apresenta um carácter homogéneo apesar de algumas pequenas diferenças inter-populacionais;

- Os parâmetros estatísticos de aplicação forense (heterozigosidade média, poder de discriminação a *priori* e poder de exclusão a *priori*) apresentaram valores elevados para os 15 *loci* autossómicos estudados, garantindo assim a sua futura aplicabilidade em estudos de identificação genética individual e de criminalística biológica, bem como a constituição de base de dados para aplicação forense.

- A comparação entre a população angolana e outras populações mundiais analisadas com os marcadores do *Identifiler* mostrou três grupos distintos: um extremo composto por países europeus e da América latina, um grupo intermédio composto pela população da Bahia (Brasil), Somália, Uganda e afro-americanos

dos EUA e um outro extremo composto por outros países africanos, entre os quais Angola;

- O conjunto dos marcadores do *PowerPlex* do cromossoma Y (11 STRs) apresentou valores de diversidade haplotípica e capacidade discriminativa haplotípica elevados nos 3 principais grupos étnico-linguísticos de Angola, incluindo os Kimbundos;

- A maior variabilidade molecular foi observada dentro de cada grupo populacional, muito mais que entre os diferentes grupos;

- O poder de discriminação na população global diminuí quando analisada a partir do somatório dos três grupos étnico-linguísticos;

- No estudo do ADNmt os três maiores grupos étnico-linguísticos mostraram uma grande variabilidade genética intra-populacional, manifestada pela grande frequência de haplótipos únicos (93.3%) na população global;

- A diversidade de sequência do ADNmt foi máxima nos três grupos, e a diversidade nucleotídica foi também elevada, sendo maior no grupo Ovimbundo;

- A população apresenta uma grande variabilidade genética manifestada pela grande frequência de haplótipos únicos;

- Os três maiores grupos étnico-linguísticos mostraram pertencer ao macro - haplogrupo mitocondrial L e seus haplogrupos, consolidando a ideia da sua origem Bantu;

- Na população angolana, o haplogrupo mitocondrial L1c3b foi o mais comum;

As frequências alélicas e haplotípicas podem agora ser utilizadas para estudos populacionais e servir para a construção de uma base de dados forense para a população de referência.

Perspectivas de Estudos Futuros

Todo o estudo científico é contínuo e não se limita apenas ao esclarecimento de algumas questões inerentes ao programa de trabalho definido.

Por outro lado, este trabalho será complementado, num futuro próximo, em Angola, com o estudo de outros polimorfismos STRs do cromossoma X e com SNPs autossómicos, do cromossoma Y e do ADN mitocondrial, que darão informações relevantes sobre haplogrupos e sub-haplogrupos.

O estudo de outros marcadores genéticos permitirá obter informações mais detalhadas acerca da afinidade filogenética e, particularmente, da variabilidade genética dos principais grupos étnico-linguísticos Bantu que constituem a população de Angola.

7. Resumo

Resumo

Diversidade genética nos principais grupos populacionais em Angola – Aplicação Forense

Este trabalho constitui um estudo genético populacional dos principais grupos étnico-linguísticos (Bakongo, Kimbundo, Kwanhama, Lunda-Tchokwe, Nganguela, Nhaneca-Humbe e Ovimbundo) presentes em Angola. Entre estes, os grupos Bakongo, Kimbundo e Ovimbundos representam ¾ da população angolana. Na sua maioria a população angolana é descendente dos povos Bantu oriundo das regiões dos grandes lagos, que migraram para sul do continente Africano.

Os marcadores autossómicos estudados foram os STRs AmpFℓSTR® Identifiler® Kit. Os resultados em 499 amostras biológicas mostraram que a população angolana se encontra em equilíbrio de Hardy-Weinberg. O poder de discriminação acumulado para os 15 *loci* foi de 0.999999999968716. A probabilidade de exclusão *a priori* variou de 0.4234 (D13S317) a 0.7810 (D2S1338) sendo o valor acumulado de 0.999999028.

No estudo do cromossoma Y foram analisados166 amostras biológicas dos três principais grupos étnico-linguísticos (Bakongo, Kimbundo e Ovimbundo) e foram amplificados 11 *loci* do cromossoma Y através do PowerPlex® Y System. Foi realizado o cálculo das frequências alélicas, da diversidade génica de cada *locus*, assim como da diversidade haplotípica e a capacidade de discriminação da população angolana e de cada grupo étnico-linguístico. No grupo étnico-linguístico Bakongo foram detectados 53/57 haplótipos diferentes, no grupo Kimbundo 53/56 e no Ovimbundo 49/53 haplótipos diferentes. Todos os grupos étnico-linguísticos apresentaram um poder de discriminação superior a 0.9245 na população Angolana. A maior variabilidade molecular foi encontrada dentro de cada grupo populacional, muito mais que entre os diferentes grupos. Os resultados obtidos foram enviados a base de dados Y-STR Haplotype Reference Database (YHRD).

No estudo do ADNmt da região controlo foram analisados 30 amostras da população angolana sendo 10 de cada um dos três principais grupos étnico-linguísticos. Na população angolana foram identificados 27 (90%) haplótipos diferentes dos quais 24 (80%) eram únicos na população estudada. A taxa de heteroplasmia foi de 23.3%, estando estas localizadas na zona poli C da região hipervariável I. O haplogrupo mais representativo na população angolana foi o L1c3b com uma taxa de 13.3%. A diversidade de sequência para os três grupos étnico-linguísticos foi de 1.000±0.0447. A região controlo total apresenta uma diversidade de sequência elevada o que permite um maior poder de distinção entre os haplótipos.

As frequências alélicas e haplotípicas podem ser utilizadas para comparação em estudos populacionais e servir para a construção de uma base de dados forense para a população Angolana.

8. Summary

Summary

GENETIC DIVERSITY IN THE MAJOR POPULATION GROUPS IN ANGOLA – FORENSIC APPLICATION

This is a population genetic study of the major ethno-linguistic groups living in Angola (Bakongo, Kimbundu, Kwanhama, Lunda Tchokwe, Nganguela, Nhaneca-Humbe and Ovimbundu). Three groups, the Bakongo, Kimbundu and Ovimbundu, represent ¾ of the global Angolan population. Most of the Angolan populations descend from the Bantu peoples, who originated in the Great Lakes region and migrated to southern Africa.

The markers studied included the autosomal short tandem repeats (STRs) of the AmpFℓSTR® Identifiler® Kit. The results (in 499 biological samples) showed the Angolan population to be in Hardy-Weinberg equilibrium for these 15 *loci*. The accumulated power of discrimination was 0.999999999968716. The *a priori* exclusion probability ranged from 0.4234 (D13S317) to 0.7810 (D2S1338), and the accumulated value was 0.999999028.

For Y chromosome polymorphisms (11 *loci* of the PowerPlex® Y System), 166 biological samples of the three main ethno-linguistic groups (Bakongo, Ovimbundu and Kimbundu) were studied. Allele frequencies were estimated to assess genetic diversity at each *locus*, as well as haplotype diversity, and their discrimination power within the Angolan population and within each ethnic group; 53/57 haplotypes were seen in the Bakongo, 53/56 in the Kimbundo and 49/53 in the Ovimbundu. The discrimination power was greater than 0.9245 in all groups. The molecular variability found was much greater within each population group than among them. These results were submitted to the Y-STR Haplotype Reference Database (YHRD).

For the study of mtDNA control region, 30 samples (10 from each major group) were analyzed. We identified 27 (90%) different haplotypes in the Angola, 24 (80%) of which are unique to this population. Heteroplasmy rate was 23.3%, variation being located in the poly C area of hypervariable region I. L1c3b was the most representative haplogroup of the Angolan population, with a frequency of 13.3%. Sequence diversity was 1.000 ± 0.0447 in the three ethnic-linguistic groups. The total control region had high sequence diversity, what allows for a greater discrimination power among the haplotypes.

These allele and haplotype frequencies may now be used for genetic population studies and for forensic investigation, in the population of Angola.

9. Referências bibliográficas

ABI PRISM® 377 and Sequencer User's Manual, (1995) Perkin Elmer Applied Biosystems. Applera Corporation.

ABI Prism Genescan 3.1. 1989-1998 Elmer Applied Biosystems. Applera Corporation.

ABI PRISM 373 DNA Sequencer with XL Upgrade User 's Manual. (2001) Applied Biosystems. Applera Corporation.

ABI PRISM® 310 Genetic Analyser. Chemistry Guide and User's Manual, (2001) Perkin Elmer Applied Biosystems. Applera Corporation.

ABI PRISM® 3100 Genetic Analyser. Chemistry Guide and User's Manual, (2001) Perkin Elmer Applied Biosystems. Applera Corporation.

ABI PRISM® Big Dye v3.0 Terminator Cycle Sequencing Ready Reaction kit protocol, (2001). Perkin.

Alberts, B.; Bray, D.; Lewis, J.; *et al*. (1994) Mollecular Biology of the Cell. 3 th edition. New York e London, Garland publishing, Inc.

Alford, R. L. et al. (1994) Rapid and efficient resolution of parentage by amplification of short tandem repeats. Am. J. Hum. Genet. 55:190-461.

Alvaréz, JC.; Entrala, C.; Lorente, JÁ.; Lorente, M.; fernández-Rosaldo, FJ.; Martinez-Espin, E.; Rodrigues, E.; Villanueva, E. (2001) Análise de ADN mitocondrial. Forensica 1 (1): 40-59.

Alves, C.; Gusmão, L.; Barbosa, J.; Amorim, A. (2003) EvaluatingtheinformativepowerofY-STRs: acomparativestudyusingEuropeanandnewAfricanhaplotypedata, Forensic Sci.Int.134:126–133.

Alves, C.; Gusmão, L.; Damascero, A.; Soares, B.; Amorim, A. (2004) Contribution for an African STR database (AmpF/STR Identifiler and Powerplex 16 System) and a reported on genotypic variations. Forensic Sci. Int. 139: 201-205.

Alves, C.; Gusmão, L.; López-Parra, A. M.; Soledad Mesa, M.; Amorim, A.; Arroyo-Pardo, E. (2005) STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpF/STS Identifiler and Powerplex 16 kits. Forensic Sci. Int. 148:239-242.

Anderson, A.; Bankier, T.; Barrel, B.G.; De Bruijn, M.H.L.; Coulson, A.R.; Drouin, J.; Eperon, I.C.; Nierlich, D.P.; Sanger, F.; Schreier, P.H.; Smith, A.J.H.; Staden, R.; Young, I.G. (1981) "Sequence and organization of the human mitochondrial genone". Nature, vol.290:457-464.

Andrews, RM.; Kubacka I.; Chinnery PF.; Lightowlers , RN. Et al., (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA *Nature Genetics* 23**:**147.

Arlequin (Version 2.000), Copyright by Laurent Excoffier, (1995-2000).

Arroyo-Pardo, E.; Gusmão, L.; Lópes-Parra, AM.; Baeza, C.; Mesa, M.; Amorim, A. (2005) Genetic variability of 16Y chromosome STRs in a sample from Equatorial Guinea (Central Africa), ForensicSci.Int.149:109-113.

Ayub, Q. *et al*. (2000) Identification and characterization of novel human Y-chromosomal microsatellites from sequences database information. Nucl. Acids Res. 28, e8.

Bailey, J. (1995) The Molecules of inheritance. Genetics and Evolution, Andromeda Oxford Ltd.

Ballard, J.W.; Whitlock, M.O. (2004) The incomplete history of mitochondria. Mol Ecol., 13 (4): 729-744.

Belle, E.M.; Piganeau, G.; Gardner, M.; Eyre-Walker, A. (2005). An investigation of the variation in the transition bias among various animal mitochondrial DNA. *Gene*. 355:58-66.

Bennett, P. (2000) Microsatellites. J Clin Pathol Mol Pathol 53:177-183.

Bernal, L.; Borjas, L.; Zabala, W.; Portillo, M.; Fernández, E.; Delgado, W.; Tovar, F.; Lander, N.; Chiurillo, M.; Ramírez, J. (2006) Genetic variation of 15 STR autossomal loci in the Maracaibo population from Venezuela. Forensic Sci. Int. 161:60-63.

Bert, F.; Corella, A.; Gene, M.; Perez-Perez, A.; Turbon, D. (2001) Major Mitochondrial DNA Haplotype and Heterogeneity in Highland and Lowland Amerindian Populations from Bolivia. Hum Biol., 73(1): 1-16.

Bini, C.; Ceccardi, S.; Calalongo, C.; Ferri, G.; Falconi, M.; Pelotti, S.; Pappalardo, G. (2003) Population data of mitochondrial DNA region HVIII in 150 individuals from Bolonga (Italy). Progress Forensic Genetic, 9, ICS, 1239, 525-528.

Bolnick, D.A.; Smith, D.G. (2003) Unexpected patterns of mitochondrial DNA variations among Native Americans from the southeastern United States. Am J Phys Anthropol., 122(4): 336-354.

Bonatto, SL.; Salzano, FM. (1997) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc Natl Acad Sci. USA 94:1866-71.

Bozzo, W.; Pena, M.; Ortiz, M.; Lojo, M. (2007) Genetic data from Powerplex® 16 system and Identifiler™ kits from Buenos Aires province (Argentina). Int. J. Legal Med 9:151-153.

Bradman, N.; Thomas, M. (1998) Why Y? The Y chromosome in the study of human evolution, migration and prehistory. Science Spectra. 14.

Brinkmann, B.; Sajantila, A.; Goegge, HW.; Matsumoto, H.; Nishi, K.; Wiegand, P. (1996). Population genetic comparisons among eight populations using allele frequency and sequence data from three microsatellite loci. Eur. J. Hum. Genet. 4:175-182.

Brinkmann, B. (1996) The STR approach. Adv. Forensic Haemogenetics 6:41-51.

Brinkmann, B.; Junge, A.; Meyer, E.; Wiegand, P. (1998) population genetic diversity in relation to microsatellite heterogeneity. Hum Mutat 11 (2):135-144.

Brión, M. J. M. (2002) Estudio de la variabilidade del cromosoma y humano: Análisis del minisatélite MSY1, loci microsatélites y polimorfismos binarios. Tese de doutoramento. Universidade de Santiago de Compostela.

Brown, W.N.; Prager, E.M.; Wang, A.; Wilson, A.C. (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol. 18(4): 225-39.

Brown, T.A. (1995) Gene Cloning an Introduction. 3 rd edition, Chapman & Hall.

Brown, T. A. (2003) Clonagem Gênica e Análise de DNA. 4ª ed. Editora Artmed.

Budowle, B.; Brown, BL. (2001) El uso del análises de ADN en la identificación forense. Forénsica 1:9-22.

Budowle, B.; Allard, M.W.; Wilson, M.R.; Chakraborty, R. (2003) Forensics and mitochondrial DNA: applications, debates, and foundations. Annu Rev Genomics Hum Genet. 4: 119-41.

Butler, J.; Levin, B.C. (1998) "Forensic applications of mitochondrial DNA" Tibtech, vol. 16: 158-162.

Butler, JM. (2001) Biology & Techonology behind STR Markers Forensic DNA Typing. Academic Press, Barcelona 2001.

Butler, JM.; Schoske, R.; Vallone, PM.; Redman, JW.; Kline, MC. (2003) Allele Frequencies for 15 Autosomal STR loci on U.S. Caucasian, African American and Hispanic populations. J. Forensic Sci. 48: 908-911.

Butler, JM.; Schoske, R.; Vallone, PM. (2003) highly multiplexed assays for measuring polymophisms on the Y-chromosome. Int Congress series 1239:301-305.

Butler, J. M. (2005) Forensic DNA Typing: Biology, Technology, and Genetic of STR Markers (2nd Edition). Elsevier Academic Press.

Calloway, C.D.; Reynolds, R.L.; Herrin, G.L.; Anderson, W.W. (2000) The Frequency of Heteroplasmy in the HVII Region of mtDNA Differs across Tissue and Increases with Age. Am J Hum Genet., 66 (4):1384-1397.

Camacho, M.V.; Benito, C.; Figueiras, A.M. (2007) Allelic frequencies of the 15 STR loci included in the AmpFlSTR Identifiler™ PCR Amplification Kit in an autochthonous Sample from Spain, Forensic Sci. Int. 173:241-245.

Cann, R.L.; Stoneking, M.; Wilson, A.C. (1987) Mitichondrial DNA and Human Evolution. Nature 325:31-36.

Carvalho-Silva, DR.; Santos, FR.; Hutz, MH.; Salzano, FM.; Pena, SDJ. (1999) Divergent human Y-chromosome microsatellite evolution rates. J Mol Evol 49:204-214.

Casanova, M.; Leroy, P.; Boucekkine, C.; Weissenbach, J.; Bishop, C.; Fellous, M.; Purrello, M.; Fiori, G.; Siniscalco, M. (1985) A Human Y-Linked DNA Polymorphism and its Potential for estimating genetic and evolutionary distance. Science 230 (4732):1403-1406.

Cavalli-Sforza, LL.; Piazza, A.; Menozzi, P.(1994) History and geography of human genes. Princ. Univ. Press. Princeton, NJ.

Cerri, N.; Ricci, U.; Sani, I.; Verzeletti, A.; De Ferrari, F. (2003) Mixed stains from sexual assault cases: autosomal or Y-chromosome short tandem repeats? Croat Med J 44(3):289-292.

Chen, Y.S.; Torroni, A.; Excoffier, L.; Santachiara-Benerecetti, A.S. *et al*. (1995) Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J Hum Genet 57: 133-149.

Chen, Y.S.; Olckers, A.; Schurr, T.G.; Kogelnik, A.M.; Huoponen, K and Wallace. (2000) mtDNA variation in the South African Kung and Khwe and their genetic relationships to other African populations. Am J Hum Genet 66:1362-1383.

Childs, M. (1949). Umbundu Kinship and character. London.

Coelho, M.; Sequeira, F.; Luiselli, D.; Beleza, S.; Rocha, J. On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. BMC Evol Biol. (2009) 9:80.

Correia, R.: (1990) Angola, datas e factos 1º vol. 1482-1652. Coimbra Editora.

Côrte-Real, F. (1999) Estudo genético-populacional de microssatélites para aplicação forense nos países lusófonos - Análise filogenética e valorização estatística como prova pericial. Dissertação de Doutoramento.

Corte-Real, F.; Andrade, L.; Carvalho, M.; Anjos, M.J.; Gamero, J.; Carracedo, A.; Vieira, D.N.; Vide, M.C. (2000) Comparative analysis STR data for Portuguese spoken countries. Progress in Forensic Genetics 8:212-214.

Darnell, J. et al. (2000) Molecular Cell Biology. 4. ed. W. H. Freeman and Company. Nova Iorque 2000.

De Kniff, P.; Kayser, M.; Caglia, A.; Corach, D.; Fretwell, N.; Gehrig, C.; Graziosi, G.; Heidorn, F.; Herrman, S.; Herzog, B.; Hidding, M. *et al*. (1997) Chromosome Y microsatellite population genetic and evolutionary aspects Int J Legal Med 110:134-140.

De knife, P. (2000) Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y-chromosome. Am J Hum Genet 67:1055-1061.

DNA COMMISSION OF THE ISFH. (1992) Second DNA recommendations, 1991 report concerning recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of DNA polymorphisms. *Int. J. Legal Med., 104:*361-364.

DNA COMMISSION OF THE ISFH. (1997), DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. *Int. J. Legal Med., 110*:175-176.

Edwards, A.; Civitello, A.; Hammond, HA.; Caskey, CT. (1991) DNA Typing and Genetic Mapping with Trimeric and Tetrameric Tandem Repeats. Am. J. Hum. Genet. 49:746-756.

Ellis, N.; Goodfellow, PN. (1989) The mammalian pseudoautosomal region. Genet. 5(12):406-10.

Excoffier, L.; Smouse, P.E.; Quattro, J.M. (1992) Analysis of molecular variante inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131(2): 479-491.

Excoffier, L.; Laval, G.; Schneider, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis Evolutionary. Bioinformatics Online 1 (2005) 47-50.

Felsenstein, J. (1986) Phylip. Phylogeny inference package (version 3.5c). Cladistics 5:164-166.

Ferreira, M.E. Grattapaglia, D. (1998) Introdução ao uso de marcadores moleculares em análise genética. 3ª edição. Brasília Embrapa-Cenargen.

Fujihara, J.; Yuasa, I.; Muro, T.; Lida, R.; Tsubota, E.; Nakamura, H. et al., (2009) Allele frequencies and haplotypes for 28 Y-STRs in Ovambo population, Legal Med. 11: 205–208.

GENEPOP (versão 3.1b), Dezembro 1997 (Update Version).

Gill, P.; Evett, I. (1995) Population genetics of short tandem repeat (STR) loci. Genetica 96:69-87.

Gill, P.; Brenner, C.; Brinkmann, B.; Budowle, B.; Carracedo, A.; Jobling, MA.; de Knijff, P.; Kayser, M.; Krawczak, M.; Mayr, WR. *et al*. (2001) DNA comission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. *Forensic Sci Int* 124: 5-10.

Gomes, V.; Sánchez-Diz, P.; Alves, C.; Gomes, I.; Amorim, A.; Carracedo, Á.; Gusmão, L. (2009) Population data defined by 15 autosomal STR loci in Karamoja population (Uganda) using AmpF[STR Identifiler Kit. Forensic Sci. Int. Genetics 3:55-58.

Gomes, V.; Alves, C.; Amorim, A.; Carracedo, Á.; Sánchez-Diz, P.; Gusmão,L. (2009) Nilotes from Karamoja,Uganda:Haplotypedatadefinedby17Y-chromosomeSTRs, Forensic Sci.Int.Genet. doi:10.1016/j.fsigen.2009.07.001.

Graves JAM (1995). The evolution of mammalian Sex chromosomes and the origin of sex determining genes. *Philos Trans R Soc LondB* 350: 305-312.

Green R.; Krause J.; Briggs A.; Maricic T. et al. (2010) A Draft Sequence of the Neandertal Genome. Sci Journals 5979(328):710-722

Guo, S.W.; Thompson, E.A. (1992) Performing the exact test of Hardy-Weinberg proportion for multiples alleles. Biometrics 48:361-372.

Gusmão, L.; Sánchez-Diz, P.; Calafell, F.; Martín, P.; Alonso, CA.; Álvarez-Fernández, F.; Alvez, C.; Borjas-Fajardo, L.; Bozzo, WR. *et al*. (2005) Mutation rates at Y chromosome specific microsatellites. Human Mutation 26(6):520-528.

Gutierrez, G.; Sanchez, D.; Marin, A. (2002) A reanalysis of the ancient mitochondrial DNA sequences recovered from Neandertal bones. Mol. Biol. Evol. 19:1359-66.

Hallenberg, C.; Simonsen, B.; Sanchez, J.; Morling, N. (2005) Y-chromosome STR haplotypes in Somalis, Forensic Sci. Int. 151:317–321.

Hammer, M. (1994). A recent insertion of na Alu element on the Y chromosome is a useful marker for human population studies. Mol Biol Evol 11:749-761.

Hammer, MF. (1995) A recent common ancestry for human Y chromosome. Nature. 378:376-378.

Hammer, MF.; Zegura, S. (1996) The role of the Y chromosomes in human evolutionary studies. Evol. Anthropol. 5:116-134.

Hammer, MF.; Karafet, TM.; Redd, AJ.; Jarjanazi, H.; Santachiara-Benerecetti, S.; Soodyall, H.; Zegura, SL. (2001) Hierarchical patterns of global human Y-chromosome diversity. Mol. Biol. Evol. 18:1189-1203.

Hammond, HA.; Caskey, CT. (1992) Personal Identification via Short Tandem Repeats. Proceedings from the Third International Symposium on Human Identification. Promega Corporation. 162-175.

Hassold, TJ.; Pettay, D.; Freeman, SB.; Grantham, M.; Takaesu, N. et al., (1991) J Med Genet 28:159-162.

Hernández-Gutiérrez, S.; Hernández-Franco, P.; Martínez-Tripp, S.; Ramos-Kuri, M.; Rangel-Villalobos, H. (2005) STR data for 15 loci in a population sample from the central region of Mexico. Forensic Sci. Int. 151: 97-100.

http://www.gep-isfg.org.

http://ghr.nlm.nih.gov/handbook.pdf (2010) Genetics Home Reference, Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine

National Institutes of Health Department of Health & Human Services Published September 26, 2010.

Inman, K.; Rudin, N. (1997) An Introduction to Forensic DNA Analysis. CRC Press. New York. 256.

Jacewi, R. *et al*. (2004) Population genetics of the Identifiler system in Poland. Int. Cong. Series 1261: 229-232.

Jarreta, M.B.M. (1999) La prueba del ADN en medicina Forense – La Genética al servicio de la Ley en el análisis de indicios criminales y en la investigación biológica de la paternidad. Barcelona- Masson SA. Barcelona.

Jeffreys, AJ.; Wilson, V.; Thein, SW. (1985a) Hypervariable minisatellite regions in human DNA. Nature. 314:67-73.

Jobling, MA.; Pandya, A.; Tyler-Smith, C. (1997) The Y chromosome in forensic analysis and paternity testing. Int. J. Legal Med. 110:118-124.

Jobling, MA.; Tyler-Smith, C. (2003) The Y Chromosome: an evolutionary marker comes age. Nature Reviews.4.

Johnson, A.A.; Johnson, K.A. (2001) Fidelity of Nucleotide Incorporation by Human Mitochondrial DNA polymerase. J Biol Chen., 276 (41): 38090-38096.

Jones, S. (2004) Y A descendência do homem. Trad. Patrícia Marques da Fonseca. Gradiva publicações pág. 277

Kayser, M.; Caglia, A.; Corach, D. et al. (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int. J Legal Med 110:125-133.

Kayser, M.; de Kniff, P.; Dieltjes, P.; Krawczak, M.; Nagy, M.; Zerjal, T. *et al*. (1997a) Applications of microssatellite-based Y chomosome haplotyping. Electrophoresis; 110:125-133.

Kayser, M.; Brauer, S.; Willuweit, S.; Schadlich, H.; Batzer, MA.; Zawacki, J.; Prinz, M.; Roewer, L.; Stoneking, M. (2002) Y-chromosomalshort tandem repeat haplotype reference database (YHRD) for U.S. populations. J Forensic Sci 47: 513-9.

Kimpton, C. *et al*. (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Appl. 3:13-22.

Kivisild, T.; Tolk, H.V.; Parik, J.; Wang, Y.; Papiha, S.S.; Bandelt, H-J. Villems, R. (2002) The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol 19:1737-1751.

Krings, M.; Stone, A.; Schmitz, RW.; Krainitzki, H.; Stoneking, M.; Paabo, S. (1997) Neandertal DNA sequences and the origin of modern humans. Cell. 90:19- 30.

Lahn, BT.; Pearson, NM.; Jegalian, K. (2001) The human Y chromosome, in the light of evolution. Nat Rev. Genet. 2:207-216.

Laréu, M.V.; *et al*. (1998) Further exploration of new STRs of interest for forensic Genetic Analysis. Progress in Forensic Genetic 7, Eds Olaisen & Brinkmann © Elsevier Science B. V. Lewontin R. C. e Hartl D. L. (1991) Population genetic in forensic DNA typing. Science 254: 1745-1750.

Lee, HC.; *et al*. (1994) DNA typing in forensic. Theory and background. Am J Forensic Mec. Pathol. 15(4):269-282.

Lee, H.Y.; *et al*. (2006) Mitochondrial DNA control region sequences in Koreans: identification of useful variable sites and phylogenetic analysis for mtDNA data quality control. Int. J. Legal Med. 120:5–14.

Lee, HY.; Song, I.; Ha, E.; Cho, S-B.; Yang, WI.; Shin, K-J. (2008) mtDNAmanager: a Web-based tool for the management and quality analysis of mitochondrial DNA control-region sequences. BMC Bioinformatics 9:483.

Iida, R.; Tsubota, E.; Matsuki, T. (2001) Identification and characterization of two novel human polymorphic STRs on the Y chromosome. *Int J Legal Med*, 115, 54 - 56.

Litt, M.; Luty, JA. (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet.44:397-401.

Lodish, H.; Baltimore, D.; Berk, S.; Zipursky, L.; Matsudaria, P.; Darnel, J. (1995). Mollecular Cell Biology 3th edition. Scientific American Books.

Longley, M.J.; Nguyen, D.; Kunkel, T.A.; Copeland, W.C. (2001) The Fidelity of Human DNA Polymerase γ with and without Exonucleolytic Proofreading and p55 Accessory Subunit. J Biol Chen., 276(42):38555-38562.

Lopes, V.A.; Serra, J.;Gamero, L.; Sampaio, F.; Balsa *et al*. (2009) Allelic frequency distribution of 17 STRs from Identifiler and PowerPlex-16 in Central Portugal area and the Azores archipelago. Forensic Sci. Int. Genetics 4: 1-7.

Lutz, S.; Weisser, H-J.; Heizmann, J.; Pollak, S. (1997) A third hypervariable region in the human mitochondrial d-loop. Hum. Genet., 101, 384.

Lutz, S.; Wittig, H.; Weisser, H.-J. Heizmann, J.; Junge, A.; Dimo-Simonin, N.; Parson, W.; Edelmann, J.; Anslinger, K.; Jung, S. Augustin, C. (2000) Is it possible to differentiate mt DNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? Forensic Sci. Int., 113: 97-101.

Maca-Meyer, N.; González, AM.; Larruga, JM.; Flores, C.; Cabrera, VM. (2001) Major genomic mitocondrial lineages delineate early human expansions. BMC Genet. 2:13-21.

Manual AmpłSTR® Profiler PlusTM PCR Amplification Kit User´s Manual (PE, Applied Biosystems, 1997.

Mateu, E.; Comas, D.; Calafell, F.; Abade, A.; Bertranpetitt, J. (1997) A tale of two islands: population history and mitochondrial DNA sequence variation of Bioko and São Tomé, Gulf of Guinea. Ann. Hum. Genet. 61:507-518.

Mathias, N.; Bayes, M.; Tyler-Smith, C. (1994) Highly informative compound haplotypes for the human Y chromosome. Hum Mol Genet. 3:115-123-123.

MCculloch, M. (1952) The Ovimbundu of Angola. London.

Melo, M.; *et al*., (2006), as Origens do grupo etnico-linguistico Bantu Ovimbundo: estudos genéticos em STRs autossomicos, comuniçao ao V Congresso de Med Legal em Portugal.

Mesa, NR.; Mondragón, MC.; Soto, ID.; Parra, MV.; Duque, C.; Ortiz-Barrientos, D.; Garcia, LF.; *et al.* (2000) Autosomal, mtDNA, and Y-chromosome diversity in amerinds: pre- and post-colombian patterns of gene flow in South America. Am. J. Hum. Genet. 67:1277-1286.

Mishmar, D.; Ruiz-Pesini, E.; Golik, P.; Macaulay, V; Clark, A.G.; Hosseini, S.; Brandon, M.; *et al*. (2003) Natural selection shaped regional mtDNA variations in humans. Proc Natl Acad Sci USA., 100(1): 171-176.

Mitchell, J.; Hammer, MF. (1996) Human evolution and the Y chromosome. Genet. and Develop. 6:737-742.

Mullis, KA.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor SympQuant Biol* 51:263-273.

Muro, T.; Fujihara, J.; Imamura, S.; Nakamura, H.; Yasuda, T.; Takeshita, H.; (2008) Allele frequencies for 15 STR loci in Ovambo population using AmpFLSTR® Identifiler Kit, Int. J Legal Med. 10:157-159.

Nei, M.; Saitou, N. (1987) The Neighbor Joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406- 425.

Nei, M.; Tajima, F. (1981) DNA polymorphism detectable by restriction endonucleases. Genetics, 97(1): 145-163.

Neto, J. (1974) Etnias and Cultures of Angola. Edition of the Institute of Investigation of Angola.

Oliver, R.; Fage, JD.(1980) Breve história da África. Editora/Sá Costa. Lisboa, Portuga.

Pakendorf, B.; Stoneking, M. (2005) Mitochondrial DNA and human evolution. *Annu. Rev. Genomics Hum. Genet* 6:165-183.

Parson, W.; Bandelt, H-J. (2007) Extended guidelines for mtDNA typing of population data in forensic science. Forensic Sci Int.: Genetics, 1 (1): 13-19.

Pascali, VL.; Dobosz, M.; Brinkmann, B. (1999) Coordinating Y-chromosomal STR research for the Courts [editorial] Int. J Legal Med. 112(1):1

Pena, SDJ.; Carvalho-Silva, DR.; Alves-Silva, J.; Prado, VF. (2000) Retrato Molecular do Brasil. Ciência Hoje 27:16-25.

Pereira, L.; Macaulay, V.; Torroni, A.; Scozzari, R.; Prata, M.J.; Amorim, A. (2001) Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade Ann Hum Genet 65:439-458.

Peric, M.; Lauc, LB.; Klaric, IM.; Janicijevic, B.; Rudan, P. (2005)Review of Croating Genetic Heritages as Revealed by Mitochondrial DNA and Y Chromosomal Lineages. Croat Med J 46(4):502-513.

Pestoni, C.; Lareu, M. V.; Rodriguez-Calvo, M.S. and Carracedo, A. (1995) The use of the STRs HUMTH01, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: Validation studies and population data for Galicia (NW Spain) Int. J. Legal Med. 107 (6) 283-290.

PHYLIP (Version 3.573c), Joseph Felsenstein, Department of Genome Sciences, University of Washington.

Prak, E.T.; Kazazian, H. Jr. (2000) Mobile elements and the human genome. *Nat Rev Genet*, 1, 134 -144.

Prinz, M.; Sansone, M.Y. (2001) Y Chromosome-specific short tandem repeats in forensic casework Croat Med J 2001;42:288-291.

Quintana-Murci, L.; Semino, O.; Bandelt, H.-J.; Passarino, G.; McElreavey, K.; Santachiara-Benerecetti, AS. (1999) Genetic evidence of an early exit of homo sapiens sapiens from Africa through eastem Africa. Nat Genet 23 437-441.

Quintana-Murci, L.; Krausz, C.; McElreavey, K. (2001) The human Y chromosome: function, evolution and disease. Forensic Sci Int 118:169-181.

Raymond, M.; Rousset, F. GENEPOP (version 3.4), 2003.

Redd, A.J.; Agellon, A.B.; Kearney, V.A.; Contreras, V.A.; Karafet, T.; Park, H.; de Knijff, P.; Butler, J.M. Hammer, M.F. (2002) Forensic value of 14 novel STRs on the human Y chromosome. *Forensic Sci Int*, 130, 97 - 111.

Redinha, J. (1974) Etnias e Culturas de Angola. Edição do Instituto de Investigação de Angola.

Richter, C. (1995) Oxidative damage to mitochondrial DNA and its relationship to ageing. Int J Biochem Cell Biol., 27(7): 647-53.

Roewer, L.; Armemann, J.; Spurr, N.K.; Grzeschik, K.H.; Epplen, J. (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum.Genet.* 89, 389-394.

Roewer, L.; Kayser, M.; Dieltjes, P.; Nagy, M.; Bakker, E.; Krawczak, M. e de Knijff, P. (1996) Analysis of molecular variance (AMOVA) of Y chromosome specific microsatellites in two closely related human populations. *Hum Mol Genet*, 5 (7), 1029-1033.

Roewer, L.; Nagy, M.; Schmidt, P.; Epplen, JT.; Schroder, G. (1993) Microsatellite and HLA class II oligonucleotide typing in a population of Yanomani Indians in: Pena , DJ. Chakraborty R. Epplen JT. Jeffreys AJ. Eds DNA fingerprinting State of the Science Birkäuser Verlag, Based, Boston, Berlin 221-230.

Roewer, L.; Kayser, M.; Dieltjes, P.; Nagy, M.; Bakker, E.; Krawczak, M.; De Knijff, P. (1996) Analysis of molecular variance (AMOVA) of Y- chromosome specific microsatellites in two closely related human populations. Hum Mol Genet 5(7) 1029-1033.

Roewer, L. Krawczak, M. Willuweit, S. Nagy, M. Alves, C. et al.(2001) Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci. Int. 118:106-113.

Roewer, L.; Croucher, PJP.; Willuweit, S.; Lu, TT.; Kayser, M.; Lessig, R.; De Knijff, P.; *et al*. (2005) Siggnature of recente historical events in the European Y-chromosomal STR haplotypes distribution. Hum Genet 116:279-291.

Rolf, B.; Keil, W.; Brinkmann, B.; Roewer, L.; Fimmers, R. (2001) Paternity testing using Y-STR haplotypes: assigning a probability for paternity in cases of mutations. Int J Legal Med. 115(1):12-15.

Rosa, A.; Brehm, A.; Kivisild, T.; Metspalu, E.; Villems, R. (2004) MtDNA profife of west Africa Guineans: towards a better understanding of the senegambia region. Ann Hum Genet 68:340-352.

Rosa, A.; Ornelas, C.; Brehm, A.; Villems, R. (2006) Population data on 11 Y-chromosome STRs from Guine-Bissau, Forensic Sci. Int. 157:210-217.

Ruitberg, C.M.; Reeder, D.J.; Butler, J.M. (2001) STR Base: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res*, 29, 320 - 322 Disponível na URL: http://www.cstl.nist.gov/biotech/strbase.

Saitou, N.; Nei, M. (1987) The Neighbor Joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Salas, A.; Richards, M.; De La Fe, T.; Lareu, M.V.; Sobrinho, B.; sanchez-Diz, P.; Macaulay, V.; Carracedo, A. (2002) The making of the African mtDNA landscape Am J Hum Genet 71:1082-1111.

Santos, M.V.; Anjos, M.J.; Andrade, L.; Vide, M.C.; Corte-Real, F.; Vieira, D.N. (2004) Population genetic data for the STR loci using the AmpFISTR® IdentifilerTM Kit in Bahia, Brazil, Int Congress Series 1261:219-222.

Schneider, PM.; Meuser, S.; Waiyawuth, W.; Seo, Y.; Rittner, C. (1998) Tande repeat structure of the duplicated Y- Chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations Forensic Sci. Int. 97 (1): 61-70.

Schneider, S.; Roessli, D.; Excoffier, L. (2000) ARLEQUIN: A Software for Population Genetics Data Analysis. Genetics and Biometry. Lab, Dept. of Antropology, University of Geneva.

Schuller, W.; Fereday, L.; Scheithauer, R. (2001) Interpol Handbook on DNA data Exchange na practice: Recommendations from the Interpol DNA Monitoring Expert group, 1:70.

Schults, T.; Hummel, S.; Herrmann, B. (1999) Amplification of Y chromosomal STR´s from ancient skeletal material. Hum Genet 104:164-166.

Seo, Y.; Stradmann-Bellinghausen, S.; Rittner, C.; Takahama, K.; Schneider, P. (1998) "Sequence polymorphism of mitochondrial DNA control region in Japanese" Forensic Sci Int. 97:155-164.

Sibille, I.; Duverneuil, C.; Grandmaison, GL.; Guerrouache, K.; Teissière, F.; Durigon, M.; *et al*., (2002) Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa Forensic Sci Int 125(2-3)212-216.

Skaletsky, H.; Kuroda-Kawaguchi, T.; Minx, PJ.; Cordum, HS.; Hillier, HD.; Brown, LG.; et al., (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classe Nature 423:825-837.

Strachan, T.; Read, A.P. (2002) Genética Molecular Humana. Trad. Henrique Bunselmeyer Ferreira et al. 2ª ed. Porto Alegre: Artmed Editora.

Su, H. and Lau, YF. (1993). Identification of the transcriptional unit, structural organization, and promoter sequence of the human Sex-determining region Y (SRY) gene, using a reverse genetic approach. *Am J Hum Genet,* 52: 1503-1507.

Tillmar, A.O.; Bäckström, G.; Montelius, K. (2009) Genetic variation of 15 autosomal STR loci in a Somali population, Forensic Sci. Int. Genetics 4:19-20.

Tishkoff, SA.; Reed, FA.; Friedlaender, FR.; Ehret, C.; *et al*. (2009) The genetic structure and history of Africans and African Americans. *Science* 324:1035–1044.

Torroni, A.; Huoponen, K.; Francalacci, P.; Petrozzi, M.; Morelli, L.; Scozzari, R.; Obinu, D.; Savontaus, M.L.; Wallace, D.C. (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835-1850.

Tully, G.; Bär, W.; Brinkmann, B.; Carracedo, A.; Gill, P.; Morling, N.; *et al*. (2001) Considerations by the European DNA profiling (EDNAP) group on the working

practices, nomenclature and interprection of mitocondrial DNA profiles. Forensic Sci Int., 124, 83-91.

TreeView(versão1.5.2)(1998)RodericD.M.Page.http//:taxonomy.zoology.gla.ac.uk/rod/rod.html.

Upholt, W.B.; Dawid, I.B. (1977) Mapping of mtDNA of individual sheep and goats: rapid evolution in the D-loop region. Cell 11 571-583.

Urquhart, A. *et al.* (1995) Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. Biotechniques 18: 116-121.

Urquhart, A.; *et al*. (1996) Selection of STR loci for forensic identification sytems. Advances in Forensic Haemogenetics 6:115-117.

Van Oven, M. Kayser, M. (2008) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mut 29: E386-E394.

Verma, R.S.; Babu, A. (1995) human chromosomes : Principles and techinique. 2ª. Ed New York.McGraw-Hill.

Vigilant, L.; Pennington, R.; Harpending, H.; Kocher, TD.; Wilson, AC. (1989) Mitochondrial DNA sequences in single hairs from a sourthern African population. Proc. Natl. Acad. Sci. USA. 86(23):9350-9354.

Vigilant, L.; StoneKing, M.; Harpending, H.; Hawkes, K.; Wilson, AC. (1991) African populations and the evolution of human mitochondrial DNA. Science. 27; 253 (5027):1503-1507.

Vogt, PH.; Affara, N.; Davey, P.; Hammer, M.; Jobling, MA.; Lau, YF.; et al. (1997) Report of the third Internacional Workshop on Y Chromosome Mapping. Cytogenet Cell Genet; 79 (1-2): 1-20.

Wallace, D.; Ye, J.; Neckelmann, S.; Singh, G.; Webster, K.; Greenberg, B. (1987) Sequence analysis of cDNAs for the human and bovine ATF synthase subunit: mitochondrial DNA genes sustain seventeen times more mutations. Curr Genet, 12, 81-90.

Wallace, DC.; Torroni, A. (1992) American indian prehistory as written in the mitochondrial DNA: a review. Hum. Biol. 64:403-416.

Walsh, P.S.; Metzger, D.A.; Higuchi, R. (1991) Chelex 100 as a Medium for Simple Extraction of DNA for PCR-based Typing from Forensic Material, Biotechniques 10, 506-513.

Walsh, PS.; Fildes, NJ.; Reynolds, R. (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus VWA. Nucleic Acids Res 24:2807-2812.

Watson, E.; Forster, P.; Richards, M.; Bandelt, H.-J. (1997) Mitochondrial footprints of human expansions in Africa Am J Hum Genet 61:691-704.

Weber, J.L.; May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction Am. J. Hum. Gen. 44: 388-396.

Weber, JL.; Wong, C. (1993) Mutation of human DNA short tandem repeats. Hum. Mol. Genet 2:1123-1128.

White, PS.; Tatum, OL.; Deaven, LL.; Longmire, JL. (1999) New male-specific microsatellite markers from the human Y chromosome Genomics 57:433-437.

Wilson, M. R.; Dizinno, J. A.; Polanksey, D.; Replogle, J.; Budowle, B. (1995) Validation of mitochondrial DNA sequencing for forensic casework analysis. Int J Legal Med. 108(2): 68-74.

Wyman, A.R.; White, R. (1980) A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Sciences of the United States of America* 77, 6754–6758.

10 – Anexos

Anexo I – Ficha preenchida no acto da colheita das amostras

Angola – grupos étnico-linguísticos

Ficha número: _____

Local de nascimento_____País_____

Grupo étnico-linguístico a que pertence_____

Residente em_____

Pai: local de nascimento:_____

Grupo étnico-linguistico:_____

Avô paterno: local de nascimento_____

Avó paterna: local de nascimento_____

Mãe: local de nascimento_____

Grupo étnico-linguístico:_____

Avô materno: local de nascimento_____

Avó materna: local de nascimento_____

Anexo II – Protocolo de extracção de ADN pelo método de *Chelex® 100* a partir de manchas de sangue (Wash et al., 1991). Antes foram numerados e identificados os tubos.

- Com o auxílio de uma pinça e tesoura cortar um fragmento de aproximadamente 3 mm^2 de área na mancha e colocar no tubo
- Num tubo eppendorf de 1.5 ml, adicionar 1ml de H_2O autoclavada
- Deixar repousar à temperatura ambiente durante 30 minutos
- Centrifugar por 3 minutos a 14 000-15 000 rotação por minuto (rpm)
- Eliminar o sobrenadante, deixando cerca de 30 µl com a mancha em cada tubo
- Adicionar 170 µl de Chelex® a 5% (vt = 200 µl) e de seguida incubar a 56 °C durante 15-30 minutos
- Agitar a amostra no vortex durante 5-10 segundos e depois submete-lá a uma fervura durante 8 minutos (as tampas dos tubos devem ser previamente perfurados com uma agulha estéril para evitar que se abram durante a fervura)
- Agitar no vortex durante 5-10 segundos e posteriormente centrifugar durante 2-3 minutos a 14000-15 000 rpm numa microcentrífuga
- O ADN extraído está em condição de ser amplificado

Todas as reacções foram realizadas utilizando controlos positivo e negativo para cada sistema.

Nota: A área a ser utilizada assim como os materiais devem ser previamente esterilizados com radiações Ultra Violeta (U.V.).

Caso se proceda de imediato à amplificação do ADN das amostras, estas devem ser armazenadas no frigorífico a 4ºC. De contrário devem ser congeladas a -20ºC. Posteriormente, para se proceder a amplificação do ADN, e após descongelamento à temperatura ambiente, as amostras devem serem agitadas e vortexadas, de acordo o referido anteriormente no protocolo de extracção.

Anexo III – Procedimentos de amplificação do ADN pelo método de PCR, de STRs autossómicos e do cromossoma Y.

Na reacção, foram utilizados um par de iniciadores (primers) complementares de ambas as extremidades do fragmento de ADN a amplificar, os flanqueadores da região-alvo.

Para o Kit multiplex AmpFℓSRT do Identifiler (Applied Biosystems) foram cumpridos os seguintes procedimentos de amplificação:

- Marcar um tubo 0.2 ml por cada amostra com número da análise e Kit de amplificação.
- Descongelar "Reaction Mix", "*Primers* Set" e "Taq Gold"; vortexar e centrifugar rapidamente.
- Preparar a "Master Mix" num tubo eppendorf de 1.5 ml
  - ✓ (n + 2) x 10.5 µl "Reaction Mix"
  - ✓ (n + 2) x 0.5µl "Taq Gold"
  - ✓ (n +2) x 5.5 µl *Primers* Set
- Vortexar e centrifugar a "Master Mix"
- Repartir 7.5 µl de Tampão TE por tubo de amostras.
- Repartir 15 µl de "Master Mix" por cada tubo de amostra
- Adicionar 2.5 µl de ADN

Em cada reacção de amplificação é fundamental a utilização de controlo positivo (PCR+) e controlo negativo (PCR-). O controlo (-) consiste apenas na adição de tampão TE em vez de ADN. As amostras foram colocadas no termociclador GenAmp® PCR System 2700, da Applied Biosystems.

Para o Kit multiplex PowerPlex Y (Promega) foram cumpridos os seguintes procedimentos:

- Marcar um tubo de 0.2 ml por cada amostra com numero da análise e marcadores a amplificar.
- Descongelar o Tampão (Gold ST *R 10 X Buffer), os primers (PowerPlex®Y 10X Primer Pair Mix) e a Taq Gold (*Applied Biosystems*); vortexar e centrifugar rapidamente (excepto os primers).
- Preparar a Master Mix num tubo eppendorf de 1.5 ml
  - ✓ (n+2) x 17.45 µl de agua desionizada autoclavada (amostras seguras).
  - ✓ (n+2) x 2.5 µl Tampão
  - ✓ (n+2) x 2.5 µl Primers
  - ✓ (n+2) x 0.55 µl Taq Gold
- Vortexar e centrifugar.
- Repartir 23 µl de master mix por cada tubo da amostra.
- Adicionar 2 µl de ADN (amostras seguras)

Em cada reacção de amplificação é fundamental a utilização de controlo positivo (PCR+) e controlo negativo (PCR-). O controlo (-) consiste apenas na adição de H_2O

em vez de ADN. As amostras foram colocadas no termociclador GenAmp® PCR System 9600, da Applied Biosystems.

Anexo IV – Detecção do produto amplificado e designação alélica

A - Para o AmpFℓSRT do Identifiler (Applied Biosystems):

- Marcar os tubos de acordo com o número total das amostras e deixa-las a descontaminar durante 20 minutos.
- Num tubo de 1.5 ml foi adicionado (N+2) x 15µℓ de Hi-Di formamida
- Adicionar a este tubo (N + 2) x 0.5 µℓ de Liz-500
- Tapar o tubo, agitar no vortex de 3-5 segundos e centrifugar a mistura
- Distribuir 15.5 µℓ da mistura por cada tubo
- Adicionar 1µℓ de produto de PCR amplificado

B – Para o PowerPlex Y (Promega):

- Num tubo de 1.5 ml foi adicionado (N+2) x 12 µℓ de Hi-Di formamida
- Adicionar a este tubo (N + 2) x 0,5 µℓ de ILS-600
- Tapar o tubo, agitar no vortex de 3-5 segundos e centrifugar a mistura
- Distribuir 12.5 µℓ da mistura por cada tubo
- Adicionar 1 µℓ de produto de PCR amplificado

Para a detecção de ambos os Kits foram segue-se o procedimento seguinte:

- Tapar cada tubo com um septo, agitar no vortex e centrifugar
- Aquecer durante 3 minutos a 95º C para desnaturação
- Colocar as amostras no suporte de 48 posições e inseri-lo no sequenciador ABI *Prism® 310 Genetic Analyser*

Durante a electroforese, após uma excitação com um laser de árgon, cada um dos fluorocromos, emite fluorescência a um cumprimento de onda diferente, detectado depois pelo sequenciador automático. Os sinais de fluorescência são separados por difracção de acordo com o comprimento de ondas emitidos.

Anexo V – Procedimentos de amplificação da região controlo.

Para a amplificação da região controlo do ADNmt foram utilizados os seguintes primers: L15971 a 100 µM, H017 a 100 µM, L16450 a 100 µM e H599 a 100 µM. Na preparação da Primers Mix adicinou-se 2 µl de cada primer [L1597 /H017/L16450 /H599=8 µl] em 92 µl de água. Todos os primers devem estar a uma concentração de 2 µM na primer Mix.

A mistura da reacção para a amplificação da região controlo foi a seguinte: 5 µl de Master Mix Qiagen, 3.5 µl de água Qiagen (2x) e 1 µl da primer Mix. No kit Qiagen estão incluídos quatro oligonucleotidos (deoxinucleotidos trifosfatados – dATP, dTTP, d GTP e dGTC), Taq polimerase, MgCl2 e Tampão.

Procedimentos:

Marcar tubos MicroAmp 0.2 com identificação da amostra. Em cada reacção de amplificação é fundamental a utilização de controlo positivo (PCR +) e controlo negativo (PCR -). O controlo (+) consiste na amplificação de uma amostra que já foi devidamente tipada e amplificada, o controlo (-) consiste apenas na amplificação da mistura. Posteriormente agitar, centrifugar o tubo com a mistura preparada e repartir 9.5 µl da mistura por cada tubo. Adicionar 0.5 µl de ADN, perfazendo um volume final de reacção de 10 µl. As amostras foram colocadas no termociclador GenAmp® PCR System 9700 (Applied Biosystems).

Anexo VI – Procedimentos de purificação dos fragmentos amplificados por ExoSap-IT

Procedimento:

- Marcar tubos MicroAmp 0.2 com identificação da amostra e em duplicado, para directos e reversos.
- Repartir 1.0 µl de ExoSap-IT por cada tubo (1.67 U de SAP mais 0.66 U de Exo diluída)
- Adicionar 1.5 µl de produto amplificado para T01F e T02F.

O volume total final de 2.5 µl por cada tudo é colocado no termociclador GenAmp® PCR System 9700 (Applied Biosystems) com um programa de purificação de 37 ºC durante 15 minutos, seguido de 85 ºC durante 15 minutos.

Anexo VII – Sequenciação cíclica (directa /reversa) dos produtos amplificados (9700)

Após amplificadas e purificadas com ExoSap as amostras foram sequenciadas. Foram realizados os seguintes procedimento:

- Adicionar 1 µl da BigDye V3.1 Terminator Reaction Mix, 1 µl de Better Buffer, 0.5 µl do Primer da região a sequenciar, directo e reverso.
- Repartir 2.5 µl da mistura por cada tubo da placa.
- Adicionar 2.5 µl de ADN (produto purificado com ExoSap).

Perfazendo volume de 5 µl e as amostras são colocadas no GenAmp® PCR System 9700 (Applied Biosystems). Todas as amostras foram sequenciadas no sentido directo e reverso.

Em cada reacção de sequenciação é fundamental a utilização de controlo (+) (Seq +) e controlo (-) (Seq -). O controlo (+) consiste na sequenciação da amostra que já foi devidamente tipada, o controlo (-) consiste na sequenciação de H_2O e BigDye Terminator Reaction Mix + Primer L15971 e os Primer H017 ou Primer H599

Anexo VIII – Procedimentos de purificação dos fragmentos sequenciados por X Terminator/SAM solution

- Retirar a placa do termociclador e centrifuga-la durante 2 minutos a 2500 rpm.
- Preparar mistura, consoante o número de amostras a purificar:
    - ✓ 16.4 µl por amostra de SAM solution – (colocar em banho Maria a 37ºC durante 15 minutos).
    - ✓ 3.7 ul por amostras de X terminator (cortar a ponta da pipeta)
    - ✓ 30 ul por amostra de água desionizada autoclavada
- Repartir 50 ul da mistura em cada poço da placa – (não deixar precipitar).
- Tapar os poços com tampas ou parafilme colante.
- Agitar no vortex durante 30 minutos.
- Centrifugar a placa durante 2 minutos a 2500 rpm.

As amostras estão prontas a ir para o sequenciador automático 3130 (Applied Biosystems). O aparelho processa 4 amostras simultaneamente.

Anexo IX – Artigo, poster e comunicação escrita

Y-STR HAPLOTYPES IN THREE ETHNIC LINGUISTIC GROUP OF ANGOLA POPULATION
MM. Melo[a,b]*, M. Carvalho[c], V. Lopes[c], MJ. Anjos[c], A. Serra[c], DN. Vieira[d], J. Sequeiros[b], F. Corte-Real[d] Forensic Sci Int: Genetics doi:10.1016/j.fsigen.2010.08.002.

GENETIC STUDY OF 15V LOCI OF IDENTIFILER SYSTEM IN ANGOLA POPULATION
MM. Melo[a,b]*, M. Carvalho[c], V. Lopes[c], MJ. Anjos[c], A. Serra[c], DN. Vieira[d], J. Sequeiros[b], F. Corte-Real[d] Forensic Sci Int: Genetics doi:10.1016/j.fsigen.2010.03.010How to Cite or Link Using DOI

MtDNA DATA of 3 ETHNIC GRROPS FROM ANGOLA
Afonso Costa H.[1], Melo M.M.[1,2], Carvalho M.[1], Anjos M.J.[1], Lopes V.[1], Serra A.[1], Vieira D.N.[3,4], SequeirosJ.[2], Côrte-Real, F.[3,4]
Presented no Haploid DNA Markers in Forensic Genetics – Charité, Berlin, Germany April 22 nd – 24 th, 2010

ESTUDO DE 11 STRs DO CROMOSSOMA Y NOS TRES PRINCIPAIS GRUPOS ETNICO-LINGUISTICOS DE - Autores: MM. Melo, V. Lopes, MJ. Anjos, M. Carvalho, DN. Vieira, J. Sequeiros, F. Corte-Real – Apresentada ao 8º Congresso de Medicina Legal em Portugal, Elvas, 6 e 7 de Novembro de 2009.

AS ORIGENS DO GRUPO ÉTNICO-LINGUÍSTICO BANTU OVIMBUNDO (ANGOLA): ESTUDO GENÉTICO EM STRs AUTOSSÓMICOS - Autores: MM. Melo, F. Corte-Real, J. Sequeiros, M.C. Vide, V. Lopes, M.J. Anjos, M. Carvalho – Apresenta ao V Congresso de Medicina Legal em Portugal, Ericeira 10 e 11 Novembro de 2006.

ESTUDO DE 15 STRs AUTOSSOMICOS NA POPULAÇAO DE ANGOLA- Autores: MM. Melo, F. Corte-Real, J. Sequeiros, M.C. Vide, V. Lopes, M.J. Anjos, M. Carvalho – Apresenta ao V Congresso de Medicina Legal em Portugal, Covilhã, 11 E 12 Novembro DE 2005.

.

Forensic Science International: Genetics xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Forensic Population Genetics – Letter to the Editor

Y-STR haplotypes in three ethnic linguistic groups of Angola population

Dear Editor,

Angola is located in the African continent, in the area of Southern Africa. The Angola population was originated from Bantu people. The ethnic linguistic groups Bakongo, Kimbundo and Ovimbundo represent of the population and the remaining are represented by other Bantu groups (Fig. 1).

The aim of this study was to characterize the three main ethnic linguistic groups in Angola using 11 *loci* [DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 and DYS385] to compare Angola population with other populations of Africa and to contribute to a population database for Y-STRs haplotypes. Therefore, gene diversity of each *locus*, haplotype diversity, discrimination capacity and allele frequencies in the overall population of Angola were analysed.

Bloodstains from 166 unrelated healthy donors (57 from Bakongo linguistic group [ABK], 56 from Kimbundo group [AKI] and 53 from Ovimbundo group [AOV]) were acquired after informed consent. Their ancestors had lived in Angola at least three generations.

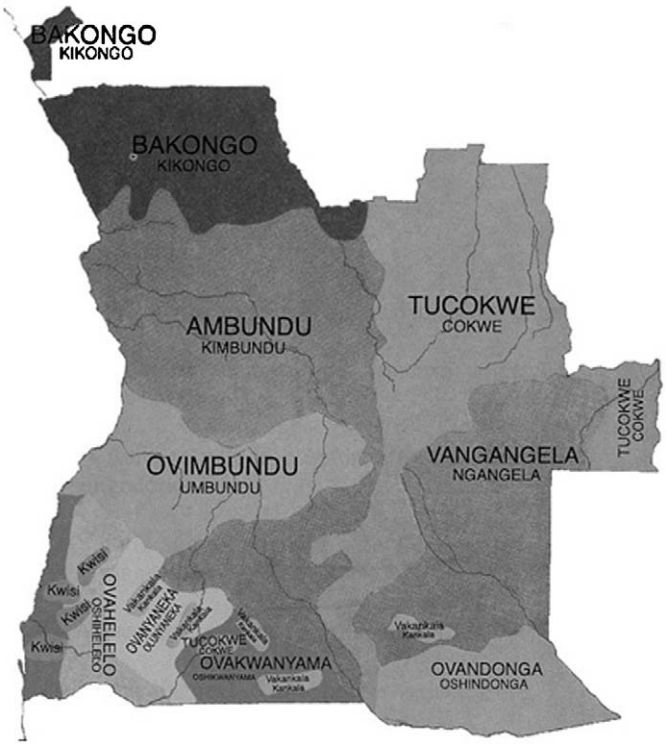


Fig. 1. Map of ethnic linguistic groups of Angola. http://www.triplov.com/letras/americo_correia_oliveira/literatura_angolana/anexo3.htm.

DNA was extracted from stains using Chelex 100 method according to Walsh et al. [1] and amplified for *PowerPlex*® Y System [DYS19, DYS389 I, DYS389 II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439 e DYS385] kit with 2700 Thermal Cycler according to manufacturer's instructions. The detection of PCR products was carried out using ABI Prism™ 310 Genetic Analyzer DNA sequencer using internal standard ILS-600 with Genescan software 3.1 [2]. The Y-STR haplotype data for the three ethnic linguistic groups of Angola population was submitted to YHRD database [3] (Accession N° YA003640 (AKI), YA003641 (ABK) and YA003642 (AOV). The genetic distance (RST) and P values were determined with program Arlequin [4] as well as the Analysis of Molecular Variance (AMOVA) test.

Technical procedures were done according to guidelines of quality control proficiency tests of the GHEP-ISFG and STRs nomenclature concerning the EDNAP group [5].In the ethnic-linguistic Bakongo group was detected a total of 53 different haplotypes of 57 individuals. A total of 49 (86%) haplotypes were unique while four haplotypes (H024, H047, H050 and H053) were observed twice. In the ethnic-linguistic Kimbundo group was detected a total of 53 different haplotypes of 56 individuals. A total of 50 (89.3%) haplotypes were unique while three 3 haplotypes (H059, H103 and H107) were observed twice. In the ethnic-linguistic Ovimbundo group was detected a total of 49 different haplotypes of 53 individuals. A total of 46 (87%) haplotypes were unique while two haplotypes (H133 and H149) were observed twice and the haplotype H118 occurred 3 times (Table 1). All ethnic groups show a discriminatory power equal or higher than 0.9245 (Table 2). Although the ethnic-linguistic Kimbundo group is closer to Ovimbundo than the Bakongo there are no significant differences ($P > 0.05$) between the three ethnic linguistic groups (Table 3). The Analysis of Molecular Variance in the three ethnic-linguistic groups of Angola shows variation within population of 98.99 and interpopulation of 1.01 suggesting a close relationship between the populations (Table 4 and Fig. 2). These results consolidate the data from the oral tradition and anthropological results according to the geographic location of each group in the context of Angola (Fig. 1).

The Angola population showed a genetic diversity in 10 *loci* ranging from 0.1478 (DYS392) to 0.7010 (DYS389 II) (Table 5). The haplotypes diversity for all 11 *loci* was 0.9969 and the overall discrimination capacity was 0.8313 (Table 2). The largest variability is found for the DYS385 *locus*. In Angola population was detected a total of 138 different haplotypes of 166 individuals and a total of 120 (72.3%) haplotypes were unique while 12 haplotypes were found twice, 2 haplotypes (H001 and H024) occurred 3 times and another 4 haplotypes (H007, H050, H053 and H103) occurred 4 times. Statistically significant differences

1872-4973/$ – see front matter © 2010 Published by Elsevier Ireland Ltd.
doi:10.1016/j.fsigen.2010.08.002

Table 1
Y-STR haplotypes found in the three ethnic linguistic groups Bakongo [ABK], Kimbundo [AKI] and Ovimbundo [AOV] of Angola ($N = 166$).

Haplotype	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H001 ABK	15	12	29	21	10	11	14	14	12	10	17–17
H002 ABK	14	12	28	26	10	11	13	14	11	13	14–21
H003 ABK	15	13	30	21	10	11	13	14	11	10	17–17
H004 ABK	16	13	30	22	10	11	14	14	11	10	17–17
H005 ABK	15	13	29	21	10	11	13	14	11	11	16–17
H006 ABK	17	13	29	21	10	11	15	14	12	11	18–18
H007 ABK	15	13	30	21	10	11	13	14	11	11	16–17
H008 ABK	17	13	30	21	10	11	14	14	11	11	17–17
H009 ABK	15	13	31	21	10	11	13	14	11	11	15–18
H010 ABK	15	13	31	21	10	11	14	14	11	11	17–18
H011 ABK	15	13	31	21	10	11	15	14	11	11	15–15
H012 ABK	16	13	29	21	10	11	15	14	11	12	16–19
H 013ABK	16	13	29	22	10	11	16	14	11	12	17–21
H 014ABK	17	13	29	21	10	11	14	14	11	12	15–17
H015 ABK	14	13	29	23	10	11	13	14	12	12	17–19
H016 ABK	14	13	29	24	10	15	13	14	12	12	11–11
H017 ABK	17	13	30	21	10	11	13	12	11	12	17–17
H018 ABK	15	13	30	21	10	11	15	13	11	12	17–17
H019 ABK	15	13	30	21	10	11	13	14	11	12	16–19
H020 ABK	15	13	30	21	10	11	14	14	11	12	15–16
H021 ABK	15	13	30	21	10	11	15	14	11	12	16–18
H022 ABK	15	13	30	21	10	11	15	14	11	12	17–17
H023 ABK	16	13	30	21	10	11	15	14	12	12	17–18
H024 ABK	15	13	31	21	10	11	13	14	11	12	15–15
H025 ABK	15	13	31	21	10	11	13	14	11	12	15–15
H026 ABK	15	13	31	21	10	11	13	14	11	12	15–20
H027ABK	16	13	31	21	10	11	13	14	11	12	15–15
H028 ABK	15	13	31	21	10	11	13	14	12	12	16–16
H029 ABK	15	13	32	21	10	12	13	14	11	12	16–17
H030 ABK	15	13	30	20	10	11	13	14	11	13	17–17
H031 ABK	15	13	30	21	10	11	14	14	11	13	17–17
H032 ABK	16	13	30	21	10	11	15	14	11	13	17–18
H033 ABK	16	13	30	21	10	11	15	14	11	13	17–19
H 034ABK	17	13	30	21	10	11	15	14	11	13	18–18
H035 ABK	16	13	31	21	10	11	15	14	11	13	17–18
H036 ABK	15	13	31	21	10	11	13	14	12	13	16–17
H037 ABK	15	14	31	21	10	11	14	14	11	11	16–18
H038 ABK	17	14	31	21	10	11	15	14	11	11	17–18
H039 ABK	17	14	31	21	10	11	15	14	11	11	17–19
H 040ABK	17	14	32	21	10	11	13	14	11	11	14–16
H041 ABK	16	14	31	21	10	11	13	14	11	12	17–18
H042 ABK	16	14	31	21	10	11	15	14	11	12	17–18
H043 ABK	17	14	31	21	10	11	14	14	11	12	16–16
H044 ABK	15	14	32	24	10	11	13	15	10	12	11–11
H045 ABK	16	14	32	24	10	11	13	15	10	12	11–11
H046 ABK	17	14	33	23	10	11	13	15	10	12	10–11
H047 ABK	15	14	32	21	10	11	13	14	11	13	16–17
H048 ABK	15	14	32	21	10	11	13	14	11	13	16–17
H049 ABK	14	12	28	25	11	13	13	14	12	12	11–14
H050 ABK	15	13	30	21	11	11	13	14	11	11	17–17
H051 ABK	15	13	30	21	11	11	13	14	11	11	17–17
H052 ABK	15	13	31	21	11	11	13	14	11	11	15–17
H053 ABK	15	13	31	21	11	11	13	14	11	11	16–17
H054 ABK	15	13	31	21	11	11	13	14	11	11	16–17
H055 ABK	14	13	29	25	11	13	13	14	12	12	11–14
H056 ABK	14	13	29	24	11	11	13	15	12	13	11–14
H057 ABK	16	13	30	21	12	11	14	13	11	13	16–17
H058 AKI	15	13	30	21	9	11	13	14	11	11	16–18
H059 AKI	15	12	29	21	10	11	14	14	12	10	17–17
H060 AKI	15	12	29	21	10	11	14	14	12	10	17–17
H061 AKI	15	13	29	21	10	11	13	14	11	11	16–17
H062 AKI	16	13	30	21	10	11	14	13	11	11	17–18
H063 AKI	14	13	30	21	10	11	14	14	11	11	15–16
H064 AKI	16	13	30	21	10	11	14	14	11	11	15–20
H065 AKI	17	13	30	21	10	11	15	14	11	11	17–18
H066 AKI	15	13	31	21	10	11	13	14	11	11	15–17
H067 AKI	15	13	31	21	10	11	13	14	11	11	15–18
H068 AKI	15	13	31	21	10	11	13	14	11	11	16–19
H069 AKI	15	13	32	21	10	11	13	14	11	11	16–17
H070 AKI	15	13	29	21	10	11	15	14	11	12	18–20
H071 AKI	16	13	29	21	10	11	16	14	11	12	16–19
H072 AKI	17	13	30	21	10	11	13	13	11	12	17–17
H073 AKI	15	13	30	21	10	11	13	14	11	12	16–17
H074 AKI	15	13	30	21	10	11	14	14	11	12	15–17
H075 AKI	15	13	30	21	10	11	14	14	11	12	16–16
H076 AKI	15	13	30	21	10	11	14	14	11	12	18–19

Please cite this article in press as: M.M. Melo, Y-STR haplotypes in three ethnic linguistic groups of Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.08.002

Table 1 (*Continued*)

Haplotype	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H077 AKI	16	13	30	21	10	11	14	14	11	12	17–18
H078 AKI	16	13	30	21	10	13	14	14	11	12	16–17
H079 AKI	17	13	30	21	10	11	13	14	11	12	17–17
H080 AKI	17	13	30	21	10	11	14	14	11	12	17–20
H081 AKI	17	13	30	21	10	11	15	14	11	12	15–20
H082 AKI	17	13	30	21	10	11	15	14	11	12	17–17
H083 AKI	15	13	31	22	10	10	11	15	9	12	12–13
H084 AKI	15	13	31	24	10	11	13	15	10	12	11–11
H085 AKI	16	13	31	24	10	11	13	15	10	12	11–11
H086 AKI	15	13	31	21	10	11	13	14	11	12	15–15
H087 AKI	16	13	31	21	10	11	13	14	11	12	15–15
H088 AKI	15	13	31	21	10	11	13	14	12	12	15–18
H089 AKI	15	13	33	21	10	12	13	14	11	12	15–17
H090 AKI	15	13	30	21	10	11	14	14	11	13	16–18
H091 AKI	16	13	31	22	10	10	11	16	9	13	12–13
H092 AKI	15	13	31	21	10	11	13	14	11	13	14–19
H093 AKI	15	13	31	21	10	11	13	14	11	13	15–18
H094 AKI	15	13	31	21	10	11	13	14	11	13	15–20
H095 AKI	15	14	31	21	10	11	13	14	11	11	16–17
H096 AKI	16	14	30	21	10	11	15	14	11	12	16–19
H097 AKI	15	14	31	20	10	11	13	14	12	12	16–16
H098 AKI	15	14	32	21	10	11	13	14	12	12	16–17
H099 AKI	17	15	32	21	10	11	14	14	11	12	16–17
H100 AKI	15	12	30	21	11	11	13	14	11	11	17–17
H101 AKI	15	13	29	21	11	11	13	14	11	11	17–17
H102 AKI	15	13	30	21	11	11	13	14	11	11	15–17
H103 AKI	15	13	30	21	11	11	13	14	11	11	16–17
H104 AKI	15	13	30	21	11	11	13	14	11	11	16–17
H105 AKI	15	13	30	21	11	11	13	14	11	11	17–17
H106 AKI	15	13	31	21	11	11	13	14	11	11	16–17
H107 AKI	15	13	31	21	11	11	13	14	11	11	17–17
H108 AKI	15	13	31	21	11	11	13	14	11	11	17–17
H109 AKI	15	13	31	21	11	11	14	14	11	11	15–17
H110 AKI	15	13	31	21	11	11	13	14	11	12	16–17
H111 AKI	14	13	29	23	11	13	13	15	12	13	11–14
H112 AKI	15	14	32	21	11	11	13	14	11	11	16–17
H113 AKI	14	14	28	24	11	13	13	15	12	13	11–13
H114 AOV	13	10	26	22	10	13	12	14	9	12	15–15
H115 AOV	14	12	29	23	10	11	13	14	11	11	14–19
H116 AOV	16	12	29	21	10	11	13	14	11	12	15–16
H117 AOV	16	13	29	21	10	11	14	14	11	11	16–18
H118 AOV	15	13	30	21	10	11	13	14	11	11	16–17
H119 AOV	15	13	30	21	10	11	13	14	11	11	16–17
H120 AOV	15	13	30	21	10	11	13	14	11	11	16–17
H121 AOV	15	13	31	21	10	11	13	14	11	11	16–16
H122 AOV	15	13	31	21	10	11	13	14	11	11	16–17
H123 AOV	16	13	30	21	10	11	15	13	11	12	16–18
H124 AOV	15	13	30	21	10	11	13	14	11	12	14–15
H125 AOV	15	13	30	21	10	11	13	14	11	12	15–16
H126 AOV	16	13	30	21	10	11	14	14	11	12	17–18
H127 AOV	17	13	30	21	10	11	14	14	11	12	17–17
H128 AOV	17	13	30	21	10	11	15	14	11	12	17–17
H129 AOV	17	13	30	21	10	11	15	14	11	12	18–18
H130 AOV	16	13	31	24	10	11	13	15	10	12	11–11
H131 AOV	15	13	31	21	10	11	13	14	11	12	16–17
H132 AOV	15	13	31	24	10	11	13	14	11	12	11–11
H133 AOV	16	13	31	21	10	11	13	14	11	12	16–18
H134 AOV	16	13	31	21	10	11	13	14	11	12	16–18
H135 AOV	17	13	31	21	10	11	14	14	11	12	17–17
H136 AOV	17	13	31	21	10	11	14	14	11	12	17–18
H137 AOV	13	13	32	24	10	11	13	14	10	13	16–16
H138 AOV	15	14	31	22	10	11	13	14	11	11	16–18
H139 AOV	16	14	31	21	10	11	15	14	11	11	16–18
H140 AOV	17	14	31	21	10	11	14	14	12	11	16–19
H141 AOV	16	14	31	23	10	11	13	15	10	12	11–11
H142 AOV	15	14	31	20	10	11	13	14	11	12	15–18
H143 AOV	15	14	31	21	10	11	14	14	11	12	16–16
H144 AOV	15	14	31	21	10	11	14	14	12	12	14–17
H145 AOV	14	12	28	25	11	11	13	14	11	11	14–21
H146 AOV	15	12	29	21	11	11	13	14	11	11	16–17
H147 AOV	14	13	29	25	11	11	13	14	11	11	14–18
H148 AOV	14	13	29	23	11	13	13	15	12	11	11–14
H149 AOV	15	13	30	21	11	11	13	14	11	11	16–17
H150 AOV	15	13	30	21	11	11	13	14	11	11	16–17
H151 AOV	15	13	30	21	11	11	13	14	11	11	17–17
H152AOV	15	13	31	21	11	11	13	14	11	11	15–17
H153 AOV	15	13	31	21	11	11	13	14	11	11	16–16

Please cite this article in press as: M.M. Melo, Y-STR haplotypes in three ethnic linguistic groups of Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.08.002

Table 1 (*Continued*)

Haplotype	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	DYS385
H154 AOV	15	13	31	21	11	11	13	14	11	11	16–17
H155 AOV	15	13	31	21	11	11	13	14	11	11	16–19
H156 AOV	15	13	31	21	11	11	14	14	11	11	15–17
H157 AOV	15	13	31	21	11	11	13	15	11	11	17–17
H158 AOV	15	13	32	21	11	11	13	14	11	11	15–18
H159 AOV	14	13	29	24	11	13	13	16	12	12	11–13
H160 AOV	15	13	31	21	11	11	13	14	9	12	15–17
H161 AOV	15	13	31	21	11	11	13	14	11	12	16–17
H162 AOV	15	13	31	21	11	11	13	14	11	12	16–18
H163 AOV	15	13	30	21	11	11	15	14	11	13	15–19
H164 AOV	15	13	31	21	11	11	13	14	11	13	16–17
H165 AOV	15	14	32	21	11	11	12	14	11	11	17–18
H166AOV	15	13	30	21	12	11	14	14	11	11	17–17

Table 2
Diversity parameters of 11 Y-STRs in three ethnic-linguistic groups of Angola.

Population	*N*	Diff Ht	*D*	*h*
Bakongo	57	53	0.9298	0.9975
Kimbundo	56	53	0.9464	0.9981
Ovimbundo	53	49	0.9245	0.9964
Total	166	138	0.8313	0.9969

Diff Ht: different haplotypes; D: discrimination capacity; h: haplotype diversity.

Table 3
Pairwise genetic distances (RST) and *P*-values in the three most important ethnic groups from Angola. The RST values are above the diagonal and *P*-values are below the diagonal.

P-values	RST		
	Bakongo	Kimbundo	Ovimbundo
Bakongo	–	0.0088	0.0201
Kimbundo	0.1696	–	−0.0001
Ovimbundo	0.0546	0.4172	–

Table 4
Analysis of molecular variance in the three ethnic-linguistic groups of Angola.

Source of variation	Sum of squares	Variance components	Percentage
Among population	15.864	0.052	1.01
Within population	825.383	5.064	98.99
Total	841.247	5.116	

Table 5
Allele/haplotype frequencies and gene diversity value for 11 Y-STRs *loci* in Angola population (166).

Allele	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplotype	DYS385
9					0.006				0.024		10–11	0.006
10					0.723	0.012			0.048	0.030	11–11	0.048
11		0.006			0.259	0.922	0.012		0.801	0.386	11–13	0.012
12		0.060			0.012	0.012	0.012	0.006	0.127	0.452	11–14	0.030
13	0.012	0.771				0.048	0.614	0.030		0.133	12–13	0.012
14	0.084	0.157					0.205	0.873			14–15	0.006
15	0.584	0.006				0.006	0.145	0.078			14–16	0.006
16	0.181						0.012	0.012			14–17	0.006
17	0.139										14–18	0.006
20				0.018							14–19	0.012
21				0.813							14–21	0.012
22				0.036							15–15	0.042
23				0.036							15–16	0.024
24				0.066							15–17	0.060
25				0.024							15–18	0.036
26			0.006	0.006							15–19	0.006
27											15–20	0.024
28			0.024								16–16	0.048
29			0.145								16–17	0.193
30			0.343								16–18	0.066

Please cite this article in press as: M.M. Melo, Y-STR haplotypes in three ethnic linguistic groups of Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.08.002

Table 5 (*Continued*)

Allele	DYS19	DYS389 I	DYS389 II	DYS390	DYS391	DYS392	DYS393	DYS437	DYS438	DYS439	Haplotype	DYS385
31			0.392								16–19	0.042
32			0.078								17–17	0.163
33			0.012								17–18	0.078
											17–19	0.018
											17–20	0.006
											17–21	0.006
											18–18	0.018
											18–19	0.006
											18–20	0.006
GD	0.5994	0.3772	0.7010	0.3307	0.4102	0.1478	0.5592	0.2298	0.3392	0.6288	–	–

GD: gene diversity.

Table 6
Pairwise genetic distances (RST) and *P*-values for Angola and other six African samples.

P-values	RST						
	Ango	Nam	Gui.-Bis	Ugan	Moza	Som	Eq. Gui.
Ango	*	−0.0064	0.0345	0.2054	0.0231	0.6346	0.0303
Nam	0.8291	*	0.0289	0.1755	0.0157	0.6237	0.0348
Gui.-Bis	0.0000	0.0040	*	0.1431	0.0254	0.6190	0.0506
Ugan	0.0000	0.0000	0.0000	*	0.1329	0.4719	0.1367
Moza	0.0025	0.0618	0.0010	0.0000	*	0.6009	0.0186
Som	0.0000	0.0000	0.0000	0.0000	0.0000	*	0.5692
Eq. Gui.	0.0004	0.0062	0.0000	0.0000	0.0114	0.0000	*

Ango: Angola present study; Nam: Namibia [6]; Gui.-Bis: Guinea-Bissau [7]; Ugan: Uganda [8]; Moza: Mozambique [9]; Som: Somalia [10]; Eq. Gui.: Equatorial Guinea [11].

($P < 0.05$) were observed between Angola population and the majority of populations analyzed, Guinea Bissau [7], Uganda [8], Mozambique [9], Somalia [10], Equatorial Guinea [11], but between Angola and Namibia [6] there was no statistically significance (RST = −0.0064; $P = 0.8291$) (Table 5). The data of this study can be applied in the reference population, to solve forensic cases and may be part of the population database (Table 6).

The authors agree and accept the guidelines for publication of population genetic papers in this journal [12].

This study was supported by Command General of police National of Angola and Forensic Genetics Service, Centre Branch, National Institute Legal of Medicine, Portugal.

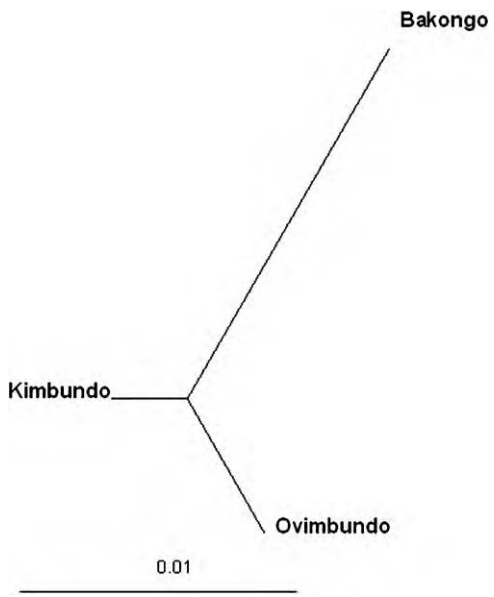


Fig. 2. Neighbour-Joining tree based on Nei genetic distances calculated between the 3 ethnic linguistic groups of Angola.

References

[1] P.S. Walsh, et al., Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, Biotechniques 10 (4) (1991) 506–513.
[2] ABI Prism Genescan® 3.1. 1989–1998 Elmer corporation.
[3] http://www.yhrd.org.
[4] S. Schneider, D. Roessli, L. Excoffier, Arlequin: A Software for Population Genetics Data Analysis, Version 2.000, Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, 2000.
[5] L. Gusmao, J.M. Butler, A. Carracedo, P. Gill, M. Kayser, W.R. Mayr, N. Morling, M. Prinz, L. Roewer, C. Tyler-smith, P.M. Schneider, DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, Forensic Sci. Int. 157 (2006) 187–197.
[6] J. Fujihara, I. Yuasa, T. Muro, R. Lida, E. Tsubota, H. Nakamura, S. Imamura, T. Yasuda, H. Takeshita, Allele frequencies and haplotypes for 28 Y-STRs in Ovambo population, Legal Med. 11 (2009) 205–208.
[7] A. Rosa, C. Ornelas, A. Brehm, R. Villems, Population data on 11 Y-chromosome STRs from Guiné-Bissau, Forensic Sci. Int. 157 (2006) 210–217.
[8] V. Gomes, C. Alves, A. Amorim, Á. Carracedo, P. Sánchez-Diz, L. Gusmão, Nilotes from Karamoja, Uganda: Haplotype data defined by 17 Y-chromosome STRs, Forensic Sci. Int. Genet. doi:10.1016/j.fsigen.2009.07.001.
[9] C. Alves, L. Gusmão, J. Barbosa, A. Amorim, Evaluating the informative power of Y-STRs: a comparative study using European and new African haplotype data, Forensic Sci. Int. 134 (2003) 126–133.
[10] C. Hallenberg, B. Simonsen, J. Sanchez, N. Morling, Y-chromosome STR haplotypes in Somalis, Forensic Sci. Int. 151 (2005) 317–321.
[11] E. Arroyo-Pardo, L. Gusmão, A.M. Lópes-Parra, C. Baeza, M. Mesa, A. Amorim, Genetic variability of 16 Y-chromosome STRs in a sample from Equatorial Guinea (Central Africa), Forensic Sci. Int. 149 (2005) 109–113.
[12] A. Carracedo, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, Publication of population data for forensic purposes, Forensic Sci. Int. Genet. April 4 (3) (2010) 145–147.

Miguel Manuel Melo[a,b*]
[a]*General Command of the Police of Angola, Department of Forensic Medicine, Angola*
[b]*Institute of Biomedical Sciences Abel Salazar, University of Oporto, Portugal*

Mónica Carvalho
Virgínia Lopes
Maria João Anjos
Armando Serra
Forensic Genetics Service, Centre Branch, National Institute of Legal Medicine, Portugal

Duarte Nuno Vieira
National Institute of Legal Medicine, Portugal, Faculty of Medicine, University of Coimbra, Portugal

Jorge Sequeiros
Institute of Biomedical Sciences Abel Salazar, University of Oporto, Portugal

Francisco Corte-Real
National Institute of Legal Medicine, Portugal, Faculty of Medicine, University of Coimbra, Portugal

*Corresponding author at: General Command of the Police of Angola, Department of Forensic Medicine, Angola
E-mail address: mig_m_melo@yahoo.com (M. M. Melo)

(25 June 2010)

Please cite this article in press as: M.M. Melo, Y-STR haplotypes in three ethnic linguistic groups of Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.08.002

G Model
FSIGEN-586; No. of Pages 5
ARTICLE IN PRESS
Forensic Science International: Genetics xxx (2010) xxx–xxx
Contents lists available at ScienceDirect
Forensic Science International: Genetics
journal homepage: www.elsevier.com/locate/fsig

Forensic Population Genetics—Letter to the Editor

Genetic study of 15 STRs loci of Identifiler system in Angola population

Dear Editor,

Angola is located in the African continent, in the area of southern Africa and has a population of approximately 14 million inhabitants. The Angola population has originated from Occidental and Southern Bantu people that came from the great lakes region, creating the most ever known African migration of our days. In their migration and formation emerged a vast array of cross-relationships [1]. The linguistic ethnic groups Bakongo, Kimbundo and Ovimbundo represent 3/4 of the population and the remaining are represented by others Bantu groups with same language, social and cultural patterns (Fig. 1).

It is intended, with this study, to characterize the genotypes of Angola population, the parametric statistic to forensic application, as well as comparison amongst other populations of other continents.

Were collected bloodstains after acquired informed consent from 480 unrelated healthy donors from North, Centre region and South of Angola, whose ancestors had lived in Angola for at least three generations.

ISFG recommendations on the analysis of the DNA polymorphisms used were strictly followed, signifying the use of recommended nomenclature and guidelines regarding quality control and statistical issues. Technical procedures were done according to guidelines of quality control proficiency tests of the EDNAP [2].

Observed and expected heterozygosity as well as unbiased estimates of Hardy–Weinberg exact *P*-values were assessed using the Markov chain method with GENEPOP [3]. Bonferroni correction leads to a significance of 0.0033. The statistical parameters evaluated were the *a priori* probability of exclusion and power of discrimination. Genetic distances between Angola and other populations were accomplished with PHYLIP [4] and the correspondent phylogenetic tree was built with TreeView [5]. *Locus* by

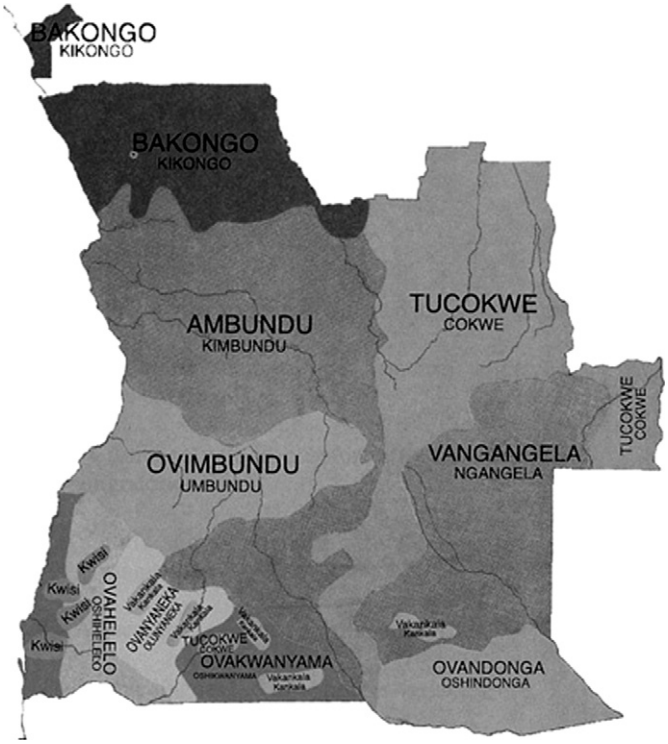


Fig. 1. Map of ethnic linguistic groups of Angola. http://www.triplov.com/letras/americo_correia_oliveira/literatura_angolana/anexo3.htm.

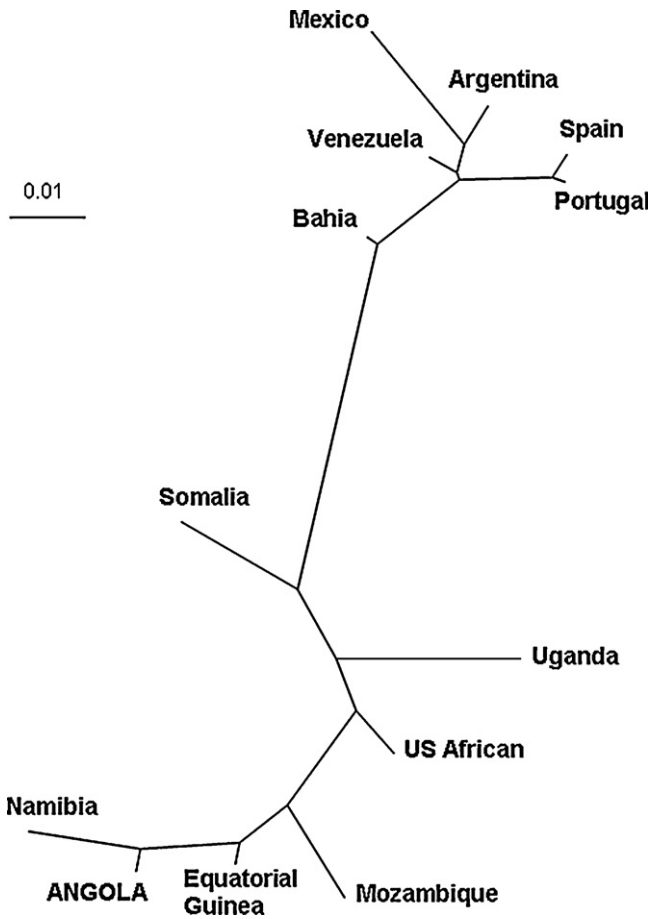


Fig. 2. Neighbour-Joining tree based on Nei genetic distances calculated between the 13 populations.

1872-4973/$ – see front matter © 2010 Published by Elsevier Ireland Ltd.
doi:10.1016/j.fsigen.2010.03.010

Please cite this article in press as: M.M. Melo, Genetic study of 15 STRs loci of Identifiler system in Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.03.010

Table 1
Allele frequencies and diversity values, HeObs (observed heterozygosity), HeExp (expected heterozygosity), HWE (Hardy–Weinberg equilibrium, *P*-values), PD (power of discrimination) and PEx (*a priori* probability of exclusion) at 15 STR *loci* in Angola ($N = 480$).

Allele	D8S1779	D21S11	D7S820	CSF1PO	D3S1358	HUMTH01	D13S317
5							
6			0.0010	0.0010		0.0656	
7			0.0052	0.0635		0.4219	
8			0.2531	0.0750		0.3229	0.0125
9			0.1385	0.0542		0.1396	0.0073
9.3						0.0417	
10	0.0021		0.3063	0.2750	0.0021	0.0083	0.0198
10.2							
11	0.0490		0.1792	0.2250			0.2938
11.2							
12	0.0990		0.0896	0.2531	0.0052		0.4594
12.2							
13	0.1948		0.0250	0.0458	0.0052		0.1615
13.2							
14	0.3979		0.0021	0.0073	0.0646		0.0427
14.2							
15	0.1927				0.2958		0.0031
15.2							
16	0.0531				0.3219		
16.1							
16.2							
17	0.0094				0.2490		
17.2							
18	0.0021				0.0542		
18.2							
19					0.0021		
19.2							
20							
20.2							
21							
21.2							
22							
23							
24		0.0010					
24.2		0.0052					
24.3		0.0010					
25		0.0010					
26		0.0010					
27		0.0771					
28		0.2406					
29		0.1656					
30		0.1833					
30.2		0.0115					
31		0.0865					
31.2		0.0490					
32		0.0177					
32.2		0.0469					
33		0.0094					
33.1		0.0010					
33.2		0.0208					
34		0.0198					
34.2		0.0010					
35		0.0469					
36		0.0104					
37		0.0021					
38		0.0010					
42.2							
43.2							
Heobs	O.7604	0.8854	0.7625	0.8083	0.7396	0.6792	0.6563
Heesp	0.7521	0.8604	0.7833	0.7958	0.7396	0.6938	0.6750
HWE	0.5797 ± 0.0168	0.4591 ± 0.0312	0.2999 ± 0.0157	0.0184 ± 0.0026	0.0543 ± 0.0043	0.4701 ± 0.0087	0.0117 ± 0.0012
PD	0.7515	0.8593	0.7822	0.7949	0.7397	0.6922	0.6742
PEx	0.5414	0.7235	0.5757	0.6008	0.5022	0.4450	0.4234

Allele	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
5	0.0031							
6	0.0010				0.0865			
7					0.0146			
8	0.0167				0.2927		0.0573	
9	0.2604		0.0031		0.2104		0.0333	
9.3								
10	0.1031		0.0135		0.0958	0.0042	0.0490	
10.2						0.0031		
11	0.3573		0.0771	0.0135	0.2781	0.0052	0.1885	
11.2			0.0021					

Please cite this article in press as: M.M. Melo, Genetic study of 15 STRs loci of Identifiler system in Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.03.010

Table 1 (*Continued*)

Allele	D16S539	D2S1338	D19S433	HUMVWA	TPOX	D18S51	D5S818	FIBRA/FGA
12	0.1260	0.0010	0.1260		0.0188	0.0260	0.3479	
12.2			0.0521					
13	0.1219	0.0010	0.2958	0.0104	0.0010	0.0365	0.2990	
13.2			0.0646			0.0010		
14	0.0094	0.0010	0.1813	0.0698		0.0625	0.0219	0.0010
14.2			0.0594			0.0031		
15	0.0010	0.0063	0.0615	0.2021	0.0021	0.1885	0.0031	
15.2			0.0417			0.0021		
16		0.0667	0.0063	0.2740		0.2094		0.0021
16.1								0.0010
16.2			0.0146			0.0010		
17		0.1073		0.2063		0.1802		0.0031
17.2			0.0010					
18		0.0458		0.1396		0.1198		0.0115
18.2								0.0146
19		0.1490		0.0615		0.0750		0.0844
19.2								0.0010
20		0.0802		0.0177		0.0458		0.0375
20.2						0.0010		
21		0.1469		0.0042		0.0219		0.0990
21.2						0.0010		
22		0.1635		0.0010		0.0094		0.1698
23		0.0583				0.0031		0.1552
24		0.0635						0.1813
24.2								0.0010
24.3								
25		0.0688						0.1188
26		0.0344						0.0344
27		0.0052						0.0458
28		0.0010						0.0083
29								0.0073
30								0.0021
30.2								0.0052
31								
31.2								0.0083
32								
32.2								0.0031
33								
33.1								
33.2								0.0010
34								
34.2								
35								
36								
37								
38								
42.2								0.0021
43.2								0.0010
Heobs	0.7729	0.8813	0.8500	0.8167	0.7646	0.8542	0.7375	0.8542
Heesp	0.7625	0.8917	0.8417	0.8125	0.7771	0.8604	0.7479	0.8792
HWE	0.1633 ± 0.0142	0.0150 ± 0.0013	0.5113 ± 0.0189	0.4963 ± 0.0146	0.0630 ± 0.0042	0.8461 ± 0.0187	0.8049 ± 0.0084	0.6654 ± 0.0302
PD	0.7628	0.8916	0.8414	0.8128	0.7755	0.8595	0.7468	0.8759
PEx	0.5514	0.7810	0.6952	0.6315	0.5644	0.7205	0.5235	0.7468

Table 2
Genetic distances for all pair of populations based on FST calculations.

	Moza	Ango	Bahia	Venez	Mexi	Arge	Soma	Ugan	E-Gui	US-Af	Spai	Port	Nami
Moza	–	0.0261	0.0837	0.1139	0.1400	0.1109	0.0783	0.0622	0.0266	0.0263	0.1352	0.1218	0.0455
Ango	0.0261	–	0.0881	0.1269	0.1547	0.1253	0.0756	0.0484	0.0101	0.0183	0.1417	0.1327	0.0123
Bahia	0.0837	0.0881	–	0.0255	0.0456	0.0257	0.0731	0.0847	0.0758	0.0499	0.0279	0.0236	0.1068
Venez	0.1139	0.1269	0.0255	–	0.0226	0.0190	0.0791	0.0974	0.1049	0.0731	0.0206	0.0196	0.1504
Mexi	0.1400	0.1547	0.0456	0.0226	–	0.0256	0.0953	0.1185	0.1339	0.0951	0.0443	0.0432	0.1735
Arge	0.1109	0.1253	0.0257	0.0190	0.0256	–	0.0890	0.1113	0.1086	0.0741	0.0235	0.0210	0.1471
Soma	0.0783	0.0756	0.0731	0.0791	0.0953	0.0890	–	0.0381	0.0659	0.0510	0.0979	0.1017	0.0938
Ugan	0.0622	0.0484	0.0847	0.0974	0.1185	0.1113	0.0381	–	0.0467	0.0317	0.1236	0.1217	0.0659
E-Gui	0.0266	0.0101	0.0758	0.1049	0.1339	0.1086	0.0659	0.0467	–	0.0163	0.1165	0.1121	0.0232
US-Af	0.0263	0.0183	0.0499	0.0731	0.0951	0.0741	0.0510	0.0317	0.0163	–	0.0869	0.0827	0.0344
Spai	0.1352	0.1417	0.0279	0.0206	0.0443	0.0235	0.0979	0.1236	0.1165	0.0869	–	0.0055	0.1584
Port	0.1218	0.1327	0.0236	0.0196	0.0432	0.0210	0.1017	0.1217	0.1121	0.0827	0.0055	–	0.1500
Nami	0.0455	0.0123	0.1068	0.1504	0.1735	0.1471	0.0938	0.0659	0.0232	0.0344	0.1584	0.1500	–

Populations: Ango: Angola [present study]; Moza: Mozambique [6]; Bahia: Bahia-Brazil [7]; Venez: Venezuela [8]; Mexi: Mexico [9]; Arge: Argentina [10]; Soma: Somalia [11]; Ugan: Uganda [12]; E-Gui: Equatorial Guinea [13]; US-Af: US African [14]; Spai: Spain [15]; Port: Portugal [16]; Nami: Namibia [17].

Please cite this article in press as: M.M. Melo, Genetic study of 15 STRs loci of Identifiler system in Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.03.010

Table 3
Locus by *locus* computation of the unbiased estimate of the exact *P*-value of the probability test between Angola and another African, European and American populations.

Loci	Angola vs. Mozambique	Angola vs. Bahia-Brazil	Angola vs. Venezuela	Angola vs. Mexico	Angola vs. Argentina	Angola vs. Somália	Angola vs. Uganda	Angola vs. Eq. Guinea	Angola vs. US African	Angola vs. Spain	Angola vs. Portugal	Angola vs. Namíbia
D8S1179	0.2520	0.0001	0.0000	0.0000	0.0000	0.0008	0.0809	0.5965	0.0296	0.0000	0.0000	0.6419
D21S11	0.0128	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.9370	0.7869	0.0000	0.0000	0.5335
D7S820	0.2386	0.0055	0.0000	0.0000	0.0000	0.0000	0.0691	0.9169	0.3525	0.0000	0.0000	0.6232
CSF1PO	0.6188	0.0019	0.0000	0.0000	0.0000	0.0000	0.0697	0.5565	0.8852	0.0000	0.0000	0.1212
D3S1358	0.3988	0.2409	0.2527	0.0000	0.0037	0.6855	0.3867	0.9210	0.8982	0.0000	0.0000	0.6728
TH01	0.0980	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4021	0.0010	0.0000	0.0000	0.7990
D13S317	0.8070	0.0000	0.0000	0.0000	0.0000	0.0000	0.3663	0.9666	0.5307	0.0000	0.0000	0.8808
D16S539	0.0193	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0131	0.1018	0.0000	0.0000	0.2979
D2S1338	0.6287	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.2198	0.5369	0.0000	0.0000	0.3447
D19S433	0.0236	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5443	0.5439	0.0000	0.0000	0.5072
VWA	0.6053	0.4262	0.0013	0.0000	0.0000	0.0000	0.0002	0.8235	0.5962	0.0895	0.0000	0.2357
TPOX	0.4572	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.2316	0.2508	0.0000	0.0000	0.4482
D18S51	0.2406	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017	0.9998	0.0469	0.0000	0.0000	0.3105
D5S818	0.0457	0.0004	0.0000	0.0000	0.0000	0.1826	0.7093	0.8890	0.7609	0.0000	0.0000	0.7761
FGA	0.0634	0.1011	0.0000	0.0149	0.0000	0.0000	0.0000	0.9404	0.0756	0.0000	0.0000	0.9143

References for the populations data used: Mozambique [6]; Bahia-Brazil [7]; Venezuela [8]; Mexico [9]; Argentina [10]; Somalia [11]; Uganda [12]; Equatorial Guinea [13]; US African [14]; Spain [15]; Portugal [16]; Namibia [17].

locus computation of the unbiased estimate of the exact *P*-value of the probability test was calculated with STRUC from GENEPOP [3].

$P < 0.05$ in the Hardy–Weinberg equilibrium analysis were observed in the Angola population for CSF1PO ($P = 0.0184$), D13S31 ($P = 0.0117$) and D2S1338 ($P = 0.0150$) (Table 1). However, if Bonferroni correction is used ($P < 0.0033$) the differences are not significant. These global data frequencies could be used as a reference in statistical calculus.

The *a priori* probability of exclusion ranged from 0.4234 (D13S317) to 0.7810 (D2S1338) (Table 1) and the combined value was 0.9999994%. The power of discrimination ranged from 0.6742 (D13S317) to 0.8916 (D2S1338) (Table 1) and the combined value was 0.99999997%.

Considering the *P*-values from *locus* by *locus* comparisons between Angola and other published population data (Tables 2 and 3) we found that the genetic distances had significant differences ($P < 0.05$) with Portugal [16], with Argentina [10] and Mexico [9] in all *loci*; with Spain [15] and Venezuela [8] in fourteen *loci*; with Somalia [11] in thirteen *loci*; with Bahia, Brazil [7] in twelve *loci*; with Uganda [12] in nine *loci*; with Mozambique [6] in four *loci*; with US African [14] in three *loci*; with Equatorial Guinea [13] in one *locus*; and with Namibia [17] did not show statistical significant differences in the 15 analysed *loci*.

This information presents great interest in forensic investigation and for the establishment of phylogenetic relation between the populations (Fig. 2). These systems are a useful tool for personal identification and can be used for routine forensic applications in the Angola population.

The first previous studies of STRs of Angola were analysed by Corte-Real et al. [18] and Beleza et al. [19]. This study represents the first genetic characterization of autosomal STRs in the most important ethnic-linguistic groups that constitute Angola population.

This study was supported by Command General of police National of Angola and Forensic Genetics Service, Centre Branch, National Institute Legal of Medicine, Portugal.

The authors agree and accept the guidelines for publication of population genetic papers in this journal [20].

References

[1] J. Neto, Etnias and Cultures of Angola, Edition of the Institute of Investigation of Angola, 1974.
[2] P. Gill, B. Brinkmann, E. d'Aloja, J. Andersen, W. Bar, A. Carracedo, B. Dupuy, B. Eriksen, M. Jangblad, V. Johnsson, A.D. Kloosterman, P. Lincoln, N. Morling, S. Rand, M. Sabatier, R. Scheithauer, P. Schneider, M.C. Vide, Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature, Forensic Sci. Int. 87 (June (3)) (1997) 185–192.
[3] M. Raymond, F. Rousset, GENEPOP (Version 3.4), 2003.
[4] J. Felsenstein (Ed.), PHYLIP (Version 3.573c), Department of Genome Sciences, University of Washington, 1995.
[5] D.M. Roderic, TreeView (v. 1.5.2), 1998 http://taxonomy.zoology.gla.ac.uk/rod/rod.html.
[6] C. Alves, L. Gusmão, A. Damascero, B. Soares, A. Amorim, Contribution for an African STR database (AmpF/STR Identifiler and Powerplex 16 System) and a reported on genotypic variations, Forensic Sci. Int. 139 (2004) 201–205.
[7] M.V. Santos, M.J. Anjos, L. Andrade, M.C. Vide, F. Corte Real, D.N. Vieira, Population genetic data for the STR loci using the AmpFISTR® Identifiler™ Kit in Bahia, Brazil, Int. Cong. Ser. 1261 (2004) 219–222.
[8] L. Bernal, L. Borjas, W. Zabala, M. Portillo, E. Fernández, W. Delgado, F. Tovar, N. Lander, M. Chiurillo, J. Ramírez, Genetic variation of 15 STR autosomal loci in the Maracaibo population from Venezuela, Forensic Sci. Int. 161 (2006) 60–63.
[9] S. Hernández-Gutiérrez, P. Hernández-Franco, S. Martínez-Tripp, M. Ramos-Kuri, H. Rangel-Villalobos, STR data for 15 loci in a population sample from the central region of Mexico, Forensic Sci. Int. 151 (2005) 97–100.
[10] W. Bozzo, M. Pena, M. Ortiz, M. Lojo, Genetic data from Powerplex® 16 system and Identifiler™ kits from Buenos Aires province (Argentina), Int. J. Legal Med. 9 (2007) 151–153.
[11] A.O. Tillmar, G. Bäckström, K. Montelius, Genetic variation of 15 autosomal STR loci in a Somali population, Forensic Sci. Int. Genet. 4 (2009) 19–20.
[12] V. Gomes, P. Sánchez-Diz, C. Alves, I. Gomes, A. Amorim, Á. Carracedo, L. Gusmão, Population data defined by 15 autosomal STR loci in Karamoja population (Uganda) using AmpFl STR Identifiler Kit, Forensic Sci. Int. Genet. 3 (2009) 55–58.

[13] C. Alves, L. Gusmão, A.M. López-Parra, M. Soledad Mesa, A. Amorim, E. Arroyo-Pardo, STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpF/STS Identifiler and Powerplex 16 kits, Forensic Sci. Int. 148 (2005) 239–242.
[14] J.M. Butler, R. Schoske, P.M. Vallone, J.W. Redman, M.C. Kline, Allele frequencies for 15 Autosomal STR loci on U.S. Caucasian, African American and Hispanic populations, J. Forensic Sci. 48 (2003) 908–911.
[15] M.V. Camacho, C. Benito, A.M. Figueiras, Allelic frequencies of the 15 STR loci included in the AmpFlSTR Identifiler™ PCR Amplification Kit in an autochthonous sample from Spain, Forensic Sci. Int. 173 (2007) 241–245.
[16] V. Lopes, A. Serra, J. Gamero, L. Sampaio, F. Balsa, et al., Allelic frequency distribution of 17 STRs from Identifiler and PowerPlex-16 in Central Portugal area and the Azores archipelago, Forensic Sci. Int. Genet. 4 (2009) 1–7.
[17] T. Muro, J. Fujihara, S. Imamura, H. Nakamura, T. Yasuda, H. Takeshita, Allele frequencies for 15 STR loci in Ovambo population using AmpFLSTR® Identifiler Kit, Int. J. Legal Med. 10 (2008) 157–159.
[18] F. Corte-Real, L. Andrade, M. Carvalho, M.J. Anjos, J. Gamero, A. Carracedo, D.N. Vieira, M.C. Vide, Comparative analysis STR data for Portuguese spoken countries, Prog. Forensic Genet. 8 (2000) 212–214.
[19] S. Beleza, C. Alves, F. Reis, A. Amorim, A. Carracedo, L. Gusmão, 17 STR data (AmpF/STR Identifiler and Powerplex 16 System) from Cabinda (Angola), Forensic Sci. Int. 141 (2004) 193–196.
[20] A. Carracedo, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, Publication of population data for forensic purposes, Forensic Sci. Int. Genet. 4 (April (3)) (2010) 145–147.

Miguel Manuel Melo[a,b,*]
[a]*General Command of the Police of Angola, Department of Forensic Medicine, Angola*
[b]*Institute of Biomedical Sciences Abel Salazar, University of Oporto, Portugal*

Mónica Carvalho
Virgínia Lopes
Maria João Anjos
Armando Serra
Forensic Genetics Service, Centre Branch, National Institute of Legal Medicine, Portugal

Duarte Nuno Vieira[a,b]
[a]*National Institute of Legal Medicine, Portugal*
[b]*Faculty of Medicine, University of Coimbra, Portugal*

Jorge Sequeiros
Institute of Biomedical Sciences Abel Salazar, University of Oporto, Portugal

Francisco Corte-Real[a,b]
[a]*National Institute of Legal Medicine, Portugal*
[b]*Faculty of Medicine, University of Coimbra, Portugal*

*Corresponding author at: General Command of the Police of Angola, Department of Forensic Medicine, Angola

E-mail address: mig_m_melo@yahoo.com (M. M. Melo)

12 February 2010

Please cite this article in press as: M.M. Melo, Genetic study of 15 STRs loci of Identifiler system in Angola population, Forensic Sci. Int. Genet. (2010), doi:10.1016/j.fsigen.2010.03.010