

Rui Miguel Moutinho Sá

PRIMATOLOGIA MOLECULAR

Contributo para a conservação de um grupo de chimpanzés (*Pan troglodytes verus*) da República da Guiné-Bissau. Diversidade mitocondrial, sexagem e aplicações forenses.

2011



UNIVERSIDADE DE COIMBRA



DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Primatologia Molecular

Contributo para a conservação de um grupo de chimpanzés (*Pan troglodytes verus*) da República da Guiné-Bissau. Diversidade mitocondrial, sexagem e aplicações forenses.

Dissertação em Biologia apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Evolução e Biologia Humanas, realizada sob a orientação científica da Professora Doutora Manuela Alvarez (Universidade de Coimbra) e da Professora Doutora Cláudia Sousa (Universidade Nova de Lisboa).

Rui Miguel Moutinho Sá

2011

DEDICATÓRIA

“Acredito que temos uma obrigação para com os nossos descendentes bem como para com os nossos antecessores: as populações de grandes símios que ainda prevalecem têm de ser protegidas. Enquanto seres humanos, temos que avançar com um novo imperativo moral de forma a assegurar a sobrevivência destes magníficos parentes, e conhece-los melhor é seguramente um bom começo.” (Richard Leakey *in* World Atlas of Great Apes and their Conservation, 2005:12).

Dedico este trabalho ao extraordinário povo da Guiné-Bissau. Dirige-se especialmente a todas as comunidades e agentes de mudança em Cantanhez. É neles que reside a capacidade de decisão última sobre a conservação destes chimpanzés.

Este trabalho é também dedicado ao falecido Régulo de Cadique, Salifo Tawel Camará, rei dos Nalús. A sua visão, poder e tolerância permitiu que os chimpanzés persistissem no seu regulado, tendo sido também um dos principais mentores para a sua conservação. Presto-lhe por isso, o meu tributo.

Foto de Capa: Chimpanzé que atravessa a estrada de Cadique, Cantanhez, 2010. Cortesia: Joana Roque de Pinho[®].

Com o apoio de:



ABSTRACT

Chimpanzees (*Pan troglodytes verus*) in Guinea-Bissau are recognized as one of the two most endangered West-African subspecies. This country represents the most western limit of the species distribution and within it they occur in the Quínara, Tombali and Gabú regions. Nonetheless, anthropogenic changes such as human population growth, unsustainable land use (for causing habitat fragmentation, degradation and loss), and poaching for pet trade are simultaneously threatening their existence and altering the landscape. Moreover, the population structure and evolutionary history of chimpanzees in the African continent continue uncertain and remarkably none study so far have incorporated individuals from this country. We investigated, for the first time, the genetic variation of chimpanzees in nine forest fragments across Cantanhez National Park (CNP) by examining 482 bp of the first domain (HVR1) of the mtDNA control region using DNA collected noninvasively. Fifty polymorphic sites, comprising 45 transitions, 2 transversions and 3 indels were observed. 16 HVR1 haplotypes were found and 42% are shared among chimpanzee communities in the south part of CNP and the remaining 58% were observed only once. Both haplotype ($h=88\%$) and nucleotide ($\pi=0.03714 \pm 0.00139$) diversities were high. The mean number of pairwise differences was 17.788 revealing varying levels of population subdivision. The phylogenetic analyses revealed the existence of two deep evolutionary lineages and the median joining network generated with this study showed five main haplogroups, yet no major groups are consistent with a defined geographical structure. The multimodal mismatch distribution observed implies that secondary contact between mtDNA lineages has occurred. Analysis of molecular variance (AMOVA) revealed genetic structure ($\Phi_{CT}=0.3918$; $p<0.001$) and the largest fraction of the variation (56.24%) is partitioned among individuals within populations ($\Phi_{ST}= 0.4376$; $p<0.05$) as expected since chimpanzee females are the dispersal sex. Implications for their conservation could be inferred and four management units were defined.

Keywords

Molecular Primatology, Chimpanzees, Guinea-Bissau, Conservation Genetics, Evolutionary History, Phylogeography

RESUMO

Os chimpanzés (*Pan troglodytes verus*) da Guiné-Bissau estão reconhecidos como os mais ameaçados das duas subespécies da África Ocidental. Este país representa o limite mais ocidental na distribuição actual da espécie e nele ocorrem nas regiões de Tombali, Quínara e Gabú. No entanto, distúrbios antropogénicos causados pelo crescimento demográfico humano, uso insustentável da terra (que causa fragmentação, degradação e perda do habitat) e caça para tráfico de animais de estimação estão simultaneamente a ameaçar a sua persistência e a alterar a paisagem. No entanto, a história evolutiva e a estrutura genética dos chimpanzés no continente africano continua incerta e notavelmente, nenhum estudo até agora incorporou indivíduos deste país. Neste estudo, investigou-se pela primeira vez, a variação genética dos chimpanzés de nove fragmentos florestais no Parque Nacional de Cantanhez (PNC) examinando 482pb do primeiro domínio (HVR1) da região de controlo do genoma mitocondrial usando ADN extraído a partir de amostras recolhidas de forma não-invasiva. Foram observados cinquenta sítios polimórficos que englobam 45 transições, 2 transversões e 3 indels. Dos 16 haplótipos encontrados, 44% são partilhados entre comunidades de chimpanzés do PNC e os restantes 56% foram observados uma única vez. A diversidade haplotípica (h) é alta (88%) e a diversidade nucleotídica é também elevada ($\pi=0.03714\pm 0.00139$). O número médio entre pares de diferenças é de 17.788 revelando existirem vários níveis de subdivisão populacional. As análises filogenéticas revelaram a existência de duas linhagens evolutivas profundas e a network gerada com o algoritmo de median-joining revelou existirem cinco haplogrupos principais, mas que não estão em conformidade com nenhuma estrutura geográfica definida. O padrão multimodal observado nos *mismatch distribution* implicam um contacto secundário entre diferentes linhagens. A análise à variância molecular (AMOVA) revelou estruturação genética ($\Phi_{CT}=0.3918$; $p<0.001$) e a maior fracção da variação (56.24%) está particionada entre indivíduos dentro das subpopulações ($\Phi_{ST}= 0.4376$; $p<0.05$) tal como esperado já que nos chimpanzés são as fêmeas o sexo que dispersa. Foi também possível inferir-se implicações para a sua

conservação e, neste sentido, definiram-se quatro unidades de gestão genética (MUs) para os chimpanzés de Cantanhez.

Palavras chave:

Primatologia Molecular, Chimpanzés, Guiné-Bissau, Genética da Conservação, História Evolutiva, Filogeografia

ÍNDICE

DEDICATÓRIA	iii
ABSTRACT.....	iv
RESUMO	vi
ÍNDICE	viii
ÍNDICE DE TABELAS.....	x
ÍNDICE DE FIGURAS.....	xi
AGRADECIMENTOS	xii
1. INTRODUÇÃO.....	13
1.1- O contributo da Primatologia Molecular: Estado da Arte	13
1.2- Os chimpanzés (<i>Pan troglodytes</i> ssp.):	16
1.3- Estudos de genética não invasiva em chimpanzés:	19
1.5- Objectivos:.....	29
2. MATERIAL E MÉTODOS	30
2.1- A área de estudo:.....	30
2.2- Amostragem não invasiva:.....	34
2.3- Extracção de ADN:.....	35
2.4- Amplificação e sequenciação de ADN:	35
2.5- Autenticidade das sequências:	36
2.6- Análise estatística dos dados:.....	37
3. RESULTADOS.....	39
3.1- Análise do sucesso de extracção de DNA e da variação genética da região hipervariável 1 (HVR1):.....	39
3.2- Análise da diversidade mitocondrial HVR1:	42
3.3- Análise filogenética:.....	43
3.4- Análise da Network filogenética:	50
3.5- Análise da história demográfica:.....	53
3.5.1- Testes de Neutralidade:	53
3.5.2- Mismatch distributions:.....	54
3.6- Análise da estrutura genética:	59
3.6.1- Análise da distância genética por F_{ST} :	59

3.6.2- Análise da distância genética pelo método F_{ST} linearizado de Slatkin:	62
3.6.3- Análise relativa da migração e do fluxo genético:	65
3.6.4-Análise da variância molecular (AMOVA):.....	67
3.6.4- Análise espacial da variância molecular (SAMOVA):	69
3.7- Análise de isolamento por distância:	71
3.8- Análise de barreiras genéticas:	73
3.9- Análise com medidas complementares de genética da conservação:	75
3.10.- Análise da divergência temporal:.....	79
3.11.- Sexagem molecular e análise do rácio sexual:	81
4. DISCUSSÃO	82
5. CONCLUSÕES	92
Implicações para a conservação dos chimpanzés de Cantanhez	93
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	95
ANEXO 1: Lista de sequencias dos haplótipos HVR1 dos chimpanzés de Cantanhez a serem submetidas ao GenBank, em formato FASTA.....	115

ÍNDICE DE TABELAS

2.1. Matos de Cantanhez e respectiva superfície.....	32
3.1. Posições polimórficas das sequências de ADNmt HVR1 dos chimpanzés de Cantanhez	40
3.2. Distribuição e frequência dos haplótipos observados nos chimpanzés de Cantanhez	41
3.3. Sumário de índices de diversidade molecular	43
3.4. Sumário dos testes de neutralidade.....	54
3.5. Distâncias genéticas F_{ST} entre pares de fragmentos de floresta para os chimpanzés de Cantanhez.....	61
3.6. Distâncias genéticas pelo modelo F_{ST} linearizado de Slatkin entre pares de fragmentos de floresta para os chimpanzés Cantanhez	64
3.7. Estimativa relativa de migração absoluta (M) entre pares de fragmentos de floresta para os chimpanzés de Cantanhez.....	66
3.8. Resultados da análise da variância molecular (AMOVA).....	69
3.9. Resultados da análise espacial da variância molecular (SAMOVA)	70
3.10 Sumário das medidas complementares de genética de conservação.....	77

ÍNDICE DE FIGURAS

1.1. Mapa de distribuição de chimpanzés e bonobos em África.....	18
1.2. Representação do genoma mitocondrial humano.....	27
2.1. Área de estudo no Parque Nacional de Cantanhez, Guiné-Bissau	30
2.2. Distribuição das principais áreas florestais em Cantanhez	33
3.1. Cladograma concatenado das relações filogenéticas entre haplótipos dos chimpanzés de Cantanhez.....	46
3.2a. Cladograma de Máxima Parcimónia.....	47
3.2b. Filograma dos haplótipos HVR1 dos chimpanzés de Cantanhez pelo método de Neighbor-Joining	48
3.2c. Filograma dos haplótipos HVR1 dos chimpanzés de Cantanhez por Inferência Bayesiana	49
3.3. Median-Joining Network dos haplótipos HVR1 dos chimpanzés de Cantanhez	51
3.4. Mapa de distribuição dos haplótipos HVR1 dos chimpanzés de Cantanhez por fragmento de floresta	52
3.5. Mismatch distributions dos chimpanzés de Cantanhez e respectivas estatísticas associadas	58
3.6. Regressão linear de isolamento por distância.....	72
3.7. Mapa das discontinuidades (barreiras) genéticas existente nas subpopulações de chimpanzés de Cantanhez.....	74
3.8. Contribuição (CT) para a diversidade haplotípica total de cada uma das subpopulações de chimpanzés de Cantanhez.....	77
3.9. Contribuição (CRT) para a riqueza haplotípica total de cada uma das subpopulações de chimpanzés de Cantanhez.....	78
3.10. Árvore de credibilidade de clades máximas para os haplótipos HVR1 dos chimpanzés de Cantanhez em milhões de anos (MYA).....	80

AGRADECIMENTOS

Um agradecimento muito especial à minha família pelo apoio e amor incondicional, sempre lá.

Um abraço imenso a todos os meus amigos espalhados pelo Mundo.

Aos meus orientadores, à Professora Doutora Cláudia Sousa da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, pela demonstração de força interior, motivação e amizade; à Professora Doutora Manuela Alvarez da Faculdade de Ciências e Tecnologia da Universidade de Coimbra pelas sugestões, comentários e longos debates sempre tão construtivos. Ao Departamento de Antropologia da Universidade de Coimbra na pessoa da Professora Doutora Eugénia Cunha pela motivação e tenacidade de que vale a pena compreendermos “Quem somos”. À Professora Doutora Catarina Casanova do Instituto Superior de Ciências Sociais e Políticas da Universidade Técnica de Lisboa. Ao Professor Mike Bruford da Universidade de Cardiff por todos os conselhos tão preciosos. Aos meus colegas e amigos de campo e a todo o Projecto Dári: Joana Sousa, André Barata, Susana Costa, Alexandra Cardoso, Ana Costa, Marta Carmo, Fernando Sousa e Joana Roque de Pinho pelos tempos passados na Guiné. Aos meus colegas da 5ª Edição do Mestrado em Evolução Humana e a todos os meus amigos no G10!

Ao Carlos Vaz Marques pela “telefonía sem fios”.

Ao Paulo Leitão por tudo!

Um abraço emotivo para a Maria Joana Silva e Tânia Minhós: “nô sta djunto!”

Um especial agradecimento vai para todos os guineenses que me ajudaram a tornar este trabalho possível: ao Zeca Dju de Lautchande, ao Djibi Indjai e Pabe Injai de Caiquene, ao Iaiá Camará de Cadique Nalu, ao Saidu Kuiaté de Iemberém, ao falecido Régulo de Cadique Nalu, ao chefe de Tabanca de Caiquene, ao sr. Abubakar Serra da ONG AD (Acção para o Desenvolvimento), ao professor Joãozinho da Costa de Cadique Nalu. À Cristima Schwarz e ao Engenheiro Alfredo Simão do IBAP .

Um agradecimento muito muito especial para os Dáris de Cantanhez!

A todos o meu sincero obrigado!

1. INTRODUÇÃO

1.1- O contributo da Primatologia Molecular: Estado da Arte

Quando em 1987, a cientista forense Cecília von Beroldingen demonstrou que uma raiz de um cabelo humano continha ADN suficiente para ser analisado, estaria longe de pensar que seria pioneira num novo campo emergente: a genotipagem não invasiva (Higuchi et al. 1988; von Beroldingen et al. 1987; Woodruff 2004a). Este novo método permitiu recolher ADN de pequenas amostras biológicas (pêlos, sangue, tecido, fezes, etc.) colmatando assim o problema de manuseamento dos animais que muitas das vezes envolvia: armadilhar, caçar, sangrar ou realizar biópsias (Woodruff 2004a). A genotipagem não invasiva usufruiu também das técnicas desenvolvidas no seio da Genética Molecular que por sua vez possibilitaram a amplificação, visualização e sequenciação da molécula de ADN (Goossens et al. 2003).

Vários estudos filogeográficos e populacionais têm vindo a basear-se em técnicas não invasivas como por exemplo: em lincos ibéricos (*Lynx pardinus*) (Pires e Fernandes 2003; Palomares et al. 2002); em ursídeos (Wasser et al. 1997); em aves (*Otis tarda*) (Idaghdour, Broderick, e Korrida 2003); em leões-marinhos (*Eumetopias jubatus*) (Deagle et al. 2005), mustelídeos (Riddle et al. 2003; Gómez-Molliner et al. 2004; Hedmark et al. 2004), raposas (*Vulpes vulpes*) (Dalén, Gotherstrom, e Angerbjorn 2004), elefantes africanos (*Loxodonta cyclotis*) (Eggert, Eggert, e Woodruff 2003), etc.

Os primeiros estudos moleculares com primatas não humanos só surgiram em meados da década de 1990 e segundo Goossens et al. (2003) este facto é resultante de que ao contrário, de outras ordens animais as populações de primatas têm sistemas sociais complexos, gerações bastante longas e, cuja obtenção de amostras genéticas antes do advento da genotipagem não invasiva era logisticamente e eticamente difícil. Hoje em dia esses estudos com primatas não humanos podem ser agregados naquilo que se designa por *Primatologia Molecular*.

Sensu lato, por Primatologia Molecular pode entender-se a aplicação das técnicas provenientes da Genética Molecular à Primatologia (Di Fiori e Gagneux 2007).

Esta designação, muito recente, surgiu com a necessidade de agrupar uma série de estudos científicos desagregados, que recorreram às técnicas moleculares. A sua designação foi proposta por Morris Goodman (Wildman et al. 2003). Mais recentemente, Anthony Di Fiori organizou uma conferência internacional (Molecular Primatology: Progress and Promise, 2006), sob os auspícios do Center for the Study of Human Origins da Universidade de Nova Iorque, e recolheu o apoio e suporte da comunidade científica para esta nova designação.

Apesar de só nos últimos anos estes estudos se terem intensificado existem alguns trabalhos que indiciam o aparecimento da Primatologia Molecular. Nuttal (1904), na sua monografia serológica: *Blood Immunity and Blood Relationship*, sugeriu que os seres humanos estão intimamente relacionados com os chimpanzés e gorilas, já que para além de possuírem os mesmos grupos sanguíneos (e.g. ABO), produzem também os mesmos antigénios.

Na década de 1960 surgiu a ideia de que seria possível reconstruir a história evolutiva dos organismos a partir das suas proteínas e ácidos nucleicos. Esta possibilidade começou a ser aplicada ao estudo dos primatas não humanos (Buettner-Janush et al. 1961; Goodman 1962; Zuckerkandl 1963). Já nos anos de 1970, surgem os primeiros estudos que fizeram uso de sequências de aminoácidos aplicadas ao estudo de filogenias de primatas não humanos e, mais concretamente, à sua socioecologia (Duggleby 1977).

É no entanto, nos anos de 1980 e 1990 que se iniciam os primeiros estudos intensivos de sequências de ADN em primatas não humanos, com o desenvolvimento de novas técnicas moleculares. De acordo com Di Fiori (Di Fiori 2003; Di Fiori e Gagneux 2007) as principais áreas científicas de estudo da Primatologia Molecular podem classificar-se da seguinte forma:

Ecologia Molecular: uso das técnicas moleculares que permitem examinar o parentesco, padrões de dispersão, organização e filogeografia [*Vide* a título

exemplificativo: (Idani 1991; Melnick, Hoelzer, e Honeycutt 1992; Morin et al. 1994a; Avise e Hamrick 1996; White 1996; Avise 1995; Roseblum, Supriatna, e Melnick 1997; Goldberg e Ruvolo 1997; Goldberg e Wrangham 1997; Avise 1998; Gerloff et al. 1999; Avise 2000; Shimada 2000; Avise 2004; Beebee e Rowe 2004).

Genética da Conservação: a aplicação das técnicas moleculares à biologia da conservação de primatas. A grande maioria das espécies de primatas não humanos habita em regiões tropicais que estão sob ameaça devido a vários factores endógenos e/ou exógenos. Por esse motivo é necessário alertar toda a 'Comunidade Global' de que por exemplo existem menos de 200 000 chimpanzés em todo o Mundo! É neste sentido que a genética da conservação se torna muito relevante já que por exemplo ajuda a caracterizar a diversidade genética das espécies ameaçadas, a perceber como estas estão geograficamente estruturadas, permitir estimar o seu tamanho efectivo populacional, ou ainda, informar sobre o acasalamento e reprodução de colónias em cativeiro. Veja-se mais detalhadamente: (Loeschcke, Tomiuk, e Jain 1994; Frankham, Ballou, e Briscoe 2005; Avise e Hamrick 1996; Smith e Wayne 1996) e mais recentemente um novo jornal: *Conservation Genetics* que fornece uma excelente perspectiva sobre as aplicações moleculares à conservação.

Diversidade Molecular e Adaptação: o significado adaptativo dos padrões da diversidade molecular *intra* e *inter* taxa de primatas não humanos. Também aqui, as técnicas moleculares oferecem um método alternativo que permite estimar directamente o grau de parentesco e afinidade entre indivíduos, baseando-se na partilha de marcadores genéticos (Pamilo e Crozier 1982; Queller e Goodnight 1989; Lynch e Ritland 1999).

Genética do Comportamento: o uso de suportes genéticos para perceber a variação comportamental inter-indivíduos ou seja, tentar perceber qual o fundamento genético para que determinado indivíduo tenha comportamentos mais agonísticos que outros, ou porque é que dentro da mesmo sexo simpátrico há indivíduos que migram mais tarde que outros, ou ainda, porque é que alguns indivíduos demonstram uma maior capacidade cognitiva que outros (Fairbanks 2001; Plomin et al. 2001; Lesch et al. 1996; Bennet et al. 2002).

Genómica Comparativa: análise e comparação dos genomas de diferentes espécies. Este termo foi cunhado por Thomas Roderick (1986) e, no ano seguinte em 1987, o director do projecto OMIM (*Online Mendelian Inheritance in Man*), Victor McKusick lançou o reputado jornal científico, *Genomics* (Di Fiori e Gagneux 2007). É também neste âmbito que se inclui a sequenciação do genoma do chimpanzé e a consequente comparação com o ser humano (The Chimpanzee Sequencing and Analysis Consortium 2005). Para uma revisão mais detalhada sobre os principais contributos realizados no âmbito da Primatologia Molecular veja-se: Di Fiori e Gagneux (2007).

O contexto deste estudo enquadra-se por isso, no domínio científico da Primatologia Molecular e pretende investigar a seguinte problemática: Que história evolutiva tiveram os chimpanzés de Cantanhez na República da Guiné Bissau?

1.2- Os chimpanzés (*Pan troglodytes ssp.*):

Os chimpanzés são os nossos parentes mais próximos em termos evolutivos, e há 7-4 milhões de anos partilharam connosco o mesmo ancestral comum (Goodman 1999). Apesar de termos um cariótipo diferente, partilhamos os mesmos grupos sanguíneos e temos pelo menos 98% de ADN semelhante, apesar dessa semelhança não se reflectir ao nível proteico (The Chimpanzee Sequencing and Analysis Consortium 2005).

Em termos de conservação, os chimpanzés ocidentais (*Pan troglodytes verus* Schartz, 1939) estão mais ameaçados do que qualquer outra subespécie de chimpanzé no continente africano (Kormos e Boesch 2003). Estima-se que existam actualmente menos de 38 000 chimpanzés ocidentais (Butynski 2003b) e em particular, entre 600>1000 indivíduos na República da Guiné Bissau, segundo a estimativa de Gippoliti et al. (2003) tendo mesmo sido dados como extintos nesse país (Lee, Thornback, e Bennett 1988).

A taxonomia mais recente, e seguida neste estudo, foi proposta por Groves (2001) e Grubb (Grubb et al. 2003). Esta taxonomia é também a utilizada pelo IUCN-SSC Primate Specialist Group (<http://www.primatesg.org>). Reconhece a existência de

duas espécies no género *Pan*: bonobos (*Pan paniscus*) e chimpanzés comuns (*Pan troglodytes*). Por sua vez os chimpanzés estão categorizados taxonomicamente em quatro subespécies: o chimpanzé da África Oriental (*Pan troglodytes schweinfurthii*), o chimpanzé da África Central (*Pan troglodytes troglodytes*), o chimpanzé da África Ocidental (*Pan troglodytes verus*) e o chimpanzé da Nigéria-Camarões (*Pan troglodytes ellioti*, anteriormente designado por *P. t. vellerosus*) (Groves 2001). O critério utilizado para a designação destas subespécies baseou-se essencialmente em evidências morfológicas e genéticas (Groves 2001).

Nesta taxonomia, os humanos anatomicamente modernos (*Homo sapiens*), os seus antecessores (*Homo* spp.; *Australopithecus* spp.; *Ardipithecus* spp.; *Paranthropus* spp.; *Kenyanthropus* sp.; *Orrorin* sp.; *Sahelanthropus* sp.), chimpanzés (*Pan* spp.) e gorilas (*Gorilla* spp.) fazem parte da família Hominidae.

A taxonomia dos chimpanzés não tem sido consensual entre a comunidade científica. Desde a classificação de Linnaeus (1758) e, posteriormente, com os postulados de Darwin (1871), os chimpanzés têm sido colocados em diversos grupos taxonómicos e reflectem a discussão em torno do conceito de espécie.

Os chimpanzés comuns estão distribuídos pelo continente africano desde o Senegal, até ao sul do rio Congo até à zona leste do Lago Tanganyika, Tanzânia (Stumpf 2007) (**Fig. 1.1**).

O chimpanzé da África Oriental (*P. troglodytes schweinfurthii*) tem uma distribuição que vai desde o norte da República Centro Africana e sudoeste do Sudão, a este com o rio Ubangi e a sul na linha do Equador na República Centro Africana. Nesta área existem muitos locais de estudo de longo termo com registos de pelo menos duas décadas, como por exemplo: Gombe, Mahale e Budongo (Inskipp 2005).

O chimpanzé da África Central (*P. t. troglodytes*) tem uma distribuição que vai desde o sul dos Camarões ao sul do rio Sanaga (Mitani 1990; Prescott, Rapley et al. 1994). Estende-se desde o este até à parte ocidental da República Centro Africana, onde se confina ao extremo sudoeste no Parque Dzanga-Ndoki (Inskipp 2005). Está também largamente distribuído pelo Gabão (Blom e Feistner 1992).

Os chimpanzés da Nigéria-Camarões (*P. t. ellioti*, nomenclatura anterior: *P. t. vellerosus*) estão distribuídos no sudeste da Nigéria em pequenas e fragmentadas

populações (e.g. em Oba Hills Forest Reserve) a sul, sudeste e este do delta do rio Níger, ao longo da fronteira com os Camarões em Gashaka Gumti (Oates, Gadsby, e Jenkins 2003).

Os chimpanzés da África Ocidental (*P. t. verus*) ocorrem a sudoeste no Mount Assirik no Senegal, sudoeste do Mali e a sudeste nas florestas de Cantanhez na Guiné-Bissau (Gippoliti, Embalo, and Sousa 2003). Distribui-se também pela República da Guiné, Serra Leoa, Libéria, Costa do Marfim, bem como a sudoeste no Gana (Inskipp 2005). Não se confirma a sua presença no Burkina-Faso (Brownell 2003b). O chimpanzé ocidental encontra-se extinto na Gâmbia, Togo e Benin (Brownell 2003a).

Segundo as estimativas de Butynski (2003a) a área geográfica que esta subespécie ocupa é de 631 000 km².

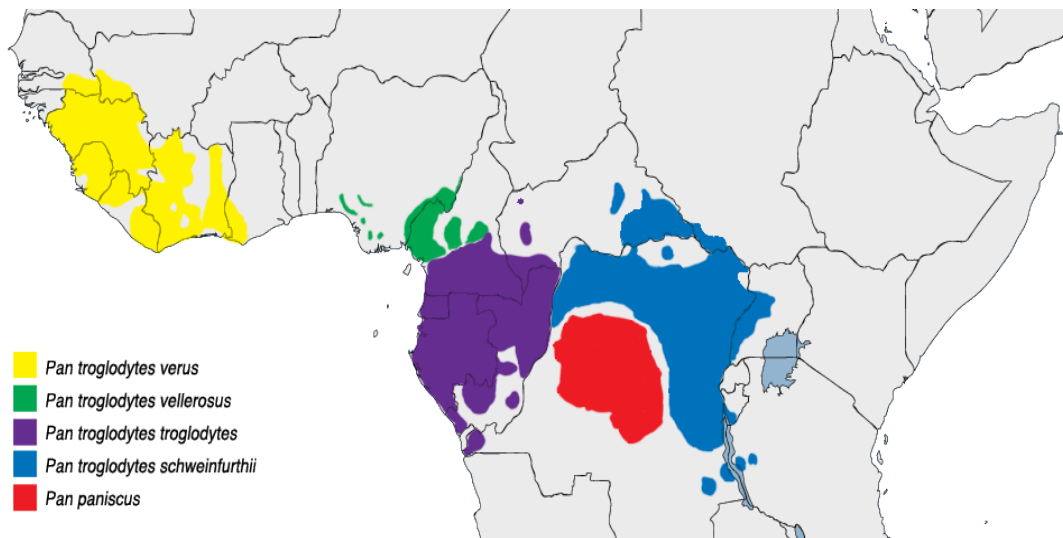


Figura 1.1: Distribuição de chimpanzés e bonobos em África. Fonte: WWF 2009.

Os chimpanzés vivem em grupos multi-macho/multi-fêmea e variam substancialmente no seu tamanho demográfico de população para população. O núcleo do grupo social encontra-se nos machos que, regra geral, são aparentados entre si (em inglês: *male-bonded society*) (Goodall 1986). Estruturam-se em comunidades de mistura-separação ou seja, uma comunidade de chimpanzés pode dividir-se em partes ou sub-unidades em função da disponibilidade de recursos (Goodall 1986). Entendo-se aqui *comunidade* como um grupo de indivíduos que partilham um território comum (Newton-Fisher 2004). No caso dos chimpanzés, a

comunidade, para além de viver num determinado espaço geográfico, demonstra também uma relação de partilha, i.e. é partilhada informação sobre a localização de recursos, alguns alimentos são partilhados entre indivíduos como o caso da carne e frutos, existem elementos culturais que todos os indivíduos partilham e há uma partilha de interações diárias entre indivíduos, nos aspectos da sua vida social e sexual (Reynolds 2005).

A composição das partes pode variar entre um a vários indivíduos ao longo do dia e, ser liderada por machos ou fêmeas. Uma comunidade de chimpanzés pode englobar duas a oito dezenas de indivíduos (Casanova 2006) e necessitar de um *home range*, i.e. parte do território que é habitualmente usada pelos efectivos comunitários (Newton-Fisher 2003) que, pode variar entre os 6 e os 20km² tendo como termos de referência os chimpanzés da floresta de Taï e de Budongo respectivamente (Boesch e Boesch-Achermann 2000; Newton-Fisher 2002). A maioria das fêmeas deixa o grupo original na adolescência, transferindo-se para outro grupo adjacente (Nishida et al. 2003; Boesch e Boesch-Achermann 2000) apesar de, haver populações em que esse facto ainda não foi verificado, como em Bossou por exemplo (Sugiyama 1999), sendo geralmente os machos, o sexo filopárico (Nishida 1979).

1.3- Estudos de genética não invasiva em chimpanzés:

Os estudos de chimpanzés iniciaram-se na década de 1960 e desde então tornaram-se modelos referenciais para a compreensão da Evolução e Biologia Humanas (Sayers e Lovejoy 2008).

Morin e Woodruff (1992) foram pioneiros em utilizar ADN não invasivo de pêlos provenientes de comunidades de chimpanzés selvagens. Os seus resultados sugeriram que existem 3 grandes linhagens mitocondriais de chimpanzés e que a sua história evolutiva deverá ser repensada (Morin e Woodruff 1992; Morin et al. 1993; Woodruff 1993; Morin et al. 1994a; Morin et al. 1994b; Morin et al. 2001).

Em 1993, Takenaka e colegas aplicaram com sucesso testes de paternidade em cativeiro, recorrendo também a amostras não invasivas e abrindo assim caminho, não

só para futuros estudos de chimpanzés em liberdade, mas também, para uma melhor compreensão das suas dinâmicas sociais (Takenaka et al. 1993).

Em 1994, baseando-se em sequências de ADN mitocondrial, Ruvolo e colegas realizaram um estudo sobre a filogenia dos hominóides e concluíram que: i) o género *Homo* e o género *Pan* formam uma clade separada do gorila; ii) a variabilidade intra-populacional dos chimpanzés é superior à observada entre espécies e por conseguinte, poderá ter efeitos importantes na gestão genética em cativeiro e; iii) a pouca diversidade mitocondrial dos povos humanos pode ser posta em perspectiva quando comparada com os outros hominóides (Ruvolo et al. 1994).

Por forma a maximizar a variabilidade genética de uma colónia de chimpanzés no Zoo de Taronga, Houlden e colegas encetaram uma análise de maternidade e paternidade dos vários indivíduos que constituíam essa colónia e comprovaram que, o macho alfa tinha efectivamente, o sucesso reprodutivo mais elevado, confirmando uma correlação entre o estatuto social na hierarquia do grupo e o seu sucesso reprodutivo. Sugeriram também, que os microsátélites são marcadores moleculares muito adequados que, permitem distinguir a origem e proveniência dos indivíduos e, podem ser considerados como um marcador de eleição para medidas de gestão e manutenção genética em cativeiro (Houlden, Woodworth, e Humphrey 1997).

Goldberg e Ruvolo, também em 1997, realizaram um estudo abrangente sobre a filogeografia e sociogenética dos chimpanzés da África Oriental demonstrando que estes não tinham um padrão clássico de subdivisão filogeográfica, possuindo uma variabilidade mitocondrial baixa e pouco agrupamento geográfico, semelhante aos estudos realizados em populações humanas. Em termos biogeográficos verificaram que os rios não impediram o seu fluxo de genes já que não foram observadas descontinuidades nas árvores filogenéticas, confirmando assim, que os chimpanzés orientais têm uma grande mobilidade, sendo capazes de se movimentar entre espaços abertos e fragmentos de floresta (Goldberg e Ruvolo 1997, Goldberg e Wrangham 1997).

A partir de uma análise filogeográfica a região HVR1 do ADNmt, Gonder e colegas descobriram a existência de uma quarta subespécie de chimpanzés na região da Nigéria e Camarões que apelidaram de *Pan troglodytes vellerosus*, relançando

assim não só o debate em torno da diversidade intraespecífica dos chimpanzés, mas também do conceito de subespécie (Gonder et al. 1997). Posteriormente, a nomenclatura desta subespécie foi revista passando a ser designada por *P. t. ellioti* (Oates, Groves e Jenkins 2009).

Gagneux e colegas em 1997, referiram as dificuldades técnicas inerentes aos métodos não invasivos, principalmente no que respeita aos falsos alelos (i.e. quando um alelo é incorrectamente genotipado devido a artefactos decorrentes da PCR ou electroforese ou devido a erros humanos na sua leitura e registo). Advertiram também que se as circunstâncias permitirem é preferível recolherem-se pêlos com raiz em detrimento de pêlos sem raiz para obtenção de uma boa qualidade de ADN (Gagneux 1998; Gagneux, Boesch, e Woodruff 1997).

Em 1999, Gagneux e colegas num estudo que reflecte a história evolutiva dos hominóides africanos apresentaram uma datação para a divergência entre humanos e chimpanzés situada nos 4,5 milhões de anos (MYA) tendo recorrido a um relógio molecular e usado sequências mitocondriais (Gagneux et al. 1999). Do mesmo modo, ao compararem sequências filogenéticas de ADNmt, verificaram que existem subclades profundas tanto em gorilas como em chimpanzés, e que as clades dos chimpanzés mostraram mais ramificações, sugerindo que estes tiveram histórias evolutivas diferentes dos humanos (Gagneux et al. 1999).

Um outro estudo, de âmbito sociogenético (Gagneux, Boesch, e Woodruff 1999), revelou que as fêmeas chimpanzés da floresta de Tai procuram furtivamente machos não pertencentes à sua comunidade, aumentando deste modo, o fluxo de genes dos machos inter comunidades, apesar destes serem filopátricos. Esta estratégia reprodutiva das fêmeas, permite-lhes escolher um maior número de machos, sem perderem os recursos que dispunham bem como usufruir da protecção que lhes era conferida pelos companheiros na comunidade original (Gagneux, Boesch, e Woodruff 1999).

No mesmo ano (1999), Taberlet e colegas publicaram um artigo sobre as vantagens e desvantagens da aplicação dos métodos não invasivos, dando assim, um grande contributo teórico através do seu modelo sobre a fiabilidade dos dados nos estudos que utilizam uma genotipagem não-invasiva, recomendado que se replicasse

a genotipagem entre três (heterozigóticos) a sete vezes (homozigóticos) no sentido de se confirmarem os genótipos minimizando a taxa de erro (Taberlet, Waits, e Luikart 1999).

Em 2000, Katherine Gonder apresentou um estudo filogeográfico extensivo sobre os chimpanzés da Nigéria e Camarões explorando extensivamente a sua estrutura genética (Gonder 2000). Surgem novas melhorias técnicas na utilização de microsátélites (Goossens et al. 2000). Meier e colegas também confirmaram a relevância da aplicação de microsátélites nas caracterizações genéticas de populações de chimpanzés em cativeiro todavia, verificaram que não há correlação entre a posição dos machos na hierarquia social e o sucesso reprodutivo a curto prazo contradizendo o estudo realizado por Houlden e colegas (Meier, Hemelrijk, e Martin 2000). Um ano mais tarde, Linda Vigilant publicou um artigo sobre a paternidade e o parentesco em comunidades selvagens (Vigilant 2001) indo ao encontro do estudo realizado por Constable e colegas sobre a paternidade dos chimpanzés de Gombe (Constable et al. 2001). Um importante estudo filogeográfico surgiu também neste ano reunindo quatro investigadores (e.g. Tony Goldberg, Katherine Gonder, Pascal Gagneux e Phillipe Morin) que mais contributos deram para a compreensão da história evolutiva dos chimpanzés (Gagneux et al. 2001).

Outra das contribuições mais relevantes na área da história evolutiva dos chimpanzés foi a publicação do estudo de Stone et al (2002). Baseando-se na região não codificante do cromossoma Y concluíram que os chimpanzés e os bonobos têm seis a sete vezes mais variação que os seres humanos e que o seu tamanho efectivo da população (N_e) era grande ou relativamente constante quando comparado com o dos humanos (Stone et al. 2002). Veja-se também o artigo de revisão de Gagneux (2002).

No entanto, Yu e colegas, relançaram a controvérsia em torno da datação do ancestral comum mais recente entre chimpanzés e humanos ao afirmarem que existe uma baixa diversidade nucleotídica em chimpanzés e bonobos (Yu et al. 2003). Apontaram para uma data mais tardia de separação entre chimpanzés e bonobos (1,8 MYA) dando como explicação a formação do rio Congo (Yu et al. 2003). Não é por isso de estranhar que no ano seguinte tenham aparecido estudos que apresentaram

resultados alternativos a esse estudo, como é o caso de Fisher e colegas que sugeriram que a história evolutiva dos chimpanzés pode ter sido bem mais complexa do que se pensava, já que se por um lado se assiste a uma diminuição do número de populações selvagens de chimpanzés, apesar destes continuarem a apresentar uma diversidade genética alta. No entanto, indicam também que este facto se pode dever a uma redução repentina da população seguida por uma recuperação rápida (Fischer et al. 2004). Atribuem ainda, uma divergência entre chimpanzés centrais e ocidentais por volta dos 430 000 a 650 000 anos (Fischer et al. 2004). McGrew e colegas, no mesmo ano, decidiram comprovar, na comunidade de chimpanzés de Mont Assirik, no Senegal que, as técnicas não invasivas podem efectivamente dar um grande contributo para o conhecimento socioecológico das comunidades de chimpanzés principalmente quando estas não estão habituadas (McGrew et al. 2004). Ainda no mesmo ano foi publicado um outro estudo filogeográfico sobre a diversidade mitocondrial dos chimpanzés de Bossou e dos Montes Nimba, onde os autores verificaram que há pouco fluxo de genes entre estas comunidades de chimpanzés na Guiné (Shimada et al. 2004).

O ano de 2005 terá sido certamente o: 'ano Internacional do chimpanzé' já que se atingira o objectivo da sequenciação do seu genoma por um consórcio internacional, permitindo deste modo, uma efectiva 'genómica comparada' (The Chimpanzee Sequencing and Analysis Consortium 2005) comprovado pelo vasto número de artigos que a seguir se publicaram (Veja-se o volume 437 da revista *Nature*).

Um relevante artigo de opinião sobre os desafios que se colocam às investigações de genética de populações de chimpanzés numa era 'pós genómica' foi apresentado por Stone e Verrelli (2006). Segundo estes autores, a maior parte dos estudos de genética evolutiva têm-se centrado nos humanos, no entanto, com a sequenciação dos genomas de outros grandes símios poderá agora levar a investigação sobre a unicidade específica de cada um para um outro patamar (Stone e Verrelli 2006). Neste mesmo ano, Gonder e colegas dão mais um contributo para a compreensão da biogeografia dos chimpanzés da Nigéria-Camarões, reforçando cada vez mais a existência desta quarta subespécie (Gonder e Disotell 2006; Gonder,

Disotell, e Oates 2006). Fischer e colegas, por outro lado, apontaram o modelo de isolamento por distância como possível explicação para as complexas histórias demográficas e evolutivas dos chimpanzés e discordaram da noção de que as subspécies sejam entidades geneticamente distintas (Fischer et al. 2006). Veja-se também o artigo de Mitani sobre as influências demográficas no comportamento dos chimpanzés (Mitani 2006).

Mais recentemente, Becquet e colegas, partindo da análise de 310 microsátélites autosomáticos provenientes de 78 chimpanzés (41 ocidentais, 16 centrais, 7 orientais, 3 híbridos e 11 desconhecidos) e 6 bonobos, chegaram aos seguintes resultados: A alta taxa de mutação dos microsátélites permite obter mais informação dos eventos históricos/ demográficos por *locus*. Em termos de análise de clusters, recorreram ao programa STRUCTURE que demonstrou a existência de 4 grandes clusters (i.e. populações descontínuas). Este resultado, conjuntamente, com o mapa de haplótipos do cromossoma Y e ADNmt confirmaram que há poucas evidências para a existência de uma quarta subespécie. No que concerne à análise principal de componentes (PCA) também demonstraram a existência de estruturação genética que se verificou através da significância estatística: uma subdivisão clara da população de chimpanzés em 4 subpopulações descontínuas agrupando-se em 4 eigenvectores muito significativos. Encetaram ainda um teste para possíveis populações adicionais, no entanto admitiram que a amostra é reduzida, que havia pouca variabilidade nas amostras de *P. t. ellioti* e que era necessário um levantamento mais sistemático. Em termos de evidência de *consanguinidade* encontraram pouca heterozigotia, principalmente nos chimpanzés ocidentais e centrais (Becquet et al. 2007). Em termos de diferenciação de frequência alélica (i.e. diversidade alélica) observaram que a diferença entre bonobos e chimpanzés ocidentais é maior do que entre chimpanzés centrais ou orientais. Este facto pode dever-se a efeitos da deriva genética via *bottleneck* dos chimpanzés ocidentais. Indicaram também, que a primeira espécie a divergir, terão sido os *P. t. verus*, apontando como possível consequência um tamanho efectivo populacional reduzido (Becquet et al. 2007) .

As conclusões principais do estudo de Becquet et al. (2007) foram:

- 1- As designações das 3 diferentes subespécies de chimpanzés correspondem a clusters de indivíduos com frequências alélicas semelhantes que podem ser definidas a partir dos dados genéticos sem ter em consideração o 'rótulo' geográfico;
- 2- Existe pouca evidência de mistura entre as diferentes subespécies e que as populações de chimpanzés orientais e centrais estão mais relacionadas entre si no tempo e no espaço do que a população de chimpanzés ocidentais/centrais;
- 3- A análise principal de componentes (PCA) deixa antever um *cluster* único na população ocidental. No entanto, reconheceram que, serão necessários mais dados, mais amostras, outros marcadores genéticos e informações sobre a sua origem geográfica para se perceber com clareza esse padrão;
- 4- Não encontraram suporte para a existência de uma quarta população de chimpanzés comuns na Nigéria-Camarões (*P.t.elliotti*);
- 5- Apesar de terem tentado incluir o maior número possível de amostras do continente africano, a amostragem geográfica não foi de todo aleatória;
- 6- O facto deste estudo ter incluído maioritariamente chimpanzés centrais e orientais, pode dar a impressão de descontinuidade. Estudos futuros que incluam amostras de chimpanzés de outras populações africanas podem em princípio, identificar populações intermédias e demonstrar outros padrões de variação (Becquet et al. 2007).

Já em 2008, Caswell e colegas, apresentam novas datações para as divergências temporais entre chimpanzés e bonobos na ordem dos 1.29 Mya; 0.51 Mya entre chimpanzés ocidentais e centrais e de pelo menos 0.50 Mya entre os chimpanzés orientais e centrais (Caswell et al. 2008). Estimaram ainda que o tamanho efectivo da população de chimpanzés centrais ronda os 100 000 indivíduos, mais do que se pensava, e que a formação do rio Congo, mais uma vez, poderá estar na origem da divergência entre chimpanzés e bonobos, já que a datação é consistente com essa hipótese (Caswell et al. 2008). Outro estudo recente sugere que o processo de especiação de chimpanzés e humanos foi complexo, podendo ter existido fluxo de genes entre ambas espécies após um período inicial de separação (Patterson et al.

2006). A história evolutiva dos chimpanzés permanece por isso, envolta de controvérsia.

1.4- O ADN mitocondrial como marcador molecular:

Os marcadores moleculares aportam uma “impressão digital” única de cada indivíduo (Parker et al. 1998) permitindo entre muitas aplicabilidades: i)-esclarecer sobre a sua estrutura populacional e divergência no espaço e no tempo (Awise 1998, Awise 2000, Awise 2004) e/ou; ii)- informar sobre as suas densidades e tamanhos efectivos da população (Beebee e Rowe 2004; Kohn e Wayne 1997).

As técnicas moleculares podem ter uma miríade de aplicações, mas são especialmente úteis quando se pretende elaborar um plano de manutenção de diversidade genética ou esclarecer sobre alguns aspectos comportamentais, demográficos, evolutivos e ecológicos das espécies ameaçadas, bem como proceder a análises forenses em casos de venda de carne proveniente de caça ilegal por exemplo (Haig 1998). Cada marcador genético tem as suas vantagens e desvantagens e é específico para os objectivos de pesquisa (Coote e Bruford 1996). Marcadores que têm uma taxa de evolução mais rápida têm uma maior aplicabilidade ao nível do indivíduo, comunidades e populações, enquanto que, outros marcadores que têm uma evolução mais lenta são úteis para estudos filogenéticos (Frankham et al. 2005). Neste estudo a região hipervariável I (HVR1) da região de controlo (também conhecida por D-loop) do ADN mitocondrial foi utilizada como marcador molecular.

As mitocôndrias são um organelo fundamental de cada unidade celular. Estão, presentes com milhares de cópias em todos os organismos eucarióticos e proporcionam à célula a energia necessária para esta realizar as suas funcionalidades básicas. Poderão ter-se aliado à célula eucariótica através de um processo evolutivo de endossimbiose (Margulis 1992). Têm também a particularidade de ter um genoma próprio, com poucos genes funcionais, que em quase toda a classe dos mamíferos se transmite por via matrilinear, ou seja, por via germinal feminina. O seu ADN, regra geral, não está também sujeito ao processo de recombinação (Awise 1998, Awise 2000).

A análise ao ADNmt é um dos métodos mais directos de estudar o ADN, ainda que não cromossomático ou nuclear, particularmente no que respeita à compreensão de alguns aspectos da variação genética em algumas populações de primatas e vertebrados, e fornece informação directa sobre os padrões de afinidade materna ou sobre a hierarquia de uma população “dominada” por um determinado sexo (Di Fiori e Gagneux 2007). Fornece ainda, informações sobre efeitos fundadores, hibridação e introgressão (Avice 2000; Avice 2004). A título ilustrativo veja-se a **Fig. 1.2** que representa o genoma mitocondrial humano, com a sua típica forma circular.

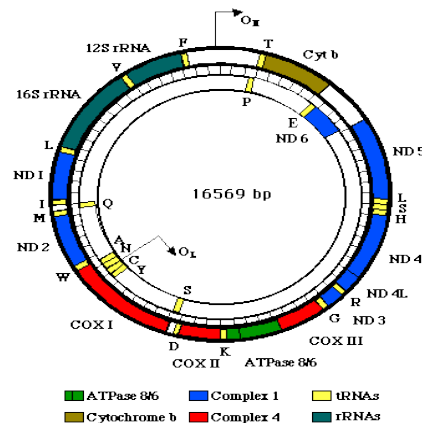


Figura 1.2: Representação do genoma mitocondrial de *Homo sapiens*. Fonte: GenBank: <http://www.ncbi.nlm.nih.gov>.

A heteroplasmia (i.e. ocorrência de dois ou mais tipos de variação molecular codominante dentro da mesma população de ADNmt, no mesmo indivíduo) é apontada como uma desvantagem (Grzybowski 2000). Por outro lado, a variação do ADNmt caracteriza-se também por um alto nível de homoplasias (i.e. um conflito filogenético/ genealógico) (Galtier et al. 2006). Já houve também alegações de transmissão por via paterna [*paternal leakage*] (Awadalla et al. 1999; Eyre-Walker et al. 1999; Hagelberg et al. 1999) mas foram posteriormente rejeitadas (Hagelberg et al. 2000; Galtier et al. 2009). Descreveram-se ainda, ocorrências de selecção positiva em primatas causando aquilo que se denomina por *selective seeps* (i.e. redução ou eliminação da variação entre nucleótidos como resultado de uma selecção natural positiva) (Grossman et al 2004). Os *selective sweeps* ocorrem geralmente numa região do genoma onde o haplótipo seleccionado é basicamente o único que existe

na população, e provoca uma redução substancial da diversidade genética total nessa região do genoma (Frankham et al. 2005).

Uma fonte importante e genuína de problemas na análise evolutiva do ADNmt prende-se com a presença de cópias translocadas de ADNmt no genoma nuclear (Bensasson et al. 2001). Depois de se transporem no núcleo, estas inserções nucleares ou '*numts*' [i.e. '*Nuclear mtDNA segments*'] evoluem de forma independente como cópias parálogas do segmento original de ADNmt (Thalmann et al. 2004). Por vezes são apelidados de *pseudogenes* porque estes fragmentos, apesar das suas homologias (i.e. semelhanças) com outras partes do genoma mitocondrial não são transcritos em proteínas funcionais (Woischnik e Moraes 2002). Estão distribuídos aleatoriamente pelos cromossomas nucleares e variam em tamanho (Tourmen et al. 2002). Os *numts* evoluem mais lentamente que o verdadeiro segmento de ADNmt e podem ser percepcionados como um registo instantâneo do ADNmt na altura em que foi transferido para o núcleo (Woischnik e Moraes 2002). Thalmann et al. (2004) aconselham que uma das chaves na análise evolutiva, para quem recorre ao ADNmt como marcador molecular, passa por determinar se há *numts* que estão inadvertidamente incluídos nas sequências em estudo.

O D-loop representa a única região não codificante do ADNmt e este domínio é composto geralmente por 1kb e contém duas regiões hipervariáveis: a HVR1 e HVR2. A função destas regiões está envolvida na iniciação e regulação da replicação e transcrição do ADNmt (Moritz et al. 1987). Estas regiões não codificantes têm uma alta taxa de substituição de nucleótidos quando comparadas com as restantes regiões codificantes mitocondriais e nucleares (Morovvati et al. 2007). Considera-se por esta razão que as regiões hipervariáveis do genoma mitocondrial são "hotspots mutacionais" (Stoneking 2000). Isto faz com que estas regiões sejam adequadas para estudos intra populacionais já que a elevada taxa de mutação significa também que as sequências polimórficas, (i.e. diferentes haplótipos) sejam mais prováveis em indivíduos não relacionados (Morovvati et al. 2007). Este estudo faz uso da HVR1, que é considerada mais polimórfica que a HVR2 (Lutz et al 1998). Morovvati et al. (2007) demonstram que a HVR1 é mais discriminante que a HVR2 em humanos devido à

presença de um elevado número de diferenças nucleotídicas. Este marcador foi por isso usado para reconstruir a história evolutiva dos chimpanzés de Cantanhez.

1.5- Objectivos:

A principal missão deste estudo é contribuir para a conservação genética dos chimpanzés de Cantanhez. A questão central é investigar a sua história evolutiva e teve como objectivos:

- 1- Reconstruir a sua filogenia intraespecífica;
- 2- Analisar e interpretar a sua diversidade e diferenciação genéticas, compando-a com outros estudos;
- 3- Avaliar a estrutura genética desta população;
- 4- Reflectir sobre a sua história demográfica e analisar a sua divergência num contexto filogeográfico;
- 5- Determinar o seu rácio sexual;
- 6- Aportar recomendações para a sua conservação.

2. MATERIAL E MÉTODOS

2.1- A área de estudo:

O trabalho de campo decorreu de Março a Abril de 2007 e de Outubro a Dezembro de 2008 nas Florestas de Cantanhez, classificadas em 2008, em Parque Nacional de Cantanhez (PNC) (**Fig. 2.1**). Esta zona é uma das principais áreas com mais biodiversidade da República da Guiné-Bissau. Situa-se no sul do país na região administrativa de Tombali, e faz fronteira com a Guiné-Conacry a Oriente. A superfície proposta é de 1 067,67 km² (IBAP 2008).

O clima de Cantanhez, é predominantemente tropical húmido com duas estações: a estação seca, de Novembro a Maio e, a estação das chuvas, de Junho a Outubro, rondando a temperatura média, entre os 28 e 31°C, com uma precipitação que varia entre os 2000 e os 2500 mm anuais, o que justifica o facto dessa área ser a mais húmida do país; testemunhando a existência de uma floresta sub-húmida e de um número elevado de espécies vegetais (IBAP 2008).

A população humana na zona de Cantanhez ronda os 20 000 habitantes com uma densidade populacional de 17,1 hab/Km² (IBAP 2008). A população humana tem uma composição multiétnica: Nalús (considerados maioritários e tradicionalmente donos da terra) Tandas, Balantas, Djacancas, Fulas e Sossos (IBAP 2008).

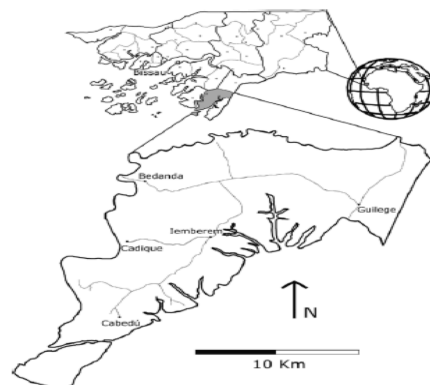


Figura 2.1: Área de estudo no Parque Nacional de Cantanhez, sul da Guiné-Bissau. Fonte: Temudo (2009).

A agricultura continua a ser a actividade fundamental com o predomínio da cultura de arroz de bolanha em solos de mangais (bolanha de água salgada ou de tarrafe); *mpampam* (i.e. culturas de sequeiro que envolvem geralmente o corte e queima dos matos para preparação das terras) nas florestas; seguido pela fruticultura diversificada, sobretudo de bananas e citrinos. A agricultura e a recolha de produtos florestais ou marinhos são a base da economia de subsistência dos residentes.

O Instituto da Biodiversidade e das Áreas Protegidas (IBAP) aponta como principais objectivos do Parque Nacional de Cantanhez: a) a preservação, conservação e defesa das manchas de floresta sub-húmida detentoras de uma grande biodiversidade, consideradas as últimas manchas de floresta primária do país e, que acolhem uma grande diversidade faunística; b) a salvaguarda das espécies animais e vegetais raras ou mesmo em vias de extinção, caso concreto dos elefantes (*Loxodonta africana*) que migram frequentemente para o outro lado da fronteira (Guiné), incluindo também outras espécies; c) a promoção do ecoturismo e; d) a valorização das actividades económicas como forma de melhoria de condições de vida das populações residentes (IBAP 2008).

Simão (1997) subdividiu a área de Cantanhez em 14 matos (**Tabela 2.1 e Fig. 2.2**) todavia, Cassamá (2006) afirma existirem 16. Os matos de Amidara, Cambéque e de Canamina são os maiores de toda a área. Alguns matos continuam bem preservados graças às restrições socioculturais e, por vezes, místicas que imperam sobre eles e que são respeitados pela população (i.e. matos com locais sagrados). Estes são utilizados como lugares de cerimónias e fanados (i.e. circuncisão). Pelo contrário, noutros matos como Capicada ou Catombói, os vestígios de degradação são mais do que evidentes devido ao uso constante a que estão sujeitos por parte das populações mais voltadas para a agricultura de planalto (Nalús e Djacancas). O mato de Flaque desapareceu devido a uma grande desflorestação (IBAP 2008).

A gestão dos 14 matos está sob a coordenação de um Comité de cogestão que reagrupa os comités de gestão criados a nível de cada mato. Cada mato tem aproximadamente 5 a 8 membros. Os membros são residentes e têm que se responsabilizar por uma boa gestão e funcionalidade da sua área. São responsáveis

directos pela fiscalização e pelo seguimento das diferentes formas de utilização dos recursos naturais.

Tabela 2.1: Matos de Cantanhez e respectiva superfície, segundo Simão (1997).

Mato	Superfície (ha)
Amidara	2.507,5
Cambeque	1.285,5
Canamina	1.067,5
Cibe de Cadique	552,5
Cafatche	457,5
Caghode	442,5
Catomboi	437,5
Lautchande	405,0
Madina	402,5
Caiquéne	362,5
Cabum	232,5
Muna	102,5
Cungha	75,0
Capicada	47,5

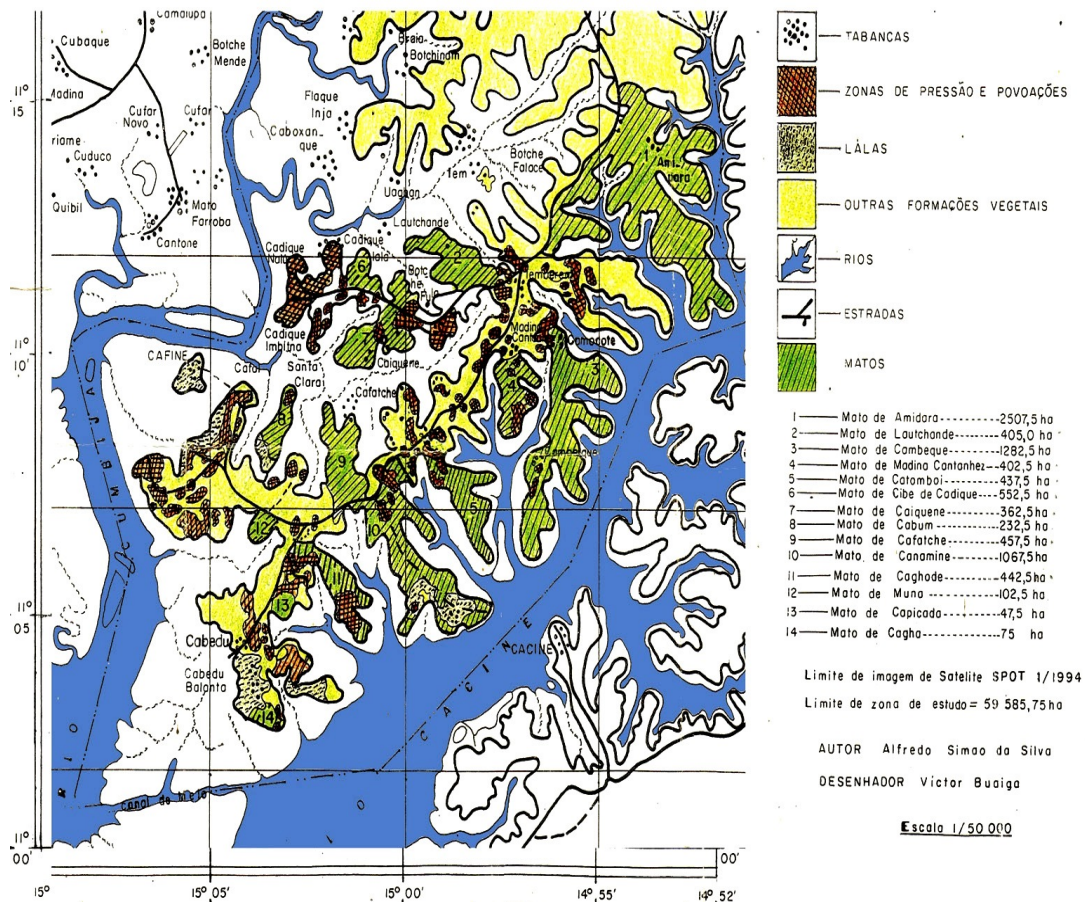


Figura 2.2: Distribuição das principais áreas florestais em Cantanhez, segundo Simão (1997).

Em termos faunísticos verifica-se a presença de várias espécies de mamíferos como o búfalo (*Syncerus caffer*), o boca branco (*Hippotragus equinus*), o porco do mato preto (*Phacochoerus aethiopicus africanus*), porco do mato vermelho (*Potamochoerus porcus*), entre outros. Inclui ainda seis espécies de primatas, dos quais se destacam: o cólobo vermelho ocidental (*Ptilocolobus badius temminckii*); o cólobo branco e preto da África Ocidental (*Colobus polykomos*); o macaco mona (*Cercopithecus mona campbelli*), o macaco vervet (*Cercopithecus aethiops sabaeus*); o babuíno da Guiné (*Papio papio*) e o chimpanzé ocidental (*Pan troglodytes verus*) (Casanova and Sousa 2007).

Sousa, Gippoliti et al. (2005) estimam que possam existir entre 600 a 1000 chimpanzés em toda a Guiné-Bissau. Não se sabe ao certo quantos chimpanzés existem em Cantanhez, no entanto um estudo realizado por Sousa (2007) e Sousa et

al. (2011) baseado na estimativa de densidade de ninhos em 4 dos principais matos de Cantanhez revelou uma densidade de 1,937 e 2,340 indivíduos que constroem ninhos por Km², quando se tem em consideração a área total desses matos (Lautchande/ Caiquéne) a mesma autora indica que poderão existir entre 33 a 40 indivíduos (Sousa 2007; Sousa et al. 2011).

2.2- Amostragem não invasiva:

As amostras não invasivas foram recolhidas em nove dos 14 matos de Cantanhez e são provenientes de chimpanzés selvagens (*Pan troglodytes verus*). A amostragem foi oportunística e envolveu o seguimento dos grupos de chimpanzés nessas áreas maximizando a maior distribuição possível nesta área do parque. As amostras foram recolhidas frescas, durante o período de manhã após os chimpanzés terem saído dos ninhos construídos na noite anterior e tendo o cuidado de guardar um intervalo de aproximadamente 1.5m para evitar a recolha de amostras do mesmo indivíduo. A parte superficial das fezes foi preferida para a recolha porque, contém a maior parte de células epiteliais. Para cada amostra recolhida, foi-lhe atribuída um código individual, um código da região, lugar, data, hora, coordenada geográfica, e algumas medidas morfométricas. Recolheram-se aproximadamente 5g de material fecal de cada amostra, e armazenaram-se à temperatura ambiente em etanol absoluto (100%) durante 24h sendo posteriormente transferidas para um outro tubo individual, contendo sílica gel (Tipo 3, Sigma Aldrich®), seguindo o protocolo de Roeder et al. (2004) e mantidas também à temperatura ambiente até o seu ADN ser extraído. Todas as precauções para evitar contaminação humana foram tomadas durante o seu manuseamento. As respectivas licenças para o transporte das amostras foram obtidas junto das autoridades nacionais guineenses e portuguesas.

2.3- Extração de ADN:

O ADN foi extraído de todas as amostras recolhidas seguindo o protocolo de Gerloff et al. (1995). Usou-se o kit de extração: QIAamp Stool mini kit (QIAGEN) com as seguintes modificações: ~2g de material fecal foi inserido num tubo de 2ml e adicionou-se 1.4ml de buffer ASL. Vortexou-se e deixou-se a misturar numa placa orbital durante a noite. Todos os tempos de centrifugação foram aumentados para 3 minutos.

Após o passo da incubação de duas horas no buffer AL, adicionou-se 1350 µg/ml de Poly-A ARN mensageiro por forma a promover uma melhor ligação do ADN antes do passo da precipitação com o etanol. Utilizou-se uma eluição de 150µl com o buffer AE em vez dos 200µl recomendados pelo fabricante, no sentido de assegurar uma boa concentração de ADN. Dois controlos negativos foram utilizados nas extrações e vários durante o processo de amplificação para monitorizar possíveis contaminações. Todo o material utilizado foi previamente esterilizado por autoclave e/ou raios UV.

2.4- Amplificação e sequenciação de ADN:

O ADN extraído das amostras foi amplificado por reacção em cadeia da polimerase (PCR). Os primers L15926 (5'- TAC ACT GGT CTT GTA AAC C-3') que corresponde às posições 15326- 15344 do genoma mitocondrial completo do chimpanzé [EMBL/GenBank nº de Acesso: D38113] e H16555 (5'- TGA TCC ATC GTG ATG TCT TA-3') que corresponde às posições 15971- 15990 também do D38113 foram usados, porque amplificam com sucesso a HVR1 em chimpanzés. Segundo Shimada et al. (2004), estes primers amplificam ~600pb do fragmento da região de controlo mitocondrial. Todas as reacções de PCR foram realizadas num termociclador Gene Amp® PCR System 9700 (Applied Biosystems®) em tubos de 200µl num volume de reacção total de 20µl. A reacção de PCR consistiu em: 9µl de master mix QIAGEN Multiplex kit (que contém: buffer, DNTPs, a taq polimerase, e 3 mM MgCl₂), 0.1µl 1x BSA; 0.1µl 1x Q-solution, 6.8µl de H₂O ultrapura; 0.2µM de cada primer e 2µl de ADN.

O perfil de PCR otimizado foi o seguinte: incubação e activação inicial de 95º durante 15 minutos seguindo-se de 45 ciclos de 94º durante 30 segundos, 51º durante um minuto e 72º durante um minuto e finalmente uma extensão final de 72º durante 10 minutos.

Para a sexagem molecular usaram-se os primers AMEL A (5'- FAM- CCT GGG CT CT GTA AAG AAT AGT G-3') e AMEL B (5'- ATC AGA GCT TAA ACT GGG AAG CTG-3'), porque amplificam de forma eficaz o gene da amelogenina nos chimpanzés e grandes símios em geral, seguindo o perfil e condições de PCR reportados por Bradley et al. (2001).

As reacções de PCR foram visualizadas através de electroforese num gel de agarose a 2% ao qual foi adicionado 1.5µl de brometo de etídio e as bandas presentes foram comparadas com a intensidade de referência da "ladder". As amostras que demonstraram ter menos 5ng/µl não foram sequenciadas porque apresentam geralmente uma má qualidade.

Os produtos PCR foram purificados com as enzimas Exonuclease I (EXO) e Shrimp Alkaline Phosphatase (SAP) (New England Biolabs, UK) a 37º durante 60 minutos e desactivadas a 80º durante 15 minutos. Os produtos PCR purificados foram enviados para sequenciar na companhia Macrogen Europe, em ambas as direcções usando os mesmos primers HVR1.

2.5- Autenticidade das sequências:

Os cromatogramas de todas as sequências foram inspeccionados e as sequências editadas, reunidas e alinhadas usando o programa SEQUENCHER v.5.0 (Gene Codes Corporation®). Todas as sequências de consenso foram submetidas a pesquisa pelo BLAST (Basic Local Alignment Search Tool- <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) usando um algoritmo de procura por nucleótido que encontra regiões de semelhança local entre sequências e compara-as indicando a percentagem de semelhança entre elas e um *e-value*. Esta técnica permitiu não só confirmar a autenticidade de chimpanzé das sequências de consenso, mas também que estas não eram "numts".

No sentido de assegurar que efectivamente as sequências deste estudo não estavam contaminadas por “numts” todos os produtos PCR foram cuidadosamente inspeccionados tanto nos géis de agarose (para presença de bandas duplas no tamanho esperado do fragmento) e, para picos duplos nas autoradiografias de sequenciação. Para além disso encontrou-se um grande nível de variação no dataset e todas as sequências alinharam em conformidade com a sequência de referência HVR1 do genoma mitocondrial de *Pan troglodytes troglodytes* [Nº de Acesso: X93335] (Arnasson et al. 1996).

Complementarmente as sequências produzidas por este estudo foram também comparadas com as sequências de numts de chimpanzé depositadas na base de dados MITOMAP (<http://www.mitomap.org/MITOMAP>). Não se encontraram evidências de presença de “numts” no dataset deste estudo. Uma sequência HVR1 de bonobo (*Pan paniscus*) (Nº de Acesso: GU189675) (Zsurka et al. 2010) foi incluída no dataset como *outgroup* para as análises filogenéticas.

2.6- Análise estatística dos dados:

Os haplótipos foram definidos analisando as sequências com o programa DNASP v. 4.01 (Rozas et al. 2003). As diversidades haplotípicas (h) nucleotídica (π) e de Watterson (θ_w) foram estimadas no programa ARLEQUIN v.3.01 (Excoffier et al. 2005) bem como as estimas de F_{ST} na análise da sua diferenciação genética. A riqueza haplotípica (H_r) foi calculada usando o método de rarefacção implementado no software CONTRIB v.1.02 (Petit et al. 1998).

As análises de variância molecular (AMOVA) usando os pares de diferenças foram também realizadas no programa ARLEQUIN e o padrão de subdivisão populacional mais provável foi estimado recorrendo à análise espacial da variância molecular no programa SAMOVA v. 1.0 (Dupanloup et al. 2002).

Aplicou-se um teste de Mantel e uma RMA (regressão linear por redução do eixo principal) entre a distância genética e a distância geográfica com o objectivo de avaliar um possível padrão de isolamento por distância no programa IBDWS v. 3.16 (Jensen et al. 2005). No mesmo sentido, para avaliar se existem barreiras vicariantes

que possam impedir o fluxo de genes usou-se o programa BARRIER v. 2.2 (Manni et al. 2004).

A reconstrução da network filogenética foi feita no programa NETWORK v. 4.6 que usa o algoritmo de median-joining (Bandelt et al. 1999).

Para as análises filogenéticas definiu-se o melhor modelo de evolução e de substituição de nucleótidos através do critério de informação Akaike implementado no programa jMODELTEST v.0.1.1 (Posada 2008). Utilizaram-se os métodos de neighbor-joining e de máxima parcimónia implementados no software MEGA v.5 (Tamura et al. 2011) bem como a inferência bayesiana implementada no programa MRBAYES v. 3.1 (Ronquist e Huelsenbeck 2003) para a reconstrução filogenética. A convergência das cadeias Markov Monte Carlo da inferência bayesiana foi avaliada no programa TRACER v.1.5 (Rambaut e Drummond 2007). Todas as árvores obtidas foram visualizadas e editadas no software FIGTree v.1.3.1 (Rambaut 2006) e concatenadas por todos os métodos numa árvore de consenso no programa TREEGraph v. 2.0 (Stöver e Müller 2010).

Usou-se um relógio molecular relaxado e uma abordagem coalescente implementada no pack do programa BEAST v. 1.6.2 (Drummond e Rambaut 2007) com o objectivo de analisar a sua divergência temporal. Realizaram-se três corridas independentes e usou-se: a mesma taxa de substituição, os mesmos parâmetros otimizados e o modelo sugerido pelo jMODELTEST. A convergência das corridas foi também avaliada no programa TRACER após os logs e árvores terem sido combinados no LogCombiner e TreeAnnotator respectivamente (Drummond e Rambaut 2007).

Por fim, a análise da história demográfica e as hipóteses de expansão súbita e espacial foram testadas usando: *mismatch distributions* (Rogers e Harpending 1992) e testes de neutralidade (Tajima D e Fu Fs) no programa ARLEQUIN e; os testes Fu e Li F* e D* no programa DNASP v.4.01.

3. RESULTADOS

3.1- Análise do sucesso de extracção de DNA e da variação genética da região hipervariável 1 (HVR1):

Determinou-se com sucesso o sexo e os haplótipos da região mitocondrial HVR1 para 99 das 110 amostras não invasivas recolhidas. A taxa de sucesso de extracção de DNA foi de 90%. De entre estas sequências, 47 demonstraram ter o mesmo sexo e o mesmo haplótipo. Por forma a assegurar que estas sequências não pertenciam ao mesmo indivíduo utilizaram-se 3 *loci* de microssatélites altamente polimórficos, inicialmente descritos para humanos mas comumente usados em estudos de chimpanzés (i.e. D1S550, D2S1326 e D16S1624). Trinta e quatro amostras, provaram ter o mesmo genótipo e foram conseqüentemente removidas do dataset porque representam amostras que foram recolhidas em replicado dos mesmos indivíduos, reduzindo assim para 65 o número total de indivíduos determinados e considerados para as análises subsequentes.

No alinhamento final de 482 pares de bases (pb) das sequências verificou-se existirem 50 sítios polimórficos de entre os quais 47 substituições ou *segretating sites* (S) e 3 inserções (*indels*) (**Tabela 3.1**). De entre as substituições 45 são transições e 2 são transversões com um rácio de 22.5. A composição nucleotídica é a seguinte: **A** (22.41%), **T** (32.12%), **C** (11.83%) e **G** (33.64%).

Tabela 3.1: Posições polimórficas no alinhamento das sequências de ADNmt da região HVR1 dos chimpanzés de Cantanhez. Os pontos indicam as posições onde as bases são idênticas ao haplótipo H1 e os traços onde ocorreram inserções.

	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3										
	2	4	4	8	8	0	0	1	1	2	2	3	3	3	3	6	6	7	8	8	9	9	9	0	0	1	1	1	2	2	3	4	5	5	6	6	0	1	1	1	2	2	
	0	2	9	7	8	0	6	2	3	5	7	0	2	4	5	9	0	4	7	3	4	4	7	8	0	4	0	4	5	7	8	9	6	5	8	3	5	9	0	5	7	2	5
H1	A	A	C	G	T	T	A	A	T	G	G	G	T	G	A	G	G	G	C	T	T	G	G	-	A	T	G	C	G	A	A	A	T	T	T	A	A	A	A	G	-	G	T
H2	G	.	.	A	C	.	.	G	G	A	.	.	A	A	G	.	.	A	.	.	C	.	-	-	.	C	A	.	.	G	G	G	.	.	C	G	.	G	G	.	A	A	.
H3	.	.	.	A	C	.	.	G	G	A	.	.	A	A	G	.	.	A	.	.	C	.	-	-	.	A	.	.	G	G	.	.	C	C	G	.	G	G	.	A	A	.	
H4	.	.	T	.	C	C	.	.	.	A	T	.	.	A	.	-	-	.	T	A	C	.	G	G	G	.	A	-	A	.		
H5	.	.	T	.	.	C	.	.	.	A	A	.	T	.	.	A	.	-	-	.	T	A	C	.	G	G	.	.	A	-	A	C			
H6	.	.	.	A	C	.	.	G	G	A	.	.	A	A	G	.	.	A	.	.	C	.	-	-	.	C	A	.	.	G	G	.	.	C	C	G	.	G	G	.	A	A	.
H7	.	.	.	A	C	.	.	G	G	A	.	.	A	A	.	.	A	.	.	C	.	G	G	.	C	A	.	.	G	G	.	.	C	C	G	.	G	G	.	A	A	.	
H8	.	G	T	.	.	C	G	G	.	A	A	.	.	A	A	.	C	.	.	-	G	.	.	.	G	.	G	.	C	.	G	
H9	.	G	T	.	.	C	G	G	.	A	A	A	.	.	A	A	.	C	.	.	-	G	.	.	.	G	.	G	.	C	.	G	
H10	.	.	.	A	C	.	.	G	G	A	.	.	A	A	.	.	A	.	.	C	.	-	-	.	C	A	.	.	G	G	.	.	C	C	G	.	G	G	.	A	A	.	
H11	.	.	.	A	C	.	.	G	G	A	.	.	A	A	G	.	.	A	.	.	C	.	G	.	C	A	.	.	G	G	.	.	C	C	G	.	G	G	.	A	A	.	
H12	.	.	.	A	C	.	.	G	G	A	.	.	A	A	G	.	.	A	.	.	C	.	-	-	.	C	A	.	.	G	G	.	.	C	G	.	G	G	.	A	A	.	
H13	.	G	T	.	.	C	G	G	.	A	A	.	.	A	A	.	C	.	.	-	G	.	.	.	G	.	G	.	C	.	G	
H14
H15	G	A	A	G	
H16	G	.	.	A	C	.	.	G	G	A	.	.	A	A	G	.	.	A	.	.	C	.	-	-	.	C	A	.	.	G	G	.	.	C	G	.	G	G	.	A	A	.	

Encontraram-se 16 haplótipos HVR1 na região sul do Parque Nacional de Cantanhez. Dos haplótipos encontrados, 44% (n=7) são partilhados por mais do que uma comunidade de chimpanzés (H1, H2, H5, H6, H8, H10, H14) e os restantes 56% (n=9) são específicos de determinado mato (H3, H4, H7, H9, H11, H12, H14, H15, H16). O mato de Cafatche possui o maior número de haplótipos exclusivos (n=2) seguido de todos os outros com apenas um haplótipo à excepção de Farim que não possui haplótipos raros (**Tabela 3.2**). A grande maioria (89%) dos haplótipos estão presentes em mais do que um indivíduo, sendo o haplótipo H2 o mais comum (28%), partilhado por 18 indivíduos ocorrendo ainda em 7 dos 9 matos amostrados, sendo os matos do Norte São Francisco (SFR) e Ponta Nova (PNO), as excepções.

Tabela 3.2: Distribuição e frequência dos haplótipos observados para os chimpanzés de Cantanhez em cada mato amostrado: CAD (Cadique), LAU (Lautchande), CAT (Catombói), FAR (Farim), CAF (Cafatche), CAN (Canghode), AMI (Amidara), SFR (São Francisco) e PNO (Ponta Nova).

	CAD	LAU	CAT	FAR	CAF	CAN	AMI	SFR	PNO	TOTAL
H1	6	-	-	-	-	-	1	-	1	8
H2	1	1	1	3	6	5	1	-	-	18
H3	1	-	-	-	-	-	-	-	-	1
H4	-	5	-	-	-	-	-	-	-	5
H5	-	2	-	-	-	-	2	-	1	5
H6	-	1	-	2	-	-	-	-	-	3
H7	-	-	1	-	-	-	-	-	-	1
H8	-	-	-	-	1	-	3	-	-	4
H9	-	-	-	-	1	-	-	-	-	1
H10	-	1	-	-	2	2	-	-	-	5
H11	-	-	-	-	1	-	-	-	-	1
H12	-	-	-	-	-	1	-	-	-	1
H13	-	1	-	-	-	-	1	2	3	7
H14	-	-	-	-	-	-	1	-	-	1
H15	-	-	-	-	-	-	-	3	-	3
H16	-	-	-	-	-	-	-	-	1	1
TOTAL	8	11	2	5	11	8	9	5	6	65
Hap. Únicos	1	1	1	0	2	1	1	1	1	9

3.2- Análise da diversidade mitocondrial HVR1:

A diversidade genética (i.e. haplotípica) (h) estimada para a globalidade do dataset foi de: 87.9% com uma variação geográfica entre 60% (FAR) a 100% (CAT). A diversidade nucleotídica (π) total estimada foi 0.037 (± 0.001) com o valor observado mais elevado em Ponta Nova (0.038 \pm 0.009) e o menor em Catombói (0.001 \pm 0.006). O número médio de pares de diferenças (k) observado para a totalidade do dataset foi de 17.788 diferenças entre pares de haplótipos com uma variação de 2.4 diferenças para os haplótipos de Farim e 18.3 diferenças para os de Ponta Nova. Uma medida alternativa à diversidade nucleotídica é a estimativa de Watterson (θ_w) (1975). Este parâmetro permite sumarizar a taxa pelo qual os processos de mutação e deriva genética criam e eliminam a variação numa determinada sequência assumindo a neutralidade ou seja, a não influência de recombinação. Baseia-se no número de sítios polimórficos numa amostra de sequências mas é independente da sua frequência ao contrário de π . A média de polimorfismos nucleotídicos estimada para o total do dataset é 2.1% com uma variação entre 0.4% (FAR) e 3.9% (PNO). Um sumário descritivo dos índices de diversidade genética é apresentado na **Tabela 3.3**.

Tabela 3.3: Sumário dos índices de diversidade molecular para o chimpanzés de Cantanhez por local amostrado. Amostra (N); número de haplótipos (H); sítios polimórficos (S); diversidade haplotípica ($h \pm SD$); número médio de pares de diferenças (K); diversidade nucleotídica ($\pi \pm SD$) e; theta *per site* a partir de S (θ_w), segundo Watterson (1975).

Origem Geográfica	Latitude-Longitude	N	H	S	$h \pm SD$	k	$\pi \pm SD$	θ_w
Cadique (CAD)	N11 14.332- W15 06.855	8	3	26	0.464±0.200	10.3	0.021±0.009	0.020
Lautchande (LAU)	N11 14.919- W15 04.340	11	6	41	0.800±0.114	16.1	0.034±0.007	0.029
Catomboi (CAT)	N11 11.638- W15 04.606	2	2	6	1.000±0.500	6	0.001±0.006	0.013
Farim (FAR)	N11 12.517- W15 04.430	5	5	4	0.600±0.175	2.4	0.005±0.001	0.004
Cafatche (CAF)	N11 19.260- W15 06.533	11	5	29	0.709±0.137	10.3	0.021±0.007	0.020
Canghode (CAN)	N11 09.678- W15 07.591	8	3	6	0.607±0.164	2.8	0.006±0.002	0.005
Amidara (AMI)	N11 16.873- W14 58.567	9	6	43	0.889±0.091	16.6	0.035±0.006	0.033
São Francisco (SFR)	N11 19.316- W14 59.487	5	2	21	0.600±0.175	12.6	0.026±0.008	0.021
Ponta Nova (PNO)	N11 18.694- W14 58.856	6	4	43	0.800±0.172	18.3	0.038±0.009	0.039
GLOBAL	-	65	16	50	0.879±0.024	17.788	0.037±0.001	0.021

3.3- Análise filogenética:

De forma a perceber a relação filogenética entre os haplótipos observados realizaram-se três análises independentes com o mesmo dataset de modo a inferir a história evolutiva dos chimpanzés de Cantanhez.

A topologia desta relação foi analisada utilizando três técnicas: i)- máxima parcimónia (MP); ii)- Neighbor-joining (NJ) e iii)- Inferência Bayesiana (BI). A MP realizada pela análise das sequências, baseia-se no menor número possível de

substituições. A árvore mais parcimoniosa foi encontrada através de uma busca heurística com 1000 réplicas de adição de sequências (i.e. *bootstrapping*). A NJ, pela análise das distâncias de Hamming entre sequências, com um factor de correcção do modelo Kimura 2-P (apropriado para o rácio de ts/tv observado) e com 10 000 bootstraps e a BI que se baseia *a priori* nas sequências e no seu modelo evolutivo de substituição de nucleótidos mais apropriado. Utilizou-se o critério de correcção Akaike (AICc) implementado no software jMODELTEST, cuja melhor probabilidade de selecção revelou ser o modelo HKI+I (-lnL = 1328.6298) (Hasegawa et al. 1985) com uma proporção de sites invariáveis de 0.6260. As probabilidades posteriores foram inferidas utilizando o algoritmo: Metropolis-Coupled das Cadeias Markov Monte Carlo (MC-MCMC) a partir de 1 000 000 de gerações, com um sampling a cada 100, 4 cadeias, 2 corridas e um *burn-in* inicial de 25%. Estes parâmetros foram implementados no software Mr. Bayes. A estacionariedade das cadeias foi estabelecida a partir de $p < 0.01$ *split frequencies*. A sua convergência foi analisada no software TRACER combinando as duas corridas no LogCombiner e tendo em atenção o valor do tamanho da amostragem estimada que recomenda que uma convergência correcta deve apresentar valores de ESS > 200: (LnL = -1445.947; ESS = 1141.57). As árvores de consenso dos três métodos foram concatenadas a partir dos seus valores de bootstrap (MP e NJ) e probabilidade posterior (BI) utilizando o software TreeGraph2. Em todos os métodos foi incorporada uma sequência de Bonobo (*Pan paniscus*) como *outgroup*.

A topologia das árvores de consenso é congruente nos três métodos (MP, NJ, BI) e caracteriza a existência de duas clades monofiléticas recíprocas ou linhagens (i.e. matrilineas) principais de ADNmt nos chimpanzés de Cantanhez. Estas linhagens estão bem suportadas pelos elevados valores de bootstrap e/ou probabilidades posteriores apresentados: Linhagem A: 72% (MP), 70% (NJ) e 95% (BI); Linhagem B: 100% (MP), 100% (NJ) e 100% (BI) respectivamente. A existência destas 2 linhagens principais pode observar-se no cladograma concatenado das três inferências (**Fig. 3.1**). A análise mais detalhada das linhagens revelou ainda existirem 5 haplogrupos. A linhagem A engloba o filograma mais resolvido determinado por inferência bayesiana (**Fig. 3.2**) que em conjunto com o cladograma concatenado mostrou ainda

a possibilidade de existirem 3 haplogrupos dentro desta linhagem. O haplogrupo A1 (constituído pelos haplótipos: H8, H13 e H9 apresenta um conflito entre métodos e não é suportado no nó interno mais ancestral sendo-o todavia suportado pelo nó intermédio (**Fig.3.1**). O haplogrupo A2 constituído pelos haplótipos H4 e H5 não demonstra qualquer conflito entre métodos e apresenta-se resolvido com suporte elevado (MP=97%; NJ= 100%; BI= 99%). O haplogrupo A3 (constituído pelos haplótipos: H15, H1 e H14) está igualmente bem suportado (MP= 89%; NJ=97%; BI=77%) e não agrupa com nenhum dos anteriores. Não obstante, a relação entre os haplogrupos A1 e A2 não é totalmente clara já que o nó interno comum mais ancestral não tem suporte (MP=38%; 51%, BI=45%), onde ocorrem também 2 conflitos entre árvores não se podendo inferir que agrupem juntos ou partilhem o mesmo ancestral, mas demonstrando claramente que também não agrupam com o haplogrupo A3.

A linhagem B contém dois haplogrupos. O haplogrupo B1, é constituído pelos haplótipos: H7 e H10, não demonstrando conflitos de topologia nos nós internos estando ainda bem resolvido pela inferência bayesiana (BI= 83%). O haplogrupo B2 é constituído pelos haplótipos: H3, H16, H11, H12, H2 e H16 estando bem suportado pela probabilidade posterior de 98% no nó interno mais ancestral, provando não agrupar com o haplogrupo B1. No entanto os ramos terminais não estão totalmente bem suportados e existe um conflito de resolução apresentado pelo baixo valor de bootstrapping e probabilidade posterior.

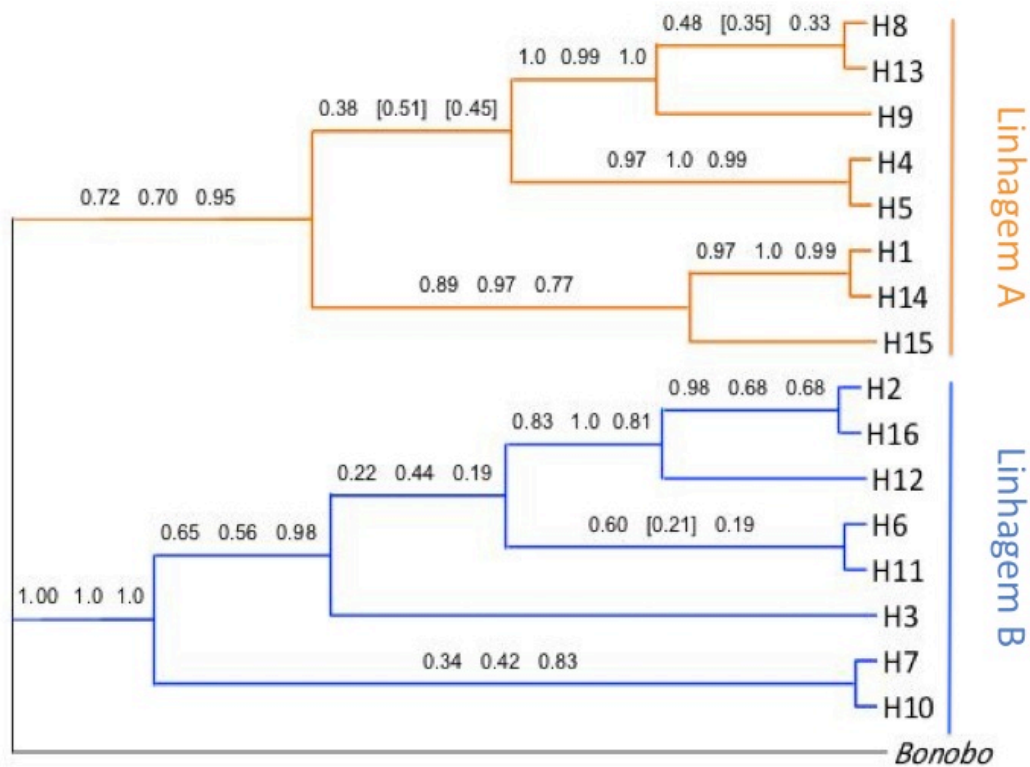


Fig. 3.1:Cladograma das relações filogenéticas entre haplótipos dos chimpanzés de Cantanhez de acordo com os métodos de Máxima Parsimónia (MP), Neighbor-Joining (NJ) e Inferência Bayesiana (BI) concatenados no software TreeGraph2. Os valores apresentados por ordem, em cada nó, referem-se aos valores de *bootstrap* para MP e NJ e probabilidades posteriores para BI, respectivamente. Os valores dentro de parenteses rectos referem-se a conflitos de resolução na topologia dos nós onde a resolução num determinado método é diferente.

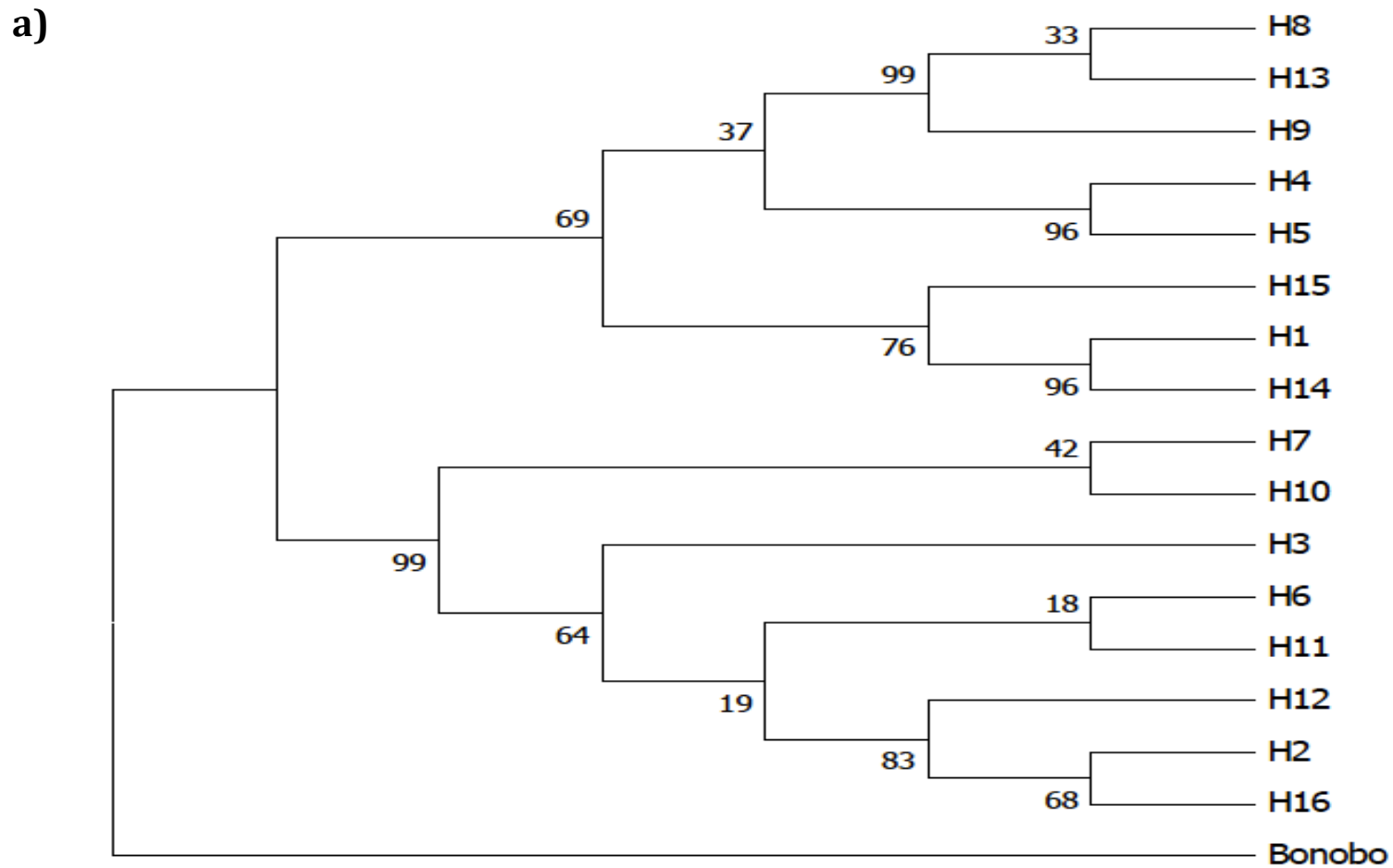


Fig. 3.2: a)- Cladograma sem raiz de máxima parcimónia (MP) com 1000 replicações e busca heurística , condensada a partir de 50% (bootstrap) de cut off.

b)

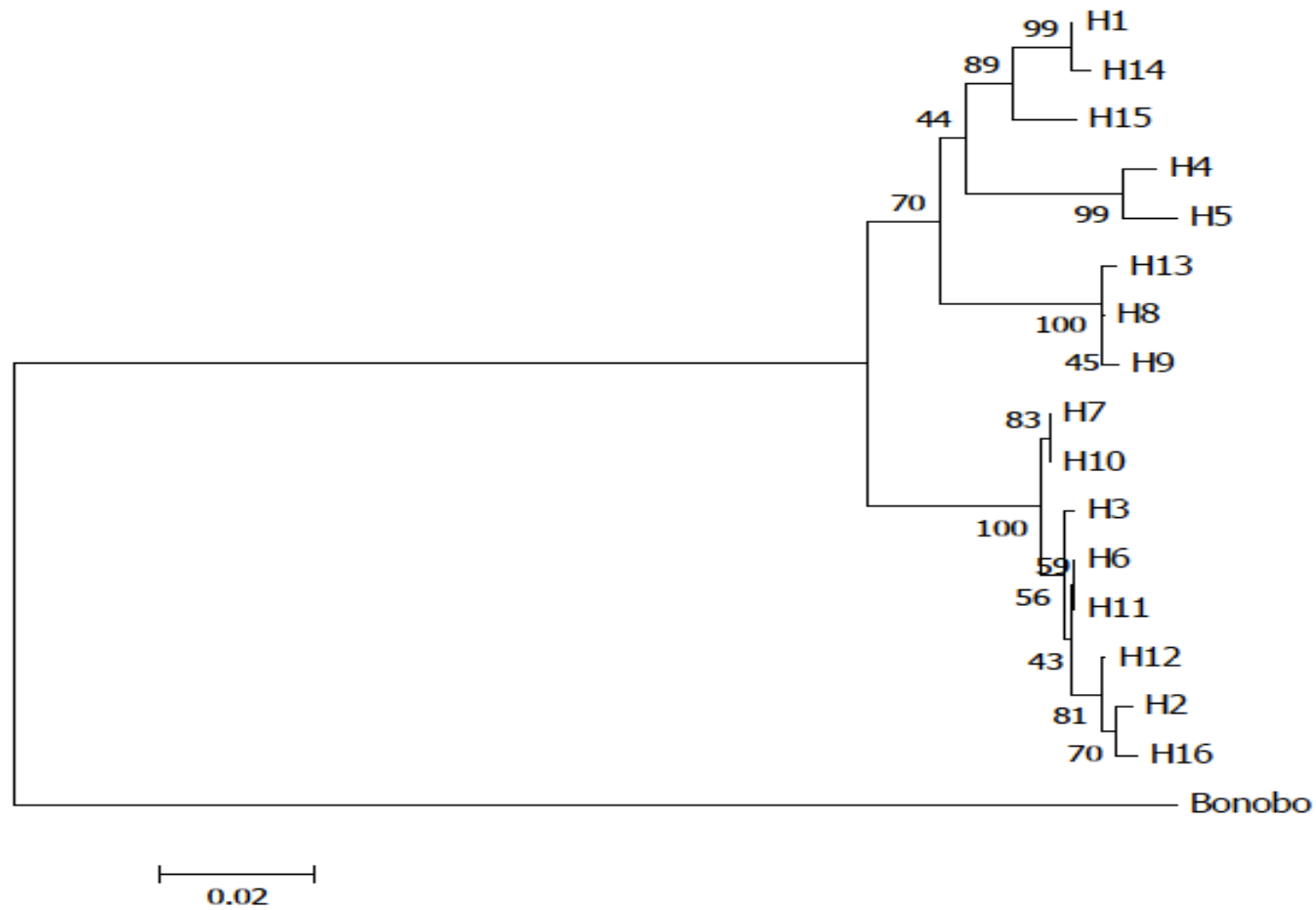


Fig. 3.2: b)- Filograma dos haplótipos mtDNA HVR1 dos chimpanzés de Cantanhez por Neighbour-joining (NJ) com 10000 reposições.

c)

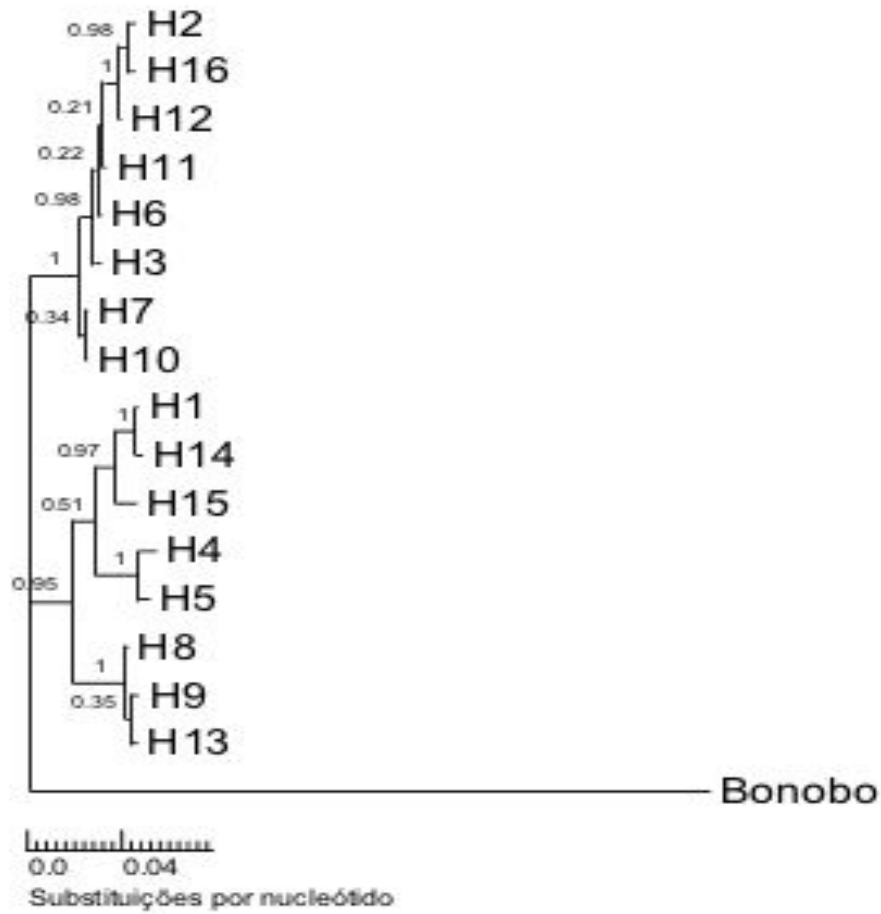


Fig. 3.2: c)- Filograma dos haplótipos mtDNA HVR1 dos chimpanzés de Cantanhez por inferência bayesiana (BI). Valores das probabilidades posteriores indicados em cada nó..

3.4- Análise da Network filogenética:

Uma alternativa complementar à representação de árvores filogenéticas são as networks filogenéticas. Estas networks têm a vantagem de perceber as relações filogenéticas entre haplótipos e ao invés das árvores filogenéticas permitem que mecanismos de evolução reticulada (e.g. transferência horizontal de genes, hibridação, recombinação e *'reassortment'* sejam visualizados). Por outro lado, através desta abordagem é possível visualizar com mais detalhe mecanismos de *'incomplete lineage sorting'* ou mesmo padrões complexos de duplicação e perda de genes que numa árvore filogenética se mostrariam não resolvidos (Huson et al. 2010). Bandelt et al. (1999) propuseram o método de *'Median-Joining Networks'* para inferir filogenias intraespecíficas que permitem fazer a reconstrução filogenética através da combinação do algoritmo utilizado para encontrar *'minimum spanning trees'* (i.e. que favorece as conexões mais curtas entre haplótipos) e o algoritmo heurístico da máxima parcimónia que adiciona vectores médios por forma a resolver a relação entre haplótipos (Huson et al. 2010). Este método tem sido particularmente usado em estudos de reconstrução filogenética humana (Kivisild et al. 1999). Na figura 3.3 pode observar-se a network filogenética para os chimpanzés de Cantanhez baseada também no algoritmo *'median-joining'*. Os 5 haplogrupos estão claramente identificados e pode verificar-se que a Linhagem A está separada da Linhagem B por 20 passos mutacionais. A network mostra ainda que os haplogrupos da linhagem B estão mais próximos entre si (separados apenas por 2 passos mutacionais) do que os haplogrupos na linhagem A (A1 distancia-se de A2 com 18 passos mutacionais enquanto que A2 se diferencia de A3 com 13 passos mutacionais). Na figura 3.4 pode observar-se a distribuição geográfica dos haplótipos, onde se verifica que H2 é o haplótipo mais frequente nos matos de Cantanhez. A distribuição geográfica dos haplótipos demonstra ainda que não existe uma aparente estrutura genética.

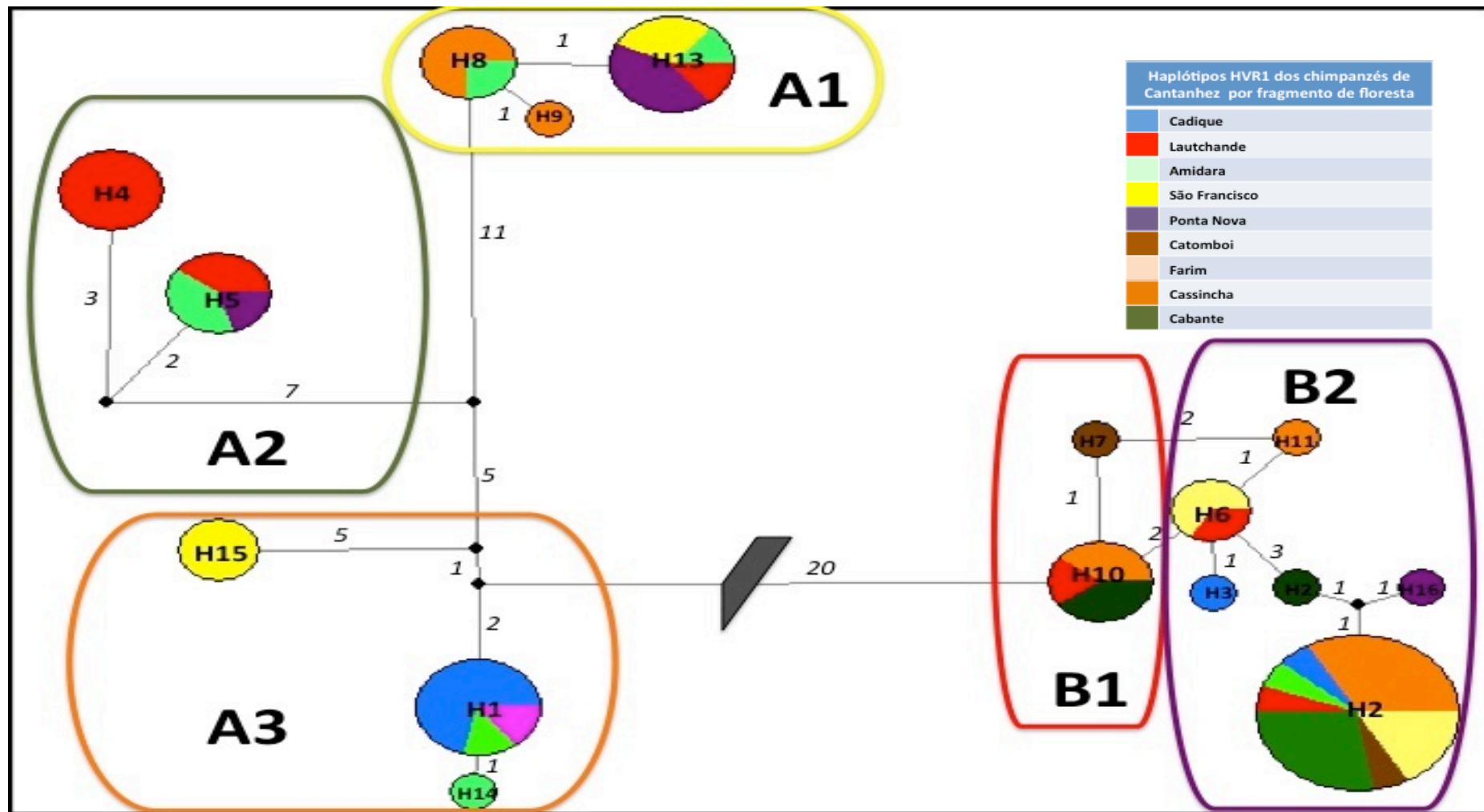


Figura 3.3- Median-Joining network da relação entre haplótipos dos chimpanzés de Cantanhez. Cada círculo representa um haplótipo e o seu tamanho é proporcional à frequência dos haplótipos. Os pequenos círculos assinalados a preto representam vectores médios não amostrados ou extintos. Cada haplogrupo está indicado com uma coloração própria (A1, A2, A3, B1 e B2) respectivamente. A separação entre a Linhagem A e B está indicada através de uma seta de intersecção.

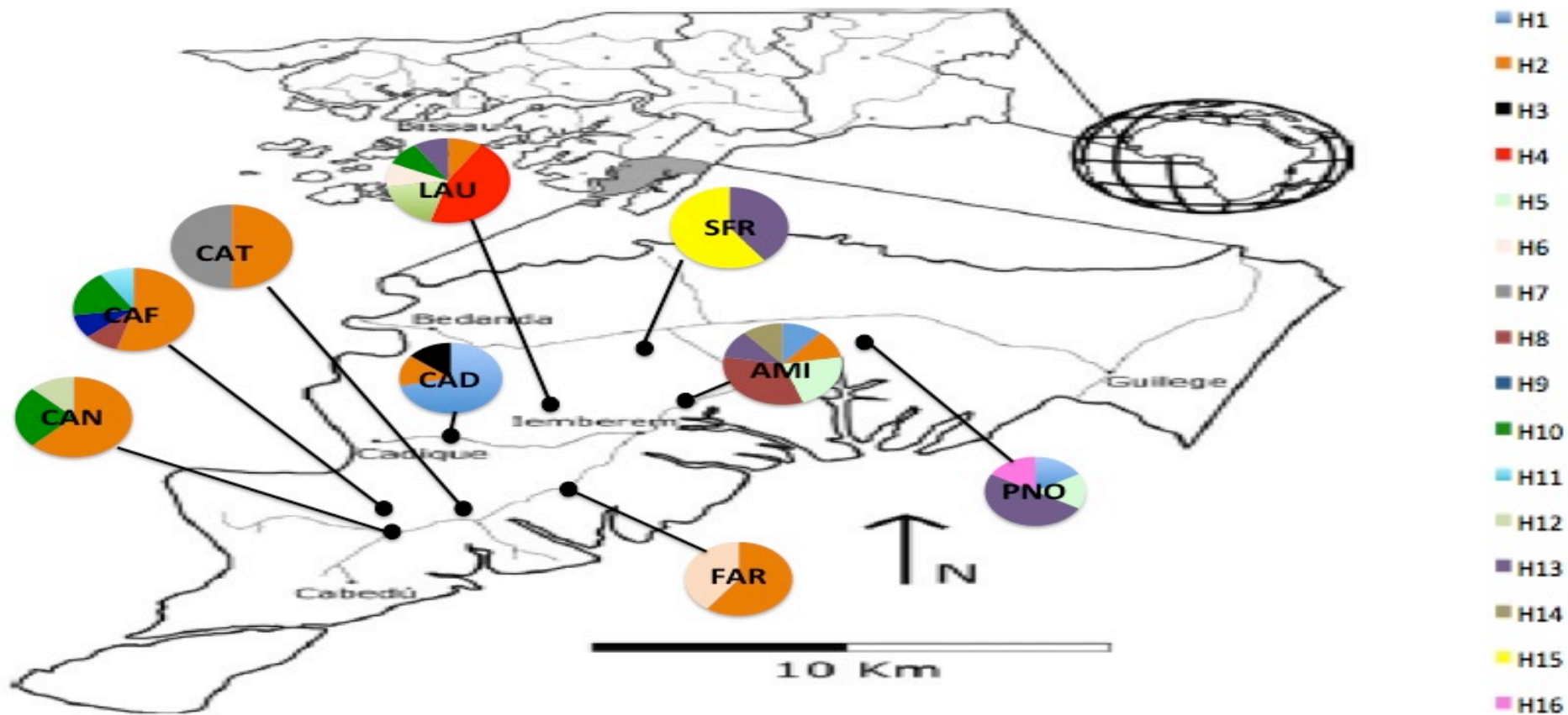


Figura 3.4: Distribuição de haplótipos HVR1 dos chimpanzés de Cantanhez por fragmento de floresta.

3.5- Análise da história demográfica:

Para além da Selecção Natural, também as alterações demográficas das populações deixam um *'footprint'* nos genomas das espécies (Avice 2000). A demografia histórica de uma espécie está intimamente relacionada com a sua filogenia. Com a forma como os seus ramos genealógicos são formados e com os seus ancestrais comuns (Avice 2000). Expansões demográficas recentes por exemplo, fazem com que os ramos terminais de uma árvore filogenética estejam mais dispersos enquanto que os nós internos estão mais comprimidos (i.e. genealogia em forma de estrela: *'star-shape'*). Desenvolveram-se assim testes estatísticos que permitem detectar se a população se manteve constante ou, se pelo contrário, houve uma expansão demográfica (Ramos-Onsins e Rozas 2002).

3.5.1- Testes de Neutralidade:

Estes testes baseiam-se na frequência de mutações ou *segregating sites*. A teoria prediz que uma expansão gera um excesso de mutações recentes nos ramos terminais de uma genealogia e por sua vez um excesso de substituições presentes numa só sequência amostrada (i.e. *singletons*). Os testes Tajima D, Fu e Li D* e F* e Fu F_S detectam estas alterações demográficas e adequam-se a dados intraespecíficos (Ramos-Onsins e Rozas 2002). O teste Tajima D assume que no modelo de evolução neutral as estimativas do parâmetro de diversidade de Watterson (θ_w) e a *'estimativa Tajima'* baseada no número médio de diferenças entre pares de sequências estão correlacionadas (Tajima 1989). O teste Fu and Li D* baseia-se na diferença entre o número de *singletons* e o número total de mutações, enquanto que o teste Fu and Li F* se baseia nas diferenças entre o número de *singletons* e o número médio de diferenças entre pares de sequências (Fu e Li 1993). Por sua vez o teste Fu F_S baseia-se na distribuição haplotípica e espera-se que tenha valores baixos decorrentes do excesso de *singletons* provocados pela expansão demográfica (Fu 1997). Todos os testes mencionados testam a hipótese nula de uma população de tamanho constante e a significância é verificada através de simulações.

No presente estudo, estes testes foram aplicados em dois níveis:

- i)- à globalidade do dataset e;
- ii)- de acordo com a agregação das sequências às duas linhagens determinadas filogeneticamente.

Os valores obtidos para os testes de neutralidade apresentam-se positivos. Os valores do teste Tajima D, bem como dos testes Fu F_s e Fu e Li F^* , para a globalidade do dataset e para a Linhagem A revelaram-se significativos. Só o teste Fu e Li D^* não apresentou resultados significativos para o total do dataset nem para as duas linhagens. Destaca-se ainda que na Linhagem B todos os valores observados pelos testes de neutralidade foram baixos mas nenhum se mostrou significativo. Os resultados dos testes de neutralidade apresentam-se a seguir na **Tabela 3.4**.

Tabela 3.4: Sumário dos testes de neutralidade.

Testes de Neutralidade	Linhagem A	Linhagem B	Total
Tajima D	2.12411*	1.07116	2.63401*
Fu F_s	9.89042**	0.48395	9.09957***
Fu e Li F^*	1.92967**	0.41744	2.25656**
Fu e Li D^*	1.38525	0.04426	1.42929

* $p < 0.05$; ** $p < 0.02$; *** $p \leq 0.001$

3.5.2- Mismatch distributions:

As alterações da história demográfica da população (expansões ou *bottlenecks*) foram também inferidas comparando os *mismatch distributions* para a totalidade do dataset. Os *mismatch distributions* são histogramas modais entre pares de diferenças genéticas, isto é, um mismatch distribution representa a distribuição do número observado de diferenças entre pares de haplótipos (Rogers e Harpending 1992). Este método, que permite caracterizar a história demográfica recente foi proposto por Rogers e Harpending (1992) e teve uma grande aplicação na caracterização e projecção da expansão demográfica humana (Marjoram e

Donnelly 1994). O facto de fornecerem uma perspectiva mais recente na janela temporal evolutiva de uma espécie do que uma árvore filogenética ou mesmo dos sumários descritivos das diversidades genéticas que, assentam sobretudo nas diferenças máximas entre pares de sequências, e não implicam também qualquer reconstrução filogenética. Este método é por isso vantajoso e complementar às análises filogenéticas porque, o tempo evolutivo, no qual as mutações informativas poderão aparecer, é curto entre linhagens relativamente próximas (Rogers e Jorde 1995). Num mismatch distribution, a forma e a posição da curva do histograma representa características populacionais que segue a coalescência dos nós mais divergentes de uma árvore filogenética (Rogers 1995). O pico da curva ocorre imediatamente antes de um evento de expansão demográfica. A posição e a forma da curva no histograma é proporcional à idade e dinâmica demográfica da população desde a sua expansão. Geralmente a forma de um mismatch distribution é multimodal em amostras retiradas de populações em equilíbrio demográfico, ou pelo contrário, uma curva com forma unimodal representa geralmente uma expansão demográfica recente, um bottleneck, ou uma série de expansões com níveis elevados de migração entre *demes* vizinhos, sendo difícil fazer a distinção entre os dois primeiros eventos demográficos (Slatkin e Hudson 1991). Três parâmetros derivam teoricamente do mismatch distribution que definem e delimitam a curva do histograma:

$$1) \Theta_0 = 2N_0u;$$

$$2) \Theta_1 = 2N_1tu;$$

$$3) \tau = 2ut, \text{ onde } u = m_t u;$$

onde, N_0 é o tamanho populacional de fêmeas reprodutoras que precedem a expansão. N_1 é o tamanho populacional pós-expansão, t é o tempo decorrido desde a expansão populacional, u é a probabilidade de que uma mutação atinja um nucleótido numa sequência em cada geração e m_t é o tamanho da sequência em estudo (Rogers and Harpending 1992). Os valores destes parâmetros definem a forma dos mismatch distribution e τ define a localização do pico da curva sendo directamente proporcional ao tempo no qual a população esteve sobre expansão.

Os mismatch distribution podem ser calculados usando três modelos diferentes: expansão demográfica, expansão populacional súbita (em inglês: sudden

expansion model) e expansão espacial (em inglês: spatial expansion model). Os primeiros dois modelos são muito semelhantes e nesse sentido, no presente estudo, só os modelos de expansão súbita e expansão espacial foram calculados.

Uma expansão espacial ocorre geralmente se o range da população está inicialmente condicionado a uma área muito pequena, onde posteriormente o range da população aumenta com o tempo e com o espaço. No modelo de expansão espacial, a forma das genealogias e o padrão geral de diversidade entre demes é influenciado pela idade da expansão e pelo nível de fluxo de genes, Nm (onde: N = tamanho do deme e m = proporção de migrantes) (Ray et al. 2003; Ray et al. 2009). Um baixo fluxo de genes (<1 migrante por geração) produz uma proporção substancial de eventos coalescentes no início da genealogia que têm geralmente um padrão em forma de estrela e um mismatch distribution multimodal. Para valores elevados de Nm , a maioria dos eventos coalescentes que ocorreram na altura da expansão espacial produzem um misto de ramos curtos e largos na filogenia resultando num mismatch distribution unimodal.

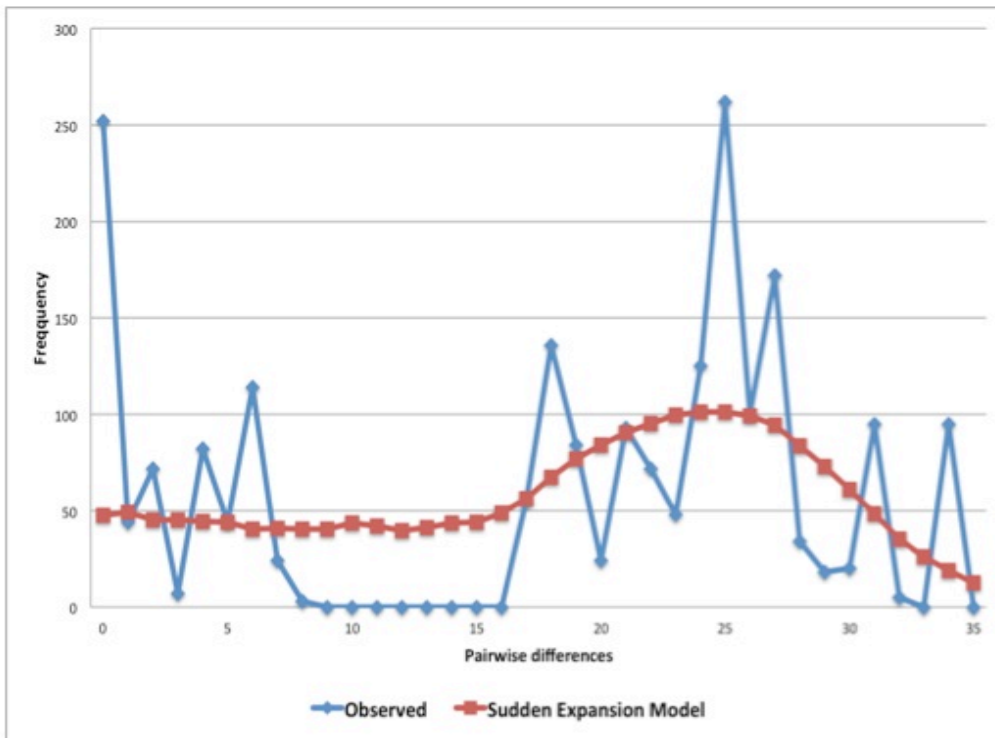
A validade global e significância do mismatch distribution testa-se obtendo a distribuição do teste estatístico SSD (soma do quadrado das diferenças) entre os mismatch distributions observados e esperados. Um valor significativo de SSD é aceite como evidência de desvio do modelo demográfico esperado (Schneider e Excoffier 1999).

Uma estatística complementar ao SSD que avalia também o desvio do mismatch distribution em relação á hipótese nula do modelo esperado é o *Índice raggedness* ou r . Em 1994, Harpending sugeriu que se utilizasse este teste *ad hoc* que calcula a soma do quadrado das diferenças entre valores sequenciais num mismatch distribution (Harpending 1994). Em termos gerais um baixo valor no índice raggedness indica uma distribuição unimodal uniforme e uma história populacional caracterizada por uma expansão recente, por outro lado, valores elevados indicam uma distribuição multimodal errática e uma história populacional caracterizada por estabilidade demográfica. A simulação de diversos modelos de mismatch distribution demonstrou ainda que populações que experienciaram um rápido crescimento populacional apresentam valores baixos de r na ordem dos 0.012 enquanto que distribuições de populações em equilíbrio apresentam valores na

ordem dos 0.26 (Harpending et al. 1993; Harpending 1994). O índice raggedness é robusto para populações subdividas, níveis diferentes de migração entre subpopulações, para dados com pouca resolução e ainda para um tamanho da amostra pequeno ou com múltiplas substituições (Rogers 1995; Rogers e Jorde 1995; Roger et al. 1996; Schneider e Excoffier 1999). No presente estudo utilizaram-se estes dois testes estatísticos para avaliar quão os dados observados se desviaram do modelo esperado.

A **Fig. 3.5** ilustra os mismatch distribution obtidos para os haplótipos dos chimpanzés de Cantanhez. O painel **3.5 a)** mostra uma expressão geométrica de mismatch distribution multimodal e sugere uma população de tamanho demográfico constante confirmado pelo teste *ad hoc* SSD muito significativo ($\rho < 0.001$) cujos dados observados não estão em conformidade com a hipótese nula do modelo esperado de expansão demográfica súbita. Não obstante os valores elevados de τ e a posição do pico da curva, que ocorre entre os 23 e 27 pares de diferenças entre os haplótipos, localizando-se mais á direita do mismatch, sugere uma expansão mais ancestral dos chimpanzés das florestas de Cantanhez, embora os valores do índice raggedness sejam relativamente baixos ($r = 0.045$) mas muito significativos ($\rho < 0.001$) rejeitando assim a hipótese nula de que a variação da curva é uniforme.

O painel **3.5 b)** demonstra um padrão quase bimodal de mismatch distribution assumindo um 'spatial expansion model' com um pico inicial localizado à esquerda entre os 0 e os 2 pares de diferenças e um pico mais ancestral localizado à direita do mismatch entre os 23 e os 27 pares de diferenças entre haplótipos. Os testes *ad hoc* (SSD e r) que avaliam a adequabilidade dos dados observados ao modelo de expansão espacial embora significativos são muito marginais ($\rho = 0.047$) e o valor de τ é também elevado ($\tau = 25.562$). Estas observações sugerem subdivisão populacional, várias expansões demográficas no passado, várias ondas de migração entre diferentes subpopulações ou uma combinação dos factores acima mencionados (Rogers and Jorde 1995; Gonder 2000). O número de fêmeas chimpanzés migrantes (M) entre subpopulações por geração estima-se em 4.065.



a)

$K=17.58$

$\Theta_0=0.23$

$\tau=25.562$

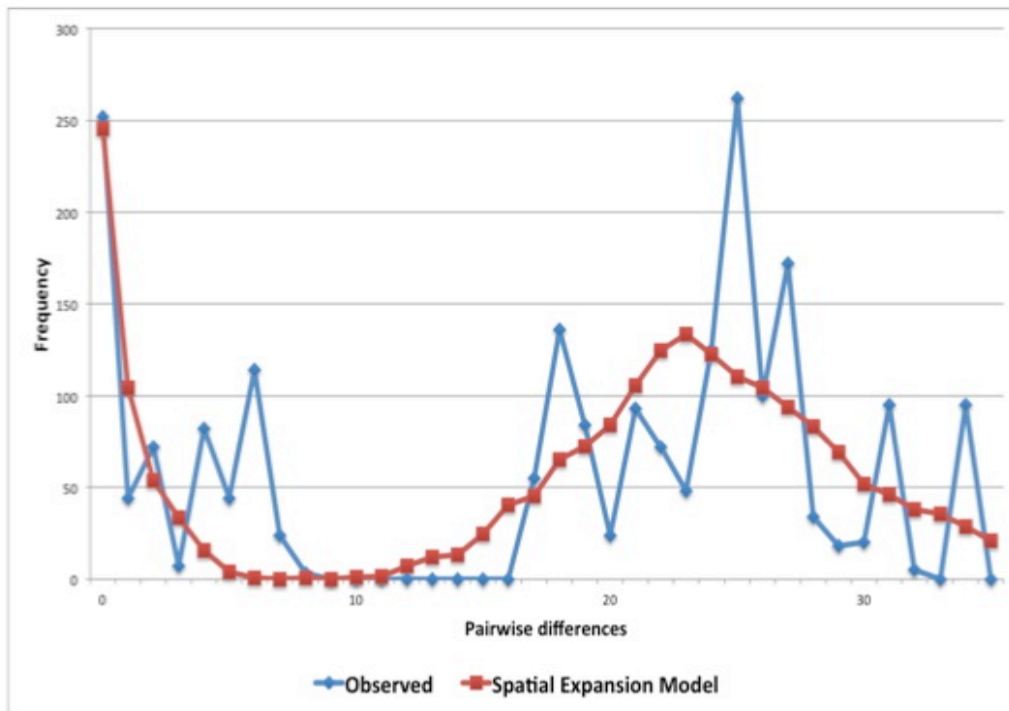
$\Theta_1=41.226$

$SSD=0.028$

$(p<0.001)$

Raggedness index: 0.045

$(p<0.001)$



b)

$K=17.58$

$\tau=25.562$

$\Theta=3.574$

$SSD=0.020$

$(p=0.47)$

Raggedness index: 0.045

$(p=0.47)$

$M=4.065$

Fig. 3.5: Mismatch distributions dos chimpanzés de Cantanhez e respectivas estatísticas associadas. **a)**-Mismatch distribution segundo o “sudden expansion model); **b)**- Mismatch distribution segundo o “spatial expansion model”. K (diferenças médias entre pares de haplótipos); Θ (diversidade genética) Θ_0 , τ , Θ_1 (parâmetros derivados do mismatch distribution). Testes de ajustamento: SSD (Soma dos quadrados das diferenças) e Índice Raggedness de Harpending. M (número de fêmeas migrantes).

3.6- Análise da estrutura genética:

As origens da variação haplotípica observada numa espécie estão relacionadas com a sua história demográfica, mas esta não explica como é que essa variação está estruturada. A estrutura genética de uma espécie tem sido tradicionalmente estudada verificando os desvios das frequências alélicas das expectativas panmíticas e para isso desenvolveram-se procedimentos derivados essencialmente dos Índices de Fixação (em inglês: *F-statistics*) propostos por Wright (1951; 1965) que permitem descrever não só o nível de subdivisão populacional mas também estabelecer uma base para estimar o fluxo de genes entre essas várias subpopulações.

3.6.1- Análise da distância genética por F_{ST} :

Um indicador clássico e comparativo da estrutura genética de uma espécie tem sido o descritor F_{ST} . Este indicador pode definir-se como a redução média da diversidade genética de uma subpopulação (relativa ao total da população) devido aos efeitos da deriva genética entre subpopulações e fornece por isso uma medida da extensão da diferenciação genética entre as várias subpopulações. Esta medida varia entre 0 (ausência de diferenciação) e 1 (diferenciação completa) (Wright 1951). Os valores de F_{ST} são apresentados sob a forma de uma matriz de distâncias entre subpopulações e a sua significância é obtida por permutações de haplótipos entre as subpopulações sob a hipótese nula de inexistência de diferenciação genética. O valor de P é a proporção de permutações que levam a valores maiores ou iguais ao observado.

A diversidade HVR1 foi utilizada para determinar as distâncias genéticas F_{ST} entre todos os pares de fragmentos de floresta onde os chimpanzés habitam em Cantanhez (**Tabela 3.5**) verifica-se que a maioria das populações são significativamente diferenciadas indicando à partida que a maior parte da variação genética se encontra ao nível intra-populacional. Os valores mais baixos de F_{ST}

observaram-se entre os pares Cafatche e Catombói ($F_{ST} = -0.24132$) indicando que não existe diferenciação genética dos chimpanzés nestes dois fragmentos de floresta. Por outro lado o valor de F_{ST} mais elevado encontrou-se entre os pares Cangode e São Francisco ($F_{ST} = 0.72832$) indicando claramente que se tratam de comunidades de chimpanzés distintas em termos genéticos.

Curiosamente pode ainda observar-se que, em fragmentos de floresta quase adjacentes, existem distâncias genéticas elevadas como por exemplo entre Cadique e Lautchande ($F_{ST} = 0.31188$) ou entre Amidara e Farim ($F_{ST} = 0.54704$) sugerindo também que se tratam de comunidades de chimpanzés distintas.

Tabela 3.5: Distâncias genéticas F_{ST} entre pares de fragmentos de floresta de Cantanhez.

	Cadique	Lautchande	Amidara	São Francisco	Ponta Nova	Catomboi	Cafatche	Canghode	Farim
Cadique	–	*	*	*			*	*	*
Lautchande	0.31188*	–		*		*	*	*	*
Amidara	0.20360*	0.11455	–				*	*	*
São Francisco	0.22836*	0.25675*	0.02727	–			*	*	*
Ponta Nova	0.19513	0.11725	-0.13834	0.00970	–		*	*	*
Catomboi	0.48110	0.35997*	0.40822	0.54667	0.34235	–			
Cafatche	0.45146*	0.39580*	0.39707*	0.49762*	0.35090*	-0.24132	–		
Canghode	0.64643*	0.54968*	0.59458*	0.72832*	0.58677*	-0.18591	0.00909	–	
Farim	0.61425*	0.5460*	0.54704*	0.69272*	0.52574*	-0.17647	-0.03862	-0.11201	–

* $p < 0.05$

3.6.2- Análise da distância genética pelo método F_{ST} linearizado de Slatkin:

A análise anterior com F_{ST} clássicos demonstrou que havia algum grau de subdivisão e diferenciação genética entre as diferentes comunidades de chimpanzés em Cantanhez. No entanto, no sentido de complementar essa análise e para se perceber com mais detalhe esse grau de divisão entre as populações, Slatkin (1991, 1993, 1994a) demonstrou que as estimativas F_{ST} podem também fornecer informação não só sobre a divergência das populações, mas também sobre isolamento por distância. Slatkin (1991, 1994) propôs que os F_{ST} podem ser expressos em termos de tempo coalescente t_1 , que por sua vez são a média do tempo coalescente de dois genes retirados de duas populações irmãs diferentes, enquanto que t_0 é a média do tempo coalescente de dois genes retirados de uma mesma população. Esta abordagem é conhecida como modelo F_{ST} linearizado de Slatkin e representa-se pela seguinte expressão:

$$F_{st} = \frac{t_1 - t_0}{t_1}$$

onde, t_0 é igual a N gerações e t_1 é igual a $\tau + N$ gerações e nesse sentido a expressão anterior pode ser reduzida a $F_{ST} = \tau / (\tau + N)$ e por conseguinte o rácio:

$$D = \frac{F_{st}}{1 - F_{st}}$$

é igual a τ/N e é directamente proporcional ao tempo de divergência entre populações. Valores baixos de D reflectem associações mais recentes entre pares de populações enquanto que valores elevados de D indicam uma maior diferenciação genética entre pares de populações e eventos coalescentes mais ancestrais.

O modelo F_{ST} linearizado de Slatkin foi aplicado à totalidade do dataset e os resultados apresentam-se na **Tabela 3.6**. Os valores mais baixos de D foram observados entre os pares: Catombói-Cafatche; Catombói-Canghode; Catombói-Farim; Cafatche-Farim; Canghode-Farim na zona sul do PNC e entre Amidara-Ponta

Nova na zona norte do PNC ($D=0$ para todos os pares mencionados) indicando que não existe diferenciação genética entres estes pares e que os eventos coalescentes entre eles são muito recentes. Em contrapartida, os valores mais elevados foram observados entre os pares: Canghode-São-Francisco ($D=2.68081$); Farim-São-Francisco ($D=2.25432$) e Cadique-Canghode ($D=1.82828$) indicando que estes pares são os mais diferenciados geneticamente e que os eventos coalescentes entre estes pares são mais ancestrais. A matriz parece sugerir ainda algum nível de isolamento por distância na medida em que, quando a distância geográfica aumenta, a distância genética também aumenta (ex: Canghode-São Francisco: $D= 2.68081$).

Tabela 3.6: Distâncias genéticas pelo modelo F_{ST} linearizado de Slatkin entre pares de fragmentos de floresta de Cantanhez.

	Cadique	Lautchande	Catomboi	Cafatche	Canghode	Amidara	Farim	São Francisco	Ponta Nova
Cadique	–								
Lautchande	0.45324	–							
Catomboi	0.92714	0.56243	–						
Cafatche	0.82303	0.65509	0	–					
Canghode	1.82828	1.22065	0	0.00918	–				
Amidara	0.25565	0.12936	0.68983	0.065858	1.46658	–			
Farim	1.59237	1.01859	0	0	0	1.20772	–		
São Francisco	0.29594	0.34545	1.20592	0.99054	2.68081	0.02804	2.25432	–	
Ponta Nova	0.24243	0.13282	0.52056	0.54059	1.41996	0	1.10854	0.00980	–

3.6.3- Análise relativa da migração e do fluxo genético:

Os F_{ST} para além de permitirem determinar o nível de divergência genética entre subpopulações permitem também fazer inferências sobre os níveis de migração. Este facto deve-se a que os valores de F_{ST} observados são resultantes da deriva genética ao longo do tempo ou da migração de indivíduos entre subpopulações ou, a uma combinação destes. Se assumirmos que a divergência entre subpopulações e os F_{ST} atingiram o equilíbrio e que todas as subpopulações partilham migrantes com todas as outras subpopulações com a mesma probabilidade [modelo de “ilhas” de estrutura genética (Wright 1931)] pode-se assim estimar uma taxa relativa de migração.

M (número total e actual de migrantes partilhados entre duas populações) é o produto do tamanho efectivo da população (N_e) e do número absoluto de migrantes por geração. Takahata e Palumbi (1985) demonstraram que o produto de N e m pode ser obtido através das estimativas de F_{ST} para populações haplóides ($M=Nm$) através da seguinte equação:

$$M = \frac{1 - F_{st}}{2F_{st}}$$

Na **Tabela 3.7** estimam-se os valores relativos de M para os pares de floresta do PNC onde se pode verificar que o par Canghode-São Francisco apresenta o valor mais baixo de migração ($M=0.18651$) cujo fluxo genético é muito limitado.

Tabela 3.7: Estimativa relativa de migração absoluta (M) entre pares de fragmentos de floresta de Cantanhez.

	Cadique	Lautchande	Catomboi	Cafatche	Canghode	Amidara	Farim	São Francisco	Ponta Nova
Cadique	–								
Lautchande	1.10317	–							
Catomboi	0.53929	0.88900	–						
Cafatche	0.60751	0.76326	∞	–					
Canghode	0.27348	0.40962	∞	54.49441	–				
Amidara	1.95583	3.86504	0.72482	0.75921	0.34093	–			
Farim	0.31400	0.49088	∞	∞	∞	0.41400	–		
São Francisco	1.68953	1.44741	0.41462	0.50478	0.18651	17.83414	0.22180	–	
Ponta Nova	2.06242	3.76441	0.96050	0.92492	0.35212	∞	0.45104	51.04615	–

3.6.4-Análise da variância molecular (AMOVA):

Por forma a perceber como é que a variação genética dos chimpanzés de Cantanhez está particionada analisou-se a sua variância molecular (AMOVA) que também permite inferir a sua estrutura genética (Excoffier et al. 1992). Este método é igualmente derivado do modelo F_{ST} de Wright (1951) que estima as estatísticas-F (que medem o desvio das frequências genotípicas em populações subdivididas assumindo um modelo nulo de panmixia). Estas estatísticas ou *índices de fixação* organizam-se geralmente em três medidas de diversidade genética: dentro das populações (F_{IT}), entre populações dentro de regiões (F_{IS}) e entre regiões (F_{ST}). Índices de fixação positivos indicam um desvio em relação ao princípio de Hardy-Weimberg. Índices de fixação negativos, próximos de zero indicam ausência de estrutura genética.

Cockerham (1969), seguindo as adaptações do modelo F_{ST} de Wright particiona de forma hierárquica a variação genética total em componentes de covariância devido a diferenças: a) entre indivíduos dentro das populações; b) entre populações dentro das regiões e c) entre regiões. Nesta última partição, as regiões são grupos de populações que são definidos *a priori*, por exemplo: divisões regionais inferidas filogeneticamente ou agrupamentos geográficos de relevância ecológica. Excoffier et al. (1992) incorporaram informação sobre as distâncias moleculares entre haplótipos na abordagem da AMOVA e sugerem as 'estatísticas Φ ' semelhantes aos tradicionais índices de fixação. Segundo esta análise Φ_{ST} , Φ_{SC} e Φ_{CT} são usados como análogos dos índices de fixação dentro das populações, entre subpopulações dentro de regiões e entre regiões, respectivamente. Para cada um destes níveis as estatísticas Φ definem-se como:

a)
$$\Phi_{CT} = \frac{\sigma_a^2}{\sigma_T^2}$$

b)
$$\Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2}$$

c)
$$\Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_T^2}$$

onde, a variância molecular total (σ_T^2) é a soma das variâncias devido a diferenças entre os haplótipos entre indivíduos dentro das subpopulações (σ_C^2); variância devido a diferenças entre haplótipos em diferentes subpopulações dentro de regiões (σ_b^2) e; variância devido a diferenças entre regiões, populações ou grupos (σ_a^2) (Excoffier et al. 1992; Schneider et al. 1997). A significância dos índices Φ é testada utilizando uma estratégia de permutação não paramétrica (Excoffier et al. 1992).

Para os propósitos da análise todo o dataset foi incluído e categorizaram-se três regiões de acordo com a sua contingência geográfica e que compreende todos os fragmentos de floresta amostrados: Norte (Amidara, São Francisco e Ponta Nova), Centro (Lautchande e Cadique) e Sul (Farim, Catombói, Cafatche e Canghode). A significância foi testada a partir de um alfa de $p < 0.05$ com 16 000 permutações.

Baseado nos resultados da AMOVA, a maioria (56.24%) da diversidade genética dos chimpanzés de Cantanhez atribui-se a diferenças entre indivíduos dentro da população indicando assim uma significativa estruturação genética ($\Phi_{ST} = 0.43764$, $p < 0.0001$). A outra parte significativa de partição da diversidade genética (39.18%) dos chimpanzés de Cantanhez atribui-se à sua divisão regional ($\Phi_{CT} = 0.39177$, $p < 0.001$) indicando que a variância da diversidade genética também se encontra ao nível inter-regional (**Tabela 3.8**).

Tabela 3.8: Análise hierárquica de variância molecular (AMOVA) para as sequências HVR1 dos chimpanzés do Parque Nacional de Cantanhez, Guiné-Bissau.

Componente de variância		Variância	% Total	d.f.	SSD	P	Índices- Φ
Entre grupos	σ_a^2	4.25237	39.18	2	202.938	0.001	$\Phi_{CT} = 0.39177$
Entre populações dentro de grupos	σ_b^2	0.49786	4.59	6	56.504	0.049	$\Phi_{SC} = 0.07541$
Entre indivíduos dentro de subpopulações	σ_c^2	6.10391	56.24	57	341.819	<0.0001	$\Phi_{ST} = 0.43764$

3.6.4- Análise espacial da variância molecular (SAMOVA):

A estrutura genética dos chimpanzés de Cantanhez foi também explorada através da análise espacial da variância molecular (SAMOVA) (Dupanloup et al. 2002). Este método baseia-se num procedimento de simulações que tem como objectivo identificar grupos de populações que são geograficamente homogéneas e maximizar a diferença em termos inter-regionais (Φ_{CT}) da variância genética total sem a necessidade *a priori* de se definir uma composição de grupos necessária para a AMOVA. O método requer que se testem a existência de várias populações (K) e gera igualmente as estatísticas- Φ (Φ_{ST} , Φ_{SC} , Φ_{CT}) usando a abordagem da AMOVA (Excoffier et al. 1992). Ao explorar-se o padrão observado dos índices Φ_{CT} para diferentes valores de K, estes entram num 'plateau' antes da variância inter-regional se diluir e, neste sentido, a média mais elevada de Φ_{CT} é indicativa da estrutura mais provável. É assim possível identificar o número óptimo de grupos populacionais presentes num set de amostras.

O programa SAMOVA v1.0 (Dupanloup et al. 2002) aplicou-se a todo o dataset de forma a identificar o número de populações mais provável dos chimpanzés de Cantanhez através de 10 000 iterações para cada 100 condições iniciais aleatórias, testando o número de populações possíveis (K) entre K= 3 até K= 8. A melhor partição da diversidade genética pelo algoritmo SAMOVA obteve-se

quando as amostras foram agrupadas em quatro populações ($K=4$, $\Phi_{CT}=0.48178$; $\rho<0.001$) (Tabela 3.9).

Tabela 3.9: Índices de fixação correspondendo aos grupos de comunidades de chimpanzés do Parque Nacional de Cantanhez inferidos pelo algoritmo SAMOVA.

K	Composição dos grupos	Index- Φ
3	1. Lautchande, Amidara, São Francisco, Ponta Nova	$\Phi_{CT}=0.44270^*$
	2. Catombói, Farim, Cafatche, Canghode	$\Phi_{SC}=0.05967^{**}$
	3. Cadique	$\Phi_{ST}=0.47595^{**}$
4	1. Cadique	$\Phi_{CT}=0.48178^{**}$
	2. Catombói, Farim, Cafatche, Canghode (Sul)	
	3. Amidara, São Francisco, Ponta Nova (Norte)	
	4. Lautchande	
5	1. São Francisco	$\Phi_{CT}=0.49910^{**}$
	2. Cadique	
	3. Catombói, Farim, Cafatche, Canghode	
	4. Lautchande	
	5. Amidara, Ponta Nova	
6	1. Catombói, Farim, Canghode	$\Phi_{CT}=0.50794^{**}$
	2. Cadique	
	3. Cafatche	
	4. São Francisco	
	5. Lautchande	
	6. Amidara, Ponta Nova	
7	1. Cadique	$\Phi_{CT}=0.52298^*$
	2. Lautchande	
	3. Catombói, Farim, Canghode	
	4. Cafatche	
	5. Amidara	
	6. São Francisco	
	7. Ponta Nova	
8	1. Cadique	$\Phi_{CT}=0.56849^*$
	2. Lautchande	
	3. Amidara	
	4. Cafatche	
	5. Canghode	
	6. Catombói, Farim	
	7. São Francisco	
	8. Ponta Nova	

* $\rho<0.05$; ** $\rho<0.001$

3.7- Análise de isolamento por distância:

A análise dos F_{ST} linearizados de Slatkin (secção 3.6.2) indicou que existe algum nível de isolamento por distância. No sentido de determinar se a diferenciação das populações de chimpanzés de Cantanhez está correlacionada com a sua distância geográfica procedeu-se à análise formal de isolamento por distância. O isolamento por distância pode definir-se como um decréscimo da semelhança genética entre populações à medida que a distância geográfica entre elas aumenta (Wright 1943). A teoria do isolamento por distância, segundo o modelo *stepping-stone*, prediz que se a dispersão está espacialmente limitada, os indivíduos de populações que estão mais próximas ou são vizinhas têm mais probabilidade de interagir e, conseqüentemente, têm também uma maior probabilidade de acasalar, do que indivíduos de populações mais distantes. Como resultado as frequências de haplótipos nas populações que estão mais próximas serão mais semelhantes do que nas populações que estão mais afastadas (Ngan 2006).

O teste de Mantel e a regressão linear por redução do eixo maior (RMA) são os métodos estatísticos que determinam a correlação de isolamento por distância. Ao representar-se as distâncias genéticas e geográficas através de duas matrizes o teste de Mantel testa essa correlação obtendo-se a estatística Z e uma correlação r . O valor de Z corresponde à correlação entre uma matriz A e uma matriz B e r^2 pode variar de -1 a 1, onde 1 representa uma correção perfeita. A significância do teste de Mantel é testada usando permutações. No mesmo sentido a RMA é uma representação linear da correlação entre as distâncias genéticas e geográficas. Os cálculos da RMA incluem a inclinação e intercepção dessa relação. O valor da correlação, r^2 , representa a força da dispersão dos pontos em relação á linha. Se todos os pontos se associarem exactamente na linha é indicativo de que existe definitivamente uma correlação entre as distâncias genéticas e geográficas.

Utilizou-se a medida de distância genética Φ_{ST} entre as populações de chimpanzés em Cantanhez com a correcção Kimura 2-parâmetros (que tem em consideração diferentes taxas de transições e transversões) e correlacionou-se com a suas distâncias geográficas (medidas no Google Earth em km) implementado no

software IBDWS v.3.16. Observou-se um significativo isolamento por distância (Teste de Mantel; $Z=159.8819$, $r^2=0.5661$, unicaudal $p=0.0080$ a partir de 30 000 permutações). No mesmo sentido a regressão linear RMA também revelou uma associação de isolamento por distância positiva ($y=0.03709x-0.08010$, $r^2=0.320$) (Fig. 3.6).

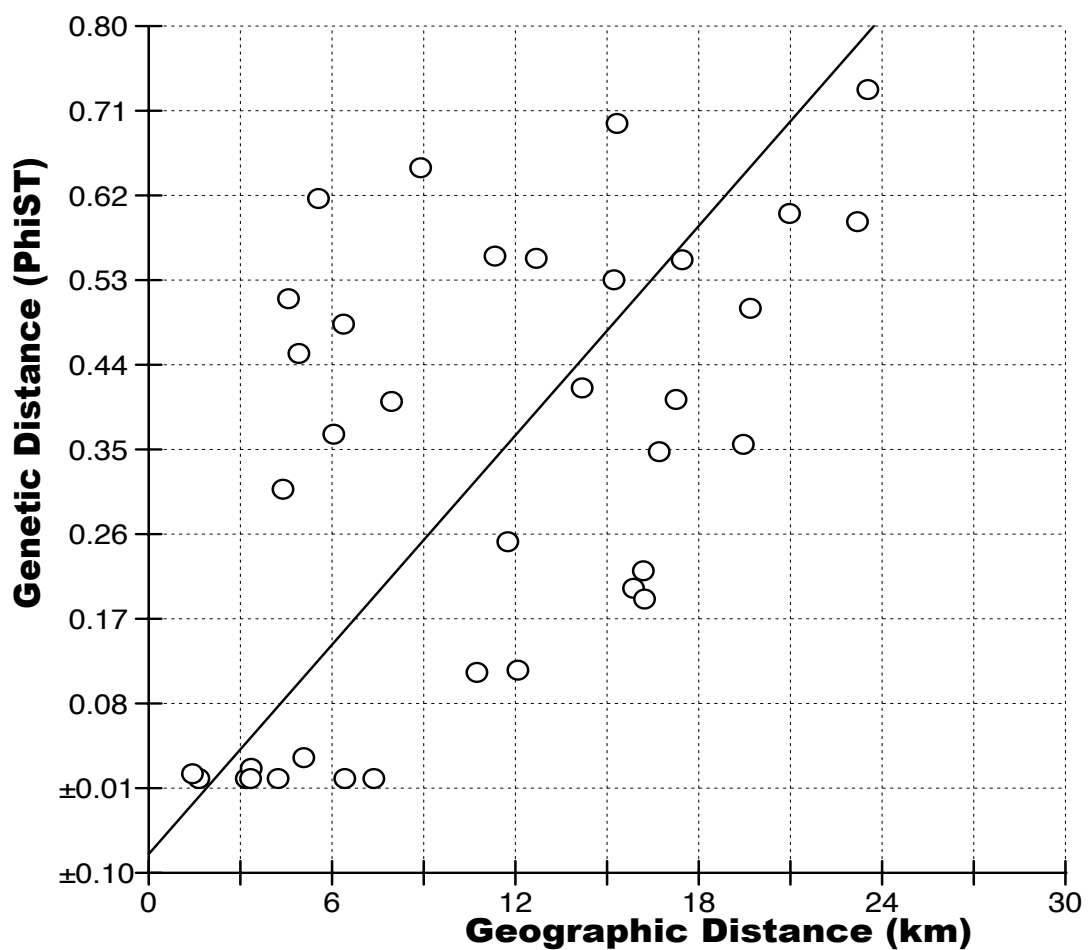


Figura 3.6: Isolamento por distância representado pela regressão linear por redução do eixo maior (RMA) entre a distância genética (Φ_{ST}) corrigida com o parâmetro Kimura 2-P e a distância geográfica (km) dos chimpanzés de Cantanhez.

3.8- Análise de barreiras genéticas:

Com o objectivo de fornecer uma representação geográfica das discontinuidades genéticas entre as populações de chimpanzés de Cantanhez procedeu-se à identificação das barreiras associadas a uma mudança genética significativa de acordo com o método proposto por Manni et al. (2004) que utiliza o algoritmo de diferença máxima de Monmonier implementado no software BARRIER. Este algoritmo permite identificar limites ou áreas onde as diferenças entre pares de populações são maiores existindo assim uma barreira entre elas. As barreiras são zonas onde ocorre uma mudança abrupta no padrão da variação genética devido a factores de isolamento de índole física, ecológica, geográfica, cultural ou morfológica. Uma das desvantagens deste método ocorre quando as populações se enquadram num modelo de isolamento por distância, na medida em que as possibilidades de se encontrarem discontinuidades no padrão da variação genética diminuem. Por outro lado, uma das suas vantagens é que a presença de factores de isolamento enfraquece o fluxo de genes fazendo aumentar as possibilidades de se encontrarem barreiras significativas. Regra geral, existem mais barreiras quando as associações entre as distâncias geográficas e genéticas são baixas e vice versa. No entanto, não quer dizer que o algoritmo de Monmonier não se possa aplicar quando se verifique um significativo isolamento por distância.

Este método utiliza também uma reconstrução geométrica que tem como base as coordenadas geográficas dos locais de amostragem criando uma rede que liga todas as populações amostradas adjacentes através de uma triangulação de Delaunay que se constrói a partir de uma tesselação de polígonos de Voronoi resultando na ligação das intersecções médias centrais de cada triângulo. As distâncias genéticas são calculadas para cada população vizinha e associadas a cada orla dessa rede.

Nesta análise usaram-se assim as distâncias genéticas F_{ST} (secção 3.6.1) como matriz de diferenciação genética entre as populações de chimpanzés de Cantanhez. O número de barreiras foi identificado recorrendo ao algoritmo de Monmonier através das orlas da rede associada com os valores mais elevados das distâncias

genéticas entre populações. Somente as barreiras que obtiveram valores mais elevados de diferenciação genética do que a média calculada para o total das amostras foram consideradas relevantes ($> \Phi_{ST} = 0.43764$).

Na **Fig. 3.7** pode observar-se que a principal barreira genética encontrada revelando uma descontinuidade genética que se situa em torno do grupo de Cadique e separa-o das restantes populações de chimpanzés em Cantanhez sugerindo uma barreira de longo termo, extrínseca ao intercâmbio genético das outras populações.

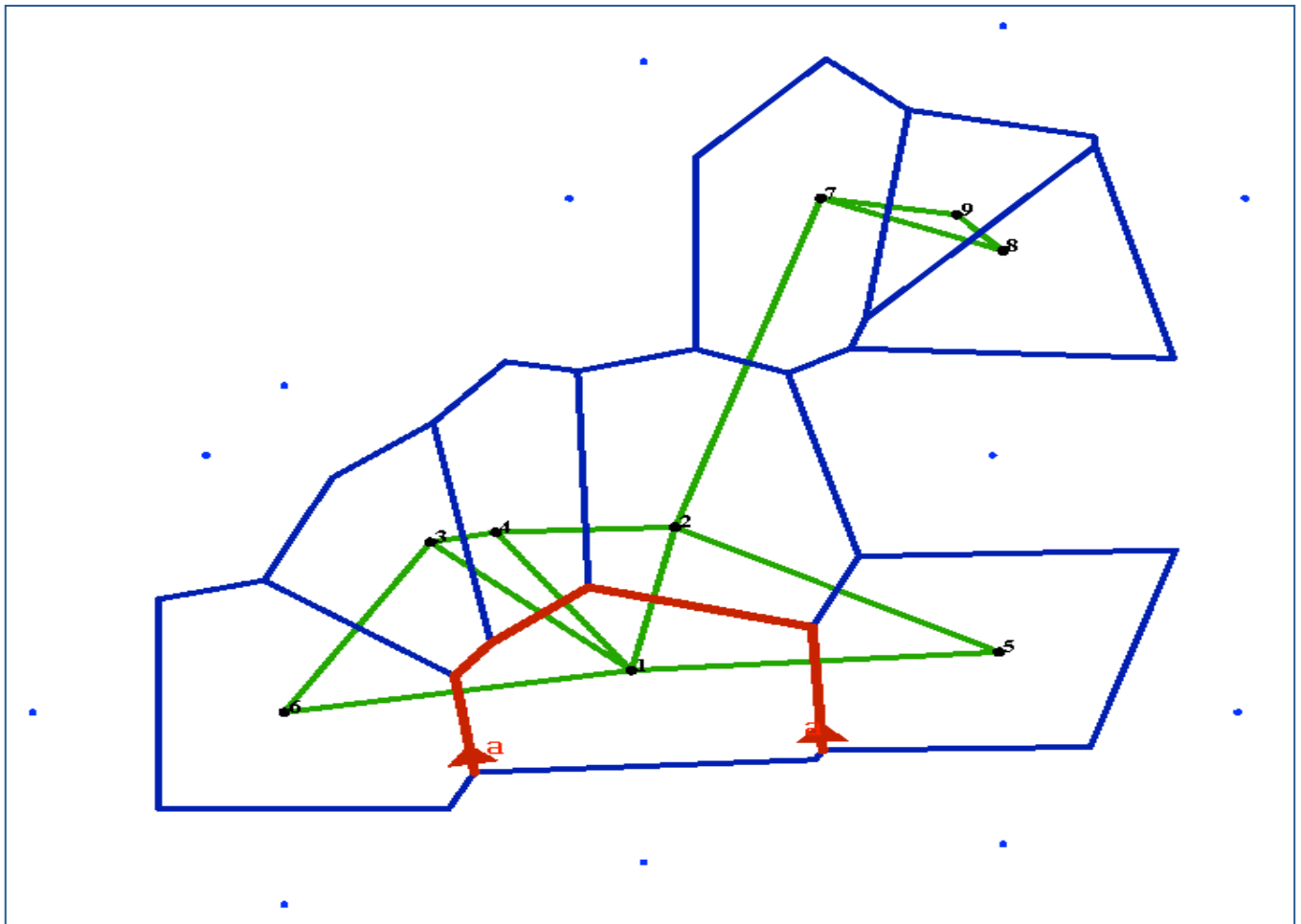


Figura 3.7: Mapa de descontinuidade genética nas subpopulações de chimpanzés de Cantanhez resultante do software BARRIER v2.2. Tesselação de Voronoi com os respectivos pontos virtuais de delimitação indicado a azul, triangulação Delaunay a verde, os números e pontos a negrito indicam os matos amostrados: 1-Cadique, 2-Lautchande, 3- Catomboi, 4-Farim, 5-Cafatche, 6-Canghode, 7-Amidara, 8- São Francisco e 9-Ponta Nova, respectivamente. Extensão da barreira genética *a* (indicada a vermelho) computada a partir da matriz de distâncias genéticas F_{ST} e da aplicação do algoritmo de Monmonier.

3.9- Análise com medidas complementares de genética da conservação:

Um dos objectivos da Genética da Conservação é usar a informação resultante de levantamentos de caracterização genética e implementá-la em medidas e políticas de conservação para uma determinada espécie identificando áreas prioritárias para a sua conservação (Millar e Libby 1991). Tendo como base este objectivo de conservação Petit et al. (1998) propuseram um conjunto de medidas adicionais de diversidade genética que permitem caracterizar as populações, medidas que segundo os autores: “estão bem adaptadas aos problemas da genética da conservação e que realçam populações individuais a quem devem ser dadas prioridades nos esforços da sua conservação” (Petit et al. 1998:845). Sugerem assim que, se incorpore a medida de riqueza haplotípica (i.e. alélica) (H_r) que permite medir o número de haplótipos por *locus* estando grandemente dependente do tamanho efectivo da população, sendo por isso um bom indicador das alterações demográficas ocorridas no passado histórico e por isso importantes para a conservação das populações de espécies ameaçadas. No entanto, uma das desvantagens da utilização desta medida é que está dependente do tamanho da amostragem para cada população. Nesse sentido, propuseram ainda que se utilizasse o método de rarefacção a partir do tamanho de amostragem populacional mais pequeno permitindo assim, normalizar a desigualdade da amostragem. Propõem também que, é possível determinar qual o contributo que cada população dá para a diversidade haplotípica total (CT) ou para a sua riqueza haplotípica total (CRT). Esse contributo pode ser estimado a partir da diferença entre a diversidade haplotípica total de todas as populações excepto da população em análise. Consequentemente advogam ainda que, essa contribuição individual de cada população pode ser subdivida ou particionada em dois componentes: uma relacionada com o seu próprio nível de diversidade (CS ou CRS) e o outro com o seu nível de unicidade, ou seja, com o seu nível de divergência das outras populações (CD ou CRD) para CT ou CRT, respectivamente (Petit et al. 1998). Algumas contribuições para a diversidade podem ser negativas, porque a diversidade que a

própria população pode ser menor do que a média da diversidade total ou porque essa população pode não ser muito diferenciada.

Com o objectivo de perceber que populações de chimpanzés mais contribuem para a diversidade genética em Cantanhez utilizaram-se estas medidas sugeridas por Petit et al. (1998) determinando-se assim a riqueza haplotípica (Hr) para as quatro subpopulações, definidas pela análise SAMOVA (secção 3.6.4) e avaliou-se o contributo que cada uma dá para a sua diversidade e riqueza haplotípicas totais calculadas no software CONTRIB v.1.02 usando o método de rarefacção, normalizado para o tamanho de amostragem mais pequeno ($n=8$, Cadique).

Aplicando o método de rarefacção para o menor número de amostras numa população, a riqueza haplotípica (Hr) dos chimpanzés de Cantanhez, é de: 5.6482 haplótipos para o HVR1. A riqueza haplotípica para cada população de chimpanzés de Cantanhez apresenta-se na **Tabela 3.10**, bem como as suas contribuições para a diversidade haplotípica total (CT) e riqueza haplotípica total (CRT) e os respectivos componentes de variação. Pode assim observar-se que a população do Norte (Amidara, São Francisco e Ponta Nova) apresenta a maior riqueza haplotípica ($Hr=4.250$). As populações de Cadique e do Norte são as que mais contribuem (CT) para a diversidade haplotípica total. Cadique pela sua diferenciação em relação às outras populações (CD) e a população do Norte pela sua própria diversidade (CS) (**Figura 3.8**). Por outro lado as populações que mais contribuem (CRT) para a riqueza haplotípica total são as populações de Lautchande, tanto pela sua divergência como pela sua diversidade (CRS e CRD), e a população do Norte pela sua diversidade (CRS) (**Figura 3.9**)

Tabela 3.10: Medidas complementares de genética de conservação (diversidade genética e diferenciação) para cada uma das populações de chimpanzés de Cantanhez.

População	Total	<i>Hr</i>	<i>CT</i>	<i>CS</i>	<i>CD</i>	<i>CRT</i>	<i>CRS</i>	<i>CRD</i>
Hap								
Cadique	3	2.000	0.000	-0.082	0.083	-0.039	-0.089	0.050
Lautchande	6	3.855	0.008	0.036	-0.029	0.075	0.044	0.031
Norte	8	4.250	-0.003	0.060	-0.064	0.036	0.072	-0.036
Sul	8	2.863	-0.005	-0.015	0.010	0.034	-0.027	0.061

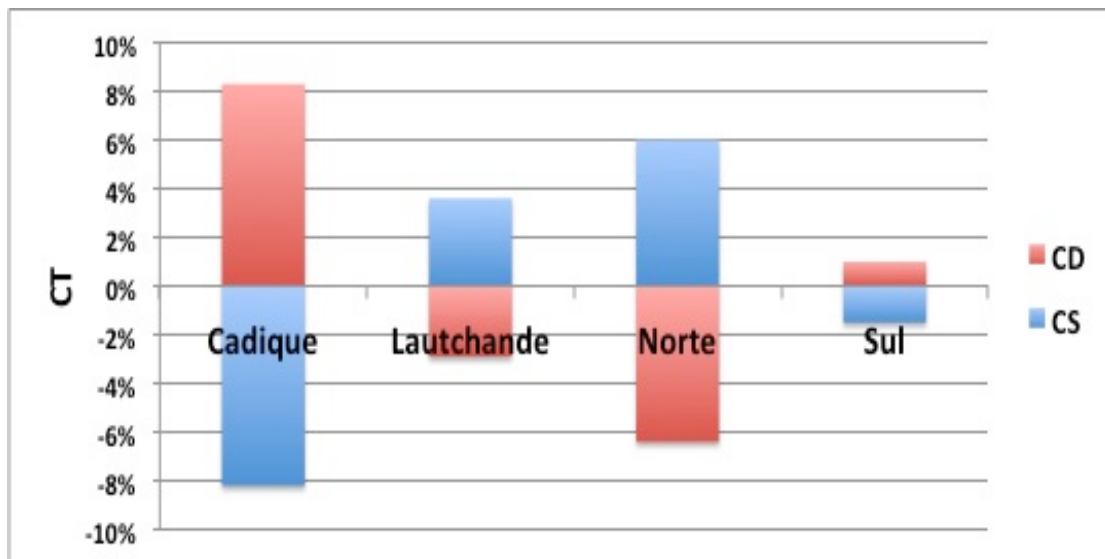


Figura 3.8: Contribuição (CT) para a diversidade haplotípica total (*h*) de cada uma das populações de chimpanzés em Cantanhez, subdividida em componentes: diversidade própria (CS) e diferenciação com as outras populações (CD). As populações que mais contribuem para a diversidade genética total são a população de Cadique devido à sua divergência com as outras populações e a população do Norte devido à sua própria diversidade.

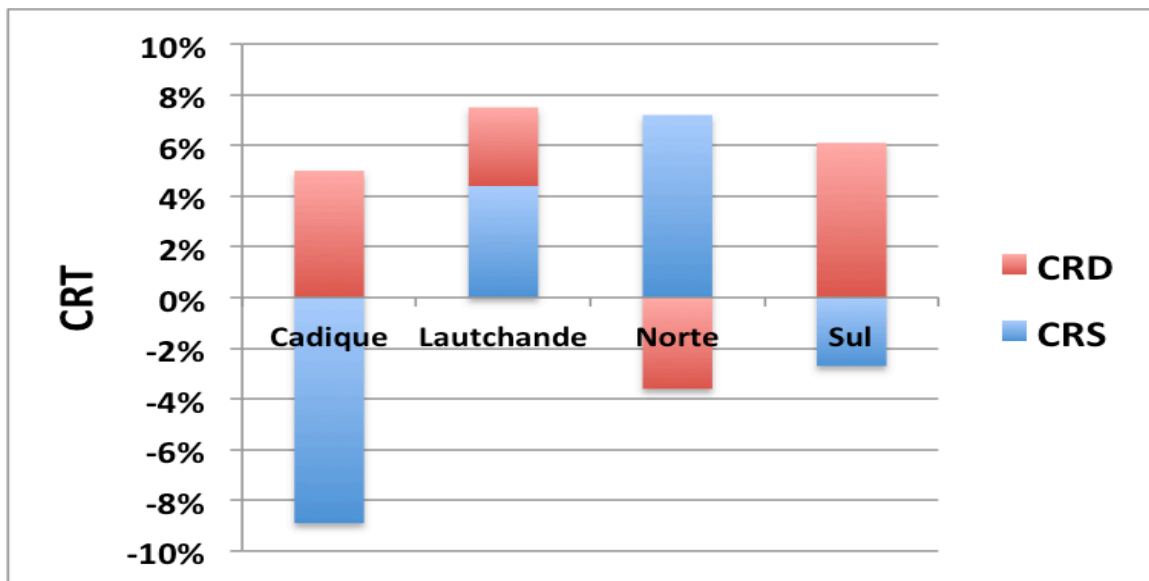


Figura 3.9: Contribuição (CRT) para a riqueza haplotípica total (H_r) de cada uma das populações de chimpanzés em Cantanhez, subdividida em componentes: diversidade própria (CRS) e diferenciação com as outras populações (CRD). As populações que mais contribuem para a riqueza haplotípica total são a população de Lautchande devido tanto à sua divergência com as outras populações, mas também à sua diversidade própria e a população do Norte devido à sua própria diversidade.

3.10.- Análise da divergência temporal:

Com o objectivo de perceber quais os tempos cronológicos de separação intraespecífica dos haplogrupos dos chimpanzés de Cantanhez e em que cenário paleontológico ocorreram procedeu-se à análise da divergência temporal dos mesmos no software BEAST v.1.6.1 (Drummond and Rambaut 2007). Este software integra um procedimento de Cadeias Markov Monte Carlo (MCMC) e uma inferência filogenética bayesiana utilizando ainda um relógio molecular “relaxed” (Drummond et al. 2006).

Considerando uma taxa de mutação (substituição) por nucleótido de 2.5% por milhão de anos (Gonder et al. 2006) e uma divergência entre humanos e chimpanzés de ~7-5 mya (milhões de anos) e entre chimpanzés e bonobos de 2.1-1.5 mya (Stone et al. 2010), como elementos de calibração, as duas linhagens mitocondriais HVR1 principais (*vide* secção 3.3) encontradas nos chimpanzés de Cantanhez (Linhagem A e B) divergiram há 1 mya [intervalo 95% HPD (High Posterior Density) de 0.34- 2.11 mya]. O antepassado comum mais recente (TMRCA) da linhagem A estima-se em 0.67 mya, enquanto que o TMRCA da linhagem B se estima em 0.43 mya (95% HPD: 0.34- 2.11 mya). O TMRCA entre o haplogrupo A1 (H13, H9, H8) e o haplogrupo A2 estima-se em 0.47 mya (95% HPD: 0.9- 1.22 mya), enquanto que o haplogrupo A3 (H15, H1, H4) divergiu dos anteriores há 0.32 mya (95% HPD: 0.03- 0.95 mya). O haplogrupo B1 (H2, H16, H12, H6, H11, H3) divergiu há 0.29 mya (95% HPD: 0.04- 0.81 mya) e o haplogrupo B2 (H10, H7) há 0.07 mya (95% HPD: 0- 0.47 mya).

Considerando a sequência de bonobo como outgroup a separação dos chimpanzés de Cantanhez (chimpanzés ocidentais) ocorreu há 1.41 mya (95% HPD: 0.44- 3.24 mya) (**Fig. 3.10**).

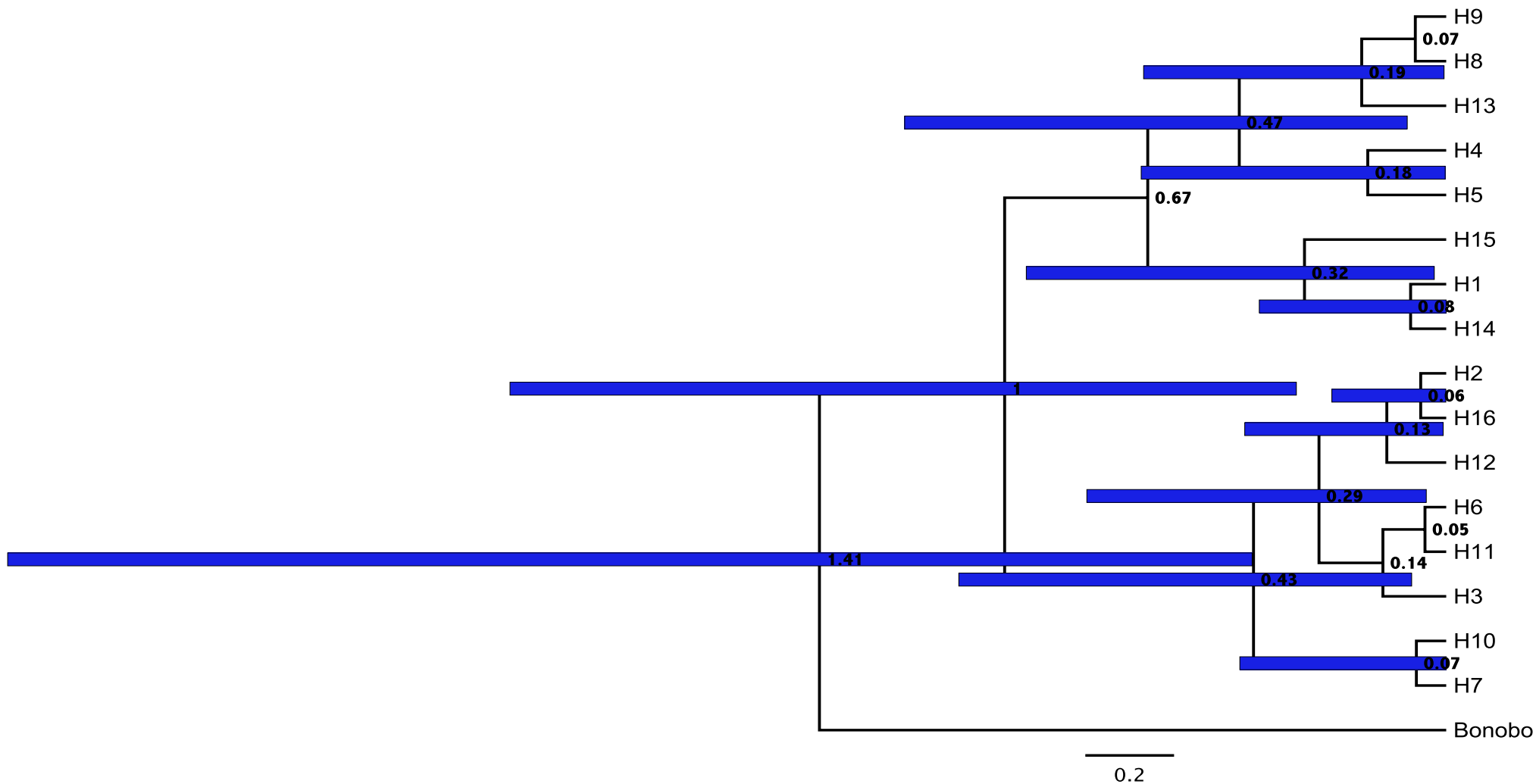


Figura 3.10:Árvore de credibilidade de clades máximas para os haplótipos HVR1 dos chimpanzés de Cantanhez em milhões de anos (mya). O intervalo de confiança 95% HPD está indicado a azul.

3.11.- Sexagem molecular e análise do rácio sexual:

A sexagem molecular é um procedimento forense que permite identificar o sexo dos indivíduos a partir da variação molecular entre os dois sexos (Cotton et al. 2000). É também um instrumento importante para a ecologia molecular e para a genética da conservação (Villessen e Fredsted 2006). Com o objectivo de estimar qual o rácio sexual de todo o dataset (n=65) da população de chimpanzés de Cantanhez procedeu-se assim à sua sexagem molecular individual, através do sistema de amelogenina desenvolvido por Sullivan et al. (1993) e optimizado por Bradley et al. (2001) em amostras não invasivas para grandes símios.

Uma das vantagens deste método é que pode ser aplicado tanto num gel de agarose como através de automatização capilar já que os alelos do cromossoma X são 6 pares de bases mais curtos que os do cromossoma Y nos chimpanzés (X:104 pb; Y:110 pb) (Bradley et al. 2001). Como nos mamíferos o sexo masculino é heterogâmico espera-se encontrar sempre 2 bandas ou dois picos e somente um para o sexo feminino.

O rácio sexual geral entre machos e fêmeas é de 2.1 (44:21), inclinado em relação aos machos (68% machos) e não é significativamente diferente das frequências esperadas ($\chi^2 = 2.409$, $df=1$, $N=65$, $p=0.1206$).

4. DISCUSSÃO

Este estudo é o primeiro a caracterizar geneticamente os chimpanzés da República da Guiné-Bissau e a contribuir em particular, para o conhecimento sobre a história evolutiva dos chimpanzés do Parque Nacional de Cantanhez.

Enquadra-se no âmbito da Primatologia Molecular (Vigilant e Guschanski 2009) e o método forense não-invasivo utilizado provou ser adequado para estabelecer informação de base relevante, que pode por isso ser aplicada na conservação da espécie neste país, indo assim ao encontro com a necessidade de monitorização identificada no Plano Nacional de Acção para os cólobos e chimpanzés da Guiné-Bissau (Casanova e Sousa 2007).

O elevado sucesso de extracção de ADN a partir de excrementos deste estudo (90%) é superior ao reportado por Shimada et al. (2004) que apresenta uma taxa de 50% nas amostras de Boussou e Nimba da Guiné-Conacry e as principais diferenças atribuem-se aos métodos de preservação e extracção das amostras utilizados (Goossens e Bruford 2009).

A partir da análise de um segmento de 482pb da região HVR1 mitocondrial, encontraram-se 50 sítios polimórficos e definiram-se 16 haplótipos. Destes últimos: 44% são partilhados e 56% são exclusivos de determinado mato. H2 é o haplótipo mais frequente (28%) ocorrendo em 17 indivíduos do número mínimo de chimpanzés definido com este estudo para Cantanhez (N=65) estando distribuído por 7 dos 9 matos amostrados, sendo os matos da zona norte do parque (Ponta Nova e São Francisco) as excepções. O mato de Cafatche é aquele que evidencia uma maior presença de haplótipos exclusivos (n=2).

Os valores observados das diversidades haplotípica, nucleotídica e de Watterson bem como o número médio de diferenças deste estudo: $h=0.879\pm 0.024$; $\pi= 0.037\pm 0.001$; $\theta_w=0.021$ e $k= 17.8$, respectivamente, são moderadamente altos. Os valores de diversidade genética deste estudo estão na mesma ordem de grandeza que os reportados para os macacos de Gibraltar (*Macaca sylvanus*, $h= 0.872$ e

$\pi=0.026$) (Modolo et al. 2005); babuíños hamadrias (*Papio hamadryas hamadryas*, $h= 0.856$ e $\pi=0.023$) (Winney et al. 2004) e macacos dourados (*Rhinopithecus roxellana*, $h= 0.872$ e $\pi=0.026$) (Li et al. 2007). Não se pode por isso afirmar que a variação genética detectada nos chimpanzés de Cantanhez seja baixa e que tenham pouca diversidade genética. Quando se tem em consideração somente a diversidade nucleotídica, os chimpanzés de Cantanhez comparam-se aos chimpanzés de Boussou e Nimba ($\pi=0.038$) (Shimada et al. 2004) e estão no intervalo de diversidade nucleotídica reportada por Gagneux et al. (1999) (*Pan troglodytes verus*, $\pi=0.07$, intervalo 0-16%) para a totalidade da subespécie, mas acima da diversidade estimada para os chimpanzés orientais (*Pan troglodytes schweinfurthii*, $\pi=0.025$, intervalo 0-9%) (Goldberg 1998) e humanos (*Homo sapiens*, $\pi=0.026$) [Yu et al. 2003]. Não obstante, quando comparados com os totais de diversidade nucleotídica HVR1 apresentados para os chimpanzés centrais (*Pan troglodytes troglodytes*, $\pi=0.056$) (Yu et al. 2003) e bonobos (*Pan paniscus*, $\pi=0.046$) (Eriksson et al. 2004) estão claramente abaixo e alguns autores argumentam que os chimpanzés ocidentais detêm uma diversidade nucleotídica relativamente baixa mas diferenciada das outras subespécies (Fischer et al. 2006).

A elevada variabilidade da região HVR1, como marcador molecular, permite resolver mais claramente as relações entre subespécies (Avice 2000). Usando os métodos de máxima parcimónia, neighbor joining e inferência bayesiana, construíram-se árvores filogenéticas e ainda, uma network através do algoritmo de median joining a partir da distância entre pares de diferenças. Esta reconstrução filogenética revelou a existência de duas linhagens profundas e cinco haplogrupos nos chimpanzés de Cantanhez, bem suportada estatisticamente em todos os métodos. Demonstrou também que estas duas linhagens não correspondem a nenhum padrão geográfico óbvio e os haplótipos dos matos do norte ao sul de Cantanhez estão praticamente distribuídos pelas duas linhagens. A existência destas duas linhagens já tinha sido anteriormente reportada por Morin et al. 1994, Eriksson et al. 2004 e Gonder et al. 2006. Morin e colegas (1994) advogam que as sequências dos chimpanzés ocidentais (*P. t. verus*) são mais divergentes que as sequências das outras subespécies e que a distância genética média entre os chimpanzés ocidentais

e as outras duas é dois-terços a distância genética entre chimpanzés e bonobos sugerindo a elevação à categoria de espécie dos chimpanzés ocidentais (Morin et al. 1994). Por outro lado Gonder et al. (2006) provou que os chimpanzés da Nigéria e Camarões *Pan troglodytes ellioti* formam uma clade monofilética independente dos restantes chimpanzés ocidentais na Alta Guiné.

A diferenciação genética e o fluxo de genes foram analisados utilizando índices F_{ST} . O número médio de pares de diferenças calculado a partir da frequência haplotípica indicam que existe uma significativa diferenciação genética entre os pares: Cadique e Lautchande ($F_{ST}=0.319$), Amidara e Cadique ($F_{ST}=0.204$) e Cadique e Cafatche ($F_{ST}=0.451$). O índice F_{ST} linearizado de Slatkin demonstra que não existe diferenciação entre os matos do sul de Cantanhez ($D=0$) e que os eventos coalescentes entre os seus haplótipos são muito recentes sugerindo que os chimpanzés dessa área podem considerar-se uma só subpopulação. O número de migrantes entre Canghode e São Francisco é muito limitado ($M=0.18651$) e que Cadique e Lautchande partilham uma 1.1 fêmea por geração.

Segundo Gagneux (1998), os factores que influenciam a estrutura genética das populações de chimpanzés são: 1)- Barreiras geográficas ou físicas (ex: rios, montanhas, desertos, etc.); 2)- Sistema de acasalamento e estrutura social; 3)- Ameaças antropogénicas e; 4)- História biogeográfica e paleoclimática. A estrutura genética dos chimpanzés de Cantanhez foi inferida pela AMOVA que mostra que a variação mitocondrial é claramente afectada pela estrutura populacional, já que 39% da variação genética está particionada entre regiões. No entanto, a maior parte da variação (56%) ocorre entre indivíduos dentro das subpopulações, como seria de esperar já que nos chimpanzés são as fêmeas o sexo que dispersa e o fluxo de genes é por isso mediado por elas, causando assim poucos níveis de diferenciação entre populações (Morin et al. 1994). Este facto fez com que o padrão filogeográfico da estrutura populacional não tenha sido claro nas reconstruções filogenéticas (árvores e network). O mesmo padrão de variância genética foi encontrado para os bonobos da República Democrática do Congo (Eriksson et al. 2004).

A análise espacial da variância molecular (SAMOVA) indicou que o melhor agrupamento significativo de subpopulações de chimpanzés em Cantanhez é de K=4, constituídos pelas subpopulações de Cadique; Lautchande; Norte (Amidara, Ponta Nova e São Francisco) e Sul (Catombói, Farim, Cafatche e Canghode) ($\Phi_{CT}=0.48178$, $\rho<0.001$).

A estrutura genética actual dos chimpanzés de Cantanhez pode predizer-se tanto por um padrão significativo de isolamento por distância como por vicarância. Os testes de Mantel ($Z=159.8819$, $r^2=0.5661$, $\rho<0.05$) e a RMA ($y=-0.08010x+0.03709$, $r^2=0.320$) demonstram que existe uma correlação positiva entre a distância genética e a distância geográfica. À medida que a última aumenta a primeira diminui e esta relação causa assim uma divergência entre as frequências haplotípicas dos diferentes matos da população contínua de Cantanhez criando por deriva genética um padrão de isolamento por distância. Explica também que as fêmeas chimpanzés de Cantanhez preferem migrar para uma subpopulação vizinha, do que fazer grandes distâncias seguindo assim, um modelo *stepping-stone* de dispersão (Kimura e Weiss 1964). Este padrão também ocorre em outras populações de chimpanzés noutros locais em África (Gonder et al. 2000; Fischer et al. 2004).

A análise de descontinuidades genéticas conduzidas no software BARRIER, que combina fundamentos de geometria estatística (Tesselação de Voronoi e triangulação de Delaunay) com o algoritmo de máxima diferença de Monmonier revelou surpreendentemente a existência de um elemento vicariante (vulgo barreira). Esta descoberta é curiosa porque, não seria de esperar encontrar-se uma barreira quando existe um padrão confirmado de isolamento por distância e vice-versa. No entanto, se por um lado o padrão de isolamento por distância enfraquece o fluxo de genes, faz aumentar também as possibilidades de se encontrarem barreiras significativas em que o programa sendo sensível a essas alterações, a partir da matriz de distâncias F_{ST} , consegue detectar essas descontinuidades revelando qual a sua localidade mais provável (Manni et al 2004). Neste estudo detectou-se uma barreira actual que causa uma descontinuidade genética na subpopulação de Cadique em relação às outras subpopulações de Cantanhez. A existência desta barreira é deveras intrigante e não é consistente com nenhum factor físico ou

ecológico evidente. Não existe nenhum curso de água ou outra característica orográfica que crie uma barreira física e isole a subpopulação de Cadique, e em termos ecológicos o habitat e os efeitos antropogénicos são os mesmos que nas restantes áreas do PNC. Existe no entanto uma estrada que liga Iemberém ao antigo porto de Cadique no rio Cumbijã, mas os chimpanzés atravessam actualmente muito facilmente essa estrada (Sá 2010, observação pessoal). A densidade humana também não parece ser um factor para que essa barreira exista, já que a localização das aldeias e o uso da terra pelas comunidades humanas são semelhantes ao resto do Parque. Avançam-se três hipóteses para a existência desta barreira em torno da comunidade de chimpanzés de Cadique. (1) A maior área do PNC está contida na pensínsula do Cubucaré, delimitada a Oeste pelo rio Cacine e a Este pelo rio Cumbijã. Devido à dimensão, profundidade, e não sazonalidade do volume de água, destes dois rios, eles são indubitavelmente fronteiras naturais à dispersão dos chimpanzés de Cantanhez. A paisagem de Cantanhez caracteriza-se por ser um mosaico composto por vários elementos e entre eles encontra-se uma extensa zona de mangal associada a uma elevação do solo muito baixa (Sousa et al. 2011). É possível que essas zonas de mangal tenham alterado o seu curso ao longo do tempo devido à flutuação paleoclimática e que assim, tenham criado durante um determinado período de tempo, uma descontinuidade no fluxo genético dos chimpanzés de Cadique mas que posteriormente tenha sido possível transpor e estes voltaram a estar em contacto com as restantes subpopulações. (2) Uma segunda hipótese pode ter a ver com o efeito da guerra colonial e posteriormente guerras civis que o país sofreu e cujo efeito tenha criado esta barreira. A estrada que liga Iemberém a Cadique foi construída na década de 50 do século passado. Durante o período mais intenso da guerra, espera-se que o tráfego militar fosse também mais intenso. No mesmo sentido, os libertadores da imposição colonial portuguesa, do Partido Africano para a Independência da Guiné e Cabo Verde (PAIGC) usavam os matos adjacentes de Caiquéne e Cibe Cadique como refúgio e área operacional e, várias vezes a estrada sofreu emboscadas (Sá 2010, observação pessoal baseada em informações etnográficas e testemunhos recolhidos). É possível que a estrada tenha servido como barreira física e os chimpanzés só agora se arrisquem a atravessá-la

devido à sua deterioração e pouco tráfego motorizado. No entanto, destaque-se que desde a sua construção a estrada tem sido constantemente usada por pessoas. Não é por isso de estranhar a proximidade de contacto entre humanos e chimpanzés nesse estrada (Sá 2010, observação pessoal). (3) Uma terceira hipótese pode dever-se a uma particularidade comportamental ou cultural dos chimpanzés de Cadique que, por algum motivo não deixam que as suas fêmeas migrem mais frequentemente para os grupos adjacentes. São por isso necessários mais estudos que permitam averiguar as hipóteses aqui avançadas.

A história demográfica dos chimpanzés de Cantanhez é complexa. As análises filogenéticas não demonstraram nenhum padrão filogeográfico em forma de “estrela” característico de acumulação de mutações devido a uma expansão demográfica recente (Avise 2000). Uma explicação para a falta deste padrão filogeográfico claro nos chimpanzés de Cantanhez pode relacionar-se com o facto de que os haplogrupos estejam ainda a sofrer um processo de “*incomplete lineage sorting*” e que ainda não passou tempo suficiente desde a sua divergência para permitir o estabelecimento desse padrão (Maddison e Knowles 2006). Tanto a diversidade haplotípica como a nucleotídica são altas (88% e 3,7% respectivamente) e de acordo com Grant e Bowen (1998), isto reflete um contacto secundário entre diferentes linhagens ou uma população larga e estável com uma história evolutiva longa. Os *mismatch distribution* demonstram que: i)- a frequência dos pares de diferenças entre haplótipos revela um padrão multimodal, associado a um tamanho efectivo da população constante e rejeita significativamente a hipótese do modelo de expansão demográfica súbita e; ii)- um padrão bimodal, segundo o modelo de expansão demográfica espacial, que pode ser marginalmente rejeitado, mas que sugere: subdivisão populacional, várias expansões demográficas no passado, várias ondas de migração entre diferentes subpopulações ou uma combinação de ambos (Rogers and Jorde 1995). Os testes de neutralidade apresentaram valores positivos e significativos (Tajima $D = 2.634$, $p < 0.05$; Fu $F = 9.099$, $p \leq 0.001$; Fu e Li $F^* = 2.256$) à excepção do teste Fu e Li D^* que apresenta um valor positivo mas não significativo (1.429, $p > 0.05$). Estes testes indicam um desvio da neutralidade e a hipótese nula de uma população constante é rejeitada sugerindo um bottleneck recente ou algum

tipo de selecção balanceadora (Schmidt e Pool 2002). A análise da história demográfica dos chimpanzés de Cantanhez sugere que a população se manteve gradual ao longo do tempo e que não sofreu flutuações no seu tamanho por forma a causar um padrão detectável de expansão ou bottleneck na distribuição da variação. Assemelha-se ao padrão referido já para os chimpanzés ocidentais e centrais (Gagneux 1998), da Nigéria e Camarões (Gonder 2000) e bonobos (Eriksson et al. 2004), mas difere do padrão unimodal dos chimpanzés orientais e humanos que sofreram claramente uma expansão demográfica (Rogers e Harpending 1992; Harpending et al. 1998; Goldberg e Ruvolo 1997).

A estimativa dos tempos de divergência é sensível não só à escolha do dataset molecular mas também aos modelos usados para descrever a estrutura das populações de chimpanzés (Stone et al. 2010) e não é por isso de estranhar que diferentes métodos apresentem diferentes datações. Este estudo assumiu um relógio molecular “relaxed” e tempos de divergência entre humanos e chimpanzés de ~7- 5 mya (milhões de anos) e entre chimpanzés e bonobos de 2.1- 1.5 mya (Stone et al. 2010). A divergência estimada para chimpanzés e bonobos deste estudo é de 1.41 milhões de anos (0.44-3.24 milhões de anos) durante o Plio-Pleistoceno, mais recente que a proposta por Stone et al. (2010) de 1.94 Mya (1.49- 2.09 Mya), mas consistente com o intervalo proposto por Yu et al. (2003) de 0.93-1.8 Mya e com a datação de 1.60 Mya de Weigman e Excoffier (2010). No entanto, quando comparada com as estimativas feitas a partir do genoma nuclear é bastante superior (e.g: 0.79-0.92 Mya) (Becquet et al. 2007). Os bonobos são endémicos da República Democrática do Congo, e a formação do rio Congo há 1.5-2 Mya, no Plio-Pleistoceno, avança-se como hipótese que determinou a divergência com os chimpanzés (que só ocorrem e estão distribuídos a norte do rio) (Caswell et al. 2008) A datação deste estudo está em conformidade também com esta predição.

Assumiu-se a divergência dos chimpanzés ocidentais com os centrais/orientais datada entre 1.6-1.1.3 Mya (Morin et al. 1994). As duas linhagens observadas nos chimpanzés de Cantanhez surgiram segundo este estudo, há 1 Mya (2.11- 0.34) após essa separação. Enquadram-se historicamente no Pleistoceno Inferior (1-0.78 Mya). A divergência dos haplogrupos dentro de cada linhagem

remonta ao Pleistocénio Médio (0.78-0.12 Mya), de 0.47-0.43 Mya (1.22-0.09; 1.1-0.08 Mya,) para os haplogrupos da linhagens A e B, respectivamente. Esta observação insere-se em termos de paleoclima durante as glaciações de Gunz (0.68-0.62 Mya) e de Mindel (0.45-0.30 Mya). Existem evidências que sugerem que nos períodos de máxima glaciação, o clima em África era mais frio e seco do que actualmente (Maley 1996). Esta possibilidade transformou a floresta tropical que sofreu expansões durante os períodos interglaciais e contracções durante os glaciares dando origem a savanas (Cowling et al. 2008). Foi muito provavelmente este o cenário evolutivo e biogeográfico dos chimpanzés de Cantanhez já que a diversidade genética que hoje detêm está intimamente relacionada com a relação histórica das florestas de onde são nativos (Gonder and Disotell 2006).

A teoria da genética da conservação postula que a manutenção da diversidade genética é crucial tanto a longo prazo (manutenção do potencial evolutivo adaptativo) como no curto prazo (manutenção do *fitness* reprodutivo) (Frankham et al. 2005). Neste estudo determinou-se que as subpopulações de Cadique e do Norte de Cantanhez são as que mais contribuem para a diversidade haplotípica total dos chimpanzés de Cantanhez e as subpopulação de Lautchande e do Norte aquelas que mais contribuem para a sua riqueza haplotípica total. Cadique pela sua diferenciação. Podem por isso, ser percepcionadas como unidades operacionais de conservação.

Um rácio sexual desequilibrado reduz o tamanho efectivo da população (N_e) (Frankham et al. 2005). Um dos resultados mais impressionantes e inesperado deste estudo é o desequilíbrio do rácio sexual total de 2.1 [68% machos (n=44) e 32% fêmeas (n=21)] em relação às fêmeas, havendo claramente um desvio do equilíbrio esperado de 1:1. Quando comparado com estudos de longo prazo de outras comunidades de chimpanzés em África, o rácio sexual apresentado neste estudo é o mais elevado reportado até agora. Em Bossou na vizinha República da Guiné o rácio sexual é de 0.33 (2.69:8.22) (Shimada et al. 2009), nas florestas de Taï na Costa do Marfim é de 0.38 (16.10:36.30) (Boesch e Boesch-Achermann 2000), em Mahale de 0.44 (16.10:36.30) e em Gombe de 0.95 (18.57:19.57) (Goodall 1986). Se se considerar o rácio sexual de forma independente do grau de parentesco os

chimpanzés machos de Cantanhez são de todos os que apresentam uma maior dificuldade em encontrar uma parceira para acasalar e pode também deduzir-se consequentemente que a competição de machos por fêmeas deverá ser intensa. A estratégia de amostragem deste estudo, tal como outras anteriores em grandes símios (Vigilant e Guschanski 2009) baseou-se nas recomendações estratégicas de Lowe et al. (2004) sobre os procedimentos amostrais para um levantamento filogeográfico ou *survey* genético e por conseguinte maximizou a amostragem por “nest site” em vez de amostragem intensiva de todas as comunidades e teve como objectivo amostrar a representação da diversidade genética dos chimpanzés de Cantanhez. A possibilidade de efeito enviesado por presença de contaminações humanas não se pode colocar, já que de acordo com o método de sexagem molecular pelo sistema de amelogenina vulgarmente utilizado pelo sistema CODIS em casos forenses, a sexagem de chimpanzés e humanos é específica. Nos chimpanzés é de: X: 104 e Y: 110, enquanto que nos humanos é de X:106 e Y: 112 havendo deste modo uma diferença de 2pb facilmente detectáveis se saírem do range esperado para os chimpanzés (Bradley et al. 2001).

Uma hipótese, pouco provável, é que os machos defequem mais que as fêmeas após deixarem os ninhos desse “nest site” e ter-se-iam recolhido mais amostras de machos devido a essa possibilidade. Outra hipótese, mais provável decorre de um efeito antropogénico. Apesar de estarem protegidos a nível nacional existe tráfico de mascotes (em inglês: *pet trade*) em Cantanhez e nesse sentido, para os caçadores conseguirem geralmente obter um chimpanzé infantil ou juvenil têm de matar as suas mães (Sá 2008, informação pessoal). Poderá haver assim uma forma de pressão antropogénica negativa em relação às fêmeas que aliada ao facto dos chimpanzés de Cantanhez se situarem num extremo da distribuição poderá estar a criar o padrão de rácio sexual observado. Sabe-se que as populações marginais são mais susceptíveis aos efeitos da fragmentação e assim mais vulneráveis à extinção (Torres et al. 2010; Oates 2006).

São necessários mais estudos que permitam comparar os padrões filogeográficos observados para os chimpanzés de Cantanhez, com uma investigação de cobertura nacional, conferindo assim um conhecimento mais abrangente sobre a

sua história evolutiva, comparando-a com a das outras regiões. A utilização de marcadores nucleares altamente polimórficos como por exemplo os microsateélites, permitirão também inferir sobre a sua estrutura genética, parentesco, consaguinidade e dinâmicas sociais, complementando também o presente estudo.

5. CONCLUSÕES

Os estudos da variação genética em primatas não humanos, em particular de chimpanzés e bonobos, no âmbito da Primatologia Molecular ajudam a esclarecer não só as suas histórias evolutivas, mas também a compara-las com a própria história evolutiva humana (Di Fiori and Gagneux 2007; Stone et al. 2010).

Este estudo determinou que o número mínimo de chimpanzés de Cantanhez é de 65 indivíduos distribuídos por quatro subpopulações sendo detentores de uma diversidade genética elevada. Revela que os seus haplogrupos derivam filogeneticamente de duas linhagens profundas e que sua variação mitocondrial é afectada pela estrutura populacional (sendo 39% particionada entre regiões) e que se pode explicar tanto por um padrão de isolamento por distância como pela existência de uma barreira vicariante. A história demográfica dos chimpanzés de Cantanhez é complexa e sugere subdivisão populacional, várias expansões demográficas no passado, várias ondas de migração entre diferentes subpopulações ou uma combinação de ambos, mas não apresenta sinais significativos de alterações drásticas ao tamanho efectivo da população (expansões ou bottlenecks). Os principais eventos de divergência ocorreram durante o Pleistocénio Médio e enquadram-se paleoclimaticamente durante os períodos de glaciação de Gunz e Mindel. Surpreendentemente, o rácio sexual dos chimpanzés está claramente desequilibrado (44:21) havendo mais machos que fêmeas e esta evidência sugere que de todos os estudos de longa duração, os chimpanzés de Cantanhez são os que apresentam a maior dificuldade de encontrar uma parceira para acasalar, deduzindo-se também que haja uma maior competição entre machos.

Implicações para a conservação dos chimpanzés de Cantanhez

A conservação dos chimpanzés é um tópico relevante e muito actual tanto ao nível internacional (Grasp 2005; Kormos et al. 2003) como ao nível nacional (IBAP 2010). Vários esforços estão a ser conduzidos por autoridades nacionais da República da Guiné-Bissau (eg. Instituto da Biodiversidade e das Áreas Protegidas e Direcção Geral de Florestas e Fauna), mas também por ONG's locais (e.g. Acção para o Desenvolvimento) e equipas multidisciplinares de investigadores (e.g: Projecto Dári) no sentido de preservar a viabilidade dos chimpanzés de Cantanhez. Este estudo teve como missão contribuir para a conservação destes chimpanzés indo ao encontro do postulado de Avise (2000) que diz que a estrutura genética histórica revelada por marcadores filogeneticamente informativos, como é o caso do HVR1 mitocondrial é muito relevante para se definirem estratégias de conservação, já que nas separações intraespecíficas profundas se registam também as maiores fontes evolutivas de diversidade genética. E, neste sentido, de acordo com o modelo de conservação de Moritz (1994a) podem assim definir-se: “unidades evolutivas significantes”- (ESUs- *Evolutionary Significant Units*) e “unidades de gestão da conservação” (MUs- *Management Units*). Segundo este autor as ESUs devem ser reciprocamente monofiléticas ao nível do ADN mitocondrial, enquanto que as Mus identificam-se pelas diferenças significativas na distribuição da frequência alélica e uma divergência significativa tanto nos *loci* mitocondriais como nucleares. A maior parte da variação genética dos chimpanzés de Cantanhez ocorre entre indivíduos dentro das subpopulações (56.24%) e nesse sentido, a criação de uma grande reserva protegida como é o caso do Parque Nacional de Cantanhez em 2008 foi uma boa solução e deverá ser suficiente para manter a a sua variação genética, se se conseguir obviamente minimizar a pressão antropogénica actual: fragmentação, perda de habitat, caça e uso insustentável dos recursos naturais. Recomenda-se também que se criem quatro unidades operacionais de gestão de conservação para os chimpanzés de Cantanhez correspondentes com os quatro clusters genéticos mais prováveis: Lautchande e a área Norte (Amidara, Ponta Nova e São Francisco) pelo contributo que estas subpopulações dão para a diversidade e riqueza haplotípica

total da população, Cadique pela grande diferenciação relativamente às restantes e, a subpopulação do Sul (Canghode, Cafatche, Farim e Catomboi) pelo número de haplótipos raros que possuem. Especial atenção deve ser dada a esta última subpopulação, já que se encontra no final da sua distribuição na península de Cantanhez e a conectividade entre as suas florestas com as restantes deve ser uma prioridade sob pena de ficar isolada por efeitos de destruição de habitat. No mesmo sentido, advoga-se especial atenção para a população de Cadique devido à sua área reduzida de habitat. Recomenda-se que sejam feitas monitorizações regulares a estas unidades operacionais e que seja criado um plano efectivo de seguimento e observação no sentido de ajudar a perceber quais os constrangimentos a que estas subpopulações estão sujeitas ao longo do ano, bem como incentivar a pesquisa sociobiológica da comunidade no sentido de inferir mais sobre as suas dinâmicas comportamentais. Não se defende contudo a sua habituação já que esta poderia facilitar ainda mais o conflito entre humanos e chimpanzés. Este estudo é por isso um contributo para a sua conservação, desejando que a população de chimpanzés mais ocidental do continente africano se mantenha estável, longa e duradoura.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Arnason, U., Xu, X., and Gullberg, A. 1996. Comparison between the complete mitochondrial DNA sequences of Homo and the common chimpanzee based on nonchimeric sequences. *J Mol Evol* 42:145–152.
- Avise, J. 1995. Mitochondrial DNA polymorphism and a connection between genetics and demography of relevance to conservation. *Conservation Biology* 9:686-690.
- Avise, J. 1998. The history and purview of phylogeography: a personal reflection. *Molecular Ecology* 7:371-379.
- Avise, J. 2000. *Phylogeography. The history and formation of species*. London: Harvard University Press.
- Avise, J. 2004. *Molecular Markers, Natural History and Evolution*, Second Edition edition. Sunderland, Massachusetts: Sinaeur Associates.
- Avise, J., and J. Hamrick. 1996. *Conservation genetics: case histories from nature*. New York: Chapman and Hall.
- Awadalla, P., Eyre-Walker, A., Maynard-Smith, J. 1999 Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science*. 286, 2524–2525.
- Bandelt, H., Forster, P., Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37-48
- Bécquet, C., Patterson, N., Stone, A., Przeworski, M., and Reich, D. 2007. Genetic structure of chimpanzee populations. *PLOS Genetics* 3(4): e66.
- Beebe, T., and Rowe, G. 2004. *An introduction to Molecular Ecology*. Oxford University Press.
- Bennet, A., Lesch, K., Heils, L., Lorenz, J., Shoaf, S., Champoux, M., Suomi, S., Linnoila, M., and Higley, D. 2002. Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Molecular Psychiatry* 7:118-122.
- Bensasson, D., Zhang, D., Hartl, D., Hewitt, G. 2001. Mitochondrial pseudogenes:

- evolution's misplaced witnesses. *Trends in Ecology and Evolution* 16:314-321.
- BLAST (Basic Local Alignment Search Tool). 2010. "<http://blast.ncbi.nlm.nih.gov/Blast.cgi>". [Acedido: 10/02/2011].
- Boesch, C., and H. Boesch-Achermann. 2000. *The Chimpanzees of the Tai Forest: Behavioural Ecology and Evolution* Oxford University Press.
- Blom, A., and Feistner, A. 1992. Primates of Gabon – current status and distribution. *Oryx* 26:223-234.
- Bradley, B. J., Chambers, K. E. and Vigilant, L. 2001. Accurate DNA-based sex identification of apes using non-invasive samples. *Conservation Genetics* 2: 179-181.
- Brownell, A. 2003a. Benin. In *West African Chimpanzees: Status Survey and Conservation Action Plan*. Edited by R. Kormos, C. Boesch, M. Bakarr, and T. Butynski, pp. 121-122: IUCN, Gland, Switzerland.
- Brownell, A. 2003b. Burkina Faso. In *West African Chimpanzees : Status Survey and Conservation Action Plan* Edited by R. Kormos, C. Boesch, M. Bakarr, and T. Butynski, pp. 121-122: IUCN, Gland. Switzerland.
- Buettner-Janush, J., Twichell, J., Wong, B., and Wagenen, G. 1961. Multiple haemoglobins and transferrins in a macaque sibship. *Nature* 192:948-950.
- Butynski, T. 2003a. The chimpanzee *Pan troglodytes*: taxonomy, distribution, abundance, and conservation status. In *The chimpanzee Pan troglodytes: taxonomy, distribution, abundance, and conservation status*. Edited by R. Kormos, C. Boesch, M. Bakarr, and T. Butynski, pp. 5-12: IUCN, Gland, Switzerland.
- Butynski, T. 2003b. The robust chimpanzee (*Pan troglodytes*: taxonomy, distribution, abundance and conservation status. In *West african chimpanzees: status survey and conservation action plan*. Edited by R. Kormos, C. Boesch, M. Bakarr, and T. Butynski: IUCN, Gland, Switzerland.
- Caldecott, J., and L. Miles. 2005. *World Atlas of Great Apes and their Conservation*: London. University of California Press.

- Casanova, C. 2006. *Introdução à Antropologia Biológica: Princípios Evolutivos, Genética e Primatologia*. Lisboa: Instituto Superior de Ciências Sociais e Políticas. Universidade Técnica de Lisboa.
- Casanova, C., and C. Sousa. 2007. *Plano de Acção Nacional para a conservação das populações de chimpanzés, cólobos vermelhos ocidentais e cólobos brancos e pretos ocidentais na República da Guiné-Bissau*. IBAP. Ministério do Desenvolvimento Rural e Agricultura, Recursos Naturais e Ambiente.
- Cassamá, V. 2006. *Alterações do coberto do solo na Mata de Cantanhez (Guiné-Bissau)*. Dissertação de Mestrado. Instituto Superior Técnico. Universidade Técnica de Lisboa.
- Caswell, J. L., S. Mallick, D. J. Richter, J. Neubauer, C. Schirmer, S. Gnerre, and D. Reich. 2008. Analysis of chimpanzee history based on genome sequence alignments. *PLoS Genet* 4:e1000057.
- Cockerham, C. 1969. Variance of gene frequencies. *Evolution* 23: 72-84.
- Constable, J., Asley, M., Goodall, J. and Pusey, A. 2001. Noninvasive paternity assignment in Gombe chimpanzees. *Molecular Ecology* 10:1279-1300.
- Coote, T., and M. Bruford. 1996. Human microsatellites applicable for analysis of genetic variation in apes and old world monkeys. *The Journal of Heredity* 87:406-410.
- Cowling, A., Cox, P., Jones, C., Maslin, M., Peros, M., and Spall, S. 2008. Simulated glacial and interglacial vegetation across Africa: implications for species phylogenies and trans-African migration of plants and animals. *Global Change Biology* 14:827-840.
- Dalén, L., Gotherstrom, A., and Angerbjorn, A. 2004. Identifying species from pieces of faeces. *Conservation Genetics* 5:109-111.
- Darwin, C. 1871. *The Descent of Man and Selection in Relation to Sex*. London. John Murray.

- Deagle, B., Tollit, D., Jarman, S., Hindel, M., Trites, A., and Gales, N. 2005. Molecular scatology as a tool to study diet of prey DNA in scats from captive Steller sea lions. *Molecular Ecology* 14:1831-1842.
- Di Fiori, A. 2003. Molecular genetic approaches to the study of primate behavior, social organization, and reproduction. *Yearbook of Physical Anthropology* 46:62-99.
- Di Fiori, A., and P. Gagneux. 2007. Molecular Primatology. In *Primates in Perspective*. Edited by C. Campbell, A. Fuentes, K. Mackinnon, M. Panger, and S. Bearder. London: Oxford University Press.
- Drummond, A. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.
- Duggleby, C. 1977. Blood group antigens and the population genetics of *Macaca mulatta* on Cayo Santiago. Effects of social group division. *Yearbook of Physical Anthropology* 20:263-271.
- Dupanloup, I., Schneider, S., Excoffier, L. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11(12): 2571-81.
- Eggert, L., Eggert, J., and Woodruff, D. 2003. Estimating population size for elusive animals: the forest elephants of Kakum National Park, Ghana. *Molecular Ecology* 12:1389-1402.
- Eriksson, J., Hohmann, G., Boesch, C., and Vigilant, L.. 2004. Rivers influence the population genetic structure of bonobos (*Pan paniscus*). *Molecular Ecology* 13:3425–3435.
- Excoffier, L., Smouse, P., E., and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
- Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

- Eyre-Walker, A. Smith, N., H., and Maynard-Smith, J. 1999 How clonal are human mitochondria? *Proceedings of the Royal Society B-Biological Sciences* 266 477–483.
- Fairbanks, L. 2001. Individual differences in response to a stranger: social impulsivity as a dimension of temperament in vervet monkeys (*Chlorocebus aethiops sabaeus*). *Journal of Complementary Psychology* 115:22-28.
- Fischer, A., Wiebe, V., Pääbo, S., and Przeworski, M. 2004. Evidence for a Complex Demographic History of Chimpanzees. *Molecular Biology and Evolution* 21:799-808.
- Fischer, A., Pollack, J., Thalmann, O., Nickel, B., and Pääbo, S. 2006. Demographic History and Genetic Differentiation in Apes. *Current Biology* 16:1133-1138.
- Frankham, R., Ballou, J., and Briscoe, D. 2005. *Introduction to Conservation Genetics*: Cambridge. Cambridge University Press.
- Fu, Y. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915-925.
- Fu, Y., and Li, W. 1993. Statistical tests of neutrality of mutations. *Genetics* 133: 693-709.
- Gagneux, P. 1998. Population Genetics of west African Chimpanzees (*Pan troglodytes verus*). PhD thesis. Basel University. Switzerland.
- Gagneux, P. 2002. The genus *Pan*: population genetics of an endangered outgroup. *TRENDS in Genetics* 18:327-330.
- Gagneux, P., Boesch C., and Woodruff, D. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* 6:861-868.
- Gagneux, P., Boesch, C., and Woodruff, D. 1999. Female reproductive strategies, paternity and community structures in wild West African chimpanzees. *Animal Behaviour* 57:19-32.
- Gagneux, P., Gonder, M., Goldberg, T., and Morin, P. 2001. Gene flow in wild chimpanzee populations: what genetic data tell us about chimpanzee movement over space and time. *Phil. Trans. R. Soc. Lond. B* 356:889-897.

- Galtier, N., Enard, D., Radondy, Y., Bazin, E., Belkhir, K. 2006. Mutation hotspots in mammalian mitochondrial DNA. *Genome Research* 16:215-222.
- Galtier, N., Nabholz, B., Glémin, S. and Hurst, G. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* 18:4541-4550.
- GenBank, 2010:"<http://www.ncbi.nlm.nih.gov>" [Acedido: 29/08/2010].
- Gerloff, U., Hartung, B., Fruth, B., Hohmann, G., and Tautz, D. 1999. Intracommunity relationships, dispersal pattern, and paternity success in a wild living community of bonobos (*Pan paniscus*) determined from DNA analysis of fecal samples. *Proceedings of the Royal Society of London [Biology]* 266:1189-1195.
- Gerloff, U., Schlötterer, C., Rassmann, K. *et al.* 1995. Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Molecular Ecology* 4: 515–518.
- Gippoliti, S., Embalo, D., and Sousa, C. 2003. Guinea-Bissau. In *West African chimpanzees: Status Survey and Conservation Action Plan*. Edited by R. Kormos, C. Boesch, M. Bakarr, and T. Butynski, pp. 55-61: IUCN, Gland, Switzerland.
- Goldberg, T. 1998. Biogeographic predictors of genetic diversity in populations of eastern African chimpanzees (*Pan troglodytes schweinfurthii*). *International Journal of Primatology* 19:237-254.
- Goldberg, T., L., and Ruvolo, M. 1997. Molecular phylogenetics and historical biogeography of east African chimpanzees. *Biological Journal of the Linnean Society* 61:301-324.
- Goldberg, T. L., and Wrangham, R. 1997. Genetic correlates of social behavior in wild chimpanzees: evidence from mitochondrial DNA. *Animal behavior* 54:559-570.
- Gonder, K. 2000. Evolutionary Genetics of Chimpanzees (*Pan troglodytes*) in Nigeria and Cameroon. PhD thesis. The City University of New York.
- Gonder, K., and Disotell, T. 2006. Contrasting Phylogeographic Histories of Chimpanzees in Nigeria and Cameroon: A Multi-Locus Genetic Analysis. In

- Primate Biogeography. Progress and Prospects*. Edited by S. Lehman and J. Fleagle, pp. 135-168: Springer.
- Gonder, K., Disotell, T., and Oates, J. 2006. New genetic evidence on the evolution of chimpanzee populations, and implications for taxonomy. *International Journal of Primatology* 27:56-68.
- Gonder, K., Oates, J., Disotell, T., Forstner, T., Morales, J., and Melnick, D.. 1997. A new west african subspecies? *Nature* 388:337.
- Gómez-Moliner, B., Cabria, M., Rubines, J., Garin, I., Madeira, M., Elejalde, A., Aihartza, J., Fournier, P., and Palazón, S. 2004. PCR-RFLP identification of mustelid species: European mink (*Mustella lutreola*), American mink (*Mustella vison*) and polecat (*Mustella putorius*) by analysis of excremental DNA. *Journal of Zoology of London* 262:311-316.
- Goodall, J. 1986. *The Chimpanzees of Gombe*. Cambridge, Massachussets. Harvard University Press.
- Goodman, M. 1962. Immunochemistry of the primates and primate evolution. *Annales of the New York Academy of Sciences* 102:219-234.
- Goodman, M. 1999. The genomic records of humankind's evolutionary roots. *American Journal of Human Genetics* 64:31-39.
- Goossens, B., and Bruford, M. W. 2009. Non-invasive genetic analyses. in conservation. In: Bertorelle, G., Bruford, M., Hauffe, H., Rizzoli, A., and Vernesi, C. (eds.) *Population genetics for animal conservation*. Cambridge. Cambridge University Press. pp. 167-201.
- Goossens, B., Anthony, N., Kathryn, J., Johnson-Bawe, M., and Bruford, M. 2003. Collection, storage and analysis of non-invasive genetic material in primate biology. In: Setchell, J. and Curtis, D. (eds.). *Field and Laboratory Methods in Primatology*. Cambridge: Cambridge University Press. pp. 295-308.
- Goossens, B., Latour, S., Vidal, S., Jamart, A., Ancrenaz, M., and Bruford, M., W. 2000. Twenty new microsatellite loci for use with hair and faecal samples in the chimpanzee (*Pan troglodytes troglodytes*). *Folia Primatologica* 71:177.

- Grant, W., S., and Bowen, B., W. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Heredity* 89:415–426.
- GRASP. 2005. Global strategy for the survival of great apes and their habitat. UNEP.
- Groves, C. P. 2001. *Primate Taxonomy*. Washington, DC. Smithsonian Institution Press.
- Grossman, L., Wildman, D., Schmidt., T. and Goodman., M. 2004. Accelerated evolution of the electron transport chain in anthropoid primates. *Trends in Genetics* 20: 578-585.
- Grubb, P., Butynski, T., Oates, J., Bearder, S., Disotell,, T., Groves, T., and Sthruhsaker, T. 2003. Assessment of African primate diversity. *International Journal of Primatology* 24:1301–1357.
- Grzybowski, T. 2000. Extremely high levels of human mitochondrial DNA heteroplasmy in single hair roots. *Electrophoresis* 21: 548-553.
- Hagelberg, E., Goldman, N., Lió, P. et al. 1999. Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proceedings of the Royal Society B-Biological Sciences*. 266, 485–492.
- Hagelberg, E., Goldman, N., Lió, P. et al. 2000. Evidence for mitochondrial DNA recombination in a human population of island Melanesia: correction. *Proceedings of the Royal Society B-Biological Sciences*. 267, 1595–1596.
- Haig, S. 1998. Molecular contribution to conservation. *Ecology* 79:413-425.
- Harpending, H. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* 66: 591-600.
- Harpending, H., Sherry, S. ,Rogers, A., and Stoneking, M.1993. The genetic structure of ancient human populations. *Curr. Anthropol.* 34:483–496.
- Harpending, H., Batzer, M., Gurven, M., Jorde, L., and Rogers, A. 1998. Genetic traces of ancient demography. *Proc. Natl. Acad. Sci. USA* 95: 1961–1967.
- Hasegawa, M., Kishino, H., and Yano, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22:160-174.

- Hedmark, E., Flagstad, O., Segerstrom, P., Persson, L., Landa, A., and Ellegren, H. 2004. DNA-based individual and sex identification from wolverine (*Gulo gulo*) faeces and urine. *Conservation Genetics* 5:405-410.
- Higuchi, R., Von Berholdingen, R., Sensabaugh, G., and Erlich, H. 1988. DNA typing from single hairs. *Nature* 332:543-546.
- Houlden, B., Woodworth, L., and Humphrey, K. 1997. Captive breeding, paternity determination, and genetic variation in chimpanzees (*Pan troglodytes*) in the Australasian Region. *Primates* 38:341-347.
- Huson, D., Regula, R., and Scornavacca, C. 2010. *Phylogenetic network: concepts, algorithms and applications*. Cambridge. Cambridge University Press.
- IBAP. 2008. Estratégia nacional para as áreas protegidas e a conservação da biodiversidade na Guiné-Bissau 2007-2011. Bissau. República da Guiné-Bissau.
- IBAP. 2010. "<http://www.ibap-gbissau.org/>." [Acedido: 27/10/2010].
- Idaghdour, Y., Broderick, D., and Korrida, A. 2003. Faeces as a source of DNA for molecular studies in a threatened population of great bustards. *Conservation Genetics* 4:789-792.
- Idani, G. 1991. Social relationships between immigrant and resident bonobo (*Pan paniscus*) females at wamba. *Folia primatologica* 57:83-95.
- IUCN/SSC. 2008. "<http://www.primate-sg.org/>." [Acedido: 24/03/2008].
- Inskipp, T. 2005. Chimpanzee (*Pan troglodytes*). In *World Atlas of Great Apes and their Conservation*,. Edited by J. Caldecott and L. Miles. London. University of California Press.
- Jensen, J., Bohonak, J., and Kelley, S. 2005. Isolation by distance, web service. *BMC Genetics* 6: 13. v.3.16. Available at: <http://ibdws.sdsu.edu/>.
- Kimura, M. And Weiss, G. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49: 561-576.
- Kivisild, T., Bamshad, M., Kaldma, K., Metspalu, M., Metspalu, E., Reidla, M., Laos, S., Parik, J., Watkins, W.S., Dixon, M. E., Papiha, S.S., Mastana, S.S., Mir, M.R., Ferak, V., Villems, R. 1999. Deep common ancestry of Indian and western Eurasian mitochondrial DNA lineages. *Current Biology* 9: 1331-1334.

- Kohn, M., and Wayne, R. 1997. Facts from feces revisited. *Trends in Ecology and Evolution* 12: 223-227.
- Kormos, R., and C. Boesch. 2003. *Regional Action Plan for the Conservation of Chimpanzees in West Africa*: IUCN.
- Leakey, R. 2005. "Introducing Great Apes" in *Atlas of Great Apes and their Conservation*. Edited by J. Caldecott and L. Miles. London. University of California Press.
- Lee, P.C., Thornback, J. and Bennett, E. 1988. Threatened Primates of Africa: The IUCN Red Data Book, ed. IUCN. Gland, Switzerland.
- Lesch, K., Bengel, D., Heils, A., Sabol, S., Greenberg, B., Petri, S., Benjamin, J., Muller, C., Hamer, D., and Murphy, D. 1996. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 271:1527-1531.
- Li, M., Liu, Z., Gou, J., Ren, B., Pan, R., Su, R., Funk, S., Wei, F. 2007. Phylogeography and population structure of the golden monkeys (*Rhinopithecus roxellana*): inferred from mitochondrial DNA sequences. *American Journal of Primatology* 69 (11): 1195-1209.
- Linnaeus, C. 1758. *Systema Naturae*. Stockholm: Laurentius Galvius.
- Loeschcke, V., Tomiuk, J., and Jain, S. 1994. *Conservation Genetics*. Basel: Birkhäuser Verlag.
- Lowe, A. et al. 2004. *Ecological genetics: design, analysis and application*. London. Blackwell.
- Lutz, S., Weisser, H., Heizmann, J., and Pollak, S. 1998. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. *International Journal of Legal Medicine*. 112, 145-150.
- Lynch, M., and Ritland, K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152:1753-1766.
- Maddison, W., and Knowles, L. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55:21-30.

- Maley, J. 1996. The african rain forest-main characteristics of change in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh* 140:31-73.
- Manni, F., Guérard, E., and Heyer, E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by “Monmonier’s algorithm”. *Human Biology* 76(2): 173-190.
- Margulis, L. 1992. *Symbiosis in Cell Evolution: Microbial communities in the Archean and Proterozoic eons*. W. H. Freeman & Company.
- Marjoram, P., and Donnelly, P. 1994. Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics* 136:673-683.
- McGrew, W., Ensminger, A., Marchant, L. Pruetz, J. and Vigilant, L. 2004. Genotyping aids field study of unhabituated wild chimpanzees. *American Journal of Primatology* 63:87-93.
- Meier, C., Hemelrijk, C., and Martin, R. 2000. Paternity determination, genetic characterization, and Social Correlates in a Captive Group of chimpanzees (*Pan troglodytes*). *Primates* 41:175-183.
- Melnick, D., Hoelzer, G., and Honeycutt, R. 1992. Mitochondrial DNA: its uses in anthropological research. In *Molecular applications in Biological Anthropology*. Edited by E. Dover, pp. 179-233. Cambridge: Cambridge University Press.
- Millar, C., I., and Libby, W., J. 1991. Strategies for conservation of clinal, ecotypic and disjunct population diversity in widespread species. Editors: Falk, D. A. and Holsinger, K. In: *Conservation of rare plants: Biology and genetics*. Oxford University Press, New York. pp: 149-170.
- Mitani, M. 1990. A note on the present situation of the primate fauna found from south-eastern Cameroon to northern Congo. *Primates* 31:625-634.
- Mitani, J. 2006. Demographic Influences on the behavior of chimpanzees. *Primates* 47:6-13.
- MITOMAP. 2011. “<http://www.mitomap.org/MITOMAP>”. [Acedido: 20/04/2011].

- Modolo, L., Salzburger, W., and Martin, R. 2005. Phylogeography of Barbary macaques (*Macaca sylvanus*) and the origin of the Gibraltar colony. *Proceedings of the National Academy of Sciences of the United States of America* 102:7392–7397.
- Morin, P., and Woodruff, D. 1992. Paternity exclusion using multiple hypervariable microsatellite loci amplified from nuclear DNA of hair cells. In *Paternity in Primates: Genetic Tests and Theories*. Edited by R. Martin, A. Dixson, and E. Wickings. Basel: Karger. pp. 63-81.
- Morin, P., Wallis, J., Moore, J., Chakraborty, R., and Woodruff, D. 1993. Non-invasive sampling and DNA amplification for paternity exclusion, community structure, and Phylogeography in wild chimpanzees. *Primates* 34:347-356.
- Morin, P., Moore, J., Chakraborty, R., Jin, L., Goodall, J., and Woodruff, D. 1994a. Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* 265:1193-1201.
- Morin, P., Wallis, J., Moore, J., and Woodruff, D. 1994b. Paternity exclusion in a community of wild chimpanzees. *Molecular Ecology* 3:469-478.
- Morin, P., K. Chambers, E., Boesch, C., and Vigilant, L. 2001. Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology* 10:1835–1844.
- Moritz, C. 1994a. Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution* 9:373-375.
- Moritz, C., Dowlin, T., and Brown, W. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*. 18: 269-292.
- Morovvati, S., Modarresi, M., Habibi, G., Kiarudi, Y., Karami, A., and Peyvandi, A. 2007. Sequence Analysis of Mitochondrial DNA Hypervariable Regions: An Approach to Personal Identification. *Archives of Medical Research*. 38, 345-349.
- Newton-Fisher, N. 2002. Ranging patterns of male chimpanzees in the Budongo Forest, Uganda: range structure and individual differences In *New*

- Perspectives in Primate Evolution and Behaviour*. Edited by C. Harcourt and B. Sherwood. London: Westbury Academic & Scientific Publishing. pp. 287-308.
- Newton-Fisher, N. 2003. The home range of the Sonso community of chimpanzees from the Budongo Forest, Uganda. *Afr. J. Ecol.* 41:150-156.
- Newton-Fisher, N. 2004. Hierarchy and social status in Budongo chimpanzees. *Primates* 45:81-87.
- Newton-Fisher, N., Reynolds, V., and Plumptre, A. 2000. Food supply and chimpanzee (*Pan troglodytes schweinfurthii*) in the Budongo Forest Reserve, Uganda. *Int. J. Primatol.* 21:615-628.
- Ngan, E. 2006. *Isolation by distance web service with implementation of DNA data sets*. MSc thesis. San Diego State University.
- Nishida, T. 1979. *The chimpanzees of the Mahale Mountains*. Tokyo. Tokyo University Press.
- Nishida, T., Corp, N., Hamai, M., Hasegawa, T., Hiraiwa-Hasegawa, M., Hosaka, K., Hunt, K., Itoh, D., Kanawaka, K., Matsumoto-Oda, A., Mitani, J., Nakamura, M., Norikosshi, K., Sakamaki, T., Turner, L., Uehara, S., and Zamma, K. 2003. Demography, female life history, and reproductive profiles among the chimpanzees of Mahale. *American Journal of Primatology* 59:99-121.
- Nuttal, G. 1904. *Blood Immunity and Blood Relationship*. Cambridge. Cambridge University Press.
- Oates, J.F. 2006. Is the chimpanzee, *Pan troglodytes*, an endangered species? It depends on what “endangered” means. *Primates* 47: 102–112.
- Oates, J., Gadsby, L., and Jenkins, P. 2003. Nigeria. In *West African Chimpanzees. Status survey and Conservation Action Plan*. Edited by R. Kormos, C. Boesch, and M. Bakarr, pp. 123-130. Gland, Switzerland: IUCN/SSC Primate Specialist Group.
- Oates, J., Groves, C., and Jenkins, P. 2009. The type locality of *Pan troglodytes vellerosus* (Gray, 1862), and implications for the nomenclature of West African chimpanzees. *Primates* 50(1): 78-80.

- Pamilo, P., and Crozier, R. 1982. Measuring genetic relatedness in natural populations: methodology. *Theoretical Populations Biology* 21:171-193.
- Palomares, F., Godoi, J., Piriz, A., O' Brien, S., and Johnson, W. 2002. Fecal genetic analysis to determinate the presence and distribution of elusive carnivores: design and feasibility for the Iberian lynx. *Molecular Ecology* 11:2171-2182.
- Parker, P., Snow, A., Schung, M., Booton, B., and Fuerst, P. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* 79:361-382.
- Patterson, N., Richter, D., Gnerre, S., Lander, E., and Reich, D. 2006. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* 441: 1103-1108.
- Petit, R., Mousadik, E., and Pons, O. 1998. Identifying populations for conservation on the basis of genetic markers. *Conserv Biol.* 12:844–855.
- Pires, A., and Fernandes, M. 2003. Last lynxes in Portugal? Molecular approaches in a pre-extinction. *Conservation Genetics* 4:525-532.
- Plomin, R., DeFries, J. McClearn, G., and McGuff, P. 2001. *Behavioral Genetics*. New York: Worth Publishers.
- Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* 25: 1253-1256.
- Prescott, J., Rapley, W., and Mewondo, J. 1994. Status and conservation of chimpanzee and gorilla in Cameroon. *Primate Conservation* 14-15:7-12.
- Queller, D., and Goodnight, K. 1989. Estimating relatedness using genetic markers. *Evolution* 43:258-275.
- Rambaut, A., and Drummond, A. 2007. Tracer v1.5, Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Rambaut, A. 2006. FigTree v.1.3.1 Available from <http://tree.bio.ed.ac.uk/software/figtree/>.
- Ramos-Onsins, S., E., and Rozas, J. 2002. Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* 19: 2092–2100.
- Ray, N., Currat, M., and Excoffier, L. 2003. Intrademe molecular diversity in spatially expanding populations. *Mol. Biol. Evol.* 20(1):76–86.

- Ray, N., and Excoffier, L. 2009. Inferring past demography using spatially explicit population genetic models. *Human Biology* 81:141-157.
- Reynolds, V. 2005. *The Chimpanzees of the Budongo Forest. Ecology, Behavior and Conservation*. Oxford University Press.
- Riddle, A., Pilgrim, K., Mills, L., Mckelvey, K., and Ruggiero, L. 2003. Identification of mustelids using mitochondrial DNA and non-invasive sampling. *Conservation Genetics* 4:241-243.
- Roeder, A., Archer, F., Poinar, H., and Morin, P. 2004. A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes* 4: 761–764.
- Rogers, A. 1995. Genetic evidence for a Pleistocene population explosion. *Evolution* 49(4): 608-615.
- Rogers, A., and Harpending, C. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9:552-569.
- Rogers, A., R., and Jorde, L., B. 1995. Genetic evidence on modern human origins. *Hum Biol* 67:1–36.
- Rogers, A., Fraley, E., Bamshad, M., Watkins, W., and Jorde, J. 1996. Mitochondrial mismatch analysis is insensitive to the mutational process. *Mol. Biol. Evol.* 13:895–902.
- Ronquist, F. and Huelsenbeck, J. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Rosemblum, L., Supriatna, J., and Melnick, D. 1997. Phylogeographic analysis of pigtail macaque populations (*Macaca nemestrina*) inferred from mitochondrial DNA. *American Journal of Physical Anthropology* 104:35-45.
- Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. and Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.

- Ruvolo, M., Pan, D., Goldberg, T., and Disotell, T. 1994. Gene trees and Hominoid Phylogeny. *PNAS* 91:8900-8904.
- Sayers, K., and Lovejoy, O. 2008. The Chimpanzee Has No Clothes. *Current Anthropology* 49:87-113.
- Schmidt, D., and Pool, J. 2002. The effect of population history on the distribution of the Tajima's D statistic. <http://www.cam.cornell.edu/Bdeena/TajimasD.pdf>.
- Schneider, S., and Excoffier, L. 1999. Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152:1079-1089.
- Schneider S., Kueffer, M., Roessli, D., and Excoffier L. 1997. Arlequin, v. 1.1: A software for population genetic data analysis. (Genetics and Biometry Laboratory, University of Geneva, Switzerland).
- Sequencher® version 5.0 sequence analysis software, Gene Codes Corporation, Ann Arbor, MI USA 2010. <http://www.genecodes.com> [Acedido: 12/01/2011].
- Shimada, M. 2000. Geographic distribution of mitochondrial DNA variations among grivet (*Cercopithecus aethiops aethiops*) populations in central Ethiopia. *Int. J. Primatol.* 21:113-129.
- Shimada, M., Hayakawa, S., Fujita, S., Sugiyama, S., and Saitou, N. 2009. Skewed matrilineal genetic composition in a small wild chimpanzee community. *Folia Primatol* 80:19-32.
- Shimada, M., Hayakawa, S., Humle, T., Fujita, S., Hirata, S., Sugiyama, Y., and Saitou, N. 2004. Mitochondrial DNA Genealogy of Chimpanzees in the Nimba Mountains and Bossou, West Africa. *American Journal of Primatology* 64:261-275.
- Simão, A. 1997. Identificação e delimitação cartográfica dos grandes maciços florestais de Cantanhez. Relatório de missão. Acção para o Desenvolvimento.
- Slatkin, M. 1991. Inbreeding coefficients and coalescence times. *Genet. Res. Camb.* 58, 167–175.

- Slatkin, M. 1993. Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, 47, 264–279.
- Slatkin, M. 1994. Gene flow and population structure. In: Real, L., A., editor. *Ecological Genetics*. Princeton. Princeton University Press. pp. 3–17.
- Slatkin, M., and Hudson, R. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129:555-562.
- Smith, T., and Wayne, R. 1996. *Molecular genetic approaches in conservation* New York: Oxford University Press.
- Sousa, J. 2007. Densidade de *Pan troglodytes verus* e Veículos de Sensibilização Ambiental: Quatro Florestas de Cantanhez, República da Guiné Bissau., Dissertação de Mestrado. Faculdade de Ciências. Universidade de Lisboa.
- Sousa, C., Gippoliti, S., and Akhlas, M. 2005. Republic of Guinea-Bissau. In *World Atlas of Great Apes and their Conservation*. Edited by J. Caldecott and L. Miles. London. University of California Press.
- Sousa, J., Barata, A., Sousa, C., Casanova, C., and Vicente, L. Chimpanzee oil-palm use in southern Cantanhez National Park, Guinea-Bissau. 2011. *American Journal of Primatology* 73:1 -13.
- Stone, A., Battistuzzi, F., Kubatko, L., Perry, G., Trudeau, E., Lin, H., and Kumar, S. 2010. More reliable estimates of divergence times in *Pan* using complete mtDNA sequences and accounting for population structure. *Phil. Trans. Soc. B* 365: 3277-3288.

- Stone, A., Griffiths, R., Zegura, L., and Hammer, M. 2002. High levels of Y-chromosome nucleotide diversity in the genus *Pan*. *PNAS* 99:43-48.
- Stone, A., and Verrelli, B. 2006. Focusing on comparative ape population genetics in the post-genomic age. *Current Opinion in Genetics & Development* 16:586–591.
- Stoneking, M. 2000. Hypervariable sites in the mtDNA control region are mutational hotspots. *The American Journal of Human Genetics*.67 (4): 1029-1032.
- Stöver, B., Müller, K. 2010. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* 11:7.
- Stumpf, R. 2005. Does promiscuity preclude choice? Female sexual strategies and mate preferences in chimpanzees of the Taï National Park, Cote d'Ivoire. *Behav. Ecol. Sociobiol.* 57:511-534.
- Sugiyama, Y. 1999. Socio-ecological factors of male chimpanzee migration at Boussou, Guinea. *American Journal of Primatology* 32:311-318.
- Sullivan, K., M., Mannucci, A., Kimpton, C., P, and Gill, P. 1993. A rapid and quantitative DNA sex test – fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *Biotechniques* 15: 636–641.
- Taberlet, P., Waits, L., and Luikart, G. 1999. Noninvasive genetic sampling: look before you leap. *Trends Ecol. Evol.* 14:323-327.
- Tajima, F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- Takahata, N., and Palumbi S., 1985. Extranuclear differentiation and gene flow in the finite island model. *Genetics* 109: 441-457.
- Takenaka, O., Kawamoto, S., Udono, T., Arakawa, M., Takasaki, H., and Takenaka, A. 1993. Chimpanzee microsatellite PCR primers applied to paternity testing in a captive colony. *Primates* 34:357-363.
- Tamura, K., Peterson, D., Peterson N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28: 2731-2739.

- Temudo, M., P. 2009. A narrativa da degradação ambiental no sul da Guiné-Bissau: uma desconstrução etnográfica. *Etnográfica* 13(2): 237-264.
- Thalmann, O., Hebler, J., Poinar, H., Pääbo, S., and Vigilant, L. 2004. Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. *Molecular Ecology* 13:321-335.
- The Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437:69-87
- Torres, J., J., C. Brito, et al. 2010. Ensemble models of habitat suitability relate chimpanzee (*Pan troglodytes*) conservation to forest and landscape dynamics in Western Africa. *Biological Conservation* 143: 416-425.
- Tourmen, Y., Baris, O., Dessen, P., Jacques, C., Malthiery, Y., and Reynier, P. 2002. Structure and chromosomal distribution of human mitochondrial pseudogenes. *Genomics* 80:71-77.
- von Berholdingen, C., Higuchi, R., Sensabaugh, G., and Erlich, H. 1987. Analysis of enzymatically amplified HLA-DQalpha DNA from single human hairs. *American Journal of Human Genetics* 41:725.
- Vigilant, L., and Guschanski, K. 2009. Using genetics to understand the dynamics of wild primate populations. *Primates* 50: 105-120.
- Vigilant, L., Hofreiter, M., Siedel, H., and Boesch, C. 2001. Paternity and relatedness in wild chimpanzee communities. *PNAS* 98:12890-12895.
- Villessen, P., and Fredsted, T. 2006. Fast and non-invasive PCR sexing of primates: apes, Old World monkeys, New World monkeys and Strepsirrhines. *BMC Ecol.* 6:8.
- Wasser, S., Houston, S., Cadd, G., and Fain, R. 1997. Techniques for application of fecal DNA methods to field studies of Ursids. *Molecular Ecology* 6:1091-1097.
- Watterson, G.A. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7 (2): 256–276.
- Winney, B., Hammond, R., Macasero, W., Flores, B., Boug, A., Biquand, V., Sylvain B., and Bruford, M.W. 2004. Crossing the Red Sea: phylogeography of the

- hamadryas baboon, *Papio hamadryas hamadryas*. *Molecular Ecology* 13(9): 2819-2827.
- White, F. 1996. "Comparative socio-ecology of *Pan paniscus*," in *Great Ape Societies*. Edited by W. McGrew, L. Marchant, and T. Nishida, pp. 29-41. Cambridge: Cambridge University Press.
- Wildman, D., Uddin, M., Liu, G., and Grossman, M. 2003. Implications of natural selection in shaping 99,4% nonsynonymous DNA identity between humans and chimpanzees: enlarging genus *Homo* *PNAS* 100:7181-7188.
- Woischnik, M., and Moraes, C. 2002. Pattern of organization of human mitochondrial pseudogenes in the nuclear genome. *Genome Research* 12:885-893.
- Woodruff, D. 1993. Non invasive genotyping of Primates. *Primates* 34:333-346.
- Woodruff, D. 2004a. Noninvasive genotyping and field studies of free-ranging nonhuman primates. In: *Kinship and Behavior in Primates*. Edited by B. Chapais and C. Berman. Oxford. Oxford University Press pp. 46-68.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16,97–159.
- Wright, S. 1943. Isolation by distance. *Genetics* 28, 114–138.
- Wright, S. 1951. The genetical structure of populations. *Ann. Eugenics* 15: 323-354.
- Wright, S. 1965. The interpretation of population structure by F-Statistics with special regard to systems of mating. *Evolution* 19:395-420.
- WWF. 2009. <http://www.panda.org> [Acedido: 27/08/2009].
- Yu, N., Jensen-Seaman, M., Chemnick, L., Kidd, J., Deinard, A., Ryder, O., Kidd, K., and Li, W. 2003. Low Nucleotide Diversity in Chimpanzees and Bonobos. *Genetics* 164:1511-1518.
- Zuckerkandl, E. 1963. "Perspectives in molecular anthropology," in *Classification and Human Evolution*. Edited by S. Washburn, pp. 243-272. Chicago: Aldine.
- Zsurka, G., Kudina, T., Peeva, V., Hallmann, K., Elger, C., Khrapko, K., and Kunz, K. 2010. Distinct patterns of mitochondrial genome diversity in bonobos (*Pan paniscus*) and humans. *BMC Evolutionary Biology* 10 (1): 270.

ANEXO 1: Lista de sequencias dos haplótipos HVR1 dos chimpanzés de Cantanhez a serem submetidas ao GenBank, em formato FASTA.

>H1

```
TCACITTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG  
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCTATCTGAGGGGGATCATCCGT  
GGGGGCGAGGATGGATTTGACTGTAATGTGCTATGTACGGTGTGTGATTGTATGTA  
CTATGTTTTGTCAAGGGGGG-TAGGTTTGTGGTATCGGGGTGGGGGAGAAGTGTCTGGAG  
TTGTGTTTTATGTTTGATAGTTAGAGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG  
AAGGTTTTAATGTGGG-GTGGGTTTTATGTGTTATAGGTAGTTGGGTAATTATGGTACCG  
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG  
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT  
TA
```

>H2

```
TCACITTTAGCTACCCCCAGGTGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG  
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCCTATCTGAGGGGGGGCATCCGT  
GGGGACGAGGAAGAGTTTACTGTAATGTGCTATGTACGGTGTATGATTGTATGTA  
CTATGTTCTGTCAAGGGGGG--TAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTCTGGGG  
TTGTGTTTTATGTTTGACAGTTGGAGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG  
AAGGTTTTGGTGTGGGAGTGGATTTTATGTATTATAGGTAGTTGGGTGATTATGGTACTG  
TACAATATTCATGGTACTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG  
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT  
TA
```

>H3

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCCTATCTGAGGGGGGGGCATCCGT
GGGGACGAGGAAGAGTTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGTA
TTCTGTCAAGGGGGG-TAGGTTTGTGATATCGGGGTGGGGGAGGGGTGTCGTTGGAG
TTGTGTTTTATGTTGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTATTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H4

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGTACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGCCAAGGGATTCCCATCTGAGGGGGATCATCCGT
GGGGCAAGGATGGATTTGACTGTAATGTGCTATGTACGGTGTGTGATTGTATGTATTAT
GTTTTGTCAAGGGAGGG-TAGGTTTGTGGTATTAGGGTGGGGGAGAAGTGTGCTTGGAG
TTGTGTTTTATGTTGATAGTTGGGGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTGATGTGAG-GTGGATTTTATGTGCTATAGGTGGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H5

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGTACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCATCTGAGGGGGATCATCCGT
GGGGCAAGGATGGATTTGACTGTAATGTGCTATGTACGATGTGTGATTGTATGTATTAT
GTTTTGTCAAGGGAGGG-TAGGTTTGTGGTATTAGGGTGGGGGAGAAGTGTGCTTGGAG
TTGTGTTTTATGTTGATAGTTGGGGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTAATGTGAG-GTGGATTCTATGTGTTATAGGTGGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG

CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H6

TCACTTTAGCTACCCCCAGATGTTATGGGCCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCCTATCTGAGGGGGGGGCATCCGT
GGGGACGAGGAAGAGTTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGACTAT
GTTCTGTCAAGGGGGGG-TAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTCGTTGGAG
TTGTGTTTTATGTTTCGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAGCATGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTATTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H7

TCACTTTAGCTACCCCCAGATGTTATGGGCCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCCTATCTGAGGGGGGGGCATCCGT
GGGGACGAGGAAGAATTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGACTAT
GTTCTGTCAAGGGGGGGgGTAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTCGTTGGAG
TTGTGTTTTATGTTTCGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAGCaTGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTGTTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H8

TCACTTTAGCTACCCCCAGATGTTATGGGCCCCGGAGCGAGGGGAGTAGTACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCATCTGGGGGGGGTCATCCGT
GGGGACAAGGATGGATTTGACTGTAATGTGCTATGTACGATGTATGATTGTATGACTAT
GTCTTGTCAAGGGGGGG-TGGGTTTGTGGTATCGGGGTGGGGGAGGAGTGTGTTGGGG

TTGTGTTTTATGTTTCGATAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTAATGTGGG-GTGGGTTTTATGTATTATAGGTGGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H9

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGGGAGTAGTACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCATCTGGGGGGGGTCATCCGT
GGGACAAGAATGGATTTGACTGTAATGTGCTATGTACGATGTATGATTGTATGACTAT
GTCTTGCAAGGGGGGG-TGGGTTTGTGGTATCGGGGTGGGGGAGGAGTGTGTTGGGG
TTGTGTTTTATGTTTCGATAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTAATGTGGG-GTGGGTTTTATGTATTATAGGTGGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H10

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCATCTGAGGGGGGGGCATCCGT
GGGGACGAGGAAGAATTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGACTAT
GTTCTGTCAAGGGGGGG-TAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTGTTGGAG
TTGTGTTTTATGTTTCGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTGTTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H11

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCATCTGAGGGGGGGGCATCCGT

GGGGACGAGGAAGAGTTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGTA
GTTCTGTCAAGGGGGGGTAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTCGTTGGA
TTGTGTTTTATGTTTCGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAGCATGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTATTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H12

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCCTATCTGAGGGGGGGCATCCGT
GGGGACGAGGAAGAGTTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGTA
GTTCTGTCAAGGGGGG--TAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTCGTTGGAG
TTGTGTTTTATGTTTCGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAGCATGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTATTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H13

TCACTTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGGGAGTAGTACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCATCTGGGGGGGGTCATCCGT
GGGGACAAGGATGGATTTGACTGTAATGTGCTATGTACGATGTATGATTGTATGTA
GTCTTGTCAAGGGGGGG-TGGGTTTGTGGTATCGGGGTGGGGGAGGAGTGTCGTTGGGG
TTGTGTTTTATGTTTCGATAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAGCATGGGGTG
AAGGTTTTAATGTGGG-GTGGGTTTTATGTATTATAGGTGGTTGGGTGATTATGGTACCG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H14

TCACITTTAGCTACCCCCAGATGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCTATCTGAGGGGGATCATCCGT
GGGGGCGAGGATGGATTTGACTGTAATGTGCTATGTACGGTGTGTGATTGTATGTACTAT
GTTTTGTCAAGGGGGG-TAGGTTTGTGGTATCGGGGTGGGGGAGAAGTGTGCGTTGGAG
TTGTGCTTTATGTTTGATAGTTAGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTAATGTGGG-GTGGGTTTTATGTGTTATAGGTAGTTGGGTAATTATGGTACCG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H15

TCACITTTAGCTACCCCCAGGTGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGGTCAAGGGATTCCCTATCTGAGGGGGATCATCCGT
GGGGGCAAGGATGGATTTAACTGTAATGTGCTATGTACGGTGTGTGATTGTATGTACTAT
GTTTTGTCAAGGGGGG--TAGGTTTGTGGTATCGGGGTGGGGGAGAGGTGTGCGTTGGAG
TTGTGTTTTATGTTTGATAGTTAGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTAATGTGGG-GTGGGTTTTATGTGCTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGCTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGAG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA

>H16

TCACITTTAGCTACCCCCAGGTGTTATGGGCCCGGAGCGAGGAGAGTAGCACTCTTGTGCG
GGATATTGATTTACGGAGGATGGTGACCAAGGGATTCCCTATCTGAGGGGGGGCATCCGT
GGGGACGAGGAAGAGTTTGACTGTAATGTGCTATGTACGGTGTATGATTGTATGTACTAT
GTTCTGTCAAGGGGGG--TAGGTCTGTTGATATCGGGGTGGGGGAGGGGTGTGCGTTGGAG
TTGTGTTTTATGTTTGACAGTTGGAGGTTGATTGTTGTGCGTGCTTGTAAAGCATGGGGTG
AAGGTTTTGGTGTGGGAGTGGATTTTATGTATTATAGGTAGTTGGGTGATTATGGTACTG
TACAATATTCATGGTGGTACTGGCAGTAATGTACGAAATACATAGCGGTTGTAATGAATGGG
CCAGTACTTAGGTGGTACTTAAATTTGCTTCCCCATGAAAGAACAGAGAATAGTTTAAAT
TA