

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA E IMUNOLOGIA**

LUCAS SECCHIM RIBEIRO

**AVALIAÇÃO DA VARIAÇÃO DO NÚMERO DE CÓPIAS DO
MICROSSATÉLITE TETRALOCAL *DYS503* EM
POPULAÇÕES DA ÁFRICA**

BELO HORIZONTE – MG

2010

LUCAS SECCHIM RIBEIRO

**AVALIAÇÃO DA VARIAÇÃO DO NÚMERO DE CÓPIAS DO
MICROSSATÉLITE TETRALOCAL *DYS503* EM
POPULAÇÕES DA ÁFRICA**

Dissertação apresentada ao Departamento de Bioquímica e Imunologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção de título de Mestre em Ciências.

Área de concentração: Biologia Molecular.

Orientador: Prof. Sérgio Danilo Junho Pena

BELO HORIZONTE – MG

2010

Ribeiro, Lucas Secchim.

Avaliação da variação do número de cópias do microssatélite tetralocal *DYS503* em populações da África. [manuscrito] / Lucas Secchim Ribeiro. – 2010.

81 f. : il. ; 29,5 cm.

Orientador: Sérgio Danilo Junho Pena.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Departamento de Bioquímica e Imunologia.

1. Microssatélites (Genética) - Teses. 2. Cromossomo Y - Teses. 3. Genética da população humana – África – Teses. 4. Biologia molecular – Teses. 5. Variabilidade genética. I. Pena, Sérgio Danilo Junho Pena. II. Universidade Federal de Minas Gerais. Departamento de Bioquímica e Imunologia. III. Título.


CDU: 577.2



Universidade Federal de Minas Gerais -
 Curso de Pós-Graduação em Bioquímica e Imunologia ICB/UFMG
 Av. Antônio Carlos, 6627 - Pampulha
 31270-901 - Belo Horizonte - MG
 e-mail: pg-biq@icb.ufmg.br (31)3409-2615




ATA DA DEFESA DA DISSERTAÇÃO DE MESTRADO DE LUCAS SECCHIM RIBEIRO. Aos trinta dias do mês de agosto de 2010 às 09:00 horas, reuniu-se no Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais a Comissão Examinadora da dissertação de Mestrado, indicada *ad referendum* do Colegiado, para julgar, em exame final, o trabalho intitulado “Avaliação de possível variação do número de cópias do microssatélite tetralocal DYS503 em populações da África” requisito final para a obtenção do grau de Mestre em Bioquímica e Imunologia. Abrindo a sessão o Presidente da Comissão, Prof. Sérgio Danilo Junho Pena, da Universidade Federal de Minas Gerais, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações: Dr. Fabrício Rodrigues Santos, da Universidade Federal de Minas Gerais, aprovado; Dr. Eduardo Tarazona Santos, da Universidade Federal de Minas Gerais, aprovado; Dr. Sérgio Danilo Junho Pena, orientador, da Universidade Federal de Minas Gerais, aprovado. Pelas indicações, o candidato foi considerado APROVADO. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente da Comissão encerrou a reunião e lavrou a presente Ata que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 30 de agosto de 2010.


 Dr. Fabrício Rodrigues Santos – UFMG


 Dr. Eduardo Tarazona Santos – UFMG


 Dr. Sérgio Danilo Junho Pena - UFMG


 Prof. Maria Elena de Lima Peres Garcia
 Coordenadora do Curso de Pós Graduação
 em Bioquímica e Imunologia
 ICB - UFMG

AGRADECIMENTOS

Ao Prof. Dr. Sérgio Danilo Junho Pena, sou grato pela oportunidade oferecida, pelo acolhimento em um grupo tão seleta de alunos e pesquisadores e pelo aprendizado que ultrapassa as paredes do laboratório. Agradeço pela orientação e pela desorientação, que muitas vezes fazia com que eu buscasse minhas próprias respostas e interpretações. Durante essa caminhada, pude compreender a importância da disciplina e competência máxima em cada experimento e que o sucesso é fruto de muita dedicação e trabalho árduo. Com seus conselhos, críticas e sugestões, cresci como pessoa e profissional. Acima de tudo, quero agradecer-lo por mostrar a Ciência que vai além da bancada.

À Fernanda de Souza Gomes Kehdy Lopes, a Ferdi, minha sincera gratidão pela amizade verdadeira e pelo apoio constante. Obrigado pelo incentivo, companheirismo, a ajuda indispensável durante toda a execução deste projeto e pelos *brownies* deliciosos!! Com você aprendi muito mais que Genética e Biologia Molecular (... *“Pega um papel e uma caneta aí! Vamos rabiscar!”*): aprendi o grande valor de uma boa amizade e como é *sensacional* cultivá-la permanentemente. Muito obrigado por tudo!

Aos professores Andréa Mara Macedo, Carlos Renato Machado e Glória Regina Franco, obrigado pela boa convivência e pelas sugestões e críticas ao meu trabalho.

Agradeço a todos os amigos do Laboratório de Genética Bioquímica, que me acolheram muito bem e desde cedo me fizeram sentir em minha própria casa. Obrigado pelas dicas de bancada e por todo conhecimento que dividiram comigo. Sou grato, com especial afeto, aos eternos mosqueteiros João Pedro e Paula, pelos momentos de desconcentração, pela companhia nos infinitos cafés e pelas ótimas recordações. Completando o time do *happy hour*. Matheus, o ás do volante de kart; Carol Furtado, dona de um carisma ímpar e Rodrigo, o futuro mago do Linux. Obrigado por proporcionarem um clima sempre agradável e divertido, não havendo tempo ruim quando estamos juntos! Ao Professor Helder, sou grato por compartilhar sua moradia e seu grande conhecimento em Biologia Molecular. Valeu, galera! Aqui... Sucesso!

À Vanessa Faria Gonçalves, agradeço demais por compartilhar informações e dicas valiosas para o aprimoramento de minhas técnicas e por contagiar todos nós com sua alegria totalmente rosa!

Um agradecimento especial à Neuza, por administrar e cuidar do nosso laboratório, sempre preocupada em manter tudo em ordem. Sua colaboração é de grande importância para todos nós.

Aos companheiros de Bases Moleculares 1, com os quais comecei essa incrível jornada. Apesar do curto tempo juntos, formamos um grupo muito unido até hoje. Agradeço a experiência compartilhada e desejo todo sucesso do mundo a cada um!

Agradeço também a atenção, eficiência e disponibilidade da secretaria da pós-graduação em Bioquímica e Imunologia, muito bem representada pela Celise, Grazielle e Ubirajara.

Um justo agradecimento ao Ubiraci (Embrapa/Sete Lagoas), ao Juliano (Núcleo de Análise Genômica/UFMG) e à Luciana Werneck (LDGH/UFMG) pelo apoio indispensável e pela troca de ideias sobre o MegaBACE 1000.

Às Irmãs da Congregação de Jesus na Eucaristia, muito obrigado por ceder o espaço, o conforto, a paz e a tranquilidade de sua morada durante os meus estudos para a prova do mestrado. Sem dúvida, a colaboração de vocês foi essencial para essa conquista!

Aos parceiros do Laboratório de Leishmanioses (LabLeish, NDI/Ufes): Aretha, Bianca, Fábio, Jáuber, Laura, Marco André, Priscila, Renata e especialmente, Marcela Segatto, pelos primeiros passos no sensacional mundo da Biologia Molecular e da Parasitologia. Foi um imenso prazer dividir com todos vocês os bons momentos de trabalho e o aprendizado de minha iniciação científica. Com grande carinho, agradeço à professora Elenice Moreira Lemos pela oportunidade de integrar seu grupo de pesquisa, pelo incentivo constante e pelas incríveis aulas de Bioquímica e Imunologia Clínica durante o curso de Farmácia. Minha gratidão a todos vocês!

Agradeço aos professores André Gomes Faraco e Suely Gomes Figueiredo por terem sido, de fato, meus primeiros orientadores. Ainda que involuntariamente, abriram meus olhos para as boas oportunidades e fizeram com que eu compreendesse o valor e a responsabilidade de trabalhar com Ciência.

O agradecimento mais especial aos meus pais, Sérgio e Rozangela, por estarem sempre ao meu lado, apoiando minhas decisões! Obrigado pela excelente educação, pelo amor incondicional, pela presença constante (apesar da distância), pelas broncas que só me fizeram crescer e por serem exemplo de valores morais. A vocês, sou grato pela herança das características que *crossing over* nenhum é capaz de superar: discernimento, perseverança, ética e determinação. Agradeço a toda minha família pelo apoio e confiança. Cada um de vocês é parte essencial de um alicerce que, construído com o tempo, torna-se mais forte e resistente contra qualquer intempérie.

Sou grato a todos que, direta ou indiretamente, contribuíram para o andamento deste projeto.

À agência financiadora, CNPq.

"A ciência é originada do desejo pelo conhecimento, e sua fonte é a curiosidade."

Luigi Luca Cavalli-Sforza

SUMÁRIO

LISTA DE FIGURAS	i
LISTA DE TABELAS	iii
LISTA DE ABREVIATURAS E SIGLAS	iv
RESUMO	vi
ABSTRACT	viii
A. INTRODUÇÃO	1
A.1 O continente africano e a origem da humanidade	1
A.2 Diversidade genética na África	3
A.3 O cromossomo Y humano	6
A.3.1 Origem evolutiva	6
A.3.2 Região não-recombinante do cromossomo Y	10
A.3.3 Estudos moleculares	14
A.3.4 Variabilidade no cromossomo Y	15
A.3.5 O microssatélite <i>DYS503</i>	16
A.4 Variações do número de cópias (<i>CNV</i>)	17
JUSTIFICATIVA E OBJETIVOS	21
B. MATERIAL E MÉTODOS	24
B.1 Construção do mapa físico do <i>DYS503</i>	24
B.2 Estudo da variação do número de cópias	28
B.2.1 Amostras de DNA	28
B.2.2 Reação em Cadeia da Polimerase Quantitativa Fluorescente	33
B.2.3 Desenvolvimento dos iniciadores	35
B.2.4 Amplificação	39
B.2.5 Genotipagem	42
B.2.5.1 Análise da eletroforese	43
B.2.5.2 Normalização e número de cópias	43
B.2.6 Análise estatística	44

C. RESULTADOS	46
C.1 Mapa físico do microssatélite <i>DYS503</i>	46
C. 2 Mapa de variantes	49
C.3 Genotipagem das amostras africanas do painel do CEPH/HGDP	51
D. DISCUSSÃO	58
D.1 Mapa físico do microssatélite <i>DYS503</i>	60
D.2 Amostragem na África	60
D.3 Genótipos estendidos	61
D.4 Deriva genética.....	63
E. CONCLUSÕES	66
F. REFERÊNCIAS	68
Sítios de Internet.....	81

LISTA DE FIGURAS

Figura 1: A migração do <i>Homo sapiens</i>	4
Figura 2: Representação gráfica do cromossomo Y humano.	8
Figura 3: Caminho proposto para a evolução dos cromossomos sexuais humanos, com a redução do cromossomo Y e a expansão de seu bloco não-recombinante (MSY).	9
Figura 4: Principais <i>amplicons</i> e palíndromos (P1, P2 e P3) de uma região de cerca de 4.5 Mb no braço longo do cromossomo Y, que contém a região AZFc.	13
Figura 5: Descoberta de CNVs através de haplótipos anômalos de microssatélites.	20
Figura 6: Fragmento do bloco genômico contendo a região variável e um pequeno pedaço da porção constante onde está o microssatélite <i>DYS503</i>	27
Figura 7: Mapa do continente africano, com indicação numérica da localização geográfica das populações estudadas.	31
Figura 8: Alinhamento gerado pelo programa online <i>MultAlin</i> , com os 5 fragmentos gerados pelo único par de iniciadores: quatro para o cromossomo Y e o fragmento referência de cópia única do cromossomo 22.	38
Figura 9: Esquema de amplificação utilizando a sequência M13-40 marcada com FAM.	41
Figura 10: Posição cromossômica dos quatro blocos do microssatélite <i>DYS503</i>	47
Figura 11: Mapa de unidades variantes encontradas dentro dos blocos de <i>DYS503</i> , segundo o banco de dados <i>Database of Genomic Variants</i>	50
Figura 12: Eletroferogramas indicando amplificação de uma amostra feminina e uma amostra masculina.	53
Figura 13: Gráficos de pontos representando a relação entre o valor normalizado das áreas e o valor normalizado das alturas dos picos encontrados para os fragmentos do cromossomo Y e 22.	54

Figura 14: Histograma da frequência dos valores normalizados de número de cópias para a população africana, em relação à amostra padrão *SP*, baseados na razão das áreas entre os picos do cromossomo Y e 22.55

Figura 15: Histograma da frequência dos valores normalizados de número de cópias para os 32 indivíduos com quatro alelos diferentes para *DYS503*, em relação à amostra padrão *SP*, baseados na razão das áreas entre os picos do cromossomo Y e 22.56

LISTA DE TABELAS

Tabela 1: Identificação, população e origem geográfica dos indivíduos estudados..	29
Tabela 2: 32 indivíduos selecionados para controle, com 4 alelos distintos na genotipagem do microssatélite <i>DYS503</i> .	32
Tabela 3: Sequências dos iniciadores utilizados na genotipagem	37
Tabela 4: Posição cromossômica dos blocos para o microssatélite <i>DYS503</i> no cromossomo Y	48

LISTA DE ABREVIATURAS E SIGLAS

ANOVA – *Analysis of Variance*

AZF (*azoospermia factor*) – fator de azoospermia

BLAST – *Basic Local Alignment Search Tool*

CEPH/HGDP (*Centre d'Etude du Polymorphisme Humain/Human Genome Diversity Project*) – Painel de Diversidade do Genoma Humano - Centro de Estudos de Polimorfismos Humanos

CNV (*copy number variation*) – variação do número de cópias

DAZ (*deleted in azoospermia*) – gene deletado em azoospermia

dATP -2'-desoxiadenosina 5'-trifosfato

dCTP- 2'-desoxicitosina 5'-trifosfato

dGTP- 2'-desoxiguanosina 5'-trifosfato

DNA – ácido desoxirribonucléico

dNTPs- desoxinucleotídeos

dTTP- 2'-desoxitimidina 5'-trifosfato

DYS – Segmento de DNA no cromossomo Y de cópia única ou moderadamente repetitivo

FAM – amidato de fluoresceína

Indels – polimorfismos de inserção ou deleção de nucleotídeos

Kb – Kilobases (1000 pares de bases)

LINEs (*long interspersed nuclear elements*) – elementos nucleares longos intercalados

LPA (*linear polyacrilamide*) – poliacrilamida linear

Mb – Megabases (1 milhão de pares de bases)

MSY (*male specific Y region*) – região específica do homem no cromossomo Y

NCBI – *National Center for Biotechnology Information*

PAR (*pseudoautosomal region*) – região pseudoautosômica

PCR (*polymerase chain reaction*) – reação em cadeia da polimerase

pb - pares de bases

QF-PCR (*Quantitative Fluorescent PCR*) – PCR quantitativa fluorescente

RNA – ácido ribonucléico

ROX – 6-carboxi-rodamina

SNP (*single nucleotide polymorphism*) – polimorfismo de nucleotídeo único

STR (*short tandem repeats*)– repetições aleatórias curtas

T_m - temperatura para a qual 50% do iniciador está corretamente anelado ao DNA molde

Y_p e Y_q – braços curto e longo do cromossomo Y, respectivamente

RESUMO

O ser humano anatomicamente moderno, *Homo sapiens*, é uma espécie recente no planeta e várias linhas de evidência genética indicam sua origem única e recente, há menos de 200 mil anos atrás, na África. A expansão da população humana a partir deste continente sofreu o efeito fundador, que ocorre quando um pequeno grupo não representativo deixa a população original estabelecida para fundar uma nova população, levando à redução da variabilidade genética do novo grupo. Com isso, uma das principais evidências da origem humana é a observação de uma maior diversidade genética na África do que em qualquer outro continente.

Contudo, recentes estudos populacionais com o microssatélite tetralocal *DYS503* do cromossomo Y revelaram uma inesperada maior diversidade haplotípica em brancos americanos em relação aos negros americanos. Considerando a natureza haplóide e a ausência de recombinação do cromossomo Y, foi averiguada a possibilidade de haver no continente africano indivíduos com múltiplas cópias para o microssatélite *DYS503*, e a que a população fundadora saiu do grupo inicial para colonizar o resto do mundo continha apenas parte dessa variabilidade, o que causaria a depressão da variabilidade quanto tipada pelos alelos presentes. Para isso, foi utilizado neste estudo um método de PCR quantitativa, na qual um único par de iniciadores é capaz de amplificar simultaneamente, produtos da região conservada próxima ao microssatélite e de um loco de referência de cópia única, com tamanhos distintos para posterior análise em sequenciador automático.

Um total de 101 amostras de homens africanos do painel do CEPH/HGDP foram quantificadas. Os resultados indicam que a grande maioria dos indivíduos possui quatro cópias para o bloco genômico onde se encontra o microssatélite *DYS503*,

não havendo, portanto, evidências que apontem a existência de indivíduos com mais ou menos de quatro alelos naquele continente. A construção do mapa físico indica assimetria entre as quatro cópias do microssatélite, levantando a suspeita de ocorrência de eventos desiguais de duplicação e deleção cromossômica. Sendo assim, é possível que outros fatores loco-específicos como a alta taxa de mutação desse microssatélite ou desvios de amostragem possam estar agindo sobre essa região e gerando a incomum queda da variabilidade.

ABSTRACT

The anatomically modern human being, *Homo sapiens*, is a recent species on the planet and several lines of genetic evidence indicate its unique and recent origin, less than 200,000 years ago in Africa. The expansion of human populations from this continent suffered the founder effect, which occurs when a small group not representative leaves the original population to found a new one, leading to reduced genetic variability of the new group. Thus, a major evidence of human origins is the observation of greater genetic diversity in Africa than in any other continent.

However, recent population studies with the tetralocal microsatellite *DYS503* revealed a higher Y-chromosome haplotype diversity among white Americans toward blacks Americans. Given the haploid nature and lack of recombination of the Y chromosome, it was investigated the possibility of the individuals from Africa have multiple copies of the microsatellite *DYS503*, and that the founding population, which left the original group to colonize the rest of the world, had only part of this variability, causing depression of variability as typed by the present alleles. Therefore, a quantitative PCR method was carried out in this study, in which a single pair of primers could simultaneously amplify both products of conserved region near the microsatellite locus and a single copy reference in another chromosome, with different sizes for subsequent analysis automated sequencer.

A group of 101 samples of African men from CEPH/HGDP panel was quantified. The results indicate that most individuals have four copies of the genomic block where the microsatellite *DYS503* is, and there is no evidence to suggest the existence of individuals with more or less than four alleles in that continent. The construction of the physical map indicates asymmetry between the four copies of the microsatellite,

triggering the suspicion of disparate events of chromosomal duplication and deletion. It is therefore possible that other site-specific factors such as high microsatellite mutation rate or deviations of sampling may be acting on this region and generating the unusual drop in variability.



Introdução



A. INTRODUÇÃO

A.1 O continente africano e a origem da humanidade

A África é uma região de considerável diversidade genética, linguística, cultural e fenotípica. Além disso, as populações africanas praticam uma ampla variedade de técnicas de subsistência, incluindo vários modos de agricultura, pastoralismo e caça. Os africanos também vivem em climas que abrangem desde o maior deserto do mundo e a segunda maior floresta tropical até as savanas, pântanos e grandes cordilheiras montanhosas (Reed e Tishkoff, 2006).

O padrão de variação genética nas populações africanas modernas é influenciado por sua história demográfica (por exemplo, mudanças no tamanho da população, eventos migratórios de curto e longo alcance e miscigenação), bem como forças loco-específicas como seleção natural, recombinação e mutação (Campbell e Tishkoff, 2008). Entender essas e demais situações que modificaram durante milhares de anos a estrutura populacional do continente africano é crucial para o entendimento de tamanha diversidade e para a compreensão do surgimento da humanidade no mundo.

Entre os vários debates existentes sobre a origem da humanidade, há consenso de que o gênero *Homo* surgiu na África e que o *Homo erectus* se espalhou por todo o há cerca de um milhão de anos (Templeton, 1997). A partir deste ponto, a origem e a história demográfica do ser humano são tópicos de frequente discussão. No final dos anos 80, os primeiros estudos sobre a diversidade molecular humana sugeriram

que todas as populações não-africanas da nossa espécie, *Homo sapiens**, evoluíram de um ancestral comum na África aproximadamente 100 a 200 mil anos atrás, e então expandiu para o resto do mundo, substituindo as populações arcaicas de hominídeos. Essa hipótese, conhecida como “Origem Africana Recente” ou “Modelo *ex Africa*” é a mais aceita atualmente, sendo muito defendida e argumentada (Goldstein *et al.*, 1995; Li *et al.*, 2008; Stringer, 2000; Tishkoff e Williams, 2002).

Acredita-se que a migração dos humanos modernos para fora da África tenha sido acompanhada de um gargalo (“*bottleneck*”) populacional. Baseado na avaliação de 783 microssatélites autossômicos genotipados para o Painel de Diversidade do Genoma Humano do Centro de Estudos de Polimorfismos Humanos (CEPH/HGDP) e por uma análise combinada de dados de diversidade nucleotídica de DNA mitocondrial e dos cromossomos X e Y, o tamanho desta população migratória foi estimado em cerca de 1000 indivíduos fundadores efetivos (Liu *et al.*, 2006). A partir dessa população, duas rotas migratórias principais foram responsáveis pelo povoamento do resto do planeta (Figura 1). A primeira delas, que ocorreu na região sul da Península Arábica, garantiu a presença do ser humano ao longo da costa sul e sudeste da Ásia, bifurcando-se para a Oceania (entre 60 e 40 mil anos atrás) e para o norte, tomando a China, o Japão e eventualmente, as Américas, pelo estreito de Bering. A segunda rota, ao Norte, definiu a povoação do Oriente Médio, Arábia, Ásia Central, alcançando a Europa, o norte da Ásia e a América há cerca de 40 mil anos (Campbell e Tishkoff, 2008 Cavalli-Sforza e Feldman, 2003; Stringer, 2000).

**Homo sapiens*: descendente do *Homo erectus*, é o ser humano anatomicamente moderno, cujos registros fósseis datam de 150 a 200 mil anos atrás (Tishkoff e Williams, 2002).

A.2 Diversidade genética na África

A diferença entre a variabilidade genética encontrada para as populações mundiais é capaz de nos fornecer preciosas informações sobre sua história. Vários estudos mostram que as populações da África subsaariana possuem maior diversidade genética (heterozigosidade e diversidade nucleotídica) que as não-africanas. Estes estudos são baseados principalmente em sequências de DNA mitocondrial (Cann *et al.*, 1987; Vigilant *et al.*, 1991), sequências *Alu* e microssatélites do cromossomo X (Santos-Lopes *et al.*, 2007), microssatélites do cromossomo Y (Hammer *et al.*, 2001; Seielstad *et al.*, 1999), SNPs (polimorfismos de nucleotídeo único) do cromossomo Y (Thomson *et al.*, 2000; Underhill *et al.*, 2000), além de sequências, microssatélites e SNPs encontrados em autossomos (Calafell *et al.*, 1998; Jorde *et al.*, 1997; Yu *et al.*, 2002).

A hipótese da Origem Africana Recente também é baseada em outros aspectos da diversidade molecular. Primeiro, a maior parte da diversidade alélica encontrada fora da África é um subconjunto daquela encontrada no continente africano. Segundo, estudos com microssatélites mostram uma gradual redução da diversidade à medida que aumenta a distância da África. Terceiro, quando foi possível reconstruir o estado ancestral de alelos polimórficos comparando sequências humanas com aquelas de primatas superiores, a linhagem mais antiga era comumente encontrada entre africanos ou o alelo ancestral tinha maior frequência nessa população (Excoffier, 2002).

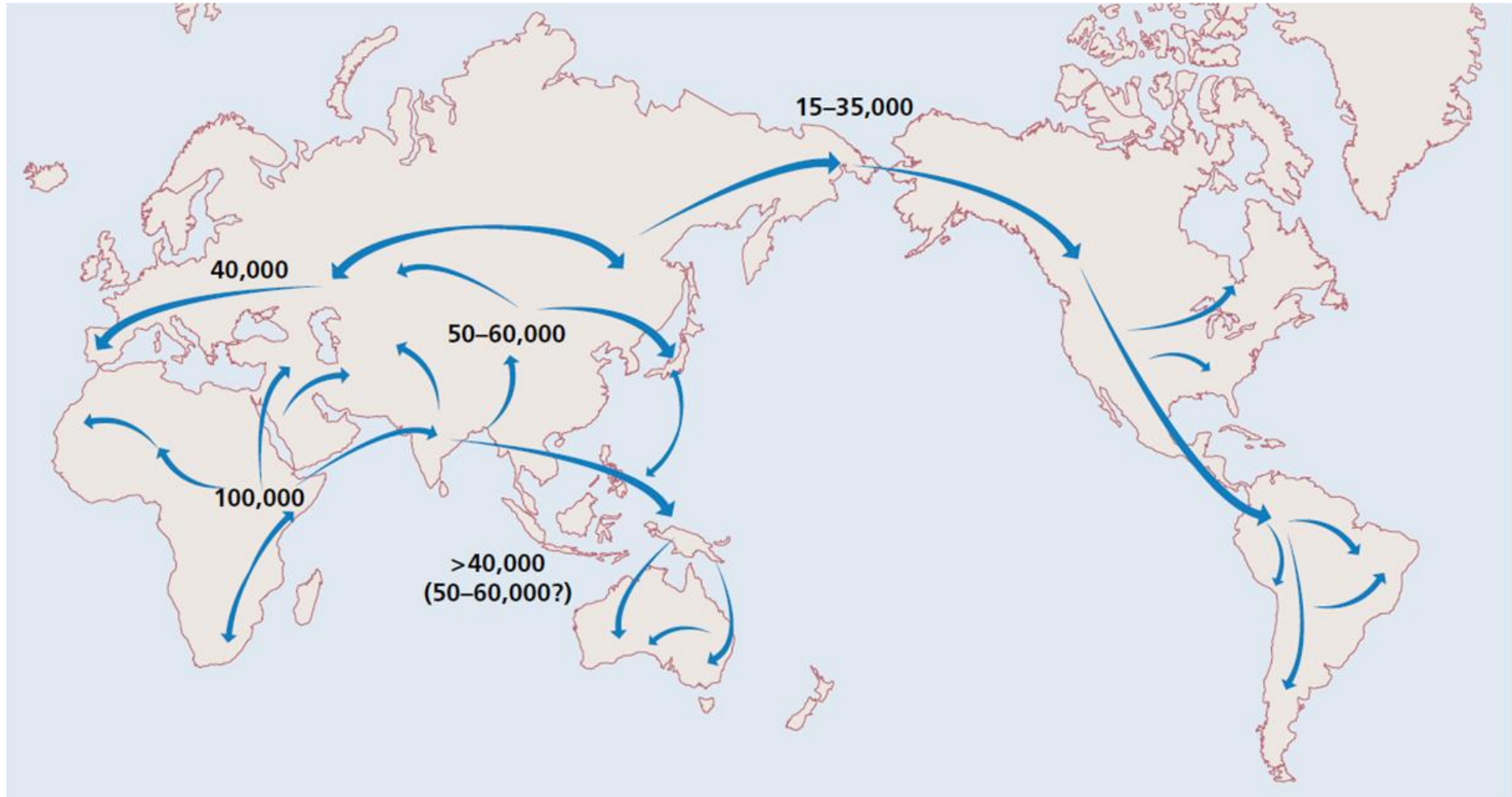


Figura 1: A migração do *Homo sapiens*. O esquema acima começa com a saída do ser humano a partir do Leste da África em direção ao resto do continente cerca de 100 mil anos atrás. Em seguida, a expansão para a Ásia, provavelmente em duas rotas, ao sul e ao norte, entre 60 e 40 mil anos atrás. A Oceania, a Europa e as Américas foram colonizadas, nesta ordem, a partir do continente asiático (Cavalli-Sforza e Feldman, 2003).

Por que uma região geográfica específica mostra níveis de diversidade genética maiores? Uma possibilidade é o tempo que uma determinada população vive em um determinado local. Quanto maior a existência dessa população, maior também o número de mutações que ocorrem e se acumulam. Além disso, o tamanho da população tem sua importância, visto que é proporcionalmente relacionado à diversidade esperada, ou seja, populações menores sofrem maiores efeitos da deriva genética e, portanto, tendem a ter sua diversidade reduzida (Relethford, 2008).

Outro importante fator a ser considerado ao comentar a diferença de diversidade entre africanos e não-africanos é o *efeito fundador*, que ocorre quando um grupo de indivíduos, não representativo da população original, migra para outro local e ali se fixa, gerando uma nova população com reduzida variabilidade. Dessa forma, alelos comuns na população parental da qual o novo grupo se originou podem ter a frequência reduzida ou até serem perdidos. Da mesma forma, alelos raros no grupo inicial podem ter sua frequência aumentada na população recém-fundada (Tishkoff e Verrelli, 2003). Estudando a relação existente entre distância geográfica e a diversidade genética das populações do painel CEPH/HGDP (Cann *et al.*, 2002), Ramachandran e colaboradores (2005) verificaram que a heterozigosidade média de um grande conjunto de microssatélites diminuía linearmente conforme aumentava a distância do Leste da África, onde a variabilidade era máxima. Desse modo, ao traçar a expansão mundial da população humana a partir da África, foi possível concluir que tais eventos aconteceram em passos, cada um envolvendo um pequeno grupo de indivíduos fundadores, levando a um gradual aumento do efeito da deriva genética e à correspondente redução da diversidade (DeGiorgio *et al.*, 2009; Deshpande *et al.*, 2009; Hellenthal *et al.*, 2008).

A.3 O cromossomo Y humano

O cromossomo Y é o menor dos cromossomos humanos, possuindo cerca de 59 megabases (Mb) de material genético. Consiste de um braço curto (Yp) e um braço longo (Yq). O braço curto e a porção proximal do braço longo consistem de eucromatina, enquanto a parte distal de Yq é constituída de heterocromatina (Hurles e Jobling, 2001). Nas pontas de cada um deles permanecem as regiões pseudoautossômicas (*PAR*, do inglês, *pseudoautosomic regions*), onde X e Y pareiam e formam o complexo sinaptonêmico durante a meiose. A *PAR1*, no braço curto do Y (2.6 Mb) é significativamente maior que a *PAR2* no braço longo (0.4 Mb). Fora das regiões pseudoautossômicas está a região não-recombinante, que consiste em diversos segmentos descontínuos de homologia X-Y entrelaçadas com regiões Y-específicas (Figura 2), sugerindo a ocorrência de muitas inversões e rearranjos durante a evolução (Yen, 1999).

A.3.1 Origem evolutiva

Entender os mecanismos responsáveis pela estrutura do cromossomo Y é fundamental para compreender sua variabilidade, visto que a ausência da recombinação constitui um dos eventos mais importantes para a determinação da história das linhagens masculinas. Entretanto, como os cromossomos X e Y evoluíram e se tornaram tão diferentes? O cromossomo Y é pequeno e praticamente não tem genes, enquanto o cromossomo X, seu parceiro meiótico, tem centenas

deles. Dada a existência de regiões de alta homologia entre os dois cromossomos, acredita-se que o cromossomo Y tenha evoluído de um ancestral comum junto com o cromossomo X (Graves, 2010; Hurles e Jobling, 2003; Lahn *et al.*, 2001; Waters *et al.*, 2007). Estudos comparativos sugerem que um autossomo tenha adquirido a função de cromossomo sexual e a supressão da recombinação entre ele e seu homólogo tenha permitido que evoluíssem independentemente (Figura 3) (Lahn e Page, 1999). A falta de recombinação em quase toda a extensão do cromossomo Y faz com que a seleção seja menos efetiva na prevenção do acúmulo de mutações deletérias e na fixação daquelas benéficas (Bachtrog e Charlesworth, 2001).

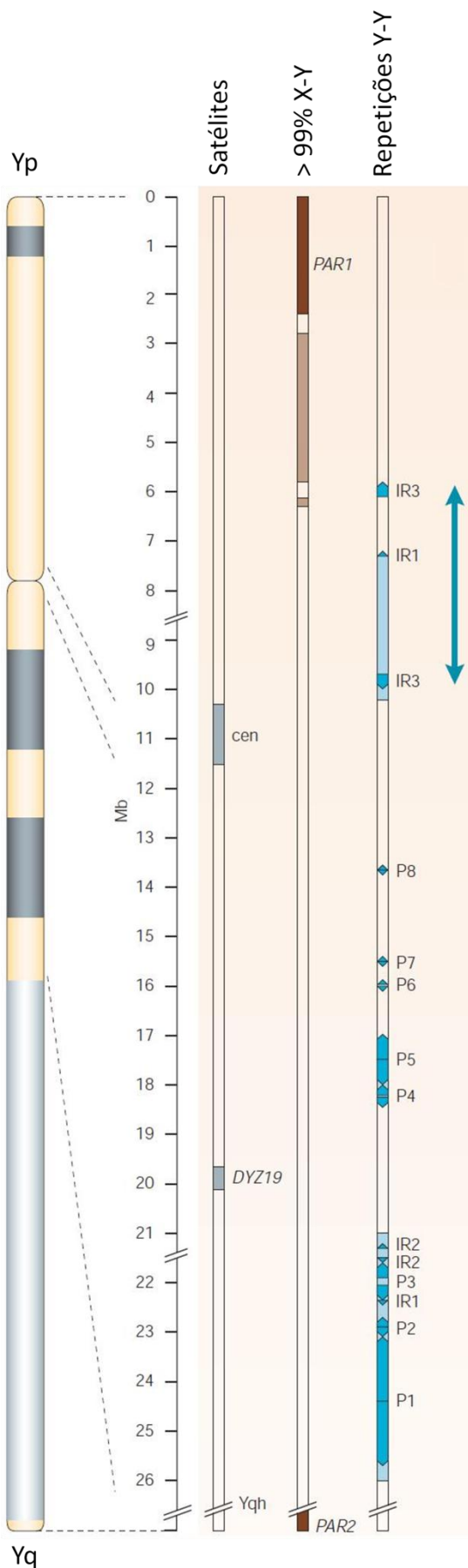


Figura 2: Representação gráfica do cromossomo Y humano. Da esquerda para a direita: características citogenéticas e suas posições aproximadas, numeradas a partir do telômero Yp. As regiões satélites do centrômero (*cen*) e *DYZ19* apontadas em cinza. Segmentos de identidade entre os cromossomos X e Y (*PAR1* e *PAR2*, marrom escuro) e de alta similaridade (marrom claro). Sequências repetitivas dentro do cromossomo Y (Y-Y) de grande homologia, designadas "IR" para *repetição invertida* e "P" para *palíndromo* (Adaptado de Jobling e Tyler-Smith, 2003).

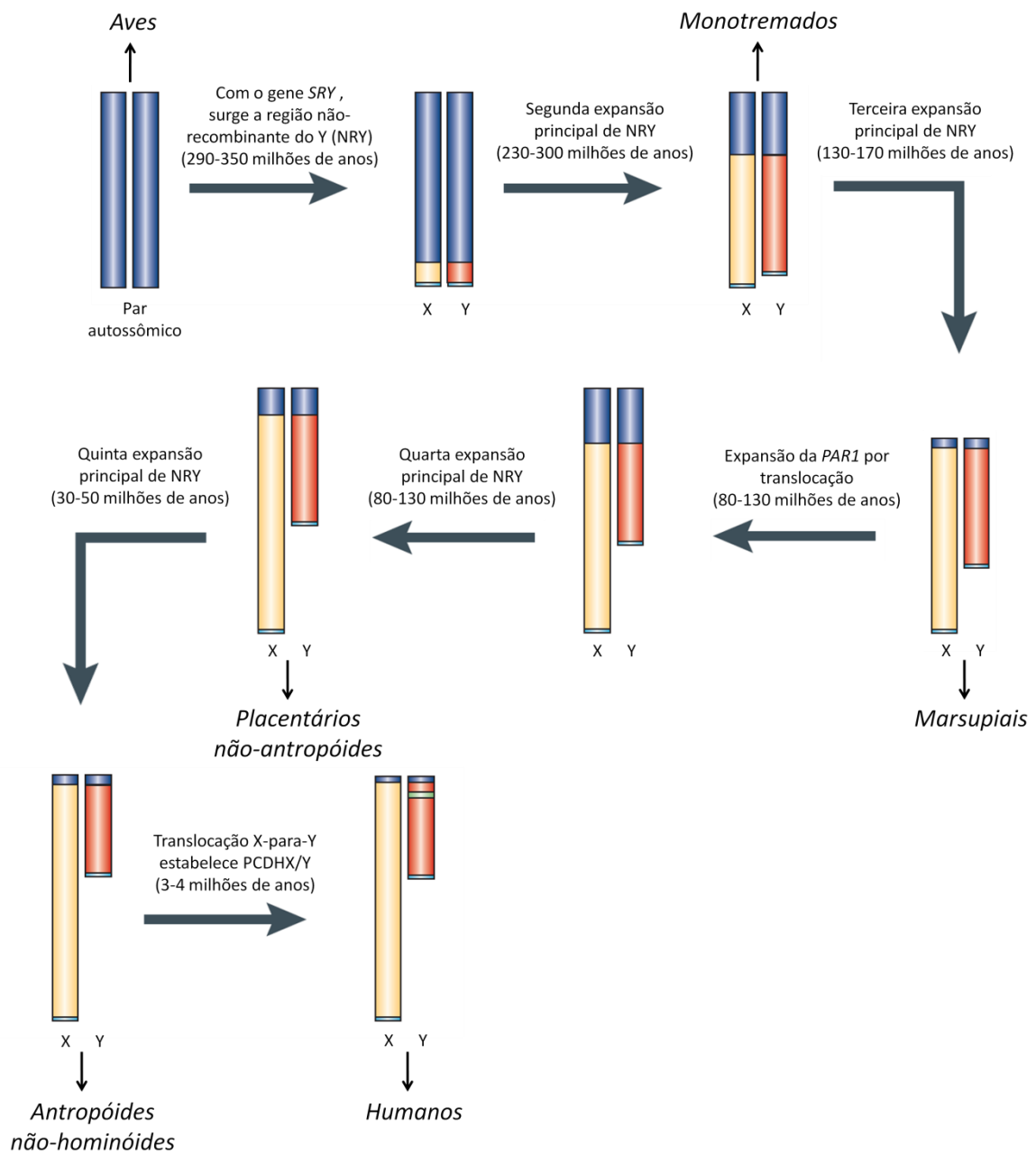


Figura 3: Caminho proposto para a evolução dos cromossomos sexuais humanos, com a redução do cromossomo Y e a expansão de seu bloco não-recombinante (MSY). Tais eventos foram provavelmente mediados pelas inversões de larga escala propostas por Lahn e Page (1999). As regiões azuis mantêm recombinação, as amarelas são X-específicas e as vermelhas, Y-específicas. A pequena porção verde aponta a sequência que contém o gene da protocaderina X/Y (PCDHY), translocada do cromossomo X para MSY (Adaptado de Lahn *et al.*, 2001).

A.3.2 Região não-recombinante do cromossomo Y

A região não-recombinante do cromossomo Y, também conhecida como região masculina específica (*male specific region*, MSY) contém um grande registro dos eventos mutacionais que ocorreram ao longo das linhagens masculinas durante a evolução, visto que é passada como estrutura haplóide de pai para filho e não está sujeita à recombinação (Jobling *et al.*, 1996; Mitchell e Hammer, 1996). Além disso, há em MSY um alto padrão geográfico de variabilidade: sua diversidade dentro das populações é menor do que encontrado para outros marcadores e a variação entre as populações é maior (Seielstad *et al.*, 1998; Wells *et al.*, 2001).

A porção eucromática de MSY possui cerca de 23 Mb (megabases), incluindo 8 Mb no braço curto e 14.5 Mb no braço longo. Está organizada em três níveis: *região transposta do X*, *região degenerada do X* e a *região amplicônica* (Skaletsky *et al.*, 2003). As sequências da região transposta do X são 99% idênticas às sequências de DNA da banda Xq21 no braço longo do cromossomo X. São assim nomeadas pois sua presença em MSY é o resultado de uma massiva transposição X-para-Y que ocorreu cerca de 3-4 milhões de anos atrás, após a divergência das linhagens do ser humano e do chimpanzé. As sequências *degeneradas* do X não participam do *crossing over* durante a meiose, distinguindo-as das regiões pseudoautossômicas nas extremidades do cromossomo. Nessa região, há poucos genes e alta densidade de elementos repetitivos intercalados, tal qual encontrado para o bloco homólogo em Xq21.

A região *amplicônica* consiste em longos segmentos repetitivos, com conjuntos de repetições invertidas e oito sequências em palíndromos, exibindo longos intervalos

de alta similaridade de nucleotídeos (> 99.9%) (Skaletsky *et al.*, 2003). Os palíndromos são grandes e seus braços alcançam de 9 kb a 1.45 Mb. Entre os braços, ao centro, há um espaçador, não duplicado, de 2-170 kb em tamanho. Juntos, atingem 5.7 Mb, abrangendo quase um quarto de toda MSY (Figura 4). Acredita-se que essas estruturas repetitivas sejam o substrato de constantes eventos de recombinação não-homóloga e conversão gênica, responsáveis por parte da variabilidade encontrada nessa região (Jobling, 2008; Lange *et al.*, 2009; Rozen *et al.*, 2003).

Além das discussões evolutivas e populacionais, há um especial interesse nessa região, justificado pela sua importância para a fertilidade masculina. A descrição citogenética de deleções no braço longo do cromossomo Y relacionadas à infertilidade (Tiepolo e Zuffardi, 1976) deu origem à hipótese da existência de *fatores de azoospermia* (AZF), onde estariam os genes indispensáveis à espermatogênese. Devido à complexidade estrutural, uma identificação precisa desses fatores levou cerca de 30 anos, quando microdeleções relacionadas à falha espermatogênica foram descritas e localizadas em três regiões definidas, denominadas AZFa, AZFb e AZFc (Vogt *et al.*, 1996), sendo a última a mais estudada.

Deleções recorrentes em AZFc são relativamente comuns, afetando um a cada 4000 homens. Essa região contém 21 genes, incluindo quatro cópias do gene *deletado em azoospermia* (DAZ), envolvido na regulação da tradução de RNA mensageiro. (Ravel *et al.*, 2009) A ausência de toda a região AZFc é causada principalmente por recombinação homóloga entre os elementos repetitivos *b2* e *b4* (Figura 4). Perdas de menor tamanho dentro dessa região podem ocorrer por recombinação intracromossômica, seja na mesma cromátide ou em cromátides irmãs (Repping *et*

al., 2003). A elucidação da sequência e estrutura do braço longo do cromossomo Y (Skaletsky *et al.*, 2003) e a descrição de vários haplótipos em diferentes linhagens humanas (The Y Chromosome Consortium, 2002) permitiram o reconhecimento de muitos tipos de deleções parciais em AZFc. De fato, essa região mostra significativa variação em termos de arquitetura genômica dentro da população masculina. Entretanto, as consequências fenotípicas para a espermatogênese para muitas dessas variantes não são claras, sendo necessários mais estudos com uma diversidade maior de amostras, a fim de determinar seu papel na fertilidade masculina.

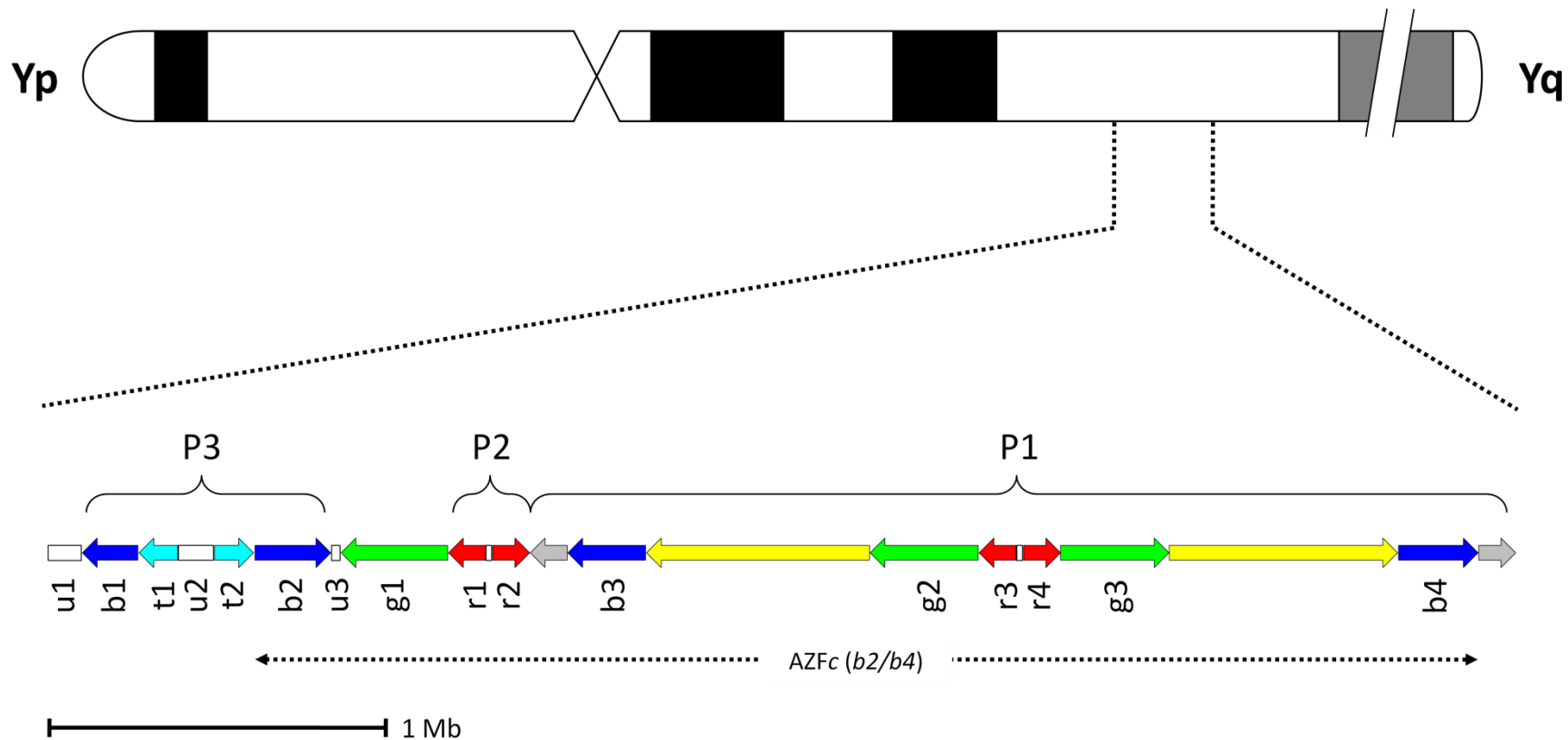


Figura 4: Principais amplicons (nomeados com letras minúsculas e números) e palíndromos (P1, P2 e P3) de uma região de cerca de 4.5 Mb no braço longo do cromossomo Y, que contém a região AZFc (*b2/b4*) (Adaptado de Kuroda-Kawaguchi, 2001).

A.3.3 Estudos moleculares

A história da pesquisa sobre o cromossomo Y pode ser traçada a partir de três partes bem distintas (Skaletsky *et al.*, 2003). Na primeira delas, estudos citogenéticos da década de 50 baseados em microscopia ótica sugeriram a existência de um cromossomo masculino específico e vários traços fenotípicos erroneamente ligados à sua presença, pela simples observação da transmissão destes caracteres de pai para filho. Mais tarde, na segunda era, investigações realizadas com portadores das síndromes de Klinefelter e Turner demonstraram que o cromossomo Y carregava um importante gene determinante de sexo. Ainda nessa época, surgiu a proposta da evolução dos cromossomos sexuais de mamíferos a partir de um par de autossomos e o subsequente “encolhimento” do cromossomo Y, restando-lhe apenas a fama de “lixeria genética” que durou até pouco tempo atrás.

Um segundo papel – o controle da espermatogênese – só foi reconhecido em meados dos anos 70, quando importantes deleções no braço longo do cromossomo Y foram diagnosticadas em seis homens inférteis (Tiepolo e Zuffardi, 1976). Devido à ausência de recombinação e ao desconhecimento de funções além da determinação do sexo, havia poucos marcadores moleculares desenvolvidos (Goodfellow *et al.*, 1985). A aplicação das modernas tecnologias genômicas e de DNA recombinante à pesquisa do cromossomo Y marcam a terceira e atual parte da história. O avanço constante das metodologias de abordagem molecular permite a compreensão de suas funções biológicas e ajudam a esclarecer seus padrões de diversidade. Com isso, houve nos últimos 15 anos um enorme crescimento na busca por novos marcadores e estratégias de tipagem que permitam a distinção de indivíduos e populações.

A.3.4 Variabilidade no cromossomo Y

Sabe-se que dois cromossomos Y escolhidos ao acaso diferem em um nucleotídeo para cada 3 a 4 mil bases. Dessa forma, o cromossomo Y representa uma grande fonte de mutações descritas como *polimorfismos de eventos únicos*, refletindo sua baixa taxa de ocorrência. Estes polimorfismos incluem marcadores binários de evolução lenta como SNPs e variações de inserção/deleção (*indels*). Por possuírem uma única origem mutacional em amostras de tamanho realístico, são utilizados para definir grupos relacionados de cromossomos, denominados *haplogrupos*. (Stumpf e Goldstein, 2001). Com a ausência de recombinação, é possível aumentar a capacidade informativa desse sistema haplotípico combinando a robusta filogenia alcançada pelos SNPs com a sensibilidade da tipagem de marcadores multialélicos de rápida mutação, como os microssatélites (Balaesque *et al.*, 2008).

Os microssatélites são sequências de unidades aleatoriamente repetitivas de 1 a 6 pares de bases, abundantes e altamente polimórficos em genomas eucarióticos. Acredita-se que os mecanismos responsáveis por tamanha variedade estejam ligados principalmente ao deslizamento da fita de DNA durante a replicação ou relacionados à conversão gênica e recombinação intracromossômica. De uma forma geral, a dinâmica populacional dos microssatélites se encaixa bem nos modelos de mutações de passo a passo (Li *et al.*, 2002). No cromossomo Y, mais de 200 microssatélites já foram identificados por vários grupos e depositados em bancos de dados públicos como o *GDB Human Genome Database* (Kayser *et al.*, 2004). Baseado nestas informações, um abrangente mapa físico foi construído, definindo o

número e a posição relativa dos microssatélites de importância forense e sua relação com genes estruturais (Hanson e Ballantyne, 2006).

Como as mutações constituem a única força que atua na diversificação dos haplótipos do cromossomo Y, entender a dinâmica mutacional é importante para a compreensão das origens de diversidade haplotípica (Jobling e Tyler-Smith, 2003). Uma taxa de mutação de 2×10^{-3} por geração foi estimada para microssatélites do cromossomo Y através de contagem direta em genealogias bem estabelecidas (Heyer *et al.*, 1997). Valores similares (3×10^{-3} /loco/geração) foram encontrados durante estudos com pares de pais e filhos de paternidade confirmada, embora os valores loco-específicos encontrados variem entre $0 - 8 \times 10^{-3}$ (Kayser *et al.*, 2000).

Todas essas propriedades fazem com que os microssatélites do cromossomo Y sejam utilizados como excelentes marcadores para análises forenses – evidências na cena do crime e casos de estupro, por exemplo (Wolf *et al.*, 2005; Roewer, 2009), investigações de pessoas desaparecidas (Coble *et al.*, 2009), testes de paternidade (Jobling *et al.*, 1997; Rolf *et al.*, 2001), pesquisas de diversidade populacional e migração (Gomes *et al.*, 2010), além de vários estudos evolutivos (Underhill *et al.*, 2000; Hammer *et al.*, 2001).

A.3.5 O microssatélite *DYS503*

O microssatélite *DYS503* foi identificado durante a construção de um mapa físico de microssatélites do cromossomo Y (Hanson e Ballantyne, 2006), estando presente em 4 cópias (*DYS503 a, b, c e d*) no braço longo. Na mesma região cromossômica,

outro microssatélite tetralocal – *DYS464* – apresentou alto poder discriminatório (Butler e Schoske, 2005). Logo, para verificar se o *DYS503* possuía igual diversidade, um novo estudo foi desenvolvido com este marcador, utilizando amostras de indivíduos americanos africanos e caucasianos (Hanson e Ballantyne, 2007). A unidade repetitiva – $[TTA]_n$ – apresentou alelos possuindo de 7 a 17 repetições, após calibração com a amostra referência sequenciada. Os resultados das frequências genotípicas revelaram maior diversidade nos caucasianos (0.92) em relação aos africanos (0.75) (Hanson e Ballantyne, 2007).

Além disso, esses curiosos dados apontam que a diversidade para o *DYS503* em caucasianos é a maior já encontrada quando comparada com qualquer outro microssatélite de cópia única do cromossomo Y conhecido. Utilizando o painel do CEPH/HGDP, Kehdy e colaboradores (2008) realizaram a genotipagem do *DYS503* para as populações parentais da África e Europa, encontrando as diversidades genotípicas de 0.63 para os africanos e 0.85 para os europeus, confirmando os dados previamente apresentados.

A.4 Variações do número de cópias (CNV)

Uma notável característica do genoma humano é a abundância de grandes duplicações caracterizadas pelo alto grau de identidade. É estimado que cerca de 5% do genoma é composto de sequências que são duplicadas no mesmo cromossomo ou em cromossomos diferentes (She *et al.*, 2004; Zhang *et al.*, 2005). Essas variações do número de cópias (CNV, do inglês, *copy number variation*) são

blocos de DNA com identidade de cerca de 90%, com tamanho maior que 1 kb (kilobase – 1000 pares de bases) e presente em múltiplas regiões em comparação com o genoma de referência (Redon *et al.*, 2006).

A presença de CNVs está fortemente associada a duplicações segmentais em regiões repetitivas do genoma, incluindo elementos *Alu* e elementos nucleares longos intercalados (LINEs). Isto provavelmente se deve às altas taxas de recombinação entre sequências flanqueadoras sujeitas a rearranjos genômicos através de recombinação homóloga não-alélica, resultando em deleções, duplicações ou inversões (Eichler, 2001; Sharp *et al.*, 2005; Trombetta *et al.*, 2010). A probabilidade de alinhamento incorreto entre regiões homólogas não-alélicas no genoma parece ser fortemente influenciada por vários fatores, como a homologia total e tamanho da sequência, bem como sua orientação (Redon *et al.*, 2006).

CNVs detectadas em estudos forenses e populacionais também podem ser avaliadas através da observação da variedade de microssatélites e do aparecimento de haplótipos anômalos (Figura 5) (Balaesque *et al.*, 2008; Butler *et al.*, 2005). Quando um microssatélite de cópia única está ausente em um determinado haplótipo e não há mutação no sítio de anelamento do iniciador, há indícios de deleção. Para a mesma situação, a presença de dois picos no eletroferograma pode indicar uma variante de duplicação, desde que não haja mosaicismos dos alelos do microssatélite, principalmente quando a origem do DNA é uma linhagem celular (Balaesque *et al.*, 2009; Jobling, 2008).

A inclusão do cromossomo Y nas buscas genômicas por CNVs e a disponibilidade dos dados do ressequenciamento do cromossomo Y forneceram um olhar mais objetivo sobre tais variantes. Como resultado, o *Database of Genomic Variants*

(lafrate *et al.*, 2004; <http://projects.tcag.ca/variation>) contém mais de 50 variantes, das quais 40 estão na região não-recombinante. Estas abrangem desde dezenas de pares de bases até centenas de kilobases, tendo sido detectadas em diferentes amostras, utilizando múltiplos métodos, com distintas sensibilidades (Jobling, 2008).

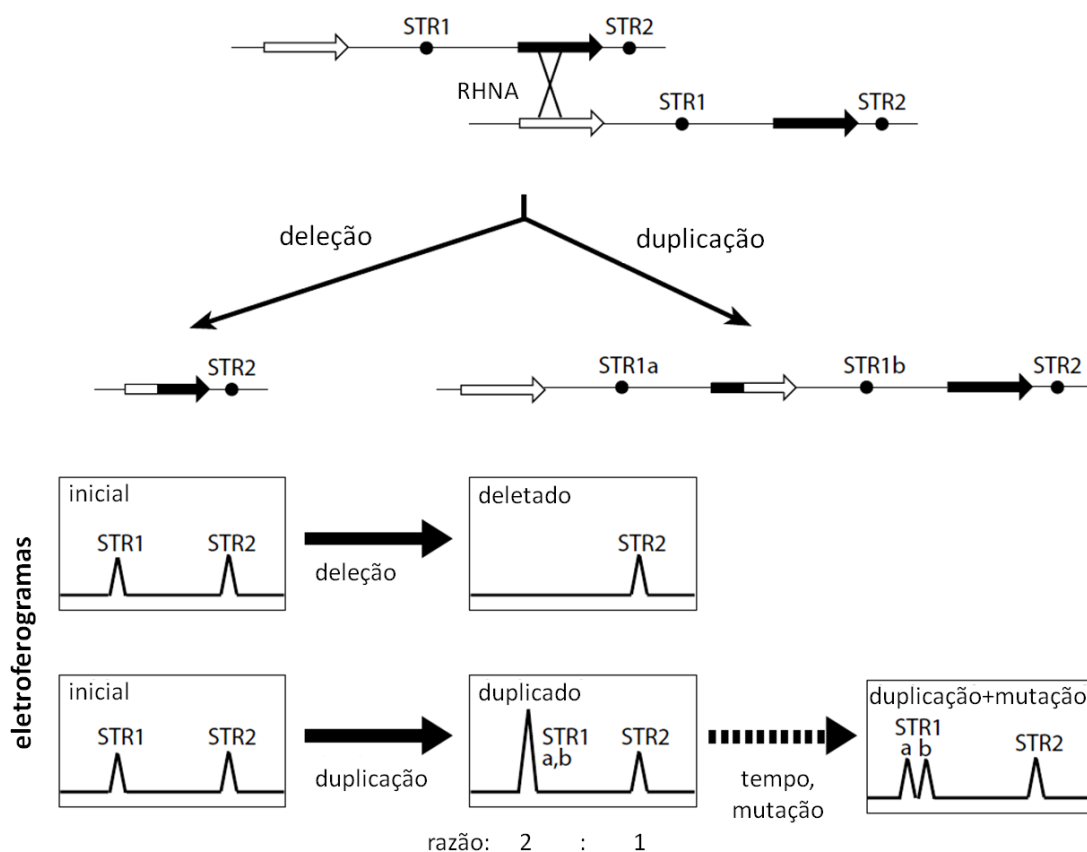


Figura 5: Descoberta de CNVs através de haplótipos anômalos de microssatélites. Neste esquema hipotético, a recombinação homóloga não-alélica (RHNA) é responsável pelos eventos iniciais de deleção e duplicação dos locos de microssatélites (STR, *short tandem repeats*). A deleção leva ao desaparecimento de um dos microssatélites, ao passo que a duplicação pode elevar a altura do respectivo pico no eletroferograma. O tempo e a variação comum das unidades repetitivas do microssatélite promovem a diferenciação dos picos original e duplicado (Adaptado de Jobling, 2008).

JUSTIFICATIVA E OBJETIVOS

Para o *Homo sapiens*, a variabilidade genética para populações de origem africana é reconhecidamente maior do que para as outras ao redor do mundo. Isso se deve ao maior acúmulo de mutações, gerado pelo tempo em que o ser humano moderno permaneceu naquele continente e principalmente ao efeito fundador das migrações para fora da África, gerando consecutivas reduções da variabilidade. Contudo, o microsatélite tetralocal *DYS503*, localizado no braço longo do cromossomo Y, apresentou menor diversidade genotípica para um grupo de afro-americanos em relação a americanos caucasianos. Esse padrão diferenciado despertou nossa curiosidade.

Devido ao caráter haplóide do cromossomo Y e por este não sofrer recombinação na maior parte de sua extensão, descartamos a possibilidade de fatores seletivos, visto que a seleção atuaria sobre todo o bloco genômico. Logo, construímos a hipótese de que há variação do número de cópias desse microsatélite na população africana e que a população fundadora, responsável pelo povoamento do resto do mundo, carregue apenas o genótipo com quatro cópias.

Para responder a essa pergunta, traçamos os seguintes objetivos:

Objetivo geral

- Testar a hipótese de existência de variação do número de cópias do loco *DYS503* na população africana.

Objetivos específicos

- Construir o mapa físico do *DYS503*, determinando os blocos repetitivos e suas posições cromossômicas;
- Desenvolver estratégia de tipagem quantitativa de número de cópias do marcador *DYS503* utilizando a técnica de QF-PCR (Quantitative Fluorescent PCR).



Material e Métodos

B. MATERIAL E MÉTODOS

B.1 Construção do mapa físico do *DYS503*

A estrutura palindrômica da região existente no braço longo do cromossomo Y, especialmente na região AZFc se deve a inversões, duplicações e deleções de grandes fragmentos de DNA. Por isso, a região onde se encontra o microssatélite *DYS503* pertence a um bloco maior, constituído da região variável – onde está o microssatélite propriamente dito – e uma região constante e conservada, que flanqueia a região variável (Figura 6). Entretanto, pouco se sabe sobre o tamanho desse bloco e como ele surgiu em várias cópias.

Com o objetivo de compreender a natureza múltipla do bloco no qual se encontra o microssatélite *DYS503* e traçar os possíveis eventos cromossômicos que possam ter gerado a variação em seu número de cópias, foi construído um mapa físico a partir das sequências genômicas existentes e da localização dos palíndromos da região AZFc (Kuroda-Kawaguchi *et al.*, 2001). Com isso também foi possível desenhar iniciadores que amplifiquem e localizem as cópias dessa região.

A partir da montagem de referência primária do genoma humano (*human build GRCh37*), disponível no banco de dados do *National Center for Biotechnology Information (NCBI, www.ncbi.nih.gov)*, a sequência total do cromossomo Y humano foi isolada e as quatro cópias do microssatélite *DYS503* foram localizadas com o auxílio dos iniciadores responsáveis por sua amplificação (Hanson e Ballantyne, 2007). Com o auxílio dos programas *PCR in silico* (<http://genome.ucsc.edu>) e *BLAST*

(*Basic Local Alignment Search Tool*, Zhang *et al.*, 2000; <http://blast.ncbi.nlm.nih.gov>), sequências de tamanhos crescentes (tendo o microssatélite como ponto central) foram comparadas à sequência do banco de dados, até onde fosse possível obter homologia máxima entre todas as quatro cópias.

O *PCR in silico* simula o anelamento dos iniciadores direto e reverso desenvolvidos pelo usuário quando comparados ao genoma de interesse. Caso haja ligação estável, o programa fornece a posição cromossômica e tamanho do(s) fragmento(s) gerado(s) pelos iniciadores, bem como o valor de T_m (temperatura para a qual 50% do iniciador está corretamente anelado ao DNA molde) dos mesmos. Utilizando essa plataforma, foram desenhados iniciadores para as extremidades das sequências constantes vicinais ao microssatélite *DYS503*. Em seguida era verificado se havia amplificação de quatro cópias: em caso positivo, o alvo dos iniciadores era afastado do ponto inicial, até que o resultado indicasse ausência das quatro cópias. Nessa região, portanto, há uma das extremidades do bloco. O mesmo procedimento foi feito no sentido oposto da sequência. Com a conhecida ferramenta *BLAST*, comparamos os fragmentos gerados pelo programa *PCR in silico* com a montagem atual do genoma humano, para confirmar a existência das quatro cópias. Com isso, são construídos grandes blocos gênicos, homólogos entre si e contendo o microssatélite.

Depois de concluída a correta montagem do mapa e confirmada a posição cromossômica dos blocos, também foi possível pesquisar a existência de CNVs descritos para nossa região de interesse. Para isso, utilizamos o *Database of Genomic Variants* (lafrate *et al.*, 2004; <http://projects.tcag.ca/variation>), que reconhece os fragmentos de PCR fornecidos pelo usuário e os compara com seu

banco de dados de variações de número de cópias. Caso haja informação correspondentes, o programa detalha a localização, tamanho da sequência variante e frequência na população estudada.

AACTCTTCCTTATTTCTGAAGTGAAAAGTAATCACTCCAGCACATTAATTGTGCAATATCTTGTACCTGC
 TTCTTGAAAATATACACATACTTGTTTTCTTCTGTTTCTATCTCTTCCAGGGAAGGCCCATATTTCAACAA
 ATTGTCTGTTCCCTCTGCATCATGGCACGGGTCCCTACATCTCAAATGTTTAAATAGACCTCATTCTCAG
 AAACAATTATGCTGACCTTTCTCTTCCAGTTCTACCCATAACCTATTTTTTATGAGTAAGCAATGCAAA
 TATACAGAAGAATTTATAGGTTCCACCAGAGGGTTCATGAACCTTTTTGCTCTGCCTTGTCTAAAGTCATC
 AGCAGGACTGAGTTAGTTTTCAAAAACAGGAGAAAAGCAACACAGTTTTTTCTGTCTCCACCCACCAT
 CTCTTCTCAATCCTTCTCATGAAATTTAAATACCTTAAGAGACAGTTAATCATTCTGCAAGTAGAAAGG
 CAACCTTCTGTGTGCTAGAAGAACAAGAAGAAAAAAGTTACTTATAAATGAGATTGCTTTCTTATATGAA
 ATAGTAATTTCAAGATAGTAATATTATTTAAGCACCTCTTTATTAAGGAAGAAAGGAAGCAAAACTACTT
 CACTGCTCTGGTTTTATATTTGTACCTACGTGGTGGAATACTGAAACCTCTTTGGACACAACGGAAAGAA
 ATGAAGTATGTCCATTGTGCCCTTGCATCAGAGAGAGCATA**TTTCAGCCTTTTCAATTTCCA**TTGAGGC
 TAGAAGAAGTGAAATGATGAATATAAAAAATTCTCAGTAACACCAATGTTACTCACTATGCTGCTAATA
 ATGTTTTTTTATATT
 ATGTTTTTTTATATT
CTGGAGTGCAGTTGAGCAGTCAGCTTAATGCAACCTCCACCCCTGGGTTCAATTGATTCTCCTTCCCTCA
 GCCTCACAAGTAGGTGGGATTACAAGCCCATGACACCATGCCTGGCTAATTTTTGTATTTTTTAGTAGAGA
 TGGTGTTCACCACGTTAGCCAGGATGGTCTCAATCTCCTGACCTCATGATGCACTCATCTCAGCCTCCA
 AAAGAGCCACTGCACCCAGTAAATAATTTCTAAATCCAGTATTATAATCTACATTCCACCTCAGAGGCAG
 AACACCAAGTTATTCAGTTGAATTCCTTTATCTTATTACCTTGTTTCAGTCCCTTTCCAGGTCTGTGAC
 AAATGGCTGCCGGTCAAACTCCCTCTTGTAGCCATGGCCCTACTCATCTGCTTCAGCCCAACTCTACT
 ACAGTCTTGGGTAATGGTTTGCTGACCATCTGCCTCAGTGTCTAGAAATTCAGGATGACACGAGCTTGGT
 TTGATTCAGTGTGTATGATCAGCCCCAAGGAAACTGCCCAGCAGAGACAGGCTCTAGATCAGGTCTTAC
 AGAGTGGTTTCCAAAACGCATTTATTCTGGAATAAGCAAAATCAAACCTCTCAAAATCCCTTGCCCTTTT
 AAGTAATTCTATATGTTACTGAATGCCATGACTATACCCAAGAAACTGCTAACCAGACATTTTGAAATA
 TTTTTAAGATGTATTTCCAATGTTTGGGAGGGAGTAAGAGGTGGCTTTCTTGAGAAAATTACTTTCATA

Figura 6: Fragmento do bloco genômico contendo a região variável (em amarelo) e um pequeno pedaço da porção constante (em azul) onde está o microssatélite *DYS503*. A região de anelamento dos iniciadores está marcada em vermelho.

B.2 Estudo da variação do número de cópias

B.2.1 Amostras de DNA

Para a execução deste projeto foram utilizadas amostras de DNA pertencentes ao painel do CEPH/HGDP - *Centre d'Etude du Polymorphisme Humain/Human Genome Diversity Project* (Cann *et al.*, 2002). Este painel foi construído a fim de padronizar os estudos genéticos, populacionais e biomédicos, tornando a comparação dos dados mais confiável e estatisticamente válida (Cavalli-Sforza, 2005). Neste painel, O DNA é extraído de linhagens de células linfoblastóides, liofilizado, acondicionado em placas de 96 poços e enviado aos centros de pesquisa. Para a execução deste projeto foram selecionados 101 indivíduos do sexo masculino (Tabela 1), de 6 populações, presentes em 7 regiões geográficas do continente africano (Figura 7).

Além disso, também foram genotipadas outras 32 amostras de homens do mesmo painel, para os quais a tipagem prévia do microssatélite *DYS503* apontou quatro alelos diferentes (Tabela 2). O objetivo dessa prática foi averiguar se há diferenças significativas do número de cópias entre esses indivíduos e as amostras africanas.

Uma amostra padrão (denominada SP) também foi amplificada em todas as placas de reações, servindo como referência normalizadora da tipagem tetralocal, ou seja, possui quatro alelos distintos para o microssatélite *DYS503*, permitindo a comparação com as amostras e sua posterior equivalência nos cálculos de número de cópias.

Tabela 1 – Identificação, população e origem geográfica dos indivíduos estudados.

Identificação CEPH	População	Origem Geográfica
HGDP00448	Pigmeus Biaka	República da África Central
HGDP00461	Pigmeus Biaka	República da África Central
HGDP00464	Pigmeus Biaka	República da África Central
HGDP00465	Pigmeus Biaka	República da África Central
HGDP00466	Pigmeus Biaka	República da África Central
HGDP00469	Pigmeus Biaka	República da África Central
HGDP00470	Pigmeus Biaka	República da África Central
HGDP00472	Pigmeus Biaka	República da África Central
HGDP00473	Pigmeus Biaka	República da África Central
HGDP00475	Pigmeus Biaka	República da África Central
HGDP00477	Pigmeus Biaka	República da África Central
HGDP00479	Pigmeus Biaka	República da África Central
HGDP00451	Pigmeus Biaka	República da África Central
HGDP00452	Pigmeus Biaka	República da África Central
HGDP00453	Pigmeus Biaka	República da África Central
HGDP00454	Pigmeus Biaka	República da África Central
HGDP00455	Pigmeus Biaka	República da África Central
HGDP00457	Pigmeus Biaka	República da África Central
HGDP00458	Pigmeus Biaka	República da África Central
HGDP00459	Pigmeus Biaka	República da África Central
HGDP00460	Pigmeus Biaka	República da África Central
HGDP00981	Pigmeus Biaka	República da África Central
HGDP00985	Pigmeus Biaka	República da África Central
HGDP00986	Pigmeus Biaka	República da África Central
HGDP01086	Pigmeus Biaka	República da África Central
HGDP01087	Pigmeus Biaka	República da África Central
HGDP01088	Pigmeus Biaka	República da África Central
HGDP01089	Pigmeus Biaka	República da África Central
HGDP01090	Pigmeus Biaka	República da África Central
HGDP01091	Pigmeus Biaka	República da África Central
HGDP01092	Pigmeus Biaka	República da África Central
HGDP01093	Pigmeus Biaka	República da África Central
HGDP01094	Pigmeus Biaka	República da África Central
HGDP00449	Pigmeus Mbuti	República Democrática do Congo
HGDP00450	Pigmeus Mbuti	República Democrática do Congo
HGDP00456	Pigmeus Mbuti	República Democrática do Congo
HGDP00462	Pigmeus Mbuti	República Democrática do Congo
HGDP00463	Pigmeus Mbuti	República Democrática do Congo
HGDP00467	Pigmeus Mbuti	República Democrática do Congo
HGDP00468	Pigmeus Mbuti	República Democrática do Congo
HGDP00474	Pigmeus Mbuti	República Democrática do Congo
HGDP00478	Pigmeus Mbuti	República Democrática do Congo
HGDP00982	Pigmeus Mbuti	República Democrática do Congo
HGDP00983	Pigmeus Mbuti	República Democrática do Congo
HGDP00984	Pigmeus Mbuti	República Democrática do Congo
HGDP01081	Pigmeus Mbuti	República Democrática do Congo
HGDP00904	Mandenka	Senegal
HGDP00905	Mandenka	Senegal
HGDP00906	Mandenka	Senegal
HGDP00907	Mandenka	Senegal
HGDP00908	Mandenka	Senegal

HGDP00911	Mandenka	Senegal
HGDP00912	Mandenka	Senegal
HGDP00913	Mandenka	Senegal
HGDP00919	Mandenka	Senegal
HGDP01199	Mandenka	Senegal
HGDP01200	Mandenka	Senegal
HGDP01202	Mandenka	Senegal
HGDP01283	Mandenka	Senegal
HGDP01284	Mandenka	Senegal
HGDP01285	Mandenka	Senegal
HGDP01286	Mandenka	Senegal
HGDP00923	Yoruba	Nigéria
HGDP00927	Yoruba	Nigéria
HGDP00929	Yoruba	Nigéria
HGDP00930	Yoruba	Nigéria
HGDP00931	Yoruba	Nigéria
HGDP00932	Yoruba	Nigéria
HGDP00936	Yoruba	Nigéria
HGDP00937	Yoruba	Nigéria
HGDP00940	Yoruba	Nigéria
HGDP00941	Yoruba	Nigéria
HGDP00942	Yoruba	Nigéria
HGDP00943	Yoruba	Nigéria
HGDP00944	Yoruba	Nigéria
HGDP01405	Bantu NE*	Quênia
HGDP01406	Bantu NE	Quênia
HGDP01408	Bantu NE	Quênia
HGDP01411	Bantu NE	Quênia
HGDP01412	Bantu NE	Quênia
HGDP01413	Bantu NE	Quênia
HGDP01415	Bantu NE	Quênia
HGDP01416	Bantu NE	Quênia
HGDP01417	Bantu NE	Quênia
HGDP01418	Bantu NE	Quênia
HGDP01419	Bantu NE	Quênia
HGDP00987	San	Namíbia
HGDP00988	San	Namíbia
HGDP00991	San	Namíbia
HGDP00992	San	Namíbia
HGDP01092	San	Namíbia
HGDP01032	San	Namíbia
HGDP01036	San	Namíbia
HGDP00993	Bantu SE* Pedi	África do Sul
HGDP00994	Bantu SE S. Sohto	África do Sul
HGDP01030	Bantu SE Tswana	África do Sul
HGDP01034	Bantu SE Tswana	África do Sul
HGDP01033	Bantu SE Zulu	África do Sul
HGDP01028	Bantu SO* Herero	África do Sul
HGDP01035	Bantu SO Herero	África do Sul
HGDP01031	Bantu SO Ovambo	África do Sul

* NE, SE e SO referem-se aos pontos cardeais *nordeste*, *sudeste* e *sudoeste*, respectivamente.



- 1 Pigmeus Biaka – República da África Central
- 2 Pigmeus Mbuti – República Democrática do Congo
- 3 Mandenka – Senegal
- 4 Yoruba – Nigéria
- 5 Bantu NE – Quênia
- 6 San – Namíbia
- 7 Bantu SE/SO – África do Sul

Figura 7: Mapa do continente africano, com indicação numérica da localização geográfica das populações estudadas.

Tabela 2 – 32 indivíduos selecionados para controle, com 4 alelos distintos na genotipagem do microssatélite *DYS503*.

Identificação CEPH	População	Origem Geográfica		Alelos <i>DYS503</i>		
HGDP00603	Druze	Israel (Carmel)	10	13	14	17
HGDP00678	Palestina	Israel (Central)	12	13	14	15
HGDP00058	Balochi	Paquistão	12	13	14	15
HGDP00068	Balochi	Paquistão	12	13	14	15
HGDP00143	Makrani	Paquistão	11	14	14	15
HGDP00181	Sindhi	Paquistão	12	13	14	15
HGDP01214	Daur	China	11	13	14	15
HGDP01227	Mongola	China	12	13	14	15
HGDP01233	Hezhen	China	12	13	14	15
HGDP01247	Xibo	China	12	13	14	15
HGDP01311	Dai	China	12	13	14	15
HGDP00755	Japonesa	Japão	11	12	13	14
HGDP00762	Japonesa	Japão	11	12	13	14
HGDP00763	Japonesa	Japão	11	12	13	14
HGDP00768	Japonesa	Japão	11	13	14	15
HGDP00711	Cambojana	Camboja	12	13	14	15
HGDP00512	Francesa	França	11	12	14	15
HGDP00533	Francesa	França	12	13	14	15
HGDP01371	Basca	França	11	12	14	15
HGDP00665	Sardenha	Itália	12	13	14	15
HGDP01155	Norte da Itália	Itália (Bergamo)	11	13	14	15
HGDP00896	Russa	Rússia	9	12	13	14
HGDP00897	Russa	Rússia	13	14	15	16
HGDP01037	Pima	México	12	13	14	15
HGDP01039	Pima	México	12	13	14	15
HGDP01040	Pima	México	12	13	14	15
HGDP01042	Pima	México	12	12	14	15
HGDP01043	Pima	México	11	12	13	14
HGDP01045	Pima	México	12	13	14	15
HGDP01057	Pima	México	12	13	14	15
HGDP01060	Pima	México	12	13	14	15
HGDP01061	Pima	México	12	13	14	15

B.2.2 Reação em Cadeia da Polimerase Quantitativa Fluorescente (QF-PCR)

A reação em cadeia da polimerase (PCR) é um método altamente sensível e específico para a detecção e amplificação de fragmentos de DNA, mas não é primariamente uma técnica quantitativa, pois pequenas variações nas condições iniciais podem afetar profundamente a quantidade de produto obtido (Pena, 1998). Desde que foi criada, na década de 80, a PCR é continuamente aprimorada, tanto em seus fundamentos como em seus métodos de análise. Tais melhorias têm como objetivo aumentar a sensibilidade e especificidade da análise, bem como proteger o operador de possíveis danos relacionados à biossegurança. Um claro exemplo é a utilização de marcadores fluorescentes para genotipagem e sequenciamento de amostras, técnica criada também nos anos 80 e que abandona o laborioso método de coloração por nitrato de prata e o uso de brometo de etídio, além de proporcionar maior sensibilidade, automatização e menor risco.

O desenvolvimento de técnicas fluorescentes na análise do DNA se iniciou em meados da mesma década, sob a direção de Lloyd M. Smith. Através da adição de grupos químicos fluorescentes aos iniciadores para a PCR e da automatização dos processos de eletroforese, foi construído o alicerce dos atuais procedimentos de sequenciamento e genotipagem (Smith *et al.*, 1986). Rapidamente, os diversos procedimentos da biologia molecular foram acoplados ao novo método de análise. Utilizando plasmídeos marcados com fluorescência e enzimas de restrição, Carrano e colaboradores (1989) aprimoraram a metodologia, aumentando a resolução de base única para até 400 pares de bases e permitindo a quantificação de fragmentos de até 2000 pares de bases.

Em busca de um procedimento que detectasse a mutação pontual 1226 no gene da glicocerebrosidase, Zimran e colaboradores (1989) desenvolveram uma técnica baseada na introdução de iniciadores com diferentes fluoróforos em sua PCR: para a sequência normal, fluoresceína; para a sequência mutada, rodamina. Com a análise da altura dos picos para as diferentes fluorescências, foram capazes de quantificar o material amplificado diferencialmente e discernir indivíduos normais daqueles afetados pela mutação, tanto homocigotos quanto heterocigotos.

A partir de então, com o advento da alta tecnologia e a melhoria dos sistemas automatizados de eletroforese e sequenciamento de DNA, a utilização desta técnica foi sensivelmente ampliada, permitindo seu imediato uso no estudo de anomalias como aneuploidias dos cromossomos 13 (Toth *et al.*, 1998), 18 (Pertl *et al.*, 1996), 21 (Lee *et al.*, 2004; Mansfield, 1993; Valero *et al.*, 1999), X (Mutter e Pomponio, 1991) e Y (Adinolfi *et al.*, 1995; Pertl *et al.*, 1996), bem como defeitos de genes únicos (Blake *et al.*, 1999).

A principal vantagem da QF-PCR sobre o método convencional é a possibilidade da amplificação simultânea da sequência-alvo e da sequência de referência, que compartilham os sítios de ligação dos mesmos iniciadores, gerando uma cinética competitiva (Pena, 1998). Além disso, outras características são de igual relevância: aumento da sensibilidade em quase 1000 vezes; menor quantidade de ciclos para obtenção de produto viável à análise; sinais fracos (imperceptíveis quando observados em gel) são facilmente observados e redução do uso de produtos tóxicos (Findlay e Quirke, 1996). Um bom exemplo das melhorias proporcionadas pela QF-PCR é, por exemplo, a tipagem de microssatélites, já que a análise de géis

de agarose e poliacrilamida é também laboriosa e confusa, já que pode haver formação de *heteroduplex* e sobreposição de fragmentos de tamanhos próximos.

Assim sendo, a *QF-PCR* é a opção mais confiável e precisa, pois sua alta resolução permite a distinção de alelos consecutivos mesmo para unidades repetitivas curtas, bem como a utilização de vários marcadores (*multiplex*) numa mesma reação, pela quantificação do produto amplificado pela análise da área ou altura dos picos gerados. Este último parâmetro – área e altura dos picos referentes aos fragmentos – constitui a base da quantificação proposta pela *QF-PCR*. A partir dos valores obtidos para um material genético de referência, é possível comparar as medidas encontradas para a amostra em questão e estabelecer relações matemáticas simples que permitam mensurar a quantidade do produto amplificado em relação ao padrão. Devido a sua rapidez e eficiência, a adição de marcadores fluoróforos aos iniciadores é rotina dos laboratórios, onde é largamente aplicada em testes de paternidade, casos forenses, diagnósticos de anomalias genéticas e estudos de diversidade populacional.

B.2.3 Desenvolvimento dos iniciadores

Utilizando a sequência do genoma do cromossomo Y, o microssatélite foi localizado, com suas porções variável e constante. Com a ferramenta *BLAST*, o bloco inteiro foi confrontado contra todo o genoma humano. Entre os vários *hits* encontrados, foram procuradas posições fora do cromossomo Y que, em comparação com a sequência original, apresentassem alguma inserção ou deleção. Com isso, é possível definir

tamanhos diferentes para os fragmentos múltiplos do cromossomo Y e o fragmento referência de cópia única no sequenciador automático.

Tão logo fossem encontradas, essas regiões eram avaliadas a fim de averiguar a possibilidade de desenho dos iniciadores. Para tanto, era necessário encontrar sequências vicinais à região diferenciada com cerca de 15 a 20 nucleotídeos, ricas em guanina e citosina (para obtenção de um T_m ótimo).

A viabilidade para amplificação foi testada com o programa online *PCR in silico*, que compara a sequência dos iniciadores com todo o genoma humano e verifica a possibilidade de anelamento estável. Em caso positivo, o programa indica a posição cromossômica e tamanho do(s) fragmento(s) gerado(s). Acoplado ao PCR in silico, o programa *Primer3* (Rozen e Skaletsky, 2000; <http://frodo.wi.mit.edu/primer3>) calcula o T_m do par de iniciadores. Finalizando, foi utilizada a plataforma *OligoAnalyzer* (*Integrated DNA Technologies*, <http://www.idtdna.com/analyzer>) para garantir a ausência de formação de heterodímeros ou grampos (“*hairpins*”) para os iniciadores.

Ao final, foi obtido um único par de iniciadores (Tabela 3), capaz de amplificar simultaneamente a região de referência de cópia única e as porções conservadas do *DYS503*, gerando fragmentos de tamanhos diferentes (Figura 8) e proporcionais à quantidade de cópias.

Tabela 3 – Sequências dos iniciadores utilizados na genotipagem

Iniciador	Sequência	Tm	Tamanho dos fragmentos
Y503-22F	5'- GTTTTCCAGTCACGACT GGTGATTAATGTCAAAGCA-3'	56.8 °C	CromY: 322 pb
Y503-22R	5'-AATATTT CAGTGTGATTATTCCTAGG -3'	56.3 °C	Crom22: 308 pb

* Em negrito, a sequência M13-40, adicionada à extremidade 5' do iniciador direto original.


```

1      10      20      30      40      50      60      70      80
|-----+-----+-----+-----+-----+-----+-----+-----|
Crom22  TGGTGATTAAATGTCAAAAGCAGGAAAAGAAATAAATATGAGTAGAAAAGTAATCAATAAATTAGAAAAACAGGTGGC
CromY_Fragmento1 TGGTGATTAAATGTCAAAAGCACA AAAAGAAA A AATACGAGTAGAAAAGTAATGAATCAATTAGAAAAACAGGTGAC
CromY_Fragmento2 TGGTGATTAAATGTCAAAAGCACA AAAAGAAA A AATACGAGTAGAAAAGTAATGAATCAATTAGAAAAACAGGTGAC
CromY_Fragmento4 TGGTGATTAAATGTCAAAAGCACA AAAAGAAA A AATACGAGTAGAAAAGTAATGAATCAATTAGAAAAACAGGTGAC
CromY_Fragmento3 TGGTGATTAAATGTCAAAAGCACA AAAAGAAA A AATACGAGTAGAAAAGTAATGAATCAATTAGAAAAACAGGTGAC
Consensus TGGTGATTAAATGTCAAAAGCacaAAAAGAAAaAATAcGAGTAGAAAAGTAATgAATAcAATTAGAAAAACAGGTGac

81     90     100    110    120    130    140    150    160
|-----+-----+-----+-----+-----+-----+-----+-----|
Crom22  AGGTGGGGAGACAGAGAGAAARACAAAGGCAGATAGACAGGAGGCRAAGGCAGAACCAA-AAAATCTAGGGTAATAAA
CromY_Fragmento1 GGGTGGGGAAGGGAGAGAGAAACAAAGGCAGATAG----GAGGCRAAGGCAGAGCCAAAGAAAATCTAGGGTAATGAA
CromY_Fragmento2 GGGTGGGGAAGGGAGAGAGAAACAAAGGCAGATAG----GAGGCRAAGGCAGAGCCAAAGAAAATCTAGGGTAATGAA
CromY_Fragmento4 GGGTGGGGAAGGGAGAGAGAAACAAAGGCAGATAG----GAGGCRAAGGCAGAGCCAAAGAAAATCTAGGGTAATGAA
CromY_Fragmento3 GGGTGGGGAAGGGAGAGAGAAACAAAGGCAGATAG----GAGGCRAAGGCAGAGCCAAAGAAAATCTAGGGTAATGAA
Consensus gGGTGGGGAaaggGAGAGAAgAACAAAgGCAGATAG...GAGGCRAAGGCAGAgcCAAAGAAAATCTAGGGTAATgAA

161    170    180    190    200    210    220    230    240
|-----+-----+-----+-----+-----+-----+-----+-----|
Crom22  AAACCTGCTAAGCAGGTATAGTTATGTAGATAAAGAGTAACCATATTAGATGTAAGCTGTTTAAACTTATGTTTAAAC
CromY_Fragmento1 AAACCTGACTGAGCAGGTGTTAGTTATTTAGATAAGAGAGTAACCATATTAGATGTAAGCTGTTTAAACTCATGTGTTAAA
CromY_Fragmento2 AAACCTGACTGAGCAGGTGTTAGTTATTTAGATAAGAGAGTAACCATATTAGATGTAAGCTGTTTAAACTCATGTGTTAAA
CromY_Fragmento4 AAACCTGACTGAGCAGGTGTTAGTTATTTAGATAAGAGAGTAACCATATTAGATGTAAGCTGTTTAAACTCATGTGTTAAA
CromY_Fragmento3 AAACCTGACTGAGCAGGTGTTAGTTATTTAGATAAGAGAGTAACCATATTAGATGTAAGCTGTTTAAACTCATGTGTTAAA
Consensus AAACCTGaCTGAGCAGGTgtLAgTTATtTAGATAAgAGAGTAACCATATTAGATGTAAGCTGTTTAAACTcATGTgtLAAa

241    250    260    270    280    290    300    309
|-----+-----+-----+-----+-----+-----+-----|
Crom22  AT-----TAGATTGGTAAAAATTTCTTCTGCCTAGGAATAATCACACTGAAATATT
CromY_Fragmento1 ATATAAATACTCTACCGTTAGATTGATAAAAAATTTGCTTTCACCTAGGAATAATCACACTGAAATATT
CromY_Fragmento2 ATATAAATACTCTACCGTTAGATTGATAAAAAATTTGCTTTCACCTAGGAATAATCACACTGAAATATT
CromY_Fragmento4 ATATAAATACTCTACCGTTAGATTGATAAAAAATTTGCTTTCACCTAGGAATAATCACACTGAAATATT
CromY_Fragmento3 ATATAAATACTCTACCGTTAGATTGATAAAAAATTTGCTTTCACCTAGGAATAATCACACTGAAATATT
Consensus ATataaaataactctaccgtTAGATTGaTAAAAATTTgCTTtcaCCTAGGAATAATCACACTGAAATATT

```

Figura 8: Alinhamento gerado pelo programa online *MultAlin*, com os 5 fragmentos gerados pelo único par de iniciadores: quatro para o cromossomo Y (CromY) e o fragmento referência de cópia única do cromossomo 22 (Crom22). Notar as variações de nucleotídeos nas posições 117-120 (nos fragmentos do cromossomo Y), 141 e 243-259 (no fragmento do cromossomo 22) que geram a diferença de tamanho necessária para a visualização no sequenciador automático.

B.2.4 Amplificação

O uso de fluorescência para a marcação dos produtos da PCR é a base de nossa técnica. Entretanto, devido ao alto custo de iniciadores marcados diretamente com fluorescência, adaptamos uma técnica mais barata e de igual eficiência (Schuelke, 2000): adicionamos à extremidade 5' do iniciador direto a sequência M13-40 (5'-GTTTTCCCAGTCACGAC-3'). Foi incluído na reação um novo iniciador, com a mesma sequência do M13-40, conjugado a uma molécula de amidato de fluoresceína (FAM). Após a desnaturação inicial, no primeiro ciclo, o iniciador direto se anela à fita molde, deixando a “cauda” com a sequência M13-40 sem qualquer ligação. No segundo ciclo, o iniciador reverso se anela à fita recém-sintetizada que contém M13-40, gerando sua sequência complementar, à qual se liga o iniciador com o componente fluoróforo, permitindo que a próxima fita a ser sintetizada possua a marcação com FAM (Figura 9).

Para a reação de PCR, 1.5 µl de DNA molde foram amplificados em um volume final de 10.0 µl de reação. Cada reação continha 200 µM de cada um dos dNTPs (dATP, dCTP, dGTP, dTTP), 0.05 µM do iniciador direto Y503-22F, 0.5 µM do iniciador reverso Y503-22R e 0.5 µM da sequência direta M13-40 marcada com o fluoróforo FAM na posição 5', 1.0 unidade de *Taq Platinum* DNA polimerase (Invitrogen) em tampão de reação com 20 mM de Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂. O ciclo de reação foi elaborado em termociclador Bio-Rad My Cycler™ e consiste em uma desnaturação inicial a 94°C por 5 minutos, seguida de 35 ciclos de desnaturação a 94°C por 45 segundos, anelamento a 60°C** por 45 segundos e extensão a 72°C por 60 segundos. A extensão final foi dada a 72°C por 45 minutos.

Para todas as reações de PCR foi feito um controle negativo de amplificação, onde água estéril foi acrescentada no lugar do DNA, bem como um controle interno, com uma amostra padrão do painel CEPH/HGDP do sexo masculino e uma amostra de origem feminina. O mesmo procedimento foi repetido para a tipagem dos 32 indivíduos com 4 alelos distintos para o microssatélite *DYS503*.

** Para verificar a temperatura ótima de anelamento dos iniciadores conforme demonstrado na Tabela 3, foi realizada uma PCR com gradiente de temperatura, para a qual a temperatura de 60 °C mostrou o melhor resultado.

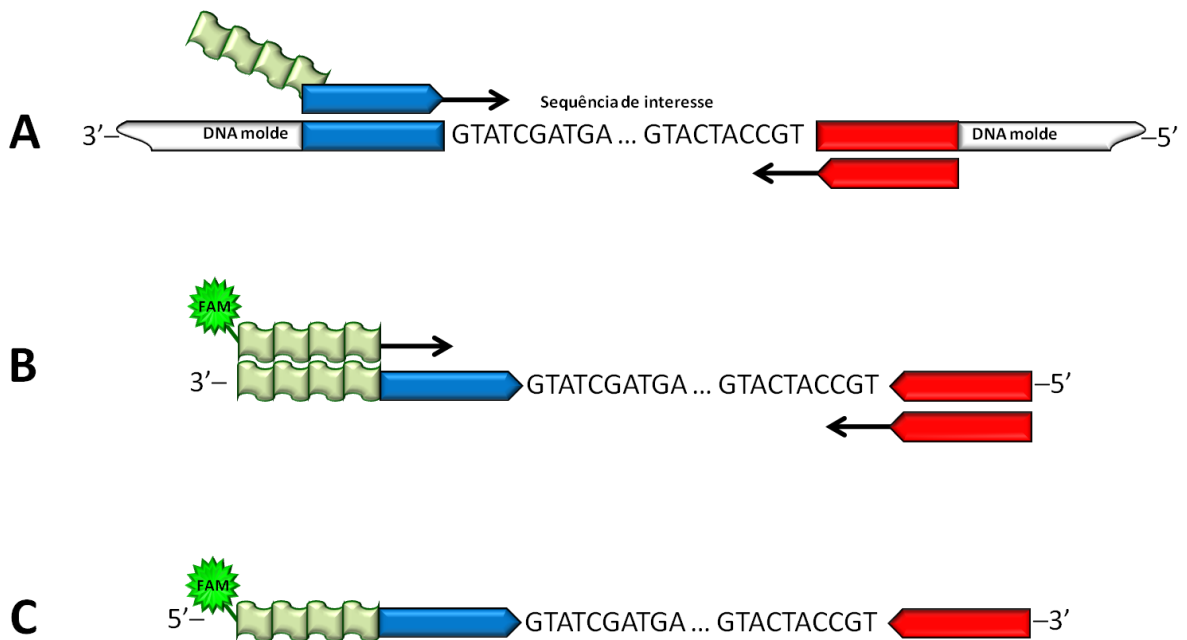
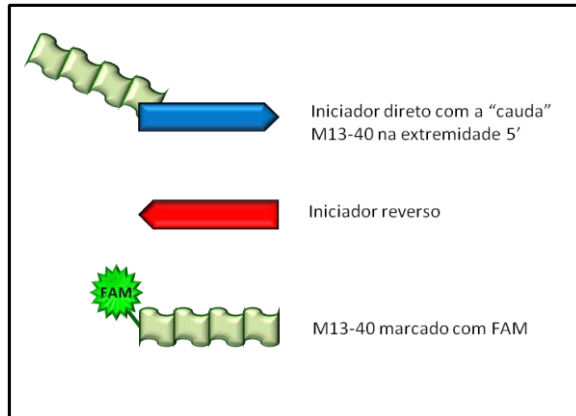


Figura 9: Esquema de amplificação utilizando a sequência M13-40 marcada com FAM. (A) No primeiro ciclo da PCR, o iniciador direto com M13-40 é incorporado aos produtos da PCR; (B) Esses produtos serão o alvo do iniciador M13-40 marcado com FAM, que é incorporado durante os ciclos subsequentes; (C) O produto final marcado, para posterior análise em sequenciadores automáticos (Adaptado de Schuelke, 2000).

B.2.5 Genotipagem

Para a genotipagem das amostras, 2.0 µL do produto da PCR foram distribuídos em placas de 96 poços de fundo cônico (GreinerBio). A cada poço também foram adicionados 8.0 µL do marcador de peso molecular *ET-550R Size Standard* (GE Healthcare), diluído a 1:32 em *Tween 20* 0.1%. As amostras foram desnaturadas a 95 °C por 3 minutos e rapidamente resfriadas a 4 °C. Em seguida, as placas foram injetadas no sequenciador automático *MegaBACE 1000 DNA Analyzer* (GE Healthcare) a 3 kV por 80 segundos. A eletroforese capilar dos fragmentos ocorreu em matriz LPA sob 10 kV de voltagem durante 75 minutos.

O aparelho *MegaBACE 1000* dispõe de um laser que emite luz à intensidade de 488 nm. Em certo ponto, durante a eletroforese, o raio laser incide sobre o capilar onde estão as amostras. O amidato de fluoresceína (FAM) presente no fragmento de PCR absorve a luz neste comprimento de onda, liberando a energia na forma de emissão de luz a 514 nm. Essa luz é captada por um tubo fotomultiplicador que amplifica o sinal, transmitindo-o ao computador, que processa os dados e os armazena. Graças à tecnologia de transferência de energia (Ju *et al.*, 1995), a luz emitida pelo FAM é capaz de excitar as moléculas de ROX adicionadas ao marcador de peso molecular, que liberam energia a um comprimento de onda de 610 nm, também captada pelo aparelho. Dessa forma, os sinais gerados pelo fragmento de PCR e pelo marcador são capturados ao mesmo tempo.

B.2.5.1 Análise da eletroforese

Após a eletroforese dos fragmentos, são gerados 96 arquivos *.rsd correspondentes a cada poço da placa. Com o programa *MegaBACE Genetic Profiler* (versão 2.2, Amersham Biosciences, 2003), cada arquivo é transformado em eletroferogramas baseados no tempo de retenção e na detecção distinta da fluorescência dos fragmentos e do marcador de peso molecular. O tamanho do produto da PCR determina sua posição no gráfico, ao passo que a quantidade de material amplificado determina a altura, largura e área de cada pico, dados automaticamente calculados pelo programa. Os valores de área dos picos referentes aos fragmentos de interesse foram transferidos para uma planilha. Em seguida, calculamos a razão do valor da área do pico referente ao cromossomo Y pelo valor encontrado para a cópia de referência do cromossomo 22.

B.2.5.2 Normalização e número de cópias

Para normalizar os dados, foi amplificada em todas as placas uma amostra padrão que apresentou 4 picos distintos pela genotipagem prévia do marcador *DYS503*. As razões obtidas para cada amostra foram matematicamente comparadas à esta amostra normalizadora.

Razão da amostra normalizadora	→	4 cópias
Razão da amostra CEPH/HGDP	→	x

onde x indica o valor normalizado de cópias em relação à amostra padrão.

B.2.6 Análise estatística

Todos os testes foram realizados em triplicata. A reprodutibilidade dos valores normalizados de área para o marcador Y503-22 entre os três testes foi validada através do teste de Correlação de Pearson (plataforma *Vassar Stats*, <http://faculty.vassar.edu/lowry/VassarStats.html>), cujo resultado – valor r – foi avaliado de acordo com a distribuição relativa ao número de amostras amplificadas (n). O resultado final é expresso na forma de probabilidade de significância (valor p), que considera estatisticamente válidas as correlações com valor menor que 0.05.

Também foi realizada a comparação entre as médias do número normalizado de cópias da população africana e o grupo de 32 indivíduos com quatro alelos distintos para o *DYS503* através do teste t de Student, com o Microsoft Excel 2007. Além disso, foi realizado o teste de “goodness of fit” para ambos os grupos acima com o programa *GraphPad Prism 5*. Esse teste é utilizado para descrever a discrepância entre os valores observados e os valores esperados para a distribuição normal dos dados. O “goodness of fit” foi quantificado através do coeficiente de determinação (R^2). Quanto mais próximo de $R^2=1$, mais ajustado à curva está o conjunto dos dados. A avaliação estatística entre os agrupamentos de número de cópias (figuras 14 e 15) foi feita através do teste de *one-way ANOVA*, com correção de Newman-Kuels para múltipla comparação pós-teste.



Resultados

C. RESULTADOS

C.1 Mapa físico do microssatélite *DYS503*

A montagem do mapa físico do *DYS503* foi realizada através da composição de grandes blocos no qual o microssatélite está inserido. Para isso, buscamos toda a região conservada que flanqueia o microssatélite e se repete com o máximo de homologia nas quatro cópias. Todo o processo foi baseado na montagem de referência primária *GRCh37* do genoma humano, disponibilizada no banco de dados do NCBI. Dessa forma, foi possível determinar o tamanho e a localização cromossômica dos blocos.

Todos os quatro blocos têm cerca de 167 mil pares de bases (pb), com mais de 99% de homologia média entre eles (Tabela 4). Estão dispostos de forma assimétrica: A distância entre os blocos A e B é de 400596 pb; entre B e C, 875011 pb; entre C e D, 2264725 pb. O sentido das cópias é o mesmo entre A e C e tem sentido oposto para B e D, conforme indica a Figura 10.

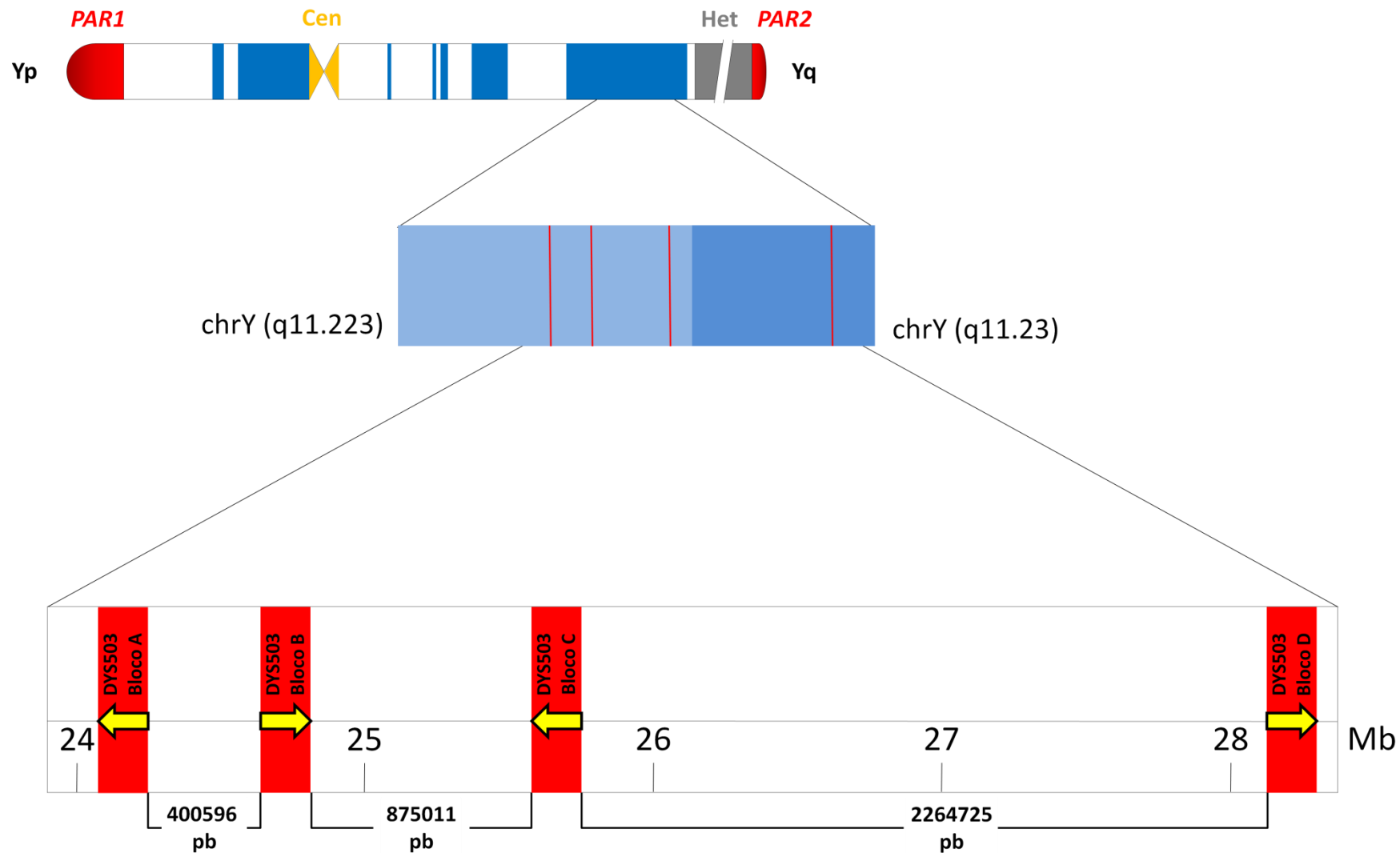


Figura 10: Posição cromossômica dos quatro blocos para o microsatélite *DYS503*. No topo da imagem, em vermelho, as regiões pseudoautossômicas 1 e 2 (PAR1/PAR2). Em amarelo, o centrômero (Cen). Em cinza, a heterocromatina (Het) foi condensada para melhor visualização gráfica. Yp, braço curto do cromossomo Y; Yq, braço longo do cromossomo Y.

Tabela 4 – Posição cromossômica dos blocos para o microssatélite *DYS503* no cromossomo Y

Bloco	Posição cromossômica			Tamanho
	<i>Banda</i>	<i>Início</i>	<i>Fim</i>	
A	q11.223	24,071,549	24,239,059	167510 pb
B	q11.223	24,639,655	24,806,678	167023 pb
C	q11.223	25,681,689	25,848,918	167229 pb
D	q11.23	28,113,643	28,281,107	167464 pb

C. 2 Mapa de variantes

O banco de dados do *Database of Genomic Variants* fornece informações sobre variações (perdas ou ganhos) de frações do genoma através de uma busca a partir do fragmento de PCR onde se encontra a região de interesse. Como resultado da busca, o *software* apresenta como resultado as variantes já descritas na literatura, com a posição cromossômica da variante e sua frequência na população estudada. Com isso, podemos averiguar se há alguma variação capaz de alterar significativamente a tipagem de nosso marcador.

A Figura 11 representa um mapa das variações construído a partir das informações do *Database of Genomic Variants*. As regiões onde estão os blocos A, B, C e D do microsatélite *DYS503* (setas) estão representadas respectivamente pelas cores vermelha, amarela, verde e azul. As variações encontradas acompanham as cores usadas para cada bloco. A variação apresentada na cor roxa engloba os blocos B, C e D.

De um modo geral, foram encontradas alterações pontuais e únicas para cada variação. As frequências para quase todas as variantes encontradas foram muito baixas (entre 0.37 – 2.22%). Entretanto duas delas – variações 4184 e 4185 – tiveram frequência relativamente elevada. Ambas foram avaliadas em um grupo de 270 indivíduos pertencentes ao painel *HapMap* (Redon, 2006). Para a primeira delas foi identificado 23.33% de perda ou ganho da região, enquanto para a segunda houve 15.18% de variação.

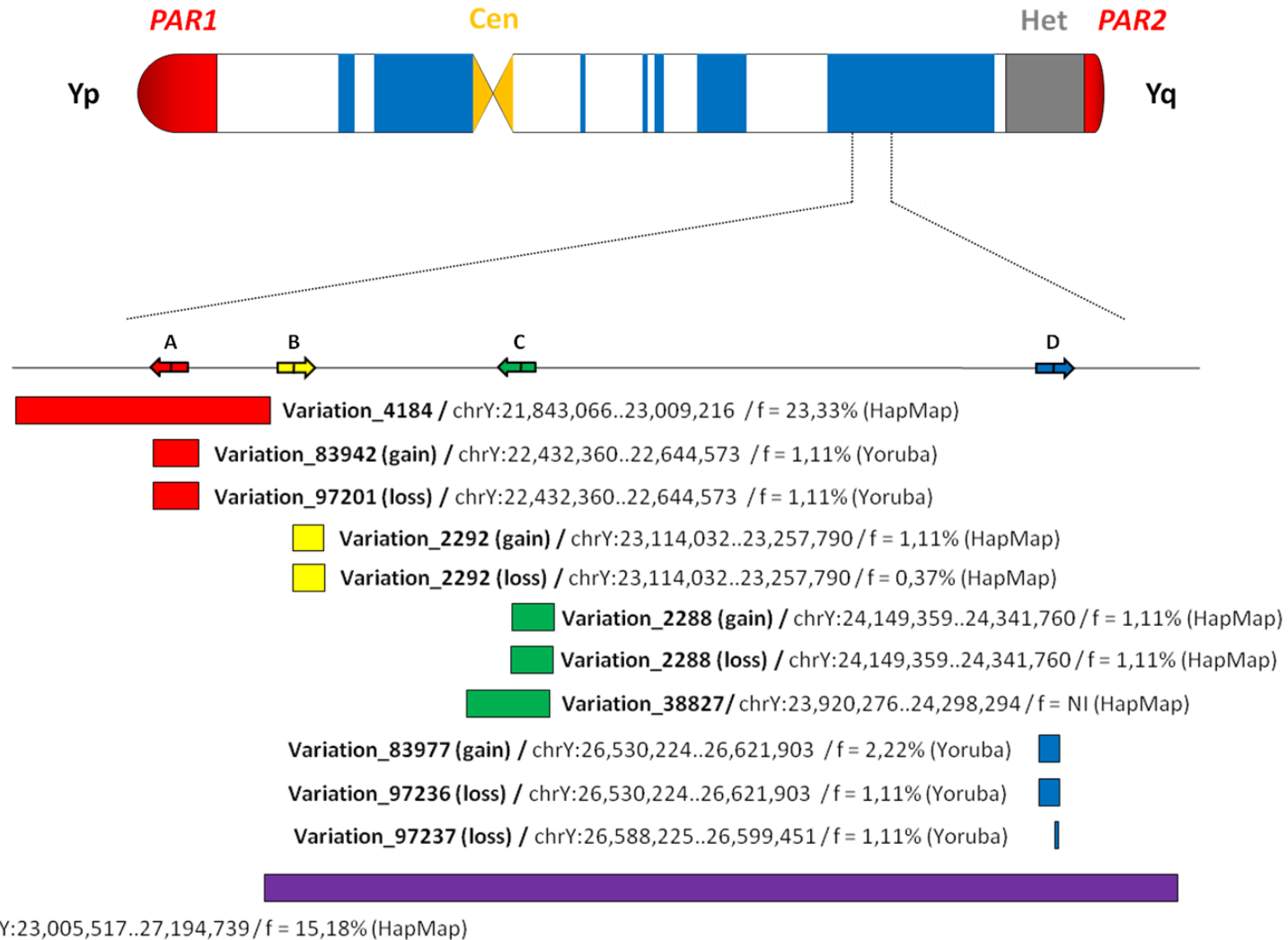


Figura 11: Mapa de unidades variantes encontradas dentro dos blocos de *DYS503*, segundo o banco de dados *Database of Genomic Variants*. A cada variante está associada a nomenclatura de acordo com o banco de dados, o tipo de variação (*gain* = ganho, *loss* = perda), a localização cromossômica e a sua frequência para a população estudada (entre parênteses).

C.3 Genotipagem das amostras africanas do painel do CEPH/HGDP

A tipagem das amostras dos homens africanos presentes no painel do CEPH/HGDP foi realizada através da análise das curvas geradas após a eletroforese dos fragmentos de PCR fluorescentes (Figura 12). Comparando os picos diferencialmente apresentados referentes aos cromossomos Y e 22, foi possível calcular a razão entre suas áreas. Devido à pouca quantidade de DNA para algumas amostras, 94 dos 101 indivíduos puderam ser apurados. Através da comparação da razão entre áreas com a razão entre alturas, obtivemos um aspecto aproximadamente linear, que confirma que qualquer uma das variáveis pode ser usada, conforme observado na Figura 13.

Os resultados expressos pela razão das áreas estão indicados na Figura 14. Os valores foram agrupados em intervalos de 0.5 para melhor visualização e agrupamento da frequência do número de cópias. Analisando o conjunto normalizado de dados, há significativa concentração de valores entre 3.0 e 4.0 e a curva de tendência não apresenta aspecto polimodal, ou seja, os dados se aglomeram em um ponto único no gráfico, afastando a possibilidade da existência significativa de variação de número de cópias na população estudada.

Para averiguar a reprodutibilidade entre os experimentos para a população africana, foram calculados os valores da correlação de Pearson (r) e a significância (p) entre as triplicatas. Em todas as comparações feitas, o valor de r foi sempre maior que 0.3 e o valor p permaneceu menor que 0.05, ou seja, todos os testes foram significativos em nível de 95% do intervalo de confiança. A média dos valores para os três testes foi, respectivamente, de 4.06, 3.56 e 3.45, ao passo que a variância foi de 5.34, 0.97

e 0.92. A quantificação do “*goodness of fit*” apontou $R^2 = 0.96$, indicando pouca discrepância dos valores apresentados em relação à curva normal.

Também foi realizada a quantificação do número de cópias de 32 indivíduos cuja tipagem para o *DYS503* resultou em quatro alelos diferentes, servindo-nos como controle. Para isso, a média dos valores normalizados de cópias para a população africana foi comparada à média dos valores normalizados de cópias desses 32 indivíduos através do teste t de Student. O valor calculado para esse teste (0.47) sugere que não há diferença significativa entre as duas populações.

O histograma da Figura 15 demonstra que a distribuição do número de cópias segue o mesmo padrão, com robusta concentração dos dados entre 3.0 e 4.0, confirmando a homogeneidade da contagem de quatro cópias para os indivíduos africanos. A média encontrada para esses valores foi de 3.55, 3.29 e 3.35, com variância de 0.61, 0.13 e 0.31, respectivamente. O cálculo de R^2 para o “*goodness of fit*” também foi calculado e resultou no valor de 0.98, apontando uma boa relação entre a distribuição esperada e observada dos dados.

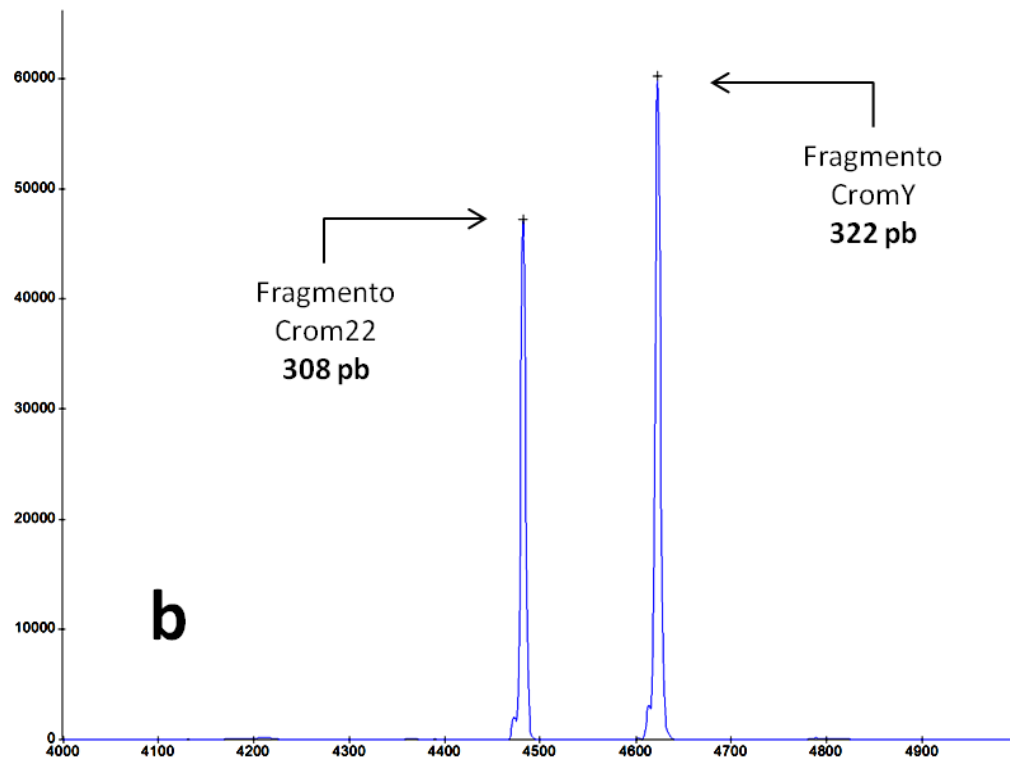
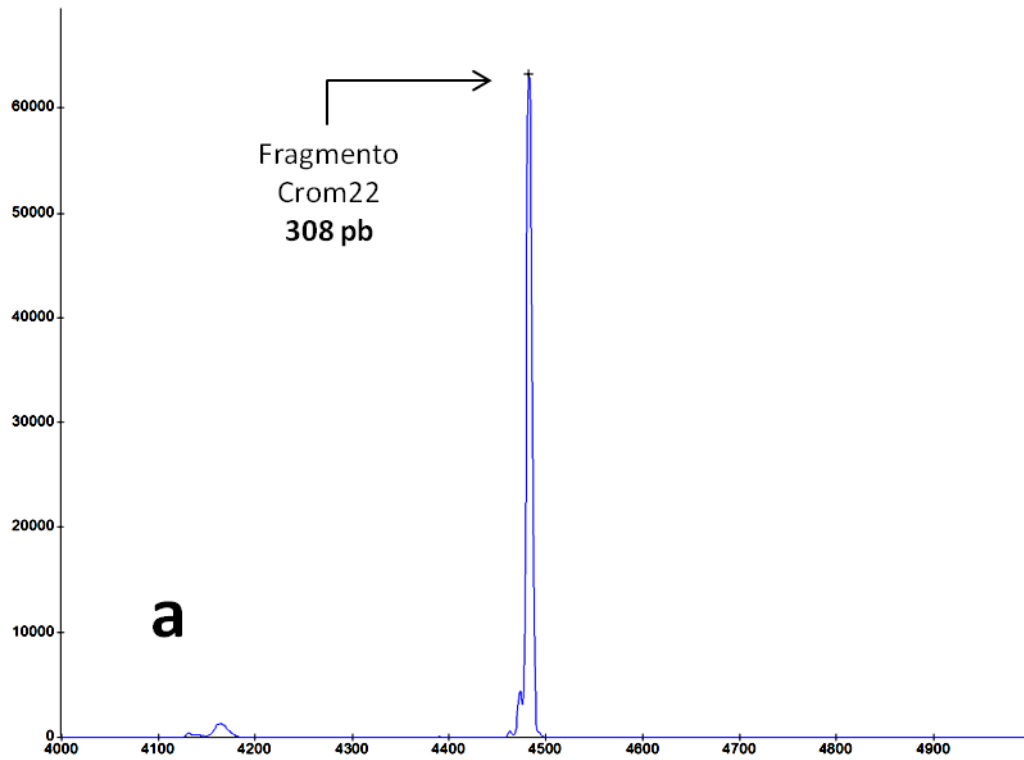


Figura 12: Eletroferogramas indicando amplificação de (a) uma amostra feminina e (b) uma amostra masculina.

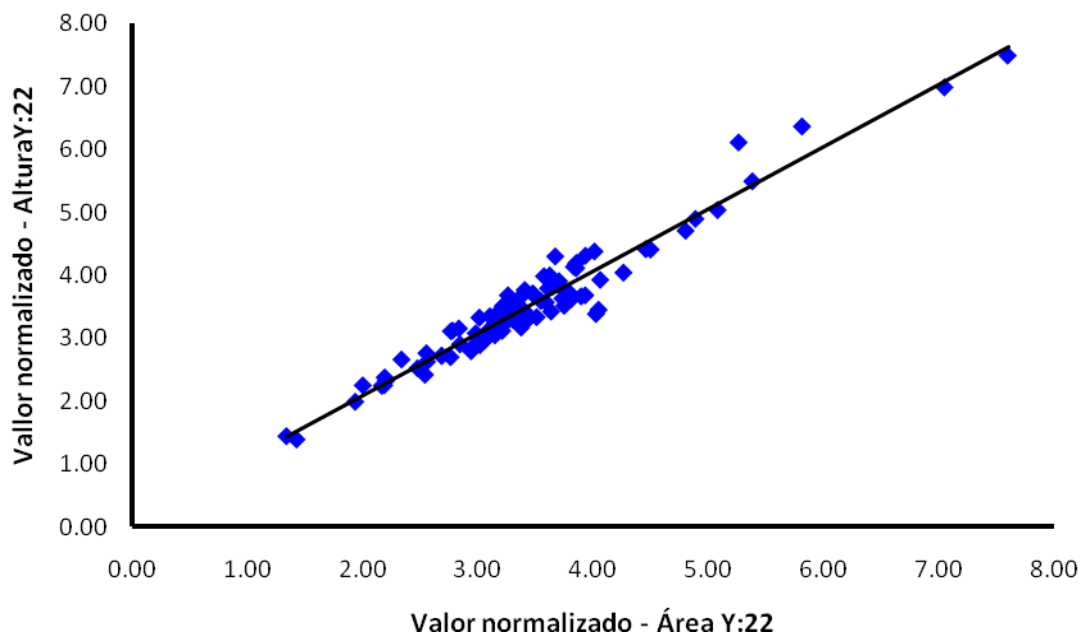


Figura 13: Gráficos de pontos representando a relação entre o valor normalizado das áreas (eixo x) e o valor normalizado das alturas (eixo y) dos picos encontrados para os fragmentos do cromossomo Y e 22.

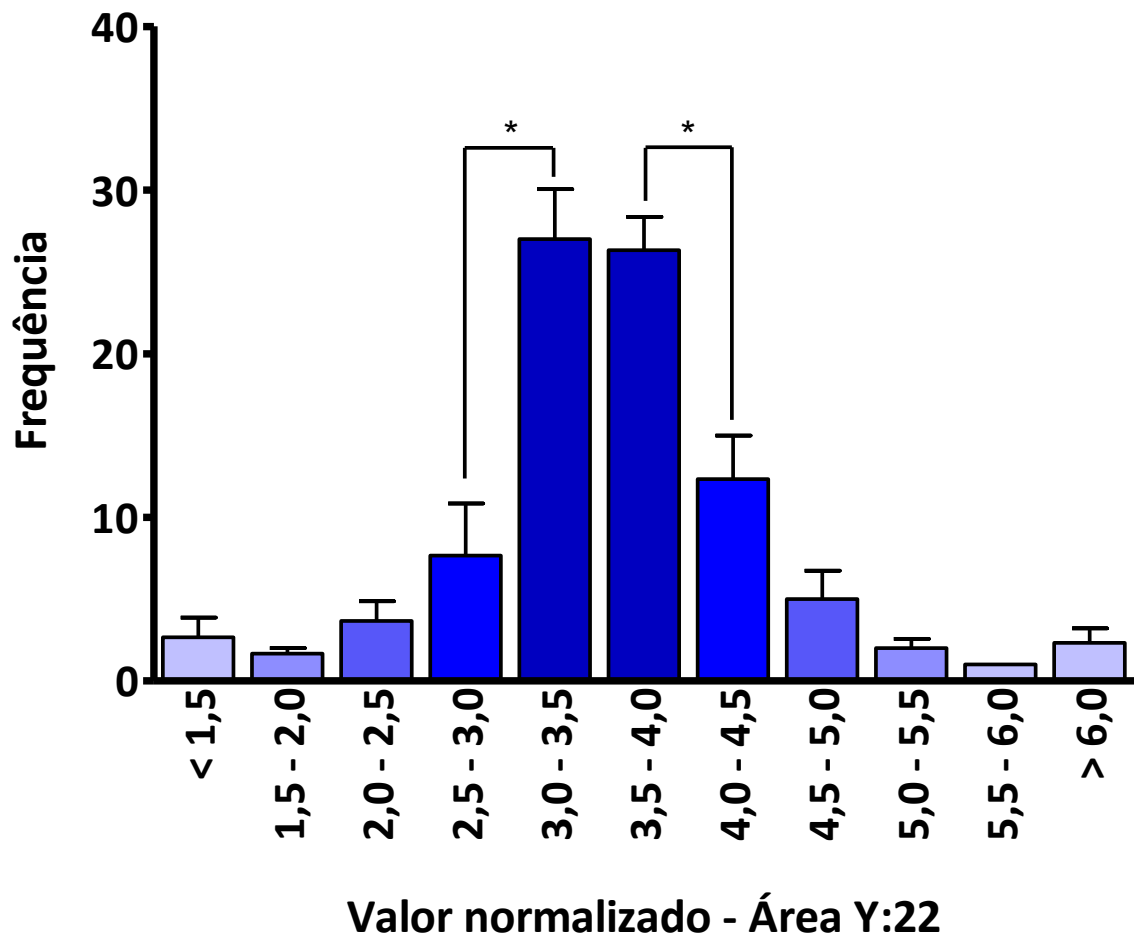


Figura 14: Histograma da frequência dos valores normalizados de número de cópias para a população africana, em relação à amostra padrão *SP*, baseados na razão das áreas entre os picos do cromossomo Y e 22. Os dados acima apontam a média dos três experimentos, com o erro padrão (* indica $p < 0,05$).

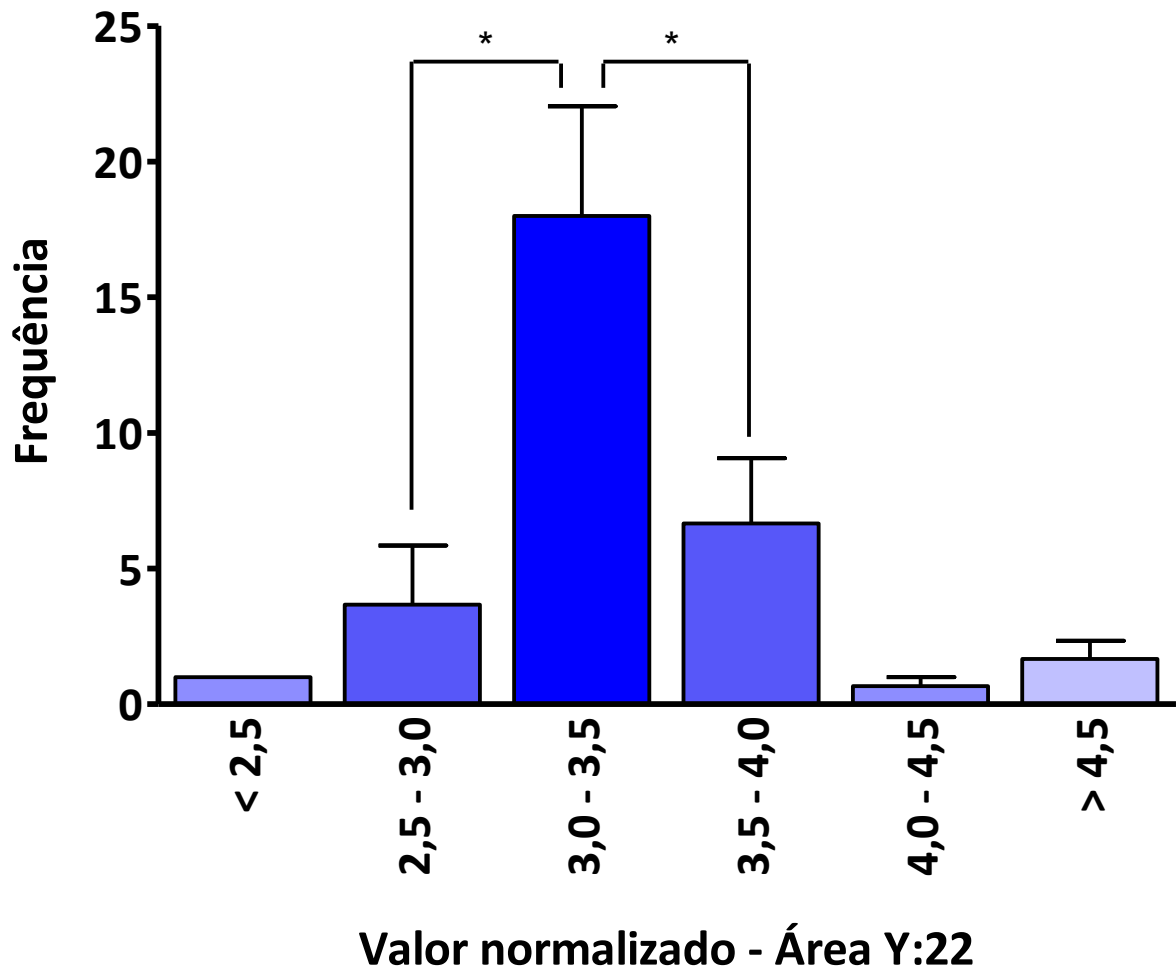


Figura 15: Histograma da frequência dos valores normalizados de número de cópias para os 32 indivíduos com quatro alelos diferentes para *DYS503*, em relação à amostra padrão *SP*, baseados na razão das áreas entre os picos do cromossomo Y e 22.



Discussão

D. DISCUSSÃO

A África possui não somente uma incrível pluralidade linguística, ambiental e cultural mas também tem os níveis mais altos de variabilidade genética do mundo. Dados paleontológicos e moleculares apontam que o *Homo sapiens* evoluiu na África cerca de 200 mil anos atrás e permaneceu nesse continente por mais tempo do que em qualquer outra região geográfica, levando ao acúmulo de mutações e ao aumento da diversidade genética. Com a migração do ser humano para o resto do mundo, ocorreram sucessivos efeitos fundadores e com isso, redução da variabilidade em relação à população africana original (Tishkoff *et al.*, 2002).

Por possuir herança uniparental, haploidia e por não sofrer recombinação na maior parte de sua extensão, o cromossomo Y é um ótimo objeto de estudo em pesquisas de evolução humana, pois guarda um registro dos eventos mutacionais que ocorreram ao longo das linhagens masculinas. O modelo de *Origem Africana Recente* (“*ex Africa*”) prediz que o ancestral de todos os cromossomos Y modernos residiu na África. A partir desta premissa, é possível desenvolver relações evolutivas entre os cromossomos Y, identificando suas origens e pontos de dispersão. Além disso, é possível definir que populações próximas têm grande chance de dividir frequências similares de um determinado padrão genético. Desta forma, a migração de um grupo pode levar ao estabelecimento de uma nova população que possuirá um subconjunto dos cromossomos Y originais. O fluxo genético proporcionado pelo deslocamento e demais eventos como a miscigenação entre as diferentes populações definem a distribuição das linhagens modernas (Jobling e Tyler-Smith, 1995).

Uma das ferramentas de investigação do cromossomo Y é a análise de microssatélites, marcadores amplamente utilizados em pesquisas forenses, evolutivas e genealógicas, pelo caráter de evolução rápida. Dentre os vários existentes, o microssatélite tetralocal *DYS503* tornou-se de especial interesse ao nosso grupo de pesquisa por apresentar uma diversidade genotípica inferior para afro-americanos em relação a caucasianos americanos.

Logo, elaboramos a hipótese de que, antes do êxodo do *Homo sapiens* para o resto do mundo, possivelmente havia no continente africano indivíduos com múltiplas cópias para o microssatélite *DYS503* e a população fundadora que saiu do grupo inicial para colonizar o resto do mundo continha apenas parte dessa variabilidade, ou seja, a nova população abrigaria apenas genótipos com os quatro alelos. Este fato justificaria a menor variabilidade na população africana em relação à caucasiana. Para investigar essa ideia, foi construído um mapa físico do *DYS503*, a partir do qual foi possível a aplicação de uma técnica de quantificação do número de cópias através de PCR quantitativa fluorescente.

Utilizando amostras de homens africanos do painel do CEPH-HGDP, a quantificação revelou a existência de uma população homogênea, sem variações significativas do número de cópias para *DYS503* (Figura 14, página 55), fato confirmado pela tipagem de 32 indivíduos com 4 alelos confirmados para o marcador (Figura 15, página 56), com a qual foi verificada a prevalência de indivíduos com as quatro cópias do microssatélite, não havendo indícios da existência de uma quantidade diferentes de cópias.

D.1 Mapa físico do microssatélite *DYS503*

A construção do mapa físico do microssatélite *DYS503* revelou quatro grandes blocos de cerca de 167 kb, alinhados em pares de sentido oposto e organizados no braço longo do cromossomo Y de forma assimétrica (Figura 10, página 47). Essa disposição incomum pode apontar efeitos cromossômicos de duplicação, deleção e inversão, comuns nessa região do genoma devido à alta densidade de sequências parálogas suscetíveis à recombinação homóloga não-alélica e conversão gênica (Rozen *et al.*, 2003).

D.2 Amostragem na África

Avaliando as possíveis causas para a existência de uma população uniforme e sem variação do número de cópias para o *DYS503*, consideramos o grupo de amostras utilizadas neste estudo. As populações africanas possuem altos níveis de diversidade genética e fenotípica, e vivem em ambientes bem distintos. É provável, portanto, que populações étnica e geograficamente diferentes tenham sofrido adaptações locais que podem ter levado a divergências em estudos moleculares. Devido a isso, acredita-se que as poucas populações do painel do CEPH-HGDP não sejam representativas de toda a variedade existente na África, sendo necessária uma amostragem mais ampla e que compreenda toda essa multiplicidade étnica e geográfica. Eventualmente, a criação de novos painéis disponíveis à comunidade científica deveria ser levada em consideração, para que, em complemento com os bancos de dados já existentes, as novas amostras tornem as análises mais

informativas e capazes de ilustrar a diversidade genética humana em todo o mundo (Campbell e Tishkoff, 2010; Reed, 2006).

D.3 Genótipos estendidos

Outra possível razão para a existência de uma população homogênea quanto ao número de cópias para o microssatélite *DYS503* são os genótipos estendidos, promovidos por esse tipo de marcador de evolução rápida. Devido à ausência de recombinação com um outro homólogo, o cromossomo Y contém um registro dos eventos mutacionais passados de compreensão relativamente simples. Entretanto, essa aparente simplicidade está comprometida pela existência de grandes regiões duplicadas, principalmente em seu braço longo (Skaletsky *et al.*, 2003). No caso específico do estudo de microssatélites, enquanto a informação contida em um marcador de cópia única é clara e inequívoca, a variação presente em microssatélites de múltiplas cópias são muito mais complexas para analisar e interpretar (Butler *et al.*, 2005). Contudo, esses microssatélites com várias cópias representam uma importante parte do cromossomo Y pois além de estarem sujeitos aos mecanismos evolutivos dos quais também participam os microssatélites de cópia única, também integram os eventos específicos da região amplicônica.

A maioria dos microssatélites duplicados como o *DYS503* e o *DYS464* estão localizados nessa região repleta de palíndromos e repetições invertidas (Figura 4, página 13). Em um total de 166 locos, Kayser e colaboradores (2004) identificaram 38 microssatélites com potencial para a geração de múltiplos fragmentos no cromossomo, cujos iniciadores mostraram 100% de identidade para mais de uma

posição no cromossomo Y. Desse grupo, 25 locos efetivamente indicaram múltiplos alelos. Do ponto de vista prático, o alto grau de similaridade encontrado entre essas porções cromossômicas torna quase impossível o desenho de iniciadores capazes de amplificar as diferentes cópias independentemente.

Dessa forma, a origem exata dos fragmentos gerados por iniciadores viáveis permanece sem distinção. A principal consequência desta amplificação simultânea é a impossibilidade de promover haplótipos corretamente ordenados, ou seja, há apenas a descrição dos alelos envolvidos ou o agrupamento dos mesmos em “genótipos estendidos”, baseados na razão da altura dos picos relativos aos fragmentos (Butler e Schoske, 2005; Hanson e Ballantyne, 2007), para os quais um mesmo “genótipo” pode corresponder a vários haplótipos diferentes. Logo, o uso de microssatélites de múltiplas cópias em estudos populacionais requer atenção especial, visto que não é possível relacionar cada fragmento gerado ao seu respectivo loco.

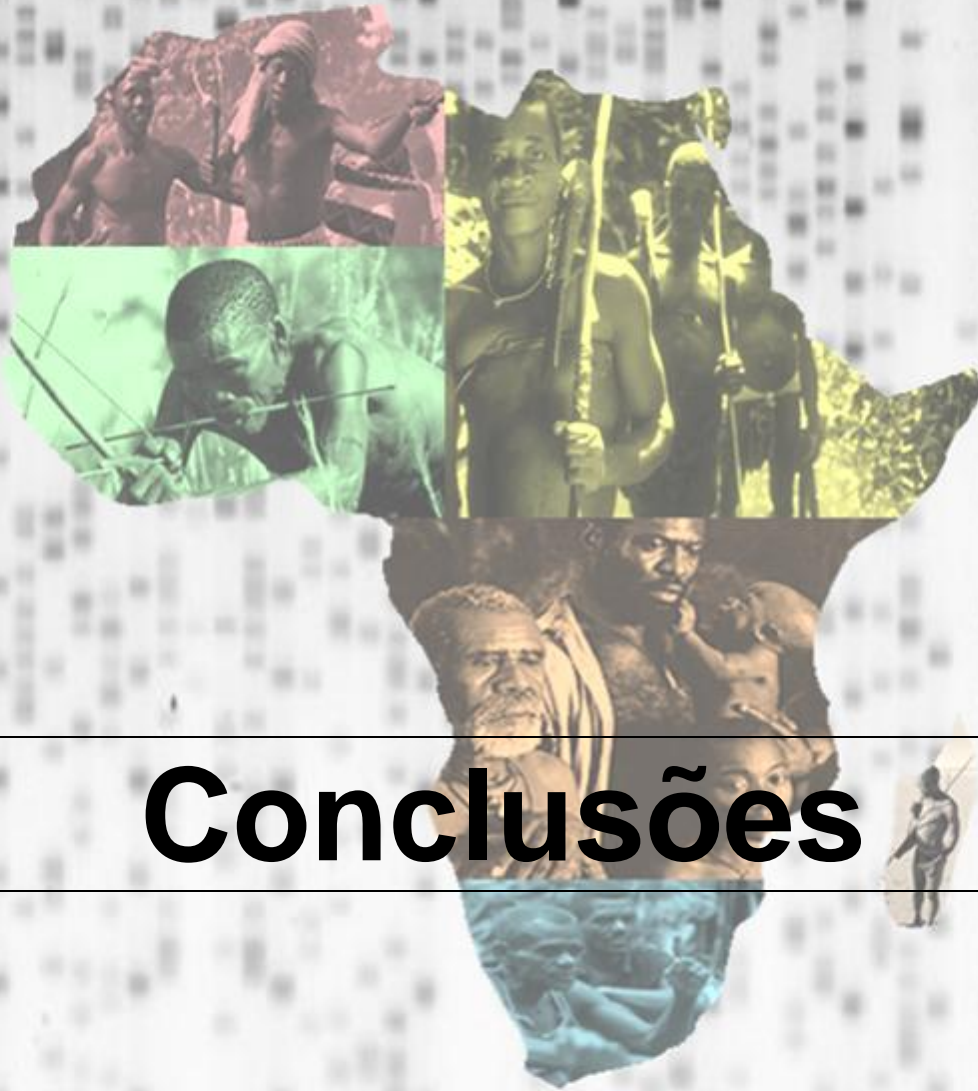
Estudos apontam que, devido à alta taxa de mutação, os microssatélites frequentemente promovem situações onde dois alelos com o mesmo número de repetições são idênticos por estado e não idênticos por descendência (Heyer *et al.*, 1997; Kayser *et al.*, 2000). Logo, não são marcadores de escolha para a construção de árvores filogenéticas ou para a inferência de relações entre populações divergentes. Sua utilidade é evidenciada nos estudos forenses, análise de paternidade e avaliações de afinidades entre populações mais intimamente relacionadas (Kayser *et al.*, 2001; Schlecht *et al.*, 2008). Além disso, a susceptibilidade dos microssatélites do cromossomo Y à conversão gênica e outros

mecanismos moleculares pode variar de acordo com as sequências repetitivas presentes. Pesquisar a formação e história evolutiva dessa região pode ser crucial para o entendimento destes processos. Dessa forma, poderia ser interessante associar a forte diferenciação geográfica proporcionada pelos haplogrupos do Y com os diferentes arranjos dos palíndromos, a fim de identificar eventos cromossômicos específicos em cada população (Balaresque *et al.*, 2007).

D.4 Deriva genética

Um importante ponto a ser considerado é a participação da deriva genética sobre a variabilidade entre as populações. Assumindo a razão entre os gêneros de 1:1, a população humana pode ser representada por um homem e uma mulher. Este casal carrega quatro cópias de cada autossomo e três cromossomos X, mas apenas um cromossomo Y. Logo, o tamanho efetivo da população do cromossomo Y é proporcional a um quarto da população de qualquer autossomo, um terço em relação ao cromossomo X e similar ao DNA haplóide mitocondrial. Assumindo que os mesmos processos mutacionais atuem em todos os cromossomos, é esperada uma menor diversidade no cromossomo Y do que em qualquer outro lugar do genoma nuclear, o que de fato é observado (Thomson *et al.*, 2000). Devido ao menor número efetivo, também se supõe que o cromossomo Y seja mais suscetível à deriva genética, que envolve mudanças aleatórias na frequência dos haplótipos devido à amostragem de uma geração para a seguinte. A deriva acelera a diferenciação entre os grupos de cromossomos Y nas diferentes populações,

constituindo uma útil propriedade para a investigação de eventos passados (Jobling e Tyler-Smith, 2003; Kayser *et al.*, 2001).



Conclusões

E. CONCLUSÕES

- O mapa físico para *DYS503* foi construído, evidenciando a posição cromossômica das extremidades dos quatro blocos que constituem a região repetitiva na qual está o microssatélite tetralocal. Esses blocos têm aproximadamente o mesmo tamanho (cerca de 167 mil pares de bases) e estão dispostos de forma assimétrica no braço longo do cromossomo Y, indicando que os eventos responsáveis pela duplicação dos locos ocorreram de forma desigual.
- A partir da construção do mapa físico, a técnica de PCR quantitativa fluorescente (QF-PCR) foi aplicada com sucesso, cumprindo o propósito da geração de fragmentos de tamanhos diferentes a partir da amplificação de distintas regiões genômicas, com o mesmo par de iniciadores.
- Essa técnica foi utilizada na quantificação relativa em indivíduos africanos, a fim de averiguar a possibilidade de variação do número de cópias em *DYS503*. Os resultados apontam a existência de uma população basicamente uniforme, contendo indivíduos com os quatro alelos. Dessa forma, não há evidência alguma que comprove a existência de indivíduos com um número diferente de cópias na África.
- Não há diferença significativa entre os valores normalizados da tipagem para a população africana em relação à realizada para indivíduos portadores de quatro cópias para o *DYS503*, o que reforça a inexistência de múltiplas cópias para o grupo estudado.



Referências

F. REFERÊNCIAS

ADINOLFI, M., SHERLOCK, J. & PERTL, B. 1995. *Rapid detection of selected aneuploidies by quantitative fluorescent PCR*. **Bioessays**, 17, 661-4.

BACHTROG, D. & CHARLESWORTH, B. 2001. *Towards a complete sequence of the human Y chromosome*. **Genome Biol**, 2, reviews 1016.1–1016.5.

BALARESQUE, P., BOWDEN, G. R., PARKIN, E. J., OMRAN, G. A., HEYER, E., QUINTANA-MURCI, L., ROEWER, L., STONEKING, M., NASIDZE, I., CARVALHO-SILVA, D. R., TYLER-SMITH, C., DE KNIJFF, P. & JOBLING, M. A. 2008. *Dynamic nature of the proximal AZFc region of the human Y chromosome: multiple independent deletion and duplication events revealed by microsatellite analysis*. **Hum Mutat**, 29, 1171-80.

BALARESQUE, P., PARKIN, E. J., ROEWER, L., CARVALHO-SILVA, D. R., MITCHELL, R. J., VAN OORSCHOT, R. A., HENKE, J., STONEKING, M., NASIDZE, I., WETTON, J., DE KNIJFF, P., TYLER-SMITH, C. & JOBLING, M. A. 2009. *Genomic complexity of the Y-STR DYS19: inversions, deletions and founder lineages carrying duplications*. **Int J Legal Med**, 123, 15-23.

BALARESQUE, P., SIBERT, A., HEYER, E. & CROUAU-ROY, B. 2007. *Unbiased interpretation of haplotypes at duplicated microsatellites*. **Ann Hum Genet**, 71, 209-19.

BLAKE, D., TAN, S. L. & AO, A. 1999. *Assessment of multiplex fluorescent PCR for screening single cells for trisomy 21 and single gene defects*. **Mol Hum Reprod**, 5, 1166-75.

BUTLER, J. M., DECKER, A. E., KLINE, M. C. & VALLONE, P. M. 2005. *Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation*. **J Forensic Sci**, 50, 853-9.

BUTLER, J. M. & SCHOSKE, R. 2005. *U.S. population data for the multi-copy Y-STR locus DYS464*. **J Forensic Sci**, 50, 975-7.

CALAFELL, F., SHUSTER, A., SPEED, W. C., KIDD, J. R. & KIDD, K. K. 1998. *Short tandem repeat polymorphism evolution in humans*. **Eur J Hum Genet**, 6, 38-49.

CAMPBELL, M. C. & TISHKOFF, S. A. 2008. *African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping*. **Annu Rev Genomics Hum Genet**, 9, 403-33.

CAMPBELL, M. C. & TISHKOFF, S. A. 2010. *The evolution of human genetic and phenotypic variation in Africa*. **Curr Biol**, 20, R166-73.

CANN, H. M., DE TOMA, C., CAZES, L., LEGRAND, M. F., MOREL, V., PIOUSFRE, L., BODMER, J., BODMER, W. F., BONNE-TAMIR, B., CAMBON-THOMSEN, A., CHEN, Z., CHU, J., CARCASSI, C., CONTU, L., DU, R., EXCOFFIER, L., FERRARA, G. B., FRIEDLAENDER, J. S., GROOT, H., GURWITZ, D., JENKINS, T., HERRERA, R. J., HUANG, X., KIDD, J., KIDD, K. K., LANGANEY, A., LIN, A. A., MEHDI, S. Q., PARHAM, P., PIAZZA, A., PISTILLO, M. P., QIAN, Y., SHU, Q., XU, J., ZHU, S., WEBER, J. L., GREELY, H. T., FELDMAN, M. W., THOMAS, G., DAUSSET, J. & CAVALLI-SFORZA, L. L. 2002. *A human genome diversity cell line panel*. **Science**, 296, 261-2.

CANN, R. L., STONEKING, M. & WILSON, A. C. 1987. *Mitochondrial DNA and human evolution*. **Nature**, 325, 31-6.

CARRANO, A. V., LAMERDIN, J., ASHWORTH, L. K., WATKINS, B., BRANSCOMB, E., SLEZAK, T., RAFF, M., DE JONG, P. J., KEITH, D., MCBRIDE, L. & ET AL. 1989. *A high-resolution, fluorescence-based, semiautomated method for DNA fingerprinting*. **Genomics**, 4, 129-36.

CAVALLI-SFORZA, L. L. 2005. *The Human Genome Diversity Project: past, present and future*. **Nat Rev Genet**, 6, 333-40.

CAVALLI-SFORZA, L. L. & FELDMAN, M. W. 2003. *The application of molecular genetic approaches to the study of human evolution*. **Nat Genet**, 33 Suppl, 266-75.

COBLE, M. D., LOREILLE, O. M., WADHAMS, M. J., EDSON, S. M., MAYNARD, K., MEYER, C. E., NIEDERSTATTER, H., BERGER, C., BERGER, B., FALSETTI, A. B., GILL, P., PARSON, W. & FINELLI, L. N. 2009. *Mystery solved: the identification of the two missing Romanov children using DNA analysis*. **PLoS One**, 4, e4838.

DEGIORGIO, M., JAKOBSSON, M. & ROSENBERG, N. A. 2009. *Out of Africa: modern human origins special feature: explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa*. **Proc Natl Acad Sci U S A**, 106, 16057-62.

DESHPANDE, O., BATZOGLOU, S., FELDMAN, M. W. & CAVALLI-SFORZA, L. L. 2009. *A serial founder effect model for human settlement out of Africa*. **Proc Biol Sci**, 276, 291-300.

EICHLER, E. E. 2001. *Recent duplication, domain accretion and the dynamic mutation of the human genome*. **Trends Genet**, 17, 661-9.

EXCOFFIER, L. 2002. *Human demographic history: refining the recent African origin model*. **Curr Opin Genet Dev**, 12, 675-82.

FINDLAY, I. & QUIRKE, P. 1996. *Fluorescent polymerase chain reaction: Part I. A new method allowing genetic diagnosis and DNA fingerprinting of single cells*. **Hum Reprod Update**, 2, 137-52.

GOLDSTEIN, D. B., RUIZ LINARES, A., CAVALLI-SFORZA, L. L. & FELDMAN, M. W. 1995. *Genetic absolute dating based on microsatellites and the origin of modern humans*. **Proc Natl Acad Sci U S A**, 92, 6723-7.

GOMES, V., SANCHEZ-DIZ, P., AMORIM, A., CARRACEDO, A. & GUSMAO, L. 2010. *Digging deeper into East African human Y chromosome lineages*. **Hum Genet**, 127, 603-13.

GOODFELLOW, P., DARLING, S. & WOLFE, J. 1985. *The human Y chromosome*. **J Med Genet**, 22, 329-44.

GRAVES, J. A. 2010. *Review: Sex chromosome evolution and the expression of sex-specific genes in the placenta*. **Placenta**, 31 Suppl, S27-32.

HAMMER, M. F., KARAFET, T. M., REDD, A. J., JARJANAZI, H., SANTACHIARA-BENERECETTI, S., SOODYALL, H. & ZEGURA, S. L. 2001. *Hierarchical patterns of global human Y-chromosome diversity*. **Mol Biol Evol**, 18, 1189-203.

HANSON, E. K. & BALLANTYNE, J. 2006. *Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications*. **Leg Med (Tokyo)**, 8, 110-20.

HANSON, E. K. & BALLANTYNE, J. 2007. *Population data for a novel, highly discriminating tetra-local Y-chromosome short tandem repeat: DYS503*. **J Forensic Sci**, 52, 498-9.

HELLENTHAL, G., AUTON, A. & FALUSH, D. 2008. *Inferring human colonization history using a copying model*. **PLoS Genet**, 4, e1000078.

HEYER, E., PUYMIRAT, J., DIELTJES, P., BAKKER, E. & DE KNIJFF, P. 1997. *Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees*. **Hum Mol Genet**, 6, 799-803.

HURLES, M. E. & JOBLING, M. A. 2001. *Haploid chromosomes in molecular ecology: lessons from the human Y*. **Mol Ecol**, 10, 1599-613.

HURLES, M. E. & JOBLING, M. A. 2003. *A singular chromosome*. **Nat Genet**, 34, 246-7.

IAFRATE, A. J., FEUK, L., RIVERA, M. N., LISTEWNIK, M. L., DONAHOE, P. K., QI, Y., SCHERER, S. W. & LEE, C. 2004. *Detection of large-scale variation in the human genome*. **Nat Genet**, 36, 949-51.

JOBLING, M. A. 2008. *Copy number variation on the human Y chromosome*. **Cytogenet Genome Res**, 123, 253-62.

JOBLING, M. A., PANDYA, A. & TYLER-SMITH, C. 1997. *The Y chromosome in forensic analysis and paternity testing*. **Int J Legal Med**, 110, 118-24.

JOBLING, M. A., SAMARA, V., PANDYA, A., FRETWELL, N., BERNASCONI, B., MITCHELL, R. J., GERELSAIKHAN, T., DASHNYAM, B., SAJANTILA, A., SALO, P. J., NAKAHORI, Y., DISTECHE, C. M., THANGARAJ, K., SINGH, L., CRAWFORD, M. H. & TYLER-SMITH, C. 1996. *Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males*. **Hum Mol Genet**, 5, 1767-75.

JOBLING, M. A. & TYLER-SMITH, C. 1995. *Fathers and sons: the Y chromosome and human evolution*. **Trends Genet**, 11, 449-56.

JOBLING, M. A. & TYLER-SMITH, C. 2003. *The human Y chromosome: an evolutionary marker comes of age*. **Nat Rev Genet**, 4, 598-612.

JORDE, L. B., ROGERS, A. R., BAMSHAD, M., WATKINS, W. S., KRAKOWIAK, P., SUNG, S., KERE, J. & HARPENDING, H. C. 1997. *Microsatellite diversity and the demographic history of modern humans*. **Proc Natl Acad Sci U S A**, 94, 3100-3.

JU, J., RUAN, C., FULLER, C. W., GLAZER, A. N. & MATHIES, R. A. 1995. *Fluorescence energy transfer dye-labeled primers for DNA sequencing and analysis*. **Proc Natl Acad Sci U S A**, 92, 4347-51.

KAYSER, M., KITTLER, R., ERLER, A., HEDMAN, M., LEE, A. C., MOHYUDDIN, A., MEHDI, S. Q., ROSSER, Z., STONEKING, M., JOBLING, M. A., SAJANTILA, A. & TYLER-SMITH, C. 2004. *A comprehensive survey of human Y-chromosomal microsatellites*. **Am J Hum Genet**, 74, 1183-97.

KAYSER, M., KRAWCZAK, M., EXCOFFIER, L., DIELTJES, P., CORACH, D., PASCALI, V., GEHRIG, C., BERNINI, L. F., JESPERSEN, J., BAKKER, E., ROEWER, L. & DE KNIJFF, P. 2001. *An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations*. **Am J Hum Genet**, 68, 990-1018.

KAYSER, M., ROEWER, L., HEDMAN, M., HENKE, L., HENKE, J., BRAUER, S., KRUGER, C., KRAWCZAK, M., NAGY, M., DOBOSZ, T., SZIBOR, R., DE KNIJFF, P., STONEKING, M. & SAJANTILA, A. 2000. *Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs*. **Am J Hum Genet**, 66, 1580-8.

KEHDY, F. S. G., RACIOPPI, C.B., SANTOS, L. S., PENA, S. D. J. *Estudo dos microssatélites tetralocais DYS464 e DYS503 em populações mundiais*. In: CONGRESSO BRASILEIRO DE GENÉTICA, 54, 2008. Salvador. Resumos... p. 24.

KURODA-KAWAGUCHI, T., SKALETSKY, H., BROWN, L. G., MINX, P. J., CORDUM, H. S., WATERSTON, R. H., WILSON, R. K., SILBER, S., OATES, R., ROZEN, S. & PAGE, D. C. 2001. *The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men*. **Nat Genet**, 29, 279-86.

LAHN, B. T. & PAGE, D. C. 1999. *Four evolutionary strata on the human X chromosome*. **Science**, 286, 964-7.

LAHN, B. T., PEARSON, N. M. & JEGALIAN, K. 2001. *The human Y chromosome, in the light of evolution*. **Nat Rev Genet**, 2, 207-16.

LANGE, J., SKALETSKY, H., VAN DAALLEN, S. K., EMBRY, S. L., KORVER, C. M., BROWN, L. G., OATES, R. D., SILBER, S., REPPING, S. & PAGE, D. C. 2009. *Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes*. **Cell**, 138, 855-69.

LEE, M. H., RYU, H. M., KIM, D. J., LEE, B. Y., CHO, E. H., YANG, J. H., KIM, M. Y., HAN, J. Y. & PARK, S. Y. 2004. *Rapid prenatal diagnosis of Down Syndrome using quantitative fluorescent PCR in uncultured amniocytes*. **J Korean Med Sci**, 19, 341-4.

LI, Y. C., KOROL, A. B., FAHIMA, T., BEILES, A. & NEVO, E. 2002. *Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review*. **Mol Ecol**, 11, 2453-65.

LI, Z., HAINES, C. J. & HAN, Y. 2008. *"Micro-deletions" of the human Y chromosome and their relationship with male infertility*. **J Genet Genomics**, 35, 193-9.

LIU, H., PRUGNOLLE, F., MANICA, A. & BALLOUX, F. 2006. *A geographically explicit genetic model of worldwide human-settlement history*. **Am J Hum Genet**, 79, 230-7.

MANSFIELD, E. S. 1993. *Diagnosis of Down syndrome and other aneuploidies using quantitative polymerase chain reaction and small tandem repeat polymorphisms*. **Hum Mol Genet**, 2, 43-50.

MITCHELL, R. J. & HAMMER, M. F. 1996. *Human evolution and the Y chromosome*. **Curr Opin Genet Dev**, 6, 737-42.

MUTTER, G. L. & POMPONIO, R. J. 1991. *Molecular diagnosis of sex chromosome aneuploidy using quantitative PCR*. **Nucleic Acids Res**, 19, 4203-7.

PENA, S. D. J. 1998. *Molecular Cytogenetics I: PCR-based diagnosis of human trisomies using computer-assisted laser densitometry*. **Genet Mol Biol**, 21 (3).

PERTL, B., WEITGASSER, U., KOPP, S., KROISEL, P. M., SHERLOCK, J. & ADINOLFI, M. 1996. *Rapid detection of trisomies 21 and 18 and sexing by quantitative fluorescent multiplex PCR.* **Hum Genet**, 98, 55-9.

RAMACHANDRAN, S., DESHPANDE, O., ROSEMAN, C. C., ROSENBERG, N. A., FELDMAN, M. W. & CAVALLI-SFORZA, L. L. 2005. *Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa.* **Proc Natl Acad Sci U S A**, 102, 15942-7.

RAVEL, C., CHANTOT-BASTARAUD, S., EL HOUATE, B., ROUBA, H., LEGENDRE, M., LORENCO, D., MANDELBAUM, J., SIFFROI, J. P. & MCELREAVEY, K. 2009. *Y-chromosome AZFc structural architecture and relationship to male fertility.* **Fertil Steril**, 92, 1924-33.

REDON, R., ISHIKAWA, S., FITCH, K. R., FEUK, L., PERRY, G. H., ANDREWS, T. D., FIEGLER, H., SHAPERO, M. H., CARSON, A. R., CHEN, W., CHO, E. K., DALLAIRE, S., FREEMAN, J. L., GONZALEZ, J. R., GRATACOS, M., HUANG, J., KALAITZOPOULOS, D., KOMURA, D., MACDONALD, J. R., MARSHALL, C. R., MEI, R., MONTGOMERY, L., NISHIMURA, K., OKAMURA, K., SHEN, F., SOMERVILLE, M. J., TCHINDA, J., VALSESIA, A., WOODWARK, C., YANG, F., ZHANG, J., ZERJAL, T., ARMENGOL, L., CONRAD, D. F., ESTIVILL, X., TYLER-SMITH, C., CARTER, N. P., ABURATANI, H., LEE, C., JONES, K. W., SCHERER, S. W. & HURLES, M. E. 2006. *Global variation in copy number in the human genome.* **Nature**, 444, 444-54.

REED, F. A. & TISHKOFF, S. A. 2006. *African human diversity, origins and migrations.* **Curr Opin Genet Dev**, 16, 597-605.

RELETFORD, J. H. 2008. *Genetic evidence and the modern human origins debate.* **Heredity**, 100, 555-63.

REPPING, S., SKALETSKY, H., BROWN, L., VAN DAALEN, S. K., KORVER, C. M., PYNTIKOVA, T., KURODA-KAWAGUCHI, T., DE VRIES, J. W., OATES, R. D.,

SILBER, S., VAN DER VEEN, F., PAGE, D. C. & ROZEN, S. 2003. *Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection.* **Nat Genet**, 35, 247-51.

ROEWER, L. 2009. *Y chromosome STR typing in crime casework.* **Forensic Sci Med Pathol**, 5, 77-84.

ROLF, B., KEIL, W., BRINKMANN, B., ROEWER, L. & FIMMERS, R. 2001. *Paternity testing using Y-STR haplotypes: assigning a probability for paternity in cases of mutations.* **Int J Legal Med**, 115, 12-5.

ROZEN, S. & SKALETSKY, H. 2000. *Primer3 on the WWW for general users and for biologist programmers.* **Methods Mol Biol**, 132, 365-86.

ROZEN, S., SKALETSKY, H., MARSZALEK, J. D., MINX, P. J., CORDUM, H. S., WATERSTON, R. H., WILSON, R. K. & PAGE, D. C. 2003. *Abundant gene conversion between arms of palindromes in human and ape Y chromosomes.* **Nature**, 423, 873-6.

SANTOS-LOPES, S. S., PEREIRA, R. W., WILSON, I. J. & PENA, S. D. 2007. *A worldwide phylogeography for the human X chromosome.* **PLoS One**, 2, e557.

SCHLECHT, J., KAPLAN, M. E., BARNARD, K., KARAFET, T., HAMMER, M. F. & MERCHANT, N. C. 2008. *Machine-learning approaches for classifying haplogroup from Y chromosome STR data.* **PLoS Comput Biol**, 4, e1000093.

SCHUELKE, M. 2000. *An economic method for the fluorescent labeling of PCR fragments.* **Nat Biotechnol**, 18, 233-4.

SEIELSTAD, M., BEKELE, E., IBRAHIM, M., TOURE, A. & TRAORE, M. 1999. *A view of modern human origins from Y chromosome microsatellite variation.* **Genome Res**, 9, 558-67.

SEIELSTAD, M. T., MINCH, E. & CAVALLI-SFORZA, L. L. 1998. *Genetic evidence for a higher female migration rate in humans.* **Nat Genet**, 20, 278-80.

SHARP, A. J., LOCKE, D. P., MCGRATH, S. D., CHENG, Z., BAILEY, J. A., VALLENTE, R. U., PERTZ, L. M., CLARK, R. A., SCHWARTZ, S., SEGRAVES, R., OSEROFF, V. V., ALBERTSON, D. G., PINKEL, D. & EICHLER, E. E. 2005. *Segmental duplications and copy-number variation in the human genome.* **Am J Hum Genet**, 77, 78-88.

SHE, X., JIANG, Z., CLARK, R. A., LIU, G., CHENG, Z., TUZUN, E., CHURCH, D. M., SUTTON, G., HALPERN, A. L. & EICHLER, E. E. 2004. *Shotgun sequence assembly and recent segmental duplications within the human genome.* **Nature**, 431, 927-30.

SKALETSKY, H., KURODA-KAWAGUCHI, T., MINX, P. J., CORDUM, H. S., HILLIER, L., BROWN, L. G., REPPING, S., PYNTIKOVA, T., ALI, J., BIERI, T., CHINWALLA, A., DELEHAUNTY, A., DELEHAUNTY, K., DU, H., FEWELL, G., FULTON, L., FULTON, R., GRAVES, T., HOU, S. F., LATRIELLE, P., LEONARD, S., MARDIS, E., MAUPIN, R., MCPHERSON, J., MINER, T., NASH, W., NGUYEN, C., OZERSKY, P., PEPIN, K., ROCK, S., ROHLFING, T., SCOTT, K., SCHULTZ, B., STRONG, C., TIN-WOLLAM, A., YANG, S. P., WATERSTON, R. H., WILSON, R. K., ROZEN, S. & PAGE, D. C. 2003. *The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes.* **Nature**, 423, 825-37.

SMITH, L. M., SANDERS, J. Z., KAISER, R. J., HUGHES, P., DODD, C., CONNELL, C. R., HEINER, C., KENT, S. B. & HOOD, L. E. 1986. *Fluorescence detection in automated DNA sequence analysis.* **Nature**, 321, 674-9.

STRINGER, C. 2000. *Palaeoanthropology. Coasting out of Africa.* **Nature**, 405, 24-5, 27.

STUMPF, M. P. & GOLDSTEIN, D. B. 2001. *Genealogical and evolutionary inference with the human Y chromosome.* **Science**, 291, 1738-42.

TEMPLETON, A. R. 1997. *Out of Africa. What do genes tell us?* **Curr Opin Genet Dev**, 7, 841-7.

THE Y CHROMOSOME CONSORTIUM. 2002. *A nomenclature system for the tree of human Y-chromosomal binary haplogroups.* **Genome Res**, 12, 339-48.

THOMSON, R., PRITCHARD, J. K., SHEN, P., OEFNER, P. J. & FELDMAN, M. W. 2000. *Recent common ancestry of human Y chromosomes: evidence from DNA sequence data.* **Proc Natl Acad Sci U S A**, 97, 7360-5.

TIEPOLO, L. & ZUFFARDI, O. 1976. *Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm.* **Hum Genet**, 34, 119-24.

TISHKOFF, S. A. & VERRELLI, B. C. 2003. *Patterns of human genetic diversity: implications for human evolutionary history and disease.* **Annu Rev Genomics Hum Genet**, 4, 293-340.

TISHKOFF, S. A. & WILLIAMS, S. M. 2002. *Genetic analysis of African populations: human evolution and complex disease.* **Nat Rev Genet**, 3, 611-21.

TOTH, T., FINDLAY, I., PAPP, C., TOTH-PAL, E., MARTON, T., NAGY, B., QUIRKE, P. & PAPP, Z. 1998. *Prenatal detection of trisomy 13 from amniotic fluid by quantitative fluorescent polymerase chain reaction.* **Prenat Diagn**, 18, 669-74.

TROMBETTA, B., CRUCIANI, F., UNDERHILL, P. A., SELBITTO, D. & SCOZZARI, R. 2010. *Footprints of X-to-Y gene conversion in recent human evolution.* **Mol Biol Evol**, 27, 714-25.

UNDERHILL, P. A., SHEN, P., LIN, A. A., JIN, L., PASSARINO, G., YANG, W. H., KAUFFMAN, E., BONNE-TAMIR, B., BERTRANPETIT, J., FRANCALACCI, P., IBRAHIM, M., JENKINS, T., KIDD, J. R., MEHDI, S. Q., SEIELSTAD, M. T., WELLS,

R. S., PIAZZA, A., DAVIS, R. W., FELDMAN, M. W., CAVALLI-SFORZA, L. L. & OEFNER, P. J. 2000. *Y chromosome sequence variation and the history of human populations*. **Nat Genet**, 26, 358-61.

VALERO, R., MARFANY, G., GIL-BENSO, R., IBANEZ, M. A., LOPEZ-PAJARES, I., PRIETO, F., RULLAN, G., SARRET, E. & GONZALEZ-DUARTE, R. 1999. *Molecular characterisation of partial chromosome 21 aneuploidies by fluorescent PCR*. **J Med Genet**, 36, 694-9.

VIGILANT, L., STONEKING, M., HARPENDING, H., HAWKES, K. & WILSON, A. C. 1991. *African populations and the evolution of human mitochondrial DNA*. **Science**, 253, 1503-7.

VOGT, P. H., EDELMANN, A., KIRSCH, S., HENEGARIU, O., HIRSCHMANN, P., KIESEWETTER, F., KOHN, F. M., SCHILL, W. B., FARAH, S., RAMOS, C., HARTMANN, M., HARTSCHUH, W., MESCHADE, D., BEHRE, H. M., CASTEL, A., NIESCHLAG, E., WEIDNER, W., GRONE, H. J., JUNG, A., ENGEL, W. & HAIDL, G. 1996. *Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11*. **Hum Mol Genet**, 5, 933-43.

WATERS, P. D., WALLIS, M. C. & MARSHALL GRAVES, J. A. 2007. *Mammalian sex--Origin and evolution of the Y chromosome and SRY*. **Semin Cell Dev Biol**, 18, 389-400.

WELLS, R. S., YULDASHEVA, N., RUZIBAKIEV, R., UNDERHILL, P. A., EVSEEVA, I., BLUE-SMITH, J., JIN, L., SU, B., PITCHAPPAN, R., SHANMUGALAKSHMI, S., BALAKRISHNAN, K., READ, M., PEARSON, N. M., ZERJAL, T., WEBSTER, M. T., ZHOLOSHVILI, I., JAMARJASHVILI, E., GAMBAROV, S., NIKBIN, B., DOSTIEV, A., AKNAZAROV, O., ZALLOUA, P., TSOY, I., KITAEV, M., MIRRAKHIMOV, M., CHARIEV, A. & BODMER, W. F. 2001. *The Eurasian heartland: a continental perspective on Y-chromosome diversity*. **Proc Natl Acad Sci U S A**, 98, 10244-9.

- WOLF, A., CALIEBE, A., JUNGE, O. & KRAWCZAK, M. 2005. *Forensic interpretation of Y-chromosomal DNA mixtures*. **Forensic Sci Int**, 152, 209-13.
- YEN, P. H. 1999. *Advances in Y chromosome mapping*. **Curr Opin Obstet Gynecol**, 11, 275-81.
- YU, N., CHEN, F. C., OTA, S., JORDE, L. B., PAMILO, P., PATTHY, L., RAMSAY, M., JENKINS, T., SHYUE, S. K. & LI, W. H. 2002. *Larger genetic differences within africans than between Africans and Eurasians*. **Genetics**, 161, 269-74.
- ZHANG, L., LU, H. H., CHUNG, W. Y., YANG, J. & LI, W. H. 2005. *Patterns of segmental duplication in the human genome*. **Mol Biol Evol**, 22, 135-41.
- ZHANG, Z., SCHWARTZ, S., WAGNER, L. & MILLER, W. 2000. *A greedy algorithm for aligning DNA sequences*. **J Comput Biol**, 7, 203-14.
- ZIMRAN, A., GLASS, C., THORPE, V. S. & BEUTLER, E. 1989. *Analysis of 'color PCR' by automatic DNA sequencer*. **Nucleic Acids Res**, 17, 7538.

Sítios de Internet

BLAST: Basic Local Alignment Search Tool

Disponível em: <<http://blast.ncbi.nlm.nih.gov/Blast.cgi>>

Acesso em 23 jun 2010.

Database of Genomic Variants

Disponível em: <<http://projects.tcag.ca/variation>>

Acesso em 23 jun 2010.

MultAlin Interface Page

Disponível em: <<http://bioinfo.genotoul.fr/multalin>>

Acesso 23 jun 2010.

NCBI - National Center for Biotechnology Information

Disponível em: <<http://www.ncbi.nlm.nih.gov>>

Acesso em 23 jun 2010.

OligoAnalyzer - Integrated DNA Technologies

Disponível em: <<http://www.idtdna.com/analyzer/applications/oligoanalyzer>>

Acesso em 23 jun 2010.

Primer3 Input (version 0.4.0)

Disponível em: <<http://frodo.wi.mit.edu/primer3>>

Acesso em 23 jun 2010.

UCSC In silico PCR

Disponível em: <<http://genome.ucsc.edu/cgi-bin/hgPcr?command=start>>

Acesso em 23 jun 2010.

Vassar Stats

Disponível em: <<http://faculty.vassar.edu/lowry/VassarStats.html>>

Acesso em 23 jun 2010.