



Ministério da Saúde
Centro de Pesquisas Gonçalo Moniz
Fundação Oswaldo Cruz
FIOCRUZ

Curso de Pós-graduação em Biotecnologia e Medicina Investigativa –
CPqGM/Fiocruz-BA

ENIO PAULO TELÓ

**Estimativa de mistura étnica avaliada por Marcadores Informativos de
Ancestralidade (AIMs) e Microssatélites (STRs)**

SALVADOR
2010

ENIO PAULO TELÓ

**Estimativa de mistura étnica avaliada por Marcadores Informativos de
Ancestralidade (AIMs) e Microssatélites (STRs)**

Dissertação apresentada ao Curso de
Pós-graduação em Biotecnologia em
Saúde e Medicina Investigativa
como requisito parcial para obtenção
do título de Mestre.

Orientadora: Prof. Dra. Kiyoko Abe Sandes

Salvador
2010

Ficha Catalográfica elaborada pela Biblioteca do
Centro de Pesquisas Gonçalo Moniz / FIOCRUZ - Salvador - Bahia.

Teló, Enio Paulo

T267e Estimativa de mistura étnica avaliada por Marcadores Informativos de Ancestralidade (AIMs) e Microssatélites (STRs) [manuscrito] / Enio Paulo Teló. - 2010.
68 f.; 30 cm

**Dissertação (mestrado) – Fundação Oswaldo Cruz. Centro de Pesquisas
Gonçalo**

**Moniz. Pós-Graduação em Biotecnologia em Saúde e Medicina
Investigativa, 2010.**

Orientadora: Profª. Dra. Kiyoko Abe Sandes.

1. Ancestralidade genômica 2. Short Tandem Repeats (STRs). 3. Marcadores Informativos de Ancestralidade (AIMs). I.Título.

CDU 575.113

Dedico esse trabalho aos meus pais, por sempre estarem ao meu lado e por terem me ensinado muitos valores que hoje estão em falta em nossas vidas. Minha noiva, familiares e amigos, a cada uma das pessoas que fizeram e fazem parte do meu cotidiano, com suas particularidades e personalidades distintas, mas que juntas acrescentaram muito na minha formação pessoal e profissional. Sem a ajuda e a presença de vocês, seria impossível seguir em frente.

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Marcadores Polimórficos	13
1.2 Diversidade genética populacional.....	15
1.3 Estimativa de mistura genética	17
1.4 Associação entre a ancestralidade africana e os sobrenomes de conotação Religiosa.....	18
2 JUSTIFICATIVA	22
3 OBJETIVOS	24
3.1 Geral	24
3.2 Específicos	24
4 MATERIAL E MÉTODOS	25
4.1 Amostra, Critérios de inclusão, autodenominação e análise fenotípica	25
4.2 Critérios de exclusão	26
4.3 Análise de sobrenomes	26
4.4 Marcadores analisados.....	26
4.4.1 Inserção/ Deleção – AT3-I/D.....	26
4.4.2 Inserções <i>Alu</i>	27
4.4.3 SNP (Polimorfismos de Nucleotídeo Único)	28
4.4.3.1 FY-Null	28
4.4.3.2 CKMM.....	29
4.4.3.3 LPL	29
4.4.3.4 CYP3A4	29
4.4.3.5 GC-1F e GC-1S	30
4.5 Reação em cadeia da polimerase (PCR).....	31
4.6 Coloração com nitrato de prata e secagem do gel	32
4.7 Genotipagem dos SNP: PCR em tempo real.....	33
4.8 Microssatélites STRs	34
4.9 Análises estatísticas	36
4.10 Considerações Éticas	37
5 RESULTADOS E DISCUSSÃO (Sumarizados no artigo ainda não publicado; e com tabelas complementares adicionadas aos apêndices)	38
- Introdução.....	39

- Material e métodos.....	42
- Resultados e discussão	43
- Referências Bibliográficas.....	50
6 CONCLUSÃO.....	54
7 REFERÊNCIAS BIBLIOGRÁFICAS	55
APÊNDICES	65

LISTA DE ABREVIATURAS

A – Adenina

AIMs- Marcadores Informativos de Ancestralidade (do inglês, Ancestry Informative Markers)

C - Citosina

DNA – Ácido desoxirribonucléico (do inglês, Deoxyribonucleic Acid)

DNAmt - Ácido desoxirribonucléico mitocondrial

FA – Frequência Alélica

FIOCRUZ – Fundação Oswaldo Cruz

G – Guanina

IBGE – Instituto Brasileiro de Geografia e Estatística

Kb – Kilobases

LASP – Laboratório Avançado de Saúde Pública

LPL – Lipoproteína Lipase

M – Marcador de peso molecular

PCR – Reação em cadeia da polimerase (do inglês, Polymerase Chain Reaction)

PSA – Alelos específicos de população (do inglês, Population Specific Alleles)

RNA - Ácido Ribonucléico (do inglês, Ribonucleic Acid)

SNP – Polimorfismo de único nucleotídeo (do inglês, Single Nucleotide Polimorphism)

STRs – Microssatélites (do ingles, Short Tandem Repeats)

T – Timina

VNTR - Número variável de sequências repetidas consecutivamente (do inglês, Variable Number of Tandem Repeats)

LISTA DE ILUSTRAÇÕES

FIGURA 1. Padrão de bandas observado para o <i>locus</i> AT3-I/D.....	32
FIGURA 2. Padrão de bandas observado para o <i>locus</i> PV92	32
FIGURA 3. Genotipagem dos marcadores SNP por PCR em Tempo Real	34
FIGURA 4. Gráfico demonstrando os alelos de três <i>locus</i> STRs: Cada linha representa o perfil genético de um indivíduo para os <i>loci</i> D8S1179, D21S11 e D18S51	36

LISTA DE QUADROS

QUADRO 1. Sobrenomes de conotação religiosa ou devocional mais frequentes no estado da Bahia.....	20
QUADRO 2. <i>Loci</i> , tipo de polimorfismo analisado, localização cromossômica e população onde o alelo *1 é mais freqüente.....	30
QUADRO 3. Seqüência dos <i>primers</i> e temperatura de pareamento para amplificação dos <i>loci</i> AIMs	31
QUADRO 4. Seqüência dos <i>primers</i> utilizados no estudo	35
QUADRO 5. Condições de amplificação da PCR multiplex de STRs	35

LISTA DE TABELAS

TABELA 1. Freqüência do alelo*1 para os marcadores de ancestralidade FY-Null, LPL, AT3-I/D, Sb19.3, APO, PV92, CYP3A4, CKMM, GC-1F e GC-1S numa amostra da população da Bahia – Brasil	43
TABELA 2. Freqüências alélicas dos oito STRs numa amostra da população do estado da Bahia – Brasil	44
TABELA 3. Alelos mais frequentes nos STRs da amostra e nas populações parentais.....	45
TABELA 4. Freqüência do alelo *1 para AIMs nas populações ancestrais e numa amostra da população da Bahia – Brasil.....	45
TABELA 5. Proporções de contribuição africana, européia e ameríndia numa amostra da população da Bahia-BA para os marcadores STRs e AIMs	46
TABELA 6. Proporções contribuição africana, européia e ameríndia utilizando AIMs na amostra em estudo quando separada pela presença ou ausência de sobrenome de conotação religiosa	47

TABELA 7. Proporções de contribuição africana, europeia e ameríndia utilizando STRs na amostra em estudo quando separada pela presença ou ausência de sobrenome de conotação religiosa	48
---	-----------

LISTA DE TABELAS DO APÊNDICE

APÊNDICE 1. Frequências dos alelos STRs nas amostras testadas e populações ancestrais	65
APÊNDICE 2. Análise de diferenciação populacional utilizando STRs em populações com e sem sobrenomes de conotação religiosa na Bahia – Brasil.....	67
APÊNDICE 3. Análise de diferenciação populacional utilizando AIMs em populações com e sem sobrenomes de conotação religiosa na Bahia – Brasil.....	68

RESUMO

A miscigenação entre os três principais grupos étnicos (ameríndios, europeus e africanos) originou a alta diversidade genética da população brasileira. Na Bahia a proporção de afrodescendentes é de 77,5%, sendo que em Salvador 79,8% se auto-denominam negros ou pardos. Poucos estudos descrevem a diversidade genética da população baiana e a contribuição de cada grupo étnico na sua formação. Diversos marcadores de DNA são atualmente utilizados para estimar mistura étnica em populações miscigenadas. Estes marcadores são denominados alelos específicos de população (PSAs) ou marcadores informativos de ancestralidade (AIMs) e apresentam alelos com grandes diferenciais de frequência, superiores a 30%, entre populações geográfica ou etnicamente definidas. Os microssatélites (STRs) são variantes genéticas úteis no mapeamento genético de espécies, na identificação de pessoas, mapeamento genético e análise de populações. Alguns STRs apresentam alelos com frequências marcantes em determinados grupos populacionais. Com objetivo de comparar a ancestralidade genômica avaliada com dois tipos de marcadores, foram estudados 8 microssatélites STRs autossômicos (*TH01*, *vWA31*, *D18S51*, *FGA*, *TPOX*, *D7S820*, *D3S1358*, *D8S1179*) e 9 AIMs (*FY-Null*, *LPL*, *AT3-ID*, *Sb19.3*, *APO*, *PV92*, *CYP3A4*, *CKMM*, *GC-1F* e *GC-1S*), em 203 indivíduos miscigenados da Bahia. A genotipagem foi realizada por PCR (*Polimerase Chain Reaction*), para deleções, inserções e para os microssatélites e PCR quantitativo em tempo real para mutações pontuais. As contribuições africana, européia e ameríndia observadas foram respectivamente 33,5%, 58,6% e 7,9% para os STRs e 45,08%, 45,16% e 9,75% para os AIMs, comprovando a miscigenação da população. O Índice Kappa, mostrou que a concordância entre as estimativas de ancestralidade utilizando os dois tipos de marcadores (AIMs e STRs), foi muito baixa ($\kappa = 0,12$). Foi observada associação entre sobrenome de conotação religiosa e ancestralidade africana.

Palavras Chave: Ancestralidade genômica, Short Tandem Repeats (STRs), Marcadores Informativos de Ancestralidade (AIMs).

ABSTRACT

The mixing between the three main ethnic groups (Amerindians, Europeans and Africans) produced a high genetic diversity of the Brazilian population. In Bahia, the proportion of African descent that call themselves black or brown is 77.5% and 79.8% in Salvador. Few studies describe the genetic diversity of the population of Bahia and the contribution of each ethnic group in its formation. Several DNA markers are currently used to estimate ethnic mix in admixed populations. These markers are called alleles specific population (PSAs) or ancestry informative markers (AIMs) and carry alleles with large differences in frequency above 30% between populations geographically or ethnically defined. Microsatellites (STRs) are useful genetic variants in the genetic mapping of species, identification of persons, genetic mapping and analysis of populations. Some STRs have alleles with frequencies marked in certain population groups. To compare the ancestry genomics evaluated with two types of markers were studied 8 microsatellite autosomal STRs (TH01, vWA31, D18S51, FGA, TPOX, D7S820, D3S1358, D8S1179) and 9 AIMs (FY-Null, LPL, AT3-I /D, Sb19.3, APO, PV92, CYP3A4, CK-MM, GC and GC-1F-1S) in 203 subjects with mixed Bahia. Genotyping was performed by PCR (Polymerase Chain Reaction), for deletions, insertions and for microsatellite and quantitative PCR in real time for mutations. The contributions of African, European and Amerindian observed were respectively 33.5%, 58.6% and 7.9% for the STRs and 45.08%, 45.16% and 9.75% for the AIMs, proving the mixing of population. The Kappa index showed that the correlation between the estimates of ancestry using both types of markers (AIMs and STRs), was very low ($\kappa = 0.12$). Association was found between devotional surnames and African ancestry.

Keywords: Genomic Ancestry, Short Tandem Repeats (STRs), Ancestry Informative Markers (AIMs).

1 INTRODUÇÃO

A formação da população brasileira é resultado de mais de 500 anos de miscigenação entre os três principais grupos étnicos que a compõem: africanos, europeus e ameríndios. Quando os portugueses chegaram ao Brasil por volta de 1500, este era ocupado por aproximadamente 2,5 milhões de indígenas que aqui viviam (RIBEIRO, 1995). A miscigenação começou logo após a chegada dos primeiros colonizadores portugueses e se estende até os dias de hoje. Por volta da metade do século XVI, começaram a chegar ao território brasileiro, africanos que iriam trabalhar como escravos nas fazendas de cana-de-açúcar, minas de ouro e diamantes e também nas plantações de café (KLEIN, 2002). Posteriormente, quando os portos brasileiros foram abertos para nações amigas, houve também a entrada de mais imigrantes europeus, dentre eles, os mais numerosos foram oriundos de países como Itália, Espanha e Alemanha (IBGE, 2000).

Salvador está situada na região nordeste do Brasil e foi a primeira capital brasileira. A história do seu povoamento é, portanto semelhante à do Brasil. A contribuição de cada grupo étnico na formação da nossa população foi diferente e, segundo Callegari-Jacques e Salzano (1999), dos imigrantes que chegaram ao Brasil entre 1500 e 1972, 58% eram europeus, 40% africanos e 2% asiáticos. Além disso, a proporção sexual entre os europeus foi também muito diferenciada, com contribuição quase que exclusivamente masculina. Como consequência desse processo, nossa população é considerada uma das mais heterogêneas do mundo (ABE-SANDES *et al.*, 2004).

Dados históricos e genéticos mostram que a distribuição desses três grupos étnicos ao longo do território brasileiro não ocorreu de forma homogênea, ou seja, a proporção de africanos, ameríndios e europeus difere significativamente a depender da região geográfica, sendo que se observa maior participação de africanos no povoamento da região nordeste, ameríndios no norte e de europeus no sul (IBGE, 2000). Abe-Sandes *et al.*, 2004, que estudaram a heterogeneidade do cromossomo Y, utilizando 16 marcadores bialélicos, em 209 indivíduos da Bahia e de São Paulo em quatro grupos afrodescendentes: dois urbanos (classificados fenotipicamente como negros e que afirmaram não conhecer história de mistura nos quatro avós) e dois de comunidades remanescente de quilombos e em descendentes de europeus e japoneses. Os resultados mostraram que o cromossomo Y da maioria dos afrodescendentes tinha origem da África Sub-saariana variando de 47,0% a 77,3% e, que a porção de origem européia era maior que a ameríndia. Nos descendentes de europeus 98,0 % dos cromossomos Y foram identificados como de origem européia e nos descendentes de

japoneses não foi observada nenhuma mistura com outros grupos étnicos, ou seja, todas as linhagens de cromossomo Y foram identificadas como asiáticas. Resultados semelhantes ao observado para os descendentes de europeus, com relação ao cromossomo Y, foram observados por Carvalho-Silva *et al.*, (2001), para os brasileiros brancos.

Já para o DNA mitocondrial Abe-Sandes (2002), sequenciou a região hipervariável I (HVS I) e analisou 29 marcadores em regiões codificantes e observou nos afrodescendentes também contribuição africana preponderante, variando de 76,0% a 94,0% e contribuição ameríndia variando de 5,7% a 21,9% e a contribuição européia foi observada apenas nos afrodescendentes de Ribeirão Preto-SP, com frequência de 4%. Nos descendentes de europeus as contribuições européia, africana e ameríndia foram 62,3%, 21,4% e 12,8% respectivamente. A frequência de linhagens asiáticas nos descendentes de japoneses foi 92,5%. Com os marcadores analisados não foi possível classificar algumas linhagens. Analisando indivíduos que se autodenominaram como brancos Alves-Silva *et al.*, 2000, observaram não haver diferença nas contribuições africana, ameríndias e européias.

Entretanto, estratificando por região do país, observou-se, que a maioria das matrilineagens no norte era de origem ameríndia (54%), no nordeste eram africanas (44%) e no sul as linhagens maternas eram principalmente européias (66%). Isto mostra que o grau de mistura varia a depender da região geográfica e dos marcadores genéticos utilizados.

1.1 Marcadores Polimórficos

Os principais marcadores biológicos utilizados na avaliação de diferenças entre os grupos humanos foram no passado os grupos sangüíneos do sistema ABO, proteínas e isoenzimas como a G6PD, hemoglobinas e a colinesterase sérica (OTTENSOOSER, 1944; TAVARES-NETO *et al.*,1978; SALZANO *et al.*, 1982; PRIMO-PARMO *et al.*,1986; BEUTLER, 1991).

Há pouco tempo a ciência da investigação de vínculo genético e de casos forenses pautava-se apenas nas análises sorológicas dos polimorfismos de proteínas, grupos sanguíneos e alguns marcadores genéticos. A utilização de amostras biológicas em exames forenses teve seu início por volta do princípio do século XX, com a aplicação dos grupos sangüíneos ABO em evidências relacionadas a crimes e à identificação de pessoas. Hoje, estes marcadores clássicos são pouco utilizados, sendo substituídos na maioria dos centros por marcadores moleculares.

As principais vantagens do DNA sobre a sorologia tradicional foram descritas por Weedn e Swarnen (1998). A primeira e principal delas, reside na possibilidade de sua aplicação sobre toda e qualquer fonte de material biológico em quantidade mínima. A segunda, e mais amplamente discutida é o seu potencial discriminatório. Em alguns casos, os estudos de DNA podem revelar a identificação positiva, comparativamente aos exames envolvendo o grupo sanguíneo ABO, que tem a capacidade de discriminar aproximadamente um em três indivíduos na população geral. A sensibilidade do exame de DNA constitui a terceira grande vantagem deste método. A quarta vantagem do DNA é sua resistência aos fatores ambientais. O DNA é uma molécula relativamente resistente aos ácidos, álcalis e detergentes, ao contrário dos determinantes protéicos, lipídicos e carboidratos. As proteínas podem ser degradadas de forma relativamente mais fácil e sua estrutura terciária conformacional, importante na tipagem, é facilmente desnaturável. A informação da tipagem do DNA, por sua vez, é encontrada na sequência nucleotídica, que independe da conformação da molécula. Conseqüentemente, os exames com DNA, diferentemente dos marcadores sorológicos tradicionais, podem ser realizados com maior segurança em amostras muito antigas e que estiveram expostas a maiores agressões ambientais. Finalmente, a quinta vantagem reside na possibilidade de separar o DNA da célula espermática de qualquer outro DNA celular.

Alguns *loci* de marcadores de DNA têm demonstrado ser população-específico por apresentarem grandes diferenciais de frequência entre populações geográfica ou etnicamente definidas. Estes marcadores foram denominados como “alelos específicos de população” (PSAs, do inglês *population specific alleles*) ou marcadores informativos de ancestralidade (AIMs) e apresentam diferenças nas suas frequências alélicas entre duas populações superiores a 30% (SHRIVER *et al.*, 1997; PARRA *et al.*, 1998). São representados por diferentes variantes genéticas, como polimorfismos de único nucleotídeo (SNP), inserções e deleções de nucleotídeos ou inserções *Alu*. As variantes genéticas que são encontradas apenas em uma população foram denominadas de “alelos únicos” (CHAKRABORTY *et al.*, 1991). Tanto os alelos únicos quanto os AIMs são úteis para investigação forense, estimativa de mistura populacional ou estudos de mapeamento gênico (REED *et al.*, 1973; CHAKRABORTY *et al.*, 1992; STEPHENS *et al.*, 1994).

Os marcadores do tipo STR (*Short Tandem Repeats*), também chamados de microssatélites ou repetições consecutivas curtas, são abundantes no genoma. Possuem sua sequência núcleo variando em número de repetições (DROBNIC *et al.*, 2000) e/ou sequência de unidades repetidas. Além disso, podem ser repetidas inúmeras vezes no genoma e são

altamente polimórficas (EDWARDS *et al.*, 1992; PENA *et al.*, 1993; LINS *et al.*, 1996). Esses marcadores são caracterizados como *loci* hipervariáveis, co-dominantes, multialélicos (FERREIRA *et al.*, 1995), têm taxa alta de mutação (KAYSER *et al.*, 2000) e heterozigosidade elevada (HAMMER *et al.*, 2002), sendo extremamente úteis no mapeamento genético de espécies, na identificação de pessoas ou de linhagens/clones e predisposições de doenças (ARMOUR *et al.*, 1996); em estudos de identificação humana (HAMMOND *et al.*, 1994; HAMMER *et al.*, 2002); análise de paternidade (ALFORD *et al.*, 1994; SCHUMM *et al.*, 1997), mapeamento genético (SCHUMM *et al.*, 1995; EDWARDS *et al.*, 1992) e análise de populações (DEKA *et al.*, 1995; MARTINEZ-JARRETA *et al.*, 1999; YAMAMOTO *et al.*, 1999; GUSMÃO *et al.*, 2001).

Alguns STRs apresentam alelos com frequências marcantes em determinados grupos populacionais, como, por exemplo, no *locus* TH01, cujo alelo 9 tem frequência de 0,198 em europeus (PÉREZ-LEZAUN *et al.*, 2000) e 0,002 para africanos (BOSH *et al.*, 2001) e ameríndios (MENDES-JR, 2001). Outro exemplo, no *locus* TPOX, o alelo 6 tem frequência de 0,105 em africanos (ALVES *et al.*, 2001) e 0,002 em europeus (ANJOS *et al.*, 2000) e ameríndios (WANDERLEY-SANTOS, 2001).

Marcadores que apresentam grandes diferenciais de frequência entre grupos étnicos são ideais para determinar estimativas de mistura populacional (PARRA *et al.*, 2001).

1.2 Diversidade genética populacional

As variantes genéticas polimórficas podem ser de mudanças únicas de nucleotídeos (por exemplo, substituições) como os SNPs (*Single Nucleotide Polymorphism*), variações no comprimento de repetições *in tandem* (VNTRs – número variável de repetições *in tandem* e STRs – repetições consecutivas curtas) e a inserção ou deleção de seqüências de DNA, os *Indels*, como as inserções *Alu* (EDWARDS *et al.*, 1991).

Dados genéticos geralmente corroboram os achados históricos. Por exemplo, de acordo com análises de mistura, utilizando marcadores clássicos de polimorfismo (grupos sanguíneos, proteínas séricas e eritrocitárias), a contribuição ameríndia estimada em populações miscigenadas varia de 55% na população de Alenquer no norte do país, a 5% (SANTOS *et al.*, 1995) em Santa Catarina no sul (DORNELLES *et al.*, 1999).

Estimativas de diversidade genética baseadas em STRs mostram maior similaridade na comparação das contribuições (africana, européia e ameríndia) nas diferentes regiões do país, porém ainda mantém o quadro global. Apesar da contribuição ameríndia ainda ser maior

na região norte (18%) e menor no sul (7%), a contribuição europeia varia pouco e a africana é similar nas demais regiões. Em estudo realizado na região centro-oeste brasileira com 1037 indivíduos utilizando 12 microssatélites autossômicos, relata situação similar àquela encontrada nos demais estados: maior presença de ancestralidade europeia, seguida da africana e ameríndia (CALLEGARI-JACQUES *et al.*, 2003).

Estudos como o de Carvalho-Silva *et al.* (2001), que analisou o cromossomo Y de 247 indivíduos de quatro das cinco regiões brasileiras, auto-definidos como brancos; Abe-Sandes *et al.* (2004), que estudou a heterogeneidade do cromossomo Y em 209 indivíduos da Bahia e de São Paulo, sendo afrodescendentes urbanos e de comunidades remanescente de quilombos e em descendentes de europeus e japoneses; Bortolini *et al.* (1997), Alves-Silva *et al.* (2000), Bandelt *et al.* (2001), Abe-Sandes (2002), que estudaram a distribuição dos halogrupos de DNA mitocondrial, contribuíram para o entendimento do processo de miscigenação no Brasil. Estes estudos mostraram também que as patrilinhagens foram essencialmente europeias, enquanto que as matrinhagens foram principalmente africanas e ameríndias (ALVES-SILVA *et al.*, 2000; CARVALHO-SILVA *et al.*, 2001; ABE-SANDES *et al.*, 2004).

A diversidade genética intrapopulacional avaliada por marcadores de DNA é maior nos africanos e menor nos ameríndios (ZAGO *et al.*, 1996; BOWCOCK *et al.*, 1994). Populações afro sul-americanas e africanas apresentam diversidade genética semelhantes avaliadas por marcadores protéicos (BORTOLINI *et al.*, 1998) e por marcadores VNTR (BORTOLINI *et al.*, 1998; SILVA-JR *et al.*, 1999) e STRs (BARBOSA *et al.*, 2006). Abe-Sandes, 2002 observou menor diversidade genética entre os descendentes de japoneses e nos afro-brasileiros do remanescente de quilombo de Barra, ($h = 0,394 \pm 0,119$), menor do que o observado em todos os afro-brasileiros. Em contraposição, a maior diversidade foi observada entre os afro-brasileiros de São Gonçalo-BA, também remanescente de quilombo ($h = 0,787 \pm 0,075$). E entre os grupos urbanos a maior diversidade genética foi observada entre os afro-brasileiros de Salvador-BA e Ribeirão Preto-SP quando comparado com os descendentes de europeus e de japoneses.

A diversidade gênica encontrada para a população de Salvador-BA foi de 0,407, variando entre 0,359 a 0,409, nas diferentes regiões estudadas. Esses valores são mais elevados que os valores encontrados nas populações ancestrais, de 0,250 em africanos, 0,254 em europeus e 0,262 nativos americanos (MACHADO, 2008).

1.3 Estimativa de mistura genética

A mistura gênica tem sido avaliada com base nas diferenças e similaridades das frequências alélicas de marcadores genéticos entre as populações de interesse e suas populações parentais.

A população da Bahia e da cidade de Salvador apresentam contingente considerável de afrodescendentes (77,5% e 79,8%, respectivamente), segundo dados do IBGE (2000) e conforme a autodenominação da população. O único viés eliminado no processo da autodenominação é o de não deixar a critério do entrevistador a decisão de classificar o entrevistado em um ou outro grupo, visto que a cor da pele depende da quantidade de melanina presente na derme, e suas diversas cores são determinadas por um conjunto de apenas quatro a seis genes, dos quais o mais importante parece ser o gene receptor do hormônio melanotrópico (STURM *et al.*, 1998; REES, 2003). Assim, o uso exclusivo dessa característica para classificação racial é fonte de vários vieses, não apenas culturais, mas também aqueles decorrentes da exposição aos raios solares. Entretanto, a maioria dos estudos na área da saúde que utilizam classificação racial, o fazem com base em características fenotípicas como, por exemplo, a cor da pele ou autodenominação.

A classificação racial conforme a cor da pele, segundo Parra *et al.* (2003), é falho, como demonstrado em seu estudo realizado na população brasileira através de um painel de alelos específicos de população, onde o índice de ancestralidade africana (IAA) demonstrou ser a cor da pele um pobre indicador de ancestralidade.

Utilizando 9 *loci* AIMs autossômicos (*APO*, *AT3-I/D*, *GC 1S*, *GC 1F*, *FY-Null*, *LPL*, *OCA 2*, *RB 2300* e *Sb19.3*), Parra *et al.* (1998) analisaram 10 populações afro-americanas dos Estados Unidos e da Jamaica. Os resultados mostram que a contribuição européia variou de 6,8% na Jamaica, a 22,5% em Nova Orleans.

Com o objetivo de avaliar a ancestralidade genômica na população brasileira, Parra *et al.* (2003) acrescentaram o *ICAM 1* e analisaram 200 indivíduos do sexo masculino das regiões norte, nordeste, sudeste e sul do Brasil, que se autodenominaram como brancos e uma amostra composta por 173 indivíduos da comunidade rural de Queixadinha (Vale do Jequitinhonha, Minas Gerais), classificados fenotipicamente como brancos, negros e intermediários. Para populações ancestrais, foram utilizados os dados de frequência das populações de Portugal, representando a população européia, da cidade de Santana (arquipélago de São Tomé) representando a africana (TOMÁS *et al.*, 2002), e de três tribos da Amazônia (Karitiana, Suruí e Ticuna), representando a ameríndia. Os resultados deste estudo

revelaram que no Brasil a cor da pele é um pobre indicador de ancestralidade genômica; e também, o quanto é arriscado equivaler os termos “cor ou raça” com ancestralidade geográfica e comparar “branco, caucasiano e europeu” com “negro, preto ou africano“, o que ocorre freqüentemente nos estudos e na literatura médica (BRASE CL, 1995; SCHWARTZ RS, 2001).

A estimativa de mistura populacional usando alelos com alto diferencial de freqüência entre as populações parentais pode auxiliar na reconstrução histórica de uma população. O estudo realizado na Bahia em uma comunidade isolada, remanescente de quilombo, analisou STRs (*CSF1PO*, *TH01*, *TPOX*, *F13A1*, *FESFPS* e *vWA*), características fenotípicas e sobrenomes de conotação religiosa. Os resultados encontrados foram 81% de contribuição africana e 19% de contribuição ameríndia. Para características fenotípicas, observou-se índice de fenótipo negróide (NPI) de 0,98, índice cultural negróide (NCI) de 0,24 (com dados de sobrenome de conotação religiosa) (BARBOSA *et al*, 2006).

Em outro estudo, foram testados oito AIMs (*FY-Null*, *RB*, *LPL*, *AT3-I/D*, *Sb19.3*, *APO*, *PV92* e *CYP1A1*2C*) na análise de três remanescentes de quilombo comparados a duas amostras de população urbana brasileira. Este trabalho demonstrou maior miscigenação nas populações urbanas quando comparadas com remanescentes de quilombos e tribos indígenas (LUIZON, 2007).

A estimativa de mistura avaliada por nove AIMs (*FY-Null*, *LPL*, *AT3-I/D*, *Sb19.3*, *APO*, *PV92*, *GC 1S*, *GC 1F* e *CYP3A4*) em mais de 1.200 indivíduos de diferentes bairros de Salvador, mostrou 49,2%, 36,3% e 14,5% de contribuições africana, europeia e ameríndia respectivamente (MACHADO, 2008).

Analisando 517 indivíduos infectados pelo HIV-1 provenientes da Bahia (BOMFIM, 2008), estimou a ancestralidade genômica e a sua relação com o nível sócio-econômico e a vulnerabilidade ao HIV-1/AIDS. Os resultados observados foram 48%, 35% e 17% de contribuição africana, europeia e ameríndia, respectivamente. A maioria dos participantes avaliados tinha renda igual ou inferior a 3 salários mínimos e esta característica está associada a maior ancestralidade africana.

1.4 Associação entre a ancestralidade africana e os sobrenomes de conotação religiosa

Darwin, em 1875, foi um dos primeiros pesquisadores a utilizar sobrenomes em pesquisas médicas ao estudar freqüência de casamentos entre primos na Inglaterra (AZEVEDO ES, 1980). Desde então, vários estudiosos têm chamado a atenção para o fato de

que os sobrenomes podem ser muito importantes nos estudos de populações devido a sua analogia com estudos genéticos de associação, como por exemplo, com grupos sanguíneos e isoenzimas (AZEVEDO *et al.*, 1975; BEAN *et al.*, 1970; CHAKRABORTY *et al.*, 1989).

Nas sociedades onde existe transmissão patrilinear dos sobrenomes, eles podem ser usados como correspondentes aos marcadores genéticos do cromossomo Y (TAVARES-NETO & AZEVEDO, 1978). O uso de sobrenomes pode ser mais econômico do que o uso de marcadores biológicos e permite o estudo de grandes populações a um custo menor (SANS, 2000). Na América Latina, o uso de sobrenomes para classificação racial foi mais usado, porque, de modo geral, cada indivíduo herda um ou dois sobrenomes de cada um dos pais, o que torna possível a análise de sua ancestralidade (SANS, 2000). Na Inglaterra, o uso dos sobrenomes foi utilizado para identificação étnica de imigrantes do sul da Ásia na investigação de padrões de mortalidade e fatores de risco para doenças (NICOLL *et al.*, 1986).

Os principais problemas encontrados com o uso desta metodologia são a adoção dos nomes dos senhores feudais por parte dos escravos (em sociedades de origem escravocrata), a modificação dos nomes dos índios após o batismo (em populações de ancestralidade indígena) e a perda de registros (SANS, 2000). Porém, estas limitações não foram encontradas em estudos realizados no Brasil por Tavares-Neto & Azevêdo (1977).

Alguns estudos já tentaram reconstruir a história biocultural da população da Bahia, e para esta tentativa, uma importante, para não dizer fundamental variável, é o sobrenome dos indivíduos (AZEVEDO *et al.*, 1982). Esta variável é rica em informações culturais, sociais e biológicas (AZEVEDO *et al.*, 1983), além de ser útil na identificação da origem racial em populações miscigenadas (TAVARES-NETO & AZEVEDO, 1977). Nesses estudos foram identificados três tipos de sobrenomes: de conotação religiosa, animal-plantas e “outros”, que são variáveis culturais associadas com ancestralidade africana, indígena e portuguesa, respectivamente (AZEVEDO *et al.*, 1982). Uma revisão do tipo de sobrenome adquirido pelos escravos nos séculos XVIII e XIX mostrou que grande proporção deles tinha significado religioso (AZEVEDO *et al.*, 1983). Na quadro 1, estão dispostos os sobrenomes de conotação religiosa mais frequentes no estado da Bahia (TAVARES-NETO & AZEVEDO, 1977).

A adoção dos nomes de conotação religiosa é uma prática antiga na Bahia, o que pode ser constatado através de investigação de documentos do período da escravidão (AZEVEDO *et al.*, 1982). Tais documentos revelaram que a maioria dos escravos permanecia sem sobrenome após adquirirem a liberdade. Estudos sugerem que a adoção de sobrenomes cresceu em resposta à demanda social e não por decisão voluntária súbita (AZEVEDO *et al.*,

1983). O método preferencial na aquisição de sobrenomes era a escolha de sobrenome de conotação religiosa diferente do nome da família do senhor feudal (TAVARES-NETO & AZEVEDO, 1977). Estes mesmos autores demonstraram que a frequência de sobrenomes de conotação religiosa aumenta com a proximidade do “fenótipo negro”.

O fenótipo negro é mais freqüente nas regiões litorâneas e com importância econômica, onde a mão-de-obra escrava foi bastante utilizada. Os sobrenomes de “animal-planta” têm sua frequência aumentada ao nos afastarmos para o interior do país (AZEVEDO *et al.*, 1983). Assim, à medida que nos afastamos do litoral, ocorre o “branqueamento” da população.

A presença de sobrenome de conotação religiosa é indício de ancestralidade africana e pode ser corroborada por parâmetros biológicos como o sistema sanguíneo ABO (JUNQUEIRA & WISHART, 1958) e o comprimento do cromossomo Y (BARBOSA *et al.*, 1997).

SOBRENOMES				
Aflitos	Bispo	Evangelista	Piedade	Santa Rita
Ajuda	Boa Morte	Hora	Prazeres	Santiago
Amor Divino	Bomfim	Jesus	Purificação	Santos
Amparo	Cardeal	Luz	Ramos	São Pedro
Anjos	Carmo	Mercês	Reis	Socorro
Anunciação	Chagas	Natividade	Ressurreição	Soledade
Arcanjo	Conceição	Nascimento	Rosário	Trindade
Assis	Cruz	Paixão	Sacramento	Virgem
Assunção	Encarnação	Palma	Santana	Virgens
Batista	Espírito Santo	Passos	Sant'Anna	Xavier

Fonte: Tavares-Neto & Azevedo, 1977.

QUADRO 1. Sobrenomes de conotação religiosa ou devocional, mais freqüentes na Bahia

Machado, 2008, analisou 1016 indivíduos em Salvador-BA e identificou os 10 sobrenomes de conotação religiosa mais freqüente e comparou com os dados de Tavares-Neto & Azevedo, 1977. Foi observando semelhança de 90% entre os sobrenomes mais freqüentes, sugerindo que os sobrenomes mais freqüentes em Salvador são também mais freqüentes na Bahia como um todo. A estimativa de mistura utilizando os marcadores informativos de

ancestralidade mostrou no grupo com sobrenome de conotação religiosa contribuição africana, europeia e ameríndia de 53,1 %, 31,2 % e 15,6 % respectivamente. Esta mesma análise no grupo sem sobrenome religioso mostrou maior contribuição europeia (44,7%), seguida da contribuição africana e ameríndia (42,7% e 12,6%). Resultados semelhantes foram observados por Bomfim, 2008, estudando uma amostra da população da Bahia. Nestes, foi observado que a contribuição africana aumenta à medida que aumenta o “fenótipo negróide”, sendo 32,0%, 37,0%, 47,0%, 61,0% e 64,0%, respectivamente para os grupos classificados fenotipicamente como branco, mulato claro, mulato médio, mulato escuro e negro.

Os dados moleculares da contribuição africana observada nos grupos com e sem sobrenome de conotação religiosa confirmam a associação deste tipo de sobrenome à origem africana.

2 JUSTIFICATIVA

O censo de 2000 realizado pelo IBGE mostrou que a população da Bahia e da cidade de Salvador apresentam contingente considerável de afrodescendentes (77,5% e 79,8%, respectivamente).

Dados da literatura apontam a existência de risco diferencial de desenvolvimento de algumas doenças, a depender do grupo étnico ou região geográfica, podendo ocorrer associação com algumas doenças como, por exemplo, doenças cardiovasculares, tromboembólicas, câncer de próstata, hiperhomocisteína, anemia falciforme, entre outras (PENA, 2005).

Alguns AIMs já são usados em estudos de associação com susceptibilidade a patologias, como por exemplo, o *CYP3A4*. O alelo*1 do marcador *CYP3A4* já foi descrito como associado ao desenvolvimento de câncer de próstata, e sua frequência encontra-se aumentada em afroamericanos (KITTLES *et al.*, 2002). Há também associação entre o alelo *GCI-F* e a susceptibilidade a doença pulmonar obstrutiva crônica (SANDFORD *et al.*, 1997), também mais frequente entre africanos. Outro alelo bastante comum entre os africanos e que demonstra associação com susceptibilidade é o *FY-Null*, onde o alelo G fornece proteção completa para infecção pelo *Plasmodium vivax* (TOURNAMILLE *et al.*, 1995) e aumenta a vulnerabilidade ao HIV-1 (HE *et al.*, 2008).

De posse destas informações, reconhecemos que é de fundamental importância analisar a contribuição de cada população ancestral na população de Salvador, baseando-se não apenas nas características fenotípicas, mas em marcadores genéticos. A caracterização será importante para fornecer informações sobre eventual risco populacional para determinadas doenças que apresentam associação com grupo étnico ou geográfico, pois através das frequências encontradas para marcadores associados com susceptibilidade, como *CYP3A4*, *GCI-F* e *FY-Null*, poderemos auxiliar na prevenção de enfermidades associadas.

Não existem estudos no estado da Bahia demonstrando o poder dos STRs na avaliação da estimativa de mistura genética em relação aos AIMs, o que poderia ser muito útil na redução de tempo e custos nas pesquisas que visam esclarecer aspectos obscuros no processo de miscigenação e formação da nossa população, visto que para genotipagem dos STRs existem kits comerciais, o que torna esta análise mais rápida e economicamente viável.

Além disso, o conhecimento da ancestralidade genética poderá ser utilizado em estudos caso/controle permitindo avaliar se as diferenças nas frequências alélicas e

associações observadas são derivadas de uma relação causal ou estruturação populacional, uma vez que vários trabalhos na área de saúde utilizam “raça” em estudos de associação.

Essas informações serão também importantes para orientar o desenvolvimento de ações na área de saúde, voltadas ao melhor atendimento da população, especialmente para aquelas ações associadas a determinados grupos, regiões geográficas e que fazem parte do nosso componente ancestral.

A pesquisa não traz risco e nem benefícios diretos ao indivíduo participante e nem para gerações futuras. O que será produzido visa colaborar com trabalhos futuros sobre ancestralidade. Será produzido conhecimento básico sobre diversidade genética e índice de mistura genética da nossa população, entretanto este conhecimento não terá impacto individual.

3 OBJETIVOS

3.1 Geral

- Estimar e comparar a mistura genética avaliada por STRs e por AIMs, com o intuito de verificar se os STRs podem ser utilizados como marcadores de ancestralidade.

3.2 Específicos

- Estimar o grau de mistura genética de indivíduos nascidos na Bahia utilizando oito *loci* de microssatélites STRs autossômicos (*TH01*, *vWA31*, *D18S51*, *FGA*, *TPOX*, *D7S820*, *D3S1358*, *D8S1179*) e dez AIMs (*FY-Null*, *LPL*, *AT3-I/D*, *Sb19.3*, *APO*, *PV92*, *CYP3A4*, *CKMM*, *GC-1F* e *GC-1S*);
- Comparar as frequências dos polimorfismos observados nesta amostra com as frequências nas populações ancestrais: africanas, européias e ameríndias;
- Observar a existência de associação entre a ancestralidade africana e os sobrenomes de conotação religiosa das pessoas que compõem a amostra;
- Descrever a diversidade genética desta amostra utilizando estes polimorfismos;
- Verificar estruturação populacional.

4 MATERIAL E MÉTODOS

4.1 Amostra, Critérios de inclusão, autodenominação e análise fenotípica

Neste estudo foram utilizadas 203 amostras de DNA extraídas que foram armazenadas de outro estudo envolvendo 517 indivíduos infectados pelo HIV-1, os quais realizaram exames periódicos no Laboratório de Retrovírus do Hospital Universitário Professor Edgard Santos (HUPES), da Universidade Federal da Bahia (UFBA), no período entre Julho a Novembro de 2006. O tamanho amostral foi estimado tomando como base os trabalhos publicados utilizando a mesma metodologia.

No momento da coleta das amostras, os indivíduos que concordavam em participar, assinaram o termo de consentimento livre e esclarecido (TCLE) e foram entrevistados para coleta dos dados. Neste questionário, os pacientes se autodenominarão (dentre as seguintes opções) como: negro, pardo, branco, indígena ou outro. Após a auto-classificação, foi realizada a caracterização fenotípica levando em consideração as seguintes características:

- Cor da pele: preta, marrom ou branca.
- Textura de cabelo: crespo, ondulado ou liso.
- Nariz: achatado, mediano ou fino.
- Formato dos lábios: grossos, medianos ou finos.

De acordo com esses critérios, os indivíduos de pele branca, cabelo liso ou ondulado, nariz mediano, lábios finos ou medianos são classificados como brancos. Os indivíduos de pele preta, cabelo crespo e lábios grossos são classificados como negros. Os demais participantes foram classificados como mulatos em três categorias a depender da quantidade de características compartilhadas (mulatos claros, mulatos médio, mulatos escuros) (AZEVEDO *et al.*, 1982).

Além disso, os indivíduos foram questionados sobre “raça/cor” de seus pais e avós e histórico de mistura, de acordo com suas próprias definições, e também relataram sobre os locais de nascimento de cada aparentado.

As amostras foram processadas e analisadas no Centro de Diagnóstico do GACC – CDG e no Laboratório Avançado de Saúde Pública – LASP - CPqGM/FIOCRUZ-BA.

4.2 Critérios de exclusão

Todos os indivíduos que não concordarem em participar do estudo se encaixam neste critério.

4.3 Análise de sobrenomes

Os sobrenomes foram analisados para identificar aqueles de conotação religiosa e relacionar tais sobrenomes com a ancestralidade, de acordo com a classificação racial por autodenominação, análise fenotípica e genotípica desses indivíduos. Vide quadro 1, onde estão dispostos os sobrenomes de conotação religiosa mais freqüente na Bahia.

4.4 Marcadores analisados

A análise dos marcadores STRs para este estudo foi realizada no laboratório do Centro de Diagnóstico do GACC – CDG. Já a análise dos AIMs, foram realizadas no Laboratório Avançado de Saúde Pública – LASP - CPqGM/FIOCRUZ-BA no estudo anterior que cedeu as amostras para as tipagens dos STRs.

No presente estudo foram analisados dez marcadores de ancestralidade, sendo um polimorfismo de inserção/deleção - *AT3-*I/D**, três inserções *Alu* - *SB19.3*, *APO* e *PV92*, e 5 *SNPs* – *FY-Null*, *CKMM*, *LPL*, *CYP3A4*, *GC-1F* e *GC-1S*. Os diferentes alelos encontrados nos marcadores de ancestralidade foram denominados de alelo*1 e alelo*2. Nos polimorfismos do tipo *in/del* e nas inserções *Alu*, o alelo *1 foi caracterizado pela presença da inserção. Nos SNP, o alelo*1 é aquele cujo nucleotídeo abole o sítio de restrição.

4.4.1 Inserção/ Deleção – *AT3-*I/D**

A antitrombina III (*AT3*) é um membro da família dos inibidores da serina, e se caracteriza por ser uma pequena molécula que inativa irreversivelmente várias enzimas da coagulação, tais como os fatores IXa, Xa, XIIa, sendo um inibidor da coagulação neutralizando a trombina. É uma glicoproteína formada por uma cadeia de 432 aminoácidos, com um peso molecular de 58 kDa (kilodaltons) produzida no fígado. O gene da *AT3* localiza-se no cromossomo 1 (1q25.1), possui 19kb e sete éxons (LIU *et al.* 1995).

Neste estudo, foi analisado o polimorfismo de comprimento de 76bp (inserção/deleção) na região 5' do éxon 1 (LIU *et al.* 1995). A presença desta inserção gera um fragmento de 572bp e caracteriza o alelo *AT3-I/D*1*. Este polimorfismo é útil em análises de populações híbridas por apresentar uma frequência da inserção de 0,858 em populações africanas, diferenciando assim, africanos de europeus e africanos de nativo-americanos.

4.4.2 Inserções *Alu*

Inserções *Alu* são assim chamadas por terem sido inicialmente descritas como uma fração repetitiva de DNA que exibia um sítio de restrição para a enzima *Alu*. (HOUCK *et al.*, 1979). Estas inserções são pequenos elementos repetitivos intercalantes (do inglês *SINE - short interspersed repetitive elements*) que possuem aproximadamente 300pb e são encontrados exclusivamente em primatas. (WATKINS *et al.* 2001). Elas são originadas por transcrição reversa de um RNA intermediário (7 SL RNA) em um processo denominado retrotransposição e inseridas em diversas partes do genoma por retroinserção. Todas as inserções em um *locus* são idênticas por descendência, devido à improbabilidade do fenômeno de inserção ocorrer duas vezes no mesmo *locus* (BATZER & DEININGER, 1991).

De acordo com a série hierárquica temporal da mutação, os elementos *Alu* foram agrupados em 3 famílias principais designadas como J, S e Y, representando a mais antiga, a intermediária e a mais recente família *Alu*, respectivamente. Essas famílias foram ainda divididas em sub-famílias específicas de acordo com a identidade nucleotídica entre elas (BATZER *et al.*, 1990; 1996; JURKA & SMITH, 1988). São estimados que aproximadamente 5000 elementos *Alu* jovens sejam específicos de humanos (BATZER & DEININGER, 1991). Isso indica que o estado ancestral desses polimorfismos é a ausência da inserção no genoma humano e que o estado derivado da mesma é o ganho do fragmento *Alu* em um *locus* específico.

Devido às inserções *Alu* serem idênticas por descendência, livres de homoplasia e por serem capazes de informar o estado ancestral da mutação, esses polimorfismos são bastante úteis em estudo de mapeamento genético e de relações entre populações. Estudos mostram que elas apresentam grandes diferenças na distribuição das frequências alélicas em populações de distintas origens geográficas (COTRIM, 2003), sugerindo que essas inserções podem ser excelentes marcadores raciais, sendo úteis para estimar a composição étnica de populações miscigenadas como a do Brasil (MENDES-JUNIOR, 2001).

As inserções *Alu* que foram analisadas nesse estudo são mutações que não determinam fenótipo alterado por estarem localizadas próximas e não dentro de genes

funcionais, apresentando assim frequências polimórficas entre as populações, são elas: *Sb19.3* pertence a subfamília Yb8 e está localizada no cromossomo 19p12 (ARCOT *et al.*, 1998). A presença dessa inserção *Alu* gera um fragmento de aproximadamente 457pb e caracteriza o alelo *Sb19.3*1*. O locus *Alu* APO está próximo ao complexo de genes da apolipoproteína AI-CIII-AIV no braço longo do cromossomo 11 (KARATHANASIS, 1985). A presença dessa inserção *Alu* gera um fragmento de aproximadamente 409pb e caracteriza o alelo APO*1. O *Alu* PV-92 localiza-se no cromossomo 16 (BATZER *et al.*, 1994) e a presença dessa inserção *Alu* gera um fragmento de aproximadamente 400pb e caracteriza o alelo PV92*1, (Figura 2).

4.4.3 SNP (Polimorfismos de Nucleotídeo Único)

Foram analisados seis polimorfismos de nucleotídeos simples – SNP (do inglês *Single Nucleotide Polymorphisms*) nos loci *FY-Null*, *CKMM*, *LPL*, *CG-1F*, *CG-1S* e *CYP3A4*. Esses polimorfismos ocorrem em regiões não codificadoras e são, portanto, menos afetados por processos de seleção natural, refletindo, desta maneira, com maior precisão a história evolutiva humana (ALVES *et al.*, 2005).

4.4.3.1 FY-Null

O gene DARC (do inglês, *Duffy antigen receptor for chemokines*) é composto por um único éxon e uma única substituição (T46C) na região promotora do gene cria um alelo FY*B silencioso, também chamado nulo (*FY-Null*) ou FY*O. Esta mutação confere resistência à malária *vivax* por abolir a expressão do RNAm, visto que o antígeno do grupo sanguíneo Duffy funciona como receptor eritrocitário para o parasita da malária *Plasmodium vivax* (TOURNAMILLE *et al.*, 1995).

O sistema Duffy foi o primeiro grupo sanguíneo a ter o locus genético atribuído ao cromossomo autossômico específico, o cromossomo 1 (DONAHUE *et al.*, 1968). Os antígenos de Duffy parecem ser proteínas multiméricas da membrana de eritrócitos compostas por diferentes subunidades. Uma glicoproteína de 35 a 45 kD nomeada GPD é a subunidade principal da proteína complexa e tem as determinantes antigênicas definidas por anti-Fy (a), anti-Fy (b), e os anticorpos anti-Fy6 (HADLEY *et al.*, 1984). Em africanos, o fenótipo dominante é o Fy (a-b-), no qual os eritrócitos não possuem os antígenos Fya e Fyb e resistem à invasão pelo *Plasmodium vivax*, sugerindo uma resposta adaptativa para resistência à malária (TOURNAMILLE *et al.*, 1995).

Neste trabalho, estudaremos a transição de uma adenina (A) para uma guanina (G) na posição -46 da região promotora deste gene. A população europeia e nativa americana

apresenta elevada frequência do alelo A, que é o alelo*1, enquanto que os africanos apresentam apenas o alelo G, alelo*2.

4.4.3.2 CKMM

A proteína denominada creatina cinase existe como enzima dímera, uma enzima encontrada no músculo formada por 2 subunidades idênticas M (MM) e outra enzima encontrada no cérebro formada por duas subunidades idênticas de B (BB) (DAWSON *et al.*, 1968). Outros tecidos mostram ainda a forma híbrida desta enzima (CKMB).

A isozima dimérica creatina cinase está envolvida na manutenção dos níveis intracelulares de ATP, particularmente em tecidos que têm elevada demanda de energia. A isozima MM é encontrada exclusivamente em músculos estriados, enquanto que a isozima BB encontra-se no cérebro, nervos e músculos liso. A enzima híbrida é encontrada no músculo cardíaco.

O gene da creatina cinase está localizado no cromossomo 19, sendo estudado o polimorfismo caracterizado por uma transição de citosina (C) por timina (T), no éxon 8 foi analisado no presente estudo. O alelo C é o prevalente nas populações europeias e africanas, e o alelo T, alelo*1, o mais freqüente em asiáticos.

4.4.3.3 LPL

A lipoproteína lipase – LPL (do inglês *lipoprotein lipase*) é uma enzima catalítica que participa do metabolismo de triglicérides através do catabolismo de partículas lipoprotéicas ricas em triglicérides, tais como quilomícrons e VLDL (do inglês, *very low density lipoproteins*) (STEPANOV & LEMZA 1993). O gene da LPL, localizado no cromossomo 8, contém 10 éxons num tamanho de 30kb. Muitas alterações têm sido descritas neste gene, sendo associadas a diversas manifestações clínicas (FUNKE *et al.*, 1987,1988; LI *et al.*, 1988). Para este trabalho, será analisado uma variante no *locus* do gene LPL, que se caracteriza por uma transição de timina (T) por citosina (C) no intron 6 do gene. Os dois alelos apresentam frequências diferenciadas em populações, sendo o alelo T, alelo*1, mais freqüente em africanos, enquanto que o alelo C é mais freqüente em ameríndios.

4.4.3.4 CYP3A4

O citocromo P450 3A4 é responsável pelo metabolismo oxidativo de uma grande variedade de xenobióticos, incluindo 60% de todas as drogas clinicamente utilizadas (LEHMANN *et al.*, 1998). Esta proteína se expressa predominantemente no fígado e é transcricionalmente ativada por diversas estruturas xenoquímicas (WRIGHTON & STEVENS, 1992).

Estudamos o polimorfismo caracterizado por uma transição de adenina (A) por guanina (G). O alelo A, alelo*1, é mais freqüente em europeus e asiáticos, enquanto que o alelo G apresenta elevada freqüência em africanos.

4.4.3.5 GC-1F e GC-1S

Estes são dois polimorfismos encontrados no gene da proteína denominada globulina ligadora de vitamina D – VDBG (do inglês *vitamin D-binding alpha-globulin*) (DAIGER *et al*, 1975) Este gene está localizado no cromossomo 4, possui 13 éxons e 3 polimorfismos já foram descritos neste gene, GC-1F, GC-1S, GC-2 (WITKE *et al*, 1993).

Estes polimorfismos são encontrados nas posições 34 e 45 da seqüência gênica e a combinação dos genótipos nessas duas posições são responsáveis pela caracterização do SNP GC. As combinações T/T, na posição 34 do gene e C/C na posição 45 determina o alelo GC-1F; o alelo GC-1S é caracterizado pela combinação G/G na posição 34 e C/C na posição 45; a combinação T/T na posição 34 e A/A na posição 45, caracteriza o alelo GC-2. As combinações de heterozigotos são - F/S - T/G (posição 34) e C/C (posição 45); F/2 – T/T (posição 34) e C/A (posição 45); S/2 – T/G (posição 34) e C/A (posição 45);

Locus	Polimorfismo	Localização	População com maior freqüência do Alelo*1
<i>AT3-ID</i>	76bp indel	1q25.1	Africana
<i>LPL</i>	T/C	8p21.3	Africana
<i>CG-F</i>	G/T	4q13.3	Africana
<i>GC-S</i>	C/A	4q13.3	Europeu
<i>FY-Null</i>	A/G	1q23.2	Europeu
<i>Sb19.3</i>	Inserção Alu	19p12	Europeu
<i>APO</i>	Inserção Alu	11q23.3	Europeu
<i>PV92</i>	Inserção Alu	16q23.3	Ameríndia
<i>CKMM</i>	C/T	19q13.32	Ameríndia
<i>CYP3A4</i>	A/ G	7q22.1	Ameríndia

QUADRO 2. *Loci*, tipo de polimorfismo analisado, localização cromossômica e população onde o alelo *1 é mais freqüente.

4.5 Reação em cadeia da polimerase (PCR)

O polimorfismo de inserção/deleção, as inserções *Alu* (AT3-I/D e Sb19.3, APO e PV92, respectivamente) foram genotipados por PCR utilizando os *primers* e condições da PCR descritas na literatura (SHRIVER *et al.*, 1997; PARRA *et al.*, 1998). As amplificações foram realizadas em termocicladores e sob as seguintes condições: 1 ciclo de 94°C por 6 minutos e temperatura de pareamento específica para cada par de primer (T.A.) por 2 minutos; 35 ciclos para extensão das fitas constituídos por 1 minuto a 72°C, 30 segundos a 94°C e temperatura de pareamento, por 1 minuto; finalizando a PCR com 1 ciclo final de 72°C e resfriamento a 4°C. As temperaturas de pareamento e os *primers* para cada *loci* estão sumarizados na Tabela 3. Os produtos amplificados foram visualizados em gel de agarose a 2%, e em gel de poliacrilamida a 6% corados com brometo de etídeo e visualizado em luz ultravioleta ou corados com nitrato de prata como mostram as figuras 1 e 2.

As reações da PCR foram realizadas com volume total de 25µl, composto por: 100 ng de DNA, 10 mM tris-HCl pH 8,5, 50 mM KCl, 1,5 mM MgCl₂ 0,01% gelatina, 200 mM de cada dNTP, 0,25 mM de cada *primer*, e 1U de *Taq* DNA polimerase.

<i>Loci</i>	Seqüência dos <i>primers</i>	Temperatura de pareamento
<i>AT3-I/D</i>	F: 5'-CCACAGGTGTAACATTGTGT-3' R: 5'-GAGATAGTGTGATCTGAGGC-3'	54°C
<i>APO</i>	F: 5'-AAGTGCTGTAGGCCATTTAGATTAG-3' R: 5'-AGTCTTCGATGACAGCGTATACAGA-3'	66°C
<i>PV 92</i>	F: 5'-AACTGGGAAAATTTGAAGAGAAAGT-3' R: 5'-TGAGTTCTCAACTCCTGTGTGTTAG-3'	56°C
<i>SB19.3</i>	F: 5'-TCTAGCCCCAGATTTATGGTAACTG-3' R: 5'-AAGCACAATTGGTTATTTTCTGAC-3'	63°C

QUADRO 3. Seqüência dos *primers* e temperatura de pareamento para amplificação dos *loci* AIMS.

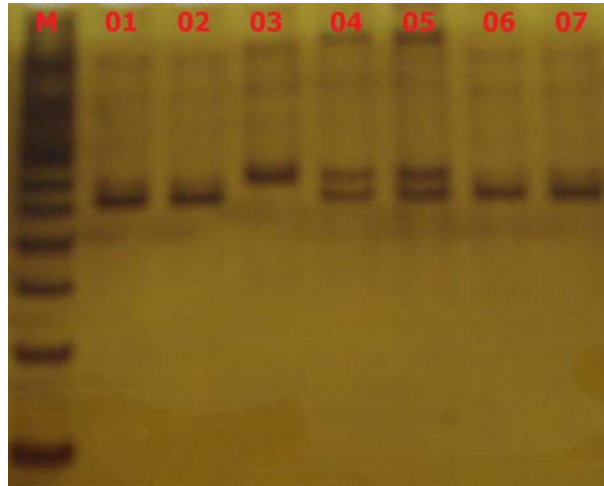


FIGURA 1. Padrão de bandas observado para o *locus* AT3-I/D. M: Marcador de peso molecular. Raias 1, 2, 6 e 7: indivíduos homocigotos para a ausência da inserção (496 pb). Raias 4 e 5 indivíduos heterocigotos (572 pb e 496 pb). Raia 3: indivíduo homocigoto para a presença da inserção (572 pb).

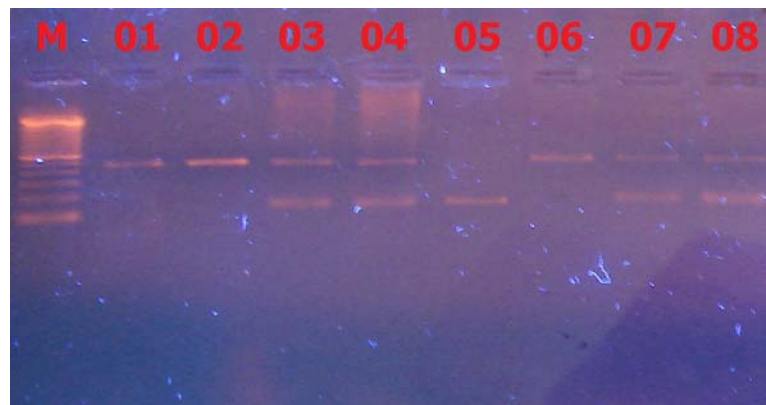


FIGURA 2. Padrão de bandas observado para o *locus* PV92. M: Marcador de peso molecular. Raias 1, 2 e 6: indivíduos homocigotos para a presença da inserção (400 pb). Raia 3, 4, 7 e 8: indivíduos heterocigotos (400 pb e 100 pb). Raia 5: indivíduo homocigoto para a ausência da inserção (100 pb).

4.6 Coloração com nitrato de prata e secagem do gel

A coloração com nitrato de prata foi realizada com 3 diferentes soluções em 3 etapas: impregnação com nitrato de prata, revelação e fixação das bandas visualizadas. As soluções usadas equivalentes à coloração de 1 gel são:

Solução de nitrato de prata: 0,3 g nitrato de prata; 2 ml de H₂O.

Solução fixadora: 25 ml etanol (PA) e 2 ml de ácido acético glacial (PA) dissolvidos em 273 ml de H₂O (volume final 300ml).

Solução reveladora: 4,5 g de NaOH; 200 ml de H₂O e 1 ml de formaldeído que foi adicionado no momento da coloração.

Este protocolo de coloração é adaptado de Sanguinetti *et al.* (1994) e consiste dos seguintes passos:

Fixação: O gel é colocado em um recipiente de vidro contendo 150 ml de solução fixadora, por 15 minutos.

Impregnação com Nitrato de prata: Adiciona-se 2,0 ml de solução de nitrato de prata, e agita-se por 5 minutos. Esta solução é então descartada e o gel lavado em água destilada por cerca de 10 segundos.

Revelação: A solução reveladora é então adicionada cuidadosamente no recipiente contendo o gel, e logo após adicionado o formaldeído. O gel é submetido à agitação por alguns minutos até a visualização das bandas.

Bloqueio da reação: A reação foi bloqueada com a adição de 150 ml de solução fixadora.

Secagem do gel: todos os géis passam por processo de secagem à temperatura ambiente entre duas folhas de papel celofane e foram armazenados para análises e confirmações posteriores.

4.7 Genotipagem dos SNP: PCR em tempo real

A técnica de discriminação alélica por PCR em tempo real foi utilizada para genotipar os polimorfismos de nucleotídeos únicos – *SNP* nos *loci* *FY-Null*, *LPL*, *CKMM*, *GC-1F*, *GC-1S* e *CYP3A4*. Esta técnica baseia-se no uso de uma sonda fluorescente, com seqüências específicas para cada alelo (uma sonda para o alelo mutante e outra para o alelo selvagem), onde cada sonda está marcada com um fluoróforo diferente. Cada sonda hibridiza com a seqüência alvo gerando sinal fluorescente proporcional à concentração dos produtos amplificados, permitindo assim, correlacionar a intensidade de sinal coletada com a quantidade de produto amplificado. O sistema utilizado para realização dessa reação foi o TaqMan™ da *Applied Biosystems* do Brasil.

Nesse sistema, a sonda se hibridiza na região complementar do DNA alvo que está localizada entre os sítios de ligação dos *primers*. Cada sonda é marcada com um fluoróforo diferente chamado *reporter* (R) na extremidade 5'. Na extremidade 3' existe outro fluoróforo chamado *quencher* (Q) cuja função é absorver a emissão de fluorescência do *reporter*, quando a sonda se encontra intacta. Durante a reação de PCR ocorre a hibridização dos *primers* e da

sonda, no fragmento de DNA, e, durante a extensão dos *primers*, a enzima, *Taq* DNA polimerase por sua atividade 5'-3' exonuclease irá também clivar a sonda a partir da extremidade 5'. Desta forma, a fluorescência emitida pelo *reporter* pode ser detectada e sua intensidade aumentará em função do número de cópias que estão sendo amplificadas. Durante a termociclagem é feita a discriminação alélica, onde se quantifica a intensidade de cada fluoróforo, correspondendo a um determinado alelo. Os genótipos selvagens ou mutantes emitiram fluorescência correspondente a cada alelo marcado; e se ambos os alelos apresentarem fluorescência teremos um genótipo heterozigoto (Figura 3).

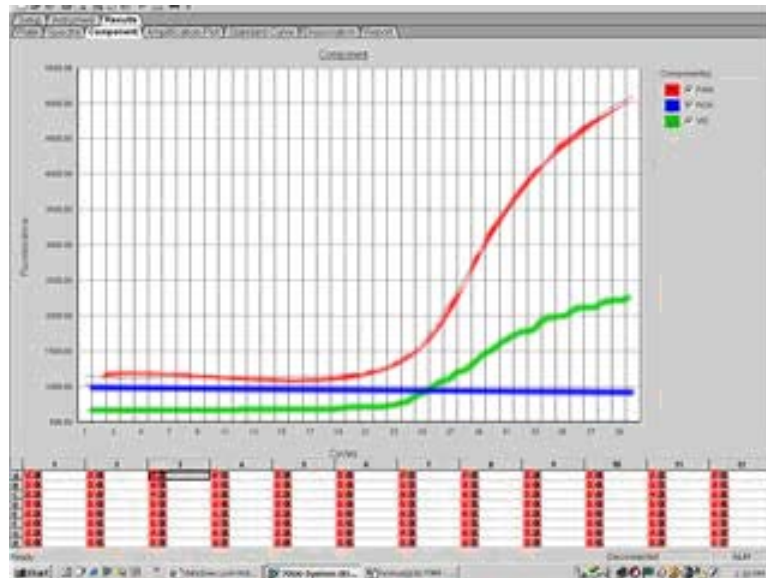


FIGURA 3. Genotipagem dos marcadores SNP por PCR em Tempo Real. Gráfico da fluorescência capturada durante a termociclagem em aparelho de PCR em Tempo Real por cada sonda. Verde: VIC e Vermelho: FAM, correspondendo a cada um dos alelos dos marcadores analisados. No exemplo acima temos um genótipo heterozigoto, pois foi detectada fluorescência para ambas as sondas.

4.8 Microssatélites STRs

Oito marcadores STRs foram analisados e genotipados utilizando o método da PCR convencional e seus alelos discriminados por eletroforese capilar (Figura 4). A reação de PCR multiplex dos STRs foi realizada em termociclador ABI 9700 da Applied Biosystems. Os primers dos marcadores STRs são marcados com três tipos de fluorescência diferentes (NED, JOE e 6FAM), para que seja possível a detecção e diferenciação entre os alelos de cada *locus* pesquisado. A Tabela 4 apresenta os marcadores STRs utilizados no estudo.

<i>Locus</i>	Seqüência 5' - 3' e fluorescência na porção 5'	Localização
<i>TPOX</i>	NED ACT GGC ACA GAA CAG GCA CTT AGG GGA GGA ACT GGG AAC CAC ACA GGT	2p25.3
<i>D8S1179</i>	NED TTT TTG TAT TTC ATG TGT ACA TTC G CGT AGC TAT AAT TAG TTC ATT TTC A	8q24.13
<i>D3S1358</i>	NED ACT GCA GTC CAA TCT GGG T ATG AAA TCA ACA GAG GCT TG	3p21.31
<i>vWA</i>	JOE CCC TAG TGG ATG ATA AGA ATA ATC GGA CAG ATG ATA AAT ACA TAG GAT GGA TGG	12p13.31
<i>FGA</i>	JOE AAG GCT GCA GGG CAT AAC ATT ATC CAG CCA CAT ACT TAC CTC CAG TCG	4q28
<i>D7S820</i>	JOE TGT CAT AGT TTA GAA CGA ACT AAC G CTG AGG TAT CAA AAA CTC AGA GG	7q21.11
<i>D18S51</i>	JOE CAA ACC CGA CTA CCA GCA AC GAG CCA TGT TCA TGC CAC TG	18q21.33
<i>TH01</i>	6FAM GTG GGC TGA AAA GCT CCC GAT TAT GTG ATT CCC ATT GGC CTG TTC CTC	11p15.5

NED, JOE e 6FAM são as fluorescências que marcam os primers da reação multiplex.

Fonte: Adaptado de http://www.cstl.nist.gov/strbase/str_fact.htm.

QUADRO 4. Seqüência dos primers utilizados no estudo.

As regiões de interesse foram amplificadas por PCR. Após a amplificação o produto de cada foi desnaturado com auxílio de formamida e baixa temperatura para que a fita dupla de DNA permanecesse em fita simples, e então a amostra foi analisada no seqüenciador. A reação de PCR multiplex é um ensaio que amplifica oito *loci* (*D8S1179*, *D7S820*, *D3S1358*, *TH01*, *VWA31*, *TPOX*, *D18S51*, *FGA*) STRs em única reação de amplificação de DNA. As condições de amplificação estão descritas na figura 4.

Etapas	Temperatura (°C)	Tempo (minutos)	Ciclos
Desnaturação inicial	95	5	1
Desnaturação	94	1	25
Pareamento	59	1	
Extensão	72	1	
Extensão final	60	60	1
Final	4 - 25	∞	

Fonte: manual de kit comercial AmpFISTR® Identiflier® (modificado).

QUADRO 5. Condições de amplificação da PCR multiplex de STRs.

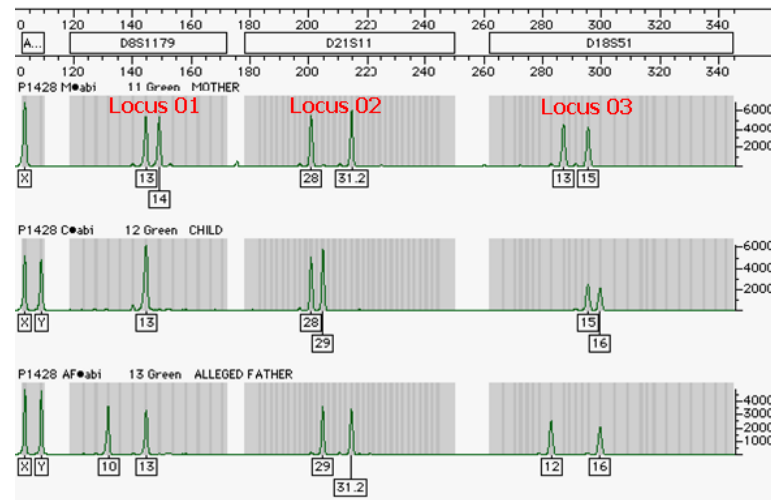


FIGURA 4. Gráfico demonstrando os alelos de três *locus* STRs: Cada linha representa o perfil genético de um indivíduo para os *loci* D8S1179, D21S11 e D18S51. O primeiro indivíduo é uma mulher, e apresenta os alelos 13 e 14 para o *locus* D8S1179, 28 e 31.2 para o *locus* D21S11 e os alelos 13 e 15 para o *locus* D18S51. Para os outros indivíduos faz-se a mesma interpretação.

4.9 Análises estatísticas

Após a genotipagem, foram testadas a aderência ao equilíbrio de Hardy-Weinberg, a análise de associações par-a-par entre *loci*, déficit de heterozigotos, diferenciação populacional e calculada as freqüências alélicas através do programa GENEPOP (RAYMOND *et al.*, 1995).

As estimativas de mistura genética foram geradas a partir do programa ADMIX, utilizando freqüências de populações ancestrais ameríndias, européias e africanas disponíveis na literatura (CHAKRABORTY *et al.*, 1991).

O índice *Kappa* foi utilizado para avaliar a concordância entre as estimativas de ancestralidade utilizando os dois tipos de marcadores (AIMs e STRs). Conforme índice *Kappa*, os testes foram avaliados seguindo os seguintes critérios: se índice de concordância for $< 0,2$ – concordância muito fraca; $> 0,2$ e $< 0,4$ – concordância fraca; $> 0,4$ e $> 0,6$ – concordância moderada; $> 0,6$ e $< 0,8$ – bom grau de concordância; se $> 0,8$ – ótimo grau de concordância. Este cálculo foi realizado no software R Projects 2.7.2 para Windows.

4.10 Considerações éticas

No momento da coleta do material biológico, todos os indivíduos assinaram o TCLE, que foi aprovado pelo Comitê de Ética em Pesquisas do Centro de Pesquisas Gonçalo Moniz/ FIOCRUZ (Parecer nº. 84/2006 – CEP/ CPqGM/ FIOCRUZ).

5 RESULTADOS E DISCUSSÃO

(Os resultados e a discussão estão sumarizados no manuscrito, ainda não publicado; tabelas complementares foram adicionadas aos apêndices)

Manuscrito

Estimativa de mistura étnica avaliada por Marcadores Informativos de Ancestralidade (AIMs) e Microssatélites (STRs) numa amostra da população da Bahia.

Teló, E.T.^{1,2}; Bomfim, T.F.¹, Machado, T.M.B.¹, Abé-Sandes, K^{1,3}

1- Laboratório Avançado de Saúde Pública/Centro de Pesquisas Gonçalo Moniz/FIOCRUZ,

2- Centro de Diagnóstico do GACC - CDG, 3- Universidade do Estado da Bahia (UNEB)

Resumo

Entre os três principais grupos étnicos que compõem a população brasileira estão os africanos, europeus e ameríndios. Com mais de 500 anos de miscigenação, nossa população é uma das mais heterogênea do mundo. Contudo a distribuição dos grupos étnicos ancestrais ao longo do território brasileiro não ocorreu de forma homogênea, diferindo significativamente a depender da região geográfica. Dados do IBGE 2000 mostram que em Salvador o percentual de afrodescendentes, por autodenominação é de 79,8%. Para estimarmos a contribuição dos grupos ancestrais nesta população foram analisados 203 indivíduos, provenientes do estado da Bahia para oito marcadores STRs (TH01, vWA31, D18S51, FGA, TPOX, D7S820, D3S1358, D8S1179) e nove AIMs (*AT3-ID*, *APO*, *SB19.3*, *PV92*, *FY-Null*, *LPL*, *CKMM*, *GC-1F*, *GC-1S* e *CYP3A4*). A miscigenação da população de Salvador foi confirmada pela diferença das frequências desta população com as ancestrais. A estimativa de mistura populacional para os AIMS encontrada foi de 45,08 % para contribuição africana, 45,16 % para europeia e 9,75 % para contribuição ameríndia. Com STRs os valores observados foram de 33,48 %, 58,58 % e 7,94 % para contribuição africana, europeia e ameríndia respectivamente. O índice de concordância entre as estimativas de ancestralidade utilizando os dois tipos de marcadores (AIMs e STRs) foi muito baixo ($\kappa = 0,12$). A observação de indivíduos com sobrenomes de conotação religiosa analisados com AIMs foi de 53,32 % de contribuição africana em seu perfil genético. A variabilidade genética observada para AIMs apresentou déficit de heterozigotos em quatro marcadores, e o coeficiente de endogamia foi significativo também para marcadores sugerindo perda de diversidade genética.

Palavras Chave: Ancestralidade genômica, Short Tandem Repeats (STRs), Marcadores Informativos de Ancestralidade (AIMs).

INTRODUÇÃO

A população brasileira apresenta alta diversidade genética resultado de mais de 500 anos de miscigenação entre os três principais grupos étnicos que a compõem: africanos, europeus e ameríndios. Quando os portugueses chegaram ao Brasil por volta de 1500, este era ocupado por aproximadamente 2,5 milhões de indígenas (RIBEIRO, 1995). A miscigenação começou logo após a chegada dos primeiros colonizadores portugueses e se estende até os dias de hoje. Por volta da metade do século XVI, chegaram ao território brasileiro, africanos para trabalhar como escravos nas fazendas de cana-de-açúcar, minas de ouro e diamantes e

também nas plantações de café (KLEIN, 2002). Posteriormente, quando os portos brasileiros foram abertos para nações amigas, houve também a entrada de mais imigrantes europeus, dentre eles, os mais numerosos foram oriundos da Itália, Espanha e Alemanha (IBGE, 2000).

A contribuição de cada grupo étnico na formação da nossa população foi diferente, e segundo Callegari-Jacques e Salzano (1999) dos imigrantes que chegaram ao Brasil entre 1500 e 1972, 58% eram europeus, 40% africanos e 2% asiáticos. Dados históricos e genéticos mostram que a distribuição desses três grupos étnicos ao longo do território brasileiro não ocorreu de forma homogênea, ou seja, a proporção de africanos, ameríndios e europeus difere significativamente a depender da região geográfica.

Abe-Sandes et al 2004, que estudaram a heterogeneidade do cromossomo Y, em seis grupos populacionais: quatro composto por afrodescendentes e um de descendentes de europeus e outro de descendentes de japoneses. Os resultados mostraram que a maioria dos cromossomos Y dos afrodescendentes são de origem africana (variando de 47,0% a 77,3%), e que a proporção de origem européia era maior que a ameríndia. Nos descendentes de europeus 98,0 % dos cromossomos Y são de origem européia e nos descendentes de japoneses não foi observado mistura com outros grupos étnicos, ou seja, todas as linhagens de cromossomo Y foram identificadas como asiáticas.

Com o objetivo de avaliar a ancestralidade genômica na população brasileira, Parra *et al.* (2003) analisaram 200 indivíduos do sexo masculino das regiões norte, nordeste, sudeste e sul do Brasil, que se autodenominaram como brancos os resultados deste estudo revelaram que no Brasil a cor da pele é um pobre indicador de ancestralidade genômica.

Alguns *loci* de marcadores de DNA têm demonstrado ser população-específico por apresentarem grandes diferenciais de frequência entre populações geográfica ou etnicamente definidas. Aqueles marcadores que apresentam diferenças nas suas frequências alélicas entre duas populações superiores a 30% foram denominados como “alelos específicos de população” (PSAs, do inglês *population specific alleles*) ou marcadores informativos de ancestralidade (AIMs) (SHRIVER *et al.*, 1997; PARRA *et al.*, 1998). Devido a esta característica estes marcadores são bastante eficazes para estimativa de mistura genética em populações miscigenadas.

Os marcadores do tipo STR (*Short Tandem Repeats*), também chamados de microssatélites ou repetições consecutivas curtas, são abundantes no genoma e são caracterizados como *loci* hipervariáveis, co-dominantes, multialélicos (FERREIRA *et al.*, 1995), têm taxa alta de mutação (KAYSER *et al.*, 2000) e heterozigosidade elevada (HAMMER *et al.*, 2002), sendo extremamente úteis no mapeamento genético de espécies, na

identificação de pessoas ou de linhagens/clones e predisposições de doenças (ARMOUR *et al.*, 1996); em estudos de identificação humana (HAMMOND *et al.*, 1994; HAMMER *et al.*, 2002); análise de paternidade (ALFORD *et al.*, 1994; SCHUMM *et al.*, 1997), mapeamento genético (SCHUMM *et al.*, 1995; EDWARDS *et al.*, 1992) e análise de populações (DEKA *et al.*, 1995; MARTINEZ-JARRETA *et al.*, 1999; YAMAMOTO *et al.*, 1999; GUSMÃO *et al.*, 2001).

A adoção dos nomes de conotação religiosa é uma prática antiga na Bahia, o que pode ser constatado através de investigação de documentos do período da escravidão (AZEVEDO *et al.*, 1982). Tais documentos revelaram que a maioria dos escravos permanecia sem sobrenome após adquirirem a liberdade. Estudos sugerem que a adoção de sobrenomes cresceu em resposta à demanda social e não por decisão voluntária súbita (AZEVEDO *et al.*, 1983). Estes mesmos autores demonstraram que a frequência de sobrenomes de conotação religiosa aumenta com a proximidade do “fenótipo negro”.

A presença de sobrenome de conotação religiosa é indício de ancestralidade africana e pode ser corroborada por parâmetros biológicos como o sistema sanguíneo ABO (JUNQUEIRA & WISHART, 1958) e o comprimento do cromossomo Y (BARBOSA *et al.*, 1997).

O censo de 2000 realizado pelo IBGE mostrou que a população da Bahia e da cidade de Salvador apresentam contingente considerável de afrodescendentes (77,5% e 79,8%, respectivamente).

Dados da literatura apontam a existência de risco diferencial de desenvolvimento de algumas doenças, a depender do grupo étnico ou região geográfica, podendo ocorrer associação com algumas doenças como, por exemplo, doenças cardiovasculares, tromboembólicas, câncer de próstata, hiperhomocisteína, anemia falciforme, entre outras (PENA, 2005).

No presente estudo, foi estimada a proporção de ancestralidade africana, européia e ameríndia em 203 indivíduos do estado da Bahia com marcadores AIMS e STRs, observando se os STRs conseguem identificar as proporções de mistura assim como os AIMS; e ainda se os indivíduos com sobrenome de conotação religiosa diferem na sua ancestralidade quando comparados os marcadores.

MATERIAL E MÉTODOS

População de estudo - foram analisadas 203 indivíduos provenientes de diversas regiões da Bahia, que realizam exames periódicos no Laboratório de Retrovírus do Hospital Universitário Professor Edgard Santos (HUPES), da Universidade Federal da Bahia (UFBA). Todos os indivíduos assinaram o termo de consentimento livre e esclarecido (TCLE). Este trabalho foi aprovado pelo Comitê de Ética em Pesquisas do Centro de Pesquisas Gonçalo Moniz/ FIOCRUZ (Parecer nº. 84/2006 – CEP/ CPqGM/ FIOCRUZ).

Extração do DNA genômico - Foram coletados 6mL de sangue em tubo *vacuntainer* e o DNA foi extraído a partir desse material biológico pela técnica de extração salina (LAHIRI & NURNBERGER, 1991).

Marcadores analisados – Foram analisados 9 marcadores de ancestralidade, sendo um polimorfismo de inserção/deleção - *AT3-I/D*, três inserções *Alu* - *SB19.3*, *APO* e *PV92*, 6 SNP – *FY-Null*, *CKMM*, *LPL*, *GC-1F*, *GC-1S* e *CYP3A4* e 8 marcadores microssatélites (TH01, vWA31, D18S51, FGA, TPOX, D7S820, D3S1358, D8S1179).

Os diferentes alelos encontrados nos marcadores de ancestralidade são denominados de alelo*1 e alelo*2. Nos polimorfismos do tipo *in/del* e nas inserções *alu*, o alelo *1 é caracterizado pela presença da inserção. Nos SNP, o alelo*1 é aquele cujo nucleotídeo abole o sítio de restrição (SHIVER *et al.*, 2003).

Análise laboratorial - O polimorfismo de inserção/deleção, as inserções *Alu* (*AT3-I/D* e *Sb19.3*, *APO* e *PV92*, respectivamente) foram genotipados por PCR utilizando os *primers* e condições da PCR descritas na literatura (SHRIVER *et al.*, 1997; PARRA *et al.*, 1998). Os SNP (*FY-Null*, *LPL*, *CKMM*, *CYP3A4* e *GC*) foram analisados pela técnica de discriminação alélica na PCR em tempo real, utilizando kits pré-sintetizados, sistema Taqman™, da Applied Biosystems. Oito marcadores STRs foram analisados e genotipados utilizando o método da PCR convencional e seus alelos discriminados por eletroforese capilar em seqüenciador automático ABI 310 Applied Biosystems.

Análises estatísticas – Utilizou-se o programa GENEPop para calcular as frequências alélicas, testar o equilíbrio de Hardy-Weinberg, avaliar a diferenciação populacional, déficit de heterozigotos e associação-par-a-par.

As estimativas de mistura genética foram geradas a partir do programa ADMIX, utilizando frequências de populações ancestrais ameríndias, européias e africanas disponíveis na literatura (CHAKRABORTY *et al.*, 1991).

A comparação das estimativas de mistura genética avaliadas por STRs e AIMs, foi realizada utilizando estatística Kappa. Este cálculo foi realizado no software R Projects 2.7.2 para Windows.

RESULTADOS E DISCUSSÃO

Este estudo utilizou a amostra de 203 indivíduos, sendo 86 mulheres e 117 homens. A caracterização fenotípica dos participantes classificou 17,7 % (36) como brancos, 73,9 % (150) como Mulatos, 5,9 % (12) como negros, 0,5 % como outros (1) e 2 % (4) indivíduos ficaram sem classificação. A maior quantidade de mulatos (73,9 %) reforça a informação de que nossa população é bastante miscigenada.

As frequências detectadas para o alelo *1 dos AIMs estão sumarizadas na tabela 1. A aderência ao equilíbrio de Hardy-Weinberg foi calculada no software GENEPOP com base nas frequências alélicas encontradas nos *loci* analisados.

Para os AIMs, Sb 19.3, PV92 e CYP3A4 não aderiram ao equilíbrio de Hardy-Weinberg. Já nos STRs, apenas o *locus* D3S1358 não está em equilíbrio. As inserções *Alu* apresentaram excesso de homozigotos tanto para a presença e para a ausência da inserção. O SNP CYP3A4 também apresentou excesso de homozigotos na amostragem. Esses resultados podem ter sido gerados devido à coleta não aleatória das amostras. Entretanto o excesso de homozigotos pode ser devido à estruturação populacional ocasionado por casamento preferencial entre indivíduos do mesmo grupo étnico aumentando dessa forma os genótipos homozigotos na população (AZEVEDO *et al*, 1986). Em relação ao D3S1358, a explicação poderia ser erro de genotipagem, visto que este *locus* já vinha apresentando problemas para o alelo 15 quando comparado com kit comercial.

TABELA 1. Frequência do alelo*1 para os marcadores de ancestralidade FY-Null, LPL, AT3-I/D, Sb19.3, APO, PV92, CYP3A4, CKMM, GC-1F-1S numa amostra da população da Bahia – Brasil.

<i>Locus/alelo</i>	Frequência do Alelo *1
<i>AT3-I/D</i>	0,517
<i>APO</i>	0,746
<i>Sb 19.3</i>	0,667
<i>PV92</i>	0,273
<i>FY-Null</i>	0,569

<i>CKMM</i>	0,298
<i>LPL</i>	0,692
<i>CYP3A4</i>	0,468
<i>GC-1F</i>	0,473
<i>GC-1S</i>	0,350

TABELA 2. Frequências alélicas dos 8 STRs numa amostra da população da Bahia – Brasil.

Alelos	D8S1179	D7S820	D3S1358	TH01	VWA31	TPOX	D18S51	FGA
6	-	0.0025	-	0.1847	-	0.0345	-	-
7	-	0.0049	-	0.2463	-	0.0049	-	-
8	0.0025	0.1970	-	0.1281	-	0.4212	-	-
9	-	0.1207	-	0.1946	-	0.1502	-	-
9.3	-	-	-	0.2463	-	-	-	-
10	0.0690	0.2685	-	-	-	0.0764	0.0074	-
11	0.0714	0.1970	-	-	0.0123	0.2734	0.0123	-
12	0.1305	0.1626	0.0049	-	0.0025	0.0394	0.0985	0.0025
13	0.2389	0.0345	0.0074	-	0.0025	-	0.0616	-
14	0.2956	0.0123	0.0862	-	0.0640	-	0.1281	-
14.2	-	-	-	-	-	-	0.0074	-
15	0.1355	-	0.3498	-	0.1675	-	0.1650	-
16	0.0542	-	0.2562	-	0.2857	-	0.1404	-
17	0.0025	-	0.1847	-	0.2217	-	0.1700	0.0099
18	-	-	0.0985	-	0.1453	-	0.0985	0.0025
18.2	-	-	-	-	-	-	0.0025	0.0099
19	-	-	0.0123	-	0.0788	-	0.0493	0.0837
20	-	-	-	-	0.0148	-	0.0369	0.0985
21	-	-	-	-	0.0049	-	0.0123	0.1429
22	-	-	-	-	-	-	0.0074	0.1552
23	-	-	-	-	-	-	0.0025	0.1626
24	-	-	-	-	-	-	-	0.1773
25	-	-	-	-	-	-	-	0.0788
26	-	-	-	-	-	-	-	0.0591
27	-	-	-	-	-	-	-	0.0074
28	-	-	-	-	-	-	-	0.0049
29	-	-	-	-	-	-	-	0.0025
31.2	-	-	-	-	-	-	-	0.0025
n	406	406	406	406	406	406	406	406
HomExp	38.7012	38.5901	48.2247	42.1778	38.2988	57.1531	23.8593	25.7679
HomObs	41	39	39	36	42	56	17	41
HetExp	164.2988	164.4099	154.7753	160.8222	164.7012	145.8469	179.1407	177.2321
HetObs	162	164	164	167	161	147	186	162
<i>Locí</i>	D8S1179	D7S820	D3S1358	TH01	VWA31	TPOX	D18S51	FGA

n = número de genes analisados; HomExp = Homozigotos esperados; HomObs = Homozigotos observados; HetExp = Heterozigotos esperados; HetObs = Heterozigotos observados. Alelos mais frequentes encontrados estão em negrito na tabela acima.

A tabela 3 apresenta os alelos mais frequentes da amostra testada comparado com os alelos mais frequentes nas populações parentais.

TABELA 3. Alelos mais frequentes nos STRs da amostra e nas populações parentais.

STR	Alelos	Amostra ^a	Europeus ^b	Africanos ^c	Ameríndios ^d
D3S1358	15	0,3500	-	-	0,5908
	16	-	0,2526	0,3355	-
D7S820	10	0,2680	0,2711	0,3569	0,4106
D8S1179	13	-	0,2992	-	-
	14	0,2960	-	0,3373	0,3538
TH01	6	-	-	-	0,3995
	7	0,2460	-	0,3733	-
	9.3	0,2460	0,2769	-	-
TPOX	8	0,4210	0,5008	-	-
	11	-	-	0,3039	0,3970
VWA	16	0,2860	-	0,2652	0,4642
	17	-	0,2637	-	-
D18S51	14	-	0,1702	-	0,2938
	16	-	-	0,1766	-
	17	0,1700	-	-	-
FGA	21	-	0,1826	-	-
	24	0,1770	-	0,1887	-
	25	-	-	-	0,3180

^a = presente estudo; ^b = Vieira-Silva *et al.*, 2006, Camacho *et al.*, 2007, Vecchio *et al.*, 2004; ^c = Alves *et al.*, 2004, Beleza *et al.*, 2004; ^d = Santos SEB *et al.*, 2009, Crossetti SG *et al.*, 2008.

A tabela 4 mostra a frequência do alelo *1 dos marcadores estudados nas populações parentais e no presente estudo. É possível observar que essas frequências na Bahia são intermediárias quando comparadas às populações parentais, confirmando a miscigenação desta população.

TABELA 4. Frequência do alelo *1 para AIMs nas populações ancestrais e numa amostra da população da Bahia – Brasil.

Alelos/ Populações	Africanos ^a	Europeus ^a	Nativo Americano ^a	Bahia ^b
<i>AT3-1/D</i>	0,858	0,282	0,061	0,517
<i>APO</i>	0,420	0,925	0,977	0,746
<i>Sb 19.2</i>	0,415	0,903	0,645	0,667
<i>PV 92</i>	0,225	0,152	0,792	0,273
<i>FY-Null</i>	0,001	0,998	1,000	0,569

<i>CKMM</i>	0,164	0,313	0,904	0,298
<i>LPL</i>	0,971	0,492	0,442	0,692
<i>CYP3A4</i>	0,198	0,958	0,959	0,468
<i>GC-1F</i>	0,853	0,156	0,339	0,473
<i>GC-IS</i>	0,069	0,607	0,542	0,350

^a Frequências segundo Shriver *et al.*, 2003; ^b = presente estudo

As frequências alélicas dos STRs e dos AIMs foram utilizadas para realizar a análise de mistura genética onde as frequências desses marcadores observadas na amostra da população da Bahia foi comparada com os dados das populações ancestrais europeia, africana e ameríndia.

A tabela 5 mostra estas informações para os STRs e para os AIMs considerando a população como não subestruturada.

TABELA 5. Proporções de contribuição africana, europeia e ameríndia numa amostra da população da Bahia-BA para os marcadores STRs e AIMs.

Contribuição	STRs#		AIMs#		AIMs *	
	m	s.e	m	s.e	m	s.e
Pop.						
Europeia	0,5858	0,0068	0,4516	0,0208	0,3554	0,0171
Africana	0,3348	0,0057	0,4508	0,0062	0,4779	0,0039
Ameríndia	0,0794	0,0024	0,0975	0,0185	0,1667	0,0154

Pop. – população parental; m – índice de mistura; s.e. – erro padrão; # presente estudo; * Bomfim, 2008.

Os marcadores informativos de ancestralidade demonstraram que a amostra utilizada da população da Bahia apresentou praticamente a mesma proporção (45 %) de contribuição europeia e africana e aproximadamente 10 % de contribuição ameríndia. A mesma análise realizada com os STRs mostra resultado diferente, apresentando maior contribuição europeia (quase 60 %), seguindo de 33 % de contribuição africana e 8 % de ameríndia.

Quando se comparou os resultados obtidos com os AIMs nesta amostra com os encontrados por Bomfim, 2008, observamos também diferenças nas contribuições das populações ancestrais, esta diferença deve ter ocorrido devido ao tamanho amostral, visto que os 203 indivíduos analisados aqui são parte da amostra de Bomfim.

Segundo dados históricos e dados de AIMs (MACHADO, 2008 e BOMFIM, 2008), era esperado que a contribuição africana, nesta amostra, fosse maior que a europeia. Esses

resultados discrepantes podem ser explicados pelo tamanho amostral e ainda no caso dos STRs pelo grande número de alelos por *locus* que leva a diluição da frequência e também pelo pequeno diferencial de frequência de vários alelos nas populações ancestrais.

A análise de concordância entre a estimativa de mistura étnica baseada nos dois tipos de marcadores, utilizando o método estatístico *Kappa*, mostrou índice de concordância global de 44% e índice *Kappa* de 0,12, o que significa concordância pequena entre os dois métodos (LANDIS & KOCK, 1977).

Vários trabalhos, como o de Junqueira & Wishart em 1958, Azevedo *et al.*, 1982, Tavares-Neto & Azevedo, 1977 mostraram que a presença de sobrenome com conotação religiosa é indício de ancestralidade africana e pôde ser corroborada por parâmetros biológicos como o sistema sanguíneo ABO (JUNQUEIRA & WISHART, 1958) e também através do comprimento do cromossomo Y (BARBOSA *et al.* em 1997).

Nossa amostra contém 203 indivíduos, destes, 56% (113 indivíduos) tinham sobrenome com conotação religiosa e o restante, 44% (90 indivíduos) não tinham esta característica no sobrenome. Com base nestes dados, dividimos a amostra em dois grupos, um com sobrenome de conotação religiosa e outro sem. Feito isto, realizamos a análise de diferenciação populacional com objetivo de verificar se haviam diferenças significativas entre elas utilizando os dois tipos de marcadores (AIMs e STRs, Tabelas 8 e 9) e a análise de mistura étnica para os AIMs e para os STRs nos dois grupos.

Não foram observadas diferenças entre os dois grupos para os marcadores STRs, porém diferenças estaticamente significantes foram observadas para os marcadores AIMs (AT3-I/D, FY-Null, CYP3A4 e GC). Isto mostra que os AIMs são mais sensíveis em detectar diferenças genéticas entre as populações.

A tabela 6 mostra a análise realizada com os AIMs. O resultado mostra que os indivíduos com sobrenome de conotação religiosa apresentam maior ancestralidade africana (53,32%). Já nos sem esta característica no sobrenome, a proporção de mistura é maior para europeus (57,38%), confirmando assim os achados de Tavares-Neto & Azevedo em 1977 em relação a esta característica. Estes resultados são estatisticamente significantes ($p = 0.023$).

TABELA 6. Proporções contribuição africana, européia e ameríndia utilizando AIMs na amostra em estudo quando separada pela presença ou ausência de sobrenome de conotação religiosa.

Contribuição	AIMs C/Rel ^a		AIMs S/Rel ^b	
	m	s.e	m	s.e
Pop.				

Européia	0,3387	0,0209	0,5738	0,0184
Africana	0,5332	0,0082	0,3604	0,0055
Ameríndia	0,1280	0,0164	0,0658	0,0174

Pop. – população parental; m – índice de mistura; s.e. – erro padrão; ^a= indivíduos com sobrenome de conotação religiosa; ^b indivíduos sem sobrenome de conotação religiosa.

TABELA 7. Proporções de contribuição africana, européia e ameríndia utilizando marcadores STRs na amostra em estudo quando separada em populações, com e sem sobrenome de conotação religiosa.

Contribuição	STRs C/Rel ^a		STRs S/Rel ^b	
	m	s.e	m	s.e
Pop.				
Européia	0,4674	0,0010	0,7226	0,0200
Africana	0,4326	0,0009	0,2198	0,0174
Ameríndia	0,1000	0,0004	0,0576	0,0059

Pop. – população parental; m – índice de mistura; s.e. – erro padrão; ^a= indivíduos com sobrenome de conotação religiosa; ^b indivíduos sem sobrenome de conotação religiosa.

A tabela 7 mostra a mesma análise da tabela anterior, porém com outros marcadores, os STRs. Neste caso, o grupo com sobrenome de conotação religiosa apresentou maior ancestralidade européia; e os indivíduos que não tinha o sobrenome de conotação religiosa, como esperado, apresentaram em sua maioria ancestralidade européia (72,3%). Estes resultados mostram que, ou os STRs utilizados neste estudo não são bons marcadores para identificar ancestralidade, ou talvez o tamanho amostral tenha sido insuficiente para detectar a mistura observada com o uso dos AIMs.

Com relação à variabilidade genética os resultados observados para os AIMs, nesta amostra, apresentou déficit de heterozigotos em quatro marcadores analisados, e o coeficiente de endogamia foi significativo para estes quatro marcadores (APO, $F_{IS} = 0,182$, Sb19.3, $F_{IS} = 0,303$, PV92, $F_{IS} = 0,221$ e CYP3A4, $F_{IS} = 0,645$; todos $p < 0,05$), sugerindo perda de diversidade genética, possivelmente por estruturação populacional, originada por casamentos preferenciais entre indivíduos do mesmo grupo racial (AZEVEDO et. al., 1986).

A análise de associações alélicas par-a-par, utilizando os marcadores STRs, não mostrou nenhuma associação significativa. Com os AIMs, 36 combinações possíveis foram feitas sendo observado 9 associações significantes ($p < 0,05$). Estes resultados indicam estruturação populacional, pois o número de associações esperadas ao acaso para 36 combinações seria 3,24. Este achado é concordante também com os dados do F_{IS} . Resultados

semelhantes já foram observados numa amostra de Salvador e da Bahia (MACHADO, 2008; BOMFIM, 2008).

A comparação da estimativa de mistura genética que realizamos com os STRs e AIMs apresentou resultados divergentes, o que nos leva a crer que os Microsatélites utilizados nesta análise não são bons marcadores para ancestralidade quando comparados com os AIMs.

REFERÊNCIAS BIBLIOGRÁFICAS

RIBEIRO D. *Povo Brasileiro: Formação e sentido do Brasil*. São Paulo, Companhia das Letras, 1995.

KLEIN HS. *As origens africanas dos escravos brasileiros*. In Pena SDJ, ed. *Homo Brasilis*, Ribeirão Preto, FUNPECRP, 93-112, 2002.

IBGE - Instituto Brasileiro de Geografia Estatística - Brasil: 500 Anos de Povoamento. IBGE, Rio de Janeiro, 2000.

CALLEGARI-JACQUES SM, Salzano FM. *Brazilian Indian/non-Indian interactions and their effects*. *Cienc Cult*, 51:166-174, 1999.

ABE-SANDES K, Silva WA Jr, Zago MA. *Heterogeneity of the Y chromosome in Afro-Brazilian populations*. *Hum Biol*, 76(1):77-86, 2004.

PARRA FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. *Color and genomic ancestry in Brazilians*. *Proc Natl Acad Sci U S A*, 100(1):177-82, 2003.

SHRIVER MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE. *Ethnic-affiliation estimation by use of population-specific DNA markers*. *Am J Hum Genet*, 60(4):957-64, 1997.

PARRA EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, Forrester T, Allison DB, Deka R, Ferrell RE, Shriver MD. *Estimating African American ture proportions by use of population-specific alleles*. *Am J Hum Genet*, 63(6):1839-51, 1998.

FERREIRA ME, Grattapaglia D. *Introdução ao uso de marcadores RAPD e RFLP em análise genética*. Lumma consultoria projetos e informática, Brasília-DF, 220pp, 1995.

KAYSER M, Roewer L, Hedman M, Henke L, Henke L, Henke J. *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*, 66:1580-1588, 2000.

HAMMER MF, Zegura S. *The human Y chromosome haplogroup tree: nomenclature and phylogeography of its major divisions*. *Annu Rev Anthropol*, 31:303-321, 2002.

ARMOUR JAL, Anttinen T, May CA, Veja EE, Sajantilla A, Kidd JR, Kidd KK, Bertranpetit J, Paabo S, Jeffreys AJ. *Minisatellite diversity supports a recent African origin for modern humans*. *Nat Genet*, 13:154-160, 1996.

- HAMMOND HA, Jin L, Zhong Y, Caskey CT, Chackraborty R. *Evaluation of 13 short tandem repeat loci for use in personal identification applications*. Am J Hum Genet, 55:175-189, 1994.
- ALFORD RL, Hammond HA, Cotel, Caskey CT. *Rapid and efficient resolution of parentage by amplification of short tandem repeats*. Am J Hum Genet, 55:190-195, 1994.
- SCHUMM JW, Lins AM, Micka KA, Sprecher CJ, Rabbach DR, Bacher JW. *Automated fluorescent detection of STR multiplexes – development of the Geneprint™ PowerPlex™ and FFL multiplexes for forensic and paternity applications*. First Europ Sympos Hum Identif, 90-104, 1997.
- SCHUMM JW, Lins AM, Puers C, Sprecher CJ. *Development of nonisotopic multiplex amplification sets for analysis of polymorphic STR loci*. First Europ Sympos Hum Identif, 177-181, 1995.
- EDWARDS A, Hammond HA, Jin L, Caskey CT, Chakraborty R. *Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups*. Genomics, 12:241-253, 1992.
- MARTINEZ-JARRETA BM, Roche PD, Abecia E. *Genetic variation at six STR loci (HUMTH01, HUMTPOX, HUMCSF1PO, HUMF13A01, HUMFES/FPS, HUMVWFA31) in Aragon (north Spain)*. Forensic Sci Intern, 100:87-92, 1999.
- YAMAMOTO T, Uchihi R, Nozawa H, Huang XL, Leong YK, Tanaka M, Mizutani M, Tamaki K, Katsumata Y. *Allele distribution at nine STR loci – D3S1358, VWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820 – in the Japanese population by multiplex PCR and capillary electroph*. J Forensic Sci, 44(1):167-170, 1999.
- GUSMÃO L, Prata MJ, Miranda C, Trovoada MJ, Amorim A. *STR data from S. Tomé e Príncipe (Gulf of Guinea, West Africa)*. Forensic Sci Intern, 116:53-54, 2001.
- AZEVEDO ES, Fortuna CMM, Silva KMC, Sousa MGF, Machado MA, Lima AMVMD, Aguiar ME, Abé-Sandes K, Eulálio MCMN, Conceição MM, Silva MCBO, Santos MG. *Spread and diversity of human populations in Bahia, Brazil*. Hum Biol, 54:329-341, 1982.
- AZEVEDO ES. *The reconstruction of cultural history and racial ture from the meaning of family names in Bahia, Brazil*. Quaderni di Semantica, 6(1):209-212, 1983.

JUNQUEIRA PC, Wishart PJ. *Distribuição dos grupos sanguíneos abo em brancos, mulatos e pretos do rio de janeiro, de acordo com a presença ou ausência de sobrenome*. Ver Clin, São Paulo 34:79-83, 1958.

BARBOSA AAL, Cavalli IJ, Abé-Sandes K, Santos MG, Azevedo ES. *Family names and the length of the Y chromosome in Brazilian blacks*. Braz J Genet, vol. 20 n. 1, Ribeirão Preto, Mar. 1997.

PENA SDJ. *Razões para banir o conceito de raça da medicina brasileira*. História, Ciências, Saúde – Manguinhos. 12(1): 321-346, 2005.

LARIHI DK, NURNBERGER JI Jr. *A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies*. Nucleic Acids Res, 19(19):5444, 1991.

RAYMOND, M. & Rousset, F. *GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism*. J Hered, 86:248-249, 1995.

CHAKRABORTY R, Kamboh MI & Ferrell RE. *“Unique” alleles in ed populations: a strategy for determining hereditary population differences of disease frequencies*. Ethn Dis, 1:245-256, 1991.

AZEVEDO ES, Chautard-Freire-Maia EA, Freire-Maia N, Fortuna CMM, Abe K, Santos MG, Barbosa AAL, Silva MET, Costa AF. *Mating types in a mixed and multicultural population of Salvador, Brazil*. Rev Brasil Genet, IX (3):487-496, 1986.

MACHADO, TMB. Bomfim, T.F , Acosta, AX, Galvão-Castro, B, Abe-Sandes, K. *Associação entre sobrenomes de conotação religiosa e ancestralidade genômica em Salvador-BA-Brasil*. FIOCRUZ - BA – Fundação Oswaldo Cruz, 2008.

MACHADO, TMB, Bomfim,TF , Acosta, AX, Galvão-Castro, B , Sousa, SMB, Abe-Sandes, K . *Estimativa de mistura genética em uma amostra da população de Salvador-BA-Brasil*. FIOCRUZ - BA – Fundação Oswaldo Cruz, 2008.

BOMFIM, TF. *Ancestralidade genômica em uma amostra de portadores do HIV-1 do Estado da Bahia*. Dissertação (Mestrado) - Curso de Pós-graduação em Biotecnologia em Saúde e Medicina Investigativa – Centro de Pesquisa Gonçalo Moniz, FIOCRUZ - BA – Fundação Oswaldo Cruz, 2008.

LANDIS, J.; KOCH, G. *The measurement of observer agreement for categorical data*. Biometrics, Washington, v.33, p.159-74, 1977.

TAVARES-NETO J, Azevêdo ES. *Racial origin and historical aspects family names in Bahia, Brazil*. Hum Biol, 49:287-289, 1977.

6 CONCLUSÕES

A estimativa de mistura genética avaliada por STRs verificada no nosso estudo foi de 58,58 % para contribuição europeia, 33,48 % para africana e 7,94 % de ameríndia. A mesma estimativa com os AIMs apresentou 45,16 % de contribuição europeia, 45,08 % de africana e 9,75 % de contribuição ameríndia. Estes dados mostram resultados diferentes entre os STRs e AIMs, o que nos leva a crer que os Microssatélites utilizados nesta análise não são bons marcadores para ancestralidade.

A concordância entre as estimativas de ancestralidade utilizando os dois tipos de marcadores (AIMs e STRs) foi muito baixa ($\kappa = 0,12$).

Comparando-se as frequências dos polimorfismos observados em nosso estudo com as frequências parentais, podemos perceber para os AIMs que os valores são intermediários, o que reflete o processo de miscigenação que ocorre até hoje com nossa população.

A observação de indivíduos com sobrenome de conotação religiosa em nossa amostra pode ser associado com ancestralidade africana como foi demonstrado com os AIMs, onde os indivíduos com esta característica apresentaram 53,32 % de contribuição africana em seu perfil genético.

A análise de variabilidade genética para os AIMs mostrou déficit de heterozigotos em quatro marcadores analisados, e o coeficiente de endogamia foi significativo para quatro marcadores sugerindo perda de diversidade genética

7 REFERÊNCIAS BIBLIOGRÁFICAS

- ABE-SANDES K. *Diversidade genética de afro-brasileiros: DNA mitocondrial e cromossomo Y*. 104 p. Tese (Doutorado – Programa de Pós-Graduação em Genética.) – Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo. 2002.
- ABE-SANDES K, Silva WA Jr, Zago MA. *Heterogeneity of the Y chromosome in AfroBrazilian populations*. Hum Biol, 76(1):77-86, 2004.
- ALFORD RL, Hammond HA, Cotel, Caskey CT. *Rapid and efficient resolution of parentage by amplification of short tandem repeats*. Am J Hum Genet, 55:190-195, 1994.
- ALVES C, Fortuna CMM, Toralles MB. *A aplicação e o conceito de raça em saúde pública: Definições, controvérsias e sugestões para uniformizar sua utilização em pesquisas biomédicas e na prática clínica*. Gazeta Médica da Bahia, 75(1):92-115, 2005.
- ALVES C, Gusmão L, Amorim A. STR data (ampflstr profiler Plus and GenePrint CTTv) from Mozambique. Forensic Sci Int, 119:131-133, 2001.
- ALVES C, Gusmão L, Damasceno A, Soares B and Amorim A. *Contribution for an African autosomic STR database (AmpF/STR Identifiler and Powerplex 16 System) and a report on genotypic variations*. Forensic Sci Int, 139:201-205, 2004.
- ALVES-SILVA J, Santos MS, Guimarães PEM, Ferreira ACS, Bandelt H, Pena SDJ, Prado VF. *The ancestry of Brazilian mtDNA lineages*. Am J Hum Genet, 67:444-461, 2000.
- ANJOS MJ, Carvalho M, Andrade L, Corte-Real F, Vieira DN, Vide MC. Allele frequencies of STR multiplex systems in two Portuguese population samples. Progr Forensic Genet, 8:208-211, 2000.
- ARCOT SS, Adamson AW, Risch GW, Lafleur J, Robichaux MB, Lamerdin JE, Carrano AV, Batzer MA. *High-resolution cartography of recently integrated human chromosome 19-specific Alu fossils*. J Mol Biol, 281(5):843-56, 1998.
- ARMOUR JAL, Anttinen T, May CA, Veja EE, Sajantilla A, Kidd JR, Kidd KK, Bertranpetit J, Paabo S, Jeffreys AJ. *Minisatellite diversity supports a recent African origin for modern humans*. Nat Genet, 13:154-160, 1996.
- AZEVEDO ES, Fortuna CMM, Silva KMC, Sousa MGF, Machado MA, Lima AMVMD, Aguiar ME, Abé-Sandes K, Eulálio MCMN, Conceição MM, Silva MCBO, Santos MG. *Spread and diversity of human populations in Bahia, Brazil*. Hum Biol, 54:329-341, 1982.

- AZEVEDO ES, Silva MCOB, Tavares-Neto J. *Human Alcohol dehydrogenase ADH1, ADH2 and ADH3 loci in a mixed population of Bahia, Brazil.* Ann Hum Genet, 39:321, 1975.
- AZEVEDO ES. *The anthropological and cultural meaning of family names in Bahia, Brazil.* Curr Anthropol, 21:360-363, 1980.
- AZEVEDO ES. *The reconstruction of cultural history and racial ture from the meaning of family names in Bahia, Brazil.* Quaderni di Semantica, 6(1):209-212, 1983.
- AZEVEDO ES, Chautard-Freire-Maia EA, Freire-Maia N, Fortuna CMM, Abe K, Santos MG, Barbosa AAL, Silva MET, Costa AF. *Mating types in a mixed and multicultural population of Salvador, Brazil.* Rev Brasil Genet, IX (3):487-496, 1986.
- BANDELT H, Alves-Silva J, Guimarães P, Santos M, Brehm A, Pereira L, Coppa A, Larruga J, Rengo C, Scozzari R, Torroni A, Prata M, Amorim A, Prado V, Pena SD. *Phylogeography of the human mitochondrial L3e: a snapshot of African prehistory and Atlantic slave trade.* Ann Hum Genet, 65:549-563, 2001.
- BARBOSA AAL, Cavalli IJ, Abé-Sandes K, Santos MG, Azevedo ES. *Family names and the length of the Y chromosome in Brazilian blacks.* Braz J Genet, vol. 20 n. 1, Ribeirão Preto, Mar. 1997.
- BARBOSA AAL, Sousa SMB, Abé-Sandes K, Alonso CA, Schneider V, Costa DCC, Cavalli IJ, Azevêdo EES. *Microsatellite studies on an isolated population of African descent in the Brazilian state of Bahia.* Genetics and Molecular Biology, 29(1):23-30, 2006.
- BATZER MA, Kilroy GE, Richard PE, Shaikh TH, Desselle TD, Hoppens CL, Deininger PL. *Structure and variability of recently inserted Alu family members.* Nucleic Acids Res, 18(23):6793-8, 1990.
- BATZER MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ. *African origin of human-specific polymorphic Alu insertions.* Proc Natl Acad Sci U S A, 91(25):12288-92, 1994.
- BEAN DF, Bradshaw BS. *Intermarriage between persons of Spanish and non-Spanish surname: Changes from the mid-nineteenth to the mid-twentieth century.* Soc Sci Q, 51: 389-395, 1970.
- BELEZA S, Alves C, Reis F, Amorim A, Carracedo A and Gusmão L. *17 STR data (AmpF/STR Identifiler and Powerplex 16 System) from Cabinda (Angola).* Forensic Sci Int, 141:193-196, 2004.

- BEUTLER E. Glucose-6-phosphate dehydrogenase deficiency. *New Engl Med* 324:169-74, 1991.
- BOMFIM, TF. *Ancestralidade genômica em uma amostra de portadores do HIV-1 do Estado da Bahia*. Dissertação (Mestrado) - Curso de Pós-graduação em Biotecnologia em Saúde e Medicina Investigativa – Centro de Pesquisa Gonçalo Moniz, FIOCRUZ - BA – Fundação Oswaldo Cruz, 2008.
- BORTOLINI MC, Zago MA, Salzano FM, Silva Júnior WA, Bonatto SL, Silva MCBO, Weiner TA. *Evolutionary and anthropological of mitochondrial DNA variation in African Brazilian populations*. *Human Biology*, 69:141:159, 1997.
- BOWCOCK AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL. *High resolution of human evolutionary trees with polymorphic microsatellites*. *Nature*, 368(6470):455-7, 1994.
- BOSCH E, Clarimón J, Pérez-Lexaun A, Cafafell F. STR data for 21 loci in northwestern África. *Forensic Sci Int*, 116:41-51, 2001.
- BRASE CL. *Region does not mean "race"--reality versus convention in forensic anthropology*. *J Forensic Sci*, 40:171-175, 1995.
- CALLEGARI-JACQUES SM, Grattapaglia D, Salzano FM, Salamoni SP, Crossetti SG, Ferreira ME, Hutz MH. *Historical genetic: spatiotemporal analysis of the formation of the brazilian population*. *Am Jour Hum Biology*, 15:824-834, 2003
- CALLEGARI-JACQUES SM, Salzano FM. *Brazilian Indian/non-Indian interactions and their effects*. *Cienc Cult*, 51:166-174, 1999.
- CAMACHO MV, Benito C, Figueiras AM. *Allelic frequencies of the 15 STR loci included in the AmpFlSTR® Identifiler™ PCR Amplification Kit in an autochthonous sample from Spain*. *Forensic Sci Int*, 173:241-245, 2007.
- CARVALHO-SILVA DR, Santos FR, Rocha J, Pena SD: *The phylogeography of Brazilian Y-chromosome lineages*. *Am J Hum Genet*, 68:284-286, 2001.
- CHAKRABORTY R, Barton SA, Ferrel RE, *et al*. *Ethnicity determination by names among the Aymara of Chile and Bolivia*. *Hum Biol*, 61:159-177, 1989.
- CHAKRABORTY R, Kamboh MI & Ferrell RE. *"Unique" alleles in ed populations: a strategy for determining hereditary population differences of disease frequencies*. *Ethn Dis*, 1:245-256, 1991.

- CHAKRABORTY R, Kamboh MI, Nwankwo M, Ferrell RE. *Caucasian genes in American blacks: new data*. Am J Hum Genet, 50:145-155, 1992.
- COTRIM NH. *Variabilidade Molecular em Populações Brasileiras: Inserções de Alu em Populações Remanescentes de Quilombos*. 2003, Dissertação de Mestrado em Ciências Biológicas, Universidade de São Paulo, USP, Brasil.
- CROSSETTI SG, Dermarchi DA, Raimann PE, Salzano FM *et al*. Autosomal STR genetic variability in the Gran Chaco Native Population: Homogeneity or Heterogeneity? Am J Hum Genet, 20:704-711, 2008.
- DAWSON DM, Eppenberger HM, Eppenberger ME. *Multiple molecular forms of creatine kinases*. Ann N Y Acad Sci, 151: 616-626, 1968.
- DEKA R, Shriver MD, Yu LM, Ferrell RE, Chackraborty R. *Intra and inter-population diversity at short tandem repeat loci in diverse populations of the world*. Electroph, 16:1659-1664, 1995.
- DORNELLES CL, Callegari-Jacques SM, Robinson WM, Weimer TA, Franco MHL, Hickmann AC, Geiger CJ, *et al*. *Genetics, surnames, grandparents' nationalities, and ethnic ture in southern Brazil – Do the patterns of variation coincide?* Genet and Mol Biol 22:151-161, 1999.
- DOUNAHUE RP, Bias WB, Renwick JH, Mckusik VA. *Probable assignment of the Duffy blood group locus to chromosome 1 in man*. Proc Nat Acad Sci, 61: 949-955, 1968.
- DROBNIC K, Budowle B. *The analysis of three short tandem repeat (STR) loci in the Slovene population by multiplex PCR*. Jour For Sci, 45:893-895, 2000.
- EDWARDS A, Civitello A, Hammond HA, Caskey CT. *DNA typing and genetic mapping with trimeric and tetrameric tandem repeats*. Am J Hum Genet, 49:746-756, 1991.
- EDWARDS A, Hammond HA, Jin L, Caskey CT, Chakraborty R. *Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups*. Genomics, 12:241-253, 1992.
- FELSENSTEIN J. *Confidence limits on phylogenies: an approach using the bootstrap*. Evolution, 39:783-791, 1985.
- FERREIRA ME, Grattapaglia D. *Introdução ao uso de marcadores RAPD e RFLP em análise genética*. Lumma consultoria projetos e informática, Brasília-DF, 220pp, 1995.

- FUNKE H, Klug J, Assmann G. *Hind III RFLP in the lipoprotein lipase gene, (LPL)*. Nucleic Acids Res, 15(21):9102, 1987.
- GUSMÃO L, Prata MJ, Miranda C, Trovoada MJ, Amorim A. STR data from S. Tomé e Príncipe (Gulf of Guinea, West Africa). Forensic Sci Intern, 116:53-54, 2001.
- HADLEY TJ, David PH, McGinniss MH, Miller LH. *Identification of an erythrocyte component carrying the Duffy blood group Fy-a antigen*. Science, 223: 597-599, 1984.
- HAMMER MF, Zegura S. *The human Y chromosome haplogroup tree: nomenclature and phylogeography of its major divisions*. Annu Rev Anthropol, 31:303-321, 2002.
- HAMMOND HA, Jin L, Zhong Y, Caskey CT, Chackraborty R. *Evaluation of 13 short tandem repeat loci for use in personal identification applications*. Am J Hum Genet, 55:175-189, 1994.
- HE W, Neil S, Kulkarni H, Wright E, Agan BK, Marconi VC, Dolan MJ, Weiss RA, Ahuja SK. *Duffy Antigen Receptor for Chemokines Mediates trans-Infection of HIV-1 from Red Blood Cells to Target Cells and Affects HIV-AIDS Susceptibility*. Cell Host Microbe, 4(1): 52-62, 2008.
- HOUCK CM, Rinehart FP, Schmid CW. *A ubiquitous family of repeated DNA sequences in the human genome*. J Mol Biol, 15;132(3):289-306, 1979.
- IBGE - Instituto Brasileiro de Geografia Estatística - Brasil: 500 Anos de Povoamento. IBGE, Rio de Janeiro, 2000.
- JUNQUEIRA PC, Wishart PJ. *Distribuição dos grupos sanguíneos abo em brancos, mulatos e pretos do rio de janeiro, de acordo com a presença ou ausência de sobrenome*. Ver Clin, São Paulo 34:79-83, 1958.
- JURKA J, Smith T. *A fundamental division in the Alu family of repeated sequences*. Proc Natl Acad Sci U S A, 85(13):4775-8, 1988.
- KARATHANASIS SK, Zannis VI, Breslow JL. *Characterization of the apolipoprotein A-I-C-III gene complex*. Methods Enzymol, 128:712-26, 1986.
- KAYSER M, Roewer L, Hedman M, Henke L, Henke L, Henke J. *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, 66:1580-1588, 2000.
- KITTLES RA, Chen W, Panguluri RK, Ahaghotu C, Jackson A, Adebamowo CA, Griffin R, Williams T, Ukoli F, Adams-Campbell L, Kwagyan J, Isaacs S, Freeman V, Dunston GM.

- CYP3A4V and prostate cancer in African Americans: causal or confounding association because of population stratification?* Hum Genet, 110:553-560, 2002.
- KLEIN HS. *As origens africanas dos escravos brasileiros*. In Pena SDJ, ed. Homo Brasilis, Ribeirão Preto, FUNPECRP, 93-112, 2002.
- LARIHI DK, NURNBERGER JI Jr. *A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies*. Nucleic Acids Res, 19(19):5444, 1991.
- LANDIS, J.; KOCH, G. The measurement of observer agreement for categorical data. *Biometrics*, Washington, v.33, p.159-74, 1977.
- LEHMANN JM, McKee DD, Watson MA, Willson TM, Moore JT, Kliewer SA. *The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions*. J Clin Invest, 102(5):1016-23, 1998.
- LEWIS PO & Zaykin D. *Genetic data analysis: software for the analysis of discrete genetic data*. Version 1.0, 1997.
- LI S, Oka K, Galton D, Stocks J. *Bst-1 RFLP at the human lipoprotein lipase (LPL) gene locus*. Nucleic Acids Res, 16(24):11856, 1988.
- LINS AM, Sprecher CJ, Puers C, Schumm JW. *Multiplex sets for the amplification of polymorphic short tandem repeat loci – silver stain and fluorescence detection*. BioTechniques, 20:882-889, 1996.
- LIU, Y; Saha N; Low PS; Tay JS. *Linkage disequilibrium between two loci (5' untranslated exon 1 and intron 5-DdeI) of the antithrombin III gene in three ethnic groups in Singapore*. Hum Hered, 45(4):192-8, 1995.
- LUIZON MR. *Dinâmica da mistura étnica em comunidades remanescentes de quilombo Brasileiras*. 2007, 115F. Tese (Doutorado – Programa de Pós-Graduação em Genética.) – Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.
- MACHADO, TMB. Bomfim, T.F , Acosta, AX, Galvão-Castro, B, Abe-Sandes, K. *Associação entre sobrenomes de conotação religiosa e ancestralidade genômica em Salvador-BA-Brasil*. FIOCRUZ - BA – Fundação Oswaldo Cruz, 2008.
- MACHADO, TMB, Bomfim, TF , Acosta, AX, Galvão-Castro, B , Sousa, SMB, Abe-Sandes, K . *Estimativa de mistura genética em uma amostra da população de Salvador-BA-Brasil*. FIOCRUZ - BA – Fundação Oswaldo Cruz, 2008.

MARTINEZ-JARRETA BM, Roche PD, Abecia E. *Genetic variation at six STR loci (HUMTH01, HUMTPOX, HUMCSF1PO, HUMF13A01, HUMFES/FPS, HUMVWFA31) in Aragon (north Spain)*. Forensic Sci Intern, 100:87-92, 1999.

MENDES-JÚNIOR CT. Frequências alélicas de polimorfismos de DNA do tipo STR em indígenas da Amazônia brasileira. Máster Thesis, Universidade de São Paulo, Ribeirão Preto, 2001.

MENDES-JUNIOR, CT. *Frequências alélicas de polimorfismos de DNA do tipo STR em indígenas da Amazônia brasileira*. 2001, Dissertação de Mestrado – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

MOURANT AE, Tills D, Domaniewska-Sobczak K. *Sunshine and the geographical distribution of the alleles of the Gc system of plasma proteins*. Hum Genet, 33(3):307-14, 1976.

NICOLL A, Basset K, Ulijaszek SJ. *What's in a name? Accuracy of using surnames and forenames in ascribing ethnic identity in English populations*. J. Epidemiol Community Health, 40:364-368, 1986.

OTA, T. *DISPAN: genetic distance and phylogenetic analysis*. Institute of Molecular Evolutionary Genetics, the University of Pennsylvania State University, University Park, PA, USA, 1993.

OTTENSOOSER F. *Cálculo do grau de mistura racial através de grupos sanguíneos*. Rev Bras Biol 4:531-37, 1944.

PARRA EJ, Kittles RA, Argyropoulos G, Pfaff CL, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey WT, Jin L, McKeigue PM, Kamboh MI, Ferrell RE, Pollitzer WS, Shriver MD. *Ancestral proportions and ture dynamics in geographically defined African Americans living in South Carolina*. Am J Phys Anthropol, 114(1):18-29, 2001.

PARRA EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, Forrester T, Allison DB, Deka R, Ferrell RE, Shriver MD. *Estimating African American ture proportions by use of population-specific alleles*. Am J Hum Genet, 63(6):1839-51, 1998.

PARRA FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. *Color and genomic ancestry in Brazilians*. Proc Natl Acad Sci U S A, 100(1):177-82, 2003.

PENA SDJ, Jeffreys AJ. *Breve introdução às impressões digitais de DNA*. Rev Bras Genet, 16(3):857-879, 1993.

- PENA SDJ. *Razões para banir o conceito de raça da medicina brasileira*. História, Ciências, Saúde – Manguinhos. 12(1): 321-346, 2005.
- PÉREZ-LEZAUN A, Calafell F, Clarimón J, Bosch E, Mateu E, Gusmão L, Amorim A, Benchemsi N, Bertranpetit J. Allele frequencies of 13 short tandem repeats in population samples from the Iberian Peninsula and Northern Africa. *Int J Legal Med*, 113:208-214, 2000.
- PRIMO-PARMO SL, Chautard-Freire-Maia EA, Lourenço MAC, Salzano FM, Freitas MJM. *Studies on serum cholinesterase (CHE1 and CHE2) in Brazilian Indian and ed populations*. *Rev Bras Genet* 9:467-8, 1986.
- RAYMOND, M. & Rousset, F. *GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism*. *J Hered*, 86:248-249, 1995.
- REED TE. *Number of gene loci required for accurate estimation of ancestral population proportions in individual human hybrids*. *Nature*, 244:575-576, 1973.
- REES JL. “*Genetics of hair and skin color*”. *Annu Rev Genet*, 37:67-90, 2003.
- REYNOLDS J, Weir BS & Cockerham CC. *Estimation of the coancestry coefficient: basis for a short term genetic distance*. *Genetics*, 105:767-779, 1983.
- RIBEIRO D. *Povo Brasileiro: Formação e sentido do Brasil*. São Paulo, Companhia das Letras, 1995.
- SAITOU N & Nei M. *The neighbor-joining method: a new method for reconstructing phylogenetic trees*. *Mol Biol Evol*, 4(4):406-425, 1987.
- SALZANO FM, Tondo CV. *Hemoglobin types of Brazilian populations*. *Hemoglobin* 6:85-97, 1982.
- SANDFORD AJ, Weir TD, Paré PD. Genetic risk factors for chronic obstructive pulmonary disease. *Eur Respir J*, 10(6):1380-1391, 1997.
- SANS M, *ture studies in Latin America: from the 20th to the 21st century*. *Hum Biol*, 72:155-177, 2000.
- SANTOS SEB, Guerreiro JF. *The indigenous contribution to the formation of the population of the Brazilian Amazon region*. *Rev Bras Genet*, 18:311-315, 1995.
- SANTOS SEB, Ribeiro-Rodrigues EM, Ribeiro-dos-Santos AKC, Hutz MH *et al*. *Autosomal STR Analyses in Native Amazonian Tribes Suggest a Population Structure Driven by Isolation by Distance*. *Human Biology*, 81(1):71-88, 2009.

- SCHUMM JW, Lins AM, Micka KA, Sprecher CJ, Rabbach DR, Bacher JW. *Automated fluorescent detection of STR multiplexes – development of the Geneprint™ PowerPlex™ and FFFL multiplexes for forensic and paternity applications*. First Europ Sympos Hum Identif, 90-104, 1997.
- SCHUMM JW, Lins AM, Puers C, Sprecher CJ. *Development of nonisotopic multiplex amplification sets for analysis of polymorphic STR loci*. First Europ Sympos Hum Identif, 177-181, 1995.
- SCHWARTZ RS. *Racial profiling in medical research* N Engl J Med, 344:1392-1393, 2001.
- SHEN MR, Batzer MA, Deininger PL. *Evolution of the master Alu gene(s)*. J Mol Evol, 33(4):311-20, 1991.
- SHRIVER MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE. *Ethnic-affiliation estimation by use of population-specific DNA markers*. Am J Hum Genet, 60(4):957-64, 1997.
- STEPANOV VA, Lemza SV. *PvuII restriction fragment length polymorphism of lipoprotein lipase in Russians*. Hum Hered, 43(6):388-90, 1993.
- STEPHENS JC, Briscoe D, O'Brien SJ. *Mapping by ture linkage disequilibrium in human populations: limits and guidelines*. Am J Hum Genet, 55:809-824, 1994.
- STURM RA, BOX NF, RAMSAY M. *“Human pigmentation genetics: The difference is only skin dee”*. Bioessays, 20:712-721, 1998.
- TAVARES-NETO J, Azevêdo ES. *Family names and ABO blood groups in a mixed population of Bahia, Brazil*. Hum Biol, 50:361-367, 1978.
- TAVARES-NETO J, Azevêdo ES. *Racial origin and historical aspects family names in Bahia, Brazil*. Hum Biol, 49:287-289, 1977.
- TOMAS G, Seco L, Seixas S, Faustino P, Lavinha J, Rocha J. *The peopling of Sao Tome (Gulf of Guinea): origins of slave settlers and ture with the Portuguese*. Hum Biol, 74(3):397-411, 2002.
- TOURNAMILLE C, Colin Y, Cartron JP, Le Van Kim C. *Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals*. Nature Genet, 10:224-228, 1995.
- WANDERLEY-SANTOS LM. STRs autossômicos e ligados ao cromossomo Y em indígenas brasileiros. PhD Thesis, Universidade de São Paulo, Ribeiro Preto, 2001.

VECCHIO C, Garofano L, Saravo L *et al.* *Allele frequencies for CODIS loci in a Sicilian population sample.* International Congress Series, 1261: 136-138, 2004.

VIEIRA-SILVA C, Cruz C, Ribeiro T, Espinheira R. *South Portugal population genetic analysis with 17 loci STRs.* International Congress Series, 1288: 367-368, 2006.

WATKINS WS, Ricker CE, Bamshad MJ, Carroll ML, Nguyen SV, Batzer MA, Harpending HC, Rogers AR, Jorde LB. *Patterns of ancestral human diversity: an analysis of Alu-insertion and restriction-site polymorphisms.* Am J Hum Genet, 68(3):738-52, 2001.

WEEDN VW, SWARNEN SL. *Exames forenses de identificação por análises do DNA.* in: Henry, J. B. *et al* *Diagnósticos Clínicos e Tratamento por Métodos Laboratoriais* 19^a ed. São Paulo: Manole, 63:1427-1438, 1998.

WITKE WF, Gibbs PE, Zielinski R, Yang F, Bowman BH, Dugaiczky A. *Complete structure of the human Gc gene: differences and similarities between members of the albumin gene family.* *Genomics.* 16(3):751-4, 1993.

WRIGHTON SA, Stevens JC. *The human hepatic cytochromes P450 involved in drug metabolism.* Crit Rev Toxicol, 22: 1-21, 1992.

YAMAMOTO T, Uchihi R, Nozawa H, Huang XL, Leong YK, Tanaka M, Mizutani M, Tamaki K, Katsumata Y. *Allele distribution at nine STR loci – D3S1358, VWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820 – in the Japanese population by multiplex PCR and capillary electroph.* J Forensic Sci, 44(1):167-170, 1999.

ZAGO MA, Silva-Jr WA, Tavela MH, Santos SEB, Guerreiro JF, Figueiredo MS. *Interpopulational and intrapopulational genetic diversity of Amerindians as revealed by six variable number of tandem repeats.* Human Heredity, 46:274-289, 1996.

APÊNDICES

APÊNDICE 1. Frequências dos alelos STRs nas amostras testadas e populações ancestrais.

Total de alelos	D3S1358	Bahia ¹	EUROPEUS ²	AFRICANOS ³	AMERÍNDIOS ³
1	10	0,0000	0,0050	0,0000	0,0000
2	11	0,0000	0,0030	0,0000	0,0000
3	12	0,0050	0,0018	0,0000	0,0000
4	13	0,0070	0,0040	0,0000	0,0150
5	14	0,0860	0,1147	0,0836	0,0590
6	15	0,3500	0,2389	0,3197	0,5908
7	16	0,2560	0,2526	0,3355	0,3198
8	17	0,1850	0,2093	0,2047	0,0857
9	18	0,0990	0,1611	0,0571	0,0610
10	19	0,0120	0,0129	0,0000	0,0000
11	20	0,0000	0,0030	0,0000	0,0000
Total de alelos	D7S820	Bahia ¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	6	0,0020	0,0000	0,0000	0,0000
2	7	0,0050	0,0233	0,0088	0,0150
3	8	0,1970	0,1652	0,1913	0,1256
4	9	0,1210	0,1440	0,1217	0,0820
5	10	0,2680	0,2711	0,3569	0,4106
6	10.3	0,0000	0,0000	0,0040	0,0000
7	11	0,1970	0,1995	0,2088	0,2295
8	12	0,1630	0,1590	0,0805	0,2725
9	13	0,0340	0,0298	0,0284	0,0160
10	14	0,0120	0,0070	0,0045	0,0000
11	15	0,0000	0,0025	0,0000	0,0000
Total de alelos	D8S1179	Bahia ¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	8	0,0020	0,0147	0,0000	0,0210
2	9	0,0000	0,0129	0,0000	0,0420
3	10	0,0690	0,0968	0,0110	0,1287
4	11	0,0710	0,0783	0,0402	0,0364
5	12	0,1310	0,1255	0,1651	0,1740
6	13	0,2390	0,2992	0,2078	0,2356
7	14	0,2960	0,2122	0,3373	0,3538
8	15	0,1350	0,1337	0,1820	0,1282
9	16	0,0540	0,0213	0,0567	0,0210
10	17	0,0020	0,0051	0,0110	0,0000
Total de alelos	TH01	Bahia ¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	4	0,0000	0,0010	0,0000	0,0340
2	5	0,0000	0,0000	0,0000	0,0470
3	6	0,1850	0,2422	0,1113	0,3995
4	7	0,2460	0,1656	0,3733	0,2973

5	8	0,1280	0,1124	0,3018	0,0165
6	9	0,1950	0,1912	0,1685	0,0163
7	9.3	0,2460	0,2769	0,0389	0,2495
8	10	0,0000	0,0171	0,0136	0,0000
Total de alelos	TPOX	Bahia¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	5	0,0000	0,0020	0,0000	0,0000
2	6	0,0340	0,0030	0,1043	0,0000
3	7	0,0050	0,0050	0,0240	0,0070
4	8	0,4210	0,5008	0,2776	0,3120
5	9	0,1500	0,1088	0,1809	0,0135
6	10	0,0760	0,0649	0,0892	0,0375
7	11	0,2730	0,2879	0,3039	0,3970
8	12	0,0390	0,0333	0,0208	0,2727
Total de alelos	VWA	Bahia¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	11	0,0120	0,0050	0,0217	0,0000
2	12	0,0020	0,0250	0,0000	0,0000
3	13	0,0020	0,0045	0,0093	0,0000
4	14	0,0640	0,1049	0,0813	0,0400
5	15	0,1670	0,1124	0,2317	0,0467
6	16	0,2860	0,2344	0,2652	0,4642
7	17	0,2220	0,2637	0,1470	0,3376
8	18	0,1450	0,1869	0,1320	0,1424
9	19	0,0790	0,0695	0,0848	0,0263
10	20	0,0150	0,0113	0,0227	0,0190
11	21	0,0050	0,0038	0,0091	0,0000
Total de alelos	D18S51	Bahia¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	9	0,0000	0,0000	0,0040	0,0000
2	10	0,0070	0,0100	0,0000	0,0000
3	10.2	0,0000	0,0010	0,0110	0,0000
4	11	0,0120	0,0132	0,0088	0,0135
5	11.2	0,0000	0,0000	0,0040	0,0000
6	12	0,0990	0,1509	0,0517	0,2320
7	12.2	0,0000	0,0000	0,0040	0,0000
8	13	0,0620	0,1098	0,0342	0,1213
9	13.2	0,0000	0,0025	0,0136	0,0000
10	14	0,1280	0,1702	0,0402	0,2938
11	14.2	0,0070	0,0000	0,0045	0,0000
12	15	0,1650	0,1433	0,1661	0,1132
13	16	0,1400	0,1409	0,1766	0,0690
14	17	0,1700	0,1045	0,1743	0,2170
15	18	0,0990	0,0702	0,1327	0,0633
16	18.2	0,0020	0,0000	0,0000	0,0000
17	19	0,0490	0,0551	0,1132	0,0320
18	20	0,0370	0,0152	0,0439	0,0135
19	21	0,0120	0,0091	0,0251	0,0120
20	21.2	0,0000	0,0000	0,0091	0,0000
21	22	0,0070	0,0028	0,0136	0,0150
22	23	0,0020	0,0030	0,0040	0,0920
23	24	0,0000	0,0050	0,0000	0,0290

Total de alelos	FGA	Bahia ¹	EUROPEUS	AFRICANOS	AMERÍNDIOS
1	12	0,0020	0,0000	0,0000	0,0000
2	15	0,0000	0,0010	0,0000	0,0000
3	16	0,0000	0,0010	0,0030	0,0000
4	17	0,0100	0,0000	0,0046	0,0000
5	17.2	0,0000	0,0000	0,0030	0,0000
6	18	0,0020	0,0094	0,0030	0,0390
7	18.2	0,0100	0,0010	0,0061	0,0000
8	19	0,0840	0,0636	0,0767	0,0856
9	19.2	0,0000	0,0000	0,0100	0,0000
10	20	0,0990	0,1302	0,0721	0,0483
11	20.2	0,0000	0,0018	0,0000	0,0000
12	21	0,1430	0,1826	0,0983	0,0545
13	21.2	0,0000	0,0105	0,0030	0,0000
14	22	0,1550	0,1725	0,1594	0,0390
15	22.2	0,0000	0,0128	0,0000	0,0000
16	23	0,1630	0,1574	0,1417	0,1630
17	23.2	0,0000	0,0035	0,0000	0,0580
18	24	0,1770	0,1452	0,1887	0,1328
19	25	0,0790	0,0737	0,1041	0,3180
20	25.2	0,0000	0,0010	0,0000	0,0000
21	26	0,0590	0,0360	0,0731	0,2006
22	26.2	0,0000	0,0010	0,0000	0,0000
23	26.3	0,0000	0,0010	0,0000	0,0000
24	27	0,0070	0,0087	0,0279	0,0263
25	27.2	0,0000	0,0025	0,0000	0,0000
26	28	0,0050	0,0000	0,0083	0,0170
27	29	0,0020	0,0050	0,0136	0,0000
28	29.2	0,0000	0,0000	0,0030	0,0000
29	30.2	0,0000	0,0000	0,0083	0,0000
30	31.2	0,0020	0,0000	0,0046	0,0000
31	42.2	0,0000	0,0000	0,0046	0,0000
32	43.2	0,0000	0,0000	0,0136	0,0000

¹ = presente estudo; ² = população européia; ³ = população africana; ⁴ = população ameríndia.

APÊNDICE 2. Análise de diferenciação populacional utilizando STRs em populações com e sem sobrenomes de conotação religiosa na Bahia – Brasil.

<i>Loci</i>	<i>Valor de P</i>
TH01	0,02913
VWA	0,87168
D7S820	0,28796
D18S51	0,73785
FGA	0,37226

D3S1358	0,15696
D8S1179	0,67135
TPOX	0,31697

Análise no Genepop - Método de Fisher

APÊNDICE 3. Análise diferenciação populacional utilizando AIMs em populações com e sem sobrenomes de conotação religiosa na Bahia – Brasil.

<i>Loci</i>	<i>Valor de P</i>
<i>AT3-I/D</i>	0,00395
<i>APO</i>	0,32123
<i>SB 19.2</i>	0,01488
<i>PV 92</i>	0,50012
<i>FY-Null</i>	0,00330
<i>CKMM</i>	0,74647
<i>LPL</i>	0,31281
<i>CYP3A4</i>	0,00041
<i>GC</i>	0,00055

Análise no Genepop - Método de Fisher