

UNIVERSIDADE ESTADUAL DE SANTA CRUZ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E
BIOLOGIA MOLECULAR



MARCADORES INFORMATIVOS DE ANCESTRALIDADE (AIMs) EM
MULHERES DO SUL DA BAHIA

Raoni Souza Ferreira

ILHÉUS-BAHIA-BRASIL

Agosto de 2012

RAONI SOUZA FERREIRA

Marcadores Informativos de Ancestralidade (AIMs) em mulheres do sul da Bahia

Dissertação apresentada à Universidade Estadual de Santa Cruz como parte das exigências para obtenção do título de Mestre em Genética e Biologia Molecular

Orientadora: Sandra Mara Bispo Sousa
Co-orientadora: Sandra Rocha Gadelha

ILHÉUS-BAHIA-BRASIL

Agosto de 2012

RAONI SOUZA FERREIRA

**MARCADORES INFORMATIVOS DE ANCESTRALIDADE (AIMs) EM
MULHERES DO SUL DA BAHIA**

Dissertação apresentada à Universidade Estadual de Santa Cruz como parte das exigências para obtenção do título de Mestre em Genética e Biologia Molecular

Área de concentração: Genética e Biologia Molecular.

APROVADA: 30 de agosto de 2012

Prof^ª. Dr^ª. Ana Angelica Leal Barbosa
(UESB)

Prof. Dr. Paulo Roberto Santana de Melo
(UESC)

Prof^ª. Dr^ª. Patrícia Santos Pereira Lima
(UESB)

Prof^ª. Dr^ª. Sandra Mara Bispo Sousa
(UESC- Orientadora)

Aos meus amigos e família, fontes de conhecimento.

Dedico

AGRADECIMENTOS

A CAPES, pela concessão da bolsa.

Ao Programa de Genética e Biologia Molecular, pela oportunidade de desenvolver o projeto.

A professora Sandra Mara Bispo Sousa, pela orientação e por ter topado esse desafio comigo. Por ter sido tão dedicada, compreensiva e humana, qualidades raras na academia. Pelo carinho e amizade.

A Fabrícia Santos, secretária do programa, pelo carinho e paciência com que sempre me ouviu, ajudou com as pendências e salvou inúmeras vezes.

Aos professores(as) Sandra Gadelha, Lauro Marin e Paulo Melo; por todo apoio que prestaram, das mais diversas formas.

Aos amigos Ronaldo Carvalho, Livia Santana, Cleiziane Bispo, Cintia Marques, Luciana Camilo e Ohana Oliveira; por me ouvirem e incentivarem. Pelas caronas e ônibus que pegamos juntos, pelos almoços, conselhos estratégicos, conselhos amorosos, arrochas, fugidas, cobranças, viagens, abraços, pastéis, hospedagens, bosques, festas, praias. Tudo que dá sentido a vida, enfim.

A amiga e companheira Jamille Oliveira, que dividiu comigo essa experiência em tempo integral. Por tudo que a devo e, felizmente, levarei algumas encarnações pra pagar. Por tudo, de coração.

A tia Márcia, que de muitas formas foi responsável pelo começo da minha vida acadêmica. Ainda assim eu a perdoo.

A painho, mainha, Rai, Rafa, Donna, vovó Ivete e vovô Wilson; pois o lar é onde o coração está.

Meus sinceros agradecimentos.

Antes de mais nada é necessário acabar com o mito de que o cientista é uma pessoa que pensa melhor do que as outras.

Rubem Alves

LISTA DE ABREVEATURAS

AFR – Africano

AIMs – *Ancestry Informative Markers*

AMR – Ameríndio

APO – Gene da apolipoproteína

AT3 – Gene da antitrombina III

BA – Estado da Bahia, Brasil

CE – Estado do Ceará, Brasil

dbSNP/NCBI – Banco de dados do *National Center for Biotechnology Information*

DNA – *Deoxyribonucleic acid*

dNTP – *Deoxyribonucleoside triphosphate*

EDTA – *Ethylenediamine tetraacetic acid*

EHW – Equilíbrio de Hardy-Weinberg

EUR – Europeu

FUNAI – Fundação Nacional do Índio

HTLV-1 – *Human T lymphotropic virus type 1*

IBGE – Instituto Brasileiro de Geografia e Estatística

Indel – Inserção-deleção

IOS – Ilhéus

ITB – Itabuna

LD – *Linkage disequilibrium*

MIAs – Marcadores Informativos de Ancestralidade

pb – Pares de base

PCR – *Polymerase Chain Reaction*

PCR-RFLP – *Polymerase Chain Reaction- Restriction Fragment Length Polymorphism*

PSAs – *Population-specific Alleles*

PV92 – Sequência repetida *Alu* PV92

RS – Estado do Rio Grande do Sul, Brasil

Sb19.3 – Sequência repetida *Alu* Yb8 Sb19.3

SBA – Sul da Bahia

SINEs – *Short Interspersed Nuclear Elements*

SNPs – *Single Nucleotide Polymorphisms*

SSRs – *Simple Sequence Repeats*

SUS – Sistema Único de Saúde

TBE – Tampão Tris/Borato/EDTA

TCLE – Termo de Consentimento Livre e Esclarecido

TEMED – Tetrametilenodiamina

USP – Universidade de São Paulo

UTI – Unidade de Terapia Intensiva

LISTA DE TABELAS

Tabela 1: Localização citogenética e sequência dos <i>primers</i> dos quatro AIMs analisados	32
Tabela 2: Condições de PCR para os <i>loci</i> analisados no presente estudo. Quantidade em μL suficiente para uma reação.....	33
Tabela 3: Frequências alélicas dos quatro <i>loci</i> avaliados nas amostras de Ilhéus (IOS) e Itabuna (ITB) seguido das médias, comparadas aos valores encontrados em africanos (AFR), europeus (EUR) e ameríndios (AMR).....	43
Tabela 4: Diferencial de frequência (δ) entre a média dos valores encontrados no presente estudo (SBA) e valores encontrados para africanos (AFR), europeus (EUR) e ameríndios (AMR).....	44
Tabela 5: Estimativas de <i>p</i> -valor segundo o teste exato para verificação de aderência ao equilíbrio de Hardy-Weinberg (Guo e Thompson, 1992). Valores significativos ($p < 0,05$) estão destacados em negrito.....	45
Tabela 6: Comparação par-a-par entre <i>loci</i> não ligados em amostra do sul da Bahia.....	46
Tabela 7: Avaliação da diferenciação gênica e genotípica das subamostras de Ilhéus e Itabuna na população do sul da Bahia através do teste exato G.....	47
Tabela 8: Diversidade gênica para cada <i>locus</i> em cada amostra. Destacados em negrito maior e menor valor.....	48
Tabela 9: Estimativa de mistura étnica na população do sul da Bahia, seguidas do respectivo erro padrão e do coeficiente de adaptação (R^2) ao teste aplicado.....	49

LISTA DE FIGURAS

- Figura 1:** Padrão eletroforético do *locus* APO em agarose 2% corado com GelRED, indicando número de raias e tamanho dos fragmentos gerados..... 39
- Figura 2:** Padrão eletroforético do *locus* Sb19.3 em agarose 1% corado com GelRED, indicando número de raias e tamanho dos fragmentos gerados..... 40
- Figura 3:** Padrão eletroforético do *locus* PV92 em agarose 2% corado com GelRED, indicando número de raias e tamanho dos fragmentos gerados..... 41
- Figura 4:** Padrão eletroforético do *locus* AT3 em agarose 3% corado com GelRED, indicando número de raias e tamanho dos fragmentos gerados. As raias 1 e 2 contem, respectivamente *ladder* de 100 pb e controle negativo da reação..... 42
- Figura 5:** Representação gráfica da ancestralidade autodeclarada (AUT) e genética (GEN) em amostra do sul da Bahia. O modelo ti-híbrido é consistente com contribuição europeia (EUR), africana (AFR) e ameríndia (AMR)..... 50
- Figura 6:** Representação gráfica das frequências do alelo *I dos sete AIMs avaliados na amostra do sul da Bahia (SBA), comparados aos valores encontrados entre africanos (AFR), europeus (EUR) e ameríndios (AMR)..... 53
- Figura 7:** Representação gráfica do diferencial de frequência (δ) entre a amostra estudada e a população africana (AFR), europeia (EUR) e ameríndia (AMR), por *locus*... 55

ÍNDICE

RESUMO	xiv
ABSTRACT	xvi
INTRODUÇÃO	18
OBJETIVOS	21
GERAL.....	21
ESPECÍCOS.....	21
REVISÃO BIBLIOGRÁFICA	22
DIVERSIDADE GENÉTICA NA ESPÉCIE HUMANA.....	22
MARCADORES INFORMATIVOS DE ANCESTRALIDADE (AIMs).....	23
FORMAÇÃO DA POPULAÇÃO BRASILEIRA.....	26
METODOLOGIA	30
AMOSTRAGEM.....	30
ASPECTOS ÉTICOS.....	30
EXTRAÇÃO DE DNA.....	31
MARCADORES MOLECULARES SELECIONADOS.....	31
REAÇÃO EM CADEIA DA POLIMERASE (PCR).....	32

ANÁLISE DO PRODUTO AMPLIFICADO.....	33
ANÁLISE DOS DADOS.....	35
Frequências Alélicas.....	35
Aderência ao Equilíbrio de Hardy-Weinberg.....	36
Associações par-a-par entre <i>loci</i>	36
Diferenciação e diversidade genética das populações.....	37
Estimativas de mistura étnica.....	37
RESULTADOS.....	39
PERFIL ELETROFORÉTICO.....	39
FREQUÊNCIAS ALÉLICAS.....	42
ADERÊNCIA AO EQUILÍBRIO DE HARDY-WEINBERG.....	44
ASSOCIAÇÕES PAR-APAR ENTRE <i>LOCI</i>	45
DIFERENCIAÇÃO GENÉTICA DAS AMOSTRAS.....	47
DIVERSIDADE GENÉTICA.....	48
MISTURA ÉTNICA.....	48
DISCUSSÃO.....	51
CONTEXTUALIZAÇÃO HISTÓRICA.....	51
FREQUÊNCIAS ALÉLICAS.....	52

EHW E ASSOCIAÇÃO PAR-APAR ENTRE <i>LOC</i> I NÃO LIGADOS.....	55
DIFERENCIAÇÃO E DIVERSIDADE GENÉTICA.....	56
MISTURA ÉTNICA.....	57
CONCLUSÕES.....	61
REFERÊNCIAS BIBLIOGRÁFICAS.....	62
ANEXO I: Mapa do estado da Bahia, destacando microrregião de Ilhéus-Itabuna.....	69
ANEXO II: Termo de Consentimento Livre e Esclarecido.....	70

RESUMO

FERREIRA, Raoni Souza, M.S., Universidade Estadual de Santa Cruz, Ilhéus, agosto de 2012. Marcadores **Informativos de Acestralidade (AIMs) em mulheres do sul da Bahia**. Orientadora: Sandra Mara Bispo Sousa. Co-orientadora: Sandra Rocha Gadelha. Conselheiro: Lauro Juliano Marin.

Assim como a brasileira, a população Baiana é o resultado de cinco séculos de mistura étnica entre europeus, ameríndios e africanos. Devido à intensa miscigenação, é presumível que qualquer caracterização física individual é arriscada. Dessa forma, a utilização de marcadores de ancestralidade - *Ancestry informative markers* (AIMs) - é fundamental para se obter uma caracterização confiável seja em nível individual ou populacional. Com o objetivo de caracterizar a composição genética da região sul da Bahia foi utilizada uma amostra de 750 mulheres, sendo 372 amostras coletadas em Ilhéus-BA e 378 em Itabuna-BA e avaliados sete AIMs (Sb19.3, PV92, APO, AT3, MID-93, MID-52 e MID-575) que apresentam alto diferencial de frequência (δ) entre as populações parentais. O DNA extraído com o uso de kits QIAGEN, foi submetido a PCR e o produto amplificado foi separado por eletroforese em gel de agarose (2% e 3%) e gel desnaturante de poliacrilamida (12%), corado com GelRed e AgNO_3 . A análise estatística foi realizada com os programas GENEPOP 4.0, FSTAT 2.9 e ADMIX3. Os *loci* Sb19.3 e APO apresentaram desvio do Equilíbrio de Hardy-Weinberg (EHW) ($p = 0$) na amostra, por apresentar déficit de heterozigotos ($p = 0$). Miscigenação e estruturação populacional podem explicar os desvios. A comparação par-a-par entre *loci* não ligados, foi significativa entre Sb19.3 x APO ($p = 0,04$) e APO x AT3 ($p = 0,04$). Não houve diferenciação gênica ou genotípica entre as subamostras quando aplicado o teste exato G, o que era esperado de populações geográfica e historicamente muito próximas, especialmente considerando limitações próprias dos marcadores. As frequências alélicas médias obtidas para os *loci* PV92, APO, AT3,

Sb19.3, MID-93, MID-52 e MID-575 foram, respectivamente, 0,26; 0,80; 0,47, 0,65, 0,49; 0,77 e 0,88. Esses marcadores foram informativos, respectivamente, para ancestralidade ameríndia ($\delta = 0,66$); africana ($\delta = 0,39$); africana ($\delta = 0,38$) ou ameríndia ($\delta = 0,42$); ameríndia ($\delta = 0,41$) e europeia ($\delta = 0,69$ e $0,87$) ou africana ($\delta = 0,41$ e $0,76$), com exceção do Sb19.3 ($\delta < 0,3$). A heterozigose média dentro das populações (H_S) variou de 0,20 no *locus* MID-575 a 0,51 no MID-93. A heterozigose média total (H_T) foi de 0,39 e o coeficiente de endogamia (F_{IS}) = 0,11; compatíveis com baixa endogamia, alto fluxo gênico e grande tamanho amostral. A estimativa de mistura étnica se adequou a um modelo tri-híbrido, apresentando contribuição europeia, africana e ameríndia de respectivamente $\approx 51\%$, 38% e 11% para a amostra total. No presente estudo, o *locus* Sb19.3 não se mostrou um bom marcador de ancestralidade. As subamostras de Ilhéus e Itabuna foram homogêneas, não havendo diferenciação genética entre ambas ($F_{ST} = 0,00$). As frequências alélicas dos AIMs analisados estão próximas às descritas na literatura, apesar dos índices de mistura e estruturação populacional. Apesar da histórica influência étnica africana na Bahia, indivíduos que se autodeclararam negros ou pardos apresentaram uma alta ancestralidade genética europeia.

Palavras-chave: miscigenação, antropogenética, etnia autodeclarada, AIMs, ancestralidade genética.

ABSTRACT

FERREIRA, Raoni Souza, M.S., Universidade Estadual de Santa Cruz, Ilhéus, August 2012. **Ancestry Informative Markers (AIMs) in women from southern Bahia.** Advisor: Sandra Mara Bispo Sousa. Advisor Committee Membres: Sandra Rocha Gadelha Mello and Lauro Juliano Marin.

As well as Brazilian, Bahian population is result of five centuries of ethnic admixture among Europeans, Amerindians and Africans. Due to intense admixture, it is presumed that any individual physical characterization is unreliable. Thus, use of ancestry-informative markers (AIMs) is essential to obtain a reliable characterization at individual or population level. In order to characterize genetic composition of southern Bahia, was used a sample of 750 women - 372 samples collected in Ilhéus-BA and 378 in Itabuna-BA - and it was analyzed seven AIMs (Sb19.3, PV92, APO, AT3, MID-93, MID-52 and MID-575) that have a high frequency differential (δ) among parental populations. DNA was extracted using QIAGEN kits and then subjected to PCR. The amplified product was separated by electrophoresis on agarose gel (2% and 3%) stained with GelRed and polyacrylamide denaturing gel stained with AgNO_3 . Statistical analysis was performed with softwares GENEPOP 4.0, ADMIX3 and FSTAT 2.9. *Loci* Sb19.3 and APO showed deviation from Hardy-Weinberg equilibrium (HWE) ($p = 0$) in sample, because the deficit of heterozygotes ($p = 0$). Admixture and population structure may explain observed deviations. Pairwise comparison between unlinked *loci* was significant between Sb19.3 x APO ($p = 0.04$) and APO x AT3 ($p = 0.04$). There was no genic or genotypic difference between subsamples when applied exact G test. This is expected for populations geographically and historically closer, especially considering limitations of markers. Average allele frequencies from *loci* PV92, APO, AT3, Sb19.3, MID-93, MID-52 e MID-575 were respectively 0.26, 0.80, 0.47, 0.65, 0.49, 0.77 and 0.88. These markers were

informative, respectively, for Amerindian ancestry ($\delta = 0.66$), African ancestry ($\delta = 0.39$) and African ($\delta = 0.38$) or Amerindian ($\delta = 0.42$) ancestry, Amerindian ancestry ($\delta = 0,41$) and European ($\delta = 0,69$ e $0,87$) or African ($\delta = 0,41$ e $0,76$) ancestry, except for Sb19. 3 ($\delta < 0.3$). Average heterozygosity within populations (H_S) ranged from 0.20 in MID-575 *locus* to 0.51 in MID-93 *locus*. Average total heterozygosity (H_T) was 0.39 and inbreeding coefficient (F_{IS}) = 0.11. These results are compatible with low inbreeding, high gene flow and large sample size. Estimated ethnic admixture has adapted to a tri-hybrid model, with European, African and Amerindian contributions respectively of $\approx 51\%$, 38% and 11% for the total sample. In present study, *locus* Sb19.3 was not a good ancestry marker. Subsamples of Ilhéus and Itabuna were homogeneous. There was no genetic differentiation between them ($F_{ST} = 0.00$). Allele frequencies of analyzed AIMs are close to literature, even though observed rates of admixture and population structure. Despite of historic African ethnic influence in Bahia, individuals who declare themselves as black or brown showed a high European genetic ancestry.

Keywords: admixture, anthropogenic, self-reported ethnicity, AIMs, genetic ancestry.

INTRODUÇÃO

A chegada dos primeiros colonizadores ao território baiano deu início a um processo de mistura entre povos que se estende há aproximadamente cinco séculos ou, em termos genéticos, vinte gerações. Três grupos figuram majoritariamente neste processo, sendo eles os (i) indígenas que habitavam o território cerca de 200 anos antes da chegada dos primeiros (ii) colonizadores europeus, em sua maioria portugueses e os (iii) africanos trazidos como escravos para trabalhar nas lavouras e, posteriormente, nas minas (ALVES-SILVA *et al.* 2000; TAVARES, 2001).

A participação dessas etnias no povoamento da região do sul da Bahia foi influenciada por pelo menos dois grandes ciclos agrícolas com resultados distintos, ambos relevantes no processo de miscigenação. O primeiro, da cana-de-açúcar, não foi bem sucedido. Já o segundo, do cacau, com os incentivos do governo ao plantio e a ideia largamente difundida de enriquecimento certo e rápido, atraiu grandes levas de imigrantes de todo o Brasil para a região, que experimentou um rápido crescimento econômico e populacional (ANDRADE, 2003; CAMPOS, 2006).

As relações socioculturais entre colonizadores, escravos e nativos também marcaram profundamente a constituição étnica do povo brasileiro e, conseqüentemente, do povo baiano. O cruzamento direcionado entre homens europeus e mulheres ameríndias ou africanas é um exemplo. Independente do aspecto cor, cada nova geração de indígenas ou negros trazia consigo uma grande contribuição genética europeia, embora a matrilinearidade (mtDNA) se mantivesse majoritariamente ameríndia ou africana (ALVES-SILVA *et al.* 2000; CARVALHO-SILVA *et al.* 2001; PENA *et al.* 2011).

A miscigenação é um processo observado em muitas populações humanas e consiste no encontro de duas ou mais populações ancestrais, as quais frequentemente são geneticamente distintas (PATTERSON *et al.* 2010). A contribuição genética de cada parental se difunde, fixando-se ou não na população derivada ao longo das

gerações. Durante esse processo a nova população pode se tornar estratificada ou subdividida, apresentando diferenças na proporção de mistura entre os grupos de indivíduos que a compõe (HOGGART *et al.* 2003).

A estratificação populacional pode agir como um fator de confusão em pesquisas de associação ou caso/controle, de forma que diversos trabalhos indicam a averiguação da composição étnica por meio de marcadores de ancestralidade (SHRIVER *et al.* 2003, GATTÁS *et al.* 2004, ZEMBRZUSKI *et al.* 2006).

Atualmente há apenas um trabalho desse tipo publicado com amostras de Ilhéus (ver Pena *et al.* 2011). Este, entretanto, não se detém na região sul da Bahia, e sim na constituição étnica de diferentes regiões do Brasil. Na análise dos dados da região nordeste, por exemplo, Pena *et al.* (2011) utilizaram 82 indivíduos de Fortaleza-CE, além dos já citados indivíduos de Ilhéus-BA.

O presente estudo conta com tamanho amostral robusto (756 indivíduos) e altamente representativo do território denominado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) como microregião de Ilhéus-Itabuna (Anexo I), referida aqui frequentemente como “população do Sul da Bahia”. As coletas foram realizadas na Maternidade Santa Helena (Ilhéus) e no Hospital e Maternidade Manoel Novaes (Itabuna), as principais unidades que realizam atendimento pelo Sistema Único de Saúde (SUS) nos referidos municípios e cidades circunvizinhas. Além da contribuição genética das populações parentais, o presente estudo comparou os resultados com a etnia autodeclarada e avaliou ainda valores de diversidade, estrutura da amostra total e diferenciação entre as subamostras.

A ferramenta molecular escolhida para as análises foram os *Ancestry Informative Markers* (AIMs). Os AIMs utilizados são de genotipagem fácil e relativamente barata, abundantes no genoma humano e possuem taxa de evolução lenta (RAY *et al.* 2005). Também possuem alto diferencial de frequência alélica (δ) entre populações mundiais e, portanto, podem ser utilizados para caracterizar a composição genética de populações miscigenadas (PARRA *et al.* 2003; BEUTEN *et al.* 2011).

O presente estudo avaliou duas hipóteses principais. A primeira diz respeito à composição étnica da população do sul da Bahia que, segundo dados históricos (ANDRADE, 2003; ANDRADE e ROCHA, 2005; CAMPOS, 2006), poderia ter contribuição africana menor do que a europeia e ameríndia em comparação à capital do estado, Salvador. A segunda diz respeito à diferenciação gênica entre as subamostras coletadas nas cidades de Ilhéus e Itabuna. Nesse caso, dois fatores indicavam que as subamostras colhidas nestes municípios poderiam ser geneticamente diferentes. O primeiro fator é a diferença histórica de fundação e povoamento dos dois municípios (ANDRADE e ROCHA, 2005; CAMPOS, 2006). O segundo fator é a prevalência diferenciada do vírus T-linfotrópico humano tipo 1 (HTLV-1), observado pela Prof^a. Sandra Rocha e colaboradores nas subamostras de Ilhéus e Itabuna (dados não publicados). O HTLV-1 possui prevalência diferenciada entre grupos étnicos (FIGUEIRÓ-FILHO *et al.* 2005, NASCIMENTO *et al.* 2009), o que poderia apontar uma contribuição parental diferenciada nas subamostras supracitadas.

OBJETIVOS

GERAL

- Caracterizar geneticamente uma amostra de mulheres do sul da Bahia, baseando-se em parâmetros populacionais mensurados por 7 marcadores informativos de ancestralidade (AIMs).

ESPECÍFICOS

- Avaliar a frequência alélica dos marcadores utilizados (Sb19.3, PV92, APO, AT3, MID-93, MID-52 e MID-575);
- Estimar a mistura étnica dessa amostra;
- Estimar parâmetros de diversidade genética inter e intrapopulacional;
- Correlacionar a etnia autodeclarada com ancestralidade genômica.

REVISÃO BIBLIOGRÁFICA

DIVERSIDADE GENÉTICA NA ESPÉCIE HUMANA

É largamente aceito no meio científico que o *Homo sapiens* moderno se originou na África entre 300-200 mil anos atrás, a partir de uma população do leste africano, embora indícios mais recentes apontem uma origem sul-africana (HENN *et al.* 2011). A cerca de 100 mil anos atrás começariam então as migrações da espécie, que colonizou não apenas o continente africano, mas virtualmente todo o globo. Esse modelo de surgimento e dispersão é conhecido como *Out of Africa* e tem graves implicações na moldagem da diversidade genética que observamos atualmente (STONEKING, 2008). De fato, os mecanismos de origem de nossa espécie tem implicações importantes para a compreensão dos padrões mundiais de diversidade genética humana observados (HARPENDING e ROGERS, 2000).

Como reflexo deste modelo, algumas tendências observadas à medida que nos afastamos do leste africano são a diminuição da heterozigosidade, o aumento do desequilíbrio de ligação ou, no inglês, *linkage disequilibrium* (LD), e um declínio na frequência de alelos ancestrais (DeGEORGIO *et al.* 2009). A grosso modo, significa dizer que cada nova população surgida com a expansão representa um subconjunto da população africana ancestral. Esse efeito fundador, somado ao acasalamento preferencial, a pressão local de seleção e a deriva gênica, durante milhares de anos foram fixando determinados alelos em determinadas populações. No entanto, a maior parte da variação é neutra, funcionalmente insignificante e pouco relevante fenotipicamente (BAMSHAD *et al.* 2004). Contudo uma pequena parte dela pode ter implicações diretas sobre o funcionamento do organismo, e mesmo a variação neutra pode estar estatisticamente associada ou em desequilíbrio de ligação com alelos de risco (GONZALES *et al.* 1999; DORIS, 2002; BEHAR *et al.* 2010; FEJERMAN *et al.* 2010).

MARCADORES INFORMATIVOS DE ANCESTRALIDADE – AIMs

É imprescindível esclarecer, de antemão, que diferenças entre indivíduos ou populações não representam, do ponto de vista genético, uma justificativa para a ideia de *raças*. Em primeiro lugar, a divergência das populações humanas é, no tempo evolutivo, recente para o surgimento de *raças*. Em segundo lugar, a diversidade observada fenotipicamente é regulada por um número ínfimo de genes comparado a todo genoma. Além disso, os valores de diversidade observados entre indivíduos de uma mesma população são normalmente pouco menores do que indivíduos de duas populações diferentes. Em suma, os valores de diversidade (F_{ST}) entre populações humanas são baixos, variando em torno de 5 a 15% (BAMSHAD *et al.* 2004; STONEKING, 2008; STURM, 2009).

O termo *raça* vem sendo categoricamente substituído no meio científico pelo termo etnia/etnicidade (AFSHARI e BHOPAL, 2010), que encerra em si a complexidade que abrange a designação de um povo. Nesse caso é levada em conta a combinação de fatores culturais e outros, incluindo língua, alimentação, religião, ascendência e as características físicas tradicionalmente associadas à *raça* (BHOPAL, 2004). De fato, a classificação por *raça* utilizada ainda em muitos estudos de população – seja por autodeclaração ou análise de traços como cor da pele, formato do nariz e lábios – apresenta uma série de desvantagens e viés que podem levar a resultados espúrios, sobretudo em países altamente miscigenados, como o Brasil (ZEMBRZUSKI *et al.* 2006).

Parte desse viés diz respeito à subjetividade e discordâncias que envolvem as definições do conceito de *raça* e a percepção de si mesmo dentro das categorias definidas. Um quadro que exemplifica bem essa questão foi observado por Bastos *et al.* (2009) que ao compararem a influência de classificação étnica entre entrevistador/entrevistado na cidade de Pelotas-RS, constataram que indivíduos a partir dos 40 anos tiveram uma chance 2,1 vezes maior de se classificarem como pardos do que como brancos, quando abordados por entrevistadoras negras, em relação às brancas. Dentro da mesma faixa etária, os entrevistados apresentaram 2,5 vezes

menos chances de se classificarem como pretos do que como brancos, ao serem abordados por entrevistadoras negras. Já as entrevistadoras negras, em comparação com as brancas, tiveram chance 2,5 vezes menor de classificar homens a partir dos 40 anos como pretos do que como brancos.

Outra desvantagem da classificação por traços fenotípicos é o pequeno número de genes envolvidos na definição dos mesmos, como mencionado anteriormente. Além disso, esses grupos de genes estão sujeitos à pressão local de seleção e poucos deles são comuns à população mundial (BAMSHAD *et al.* 2004; McEVOY *et al.* 2006; STURM, 2009). Em determinados grupos étnicos, certo conjunto de genes está envolvido na definição da cor da pele, por exemplo, e em outro grupo um conjunto diferente vai atuar sobre a mesma característica, exibindo nesse universo apenas poucos genes em comum. De um conjunto de onze genes candidatos (TYR, TYRP1, OCA2, SLC45A2, SLC24A5, MC1R, ASIP, KITLG, SLC24A4, IRF4, TPCN2) observado por Sturm (2009), apenas três (TYR, SLC45A2 e SLC24A5) explicariam as diferenças no tom da pele entre sul asiáticos, por exemplo, e enquanto um grande conjunto de genes estaria envolvido na relação entre cor de pele e cabelo, um único SNP seria responsável por regular a expressão do gene OCA2 para cor dos olhos.

Não obstante, pesquisas em áreas como saúde, genética de populações, antropologia, forense e outras precisam de mecanismos confiáveis para agrupar pessoas, além de averiguar qual a contribuição genética ancestral ou provável origem geográfica de dadas populações ou indivíduos e sua dinâmica. Nesse caso, sobretudo em países miscigenados como o Brasil, é aconselhável utilizar-se de marcadores moleculares imparciais, de fácil genotipagem, abundantes no genoma e população-específicos (SHRIVER *et al.* 1997; PARRA *et al.* 1998; SHRIVER *et al.* 2003; ZEMBRZUSKI *et al.* 2006; LUIZON *et al.* 2008). É o caso dos MIAs - Marcadores Informativos de Ancestralidade, ou no inglês, *AIMs – Ancestry Informative Markers*. Inicialmente denominados *PSAs - Population-specific alleles*, por Shriver *et al.* 1997, estes podem ser, virtualmente, qualquer região polimórfica do genoma que atinja diferencial de frequência entre populações parentais igual ou maior a 30% ($\delta \geq 30\%$) (BONILLA, 2004).

No presente estudo foram utilizados dois tipos de marcadores: (i) inserções *Alu* e (ii) *indels*.

O nome inserção *Alu* (i) vem das primeiras tentativas de caracterizar a sequência com utilização da nuclease de restrição *AluI*. São da família de elementos transponíveis repetitivos dispersos pelo genoma, mais especificamente os elementos nucleares intercalares curtos, ou SINEs (do inglês *short interspersed nuclear elements*). Fazem parte da linhagem não-viral, ou seja, são incapazes de codificar a transcriptase reversa. Apesar disso é o elemento mais abundante na família, alcançando cerca de um milhão de cópias. A repetição ocorre cerca de uma vez a cada 3 kb no genoma, e inclui exemplos transcritos (STRACHAN, 2002).

Caso duas pessoas compartilhem a mesma sequência *Alu* no mesmo local do seu genoma, elas descendem de um antepassado comum que lhes forneceu aquele segmento específico de DNA (BATZER e DEININGER, 2002). Isto porque é improvável que o evento de inserção ocorra duas vezes em um mesmo *locus*. A recente integração dos elementos *Alu* serve como marco temporal na evolução do genoma, e muitos deles provaram ser úteis no estudo da evolução humana e no estudo da história natural de diferentes regiões do genoma (CARROL *et al.* 2001). Uma vez que possui frequências diferenciadas por população, podem ser usados para estimar ancestralidade biogeográfica em nível de população, subgrupo e individual. Estas inserções podem ser excelentes marcadores étnicos, de grande uso para estimar a composição de populações híbridas, como a brasileira. (MENDES-JR. e SIMÕES, 2001; SHRIVER *et al.* 2003; RAY *et al.* 2005).

O termo *indel* (ii) é uma contração das palavras inserção/deleção e pode ser definido como uma interrupção no alinhamento da sequência de DNA causada por uma mutação de inserção ou de eliminação (BARTHE *et al.* 2012). Essas sequências normalmente possuem um tamanho reduzido de pares de base e podem ser formadas a partir da inserção de transposons e outros elementos móveis, por deslizamento da DNA polimerase na replicação de sequências simples ou por *crossing-over* desigual entre cópias repetidas similares, tais como sequências *Alu* (BRITTEN *et al.* 2003). A

comparação entre regiões genômicas de organismos proximalmente relacionados, como humanos e chimpanzés, demonstrou que essas sequências dominam o processo de divergência desde o início, indicando que são fontes significativas de mudança evolutiva. As *indels* estão ainda relacionadas com doenças humanas, principalmente neurodegenerativas (BRITTEN *et al.* 2003). Esses elementos começaram a ser utilizados como marcadores posteriormente às *Simple Sequence Repeats* (SSRs) e *Single Nucleotide Polymorphism* (SNPs) e apresentam vantagens como padrão mais homogêneo de mutação, genotipagem mais fácil e barata. Dessa forma são uma importante fonte de marcadores genéticos para estudos de populações naturais (VÄLI *et al.* 2008).

FORMAÇÃO DA POPULAÇÃO BRASILEIRA

A América do Sul já era povoada cerca de 12 mil anos antes da chegada dos primeiros colonizadores europeus (CUNHA, 1998), sendo estimada em 2,5 milhões a população ameríndia no território brasileiro no início da colonização (IBGE, 2012). À época do “achamento”, aproximadamente 500 mil portugueses, em sua maioria homens, desembarcaram nessas terras (RIBEIRO, 1995). O contato entre essas civilizações envolveu extremos que vão da ajuda indígena na construção das primeiras vilas e defesa do território, à escravização e disseminação direta ou indireta (doenças) pelos colonos (TAVARES, 2001; FUNAI, 2012). Apesar dos conflitos, cruzamentos entre portugueses e ameríndias ocorriam amplamente, em parte pela escassez de mulheres europeias, mas também por incentivo da Coroa como estratégia para facilitar o povoamento (PENA *et al.* 2000).

Os africanos foram levados ao Brasil, enquanto escravos, durante a colonização. Entre 1502 e 1860, mais de nove milhões e meio de africanos foram transportados para as Américas, 38% deles para o Brasil (MATTOSE, 1992). Os principais destinos eram as cidades de maior importância da época, Salvador, Recife e Rio de Janeiro. A mão de obra escrava era aplicada nas lavouras e, posteriormente, na exploração das minas de ouro e diamante (ALVES-SILVA *et al.* 2000).

A sobreposição destes povos gerou as bases da população brasileira, sendo por esse motivo considerada tri-híbrida. Outros eventos marcantes de imigração para o Brasil ocorreram a partir do século XIX, com a abertura dos portos as nações amigas. Entre 1872 e 1975, aproximadamente 5,5 milhões de imigrantes desembarcaram no país, vindos principalmente da Europa (Itália, Portugal, Espanha e Alemanha) – 81%, Japão – 5%, Líbano e Síria – 2% e outros – 12% (IBGE, 2012). Esse fenômeno é conhecido como “embranquecimento do Brasil” (SEYFERTH, 1985; SANTOS, 2002; PENA *et al.* 2009).

O somatório dos eventos de migração cunhou uma população extremamente heterogênea. Entretanto, razões históricas, geográficas, políticas e sócio-culturais fizeram com que o povoamento do território fosse desigual, marcando cada região com diferentes contribuições ameríndia, européia e africana. De fato, o reflexo dessa dinâmica pode ser percebido ainda atualmente. Analisando uma amostra de indivíduos das cinco regiões sócio-geográficas brasileiras mais o Distrito Federal, Callegari-Jacques *et al.* 2003, observou uma pequena, porém significativa diferença genética entre as populações.

O povoamento da Bahia se insere nesse contexto. Um quadro pertinente da distribuição dos primeiros habitantes no estado seria: por toda costa *tupinambás* e *tupiniquins*; no centro do estado os *maracás* e *paiaiás*; ao sul os *aimorés*, *camacãs* e *pataxós*; ao norte do estado *kariris*; no extremo oeste *acroás* e *anaiós* (TAVARES, 2001). Não obstante a diversidade e importância desses povos, muito pouco se sabe sobre sua estrutura e costumes originais e atualmente apenas 0,4% da população baiana se auto-declara indígena (IBGE, 2012). Isso porque os impactos da colonização sobre estas culturas foi devastador. Um fato decisivo foi a inadequação do índio enquanto escravo devido a fatores como baixa densidade das populações indígenas e ao próprio agravamento da crise entre colonizadores e nativos, além de o tráfico negreiro ser mais vantajoso economicamente à Coroa (PINSKY, 2000).

Assim sendo, a influência dos africanos na Bahia tem papel de destaque. Devido à importância do porto de Salvador durante todo o tráfico negreiro, a vinda

destes à Bahia se confunde com a história dos mesmos no Brasil (TAVARES, 2001). De acordo com dados da época, como o censo de 1823, o contingente de negros e mestiços suplantava numericamente os europeus presentes no país. Atualmente cerca de 75% da população do estado é composta de negros ou pardos (GATTÁS *et al.* 2004).

O povoamento da região sul da Bahia, em particular, é bastante antigo e remete a época das capitanias hereditárias. A vila de Ilhéus foi fundada entre 1535-6, como sede administrativa da capitania. Essa experimentou um breve período de prosperidade, decaindo por desentendimentos dos colonos com o administrador e por conflitos contra os *tupinambás* e *aimorés*. Como a introdução de escravos estava diretamente relacionada ao desenvolvimento econômico, durante essa época o fluxo de africanos para a região foi significativamente reduzido. Por outro lado, os índios vencidos e/ou catequizados pelas missões jesuíticas, além dos tupiniquins, eram inseridos – não entrando aqui no mérito das relações sociais – na sociedade e somavam, incluindo os mestiços, maior número que os portugueses. No século XIX a região recebeu um fluxo de franceses, colonos alemães e suíços, bem como irlandeses (CAMPOS, 2006). Já no início do século XX, a cultura de cacau atraía trabalhadores e profissionais de diversas partes do país, além de estrangeiros. Atualmente a região sul da Bahia tem um relevante apelo turístico, recebendo visitantes brasileiros e estrangeiros. Esses episódios podem fazer com que a região tenha uma composição étnica diferente daquela observada em outras partes do estado.

Embora o conceito de *raça* não se aplique do ponto de vista biológico (BAMSHAD e OLSON, 2004; HUNLEY *et al.* 2009), como já discutido anteriormente, percebe-se no Brasil uma relação entre cor da pele e desigualdades sociais, que se estende a campos como educação e saúde (HASENBALG, 1995). Significa dizer que designação por características físicas é um assunto simultaneamente bastante controverso e relevante, inclusive no que tange à genética.

No Brasil, fica claro que, devido à intensa miscigenação, qualquer caracterização física individual é arriscada (GATTÁS *et al.* 2004; PIMENTA *et al.*

2006) e mesmo a cor da pele é um baixo indicativo de ancestralidade africana (PARRA *et al.* 2003). Dessa forma, a utilização de marcadores de ancestralidade é fundamental para se obter uma caracterização confiável seja em nível individual ou populacional (SHRIVER *et al.* 2003).

É importante, portanto, conhecer a contribuição genética dos principais grupos étnicos formadores em populações do sul da Bahia com o uso de AIMS, visto que há poucas informações a esse respeito e pela grande relevância desse conhecimento do ponto de vista genético, forense, social e médico.

METODOLOGIA

AMOSTRAGEM

Foram selecionadas aleatoriamente 800 amostras de mulheres, sendo 400 colhidas em Ilhéus e 400 em Itabuna, utilizando o *software* SPSS 10.0. Estas fazem parte de um banco de cerca de 3.000 mulheres criado no projeto coordenado pela Prof^a Sandra Rocha Gadelha. Esse material foi coletado de gestantes na maternidade Santa Helena em Ilhéus e Hospital e Maternidade Manoel Novaes em Itabuna, Bahia. É importante ressaltar que essas unidades são referências na região, sobretudo por serem a única de Ilhéus e principal de Itabuna a atender pelo Sistema Único de Saúde (SUS) e a dispor de Unidade de Terapia Intensiva (UTI) neonatal. Dessa forma, representam bem a região sul, recebendo pacientes de cidades circunvizinhas como Una, Canavieiras, Ibicaraí, entre outras.

ASPECTOS ÉTICOS

O presente estudo ajuda a responder questões relevantes ao projeto da Prof^a Sandra Rocha Gadelha intitulado “Soroprevalência de agentes de transmissão materno-fetal (vírus linfotrópico de células T humanas (HTLV), citomegalovírus (CMV), toxoplasma gondii e rubéola), subtipagem do HTLV e avaliação da transmissão materno-fetal em gestantes do sul da Bahia”. Dessa forma, após análise, a caracterização genética da população foi entendida pelo Comitê de Ética em Pesquisa da UESC CEP-UESC como adendo do referido projeto, aprovado sob número de protocolo 196/08. Todos os sujeitos de pesquisa assinaram, previamente à coleta de sangue, um Termo de Consentimento Livre e Esclarecido – TCLE (Anexo II).

EXTRAÇÃO DE DNA

O DNA foi extraído a partir de sangue total, armazenado em freezer -24°C . Foram utilizados kits QIAGEN de purificação por coluna, modelo QIAamp, procedendo segundo instruções do fabricante.

Devido à má qualidade de algumas amostras, determinadas extrações não puderam ser concluídas ou obtiveram resultado insatisfatório e, portanto, foram excluídas. Dessa forma, a amostra final foi reduzida para 750 mulheres, sendo 372 coletadas em Ilhéus e 378 em Itabuna.

MARCADORES MOLECULARES SELECIONADOS

Neste trabalho foram analisados sete AIMs, três (AT3, Sb19.3 e APO) com $\delta > 40\%$ entre africanos e europeus ou ameríndios; dois (MID-52 e MID-575) com $\delta > 45\%$ entre ameríndios e africanos ou europeus; um (MID-93) com $\delta > 50\%$ entre europeus e africanos ou ameríndios e um (PV92) com $\delta > 40\%$ entre ameríndios e africanos ou europeus (PARRA *et al.* 1998, SHRIVER *et al.* 2003).

Para os *loci* analisados no presente estudo, foi denominado de alelo *1 (ex.: APO * 1) o fragmento correspondente à banda de maior peso molecular observada no gel agarose, devido à presença de uma inserção, seguindo a convenção estabelecida por Parra *et al.* (1998). Os *primers* e a localização citogenética dos *loci* analisados no presente estudo estão relacionados na Tabela 1.

Tabela 1: Localização citogenética e sequência dos *primers* dos quatro AIMs analisados.

<i>Locus</i>	Tipo	LOCALIZAÇÃO	SEQUÊNCIA DOS <i>PRIMERS</i>
AT3	<i>indel</i>	1q25.1	CCACAGGTGTAACATTGTGT GAGATAGTGTGATCTGAGGC
Sb19.3	Inserção <i>Alu</i>	19p12	TCTAGCCCCAGATTTATGGTAACTG AAGCACAAATTGGTTATTTTCTGAC
APO	Inserção <i>Alu</i>	11q23	AAGTGCTGTAGGCCATTTAGATTAG AGTCTTCGATGACAGCGTATACAGA
PV92	Inserção <i>Alu</i>	16q23.3	AACTGGGAAAATTTGAAGAGAAAGT TGAGTTCTCAACTCCTGTGTGTTAG
MID-93	<i>indel</i>	22q13.2	CGAGTCAAATCTGGGGTGT TGCTGAGATTACAGGCGTGA
MID-52	<i>indel</i>	4q24	TGGAAAATTGGTTGCAAATTC TTGTTCCAGTTGTTGGTGTGA
MID-575	<i>indel</i>	1p34.3	ACCACAGGCAAATCCTGAAG CAAGGGTTAGGGAGGTTGGT

REAÇÃO EM CADEIA DA POLIMERASE (PCR)

Os ensaios da PCR foram realizados em um volume total de 25 μ L. Todos os reagentes (água ultrapura, tampão, dNTP – solução trabalho, *primers* – solução trabalho, MgCl₂ e Taq DNA polimerase), com exceção do DNA, foram misturados em quantidade específica para cada *locus* em um único tubo (mix) para garantir a homogeneidade das reações. No caso dos *loci* MID-93, MID-52 e MID-575, um conjunto de três pares de *primers* foi adicionado em um mesmo tubo (multiplex). O mix foi então aliquotado em microtubos individuais de 200 μ L, onde em um ambiente separado aplicou-se 4 μ L de DNA equivalente as amostras (Tabela 2).

Tabela 2. Condições de PCR para os *loci* analisados no presente trabalho. Quantidade em μL suficiente para uma reação.

<i>Locus</i>	Água	Tampão PCR livre de cloreto	MgCl_2 (50mM)	dNTP (20mM)	Primers (2,5mM)	Taq 5U
Sb19.3	14,2	2,5	1,0	0,25	3,0	0,05
APO	14,2	2,5	1,0	0,25	3,0	0,05
PV92	14,2	2,5	1,0	0,25	3,0	0,05
AT3	14,2	2,5	1,0	0,25	3,0	0,05
MID-93					2,0	
MID-52	9,75	2,5	1,0	0,25	3,0	0,05
MID-575					2,0	

O termociclador utilizado foi o Biocycler. Para cada análise foi usado controle negativo contendo água no lugar do DNA genômico. Após o término da reação de PCR o produto foi guardado em freezer comum até sua utilização.

ANÁLISE DO PRODUTO AMPLIFICADO

Os *loci* MID-93, MID-52 e MID-575 foram genotipados no laboratório coordenado pelo Prof. Aguinaldo Luiz Simões, da Universidade de São Paulo (USP), campus de Ribeirão Preto. A eletroforese foi realizada em gel desnaturante de poli-acrilamida a 12% por 03h40min e então corado com nitrato de prata (AgNO_3).

Os produtos dos demais marcadores foram separados em gel de agarose a 2% (Sb19.2, PV92 e APO) e 3% (AT3), segundo protocolo abaixo.

Reagentes e soluções para os géis de Agarose:

Agarose: AGR-LE-500 (Uniscience)

Tampão TBE 1X: 100 mL do tampão TBE (10X); 900mL de H_2O .

GelRed: GelRed 10.000x *in Water*.

GelRed 20X: 998 μL de H_2O ; 2 μL de GelRed 10.000X

Azul de Bromofenol (corante tipo IV): 8g de sacarose; 0,05g de bromofenol; 20 mL de água destilada.

Tampão da amostra: 2 μL de GelRed 20X; 1 μL de azul de bromofenol, para cada amostra.

Gel a 2%: 2g de agarose; 100ml de TBE 1X; 1 μL de GelRed 10X.

Gel a 3%: 3g de agarose; 100ml de TBE 1X; 1 μL de GelRed 10X.

Procedimento:

Os géis a 2 e 3% foram preparados como no protocolo acima determinado. A mistura de agarose e tampão foi levada ao microondas por aproximadamente 1 minuto, até total dissolução. O GelRed foi adicionado à mistura do gel imediatamente antes de vertê-la em um cassete previamente montado, composto de uma bandeja de 100 mL, um suporte para gel e pentes colocados na borda superior e no meio. Aguardou-se a polimerização por cerca de 15 minutos. Nos poços formados pelos pentes foram aplicadas as amostras de DNA amplificado por PCR

Após a polimerização do gel, o suporte foi retirado da bandeja, seguido do pente. O gel polimerizado foi montado em cuba de eletroforese horizontal contendo tampão TBE (1X). Esta cuba foi conectada a uma fonte de eletroforese e ajustada à voltagem constante de 100V, necessária para uma separação dos fragmentos amplificados.

Aproximadamente 7 μL de DNA amplificado oriundo da PCR, juntamente com 3 μL de tampão amostra, foram aplicados nos géis e o sistema ligado por 30 a 45 minutos. Com o término da corrida eletroforética, os géis foram retirados cuidadosamente da cuba e submetidos aos procedimentos de documentação através da captura de imagem realizada pelo fotodocumentador L-PIX ST (Loccus Biotecnologia).

ANÁLISES DOS DADOS

Frequências Alélicas

As frequências alélicas para cada *locus* em cada amostra foram estimadas por contagem direta, utilizando-se o programa FSTAT versão 2.8 (GOUDET, 1995). Representando-se a frequência do genótipo A_iA_j por X_{ij} , a frequência do alelo “i” é dada por:

$$x_i = X_{ii} + \frac{1}{2} \sum_{j \neq i} X_{ij}$$

onde $\sum_{j \neq i}$ indica a somatória de X_{ij} (frequência de todos os genótipos que apresentam o alelo “i”), exceto quando $i = j$.

A partir de dados de frequências alélicas disponíveis na literatura foi elaborada uma tabela de frequências dos AIMS estudados para populações da África, Europa e Ameríndios (Tabela 2). Conforme exposto anteriormente, estão listados somente as frequências do alelo *1 para cada *locus*, de acordo com a convenção descrita em Parra *et al.* (1998).

Os dados referentes a valores mundiais de frequência e diferencial de frequência das populações parentais foram obtidos a partir da média ponderada daqueles observados na literatura. No caso da população europeia (EUR), foram usados valores da Inglaterra, Irlanda, Alemanha (PARRA *et al.* 1998), Portugal (TOMÁS *et al.* 2002) e Espanha (dbSNP/NCBI). No caso da população africana (AFR) foram usados valores da República Africana, Nigéria (Benin e Ibadan) (PARRA *et al.* 1998) e Serra Leoa (Mende e Temne) (PARRA *et al.* 2001). Para a população ameríndia (AMR) foram utilizados valores de tribos indígenas da Amazônia brasileira (LUIZON *et al.* 2008). Os MIDs (93, 52 e 575) Os cálculos foram feitos no programa Microsoft Excel®. As populações foram escolhidas buscando manter proximidade com dados históricos de colonização e migração.

Aderência ao Equilíbrio de Hardy-Weinberg

A aderência da amostra ao equilíbrio de Hardy-Weinberg foi avaliada pelo teste exato de Fisher, assumindo-se independência entre os *loci* utilizados. O teste *multi-locus* analisa as frequências genotípicas de todos os *loci* em uma única população e o teste multi-populacional analisa as frequências genotípicas de um único *locus* em todas as populações. Os dados foram analisados no GENEPOP 2.0 (RAYMOND e ROUSSET, 1995), e a hipótese nula (H_0) é de que a união dos gametas ocorra de maneira aleatória.

Foram também aplicados testes estatísticos com uma maior sensibilidade ao excesso ou à deficiência de heterozigotos. Nestes dois casos, a hipótese nula é a união aleatória dos gametas e a hipótese alternativa (H_1), o excesso ou deficiência de heterozigotos. Novamente, as análises foram realizadas no GENEPOP 2.0 (RAYMOND e ROUSSET, 1995).

Associações par-a-par entre *loci*

A análise de associações par-a-par entre *loci* não ligados foi realizada utilizando-se o programa GENEPOP 2.0 (RAYMOND e ROUSSET, 1995). A hipótese nula é a de que a distribuição genotípica em um *locus* é independente da distribuição em outro *locus*. Esta análise foi aplicada para verificar desvios do esperado pela regra de multiplicação entre pares de *loci* localizados em diferentes cromossomos. A palavra “ligação” neste caso não está relacionada com associação física entre alelos de *loci* de um mesmo cromossomo. De fato, qualquer associação encontrada entre as frequências gênicas de *loci* diferentes pode ser denominada como desequilíbrio de ligação, ainda que as associações não sejam decorrentes de ligação física entre eles (WEIR, 1996).

Diferenciação e diversidade genética das populações

Os testes exatos para diferenciação populacional foram realizados com o uso do programa GENEPOP 2.0 (RAYMOND e ROUSSET, 1995). Este utiliza tabelas de contingência RxC geradas automaticamente para cada *locus*, em que R é o número de populações e C é o número de alelos no *locus*.

Este procedimento compara cada *locus* em pares de populações, para determinar se existem diferenças nas frequências alélicas observadas. Nesse caso, a hipótese nula testada é a de que a distribuição alélica é idêntica entre as populações (RAYMOND e ROUSSET, 1995).

O coeficiente de endogamia (F_{IS}) e o índice de fixação (F_{ST}) foram calculados pelo *software* FSTAT 2.8. As estimativas por *locus* foram rodadas por ANOVA padrão, segundo Weir e Cockerham (1984).

Estimativas de mistura étnica

As estimativas das proporções étnicas foram obtidas segundo o método de identidade gênica (CHAKRABORTY, 1985) e foram realizadas com o uso do programa ADMIX 3 e 2. O ajuste a este modelo é avaliado pelo coeficiente de correlação múltipla (R^2) entre as frequências alélicas nas populações híbridas e aquelas das populações ancestrais (CHAKRABORTY, 1986).

Enquanto o ADMIX 2 pressupõe um modelo populacional di-híbrido, no qual são inseridas frequências alélicas de duas populações ancestrais, o ADMIX 3 pressupõe uma população formada por três parentais, ou seja, um modelo tri-híbrido. Em caso de inconsistência com algum dos modelos, o programa não roda. Contudo, caso ambos sejam aceitos pelo *software*, prevalece aquele com maior R^2 (coeficiente de adaptação ao modelo de mistura).

A utilização simultânea de diversos *loci* genéticos, no cálculo de mistura étnica, teve por objetivo minimizar os efeitos dos fatores limitantes inerentes ao método de

investigação, tais como desvios de amostragem, frequências ancestrais representadas por médias ponderadas e forças microevolutivas atuando junto com a mistura (BORTOLINI *et al.* 1992).

RESULTADOS

PERFIL ELETROFORÉTICO

APO

O padrão de bandas observado para o *locus* APO é produzido pela inserção de uma sequência *Alu* com 290 pares de base (pb) na região não codificadora próxima ao complexo de genes da apolipoproteína AI-CIII-AIV, braço longo do cromossomo 11 (BATZER *et al.* 1991). Como observado na Figura 1, indivíduos homozigotos para ausência da inserção (1) produzem apenas fragmentos de 110 pb após amplificação, enquanto indivíduos homozigotos para presença da inserção (2 e 4) produzem apenas fragmentos de 400 pb. Heterozigotos (3) apresentam fragmentos de ambos tamanhos. Para o *locus* APO, todos os alelos foram observados no presente estudo.

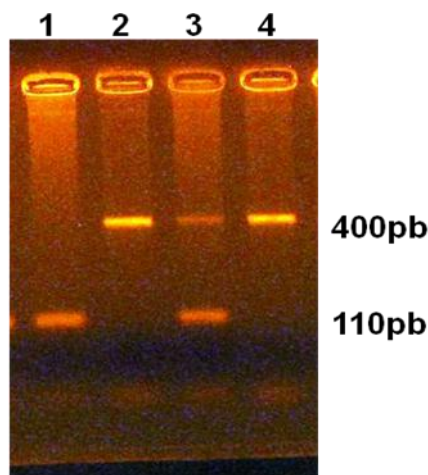


Figura 1. Padrão eletroforético do *locus* APO em agarose 2% corado com GelRed, indicando número das raias e tamanho dos fragmentos gerados.

SB19.3

O padrão de bandas observado para o *locus* Sb19.3 é produzido pela inserção de uma sequência *Alu* com 311 pb no cromossomo 19p12 (ARCOT *et al.* 1997). Como

observado na Figura 2, indivíduos homozigotos para ausência da inserção (2) produzem apenas fragmentos de 145 pb após amplificação, enquanto indivíduos homozigotos para presença da inserção (4) produzem apenas fragmentos de 456 pb. Os heterozigotos (1 e 3) apresentam fragmentos de ambos tamanhos. Para o *locus* Sb19.3, todos os alelos foram observados no presente estudo.

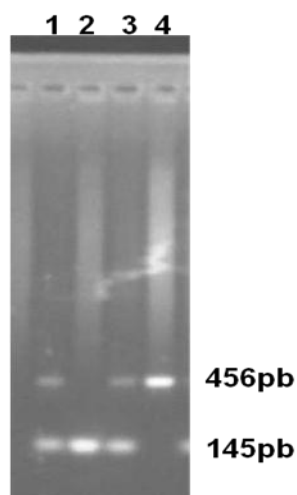


Figura 2. Padrão eletroforético do *locus* Sb19.3 em agarose 1% corado com GelRed, indicando número das raias e tamanho dos fragmentos gerados.

PV92

O padrão de bandas observado para o *locus* PV92 é produzido pela inserção de uma sequência *Alu* com 290 pb no cromossomo 16 (BATZER *et al.* 1991). Como observado na Figura 3, indivíduos homozigotos para ausência da inserção (2) produzem apenas fragmentos de 110 pb após amplificação, enquanto indivíduos homozigotos para presença da inserção (1) produzem apenas fragmentos de 400 pb. Heterozigotos (3 e 4) apresentam fragmentos de ambos tamanhos. Para o *locus* PV92, todos os alelos foram observados no presente estudo.

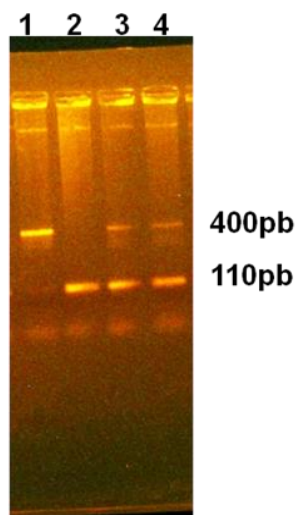


Figura 3. Padrão eletroforético do *locus* PV92 em agarose 2% corado com GelRed, indicando número das raias e tamanho dos fragmentos gerados.

AT3

O padrão de bandas observado para o *locus* AT3 é produzido pela inserção ou deleção (*indel*) de uma sequência 76 pb no exón 1 do gene da antitrombina III. Como observado na Figura 4, indivíduos homocigotos para ausência da inserção (5) produzem apenas fragmentos de 496 pb após amplificação, enquanto indivíduos homocigotos para presença da inserção (3) produzem apenas fragmentos de 572 pb. Heterocigotos (4) apresentam fragmentos de ambos tamanhos. Para o *locus* AT3, todos os alelos foram observados no presente estudo.

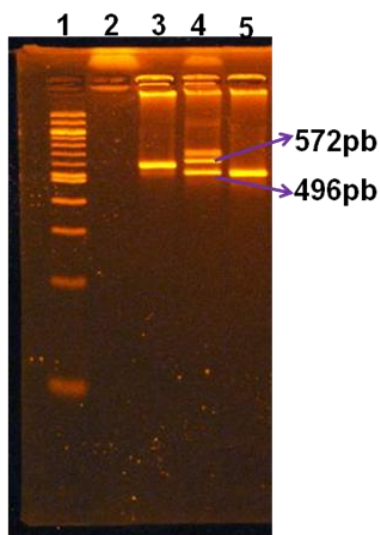


Figura 4. Padrão eletroforético do *locus* PV92 em agarose 3% corado com GelRed, indicando número das raias e tamanho dos fragmentos gerados. As raias 1 e 2 contêm, respectivamente, *ladder* de 100 pb e controle negativo da reação.

MIDs

O padrão de bandas observado no multiplex dos *loci* MID-93, MID-52 e MID-575 é produzido pela inserção ou deleção de sequências com aproximadamente 3 pb na região amplificada pelos pares de *primers*. Cada *locus* pode apresentar dois alelos com tamanhos distintos, sendo estes respectivamente 240 ou 243 pb, 226 ou 229 pb e 212 ou 215 pb (SHRIVER *et al.* 2003). Todos os alelos foram observados no presente estudo.

FREQUÊNCIAS ALÉLICAS

Seguindo convenção de Parra *et al.* 1998, para comparação e diferencial de frequências alélicas será considerado apenas o alelo *1. As frequências nas populações parentais bem como os valores encontrados nas amostras de Ilhéus (IOS) e Itabuna (ITB) seguidos da média, estão dispostos na Tabela 2.

Tabela 3. Frequências alélicas dos quatro *loci* avaliados nas amostras de Ilhéus (IOS) e Itabuna (ITB) seguido das médias, comparadas aos valores encontrados em africanos (AFR), europeus (EUR) e ameríndios (AMR).

População	Frequências alélicas					
	Populações mundiais			Sul da Bahia		
Alelo	AFR	EUR	AMR	IOS	ITB	Média
Sb19.3 *I	0,42	0,93	0,71	0,64	0,66	0,65
PV92 *I	0,22	0,15	0,92	0,25	0,27	0,26
APO *I	0,41	0,94	1,00	0,81	0,80	0,80
AT3 *I	0,85	0,26	0,05	0,50	0,44	0,47
MID-93 *I	0,26	0,81	0,08	0,49	0,54	0,51
MID-52 *I	0,36	0,08	0,76	0,77	0,79	0,78
MID-575 *I	0,12	0,01	0,58	0,89	0,82	0,88

O MID-575 e o PV92 são marcadores freqüentes em ameríndios (SHRIVER *et al.* 2003; LUIZON *et al.* 2008), porém no presentes estudo apresentaram resultados controversos. Enquanto o primeiro foi o mais freqüente na população estudada (0,88) o segundo foi o menos freqüente (0,26).

O diferencial de frequência (δ) entre as amostras analisadas e os grupos mundiais estão apresentados na Tabela 3. Os valores significativos ($> 0,35$) estão destacados em negrito. Dentre os sete marcadores utilizados, apenas um (Sb19.3) não atingiu δ significativo quando comparadas quaisquer populações. Os demais (PV92, APO, AT3, MID-93, MID-52 e MID-575) se mostraram indicativos de ancestralidade.

Tabela 4. Diferencial de frequência (δ) entre a média dos valores encontrados no presente estudo (SBA) e valores encontrados para africanos (AFR), europeus (EUR) e ameríndios (AMR).

Pop.	Diferencial de Frequência (δ)		
	AFR/ SBA	EUR/ SBA	AMR/ SBA
<i>Loci</i>			
Sb19.3	0,23	0,28	0,06
PV92	0,04	0,11	0,66
APO	0,39	0,14	0,20
AT3	0,38	0,21	0,42
MID-93	0,23	0,32	0,41
MID-52	0,41	0,69	0,01
MID-575	0,76	0,87	0,3

ADERÊNCIA AO EQUILÍBRIO DE HARDY-WEINBERG

As frequências genotípicas observadas nos quatro *loci* foram comparadas aos valores esperados segundo o equilíbrio de Hardy-Weinberg pelo teste exato G. Os valores de p encontrados estão dispostos na Tabela 4. Os *loci* Sb19.3 e APO estão em desvio ao esperado pelo EHW, tanto na amostra de Ilhéus quanto de Itabuna. Os demais resultados não foram significativos, de forma que aderem ao esperado pelo EHW.

Testes mais sensíveis foram utilizados para detectar se os desvios observados ocorreram pelo déficit ou pelo excesso de heterozigotos. Foi detectado então que os desvios observados nos *loci* Sb19.3 e APO são causados por déficit de heterozigotos ($p = 0$) nas amostras.

Tabela 5. Estimativas de p -valor segundo teste exato para verificação de aderência ao equilíbrio de Hardy-Weinberg (Guo e Thompson, 1992). Valores significativos ($p < 0,05$) estão destacados em negrito.

Aderência ao EHW ¹		
Amostra	Ilhéus	Itabuna
<i>Locus</i>	<i>p</i> -valor	<i>p</i> -valor
Sb19.3	0,000	0,000
PV92	0,177	0,793
APO	0,000	0,000
AT3	0,145	0,565
MID-93	0,328	0,102
MID-52	0,728	0,488
MID-575	0,001	0,347

¹ Equilíbrio de Hardy-Weinberg

ASSOCIAÇÕES PAR-A-PAR ENTRE *LOCI*

O desequilíbrio genotípico foi avaliado em um total de 420 permutas. Foi observado desequilíbrio de ligação entre os *loci* Sb19.3 x APO ($p = 0,038$) e APO x AT3 ($p = 0,041$) (Tabela 05).

Tabela 6. Comparação par-a-par entre *loci* não ligados em amostra do sul da Bahia.

Desequilíbrio de ligação	
<i>Loci</i>	<i>p</i> -valor
Sb19.3 X PV92	0,657
Sb19.3 X APO	0,038
Sb19.3 X AT3	0,214
Sb19.3 X MID-93	0,121
Sb19.3 X MID-52	0,595
Sb19.3 X MID-575	0,112
PV92 X APO	0,341
PV92 X AT3	0,693
PV92 X MID-93	0,443
PV92 X MID-52	0,431
PV92 X MID-575	0,183
APO X AT3	0,041
APO X MID-93	0,705
APO X MID-52	0,541
APO X MID-575	0,521
AT3 X MID-93	0,8177
AT3 X MID-52	0,888
AT3 X MID-575	0,945
MID-93 X MID-52	0,326
MID-93 X MID-575	0,433
MID-52 X MID-575	0,491

DIFERENCIAÇÃO GENÉTICA DAS AMOSTRAS

A diferenciação gênica e genotípica foi avaliada pelo teste exato G. Enquanto o poder do conjunto de marcadores foi avaliado pelo método de Fisher, o poder de cada AIM separar as subamostras (Ilhéus/Itabuna) na população do sul da Bahia foi analisado pelo processo das cadeias de Markov. No presente estudo, nenhum marcador isoladamente ou em conjunto foi capaz de diferenciar as subamostras do sul da Bahia (Tabela 6).

Para a interpretação dos valores de F_{ST} foi sugerido que valores entre 0 - 0,05 indicam pouca diferenciação, de 0,05 a 0,15 uma diferenciação moderada, de 0,15 a 0,25 uma grande diferenciação, e acima deste valor uma diferenciação muito forte (WRIGHT, 1978). Novamente não foi observada diferenciação entre as subamostras ($F_{ST} = 0,00$).

Tabela 7. Avaliação da diferenciação gênica e genotípica das subamostras de Ilhéus e Itabuna na população do sul da Bahia através do teste exato G.

<i>Locus</i>	Diferenciação genotípica	Diferenciação gênica
	<i>p</i> -valor	<i>p</i> -valor
Sb19.3	0,712	0,664
PV92	0,432	0,438
APO	0,690	0,648
AT3	0,063	0,063
MID-93	0,705	0,681
MID-52	1,000	1,000
MID-575	0,635	0,674
Todos os <i>loci</i>	0,375	0,354

DIVERSIDADE GENÉTICA

Foi analisada a diversidade de cada *locus* em cada subamostra (Tabela 7), além das estimativas de Nei (H_S e H_T), independentes do número de populações. A diversidade gênica variou de 0,20 no *locus* MID-575 a 0,51 no *locus* MID-93. A heterozigose média dentro das amostras (H_S) e a heterozigosidade total (H_T) foram de 0,39. O coeficiente de endogamia (F_{IS}) foi de 0,11.

Tabela 8. Diversidade gênica para cada *locus* em cada amostra. Destacados em negrito maior e menor valor.

<i>Locus</i>	Diversidade gênica	
	Ilhéus	Itabuna
Sb19.3	0,46	0,45
PV92	0,38	0,40
APO	0,31	0,32
AT3	0,50	0,49
MID-93	0,50	0,51
MID-52	0,35	0,35
MID-575	0,20	0,31

MISTURA ÉTNICA

Os valores percentuais das estimativas de mistura étnica foram muito semelhantes entre as subamostras de Ilhéus e Itabuna. Em ambos os casos foi identificada uma maior contribuição europeia (Tabela 8).

Tabela 9. Estimativa de mistura étnica na população do sul da Bahia, seguidas do respectivo erro padrão e do coeficiente de adaptação (R^2) ao teste aplicado.

Pop. Parentais	Componentes da mistura	
	Sul da Bahia	EP ¹
Africana	0,38	± 0,07
Européia	0,51	± 0,07
Ameríndia	0,11	± 0,05
R²	0,999	

¹ Erro padrão

Com base nos quatro AIMs utilizados, tanto as subamostras quanto a amostra total apresentaram valores consistentes com o modelo tri-híbrido. Em todos os casos a contribuição europeia foi superior, seguida da africana e ameríndia (Figura 5).

Para comparação da etnia autodeclarada foram utilizados dados do banco criado em projeto da Prof^a Sandra Rocha Gadelha, intitulado “Soroprevalência de agentes de transmissão materno-fetal (vírus linfotrópico de células T humanas (HTLV), citomegalovírus (CMV), toxoplasma gondii e rubéola), subtipagem do HTLV e avaliação da transmissão materno-fetal em gestantes do sul da Bahia”. Entretanto, o questionário continha cinco classes para designar etnia, a saber: branco, preto, pardo, amarelo e indígena. Para facilitar a comparação, as classes originais foram reagrupadas em três. Autodeclarados brancos foram classificados como eurodescendentes, em comparação a contribuição europeia (EUR). Autodeclarados pretos ou pardos foram classificados como afrodescendentes, em comparação a contribuição africana (AFR). Finalmente, autodeclarados amarelos ou indígenas foram agrupados na mesma classe, em comparação aos ameríndios (AMR) (Figura 5).

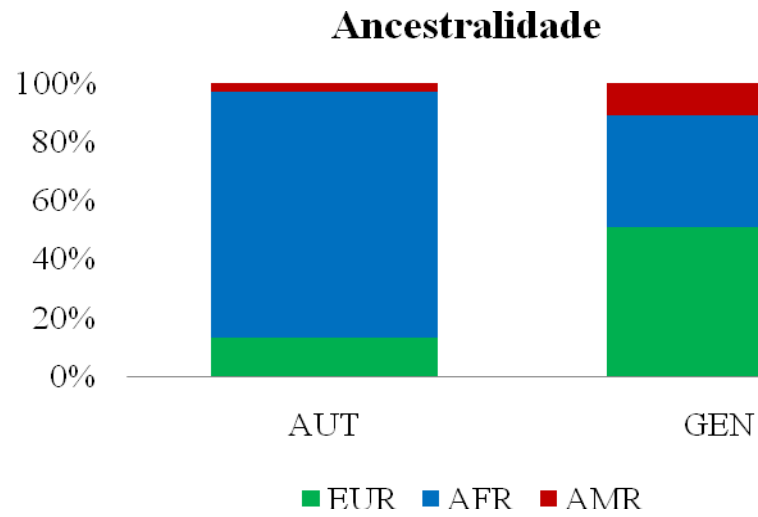


Figura 5. Representação gráfica da ancestralidade autodeclarada (AUT) e genética (GEN) em amostra do sul da Bahia. O modelo tri-híbrido é consistente com contribuição europeia (EUR), africana (AFR) e ameríndia (AMR).

DISCUSSÃO

CONTEXTUALIZAÇÃO HISTÓRICA

O Brasil é amplamente reconhecido como um país multicultural, a despeito das políticas sociais aplicadas nesse sentido (GRIN, 2006). Em parte a afirmação está embasada pelas proporções continentais do país, com ao menos cinco extensas regiões geopolíticas distintas – a saber: Sul, Sudeste, Centro-oeste, Nordeste e Norte. Também pela diversidade cultural entre ameríndios, europeus e africanos, as populações formadoras do Brasil. A região Nordeste, onde a Bahia está inserida, é historicamente conhecida pela influência africana. Os portos de Recife e Salvador foram portas de entrada para milhares de africanos trazidos como escravos da região oeste da África entre 1502 e 1860 (MATTOSO, 1992).

A distribuição desses africanos no território não foi homogênea, estando fortemente vinculada ao grau de desenvolvimento local. A ligação entre a quantidade de açúcar produzida em uma região, o número de engenhos e conseqüentemente escravos necessários para mão de obra era diretamente proporcional. Além disso, comprar e manter escravos exigia um alto custo. Sob este aspecto a região sul da Bahia foi desfavorecida (CAMPOS, 2006).

Após um breve período de prosperidade, a sede da capitania de Ilhéus possuía um número baixo de engenhos. Com uma produção açucareira menor, menos recursos e escravos chegavam àquela região. O número de colonos europeus também era comparativamente reduzido, afinal o interesse em ocupar a região diminuía em frente à tamanha adversidade. O pequeno contingente era insatisfatório para defender o território contra as investidas de reconquista das tribos indígenas, sobretudo aqueles conhecidos como *Aimorés*. De fato, a ofensiva indígena foi um grande empecilho à ocupação europeia na região sul da Bahia, principalmente porque os nativos suplantavam numericamente os colonos (CAMPOS, 2006).

A hostilidade entre colonizadores e nativos causou, além das inúmeras baixas entre ambos, o comprometimento de plantações e prédios. Um quadro semelhante a esse se estendeu entre meados do século XVI até meados do século XVIII, quando colonos alemães introduziram a cultura do cacau na região sul da Bahia (ANDRADE, 2003; CAMPOS, 2006).

Até esse período, os dados históricos mostram que a região sul da Bahia era fortemente influenciada pela presença de nativos indígenas, em comparação ao número de colonos europeus e escravos africanos. Contudo, a partir do sucesso da cultura cacauífera – iniciada na Bahia por alemães a partir de sementes trazidas da Amazônia – esse quadro mudou (ANDRADE, 2003; CAMPOS, 2006).

A experiência pioneira dos europeus se mostrou bastante lucrativa, de forma que indivíduos de todos os cantos do estado e diversas regiões do país migraram para a região sul, com a expectativa de mudarem de vida com o plantio do cacau ou com trabalhos relacionados a este. Neste período de desenvolvimento, foi construído o porto de Ilhéus em 1924, que facilitou ainda mais a entrada de imigrantes. Em meados dos anos 80 do século XIX, a baixa no preço do cacau, além da queda na produção – influenciada por fenômenos climáticos e pela vassoura de bruxa, levou a decadência da cultura cacauífera (ANDRADE, 2003; ROCHA, 2008).

Esse fenômeno de povoamento mais recente, além da evidente relevância histórica, parece ter influenciado de forma definitiva também a composição genética de indivíduos do sul da Bahia no que tange à ancestralidade.

FREQUÊNCIAS ALÉLICAS

A frequência alélica dos marcadores utilizados, obtidas por contagem direta, permitiu uma análise comparativa entre a amostra avaliada no presente estudo e populações mundiais, incluindo a brasileira (Figura 6). O alelo PV92 *1, por exemplo, atinge frequências superiores a 90% em populações ameríndias (LUIZON *et al.* 2008), no entanto a frequência do mesmo na amostra avaliada não passou de 30%. Apesar da

histórica presença indígena na região, o PV92 *I apresentou frequência bastante semelhante a de europeus (15%) e africanos (22%). A frequência do alelo APO *I chegou a 80%, ficando mais próxima a encontrada entre europeus (94%) e ameríndios (100%). O alelo AT3 *I apresentou frequência de 47% na amostra, sensivelmente diferente dos grupos mundiais, sendo o valor mais próximo encontrado entre europeus (26%). Contudo, a frequência do AT3 foi bastante semelhante ao encontrado por Abe-Sandes *et al.* (2010) em amostra de Salvador (55%). O alelo Sb19.3 *I em particular apresentou frequência de 65%, muito próxima a observada entre ameríndios (71%) e está coerente aos 67% encontrados por Abe-Sandes *et al.* (2010). Esse marcador é capaz de diferenciar africanos de europeus e ameríndios, em nível global. Porém, embora tenha apresentado um alto diferencial de frequência entre a amostra do presente estudo e europeus ($\delta = 0,28$), o Sb19.3 *I, não atingiu o valor mínimo para ser considerado um AIM ($\delta \geq 0,30$) segundo Bonilla *et al.* 2004.

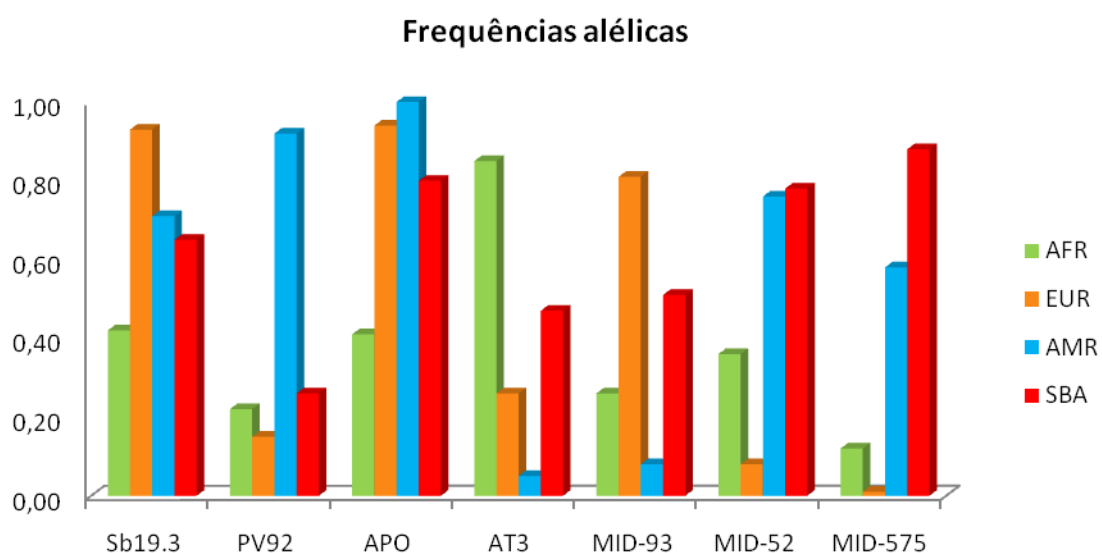


Figura 6. Representação gráfica das frequências do alelo *I dos sete AIMs avaliados na amostra do sul da Bahia (SBA), comparados aos valores encontrados entre africanos (AFR), europeus (EUR) e ameríndios (AMR).

O módulo do valor gerado pelo diferencial entre a frequência de um mesmo marcador em populações distintas (δ) trará parte da informação do quanto estas se aproximam/distanciam geneticamente. Na prática, significa dizer que um grande diferencial de frequência sugere grupos geneticamente distantes, enquanto um pequeno diferencial de frequência sugere grupos geneticamente próximos. Com exceção do Sb19.3, os demais *loci* mostraram-se bons indicativos de ancestralidade. Os *loci* PV92 e MID-93 conseguiram discriminar expressivamente a população ameríndia da amostra ($\delta > 40\%$), portanto podem ser considerados marcadores ameríndios, em acordo ao observado por Luizon *et al.* 2008 entre indígenas da Amazônia. O *locus* APO conseguiu discriminar bem a população africana da amostra ($\delta > 35\%$), dessa forma foi validado para estimar a contribuição gênica africana na população avaliada. O *locus* AT3 conseguiu discriminar tanto a população africana ($\delta > 35\%$) quanto a ameríndia ($\delta > 40\%$) da amostra, portanto foi validado para estimar a contribuição gênica africana ou ameríndia na população avaliada. Os *loci* MID-52 e MID-575 conseguiram discriminar as populações africana e europeia da amostra ($\delta > 40\%$), portanto foram validadas no presente estudo como marcadores africanos ou europeus (Figura 7). De fato, estes *loci* podem ser considerados bons AIMS na população do sul da Bahia, uma vez que o δ variou de 38% contra africanos (AT3) a 87% contra europeus (MID-575)(Tabela 3).

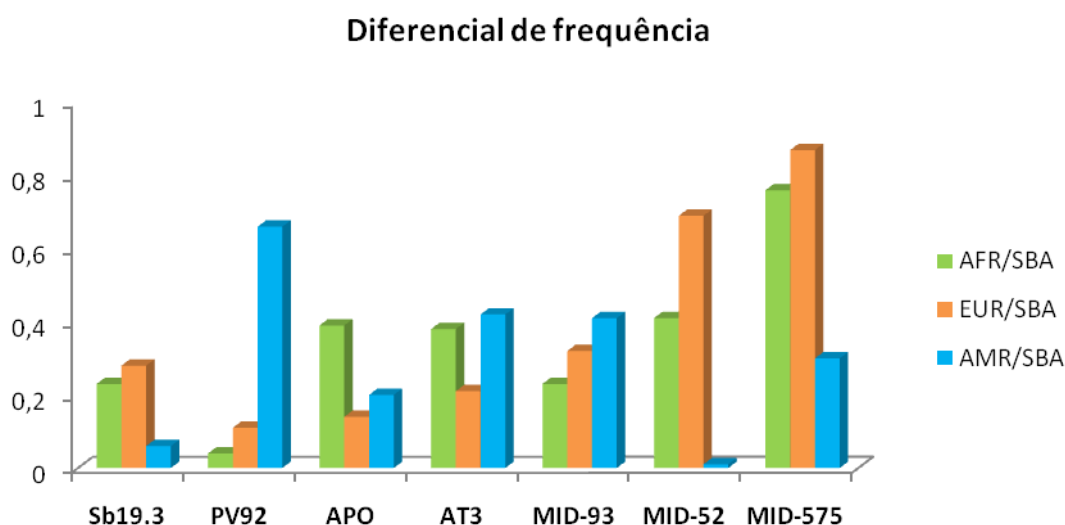


Figura 7. Representação gráfica do diferencial de frequência (δ) entre a amostra estudada e a população africana (AFR), europeia (EUR) e ameríndia (AMR), por *locus*.

EHW E ASSOCIAÇÃO PAR-A-PAR ENTRE *LOCI* NÃO LIGADOS

Foram observados desvios ao esperado no Equilíbrio de Hardy-Weinberg (EHW) nos *loci* Sb19.3 e APO tanto nas subamostras quanto na amostra total. Em ambos os casos, testes mais sensíveis detectaram déficit de heterozigotos. Desvios são observados quando há problemas de amostragem, mas também quando a estruturação da população é real (FALUSH *et al.* 2003). A estruturação populacional é um fenômeno observado entre populações miscigenadas – sobretudo quando se trata de um evento recente – uma vez que cada parental traz consigo um conjunto de alelos distinto dos demais. Levarão algumas gerações até que esses novos alelos sejam fixados na população, além do que a distribuição espacial destes pode não ser homogênea (HOGGART *et al.* 2003; PATTERSON *et al.* 2010).

Outro indício de estruturação pode ser notado quando observamos os resultados da associação par-a-par entre *loci* não ligados, onde a hipótese testada é a de que os alelos estão segregando independentemente. Os mesmos *loci* que apresentaram desvio no EHW, também aparecem em desequilíbrio de ligação na amostra (APO/Sb19.3; $p = 0,04$ e APO/AT3; $p = 0,04$). Dentre os fatores que podem aumentar o desequilíbrio de ligação – como deriva genética, seleção e ligação gênica – mistura recente e

estruturação populacional parecem ser os mais plausíveis nesse caso (HUNTER-ZINCK *et al.* 2010; NARANG *et al.* 2011).

A avaliação da estruturação populacional em uma determinada amostra é relevante, uma vez que esta pode levar a resultados espúrios em trabalhos do tipo caso/controle, onde uma similaridade entre casos e controles é necessária (SHRIVER *et al.* 2003, GATTÁS *et al.* 2004, ZEMBRZUSKI *et al.* 2006). É conveniente salientar que não foram observados *loci* em desequilíbrio na subamostra de Itabuna.

DIFERENCIAÇÃO E DIVERSIDADE GENÉTICA

A amostra do sul da Bahia avaliada no presente estudo é composta de duas subamostras, como explicado inicialmente. Essa medida foi tomada depois de observado pela Dr^a. Sandra Gadelha uma prevalência diferenciada do HTLV-1 entre as amostras coletadas nos municípios de Ilhéus e Itabuna (dados não publicados). Encontramos dois momentos diferentes de fundação e povoamento na região sul da Bahia. O primeiro remete às Capitânicas Hereditárias, enquanto o segundo – já em meados do século XIX – diz respeito ao surgimento de entrepostos comerciais e crescimento de propriedades agrárias (CAMPOS, 2006; ANDRADE e ROCHA, 2005), o que poderia gerar divergências quanto à constituição genética das populações.

Entretanto, nenhum dos marcadores utilizados conseguiu, isoladamente ou no conjunto, diferenciar geneticamente as duas subamostras (Tabela 6). Esse resultado é reforçado pela interpretação do índice de fixação ($F_{ST} = 0,00$). A ausência de barreiras geográficas ou culturais bem definidas por toda a microrregião de Ilhéus-Itabua aparentemente impediu o isolamento genético, tornando a população do sul da Bahia geneticamente homogênea (GAIBAR *et al.* 2011).

A diversidade gênica por *loci* variou de 0,20 (MID-575) a 0,51 (MID-93) (Tabela 7), ambos na amostra de Ilhéus. A heterozigosidade média dentro das subamostras foi de $H_s = 0,39$. Valores altos de diversidade como os encontrados no presente estudo, somado ao baixo coeficiente de endogamia observado ($F_{IS} = 0,11$)

corroboram com a presença de fluxo gênico entre as populações, favorecendo a similaridade genética entre ambas no decorrer das gerações (WANG *et al.* 2007).

MISTURA ÉTNICA

As estimativas de mistura étnica foram obtidas utilizando o método da identidade gênica. O coeficiente de identidade gênica estima a probabilidade de dois genes escolhidos ao acaso em uma ou mais populações serem idênticos por natureza (CHAKRABORTY, 1985). Podem ser utilizados modelos onde há duas (ADMIX2) ou mais (ADMIX3) populações parentais. Tanto as subamostras quanto a amostra total se adequaram ao modelo tri-híbrido com coeficientes de adaptação (R^2) de quase 1 ($\approx 0,99$). Dessa forma podemos dizer que a população do sul da Bahia é histórica e geneticamente tri-híbrida, formada com a contribuição de ameríndios, europeus e africanos.

Em todos os casos a maior contribuição genética foi europeia, seguida da africana e ameríndia. Na amostra total esses valores foram respectivamente, 51, 38 e 11%. Os resultados obtidos no sul da Bahia não estão distantes dos encontrados na capital do estado, com a diferença que nessa última a maior contribuição observada foi africana. Abe-Sandes *et al.* (2010) encontraram uma contribuição africana de 47%, seguida da europeia (37%) e da ameríndia (16%). Já Felix *et al.* (2010) observaram aproximadamente 49% de contribuição africana, 44% de contribuição europeia e 7% de contribuição ameríndia. De fato, em 1982 Azevedo e colaboradores já observavam um processo de “embranquecimento” à medida que se afasta da capital, em direção ao interior do estado. A cidade de Salvador é conhecida historicamente como uma das principais vias de entrada de escravos africanos no Brasil (ALVES-SILVA *et al.* 2000), de forma que não é difícil presumir um gradiente de influência africana a partir dela.

De toda forma, mesmo indivíduos autodeclarados pardos ou negros possuem um alto grau de ancestralidade europeia, no Brasil. Foi o que observou Pena *et al.*

2011 ao avaliar uma amostra total proveniente de quatro regiões do país, incluindo 147 indivíduos de Ilhéus. Entre os nordestinos que se autodeclararam pretos ou pardos, foi encontrada uma contribuição genética europeia $> 50\%$. Para a região nordeste como um todo, a ancestralidade média europeia observada foi $\approx 61\%$, seguida da africana (30%) e da ameríndia (9%).

Em todos os casos, a contribuição genética ameríndia é a mais baixa. É impreciso inferir sobre quantidade, distribuição espacial e organização social dos nativos que ocupavam o território baiano à época do “achamento”, contudo estima-se que já o habitavam a cerca de 200 anos antes da chegada dos primeiros europeus (TAVARES, 2001). Os relatos históricos a respeito da capitania de Ilhéus são de que o número de colonos era insuficiente para conter a ofensiva de alguns grupos indígenas, de forma que a ocupação europeia no território encontrou forte resistência (CAMPOS, 2006). Segundo a Fundação Nacional do Índio (FUNAI), entre 1 a 10 milhões de indígenas ocupavam o território brasileiro antes da chegada dos portugueses. Hoje, segundo o censo de 2010 (IBGE), esse número chega a 817 mil, o que corresponde a cerca de 0,4% da população Brasileira. Além das baixas bélicas, doenças com as quais os nativos nunca haviam tido contato dizimaram grande parte da população ameríndia (FUNAI, 2012).

Dessa forma, a maior contribuição genética indígena parece estar restrita ao DNA mitocondrial, de linhagem materna, onde se observa cerca de 30% de origem ameríndia (ALVES-SILVA *et al.* 2000). Marcadores no cromossomo Y, por sua vez, registraram pouca ou nenhuma influência ameríndia (CARVALHO-SILVA *et al.* 2001). Dados históricos a respeito dos fundadores e do incentivo da Coroa portuguesa à procriação com nativas no início da colonização, visando facilitar o povoamento, corroboram com o cruzamento direcionado (PENA, 2000).

No censo de 2010 (IBGE) apenas 3% da população do sul da Bahia se autodeclarou indígena/amarela. Novamente a menor proporção comparada aos 84% de afrodescendentes (pretos ou pardos) e 13% de eurodescendentes (brancos). A semelhança entre etnia autodeclarada e genética, no entanto, se restringe a menor

proporção indígena. Ainda assim, o presente estudo demonstrou que a contribuição genética ameríndia é quatro vezes maior ($\approx 12\%$) do que observado pela autodeclaração de *cor/raça* (Figura 8). Com relação à contribuição africana e europeia, esses valores se invertem. Enquanto segundo autodeclaração a proporção de afrodescendentes é maior, de acordo com o presente estudo a proporção de eurodescendentes é superior (Figura 5).

No Brasil, o termo *cor/raça* ainda é utilizado pelo IBGE para agrupar indivíduos que devem se autodeclarar como brancos, pardos, pretos, amarelos ou indígenas. Mundialmente, entretanto, há uma clara tendência em abandonar o termo *raça* em detrimento do termo *etnia*. Este último engloba elementos como língua, alimentação e religião, que correspondem à complexidade em torno da expressão e, portanto, é não só mais eficiente para pesquisas como mais realístico (AFSHARI e BHOPAL, 2010).

No Brasil, se percebe uma relação entre cor da pele e desigualdades sociais, que se estende a campos como educação e saúde (HASENBALG, 1995). Nesses casos, como foi observado por Bastos *et al.* (2009), o simples fato do entrevistador/entrevistado ser visualmente classificado como preto/branco pode aumentar em até mais de duas vezes a chance de uma resposta diferente na autodeclaração. Foi observado ainda por Pena *et al.* (2011) que mesmo indivíduos que se autodeclararam pretos ou pardos possuem grande contribuição genética europeia, o que corrobora com o presente estudo.

Como observado, o termo *raça/etnia* sobrepõe uma série de questões, as quais a genética só pode ajudar a resolver em parte. Em se tratando de políticas afirmativas, por exemplo, há de se destacar no debate a forte demanda sociocultural e econômica que afeta diretamente grupos entendidos como minorias, aos quais muito foi negado no decorrer da história Brasileira. O mesmo se aplica a questões como reconhecimento de comunidades quilombolas, indígenas e ao censo, entre outros. Porém, em pesquisas na área de saúde (ZEMBRZUSKI *et al.* 2006; BIRMANN *et al.* 2009), antropogenética (CORACH *et al.* 2010; REICH *et al.* 2012) e forense (PHILLIPS *et*

al. 2009), por exemplo, a utilização de marcadores moleculares para agrupamento e classificação de indivíduos é altamente indicado, sobretudo em populações miscigenadas como o Brasil, no qual a cor de pele é um fraco indicativo de ancestralidade (PARRA *et al.* 2003; BEUTEN *et al.* 2011).

CONCLUSÕES

No presente estudo, foi verificado que a população do sul da Bahia é de fato geneticamente tri-híbrida. As populações parentais são, em ordem crescente de contribuição, europeia, africana e ameríndia. Os valores encontrados em Ilhéus e Itabuna estão bastante próximos, e como nenhum marcador conseguiu diferenciar as amostras, pode-se dizer que estas são geneticamente muito similares. Com exceção do Sb19.3, os demais marcadores (APO, AT3, PV92, MID-93, MID-52 e MID-575) se mostraram bons indicativos de ancestralidade para a população do sul da Bahia. A grande diferença encontrada entre as estimativas de mistura étnica obtidas utilizando marcadores moleculares e *cor/raça* autodeclarada indicam a importância de avaliar a constituição étnica de uma população utilizando AIMs, sobretudo pela relevância do tema em questões socioculturais, de saúde e forense.

REFERÊNCIAS BIBLIOGRÁFICAS

ABÉ-SANDES, K.; BOMFIM, T. F.; ABE-SANDES, C.; ACOSTA, A. X.; ALVES, C. R. B. Ancestralidade Genômica, Nível Socioeconômico e Vulnerabilidade ao HIV/AIDS na Bahia, Brasil. *Saúde Soc.* São Paulo, v.19, supl.2, p.75-84, 2010.

AFSHARI, R.; BHOPAL, R. S. Ethnicity has overtaken race in medical science: MEDLINE-based comparison of trends in the USA and the rest of the world, 1965–2005. *Int J Epidemiol*, p. 1-3, 2010.

ALVES-SILVA, J.; SANTOS, M. S.; GUIMARÃES, P. E.; FERREIRA, A. C.; BANDELT, H. J.; PENNA, S. D.; PRADO, V. F. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*, v. 67, p. 444-61, 2000.

ANDRADE, M. P. e ROCHA, L. B. *De Tabocas a Itabuna: um estudo histórico-geográfico*. Ilhéus: Editos, 2005.

ANDRADE, M. P. *Ilhéus: passado e presente*. 2ed. Ilhéus: EDITUS, 2003.

ARCOT, S. S.; DeANGELIS, M. M.; SHERRY, S. T.; et al. Identification and characterization of two polymorphic Ya5 Alu repeats. *Science*, v. 382, p. 1-6, 1997.

AZEVEDO, E. S.; FORTUNA, C. M.; SILVA, K. M.; SOUSA, M. G.; MACHADO, M. A.; LIMA, A. M.; AGUIAR, M. E.; ABÉ, K.; EULÁLIO, M. C.; CONCEIÇÃO, M. M.; SILVA, M. C.; SANTOS, M. G. Spread and diversity of human populations in Bahia, Brazil. *Hum Biol*, v. 54, n. 2, p. 329-41, 1982.

BAMSHAD, M. J.; OLSON, S. E. Ambiguidades que limitam uma definição de raça. *Sci Am*, v. 20, 2004.

BAMSHAD, M.; WOODING, S.; SALISBURY, B. A.; STEPHENS, J. C. Deconstructing the relationship between genetics and race. *Nature reviews*, v. 5, n. 8, 598-609, 2004.

BARTHE, S.; GUGERLI, F.; BARKLEY, N. A.; MAGGIA, L.; CARDI, C.; SCOTTI, I. Always Look on Both Sides: Phylogenetic Information Conveyed by Simple Sequence Repeat Allele Sequences. *PLoS ONE*, v. 7, n. 7, 2012.

BASTOS, J. L.; DUMITH, S. C.; SANTOS, R. V.; BARROS, A. J. D.; DUCA, G. F. D.; GONÇALVES, H.; NUNES, A. P. Does the way I see you affect the way I see myself? Associations between interviewers' and interviewees' "color/race" in southern Brazil. *Cad de Saúde Pública*, v. 25, n. 10, p. 2111-24, 2009.

BATZER M.A.; DEININGER P.L. A human-specific subfamily of Alu sequences. *Genomics*, v. 9, p. 481-7, 1991.

- BATZER M.A; DEININGER P.L. Repeats of Alu and Human Diversity of Genomic. *Genetics*, v. 3, 370-9, 2002.
- BATZER, M. A; STONEKING, M.; ALEGRIA-HARTMAN, M.; BAZAN, H.; KASS, D. H.; SHAIKH, T. H.; NOVICK, G. E.; et al.. African origin of human-specific polymorphic Alu insertions. *Proc Natl Acad Sci USA*, v. 91, n. 25, p. 12288-92, 1994.
- BEHAR, D. M.; ROSSET, S.; TZUR, S.; SELIG, S.; YUDKOVSKY, G.; BERCOVICI, S.; KOPP, J. B.; et al. African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans. *Hum Mol Genet*, v. 19, n. 9, p. 1816-27, 2010.
- BEUTEN, J.; HALDER, I.; FOWLER, S. P.; GORING, H. H.H.; DUGGIRALA, R.; ARYA, R.; THOMPSON, I. M.; LEACH, R. J.; AND. LEHMAN, D. M. Wide Disparity in Genetic Admixture Among Mexican Americans from San Antonio, TX. *Ann Hum Genet*, v. 75, p. 529–38, 2011.
- BHOPAL, R. Glossary of terms relating to ethnicity and race: for reflection and debate. *J Epidemiol Community Health*,v. 58, n. 6, p. 441-5, 2004.
- BIRMANN, B. M.; BREEN, E. C.; STUVER, S.; CRANSTON, B.; FALK, K. I.; OKAYAMA, A.; HANCHARD, B.; et al. Population differences in immune marker profiles associated with human T-lymphotropic virus type I infection in Japan and Jamaica. *Int J Cancer*, v. 124, n. 3, p. 614–21, 2009.
- BONILLA, C.; PARRA, E. J.; PFAFF, C. L.; DIOS, S.; MARSHALL, J. A.; HAMMAN, R. F.; FERRELL, R. E.; et al. Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann Hum Genet*, v. 68, n. 2, 139-53, 2004.
- BORTOLINI, M.C.; WEIMER, T. A.; FRANCO, M. H.; SALZANO, F. M.; LAYRISSE, Z.; SCHNEIDER, H.; SCHNEIDER, M. P.; HARADA, M. L. Genetic studies in three South American black populations. *Gene Geogr*, v. 6, p. 1-16, 1992.
- BRITTEN, R. J.; ROWEN, L.; WILLIAMS, J.; CAMERON, R. A. Majority of divergence between closely related DNA samples is due to indels. *PNAS*, v. 100, n. 8, p. 4661-5, 2003.
- CALLEGARI-JACQUES, S. M.; GRATAPAGLIA, D.; SALZANO, F. M.; SALAMONI, S. P.; CROSSETTI, S. G.; FERREIRA, M. E.; HUTZ, M. H. Historical genetics: spatiotemporal analysis of the formation of the Brazilian population. *Am J Hum Biol*, v. 15, n. 6, p. 824-34, 2003.
- CAMPOS, J. S. *Crônica da capitania de São Jorge dos Ilhéus*. 3 ed. Ilhéus: EDITUS, 2006.
- CARROLL, M. L.; ROY-ENGEL, A. M.; NGUYEN, S. V.; SALEM, A. H.; VOGEL, E.; VINCENT, B.; MYERS, J.; AHMAD, Z.; NGUYEN, L.; SAMMARCO, M.;

WATKINS, W. S.; HENKE, J.; MAKALOWSKI, W.; JORDE, L. B.; DEININGER, P. L.; BATZER, M. A. Large-scale analysis of the Alu Ya5 and Yb8 subfamilies and their contribution to human genomic diversity. *J Mol Biol*, v. 311, p. 17-40, 2001.

CARVALHO-SILVA, D. R.; SANTOS, F. R.; ROCHA, J.; PENA, S. D. The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet*, v. 68, n. 1, p. 281-6, 2001.

CHAKRABORTY, R. Gene identity in racial hybrids and estimation of admixture rates. In AHUJA Y, R.; NEEL, J. V. *Genetic differentiation in human and others animal populations. Indian Anthropological Association.* p, 171-80, 1985.

CORACH, D.; LAO, O.; BOBILLO, C.; GAAG, K.; ZUNIGA, S.; VERMEULEN, M.; DUIJN, K.; et al. Inferring continental ancestry of argentineans from Autosomal, Y-chromosomal and mitochondrial DNA. *Ann Hum Genet*, v. 74, n. 1, p. 65–76, 2010.

CUNHA, M. C. *História dos Índios no Brasil.* 2a Ed. São Paulo: Companhia das Letras, 1998.

DeGIORGIO, M.; JAKOBSSON, M.; ROSENBERG, N. A. Explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *PNAS*, v. 106, n. 38, p. 16057–62, 2009.

DORIS, P. A. Hypertension Genetics, Single Nucleotide Polymorphisms, and the Common Disease: Common Variant Hypothesis. *Hypertension*, v. 39, n. 2, p. 323-31, 2002.

DOURADO, I.; ALCANTARA, L. C. J.; BARRETO, M. L.; TEIXEIRA, M. G.; GALVÃO-CASTRO, B. HTLV- I in the General Population of Salvador , Brazil. *J Acquir Immune Defic Syndr*, v. 34, n. 5, p. 527-31, 2003.

FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, v. 164, p. 1567–87, 2003.

FEJERMAN, L.; ROMIEU, I.; JOHN, E. M.; LAZCANO-PONCE, E.; HUNTSMAN, S.; BECKMAN, K. B.; PÉREZ-STABLE, E. J.; et al. European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiol Biomarkers Prev*, v. 19, n. 4, p. 1074-82, 2010.

FELIX, G. E. S.; ABE-SANDES, K.; BONFIM, T. M.; BENDICHO, M. T.; CISNEIROS, P.; GUEDES, R.; BRANDÃO, C. J. F.; TORRES, A. J. L.; BRITES, C.; NETTO, E. M.; MEYER, R.; FREIRE, S. M. Ancestry informative markers and complete blood count parameters in Brazilian blood donors. *Rev. Bras. Hematol. Hemoter.* 32(4):282-285, 2010.

FUNAI – SITE DA FUNDAÇÃO NACIONAL DO ÍNDIO. Disponível na Internet via <http://www.funai.gov.br/>. Acesso em março de 2012.

- GAIBAR, M.; ESTEBAN, E.; HARICH, N.; KANDIL, M.; FERNÁNDEZ-SANTANDER, A. (). Genetic differences among North African Berber and Arab-speaking populations revealed by Y-STR diversity. *Ann Hum Biol*, v. 38, n. 2, p. 228-36, 2011.
- GATTÁS, G. J. F.; KATO, M.; SOARES-VIEIRA, J. A.; SIRAQUE, M. S.; KOHLER, P.; GOMES, L.; REGO, M. A. V.; et al. Ethnicity and glutathione S-transferase (GSTM1/GSTT1) polymorphisms in a Brazilian population. *Braz J Med Biol Res*, v. 37; n. 4; p. 451-8, 2004.
- GONZALEZ, E.; BAMSHAD, M.; SATO, N.; MUMMIDI, S.; DHANDA, R.; CATANO, G.; CABRERA, S.; et al. Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *PNAS*, v. 96, n. 21, p. 12004–9, 1999.
- GOUDET, J. *A program to estimate and test gene diversities and fixation indices FSTAT (vers.2.8)*, 1996. (Updated from GOUDET, J. FSTAT (vers.1.2): a computer).
- GRIN, M. A celebração oficial da diversidade no Brasil. *Revista USP*, p. 36-45, 2006.
- GUO S. W.; THOMPSON, E. A. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, v. 48, n. 2, p. 361-72, 1992.
- HARPENDING, H.; ROGERS, A. Genetic perspectives on human origins and differentiation. *Annu Rev Genomics Hum Genet*, v. 1, n. 6, p. 361-85, 2000.
- HASENBALG, C. Entre o mito e os fatos: racismo e relações raciais no Brasil. *Revista de Ciências Sociais*, v. 38, p. 355-74, 1995.
- HENN, B. M.; GIGNOUX, C. R.; JOBIN, M.; GRANKA, J. M.; MACPHERSON, J. M.; KIDD, J. M.; FELDMAN, M. W. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *PNAS*, v. 108, n. 13, p. 5154–62, 2011.
- HOGGART, C. J.; PARRA, E. J.; SHRIVER, M. D.; BONILLA, C.; KITTLER, R. A.; CLAYTON, D. G.; MCKEIGUE, P. M. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet*, v. 72, p. 1492–1504, 2003.
- HUNLEY, K. L.; HEALY, M. E.; LONG, J. C. The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: implications for biological race. *Am J Phys Anthropol*, v. 139, n. 1, p. 35-46, 2009.
- HUNTER-ZINCK, H.; MUSHAROFF, S.; SALIT, J.; AL-ALI, K. A.; CHOUCANE, L.; GOHAR, A.; MATTHEWS, R.; et al. Population genetic structure of the people of Qatar. *Am J Hum Genet*, v. 87, n. 1, p. 17-25, 2010.
- IBGE – CENSO 2010. SITE DO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Disponível na Internet via <http://www.ibge.gov.br/cidadesat/topwindow.htm?1%3E>. Acesso em março de 2012.

- LUIZON, M. R.; MENDES-JUNIOR, C. T.; OLIVEIRA, S. F. D. E.; SIMÕES, A. L. Ancestry Informative Markers in Amerindians from Brazilian Amazon. *Am J Hum Biol*, v. 20, p. 86-90, 2008.
- MATTOSO, K. Q. *Ser escravo no Brasil*. 3a Ed. São Paulo: Brasiliense, 1992.
- McEVOY, B.; BELEZA, S.; SHRIVER, M. D. The genetic architecture of normal variation in human pigmentation: an evolutionary perspective and model. *Hum Mol Genet*, v. 15, n. 2, p. R176-81, 2006.
- MENDES-JR, C. T.; SIMÕES, A. L. *Alu* Insertions and Ethnic Composition in a Brazilian Population Sample. *Int J Hum Genet*, v. 4, p. 249-54, 2001.
- NARANG, A.; JHA, P.; RAWAT, V.; MUKHOPADHAYAY, A.; DASH, D.; BASU, A.; MUKERJI, M. Recent admixture in an Indian population of African ancestry. *Am J Hum Genet*, v. 89, n. 1, p. 111-20, 2011.
- NASCIMENTO, L. B.; CARNEIRO, M. A. S.; TELES, S. A.; LOPES, C. L. R.; REIS, N. R. S.; SILVA, A. M. C. E.; MOTTA-CASTRO, A. R. C.; et al. Prevalência da infecção pelo HTLV-1, em remanescentes de quilombos no Brasil Central. *Rev Soc Bras Med Trop*, v. 42, n. 6, p. 657-660, 2009.
- PARRA, E. J.; KITTLES, R. A.; ARGYROPOULOS, G.; PFAFF, C. L.; HIESTER, K.; BONILLA, C.; SYLVESTER, N.; et al. Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am J Phys Anthropol*, v. 114, n. 1, p. 18-29, 2001.
- PARRA, E. J.; MARCINI, A.; AKEY, J.; MARTINSON, J.; BATZER, M. A.; COOPER, R.; FORRESTER, T.; et al. Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet*, v. 63, n. 6, p. 1839-51, 1998.
- PARRA, F. C.; AMADO, R. C.; LAMBERTUCCI, J. R.; ROCHA, J.; ANTUNES, C. M.; PENA, S. D. J. Color and genomic ancestry in Brazilians. *PNAS*, v. 100, n. 1, p. 177-82, 2003.
- PATTERSON, N.; PETERSEN, D. C.; ROSS, R. E.; SUDOYO, H.; GLASHOFF, R. H.; MARZUKI, S.; REICH, D.; HAYES, V. M. Genetic structure of a unique admixed population: implications for medical research. *Hum Mol Genets*, v. 19, n. 3, 2010.
- PENA, S. D. J.; BASTOS-RODRIGUES, L.; PIMENTA, J. R.; BYDLOWSKI, S. P. DNA tests probe the genomic ancestry of Brazilians. *Braz J Med Biol Res*, v. 42, n. 10, p. 870-6, 2009.
- PENA, S. D. J.; CARVALHO-SILVA, D.R.; ALVES-SILVA, J.; PRADO, V. F. Retrato molecular do Brasil. *Ciência Hoje*, v. 27, p. 16-25, 2000.
- PENA, S. D. J.; Di PIETRO, G.; FUCHSHUBER-MORAES, M.; GENRO, J. P.; HUTZ, M. H.; KEHDY, F. D. S. G.; KOHLRAUSCH, F.; et al. The genomic ancestry

of individuals from different geographical regions of Brazil is more uniform than expected. *PloS ONE*, v. 6, n. 2, p. e17063, 2011.

PHILLIPS, C.; PRIETO, L.; FONDEVILA, M.; SALAS, A.; GÓMEZ-TATO, A.; ALVAREZ-DIOS, J.; ALONSO, A.; et al. Ancestry analysis in the 11-M Madrid bomb attack investigation. *PloS ONE*, v. 4, n. 8, p. e6583, 2009.

PIMENTA, J. R.; ZUCCHERATO, L. W.; DEBES, A. A.; MASELLI, L.; SOARES, R. P.; MOURA-NETO, R. S.; ROCHA, J.; et al. Color and genomic ancestry in Brazilians: a study with forensic microsatellites. *Hum Hered*, v. 62, p. 190-5, 2006.

PINSKY, J. *A escravidão no Brasil*. São Paulo: Editora Contexto, 2000.

RAY, D. A.; WALKER, J. A.; HALL, A.; LLEWELLYN, B.; BALLANTYNE, J.; CHRISTIAN, A. T.; TURTELTAUB, K.; et al. Inference of human geographic origins using Alu insertion polymorphisms. *Forensic Science International*, v. 153, p. 117-24, 2005.

RAYMOND, M.; ROUSSET, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenism. *J Hered*, v. 86, p. 248-9, 1995.

REICH, D.; PATTERSON, N.; CAMPBELL, D.; TANDON, A.; MAZIERES, S.; RAY, N.; PARRA, M. V.; et al. Reconstructing Native American population history. *Nature*, p. 1-6, 2012.

RIBEIRO, D. *O povo brasileiro: a formação e o sentido do Brasil*. São Paulo: Companhia das Letras, 1995.

ROCHA, L. B. *A região cacauzeira da Bahia: dos coronéis à vassoura-de-bruxa*. 1ª Ed, Ilhéus: EDITUS, 2008.

SANTOS, A. S. Historical roots of the “Whitening” of Brazil. *Lat Am Perspect*, v. 28, p. 61-82, 2002.

SEYFERTH G. A antropologia e a teoria do branqueamento de raça no Brasil: a tese de João Batista de Lacerda. *Rev Mus Paulista*, v. 30, p. 81-98, 1985.

SHRIVER, M. D.; PARRA, E. J.; DIOS, S.; BONILLA, C.; NORTON, H.; JOVEL, C.; PFAFF, C.; et al. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet*, v. 112, n. 4, p. 387-99, 2003.

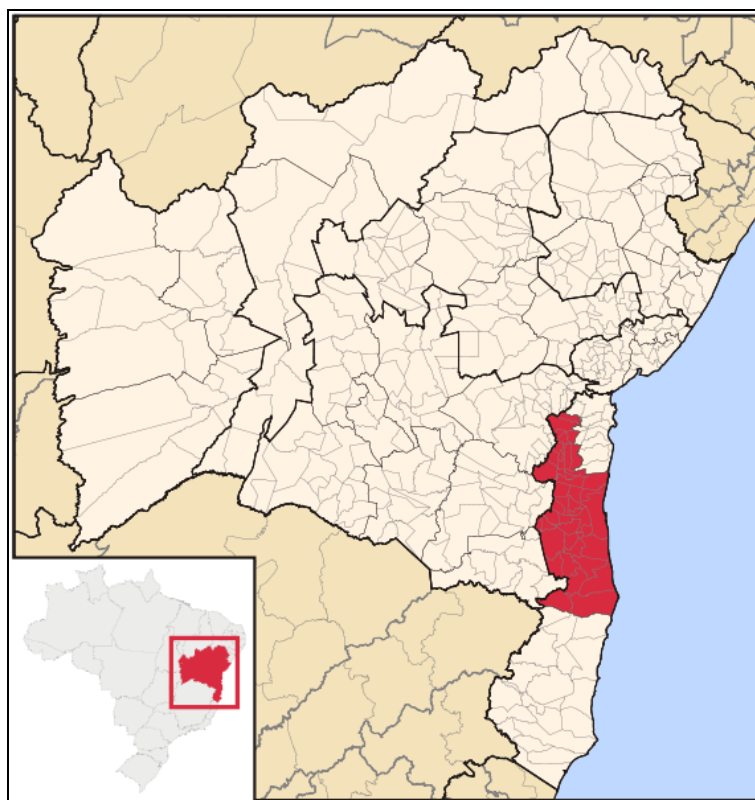
SHRIVER, M. D.; SMITH, M. W.; JIN, L.; MARCINI, A.; AKEY, J. M.; DEKA, R.; FERRELL, R. E. Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet*, v. 60, n. 4, p. 957-64, 1997.

STONEKING, M. Human origins. *EMBO reports*, v. 9, p. S46-50, 2008.

STRACHAN, T.; READ, A. P. *Genética Molecular Humana*. 2ª Ed. Porto Alegre: Editora Artmed, 2002.

- STURM, R.A. Molecular genetics of human pigmentation diversity. *Hum Mol Genet*, v. 18, n. 1, p. R9-17, 2009.
- TAVARES, L. H. D. *História da Bahia*. 10 ed. Salvador: Edufba/Editora Unesp, 2001.
- TOMAS, G.; SECO, L.; SEIXAS, S.; FAUSTINO, P.; LAVINHA, J.; ROCHA, J. The peopling of São Tomé (Gulf of Guinea): origins of slave settlers and admixture with the Portuguese. *Hum Biol*, v. 74, p. 397–411, 2002.
- VÄLI, U.; BRANDSTRÖM, M.; JOHANSSON, M.; ELLEGREN, H. Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC Genetics*, v. 9, n. 8, 2008.
- WANG, S.; LEWIS, C. M.; JAKOBSSON, M.; RAMACHANDRAN, S.; RAY, N.; BEDOYA, G.; ROJAS, W.; et al. Genetic variation and population structure in native Americans. *PLoS Genetics*, v. 3, n. 11, p. e185, 2007.
- WEIR, B.S. *Genetic data analysis II*. Sunderland: Sinauer Associates, 1996.
- WRIGHT, S. *Evolution and the genetics of population, variability within and among natural populations*. The University of Chicago Press: Chicago, 1978.
- ZEMBRZUSKI, V. M.; CALLEGARI-JACQUES, S. M.; HUTZ, M. H. Application of an African Ancestry Index as a genomic control approach in a Brazilian population. *Ann Hum Genet*, v. 70, p. 822-8, 2006.

ANEXO I



Anexo I. Mapa do estado da Bahia, destacando microrregião de Ilhéus-Itabuna

Fonte: Wikipédia; Autor: Raphael Lorenzeto de Abreu

Lista dos municípios da microrregião de Ilhéus-Itabuna:

- | | | | |
|-------------------|-------------------|---------------------------|-------------------------|
| 1. Almadina | 12. Floresta Azul | 23. Itaju do Colônia | 34. Santa Luzia |
| 2. Arataca | 13. Gandu | 24. Itajuípe | 35. São José da Vitória |
| 3. Aurelino Leal | 14. Gongogi | 25. Itamari | 36. Teolândia |
| 4. Barra do Rocha | 15. Ibicarai | 26. Itapé | 37. Ubaitaba |
| 5. Barro Preto | 16. Ibirapitanga | 27. Itapebi | 38. Ubatã |
| 6. Belmonte | 17. Ibirataia | 28. Itapitanga | 39. Una |
| 7. Buerarema | 18. Ilhéus | 29. Jussari | 40. Uruçuca |
| 8. Camacan | 19. Ipiauí | 30. Mascote | 41. Wenceslau Guimarães |
| 9. Canavieiras | 20. Itabuna | 31. Nova Ibiá | |
| 10. Coaraci | 21. Itacaré | 32. Pau Brasil | |
| 11. Firmino Alves | 22. Itagibá | 33. Santa Cruz da Vitória | |

ANEXO II



www.acipes.com.br

Soroprevalência de agentes de transmissão materno-fetal (vírus linfotrópico de células T humanas (HTLV), citomegalovírus (CMV), toxoplasma gondii e rubéola), subtipagem do HTLV e avaliação da transmissão materno-fetal em gestantes do sul da Bahia.



www.colsan.org

Termo de consentimento livre e esclarecido

(de acordo com Resolução nº 196/96, Conselho Nacional de Saúde)

A senhora está sendo convidada a participar de uma pesquisa que tem como finalidade pesquisar vários agentes que podem ser transmitidos da mãe para o feto em gestantes de Ilhéus e Itabuna. Iremos pesquisar: o HTLV (vírus linfotrópico de células T humanas), o CMV (citomegalovírus), o vírus da rubéola e o toxoplasma gondii (responsável pela toxoplasmose). Todos eles podem ser passados da mãe para o feto ou para o bebê dentro da barriga, ou durante o parto, ou, ainda, pela amamentação. Pretendemos também, neste trabalho, acompanhar as mães que forem positivas para o HTLV e os seus filhos.

A maioria dos indivíduos que tem o HTLV e o CMV não apresenta qualquer doença, porém podem passar o vírus para seu filho, e só descobre que têm o vírus quando fazem testes específicos no sangue. A transmissão pode também ser por transfusão de sangue, ato sexual, uso de drogas injetáveis, e não está relacionada ao HIV, vírus que causa a AIDS. O HTLV causa doenças diferentes da AIDS, porém também causa doenças sérias e para as quais ainda não existe tratamento estabelecido e nem cura.

No caso da rubéola, quando o vírus é adquirido na gravidez, ele pode passar para o feto e causar vários problemas, incluindo surdez e retardo mental. Estes problemas também podem ocorrer quando a grávida se infecta com o toxoplasma, e, neste caso, a infecção normalmente está associada ao consumo de carne mal cozida.

A senhora será convidada a participar como voluntária. Caso aceite, será realizado um exame de sangue. Esse exame poderá provocar um leve ardor causado pela picada da agulha, e, muito raramente, hematoma (mancha roxa). Esses são os mesmos efeitos que qualquer exame de sangue pode causar. Todos os testes serão acompanhados por profissional habilitado e medidas para diminuir os problemas citados serão realizadas.

Caso o seu teste seja positivo para o HTLV, você e seu filho serão convidados a voltar ao hospital para fazer outro exame de sangue. Seu filho será acompanhado pela equipe (médico, bioquímico e pesquisadores) até os dois anos e fará exames de sangue duas vezes por ano, para verificar se o vírus foi passado para ele. A senhora tem liberdade de se recusar a participar e ainda se recusar a continuar participando em qualquer momento da pesquisa, sem qualquer prejuízo. A pesquisa será realizada na Universidade Estadual de Santa Cruz e qualquer dúvida você poderá entrar em contato com a pesquisadora responsável Profa. Dra. Sandra Rocha Gadelha Mello, através do telefone 3680 5307.

