

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



COMPOSIÇÃO GENÔMICA DE UMA AMOSTRA DE
HOMENS DO SUL DA BAHIA.

JAMILLE SILVA OLIVEIRA

ILHÉUS-BAHIA-BRASIL

Fevereiro de 2012

JAMILLE SILVA OLIVEIRA

**COMPOSIÇÃO GENÔMICA DE UMA AMOSTRA DE HOMENS 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

ILHÉUS-BAHIA-BRASIL

Fevereiro de 2012

O48

Oliveira, Jamille Silva.

Composição genômica de uma amostra de homens do sul da Bahia / Jamille Silva Oliveira. – Ilhéus, BA: UESC, 2012.

xiv, 58f. : Il.

Orientadora: Sandra Mara Bispo Sousa.

Dissertação (Mestrado) – Universidade Estadual de Santa Cruz. Programa de Pós-Graduação em Genética e Biologia Molecular.

Inclui bibliografia.

1. Genética da população humana. 2. Marcadores genéticos . 3. Próstata - Câncer. 4. Miscigenação. I. Título.

CDD 576.5

JAMILLE SILVA OLIVEIRA

**COMPOSIÇÃO GENÔMICA DE UMA AMOSTRA DE HOMENS 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: 24 de fevereiro de 2012

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

Prof. Dr. Jerônimo Pereira de França
(UESC)

Prof. Dr. Lauro Juliano Marin
(UESC)

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

A toda minha família. Suas existências me fazem feliz, me realiza, suas presenças me completam. Vocês são minha razão e inspiração de estar aqui e de viver esta aventura.

Dedico

AGRADECIMENTOS

A Prof.^a Dr.^a Sandra Mara Bispo Sousa, pela orientação, incentivo, exemplo de dedicação à pesquisa, interesse, disponibilidade com que sempre me atendeu e amizade.

A Prof.^a Dr.^a Sandra Gadelha pela orientação nas pesquisas e no grupo de discussão em genética humana.

Ao Prof. Dr. Paulo Melo pela participação na comissão de orientação.

Ao Prof. Dr. Fabrício Santos Rios, pelo apoio no Laboratório de Farmacogenética e Epidemiologia Molecular.

A todos os professores do PPGGBM que direta ou indiretamente contribuíram na minha formação acadêmica e construção desse trabalho.

Aos amigos Cintia Marques, Thiago Magalhães, Bruna Luise, Livia Santana e Fernanda Ancelmo pelo apoio e incentivo nos momentos de desespero, onde quase desisti, só não perdi a esperança, porque tive vocês ao meu lado; **MUITO OBRIGADA!**

Aos amigos Raoni e Ronaldo pelo companheirismo, por mudarem meu modo de ver a vida e por tornarem o cotidiano na Universidade tão simples e divertido, que com certeza terei boas lembranças e saudades.

A minha família: **Geraldão** por sempre estar ao meu lado, me apoiar em tudo e por ter proporcionado a realização deste desejo; **Cabeção** pelas meigas palavras e por me dar forças; **Jack** por ter segurado a barra em casa e pelo incentivo de sempre; e por fim, **Yan** por ter trazido alegria e esperança a nossa família. Muito obrigada por compreenderem a distância e a falta de tempo para ficarmos juntos; não poderia deixar de agradecer também a **mainha** (*in memorian*) por ter me ensinado a aprender com as dificuldades e torná-las uma ferramenta favorável, mãe muito obrigada por sempre estar ao meu lado e por ter me ajudado a superar tantas dificuldades através de sua energia positiva. **AMO VOCÊS.**

Ao CNPq pela concessão de dois anos da Bolsa de mestrado e apoio financeiro de suma importância para o desenvolvimento do trabalho aqui apresentado.

A todos que de alguma forma contribuíram para a realização deste trabalho.

Meus sinceros agradecimentos

“Ciência: busca incessante da verdade que é inatingível”

Albert Einstein

ÍNDICE

1	INTRODUÇÃO.....	1
2	REVISÃO BIBLIOGRÁFICA.....	3
2.1	CARACTERÍSTICAS DA POPULAÇÃO	3
2.2	MARCADORES MOLECULARES	8
2.3	CÂNCER DE PRÓSTATA E ETNIA.....	11
3	METODOLOGIA	13
3.1	AMOSTRAGEM	13
3.2	EXTRAÇÃO DE DNA.....	13
3.3	ASPECTOS ÉTICOS.....	14
3.4	MARCADORES MOLECULARES ANALISADOS.....	14
3.5	REAÇÃO EM CADEIA DA POLIMERASE (PCR).....	15
3.6	REAÇÃO DE RESTRIÇÃO.....	16
3.7	ANÁLISE DO PRODUTO AMPLIFICADO	18
3.7.1	Reagentes e Soluções para os géis de acrilamida.....	18
3.7.2	Reagentes e Soluções para os géis de AGAROSE.....	19
3.7.3	Coloração com nitrato de prata e secagem do gel	20
3.8	PERFIL ELETROFORÉTICO	22
3.8.1	APO	22
3.8.2	SB19.3	23
3.8.3	PV92	23
3.8.4	AT3.....	24
3.8.5	LPL	24
3.8.6	RB2300.....	25
3.8.7	DRD2.....	25
3.8.8	CKMM.....	26
3.9	ANÁLISES DOS DADOS	26
3.9.1	Frequências Alélicas.....	26
3.9.2	Aderência ao Equilíbrio de Hardy-Weinberg.....	27
3.9.3	Associações par-a-par entre <i>loci</i>	27
3.9.4	Diferenciação genética das populações	28
3.9.5	Estimativas de mistura étnica	28
4	RESULTADOS	29

4.1	FREQUÊNCIAS ALÉLICAS.....	29
4.2	ADERÊNCIA AO EQUILÍBRIO DE HARDY-WEINBERG.....	34
4.3	ASSOCIAÇÕES PAR A PAR ENTRE <i>LOCI</i>	35
4.4	DIFERENCIAÇÃO GENÉTICA DAS POPULAÇÕES	35
4.5	MISTURA ÉTNICA	38
5	DISCUSSÃO.....	40
5.1	FREQUÊNCIAS ALÉLICAS.....	40
5.2	EQUILÍBRIO DE HARDY-WEINBERG.....	44
5.3	ASSOCIAÇÕES PAR A PAR ENTRE <i>LOCI</i>	44
5.4	DIFERENCIAÇÃO GENÉTICA DAS POPULAÇÕES	45
5.5	MISTURA ÉTNICA	45
6	CONCLUSÕES.....	49
	REFERÊNCIAS BIBLIOGRÁFICAS	50

EXTRATO

OLIVEIRA, Jamille Silva, M.S., Universidade Estadual de Santa Cruz, Ilhéus, fevereiro de 2012. **Composição genômica de uma amostra de homens do sul da Bahia.** Orientadora: Sandra Mara Bispo Sousa. Co-orientadora: Sandra Rocha Gadelha Mello. Conselheiro: Paulo Roberto Melo.

A população brasileira é o resultado de cinco séculos de uniões interétnicas entre ameríndios, europeus e africanos. Muitas populações humanas miscigenadas têm sido formadas pela mistura de duas ou mais populações ancestrais, as quais são geneticamente distintas. Em estudos de associação do tipo caso-controle com amostras miscigenadas é comum a ocorrência de associações espúrias por conta da estratificação populacional. Esse efeito pode levar a resultados falso-positivos ou falso-negativos em estudos de associação, não permitindo detectar a real relação entre genes candidatos e o fenótipo em questão. Uma estratégia que minimiza tais discrepâncias é utilizar Marcadores Informativos de Ancestralidade (MIAs) para conhecer a composição étnica das amostras e se necessário realizar o controle genômico e assim homogeneizar as subamostras. Os MIAs apresentam alto diferencial de frequência alélica (δ) entre populações ancestrais, portanto, podem ser utilizados para caracterizar a composição genética de populações miscigenadas. Com o objetivo de estimar as contribuições africanas, europeias e ameríndias em uma amostra caso-controle da região sul da Bahia foram utilizados oito MIAs (Sb19.3, APO, AT3, RB2003 e LPL, CKM, PV92 e DRD2). Foram selecionados 102 homens diagnosticados com câncer de próstata e 109 controles. Os resultados desse estudo mostraram os alelos **I* para sete (APO, SB19.3, PV92, AT3, LPL, RB2300 e CKM) dos oito *loci* utilizados, exceto o *loci* DRD2. Os desvios encontrados do Equilíbrio de Hardy-Weinberg para os *loci* AT3, Sb19.3, APO podem ser explicados pela falta de heterozigotos e para o *locus* RB2300 pelo excesso. O teste exato G de diferenciação populacional mostrou heterogeneidade entre Controles Ilhéus (ILH) e Casos Itabuna (ITB) para o teste multi-locus. Na análise multipopulacional apenas o *locus* APO diferenciou as subamostras. As estimativas de mistura étnica mostraram que a contribuição média para os casos foi de 52% europeu, 43% africano e 5% ameríndio, e para os controles 63% europeu, 32% africano e 5% ameríndio. Em todas as amostras a maior contribuição foi europeia, demonstrando que há diferenças nas contribuições étnicas das amostras da região sul quando comparadas as encontradas em amostra populacional de Salvador. A diferença observada entre as estimativas de mistura étnica obtidas com os MIAs e as referidas através da autodeclaração indicam a importância de estimar a ancestralidade utilizando os marcadores genéticos. A diferenciação gênica populacional observada no *locus* APO, desvios no EHW, associações par a par entre *loci* não ligados, contribuições variáveis nas estimativas de mistura étnica entre as subamostras e por fim discordâncias entre ancestralidade genômica e cor da pele autodeclarada indicam subestruturação populacional e a necessidade de realização do controle genômico.

Palavras-chave: Miscigenação. Caso/controles. Câncer de próstata. MIAs. Ancestralidade genética

ABSTRACT

OLIVEIRA, Jamille Silva, M.S., Universidade Estadual de Santa Cruz, Ilhéus, fevereiro de 2012. **Genomic composition of a sample the men in the southern region of Bahia.** Advisor: Sandra Mara Bispo Sousa. Advisor Committee Membres: Sandra Rocha Gadelha Mello and Paulo Roberto Melo.

Brazilian population is result of five centuries of interethnic unions among Amerindians, Europeans and Africans. Many admixed human populations have been being formed by the mixture of two or more ancestral population, which are genetically distinct. Association studies of the type case/controls in samples admixed is common incidence of spurious associations due to population stratification. This effect can lead to false-positive or false-negative results in association studies, not detecting the real relationships between candidate genes and the phenotype in question. A strategy that reduces such discrepancies is use Ancestry Informative Markers (AIM) to know the ethnic composition of the samples and if necessary accomplish the genomic control and thus homogenize the samples. The AIM showing alleles with large frequency differentials (δ) between ancestral populations. AIMs can be used to characterize the genetic composition of admixed population. We used eight AIMs to estimate African, European and Amerindian contributions in a case/control sample of men with prostate cancer in southern region of Bahia. We selected 102 men diagnosed with prostate cancer and 109 controls. Results of this study showed the alleles *1 for seven (APO, SB19.3, PV92, AT3, LPL, RB2300 and CKM) of the eight *loci* used, except *loci* DRD2. Found deviations from Hardy-Weinberg equilibrium for *loci* AT3, Sb19.3, APO can be explained by heterozygote deficiency and to *locus* RB2300 by the excess. G test of population differentiation showed heterogeneity between Ilhéus (ILH) controls and Itabuna (ITB) cases for the multi-locus test. In multi-population analysis only *locus* APO differed samples. Admixture estimation showed 52% European, 43% African and 5% Amerindian average contribution for cases. Controls were in average 63% European, 32% African and 5% Amerindian. We find major European contribution for all samples. Results differ from ethnic contribution found in a population of Salvador, Bahia. The differences observed between admixture estimation estimated data form AIMs and self-classification based criteria indicate the importance of to estimate the ancestry using the genetic markers. The population genetic differentiation observed in APO *locus*, HWE deviations, pairwise association between *loci* not linked variable contributions in the admixture estimation among samples and finally disagreements between genomic ancestry and skin color self-declared indicate structuring population and the need for the performance of genomic control.

Keywords: Admixture. Case/control. Prostate cancer. AIMs. Genetic ancestry

LISTA DE TABELAS

Tabela 1. Localização citogenética, enzimas de restrição e sequência dos <i>primers</i> dos oito MIAs analisados.	15
Tabela 2. Condições de PCR para os <i>loci</i> analisados no presente trabalho. Quantidade em μ L suficiente para uma reação.	16
Tabela 3. Frequências genótípicas e alélicas dos oito <i>loci</i> nas quatro amostras analisadas- Ilhéus (ILH) e Itabuna (ITB).	29
Tabela 4. Diferencial de frequência (δ) dos loci MIAs entre grupos étnicos mundiais – Africano (AFR)1, Europeu (EUR)1 e Ameríndio (AMR)1,2 e os casos e controles em Itabuna (ITB) e Ilhéus (ILH). Os valores >30% estão destacados em negrito.	33
Tabela 5. Probabilidades de desvio casual segundo teste exato (Guo e Thompson, 1992) para a verificação do equilíbrio de Hardy-Weinberg (excesso de heterozigotos). Os valores significativos ($p < 0,05$) estão destacados em negrito.	34
Tabela 6. Probabilidades de desvio casual segundo teste exato (Guo e Thompson, 1992) para a verificação do equilíbrio de Hardy-Weinberg (déficit de heterozigotos). Os valores significativos ($p < 0,05$) estão destacados em negrito.	35
Tabela 7. Associações par a par entre <i>loci</i> em cada uma das quatro subamostras. Os valores estatisticamente significativos ($p < 0,05$) estão destacados em negrito.	36
Tabela 8. Diferenciação gênica baseada nos oito <i>loci</i> MIAs entre as quatro subamostras: Casos ILH (1), Casos ITB (2), Controles ITB (3) e Controles ILH (4). Os valores significativos ($p < 0,05$) estão destacados em negrito.	37
Tabela 9. Estimativa de mistura étnica nas amostras analisadas, segundo o método de identidade gênica. Valores em percentuais.	38
Tabela 10. Estimativas das contribuições étnicas da amostra de casos e controles obtidas com os MIAs e as avaliadas com a autodeclaração obtidas por Santos (2010).	39

LISTA DE FIGURAS

- Figura 1.** Padrão de bandas observado para o marcador APO. Gel de agarose 2%, corado com GelRed. Na raia 1 o *Ladder* de 50pb, nas raias 2, 4 5 e 6 encontram-se os homozigotos para a inserção (+/+) e na raia 3 heterozigoto (-/+).22
- Figura 2.** Padrão de bandas observado para o marcador SB19.3. Gel de agarose 2%, corado com GelRed. Nas raias 1, 8 e 9 encontram-se homozigotos para inserção (+/+), na raia 4, homozigotos para ausência da inserção (-/-), e nas raias 2, 3, 5, 6 e 7 heterozigotos (-/+).23
- Figura 3.** Padrão de bandas observado para o marcador PV92. Gel de agarose 2%, corado com GelRed. Na raia 1, o *Ladder* de 50pb, nas raias 2 e 6 encontram-se homozigotos para a inserção (+/+), nas raias 3, 4 e 5, homozigotos para a ausência da inserção (-/-), e nas raias 7 e 8 heterozigotos (-/+).23
- Figura 4.** Padrão de bandas observado para o marcador AT3. Gel de agarose 3%, corado com GelRed. Nas raias 1 e 5 encontram-se homozigotos para a inserção (+/+), nas raias 3 e 4 homozigotos para a deleção (-/-), e na raia 2 heterozigotos (-/+).24
- Figura 5.** Padrão de bandas observado para o marcador LPL. Gel de agarose 2%, corado com GelRed. Nas raias 1 e 3 encontram-se homozigotos para ausência do sítio de restrição (+/+), na raia 4 , homozigoto para a presença do sítio de restrição (-/-), e na raia 2 heterozigoto (-/+).24
- Figura 6.** Padrão de bandas observado para o marcador RB200. Gel de agarose 2%, corado com GelRed. Nas raias 3 e 6 encontram-se homozigotos para ausência do sítio de restrição (+/+), na raia 2, 4 e 5 heterozigotos o sítio de restrição (-/+).25
- Figura 7.** Padrão de bandas observado para o marcador DRD2. Gel de agarose 3%, corado com GelRed. Na raia 1 o *Ladder*, de 50pb, nas raias 2, 4 e 5 encontram-se homozigotos para o sítio de corte, nas raias 3 e 6, amostras não digeridas.25
- Figura 8.** Padrão de bandas observado para o marcador CKM. Gel de agarose 2%, corado com GelRed. Nas raias 1 e 2 encontram-se homozigotos para ausência do sítio de restrição (+/+), na raia 3 e 5, homozigotos para o sítio de restrição (-/-), e nas raias 2 e 5, heterozigotos (-/+).26
- Figura 9.** Comparação das frequências dos alelos *APO*1*, *PV92*1*, *DRD2*1*, *RB2300*1*, *LPL*1*, *AT3*1* e *Sb19.3*1* nos casos e controles de Ilhéus.31
- Figura 10.** Comparação das frequências dos alelos *APO*1* *PV92*1* *DRD2*1*, *RB2300*1*, *LPL*1*, *AT3*1* e *Sb19.3*1* nos casos e controles de Itabuna.32
- Figura 11.** Comparação gráfica das estimativas de mistura étnica obtida a partir de oito *loci* (*Sb19.3*, *APO*, *PV92*, *AT3*, *LPL*, *DRD2*, *RB2300* e *CKM*).39
- Figura 12.** Diferencial de frequência (δ) dos alelos *Sb19.3*1*, *APO*1*, *PV92*1*, *AT3*1*, *LPL*1*, *DRD2*1*, *RB2300*1*, *CKM*1* entre os grupos étnicos mundiais.43

LISTA ABREVEATURAS

MIAs- Marcadores Informativos de Ancestralidade
IBGE- Instituto Brasileiro de Geografia e Estatística
DNA- Ácido Desoxirribonucléico
PSAs- *Population Especific Aleles*
AIMs- *Ancestry Informative Markers*
SNPs- *Single Nucleotide Polymorphisms*
SINE- *Short Interspersed Nuclear Elements*
LINE- *Long Interspersed Nuclear Elements*
Ca- Câncer
CaP- Câncer de Próstata
INCA- Instituto Nacional do Câncer
PSA- *Prostate Specific Antigen*
TCLE- Termo de Consentimento Livre e Esclarecido
RB2300- gene do Retinoblastoma
LPL- gene da lipoproteína lipase
AT3- gene da antitrombina III
Sb19.3- sequência repetida *Alu Yb8 Sb19.3*
APO- gene da apolipoproteína
PV92- sequência repetida *Alu PV92*
CKMM- gene da Creatina kinase do músculo
DRD2- gene 2 do receptor de dopamina
PCR- *Polymerase Chain Reaction*

dNTP- *deoxyribonucleoside triphosphate*

DMSO- Dimetilsulfóxido

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

PAGE- gel de poliacrilamida

EDTA- *Ethylenediamine tetraacetic acid*

TEMED: tetrametilenodiamina

TBE- Tampão Tris/Borato/EDTA

EHW- Equilíbrio de Hardy-Weinberg

ILH- Ilhéus

ITB- Itabuna

AFR- Africano

EUR- Europeu

AMR- Ameríndio

1 INTRODUÇÃO

A população brasileira é uma das mais heterogêneas do mundo, sendo resultado de cinco séculos de cruzamentos interétnicos entre ameríndios, europeus e africanos (PENA *et al.*, 2009). A Bahia possui a mesma composição étnica do Brasil, já que foi o berço da colonização do país. Este estado possui uma população altamente diversa, também baseada em três principais grupos formadores: ameríndios, africanos e europeus (ABÉ-SANDES *et al.*, 2010; PENA *et al.*, 2011).

Muitas populações humanas miscigenadas têm sido formadas pela mistura de duas ou mais populações ancestrais, as quais são geneticamente distintas (PATTERSON *et al.*, 2010). Esse processo de miscigenação torna a população estratificada ou subdividida, na qual há diferenças na proporção de mistura entre grupos de indivíduos dentro da população (HOGGART *et al.*, 2003).

Dessa forma, em populações como a brasileira, a existência de estruturação populacional pode levar a resultados falso-positivos ou falso-negativos, não permitindo detectar a real associação existente entre os genes candidatos e o fenótipo em questão. Por isso, o controle genômico é um passo indispensável nesses estudos populacionais. Entende-se por controle genômico a caracterização da proporção de ancestralidade individual e global de uma amostra populacional caso/controles (SANTOS *et al.*, 2009), tornando-as a mais homogênea possível.

Os estudos de associação do tipo caso/controles são amplamente utilizados em trabalhos que visam entender a etiologia genética e ambiental de muitas doenças em populações humanas. A estrutura da população, entretanto, torna-se um obstáculo para esses estudos, podendo surgir associações espúrias entre um marcador e um determinado fenótipo. Este tipo de delineamento justifica a importância de caracterizar geneticamente a etnia das amostras estudadas, com o intuito de evitar falsas associações. Isto pode ser observado nos casos em que o risco da doença é diferente entre os grupos com ancestralidade diferente. Diversos estudos têm demonstrado que determinados polimorfismos genéticos candidatos para o câncer de

próstata são mais frequentes em indivíduos de ascendência africana em comparação com indivíduos de ascendência europeia (GIRI *et al.*, 2009; BONILLA *et al.*, 2011; OKOBIA *et al.*, 2011).

Uma estratégia para eliminar esse viés de confundimento é a utilização de Marcadores Informativos de Ancestralidade (MIAs) para realizar o controle genômico e assim homogeneizar as subamostras. Os MIAs apresentam alto diferencial de frequência alélica (δ) entre populações parentais, portanto, podem ser utilizados para caracterizar a composição genética de populações miscigenadas (PARRA *et al.*, 2003; BEUTEN *et al.*, 2011; PENA *et al.*, 2011), visto que a cor da pele é um marcador insuficiente de ancestralidade genômica. Esses marcadores moleculares também podem ser utilizados em estudos com a finalidade de identificar a estrutura populacional atual e identificar os principais povos formadores da população (BRYC *et al.*, 2010).

Uma hipótese avaliada neste estudo foi se as amostras caso/controles para câncer de próstata da região sul da Bahia são diferentes em sua ancestralidade genômica. Dessa forma, objetivou-se estimar as contribuições africanas, europeias e ameríndias nas amostras analisando oito MIAs (Sb19.3, APO, AT3, RB2003 e LPL, CKM, PV92 e DRD2) e correlacionar a etnia autodeclarada com ancestralidade genômica.

2 REVISÃO BIBLIOGRÁFICA

2.1 CARACTERÍSTICAS DA POPULAÇÃO

A população brasileira é o resultado de cinco séculos de mistura étnica entre populações, formada pelos subgrupos étnicos: ameríndios, imigrantes da Europa e da África, tornando-se assim altamente heterogênea (PENA *et al.*, 2009). A forma como esses grupos distribuíram-se e contribuíram para a formação das populações das diversas regiões do país e em comunidades específicas reflete na distribuição diferencial de fenótipos entre cada um (BATTILANA *et al.*, 2006; LUIZON *et al.*, 2008; LINS *et al.*, 2010; PENA *et al.*, 2011).

Quando aproximadamente 500 mil portugueses, em sua maioria homens, aportaram na costa brasileira em 1500 (RIBEIRO, 1995), perto 2 milhões de índios aqui viviam (CALLEGARI-JACQUES; SALZANO, 1999). O contato entre as duas etnias gerou a redução da civilização indígena, a qual havia em 2000 aproximadamente 460 mil indivíduos distribuídos em 225 grupos indígenas diferentes (FUNAI, 2010). O convívio instável entre as civilizações causou o extermínio dos índios por estes resistirem ao trabalho na construção das primeiras vilas e na defesa do território, além de sofrerem com as doenças dos europeus (TAVERES, 2001; FUNAI, 2010). Em contrapartida dos conflitos, cruzamentos entre portugueses e ameríndias ocorriam em larga escala, em parte pela escassez de mulheres europeias, mas também por incentivo da Coroa portuguesa como estratégia para facilitar o povoamento (PENA *et al.*, 2000).

Devido à diminuição da população indígena, provocada pelos conflitos, foi necessária a introdução de escravos trazidos de diferentes regiões da África como mão de obra alternativa (TAVERES, 2001; KLEIN, 2002). 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 (GATTÁS; GOMES;

KOHLER, 2004). 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 origem dos africanos que foram trazidos para o Brasil foi bastante diversa, envolvendo pessoas vindas da costa ocidental (do atual Senegal até a atual Angola), o interior, vales e terras das bacias dos rios: Congo, Níger e Benin e da costa oriental (do atual Moçambique até a Etiópia) (TAVERES, 2001), sendo que a maioria originou-se das regiões atuais do Congo e Angola (KLEIN, 2002).

Além dos africanos, ameríndios e portugueses, a população brasileira foi constituída em menor proporção por outros grupos de imigrantes. 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 denominado “embranquecimento do Brasil” (PENA *et al.*, 2009).

Todos os eventos de imigração descritos criaram uma população extremamente heterogênea. Entretanto, episódios históricos, geográficos, políticos e socioculturais fizeram com que o povoamento do território fosse desigual, marcando cada região com diferentes contribuições ameríndia, europeia e africana (BATTILANA *et al.*, 2006; LUIZON *et al.*, 2008; LINS *et al.*, 2010; PENA *et al.*, 2011). 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 na contribuição genética entre as populações.

A Bahia possui a mesma composição étnica do Brasil, já que foi o berço da colonização do país. Este estado possui uma população altamente diversa, composta por três grupos principais: ameríndios, africanos e europeus. Os escravos foram trazidos para a Bahia, por volta de 1549, após o estabelecimento dos primeiros engenhos de açúcar pertencentes aos colonos europeus (TAVERES, 2001). Contudo, ocorreu uma miscigenação mais intensa por parte de europeus e africanos, uma vez que a população de índios da região foi muito reduzida ao longo dos anos

impossibilitando, assim a união entre indivíduos dessa etnia com descendentes de africanos e europeus. De acordo com o IBGE (censo 2010), a Bahia é o estado com maior população de afrodescendentes, 77,4% (cor- preta, parda, amarela), porém, estes dados são de acordo com a autodeclaração, desta forma, a cor da pele autodeclarada é um caráter insuficiente para inferir ancestralidade populacional/individual (PARRA *et al.*, 2003).

A região Sul da Bahia é constituída por 26 municípios, sendo eles: Almadina, Arataca, Aurelino Leal, Barro Preto, Buerarema, Camacã, Canavieiras, Coaraci, Floresta Azul, Gongogi, Ibicaraí, Ilhéus, Itabuna, Itacaré, Itajú do Colônia, Itajuípe, Itapé, Itapitanga, Jussari, Mascote, Pau Brasil, Santa Cruz da Vitória, Santa Luzia, São José da Vitória, Ubaitaba e Una (IBGE, 2009). Sendo que cada uma dessas cidades teve uma história de formação populacional diferente, bem como um padrão de mistura étnica distinto. Dessa forma, será descrita sucintamente a história de formação dos principais municípios da região sul da Bahia (Ilhéus e Itabuna), por suas populações serem objeto de estudo do presente trabalho.

A fundação do povoado que originou a cidade de Ilhéus está ligada à introdução dos sistemas de capitânicas hereditárias no Brasil durante o século XVI. A capitania de São Jorge dos Ilhéus foi doada em 1536 ao donatário Jorge de Figueiredo Correia (SALES, 1981). Os grupos étnicos que deram origem à formação da população de Ilhéus são os índios que habitavam o Brasil na época da colonização dos portugueses, africanos e europeus. Os grupos indígenas que habitavam a capitania eram: os *tupis* (tribo dos tupiniquins) que moravam no litoral, e os *jês* ou *tapuia* (tribo dos aimorés), que viviam no interior (BARBOSA, 1994). Mais tarde, outras tribos indígenas descendentes dos tapuias foram aparecendo no litoral, como os *gueren*, *pataxó* e *kamakã* (ANDRADE, 2003). Os portugueses não foram o único povo europeu a contribuir com a formação de Ilhéus. Houve influência de espanhóis, alemães, franceses, árabes, entre outros (BARBOSA, 1994).

A cidade de Itabuna teve surgimento mais recente como um povoado em 1818, chamado Ferradas, nele havia aproximadamente 13 famílias de índios Gueren e 120

Camacan. Em meados de 1829 os índios em Ferradas não passavam de 90 (ANDRADE; ROCHA, 2005). Nessa época teve início o cultivo do cacau, o que provou conflitos entre os indígenas e colonizadores. Os negros, descendentes de escravos que chegaram a essa região, foram atraídos pela necessidade de mão de obra nas roças de cacau e para os serviços domésticos dos fazendeiros. Os mulatos e negros vindos de Sergipe e do sertão da Bahia, em 1850, foram atraídos pela riqueza da região (ANDRADE; ROCHA, 2005).

Por conta do processo de miscigenação intenso que ocorreu em todo o Brasil, a cor da pele, autorreferida, tem uma correlação muito fraca com o grau de ancestralidade africana. No nível individual, qualquer tentativa de previsão torna-se impossível, ou seja, pela observação da aparência física de um brasileiro, não se pode chegar a nenhuma conclusão confiável sobre seu grau de ancestralidade africana. Portanto a cor da pele é um marcador insuficiente de ancestralidade (PARRA *et al.*, 2003).

Diante de todo o contexto histórico citado anteriormente, é importante esclarecer que muitas populações humanas miscigenadas têm sido formadas pela mistura de duas ou mais populações ancestrais, as quais são geneticamente distintas (PATTERSON *et al.*, 2010). Esse processo de miscigenação torna a população estratificada ou subdividida, onde a proporção de mistura varia entre indivíduos (HOGGART *et al.*, 2003). Assim a frequência de determinados alelos varia de acordo com o padrão e o tempo de miscigenação durante a formação de uma população. Logo, quando não se identifica a estratificação populacional podem surgir associações espúrias entre um marcador e um determinado fenótipo (HOGGART *et al.*, 2003; HINDS *et al.*, 2004).

Dessa forma, em populações como a brasileira, a existência de estruturação populacional pode levar a resultados falso-positivos ou falso-negativos, não permitindo detectar a real associação existente entre os genes candidatos e o fenótipo em questão. Visto que, tal gene pode estar associado a uma etnia específica e que esta não seja a mesma da amostra analisada, assim não será encontrada nenhuma associação significativa entre o gene candidato e a doença (BARNHOLTZ-SLOAN *et*

al., 2005). Por isso, o controle genômico é um passo indispensável nesses estudos populacionais. Entende-se por controle genômico a caracterização da proporção de ancestralidade individual e global de uma amostra populacional caso/controles (SANTOS *et al.*, 2009), tornadas as mais homogêneas possíveis.

Os estudos caso/controles que investigam frequências aumentadas de alelos específicos em pessoas afetadas em comparação com as não afetadas na população (NUSSBAUM *et al.*, 2002) têm sido amplamente empregados em populações miscigenadas, as quais apresentam uma formação histórica genética complexa (PATTERSON *et al.*, 2010).

Como mencionado, a estratificação da população pode ocorrer quando há diferenças nas frequências alélicas entre casos e controles relacionados à ancestralidade, em vez do “risco de doença” (MYLES; STONEKING; TIMPSON, 2009). Isto pode ser observado nos casos em que o risco da doença é diferente entre os grupos com ancestralidade diferente. Por exemplo, estudos demonstram que determinados polimorfismo genéticos candidatos para o câncer de próstata são mais frequentes em indivíduos de ascendência africana em comparação com indivíduos de ascendência europeia (GIRI *et al.*, 2009; BONILLA *et al.*, 2011; OKOBIA *et al.*, 2011).

As características fenotípicas como a cor da pele, classificadas de acordo com o IBGE é um indicador insuficiente de ancestralidade genômica africana (PARRA *et al.*, 2003). Sendo assim, conhecer a ancestralidade genômica de amostras miscigenadas torna-se importante, pois permite inferir sobre o padrão de distribuição de alelos associados a doenças entre os indivíduos de etnias diferentes.

O método inicialmente proposto por Pritchard e Rosenberg (1999) no qual os marcadores genéticos não ligados, entre si e com a doença são largamente utilizados em trabalhos com amostras de casos e controles. Esses marcadores são empregados para identificar diferenças não esperadas nas frequências gênicas, como as diferenças produzidas pela etnia. Assim, essa metodologia pode evidenciar uma possível estratificação populacional (ZEMBRZUSKI; CALLEGARI-JACQUES; HUTZ, 2006).

2.2 MARCADORES MOLECULARES

Até a década de 80, os estudos sobre a variabilidade e a estrutura das populações baseavam-se nas análises dos parâmetros demográficos e dos polimorfismos proteicos. Com o desenvolvimento das técnicas de Biologia Molecular tornou-se possível a análise direta do material genético, e vários métodos tem sido empregados para detectar alterações do genoma em regiões codificadoras e não-codificadoras. O estudo do DNA das populações atuais é uma poderosa fonte de informação que permite fazer inferências sobre eventos que ocorreram no passado, como migrações, gargalos populacionais e expansões demográficas (CAVALLI-SFORZA; MENOZZI; PIAZZA, 1996).

Existem vários tipos de polimorfismos genéticos no DNA e estes são classificados de acordo com a sua natureza molecular e localização no genoma. A maioria (de 80% a 90%) da variação genética entre humanos é intrapopulacional e somente 10% a 20% da variação total é devida as diferenças populacionais (CAVALLI-SFORZA; MENOZZI; PIAZZA, 1994). Os valores de diferencial de frequência (δ) médios para *loci* bialélicos entre os principais grupos étnicos estão entre 15-20%, e a vasta maioria (>95%) dos *loci* marcadores bialélicos arbitrariamente identificados possuem $\delta < 50\%$ (DEAN *et al.*, 1994).

Alelos com valor de $\delta > 50\%$ entre grupos populacionais definidos geográfica e/ou etnicamente passaram a ser denominados de PSAs (do inglês *Population Specific Alleles*) (SHRIVER *et al.*, 1997). Estes marcadores apresentam potencial evidente e elevado para discriminar grupos populacionais e são, portanto, ferramentas eficientes para estimativas de mistura étnica (LINS *et al.*, 2010; BONILLA *et al.*, 2011). Embora existisse uma limitação para esse tipo de estudo devido ao número limitado de marcadores clássicos ou de DNA apropriados, com o advento dos PSAs tornou-se possível gerar estimativas mais precisas das proporções ancestrais de uma população miscigenada (PARRA *et al.*, 1998; SHRIVER *et al.*, 2003; ERDEI *et al.*, 2011; PENA *et al.*, 2011).

Esses marcadores são atualmente denominados de AIMS (do inglês *Ancestry Informative Markers*) (BONILLA *et al.*, 2004). Além disso, algumas pesquisas veem demonstrando que é possível utilizar um número menor de marcadores com um diferencial <50%. Bonilla *e cols.* (2004) em seu trabalho utilizaram 22 MIAs para descrever o processo e a dinâmica de mistura na população hispânica de San Luis Valley (Colorado, Estados Unidos), estabelecendo um diferencial de frequência $\geq 30\%$ entre os grupos populacionais.

Os MIAs (Marcadores Informativos de Ancestralidade) apresentam grandes diferenças de frequência alélica entre populações europeias, africanas e asiáticas e são por isso considerados ideais para alocar amostras individuais em seus grupos étnicos correspondentes (LINS *et al.*, 2010). Os MIAs podem pertencer a qualquer classe molecular, como inserções/deleções, inserções *Alu* e polimorfismos de um único nucleotídeo (SNPs), todos com taxas de mutação diferentes e, portanto, com valores informativos diferentes.

Os elementos *Alu* são um tipo de elementos transponíveis, que existe apenas em primatas (GROVER *et al.*, 2004). As sequências *Alu* são os mais abundantes elementos móveis do genoma humano, compreendendo até 28% deste (GASIOR *et al.*, 2007). Esses fazem parte da família de DNA repetitivo SINE- *Short Interspersed Nuclear Elements* (ROUZIC; DECELIERE, 2005), que representam uma das muitas famílias de retrotransposons não autônomos que dependem da maquinaria enzimática dos elementos autônomos LINE- *Long Interspersed Nuclear Elements* (JURKA *et al.*, 2004).

Essas inserções polimórficas têm sido utilizadas como marcadores genéticos em estudos de evolução humana (BATZER; DEININGER, 1991), devido às suas propriedades particulares: como rapidez e facilidade na genotipagem, serem seletivamente neutros e terem um estado ancestral conhecido (COMAS *et al.*, 2001). Estes polimorfismos são altamente informativos em estudos de genética de populações, permitindo sua utilização como instrumento de investigação para caracterizar a composição genética de diferentes povos, uma vez que possui frequências diferenciadas para cada etnia. Segundo Batzer e Deininger (2002), se duas pessoas têm a mesma sequência *Alu* no mesmo local do seu genoma, elas descendem

de um antepassado comum que lhes forneceu aquele segmento específico de DNA. Isto porque, é improvável que o evento de inserção ocorra duas vezes em um mesmo *locus*.

Polimorfismos de nucleotídeo único (SNPs) são pequenas variações na sequência de DNA causadas por mutações pontuais do tipo substituição. É a classe de marcadores mais comum no genoma humano, representando cerca de 90% das diferenças de sequência (COLLINS; BROOKS; CHAKRAVARTI, 1998). Existe aproximadamente um nucleotídeo diferente a cada 1200 nucleotídeos entre dois cromossomos homólogos. Logo, há 3 milhões de diferenças de uma base e cerca de 10.0000 aminoácidos diferentes entre quaisquer duas pessoas (MICKLOS; FREYER; CROTTY, 2005). Entretanto, a maior parte das mutações de um nucleotídeo único é rara em uma população, para ser considerado SNP, o polimorfismo deve apresentar uma frequência de pelo menos 1% na população.

Populações diferentes têm acumulado diferentes SNPs dentro de uma dada região por conta do desequilíbrio de ligação, formando haplótipos. Assim sendo, alguns marcadores são característicos de populações específicas (MICKLOS; FREYER; CROTTY, 2005). Um dos principais objetivos de pesquisas com SNP é compreender a genética da variação fenotípica humana e, especialmente, a base genética das doenças humanas complexas (RAMENSKY; BORK; SUNYAEV, 2002; GARCIA *et al.*, 2007; PALMER; SCHUTTE; HUISMAN, 2009; STEWART *et al.*, 2009; ABU-AMERO *et al.*, 2011).

A análise de populações miscigenadas utilizando os MIAs permite caracterizar a ancestralidade genômica de populações e indivíduos, assim tornando útil para homogeneizar amostras caso/controles que apresentam manifestação variável de uma determinada doença em diferentes grupos étnicos (BOCK *et al.*, 2009; ERDEI *et al.*, 2011). Uma dessas evidências foi observada no trabalho POOKOT *e cols.* (2006), que mostrou que a variação alélica do SNP -160 C/A do gene da E-caderina pode ser associada com a incidência e a severidade do câncer de próstata de acordo com as diferenças “raciais”.

2.3 CÂNCER DE PRÓSTATA E ETNIA

O câncer (Ca) neoplasia maligna é a segunda principal causa de morte em países economicamente desenvolvidos perdendo apenas para as doenças cardíacas, e a terceira principal causa de morte nos países em desenvolvimento, sendo antecipado pelas doenças cardíacas e diarreicas (GARCIA *et al.*, 2007).

O Ca é uma doença multifatorial que pode ser desencadeada por fatores genéticos e/ou ambientais. As causas do Ca podem ser tanto por fatores externos: químicos, físicos ou biológicos, como fatores internos: mutações herdadas, hormônios, variação genética no processo inflamatório (GARCIA *et al.*, 2007; MASON *et al.*, 2010).

De acordo com o INCA- Instituto Nacional do Câncer (2011):

O câncer de próstata (CaP) é o segundo mais comum entre os homens, sendo o primeiro Ca de pele não-melanoma, no Brasil, e o sexto tipo mais comum no mundo e o mais prevalente em homens, representando cerca de 10% do total de cânceres. Sua taxa de incidência é cerca de seis vezes maior nos países desenvolvidos em comparação aos países em desenvolvimento.

No Brasil, o número de mortes por CaP foi de aproximadamente 12.274 no ano de 2009. Em 2012, o número de casos novos estimado será de 60.180.

O câncer de próstata (CaP) é um dos principais problemas de saúde na população dos Estados Unidos (BONILLA *et al.*, 2011). Os afro-americanos são desproporcionalmente afetados por CaP (KAMOTO *et al.*, 2005; GIRI *et al.*, 2009; TROTTIER *et al.*, 2010; BONILLA *et al.*, 2011; OKOBIA *et al.*, 2011), apresentando um prognóstico da doença mais tardio e severo que euro-americano (REDDY *et al.*, 2003). Vários fatores genéticos e ambientais influenciam as diferenças observadas entre os cânceres, entre elas destacam-se: etnia, idade, história familiar da doença, condição socioeconômica e acesso à saúde (WHITMAN *et al.*, 2010).

Em análise com amostra de eurodescendentes e afrodescendentes norte Americanos, Pookot *e col.* (2006) mostraram que a variação alélica do polimorfismo -160 C/A da E-caderina pode estar associado com as diferenças étnicas na incidência e

severidade do CaP. Na população caucasiana a frequência do alelo A foi maior nos pacientes com CaP em relação aos controles. Já na população afrodescendente este quadro se inverteu e a presença do alelo A foi associada a uma diminuição de 2,4 vezes o risco para o CaP. Com base nesses resultados os autores concluíram que o alelo A é fator de risco para homens com ascendência europeia e fator de proteção para homens com ascendência africana.

Considerando a heterogeneidade da população baiana também é interessante a utilização de marcadores genéticos neutros para caracterização de ancestralidade e realização do controle genômico em estudos de associação do tipo caso/controles em amostras miscigenadas para o câncer de próstata da região sul da Bahia.

3 METODOLOGIA

3.1 AMOSTRAGEM

Foi utilizada amostra previamente coletada pelo Msc. Leonardo Moreira Santos durante seu projeto de mestrado (SANTOS, 2010). A amostra foi composta de 102 homens, diagnosticados com câncer de próstata, na clínica ONCOSUL em Itabuna e na Clínica Oncológica de Ilhéus – CLIONI. Foram realizadas três campanhas públicas nas cidades de Itabuna e Ilhéus em dias alternados, com a finalidade de conscientização e prevenção ao câncer, fornecendo exames de PSA gratuito e aproveitando para convidar pessoas a aderirem como voluntários nesta pesquisa bem como para explicá-la. Dessas campanhas selecionou-se somente 109 para compor a amostra controle. O critério adotado para inclusão nos controles foi selecionar os indivíduos sem histórico de câncer na família com parentesco de até 2º grau e que apresentavam PSA (do inglês *Prostate Specific Antigen*) abaixo de 3 ng/mL.

3.2 EXTRAÇÃO DE DNA

Cinco mililitros de sangue periférico foram obtidos de uma única punção e separados em dois tubos a vácuo, sendo que um deles continha anticoagulante EDTA em seu interior. A amostra de sangue do tubo seco (sem EDTA) foi utilizada nos respectivos laboratórios para realização da dosagem de PSA total e o tubo contendo EDTA foi levado ao Laboratório de Farmacogenômica e Epidemiologia Molecular da UESC para extração de DNA. O EDTA foi usado para impedir a coagulação do sangue no interior dos tubos, facilitando assim as extrações do DNA. Essa extração foi realizada de acordo com as especificações do kit de extração FlexiGene da Qiagen e o DNA foi quantificado no aparelho Gene-Quant®.

3.3 ASPECTOS ÉTICOS

O projeto de pesquisa do presente trabalho foi aprovado em seus aspectos éticos por meio de um adendo pela Comissão de Ética em Pesquisa da Universidade Estadual de Santa Cruz de acordo com o processo n° 11/2011. A coleta de informações socioeconômicas e de material biológico das populações envolvidas ocorreu anteriormente durante o desenvolvimento do projeto: "*Genotipagem de polimorfismos no gene de E-caderina em pacientes portadores de câncer de próstata, em Itabuna, Itapetinga e Vitória da Conquista, Bahia*", seguindo rigorosamente os preceitos de conscientização e autorização dos participantes, aquiescência voluntária e garantia de anonimato dos participantes, segundo processo aprovado pelo comitê de ética da UESC sob o número 180/2008.

3.4 MARCADORES MOLECULARES ANALISADOS

Neste trabalho foram analisados oito marcadores que se enquadram na definição de MIAs, cinco (RB, LPL, AT3, Sb19.3 e APO) com valores de $\delta > 40\%$ entre Africanos e Europeus ou Ameríndios (PARRA *et al.*, 1998; SHRIVER *et al.*, 2003) e três (PV92, CKMM e DRD2) com $\delta > 40\%$ entre Ameríndios e Africanos ou Europeus (SHRIVER *et al.*, 2003).

Para os *loci* dos marcadores analisados no presente estudo, foi denominado de alelo *1 (ex., APO * 1) o fragmento correspondente à banda de maior peso molecular observada no gel de poliacrilamida ou agarose, devido à presença de uma inserção (para os *loci* AT3-I/D, APO, PV92 e Sb19.3) ou à ausência de um sítio de restrição (para os *loci*, RB2300, DRD2,CKMM e LPL), seguindo a convenção estabelecida em Parra *et al.*,(1998).

Os *primers*, enzimas de restrição, a localização citogenética dos *loci* analisados no presente estudo estão relacionados na Tabela 1.

Tabela 1: Localização citogenética, enzimas de restrição e sequência dos *primers* dos oito MIAs analisados.

Locus	Tipo	LOCALIZAÇÃO	SEQUÊNCIA DOS PRIMERS
AT3-I/D	76pb indel	1q25.1	CCACAGGTGTAACATTGTGT GAGATAGTGTGATCTGAGGC
LPL	<i>PvuII</i>	8p21.3	AGGCTTCACTCATCCGTGCCTCC TTATGCTGCTTTAGACTCTTGTC
RB2300	<i>BamHI</i>	13q14.2	CAGGACAGCGGCCCGGAG CTGCAGACGCTCCGCCGT
Sb19.3	Inserção <i>Alu</i>	19p12	TCTAGCCCCAGATTTATGGTAACTG AAGCACAATTGGTTATTTTCTGAC
APO	Inserção <i>Alu</i>	11q23	AAGTGCTGTAGGCCATTTAGATTAG AGTCTTCGATGACAGCGTATACAGA
PV92	Inserção <i>Alu</i>	16q23.3	AACTGGGAAAATTTGAAGAGAAAGT TGAGTTCTCAACTCCTGTGTGTTAG
CKMM	<i>TaqI</i>	19q13.32	GCAGGCGCCTACTTCTGG AGCTCATGGTGGAAATGGAG
DRD2	<i>Bcl I</i>	11q23.1	CACCAGCAGCCCTGCCTATC TCTGGAGGCCCTGTTTCTT

3.5 REAÇÃO EM CADEIA DA POLIMERASE (PCR)

Os ensaios da PCR foram realizados em um volume total de 25 μ L. Todos os reagentes (Água, Tampão, dNTP – solução trabalho, *primers* – solução trabalho, MgCl₂ e a Taq DNA polimerase), com exceção do DNA, foram misturados em quantidade específica para cada *locus* em um único tubo (mistura de reação) para garantir a homogeneidade das reações. Foram utilizados 16 *primers* para os oito marcadores.

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	DMSO	MgCl (50mM)	dNTP (20mM)	Primers (solução trabalho)	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 I/D	14,2	2,5	-	1,0	0,25	3,0	0,05
LPL	14,2	2,5	-	1,0	0,25	3,0	0,05
DRD2	14,2	2,5	-	1,0	0,25	3,0	0,05
RB 2300	11,9	2,5	2,5	0,75	0,25	3,0	0,05
CKMM	14,2	2,5	-	1,0	0,25	3,0	0,05

Em cada microtubo de 0,2 mL foram pipetados 4 μL do DNA genômico, previamente extraído. Para cada análise foi usado controle negativo contendo água no lugar do DNA genômico. Em cada microtubo, sob a amostra, foram pipetados 21 μL da mistura de reação. O termociclador utilizado foi o Biocycler. A este passo seguiu-se o programa correspondente a cada *locus*. Após o término da reação de PCR o produto da PCR foi guardado em geladeira (4C°) até sua utilização.

3.6 REAÇÃO DE RESTRIÇÃO

Para a detecção dos SNPs analisados no presente estudo (LPL, CKM, DRD2, RB2300,) foi utilizada a técnica da PCR-RFLP. Esta consiste basicamente de um ensaio onde o produto da PCR é posteriormente submetido à ação de endonuclease de restrição. Os *primers* utilizados na PCR flanqueiam a região do gene que contém, ou não, o sítio de restrição específico da enzima utilizada na reação de restrição para cada *locus*.

O indivíduo pode ser homozigoto para a presença do sítio de restrição da enzima (genótipo *2/*2), heterozigoto (genótipo *1/*2), ou homozigoto para a ausência do sítio de restrição (genótipo *1/*1).

A reação de restrição foi feita em microtubos tipo *ependorf* de 500µL para cada amostra, ambas as séries contendo 5 µL de DNA previamente amplificado. O controle de reação foi o DNA amplificado de um indivíduo ao qual foi submetido somente água e o tampão comercial específico da enzima. O controle de leitura é o DNA amplificado de um indivíduo de genótipo conhecido.

Em microtubos tipo *ependorf* de 1,5mL foi preparada a mistura de reação, onde o volume de cada componente utilizado foi multiplicado pelo número de mostras utilizadas na reação. A mistura de reação foi constituída por 0,3 µL (3U) da enzima de restrição (10U/ µL), 1,0 µL do tampão de reação (específico para cada enzima) e 5,7 µL de água deionizada autoclavada. O tampão de reação continha as condições de pH e salinidade ótimas para a atividade da enzima de restrição, sendo estas condições específicas para cada enzima.

Aplicou-se 7 µL da mistura de Reação em cada *ependorf* com o auxílio de uma micropipeta e ponteiras estéreis. Estes foram então levados a um banho Maria ou ao Termociclador, à temperatura de atividade específica para a enzima de restrição utilizada (geralmente 37°C), por um período de 1 a 2 horas.

Após este período os tubos foram retirados do banho Maria e realizou-se um choque térmico para inativação da enzima com água e gelo. Os tubos foram conservados em geladeira, sendo os fragmentos resultantes da reação de Restrição submetidos à eletroforese para detecção da presença ou ausência do sítio de restrição investigado.

3.7 ANÁLISE DO PRODUTO AMPLIFICADO

Os produtos amplificados para os marcadores de inserção *Alu* (Sb19.3, APO, AT3, PV92) e os SNPs (DRD2 e LPL) foram separados em gel de agarose a 2 e 3%. Os outros SNPs (CKMM e RB2300) oriundos das reações de restrição foram separados por eletroforese em géis de poliacrilamida (PAGE) não- desnaturante a 12% de concentração.

3.7.1 Reagentes e Soluções para os géis de acrilamida

Solução de acrilamida/bis- acrilamida (29:1): 29g de acrilamida; 1g de bis-acrilamida diluídas em 100mL de água deionizada.

Solução de EDTA pH 8,0: 186g de EDTA; 1L de H₂O. Acertar o pH com pastilhas de hidróxido de sódio (NaOH).

Solução saturada de persulfato de amônio (20%): 650mg de amônia de potássio; 6,5mL de H₂O.

TEMED: tetrametilenodiamina

Tampão TBE 10X (0,9M) pH 8,0: 108g de Tris (PM=121,1); 53 g de ácido bórico; 40mL de solução de EDTA; H₂O qsp 1L.

Tampão TBE cubas (1X): 100mL do tampão TBE (10X); 900mL de H₂O.

Gel não desnaturante 6%: 21mL de H₂O; 6mL solução de bis-acrilamida; 3mL TBE (10X); 20 µL TEMED; 250 µL solução de persulfato de amônia.

Gel não desnaturante 12%: 15mL de H₂O; 12mL solução de bis-acrilamida; 3mL TBE (10X); 20 µL TEMED; 250 µL solução de persulfato de amônia.

Gel desnaturante 12%: 9,6g de ureia; 4mL de H₂O; 8mL de solução de bis-acrilamida; 2mL de TBE (10X); 15 µL TEMED; 300 µL solução de persulfato de amônia.

Procedimento:

Os géis não-desnaturantes 12% foram feitos como no protocolo acima determinado. Os catalisadores da reação de polimerização do gel, TEMED e persulfato de potássio foram adicionados à mistura do gel imediatamente antes de vertê-la em um cassete previamente montado, composto de duas placas de vidro de 12 cm de altura por 16,5 cm de largura, separadas por espaçadores de teflon e presas com grampos. Logo após, um pente de teflon foi colocado na borda superior, formando poços no gel,

onde posteriormente foram aplicadas as amostras de DNA amplificado por PCR ou oriundos das reações de restrição. Aguardou-se a polimerização por cerca de 30 minutos.

Após a polimerização de gel o pente foi retirado e os poços foram lavados com água. O gel polimerizado foi montado em cuba de eletroforese vertical contendo tampão TBE cubas (1X), em ambos os polos (porção superior e inferior). Esta cuba foi conectada a uma fonte de voltagem e ajustada à voltagem constante de 200V, necessária para uma separação dos fragmentos amplificados.

Aproximadamente 7 μ L de amostra de DNA amplificado oriundo da PCR ou Reação de restrição, juntamente com 4 μ L de tampão amostra, foram aplicados nos géis. Para os MIAs que são SNPs (LPL) uma eletroforese prévia era realizada, no intuito de confirmar a amplificação positiva do fragmento de interesse, antes de submetê-lo à ação de enzima de restrição adequada.

Após a aplicação das amostras, as fontes foram novamente ligadas e a eletroforese prosseguiu da maneira descrita acima. Com término da corrida eletroforética, o gel foi retirado cuidadosamente das placas de vidro e submetidos aos procedimentos de coloração e secagem.

3.7.2 Reagentes e Soluções para os géis de AGAROSE

Agarose: AGR-LE-500 (Uniscience)

Tampão TBE 1X: 100mL do tampão TBE (10X); 900mL de H₂O.

GelRed: GelRed 10.000x in Water.

GelRed 10X: 1mL de H₂O; 1 μ L de GelRed 10.000X

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

Tampão da amostra: 2 μ L de GelRed 10X; 1 μ L de azul de bromofenol, para cada amostra.

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

Gel a 3%: 3g de agarose; 100ml de TBE 1X; 1 μ L 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 micro-ondas por aproximadamente 1 minuto, até a 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 100mL, um suporte para gel e um pente colocado na borda superior e no meio, formando poços no gel, onde posteriormente foram aplicadas as amostras de DNA amplificado por PCR ou oriundos das reações de restrição. Aguardou-se a polimerização por 15 minutos.

Após a polimerização do gel o suporte foi retirado da bandeja, em seguida o pente. O gel polimerizado foi montado em cuba de eletroforese horizontal contendo tampão TBE cubas (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 μ L de amostra de DNA amplificado oriundo da PCR ou Reação de restrição, juntamente com 3 μ L de tampão amostra, foram aplicados nos géis. Para os MIAs que são SNPs (RB2300, CKMM, DRD2 e LPL) uma eletroforese prévia era realizada, no intuito de confirmar a amplificação positiva do fragmento de interesse, antes de submetê-lo à ação de enzima de restrição adequada.

Após a aplicação das amostras, as fontes foram ligadas e a eletroforese prosseguiu da maneira descrita acima. 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).

3.7.3 Coloração com nitrato de prata e secagem do gel**Regentes e Soluções:**

Solução de nitrato de prata: 10g nitrato de prata; 100mL de H₂O . Dissolver a prata em uma parte da água e depois completar com o restante, manter a solução ao abrigo da luz (volume final 100mL).

Solução fixadora: 160mL etanol (PA) e 7mL de ácido acético glacial (PA); 833mL de H₂O (volume final 1L).

Solução reveladora: 22,5g de NaOH; 1L de H₂O. Dissolver em um agitador o hidróxido de sódio em uma parte da água e depois completar com o restante (volume final 1L). Na hora da coloração adicionar 1mL de formaldeído par cada 100mL da solução.

Procedimento:

A coloração do gel foi feita de acordo com protocolo adaptado de Sanguinetti *et al.*,(1994):

Fixação: após a retirada das placas de vidro e dos espaçadores o gel foi colocado em um recipiente de vidro contendo 100mL de solução fixadora.

Impregnação com nitrato de prata: adicionou-se 2,0mL de solução de nitrato de prata, e agitou-se por 5 minutos. A solução foi então descartada e o gel lavado em água quente por cerca de 10 segundos, agitando levemente e, ao final, descartando a água.

Revelação: a solução reveladora foi despejada cuidadosamente no recipiente contendo o gel, que foi submetido à agitação por alguns minutos até que as bandas aparecessem nitidamente. A solução foi pré-aquecida em estufa a 65°C, par facilitar a reação de coloração.

Bloqueio da reação: após ter sido revelado, a solução reveladora foi descartada e a reação bloqueada com a lavagem direta do gel em 100mL de solução fixadora.

Secagem do gel: após a leitura, todos os géis passaram por um simples processo de secagem para que pudessem ser armazenados para análises e confirmações posteriores. Duas folhas de papel celofane foram molhadas; uma placa de vidro, com a área maior que a do gel, foi coberta com uma das folhas; o gel foi colocado sobre a placa com o celofane sem deixar bolhas; o gel foi então bem molhado e coberto com a outra folha de celofane, também com cuidado de não deixar bolhas; este gel foi deixado secando à temperatura ambiente por dois ou três dias até secagem completa, sendo então devidamente identificado e arquivado.

3.8 PERFIL ELETROFORÉTICO

3.8.1 APO

O *locus* para a inserção *Alu* APO se situa no braço longo do cromossomo 11 (BATZER *et al.*, 1999), na região não codificadora próxima ao complexo de genes da apolipoproteína AI-CIII-AIV (KARATHANASIS, 1985). Sua amplificação pela PCR produz um fragmento de 400pb (presença da inserção) e outro de 110pb (ausência da inserção) (NOVICK *et al.*, 1995).

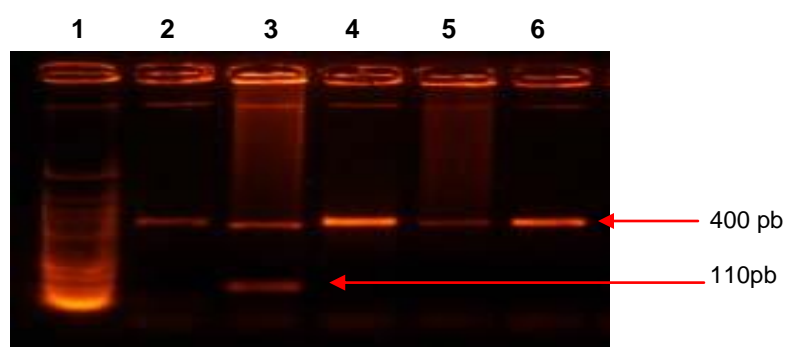


Figura 1. Padrão de bandas observado para o marcador APO. Gel de agarose 2%, corado com GelRed. Na raia 1 o *Ladder* de 50pb, nas raias 2, 4 5 e 6 encontram-se os homocigotos para a inserção (+/+) e na raia 3 heterocigoto (-/+).

3.8.2 SB19.3

A inserção *Alu* Sb19.3 pertence a subfamília Yb8 e está localizada no cromossomo 19p 12 (ARCOT *et al.*, 1998). A presença da inserção *Alu* gera um fragmento de ~450pb (Figura 7) e caracteriza o alelo *Sb19.3 *1*.

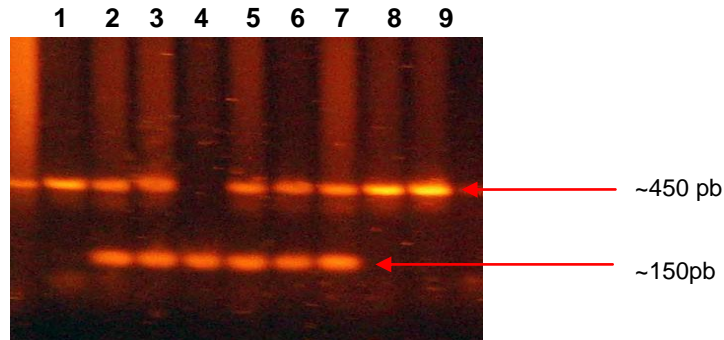


Figura 2. Padrão de bandas observado para o marcador SB19.3. Gel de agarose 2%, corado com GelRed. Nas raias 1, 8 e 9 encontram-se homozigotos para inserção (+/+), na raia 4, homozigotos para ausência da inserção (-/-), e nas raias 2, 3, 5, 6 e 7 heterozigotos (-/+).

3.8.3 PV92

O *locus* para essa inserção *Alu* encontra-se no cromossomo 16 (BATZER *et al.*, 1994). Sua amplificação por PCR produz um fragmento de aproximadamente 400pb (presença de inserção) e um de 110pb (ausência da inserção) (NOVICK *et al.*, 1995).

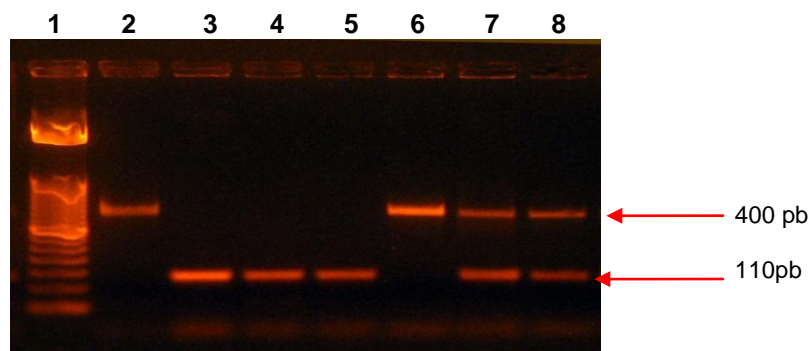


Figura 3. Padrão de bandas observado para o marcador PV92. Gel de agarose 2%, corado com GelRed. Na raia 1, o *Ladder* de 50pb, nas raias 2 e 6 encontram-se homozigotos para a inserção (+/+), nas raias 3, 4 e 5, homozigotos para a ausência da inserção (-/-), e nas raias 7 e 8 heterozigotos (-/+).

3.8.4 AT3

O polimorfismo do *locus* AT3 aqui analisado se caracteriza pela presença ou ausência de uma inserção de 76pb no exón 1 de gene da antitrombina III. O alelo*1 é observado pela presença da inserção, a mais pesada (572pb), e a mais leve, o alelo AT3*2.

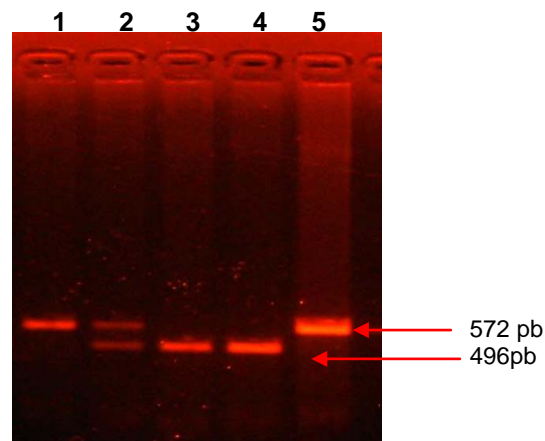


Figura 4. Padrão de bandas observado para o marcador AT3. Gel de agarose 3%, corado com GelRed. Nas raias 1 e 5 encontram-se homocigotos para a inserção (+/+), nas raias 3 e 4 homocigotos para a deleção (-/-), e na raia 2 heterocigotos (-/+).

3.8.5 LPL

Os *primers* utilizados para o *locus* LPL (Tabela 1) flanqueiam uma região de 319pb no intron 6 do gene da lipoproteína lipase (LPL). A presença de um sítio de restrição da enzima *PvuII* nesta região permite a detecção de dois fragmentos após a reação de restrição: 161pb e 158pb. A ausência do sítio da *PvuII*, e conseqüentemente a presença do fragmento de 319pb após a reação de restrição caracterizam o alelo *LPL* *1.

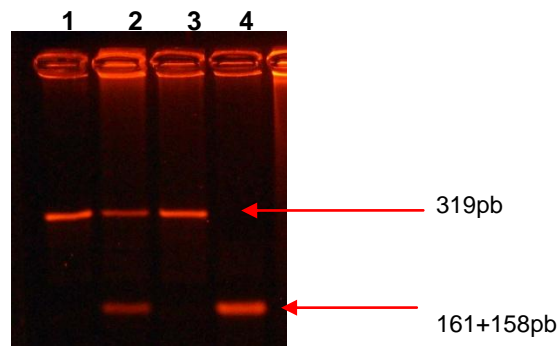


Figura 5. Padrão de bandas observado para o marcador LPL. Gel de agarose 2%, corado com GelRed. Nas raias 1 e 3 encontram-se homocigotos para ausência do sítio de restrição (+/+), na raia 4, homocigoto para a presença do sítio de restrição (-/-), e na raia 2 heterocigoto (-/+).

3.8.6 RB2300

Os *primers* utilizados para o *loci* RB2300 (Tabela 1) flanqueiam uma região de 180pb no intron 1 do gene do Retinoblastoma (RB) (BOOKSTEIN *et al.*, 1990). A presença de um sítio de restrição da enzima *Bam*HI nesta região permite a detecção de dois fragmentos após a reação de restrição: 130pb e 50pb.

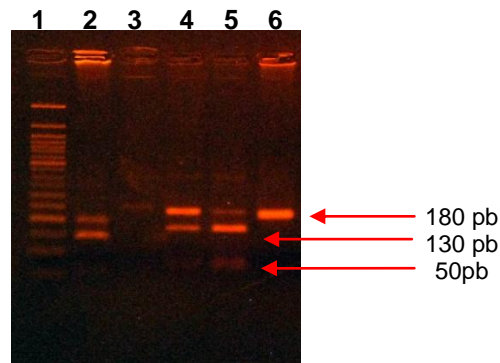


Figura 6. Padrão de bandas observado para o marcador RB200. Gel de agarose 2%, corado com GelRed. Nas raias 3 e 6 encontram-se homozigotos para ausência do sítio de restrição (+/+), na raia 2, 4 e 5 heterozigotos o sítio de restrição (-/+).

3.8.7 DRD2

Os *primers* utilizados para o *locus* DRD2 (Tabela 1) flanqueiam uma região de 291pb no gene 2 do receptor de dopamina (DRD2). A presença de um sítio de restrição da enzima *Bcl*I nesta região permite a detecção de dois fragmentos após a reação de restrição: 155pb e 136pb.

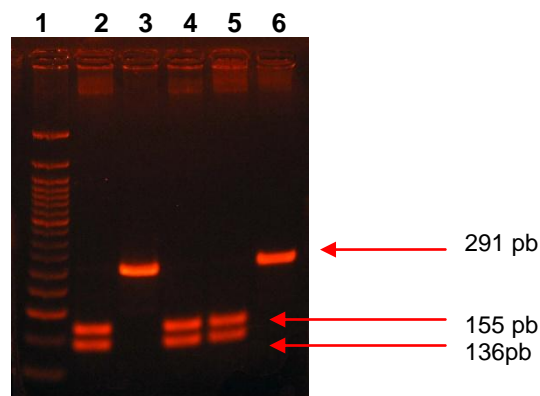


Figura 7. Padrão de bandas observado para o marcador DRD2. Gel de agarose 3%, corado com GelRed. Na raia 1 o *Ladder*, de 50pb, nas raias 2, 4 e 5 encontram-se homozigotos para o sítio de corte, nas raias 3 e 6, amostras não digeridas.

3.8.8 CKMM

Os *primers* utilizados para o *loci* CKMM (Tabela 1) flanqueiam uma região de 216pb do gene da Creatina cinase do músculo (CKM). A presença de um sítio de restrição da enzima *TaqI* nesta região permite a detecção de dois fragmentos após a reação de restrição: 170pb e 46pb.

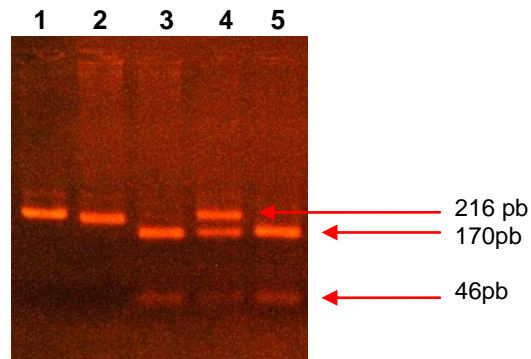


Figura 8. Padrão de bandas observado para o marcador CKM. Gel de agarose 2%, corado com GelRed. Nas raias 1 e 2 encontram-se homocigotos para ausência do sítio de restrição (+/+), na raia 3 e 5, homocigotos para o sítio de restrição (-/-), e nas raias 2 e 5, heterocigotos (-/+).

3.9 ANÁLISES DOS DADOS

3.9.1 Frequências Alélicas

As frequências alélicas de 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 MIAs estudados para populações da África, Europa e Ameríndios (Tabela 11). Conforme salientado anteriormente estão listados somente as frequências do alelo *1 para cada *locus*, de acordo com a convenção descrita em Parra *et cols.* (1998).

Foi construído um gráfico e uma tabela apresentando o diferencial de frequência (δ) dos *loci* MIAs estudados entre os grandes grupos mundiais (Tabela 11 e figura 12). A partir da média ponderada das frequências alélicas de Inglaterra, Irlanda, Alemanha (PARRA *et al.*, 1998), Portugal (TOMÁS *et al.*, 2002) e Espanha (dbSNPINCBI) para a Europa; da média ponderada das frequências alélicas de República Centro Africana, Nigéria (Benin e Ibadan) (PARRA *et al.*, 1998), e Serra Leoa (Mende e Temne) (PARRA *et al.*, 2001), para a África. E da média ponderada para as tribos indígenas da Amazônia (LUIZON, 2003). Para isso foi utilizado o programa Microsoft Excel®.

3.9.2 Aderência ao Equilíbrio de Hardy-Weinberg

Foram realizados dois tipos de testes para verificar a aderência ao equilíbrio de Hardy-Weinberg.

Testes globais realizados pelo teste exato de Fisher, assumindo-se independência entre os *loci* utilizados: o teste *multi-locus* (que analisa as frequências genotípicas de todos os *loci* em uma única população) e o teste multipopulacional (que analisa as frequências genotípicas de um único *locus*, em todas as populações). Para isto foi utilizado o programa GENEPOP (RAYMOND & ROUSSET 1995), e a hipótese nula 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) é excesso e deficiência de heterozigotos, respectivamente. Estas análises também foram feitas utilizando-se o programa GENEPOP (RAYMOND & ROUSSET 1995).

3.9.3 Associações par-a-par entre *loci*

A análise de associações par-a-par entre *loci* foi realizada utilizando-se o programa GENEPOP 2.0 (RAYMOND & ROUSSET 1995b). A hipótese nula é 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.

3.9.4 Diferenciação genética das populações

Os testes exatos para diferenciação populacional foram realizados com o uso do programa GENEPOP 2.0 (RAYMOND & ROUSSET 1995b). 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, onde hipótese nula testada é a de que a distribuição alélica é idêntica entre as populações (RAYMOND & ROUSSET 1995a).

3.9.5 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).

A estimativa foi realizada utilizando as frequências nas três populações ancestrais utilizando o ADMIX 3. Nos casos onde houve inconsistência com alguma das populações consideradas como ancestral esta foi retirada e se refez a estimativa com somente duas populações ancestrais utilizando o ADMIX 2. 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, forças microevolutivas atuando junto com a mistura (BORTOLINI, 1991).

4 RESULTADOS

4.1 FREQUÊNCIAS ALÉLICAS

Foram observados todos os alelos já identificados para sete (APO, SB19.3, PV92, AT3, LPL, RB2300 e CKM) dos oito *loci* utilizados, exceto o *loci* DRD2, no qual o alelo *DRD2*1* não foi observado entre as subamostras. Portanto, a frequência do mesmo foi de 0% para a amostra populacional (Tabela 3).

A comparação das distribuições das frequências alélicas entre as quatro subamostras (Tabela 3 e Figuras 9 e 10), mostrou que os alelos *Sb19.3*1* e *APO*1* foram mais frequente nas amostras controles tanto em Ilhéus quanto em Itabuna comparadas com os casos. Enquanto isso, os alelos *AT3*1*, *LPL*1*, *RB2300*1*, *CKM*1* tiveram suas frequências maiores entre os casos. Já o alelo *PV92*1*, comportou-se diferente entre as amostras, em Ilhéus os casos obtiveram uma frequência maior, em Itabuna foram os controles que apresentaram a maior frequência (Tabela 3 e Figuras 9 e 10).

Os valores de δ entre as amostras analisadas e os grupos mundiais estão apresentados na Tabela 4. Os valores de sete dos oito *loci* analisados foram eficazes para a separação entre os grupos, apenas o *locus* SB19.3 não foi eficiente.

Tabela 3. Frequências genótípicas e alélicas dos oito *loci* nas quatro amostras analisadas-Ilhéus (ILH) e Itabuna (ITB).

	Casos ILH	Controles ILH	Casos ITB	Controles ITB	Total Geral
APO					
1/1	27	34	44	55	160
1/2	7	3	13	9	32
2/2	0	0	4	5	9
N	34	37	61	69	201
APO*1	0,897	0,959	0,828	0,862	0,887
PV92					
1/1	1	0	3	8	12
1/2	15	13	18	27	73

2/2	22	23	42	35	122
N	38	36	63	70	207
PV92*1	0,224	0,181	0,190	0,307	0,225
AT3					
1/1	9	6	19	8	42
1/2	16	18	24	32	90
2/2	10	11	20	29	70
N	35	35	63	69	202
AT3*1	0,486	0,429	0,492	0,348	0,439
LPL					
1/1	14	14	20	24	72
1/2	17	15	36	31	99
2/2	2	7	7	14	30
N	33	36	63	69	201
LPL*1	0,682	0,597	0,603	0,572	0,614
DRD2					
1/1					
1/2					
2/2	38	37	56	70	201
N	38	37	56	70	201
DRD2*1	0	0	0	0	0
RB2300					
1/1	10	11	26	21	68
1/2	28	23	35	48	134
2/2	0	3	1	1	5
N	38	37	62	70	207
RB2300*1	0,632	0,608	0,702	0,643	0,646
CKM					
1/1	5	2	6	10	23
1/2	19	14	34	25	92
2/2	14	21	24	33	92
N	38	37	64	68	207
CKM*1	0,382	0,243	0,359	0,331	0,329
SB19.3					
1/1	16	20	22	37	95
1/2	19	15	32	17	83
2/2	3	3	11	15	32
N	38	38	65	69	210
SB19.3*1	0,671	0,724	0,585	0,659	0,660

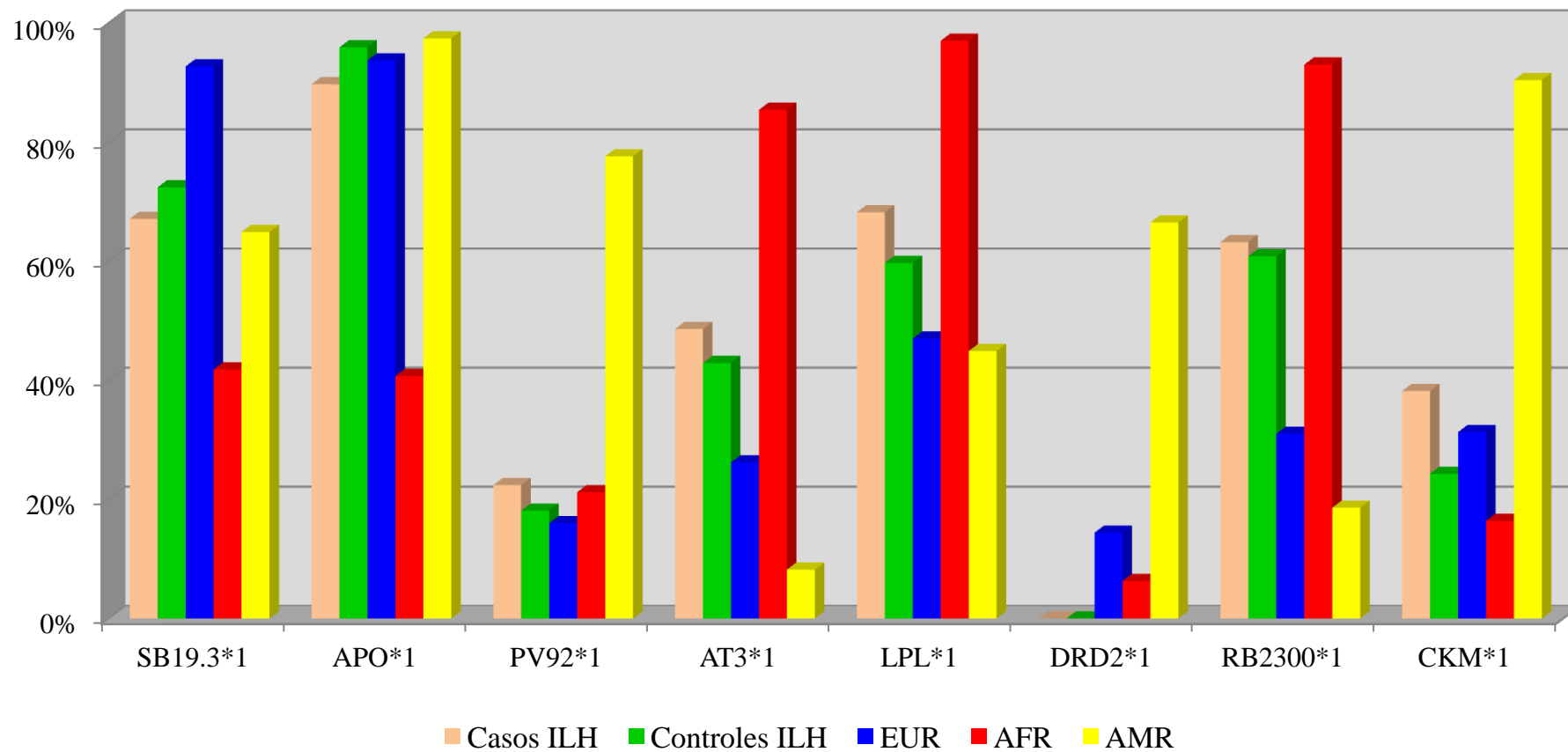


Figura 9. Comparação das frequências dos alelos *APO*1*, *PV92*1*, *DRD2*1*, *RB2300*1*, *LPL*1*, *AT3*1* e *Sb19.3*1* nos casos e controles de Ilhéus.

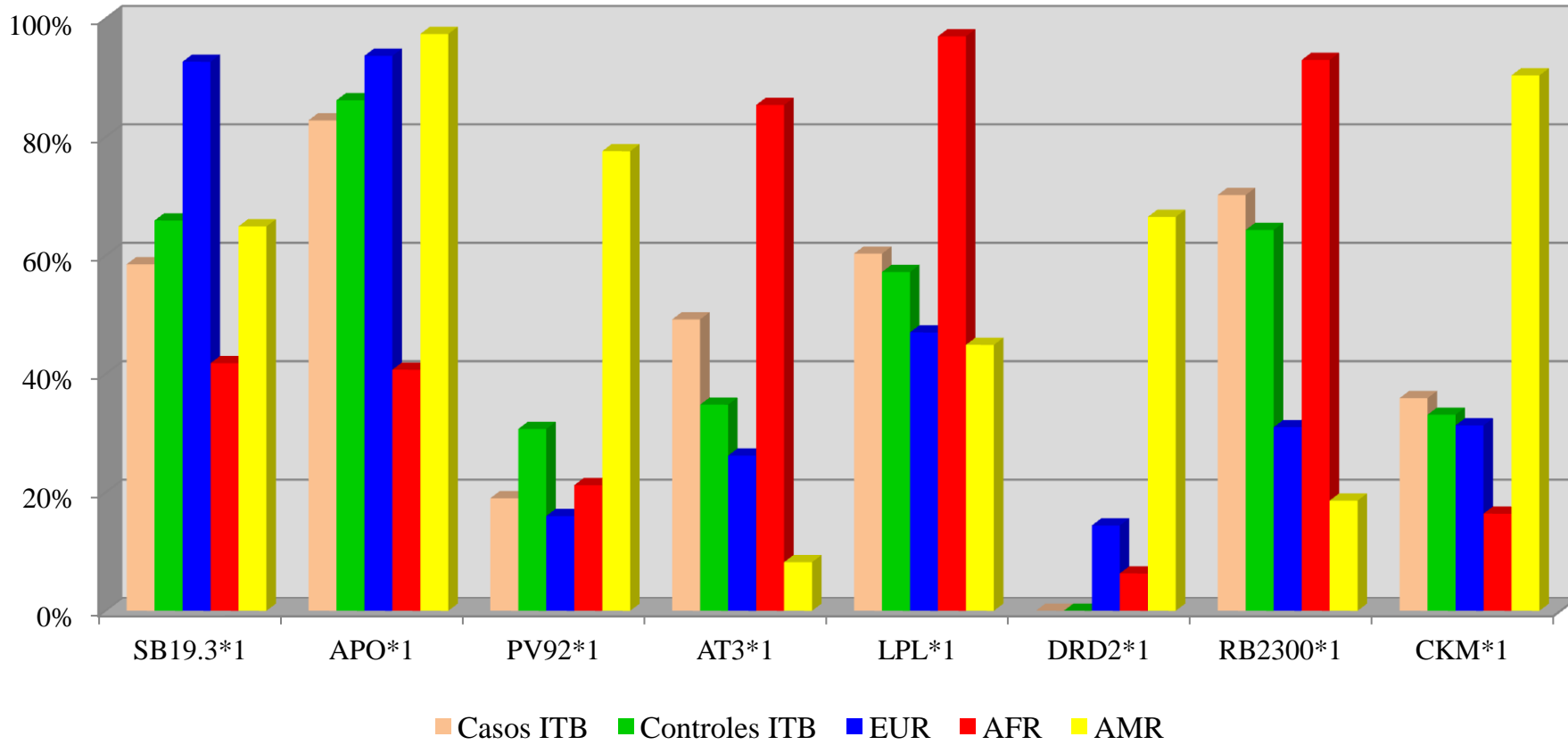


Figura 10. Comparação das frequências dos alelos *APO*1* *PV92*1* *DRD2*1*, *RB2300*1*, *LPL*1*, *AT3*1* e *Sb19.3*1* nos casos e controles de Itabuna.

Tabela 4. Diferencial de frequência (δ) dos loci MIAs entre grupos étnicos mundiais – Africano (AFR)¹, Europeu (EUR)¹ e Ameríndio (AMR)^{1,2} e os casos e controles em Itabuna (ITB) e Ilhéus (ILH). Os valores >30% estão destacados em negrito.

Alelos	ILH						ITB					
	Casos			Controles			Casos			Controles		
	AFR	EUR	AMR	AFR	EUR	AMR	AFR	EUR	AMR	AFR	EUR	AMR
<i>SB19.3*1</i>	0,253	0,256	0,022	0,306	0,203	0,075	0,167	0,342	0,064	0,241	0,268	0,010
<i>APO*1</i>	0,490	0,040	0,077	0,552	0,022	0,015	0,421	0,109	0,146	0,455	0,075	0,112
<i>PV92*1</i>	0,012	0,064	0,552	0,031	0,021	0,595	0,022	0,030	0,586	0,095	0,147	0,469
<i>AT3*1</i>	0,368	0,224	0,404	0,425	0,167	0,347	0,362	0,230	0,410	0,506	0,086	0,266
<i>LPL*1</i>	0,288	0,212	0,233	0,373	0,127	0,148	0,367	0,133	0,154	0,398	0,102	0,123
<i>DRD2*1</i>	0,063	0,144	0,665	0,063	0,144	0,665	0,063	0,144	0,665	0,063	0,144	0,665
<i>RB2300*1</i>	0,298	0,322	0,446	0,322	0,298	0,422	0,228	0,392	0,516	0,287	0,333	0,457
<i>CKM*1</i>	0,218	0,069	0,522	0,079	0,070	0,661	0,195	0,046	0,545	0,167	0,018	0,573

¹ SHRIVER *et al.*, 2003

² LUIZON *et al.*, 2008

4.2 ADERÊNCIA AO EQUILÍBRIO DE HARDY-WEINBERG

As frequências genóticas observadas dos oito *loci* (Tabela 3) foram comparadas aos valores esperados segundo o equilíbrio de Hardy-Weinberg pelo teste exato G. Observou-se seis desvios, um na subamostra de casos ITB para o *locus* AT3, dois na subamostra de controles ITB para os *loci* Sb19.3 e APO, um no teste multi-*locus* da subamostra controles ITB e dois nos testes multipopulacionais para os *loci* Sb19.3 e APO (Tabela 5 e 6).

Testes mais sensíveis podem ser utilizados para detectar se os desvios observados ocorreram pela ausência ou pelo excesso de heterozigotos. Nas subamostras Casos ITB para o *locus* AT3 ($p=0,041$) e Controles ITB para os Sb19.3 ($p=0,0002$) e APO ($p=0,002$) os desvios podem ser explicados pela falta de heterozigotos (Tabela 6). Enquanto que o *locus* RB2300 mostrou diferença significativa pelo excesso de heterozigotos (Tabela 5). Não foi possível a realização do teste exato para o *locus* DRD2, pois este só apresenta o genótipo homozigoto (Tabelas 5 e 6).

Tabela 5. Probabilidades de desvio casual segundo teste exato (Guo e Thompson, 1992) para a verificação do equilíbrio de Hardy-Weinberg (excesso de heterozigotos). Os valores significativos ($p < 0,05$) estão destacados em negrito.

Amostra	Sb19.3	APO	PV92	AT3	LPL	DRD2	RB2300	CKM	Multi- <i>locus</i>
Casos ILH	0,365	0,714	0,404	0,823	0,280	-	0,0001	0,525	0,006
Controles ILH	0,717	0,959	0,266	0,552	0,894	-	0,075	0,660	0,393
Casos ITB	0,578	0,990	0,856	0,988	0,114	-	0,0050	0,185	0,990
Controles ITB	1,00	1,00	0,875	0,555	0,842	-	0,0000	0,958	0,173
Multi- populacional	0,969	0,996	0,510	0,865	0,318	-	0,0000	0,520	0,447

Tabela 6. Probabilidades de desvio casual segundo teste exato (Guo e Thompson. 1992) para a verificação do equilíbrio de Hardy-Weinberg (déficit de heterozigotos). Os valores significativos ($p < 0,05$) estão destacados em negrito.

Amostra	Sb19.3	APO	PV92	AT3	LPL	DRD2	RB2300	CKM	Multi-locus
Casos ILH	0,868	1,000	0,901	0,401	0,924	-	1,000	0,735	0,994
Controles ILH	0,586	1,000	1,000	0,711	0,292	-	0,985	0,692	0,827
Casos ITB	0,618	0,061	0,385	0,041	0,961	-	1,000	0,928	0,607
Controles ITB	0,0002	0,002	0,282	0,649	0,309	-	1,000	0,117	0,010
Multi-populacional	0,030	0,004	0,490	0,135	0,682	-	1,000	0,480	0,553

4.3 ASSOCIAÇÕES PAR A PAR ENTRE *LOCI*

Para cada uma das subamostras foi realizada comparações par a par entre os oito *loci* MIAs (Tabela 7). De 28 análises realizadas foram observadas associações par a par entre os *loci* LPL/RB2300 (Casos ILH- $p=0,017$), Sb19.3/CKM (Casos ITB- $p=0,015$, Controles ILH- $p=0,04$), AT3/RB2300 (Controles ITB- $p=0,0013$), PV92/AT3 (Controles ILH- $p=0,04$), RB2300/CKM (Controles ILH- $p=0,007$). Quando agrupadas todas as subamostras identificou-se valores significativos para os pares de *loci* AT3/RB2300 ($p=0,014$) e Sb19.3/CKM ($p=0,03$).

4.4 DIFERENCIAÇÃO GENÉTICA DAS POPULAÇÕES

Na análise de diferenciação gênica baseada nos oito *loci* MIAs entre as quatro subamostras comparou-se cada *locus* em pares de populações (Tabela 8). Dessa forma, encontrou-se valor significativo ($p < 0,05$) entre as subamostras: Controles ITB/Casos ITB (*loci* PV92 e AT3), Controles ILH/Casos ITB (*locus* APO) e Controles ILH/Controles ITB (*locus* APO). Considerando todos os *loci* somente um par de subamostras (Controles ILH/Casos ITB) apresentou diferença significativa. Na análise multipopulacional apenas o *locus* APO diferenciou as subamostras (Tabela 8).

Tabela 7. Associações par a par entre *loci* em cada uma das quatro subamostras. Os valores estatisticamente significativos ($p < 0,05$) estão destacados em negrito.

Par de <i>loci</i>	Casos ILH	Casos ITB	Controles ITB	Controles ILH	Todas
SB193/APO	1,000	0,298	0,544	0,164	0,510
SB193/PV92	0,665	0,983	0,790	0,254	0,852
APO/PV92	0,063	0,578	0,557	0,537	0,339
SB193/AT3	0,820	0,895	0,083	0,485	0,532
APO/AT3	0,845	0,576	0,124	0,563	0,563
PV92/AT3	0,522	0,472	0,257	0,043	0,160
SB193/LPL	0,334	0,214	0,885	0,295	0,437
APOLPL	0,355	0,334	0,662	1,000	0,748
PV92/LPL	0,694	1,000	0,293	0,565	0,827
AT3/LPL	0,958	0,070	0,470	0,262	0,295
SB193/DRD2	STC	STC	STC	STC	STC
APO/DRD2	STC	STC	STC	STC	STC
PV92/DRD2	STC	STC	STC	STC	STC
AT3/DRD2	STC	STC	STC	STC	STC
LPL/DRD2	STC	STC	STC	STC	STC
SB193/RB2300	0,866	0,819	0,366	0,571	0,873
APO/RB2300	0,646	1,000	0,154	1,000	0,797
PV92/RB2300	0,449	0,698	0,567	1,000	0,903
AT3/RB2300	0,349	0,263	0,001	0,572	0,014
LPL/RB2300	0,017	0,275	0,645	0,197	0,062
DRD2/RB2300	STC	STC	STC	STC	STC
SB193/CKM	0,556	0,016	0,566	0,042	0,030
APO/CKM	0,718	0,092	0,569	1,000	0,585
PV92/CKM	0,493	0,804	0,522	1,000	0,925
AT3/CKM	0,339	0,417	0,679	0,476	0,628
LPL/CKM	0,749	0,058	0,688	0,457	0,379
DRD2/CKM	STC	STC	STC	STC	STC
RB2300/CKM	0,877	0,182	0,693	0,007	0,074

STC- Sem tabela de contingência.

Tabela 8. Diferenciação gênica baseada nos oito *loci* MIAs entre as quatro subamostras: Casos ILH (1), Casos ITB (2), Controles ITB (3) e Controles ILH (4). Os valores significativos ($p < 0,05$) estão destacados em negrito.

Amostras	SB19.3	APO	PV92	AT3	LPL	DRD2	RB2300	CKM	Total
2 X 1	0,239	0,204	0,593	1,000	0,338	NI	0,357	0,766	0,619
3 X 1	0,883	0,515	0,212	0,068	0,169	NI	0,884	0,539	0,372
3 X 2	0,218	0,492	0,034	0,024	0,621	NI	0,353	0,703	0,071
4 X 1	0,597	30,198	W	0,609	0,383	NI	0,865	0,080	0,469
4 X 2	0,054	0,006	1,000	0,455	1,000	NI	0,216	0,123	0,035
4 X 3	0,368	0,034	0,051	0,295	0,773	NI	0,666	0,213	0,088
Multi-pop	0,234	0,031	0,091	0,092	0,504	NI	0,534	0,263	0,038

NI – não informativo;

Multi-pop – Multipopulacional.

4.5 MISTURA ÉTNICA

Os valores percentuais das estimativas de mistura étnica foram diferentes em cada um das amostras populacionais, além de indicar uma maior contribuição europeia em todas as amostras (Casos ILH- 58%, Controles ILH- 68%, Casos ITB- 49% e Controles ITB- 56%) (Tabela 9 e Figura 11).

Com base nos oito *loci* MIAs as subamostras Casos ILH, Casos ITB e Controles ITB apresentaram valores consistentes com o modelo tri-híbrido. Sendo que a amostra Controles ILH adaptou-se ao modelo di-híbrido (Tabela 9). Todas as subamostras apresentaram a contribuição ameríndia inferior à contribuição europeia e africana (Figura 11).

Tabela 9. Estimativa de mistura étnica nas amostras analisadas, segundo o método de identidade gênica. Valores em percentuais.

	Casos (%)			Controles (%)		
	ILH	ITB	Total	ILH	ITB	Total
Europeu	58±2,7	49±0,7	52±4	68±1,7	56±0,9	63±3
Africano	38±1,4	45±0,4	43±2	32±1,7	34,10±0,5	32±1
Ameríndio	4±1,8	6±0,5	5±3	-	10±0,6	5±2
R²	99,88	99,99	99,82	97,96	99,98	99,87

R²- indica a adaptação dos dados ao modelo de mistura.

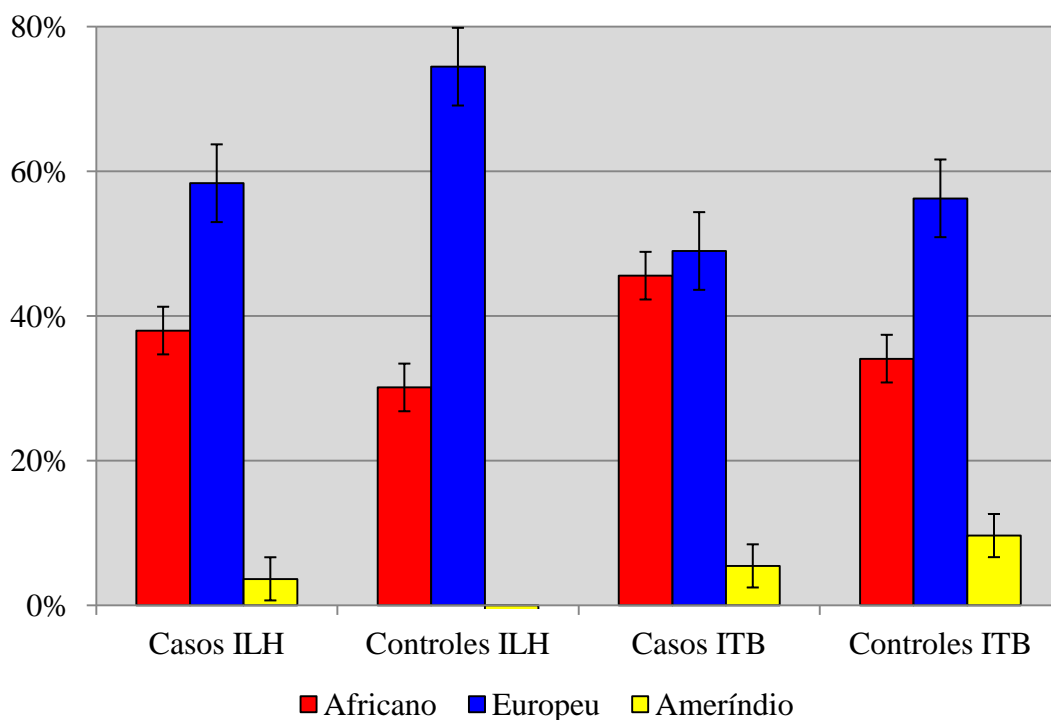


Figura 11. Comparação gráfica das estimativas de mistura étnica obtida a partir de oito *loci* (Sb19.3, APO, PV92, AT3, LPL, DRD2, RB2300 e CKM).

A Tabela 10 apresenta os dados agrupados das quatro subamostras de Ilhéus e Itabuna em casos e controles. Esta apresenta as estimativas de mistura estimadas utilizando os MIAs e obtidas com autodeclaração. A autodeclaração foi classificada apenas em branco e não branco, desta forma os brancos foram considerados europeus e não brancos como africanos e ameríndios.

Tabela 10. Estimativas das contribuições étnicas da amostra de casos e controles obtidas com os MIAs e as avaliadas com a autodeclaração obtidas por Santos (2010).

Contribuições étnicas	Casos (%)		Controles (%)	
	MIAs	Autodeclarada	MIAs	Autodeclarada
EUR / branco	52	26	63	33,3
AFR+AMR / não branco	48	74	37	66,7

5 DISCUSSÃO

O senso realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) é uma ferramenta importante para conhecer o perfil étnico-racial da população através da autodeclaração. A pesquisa é feita com base na cor da pele, nas categorias branca, preta, amarela, parda e indígena. De acordo com o censo 2010, a população baiana se declarou 22,2% branca, 17,1% preta, 1,1% amarela, 59,2% parda e 0,4% indígena. No município de Itabuna esses valores foram de, 22%, 14%, 1%, 62,6% e 0,4%, enquanto que em Ilhéus foi de 19%, 19%, 1%, 59% e 2% para as contribuições branca, preta, amarela e indígena, respectivamente. Assim, de acordo com essas informações todas possuem o mesmo padrão de contribuição, sendo a parda a maior entre os grupos. Entretanto, este tipo de abordagem não reflete a real contribuição genética, visto que no Brasil, país altamente miscigenado, a cor da pele tem uma fraca correlação com o grau de ancestralidade africana (PARRA *et al.*, 2003; PENA *et al.*, 2009; BEUTEN *et al.*, 2011).

Os brasileiros formam uma das populações mais heterogêneas do mundo, por isso é interessante estudar como os três grupos (africano, europeu e ameríndio) contribuíram para sua formação. Com a análise dos MIAs foi possível estimar a composição étnica de amostras do sul do estado da Bahia (Ilhéus e Itabuna) onde o povoamento e introdução dos africanos foi posterior a de Salvador capital do estado da Bahia. (TAVERES, 2001; ANDRADE; ROCHA, 2005).

5.1 FREQUÊNCIAS ALÉLICAS

São considerados MIAs os marcadores que apresentam diferenciais de frequência (δ) $>30\%$ entre quaisquer grupos étnicos ou geográficos (BONILLA *et al.*, 2004). Os valores de δ obtidos para alelos *APO*1*, *RB2300*1*, *LPL*1* e *AT3*1* foram $> 50\%$ entre Africanos e Europeus e entre Africanos Ameríndios, dessa forma podem ser utilizados para estimar a contribuição gênica Africana. Os *loci* PV92 e CKM

apresentaram $\delta > 50\%$ entre Europeus e Ameríndios ou entre Africano e Ameríndio, assim com poder para discriminar populações ameríndias da africana e europeia. O *locus* DRD2 obteve o $\delta \geq 40\%$ entre Africano e Ameríndio e Europeu e Ameríndio, deste modo pode ser usado para estimar a contribuição ameríndia. O *locus* Sb19.3 teve o $\delta > 50\%$ apenas entre Africanos e Europeus, sendo capaz apenas de diferenciar entre as duas populações. (Figura 12 e a Tabela 11)

Entre as amostras do presente estudo a frequência do alelo *LPL *1* variou de 57% (Controles ITB) a 68% (Casos ILH) com média de 61% (Tabela 3 e Figura 9), mais próxima à frequência relatada para africanos, 97% (Tabela 3). Em Salvador este alelo apresentou a frequência de 74% (ABÉ-SANDES *et al.*, 2010).

A frequência do alelo *AT3-I/D* 1* variou de 35% (Controles ILH) a 49% (Casos ILH) (Tabela 5 e Figura 8), com média de 44%. Este marcador, em populações mundiais é mais frequente (85%) em africanos (Figura 3). Em estudo realizado em uma amostra populacional de Salvador foi encontrada a frequência de 55% (ABÉ-SANDES *et al.*, 2010).

Para o alelo *Sb19.3*1* (Tabela 3) a frequência variou de 58% (Casos ILH) a 72% (Controles ITB) entre as subamostras, com média de 65,5%. Este valor médio posiciona-se entre a frequência de ameríndios (65%) e europeus (92%), nas populações mundiais (Tabela 11). Na população de Salvador foi encontrada uma frequência de 67% (ABÉ-SANDES *et al.*, 2010).

Esta variação das frequências alélicas das subamostras em relação às populações mundiais pode ser explicada pela miscigenação ou pelo tamanho amostral, visto que há uma grande dificuldade na coleta dos indivíduos com câncer de próstata e dos pareados para esta doença. Além do mais, como a etnia pode estar associada à manifestação da patologia, a seleção e coleta das amostras não são processos aleatórios.

Tabela 11. Frequências e diferencial de frequência (δ) dos *loci* MIAs entre grupos étnicos mundiais. Os valores de $\delta > 50\%$ estão destacados em negrito.

Alelos/Pop	Frequência do alelo *1			Diferencial de frequência (δ)		
	AFR	EUR	AMR	AFR / EUR	AFR / AMR	EUR / AMR
<i>Sb19.3*1</i>	0,418	0,927	0,649	0,509	0,295	0,214
<i>APO*1</i>	0,407	0,937	0,974	0,530	0,590	0,060
<i>PV92 *1</i>	0,212	0,160	0,776	0,052	0,712	0,764
<i>AT3*1</i>	0,854	0,262	0,082	0,592	0,808	0,216
<i>LPL*1</i>	0,970	0,470	0,449	0,500	0,503	0,003
<i>DRD2*1</i>	0,063	0,144	0,665	0,080	0,409	0,446
<i>RB2300*1</i>	0,930	0,310	0,186	0,620	0,824	0,204
<i>CKM*1</i>	0,164	0,313	0,904	0,150	0,740	0,590

AFR: Média Africana Ponderada (República Centro Africana e Nigéria (Benin e Ibadan), Parra *et al.*, 1998; e Serra Leoa (Mende e Temne), Parra *et al.*, 2001);

EUR: Média Europeia Ponderada (Inglaterra, Irlanda e Alemanha, Parra *et al.*, 1998; Portugueses, Tomás *et al.*, 2002; e Espanhóis, dbSNPINCBI);

AMR: Indígenas da Amazônia para os *loci* APO, SB19.3, AT3, PV92, LPL e RB2300 (LUIZON *et al.* 2008); Americanos Nativos para os *loci* CKM e DRD2 (SHRIVER, 2003)

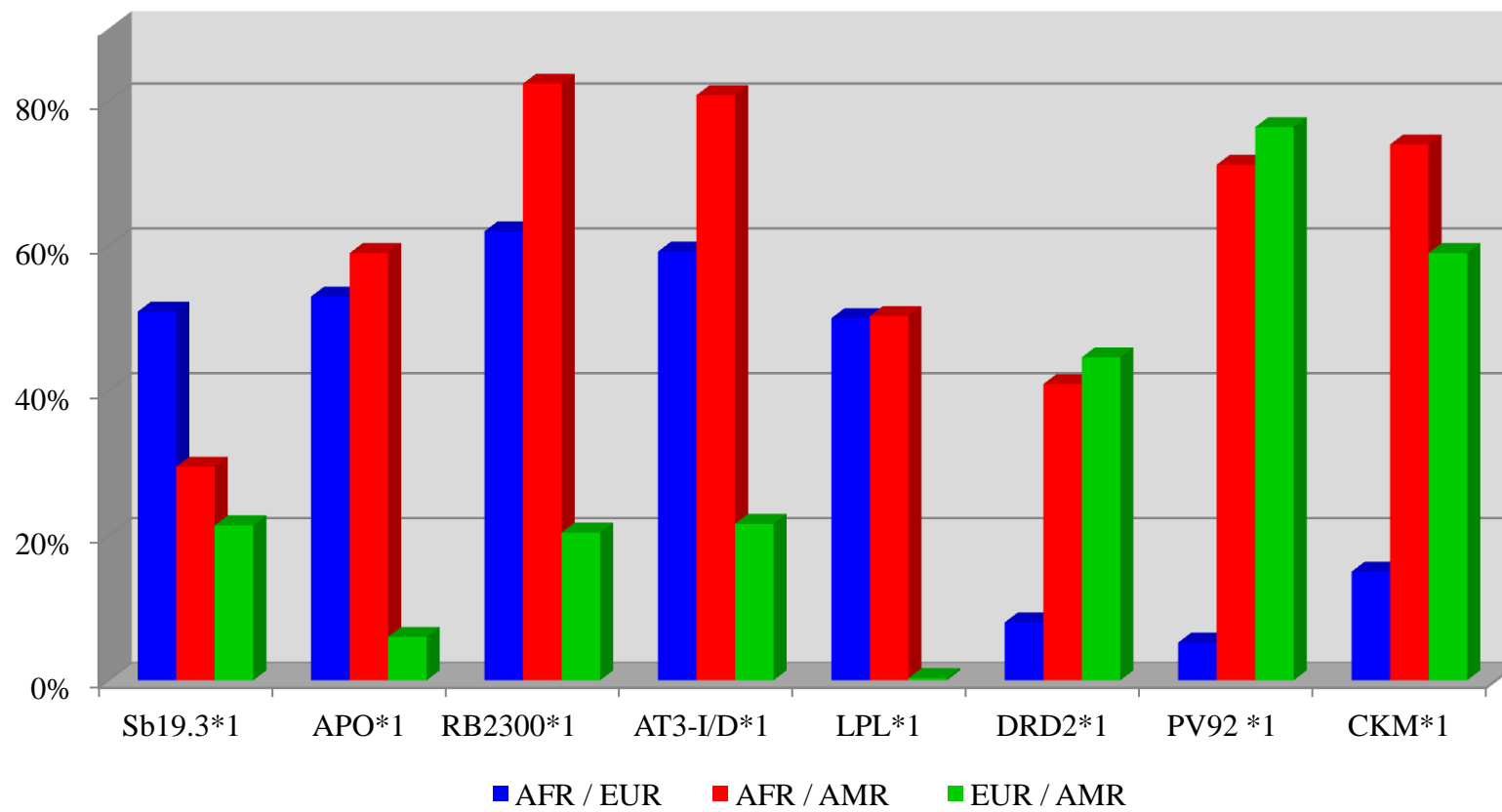


Figura 12. Diferencial de frequência (δ) dos alelos Sb19.3*1, APO*1, PV92*1, AT3*1, LPL*1, DRD2*1, RB2300*1, CKM*1 entre os grupos étnicos mundiais.

5.2 EQUILÍBRIO DE HARDY-WEINBERG

Foram observados desvios do Equilíbrio de Hardy-Weinberg (EHW) nos *loci* Sb19.3 e APO nos Controles ITB e no *locus* AT3 nos Casos ITB. (Tabela 6). Quando aplicado o teste multi-locus apenas a subamostra Controles ITB, indicou a falta de heterozigotos. Na análise calculada por *locus*, teste multipopulacional, sugeriu que os *loci* Sb19.3 e APO também apresentaram a falta de heterozigotos (Tabela 6).

Enquanto que os desvios do EHW encontrados no *locus* RB2300 nos Casos ILH, Casos ITB e Controles ITB podem ser explicados pelo excesso de heterozigotos. O teste multi-locus aplicado indicou o excesso de heterozigotos nos Casos ILH, já a análise multipopulacional sugeriu excesso para o *locus* RB2300 (Tabela 5).

Desvios no EHW em alguns ou vários *loci* serão detectados se a estruturação da população for real (FALUSH; STEPHENS; PRITCHARD, 2003). Como a subdivisão populacional pode surgir por causa da miscigenação, estes resultados sugerem estruturação nas amostras analisadas.

5.3 ASSOCIAÇÕES PAR A PAR ENTRE *LOCI*

O padrão de associação par a par entre os oito *loci* foi analisado com o intuito de averiguar a existência de estrutura genética nas subamostras, o que seria sugerido pela presença de desvios significativos entre *loci* não ligados.

As associações observadas entre os *loci* LPL/RB2300 (Casos ILH- $p=0,017$), Sb19.3/CKM (Casos ITB- $p=0,015$, Controles ILH- $p=0,04$), AT3/RB2300 (Controles ITB- $p=0,0013$), PV92/AT3 (Controles ILH- $p=0,04$), RB2300/CKM (Controles ILH- $p=0,007$) seriam indicativas de subestruturação por miscigenação (Tabela 7).

5.4 DIFERENCIAÇÃO GENÉTICA DAS POPULAÇÕES

A comparação dos oito *loci* em pares das quatro subamostras indicou diferenças das frequências alélicas (Tabela 8). Os valores significativos ($p < 0,05$) mostraram que a distribuição alélica foi diferente entre as subamostras e na análise multipopulacional (*locus* APO), estes dados são condizentes com a história de formação das subpopulações e também evidenciam diferenças entre as amostras casos/controle para o câncer de próstata, sugerindo a necessidade do controle genômico.

5.5 MISTURA ÉTNICA

As estimativas de mistura étnica foram obtidas utilizando o método da identidade gênica (CHAKRABORTY 1985), o coeficiente de identidade gênica estima a probabilidade de dois genes escolhidos ao acaso (em uma ou mais populações) serem idêntico por natureza. Nas subamostras Casos ILH, casos ITB e Controles ITB, utilizando os oitos *loci* MIAs a estimativa de mistura étnica se adequou ao modelo tri-híbrido. Para a subamostra Controles ILH o modelo sugerido foi o di-híbrido, sendo inconsistente para o componente ameríndio (Tabela 9 e Figura12).

Em todas as amostras a maior contribuição observada foi europeia, nos Casos ITB (49%), Controles ITB (56%), Casos ILH (58%) e Controles ILH (68%). A contribuição ameríndia não foi estimada nos Controles ILH, por conta da ausência de contribuição ameríndia (modelo di-híbrido) e a maior contribuição africana foi observada nos Casos ITB (45%). Observa-se também que a amostra controle apresentou uma frequência europeia maior quando comparada aos casos (Tabela 9). Nas análises realizadas os valores de R^2 foram altos ($> 97\%$) e erros-padrão baixos, o que indicam uma boa adaptação dos dados ao modelo de mistura.

Estes resultados estão de acordo com a história da formação da população de Ilhéus e Itabuna, inicialmente de constituição ameríndia, seguida pela europeia e africana. (BARBOSA, 1994; ANDRADE, 2003; ANDRADE; ROCHA, 2005). A pequena contribuição ameríndia deve-se principalmente, pela baixa união deste grupo

com outros, visto que desde o início da colonização e povoamento da região, a população indígena sofreu uma grande redução, como também às uniões preferenciais entre homens africanos e europeus e mulheres ameríndias (CARVALHO-SILVA *et al.*, 2001; ABE-SANDES; SILVA; ZAGO, 2004; MARRERO *et al.*, 2005; RESQUE *et al.*, 2010).

No presente trabalho o componente europeu apresenta maior contribuição, esta estimativa corrobora com a encontrada em outros trabalhos realizados em amostras da Bahia e do Nordeste (PENA *et al.*, 2009; LINS *et al.*, 2010). Analisando um conjunto de 40 *indels* em amostras de várias regiões do Brasil, dentre elas 147 indivíduos doadores de sangue proveniente de Ilhéus-BA, Pena *e cols.* (PENA *et al.*, 2011) observaram que apesar das diferenças regionais, a ancestralidade dos brasileiros é relativamente uniforme e os indivíduos que se autodeclararam negros ou pardos possuem uma alta ancestralidade europeia.

Análise realizada em amostra populacional de Salvador apresentou contribuições africana, europeia e ameríndia estimadas em $49.0 \pm 3.0\%$, $44.0 \pm 9.0\%$, $7.0 \pm 9.0\%$, respectivamente (FELIX *et al.*, 2010). Assim, os resultados encontrados para presente amostra indicam que a região sul apresenta suas contribuições étnicas moderadamente diferentes quando comparadas as encontradas na capital. Estes dados estão de acordo com o sugerido por Azevedo *e cols.* (1982) que ocorre um processo de “branqueamento” da costa para o interior do estado ou da capital para o litoral.

As estimativas de mistura obtidas com marcadores genéticos autossômicos em outra amostra da população de Salvador para indivíduos infectados pelo HIV-1 indicaram 47% de contribuição africana, 37% europeia e 16% ameríndia, sendo semelhante às observadas na população de não infectados. Deste modo, nessas amostras a ancestralidade genômica africana também foi maior, seguida da europeia e ameríndia (ABÉ-SANDES *et al.*, 2010).

Em análises que utilizam marcadores uniparentais, do cromossomo Y e do DNA mitocondrial (DNAMt) os resultados indicaram que a patrilinhagem brasileira é principalmente de origem europeia, enquanto que a matrinhagem é ameríndia ou

africana (ALVES-SILVA *et al.*, 2000; CARVALHO-SILVA *et al.*, 2001; ABE-SANDES; SILVA; ZAGO, 2004; MARRERO *et al.*, 2005; RESQUE *et al.*, 2010).

Quando observadas as estimativas de mistura étnica obtidas utilizando os marcadores moleculares e as avaliadas pela cor da pele autodeclarada segundo critérios do IBGE, uma grande diferença foi notada (Tabela 10). As estimativas obtidas utilizando os MIAs indicou uma maior contribuição europeia, sendo que nas estimativas encontradas com base na autodeclaração, a maior contribuição foi de não brancos, já que a autodeclaração não separa africanos, de pardos e índios. Porém, não foram encontradas diferenças expressivas entre os caso e controles para a amostra analisada.

Na análise realizada com MIAs os casos apresentaram uma frequência de 52% para a contribuição europeia e 43% para a africana. Na autoclassificação 26% declaram-se brancos e 74% não brancos (negros e pardos). Paralelamente, na amostra controle os MIAs indicaram 63% para a ancestralidade europeia e 32% africana. Já na autoclassificação 33,3% declaram-se brancos e 66,7% não brancos (negros e pardos). Estes valores divergentes poderiam ser explicados pela não confiabilidade da classificação com base na cor da pele, sendo que entre os pardos podem ser encontrados indivíduos de composição africana ou europeia, o que leva a subestimativa principalmente da ancestralidade europeia. Visto que, a correlação entre a cor da pele e ancestralidade genômica é falha em populações miscigenadas (PENA *et al.*, 2009).

Os dados do presente trabalho mostram que com a aplicação dos MIAs os indivíduos com câncer de próstata apresentaram contribuição europeia maior, contrapondo com as estimativas obtidas pela classificação fenotípica autorreferida, na qual foi quantificada uma maior contribuição africana. Estes resultados indicam a importância de avaliar o grau de mistura genética baseados em marcadores genéticos, visto que estes fornecem resultados mais precisos sobre a composição genética da população.

No trabalho de MASON *et al.* (2010), afro-americanos com genótipos CC do CD14 (variante $_260$, C>T) agrupados por raça auto-declarada, apresentaram maior risco de câncer de próstata. Esses resultados demonstram, que a variação na região promotora do CD14 (gene que regula o processo inflamatório) pode contribuir para a suscetibilidade ao câncer de próstata em homens afro-americanos.

Outro estudo realizado com amostra de norte americanos mostrou que a variação alélica do SNP -160 C/A do gene da E-caderina conferi um fator de risco para o câncer de próstata maior em euro-americanos do que em afro-americanos (POOKOT *et al.*, 2006). Entretanto, uma análise realizada com o mesmo SNP na amostra do presente trabalho não entrou associação com a etnia, provavelmente porque esta usou inicialmente o critério da autodeclaração (SANTOS, 2010). Dessa forma, a classificação através da cor da pele pode ter inserido um viés impedindo assim uma associação entre a etnia e o câncer de próstata.

Levando isso em conta, a estratificação deve ser avaliada em estudos de associação caso/controles em populações miscigenadas. Assim, a aplicação da análise com MIAs e a realização do controle genômico permite reduzir as associações espúrias devido à estratificação populacional. Esse tipo de estudo viabiliza a retirada dos indivíduos que não possuam a proporção de mistura necessária para determinada análise. Dessa forma, tornar a amostra mais homogênea possível, estabelecendo “a mesma” proporção de amostras miscigenadas entre os casos e controles.

Para a realização do controle genômico pode ser aplicado o programa STRUCTURE que investiga a estrutura da população. Seus usos incluem inferir a presença de populações distintas, atribuindo indivíduos às populações, estudando zonas híbridas, identificando os migrantes e indivíduos miscigenados, e estimar as frequências alélicas da população em situações em que muitos indivíduos são migrantes ou miscigenados (PRITCHARD; STEPHENS; DONNELLY, 2000). Assim, é possível comparar a probabilidade individual inferida entre casos e controles e fazer as correções utilizando o método da regressão logística multivariada.

6 CONCLUSÕES

As frequências alélicas para os *loci* MIAs (APO, Sb19.3, PV92, AT3, LPL, DRD2, RB2300 e CKM) apresentaram diferenças entre subamostras analisadas em relação às populações mundiais.

Os desvios no Equilíbrio de Hardy-Weinberg detectados em alguns *loci* sugerem estruturação nas amostras analisadas especialmente por causa da miscigenação.

As estimativas de mistura étnica se adequaram ao modelo tri-híbrido, exceto para a subamostra Controles ILH. Em todas as amostras, a maior contribuição observada foi a europeia. Os resultados encontrados na presente amostra foram diferentes dos observados em amostra populacional de Salvador.

A diferença encontrada entre as estimativas de mistura étnica obtidas utilizando os marcadores moleculares e as analisadas pela cor da pele autodeclarada indicam a importância de avaliar o grau de mistura genética, utilizando MIAs.

A diferenciação gênica populacional observada no locus APO, desvios no EHW, associações par a par entre *loci* não ligados, contribuições variáveis nas estimativas de mistura étnica entre as subamostras e por fim discordâncias entre ancestralidade genômica e cor da pele autodeclarada indicam subestruturação populacional e a necessidade de realização do controle genômico.

REFERÊNCIAS BIBLIOGRÁFICAS

ABÉ-SANDES, K. et al. Ancestralidade Genômica, Nível Socioeconômico e Vulnerabilidade ao HIV/AIDS na Bahia, Brasil. **Saúde Soc**, v. 19, n. 2, p. 75-84, 2010.

ABE-SANDES, K.; SILVA, W. A., JR.; ZAGO, M. A. Heterogeneity of the Y chromosome in Afro-Brazilian populations. **Hum Biol**, v. 76, n. 1, p. 77-86, Feb 2004. ISSN 0018-7143 (Print)
0018-7143 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15222681 >.

ABU-AMERO, K. K. et al. Susceptibility to primary angle closure glaucoma in Saudi Arabia: the possible role of mitochondrial DNA ancestry informative haplogroups. **Mol Vis**, v. 17, p. 2171-2176, 2011.

ALVES-SILVA, J. et al. The ancestry of Brazilian mtDNA lineages. **Am J Hum Genet**, v. 67, n. 2, p. 444-61, Aug 2000. ISSN 0002-9297 (Print)
0002-9297 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10873790 >.

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

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

AZEVÊDO, E. et al. Spread and diversity of human populations in Bahia, Brazil. **Hum Biol**, v. 54, n. 2, p. 329-41 May, 1982.

BARBOSA, C. R. A. **Notícia histórico de Ilhéus**. Itabuna: COLOGRAF, 1994.

BARNHOLTZ-SLOAN, J. S. et al. Examining Population Stratification via Individual Ancestry Estimates versus Self-Reported Race. **Cancer Epidemiol Biomarkers Prev**, v. 14, n. 6, 2005.

BATTILANA, J. et al. Alu insertion polymorphisms in Native Americans and related Asian populations. **Ann Hum Biol**, v. 33, n. 2, p. 142-160, March–April 2006.

BATZER, M. A.; DEININGER, P. L. A human-specific subfamily of Alu sequences. **Genomics**, v. 9, p. 481-487, 1991.

BATZER, M. A.; DEININGER, P. L. Repeats of Alu and Human Diversity of Genomic. **Genetics**, v. 3, p. 370-9, 2002.

BEUTEN, J. et al. Wide disparity in genetic admixture among Mexican Americans from San Antonio, TX. **Ann Hum Genet**, v. 75, n. 4, p. 529-38, Jul 2011. ISSN 1469-1809 (Electronic)

- 0003-4800 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21592109 >.
- BOCK, C. A. H. V. et al. Results from a prostate cancer admixture mapping study in African-American men. **Hum Genet**, v. 126, p. 637-642, 2009.
- BONILLA, C. et al. Prostate cancer susceptibility Loci identified on chromosome 12 in African Americans. **PLoS One**, v. 6, n. 2, p. e16044, 2011. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21358824 >.
- BONILLA, C. 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, p. 139-153, 2004.
- BRYC, K. et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. **PNAS**, v. 107, n. 2, 2010.
- CALLEGARI-JACQUES, S. M. et al. Historical Genetics: Spatiotemporal Analysis of the Formation of the Brazilian Population. **Am J Hum Biol**, v. 15, p. 824-34, 2003.
- CALLEGARI-JACQUES, S. M.; SALZANO, F. M. Brazilian Indian/non-Indian interactions and their effects. **Ciência e Cultura: J Braz Assoc Advancement Sci**, v. 51, p. 166-174, 1999.
- CARVALHO-SILVA, D. R. et al. The phylogeography of Brazilian Y-chromosome lineages. **Am J Hum Genet**, v. 68, n. 1, p. 281-6, Jan 2001. ISSN 0002-9297 (Print) 0002-9297 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11090340 >.
- CAVALLI-SFORZA, L. L.; MENOZZI, P.; PIAZZA, A. The History and Geography of Human Genes. **Princeton University Press**, 1996.
- CAVALLI-SFORZA, L. L.; MENOZZI, P.; PIAZZA, A. History and Geography of Human Genes. **Princeton University Press**, 1994.
- COLLINS, F. S.; BROOKS, L. D.; CHAKRAVARTI, A. A DNA polymorphism discovery resource for research on human genetic variation. **Genome**, v. 8, p. 1229-1231, 1998.
- COMAS, D. et al. Recent insertion of an Alu element within a polymorphic human-specific Alu insertion. **Mol Biol Evol**, v. 18, n. 1, p. 85-8, 2001.
- DEAN, M. et al. Polymorphic admixture typing in human ethnic populations. **Am J Hum Genet**, v. 55, p. 788-808, 1994.
- ERDEI, E. et al. Self-Reported Ethnicity and Genetic Ancestry in Relation to Oral Cancer and Pre-Cancer in Puerto Rico. **PLoS ONE**, v. 6, n. 8, 2011.

- 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. 1567-1587, 2003.
- FELIX, G. E. S. et al. Ancestry informative markers and complete blood count parameters in Brazilian blood donors. **REVISTA BRASILEIRA DE HEMATOLOGIA E HEMOTERAPIA**, v. 32, p. 282-285, 2010.
- FUNAI. Quem São. Brasília, 2010. Disponível em: < <http://www.funai.gov.br/index.php/indios-no-brasil/quem-sao> >. Acesso em: Mar. 2010.
- GARCIA, M. et al. Global Cancer Facts. **Am Cancer Society**, 2007.
- GATTÁS, G. J. F.; GOMES, L.; KOHLER, P. Ethnicity and glutathione S-transferase (GSTM1/GSTT1) polymorphisms in a Brazilian population. **Braz J Med Biol Res**, v. 37, p. 451-458, 2004.
- GIRI, V. N. et al. Race, genetic West African ancestry, and prostate cancer prediction by prostate-specific antigen in prospectively screened high-risk men. **Cancer Prev Res (Phila)**, v. 2, n. 3, p. 244-50, Mar 2009. ISSN 1940-6215 (Electronic) 1940-6215 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19240249 >.
- GROVER, D. et al. Alu repeat analysis in the complete humana genome: trends and variation with respect to genomic composition. **Bioinformatics. Oxford University Press**, v. 20, n. 6, p. 813-817, 2004.
- HINDS, D. A. et al. Matching strategies for genetic association studies in structured population. **Am J Hum Genet**, v. 74, p. 317-325, 2004.
- HOGGART, C. J. et al. Control of confounding of genetic associations in stratified populations. **Am J Hum Genet**, v. 72, n. 6, p. 1492-1504, Jun 2003. ISSN 0002-9297 (Print) 0002-9297 (Linking). Disponível em: < http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12817591 >.
- IBGE. CENSO 2009. Brasília, 2009. Disponível em: < <http://www.ibge.gov.br/cidadesat/topwindow.htm?1> >. Acesso em: setembro. 2010.
- IBGE. INSTITUTO BRASILEIRO DE GEOPGRAFIA E ESTATÍSTICA. 2012. Disponível em: < www.ibge.gov.br/ >.
- INCA. Instituto Nacional de Câncer José Alencar Gomes da Silva. 2011. Disponível em: < <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/prostata/definicao> >. Acesso em: 15 de dezembro. 2011.
- JURKA, J. et al. Duplication, coclustering, and selection of human Alu retrotransposons. **Proc Natl Acad Sci U S A**, v. 101, p. 1268-1272, 2004.

- KAMOTO, T. et al. Association of a genetic polymorphism of the E-cadherin gene with prostate cancer in a Japanese population. **Jpn J Clin Oncol**, v. 35, n. 3, p. 158-61, Mar 2005. ISSN 0368-2811 (Print)
0368-2811 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15741307 >.
- KLEIN, H. S. **As origens africanas dos escravos brasileiros. Em Homo brasiliis.** São Paulo: Ed. Funpec, 2002.
- LINS, T. C. et al. Genetic composition of Brazilian population samples based on a set of twenty-eight ancestry informative SNPs. **Am J Hum Biol**, v. 22, n. 2, p. 187-92, Mar-Apr 2010. ISSN 1520-6300 (Electronic)
1042-0533 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19639555 >.
- LUIZON, M. R. et al. Ancestry informative markers in Amerindians from Brazilian Amazon. **Am J Hum Biol**, v. 20, n. 1, p. 86-90, Jan-Feb 2008. ISSN 1042-0533 (Print)
1042-0533 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17963227 >.
- MARRERO, A. R. et al. Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. **Am J Hum Biol**, v. 17, n. 4, p. 496-506, Jul-Aug 2005. ISSN 1042-0533 (Print)
1042-0533 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15981186 >.
- MASON, T. E. et al. Association of CD14 Variant with Prostate Cancer in African American Men. **Prostate cancer and prostatic diseases**, v. 70, p. 262-269, 2010.
- MATTOSO, K. Q. **Ser escravo no Brasil.** São Paulo: Brasiliense, 1992.
- MICKLOS, D. A.; FREYER, G. A.; CROTTY, D. A. **A ciência do DNA.** Porto Alegre: Artmed, 2005.
- MYLES, S.; STONEKING, M.; TIMPSON, N. An assessment of the portability of ancestry informative markers between human populations. **BMC Med Genomics**, v. 2, n. 45, 2009.
- NUSSBAUM, R. L. et al. **Genética Médica.** Rio de Janeiro: Editora Guanabara Koogan S.A., 2002.
- OKOBIA, M. N. et al. Chromosome 8q24 variants are associated with prostate cancer risk in a high risk population of African ancestry. **Prostate**, Jan 12 2011. ISSN 1097-0045 (Electronic)
0270-4137 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21229559 >.

- PALMER, I. M.; SCHUTTE, A. E.; HUISMAN, H. W. Ethnic and gender differences regarding the insulin-blood pressure relationship. **Diabetes Res Clin Pract**, v. 85, p. 102-10, 2009.
- PARRA, E. J. et al. Estimating African American admixture proportions by use of population-specific alleles. **Am J Hum Genet**, v. 63, n. 6, p. 1839-51, Dec 1998. ISSN 0002-9297 (Print)
0002-9297 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9837836 >.
- PARRA, F. C. et al. Color and genomic ancestry in Brazilians. **Proceedings of the National Academy of Sciences of the United States of America** v. 100, p. 177-182, 2003.
- PATTERSON, N. et al. Genetic structure of a unique admixed population: implications for medical research. **Hum Mol Genets**, v. 19, n. 3, 2010.
- PENA, S. D. et al. DNA tests probe the genomic ancestry of Brazilians. **Braz J Med Biol Res**, v. 42, n. 10, p. 870-6, Oct 2009. ISSN 1414-431X (Electronic)
0100-879X (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19738982 >.
- PENA, S. D. 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, Feb 16 2011.
- PENA, S. D. J. et al. **Retrato molecular do Brasil**. *Ciência Hoje*. 27: 16-25 p. 2000.
- POOKOT, D. et al. The E-Cadherin - 160 C/A Polymorphism and Prostate Cancer Risk in White and Black American Men. **J Urol**, v. 176, p. 793-796, 2006.
- PRITCHARD, J. K.; ROSENBERG, N. A. Use of unlinked genetic markers to detect population stratification in association studies. **Am J Hum Genet**, v. 65, p. 220-228, 1999.
- PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945-59, 2000.
- RAMENSKY, V.; BORK, P.; SUNYAEV, S. Human non-synonymous SNPs: server and survey. **Nucleic Acids Res**, v. 30, n. 17, 2002.
- REDDY, S. et al. Prostate cancer in black and white Americans. **Cancer Metastasis Rev**, v. 22, n. 1, p. 83-6, Mar 2003. ISSN 0167-7659 (Print)
0167-7659 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12716039 >.
- RESQUE, R. L. et al. Estimates of interethnic admixture in the Brazilian population using a panel of 24 X-linked insertion/deletion markers. **Am J Hum Biol**, v. 22, n. 6, p. 849-52, Nov-Dec 2010. ISSN 1520-6300 (Electronic)

1042-0533 (Linking). Disponível em: <
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20865761 >.

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

ROUZIC, A. L. E.; DECELIERE, G. Models of the population genetics of transposable elements. **Genet Res**, v. 85, p. 171-181, 2005.

SALES, F. **Memória de Ilhéus: em comemoração do centenário de sua elevação a cidade**. São Paulo: Edições GRD, 1981.

SANTOS, L. M. **Polimorfismos do gene de E-caderina e o risco de câncer de próstata em uma população da Bahia**. 2010. Dissertação (Mestre). Programa de Pós- graduação em Genética e Biologia Molecular, Universidade Estadual de Santa Cruz, Ilhéus, 2010.

SANTOS, R. V. et al. Color, race, and genomic ancestry in Brazil: dialogues between anthropology and genetics. **Curr Anthropol**, v. 50, n. 6, p. 787-819, Dec 2009. ISSN 0011-3204 (Print)

0011-3204 (Linking). Disponível em: <

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20614657 >.

SHRIVER, M. D. et al. Skin pigmentation, biogeographical ancestry and admixture mapping. **Hum Genet**, v. 112, n. 4, p. 387-99, Apr 2003. ISSN 0340-6717 (Print)

0340-6717 (Linking). Disponível em: <

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12579416 >.

SHRIVER, M. D. et al. Ethnic-affiliation estimation by use of population-specific DNA markers. **Am J Hum Genet**, v. 60, n. 4, p. 957-64, 1997.

STEWART, S. et al. The clinical consequences and challenges of hypertension in urban dwelling black Africans: Insights from the Heart of Soweto Study. **Int J Cardiol**, 2009.

TAVERES, L. H. D. **História da Bahia**. São Paulo, Editora UNESP. Salvador, BA: EDUFBA, 2001.

TROTTIER, G. et al. Nutraceuticals and prostate cancer prevention: a current review. **Nat Rev Urol**, v. 7, n. 1, p. 21-30, Jan 2010. ISSN 1759-4820 (Electronic)

1759-4812 (Linking). Disponível em: <

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19997071 >.

WHITMAN, E. J. et al. Prostate cancer risk allele specific for African descent associates with pathologic stage at prostatectomy. **Cancer Epidemiol Biomarkers Prev**, v. 19, n. 1, p. 1-8, Jan 2010. ISSN 1538-7755 (Electronic)

1055-9965 (Linking). Disponível em: <

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20056617 >.

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. 822-828, 2006.