

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL



Composição Genética de Duas Populações Afro-derivadas Brasileiras
Inferida a partir de Marcadores Informativos de Ancestralidade

Carolina Carvalho Gontijo

Brasília

2008

Composição Genética de Duas Populações Afro-derivadas Brasileiras
Inferida a partir de Marcadores Informativos de Ancestralidade

Dissertação apresentada ao Programa de Pós
Graduação em Biologia Animal da Universidade
de Brasília para a obtenção do título de Mestre.

Carolina Carvalho Gontijo
Orientadora: Prof. Dra. Silviene Fabiana de Oliveira

Brasília

2008

Trabalho desenvolvido no Laboratório de Genética da Universidade de Brasília – UnB, com apoio do Centro Interdepartamental de Biologia Experimental e Biotecnologia da UniR e da Faculdade Adventista do Recôncavo Baiano e suporte financeiro da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e FINATEC (Fundação de Empreendimentos Científicos e Tecnológicos).



Dissertação de mestrado

Carolina Carvalho Gontijo

Composição Genética de Duas Populações Afro-derivadas Brasileiras Inferida
a partir de Marcadores Informativos de Ancestralidade

Comissão examinadora

Prof. Dra. Silviene Fabiana de Oliveira

Presidente/Orientadora

UnB

Prof. Dr. Celso Teixeira Mendes Júnior

Membro titular externo não vinculado ao programa

USP

Prof. Dr. Luzitano Brandão Ferreira

Membro titular externo não vinculado ao programa

UniCEUB

Prof^a. Dr^a. Zulmira Guerrero Marques Lacava

Membro suplente interno vinculado ao programa

UnB

Brasília, 29 de maio de 2008

A minha família,

Dedico.

Agradecimentos

À minha orientadora, Dra. Silviene Fabiana de Oliveira, pelo apoio, confiança e amizade;

Aos membros da banca examinadora, pela disponibilidade em avaliar esse trabalho e, assim, contribuir para meu crescimento acadêmico;

À equipe do Centro Interdepartamental de Biologia Experimental e Biotecnologia da UniR, em especial a Adriana Nunes, e ao Dr. Wellington Silva, da Faculdade Adventista do Recôncavo Baiano, pela parceria firmada que possibilitou a execução desse trabalho;

Às agências de fomento CAPES, CNPq e FINATEC pelo apoio financeiro.

Aos laboratórios de Microbiologia e Biologia Molecular, pelo apoio laboratorial.

A toda a equipe do Laboratório de Genética da UnB, por tornarem o ambiente de trabalho sempre agradável;

Aos amigos do Laboratório de Genética da UnB pela amizade e grande ajuda. Mila, Neide, Dudu e Arthur, seu apoio me manteve de pé em vários momentos. Rafaela, Diana e Gabriel que, além da amizade, doaram seu tempo e conhecimento;

Ao André, pelo amor, carinho e apoio incondicionais, por estar sempre ao meu lado nos momentos mais difíceis e pela ajuda com o mapa e figuras;

A minha família e amigos, por compreenderem minha ausência e mau humor durante os últimos dois anos;

Agradeço.

Índice

Resumo 1

Abstract 2

Introdução 3

A constituição da população humana brasileira 3

Remanescentes de quilombo e comunidades afro-derivadas 6

Composição genética de remanescentes de quilombo brasileiros 6

Marcadores genéticos 9

Marcadores informativos de ancestralidade 10

Objetivos 13

Material e Métodos 14

Populações analisadas 14

 Santo Antônio do Guaporé 14

 Santiago do Iguape 15

 Distrito Federal 16

Coleta e processamento de material biológico 16

 Santo Antônio do Guaporé 16

 Santiago do Iguape 17

Extração de DNA 18

Marcadores moleculares selecionados 18

Análise dos marcadores genéticos moleculares 21

Análise estatística 24

 Análises descritivas 24

 Frequências gênicas e genotípicas 24

 Aderência ao Equilíbrio de Hardy-Weinberg 25

 Desequilíbrio de ligação 25

 Comparação das frequências gênicas e genotípicas 26

 Estatísticas F 26

 Definição dos intervalos de confiança 26

 Estimativa de mistura 27

 Componente principal 27

Resultados 29

Descrição das populações 29

 Santo Antônio do Guaporé 29

 Santiago do Iguape 32

Comparação das frequências gênicas e genotípicas 35

Estatísticas F 37

Fis – Coeficiente de endogamia 37

Fst – Diversidade inter-populacional 37

Estimativa de mistura 38

Análise de componente principal 39

Discussão 44

Equilíbrio de Hardy-Weinberg e Desequilíbrio de Ligação 44

Frequências alélicas e genotípicas 46

Estimativas de mistura 48

Estatísticas F 49

Análise de Coordenada principal 49

Conclusões 51

Referência bibliográfica 52

Referências da Internet 62

Anexo 1: Protocolos dos Procedimentos Laboratoriais 63

Extração de DNA 63

Reação em Cadeia da Polimerase 64

Digestão enzimática 64

Eletroforese 65

Coloração 66

Leitura dos géis 66

Índice de tabelas

- Tabela 1.** Estimativas de mistura étnica em várias comunidades afro-derivadas brasileiras (referências listadas na tabela 2). 8
- Tabela 2.** Referências das estimativas de mistura compiladas na tabela 1. 9
- Tabela 3.** Diferenciais de frequência médios - δ - encontrados entre as frequências do alelo *1 nas populações parentais. para o marcador DRD2, o δ foi calculado para cada haplótipo e para o marcador GC, para os alelos *1F e *1S. 19
- Tabela 4.** Iniciadores, temperatura padronizada para pareamento dos iniciadores (T) e enzimas de restrição utilizadas para a análise dos aims selecionados. 22
- Tabela 5.** Misturas de reagentes utilizadas nas reações de PCR para cada sistema. 23
- Tabela 6.** Misturas de reagentes, temperatura e tempo utilizados nas reações de restrição para cada sistema. 23
- Tabela 7.** Haplótipos possíveis considerando três sítios de restrição, *loci* DRD2A, DRD2B e DRD2D, presentes no gene DRD2. 24
- Tabela 8.** Médias ponderadas das frequências alélicas dos marcadores AIMS encontradas na literatura para as populações parentais utilizadas na análise de mistura gênica. 28
- Tabela 9.** Distribuição alélica dos 15 loci aims em Santo Antônio do Guaporé ($n = 31$): frequências alélicas e genotípicas e heterozigoses esperada (H_e) e observada (H_o). 29
- Tabela 10.** Frequências dos haplótipos DRD2*TaqI* observadas em Santo Antônio do Guaporé. 30
- Tabela 11.** Santo Antônio do Guaporé: valores de p com respectivos erros padrão (ep) para os teste de aderência ao Equilíbrio de Hardy-Weinberg (EHW), teste para deficiência de heterozigotos (DH) e teste para excesso de heterozigotos (EH). 30
- Tabela 12.** Teste de desequilíbrio de ligação *locus a locus*. em destaque, valores que indicam existência de desequilíbrio de ligação antes de aplicada a correção de Bonferroni. EP: erro padrão. 31
- Tabela 13.** Distribuição alélica dos 11 loci AIMS em Santiago do Iguape (n total = 37): frequências alélicas e genotípicas e heterozigoses esperada (H_e) e observada (H_o). 32
- Tabela 14.** Frequências dos haplótipos DRD2*TaqI* observadas em Santiago do Iguape. 33
- Tabela 15.** Santiago do Iguape: valores de P com respectivos erros padrão (EP). EHW - teste de aderência ao Equilíbrio de Hardy-Weinberg; DH - teste para deficiência de heterozigotos; EH - teste para excesso de heterozigotos. 33
- Tabela 16.** Santiago do Iguape: teste de desequilíbrio de ligação *locus a locus*. em destaque, valores que indicam existência de desequilíbrio de ligação antes de aplicada a correção de bonferroni. EP: erro padrão. 34

Tabela 17. Valores de P obtidos pelo Teste Exato de Fisher para diferenciação gênica e genotípica entre os pares de populações analisadas. 35

Tabela 18. Valores de Fis e intervalos de confiança (95% e 99%) nas populações de Santo Antônio do Guaporé e Santiago do Iguape e comparação com o observado para o Distrito Federal. 37

Tabela 19. Valores de Fst e intervalos de confiança (95% e 99%) para os três pares de populações e para as três populações em conjunto. 37

Tabela 20. Estimativas de mistura (m) para Santo Antônio do Guaporé e Santiago do Iguape calculadas a partir de todos os marcadores analisados para cada. 38

Tabela 21. Estimativas de mistura estimadas a partir de 12 dos marcadores aqui analisados para as populações de Santo Antônio do Guaporé, Santiago do Iguape e Distrito Federal. 38

Índice de figuras

- Figura 1.** Proporções de indivíduos auto-declarados como pretos, pardos, brancos ou outros levantado pelo PNAD 2006 em todas as regiões do Brasil. 3
- Figura 2.** Estimativas de mistura genética obtidas a partir de dados de marcadores STRs para as cinco regiões brasileiras (Godinho, 2008). 4
- Figura 3.** Mapa mostrando a localização das comunidades de Santo Antônio do Guaporé, Santiago do Iguape e das outras comunidades urbanas e quilombolas com as quais foram comparadas. 14
- Figura 4.** Frequências dos marcadores disponíveis na literatura (entre os aqui analisados) para as populações afro-derivadas Santo Antônio do Guaporé, Santiago do Iguape, Kalunga, Rio das Rãs, Riacho de Sacutiaba, Mocambo, Barra, São Gonçalo e Valongo. 36
- Figura 5.** Distância de Nei: estimativas obtidas entre as populações em análise (SAG e STI), as populações parentais utilizadas no cálculo de mistura (AME, AFR e EUR) e do Distrito Federal (DF). 39
- Figura 6.** Distância de Nei: estimativas obtidas entre as populações em análise, as populações parentais utilizadas no cálculo de mistura, Distrito Federal e os remanescentes de quilombo Kalunga (KAL), Rio das Rãs (RAS), Riacho de Sacutiaba (SAC) e Mocambo (MOC). 40
- Figura 7.** Distância de Nei: estimativas obtidas entre as populações em análise, as populações parentais utilizadas no cálculo de mistura, Distrito Federal e os remanescentes de quilombo São Gonçalo (SGO), Barra (BAR), Valongo (VAL) e Kalunga (KAL). 40
- Figura 8.** Componente principal gerado a partir de uma matriz de distância (Nei) entre Santo Antônio do Guaporé, Santiago do Iguape, Distrito Federal e as três populações parentais utilizadas nas análises de mistura. 41
- Figura 9.** Componente principal gerado a partir de uma matriz de distância (Nei) entre Santo Antônio do Guaporé, Santiago do Iguape, Distrito Federal, os remanescentes de quilombo Kalunga (KAL), Rio das Rãs (RAS), Riacho de Sacutiaba (SAC), Mocambo (MOC) e as três populações parentais utilizadas nas análises de mistura. 42
- Figura 10.** Componente principal gerado a partir de uma matriz de distância (Nei) entre Santo Antônio do Guaporé, Santiago do Iguape, Distrito Federal, os remanescentes de quilombo São Gonçalo (SGO), Barra (BAR), Valongo (VAL), Kalunga (KAL) e as três populações parentais utilizadas nas análises de mistura. 43
- Figura 11.** Modelos de formação propostos por Pfaff e colaboradores (2001): hybrid-isolation (HI) e continuous-gene-flow (CGF). Modificado de Pfaff e colaboradores (2001). 45
- Figura 12.** Quantidade de ALD esperada sob cada modelo de mistura (HI e CGF) para *loci* não ligados e para *loci* ligados a 5cM. HI: modelo *hybrid-isolation*; CGF: modelo *continuous-gene-*

flow; DL: desequilíbrio de ligação. 5cM: modelo para marcadores ligados. *Unlinked*: modelo para marcadores sem ligação. Figura retirada de Pfaff *et al.* (2001). 45

Resumo

No Brasil, uma das principais formas de resistência à escravidão pelos escravos de ancestralidade africana foi a fuga e concentração em locais de difícil acesso, formando comunidades quilombolas. Hoje, remanescentes de quilombo, originados ou não dos antigos quilombos, se distribuem por todo o país. De forma geral, essas comunidades apresentam em sua composição genética grande contribuição africana, seguida, em graus variados, das contribuições ameríndia e européia. No presente trabalho, duas comunidades afro-derivadas, Santo Antônio do Guaporé (SAG) em Rondônia e Santiago do Iguape (STI) na Bahia, foram analisadas quanto a 13 marcadores informativos de ancestralidade – AIMS, i.e.: marcadores que apresentam diferenciais de frequência maiores que 30% entre grupos definidos geográfica ou etnicamente. Os dados gerados foram utilizados para descrever SAG e STI quanto à adequação da distribuição genotípica ao Equilíbrio de Hardy-Weinberg, deficiência e excesso de heterozigotos, desequilíbrio de ligação e mistura; e para compará-las a outras populações urbanas e afro-derivadas. Os resultados indicaram que as duas populações, embora apresentem semelhanças, como a alta contribuição do grupo parental ameríndio em suas composições genéticas (37% em SAG e 56% em STI), apresentam diferenças significativas quando comparadas entre si pelo teste Exato de Fisher para diferenciação gênica e genotípica e pela estimativa de F_{st} . SAG e STI são também diferentes da população do Distrito Federal, com quem foram comparadas pelos mesmos testes. As análises de componente principal mostraram que as duas populações analisadas posicionam-se entre as parentais africana e ameríndia, o que vai ao encontro das estimativas de mistura. Em suma, as duas população aqui averiguadas apresentam alta contribuição africana e ameríndia, como outros remanescentes do país, mas se distinguem em decorrência dos inúmeros fatores associados a suas diferentes histórias.

Abstract

In Brazil, many African descendant slaves avoided slavery by running away and hiding in communities called *quilombos*. Nowadays, *quilombo* remnants, either originated from old *quilombos* or defined by anthropological traces, are spread across the whole country. The present work analyzed 13 ancestry informative markers (AIMs) in order to detect the amount of contribution from the parental populations (Amerindians, Africans and Europeans) in the constitution of two afro-derived communities (*Santo Antônio do Guaporé* – SAG and *Santiago do Iguape* – STI). AIMs are markers whose frequencies differ by more than 30% between groups defined by ethnicity or geography. Statistical analysis such as genic and genotypic differentiation, F statistics and principal component analysis (PCA) were performed to compare SAG and STI to one another and to other Brazilian populations, mainly to *Distrito Federal* (DF). Our results pointed out that SAG and STI, despite their high Amerindian contributions (37% and 56%, respectively), are significantly different from each other and from DF, according to the Fisher Exact Test for genic and genotypic differentiation and by F_{st} estimates. In PCA analysis, both populations were located between the African and Amerindian parental groups, matching the admixture estimates. In summary, SAG and STI showed high African and Amerindian contributions, like other Brazilian afro-derived populations, but are different from each other in their genetic constitutions as a consequence of their different histories.

Introdução

A constituição da população humana brasileira

A constituição da população brasileira se deu pela miscigenação de três grupos étnicos principais: europeus, ameríndios e africanos. A forma como esses grupos se distribuíram e contribuíram para a formação das populações das diversas regiões do país se reflete na distribuição diferencial de fenótipos entre elas. De acordo com a Pesquisa Nacional por Amostra Domiciliar (PNAD) realizada em 2006 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), cerca de 49,5% dos mais de 180 milhões de brasileiros devem ter indicações morfológicas de contribuição africana (auto-definidos como pretos ou pardos, de acordo com a nomenclatura utilizada no levantamento). Porém, a porcentagem desta contribuição varia entre as cinco regiões brasileiras, sendo a maior encontrada na região Nordeste (figura 1).

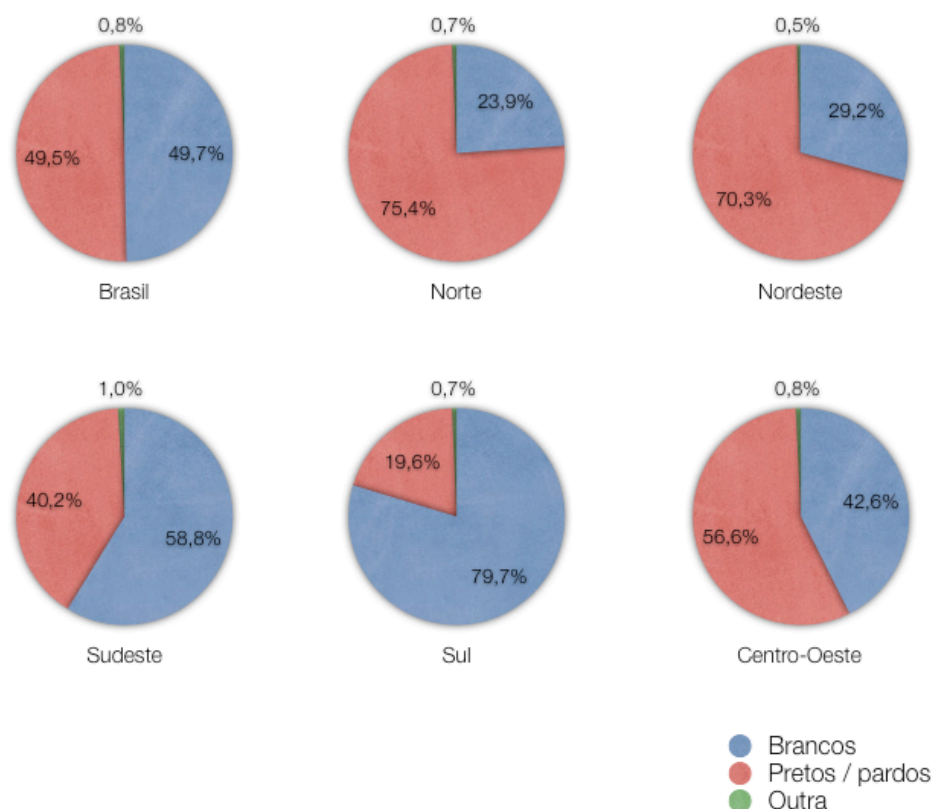


Figura 1. Proporções de indivíduos auto-declarados como pretos, pardos, brancos ou outros levantado pelo PNAD 2006 em todas as regiões do Brasil.

Godinho (2008) estimou a contribuição dos grupos parentais africano, europeu e ameríndio a partir de marcadores microssatélites para as cinco regiões brasileiras. Os resultados obtidos (figura 2) mostram proporções diferentes daquelas indicadas pela auto-classificação em raças (figura 1). Em geral, as contribuições ameríndia e europeia são subestimadas pela auto-classificação e a africana, superestimada. É provável que isso se deva à grande proporção de pardos observada em todo o país: essa categoria inclui indivíduos cuja ancestralidade inclui mais de um grupo parental, diluindo as contribuições ameríndia e europeia.

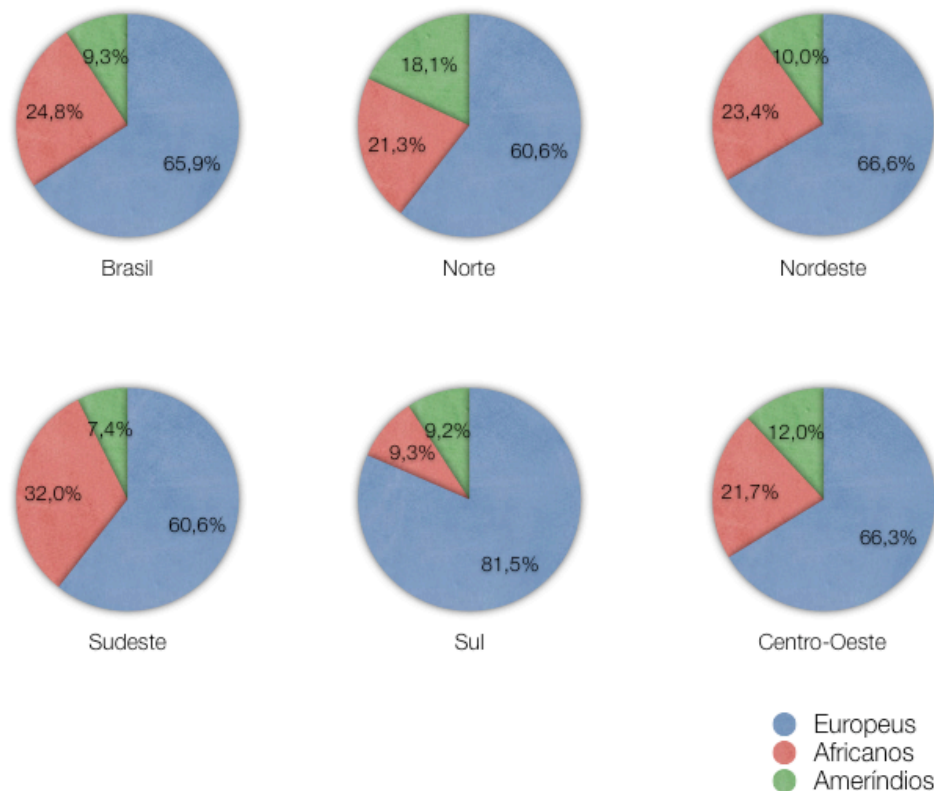


Figura 2. Estimativas de mistura genética obtidas a partir de dados de marcadores STRs para as cinco regiões brasileiras (Godinho, 2008).

O fluxo de europeus para o Brasil se iniciou em 1500 com a chegada dos primeiros navios portugueses. Portugal foi quase sempre a principal origem de imigrantes, ainda que a partir de meados do século XIX, com o incentivo governamental para a entrada de imigrantes como alternativa à mão de obra escrava, o fluxo de imigrantes oriundos de vários países, em especial da Itália, tenha se intensificado e mantido até o início do século XX (Fausto, 2002). Esses novos imigrantes se fixaram principalmente no Sul e Sudeste do país em áreas urbanas e rurais (Fausto, 2002).

O tamanho da população ameríndia presente no território brasileiro à época da chegada dos portugueses é estimado em 1 a 10 milhões de indivíduos (Fundação Nacional do Índio - FUNAI, 2006), principalmente Tupi-Guaranis, grupos que compartilham semelhanças culturais e lingüísticas (Fausto, 2001). O contato português, pacífico em alguns casos e hostil em outros, gerou miscigenação e a redução drástica da população ameríndia brasileira para os hoje estimados 345 mil indivíduos (cerca de 0,2% da população do país) que habitam as aproximadamente 215 sociedades indígenas e outros 100 a 190 mil vivendo fora dessas terras, incluindo áreas urbanas (FUNAI, 2006). Há também grupos que estão requerendo seu reconhecimento como indígena junto à FUNAI, além de indícios da existência de cerca de 50 grupos ainda isolados (FUNAI, 2006). Diversas etnias foram extintas, mas a contribuição genética gerada pela miscigenação com europeus e africanos é significativa.

Estima-se que 40% dos escravos africanos trazidos para a América (entre 3,6 e 10 milhões de indivíduos) tiveram o Brasil como destino. Eram principalmente bantos e sudaneses capturados nas costas leste e oeste da África (Salzano e Freire-Maia, 1967). De acordo com Mello e Souza (2006), ao longo dos séculos em que foram trazidos para o Brasil, esses escravos eram capturados e embarcados em diferentes regiões da África sub-saariana: no início do século XVI, sua principal origem eram os portos de Congo e, em menor grau, Angola. Com a crescente influência e dominação holandesa, inglesa e francesa na região do Congo, a coroa portuguesa estabeleceu nos séculos seguintes portos e comércio em regiões mais ao sul de Angola e tornou Benguela seu principal porto. No início do século XVIII, como consequência do aumento na intensidade do comércio de escravos e da modificação do cenário político, houve uma interiorização da origem dos escravos, que passaram a ser capturados em regiões mais internas, próximas aos grandes lagos.

Na América, eles apresentaram várias formas de resistência como suicídio, agressão aos senhores de escravos e, especialmente, a fuga e concentração em locais de difícil acesso, que deu origem às comunidades isoladas geográfica e culturalmente denominadas no Brasil de *Quilombos* ou *Mocambos* (Vila Real, 1996; Neme e Andrade, 1987). Existem registros da presença de quilombos em todas as regiões do país, com exceção dos estados do Acre e Roraima (Anjos, 2006).

Remanescentes de quilombo e comunidades afro-derivadas

Nos dias atuais, no Brasil, assim como em vários outros países da América Latina, existem comunidades rurais cujas histórias e tradições orais as identificam como antigos *Quilombos* e que são denominadas no Brasil *Remanescentes de Quilombos*. Até o momento, mais de 2000 comunidades com o perfil de ascendência de escravos africanos e seus dois milhões de descendentes foram registradas em todas as regiões brasileiras (Anjos, 2006). A maioria destas comunidades não foi ainda reconhecida pelo governo como comunidade quilombola por faltarem estudos antropológicos que comprovem sua origem ou condição (já que a classificação como remanescente de quilombo não está necessariamente atrelada à uma origem em um quilombo - Associação Brasileira de Antropologia - <http://www.abant.org.br/>). O não reconhecimento traz prejuízos para estas comunidades, já que estas não têm o direito de posse de suas terras assegurado (Fundação Cultural Palmares, 2006).

Apesar de possuírem diferentes histórias desde suas fundações no século passado, os remanescentes de *quilombo* e outras comunidades afro-derivadas, como Salvador, por exemplo, mantêm ainda hoje uma contribuição africana genética e cultural maior do que as comunidades brasileiras não afro-derivadas (Abe-Sandes, 2002; Lima, 2002; Souza, 2001). No entanto, o interesse pelas terras ocupadas por remanescentes de quilombo rurais e a proximidade e sua dependência cada vez maior dos centros urbanos pode vir a ocasionar a perda do patrimônio antropológico e biológico, como diversidade alélica, que caracterizam os indivíduos e as populações, devido, principalmente, à miscigenação, mas também à emigração. Tudo isso já é realidade para várias dessas comunidades, e por isso é de grande importância que sejam estudadas o quanto antes, em uma tentativa de resgate de uma parte relevante da história da formação da população brasileira.

Composição genética de remanescentes de quilombo brasileiros

Várias comunidades afro-derivadas brasileiras, remanescentes de quilombo ou não, já foram estudadas com relação à composição populacional acessada a partir da análise de marcadores genéticos clássicos e moleculares autossômicos e uniparentais. Exemplos dessas estimativas estão listados na Tabela 1. Na Tabela 2 estão apresentadas as referências bibliográficas utilizadas para a elaboração da Tabela 1.

De forma geral, os resultados da análise de contribuição das populações parentais na formação dessas populações indicaram uma grande contribuição da parental africana e contribuições variáveis das parentais ameríndia e europeia, embora existam diferenças entre a região Norte e as demais regiões e entre os quilombos de cada região. Na região Norte, a maioria das estimativas mostra contribuições africana e ameríndia altas, enquanto nas demais regiões a contribuição ameríndia é quase sempre inferior à europeia.

Outra diferença notável ocorre entre as estimativas obtidas para marcadores mitocondriais e do cromossomo Y: a contribuição materna ameríndia na formação dessas comunidades foi, em geral, maior que a europeia e a paterna ameríndia, menor que a europeia. Tal discrepância se explica por razões históricas, já que houve, desde o início da colonização do Brasil, reprodução direcional entre homens europeus e mulheres ameríndias e africanas (Alencastro, 2000).

Tabela 1. Estimativas de mistura étnica em várias comunidades afro-derivadas brasileiras (referências listadas na tabela 2).

	Comunidade	Tipo de marcador	Estimativa de contribuição (%)			Referência		Comunidade	Tipo de marcador	Estimativa de contribuição (%)			Referência	
			Ameríndia	Africana	Européia					Ameríndia	Africana	Européia		
Sul	Paredão - RS	Clássicos	2,80	79,10	18,10	1	Rio das Rãs - BA	AIMs	4,60	95,40	0,00	5		
		Clássicos	0,00	80,20	19,80	2		Clássicos e <i>Alu</i>	13,72	27,69	58,59	6		
		STR-Y	0,00	65,00	35,00	3		STR-Y	0,00	38,00	62,00	7		
		Mitocondriais	27,00	58,00	15,00	3		Mitocondriais	14,0	86,00	0,00	8		
		STRs e VNTRs	12,80	49,20	38,00	3		STRs	3,43	53,88	42,69	9		
Valongo - SC		Clássicos	0,00	97,30	2,70	4	AIMs	8,66	60,56	30,78	9			
		AIMs	0,00	68,10	31,90	5	Rio de Sa-cutiaba - BA	Clássicos e <i>Alu</i>	0,00	72,29	27,71	6		
Centro-Oeste	Kalunga - GO	STR-Y	6,00	27,00	67,00	7		STR-Y	0,00	5,00	95,00	7		
		Mitocondriais	13,3	86,7	0,00	8		Mitocondriais	15,4	84,6	0,00	8		
		STRs	3,33	57,38	39,29	9		STRs	0,61	35,08	64,31	9		
		AIMs	9,18	61,05	29,77	9		AIMs	2,73	53,94	43,33	9		
		Clássicos	32,30	49,70	18,00	1	Nordeste	São Gonçalo - BA	Clássicos e STRs	14,12	53,49	32,39	10	
Cameté - PA		Clássicos	34,10	48,00	17,90	2			STR-Y	0,00	59,72	40,28	10	
		STR-Y	18,00	28,00	54,00	3			AIMs	10,30	68,10	21,60	5	
		Mitocondriais	24,00	59,00	17,00	3			Cajueiro - MA	Clássicos	0,00	67,40	32,60	11
		STRs e VNTRs	24,20	53,40	22,40	3				STR-Y	0,00	78,00	22,00	3
Norte	Trombetas - PA	Clássicos	11,00	62,00	27,00	13	Mitocondrial	24,00		59,00	17,00	3		
		Clássicos	23,80	56,40	19,80	2	STRs e VNTRS	25,00		48,80	26,20	3		
		STR-Y	13,00	84,00	3,00	3	Mimbó - PI	Clássicos		17,00	61,00	22,00	12	
		Mitocondriais	13,00	80,00	7,00	3								
		Mitocondriais	43,70	56,3	0,0	14								
		STRs e VNTRs	9,90	57,60	32,50	3	Sítio Velho - PI	Clássicos	16,00	72,00	12,00	12		
		Curiaú - AM		Clássicos	0,00	73,60	26,40	15	Mocambo - SE	Clássicos e <i>Alu</i>	10,32	32,00	57,57	6
Mitocondriais	47,00			53,00	0,00	16	STR-Y	7,00		7,00	86,00	7		
STRs e VNTRs	24,00			42,80	33,20	17	Mitocondriais	21,4		78,6	0,00	8		
Pacoval - AM		Clássicos	28,30	44,30	27,40	15	STR	13,62		46,57	39,81	9		
		STRs e VNTRs	37,50	51,30	11,20	17	AIMs	18,81		43,42	37,77	9		

Tabela 2. Referências das estimativas de mistura compiladas na tabela 1.

Número	Referência
1	Bortolini <i>et al.</i> , 1992
2	Bortolini <i>et al.</i> , 1995
3	Bortolini <i>et al.</i> , 1999
4	Souza e Culpí, 2005
5	Luizon, 2007
6	Pedrosa e Oliveira, 2004
7	Ribeiro, 2005
8	Oliveira <i>et al.</i> , 2005
9	Pedrosa, 2006
10	Souza, 2001
11	Bortolini <i>et al.</i> , 1998a
12	Arpini-Sampaio <i>et al.</i> , 1999
13	Schneider <i>et al.</i> , 1987
14	Carvalho <i>et al.</i> , 2008
15	Guerreiro <i>et al.</i> , 1999
16	Ribeiro-dos-Santos <i>et al.</i> , 2002
17	Vallinoto <i>et al.</i> , 2003

Marcadores genéticos

Marcadores genéticos são regiões do genoma que apresentam variabilidade e possibilitam a diferenciação entre indivíduos e/ou populações. O estudo desses marcadores auxilia no entendimento das relações evolutivas entre populações e na reconstrução das rotas e origens dos povoamentos (Salzano e Callegari-Jacques, 1988; Saitou, 1995). Estudos sobre a evolução humana baseados em genética de populações baseiam-se na comparação de frequências gênicas entre populações, buscando inferir sua história. Para esse tipo de estudo é particularmente importante incluir *loci* polimórficos (Saitou, 1995), marcadores genéticos cujo alelo mais comum apresenta uma frequência populacional de no máximo 99%.

Nos primeiros trabalhos dessa natureza, o número de *loci* utilizados para a comparação de duas populações era muito baixo. Com o desenvolvimento da eletroforese em suporte sólido (Smithies, 1955), a descoberta das endonucleases (Market e Müller, 1959) e de inúmeros polimorfismos protéicos, a quantidade de dados relevantes acerca de frequências gênicas aumentou substancialmente. Nas duas últimas décadas foram desenvolvidas novas tecnologias para a avaliação de polimorfismos de seqüências de DNA que possibilitaram um avanço no estudo da variabilidade genética, inclusive dos seres humanos. Esses novos estudos revelaram incongruências entre existir uma variabilidade muito maior que a determinada por estudos de marcadores clássicos (Jorde *et al.*, 1995; Bortolini *et al.*, 1998b; Hutz *et al.*, 1999).

Além da dificuldade de compreensão da história evolutiva de populações humanas decorrente das diferentes respostas obtidas da análise de marcadores genéticos de diferentes naturezas – como clássicos, moleculares e uniparentais (Lum *et al.*, 1998), as relações entre elas não seguem padrões rígidos e, conseqüentemente, não geram padrões filogenéticos simples, baseados apenas em distâncias geográficas. Isso resulta de eventos demográficos como a migração, que tende a homogeneizar populações (Nei e Roychoudhury, 1982), eventos de fusão e fissão (Neel e Weiss, 1975) e conseqüentes alterações drásticas do tamanho populacional (gargalos populacionais ou expansões rápidas) causadas também por guerras, catástrofes e inovações tecnológicas, por exemplo. De todos esses eventos podem resultar fatores estocásticos como a deriva genética e fluxo gênico que tornam mais complexos os processos demográficos que geram a variabilidade e estrutura genética observados em um dado momento da história de uma população (Thompson, 1979).

Marcadores informativos de ancestralidade

Na reconstrução da história evolutiva da espécie, considera-se, desde a descrição dos primeiros marcadores genéticos para humanos, que os melhores locais do genoma a serem analisados são aqueles em que podemos diferenciar uma população de outra, seja pela frequência diferencial ou pela ocorrência de um alelo em uma população e em outra não. Os primeiros marcadores com descrições de que apresentavam alelos específicos em uma dada população foram denominados “polimorfismos privados” (Neel, 1973). Exemplos dessa classe de marcadores são os alelos: S da hemoglobina, de origem africana, $\Delta F508$ do gene do CFTR (*Cystic Fibrosis Transmembrane Conductance Regulator*), de origem européia, anidrase carbônica II *Baniwa*, específico da tribo *Baniwa* e o DYS19, que define um haplótipo de Y específico de ameríndios.

Porém, marcadores com alelos específicos de populações são muito raros. Por isso, a busca por marcadores que apresentavam diferenciais de frequência se intensificou a partir da década de 90. Hoje são denominados PSAs (*Population-specific Alleles*) ou AIMS (*Ancestry Informative Markers*) e são conceituados como marcadores moleculares que apresentam grande diferença de frequência alélica entre populações, podendo chegar à ausência ou presença exclusiva de um determinado alelo em uma dada população (Shriver *et al.*, 1997). Os AIMS apresentam grandes diferenciais de frequência entre

grupos étnicos: $\delta > 50\%$ (Shriver *et al.*, 1997), ou $\delta > 30\%$ (Bonilla *et al.*, 2004), enquanto o valor médio encontrado para o diferencial de frequência (δ) para loci bialélicos entre os grupos étnicos principais varia entre 15 e 20% (Dean *et al.*, 1994). Por isso, permitem estimar a contribuição de diferentes grupos na formação de uma população híbrida, ou seja, inferir sua ancestralidade. Como o valor de δ determina diretamente a precisão das estimativas de mistura (Chakraborty *et al.*, 1991), AIMS são os marcadores ideais para tais análises.

Eles podem pertencer a qualquer classe molecular, como inserções/deleções (indels), inserções *Alu* e polimorfismos de um único nucleotídeo (SNPs), todos com taxas de mutação diferentes e, portanto, com valores informativos diferentes.

Indels são marcadores genéticos bialélicos caracterizados pela inserção ou pela deleção de um fragmento ou par de base. Estima-se que representem 20% dos polimorfismos humanos, e podem ser agrupados em: a. originados de um evento de retrotransposição, com as inserções *Alu*, e b. originados por outros eventos mutacionais, como a indel AT3 (Weber *et al.*, 2002).

Inserções *Alu* são a maior família de repetições intercalantes curtas (*short interspersed elements* - SINEs), sendo exclusivamente encontradas no genoma de primatas. Representam cerca de 5% do genoma humano, somando cerca de um milhão de elementos, sendo que aproximadamente 1000 destas inserções pertencem a uma subfamília humano-específica (HS) com alto grau de identidade nucleotídica. Tipicamente, possuem 290 pb de comprimento e estima-se que essas inserções tenham surgido há 65 milhões de anos pela retrotransposição de um ou mais genes mestres. Elas não são transcritas e são muito estáveis no genoma. Duas características as tornam excelentes marcadores para o estudo de populações: 1. indivíduos que portam a mesma inserção o fazem por ancestralidade comum; 2. o estado ancestral (ausência da inserção) pode ser inferido (Arcot *et al.*, 1995; Novick *et al.*, 1995; Batzer *et al.*, 1994; Batzer *et al.*, 1991).

SNPs (polimorfismos de nucleotídeo único – do inglês, *single nucleotide polymorphism*) são variações na sequência de DNA ocasionadas por mutações pontuais do tipo substituição (mais comumente transições), e são, em geral, bialélicos, embora possam ser também tri- ou tetra-alelicos (Brookes, 1999). É a classe de marcadores mais comum no genoma humano, sendo que sua frequência é estimada em 1 a cada 200 a 300pb. Acredita-se que perfaçam cerca de 90% das variações de sequência em humanos

(Collins *et al.*, 1998), e que apenas 15% dos SNPs sejam exclusivos de população (Barbujani *et al.*, 1997), o que corrobora a idéia de que a maior parte da variação genética humana esteja dentro das populações e não entre elas (Brookes, 1999).

Esses marcadores têm sido amplamente utilizados em análises de mistura e de equilíbrio de ligação por terem maior poder na detecção de estruturação populacional advinda de mistura, como avaliado por Enoch e colaboradores (2006), Tsai e colaboradores (2005), Bonilla e colaboradores (2004) e Collins-Schramm e colaboradores (2004). Entre os estudos que utilizaram esses marcadores, alguns analisaram populações da América Latina (Wang *et al.*, 2008; Tang *et al.*, 2007; Martinez *et al.*, 2007; González-Andrade *et al.*, 2007 e Marignac *et al.*, 2004).

Objetivos

O objetivo geral desse trabalho foi avaliar a estrutura genética do remanescente de *quilombo* Santo Antônio do Guaporé e da comunidade afro-derivada Santiago do Iguape, utilizando para tanto análises de genética de populações para melhor compreender o processo de formação da população em questão e, com isso, contribuir para um maior entendimento das populações das regiões Norte e Nordeste e, conseqüentemente, do Brasil.

Os objetivos específicos foram:

1. Verificar a existência de polimorfismos para treze *loci* AIM (AT3, LPL, GC, OCA2, RB2300, DRD2*TaqI*, *AluSb19.3*, *AluAPO*, *AluECA*, *AluPV92*, *AluT-PA25*, *AluD1*, *AluFXIIIB*);
2. Estimar as freqüências gênicas e genóticas para os *loci* polimórficos e verificar se a população está em equilíbrio de Hardy-Weinberg para os mesmos;
3. Estimar as contribuições parentais (africana, européia e ameríndia) na formação dessas populações;
4. Comparar essas populações com os prováveis grupos parentais (africanos, ameríndios e europeus) que teriam contribuído para a formação da população brasileira;
5. Comparar essas populações com outras populações afro-derivadas e com populações urbanas brasileiras.

Material e Métodos

Populações analisadas

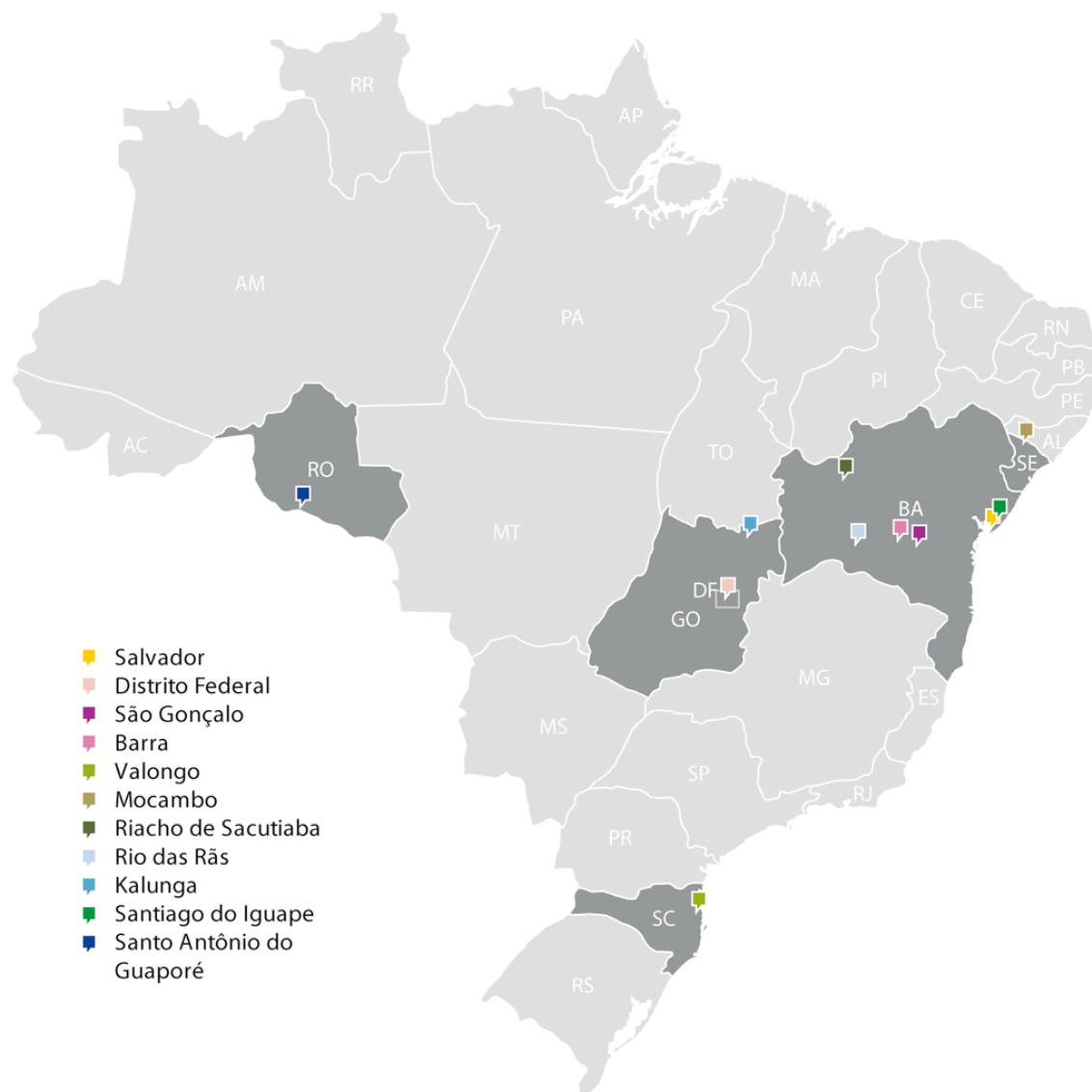


Figura 3. Mapa mostrando a localização das comunidades de Santo Antônio do Guaporé, Santiago do Iguape e das outras comunidades urbanas e quilombolas com as quais foram comparadas.

Santo Antônio do Guaporé

A comunidade remanescente de quilombo Santo Antônio do Guaporé situa-se na margem direita do Rio Guaporé em uma área pertencente ao município de Costa Marques, Rondônia, distante cerca de 140 km de sua área urbana e 800 km de Porto Velho.

Sua localização geográfica tem as seguintes coordenadas: Latitude 12°32'33'' e Longitude 63°31'37'', como determinado por pesquisadores da Universidade Federal de Rondônia – UniR.

No início do século XVIII, a expansão das fronteiras agrícolas e a mineração do ouro trouxeram um afluxo populacional para o Vale do Guaporé e levou a coroa portuguesa a construir o Real Forte Príncipe da Beira (Teixeira, 2004). Por esse motivo, escravos de ancestralidade africana foram levados para a região pela Companhia Geral de Comércio do Grão Pará e do Maranhão pela via fluvial do Rio Madeira e seus afluentes (Teixeira, 2004). Nessa época, vários quilombos foram formados. A convivência entre quilombolas e indígenas incluiu a captura de mulheres da tribo Cabixi (Rocha Pombo, 1935).

Relatos orais dos habitantes de Santo Antônio do Guaporé (Teixeira, 1997) indicam que a comunidade foi formada por duas famílias – Silva e Calazans – e por outros escravos fugidos e/ou abandonados quando as minas da região deixaram de ser produtivas. Os moradores relatam também a grande redução populacional desde a década de 80, quando da criação da Reserva Biológica do Guaporé e conseqüente expulsão das comunidades quilombolas, indígenas e ribeirinhas do local: a população, estimada antes em 400 habitantes, está hoje reduzida a cerca de 100 indivíduos (Teixeira, 1997).

Santiago do Iguape

A vila de Santiago do Iguape situa-se no distrito de mesmo nome no município de Cachoeira, no Recôncavo Baiano. Sua localização geográfica tem as seguintes coordenadas: S 12° 41',05 W 38° 51',35.

No município de Cachoeira, estão as ruínas do mais antigo engenho exportador de açúcar do país (Barickman, 1999; Barickman, 2003), o que mostra que a presença de escravos africanos na região é muito antiga.

A Secretaria de Agricultura, Irrigação e Reforma Agrária – Superintendência de Agricultura Familiar do Governo da Bahia - reconhece a existência de onze remanescentes de quilombo no município (Bahia – Municípios com Registro de Comunidades Quilombolas, 2005), embora apenas um, Paraguaçu, tenha sido oficialmente reconhecido como tal (Fundação Cultural Palmares). Dessa lista, consta a vila de Santiago do Iguape, de marcada ancestralidade africana, mas fundada em território anteriormente o-

cupado por ameríndios Tapuia e Guaranis, estes últimos levados como escravos para o Recôncavo Baiano (Alencastro, 2000).

A vila de Santiago do Iguape foi fundada em 1561 por padres jesuítas na área do Engenho São Domingos da Ponta, um dos mais fortes produtores de açúcar da época (Barickman, 1999). Hoje, seus 3000 habitantes vivem principalmente de agricultura familiar e pequenos comércios.

Distrito Federal

A população do Distrito Federal começou a ser formada efetivamente na década de 60 com a transferência da capital federal do Rio de Janeiro para o Planalto Central. Diferente das demais regiões brasileiras, seu povoamento se deu pela migração rápida de indivíduos provenientes de todas as regiões do país (em especial, sudeste, nordeste e sul), onde a população já era miscigenada.

Segundo dados do Companhia de Planejamento do Distrito Federal (2008), apenas metade da população do DF é proveniente da região Centro-Oeste, sendo a outra metade proveniente das outras quatro regiões brasileiras. Por esses motivos, supõe-se que a população do DF, hoje estimada em 2 milhões de habitantes, represente um quadro geral da população brasileira (Barcelos, 2006).

Dados genotípicos dessa população, utilizados em várias análises estatísticas que buscaram comparar as populações de SAG e STI com uma população brasileira urbana, foram gentilmente cedidos por Neide Maria de Oliveira Godinho. Esses dados são parte integrante de sua tese de doutorado (Godinho, 2008).

Coleta e processamento de material biológico

Aspectos éticos

A coleta e análise das amostras das duas populações em estudo foi aprovada pelos Comitês de Ética em Pesquisa (CEPs) das instituições em que os projetos foram propostos. Todos os doadores preencheram, previamente à coleta, Termos de Consentimento Livre e Esclarecido. A coleta e uso das amostras de Santo Antônio do Guaporé foi aprovada pelo CEP da Universidade Federal de Rondônia (018/CEP/NUSAU). Já a

da comunidade de Santiago do Iguape, foi aprovada pelo CEP da Faculdade de Ciências da Saúde da UnB e pela CONEP (CAAE 0021.0.000.012-04).

Santo Antônio do Guaporé

No ano de 2006, uma expedição composta por pesquisadores da Universidade Federal de Rondônia - UniR visitou a comunidade de Santo Antônio do Guaporé. Lá, 84 voluntários previamente informados dos objetivos da pesquisa e que assinaram um termo de consentimento livre e esclarecido tiveram amostras de sangue venoso coletadas utilizando o sistema de coleta a vácuo com tubos contendo EDTA como anticoagulante.

O material foi processado (separado nas frações plasma, leucócitos e células vermelhas) e estocado sob congelamento em microtubos devidamente etiquetados de forma codificada para manter a privacidade do doador no laboratório do Centro Interdepartamental de Biologia Experimental e Biotecnologia – CIBEBI – na UniR, localizada em Porto Velho. Posteriormente, parte desse material foi levado para a UnB, onde foi estocado em freezer a -20°C. Por problemas no processamento e estocagem das amostras, o número de amostras disponíveis para análise foi reduzido a 31.

Santiago do Iguape

As amostras de Santiago do Iguape foram coletadas no ano de 2006 por uma equipe de pesquisadores da Faculdade Adventista da Bahia em programas comunitários de saúde realizados em parceria com a Quarta Divisão Regional de Saúde do Estado da Bahia. Na época, as amostras foram coletadas para um trabalho que se iniciava objetivando levantar a ocorrência, no Recôncavo Baiano, de haplótipos da β -globina relacionados à anemia falciforme (Silva, 2007). As amostras foram coletadas de 37 indivíduos não aparentados que habitam a vila de Santiago do Iguape, no Município de Cachoeira. Todos os voluntários foram informados dos objetivos da pesquisa e assinaram um termo de consentimento livre e esclarecido previamente à coleta.

Os doadores tiveram 5mL de sangue retirado por via parenteral em tubos contendo EDTA como anticoagulante. O sangue foi processado e estocado no Laboratório de Genética da Faculdade Adventista da Bahia em microtubos etiquetados de forma codificada em freezer a -20°C.

Extração de DNA

50 μ L da fração leucocitária das amostras de SAG foram utilizados para a extração de DNA no Laboratório de Genética da UnB utilizando o kit comercial *blood genomicPrep Mini Spin Kit* da *GE Healthcare*, de acordo com o protocolo do fabricante. O DNA extraído foi quantificado no equipamento *DNAQuant* da *Pharmacia* no Laboratório de Biologia Molecular da UnB e armazenado em microtubos devidamente etiquetados com códigos de forma a manter a privacidade sem, entretanto, dissociar as amostras dos dados de identificação individual dos doadores e mantidos em freezer a -80°C no Laboratório de Genética da Universidade de Brasília em um banco de amostras. Aliquotas para uso foram diluídas a 20 η g/ μ L e armazenadas também em microtubos devidamente identificados em freezer a -20°C no mesmo laboratório.

100 μ L de sangue total das amostras coletadas em STI foram utilizados para a extração de DNA genômico pelo *GFXTM Genomic Blood DNA Purification Kit* da *Amer sham Pharmacia Biotech*. A extração foi realizada pela equipe do Laboratório de Genética da Faculdade Adventista da Bahia de acordo com o protocolo do fabricante e uma fração, levada para o Laboratório de Genética da UnB. O DNA recebido foi quantificado no equipamento *Gene Quant* da *Pharmacia* no Laboratório de Microbiologia da UnB, diluído a 20 η g/ μ L e armazenado em microtubos codificados em freezer a -20°C no Laboratório de Genética da UnB.

Marcadores moleculares selecionados

Foram selecionados 13 marcadores que se enquadram na definição de AIM, de acordo com seus δ (tabela 3) entre as populações parentais utilizadas nesse estudo (anexo 1), todos SNPs ou Indels autossômicos. A seguir, uma breve descrição dos marcadores.

TPA25: o *locus* para a inserção *Alu* TPA-25 situa-se no cromossomo 8, no gene do ativador de plasminogênio tissular (Tishkoff *et al.*, 1996). A inserção tem cerca de 290pb e a amplificação por PCR da região onde se insere produz um fragmento de 400pb (presença da inserção) e outro de 110pb (ausência da inserção) (Novick *et al.*, 1995).

APO: o *locus* para a inserção Alu APO se situa no braço longo do cromossomo 11 (Batzer *et al.*, 1994; Karathanasis, 1985), na região não codificadora próxima ao complexo de genes da apolipoproteína AI-CIII-AIV (Karathanasis, 1985). Sua amplificação por PCR produz um fragmento de 400pb (presença da inserção) e outro de 110pb (ausência da inserção) (Novick *et al.*, 1995).

Tabela 3. Diferenciais de frequência médios - δ - encontrados entre as frequências do alelo *1 nas populações parentais. Para o marcador DRD2, o δ foi calculado para cada haplótipo e para o marcador GC, para os alelos *1F e *1S.

Marcador	África	África	Europa
	x Europa	x Ameríndio	x Ameríndio
TPA25	23%	30%	7%
APO	50%	53%	4%
PV92	0%	66%	66%
ECA	9%	47%	37%
FXIIIB	33%	84%	51%
D1	39%	54%	15%
Sb19.3	39%	21%	19%
AT3	59%	80%	22%
DRD2TaqI *1	1%	1%	0%
DRD2TaqI *2	0%	0%	1%
DRD2TaqI *3	2%	37%	35%
DRD2TaqI *4	1%	2%	3%
DRD2TaqI *5	6%	4%	2%
DRD2TaqI *6	19%	18%	1%
DRD2TaqI *7	31%	15%	16%
DRD2TaqI *8	55%	3%	58%
LPL	47%	51%	3%
RB2300	57%	75%	18%
OCA2	62%	32%	30%
GC *1S	68%	49%	19%
GC *1F	49%	44%	6%

PV92: o *locus* para essa inserção Alu encontra-se no cromossomo 16 (Batzer *et al.*, 1994). Sua amplificação por PCR produz um fragmento de 400pb (presença da inserção) e um de 110pb (ausência da inserção) (Novick *et al.*, 1995).

ECA: o *locus* para a inserção AluACE se situa no braço longo do cromossomo 17, em uma região intrônica do gene da enzima conversora da angiotensina (Nakai *et*

al., 1994). Sua amplificação por PCR produz um fragmento de 490pb (presença da inserção) e outro de 190pb (ausência da inserção) (Nakai *et al.*, 1994).

FXIIIB: o *locus* para a inserção *Alu*FXIIIB se situa em um íntron do gene da subunidade B do fator XIII, uma glicoproteína formada por duas subunidades (A e B) envolvida na coagulação sanguínea (Kass *et al.*, 1994). O gene da subunidade B se localiza no cromossomo 1 (Map Viewer).

D1: o *locus* para a inserção *Alu*D1 está situado no cromossomo 3 (Batzer *et al.*, 1995; Arcot *et al.*, 1995). A amplificação por PCR produz dois fragmentos: 400 pb para a ausência e 700 pb para a presença da inserção.

Sb19.3: esse marcador é uma inserção *Alu* pertencente à subfamília Yb8 situada no braço curto do cromossomo 19 (19p12), como descrito por Arcot e colaboradores (1995). A presença da inserção caracteriza o alelo *Sb19.3*1* e sua ausência, o alelo *Sb19.3*2*. A amplificação do fragmento de interesse gera um produto de 450pb, caso a inserção esteja presente, ou um produto de 150pb, caso ela esteja ausente.

AT3-I/D: o polimorfismo do *locus* AT3-I/D aqui analisado se caracteriza pela presença ou ausência de uma inserção de 76pb no éxon 1 do gene da antitrombina III, 1q25.1 (Liu *et al.* 1995). O alelo AT3*1 se caracteriza pela presença da inserção, e o alelo AT3*2, pela ausência da inserção, o que conseqüentemente resulta na presença de duas bandas de pesos diferentes visualizadas em eletroforese: a mais pesada corresponde ao alelo AT3*1 (572pb), e a mais leve, ao alelo AT3*2 (Luizon 2003).

DRD2*TaqI* A, B e D: os três polimorfismos analisados no gene 2 do receptor de dopamina caracterizam-se por mutações de ponto que geram sítios de restrição para a enzima *TaqI* (Gelernter *et al.*, 1998). Cada um dos três loci é analisado individualmente utilizando pares de iniciadores específicos. A digestão do produto de PCR de cada *locus* gera dois fragmentos de tamanhos diferentes que caracterizam o alelo *2 e a não digestão, o alelo *1. Por sua proximidade física, as três mutações são analisadas como haplótipo construído por um algoritmo com base nos genótipos *multilocus* de cada indivíduo.

LPL: os iniciadores utilizados para a análise do *locus* LPL flanqueiam uma região de 319pb no íntron 6 da Lipoproteína Lipase (LPL). Sua localização citogenética é 8p22 (Sparkes *et al.*, 1987). O alelo LPL*1 é caracterizado pela ausência de um sítio de restrição para a enzima *PvuII*, e o alelo LPL*2, pela presença desse sítio (Gotoda *et al.*, 1992).

RB2300: os iniciadores utilizados na análise desse *locus* flanqueiam uma região de 180pb no íntron 1 do gene do Retinoblastoma localizado no cromossomo 13 (Bookstein *et al.*, 1990). O alelo RB*1 se caracteriza pela ausência do sítio de restrição para a enzima *Bam*HI, e o alelo RB*2, pela presença de um sítio.

OCA2: o polimorfismo do *locus* OCA2 se caracteriza por uma mutação do tipo SNP (single nucleotide polymorphism) no gene P humano, 15q11-q13, (Ramsay *et al.*, 1992). Esse gene é formado por 25 éxons e diversos polimorfismos já foram descritos, sendo que cinco deles causam albinismo oculocutâneo tirosinase-positivo (Lee *et al.*, 1995). O SNP aqui analisado é uma mutação neutra que cria um sítio de restrição para a enzima *Hae*III e o produto de PCR, após digestão por essa enzima, resulta em dois fragmentos de tamanhos diferentes. O alelo OCA2*1 é caracterizado pela ausência de um sítio de restrição para a enzima *Hae*III, gerando um fragmento mais pesado, e o alelo OCA2*2, pela presença desse sítio, gerando duas bandas menores.

GC: esse marcador se caracteriza pela presença de: a) uma mutação de ponto que cria um sítio de restrição para a enzima *Hae*III (alelo 1S) e b) uma mutação de ponto que cria um sítio de restrição para a enzima *Sty*I (alelo 1F), ambas no gene do componente de grupo-específico humano. A ausência dessas mutações caracteriza o alelo 2 (Kofler *et al.*, 1995; Parra *et al.*, 1998).

Análise dos marcadores genéticos moleculares

Para a análise dos marcadores genéticos, os fragmentos de interesse nesse estudo foram amplificados por PCR (tabelas 4 e 5) em um termociclador *MJ Research PTC-100*® e visualizados em gel vertical de poliacrilamida não desnaturante com glicerol a 6% ou 10% corado com nitrato de prata. Previamente à eletroforese, os marcadores do tipo SNP foram submetidos a digestão por enzimas de restrição específicas (tabela 6). Já o produto de PCR dos marcadores do tipo indel foi submetido diretamente a eletroforese.

Todos os marcadores foram analisados em gel de poliacrilamida não desnaturante com glicerol a 6%, com exceção dos loci DRD2A e DRD2D, analisados em gel a 10%. O tempo de corrida foi de 2h, com exceção dos loci LPL (2h30) e DRD2A, DRD2B e DRD2D (3h) e a voltagem, 200V, com exceção dos loci DRD2A, DRD2B e DRD2D (150V).

Tabela 4. Iniciadores, temperatura padronizada para pareamento dos iniciadores (T) e enzimas de restrição utilizadas para a análise dos AIMS selecionados.

<i>Locus</i>	Iniciadores (5'/3')	Referência	T
TPA25	GTAAGAGTTCCGTAACAGGACAGCT CCCCACCCTAGGAGAACTTCTCTTT	Novick <i>et al.</i> , 1995	58°C
APO	AAGTGCTGTAGGCCATTTAGATTAG AGTCTTCGATGACAGAGTATACAGA	Batzer <i>et al.</i> , 1996	50°C
ECA	CTGCAGACCACTCCCATCCTTTCT GATGTGGCCATCACATTCGTCAGAT	Nakai <i>et al.</i> , 1994	58°C
PV92	AACTGGGAAAATTTGAAGAGAAAGT TGAGTTCTCAACTCCTGTGTGTTAG	Novick <i>et al.</i> , 1995	54 °C
FXIII B	AAGTGCTGTAGGCCATTTAGATTAG AGTCTTCGATGACAGCGTATAACAGA	Kass <i>et al.</i> , 1994	54°C
D1	TGCTGATGCCAGGGTTAGTAAA TTTCTGCTATGCTCTTCCCTCTC	Batzer <i>et al.</i> , 1995	56°C
Sb 19.3	TCTAGCCCCAGATTTATGGTAACTG AAGCACAATTGGTTATTTTCTGAC	Parra <i>et al.</i> , 1998	58 °C
AT3-I/D	CCACAGGTGTAACATTGTGT GAGATAGTGTGATCTGAGGC	Liu <i>et al.</i> , 1995	55 °C
LPL	AGGCTTCACTCATCCGTGCCTCC TTATGCTGCTTTAGACTCTTGTC	Gotoda <i>et al.</i> , 1992	60 °C
OCA2	CTTTCGTGTGTGCTAACTCC ACCTCTAGCATGGTTCTTGGGC	Lee <i>et al.</i> , 1995	60 °C
RB 2300	CAGGACAGCGGCCCGGAG CTGCAGACGCTCCGCCGT	Bookstein <i>et al.</i> , 1990	60°C
DRD2A	CCG TCG ACG GCT GGC CAA GTT GTC TA CCG TCG ACC CTT CCT GAG TGT CAT CA	Gelernter <i>et al.</i> , 1998	60°C
DRD2B	GAT GTG TAG GAA TTA GCC AGG GAT ACC CAC TTC AGG AAG TC	Gelernter <i>et al.</i> , 1998	64 °C
DRD2D	CCCAGCAGGGAGAGGGAGTA GACAAGTACTTGTAAGCATG	Gelernter <i>et al.</i> , 1998	64 °C
GC	AGATCTGAAATGGCTATTATTTTGC GGAAGGTGAGTTTATGGAACAGC	Parra <i>et al.</i> , 1998	55 °C

Tabela 5. Misturas de reagentes utilizadas nas reações de PCR para cada sistema.

Reagentes	TPA25	APO	PV92	ECA	FXIIIB	D1	Sb19.3	AT3
H ₂ O ^a (μL)	6,20	5,80	7,40	6,60	6,60	6,40	6,45	7,20
Tampão ^b (μL)	1,50	1,25	2,00	1,25	1,25	1,25	1,40	1,30
dNTP (μL) ^c	1,25	1,25	1,25	1,25	1,25	1,00	1,25	1,25
iniciadores ^d (μL)	1,25	2,00	0,50	1,00	1,25	1,25	1,50	1,25
MgCl ₂ ^e (μL)	-	0,25	-	0,50	0,15	0,50	-	-
DMSO ^f (μL)	0,32	-	0,32	0,32	-	0,32	-	-
Taq ^g (U)	1	0,80	1,00	3,00	1,00	0,50	2,50	2,50
DNA (ng)	20	20	20	30	20	10	30	20
Reagentes	LPL	DRD2A	DRD2B	DRD2D	RB2300	OCA2	GC	
H ₂ O ^a (μL)	7,38	6,30	6,30	6,00	5,95	7,65	6,90	
Tampão ^b (μL)	1,25	1,30	1,30	1,30	1,75	1,13	1,25	
dNTP (μL) ^c	0,63	1,60	1,25	1,25	1,25	0,63	1,25	
iniciadores ^d (μL)	1,50	1,25	1,60	1,60	1,25	1,50	1,50	
MgCl ₂ ^e (μL)	-	-	-	-	0,80	-	-	
DMSO ^f (μL)	-	-	-	-	-	-	-	
Taq ^g (U)	0,50	2,50	2,50	2,50	2,50	0,50	0,60	
DNA (ng)	20	24	30	24	20	30	20	

a: H₂O deionizada autoclavada; b: tampão com MgCl₂ 2mM fornecido pelo fabricante da Taq DNA Polimerase; c: solução de dATP, dTTP, dCTP e dGTP a 5Mm; d: solução dos iniciadores *forward* e *reverse* a 5μM; e: MgCl₂ a 2Mm; f: DMSO a 10%; g: Taq DNA Polimerase.

Tabela 6. Misturas de reagentes, temperatura e tempo utilizados nas reações de restrição para cada sistema.

Reagentes	DRD2A	DRD2B	DRD2D	LPL	OCA2	RB2300	GC	
H ₂ O ^a (μL)	5,2	5,2	5,2	5,7	5,5	5,7	5,5	5,6
Tampão ^b (μL)	1,2	1,2	1,2	1,0	1,2	1,0	1,2	1,0
BSA ^c (μL)	0,1	0,1	0,1	-	-	-	-	0,1
enzima: 3U	TaqI	TaqI	TaqI	PvuII	HaeIII	BamHI	HaeIII	StyI
produto PCR (μL)	5,0	5,0	5,0	5,0	5,0	5,0	5,0	5,0
Temperatura °C	65	65	65	37	37	37	37	37
Tempo (h)	16	16	16	4	4	16	4	4

a: H₂O deionizada autoclavada; b: tampão fornecido pelo fabricante da enzima de restrição; c: BSA (*bovine serum albumine*) fornecida pelo fabricante da enzima de restrição.

Análise estatística

As análises estatísticas tiveram a finalidade principal de descrever as populações quanto a aderência ao Equilíbrio de Hardy-Weinberg, estruturação genética, diversidade interpopulacional e mistura genética em sua formação. Os dados obtidos foram também comparados aos gerados previamente para outras populações afro-derivadas brasileiras, para a população do Distrito Federal (DF), considerada aqui como população brasileira urbana, e para as populações parentais. Sete populações remanescentes de quilombo (Kalunga, Rio das Rãs, Riacho de Sacutiaba, Mocambo, São Gonçalo, Barra e Valongo) foram incluídas nas análises de coordenada principal.

Análises descritivas

Frequências gênicas e genotípicas

As frequências gênicas e genotípicas foram obtidas por contagem direta utilizando o programa GENEPOP 3.4 (Raymond e Rousset, 1995), com exceção dos haplótipos obtidos com a análise da presença ou ausência de três sítios de restrição situados no gene DRD2 (DRD2A, B e D).

Tabela 7. Haplótipos possíveis considerando três sítios de restrição, *loci* DRD2A, DRD2B e DRD2D, presentes no gene DRD2.

Haplótipo	<i>Locus</i>		
	DRD2A	DRD2B	DRD2D
1	*1	*1	*1
2	*1	*1	*2
3	*1	*2	*2
4	*1	*2	*1
5	*2	*2	*2
6	*2	*2	*1
7	*2	*1	*1
8	*2	*1	*2

A definição das frequências haplotípicas das mutações do DRD2 foi realizada utilizando o programa Arlequin 3.1.1 (Excoffier *et al.*, 2005). O algoritmo empregado foi o EM-*zipper*, extensão do algoritmo EM (Excoffier *et al.*, 2005). Considerando as três mutações possíveis, foram definidos oito haplótipos DRD2 que estão listados na tabela 7.

Aderência ao Equilíbrio de Hardy-Weinberg

A distribuição dos alelos de um dado *locus* segue, obedecidas a certas premissas, a expansão de polinômios (Weir, 1996), de acordo com o número de alelos desse *locus*. A essa situação chamamos Equilíbrio de Hardy-Weinberg (EHW) e espera-se que os alelos sejam assim distribuídos em situações normais.

Estabelecendo como H_0 a situação de equilíbrio (distribuição aleatória dos gametas), para verificação da concordância com essa situação, foi utilizado o teste descrito por Guo e Thompson (1992) executado pelo programa GENEPOP 3.4 (Raymond e Rousset, 1995) utilizando o método da cadeia de Markov.

Foram realizados testes para a verificação de excesso e deficiência de heterozigotos, também pelo programa GENEPOP 3.4 (Raymond e Rousset, 1995). Esses testes verificam a hipótese alternativa H_A de excesso ou deficiência de heterozigotos em comparação com o esperado pelo EHW, também utilizando o método da cadeia de Markov.

Desequilíbrio de ligação

O desequilíbrio de ligação, situação em que a herança de um *locus* não é independente da de outro *locus*, pode ser gerado por inúmeros fatores, entre eles a proximidade física dos *loci*, miscigenação e estruturação populacional. Para verificar a ocorrência de desequilíbrio de ligação (DL) entre os *loci* aqui analisados, foi utilizado o programa GENEPOP 3.4 (Raymond e Rousset, 1995). Estabelecendo como H_0 a ausência de DL, o programa realiza o teste exato de Fisher para tabelas de contingência construídas para todos os pares de *loci* utilizando a cadeia de Markov.

Com a realização de múltiplos testes, no caso um para cada par de *loci* sendo testado, sob a mesma H_0 utilizando os mesmos *loci*, aumentam as chances de se cometer um erro do tipo I (Ha, 2008) e, com isso, um aumento no α real em aproximadamente o

número de testes independentes realizados. O nível de significância foi estabelecido em 0,05, mas foram incluídos também os resultados da correção de Bonferroni, que consiste em dividir o α pretendido pelo número de testes independentes realizados (Weir e Cockerham, 1996). Para a população de SAG, o novo α (α_{BONF}) foi estabelecido em $0,05/78 = 0,0006$ e para STI, em $0,05/55 = 0,0009$.

Comparação das frequências gênicas e genotípicas

As frequências gênicas obtidas para SAG e STI foram comparadas entre si e entre cada uma delas e DF pelo teste exato de Fisher, realizado pelo programa GENEPOP 3.4 (Raymond e Rousset, 1995). As duas populações foram também comparadas quanto às frequências genotípicas pelo teste descrito por Goudet e colaboradores (1996). A hipótese nula considerada nos dois casos é de que a distribuição alélica/genotípica é idêntica nos pares de populações analisados. O mesmo teste foi realizado para os haplótipos DRD2TaqI entre SAG e STI.

Estatísticas F

As estatísticas F foram calculadas usando o programa *Genetic Data Analysis – GDA* (Lewis e Zaykin, 1997) de acordo com o descrito por Weir (1984).

Os parâmetros, definidos como índices de fixação baseados na heterozigose média observada e esperada em cada amostra e na amostra como um todo, foram 1) Fis (coeficiente de endogamia em uma população); 2) Fit (coeficiente global de endogamia, mede a existência de endogamia devida a sub-estruturação na amostra global) e 3) Fst (mede a diferenciação das amostras) (Weir, 1996).

A análise desses parâmetros levou em consideração os dez marcadores tipados para as três populações avaliadas: GC, APO, AT3, D1, LPL, OCA2, PV92, RB2300, Sb19.3 e TPA25.

Definição dos intervalos de confiança

Para determinar intervalos de confiança (IC) para essas estimativas e, assim, poder extrapolar os resultados obtidos a partir das amostras para as populações (Weir,

1996), foi utilizado o *Bootstrapping* como metodologia de re-amostragem. Essa metodologia simula a obtenção de diferentes amostras da população amostrando vários conjuntos a partir da amostra original e analisando separadamente cada conjunto.

Dessa forma, são definidos ICs em que X% das estimativas estejam inseridas, ou seja, é determinada a probabilidade de os ICs obtidos incluírem a estimativa obtida, ou seja, de a amostra representar a medida real para a população (Ha, 2008). Assim, um IC de 95% representa a faixa de valores encontrados para 95% dos ICs obtidos. No presente estudo, foram analisados ICs de 95%, mais estreitos, e de 99%, mais amplos, para as estatísticas F obtidos por *Bootstrapping* com 1000 replicações.

Estimativa de mistura

As estimativas de contribuição parental na formação de SAG e STI foram estimadas pelo método de identidade gênica (Chakraborty, 1985) utilizando o programa ADMIX3 e considerando o número total de marcadores analisados para cada população. Assumiu-se um modelo tri-híbrido coerente com a história de formação da população brasileira, de acordo com o qual foram escolhidos ameríndios (brasileiros e Quechua do Peru), africanos (subsaarianos) e europeus (da península ibérica e italianos) como parentais. As frequências parentais utilizadas (tabela 8) foram as médias ponderadas das frequências disponíveis na literatura para populações pertencentes a cada um desses grupos.

Uma segunda análise, visando a comparação entre as populações de SAG, STI e DF, foi realizada utilizando as mesmas parentais, mas excluindo os dados dos marcadores ECA e FXIIB e do haplótipo DRD2*TaqI* para igualar os bancos de dados.

Componente principal

A análise de componente principal (PCA) e outras dela derivadas, como a análise de coordenada principal, são formas de simplificação de dados multivariados (Cavalli-Sforza et al., 1996). Aqui, matrizes de distância foram convertidas em gráficos de coordenada principal com o intuito de simplificar a interpretação das distâncias obtidas entre todas as populações.

As frequências alélicas e haplotípicas – no caso do DRD2*TaqI* – encontradas para as populações em análise, para DF, para as populações parentais e para populações remanescentes de quilombo (encontradas na literatura) foram utilizadas para gerar matrizes de distância. Foi escolhida a distância de Nei, que mostra menores desvios dos valores reais em experimentos de simulação para marcadores desse tipo (Batillana *et al.*, 2002). As matrizes foram construídas pelo programa GENDIST do pacote PHYLIP (Felsenstein, 1989) e utilizadas para a construção de gráficos de coordenada principal pelo programa NTSYS (Rohlf, 1995).

Tabela 8. Médias ponderadas das frequências alélicas dos marcadores AIMs encontradas na literatura para as populações parentais utilizadas na análise de mistura gênica.

Marcador	África	Europa	Ameríndio
TPA25 *1	0,357	0,586	0,657
APO *1	0,455	0,951	0,989
PV92 *1	0,183	0,182	0,845
ECA *1	0,321	0,411	0,786
FXIII B *1	0,082	0,417	0,923
D1 *1	0,000	0,389	0,544
Sb19.3 *1	0,458	0,853	0,664
AT3 *1	0,854	0,268	0,050
hDRD2 *1	0,002	0,007	0,011
hDRD2 *2	0,004	0,000	0,006
hDRD2 *3	0,135	0,153	0,503
hDRD2 *4	0,018	0,011	0,042
hDRD2 *5	0,079	0,015	0,038
hDRD2 *6	0,200	0,013	0,019
hDRD2 *7	0,497	0,183	0,344
hDRD2 *8	0,065	0,617	0,037
LPL *1	0,970	0,497	0,463
RB2300 *1	0,894	0,325	0,145
OCA2 *1	0,117	0,737	0,434
GC *1S	0,825	0,152	0,339
GC *1F	0,089	0,598	0,542

Resultados

Descrição das populações

Santo Antônio do Guaporé

A amostra de Santo Antônio do Guaporé foi analisada para os 13 marcadores selecionados e todos os marcadores apresentaram-se polimórficos. As frequências gênicas e genóticas observadas, número de indivíduos analisados (n), heterozigose observada e esperada e valores de P obtidos pelo teste para a distribuição dos alelos de acordo com o esperado pelo Equilíbrio de Hardy-Weinberg encontram-se listados nas tabelas 9, 10 e 11.

Tabela 9. Distribuição alélica dos 15 loci AIMS em Santo Antônio do Guaporé ($n = 31$): frequências alélicas e genóticas e heterozigoses esperada (H_e) e observada (H_o).

<i>Locus</i>	Alelo			Genótipo					
	1	2	3	11	12	22	13	23	33
GC	0,226	0,516	0,258	0,097	0,194	0,258	0,065	0,323	0,065
APO	0,887	0,113	..	0,774	0,226	0,000
AT3	0,500	0,500	..	0,258	0,484	0,258
D1	0,387	0,613	..	0,129	0,516	0,355
ECA	0,500	0,500	..	0,290	0,419	0,290
FXIIB	0,468	0,532	..	0,290	0,355	0,355
LPL	0,613	0,387	..	0,355	0,516	0,129
OCA2	0,468	0,532	..	0,194	0,548	0,258
PV92	0,500	0,500	..	0,226	0,548	0,226
RB2300	0,468	0,532	..	0,226	0,484	0,290
Sb19.3	0,500	0,500	..	0,194	0,613	0,194
DRD2TaqI
DRD2A	0,403	0,597	..	0,129	0,548	0,323
DRD2B	0,242	0,758	..	0,032	0,419	0,548
DRD2D	0,210	0,790	..	0,032	0,355	0,613
TPA25	0,371	0,629	..	0,161	0,419	0,419
H_o	0,469								
H_e	0,468								

GC*1: GC*1S; GC*2: GC*1F; GC*3: GC*2.

Como os *loci* DRD2A, DRD2B e DRD2D são fisicamente ligados, estes foram analisados em conjunto. As frequências apresentadas dos haplótipos DRD2TaqI (tabela

10) referem-se à melhor estimativa obtida para a população. Foram observados nessa população sete dos oito haplótipos possíveis, sendo que o haplótipo modal foi o 7 (DRD2A*2, DRD2B*2, DRD2*2).

Tabela 10. Frequências dos haplótipos DRD2*TaqI* observadas em Santo Antônio do Guaporé.

haplótipo	descrição	SAG
1	111	0,000
2	112	0,206
3	121	0,098
4	122	0,099
5	211	0,011
6	221	0,101
7	222	0,460
8	212	0,025

Os genótipos de todos os marcadores encontraram-se distribuídos de acordo com o esperado pelo EHW ($P > 0,05$). Os resultados apresentados na tabela 11 mostraram também não haver excesso ou deficiência de heterozigotos nessa população ($P > 0,05$).

Tabela 11. Santo Antônio do Guaporé: valores de P com respectivos erros padrão (EP) para os teste de aderência ao Equilíbrio de Hardy-Weinberg (EHW), teste para deficiência de heterozigotos (DH) e teste para excesso de heterozigotos (EH).

<i>Locus</i>	EHW		DH		EH	
	P	EP	P	EP	P	EP
GC	0,438	0,002	0,202	0,002	0,811	0,002
APO	1,000	0,000	1,000	0,000	0,684	0,001
AT3	1,000	0,000	0,536	0,002	0,739	0,002
D1	1,000	0,000	0,781	0,001	0,497	0,002
ECA	0,472	0,001	0,265	0,002	0,914	0,001
FXIIB	0,145	0,001	0,090	0,001	0,982	0,000
hDRD2	0,787	0,006	0,564	0,007	0,444	0,008
LPL	1,000	0,000	0,782	0,001	0,494	0,002
OCA2	0,725	0,001	0,801	0,001	0,455	0,002
PV92	0,728	0,001	0,791	0,002	0,464	0,002
RB2300	1,000	0,000	0,546	0,002	0,731	0,002
Sb19.3	0,295	0,002	0,939	0,001	0,207	0,001
TPA25	0,700	0,001	0,390	0,002	0,854	0,002

Tabela 12. Santo Antônio do Guaporé: teste de desequilíbrio de ligação *locus a locus*. Em destaque, valores que indicam existência de desequilíbrio de ligação. EP: erro padrão.

<i>Locus x Locus</i>	P	EP	<i>Locus x Locus</i>	P	EP		
APO	0,633	0,001	PV92	0,733	0,001		
PV92	0,072	0,001	ECA	0,503	0,001		
ECA	1,000	0,000	FXIIIB	0,165	0,002		
FXIIIB	0,480	0,002	D1	0,848	0,001		
D1	1,000	0,000	Sb19.3	0,147	0,001		
TPA25	Sb19.3	0,461	0,002	APO	AT3	0,064	0,001
AT3	0,082	0,001	DRD2	0,013	0,001		
DRD2	0,466	0,005	LPL	0,709	0,001		
LPL	0,221	0,002	RB2300	1,000	0,000		
RB2300	0,437	0,002	OCA2	1,000	0,000		
OCA2	0,061	0,001	GC	0,935	0,001		
GC	0,245	0,003					
ECA	0,513	0,002	FXIIIB	0,680	0,002		
FXIIIB	0,266	0,002	D1	0,656	0,002		
D1	0,711	0,002	Sb19.3	0,170	0,002		
Sb19,3	0,807	0,001	AT3	0,616	0,002		
PV92	AT3	0,127	0,001	ECA	DRD2	0,666	0,005
DRD2	0,721	0,005	LPL	0,597	0,002		
LPL	0,191	0,002	RB2300	0,571	0,002		
RB2300	0,899	0,001	OCA2	0,982	0,000		
OCA2	1,000	0,000	GC	0,379	0,004		
GC	0,380	0,004					
D1	0,068	0,001	Sb19,3	0,242	0,002		
Sb19.3	0,205	0,002	AT3	0,808	0,001		
AT3	0,001	0,000	DRD2	0,602	0,005		
FXIIIB	DRD2	0,794	0,005	D1	LPL	0,368	0,002
LPL	0,580	0,002	RB2300	0,497	0,002		
RB2300	0,930	0,001	OCA2	0,405	0,002		
OCA2	0,005	0,000	GC	0,048	0,002		
GC	0,764	0,003					
AT3	0,025	0,001	DRD2	0,633	0,005		
DRD2	0,011	0,001	LPL	0,974	0,000		
Sb19.3	LPL	0,875	0,001	AT3	RB2300	1,000	0,000
RB2300	0,806	0,001	OCA2	0,328	0,002		
OCA2	1,000	0,000	GC	0,908	0,002		
GC	0,211	0,003					
LPL	0,009	0,001	RB2300	0,104	0,002		
DRD2	RB2300	0,333	0,006	LPL	OCA2	0,892	0,001
OCA2	0,817	0,004	GC	0,155	0,003		
GC	0,181	0,007					
OCA2	0,725	0,002	OCA2	GC	0,087	0,002	
GC	0,664	0,003					

O teste de desequilíbrio de ligação mostrou que alguns pares de *loci* apresentaram desvios quanto ao esperado, no caso de segregação independente, como apresentado na tabela 12. Os pares dos *loci* FXIII^B/AT3, D1/GC, Sb19.3/AT3, DRD2*TaqI*/APO e DRD2*TaqI*/LPL apresentaram desequilíbrio de ligação ($P < 0,05$). Ao considerar $\alpha_{\text{BONF}} = 0,0006$, entretanto, nenhuma associação significativa foi encontrada.

Santiago do Iguape

Foram analisados para a população de Santiago do Iguape 11 marcadores do tipo AIM: TPA25, APO, PV92, D1, Sb19.3, AT3, DRD2*TaqI*, LPL, RB2300, OCA2 e GC e todos apresentaram polimorfismo na população. As frequências gênicas e genotípicas, n , heterozigose observada e esperada e valores de P obtidos pelo teste para a distribuição dos alelos de acordo com o esperado pelo Equilíbrio de Hardy-Weinberg encontram-se listados nas tabelas 13, 14 e 15.

Tabela 13. Distribuição alélica dos 11 loci AIMS em Santiago do Iguape (n total = 37): frequências alélicas e genotípicas e heterozigoses esperada (H_e) e observada (H_o).

<i>Loci</i>	Alelo			Genótipo					
	1	2	3	11	12	22	13	23	33
GC	0,270	0,635	0,095	0,000	0,405	0,405	0,135	0,054	0,000
APO	0,765	0,235	..	0,559	0,412	0,029
AT3	0,742	0,258	..	0,545	0,394	0,061
D1	0,236	0,764	..	0,000	0,472	0,528
LPL	0,803	0,197	..	0,667	0,273	0,061
OCA2	0,210	0,790	..	0,032	0,355	0,613
PV92	0,568	0,432	..	0,432	0,270	0,297
RB2300	0,514	0,486	..	0,222	0,583	0,195
Sb19.3	0,581	0,419	..	0,378	0,405	0,216
TPA25	0,378	0,622	..	0,243	0,270	0,487
DRD2 <i>TaqI</i>
DRD2A	0,319	0,681	..	0,083	0,472	0,444
DRD2B	0,167	0,833	..	0,028	0,278	0,694
DRD2D	0,303	0,697	..	0,152	0,303	0,546
<i>n</i> médio	35,08								
H_o	0,391								
H_e	0,417								

GC*1: GC*1S; GC*2: GC*1F; GC*3: GC*2.

As frequências dos sete haplótipos DRD2*TaqI* encontrados encontram-se listadas na tabela 14. O haplótipo modal nessa população também foi o 7 (DRD2A*2, DRD2B*2, DRD2D*2).

Tabela 14. Frequências dos haplótipos DRD2*TaqI* observadas em Santiago do Iguape.

haplótipo	descrição	STI
1	111	0,020
2	112	0,100
3	121	0,074
4	122	0,108
5	211	0,016
6	221	0,193
7	222	0,488
8	212	0,000

Com exceção dos *loci* PV92, TPA25 e GC, todos os marcadores apresentaram distribuição genotípica de acordo com o esperado pelo EHW (tabela 15), considerando $\alpha = 0,05$. Entretanto, considerando $\alpha = 0,01$, apenas o *locus* PV92 se manteve fora de equilíbrio.

Tabela 15. Santiago do Iguape: valores de P com respectivos erros padrão (EP). EHW - teste de aderência ao Equilíbrio de Hardy-Weinberg; DH - teste para deficiência de heterozigotos; EH - teste para excesso de heterozigotos.

Marcador	EHW		DH		EH	
	P	EP	P	EP	P	EP
GC	0,017	0,001	0,958	0,001	0,041	0,001
APO	0,649	0,002	0,901	0,001	0,407	0,001
AT3	1,000	0,000	0,696	0,001	0,658	0,001
D1	0,157	0,002	1,000	0,000	0,084	0,001
DRD2A	1,000	0,000	0,794	0,001	0,490	0,002
DRD2B	1,000	0,000	0,667	0,001	0,774	0,001
DRD2D	0,112	0,002	0,098	0,001	0,983	0,000
LPL	0,576	0,002	0,343	0,001	0,922	0,001
OCA2	1,000	0,000	0,797	0,001	0,621	0,001
PV92	0,008	0,001	0,006	0,000	0,999	0,000
RB2300	0,504	0,003	0,895	0,001	0,280	0,002
Sb19.3	0,317	0,003	0,221	0,002	0,928	0,001
TPA25	0,012	0,001	0,096	0,000	0,999	0,000

Como observado na tabela 15, foi detectada deficiência de heterozigotos para os *loci* TPA25 e PV92 considerando $\alpha = 0,05$, porém quando se considerado um $\alpha = 0,01$ esse resultado foi mantido apenas para o *locus* PV92. O mesmo teste detectou excesso de heterozigotos para o *locus* GC, mas apenas quando considerado $\alpha = 0,05$.

Tabela 16. Santiago do Iguape: teste de desequilíbrio de ligação *locus a locus*. Em destaque, valores que indicam existência de desequilíbrio de ligação antes de aplicada a correção de Bonferroni. EP: erro padrão.

<i>Locus x Locus</i>	P	EP	<i>Locus x Locus</i>	P	EP
TPA25	APO	1,000	APO	PV92	0,683
	PV92	0,478		D1	0,591
	D1	0,238		Sb19.3	0,478
	Sb19.3	0,749		AT3	0,487
	AT3	0,043		DRD2	0,199
	DRD2	0,200		LPL	0,606
	LPL	0,158		RB2300	0,413
	RB2300	0,619		OCA2	0,540
	OCA2	0,069		GC	0,792
	GC	0,106			
PV92	D1	0,758	D1	Sb19,3	0,425
	Sb19,3	0,333		AT3	0,488
	AT3	0,704		DRD2	0,584
	DRD2	0,976		LPL	0,178
	LPL	0,550		RB2300	0,344
	RB2300	0,752		OCA2	0,095
	OCA2	0,587		GC	0,818
	GC	0,653			
Sb19.3	AT3	0,685	AT3	DRD2	0,836
	DRD2	0,872		LPL	0,629
	LPL	0,111		RB2300	0,038
	RB2300	0,782		OCA2	0,619
	OCA2	1,000		GC	0,922
	GC	0,499			
DRD2	LPL	0,985	LPL	RB2300	0,106
	RB2300	0,712		OCA2	0,839
	OCA2	0,067		GC	0,981
	GC	0,324			
RB2300	OCA2	0,323	OCA2	GC	0,884
	GC	0,766			

O teste de desequilíbrio de ligação mostrou que, considerando $\alpha = 0,05$, os pares de *loci* em destaque na tabela 16 apresentam desvios quanto ao esperado no caso de se-

gregação independente. Quando considerado $\alpha_{\text{BONF}}=0,0009$, nenhum par de *loci* apresentou P indicativo de desvio quanto ao esperado.

Comparação das frequências gênicas e genotípicas

A Figura 4 apresenta uma comparação entre as frequências alélicas para os marcadores do tipo AIM que foram analisados nas populações de SAG, STI e Distrito Federal.

As comparações das frequências alélicas (diferenciação gênica e genotípica) observadas para todos os marcadores nas populações de SAG, STI e Distrito Federal mostraram diferenças altamente significativas ($P < 0,01$) entre os três pares de populações. A tabela 17 mostra os valores de P obtidos nesses testes.

Tabela 17. Valores de P obtidos pelo teste para diferenciação gênica e genotípica entre os pares de populações analisadas.

Populações	Diferenciação	
	Gênica (P)	Genotípica (P)
Santo Antônio do Guaporé Santiago do Iguape	0,0004	0,0003
Santo Antônio do Guaporé Brasília	$<10^{-4}$	$<10^{-4}$
Santiago do Iguape Brasília	$<10^{-4}$	$<10^{-4}$

O haplótipo 1 do gene DRD2 apareceu apenas em STI e o haplótipo 8, apenas em SAG. Apesar de não compartilharem todos os haplótipos, o teste para diferenciação gênica e genotípica mostrou não haver diferenças quanto à distribuição dos haplótipos nas duas populações ($P > 0,05$).

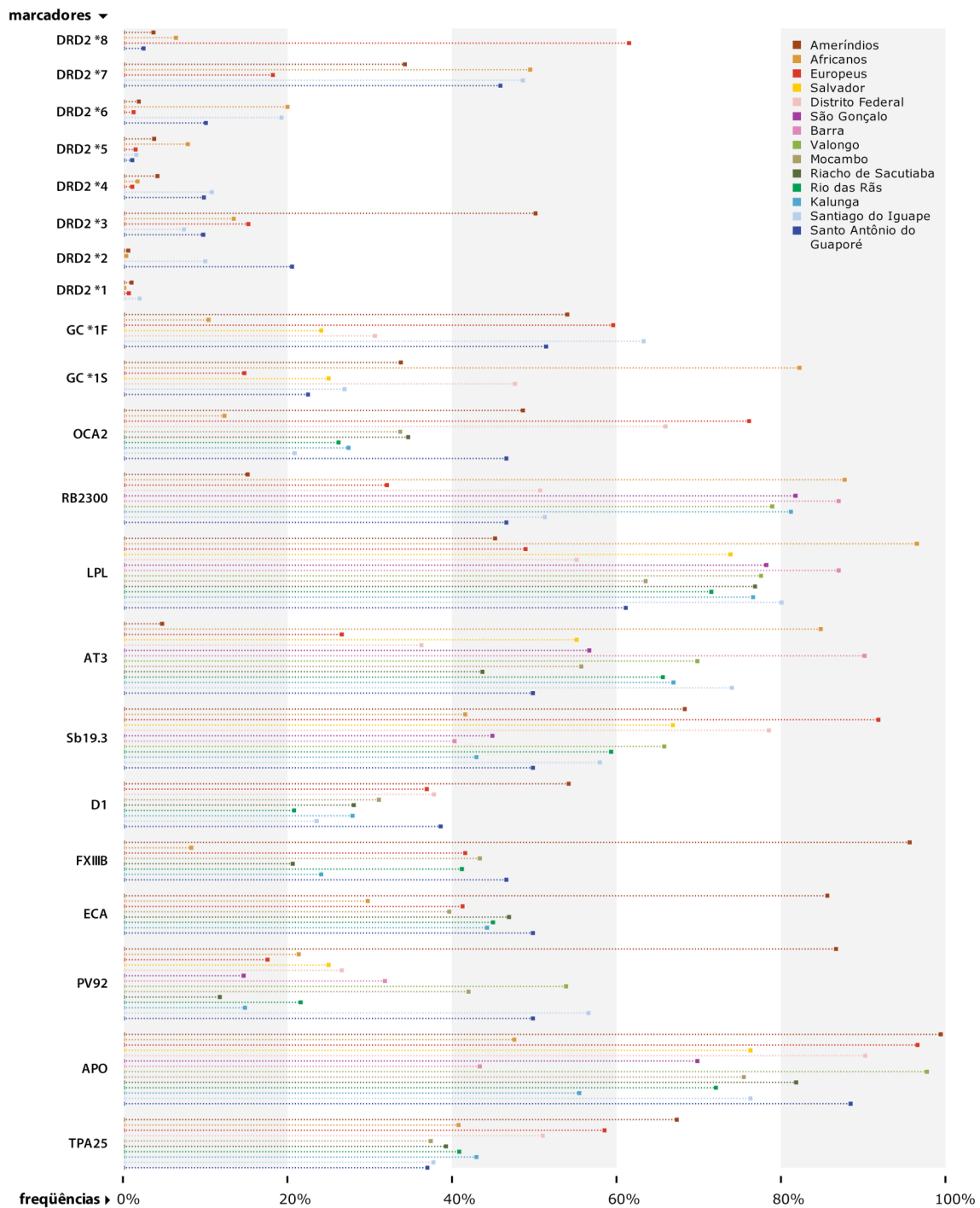


Figura 4. Freqüências dos marcadores disponíveis na literatura (entre os aqui analisados) para as populações afro-derivadas Santo Antônio do Guaporé, Santiago do Iguape, Kalunga, Rio das Rãs, Riacho de Sacutiaba, Mocambo, Barra, São Gonçalo e Valongo.

Estatísticas F

Fis – Coeficiente de endogamia

O coeficiente de endogamia – Fis – encontrado para as duas populações não foi significativo, tanto quando considerado um intervalo de confiança de 99%, quanto de 95%, e não foi muito distinto do observado no Distrito Federal (tabela 18), visto que os intervalos incluíram o zero.

Tabela 18. Valores de Fis e intervalos de confiança (95% e 99%) nas populações de Santo Antonio do Guaporé e Santiago do Iguape e comparação com o observado para o Distrito Federal.

População	Fis	IC	
		99%	95%
Santo Antônio do Guaporé	-0,002	-0,096 a 0,091	-0,069 a 0,068
Santiago do Iguape	0,064	-0,108 a 0,245	-0,064 a 0,194
Distrito Federal	0,015	-0,061 a 0,074	-0,040 a 0,066

Fst – Diversidade inter-populacional

Foram encontrados valores de Fst significativos entre as três populações e entre os três pares de populações: SAG e STI, SAG e DF e STI e DF (tabela 19), considerando IC de 95% (ICs não incluíram o zero). Para os pares SAG/STI e SAG/DF, os valores encontrados foram mais baixos que para o par STI/DF.

Tabela 19. Valores de Fst e intervalos de confiança (95% e 99%) para os três pares de populações e para as três populações em conjunto considerando 10 marcadores AIMs.

Pares de populações	Fst	IC	
		95%	99%
Santo Antônio do Guaporé Santiago do Iguape	0,031	0,006 a 0,059	-0,001 a 0,065
Santo Antônio do Guaporé Distrito Federal	0,040	0,013 a 0,076	0,003 a 0,084
Santiago do Iguape Distrito Federal	0,128	0,064 a 0,192	0,050 a 0,216
Santo Antônio do Guaporé Santiago do Iguape Distrito Federal	0,079	0,040 a 0,116	0,029 a 0,129

Estimativa de mistura

As contribuições parentais estimadas para as populações de SAG e STI, considerando todos os marcadores genéticos analisados para essas populações, encontram-se na tabela 20. A maior contribuição parental na constituição genética da população de SAG foi a ameríndia (42%), seguida da africana (37,6%) e da europeia (20,5%). Em STI, foi observada uma maior contribuição africana (56,8%), seguida da ameríndia (40,7%). A estimativa de contribuição europeia foi de 2,5%.

Tabela 20. Estimativas de mistura (m) para Santo Antônio do Guaporé e Santiago do Iguape calculadas a partir de todos os marcadores analisados para cada.

Parental	SAG		STI	
	m	EP	m	EP
Ameríndia	0,420	0,026	0,407	0,003
Africana	0,376	0,014	0,568	0,001
Europeia	0,205	0,024	0,025	0,003
R. Square	0,985		0,999	

Além da análise considerando todos os marcadores possíveis, as contribuições parentais para SAG foi também estimada a partir das frequências alélicas dos marcadores também analisados em STI e DF: TPA25, APO, PV92, D1, Sb19.3, AT3, LPL, RB2300, OCA2 e GC (tabela 21). Com a retirada dos marcadores DRD2*TaqI*, ECA e FXIIIB, a contribuição europeia na população de STI subiu de 2,5% para 6,3%. O perfil da população do Distrito Federal foi apresentado para fins de comparação.

Tabela 21. Estimativas de mistura estimadas a partir de 12 dos marcadores aqui analisados para as populações de Santo Antônio do Guaporé, Santiago do Iguape e Distrito Federal.

Parental	SAG		STI		DF	
	%	EP	%	EP	%	EP
Ameríndia	0,434	0,076	0,385	0,010	0,134	0,018
Africana	0,350	0,024	0,552	0,003	0,234	0,008
Europeia	0,216	0,071	0,063	0,009	0,632	0,019
R. Square	0,964		0,999		0,997	

As duas estimativas com e sem os marcadores DRD2*TaqI*, ECA e FXIIIIB para SAG e STI foram muito próximas. A contribuição africana em SAG e STI foi maior que em DF, como a contribuição ameríndia. Já a contribuição europeia foi muito maior em DF que nas outras duas populações.

Análise de componente principal

As frequências alélicas obtidas para SAG e STI e as de populações previamente estudadas para alguns dos marcadores aqui analisados foram utilizadas para gerar matrizes de distância. Essas matrizes (figuras 5 a 7) foram utilizadas na construção de gráficos de coordenada principal (figuras 8 a 10).

Para a construção da primeira matriz, foram considerados os dez marcadores analisados para as populações afro-derivadas em análise e para a do DF. Para a construção das outras duas, foi considerado o maior número de marcadores entre os aqui descritos disponíveis para todas as populações analisadas, que incluíram a população urbana de Salvador, Bahia, e sete remanescentes de quilombo brasileiros:

- matriz da figura 5: GC, APO, AT3, D1, LPL, OCA2, PV92, RB2300, Sb19.3 e TPA25;
- matriz da figura 6: TPA25, PV92, APO, D1, AT3, LPL e OCA2;
- matriz da figura 7: APO, PV92, LPL, Sb19.3, RB2300 e AT3.

	SAG	STI	DF	EUR	AFR	AME
SAG	0,000					
STI	0,037	0,000				
DF	0,055	0,112	0,000			
EUR	0,173	0,107	0,221	0,000		
AFR	0,097	0,187	0,136	0,448	0,000	
AME	0,117	0,183	0,039	0,412	0,182	0,000

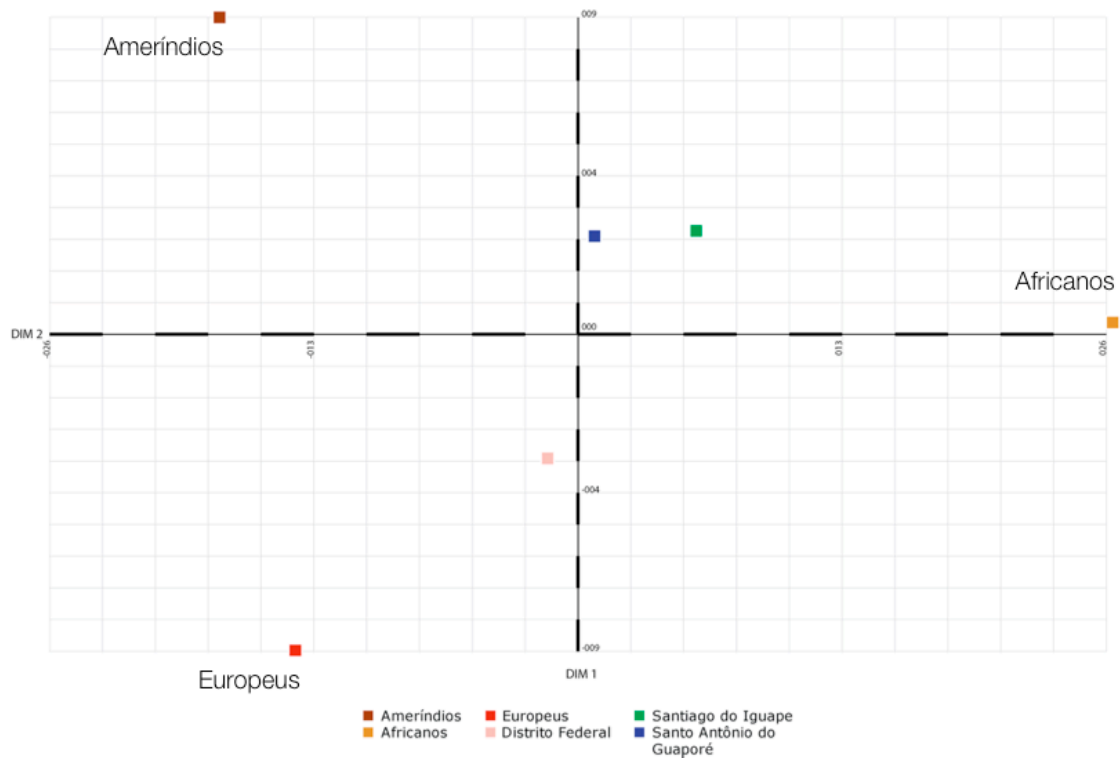
Figura 5. Distância de Nei: estimativas obtidas entre as populações em análise (SAG e STI), as populações parentais utilizadas no cálculo de mistura (EUR, AFR e AME) e do Distrito Federal (DF).

	SAG	STI	KAL	RAS	SAC	MOC	EUR	AFR	AME	DF
SAG	0,000									
STI	0,052	0,000								
KAL	0,099	0,066	0,000							
RAS	0,055	0,041	0,010	0,000						
SAC	0,055	0,090	0,037	0,021	0,000					
MOC	0,013	0,028	0,044	0,019	0,036	0,000				
EUR	0,086	0,260	0,198	0,158	0,094	0,126	0,000			
AFR	0,175	0,066	0,025	0,045	0,102	0,096	0,372	0,000		
AME	0,125	0,261	0,403	0,315	0,269	0,187	0,162	0,566	0,000	
DF	0,039	0,162	0,130	0,094	0,055	0,064	0,010	0,262	0,145	0,000

Figura 6. Distância de Nei: estimativas obtidas entre as populações em análise, as populações parentais utilizadas no cálculo de mistura, Distrito Federal e os remanescentes de quilombo Kalunga (KAL), Rio das Rãs (RAS), Riacho de Sacutiaba (SAC) e Mocambo (MOC).

	SAG	STI	SGO	BAR	VAL	KAL	EUR	AFR	AME	DF
SAG	0,000									
STI	0,037	0,000								
SGO	0,094	0,093	0,000							
BAR	0,189	0,010	0,060	0,000						
VAL	0,056	0,035	0,080	0,126	0,000					
KAL	0,124	0,099	0,009	0,029	0,107	0,000				
EUR	0,100	0,189	0,200	0,431	0,187	0,257	0,000			
AFR	0,193	0,109	0,043	0,006	0,128	0,021	0,402	0,000		
AME	0,129	0,245	0,448	0,696	0,285	0,536	0,152	0,716	0,000	
DF	0,046	0,106	0,104	0,267	0,093	0,146	0,018	0,250	0,160	0,000

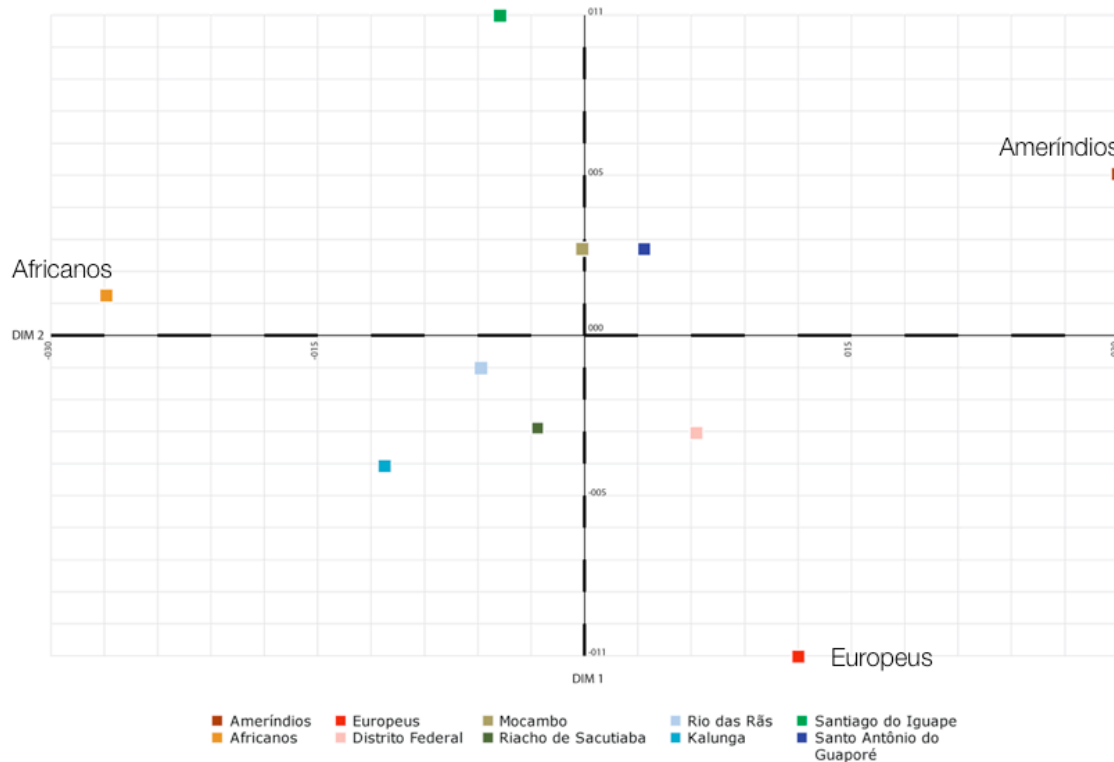
Figura 7. Distância de Nei: estimativas obtidas entre as populações em análise, as populações parentais utilizadas no cálculo de mistura, Distrito Federal e os remanescentes de quilombo São Gonçalo (SGO), Barra (BAR), Valongo (VAL) e Kalunga (KAL).



Componente	Eigenvalue	% explicada	% cumulativa
eixo 1	0,108	80,5	80,5
eixo 2	0,026	19,5	100

Figura 8. Componente principal gerado a partir de uma matriz de distância (Nei) entre Santo Antônio do Guaporé, Santiago do Iguape, Distrito Federal e as três populações parentais utilizadas nas análises de mistura.

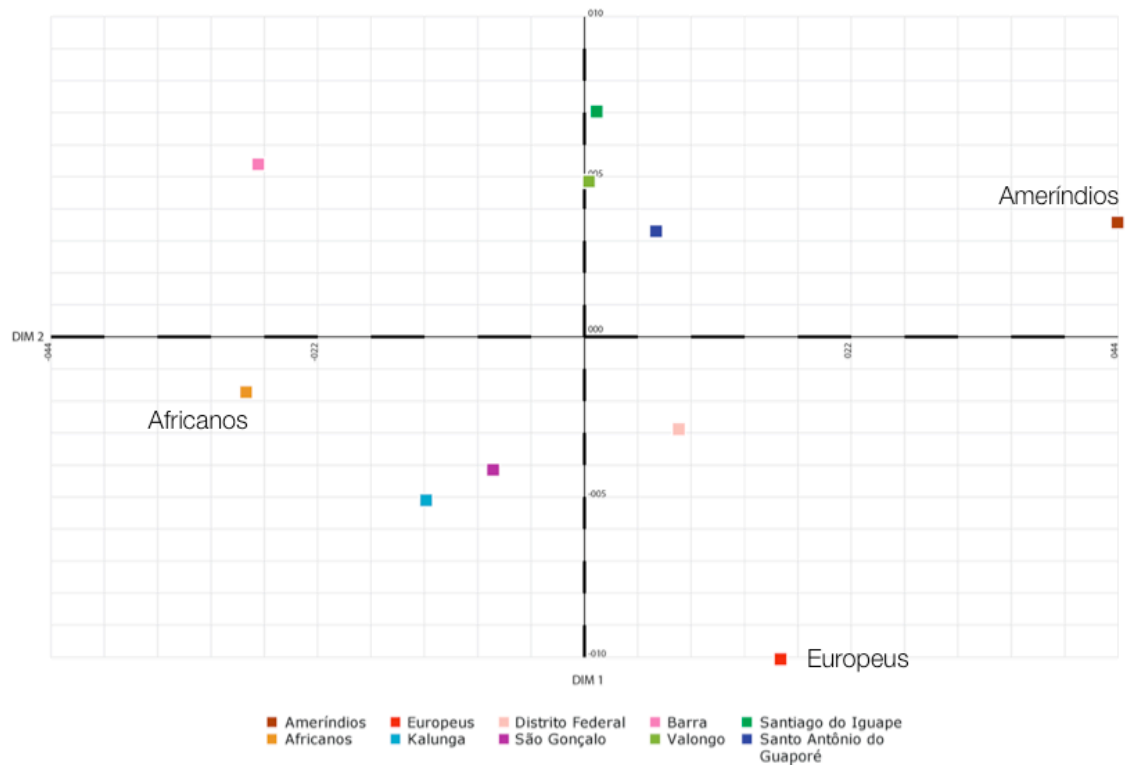
A figura 8 mostra o agrupamento das comunidades afro-derivadas SAG e STI e seu posicionamento intermediário entre as populações parentais ameríndia e africana, mais afastadas da parental europeia. Já a população do DF, posicionou-se mais próxima à parental europeia. O primeiro componente foi capaz de separar africanos dos demais grupos, mas ameríndios e europeus ficaram muito próximos. Já o segundo, diferenciou ameríndios e europeus, e posicionou SAG e STI, conjuntamente, entre africanos e ameríndios. A população do DF posicionou-se mais próxima da parental europeia. Os dois primeiros componentes explicaram 100% da variação observada.



Componente	eigenvalue	% explicada	% cumulativa
eixo 1	0,200	86,23	86,23
eixo 2	0,032	13,65	99,90
eixo 3	0,000	0,10	100,00

Figura 9. Componente principal gerado a partir de uma matriz de distância (Nei) entre Santo Antônio do Guaporé, Santiago do Iguape, Distrito Federal, os remanescentes de quilombo Kalunga (KAL), Rio das Rãs (RAS), Riacho de Sacutiaba (SAC), Mocambo (MOC) e as três populações parentais utilizadas nas análises de mistura.

A Figura 9 mostra que todas as populações híbridas posicionaram-se entre as três parentais, mas com STI mais afastada das demais. O primeiro componente separou as populações parentais, posicionando as híbridas entre africanos e ameríndios e europeus. O segundo, posicionou STI, SAG e MOC entre africanos e ameríndios e as demais populações entre africanos e europeus. A população do DF posicionou-se mais próxima da parental europeia. Os dois primeiros componentes explicaram 99,9% da variação observada.



Componente	eigenvalue	% explicada	% cumulativa
eixo 1	0,401	92,90%	92,90%
eixo 2	0,028	6,50%	99,40%
eixo 3	0,002	0,40%	99,80%

Figura 10. Componente principal gerado a partir de uma matriz de distância (Nei) entre Santo Antônio do Guaporé, Santiago do Iguape, Distrito Federal, os remanescentes de quilombo São Gonçalo (SGO), Barra (BAR), Valongo (VAL) Kalunga (KAL) e as três populações parentais utilizadas nas análises de mistura.

Na figura 10, as três populações parentais foram diferenciadas pelo primeiro componente, que não diferenciou a população de BAR da parental africana. SAG, STI e VAL agruparam-se entre as parentais africana e ameríndia. As demais populações híbridas posicionaram-se entre as parentais e DF, mais próxima da parental europeia. O segundo componente diferenciou BAR de AFR. Os dois primeiros componentes explicaram 99,4% da variação observada.

Discussão

Equilíbrio de Hardy-Weinberg e Desequilíbrio de Ligação

Foi observado desvio na distribuição genotípica do marcador PV-92, TPA-25 e GC com α de 5% e apenas PV-92 com α de 1% na população de Santiago do Iguape. Nenhum desvio na população de Santo Antonio do Guaporé foi observado.

Como a população de SAG (cerca de 100 habitantes) é consideravelmente menor que a de STI (o distrito de STI é povoado por cerca de 3000 indivíduos), poderia ser esperado, a princípio, que fossem encontrados desvios do EHW na primeira, já que o pequeno tamanho é um dos fatores que deixam uma população mais suscetível aos efeitos da deriva genética (Futuyma, 1998). Apesar de pequena, a amostra de SAG representa cerca de 30% da população total e é, portanto, mais provável que represente a variabilidade genética total de seus habitantes e sua distribuição. Já a amostra de STI, representa uma porcentagem menor da população. Os desvios quanto ao EHW em STI devem ter sido causados por erro amostral, apesar de ser possível que tenha havido a ação de outros fatores evolutivos perturbadores da situação de equilíbrio como fluxo gênico e deriva genética.

Foi encontrado DL entre sete pares de marcadores para SAG e dois, para STI. Tais desvios podem ser consequência de desequilíbrio de ligação gerado por mistura (ALD – *admixture linkage disequilibrium* - Hartl e Clark, 2007), já que o processo de miscigenação entre populações com frequências distintas dos marcadores gera DL (Pfaff *et al.*, 2001).

Pfaff e colaboradores (2001) propuseram dois modelos de formação de populações híbridas (figura 11), *hybrid-isolation* (HI), em que há mistura nas primeiras gerações e subsequente isolamento, e *continuous-gene-flow* (CGF), em que ocorre influxo gênico ao longo das gerações, e testaram o comportamento do DL ao longo de gerações. Nos dois casos, o DL tende a diminuir ao longo das gerações (figura 12), mesmo entre *loci* fisicamente próximos, ainda que com velocidades diferentes.

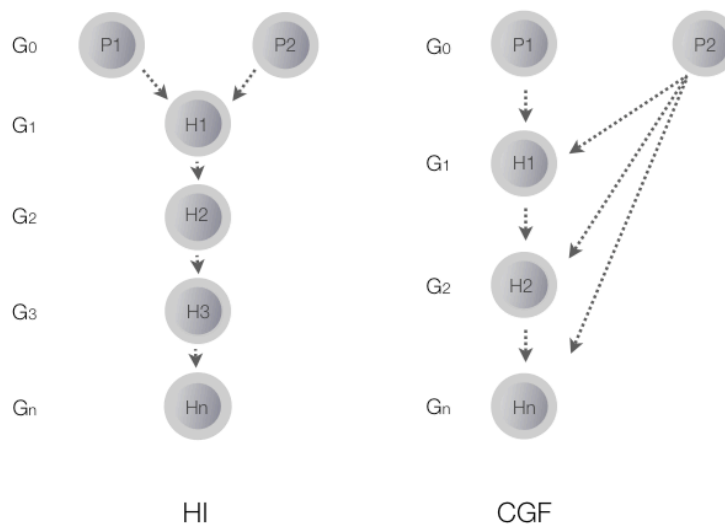


Figura 11. Modelos de formação propostos por Pfaff e colaboradores (2001): hybrid-isolation (HI) e continuous-gene-flow (CGF). Modificado de Pfaff e colaboradores (2001).

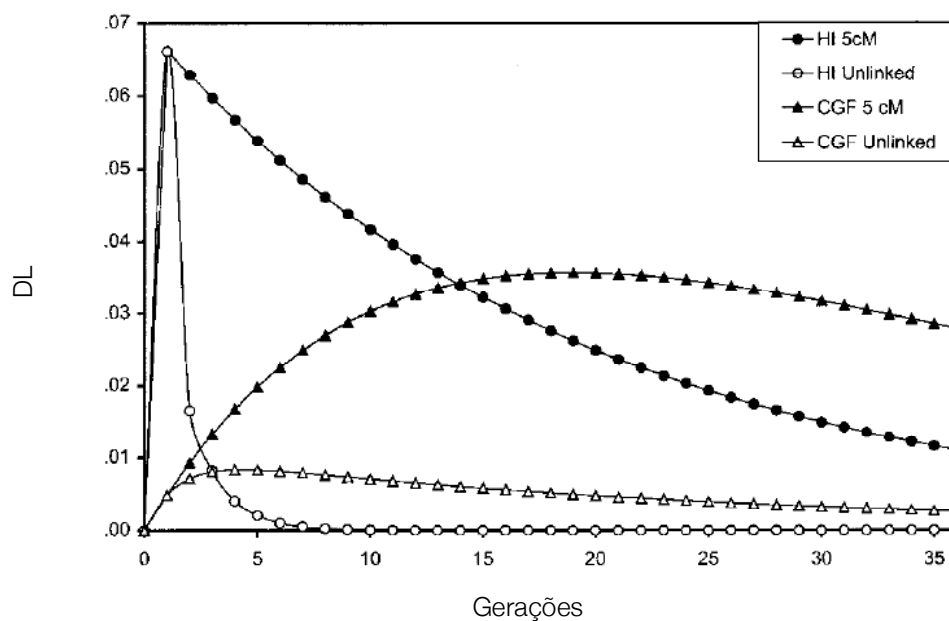


Figura 12. Quantidade de ALD esperada sob cada modelo de mistura (HI e CGF) para *loci* não ligados e para *loci* ligados a 5cM. HI: modelo *hybrid-isolation*; CGF: modelo *continuous-gene-flow*; DL: desequilíbrio de ligação. 5cM: modelo para marcadores ligados. *Unlinked*: modelo para marcadores sem ligação. Figura retirada de Pfaff *et al.* (2001).

Apesar de a história e localização de SAG sugerirem que sua história demográfica se assemelhe mais ao modelo HI, os dados genéticos, sugerem maior aproximação com o modelo CGF. Após apenas cinco gerações, o DL deveria deixar de existir em uma população descrita adequadamente pelo modelo HI, mas foi encontrado DL entre alguns pares de *loci* em SAG. Além disso, a alta contribuição ameríndia e a baixa, mas significativa contribuição européia sugerem a existência de fluxo gênico entre ela e populações vizinhas, em especial ameríndios.

Já a história e localização de STI sugerem que sua história demográfica se assemelhe ao modelo CGF. Novamente os dados genéticos sugerem adequação a outro modelo, no caso HI: foi encontrado DL em apenas dois dos pares de *loci* testados para essa população, o que, como decorreram mais de cinco gerações desde a formação dessa comunidade, não é condizente com o modelo CGF. A baixa estimativa de contribuição européia reforça o modelo de HI, já que indica baixo fluxo gênico com populações vizinhas.

A proposta de que há diferenças na história demográfica entre as duas populações é condizente com a localização geográfica das mesmas. Entretanto SAG está localizada no estado de Rondônia em uma região que apresenta dificuldade de acesso à área urbana mais próxima, que é o município de Costa Marques, e STI, no recôncavo baiano, em um local de acesso mais facilitado a Cachoeira, zona urbana mais próxima desse povoado. A diferença no acesso a outras populações é fator determinante na manutenção de fluxo gênico, mas aparentemente algum outro fator influenciou essa característica tanto em SAG, quanto em STI, já que os dados genéticos sugerem ter havido fluxo gênico constante em SAG e não em STI.

Frequências alélicas e genotípicas

Com a formação de uma população miscigenada, espera-se que a frequência de qualquer *locus*, e em especial de *loci* AIMS, definidos como marcadores que apresentam frequências extremas e distintivas entre populações diferenciadas por fatores étnicos ou geográficos, atinja valor intermediário ao encontrado nas populações parentais. As frequências alélicas e haplotípicas encontradas nas duas populações afro-derivadas SAG e STI seguem, de forma geral, essa previsão, como mostrado na figura 4 que compara,

além das populações aqui analisadas, as frequências de outras populações urbanas e afro-derivadas quanto às frequências dos marcadores disponíveis na literatura.

Em SAG e STI, alguns haplótipos do marcador DRD2*TaqI* e, apenas em STI, o alelo GC*1F, apresentaram frequências extremas (mais altas ou mais baixas que nas populações parentais). Tal discrepância quanto ao esperado pode indicar efeito fundador, estruturação populacional, sucesso reprodutivo diferencial e/ou deriva genética, especialmente em populações pequenas como SAG, que tendem ao aumento da homozigose gerada por deriva genética, endogamia e fluxo gênico, apesar de não ter sido identificado nessa população excesso de homozigotos para os *loci* aqui analisados.

A grande distância que separa SAG e STI, aliado às diferentes dinâmicas demográficas entre elas e as populações vizinhas, tornaram essas comunidades extremamente diferentes quanto às frequências gênicas e à distribuição genotípica, apesar da história em comum de formação (já que as duas foram formadas por escravos fugitivos). Outro fator que pode ter gerado essa diferença é a origem dos escravos que as formaram: como as duas comunidades foram fundadas em épocas diferentes, a origem africana de seus fundadores provavelmente não é a mesma, já que ao longo da história, escravos foram trazidos de diferentes regiões da África. Como dito anteriormente, durante o século XVI, quando foi fundada STI, a maioria dos escravos trazidos para o Brasil era capturada e embarcada no Congo e norte de Angola, enquanto no século XVIII, quando foi fundada SAG, eles eram capturados principalmente no sul de Angola e regiões internas próximas aos grandes lagos.

As análises de comparação gênica e genotípica par a par mostraram haver, entre os pares de populações SAG/STI, SAG/DF e STI/DF, diferenças significativas que indicam que, além de serem diferentes uma da outra, SAG e STI são diferentes de uma população brasileira urbana comum formada por fluxo migratório de todas as regiões do país, que é o Distrito Federal.

Diferenças significativas, como as observadas entre SAG e STI, foram também encontradas entre outras populações afro-derivadas analisadas para marcadores AIMs, como os remanescentes de quilombo analisados por Luizon (2007) e Pedrosa (2006). Essas comunidades também são diferentes entre si e com relação a populações urbanas, como indicado pelo mesmo teste e pelo teste de *F_{st}*.

Estimativas de mistura

As estimativas de mistura mostram que, para as duas populações afro-derivadas, existe uma grande contribuição das parentais ameríndia e africana e uma menor da parental europeia.

Na população de SAG, a contribuição ameríndia foi maior que a africana. Essa população habita uma área ocupada por várias aldeias indígenas, com quem se relacionam desde a formação da comunidade. A menor contribuição europeia observada reflete o grau de isolamento dessa comunidade com relação a áreas urbanas, já que, além do isolamento gerado por sua localização, na década de 80, com a criação da Reserva Biológica do Guaporé, outras populações que residiam na área, inclusive indígenas, foram removidas.

A população de STI também apresentou contribuição ameríndia elevada, mas a africana foi mais alta nessa população. A contribuição ameríndia em remanescentes de quilombo tem sido frequentemente descrita, visto o relacionamento entre essas comunidades e aldeias indígenas próximas (Ferreira, 2007; Lima, 2002).

Os sete remanescentes de quilombo analisados para marcadores AIMs por Luizon (2007) e Pedrosa (2006) também apresentam, como STI e SAG, alta contribuição africana. Essas populações, entretanto, apresentaram contribuição ameríndia mais baixa e europeia mais alta, com exceção de Barra, para a qual não foi possível detectar contribuição europeia. Entre elas, Mocambo, que manteve ao longo de sua história uma intensa relação com a população ameríndia Xocó, apresentou a estimativa de contribuição ameríndia mais alta (18%), mas que ainda assim é consideravelmente menor que a estimada para SAG (42%) e STI (40,7%).

Para a população de Salvador, uma população urbana de marcada ancestralidade africana, as contribuições parentais foram estimadas em 49,2% de africana, 36,3% europeia e 14,5% de ameríndia (Machado, 2008). Apesar da alta contribuição africana, similar à estimada para SAG, STI e os remanescentes de quilombo estudados por Luizon (2007) e Pedrosa (2006), a alta contribuição europeia diferencia essa população de SAG, STI, BAR e SGO, nas quais a mais alta estimativa europeia foi encontrada em SGO (21,6%). A contribuição ameríndia, estimada em 14,5%, foi similar às encontrada em SGO (10,3%) e MOC (18,81%), mas consideravelmente mais alta que as encontradas em BAR (4,6%), SAC (2,73%) e VAL (não detectada).

Estatísticas F

Apesar do pequeno tamanho amostral e populacional, este último especialmente em SAG, as estimativas de Fis não mostraram existir endogamia em nenhuma das duas populações. Resultado semelhante foi encontrado para os remanescentes de quilombo Valongo (SC) e São Gonçalo (BA), mas diferente para o remanescente de Barra (BA), (Luizon, 2007).

O valores de Fst encontrados foram significativos (ICs não incluem o zero) e mostram, assim como os testes de diferenciação gênica e genotípica par a par e a comparação das análises de mistura, existirem diferenças significativas entre as duas populações afroderivadas e destas com relação a população urbana do Distrito Federal. Essas diferenças eram, de fato, esperadas visto a localização geográfica destas e a expectativa de que a história das comunidades seja bastante distinta, inclusive quanto às proporções das populações parentais em sua formação.

Análise de Componente principal

Com relação à análise de componente principal, a estrutura de agrupamento das populações e destas com as parentais reflete os resultados de mistura e é condizente com as frequências alélicas aqui descritas. SAG e STI tenderam a ficar mais próximas e localizadas entre as três parentais, embora mais próximas às parentais africana e ameríndia. Já DF, posicionou-se mais próxima à parental europeia. O primeiro componente posicionou as duas populações afro-derivadas e DF entre a parental africana e o agrupamento ameríndio-europeu (gerado pelos δ existentes entre as frequências dos marcadores selecionados). Já o segundo componente separou claramente ameríndios de europeus e posicionou SAG e STI entre ameríndios e africanos e DF, entre africanos e europeus.

Nos gráficos comparando 1. as populações aqui analisadas, SAG, STI, outros remanescentes, KAL, RAS, SAC, MOC, a população urbana do DF e as populações parentais e 2. SAG, STI, SGO, BAR, VAL, KAL e as parentais, foi possível observar que todas as populações híbridas posicionaram-se entre as parentais. De forma geral, as posições relativas das populações nos dois gráficos foram coerentes com os resultados de mistura. No primeiro (figura 9), SAG, STI e MOC posicionaram-se entre africanos e ameríndios, KAL, RAS e SAC, entre africanos e europeus e DF, mais próximo a europeus. No segundo (figura 10), KAL e SGO posicionaram-se entre africanos e europeus,

DF, mais próximo a africanos e as demais populações, entre africanos e ameríndios, apesar de o primeiro componente não ter diferenciado BAR de africanos. Dessa forma, foi possível verificar que essa técnica de comparação é eficiente mesmo quando as populações comparadas apresentam mistura genética. Nesse caso, a coordenada principal mostrou agrupamentos coerentes com a contribuição de cada parental na composição genética atual dessas comunidades revelando que, mesmo após gerações de possível fluxo gênico, a maioria delas ainda mantém uma contribuição africana e ameríndia acima da observada nas regiões urbanas.

Conclusões

A análise de 13 marcadores informativos de ancestralidade (AIMs) autossômicos em duas comunidades afro-derivadas brasileiras - Santo Antônio do Guaporé e Santiago do Iguape – e a comparação com uma população brasileira urbana - Distrito Federal - e outras comunidades previamente descritas na literatura permitiram concluir que:

- As duas comunidades diferiram significativamente entre si com relação às frequências gênicas e genotípicas, assim como com relação à população do Distrito Federal;
- As duas comunidades apresentam diferenciação genética de acordo com a estimativa de F_{st} . Essa observação reflete a grande distância geográfica que as separa e as diferentes origens dos indivíduos que as formaram e, conseqüentemente, suas diferentes histórias e relações com populações próximas;
- Foi observada diferença estatisticamente significativa (F_{st}) entre as duas comunidades e a população do DF;
- Observou-se expressiva contribuição africana e ameríndia para a formação dessas comunidades. Essa observação, especialmente para Santo Antônio do Guaporé, é condizente com relatos históricos de relacionamento entre essa população e populações ameríndias que habitaram a mesma região;
- A contribuição européia na formação dessas comunidades foi baixa, especialmente em Santiago do Iguape;
- As duas comunidades apresentaram composição genética similar à de outras populações afro-derivadas no que diz respeito à alta contribuição africana, mas apresentaram, de forma geral, maior contribuição ameríndia que as demais. Essas observações podem ser observadas refletidas nos gráficos de coordenadas principais.

Referência bibliográfica

- Abe-Sandes, K. *Diversidade genética de afro-brasileiros: DNA mitocondrial e cromossomo Y*. 2002. Tese (Doutorado em Genética de Populações) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (USP), 2002.
- Alencastro, LF. (2000) *O trato dos Viventes: A formação do Brasil no Atlântico Sul*. 1ed. Companhia das Letras, São Paulo, SP.
- Arcot, SS; Shaikh, TH; Kim, J; Bennet, L; Alegria-Hartman, M; Nelson, DO; Deininger, PL; Batzer, MA. Sequence diversity and chromosomal distribution of young *Alu* repeats. *Gene*, 163 (1995) 273 – 278.
- Arpini-Sampaio, Z.; Costa, M.C.; Melo, A.A.; Carvalho, M.F.; Deus, M.S.; Simões, A.L. (1999). Genetic polymorphisms and admixture in African-derived black communities of northeastern Brazil. *Hum Biol* 71(1):69-85.
- Barbujani, G.; Magagni, A.; Minch, E.; Cavalli-Sforza, L.L. (1997). An apportionment of human DNA diversity. *Proc Natl Acad Sci USA* 94:4516–4519.
- Barcelos, R.S.S. (2005) *Constituição genética de populações urbanas do Centro-Oeste brasileiro (Goiás e Distrito Federal) estimada por marcadores uniparentais Y-específicos e do DNAm*. Tese de doutorado. Programa de Pós-Graduação em Biologia Animal. Universidade de Brasília. Brasília – DF.
- Barickman, B.J. (1999) Até A Véspera: O Trabalho Escravo e a Produção de Açúcar nos Engenhos do Recôncavo Baiano (1850-1881). *Afro-Ásia*, 21-22, 177-238.
- Barickman, B.J. (2003) E se a casa não fosse tão grande? Uma freguesia açucareira no Recôncavo Baiano em 1835. *Afro-Ásia*, 20-31, 79-132
- Battilana J, Bonatto SL, Freitas LB, Hutz MH, Weimer TA, Callegari-Jacques SM, Batzer MA, Hill K, Hurtado AM, Tsuneto LT, Petzl-Erler ML, Salzano FM (2002) *Alu* insertions versus blood group plus protein genetic variability in four Amerindian populations. *Ann Hum Biol*. May-Jun;29(3):334-47.
- Batzer, M.A.; Rubin, C.M., Hellmann-Blumberg, U.; Alegria-Hartman, M.; Leeflang, E.P.; Stern, J.D.; Bazan, H.A., Shaikh, T.H.; Deininger, p.L.; Schmid, C.W. (1995) Dispersion and Insertion Polymorphism in Two Small Subfamilies of

- Recently Amplified Human Alu Repeats. *J. Mol. Biol.* (1995) 247, 418–427
- Batzer, MA; Arcot, SS; Phinney, JW; Alegria-Hartman, M; Kass, DH; Milligan, SM; Kimpton, C; Gill, P; Hochmeister, M; Ioannou, PA; Herrera, RJ; Boudreau, DA; Scheer, WD; Keats, BJB; Deininger, PL; Stoneking, M. Genetic Variation of Recent *Alu* Insertions in Human Populations. *J Mol Evol* (1996) 42:22 – 29.
- Batzer, MA; Deininger, PL. A human-specific subfamily of Alu sequences. *Genomics*. 9(3):481-7. 1991.
- Batzer, MA; Stoneking, M; Alegria-Hartman, M; Bazan, H; Kass, DH; Shaikh, TH; Novick, GE; Ioannou, PA; Scheer, WD; Herrera, RJ; Deininger, PL. African origin of human-specific polymorphic *Alu* insertions. *Proc. Natl. Acad. Sci.* 91:12288-12292. 1994.
- Bonilla, C.; Parra, E.J.; Dios, S.; Marshall, J.A.; Hamman, R.F.; Ferrel, R.E.; Hoggart, C.L.; Mckeigue, P.M.; Shriver, M.D. Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann. Hum. Genet.* 68: 139-153, 2004.
- Bookstein, R.; Lai, C. C.; To, H. & Lee, W. H. PCR-based detection of a polymorphic *Bam*HI site in intron 1 of the human retinoblastoma (RB) gene. *Nucleic Acids Res*, 18:1666, 1990.
- Bortolini, M. C.; Weimer, T. A.; Salzano, F. M.; Callegari-Jacques; Schneider, H.; Layrisse, Z.; Bonatto, S. L. Evolutionary relationships between Black South American and African Populations. *Hum Biol*, 67: 547-559, 1995.
- Bortolini, M.C.; Silva-Jr, W.A.; Guerra, D.C.; Remonato, G.; Mirandola, R.; Hutz, M.H.; Weimer, T.A.; Silva, M.C.B.O.; Zago, M.A.; Salzano, F.M. (1999). African-derived South American Populations: A History of Symmetrical and Asymmetrical Matings According to Sex Revealed by Bi- and Uni-parental Genetic Markers. *Am J Hum Biol* 11:551- 563.
- Bortolini, M.C.; Silva-Jr, W.A.; Weimer, T.A.; Zago, M.A.; Guerra, D.C.; Schneider, M.P.; Layrisse, Z.; Castellano, H.M.; Salzano, F.M. (1998a). Protein and hyper-variable tandem repeat diversity in eight African-derived South American populations: inferred relationships do not coincide. *Hum Biol* 70(3):443-461

- Bortolini, M.C.; Weimer, T.A.; Franco, M.H.L.P.; Salzano, F.M.; Layrisse, Z.; Schneider, H.; Schneider, M.P.C.; Harada, M.L. (1992). Genetic studies in three American black populations. *Gene Geogr* 6:1-16.
- Bortolini, MC; Baptista, C; Callegari-Jacques, SM; Weimer, TA; Salzano, FM. Diversity in protein, nuclear DNA, and mtDNA in South Amerinds – agreement or discrepancy? *Ann. Hum. Genet.* 62: 133-45. 1998b.
- Brookes, A.J. (1999). The essence of SNPs. *Gene.* 234:177-186.
- Carvalho, B.M.; Bortolini, M.C.; Santos, S.E.B.; Ribeiro-dos-Santos, A.K.C. (2008) Mitochondrial DNA mapping of social-biological interactions in Brazilian Amazonian African-descendant populations. *Genet. Mol. Biol.* Vol.31 no.1. São Paulo 2008
- Cavalli-Sforza, L. L.; Menozzi, P.; Piazza, A. *History and geography of human genes.* Princeton: Princeton University Press, 1994.
- Chakraborty, R. Gene identity in racial hybrids and estimation of admixture rates. Em.: Ahuja YR, Neel JV, editors. Genetic differentiation in human and others animal populations. *Indian Anthropological Association.* 1985. p, 171-180.
- Chakraborty, R.; Kamboh, M.I.; Ferrell, R.E. (1991). Unique alleles in admixed populations: a strategy for determining hereditary population differences of disease frequencies. *Ethn Dis* 1:245-256.
- Collins, F.S.; Brooks, L.D.; Chakravarti, A. (1998). A DNA Polymorphism Discovery Resource for Research on Human Genetic Variation. *Genome Res* 8:1229-1231.
- Collins-Schramm, H. E.; Chima, B.; Morii, T.; Wah, K.; Figueroa, Y.; Criswell, L.A.; Hanson, R. L.; Knowler, W. C.; Silva, G.; Belmont, J. W.; Seldin, M.F. Mexican American ancestry-informative markers: Examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians. *Hum Genet* 114:263–271, 2004.
- Dean, M.; Stephens, J.C.; Winkler, C.; Lomb, D.A.; Ramsburg, M.; Boaze, R.; Stewart, C.; Charbonneau, L.; Goldman, D.; Albaugh, B.J.; Goedert, J.J.; Beasley, R.P.; Hwang, L.; Buchbinder, S.; Weedon, M.; Johnson, P.A.; Eichelberger, M.; O'Brien, S.J. (1994). Polymorphic admixture typing in human ethnic populations. *Am J Hum Genet* 55:788-808.

- Enoch, MA., Shen, PH.; Hodgkinson, C.; Goldman, D. (2006) Using ancestry-informative markers to define populations and detect population stratification. *Journal of Psychopharmacology* 20(4) Supplement (2006) 19–26
- Excoffier, L.G.L. & Schneider, S. (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.
- Fausto, B. (2001) *História Concisa do Brasil*. São Paulo, SP.
- Fausto, B. (2002) *História do Brasil*. 10 ed. Editora da Universidade de São Paulo. São Paulo, SP.
- Felsenstein J (1989) PHYLIP – phylogeny inference package. *Cladistics* 5:164–166.
- Futuyma, D.J. (1998). *Evolutionary Biology*. 3.ed. Sinauer, Sunderland, MA. 751pp.
- Gelernter, J; Kranzler, H; Cubells, JF; Ichinose, H; Nagatsu, T. DRD2 Allele Frequencies and Linkage Disequilibria, Including the -141C/Del Promoter Polymorphism, in European-American, and Japanese Subjects. *Genomics* 51, 21 – 26 (1998).
- Godinho, N.M.O. (2008) *O Impacto das migrações na constituição genética de populações latino-americanas*. Tese de doutorado. Programa de Pós Graduação em Biologia Animal. Universidade de Brasília. Brasília – DF.
- González-Andrade, F.; Sánchez, D.; González-Solórzano, J.; Gascón, S.; Martínez-Jarreta, B. (2007) Sex-specific genetic admixture of mestizos, amerindian kichwas, and afro-ecuadorans from Ecuador. *Human Biology*, v.79, no. 1, pp. 51-77
- Gotoda T, Yamada N, Murase T, Shimano H, Shimada M, Harada K, Kawamura M, Kozaki K, Yazaki Y (1992) Detection of three separate DNA polymorphisms in the human lipoprotein lipase gene by gene amplification and restriction endonuclease digestion. *J Lipid Res*. 33(7):1006-72.
- Guerreiro, J.F.; Ribeiro-Dos-Santos, A.K.C.; Santo, E.F.M.; Vallinoto, A.C.R.; Cayres-Vallinoto, I.M.V.; Aguiar, G.F.S.; Santos, S.E.B. (1999). Genetic-demographic data from two Amazonian populations composed of descendants of African slaves: Pacoval and Curiau. *Genet Mol Biol* 22:163-167.

- Ha, JC; HA, RR. *Integrative Statistics for Behavior Science*. In Press.
- Hartl, D.L.; Clark, A.G. *Principles of population genetics*. 4th ed. Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts, 2007.
- Hui-Ju Tsai, HJ.; Choudhry, S.; Naqvi, M.; Rodriguez-Cintron, W.; Burchard, E.G.; Ziv, E. (2005) Comparison of three methods to estimate genetic ancestry and control for stratification in genetic association studies among admixed populations. *Hum Genet* (2005) 118: 424–433
- Hutz, M.H.; Callegari-Jacques, S. M., Bortolini, M. C.; Salzano, F. M. Variability in nDNA, mtDNA and proteins: a test case. *In Genomic Diversity: Applications in Human Population Genetics*. Kluwer Academic/Plenum Publishers, New York, p.23-32, 1999.
- Jorde, LB; Bamshad, MJ; Watkins, WS; Zenger, R; Fraley, AE; Krakowiak, PA; Carpenter, KD; Soodyall, H; Jenkins, T; Rogers, AR. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. *Am. J. Hum. Genet.* 57: 523-38. 1995.
- Karathanasis *et al.*, 1985 em Cotrim, NH. Variabilidade Genética das Inserções de *Alu* em Remanescentes de Quilombos. Dissertação de mestrado apresentada ao Instituto de Biociências da Universidade de São Paulo, São Paulo. 2003.
- Kidd, JR; Black, FL; Weiss, KM; Balazs, I; Kidd, KK. Studies of three Amerindian populations using nuclear DNA polymorphisms. *Hum. Biol.* 63 (6): 775-794. 1991.
- Kass, DH; Aleman, C; Batzer, MA; Deininger, PL. Identification of a human specific *Alu* insertion in the factor XIII B gene. *Genetica.* 94(1): 1-8. 1994.
- Kofler, A.; Braun, A.; Jenkins, T.; Serjeantson, S. W.; Cleve, H. :Characterization of mutants of the vitamin-D-binding protein/group specific component: GC Aborigine (1A1) from Australian Aborigines and South African blacks, and 2A9 from South Germany. *Vox Sang.* 68: 50-54, 1995.
- Lee, ST; Nicholls, RD; Jong, MTC; Fukai, K; Spritz, RA. Organization and sequence of the human *P* gene and identification of transport proteins. *Genomics* 26, 354 – 363 (1995)
- Lewis, P. O.; Zaykin, D. *Genetic data analysis: software for the analysis of discrete*

- genetic data*. Version 1.0. 1997.
- Lima, R.R. *A contribuição masculina na formação de três remanescentes de quilombos do Vale do São Francisco (nordeste do Brasil) avaliada por marcadores do cromossomo Y*. 2002. Dissertação (Mestrado em Genética de Populações) - Instituto de Ciências Biológicas, Universidade de Brasília (UnB), Brasília, 2002.
- Liu, Y.; Saha, N.; Low, P. S.; Tay, J. S. Linkage disequilibrium between two loci (5'untranslated exon 1 and intron 5-DdeI) of the antithrombin III gene in three ethnic groups in Singapore. *Hum Hered*, 45:192-198, 1995.
- Luizon, M.R. (2003) *Polimorfismos de DNA população-específicos em indígenas da Amazônia Brasileira*. Dissertação de mestrado. Faculdade de Medicina de Ribeirão Preto. USP. Ribeirão Preto – SP.
- Luizon, M.R. (2007) *Dinâmica da Mistura Étnica em Comunidades Remanescentes de Quilombo Brasileiras*. Tese de Doutorado. Faculdade de Medicina de Ribeirão Preto. USP. Ribeirão Preto – SP.
- Lum, JK; Cann, RL; Martinson, JJ; Jorde, LB. Mitochondrial and nuclear genetic relationships among Pacific Island and Asian populations. *Am J Hum Genet* 63(2):613-24. 1998.
- Machado, T.M.B.M. (2008) *Ancestralidade em Salvador-BA*. Dissertação de Mestrado em Biotecnologia em Saúde e Medicina Investigativa - Centro de Pesquisa Gonçalo Moniz - FIOCRUZ - BA,
- Market, CL; Müller, F. Multiple forms of enzyme: tissue, ontogenetic, and species specific patterns. *Proc. Natl. Acad. Scien.* 45: 753-63. 1959.
- Martínez, H.; Rodríguez-Larralde, A.; Izaguirre, M.H.; de Guerra, D.C. (2007) Admixture estimates fo Caracas, Venezuela, based on autosomal, Y-chromosome, and mtDNA Markers. *Human Biology*, v. 79, no. 2, pp. 201-213
- Martínez-Marignac, V.; Bertoni, B.; Parra, E.J.; Bianchi, N.O. (2004) Characterization of admixture in an urban sample from Buenos Aires, Argentina, using ui-parentally and biparentally inherited genetic markers. *Human Biology* v. 76, no 4, pp. 543-557
- Mello e Souza, M. (2006) *Reis Negros no Brasil Escravista: História da Festa de Coroação de Rei Congo*. 1 ed. Editora UFMG, Belo Horizonte, MG – Brasil.

- Nakai, KMD; Itoh, CMD; Miura, Y; Hotta, K; Musha, T; Itoh, T; Miyakawa, T; Iwasaki, R; Hiramori, MD. Deletion polymorphism of the angiotensin I-converting enzyme is associated with serum ACE concentration and increase risk for CAD in the Japanese. *Circulation*. 90: 2199-2202. 1994.
- Neel, JV; Weiss, KM. The genetic structure of a tribal population, the Yanomama Indians. XII. Biodemographic studies. *Am J Phys Anthropol* 42(1):25-51. 1975.
- Nei, M; Roychoudhury, AK. Genetic relationship and evolution of human races. *Evolutionary Biology* - vol. 14: 1-59, New York, Appleton - Century - Crafts. Plenum Press. ed. Dobzhanski, T. *et al.* 1982.
- Neme, S.; Andrade, C.O. Quilombo: forma de resistência. Proposta histórico-arqueológica. In: *Insurreição Negra e Justiça* (eds. G. HUBER e F.B. DE SOUZA). Rio de Janeiro, OAB, 1987.
- Noble, E. P. : The D2 dopamine receptor gene: a review of association studies in alcoholism and phenotypes. *Alcohol* 16: 33-45, 1998.
- Novick, GE; Novick, CC; Yunis, J; Yunis, E; Martinez, K; Duncan, GG; Troup, GM; Deininger, PL; Stoneking, M; Batzer, MA; Herrera, RJ. Polymorphic human-specific *Alu* insertions as markers for human identification. *Electrophoresis* 16: 1596-601. 1995.
- Oliveira, S.F.; Ribeiro, G.G.B.L.; Ferreira, L.B.; Klautau-Guimarães, M.N.; Simões, A.L. (2005). History reconstruction of afro-derived isolated Brazilian populations: the contrast among female and male genetic contribution. Espanha. *Diversidad Biológica y Salud Humana*.
- Parra, F.C.; Marcini, A.; Akey, J.; Martinson, J.; Batzer, M.A.; Cooper, R.; Forrester, T.; Allison, D.B.; Deka, R.; Ferrell, R.E. & Shriver, M.D. Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet*, 63:1839-1851, 1998.
- Pedrosa, M.A.F. & Oliveira, S.F. (2004). *Mixtura étnica en poblaciones afrodescendientes semi-aisladas de Brasil*. Caracas. Anais do VIII Congresso de la Asociación Latinoamericana de Antropología Biológica.

- Pedrosa, M.A.F. Composição . *Composição Genética de Quatro Populações Remanescentes de Quilombo do Brasil com Base em Microssatélites e Marcadores de Ancestralidade*. 2006. Dissertação (Mestrado em Biologia Molecular). Instituto de Ciências Biológicas – Universidade de Brasília, Brasília – DF.
- Pfaff, C. L.; Parra, E. J.; Bonilla, C.; Hiester, K.; Mckeigue, P. M.; Kamboh, M. I.; Hutchinson, R. G.; Ferrell, R. E.; Boerwinkle, E.; Shriver, M. D. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet*, 68:198-207, 2001.
- Ramsay, M., Colman, M.-A., Stevens, G., Zwane, E., Kromberg, J., Farrall, M., and Jenkins, T. (1992). The tyrosinase-positive oculocutaneous albinism *locus* maps to chromosome 15qll.2-q12. *Am. J. Hum. Genet.* 51: 879-884.
- Raymond, M. and Rousset, F., Genepop Version 3.3: A Population Genetics Software for Exact Tests and Ecumenicism, *J. Hered.*, 1995, vol. 86, pp. 248–249.
- Ribeiro, G.G.B.L. (2005). *Contribuição genética masculina, baseada em marcadores Y-específicos na constituição de quatro comunidades afrodescendentes brasileiras*. Dissertação de Mestrado. Programa de Pós-Graduação em Biologia Animal. Universidade de Brasília.
- Ribeiro-Dos-Santos, A.K.; Pereira, J.M.; Lobato, M.R.L.; Carvalho, B.M.; Guerreiro, J.F.; Batista-Dos-Santos, S.E. (2002). Dissimilarities in the process of formation of Curiau, a semi-isolated Afro-brazilian population of the amazon Region. *Am J Hum Biol* 14:440-447.
- Rocha Pombo, 1935 In: Teixeira, Marco Antônio Domingues. *Quilombolas de Santo Antônio do Guaporé: conflitos nas relações socioambientais em área de reserva ecológica*. Porto Velho, Work-paper/UNIR. 2006.
- Rohlf J (2001) *Ntsys: numerical taxonomy and multivariate analysis system version 2.1*. Exeter Software, Setauket.
- Saitou, N. A genetic affinity analysis of human populations. *Hum. Evol.* 10(1): 17-33. 1995.

- Salzano, FM; Callegari-Jacques, SM. South American Indians: a case study in evolution. *Research Monographs on Human Population Biology*. Number 6. Clarendon Press. Oxford. 1988.
- Salzano, FM; Freire-Maia, N. (1967) *Populações Brasileiras*. Companhia Editora Nacional. São Paulo, SP.
- Schneider S, Kueffer JM, Roessli D, Excoffier L (1996) Arlequin (ver. 1.0): a software environment for the analysis of population genetics data. *Genetics and Biometry Lab, University of Geneva, Geneva, Switzerland*.
- Schneider, H.; Guerreiro, J.F.; Santos, S.E.B.; Weimer, T.A.; Schneider, M.P.C.; Salzano, F.M. (1987) Isolate breakdown in Amazonian – The blacks of the Trombeta river. *Rev Bras Genet* 10:565-574.
- Shriver, M.D.; Smith, M.W.; Jin, L.; Marcini, A.; Akey, J.M.; Deka, R. & Ferrell, R.E. Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet*, 60:957-964, 1997.
- Silva, W.S. (2007) Polimorfismos de DNA no complexo de genes da Beta-Globina em indivíduos normais e portadores de hemoglobinopatias do Recôncavo Baiano. Tese de doutorado. Programa de Pós-Graduação em Patologia Molecular. Universidade de Brasília. Brasília – DF
- Smithies, O. Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *Biochem. J.* 61: 629-41. 1955.
- Souza, I.R. & Culp, L. (2005). Valongo, genetic studies on an isolated Afro-Brazilian community. *Genet Mol Biol.* 28(3):402-406.
- Souza, S.M.B. *Estrutura Genética de uma Comunidade Afro-Brasileira, São Gonçalo (BA)*. 2001. Dissertação (Mestrado em Genética de Populações) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (USP), 2001.
- Sparkes, R. S.; Zollman, S.; Klisak, I.; Kirchgessner, T. G.; Komaromy, M. C.; Mohandas, T.; Schotz, M. C.; Lusis, A. J. Human genes involved in lipolysis of plasma lipoproteins: mapping of *loci* for lipoprotein lipase to 8p22 and hepatic lipase to 15q21. *Genomics* 1: 138-144, 1987.

- Tang, H.; Choudhry, S.; Mei, R.; Morgan, M.; Rodriguez-cintron, W.; Burchard, E.G.; Risch, N.J. (2007) *The American Journal of Human Genetics*. v.81
- Teixeira, Marco Antônio Domingues. *Campesinato negro de Santo Antônio do Guaporé, identidade, territorialidade e sustentabilidade*. Tese de doutorado, NAEA/UFPA, Belém, 2004.
- Teixeira, Marco Antônio Domingues. *Dos campos d'ouro à cidade das ruínas, grandeza e decadência do colonialismo português no Vale do Guaporé, séculos XVIII e XIX*. Dissertação de Mestrado, FFLCH/Mestrado em História/UFPE, Recife, 1997.
- Thompson, EA. Fission models of population variability. *Genetics* 93: 479 - 95. 1979.
- Tishkoff, AS; Ruano, G; Kidd, JR; Kidd, KK. Distribution and frequency of a polymorphic Alu insertion at the plasminogen activator locus in humans. *Hum. Genet.* (1996) 97: 759 – 764.
- Tournamille, C.; Colin, Y.; Cartron, J. P. & Le Van Kim C. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet*, 10:224-228, 1995.
- Vallinoto, I.M.V.C; Antonio Carlos Rosário Vallinoto, A.C.R.; Valente C.M.D.; Guerreiro J.F. Allele frequency distributions of six hypervariable loci (D1S80, APOB, D4S43, vW1, F13A and DYS19) in two African-Brazilian communities from the Amazon region. *Genet. Mol. Biol.* Vol.26, no.3. São Paulo – 2003
- Vila Real, R.N.S. *Cultura e Currículo: Um estudo da escola Kalunga*. 1996. *Dissertação (Mestrado)* - Universidade Federal de Goiás, Goiânia, 1996.
- Wang, S.; Ray, N.; Rojas, W.; Parra, M.V.; Bedoya, G.; Gallo, C.; Poletti, G.; Mazzotti, G.; Hill, K.; Hurtado, A.M.; Camrena, B.; Nicolini, H.; Klitz, W.; Barrantes, R.; Molina, J.A.; Freimer, N.B.; Bortolini, M.C.; Salzano, F.M.; Petzl-Erler, M.L.; Tsuneto, L.T.; Dipierri, J.E.; Aro, E.L.; Bailliet, G.; Bianchi, N.O.; Llop, E.; Rothhammer, F.; Excoffier, L.; Ruiz-Linares, A (2008) Geographic Patterns of Genome Admixture in Latin American Mestizos. *PLoS Genet* 4(3):e1000037.doi:10.1371/journal.pgen.1000037

Weber, J.L.; David, D.; Heil, J.; Fan, Y.; Zhao, C.; Marth, G. (2002). Human Diallelic Insertion/Deletion Polymorphisms. *Am J Hum Genet* 71:854–862.

Weir, B. S.; Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6):1358-1370, 1984.

Referências da Internet

Associação Brasileira de Antropologia – ABA:

<http://www.abant.org.br/> - Acessado em julho de 2007.

CODEPLAN:

<http://www.codeplan.df.gov.br> - Acessado em janeiro de 2008.

Funai:

<http://www.funai.gov.br> - Acessado em maio de 2006.

Fundação Cultural Palmares:

<http://www.palmares.gov.br> - Acessado em dezembro de 2007.

Instituto Brasileiro de Geografia e Estatística – PNAD 2007:

<http://www.ibge.gov.br/home/estatística/população> - Acessado em abril de 2008.

Anexo 1: Protocolos dos Procedimentos Laboratoriais

Extração de DNA

O DNA genômico da população de SAG foi extraído utilizando o kit comercial blood genomicPrep Mini Spin Kit da GE Healthcare, de acordo com o protocolo do fabricante:

Soluções:

- Proteinase K: diluir em 3mL de H₂O miliQ autoclavada.
- *Wash Buffer*: adicionar 120mL de etanol.
- *Elution Buffer*: aquecer a 70°C em banho-maria.

Procedimento:

- Colocar 20uL de proteinase K nos microtubos e, em seguida, 50µL de sangue total ou fração de leucócitos
- Adicionar 400uL de solução de lise e vortexar por 15seg. Incubar em temperatura ambiente por 10min, vortexando eventualmente durante esse tempo.
- Centrifugar rapidamente para que todo o conteúdo vá para o fundo do tubo
- Colocar as colunas nos tubos de coleta previamente identificados. Colocar o produto da lise nas colunas.
- Centrifugar a 11000G por 1min
- Descartar o líquido dos tubos de coleta e recolocar neles as colunas
- Adicionar 500µL de solução de lise à coluna
- Centrifugar a 11000G por 1min.
- Descartar o líquido dos tubos de coleta e recolocar neles as colunas
- Adicionar 500µL de *Wash Buffer* às colunas e centrifugar a 11000G por 3min
- Descartar os tubos de coleta com seus conteúdos.
- Transferir as colunas para os microtubos identificados
- Adicionar 200µL de *Elution Buffer* pré-aquecido a 70°C às colunas
- Incubar a temperatura ambiente por 1min
- Centrifugar a 11000G por 1min

Reação em Cadeia da Polimerase

Reagentes e soluções:

- DNA genômico a 20ng/μL
- Tampão de reação (Tris/HCl 0,2 M pH 8,5; KCl 0,5 M; MgCl₂ 0,02 M)
- Solução de dNTP a 20μM
- Taq DNA Polimerase
- Solução de iniciadores a 10 μM, cada

Procedimentos:

- Descongelar e aliquotar as amostras de DNA genômico em microtubos de polipropileno de 0,5mL identificados com o número da amostra, marcador analisado e número de registro da PCR (de uso interno do laboratório).
- Preparar uma mistura de reagentes da PCR proporcional ao número de amostras em análise
- Distribuir a mistura de reagentes nos microtubos contendo o DNA aliquotado
- Adicionar óleo mineral sobre o líquido
- Levar os microtubos aos termocicladores com o programa adequado ao sistema em análise.

Digestão enzimática

Os marcadores do tipo SNP foram analisados por PCR/RFLP, que consiste em amplificar fragmentos de interesse e submetê-los a digestão enzimática para detectar a presença de uma mutação específica.

Reagentes e soluções:

- Tampão da enzima de restrição
- Enzima de restrição
- H₂O miliQ autoclavada

Procedimentos:

- Descongelar os reagentes e preparar uma mistura adequada à análise do sistema em questão

- Aliquotar em microtubos de polipropileno, 5,0 μ L de produto da PCR de cada amostra;
- Distribuir a mistura de reagentes nos tubos contendo produto de PCR
- Incubar, por um tempo suficiente (4h a 16h, a depender do sistema), o tubos em banho-maria em temperatura adequada ao funcionamento da enzima
- Ao final da reação, submeter os microtubos a choque térmico

Eletroforese

Preparo do gel de Poliacrilamida

Soluções:

- Tampão TBE 10X pH 8,0 (Tris/HCl 0,89 M, EDTA 0,08 M, Ácido Bórico 0,89M)
- Solução acrilamida 29% e bis-acrilamida 1%
- Glicerol
- TEMED (N,N,N',N', tetrametiletenodiamina)
- Persulfato de potássio ou amônia a 0,1%

Preparo:

- Limpar as placas de vidro com etanol a 70%
- Montar o suporte para o gel separando as placas de vidro com espaçadores e prendendo o conjunto com garras metálicas
- Misturar, em um béquer, os volumes de cada reagente (tabela 1)
- Verter a mistura de reagentes entre as placas de vidro
- Colocar um pente entre as placas na extremidade sem espaçadores

Tabela 1. Reagentes usados no preparo da solução gel de poliacrilamida.

Reagentes (mL)	[] do gel	
	6%	10%
Acrilamida + Bisacrilamida	4,0	6,6
Glicerol	1,4	1,4
H ₂ O destilada	12,4	9,6
TBE 10X	2,0	2,0
Catalisadores (μL)		
TEMED	15,0	15,0
APS	300,0	300,0

Preparo das amostras para eletroforese

- Adicionar a 10 μ L de produto de PCR ou restrição, tampão de corrida contendo o corante bromofenol.

Procedimentos para eletroforese:

- Montar a cuba de eletroforese colocando tampão TBE 1x nos compartimento inferior e superior
- Prender o conjunto de placas (montado previamente) à cuba
- Retirar o pente e lavar os poços com TBE 1x com o auxílio de uma seringa
- Aplicar as amostras preparadas nos poços
- Submeter as placas a eletroforese a 150 ou 200V por 1h30min a 3h conforme o sistema em análise,

Coloração**Soluções:**

- Solução Fixadora (750ml de H₂O destilada; 144ml de etanol; 6,0ml de ácido acético)
- Solução Reveladora (22,5g de NaOH; H₂O destilada vsp 1L)
- Solução de nitrato de prata a 10%
- Formaldeído

Procedimentos:

- Colocar o gel em um recipiente contendo solução fixadora suficiente para cobri-lo
- Adicionar 1mL de solução de nitrato de prata e agitar por 5min
- Descartar o líquido e lavar rapidamente com H₂O destilada
- Colocar 200mL de solução fixadora aquecida acrescida de 1mL de formaldeído
- Agitar até que as bandas se tornem visíveis

Leitura dos géis

A determinação dos alelos presentes foi feita pela comparação visual com uma banda de tamanho conhecido (como produto de PCR) ou de marcador de peso molecular.