

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA**

Claudia Caixeta Franco Andrade

**Estrutura genética e Desequilíbrio de Ligação em
Africanos, Ameríndios e Remanescentes de Quilombos
Brasileiros estimados por novos STRs-X**

**RIBEIRÃO PRETO
2013**

Claudia Caixeta Franco Andrade

**Estrutura genética e Desequilíbrio de Ligação em
Africanos, Ameríndios e Remanescentes de Quilombos
Brasileiros estimados por novos STRs-X**

Tese apresentada à Faculdade de Medicina de
Ribeirão Preto da Universidade de São Paulo,
como requisito parcial para a obtenção do
título de Doutora em Ciências – Área de
concentração: Genética.

Orientador: Prof. Dr. Aguinaldo Luiz Simões

**RIBEIRÃO PRETO
2013**

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

Andrade, Claudia Caixeta Franco.

Estrutura genética e Desequilíbrio de Ligação em Africanos, Ameríndios e Remanescentes de Quilombos Brasileiros estimados por novos STRs-X. / Claudia Caixeta Franco Andrade; orientador Aguinaldo Luiz Simões. – Ribeirão Preto, 2013.

118p f. il. 30cm.

Tese (Doutorado – Programa de Pós-Graduação em Genética.) – Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

1. Ameríndios, 2. Quilombos, 3. População Africana, 4. STRs-X,
5. Genética populacional.

FOLHA DE APROVAÇÃO

Claudia Caixeta Franco Andrade

Estrutura genética e Desequilíbrio de Ligação em Africanos, Ameríndios e Remanescentes de Quilombos Brasileiros estimados por novos STRs-X

Tese apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como requisito parcial para a obtenção do título de Doutora em Ciências – Área de concentração: Genética.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

“Porque o SENHOR dá a sabedoria; da sua boca é
que vem o conhecimento e o entendimento.”

Provérbios 2:6

"Não sei se estou perto ou longe demais, se peguei o
rumo certo ou errado. Sei apenas que sigo em
frente, vivendo dias iguais de forma diferente. Já
não caminho mais sozinho, levo comigo cada
recordação, cada vivência, cada lição. E, mesmo
que tudo não ande da forma que eu gostaria, saber
que já não sou o mesmo de ontem me faz perceber
que valeu a pena."

Professor Galvão

À minha família, **Dalcio, Mary e Raquel**,
por serem meu amparo, meu estímulo
e a minha maior alegria!

Dedico

AGRADECIMENTOS

À **Deus**, pelo seu amor incondicional, sua fidelidade, por ser meu amparo e sempre me dar condições de prosseguir em meio as dificuldades.

Ao Prof. Dr. **Aguinaldo Luiz Simões**, pela oportunidade deste trabalho, pelos muitos conhecimentos divididos, pela confiança, compreensão e paciência em todos os momentos.

Ao Prof. Dr. **Enrique Medina-Acosta**, pela possibilidade do trabalho em colaboração, pela ajuda em muitos momentos, pelos conhecimentos divididos, pela alegria e amizade.

Aos **professores membros da banca examinadora**, por se mostrarem disponíveis para avaliar este trabalho.

Aos **funcionários do Departamento de Genética** pela acolhida e ajuda. Em especial à **Susie** e à **Sílvia**, secretárias do departamento, pela paciência, atenção, carinho e apoio.

À **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)**, pela bolsa concedida para o desenvolvimento desta tese.

À **Ana Lúcia**, e **Maria do Carmo**, por serem ajudadoras em todos os momentos, pelas palavras de carinho, de ânimo e, principalmente, pela ajuda com a extração de DNA.

À **Cláudia Wiesel**, pela grande amizade, pelas conversas sérias e também pelas descontraídas, pelo apoio em todos os momentos, caronas e por toda a ajuda e conhecimento compartilhado durante esses quatro anos. “Tudo isso é impagável!”

À todos os amigos do Bloco B: **Edna, Sabrina, Fernanda, Marcela, Natália, Juliana Massaro, Rosana, Bete, Geusa, Rômulo** e **Prof. Dr. Ana Lílian**, pela amizade, companhia no dia a dia, pelos momentos alegres no café e por me proporcionarem um excelente ambiente de trabalho.

Aos amigos do NUDIM (Núcleo de Diagnóstico e Investigação Molecular): **Laís, Antônio, Filipe, Graziela, Fabrício, Thaís, Priscila, Milena, Thiago** e **Maria**, por me receberem como um de vocês, pela alegria, ajuda em diversos momentos, companhia nos fins de semana de trabalho, pelos momentos de conversa durante os cafés e pelas inúmeras risadas. É sempre bom trabalhar com vocês!

Às minhas amigas de república: **Vanessa, Glauce** e **Mayra**, por me proporcionarem um ambiente “de família”, pela amizade gostosa, pelos muitos momentos de alegria, pelas conversas, conselhos, puxões de orelha, por se preocuparem comigo e me amparem nos momentos mais difíceis.

Aos meus grandes amigos e irmãos em Cristo: **Cristiana, Gilvan, Giselle, Rafael, Amani, Rachel, Terezinha (vó), Lâmia, Bruna, Victor, Renata, Vagner, Tamara, Ramati, Marília, Célia, Rosana, Ari, Tiago, Lucas, Liliane, Sandra, Izenildo, Janaina, Ilton, Fábio, Luciana** por serem minha família nesta cidade, pelas orações, pelo socorro e conselhos em diversos momentos, pelos almoços de domingo, saídas em amigos, festas e pela imensa alegria que os momentos com vocês sempre me proporcionam. Vou sentir muita saudade!

Aos meus **amigos de Campos** (são tanto que é melhor nem citar!): por mesmo na distância serem amigos verdadeiros, acreditarem em mim, sempre me incentivarem, desejarem minha felicidade e por tornarem minhas idas a Campos sempre maravilhosas!

Aos meus pais, **Dalcio e Mary**, e minha irmã **Raquel**, por nunca terem poupado esforços pra me ajudar, por sempre acreditarem em mim, me “carregarem” nos piores momentos, por serem meu referencial e por juntos, construirmos a melhor família que eu poderia ter. Amo vocês!

À **minha família**, por sempre me incentivar, apoiar, pelas orações, conselhos e momentos de descontração e alegria. À minha prima **Lilian**, por ser como uma irmã, por ter cuidado de mim, pelas caronas, conversas, amizade e confiança que se solidificaram quando moramos juntas. Ribeirão sem você não é a mesma...

RESUMO

Andrade, CCF. **Estrutura genética e Desequilíbrio de Ligação em Africanos, Ameríndios e Remanescentes de Quilombos Brasileiros estimados por novos STRs-X**. 2013. 118p. Tese (Doutorado) – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

O cromossomo X possui características que o tornam um bom marcador em estudos de genética populacional. A formação da população brasileira é resultado de cinco séculos de mistura étnica entre populações de três diferentes populações: europeus, africanos e ameríndios. Sendo assim, a população brasileira atual é considerada tri-híbrida embora as proporções dos três grandes componentes étnicos variem consideravelmente conforme a região geográfica. O objetivo deste trabalho é obter informações sobre a estrutura genética de populações urbanas brasileiras, indígena, remanescentes de quilombos e uma amostra da população africana (Congo). Foram utilizados 20 STRs localizados em três regiões do cromossomo X: Xp21, Xp-q11.1, Xq28. O total de alelos obtidos foi 169. Todas as três regiões apresentaram alelos privados, totalizando 26. A população ameríndia, *Tikúna*, foi a que apresentou a menor diversidade genética. Além disso, quando comparada com as outras populações, *Tikúna* foi a que mais se diferenciou. Era esperado que *Tikúna* fosse a população mais divergente e com menor diversidade, pois esta é uma população pequena e isolada que sofre as consequências da ocorrência de deriva genética. Nas mulheres, as populações mais semelhantes foram Sítio Velho-Teresina, Sítio Velho-Ribeirão Preto e Ribeirão Preto-Teresina, pois tiveram valores de F_{ST} não significativos. Nos homens, apenas Mimbó-Sítio Velho são as populações mais semelhantes (F_{ST} não significativo). Mimbó e Sítio Velho são os dois remanescentes de quilombos utilizados neste estudo e o histórico de formação destas populações é semelhante, o que justifica a proximidade genética entre elas. Todas as populações foram agrupadas em três clusters, tanto no grupo dos homens quanto no das mulheres. O primeiro *cluster* é formado, em sua grande maioria, por indivíduos da população Africana, Mimbó e Sítio Velho, representando o componente africano na população brasileira. O segundo *cluster* foi formado por mais de 90% da população *Tikúna*, sendo assim indicativo do componente ameríndio. O terceiro *cluster* agrupou as populações urbanas, Ribeirão Preto e *Tikúna*, que possuem um maior componente europeu na sua formação. DL ocorreu entre os marcadores de todos os haploblocos (DMD, PC e HEMA). HEMA foi o haplobloco com maior DL. Semelhanças e diferenças entre as populações foram encontradas de acordo com o esperado pelo histórico de formação de cada uma delas. Os três componentes genéticos (africano, ameríndio e europeu) da população brasileira foram claramente identificados nas amostras analisadas.

ABSTRACT

Andrade, CCF. **Genetic Structure and Linkage Disequilibrium in African, Amerindian and Quilombo Remnants of Brazilian estimated by new X-STRs.** 2013. 118p. Thesis (Doctoral) – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

The X chromosome has characteristics that make it a good marker for studies of population genetics. Brazilian population results from five centuries of admixture of European, Africans and Amerindians. Thus, the current Brazilian population is considered trihybrid although the proportions of the three main ethnic components vary considerably according to the geographical region. The aim of this work was to obtain information about the genetic structure of Brazilian urban populations, Amerindians, quilombo remnants and a sample of African (Congo). We used 20 STRs located in three regions of the X chromosome: Xp21, Xp-q11.1 and Xq28. The total of alleles obtained was 169. All three chromosomal regions have private alleles, totaling 26. The Amerindian population (*Tikúna*) had lower genetic diversity and higher F_{ST} values, when compared to other populations. This is expected given the genetic isolation and small population size, which make them most sensitive to the effects of genetic drift. In women, by pairs of populations Sítio Velho-Teresina, Sítio Velho-Ribeirão Preto and Ribeirão Preto-Teresina, were similar (F_{ST} values not significantly different). In men, only the pair Mimbó-Sítio Velho rendered no significantly different the individuals analyzed were grouped into three clusters, even when considered separately for men and women. One cluster was termed 'African' to be formed largely by individuals from Congo, quilombo remnants and Sítio Velho. Another cluster termed 'Amerind' was formed by >90% of *Tikúna* population. In the third one, were grouped urban populations of Ribeirão Preto and Teresina. In this cluster are observed components African, Amerindian, and European with predominance of the latter. DL occurred between markers of all blocks (DMD, PC e HEMA). HEMA had the highest DL. There were differences between DL in men and women, which may have been caused by sampling problems. The observed similarities and differences between clusters and populations subsets match expectation from the population histories. The results presented bring new elements of analysis to the structure of the Brazilian population.

LISTA DE TABELAS

Tabela 1. Informações sobre os 20 loci utilizados. N°: número de ordem dado para os loci dentro de cada haplobloco. cM: Distância genética em centimorgans. pb: número de pares de bases do amplicon. A distância genética foi obtida em: http://compgen.rutgers.edu/old/map-interpolator/	29
Tabela 2. Temperatura de pareamento e concentração dos reagentes utilizados em cada uma das três PCRs multiplex	31
Tabela 3. Frequências alélicas dos 20 STRs-X em todas as populações (Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e <i>Tikúna</i>). N- número de cromossomos X analisados em cada locus.....	39
Tabela 4. Teste exato de Fisher para a verificação do Equilíbrio de Hardy-Weinberg nas populações de mulheres. * Locus fora do Equilíbrio de Hardy-Weinberg depois da correção de Bonferroni ($\alpha_{Bonf} < 0,0025$).....	50
Tabela 5. Diferenciação Gênica entre as amostras populacionais (mulheres) comparadas par a par. AF: Africana; MB: Mimbó; SV: Sítio Velho; TR: Teresina; RP: Ribeirão Preto; TK: <i>Tikúna</i> . Valores significativos ($p < 0,05$) em negrito sublinhado. *Valores significativos após a correção de Bonferrone ($\alpha_{Bonf} = 0,00016$).....	52
Tabela 6. Diferenciação Gênica, entre as amostras populacionais (homens) comparadas par a par. AF: Africana; MB: Mimbó; SV: Sítio Velho; TR: Teresina; RP: Ribeirão Preto; TK: <i>Tikúna</i> . Valores significativos ($p < 0,05$) em negrito sublinhado. *Valores significativos após a correção de Bonferrone ($\alpha_{Bonf} = 0,00016$).....	54
Tabela 7. Diferenciação Genotípica entre as mulheres de todas as amostras populacionais comparadas par a par. AF: Africana; MB: Mimbó; SV: Sítio Velho; TR: Teresina; RP: Ribeirão Preto; TK: <i>Tikúna</i> . Valores significativos ($p < 0,05$) em negrito sublinhado.*Valores significativos após a correção de Bonferrone ($\alpha_{Bonf} = 0,00016$).	56
Tabela 8. Heterozigose média observada e esperada nas mulheres de todas as populações.	57
Tabela 9. Valores de <i>FST</i> em todos os loci nos homens (acima da diagonal) e nas mulheres (abaixo da diagonal). **Valores não significativos depois da correção de Bonferroni ($\alpha_{Bonf} < 0,0033$).....	57
Tabela 10. Proporção da composição genética da amostra de mulheres e homens de todas as populações divididas em três clusters	58
Tabela 11. Diversidade haplotípica dos três blocos em todas as populações).....	60
Tabela 12. Haplótipos do bloco DMD encontrados nas amostras populacionais de Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e <i>Tikúna</i>	61
Tabela 13. Haplótipos do bloco PC encontrados nas amostras populacionais de Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e <i>Tikúna</i>	63
Tabela 14. Haplótipos do bloco HEMA encontrados nas amostras populacionais de Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e <i>Tikúna</i>	65
Tabela 15. Proporção (em %) do Desequilíbrio de Ligação em cada bloco para todas as populações, antes de depois da correção de Bonferroni.	67

LISTA DE FIGURAS

Figura 1. Localização no cromossomo X dos lócus STRs utilizados nas populações brasileiras. Adaptado de http://www.bgbx.com.br/?pagina=idiograma	20
Figura 2. Localização dos três blocos de lócus STRs no cromossomo X que foram utilizados neste trabalho.	28
Figura 3. Eletroferograma do Heptaplex – DMD – de uma mulher.....	32
Figura 4. Eletroferograma do Hexaplex – PC – de uma mulher.....	32
Figura 5. Eletroferograma do Heptaplex – HEMA – de uma mulher.....	32
Figura 6. Proporção genética dos homens de todas as populações divididas em <i>clusters</i> (K=3). 1: Africanos; 2: Mimbó; 3: Sítio Velho; 4: Teresina; 5: <i>Tikúna</i> ; 6: Ribeirão Preto.	59
Figura 7. Proporção genética das mulheres de todas as populações divididas em <i>clusters</i> (K=3). 1: Africanos; 2: Mimbó; 3: Sítio Velho; 4: Teresina; 5: <i>Tikúna</i> ; 6: Ribeirão Preto\	59

CONTEÚDO

INTRODUÇÃO	13
Estrutura tri-híbrida da população do Brasil	13
Ancestralidade da população brasileira.....	15
Informatividade dos marcadores ligados ao Cromossomo X.....	16
Estudos populacionais com cromossomo X no Brasil.....	17
JUSTIFICATIVA	21
HIPÓTESE.....	22
OBJETIVOS.....	23
MATERIAIS E MÉTODOS.....	24
Desenho experimental	24
Descrição das amostras populacionais	24
Coleta das amostras	26
Estocagem do material e extração de DNA	27
Marcadores genéticos	28
Reação em cadeia da polimerase – PCR.....	30
Eletroforese Capilar.....	30
Leitura dos alelos	32
Análise estatística	33
RESULTADOS	38
Alelos Novos	38
Frequências Alélicas	49
Aderência ao equilíbrio de Hardy-Weinberg.....	50
Diferenciação genética das populações	51
Diversidade Gênica.....	57
Análise de Variância Molecular (AMOVA)	58
Inferência de Estrutura Populacional.....	58
Diversidade Haplotípica.....	60
Desequilíbrio de ligação entre <i>locus</i> (DL)	67
DISCUSSÃO	68
Alelos Novos	68
Frequências Alélicas	68
Aderência ao equilíbrio de Hardy-Weinberg.....	69
Diferenciação genética das populações	70
Diversidade Gênica.....	71
Análise de Variância Molecular (AMOVA)	73
Inferência de Estrutura Populacional.....	73
Diversidade Haplotípica.....	74
Desequilíbrio de ligação entre <i>locus</i> (DL)	76
CONCLUSÕES.....	78
BIBLIOGRAFIA	79
COMPROVANTE DE SUBMISSÃO	86
MANUSCRITO	87

INTRODUÇÃO

Estrutura tri-híbrida da população do Brasil

A complexidade da constituição da população brasileira se atribui à sua formação tri-híbrida (Leite *et al.* 2009): a aculturação imposta aos indígenas fez com que grandes contingentes deles fossem aos poucos sendo ‘absorvidos’ pela sociedade neo-brasileira composta inicialmente por portugueses e, posteriormente, por grandes contingentes de africanos trazidos como escravos (Salzano e Bortolini, 2002).

Nos primeiros 350 anos da colonização do Brasil, entre 1500 e 1850, chegaram aproximadamente 500.000 portugueses e quatro milhões de africanos (oriundos de diferentes regiões Africanas sub-saarianas, predominantemente de regiões hoje pertencentes a Angola, Congo e Moçambique) que dividiram espaço na nova terra com 2,4 milhões de ameríndios então existentes (Curtin, 1969; IBGE, 2000). No período seguinte, até o presente, chegaram oficialmente no Brasil mais de cinco milhões de imigrantes, a maioria portugueses (espanhóis, alemães, sírios, libaneses, japoneses e outros não ultrapassaram 30% deste total) (IBGE, 2000).

Assim, a formação da população brasileira é resultado de cinco séculos de mistura entre populações de três diferentes continentes: colonizadores europeus, escravos africanos e ameríndios. Isto caracteriza a população brasileira como uma das que possui maior diversidade (Parra *et al.*, 2003; Pena *et al.*, 2011). As migrações inter-regionais em massa durante o período colonial do Brasil promoveram a dispersão e a miscigenação desses três componentes ancestrais fundadores em diferentes proporções ao longo do país. A proporção da contribuição desses componentes étnicos é diferente entre as regiões geopolíticas do Brasil. De acordo com o censo realizados no ano de 2000 pelo IBGE, 53% dos indivíduos declararam ser brancos, 38% pardos, 6% negros e 3% às demais categorias. No entanto, esses valores variam de acordo com a região do Brasil e também com o história da sua colonização.

Populações urbanas (Ribeirão Preto e Teresina)

No estado de São Paulo as proporções de indivíduos que se autodeclararam brancos, pardos, negros entre outras categorias é diferente da média apresentada pelo IBGE, sendo 70%, 23%, 4% e 3%, respectivamente.

A população de Ribeirão Preto, no Estado de São Paulo, é altamente miscigenada. A contribuição de Europeus, Africanos e Ameríndios desta população foi inferida a partir de dados obtidos através de marcadores moleculares (Ferreira *et al.*, 2006 e Muniz *et al.*, 2008). Essa contribuição é diferentes entre os indivíduos que se consideram brancos, mulatos e negros. Nos brancos, 79% da contribuição é europeia, 14% africana e 7% de ameríndios

(Ferreira *et al.*, 2006). Nos mulatos, como era esperado, a contribuição africana aumentou para 26%, a europeia foi 62% e a ameríndia, 12%. Os negros foram caracterizados com apenas os componentes europeu (37%) e africano (63%).

A população de Teresina no estado do Piauí foi analisada por 46 marcadores de inserção-deleção (INDELS) e o resultado comparado com outras populações brasileiras e mundiais (Lopes *et al.*, 2013). A população de Teresina é mais próxima de Macapá e Belém do que das ameríndias e africanas. A contribuição europeia, africana e ameríndia foi de 60%, 21,5% e 18,5%, respectivamente (Lopes *et al.*, 2013).

População Africana – Congo

A África é uma população que possui diversidade genética alta, são conhecidos mais de 2000 grupos étnicos. Além disso, a população africana possui uma extensa estrutura genética entre as populações e valores de DLs mais baixos (Campbell *et al.*, 2008). A contribuição africana na população brasileira se deu pela migração de escravos africanos. Estes são oriundos, principalmente, da região Niger-Congo da África Equatorial (França, 2001).

Os remanescentes de quilombos no Brasil

Os quilombos surgiram na época da escravatura e embora a maior parte dos habitantes fosse composta por escravos fugidos das fazendas, essas comunidades frequentemente abrigavam mestiços (crioulos), homens livres e até índios, formando uma população com variado grau de heterogeneidade. O difícil acesso aos locais em que tais comunidades se instalaram não só decorreu da necessidade de defesa e proteção, mas principalmente por não provocar o interesse de donos da terra em tê-las de volta (Mattoso, 1982). Em consequência, existem ainda em toda a América Latina comunidades remanescentes de quilombos com algum grau de isolamento genético. No Brasil, até hoje, foram identificadas oficialmente 743 comunidades remanescentes de quilombos, 42 reconhecidas e 29 tituladas (www.palmares.gov.br).

Apesar de certo grau de heterogeneidade, como acima já referido, a predominância de indivíduos oriundos do continente africano, ou de seus descendentes diretos, faz com que estas comunidades se constituam em amostra viva da contribuição africana para a formação da população brasileira. Isto torna obrigatório o estudo detalhado das frequências de seus genes (autossômicos ou não) para quantificar a contribuição africana na formação da população brasileira.

População Ameríndia no Brasil

Os ameríndios brasileiros tem como ancestrais, populações de caçadores-coletores que alcançaram a América do Sul através da Colômbia (Keyeux *et al.*, 2002). A população

ameríndia que povoou o Brasil era bastante homogênea, no que diz respeito à sua cultura e linguística. No ano de 1500, era formada por dois milhões de indivíduos (Salzano, 1986). Apesar da população ameríndia no Brasil ter crescido aproximadamente 205% nos últimos 20 anos, ela ainda é muito reduzida em relação à população original. Atualmente existem 896,9 mil indígenas no Brasil, e foi relatado um equilíbrio entre o número de homens e mulheres nas populações ameríndias (IBGE, 2010). A diminuição da população ameríndia teve como principais causas epidemias, assassinatos e guerras, mas também a simples aculturação que provocou o ‘desaparecimento’ pela absorção de comunidades indígenas que, progressivamente, deixaram de denominar-se como tal (Cunha, 1998). As primeiras populações ameríndias a sofrer esta diminuição foram as do litoral, pois estas eram as primeiras a serem encontradas pelos europeus que chegavam ao Brasil. Posteriormente, com o avanço da colonização, os ameríndios que se encontravam próximos aos grandes rios, especialmente na Amazônia, migraram para o interior das florestas em busca de maior isolamento e conseqüente proteção (Mohrenweiser *et al.*, 1979).

Ancestralidade da população brasileira

Como visto acima, a população brasileira possui uma tríplice ancestralidade (africana, ameríndia e europeia), decorrente do seu processo de formação. Como conseqüência, a população brasileira tem alto nível de variabilidade.

Para estimar as proporções dos diferentes componentes ancestrais da população brasileira, diferentes métodos que empregam frequências gênicas e\ou genotípicas têm sido utilizados. Os resultados destas estimativas, a partir de frequências de marcadores autossômicos obtidas por alguns trabalhos, demonstraram predominância semelhante entre o componente europeu (8,83 a 87,98%) e o componente africano (3,2 a 95,4%). O componente indígena é o menor (4,6 a 27,96%) (Wiesel, 2003; Mendes Jr, 2005; Luizon, 2007; Massaro, 2009).

O maior valor do componente europeu é encontrado nas populações do sul do Brasil e componente africano, como era esperado, em remanescentes de quilombos. O componente ameríndio é sempre o menor componente, em todas as regiões. Seus maiores valores são observados nas populações do norte-nordeste

Estimativas feitas a partir de frequências gênicas de marcadores do cromossomo Y (Wiesel, 2003) mostram valores maiores do componente europeu, quando comparados àqueles obtidos a partir de marcadores autossômicos, exceto no remanescente de quilombo Gaucinha. Por outro lado, a análise a partir de marcadores mitocondriais (Mendes Jr, 2005) mostram valores mais elevados da contribuição indígena

As peculiaridades do cromossomo X implicam em uma dinâmica de transmissão dos marcadores aí localizados diferente daquela dos outros marcadores. Por esta razão, exigem uma abordagem mais detalhada.

Informatividade dos marcadores ligados ao Cromossomo X

O cromossomo X pode ser usado como ferramenta para gerar informações tanto sobre a linhagem feminina quanto sobre a linhagem masculina de uma determinada população (Santos-Lopes *et al.*, 2007). Nas mulheres, metade da informação genética existente no cromossomo X é proveniente da linhagem materna e a outra metade da linhagem paterna. Nos homens o cromossomo X (linhagem materna) está presente em apenas uma cópia, o que favorece a identificação de haplótipos. Por estar sujeito a uma menor taxa de recombinação, mutação, assim como uma menor deriva genética, o cromossomo X reflete por um número maior de gerações, a história de uma população.

O cromossomo X, na espécie humana, possui 150 milhões de pares de base e apresenta taxa de mutação e recombinação menor do que nos cromossomos autossômicos (Schaffner, 2004). Assim como o cromossomo Y, ele divergiu de um único cromossomo autossômico há aproximadamente 300 milhões de anos (Lahn e Page, 1999) e ainda mantém algumas características autossômicas, principalmente a ocorrência de um par nas mulheres. Isto permite a recombinação de toda a extensão do cromossomo, tal como entre os autossomos.

A ocorrência assimétrica do cromossomo X em homens e mulheres determina condições biológicas que levam à ausência de recombinação neste cromossomo no sexo masculino, exceto em uma pequena parte – a região pseudoautossômica, constituída por 2.920 Mb e está localizada na extremidade do braço curto do cromossomo X (Mangs & Moris, 2007). Esta restrição à recombinação também leva à menor taxa de diversidade, deriva genética mais rápida e a estimativas menores do número efetivo populacional (N_e). Estas estimativas, calculadas com base em marcadores ligados a este cromossomo corresponde a $\frac{3}{4}$ do mesmo parâmetro calculado a partir de marcadores autossômicos (Schaffner, 2004). A cada geração, a recombinação ocorre em apenas 75% dos cromossomos X, o que, aliado ao baixo índice de mutação (Li *et al.*, 2002), confere a este cromossomo uma menor diversidade, quando comparado aos autossomos. Apesar disso, a diversidade do cromossomo X ainda é o dobro da do cromossomo Y (Schaffner, 2004).

A ausência de recombinação no sexo masculino facilita o reconhecimento do haplótipo que o pai obrigatoriamente transmitirá a todos filhos. Daí decorre também a necessidade de adaptar fórmulas para cálculos estatísticos à situação haplo-diplóide, diferente do que ocorre com os autossomos. Com o advento da análise molecular, esta característica peculiar do cromossomo X passou a ser mais explorada evidenciando informações populacionais diferentes das obtidas por marcadores genéticos autossômicos. O primeiro trabalho com informações sobre as sequências do cromossomo X surgiu em 1997 (Schaffner, 2004).

As características que tornam o cromossomo X um bom marcador em estudos de genética populacional são: taxa de mutação e recombinação menores do que de outros cromossomos;

facilidade na determinação dos haplótipos; deriva genética mais rápida; e consequentemente, uma menor diversidade (Leite *et al.*, 2009; Amorim *et al.*, 2010).

Como a recombinação só ocorre na meiose feminina (Pereira *et al.*, 2007; Zarrabeitia *et al.*, 2009) os níveis de DL ligado ao cromossomo X tendem a ser correspondentemente maiores do que os valores relativos a segmentos autossômicos de igual tamanho. Não só a intensidade do DL, mas sua extensão também aumenta.

Comparado aos autossomos, o cromossomo X reflete os efeitos de eventos evolutivos (por exemplo, diminuição de sua diversidade por um efeito 'gargalo') durante um número maior de gerações, porque as menores taxas de mutação e recombinação (Li *et al.*, 2002) demandarão mais tempo para recuperar a diversidade.

Marcadores genéticos ligados ao X também são mais sensíveis aos efeitos da deriva quando comparados aos marcadores autossômicos, pois para uma mesma situação, o N_e baseado nestes últimos é sempre maior. Por isso, também a diferenciação entre populações é melhor detectada pelo uso de marcadores ligados ao X (Schaffner, 2004)

Em conclusão, a informação genética que se obtém de marcadores ligados ao cromossomo X pode fornecer estatísticas distintas daquelas obtidas a partir de marcadores autossômicos, mitocondriais ou ligados ao cromossomo Y. Às vezes, tais resultados são concordantes entre si. Outras, se não chegam a ser conflitantes, contam aspectos da história demográfica que não são observados de outra maneira e, por isso, se tornam necessários à completa descrição da história demográfica de uma população. Daí decorre a necessidade de seu estudo para o entendimento da complexa estrutura da população brasileira, cujo histórico deixa clara a contribuição de contingentes populacionais indígenas, africanos e europeus.

Estudos populacionais com cromossomo X no Brasil

Uma amostra da população urbana de Santa Catarina e remanescentes de quilombos (Mimbó, Sítio Velho e Gaucinha) foram analisadas, através de STRs do cromossomo X (Wiesel, 2003). A maior diferenciação foi observada entre os três remanescentes de quilombos, enquanto a menor diferenciação surgiu quando a população de Teresina foi comparada com os três quilombos agrupados.

Amostras das populações da Argentina, Brasil (São Paulo, Rio de Janeiro, Paraná e Mato Grosso), Costa Rica, Portugal e Espanha foram analisados com 10 STRs-X (Gusmão *et al.*, 2009). A diversidade genica em todas as populações foi maior que 0,56. A população do Rio de Janeiro apresentou diferenças significativas quando comparada com São Paulo, Paraná e Mato Grosso, assim como São Paulo se diferenciou do Paraná. A maior diferenciação encontrada foi entre Paraná e Rio de Janeiro, e as populações mais semelhantes foram São Paulo e Mato Grosso do Sul.

Outro trabalho com STRs do cromossomo X em populações ameríndias (Caingang, Guarani, Wai Wai, Xavante, Gavião, Zoró, Suruí, Guarany Mbya, Pilagá, Toba Formosa, Toba Chaco, Wichi Chaco e Wichi Formosa) da América do Sul e população urbana do sul do Brasil (Rio Grande do Sul) foi realizado por Leite *et al.* (2009). Foi observada uma menor diversidade genética associada a diferenças entre as populações estudadas, o que reflete uma estrutura populacional que pode ser responsável pelo maior valor de DL nas populações ameríndias. Puderam concluir que diferentes histórias populacionais no Brasil e na América do Sul têm contribuído significativamente para a diferença entre os valores e extensão de DL encontrado. Além disso, foi observado que as populações quando agrupadas pelo geográfico são mais semelhantes do que quando agrupadas por características linguísticas.

A população de Santa Catarina, analisadas por 10 STRs do cromossomo X (Cainé *et al.*, 2010), exibiu componente europeu em maior proporção do que o africano e ameríndio. Além disso mostrou-se mais semelhante às populações da Europa (Portugal, Península Ibérica, Alemanha e Itália).

A população de Alagoas e do Rio de Janeiro foram analisadas por cinco STRs do cromossomo X (Ferreira da Silva *et al.*, 2010). Nestes cinco STRs foram encontrados dois blocos haplotípicos que não apresentaram ligação entre si. Estas duas populações, comparadas com uma amostra da população da Alemanha e outra de Gana, mostraram semelhanças apenas entre as duas brasileiras e com a amostra europeia.

Outras amostras populações do sudeste brasileiro, São Paulo, Rio de Janeiro, Belo Horizonte e Vitória, apresentaram uma diversidade entre 68% e 83% nos 10 STRs-X analisados, e estruturação quando comparadas entre si (Martins *et al.*, 2010a). Estas são mais similares às populações europeias (Portugal e Espanha), em seguida às populações da América Latina (Argentina, Costa Rica e Colômbia), e são mais divergentes da população da África (Uganda). Foi encontrado desequilíbrio de ligação em apenas um par de loci na população de Belo Horizonte, o que foi relacionado a problemas na amostragem.

Em um terceiro trabalho a população de São Paulo foi estudada com outros cinco STRs-X (Martins *et al.*, 2010b) e se mostrou mais próxima geneticamente a grupos de brasileiros com descendência europeia e africana do que a grupos com descendência asiática. Foi encontrado um par de loci em DL, que pode ser resultado da miscigenação existente na população de São Paulo ou até mesmo de efeitos de amostragem.

Seis populações nativas americanas, Wayuu, Ingano, Kogi, Zenu, *Tikúna* (Colômbia) e Cree, foram caracterizadas por Wang *et al.* (2010) utilizando sete STRs localizados no cromossomo Xq13. As populações analisadas apresentaram baixa diversidade e alto DL, o que era esperado para populações que passaram pelo efeito fundador e em seguida sofreram deriva genética. Em relação a diferenciação genética, as populações de Nativos Americanos foram

mais semelhantes entre si. A maior divergência foi observada entre as os Nativos Americanos e as populações da Europa juntamente com as do Leste Asiático.

Amostras de populações de 16 estados brasileiros foram analisadas através de 12 STRs-X (Ribeiro-Rodrigues *et al.*, 2011). Estes representavam as cinco regiões geopolíticas do Brasil. A diversidade encontrada variou de 0,59 no estado do Amazonas a 0,96 no Pará. Apenas cinco pares de locus estão em desequilíbrio de ligação. Foi observado que existem diferenças genéticas significativas ($F_{ST}=0.00261$; $p=0.0000$) entre as cinco regiões geopolíticas do país, que em parte, é consequência da diferente contribuição de cada um destes três componentes pelo Brasil. Estas 16 populações são significativamente diferentes entre si. Além disso, as populações brasileiras foram mais semelhantes às populações da América Latina do que às da Europa, África e nativos americanos.

O estado do Paraná também foi analisado através dos mesmos 10 STRs do cromossomo X (Kobachuk *et al.*, 2012) e apresentou diferenciação significativa com várias outras populações brasileiras (São Paulo, Rio de Janeiro, Minas Gerais, Mato Grosso, Santa Catarina, Rio Grande do Sul, Pará e Amazônia). Apenas um par de locus, na população do Paraná, apresentou desequilíbrio de ligação, que foi considerado espúrio (Kobachuk *et al.*, 2012).

Uma amostra da população urbana do Rio Grande do sul foi analisada com 14 STRs-X (Penna *et al.*, 2012) e os resultados obtidos foram comparados com populações europeias, asiáticas, afro-americanos, americanos asiáticos e também com populações africanas. A população do Rio Grande do Sul apresentou maior similaridade ($F_{ST} < 0,015$) com populações europeias do que com populações asiáticas e africanas ($F_{ST} > 0,015$), o que ressalta mais uma vez a grande proporção do componente europeu nesta população.

Resumidamente, pode-se concluir que comparações entre populações podem ser realizadas através do cromossomo X. Estas comparações mostram uma alta diversidade e estruturação entre as populações brasileiras e foram observadas diferenças entre as regiões geopolíticas do Brasil. Quando comparadas com outras populações mundiais pode ser observado que as populações da América Latina são as mais semelhantes às populações brasileiras. Além da distância genética entre as populações, o desequilíbrio de ligação também tem sido estudado em populações brasileiras. O resultado destas análises do desequilíbrio de ligação, em alguns casos como o de Belo Horizonte e Paraná, tem sido considerado espúrio devido a problemas de amostragem.

Comparações entre populações ameríndias e urbanas brasileiras são mais raras. Populações ameríndias da Colômbia (Kogy, Wayuu e Zebu) e não-ameríndias da América do Sul (Costa Rica e Rio Grande do Sul) foram analisadas com 17 STR-X (Amorim *et al.*, 2011). Estes autores observaram que a população ameríndia da Colômbia possui menor diversidade genética e um DL maior quando comparada com populações urbanas da Costa Rica e do Sul

do Brasil. Eles concluíram que as diferenças entre as populações podem ter sido causadas pela deriva na população ameríndia e pelo modelo de mistura seguido de isolamento ao qual as populações urbanas foram submetidas.

Como se vê, as comparações entre populações ameríndias e urbanas brasileiras são menos numerosas. Além disso, DL tem sido menos calculado para populações urbanas.

A comparação entre remanescentes de quilombos, populações urbanas e ameríndias é ainda mais rara. É relatado apenas um trabalho que os compara, através de lócus do cromossomo X (Xq28) (Massaro, 2009). Foram analisados 5 STRs (IVS 1, IVS 13, IVS 22, IVS 25.3 e IKBKG) em populações brasileiras urbanas (indivíduos normais de São Paulo, Rio Grande do Sul e Pernambuco), remanescentes de quilombos (Mimbó, Sítio Velho e Gaucinha) e ameríndios (Tikúna, Baniwa e Kashinawa). A menor diversidade foi encontrada nos quilombos e ameríndios, sendo que neste último a diversidade foi ainda menor. A maior parte da diversidade encontrada é intrapopulacional. Foi observado um alto índice de estruturação genética entre os grupos de populações (urbanas, quilombos e ameríndias) analisados. Sendo assim, a maior diferenciação ocorreu entre uma tribo indígena (Kashinawa) e um remanescente de quilombo, e as populações mais semelhantes foram Rio Grande do Sul e São Paulo. A análise de desequilíbrio de ligação entre os lócus IVS 1, IVS 13, IVS 22, IVS 25.3 apresentou resultado significativo para todas as comparações.

Nos trabalhos citados acima foram utilizados, ao todo, 53 lócus de STRs para realizar comparações entre populações. Estes lócus estão distribuídos ao longo do cromossomo X (Figura 1) e a maior parte (35) está localizada no braço longo (Xq). No braço curto (Xp), apenas doze lócus foram utilizados em estudos com populações brasileiras até o momento. É importante evidenciar que as regiões Xp21 e Xp11.1-Xq11.21 foram pouco exploradas em estudos de genética populacional no Brasil. Apenas um lócus foi utilizado na região Xp21 e seis na Xp11.1-Xq11.21.

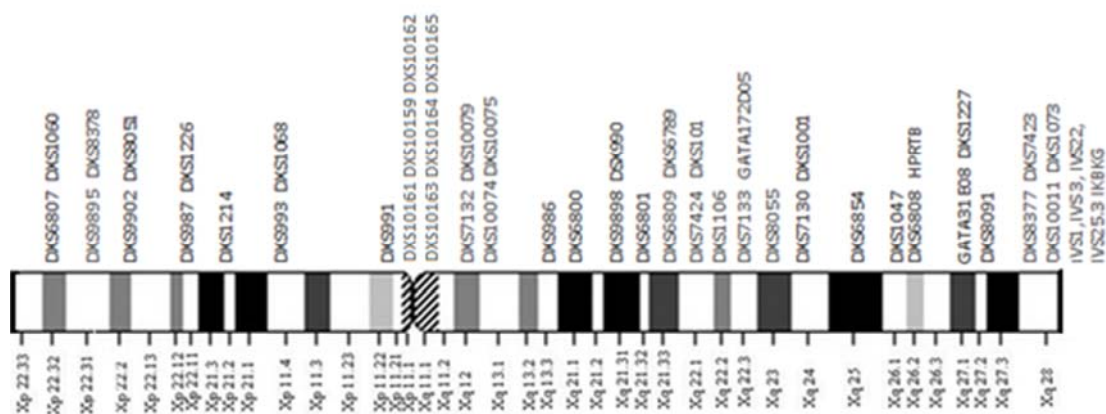


Figura 1: Localização no cromossomo X dos lócus STRs utilizados nas populações brasileiras. Adaptado de <http://www.bgbx.com.br/?pagina=idiograma>

JUSTIFICATIVA

Como pôde ser observado acima, populações ameríndias da América do sul têm sido comparadas com populações não ameríndias do Brasil.

Esses estudos abordaram populações ameríndias de vários outros países além do Brasil e como população urbana miscigenada, na grande maioria, utilizaram populações do sul do Brasil.

Neste trabalho, os indivíduos analisados são provenientes de amostras da tribo *Tikúna*, situada em território brasileiro, na Amazônia central, de duas populações urbanas brasileiras (Teresina e Ribeirão Preto) de diferentes regiões geográficas do Brasil.

Além do mais, não têm sido analisadas amostras que representam o componente africano existente na formação da população brasileira. Neste trabalho o componente africano é representado de duas formas: por dois remanescentes de quilombos (Mimbó e Sitio Velho) e por uma população africana oriunda do Congo.

Nos trabalhos citados no item acima, os marcadores empregados, totalizam 53 loci (12 dos quais no braço curto e apenas seis na região pericentromérica). Na maioria destes trabalhos, os loci analisados não estavam agrupados em blocos que permitissem cálculo robusto do desequilíbrio de ligação. Neste trabalho, os 20 loci analisados foram escolhidos por constituírem três haploblocos: seis loci agrupados na região Xp11.1-Xq11.21 (0,65cM de comprimento), outros sete na região Xp21 (8,24cM, próximo ao locus DMD) e sete restantes na região Xq28 (0,23cM de comprimento, próximo ao gene HEMA).

O haplobloco DMD se localiza em uma região onde apenas um locus foi analisado em estudos de populações brasileiras. O PC está localizado em uma região cromossômica que ainda não foi empregada em estudos deste tipo. O haplobloco HEMA, por sua vez, se encontra em uma região que tem sido utilizada nestes estudos, mas os loci selecionados para este trabalho foram utilizados pela primeira vez com o intuito de comparar populações brasileiras.

É importante ressaltar que os seis loci da região pericentromérica (PC) são inéditos, portanto, estão sendo pela primeira vez.

HIPÓTESE

A hipótese central deste trabalho é que a distribuição dos loci em três blocos de ligação produzirá um conjunto de resultados dos parâmetros básicos da genética de população, incluindo cálculos de desequilíbrio de ligação.

Com isto, espera-se não só estabelecer parâmetros de diferenciação genética entre estas populações, como também estimar o desequilíbrio de ligação (DL), cujo perfil poderá refletir os diferentes graus de isolamento genético e deriva peculiares a cada uma.

OBJETIVOS

O objetivo geral de trabalho é inferir a estrutura genética populacional entre as seis populações estudadas (Africana, Mimbó, Sítio Velho, Teresina, Tikúna e Ribeirão preto) a partir da genotipagem de 20 STRs distribuídos ao longo do cromossomo X.

Como objetivos específicos podem ser citados:

- Descrever seis novos loci STRs no cromossomo X, suas condições de amplificação e genotipagem;
- Obter informações sobre os parâmetros formais da genética de populações (frequências, equilíbrio H-W, diversidade e diferenciação) para os novos loci e também para os 14 loci descritos anteriormente;
- De posse dos parâmetros formais da genética de população, realizar comparações entre as populações estudadas para conhecer a diversidade e a diferenciação entre cada uma delas, bem como o desequilíbrio de ligação.

MATERIAIS E MÉTODOS

Desenho experimental

Neste trabalho foram obtidas as frequências gênicas de 20 loci STRs ligados ao cromossomo X em seis amostras populacionais (Africana-AF, Mimbó-MB, Sítio Velho-SV, Teresina-TR, Ribeirão Preto-RP e *Tikúna*-TK).

Estes 20 loci estão distribuídos em três blocos de ligação (PC, DMD e HEMA) ao longo do cromossomo X. O primeiro grupo, PC, é composto por seis loci na região pericentromérica (Xq11.21-Xp11.1) que possui 0,65cM de extensão. O segundo grupo é denominado DMD, pela sua proximidade e/ou sobreposição com o gene da distrofina (Xp21), e engloba sete loci dessa região. Este é o bloco mais extenso, com 8,24cM entre o primeiro e o último loci. O terceiro grupo, HEMA, tem proximidade e/ou sobreposição ao loci do Fator VIII (Xq28). Este grupo também é composto por sete loci e possui uma extensão de 0,23cM.

As seis amostras populacionais utilizadas neste trabalho podem ser divididas em quatro grupos: Africanos (Congo, n=50), remanescentes de quilombos (Mimbó, n=25 e Sítio Velho, n=25), Indígenas (tribo *Tikúna*, n=50) e Populações urbanas (Ribeirão Preto, n=51 e Teresina, n=50).

A partir das frequências alélicas dos loci descritos acima, foram calculados vários parâmetros estatísticos: diversidade haplotípica, aderência ao Equilíbrio de Hardy-Weinberg, diferenciação gênica e genotípica, diversidade gênica, estrutura populacional (ancestralidade) e desequilíbrio de ligação.

Descrição das amostras populacionais

População Africana (Congo)

Através de uma colaboração existente com a equipe do Dr. André Garcia do “Mère et Enfant face aux infections tropicales, Institut de Recherche pour le Développement” et “Laboratoire de Parasitologie”, Université Paris Descartes, Paris, França foi possível obter amostras de africanos pertencentes a regiões (República Democrática do Congo) de onde ocorreram eventos de migração para o Brasil. Estas amostras de Africanos da República Democrática do Congo passaram a fazer parte de um banco de amostras “BANFRI”, sob a responsabilidade do Prof. Dr. Antonio Eduardo Donadi, aprovado pelo Comitê de Ética em 18 de Julho de 2012, conforme o processo HCRP nº7581/2007.

A República Democrática do Congo possui uma população em torno de 5 milhões de habitantes. A amostra proveniente deste país foi coletada em uma população da rural que vive próximo ao rio Congo, na província de Bandundu. Esta província está localizada na região sudoeste do país e possui uma população em torno de 120 mil habitantes. Foram amostradas 143 famílias,

totalizando 354 indivíduos (Courtin *et al.*, 2006); e destas, 50 amostras foram selecionadas dentre indivíduos não aparentados para serem utilizadas neste trabalho.

Comunidades Remanescentes de Quilombos.

1. **Mimbó.** Localiza-se a 6° 14' latitude Sul e 42° 50' longitude Oeste a 22 Km da sede do município de Amarante e a 170 Km de Teresina, capital do Estado do Piauí. Segundo moradores mais antigos, o povoado originou-se na época da abolição da escravatura. Grupos de negros abandonaram as fazendas e migraram em busca de um local para se fixarem e viverem sua liberdade. De um desses locais, próximo ao município de Oeiras, no início do século 20, desmembraram-se do grupo três casais e se fixaram às margens do riacho Mimbó, onde gradativamente, foram incorporados ao grupo outros migrantes do povoado original ou negros oriundos de outras fazendas do Estado. A liderança é sempre mantida pelos membros mais velhos e embora não existam leis escritas, vigora o bom senso na decisão sobre o que é permitido ou não. Utilizam a agricultura de subsistência, em especial a de mandioca. Os produtos obtidos na agricultura são divididos entre todos os moradores, que se consideram uma grande família.
2. **Sítio Velho.** Localiza-se a 5° 20' latitude Sul e 41° 15' longitude Oeste a 280 Km de Teresina - PI. Segundo moradores mais antigos, a comunidade surgiu no início do século 20 por aglomeração, nas margens de um riacho onde a caça era abundante, de africanos oriundos de diferentes pontos de uma região relativamente ampla da fronteira entre os Estados do Piauí e Ceará. Sua existência foi registrada por volta de 1984, quando cearenses, em busca de terras, chegaram ao povoado desencadeando uma guerrilha com seus moradores. Praticam a agricultura de subsistência.

Indígenas da Amazônia Brasileira (Tikúna)

Os Tikúna habitam o Estado do Amazonas, mas representantes também podem ser encontrados em território peruano e colombiano. Na época da coleta havia cerca de 11.000 membros desta tribo em território brasileiro. No entanto, o censo de 1994 indicou 23.000 membros, também em território brasileiro. Embora apresentem contatos com não-indígenas, esta tribo ainda mantém a sua identidade étnica e casamentos com pessoas de fora são raros (Simões, 1980).

Populações Urbanas

1. **Teresina** A amostra de Teresina é utilizada como amostra urbana de referência para os remanescentes de quilombos Mimbó, Sítio Velho. É composta por 87 indivíduos nascidos no Estado do Piauí, selecionados entre os doadores de sangue do Hemocentro de Teresina. Apenas 50 amostras foram selecionadas e nesta população a maioria é de homens.

2. **Ribeirão Preto.** A amostra de Ribeirão Preto é constituída por 300 indivíduos residentes na cidade de Ribeirão Preto. Estes foram amostrados para realização do projeto “Dinâmica populacional e uso forense de polimorfismos genéticos associados a fenótipos morfológicos no Brasil”, em execução, sob orientação do Prof. Celso Teixeira Mendes Jr. Destas, 51 amostras foram selecionadas aleatoriamente para serem utilizadas.

Coleta das amostras

Africanos

As amostras de Africanos foram coletadas em famílias da população rural que vive próximo ao rio Congo na província de Bandundu. A amostragem foi realizada para utilização em projetos de imunologia, mais precisamente para estudos de comparação dos níveis de citocinas plasmáticas em indivíduos Africanos com tripanossomiase por Courtin *et al.* (2006) no qual mais detalhes podem ser obtidos. Para a utilização nesta tese, apenas um indivíduo de algumas das famílias amostradas foi selecionada, visando o menor grau de parentesco entre as amostras.

Mimbó, Sítio Velho e Teresina

As amostras analisadas são provenientes de duas comunidades afro-derivadas (Mimbó e Sítio Velho), situadas no Estado do Piauí além de uma amostra urbana do Hemocentro de Teresina.

As amostras foram coletadas em 1993 por professores da Universidade Federal do Piauí, liderados pela professora Dra. Zélia Arpini Sampaio (Arpini-Sampaio *et al.*, 1999), em projeto que incluía atendimento médico e alguns exames laboratoriais de assistência às comunidades estudadas.

Fez-se uma visita a todos os domicílios das três comunidades. Para cada indivíduo com idade superior a três meses, preencheu-se uma ficha para a identificação e registros de dados pessoais e genealógicos. Tais amostras foram caracterizadas com base em marcadores clássicos e dois STRs, para vários parâmetros, incluindo levantamento demográfico (Arpini-Sampaio *et al.*, 1999).

Tikúna

As amostras de sangue venoso foram coletadas no ano de 1976 durante a expedição “Alpha-Hélix” (Neel *et al.*, 1980; Mestriner *et al.*, 1980), destinada à Amazônia oriental brasileira e aprovada segundo processo FUNAI/BSB/4854/75 e autorizações 26/76, 27/76, 74/76 e 75/76. Descrição detalhada encontra-se em Simões (1980). O critério para a seleção dos indivíduos da amostra analisada foi o menor grau de parentesco possível dentro da amostra total. A partir das fichas de coleta, que incluem o número de identificação, o nome do indivíduo, sexo, idade, notas sobre parentesco, mistura com brancos e o nome dos pais e/ou esposos, foi possível excluir primeiramente os filhos e netos da amostra. Em uma segunda triagem foi escolhido, de maneira aleatória, apenas um indivíduo dentre os que

apresentavam irmãos ou o mesmo sobrenome, o que resultou em uma amostra de 323 indígenas (Mendes Jr., 2005).

Ribeirão Preto

Estas amostras foram coletadas para o projeto de dissertação “Análise de polimorfismos do gene MC1R associados a fenótipos humanos de pigmentação na população brasileira” coordenado pelo Prof. Dr Celso Mendes Jr. e cedidas em colaboração com este trabalho. Na cidade de Ribeirão Preto (SP) foram coletadas amostras de sangue de indivíduos voluntários tanto homens como mulheres, com idade entre 18 e 40 anos, não aparentados e que pertenciam a diferentes grupos étnicos da população brasileira.

Estocagem do material e extração de DNA

Africanos

As amostras coletadas pela equipe do Dr. André Garcia do Laboratório de Parasitologia da Universidade de Paris Descartes (Paris, França) de acordo com Courti *et al.* (2006). O DNA foi extraído através do método “salting-out” (Miller *et al.*, 1988), pelo grupo do Dr. André Garcia, e foram enviados 20 µL para o Brasil. Estas amostras foram devidamente registradas no Comitê de Ética (processo HCRP n°7581/2007) e estão armazenadas sob responsabilidade do Prof. Dr. Antonio Eduardo Donadi (Depto. De Clínica Médica / FMRP-USP) que cedeu parte delas para o presente trabalho.

Mimbo, Sítio Velho e Teresina

Coletou-se 5 mL de sangue em tubo VacuTainer, com EDTA, os quais foram mantidos e transportados em caixas de isopor e transferidos para a geladeira, até 12 horas após a coleta.

O total de amostras coletadas para cada uma destas populações foi: 91 amostras do Hemocentro de Teresina, 183 na comunidade de Mimbó, 182 na comunidade de Sítio Velho.

Todas as amostras foram processadas no laboratório de Biologia Geral da Universidade Federal do Piauí, em Teresina, em até 72 horas após a coleta. O plasma foi estocado diretamente a -20°C, o *buffy coat* (cerca de 500 µL) foi recolhido em tubo de polipropileno (*eppendorf*) de 1,5 mL e mantido em geladeira para a extração de DNA. Os eritrócitos restantes foram lavados em solução salina e estocados - 20°C em glicerol 40% v/v em tampão contendo Citrato tripotássico KH₂PO₄, K₂HPO₄, H₂O. As amostras devidamente processadas e congeladas foram transportadas por via aérea para o laboratório de Genética Bioquímica do Departamento de Genética da FMRP. Todos esses procedimentos foram realizados pela equipe liderada pela professora Dra. Zélia Arpini Sampaio. O DNA precisou ser extraído do restante de células misturadas a esses eritrócitos. Por isso foi necessário utilizar metodologias mais refinadas de extração, como o ‘QIAamp DNA Investigator Kit’ da

Qiagen. O DNA foi extraído de acordo com o protocolo estabelecido no Kit. A extração dessas amostras foi realizada pela responsável desta tese, durante a execução da mesma.

Tikúna

As amostras de sangue dos indivíduos de *Tikúna* foram coletadas em tubos contendo anticoagulante ACD e mantidas em refrigeração por até 12 horas. O processamento deste material para a estocagem ocorreu depois de três dias. O plasma foi retirado e as hemácias foram lavadas três vezes em solução salina isotônica. Após centrifugação, o sedimento resultante desse processo foi suspenso em solução de glicerol a 40% em tampão de estocagem. As amostras são mantidas até hoje, congeladas a -20°C . Informações mais detalhadas estão em Simões, 1980. A extração de DNA destas amostras também foi realizada pela responsável desta tese através do Kit 'QIAamp DNA Investigator Kit' da Qiagen.

Ribeirão Preto

O sangue foi coletado (10ml) em tubos com anticoagulante (EDTA). As amostras foram submetidas à centrifugação para retirada do plasma sanguíneo. A extração de DNA foi realizada pelo método *salting-out* (Miller et al., 1988). O DNA extraído foi eluído em $400\mu\text{L}$ de água estéril e armazenado a -20°C . Essa extração foi realizada pelo pesquisador Leonardo Arduino Marano, durante seu mestrado (Marano, 2011).

Marcadores genéticos

Foi analisado um grupo de 20 loci em diferentes localizações genéticas. Estes estão distribuídos em três regiões do cromossomo X (Figura 2, Tabela 1): Xp11.1-Xq11.21 (pericentroméricos – PC) com seis loci, Xp21 (Distrofia Muscular de Duchenne – DMD) e Xq28 (Hemofilia - HEMA), ambos composto por 7 loci.

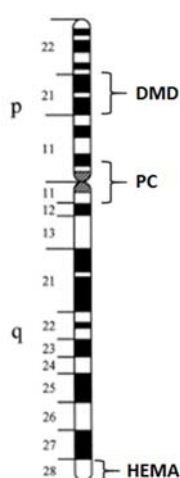


Figura 2: Localização dos três blocos de loci STRs no cromossomo X que foram utilizados neste trabalho.

Todos os iniciadores (*primers*) para detecção destes STRs foram descritos pelo grupo liderado pelo Prof. Enrique Medina-Acosta, da Universidade Federal do Norte Fluminense, e foram cedidos em colaboração com o presente estudo. Os iniciadores desenvolvidos para a região pericentromérica são inéditos. Dos sete marcadores da região Hemofilia, cinco (REN90833, F8int25.2, HEMA154311.3, TMLHInt5 e HEMA154507.3) foram recentemente descritos (Machado, 2008) e usados para detecção de portadores de Hemofilia A (Machado, 2009a; Machado, 2009b; Machado, 2009c). Os outros dois marcadores, F8int22 (Lalloz et al., 1994) e F8int13.2 (Lalloz et al., 1991) foram descritos anteriormente. Os sete marcadores da região DMD foram descritos e utilizados para a identificação indireta de portadores de mutação no gene DMD por Sarlo (2011).

Tabela 1: Informações sobre os 20 loci utilizados. N°: número de ordem dado para os loci dentro de cada haplobloco. cM: Distância genética em centimorgans. pb: número de pares de bases do amplicon. A distância genética foi obtida em: <http://compgen.rutgers.edu/old/map-interpolator/>

N°	Primer	Locus	Repetição (UR)	Localização (pb)	(cM)	(pb)	Fluoróforo
Xp21 – Distrofia Muscula de Duchene (DMD) – 8,25cM							
1	PR1	DXS31212596	CTT	31302625--31302831	53,327	207	FAM-6
2	PR4	DXS31841757	AAAAC	31931802--31932007	54,7628	206	VIC
3	PR8	DXS32069474	CAAAAA	32159403--32159628	55,5868	226	NED
4	PR9	DXS32064591	AAACA	32154641--32154744	55,5646	274	FAM-6
5	PR10	DXS32161638	TTTA	32251680--32251841	56,1903	136	NED
6	PR13	DXS32336904	ATCT	32426703--32427069	57,3406	367	PET
7	PR14	DXS33001871	AAAT	33091924--33092130	61,5753	207	PET
Xp11.1-Xq11.21 – Pericentroméricos (PC) – 0,66cM							
1	PC1	DXS56749371	GAAA	56749371--56749571	90,0056	230	PET
2	PC2	DXS56909198	ATTT	56909000--56909298	90,1024	126	NED
3	PC3	DXS57434969	ATTT	57434869--57435069	90,4398	125	FAM-6
4	PC4	DXS61920979	CTTT	61920879--61921079	90,6504	209	VIC
5	PC5	DXS62085977	TAAAA	62085877--62086077	90,6569	325	NED
6	PC6	DXS62290471	GAAA	62290371--62290671	90,6642	334	FAM-6
Xq28 – Hemofilia (HEMA) – 0,23cM							
1	PH1	REN 90833	TTAT	153694353--153694393	195,7527	138	PET
2	PH2	F8Int25.2	TG	15373606--153736088	195,7711	143	FAM-6
3	PH3	F8Int22	GT	153757206--153757243	195,7773	214	FAM-6
4	PH4	F8Int13.2	AC	153817451--153817493	195,7957	134	NED
5	PH5	HEMA154311.3	GAAGA	154311336--154311361	195,9509	246	VIC
6	PH6	TMLHInt5	TTTTA	154393088--154393119	195,9664	219	NED
7	PH7	HEMA154507.3	ATCT	154507351--154507428	195,9826	334	VIC

Reação em cadeia da polimerase – PCR

Os 20 lócus utilizados (Tabela 1) foram amplificados através da reação em cadeia da polimerase – PCR. Três PCRs multiplex foram realizadas, a primeira para a amplificação dos lócus da região Xp21 (DMD). Esses lócus foram amplificados em heptaplex, ou seja, os sete lócus amplificados de uma só vez na PCR. A segunda PCR foi realizada para a amplificação do hexaplex da região Xp11.1 (Pericentromérica). E a terceira, para a amplificação do heptaplex da região Xq28 (Hemofilia).

A reação PCR multiplex para as regiões DMD (Sarlo, 2011) e HEMA (Machado, 2008) já estavam otimizadas. Para a região Pericentromérica foi necessário testar os iniciadores e otimizar a PCR multiplex de acordo com o descrito a seguir.

Para a otimização da PCR multiplex foram avaliados os efeitos de algumas variáveis como: concentração do DNA alvo, concentração dos iniciadores, e a ciclagem. Foram realizados ensaios com várias temperaturas (PCR gradiente) com as seguintes variáveis: (1) temperatura de pareamento, (2) concentração de iniciadores; (3) concentração de $MgCl_2$; e (4) concentração de DNA. A reação utilizada teve a seguinte formulação: 0.2 mM dNTP, 2mM $MgCl_2$, 1.25U Taq Polimerase (Alves Da Silva, 2007). Na primeira etapa a PCR foi realizada em um termociclador gradiente (*96-Well Veriti Thermal Cycler - Applied Biosystems, USA*) para identificar a melhor temperatura para o funcionamento dos iniciadores, em reação uniplex. Na segunda etapa foi realizada a junção dos iniciadores com temperatura de pareamento iguais. A termociclagem otimizada para a PCR multiplex PC teve as seguintes etapas: um aquecimento inicial a 95°C por 11 segundos, e em seguida 28 ciclos, sendo: 94°C por um minuto para desnaturar o DNA, 60/63°C por um minuto para o pareamento dos iniciadores, 72°C por um minuto para a extensão do fragmento, e no final dos ciclos, 60°C por 60 minutos para finalizar a reação.

Nas três PCRs realizadas (DMD, PC, HEMA) a concentração de DNA foi 37,5ng. Para algumas amostras (Mimbó, Sítio Velho, Teresina e Tikuna) o número de ciclos da PCR foi de 40. Estas amostras foram as que apresentaram maior dificuldade na extração de DNA, como relatado anteriormente. Na tabela 2 estão detalhados todas as concentrações e temperaturas utilizadas para cada uma das três regiões.

Eletroforese Capilar

Aos produtos de amplificação foi adicionado 0,01µL formamida (Hi-Di Formamida, *Applied Biosystems*) e 0,15µL do padrão de massa molecular *GeneScan LIZ 500*, marcado com o fluorocromo LIZ™ (fluorescência laranja) (*Applied Biosystems*). Os produtos de amplificação foram separados por eletroforese capilar com o polímero *POP 4* (contendo 4% de acrilamida), utilizando a plataforma *ABI Prism™ 310 Genetic Analyzer*.

Tabela 2: Temperatura de pareamento e concentração dos reagentes utilizados em cada uma das três PCRs multiplex.

	DMD	PC	HEMA
dNTP	0,2 mM	0,2 mM	0,2 mM
MgCl₂	2mM	2mM	2mM
Taq	0,0125 U	0,0125 U	0,0125 U
DNA	37,5 ng	37,5 ng	37,5 ng
Ciclos	28/40	28/40	28/40
T°	60°C	60°C	63°C
Primers (µmol)			
PR1	8		
PR2	8		
PR3	4		
PR4	4		
PR5	3		
PR6	15		
PR7	10		
PC1		10	
PC2		5	
PC3		7	
PC4		3	
PC5		20	
PC6		10	
PH1			8
PH2			10
PH3			10
PH4			10
PH5			7
PH6			6
PH7			4

Leitura dos alelos

Os dados gerados foram interpretados pelo programa *GeneScan™* e também pelo *Genotyper™* (Applied Biosystems). Estes programas geram uma imagem do pico, que representa cada alelo e realiza o reconhecimento deste. Os alelos foram nomeados com o valor em pares de base, dos produtos da PCR. Para as três PCRs multiplex foi possível visualizar todos os alelos de uma só vez (Figuras 3, 4 e 5).

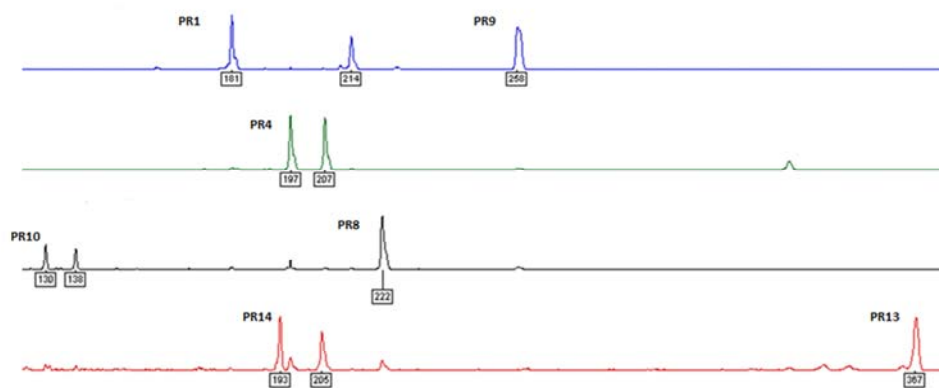


Figura 3: Eleetroferograma do Heptaplex – DMD – de uma mulher.

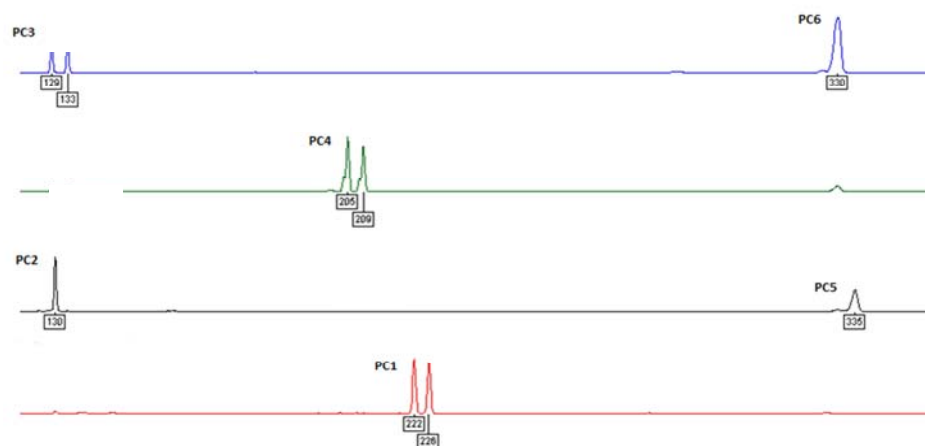


Figura 4: Eleetroferograma do Hexaplex – PC – de uma mulher.

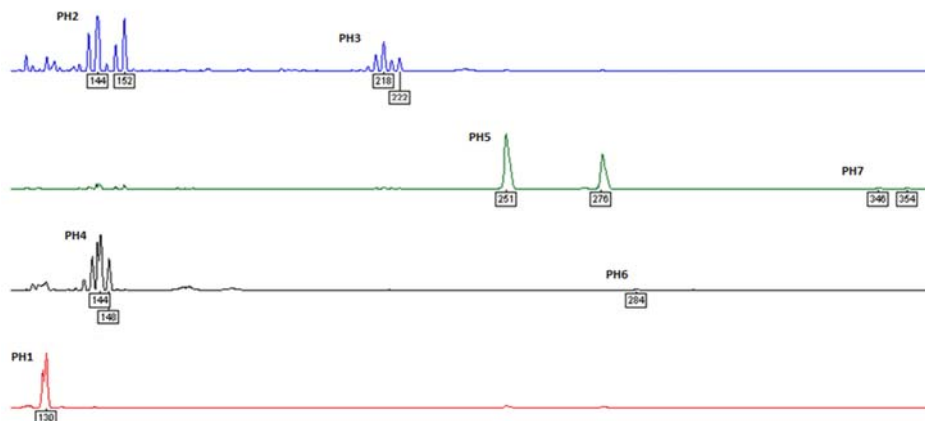


Figura 5: Eleetroferograma do Heptaplex – HEMA – de uma mulher.

Análise estatística

Estimativas das frequências alélicas

Os microssatélites analisados apresentam alelos codominantes, o que permite inferir os genótipos a partir dos respectivos fenótipos. As frequências alélicas (x_i) de cada *locus* em cada amostra foram estimadas por contagem direta, utilizando-se o programa GENEPOP versão 3.4 (Raymond e Rousset, 1995a), segundo as equações:

$$x_i = \frac{2n_{ii} + \sum n_{ij}}{2n}$$

Em que:

x_i é a frequência do alelo i ;

n_{ii} e n_{ij} correspondem ao número de homocigotos e heterocigotos observados para o alelo i , respectivamente;

n corresponde ao número de indivíduos analisados.

Diversidade haplotípica

A diversidade haplotípica (h), equivalente à heterocigose em dados diplóides, foi estimada considerando a frequência dos haplótipos como definida por Nei (1987). No caso dos genomas haplóides, h representa a probabilidade de que dois haplótipos escolhidos aleatoriamente sejam diferentes na amostra. As fórmulas da diversidade haplotípica (h) e seu desvio padrão (DP) são:

$$h = \frac{n}{n-1} \left(1 - \sum_{i=1}^k p_i^2 \right)$$

$$DP = \left\{ \frac{2}{n(n-1)} \left[2(n-2) \left[\sum p_i^3 - \left(\sum p_i^2 \right)^2 \right] + \sum p_i^2 - \left(\sum p_i^2 \right)^2 \right] \right\}^{\frac{1}{2}}$$

Em que:

p_i é a frequência do haplótipo i ;

n é o número de cromossomos da amostra;

k é o número de haplótipos.

Além da diversidade haplotípica também foram identificados os haplótipos existentes em cada população.

Os cálculos foram realizados com o programa ARLEQUIN (Schneider *et al.*, 2000). Esta análise foi realizada para cada um dos grupos de loci separados (DMD, PC e HEMA). Para

a realização desta análise foram retirados os indivíduos que apresentaram falha na amplificação de algum marcador.

Aderência ao equilíbrio de Hardy-Weinberg

Segundo o teorema de Hardy-Weinberg, as frequências genóticas esperadas no equilíbrio podem ser estimadas a partir da expansão do seguinte binômio:

$$(x_i + x_j)^2 = x_i^2 + 2x_i x_j + 2x_j^2$$

Em que:

x_i^2 é a frequência esperada dos homozigotos do alelo i;

$2 x_i x_j$ é a frequência esperada do heterozigoto ij;

$2x_j^2$ é a frequência esperada dos homozigotos para o alelo j.

A aderência das frequências genóticas observadas às proporções teóricas de Hardy-Weinberg foi verificada com o emprego do programa GENEPOP (Raymond & Rousset, 1995a) versão 3.4, disponível em <http://wbiomed.curtin.edu.au/genepop>. Foram realizados três testes baseados na hipótese nula de união aleatória dos gametas: teste exato de probabilidade, teste para detecção da deficiência e para detecção do excesso de heterozigotos.

No teste exato de probabilidade, o valor de P corresponde à soma de probabilidades de todas as tabelas com probabilidade menor ou igual ao observado.

O segundo e o terceiro são testes mais sensíveis do que o de probabilidade e utilizam uma hipótese alternativa (H1) de excesso ou de deficiência de heterozigotos, respectivamente.

Diferenciação genética das populações

Os testes exatos para diferenciação populacional foram realizados com o uso do programa GENEPOP 3.4 (Raymond & Rousset, 1995a). Este utiliza tabelas de contingência RxC geradas automaticamente para cada locus, em que R é o número de populações e C é o número de alelos no locus. Este procedimento compara cada locus em pares de populações, para determinar se existem diferenças nas frequências alélicas observadas, onde a hipótese nula testada é a de que a distribuição alélica é idêntica entre as populações (Raymond & Rousset, 1995b).

As análises de diferenciação gênica deste tópico foram realizadas em dois grupos de amostras: apenas as amostras de mulheres e depois as amostras de homens. A análise de

diferenciação genotípica so pode ser realizada com as mulheres, por se tratar de dados do cromossomo X.

Diversidade Gênica

Heterosigoze média

A diversidade gênica média (\bar{H}_S) com o respectivo erro padrão foi calculada para cada amostra populacional utilizando o programa ARLEQUIN (Schneider *et al.*, 2000) conforme a equação 8.6 apresentada em NEI (1987):

$$H = \sum_{j=1}^r h_j / r$$

Em que:

r é o número de loci utilizados;

h_j, de acordo com a equação 8.1 de Nei (1987), é a heterozigose esperada para cada lócus na j-ésima população, estimada por:

$$h = 1 - \sum_{i=1}^m x_i^2$$

Em que:

m é o número de alelos.

Esta medida de heterozigose média é equivalente à proporção média de heterozigotos por *locus* em uma população com padrão de acasalamento aleatório e, também, é igual à proporção de *loci* heterozigotos em um indivíduo escolhido aleatoriamente. O desvio padrão desta estimativa é descrito pela seguinte equação adaptada (Nei, 1987):

$$H = \left[\sum_{j=1}^r h_j - H \right]^2 / (r-1)r^{1/2}$$

Outro parâmetro para medir a diversidade gênica é o *F_{ST}*. Este é o coeficiente de endogamia em um par de populações, dado pela probabilidade de dois alelos tomados ao acaso em duas populações serem idênticos por descendência (déficit de heterozigotos entre populações). Foi utilizado para se estimar a diversidade genética entre duas ou mais populações. Os cálculos, feitos por meio do programa ARLEQUIN® (Schneider *et al.*, 2000), utilizando-se método baseado na correlação entre alelos (Weir e Cockerham, 1984). Embora

os resultados não sejam exatamente os mesmos, o F_{ST} também pode ser definido em termos de probabilidade de identidade por descendência entre pares de alelos (Nei, 1987):

$$F_{ST} = \frac{(J_S - J_T)}{(1 - J_T)}$$

onde, considerando-se uma população total dividida em k sub populações S , sendo p_{ki} a frequência do alelo i na sub população k e p_i a frequência do alelo i na população total,

temos: $J_S = \frac{\sum_k J_k}{S} = \frac{\sum_k \left(\sum_i p_{ki}^2 \right)}{S}$ correspondente à probabilidade média de que dois alelos

amostrados ao acaso em uma mesma sub população sejam idênticos por descendência e

$J_T = \sum_i p_i^2$ à probabilidade de que dois alelos amostrados ao acaso na população total

sejam idênticos por descendência (Nei, 1987).

Análise de Variância Molecular (AMOVA)

O Programa ARLEQUIN (Schneider *et al.*, 2000) foi usado para estimar a diferenciação genética das populações pela análise de variância molecular (AMOVA), o que permite análises hierárquicas de três componentes da variância genética, ou seja, aquela devido a diferenças genéticas entre indivíduos dentro das populações, entre populações dentro dos grupos e entre grupos. O teste de significância dos valores de variância genética foi estimado com o uso de 10.000 permutações.

Inferência de Estrutura Populacional

O programa *Structure* (<http://pritch.bsd.uchicago.edu/structure.html>) utiliza um método de agrupamento, baseado no modelo desenvolvido por Pritchard *et al.* (2000), para a inferência de estrutura populacional utilizando dados genotípicos. As análises foram realizadas com $K = 3$ como parâmetro predefinido para o número de populações assumidas como parentais, com 30.000 interações para o período *burn-in* e 100.000 interações adicionais para obter as estimativas dos parâmetros.

Desequilíbrio de ligação entre loci

A análise de associações par-a-par entre *loci* foi realizada utilizando-se o programa ARLEQUIN (Schneider *et al.*, 2000). A hipótese nula é a de que a distribuição genotípica em um *locus* é independente da distribuição em outro *locus*. Esta análise foi aplicada para verificar desvios do esperado pela regra de multiplicação entre pares de *loci* localizados em diferentes cromossomos. A palavra “ligação” neste caso não está relacionada com associação física entre alelos de *loci* de um mesmo cromossomo.

As amostras das mulheres tiveram a fase inferida pelo algoritmo EM existente no programa ARLEQUIN (Schneider *et al.*, 2000). Com a fase inferida, os haplótipos das mulheres foram utilizados juntamente com os dos homens, como dado haploide de fase conhecida. Os marcadores PR4 e PR14 foram homozigotos na amostra de homens da população *Tikúna*, por isso, nesta população, foram retirados da análise de DL.

Correção de Bonferroni

A correção de Bonferroni foi aplicada para as análises que se baseiam no valor de significância (p). Sendo assim, a significância foi recalculada de acordo com o número de testes independentes (Hartl e Clark, 2007) em cada análise.

A aderência ao Equilíbrio de Hardy-Weinberg teve seu nível de significância recalculado da seguinte forma: $\alpha=0,05$ dividido pelo número de testes independentes (20), que resultou no valor corrigido de $\alpha_{\text{Bonf}}=0,0025$. As análises de diferenciação gênica e genotípica assim como o F_{ST} e DL também tiveram o nível de significância recalculado de acordo com a correção de Bonferroni.

RESULTADOS

Alelos Novos

Neste estudo foi descrito pela primeira vez, o procedimento para detecção dos alelos de seis loci do tipo STR localizados na região pericentromérica (Xp11.1-Xq11.21) do cromossomo X humano. Estes loci compõem o bloco PC.

São também apresentadas as frequências alélicas de 14 outros loci, sete deles formando um bloco (DMD) ligado ao gene da Distrofina (Xp21) e os outros (bloco HEMA) ligado ao gene do Fator VIII (Xq28). Os procedimentos para classificação fenotípica destes dois blocos já foram previamente descritos (Sarło, 2011 e Machado *et al.*, 2009c).

Os seis loci inéditos (PC) apresentaram 55 alelos (Tabela 3), com média de 9,16 alelos. Nos outros 14 loci, pertencentes ao bloco DMD e HEMA, foram identificados 34 alelos novos, que ainda não tinham sido descritos, distribuídos em 11 loci (um nos loci PR14, PH2 e, PH3; dois em PR9 e PR10; três em PR1; cinco em PR13 e PH5; seis em PH6 e sete em PH1). Todos estes 89 (55 + 34) alelos descritos pela primeira vez estão sublinhados e em negrito na primeira coluna da Tabela 3. Os loci PR4, PR8 e PH7 são os únicos que não apresentaram alelos novos. A população Africana foi a que teve maior número de alelos novos (oito alelos: *PC1*238*, *PC2*110*, *PC2*114*, *PC4*223*, *PC6*314*, *PC6*332*, *PH4*138*, *PH6*279*).

Tabela 3: Freqüências alélicas dos 20 STRs-X em todas as populações (Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e *Tikúna*). N-número de cromossomos X analisados em cada lócus.

Lócus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			<i>Tikúna</i>		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
PR1																		
N	14	58	72	13	24	37	12	24	36	38	24	62	36	28	64	27	46	73
181	0,071	0,121	0,111				0,083	0,167	0,139	0,079	0,083	0,081						
190		0,034	0,028					0,042	0,028	0,079	0,083	0,081	0,028	0,143	0,078			
193	0,214	0,103	0,125	0,308	0,542	0,459				0,079	0,208	0,129	0,083	0,143	0,109			
196	0,071	0,052	0,056	0,231		0,081	0,167	0,083	0,111	0,053		0,032	0,028	0,071	0,047	0,037	0,087	0,068
199	0,143	0,017	0,042					0,042	0,028	0,026		0,016	0,028	0,036	0,031	0,037		0,014
202	0,071	0,138	0,125	0,231	0,167	0,189	0,083	0,083	0,083	0,106	0,167	0,129	0,083	0,036	0,063	0,074	0,304	0,219
205		0,103	0,083	0,154	0,042	0,081	0,083	0,042	0,056	0,026	0,083	0,048	0,028		0,016	0,111	0,152	0,137
208	0,143	0,052	0,069		0,042	0,027		0,042	0,028	0,026	0,042	0,032	0,056	0,036	0,047	0,148	0,065	0,096
211		0,034	0,028	0,077	0,042	0,054				0,079	0,042	0,065		0,071	0,031	0,074		0,027
214	0,071	0,138	0,125							0,026		0,016	0,056	0,036	0,047			
217		0,017	0,014					0,250	0,167				0,056		0,031	0,037	0,109	0,082
220		0,034	0,028				0,083	0,042	0,056	0,053	0,042	0,048	0,167	0,179	0,172		0,022	0,014
223					0,083	0,054				0,079	0,042	0,065	0,083	0,071	0,078			
226		0,034	0,028				0,083		0,028	0,079	0,125	0,097	0,056	0,107	0,079	0,074	0,109	0,096
229		0,017	0,014		0,017	0,027	0,083	0,083	0,083	0,053	0,042	0,048	0,056	0,036	0,047	0,037	0,022	0,027
232		0,069	0,056		0,042	0,027	0,083	0,042	0,056	0,079	0,042	0,065	0,111		0,063	0,111	0,022	0,055
235	0,071	0,034	0,042							0,053		0,032	0,028		0,016	0,074	0,022	0,041
238							0,083	0,083	0,083	0,026		0,016		0,036	0,016	0,074		0,027
<u>241</u>							<u>0,167</u>		0,056				<u>0,028</u>		0,016			

Locus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
<u>244</u>	0,143		0,028													0,111	0,065	0,082
<u>250</u>													0,028		0,016		0,022	0,014
PR4																		
N	19	62	81	13	24	37	13	24	37	38	24	62	37	28	75	27	46	73
197	0,579	0,532	0,543	0,385	0,208	0,270	0,462	0,458	0,459	0,237	0,083	0,177	0,270	0,214	0,246			
202	0,105	0,226	0,198	0,462	0,583	0,541	0,385	0,458	0,432	0,605	0,875	0,710	0,649	0,536	0,600	<u>1,000</u>	<u>1,000</u>	<u>1,000</u>
207	0,316	0,242	0,259	0,154	0,208	0,189	0,154	0,083	0,108	0,158	0,042	0,113	0,081	0,250	0,154			
PR8																		
N	19	62	81	13	24	37	13	24	37	38	24	62	37	28	75	27	46	73
210													0,027		0,015		0,022	0,014
216	0,105	0,242	0,210	0,615	0,458	0,514	0,308	0,458	0,405	0,526	0,583	0,548	0,649	0,643	0,646	0,630	0,630	0,630
222	0,421	0,532	0,506	0,385	0,458	0,432	0,462	0,292	0,351	0,368	0,333	0,355	0,189	0,321	0,246	0,370	0,348	0,356
228	0,474	0,226	0,284		0,083	0,054	0,231	0,250	0,243	0,105	0,083	0,097	0,135	0,036	0,092			
PR9																		
N	19	62	81	13	24	37	13	24	37	38	24	62	37	28	75	27	46	73
254											0,167	0,065	0,054		0,031			
<u>257</u>		0,032	0,111	0,231		0,243	0,538		0,378	0,105		0,097	0,108		0,108	0,259		0,247
259	0,526	0,355	0,309		0,250			0,292		0,158	0,166	0,129	0,027	0,143	0,031			
264	0,421	0,565	0,531	0,692	0,750	0,730	0,385	0,542	0,486	0,658	0,667	0,661	0,649	0,714	0,677	0,741	0,739	0,740
<u>267</u>								0,167	0,135	0,053		0,032	0,027	0,107	0,107		0,022	0,014

Locus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
269	0,053	0,048	0,037	0,077			0,077			0,026		0,016	0,135	0,036	0,046			
PR10																		
N	19	62	81	13	24	37	13	24	37	35	24	57	37	28	65	26	46	72
<u>114</u>										<u>0,029</u>		<u>0,018</u>						
122		0,016	0,012	0,077		0,027		0,083	0,054		0,136	0,053	0,027	0,143	0,077			
126	0,105	0,032	0,049	0,077	0,042	0,054		0,083	0,054	0,029		0,018	0,054	0,071	0,062		0,022	0,014
130	0,316	0,323	0,333	0,385	0,500	0,459	0,154	0,208	0,189	0,171	0,227	0,193	0,108	0,250	0,169		0,109	0,069
134	0,053	0,387	0,309	0,154	0,292	0,243	0,462	0,333	0,378	0,600	0,318	0,491	0,595	0,214	0,431	0,654	0,543	0,583
138	0,526	0,194	0,259	0,308	0,167	0,216	0,385	0,292	0,324	0,171	0,318	0,228	0,162	0,214	0,185	0,346	0,326	0,333
142		0,048	0,037											0,107	0,046			
<u>146</u>													<u>0,027</u>		<u>0,015</u>			
148													<u>0,027</u>		<u>0,015</u>			
PR13																		
N	19	62	81	13	24	37	13	22	35	34	20	54	37	28	65	27	46	73
356	0,105	0,081	0,086	0,308	0,292	0,297	0,077	0,045	0,057	0,029		0,019						
358											0,100	0,037	0,054				0,022	0,014
360	0,105	0,081	0,086		0,083	0,054				0,059		0,037			0,031		0,022	0,014
362										<u>0,088</u>	<u>0,050</u>	<u>0,074</u>						
<u>364</u>	0,211	0,145	0,160		0,083	0,054	0,154	0,318	0,257	0,059	0,050	0,056	0,189	0,179	0,185		0,022	0,014
<u>366</u>								0,045	0,029	0,265	0,200	0,241				0,037	0,022	0,027
368	0,211	0,323	0,296	0,462	0,208	0,297	0,462	0,182	0,286	0,118	0,050	0,093	0,270	0,393	0,323	0,481	0,413	0,438
<u>370</u>										<u>0,206</u>	<u>0,250</u>	<u>0,222</u>						

Locus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
372	0,368	0,290	0,309	0,154	0,292	0,243	0,308	0,227	0,257	0,088	0,150	0,111	0,351	0,357	0,354	0,370	0,478	0,438
<u>374</u>								0,045	0,029	0,029	0,050	0,037						
376		0,081	0,062	0,077	0,042	0,054		0,136	0,086	0,029	0,100	0,056	0,135	0,036	0,092	0,111	0,022	0,055
<u>378</u>										0,029		0,019		0,036	0,015			
PR14																		
N	19	62	81	13	24	37	13	24	37	36	24	60	37	28	65	27	46	73
189	0,105	0,081	0,086											0,107	0,046			
193	0,053	0,081	0,074										0,054	0,036	0,046			
197	0,105	0,065	0,074	0,385	0,125	0,216	0,462	0,125	0,243	0,167	0,042	0,117						
201	0,421	0,565	0,531	0,462	0,667	0,595	0,385	0,667	0,568	0,694	0,833	0,750	0,459	0,214	0,354	<u>1,000</u>	<u>1,000</u>	<u>1,000</u>
205	0,316	0,194	0,222	0,154	0,208	0,189	0,154	0,208	0,189	0,139	0,083	0,117	0,459	0,500	0,477			
209		0,016	0,012								0,042	0,017	0,027	0,107	0,062			
<u>213</u>														<u>0,036</u>	<u>0,015</u>			
PC1																		
N	19	62	81	13	24	37	13	24	37	38	24	62	36	28	64	26	46	72
<u>210</u>										<u>0,026</u>		0,016						
<u>214</u>	0,211	0,081	0,111		0,042	0,027				0,053		0,032						
<u>218</u>	0,105	0,129	0,123	0,077	0,167	0,135	0,077	0,167	0,135	0,158	0,042	0,113	0,056	0,036	0,047		0,022	0,014
<u>222</u>	0,211	0,258	0,247	0,462	0,500	0,486	0,385	0,208	0,270	0,316	0,125	0,242	0,222	0,357	0,281	0,462	0,478	0,472
<u>226</u>	0,263	0,177	0,198	0,077	0,042	0,054	0,077	0,167	0,135	0,184	0,125	0,161	0,250	0,179	0,219	0,154	0,261	0,222
<u>230</u>	0,105	0,194	0,173	0,154	0,083	0,108	0,231	0,292	0,270	0,211	0,500	0,323	0,333	0,250	0,297	0,077	0,087	0,083
<u>234</u>	0,105	0,065	0,074	0,154	0,125	0,135	0,231	0,167	0,189	0,026	0,125	0,065	0,111	0,179	0,141	0,115	0,065	0,083

Locus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
<u>236</u>		0,016	0,012	0,077	0,042	0,054					0,083	0,032				0,154	0,065	0,097
<u>238</u>		<u>0,048</u>	<u>0,037</u>															
<u>240</u>		0,032	0,025							0,026		0,016	0,028		0,016	0,038	0,022	0,028
PC2																		
N	19	62	81	13	24	37	13	24	37	37	24	61	36	28	64	26	46	72
<u>110</u>		<u>0,016</u>	<u>0,012</u>															
<u>114</u>		<u>0,032</u>	<u>0,025</u>															
<u>118</u>		0,016	0,012								0,042	0,016						
<u>122</u>	0,263	0,323	0,309	0,385	0,250	0,297	0,385	0,292	0,324	0,189	0,333	0,246	0,389	0,357	0,375	0,385	0,174	0,250
<u>124</u>											<u>0,042</u>	<u>0,027</u>						
<u>126</u>	0,526	0,355	0,395	0,231	0,333	0,297	0,154	0,208	0,189	0,108	0,292	0,180	0,111	0,143	0,125		0,022	0,014
<u>130</u>	0,158	0,242	0,222	0,385	0,292	0,324	0,385	0,417	0,405	0,622	0,333	0,508	0,500	0,464	0,484	0,615	0,804	0,736
<u>134</u>	0,053	0,016	0,025		0,125	0,081	0,077	0,042	0,054	0,081		0,049		0,036	0,016			
PC3																		
N	19	62	81	13	24	37	13	24	37	38	24	62	36	28	64	26	46	72
<u>121</u>		0,016	0,012										0,028		0,016			
<u>125</u>	0,632	0,774	0,741	0,615	0,542	0,568	0,846	0,833	0,838	0,421	0,625	0,500	0,472	0,536	0,500	0,192	0,109	0,139
<u>129</u>	0,368	0,161	0,210	0,385	0,417	0,405	0,154	0,167	0,162	0,500	0,375	0,452	0,500	0,321	0,422	0,769	0,891	0,847
<u>133</u>		0,048	0,037		0,042	0,027				0,079		0,048		0,143	0,063	0,038		0,014

Lócus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
PC4																		
N	19	60	79	13	24	37	13	24	37	38	24	62	35	26	61	26	46	72
193		0,017	0,013								0,042	0,016		0,038	0,016			
197		0,033	0,025							0,053	0,042	0,048						
201		0,017	0,013	0,077	0,042	0,054				0,079	0,167	0,113	0,029	0,038	0,033	0,038		0,014
205	0,105	0,183	0,165	0,154	0,125	0,135				0,026	0,167	0,081	0,200	0,115	0,164	0,038		0,014
209	0,158	0,400	0,342	0,231	0,333	0,297	0,538	0,375	0,432	0,342	0,167	0,274	0,371	0,385	0,377	0,192	0,283	0,250
213	0,211	0,150	0,165	0,308	0,292	0,297	0,308	0,208	0,243	0,316	0,250	0,290	0,171	0,308	0,230	0,269	0,196	0,222
217	0,368	0,133	0,190	0,231	0,208	0,216	0,154	0,250	0,216	0,079	0,083	0,081	0,143	0,077	0,115	0,462	0,435	0,444
221	0,053	0,050	0,051					0,167	0,108	0,053	0,083	0,065	0,086	0,038	0,066		0,087	0,056
223	0,053		0,013															
225	0,053	0,017	0,025							0,053		0,032						
PC5																		
N	19	62	81	12	18	30	13	24	37	37	24	61	36	28	64	26	44	70
303																0,038		0,014
305	0,158	0,065	0,086	0,083	0,056	0,067				0,054	0,125	0,082		0,036	0,016			
310	0,105	0,032	0,049				0,077	0,083	0,081	0,027		0,016	0,028	0,036	0,031			
315	0,105	0,355	0,296	0,250	0,278	0,267	0,231	0,375	0,324	0,135	0,292	0,197	0,111	0,321	0,203	0,038	0,023	0,029
320	0,474	0,274	0,321	0,417	0,444	0,433	0,231	0,250	0,243	0,216	0,250	0,230	0,111		0,063	0,308	0,341	0,329
325	0,053	0,145	0,123	0,083	0,056	0,067				0,027	0,042	0,033	0,111	0,036	0,078	0,231	0,136	0,171
330	0,105	0,129	0,123	0,083	0,111	0,100	0,077	0,208	0,162	0,270	0,208	0,246	0,222	0,179	0,203	0,192	0,114	0,143
332																0,115	0,045	0,072

Lócus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
<u>335</u>				0,083		0,033	0,385	0,083	0,189	0,135	0,042	0,098	0,139	0,286	0,203		0,205	0,129
<u>340</u>					0,056	0,033				0,108	0,042	0,082	0,278	0,107	0,203	0,077	0,136	0,114
<u>350</u>										<u>0,027</u>		<u>0,016</u>						
PC6																		
N	19	60	79	12	22	34	13	22	35	33	24	57	36	28	64	26	46	72
<u>310</u>																	<u>0,065</u>	<u>0,042</u>
<u>314</u>		<u>0,017</u>	<u>0,013</u>															
<u>318</u>		0,033	0,025				0,077	0,091	0,086									
<u>322</u>	0,211	0,067	0,101	0,083	0,227	0,176	0,154		0,057	0,154		0,018		0,107	0,047	0,077	0,109	0,097
<u>326</u>	0,105	0,067	0,076	0,083		0,029	0,077	0,227	0,171	0,182	0,125	0,158	0,167	0,179	0,172	0,077	0,022	0,042
<u>330</u>	0,211	0,350	0,316	0,250	0,182	0,206	0,077	0,364	0,257	0,273	0,333	0,298	0,222	0,250	0,234	0,077	0,174	0,139
<u>332</u>		<u>0,053</u>	<u>0,013</u>															
<u>334</u>	0,316	0,250	0,266	0,333	0,227	0,265	0,538	0,182	0,314	0,273	0,333	0,298	0,417	0,286	0,375	0,346	0,196	0,250
<u>338</u>		0,100	0,076	0,083	0,136	0,118	0,077	0,045	0,057	0,212	0,125	0,175	0,139	0,143	0,141	0,192	0,130	0,153
<u>342</u>	0,053	0,100	0,089	0,167	0,227	0,206		0,045	0,029		0,042	0,018				0,231	0,239	0,236
<u>344</u>	0,053	0,017	0,025										0,056	0,036	0,031			
<u>346</u>							0,045	0,029	0,030	0,042	0,035						0,065	0,042
PH1																		
N	18	62	80	13	22	35	13	22	35	38	22	60	37	26	63	27	46	73
<u>114</u>	0,111	0,177	0,163	0,231	0,182	0,200	0,308	0,136	0,200	0,026	0,091	0,050	0,054		0,032			
<u>118</u>	0,222	0,113	0,138	0,077	0,091	0,086	0,231	0,273	0,257	0,526	0,409	0,483	0,622	0,538	0,587	0,259	0,261	0,260

Locus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
<u>122</u>		0,081	0,063		0,045	0,029				0,079		0,050	0,027	0,038	0,032	0,333	0,239	0,274
<u>126</u>	0,500	0,355	0,388	0,231	0,182	0,200	0,231	0,273	0,257	0,132	0,091	0,117	0,081	0,231	0,143			
<u>130</u>	0,111	0,210	0,188	0,385	0,500	0,457	0,154	0,273	0,229	0,211	0,318	0,250	0,189	0,192	0,190	0,407	0,391	0,397
<u>132</u>																	<u>0,043</u>	<u>0,027</u>
<u>134</u>	0,056	0,065	0,063	0,077		0,029	0,077	0,045	0,057	0,026	0,045	0,033	0,027		0,016		0,065	0,041
138											<u>0,045</u>	<u>0,017</u>						
PH2																		
N	18	62	80	13	22	35	13	22	35	38	22	60	37	26	63	27	46	73
143	0,889	0,887	0,887	0,923	0,818	0,857	0,462	0,727	0,629	0,342	0,500	0,400	0,351	0,346	0,349	0,444	0,500	0,493
146	0,056	0,065	0,063		0,045	0,029	0,385	0,091	0,200	0,079	0,182	0,117	0,108	0,077	0,095			
148	0,056	0,048	0,050	0,077	0,045	0,057	0,154	0,182	0,171	0,579	0,318	0,483	0,541	0,577	0,556	0,556	0,500	0,507
<u>152</u>					<u>0,090</u>	<u>0,058</u>												
PH3																		
N	18	62	80	13	24	37	13	24	37	28	22	60	37	26	63	27	46	73
<u>212</u>													<u>0,027</u>		<u>0,016</u>			
214		0,065	0,050		0,042	0,027				0,026		0,017	0,027		0,016			
216	0,444	0,242	0,287	0,077	0,250	0,189	0,154		0,054	0,079	0,091	0,083	0,108	0,192	0,143			
218	0,056	0,210	0,175	0,154	0,167	0,162	0,308	0,250	0,270	0,026	0,045	0,033	0,054		0,032			
220	0,500	0,355	0,388	0,692	0,333	0,459	0,385	0,542	0,486	0,289	0,545	0,383	0,243	0,423	0,317	0,407	0,522	0,479
222		0,097	0,075	0,077	0,208	0,162	0,154	0,167	0,162	0,526	0,318	0,450	0,486	0,346	0,429	0,593	0,478	0,521
224		0,016	0,013							0,053		0,033						

Lócus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
226		0,016	0,013					0,042	0,027				0,054	0,038	0,048			
PH4																		
N	18	62	80	13	22	35	13	24	37	37	20	57	37	26	63	27	42	69
138		0,016	0,013															
140	0,056	0,145	0,125	0,077	0,318	0,229		0,083	0,054	0,027	0,050	0,035				0,074	0,024	0,043
142	0,333	0,129	0,175	0,154	0,045	0,086	0,154	0,500	0,378	0,595	0,200	0,456	0,378	0,423	0,397	0,926	0,595	0,725
144	0,389	0,177	0,225	0,385	0,455	0,429	0,385	0,083	0,189	0,162	0,150	0,158	0,351	0,154	0,270		0,333	0,203
146	0,056	0,274	0,225	0,154	0,045	0,086	0,154	0,042	0,081	0,054	0,100	0,070	0,135	0,038	0,095		0,048	0,029
148		0,048	0,037	0,154	0,045	0,086	0,154	0,125	0,135	0,135	0,350	0,211	0,027	0,154	0,079			
150					0,045	0,029		0,167	0,108	0,027	0,050	0,018	0,054	0,038	0,048			
152	0,111	0,048	0,063									0,018	0,054	0,192	0,111			
154		0,113	0,087	0,077	0,045	0,057												
156	0,056	0,048	0,050				0,154		0,054		0,100	0,035						
PH5																		
N	18	62	80	13	22	35	13	22	35	38	22	60	37	26	63	27	46	73
221					0,045	0,029												
246	0,111	0,032	0,050										0,027	0,077	0,048			
251	0,833	0,887	0,875	0,462	0,409	0,429	0,846	0,636	0,714	0,421	0,545	0,467	0,405	0,462	0,429	0,481	0,478	0,479
254													0,027	0,038	0,032	0,074		0,027
256																	0,022	0,014
261		0,032	0,025	0,385	0,273	0,314	0,154	0,364	0,286	0,579	0,455	0,533	0,514	0,423	0,476	0,444	0,478	0,466

Locus/ Alelos	Africanos			Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total	H	M	Total
266	0,056	0,048	0,050	0,077	0,227	0,171							0,027		0,016			
<u>276</u>					<u>0,045</u>	<u>0,029</u>												
<u>346</u>				<u>0,077</u>		<u>0,029</u>												
<u>354</u>																	<u>0,022</u>	<u>0,014</u>
PH6																		
N	18	62	80	12	12	24	13	20	33	35	18	53	37	26	63	27	44	71
<u>274</u>		0,081	0,075	0,083	0,167	0,125												
<u>279</u>		<u>0,032</u>	<u>0,025</u>															
<u>284</u>	0,667	0,548	0,563	0,417	0,417	0,417	0,538	0,600	0,576	0,343	0,500	0,396	0,270	0,346	0,302	0,481	0,477	0,479
<u>289</u>	0,056	0,048	0,050	0,083	0,250	0,167				0,029	0,056	0,038	0,081		0,048			
<u>294</u>	0,278	0,290	0,287				0,462	0,300	0,364	0,429	0,333	0,396	0,595	0,654	0,619		0,023	0,014
<u>299</u>				0,417	0,167	0,292		0,100	0,061	0,200	0,111	0,170	0,054		0,032	0,519	0,500	0,507
PH7																		
N	18	62	80	11	12	23	13	16	29	32	16	48	37	26	63	27	44	71
342																	<u>0,023</u>	<u>0,014</u>
346		0,081	0,063	0,273	0,167	0,217	0,154	0,125	0,138	0,063	0,188	0,104	0,081	0,154	0,111		0,045	0,028
350	0,556	0,387	0,425	0,182	0,167	0,174	0,231	0,438	0,345	0,250	0,250	0,250	0,216	0,308	0,254	0,444	0,318	0,366
354	0,222	0,306	0,287	0,091	0,250	0,174	0,154	0,313	0,241	0,281	0,125	0,229	0,378	0,423	0,397	0,037	0,023	0,028
358	0,222	0,177	0,188	0,455	0,333	0,391	0,462	0,125	0,276	0,375	0,438	0,396	0,243	0,038	0,159	0,407	0,477	0,451
362		0,048	0,037		0,083	0,043				0,031		0,021	0,081	0,077	0,079	0,111	0,114	0,113

Frequências Alélicas

A genotipagem dos 20 loci, agrupados em três blocos (DMD, PC e HEMA), exibiu um total de 169 alelos, média de 8,45 alelos por locus, cujas frequências estão descritas na Tabela 3.

O locus que teve maior número de alelos foi o PR13, com 21 alelos, e o locus com menor número de alelos (três) foi o PR4. O bloco DMD apresentou o maior número de alelos (63), enquanto que nos blocos PC e HEMA ocorreram 55 e 52, respectivamente.

Considerando o conjunto de todas as populações estudadas, todos os 20 loci analisados apresentaram frequências polimórficas (Tabela 3). Entretanto, os loci PR4 e PR14, ambos do bloco DMD, foram monomórficos na população *Tikúna*, tendo sido fixados os alelos *PR4*202* e *PR14*201*. Ainda na população *Tikúna*, o alelo *PH4*142* apresentou frequência superior a 0,9. Em 18 dos 20 loci, o alelo de maior frequência na população Africana foi compartilhado com todas as outras populações estudadas.

Alguns alelos foram encontrados apenas nas amostras de mulheres (*PR10*146*, *PR10*148*, *PR14*213*, *PC1*238*, *PC2*110*, *PC2*114*, *PC2*124*, *PC2*118*, *PC4*193*, *PC6*310*, *PC6*314*, *PH1*138*, *PH2*152*, *PH4*138*, *PH5*221*, *PH5*254*, *PH5*276*, *PH5*354*, *PH6*279* e *PH7*342*) e outros, apenas nos homens (*PR1*241*, *PR1*244*, *PR10*114*, *PR10*146*, *PR10*148*, *PC1*210*, *PC4*223*, *PC5*303*, *PC5*350*, *PC6*332*, *PH3*212* e *PH5*346*). Essas diferenças existentes entre as amostras de homens e mulheres não é significativa ($P=0,0637$).

Alelos Privados

Os alelos que ocorrem em apenas uma população são chamados de alelos privados. Estes estão sublinhados e em negrito no corpo da Tabela 3. No bloco PC, ocorreram 26 alelos privados. A população Africana apresentou o maior número (sete) destes alelos: *PC1*238*, *PC2*110*, *PC2*114*, *PC4*223*, *PC6*314*, *PH4*138*, *PH6*279*.

No bloco DMD, ocorreram cinco alelos privados: *PR10*114* e *PR13*370* em Teresina e *PR10*146*, *PR10*148* e *PR14*213* em Ribeirão Preto. No bloco HEMA foram encontrados oito alelos privados em quatro das populações estudadas: Africana (*PH6*279*), Mimbó (*PH2*152*, *PH5*221*, *PH5*276*, *PH5*346*), *Tikúna* (*PH1*132*, *PH5*354*) e Ribeirão Preto (*PH3*212*).

Analisando os alelos privados em cada população, a Africana foi a que apresentou o maior número entre todas as populações (*PC1*238*, *PC2*110*, *PC2*114*, *PC4*223*, *PC6*314*, *PC6*332*, *PH4*138*, *PH6*279*). *Tikúna* e Teresina tiveram cinco alelos privados cada. Os alelos privados de Teresina são *PR10*114*, *PR13*362* e *PR13*370*, *PC5*350*, *PH1*138* e os de *Tikúna* são *PC5*303*, *PC5*332*, *PC6*310*, *PH5*256*, *PH5*354*. Duas populações possuem apenas quatro alelos privados: Mimbó (*PH2*152*, *PH5*122*, *PH5*276*, *PH5*346*) e Ribeirão

Preto (*PR10*146*, *PR10*148**, *PR14*213*, *PH3*212*). Sítio Velho foi a população com menor número deste tipo de alelo, apenas um (*PC2*124*).

Aderência ao equilíbrio de Hardy-Weinberg

O teste de equilíbrio de Hardy-Weinberg foi realizado nos 20 STRs nas seis populações, totalizando 120 testes. Destes, 12 indicaram desequilíbrio de Hardy-Weinberg (Tabela 4). Após a correção de Bonferroni apenas três loci continuaram fora do desequilíbrio.

Tabela 4: Teste exato de Fisher para a verificação do Equilíbrio de Hardy-Weinberg nas populações de mulheres. * Locus fora do Equilíbrio de Hardy-Weinberg depois da correção de Bonferroni ($\alpha_{Bonf} < 0,0025$).

Locus	Africana	Mimbó	Sítio Velho	Teresina	Ribeirão Preto	Tikúna
PR1	0,006	0,59	0*	0,528	0,643	0,484
PR4	0,785	0,361	0,192	1	0,397	-
PR8	0,846	1	0,257	1	0,472	1
PR9	0,002*	0,529	0,827	1	1	0,448
PR10	0,798	1	0,892	0,158	0,577	0,602
PR13	0,866	0,227	0,06	0,646	1	0,094
PR14	0,357	0,741	0,741	0,254	0,012	-
PC1	0,006	0,924	0,019	0,292	1	0,448
PC2	0,21	1	0,338	0,753	0,931	0,081
PC3	0,075	1	0,254	0,592	0,333	0,214
PC4	0,902	0,619	0,148	0,417	0,329	0,444
PC5	0,656	0,965	0,056	0,34	0,937	0,203
PC6	0,524	0,571	0,004	0,048	0,97	0,117
PH1	0,835	0,38	0,301	0,107	0,191	0,437
PH2	1	1	1	1	0,12	1
PH3	0,279	0,95	0,309	1	0,01	0,675
PH4	0,166	1	0,014	0,015	0*	0,724
PH5	1	0,79	0,537	1	0,485	0,675
PH6	0,21	0,053	0,056	0,403	0,578	0,668
PH7	0,815	0,183	0,477	0,582	0,027	0,127

Na população Africana apenas o locus PR9 continuou em desequilíbrio com déficit de heterozigotos. Sítio Velho manteve o desequilíbrio apenas no locus PR1 que também apresentou déficit de heterozigotos. Após a correção de Bonferroni, a população de Teresina não apresentou nenhum locus fora do equilíbrio de Hardy-Weinberg, e Ribeirão Preto manteve apenas um locus em desequilíbrio (PH4), onde foi identificado déficit de heterozigoto. Em Mimbó e Tikúna, todos os 20 loci ficaram em equilíbrio de Hardy-Weinberg.

Diferenciação genética das populações

A diferenciação genética das populações foi obtida através do Teste G de diferenciação gênica (Tabelas 5 e 6) e genotípica (Tabela 7). O teste G de diferenciação gênica foi realizado para homens e mulheres separadamente e fez comparações para a par de todas as populações para cada locus analisado. Os resultados encontrados nos homens e nas mulheres foram semelhantes, mas no geral o grupo dos homens apresentou menor diferenciação do que o grupo de mulheres.

Quando a diferenciação gênica foi analisada considerando apenas a população de mulheres (Tabela 5), a porcentagem de diferenças significativas foi de 60,4%. Depois da correção de Bonferroni esse valor caiu para 20%. *Tikúna* foi a população que mais divergiu de todas as outras, com 36% das diferenciações significativas. As duas populações mais divergentes entre si foram novamente *Tikúna* e Africana, com valores significativos em 65% dos loci. Já o par de populações mais semelhantes foi Teresina e Sitio Velho.

Tabela 5: Diferenciação Gênica entre as amostras populacionais (mulheres) comparadas par a par. AF: Africana; MB: Mimbó; SV: Sítio Velho; TR: Teresina; RP: Ribeirão Preto; TK: *Tikúna*. Valores significativos ($p < 0,05$) em negrito sublinhado. * Valores significativos após a correção de Bonferrone ($\alpha_{Bonf}=0,00016$).

Lócus	AFxMB	AFxSV	AFxTR	AFxRP	AFxTK	MBxSV	MBxTR	MBxRP	MBxTK	SVxTR	SVxRP	SVxTK	TRxRP	TRxTK	RPxTK
PR1	0,124	0,153	0,263	0,285	0,21	0,151	0,133	0,077	<u>0,017</u>	0,706	0,647	0,376	0,62	0,155	0,1
PR4	0,09	0,197	<u>0,001</u>	<u>0*</u>	<u>0*</u>	1	0,621	0,563	<u>0*</u>	0,335	0,421	<u>0*</u>	0,656	<u>0*</u>	<u>0*</u>
PR8	<u>0*</u>	0,26	<u>0,001</u>	<u>0*</u>	<u>0*</u>	0,125	0,416	0,24	1	0,331	0,176	<u>0,016</u>	0,276	0,157	<u>0,038</u>
PR9	<u>0,001</u>	<u>0*</u>	<u>0,009</u>	<u>0*</u>	<u>0*</u>	0,595	0,139	0,831	0,499	<u>0,003</u>	0,084	<u>0,049</u>	0,057	<u>0,017</u>	0,091
PR10	0,608	<u>0,037</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,217	<u>0,063</u>	0,106	<u>0*</u>	0,652	0,67	0,118	0,918	<u>0,025</u>	0,086
PR13	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,161	<u>0*</u>	0,222	<u>0,001</u>	<u>0*</u>	<u>0,028</u>	<u>0,01</u>	<u>0*</u>	0,097	<u>0*</u>	<u>0*</u>	<u>0*</u>
PR4	0,22	0,145	<u>0,043</u>	0,1	<u>0*</u>	1	0,298	<u>0,001</u>	<u>0*</u>	0,161	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0,002</u>	<u>0*</u>
PC1	0,255	0,241	0,482	0,075	<u>0,021</u>	1	0,448	0,309	0,754	0,447	0,611	0,29	0,293	<u>0,011</u>	<u>0,011</u>
PC2	0,205	0,21	<u>0,001</u>	<u>0,003</u>	<u>0*</u>	1	0,221	0,612	<u>0,039</u>	0,604	0,473	0,063	0,086	<u>0,037</u>	0,163
PC3	1	0,248	0,163	0,605	<u>0,004</u>	0,376	0,332	0,647	<u>0,031</u>	<u>0,026</u>	0,051	<u>0*</u>	0,267	0,119	<u>0,023</u>
PC4	0,882	0,147	0,089	0,189	0,662	0,313	0,496	0,636	0,681	0,581	0,199	0,118	0,094	<u>0,018</u>	<u>0,019</u>
PC5	0,755	<u>0,035</u>	0,09	<u>0*</u>	<u>0,024</u>	0,589	0,612	<u>0,048</u>	0,128	0,336	<u>0,042</u>	<u>0*</u>	0,366	<u>0,008</u>	<u>0,006</u>
PC6	0,82	0,703	<u>0,025</u>	<u>0,018</u>	0,07	0,683	0,327	0,163	0,869	0,144	0,104	0,277	0,444	<u>0,018</u>	<u>0,005</u>
PH1	0,298	0,685	<u>0,026</u>	<u>0,009</u>	<u>0*</u>	0,751	<u>0,015</u>	<u>0,011</u>	<u>0*</u>	0,059	0,081	<u>0*</u>	0,866	<u>0,002</u>	<u>0*</u>
PH2	1	0,049	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0,019</u>	<u>0*</u>	<u>0,001</u>	<u>0,005</u>	<u>0,009</u>	<u>0,037</u>	<u>0*</u>	0,943	0,221	0,151
PH3	0,099	0,062	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,536	<u>0,014</u>	<u>0,031</u>	<u>0,001</u>	<u>0,026</u>	0,115	<u>0*</u>	0,629	0,225	0,075
PH4	0,379	0,32	0,092	0,503	<u>0*</u>	0,702	0,091	0,147	<u>0*</u>	<u>0,024</u>	0,092	<u>0*</u>	0,067	<u>0,001</u>	<u>0*</u>
PH5	<u>0,003</u>	0,144	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,132	0,085	0,652	0,234	<u>0,011</u>	0,051	0,077	0,765	0,178	0,769
PH6	<u>0,002</u>	0,562	<u>0,035</u>	<u>0,039</u>	<u>0*</u>	<u>0,001</u>	<u>0,008</u>	<u>0*</u>	0,193	0,146	0,171	<u>0*</u>	0,207	<u>0*</u>	<u>0*</u>
PH7	<u>0,03</u>	0,113	0,193	0,066	0,071	1	0,334	0,141	<u>0,028</u>	0,791	0,331	0,084	0,684	<u>0,029</u>	<u>0,002</u>

Na análise apenas dos homens, as populações foram novamente comparadas par a par. Após a correção de Bonferroni 24,5% dessas comparações foram significativas (Tabela 6). No caso dos homens, as populações Sítio Velho X Mimbó foram as mais semelhantes, com apenas dois valores significativos antes da correção de Bonferroni e nenhuma depois dela. As populações que foram mais divergentes na comparação par a par foram Mimbó X *Tikúna* e Ribeirão Preto X *Tikúna*, com 11 diferenças significativas após a correção de Bonferroni. Comparando *Tikúna* com todas as outras populações, ela teve uma porcentagem de diferenciação menor nas mulheres, mas continuou sendo a população que mais se diferenciou (Tabela 6). No geral o grupo dos homens apresentou menor diferenciação do que o grupo de mulheres.

Tabela 6: Diferenciação Gênica, entre as amostras populacionais (homens) comparadas par a par. AF: Africana; MB: Mimbó; SV: Sítio Velho; TR: Teresina; RP: Ribeirão Preto; TK: *Tikúna*. Valores significativos ($p < 0,05$) em negrito sublinhado. * Valores significativos após a correção de Bonferrone ($\alpha_{Bonf}=0,00016$)

Lócus	AFxMB	AFxSV	AFxTR	AFxRP	AFxTK	MBxSV	MBxTR	MBxRP	MBxTK	SVxTR	SVxRP	SVxTK	TRxRP	TRxTK	RPxTK
PR1	<u>0,001</u>	<u>0,026</u>	0,348	<u>0,009</u>	<u>0*</u>	<u>0*</u>	0,299	<u>0,004</u>	<u>0*</u>	<u>0,011</u>	<u>0,004</u>	<u>0,004</u>	0,36	<u>0*</u>	<u>0*</u>
PR4	<u>0,005</u>	0,065	<u>0*</u>	<u>0,007</u>	<u>0*</u>	0,19	0,08	0,933	<u>0*</u>	<u>0,006</u>	0,124	<u>0*</u>	<u>0,029</u>	<u>0,037</u>	<u>0*</u>
PR8	0,089	0,101	<u>0,013</u>	<u>0*</u>	<u>0*</u>	0,276	0,822	0,438	0,103	0,393	0,063	<u>0,002</u>	0,818	0,212	0,745
PR9	<u>0,003</u>	<u>0*</u>	<u>0,007</u>	<u>0,011</u>	<u>0*</u>	0,067	<u>0,025</u>	0,155	1	<u>0,002</u>	0,261	0,074	0,078	<u>0,001</u>	0,095
PR10	0,64	0,353	0,137	0,165	<u>0,016</u>	0,235	0,065	0,071	<u>0,004</u>	0,711	0,499	0,119	0,357	<u>0,038</u>	<u>0*</u>
PR13	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,191	<u>0*</u>	0,054	<u>0*</u>	<u>0*</u>	<u>0,001</u>	<u>0,018</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>
PR14	0,206	0,203	0,094	<u>0,001</u>	<u>0*</u>	1	0,265	<u>0*</u>	<u>0*</u>	0,26	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0,011</u>	<u>0*</u>
PC1	0,226	0,362	<u>0,041</u>	0,145	<u>0,011</u>	0,14	<u>0,007</u>	0,119	0,096	0,324	0,525	<u>0,011</u>	0,168	<u>0,002</u>	0,178
PC2	0,439	0,396	0,905	0,171	<u>0*</u>	0,543	0,339	0,247	<u>0*</u>	0,799	0,882	<u>0,003</u>	0,42	<u>0*</u>	<u>0,015</u>
PC3	0,084	0,656	0,095	0,104	<u>0*</u>	0,061	0,769	0,431	<u>0*</u>	0,187	0,028	<u>0*</u>	0,142	<u>0*</u>	<u>0*</u>
PC4	0,588	0,072	0,256	0,821	<u>0*</u>	0,068	0,259	0,862	<u>0,011</u>	<u>0,015</u>	0,101	0,458	0,637	<u>0*</u>	<u>0,002</u>
PC5	0,466	<u>0,027</u>	0,256	<u>0*</u>	<u>0*</u>	0,189	0,892	<u>0*</u>	<u>0,01</u>	0,343	<u>0,011</u>	<u>0*</u>	<u>0,019</u>	<u>0,001</u>	<u>0*</u>
PC6	0,218	0,291	0,509	0,309	<u>0,016</u>	<u>0,002</u>	<u>0,015</u>	<u>0,02</u>	0,578	0,683	0,157	<u>0,002</u>	0,445	<u>0,034</u>	<u>0,002</u>
PH1	0,156	0,349	<u>0,006</u>	<u>0*</u>	<u>0*</u>	0,278	0,086	<u>0,001</u>	<u>0*</u>	0,729	0,088	<u>0*</u>	0,156	<u>0,001</u>	<u>0*</u>
PH2	0,37	0,315	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,461	<u>0,022</u>	<u>0*</u>	<u>0*</u>	0,37	<u>0,018</u>	<u>0,006</u>	0,206	<u>0,009</u>	0,109
PH3	0,892	<u>0,021</u>	<u>0,029</u>	<u>0,004</u>	<u>0*</u>	<u>0,034</u>	0,199	0,104	<u>0*</u>	0,07	<u>0,002</u>	<u>0*</u>	0,55	<u>0,046</u>	<u>0,003</u>
PH4	<u>0,017</u>	<u>0*</u>	<u>0,009</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0,007</u>	<u>0*</u>	<u>0*</u>	0,168	0,099	<u>0,001</u>	0,061	<u>0*</u>	<u>0*</u>
PH5	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0,029</u>	<u>0,029</u>	<u>0,015</u>	<u>0*</u>	0,759	0,396	0,687	0,564	0,908	0,201
PH6	<u>0,001</u>	0,061	0,106	<u>0,008</u>	<u>0*</u>	<u>0,005</u>	<u>0,032</u>	<u>0*</u>	<u>0*</u>	0,839	<u>0,018</u>	<u>0*</u>	<u>0,042</u>	<u>0*</u>	<u>0*</u>
PH7	0,553	0,841	0,102	0,291	<u>0*</u>	0,373	0,817	0,261	0,152	0,173	0,577	<u>0,003</u>	<u>0,012</u>	0,172	<u>0*</u>

A diferenciação genotípica entre as populações foi calculada apenas nas amostras de mulheres (Tabela 7). Aproximadamente 28% das diferenciações foram estatisticamente significativas após a correção de Bonferroni. Semelhante ao que foi encontrado na diferenciação gênica das mulheres, as populações mais divergentes foram Africana e *Tikúna*, com 80% das comparações significativamente diferentes e as mais semelhantes foram novamente Sítio Velho e Teresina, com nenhuma diferença significativa após a correção de Bonferroni.

Tabela 7: Diferenciação Genotípica entre as mulheres de todas as amostras populacionais comparadas par a par. AF: Africana; MB: Mimbó; SV: Sítio Velho; TR: Teresina; RP: Ribeirão Preto; TK: *Tikúna*. Valores significativos ($p < 0,05$) em negrito sublinhado.* Valores significativos após a correção de Bonferrone ($\alpha_{Bonf}=0,00016$)

Lócus	AFxMB	AFxSV	AFxTR	AFxRP	AFxTK	MBxSV	MBxTR	MBxRP	MBxTK	SVxTR	SVxRP	SVxTK	TRxRP	TRxTK	RPxTK
PR1	<u>0,01</u>	0,176	0,557	<u>0,027</u>	<u>0,002</u>	<u>0*</u>	0,362	<u>0,004</u>	<u>0*</u>	0,184	0,069	<u>0,049</u>	0,361	<u>0,001</u>	<u>0*</u>
PR4	<u>0,004</u>	0,08	<u>0*</u>	<u>0,007</u>	<u>0*</u>	0,247	0,081	0,929	<u>0*</u>	<u>0,02</u>	0,167	<u>0*</u>	<u>0,032</u>	<u>0,033</u>	<u>0*</u>
PR8	0,081	0,117	<u>0,011</u>	<u>0*</u>	<u>0*</u>	0,307	0,804	0,418	0,099	0,434	0,089	<u>0,003</u>	0,812	0,214	0,739
PR9	<u>0,013</u>	<u>0*</u>	<u>0,023</u>	0,032	<u>0*</u>	0,062	<u>0,016</u>	0,118	1	<u>0,002</u>	0,234	<u>0,048</u>	0,059	<u>0,001</u>	0,065
PR10	0,645	0,365	0,14	0,218	<u>0,025</u>	0,224	<u>0,043</u>	0,108	<u>0,005</u>	0,677	0,64	0,123	0,462	<u>0,034</u>	<u>0,002</u>
PR13	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,18	<u>0*</u>	0,107	<u>0*</u>	<u>0*</u>	<u>0,002</u>	0,054	<u>0*</u>	<u>0,003</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>
PR14	0,262	0,257	0,151	<u>0,012</u>	<u>0*</u>	1	0,268	<u>0,002</u>	<u>0*</u>	0,271	<u>0,001</u>	<u>0*</u>	<u>0*</u>	<u>0,035</u>	<u>0*</u>
PC1	0,294	0,477	0,079	0,217	<u>0,027</u>	0,148	<u>0,009</u>	0,089	0,088	0,281	0,474	<u>0,019</u>	0,143	<u>0,004</u>	0,179
PC2	0,407	0,361	0,892	0,162	<u>0*</u>	0,453	0,286	0,209	<u>0*</u>	0,741	0,858	<u>0,001</u>	0,356	<u>0*</u>	<u>0,014</u>
PC3	0,118	0,681	0,131	0,189	<u>0*</u>	0,077	0,751	0,534	<u>0*</u>	0,217	0,075	<u>0*</u>	0,252	<u>0*</u>	<u>0*</u>
PC4	0,732	0,159	0,38	0,898	<u>0*</u>	0,071	0,266	0,842	<u>0,005</u>	<u>0,038</u>	0,141	0,451	0,691	<u>0*</u>	<u>0,001</u>
PC5	0,475	0,056	0,321	<u>0*</u>	<u>0*</u>	0,297	0,918	<u>0*</u>	<u>0,018</u>	0,598	<u>0,021</u>	<u>0*</u>	0,051	<u>0,005</u>	<u>0*</u>
PC6	0,225	0,306	0,549	0,332	<u>0,026</u>	<u>0,003</u>	<u>0,012</u>	<u>0,011</u>	0,598	0,697	0,134	<u>0,007</u>	0,398	0,053	<u>0,006</u>
PH1	0,212	0,437	<u>0,013</u>	<u>0*</u>	<u>0*</u>	0,325	0,145	<u>0,003</u>	<u>0*</u>	0,761	0,13	<u>0*</u>	0,2	<u>0,004</u>	<u>0*</u>
PH2	0,351	0,28	<u>0*</u>	<u>0*</u>	<u>0*</u>	0,409	<u>0,023</u>	<u>0,002</u>	<u>0*</u>	0,341	<u>0,036</u>	<u>0,008</u>	0,224	<u>0,007</u>	0,135
PH3	0,894	<u>0,028</u>	<u>0,032</u>	<u>0,014</u>	<u>0*</u>	<u>0,027</u>	0,18	0,144	<u>0*</u>	0,067	<u>0,013</u>	<u>0*</u>	0,658	0,051	<u>0,026</u>
PH4	<u>0,032</u>	<u>0*</u>	<u>0,025</u>	<u>0*</u>	<u>0*</u>	<u>0,003</u>	<u>0,027</u>	<u>0,001</u>	<u>0*</u>	0,353	0,293	<u>0,004</u>	0,205	<u>0*</u>	<u>0,003</u>
PH5	<u>0*</u>	<u>0,001</u>	<u>0*</u>	<u>0*</u>	<u>0*</u>	<u>0,025</u>	<u>0,017</u>	<u>0,012</u>	<u>0*</u>	0,774	0,423	0,679	0,56	0,905	0,194
PH6	<u>0,002</u>	0,096	0,087	<u>0,007</u>	<u>0*</u>	<u>0,025</u>	0,087	<u>0*</u>	<u>0*</u>	0,892	0,055	<u>0*</u>	0,057	<u>0*</u>	<u>0*</u>
PH7	0,501	0,804	0,077	0,324	<u>0*</u>	0,367	0,853	0,453	0,145	0,175	0,818	<u>0,001</u>	<u>0,041</u>	0,197	<u>0*</u>

Diversidade Gênica

A heterozigose observada (H_o) média e a heterozigose esperada média (H_s) foram obtidas nas seis populações (Tabela 8).

Tabela 8: Heterozigose média observada e esperada nas mulheres de todas as populações.

População	H_o Média	H_e Média (H_s)
Africana	0,63065 ± 0,17887	0,67084 ± 0,19208
Mimbó	0,65833 ± 0,19167	0,68370 ± 0,12408
Sítio Velho	0,53333 ± 0,17361	0,69457 ± 0,14209
Teresina	0,60417 ± 0,17255	0,69149 ± 0,18215
Ribeirão Preto	0,58214 ± 0,19984	0,70146 ± 0,11203
Tikúna	0,49348 ± 0,22351	0,53836 ± 0,24270

A H_s nas populações Africana, Mimbó, Sítio Velho, Teresina e Ribeirão Preto foram semelhantes e variou entre 0,67084 (± 0,19208) e 0,70146 (± 0,11203). O menor valor de H_s (0,53836 ± 0,24270) foi o que mais diferenciou dos outros valores e pertence à *Tikúna*. A população com maior H_s é Ribeirão Preto (0,70146 ± 0,11203).

O menor valor da heterozigose observada (H_o) média foi em *Tikúna* (0,49348 ± 0,22351) e o maior valor foi encontrado em Mimbó (0,65833 ± 0,19167).

Análise da distância genética entre as populações

A diversidade gênica também foi medida entre os pares de amostras populacionais pela estatística F_{ST} na população feminina (valores abaixo da diagonal na Tabela 9).

Tabela 9: Valores de F_{ST} em todos os loci nos homens (acima da diagonal) e nas mulheres (abaixo da diagonal). **Valores não significativos depois da correção de Bonferroni ($\alpha_{Bonf} < 0,0033$)

Populações	Africana	Mimbó	Sítio Velho	Teresina	Rib. Preto	<i>Tikúna</i>
Africana	---	0,06883	0,07212	0,12595	0,12442	0,22397
Mimbó	0,07930	---	0,01958**	0,06220	0,08945	0,11933
Sítio Velho	0,04337	0,05231	---	0,06818	0,07833	0,16609
Teresina	0,08222	0,04508	0,02233**	---	0,01864	0,06280
Rib. Preto	0,08000	0,09649	0,05254**	0,04933**	---	0,15162
<i>Tikúna</i>	0,19077	0,12992	0,12650	0,09662	0,11321	---

Dentre os valores significativos de F_{ST} ($P < 0,05$) na amostra das mulheres (valores abaixo da diagonal na Tabela 9), foi possível observar que as populações mais semelhantes entre si foram Sítio Velho e Teresina, pois a comparação entre elas gerou o menor F_{ST} (0,02233). Já o maior F_{ST} (0,19077), foi gerado na comparação entre as populações *Tikúna* e Africana, sendo estas as populações mais divergentes como mostrado anteriormente pelo teste de diferenciação populacional (Teste G).

Após esta correção, a comparação par a par entre as populações Sítio Velho X Teresina, Sítio Velho X Ribeirão Preto e Teresina X Ribeirão Preto não foram estatisticamente significativas ($P > 0,0033$), ou seja, estas podem ser consideradas populações semelhantes entre si. Os valores mais altos de F_{ST} foram obtidos quando *Tikúna* foi comparada par a par com as outras populações exceto com Teresina. Estes valores variaram de 0,1265 (*Tikúna* x Sítio Velho) e 0,19077 (*Tikúna* x Africana).

Na análise dos homens (valores acima da diagonal na Tabela 9), o F_{ST} entre Mimbó e Sítio velho foi o único não significativo e estas populações foram então consideradas as mais semelhantes. Não houve nenhuma alteração na significância dos valores de F_{ST} após a correção de Bonferroni.

Os valores de F_{ST} geral para homens (0,09101) e mulheres (0,10193) foram muito semelhantes.

Análise de Variância Molecular (AMOVA)

Em um primeiro momento, a AMOVA foi realizada colocando as amostras de remanescentes de quilombos em um grupo, as populações urbanas em outro e os ameríndios em um terceiro. A variância entre estes grupos não foi significativa ($P > 0,001$). Quando as cinco populações foram colocadas em um único grupo, a análise revelou que a variação entre os indivíduos de cada população foi maior (89.81% e 90.98%) do que a variação entre as populações (10.19% e 9.1%) tanto em mulheres quanto em homens, respectivamente.

Inferência de Estrutura Populacional

A inferência da estrutura populacional foi realizada com mulheres e homens separadamente (Tabela 10 e Figura 6 e 7). Nos dois casos o número de grupos com semelhança genética (*clusters*) foi três. A divisão das populações em cada *cluster* e os valores da proporção da composição genética em cada um deles também foi semelhante entre homens e mulheres (Tabela 10).

Tabela 10: Proporção da composição genética da amostra de mulheres e homens de todas as populações divididas em três clusters.

População	Mulheres			Homens		
	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
Africana	0,836	0,029	0,134	0,948	0,023	0,029
Mimbó	0,674	0,226	0,100	0,669	0,237	0,094
Sítio Velho	0,423	0,251	0,326	0,763	0,086	0,151
Teresina	0,273	0,247	0,480	0,255	0,294	0,451
Ribeirão Preto	0,248	0,050	0,702	0,354	0,076	0,570
<i>Tikúna</i>	0,025	0,916	0,059	0,035	0,858	0,106

Nas mulheres as populações que formaram, em grande parte, o primeiro *cluster* (*cluster* 1) foram: Africana (83,6% dos indivíduos africanos foram alocados neste *cluster*), Mimbó (67,4%) e Sítio Velho (42,3%) (Tabela 10). O segundo *cluster* é formado pela maior parte dos indivíduos de *Tikúna* (91,6%). A composição mais significativa no terceiro *cluster* foi das populações urbanas Ribeirão Preto (70,2% dos indivíduos foram agrupados no *cluster* 3) e Teresina (48%).

Na análise realizada apenas com os homens (Tabela 10) a divisão de cada *cluster* foi semelhante ao resultado encontrado nas mulheres. O *cluster* 1 foi composto por maior parte dos indivíduos da população Africana (94,8%), Mimbó (66,9%) e Sítio Velho (76,3%), O *cluster* 2 foi em maior parte formado por indivíduos de *Tikúna* (85,8% dos indivíduos de *Tikúna* foram agrupados no *cluster* 2). O terceiro *cluster* foi composto em sua maioria por indivíduos de Ribeirão Preto (57%) e Teresina (45,1%) (Tabela 10).

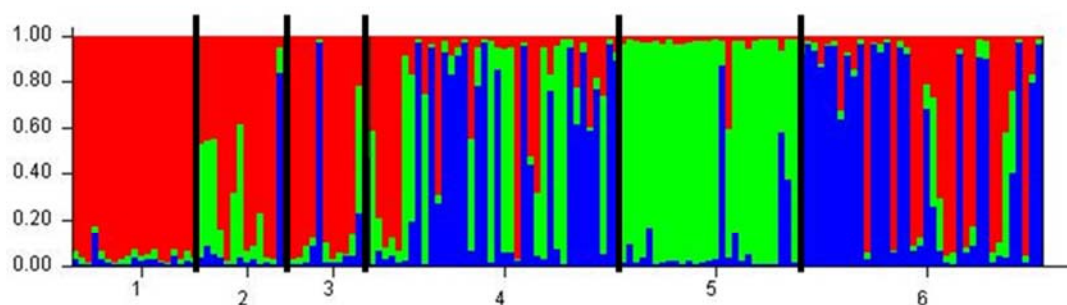


Figura 6: Proporção genética dos homens de todas as populações divididas em *clusters* (K=3). 1: Africanos; 2: Mimbó; 3: Sítio Velho; 4: Teresina; 5: *Tikúna*; 6: Ribeirão Preto

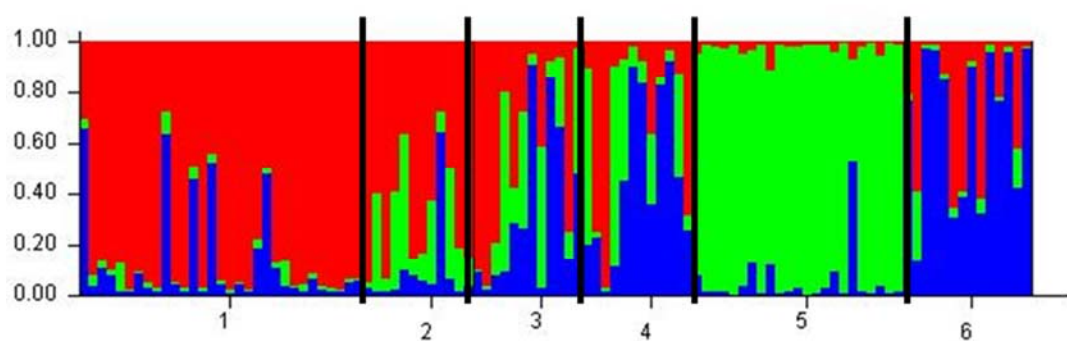


Figura 7: Proporção genética das mulheres de todas as populações divididas em *clusters* (K=3). 1: Africanos; 2: Mimbó; 3: Sítio Velho; 4: Teresina; 5: *Tikúna*; 6: Ribeirão Preto

Na figura 6 e 7 podemos ver o diagrama da proporção genética dos três *clusters* nos homens e mulheres, respectivamente. Nos dois grupos de amostras (homens e mulheres), a interpretação dos diagramas é a mesma, a única variação é na proporção da composição de cada *cluster*. A cor vermelha representa o componente africano (*cluster* 1), a verde representa o componente ameríndio (*cluster* 2) e a azul, o europeu (*cluster* 3). As populações

Africana, Mimbó e Sítio Velho são compostas, na sua maior parte, pelo componente africano (cor vermelha). A população Tikúna é a que, em sua composição, possui maior parte do componente ameríndio (cor verde) e Teresina e Ribeirão Preto apresentam uma maior proporção do terceiro componente, que seria o componente europeu (cor azul).

Diversidade Haplotípica

A análise da diversidade haplotípica (h) foi determinada para cada bloco: DMD, PC HEMA (Tabela 11).

Tabela 11: Diversidade haplotípica dos três blocos em todas as populações

Populações	DMD	PC	HEMA
Africana	0,683 ± 0,398	0,736 ± 0,424	0,529 ± 0,314
Mimbó	0,668 ± 0,393	0,750 ± 0,447	0,618 ± 0,373
Sítio Velho	0,719 ± 0,423	0,660 ± 0,397	0,665 ± 0,392
Teresina	0,666 ± 0,373	0,731 ± 0,412	0,627 ± 0,354
Rib. Preto	0,646 ± 0,362	0,720 ± 0,405	0,647 ± 0,362
Tikúna	0,419 ± 0,253	0,659 ± 0,380	0,513 ± 0,299

O bloco PC apresentou os maiores valores de diversidade haplotípica e o bloco HEMA, os menores (Tabela 11). Na análise do bloco PC, foi possível verificar que Mimbó foi a população com maior h (0,750 ± 0,447) e Tikúna a que apresentou menor h (0,659 ± 0,380). No bloco DMD, a população que teve a maior diversidade haplotípica foi Sítio Velho (0,719 ± 0,423) e a menor h foi Tikúna (0,419 ± 0,253). No bloco HEMA, os valores da diversidade haplotípica foram mais semelhantes aos do DMD, mas um pouco menores. A população que apresentou maior h também foi Sítio Velho (0,665 ± 0,392) e como nos outros dois blocos, Tikúna, mais uma vez, foi a população que apresentou o menor h (0,513 ± 0,299).

Os haplótipos foram determinados para cada bloco separadamente e apenas na amostra de homens, (Tabelas 12-14).

Nos 133 homens utilizados na análise do bloco DMD (Tabela 12), foram encontrados 123 haplótipos. Este bloco foi o que teve a menor proporção entre indivíduos e haplótipos (1,081). Oito haplótipos (D17, D20, D42, D70, D71, D72, D77, D82) foram compartilhados entre dois indivíduos cada um. Com exceção do haplótipo D20 que foi compartilhado entre um indivíduo de Mimbó e outro de Teresina, os outros foram compartilhados entre dois indivíduos de uma mesma população. O haplótipo D75 foi o único compartilhado por três indivíduos (TK07, TK11, TK17), todos da tribo Tikúna.

Tabela 12: Haplótipos do bloco DMD encontrados nas amostras populacionais de Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e *Tikúna*.

Indivíduos	Haplótipo	PR1	PR4	PR8	PR9	PR10	PR13	PR14
AF08	D1	235	197	228	268	130	363	197
AF18	D2	244	207	228	259	130	371	205
AF20	D3	193	197	228	259	130	363	205
AF23	D4	196	197	228	264	138	359	193
AF24	D5	193	202	222	259	126	367	205
AF26	D6	181	197	222	264	138	367	201
AF29	D7	199	197	228	259	130	363	201
AF33	D8	199	207	222	264	138	367	201
AF39	D9	208	207	228	259	138	371	201
AF40	D10	214	207	228	259	138	363	201
AF42	D11	208	197	228	264	138	367	205
AF45	D12	244	207	228	259	138	371	201
AF46	D13	202	197	216	259	134	371	205
AF48	D14	193	207	222	264	130	356	201
M01	D15	193	202	216	264	130	356	201
M02	D16	202	197	222	264	134	368	201
M03; M04	D17	193	202	216	264	130	356	197
M05	D18	205	207	216	258	130	368	205
M06	D19	196	197	216	268	138	376	201
M07; T22	D20	202	202	216	264	134	372	201
M08	D21	205	207	222	264	138	372	205
M09	D22	202	202	216	264	126	368	201
M10	D23	211	202	222	258	138	356	201
M11	D24	193	197	222	264	138	368	197
M12	D25	196	197	222	258	130	368	197
M13	D26	196	197	216	264	122	368	197
SV01	D27	241	197	222	258	138	368	197
SV02	D28	241	197	222	258	138	368	201
SV03	D29	196	207	216	264	130	372	197
SV04	D30	205	207	216	258	134	368	201
SV05	D31	202	202	216	258	134	368	205
SV06	D32	229	202	216	264	134	364	197
SV08	D33	181	197	222	264	138	356	205
SV09	D34	226	197	222	258	138	368	201
SV10	D35	196	197	228	268	130	364	201
SV11	D36	220	202	222	264	138	372	197
SV12	D37	232	202	222	264	134	372	197
SV13	D38	238	202	228	258	134	372	197
T01	D39	232	202	216	264	134	360	201
T02	D40	226	202	222	264	130	370	201
T03	D41	199	202	216	267	134	364	201
T04; T41	D42	190	202	216	264	134	368	201
T05	D43	232	207	222	264	138	366	205
T06	D44	214	202	222	258	134	360	201
T07	D45	205	202	216	264	134	366	201
T08	D46	220	197	216	264	134	362	197
T09	D47	235	202	216	264	130	378	201
T10	D48	193	197	216	264	134	370	205
T11	D49	211	197	222	264	130	370	197
T13	D50	211	207	222	264	130	372	205
T14	D51	202	207	222	264	134	362	205
T15	D52	223	197	216	264	134	370	201
T17	D53	226	202	222	259	138	374	201
T18	D54	229	202	216	264	134	370	201
T19	D55	238	202	216	264	130	366	201
T20	D56	196	202	216	269	138	368	201
T21	D57	232	197	216	264	134	366	197
T23	D58	196	202	228	259	126	376	201
T24	D59	190	202	216	264	134	364	201
T38	D60	181	207	228	259	138	356	201
T40	D61	211	207	222	264	134	366	205
T42	D62	235	202	222	258	130	368	201

Indivíduos	Haplótipo	PR1	PR4	PR8	PR9	PR10	PR13	PR14
T43	D63	208	197	222	264	134	366	197
T44	D64	193	197	222	259	134	372	197
T45	D65	229	202	216	264	134	366	201
T46	D66	181	197	228	259	138	370	201
T48	D67	220	202	216	267	134	366	201
T50	D68	202	202	228	258	138	362	201
RP01	D69	232	202	216	264	134	363	205
RP03	D70	193	202	216	264	134	371	201
RP04	D71	208	197	222	259	134	371	201
RP05	D72	220	202	216	254	134	363	205
RP06	D73	223	202	216	264	134	375	205
RP09	D74	223	202	228	258	138	375	201
RP12	D75	190	202	216	264	134	367	205
RP13	D76	208	207	222	264	134	359	193
RP15	D77	226	202	216	264	134	375	205
RP16	D78	220	202	216	264	134	367	205
RP17	D79	199	197	216	269	126	367	201
RP19	D80	220	202	216	254	134	371	205
RP20	D81	235	197	222	264	134	371	205
RP22	D82	220	202	228	264	134	371	201
RP23	D83	193	197	228	258	134	363	205
RP27	D84	202	202	216	264	138	367	201
RP36	D85	214	202	216	264	134	367	205
RP37	D86	214	202	228	258	126	375	193
RP39	D87	223	207	216	264	134	371	201
RP41	D88	217	197	228	258	138	367	201
RP42	D89	232	202	216	268	134	367	205
RP43	D90	205	197	216	264	134	367	209
RP66	D91	193	202	222	264	146	363	201
RP68	D92	202	197	222	266	148	363	201
RP70	D93	229	202	210	268	138	371	205
RP78	D94	232	202	216	264	138	367	205
RP79	D95	202	202	216	264	134	371	201
RP80	D96	226	202	216	264	134	371	201
RP81	D97	241	202	216	264	130	359	205
RP82	D98	196	197	222	264	134	371	201
RP83	D99	250	202	216	264	130	375	201
RP84	D100	232	202	216	264	130	363	205
RP85	D101	220	197	216	264	134	363	205
RP86	D102	217	197	216	264	130	371	201
RP89	D103	220	202	222	269	138	371	205
RP90	D104	229	202	216	264	122	371	201
TK01	D105	196	202	216	264	134	372	201
TK02; TK15	D106	232	202	216	264	134	368	201
TK03; TK27	D107	211	202	222	258	134	376	201
TK04; TK26	D108	208	202	222	258	134	372	201
TK05	D109	229	202	216	264	134	368	201
TK06	D110	202	202	222	264	134	372	201
TK07; TK11; K17	D111	244	202	216	264	138	368	201
TK08	D112	235	202	216	264	134	368	201
TK09; TK23	D113	205	202	216	264	134	372	201
TK10	D114	199	202	216	264	134	368	201
TK12	D115	202	202	222	258	134	372	201
TK13	D116	226	202	222	258	134	368	201
TK14	D117	238	202	222	264	134	366	201
TK16; TK20	D118	208	202	216	264	138	368	201
TK18	D119	217	202	216	264	138	372	201
TK19	D120	238	202	216	264	134	372	201
TK21	D121	235	202	222	258	138	368	201
TK22	D122	232	202	222	264	138	372	201
TK24	D123	205	202	216	264	138	368	201

Na análise de haplótipos do bloco de lócus pericentroméricos (PC) foram utilizados 137 homens, foram encontrados 124 haplótipos (Tabela 13), e a proporção indivíduos/haplótipos foi 1,104. Dos 12 haplótipos compartilhados, apenas um apareceu em três indivíduos e os outros 11 foram compartilhados por dois. Apenas três destes haplótipos foram compartilhados entre indivíduos de populações diferentes: C19 compartilhado entre indivíduos de Mimbó e Teresina; C59 entre indivíduos de Teresina e *Tikúna* e C81 foi compartilhado entre indivíduos de *Tikúna* e Ribeirão Preto. O único haplótipo que foi compartilhado por três indivíduos foi o C5 e ocorreu apenas na população Africana.

Tabela 13: Haplótipos do bloco PC encontrados nas amostras populacionais de Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e *Tikúna*.

Indivíduos	Haplótipos	PC1	PC2	PC3	PC4	PC5	PC6
AF02; AF17	C1	226	130	125	213	330	326
AF08	C2	222	130	125	213	320	322
AF10	C3	218	122	125	221	320	334
AF15	C4	226	122	125	209	310	342
AF18; AF26; AF45	C5	214	126	129	217	320	334
AF20	C6	226	126	125	225	315	330
AF23	C7	222	126	125	213	320	344
AF24	C8	226	126	129	217	320	330
AF29	C9	222	122	125	205	320	332
AF33	C10	214	126	129	205	310	334
AF35	C11	222	126	125	217	305	322
AF39	C12	234	122	125	209	315	330
AF40	C13	230	126	125	223	305	334
AF42	C14	218	126	129	209	325	322
AF46	C15	234	134	129	217	320	330
AF48	C16	230	122	125	217	305	322
M01; M03	C17	222	130	129	209	315	342
M02	C18	222	122	125	217	320	334
M04; T46	C19	218	126	125	213	305	330
M05	C20	230	126	125	205	320	322
M07	C21	236	126	125	213	330	334
M08	C22	226	122	125	201	320	334
M09	C23	234	122	125	209	315	326
M10	C24	222	130	129	213	320	330
M11	C25	234	122	125	217	325	334
M12	C26	230	130	129	213	335	330
M13	C27	222	122	125	205	320	338
SV01	C28	222	130	129	213	320	334
SV02; SV09	C29	234	122	125	209	335	334
SV03	C30	230	130	125	217	310	318
SV04	C31	222	122	125	209	315	322
SV05	C32	234	130	125	209	315	334
SV06	C33	222	130	125	209	315	326
SV07	C34	226	130	129	209	335	334
SV08	C35	222	122	125	213	320	334
SV10	C36	230	122	125	209	335	338
SV11	C37	218	126	125	213	330	330
SV12	C38	222	126	125	213	320	322
SV13	C39	230	134	125	217	335	334
T01	C40	226	130	129	221	320	338
T02	C41	222	134	133	217	320	334
T03	C42	230	122	125	213	340	334
T06	C43	218	130	129	217	320	330
T07	C44	218	130	125	209	330	338
T10	C45	222	126	129	209	320	326
T13	C46	230	130	129	201	330	334
T14	C47	240	130	125	225	325	326
T15	C48	230	122	125	209	335	330

Indivíduos	Haplótipos	PC1	PC2	PC3	PC4	PC5	PC6
T16	C49	230	130	125	213	330	330
T17	C50	222	130	129	197	330	338
T18	C51	226	126	125	209	340	334
T19	C52	222	130	129	213	335	346
T20	C53	222	130	129	209	310	326
T21; T39	C54	218	130	129	213	330	334
T22	C55	226	130	129	209	335	326
T23	C56	234	134	125	197	315	330
T24	C57	210	130	125	209	315	330
T25	C58	226	122	125	209	320	330
T38; TK11	C59	222	130	129	209	340	334
T40; T50	C60	222	122	125	209	315	326
T41	C61	226	130	129	213	335	334
T42	C62	222	130	129	213	330	338
T43	C63	230	122	125	209	330	338
T44	C64	214	130	129	217	320	334
T45	C65	214	126	129	205	320	330
T47	C66	226	134	133	213	340	322
T48	C67	222	130	129	213	315	338
T49	C68	230	130	129	209	335	330
RP03	C69	230	122	125	217	340	334
RP04	C70	226	122	125	209	330	334
RP05	C71	230	130	125	209	340	326
RP06	C72	230	122	125	213	310	330
RP09	C73	240	122	125	213	340	338
RP12	C74	230	122	125	209	340	338
RP13	C75	230	122	121	209	330	326
RP15	C76	230	122	125	217	330	330
RP16	C77	230	130	129	213	340	334
RP17	C78	226	130	129	205	340	334
RP19	C79	234	122	125	205	335	330
RP20	C80	222	122	125	201	320	334
RP22	C81	230	122	125	205	340	334
RP23	C82	234	122	125	205	315	330
RP27	C83	226	130	129	205	325	338
RP36	C84	222	130	129	209	330	344
RP37	C85	230	130	129	209	340	326
RP39	C86	230	122	125	213	340	344
RP41	C87	218	130	129	221	330	334
RP42	C88	222	130	129	209	340	326
RP59	C89	234	126	125	213	320	334
RP66	C90	226	130	129	221	335	334
RP68	C91	226	126	125	217	315	326
RP70	C92	226	126	129	209	335	330
RP78	C93	234	122	125	209	325	334
RP79	C94	226	130	129	217	330	334
RP81	C95	222	126	129	213	315	334
RP82	C96	230	130	129	209	330	338
RP83	C97	222	130	129	217	320	330
RP84	C98	222	130	129	205	335	338
RP85	C99	226	130	129	209	325	326
RP86	C100	218	130	129	221	320	330
RP89	C101	226	130	125	209	315	330
RP90	C102	222	130	129	205	335	334
TK01; TK12	C103	222	130	129	217	320	322
TK02; TK03	C104	222	130	129	217	320	326
TK04	C105	236	122	125	213	330	334
TK05; TK10	C106	226	130	129	217	325	334
TK06	C107	234	122	129	213	332	342
TK07	C108	234	122	129	213	303	342
TK08	C109	222	130	129	217	332	338
TK09	C110	226	130	129	217	315	330
TK13	C111	222	130	129	213	320	342
TK14	C112	222	130	129	209	325	330
TK15	C113	236	122	129	209	340	338
TK16	C114	240	122	125	201	330	334

Indivíduos	Haplótipos	PC1	PC2	PC3	PC4	PC5	PC6
TK17; RP01	C115	230	122	125	209	325	334
TK18	C116	222	130	129	217	325	334
TK19	C117	226	130	133	213	320	334
TK20	C118	234	122	125	209	325	342
TK21	C119	222	130	129	217	320	334
TK22	C120	236	122	129	213	330	342
TK23	C121	222	130	129	217	332	342
TK24	C122	222	130	129	217	320	338
TK25	C123	230	122	125	205	330	338
TK27	C124	236	122	129	213	330	338

Para o bloco HEMA foram utilizados 137 indivíduos e encontrados 82 haplótipos (Tabela 14). A proporção de indivíduos por haplótipos foi 1,691. Vários haplótipos foram compartilhados entre indivíduos da mesma população e também entre indivíduos de diferentes populações. Vinte haplótipos foram compartilhados entre dois ou mais indivíduos. O haplótipo H50 foi o mais compartilhado (10 indivíduos). Sete haplótipos foram compartilhados apenas entre indivíduos que pertencem à mesma população. Os outros foram compartilhados entre indivíduos de duas ou três populações diferentes.

Tabela 14: Haplótipos do bloco HEMA encontrados nas amostras populacionais de Africanos, Mimbó, Sítio Velho, Teresina, Ribeirão Preto e *Tikúna*.

Indivíduos	Haplótipos	PH1	PH2	PH3	PH4	PH5	PH6	PH7
AF002; AF017	H1	126	144	220	144	251	284	358
AF008	H2	130	144	216	156	251	284	354
AF010	H3	126	144	220	144	251	284	354
AF015	H4	118	148	216	152	246	294	350
AF018; AF045	H5	118	144	216	144	251	284	350
AF020	H6	126	144	220	144	251	289	350
AF023	H7	134	144	216	142	251	284	350
AF024; M06; T09; T48	H8	126	144	220	142	251	284	350
AF029	H9	126	144	220	142	251	294	358
AF033	H10	126	144	220	142	251	294	354
AF035	H11	114	144	216	140	266	294	358
AF039; SV03	H12	126	144	220	146	251	284	350
AF040	H13	126	144	216	152	251	284	354
AF042	H14	118	146	220	144	251	284	350
AF046	H15	114	144	218	142	251	284	350
AF048	H16	130	144	216	142	246	294	350
M01	H17	130	144	218	144	261	299	358
M02	H18	114	144	220	144	261	299	358
M03; M04; RP42	H19	130	144	220	144	261	299	358
M07	H20	130	144	220	148	251	284	346
M08	H21	130	144	220	154	251	284	346
M09; SV04	H22	134	144	220	148	251	284	346
M10	H23	126	144	220	144	346	284	350
M11	H24	114	144	216	140	266	289	354
M13; TK16; RP04	H25	118	148	222	142	261	299	358
SV01	H26	114	146	218	146	251	294	358
SV02; SV07	H27	126	144	220	144	251	284	350
SV05; SV10	H28	114	146	218	144	251	294	358
SV06; T20; T46; RP09; RP13; RP79; RP85; RP90	H29	118	148	222	142	261	294	354
SV08; SV09	H30	114	146	218	148	251	294	358
SV11	H31	130	144	216	156	251	284	358
SV12; T40; RP89	H32	118	148	222	142	251	284	346
SV13	H33	118	146	220	144	261	294	354
T02	H34	126	144	214	152	251	299	354

Indivíduos	Haplótipos	PH1	PH2	PH3	PH4	PH5	PH6	PH7
T03	H35	126	144	220	148	251	284	350
T04	H36	130	144	220	146	251	289	354
T06; TK12; TK15	H37	122	148	222	140	251	284	350
T07	H38	118	146	220	142	251	284	354
T08; T18; T45; RP22	H39	118	148	222	142	261	294	358
T10; T13	H40	118	148	222	148	261	294	358
T14	H41	130	148	222	142	261	294	354
T15; RP16; RP20; RP36	H42	118	148	222	144	261	294	358
T16	H43	130	144	216	142	251	284	358
T19; T21; TK03; TK066; TK07; TK14; TK22; TK24	H44	122	148	222	142	251	284	350
T22	H45	134	144	220	142	261	299	358
T23	H46	126	146	218	144	251	294	358
T24	H47	118	148	222	144	261	294	354
T25	H48	118	148	222	144	251	284	346
T38; T39; T42; TK04; TK10; TK11; TK19; TK20; TK21; TK27	H49	130	144	220	142	261	299	358
T41; TK13	H50	130	144	220	142	261	299	354
T43	H51	118	148	222	142	261	299	354
T49	H52	118	148	222	148	261	294	362
T50	H53	118	148	224	144	261	294	350
TK01; TK02; TK23	H54	130	144	220	142	261	299	362
TK18	H55	118	148	222	142	254	299	358
TK05; TK08; TK25	H56	118	148	222	142	251	284	350
TK09	H57	122	144	222	142	251	284	358
TK17	H58	130	144	220	142	251	284	350
TK26	H59	118	148	222	142	261	299	362
RP01	H60	118	148	222	144	261	289	358
RP03; RP68	H61	118	148	222	144	261	294	362
RP05	H62	118	148	216	150	246	294	350
RP06	H63	118	148	222	142	266	294	358
RP12	H64	118	146	220	144	251	284	354
RP15	H65	118	146	220	146	261	294	354
RP17; RP59	H66	114	144	216	142	251	284	350
RP19; RP27	H67	118	148	226	144	261	294	354
RP23	H68	126	146	216	144	251	294	354
RP37	H69	126	146	218	142	251	289	350
RP39	H70	130	144	220	146	251	284	350
RP41	H71	118	148	222	142	261	294	362
RP43	H72	122	144	214	152	251	294	354
RP66	H73	130	144	220	146	251	284	346
RP70	H74	130	144	212	144	251	289	354
RP78	H75	118	144	220	144	251	284	350
RP80	H76	118	148	222	142	254	294	354
RP81	H77	130	144	220	148	251	294	350

Desequilíbrio de ligação entre *lócus* (DL)

A análise do desequilíbrio de ligação foi realizada através da proporção de DL em cada população e em cada bloco separadamente (Tabela 15). As mulheres tiveram a fase inferida e foram utilizadas como dados haplóides de fase conhecida juntamente com os homens. Após aplicada a correção de Bonferroni, as proporções do DL em cada população variaram consideravelmente.

Tabela 15: Proporção (em %) do Desequilíbrio de Ligação em cada bloco para todas as populações, antes de depois da correção de Bonferroni

Blocos	Mimbó	Sítio Velho	Teresina	Ribeirão Preto	<i>Tikúna</i>
Proporção (%) de <i>lócus</i> em DL (Sem correção)					
DMD (8,23 cM)	52,38	57,14	28,57	14,29	14,29
PC (0,66 cM)	33,33	60,00	26,67	20,00	86,67
HEMA (0,23 cM)	38,10	42,86	76,19	95,24	100,00
Proporção (%) de <i>lócus</i> em DL (Com correção)					
DMD(8,23 cM)	38,10	9,52	9,52	4,76	9,52
PC (0,66 cM)	20,00	33,33	13,33	6,67	80,00
HEMA (0,23 cM)	33,33	23,81	47,62	76,19	95,24

Comparando os três blocos em relação à proporção do DL, pode-se observar que HEMA foi o bloco que apresentou maior valor, seguido pelo PC e DMD (de acordo com o esperado pela extensão dos blocos) nas populações TR, RP e TK. No entanto, o bloco DMD em MB e o bloco PC em SV apresentaram proporção de DL maior do que o esperado (Tabela 15).

Em relação às populações, *Tikúna* foi a que apresentou a maior proporção de DL nos blocos PC e HEMA. O mesmo não ocorreu no bloco DMD. Neste, Mimbó apresentou a maior proporção de DL (Tabela 15). Era esperado que TK apresentasse maior DL em todos os blocos.

A população Africana apresentou proporção de DL intermediária, diferente do esperado. A proporção de DL também não ocorreu como o esperado nos dois remanescentes de quilombos: ela foi diferente entre MB e SV. Este último apresentou valores menores do que os encontrados em MB. Se agruparmos os remanescentes de quilombos (MB e SV) e em outro grupo as populações urbanas (TR e RP), é possível observar que a média da proporção de DL dentro de cada um desses dois grupos é a mesma (79,05%). Essa média, se lavada em consideração, colocaria essas populações em terceiro lugar na proporção de DL.

DISCUSSÃO

Alelos Novos

Os loci do bloco PC ainda não tinham sido descritos. As sequências dos iniciadores, as condições da amplificação multiplex e identificação eletroforética de cada alelo foram descritos pelo grupo liderado por Enrique Medina-Acosta em cujo laboratório foi desenvolvida parte do presente trabalho.

Neste bloco PC, a média de 9,16 alelos por locus pode ser considerada alta tendo em vista a localização pericentromérica, região que possui uma baixa taxa de recombinação. Esse valor foi mais alto do que o encontrado na população da Alemanha (6,33), analisada por outros seis STRs da mesma região cromossômica (Edelman *et al.*, 2010). Grande parte dos alelos novos que foram encontrados está presente nas seis populações analisadas, inclusive na Africana; o que mostra que estes podem ser alelos mais antigos.

Os alelos novos dos blocos DMD e HEMA sugerem a existência de diferenças genéticas entre as populações analisadas neste trabalho e a população do estado do Rio de Janeiro, analisada nos trabalhos anteriores (Sarlo, 2011 e Machado, 2008, respectivamente). Os alelos novos não identificados na população do Rio de Janeiro têm baixa frequência e sua eventual detecção nas amostras descritas exigiria amostras de maior tamanho.

Frequências Alélicas

Todos os STRs dos três blocos utilizados neste trabalho são polimórficos. No bloco DMD, dois loci PR4 e PR14 foram monomórficos na população *Tikúna* (Tabela 3). Esta população é uma tribo indígena da Amazônia, e na época que foi amostrada, a mistura com outras populações brasileiras não era frequente (<5%) (Neel *et al.*, 1980). Ou seja, *Tikúna* pode ser considerada uma população sob isolamento genético e endogamia, o que diminui a diversidade e justifica estes dois loci (PR4 e PR14) serem monomórficos.

Em 90% dos loci o alelo mais frequente da população Africana está presente em todas as outras populações. De acordo com o que é conhecido sobre a colonização do mundo, a população africana, por ser a mais antiga, é a que apresenta maior diversidade. De acordo com o modelo de dispersão “Out of Africa”, a população africana foi colonizando novos territórios e, conseqüente, distribuindo seus alelos (Campbell *et al.*, 2008). Os alelos com maior frequência na população Africana teriam as maiores chances de passar pelo “efeito gargalo”. Sendo assim, os alelos mais antigos, que tendem a ser os mais frequentes, estariam presentes na maioria das populações, como ocorreu nas populações estudadas neste trabalho.

As diferenças observadas entre as amostras de homens e de mulheres não são reais, uma vez que o teste de diferenciação entre esses grupos não foi significativo ($P > 0,05$). A população do Rio de Janeiro e de Alagoas (Ferreira da Silva *et al.*, 2010) e também dos estados de São Paulo, Minas Gerais, Espírito Santo, e novamente Rio de Janeiro (Martins *et al.*, 2010a) apresentaram amostras de homens e de mulheres semelhantes, da mesma forma que neste trabalho.

Alelos Privados

Os alelos que ocorrem em uma única população são chamados de alelos privados. As populações isoladas possuem pouco fluxo gênico e conseqüentemente sofrem menor proporção de mistura. Sendo assim, estão sujeitas a possuírem alelos privados, pois alelos novos que surgem por mutações nestas populações ficam restritos às mesmas.

Neste trabalho foi encontrado um total de 27 alelos privados nos 20 loci analisados (Tabela 3). A maior parte dos alelos privados foi encontrada nas populações Africana, Teresina e *Tikúna*. Dentre as populações analisadas, Africana e *Tikúna* são as que possuem maior isolamento e também são as mais divergentes geneticamente. Isto justificaria o fato de possuírem maior número de alelos privados. Em outras populações de ameríndios da América do Sul também foram identificados alelos privados (Schroeder *et al.*, 2007; Schroeder *et al.*, 2009; Acuña-Alonzo *et al.*, 2010). Este achado foi justificado por estas populações serem isoladas, pequenas e endogâmicas, o que é semelhante na população Africana e *Tikúna*.

A ocorrência de alelos restritos às amostras de homens e às de mulheres deve ter surgido por causa da baixa frequência existentes na maioria destes alelos, uma vez que o teste de diferenciação genética entre homens e mulheres não foi significativo ($P > 0,05$).

Nos loci inéditos (bloco PC) foram encontrados 12 alelos privados, o que corresponde a aproximadamente 45% do total destes alelos. Dos alelos privados no bloco PC, cinco ocorreram na população Africana, que é uma população rural próxima ao rio Congo. Essa grande proporção de alelos privados encontrados na população Africana pode ter ocorrido por dois fatores: 1) por esta ser uma população isolada e endogâmica como explicado acima, e 2) devido à diferença na composição genética dessa população em relação às populações brasileiras.

Aderência ao equilíbrio de Hardy-Weinberg

Dos 120 testes realizados para verificar a aderência ao equilíbrio de Hardy-Weinberg, apenas três resultaram em desvios após a correção de Bonferroni (Tabela 4). Estes desvios foram encontrados nas populações Africana, Sítio Velho e Ribeirão Preto, sendo que todos podem ser explicados pelo déficit de heterozigotos. Esse tipo de desvio sugere que a

heterozigose esteja abaixo do esperado nas populações em questão, o que pode ter sido causado por endogamia ou acasalamento preferencial. Os desvios encontrados na população Africana e em Sítio Velho podem ser explicados por esta hipótese. Os desvios nestas duas populações são semelhantes ao ocorrido nos quilombos Barra (Luizon, 2007) e Gaucinha (Wiesel, 2003) que também podem ser explicados por casamentos não-aleatórios, pois estas são populações pequenas e isoladas. Outro fator que pode acarretar o déficit de heterozigotos e conseqüentemente o desequilíbrio de Hardy-Weinberg é o viés amostral causado pelo pequeno tamanho da amostra, sendo este o único fator que justifica o desequilíbrio encontrado em Ribeirão Preto.

Diferenciação genética das populações

A diferenciação genética foi estimada nas mulheres e homens separadamente. Na análise do grupo das mulheres (Tabela 5), as populações mais divergentes foram Africana e *Tikúna*. Isso era esperado, pois estas representam duas das três populações parentais no Brasil. Além disso, são as mais isoladas dentre as seis estudadas. *Tikúna* foi a população mais divergente deste trabalho, semelhante ao encontrado anteriormente para outras populações ameríndias (Massaro *et al.*, 2011; Wang *et al.*, 2010; Amorim *et al.*, 2011). Em todas essas populações ameríndias já analisadas, a maior divergência genética observada foi justificada pelo isolamento genético das populações, que é o mesmo fator que explica a divergência de *Tikúna*.

Ainda nas mulheres (Tabela 5), a semelhança entre SV e TR só pode ser explicada pelo efeito da pequena amostra de mulheres analisadas em TR e em Sítio Velho (apenas 12 e 13 mulheres, respectivamente), uma vez que há grande diferença entre as proporções da ancestralidade africana de ambas as populações. A contribuição africana em SV, baseada em STRs-X, foi estimada em 96,39% (Wiesel, 2003), enquanto que em TR, baseados em marcadores do tipo INDEL, esta estimativa foi de 21,5% (Lopes *et al.*, 2013). Em relação aos STRs-X, a contribuição africana na população TR é 41,20%, muito menor do que em Sítio Velho (Wiesel, 2003). A diferença na contribuição africana em SV e TR reforça a hipótese de que esse resultado foi ocasionado por um viés do pequeno número amostral.

Os homens mais semelhantes entre si foram os de Sítio Velho e Mimbó (Tabela 6). Através da observação da representação de cada ancestralidade (Figura 7) nos homens de Sítio Velho em Mimbó é possível observar que o componente africano (vermelho) é mais presente nessas duas populações. Em estimativas de ancestralidade obtidas pelo cromossomo Y (Wiesel *et al.*, 2013), estas duas populações apresentam proporções da ancestralidade africana semelhantes (47,15% e 39,84%, respectivamente) e maior do que as outras.

Na diferenciação genotípica (Tabela 7), as populações mais semelhantes (Sítio Velho e Teresina) e as mais divergentes (Africana e *Tikúna*) são as mesmas encontradas na

diferenciação gênica das mulheres. Por isso, este resultado reafirma o resultado encontrado anteriormente. Além disso, esses dados reforçam as justificativas apresentadas acima para a semelhança entre Sítio Velho e Teresina e também para a divergência entre Africana e *Tikúna*.

Diversidade Gênica

O menor H_s ($0,53836 \pm 0,24270$) encontrado nas seis populações foi o de *Tikúna* e o maior ($0,70146 \pm 0,11203$) o de Ribeirão Preto (Tabela 8). Pode-se considerar que *Tikúna* é a população que apresenta menor diversidade e Ribeirão Preto a maior. A primeira é uma população pequena e isolada geneticamente com pouco ou talvez nenhum fluxo gênico entre outros grupos, o que leva à deriva genética. A heterozigose média encontrada em *Tikúna* foi semelhante à encontrada na população ameríndia Kogi ($H_s=0,52$) quando foi analisada por 17 STRs (Amorim *et al.*, 2011). Kogi também é uma população ameríndia que sofre efeito da deriva genética. Outras populações de ameríndios da América do Sul, inclusive *Tikúna*, também apresentaram baixa heterozigose (Leite *et al.*, 2009 e Wang *et al.*, 2010), o que é concordante com o nossos resultados.

A população de Ribeirão Preto possui a maior diversidade (0,70) entre as seis populações estudadas. Esta é uma população urbana que recebe fluxo gênico contínuo (população aberta), o que aumenta a diversidade, por isso é plausível que tenha um H_s mais alto. Além disso, em Ribeirão Preto pôde ser observada a presença dos três componentes ancestrais (Figura 6 e 7) que formaram a população brasileira. Outras populações urbanas analisadas por outros STRs apresentaram valores de diversidade gênica semelhantes ao de Ribeirão Preto. As populações urbanas da Costa Rica e a Gaúcho (Amorim *et al.*, 2011) tiveram diversidade gênica (0,72 nas duas) muito semelhantes à de Ribeirão Preto (0,70), e a da população do Rio de Janeiro (Martins *et al.*, 2010a) foi apenas um pouco maior (0,76).

Os remanescentes de quilombos (Mimbó e Sítio Velho) possuem H_s semelhante ao das populações urbanas (Tabela 8). O valor da diversidade gênica de Sítio Velho ($0,695 \pm 0,142$) foi semelhante ao encontrado por Wiezel (2003) que foi de $0,615 (\pm 0,042)$. Estes valores são mais altos do que o esperado quando se trata de uma população com certo grau de isolamento, como é o caso dos quilombos. No entanto, essa heterozigose ainda reflete a mistura que ocorreu no momento da formação desses quilombos, uma vez que se passaram apenas 5-7 gerações, tempo evolutivamente curto para apagar os traços dessa mistura.

A população Africana também possui um H_s mais alto do que o esperado (Tabela 8). A população africana que analisamos neste trabalho é uma população ribeirinha, pequena e isolada, e por isso deveria apresentar diversidade gênica mais baixa. O pequeno tamanho amostral analisado pode ser uma explicação para este achado, que difere do esperado. No entanto existem relatos (Campbell *et al.*, 2008) de que as populações africanas, por serem

mais antigas e estarem sujeitas há mais tempo aos fatores que causam a diversidade (mutação e recombinação), possuem diversidade gênica mais alta, o que por sua vez, reforça a veracidade da *Hs* encontrado por este trabalho.

Análise da distância genética entre as populações

Concordante com o já descrito nas outras análises deste trabalho, os valores de F_{ST} (Tabela 9) mostraram que *Tikúna* foi novamente a população mais divergente. Outras populações nativo americanas analisadas também por STRs do cromossomo X, apresentaram alto valor de F_{ST} (0,17) quando comparadas umas com as outras (Wang et al., 2010). A população ameríndia Kashináwa também foi comparada com populações urbanas e com remanescentes de quilombo (Massaro, 2009), e apresentou valores de F_{ST} altos (0,104 entre KA e SP; 0,12 entre KA e Sítio Velho). Todos estes relatos puderam ser explicados pela ocorrência de deriva genética em populações pequenas e isoladas, como as ameríndias, que é a mesma explicação encontrada para o caso de *Tikúna*.

A população Africana deste trabalho também passa por isolamento genético, o que provoca diferenciação genética entre as populações. Este fator explica mais uma vez as diferenças genéticas entre as duas populações mais divergentes (Africana e *Tikúna*). Elas passaram por processos semelhantes, mas como esses processos são aleatórios, elas apresentam diferentes composições genéticas.

As populações mais semelhantes entre as mulheres foram Sítio Velho-Teresina, Ribeirão Preto-Sítio Velho e também Ribeirão Preto-Teresina (Tabela 9), pois a comparação entre essas populações apresentou valores de F_{ST} não significativos.

A semelhança entre Sítio Velho e Teresina foi observada também pela análise de diferenciação gênica e genotípica das mulheres. O mesmo motivo que justifica essa semelhança nos itens acima é o mesmo a ser empregado aqui. Estas populações apresentam diferentes contribuições ancestrais (Wiesel et al., 2013; Lopes et al., 2013) e apenas o pequeno tamanho amostral deve ser o responsável por esse resultado.

A semelhança identificada entre SV e RP pode, da mesma forma, ter sido causada pelo pequeno tamanho amostral em SV (25 indivíduos), uma vez que a ancestralidade presente nessas populações é diferente. A estimativa de ancestralidade de RP é: 0,49 de europeus, 0,3 de africanos e 0,21 de ameríndios (Massaro et al., 2011). Em SV esta estimativa é composta apenas pela ancestralidade europeia e africana (aproximadamente 0,4 e 0,6) (Wiesel et al., 2013). Além disso, estas são populações muito distantes geograficamente, o que impediria o fluxo gênico entre elas.

A semelhança entre RP e TR pode ser facilmente explicada pelo fato destas serem populações urbanas que, apesar da grande distância que as separa, passaram pelo mesmo

histórico de formação. O fato da formação destas populações serem parecidas é evidenciado por proporções de ancestralidade semelhantes: alta contribuição europeia. Essa ancestralidade é alta em TR tanto em análises de INDEL (0,6) (Lopes *et al.*, 2013) quanto em análises do cromossomo Y (0,795) (Wiezel *et al.*, 2013). No estado de SP (Manta *et al.*, 2013), a ancestralidade europeia também foi alta (0,63). Além disso, mais precisamente em RP, a ancestralidade europeia, obtida a partir do cromossomo X (Massaro *et al.*, 2011) foi menor, mas também pode ser considerada alta (0,49).

Na análise dos homens, apenas a comparação entre Mimbó e Sítio Velho apresentou valor de F_{ST} não significativo (Tabela 9), sendo estas as populações mais semelhantes. Sítio Velho e Mimbó são remanescentes de quilombo e possuem histórico de formação populacional semelhante. Além disso, foram consideradas populações que possuem similaridade genética em trabalhos anteriores (Wiezel, 2003).

Análise de Variância Molecular (AMOVA)

A diferença genética entre as populações não é tão grande a ponto de separá-las, por isso o resultado da AMOVA não foi significativo ($P < 0,01$) quando as populações foram colocadas em três grupos (quilombos, urbanas e ameríndia). Apesar destas populações não serem muito diferentes, existe variação entre elas. A amostra de mulheres apresenta uma divergência entre as populações (10,19%) semelhantes às dos homens (9,1%), porém levemente mais alta, o que ressalta a semelhança entre essas amostras. A maior parte da diversidade destas populações foi representada pela variação dentro das mesmas, ou seja, pela variação entre os indivíduos (89,81% nas mulheres e 90,98% nos homens). Populações afro-derivadas e urbana (Teresina) também apresentaram maior variação (92,18%) dentro dos grupos, ou seja, entre os indivíduos (Wiezel, 2003).

Inferência de Estrutura Populacional

A inferência da estrutura populacional foi realizada no grupo dos homens e mulheres separadamente pelo método de agrupamento, baseado no modelo desenvolvido por Pritchard *et al.* (2000). Tanto para as mulheres quanto para os homens, as populações foram agrupadas em três *clusters* (Tabela 10).

No caso das mulheres o *cluster* 1 é composto principalmente por amostras da população Africana. Aproximadamente 83% das amostras Africanas foram agrupadas no *cluster* 1. Este ainda é formado por parte dos indivíduos de Mimbó (67%) e Sítio Velho (42%). Portanto, pode-se inferir que este *cluster* representa o componente Africano na população brasileira. O segundo *cluster* (2) é o que tem sua composição melhor definida. Ele é formado por 91% da população *Tikúna*, sendo assim o *cluster* representante do componente ameríndio na população brasileira. O *cluster* 3 representa o componente europeu. Este *cluster* é formado pela maior parte das populações Teresina e Ribeirão Preto. Estas são populações

miscigenadas e que possui um alto componente europeu, como já foi explicado acima (Lopes *et al.*, 2013; Wiezel *et al.*, 2013; Manta *et al.*, 2013; Massaro *et al.*, 2011). Como o componente europeu não está representado por uma amostra populacional neste trabalho, ele pôde ser inferido como o terceiro componente das populações urbanas.

Para os homens, essa divisão em *cluster* foi muito semelhante ao que aconteceu com as mulheres. Foram encontrados os mesmos três *clusters*, com pequenas diferenças nas proporções dos indivíduos de cada população (Tabela 10). Essa semelhança era esperada, uma vez, que não existem diferenças significativas entre a amostra de homens e de mulheres ($P > 0,01$)

Nas figuras 6 e 7 tem-se a proporção da composição ancestral de cada população nos homens e mulheres, respectivamente. Nas duas figuras a população Africana foi representada em grande parte pelo componente africano (vermelho). Mimbó e Sítio Velho possuem em sua maior parte o componente africano, mas apresentam também uma proporção de mistura com ameríndio (verde) e europeu (azul), semelhante ao que ocorreu com os quilombos Barra e São Gonçalo analisados por Luizon (2007). A existência da ancestralidade ameríndia e europeia, mesmo que pouca, nos quilombos já era esperada, pois estas são populações com certo grau de miscigenação. Essa mistura é muito mais evidenciada na amostra das mulheres do que dos homens.

Tikúna é a população que apresentou menor proporção de mistura (Figura 6 e 7), ou seja, é a população mais pura, o que está de acordo com achados anteriores que demonstraram uma miscigenação menor do que 5% nesta população (Neel *et al.*, 1980). Em sua quase totalidade, é formada pelo componente ameríndio, que é representado pela cor verde.

As populações urbanas, Teresina e Ribeirão Preto, são as que possuem maior proporção de mistura entre os três grupos étnicos (africano, europeu e ameríndio) (Tabela 10 e Figuras 6 e 7), o que era esperado de acordo com o histórico da formação destas populações. Comparando estas duas, Teresina possui uma proporção maior do componente ameríndio do que Ribeirão Preto. Isto pode ser justificado por ela estar localizada geograficamente mais próxima às tribos indígenas do que a população de Ribeirão Preto. A predominância do componente europeu nestas duas populações é resultado da semelhança no histórico de formação delas e já foi demonstrada acima (Lopes *et al.*, 2013; Wiezel *et al.*, 2013; Manta *et al.*, 2013; Massaro *et al.*, 2011).

Diversidade Haplotípica

Os valores da diversidade haplotípica foram obtidos para cada bloco (DMD, PC e HEMA). O número de indivíduos utilizados para a análise em cada bloco foi muito semelhante.

O maior valor de diversidade haplotípica (h) era esperado para o bloco DMD, que é o mais extenso (8,24cM) (Tabela 11). Entretanto, este bloco apresentou valores de h intermediários aos dos outros dois blocos. O bloco DMD foi o que apresentou a menor proporção de indivíduos por haplótipos (1,081), o que significa que foi o bloco com maior quantidade relativa de haplótipos. Conseqüentemente, este foi o bloco que apresentou o menor compartilhamento de haplótipos (Tabela 12). Isto era esperado pelo fato deste ser o bloco mais extenso, o que aumenta a probabilidade de recombinação entre os loci do mesmo.

O bloco PC, apesar de localizado em região de baixa recombinação, foi o que teve os maiores valores de diversidade haplotípica (Tabela 11) e é o segundo bloco mais extenso (0,658cM). Um alto número de haplótipos em loci da região pericentromérica também foi encontrado entre indivíduos da Alemanha, Etiópia e Egito (Edelman *et al.*, 2010). Os autores analisaram 964 indivíduos ao todo e encontraram 695 haplótipos, uma proporção de indivíduos por haplótipos (1,387) apenas um pouco maior do que a encontrada neste trabalho (1,104). A diversidade haplotípica encontrada pode ser explicada pelas diferenças genéticas entre as populações analisadas neste trabalho, diferenças também evidenciadas pela utilização de outros parâmetros que compararam as populações, tais como diferenciação gênica, genotípica (Tabelas 5 a 7) e F_{ST} (Tabela 9).

No bloco PC existem haplótipos compartilhados entre alguns indivíduos de uma mesma população e também de populações diferentes (Tabela 13). Por exemplo, o haplótipo C19 foi compartilhado por um indivíduo de Mimbó e outro de Teresina. O compartilhamento de um único haplótipo entre duas populações não afirma que estas sejam populações geneticamente idênticas, mas permite inferir que esses indivíduos possuem semelhanças genéticas.

O bloco que apresentou a menor diversidade haplotípica foi o HEMA (Tabela 11). A distância entre primeiro e o último STR deste bloco é de apenas 0,23 cM, sendo assim, é o bloco com menor extensão utilizado neste trabalho. Em blocos de menor extensão, a probabilidade de ocorrer recombinação entre os loci diminui, o que justifica a menor diversidade haplotípica encontrada. Este também foi o bloco com maior proporção de indivíduos por haplótipos (1,691) e conseqüentemente maior número de haplótipos compartilhados (Tabela 14). Como neste bloco a distância entre os loci é muito pequena, a chance de ocorrer recombinação entre eles diminui, o que justifica também, o menor número de haplótipos. A diversidade haplotípica encontrada para o bloco HEMA na amostra da população do Rio de Janeiro (Machado *et al.*, 2009) foi maior (0,9786) do que a encontrada nas cinco populações analisadas neste trabalho.

Comparando os valores da diversidade haplotípica entre as populações, foi possível verificar que os menores valores foram encontrados na população *Tikúna* (Tabela 11). Este resultado, assim como os que foram discutidos nos itens anteriores, está dentro do esperado para

populações que possuem isolamento genético. Como já foi dito, este isolamento se confirma pelo registro feito por Neel *et al.* (1980), de que esta população ameríndia, quando amostrada, praticamente não apresentava mistura genética com outras populações não ameríndias.

A maior diversidade haplotípica foi obtida nas populações Mimbó (para o bloco PC) e Sítio Velho (para DMD e HEMA) (Tabela 11), refletindo assim a mistura (entre africanos, ameríndios e europeus) que ocorreu durante o processo de formação destes quilombos. Alta diversidade haplotípica também já foi encontrada anteriormente nestes dois quilombos (Wiesel, 2003).

Desequilíbrio de ligação entre *locus* (DL)

A proporção do desequilíbrio de ligação variou tanto entre os blocos quanto entre as populações. O bloco que apresentou maior quantidade de DL foi o HEMA. Este bloco está localizado na região Xq28 e possui apenas 0,229 cM de extensão (Tabela 15). Pelo fato dos marcadores serem tão próximos um dos outros, era esperado que este bloco apresentasse o maior DL. Em outras análises nas famílias de pacientes com hemofilia e na população do Estado do Rio de Janeiro (Machado *et al.*, 2009c) observou-se um alto DL, em concordância com os nossos resultados.

Foi observada uma proporção de DL maior do que o esperado nos blocos DMD e PC (em MB e SV, respectivamente). Este fenômeno pode ser mais facilmente explicado pela variação casual em amostras de pequeno tamanho (25 indivíduos).

Analisando o DL em cada uma das populações (Tabela 15), *Tikúna* foi a que apresentou maior proporção de DL. De acordo com os resultados citados e comentados acima, esta é uma população com baixa diversidade e alta diferenciação causadas pelo isolamento genético e endogamia o que favorece a deriva genética. Isto, por sua vez provoca o aumento do DL e, assim, explica a alta proporção de DL encontrada em TK. Em concordância com os nossos resultados, populações ameríndias da América Latina, analisadas por outros STRs do cromossomo X (Leite *et al.*, 2009; Wang *et al.*, 2010; Amorim *et al.*, 2011), apresentaram maior quantidade de DL quando comparadas com populações urbanas.

A população Africana apresentou uma proporção de DL intermediária, porém bem mais baixa do que a encontrada em *Tikúna* (Tabela 15). Tanto a população Africana quanto *Tikúna*, são pequenas e isoladas, por isso, era esperado que elas apresentassem valores de DL semelhantes. No entanto, é conhecido que populações mais antigas apresentam menor DL (Pfaff *et al.* 2001; Slatkin, 2008). A população Africana é mais antiga do que *Tikúna*, e por isso é possível que com o passar das gerações os eventos de recombinação quebrem os haplótipos gerados pela endogamia e diminua a proporção de DL.

SV e MB possuem o mesmo histórico de formação populacional, por isso deveriam apresentar proporções de DL semelhantes. No entanto, isso não aconteceu: SV apresentou proporção de DL menor do que MB (Tabela 15). O único fator que explica esse achado é, novamente, a existências dos efeitos causados pelo tamanho amostral menor (25) nos dois remanescentes de quilombo.

O modelo de mistura contínuo (CGF) onde a miscigenação ocorre em várias gerações, como é o caso das populações urbanas, apresenta uma queda de DL mais lenta do que no modelo em que a miscigenação acontece em apenas uma geração seguida de isolamento (HI) (Pfaff *et al.*, 2011; Amorim *et al.*, 2011). Por isso, era esperado que as populações urbanas (modelo CGF) apresentassem uma proporção de DL maior do que os quilombos (modelo HI). No entanto, a proporção de DL foi semelhantes entre esses dois grupos populacionais, e isto pode ser consequência, mais uma vez, do pequeno tamanho amostral utilizado em cada quilombo.

CONCLUSÕES

- Os seis novos loci da região pericentromérica foram amplificados com sucesso em PCR hexaplex. Estes loci são polimórficos (totalizam 169 alelos) e constituem uma nova opção para estimar a diversidade genética de uma população.
- As amostras de população Africana e de população ameríndia (*Tikúna*) constituíram o par de amostras com o maior índice de divergência entre si. *Tikúna*, de acordo com o esperado, foi a amostra que apresentou a menor diversidade e a maior proporção DL.
- Os três componentes ancestrais da população brasileira (europeu, africano e ameríndio) foram identificados, com diferentes proporções, em todas as populações aqui amostradas. As populações Africana e *Tikúna*, como era esperado, apresentaram o menor grau de miscigenação.
- A diversidade haplotípica foi diferente entre os três haploblocos analisados (DMD, PC e HEMA). O haplobloco HEMA, por ser o de menor extensão, foi o que apresentou a menor diversidade haplotípica, o menor número de haplótipos, assim como a maior proporção de DL.
- O conjunto dos 20 loci escolhidos se mostrou eficiente, tanto quanto para identificar a estrutura genética existente entre as populações analisadas, quanto para gerar informações sobre o DL e a diversidade haplotípica existente em cada região do cromossomo X analisado.

BIBLIOGRAFIA

1. Acuña-Alonzo V, Flores-Dorantes T, Kruit JK, Villarreal-Molina T, Arellano-Campos O, Hunemeier T, Moreno-Estrada A, Ortiz-López MG, Villamil-Ramírez H, León-Mimila P, Villalobos-Comparan M, Jacobo-Albavera L, Ramírez-Jiménez S, Sikora M, Zhang LH, Pape TD, Granados-Silvestre M de A, Montufar-Robles I, Tito-Alvarez AM, Zurita-Salinas C, Bustos-Arriaga J, Cedillo-Barrón L, Gómez-Trejo C, Barquera-Lozano R, Vieira-Filho JP, Granados J, Romero-Hidalgo S, Huertas-Vázquez A, González-Martín A, Gorostiza A, Bonatto SL, Rodríguez-Cruz M, Wang L, Tusie-Luna T, Aguilar-Salinas CA, Lisker R, Moises RS, Menjivar M, Salzano FM, Knowler WC, Bortolini MC, Hayden MR, Baier LJ, Canizales-Quinteros S. A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. *Hum Mol Genet* 19:2877–2885. 2010.
2. Alves da Silva, AF. Recombinação em 21q dificulta a determinação da origem meiótica da não disjunção em portadores da Síndrome de Down. Monografia de Conclusão de Curso Ciências Biológicas (ênfase em Biotecnologia). UENF, RJ. 2007
3. Amorim CEG, Wang S, Marrero AR, Salzano FM, Ruiz-Linares A, Bortolini MC. X-chromosomal Genetic Diversity and Linkage Disequilibrium Patterns in Amerindians and non-Amerindian Populations. *American Journal of Human Biology* 23: 299–304. 2011.
4. Angius A, Hyland FC, Persico I, Pirastu N, Woodage T, Pirastu M, De la Vega FM. Patterns of linkage disequilibrium between SNPs in a Sardinian population isolate and the selection of markers for association studies. *Hum Hered* 65:9-22. 2008.
5. Arpini-Sampaio Z, Costa MC, Melo AA, Carvalho MF, Deus MS, Simões AL. Genetic polymorphisms and ethnic admixture in African-derived black communities of northeastern Brazil. *Hum Biol* 71: 69-85, 1999.
6. Barbaro A. Distribution of 8 X-chromosomal STR loci in an Italian population sample (Calabria) *Forensic Science International: Genetics* 6: e174–e175. 2012.
7. Bortolini MC, Salzano FM, Bau CH, Layrisse Z, Petzl-Erler ML, Tsuneto LT, Hill K, Hurtado AM, Castro-De-Guerra D, Bedoya G, Ruiz-Linares A. Y-chromosome biallelic polymorphisms and Native American population structure. *Ann Hum Genet* 66:255-259. 2002.
8. Cainé LM, Zarrabeitia MT, Riancho JA, Pinheiro MF. Genetic data of a Brazilian population sample (Santa Catarina) using an X-STR decaplex *Journal of Forensic and Legal Medicine* 17: 272-274. 2010.
9. Campbell MC, Tishkoff SA. African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. *Annu. Rev. Genom. Human Genet.* 9:403-433. 2008
10. Chakraborty, R. Gene admixture in human populations: Models and predictions. *American Journal of Physical Anthropology*, 29: 1–43. 1986.
11. Courtin D, Jamonneau V, Mathieu J-F, Koffi M, Milet J, Yeminanga CS, Kumeso VKB, Cuny G, Bilengue CMM, Garcia A. Comparison of cytokine plasma levels in human African trypanosomiasis. *Trop Med Int Health* 11: 647–653. 2006.

12. Cunha, MC. Política indigenista no século XIX. In. CUNHA, Manuela Carneiro da (Org.) História dos Índios no Brasil. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura: FAPESP, 1992.
13. Curtin PD. The Atlantic Slave Trade: A Census. University of Winsconsin Press, Madison, 357 pp. 1969.
14. Edelman J, Hering S, Augustin C, Immel UD, Szibor R. Chromosome X centromere region—Haplotype frequencies for different Populations. Forensic Science International: Genetics Supplement Series 2(1): 398-399. 2009.
15. França, GVD. Medicina Legal. 6a edição. Rio de Janeiro, RJ: Guanabara Koogan, 2001. xiii, 579 pp.
16. Ferreira LB, Mendes-Junior CT, Wiezel CE, Luizon MR, Simões AL. Genomic ancestry of a sample Population from the state of São Paulo, Brazil. Am J Hum Biol.:18(5):702-5. 2006.
17. Ferreira da Silva IH, Barbosa AG, Azevedo DA, Sánchez-Diz P, Gusmão L, Tavares CC, Carvalho EF, Ferreira da Silva LA. An X-chromosome pentaplex in two linkage groups: haplotype data in Alagoas and Rio de Janeiro populations from Brazil. Forensic Sci Int Genet.4(4): 95-100. 2010.
18. Goldstein DB & Schlotterer C. Microsatellites: Evolution and Applications. Oxford University press. New York. 352 páginas. 1999.
19. Gusmão L, Sánchez-Diz P, Alves C, Gomes I, Zarrabeitia MT, Abovich M, Atmetlla I, Bobillo C, Bravo L, Builes J, Cainé L, Calvo R, Carvalho E, Carvalho M, Cicarelli R, Catelli L, Corach D, Espinoza M, García O, Malaghini M, Martins J, Pinheiro F, Porto MJ, Raimondi E, Riancho JÁ, Rodríguez A, Rodríguez A, Cardozo BR, Schneider V, Silva S, Tavares C, Toscanini U, Vullo C, Whittle M, Yurrebaso I, Carracedo A, Amorim A. A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. Int J Legal Med 123:227-234. 2009.
20. Hartl DL, Clark AG. Principles of population genetics. 4th ed. SinauerAssociates, Inc. Publishers. Sunderland, Massachusetts. 2007.
21. IBGE – CENSO 2000. Site do Instituto Brasileiro de Geografia e Estatística. Disponível na Internet via WWW. URL: www.ibge.gov.br. Acesso em setembro de 2009.
22. IBGE – CENSO 2010. Site do Instituto Brasileiro de Geografia e Estatística. Disponível na Internet via WWW. URL: www.ibge.gov.br. Acesso em Junho de 2013.
23. Jorgensen TH, Degn B, Wang AG, Vang M, Gurling H, Kalsi G, McQuillin A, Kruse TA, Mors O, Ewald H. Linkage disequilibrium and demographic history of the isolated population of the Faroe Islands. Eur J Hum Genet 10:381-387. 2002.
24. Kasai, K.; Nakamura, Y.; White, R. Amplification of a Variable Number of Tandem Repeats (VNTR) Locus (pMCT118) by the Polymerase Chain Reaction (PCR) and its application to Forensic. Science. J Forensic Sci, 35: 1196-1200. 1990.
25. Kashi, Y, King, DG. Simple sequence repeats as advantageous mutators in evolution. Trends in Genetics 22 (5): 253-259. 2006.
26. Kobachuk LAG, Malaghini M, Silva JS, Gusmão L, Bicalho MG. Genetic profile characterization of ten X-STRs in a sample from Paraná, Brazil Int J Legal Med: DOI 10.1007/s00414-012-0739-x. 2012.

27. Keyeux KY & Stephan W. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* 160, 765–777. 2002.
28. Lahn BT & Page DC. Four evolutionary strata on the human Y chromosome. *Science* 286, 964–967. 1999.
29. Lalloz MRA, Mcvey JH, Pattinson JK, Tuddenham EGD. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *Lancet*. 338: 207-11. 1991.
30. Lalloz MRA, Schwaab R, Mcvey JH, Michaelides K, Tuddenham EGD. Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats within the factor VIII gene. *Br J Haematol*. 86:804-9. 1994.
31. Leite FP, Santos SE, Rodriguez EM, Callegari-Jacques SM, Demarchi DA, Tsuneto LT, Petzl-Erler ML, Salzano FM, Hutz MH. Linkage disequilibrium patterns and genetic structure of Amerindian and non-Amerindian Brazilian populations revealed by long-range X-STR markers. *Am J Phys Anthropol* 139:404-412. 2009.
32. Lewontin, RC & Kojima, K. The evolutionary dynamics of complex polymorphisms. *Evolution* 14,458–472. 1960.
33. Lewis, PO.; Zaykin, D. Genetic data analysis: software for the analysis of discrete genetic data. Version 1.0. Sunderland: Sinauer Associates Inc. 1997.
34. Li, WH, Yi, S. & Makova, K. Male-driven evolution. *Curr. Opin. Genet. Dev.* **12**, 650–656 (2002).
35. Ling Liu Q, Zhao H, Chen JD, Wang G, De-Jian L, Quan, L. Development and population study of the 12 X-STR loci multiplexes PCR systems. *Int J Legal Med* 126:665–670. 2012.
36. Long JC. The genetic structure of admixed populations. *Genetics* 127:417–428. 1991.
37. Lopes TR, Santos S, Pinto, GR. Population data of the 46 insertion–deletion (INDEL) loci in population in Piauí State, Northeastern Brazil. *Forensic Sci. Int. Genet.* 8(9):e75145. 2013. doi: 10.1371/journal.pone.0075145.
38. Luizon MR, Mendes-Junior CT, De Oliveira SF, Simoes AL. Ancestry informative markers in Amerindians from Brazilian Amazon. *Am J Hum Biol* 20:86-90. 2008.
39. Machado FB. Análise compreensiva de marcadores STR no cromossomo X humano: aplicações na genética médica e forense. Dissertação de mestrado apresentada no Curso de Biotecnologia, UENF, RJ. 2008.
40. Machado FB, Medina-Acosta E. High-resolution combined linkage physical map of short tandem repeat loci on human chromosome band Xq28 for indirect haemophilia A carrier detection. *Haemophilia*, 15, 297–308. 2009a
41. Machado FB, Duarte LP, Medina-Acosta E. A novel informative dinucleotide microsatellite marker located on human factor VIII intron 25. *Haemophilia*, 15, 613–614. 2009b.
42. Machado FB, Duarte LP, Medina-Acosta E. Improved criterion-referenced assessment in indirect tracking of haemophilia A using a 0.23 cM-resolution dense polymorphic marker set. *Haemophilia*, 15, 1135–1142. 2009c.
43. Mangs HA, Morris BJ. The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Curr. Genomics* 8(2): 129–36. 2007.


44. Manta FSN, Pereira R, Vianna R, Araújo AFB, Gitaí DLG, Silva DA, Wolfgramm EV, Pontes IM, Aguiar JI, Moraes MO, Carvalho EF, Gusmão L. Revisiting the genetic ancestry of brazilians using autosomal AIM-indels. *Plos One* 8(9): e75145. 2013.
45. Martins JA, Costa JC, Paneto GG, Figueiredo RF, Gusmão L, Sánchez-Diz P, Carracedo A, Cicarelli RMB. Genetic profile characterization of 10 X-STRs in four populations of the southeastern region of Brazil *Int J Legal Med* 124:427–432. 2010a.
46. Martins JA, Silva RHA, Freschi A, Paneto GG, Oliveira RN, R. M. B. Cicarelli RMB. X-chromosome genetic variation in São Paulo State (Brazil) population. *Annals of Human Biology*, July–August 37(4): 598–603. 2010b.
47. Mattoso, KMQ. *Ser escravo no Brasil*. São Paulo: Brasiliense, 1982.
48. Marano, LA. *Análise de polimorfismos do gene MC1R associados á fenótipos humanos de pigmentação na população brasileira*. Dissertação de mestrado apresentada no Curso de Genética, FMRP, USP, SP. 2011.
49. Massaro, JD. *DINÂMICA de polimorfismos genéticos ligados ao gene da hemofilia a ($f8$) na população brasileira*. Tese de Doutorado apresentada no Curso de Genética, FMRP, USP, SP. 2009.
50. Massaro JD, Wiezel CE, Muniz YC, Rego EM, de Oliveira LC, Mendes-Junior CT, Simões AL. Analysis of five polymorphic DNA markers for indirect genetic diagnosis of haemophilia a in the brazilian population. *Haemophilia* 17(5):936-43. 2011.
51. Matoso KMQ. *Ser escravo no Brasil*. São Paulo: Editora Brasiliense. 1982.
52. Mendes Jr CT. *DNA mitocondrial na Amazônia brasileira: Estrutura genética regional e inferências continentais*. Tese de doutorado apresentada no curso de Genética, FMRP, USP, SP. 2005.
53. Mestriner MA, Simoes AL, Salzano FM. New studies on the esterase D polymorphism in South American Indians. *Am J Phys Anthropol* 52:95–101. 1980.
54. Miller, S. A; Dykes, D. D; Polesky, H. F. A simples salting out procedure for extracting DNA from human nucleates cells. *Nucleic Acids Res* 16 (3): p.1215, 1988.
55. Mohrenweiser H, Neel JV, Mestriner MA, Salzano FM, Migliazza E, Simões AL, Yoshihara CM, Electroforesis variants in three amerindians tribes: the Baniwa, Kanamari and Central Pano of Western Brazil. *Am. J Phys. Anthropol* 50: 237-246. 1979.
56. Muniz YC, Ferreira LB, Mendes-Junior CT, Wiezel CE, Simões AL. Genomic ancestry in urban Afro-Brazilians. *Ann Hum Biol.*:35(1):104-11. 2008.
57. Nakamura Y, Leppert M, O'connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R. Variable Number of Tandem Repeat (VNTR) markers for human gene mapping. *Science*, 235:1616-1622. 1987.
58. Neel JV, Gershowitz H, Mohrenweiser HW, Amos B, Kostyu DD, Salzano F M, Mestriner MA, Lawrence D, Simoes AL, Smouse PE, Oliver WJ, Spielman RS, Neel Jr JV. Genetic studies on the Ticuna, an enigmatic tribe of Central Amazonas. *Ann Hum Genet.* 44 (1): 37-54, 1980.
59. Nei M. *Molecular evolutionary genetics*. New York: Columbia University Press. 1987.

60. Nishi T. Application of a novel multiplex polymerase chain reaction system for 12X-chromosomal short tandem repeats to a Japanese population study *Legal Medicine* 15: 43–46. 2013.
61. Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, Chen F, Platt D, Paabo S, Pritchard JK, Rubin EM. Sequencing and analysis of Neanderthal genomic DNA. *Science* 314:1113–1118. 2006.
62. Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, et al. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA* 100: 177–182. 2003.
63. Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLOS ONE* 6: e17063. doi:10.1371/journal.pone.0017063. PubMed:21359226. 2011.
64. Penna LS, Silva FG, Salim PH, Ewald G, Jobim M, Magalhães JA, Jobim LF. Development of two multiplex PCR systems for the analysis of 14 X-chromosomal STR loci in a southern Brazilian population sample. *Int J Legal Med.* 126(2):327–30. 2012.
65. Pereira R, Gomes I, Amorim A, Gusmão L. Genetic diversity of 10 X chromosome STRs in northern Portugal. *International Journal of Legal Medicine*, v. 121, p. 192–197. 2007.
66. Pfaff CL, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet* 68:198–207. 2001.
67. Pistis G, Piras I, Pirastu N, Persico I, Sassu A, Picciau A, Prodi D, Fraumene C, Mocci E, Manias MT, Atzeni R, Cosso M, Pirastu M, Angius A. High differentiation among eight villages in a secluded area of Sardinia revealed by genome-wide high density SNPs analysis. *PLoS One* 4:e4654. 2009.
68. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–59. 2000.
69. Raymond M, Rousset F. GENEPOP (version 2.0): population genetics software for exact tests and ecumenism. *J Hered*, 86:248–249, 1995a.
70. Raymond M, Rousset F. An exact test for population differentiation. *Evolution*, 49: 1280–1283. 1995b.
71. Ribeiro-Rodrigues EM. A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region *Forensic Science International: Genetics* 2: 154–158. 2008.
72. Ribeiro-Rodrigues EM, Palha TJB, Bittencourt EA, Ribeiro-dos-Santos AR, Santos S. Extensive survey of 12 X-STRs reveals genetic heterogeneity among Brazilian populations. *Int J Legal Med* 125:445–452. 2011.
73. Salzano FM, Gershowitz H, Mohrenweiser H, Nee JV, Smouse PE, Mestriner MA, Weimer TA, Franco MHL, Simões AL, Constans J, Oliveira AE, De Melo e Freitas MJ, Gene flow across tribal barriers and its effect among the Amazonian Içana river indians. *Am. J. Phys. Anthropol.* 69:3–14. 1986.
74. Salzano FM & Bortolini MC. *The Evolution and Genetics of Latin American Populations.* Cambridge, Cambridge University Press. 2002.

75. Santos-Lopes SS, Pereira RW, Wilson IJ, Pena SDJ. A Worldwide Phylogeography for the Human X Chromosome. *PLoS ONE* 2(6): e557. 2007.
76. Sarlo, GL. Desenvolvimento de marcadores de microssatélites para o rastreamento indireto de portadores de mutações no gene DMD causadoras das distrofias musculares de Duchenne e Becker. 2011. Monografia de Conclusão de Curso Ciências Biológicas (ênfase em Biotecnologia), UENF, RJ.
77. Schaffner, SF. The X chromosome in population genetics. *Nature Reviews Genetics*, v. 5, p. 43-51, 2004.
78. Schneider S, Roessli D, Excoffier L. Arlequin ver. 2000: A software for population genetics data analysis. *Genetics and Biometry Laboratory*, University of Geneva, Switzerland, 2000.
79. Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. A private allele ubiquitous in the Americas. *Biol Lett* 3:218-223. 2007.
80. Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, et al. Haplotypic background of a private allele at high frequency in the Americas. *Mol Biol Evol* 26: 995-1016. 2009.
81. Schum JW. Genetic identity: new approaches to DNA fingerprint analysis. *Promega Notes* 58:12-20. 1996.
82. Service S, DeYoung J, Karayiorgou M, Roos JL, Pretorius H, Bedoya G, Ospina J, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet* 38: 556-560. 2006.
83. Simões AL. O polimorfismo da esterase D em indígenas sul-americanos. Tese de Doutorado apresentada no curso de Genética, FMRP, USP, SP. 1980.
84. Szibor R, Krawczak M, Hering S, Edelmann J, Kuhlisch E, Krause D. Use of X-linked markers for forensic purposes. *Int J Legal Med* 117:67-74. 2003.
85. Slatkin M. Linkage disequilibrium understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 9:477-485. 2008.
86. Stenzel A, Lu T, Koch WA, Hampe J, Guenther SM, De La Vega FM, Krawczak M, and Schreiber S. Patterns of linkage disequilibrium in the MHC region on human chromosome 6p. *Hum. Genet.* 114: 377-385. 2004.
87. Stephens M, Smith N, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978-989. 2001.
88. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* 76:449-462. 2005.
89. Wall, J. D. & Pritchard, J. K. Haplotype blocks and linkage disequilibrium in the human genome. *Nature Rev. Genet.* 4, 587-597. 2003.
90. Wang S, Bedoya G, Labuda D, Ruiz-Linares A. Brief communication: patterns of linkage disequilibrium and haplotype diversity at Xq13 in six Native American populations. *Am J Phys Anthropol* 142:476-480. 2010.
91. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6):1358-1370, 1984.

92. Wiezel, CEV. Diversidade genética e composição étnica em remanescentes de quilombos do estado do Piauí. 2003. Tese de Doutorado. Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, USP, SP.
93. Wiezel CEV, Luizon MR, Sousa SMB, Santos LMW, Muniz YCN, Mendes-Junior CT, Simões AL. Y-Linked microsatellites in Amazonian Amerindians applied to ancestry estimates in Brazilian Afro-derived populations. *American Journal of Human Biology* 25:313-317, 2013.
94. www.bgbx.com.br Banco Genético Brasileiro do Cromossomo X. Disponível na Internet. Acesso em Julho de 2013.
95. www.palmares.gov.br Fundação Cultural Palmares. Disponível na Internet. Acesso em Junho de 2013.
96. Zarrabeitia MT, Pinheiro F, Pancorbo MM, Cainé L, Cardoso S, Gusmão L, Riancho JA. Analysis of 10 X-linked tetranucleotide markers in mixed and isolated populations. *Forensic Science International: Genetics*, v. 3, p. 63-66. 2009.

COMPROVANTE DE SUBMISSÃO

 American Journal of Human Biology Edit Account | Instructions & Forms | Log Out | [Get Help Now](#)
SCHOLARONE™
Manuscripts

[Main Menu](#) → [Authoring Dashboard](#) → Submission Confirmation You are logged in as Aguinaldo Simões

Submission Confirmation



Thank you for submitting your manuscript to *American Journal of Human Biology*.

Manuscript ID: AJHB-13-0248

Title: GENETIC STRUCTURE AND LINKAGE DISEQUILIBRIUM IN BRAZILIAN AMERINDIAN AND QUILOMBO REMNANT COMMUNITIES ESTIMATED BY NEW X-CHROMOSOME STRS

Authors: Andrade, Claudia
Wiezel, Claudia
Sarfo, Laís
Machado, Filipe
Silva, Antônio
Mendes-Junior, Celso
Medina-Acosta, Enrique
Simões, Aguinaldo

Date Submitted: 23-Oct-2013

 Print  Return to Dashboard

MANUSCRITO

TITLE PAGE

GENETIC STRUCTURE AND LINKAGE DISEQUILIBRIUM IN BRAZILIAN AMERINDIAN AND QUILOMBO REMNANT COMMUNITIES ESTIMATED BY NEW X-CHROMOSOME STRS

Claudia C. F. Andrade¹, Cláudia E. V. Wiesel¹, Laís G. Sarlo², Filipe B. Machado², Antônio F. Alves da Silva², Celso T. Mendes-Junior³, Enrique M. A-costa^{2#}, Aguinaldo L. Simões^{1#}

¹ Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil.

² Núcleo de Diagnóstico e Investigação Molecular, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Avenida Alberto Lamego 2000, Parque Califórnia, Campos dos Goytacazes, RJ, Brazil.

³ Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil.

Abbreviated title: X-STRs in Amerindian, Afro-derived Brazilian and Africa populations

Correspondence to:

Aguinaldo Luiz Simões, Av. Bandeirantes 3.900, Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, São Paulo, ZIP 14049-900, Brazil. Telephone +55 (16) 3602-3050, Fax +55 (16) 3602-3157. E-mail: alsimoes@fmrp.usp.br

Enrique Medina-Acosta, Laboratório de Biotecnologia, Centro de Biociências e Biotecnologia, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Avenida Alberto Lamego 2000, Parque Califórnia, Campos dos Goytacazes, RJ, ZIP 28013-602, Brazil.

Tel/fax: +55 (22) 2739-7086; E-mail: quique@uenf.br

Grant sponsors: Fundação de Apoio ao Ensino Pesquisa e Assistência do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (FAEPA), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-Brazil).

ABSTRACT

OBJECTIVES: The aim of this study was to obtain information on the genetic structure and linkage disequilibrium of two quilombo remnant communities, two Brazilian urban populations, and one indigenous population. Furthermore, describing thirteen new STRs loci (Xp21 and Xp11.21-Xq11.1). Herein we have genotyped twenty short tandem repeats (STRs) located in three regions of the X-chromosome (Xp21, Xp11.21-Xq11.1, and Xq28).

METHODS: The frequencies for 20 X-chromosome STRs were characterized in four Brazilian populations analyzed. We then estimated the genetic structure among these four population through the performed analysis of molecular variance (AMOVA), F_{ST} , clustering, haplotype diversity, and linkage disequilibrium.

RESULTS: The most divergent population was the Amerindian population Tikúna. In addition, this population exhibited the lowest diversity (0.5394 ± 0.2751) as well as the highest proportion of linkage disequilibrium (LD). The populations were grouped in three clusters that represent the each of ancestral components of the Brazilian population. The block HEMA has the lowest values of haplotype diversity and proportion of LD on five populations analyzed.

CONCLUSIONS: The new loci were successfully described for population genetics studies. The set of 20 markers used in the present study were able

to provide information and genetically distinguish the five populations under study. The genetic parameters analyzed showed that the population size, isolation, and genetic drift have a direct influence on diversity, differentiation and the proportion of LD in a population. In addition, we found that lower extension blocks have greater proportion of LD and lower haplotype diversity.

KEYWORDS: Amerindians, Quilombo remnant communities, X-STR, Genetic structure, Linkage disequilibrium.

INTRODUCTION

The formation of the Brazilian population is the result of five centuries of admixture between populations originating from three continents: European colonizers, African slaves, and Amerindians (Salzano and Bortolini, 2002). The interregional migrations that occurred during the colonial period in Brazil promoted an increasing gradient of the European contribution in a north–south direction, thus reflecting the historical differences of the colonization of each region (Wiesel et al., 2013; Luizon, 2007). As a result, Brazil has one of the most diverse populations (Pena et al., 2011); this diversity can be estimated using various classes of genetic markers.

X-chromosome genetic markers generate information regarding the female and male lineages of a population (Santos-Lopes et al., 2007), and because they exhibit a lower recombination rate and a faster genetic drift (Leite et al., 2009), they reflect the evolutionary history of the population over more generations. In addition, determination of haplotypes is facilitated by the occurrence of a single X-chromosome in males (Leite et al., 2009; Amorim et al., 2011).

X-chromosome short tandem repeats (STRs) have been studied in different Latin American populations (Leite et al., 2009; Massaro et al., 2011; Wang et al., 2010; Amorim et al., 2011; Penna et al., 2012; Lopes et al., 2013). These

studies have demonstrated that the Amerindian populations exhibit lower diversity and higher linkage disequilibrium (LD), the urban populations exhibit higher diversity and lower LD, and the quilombo remnant communities exhibit intermediate values with respect to both parameters.

In the present study, allele frequencies of 20 X-chromosome STR loci distributed in 3 haploblocks are described (one in the pericentromeric region and one in each arm of the chromosome) to infer the genetic structure as well as the LD pattern in five Brazilian populations (two quilombo remnant communities, two urban populations, and one Amerindian population) with distinct histories of formation. Seven of these loci constitute a haploblock whose genotyping conditions have been described previously (Machado et al., 2009a). The remaining 13 loci are described here for the first time.

MATERIALS AND METHODS

Populations

We analyzed 201 samples from unrelated individuals from two quilombo remnant communities, Mimbó (MB; $n = 25$) and Sítio Velho (SV; $n = 25$); two urban populations, Teresina (TR; $n = 50$) and Ribeirão Preto (RP; $n = 51$); and one Amerindian population, *Tikúna* (TK; $n = 50$).

MB and SV are quilombo remnant communities in the state of Piauí (Arpini-Sampaio et al., 1999) located at $6^{\circ}14''\text{S } 42^{\circ}50''\text{W}$ and $5^{\circ}20''\text{S } 41^{\circ}15''\text{W}$, respectively. TR ($5^{\circ}1''\text{S } 42^{\circ}49''\text{W}$) and RP ($21^{\circ}33''\text{S } 47^{\circ}43''\text{W}$) are cities with urban admixed population. TK is an Amerindian population located in the Amazon region at $4^{\circ}15'\text{S } 69^{\circ}35'\text{W}$, which was sampled during the Alpha Helix expedition in 1976 (Neel et al., 1980).

Laboratory methods

Blood samples were collected and processed according to the procedure described previously (Mendes-Junior et al., 2007). Genomic DNA was extracted using the QIAamp DNA Investigator Kit (Qiagen, Gaithersburg, MD, USA). Twenty microsatellites grouped into three blocks distributed along

the X-chromosome were analyzed. Seven of these (DXS31302675, DXS31931836, DXS32159553, DXS32154670, DXS32251717, DXS32426983, and DXS33091950) form the DMD block in the Xp21 region, internal to the Duchenne muscular dystrophy (DMD) gene. The PC block, located in the pericentromeric region (Xp11.21-Xq11.1) is composed of six loci (DXS56749371, DXS56909198, DXS57434969, DXS61920979, DXS62085977, and DXS62290471). The remaining seven loci (REN90833, F8Int25.2, F8Int22, F8Int13.2, HEMA154311.3, TMLHInt5, and HEMA154507.3) are located in the F8 gene (Xq28) and comprise the HEMA block (Machado et al., 2009a; Machado et al., 2009b, Machado et al., 2009c).

The 13 short tandem repeats (STR) tested-and-tried for the first time (blocks DMD and PC) in this study were computationally identified using the same microsatellite mining and validation strategy previously outline by us to determine the high-resolution combined linkage physical map of STR loci on Xq28 for indirect haemophilia A carrier detection with minor modifications (Machado et al., 2009b). Briefly, we restricted our analysis to tri-, tetra-, penta- and hexanucleotide STR loci with score >80 and repeat identity >90% using Tandem Repeat Finder (Benson, 1999) for human chromosomal bands Xp21.2-Xp21.1 (spanning the *DMD* gene) and Xp11.21-Xq11.1 (pericentromeric region encompassing 19 annotated at the NCBI Map Viewer (Wolfsberg, 2007) and the UCSC Genome Browser (<http://genome.ucsc.edu>) (Kent et al., 2002). The STR loci were firstly predicted *in silico* as informative using Multiple Sequence Alignment by CLUSTALW to identify allelic variants on three human X chromosome reference sequences (NC_000023, AC_000066, AC_000155). Genetic distances for unmapped markers were interpolated on the Rutgers map of the human genome (Matise et al., 2007) (Table 1).

Locus-specific primer pairs were designed using the online software *OligoPerfect™ Designer* from *Invitrogen™* (<http://tools.lifetechnologies.com/content.cfm?pageid=9716>), and conferred free of structure constraints (potential hairpin formation, 3' complementary, potential self-annealing sites) by *OligoCalc Oligonucleotide Properties*

Calculator (Kibbe, 2007) and *Autodimer* (Vallone, 2004) (Table 2). The uniqueness of each amplicon was validated by Basic Local Alignment Search Tool (BLAST) (Cummings et al., 2002).

The concentration of the primers in multiplex PCR was adjusted for each locus: 8, 8, 4, 4, 3, 15, and 10 μmol , were used for the amplification of the loci DXS31302675, DXS31931836, DXS32159553, DXS32154670, DXS32251717, DXS32426983, and DXS33091950, respectively, of the DMD block; 10, 5, 7, 3, 20 and 10 μmol were used for the loci DXS56749371, DXS56909198, DXS57434969, DXS61920979, DXS62085977, and DXS62290471, respectively, of the PC block; and 8, 10, 10, 10, 7, 6, and 4 μmol were used for the loci REN90833, F8Int25.2, F8Int22, F8Int13.2, HEMA154311.3, TMLHInt5, and HEMA154507.3, respectively, of the HEMA block. The remaining conditions were the same as those described by Machado et al. (2009), with the exception of the annealing temperature: 63°C for the HEMA block and 60°C for the other blocks.

Electrophoresis of the amplified fragments was performed using the automatic analyzer ABI PRISM 310 (Applied Biosystems, Foster City, CA). The size of the fragments was obtained using the software GeneScan v3.2.1 and Genotyper v2.5.2 that identifies the alleles by comparison with DNA fragments of specific sizes obtained using The GeneScan™ 600 LIZ Size Standard v2.0 (Applied Biosystems, Foster City, CA).

Statistical analysis

The allele frequencies and adherence to the Hardy–Weinberg equilibrium were estimated using the method of gene counting with the GENEPOP 4.1 software (Rousset, 2008).

The mean heterozygosity, F_{ST} , molecular variance analysis (AMOVA), and (LD) were estimated using the ARLEQUIN software, version 3.1 (Schneider et al., 2000). The estimation of haplotype diversity was performed only in men and separately for each group of loci (DMD, PC, and HEMA). In the analysis of

the DL, the phase of the women was inferred and then calculations were performed in men and women together.

The inference of the population structure was performed using the STRUCTURE program, version 2.3.3, according to the grouping method, based on Pritchard et al. (2000).

Bonferroni correction was applied to analyses that were based on the significance value (P). Significance was thus recalculated according to the number of independent tests (Hartl and Clark, 2007) in each analysis.

RESULTS

Basic population parameters (new alleles, alleles per locus, private alleles, and adherence to the Hardy–Weinberg equilibrium)

We describe for first time the primers and PCR conditions for 13 X-chromosome STR loci (7 in the DMD haploblock and 6 in the PC centromeric haploblock). Distribution of the allele frequency of the 20 loci (161 alleles) is shown in Supplementary Table 1. Considering the previously described loci of the HEMA block, 20 new alleles were identified (underlined in underlined in Supplementary Table 1). The loci REN90833, HEMA154311.3, and TMLHInt5 exhibited a high number of new alleles (7, 5, and 5, respectively). The loci F8Int25.2, F8Int22, and F8Int13.2 exhibited only 1 new allele each. The differences in frequency between men and women were not significant.

Only two loci (DXS31302675 in SV and F8Int13.2 in RP) were in Hardy–Weinberg disequilibrium after Bonferroni correction ($P < 0.0025$).

The number of alleles per locus varied between three and 21 (Supplementary Table 1), and the number of alleles per locus per population varied between one (in TK, loci DXS31931836 and DXS33091950) and 19 (in RP, locus DXS31212596). The average number of alleles per population (Table 3) was higher (6.4) in TR and lower (4.8) in SV and TK.

Eight alleles, known as private alleles, occurred with a frequency higher than 5% in a single population. Five of them (DXS62290471*318, F8Int13.2*154, F8Int25.2*152, TMLHInt5*274, and HEMA154311.3*266) were found in the quilombo remnant communities, two (DXS32426983*362, DXS32426983*370) in the urban population and one (DXS62085977*332) in the Amerindian sample. The rare alleles (frequency <5%) may not have been identified in other populations because of the small sample size.

Population structure (F_{ST} , AMOVA, Clustering)

TK was the only population that was distinct from all others, both in men and women. By comparing the male population alone, MB and SV were the only communities in which F_{ST} showed similarity between them ($P > 0.14$). Amongst females, the number of genetically similar samples was higher (SV \times TR, SV \times RP, and RP \times TR with $P = 0.14$; 0.018, and 0.09 respectively) than that in men.

AMOVA was initially performed by creating three distinct groups with the samples of the quilombo remnant communities, urban populations, and Amerindian population. The variance between these groups was not significant ($P > 0.05$). When the five populations were clustered in one single group, the analysis showed that the variation between the individuals of each population was higher (89.81% and 90.98%) than the variation between the populations (10.19% and 9.1%) both in women and men, respectively.

The Bayesian structure analysis regrouped the individuals of each population in three clusters both in men and women (Table 4). The individuals of the TK population (men and women) were almost exclusively grouped into a single cluster, which revealed the genetic homogeneity of this population. Men and women of the quilombo remnant communities were grouped in another cluster. Both men and women of the RP urban population were grouped in a third cluster. It was expected that the TR sample would be grouped in this same cluster because it is also an urban population. However, only women were grouped (45%) within this cluster.

Men were grouped (54%) in the same cluster as that of the quilombo remnant communities.

Linkage Disequilibrium and Haplotype diversity

With respect to the degree of LD, the HEMA block exhibited the highest LD, followed by the PC and DMD blocks (in agreement with what was expected taking into consideration the extension of the blocks) in the TR, RP, and TK populations. However, the DMD block in MB and the PC block in SV exhibited higher LD than expected (Table 3).

With regard to the degree of LD in the populations, TK exhibited the highest LD in the PC and HEMA blocks. In contrast, in the DMD block, MB exhibited the highest LD (Table 3). It was expected that TK would exhibit higher LD in all blocks, due to its evolutionary and demographic histories. Moreover, the different degrees of LD in the two quilombo remnant communities, MB and SV, was unexpected due to their similar demographic histories. The urban populations and quilombo remnant communities exhibited similar degrees of LD.

Due to the lack of LD between haplotype blocks and the low LD observed inside a given haplotype block in most of the populations, haplotype diversity was estimated solely on the basis of the male population for each block in each sample. The PC block exhibited the highest haplotype diversity, while the HEMA block exhibited the lowest (Table 3). The TK Amerindian sample exhibited the lowest haplotype diversity in all three blocks.

In the DMD block, 8 of 109 haplotypes were shared by two individuals; one of these haplotypes was shared among individuals from two very different populations (MB and TR); another haplotype was shared among three individuals. In the PC block, a similar number of haplotypes was found (108); 10 of these were shared by two or three individuals and only three were shared among individuals belonging to distinct populations. The HEMA block exhibited the lowest number of haplotypes (81), 20 of which were shared by two or three individuals. Of these, only eight were shared among

individuals of the same population. Therefore, HEMA was the block that shared the most haplotypes among individuals, which is in accordance to its lower physical extension (0.23cM).

DISCUSSION

In the present study, 20 STR loci distributed into three linkage blocks were described with the aim of comparing population parameters within five Brazilian populations with distinct demographic and evolutionary histories. The selection of these three regions was based on three criteria. (1) Primers were designed for the amplification of STR regions in multiplex reactions to facilitate the genotyping of each block. The three selected regions allowed an appropriate designing of these primers. (2) In addition, the dystrophin and factor VIII genes are found in two of these regions (Xp21 and Xq28, respectively), which are of interest in medical genetics studies. (3) Another factor that influenced the selection of these regions was the possibility of comparing LD along the X-chromosome, i.e., to analyze how these three regions behave with regard to recombination and mutation and how they are affected by events such as admixture and endogamy. In addition, the three blocks are distinct with regard to their extensions (DMD-8.24cM; PC-0.66cM, and HEMA-0.23cM), which also affected LD.

Of these three selected blocks, only one (HEMA) had already been described in the literature (Machado et al., 2009a). The other two are novel blocks and are thus important because they will allow the study of 13 new STR regions in the X-chromosome. Till date, these two regions (Xp21 and Xp11-q11) have been little explored in population genetics studies.

Basic population parameters (new alleles, alleles per locus, private alleles, and adherence to the Hardy–Weinberg equilibrium)

The HEMA block was studied in the population of the state of Rio de Janeiro (Machado et al., 2009a). Twenty new alleles were found in the populations from the present study (underlined in Supplementary Table 1). Some of these are low frequency alleles, which may explain why they do not appear in

the population of Rio de Janeiro. The highest numbers of new alleles occurred in the loci REN90833, HEMA154311.3, and TMLHInt5 (7, 5, and 5, respectively). These loci are extragenic (Machado et al., 2009a) and are therefore more prone to mutation, which can lead to the generation of new alleles.

Of the 20 loci that were analyzed, two (DXS31931836 and DXS33091950) were monomorphic only in TK, the only population of Amerindians investigated in the present study. This lack of diversity may be a consequence of the genetic drift present in small, isolated, and endogamous populations.

The average number of alleles/loci is considered as an indication of the intrapopulation diversity when the size of the sample is similar among all the populations under analysis (Nei, 1987). The lowest number of alleles/loci of X-STRs occurred in the TK Amerindian population and in the quilombo remnant communities MB and SV. The average number of alleles/loci obtained for the Amerindian populations in Colombia (Amorim et al., 2011) and in Paraguay (Kohlrausch et al., 2005) was similar to the values obtained for TK, SV, and MB in the present study. This reduced diversity in TK is explained by the genetic drift that occurs in small and isolated populations, as is the case for TK and the quilombo remnant communities. The urban populations exhibited the highest average number of alleles/loci. These are larger, populations that have experienced extensive admixture events, which explains the higher number of alleles/locus (Kohlrausch et al., 2005).

The highest number of private alleles occurred in the quilombo remnant communities. These alleles can be characteristic of African populations and appear in the quilombos because of the large contribution of the African component. Because these populations have a certain degree of isolation, these alleles have become unique to these populations. On the other hand, we expected to observe the highest number of private alleles in the TK population (Schroeder et al., 2007; Schroeder et al., 2009; Acuña-Alonzo et

al., 2010), since Amerindian populations are endogamous, isolated, and small.

The only two loci that were in Hardy–Weinberg disequilibrium (DXS31302675 in SV and F8Int13.2 in RP) after Bonferroni correction exhibited heterozygote deficiency. The first locus can be attributed to preferential marriages, drift, and/or the small size of the sample. With regard to the second locus, the only plausible cause is the bias caused by the sample size.

Population structure (F_{ST} , AMOVA, Clustering)

Analysis of the population structure on the basis of F_{ST} showed that in the male sample, similarity occurred only between MB and SV, which is in agreement with the results of Wiezel et al. (2013) who studied autosomal STRs related with the Y-chromosome. This similarity may be attributed to a similar history of formation of these two quilombos.

In the women's sample, the number of similar populations was higher: SV × TR, SV × RP, and RP × TR. The similarity between SV and TR can only be explained by the effect of the small size of the women's sample analyzed in TR (only 12 individuals) because there was a large difference between the contributions of the African ancestry in both populations. The African contribution in SV, based on Y-chromosome STRs, was estimated at 60.16% (Wiezel et al., 2013), whereas in TR, this estimate (based on INDEL-type markers) was 21.5% (Lopes et al., 2013). With regard to the Y-chromosome STRs (Wiezel et al., 2013), the African contribution in the TR population is larger (20.46%). The difference in the African contribution between SV and TR reinforces the hypothesis that this result was caused by a bias related to the small sample size.

The similarity between SV and RP may also have been caused by the small sample size in SV (25 individuals) because the ancestries of these populations are distinct. The estimate of ancestry of RP is as follows: 0.49 European, 0.3 African, and 0.21 Amerindian (Massaro et al., 2011). In SV,

this estimate includes only European and African ancestries (0.4 and 0.6) (Wiezel et al., 2013). Moreover, these populations are geographically very distant, which would prevent the genetic flow between them.

The similarity between RP and TR can be easily explained by the fact that these are urban populations that, although separated by a great distance, have similar histories of formation, which is revealed by the similar contributions of ancestry: a high European contribution. This ancestry is high in TR both in INDEL analysis (0.6) (Lopes et al., 2013) and Y-chromosome analysis (0.795) (Wiezel et al., 2013). In the state of SP (Manta et al., 2013), the European ancestry was similar (0.63). In RP, the European ancestry obtained from the X-chromosome (Massaro et al., 2011) was also high (0.49). TK was the most distinct population because it exhibited little admixture (<5%) (Neel et al., 1980; Mendes-Junior and Simões, 2009).

AMOVA analysis showed that the populations are similar because it was statistically significant only when the populations were assigned to the same group. This result was in accordance with F_{ST} , which showed similarity among some populations (MB × SV, SV × TR, SV × RP, and TR × RP) and a slight variation among the others. In the results by Wiezel et al. (2013), the same pattern of variation was observed between the quilombo remnant communities and TR.

The clustering of samples identified three clusters, which is in accordance with the trihybrid composition of the Brazilian population. MB and SV were grouped in the first cluster (Table 4), which can thus be considered as representative of the African component, although this cluster also included women from the TR sample. RP and TR men were grouped in the second cluster. The fact that TR women were grouped within the first cluster is related to the small number (12) of women in this population and not to the real difference between the urban samples of TR and RP. The absence of significance in the difference between genders as well as the high contribution of the European component in TR that was estimated by previous studies (Lopes et al., 2013; Wiezel et al., 2013) confirms this

hypothesis. The third cluster is considered to be representative of the Amerindian component because it concentrated most of the TK sample.

Linkage Disequilibrium and Haplotype diversity

A lower haplotype diversity value (h) was expected for the DMD block, which is the most extensive of all (8.24 cM) (Table 3). However, it exhibited h values intermediate to those of the other two blocks. This suggests that this region has higher chromosome stability, with regard to recombination and mutation, than the other two regions under analysis in the present study.

Although the PC block is located in a low recombination region, it exhibited the highest haplotype diversity (Table 3) and was the second most extensive block (0.658 cM). The ratio of 1.1 individuals per haplotype represents a haplotype diversity slightly higher than that described by Edelman et al. (2010) who reported 1.4 individuals per haplotype.

The HEMA block exhibited the lowest number of haplotypes and haplotype diversity (Table 3). The distance between the first and the last STR of this block was only 0.23 cM; therefore, this was the shortest block that was used. In shorter blocks, the possibility of recombination occurring between loci decreases, which explains the lower haplotype diversity that was observed. The haplotype diversity found in the HEMA block in the Rio de Janeiro population sample (Machado et al., 2009a) was higher (0.9786) than that found in the five populations investigated in the present study.

With regard to the populations, the lowest haplotype diversity was observed in TK (Table 3). This result, as well as those that were discussed in the previous sections, are in accordance with the expected results for genetically isolated populations. In TK, this isolation is confirmed by the report of Neel et al. (1980) in which a sample of this Amerindian population did not exhibit more than 5% of other ancestries.

The degree of LD varied both between blocks and between populations (Table 3). The HEMA block exhibited the highest LD. This block is located in the Xq28 region and is only 0.23 cM long (Table 3). Because the markers are so

close to each other, it would be expected that this block would exhibit the highest LD. In other family studies of patients with hemophilia and in the population of the state of Rio de Janeiro (Machado et al., 2009a), a high LD was observed, which is in agreement with our results.

The degree of LD observed in the blocks DMD and PC was higher than expected (in MB and SV, respectively). The most plausible explanation for this phenomenon is the casual variation in samples of this size (25 individuals).

With regard to LD in each population (Table 3), TK presented the highest LD. According to the results mentioned and discussed above, this is the most divergent population among those investigated in the present study. It has a low level of admixture (Neel et al., 1980), in addition to being a genetically isolated and endogamous population, which favors genetic drift. This in turn results in an increase in LD, which explains the high LD observed in TK. The Amerindian populations of Latin America that were also investigated via X-chromosome STR analysis (Leite et al., 2009; Wang et al., 2010; Amorim et al., 2011) exhibited higher LD than urban populations, which is in line with our results.

SV and MB have the same history of population formation; therefore they should exhibit similar LD. However, this was not the case: SV exhibited lower LD than MB. The only factor that explains this finding is, again, the effect caused by the small sample size (25 individuals) in the two quilombo remnant communities.

The continuous gene flow (CGF) model, in which admixture occurs for several generations, as is the case for urban populations, exhibits a slower reduction in LD than the hybrid isolation (HI) model, in which admixture occurs in a single generation and is followed by isolation (Pfaff et al., 2011; Amorim et al., 2011). Therefore, we expected urban populations (CGF model) to exhibit higher LD than the quilombo populations (HI model). However, the degree of LD was similar between these two population groups, which may be a consequence of the small sample size used in each quilombo.

In conclusion, the similarities and differences observed among the populations agree with the results expected for each of them. The populations were grouped into three clusters, which were representative of Brazilian ancestries (European, African, and Amerindian). As expected, TK was the population that exhibited the highest differentiation as well as the lowest diversity and the highest LD. Furthermore, the set of markers that was used generated information that revealed the genetic structure among the populations under analysis as well as the degrees of LD, both among the loci of each block and among the populations.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The study of X-STRs in different population samples was approved by the Comitê de Ética em Pesquisa of the Institution Hospital das Clínicas-FMRP/USP (nº 8711/2011), and by the Comissão Nacional de Ética em Pesquisa, according to processes CONEP 25000.016538/00-11, 25000.003005/00-99 and 25000.097241.2001-83. We are grateful to Mrs. Maria do Carmo Tomitão Canas and Mrs. Ana Lucia Pimentel for the technical assistance in the laboratory analyses.

LITERATURE CITED

- Acuña-Alonzo V, Flores-Dorantes T, Kruit JK, Villarreal-Molina T, Szibor R, et al. 2010. Chromosome X centromere region—Haplotype frequencies for different populations. *Forensic Sci Int Genet Supplement Series* 2:398–399.
- Amorim CEG, Wang S, Marrero AR, Salzano FM, Ruiz-Linares A, Bortolini MC. 2011. X-chromosomal genetic diversity and linkage disequilibrium patterns in Amerindians and non-Amerindian populations. *Am J Hum Biol* 23:299–304.
- Machado FB, Duarte LP, Medina-Acosta E. 2009a. Improved criterion-referenced assessment in indirect tracking of haemophilia A using a 0.23 cM-resolution dense polymorphic marker set. *Haemophilia* 15:1135–1142.
- Arpini-Sampaio Z, Costa MC, Melo AA, Carvalho MF, Deus MS, Simões AL. 1999. Genetic polymorphisms and ethnic admixture in African-derived black communities of northeastern Brazil. *Hum Biol* 71:69–85.
- Machado FB, Medina-Acosta E. 2009b. High-resolution combined linkage physical map of short tandem repeat loci on human chromosome band Xq28 for indirect haemophilia A carrier detection. *Haemophilia* 15: 297–308.
- Machado FB, Duarte LP, Medina-Acosta E. 2009c. A novel informative dinucleotide microsatellite marker located on human factor VIII intron 25. *Haemophilia* 15: 613–614.
- Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580.
- Wolfsberg TG. 2007. Using the NCBI Map Viewer to browse genomic sequence data. *Curr Protoc Bioinformatics Chapter 1:Unit 1 5*.
- Kent WJ, Sugnet CW, Furey TS, et al. 2002. The human genome browser at UCSC. *Genome Res* 12:996–1006.
- Matisse TC, Chen F, Chen W, et al. 2007. A second-generation combined linkage physical map of the human genome. *Genome Res* 17:1783–1786.
- Kibbe WA. 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* 35:W43–46.
- Vallone PM, Butler JM. 2004. AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 37:226–231.
- Cummings L, Riley L, Black L, et al. 2002. Genomic BLAST: custom-defined virtual databases for complete and unfinished genomes. *FEMS Microbiol Lett* 216:133–138.
- Hill CR, Kline MC, Coble MD, Butler JM. 2008. Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J Forensic Sci* 53:73–80.
- Edelmann J, Hering S, Augustin C, Immel UD, Szibor R. 2009. Chromosome X centromere region—Haplotype frequencies for different populations. *Forensic Sci Int Genet Supplement Series* 2:398–399.
- Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47–50.
- Hartl DL, Clark AG. 2007. *Principles of population genetics*. 4th ed. Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts.

- Kohlrausch FB, Callegari-Jacques SM, Tsuneto LT, Petzl-Erler ML, Hill K, Hurtado AM, Salzano FM, Hutz MH. 2005. Geography influences microsatellite polymorphism diversity in Amerindians. *Am J Phys Anthropol* 126:463–470.
- Leite FP, Santos SE, Rodriguez EM, Callegari-Jacques SM, Demarchi DA, Tsuneto LT, Petzl-Erler ML, Salzano FM, Hutz MH. 2009. Linkage disequilibrium patterns and genetic structure of Amerindian and non-Amerindian Brazilian populations revealed by long-range X-STR markers. *Am J Phys Anthropol* 139:404–412.
- Lopes TR, Santos S, Ribeiro-Dos-Santos A, Resque RL, Pinto GR, Yoshioka FK. 2013. Population data of the 46 insertion–deletion (INDEL) loci in population in Piauí State, Northeastern Brazil. *Forensic Sci Int Genet* 8(9):e75145.
- Luizon MR, Mendes-Junior CT, De Oliveira SF, Simoes AL. 2008. Ancestry informative markers in Amerindians from Brazilian Amazon. *Am J Hum Biol* 20:86–90.
- Manta FSN, Pereira R, Vianna R, Araújo AFB, Gitaí DLG, Silva DA, Wolfgramm EV, Pontes IM, Aguiar JI, Moraes MO, Carvalho EF, Gusmão L. 2013. Revisiting the genetic ancestry of Brazilians using autosomal AIM-indels. *PLoS One* 8:e75145.
- Massaro JD, Wiezel CE, Muniz YC, Rego EM, de Oliveira LC, Mendes-Junior CT, Simões AL. 2011. Analysis of five polymorphic DNA markers for indirect genetic diagnosis of haemophilia A in the Brazilian population. *Haemophilia* 17:936–943.
- Mendes-Junior CT, Castelli EC, Simoes AL, Donadi EA. 2007. Absence of the HLA-G*0105N allele in Amerindian populations from the Brazilian Amazon Region: a possible role of natural selection. *Tissue Antigens* 70:330–334.
- Mendes-Junior CT, Simões AL. 2009. Mitochondrial DNA variability among eight Tikúna villages: evidence for an intratribal genetic heterogeneity pattern. *Am J Phys Anthropol* 140:526–531.
- Neel JV, Gershowitz H, Mohrenweiser HW, Amos B, Kostyu DD, Salzano FM, Mestriner MA, Lawrence D, Simões AL, Smouse PE, Oliver WJ, Spielman RS, Neel JV Jr. 1980. Genetic studies on the Ticuna, an enigmatic tribe of Central Amazonas. *Ann Hum Genet* 44:37–54.
- Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.
- Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy Fde S, Kohlrausch F, Magno LA, Montenegro RC, Moraes MO, de Moraes ME, de Moraes MR, Ojopi EB, Perini JA, Racciopi C, Ribeiro-Dos-Santos AK, Rios-Santos F, Romano-Silva MA, Sortica VA, Suarez-Kurtz G. 2011. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One* 6:e17063.
- Penna LS, Silva FG, Salim PH, Ewald G, Jobim M, Magalhães JA, Jobim LF. 2012. Development of two multiplex PCR systems for the analysis of 14 X-chromosomal STR loci in a southern Brazilian population sample. *Int J Legal Med* 126:327–330.
- Pfaff CL, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD. 2001. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet* 68:198–207.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Rousset F. 2008. GENEPOP'007: a complete reimplement of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8:103–106.

- Salzano FM, Bortolini MC. 2002. The evolution and genetics of Latin American populations. Cambridge: Cambridge University Press.
- Santos-Lopes SS, Pereira RW, Wilson IJ, Pena SDJ. 2007. A worldwide phylogeography for the human X chromosome. *PLoS One* 2:e557.
- Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, Conrad DF, Tito RY, Osipova LP, Tarskaia LA, Zhadanov SI, Wall JD, Pritchard JK, Malhi RS, Smith DG, Rosenberg NA. 2009. Haplotypic background of a private allele at high frequency in the Americas. *Mol Biol Evol* 26:995–1016.
- Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. A private allele ubiquitous in the Americas. 2007. *Biol Lett* 3:218–223.
- Wang S, Bedoya G, Labuda D, Ruiz-Linares A. 2010. Brief communication: patterns of linkage disequilibrium and haplotype diversity at Xq13 in six Native American populations. *Am J Phys Anthropol* 142:476–480.
- Wiesel CEV, Luizon MR, Sousa SMB, Santos LMW, Muniz YCN, Mendes-Junior CT, Simões AL. 2013. Y-Linked microsatellites in Amazonian Amerindians applied to ancestry estimates in Brazilian Afro-derived populations. *Am J Hum Biol* 25:313–317.

TABLES

TABLE 1. Novel STR loci markers tested-and-tried in this study.

Locus ^a	Repeat ^b	Physical interval ^c	Gene location	Copies ^d	Map length ^e
DXS31302675	TTC	31302625--31302831	<i>DMDIntn78.2</i>	16.7	53.3270
DXS31931836	AAAAC	31931802--31932007	<i>DMDInt62</i>	11.8	54.7628
DXS32159553	AAAACA	32159403--32159628	<i>DMDInt36.3</i>	7.5	55.5868
DXS32154670	AAACA	32154588--32154861	<i>DMDInt36.2</i>	9.4	55.5646
DXS32251717	TTTA	32251680--32251815	<i>DMDInt36.1</i>	12.5	56.1903
DXS32426983	ATCT	32426703--32427069	<i>DMDInt35.1</i>	11.8	57.3406
DXS33091950	AAAT	33091924--33092130	<i>DMDInt19</i>	10.8	61.5753
DXS56749371	AGAA	56749371--56749571	intergenic	31.3	90.0057
DXS56909198	TTTA	56909000--56909298	intergenic	12.8	90.1025
DXS57434969	TATT	57434869--57435069	<i>FAAH2Int7</i>	8.3	90.4398
DXS61920979	TTTC	61920879--61921079	intergenic	18.8	90.6504
DXS62085977	AAATA	62085877--62086077	intergenic	18.2	90.6569
DXS62290471	AAAG	62290371--62290671	intergenic	29.3	90.6642

- ^a Newly described STR loci were named alphanumerically with DXS followed by the physical distance in base pairs from Xp-tel (assembly Feb. 2009 GRCh37/hg19, NC_000023) relative to the first nucleotide of the repeat locus as identified by TRF (Benson, 1999).
- ^b Core, consensus sequence pattern of the repeat element (upper strand) (assembly Feb. 2009 GRCh37/hg19, NC_000023).
- ^c Distance in base pairs from Xp-tel, as determined using the human X-chromosome reference sequence (GRCh37/hg19, NC_000023). Shown is the interval distance for each STR-containing amplicon.
- ^d Repeat copy number as determined by TRF (Benson, 1999).
- ^e Interpolated genetic distances in centiMorgans (females), determined using the online regression-based smoothed Rutgers map v.2 interpolator of the human genome with human X-chromosome reference sequence NCBI36/hg18 coordinates (Matisse et al. 2007).

TABLE 2. Primer sequences.

Marker	Forward	Reverse ^a	Amplimer ^b	Fluorochrome
DXS31302675	TCAGAAAGGAAAACACAAACTCC	CCGGTTATCCGAACACTTTG	207	FAM-6
DXS31931836	ACAAGGTCAGCATCCATTCC	TGAAGGCAAAGCTCACTGAA	206	VIC
DXS32159553	AATAGCCTGGCGTGGTTGT	CACTGAGGCCATTTAGAAGAGAA	226	NED
DXS32154670	CTGGCAGCCGAAGTCAAT	CCATCGGTGGATTAGAATTTGT	274	FAM-6
DXS32251717	CATTCCAATCCACTTCCTG	ATCACACCATTGCACTCCAG	136	NED
DXS32426983	TGGAGTCAGACCCCACAAA	GTGGGGTGGATAGATGGATG ^a	367	PET
DXS33091950	GGACAACGAGAGCAAAACTCT	GAGTCCTTCTCAGGCCGTATT	207	PET
DXS56749371	TCACTTGAACCCAGGAGGAG	GTGACATGTTGCTGGATTTGG ^a	237	PET
DXS56909198	TGATTTTCTCTGGTGGTAGGC	GTGCAAGCATCACGGTAACAT ^a	123	NED
DXS57434969	GGATTTTTACGTATGAGAACATGTCA	GTGGAAGTCTTAGCCAGAGCAA ^a	128	FAM-6
DXS61920979	CCAACCTGGTGAACCAAAAG	GAGGGAAATTTGGGAGACCAT ^a	212	VIC
DXS62085977	CTGGGCAATGTAGCCAAACT	GAAACAAAATTGACAGAGCACTAGC ^a	311	NED
DXS62290471	TTCCTGCTGATTGCTGTTTG	GTCAGGGGTTTCTGCTTTTGA ^a	342	FAM-6

^a The non-templated 5' guanine residue in these reverse primers was added to promote amplimer adenylation by Taq DNA polymerase (Hill et al., 2008).

^b *In silico* reference genome length in base pairs at the UCSC Genome Browser (<http://genome.ucsc.edu>) (Kent et al., 2002)

TABLE 3. Genetic information regarding the five populations analyzed (MB: Mimbó, SV: Sítio Velho, TR: Teresina, RP: Ribeirão Preto, and TK: Tikúna) on 20 loci distributed in three linkage blocks (DMD, PC, and HEMA). The numbers in parenthesis indicate the distance (cM) between the first and last loci of each haploblock.

Characteristic	MB	SV	TR	RP	TK
Average gene diversity (\pm SD)	0.6876 +/- 0.3507	0.6967 +/- 0.3552	0.6959 +/- 0.3510	0.6847 +/- 0.3453	0.5394 +/- 0.2751
Average no. of alleles per locus	5,0	4,8	6,4	6,15	4,8
Proportion (%) of loci in LD before correction					
DMD (8,23 cM)	52,38	57,14	28,57	14,29	14,29
PC (0,66 cM)	33,33	60,00	26,67	20,00	86,67
HEMA (0,23 cM)	38,10	42,86	76,19	95,24	100,00
Proportion (%) of loci in LD after correction					
DMD(8,23 cM)	38,10	9,52	9,52	4,76	9,52
PC (0,66 cM)	20,00	33,33	13,33	6,67	80,00
HEMA (0,23 cM)	33,33	23,81	47,62	76,19	95,24
Haplotype Diversity (\pmSD)					
DMD	0,668 \pm 0,393	0,719 \pm 0,423	0,666 \pm 0,373	0,646 \pm 0,362	0,419 \pm 0,253
PC	0,750 \pm 0,447	0,660 \pm 0,397	0,731 \pm 0,412	0,720 \pm 0,405	0,659 \pm 0,380
HEMA	0,618 \pm 0,373	0,665 \pm 0,392	0,627 \pm 0,354	0,647 \pm 0,362	0,513 \pm 0,299

TABLE 4. Proportion of the genetic composition of the samples of women and men from all the populations divided into three clusters.

Population	Males			Females		
	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 3	Cluster 2
MB	0.115	0.835	0.050	0.161	0.750	0.089
SV	0.152	0.631	0.217	0.074	0.792	0.134
TR	0.180	0.541	0.279	0.261	0.287	0.452
RP	0.026	0.292	0.682	0.050	0.039	0.551
TK	0.915	0.045	0.040	0.862	0.039	0.099

SUPPLEMENTARY TABLE 1. Allele frequencies of 20 X-chromosome STRs in samples of the Mimbó, Sítio Velho, Teresina, Ribeirão Preto, and Tikúna populations. The alleles were identified via the number of base pairs (bp). N: number of X-chromosomes analyzed; F: females; M: males.

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
Block DMD															
DXS31302675															
N	13	24	37	12	24	36	38	24	62	36	28	64	27	46	73
181				0,083	0,167	0,139	0,079	0,083	0,081						
190					0,042	0,028	0,079	0,083	0,081	0,028	0,143	0,078			
193	0,308	0,542	0,459				0,079	0,208	0,129	0,083	0,143	0,109			
196	0,231		0,081	0,167	0,083	0,111	0,053		0,032	0,028	0,071	0,047	0,037	0,087	0,068
199					0,042	0,028	0,026		0,016	0,028	0,036	0,031	0,037		0,014
202	0,231	0,167	0,189	0,083	0,083	0,083	0,106	0,167	0,129	0,083	0,036	0,063	0,074	0,304	0,219
205	0,154	0,042	0,081	0,083	0,042	0,056	0,026	0,083	0,048	0,028		0,016	0,111	0,152	0,137
208		0,042	0,027		0,042	0,028	0,026	0,042	0,032	0,056	0,036	0,047	0,148	0,065	0,096
211	0,077	0,042	0,054				0,079	0,042	0,065		0,071	0,031	0,074		0,027
214							0,026		0,016	0,056	0,036	0,047			
217					0,250	0,167				0,056		0,031	0,037	0,109	0,082
220				0,083	0,042	0,056	0,053	0,042	0,048	0,167	0,179	0,172		0,022	0,014
223		0,083	0,054				0,079	0,042	0,065	0,083	0,071	0,078			
226				0,083		0,028	0,079	0,125	0,097	0,056	0,107	0,079	0,074	0,109	0,096
229		0,017	0,027	0,083	0,083	0,083	0,053	0,042	0,048	0,056	0,036	0,047	0,037	0,022	0,027
232		0,042	0,027	0,083	0,042	0,056	0,079	0,042	0,065	0,111		0,063	0,111	0,022	0,055
235							0,053		0,032	0,028		0,016	0,074	0,022	0,041

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
238				0,083	0,083	0,083	0,026		0,016		0,036	0,016	0,074		0,027
241				0,167		0,056				0,028		0,016			
244													0,111	0,065	0,082
250										0,028		0,016		0,022	0,014
DXS31931836															
N	13	24	37	13	24	37	38	24	62	37	28	75	27	46	73
197	0,385	0,208	0,270	0,462	0,458	0,459	0,237	0,083	0,177	0,270	0,214	0,246			
202	0,462	0,583	0,541	0,385	0,458	0,432	0,605	0,875	0,710	0,649	0,536	0,600	1,000	1,000	1,000
207	0,154	0,208	0,189	0,154	0,083	0,108	0,158	0,042	0,113	0,081	0,250	0,154			
DXS32159553															
N	13	24	37	13	24	37	38	24	62	37	28	75	27	46	73
210										0,027		0,015		0,022	0,014
216	0,615	0,458	0,514	0,308	0,458	0,405	0,526	0,583	0,548	0,649	0,643	0,646	0,630	0,630	0,630
222	0,385	0,458	0,432	0,462	0,292	0,351	0,368	0,333	0,355	0,189	0,321	0,246	0,370	0,348	0,356
228		0,083	0,054	0,231	0,250	0,243	0,105	0,083	0,097	0,135	0,036	0,092			
DXS32154670															
N	13	24	37	13	24	37	38	24	62	37	28	65	27	46	73
254								0,167	0,065	0,054		0,031			
257	0,231	0,250	0,243	0,538	0,292	0,378	0,105		0,097	0,108		0,108	0,259	0,239	0,247
259							0,158	0,166	0,129	0,027	0,143	0,031			

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
264	0,692	0,750	0,730	0,385	0,542	0,486	0,658	0,667	0,661	0,649	0,714	0,677	0,741	0,739	0,740
267					0,167	0,135	0,053		0,032	0,027	0,107	0,107		0,022	0,014
269	0,077		0,27	0,077		0,27	0,026		0,016	0,135	0,036	0,046			
DXS32251717															
N	13	24	37	13	24	37	35	24	57	37	28	65	26	46	72
114							0,029		0,018						
122	0,077		0,027		0,083	0,054		0,136	0,053	0,027	0,143	0,077			
126	0,077	0,042	0,054		0,083	0,054	0,029		0,018	0,054	0,071	0,062		0,022	0,014
130	0,385	0,500	0,459	0,154	0,208	0,189	0,171	0,227	0,193	0,108	0,250	0,169		0,109	0,069
134	0,154	0,292	0,243	0,462	0,333	0,378	0,600	0,318	0,491	0,595	0,214	0,431	0,654	0,543	0,583
138	0,308	0,167	0,216	0,385	0,292	0,324	0,171	0,318	0,228	0,162	0,214	0,185	0,346	0,326	0,333
142											0,107	0,046			
146										0,027		0,015			
148										0,027		0,015			
DXS32426983															
N	13	24	37	13	22	35	34	20	54	37	28	65	27	46	73
356	0,308	0,292	0,297	0,077	0,045	0,057	0,029		0,019						
358								0,100	0,037					0,022	0,014
360		0,083	0,054				0,059		0,037	0,054		0,031		0,022	0,014
362							0,088	0,050	0,074						
364		0,083	0,054	0,154	0,318	0,257	0,059	0,050	0,056	0,189	0,179	0,185		0,022	0,014

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
366					0,045	0,029	0,265	0,200	0,241				0,037	0,022	0,027
368	0,462	0,208	0,297	0,462	0,182	0,286	0,118	0,050	0,093	0,270	0,393	0,323	0,481	0,413	0,438
370							0,206	0,250	0,222						
372	0,154	0,292	0,243	0,308	0,227	0,257	0,088	0,150	0,111	0,351	0,357	0,354	0,370	0,478	0,438
374					0,045	0,029	0,029	0,050	0,037						
376	0,077	0,042	0,054		0,136	0,086	0,029	0,100	0,056	0,135	0,036	0,092	0,111	0,022	0,055
378							0,029		0,019		0,036	0,015			
DXS33091950															
N	13	24	37	13	24	37	36	24	60	37	28	65	27	46	73
189											0,107	0,046			
193										0,054	0,036	0,046			
197	0,385	0,125	0,216	0,462	0,125	0,243	0,167	0,042	0,117						
201	0,462	0,667	0,595	0,385	0,667	0,568	0,694	0,833	0,750	0,459	0,214	0,354	1,000	1,000	1,000
205	0,154	0,208	0,189	0,154	0,208	0,189	0,139	0,083	0,117	0,459	0,500	0,477			
209								0,042	0,017	0,027	0,107	0,062			
213											0,036	0,015			
Block PC															
DXS56749371															
N	13	24	37	13	24	37	38	24	62	36	28	64	26	46	72
210							0,026		0,016						
214		0,042	0,027				0,053		0,032						

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
218	0,077	0,167	0,135	0,077	0,167	0,135	0,158	0,042	0,113	0,056	0,036	0,047		0,022	0,014
222	0,462	0,500	0,486	0,385	0,208	0,270	0,316	0,125	0,242	0,222	0,357	0,281	0,462	0,478	0,472
226	0,077	0,042	0,054	0,077	0,167	0,135	0,184	0,125	0,161	0,250	0,179	0,219	0,154	0,261	0,222
230	0,154	0,083	0,108	0,231	0,292	0,270	0,211	0,500	0,323	0,333	0,250	0,297	0,077	0,087	0,083
234	0,154	0,125	0,135	0,231	0,167	0,189	0,026	0,125	0,065	0,111	0,179	0,141	0,115	0,065	0,083
236	0,077	0,042	0,054					0,083	0,032				0,154	0,065	0,097
240							0,026		0,016	0,028		0,016	0,038	0,022	0,028
DXS56909198															
N	13	24	37	13	24	37	37	24	61	36	28	64	26	46	72
118								0,042	0,016						
122	0,385	0,250	0,297	0,385	0,292	0,324	0,189	0,333	0,246	0,389	0,357	0,375	0,385	0,174	0,250
124					0,042	0,027									
126	0,231	0,333	0,297	0,154	0,208	0,189	0,108	0,292	0,180	0,111	0,143	0,125		0,022	0,014
130	0,385	0,292	0,324	0,385	0,417	0,405	0,622	0,333	0,508	0,500	0,464	0,484	0,615	0,804	0,736
134		0,125	0,081	0,077	0,042	0,054	0,081		0,049		0,036	0,016			
DXS57434969															
N	13	24	37	13	24	37	38	24	62	36	28	64	26	46	72
121										0,028		0,016			
125	0,615	0,542	0,568	0,846	0,833	0,838	0,421	0,625	0,500	0,472	0,536	0,500	0,192	0,109	0,139
129	0,385	0,417	0,405	0,154	0,167	0,162	0,500	0,375	0,452	0,500	0,321	0,422	0,769	0,891	0,847
133		0,042	0,027				0,079		0,048		0,143	0,063	0,038		0,014

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
DXS61920979															
N	13	24	37	13	24	37	38	24	62	35	26	61	26	46	72
193								0,042	0,016		0,038	0,016			
197							0,053	0,042	0,048						
201	0,077	0,042	0,054				0,079	0,167	0,113	0,029	0,038	0,033	0,038		0,014
205	0,154	0,125	0,135				0,026	0,167	0,081	0,200	0,115	0,164	0,038		0,014
209	0,231	0,333	0,297	0,538	0,375	0,432	0,342	0,167	0,274	0,371	0,385	0,377	0,192	0,283	0,250
213	0,308	0,292	0,297	0,308	0,208	0,243	0,316	0,250	0,290	0,171	0,308	0,230	0,269	0,196	0,222
217	0,231	0,208	0,216	0,154	0,250	0,216	0,079	0,083	0,081	0,143	0,077	0,115	0,462	0,435	0,444
221					0,167	0,108	0,053	0,083	0,065	0,086	0,038	0,066		0,087	0,056
225							0,053		0,032						
DXS62085977															
N	12	18	30	13	24	37	37	24	61	36	28	64	26	44	70
303													0,038		0,014
305	0,083	0,056	0,067				0,054	0,125	0,082		0,036	0,016			
310				0,077	0,083	0,081	0,027		0,016	0,028	0,036	0,031			
315	0,250	0,278	0,267	0,231	0,375	0,324	0,135	0,292	0,197	0,111	0,321	0,203	0,038	0,023	0,029
320	0,417	0,444	0,433	0,231	0,250	0,243	0,216	0,250	0,230	0,111		0,063	0,308	0,341	0,329
325	0,083	0,056	0,067				0,027	0,042	0,033	0,111	0,036	0,078	0,231	0,136	0,171
330	0,083	0,111	0,100	0,077	0,208	0,162	0,270	0,208	0,246	0,222	0,179	0,203	0,192	0,114	0,143
332													0,115	0,045	0,072

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
335	0,083		0,033	0,385	0,083	0,189	0,135	0,042	0,098	0,139	0,286	0,203		0,205	0,129
340		0,056	0,033				0,108	0,042	0,082	0,278	0,107	0,203	0,077	0,136	0,114
350							0,027		0,016						
DXS62290471															
N	12	22	34	13	22	35	33	24	57	36	28	64	26	46	72
310														0,065	0,042
318				0,077	0,091	0,086									
322	0,083	0,227	0,176	0,154		0,057	0,154		0,018		0,107	0,047	0,077	0,109	0,097
326	0,083		0,029	0,077	0,227	0,171	0,182	0,125	0,158	0,167	0,179	0,172	0,077	0,022	0,042
330	0,250	0,182	0,206	0,077	0,364	0,257	0,273	0,333	0,298	0,222	0,250	0,234	0,077	0,174	0,139
334	0,333	0,227	0,265	0,538	0,182	0,314	0,273	0,333	0,298	0,417	0,286	0,375	0,346	0,196	0,250
338	0,083	0,136	0,118	0,077	0,045	0,057	0,212	0,125	0,175	0,139	0,143	0,141	0,192	0,130	0,153
342	0,167	0,227	0,206		0,045	0,029		0,042	0,018				0,231	0,239	0,236
344										0,056	0,036	0,031			
346					0,045	0,029	0,030	0,042	0,035					0,065	0,042
Block HEMA															
REN90833															
N	13	22	35	13	22	35	38	22	60	37	26	63	27	46	73
<u>114</u>	0,231	0,182	0,200	0,308	0,136	0,200	0,026	0,091	0,050	0,054		0,032			
<u>118</u>	0,077	0,091	0,086	0,231	0,273	0,257	0,526	0,409	0,483	0,622	0,538	0,587	0,259	0,261	0,260
<u>122</u>		0,045	0,029				0,079		0,050	0,027	0,038	0,032	0,333	0,239	0,274

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
<u>126</u>	0,231	0,182	0,200	0,231	0,273	0,257	0,132	0,091	0,117	0,081	0,231	0,143			
<u>130</u>	0,385	0,500	0,457	0,154	0,273	0,229	0,211	0,318	0,250	0,189	0,192	0,190	0,407	0,391	0,397
<u>132</u>														0,043	0,027
<u>134</u>	0,077		0,029	0,077	0,045	0,057	0,026	0,045	0,033	0,027		0,016		0,065	0,041
138								0,045	0,017						
F8Int25.2															
N	13	22	35	13	22	35	38	22	60	37	26	63	27	46	73
143	0,923	0,818	0,857	0,462	0,727	0,629	0,342	0,500	0,400	0,351	0,346	0,349	0,444	0,500	0,493
146		0,045	0,029	0,385	0,091	0,200	0,079	0,182	0,117	0,108	0,077	0,095			
148	0,077	0,045	0,057	0,154	0,182	0,171	0,579	0,318	0,483	0,541	0,577	0,556	0,556	0,500	0,507
<u>152</u>		0,090	0,058												
F8Int22															
N	13	24	37	13	24	37	28	22	50	37	26	63	27	46	73
<u>212</u>										0,027		0,016			
214		0,042	0,027				0,026		0,017	0,027		0,016			
216	0,077	0,250	0,189	0,154		0,054	0,079	0,091	0,083	0,108	0,192	0,143			
218	0,154	0,167	0,162	0,308	0,250	0,270	0,026	0,045	0,033	0,054		0,032			
220	0,692	0,333	0,459	0,385	0,542	0,486	0,289	0,545	0,383	0,243	0,423	0,317	0,407	0,522	0,479
222	0,077	0,208	0,162	0,154	0,167	0,162	0,526	0,318	0,450	0,486	0,346	0,429	0,593	0,478	0,521
224							0,053		0,033						
226					0,042	0,027				0,054	0,038	0,048			

Block/ Locus / Allele	Mimbó			Sítio Velho			Teresina			Ribeirão Preto			Tikúna		
	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL	M	F	TOTAL
<u>346</u>	0,077		0,029												
<u>354</u>													0,022	0,014	
TMLHInt5															
N	12	12	24	13	20	33	35	18	53	37	26	63	27	44	71
<u>274</u>	0,083	0,167	0,125												
<u>284</u>	0,417	0,417	0,417	0,538	0,600	0,576	0,343	0,500	0,396	0,270	0,346	0,302	0,481	0,477	0,479
<u>289</u>	0,083	0,250	0,167				0,029	0,056	0,038	0,081		0,048			
<u>294</u>				0,462	0,300	0,364	0,429	0,333	0,396	0,595	0,654	0,619		0,023	0,014
<u>299</u>	0,417	0,167	0,292		0,100	0,061	0,200	0,111	0,170	0,054		0,032	0,519	0,500	0,507
HEMA154507.3															
N	11	12	23	13	16	29	32	16	48	37	26	63	27	44	71
342														0,023	0,014
346	0,273	0,167	0,217	0,154	0,125	0,138	0,063	0,188	0,104	0,081	0,154	0,111		0,045	0,028
350	0,182	0,167	0,174	0,231	0,438	0,345	0,250	0,250	0,250	0,216	0,308	0,254	0,444	0,318	0,366
354	0,091	0,250	0,174	0,154	0,313	0,241	0,281	0,125	0,229	0,378	0,423	0,397	0,037	0,023	0,028
358	0,455	0,333	0,391	0,462	0,125	0,276	0,375	0,438	0,396	0,243	0,038	0,159	0,407	0,477	0,451
362		0,083	0,043				0,031		0,021	0,081	0,077	0,079	0,111	0,114	0,113