Universidade de São Paulo
Faculdade de Medicina de Ribeirão Preto
Departamento de Genética

AYLING MARTINS NG

Validação de marcadores
inserção/deleção para genotipagem fetal
não invasiva

Ribeirão Preto
2015

AYLING MARTINS NG

Validação de marcadores
inserção/deleção para genotipagem fetal
não invasiva

Dissertação apresentada à
Faculdade de Medicina de
Ribeirão Preto da Universidade
de São Paulo para obtenção
do título de Mestre em Ciências

Área de concentração:
Genética

Orientador:
Prof. Dr. Aguinaldo Luiz Simões

Ribeirão Preto
2015

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

Ng, Ayling Martins.

Validação de marcadores inserção/deleção para genotipagem fetal não invasiva/ Ayling Martins Ng; orientador: Aguinaldo Luiz Simões — Ribeirão Preto, 2015.

68f. il. 30cm

Dissertação de Mestrado – Programa de Pós Graduação em Genética – Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

1. InDel; 2. Frequência alélica; 3. PCR; 4. Primer inserção-específico; 5. Genotipagem fetal não invasiva.

FOLHA DE APROVAÇÃO

NG, Ayling Martins

Validação de marcadores inserção/deleção para genotipagem fetal não invasiva

Dissertação apresentada à
Faculdade de Medicina de
Ribeirão Preto da Universidade
de São Paulo para obtenção do
título de Mestre em Ciências

Aprovado em: ____/____/____

Banca examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

*Aos alunos cientistas que cutilizarão
esse trabalho, que esteja claro e
compreensível.*

*Aos meus sobrinhos, para que eles se inspirem
e se sintam capazes de fazer muito mais.*

Agradecimentos

À minha família pelo apoio e torcida. Em especial, à May e a Tay pela atenção e a Pedro Dantas por estarmos conseguindo continuar o nosso caminho de mãos dadas e com amor. À minha mãe, D. Nilda, por enxergar o que há de melhor em mim e pra mim. Se “a lenda pessoal é aquilo que você sempre desejou fazer” e a minha é ser mais e mais feliz, ela é a grande responsável por isso! Mãe, mais uma vitória nossa!

Ao meu orientador, professor Aguinaldo Simões, por ter acreditado no meu potencial, pelos momentos de orientação e reflexão e pelos ensinamentos.

À Claudia Wiesel, pela ajuda essencial no desenvolvimento do trabalho, sem a qual não teria conseguido atender meus prazos. E à Vanessa e Polyana, ICs que me ajudaram na reta final.

À Maria Carmo e Ana Lúcia pelo carinho e apoio nos momentos de dúvida. Edna, Nádia, Sabrina, Marcela, Igor, Gêssica e aos outros colegas do bloco B. À galera do bloco G e à professora Catarina, meu muito obrigado por todo acolhimento e carinho.

Agradeço também ao professor Tiago Campos Pereira pelas conversas e orientações da pesquisa à docência e ao professor de corrida Filipe Oliveira, que me ajudou a ser cada dia mais saudável. *Mens sana in corpore sano.*

Aos amigos: Isabela, Fernanda, Estela, Júlia, Jaque, João, Jorge, Saulo e principalmente Natália, Simone, Melina e Duda por terem se tornado a minha família ribeirão-pretana na alegria e na tristeza, na saúde e na doença. Sem vocês, teria sido tudo mais difícil.

Obrigada família B05 e os outros amigos baianos, principalmente Dinho, Nava, Marquinhos, Ohara, Ramon, Juli, as meninas Marcela, Hanna, Flavia e Nanda pelas preocupações e conselhos, por se alegrarem com minhas conquistas e suportarem minhas reclamações.

Agradeço também a todos aqueles que não foram citados, mas que se sentem fazendo parte desta conquista.

Pelo apoio financeiro indispensável à produção acadêmica, agradeço a: Programa de Excelência Acadêmica da Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES/Proex) e Fundação de Apoio ao Ensino, Pesquisa e Assistência (FAEPA) do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto – USP.

Non scholæ sed vitæ discimus

Séneca

Sumário

RESUMO	10
ABSTRACT	11
LISTA DE FIGURAS	12
LISTA DE TABELAS	13
LISTA DE ABREVIATURAS E SIGLAS	14
INTRODUÇÃO	16
POLIMORFISMOS GENÉTICOS	16
MARCADORES MOLECULARES NA IDENTIFICAÇÃO HUMANA	17
STRs	17
SNPs	17
INDELS	18
GENOTIPAGEM FETAL NÃO INVASIVA	19
DNA FETAL LIVRE DE CÉLULAS	19
DIAGNÓSTICO PRÉ-NATAL NÃO INVASIVO	21
IDENTIFICAÇÃO HUMANA PRÉ-NATAL NÃO INVASIVA	21
HIPÓTESE, JUSTIFICATIVA E OBJETIVO	23
OBJETIVOS OPERACIONAIS	24
MATERIAL E MÉTODOS	25
DESENHO DO ESTUDO	25
AMOSTRAS BIOLÓGICAS	25
COLETA E CONSERVAÇÃO	25
ASPECTOS ÉTICOS	26
MARCADORES GENÉTICOS ANALISADOS	26
DESENHO DE INICIADORES	27
ANÁLISE LABORATORIAL	29
REAÇÃO EM CADEIA DA POLIMERASE (PCR)	30
ANÁLISE DO PRODUTO AMPLIFICADO	31
DETERMINAÇÃO FENOTÍPICA	32
ANÁLISE ESTATÍSTICA	32
DISTRIBUIÇÃO DE FREQUÊNCIAS	32
ADERÊNCIA AO EQUILÍBRIO DE HARDY-WEINBERG	33
HETEROZIGOSE	34
PARÂMETROS FORENSES	34
CONTEÚDO DE INFORMAÇÃO DO POLIMORFISMO	34
PODER DE DISCRIMINAÇÃO E PROBABILIDADE DE IDENTIDADE	35

PODER DE EXCLUSÃO	35
RESULTADOS	37
CLASSIFICAÇÃO FENOTÍPICA	37
DNA NÃO AMPLIFICADO (NA)	38
ANÁLISE ESTATÍSTICA	40
PARÂMETROS FORENSES	41
DISCUSSÃO	43
AUSÊNCIA DE RESULTADOS	43
ESCOLHA DOS LÓCUS	43
ANÁLISE ESTATÍSTICA	45
PARÂMETROS FORENSES	47
GENOTIPAGEM FETAL NÃO INVASIVA	48
CONCLUSÃO	49
REFERÊNCIAS BIBLIOGRÁFICAS	50
APÊNDICES	55

Resumo

NG, A. M. **Validação de marcadores inserção/deleção para genotipagem fetal não invasiva.** 2015. 68p. Dissertação (Mestrado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

A presença de DNA fetal livre de células no plasma materno possibilitou o surgimento de novas tecnologias para o diagnóstico pré natal não invasivo. Existem técnicas de genotipagem fetal já estabelecidas em investigações de sequências exclusivas fetais, mas a sua aplicação para a identificação humana em testes de paternidade ainda é pouco conhecida. A metodologia de genotipagem fetal não invasiva foi padronizada com o uso de três loci InDels e iniciadores inserção-específicos. Entretanto, para que se alcance o poder satisfatório, é necessário investigar elevado número de loci informativos. Para suprir a ausência de informações suficientes sobre loci deste tipo, são descritas, no presente trabalho, as condições laboratoriais para PCR e as frequências alélicas de um conjunto de loci InDels, ainda inéditos, em uma amostra representativa da população urbana brasileira, visando aumentar a robustez e a confiabilidade da genotipagem fetal não invasiva na investigação de paternidade. Foram selecionados 20 loci, um em cada cromossomo. Em um dos loci polimórficos foi encontrado um alelo novo, não registrado no banco de dados do NCBI. O poder de discriminação e o poder de exclusão dos 16 loci polimórficos combinados (0,999 e 0,937, respectivamente) tiveram valores dentro do esperado para um conjunto com essa quantidade de loci bialélicos. Alguns loci apresentaram elevado número de indivíduos que não responderam à amplificação, indicando a necessidade de reajustes na padronização dos procedimentos laboratoriais e a possibilidade de variabilidade das sequências complementares aos iniciadores empregados, o que exigirá o sequenciamento completo das regiões. Para a maioria dos loci, entretanto, os parâmetros forenses atingiram valores esperados e sugerem que podem ser úteis na composição de painel robusto e eficiente para ser usado na genotipagem fetal não invasiva.

Palavras-chave: 1. InDel; 2. Frequência alélica; 3. PCR; 4. Primer inserção-específico; 5. Genotipagem fetal não invasiva.

Abstract

NG, A. M. **Insertion/deletion markers validation for non invasive fetal genotyping.** 2015. 68p. Dissertação (Mestrado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

The presence of cell-free fetal DNA in maternal plasma allows for new non-invasive prenatal diagnosis technologies to develop. Even though there are already well-established fetal genotyping techniques in fetal-exclusive sequence investigations, their application to human identification in paternity tests is not yet well-known. Non-invasive fetal genotyping methodology was previously standardized using three InDel loci and insertion-specific primers. However, in order to attain satisfactory power, it is necessary to investigate large number of informative loci. To compensate for the absence of sufficient information about this type of locus, this work describes the PCR conditions and determinates allelic frequencies for a set of unpublished InDel loci, in a representative group of Brazilian population, raising the robustness and reliability of non-invasive fetal genotyping in paternity investigation. We selected 20 loci, one in each chromosome. Not previously registered on NCBI database allele was found in one of the polymorphic loci. Discrimination and exclusion power of the 16 polymorphic loci combined (0.999 and 0.937, respectively) had expected values for an ensemble with such an amount of biallelic loci. Some loci showed many non-amplified individuals, vining the need of corrections in the standardizing process and the possible variability of the complementary sequences to the primers used. However, for most loci, the forensic parameters reached the expected values and suggest that they may be useful in the more robust panel for utilization in non-invasive fetal genotyping.

Keywords: 1. InDel; 2. Allelic frequency; 3. PCR; 4. Insertion-specific primer; 5. Non invasive fetal genotyping.

Lista de figuras

Figura 1: Métodos de coleta de material genético fetal de forma invasiva (A e B) e não invasiva (C)_____ **21**

Figura 2: Esquema de posicionamento par a par do trio de iniciadores dos loci InDels. (A) Os dois iniciadores externos, nas regiões flanqueadoras em ambos os sentidos. (B) Um iniciador em sentido direto interno, complementar à inserção e um em sentido inverso externo, complementar à região flanqueadora. (C) Um iniciador em sentido direto externo, complementar à região flanqueadora e um inverso interno, complementar à inserção_____ **30**

Figura 3: Fotos do padrão eletroforético visualizado para os loci InDels amplificados individualmente. Os números indicam o peso molecular dos alelos representados pelas bandas. (*) No locus L2, o alelo de maior peso molecular ainda não foi descrito e seu tamanho é aproximado_____ **38**

Figura 4: Fotos do padrão eletroforético de cada InDel amplificado em sistema duplex. Os números indicam o peso molecular dos alelos representados pelas bandas_____ **39**

Figura 5: Esquema mostrando possíveis alelos do locus 2 (rs3842570). Alelo deleção com 132 pares de bases (pb); alelo inserção com 164 pb e alelo com a inserção repetida uma vez (196 pb)_____ **46**

Lista de tabelas

Tabela 1: Localização cromossômica, sequência e tamanho (em pb) da inserção dos 20 lócus selecionados listados segundo o número de rs de identificação no dbSNP (NCBI)_____ **27**

Tabela 2: Sequências do trio de iniciadores e temperatura teórica de pareamento dos pares de iniciadores flanqueadores para cada lócus_____ **28**

Tabela 3: Quantidade de iniciador, quantidade de *Taq* (DNA polimerase), temperatura de pareamento dos iniciadores flanqueadores e tamanho do produto amplificado em pares de bases para cada lócus amplificado. Concentrações descritas no Protocolo 1 _____ **30**

Tabela 4: Quantidade de iniciador, quantidade de *Taq* (DNA polimerase), temperatura prática de pareamento dos iniciadores flanqueadores e tamanho do produto amplificado em pares de bases para os lócus amplificados em duplex _____ **31**

Tabela 5: Quantidade de amostras não amplificadas (NA) em cada lócus_____ **39**

Tabela 6: Quantidade de lócus não amplificados (NA) por indivíduo. Estão listados apenas aqueles que apresentaram pelo menos 1 NA_____ **39**

Tabela 7: Frequências genotípicas, frequência do alelo inserção, heterozigose observada, heterozigose esperada e valor de p no Equilíbrio de Hardy-Weinberg para os 16 lócus polimórficos (n = 100)_____ **40**

Tabela 8: Conteúdo informativo do polimorfismo (PIC), Probabilidade de Identidade (PI), Poder de discriminação (PD) e Poder de exclusão (PE)_____ **41**

Tabela 9: Frequência do alelo inserção em diferentes populações_____ **46**

Lista de abreviaturas e siglas

°C	graus Celsius
(NH₄)₂SO₄	Sulfato de amônio
a	número de alelos
AFR	População africana
AMR	População americana
ASN	População asiática
CEP	Comitê de Ética em Pesquisa
cffDNA	DNA fetal livre de célula
cm	Centímetros
d	alelo deleção
dbSNP	do inglês, <i>data base SNP</i>
dd	genótipo homozigoto deleção/deleção
DNA	Ácido Desoxiribonucleico
dNTP	Desoxirribonucleotídeo trifosfato
EDTA	Ácido etilenodiamino tetra-acético
EHW	Equilíbrio de Hardy-Weinberg
EUR	População europeia
g	Grama
GP	Geração Parental
HCFMRP	Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto
HCl	Ácido Clorídrico
HCRP	Hospital das Clínicas de Ribeirão Preto
He	Heterozigose esperada
Ho	Heterozigose observada
i	alelo inserção
id	genótipo heterozigoto inserção/deleção
ii	genótipo homozigoto inserção/inserção
InDels	Polimorfismo Inserção/Deleção
k	número de loci
kb	quilobases
KCl	Cloreto de Potássio
L	Litro
mA	miliampère
mg	Miligrama
MgCl₂	Cloreto de Magnésio
mL	Mililitro
n	Número de indivíduos
NA	Não amplificado
NaOH	Hidróxido de sódio
NCBI	do inglês, <i>National Center for Biotechnology Information</i>
n_{id}	Número observado de heterozigotos

n_{ii}	Número observado de homozigotos para o alelo inserção
NIPD	do inglês, <i>Non-invasive prenatal diagnosis</i>
NIPT	do inglês, <i>Non-invasive prenatal testing</i>
p ou p_i	Frequência do alelo inserção
PAGE	do inglês, <i>Polyacrilamide gel electrophoresis</i>
pb	par de bases
PCR	do inglês, <i>Polymerase Chain Reaction</i>
PD	Poder de discriminação
PDC	Poder de discriminação cumulativo
PE	Poder de exclusão
PEC	Poder de exclusão cumulativo
pH	Potencial de hidrogênio
PI	Probabilidade de identidade
PIC	do inglês <i>Polymorphism Information Content</i>
P_{ii}	Frequência do genótipo inserção/inserção
PM	Peso molecular
Rh	Fator Rhesus de antígenos eritrocitários
RhD	Antígeno D do grupo Rh
rs	do inglês, <i>Reference SNP</i>
STRs	do inglês, <i>Short Tandem Repeat</i>
SUS	Sistema Único de Saúde
Taq	do inglês, <i>Thermus aquaticus</i>
TBE	Tris-Borato-EDTA
TEMED	N,N,N',N'-tetrametiletilenodiamina
U	Unidade
USP	Universidade de São Paulo
V	Volts
μL	Microlitro
μM	Micromolar

Introdução

POLIMORFISMOS GENÉTICOS

O genoma humano, bem como o de outros organismos, possui variações genéticas que caracterizam cada um dos indivíduos. O avanço científico permitiu a ampliação das análises dessas variações e com o conhecimento das várias formas da sequência de nucleotídeos concluiu-se que os dados proteicos e morfológicos estudados consistiam apenas em uma subestimativa da variabilidade genética real. O Projeto Genoma Humano identificou 3 bilhões de pares de bases, possibilitando que as variações na molécula de DNA passassem a ser cada vez mais investigadas no sentido de entender as bases da diversidade genética entre os indivíduos.

As variações na sequência de DNA são conhecidas como polimorfismos genéticos, ou seja, é a ocorrência simultânea de duas ou mais variantes (alelos) em um locus. Entretanto, conceitualmente, um locus apenas é considerado polimórfico se seu alelo mais comum apresentar frequência na população igual ou inferior a 99% (Cavalli-Sforza *et al.*, 1994). Os polimorfismos moleculares são resultantes de diferenças na sequência de DNA provocadas por variações nucleotídicas ocorridas ao longo da evolução, que garantem o alto grau de diversidade e individualidade genética (Thompson e Thompson, 2002). A análise da distribuição de frequências dos alelos na população possibilita o acesso à constituição genética da mesma, sendo essa utilizada para estimar o nível de variação genética entre populações e entre indivíduos, entender mecanismos evolutivos, analisar filogenias, utilizar as diferenças genéticas para identificação individual, rastreamento de doenças, etc (Hartl e Clark, 2010).

Com o advento das tecnologias de análise do DNA, os polimorfismos puderam ser aplicados como marcadores moleculares na identificação de indivíduos, resultando no aumento da capacidade para realizar testes de identificação humana (Butler, 2005). Esses marcadores, em geral, podem ser classificados como VNTR (Variações no número de repetições em tandem), que incluem os mini e microssatélites ou STRs (*Short Tandem Repeat*), SNP (*Single Nucleotide Polymorphism*) que são as substituições de um único nucleotídeo e InDels, polimorfismos caracterizados pela inserção ou deleção de um ou mais nucleotídeos. Mais recentemente, foram iniciados estudos sistemáticos de outros tipos de variantes, os CNVs (*Copy Number Variation*), grandes segmentos genômicos com variações no número de cópias em comparação com o genoma de referência (Alkan *et al.*, 2001), que começaram a ser aplicados para estudos de frequência, distribuição populacional, padrões de desequilíbrio de ligação e outras.

MARCADORES MOLECULARES NA IDENTIFICAÇÃO HUMANA

A identificação humana sempre foi investigada através dos polimorfismos, passando dos grupos sanguíneos aos polimorfismos proteicos. Entretanto, foi a descoberta do “DNA *fingerprint*” (Jeffreys *et al.*, 1985) que alavancou a eficiência das técnicas aplicadas possibilitando que os marcadores moleculares pudessem ser utilizados na genética forense. Um bom marcador molecular é suficientemente informativo quando apresenta elevado poder de discriminação combinado à análise do menor número de loci possível. Essas características, associadas à possibilidade de se aplicar metodologias simples como a PCR (*Polymerase Chain Reaction*) convencional, fazem dos microssatélites ou STRs os marcadores mais estudados e utilizados atualmente (Butler, 2005).

STRs

Os STRs ou microssatélites são variações repetidas em tandem (Pena e Chakraborty, 1994, Pena e Jeffreys, 1993), onde sequências curtas de 2 a 9 pares de bases se repetem em um ou mais lugares do genoma (Hartl e Clark, 2010). Possui taxa de evolução rápida (Kayser *et al.*, 2000), que permite uma maior quantidade de alternativas alélicas tornando-os úteis tanto para estudos de eventos populacionais recentes, quanto para identificação. O uso dos STRs em vários estudos genéticos, sobretudo na genética forense, deve-se a um conjunto de características como o alto grau de polimorfismo, dispersão uniforme, altos níveis de reprodutibilidade e codominância.

Apesar de o uso universal forense dos STRs ter permitido o estabelecimento de bases de dados em vários países, metodologias simplificadas e avaliação de parâmetros estatísticos relevantes, o extenso tamanho dos produtos amplificados de alguns deles e as limitações com relação à quantidade e qualidade da amostra analisada podem representar desvantagens para sua aplicação (Manta *et al.*, 2012). Essa metodologia ainda tem uso prioritário na genética forense, entretanto, em certas situações, a falha na obtenção do perfil genético exige o uso de outras ferramentas que a complementem ou substitua.

SNPs

Apenas mais recentemente, a análise genética se estendeu aos marcadores bialélicos, que, devido a sua quantidade de alelos limitada, necessitam da análise de um número maior de loci em comparação com os de STRs, para que a confiabilidade do perfil genético seja aceitável (Chakraborty *et al.*, 1999).

Os SNPs bialélicos são os mais comuns (Brookes, 1999), sendo encontrados tanto em regiões codificantes quanto não-codificantes, a cada 1000 a 1200 pares de bases aproximadamente (Venter *et al.*, 2001). Podem ser úteis como marcadores moleculares, pois tem baixa taxa de mutação, o que é interessante para teste de paternidade. Em geral, tem produtos de PCR curtos, desejáveis para análise bem sucedida de amostras degradadas e são apropriados para análises em novas tecnologias com alto rendimento que facilitam a automatização e reduzem os erros (Sobrinho *et al.*, 2005). Contudo, a tipagem de SNPs é relativamente difícil, envolvendo protocolos de genotipagem complexos, que exigem várias etapas e depende de tecnologias mais caras (Huang *et al.*, 2014). Essa desvantagem permitiu que os InDels passassem a ser mais investigados para o uso na identificação humana.

InDels

InDels são polimorfismos gerados pela inserção ou deleção de um ou mais nucleotídeos (Francez *et al.*, 2011). Estes são classificados como marcadores de evolução lenta, constituídos por fragmentos de DNA que podem ter sido acrescentados ou eliminados de uma sequência do genoma (Bastos-Rodrigues *et al.*, 2006, Weber *et al.*, 2002). Podem ser encontrados numa forma multialélica, na qual há números variáveis de um segmento de DNA que podem ocupar determinado locus, entretanto, aproximadamente, metade de todos os InDels é referida como um polimorfismo bialélico, composto pela presença (alelo longo) ou ausência (alelo curto) do segmento (Thompson e Thompson, 2002). Nesse caso, são possíveis três genótipos: homozigose do alelo inserção (*ii*) ou do alelo deleção (*dd*) e heterozigose (*id*).

InDels bialélicos são os polimorfismos mais abundantes depois dos SNPs (Huang *et al.*, 2014), compreendendo até cerca de 8% de todos os polimorfismos humanos (Weber *et al.*, 2002). Cerca de 400 mil destes polimorfismos tem extensão que não ultrapassa 10000 pb, aproximadamente 1/3 foi descrito em genes já conhecidos e agrupam-se em cinco classes: (a) inserções/deleções de uma base; (b) expansão monomérica de uma base; (c) expansões múltiplas de unidades repetitivas de 2-15 pb; (d) inserção de transposons e (e) InDels contendo sequências randômicas, cujas bases não seguem uma ordem ou padrão (Mills *et al.*, 2006).

O poder de discriminação obtido pelo uso dos 13 a 15 locus STRs rotineiramente utilizados na identificação humana pode ser alcançado pelo emprego de aproximadamente 60 pequenos InDels. Apesar da grande quantidade, o conjunto de 60 locus analisados, organizados em dois grupos de multiplex, alcança potencial considerável na identificação de pessoas (Fondevila *et al.*, 2012).

O crescimento do interesse pelo trabalho com esses polimorfismos no campo forense se deve a algumas propriedades compartilhadas com os SNPs como: ampla distribuição por todo o genoma (Weber *et al.*, 2002), incluindo os cromossomos sexuais; podem apresentar diferenças significativas nas frequências alélicas entre populações distintas geograficamente; podem ser genotipados em pequenos produtos de amplificação, aumentando o sucesso nas análises de DNA degradado (Pereira *et al.*, 2009, Ribeiro-Rodrigues *et al.*, 2009). Além disso, e diferentemente dos SNPs, são detectados por PCR simples e a análise dos produtos amplificados por eletroforese em gel de poliacrilamida (PAGE), seguida de coloração por prata, aumenta a praticidade da técnica para uso na rotina.

GENOTIPAGEM FETAL NÃO INVASIVA

DNA fetal livre de células

A primeira descrição de DNA fetal na circulação materna foi em células nucleadas de trofoblastos associadas à eclâmpsia (Attwood e Park, 1961). O primeiro cariótipo fetal foi obtido através da cultura de células do líquido amniótico em 1966 (Steele e Breg, 1966). Desde então, a coleta invasiva de tecido placentário por punção (transabdominal ou transcervical) ou de células do líquido amniótico por amniocentese (Figura 1) passou a ser realizada para fins de diagnóstico pré-natal de anomalias congênitas. Entretanto, ambas as técnicas geram riscos para a gestação, podendo alcançar uma taxa de até 1% de aborto (Evans e Andriole, 2008).

Figura 1. Métodos de coleta de material genético fetal de forma invasiva (A e B) e não invasiva (C). Adaptado de HUI; BIANCHI, 2013.

O avanço das tecnologias na área, visando redução de riscos associados, permitiu que técnicas mais seguras surgissem e, em 1997, foi descrita pela primeira vez a presença de DNA fetal livre de célula (*cell-free fetal DNA* – cffDNA) circulante no plasma sanguíneo materno (Lo *et al.*, 1997). Sequências do cromossomo Y foram detectadas em 80% das amostras de gestantes de fetos masculinos.

Existem inúmeras hipóteses de origem do cffDNA na circulação materna. Alguns autores apontam os trofoblastos como principal fonte (Alberry *et al.*, 2007). A migração dessas células é possível, pois as circulações envolvidas (uteroplacentária e fetoplacentária), apesar de separadas pela membrana placentária, se tornam mais próximas após a 20ª semana (Gude *et al.*, 2004). A estreita justaposição entre elas possibilita a migração de células do feto para a mãe. Pequenos fragmentos de DNA fetal, com origem em diferentes tecidos, também podem passar do líquido amniótico para a circulação materna por difusão transplacentária (Vodicka *et al.*, 2011). Em menor proporção, o DNA fetal pode ser originado também pela desintegração de células da placenta (Wataganara e Bianchi, 2004) ou de células hematopoiéticas fetais (Bianchi, 2004).

O DNA nuclear das células fetais é liberado por apoptose, logo após o nascimento (Lo *et al.*, 1998), causada por reação do sistema imune materno (Pertl *et al.*, 2000). Isso explica a maior concentração relativa de cffDNA em preparações com plasma sanguíneo materno em relação às preparações com o sangue total (Bianchi, 2004, Bischoff *et al.*, 2005). A apoptose ocorre de maneira rápida, levando de duas a três horas (Kolialexi *et al.*, 2006), e o microquimerismo provocado por células fetais persistentes está associado a doenças autoimunes, mesmo havendo alguns benefícios para a mãe, devido às habilidades de diferenciação e proliferação das células fetais (Bianchi, 2004).

No plasma materno, são encontradas moléculas de DNA livres circulantes tanto da mãe quanto do feto. O genoma fetal está presente por completo e em uma proporção relativa constante ao genoma materno (Lo *et al.*, 2010). Os fragmentos maternos são maiores, acima de 1 kilobases, e o cffDNA está fragmentado em pequenas moléculas de 100 a 700 pares de bases (Koide *et al.*, 2005, Li *et al.*, 2004, Zhong *et al.*, 2000).

Em geral, a concentração de cffDNA aumenta com o tempo de gestação, podendo alcançar até 12% do DNA total (Lo *et al.*, 1998) e com um aumento agudo nas últimas oito semanas antes do parto (Bianchi, 2000). Os fragmentos maiores – entre 500 e 700 pb – aumentam sua concentração ao longo da gestação, enquanto que a concentração de fragmentos menores não apresenta mudança significativa (Vodicka *et al.*, 2010).

Diagnóstico pré-natal não invasivo

Com a descoberta do DNA fetal na circulação materna, foram desenvolvidas novas metodologias para o diagnóstico pré-natal de forma não invasiva (Non-invasive prenatal diagnosis or testing – NIPD ou NIPT), não só por meio da citogenética clássica, mas também com a análise de marcadores moleculares de forma qualitativa e quantitativa. O primeiro procedimento de NIPD relatado na literatura foi a determinação do sexo através da amplificação por PCR de um fragmento de sequência repetitiva do cromossomo Y, específica do feto, realizada com amostras de sangue periférico de gestantes com 9-14 semanas (Lo *et al.*, 1989).

Atualmente, a análise do cfDNA está disponível como método para diagnosticar doenças ligadas ao X, como hemofilia e síndrome do X frágil, determinação do sexo e do fator Rh, a fim de evitar a doença hemolítica do recém nascido (Chiu e Lo, 2002). Em gestações de fetos acometidos de certas aneuploidias cromossômicas (Lo, 2009) ou com ocorrência de pré-eclâmpsia (Lo *et al.*, 1999, Yu *et al.*, 2013), a quantidade total de cfDNA observada no plasma materno é elevada. A sua concentração parece estar relacionada com a severidade da eclâmpsia e por isso, a sua quantificação pode ser usada no prognóstico do quadro clínico (Bischoff *et al.*, 2005, Zhong *et al.*, 2001).

No Brasil, o primeiro procedimento com cfDNA foi descrito em 2003 (Levi *et al.*, 2003) para determinação do sexo fetal por detecção do cromossomo Y. Em 2011, foi demonstrada a aplicabilidade da técnica para detecção precoce do fator RhD fetal, com 100% de acurácia, em uma amostra da população brasileira com alto percentual de mistura étnica (Amaral *et al.*, 2011). Embora seja uma importante alternativa aos métodos invasivos, apenas alguns laboratórios privados disponibilizam exames de diagnóstico pré-natal não invasivo. Seu uso ainda é inovador, limitado pelos custos e não há implementação no Sistema Único de Saúde (SUS) (Romão *et al.*, 2012).

Identificação humana pré-natal não invasiva

Alguns métodos de NIPD, baseados na genotipagem através de marcadores moleculares, visam isolar o material genético fetal do sangue materno durante a gravidez. Existem técnicas de genotipagem fetal já estabelecidas em investigações de sequências exclusivas fetais, entretanto, a sua aplicação para a identificação humana em testes de paternidade ainda é pouco conhecida.

A primeira análise de cfDNA para investigação de paternidade foi por amplificação de alelos STR fetais em amostra de plasma materno, relatada em 2009 (Wagner *et al.*, 2009). Embora os STRs sejam marcadores adequados para identificação humana, seu uso na genotipagem fetal não invasiva é limitado pela quantidade e qualidade do cfDNA. A integridade do material genético fetal é reduzida, sendo a maioria altamente

fragmentada (Vodicka *et al.*, 2011). Além disso, visto que a proporção de DNA fetal é baixa com relação ao DNA materno (1:10⁶) (Alberry *et al.*, 2009), a concorrência pelos iniciadores utilizados na PCR faz com que haja amplificação preferencial do DNA materno.

Em trabalho recente ainda não publicado, Santos (2014) utilizou três loci de marcadores InDels para padronização da metodologia de genotipagem fetal não invasiva com o uso de iniciadores inserção-específicos. O DNA fetal foi detectado em amostras de sangue materno por PCR quantitativa em tempo real e o fenótipo fetal confirmado por amplificação de DNA extraído diretamente de amostras do feto.

Entretanto, para que se possa determinar um perfil genético com o uso de InDels, é necessário aumentar o número de loci investigados. Com um painel de InDels mais robusto, essa nova metodologia poderá apresentar o poder de discriminação satisfatório para uso na genotipagem fetal não invasiva.

HIPÓTESE, JUSTIFICATIVA E OBJETIVO

A metodologia de genotipagem fetal não invasiva com o uso de iniciadores inserção-específicos de marcadores InDels foi estabelecida, inicialmente, com o uso de três loci e validada por PCR quantitativa (Santos, 2014).

Sabendo-se que, para a determinação de um perfil genético com alto poder de discriminação, o número de loci InDels investigado deve ser aproximadamente 60, como sugerido por Fondevilla *et al.* (2012), faz-se necessário descrever as condições para detecção laboratorial e frequências alélicas de uma maior quantidade de loci deste tipo de marcador.

Propõe-se, neste projeto, a seleção de um conjunto de loci do tipo InDel que atenda basicamente aos critérios seguintes:

1. Que seja possível desenhar um par de iniciadores flanqueador à inserção;
2. Que seja possível desenhar iniciador inserção-específico, isto é, um dos iniciadores é complementar à inserção e o outro seria um dos dois iniciadores flanqueadores.
3. Tenha elevada heterozigose (grau de polimorfismo).
4. Não haja desequilíbrio de ligação (DL) entre eles.

Os dados aqui gerados não só servirão para compor o conjunto de loci necessários no futuro, mas também para os procedimentos de validação, necessariamente compostos por exames individualizados por loci. Além disso, as informações aqui obtidas são necessárias para indicar loci que eventualmente serão incluídos em plataformas automatizadas de análise genômica.

Para que o loci seja considerado útil para o diagnóstico fetal não invasivo, pelos nossos critérios, além de ser passível de amplificação a partir do trio de iniciadores, é necessário que seja polimórfico e altamente informativo, isto é, apresente elevada heterozigose.

Pelo exposto acima, o objetivo principal do presente trabalho consiste na determinação das frequências alélicas de loci de polimorfismos do tipo InDel em um grupo representativo da população urbana brasileira, para a seleção daqueles que possam ser usados na genotipagem fetal não invasiva.

Visando compor um painel de locus InDel útil na determinação de perfil genético fetal de forma não invasiva, constitui objetivo geral do presente trabalho selecionar um conjunto de 20 locus do tipo InDel, com alta heterozigose, um em cada cromossomo (garantia de ausência de desequilíbrio de ligação entre eles), potencialmente amplificados por iniciadores flanqueadores à inserção e por iniciadores, um deles flanqueador e outro complementar à inserção.

OBJETIVOS OPERACIONAIS

1. Selecionar um conjunto de locus InDel adequados pela procura em bancos de dados e na literatura;
2. Padronizar as condições de PCR dos iniciadores e, eventualmente, estabelecer as condições necessárias para a multiplexação.
3. Caracterizar o perfil genético de uma amostra da população urbana brasileira com o uso dos iniciadores complementares à região flanqueadora;
4. Determinar a frequência alélica dos locus para estimar parâmetros populacionais e de interesse forense;
5. Selecionar, a partir dos dados populacionais obtidos, os locus adequados para uso posterior na genotipagem fetal não invasiva.

Material e Métodos

DESENHO DO ESTUDO

Um conjunto de 20 loci InDel, um em cada cromossomo, foi selecionado a partir do tamanho da inserção, de maneira que fosse possível desenhar dois iniciadores externos, flanqueando a inserção, e um interno, pareando com ela. O desenho de iniciadores priorizou produtos amplificados pequenos, mas de tamanhos variados para que fosse possível a tentativa de multiplexação. As condições de PCR dos iniciadores externos foram padronizadas e os fenótipos de 100 indivíduos de uma amostra da população urbana brasileira foram analisados com o objetivo de caracterizar os alelos e estimar suas frequências.

AMOSTRAS BIOLÓGICAS

As amostras de DNA são provenientes de 100 indivíduos não aparentados do Estado de São Paulo, que procuraram o Serviço de Investigação de Paternidade do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto – USP (HCFMRP), constituindo o grupo parental (GP) de parte do banco de amostras “Populações Urbanas”, alocado no Laboratório de Genética Bioquímica do Departamento de Genética.

COLETA E CONSERVAÇÃO

À época da coleta, as amostras de sangue ou de *swab* bucal foram armazenadas em tubos VacuTainer (com anticoagulante EDTA para as amostras de sangue) devidamente identificados. A fração celular, obtida por centrifugação e posterior lavagem em solução salina, permaneceu estocada em Tampão de Estocagem de Eritrócitos e congelada a -20°C até o momento da extração de DNA.

O DNA das amostras de sangue coletadas dos indivíduos foi anteriormente extraído com adaptação à metodologia proposta por Higuchi (Higuchi, 1989) pelos funcionários do Serviço de Investigação de Paternidade do HCFMRP e encontra-se estocado e mantido a -20°C no banco de amostras “Populações Urbanas”.

ASPECTOS ÉTICOS

As amostras oriundas do Serviço de Investigação de Paternidade do HCFMRP, vinculado ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto – USP, integram o banco de amostras intitulado “Populações Urbanas”. Além das amostras de São Paulo que serão utilizadas, compõem o mesmo banco amostras de Pernambuco e do Rio Grande do Sul.

A coleta do material biológico e as informações geradas a partir dele estão em conformidade com o Comitê de Ética em Pesquisa (CEP) do HCFMRP, com aprovação da criação do Banco de Amostras e do seu uso neste projeto registrada nos processos HCRP nº8079/2004 e HCRP nº11015/2014, respectivamente (Apêndice I).

MARCADORES GENÉTICOS ANALISADOS

As primeiras buscas foram feitas no banco de dados da Marshfield Clinic (<http://www3.marshfieldclinic.org/mgs/default.aspx>), onde é possível buscar sequências a partir do tipo de marcador e do tamanho da inserção dentre outros critérios.

Uma segunda busca foi feita a partir das sequências depositadas no banco de dados de SNPs (dbSNP) do National Center for Biotechnology Information – NCBI (<http://www.ncbi.nlm.nih.gov/SNP/>). Para isso, foi salvo um arquivo contendo todas as InDels registradas, sendo um arquivo para cada cromossomo. O banco de dados disponibiliza as informações de todos os loci em formato texto (FASTA) e a sua conversão em tabela do programa Microsoft Excel permitiu que fossem filtradas apenas as informações necessárias para a seleção das sequências. Os dados foram organizados com o número de identificação (rs) e a sequência das inserções. Visto que muitos dos loci não eram bialélicos, aqueles que possuíam, além da deleção, mais de uma inserção foram excluídos, bem como aqueles com inserção menor do que 17 pares de bases e mais de 4 repetições na sequência. A partir de então, as sequências foram escolhidas manualmente visando o desenho dos iniciadores.

A análise foi feita em um conjunto de 20 loci de marcadores genéticos do tipo inserção/deleção localizados um em cada cromossomo (tabela 1). A seleção dos loci teve como critério principal o tamanho da inserção, que deveria ser igual ou superior a 17 pares de bases para que coubesse um iniciador interno (inserção-específico).

Tabela 1. Localização cromossômica, sequência e tamanho (em pb) da inserção dos 20 loci selecionados listados segundo o número de rs de identificação no dbSNP (NCBI).

Locus*	Rs	Local. crom.	Tam (pb)	Sequência da inserção (5'-3')
L1	rs6143647	1q42.11	25	AGCTGCGGAGCTGGGGTTACTTCCT
L2	rs3842570	2q37.3	32	CGGGAGGAGGGTGATGATTCTGTCCCAGGAGC
L3	rs67696659	3q22.2	23	TCCTATAAGGCCGTGCAATGACC
L4	rs33926123	4q13.3	24	ACCATGGATGTTGGTGCCTCCTAC
L5	rs34156041	5p15.31	23	AGTCATCTGGAAGCTGCTACTGC
L7	rs74315104	7q21.3	56	TCATTACCTCAAGGTAAATGAGGCAAAGTTCTACAA TATCAGTTTTGTCCCTTCGA
L8	rs1305049	8q24.21	29	TCCCTCTTCTCTAGAACTCTGCCCTAAAA
L9	rs118203674	9q34.13	23	GCAGCGTGACACTATGGTAACCA
L10	rs34656661	10p14	29	AAGATAGAACATACTCCACTCTGTATAGC
L11	rs80356802	11q13.3	84	GTAAGTCCGGATTATCCAACCATTTGTTACTTACCTG GGATGGGAAACACATTTGAAGTCTTACAGATGAATT TTCCAACCGTAG
L12	rs267607475	12q13.13	27	ATCGCCACCTACCGCAAGCTGCTGGAG
L13	rs67616458	13q14.11	27	TCTAAGAACAGTGGGACCTTTGGAAGG
L15	rs11275778	15q26.1	27	TTTACCCAGCCCTGAGAAGGCTGGGTA
L16	rs138481647	16q23.1	28	GTCTGCTTGGCACAAAGCCTCGCACGAG
L17	rs59137132	17q24.3	28	CTGGCTTCCTTCCCCAGTCCCCTGTGC
L18	rs140462903	18p11.21	28	TTACTCTGACATAGAGGAGAAGCAGCAG
L19	rs148112183	19p13.3	26	CCCAGCGTCTCCCTGTCCGCACTGAT
L20	rs71311113	20p13	51	GCAGATCAGAGGCAGGATGTGGCTACCGAGCAGG CAAAGATGATGGGAAAT
L21	rs71330398	21q22.2	25	AGAGGAAAATCAAGAGAAGTGGCAG
L22	rs67519581	22q12.2	32	GCGTGGCCTGATAACAGGGTAAGAGTCCAGAG

(*) Com o objetivo de facilitar a referência aos loci, adotamos o nome composto por “L” seguido pelo número do cromossomo em que se localiza.

DESENHO DE INICIADORES

Um trio de iniciadores, sendo dois complementares às regiões flangeadoras (externos) e um complementar à inserção (interno), foi desenhado para cada sequência de modo que pudessem ser combinados dois a dois para padronização das condições de PCR (Tabela 2).

Tabela 2. Sequências do trio de iniciadores e temperatura teórica de pareamento dos pares de iniciadores flanqueadores para cada locus.

Locus	Iniciadores (5'-3')			°C
	Direto	Interno*	Inverso	
L1	AACCTCTCCCTTGGCGTCAG	GGAGCTGGGGTTACTTCCT	AAGACCAACTCCGCTACCC	55
L2	TTGGTTCTCTTCAGCGTGGA	GAATCATCACCTCCTCCCG*	CATGGTGGCCTCATGAACC	55
L3	ACAAGCGGCTATGGAAGTGT	TATAAGGCCCGTGCAATGACC	CTCTGGTCAACCCCACTGAT	56
L4	CCCACCACCTTCAATTCCCT	CATGGATGTTGGTGCCTCC	GTATGCGAAGTAGTGCCTGC	57
L5	GGAAGCAAGTACCAGATGCC	AGTCATCTGGAAGCTGCTACT	ACTCTCACACTATACCTTCACTCA	60
L7	CAGACATGCCCTTCCTGTTA	CAAGGTAAATGAGGCAAAGTTCTAC	TGTTGAAACTCTCTTGACAGACA	56
L8	TTTAGACCAAGGCTCAAGG	CTTTCCTAGAACTCTGCCCT	CTGGAGCACATTCTGAACC	52
L9	TGTGGAAGGTTAGTCTGCAGA	GCAGCGTGACACTATGGTAA	TCCTGGCTCTGGTTGTAGAA	56
L10	TGACACTTTGACTCAGCGA	GAACATACTCCACTCTGTATAGC	GCTTTTGATGCTAATTCCCATG	55
L11	GGCGACATCAATCCGAACAT	ATTGTTACTTACCTGGGATGGG	GCAAACAACTTTCAGCCTCAG	56
L12	GATGAATGTCAAGCTGGCCC	ATCGCCACCTACCGCAA	GTCGCGTCAGTTACCTACCT	57
L13	TGGAGTCAATGCTTGGGTGA	AGAACAGTGGGACCTTTGGA	TGATGAGAAGGAGGAGGAAGG	57
L15	GCAAATCTCACCCGTGTCTC	TTTACCCAGCCCTGAGAAGG	GATCTTAGGGCAGCCTGACA	57
L16	TCAGCAAGACGAGAGCAGAA	CACAAAGCCTCGCACGA	CCCTCTGCCTCGGAACCTTTA	56
L17	AAACTCTTTCCTCCTCCCC	TTCTTCCCCAGTCCCACT	CATAGTGGGAGAGCTGGGAG	58
L18	GACAACTCAGGATGTGCAGC	TGACATAGAGGAGAAGCAGCA	ACAGATGAGGTGGGTGACTG	57
L19	CCTCTCTGCTGTGTGGACTT	TCTCCCTGTCCGCACTGAT	CTGGACCTCACTCTCCTCAC	58
L20	TGCCACTCCTCTTAAGCCACT	CTACCGAGCAGGCAAAGATG	TTCCAAGGAAGGTAGCCACA	57
L21	CAGGGGAAAGTGTGCATCTG	GAGGAAAATCAAGAGAAGTGGCA	CTCAAGTCCCACCTTCCCAT	57
L22	GTGAGTGAGGGGACAGTCGT	TGGCCTGATAACAGGGTAAGA	TGCAGGCCTGACATAGACC	57

*Todos os iniciadores internos foram desenhados no sentido direto pareando com o iniciador flanqueador inverso, exceto o interno do locus L2 que é no sentido inverso e faz par com o flanqueador direto.

A amplificação com o par de iniciadores externos é capaz de identificar os três fenótipos possíveis para um InDel bialélico. Por sua vez, o uso do iniciador interno possibilita a detecção apenas da inserção. Esse pode ser direto (complementar à fita 3'>5') ou reverso (complementar à fita 5'>3'), desde que o seu par referente à região flanqueadora esteja em sentido inverso e contrário ao seu (Figura 2).

Embora não tenha sido o objetivo deste trabalho, o desenho do iniciador interno foi um dos critérios para a escolha dos loci e para o desenho dos iniciadores, uma vez que possibilita o uso futuro desses loci em estudos em que seja necessária a amplificação a partir da inserção.

Figura 2. Esquema de posicionamento par a par do trio de iniciadores dos loci InDels. (A) Os dois iniciadores externos, nas regiões flangeadoras em ambos os sentidos. (B) Um iniciador em sentido direto interno, complementar à inserção e um em sentido inverso externo, complementar à região flangeadora. (C) Um iniciador em sentido direto externo, complementar à região flangeadora e um inverso interno, complementar à inserção.

Dois programas foram usados: Primer3 (<http://bioinfo.ut.ee/primer3-0.4.0/>) e GeneRunner (<http://www.generunner.net/>). Em ambos, as condições foram ajustadas para aproximar as temperaturas de melting dos trios de iniciadores, analisando-os par a par e evitando aqueles com formação de estruturas secundárias como dímeros e grampos. A porcentagem de bases C e G ficou em torno de 50%.

Para avaliação da especificidade das sequências e dos iniciadores, realizou-se PCR *in silico* e BLAT, disponível online no site UCSC Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgPcr?command=start>).

ANÁLISE LABORATORIAL

Em geral, as técnicas laboratoriais aqui utilizadas são bastante conhecidas e descritas na literatura, a partir do que protocolos foram elaborados em nosso laboratório (Laboratório de Genética Bioquímica do Departamento de Genética da Faculdade de Medicina de Ribeirão Preto, Univ. de São Paulo) para uso em todos os projetos aqui executados.

Por esta razão, tais protocolos são apresentados em no Apêndice I e, no texto, são descritos apenas as modificações específicas ao presente trabalho.

Reação em cadeia da polimerase (PCR)

Padronização dos novos iniciadores

Para padronização das condições de PCR, os iniciadores sintetizados (SIGMA-ALDRICH) foram diluídos em água destilada ultrafiltrada para uma solução estoque a 50 μ M. Os pares de flaqueadores foram empregados em PCR individual como descrito no Protocolo 1 (Apêndice II), utilizando inicialmente as temperaturas teóricas de pareamento para cada locus (vide Tabela 2).

Após a padronização, as quantidades de iniciador, de *Taq*-polimerase e as temperaturas para amplificação individual, variaram para alguns locus e estão descritas nas Tabelas 3 (para os locus amplificados individualmente) e 4 (para os locus amplificados em reação duplex), onde também consta o tamanho dos produtos amplificados por cada par de iniciador flaqueador amplificado individualmente.

Tabela 3. Quantidade de iniciador, quantidade de *Taq* (DNA polimerase), temperatura de pareamento dos iniciadores flaqueadores e tamanho do produto amplificado em pares de bases para cada locus amplificado. Concentrações descritas no Protocolo 1.

Locus	Quantid. de iniciador (μ L)	Quantid. de <i>Taq</i> (μ L)	°C	Tamanho do amplificado (pb)	
				in	del
L2	2	0,1	55	164	132
L4	0,7	0,1	55	160	136
L11	1,7	0,1	57	198	114
L15	2,5	0,1	59	159	132
L16	2	0,1	57	151	123
L17	2,5	0,3	59	197	169
L19	2	0,1	57	171	145
L20	1,5	0,1	57	169	118

O termociclador utilizado para as reações de amplificação do DNA foi o modelo GeneAmp® PCR System 9700 (Applied Biosystems).

A especificidade dos iniciadores verificada *in silico* foi confirmada pelo tamanho dos produtos amplificados por observação dos alelos pretendidos.

Sistemas Duplex

Na tentativa de minimizar o esforço (tempo e custo) laboratorial, procurou-se estabelecer as condições de PCR para análise conjunta de mais de um par de iniciadores em uma mesma reação. Para tal, os pares de iniciadores com características menos discrepantes foram incorporados de forma progressiva e com ajuste das quantidades de reagentes. Alguns deles foram agrupados em duplex com amplificação simultânea dos fragmentos referentes à inserção e à deleção de ambos os locus, com resultados claros e semelhantes aos das amplificações individuais. As condições de PCR duplex estão descritas na Tabela 4.

Tabela 4. Quantidade de iniciador, quantidade de *Taq* (DNA polimerase), temperatura prática de pareamento dos iniciadores flanqueadores e tamanho do produto amplificado em pares de bases para os locus amplificados em duplex.

Locus	Quantid. de iniciador (μL)	Quantid. de <i>Taq</i> (μL)	°C	Tamanho do amplificado (pb)			
				in	del	in	del
L1-3	1 e 1	0,1	55	182	157	115	92
L5-10	1,5 e 1,5	0,1	55	101	78	168	139
L7-12	2 e 1	0,1	55	209	153	90	63
L8-13	1 e 1,5	0,1	55	188	159	140	113
L9-18	1 e 1	0,1	55	139	116	97	69
L21-22	0,5 e 1	0,2	57	191	124	103	71

Análise do produto amplificado

O produto amplificado foi analisado por eletroforese em gel de poliacrilamida não-desnaturante descrito no Protocolo 2 (Apêndice II).

Modificações em relação ao protocolo:

- O cassete utilizado possui placas de tamanho 16 cm x 16 cm e pente para 30 poços;

- b) A eletroforese teve duração aproximada de 2 horas, com fonte de voltagem constante ajustada a 200 V e 400 mA.

Coloração com nitrato de prata e secagem do gel

Para observação do padrão de bandas, a coloração do gel foi feita com nitrato de prata seguindo o Protocolo 3 (Apêndice II).

A secagem do gel foi feita forrando-se uma placa de vidro com papel celofane culinário umedecido. Sobre a placa forrada, coloca-se o gel e, por cima, outra folha de papel celofane umedecido, de maneira que não haja bolhas de ar entre as camadas sobre o vidro. Em condições ambiente, o conjunto de celofane e gel ressecará completamente em um ou dois dias e o gel poderá ser identificado e conservado em coleção para registro dos resultados e para consultas posteriores.

Determinação fenotípica

O DNA dos 100 indivíduos foi genotipado para cada locus, a partir de seus programas de PCR (individual e duplex). Para aquelas amostras com resultados inconclusivos (bandas espúrias, alteradas ou claras) e aquelas que não apresentaram amplificação, o ensaio foi repetido até, no máximo, quatro vezes com mudanças no protocolo antes de ser assumida a ausência de resultados.

ANÁLISE ESTATÍSTICA

Distribuição de frequências

As frequências genotípicas e alélicas foram determinadas por contagem direta com uso do programa GenAlEx versão 6.0 (Peakall e Smouse, 2006) disponível em <http://www.anu.edu.au/BoZo/GenAlEx/>. Adiante, as equações para a contagem em organismos diploides.

$$p_i = \frac{2n_{ii} + n_{id}}{2n}$$

$$P_{ii} = \frac{n_{ii}}{n}$$

Em que:

p_i = frequência do alelo inserção;

P_{ii} = frequência do genótipo inserção/inserção;

n_{ii} = número observado de homozigotos para o alelo inserção;

n_{id} = número observado de heterozigotos

n = número de indivíduos

Aderência ao Equilíbrio de Hardy-Weinberg

No Equilíbrio de Hardy-Weinberg (EHW), espera-se que as frequências genotípicas possam ser estimadas a partir do binômio expandido:

$$(p + q)^2 = p^2 + 2pq + q^2$$

Em que:

p = frequência do alelo inserção;

q = frequência do alelo deleção;

p^2 = frequência esperada de homozigotos inserção/inserção;

$2pq$ = frequência esperada de heterozigotos;

q^2 = frequência esperada de homozigotos deleção/deleção.

Para verificar a aderência das frequências genotípicas observadas às proporções teóricas propostas para o EHW foi utilizado o programa Genepop versão 4.3 (Rousset, 2008) disponível em <http://kimura.univ-montp2.fr/~rousset/Genepop.htm>. Primeiramente, foi realizado o teste exato do EHW (Weir, 1996), que faz os cálculos para todas as configurações de amostras possíveis por Cadeia de Markov, considerando a hipótese nula de combinação aleatória dos gametas. Para os loci que apresentaram desvio do EHW, foi realizado um segundo teste, no qual a hipótese nula é a deficiência de heterozigotos.

Heterozigose

A heterozigose esperada H_e estima a diversidade gênica intra-populacional, definida como a probabilidade de dois alelos escolhidos aleatoriamente em uma população serem diferentes (Weir, 1996). A heterozigose observada é estimada por contagem direta. Ambas podem ser calculadas segundo as equações:

$$H_o = \frac{n_{id}}{n} \qquad H_e = 2pq$$

Em que:

n_{id} = número observado de heterozigotos;

n = número total de indivíduos analisados;

p = frequência do alelo inserção;

q = frequência do alelo deleção.

PARÂMETROS FORENSES

Conteúdo de informação do polimorfismo

O conteúdo de informação do polimorfismo (PIC do inglês *Polymorphism Information Content*), segundo Hildebrand (1992), é a probabilidade de que o genótipo de uma prole de pais que carregam um alelo raro permita a dedução do genótipo parental e pode ser calculada de acordo com a fórmula expandida (Botstein *et al.*, 1980):

$$PIC = 1 - \sum_{i=1}^a p_i^2 - \left(\sum_{i=1}^a p_i^2 \right)^2 + \sum_{i=1}^a p_i^4$$

Em que:

a = número de alelos;

p_i = frequência do i-ésimo alelo

Quando PIC é inferior a 0,25, o polimorfismo é considerado pouco informativo. Valores entre 0,25 e 0,5 indicam informatividade mediana e acima de 0,5 o polimorfismo pode ser considerado muito informativo.

Poder de discriminação e Probabilidade de identidade

O poder de discriminação (PD) fornece a probabilidade de dois indivíduos escolhidos ao acaso em uma população apresentarem genótipos diferentes (Desmarais *et al.*, 1998). A probabilidade de identidade (PI) é a probabilidade de dois indivíduos escolhidos ao acaso em uma população apresentarem genótipos idênticos para um locus (Sensabaugh, 1982). Sendo P_i a soma do quadrado das frequências genotípicas e $PD = 1 - PI$ (Butler, 2005):

$$PD = 1 - PI$$

$$PD = 1 - (p^4 + q^4 + (2pq)^2)$$

$$PDC = 1 - (1 - PD_1)(1 - PD_2)(1 - PD_3) \dots (1 - PD_k)$$

Em que:

p = frequência do alelo inserção;

q = frequência do alelo deleção.

PDC: Poder de discriminação cumulativo

k = número de locus

O PI foi calculado pelo programa GenAlEx 6.0, em seguida, foi aplicada a fórmula proposta para o PD e PDC.

Poder de exclusão

O poder de exclusão (PE) é a probabilidade de excluir um indivíduo que foi erroneamente acusado em uma investigação de paternidade (Desmarais *et al.*, 1998). O cálculo do poder para cada locus e do poder combinado (PEC) para o conjunto de todos os locus foi feito pelo programa GenAlEx 6.0, com a ferramenta Probabilidade de Exclusão P1, que segue as respectivas equações (Jamieson e Taylor, 1997):

$$PE = 1 - 2 \sum_{i=1}^a p_i^2 + \sum_{i=1}^a p_i^3 + 2 \sum_{i=1}^a p_i^4 - 3 \sum_{i=1}^a p_i^5 - 2 \left(\sum_{i=1}^a p_i^2 \right)^2 + 3 \sum_{i=1}^a p_i^2 \sum_{i=1}^a p_i^3$$

$$PEC = 1 - (1 - PE_1)(1 - PE_2)(1 - PE_3) \dots (1 - PE_k)$$

Em que:

a = número de alelos;

p_i = frequência do i-ésimo alelo;

k = número de locos.

Resultados

CLASSIFICAÇÃO FENOTÍPICA

No plano de uma primeira abordagem de triagem do conjunto de lócus, os 20 selecionados apresentaram resposta positiva às reações de amplificação e seus resultados em cada um dos 100 indivíduos analisados foram descritos e registrados no Apêndice III.

As amplificações por PCR geraram bandas nítidas representando alelos com migração eletroforética dentro da região esperada para o tamanho estimado de cada produto amplificado.

Na figura 3, estão representados os oito lócus amplificados individualmente. Observou-se o padrão de bandas: a banda maior e mais acima no gel representa o alelo inserção, enquanto que a banda menor e mais abaixo representa o alelo deleção. São possíveis três fenótipos: ii (homozigose para o alelo inserção); id (heterozigose) e dd (homozigose para o alelo deleção). Exceto para o lócus 11 (rs80356802), que apresentou monomorfismo com representação apenas da banda referente ao alelo inserção.

Figura 3. Fotos do padrão eletroforético visualizado para os lócus InDels amplificados individualmente. Os números indicam o peso molecular dos alelos representados pelas bandas. (*) No lócus L2, o alelo de maior peso molecular ainda não foi descrito e seu tamanho é aproximado.

Os outros 12 lócus foram amplificados aos pares em reação de PCR duplex (Figura 4). O mesmo padrão acima descrito pode aqui também ser observado, porém em duas regiões do gel. Nos duplex, os lócus 7 (rs74315104), 9 (rs118203674) e 12 (rs267607475), foram monomórficos, apresentando apenas a banda referente ao alelo inserção.

Figura 4. Fotos do padrão eletroforético de cada InDel amplificado em sistema duplex. Os números indicam o peso molecular dos alelos representados pelas bandas.

DNA não amplificado (NA)

Muitos indivíduos não apresentaram amplificação. A quantidade de indivíduos não amplificados (NA) por lócus está descrita nas tabelas 5.

Tabela 5. Quantidade de amostras não amplificadas (NA) em cada locus.

Locus	NA	Locus	NA	Locus	NA
L1	2	L9	0	L17	18
L2	11	L10	5	L18	1
L3	0	L11	5	L19	1
L4	10	L12	1	L20	3
L5	3	L13	3	L21	13
L7	17	L15	4	L22	6
L8	11	L16	2		

Dois locus tiveram mais de 15% dos indivíduos não amplificados. O locus 17 (rs59137132) teve menor eficiência (18 indivíduos não amplificados), representando 18% dos indivíduos amostrados. Resposta parecida foi observada no locus 7 (rs74315104) que, além de monomórfico, apresentou 17 indivíduos não amplificados.

Os NA também foram contabilizados em cada indivíduo. A quantidade de locus não amplificado por indivíduo está descrita na Tabela 6.

Tabela 6. Quantidade de locus não amplificados (NA) por indivíduo. Estão listados apenas aqueles que apresentaram pelo menos 1 NA.

Indivíduo	NA	Indivíduo	NA	Indivíduo	NA	Indivíduo	NA
1	14	23	4	37	2	74	1
3	3	24	5	38	1	79	1
6	3	25	1	39	1	80	1
7	3	26	1	41	1	86	1
10	7	28	4	42	1	89	5
12	2	29	1	43	2	91	1
13	3	30	2	44	3	95	3
18	2	31	1	45	8	96	1
19	1	32	2	46	1	97	1
20	1	33	1	47	2	98	1
21	1	34	9	62	1		
22	2	36	2	73	2		

Dos 100 indivíduos, 46 não amplificaram para pelo menos um locus. O indivíduo 1 foi o menos amplificado com 14 locus sem resultados, seguido pelos indivíduos 34 (9 locus), e 45 (8 locus).

ANÁLISE ESTATÍSTICA

Para a análise estatística foram excluídos os quatro loci que apresentaram monomorfismo, restando 16 loci bialélicos. Os indivíduos estão representados como GP.

As frequências genotípicas, frequência do alelo inserção, heterozigose observada e esperada e o valor de p do teste exato para o Equilíbrio de Hardy-Weinberg (EHW) estão descritos na tabela 7.

Tabela 7. Frequências genotípicas, frequência do alelo inserção, heterozigose observada, heterozigose esperada e valor de p no Equilíbrio de Hardy-Weinberg para os 16 loci polimórficos ($n = 100$).

Locus	GP							
	n	ii	id	dd	p_i	Ho	He	p
L1	98	21	53	24	0,485	0,541	0,500	0,544
L2	89	29	35	25	0,522	0,393	0,499	0,055
L3	100	39	49	12	0,635	0,490	0,464	0,669
L4	90	11	40	39	0,344	0,444	0,452	1,000
L5	97	79	18	0	0,907	0,186	0,168	1,000
L8	89	37	38	14	0,629	0,427	0,467	0,495
L10	95	49	33	13	0,689	0,347	0,428	0,090
L13	97	16	33	48	0,335	0,340	0,446	0,023
L15	96	70	21	5	0,839	0,219	0,271	0,064
L16	99	57	36	6	0,758	0,364	0,367	1,000
L17	82	1	15	66	0,104	0,183	0,186	1,000
L18	99	42	43	14	0,641	0,434	0,460	0,662
L19	99	27	48	24	0,515	0,485	0,500	0,841
L20	97	9	35	53	0,273	0,361	0,397	0,441
L21	87	3	36	48	0,241	0,414	0,366	0,377
L22	94	24	51	19	0,527	0,543	0,499	0,534

O locus 5 (rs34156041) não apresentou nenhum indivíduo homozigoto para o alelo deleção ($n = 97$), o locus 17 (rs59137132) apenas um indivíduo homozigoto para o alelo inserção ($n = 82$) e o locus 21 (rs71330398) três homozigotos para a inserção ($n = 94$).

No teste para o EHW, os loci 4 (rs33926123), 5 (rs34156041), o 16 (rs138481647) e o 17 (rs59137132) apresentaram valor de p igual a 1. O locus 13 (rs67616458) foi o único a apresentar valor de p menor que o nível de significância de 0,05 ($p = 0,023$). O teste sensível à deficiência de heterozigotos indicou a possibilidade de desequilíbrio provocado pelo déficit de heterozigotos ($p = 0,01$).

PARÂMETROS FORENSES

O poder de informação do polimorfismo (PIC), a probabilidade de identidade (PI), o poder de discriminação (PD) e o poder de exclusão (PE) estão descritos na tabela 8.

Tabela 8. Conteúdo informativo do polimorfismo (PIC), Probabilidade de Identidade (PI), Poder de discriminação (PD) e Poder de exclusão (PE).

Lócus	PIC	PI	PD	PE
L1	0,3748	0,3752	0,6248	0,1874
L2	0,3739	0,3761	0,6239	0,1869
L3	0,3561	0,3952	0,6048	0,1781
L4	0,3496	0,4027	0,5973	0,1748
L5	0,1542	0,7058	0,2942	0,0771
L8	0,3577	0,3934	0,6066	0,1789
L10	0,3365	0,4186	0,5814	0,1683
L13	0,3463	0,4066	0,5934	0,1732
L15	0,2341	0,5684	0,4316	0,1171
L16	0,2947	0,4751	0,5249	0,1474
L17	0,1686	0,6801	0,3199	0,0843
L18	0,3542	0,3974	0,6026	0,1771
L19	0,3748	0,3752	0,6248	0,1874
L20	0,3183	0,4423	0,5577	0,1591
L21	0,2992	0,4687	0,5313	0,1496
L22	0,3743	0,3757	0,6243	0,1871
Valor cumulativo	-	-	0,9990	0,9374

Os lócus apresentaram PIC menores que 0,5, sendo considerados de informatividade baixa e mediana. O lócus 1 (rs6143647) e o 19 (rs148112183) apresentaram o maior valor de 0,37, o lócus 5 (rs34156041) foi o menos informativo com PIC de 0,15.

Sendo PI a probabilidade de dois indivíduos aleatórios apresentarem genótipos idênticos para um dado lócus, o PD é o seu complementar e indica a probabilidade de dois indivíduos aleatórios apresentarem genótipos diferentes. Os maiores valores foram alcançados pelos lócus 1 (rs6143647) e 19 (rs148112183), ambos com 0,6248 de poder e o menor valor foi do lócus 5 (rs34156041) 0,2942. O poder de discriminação alcançado para os 16 lócus polimórficos combinados foi de 0,9990 tanto para grupo parental quanto para o filial.

O PE, que indica a probabilidade de exclusão de um indivíduo que foi erroneamente acusado, apresentou valores altos para todos os loci, considerando o valor máximo de 0,19 para loci bialélicos. O maior valor foi 0,1874 nos loci 1 (rs6143647) e 19 (rs148112183). O menor valor foi do locus 17 (rs59137132) com 0,08.

Discussão

AUSÊNCIA DE RESULTADOS

DNA não amplificado (NA)

Alguns indivíduos não apresentaram amplificação mesmo com as repetições das reações de PCR (tabela 8). Os loci 7 (rs74315104) e 17 (rs59137132) obtiveram 17% e 18% de NA no total de 100 indivíduos testados. É possível que alguns iniciadores não sejam exatamente adequados, não por um erro no seu desenho, que foi confirmado por ferramentas usuais e publicamente disponíveis, mas pela escassez de informações no registro das sequências correspondentes à zona de pareamento. Diante da complexidade possível para solução do problema e diante do fato de que esta fase do trabalho consiste basicamente em triagem, também é considerada a possibilidade de algumas reações não estarem completamente padronizadas.

ESCOLHA DOS LÓCUS

A maioria dos loci InDels publicados não eram adequados para o nosso trabalho, principalmente pelo pequeno tamanho das inserções, o que não permitiria o desenho de iniciadores complementares a elas. Para citar dados sobre a população brasileira, dos 48 loci descritos por Santos *et al.* (2009), somente três se adequaram para padronização da técnica de genotipagem fetal não invasiva com uso de InDels e iniciadores inserção-específicos (Santos, 2014). Por isso, foi considerado mais eficaz a busca direta nos bancos de dados do que a consulta à literatura.

Além da possibilidade do desenho de iniciadores internos, a escolha visou loci que fossem polimórficos, isto é, seus dois alelos deveriam ter frequências entre 1% e 99% (Cavalli-Sforza *et al.*, 1994). Entretanto, as informações neste sentido eram e ainda são muito poucas e nos obrigaram a escolher loci sem informações populacionais. Por esta razão, era esperado que alguns, embora registrados como variantes de inserção/deleção, fossem monomórficos.

Locus monomórficos

Os locus 7 (rs74315104), 9 (rs118203674), 11 (rs80356802) e 12 (rs267607475) apresentaram monomorfismo e apenas o alelo inserção esteve presente em todos os indivíduos analisados. Todos eles estão registrados no banco de dados do NCBI como pequenos InDels intragênicos, mas até o momento não há registro sobre associação, patogenicidade ou dados populacionais que pudessem ser comparados com os resultados aqui obtidos.

A confirmação do locus como polimorfismo é uma característica que deve ser usada futuramente como critério de seleção dos próximos locus a serem analisados para uso na genotipagem fetal não invasiva. Isso, porém, não invalida os resultados aqui encontrados, que permitirão um novo registro desses locus no banco de dados com acréscimo de informação sobre as frequências populacionais na amostra da população urbana brasileira.

O locus 2 (rs3842570)

O locus 2 (rs3842570) apresentou um padrão com duas bandas, entretanto não como esperado a partir do desenho dos iniciadores e da PCR *in silico*.

A PCR *in silico* e o BLAT indicaram a especificidade do iniciador, apresentando um produto amplificado de 164 pares de bases (pb) na presença da inserção (alelo i) e de 132 pares de bases na ausência dela (alelo d). Entretanto, em toda amostra amplificada o padrão de bandas mostrou uma banda na altura de 164 pb e outra numa altura aproximada à banda de 200 pb do marcador padrão de peso molecular.

É possível que os iniciadores flanqueadores amplifiquem essa região polimórfica contendo a sequência de 32 pb da inserção, mas que o alelo alternativo seja a presença de uma repetição dessa sequência, gerando então um produto amplificado de 196 pb. Já é sabido que pequenas InDels (até 100 pb) podem gerar inserções de duplicações em tandem (Levinson e Gutman, 1987) e no caso aqui encontrado, a região registrada como um polimorfismo inserção/deleção poderá caracteriza-se como um polimorfismo de repetição, onde os alelos podem ser: de 196 pb, de 164 pb e outro de 132 pb (Figura 5). O locus seria descartado da análise se a sequência estivesse repetida 2,5 vezes ou mais, sendo considerado como um possível STR (LaRue *et al.*, 2014).

Figura 5. Esquema mostrando possíveis alelos do locus 2 (rs3842570). Alelo deleção com 132 pares de bases (pb); alelo inserção com 164 pb e alelo com a inserção repetida uma vez (196 pb).

Dados preliminares de sequenciamento deste locus confirmam o tamanho de 164 pb do fragmento inserção e indica a possibilidade de que o fragmento maior tenha 196 pb.

ANÁLISE ESTATÍSTICA

A análise estatística foi realizada apenas com os dados relativos aos loci polimórficos.

Dezesseis loci foram analisados e apenas o locus 13 (rs67616458) não apresentou proporções fenotípicas de acordo com o esperado pelo EHW ($p = 0,02$). O teste para o déficit de heterozigotos indicou que essa pode ser a causa ($p = 0,01$). Considerando que as proporções no EHW são teóricas e que as suas condições perfeitas são de uma população modelo, o desvio do equilíbrio pode ocorrer, além do desbalanço da quantidade de heterozigotos, por subestruturação populacional, fluxo gênico, seleção, deriva genética, endogamia ou casamento não aleatório.

Os loci 4 (rs33926123), 5 (rs34156041), 16 (rs138481647) e 17 (rs59137132) apresentaram valor de p para o teste do EHW igual a 1. Os testes de hipótese foram feitos para controle de qualidade dos dados, mas sem assumir o valor de p como evidência. O locus 13 não será desconsiderado nas análises seguintes, pois um valor de p abaixo do nível de significância de 0,05 não significa que esse locus deva ser evitado (Foreman e Evett, 2001).

A maior frequência do alelo inserção (i) foi 0,907 no locus 5 (rs34156041), o que pode ser explicado pela ausência de indivíduos homozigotos para deleção (dd). No locus 17 (rs59137132), além da grande quantidade de indivíduos não amplificados, foi

observado apenas um indivíduo homozigoto ii, resultando numa frequência 0,1 do alelo i. Essa baixa frequência alélica nos loci mencionados indica um baixo grau de polimorfismo, mas ainda dentro do percentual de 99% que caracteriza um locus como polimórfico segundo definição de Cavalli-Sforza (1994).

Alguns loci tiveram dados de frequências populacionais determinados pela fase 3 do Projeto 1000 Genomas (disponível em: www.1000genomes.org), que estudou populações de origem Africana (AFR), Americana (AMR), Asiática (ASN) e Europeia (EUR). A distribuição das frequências está descrita na tabela 9, onde foram colocadas também as frequências da população urbana brasileira (GP) para melhor visualização.

Tabela 9. Frequência do alelo inserção em diferentes populações.

Locus	Rs	Geral n = 2504	AFR n = 661	AMR n = 347	ASN n = 504	EUR n = 503	GP n = 100
L1	rs6143647	-	-	-	-	-	0,485
L2	rs3842570	-	-	-	-	-	0,522
L3	rs67696659	0,670	0,775	0,661	0,691	0,538	0,635
L4	rs33926123	0,005	0,017	0,003	0,000	0,000	0,344
L5	rs34156041	0,900	0,859	0,911	0,982	0,895	0,907
L8	rs1305049	0,580	0,276	0,650	0,550	0,803	0,629
L10	rs34656661	-	-	-	-	-	0,689
L13	rs67616458	0,340	0,461	0,232	0,286	0,325	0,335
L15	rs11275778	0,000	0,000	0,000	0,000	0,000	0,839
L16	rs138481647	0,843	0,801	0,752	0,944	0,862	0,758
L17	rs59137132	0,179	0,168	0,110	0,448	0,081	0,104
L18	rs140462903	0,606	0,355	0,644	0,720	0,794	0,641
L19	rs148112183	0,571	0,841	0,522	0,643	0,361	0,515
L20	rs71311113	0,000	0,000	0,000	0,000	0,000	0,273
L21	rs71330398	0,172	0,098	0,269	0,188	0,146	0,241
L22	rs67519581	0,636	0,690	0,523	0,788	0,474	0,527

A alta frequência do alelo i encontrada no locus 5 (rs34156041) se confirma nas populações estudadas, sendo mais próxima da frequência da população americana (AMR), bem como a baixa frequência da inserção no locus 17 (rs59137132). Os loci 3 (rs67696659), 8 (rs1305049), 16 (rs138481647), 18 (rs140462903), 19 (rs148112183), 21 (rs71330398) e 22 (rs67519581) também tiveram frequências do alelo inserção similares às encontradas na população americana.

Na população urbana brasileira estudada no presente trabalho, a frequência do alelo i no 4 (rs33926123) foi de 0,344. Esse valor não foi confirmado nem na população americana, nem nas outras estudadas no 1000 Genomas, onde a frequência foi muito próxima a zero. O mesmo ocorreu para os loci 15 (rs11275778) e 20 (rs71311113) que não apresentaram frequência significativa do alelo i no projeto citado. Entretanto, na população brasileira estudada, a frequência do alelo i no locus 15 (rs11275778) foi 0,839, considerada elevada, e 0,273 no locus 20 (rs71311113).

Esses dados não foram consultados anteriormente para a seleção dos loci, pois não estavam disponíveis no banco de dados público. A última atualização da fase 3 do Projeto 1000 Genomas foi feita em outubro de 2014 e algumas variantes estudadas nessa fase ainda estão em processo de arquivamento no banco de dados do NCBI.

PARÂMETROS FORENSES

os valores de PIC vão de 0 a 1. Aqui, os loci apresentaram PIC menores que 0,5 sendo considerados de informatividade baixa a mediana. O maior valor alcançado foi 0,3748 nos loci 1 (rs6143647) e 19 (rs148112183). Isso significa que 37,48% da prole deve ser informativa sobre loci parentais com alelos raros e em desequilíbrio de ligação com o marcador em questão. O valor de PIC é diretamente proporcional ao número de alelos e suas frequências e para um marcador com dois alelos o seu valor máximo é 0,375 (Hildebrand *et al.*, 1992). Considerando que em sistemas bialélicos os marcadores são mais informativos quando as frequências dos alelos se igualam, os dois loci que aqui obtiveram o maior valor de PIC possível são exatamente aqueles com frequências alélicas próximas a 0,5.

A eficácia dos marcadores moleculares para uso na identificação humana é alcançada com o uso de marcadores combinados. O poder de discriminação dos 16 loci InDels combinados foi de 0,9990, ou seja, a probabilidade de se diferenciar dois indivíduos não aparentados em uma população é de 99,90%. De forma complementar, a probabilidade de identificar dois indivíduos com o mesmo genótipo com a análise desses 16 loci é de aproximadamente 0,001. O PD ficou próximo de 0,9999 encontrado em um painel com 20 InDels bialélicos combinados (Huang *et al.*, 2014) e outro com 30 InDels (Torres *et al.*, 2014).

Para o poder de exclusão, o maior valor alcançado foi de 0,1874 nos loci 1 (rs6143647), 2 (rs3842570) e 19 (rs148112183), alcançando 98,63% do valor máximo esperado de 0,19 para um locus bialélico (Jamieson, 1997). O poder de exclusão dos

16 loci combinados (PEC) foi de 0,9374, indicando que a probabilidade de excluir um indivíduo falsamente acusado em pelo menos um dos 16 loci é de 93,74%.

Se, hipoteticamente, mantivermos os valores de PEC encontrado em GP e duplicarmos o número de loci, o PEC alcançado seria de 99,50%. Se o número for triplicado chegando a 48 loci, o PEC seria de 99,96% ainda não alcançando os 99,99% rotineiramente encontrados nas análises com os microssatélites. Esses dados apontam que é necessário um conjunto com mais ou menos 60 InDels, como sugerido por Fondevila (2012), para que se alcance o poder de exclusão satisfatório.

GENOTIPAGEM FETAL NÃO INVASIVA

Como visto por Santos (2014) o uso dos iniciadores internos permite uma fácil detecção do cffDNA, bem como sua classificação genotípica.

A caracterização do cffDNA com esses loci permitirá a determinação da paternidade de forma não invasiva ainda durante a gravidez. Uma aplicação imediata seria a dosagem de cffDNA para diagnóstico de pré-eclâmpsia. Em revisão sobre a relação da concentração de cffDNA com a pré-eclâmpsia (Martin *et al.*, 2014), nenhum trabalho revisado utilizou marcador do tipo InDel. Outra aplicação seria a dosagem de DNA em pequenas quantidades oriundo de células de pacientes receptores de transplante de medula óssea.

Os loci aqui estudados tiveram suas características descritas na população urbana brasileira pela primeira vez. Esses resultados são fundamentais para prosseguir com a padronização dos iniciadores internos, complementares à inserção, que poderão ser usados para detecção de cffDNA.

Conclusão

No presente trabalho é apresentada a análise de 20 loci InDels, um em cada cromossomo, com o objetivo de determinar suas frequências em uma amostra da população urbana brasileira, da qual resultou na descrição de 16 deles.

Com a amplificação dos 16 loci polimórficos no tamanho previsto, consideramos a confirmação das sequências complementares aos iniciadores, registradas no banco de dados do NCBI e descrevemos as frequências de seus alelos numa amostra da população urbana brasileira.

Observamos que o locus 2 (rs3842570) possui um alelo ainda não registrado. O novo alelo parece ser caracterizado pela repetição da sequência inserida, o que não invalida sua classificação como polimorfismo de inserção/deleção, mas o torna um locus InDel trialélico.

Acreditamos que a hipótese de que um conjunto de loci inserção/deleção atenderia a critérios populacionais para a determinação de um perfil genético permitindo o seu uso na genotipagem fetal não invasiva possa ser confirmada. No total, 16 loci apresentaram resultados de análise estatística e forense satisfatórios para compor um painel com outros InDels e prosseguir para a etapa de padronização dos iniciadores internos, que permitirão seu uso para genotipagem fetal não invasiva.

Referências Bibliográficas

- Alberry, M., Maddocks, D., Jones, M., Abdel Hadi, M., Abdel-Fattah, S., Avent, N. *et al.* Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast. *Prenatal Diagnosis*. 2007;27:415-418.
- Alberry, M. S., Maddocks, D. G., Hadi, M. A., Metawi, H., Hunt, L. P., Abdel-Fattah, S. A. *et al.* Quantification of cell free fetal DNA in maternal plasma in normal pregnancies and in pregnancies with placental dysfunction. *American Journal of Obstetrics & Gynecology*. 2009;200:981-986.
- Alkan, C., Coe, B. P., Eichler, E. E. Genome structural variation discovery and genotyping. *Nature Reviews Genetics*. 2001 may 2011;12:363-376.
- Amaral, D. R. T., Credidio, D. C., Pellegrino-Jr, J., Castilho, L. Fetal RHD genotyping by analysis of maternal plasma in a mixed population. *Journal of Clinical Laboratory Analysis*. 2011;25:100-104.
- Attwood, H. D., Park, W. W. Embolism to the lungs by trophoblast. *J Obstet Gynaecol Br Commonw*. 1961;68:611-617.
- Bastos-Rodrigues, L., Pimenta, J. R., Pena, S. D. J. The Genetic Structure of Human Populations Studied Through Short Insertion-Deletion Polymorphisms. *Annals of Human Genetics*. 2006;70:658-665.
- Bianchi, D. W. Fetomaternal cell trafficking: a new cause of disease? *American Journal of Medical Genetics*. 2000;91:22-28.
- Bianchi, D. W. Fetomaternal cell traffic, pregnancy-associated progenitor cells, and autoimmune disease. *Best Practice & Research Clinical Obstetrics and Gynaecology*. 2004;18:959-975.
- Bischoff, F. Z., Lewis, D. E., Simpson, J. L. Cell-free fetal DNA in maternal blood: kinetics, source and structure. 2005;11:59-67.
- Botstein, D., White, R. L., Skolnick, M., Davis, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetic*. 1980;32:314-331.
- Brookes, A. The essence of SNPs. *Gene*. 1999;234:177-186.
- Butler, J. M. *Forensic DNA typing Biology, technology and genetics of STR markers*. 2 ed: Academic Press 2005.
- Cavalli-Sforza, L. L., Menozzi, P., Piazza, A. *The history and geography of human genes*. Princeton, USA: Princeton University Press 1994.
- Chakraborty, R., Stivers, D. N., Su, B., Zhong, Y., Budowle, B. The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems. *Electrophoresis*. 1999;20:1682-1696.
- Chiu, R. W. K., Lo, Y.-M. D. Application of fetal DNA in maternal plasma for noninvasive prenatal diagnosis. *Expert Review Molecular Diagnostics*. 2002;2:32-40.
- Desmarais, D., Zhong, Y., Chakraborty, R., Perreault, C., Busque, L. Development of a highly polymorphic STR marker for identity test purpose at the Human Androgen Receptor Gene (Humara). *Journal Forensic Science*. 1998;5:1046-1049.

- Evans, M. I., Andriole, S. Chorionic villus sampling and amniocentesis in 2008. *Current Opinion in Obstetrics and Gynecology*. 2008;20:164-168.
- Fondevila, M., Phillips, C., Santos, C., Pereira, R., Gusmão, L., Carracedo, Á. *et al.* Forensic performance of two insertion-deletion marker assays. *International Journal of Legal Medicine*. 2012;126:725-737.
- Foreman, L. A., Evett, I. W. Statistical analyses to support forensic interpretation for a new ten-locus STR profiling system. *International Journal Legal Medicine*. 2001;114:147-155.
- Francez, P. A. C., Ribeiro-Rodrigues, E. M., Velasco, A. M., Santos, S. E. Insertion-deletion polymorphisms—utilization on forensic analysis. *International Journal of Legal Medicine*. 2011;126:491-496.
- Gude, N. M., Roberts, C. T., Kalionis, B., King, R. G. Growth and function of the normal human placenta. *Thrombosis Research*. 2004;14:397-407.
- Hartl, D. L., Clark, A. G. *Princípios de Genética de Populações*. 4 ed: Artmed 2010.
- Higuchi, R. Simple and rapid preparation of samples for PCR. In: Erlich H, ed. *PCR technology - principles and applications for DNA amplification*. New York: Stockton press 1989:31-38.
- Hildebrand, C. E., Torney, D. C., Wagner, R. P. Informativeness of polymorphic DNA markers. *Los Alamos Science*. 1992;20:100-102.
- Huang, J., Luo, H., Wei, W., Hou, Y. A novel method for the analysis of 20 multi-Indel polymorphisms and its forensic application. *Electrophoresis*. 2014;35:487-493.
- Jamieson, A., Taylor, S. C. S. Comparisons of three probability formulae for parentage exclusion. *Animal Genetics*. 1997;28:397-400.
- Jeffreys, A. J., Wilson, V., S.L., T. Individual-specific "fingerprints" of human DNA. *Nature*. 1985;316:76-79.
- Kayser, M., Roewer, L., Hedman, M., Henke, L., Henke, J., Sajantila, A. Characteristics and Frequency of Germline Mutations at Microsatellite Loci from the Human Y Chromosome, as Revealed by Direct Observation in Father/Son Pairs. *American Journal of Human Genetic*. 2000;66:1580-1588.
- Koide, K., Sekizawa, A., Iwasaki, M., Matsuoka, E., Honma, S., Farina, A. *et al.* Fragmentation of cell-free fetal DNA in plasma and urine of pregnant women. *Prenatal Diagnosis*. 2005;25:604-607.
- Kolialexi, A., Tsangaris, G. T. H., Antsaklis, A., Mavrou, A. Rapid clearance of fetal cells from maternal circulation after delivery. *Annals of New York Academy Science*. 2006;1022:113-118.
- LaRue, B. L., Lagacé, R., Chang, C. W., Holt, A., Hennessy, L., Ge, J. *et al.* Characterization of 114 insertion/deletion (INDEL) polymorphisms, and selection for a global INDEL panel for human identification. *Legal Medicine*. 2014;16:26-32.
- Levi, J. E., Wendel, S., Takaoka, D. T. Determinação pré-natal do sexo fetal por meio da análise do DNA no plasma materno. *Revista Brasileira de Ginecologia e Obstetrícia*. 2003;25:687-690.
- Levinson, G., Gutman, G. A. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Molecular Biology Evolution*. 1987;4(3):203-221.

- Li, Y., Zimmermann, B., Rusterholz, C., Kang, A., Holzgreve, W., Hahn, S. Size Separation of Circulatory DNA in Maternal Plasma Permits Ready Detection of Fetal DNA Polymorphisms. *Clinical Chemistry*. 2004;50:1002-1011.
- Lo, Y.-M. D. Noninvasive prenatal detection of fetal chromosomal aneuploidies by maternal plasma nucleic acid analysis: a review of the current state of the art. *BJOG: An International Journal of Obstetrics & Gynecology*. 2009;116:152-157.
- Lo, Y.-M. D., Chan, K. C. A., Sun, H., Chen, E. Z., Jiang, P., Lun, F. M. F. *et al*. Maternal Plasma DNA Sequencing Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus. *Prenatal Diagnosis*. 2010;2:61-91.
- Lo, Y.-M. D., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W. G. *et al*. Presence of fetal DNA in maternal plasma and serum. *The Lancet*. 1997;350:485-487.
- Lo, Y.-M. D., Leung, T. N., Tein, M. S. C., Sargent, I. L., Zhang, J., Lau, T. K. *et al*. Quantitative Abnormalities of Fetal DNA in Maternal Serum in Preeclampsia. *Clinical Chemistry*. 1999;45:184-188.
- Lo, Y.-M. D., Tein, M. S. C., Haines, C. J., Leung, T. N., Poon, P. M. K., Wainscoat, J. S. *et al*. Quantitative Analysis of Fetal DNA in Maternal Plasma and Serum: Implications for Noninvasive Prenatal Diagnosis. *American Journal of Human Genetic*. 1998;62:768-775.
- Lo, Y.-M. D., Wainscoat, J. S., Gillmer, M. D. G., Patel, P., Sampietro, M., Fleming, K. A. Prenatal sex determination by DNA amplification from maternal peripheral blood. *The Lancet*. 1989:1363-1365.
- Manta, F., Caiafa, A., Pereira, R., Silva, D., Amorim, A., Carvalho, E. F. *et al*. Indel markers: Genetic diversity of 38 polymorphisms in Brazilian populations and application in a paternity investigation with post mortem material. *Forensic Science International: Genetics*. 2012;6:658-661.
- Martin, A., Krishna, I., Martina, B., Samuel, A. Can the quantity of cell-free fetal DNA predict preeclampsia: a systematic review. *Prenatal Diagnosis*. 2014;34:1-7.
- Mills, R. E., Luttig, C. T., Larkins, C. E., Beauchamp, A., Tsui, C., Pittard, W. S. *et al*. An initial map of insertion and deletion (INDEL) variation in the human genome. *Genome Research*. 2006;16:1182-1190.
- Peakall, R., Smouse, P. E. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 2006;6:288-295.
- Pena, S. D. J., Chakraborty, R. Paternity testing in DNA era. *Trends in Genetics*. 1994;10:204-209.
- Pena, S. D. J., Jeffreys, A. J. Breve Introdução Às impressões digitais de DNA. *Revista Brasileira de Genética*. 1993;16:857-879.
- Pereira, R., Phillips, C., Alves, C., Amorim, A., Carracedo, Á., Gusmão, L. A new multiplex for human identification using insertion/deletion polymorphisms. *Electrophoresis*. 2009;30:3682-3690.
- Pertl, B., Sekizawa, A., Samura, O., Orescovic, I., Rahaim, P. T., Bianchi, D. W. Detection of male and female fetal DNA in maternal plasma by multiplex fluorescent polymerase chain reaction amplification of short tandem repeats. *Human Genetics*. 2000;106:45-49.

Ribeiro-Rodrigues, E. M., Santos, N. P. C., Ribeiro-dos-Santos, A. K. C., Pereira, R., Amorim, A., Gusmão, L. *et al.* Assessing Interethnic Admixture Using an X-Linked Insertion-Deletion Multiplex. *American Journal of Human Biology*. 2009;21:707-709.

Romão, R. M., Levi, J. E., Carvalho, M. H. B., Francisco, R. P. V., Amorim, A. G. d., Zugaib, M. Use of cell-free fetal nucleic acids in maternal blood for prenatal diagnosis: the reality of this scenario in Brazil. *Revista da Associação Médica Brasileira*. 2012;58:615-619.

Rousset, F. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*. 2008;8(1):103-106.

Santos, M. D. Marcadores inserção/deleção na identificação humana pré-natal não invasiva. Ribeirão Preto: Universidade de São Paulo; 2014.

Sensabaugh, G. Biochemical markers of individuality. In: Saferstein R, ed. *Forensic Science Handbook*. Nova Iorque: Prentice-Hall 1982:338-415.

Sobrino, B., Brión, M., Carracedo, Á. SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Science International*. 2005;154:181-194.

Steele, M. W., Breg, W. R. Chromosome analysis of human amniotic-fluid cells. *Lancet*. 1966;1:383-385.

Thompson, J., Thompson, M. *Genética Médica*. 6 ed. Rio de Janeiro: Guanabara Koogan 2002.

Torres, S. R. R., Uehara, C. J. S., Sutter-Latorre, A. F., Almeida, B. S., Sauerbier, T. S., Muniz, Y. C. N. *et al.* Population genetic data and forensic parameters of 30 autosomal InDel markers in Santa Catarina State population, Southern Brazil. *Molecular Biology Reports*. 2014;41:5429-5433.

Venter, C. J., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G. *et al.* The Sequence of the Human Genome. *Science*. 2001;291:1304-1351.

Vodicka, R., Böhmová, J., Dhaifalah, I., Blumenthalová, J., Kratochvilová, R., Procházka, M. *et al.* Analysis of cell free fetal DNA fragmentation rate in pregnant women during the course of gravidit. *Ceská gynekologie*. 2010;75:312-316.

Vodicka, R., Vrtěl, R., Böhmová, J., Kratochvilova, R., Dusek, L., Dhaifalah, I. *et al.* Noninvasive Prenatal Nucleic Acid Diagnostics of Down Syndrome. In: Dey PS, ed. *Genetics and Etiology of Down Syndrome* 2011:313-328.

Wagner, J., Džijan, S., Marjanović, D., Lauc, G. Non-invasive prenatal paternity testing from maternal blood. *International Journal Legal Medicine*. 2009;123:75-79.

Watananara, T., Bianchi, D. W. Fetal cell-free nucleic acids in the maternal circulation. *Annals of New York Academy Science*. 2004;1022:90-99.

Weber, J. L., David, D., Heil, J., Fan, Y., Zhang, C., Marth, G. Human Diallelic Insertion/Deletion Polymorphisms. *American Journal of Human Genetic*. 2002;71:854-862.

Weir, B. S. *Genetic data analysis: methods for discrete population genetic data*. 2a ed. Massachusetts: Sinauer Associates 1996.

Yu, H., Shen, Y., Ge, J., He, Y., Qiao, D., Ren, M. *et al.* Quantification of maternal serum cell-free fetal DNA in early-onset preeclampsia. *International Journal of Molecular Sciences*. 2013;14:7571-7582.

Zhong, X. Y., Bürk, M. R., Troeger, C., Jackson, L. R., Holzgreve, W., Hahn, S. Fetal DNA in maternal plasma is elevated in pregnancies with aneuploid fetuses. *Prenatal Diagnosis*. 2000;20:795-798.

Zhong, X. Y., Laivuori, H., Livingston, J. C., Ylikorkala, O., Sibai, B. M., Holzgreve, W. *et al.* Elevation of both maternal and fetal extracellular circulating deoxyribonucleic acid concentrations in the plasma of pregnant women with preeclampsia. *American Journal of Obstetrics & Gynecology*. 2001;184:414-419.

Apêndices

Apêndice I – Ofícios Comitê de Ética

Aprovação do Banco de Amostras *Populações Urbanas*

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

CAMPUS UNIVERSITÁRIO - MONTE ALEGRE
FONE: 602-1000 - FAX (016) 633-1144

Ribeirão Preto, 02 de março de 2005

Ofício nº 570/2005
CEP/SPC

Senhor Professor:

O Comitê de Ética em Pesquisa recebeu carta, onde Vossa Senhoria encaminha a proposta de criação e reconhecimento de Banco de Amostras designado **“Populações Urbanas”** localizado no Laboratório de Genética e Bioquímica, do Departamento de Genética – FMRP-USP.

Sobre o assunto, o Comitê de Ética em Pesquisa em sua 196ª Reunião Ordinária realizada em 28.02.2005, analisou a proposta de criação e a enquadrou na categoria: **APROVADO**, de acordo com o Processo HCRP nº 8079/2004.

Aproveito a oportunidade para apresentar a Vossa Senhoria protestos de estima e consideração.

PROF. DR. SÉRGIO PEREIRA DA CUNHA
Coordenador do Comitê de Ética
em Pesquisa do HCFMRP-USP e da FMRP-USP

Ilustríssimo Senhor
PROF. DR. AGUINALDO LUIZ SIMÕES
Depto. de Genética – FMRP-USP
Em mãos

Aprovação do projeto

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

Ribeirão Preto, 02 de outubro de 2014

Ofício nº 3622/2014
CEP/FC

Prezados Senhores,

O trabalho intitulado “**VALIDAÇÃO DE MARCADORES INSERÇÃO/DELEÇÃO PARA GENOTIPAGEM FETAL NÃO INVASIVA**” – de 24/09/2014 foi analisado pelo Comitê de Ética em Pesquisa, em sua 395ª Reunião Ordinária realizada em 29/09/2014 e enquadrado na categoria: APROVADO, bem como a solicitação de dispensa do Termo de Consentimento Livre e Esclarecido, de acordo com o Processo HCRP nº 11015/2014.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (IGH-GCP), bem como a Resolução nº 466/12 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente.

PROF. DR. FÁBIO CARMONA
Vice-Coordenador do Comitê de Ética em
Pesquisa do HCRP e da FMRP-USP

Ilustríssimos Senhores
AYLING MARTINS NG
PROF. DR. AGUINALDO LUIZ SIMÕES (Orientador)
Depto. de Genética – FMRP-USP

Campus Universitário – Monte Alegre
14048-900 Ribeirão Preto SP

Comitê de Ética em Pesquisa HCRP e FMRP-USP
FWA-00002733; IRB-00002186 e Registro PB/CONEP nº 5440
(16) 3602-2228
cep@hcrp.usp.br

www.hcrp.usp.br

Apêndice II – Protocolos de análise laboratorial

Protocolo 1 – Reação em cadeia da polimerase

Reagentes e soluções

Tampão de reação 10X livre de $MgCl_2$: Tris/HCl 75 mM pH 9,0; KCl 50 mM; $(NH_4)_2SO_4$ 20 mM (BIOTOOLS-B&M Labs, SA).

$MgCl_2$: 50mM (BIOTOOLS-B&M Labs, SA).

Solução de iniciadores: solução trabalho a 2,5 μM , sendo 10 μL do iniciador direto, 10 μL do iniciador inverso e 180 μL de água destilada.

dNTP: 20 mM (BIOTOOLS-B&M Labs, SA).

DNA polimerase (Taq): 5 U/ μL (BIOTOOLS-B&M Labs, SA).

Água destilada ultrafiltrada (MilliQ - Millipore Corporation).

Procedimento

Ensaio de PCR convencional com volume final de 25 μL :

- Misture os reagentes acima em quantidades específicas: 2,5 μL de tampão de reação; 1 μL de $MgCl_2$; 0,25 μL de dNTP; 2 μL de iniciador; 0,1 μL de Taq

As quantidades de iniciador e Taq podem ser variáveis para cada locus (tabela 3).

- Acrescente água destilada para completar o volume final de 21 μL para o mix de reagentes.

Em microtubos de 0,5 mL (ou 0,2 mL quando em placa de 96 poços):

- Pipete 4 μL de DNA genômico, já extraído das amostras, em seu respectivo microtubo;

- Pipete 4 μL de água destilada em um microtubo como controle negativo de amplificação;
- Acrescente em cada microtubo os 21 μL do mix de reagentes.

Cada reação deve ter um controle negativo.

Protocolo 2 – Eletroforese em gel de poliacrilamida não-desnaturante

Reagentes e soluções

Solução de acrilamida/bisacrilamida (29:1): 29 g de acrilamida (PM= 71,08); 1 g de bisacrilamida (N,N'-metileno-bis-acrilamida; PM= 154,17), 100 mL de água destilada.

Solução de EDTA 0,5 M pH 8,0: 186 g de EDTA (PM = 372,24); 1 L de água destilada; pastilhas de hidróxido de sódio (NaOH) para ajuste do pH.

Tampão TBE 10X 0,9 M pH 8,0: 108 g de Tris (PM = 121,1); 53 g de ácido bórico (PM = 61,83); 40 mL de solução de EDTA 0,5 M pH 8,0; água destilada para completar volume final de 1 L.

Tampão de corrida TBE 1X 0,9 M pH 8,0: 100 mL de Tampão TBE 10X; 900 mL de água destilada.

TEMED (N,N,N',N'-tetrametiletenodiamina) (SIGMA-ALDRICH).

Solução saturada de persulfato de potássio: 650 mg de persulfato de potássio (PM = 270,32); 6,5 mL de água destilada.

Tampão de amostra: 900 µL de bromofenol; 900 µL de xilenocianol; 900 µL de TBE 10X; 4,5 mL de Ficol 30% diluído em água; 1,8 mL de EDTA 0,5 M pH 8,0; 3,6 g de sacarose (completamente diluída na solução final).

Mix para gel não-desnaturante 10%: 66,66 mL de solução acrilamida/bisacrilamida; 14 mL de glicerol; 97 mL de água destilada; 20 mL de TBE 10%.

Marcador padrão de peso molecular 50 pb: 5 µL de solução estoque do marcador padrão de peso molecular 500 µg/µL; 150 µL de água destilada; 50 µL de tampão de amostra.

Procedimento

Montagem do cassete de vidro:

- Coloque, entre duas placas de vidro (16 cm x 16 cm) próprias para eletroforese, dois espaçadores de teflon em cada borda lateral;

- Prenda o conjunto com grampos nas bordas laterais;

Preparação do gel:

- Misture 20 mL do mix para gel mais 20 μ L de TEMED e 500 μ L de persulfato de potássio (catalisadores da polimerização);
- Imediatamente depois, coloque cuidadosamente a mistura no cassete de vidro montado;
- Encaixe um pente de teflon no limite superior do cassete com o gel ainda não polimerizado para que sejam formados os poços, onde serão aplicadas as amostras do DNA amplificado.

O pente deve ser colocado o mais rápido possível evitando que posteriormente fiquem resquícios de gel polimerizado nos poços.

A polimerização do gel ocorre em aproximadamente 20 minutos.

- Remova o pente e os grampos cuidadosamente e lave os poços com água sem desmontar o cassete.

A água acumulada nos poços pode ser removida vertendo-se repetidamente o cassete com os poços para baixo, de maneira cuidadosa.

Eletroforese:

- Posicione o cassete com o gel na cuba de eletroforese vertical, contendo tampão de corrida (TBE 1X) em ambos os polos (superior e inferior), de maneira que os poços se encham com o tampão;
- Conecte a cuba a uma fonte de voltagem constante.

Por segurança, é recomendado fazer uma pré-corrida ligando-se a fonte antes da aplicação das amostras para checar o correto funcionamento do aparato.

- Pipete em cada poço um mix de 5 μ L da amostra de DNA amplificado e aproximadamente 5 μ L de tampão de amostra.
- Pipete 3 μ L do marcador padrão de peso molecular (50 pb) em poço separado, afim de comparar o padrão de altura das bandas.

É recomendado que se coloque um poço com o marcador padrão de peso molecular ao lado de cada locus analisado num mesmo gel.

- Após aplicação das amostras, inicie a eletroforese ligando a fonte na voltagem padronizada.

A duração da corrida eletroforética é baseada no tamanho das bandas e sua posição desejada no gel.

Protocolo 3 – Coloração com nitrato de prata

Reagentes e soluções

Solução fixadora: 166,67 mL de álcool etílico; 6,67 mL de ácido acético glacial e 826,66 mL de H₂O.

Solução reveladora: 22,5g de NaOH (PM= 40,00); 1L de H₂O. No momento da coloração, adicionar 3 mL de formaldeído (PM = 30,05) para cada 100 mL de solução reveladora utilizada.

Solução de nitrato de prata: 100 g de nitrato de prata; 1 L de água (armazenar a solução sob abrigo da luz).

Procedimento

- Retire os espaçadores do cassete e separe as placas de vidro, cuidadosamente com ajuda de uma espátula, para remoção do gel;
- Mergulhe o gel em recipiente contendo 100 mL de solução fixadora
- Logo após, acrescente 4 mL da solução de nitrato de prata;
- Mantenha o gel em agitação nessa solução por 5 minutos para uma impregnação uniforme da prata;
- Descarte a solução anterior e adicione 100 mL de água destilada a 60°C para lavagem do gel;
- Descarte a água com resquícios de nitrato de prata e adicione 100 mL da solução reveladora levemente aquecida;
- Acrescente de maneira uniforme 3 mL de formaldeído.

O recipiente deve ser agitado por alguns minutos até que as bandas estejam nítidas.

Após descarte da solução reveladora, a reação de revelação é bloqueada com um pouco de solução fixadora.

Apêndice III – Classificação fenotípica

Indiv.	L1	L2	L3	L4	L5	L7	L8	L9	L10	L11	L12	L13	L15	L16	L17	L18	L19	L20	L21	L22	Total NA
1	ii	na	id	na	na	na	na	ii	na	na	na	na	na	na	na	ii	na	na	id	dd	14
2	id	ii	ii	id	ii	ii	ii	ii	ii	ii	ii	dd	ii	ii	dd	id	dd	id	id	id	0
3	ii	na	id	na	ii	na	id	ii	ii	ii	ii	dd	ii	id	dd	id	id	dd	id	ii	3
4	id	ii	id	id	ii	ii	id	ii	ii	ii	ii	dd	ii	id	id	ii	ii	id	dd	ii	0
5	id	ii	dd	dd	ii	ii	ii	ii	id	ii	ii	dd	ii	ii	dd	dd	id	dd	id	id	0
6	id	na	id	id	id	ii	na	ii	na	ii	ii	ii	ii	ii	id	id	dd	dd	id	dd	3
7	id	na	id	na	ii	ii	dd	ii	ii	ii	ii	ii	ii	ii	dd	na	id	dd	id	ii	3
8	id	ii	ii	dd	id	ii	id	ii	id	ii	ii	dd	ii	ii	dd	dd	ii	ii	dd	id	0
9	id	ii	id	dd	ii	ii	id	ii	id	ii	ii	dd	id	ii	dd	id	id	dd	id	ii	0
10	na	id	id	dd	na	ii	na	ii	na	ii	ii	dd	id	id	na	ii	id	dd	na	na	7
11	id	dd	ii	dd	id	ii	id	ii	dd	ii	ii	dd	ii	ii	dd	dd	id	id	dd	ii	0
12	id	id	id	dd	ii	ii	id	ii	na	ii	ii	id	ii	id	na	id	id	dd	id	id	2
13	dd	na	id	dd	ii	ii	ii	ii	id	ii	ii	ii	ii	id	na	ii	id	dd	na	id	3
14	id	ii	ii	dd	ii	ii	ii	ii	ii	ii	ii	dd	ii	id	dd	ii	ii	dd	dd	ii	0
15	id	ii	id	id	ii	ii	dd	ii	id	ii	ii	id	dd	id	dd	id	id	dd	dd	id	0
16	ii	ii	id	id	id	ii	ii	ii	ii	ii	ii	id	id	dd	dd	id	dd	dd	dd	id	0
17	id	ii	dd	id	id	ii	id	ii	ii	ii	ii	dd	ii	id	dd	ii	id	id	id	dd	0
18	id	na	ii	id	ii	ii	id	ii	id	ii	ii	dd	ii	ii	dd	id	id	dd	na	id	2
19	ii	dd	id	dd	ii	na	ii	ii	id	ii	ii	id	ii	ii	dd	ii	id	id	dd	id	1
20	dd	ii	ii	id	ii	na	ii	ii	id	ii	ii	ii	ii	ii	dd	id	dd	dd	dd	id	1

ii = homozigoto inserção/inserção; id = heterozigoto; dd = homozigoto deleção/deleção; na = não amplificado.

Indiv.	L1	L2	L3	L4	L5	L7	L8	L9	L10	L11	L12	L13	L15	L16	L17	L18	L19	L20	L21	L22	Total NA
21	id	id	ii	ii	id	na	id	ii	ii	ii	ii	dd	id	id	id	ii	id	dd	id	id	1
22	dd	id	dd	ii	ii	na	dd	ii	id	ii	ii	id	ii	id	dd	dd	dd	ii	na	ii	2
23	id	id	ii	dd	ii	na	na	ii	id	ii	ii	ii	ii	ii	na	ii	id	dd	na	id	4
24	id	na	id	id	ii	na	na	ii	ii	na	ii	dd	ii	ii	na	ii	id	dd	dd	dd	5
25	ii	ii	ii	ii	ii	na	ii	ii	ii	ii	ii	dd	id	ii	dd	id	dd	dd	dd	id	1
26	dd	id	ii	id	ii	na	id	ii	id	ii	ii	dd	ii	id	dd	ii	li	dd	id	id	1
27	dd	ii	dd	id	ii	ii	ii	ii	id	ii	ii	ii	id	ii	dd	id	li	ii	dd	dd	0
28	ii	dd	ii	dd	ii	na	id	ii	ii	ii	ii	id	id	id	na	dd	dd	dd	na	na	4
29	id	ii	id	id	ii	na	id	ii	dd	ii	ii	dd	ii	ii	dd	id	dd	dd	dd	dd	1
30	id	id	id	ii	ii	na	dd	ii	id	ii	ii	dd	ii	ii	na	dd	id	id	dd	id	2
31	ii	ii	id	dd	ii	na	dd	ii	dd	ii	ii	id	ii	dd	dd	id	id	id	dd	id	1
32	id	dd	ii	dd	ii	na	dd	ii	dd	ii	ii	dd	ii	ii	na	dd	li	ii	dd	ii	2
33	dd	dd	ii	dd	ii	na	dd	ii	dd	ii	ii	dd	id	id	dd	id	id	id	dd	id	1
34	na	na	id	na	ii	ii	na	ii	ii	ii	ii	dd	na	ii	na	id	id	na	na	na	9
35	dd	id	id	id	id	ii	ii	ii	ii	ii	ii	ii	ii	id	dd	dd	dd	dd	id	id	0
36	id	id	id	na	ii	ii	na	ii	id	ii	ii	id	ii	id	dd	id	id	id	ii	id	2
37	ii	na	dd	id	id	ii	dd	ii	id	ii	ii	id	ii	ii	na	id	li	dd	dd	id	2
38	dd	ii	id	dd	ii	ii	ii	ii	ii	ii	ii	ii	id	ii	na	id	li	id	dd	ii	1
39	ii	ii	ii	id	ii	ii	ii	ii	id	ii	ii	ii	ii	na	id	ii	li	id	dd	ii	1
40	dd	id	id	id	ii	ii	id	ii	ii	ii	ii	id	dd	id	dd	ii	li	id	id	id	0

ii = homozigoto inserção/inserção; id = heterozigoto; dd = homozigoto deleção/deleção; na = não amplificado.

Indiv.	L1	L2	L3	L4	L5	L7	L8	L9	L10	L11	L12	L13	L15	L16	L17	L18	L19	L20	L21	L22	Total NA
41	id	dd	id	id	ii	ii	id	ii	ii	ii	ii	dd	ii	ii	na	ii	ii	id	id	ii	1
42	id	id	id	id	ii	ii	id	ii	id	ii	ii	dd	ii	id	na	id	id	dd	dd	id	1
43	id	ii	ii	id	id	ii	na	ii	id	ii	ii	ii	ii	ii	dd	ii	id	dd	na	id	2
44	id	na	id	id	id	ii	ii	ii	ii	ii	ii	dd	ii	id	na	ii	ii	id	na	id	3
45	id	dd	id	id	na	ii	na	ii	na	ii	ii	na	na	ii	dd	id	ii	na	na	na	8
46	id	na	ii	ii	ii	ii	dd	ii	ii	ii	ii	id	ii	ii	dd	ii	id	dd	dd	id	1
47	dd	id	ii	na	ii	ii	ii	ii	dd	ii	ii	id	id	id	dd	ii	ii	dd	na	dd	2
48	id	id	ii	dd	ii	ii	id	ii	ii	ii	ii	id	ii	ii	dd	ii	dd	dd	dd	dd	0
49	id	id	id	ii	ii	ii	id	ii	id	ii	ii	id	ii	id	dd	ii	dd	id	dd	dd	0
50	dd	ii	id	id	ii	ii	id	ii	ii	ii	ii	id	id	ii	dd	ii	id	dd	id	ii	0
51	dd	dd	id	dd	id	ii	id	ii	ii	ii	ii	id	id	ii	id	ii	id	id	id	id	0
52	ii	ii	dd	id	id	ii	ii	ii	ii	ii	ii	dd	ii	ii	dd	ii	id	id	id	ii	0
53	dd	dd	id	id	id	ii	id	ii	ii	ii	ii	dd	id	ii	dd	dd	id	id	id	dd	0
54	ii	id	id	id	ii	ii	id	ii	ii	ii	ii	dd	id	dd	dd	ii	id	dd	id	id	0
55	dd	dd	id	ii	ii	ii	id	ii	ii	ii	ii	ii	ii	ii	dd	ii	ii	dd	id	ii	0
56	id	id	ii	dd	ii	ii	ii	ii	ii	ii	ii	dd	ii	ii	dd	dd	ii	id	id	id	0
57	dd	id	id	dd	ii	ii	ii	ii	ii	ii	ii	ii	id	ii	dd	id	dd	id	id	dd	0
58	ii	ii	id	dd	ii	ii	id	ii	ii	ii	ii	dd	ii	ii	dd	id	ii	dd	dd	id	0
59	ii	id	id	dd	ii	ii	id	ii	ii	ii	ii	dd	ii	ii	dd	id	id	dd	id	ii	0
60	id	ii	id	id	ii	ii	id	ii	dd	ii	ii	ii	ii	id	id	dd	id	ii	dd	ii	0

ii = homozigoto inserção/inserção; id = heterozigoto; dd = homozigoto deleção/deleção; na = não amplificado.

Indiv.	L1	L2	L3	L4	L5	L7	L8	L9	L10	L11	L12	L13	L15	L16	L17	L18	L19	L20	L21	L22	Total NA
61	dd	dd	id	id	ii	ii	ii	ii	id	ii	ii	dd	dd	ii	id	id	id	id	dd	ii	0
62	dd	id	ii	ii	ii	ii	id	ii	id	ii	ii	dd	ii	ii	id	ii	id	id	na	id	1
63	ii	dd	ii	ii	ii	ii	id	ii	ii	ii	ii	id	ii	id	dd	id	dd	dd	id	dd	0
64	id	ii	id	id	id	ii	ii	ii	id	ii	ii	dd	ii	id	id	id	id	id	dd	dd	0
65	ii	dd	ii	id	ii	ii	ii	ii	ii	ii	ii	id	ii	id	dd	id	dd	dd	ii	id	0
66	id	dd	ii	dd	ii	ii	id	ii	id	ii	ii	dd	id	id	dd	ii	dd	dd	id	dd	0
67	id	id	dd	dd	id	ii	ii	ii	ii	ii	ii	id	id	ii	id	id	ii	dd	dd	ii	0
68	id	ii	ii	id	ii	ii	ii	ii	id	ii	ii	dd	id	id	dd	id	id	id	dd	id	0
69	id	id	id	ii	ii	ii	id	ii	ii	ii	ii	dd	ii	id	id	ii	id	dd	dd	id	0
70	id	dd	dd	id	ii	ii	id	ii	ii	ii	ii	dd	ii	id	dd	id	id	dd	dd	id	0
71	ii	dd	ii	id	ii	ii	ii	ii	ii	ii	ii	id	ii	id	dd	id	dd	dd	ii	id	0
72	id	id	id	dd	ii	ii	id	ii	id	ii	ii	id	ii	ii	dd	ii	ii	dd	dd	ii	0
73	id	id	id	dd	ii	ii	na	ii	dd	ii	ii	ii	na	ii	dd	ii	id	dd	id	ii	2
74	id	id	ii	na	ii	ii	id	ii	ii	ii	ii	id	ii	ii	dd	ii	dd	dd	dd	id	1
75	dd	dd	id	id	id	ii	ii	ii	id	ii	ii	id	ii	id	dd	dd	dd	dd	id	id	0
76	dd	id	ii	dd	ii	ii	id	ii	ii	ii	ii	dd	ii	ii	dd	dd	id	dd	dd	dd	0
77	id	dd	ii	dd	ii	ii	dd	ii	dd	ii	ii	dd	ii	ii	dd	id	id	ii	dd	id	0
78	id	id	ii	id	ii	ii	dd	ii	id	ii	ii	id	ii	ii	dd	id	id	ii	id	id	0
79	id	ii	ii	ii	ii	ii	ii	ii	id	ii	ii	id	ii	id	na	id	id	dd	dd	id	1

ii = homozigoto inserção/inserção; id = heterozigoto; dd = homozigoto deleção/deleção; na = não amplificado.

Indiv.	L1	L2	L3	L4	L5	L7	L8	L9	L10	L11	L12	L13	L15	L16	L17	L18	L19	L20	L21	L22	Total NA
80	dd	id	dd	id	ii	ii	ii	ii	ii	ii	ii	dd	ii	ii	na	ii	ii	dd	dd	dd	1
81	ii	id	ii	dd	ii	ii	id	ii	id	ii	ii	dd	id	id	dd	ii	ii	dd	dd	dd	0
82	ii	id	id	dd	id	ii	ii	ii	id	ii	ii	id	ii	dd	dd	id	dd	dd	id	id	0
83	dd	id	id	dd	ii	ii	id	ii	ii	ii	ii	ii	ii	ii	dd	ii	ii	dd	dd	id	0
84	dd	ii	ii	id	ii	ii	ii	ii	ii	ii	ii	ii	ii	id	dd	ii	id	dd	id	id	0
85	id	id	ii	dd	id	ii	ii	ii	ii	ii	ii	dd	id	id	ii	ii	ii	dd	dd	id	0
86	id	ii	ii	na	ii	ii	ii	ii	ii	ii	ii	dd	id	ii	dd	ii	dd	ii	dd	id	1
87	id	id	ii	dd	ii	ii	id	ii	dd	ii	ii	dd	ii	ii	id	id	id	id	id	id	0
88	id	dd	ii	dd	ii	ii	dd	ii	dd	ii	ii	id	dd	ii	dd	id	ii	id	dd	id	0
89	id	dd	dd	dd	ii	na	na	ii	ii	ii	ii	na	ii	ii	na	id	id	dd	dd	na	5
90	ii	dd	dd	dd	ii	ii	ii	ii	ii	ii	ii	id	ii	ii	dd	ii	id	id	id	ii	0
91	id	dd	id	dd	ii	ii	dd	ii	dd	na	ii	id	ii	ii	id	id	ii	id	dd	dd	1
92	id	id	ii	dd	ii	ii	ii	ii	ii	ii	ii	id	ii	ii	id	dd	ii	id	id	id	0
93	id	id	id	dd	ii	ii	id	ii	ii	ii	ii	dd	ii	id	id	id	dd	id	id	ii	0
94	id	id	id	id	ii	ii	id	ii	id	ii	ii	id	ii	ii	dd	ii	dd	ii	dd	id	0
95	ii	ii	id	id	ii	ii	ii	ii	dd	na	ii	dd	dd	ii	dd	id	id	id	na	na	3
96	id	ii	dd	na	ii	ii	ii	ii	id	ii	ii	id	ii	ii	dd	id	dd	id	dd	id	1
97	id	ii	ii	id	ii	ii	ii	ii	ii	na	ii	id	ii	ii	dd	ii	id	id	id	ii	1
98	dd	id	id	na	ii	ii	ii	ii	ii	ii	ii	dd	ii	ii	dd	ii	dd	id	dd	dd	1
99	dd	id	id	dd	ii	ii	id	ii	ii	ii	ii	dd	ii	dd	dd	id	id	dd	dd	ii	0
100	ii	dd	ii	id	ii	ii	ii	ii	id	ii	ii	dd	ii	id	dd	ii	ii	id	id	id	0
Total NA	2	11	0	10	3	17	11	0	5	5	1	3	4	2	18	1	1	3	13	6	

ii = homozigoto inserção/inserção; id = heterozigoto; dd = homozigoto deleção/deleção; na = não amplificado.