

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
CAMPUS DE BOTUCATU
INSTITUTO DE BIOCIÊNCIAS**

Análise de casos de identificação humana em situações especiais.

Mirella Perruccio Soler

Orientadora: Prof^ª. Dra. Edna Sadayo Miazato Iwamura

Trabalho de Conclusão de Curso
apresentado como requisito para
a obtenção do grau de Bacharel
em Ciências Biológicas no
Instituto de Biociências da
Universidade Estadual Paulista
“Júlio de Mesquita Filho” -
Campus de Botucatu.

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO
DA INFORMAÇÃO.

DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - CAMPUS DE BOTUCATU - UNESP

BIBLIOTECÁRIA RESPONSÁVEL: SELMA MARIA DE JESUS

Soler, Mirella Perruccio.

Análise de casos de identificação humana em situações especiais / Mirella Perruccio Soler. – Botucatu : [s.n.], 2009.

Trabalho de conclusão (bacharelado – Ciências Biológicas) – Universidade Estadual Paulista, Instituto de Biociências de Botucatu, 2009

Orientadora: Edna Sadayo Miazato Iwamura

1. Genética 2. Programas de computador 3. Identificação humana

Palavras-chave: DNA; Programas computacionais; STR

Análise de casos de identificação humana em situações especiais.

Mirella Perruccio Soler¹ e Edna Sadayo Miazato Iwamura²

RESUMO - A utilização de programas computacionais específicos como ferramentas para análise estatística em identificação humana e de vínculos genéticos tem se tornado imprescindível, principalmente em casos de maior complexidade ou em grande escala. A análise de casos em que há genealogia incompleta gera resultados inconclusivos tornando os cálculos estatísticos mais laboriosos, consumindo tempo e pessoal qualificado. Os objetivos desse estudo são: o estudo simulado de casos reais de identificação humana e vínculos genéticos em situações especiais (meios irmãos na ausência dos pais, reconstrução), utilizando programas computacionais livres de cálculos estatísticos como ferramenta; a verificação do melhor programa livre de análise estatística disponível para utilização nos centros públicos brasileiros; e o levantamento estatístico dos casos reais brasileiros (vínculos genéticos, negatória, exclusão) para se obter uma visão da realidade brasileira em relação às suas relações familiares. Os principais métodos foram: análise das informações contidas nas fichas de registro e documentos anexados aos exames de vínculo genético dos participantes; e inserção nos softwares dos resultados de perfis genéticos obtidos de diferentes casos, de acordo com as especificações de uso. A amostra populacional dos casos analisados nesse trabalho incluiu 3.871 indivíduos. Em relação aos casos reais brasileiros, o exame de vínculo genético mais requisitado foi o de paternidade tipo trio; 68% dos casos incluíram o suposto pai como o pai biológico e 27% dos supostos pais estavam registrados na certidão de nascimento do filho e requisitaram o exame de paternidade para confirmá-la. A taxa de mutação encontrada variou de 0,4 a 2×10^{-3} *per locus per gamete per generation*. Após os estudos simulados, verificou-se que os melhores programas computacionais livres são os softwares PATCAM e FAMILIAS. O software PATCAM é adequado e eficaz para o uso em grande escala nos casos mais simples, enquanto software FAMILIAS é indicado para os casos complexos envolvendo relacionamentos familiares complicados ou com muitos indivíduos.

Palavras- Chave: DNA, STR, Programas Computacionais

¹ Instituto de Biociências, UNESP, Botucatu, SP.

² Laboratório de Patologia Molecular, Departamento de Patologia da Escola Paulista de Medicina – UNIFESP, São Paulo, SP.

1. Introdução

Recentemente, as investigações de vínculos genéticos e de identificação humana através da análise de DNA têm sido requisitadas para os mais diversos tipos de casos. Essas situações ocorrem tanto em casos familiares por questões sociais, financeiras ou de saúde; como em casos de acidentes em massa, os quais abrangem acidentes aéreos (TAM 1996, GOL 2006, TAM 2007), desastres naturais (Tsunami 2004) e conflitos armados (World Trade Center 2001, Madrid 2004) [1,2].

Entretanto, nos últimos anos, as análises de investigação tornaram-se mais complexas visto que a sociedade tem formado relações familiares cada vez mais amplas. Exemplos desses relacionamentos são os casos de meios irmãos na ausência dos pais, casos de reconstrução, casos de identificação através de tios e primos, entre outros [3].

No processo de investigação, após a obtenção do perfil genético dos envolvidos em laboratório, através da técnica da Reação em Cadeia da Polimerase (PCR), a estatística funciona como uma importante ferramenta para a obtenção de um resultado efetivo e conclusivo a respeito do relacionamento dos envolvidos [4].

A razão de verossimilhança dos relacionamentos pode ser calculada manualmente, todavia, em casos de maior complexidade (em que há genealogia incompleta ou em grande escala), os cálculos estatísticos tornam-se mais laboriosos, consumindo tempo e pessoal qualificado. Dessa forma, é imprescindível a utilização de programas computacionais específicos como ferramenta para esses cálculos [5].

Atualmente, existem programas computacionais de análise estatística tanto para casos simples, como para casos complexos. Porém, a maioria deles não é gratuita ou tem uma interface de uso não compatível ou “user-friendly”[5], o que torna difícil o acesso a esses aparatos. Assim, é inegável a necessidade da escolha de um software de análises estatísticas efetivo, ou seja, com alto grau de confiabilidade, que seja de fácil manuseio e de

rápida análise, além de ter acesso fácil pelos laboratórios [4,5].

O objetivo desse trabalho é o de encontrar e verificar o software livre (não pago) com o modo de análise de dados de vínculos genéticos e identificação humana mais eficiente e prático, que facilite e agilize as análises estatísticas de casos que possuem um maior grau de dificuldade. Para avaliar os programas, será feito o estudo simulado utilizando dados de casos encontrados na literatura ou de casos reais da população brasileira tendo como base os resultados já obtidos através da genotipagem.

Também será realizado o levantamento estatístico dos casos da população brasileira utilizados no estudo para ter-se uma visão da realidade brasileira em relação às suas relações familiares e os tipos de exames solicitados.

2. Materiais e Métodos

2.1. Para o levantamento estatístico dos casos:

Foi realizado o levantamento dos casos reais de vínculo genético dos anos de 2005 e 2006 de um laboratório particular da cidade de São Paulo. Nesse período, o laboratório recebeu exames com amostras de diferentes regiões do país (compondo 23 estados) para serem genotipadas. A análise estatística foi possível a partir das informações genéticas e pessoais obtidas dos dados contidos nas fichas de registro e documentos dos participantes dos exames. Essa amostra populacional incluiu 3.834 indivíduos, em 1.300 casos.

Para isto, utilizou-se uma planilha do Excel como banco de dados e foram registradas as seguintes variáveis:

- Número de identificação de cada exame
- Estado de origem materna
- Estado de origem do suposto pai
- Idade materna na data de concepção do filho do exame
- Idade materna no nascimento do filho do exame
- Idade do suposto pai na data de concepção do filho do exame

- Idade do suposto pai no nascimento do filho do exame
- Idade do filho na data em que o exame foi realizado
- Cor da pele declarada pela mãe
- Cor da pele declarada pelo suposto pai
- Se negatória (quando o pai registrado na certidão de nascimento participa do exame para confirmação de paternidade)
- Se inclusão (quando o suposto pai é o pai biológico)
- Se exclusão (quando o suposto pai não é o pai biológico)
- Se o pai registrado na certidão de nascimento não foi confirmado como biológico (quando o pai biológico é outro homem que não aquele da certidão)
- Se o exame foi realizado para confirmação de maternidade biológica
- Se o exame foi realizado para confirmação de paternidade
- Se o exame foi de reconstrução (utilizam-se dados genéticos dos avós)
- Se o exame foi Trio (utilizam-se dados genéticos do suposto pai, mãe e filho).
- Se o exame foi Duo (utilizam-se dados genéticos do filho e do suposto parente, sem a análise dos dados da mãe)
- Se houve mutação
- Que *locus* ocorreu à mutação
- Que alelo ocorreu à mutação
- Qual a origem da mutação (assume-se que o alelo parental mais similar em tamanho é o que originou a mutação no descendente)
- Quantas repetições a mais ou a menos ocorreram na mutação

2.2. Escolha dos programas computacionais:

A pesquisa da bibliografia referente às ferramentas computacionais (softwares) foi realizada através da busca por esse tema nos bancos de dados dos sites PubMed e ISI Web of Knowledge [6,7].

2.3. Análise de dados de identificação humana e vínculo genético utilizando softwares.

Foram inseridas, de acordo com as especificações de uso dos softwares, as seguintes informações:

- Perfis genéticos dos envolvidos, contidos nos exames dos laudos de filiação dos anos de 2005 e 2006, utilizando 13 ou 15 marcadores STR autossômicos.
- Perfis genéticos de artigos científicos e relatos de casos utilizando STR autossômicos. [8-12]

2.4. Tipos de casos analisados:

a) Casos teste 1: Trio Clássico

Para o estudo simulado dos casos de trio foram utilizados os dados genéticos de 40 supostos pais, mães e filhos dos exames de filiação.

b) Casos teste 2: Duo

Para o estudo simulado dos casos de duo foram utilizados os dados genéticos de 40 supostos pais e filhos dos exames de filiação.

c) Casos teste complexos 3: Parentesco

Para o estudo simulado dos casos de parentesco foram utilizados os dados genéticos de 20 casos de: meio-irmãos (6), irmãos (13) e de tio-sobrinho (1) dos exames de filiação.

d) Casos teste complexos 4: Reconstrução

Para o estudo simulado dos casos de reconstrução foram utilizados os dados genéticos de 4 avós/avôs maternos e/ou paternos, filhos, mães e supostos pais dos exames de filiação.

e) Casos teste complexos 5: Mutações e Alelos Nulos

Para esse estudo simulado foram utilizados os dados genéticos dos exames de filiação de 8 supostos pais, mães e filhos em que foram encontradas mutações, e para as simulações de alelos nulos, alguns dados genéticos foram retirados.

f) Casos teste complexos 6: Casos complexos variados

Para esse estudo simulado foram utilizados dados genéticos obtidos de 6 artigos científicos ou relatos de caso [8-12], como segue:

- I. Caso em que temos 4 dados genéticos: da criança 1, da mãe da criança 1, da criança 2 que é filha dessa mãe com o pai X, e da criança 3 que é filha do pai X com outra mulher. A criança 1 é filha do pai X? (7 STRs analisados). [8]
- II. Caso em que temos 5 dados genéticos: da criança 1, da mãe da criança 1, da criança 2, da criança 3, da mãe das crianças 2 e 3. Sendo o pai X o pai das crianças 2 e 3, a criança 1 também é filha desse homem? (7 STRs analisados). [8]
- III. Caso em que uma criança tem duas mutações de origem materna em dois *locus* diferentes. Temos os dados do suposto pai, do filho e da mãe. (14 STRs analisados). [9]
- IV. Caso em que temos 9 dados genéticos: de um feto, da mãe desse feto, do irmão dessa mãe, do pai 1 que é pai da mãe e do irmão, do avô que é pai do pai 1, e de quatro homens não relacionados com essa família. Quem é o pai do feto? (8 STRs analisados) [10]
- V. Caso em o homem X está desaparecido e foi encontrado o cadáver X. Queremos saber se o cadáver X é o do homem X que estava desaparecido. Temos os dados do cadáver X, de 3 irmãos do homem X, dois filhos do homem X, do sogro do homem X e de duas cunhadas do homem X, filhas de seu sogro. (9 STRs analisados). [11]

- VI. Caso em que foram encontradas ossadas de 5 crianças e 2 adultos em duas covas distintas. Com os dados genéticos desses ossos quer-se saber se estes indivíduos tinham relação familiar e qual era essa relação. (15 STRs analisados). [12]

Dessa forma os softwares puderam ser estudados e avaliados em relação a:

- Simplicidade da interface
- Tipos de casos que realizam
- Introdução das frequências populacionais
- Tempo que necessitam para realização dos casos
- Armazenamento de dados
- Desempenho na realização dos casos complexos relacionando o tipo de caso complexo em questão
- Desempenho em relação à análise estatística de cada tipo de caso
- Viabilidade para uso em grande escala

3. Resultados

3.1 Levantamentos estatísticos dos casos

78% dos indivíduos adultos participantes dos exames nasceram nas regiões Sul (19%) e Sudeste (59%) (Figura 1), nos estados de São Paulo com 47,73% e Paraná com 8,88% (Tabela 1), respectivamente.

A faixa de idade materna predominante quando o filho submetido ao exame nasceu foi de 20 a 24 anos, seguida da faixa de idade de 15 a 19 anos. A faixa de idade do suposto pai predominante nesse período foi também de 20 a 24 anos, seguida de 25 a 29 anos (Figura 2).

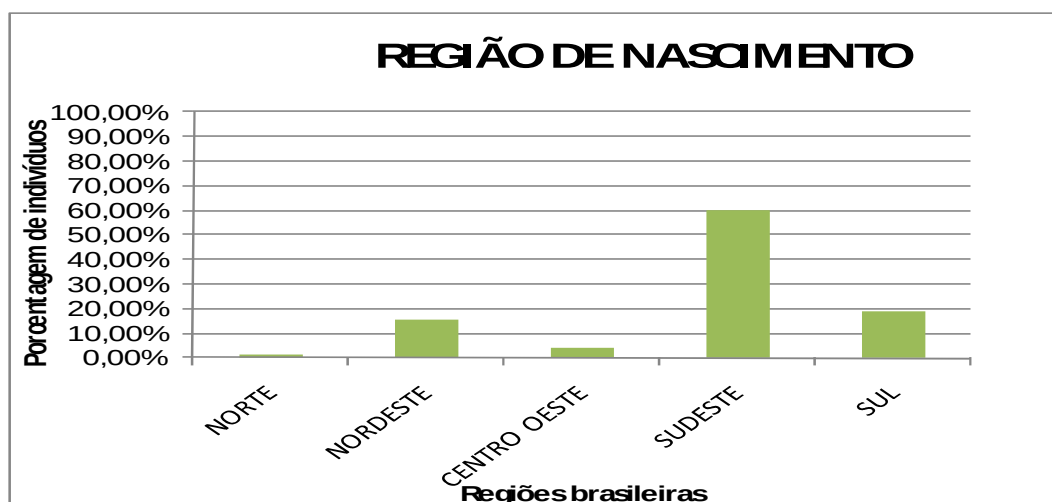


Figura 1. Gráfico da porcentagem de indivíduos por região brasileira.

Tabela 1. Porcentagem de indivíduos adultos do estudo por estado.

ESTADO	NÚMERO DE INDIVÍDUOS	% DE INDIVÍDUOS
SP	1241	47,73%
PR	231	8,88%
MG	168	6,46%
SC	154	5,92%
PE	110	4,23%
RS	102	3,92%
ES	82	3,15%
BA	80	3,08%
MS	79	3,04%
MA	70	2,69%
RJ	56	2,15%
PI	35	1,35%
CE	24	0,92%
PB	20	0,77%
AL	19	0,73%
SE	19	0,73%
DF	10	0,38%
GO	5	0,19%
PA	4	0,15%
MT	3	0,12%
RO	3	0,12%
AM	2	0,08%
RN	2	0,08%
Não declarado	81	3,12%
TOTAL	2.600	100,00%

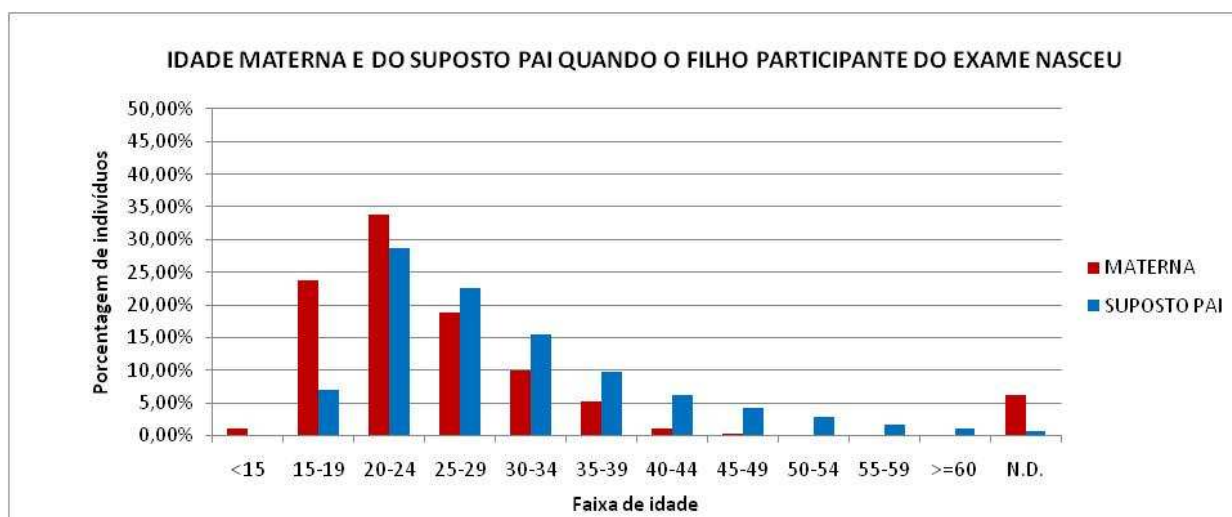


Figura 2. Gráfico da porcentagem quanto à faixa de idade materna e do suposto pai quando o filho submetido ao exame nasceu.

A idade mínima encontrada entre os participantes dos exames quando o filho participante do exame nasceu foi de 11 e de 15 anos para as mães e supostos pais, respectivamente. Os exames de filiação, em sua maioria (64%), foram aplicados em filhos que estavam na faixa de idade menor que 5 anos (Figura 3). Em 41% dos casos, o filho em questão não possuía registro de filiação paterna na certidão de nascimento.

Grande parte (74%) dos indivíduos adultos se auto-declarou ou foi declarada branca, com minoria parda, negra e oriental, respectivamente (Figura 4).

Em relação ao tipo de exame, em quase todos os exames foi requerido o exame de paternidade e na maioria destes participaram os supostos pais, filhos e mães (trio). Poucos casos se trataram de exames em busca da maternidade ou de reconstrução (Figura 5).

Quanto ao resultado do exame, cerca de 70% incluiu o suposto pai como o pai biológico e 27% dos pais que registraram os filhos na certidão de nascimento duvidaram dessa paternidade (negatória) (Figura 5).

Dos filhos que possuem registro paterno na certidão de nascimento, não tiveram esses pais confirmados como biológicos pelo exame 31,46% (Tabela 2).

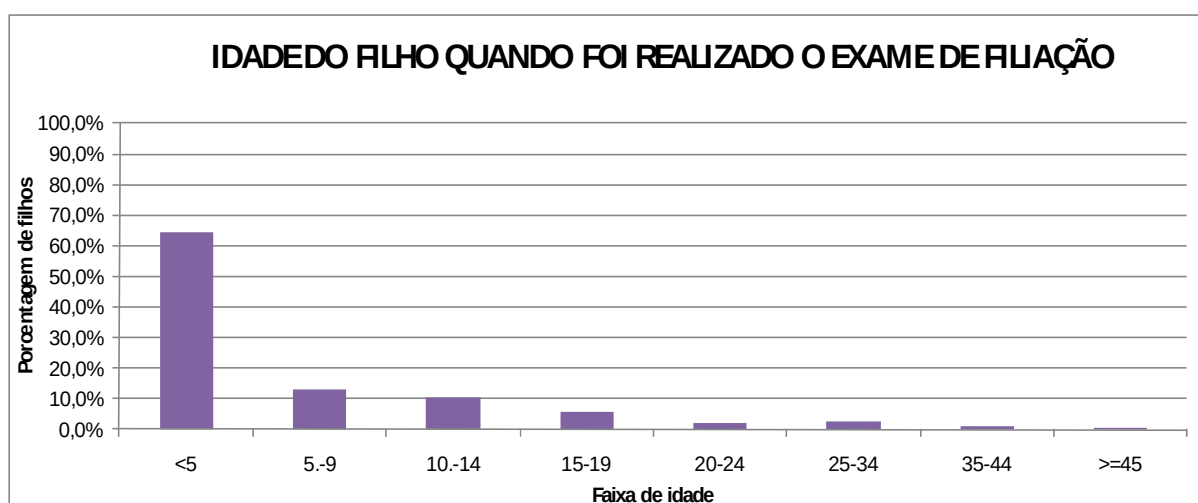


Figura 3. Gráfico da porcentagem quanto à faixa de idade em que o filho submetido ao exame tinha quando o realizou.

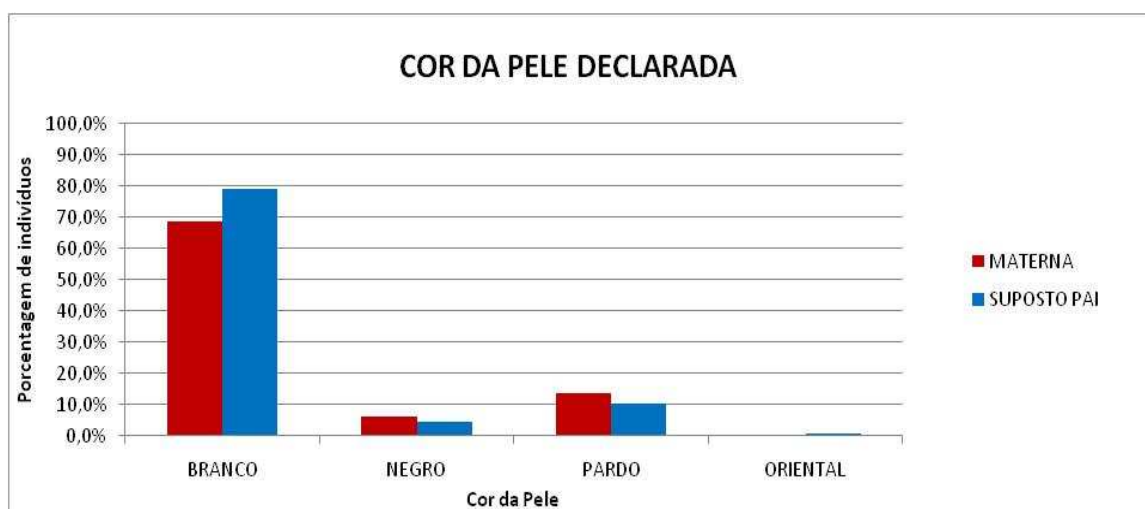


Figura 4. Gráfico da porcentagem de indivíduos adultos quanto à cor da pele declarada no exame.

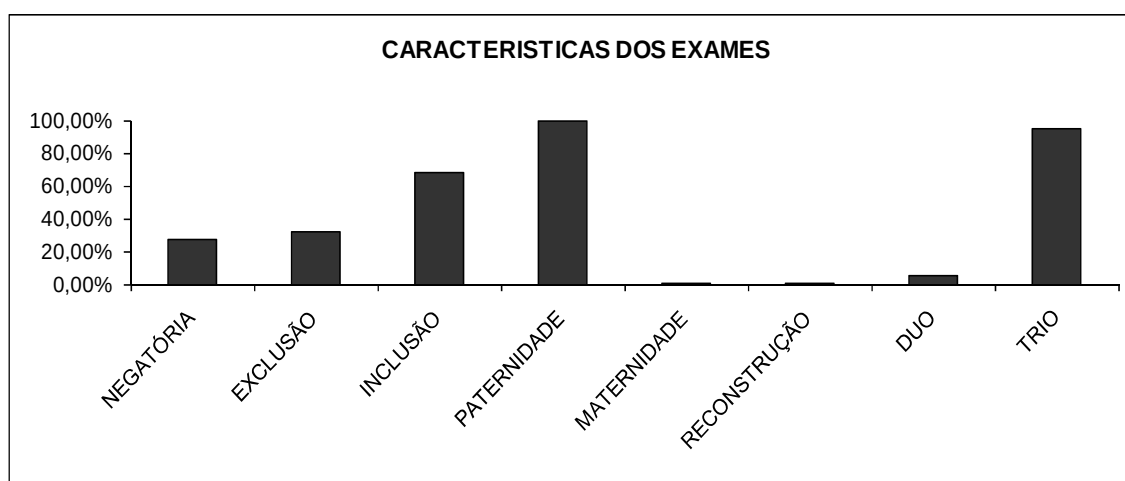


Figura 5. Gráfico da porcentagem dos exames em relação às suas características.

Tabela 2. Número total e porcentagem dos filhos que possuem registro paterno na certidão de nascimento e não tiveram esses pais confirmados como biológicos pelo exame.

Possui outro pai na certidão	% do total de exames	% sobre os exames do tipo negatória
191	14,69%	31,46%

Quanto às mutações, foram encontradas 23 mutações, sendo a maioria de origem paterna (Tabela 3). A maioria das mutações ocorreu nos *loci* vWA, FGA e 18S51 (Tabela 4) e se tratou do ganho ou perda de apenas 1 repetição(Tabela 5). A maioria das mutações

ocorreu em alelos com alto número de repetições (Tabela 6). A taxa de mutação encontrada varia de 0,4 a 2×10^{-3} *per locus per gamete per generation*. A idade paterna que cedeu a mutação, em 16 dos 20 casos, era maior ou igual a 30 anos.

Tabela 3. Porcentagem de mutações encontradas nos exames em relação a sua origem.

ORIGEM	MUTAÇÃO	%
PATERNA	20	1,54%
MATERNA	1	0,08%
INDEFINIDA	2	0,15%
GERAL	23	1,77%

Tabela 4. Porcentagem de mutações encontradas em cada *locus* dos STRs.

<i>locus</i>	NÚMERO	%
D21S11	2	8,70%
D8S1179	1	4,35%
vWA	5	21,74%
FGA	3	13,04%
D18S51	3	13,04%
D5S818	1	4,35%
D16S539	1	4,35%
CSF1PO	2	8,70%
D19S253	1	4,35%
D3S1358	1	4,35%
D19S433	1	4,35%
D13S317	1	4,35%
D7S820	1	4,35%
TOTAL	23	100,00%

Tabela 5. Distribuição das mutações por número de ganho/perda de repetições.

número de repetições	número de mutações
3	1
2	1
1	12
-1	9

Tabela 6. Relação dos alelos em que a mutação ocorreu com o intervalo de alelos existentes no respectivo *locus*.

alelo que mudou	intervalo de alelos do <i>locus</i>
11	6 até 14
23	18 até 32
14	5 até 15
15	11 até 21
22	18 até 32
14	7 até 15
14	9 até 24
23	18 até 32
20	11 até 21
15	8 até 17
12	6 até 14
9	9 até 18.2
12	8 até 15
16	11 até 21
21	9 até 24
12	9 até 17
30	24 até 38
16	11 até 20
16	11 até 21
9	7 até 14
19	9 até 24
32,2	25 até 36
15	11 até 21

3.2 Softwares

Durante a revisão bibliográfica, foram encontradas referências a respeito de softwares para análise estatística em vínculos genéticos e identificação humana e para a análise de vínculos genéticos em não humanos [13-24].

Dentre os utilizados em humanos destacam-se 12. Desses, apenas metade, ou seja, 6 programas são de livre acesso (não pagos), (Tabela 7). Dessa forma, esses 6 programas computacionais foram analisados e estudados mais detalhadamente através do download de seus manuais e programas:

a) O programa DNASTat, criado por um pesquisador polonês, contém manual apenas em polonês, sendo assim de difícil utilização e entendimento no Brasil. [15]

b) Os programas Patern e EasyPat criados por um pesquisador alemão, nos anos de 1996 e 2000 respectivamente, contém uma

interface ultrapassada, requerem um manuseio laborioso e demorado e usam o sistema operacional DOS. Assim, não estão de acordo com a proposta de obter um programa de análise estatística para utilização nos centros públicos brasileiros que realizam exames em grande escala e só utilizam sistemas operacionais atuais. [16]

c) O programa PatPCR criado por um pesquisador espanhol só pode ser utilizado com as frequências da população espanhola. Ele é bloqueado para outra base de dados de frequências populacionais, assim não pode ser utilizado para casos brasileiros. [17]

Assim, somente os programas computacionais FAMILIAS e PATCAM contém interface e funcionamento de acordo com a proposta do projeto e foram avaliados e comparados nesse estudo, a seguir.

Tabela 7. Softwares disponíveis. * Base julho de 2009

SOFTWARE	País	Ano que foi criado	Preço em	CASOS	CASOS	RECONSTRUÇÃO	CASOS MEIC	CASOS MAIS
			Dólares*	TRIO	DUO		IRMÃOS	COMPLEXOS
DNA Stat	Polônia	2007	Free	SIM	SIM	NÃO	NÃO	Quando há mutações e alelos nulos
DNA-View	E.U.A	2008	6.601,17	SIM	SIM	SIM	SIM	SIM
EasyDNA	China	2008	3.667,32	SIM	SIM	SIM	SIM	SIM
EasyPat	Alemanha	2000	Free	SIM	SIM	NÃO	SIM	NÃO
Familias	Holanda	2006	Free	SIM	SIM	SIM	SIM	SIM
GenoProof	Alemanha	2008	5.134,25	SIM	SIM	SIM	SIM	Quando há mutações
Genotype	Eslováquia	2008	5.733,47	SIM	SIM	SIM	SIM	SIM
Hugin	Dinamarca	2004	3.675,30	SIM	SIM	SIM	SIM	SIM
PATCAM	Espanha	2003	Free	SIM	SIM	SIM	SIM	Quando há mutações e alelos nulos
Patern	Alemanha	1996	Free	SIM	NÃO	NÃO	NÃO	NÃO
Paternity Index	Alemanha	2007	2.940,24	SIM	SIM	SIM	SIM	SIM
PatPCR	Espanha	2003	Free	SIM	SIM	SIM	NÃO	NÃO

Tanto nos exames do tipo trio, quanto nos exames do tipo duo que são exames mais básicos, a diferença de resultado entre os dois programas é mínima ou nula (Tabelas 8 e 9). Nos casos mais complexos de relação de parentesco, o programa PATCAM obteve, em 90% dos casos, resultados com valores maiores do que os obtidos com o programa FAMILIAS. O programa PATCAM, no momento do cálculo, mostra um quadro de relações possíveis entre os dois indivíduos, e suas devidas probabilidades (ex. meio-irmãos, primos, tio-

sobrinho, irmãos, avô-neto). Entretanto, o software PATCAM, em 35% dos casos, ao calcular as possibilidades e probabilidades de parentesco entre os dois indivíduos, mostrou uma relação de parentesco diferente da real como sendo a de maior probabilidade. Já no software FAMILIAS, que depende do usuário adicionar as relações (entre os dois indivíduos), sobre as quais ele quer saber a probabilidade, isso só ocorreu em 15% dos casos. (Tabelas 10 e 11).

Tabela 8. Resultados dos cálculos estatísticos dos trios nos softwares PATCAM e FAMILIAS.
PP: Probabilidade de Paternidade. IPC: Índice de paternidade cumulativo.

CASO	PATCAM PP	PATCAM IPC	FAMILIAS PP	FAMILIAS IPC
1	99,999921%	1.265.507	99,999921%	1.265.507
2	99,999920%	1.246.609	99,999920%	1.246.609
3	99,999785%	464.784	99,999785%	464.784
4	99,999999%	72.335.820	99,999999%	72.335.822
5	99,999986%	7.274.996	99,999986%	7.274.996
6	99,999991%	11.609.043	99,999991%	11.609.043
7	99,999380%	161.372	99,999380%	161.372
8	99,996742%	30.691	99,996742%	30.691
9	99,999728%	367.036	99,999728%	367.036
10	99,999981%	5.347.740	99,999981%	5.347.740
11	99,999997%	33.413.624	99,999997%	33.413.624
12	99,999709%	343.206	99,999709%	343.205
13	99,999417%	171.409	99,999417%	171.580
14	99,999815%	541.004	99,999815%	541.003
15	99,999935%	1.529.552	99,999935%	1.529.551
16	99,998370%	61.345	99,998372%	61.405
17	99,999989%	8.859.032	99,999989%	8.859.031
18	99,999948%	1.919.064	99,999948%	1.919.063
19	99,999963%	2.670.716	99,999963%	2.670.715
20	99,999979%	4.749.308	99,999979%	4.749.308
21	99,999996%	22.856.290	99,999996%	22.856.289
22	99,999979%	4.710.893	99,999979%	4.710.892
23	99,999994%	17.976.472	99,999994%	17.976.471
24	99,999978%	4.511.379	99,999978%	4.511.379
25	99,999997%	33.557.813	99,999997%	33.591.371
26	99,998523%	67.710	99,998523%	67.710
27	99,999998%	60.905.512	99,999998%	60.905.511
28	99,999232%	130.208	99,999232%	130.207
29	99,999373%	159.362	99,999373%	159.362
30	99,999693%	325.706	99,999693%	325.706
31	99,999615%	259.651	99,999615%	259.651
32	99,998591%	70.994	99,998591%	70.994
33	99,994614%	18.564	99,994614%	18.564
34	99,999842%	634.299	99,999842%	634.298
35	99,999997%	32.472.462	99,999997%	32.472.462
36	99,999952%	2.072.109	99,999952%	2.072.108
37	99,999949%	1.976.732	99,999949%	1.976.731
38	99,999621%	263.869	99,999621%	263.868
39	100,000000%	1.043.579.787	100,000000%	1.043.579.849
40	0,999999838	6189785	0,999999838	6189785

Tabela 9. Resultados dos cálculos estatísticos dos duos nos softwares PATCAM e FAMILIAS.
PP: Probabilidade de Paternidade. IPC: Índice de paternidade cumulativo.

CASO	PATCAM PP	PATCAM IPC	FAMILIAS PP	FAMILIAS IPC
1	99,933959%	1513	99,933959%	1513
2	99,998542%	68586	99,998542%	68586
3	99,980845%	5220	99,980845%	5219
4	99,999401%	167067	99,999401%	167066
5	99,999024%	102458	99,999024%	102458
6	99,994249%	17388	99,994249%	17388
7	99,970773%	3421	99,970773%	3420
8	99,807108%	517	99,807108%	517
9	99,917161%	1206	99,917161%	1206
10	99,999866%	745964	99,999866%	745963
11	99,999714%	349667	99,999714%	349666
12	99,973766%	3811	99,973766%	3810
13	99,978190%	4584	99,978212%	4588
14	99,947343%	1898	99,947343%	1898
15	99,805310%	513	99,805310%	512
16	99,962670%	2678	99,962707%	2680
17	99,998761%	80720	99,998761%	80720
18	99,992911%	14105	99,992911%	14104
19	99,993032%	14350	99,993032%	14349
20	99,997796%	45377	99,997796%	45376
21	99,996613%	29521	99,996613%	29521
22	99,989189%	9249	99,989189%	9249
23	99,998811%	84123	99,998957%	84122
24	99,986288%	7287	99,986278%	7286
25	99,999508%	203242	99,999508%	203445
26	99,991842%	12256	99,991842%	12256
27	99,999686%	318852	99,999686%	318851
28	99,612561%	257	99,612561%	257
29	99,856912%	698	99,856912%	697
30	99,981528%	5413	99,981528%	5412
31	99,988951%	9049	99,988951%	9049
32	99,904978%	1051	99,904978%	1051
33	99,810313%	526	99,810313%	526
34	99,957265%	2339	99,957265%	2338
35	99,998073%	51903	99,998073%	51902
36	99,985567%	6928	99,985567%	6927
37	99,994215%	17286	99,994215%	17285
38	99,987448%	7966	99,987448%	7965
39	99,999979%	4690431	99,999979%	4690430
40	99,989510%	9532	99,989510%	9532

Tabela 10. Resultados dos cálculos estatísticos de diferentes parentescos nos softwares PATCAM e FAMILIAS. PS. Probabilidade de parentesco. CSI: Índice de irmãos completos. HSI Índice de meios irmãos.

CASO	TIPO	PATCAM PS	PATCAM CSI/HSI	FAMILIAS PS	FAMILIAS CSI/HSI
1	MEIO IRMÃO	81,442330%	4	47,590193%	1
2	MEIO IRMÃO	91,122660%	10	69,278699%	2
3	MEIO IRMÃO	98,417190%	62	66,331938%	2
4	MEIO IRMÃO	98,417190%	62	99,340061%	151
5	MEIO IRMÃO	74,315560%	3	53,051654%	1
6	MEIO IRMÃO	62,937170%	2	32,433393%	0,48
7	IRMÃO COMPLETO	99,996960%	32.976	93,514967%	14
8	IRMÃO COMPLETO	99,949750%	1.989	92,277087%	12
9	IRMÃO COMPLETO	99,996400%	27.853	98,794809%	82
10	IRMÃO COMPLETO	99,965470%	2.895	94,435537%	17
11	IRMÃO COMPLETO	99,999930%	1.457.318	67,053850%	2
12	IRMÃO COMPLETO	8,993000%	1	3,576820%	0,04
13	IRMÃO COMPLETO	99,999410%	171.710	97,011768%	32
14	IRMÃO COMPLETO	98,348400%	59	55,362644%	1
15	IRMÃO COMPLETO	99,999990%	3.000.000.000	99,999960%	24.808.007
16	IRMÃO COMPLETO	99,999990%	253.940.461	99,968526%	3.176
17	IRMÃO COMPLETO	69,585100%	2	52,715071%	1
18	IRMÃO COMPLETO	51,846540%	1	11,575718%	0,13
19	IRMÃO COMPLETO	99,994630%	18.655	98,567093%	69
20	TIO-SOBRINHO	69,280090%	2	71,440931%	3

Tabela 11. Relação fornecida pelos softwares PATCAM e FAMILIAS como sendo a de maior probabilidade.

CASO	TIPO	PATCAM	FAMILIAS
1	MEIO IRMÃO	MAIOR PARA PRIMOS	
2	MEIO IRMÃO		
3	MEIO IRMÃO		
4	MEIO IRMÃO		
5	MEIO IRMÃO	MAIOR PARA PRIMOS	
6	MEIO IRMÃO	MAIOR PARA PRIMOS	MAIOR PARA PRIMOS
7	IRMÃO COMPLETO		
8	IRMÃO COMPLETO		
9	IRMÃO COMPLETO		
10	IRMÃO COMPLETO		
11	IRMÃO COMPLETO		
12	IRMÃO COMPLETO	MAIOR PARA PRIMOS E MEIO IRMÃOS	MAIOR PARA PRIMOS
13	IRMÃO COMPLETO		
14	IRMÃO COMPLETO		
15	IRMÃO COMPLETO		
16	IRMÃO COMPLETO		
17	IRMÃO COMPLETO	MAIOR PARA PRIMOS	
18	IRMÃO COMPLETO	MAIOR PARA PRIMOS E MEIO IRMÃOS	MAIOR PARA MEIO IRMÃOS
19	IRMÃO COMPLETO		
20	TIO-SOBRINHO	MAIOR PARA PRIMOS	

Nos exames complexos do tipo reconstrução, em 3 dos 4 casos, o programa FAMILIAS obteve uma probabilidade maior, sendo que o programa PATCAM não resolveu o

caso 2 em que só havia os dados genéticos da mãe, do filho e da avó paterna (Tabela 12).

Quando testados os casos em que foram encontradas mutações, em 87% dos casos, o

programa PATCAM forneceu valores de probabilidade maiores do que as fornecidas pelo programa FAMILIAS (Tabela 13). Nos casos de alelos nulos, o software FAMILIAS não resolveu esse tipo de caso, visto que não há como deixarmos um alelo em branco, ao gravar os dados genéticos para um marcador. Nas simulações dos casos de maior grau de dificuldade, o programa PATCAM não foi capaz de resolver nenhum.

Porém, o programa FAMILIAS resolveu quase todos os casos e forneceu probabilidades com valores altos para os específicos relacionamentos. Para o caso III, a única probabilidade que o programa calculou foi a do suposto pai ser o pai verdadeiro (Tabela 14), ele não incluiu a mãe no cálculo. Assim, para esse caso de duas mutações em dois *locus* diferentes não foi possível com o programa FAMILIAS obter o resultado.

Tabela 12. Resultados dos cálculos estatísticos dos casos de reconstrução nos softwares PATCAM e FAMILIAS. PP: Probabilidade de Paternidade. IPC: Índice de paternidade cumulativo.

CASO	PATCAM PP	PATCAM IPC	FAMILIAS PP	FAMILIAS IPC
1	99,993284%	14.889	99,999809%	523.088
2	NÃO FAZ	NÃO FAZ	56,609191%	1
3	99,230289%	128	99,754744%	407
4	99,967074%	3.036	99,964707%	2.832

Tabela 13. Resultados dos cálculos estatísticos dos casos de mutação nos softwares PATCAM e FAMILIAS. PP: Probabilidade de Paternidade. IPC: Índice de paternidade cumulativo.

CASO	PATCAM PP	PATCAM IPC	FAMILIAS PP	FAMILIAS IPC
1	99,999903%	1.031.216	99,821525%	559
2	99,999989%	8.773.577	99,976332%	4.224
3	99,997097%	34.442	99,998818%	84.636
4	99,999987%	7.942.891	99,999210%	126.641
5	99,999788%	472.628	99,597920%	248
6	100,000000%	5.540.820.997	99,999996%	27.338.898
7	99,999999%	82.691.203	99,999562%	228.448
8	99,999758%	412.396	99,426013%	173

Tabela 14. Resultados dos casos complexos variados no software FAMILIAS

CASO	FAMILIAS Probabilidade	FAMILIAS Índice
I	99,9646943763374%	2.831
II	77,6222836898552%	3
III	99,9962751897130%	26.846
IV	99,9999999999986%	96.889.010.407
V	98,3778019827098%	60
VI	99,9999349717365%	1.537.792

Ao compararmos os dois softwares, o programa PATCAM realiza casos básicos e complexos, que envolvem até 6 indivíduos (reconstrução), enquanto o programa FAMILIAS realiza casos simples e de maior dificuldade em que estão envolvidos mais de 6

indivíduos. O software PATCAM, na maioria dos casos, só permite calcular probabilidades de relacionamentos entre indivíduos que possuem um perfil genético. O programa FAMILIAS dá ao usuário a possibilidade de determinar a relação familiar mais provável entre um número

grande de indivíduos, podendo ser necessário adicionar pessoas para as quais não se tem um perfil genético. Por exemplo: para descrever que uma mulher tem 3 filhos com o mesmo homem, é necessário incluir este na genealogia, mesmo se os dados moleculares estiverem indisponíveis.

As áreas de trabalho dos dois programas, para os diferentes tipos de casos, possuem uma interface muito fácil de ser utilizada. Para o software PATCAM, o formato em planilhas de Excel facilita a inserção e a leitura de dados. Na mesma planilha há a opção de escolher, dentre todos os casos gravados, o caso que se quer o resultado, podendo assim visualizar rapidamente os resultados de qualquer caso que já teve seus dados genéticos inseridos na base de dados. O software FAMILIAS tem uma interface mais organizada, com ícones que são figuras representativas das variadas funções. Todavia, só se pode visualizar um caso por vez e, para interpretar os resultados, é necessário um conhecimento mais detalhado do programa.

Em relação à introdução e utilização de frequências populacionais, nos dois softwares o processo demanda tempo, já que as frequências precisam ser inseridas manualmente no local apropriado e os alelos com uma casa decimal precisam ser inseridos com um zero na frente. Esse método aumenta os riscos de um erro no momento da digitação.

No PATCAM, podem ser inseridos dados de muitas populações, uma em cada planilha da área de trabalho. Essa característica é uma grande vantagem, visto que existem frequências populacionais diferentes nas regiões brasileiras [25-30], sendo possível uma flexibilidade na escolha da população a ser utilizada. Outra vantagem é a possibilidade de inserir nas linhas “other”, ou nas que estão em branco, novos nomes de alelos e suas frequências. Dessa maneira, a atualização dos dados das frequências populacionais é facilitada quando são estabelecidos novos alelos. No programa FAMILIAS, os marcadores e suas frequências alélicas são inseridos independentemente da tabela de frequência a

que pertencem, assim, deve-se identificar de modo uniforme os marcadores pertencentes a uma mesma tabela de frequência. Esse modo de inserção de frequências populacionais pode causar confusão ou erro no momento de introduzir os dados genéticos, fazendo com que marcadores de diferentes populações participem do mesmo caso.

Na função de armazenamento de dados, os dados genéticos no programa PATCAM são inseridos em uma Base de dados que está relacionada com as áreas de trabalho específicas a cada tipo de caso. Esses valores são gravados e armazenados, permitindo futuras consultas nos casos realizados anteriormente. No programa FAMILIAS, cada caso é gravado separadamente, sendo que a cada novo caso, as frequências alélicas e os marcadores precisam ser inseridos novamente.

Em relação aos cálculos estatísticos, os dois programas não disponibilizam as fórmulas algébricas utilizadas para os cálculos. O PATCAM permite que sejam calculadas as frequências alélicas mínimas. O programa FAMILIAS utiliza parâmetros para genealogias que possuem várias gerações, relacionamentos incestuosos e genealogias envolvendo comportamento promíscuo; além de disponibilizar variados modelos de mutação, diferenciando as taxas de mutação femininas e masculinas. Isso permite que ele resolva os mais variados casos complexos.

Nos casos em que há exclusão, o programa FAMILIAS não calcula nenhuma probabilidade, enquanto que o programa PATCAM, ao calcular as probabilidades nesses casos, considera os índices de paternidade dos marcadores que não foram excluídos.

Os dois softwares resolvem os casos com rapidez. O tempo necessário entre a introdução dos dados genéticos (sem contar a introdução das frequências alélicas) até o cálculo dos resultados para o PATCAM é de 7 minutos (casos mais básicos) até 10 minutos (casos de reconstrução). Para o programa FAMILIAS o tempo varia de 5 minutos (casos mais básicos) até 15 minutos (casos envolvendo mais de 6 indivíduos) (Tabela 15).

Tabela 15. Comparação entre os dois softwares

Parâmetro	PATCAM	FAMILIAS
Simplicidade da interface	A área de trabalho dos diferentes tipos de casos possui uma estrutura muito simples de ser utilizada. O formato em planilhas de Excel facilita a inserção e a leitura de dados. Há a opção de, numa mesma planilha, escolher e mudar de caso rapidamente. Os resultados podem ser analisados e interpretados rapidamente.	Tem uma interface organizada, “clean”, fácil de ser manipulada, com ferramentas que são diferenciadas através de desenhos representativos. Os dados genéticos são inseridos facilmente. Para interpretar os resultados é necessário um conhecimento mais detalhado do programa. Apenas um caso pode ser visualizado no programa.
Introdução das frequências populacionais	As frequências precisam ser inseridas manualmente na planilha do Excel. Podem ser guardados dados de muitas populações, em diferentes planilhas da área de trabalho. Há a possibilidade de inserir nas linhas que estão em branco, novos nomes de alelos e suas frequências.	Insere-se manualmente, para cada marcador, suas frequências específicas. Os dados de diferentes populações são armazenados numa mesma plataforma, assim, sua diferenciação depende de uma nomeação uniforme para marcadores de uma mesma população.
Tipos de casos que realizam	Trios, Duos, Parentesco, Reconstrução (quando há os dados de pelo menos 2 avós), Mutação e Alelos Nulos.	Trios, Duos, Parentesco, Reconstrução, Mutação e Casos mais complicados envolvendo mais de 6 indivíduos.
Tempo que necessitam para realização dos casos	De 7 minutos (casos mais básicos) até 10 minutos (casos de reconstrução).	De 5 minutos (casos mais básicos) até 15 minutos (casos complexos ou envolvendo mais de 6 indivíduos).
Armazenamento de dados	Os dados genéticos de todos os casos são inseridos em uma base de dados relacionada com as áreas de trabalho específicas. Os valores são gravados e armazenados, permitindo futuras consultas.	Só pode ser gravado, ou aberto para consulta, um caso por vez. Os dados gravados podem ser re consultados. A cada caso novo devem-se inserir novamente as frequências populacionais relevantes.
Desempenho em relação à análise estatística de cada tipo de caso	Não disponibiliza as fórmulas algébricas utilizadas para os cálculos. Permite cálculo das frequências alélicas mínimas. Não disponibiliza variados modelos de mutação. Não diferencia as taxas de mutação femininas e masculinas. Não utiliza parâmetros para genealogias que possuem várias gerações, relacionamentos incestuosos e genealogias envolvendo comportamento promíscuo. Utiliza coeficiente de coancestralidade.	Não disponibiliza as fórmulas algébricas utilizadas para os cálculos. Não realiza cálculo das frequências alélicas mínimas. Disponibiliza variados modelos de mutação. Diferencia as taxas de mutação femininas e masculinas. Utiliza parâmetros para genealogias que possuem várias gerações, relacionamentos incestuosos e genealogias envolvendo comportamento promíscuo. Utiliza coeficiente de coancestralidade.

No software FAMILIAS, o usuário pode fornecer ao programa as prováveis relações entre os indivíduos e calcular a relação mais provável dentre essas possibilidades. Se não houver nenhuma informação prévia a respeito dessa relação, o programa acaba fornecendo centenas ou até milhares de relações entre os indivíduos.

Utilizar a ferramenta das *prioris*, informar a idade e o gênero dos indivíduos, e identificar se alguma pessoa é criança (ainda não tem filhos), são informações que auxiliam a desconsiderar algumas alternativas improváveis e limitar o número de genealogias. Uma das ferramentas do programa permite a introdução das relações previamente conhecidas, sendo outro mecanismo limitante do número de relações geradas.

4. Discussão e Conclusão

Exceto nos casos retirados da literatura, a amostragem avaliada compreende predominantemente a população do Sudeste do Brasil (78%).

As análises revelam que grande parte das mulheres tornou-se mãe entre 20 e 24 anos de idade (34%), seguido da faixa de idade de 15 a 19 anos (24%), isto é, eram jovens no nascimento do filho para o qual requisitaram o exame de filiação. Figura 2.

A idade muito precoce em que o filho se encontrava no momento do exame de paternidade (mais de 64% dos filhos tinham menos que 5 anos de idade) mostra que essa busca se faz nos primeiros anos de vida. Isso possivelmente ocorre para que a criança possa receber o direito a filiação, uma vez que 66% dos filhos menores que 5 anos possuíam como documento de identificação apenas a Declaração de Nascido Vivo. A segunda maior faixa de idade do filho no momento do exame foi de 5 a 9 anos (13%). Isso pode estar relacionado à quando a criança está em idade escolar e questiona a respeito de sua genealogia.

A porcentagem de declarados brancos (74%) predominou sobre as outras. Esse dado é esperado, considerando que o exame de filiação

ainda é caro para as classes sociais mais baixas, e realizado em umas das cidades mais desenvolvidas do país.

O vínculo genético requerido em quase todos os exames foi o de paternidade e ainda é muito pequena a proporção de exames de reconstrução. Esse fato se deve ao maior prazo necessário à elaboração desse tipo de exame, ao custo mais elevado, e porque esse tipo de solicitação é normalmente requerido em casos cíveis em processos judiciais. Os exames com resultado de exclusão da paternidade representaram um índice considerável (32%), mostrando as dúvidas geradas pela própria dinâmica da sociedade moderna.

Em relação aos exames considerados negatória, ou seja, em que o suposto pai participante do exame constava no Registro de Nascimento ou Identidade; não foi confirmada a paternidade em aproximadamente 32% dos casos.

As mutações encontradas mantiveram-se dentro das taxas encontradas na literatura [31-38]. Os *loci* que mais mutaram FGA e 18S51, são também os mais polimórficos dentre os STR do estudo. A maioria (91%) das mutações teve a modificação de uma ou poucas repetições em relação à sequência original assim como já descrito na literatura [32-38]. 61% das mutações ocorreram em alelos com alto número de repetições, visto que a estabilidade das sequências de repetições em um alelo é inversamente correlacionada com o número de repetições que ela contém, isto é, a taxa de mutação aumenta com o tamanho do alelo [31,35]. O fato de 87% das mutações possuírem origem paterna condiz com a premissa de que as taxas de mutações em STR humanos variam com o gênero. O sexo masculino muta mais freqüentemente do que o feminino. Essas maiores taxas de mutação em homens provêm do fato de que eles contêm um número muito maior de células em mitose durante a espermatogênese do que as mulheres contêm na oogênese, se for assumido que a maioria das mutações é proveniente da replicação [34,36]. A idade paterna que cedeu a mutação, em 80% dos casos, era maior ou igual a 30 anos. Esse resultado está de acordo com

[36] que mostra que as taxas de mutação estão relacionadas à idade do homem na reprodução, visto o acúmulo do número de divisões celulares mitóticas na linhagem germinativa masculina que aumenta até a fase adulta.

Em relação à avaliação dos softwares, nos exames mais básicos tipo trio e duo os dois programas resolveram os casos rapidamente e a diferença de resultado entre os programas, nos 40 casos testados, não chegou a 1%. O programa PATCAM, na área de trabalho dos trios apresenta a possibilidade de analisar e comparar o índice de paternidade para outros tipos de relação entre o suposto pai e o filho em questão como: o suposto pai é o irmão do verdadeiro pai, o suposto pai é meio irmão ou avô do verdadeiro pai, o suposto pai é sobrinho do verdadeiro pai e o suposto pai é primo do verdadeiro pai. Esse mecanismo, além de permitir saber se o suposto pai é o biológico ou não, permite saber qual é a hipótese com maior índice de paternidade dentre diferentes tipos de relação de parentesco.

Nos exames mais complexos de parentesco, o programa PATCAM resolveu rapidamente necessitando, apenas, da inserção dos dados genéticos dos dois indivíduos. Já o programa FAMILIAS requer que as relações possíveis entre os dois indivíduos sejam fornecidas pelo usuário após uma análise prévia do caso. Ele requer também a adição de mais 5 indivíduos (sem perfil genético) para resolver efetivamente o caso. Em relação aos resultados desses casos, apesar do programa PATCAM fornecer a maioria dos resultados com probabilidades maiores do que o FAMILIAS, ao calcular as possibilidades e probabilidades de parentesco entre os dois indivíduos, ele apresentou outros casos com uma relação de parentesco diferente da real como sendo a de maior probabilidade. Como a literatura mostra [39-44], os cálculos para probabilidades de parentesco tem um alto índice de erro. Assim, para esses tipos de exame, os softwares não devem ser as únicas ferramentas. Para concluir efetivamente esses casos, a literatura sugere o acréscimo de outras ferramentas moleculares, assim a relação de parentesco poderá ser confirmada [39-44].

Todos os exames de reconstrução só foram resolvidos no programa FAMILIAS, visto que o programa PATCAM não resolveu o caso em que só há um dos avós. Não há como realizar casos em que só há uma avó ou um avô, apenas quando existem dados dos dois juntos. As probabilidades fornecidas também foram maiores no FAMILIAS.

Os exames tipo trio com mutação foram resolvidos pelos dois softwares, todavia no PATCAM esses casos só são resolvidos, e apresentam um índice de paternidade acumulado alto, quando o programa ignora o cálculo do *locus* que mostra a exclusão (mutação) e calcula o índice de paternidade considerando só os *loci* compatíveis. Ao utilizar a planilha chamada “Mutation” para calcular o índice de paternidade do *locus* que mostrou a exclusão e multiplicá-lo aos índices dos outros *loci*, os índices de paternidade acumulados dos casos diminuíram bastante, sendo que alguns ficaram abaixo de 300. No software FAMILIAS, 50% dos casos tiveram probabilidades de paternidade abaixo de 99,99% utilizando-se o modelo de mutação “Equal model”, o qual assume que a probabilidade de ocorrer uma mutação de um alelo a outro independe da frequência desse alelo na população ou da distância, em número de repetições, entre o alelo mutado e o alelo original.

Os casos em que foram simulados alelos nulos só foram resolvidos pelo PATCAM. O programa FAMILIAS não permite a retirada de um alelo no marcador. Dessa forma, quando existirem alelos nulos, o marcador em que ele ocorreu deverá ser ignorado para o cálculo.

Em relação aos casos com alto grau de dificuldade retirados da literatura, somente o software FAMILIAS conseguiu os resolver. Nos casos I, II, V as relações previamente conhecidas entre os indivíduos foram inseridas e o programa gerou as probabilidades solucionando o caso. No caso IV não foram estabelecidas relações previamente conhecidas. O usuário precisa analisar previamente o caso e adicionar relações possíveis. Se for o próprio programa que gerar as relações, elas chegam a milhares. Após o cálculo, o programa forneceu

a maior probabilidade resolvendo o caso. No caso VI também não foram estabelecidas relações previamente conhecidas. O programa gerou 1.024 possibilidades e calculou suas probabilidades, resolvendo o caso. Todavia, no caso III o programa não conseguiu mostrar a relação entre a mãe e o filho, visto que ela cedeu para o filho duas mutações em diferentes *loci*. A mãe foi excluída, o suposto pai foi incluído e sua probabilidade de paternidade calculada.

As *prioris* foram selecionadas dependendo do tipo de caso, conforme as instruções do programa.

Portanto, após a avaliação realizada nos dois programas, percebe-se que o programa computacional PATCAM é adequado e eficaz para o uso em grande escala nos casos de trios, duos e reconstrução simples, visto que suas características unem simplicidade de interface, rapidez na utilização e armazenamento de muitos dados. Em casos de maior complexidade como ausência de familiares, desastres em massa e vínculos familiares distantes, o programa não tem aplicabilidade.

O programa computacional FAMILIAS é indicado para os casos envolvendo relações familiares complexas ou muitos indivíduos, entretanto deve ser usado com cautela nos casos de parentesco. Suas ferramentas permitem que todos os tipos de caso sejam resolvidos. O software considera todas as possibilidades envolvendo os indivíduos relacionados, e por isso é de grande ajuda em casos especiais como os de incesto. Além de realizar casos com dados genéticos de marcadores STR autossômicos, o programa FAMILIAS pode ser utilizado com dados genéticos de STR-X, STR-Y ou de SNPs.

Em relação aos casos de acidente em massa, como em desastres aéreos que envolvem quase 200 mortos, são realizadas as tipagens genéticas tanto dos restos biológicos encontrados, quanto dos familiares das vítimas. Dessa maneira, um software que manipule e resolva um caso envolvendo tantos perfis genéticos é necessário. O programa FAMILIAS só consegue lidar com um caso como esse, se o separarmos em muitos exames de identificação menores. Para isso, devemos inserir no software

apenas um perfil genético de material biológico encontrado junto com os perfis genéticos dos familiares das vítimas, e não todos os perfis genéticos de materiais biológicos encontrados junto com todos os perfis de familiares das vítimas.

Como consideração final, não deve ser negligenciado o fato de que as análises laboratoriais utilizando ferramentas moleculares e computacionais, assim como a qualificação profissional constante, constituem hoje em dia uma parte importante e imprescindível no processo de investigação humana.

Agradecimentos

Agradeço ao grupo da Patologia Molecular da UNIFESP pelos ensinamentos, ajudas e pela boa vontade durante todo o período; ao Mark pela paciência; à Prof^a. Dra. Edna Miazato Sadayo Iwamura pela confiança; e a todos os que me auxiliaram nesse projeto.

Referências Bibliográficas

- [1]Alonso A, Martin P, Albarrán C, Garcia P, Fernandez de Simon L, Jesús Iturralde M, Fernández-Rodriguez A, Atienza I, Capilla J, García-Hirschfeld J, Martinez P, Vallejo G, García O, García E, Real P, Alvarez D, León A, Sancho M. Challenges of DNA profiling in mass disaster investigations. Croatian Medical Journal. 2005; 46(4): 540-548.
- [2]Brenner C.H.. Symbolic Kinship Program. Genetics. 1997; 145: 535-542
- [3]Allen R.W., Fu J, Reid TM, Baird M. Considerations for the interpretation of STR results in cases of questioned half-sibship. Transfusion. 2007; 47(3): 515-9.
- [4]Drabek J. Validation of software for calculating the likelihood ratio for parentage and kinship. Forensic Science International. 2009; 3(2): 112-118
- [5] Riancho J. A. , Zarrabeitia M. T. A Windows-based software for common paternity and sibling analyses. Forensic Science International. 2003; 135: 232-234.
- [6] www.isiknowledge.com/. Acesso em: março, 2009.
- [7]<http://www.ncbi.nlm.nih.gov/pubmed/>. Acesso em: março, 2009.
- [8]Narkuti V, Vellanki RN, Anubrolu N, Doddapaneni KK, Gandhi Kaza PC, Mangamoori LN. Single and double incompatibility at vWA and D8S1179/D21S11

loci between mother and child: implications in kinship analysis. *Clinica Chimica Acta*. 2008; 395(1-2): 162-5.

[9]Vicard P, Dawid AP. Remarks on: "Paternity analysis in special fatherless cases without direct testing of alleged father". *Forensic Science International*. 2006; 163(1-2): 158-60.

[10] Fung WK, Hu YQ, Chung YK..On statistical analysis of forensic DNA: theory, methods and computer programs. *Forensic Science International*. 2006; 162(1-3): 17-23.

[11]Coble MD, Loreille OM, Wadhams MJ, Edson SM, Maynard K, Meyer CE, Niederstätter H, Berger C, Berger B, Falsetti AB, Gill P, Parson W, Finelli LN. Mystery solved: the identification of the two missing Romanov children using DNA analysis. *PLoS One*. 2009;4(3):4838.

[12]Tamura A, Tsuji H, Miyazaki T, Iwata M, Nishio H, Hashimoto T, Kamiyama K, Suzuki K. Sibling incest and formulation of paternity probability: case report. *Legal Medicine*. 2000;2(4):189-96.

[13]Disponível em: <http://www.paternityindex.com/>. Acesso em: março, 2009.

[14]Disponível em: http://www.hugin.com/Products_Services/Products/Demo/Lite/. Acesso em: março, 2009.

[15]Disponível em: <http://www.umed.lodz.pl/ou/zms/>. Acesso em: março, 2009.

[16]Disponível em: <http://www.unikiel.de/medinfo/mitarbeiter/krawczak/download/>. Acesso em: março, 2009.

[17]Disponível em: vestad@telepolis.com.

[18]Disponível em: http://www.dip.sk/typo3/dip.sk/index.php?id=9&no_cache=1&L=1. Acesso em: março, 2009.

[19]Disponível em: <http://dna.view.com/index.html>. Acesso em: março, 2009.

[20]Disponível em: <http://www.hku.hk/statistics/EasyDNA/>. Acesso em: março, 2009.

[21]Disponível em: <http://qualitype.de/genoproof/index.jsp?lang=en>. Acesso em: março, 2009.

[22]Disponível em: <http://www.math.chalmers.se/~mjl/mostad/familias/>. Acesso em: março, 2009.

[23]Fung WK. User-friendly programs for easy calculations in paternity testing and kinship determinations. *Forensic Science International*. 2003; 136(1-3): 22-34.

[24]Egeland T, Mostad PF, Olaisen B. A computerised method for calculating the probability of pedigrees from genetic data. *Science Justice*. 1997; 37(4): 269-74.

[25]Grattapaglia D, Schmidt AB, Costa e Silva C, Stringher C, Fernandes AP, Ferreira ME. Brazilian population database for the 13 STR loci of the AmpFlSTR Profiler Plus and Cofiler multiplex kits. *Forensic Science International*. 2001; 118(1): 91-4.

[26]Silva DA, Crouse CA, Chakraborty R, Góes AC, Carvalho EF. Statistical analyses of 14 short tandem

repeat loci in Brazilian populations from Rio de Janeiro and Mato Grosso do Sul states for forensic and identity testing purposes. *Forensic Science International*. 2004; 139(2-3): 173-6.

[27]Dellalibera E, Havro ML, Souza M, Kajihara K, Mauricio-da-Silva L, dos Santos Silva R. Genetic analysis of 13 STR loci in the population from the State of Pernambuco, northeast Brazil. *Forensic Science International*. 2004; 146(1): 57-9.

[28]Whittle, MR; Romano, NL; Negreiros, VAC. Updated Brazilian genetic data, together with mutation rates, on 19 STR loci, including D10S1237. *Forensic Science International*. 2004; 139(2-3): 207-210.

[29]Rodrigues EMR, Palha TDBF, dos Santos SEB. Allele frequencies data and statistic parameters for 13 STR loci in a population of the Brazilian Amazon Region. *Forensic Science International*. 2007; 168 (2-3): 244-247.

[30]Gomes AV, Mauricio-Da-Silva L, Raposo G, et al. 13 STR loci frequencies in the population from Paraiba, Northeast Brazil *Forensic Science International*. 2007; 173 (2-3): 231-234.

[31]Brinkmaan B, Klitschar M, Neuhuber F, Huhne J, Rolf B. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *The American Journal of Human Genetics*. 1998; 62:1408-15.

[32]Weber JL, Wong C. Mutation of human short tandem repeats. *Human Molecular Genetics*. 1993; 2:11238.

[33]Chakraborty R, Kimmel M, Stivers D.N., Davison L.J., Deka R. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc. National Academy of Sciences*. 1997; 94:1041-6.

[34]Hurst L.D. & Elegenden H. Sex biases in the mutation rate. *Trends in Genetics*. 1998; 14:446-452.

[35]Xu X, Peng M, Fang Z, Xu X. The direction of microsatellite mutations is dependent upon allele length. *Nature Genetics*. 2000; 24:396-9.

[36]Elegenden H. Heterogeneous mutation process in human microsatellite DNA sequences. *Nature Genetics*. 2000; 24:400-2.

[37]Edwards M, Allen R.W. Characteristics of mutations at the D5S818 locus studied with a tightly linked marker. *Transfusion*. 2004; 44:83-90.

[38]Asicioglu F, Oguz-Savran F, Ozbek U. Mutation rate at commonly used forensic STR loci: Paternity testing experience. *Disease Markers*. 2004; 20:313-5.

[39]Pu CE, Linacre A. Systematic evaluation of sensitivity and specificity of sibship determination by using 15 STR loci. *Journal of Forensic and Legal Medicine*. 2008; 15(5): 329-34.

[40]R.E. Wenk, M. Traver and F.A. Chiafari, Determination of sibship in any two persons. *Transfusion*. 1996; (36): 259-262.

- [41]Pu CE, Linacre A. Increasing the confidence in half-sibship determination based upon 15 STR loci. *Journal of Forensic and Legal Medicine*. 2008; (6): 373-7.
- [42]Reid TM, Baird ML, Reid JP, Lee SC, Lee RF. Use of sibling pairs to determine the familial searching efficiency of forensic databases. *Forensic Science International Genetics*. 2008;2(4): 340-2.
- [43] Mályusz V, Poetsch M, Simeoni E, Schwark T, Manfred O, von Wurmb-Schwark N. Problems of assessing sibship probabilities by means of genetic analysis. *Arch Kriminol*. 2006; 218(1-2): 44-54.
- [44]Blouin MS. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution*. 2003; 18:503–511.