Luciano Abreu Brito

Identificação de Genes de Suscetibilidade às Fissuras
Labiopalatinas Não Sindrômicas: Influência da
Epidemiologia e da Estratificação Populacional

Identification of Susceptibility Genes to Nonsyndromic Cleft Lip /
Palate: epidemiology and population stratification influences

São Paulo

2011

Luciano Abreu Brito

Identificação de Genes de Suscetibilidade às Fissuras
Labiopalatinas Não Sindrômicas: Influência da
Epidemiologia e da Estratificação Populacional

Identification of Susceptibility Genes to Nonsyndromic Cleft Lip /
Palate: epidemiology and population stratification influences

Dissertação apresentada ao Instituto de
Biociências da Universidade de São Paulo, para
a obtenção de Título de Mestre em Ciências, na
Área de Biologia/Genética.

Orientadora: Maria Rita Passos-Bueno

São Paulo

2011

Ficha Catalográfica

Brito, Luciano A.

Identificação de Genes de
Suscetibilidade às Fissuras Labiopalatinas Não
Sindrômicas: Influência da Epidemiologia e da
Estratificação Populacional

79 páginas

Dissertação (Mestrado) - Instituto de
Biociências da Universidade de São Paulo.
Departamento de Genética e Biologia
Evolutiva

1.Fissura labiopalatina 2.Estudo de associação
3.Ancestralidade genômica

Comissão Julgadora:

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a). Maria Rita Passos Bueno

Orientadora

A todas as famílias portadoras
de fissuras labiopalatinas
atendidas durante
este projeto

Agradecimentos

À minha orientadora, Maria Rita Passos Bueno, pela oportunidade de desenvolver um trabalho no laboratório e por toda a orientação dispensada desde o meu ingresso no laboratório.

À minha família, em especial meus pais meu irmão e meu cachorro, com quem convivo diariamente e de quem sempre recebi todo o apoio necessário.

Aos amigos do laboratório, que muitas vezes transformam o ambiente de trabalho em um agradável circo, na real acepção da palavra. Agradeço em especial a todos aqueles que contribuíram de alguma maneira para este projeto, incluindo aqui as intermináveis jornadas de extração de DNA: Camila, Lucas, Lígia, Gerson, Melina, Carol, Dani Yumi, Dani Bueno, Cibele, Eric, Joel, Lara, Letícia, Bruno, May, Vanessa, Joanna, Felipe, Dani Moreira, Atique, Cíntia, Érika, Karina, Deinha, Roberto, Larissa, Simone e Meire.

Aos funcionários do Genoma, em especial ao pessoal do sequenciamento, Martha e Camila, que aturaram milhares de placas de genotipagem, algumas delas formatadas corretamente, e Kátia, excelente coorientadora durante todo este projeto.

A todos os pacientes atendidos, solícitos em contribuir com a pesquisa.

E à equipe da Operação Sorriso, pela oportunidade de participar dos programas de reabilitação, que me proporcionaram uma experiência única de contato com os pacientes.

Este trabalho contou com o apoio financeiro da FAPESP, do CNPq e do Ministério da Ciência e Tecnologia do Brasil.

Notas

Esta dissertação de mestrado compreende um trabalho desenvolvido durante os anos de 2009 a 2011 no Laboratório de Genética do Desenvolvimento do Centro de Estudos do Genoma Humano do Instituto de Biociências da Universidade de São Paulo (CEGH/IB-USP).

As referências citadas nos artigos (capítulos 3 e 4) encontram-se ao final de cada um, no formato da revista *American Journal of Medical Genetics*, na qual um dos artigos foi aceito para publicação. As referências bibliográficas dos capítulos 1, 2 e 5 (Introdução, Metodologia Geral e Discussão Geral / Conclusões, respectivamente) estão relacionadas ao final da dissertação.

Ambos os artigos estão redigidos em inglês, por motivos referentes à publicação.

Para suprir informações omitidas nas metodologias dos artigos, optei por criar uma breve seção de Metodologia Geral.

Lista de Abreviações

AIM: marcador informativo de ancestralidade

CI: intervalo de confiança (*confidence interval*)

FL: fissura de lábio somente

FL+P: fissura de lábio com fissura de palato

FL±P: fissuras labial com ou sem fissura no palato

FP: fissura de palato somente

H²: herdabilidade

Indel: polimorfismo de inserção-deleção

IRF6: *Interferon Regulatory Factor 6*

NS: não sindrômico (a)

QTL: *quantitative trait locus*

SNP: *single nucleotide polymorphism*

SVW: síndrome de van der Woude

TDT: teste de desequilíbrio de transmissão

Índice

Capítulo 1.	Introdução Geral	1
1.1.	Aspectos Gerais das Fisuras Labiopalatinas	1
1.1.1.	<i>Clínica</i>	1
1.1.2.	<i>Epidemiologia</i>	3
1.1.3.	<i>Embriologia</i>	4
1.1.4.	<i>Genéticas das Fissuras Não Sindrômicas</i>	6
1.1.5.	<i>Fatores Ambientais</i>	10
1.1.6.	<i>Reabilitação</i>	11
1.2.	Estratégias de Identificação de Genes Associados às Fissuras	12
1.2.1.	<i>Estudos de Ligação</i>	12
1.2.2.	<i>Estudos de Associação e as Dificuldades Geradas pela Estratificação Populacional</i>	13
1.2.3.	<i>Estudos de Expressão Gênica</i>	14
1.3.	O uso de Marcadores de Ancestralidade na Caracterização de Populações	14
1.4.	Objetivos	16
Capítulo 2.	Metodologia Geral	18
2.1.	Atendimento aos Pacientes	18
2.2.	Coleta de Material Biológico e Extração de DNA	19
Capítulo 3.	Genetic Contribution for Non-Syndromic Cleft Lip With or Without Cleft Palate (NS CL/P) in Different Regions of Brazil and Implications for Association Studies	20
Capítulo 4.	<i>IRF6</i> and 8q24 Roles in Susceptibility to Nonsyndromic Cleft Lip/Palate in Brazilian Population	29
Capítulo 5.	Discussão Geral e Conclusões	54
Resumo		59
Abstract		61
Referências Bibliográficas		63

Capítulo 1. Introdução Geral

1.1 Aspectos Gerais das Fissuras labiopalatinas

1.1.1 Clínica

As fissuras labiopalatinas caracterizam-se essencialmente pela separação incompleta das cavidades oral e nasal, admitindo um grande espectro de variabilidade clínica. O acometimento do lábio pode ser unilateral ou bilateral, com a fissura atingindo a(s) narina(s) ou não. O palato pode ser acometido por fissura pré-forame incisivo anterior e pós-forame; quando a fissura atinge as duas estruturas, a fissura é dita transforame. Também podem ser classificadas como uni ou bilaterais (Figs. 1 e 2; Schutte e Murray, 1999; Gorlin *et al.*, 2001).

Figura 1 – Variabilidade clínica das fissuras labiopalatinas, em crianças atendidas durante a realização do projeto. A) Fissura labial unilateral direita; B) Fissura labial bilateral; C) Fissura de palato; D) Detalhe de fissura de palato submucosa com úvula bífida.

Figura 2 – Esquema da variabilidade clínica das fissuras palatinas. A) Fissura pré-forame unilateral. B) Fissura pré-forame bilateral; C) Fissura transforame unilateral; D) Fissura transforame bilateral; E) Fissura pós-forame. No centro, divisões do palato (duro / mole e primário / secundário, esta última classificação derivada da formação embriológica) e posições das estruturas (Kim et al., 2010).

Além das características clinicamente mais óbvias de serem notadas, as fissuras labiopalatinas podem carregar consigo microformas mais brandas, de difícil percepção, como, por exemplo, a descontinuidade do músculo orbicular do lábio (Rogers *et al.*, 2008; Weinberg *et al.*, 2008; Jugessur *et al.*, 2009), a fissura de palato submucosa (união imperfeita do músculo no véu palatino, mas com superfície da mucosa intacta; Gorlin *et al.*, 2001), e úvula bífida (Fig. 1).

A forma mais usual de classificação das fissuras labiopalatinas entre os geneticistas é o agrupamento em formas sindrômicas, quando outras anomalias estão presentes além da fissura, ou não sindrômicas (NS), quando estas ocorrem isoladamente.

1.1.2 Epidemiologia

As fissuras labiopalatinas são as malformações faciais mais comuns ao nascimento, de incidência frequentemente citada como 1:700 (Stanier e Moore, 2004), apresentando grande variação entre etnias, origens geográficas e níveis socioeconômicos. Em populações europeias, a incidência é, em média, 1:1.000 nascimentos (varia de 0,7 a 1,3). A incidência é considerada maior em nativos americanos (3,6:1.000), japoneses (2,1:1.000) e chineses (1,4:1.000); populações africanas, por outro lado, apresentam a menor incidência (0,3:1.000; Vanderas, 1987; Tolarova e Cervenka, 1998; Gorlin *et al.*, 2001). No Brasil, a incidência já foi estimada entre 0,28 a 1,54:1.000 (Nagem Filho e Martins, 1968; Menegotto e Salzano, 1991b; Loffredo *et al.*, 2001).

As formas sindrômicas de fissura respondem por, aproximadamente, 30% do total de casos. Existem descritas mais de 400 síndromes, de diferentes etiologias (*Online Mendelian Inheritance in Man*, OMIM). Podem estar relacionadas a este grupo alterações gênicas, cromossômicas ou ainda fatores teratogênicos em exposição no primeiro trimestre do desenvolvimento embrionário (Gorlin *et al.*, 2001), descritos na seção 1.4.

Dentre as formas sindrômicas, a mais comum é a síndrome de van der Woude (SVW), respondendo por 2% dos casos totais de fissura (e incidência de 1:35.000). É uma condição de herança autossômica dominante, com penetrância de 92%, que se caracteriza pela presença de comissuras no lábio inferior (que são glândulas salivares ectópicas, também chamadas de *pits* ou fossetas labiais), hipodontia, além da fissura, que pode envolver lábio, palato ou ambos. É a presença dos *pits* que permite, na

maioria dos casos, distinguir clinicamente a SVW das fissuras não sindrômicas (Van Der Woude, 1954; Gorlin *et al.*, 2001).

A fissura não sindrômica é uma doença complexa com herança do tipo multifatorial (a ocorrência envolve componentes genéticos e ambientais) e somam 70% dos casos totais de fissura (Schutte e Murray, 1999). São historicamente divididas em fissura de palato isolada (FP) e fissura labial com ou sem fissura de palato (FL±P), devido a uma provável diferença etiológica entre elas. Entre as evidências epidemiológicas sustentando essa separação, estão as diferentes incidências e riscos de recorrência observados em cada grupo (Tabela I).

Tabela I - Diferenças de incidência média entre fissuras labiais com ou sem fissura de palato (FL±P) e fissura de palato isolada (FP) e riscos de recorrência aproximados para futuros irmãos de uma criança afetada (Gorlin *et al.*, 2001).

	FL±P	FP
Incidência por etnia	1:2400 africanos 1:1000 caucasianos 1:500 asiáticos e nativos americanos	1:2000, sem heterogeneidade étnica
Incidência por sexo	60-80% homens	Predominância em mulheres
Risco de recorrência	3-4%	2%

1.1.3 Embriologia

O início da formação do lábio se dá na quarta semana de gestação, quando os primórdios faciais são originados a partir migração de células da crista neural e de

células mesodérmicas. Nesta etapa, podem ser distinguidas cinco proeminências (frontonasal, duas mandibulares e duas maxilares; Fig. 3). A partir da quarta semana de desenvolvimento, forma-se a proeminência nasal medial a partir do crescimento da proeminência frontonasal. O lábio superior e o palato primário são então formados quando as proeminências maxilares encontram a proeminência nasal, processo que ocorre até a sétima semana (Jiang *et al.*, 2006).

A formação do palato secundário inicia-se na sexta semana, a partir do crescimento das lâminas palatinas, estruturas derivadas das proeminências maxilares e localizadas lateralmente à língua em desenvolvimento. Em seguida, as lâminas palatinas crescem horizontalmente e, após entrarem em contato uma com a outra, segue a fusão completa até o final da 12ª semana (Fig. 3; Kerrigan *et al.*, 2000).

A fissura, portanto, pode ser decorrente de qualquer distúrbio nos mecanismos biológicos que regulam estes processos durante esta etapa da embriogênese, sejam eles de natureza intrínseca (genéticos) ou extrínseca (ambientais)

Figura 3 - Esquema da formação do lábio e palato. a) Proeminência frontonasal, processos maxilares e processos mandibulares circundando a cavidade oral na 4ª semana do desenvolvimento embrionário. b) Por volta da 5ª semana, a fosseta nasal é formada, além dos processos nasais laterais e mediais. c) Processos nasais mediais fundem-se com os maxilares para a formação do lábio superior e palato primário no fim da 6ª semana. Os processos nasais laterais formam as abas nasais, e os mandibulares formam a mandíbula. d) Na 6ª semana, o palato secundário se desenvolve a partir dos processos maxilares, que crescem verticalmente. e) As lâminas palatinas elevam-se para uma posição horizontal acima da língua e encostam uma na outra, iniciando a fusão. f) A fusão completa das lâminas palatinas divide completamente as cavidades oral e nasal. Figura modificada de Dixon *et al.*, (2011).

1.1.4 Genética das Fissuras Não Sindrômicas

Dentre as evidências frequentemente observadas em favor de um importante componente genético para essas formas em diferentes estudos, está a agregação familiar (observada em 20-30% dos casos; Lie *et al.*, 1994; Carinci *et al.*, 2007), e a concordância em gêmeos (monozigóticos: 40-60%; dizigóticos: 3-5%; Christensen e Fogh-Andersen, 1993).

Estudos de herdabilidade (proporção da variância total da doença que é devida à variância genética) das FL±Ps NS também evidenciam a participação da genética nessas malformações e revelam diferentes contribuições genéticas para a malformação entre as populações. Estes estudos têm observado uma variação de herdabilidade de 17% a 84%, utilizando diferentes metodologias (Tabela II).

Ao contrário da maioria dos genes das formas síndrômicas, a associação de genes com FL±Ps NS permanece bastante controversa. Muitos estudos têm se dedicado a encontrá-los, incluindo estudos de ligação ou associação, de gene candidato ou varredura genômica, além de estudos de expressão gênica em células humanas ou em modelos animais (ver seção 1.2).

Os esforços para a descoberta de variantes de suscetibilidade às FL±Ps NS foram primeiramente baseados em estudos de associação com genes candidatos ou em estudos de ligação com famílias. Entretanto, muitas das associações encontradas nesses estudos iniciais eram de difícil replicação, dada a heterogeneidade dos fatores genéticos envolvidos e também à penetrância deles, presumidamente menor do que a dos genes responsáveis pelas formas síndrômicas de padrão mendeliano de herança. O estudo molecular destas, aliás, foi de grande valor para a elucidação dos genes responsáveis pelas fissuras não síndrômicas (Carinci et al., 2007). Recentemente, resultados mais consistentes têm sido produzidos por estudos de associação e ligação de varredura genômica com grande densidade de marcadores (Tabela III).

Tabela II - Estimativas de herdabilidade (H^2) realizadas em diferentes países, baseadas em estudos de gêmeos ou parentes em primeiro, segundo e terceiro grau

Local	Desenho do estudo	H^2	Referência
Finlândia	Gêmeos	17%	Nordstrom <i>et al.</i> , 1996
Dinamarca	Gêmeos	45%	Nordstrom <i>et al.</i> , 1996
Taiwan	Gêmeos	53%	Lin <i>et al.</i> , 1999
América do Sul	Parentes em 1º grau	74%	Menegotto e Salzano, 1991a
China	Parentes em 1º a 3º grau	77.6%	Hu <i>et al.</i> , 1982
França	Parentes em 1º grau	81%	Stoll <i>et al.</i> , 1991
Itália	Parentes em 1º a 3º grau	82%	Tenconi <i>et al.</i> , 1988
Itália	Parentes em 1º grau	84%	Calzolari <i>et al.</i> , 1988

Tabela III - Loci recentemente associados com FL±Ps NS em estudos de varredura genômica, genes candidatos e suas funções.

Região	Gene candidato	Produto	Referência
1q32-q41	<i>IRF6</i>	Fator de transcrição	Zuccherro <i>et al.</i> , 2004; Beaty <i>et al.</i> , 2010
8q24	Não há		Birnbaum <i>et al.</i> , 2009; Chiquet <i>et al.</i> , 2009; Grant <i>et al.</i> , 2009; Beaty <i>et al.</i> , 2010
10q25.3	<i>VAX1</i>	Fator de transcrição	Mangold <i>et al.</i> , 2009; Beaty <i>et al.</i> , 2010
6q14.2	<i>PRSS35</i>	Peptidase	Marazita <i>et al.</i> , 2002; Letra <i>et al.</i> , 2010
9q22	<i>FOXE1</i>	Fator de transcrição	Moreno <i>et al.</i> , 2009
1q22.1	<i>ABCA4</i>	Transportador transmembrana	Beaty <i>et al.</i> , 2010
20q12	<i>MAFB</i>	Fator de transcrição	Beaty <i>et al.</i> , 2010
4p16	<i>MSX1</i> <i>STK32B</i> <i>EVC</i> <i>EVC2</i>	Fator de transcrição Cinase Transdução de sinal Transdução de sinal	Van Den Boogaard <i>et al.</i> , 2000; Ingersoll <i>et al.</i> , 2010
1q36	<i>PAX7</i>	Fator de transcrição	Beaty <i>et al.</i> , 2010

Dentre os *loci* que surgiram dessas análises, marcadores no gene *IRF6* e em um deserto gênico em 8q24 são os mais corroborados em diferentes populações.

O gene *IRF6* pertence a uma família de nove fatores de transcrição que compartilham um domínio hélice-volta-hélice ligante do DNA altamente conservado e um domínio menos conservado ligante à proteína. Apesar da maioria dos IRFs desempenhar um papel conhecido na regulação da expressão dos interferons alfa e beta na resposta a infecções virais, o papel do *IRF6*, em particular, permanece desconhecido. Além da SVW, causada por mutações nesse gene, há indícios cada vez mais consistentes de que variantes mais comuns nesse gene estejam envolvidas com a predisposição às FL±Ps NS (Kondo *et al.*, 2002; Zuccherro *et al.*, 2004; Carinci *et al.*, 2007).

1.1.5. Fatores Ambientais

Muitos fatores ambientais têm sido associados à ocorrência das fissuras, tais como anticonvulsivantes, corticosteroides, benzodiazepinas, antagonistas do ácido fólico, talidomida, exposição a pesticidas e solventes orgânicos, ingestão em excesso de vitamina A, diabetes mellitus, diabetes gestacional, estresse, obesidade, consumo de álcool e tabagismo, entre outros (Wyszynski e Beaty, 1996; Schutte e Murray, 1999; Bender, 2000; Hernandez-Diaz *et al.*, 2000; Pradat *et al.*, 2003; Jugessur e Murray, 2005; Chevrier *et al.*, 2006; Krapels *et al.*, 2006; Deroo *et al.*, 2008; Dolk *et al.*, 2008). O ácido fólico, por outro lado, parece atuar na prevenção da ocorrência das FL/Ps (Wilcox *et al.*, 2007). Os resultados dos estudos dos fatores ambientais, entretanto, são bastante conflitantes.

1.1.6. Reabilitação

A grande maioria dos pacientes com fissura labiopalatina apresenta desenvolvimento cognitivo normal; contudo, a reabilitação bem sucedida representa um custo elevado para a sociedade e consiste em um processo bastante penoso para o afetado e seus familiares, uma vez que o tratamento envolve vários procedimentos cirúrgicos e intervenção de uma equipe multidisciplinar, podendo incluir médicos, geneticistas, fonoaudiólogos, dentistas, nutricionistas e psicólogos, entre outras especialidades. Como frequentemente os indivíduos afetados também são vítimas de certo isolamento social, a reabilitação é considerada completa quando a total reinserção do afetado na sociedade é alcançada.

Este acompanhamento multidisciplinar frequentemente dura até a idade adulta e tem custos estimados em mais de US\$ 100,000 por paciente (*Centers for Disease and Prevention* [CDC], 1995, 2007), colocando esse grupo de malformações como um relevante problema de saúde pública. Com o intuito de contribuir para a prevenção e reabilitação dos indivíduos acometidos pelas fissuras, existe um grande esforço na tentativa de identificar os fatores de predisposição, tanto genéticos quanto ambientais, a estas malformações.

1.2. Estratégias de Identificação de Genes Associados às Fissuras

1.2.1. Estudos de Ligação

Estudos de ligação são abordagens de mapeamento genético baseadas na observação da segregação de marcadores genéticos de localizações cromossômicas conhecidas com a doença em várias meioses em indivíduos de uma ou mais famílias. Esta abordagem parte do princípio de que dois *loci* próximos estão ligados e que tendem a ser transmitidos juntos nas meioses. Assim sendo, quando for identificado um marcador que segregue junto com a doença em estudo, sugere-se que o *locus* da doença esteja localizado próximo deste marcador (Altschuler *et al.*, 2008). Genes de grande efeito, geralmente associados a síndromes mendelianas, são mais propícios a serem identificados com essa abordagem, para a qual um dos requisitos é justamente a presença de famílias grandes segregando a doença. De fato, muitos genes mendelianos responsáveis por fissuras síndrômicas foram identificados dessa maneira, como o *IRF6*, responsável pela SVW (Kondo *et al.*, 2002), e o *COL2A1*, responsável pela síndrome de Stickler tipo II (Francomano *et al.*, 1987), forma síndrômica de fissura de palato, miopia e perda de audição de herança autossômica dominante.

Entretanto, as fissuras não síndrômicas, assim como a maioria das doenças comuns, parecem não ser causadas por variantes de grande efeito. Dado o padrão de herança multifatorial, a ocorrência de muitos indivíduos afetados em uma família também é menos frequente. Essas características, portanto, acabam dificultando a abordagem de ligação na busca dos genes associados a fissuras não síndrômicas.

Apesar dessas limitações, muitos estudos de ligação já foram realizados em fissura não síndrômica, mas os resultados em geral não são bem replicados.

1.2.2. Estudos de Associação e as Dificuldades Geradas pela Estratificação Populacional

A abordagem geralmente mais utilizada para doenças complexas é o estudo de caso-controle, no qual se procura verificar se a frequência de variantes polimórficas difere entre o grupo de pacientes e controles. Muitos estudos de caso-controle têm sido realizados para encontrar fatores de suscetibilidade às fissuras não síndrômicas, seja investigando um gene candidato, ou região candidata, ou o genoma inteiro através de densos painéis de marcadores genéticos. Algumas variantes encontradas por essa abordagem foram descritas na tabela III.

O maior problema dessa abordagem, entretanto, é a estratificação populacional decorrente da mistura étnica, ou seja, a presença de subgrupos relativamente distintos dentro da população aparentemente homogênea sob estudo. Subgrupos ancestralmente distintos tendem a ter representações majoritárias de diferentes ancestralidades no nível genético e, assim, uma associação de um alelo super-representado nos pacientes em relação aos controles pode ser decorrente da maior frequência desse alelo na população ancestral predominante nos casos.

Uma maneira de corrigir o efeito da estratificação é a utilização dos pais de afetados como um controle interno, no intuito de detectar desequilíbrio na transmissão dos alelos (aplica-se nessa abordagem o teste TDT, do inglês, *transmission disequilibrium test*). Contudo, na prática é bastante difícil conseguir amostra dos pais dos pacientes (Tiwari *et al.*, 2008).

1.2.3. Estudos de expressão gênica

Os estudos de expressão baseiam-se na análise dos transcritos para o estudo de genes candidatos ou para a busca de novas vias gênicas diferencialmente reguladas em indivíduos com FL±P NS.

Diversos modelos animais vêm sendo empregados nos estudos de expressão, como camundongo, galinha e *zebrafish* (ou “paulistinha”), entre outros, e diversos genes foram caracterizados como de grande importância no processo da formação labial ou do fechamento do palato (Kondo *et al.*, 2002; Brown *et al.*, 2003; Mukhopadhyay *et al.*, 2004). Estudos em cultura celular de tecidos da região afetada de pacientes ou de células tronco também vêm sendo empregados para a investigação de processos biológicos importantes para a etiologia das FL±Ps NS (Bosi *et al.*, 1998; Jakobsen *et al.*, 2009; Baroni *et al.*, 2010; Bueno *et al.*, 2010).

1.3. O uso de Marcadores de Ancestralidade na Caracterização de Populações

Muitos estudos têm se dedicado na caracterização de populações com o uso de marcadores informativos de ancestralidade (ou AIMs, do inglês *ancestry informative marker*), polimorfismos de frequências alélicas são substancialmente diferentes entre as populações. Entretanto, a quantidade de marcadores necessários para caracterizar ancestralmente uma população é alvo de debate. Recentemente, alguns estudos esforçaram-se em esclarecer essa questão (Tabela IV).

Tabela IV- Estudos envolvendo a caracterização de populações baseados em marcadores de ancestralidade. Embora os estudos tenham obtidos resultados semelhantes no agrupamento de populações nas 5 regiões maiores (no caso, continentes), houve algumas discrepâncias na estimativa da variabilidade genética entre indivíduos dentro de uma mesma população, entre populações dentro de uma região e entre regiões. Os indels foram os marcadores que mais detectaram a variabilidade entre continentes. STR: *Short Tandem Repeats* (microssatélites); RFLP: *restriction fragment length polymorphism* (polimorfismo de fragmento de restrição).

Referência	Barbujani <i>et al.</i> , 1997	Barbujani <i>et al.</i> , 1997	Romualdi <i>et al.</i> , 2002	Rosenberg <i>et al.</i> , 2002	Excoffier <i>et al.</i> , 2003	Bastos- Rodrigues <i>et al.</i> , 2006
Marcadores	RFLPs	STR	Inserções Alu	STR	STR	Indels
Nº <i>loci</i>	70	30	21	377	377	40
Nº populações	11	14	32	52	52	52
Nº regiões	5	5	5	5	5	5
Variabilidade entre indivíduos dentro da população (%)	84,5	84,5	82,9	93,2	87,6	85,7
Variabilidade entre populações de mesma região (%)	3,9	5,5	8,2	2,5	3,1	2,3
Variabilidade entre regiões (%)	11,7	10,0	8,9	4,3	9,2	12,1

Barbujani *et al.*, (1997) analisou 109 marcadores (30 microssatélites e 79 polimorfismos de fragmentos de restrição) em indivíduos de 16 populações. Romualdi *et al.*, (2002) analisou 21 inserções Alu em 32 populações. Rosenberg *et al.*, (2002) e Excoffier e Hamilton, (2003) agruparam 52 populações em 5 grandes

grupos geográficos (coincidentes com os continentes) utilizando-se de 377 microssatélites autossômicos. Bastos-Rodrigues *et al.*, (2006), utilizando apenas 40 marcadores indels dialélicos espalhados pelo genoma, caracterizou de forma muito satisfatória, em relação aos estudos anteriores, a estrutura do genoma das mesmas populações estudadas por Rosenberg *et al.* (2002) e Excoffier *et al.* (2003), obtendo a maior diferença genética entre regiões dentre esses estudos. Além do sucesso em caracterizar as populações, o uso dos marcadores indels apresentou também a vantagem da genotipagem por PCR-multiplex, com alelos de diferentes tamanhos amplificados em uma mesma reação. Outro grupo posteriormente implementou outro painel de 40 marcadores de inserção-deleção para a caracterização de populações brasileiras (Santos *et al.*, 2010).

A caracterização de marcadores capazes de identificar estrutura populacional e a validação de metodologias que contornem esse obstáculo são avanços especialmente úteis para a busca de genes associados às fissuras, pois, além da heterogeneidade e do padrão de herança multifatorial das FL±P NS, a estrutura populacional decorrente da mistura étnica também é um fator que dificulta essa busca, principalmente em países com histórico de miscigenação.

1.4. Objetivo

Como foi explicitado, a busca de genes de suscetibilidade às FL±Ps NS tem encontrado bastantes complicações, decorrentes da heterogeneidade da doença e seu padrão multifatorial de herança. Somada a elas, a estratificação populacional constitui uma dificuldade à parte especialmente em populações recentemente miscigenadas.

Sendo assim, é extremamente importante que possamos identificar variáveis e metodologias que contornem esses obstáculos.

O nosso objetivo principal é o de identificar genes de predisposição as FL±Ps NS. Considerando-se a heterogeneidade da doença, o padrão de herança multifatorial e as características da estrutura da nossa população, nos propusemos neste projeto de mestrado a:

- a) Caracterizar variáveis que possam interferir na busca dos fatores de risco genético desta malformação. Para isto, realizamos um estudo epidemiológico no qual estimamos, entre outras coisas, a herdabilidade e grau de consanguinidade de famílias com afetados por FL±Ps NS de várias regiões do país;
 - b) Avaliamos o uso de marcadores de ancestralidade do tipo inserção-deleção para corrigir o efeito da mistura étnica brasileira em estudos de caso-controle, em uma abordagem conhecida como *structured association*. Por meio dela, testamos duas regiões candidatas, as quais têm sido mais validadas: o gene *IRF6* e a região 8q24.
-

Capítulo 2. Metodologia Geral

2.1. Atendimento aos Pacientes

Os pacientes alvos deste estudo foram averiguados por meio de programas médicos da organização não governamental Operação Sorriso, que oferece atendimento multidisciplinar a crianças portadoras de fissuras. Os programas são realizados anualmente em cidades brasileiras parceiras da organização: Santarém-PA; Barbalha-CE (interior do estado), Fortaleza-CE, Maceió-AL e Rio de Janeiro-RJ. A divulgação dos programas é realizada não somente na cidade sede, mas também nas cidades vizinhas, de modo que essas cinco cidades tornam-se polos de encaminhamento de crianças afetadas de toda a localidade durante a realização do programa.

Durante o atendimento ao paciente, foi aplicado um questionário para levantar informações sobre a família (consanguinidade, presença de outros afetados, número de irmãos) e gestação. A fissura foi caracterizada em labial e/ou de palato, e, na presença de qualquer outra malformação, o paciente foi classificado como sindrômico. Atenção especial foi aplicada na identificação de *pits* ou fossetas labiais, para evitar a inclusão de portadores da síndrome de van der Woude no grupo de pacientes não sindrômicos. Termo de consentimento livre e esclarecido foi assinado pelo paciente ou responsável e o trabalho foi aprovado pelo Comitê de Ética e Pesquisa do Instituto de Biociências da Universidade de São Paulo.

2.2. Coleta de Material Biológico e Extração de DNA

Durante as avaliações dos pacientes, foram coletadas amostras de células de mucosa oral, esfregando uma escova cervical em cada lado da bochecha do paciente. Após a secagem a temperatura ambiente do excesso de saliva, as escovinhas foram guardadas em solução de etanol 70% até o momento da extração do DNA. Foram utilizados para a extração de DNA genômico os *kits* Nucleospin® Tissue (Macherey-Nagel, Alemanha) e QIAamp DNA Mini Kit (QIAGEN, Alemanha). Após a eluição, as amostras são estocadas a -80°C em concentração média de 25ng/ul.

Capítulo 3. Contribuição Genética para Fissuras Labiopalatinas não Sindrômicas em Diferentes Regiões do Brasil e Implicações para Estudos de Associação

Artigo submetido à *American Journal of Medical Genetics* e aprovado abril de 2011 para publicação

Resumo

Fissura labial com ou sem fissura de palato não sindrômica (FL/P NS) é uma doença complexa cujas estimativas de herdabilidade variam grandemente dependendo da população estudada. Para avaliar a importância da contribuição genética para as FL/P NS na população brasileira, nós conduzimos um estudo com 1042 famílias provenientes de cinco localidades (Santarém, Fortaleza, Barbalha, Maceió, e Rio de Janeiro). Nós também avaliamos o papel exercido pela consanguinidade e composição étnica. A proporção de casos familiares variou significativamente entre as localidades, com os valores mais altos em Santarém (44%) e os mais baixos em Maceió (23%). Estimativas de herdabilidade mostraram maior contribuição genética para a malformação em Barbalha (85%), seguido por Santarém (71%), Rio de Janeiro (70%), Fortaleza (64%) e Maceió (45%). Ancestralidade não foi correlacionada com a variação de herdabilidade encontrada. Apenas no Rio de Janeiro o coeficiente de endocruzamento foi significativamente maior nas famílias com FL/P NS do que na população geral. Risco de recorrência baseado na amostra total foi de 1,5-1,6%, variando entre as localidades (de 0,6-0,7% em Maceió a 2,2-2,8% em Barbalha). Este

trabalho mostra que o grau de determinação genética para FL/P NS varia de acordo com a região geográfica estudada, e essa diferença não pode ser atribuída à consanguinidade nem à ancestralidade. Esses resultados indicam Barbalha como a região mais propícia para estudos genéticos. Os dados apresentados aqui serão de grande utilidade para a interpretação de resultados de análises moleculares, e mostram que se deve tomar cuidado ao amostrar indivíduos de populações diferentes em estudos de associação.

RESEARCH ARTICLE

AMERICAN JOURNAL OF
medical genetics PART A

Genetic Contribution for Non-Syndromic Cleft Lip With or Without Cleft Palate (NS CL/P) in Different Regions of Brazil and Implications for Association Studies

Luciano A. Brito,¹ Lucas A. Cruz,¹ Kátia M Rocha,¹ Ligia K. Barbara,¹ Camila B.F. Silva,¹ Daniela F. Bueno,¹ Meire Agüena,¹ Débora R. Bertola,^{1,2} Diogo Franco,³ André M. Costa,⁴ Nivaldo Alonso,⁵ Paulo A. Otto,¹ and Maria Rita S. Passos-Bueno^{1*}¹Human Genome Research Center, Institute de Biosciences, University of São Paulo, São Paulo, Brazil²Genetics Unit, Department of Pediatrics, Children Institute, School of Medicine, University of São Paulo, São Paulo, Brazil³Department of Plastic Surgery, School of Medicine, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil⁴Division of Plastic Surgery, University of Health Sciences of Alagoas, Maceió, Brazil⁵Division of Plastic Surgery, Department of Surgery, School of Medicine of São Paulo University, São Paulo, Brazil

Received 27 July 2010; Accepted 16 March 2011

Non-syndromic cleft lip with or without cleft palate (NS CL/P) is a complex disease in which heritability estimates vary widely depending on the population studied. To evaluate the importance of genetic contribution to NS CL/P in the Brazilian population, we conducted a study with 1,042 families from five different locations (Santarém, Fortaleza, Barbalha, Maceió, and Rio de Janeiro). We also evaluated the role of consanguinity and ethnic background. The proportion of familial cases varied significantly across locations, with the highest values found in Santarém (44%) and the lowest in Maceió (23%). Heritability estimates showed a higher genetic contribution to NS CL/P in Barbalha (85%), followed by Santarém (71%), Rio de Janeiro (70%), Fortaleza (64%), and Maceió (45%). Ancestry was not correlated with the occurrence of NS CL/P or with the variability in heritability. Only in Rio de Janeiro was the coefficient of inbreeding significantly larger in NS CL/P families than in the local population. Recurrence risk for the total sample was approximately 1.5–1.6%, varying according to the location studied (0.6–0.7% in Maceió to 2.2–2.8% in Barbalha). Our findings show that the degree of genetic contribution to NS CL/P varies according to the geographic region studied, and this difference cannot be attributed to consanguinity or ancestry. These findings suggest that Barbalha is a promising region for genetic studies. The data presented here will be useful in interpreting results from molecular analyses and show that care must be taken when pooling samples from different populations for association studies. © 2011 Wiley-Liss, Inc.

Key words: non-syndromic cleft lip with or without cleft palate; complex disease; heritability; consanguinity; ancestry contribution; recurrence risk

How to Cite this Article:

Brito LA, Cruz LA, Rocha KM, Barbara LK, Silva CBF, Bueno DF, Agüena M, Bertola DR, Franco D, Costa AM, Alonso N, Otto PA, Passos-Bueno MRS. 2011. Genetic contribution for non-syndromic cleft lip with or without cleft palate (NS CL/P) in different regions of Brazil and implications for association studies.

Am J Med Genet Part A 9999:1–7.

INTRODUCTION

Non-syndromic cleft lip with or without cleft palate (NS CL/P) is the most prevalent congenital facial defect, with a large degree of clinical expressivity, ranging from a lip scar to a complete bilateral cleft lip with dental–alveolar and palate involvement, without any other associated malformation [Gorlin et al., 2001]. It is still unclear whether all these subtypes represent the spectrum of clinical variability under a common etiological basis or underlie a genetic heterogeneity [Harville et al., 2005; Rahimov et al., 2008].

The worldwide birth prevalence of NS CL/P has been estimated as 1:1,000, but it varies according to ethnicity (African: 0.3:1,000;

Grant sponsor: [FAPESP](#)^{Q1}; Grant sponsor: CNPq; Grant sponsor: Brazilian Ministry of Health.

*Correspondence to:

Maria Rita S. Passos-Bueno, Rua do Matão, 277, sala 200, Cidade Universitária, São Paulo, SP 05508-090, Brazil.

E-mail: passos@ib.usp.br, passosbueno@gmail.com

Published online 00 Month 2011 in Wiley Online Library (wileyonlinelibrary.com).

DOI 10.1002/ajmg.a.34036

European: 0.7–1.3:1,000; Chinese: 1.4:1,000; Japanese: 2.1:1,000; Amerindian: 3.6:100), geographic origin, and socioeconomic level [Gorlin et al., 2001; Carinci et al., 2007]. In Brazil, birth prevalence estimates for NS CL/P have ranged from 0.28 to 1.54:1,000 [Menegotto and Salzano, 1991a; Loffredo et al., 2001].

A strong genetic basis for NS CL/P has been inferred due to the higher recurrence risk for first-degree relatives of affected subjects, and also to the large proportion of affected subjects with positive family history (about 20–30% [Melnick, 1992; Lie et al., 1994; Carinci et al., 2007]). Twin studies have shown a significantly higher concordance for NS CL/P in monozygotic twins (40–60%) as compared to dizygotic twins (3–5%), further reinforcing the importance of genetic factors in the etiology of NS CL/P [Christensen and Fogh-Andersen, 1993; Jugessur et al., 2009]. A multifactorial inheritance model, with additive effects of genetic and environment factors triggering the malformation, best fits most of the epidemiological findings on NS CL/P [Fraser, 1974; Schutte and Murray, 1999; Gorlin et al., 2001; Grosen et al., 2010].

Many NS CL/P candidate genes have been identified through different methods, but only a few loci, particularly those recently identified through genome-wide association studies (GWAS), have been consistently corroborated in different populations [Zucchero et al., 2004; Rahimov et al., 2008; Birnbaum et al., 2009a, 2009b; Mangold et al., 2009; Beaty et al., 2010]. One of the reasons underlying this difficulty of replication may be the importance of genetic contribution for the malformation. Genetic contribution to NS CL/P appears to vary according to the population studied, with estimates ranging from as low as 17% in Finland to 84% in Italy [Hu et al., 1982; Calzolari et al., 1988; Tenconi et al., 1988; Stoll et al., 1991; Christensen and Fogh-Andersen, 1993; Nordstrom et al., 1996; Lin et al., 1999]. In South America, genetic contribution to NS CL/P was estimated to be 70–74% [Menegotto and Salzano, 1991b]. However, the degree of genetic contribution can be affected by ethnicity, local environment, consanguinity, and geographical origin. In countries with a recent history of immigration and admixture, such as Brazil, ethnicity might have a strong effect on the genetic contribution to multifactorial diseases as it the case of NS CL/P.

Brazil has an extense territory with significant regional differences in its cultural habits and population origins. The country is divided in five large geographical regions: North, Northeast, Center-West, Southeast, and South. The prevalence of consanguineous marriages, a known risk factor in multifactorial diseases, is very heterogeneous in different regions of Brazil, being more common in the Northeastern part of the country, and increasing in frequency as one travels inland. The North and Southeast regions of Brazil are considered weak inbreeding zones [Salzano and Freire-Maia, 1967].

In this article, we estimate the importance of genetic contribution to NS CL/P susceptibility in Brazilian populations from five locations, in three different geographical regions. We also ascertained whether ethnicity and consanguinity play a role in the etiology of this malformation in our population. This pioneer work will be relevant to future studies that utilize molecular analyses.

PATIENTS AND METHODS

We ascertained NS CL/P families treated under the medical program of a non-governmental organization—Operation Smile

(<http://www.operationsmile.org/>), with the collaboration of local hospitals. Between the years 2007 and 2009, we visited five cities in three different Brazilian regions: Santarém-PA (Northern region—N); Fortaleza-CE, Barbalha-CE, and Maceió-AL (Northeastern region—NE); and Rio de Janeiro-RJ (Southeastern region—SE). The locations of the cities studied are shown in Figure 1.

For each family, a pedigree with at least two ascendant generations of the proband was created, based on a standardized interview of the parents (available on request). We also recorded cleft classification [unilateral or bilateral cleft lip with or without cleft palate (CP)], sex, age, and information about the pregnancy.

We have included only non-syndromic CL/P cases. Therefore, patients who had another malformation or syndrome in addition to the CL/P were excluded. Patients with only CP were also excluded, as in most of the cases CP constitutes a different genetic entity from NS CL/P [Fraser, 1974; Jehee et al., 2009]. A final sample of 1,042 families was obtained: 121 from Santarém, 331 from Fortaleza, 86 from Barbalha, 188 from Maceió, and 316 from Rio de Janeiro.

Skin color as phenotypic classification for ethnicity correlates poorly with ancestry [Suarez-Kurtz et al., 2007], and therefore, we used molecular methods to determine the ancestral composition of our sample. We genotyped 40 autosomal insertion–deletion ancestry-informative polymorphisms spread across the genome (following the criteria adopted in Bastos-Rodrigues et al. [2006]; primer sequences available on request) through multiplex PCR in 602 subjects. Fragment analysis was performed with Gene Mapper software after capillary electrophoresis in ABI 3730 DNA Analyzer

FIG. 1. Locations analyzed in this study: Santarém, in the state of Para-PA (Northern region—N); Fortaleza and Barbalha, in the state of Ceará-CE, and Maceió, in the state of Alagoas-AL (Northeastern region—NE), and Rio de Janeiro, in the state of Rio de Janeiro-RJ (Southeastern region—SE).

(Applied Biosystems^{Q2}). Ancestry contributions of Europeans, West Africans, and Amerindians were calculated with the software Structure 2.3.3 [Pritchard et al., 2000].

Subjects were grouped in familial or non-familial cases, based on the report of at least one other affected member in the family. We did not identify any affected twin pairs. The proportion of affected first-degree relatives (parents and siblings) of the proband was also calculated for each location. Frequencies of familial cases, consanguineous cases, and affected first-degree relatives were compared by chi-squared tests followed by the analysis of adjusted residuals (Haberman's test).

The heritability values for NS CL/P were estimated graphically for each location using a diagram as proposed by Professor C.A.B. Smith [Emery, 1986]. The graph correlates the malformation prevalence among first-degree relatives of the probands and in the general population; the obtained heritability is expressed as the proportion of the phenotypic variation attributable to the additive genetic variation. Therefore, high heritability implies high genetic determination. Tests of trend were performed for the average inbreeding coefficient (F), frequency of consanguineous cases, proportion of genetic determination, and frequency of familial cases to verify their association (negative results were not reported).

When the proband was the offspring of a consanguineous marriage, the family was classified as a consanguineous case and a corresponding F was calculated. For a given type of mating, this coefficient is obtained simply by multiplying by one-half the parent's coefficient of relationship (which in turn reveals the proportion of genes shared by two relatives); a child of a first-cousin mating, for instance, has $F = 1/16$ [Cavalli-Sforza and Bodmer, 1971]. A case was considered consanguineous only when $F \geq 1/64$, that is, offspring of a second-cousin marriage or closer kinships. An average coefficient of inbreeding was calculated for each population by weighing the various F by their corresponding frequencies in the populations, including the class $F = 0$ (non-consanguineous marriages [Hedrick, 1986]).

Empiric recurrence risks were calculated for the siblings of affected children. Since all families were ascertained through affected children in sibships, the recurrence risk was obtained from the formula $\text{recurrence risk} = (\text{nao} - \text{nsf}) / (\text{nao} + \text{nno} - \text{nsf})$, where nao is the number of affected offspring, nno is the number of normal offspring, and nsf is the number of sibships or families; $\text{nao} - \text{nsf}$ is therefore the number of affected children in excess of one per sibship. The correction shown above is essentially the same used for correcting segregation ratios from material collected with bias ascertainment [Emery, 1986]. Recurrence risks for children

born to affected mothers or fathers could not be determined in this study due to the small sample size.

RESULTS

We have evaluated 1,042 NS CL/P families ascertained between the years 2007 and 2009 at Operation Smile's programs in five different locations in Brazil: Santarém, Fortaleza, Barbalha, Maceió, and Rio de Janeiro. The proportion of affected males to females was 1.6:1. Most of the cases had cleft lip and palate (82.4%), while the remaining cases (17.6%) had cleft lip only.

After genotyping 40 ancestry-informative markers, we observed that the ancestry contributions of our subjects differs significantly among the regions studied ($P = 0.014$). Fortaleza, Barbalha, and Maceió (NE) presented similar ancestry contributions, whereas Rio de Janeiro (SE) was slightly enriched with African ancestry, compared to the others. Subjects from Santarém (N), on the other hand, presented a higher Amerindian contribution than the other regions (Table I).

To evaluate the genetic contribution to NS CL/P in this sample, we first estimated the frequency of familial cases and the number of affected first-degree relatives of the proband, as shown in Table II. We observed that the proportion of familial cases was significantly different among the five locations ($P < 0.0005$) with the lowest proportions in Maceió (23%) and Rio de Janeiro (24%), and the highest in Santarém (44%), followed by Fortaleza (40%). Comparison of the number of affected first-degree relatives of probands revealed statistically significant differences ($P = 0.02$), for which Barbalha and Maceió contributed the most (Table II). Heritability of NS CL/P, estimated based on these data, revealed a major genetic contribution in Barbalha (85%). Santarém, Rio de Janeiro, and Fortaleza presented intermediate genetic contributions (71%, 70%, and 64%, respectively), whereas Maceió was the location with less genetic contribution to NS CL/P (45%).

The rates of consanguinity among parents of probands were also significantly different among the five studied locations ($P = 0.001$), with the highest values in families from Fortaleza and Barbalha and lower estimates in Santarém, Maceió, and Rio de Janeiro. The same trends were observed for the respective coefficients of inbreeding (Table III). As expected, average F is positively correlated with the frequency of consanguineous cases (correlation coefficient = 0.97 at the 5% critical level) and, from now on, we will refer to F.

Recurrence risk based on the existence of an affected member in the sibship was 1.6% (1.5% if consanguinity is not considered) and varied according to the location, with the highest values in Barbalha

TABLE I. Ancestral Contributions to Patients From Five Regions in Brazil and Total Sample

	Santarém	Fortaleza	Barbalha	Maceió	Rio de Janeiro	Total
Genotyped patients	90	216	60	120	116	602
European	60%	71%	68%	67%	65%	67%
West African	11%	13%	19%	19%	23%	17%
Amerindian	29%	16%	13%	14%	12%	16%

TABLE II. Frequencies of Familial Cases and of Affected First-Degree Relatives and Heritability of Non-Syndromic Cleft Lip With or Without Cleft Palate for Different Regions in Brazil

	Proportion of familial cases	Proportion of affected first-degree relatives	Heritability
Santarém	49/111 = 44%	13/619 = 2.1%	71%
Fortaleza	131/330 = 40%	23/1358 = 1.7%	64%
Barbalha	32/86 = 37%	14/414 = 3.4%	85%
Maceió	42/183 = 23%	8/973 = 0.8%	45%
Rio de Janeiro	75/312 = 24%	30/1492 = 2.0%	70%
Total	329/1022 = 32%	88/4856 = 1.8%	—

TABLE III. Frequencies of Consanguineous Cases and Coefficients of Inbreeding [F] of Affected and General Populations [Salzano and Freire-Maia, 1967] From Five Regions in Brazil

	Proportion of consanguineous cases	F (present study)	F (general population)
Santarém	7/121 = 5.8%	0.002	0.002
Fortaleza	38/331 = 11.5%	0.004	0.004
Barbalha	12/86 = 14.0%	0.004	0.004
Maceió	8/188 = 4.3%	0.001	0.001
Rio de Janeiro	15/316 = 4.8%	0.002	0.001

(2.2% excluding consanguineous cases and 2.9% with the total sample) and the lowest in Maceió (0.7% without consanguineous cases and 0.6% with the total sample; Table IV).

DISCUSSION

The present Brazilian population is the result of centuries of ethnic admixture from three main groups: Europeans, Sub-Saharan Africans, and Amerindians. Because the patterns of immigration flow and interethnic crosses were not equal in different areas of the country, population composition may vary even among geographically close regions [Parra et al., 2003; Pena et al., 2009]. We would also expect a higher contribution of Europeans and Amerindians than Africans to our NS CL/P sample, compared to the overall

population, if we take into account the prevalence of NS CL/P in the populations from which Brazilians originate. However, ancestry contributions were similar to previous estimates for the corresponding local control populations (60% European, 12% African, and 28% Amerindian for Santarém (N) and 75% European, 15% African, and 10% Amerindian for Fortaleza, Barbalha, and Maceió (NE); no control data are available for Rio de Janeiro (SE) [Santos et al., 2010]). Although we are aware that the use of 40 markers allows the discrimination of only three ancestries, our data suggest that there is no strong association between ancestry and NS CL/P in our population.

Most studies have supported a genetic contribution to NS CL/P, reporting 20–30% of cases as familial [Schutte and Murray, 1999; Gorlin et al., 2001; Jugessur et al., 2009]. Considering the five

TABLE IV. Number of Sibships and Recurrence Risks [Presented as Number of Affected Sibs/Total Sibs] Excluding and Including Consanguineous Cases in Five Regions in Brazil

	Excluding consanguineous cases		Including consanguineous cases	
	Number of sibships	Recurrence risk (95% CI)	Number of sibships	Recurrence risk (95% CI)
Santarém	114	5/373 = 1.3% [0.4–3.1]	121	5/382 = 1.3% [0.4–3.0]
Fortaleza	290	9/629 = 1.4% [0.7–2.7]	328	13/706 = 1.8% [1.0–3.1]
Barbalha	73	4/184 = 2.2% [0.6–5.5]	87	7/242 = 2.9% [1.2–5.9]
Maceió	164	4/578 = 0.7% [0.2–1.8]	172	4/612 = 0.6% [0.2–1.7]
Rio de Janeiro	261	17/915 = 1.9% [1.1–3.0]	273	19/971 = 2.0% [1.2–3.0]
Total	902	39/2679 = 1.5% [1.0–2.0]	981	48/2913 = 1.6% [1.2–2.2]

CI, confidence intervals.

locations together, we found 32% of cases were familial; however, this proportion varied significantly accordingly to the population ascertained, with the highest proportion of familial cases in Santarém (44%) and the lowest in Maceió (23%). Consistent with these findings, heritability analysis revealed a lower genetic contribution in Maceió (45%) than in the other locations. The higher heritability value in Barbalha (85%, even though it did not have the highest proportion of familial cases) is explained by the stronger aggregation of affected first-degree relatives in this location. Although heritability estimated through affected first-degree relatives is not an independent line of evidence in favor of increased genetic etiology from the proportion of familial cases, it appears to more accurately characterize populations with higher contributions from genetic factors. With the exception of Maceió, our heritability values are close to those previously reported from South America (70% [Menegotto and Salzano, 1991b]) or European countries such as Italy and France (81–84% [Calzolari et al., 1988; Tenconi et al., 1988; Stoll et al., 1991]).

Although we observed different *F* values among the five studied populations, the coefficients of inbreeding obtained in our study are very similar to those found by Salzano and Freire-Maia [1967] in local Brazilian control populations. The one exception is Rio de Janeiro, in which *F* was twice as high in NS CL/P families as in the local population (Table III). Therefore, it seems that consanguinity is not a major contributing factor for NS CL/P in most of these populations and does not appear to be sufficient to explain the higher heritability in Barbalha.

The recurrence risk for the total sample, whether consanguineous couples were included or excluded, was approximately 1.5–1.6%, which is lower than values estimated for other populations (Italy, 2.4% [Tenconi et al., 1988]; Denmark, 3.5% [Grosen et al., 2010], and Norway, 7.0% [Sivertsen et al., 2008]). The recurrence risk varied according to the ascertained location, with the highest values in Barbalha (2.2–2.8%), which further confirms that genetic factors play a more prominent role in this group.

Although the reasons for the higher degree of genetic contribution to NS CL/P in Barbalha are unknown, these results clearly indicate that this is a promising location to study candidate loci for NS CL/P. On the other hand, the lowest heritability value observed (45% in Maceió) does not seem to be related to differences of ethnic constitution, as no obvious differences were observed in the contribution of European and African ancestries to affected subjects and local population from Maceió and Barbalha (Table I). Environmental factors may have a larger impact in the etiology of NS CL/P in Maceió than in the other locations.

In summary, we have shown that the contribution of genetic factors to NS CL/P varies across different regions of Brazil, and that neither consanguinity nor ancestry is the main factor influencing this finding. However, we cannot rule out that the effect of consanguinity was not evidenced because the *F* values of the control population used in the analysis might have changed from the time that they were calculated. The recurrence risks obtained for our populations were much lower than those previously reported in other countries, and consanguinity played a small role only in Fortaleza and Barbalha. These results have direct implications for genetic counseling, as they provide more accurate information about the genetic risks for Brazilian families with NS CL/P. These

data can provide the basis for delineation of public health research strategies to prioritize regions to identify genetic or environmental factors of susceptibility to NS CL/P. Identification of such factors may in the future contribute to lowering the prevalence of NS CL/P. In most association studies thus far conducted for NS CL/P, the proportion of genetic contribution has not been considered. Our data indicate that at-risk alleles will be underrepresented in populations in which heritability for NS CL/P is lower, compared to those with higher heritability values. This effect should be even more dramatic for alleles that confer only a modest increase in disease susceptibility. Recent genetic studies including GWAS have replicated findings identifying *IRF6*, *MSX1*, and loci in chromosomes 1p22.1, 8q24, 9q21, 10q25.3, 17q22, and 20q12 as susceptibility loci for NS CL/P [Zuccherro et al., 2004; Birnbaum et al., 2009b; Mangold et al., 2009; Moreno et al., 2009; Beaty et al., 2010; Ingersoll et al., 2010]. Most of these studies have included very large cohorts of patients and controls. If our assumption is correct, we would expect to find a correlation between positive association for some of these loci and higher heritability. In addition, we would expect that in a group of patients with higher heritability, a smaller sample size will be sufficient to detect the association. We intend to test this hypothesis in the near future. If confirmed, it will indicate that pooling of samples from different populations is not necessarily the best strategy to study the genetics of NS CL/P, and will also represent a significant decrease in the costs of genetic analysis.

ACKNOWLEDGMENTS

We want to thank the Operation Smile Brazil, its executive director, Clovis Brito, for the opportunity of joining the rehabilitation programs and developing a genetic assistance program for the patients, and the hospitals which hosted Operation Smile's programs: São Camilo (Santarém-PA), Albert Sabin (Fortaleza-CE), São Vicente de Paulo (Barbalha-CE), Santa Casa de Misericórdia (Maceió-AL), and Clementino Fraga Filho (Rio de Janeiro-RJ). We would like to thank to the colleagues Beatriz Burin, Camila Ornelas, Cintia Ribeiro, and Fernanda Jehee for patient consultations and Diogo Meyer and David Schlesinger for assistance in the ancestry analysis. We also appreciate Mrs. Constança Urbani's secretarial assistance and the financial support of FAPESP, CNPq, and the Brazilian Ministry of Health.

REFERENCES

- Bastos-Rodrigues L, Pimenta JR, Pena SD. 2006. The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Ann Hum Genet* 70:658–665.
- Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. 2010. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet* 42:525–529.

- Birnbaum S, Ludwig KU, Reutter H, Herms S, de Assis NA, Diaz-Lacava A, Barth S, Lauster C, Schmidt G, Scheer M, Saffar M, Martini M, Reich RH, Schiefke F, Hemprich A, Potzsch S, Potzsch B, Wienker TF, Hoffmann P, Knapp M, Kramer FJ, Nothen MM, Mangold E. 2009a. IRF6 gene variants in Central European patients with non-syndromic cleft lip with or without cleft palate. *Eur J Oral Sci* 117:766–769.
- Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Berge SJ, Reich RH, Schiefke F, Hemprich A, Potzsch S, Steegers-Theunissen RP, Potzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nothen MM, Mangold E. 2009b. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet* 41:473–477.
- Calzolari E, Milan M, Cavazzuti GB, Cocchi G, Gandini E, Magnani C, Moretti M, Garani GP, Salvioli GP, Volpato S. 1988. Epidemiological and genetic study of 200 cases of oral cleft in the Emilia Romagna region of northern Italy. *Teratology* 38:559–564.
- Carinci F, Scapoli L, Palmieri A, Zollino I, Pezzetti F. 2007. Human genetic factors in nonsyndromic cleft lip and palate: An update. *Int J Pediatr Otorhinolaryngol* 71:1509–1519.
- Cavalli-Sforza LL, Bodmer WF. 1971. The genetics of human populations. San Francisco: W.H. Freeman. p 966.
- Christensen K, Fogh-Andersen P. 1993. Cleft lip (+/- cleft palate) in Danish twins, 1970–1990. *Am J Med Genet* 47:910–916.
- Emery AEH. 1986. Methodology in medical genetics. Edinburgh: Churchill Livingstone. p 198.
- Fraser FC. 1974. Updating the genetics of cleft lip and palate. *Birth Defects Orig Artic Ser* 10:107–111.
- Gorlin RJ, Cohen MM Jr, Hennekam RCM. 2001. Syndromes of the head and neck, 4th edition. New York: Oxford University Press. p 1283.
- Grosen D, Chevrier C, Skytthe A, Bille C, Molsted K, Sivertsen A, Murray JC, Christensen K. 2010. A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: Support for the multifactorial threshold model of inheritance. *J Med Genet* 47:162–168.
- Harville EW, Wilcox AJ, Lie RT, Vindenes H, Abyholm F. 2005. Cleft lip and palate versus cleft lip only: Are they distinct defects? *Am J Epidemiol* 162:448–453.
- Hedrick PW. 1986. Average inbreeding or equilibrium inbreeding? *Am J Hum Genet* 38:965–970.
- Hu DN, Li JH, Chen HY, Chang HS, Wu BX, Lu ZK, Wang DZ, Liu XG. 1982. Genetics of cleft lip and cleft palate in China. *Am J Hum Genet* 34:999–1002.
- Ingersoll RG, Hetmanskij J, Park JW, Fallin MD, McIntosh I, Wu-Chou YH, Chen PK, Yeow V, Chong SS, Cheah F, Sull JW, Jee SH, Wang H, Wu T, Murray T, Huang S, Ye X, Jabs EW, Redett R, Raymond G, Scott AF, Beaty TH. 2010. Association between genes on chromosome 4p16 and non-syndromic oral clefts in four populations. *Eur J Hum Genet* 18:726–732.
- Jehee FS, Burin BA, Rocha KM, Zechi-Ceide R, Bueno DF, Brito L, Souza J, Leal GF, Richieri-Costa A, Alonso N, Otto PA, Passos-Bueno MR. 2009. Novel mutations in IRF6 in nonsyndromic cleft lip with or without cleft palate: When should IRF6 mutational screening be done? *Am J Med Genet Part A* 149A:1319–1322.
- Jugessur A, Farlie PG, Kilpatrick N. 2009. The genetics of isolated orofacial clefts: From genotypes to subphenotypes. *Oral Dis* 15:437–453.
- Lie RT, Wilcox AJ, Skjaerven R. 1994. A population-based study of the risk of recurrence of birth defects. *N Engl J Med* 331:1–4.
- Lin YC, Lo LJ, Noordhoff MS, Chen YR. 1999. Cleft of the lip and palate in twins. *Changcheng Yi Xue Za Zhi* 22:61–67.
- Loffredo LCM, Freitas JAS, Grigolli AAG. 2001. Prevalence of oral clefts from 1975 to 1994, Brazil. *Rev Saúde Pública* 35:571–575.
- Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B, Scheer M, Reich RH, Hemprich A, Potzsch S, Blaumeiser B, Moebus S, Krawczak M, Schreiber S, Meitinger T, Wichmann HE, Steegers-Theunissen RP, Kramer FJ, Cichon S, Propping P, Wienker TF, Knapp M, Rubini M, Mossey PA, Hoffmann P, Nothen MM. 2009. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nat Genet* 42:24–26.
- Melnick M. 1992. Cleft lip (+/- cleft palate) etiology: A search for solutions. *Am J Med Genet* 42:10–14.
- Menegotto BG, Salzano FM. 1991a. Epidemiology of oral clefts in a large South American sample. *Cleft Palate Craniofac J* 28:373–376.
- Menegotto BG, Salzano FM. 1991b. Clustering of malformations in the families of South American oral cleft neonates. *J Med Genet* 28:110–113.
- Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, Johnson MK, Brauer D, Krahn K, Daack-Hirsch S, L'Heureux J, Valencia-Ramirez C, Rivera D, Lopez AM, Moreno MA, Hing A, Lammer EJ, Jones M, Christensen K, Lie RT, Jugessur A, Wilcox AJ, Chines P, Pugh E, Doheny K, Arcos-Burgos M, Marazita ML, Murray JC, Lidral AC. 2009. FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. *Hum Mol Genet* 18:4879–4896.
- Nordstrom RE, Laatikainen T, Juvonen TO, Ranta RE. 1996. Cleft-twin sets in Finland 1948–1987. *Cleft Palate Craniofac J* 33:340–347.
- Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. 2003. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA* 100:177–182.
- Pena SD, Bastos-Rodrigues L, Pimenta JR, Bydlowski SP. 2009. DNA tests probe the genomic ancestry of Brazilians. *Braz J Med Biol Res* 42:870–876.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M, Christensen K, Bille C, Melbye M, Jugessur A, Lie RT, Wilcox AJ, Fitzpatrick DR, Green ED, Mossey PA, Little J, Steegers-Theunissen RP, Pennacchio LA, Schutte BC, Murray JC. 2008. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. *Nat Genet* 40:1341–1347.
- Salzano FM, Freire-Maia N. 1967. *Populações brasileiras*. São Paulo: Companhia Editora Nacional. p 178.
- Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, Pereira R, Gusmao L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, Santos SE. 2010. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat* 31:184–190.
- Schutte BC, Murray JC. 1999. The many faces and factors of orofacial clefts. *Hum Mol Genet* 8:1853–1859.
- Sivertsen A, Wilcox AJ, Skjaerven R, Vindenes HA, Abyholm F, Harville E, Lie RT. 2008. Familial risk of oral clefts by morphological type and severity: Population based cohort study of first degree relatives. *BMJ* 336:432–434.
- Stoll C, Alembik Y, Dott B, Roth MP. 1991. Epidemiological and genetic study in 207 cases of oral clefts in Alsace, north-eastern France. *J Med Genet* 28:325–329.

Suarez-Kurtz G, Vargens DD, Struchiner CJ, Bastos-Rodrigues L, Pena SD. 2007. Self-reported skin color, genomic ancestry and the distribution of GST polymorphisms. *Pharmacogenet Genomics* 17:765–771.

Tenconi R, Clementi M, Turolla L. 1988. Theoretical recurrence risks for cleft lip derived from a population of consecutive newborns. *J Med Genet* 25:243–246.

Zuccherro TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. 2004. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med* 351:769–780.

Q1: Please check the grant sponsors.

Q2: Please provide complete location.

Capítulo 4. *IRF6* and 8q24 roles in Susceptibility to Nonsyndromic Cleft Lip / Palate in Brazilian Population

Luciano A. Brito, Camila B. F. Silva, Kátia M. Rocha, David Schlesinger, Lucas A. Cruz, Ligia K. Bárbara, Cibele Masotti, Daniela F. Bueno, Diogo Meyer, Maria Rita Passos-Bueno

Papel do *IRF6* e da Região 8q24 na Suscetibilidade às Fissuras

Labiopalatinas Não Sindrômicas na População Brasileira

Resumo

Fissura labial com ou sem fissura de palato não sindrômica (FL±P NS) é uma doença complexa com incidência mundial estimada em 1:700, variando de acordo com a população estudada. Dentre os *loci* de suscetibilidade propostos, os marcadores rs642961 no gene *IRF6* e rs987535 em um deserto gênico na região 8q24 são os mais corroborados em diferentes populações. Para testar os papéis desses marcadores e de um marcador adicional (rs2236906, em um íntron do gene *IRF6*) na população brasileira, nós realizamos um estudo de *structured association*, abordagem que infere a estrutura populacional gerada pela mistura étnica para testar associação das variantes. Além dos SNPs, foram genotipados 40 marcadores de inserção-deleção informativos de ancestralidade em 563 pacientes de 5 localizações diferentes (Santarém-PA, Barbalha-CE, Fortaleza-CE, Maceió-AL e Rio de Janeiro-RJ) e 336 controles de São Paulo-SP. As estimativas de ancestralidade foram realizadas com o

programa *Structure*, e os testes estatísticos de associação com o programa STRAT. As contribuições das populações ancestrais foram de 16,5% africana, 70,7% europeia e 12,8% ameríndia nos controles e 18,7% africana, 62,2% europeia e 19,1% ameríndia nos pacientes. O SNP rs2236906 apresentou desvio de equilíbrio de Hardy-Weinberg e foi excluído da análise. Barbalha foi a única região onde associação significativa foi encontrada para os SNPs (rs642961: $P=0,003$; $OR_{het}=1,94$; $OR_{hom}=4,16$; rs987525: $P=0,005$; $OR_{het}=1,91$; $OR_{hom}=0,86$). Este é o primeiro estudo em população brasileira que detectou associação para esses SNPs, embora ela só tenha sido notada em uma região (Barbalha). Essa região foi anteriormente caracterizada como a de maior herdabilidade para as FL $\pm$ S NS (85%), o que é consistente com a associação dos SNPs. A falta de associação observada nas outras regiões sugere que estes SNPs não têm um fator tão decisivo na população brasileira. Os dados apresentados aqui também reforçam a importância do uso de abordagens que contornem o problema da estratificação populacional.

Abstract

Nonsyndromic cleft lip with or without cleft palate (NS CL $\pm$ P) is a complex disease with worldwide incidence estimated as 1:700, varying according to the populations. Among the putative susceptibility *loci*, rs642961 in *IRF6* gene and rs987525 in 8q24 gene desert are the most corroborated in different populations. In order to test the roles of these two markers plus an additional marker in *IRF6*, rs2236906, in Brazilian population, we carried out a structured association study, in which details of population structure are inferred to testing for association. In addition to the SNPs,

we also genotyped a panel of 40 ancestry informative insertion-deletion markers in 563 patients from 5 different locations (Santarém-PA, Barbalha-CE, Fortaleza-CE, Maceió-AL and Rio de Janeiro-RJ) and 336 controls. Ancestry analysis was performed with Structure software, and statistical test for association was performed with STRAT software. Ancestry contribution was estimated as 16.5% African, 70.8% European and 12.8% Amerindian in controls and 18.7% African, 62.2% European and 19.1% Amerindian in cases. The SNP rs2236906 presented deviation from Hardy-Weinberg equilibrium and was excluded from the analysis. Barbalha was the only region with significant association for any of the two remaining SNPs (rs642961: $P=0.003$; $OR_{het}=1.94$; $OR_{hom}=4.16$; rs987525: $P=0.005$; $OR_{het}=1.91$; $OR_{hom}=0.86$). This is the first time that association for these SNPs is found in Brazil, although it could only be detected in one region (Barbalha). This region was previously characterized as the most influenced by genetic influences (heritability for NS CL $\pm$ P = 85%), which is consistent with these findings. The lack of association observed in the other regions suggests that these SNPs may not play a major role in Brazil. The data presented here also reinforce the importance of approaches for correcting case-control studies in admixed populations such as Brazilian.

Introduction

Nonsyndromic cleft lip with or without cleft palate (NS CL $\pm$ P) is one of the most common congenital defects, with worldwide prevalence estimated as 1:700, which seems to vary according to population, being rarer in Africans and more frequent in Asians and Amerindians [Vanderas 1987]. NS CL $\pm$ P is considered a

multifactorial disease, and its occurrence depends on interactions between genetic and environmental components [Mitchell and Risch 1992]. Incidence among Brazilians is estimated in 0.65-0.85:1,000 [Loffredo et al. 2001; Menegotto and Salzano 1991; Nagem Filho and Martins 1968].

Many susceptibility *loci* have been discovered in the past two decades, such as *IRF6* (Interferon Regulatory Factor 6, 1q32-q41), 8q24, 4p16, 10q25.3, 6q14.2 and 9q22-q33 [Beaty et al. 2010; Birnbaum et al. 2009b; Chiquet et al. 2009; Ingersoll et al. 2010; Mangold et al. 2009; Marazita et al. 2002; Moreno et al. 2009; van den Boogaard et al. 2000]. However, possibly due to genetic heterogeneity, most of them have been difficult to be replicated in different populations. Of these *loci*, *IRF6* has been appointed as the most determinant to NS CL±P, and replication in different populations has been observed [Blanton et al. 2005; Ghassibe et al. 2005; Jugessur et al. 2008; Park et al. 2007; Pegelow et al. 2008; Scapoli et al. 2005; Zuccherro et al. 2004].

IRF6 belongs to a family of nine transcription factors sharing a much conserved DNA-binding domain and a less conserved protein-binding domain. Although many IRFs play role in regulation of host defense [Taniguchi et al. 2001], *IRF6* function remains obscure. *IRF6* is known to be expressed in the medial edge of the mouse fusing palate, tooth germ, hair follicle, genitalia and skin [Kondo et al. 2002], and *Irf6*-null mice present abnormal skin, limb and craniofacial development [Ingraham et al. 2006; Richardson et al. 2006]. Mutations in *IRF6* also lead to van der Woude Syndrome (VWS), the most common type of syndromic CL±P, and Popliteal Pterygium Syndrome (PPS), in whom pterygium in the members is present in addition to facial clefts and lip pits [Kondo et al. 2002]. Within the *IRF6 locus*, a SNP in an enhancer (rs642961, G>A, Fig. 1) emerges as a functional candidate to NS

CL±P, with the most robust signal among other SNPs in the region in different studies (OR_{GA}=1.75; OR_{AA}=1.94; relative risk_{GA} = 1.68; relative risk_{AA} = 2.40 [Birnbaum et al. 2009a; Rahimov et al. 2008]. The polymorphic allele (A) was proposed to disrupt an AP-2alpha-binding site, which plays a role as an enhancer for *IRF6*. However, the association of rs642961 has not been replicated in a recent study with Brazilian population [Paranaíba et al. 2010].

Another promising *locus* is a gene desert in 8q24. The reported associations has been replicated in European and South American populations, with the strongest signal at the SNP rs987525 (C>A; (OR_{CA} =1.58 to 2.57; OR_{AA} =3.17 to 6.05; [Beaty et al. 2010; Birnbaum et al. 2009b; Nikopensius et al. 2009]. Due to the inexistence of known genes along the haplotypic block of this SNP (Fig. 2), the importance of this region is believed to be regulatory. Nonetheless, its biological role remains unclear.

In Brazilian population, the association of none of these variants was corroborated or even described so far. Brazilians constitute a tri-hybrid population, composed primarily by West Africans (brought as slaves between 16th and 19th centuries), Europeans (settlers and immigrants) and Amerindians (native inhabitants; [Centro de Documentação e Disseminação de Informações (Brazil) 2007]). Admixed populations are known to be problematic in association studies, since admixture can generate population substructure, leading to false correlations between cases and controls. In order to contribute to a better understanding on the controversial data obtained in the literature for NS CL±P at risk variants, we conducted this study to test the structure effects of our population and the heritability impact on the detection of at-risk alleles for clefting (heritability indexes of NS CL±P in Brazilian populations previously described in Brito et al., 2011, in press). To achieve these goals, we tested the association of rs642961 and rs2236906 in *IRF6 locus* (Fig. 1), and rs987525 in

8q24 in Brazilian population, through a structured association approach. Rs642961 and rs987525 are SNPs previously associated in other populations, whereas rs2236906 is an *IRF6* intronic SNP, recently characterized as a QTL (quantitative trait *locus*; [Dimas et al. 2009]), thus a potential candidate SNP, which is in linkage disequilibrium with rs642961 in HapMap reference populations (Fig. 3).

Figure 1 – Scheme of *IRF6* gene (1q32-141), not drawn in scale. Exons are numbered from 1 to 10; rs2236906 is located in intron 3 whereas rs642961 is within a 5' enhancer. Modified from Jugessur et al., 2008.

Figure 2 – Linkage disequilibrium patterns of 8q24 desert gene region in HapMap reference populations: YRI (Yoruba, Nigeria), CEU (Central Europeans from Utah) and JPT/CHB (Japanese from Tokyo and Han Chinese from Beijing). No genes are annotated in this region (<http://genome.ucsc.edu/>).

Figure 3 – Linkage disequilibrium patterns in *IRF6* locus in HapMap reference populations: YRI (Yoruba, Nigeria), CEU (Central Europeans from Utah) and JPT/CHB (Japanese from Tokyo and Han Chinese from Beijing). On the top, *IRF6*, rs2236906 and rs642961 positions (<http://genome.ucsc.edu/>).

Methods

Affected individuals were ascertained from medical programs of a non-governmental organization - Operation Smile (<http://www.operationsmile.org/>) – in 5 different Brazilian cities (Santarém-PA, Fortaleza-CE, Barbalha-CE, Maceió-AL and Rio de Janeiro-RJ). All control individuals were sampled from São Paulo-SP, since we were not able to ascertain local controls. Syndromic patients or patients with cleft palate only were removed from the sample. Patient ascertainment method and inclusion *criteria* were previously reported in Brito et al., [2011, in press], and non-related individuals were pooled from that sample to carry out this study. This study was approved by the Research Ethics Committee of the Institute of Biosciences, University of São Paulo, Brazil and informed consent was obtained from the patients or their parents.

In order to characterize the regional and individual ancestries, we genotyped a panel of 40 ancestry informative insertion-deletion polymorphisms through multiplex PCR with 4 to 7 markers per reaction (primer sequences available upon request). M13 tails (with 18 nucleotides) labeled with FAM fluorochrome were added in every 5' end of forward primers, so that PCR products incorporated the tails. These markers are spread across the chromosomes and were accessed from Marshfield database (<http://research.marshfieldclinic.org/>), which holds a set of indel markers described by Weber et al. [2002]. Marker selection *criteria* were the same adopted in Bastos-Rodrigues et al. [2006]. Fragment analysis was performed with Gene Mapper

software (Applied Biosystems, USA), after capillary electrophoresis targeting the PCR products for the FAM-labeled M13 tail. After excluding individuals with low genotype assignment (<80%), we obtained the final sample of 563 patients (Santarém, n=88; Barbalha, n=57; Fortaleza, n=225; Maceió, n=77 and Rio de Janeiro, n=116) and 336 controls. The SNPs (rs2236906, rs642961 and rs987525) were genotyped with TaqMan® method (Applied Biosystems, USA) which uses allele-specific probes labeled with fluorochromes in PCR amplification for allele discrimination [Shen et al. 2009].

For statistical analyses, Structure software was used to infer population stratification due to admixture [Falush et al. 2003; Falush et al. 2007; Pritchard et al. 2000a]. We assumed K=3 populations in our analysis, based on the well-known tri-hybrid origin of Brazilians, and incorporated “learning samples” in the run, with pre-specified population of origin to assist the software to estimate ancestry of the admixed individuals. Following the ancestry contribution estimations, STRAT software was used to test each SNP for association, assuming no stratification and also conditioning on the individual ancestries (structured association; Pritchard et al. 2000b).

We also performed a haplotype-based analysis of SNPs in *IRF6 locus* with PHASE 2.1 software (<http://www.stat.washington.edu/stephens/software.html>), which reconstructs the haplotypes based on a Bayesian statistical method [Stephens and Donnelly 2003; Stephens et al. 2001]. A permutation test was then applied to test deviations from random haplotypic distribution between cases and controls. We also performed a likelihood ratio test of linkage disequilibrium based on the genotypic data and estimated the pairwise linkage disequilibrium values D' and r^2 based on the deduced haplotypes at 0.05 significance level with Arlequin 3.5 software [Excoffier

and Lischer 2010]. Odds ratio for the genotypes with one or two copies of the risk allele were calculated whenever a positive association was found.

Results

Average ancestry contribution was estimated as 18.7% African, 62.2% European and 19.1% Amerindian for cases and 16.5% African, 70.7% European and 12.8% Amerindian for controls. Table 1 summarizes ancestry contributions estimated for patients (altogether and splitting the analysis by patient's geographic region of origin) and controls. Individual ancestry contributions are plotted in Fig. 4.

Table 1 - Ancestry contributions for controls and patients. Patients were firstly analyzed altogether and secondly assorted by the five geographical regions of origin of patients

	African contribution (%)	European contribution (%)	Amerindian contribution (%)
Controls	16.5	70.7	12.8
All patients	18.7	62.2	19.1
Santarém	12.3	56.5	31.2
Barbalha	21.2	62.5	16.3
Fortaleza	16.6	65.6	17.8
Maceió	19.8	61.5	18.6
Rio de Janeiro	26.4	59.0	14.6

Figure 2 - Graphic representations of individual ancestry contributions generated by Structure software in run assuming $K=3$ populations. A) Bar plot displaying African (red), European (green) and Amerindian (blue) contributions. Every single column represents an individual, which is grouped into their subpopulations (Santarém-PA, Barbalha-CE, Fortaleza-CE, Maceió-AL, Rio de Janeiro-RJ and controls from São Paulo-SP). The last three groups represent the parental populations (EUropeans, AFRicans and AMerIndians), used as “learning samples” to assist the software to estimate ancestry of admixed individuals; B) Triangle plot showing parental populations clustered in the vertices and cases and controls spread into the triangle according to their ancestralities. Each circle represents an individual. Cases are in purple, and controls in orange.

Allele and genotype frequencies of rs2236906 (*IRF6* intron), rs642961 (*IRF6* enhancer), and rs987525 (8q24 region) are shown in Tables II, III, and IV, respectively, with tests for Hardy-Weinberg deviations and association tests at 95% confidence assuming no stratification and correcting for stratification effects.

Table II – Genotype frequencies and allele G frequency of SNP rs2236906, localized in *IRF6* intron 3, estimated for patients and controls; P-values obtained from Hardy-Weinberg equilibrium test, association test assuming no population structure and after the correction for stratification.

rs2236906	Genotype frequencies (%) GG/GT/TT	f(G) (%)	H-Weq. P-value	P-value (no structure)	Corrected P-value
Controls	33.0/17.1/49.8	58.4	<1E-6	-	-
All patients	41.2/16.6/42.2	50.4	<1E-6	0.001	0.3
Santarém	43.8/21.2/35.0	45.6	<1E-6	0.003	0.52
Barbalha	50.9/14.6/34.5	41.8	<1E-6	0.001	0.003
Fortaleza	39.5/15.3/45.1	52.8	3E-5	0.068	0.24
Maceió	44.0/14.7/41.3	47.4	<1E-6	0.03	0.36
Rio	35.6/18.3/46.1	55.3	<1E-6	0.427	0.59

Table III - Absolute genotype frequencies and relative minor allele frequencies of SNP rs642961, located within an *IRF6* enhancer, estimated for patients and controls; P-values obtained from Hardy-Weinberg equilibrium test, association test assuming no population structure and after the correction for stratification.

rs642961	Genotype frequencies (%) AA/AG/GG	MAF (%)	H-Weq. P- value	P-value (no structure)	Corrected P-value
Controls	2.7/20.5/76.8	13.0	0.18	-	-
All patients	5.0/26.5/68.5	18.2	0.01	0.003	0.06
Santarém	4.5/34.1/61.4	21.6	0.37	0.004	0.17
Barbalha	11.3/30.2/58.5	26.4	0.10	0.0002	0.003
Fortaleza	5.3/25.3/69.3	18.0	0.34	0.021	0.14
Maceió	1.3/28.6/70.1	15.6	0.42	0.388	0.51
Rio	4.3/19.8/75.9	14.2	0.04	0.623	0.43

Table IV - Absolute genotype frequencies and relative minor allele frequencies of SNP rs987525, located in 8q24 desert gene, estimated for patients and controls; P-values obtained from Hardy-Weinberg equilibrium test, association test assuming no population structure and after the correction for stratification.

rs987525	Genotypes frequencies (%) AA/AC/CC	MAF (%)	H-Weq. P- value	P-value (no structure)	Corrected P-value
Controls	13.1/43.8/43.1	35.0	0.48	-	-
All patients	11.5/44.6/43.9	33.8	0.92	0.62	0.006
Santarém	6.8/38.6/54.6	26.1	0.99	0.03	0.07
Barbalha	12.3/57.9/29.8	41.2	0.09	0.198	0.005
Fortaleza	12.0/41.8/46.2	32.9	0.42	0.471	0.33
Maceió	11.7/44.1/44.1	33.8	0.92	0.777	0.21
Rio	13.8/48.3/37.9	37.9	0.79	0.417	0.19

Assuming no population stratification, both SNPs in *IRF6* locus are significantly associated in whole patient sample analysis. Testing each region separately, Rio de Janeiro was the only region that did not reach statistical significance for any of them; rs2236906 was significantly associated in Santarém, Barbalha and Maceió, and rs642961 was significantly associated in Santarém, Barbalha and Fortaleza. Considering the haplotype of these two SNPs, we found a differential distribution between the full sample of patients cases and controls ($P=0.02$) and between Barbalha and controls ($P=0.01$), with overrepresentation of G-A haplotype (haplotype frequencies described in Table V).

Table V - Haplotype frequencies for SNPs at IRF6 in cases and controls, and significance of the permutation test of the null hypothesis that haplotypes are randomly distributed between cases and controls.

	G-G	G-A	T-G	T-A	P-value
Controls	58.4	0.2	28.4	12.9	-
All patients	50.2	0.6	31.4	17.7	0.02
Santarém	46.1	0.1	32.3	21.5	0.06
Barbalha	39.9	2.3	33.9	23.9	0.01
Fortaleza	52.2	0.8	29.8	17.2	0.22
Maceió	48.9	0	15.5	35.6	0.18
Rio	55.5	0.5	30.3	13.7	0.93

When the test is performed correcting for stratification effects, rs642961 becomes marginally significant ($P=0.06$), whereas Barbalha presented the most robust association signal ($P=0.003$; $OR_{het}=1.94$ (95% CI 1.0-3.7) and $OR_{hom}=4.16$ (95% CI 1.4-12.2)); *IRF6* intronic SNP was significantly associated only in Barbalha after the correction.

Likelihood test of linkage disequilibrium revealed strong linkage disequilibrium between these *IRF6* SNPs in patient and control samples ($r^2 < .05$; $D'_{cases} = 0.94$; $D'_{controls} = 0.98$; Table VI).

Table VI - Linkage disequilibrium between the two SNPs in *IRF6*. Linkage disequilibrium coefficients (D, D' and r^2).

	D	D'	r^2
Controls	0.076	0.98	0.2043
All patients	0.087	0.94	0.2008
Santarém	0.098	1	0.2295
Barbalha	0.093	0.84	0.1840
Fortaleza	0.088	0.93	0.2091
Maceió	0.076	1	0.1753
Rio	0.074	0.95	0.1790

The SNP in 8q24 (rs987525) was significantly associated in Santarém without correction for stratification; after the correction, it weakened ($P=0.07$), whereas Barbalha and the whole sample associations became stronger (Barbalha: $P=0.005$; $OR_{het}=1.91$ (1.0-3.6) and $OR_{hom}=0.93$ (0.4-2.2); whole sample: $OR_{het}=1.0$ (0.7-1.3), $OR_{hom}=0.86$ (0.6-1.3)).

Discussion

Brazilians are highly admixed, and some deviation in Hardy-Weinberg equilibrium is expected if the population under study is structured. For that reason, Hardy-Weinberg deviation analyses were performed for each region and for the whole

sample. The 8q24 SNP was in Hardy-Weinberg equilibrium in all regions and *IRF6* enhancer SNP presented discrete deviation in Rio de Janeiro ($P=0.04$). The *IRF6* intronic SNP, on the other hand, was in strong disequilibrium in all samples. Although this finding strongly suggests genotyping error, our sample has presented a high rate of genotyping concordance ($>90\%$). A limitation of TaqMan method cannot be discarded, and further genotype confirmation will be performed sequencing part of the sample. Considering the potential bias that this SNP could introduce in the study, it was removed from further analysis.

Differences observed between the tests assuming no population structure and correcting for this effect can be attributed to the differences of genetic background among patients and among regions, and also due to allele frequencies in parental populations (HapMap reference populations CEU and YRI). These differences reflect the importance of such corrections in order to avoid false correlations. It is interesting to observe that both positive and negative associations were observed in the non-corrected analysis.

We found positive association for rs642961 for the first time in a Brazilian population, but only in Barbalha sample, which remained with a significant signal after the correction for stratification ($P=0.003$, risk allele = A). In this region, the effect of this SNP ($OR_{CA} = 1.94$ (95% CI 1.0-3.7) and $OR_{AA} = 4.16$ (95% CI 1.4-12.2)) is similar to those described in other studies. Since the first report of the *IRF6* enhancer SNP association [Rahimov et al. 2008], many studies were carried out aiming a validation in other populations. Most of them have confirmed a significant association; however, the odds ratio or relative risk attributed to the risk allele (A) are frequently lower than the reported by Rahimov et al. (2008). In addition, some studies have not found association for this SNP. Blanton et al. [2010a], investigating this

SNP in Hispanic and non-Hispanic white families in USA, could not find a significant association for it alone; however, haplotype analysis with a silent mutation in *IRF6* (rs2235371) supported the role of this gene NS CL±P etiology. The fact that the analysis with all regions could not reach statistical significance for *IRF6*, in addition to a previous study which fail to detect this association in Brazilians [Paranaíba et al. 2010], suggests that this SNP may not play a major role in our population; thus, other variants in *IRF6* region or other genes may be more involved in NS CL±P susceptibility in our population.

Similarly, a strong association of rs987525 in 8q24 region could be detected after the correction for population stratification. Barbalha sample reached significant association after the test for stratification (0.005), whereas only Santarém presented significant association assuming no structure ($P=0.03$). Again, Barbalha arises with the most robust signal, reinforcing our assumption of stronger genetic effects in this region. In contrast to *IRF6*, association for this SNP also could be detected for the whole sample analysis, suggesting a stronger effect of 8q24 region. Evidences of a major role for this region also have been collected in Europeans populations [Blanton et al. 2010b; Grant et al. 2009; Nikopensius et al. 2009], but the risk conferred by this SNP in Brazilians were lower than those found in the other studies.

These results are consistent with heritability ratings (H^2) of NS CL±P previously estimated in these regions (Brito et al., 2011 – in press). In that study, the highest genetic contribution was found in Barbalha (85%), followed by Santarém (71%), Rio de Janeiro (70%), Fortaleza (64%) and Maceió (45%). In general, the association strength found in this study seems to be related with the H^2 estimates, and these results support our hypothesis that low H^2 populations might dilute the genetic effect in an association analysis in which different populations are pooled together,

requiring thus larger sample size to reach statistical significance; populations with high H^2 , on the other hand, seem to be more sensible for detection of association. The data presented here illustrates that a previous characterization of the sample regarding to disease heritability or familial aggregation can reveal a promising subgroup, in which association study has greater chances of finding low risk variants which signal is lost in larger samples with individuals from different regions, even if within a same country supposedly with homogeneous ancestry.

The data presented here related to *IRF6* intronic SNP will be reviewed as soon as its genotypes are confirmed through sequencing.

References

- Bastos-Rodrigues L, Pimenta JR, Pena SD. 2006. The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Ann Hum Genet* 70(Pt 5):658-65.
- Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. 2010. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet* 42(6):525-9.
- Birnbaum S, Ludwig KU, Reutter H, Herms S, de Assis NA, Diaz-Lacava A, Barth S, Lauster C, Schmidt G, Scheer M, Saffar M, Martini M, Reich RH, Schiefke F, Hemprich A, Potzsch S, Potzsch B, Wienker TF, Hoffmann P, Knapp M, Kramer FJ, Nothen MM, Mangold E. 2009a. *IRF6* gene variants in Central European patients with non-syndromic cleft lip with or without cleft palate. *Eur J Oral Sci* 117(6):766-9.
- Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster

- C, Schmidt G, Scheer M, Braumann B, Berge SJ, Reich RH, Schiefke F, Hemprich A, Potzsch S, Steegers-Theunissen RP, Potzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nothen MM, Mangold E. 2009b. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet* 41(4):473-7.
- Blanton SH, Burt A, Garcia E, Mulliken JB, Stal S, Hecht JT. 2010a. Ethnic Heterogeneity of IRF6 AP-2a Binding Site Promoter SNP Association With Nonsyndromic Cleft Lip and Palate. *Cleft Palate Craniofac J* 47(6):574-7.
- Blanton SH, Burt A, Stal S, Mulliken JB, Garcia E, Hecht JT. 2010b. Family-based study shows heterogeneity of a susceptibility locus on chromosome 8q24 for nonsyndromic cleft lip and palate. *Birth Defects Res A Clin Mol Teratol* 88(4):256-9.
- Blanton SH, Cortez A, Stal S, Mulliken JB, Finnell RH, Hecht JT. 2005. Variation in IRF6 contributes to nonsyndromic cleft lip and palate. *Am J Med Genet A* 137A(3):259-62.
- Centro de Documentação e Disseminação de Informações (Brazil). 2007. Brazil, 500 years of settlement. Rio de Janeiro: IBGE-Instituto Brasileiro de Geografia e Estatística. 229 p. p.
- Chiquet BT, Hashmi SS, Henry R, Burt A, Mulliken JB, Stal S, Bray M, Blanton SH, Hecht JT. 2009. Genomic screening identifies novel linkages and provides further evidence for a role of MYH9 in nonsyndromic cleft lip and palate. *Eur J Hum Genet* 17(2):195-204.
- Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, Ingle C, Beazley C, Gutierrez Arcelus M, Sekowska M, Gagnebin M, Nisbett J, Deloukas P, Dermitzakis ET, Antonarakis SE. 2009. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 325(5945):1246-50.
- Excoffier L, Lischer HEL. 2010 Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10:564-567.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164(4):1567-87.
- Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7(4):574-578.
- Ghassibe M, Bayet B, Revencu N, Verellen-Dumoulin C, Gillerot Y, Vanwijck R, Vikkula M. 2005. Interferon regulatory factor-6: a gene predisposing to

- isolated cleft lip with or without cleft palate in the Belgian population. *Eur J Hum Genet* 13(11):1239-42.
- Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield JP, Glessner JT, Thomas KA, Garriss M, Frackelton EC, Otieno FG, Chiavacci RM, Nah HD, Kirschner RE, Hakonarson H. 2009. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *J Pediatr* 155(6):909-13.
- Ingersoll RG, Hetmanski J, Park JW, Fallin MD, McIntosh I, Wu-Chou YH, Chen PK, Yeow V, Chong SS, Cheah F, Sull JW, Jee SH, Wang H, Wu T, Murray T, Huang S, Ye X, Jabs EW, Redett R, Raymond G, Scott AF, Beaty TH. 2010. Association between genes on chromosome 4p16 and non-syndromic oral clefts in four populations. *Eur J Hum Genet* 18(6):726-32.
- Ingraham CR, Kinoshita A, Kondo S, Yang B, Sajan S, Trout KJ, Malik MI, Dunnwald M, Goudy SL, Lovett M, Murray JC, Schutte BC. 2006. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). *Nat Genet* 38(11):1335-40.
- Jugessur A, Rahimov F, Lie RT, Wilcox AJ, Gjessing HK, Nilsen RM, Nguyen TT, Murray JC. 2008. Genetic variants in IRF6 and the risk of facial clefts: single-marker and haplotype-based analyses in a population-based case-control study of facial clefts in Norway. *Genet Epidemiol* 32(5):413-24.
- Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. 2002. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat Genet* 32(2):285-9.
- Loffredo LCM, Freitas JAS, Grigolli AAG. 2001. Prevalence of oral clefts from 1975 to 1994, Brazil. *Rev Saúde Pública* 35(6):571-5.
- Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B, Scheer M, Reich RH, Hemprich A, Potzsch S, Blaumeiser B, Moebus S, Krawczak M, Schreiber S, Meitinger T, Wichmann HE, Steegers-Theunissen RP, Kramer FJ, Cichon S, Propping P, Wienker TF, Knapp M, Rubini M, Mossey PA, Hoffmann P, Nothen MM. 2009. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nat Genet* 42(1):24-6.
- Marazita ML, Field LL, Cooper ME, Tobias R, Maher BS, Peanchitlertkajorn S, Liu YE. 2002. Genome scan for loci involved in cleft lip with or without cleft palate, in Chinese multiplex families. *Am J Hum Genet* 71(2):349-64.

- Menegotto BG, Salzano FM. 1991. Epidemiology of oral clefts in a large South American sample. *Cleft Palate Craniofac J* 28(4):373-6; discussion 376-7.
- Mitchell LE, Risch N. 1992. Mode of inheritance of nonsyndromic cleft lip with or without cleft palate: a reanalysis. *Am J Hum Genet* 51(2):323-32.
- Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, Johnson MK, Brauer D, Krahm K, Daack-Hirsch S, L'Heureux J, Valencia-Ramirez C, Rivera D, Lopez AM, Moreno MA, Hing A, Lammer EJ, Jones M, Christensen K, Lie RT, Jugessur A, Wilcox AJ, Chines P, Pugh E, Doheny K, Arcos-Burgos M, Marazita ML, Murray JC, Lidral AC. 2009. FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. *Hum Mol Genet* 18(24):4879-96.
- Nagem Filho H, Martins DR. 1968. [Teamwork in the rehabilitation of patients with lesions of lip and palate]. *Estomatol Cult* 2(1):127-32.
- Nikopensius T, Ambrozaityte L, Ludwig KU, Birnbaum S, Jagomagi T, Saag M, Matuleviciene A, Linkeviciene L, Herms S, Knapp M, Hoffmann P, Nothen MM, Kucinskas V, Metspalu A, Mangold E. 2009. Replication of novel susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24 in Estonian and Lithuanian patients. *Am J Med Genet A* 149A(11):2551-3.
- Paranaíba LM, Bufalino A, Martelli-Junior H, de Barros LM, Graner E, Coletta RD. 2010. Lack of association between IRF6 polymorphisms (rs2235371 and rs642961) and non-syndromic cleft lip and/or palate in a Brazilian population. *Oral Dis* 16(2):193-7.
- Park JW, McIntosh I, Hetmanski JB, Jabs EW, Vander Kolk CA, Wu-Chou YH, Chen PK, Chong SS, Yeow V, Jee SH, Park BY, Fallin MD, Ingersoll R, Scott AF, Beaty TH. 2007. Association between IRF6 and nonsyndromic cleft lip with or without cleft palate in four populations. *Genet Med* 9(4):219-27.
- Pegelow M, Peyrard-Janvid M, Zucchelli M, Fransson I, Larson O, Kere J, Larsson C, Karsten A. 2008. Familial non-syndromic cleft lip and palate--analysis of the IRF6 gene and clinical phenotypes. *Eur J Orthod* 30(2):169-75.
- Pritchard JK, Stephens M, Donnelly P. 2000a. Inference of population structure using multilocus genotype data. *Genetics* 155(2):945-59.
- Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. 2000b. Association mapping in structured populations. *Am J Hum Genet* 67(1):170-81.
- Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M, Christensen K, Bille C, Melbye M, Jugessur A, Lie RT, Wilcox AJ, Fitzpatrick DR, Green ED, Mossey PA, Little J, Steegers-Theunissen RP, Pennacchio LA, Schutte BC, Murray JC. 2008. Disruption of an AP-2alpha

- binding site in an IRF6 enhancer is associated with cleft lip. *Nat Genet* 40(11):1341-7.
- Richardson RJ, Dixon J, Malhotra S, Hardman MJ, Knowles L, Boot-Handford RP, Shore P, Whitmarsh A, Dixon MJ. 2006. *Irf6* is a key determinant of the keratinocyte proliferation-differentiation switch. *Nat Genet* 38(11):1329-34.
- Scapoli L, Palmieri A, Martinelli M, Pezzetti F, Carinci P, Tognon M, Carinci F. 2005. Strong evidence of linkage disequilibrium between polymorphisms at the IRF6 locus and nonsyndromic cleft lip with or without cleft palate, in an Italian population. *Am J Hum Genet* 76(1):180-3.
- Shen GQ, Abdullah KG, Wang QK. 2009. The TaqMan method for SNP genotyping. *Methods Mol Biol* 578:293-306.
- Stephens M, Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 73(5):1162-9.
- Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978-89.
- Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. 2001. IRF family of transcription factors as regulators of host defense. *Annu Rev Immunol* 19:623-55.
- van den Boogaard MJ, Dorland M, Beemer FA, van Amstel HK. 2000. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. *Nat Genet* 24(4):342-3.
- Vanderas AP. 1987. Incidence of cleft lip, cleft palate, and cleft lip and palate among races: a review. *Cleft Palate J* 24(3):216-25.
- Weber JL, David D, Heil J, Fan Y, Zhao C, Marth G. 2002. Human diallelic insertion/deletion polymorphisms. *Am J Hum Genet* 71(4):854-62.
- Zuccherro TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. 2004. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med* 351(8):769-80.

Capítulo 5. Discussão Geral e Conclusões

O trabalho aqui apresentado visa à compreensão de parte do mecanismo etiológico das FL±Ps NS na população brasileira e integra duas importantes áreas da genética: a epidemiologia e a análise molecular. Tanto o estudo de herdabilidade quanto o da associação dos polimorfismos colocaram Barbalha como uma região promissora para o estudo das fissuras, e os pacientes dessa região devem ser foco mais detalhado de futuros estudos. Essa integração é especialmente valiosa no momento em que muitos dos recentes estudos de associação de varredura genômica têm dificuldades para encontrar variantes de grandes efeitos (individuais e cumulativos) que possam explicar o total do componente genético das FL±Ps NS. Nesses estudos, é bastante comum a utilização de uma amostra grande (frequentemente milhares de pacientes de diferentes populações) para aumentar o poder da abordagem de detectar variantes de efeitos menores. Porém, a presença de variantes de efeitos específicos para determinadas populações, bem como as diferentes contribuições dos fatores genéticos em cada população, pode fazer com que esses estudos busquem variantes comuns entre indivíduos que não compartilham tantas. Assim, restringir a análise a um grupo mais homogêneo e enriquecido de componentes genéticos pode ser uma alternativa interessante à mistura indiscriminada de populações. Essa alternativa é especialmente atraente considerando o alto custo de uma análise de varredura genômica em milhares de indivíduos; as informações epidemiológicas obtidas neste trabalho contribuirão de maneira significativa para um direcionamento do estudo aos pacientes mais promissores de se identificar variantes

de suscetibilidade. A ideia, portanto, é a de que uma amostra menor, porém qualitativamente melhor do ponto de vista genético, ofereça um grande poder estatístico para detectar variantes de efeitos menores. Para a identificação destas, um estudo de associação de varredura genômica neste momento mostra-se extremamente importante de ser realizado, visto que é uma abordagem que vem produzindo resultados importantes em outras populações.

O presente trabalho ainda admite duas análises importantes de serem realizadas em um futuro próximo, além das que foram apresentadas aqui. Uma referente ao polimorfismo rs2236906, em um íntron do *IRF6*, e outra ao fenótipo das FL±Ps NS.

O polimorfismo rs2236906 (intrônico do *IRF6*), genotipado para toda a amostra descrita no capítulo 4, apresentou-se em forte desequilíbrio de Hardy-Weinberg. Por conta disso, optamos por excluí-lo provisoriamente da interpretação dos resultados, para não correremos o risco de enviesarmos os resultados devido a um possível problema metodológico. A análise desse SNP é importante principalmente por dois motivos. Primeiramente, ele próprio é um SNP funcional, já caracterizado como um *QTL* (*Quantitative trait locus*), dentro de um gene altamente correlacionado com as fissuras não sindrômicas, sendo assim um potencial SNP candidato. Segundo, a associação deste SNP funcionaria como um controle para os resultados obtidos do SNP do *enhancer* do *IRF6* (rs642961), pois, uma vez que esse último é constantemente descrito como a principal variante de suscetibilidade às FL±S NS, esperaríamos que a associação detectada para este fosse maior do que aquela detectada para o SNP intrônico (porém, no caso de haver forte desequilíbrio de ligação entre os dois SNPs na nossa amostra, o papel do SNP intrônico não poderá ser testado por essa abordagem, visto que o sinal de associação captado por ele poderia

refletir o efeito do SNP do *enhancer* do *IRF6*). No caso de não haver desequilíbrio de ligação entre eles, um resultado oposto indicaria que, ao contrário do que a maioria dos trabalhos propõe, esse SNP não é o mais determinante da região. Por isso, investiremos na busca do motivo pelo qual esse polimorfismo encontra-se em desequilíbrio de Hardy-Weinberg e na confirmação dos genótipos obtidos por sequenciamento direto da região, para posterior repetição das análises.

Estudos epidemiológicos realizados desde a década de 1960 apontam para uma diferença etiológica entre fissura de palato e fissura labial com ou sem fissura de palato. Alguns estudos moleculares e epidemiológicos recentes, entretanto, sugerem que o grupo “fissura labial com ou sem fissura de palato não sindrômica” (FL±P NS) pode, na verdade, conter dois grupos etiológicamente distintos (mas admitindo certa sobreposição de causas): fissura de lábio com fissura de palato (FL+P) e fissura de lábio isolada (FL). Evidências nesse sentido foram sugeridas em um estudo epidemiológico na Noruega (Harville *et al.*, 2005). Embora os próprios autores deste estudo considerem que muitos dos resultados obtidos são explicados simplesmente pela ideia de que FL+P seja somente uma forma mais grave de FL, alguns resultados foram curiosos, como, por exemplo, a predominância masculina em FL+P comparada a FL, e o risco maior de pais consanguíneos terem um filho com FL em relação à FL+P. Evidências moleculares para essa divisão foram observadas em Rahimov *et al.*, (2008), quando foi encontrada associação diferenciada do SNP do *enhancer* do *IRF6* quando a amostra de casos foi estratificada pelo fenótipo (FL+P e FL). Jugessur *et al.*, (2008), ao analisar seis SNPs na região do *IRF6*, encontrou forte associação na amostra de FL+P, mas não na de FL, reforçando a hipótese. Um estudo molecular mais recente na Noruega também encontrou associações diferenciadas entre os grupos: os genes *FGF12*, *IRF6*, *CX43* e *VCL* tiveram sinal detectado somente na

amostra de FL, e não na de FL+P, e os autores sugeriram que ao menos uma parte dos casos de FL é etiologicamente distinta de FL+P (Jugessur *et al.*, 2011). Com base nessa nova divisão etiológica, ainda pouco corroborada, é de nosso interesse futuro investigar na nossa amostra uma possível variação entre esses grupos na força da associação detectada e no risco que os SNPs conferem para a malformação. A importância da validação ou refutação dessa nova divisão fenotípica se dá tanto na análise molecular, uma vez que a fenotipagem é um passo de fundamental importância e indivíduos mal categorizados podem comprometer a análise de resultados, quanto na análise epidemiológica, uma vez que a maioria dos estudos realizados até então não separam esses dois grupos. Haverá ainda, caso confirmada essa divisão, um impacto no aconselhamento genético, uma vez que diferentes riscos de recorrências poderão ser revelados.

Por fim, este trabalho se utilizou de técnicas de discriminação de ancestralidade para testar a associação das regiões candidatas, procedimento que parece ser essencial de ser realizado na nossa população, haja vista os diferentes resultados produzidos pelos testes de qui-quadrado admitindo ausência de estratificação e posteriormente corrigindo para a estratificação. A discriminação ancestral foi realizada a partir de um painel de marcadores de inserção-deleção criado a partir de critérios estabelecidos em Bastos-Rodrigues *et al.* (2006), como diferentes frequências alélicas entre as populações ancestrais e frequências alélicas em europeus próximas de 0.5 (medida para evitar marcadores pouco polimórficos). Esse sistema revelou-se satisfatório na separação das ancestralidades e produziu resultados concordantes com trabalhos já realizados com esse intuito. Santarém, por exemplo, cidade localizada em uma região onde ainda existem muitas comunidades indígenas e onde o componente genômico indígena é maior comparado a outras regiões

brasileiras, apresentou a maior contribuição ameríndia (31%). Em um estudo também com indels, Santos *et al.* (2010) encontrou em Belém-PA uma contribuição ameríndia similar (27%), que também foi a maior daquele estudo, que contou também com populações do Sul do Brasil. A utilidade da genotipagem desse painel de indels nos pacientes vai além do presente trabalho, uma vez que qualquer variante candidata às fissuras pode ser rapidamente testada na amostra aqui descrita.

Resumo

Fissura labial com ou sem fissura de palato não sindrômica (FL±P NS) é uma doença complexa que afeta 1:700 indivíduos no mundo. A busca das causas genéticas dessa malformação é dificultada pelo padrão multifatorial de herança e pela heterogeneidade genética, sendo que o gene *IRF6* e a região 8q24 são os *loci* de associação mais corroborada. A estratificação populacional é um problema adicional a ser considerado em estudos de caso-controle na população brasileira. No intuito de caracterizar variáveis que possam interferir na busca dos fatores de risco, realizamos em um primeiro estudo uma avaliação epidemiológica de pacientes de cinco localidades do país (Santarém-PA, Barbalha-CE, Fortaleza-CE, Maceió-AL e Rio de Janeiro-RJ) . Este estudo revelou Barbalha como a região onde a genética desempenha papel mais determinante (herdabilidade = 85%; risco de recorrência = 2,2-2,8%); Maceió, por outro lado, foi a região de menor influência genética (herdabilidade = 45%; risco de recorrência = 0,6-0,7%). Ainda, a consangüinidade não mostrou um mecanismo importante para explicar estes resultados. Em um segundo estudo, realizamos a caracterização da ancestralidade da amostra, com o intuito de estabelecermos parâmetros para serem utilizados em futuros estudos de associação na nossa população. Para testarmos as nossas hipóteses realizamos um estudo de caso-controle com os SNPs mais corroborados nos dois *loci* em outras populações: rs642961 em um enhancer do gene *IRF6* e rs987525 na região 8q24. Verificamos que quando realizamos um teste de associação para os SNPs com correção para estrutura populacional obtivemos resultados consistentes com as estimativas de herdabilidade, uma vez que Barbalha foi a única região de associação

positiva para os SNPs. Apesar de estes SNPs terem sido estudados em outras populações, este é o primeiro relato de associação destes SNPs na população brasileira. Ainda, o estudo molecular revelou a importância da caracterização da estrutura populacional por meio de marcadores informativos de ancestralidade em estudos de caso-controle na nossa população, uma vez que resultados diferentes puderam ser observados em análise assumindo ausência de estratificação. Este trabalho fornece importantes bases para a identificação de novos genes de predisposição às FL±P NS na população brasileira, pois permite um direcionamento para as populações de maior contribuição genética nas abordagens que virão a ser realizadas.

Abstract

Nonsyndromic cleft lip with or without cleft palate (NS CL±P) is a complex disease with worldwide incidence estimated as 1:700. The multifactorial model of inheritance and the genetic heterogeneity difficult the search for the genetic causes of NS CL±P, and, of all *loci*, *IRF6* gene and 8q24 gene desert are the two most associated. The population stratification constitutes an additional problem to be considered in case-control studies in Brazil. In order to identify the factors that may interfere in the hunting of risk factors, we carried out, in a first moment, an epidemiologic evaluation of patients from five locations in Brazil (Santarém-PA, Barbalha-CE, Fortaleza-CE, Maceió-AL and Rio de Janeiro-RJ). This study put Barbalha as the region where genetic factors play the more determinant role (heritability = 85%; recurrence risk = 2.2-2.8%); Maceió, on the other hand, was the region of less genetic contribution to the disease (heritability = 45%; recurrence risk = 0.6-0.7%). In addition, consanguinity did not appear to influence these results. In a second study, we characterized the sample ancestry, in order to establish parameters for future association studies in our population. To test our hypothesis, we carried out a case-control study with the SNPs which are the most corroborated in other populations: rs642961 (in an *IRF6* enhancer), and rs987525 (in 8q24). We verified that, when a structured association test was performed, we obtained results that are consistent with heritability estimates, since Barbalha was the only region with positive association for both SNPs. This was the first time that a positive association for these markers was reported in a Brazilian population. In addition, the molecular analysis evidenced the importance of an individual characterization with ancestry informative markers when

performing a case-control study in this population, since different results were obtained from the analyses assuming no stratification and correcting its effect. This study provides important bases for the identification of new susceptibility variants to NS CL±P in Brazilian population, since targeting in the populations of highest genetic contribution to the disease will be possible in the forthcoming studies, increasing the power of the study.

Referências Bibliográficas

Altschuler, D., Daly, M. J. e Lander, E. S. Genetic Mapping in Human Disease. *Science*, v.322, n. 5903, Nov 7, p.881-8. 2008

Barbujani, G., A. Magagni, E. Minch e L. L. Cavalli-Sforza. An apportionment of human DNA diversity. *Proc Natl Acad Sci U S A*, v.94, n.9, Apr 29, p.4516-9. 1997.

Baroni, T., C. Bellucci, C. Lilli, F. Pezzetti, F. Carinci, E. Lumare, A. Palmieri, G. Stabellini e M. Bodo. Human cleft lip and palate fibroblasts and normal nicotine-treated fibroblasts show altered in vitro expressions of genes related to molecular signaling pathways and extracellular matrix metabolism. *J Cell Physiol*, v.222, n.3, Mar, p.748-56. 2010.

Bastos-Rodrigues, L., J. R. Pimenta e S. D. Pena. The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Ann Hum Genet*, v.70, n.Pt 5, Sep, p.658-65. 2006.

Beaty, T. H., J. C. Murray, M. L. Marazita, R. G. Munger, I. Ruczinski, J. B. Hetmanski, K. Y. Liang, T. Wu, T. Murray, M. D. Fallin, R. A. Redett, G. Raymond, H. Schwender, S. C. Jin, M. E. Cooper, M. Dunnwald, M. A. Mansilla, E. Leslie, S. Bullard, A. C. Lidral, L. M. Moreno, R. Menezes, A. R. Vieira, A. Petrin, A. J. Wilcox, R. T. Lie, E. W. Jabs, Y. H. Wu-Chou, P. K. Chen, H. Wang, X. Ye, S. Huang, V. Yeow, S. S. Chong, S. H. Jee, B. Shi, K. Christensen, M. Melbye, K. F. Doheny, E. W. Pugh, H. Ling, E. E. Castilla, A. E. Czeizel, L. Ma, L. L. Field, L. Brody, F. Pangilinan, J. L. Mills, A. M. Molloy, P. N. Kirke, J. M. Scott, M. Arcos-Burgos e A. F. Scott. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet*, v.42, n.6, Jun, p.525-9. 2010.

Bender, P. L. Genetics of cleft lip and palate. *J Pediatr Nurs*, v.15, n.4, Aug, p.242-9. 2000.

Birnbaum, S., K. U. Ludwig, H. Reutter, S. Herms, M. Steffens, M. Rubini, C. Baluardo, M. Ferrian, N. Almeida De Assis, M. A. Alblas, S. Barth, J. Freudenberg, C. Lauster, G. Schmidt, M. Scheer, B. Braumann, S. J. Berge, R. H. Reich, F. Schiefke, A. Hemprich, S. Potzsch, R. P. Steegers-Theunissen, B. Potzsch, S. Moebus, B. Horsthemke, F. J. Kramer, T. F. Wienker, P. A. Mossey, P. Propping, S. Cichon, P. Hoffmann, M. Knapp, M. M. Nothen e E. Mangold. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet*, v.41, n.4, Apr, p.473-7. 2009.

Bosi, G., R. Evangelisti, V. Valeno, F. Carinci, F. Pezzetti, C. Calastrini, M. Bodo e P. Carinci. Diphenylhydantoin affects glycosaminoglycans and collagen production by human fibroblasts from cleft palate patients. J Dent Res, v.77, n.8, Aug, p.1613-21. 1998.

Brown, N. L., L. Knott, E. Halligan, S. J. Yarram, J. P. Mansell e J. R. Sandy. Microarray analysis of murine palatogenesis: temporal expression of genes during normal palate development. Dev Growth Differ, v.45, n.2, Apr, p.153-65. 2003.

Bueno, D. F., D. Y. Sunaga, G. S. Kobayashi, M. Agüena, C. E. Raposo-Amaral, C. Masotti, L. A. Cruz, P. L. Pearson e M. R. Passos-Bueno. Human stem cell cultures from cleft lip/palate patients show enrichment of transcripts involved in extracellular matrix modeling by comparison to controls. Stem Cell Rev, v.7, n.2, Jun, p.446-57. 2010.

Calzolari, E., M. Milan, G. B. Cavazzuti, G. Cocchi, E. Gandini, C. Magnani, M. Moretti, G. P. Garani, G. P. Salvioli e S. Volpato. Epidemiological and genetic study of 200 cases of oral cleft in the Emilia Romagna region of northern Italy. Teratology, v.38, n.6, Dec, p.559-64. 1988.

Carinci, F., L. Scapoli, A. Palmieri, I. Zollino e F. Pezzetti. Human genetic factors in nonsyndromic cleft lip and palate: an update. Int J Pediatr Otorhinolaryngol, v.71, n.10, Oct, p.1509-19. 2007.

Centers of Disease control and prevention. Economic costs of birth defects and cerebral palsy – United States, 1992. MMWR Morb Mortal Wkly Rep. 1995;44:694-699

Chevrier, C., B. Dananche, M. Bahuau, A. Nelva, C. Herman, C. Francannet, E. Robert-Gnansia e S. Cordier. Occupational exposure to organic solvent mixtures during pregnancy and the risk of non-syndromic oral clefts. Occup Environ Med, v.63, n.9, Sep, p.617-23. 2006.

Chiquet, B. T., S. S. Hashmi, R. Henry, A. Burt, J. B. Mulliken, S. Stal, M. Bray, S. H. Blanton e J. T. Hecht. Genomic screening identifies novel linkages and provides further evidence for a role of MYH9 in nonsyndromic cleft lip and palate. Eur J Hum Genet, v.17, n.2, Feb, p.195-204. 2009.

Christensen, K. e P. Fogh-Andersen. Cleft lip (+/- cleft palate) in Danish twins, 1970-1990. Am J Med Genet, v.47, n.6, Nov 1, p.910-6. 1993.

Deroo, L. A., A. J. Wilcox, C. A. Drevon e R. T. Lie. First-trimester maternal alcohol consumption and the risk of infant oral clefts in Norway: a population-based case-control study. Am J Epidemiol, v.168, n.6, Sep 15, p.638-46. 2008.

Dixon, M. J., M. L. Marazita, T. H. Beaty e J. C. Murray. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet, v.12, n.3, Mar, p.167-78. 2011.

Dolk, H., J. Jentink, M. Loane, J. Morris e L. T. De Jong-Van Den Berg. Does lamotrigine use in pregnancy increase orofacial cleft risk relative to other malformations? Neurology, v.71, n.10, Sep 2, p.714-22. 2008.

Excoffier, L. e G. Hamilton. Comment on "Genetic structure of human populations". Science, v.300, n.5627, Jun 20, p.1877; author reply 1877. 2003.

Francomano, C. A., R. M. Liberfarb, T. Hirose, I. H. Maumenee, E. A. Streeten, D. A. Meyers e R. E. Pyeritz. The Stickler syndrome: evidence for close linkage to the structural gene for type II collagen. Genomics, v.1, n.4, Dec, p.293-6. 1987.

Gorlin, R. J., M. M. Cohen Jr e R. C. M. Hennekam. Syndromes of the Head and Neck, 4th edition. New York: Oxford University Press. 2001. 1283 p.

Grant, S. F., K. Wang, H. Zhang, W. Glaberson, K. Annaiah, C. E. Kim, J. P. Bradfield, J. T. Glessner, K. A. Thomas, M. Garriss, E. C. Frackelton, F. G. Otieno, R. M. Chiavacci, H. D. Nah, R. E. Kirschner e H. Hakonarson. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr, v.155, n.6, Dec, p.909-13. 2009.

Harville, E. W., A. J. Wilcox, R. T. Lie, H. Vindenes e F. Abyholm. Cleft lip and palate versus cleft lip only: are they distinct defects? Am J Epidemiol, v.162, n.5, Sep 1, p.448-53. 2005.

Hernandez-Diaz, S., M. M. Werler, A. M. Walker e A. A. Mitchell. Folic acid antagonists during pregnancy and the risk of birth defects. N Engl J Med, v.343, n.22, Nov 30, p.1608-14. 2000.

Hu, D. N., J. H. Li, H. Y. Chen, H. S. Chang, B. X. Wu, Z. K. Lu, D. Z. Wang e X. G. Liu. Genetics of cleft lip and cleft palate in China. Am J Hum Genet, v.34, n.6, Nov, p.999-1002. 1982.

Ingersoll, R. G., J. Hetmanski, J. W. Park, M. D. Fallin, I. McIntosh, Y. H. Wu-Chou, P. K. Chen, V. Yeow, S. S. Chong, F. Cheah, J. W. Sull, S. H. Jee, H. Wang, T. Wu, T. Murray, S. Huang, X. Ye, E. W. Jabs, R. Redett, G. Raymond, A. F. Scott e T. H. Beaty. Association between genes on chromosome 4p16 and non-syndromic oral clefts in four populations. Eur J Hum Genet, v.18, n.6, Jun, p.726-32. 2010.

Jakobsen, L. P., R. Borup, J. Vestergaard, L. A. Larsen, K. Lage, L. L. Maroun, I. Kjaer, C. U. Niemann, M. Andersen, M. A. Knudsen, K. Mollgard e N. Tommerup. Expression analyses of human cleft palate tissue suggest a role for osteopontin and immune related factors in palatal development. Exp Mol Med, v.41, n.2, Feb 28, p.77-85. 2009.

Jiang, R., J. O. Bush e A. C. Lidral. Development of the upper lip: morphogenetic and molecular mechanisms. Dev Dyn, v.235, n.5, May, p.1152-66. 2006.

Jugessur, A., P. G. Farlie e N. Kilpatrick. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. Oral Dis, v.15, n.7, Oct, p.437-53. 2009.

Jugessur, A. e J. C. Murray. Orofacial clefting: recent insights into a complex trait. Curr Opin Genet Dev, v.15, n.3, Jun, p.270-8. 2005.

Jugessur, A., F. Rahimov, R. T. Lie, A. J. Wilcox, H. K. Gjessing, R. M. Nilsen, T. T. Nguyen e J. C. Murray. Genetic variants in IRF6 and the risk of facial clefts: single-marker and haplotype-based analyses in a population-based case-control study of facial clefts in Norway. Genet Epidemiol, v.32, n.5, Jul, p.413-24. 2008.

Jugessur, A., M. Shi, H. K. Gjessing, R. T. Lie, A. J. Wilcox, C. R. Weinberg, K. Christensen, A. L. Boyles, S. Daack-Hirsch, T. T. Nguyen, L. Christiansen, A. C. Lidral e J. C. Murray. Fetal genetic risk of isolated cleft lip only versus isolated cleft lip and palate: a subphenotype analysis using two population-based studies of orofacial clefts in scandinavia. Birth Defects Res A Clin Mol Teratol, v.91, n.2, Feb, p.85-92. 2011.

Kerrigan, J. J., J. P. Mansell, A. Sengupta, N. Brown e J. R. Sandy. Palatogenesis and potential mechanisms for clefting. J R Coll Surg Edinb, v.45, n.6, Dec, p.351-8. 2000.

Kim, C. A., Bertola, D. R., Albano, L. M. J. Genética na Prática Pediátrica. Ed Manole. 620p. 2010.

Kondo, S., B. C. Schutte, R. J. Richardson, B. C. Bjork, A. S. Knight, Y. Watanabe, E. Howard, R. L. De Lima, S. Daack-Hirsch, A. Sander, D. M. McDonald-McGinn, E. H. Zackai, E. J. Lammer, A. S. Aylsworth, H. H. Ardinger, A. C. Lidral, B. R. Pober, L. Moreno, M. Arcos-Burgos, C. Valencia, C. Houdayer, M. Bahau, D. Moretti-Ferreira, A. Richieri-Costa, M. J. Dixon e J. C. Murray. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet, v.32, n.2, Oct, p.285-9. 2002.

Krapels, I. P., C. Vermeij-Keers, M. Muller, A. De Klein e R. P. Steegers-Theunissen. Nutrition and genes in the development of orofacial clefting. Nutr Rev, v.64, n.6, Jun, p.280-8. 2006.

Letra, A., R. Menezes, R. F. Fonseca, M. Govil, T. Mchenry, M. J. Murphy, J. D. Hennebold, J. M. Granjeiro, E. E. Castilla, I. M. Orioli, R. Martin, M. L. Marazita, B. C. Bjork e A. R. Vieira. Novel cleft susceptibility genes in chromosome 6q. J Dent Res, v.89, n.9, Sep, p.927-32. 2010.

Lie, R. T., A. J. Wilcox e R. Skjaerven. A population-based study of the risk of recurrence of birth defects. N Engl J Med, v.331, n.1, Jul 7, p.1-4. 1994.

Lin, Y. C., L. J. Lo, M. S. Noordhoff e Y. R. Chen. Cleft of the lip and palate in twins. Changgang Yi Xue Za Zhi, v.22, n.1, Mar, p.61-7. 1999.

Loffredo, L. C. M., J. A. S. Freitas e A. A. G. Grigolli. Prevalence of oral clefts from 1975 to 1994, Brazil. Rev Saúde Pública, v.35, n.6, p.571-5. 2001.

Mangold, E., K. U. Ludwig, S. Birnbaum, C. Baluardo, M. Ferrian, S. Herms, H. Reutter, N. A. De Assis, T. A. Chawa, M. Mattheisen, M. Steffens, S. Barth, N. Kluck, A. Paul, J. Becker, C. Lauster, G. Schmidt, B. Braumann, M. Scheer, R. H. Reich, A. Hemprich, S. Potzsch, B. Blaumeiser, S. Moebus, M. Krawczak, S. Schreiber, T. Meitinger, H. E. Wichmann, R. P. Steegers-Theunissen, F. J. Kramer, S. Cichon, P. Propping, T. F. Wienker, M. Knapp, M. Rubini, P. A. Mossey, P. Hoffmann e M. M. Nothen. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet, v.42, n.1, Jan, p.24-6. 2009.

Marazita, M. L., L. L. Field, M. E. Cooper, R. Tobias, B. S. Maher, S. Peanchitlertkajorn e Y. E. Liu. Genome scan for loci involved in cleft lip with or without cleft palate, in Chinese multiplex families. Am J Hum Genet, v.71, n.2, Aug, p.349-64. 2002.

Menegotto, B. G. e F. M. Salzano. Clustering of malformations in the families of South American oral cleft neonates. J Med Genet, v.28, n.2, Feb, p.110-3. 1991a.

Menegotto, B. G. e F. M. Salzano. Epidemiology of oral clefts in a large South American sample. Cleft Palate Craniofac J, v.28, n.4, Oct, p.373-6; discussion 376-7. 1991b.

Moreno, L. M., M. A. Mansilla, S. A. Bullard, M. E. Cooper, T. D. Busch, J. Machida, M. K. Johnson, D. Brauer, K. Krahn, S. Daack-Hirsch, J. L'heureux, C. Valencia-Ramirez, D. Rivera, A. M. Lopez, M. A. Moreno, A. Hing, E. J. Lammer, M. Jones, K. Christensen, R. T. Lie, A. Jugessur, A. J. Wilcox, P. Chines, E. Pugh, K. Doheny, M. Arcos-Burgos, M. L. Marazita, J. C. Murray e A. C. Lidral. FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. Hum Mol Genet, v.18, n.24, Dec 15, p.4879-96. 2009.

Mukhopadhyay, P., R. M. Greene, W. Zacharias, M. C. Weinrich, S. Singh, W. W. Young, Jr. e M. M. Pisano. Developmental gene expression profiling of mammalian, fetal orofacial tissue. Birth Defects Res A Clin Mol Teratol, v.70, n.12, Dec, p.912-26. 2004.

Nagem Filho, H. e D. R. Martins. [Teamwork in the rehabilitation of patients with lesions of lip and palate]. Estomatol Cult, v.2, n.1, Jan-Jun, p.127-32. 1968.

Nordstrom, R. E., T. Laatikainen, T. O. Juvonen e R. E. Ranta. Cleft-twin sets in Finland 1948-1987. Cleft Palate Craniofac J, v.33, n.4, Jul, p.340-7. 1996.

Pradat, P., E. Robert-Gnansia, G. L. Di Tanna, A. Rosano, A. Lisi e P. Mastroiacovo. First trimester exposure to corticosteroids and oral clefts. Birth Defects Res A Clin Mol Teratol, v.67, n.12, Dec, p.968-70. 2003.

Rahimov, F., M. L. Marazita, A. Visel, M. E. Cooper, M. J. Hitchler, M. Rubini, F. E. Domann, M. Govil, K. Christensen, C. Bille, M. Melbye, A. Jugessur, R. T. Lie, A. J. Wilcox, D. R. Fitzpatrick, E. D. Green, P. A. Mossey, J. Little, R. P. Steegers-Theunissen, L. A. Pennacchio, B. C. Schutte e J. C. Murray. Disruption of an AP-

2alpha binding site in an IRF6 enhancer is associated with cleft lip. Nat Genet, v.40, n.11, Nov, p.1341-7. 2008.

Rogers, C. R., S. M. Weinberg, T. D. Smith, F. W. Deleyiannis, M. P. Mooney e M. L. Marazita. Anatomical basis for apparent subepithelial cleft lip: a histological and ultrasonographic survey of the orbicularis oris muscle. Cleft Palate Craniofac J, v.45, n.5, Sep, p.518-24. 2008.

Romualdi, C., D. Balding, I. S. Nasidze, G. Risch, M. Robichaux, S. T. Sherry, M. Stoneking, M. A. Batzer e G. Barbujani. Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. Genome Res, v.12, n.4, Apr, p.602-12. 2002.

Rosenberg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky e M. W. Feldman. Genetic structure of human populations. Science, v.298, n.5602, Dec 20, p.2381-5. 2002.

Santos, N. P., E. M. Ribeiro-Rodrigues, A. K. Ribeiro-Dos-Santos, R. Pereira, L. Gusmao, A. Amorim, J. F. Guerreiro, M. A. Zago, C. Matte, M. H. Hutz e S. E. Santos. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. Hum Mutat, v.31, n.2, Feb, p.184-90. 2010.

Schutte, B. C. e J. C. Murray. The many faces and factors of orofacial clefts. Hum Mol Genet, v.8, n.10, p.1853-9. 1999.

Stanier, P. e G. E. Moore. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. Hum Mol Genet, v.13 Spec No 1, Apr 1, p.R73-81. 2004.

Stoll, C., Y. Alembik, B. Dott e M. P. Roth. Epidemiological and genetic study in 207 cases of oral clefts in Alsace, north-eastern France. J Med Genet, v.28, n.5, May, p.325-9. 1991.

Tenconi, R., M. Clementi e L. Turolla. Theoretical recurrence risks for cleft lip derived from a population of consecutive newborns. J Med Genet, v.25, n.4, Apr, p.243-6. 1988.

Tiwari, H. K., J. Barnholtz-Sloan, N. Wineinger, M. A. Padilla, L. K. Vaughan e D. B. Allison. Review and evaluation of methods correcting for population stratification with a focus on underlying statistical principles. Hum Hered, v.66, n.2, p.67-86. 2008.

Tolarova, M. M. e J. Cervenka. Classification and birth prevalence of orofacial clefts. Am J Med Genet, v.75, n.2, Jan 13, p.126-37. 1998.

Van Den Boogaard, M. J., M. Dorland, F. A. Beemer e H. K. Van Amstel. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. Nat Genet, v.24, n.4, Apr, p.342-3. 2000.

Van Der Woude, A. Fistula Labii Inferioris Congenita and Its Association with Cleft Lip and Palate. Am. J. Hum. Genet. 6: 244-56 p. 1954.

Vanderas, A. P. Incidence of cleft lip, cleft palate, and cleft lip and palate among races: a review. Cleft Palate J, v.24, n.3, Jul, p.216-25. 1987.

Weinberg, S. M., C. A. Brandon, T. H. Mchenry, K. Neiswanger, F. W. Deleyiannis, J. E. De Salamanca, E. E. Castilla, A. E. Czeizel, A. R. Vieira e M. L. Marazita. Rethinking isolated cleft palate: evidence of occult lip defects in a subset of cases. Am J Med Genet A, v.146A, n.13, Jul 1, p.1670-5. 2008.

Wilcox, A. J., R. T. Lie, K. Solvoll, J. Taylor, D. R. Mcconnaughey, F. Abyholm, H. Vindenes, S. E. Vollset e C. A. Drevon. Folic acid supplements and risk of facial clefts: national population based case-control study. Bmj, v.334, n.7591, Mar 3, p.464. 2007.

Wyszynski, D. F. e T. H. Beaty. Review of the role of potential teratogens in the origin of human nonsyndromic oral clefts. Teratology, v.53, n.5, May, p.309-17. 1996.

Zuccherro, T. M., M. E. Cooper, B. S. Maher, S. Daack-Hirsch, B. Nepomuceno, L. Ribeiro, D. Caprau, K. Christensen, Y. Suzuki, J. Machida, N. Natsume, K. Yoshiura, A. R. Vieira, I. M. Orioli, E. E. Castilla, L. Moreno, M. Arcos-Burgos, A. C. Lidral, L. L. Field, Y. E. Liu, A. Ray, T. H. Goldstein, R. E. Schultz, M. Shi, M. K. Johnson, S. Kondo, B. C. Schutte, M. L. Marazita e J. C. Murray. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. N Engl J Med, v.351, n.8, Aug 19, p.769-80. 2004.

Web Sites

Arlequin 3.5 (software): <http://cmpg.unibe.ch/software/arlequin35/>

Marshfield database: <http://research.marshfieldclinic.org/genetics/home/index.asp>

OMIM: <http://www.ncbi.nlm.nih.gov/omim>

Operação Sorriso: <http://www.operationsmile.org.br/novo/>

Phase (software): <http://stephenslab.uchicago.edu/software.html>

STRAT (software): <http://pritch.bsd.uchicago.edu/software/STRAT.html>

Structure (software): <http://pritch.bsd.uchicago.edu/software.html>