Carolina Malcher Amorim de Carvalho Silva

Investigação do papel de SNVs (*single nucleotide variants*) na etiologia da fissura lábio-palatina não
sindrômica

*Investigation of the role of SNVs (single nucleotide variants) in
the etiology of nonsyndromic cleft lip with or without cleft
palate*

São Paulo

2013

Carolina Malcher Amorim de Carvalho Silva

Investigação do papel de SNVs (*single nucleotide variants*) na etiologia da fissura lábio-palatina não
sindrômica

*Investigation of the role of SNVs (single nucleotide variants) in
the etiology of nonsyndromic cleft lip with or without cleft
palate*

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.

Orientador(a): Maria Rita dos Santos
e Passos Bueno

São Paulo

2013

Ficha Catalográfica

Silva, Carolina Malcher Amorim de Carvalho

Investigação do papel de SNVs (*single nucleotide variants*) na etiologia da FL/P NS / Carolina Malcher Amorim de Carvalho ; orientadora Maria Rita dos Santos Passos-Bueno. – São Paulo, 2013. 67 f.

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

1. FL/P NS. 2. Estudo e Associação. 3. eQTL. I. Universidade de São Paulo. Instituto de Biociências. Departamento de Genética e Biologia Evolutiva. II. Título.

Comissão Julgadora:

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr.(a).

Orientador(a)

Dedicatória

À minha família,
em especial à minha mãe.

Agradecimentos

Primeiramente gostaria de agradecer à Rita, minha orientadora, que desde o primeiro contato me recebeu de braços abertos no laboratório. Por sempre me incentivar e apoiar no que quero fazer.

À minha mãe, melhor amiga, terapeuta, que sempre me faz ver o lado bom de tudo. Pelo amor, confiança, carinho, incentivo e suporte infinitos. Aos meus irmãos, meus amores, que me completam. Ao Maurício, meu pai, por sempre estar ao meu lado.

Ao Léo, que em mais uma etapa concluída está ao meu lado, me incentivando, consolando e apoiando meus devaneios.

Ao meu pai João, pela confiança e oportunidades.

Minhas avós, e em especial ao meu avô, meu exemplo.

Às minhas tias e tios, pelo eterno apoio. Em especial Ana e Gerson, pelo teto, comida e roupa lavada e todo carinho de pais que sempre tiveram comigo. E o mais importante: o vinhozinho.

Aos meus amigos, por sempre me lembrar do que é realmente importante na vida.

À todos do lab200, por tornarem o ambiente de trabalho tão agradável. Em especial à Ka e Lu (minha vida seria muito mais difícil sem os sábios conselhos e as inúmeras caronas), Jozinha, Gelsu, Dani Bueno, May, Andressa, Simone, Luquinhas, Brunão...

À Ciba, minha co-orientadora, por todo carinho, amizade e conhecimento dispensados. Por sempre me apoiar de todas as formas possíveis e imagináveis.

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 2011 e 2012 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).

A apresentação do trabalho seguiu o modelo de divisão em capítulos/artigo, modelo este aceito atualmente pelos programas de pós-graduação do Instituto de Biociências. Desta maneira, os artigos estão redigidos em inglês, por motivos referentes à publicação. As seções referentes à Introdução e Discussão gerais estão escritas em português, conforme especificado pelas regras deste Instituto.

As referências bibliográficas referentes a cada um dos artigos estão relacionadas ao final dos mesmos. Já as referências que constam na Introdução e Discussão gerais encontram-se ao final na dissertação.

Embora o novo acordo ortográfico da língua portuguesa esteja em vigência, sua obrigatoriedade será apenas em 2016. Sendo assim, optei por utilizar a regra antiga.

Índice

Capítulo 1.	Introdução Geral e Objetivos	1
1.1.	Fissura de lábio com ou sem fissura de palato	1
1.1.2.	<i>Embriologia</i>	2
1.1.3.	<i>Classificação</i>	3
1.2.	Fissura labial com ou sem fissura de palato não-sindrômica	4
1.2.1.	<i>Epidemiologia</i>	4
1.2.2.	<i>Modelo de Herança - Genética e Ambiente</i>	5
1.2.3.	<i>Fatores de risco ambientais</i>	6
1.2.4.	<i>Fatores de risco genéticos</i>	7
1.2.5.	<i>Estudos de Associação</i>	7
1.2.5.1.	<i>Loci associados à FL/P NS</i>	8
1.2.6.	<i>Expressão gênica nas doenças complexas</i>	11
1.3.	Objetivos	13
Capítulo 2.	<i>IRF6 é um fator de risco para fissura de lábio não-sindrômica na população brasileira / IRF6 is a Risk Factor for Nonsyndromic Cleft Lip in the Brazilian Population</i>	14
Capítulo 3.	<i>Região 8q24 é um locus de susceptibilidade para fissura oral não-sindrômica no Brasil / Region 8q24 is a Susceptibility Locus for Nonsyndromic Oral Clefting in Brazil</i>	22
Capítulo 4.	<i>Uso de eQTLs para elucidar o efeito de SNVs na etiologia da fissura oral não-sindrômica / Use of eQTLs to Elucidate the Effect of SNVs in the Etiology of Nonsyndromic Oral Clefting</i>	29
Capítulo 5.	Discussão Geral e Conclusões	45
	Resumo	48
	Abstract	49
	Referências Bibliográficas	50

Capítulo 1. Introdução Geral e Objetivos

1.1. Fissura de lábio com ou sem fissura de palato

Fissuras de lábio associadas ou não a fissuras de palato (FL/P) estão entre as malformações congênitas mais comuns, e caracterizam-se essencialmente pela separação incompleta das cavidades oral e nasal, admitindo um grande espectro de variabilidade clínica. A fissura pode afetar desde apenas o lábio (FL), ou até o palato (FLP). O acometimento do lábio pode ser unilateral ou bilateral, enquanto o palato pode ser acometido apenas no palato mole, ou até o palato duro (Figura 1).

Figura 1: Desenho esquemático do lábio e palato, demonstrando as variações clínicas da FL/P, envolvendo desde apenas o lábio (a,d), afetando o palato duro (b,e), até o palato mole (c, f); As figuras a, b, c mostram variações de fissura unilateral; As figuras d, e, f mostram variações de fissura bilateral (Mod. de DIXON *et al.*, 2011).

As fissuras lábio-palatinas também podem se apresentar como microformas mais brandas, de difícil percepção, como, por exemplo, uma simples cicatriz no lábio, uma descontinuidade do músculo orbicular do lábio (detectada através de ultrassonografia, WEINBERG *et al.*, 2008), fissura de palato submucosa (união imperfeita do músculo no véu palatino, mas com superfície da mucosa intacta; GORLIN *et al.*, 2001), e úvula bífida.

Como o lábio e o palato primário possuem origens embriológicas distintas do palato secundário, as FL/P são consideradas entidades etiologicamente distintas das fissuras palatinas (FP), nas quais o lábio não é afetado (MOSSEY *et al.*, 2009). Essa distinção é reforçada por dados epidemiológicos e pelo fato das duas formas geralmente não segregarem na mesma família (DIXON *et al.*, 2011).

1.1.2. Embriologia

As fissuras lábio-palatinas resultam de uma falha na fusão dos processos maxilares e das lâminas palatinas, na formação do lábio e palato, o que ocorre entre a quarta e décima segunda semanas de gestação, conforme representado na Figura 2.

Figura 2: Desenho esquemático do desenvolvimento do lábio e palato em humanos. a) Ao redor da quarta semana de desenvolvimento embrionário ocorre o desenvolvimento da proeminência frontonasal (observada em azul), os processos maxilares (verde), os processos mandibulares pareados (rosa), e a cavidade oral primitiva ao centro (roxo); b) Ao redor da quinta semana, as fossas nasais são formadas (roxo), o que leva à formação dos processos nasais pareados mediais (amarelo) e laterais (azul); c) Ao fim da sexta semana os processos mediais nasais se fundem com os processos maxilares para a formação do lábio superior e palato primário. Os processos laterais nasais formam a ala nasal (amarelo). De forma similar, os processos mandibulares se fundem para formar a maxila inferior (rosa); d) Durante a sexta semana de embriogênese, o palato secundário se desenvolve com o crescimento bilateral dos processos maxilares, que crescem verticalmente ao lado da língua (verde); e) Em sequência as lâminas palatinas se elevam a uma posição horizontal sobre a língua, entrando em contato uma com a outra e começando a fusão; f) Por fim, temos a fusão das lâminas palatinas, dividindo o espaço oronasal em cavidades oral e nasal separadas (Mod. de DIXON *et al.*, 2011).

1.1.3. Classificação

As fissuras podem ser classificadas a depender da especialidade médica. Entre os cirurgiões, uma das classificações mais utilizadas é a de Spina e colaboradores (1972), que

leva em consideração a embriologia das estruturas faciais divididas em palato primário e secundário, e sua convergência em torno do forâmen incisivo. Já entre os geneticistas, a forma mais usual de classificação das FL/P é o agrupamento em formas síndrômicas, quando outras anomalias estão presentes além da fissura, ou não síndrômicas (FL/P NS), quando estas ocorrem isoladamente. (MOSSEY *et al.*, 2009; PASSOS-BUENO *et al.*, 2010; PEGELOW *et al.*, 2008; WONG & HÄGG, 2004).

Já existem mais de 500 síndromes associadas à FL/P e, dentre elas, destaca-se a síndrome de van der Woude (SVW, OMIM #119300) pela frequência relativamente alta (cerca de 2% dos casos de fissura). A SVW distingue das formas não-síndrômicas apenas pela presença de fossetas no lábio inferior (presente em 85% dos casos) e hipodontia (presente em 25% dos casos), e, em muitos casos, pacientes sem essas características são indistinguíveis de indivíduos afetados por formas não-síndrômicas. O padrão de herança da síndrome é autossômico dominante com penetrância alta (89-99%, BURDICK *et al.*, 1985), e mutações no gene *IRF6* (*interferon regulatory factor 6*), presente na região cromossômica 1q32.2, têm sido associadas a casos de SVW (KONDO *et al.*, 2002). Uma forma alélica da SVW é a Síndrome do Pterigio Poplíteo (OMIM #119500), na qual além de comprometimento craniofacial pode haver envolvimento dos membros inferiores e superiores (LEES *et al.*, 1999). As formas não-síndrômicas correspondem a aproximadamente 70% dos casos de FL/P, e serão o foco do presente trabalho (MOSSEY *et al.*, 2009; PASSOS-BUENO *et al.*, 2010).

1.2. Fissura labial com ou sem fissura de palato não-síndrômica

1.2.1. Epidemiologia

A incidência da FL/P é de cerca de 1:700 nascimentos, com variação étnica e geográfica, sendo mais frequente em asiáticos e nativos americanos (3,6:1000), intermediária em populações européias (1:1000) e menor em africanos (0,3:1000) (MOSSEY & CASTILLIA, 2003; VANDERAS, 1987). No Brasil, a prevalência de FL/P NS tem sido estimada

em 0.28-1.54 a cada 1000 nascimentos (LOFFREDO *et al.*, 2001; MENEGOTTO & SALZANO, 1991).

O risco de recorrência para FL/P NS é calculado de forma empírica, e varia de acordo com alguns fatores. O risco de recorrência aumenta de acordo com o grau de parentesco em relação a um afetado, sendo maior para indivíduos de 1º grau (3,5%), diminuindo para 2º e 3º graus (0,8% e 0,6% respectivamente). O risco também difere com a gravidade do defeito, sendo que em fissura de lábio e palato unilateral (FLPU) é de 2,5%, aumentando para 4,6% quando se trata de fissura de lábio e palato bilateral (FLPB) (GROSEN *et al.*, 2010).

Esses dados, juntamente com fatores como altas taxas de agregação familiar, altas estimativas de herdabilidade (variando de 45 a 85% no Brasil) e elevadas taxas de concordância de gêmeos monozigóticos (40-60%) versus dizigóticos (3-5%), evidenciam um forte componente genético na etiologia da FL/P NS (BRITO *et al.*, 2011; CHRISTENSEN & FOGH-ANDERSEN, 1993; GROSEN *et al.*, 2010; HECHT *et al.*, 1991;).

1.2.2. Modelo de Herança - Genética e Ambiente

O modelo de herança mais aceito é o multifatorial com efeito de limiar, isto é, a manifestação da doença resulta da interação entre diversos fatores genéticos e ambientais, de modo que ao atingir uma determinada concentração desses fatores de risco, ocorre a manifestação da doença (GROSEN *et al.*, 2010) (Figura 3).

O modelo multifatorial se aplica a traços quantitativos, onde há uma distribuição contínua do fenótipo (ex: altura, peso, colesterol), diferentemente da herança mendeliana, onde os fenótipos são distribuídos em classes distintas.

Figura 3: Modelo multifatorial com efeito de limiar. Curva de distribuição da susceptibilidade na população geral, composta por alelos de risco e fatores ambientais (Mod. de STRACHAN & READ, 1999).

Visto que a fissura ocorre no início do desenvolvimento embrionário, possui uma etiologia complexa e taxas de recorrência modestas, a identificação dos fatores de risco tem sido difícil (DIXON *et al.*, 2011).

1.2.3. Fatores de risco ambientais

Diversos fatores ambientais têm sido apontados como de predisposição à FL/P NS, porém a associação é inconsistente para a maioria deles. O fumo materno é o fator de risco mais consistente até o momento, e já foi associado a um aumento do risco de 30% (LITTLE *et al.*, 2002). Esse aumento no risco levanta a hipótese de que determinadas vias metabólicas estejam envolvidas no desenvolvimento de fissura. Neste contexto, observou-se associação entre variantes nos genes *GSTT1* (*glutathione s-transferase-θ1*) e *NOS3* (*nitric oxide synthase 3*) e susceptibilidade à FL/P NS na presença de fumo materno (SHAW *et al.*, 2005; SHI *et al.*, 2007; VAN ROOIJ *et al.*, 2001), porém a associação dessas variantes ainda é controversa (CHEVRIER *et al.*, 2008; MALLOY *et al.*, 1989; ROMITTI *et al.*, 1999).

O consumo materno de álcool também já foi associado a um aumento do risco de o feto desenvolver FL/P NS. Aparentemente, este aumento no risco está associado

particularmente à presença de variante responsável pela metabolização lenta do álcool no gene *ADH1C* (álcool desidrogenase 1C) (BOYLES *et al.*, 2010; DEROO *et al.*, 2008).

A ingestão de ácido fólico já foi descrita como fator protetor na prevenção de FL/P NS (VAN ROOIJ *et al.*, 2004; WILCOX *et al.*, 2007), sendo que variantes maternas em genes relacionados ao metabolismo de ácido fólico já foram associadas a um aumento no risco de FL/P (BUFALINO *et al.*, 2010). Essas associações, porém, nem sempre são replicadas (BHASKAR *et al.*, 2011; BLANTON *et al.*, 2011; SHAW *et al.*, 2006).

Outros fatores, como deficiência de multivitaminas, obesidade materna e infecções já foram apontadas como fatores de risco, porém ainda não há consenso sobre a contribuição destes fatores na predisposição à FL/P NS (DIXON *et al.*, 2011; MOSSEY *et al.*, 2009).

Estes achados provêm evidência de que fatores ambientais e interações gene-ambiente são importantes para o desenvolvimento de FL/P NS, e estudos prospectivos com poder suficiente para detectar esses efeitos serão importantes para a elucidação desses fatores.

1.2.4. Fatores de risco genéticos

Além dos fatores ambientais, fatores genéticos estão envolvidos no risco à FL/P NS, segundo o modelo multifatorial. Algumas hipóteses foram levantadas a respeito da natureza desses fatores.

A hipótese doença comum - variante comum assume que doenças comuns são causadas por variantes comuns (*minor allele frequency*, MAF>5%), as quais individualmente possuem um efeito menor, porém, quando combinadas, conferem um grande risco à doença. O aparecimento dessas variantes teria ocorrido uma vez em um ancestral comum, as quais com o passar das gerações, tornaram-se frequentes na população. Sendo frequentes, espera-se que seu efeito seja pequeno e cumulativo, a ponto de não terem sido removidas por seleção natural. Algumas variantes comuns de suscetibilidade às FL/P NS têm sido identificadas por meio de estudos de associação, como se verá a seguir.

Já a hipótese doença comum - variante rara assume a presença de variantes raras (MAF<1%), que conferem efeito moderado/alto ao risco da doença. Tais variantes raras

seriam em sua maioria população-específica, recentemente derivadas na população (BODMER & BONILA, 2008; MARIAN *et al.*, 2012). Essa hipótese está começando a ser testada através de sequenciamentos de nova geração de exomas ou genomas inteiros de afetados por FL/P NS (LESLIE & MURRAY, 2012).

1.2.5. Estudos de Associação

Estudos de associação, nos quais são comparadas as frequências de variantes entre dois grupos de indivíduos (como casos e controles), têm sido considerados a melhor abordagem para a identificação de variantes comuns associadas a doenças multifatoriais. Eles podem ser realizados utilizando a abordagem de genes candidatos, na qual é necessário um conhecimento prévio da etiologia da doença para escolher um gene ou uma via a ser testada para associação. Essa abordagem foi inicialmente utilizada pelos pesquisadores, e resultou na identificação de genes como *TGF α* e *MSX1* (envolvidos na formação do lábio e palato em estudos em modelos animais, ARDINGER *et al.* 1989; SATOKATA & MAAS, 1994), *MTHFR* (envolvido no metabolismo do folato, TOLAROVA *et al.*, 1998), e *IRF6* (associado à SVW, forma sindrômica de fissura, ZUCCHERO *et al.*, 2004).

Recentemente, estudos de associação de genoma inteiro (GWAS – *Genome Wide Association Study*) vêm sendo utilizados, os quais são bastante interessantes devido à possibilidade de se testar o genoma sem necessidade de uma hipótese *a priori* de gene candidato, possibilitando assim a descoberta de novos *loci* e mecanismos (MARIAN *et al.*, 2012). O GWAS foi possibilitado devido a recentes esforços de projetos como o HapMap (INTERNATIONAL HAPMAP CONSORTIUM, 2003), que caracterizou, em diferentes populações, frequências de variantes genéticas de base única (SNVs – *single nucleotide variant*) e a estrutura de blocos haplotípicos ao longo do genoma. Aproximadamente 10 milhões de variantes comuns (MAF>5%) são transmitidas através das gerações em blocos haplotípicos, o que permite que alguns SNVs (chamados de *tagSNVs*) capturem a maioria da variação de cada bloco. A estrutura haplotípica do genoma, aliada aos rápidos avanços na tecnologia e constante queda dos custos dos *arrays* de genotipagem, tem permitido um

avanço na cobertura do genoma de cada indivíduo, chegando a mais de um milhão de SNVs nos *arrays* atuais (HARDY & SINGLETON, 2009; MANOLIO, 2010).

1.2.5.1 *Loci* associados à FL/P NS

Considerando os estudos de associação realizados para FL/P NS, que incluem quatro GWAS, os principais *loci* associados a FL/P NS são o gene *IRF6* e as regiões 8q24.21 e 10q25.3, representados na Figura 4 (BIRNBAUN *et al.*, 2009a, MANGOLD *et al.*, 2009; ZUCCHERO *et al.*, 2004). Uma meta-análise recente dos dois maiores GWAS realizados para FL/P NS (BEATY *et al.*, 2010; MANGOLD *et al.*, 2009) culminou na identificação de seis novos *loci* (1p36, 2p21, 3p11.1, 8q21.3, 13q31.1, 15q22), sendo que 1p36 atingiu significância marginal um dos GWAS anteriores (BEATY *et al.*, 2010). Essa meta-análise também corroborou a associação de *loci* anteriormente identificados (1p22.1, 1q32.2 - *IRF6*, 8q24.21, 10q25, 17q22, 20q12) (LUDWIG *et al.*, 2012).

Figura 4: Diagrama representando os principais *loci* associados a FL/P NS por GWAS. Linhas pontilhadas representam associações marginais, enquanto linhas sólidas representam associações que são comumente aceitas em GWAS ($p < 10^{-7}$). (*) MANGOLD e colaboradores (2009) encontraram evidência de interação entre *IRF6* e *GREM1*, um gene localizado na região 15q13.3, na susceptibilidade a FL/P NS (Mod. de BRITO *et al.*, 2012c).

O primeiro estudo a sugerir que variações polimórficas no gene *IRF6* estavam associadas à predisposição a FL/P NS foi realizado por Zuccherro e colaboradores (2004). O gene *IRF6* foi considerado um gene candidato devido ao fato de ser responsável por uma forma sindrômica de FL/P (SVW) que possui uma grande sobreposição fenotípica com a FL/P NS. A associação de diversos SNVs de *IRF6* com a doença tem sido identificada e confirmada através de diversos estudos, incluindo GWAS (BEATY *et al.*, 2010a; BIRNBAUM *et al.*, 2009b; BLANTON *et al.*, 2005; BRITO *et al.*, 2012a; GHASSIBÉ *et al.*, 2005; HUANG *et al.*, 2009; JAGOMÄGI *et al.*, 2010; LUDWIG *et al.*, 2012; PAN *et al.*, 2010; PARK *et al.*, 2007; PEGELOW *et al.*, 2008; RAHIMOV *et al.*, 2008; SCAPOLI *et al.*, 2005; VIEIRA *et al.*, 2008).

A maioria das variantes de *IRF6* associadas à FL/P NS encontra-se em região não-codificante do gene, com um possível papel na regulação gênica ainda controverso. Rahimov e colaboradores (2008) identificaram o SNV rs642961 (G>A), localizado em um *enhancer* de *IRF6* (aproximadamente 10kb *upstream* do sítio de início de transcrição do gene). Essa variante altera o sítio de ligação com o fator de transcrição AP2 α , o qual, por sua vez, quando produzido em quantidades insuficientes em decorrência de mutações patogênicas, causa a síndrome Branchiooculofacial (OMIM #113620), uma forma sindrômica de fissura (MILUNSKY *et al.*, 2008). Estes resultados sugerem que o SNV rs642961 seja uma variante funcional. De fato, o alelo de risco desse SNV (A) está associado à diminuição dos níveis de expressão de *IRF6* em tecido de lábio, embora não tenha sido detectada diferença nos níveis de expressão do gene repórter (luciferase) em cultura de queratinócitos (PAN *et al.*, 2010; RAHIMOV *et al.*, 2008).

Outros dois SNVs do gene *IRF6* associados a predisposição à FL/P NS foram também correlacionados com variações nos níveis de transcritos de *IRF6*: rs2236906 (T>G) e rs590223 (A>G). O primeiro está localizado no íntron 3 do gene *IRF6* (JUGESSUR *et al.*, 2009) e foi descrito como regulatório em culturas primárias de células T (DIMAS *et al.*, 2009). Já o rs590223 (A>G) está localizado no íntron 6 do gene *TRAF3IP3* (adjacente a 3' do *IRF6*) (JAGOMÄGI *et al.*, 2010) e altera os níveis de expressão do gene *IRF6* em amostras de fígado (SCHADT *et al.*, 2008).

Um GWAS realizado por Birnbaum e colaboradores (2009a) utilizando população europeia detectou associação entre FL/P NS e diversos SNVs na região 8q24.21, sendo que a variante com sinal de associação mais significativo é a rs987525. A região de associação ocupa uma extensão de 640kb (chr8:129,700,000–130,340,000, NCBI build 36.1). Este

resultado tem sido replicado em diversos estudos, em sua maioria em populações de origem européia (GRANT *et al.*, 2009; MOSTOWSKA *et al.*, 2010; NIKOPENSIUS *et al.*, 2009). A região de associação está localizada em um deserto gênico, o qual é flanqueado pelo proto-oncogene *PVT1* (localizado aproximadamente a 900kb *upstream* a rs987525) e pelo gene *LOC728724* (aproximadamente a 280kb *downstream* a rs987525), o qual é transcrito em um RNA não-codificante.

Já foi demonstrado que outra região no deserto gênico 8q24.21, fora do intervalo de associação com FL/P NS, possui um papel regulatório importante, contendo elementos regulatórios que controlam genes mais distantes, como *MYC* e *PVT1*, além de variantes associadas a diversos tipos de câncer (HUPPI *et al.*, 2012; SOTELLO *et al.*, 2010; WASSERMAN *et al.*, 2010).

Na região 10q25.3 quatro variantes foram associadas à FL/P NS, e ocupam uma extensão de 30kb. O principal gene candidato da região é o *VAX1*, um fator de transcrição aparentemente envolvido no desenvolvimento do prosencéfalo, e está localizado a 53kb da região de associação (MANGOLD *et al.*, 2009). Camundongos homozigotos para mutações em *Vax1* apresentam malformações craniofaciais, incluindo fissura de palato (HALLONET *et al.*, 1999).

Ao contrário de *IRF6*, no qual um mecanismo molecular subjacente à variante rs642961 foi identificado, os papéis funcionais das regiões 8q24.21 e 10q25.3 na susceptibilidade à FL/P NS permanecem desconhecidos.

1.2.6. Expressão gênica nas doenças complexas

Apesar de diversos *loci* já terem sido identificados (BEATY *et al.*, 2010; BIRNBAUN *et al.*, 2009a; GRANT *et al.*, 2009; LUDWIG *et al.*, 2012; MANGOLD *et al.*, 2009), estes não explicam totalmente a herdabilidade da FL/P NS. Os altos níveis de significância dos sinais de associação ainda precisam ser traduzidos em um entendimento completo dos genes ou elementos genéticos dos *loci* que estão mediando a susceptibilidade à doença.

É interessante notar que, diferentemente das doenças mendelianas, nas quais as mutações geralmente afetam a estrutura das proteínas, nas doenças complexas a grande

maioria das variantes associadas encontra-se em regiões não-codificantes, o que sugere que os mecanismos subjacentes à susceptibilidade às doenças complexas envolvem segmentos genômicos envolvidos com regulação da expressão gênica (HINDORFF *et al.*, 2009).

Os níveis de expressão gênica são caracteres quantitativos e herdáveis, e um dos mecanismos que modula esses níveis de expressão são variantes presentes no DNA em regiões regulatórias (CHEUNG *et al.*, 2003; HARDY & SINGLETON, 2009; SCHADT *et al.*, 2003;). Esses *loci* associados a traços quantitativos de expressão, denominados eQTLs (*expression quantitative trait loci*), podem ser mapeados ao longo do genoma e podem ajudar na elucidação dos mecanismos de susceptibilidade à doença, identificando regiões do genoma com efeito funcional na expressão gênica (SCHADT *et al.*, 2003, 2008). Essas podem ser regiões candidatas para verificação da associação com a doença, auxiliando assim no entendimento do papel das variantes na etiologia da doença.

A variação da expressão gênica é um importante mecanismo de susceptibilidade a doenças complexas, e diversos trabalhos já obtiveram sucesso ao aliar estudos de associação e eQTLs como estratégia para identificação de novos *loci* de risco, como já demonstrado para diversas doenças como esquizofrenia, asma e doença de Crohn (DE JONG *et al.*, 2012; FRANSEN *et al.*, 2010; LIBIOULLE *et al.*, 2007; MOFFAT *et al.* 2007). Para asma, diversos SNVs apresentando forte associação com a doença foram detectados em uma região de 206 kb na região 17q21. Essas variantes estão relacionadas à variação nos níveis de transcrição do gene *ORMDL3*, o qual pertence a uma classe de genes que codifica proteínas transmembrana ancoradas ao retículo endoplasmático, porém de função desconhecida (MOFFATT *et al.*, 2007). O papel regulatório dessas variantes indica que a variação na expressão desse gene está envolvida na susceptibilidade à asma. O mesmo acontece para doença de Crohn, onde uma região de associação de 1,25 Mb em um deserto gênico na região 5p13.1 está correlacionada com os níveis de expressão do gene *PTGER4*, um receptor de prostaglandina, para o qual camundongos *knock-out* desenvolvem colite (LIBIOULLE *et al.*, 2007).

Estudos já demonstraram que há um enriquecimento de eQTLs dentre os *loci* associados a doenças complexas (KNIGHT, 2005; NICA *et al.*, 2010; NICOLAE *et al.*, 2010), e também já foi observado que eQTLs mapeados em tecidos envolvidos na doença estão enriquecidos de variantes associadas à mesma, o que reforça a importância de integrar as duas abordagens, utilizando, se possível um tecido acometido pela doença (DING *et al.*, 2010; KANG *et al.*, 2012a,b; RICHARDS *et al.*, 2012).

Para FL/P NS, as variantes com maior evidência de associação com a doença, rs641961 (gene *IRF6*) e rs987525 (*locus* 8q24.21), estão localizadas em regiões não-codificantes do genoma, o que sugere que a variação na expressão gênica também seja importante para susceptibilidade à FL/P NS, como já mencionado acima.

1.3. Objetivos

Conforme explicitado, a integração das abordagens de estudo de associação e mapeamento de eQTLs é uma ferramenta promissora na elucidação da etiologia de doenças complexas, visto que propicia descobertas de novas vias biológicas e entendimento funcional das variantes já associadas.

O objetivo principal do presente trabalho é elucidar o papel de variantes na etiologia da FL/P NS. Para tanto, desenvolvemos os seguintes objetivos específicos:

- 1) Verificar se variantes previamente associadas à fissura possuem um papel regulatório em células-tronco mesenquimais provenientes de músculo orbicular do lábio (OOMMSC) de pacientes FL/P NS;
- 2) Verificar se eQTLs mapeados em OOMMSC de pacientes FL/P NS estão associados à doença.

Capítulo 2. *IRF6 is a Risk Factor for Nonsyndromic Cleft Lip in the Brazilian Population*

Este capítulo se refere ao objetivo 1, segundo o qual verificamos se variantes previamente associadas à fissura possuem um papel regulatório em células-tronco mesenquimais provenientes de músculo orbicular do lábio (OOMMSC) de pacientes FL/P NS. Para isso, foi verificada a correlação dos genótipos das variantes rs642961 e rs590223 com os níveis de expressão do gene *IRF6*.

***IRF6* é um fator de risco para fissura de lábio não-sindrômica na população brasileira**

Luciano A. Brito, Camila F. S. Bassi, Cibele Masotti, Carolina Malcher, Kátia M. Rocha, David Schlesinger, Daniela F. Bueno, Lucas A. Cruz, Ligia K. Barbara, Débora R. Bertola, Diogo Meyer, Diogo Franco, Nivaldo Alonso, Maria Rita Passos-Bueno

American Journal of Medical Genetics Part A. Setembro, 158A (9): 2170–2175 (2012).
Publicado online: 6 Agosto 2012

Resumo

Fissura de lábio com ou sem fissura de palato não-sindrômica (FL/P NS) é uma doença complexa com incidência mundial estimada em 1:700. Dentre os *loci* de susceptibilidade propostos, o gene *IRF6* e a região 8q24.21 têm sido corroborados em diferentes populações. Para testar o papel do *IRF6* na predisposição à FL/P NS na população brasileira, nós conduzimos um estudo de associação com correção para ancestralidade com os SNPs rs642961 e rs590223, respectivamente, localizados a 5' e 3' do gene *IRF6*, e sem desequilíbrio de ligação, em pacientes provenientes de cinco diferentes localidades do

Brasil. Nós também avaliamos o efeito desses SNPs na expressão de *IRF6* em células-tronco mesenquimais. Observamos associação entre rs642961 e fissura de lábio isolada (FL) ($p=0,009$; *odds ratio* para o genótipo AA = 1,83 [Intervalo de confiança 95%, 1,03 – 2,84]). Os resultados indicam que essa associação é decorrente dos pacientes afetados de Barbalha, uma localidade que apresenta a maior herdabilidade estimada ($H^2 = 0,85$), sendo que o alelo A de rs642961 está agindo por meio de um modelo dominante. Nenhuma associação foi detectada para o SNP rs590223. Nós não encontramos nenhuma correlação entre os níveis de expressão e os genótipos dos dois *loci*, e é possível que esses SNPs possuam um papel funcional em algum período específico da embriogênese.

IRF6 is a Risk Factor for Nonsyndromic Cleft Lip in the Brazilian Population

Luciano A. Brito,¹ Camila F.S. Bassi,¹ Cibele Masotti,¹ Carolina Malcher,¹ Kátia M. Rocha,¹ David Schlesinger,¹ Daniela F. Bueno,¹ Lucas A. Cruz,¹ Ligia K. Barbara,¹ Débora R. Bertola,¹ Diogo Meyer,² Diogo Franco,³ Nivaldo Alonso,⁴ and Maria Rita Passos-Bueno^{1*}

¹Human Genome Research Center, Institute de Biosciences, University of São Paulo, São Paulo, Brazil

²Department of Genetics, Institute de Biosciences, University of São Paulo, São Paulo, Brazil

³Department of Plastic Surgery, School of Medicine, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

⁴Division of Plastic Surgery, Department of Surgery, School of Medicine of São Paulo University, São Paulo, Brazil

Manuscript Received: 27 October 2011; Manuscript Accepted: 20 May 2012

Nonsyndromic cleft lip with or without cleft palate (NSCL/P) is a complex disorder with a worldwide incidence estimated at 1:700. Among the putative susceptibility loci, the *IRF6* gene and a region at 8q24.21 have been corroborated in different populations. To test the role of *IRF6* in NSCL/P predisposition in the Brazilian population, we conducted a structured association study with the SNPs rs642961 and rs590223, respectively, located at 5' and 3' of the *IRF6* gene and not in strong linkage disequilibrium (LD), in patients from five different Brazilian locations. We also evaluated the effect of these SNPs in *IRF6* expression in mesenchymal stem cells (MSC). We observed association between rs642961 and cleft lip only (CLO) ($P = 0.009$; odds ratio (OR) for AA genotype = 1.83 [95% Confidence interval (CI), 0.64–5.31]; OR for AG genotype = 1.72 [95% CI, 1.03–2.84]). This association seems to be driven by the affected patients from Barbalha, a location which presents the highest heritability estimate ($H^2 = 0.85$), and the A allele at rs642961 is acting through a dominant model. No association was detected for the SNP rs590223. We did not find any correlation between expression levels and genotypes of the two loci, and it is possible that these SNPs have a functional role in some specific period of embryogenesis. © 2012 Wiley Periodicals, Inc.

Key words: cleft lip/palate; *IRF6* transcription levels; rs642961; rs590223; mesenchymal stem cell; structured association; common disease-common variant; admixture population; heritability

INTRODUCTION

Nonsyndromic cleft lip with or without cleft palate (NSCL/P) are among the most common congenital malformations [Gorlin et al., 2001], with an estimated worldwide prevalence of 1:700, varying across ethnicities [Vanderas, 1987; Loffredo et al., 2001; Mossey et al., 2009]. NSCL/P represents a significant burden for family and society, as affected individuals require multidisciplinary treatment until adult life. High heritability estimates and a 20–40 times

How to Cite this Article:

Brito LA, Bassi CFS, Masotti C, Malcher C, Rocha KM, Schlesinger D, Bueno DF, Cruz LA, Barbara LK, Bertola DR, Meyer D, Franco D, Alonso N, Passos-Bueno MR. 2012. *IRF6* is a risk factor for nonsyndromic cleft lip in the Brazilian population.

Am J Med Genet Part A 158A:2170–2175.

increased recurrence risk in families with one affected patient suggest an important genetic component in NSCL/P etiology [Menegotto and Salzano, 1991; Christensen and Fogh-Andersen, 1993; Brito et al., 2011]. Worldwide studies have been conducted in order to dissect the genetics of NSCL/P, assuming a multifactorial model of inheritance with a common origin for the risk alleles (reviewed in Dixon et al., 2011). The two loci more consistently associated are the *IRF6* gene at 1q32.2 and a locus mapped to a gene desert at 8q24.21. The SNP rs2235371 (V274I) was the first marker in *IRF6* shown to be associated with NSCL/P, notably in Asian and South American populations [Zuccherro et al., 2004]. This association was further reproduced in other populations, including those of European descent, and with other *IRF6* SNPs in linkage disequilibrium (LD) with V274I [Park et al., 2007; Jugessur et al., 2008; Huang et al., 2009; Blanton et al., 2010]. A second SNP, rs642961 (G>A), located in an *IRF6* enhancer, was reported as a likely cause

Additional supporting information may be found in the online version of this article.

Grant sponsor: CNPq, FAPESP and the Brazilian Ministry of Health.

*Correspondence to:

Maria Rita Passos-Bueno, PhD, Rua do Matão, 277, sala 200, Cidade Universitária, São Paulo, SP 05508-090, Brazil. E-mail: passos@ib.usp.br
Article first published online in Wiley Online Library

(wileyonlinelibrary.com): 6 August 2012

DOI 10.1002/ajmg.a.35526

of the association between the previously tested *IRF6* SNPs and NSCL/P and a more pronounced effect was observed in cleft lip only (CPO) patients than in cleft lip with cleft palate (CLP) patients. [Rahimov et al., 2008]. Only one in vivo study, conducted in a small sample, suggested that rs642961 genotypes modulate *IRF6* expression in an allele-dosage manner in skin lip from affected patients [Pan et al., 2010]. The SNP rs590223 (A>G) also is predicted to modulate in vivo the transcriptional levels of *IRF6* in liver cells [Schadt et al., 2008], and its association with NSCL/P is controversial [Jagomägi et al., 2010; Mostowska et al., 2010].

The extent of the contribution of *IRF6* SNPs for NSCL/P susceptibility in Brazilian patients is still unknown. A single previous study failed to find association of rs642961 in patients from the Southeastern region of Brazil [Paranaíba et al., 2010], while the SNP rs590223 has never been tested for association in Brazilian patients. The Brazilian population has a tri-hybrid origin, composed primarily by West Africans (brought as slaves between the 16th and 19th centuries), Europeans (settlers and immigrants), and Amerindians (native inhabitants; [Centro de Documentação e Disseminação de Informações (Brazil), 2007]), a factor that may introduce bias if not taken into account. Heritability of NSCL/P can be another confounding factor in our population, as it varies from 45 to 85% depending on the patient's geographical origin [Brito et al., 2011].

In the present study, our goal is to verify whether *IRF6* is associated with NSCL/P in Brazilian patients through a structured analysis of the SNPs rs642961 and rs590223 to control population admixture. We also discuss whether heritability of NSCL/P and clinical variability interfere in the detection of the association. Finally, we evaluate whether the genotype of the two SNPs were correlated with *IRF6* transcriptional expression levels in mesenchymal stem cells (MSC) derived from *orbicularis oris* muscle, as their transcriptome seems to reflect, at least in part, the dysregulated pathways of NSCL/P patients [Bueno et al., 2009].

MATERIALS AND METHODS

Patients and Controls

Patients were ascertained from medical programs of a non-governmental organization—Operation Smile (<http://www.operationsmile.org/>)—in five different Brazilian cities: Santarém-PA ($n = 88$), Barbalha-CE ($n = 57$), Fortaleza-CE ($n = 225$), Maceió-AL ($n = 77$), and Rio de Janeiro-RJ ($n = 116$), totaling 563 patients. Patients were classified as CPO if the cleft involved only the lip (unilateral or bilateral) with no involvement of the palate, but including those with alveolar defect, while those classified as CLP included unilateral and bilateral cleft lip associated with cleft palate. Detailed patients' ascertainment method and inclusion criteria were previously reported elsewhere [Brito et al., 2011]. We clinically excluded syndromic and cleft palate only patients. Only unrelated individuals were included in our sample.

The control group was provided by the Biobank of our institute, and consisted of 401 individuals (336 individuals ascertained in São Paulo and 65 from Fortaleza-CE, as detailed in Errera et al. [2006] and Orabona et al. [2009], respectively). Ethical issues prevented

the collection of control samples in other locations. 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 individuals or their legal tutors.

Association Analysis

To correct for the confounding effect of population structure, we used a panel of 40 ancestry-informative markers (AIMs) to characterize regional and individual contributions of African, European, and Amerindian ancestries. The AIMs consist of indel (insertion-deletion) polymorphisms spread across the autosomal 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 described in Bastos-Rodrigues et al. [2006], and also adopted in Santos et al. [2010] and Brito et al. [2011]. Our ancestry panel were previously validated by the verification that they allow the ancestry discrimination of the three parental populations (data available under request). An M13 tail labeled with FAM fluorochrome was added to every 5' end of forward primers, so that each PCR product incorporated one tail, allowing multiplex PCR with 4 to 7 markers per reaction (primer sequences available upon request). Fragment analysis was performed with Gene Mapper software (Applied Biosystems, Foster City, CA), after capillary electrophoresis targeting the PCR products for the FAM-labeled M13 tail.

We next ran Structure 2.3.3 to estimate ancestry components of each sample, and thus infer possible roles of population stratification [Falush et al., 2007]. We assumed $K = 3$ (where K is the number of ancestry components that Structure estimates) based on the well-known tri-hybrid origin of Brazilians, and incorporated "learning samples" in the run, with pre-specified population of origin (European, African, and Amerindian) according to the Mammalian Genotyping Service database (Marshfield Clinic, Marshfield-WI), in order to assist the software to estimate ancestry of the admixed individuals.

The SNPs rs642961 and rs590223, located 42,563 bp apart (NCBI build 18) and in weak LD ($r^2 = 0.30$ in patients and 0.36 in controls) were genotyped with TaqMan[®] method (Applied Biosystems). Following the ancestry contribution estimations, we used STRAT to test each SNP for association, conditioning on the individual ancestries [Pritchard et al., 2000]. Considering the multiple testing, we accepted 0.0083 as threshold for significance in the analysis stratified by clinical variability, after Bonferroni correction. Odds ratios (OR) with 95% confidence interval (CI) were estimated under additive model for the genotypes with one or two copies of the risk allele. Haplotype inferences were performed with Arlequin 3.5 software [Excoffier and Lischer, 2010].

Expression Analysis

The effect of both SNPs in *IRF6* expression levels was tested in MSC cultures established from *orbicularis oris* muscle fragments from an independent sample of 46 individuals (42 NSCL/P patients and 4 controls; median and average = 4 and 6.9 years old, respectively). Tissue samples were obtained from three different centers (Division of Plastic Surgery of University of São Paulo, SOBRAPAR

and Operation Smile). Informed consent was obtained from the patient's legal tutors.

Orbicularis oris muscle MSC cultures were established according to previously published protocol [Bueno et al., 2009], and expanded to 80% confluence. RNA and DNA extraction followed manufacturer's protocols (Macherey-Nagel, Germany). We confirmed the homogeneous mesenchymal immunophenotype of 10 cell cultures through flow cytometer analysis (EasyCyte Flow cytometer, Guava Technologies, Hayward, CA).

We used the Human Gene 1.0 ST microarray chips (Affymetrix, Santa Clara, CA) to measure whole gene expression of the 46 individuals, according to the manufacturer's protocol. Gene expression values were obtained with the three-step Robust Multi-array Average pre-processing method implemented in the software Expression Console (Affymetrix). As a validation, we also measured the relative expression of *IRF6* in 22/46 individuals through real-time PCR, and we observed a strong correlation between the two methods ($P = 0.001$, $r = 0.67$). These individuals were also genotyped for the two SNPs with TaqMan® method, and presented 100% of genotype concordance with those obtained through microarray.

RESULTS

Association Analysis

After excluding individuals with low AIM genotype assignment (<80%), we obtained a final sample of 471 patients (Santarém, $n = 82$; Barbalha, $n = 52$; Fortaleza, $n = 163$; Maceió, $n = 62$, and Rio de Janeiro, $n = 112$) and 391 controls. The two SNPs (Table I) and the indel markers (data not shown) were in Hardy–Weinberg equilibrium in controls.

The population structure analysis of our sample revealed a high contribution of European ancestry both in patients (61.2%) and in controls (71.1%; Table I). We also observed some variation in the ancestry contribution among the patient's groups, e.g., the highest Amerindian ancestry in Santarém (31.6%) and the highest African contribution in Rio de Janeiro (30.8%; Supplementary eTable I—See Supporting Information online). Individual ancestry contributions inferred by Structure are plotted in Figure 1.

Minor allele (A) frequency of rs642861 was higher in patients (0.19) than in controls (0.15) but no significant allelic association was observed in the total sample of patients ($P = 0.48$). We next stratified the analysis by “subphenotype” (CLO and CLP), and we observed a marginal association between rs642961 and CLO (90 patients; $P = 0.009$; $OR_{AA/GG} = 1.83$ [95% CI 0.64–5.31], $OR_{AG/GG} = 1.72$ [95% CI 1.03–2.84]), but no evidence of association with CLP (381 patients; $P = 0.62$, $OR_{AA/GG} = 1.62$ [95% CI 0.81–3.25], $OR_{AG/GG} = 1.12$ [95% CI 0.80–1.56]). The rs590223 SNP was not associated (NSCL/P, $P = 0.99$; CLO, $P = 0.85$; CLP, $P = 0.99$; Table I). The haplotype based analysis for rs642961–rs590223 did not present significant deviation from random distribution among cases and controls ($P = 0.95$).

Expression Analysis

We did not observe any significant effect neither for rs642961 (Spearman correlation test, $P = 0.15$, $r = 0.21$) nor for rs590223 ($P = 0.07$, $r = 0.27$) in the *IRF6* expression levels in *orbicularis oris* muscle MSC (Supplementary eFigure 1; genotyping information in Supplementary eTable II – See Supporting Information online). Inter-individual variation of *IRF6* expression levels was not correlated with age in our sample (Spearman correlation, two-tailed $P = 0.74$, $r = 0.05$).

DISCUSSION

Recent studies have reported association of NSCL/P with many loci throughout the genome, but the majority of the studies conducted so far has analyzed non-Latin populations. The replication of those results in other populations and the investigation of the mechanisms by which the arising variants interfere in clefting susceptibility are crucial for a better understanding of the genetic architecture of NSCL/P.

In the present study, we did not detect association between rs642961 and NSCL/P, however, stratifying the analysis by clinical variability, we observed association in the CLO group ($P = 0.009$), thus confirming the previous observation of Rahimov et al. [2008] that the effect of this allele is more pronounced among CLO than in CLP patients. The A allele seems to be acting in a dominant model of

TABLE I. Estimated Ancestry Proportions of European, African, and Amerindian Parental Populations

	N	Ancestry contributions Eur/Afr/Am [%]	rs642961				rs590223			
			Observed genotype frequencies [%] AA/AG/GG	f[A] [%]	H–W P-value	Association test P-value	Observed genotype frequencies [%] GG/GA/AA	f[G] [%]	H–W P-value	Association test P-value
Controls	391	71/18/11	4/23/73	15	0.05	—	11/46/43	34	0.65	—
NSCL/P patients										
Total	471	61/22/17	6/26/68	19	0.004	0.48	11/47/42	34	0.43	0.99
CLO	90	61/22/17	6/33/61	22	0.74	0.009	10/53/37	37	0.16	0.85
CLP	381	61/22/17	5/25/70	18	0.001	0.68	44/45/11	34	0.82	0.99

Genotypic and allelic frequencies of rs642961 and rs590223, and P-values of Hardy–Weinberg test, and structured association test (at 95% confidence).

H–W, Hardy–Weinberg; NSCL/P, nonsyndromic cleft lip with or without cleft palate; CLO, cleft lip only; CLP, cleft lip and palate.

FIG. 1. Bar plot of contribution of European (blue), African (green), and Amerindian (red) ancestries to our sample, in which single columns represent an individual. NSCL/P patients in the upper panel; in the lower panel, controls, and parental populations (EUR, Europeans, AFR, Africans, AMI, Amerindians—used as “learning samples” to assist the ancestry estimations of admixed individuals in Structure, run assuming $K = 3$ subpopulations).

susceptibility, given the similar OR observed for homozygous (1.83 [95% CI 0.64–5.31]) and heterozygous (1.72 [95% CI 1.03–2.84]) genotypes, in accordance to previous data obtained among European ancestry NS CL/P patients [Rahimov et al., 2008; Birnbaum et al., 2009]. Although these findings might suggest different etiologies for CLO and CLP, their co-occurrence within families is not uncommon, which in turn favors a common etiology for them. It is thus possible that CLO etiology might be genetically more homogenous than CLP, that is, only one or a few molecular pathways are deregulated among CLO patients.

We previously reported on a positive association of rs987525 SNP (8q24.21) in Brazil, driven by a subpopulation of patients with the highest heritability estimates (from Barbalha; $H^2 = 85\%$) as compared to other four regions of Brazil (Santarém, Fortaleza, Maceió, and Rio de Janeiro presented $H^2 = 45\text{--}71\%$; [Brito et al., 2012]). A stratified analysis by region of origin in the present study did not reveal a strong association with any group (Supplementary eTable I—See Supporting Information online), however, we observed a trend of positive association only in the Barbalha subset (Barbalha, $P = 0.02$; Santarém, $P = 0.59$; Fortaleza, $P = 0.66$; Maceió, $P = 0.44$; Rio de Janeiro, $P = 0.69$). In Barbalha, the frequency of the A allele/rs642961 among CLO patients is 0.5, which is much higher than among CLO patients from the other regions (Santarém = 0.21, Fortaleza = 0.20, Maceió = 0.22, Rio = 0.15), and it might be the contributing factor to the trend of association observed only in Barbalha. In support to this finding, we did not observe such substantial deviation of the rs642961 A allele frequency comparing CLP patients (Barbalha: 0.23, Santarém: 0.22, Fortaleza: 0.18, Maceió: 0.14, Rio: 0.10) and CLO patients. These

results further support the hypothesis that high-heritability populations are more suitable to find moderate-to-low-effect alleles in association mapping approaches.

The marginal association observed in the small CLO sample is in accordance with a common *IRF6* risk allele in populations across the world. Although the rs590223 SNP failed to reach positive association, it is still possible that other susceptibility variants at *IRF6* locus are also involved with NSCL/P susceptibility, independently of rs642961. For example, SNP rs2235371 but not rs642961, has been associated with NSCL/P in some population groups [Blanton et al., 2010]. Giving the low frequency of rs2235371 in our population (previously estimated at 0.05 in a set of 108 NS CL/P patients, unpublished data), we were not able to investigate its role in the orofacial clefting predisposition in our sample.

We verified a lack of correlation between *IRF6* expression levels and the SNPs rs642961 and rs590223 in the MSCs here tested, which in turn might be the consequence of the small sample size or the studied cell type might not be the best one to address this issue. It is possible that these alleles interfere in the proper expression of *IRF6* in critical moments of embryonic craniofacial development and different functional analysis will be necessary to conduct, in order to understand how they contribute to the occurrence of CLO or CLP.

In conclusion, despite the limitations of the small sample size of the CLO group, we found a weak association between the CLO group and rs642961, which risk allele seems to act in a dominant model. The CLO patients also explain the trend of association observed in Barbalha, the population subset with the highest heritability of NSCL/P. Our findings, together with the literature, support the model of a common risk allele in CLO in several

populational groups. Although the structured association approach minimizes the effects of population structure, we can never discard the possibility of false positive results in such heterogeneous population. We expect that the elucidation of the mechanisms through which rs642961 leads to NSCL/P predisposition will impact the translation of the association studies results into the clinic.

ACKNOWLEDGMENTS

We are grateful to all the Operation Smile Brazil crew for the opportunity of joining rehabilitation programs, and to Drs. Simone Ferreira and Andressa Morales for helping with the genotyping. We also appreciate the secretarial assistance of Mrs. Constança Urbani and Mrs. Marcia Góes Teixeira. This work was supported by CNPq, FAPESP, 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.
- Birnbaum S, Ludwig KU, Reutter H, Herms S, de Assis NA, Díaz-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. 2009. IRF6 gene variants in Central European patients with non-syndromic cleft lip with or without cleft palate. *Eur J Oral Sci* 117:766–769.
- Blanton SH, Burt A, Garcia E, Mulliken JB, Stal S, Hecht JT. 2010. Ethnic heterogeneity of IRF6 AP-2a binding site promoter SNP association with nonsyndromic cleft lip and palate. *Cleft Palate Craniofac J* 47:574–577.
- Brito LA, Cruz LA, Rocha KM, Barbara LK, Silva CB, Bueno DF, Aguiar M, Bertola DR, Franco D, Costa AM, Alonso N, Otto PA, Passos-Bueno MR. 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* 155A:1581–1587.
- Brito LA, Paranaíba LMR, Bassi CFS, Masotti C, Malcher C, Schlesinger D, Rocha KM, Cruz LA, Bárbara LK, Alonso N, Franco D, Bagordakis E, Martelli-Junior H, Meyer D, Coletta RD, Passos-Bueno MR. 2012. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. *Birth Defects Res A Clin Mol Teratol* In press.
- Bueno DF, Kerkis I, Costa AM, Martins MT, Kobayashi GS, Zucconi E, Fanganiello RD, Salles FT, Almeida AB, do Amaral CE, Alonso N, Passos-Bueno MR. 2009. New source of muscle-derived stem cells with potential for alveolar bone reconstruction in cleft lip and/or palate patients. *Tissue Eng Part A* 15:427–435.
- Centro de Documentação e Disseminação de Informações (Brazil). 2007. Brazil, 500 years of settlement. Rio de Janeiro: IBGE-Instituto Brasileiro de Geografia e Estatística. p 229.
- Christensen K, Fogh-Andersen P. 1993. Cleft lip (+/– cleft palate) in Danish twins, 1970–1990. *Am J Med Genet* 47:910–916.
- Dixon MJ, Marazita ML, Beaty TH, Murray JC. 2011. Cleft lip and palate: Understanding genetic and environmental influences. *Nat Rev Genet* 12:167–178.
- Errera FI, Silva ME, Yeh E, Maranduba CM, Folco B, Takahashi W, Pereira AC, Krieger JE, Passos-Bueno MR. 2006. Effect of polymorphisms of the MTHFR and APOE genes on susceptibility to diabetes and severity of diabetic retinopathy in Brazilian patients. *Braz J Med Biol Res* 39: 883–888.
- Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res* 10:564–567.
- Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Mol Ecol Notes* 7:574–578.
- Gorlin RJ, Cohen MM Jr, Hennekam RCM. 2001. Syndromes of the head and neck, 4th edition. New York: Oxford University Press. p 1283.
- Huang Y, Wu J, Ma J, Beaty TH, Sull JW, Zhu L, Lu D, Wang Y, Meng T, Shi B. 2009. Association between IRF6 SNPs and oral clefts in West China. *J Dent Res* 88:715–718.
- Jagomägi T, Nikopensius T, Krjutskov K, Tammekivi V, Viltrop T, Saag M, Metspalu A. 2010. MTHFR and MSX1 contribute to the risk of nonsyndromic cleft lip/palate. *Eur J Oral Sci* 118:213–220.
- 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* 32:413–424.
- Loffredo LCM, Freitas JAS, Grigolli AAG. 2001. Prevalence of oral clefts from 1975 to 1994, Brazil. *Rev Saúde Pública* 35:571–575.
- Menegotto BG, Salzano FM. 1991. Clustering of malformations in the families of South American oral cleft neonates. *J Med Genet* 28:110–113.
- Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. 2009. Cleft Lip and palate. *Lancet* 374:1773–1785.
- Mostowska A, Hozyasz KK, Wojcicki P, Biedziak B, Paradowska P, Jagodzinski PP. 2010. Association between genetic variants of reported candidate genes or regions and risk of cleft lip with or without cleft palate in the polish population. *Birth Defects Res A Clin Mol Teratol* 88: 538–545.
- Orabona GM, Griesi-Oliveira K, Vadasz E, Bulcão VL, Takahashi VN, Moreira ES, Furia-Silva M, Ros-Melo AM, Dourado F, Matioli SR, Otto P, Passos-Bueno MR. 2009. HTR1B and HTR2C in autism spectrum disorders in Brazilian families. *Brain Res* 1250:14–19.
- Pan Y, Ma J, Zhang W, Du Y, Niu Y, Wang M, Zhang Z, Wang L. 2010. IRF6 polymorphisms are associated with nonsyndromic orofacial clefts in a Chinese Han population. *Am J Med Genet Part A* 152A:2505–2511.
- 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* 16:193–197.
- 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: 219–227.
- Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. 2000. Association mapping in structured populations. *Am J Hum Genet* 67:170–181.
- 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.
- 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.

- Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, GuhaThakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusis AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich R. 2008. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 6:e107.
- Vanderas AP. 1987. Incidence of cleft lip, cleft palate, and cleft lip and palate among races: A review. *Cleft Palate J* 24:216–225.
- Weber JL, David D, Heil J, Fan Y, Zhao C, Marth G. 2002. Human diallelic insertion/deletion polymorphisms. *Am J Hum Genet* 71:854–862.
- 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.

Capítulo 3. *Region 8q24 is a Susceptibility Locus for Nonsyndromic Oral Clefting in Brazil*

Este capítulo se refere também ao objetivo 1, segundo o qual verificamos se variantes previamente associadas à fissura possuem um papel regulatório em células-tronco mesenquimais provenientes de músculo orbicular do lábio (OOMMSC) de pacientes FL/P NS. Para isso, foi verificada a correlação do genótipo da variante rs987525, localizada no deserto gênico 8q24.21, com os níveis de expressão do gene *MYC*.

Região 8q24 é um *locus* de susceptibilidade para fissura oral não-sindrômica no Brasil

Luciano Abreu Brito, Livia Máris Ribeiro Paranaíba, Camila Fernandes Silva Bassi, Cibele Masotti, Carolina Malcher, David Schlesinger, Katia Maria Rocha, Lucas Alvizi Cruz, Lígia Kobayashi Bárbara, Nivaldo Alonso, Diogo Franco, Elizabete Bagordakis, Hercílio Martelli Júnior, Diogo Meyer, Ricardo D. Coletta, Maria Rita Passos-Bueno

Birth Defects Research Part A: Clinical and Molecular Teratology. Junho, 94 (6): 464-468 (2012). Publicado online: 18 Abril 2012

Resumo

Fissura de lábio com ou sem fissura de palato não-sindrômica é um defeito craniofacial relativamente comum com herança multifatorial. A associação da variante de nucleotídeo único rs987525, localizada em um deserto gênico na região 8q24.21, tem sido consistentemente replicada em populações europeias. Nós realizamos uma abordagem de estudo de associação com correção para ancestralidade combinada com análise

transcricional do gene *MYC* para investigar o papel do rs987525 na susceptibilidade à fissura orofacial na população brasileira etnicamente miscigenada. MÉTODOS: Nós realizamos um estudo de associação condicionado às proporções ancestrais individuais em uma amostra de 563 pacientes e 336 controles, e em uma amostra independente de 221 pacientes e 261 controles. A correlação entre os genótipos de rs987525 e os níveis de transcrição de *MYC* em células-tronco mesenquimais de músculo orbicular do lábio também foi investigada em 42 pacientes e 4 controles. RESULTADOS: Nós encontramos uma associação significativa na amostra maior ($p=0,0016$; OR=1,8 [Intervalo de confiança 95% {IC}, 1,21 – 2,69], para o genótipo heterozigoto, e 2,71 [IC 95%, 1,47 – 4,96] para o genótipo homozigoto). Não encontramos uma correlação significativa entre os genótipos de rs987525 e os níveis de transcrição de *MYC* ($p=0,14$; $r= -0,22$, Correlação de Spearman). CONCLUSÕES: Nós apresentamos uma associação positiva para rs987525 na população brasileira pela primeira vez, e é provável que a contribuição européia para a nossa população esteja dirigindo essa associação. Nós também não podemos descartar um papel de rs987525 na regulação de *MYC*, porque esse *locus* se comporta como um eQTL (*locus* associado a traço quantitativo de expressão) de *MYC* em outro tecido.

Region 8q24 Is a Susceptibility Locus for Nonsyndromic Oral Clefting in Brazil

Luciano Abreu Brito,¹ Lívia Máris Ribeiro Paranaíba,^{2,3} Camila Fernandes Silva Bassi,¹ Cibeles Masotti,¹ Carolina Malcher,¹ David Schlesinger,^{1,4} Katia Maria Rocha,¹ Lucas Alvizi Cruz,¹ Lígia Kobayashi Bárbara,¹ Nivaldo Alonso,⁵ Diogo Franco,⁶ Elizabete Bagordakis,² Hercílio Martelli Jr.,^{3,7} Diogo Meyer,⁸ Ricardo D. Coletta,² and Maria Rita Passos-Bueno^{1*}

¹Human Genome Research Center, Institute of Biosciences, University of São Paulo, São Paulo, Brazil

²Department of Oral Diagnosis, School of Dentistry, State University of Campinas, Piracicaba, São Paulo, Brazil

³Stomatology Clinic, Dental School, State University of Montes Claros, Montes Claros, Minas Gerais, Brazil

⁴Instituto do Cérebro, Instituto Israelita de Ensino e Pesquisa Albert Einstein, São Paulo, Brazil

⁵Department of Surgery, School of Medicine, University of São Paulo, São Paulo, Brazil

⁶Department of Plastic Surgery, University of Rio de Janeiro, Rio de Janeiro, Brazil

⁷Center for Rehabilitation of Craniofacial Anomalies, Dental School, University of Alfenas, Minas Gerais, Brazil

⁸Department of Genetics, Institute of Biosciences, University of São Paulo, São Paulo, Brazil

Received 28 January 2012; Revised 2 March 2012; Accepted 6 March 2012

BACKGROUND: Nonsyndromic cleft lip with or without cleft palate is a relatively common craniofacial defect with multifactorial inheritance. The association of the rs987525 single nucleotide variant, located in a gene desert at 8q24.21 region, has been consistently replicated in European populations. We performed a structured association approach combined with transcriptional analysis of the MYC gene to dissect the role of rs987525 in oral clefting susceptibility in the ethnically admixed Brazilian population. **METHODS:** We performed the association study conditioned on the individual ancestry proportions in a sample of 563 patients and 336 controls, and in an independent sample of 221 patients and 261 controls. The correlation between rs987525 genotypes and MYC transcriptional levels in orbicularis oris muscle mesenchymal stem cells was also investigated in 42 patients and 4 controls. **RESULTS:** We found a significant association in the larger sample ($p = 0.0016$; OR = 1.80 [95% confidence interval (CI), 1.21–2.69], for heterozygous genotype, and 2.71 [95% CI, 1.47–4.96] for homozygous genotype). We did not find a significant correlation between rs987525 genotypes and MYC transcriptional levels ($p = 0.14$; $r = -0.22$, Spearman Correlation). **CONCLUSIONS:** We present a positive association of rs987525 in the Brazilian population for the first time, and it is likely that the European contribution to our population is driving this association. We also cannot discard a role of rs987515 in MYC regulation, because this locus behaves as an expression quantitative locus of MYC in another tissue. *Birth Defects Research (Part A) 94:464–468, 2012.* © 2012 Wiley Periodicals, Inc.

Key words: rs987525; rs1476165; rs2099897; heritability; cleft lip; cleft palate; 8q24 gene desert; MYC transcriptional levels; common disease-common variant; mesenchymal stem cells

INTRODUCTION

Nonsyndromic cleft lip with or without cleft palate (NS CL/P; OMIM 119530), which represents 70% of orofacial clefts and is the most common craniofacial anomaly, is always accompanied by physical and psychological impairments to affected children, with great family burden (Wehby and Cassell, 2010).

NS CL/P has a worldwide birth prevalence of approximately 1:700 live births (Dixon et al., 2011), varying according to ethnicity (African, 0.3:1000; European, 1.0:1000; Native American, 3.6:1000), geographic origin,

Additional Supporting Information may be found in the online version of this article.

Financial support from FAPESP/CEPID (98/14254-2), CNPq (401952/2010-S) and the Brazilian Ministry of Health.

*Correspondence to: Maria Rita Passos-Bueno, Rua do Matão, 277, sala 200, Cidade Universitária, São Paulo, SP 05508-090, Brazil. E-mail: passos@i-b.usp.br

Published online 18 April 2012 in Wiley Online Library (wileyonlinelibrary.com).

DOI: 10.1002/bdra.23011

and socioeconomic status (Carinci et al., 2007; Jugessur et al., 2009). In Brazil, NS CL/P prevalence has ranged from 0.28:1000 to 1.54:1000 (Menegotto and Salzano, 1991; Loffredo et al., 2001; Martelli-Junior et al., 2006). NS CL/P etiology is multifactorial and is determined by numerous interacting loci with additional environmental covariates (Mossey et al., 2009). The identification of genetic risk factors for NS CL/P has been the subject of intensive research, and genome-wide scans have shown a range of plausible candidate genes or chromosomal regions (Dixon et al., 2011). Of those, the most remarkable loci are the *IRF6* gene (Zuccherro et al., 2004; Rahimov et al., 2008) and the rs987525 single nucleotide variant (SNV) at 8q24.21 (Birnbbaum et al., 2009), as their association with NS CL/P has been largely replicated (Park et al., 2007; Jugessur et al., 2008; Grant et al., 2009; Huang et al., 2009; Nikopensius et al., 2009; Beaty et al., 2010; Blanton et al., 2010a, 2010b). Despite the several reported associations between rs987525 and NS CL/P, it is unknown whether this SNV is the cause of the association or if it is in linkage disequilibrium with still some unknown gene or region; nevertheless, a regulatory function has been credited for this region (Dixon et al., 2011). Furthermore, patients of European origin seem to be more at risk from this SNV (Beaty et al., 2010), and it is certainly important to evaluate its worldwide effect. Although the Brazilian population has a trihybrid constitution of European, African, and Native American parental populations, European contribution is as high as 80% (Santos et al., 2010), and it is possible that rs987525 contributes to oral clefting in our population. To test this hypothesis, and to avoid population stratification bias, we have conducted a structured association analysis in a large sample of patients ascertained in different localities of the country (Sample 1) and in a minor replication sample ascertained in the state of Minas Gerais (southeastern region; Sample 2). To provide new insights into the functional role of rs987525, we have also investigated the correlation between its genotypes and the expression levels of the *MYC* gene, which is an important regulator of a variety of cellular processes necessary for normal cell growth and differentiation (Secombe et al., 2004), located at approximately 1.1 Mb of rs987525.

MATERIALS AND METHODS

Ethics

All samples were collected in accordance with the rules of the Research Ethics Committee of the Institute of Biosciences, University of São Paulo, Brazil, and the Human Research Ethics Committee of the University of Alfenas, Brazil. Informed consent was obtained from the patients or legal tutors.

Patients and Controls

We ascertained two independent samples of unrelated cases and controls. Sample 1 comprises 563 NS CL/P patients from medical programs of Operation Smile (<http://www.operationsmile.org>) between 2007 and 2010 in five different Brazilian sites (Santarém, Fortaleza, Barbalha, Maceió, and Rio de Janeiro) and 336 controls from São Paulo (previously described by Errera et al., 2006), provided by our Biobank, because we were not able to ascertain local controls at the corresponding period.

Patients affected by cleft palate only were removed from the sample. Patient ascertainment methods were previously reported by Brito et al., (2011a). DNA extraction from an oral swab followed the manufacturer's protocols (Macherey-Nagel, Düren, Germany).

To replicate our findings, we recruited an independent sample of 221 patients and 261 controls (Sample 2) from the Center for Rehabilitation of Craniofacial Anomalies–Dental School, University of Alfenas, State of Minas Gerais (southeastern region of Brazil). Control individuals were subjects admitted as in-patients, but with conditions unrelated to clefting disorders and their endophenotypes. Genomic DNA of Sample 2 was extracted from oral mucosa cells, as described previously (Aidar and Line, 2007).

To verify the effect of rs987525 in the expression levels of the *MYC* gene, we established mesenchymal stem cell cultures (MSCs) from orbicularis oris muscle fragments (OOMF), discarded during surgical procedures of cheiloplasty, of an additional 42 NS CL/P patients and 4 controls submitted to surgery for other reasons than CL/P correction. These surgeries were performed in three different centers (Division of Plastic Surgery of University of São Paulo, SOBRAPAR Hospital–Campinas, and Operation Smile, Rio de Janeiro, Brazil).

Genotyping

All individuals were genotyped for the SNV rs987525 with the TaqMan method (Applied Biosystems, Foster City, CA, USA). To characterize the European, African, and Native American ancestry contributions, Samples 1 and 2 were genotyped for independent panels of 40 ancestry informative indels. FAM-labeled M13 tail was incorporated to forward primers, allowing multiplex PCR, with four to seven markers per reaction (sequences available upon request). These markers are spread across the genome and were accessed from the Marshfield database (<http://research.marshfieldclinic.org>), which holds a set of indel markers described by Weber et al. (2002). Sample 2 was genotyped for a panel extensively used for characterizing Brazilian populations (Bastos-Rodrigues et al., 2006), and Sample 1 was genotyped for an independent panel with matched marker selection criteria, as published previously (Brito et al., 2011a). Fragment analysis was performed with Gene Mapper software (Applied Biosystems) after capillary electrophoresis targeting the PCR products for the FAM-labeled M13 tail.

Cell Culture

The MSCs were established according to previously published protocol (Bueno et al., 2009) and expanded to 80% confluence. RNA and DNA extraction followed the manufacturer's protocols (Macherey-Nagel). We confirmed the homogeneous mesenchymal status of 10 cell cultures using five mesenchymal cell markers (CD29, CD90, CD105, SH3, and SH4), one hematopoietic (CD45), and one endothelial (CD31) through flow cytometric analysis (EasyCyte Flow Cytometer; Guava Technologies, Hayward, CA) according to standard protocol in our laboratory (Bueno et al., 2010).

Statistical Analyses

We used Structure 2.3.3 to infer ancestry components of each individual (Pritchard et al., 2000a; Falush et al.,

Figure 1. European (blue), African (green), and Native American (red) ancestry contributions for Samples 1 and 2. Each individual is represented by a single column. In the lower panel, EUR, AFR, and AMI represent the parental populations used as “learning samples” to assist Structure in estimating ancestry of the admixed individuals.

2003, 2007), in a run assuming $K = 3$ parental populations, based on the trihybrid origin of Brazilians; we also incorporated “learning samples” in the run, with prespecified population of origin (European, African, and Native American reference populations from Marshfield Clinic collection) to assist the software to estimate ancestry of the admixed individuals. Following the ancestry inference, STRAT was used to test the SNV for association, conditioning on the individual ancestry proportions (Pritchard et al., 2000b).

We obtained *MYC* expression values from microarray gene expression data (Human Gene 1.0 ST microarray chips; Affymetrix, Santa Clara-CA, USA), with the three-step robust multiarray average preprocessing method implemented in the software Expression Console (Affymetrix). We assessed the correlation between the rs987525 genotype and *MYC* expression levels with a Spearman correlation test (95% confidence interval [CI]).

RESULTS

Association Analyses

After excluding individuals with low genotype assignment (<80%), Sample 1 comprised 471 patients (18% presenting cleft lip only [CLO] and 82% presenting cleft lip and cleft palate [CLP]) and 326 controls. Average ancestry contributions were estimated at 22.2% African, 60.3% European, and 17.5% Native American for patients, and 18.8% African, 70.9% European, and 10.3% Native American for controls (individual ancestry profile is shown in Fig. 1). Ancestry analysis stratifying the patients for region of origin is shown in Table 1. Minor allele frequency (MAF) of rs987525 was similar in cases (34%) and controls (35%), and genotype distributions followed the Hardy-Weinberg expected proportions in both groups ($p > 0.48$). Despite the similar MAF between cases and controls, a significant association with the A allele was observed ($p = 0.006$) in the structured association test (Table 1). We further investigated whether the European ancestry would be driving this association by stratifying

the analysis into two groups: a European group with the 361 patients presenting more than 50% of European ancestry, and a non-European group with 110 patients with less than 50% of European ancestry, and then comparing both groups with the same 326 controls. We found a stronger association in the European group ($p = 0.008$) than in the non-European group ($p = 0.21$).

Sample 2 consisted of 196 patients (28% CLO and 72% CLP) and 263 controls, after excluding low-genotype assignment individuals. Average ancestry contributions of this sample were estimated at 13.4% African, 84.1% European, and 2.5% Native American for patients and 8.1% African, 89.0% European, and 2.9% Native American for controls (Fig. 1). Minor allele frequency was higher in patients (39.5%) than in controls (27.4%), and genotypes matched Hardy-Weinberg proportions in both groups ($p > 0.31$). A significant association between rs987525 and NS CL/P group was also observed in this sample ($p = 0.0016$). Because rs987525 allele frequencies vary greatly among continents, and differences in ancestry contribution between cases and controls are more pronounced in Sample 1, we considered that we would obtain a more reliable odds ratio (OR) estimates only for Sample 2, from where cases and controls come from a single location. For this group, OR was estimated at 1.80 (95% CI, 1.21–2.69) for the heterozygous genotype (AC versus CC) and 2.70 (95% CI, 1.47–4.96) for the homozygous genotype (AA versus CC), under an additive model.

MYC Expression Analyses

We found a nonsignificant correlation between rs987525 genotypes (Supporting Information Table S1) and *MYC* transcriptional levels in MSCs (Supporting Information Fig. S1; $p = 0.14$; $r = -0.22$, Spearman correlation).

DISCUSSION

It has been postulated that part of the susceptibility alleles for complex disorders might have a common origin

Table 1
Estimates of Average Ancestry Contributions of Parental Populations (European, African, and Native American), rs987525 Genotype Frequencies, Risk Allele (A) Frequency, and *p* Value of the Structured Association Tests

Sample	Subsamples	N	Ancestry contribution (%)			Genotype frequencies (%)				<i>p</i> value (H-W test)	<i>p</i> value (structured association)	OR (95% CI)
			European	African	Native American	AA	AC	CC	f(A)			
Sample 1	All Patients	471	60	22	18	12	45	44	34	0.92	0.006	
	Santarém	82	54	14	32	7	38	55	26	0.99	0.07	
	Barbalha	52	64	24	12	12	58	30	41	0.09	0.005	
	Fortaleza	163	64	20	16	12	42	46	33	0.42	0.33	
	Maceió	62	59	23	19	12	44	44	34	0.92	0.21	
	Rio de Janeiro	112	58	31	11	14	48	40	38	0.79	0.19	
Sample 2	Controls	326	71	19	10	13	44	42	35	0.48	—	
	Patients	196	84	13	3	16	47	37	40	0.69	0.0016	AC: 1.80 (1.21–2.69) AA: 2.70 (1.47–4.96)
	Controls	263	89	8	3	9	37	57	21	0.31		

The results are presented for Sample 1 as a whole and stratified by location of origin of the patients (Santarém, Barbalha, Fortaleza, Maceió, and Rio de Janeiro). The odds ratio was only estimated for Sample 2, since cases and controls from this sample are less stratified. The odds ratios for heterozygous (AC; $p = 0.004$) and homozygous (AA; $p = 0.001$) genotypes were estimated by comparing with CC genotype, under an additive model. CA, frequency of the risk allele (A); OR, odds ratio; CI, confidence interval; H-W, Hardy-Weinberg.

(Lander, 2011). Therefore, replicating the NS CL/P association studies results in other populations is relevant, not only to define the common or population-specific risk alleles, but to construct the best model to explain the disorder.

In this study, we show a positive association between rs987525 and NS CL/P in the Brazilian population through a structured association approach. It is interesting to note that the level of significance was much higher in Sample 2, which has a greater contribution of European ancestry compared with Sample 1. The higher OR for homozygous genotype (AA) suggested an additive model, similar to those observed in previous studies with European populations (Birnbbaum et al., 2009; Nikopienius et al., 2009).

Although we did not observe a significant correlation between rs987525 genotypes and the expression levels of MYC (~ 1.1 Mb apart from rs987525) in MSCs, this locus behaves as an expression quantitative trait locus of MYC on adipose tissue according to the GENEVAR database ($p = 0.01$; MuTHER study, Yang et al., 2010; Nica et al., 2011). It is of note that other SNVs at 8q24, rs1476165 and rs2099897 (802 and 831 kb apart from the 5' end of MYC, respectively), were found to modulate MYC expression tissue-specifically, according to GENEVAR (Yang et al., 2010; Nica et al., 2011), suggesting that this region might be of great relevance to MYC regulation. Given the importance of MYC on cell proliferation control and the absence of other known genes within this region playing a role in craniofacial development, it would be of interest to conduct additional studies on the effect of rs987525 in MYC transcriptional regulation. It is possible that we did not detect the effect of rs987525 genotypes in the expression levels of MYC because of our limited sample size or, alternatively, MSCs of OOMF might not be the best cell type to conduct such studies.

Our results confirm the importance of this locus to NS CL/P predisposition and provide additional evidence that the true causal variant beyond this association might be of European origin, because a positive association between rs987525 and clefting patients has been shown in a variety of European patients of different regions of the world, such as Germany, Estonia, Lithuania, and Poland, and European descents in United States (Birnbbaum et al., 2009; Grant et al., 2009; Nikopienius et al., 2009; Beaty et al., 2010; Mostowska et al., 2010), in contrast to the lack of association in patients of East Asian origin (Beaty et al., 2010). Considering the trihybrid origin of Brazilians and that the association observed in Sample 1 was driven by the subset with more than 50% of European contribution, it is also likely that this allele is not so relevant in populations of African descent. Indeed, a recent report by (Weatherley-White et al., 2011) showed no evidence of association of rs987525 in patients from Kenya; therefore, this SNV might track a functionally relevant variant only in the European haplotype block.

We have previously shown that heritability of NS CL/P in Brazil is high (up to 85%) and varies according to the geographic region. We have also predicted that heritability might be a confounding factor in association studies of heterogeneous diseases such as NS CL/P (Brito et al., 2011a). In this regard, we have recently observed a positive association between the traditional GWAS hit rs642961, in the *IRF6* gene, and NS CL/P only in the

sample with the highest heritability, which comes from Barbalha (Brito et al., 2012). We have therefore stratified the present association analysis by location of origin, and we have verified that the Barbalha group reached the highest significance ($p = 0.005$), compared with samples from Santarém ($p = 0.07$), Fortaleza ($p = 0.33$), Maceió ($p = 0.21$), and Rio de Janeiro ($p = 0.19$; Table 1). These results provide confirm heritability as a confounding factor in association analyses.

In summary, our results provided further evidence that rs987525 is involved with NS CL/P susceptibility in patients of European descent, and heritability seemed to be an important confounding factor. In addition, we did not find evidence of rs987525 modulating MYC expression.

ACKNOWLEDGMENTS

We thank the Operation Smile team and its executive director Clovis Brito, Constância Urbani and Márcia Teixeira for secretarial assistance, and the Brazilian Ministry of Health, FAPESP, and CNPq for financial support.

REFERENCES

- Aidar M, Line SR. 2007. A simple and cost-effective protocol for DNA isolation from buccal epithelial cells. *Braz Dent J* 18:148–152.
- Bastos-Rodrigues L, Pimenta JR, Pena SD, et al. 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, et al. 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, et al. 2009. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet* 41:473–477.
- Blanton SH, Burt A, Garcia E, et al. 2010a. Ethnic heterogeneity of IRF6 AP-2a binding site promoter SNP association with nonsyndromic cleft lip and palate. *Cleft Palate Craniofac J* 47:574–577.
- Blanton SH, Burt A, Stal S, et al. 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* 88:256–259.
- Brito LA, Cruz LA, Rocha KM, et al. 2011a. Genetic contribution for nonsyndromic cleft lip with or without cleft palate (NS CL/P) in different regions of Brazil and implications for association studies. *Am J Med Genet A* 155A:1581–1587.
- Brito LA, Bassi CFS, Masotti C, et al. 2012. IRF6 is a risk factor for nonsyndromic cleft lip in the Brazilian Population. *Am J Med Genet A*, in press.
- Bueno DF, Kerkis I, Costa AM, et al. 2009. New source of muscle-derived stem cells with potential for alveolar bone reconstruction in cleft lip and/or palate patients. *Tissue Eng Part A* 15:427–435.
- Bueno DF, Sunaga DY, Kobayashi GS, et al. 2010. 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* 7:446–457.
- Carinci F, Scapoli L, Palmieri A, et al. 2007. Human genetic factors in nonsyndromic cleft lip and palate: an update. *Int J Pediatr Otorhinolaryngol* 71:1509–1519.
- Dixon MJ, Marazita ML, Beaty TH, et al. 2011. Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev* 12:167–178.
- Errera FI, Silva ME, Yeh E, et al. 2006. Effect of polymorphisms of the MTHFR and APOE genes on susceptibility to diabetes and severity of diabetic retinopathy in Brazilian patients. *Braz J Med Biol Res* 39:883–888.
- Falush D, Stephens M, Pritchard JK, et al. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.
- Falush D, Stephens M, Pritchard JK, et al. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7:574–578.
- Grant SF, Wang K, Zhang H, et al. 2009. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *J Pediatr* 155:909–913.
- Huang Y, Wu J, Ma J, et al. 2009. Association between IRF6 SNPs and oral clefts in West China. *J Dent Res* 88:715–718.
- Jugessur A, Farlie PG, Kilpatrick N, et al. 2009. The genetics of isolated orofacial clefts: from genotypes to subphenotypes. *Oral Dis* 15:437–453.
- Jugessur A, Rahimov F, Lie RT, et al. 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* 32:413–424.
- Lander ES. 2011. Initial impact of the sequencing of the human genome. *Nature* 470:187–197.
- Loffredo LCM, Freitas JAS, Grigolli AAG, et al. 2001. Prevalence of oral clefts from 1975 to 1994, Brazil. *Rev Saúde Pública* 35:571–575.
- Menegotto BG, Salzano FM. 1991. Epidemiology of oral clefts in a large South American sample. *Cleft Palate Craniofac J* 28:373–376, discussion 376–377.
- Martelli-Junior H, Orsi-Junior J, Chaves MR, et al. 2006. Estudo epidemiológico das fissuras labiais e palatais em Alfenas - Minas Gerais - de 1986 a 1998. *RPG Rev Pos-grad* 13:31–35.
- Mossey PA, Little J, Munger RG, et al. 2009. Cleft lip and palate. *Lancet* 374:1773–1785.
- Mostowska A, Hozyasz KK, Wojcicki P, et al. 2010. Association between genetic variants of reported candidate genes or regions and risk of cleft lip with or without cleft palate in the Polish population. *Birth Defects Res A Clin Mol Teratol* 88:538–545.
- Nica AC, Parts L, Glass D, et al. 2011. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS genetics* 7:e1002003.
- Nikopensius T, Ambrozaityte L, Ludwig KU, et al. 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:2551–2553.
- Park JW, McIntosh I, Hetmanski JB, et al. 2007. Association between IRF6 and nonsyndromic cleft lip with or without cleft palate in four populations. *Genet Med* 9:219–227.
- Pritchard JK, Stephens M, Donnelly P, et al. 2000a. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Pritchard JK, Stephens M, Rosenberg NA, et al. 2000b. Association mapping in structured populations. *Am J Hum Genet* 67:170–181.
- Rahimov F, Marazita ML, Visel A, et al. 2008. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. *Nat Genet* 40:1341–1347.
- Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, et al. 2010. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat* 31:184–190.
- Secombe J, Pierce SB, Eisenman RN, et al. 2004. Myc: a weapon of mass destruction. *Cell* 117:153–156.
- Weatherley-White RC, Ben S, Jin Y, et al. 2011. Analysis of genome-wide association signals for nonsyndromic cleft lip/palate in a Kenya African Cohort. *Am J Med Genet A* 155A:2422–2425.
- Weber JL, David D, Heil J, et al. 2002. Human diallelic insertion/deletion polymorphisms. *Am J Hum Genet* 71:854–862.
- Wehby GL, Cassell CH. 2010. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral Dis* 16:3–10.
- Yang TP, Beazley C, Montgomery SB, et al. 2010. Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics (Oxford, England)* 26:2474–2476.
- Zuccherro TM, Cooper ME, Maher BS, et al. 2004. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *New Engl J Med* 351:769–780.

Capítulo 4. *Use of eQTLs to Elucidate the Effect of SNVs in the Etiology of Nonsyndromic Oral Clefting*

Esse capítulo se refere ao objetivo 2 de verificar se eQTLs mapeados em OOMMSC de pacientes FL/P NS estão associados à doença. Para isso, foram testados cinco eQTLs para associação com fissura: rs5011163, rs1505443, rs4793213, rs4793229 e rs1242500.

Uso de eQTLs para elucidar o efeito de SNVs na etiologia da fissura oral não-sindrômica

Carolina Malcher¹, Cibeles Masotti¹, Luciano Abreu Brito¹, Camila F. S. Bassi¹, Andressa Gois Morales¹, Simone Gomes Ferreira¹, Daniela Franco Bueno¹, Gerson Shigeru Kobayashi¹, Nivaldo Alonso², and Maria Rita Passos-Bueno^{1*}.

¹Human Genome Research Center, Institute de Biosciences, University of São Paulo, São Paulo, Brazil

²Division of Plastic Surgery, Department of Surgery, School of Medicine of São Paulo University, São Paulo, Brazil

Resumo:

Fissura de lábio com ou sem fissura de palato não-sindrômica (FL/P NS) é uma doença complexa, determinada pela interação entre fatores de susceptibilidade genéticos e ambientais. Dentre os fatores genéticos, diversos *loci* de susceptibilidade têm sido encontrados, porém o mecanismo molecular permanece desconhecido. Visto que a variação da expressão gênica é um importante mecanismo em doenças complexas, nós investigamos

se *loci* associados a traços quantitativos de expressão (eQTL - *expression quantitative trait loci*), mapeados em amostras de células-tronco mesenquimais derivadas de músculo orbicular do lábio, um tecido afetado em FL/P NS, são associados à doença. Apesar do fato de não termos encontrado associação para as cinco variantes testadas, nossos resultados excluem alguns mecanismos potenciais através dos quais essas variantes poderiam estar envolvidas na etiologia da FL/P NS. Uma abordagem genômica será necessária para avaliar o papel dos eQTLs na susceptibilidade à FL/P NS.

Palavras-chave: FL/P NS, eQTL, estudos de associação

Abstract:

Nonsyndromic cleft lip with or without cleft palate (NSCL/P) is a complex disorder determined by the interaction between genetic and environmental susceptibility factors. Among the genetic factors, several susceptibility loci have been found, but the underlying molecular mechanisms remains elusive. Considering that variation of gene expression is an important mechanism in complex diseases, we investigated if expression quantitative trait loci (eQTLs), mapped in mesenchymal stem cells derived from *orbicular oris* muscle, a tissue affected in the NSCL/P phenotype, are associated with the disease itself. Despite the fact that we did not find association for the 5 variants tested, our results exclude some potential mechanisms through which these variants could be involved in the etiology of NSCL/P. A genomic approach will be necessary in order to evaluate the role of the eQTLs in NSCL/P susceptibility.

Key words: NSCL/P, eQTL, association studies

Introduction

Cleft lip with or without palate (CL/P) is one of the most frequent congenital birth defects (GORLIN *et al.*, 2001). It affects approximately 1:700 live births, with ethnic and geographic variability (DIXON *et al.*, 2011; MOSSEY & CASTILLIA, 2003; VANDERAS, 1987). The nonsyndromic forms (NSCL/P) are characterized by CL/P only and represent the majority of the cases (70%). It is still a challenge to disentangle the genetic components involved with their etiology. The most accepted model of inheritance is the multifactorial one, in which both genetic and environmental factors play a role (DIXON *et al.*, 2011; GROSEN *et al.*, 2010).

With the advance of genome-wide studies, several susceptibility variants have been found, mostly in noncoding regions (about 90% of all variants) (HARDY & SINGLETON, 2009; HINDORFF *et al.*, 2009). Variants in noncoding regions may be involved in regulation of gene expression, and the identification of the genes being targeted by these variants in a genome-wide scale has been approached through expression quantitative trait loci (eQTL) mapping (COOKSON *et al.*, 2009; NICA *et al.*, 2010). A proportion of eQTLs has been demonstrated to be associated with susceptibility mechanisms for complex diseases, such as asthma and Crohn's disease (LIBIOULLE *et al.*, 2007; MOFFATT *et al.*, 2007; NICA *et al.*, 2010). It also has been shown that trait-associated SNVs (single nucleotide variant) are more likely to be eQTLs, and also that eQTLs found in tissues that play a role in the disease are more likely to be associated with the disease itself (KANG *et al.*, 2012; NICOLAE *et al.*, 2010). Therefore, using eQTLs as a strategy to prioritize variants for association studies is a valuable strategy to accurately find disease-associated variants.

We selected 5 eQTLs previously mapped in mesenchymal stem cells derived from primary cultures of *orbicularis oris* muscle (OOMMSC) of NSCL/P patients and controls, and tested these variants for association with NSCL/P susceptibility. These variants have a regulatory role in other tissues and regulate genes known or suspected to be important for development or the pathogenesis of orofacial clefts.

Methodology

Subjects

The sample consisted of 627 patients ascertained from the medical program of a nongovernmental organization – Operation Smile (<http://www.operationsmile.org/>) - in five different Brazilian cities: Barbalha-CE (n=66), Santarém-PA (n=91), Rio de Janeiro-RJ (n=107), Fortaleza-CE (n=241), and Maceió-AL (n=122). The control group consisted of 470 individuals from São Paulo (n=405) and Fortaleza (n=65). Detailed subject ascertainment method and inclusion criteria have been previously described by our group (BRITO *et al.*, 2012a, b; ERRERA *et al.*, 2006; ORABONA *et al.*, 2009).

Informed consent was obtained from each participant included in this study, as approved by the Research Ethics Committee of the Institute of Biosciences, University of São Paulo, Brazil. Only patients with nonsyndromic cleft lip with or without cleft palate were included in this study. Genomic DNA was purified from oral swabs using an extraction kit (Macherey-Nagel, Duren, Germany), according to the manufacturer's recommendations.

Single-nucleotide variant (SNV) selection and genotyping

We selected 5 putative functional SNVs (Table 1) from the eQTL mapping study conducted in *orbicular oris* muscle mesenchymal stem cells (OOMMSC) derived from NSCL/P patients and controls (MASOTTI *et al.*, manuscript in preparation). We used as criteria for SNV selection their regulatory role in other tissues and their relevance for NSCL/P susceptibility or embryonic development.

Table 1: SNVs tested for association.

SNV	Alleles*	Cytoband	Physical location (hg19)
rs5011163	a/G	8q24.21	chr8:128,675,862
rs1505443	a/T	3q29	chr3:192,433,199
rs4793213	c/T	17q21.31	chr17:41,307,101
rs4793229	A/g	17q21.31	chr17:41,418,334
rs1242500	c/T	17p11.2	chr17:17,379,382

*Lower case: Minor Allele.

The SNV rs5011163 is located in the 8q24.21 region, previously associated with NSCL/P. It is approximately 1.2 Mb distant from rs987525, which is the marker with the strongest association with NSCL/P (BIRNBAUM *et al.*, 2009). This variant is an eQTL of *U1* in OOMMSC and also of *MYC* in lymphoblastoid cell lines (STRANGER *et al.*, 2012) according to Genevar (YANG *et al.*, 2010). *U1* is a small nuclear RNA, part of the spliceosome, and also the small nuclear RNA *U5E* is found to be downregulated in NSCL/P individuals (KOBAYASHI *et al.*, manuscript submitted for publication), which make genes from this family interesting candidates to be tested.

rs1505443 is located in the intron 2 of *FGF12* and, besides being an eQTL of *HRALS* in OOMMSC, it is also an eQTL of *FGF12* in lymphoblastoid cell lines (STRANGER *et al.*, 2012; YANG *et al.*, 2010). *FGF12* has been associated with NSCL/P and with cancer, and *HRALS* is a gene highly expressed in cancer (CHEN *et al.*, 2008; JUGESSUR *et al.*, 2009).

We selected the SNVs rs4793213 and rs4793229 ($D'=1$, $r^2=0.694$) as tagSNVs of four other eQTLs of *NBR2* (neighbor of *BRCA1* gene 2) in OOMMSC (Supplementary Table 1). For that we used the software Haploview (BARRET *et al.*, 2005), using as cutoff parameter for linkage disequilibrium $r^2>0.8$. These two SNVs are 29,501bp and 242,734bp away from the *NBR2* gene, respectively. These SNVs are also eQTLs of *BRCA1* in monocytes and lymphoblastoid cell lines (STRANGER *et al.*, 2012; YANG *et al.*, 2010; ZELLER *et al.*, 2010;). *BRCA1* encodes a nuclear phosphoprotein that plays a role in maintaining genomic stability, and acts as a tumor suppressor (WU *et al.*, 2010), and dysregulation of *BRCA1* pathway has

been recently observed in NSCL/P individuals (KOBAYASHI *et al.*, manuscript submitted for publication).

Finally, we decided to test rs1242500, an eQTL of *DRG2* in both OOMMSC and lymphoblastoid cells (STRANGER *et al.*, 2012; YANG *et al.*, 2010). *DRG2* (developmentally regulated GTP binding protein 2) plays an essential role in the control of cell growth and differentiation, and it is located in region 17p11.2, which has been previously associated with cleft palate predisposition (ANDRIEUX *et al.*, 2007; SERTIÉ *et al.*, 1999; SONG *et al.*, 2004; VLANGOS *et al.*, 2000).

All individuals were genotyped with the Taqman method (Applied Biosystems, Foster City, CA, USA). To correct for population structure, the individuals were also genotyped for a panel of 40 ancestry-informative markers (AIMs) to characterize the individual proportions of European, African and Amerindian ancestry (data available under request). The AIMs consisted of indel (insertion/deletion) polymorphisms spread across the autosomal genome, and were retrieved from the Marshfield database (<http://research.marshfieldclinic.org>), which holds a set of indel markers described by Weber *et al.* (2002). Marker selection was conducted as described by Brito *et al.* (2012a, b). An M13 tail labeled with FAM fluorochrome was added to the 5' end of each forward primer, so that PCR products incorporated one tail, allowing for multiplex PCR with 4 to 7 markers per reaction (primer sequences available upon request). Fragment analysis was performed with the Gene Mapper software (Applied Biosystems, Foster City, CA), after capillary electrophoresis targeting the PCR products with the FAM-labeled M13 tail.

Statistical analysis

Allele and genotype frequencies were calculated for each locus and tested for Hardy-Weinberg Equilibrium (HWE) among controls. Prior to the association test, we used the Structure 2.3.3 software to infer the ancestry proportion for each individual (FALUSH *et al.*, 2003, 2007; PRITCHARD *et al.*, 2000a), assuming K=3 subpopulations based on the well-known tri-hybrid origin of Brazilians, and incorporated “learning samples” in the run, with pre-specified population of origin (European, African, and Amerindian) according to the

Mammalian Genotyping Service database (Marshfield Clinic, Marshfield-WI) to assist the ancestry discrimination. We then used the ancestry information as a co-variable to make a structured association test using STRAT software (PRITCHARD *et al.*, 2000b). We assume 0.01 as significance threshold after Bonferroni adjustment for multiple testing.

Results

The final sample comprised 454 patients and 379 controls, all with high genotype assignment (>70%). The mean ancestry contribution of European, African and Amerindian in the case group was 0.6/0.22/0.18, while it was 0.7/0.19/0.11 in the control group (Figure 1).

Fig. 1: Bar plot representing the contribution of European (blue), African (red), and Amerindian (green) ancestries to our sample, in which each single column represents an individual. The panel is divided into NSCL/P patients, controls, and parental populations (EUR, Europeans, AFR, Africans, AMI, Amerindians), which were used as “learning samples” to assist the software in the ancestry estimations.

Chi-square analysis showed no evidence of deviation from HWE for all of the SNVs in our control sample. The observed genotypes and minor allele frequencies, as well as the HWE and p-values are shown in Table 2.

Table 2: Values of genotype and allele frequency, HWE and association p-values for the 5 SNVs.

		Patients	Controls
	N	454	379
rs5011163	Genotype frequency (AA/AG/GG)	11.3/51/37.7 (%)	14.7/47.6/37.7 (%)
	MAF (A)	0.36	0.38
	HWE p-value	-	0.92
	Association p-value	0.057	
rs1505443	Genotype frequency (AA/AT/TT)	4.2/31.8/64 (%)	5.2/31.2/63.6 (%)
	MAF (A)	0.2	0.21
	HWE p-value	-	0.28
	Association p-value	0.664	
rs4793213	Genotype frequency (CC/CT/TT)	7/39/54 (%)	4.8/39.6/55.6 (%)
	MAF (C)	0.26	0.24
	HWE p-value	-	0.2
	Association p-value	0.6	
rs4793229	Genotype frequency (AA/AG/GG)	44/44/12 (%)	42.5/46.5/11 (%)
	MAF (G)	0.34	0.34
	HWE p-value	-	0.52
	Association p-value	0.9	
rs1242500	Genotype frequency (CC/CT/TT)	12/47/41 (%)	9.5/43.8/46.7 (%)
	MAF (C)	0.35	0.31
	HWE p-value	-	0.74
	Association p-value	0.4	

MAF = minor allele frequency.

We did not observe any significant association between the tested SNVs and NSCL/P. A marginal association between rs5011163 and NSCL/P in the Rio de Janeiro population ($p=0.045$) and between rs5011163 and NSCL/P familial cases ($p=0.04$) was observed. However, these associations were lost after correcting for multiple testing (Table 3).

Table 3: Association studies p-values after stratifying the affected sample by ascertainment region, type of cleft and familial recurrence.

		rs5011163	rs1505443	rs4793213	rs4793229	rs1242500
Ascertainment region	Barbalha (N=51)	p = 0.55	p = 0.95	p = 0.9	p = 0.64	p = 0.36
	Santarem (N=81)	p = 0.28	p = 0.28	p = 0.48	p = 0.86	p = 0.66
	Rio de Janeiro (N=104)	p = 0.045*	p = 0.29	p = 0.95	p = 0.87	p = 0.5
	Fortaleza (N=158)	p = 0.29	p = 0.85	p = 0.64	p = 0.76	p = 0.55
	Maceió (N=60)	p = 0.23	p = 0.77	p = 0.25	p = 0.42	p = 0.7
Type of cleft	CLO (N=79)**	p = 0.52	p = 0.52	p = 0.12	p = 0.15	p = 0.95
	CLP (N=373)**	p = 0.11	p = 0.76	p = 0.67	p = 0.95	p = 0.17
Familial recurrence	Isolated (N=284)	p = 0.1	p = 0.53	p = 0.87	p = 0.99	p = 0.35
	Familial (N=164)	p = 0.04*	p = 0.63	p = 0.46	p = 0.78	p = 0.8

*Not significant after Bonferroni adjustment

**CLO: cleft lip only; CLP: cleft lip and palate

Discussion

Attributing a functional role to the many variants associated with NSCL/P has been proven to be a difficult task. One possible strategy that might contribute to translate the effect of SNVs into how they mediate disease susceptibility is studying functional variants mapped in an affected tissue, and assessing their role in regulating candidate genes to the phenotype.

rs5011163 is an eQTL of *U1*, a 164-nt-long spliceosomal small nuclear RNA, required at the very first stage of the splicing process (REED, 1996; SCHELLENBERG *et al.*, 2008). The involvement of other spliceosome components in craniofacial disorders has been demonstrated, as for Mandibulofacial Dysostosis with Microcephaly (MFDM, OMIM #610536) and Nager Syndrome (OMIM #154400). In MFDM, which includes cleft palate in a

significant proportion of the cases, the causative gene (*EFTUD2*) encodes U5-116KD, a highly conserved ribonucleoproteic spliceosomal GTPase with a central regulatory role in catalytic splicing and post-splicing-complex disassembly (LINES *et al.*, 2012). In Nager syndrome, which is characterized by malformation of the craniofacial skeleton and the limbs, the mutated gene (*SF3B4*) is a component of the U2 pre-mRNA spliceosomal complex (BERNIER *et al.*, 2012). Also, the small nuclear RNA U5E is found to be downregulated in NSCL/P individuals (KOBAYASHI *et al.*, manuscript submitted for publication). The SNV rs5011163 is also an eQTL of *MYC* in lymphoblastoid cell lines (STRANGER *et al.*, 2012; YANG *et al.*, 2010), whereas *MYC* regulates several developmental processes, including palate growth (JI *et al.*, 2011; ZHU *et al.*, 2012). However, we did not find association between this variant and NSCL/P, which suggests that *U1* is not involved in NSCL/P etiology.

The eQTL rs1505443 is located within intron 2 of *FGF12* (ENST00000445105). Besides being an eQTL of *HRASLS* in OOMMSC, it has been reported to be an eQTL of *FGF12* in lymphoblastoid cell lines (STRANGER *et al.*, 2012; YANG *et al.*, 2010). *FGF12* has been previously associated with NSCL/P (JUGESSUR *et al.*, 2009). However, the previously associated SNVs are not in LD with the eQTL rs1505443 tested here. This result suggests that this eQTL is not related to the mechanism through which *FGF12* plays a role in NSCL/P susceptibility.

The eQTLs rs4793213 and rs4793229 ($r^2=0.694$) regulate the expression levels of *NBR2* in OOMMSC as well as *BRCA1* in monocytes and lymphoblastoid cell lines (STRANGER *et al.*, 2012; YANG *et al.*, 2010; ZELLER *et al.*, 2010), and *NBR2* and *BRCA1* share a bi-directional promoter (SUEN *et al.*, 2005). Transcriptional dysregulation of the *BRCA1* repair pathway among NSCL/P individuals has been observed by our group (KOBAYASHI *et al.*, manuscript submitted for publication), which could be determined, at least in part, by these SNVs. Our negative findings, however, do not support this hypothesis and exclude rs4793213 and rs4793229 as variants directly involved with NSCL/P susceptibility.

We also tested rs1242500, an eQTL of *DRG2* in both OOMMSC and lymphoblastoid cells (STRANGER *et al.*, 2012; YANG *et al.*, 2010). *DRG2*, which encodes a GTP-binding protein known to function in the regulation of cell growth and differentiation, is mapped at 17p11.2, a region previously pointed out to contain candidate genes for cleft palate (CP) (ANDRIEUX *et al.*, 2007; SERTIÉ *et al.*, 1999; SONG *et al.*, 2004; VLANGOS *et al.*, 2000). We did not find association between this SNV and NSCL/P. There are other potential candidates in this

region and it is possible that these genes are relevant mainly to CP, which is known to have a different etiological mechanism from CL/P.

Even though we were not able to find association among any of the selected eQTLs and the disease, we have excluded some potential mechanisms through which the suggested candidate SNVs could lead to NSCL/P. Importantly, a genomic analysis will be necessary in order to evaluate the relevance of eQTLs in NSCL/P susceptibility.

References

ANDRIEUX, J.; VILLENET, C.; QUIEF, S.; LIGNON, S.; GEFFROY, S.; ROUMIER, C.; DE LEERSNYDER, H.; DE BLOIS, M. C.; MANOUVRIER, S.; DELOBEL, B.; BENZACKEN, B.; BITOUN, P.; ATTIE-BITACH, T.; THOMAS, S.; LYONNET, S.; VEKEMANS, M.; KERCKAERT, J. P. Genotype–phenotype correlation of 30 patients with Smith–Magenis syndrome (SMS) using comparative genome hybridisation array: cleft palate in SMS is associated with larger deletions. **J Med Genet.** v.; 44, p. 537–540, 2007

BARRETT, J. C.; FRY, B.; MALLER, J.; DALY, M. J. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics.** n. 21, p. 263–265, 2005.

BERNIER, F. P.; CALUSERIU, O.; NG, S.; SCHWARTZENTRUBER, J.; BUCKINGHAM, K. J.; INNES, A. M.; JABS, E. W.; INNIS, J. W.; SCHUETTE, J. L.; GORSKI, J. L.; BYERS, P. H.; ANDELFINGER, G.; SIU, V.; LAUZON, J.; FERNANDEZ, B. A.; MCMILLIN, M.; SCOTT, R. H.; RACHER, H.; FORGE CANADA CONSORTIUM; MAJEWSKI, J.; NICKERSON, D. A.; SHENDURE, J.; BAMSHAD, M. J.; PARBOOSINGH, J. S. Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. **Am J Hum Genet.** v. 90, n. 5, p. 925-33, 2012.

BIRNBAUM, S.; LUDWIG, K.; REUTTER, H.; HERMS, S.; STEFFENS, M.; RUBINI, M.; BALUARDO, C.; FERRIAN, M.; ASSIS, N.; ALBLAS, M.; BARTH, S.; FREUDENBERG, J.; LAUSTER, C.; SCHMIDT, G.; SCHEER, M.; BRAUMANN, B.; BERGÉ, S.; REICH, R.; SCHIEFKE, F.; HEMPRICH, A.; POTZSCH, S.; STEEGERS-THEUNISSEN, R.; POTZSCH, B.; MOEBUS, S.; HORSTHEMKE, B.; KRAMER, F.; WIENKER, T.; MOSSEY, P.; PROPPING, P.; CICHON, S.; HOFFMAN, P.; KNAPP, M.; NOTHEN, M.; MANGOLD, E. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. **Nature Genetics.** v. 41, n. 4, p. 473-477, 2009.

BRITO, L. A.; BASSI, C.; F. S.; MASOTTI, C.; MALCHER, C.; ROCHA, K. M.; SCHLESINGER, D.; BUENO, D. F.; CRUZ, L. A.; BÁRBARA, L. K.; BERTOLA, D. R.; MEYER, D.; FRANCO, D.; ALONSO, N.; PASSOS-BUENO, M. R. IRF6 is a risk factor for nonsyndromic cleft lip in the Brazilian population. **American Journal of Medical Genetics: Part A**. v. 158A, p. 2170–2175, 2012a.

BRITO, L. A.; PARANAIBA, L. M.; BASSI, C. F. S.; MASOTTI, C.; MALCHER, C.; SCHLESINGER, D.; ROCHA, K., M.; CRUZ, L. A.; BÁRBARA, L. K.; ALONSO, N.; FRANCO, D.; BAGORDAKIS, E.; MARTELLI JÚNIOR, H.; MEYER, D.; COLETTA, R. D.; PASSOS-BUENO, M. R. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. **Birth Defects Research Part A: Clinical and Molecular Teratology**. v. 94, p. 464-468, 2012b.

BRITO, A. B.; MEIRA, J. G. C.; KOBAYASHI, G. S.; PASSOS-BUENO, M. R. Genetics and Management of the Patient with Orofacial Cleft. **Plast Surg Int**. 2012c.

CHEN, J.; GUO, L.; PEIFFER, D. A.; ZHOU, L.; CHAN, O. T. M.; BIBIKOVA, M.; WICKHAM-GARCIA, E.; LU, S.; ZHAN, Q.; WANG-RODRIGUEZ, J.; JIANG, W.; FAN, J. Genomic profiling of 766 cancer-related genes in archived esophageal normal and carcinoma tissues. **Int. J. Cancer**. v. 122, p. 2249–2254, 2008.

COOKSON, W.; LIANG, L.; ABECASIS, G.; MOFFATT, M.; LATHROP, M. Mapping complex disease traits with global gene expression. **Nature Review Genetics**. v. 10, p. 184-94, 2009.

DIXON, M.; MARAZITA, M.; BEATY T.; MURRAY J. Cleft lip and palate: understanding genetic and environmental influences. **Nat Rev Genet**, v. 12, p.167-178, 2011.

ERRERA, F. I.; SILVA, M. E.; YEH, E.; MARANDUBA, C. M.; FOLCO, B.; TAKAHASHI, W.; PEREIRA, A. C.; KRIEGER, J. E.; PASSOS-BUENO MR. Effect of polymorphisms of the MTHFR and APOE genes on susceptibility to diabetes and severity of diabetic retinopathy in Brazilian patients. **Braz J Med Biol Res**. v. 39, n. 7, p. 883–888, 2006.

FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. **Genetics**. v. 164, n. 4, p. 1567–1587, 2003.

FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: dominant markers and null alleles. **Mol Ecol Notes**. v. 7, n. 4, p. 574–578, 2007.

GORLIN, R.; COHEN, M.; HENNEKAM, R. Orofacial Clefting Syndromes: General Aspects In:_____. **Syndromes of the head and neck**. Oxford: Oxford University Press, 2001. Cap. 20, p. 850.

GROSEN, D.; CHEVRIER, C.; SKYTTE, A.; BILLE, C.; MØLSTED, K.; SIVERTSEN, A.; MURRAY, J.; CHRISTENSEN, K. 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*. v. 47, p. 162-168, 2010.

HALLONET, M.; HOLLEMANN, T.; PIELER, T.; GRUSS P. Vax1, a novel homeobox-containing gene, directs development of the basal forebrain and visual system. **Genes Dev**. v. 13, n. 23, p. 3106-14, 1999.

HARDY, J.; SINGLETON, A. Genomewide Association Studies and Human Disease. **The New England Journal of Medicine**. v. 360, n. 17, p. 1759-1768, 2009.

HINDORFF, L. A.; SETHUPATHY, P.; JUNKINS, H. A.; RAMOS, E. M.; MEHTA, J. P.; COLLINS, F. S.; MANOLIO, T. A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. **PNAS**. v. 106, n. 23, p. 9362–9367, 2009.

JI, H.; WU, G.; ZHAN, X.; NOLAN, A.; KOH, C.; DE MARZO, A.; DOAN, H. M.; FAN, J.; CHEADLE, C.; FALLAHI, M.; CLEVELAND, J. L.; DANG, C. V.; ZELLER, K. I. Cell-type independent MYC target genes reveal a primordial signature involved in biomass accumulation. **PLoS One**. v. 6, n. 10, 2011.

JUGESSUR, A.; GJESSING, K.; LIE, R.; WILCOX, A.; SHI, M.; WEINBERG, C.; CHRISTENSEN, K.; BOYLES, A.; DAACK-HIRSCH, S.; NGUYEN, T.; BILLE, C.; LIDRAL, A.; MURRAY, J. GENETIC Determinants of Facial Clefting : Analysis of 357 Candidate Genes Using Two National Cleft Studies from Scandinavia. **PLoS ONE**. v. 4, n. 4, 2009.

KANG H. P.; MORGAN, A.A.; CHEN, R.; SCHADT, E. E.; BUTTE, A J. Coanalysis of GWAS with eQTLs reveals disease-tissue associations. **AMIA Summits Transl Sci Proc**. v. 2012 p. 35–41, 2012.

LIBIOULLE, C.; LOUIS, E.; HANSOUL, S.; SANDOR, C.; FARNIR, F.; FRANCHIMONT, D.; VERMEIRE, S.; DEWIT, O.; DE VOS, M.; DIXON, A.; DEMARCHE, B.; GUT, I.; HEATH, S.; FOGGIO, M.; LIANG, L.; LAUKENS, D.; MNI, M.; ZELENIKA, D.; VAN GOSSUM, A.; RUTGEERTS, P.; BELAICHE, J.; LATHROP, M.; GEORGES, M. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. **PLoS Genet**. v. 3, n 4, p. 538-543, 2007.

LINES, M. A.; HUANG, L.; SCHWARTZENTRUBER, J.; DOUGLAS, S. L.; LYNCH, D. C.; BEAULIEU, C.; GUION-ALMEIDA, M. L.; ZECHI-CEIDE, R. M.; GENER, B.; GILLESSEN-KAESBACH, G.; NAVA, C.; BAUJAT, G.; HORN, D.; KINI, U.; CALIEBE, A.; ALANAY, Y.; UTINE, G. E.; LEV, D.; KOHLHASE, J.; GRIX, A. W.; LOHMANN, D. R.; HEHR, U.; BÖHM, D.; FORGE CANADA CONSORTIUM; MAJEWSKI, J.; BULMAN, D. E.; WIECZOREK, D.; BOYCOTT KM. Haploinsufficiency of a spliceosomal GTPase encoded by EFTUD2 causes mandibulofacial dysostosis with microcephaly. **Am J Hum Genet.** v. 90, n. 2, p. 369-77, 2012.

MOFFATT, M. F.; KABESCH, M.; LIANG, L.; DIXON, A. L.; STRACHAN, D.; HEATH, S.; DEPNER, M.; VON BERG, A.; BUFE, A.; RIETSCHEL, E.; HEINZMANN, A.; SIMMA, B.; FRISCHER, T.; WILLIS-OWEN, S. A. G.; WONG, K. C. C.; ILLIG, T.; VOGELBERG, C.; WEILAND, S. K.; VON MUTIUS, E.; ABECASIS, G. R.; FARRALL, M.; GUT, I. G.; LATHROP, G. M.; COOKSON, W. O. C. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. **Nature.** v. 448, p. 470-474, 2007.

MOSSEY, P.; CASTILLIA, E. **Global registry and database on craniofacial anomalies...**2003. Relatório da Reunião da Organização Mundial de Saúde em Anomalias Craniofaciais, Geneva, 2003.

NICA, A. C.; MONTGOMERY, S. B.; DIMAS, A. S.; STRANGER, B. E.; BEAZLEY, C.; BARROSO, I.; DERMITZAKIS, E. T. Candidate Causal Regulatory Effects by Integration of Expression QTLs with Complex Trait Genetic Associations. **PLoS Genet.** v. 6, n. 4, 2010.

NICOLAE, D. L.; GAMAZON, E.; ZHANG, W.; DUAN, S.; DOLAN, M. E. COX, N. J. Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS. **PLoS Genet.** v. 6, n. 4, 2010.

ORABONA, G. M.; GRIESI-OLIVEIRA, K.; VADASZ, E.; BULCAO, V. L.; TAKAHASHI, V. N.; MOREIRA, E. S.; FURIA-SILVA, M.; ROS-MELO, A. M.; DOURADO, F.; MATIOLI, S. R.; OTTO, P.; PASSOS-BUENO, M. R. HTR1B and HTR2C in autism spectrum disorders in Brazilian families. **Brain Res.** v. 1250, p. 14–19, 2009.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics.** v. 155, p. 945– 959, 2000a.

PRITCHARD, J. K.; STEPHENS, M.; ROSENBERG, N. A.; DONNELLY, P. Association mapping in structured populations. **Am J Hum Genet.** v. 67, p. 170–181, 2000b.

REED, R. Initial splice-site recognition and pairing during pre-mRNA splicing. **Curr Opin Genet Dev.** v. 6, n. 2, p. 215-20, 1996.

SCHELLENBERG, M. J.; RITCHIE, D. B.; MACMILLAN, A. M. Pre-mRNA splicing: a complex picture in higher definition. **Trends Biochem. Sci.** v, 33, p. 243–246, 2008.

SERTIÉ, A. L.; SOUSA, A. V.; STEMAN, S.; PAVANELLO, R. C.; PASSOS-BUENO, M. R. Linkage analysis in a large Brazilian family with van der Woude syndrome suggests the existence of a susceptibility locus for cleft palate at 17p11.2-11.1. **Am J Hum Genet.** v. 65, n. 2, p. 433-40, 1999.

SONG, H.; KIM, S. I.; KO, M. S.; KIM, H. J.; HEO, J. C.; LEE, H. J.; LEE, H. S.; HAN, I. S.; KWACK, K.; PARK, J. W. Overexpression of DRG2 increases G2/M phase cells and decreases sensitivity to nocodazole-induced apoptosis. **J Biochem.** v. 135, n. 3, p. 331-5, 2004.

STRANGER, B. E.; MONTGOMERY, S. B.; DIMAS, A. S.; PARTS, L.; STEGLE, O.; INGLE, C. E.; SEKOWSKA, M.; SMITH, G. D.; EVANS, D.; GUTIERREZ-ARCELUS, M.; PRICE, A.; RAJ, T.; NISBETT, J.; NICA, A. C.; BEAZLEY, C.; DURBIN, R.; DELOUKAS, P.; DERMITZAKIS, E. T. Patterns of cis regulatory variation in diverse human populations. **PLoS Genet.** v. 8, n. 4, 2012.

SUEN, T. C.; TANG, M. S.; GOSS, P. E. Model of transcriptional regulation of the BRCA1–NBR2 bi-directional transcriptional unit. **Biochimica et Biophysica Acta.** v. 1728, n. 3, p. 126– 134, 2005.

VANDERAS, A. P. Incidence of Cleft Lip, Cleft Palate, and Cleft Lip and Palate Among Races: A Review. **Cleft Palate Journal.** v. 24, n. 3, 1987.

VLANGOS, C. N.; DAS, P.; PATEL, P. I.; ELSEA, S. H. Assignment of developmentally regulated GTP-binding protein (DRG2) to human chromosome band 17p11.2 with somatic cell hybrids and localization to the Smith-Magenis syndrome critical interval. **Cytogenet Cell Genet.** v. 88, p. 283-5, 2000.

WEBER, J. L.; DAVID, D.; HEIL, J.; FAN, Y.; ZHAO, C.; MARTH, G. Human diallelic insertion/deletion polymorphisms. **Am J Hum Genet.** v. 71, p. 854–862, 2002.

WU, J.; LU, L. Y.; YU, X. The role of BRCA1 in DNA damage response. **Protein Cell.** v. 1, n. 2 p. 117-23, 2010.

YANG, T. P.; BEAZLEY, C.; MONTGOMERY, S. B.; DIMAS, A. S.; GUTIERREZ-ARCELUS, M.; STRANGER, B. E.; DELOUKAS, P.; DERMITZAKIS, E. T. Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. **Bioinformatics**. v. 26, n. 19, p. 2474-6, 2010.

ZELLER, T.; WILD, P.; SZYMCZAK, S.; ROTIVAL, M.; SCHILLERT, A.; CASTAGNE, R.; MAOUCHE, S.; GERMAIN, M.; LACKNER, K.; ROSSMANN, H.; ELEFThERIADIS, M.; SINNING, C. R.; SCHNABEL, R. B.; LUBOS, E.; MENNERICH, D.; RUST, W.; PERRET, C.; PROUST, C.; NICAUD, V.; LOSCALZO, J.; HÜBNER, N.; TREGOUET, D.; MÜNZEL, T.; ZIEGLER, A.; TIRET, L.; BLANKENBERG, S.; CAMBIEN F. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. **PLoS One**. v. 5, n. 5, 2010.

ZHU, X.; OZTURK, F.; LIU, C.; OAKLEY, G. G.; NAWSHAD, A. Transforming growth factor- β activates c-Myc to promote palatal growth. **J Cell Biochem**. v. 113, n. 10, p. 3069-85, 2012.

Supplementary Material

Supplementary Table 1: eQTLs of *NBR2* and linkage disequilibrium (LD) coefficients between them. Top: r^2 ; Bottom: D' .

$D' \setminus r^2$	rs4793229	rs4793213	rs4793197	rs4520874
rs4793229		0.694	1	1
rs4793213	1		0.694	0.687
rs4793197	1	1		1
rs4520874	1	1	1	

Capítulo 5. Discussão Geral e Conclusões

Com o advento dos estudos de associação de genoma inteiro, diversas variantes têm sido associadas à FL/P NS. Entretanto, a contribuição biológica de cada um desses *loci* para a determinação da susceptibilidade à doença ainda é pouco esclarecida. A elucidação de funções biológicas chave envolvidas na susceptibilidade à doença provê um importante avanço para um completo entendimento da etiologia da doença.

A integração de dados de expressão gênica com resultados de estudos de associação pode ser informativa para auxiliar na descoberta de genes e vias biológicas envolvidas na susceptibilidade à doença. Visto que há um enriquecimento de eQTLs dentre variantes associadas a doenças complexas (KNIGHT, 2005; NICA *et al.*, 2010; NICOLAE *et al.*, 2010), e enriquecimento de variantes associadas a doenças complexas dentre eQTLs tecido-específicos (DING *et al.*, 2010; KANG *et al.*, 2012a,b; RICHARDS *et al.*, 2012), o presente trabalho integra estudo de associação e análise de eQTLs de duas maneiras: (1) verificando se variantes associadas à FL/P NS possuem efeito regulatório em células-tronco de músculo orbicular do lábio (OOMMSC), e (2) verificando se eQTLs mapeados em OOMMSC estão associados à FL/P NS.

Alguns trabalhos já demonstraram a utilidade da integração das abordagens de estudos de associação e eQTLs na elucidação da etiologia de doenças complexas, porém nenhum foi realizado para FL/P NS. Fransen e colaboradores (2010) utilizaram como estratégia para a descoberta de *loci* associados à doença de Crohn a priorização dos SNVs com base no seu efeito funcional. Dentre as 500 variantes com maior sinal de associação, selecionaram para uma análise subsequente aquelas que tinham correlação com níveis de expressão. Essa abordagem culminou na identificação de dois novos genes de risco para a doença: *UBE2L3* e *BCL3*. *UBE2L3* já foi associado a outras doença auto-imunes (doença celíaca, artrite reumatóide e lupus eritematoso sistêmico, DUBOIS *et al.*, 2010; HAN *et al.*, 2009; STAHL *et al.*, 2010), as quais sabidamente compartilham *loci* de risco com a doença de Crohn. O eQTL de *UBE2L3* associado à doença está correlacionado a maiores níveis de expressão do mesmo gene. Teoricamente, uma superexpressão de *UBE2L3* poderia levar a uma degradação mais rápida do precursor de NF- κ B, e assim a uma menor produção de NF-

kB, consequentemente diminuindo a resposta imune inata. Um efeito similar é visto para de variantes de risco do gene *NOD2*, o *locus* com maior risco para doença de Crohn. Dessa forma, essa estratégia se mostrou válida para a identificação de *loci* de risco que tenham um efeito funcional para a doença.

eQTLs são em sua maioria tecido-específicos (69-80%, DIMAS *et al.*, 2009), o que corrobora a importância de estudar expressão gênica em um tecido afetado pela doença. Sendo assim, escolhemos utilizar células-tronco mesenquimais de músculo orbicular do lábio (OOMMSC), o qual é um tecido afetado na doença, mesmo em formas subclínicas (WEINBERG *et al.*, 2008). Visto que a formação do lábio e palato ocorre entre a quarta e sexta semanas de gestação (GORLIN *et al.*, 2001), esse tipo celular seria o representante disponível mais próximo de uma célula embrionária, proveniente de um tecido afetado pela doença, o que o torna um modelo interessante para o estudo de FL/P NS.

Porém, ao analisar as variantes rs642961, rs590223 e rs987525, já associadas à FL/P NS e com papel regulatório anteriormente sugerido, não observamos correlação entre genótipo e níveis de expressão dos genes analisados. Isso pode ser devido ao pequeno número amostral analisado (N=46), uma vez que a frequência do alelo de menor frequência desses SNPs é 17,19% (A), 34% (G), 32,47% (A), respectivamente. Alternativamente, é possível que essas variantes somente interfiram em momentos críticos do desenvolvimento embrionário, sendo necessário observar os níveis de expressão em células embrionárias envolvidas na formação das estruturas orofaciais afetadas nas FL/P. Para isso, uma alternativa seria o uso da reprogramação celular (iPSC, *induced Pluripotent Stem Cells*, TAKAHASHI *et al.*, 2007), o que permitiria avaliar a expressão gênica de células indiferenciadas do próprio paciente, as quais tem potencial de originar as estruturas embrionárias afetadas na doença.s

Ao verificar se variantes regulatórias estariam associadas à FL/P NS, utilizamos eQTLs mapeados em OOMMSC provenientes de um projeto de pós-doutorado realizado no nosso grupo, o qual mapeou 119 eQTLs, relacionados a 18 genes ($p < 0.0001$; FDR=14%; Masotti *et al.*, manuscrito em preparação). Para o estudo de associação, não priorizamos os eQTLs mais significantes; escolhemos aqueles localizados em regiões genômicas com associação prévia à doença ou que estivessem correlacionados com genes de desenvolvimento, conforme explicitado no Capítulo 4.

Não detectamos associação de nenhuma das cinco variantes testadas com FL/P NS, as quais não foram associadas a nenhum outro fenótipo por GWAS (Disponível em: www.genome.gov/gwastudies). Isso pode ter sido devido à significância marginal dos eQTLs escolhidos para estudo de associação caso-controle (FDR>42.5%). Além disso, a casuística utilizada (N=46) para o mapeamento é bastante pequena em comparação com estudos de eQTLs que vêm sendo realizados (GRUNDBERG *et al.*, 2012). Em estudos com baixo poder, os eQTLs detectados são aqueles de maior efeito, e esses geralmente são compartilhados entre diferentes tecidos (NICA *et al.*, 2011; GRUNDBERG *et al.*, 2012). Sendo assim, esses cinco eQTLs, podem não ter relevância para a doença em si.

O presente trabalho vem a contribuir com a integração das abordagens de estudo de associação e eQTLs pela primeira vez em FL/P NS. Essa integração já demonstrou sucesso para outras doenças e nossos resultados ainda não são conclusivos para FL/P NS. Para julgarmos a relevância dessa abordagem para a identificação de variantes de suscetibilidade para a doença, está em andamento o estudo de associação do tipo caso-controle para outros eQTLs de OOMMSC. Cem eQTLs foram escolhidos e priorizamos aqueles mais significantes. Além disso, escolhemos eQTLs que regulam genes associados à FL/P NS que por sua vez também estão associados a módulos de co-expressão gênica correlacionados à variabilidade clínica de FL/P NS (MASOTTI *et al.*, manuscrito em preparação). Ainda, testaremos eQTLs que regulam ou que estão localizados em regiões regulatórias de genes associados a vias de reparo de dano de DNA, recentemente relacionada à FL/P NS por nosso grupo (KOBAYASHI *et al.*, manuscrito submetido).

Resumo

Fissura de lábio com ou sem fissura de palato não-sindrômica (FL/P NS) é uma malformação craniofacial frequente, com modelo de herança multifatorial, onde fatores de risco genéticos e ambientais atuam na manifestação da doença. Variações nos níveis de expressão gênica têm sido apontadas como um importante mecanismo de susceptibilidade a doenças complexas, e variantes no DNA que regulam esses níveis de expressão (eQTL) têm sido combinadas a estudos de associação para auxiliar no entendimento da etiologia de algumas doenças. No presente trabalho, integramos eQTLs e estudo de associação para 1) verificar se variantes já associadas com FL/P NS possuem um papel regulatório em células-tronco de músculo orbicular do lábio (OOMMSC, um tecido afetado em FL/P NS), e 2) verificar se eQTLs mapeados em OOMMSC teriam associação com a mesma. Para o primeiro objetivo, verificamos a correlação entre os genótipos das variantes rs642961 e rs590223 e os níveis de expressão de *IRF6*, e também entre rs987525 e os níveis de expressão de *MYC*. Não encontramos correlação para nenhuma das três variantes testadas. É possível que essas variantes possuam um papel funcional em algum momento específico da embriogênese, ou mesmo que não tenhamos detectado essa correlação devido ao número amostral analisado (N=46). Para o segundo objetivo, realizamos um estudo de associação do tipo caso-controle dos eQTLs rs5011163, rs1505443, rs4793213, rs4793229 e rs1242500. Não encontramos associação entre nenhuma das cinco variantes e FL/P NS. Uma possível explicação para a associação negativa seria a significância marginal dessas variantes como eQTLs em OOMMSC. Além disso, estudos com baixo poder, como o mapeamento de eQTLs em OOMMSC realizado em outro projeto pelo nosso grupo, geralmente detectam os eQTLs de maior efeito, sendo esses frequentemente compartilhados entre tecidos, e, assim, podem não ter relevância para a doença em si. Outros eQTLs de OOMMSC, selecionados por critérios diferentes do presente estudo, estão sendo testados para associação com FL/P NS, o que nos permitirá avaliar a relevância dessa abordagem para detectar variantes de susceptibilidade a FL/P NS.

Abstract

Nonsyndromic cleft lip with or without cleft palate (NSCL/P) is a frequent craniofacial malformation, with a multifactorial model of inheritance, in which genetic and environmental risk factors act in disease manifestation. Variation of gene expression has been pointed as an important susceptibility mechanism to complex diseases, and DNA variants that regulate expression levels (eQTLs) have been combined with association studies to help elucidate the etiology of some diseases. In the present work, we integrate eQTL and association studies to 1) verify if variants associated with NSCL/P have a regulatory role in *orbicularis oris* muscle mesenchymal stem cells (OOMMSC, a tissue affected by NSCL/P); and 2) verify if eQTLs mapped in OOMMSC are associated with the disease. For the first goal, we verified the correlation of the rs642961 and rs590223 genotype variants with *IRF6* expression levels, and also between the rs987525 genotype and *MYC* expression levels. We did not find correlation for any of the three variants tested. Possibly, these variants have a functional role in specific moments of embryogenesis, or sample size (N=46) was insufficient to detect correlation. For the second goal, we did a case-control association study for eQTLs rs5011163, rs1505443, rs4793213, rs4793229 and rs1242500. We did not find association between these variants and NSCL/P. The negative association could be explained by the marginal significance of these variants as eQTLs in OOMMSC. Besides, low-power studies, as the OOMMSC eQTL mapping performed in another project by our group, usually detect eQTLs of larger effect, which are frequently shared among tissues; therefore, they may not be relevant for the disease itself. Other eQTLs, selected under different criteria, are currently being tested for association with NSCL/P, which will enable us to evaluate the relevance of this approach to detect susceptibility variants for NSCL/P.

Referências Bibliográficas

ARDINGER, H. H.; BUETOW, K. H.; BELL, G. I.; BARDACH, J.; VAN DEMARK, D. R.; MURRAY, D. R. Association of genetic variation of the transforming growth factor-alpha gene with cleft lip and palate. **American Journal of Human Genetics**. v. 45, n. 3, p. 348–353, 1989.

BEATY, T.; MURRAY, J.; MARAZITA, M. MUNGER, R. .; RUCZINSKI, I.; HETMANSKI, J.; LIANG, K.; WU, T.; MURRAY, T.; FALLIN, M.; REDETT, R.; RAYMOND, G.; SCHWENDER, H.; JIN, S.; COOPER, M.; DUNNWALD, M.; MANSILLA, M.; LESLIE, E.; BULLARD, S.; LIDRAL, A.; MORENO, L.; MENEZES, R.; VIEIRA, A.; PETRIN, A.; WILCOX, A.; LIE, R.; JABS, E.; WU-CHOU, Y.; CHEN, P.; WANG, H.; YE, X.; HUANG, S.; YEOW, V.; CHONG, S.; JEE, S.; SHI, B.; CHRISTENSEN, K.; MELBYE, M.; DOHENY, K.; PUGH, E.; LING, H.; CASTILLA, E.; CZEIZEL, A.; MA, L.; FIELD, L.; BRODY, L.; PANGILINAN, F.; MILLS, J.; MOLLOY, A.; KIRKE, P.; SCOTT, J.; ARCOS-BURGOS, M.; SCOTT, A. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. **Nature Genetics**. v. 42, n. 6, p. 525-529, 2010.

BLANTON, S.; CORTEZ, A.; STAL, S.; MULLIKEN, J.; FINNELL, R.; HECHT, J. Variation in IRF6 contributes to nonsyndromic cleft lip and palate. **American Journal of Medical Genetics**, v. 137, p. 259–262, 2005.

BHASKAR, L. V.; MURTHY, J.; VENKATESH BABU, G. Polymorphisms in genes involved in folate metabolism and orofacial clefts. **Arch Oral Biol**. v. 56, n. 8, p. 723-37, 2011.

BIRNBAUM, S.; LUDWIG, K.; REUTTER, H.; HERMS, S.; STEFFENS, M.; RUBINI, M.; BALUARDO, C.; FERRIAN, M.; ASSIS, N.; ALBLAS, M.; BARTH, S.; FREUDENBERG, J.; LAUSTER, C.; SCHMIDT, G.; SCHEER, M.; BRAUMANN, B.; BERGÉ, S.; REICH, R.; SCHIEFKE, F.; HEMPRICH, A.; POTZSCH, S.; STEEGERS-THEUNISSEN, R.; POTZSCH, B.; MOEBUS, S.; HORSTHEMKE, B.; KRAMER, F.; WIENKER, T.; MOSSEY, P.; PROPPING, P.; CICHON, S.; HOFFMAN, P.; KNAPP, M.; NOTHEN, M.; MANGOLD, E. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. **Nature Genetics**. v. 41, n. 4, p. 473-477, 2009a.

BIRNBAUM, S.; LUDWIG, K.; REUTTER, H.; HERMS, S.; DE ASSIS, N. A.; DIAZ-LACAVA, A.; BARTH, S.; LAUSTER, C.; SCHMIDT, G.; SCHEER, M.; SAFFAR, M.; MARTINI, M.; REICH, R. H.; SCHIEFKE, F.; HEMPRICH, A.; POTZSCH, S.; POTZSCH, B.; WIENKER, T. F.; HOFFMANN, P.; KNAPP, M.; KRAMER, F.-J.; NOTHEN, M. M.; MANGOLD, E. IRF6 gene variants in Central European patients with non-syndromic cleft lip with or without cleft palate. **Eur J Oral Sci**. v. 117, p. 766-769, 2009b.

BLANTON, S. H.; HENRY, R. R.; YUAN, Q.; MULLIKEN, J. B.; STAL, S.; FINNELL, R. H.; HECHT, J. T. Folate pathway and nonsyndromic cleft lip and palate. **Birth Defects Res A Clin Mol Teratol.** v. 91, n. 1, p. 50-60, 2011.

BODMER, W.; BONILLA, C. Common and rare variants in multifactorial susceptibility to common diseases. **Nat. Genet.** v. 40, n. 6, p. 695-701, 2008

BOYLES, A. L.; DEROO, L. A.; LIE, R. T.; TAYLOR, J. A.; JUGESSUR, A.; MURRAY, J. C.; WILCOX, A. J. Maternal alcohol consumption, alcohol metabolism genes, and the risk of oral clefts: a population-based case-control study in Norway, 1996–2001. **Am. J. Epidemiol.** v. 172, p. 924–931, 2010.

BRITO, L. A.; CRUZ, L. A.; ROCHA, K. M.; BÁRBARA, L. K.; SILVA, C. B. F.; BUENO, D. F.; AGUENA, M.; BERTOLA, D. R.; FRANCO, D.; COSTA, A. M.; ALONSO, N.; OTTO, P. A.; PASSOS-BUENO, M. R. 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. **American Journal of Medical Genetics: Part A.** v.155, p.1581–1587, 2011.

BRITO, L. A.; BASSI, C.; F. S.; MASOTTI, C.; MALCHER, C.; ROCHA, K. M.; SCHLESINGER, D.; BUENO, D. F.; CRUZ, L. A.; BÁRBARA, L. K.; BERTOLA, D. R.; MEYER, D.; FRANCO, D.; ALONSO, N.; PASSOS-BUENO, M. R. IRF6 is a risk factor for nonsyndromic cleft lip in the Brazilian population. **American Journal of Medical Genetics: Part A.** v. 158A, p. 2170–2175, 2012a.

BRITO, L. A.; PARANAIBA, L. M.; BASSI, C. F. S.; MASOTTI, C.; MALCHER, C.; SCHLESINGER, D.; ROCHA, K. M.; CRUZ, L. A.; BÁRBARA, L. K.; ALONSO, N.; FRANCO, D.; BAGORDAKIS, E.; MARTELLI JÚNIOR, H.; MEYER, D.; COLETTA, R. D.; PASSOS-BUENO, M. R. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. **Birth Defects Research Part A: Clinical and Molecular Teratology.** v. 94, p. 464-468, 2012b.

BUFALINO, A.; RIBEIRO, PARANAÍBA, L. M.; NASCIMENTO DE AQUINO, S.; MARTELLI-JÚNIOR, H.; OLIVEIRA SWERTS, M. S.; COLETTA, R. D. Maternal polymorphisms in folic acid metabolic genes are associated with nonsyndromic cleft lip and/or palate in the Brazilian population. **Birth Defects Res Part A.** v. 88, p. 980–6, 2010.

BURDICK, A. B.; BIXLER, D.; PUCKETT, C. I. Genetic analysis in families with Van Der Woude syndrome. **Journal of Craniofacial Genetics and Developmental Biology.** v. 5, n. 2, p. 181–208, 1985.

CHEUNG, V. G.; CONLIN, L. K.; WEBER, T. M.; ARCARO, M.; JEN, K.; MORLEY, M.; SPIELMAN, R. S. Natural variation in human gene expression assessed in lymphoblastoid cells. **Nat. Genet.** v. 33, p. 422–425, 2003.

CHEVRIER, C.; BAHUAU, M.; PERRET, C.; IOVANNISCI, D. M.; NELVA, A.; HERMAN, C.; VAZQUEZ, M. P.; FRANCCANNET, C.; ROBERT-GNANSIA, E.; LAMMER, E. J.; CORDIER S. Genetic susceptibilities in the association between maternal exposure to tobacco smoke and the risk of nonsyndromic oral cleft. **Am J Med Genet A**. v. 146A, n. 18, p. 2396-406, 2008.

CHRISTENSEN, K.; FOGH-ANDERSEN, P. Cleft lip (+/- cleft palate) in Danish twins. **Am J Med Genet**. v. 47, p. 10-916, 1993.

DE JONG, S.; VAN EIJK, K. R.; ZEEGERS, D. W.; STRENGMAN, E.; JANSON, E.; VELDINK, J. H.; VAN DEN BERG, L. H.; CAHN, W.; KAHN, R. S.; BOKS, M. P. M.; OPHOFF, R. A.; THE PGC SCHIZOPHRENIA (GWAS) CONSORTIUM. Expression QTL analysis of top loci from GWAS meta-analysis highlights additional schizophrenia candidate genes. **European Journal of Human Genetics**. v. 20, p. 1004–1008, 2012.

DEROO, L. A.; WILCOX, A. J.; DREVON, C. A.; LIE, R. T. 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, p. 638–646, 2008.

DIMAS, A.; DEUTSCH, S.; STRANGER, B.; MONTGOMERY, S.; BOREL, C.; ATTAR-COHEN, H.; INGLE, C.; BEAZLEY, C.; ARCELUS, M.; SEKOWSKA, M.; GAGNEBIN, M.; NISBETT, J.; DELOUKAS, P.; DERMITZAKIS, E.; ANTONARAKIS, S. Common regulatory variation impacts gene expression in a cell type-dependent manner. **Science**. v. 325, p. 1246-1250, 2009.

DING, J.; GUDJONSSON, J. E.; LIANG, L.; STUART, P. E.; LI, Y.; CHEN, W.; WEICHENTHAL, M.; ELLINGHAUS, E.; FRANKE, A.; COOKSON, W.; NAIR, R. P.; ELDER, J. T.; ABECASIS, G. R. Gene expression in skin and lymphoblastoid cells: refined statistical method reveals extensive overlap in cis-eQTL signals. **Am J Hum Genet**. v. 87, p. 779–789, 2010.

DIXON, M.; MARAZITA, M.; BEATY T.; MURRAY J. Cleft lip and palate: understanding genetic and environmental influences. **Nat Rev Genet**, v. 12, p.167-178, 2011.

DUBOIS, P. C.; TRYNKA, G.; FRANKE, L.; HUNT, K. A.; ROMANOS, J.; CURTOTTI, A.; ZHERNAKOVA, A.; HEAP, G. A.; ADÁNY, R.; AROMAA, A.; BARDELLA, M. T.; VAN DEN BERG, L. H.; BOCKETT, N. A.; DE LA CONCHA, E. G.; DEMA, B.; FEHRMANN, R. S.; FERNÁNDEZ-ARQUERO, M.; FIATAL, S.; GRANDONE, E.; GREEN, P. M.; GROEN, H. J.; GWILLIAM, R.; HOUWEN, R. H.; HUNT, S. E.; KAUKINEN, K.; KELLEHER, D.; KORPONAY-SZABO, I.; KURPPA, K.; MACMATHUNA, P.; MÄKI, M.; MAZZILLI, M. C.; MCCANN, O. T.; MEARIN, M. L.; MEIN, C. A.; MIRZA, M. M.; MISTRY, V.; MORA, B.; MORLEY, K. I.; MULDER, C. J.; MURRAY, J. A.; NÚÑEZ, C.; OOSTEROM, E.; OPHOFF, R. A.; POLANCO, I.; PELTONEN, L.; PLATTEEL, M.; RYBAK, A.; SALOMAA, V.; SCHWEIZER, J. J.; SPERANDEO, M. P.; TACK, G. J.; TURNER, G.; VELDINK, J. H.;

VERBEEK, W. H.; WEERSMA, R. K.; WOLTERS, V. M.; URCELAY, E.; CUKROWSKA, B.; GRECO, L.; NEUHAUSEN, S. L.; MCMANUS, R.; BARISANI, D.; DELOUKAS, P.; BARRETT, J. C.; SAAVALAINEN, P.; WIJMENG, C.; VAN HEEL, D. A. Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* v. 42, n. 4, p. 295 – 302, 2010.

FRANSEN, K.; VISSCHEDIJK, M. C.; VAN SOMMEREN, S.; FU, J. Y.; FRANKE, L.; FESTEN, E. A. M.; STOKKERS, P. C. F.; VAN BODEGRAVEN, A. A.; CRUSIUS, J. B. A.; HOMMES, D. W.; ZANEN, P.; DE JONG, D. J.; WIJMENG, C.; VAN DIEMEN, C. C.; WEERSMA, R. K. Analysis of SNPs with an effect on gene expression identifies UBE2L3 and BCL3 as potential new risk genes for Crohn's disease. **Human Molecular Genetics.** v. 19, n. 17, 2010.

GHASSIBÉ, M.; BAYET, B.; REVENCU, N.; VERELLEN-DUMOULIN, C.; GILLEROT, Y.; VANWIJCK, R.; VIKKULA, M. Interferon regulatory factor-6: a gene predisposing to isolated cleft lip with or without cleft palate in the Belgian population. **European journal of human genetics.** v. 13, p. 1239–1242, 2005.

GORLIN, R.; COHEN, M.; HENNEKAM, R. Orofacial Clefting Syndromes: General Aspects In:_____. **Syndromes of the head and neck.** Oxford: Oxford University Press, 2001. Cap. 20, p. 850.

GRANT, S.; WANG, K.; ZHANG, H.; GLABERSON, W.; ANNAIAH, K.; KIM, C.; BRADFIELD, J.; GLESSNER, J.; THOMAS, K.; GARRIS, M.; FRACKELTON, E.; OTIENO, F. GEORGE, C.; ROSETTA, M.; NAH, H.; KIRSCHNER, R.; HAKONARSON, H. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. **The Journal of pediatrics.** v. 155, n. 6, p. 909-13, 2009.

GROSEN, D.; CHEVRIER, C.; SKYTTE, A.; BILLE, C.; MØLSTED, K.; SIVERTSEN, A.; MURRAY, J.; CHRISTENSEN, K. 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.** v. 47, p. 162-168, 2010.

GRUNDBERG, E.; SMALL, K. S.; HEDMAN, Å. K.; NICA, A. C.; BUIL, A.; KEILDSON, S.; BELL, J. T.; YANG, T. P.; MEDURI, E.; BARRETT, A.; NISBETT, J.; SEKOWSKA, M.; WILK, A.; SHIN, S. Y.; GLASS, D.; TRAVERS, M.; MIN, J. L.; RING, S.; HO, K.; THORLEIFSSON, G.; KONG, A.; THORSTEINDOTTIR, U.; AINALI, C.; DIMAS, A. S.; HASSANALI, N.; INGLE, C.; KNOWLES, D.; KRESTYANINOVA, M.; LOWE, C. E.; DI MEGLIO, P.; MONTGOMERY, S. B.; PARTS, L.; POTTER, S.; SURDULESCU, G.; TSAPROUNI, L.; TSOKA, S.; BATAILLE, V.; DURBIN, R.; NESTLE, F. O.; O'RAHILLY, S.; SORANZO, N.; LINDGREN, C. M.; ZONDERVAN, K. T.; AHMADI, K. R.; SCHADT, E. E.; STEFANSSON, K.; SMITH, G. D.; MCCARTHY, M. I.; DELOUKAS, P.; DERMITZAKIS, E. T.; SPECTOR, T. D.; MULTIPLE TISSUE HUMAN EXPRESSION RESOURCE (MUTHER) CONSORTIUM. Mapping cis- and trans-regulatory effects across multiple tissues in twins. **Nat Genet.** v. 44, n. 10, p. 1084-9, 2012.

HARDY, J.; SINGLETON, A. Genomewide Association Studies and Human Disease. **The New England Journal of Medicine**. v. 360, n. 17, p. 1759-1768, 2009.

HECHT, J.; YANG, P.; MICHELST, V.; BUETOWT, K. Complex Segregation Analysis of Nonsyndromic Cleft Lip and Palate. **Am. J. Hum. Genet.** v. 49, p. 674-681, 1991.

HINDORFF, L. A.; SETHUPATHY, P.; JUNKINS, H. A.; RAMOS, E. M.; MEHTA, J. P.; COLLINS, F. S.; MANOLIO, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. **PNAS**. v. 106, n. 23, p. 9362–9367, 2009.

HAN, J. W.; ZHENG, H. F.; CUI, Y.; SUN, L. D.; YE, D. Q.; HU, Z.; XU, J. H.; CAI, Z. M.; HUANG, W.; ZHAO, G. P.; XIE, H. F.; FANG, H.; LU, Q. J.; XU, J. H.; LI, X. P.; PAN, Y. F.; DENG, D. Q.; ZENG, F. Q.; YE, Z. Z.; ZHANG, X. Y.; WANG, Q. W.; HAO, F.; MA, L.; ZUO, X. B.; ZHOU, F. S.; DU, W. H.; CHENG, Y. L.; YANG, J. Q.; SHEN, S. K.; LI, J.; SHENG, Y. J.; ZUO, X. X.; ZHU, W. F.; GAO, F.; ZHANG, P. L.; GUO, Q.; LI, B.; GAO, M.; XIAO, F. L.; QUAN, C.; ZHANG, C.; ZHANG, Z.; ZHU, K. J.; LI, Y.; HU, D. Y.; LU, W. S.; HUANG, J. L.; LIU, S. X.; LI, H.; REN, Y. Q.; WANG, Z. X.; YANG, C. J.; WANG, P. G.; ZHOU, W. M.; LV, Y. M.; ZHANG, A. P.; ZHANG, S. Q.; LIN, D.; LI, Y.; LOW, H. Q.; SHEN, M.; ZHAI, Z. F.; WANG, Y.; ZHANG, F. Y.; YANG, S.; LIU, J. J.; ZHANG, X. J. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. **Nat. Genet.** v. 41, n. 11, p. 1234–1237, 2009.

HUANG, Y.; WU, J.; MA, J.; BEATY, T.; SULL, J.; ZHU, L.; LU, D.; WANG, Y.; MENG, T.; SHI, B. Association between IRF6 SNPs and Oral Clefts in West China. **Journal of Dental Research**. v. 88, n. 8, p. 715-718, 2009.

HUPPI, K.; PITT, J. J.; WAHLBERG, B. M.; CAPLEN, N. J. The 8q24 gene desert: an oasis of non-coding transcriptional activity. **Frontiers in Genetics**. v. 3, n. 69, 2012.

INTERNATIONAL HAPMAP CONSORTIUM. The International HapMap Project. **Nature**. v. 426, n. 6968, p. 789-96, 2003.

JAGOMÄGI, T.; NIKOPENSIUS, T.; KRJUTSKOV, K.; TAMMEKIVI, V.; VILTROP, T.; SAAG, M.; METSPALU, A. MTHFR and MSX1 contribute to the risk of nonsyndromic cleft lip/palate. **European journal of oral sciences**. v. 118, p. 213–220, 2010.

JUGESSUR, A.; GJESSING, K.; LIE, R.; WILCOX, A.; SHI, M.; WEINBERG, C.; CHRISTENSEN, K.; BOYLES, A.; DAACK-HIRSCH, S.; NGUYEN, T.; BILLE, C.; LIDRAL, A.; MURRAY, J. GENETIC Determinants of Facial Clefting : Analysis of 357 Candidate Genes Using Two National Cleft Studies from Scandinavia. **PLoS ONE**. v. 4, n. 4, 2009.

KANG H. P.; MORGAN, A.A.; CHEN, R.; SCHADT, E. E.; BUTTE, A J. Coanalysis of GWAS with eQTLs reveals disease-tissue associations. **AMIA Summits Transl Sci Proc.** v. 2012 p. 35–41, 2012a.

KANG, H. P.; YANG, X.; CHEN, R.; ZHANG, B.; CORONA, E.; SCHADT, E. E.; BUTTE, A. J. Integration of disease-specific single nucleotide polymorphisms, expression quantitative trait loci and coexpression networks reveal novel candidate genes for type 2 diabetes. **Diabetologia.** v. 55, p. 2205–2213, 2012b.

KNIGHT, J. C. Regulatory polymorphisms underlying complex disease traits. **J Mol Med.** v. 83, p. 97–109, 2005.

KONDO, S.; SCHUTTE, B.; RICHARDSON, R.; BJORK, B.; KNIGHT, A.; WATANABE, Y.; HOWARD, E.; FERREIRA, R.; DAACK-HIRSCH, S.; SANDER, A.; MCDONALD-MCGINN, D.; ZACKAI, E.; LAMMER, E.; AYLSWORTH, A.; ARDINGER, H.; LIDRAL, A.; POBER, B.; MORENO, L.; ARCOS-BURGOS, M.; VALENCIA, C.; HOUDAYER, C.; BAHUAU, M.; MORETTI-FERREIRA, D.; RICHIERI-COSTA, A.; DIXON, M.; MURRAY, J. Mutations In *IRF6* cause Van Der Woude And Popliteal Pterygium Syndromes. **Nature Genetics.** V. 32, P. 285-289, 2002.

LEES, M.; WINTER, R.; MALCOLM, S.; SAAL, H.; CHITTY, L. Popliteal pterygium syndrome: a clinical study of three families and report of linkage to the Van der Woude syndrome locus on 1q32. **Journal of Medical Genetics.** v. 36, p. 888–892, 1999.

LESLIE, E.; MURRAY, J. Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate. **Clin Genet.** 2012 [Epub ahead of print].

LIBIOULLE, C.; LOUIS, E.; HANSOUL, S.; SANDOR, C.; FARNIR, F.; FRANCHIMONT, D.; VERMEIRE, S.; DEWIT, O.; DE VOS, M.; DIXON, A.; DEMARCHE, B.; GUT, I.; HEATH, S.; FOGLIO, M.; LIANG, L.; LAUKENS, D.; MNI, M.; ZELENKA, D.; VAN GOSSUM, A.; RUTGEERTS, P.; BELAICHE, J.; LATHROP, M.; GEORGES, M. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. **PLoS Genet.** v. 3, n 4, p. 538-543, 2007.

LITTLE, J.; CARDY, A.; MUNGER, R. G. Tobacco smoking and oral clefts: a meta-analysis. **Bull. World Health Organ.** v. 82, p. 213–218, 2002.

LOFFREDO, L. C. M.; FREITAS, J. A. S.; GRIGOLLI, A. A. G. Prevalence of oral clefts from 1975 to 1994, Brazil. **Rev Saúde Pública.** v.35, n.6, p.571-5. 2001.

LUDWIG, K. U.; MANGOLD, E.; HERMS, S.; NOWAK, S.; REUTTER, H.; PAUL, A.; BECKER, J.; HERBERZ, R.; ALCHAWA, T.; NASSER, E.; BÖHMER, A. C.; MATTHEISEN, M.; ALBLAS, M. A.; BARTH, S.; KLUCK, N.; LAUSTER, C.; BRAUMANN, B.; REICH, R. H.; HEMPRICH, A.; PÖTZSCH, S.; BLAUMEISER, B.; DARATSIANOS, N.; KREUSCH, T.; MURRAY, J. C.; MARAZITA, M. L.; RUCZINSKI, I.; SCOTT, A. F.; BEATY, T. H.; KRAMER, F.; WIENKER, T. F.; STEEGERS-THEUNISSEN, R. P.; RUBINI, M.; MOSSEY, P. A.; HOFFMANN, P.; LANGE, C.; CICHON, S.; PROPPING, P.; KNAPP, M.; NÖTHEN, M. M. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. **Nature Genetics**. v. 44, n. 9, p. 968-971, 2012.

MALLOY, M. H.; KLEINMAN, J. C.; BAKEWELL, J. M.; SCHRAMM, W. F.; LAND, G. H. Maternal smoking during pregnancy: No association with congenital malformations in Missouri 1980-83. **Am J Public Health**. v. 79, p. 1234-6, 1989.

MANGOLD, E.; LUDWIG, K. U.; BIRNBAUM, S.; BALUARDO, C.; FERRIAN, M.; HERMS, S.; REUTTER, H.; DE ASSIS, N. A.; CHAWA, T. A.; MATTHEISEN, M.; STEFFENS, M.; BARTH, S.; KLUCK, N.; PAUL, A.; BECKER, J.; LAUSTER, C.; SCHMIDT, G.; BRAUMANN, B.; SCHEER, M.; REICH, R. H.; HEMPRICH, A.; POTZSCH, S.; BLAUMEISER, B.; MOEBUS, S.; KRAWCZAK, M.; SCHREIBER, S.; MEITINGER, T.; WICHMANN, H. E.; STEEGERS-THEUNISSEN, R. P.; KRAMER, F. J.; CICHON, S.; PROPPING, P.; WIENKER, T. F.; KNAPP, M.; RUBINI, M.; MOSSEY, P. A.; HOFFMANN, P.; NOTHEN, M. M. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. **Nat Genet**. v. 42, n. 1, p. 24-6, 2009.

MANOLIO, T. A. Genomewide Association Studies and Assessment of the Risk of Disease. **The New England Journal of Medicine**. v. 363, p. 166-76, 2010.

MARIAN, A. Molecular genetic studies of complex phenotypes. **Translational Research**. v. 159, n. 2, p. 64-78, 2012.

MENEGOTTO, B. G.; SALZANO, F. M. Epidemiology of oral clefts in a large South American sample. **Cleft Palate Craniofac J**, v.28, n.4, p.373-6, 1991.

MILUNSKY, J. M.; MAHER, T. A.; ZHAO, G.; ROBERTS, A. E.; STALKER, H. J.; ZORI, R. T.; BURCH, M. N.; CLEMENS, M.; MULLIKEN, J. B.; SMITH, R.; LIN, A. E. TFAP2A mutations result in branchio-oculo-facial syndrome. **Am. J. Hum. Genet**. 82, 1171–1177 (2008).

MOFFATT, M. F.; KABESCH, M.; LIANG, L.; DIXON, A. L.; STRACHAN, D.; HEATH, S.; DEPNER, M.; VON BERG, A.; BUFE, A.; RIETSCHER, E.; HEINZMANN, A.; SIMMA, B.; FRISCHER, T.; WILLIS-OWEN, S. A. G.; WONG, K. C. C.; ILLIG, T.; VOGELBERG, C.; WEILAND, S. K.; VON MUTIUS, E.; ABECASIS, G. R.; FARRALL, M.; GUT, I. G.; LATHROP, G. M.; COOKSON, W. O. C. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. **Nature**. v. 448, p. 470-474, 2007.

MOSSEY, P.; CASTILLIA, E. **Global registry and database on craniofacial anomalies...**2003. Relatório da Reunião da Organização Mundial de Saúde em Anomalias Craniofaciais, Geneva, 2003.

MOSSEY, P.; LITTLE, J.; MUNGER, R.; DIXON, M.; SHAW, W. Cleft lip and palate. **The Lancet**. v. 374, p. 1-13, 2009.

MOSTOWSKA, A.; HOZYASZ, K.; WOJCICKI, P.; BIEDZIAK, B.; PARADOWSKA, P.; JAGODZINSKI, P. Association between Genetic Variants of Reported Candidate Genes or Regions and Risk of Cleft Lip with or without Cleft Palate in the Polish Population. **Birth Defects Research (Part A)**. v. 88, p. 538-545, 2010.

NICA, A. C.; MONTGOMERY, S. B.; DIMAS, A. S.; STRANGER, B. E.; BEAZLEY, C.; BARROSO, I.; DERMITZAKIS, E. T. Candidate Causal Regulatory Effects by Integration of Expression QTLs with Complex Trait Genetic Associations. **PLoS Genet**. v. 6, n. 4, 2010.

NICA, A. C.; PARTS, L.; GLASS, D.; NISBET, J.; BARRETT, A.; SEKOWSKA, M.; TRAVERS, M.; POTTER, S.; GRUNDBERG, E.; SMALL, K.; HEDMAN, A. K.; BATAILLE, V.; TZENOVA, B. J.; SURDULESCU, G.; DIMAS, A. S.; INGLE, C.; NESTLE, F. O.; DI MEGLIO, P.; MIN, J. L.; WILK, A.; HAMMOND, C. J.; HASSANALI, N.; YANG, T. P.; MONTGOMERY, S. B.; O'RAHILLY, S.; LINDGREN, C. M.; ZONDERVAN, K. T.; SORANZO, N.; BARROSO, I.; DURBIN, R.; AHMADI, K.; DELOUKAS, P.; MCCARTHY, M. I.; DERMITZAKIS, E. T.; SPECTOR, T. D.; MUTHER CONSORTIUM. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. **PLoS Genet**. v. 7, n. 2, 2011.

NICOLAE, D. L.; GAMAZON, E.; ZHANG, W.; DUAN, S.; DOLAN, M. E. COX, N. J. Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS. **PLoS Genet**. v. 6, n. 4, 2010.

NIKOPENSIUS, T.; AMBROZAITYTE, L.; LUDWIG, K. U.; BIRNBAUM, S.; JAGOMAGI, T.; SAAG, M.; MATULEVICIENE, A.; LINKEVICIENE, L.; HERMS, S.; KNAPP, M.; HOFFMANN, P.; NOTHEN, M. M.; KUCINSKAS, V.; METSPALU, A.; MANGOLD, E. 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**. v. 149A, p. 2551–2553, 2009.

PASSOS-BUENO, M. R.; ALONSO, N.; BRITO, L. A. Fissura labiopalatina: genética, reabilitação e aconselhamento genético. In: KIM, C. A; ALBANO, L. M. J; BERTOLA, D. R. (Org.). **Genética na Prática Pediátrica**. 1 ed. Barueri: Manole, 2010, v. 1, p. 159-173.

PAN, Y.; MA, J.; ZHANG, W.; DU, Y.; NIU, Y.; WANG, M.; ZHANG, Z.; WANG, L. IRF6 polymorphisms are associated with nonsyndromic orofacial clefts in a Chinese Han population. **Am J Med Genet.** v. 152A, p. 2505–2511, 2010.

PARK, J.; MCINTOSH, I.; HETMANSKI, J.; JABS, E.; KOLK, C.; WU-CHOU, Y.; CHEN, P.; CHONG, S.; YEOW, V.; JEE, S.; PARK, B.; FALLIN, D.; INGERSOLL, R.; SCOTT, A.; BEATY, T. Association between IRF6 and nonsyndromic cleft lip with or without cleft palate in four populations. **Genet Med.** v. 9, p. 219–227, 2007.

PEGELOW, M.; PEYRARD-JANVID, M.; ZUCHELLI, M.; FRANSSON, I.; LARSON, O.; KERE, J.; LARSSON, C.; KARSTEN, A. Familial non-syndromic cleft lip and palate-analysis of the *IRF6* gene and clinical phenotypes. **European Journal of Orthodontics.** v. 30, p. 169–175, 2008.

RAHIMOV, F.; MARAZITA, M.; VISEL, A.; COOPER, M.; HITCHLER, M.; RUBINI, M.; DOMANN, F.; GOVIL, M.; CHRISTENSEN, K.; BILLE, C.; MELBYE, M.; JUGESSUR, A.; LIE, R.; WILCOX, A.; FITZPATRICK, D.; GREEN, E.; MOSSEY, P.; LITTLE, J.; STEEGERS-THEUNISSEN, R.; PENNACCHIO, L.; SCHUTTE, B.; MURRAY, J. Disruption of an AP-2a binding site in an IRF6 enhancer is associated with cleft lip. **Nature Genetics.** v. 40, p. 1341-1347, 2008.

RICHARDS, A. L.; JONES, L.; MOSKVINA, V.; KIROV, G.; GEJMAN, P. V.; LEVINSON, D. F.; SANDERS, A. R.; MOLECULAR GENETICS OF SCHIZOPHRENIA COLLABORATION (MGS), INTERNATIONAL SCHIZOPHRENIA CONSORTIUM (ISC), PURCELL, S.; VISSCHER, P. M.; CRADDOCK, N.; OWEN, M. J.; HOLMANS, P.; O'DONOVAN, M. C. Schizophrenia susceptibility alleles are enriched for alleles that affect gene expression in adult human brain. **Molecular Psychiatry.** v. 17, p. 193-201, 2012.

ROMITTI, P. A.; LIDRAL, A. C.; MUNGER, R. G.; DAACK-HIRSCH, S.; BURNS, T. L.; MURRAY, J. C. Candidate Genes For Nonsyndromic Cleft Lip and Palate and Maternal Cigarette Smoking and Alcohol Consumption: Evaluation of Genotype-Environment Interactions from a Population-Based Case-Control Study of Orofacial Clefts. **TERATOLOGY.** v. 59, p. 39-50, 1999.

SATOKATA, I.; MAAS, R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. **Nature Genetics.** v. 6, n. 4, p. 348–356, 1994.

SCAPOLI, L.; PALMIERI, A.; MARTINELLI, M.; PEZZETTI, F.; CARINCI, P.; TOGNON, M.; CARINCI, F. 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.** v. 76, p. 180–183, 2005.

SCHADT, E. E.; MONKS, S. A.; DRAKE, T. A.; LUSISK, A. J.; CHEK, N.; COLINAYOK, V.; RUFF, T. G.; MILLIGAN, S. B.; LAMB, J. R.; CAVET, G.; LINSLEY, P. S.; MAO, M.; STOUGHTON, R. B.; FRIEND, S. H. Genetics of gene expression surveyed in maize, mouse and man. **NATURE**. v. 422, 2003.

SCHADT, E.; MOLONY, C.; CHUDIN, E.; HAO, K.; YANG, X.; LUM, P.; KASARSKIS, A.; ZHANG, B.; WANG, S.; SUVER, C.; ZHU, J.; MILLSTEIN, J.; SIEBERTS, S.; LAMB, J.; THAKURTA, D.; DERRY, J.; STOREY, J.; AVILA-CAMPILLO, I.; KRUGER, M.; JOHNSON, J.; ROHL, C.; VAN NAS, A.; MEHRABIAN, M.; DRAKE, T.; LUSIS, A.; SMITH, R.; GUENGERICH, F.; STROM, S.; SCHUETZ, E.; RUSHMORE, T.; ULRICH, R. Mapping the Genetic Architecture of Gene Expression in Human Liver. **PLoS Biology**. v. 6, p. 1020-1032, 2008.

SHAW, G. M.; IOVANNISCI, D. M.; YANG, W.; FINNELL, R. H.; CARMICHAEL, S. L.; CHENG, S.; LAMMER, E. J. Endothelial nitric oxide synthase (NOS3) genetic variants, maternal smoking, vitamin use, and risk of human orofacial clefts. **Am J Epidemiol**. v. 162, n. 12, p. 1207-14, 2005.

SHAW, G. M.; CARMICHAEL, S. L.; LAURENT, C.; RASMUSSEN, S. A. Maternal nutrient intakes and risk of orofacial clefts. **Epidemiology**. v. 17, p. 285–91, 2006.

SHI M, CHRISTENSEN K, WEINBERG CR, ROMITTI P, BATHUM L, LOZADA A, MORRIS RW, LOVETT M, MURRAY JC. **Am J Hum Genet**. Orofacial cleft risk is increased with maternal smoking and specific detoxification-gene variants. v. 80, n. 1, p. 76-90, 2007.

SOTELO, J.; ESPOSITO, D.; DUHAGON, M.A .; BANFIELD, K.; MEHALKO, J.; LIAO, H.; STEPHENS, R. M.; HARRIS, T. J. R.; MUNROE, D. J.; WU, X. Long-range enhancers on 8q24 regulate c-Myc. **PNAS**. v. 107, n. 7, p. 3001–3005, 2010.

SPINA, V.; PSILLAKIS, J. M.; LAPA, F. S.; FERREIRA, M. C. Classificação das Fissuras lábio-palatinas. Sugestão de modificação. **Rev. Hosp. Clin Fac Med S. Paulo**. v. 27, p. 5-6, 1972.

STAHL, E. A.; RAYCHAUDHURI, S.; REMMERS, E. F.; XIE, G.; EYRE, S.; THOMSON, B. P.; LI, Y.; KURREEMAN, F. A. S.; ZHERNAKOVA, A.; HINKS, A.; GUIDUCCI, C.; CHEN, R.; ALFREDSSON, L.; AMOS, C. I.; ARDLIE, K. G.; BIRAC CONSORTIUM; BARTON, A.; BOWES, J.; BROUWER, E.; BURTT, N. P.; CATANESE, J. J.; COBLYN, J.; COENEN, M. J.; COSTENBADER, K. H.; CRISWELL, L. A.; CRUSIUS, J. B.; CUI, J.; DE BAKKER, P. I.; DE JAGER, P. L.; DING, B.; EMERY, P.; FLYNN, E.; HARRISON, P.; HOCKING, L. J.; HUIZINGA, T. W.; KASTNER, D. L.; KE, X.; LEE, A. T.; LIU, X.; MARTIN, P.; MORGAN, A. W.; PADYUKOV, L.; POSTHUMUS, M. D.; RADSTAKE, T. R.; REID, D. M.; SEIELSTAD, M.; SELDIN, M. F.; SHADICK, N. A.; STEER, S.; TAK, P. P.; THOMSON, W.; VAN

DER HELM-VAN MIL, A. H.; VAN DER HORST-BRUIJNSMA, I. E.; VAN DER SCHOOT, C. E.; VAN RIEL, P. L.; WEINBLATT, M. E.; WILSON, A. G.; WOLBINK, G. J.; WORDSWORTH, B. P.; YEAR CONSORTIUM; WIJMENGA, C.; KARLSON, E. W.; TOES, R. E.; DE VRIES, N.; BEGOVICH, A. B.; WORTHINGTON, J.; SIMINOVITCH, K. A.; GREGERSEN, P. K.; KLARESKOG, L.; PLENGE, R. M. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. **Nat. Genet.** v. 42, n. 6, p. 508–514, 2010.

STRACHAN, T.; READ, A. P. Complex diseases: theory and results. In:_____. **Human Molecular Genetics**. New York: Wiley-Liss, 1999. Cap. 19, p. 451.

STRANGER, B. E.; DERMITZAKIS, E. T. The genetics of regulatory variation in the human genome. **Human Genomics**. v 2, n. 2, p. 126–131, 2005.

TAKAHASHI, K.; TANABE, K.; OHNUKI, M.; NARITA, M.; ICHISAKA, T.; TOMODA, K.; YAMANAKA, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. **Cell**. v. 131, n. 5, p. 861-72, 2007.

TOLAROVA, M.; VAN ROOIJ, I.; PASTOR, M. A common mutation in the MTHFR gene is a risk factor for nonsyndromic cleft lip and palate anomalies. **The American Journal of Human Genetics**. v. 63, p. 27, 1998.

VAN ROOIJ, I. A.; WEGERIF, M. J.; ROELOFS, H. M.; PETERS, W. H.; KUIJPERS-JAGTMAN, A. M.; ZIELHUIS, G. A.; MERKUS, H. M.; STEEGERS-THEUNISSEN, R. P. Smoking, genetic polymorphisms in biotransformation enzymes, and nonsyndromic oral clefting: a gene-environment interaction. **Epidemiology**. v. 12, n. 5, p. 502-7, 2001.

VAN ROOIJ, I. A.; OCKE, M. C.; STRAATMAN, H.; ZIELHUIS, G. A.; MERKUS, H. M.; STEEGERS-THEUNISSEN, R. P. Periconceptional folate intake by supplement and food reduces the risk of nonsyndromic cleft lip with or without cleft palate. **Prev Med**. v. 39, p. 689–94, 2004.

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, p. 216-25, 1987.

VIEIRA, A.; MCHENRY, T.; DAACK-HIRSCH, S.; MURRAY, J.; MARAZITA, M. A Genome Wide Linkage Scan for Cleft Lip and Palate and Dental Anomalies. **American Journal of Medical Genetics**. v. 146, n. 11, p. 1406–1413, 2008.

WASSERMAN, N. F.; ANEAS, I.; NOBREGA, M. A. An 8q24 gene desert variant associated with prostate cancer risk confers differential in vivo activity to a MYC enhancer. **Genome Res.** v. 20, n. 9, p. 1191-7, 2010.

WEINBERG, S. M.; BRANDON, C. A.; MCHENRY, T. H.; NEISWANGER, K.; DELEYIANNIS, F. W. B.; DE SALAMANCA, J. E.; CASTILLA, E. E.; CZEIZEL, A. E.; VIEIRA, A. R.; MARAZITA, M. L. Rethinking Isolated Cleft Palate: Evidence of Occult Lip Defects in a Subset of Cases. **American Journal of Medical Genetics Part A.** v. 146A, p. 1670–1675, 2008.

WILCOX, A. J.; LIE, R. T.; SOLVOLL, K.; TAYLOR, J.; MCCONNAUGHEY, D. R.; ABYHOLM, F.; VINDENES, H.; VOLLSET, S. E.; DREVON, C. A. Folic acid supplements and risk of facial clefts: national population based case–control study. **Br Med J.** v. 334, p. 464–70, 2007.

WONG, F.; HÄGG, U. An update on the aetiology of orofacial clefts. **Hong Kong Med J.** v. 10, n. 5, p. 331-336, 2004.

ZUCCHERO, T.; COOPER, M.; MAHER, B.; DAACK-HIRSCH, S.; NEPOMUCENO, B.; RIBEIRO, L.; CAPRAU, D.; CHRISTENSEN, K.; SUZUKI, Y.; MACHIDA, J.; NATSUME, N.; YOSHIURA, K.; VIEIRA, A.; ORIOLI, I.; CASTILLA, E.; MORENO, L.; ARCOS-BURGOS, M.; LIDRAL, A.; FIELD, L.; LIU, Y.; RAY, A.; GOLDSTEIN, T.; SCHULTZ, R.; SHI, M.; JOHNSON, M.; KONDO, S.; SCHUTTE, B.; MARAZITA, M.; MURRAY, J. Interferon regulatory factor 6 (*IRF6*) gene variants and the risk of isolated cleft lip or palate. **The New England Journal of Medicine.** v. 351, p. 769-780, 2004.