UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ODONTOLOGIA DE PIRACICABA

RENATO ASSIS MACHADO

SUSCETIBILIDADE DE VARIANTES GENÉTICAS E INTERAÇÕES GENE-GENE E GENE-FATORES AMBIENTAIS NA ETIOLOGIA DAS FISSURAS ORAIS NÃO-SINDRÔMICAS NA POPULAÇÃO BRASILEIRA

SUSCEPTIBILITY OF GENETIC VARIANTS AND GENE-GENE AND GENE-ENVIRONMENT FACTOR INTERACTIONS IN THE ETIOLOGY OF NONSYNDROMIC ORAL CLEFTS IN THE BRAZILIAN POPULATION

PIRACICABA

2018

RENATO ASSIS MACHADO

SUSCETIBILIDADE DE VARIANTES GENÉTICAS E INTERAÇÕES GENE-GENE E GENE-FATORES AMBIENTAIS NA ETIOLOGIA DAS FISSURAS ORAIS NÃO-SINDRÔMICAS NA POPULAÇÃO BRASILEIRA

SUSCEPTIBILITY OF GENETIC VARIANTS AND GENE-GENE AND GENE-ENVIRONMENT FACTOR INTERACTIONS IN THE ETIOLOGY OF NONSYNDROMIC ORAL CLEFTS IN THE BRAZILIAN POPULATION

Tese apresentada à Faculdade de Odontologia de Piracicaba da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Estomatopatologia, na Área de Patologia.

Thesis presented to the Piracicaba Dental School of the University of Campinas to partial fulfillment of the requirements for the degree of Doctor in Oral Medicine and Oral Pahtology, in Pathology area.

Orientador: Prof. Dr. Ricardo Della Coletta
Coorientador: Prof. Dr. Hercílio Martelli Junior

Este exemplar corresponde à versão final da tese defendida pelo aluno Renato Assis Machado e orientada pelo Prof. Dr. Ricardo Della Coletta.

PIRACICABA
2018

Agência(s) de fomento e nº(s) de processo(s): CAPES; FAPESP, 2016/02667-0
ORCID: https://orcid.org/0000-0002-1697-3662

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Odontologia de Piracicaba
Heloisa Maria Ceccotti - CRB 8/6403

M18s

Machado, Renato Assis, 1989-
Suscetibilidade de variantes genéticas e interações gene-gene e gene-ambiente na etiologia das fissuras orais não-sindrômicas na população brasileira / Renato Assis Machado. – Piracicaba, SP : [s.n.], 2018.

Orientador: Ricardo Della Coletta.
Coorientador: Hercílio Martelli Junior.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Odontologia de Piracicaba.

1. Fenda labial. 2. Fenda palatina. 3. Polimorfismo de nucleotídeo único. 4. Genética populacional. 5. Anormalidades congênitas. I. Della Coletta, Ricardo, 1972-. II. Martelli Junior, Hercílio. III. Universidade Estadual de Campinas. Faculdade de Odontologia de Piracicaba. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Susceptibility of genetic variants and gene-gene and gene-environment factor interactions in the etiology of nonsyndromic oral clefts in the Brazilian population
Palavras-chave em inglês:
Cleft lip
Cleft palate
Single nucleotide polymorphism
Genetics, population
Congenital abnormalities
Área de concentração: Patologia
Titulação: Doutor em Estomatopatologia
Banca examinadora:
Ricardo Della Coletta [Orientador]
Sergio Roberto Peres Line
Débora Campanella Bastos
Raquel Mantuaneli Scarel Caminaga
Erika Calvano Kuchler
Data de defesa: 16-03-2018
Programa de Pós-Graduação: Estomatopatologia

UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Odontologia de Piracicaba

A Comissão Julgadora dos trabalhos de Defesa de Tese de Doutorado, em sessão pública realizada em 16 de Março de 2018, considerou o candidato RENATO ASSIS MACHADO aprovado.

PROF. DR. RICARDO DELLA COLETTA

PROFª. DRª. RAQUEL MANTUANELI SCAREL CAMINAGA

PROFª. DRª. ERIKA CALVANO KUCHLER

PROFª. DRª. DÉBORA CAMPANELLA BASTOS

PROF. DR. SERGIO ROBERTO PERES LINE

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

DEDICATÓRIA

Aos meus pais, Carlos e Lucimar, que com muito esforço conseguiram formar seus dois filhos. Vocês são meus exemplos de respeito, humildade, dignidade, honestidade, dedicação e perseverança. Graças ao apoio irrestrito, paciência e compreensão pela minha ausência em muitos momentos, finalizo mais esta etapa de minha vida.

Aos pacientes com fissura oral não-sindrômica e suas famílias, razão de nosso estudo, que gentilmente colaboraram, confiaram e se prontificaram em participar desta pesquisa. Pacientes que nos ensinaram o valor do nosso trabalho e depositaram em nós suas esperanças em busca de uma melhor qualidade de vida. Que este estudo, em sua pequena contribuição e buscando reduzir o universo de desconhecimento sobre o assunto, possa ajudar-nos a oferecer-lhes uma melhor qualidade em suas jornadas diárias.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me dar força para superar as dificuldades e mostrar os caminhos certos nas horas difíceis.

À Faculdade de Odontologia de Piracicaba da Universidade Estadual de Campinas, na pessoa de seu diretor, Professor Doutor Guilherme Elias Pessanha Henriques.

Ao Professor Doutor Márcio Ajudarte Lopes, coordenador do Programa de Pós-Graduação em Estomatopatologia da Faculdade de Odontologia de Piracicaba da Universidade Estadual de Campinas.

Ao meu orientador Professor Doutor Ricardo Della Coletta, pela dedicação irrestrita e apoio direto pela missão que se cumpre. As indicações, as dicas, as correções e até mesmo alguns entreveros que passaram em nossa relação. Tudo isso compôs um somatório fundamental não só para construção do pensamento que se traduz nas páginas desse longo texto, mas para a maturidade de toda uma vida a seguir. Registro aqui meu profundo agradecimento.

Ao meu coorientador Professor Doutor Hercílio Martelli Júnior, pelo incentivo aos trabalhos científicos desde a graduação e por compartilhar seus conhecimentos com tranquilidade e sabedoria.

Aos Professores Doutores do Departamento de Diagnóstico Oral, áreas de Patologia e Semiologia da Faculdade de Odontologia de Piracicaba, Alan Roger dos Santos Silva, Edgard Graner, Jacks Jorge Júnior, Márcio Ajudarte Lopes, Oslei Paes de Almeida, Pablo Vargas e Ricardo Della Coletta. Obrigado pelos ensinamentos repassados durante a realização dos créditos.

Ao coordenador do comitê de ética da FOP Jacks Jorge Júnior pelo auxílio prestado durante o processo de submissão do trabalho, pela sua disposição, solidariedade e auxílio mostrando em tudo dedicação e responsabilidade.

Aos funcionários da FOP-UNICAMP Emílio Salles, Adriano Luis Martins e Fabiana Facco Cassarotti pela amabilidade e paciência, e ao Fábio Haach Téo pelo apoio e valiosa ajuda no laboratório de biologia molecular.

Aos centros de colaboração e parceria neste estudo: Centro Pró-Sorriso da Universidade José do Rosário Vellano – Unifenas, Alfenas, Minas Gerais, na pessoa do Prof. Dr. Mário Sérgio Oliveira Swerts; Centrinho do Hospital Santo Antônio das Obras Assistenciais Irmã Dulce, Salvador, Bahia, na pessoa da Profª. Drª. Sílvia Reis; Associação Portadores de Fissura Labial localizada na cidade de Cascavel, Paraná (APOFILAB), na pessoa da Profª. Drª. Ana Lúcia Carrinho Ayroza Rangel; o Hospital Universitário Lauro Wanderley da Universidade Federal da Paraíba-UFPB, João Pessoa, Paraíba, na pessoa da Profª. Drª. Darlene Camati Persuhn; e o Hospital Ophir Loyola, Belém, Pará, na pessoa do Dr. David Normando, que foram de fundamental importância para a concretização deste trabalho.

Ao professor Dr. Sérgio Line pelos ensinamentos passados nas aulas de introdução a programação do software R.

À professora Drª. Tao Wu e suas alunas Yolanda, Yuan e Rita pela recepção na Universidade de Pequim e por todo conhecimento adquirido na programação do package TRIO no programa RStudio.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo auxílio financeiro.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudo durante os três anos de meu doutorado.

À todos os amigos e colegas do programa de Pós-Graduação em Estomatopatologia: Amanda Leite, Anna Luiza Araujo, Bruno Augusto Mariz, Carla Fernandes, Carolina Carneiro, Celeste Romero, Cinthia Veronica Cáceres, Ciro Soares, Débora Pereira, Diego Fernandes, Felipe Martins, Gleyson do Amaral, Isabel Schausltz, Jamile de Sá, Jessica Montenegro, João Scarini, Juliana Nascimento, Lady Paola Arboleda, Leonardo dos Reis, Ligia Miyahara, Luan da Silva, Maria Cecilia de Oliveira, Mariana Paglioni, Marisol Galvis, Natália Palmier, Patrícia Fernandes, Pedro Curioso, Priscilla Diniz, Rachel Lamarck, Raísa de Sá, Renata Markman, Rodrigo de Andrade, Thayná Morais, Vinicius Torregrossa, Wagner da Silva e especialmente aos amigos do Laboratório de Biologia Celular e Molecular Débora Bastos, Florence Cuadra e Iara de Aquino pelos conselhos e ensinamentos transmitidos.

Aos amigos do Orocentro Camila Borges, Karina Morais, Renata Markman, Vinícius Rabelo, Wagner Gomes, Aparecida Campion, Daniele Morelli e Rogério Elias, pelos ensinamentos compartilhados.

Aos amigos da sociedade do café Eduardo Castro, Paulo Campos, Gabriell Borgato e Gabriel Abuna pelos momentos de descontração e conselhos de vida.

Aos amigos do Departamento de Dentística da FOP Bruna Resende, Janaina dos Santos, Josué Pierote, Marco Túlio, Mariana Dias Flor, Matheus Kuri, Priscila Regis, Rodrigo Lins e Suelem Chasse por todo o companheirismo durante todos estes anos de vida em Piracicaba.

Aos amigos que a FOP me deu Andreia Bufalino, Fernanda Moreira, Florence Cuadra, Lívia Paranaíba, Luciana Yamamoto, Rodrigo Neves e Sibele de Aquino. Serão levados por toda vida.

Aos meus pais Carlos Roberto Machado e Lucimar Aparecida de Assis Machado por me apoiarem sempre em todas as minhas escolhas. Obrigado pelo amor, dedicação

e exemplo de vida que sempre foram para mim. Não tenho palavras para agradecer! Ao meu querido irmão Roberto Assis Machado pelo companheirismo de sempre e exemplo de pessoa a seguir. Obrigado a toda família, cada um de vocês é especial para mim.

À todos que colaboraram direta ou indiretamente durante a realização deste trabalho.

“Ação é a chave fundamental para todo o sucesso.”

Pablo Picasso

RESUMO

A etiologia das fissuras orais não-sindrômicas (FONS) é complexa e fortemente influenciada pelos fatores genéticos e ambientais específicos de cada população. No Brasil, a elevada miscigenação é considerada uma característica que pode influenciar na suscetibilidade das FONS. Para melhor compreender os aspectos genéticos associados as FONS na população brasileira, este estudo compilou resultados de 4 estudos específicos. O primeiro estudo foi uma revisão sistemática e meta-análise de marcadores genéticos avaliados na população brasileira. Este estudo revelou possíveis associações dos polimorfismos de nucleotídeo único (SNP) rs642961 (*IRF6*), rs987525 e rs1530300 (8q24), rs1801133 (*MTHFR*) e rs17563 (*BMP4*) com a etiologia das fissuras labiais com ou sem fissuras palatinas não-sindrômica (FL±PNS). Contudo, frente ao pequeno número de estudos que analisou cada um destes marcadores, mais estudos com amostras robustas e que levem em consideração a elevada miscigenação da população brasileira são necessários. O segundo estudo optou por validar 7 SNPs (rs7552 em 2q24.2, rs8049367 em 16p13.3, rs1880646, rs7406226 e rs9891446 em 17p13, rs1588366 em 17q23.2 e rs73039426 em 19q13.11), localizados em regiões associadas com FL±PNS em estudos de larga escala genômica, em 831 pacientes com FL±PNS e 866 controles. A análise de regressão logística levando em consideração as diferenças na ancestralidade genômica e no gênero entre os grupos revelou que o SNP rs7552 é um marcador de risco para o desenvolvimento das FL±PNS. Interações gene-gene (GxG) entre rs7552 com rs8049367, rs1880646, rs9891446, rs1588366 e rs73039426 foram também associadas com risco aumentado para o desenvolvimento das FL±PNS. Embora os SNPs rs1880646 e rs9891446 não foram individualmente associados com FL±PNS, o haplótipo AG (alelo A de rs1880646 e alelo G de rs9891446) foi mais frequente entre os pacientes com FL±PNS em comparação com os controles, exibindo um risco aumentado para o desenvolvimento da fissura. O terceiro estudo avaliou a influência de SNP em genes associados com a neutralização do estresse oxidativo (famílias de genes superóxido dismutase-SOD e paraoxinase-PON) no risco das FL±PNS na população brasileira, considerando interações GxG e gene-fatores ambientais (GxE). Os resultados demonstraram que o alelo C e o genótipo CT do SNP rs2237583 em *PON1* evocam efeitos protetores para as FL±PNS, enquanto rs3917490 apresentou significância apenas

com amostra composta por pacientes com alta ascendência africana. Várias interações GxG contendo os SNP rs2237583 em *PON1* e rs17166879 em *PON2* atingiram significância após o ajuste para múltiplos testes. Por fim, o quarto estudo utilizou a estratégia de tag-SNP para verificar a participação de variantes (rs1169, rs7153, rs9968051, rs9819530 e rs6794341) em *GOLGB1* na patogênese das fissuras palatinas isoladas não-sindrômicas (FPNS). Neste estudo contendo 270 pacientes com FPNS e 284 controles, nenhuma associação significante foi observada entre as variantes em *GOLGB1* e as FPNS. Em conclusão, este estudo revela alguns potenciais marcadores genéticos associados ao desenvolvimento das FONS na população brasileira e reforça a importância de considerar interações GxG na patogênese desta malformação congênita.

Palavras-Chave: Fissura de lábio. Fissura de palato. Polimorfismo de nucleotídeo único. Genética populacional. Anormalidades congênitas.

ABSTRACT

The nonsyndromic oral cleft (NOC) etiology is complex and strongly influenced by specific genetic and environmental factors within each population. In Brazil, the high miscegenation is considered a characteristic that may influence NOC susceptibility. To improve our understanding in the genetic aspects associated with NOC in the Brazilian population, this study compiled results from 4 specific studies. The first study was a systematic review and meta-analysis of genetic markers evaluated in the Brazilian population. This study revealed possible associations of single nucleotide polymorphisms (SNP) rs642961 (*IRF6*), rs987525 and rs1530300 (8q24), rs1801133 (*MTHFR*) and rs17563 (*BMP4*) with the etiology of nonsyndromic cleft lip with or without cleft palate (NSCL±P). However, in the view of the small number of studies that analyzed each of these markers, more studies with robust samples taking into account the high miscegenation of the Brazilian population are necessary. The second study opted to validate 7 SNP (rs7552 in 2q24.2, rs8049367 in 16p13.3, rs1880646, rs7406226 and rs9891446 in 17p13, rs1588366 in 17q23.2 and rs73039426 in 19q13.11), located in regions associated with NSCL±P in large-scale genomic studies, in 831 patients with NSCL±P and 866 controls. Logistic regression analysis adjusted to differences in the genomic ancestry and gender of the groups revealed that the SNP rs7552 is a marker of risk for the development of NSCL±P. Gene-gene (GxG) interactions between rs7552 with rs8049367, rs1880646, rs9891446, rs1588366 and rs73039426 were also associated with increased risk for the development of NSCL±P. Although rs1880646 and rs9891446 were not individually associated with NSCL±P, the A-G haplotype (A allele of rs1880646 and G allele of rs9891446) was more frequent among NSCL±P patients as compared to controls, exhibiting an increased risk to NSCL±P. The third study evaluated the influence of SNP in genes associated with neutralization of stress oxidative (superoxide dismutase-*SOD* and paraoxonase-*PON* gene families) in the risk of NSCL±P in the Brazilian population, considering GxG and gene-environment factor (GxE) interactions. The results showed that the C allele and the CC genotype of *PON1* rs2237583 evoke significant protective effects against NSCL±P, while rs3917490 showed a significant association only in the sample composed of patients displaying high African ancestry. Some GxG interactions containing *PON1* rs2237583 and *PON2* rs17166879 reached significance after

adjustment for multiple tests. Finally, the fourth study used the tag-SNP strategy to verify the participation of variants (rs1169, rs7153, rs9968051, rs9819530 and rs6794341) in *GOLGB1* in the pathogenesis of nonsyndromic cleft palate only (NSCPO). In this study containing 270 patients with NSCPO and 284 controls, no significant associations were observed between variants in *GOLGB1* and NSCPO. In conclusion, this study reveals some potential genetic markers associated with the development of NOC in the Brazilian population, and reinforces the importance of considering GxG interactions in the pathogenesis of this congenital malformation.

Key words: Cleft lip. Cleft palate. Single nucleotide polymorphism. Genetics, population. Congenital abnormalities.

SUMÁRIO

1 INTRODUÇÃO

As fissuras orais (FO), principalmente as caracterizadas por áreas de descontinuidade no lábio e/ou palato, representam o defeito craniofacial congênito mais frequente em humanos (Dixon *et al*., 2011; Marazita, 2012). Aproximadamente 70% das FO se manifestam de forma não-sindrômica (FONS), ou seja, sem malformações ou alterações adicionais (Jones, 1988; Marazita *et al*., 2002a). Os demais casos representam como uma característica fenotípica no espectro clínico de mais de 500 síndromes relatadas (Dixon *et al*., 2011). A classificação das FO depende da região anatômica envolvida e se divide basicamente em 3 grupos: fissura pré-forame incisivo ou fissura labial (FL), fissura pós-forame incisivo ou fissura palatina (FP) e fissuras transforme incisivo ou fissura lábio-palatina (FLP). A FL é o resultado da ausência de fusão das proeminências nasais e maxilares, a FP ocorre quando os processos palatinos deixam de se fundir. Quando surge a falha concomitante nos dois processos ocorre a FLP (Shkoukani *et al*., 2013). Contudo, há uma tendência mundial em considerar a FL uma variação da FLP, colocando-as no grupo das fissuras labial com ou sem fissura palatina (FL±P).

As FONS afetam aproximadamente 1 em cada 500 a 2.500 nascidos vivos, com grande variabilidade de acordo com a origem étnica, gênero do indivíduo e fatores ambientais (Marazita *et al*., 2002b; Dixon *et al*., 2011). Em geral, as populações asiáticas e ameríndias possuem uma alta prevalência (1:500), as populações européias possuem prevalência intermediária (1:1.000) e as menores taxas de prevalência são observadas em africanos e descendentes de africanos (1:2.500) (Mossey e Little, 2002; Murthy e Bhaskar, 2009; Mossey *et al*., 2009). No Brasil, por ter uma população altamente miscigenada, baseada principalmente entre ancestralidade europeia, afriana e ameríndia, a prevalência varia entre 1:685 a 1:2.800 nascidos vivos (Martelli-Junior *et al*., 2007; Rodrigues *et al*., 2009).

Além de ser uma causa importante de mortalidade, chegando até 30% em países subdesenvolvidos e em áreas pontuais em países desenvolvidos, as FONS são associadas também a uma significante morbidade (Wehby *et al*., 2011). Efeitos sobre a fala, audição e estética geram resultados adversos sobre a saúde e integração social (Nopoulos *et al*., 2007; Mossey *et al*., 2009). Entre as primeiras complicações está à dificuldade no aleitamento materno, resultando em dificuldades no ganho de peso e no desenvolvimento da criança (Montagnoli *et al*., 2005). Pesquisas com

pacientes fissurados em fase escolar registram alterações cognitivas, que em sua maioria foram associadas a micro-alterações na trajetória do crescimento e desenvolvimento das estruturas cerebrais (Broder *et al*., 1998; Weinberg *et al*., 2013). Embora a literatura tenha apontado um maior risco na incidência de alguns tipos de câncer nos indivíduos afetados por FONS e seus familiares (Zhu *et al*., 2002; Bille *et al*., 2005; Taioli *et al*., 2010; Jindal e Vieira, 2012; Vieira *et al*., 2012), alguns estudos realizados pelo nosso grupo não confirmaram tal achado (Gonçalves *et al*., 2014; Martelli *et al*., 2014; Cardoso *et al*., 2018). No que diz respeito às alterações no desenvolvimento da dentição, as anomalias dentais também têm sido cada vez mais investigadas e detectadas com maior frequência em indivíduos com FONS (Paranaiba *et al*., 2013; Melo-Filho *et al*., 2015; Sá *et al*., 2016a; Sá *et al*., 2016b; Fernandez *et al*., 2018). Além da saúde física, já se tem apontado o alto impacto das FONS na saúde mental dos afetados, uma vez que estes podem apresentar um maior risco em desenvolver distúrbios psiquiátricos que interferem na vida social (Trindade *et al*., 2007; Lima *et al*., 2015). Todas estas complicações, vastamente estudadas, reforçam o impacto negativo que as FONS podem acarretar ao indivíduo, trazendo inúmeras consequências diretas e indiretas na sua qualidade de vida, tanto nos aspectos individuais quanto familiares e sociais (Wehby e Cassell, 2010). Por consequência da complexidade das manifestações clínicas, os indivíduos afetados por FONS necessitam de acompanhamento especializado e assistência integral a longo prazo em centros de referência de alta complexidade (Wehby e Cassell, 2010). Embora a reabilitação seja possível com o atendimento de boa qualidade, as FONS inevitavelmente constituem um ônus para o indivíduo, para a família e para a sociedade, com um custo substancial em termos de saúde e serviços relacionados (Dixon *et al*., 2011).

O desenvolvimento normal da face é marcado por uma sequência de eventos complexos, coordenados por interações entre fatores de transcrição e sinalizadores moleculares, juntamente com interações célula-célula e aquisição de polarização celular (Stanier e Moore, 2004). Até o final da quarta semana, as células da crista neural, oriundas do tubo neural anterior, migram para formar o primórdio da face, surgindo dela os processos nasal medial e nasal lateral, que se fundem ao processo maxilar, para formar a parte central do lábio superior, o palato primário e o nariz (Marazita e Mooney, 2004). O palato primário aloja os dentes incisivos da maxila, dando origem à parte anterior do forame incisivo e também contribuindo para a

formação do lábio e a porção anterior da maxila (Rice, 2005). O palato secundário se desenvolve após o palato primário durante a sexta e a décima segunda semanas. Os processos palatinos se elevam acima da língua, fundindo medialmente na linha média, anteriormente com o palato primário e superiormente com o septo nasal. O forame incisivo marca a extensão anterior do palato secundário. A formação dos palatos primário e secundário completa a separação das cavidades nasal e oral, permitindo a respiração simultaneamente à mastigação (Dixon *et al*., 2011).

A diversidade de eventos embriológicos que contribuem para a formação das estruturas faciais reflete nos inúmeros fatores envolvidos na formação das FONS (Jugessur *et al*., 2009; Rojas-Martinez *et al*., 2010). Diante da complexidade destes processos, pode-se perceber o significado biológico dos mecanismos de desenvolvimento embrionário e a sua importância, pelo fato da ocorrência de alguma falha neste processo poder contribuir para possíveis alterações congênitas. Nos últimos anos têm ocorrido uma evolução no entendimento dos fatores causais, com a identificação de novas variantes genéticas, de fatores de risco ambientais, a interação de fatores de risco ambientais e fatores genéticos, bem como a associação entre os genes (Dixon *et al*., 2011; Mangold *et al*., 2011; Wu *et al*., 2014; Machado *et al*., 2016; Wang *et al*., 2018). Há uma enorme variedade de agentes teratogênicos externos que podem influenciar o desenvolvimento do lábio e do palato, embora poucos deles estejam comprovados. Tem sido sugerido que a exposição materna a alguns fatores teratogênicos como medicamentos, álcool e tabaco e as deficiências vitamínicas, principalmente ácido fólico, durante o primeiro trimestre de gestação está intimamente associada com a ocorrência das FONS (Jianyan *et al*., 2010; Wehby e Murray, 2010). As complexidades dos mecanismos envolvidos se diferem quanto ao tipo, frequência e gravidade, resultando na diversidade das manifestações clínicas das fissuras (Economou *et al*., 2012).

Após Fogh-Andersen (1942) observar em um estudo populacional a existência de um componente hereditário associado ao desenvolvimento das FONS, uma variedade de estudos tem sido utilizada para identificar vias e genes envolvidos na etiologia das FONS. Parte dos genes candidatos foi sugerida por meio de estudos com modelos experimentais em camundongos *knockouts* (Juriloff e Harris, 2008), citogenética (Brewer *et al*., 1999; Higgins *et al*., 2008), estudos de fissuras associadas a síndromes mendelianas (Kondo *et al*., 2002; Zucchero *et al*., 2004) e análises da expressão gênica em tecidos embrionários (Mukhopadhyay *et al*., 2004; Gong et al.,

2005). As abordagens mais recentes são baseadas em estudos de associação de larga escala genômica (GWAS), no qual polimorfismos distribuídos pelo genoma são analisados simultaneamente em pacientes afetados ou não pelas FONS. Seis GWAS (Birnbaum *et al*., 2009; Grant *et al*., 2009; Beaty *et al*., 2010; Mangold *et al*., 2010; Sun *et al*., 2015; Leslie *et al*., 2016a) e 1 meta-análise de 2 destes GWAS (Ludwig *et al*., 2012) foram realizados com amostras de FONS, identificando 15 regiões genéticas de risco. No ano passado, um novo GWAS com uma ampla população chinesa e um grupo de validação com amostras de pacientes de populações europeias foi publicado, revelando 14 loci novos e confirmando outros 12 (Yu *et al*., 2017). Em adição, 3 loci foram identificados em estudos de validação dos resultados dos GWAS (Beaty *et al*., 2013; Ludwig *et al*., 2016) e em associação com estudos de ligação e análise de genes candidatos, os quais confirmaram a participação de *IRF6* (interferon regulatory factor 6) (Zucchero *et al*., 2004; Rahimov *et al*., 2008) e o lócus 9q22 o qual contém o gene *FOXE1* (forkhead box E1) (Marazita *et al*., 2009; Moreno *et al*., 2009). Então, 34 loci são atualmente associados com FONS, no entanto, diferenças de ancestralidade determinam a predisposição genética, onde regiões que são fortemente associadas com uma população podem não ser com outras. Exemplo disso, nota-se nos dois últimos GWAS que identificaram a região 17q23 em indivíduos com ascendência europeia (Leslie *et al*., 2016a) ou com as novas regiões identificadas na população chinesa 14q32.13, 17q21.32 e 9q22.32 (Yu *et al*., 2017).

Como forma de delineamento experimental, as abordagens de mapeamento genético mais utilizadas têm sido os estudos de ligação e os estudos de associação. A análise de ligação genética é baseada pelo fato que duas regiões genéticas no mesmo cromossomo e muito próxima uma da outra, tendem a ser herdadas conjuntamente (ligadas). Nesse tipo de investigação, utilizam-se marcadores moleculares de determinadas regiões cromossômicas. Tais marcadores sinalizariam que na vizinhança dessas regiões existiriam genes relacionados ao transtorno, já que os *loci* dos marcadores utilizados e dos genes estariam ligados. Para essa modalidade de investigação, em geral, necessita-se de famílias grandes e com múltiplos afetados. Diferentemente, no estudo por associação levanta a hipótese de que um determinado gene pode estar envolvido na etiopatogenia do transtorno e, então, verifica se a frequência de uma determinada variante genética é significativamente diferente entre uma população de afetados e outra de não-afetados. Busca-se, portanto, determinar se há associação entre a condição de afetado e o polimorfismo investigado (Feitosa e

Krieger, 2002). Dois tipos de desenhos são frequentemente utilizados nos estudos de associação e são classificados como estudos caso-controle e de núcleo familiar. Pode-se definir o estudo caso-controle como uma forma de pesquisa na qual são comparados dois grupos, onde o primeiro grupo de indivíduos apresenta o fenótipo (caso) e o segundo grupo não apresenta tal condição (controle). Este tipo de estudo é muito utilizado por permitir um recrutamento de um grande número de indivíduos afetados, sem a necessidade de incluir seus familiares (Risch, 2000). Por outro lado, os estudos caso-controle apresentam a desvantagem de se basearem na diferença entre a frequência das variações genéticas entre os grupos, que pode ser influenciada pela heterogeneidade populacional. Visto que a população brasileira resulta principalmente da mistura de 3 populações ancestrais (Europeia, Africana e Ameríndia) e exibe níveis muito altos de diversidade genômica (Pena *et al*., 2011), para evitar resultados tendenciosos e incorretos, é recomendado que estudos com abordagem caso-controle envolvendo a população brasileira, e outras populações miscigenadas encontradas ao redor do mundo, levem em consideração a variação ancestral genética de cada indivíduo dos grupos.

No estudo do núcleo familiar (pai, mãe e filho afetado), a ideia é coletar uma amostra de indivíduos afetados juntamente com seus pais (não afetados), ou seja; a base da análise é a segregação alélica em um mesmo núcleo ancestral. O genótipo do filho afetado é considerado como o ponto amostral do grupo “caso” e os dois alelos (um materno e um paterno), que não foram transmitidos para o filho afetado, são considerados como ponto amostral do grupo “controle”. Desta maneira, têm-se as amostras de uma mesma população genética (Purcell *et al*., 2005). A principal vantagem do estudo de núcleos familiares é o controle de possíveis efeitos da estratificação populacional (Sebro e Rogus, 2010). Para a análise utiliza-se o teste de desequilíbrio de transmissão (TDT), avaliando-se a transmissão dos alelos do polimorfismo. O TDT usa um teste x2 de Mc Nermar (1947) para verificar a hipótese nula em que o alelo tido como associado à característica é transmitido em 50% das vezes pelo progenitor heterozigoto (Feitosa e Krieger, 2002).

Diferentemente das fissuras orais sindrômicas, que apresentam um padrão de transmissão estritamente mendeliano ou ligado ao X, a análise de segregação das fissuras orais não-sindrômicas parte de um modelo multifatorial e complexo, com participação de fatores genéticos e ambientais múltiplos. Em relação ao componente genético é importante considerar que a distribuição populacional dos genótipos siga

os preceitos do teorema do equilíbrio de Hardy-Weinberg que afirma que, em uma população mendeliana, dentro de determinadas condições, as frequências alélicas permanecerão constantes ao passar das gerações. A posição da média do genótipo heterozigoto “Aa” em relação às médias dos homozigotos é designada por um grau de dominância. No caso de um dos fenótipos em estudo ser completamente dominante, “AA” será igual a “Aa”, no entanto, no modelo recessivo, apenas o homozigoto “aa” promove a variação da característica (Pierce, 2013). No modelo aditivo cada par de gene possui efeito próprio e independente dos outros que se encontram presentes no genótipo do indivíduo. Assim, a ação total do genótipo sobre o fenótipo será igual a soma dos efeitos de cada par de gene e a simples substituição de um alelo por outro em um gene impacta o resultado total dos efeitos gênicos. Sob essa hipótese, a relação entre o genótipo e o fenótipo é linear (Clarke *et al*., 2011). Em geral, os genes exibem distribuição independente, mas não atuam independentemente em sua expressão fenotípica, ou seja, os efeitos de um gene dependem da presença de outro (Pierce, 2013). Esse tipo de interação entre genes em regiões distintas sobre uma caracterítica é denominado interação gênica ou epistasia (Moore e Williams, 2015). No que tange à complexidade na relação genótipo-fenótipo das FONS, uma segunda fonte de variação surge quando o efeito de um gene depende do ambiente é encontrado. Moore e Williams (2009) argumentam que o modelo de regressão linear simples ignora o contexto genômico (interação gene-gene) e ambiental (interação gene-ambiente) de cada SNP e a utilização do modelo de regressão multifatorial se faz necessário para estudo com a abordagem de co-variantes influênciando no fenótipo.

A correlação ou associação não aleatória entre alelos em dois ou mais polimorfismos de nucleotídeo único (SNP, do inglês single-nucleotide polymorpohism) é referida como desequilíbrio de ligação (LD, do inglês linkage desequilibrium) e a utilização dessa abordagem para tratar a redundância de informação dada pelo alto LD entre pares de SNPs se faz com a construção de blocos de SNPs. Uma possível forma de montar tais blocos é separar os grupos de marcadores identificando regiões *hotspots*, as quais tem grande chance de sofrer recombinação. Para computar o LD desse bloco basta calcular a média do LD entre todos os pares que compõem o mesmo. Feito isso, escolhe-se um marcador representante, denominado tag-SNP, para esse bloco e utiliza-o juntamente com informações fenotípicas na tentativa de encontrar o subconjunto de tag-SNPs que estão associados ao fenótipo. Com esse

procedimento grande parte da redundância é eliminada. O entendimento e o uso do LD entre marcadores podem melhorar o desempenho de métodos ou técnicas usadas em estudos genéticos e fornecer melhor compreensão sobre a estrutura de associação não-aleatória entre marcadores encontrados ao final do processo de seleção. Como consequência, um método de seleção de SNPs deve selecionar os marcadores mais informativos com baixo ou nenhum LD, pois, desta forma, a redudância será reduzida ou até mesmo eliminada. De acordo com Foulkes (2009), os blocos LD variam substancialmente entre diferentes grupos étnicos. Como resultado, um conjunto de tag-SNPs podem capturar informações de variantes causais de uma doença em um grupo, mas não em outro. Logo, considerar esse fenômeno e a aplicação de abordagens apropriadas é crucial para estudos genéticos baseados em populações miscigenadas.

Este estudo teve como objetivo compreender a participação de fatores genéticos na etiologia das FONS na população brasileira. Para tanto, compilou 4 estudos específicos: 1) realizou uma revisão sistemática e meta-análise da literatura incluindo estudos que avaliaram marcadores polimórficos de risco para as FONS na população brasileira; 2) determinou o papel de variantes polimórficas recentemente descritas na literatura como de susceptibilidade para as fissuras orais não-sindrômicas, incluindo os SNP rs7552 em 2q24.2, rs8049367 em 16p13.3, rs1880646, rs7406226 e rs9891446 em 17p13, rs1588366 em 17q23.2 e rs73039426 em 19q13.11, na população brasileira com FL±PNS; 3) avaliou 28 SNP em genes relacionados ao estresse oxidativo (*SOD1*, *SOD2*, *SOD3, PON1*, *PON2* e *PON3*), bem como as interações genes-fatores ambientais maternos, no risco das FL±PNS combinando uma abordagem de núcleo familiar complementada por análise caso-controle; e 4) analisou a relação dos tag-SNP rs1169, rs7153, rs9968051, rs9819530 e rs6794341 no gene *GOLGB1* com o risco de desenvolvimento de FPNS. Nos estudos 2, 3 e 4, a análise caso-controle foi estruturada com as proporções de ancestralidade europeia, africana e ameríndia de cada individuo.

2 ARTIGOS

2.1 Artigo

Potential genetic markers for nonsyndromic oral clefts in the Brazilian population: a systematic review and meta-analysis

Artigo aceito para publicação no periódico *Birth Defects Research, 2018, doi: 10.1002/bdr2.1208* (Anexo 1)

Renato Assis Machado[1], Isabela Porto de Toledo[2], Hercilio Martelli-Júnior[3], Silvia Regina Reis[4], Eliete Neves Silva Guerra[2], Ricardo D. Coletta[1]

[1]Department of Oral Diagnosis, School of Dentistry, University of Campinas, Piracicaba-SP, Brazil
[2]Laboratory of Oral Histopathology, Health Sciences Faculty, University of Brasília, Brasília-DF, Brazil
[3]University of Montes Claros, Montes Claros-MG, Brazil
[4]Bahiana School of Medicine and Public Health, Salvador-BA, Brazil.

Correspondence to: Ricardo D. Coletta, Department of Oral Diagnosis, School of Dentistry, University of Campinas, Av. Limeira 901, CEP 13414-018, Piracicaba-SP, Brazil. E-mail: coletta@fop.unicamp.br

Running title: Genetic markers for Brazilian NOC.

Manuscript information: 5589 words in the text, 4 figures, 1 table, 1 supplementary figure, 4 supplementary tables and 73 references.

Abstract

Background: Although various genes and genomic regions were described as of susceptibility for nonsyndromic oral clefts (NOC), recent reports have demonstrated significant inter-ethnic variations in the genetic predisposition, a situation that affect the Brazilian population, one of the most admixed populations in world. Therefore, the purpose of this review was to describe the available information on genetic risk markers for NOC in the Brazilian population. **Methods:** A systematic search of the literature was performed using LILACS, LIVIVO, PubMed, Scopus and Web of Science databases, and studies that investigated genetic susceptibility markers for NOC in the Brazilian population were retrieved. Markers with enough statistical data were subjected to meta-analysis using random- or fixed-effects model with odds ratio (OR) and 95% confidence intervals (95% CI) as effect measures. **Results:** Forty-nine studies conducted since 1999 were found, and in these 114 markers were evaluated throughout case-control or family-based approaches. Most of the studies were conducted with patients affected by nonsyndromic cleft lip with or without cleft palate (NSCL±P), and 79 markers (69.3%) were evaluated by a single study only. Meta-analysis was performed with 9 markers, and the most promising results were obtained for *IRF6* (rs642961), 8q24 (rs987525 and rs1530300) and *MTHFR* (rs1801133), which were associated with increased risk for NSCL±P, and for *BMP4* (rs17563) that showed a protective effect for NSCL±P. **Conclusion:** A large number of genetic markers distributed in several genes/loci was associated with NOC in the Brazilian population, but in general the original studies included limited number of samples and unsatisfactory protocols. The classical risk markers located in *IRF6* and 8q24 showed promising results as well as rs1801133 in *MTHFR* and rs17563 in *BMP4*, and they should be validated in larger and multicenter studies taking in consideration the variations in the miscegenation of Brazilian population.

Key words: nonsyndromic oral cleft; susceptibility; genetic marker; Brazil; systematic review; meta-analysis.

Introduction

Nonsyndromic oral cleft (NOC), a developmental defect *characterized* by *lack* of complete *fusion* of the craniofacial embryonic processes, is considered the most common birth defect worldwide with a prevalence of 1.43:1000 live births (Dixon et al., 2011). The incidence varies among the different areas of world, with high incidences in Asia and Latin America countries, moderate in Europeans and low in African countries (Mossey and Modell, 2012). Among Brazilian newborns, the prevalence has been reported between 0.36 to 1.54:1000 live births, with approximately 4,000 new cases every year (Martelli-Junior et al., 2007; Rodrigues et al., 2009). As in many other developing/undeveloped countries, NOC is considered an important problem of public health to be addressed in Brazil. Taking the incisive foramen as reference, NOC are traditionally divided in cleft lip only (NSCLO), cleft lip and palate (NSCLP) and cleft palate only (NSCPO), however, as there are similarities in both epidemiologic features and embryologic timing for both CLO and CLP, they are considered variants of the same defect and grouped together to form the group cleft lip with or without cleft palate (NSCL±P). Very few studies have been focused on genetic risk factors for NSCPO, limiting the knowledge about this type of oral cleft.

It is widely accepted that NOC is a multifactorial disease dependent of a complex interplay between environmental exposures, genetic and epigenetic factors (Machado et al., 2016a). Although the specific underlying factors remain largely unclear, several environmental exposures comprising maternal smoking, drinking and vitamin deficiency during the gestation, and some disease-susceptibility genes, specifically interferon regulatory factor 6 (*IRF6*) and 8q24 locus, have been considered the most reliable susceptibility factors for NOC (Dixon et al., 2011; Beaty et al., 2016). The genetic predisposition to NOC is ethnicity-dependent, and the genetic basis of susceptibility varies among different populations (Dixon et al., 2011). The population of Brazil is highly admixed, with each individual showing variable ancestry proportions of Amerindians, Europeans and sub-Saharan Africans, which has been shown to affect the specific genetic susceptibility of genes and loci previously associated with NOC in other populations (Brito et al., 2012a; Bagordakis et al., 2013; de Aquino et al., 2014a; do Rego Borges et al., 2015). Identifying the common and specific risk markers for the Brazilian population will allow preventive actions and will have important impacts on

genetic counseling, besides the understanding of the underlying biological mechanisms behind this common developmental disease.

In order to summarize the current understanding of the potential genetic markers related to NOC in the Brazilian population, we carried out this systematic review of literature and meta-analyzed the most repeatedly reported markers. We also highlight the main shortcomings of the published studies to improve future research in the field.

Materials and Methods

PROTOCOL AND REGISTRATION

This review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Moher et al., 2010), and was registered at the International Prospective Register of Systematic Reviews (PROSPERO; https://www.crd.york.ac.uk/prospero/#myprospero) under the number CRD42017077272.

STUDY DESIGN

The study was undertaken to systematically review the susceptibility genetic markers for NOC in the Brazilian population, and to meta-analyze those that were more frequently reported.

ELIGIBILITY CRITERIA

Inclusion criteria: Studies that investigated the susceptibility of genetic markers for NOC in the Brazilian population. The search was conducted without time and language restrictions. The PICOS (population, intervention, comparison, outcome, study design) format was used to construct the research question with the following inclusion criteria: (i) Population: patients with NOC in the Brazilian population; (ii) Intervention: analysis of genetic markers; (iii) Comparison: inclusion of unaffected individuals (unrelated subjects for case-control studies and family members for family-based studies); (iv) Outcome: frequency and potential association of genetic markers; (v) Study Design: observational studies (case-control or family-based approaches).

Exclusion criteria: Studies were excluded for the following reasons: 1) patients with syndromic oral cleft, 2) studies that did not include Brazilian samples or mixed the Brazilian samples with those of other countries and was not possible the discrimination, 3) studies that did not include genetic markers in the analysis, 4) lack of a control group, and 5) Reviews, letters, personal opinions, book chapters and conference abstracts.

INFORMATION SOURCES AND SEARCH STRATEGY

Search strategies in the LILACS, LIVIVO, PubMed, Scopus and Web of Science databases included the following terms: "polymorphism, genetic" OR polymorphism OR polymorphisms OR "single nucleotide polymorphism" OR "SNP" OR "nucleotide variant" OR "single nucleotide variant" OR "SNV" OR "genetic variant" OR "coding variant" OR "genetic marker" OR "polymorphisms, genetic" OR "genetic polymorphism" OR "polymorphism (genetics)" OR "genetic polymorphisms" OR "genetic markers" OR "genetic linkage" OR "genetic frequency" OR genes OR genetics OR "fine mapping" OR "gene variants" OR "gene variant" OR "genetic factors" OR "mutational screening" OR "rare variant" AND "cleft lip" OR "cleft palate" OR cleft OR "oral cleft" OR clefting OR "oral clefting" OR "cleft lip palate" OR "cleft lip only" OR "cleft palate only" OR "CLP" OR "CL/P" OR "CLO" OR "CPO" OR "CL±P" OR "cleft lip and palate" OR "cleft lip and cleft palate" OR "nonsyndromic cleft lip and cleft palate" OR "oral clefts" OR "cleft lip/palate" OR "cleft lip or palate" OR "orofacial clefts" OR "orofacial cleft" AND nonsyndromic OR non-syndromic. The full record of the searches performed can be found in the Supplementary Table 1.

The search in the gray literature was conducted on August 23th, 2017 and the search on the 5 chosen databases was performed on September 11th, 2017. The partial grey literature search was performed using Google Scholar and ProQuest. The selected references were checked and managed by a reference manager software (EndNote, Thomson Reuters, Virginia, USA). In addition, the reference lists of the selected articles were hand screened for potential relevant studies that could have been missed during the electronic database searches.

STUDY SELECTION

The articles were selected in two phases. In phase 1, 2 authors (RAM and RDC) independently reviewed the titles and abstracts, and selected those that apparently met the inclusion criteria. In phase 2, the same authors read the full texts of the

selected articles at phase 1 and excluded those that did not meet the inclusion criteria. Any disagreements in the first or second phases were resolved by discussion and mutual agreement between the 2 authors.

DATA COLLECTION PROCESS AND DATA ITEMS

One author (RDC) collected the following information from the included articles: authors, year of publication, markers, study design, genotyping strategy, number of samples, origin of the samples by Brazilian states and results. The second author (RAM) crosschecked all the retrieved information. Disagreements were resolved by discussion and mutual agreement between authors.

RISK OF BIAS IN INDIVIDUAL STUDIES

Methodologically, the authors appraised all included studies according to a checklist based in Meta-Analysis of Statistics Assessment and Review Instrument (MAStARI) (The JBI, 2014). The reviewers (RAM and RDC) independently answered 9 questions for descriptive studies as Y for “yes”, N for “no”, U for “unclear” and NA for “not applicable”. After that, the risk of bias was categorized as high when the study reached up to 49% of a “yes” score, moderate when the study reached 50% to 69% of a “yes” score, and low when the study reached more than 70% of a “yes” score. Disagreements were solved by discussion between the 2 authors.

RISK OF BIAS ACROSS STUDIES

Clinical (by comparing variability among the participant’s characteristics and outcomes studied), methodological (by comparing the variability in study design and risk of bias) and statistical heterogeneities were considered.

SYNTHESIS OF RESULTS

Proportion meta-analysis based in the frequency of the genetic markers was performed using StataCorp Software version 13 (StataCorp. 2013, *Stata Statistical Software: Release 13.* College Station, TX: StataCorp LP, Texas, USA). Heterogeneity was calculated by inconsistency indexes (I^2) and a value greater than 50% was considered an indicator of substantial heterogeneity between studies (Higgins and Green, 2011). The significance level was set at 5%.

CONFIDENCE IN CUMULATIVE EVIDENCE

The Grading of Recommendation, Assessment, Development and Evaluation (GRADE) instrument (Balshem et al., 2011) was used to assess evidence quality and grading of recommendation strength in the studies included in the quantitative analysis. This assessment was based on the study design, risk of bias, inconsistency, indirectness, imprecision and other considerations. Evidence quality was characterized as high, moderate, low or very low. The GRADE was assessed using http://gradepro.org.

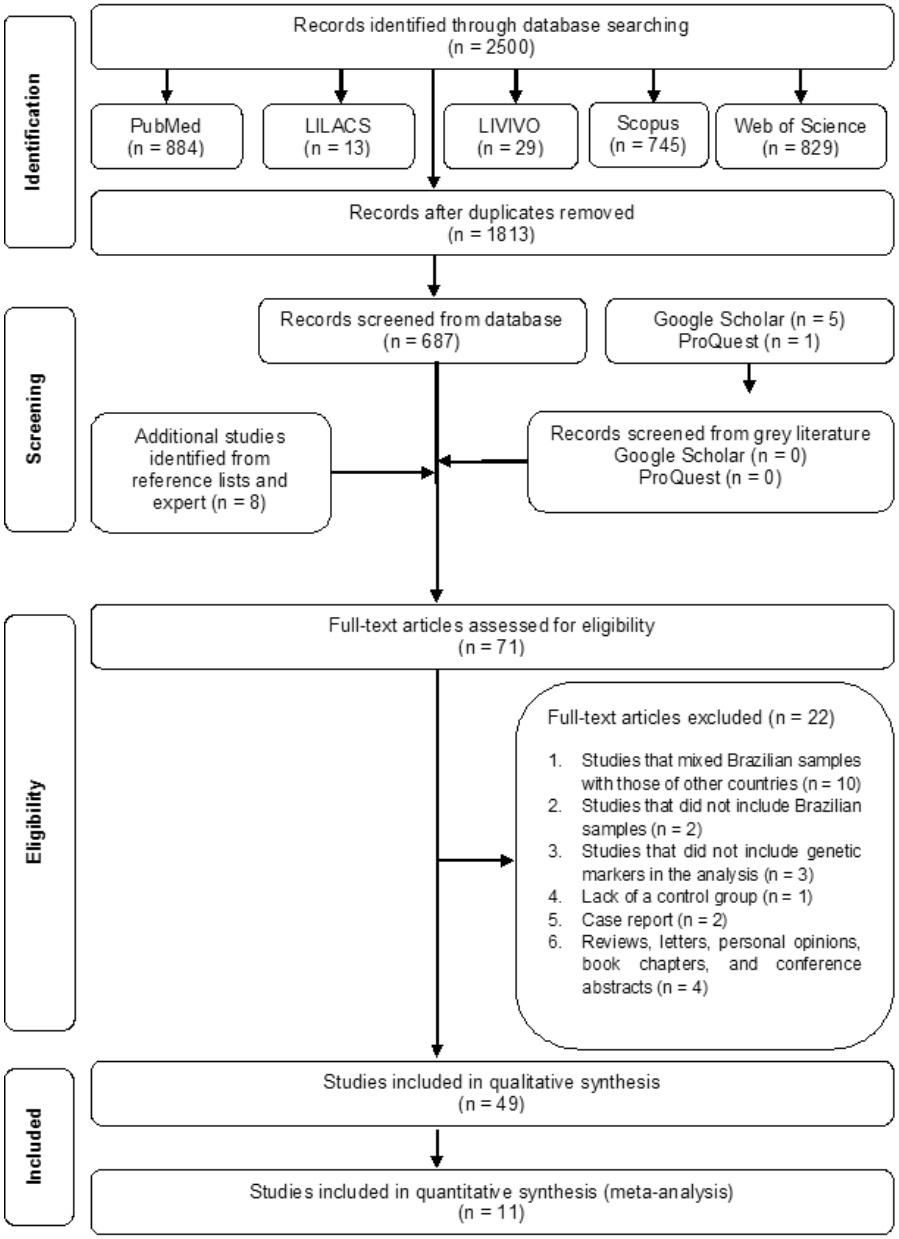
Results

STUDY SELECTION

A flow-chart detailing the processes of identification, inclusion and exclusion of the studies is depicted in Fig. 1. In the first phase, 2500 studies were selected in the 5 electronic databases. The duplicate studies were removed and 687 different citations remained. Subsequently, the comprehensive evaluation of the titles and abstracts resulted in the exclusion of 624 citations, thereby remaining 63 studies for consideration into the second phase. Moreover, 6 citations were identified in the grey literature (Google Scholar and ProQuest), but they were not included in the second phase. The experts identified 8 additional studies. In the second phase, the full-text review was then conducted on the 71 first-phase selected citations, which lead to the exclusion of 22 studies. In the end of the two phases, 49 studies fulfilled the inclusion criteria (Gaspar et al., 1999; Gaspar et al., 2002; Gaspar et al., 2004; Passos-Bueno et al., 2004; Zucchero et al., 2004; da Silva et al., 2006; Brandalize et al., 2007; Letra et al., 2007a; Letra et al., 2007b; Ehlers Bertoja et al., 2008; Menezes et al., 2008; Choi et al., 2009; Jehee et al., 2009; Letra et al., 2009; Bufalino et al., 2010; Menezes et al., 2010; Paranaíba et al., 2010; Letra et at., 2010; Fontoura et al., 2012; Araújo et al., 2012; Souza et al., 2012; Brito et al., 2012a; Brito et al., 2012b; Letra et al., 2012a; Letra et al., 2012b; Filézio et al., 2013; Antunes et al., 2013; Bagordakis et al., 2013; Paranaíba et al., 2013; Cardoso et al., 2013; Souza et al., 2013; de Aquino et al., 2013; de Aquino et al., 2014a; de Aquino et al., 2014b; Küchler et al., 2014; Bezerra et al., 2015; de Aguiar et al., 2015; do Rego Borges et al., 2015; Fontoura et al., 2015; Falagan-Lotsch et al., 2015; Waltrick-Zambuzzi at al., 2015; Brito et al., 2015; Sabóia

et al., 2015; de Souza et al., 2016; de Araújo et al., 2016; Machado et al., 2016a; Machado et al., 2016b; Machado et al., 2017; Messetti et al., 2017), but only 11 articles were used in the meta-analysis (Gaspar et al., 1999; Gaspar et al., 2004; Paranaíba et al., 2010; Araújo et al., 2012; Brito et al., 2012a; Brito et al., 2012b; Fontoura et al., 2012; Antunes et al., 2013; Bagordakis et al., 2013; de Aguiar et al., 2015; do Rego Borges et al., 2015).

Figure 1. Flow diagram of literature search and selection criteria adapted from PRISMA.



STUDY CHARACTERISTICS

The main features and findings of the studies are presented in Table 1. The selected studies were published between 1999 and 2017 and were all written in the English language. Samples from all Brazilian regions, except the Central-West region, were analyzed, but those from the Southeast region (São Paulo, Minas Gerais and Rio de Janeiro states) were the most frequent. The selected articles were observational studies.

Table 1. Overview of the selected studies (n= 49).

Studies	Markers (gene/locus)	Study Design	Genotyping strategy	Samples	Origen of the samples (Brazilian states)	Summary of Results
Messetti et al (2017)	*CRISPLD2*, *JARID2*	Case-control	Taqman assay	549 NSCL±P 236 NSCPO 693 Control	BA, MG, PB, PR	Marginal associations between *CRISPLD2* rs4783099 T allele and increased risk for NSCPO and between *JARID2* rs2237138 and decreased NSCL±P risk
Machado et al (2017)	*AXIN2*, *CDH1*	Case-parent trios	Taqman assay	223 trios of NSCL±P	BA, MG, PB, PR	A allele of *AXIN2* rs7210356 and the haplotype containing it were under-transmitted to NSCL±P patients, and haplotypes of *CDH1* were both over-transmitted and under-transmitted from parents to the children with NSCL±P
Machado et al (2016)	*TNP1*, *MSX1*, *TCOF1*, *FGFR1*, *COL2A1*, *WNT3*, *TIMP3*	Case-parent trios followed by case-control	Taqman assay	189 trios of NSCL±P 107 trios of NSCPO 318 NSCL±P 189 NSCPO 599 Control	BA, MG, PB, PR	The significant associations at TDT analysis were not supported by case-control study
de Araujo et al (2016)	*LHX8*, *IRF6*, *TCEB3*, *WNT3A*, *SUMO1*, *WNT5A*, *MSX1*, *SPRY1*, *WNT8A*, *MSX2*, *TFAP2A*, *PRSS35*, *HOXA2*, *SHH*, *SOX7*, *FOEX1*, *PTCH1*, *VAX1*, *TBX10*, *WNT1*, *SPRY2*, *PAX9*, *TGFB3*, *BMP4*, *JAG2*, *GREM1*, *KIF7*,	Case-control	Open array technology	182 NSCLP 355 Control	AL, CE, PR, RN, RS, SC, SP	SNPs in *TCEB3*, *MSX1*, *SPRY1*, *SHH*, *VAX1*, *TBX10*, *WNT11*, *PAX9*, *KIF7* and *AXIN2* increased risk for cleft, whereas SNPs in *MSX2*, *BMP4*, *JAG2* and *DVL2* decreased it. SNPs in *PRSS35* and *TFAP2A* showed both effect in

	DVL2, *RARA*, *ERBB2*, *WNT9B*, *AXIN2*, *FGF22*, *TGFB1*, *APOC2*, *CLPTM1*, *BCL3*, *TBX1*, *TBX22*					increase and decrease the cleft risk.
de Souza et al (2016)	*IRF6*, 8q24	Case-parent trios	Taqman assay	186 trios of NSCLP 32 trios of NSCLO 41 trios of NSCPO	AL, CE, PR, RS, SP	For *IRF6*, the A risk allele of rs2235371 was under-transmitted to NSCLP patients in the sample of European ancestry, and not significant associations with rs642961 were observed. Only a borderline association between rs987525 and patients with European ancestry was observed.
Machado et al (2016)	*ADPRT*, *OGG1*, *MLH1*, *APEX1*, *XRCC3*, *RAD51*, *XRCC1*, *ERCC2*	Case-parent trios	Taqman assay	223 NSCL±P	BA, MG, PB, PR	Maternal cigarette smoking interacts with *RAD51* rs1801321 genotypes to increase the risk of NSCL±P.
Saboia et al (2015)	*AXIN2*, *BMP2*, *BMP4*, *BMP7*, *DLX1*, *MMP3*	Case-parent trios	Taqman assay	98 NSCLP 27 NSCLO 23 NSCPO 162 Control	RJ	Marginal associations between NSCLP and rs788173 in *DLX1* and NSCPO and rs522616 in *MMP3*.
Brito et al (2015)	*CDH1*	Multiplex families	Sequencing	189 NSCL±P 32 NSCPO 609 Control	RS, SP, other states	4 rare and moderately penetrant variants in *CDH1* were identified. The frequency was significantly higher compared to controls.
Waltrick-Zambuzzi et al (2015)	*TCN2*, *MTRR*	Case-control	Taqman assay	270 NSCLP 71 NSCLO 60 NSCPO 466 Control	RJ	No association with *TCN2*, but logistic regression adjusted for maternal smoking reveled that *MTRR* AG genotype

						is a risk factor for NSCL±P.
Falagan-Lotsch et al (2015)	*EGF*	Case-control	RFLP	152 NSCLP 39 NSCLO 27 NSCPO 253 Control	RJ	No significant association.
Fontoura et al (2015)	*WNT9B*, *WNT3*	Case-parent trios	Taqman assay	70 trios of NSCL±P	RJ	*WNT3* was not associated, but the G allele of *WNT9B* rs1530364 was overtransmitted. One haplotype between *WNT3* and *WNT9B* was also overtransmitted at significant level.
do Rego Borges et al (2015)	*ABCA4*, *IRF6*, 8q24, *FOXE1*, *VAX1*, 18q22, *MAFB*	Case-control	Taqman assay	293 NSCL±P 352 Control	BA	Minor alleles of *FOEX1* (rs3758249) and *VAX1* (rs7078160) were significantly associated with risk of NSCL±P. Only marginal associations for *IRF6* and 8q24.
de Aguiar et al (2015)	*MTHFR*	Case-parent trios followed by case-control	Taqman assay	197 trios of NSCL±P 318 NSCL±P 598 Control	BA, MG, PB, PR	rs1801133 T allele was overtransmitted, and case-control analysis confirmed association with risk allele.
Bezerra et al (2015)	*MTHFR*, *MTR*, *MTRR*, *RFC1*	Case-control (mother and offspring)	RFLP	81 NSCLP 25 NSCLO 30 NSCPO 175 Control	RN	*MTHFR* rs1801131 C allele was associated with risk of a mother has a child with NSCLP after adjustment for alcohol consumption.
Kuchler et al (2014)	*AXIN2*, *DLX1*, *DLX2*, *EDAR*, *FGF3*, *FGF10*, *FGFR2*, *GLI2*, *GLI3*, *LHX6*, *MSX1*, *PAX9*, *PITX2*	Case-control	Taqman assay	382 NSCLP 90 NSCLO 75 NSCPO 823 Control	RJ	rs4980700 in *FGF3* showed a borderline association with unilateral clefts and tendency of association with NSCLP and bilateral clefts.

de Aquino et al (2014)	*MTHFR*, *MTHFD1*	Case-parent trios followed by case-control	Taqman assay	147 trios of NSCL±P 181 NSCL±P 478 Control	BA, MG, PR	A allele of rs2274976 was significantly associated with NSCL±P risk.
de Aquino et al (2014)	*PAX7*, *THADA*, 3p11.1, 8q21.2, 13q31.1, 15q22.2, 17q22	Case-control	Taqman assay	343 NSCLP 162 NSCLO 594 Control	BA, MG	SNP in 17q22 (rs227731) was associated with cleft risk in population with European ancestry, and 15q22.2 (rs1873147) was associated with cleft patients displaying high African ancestry. The association between *PAX7* (rs742071) and cleft risk was only detected in the combined sample.
de Aquino et al (2013)	*FGF12*, *VCL*, *CX43*, *VAX1*	Case-control	Taqman assay	195 NSCLP 105 NSCLO 385 Control	MG	As single markers, no significant associations. A *VAX1* haplotype showed marginal association.
Souza et al (2013)	*MSX1*	Case-parent trios	Fragment size analysis	156 trios of NSCL±P 26 trios of NSCPO	RS	169 bp allele was overtransmitted for NSCL±P patients.
Cardoso et al (2013)	*MSX1*	Case-control	Taqman assay	96 NSCLP 35 NSCLO 27 NSCPO 200 Control	RN	No significant association.
Paranaiba et al (2013)	*TBX1*, *PVRL1*, *MID1*, *RUNX2*, *TP63*, *TGFB3*, *MSX1*, *MYH9*, *JAG2*	Case-control	RFLP	198 NSCLP 88 NSCLO 81 NSCPO 413 Control	MG	G allele of *TBX1* rs28649236 showed protective effect.
Bagordakis et al (2013)	*ABCA4*, *IRF6*, 8q24, *FOXE1*, *VAX1*, 18q22, *MAFB*	Case-control	Taqman assay	194 NSCLP 105 NSCLO 384 Control	MG	rs560426 in *ABCA4* is a risk factor for NSCLP, whereas 8q24 locus is associated with both NSCLP and NSCLO.
Antunes et al (2013)	*TGFB3*, *BMP4*	Case-control	Taqman assay	253 NSCLP 71 NSCLO 59 NSCPO	RJ	No association with *TGFB3*, but *BMP4* was significantly associated

				450 Control		with NSCLO and oral cleft group.
Filezio et al (2013)	*GABRG3*	Case-control	Taqman assay	147 NSCLP 82 NSCLO 314 Control	MG	No significant associations after Bonferroni correction.
Letra et al (2012)	*IRF6*, *TGFA*	Case-Control	Taqman assay, Allele-specific PCR	353 NSCL±P 53 NSCPO 285 Control	SP	rs2902345 in *TGFA* was associated with NSCL±P, and rs2235371 and rs2073487, both in *IRF6*, were associated with complete left NSCL±P. Significant gene-gene interaction between *TGFA* and *IRF6*.
Letra et al (2012)	*MMP2*, *MMP3*, *MMP7*, *MMP9*, *MMP10*, *MMP13*, *MMP14*, *MMP16*, *MMP25*, *MMP27*, *TIMP1*, *TIMP2*, *TIMP3*, *TIMP4*	Case-Control	Taqman assay	411 NSCLP 10 NSCLO 73 NSCPO 413 Control	RJ, SP	rs522616 in *MMP3* showed a protective effect against NSCLP and all clefts, and rs8179096 in *TIMP2* showed protective effect against NSCLP, NSCPO and all clefts. Significant gene-gene interaction between *MMP3* and *TIMP2*.
Brito et al (2012)	8q24	Case-Control	Taqman assay	667 NSCL±P 589 Control	AL, CE, MG, PA, RJ, SP	8q24 locus is associated with NSCL±P of European ancestry.
Souza et al (2012)	*TGFA*	Case-parent trios	RFLP	147 trios of NSCL±P 28 trios of NSCPO	RS	No association with oral cleft.
Araujo et al (2012)	*BMP4*	Case-control	RFLP	98 NSCLP 25 NSCLO 246 Control	AL, PR, RS, SP	C allele of *BMP4* rs17563 showed a protective effect.
Fontoura et al (2012)	*ABCA4*, *MAFB*	Case-control	Taqman assay	400 NSCL±P 412 Control	SP	*ABCA4* SNPs showed risk effects and no significant associations with *MAFB* were observed.
Brito et al (2012)	*IRF6*	Case-control	Taqman assay	381 NSCLP 90 NSCLO 391 Control	AL, CE, PA, RJ, SP	A allele of rs642961 was significantly associated with NSCLO.

Letra et al (2010)	6q14.2-14.3	Case-control	Taqman assay	324 NSCL±P 4 NSCLO 282 Control	NA	rs7753918 in *PRSS35* was significantly associated with oral cleft.
Paranaiba et al (2010)	*IRF6*	Case-control	RFLP	177 NSCL±P 51 NSCPO 126 Control	MG	No significant associations between IRF6 and oral clefts.
Menezes et al (2010)	*WNT3A*, *WNT5A*, *WNT8A*, *WNT11*, *WNT3*, *WNT9B*	Case-control	Taqman assay	372 NSCL±P 69 NSCPO 303 Control	RJ, SP	*WNT3* rs142167 is a risk factor for NSCL±P at Bonferroni level, whereas rs9890413 risk a nominal level only. Some haplotypes of *WNT3* and *WNT3-WNT9B* were associated with oral clefts.
Bufalino et al (2010)	*MTHFR*, *MTHFD1*, *MTR*, *SLC19A1*	Case-Control (mothers of cleft patient only)	RFLP	50 NSCLP 31 NSCLO 26 NSCPO 184 Control	MG	*MTHFR* rs2274976 increased the risk of oral cleft. The risk was even increased if mothers did not take vitamins.
Letra et al (2009)	*AXIN2*, *CDH1*	Case-control	Taqman assay	323 NSCL±P 53 NSCPO 252 Control	SP	Significant association between rs7591 in *AXIN2* and incomplete NSCPO. Marginal associations between *CDH1* SNPs and NSCL±P.
Jehee et al (2009)	*IRF6*	Multiplex families	Sequencing	108 families with at least 2 affected members 100 Controls	NA	4 mutations in *IRF6* out of 108 families with NSCL±P. Mutations were not found in 100 controls.
Choi et al (2009)	*PDGF-C*	Case-control	Taqman assay	404 NSCL±P 66 NSCPO 500 Control	SP	No significant associations.
Menezes et al (2008)	*FGF3*, *FGF7*, *FGF10*, *FGF18*, *FGFR1*, *FGFR2*	Case-control	Taqman assay	326 NSCLP 53 NSCPO 281 Control	SP	rs1448037 in *FGF10* variant allele increased risk for unilateral NSCLP, and rs4631909 in *FGF3* increased risk for unilateral right NSCLP

Ehlers Bertoja et al (2008)	*TGFA*	Case-control	RFLP	133 NSCL±P 7 NSCPO 142 Control	RS	No significant associations.
Letra et al (2007)	*MMP1*, *MMP3*	Case-control	RFLP, Allele-specific PCR	139 NSCL±P 24 NSCP 255 Control	SP	No significant association with *MMP1*, but protective effect for *MMP3* SNP.
Brandalize et al (2007)	*MTHFR*, *MTR*, *MTRR*	Case-control (mother and offspring)	RFLP	87 NSCL±P 27 NSCPO 100 Control 110 mother of cleft patient 100 Control mother	RS	No significant associations.
Letra et al (2007)	*MMP9*	Case-control	RFLP	96 NSCLP 7 NSCLO 22 NSCPO 173 Control	SP	No significant associations.
da Silva et al (2006)	*MTHFR*, *TGFB3*, *MSX1*	Case-parent trios	RFLP, Fragment size analysis	60 without cleft classification	SP	No significant associations for *TGFB3* and *MSX1*, but a tendency of overtransmission of rs1801133 C allele to affected offspring.
Zucchero et al (2004)	*IRF6*	Multiplex families	Allele-specific kinetic PCR, Taqman assay, Sequencing	161 NSCLP 88 NSCLO 51 NSCPO 3 unknown	NA	Significant overtransmission of the risk allele of rs2235371.
Passos-Bueno et al (2004)	*TGFA*	Case-Control	RFLP	398 NSCL±P 138 NSCLO 385 Control	CE, SP	No significant associations.
Gaspar et al (2004)	*MTHFR*	Multiplex families, Case-control	RFLP	102 families with NSCL±P 424 NSCL±P 644 Control	CE, SP	rs1801133 in *MTHFR* was not associated with cleft risk.
Gaspar et al (2002)	*BCL3*	Multiplex families	Fragment size analysis	98 families with at least one NSCL±P patient	CE, SC, SP	Marginal association with the allele of 135 bp, with a major effect in females and in familial cases.

Gaspar et al (1999)	*MTHFR*	Case-parent trios (offspring and mother only) followed by case-control	RFLP	77 NSCL±P 59 mothers of NSCL±P 113 Control 90 Control mothers	NA	No significant associations.

NSCL±P: nonsyndromic cleft lip with or without cleft palate; NSCLO: nonsyndromic cleft lip only; NSCPO: nonsyndromic cleft palate only; NSCLP: nonsyndromic cleft lip and palate; RFLP: restriction fragment length polymorphism; NA: not available.

Brazilian states: AL: Alagoas; BA: Bahia; CE: Ceará; MG: Minas Gerais; PA: Pará; PB: Paraíba; PR: Paraná; RJ: Rio de Janeiro; RN: Rio Grande do Norte; RS: Rio Grande do Sul; SC: Santa Catarina; SP: São Paulo.

RISK OF BIAS WITHIN STUDIES

Based on the MAStARI assessment, 20 (40.8%) articles were classified as carrying a low risk of bias, and 29 (59.2%) were classified as with moderate risk for bias (Supplementary Table 2). No study was classified as high risk. The main differences were due to guideline 9, which is related to statistical analysis. The main drawbacks included 1) the lack of baseline data, not allowing proportion analysis and calculation of OR and 95% CI; 2) the inclusion of several markers without adequate correction of p value for multiple tests; 3) the combination of cleft subtypes in a unique group, not respecting the differences among them; and 4) the absence of any correction in the case-control studies for differences on ancestry contribution of the patients. Guidelines 7 (related to the outcome of samples that were withdraw or included in the final analysis) and 8 (associated with measurement of data) were classified as unclear for most of the studies because no information were available (Supplementary Table 2).

SYNTHESIS OF RESULTS

A total of 114 genes/loci was analyzed. Of these, 79 (69.3%) were reported in only one study. For various genes/loci, different markers (single nucleotide polymorphisms, SNPs) in or near, both at 5'- and 3'-UTR, were evaluated. Supplementary Table 3 depicts the markers evaluated in only one study. Among the markers reported in more than 1 study (n=35, 30.7%), more than haft was described in only 2 studies (Supplementary Fig. 1).

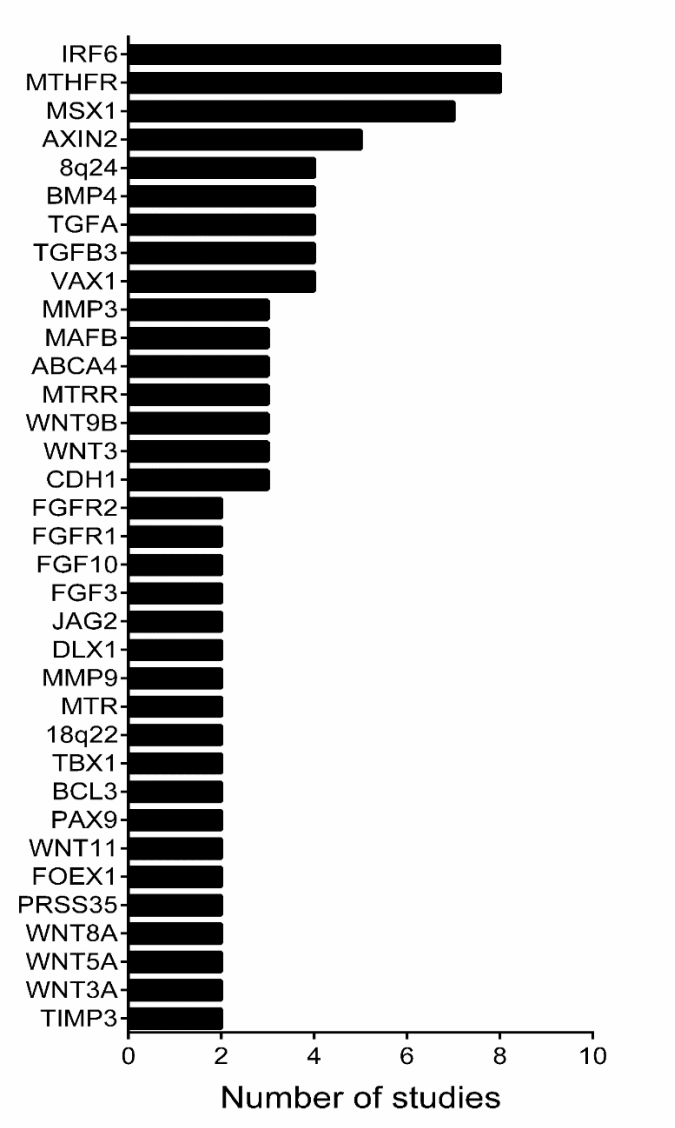
A clear majority (n=34, 69.4%) of the studies reported at least 1 associated marker with NOC, based on obtaining a statistically significant result for the association of that marker with at least 1 subtype of oral cleft or 1 considered genetic models (Table 1). Regarding the subtype of cleft, NSCL±P was analyzed in all studies, whereas NSCPO was included in only 29 studies. The size of samples of NSCPO was always smaller than NSCL±P. Some studies have combined the subtypes of clefts in a unique group, termed oral cleft group, no respecting the differences on developmental structures, epidemiologic features and embryologic timing.

RISK OF BIAS ACROSS STUDIES

The included studies used quite similar methodology, which reduced the possibility of misinterpretation. All studies showed relative homogeneity, since all of them were

observational studies. Besides this particular issue, in the meta-analysis, high heterogeneity was found for some markers, possibly due to the sample size and the results that varied widely among them.

Supplementary Figure 1. Markers evaluated in more than one study.



RESULTS OF META-ANALYSES

Meta-analysis for allele distribution was performed for each of the markers evaluated in at least 2 studies that reported necessary statistical data, as summarized in Fig. 2 and 3. The number of eligible studies was very small ranging from 2 for 8q24 (rs987525 and rs1530300), *BMP4* (rs17563), *VAX1* (rs7078160) and 18q22 (rs17085106) to 3 for *IRF6* (rs642961), *MTHFR* (rs1801133), *ABCA4* (rs560426) and *MAFB* (rs13041247).

A high heterogeneity in the frequencies of the alleles and ORs across the individual studies (I^2>50%) was observed for *BMP*4 (rs17563), *VAX1* (rs7078160) and *ABCA4* (rs560426), whereas the pertinent measures (I^2 score) had low value for *MTHFR* (rs1801133) and *MAFB4* (rs13041247) and very low value for *IRF6* (rs642961), 8q24 (rs987525 and rs1530300) and 18q22 (rs17085106). Thus, the fixed-effects model was applied in the analysis of those markers with low I^2 values.

Three studies reported data regarding the association between rs642961 in *IRF6* and NSCL±P (Paranaíba et al., 2010; Brito et al., 2012b; do Rego Borges et al., 2015), with 2 of them (Brito et al., 2012b; do Rego Borges et al., 2015) reporting high odds for the presence of the A allele in NSCL±P. The pooled OR was 1.30 (95% CI: 1.08-1.56, p=0.006), indicating that the A allele of rs642961 in *IRF6* is a susceptibility marker for NSCL±P in the Brazilian population (Fig. 2A). As regards 8q24, both SNPs (rs987525 and rs1530300) were shown to be associated with increased and significant odds to NSCL±P risk. For each SNP, 2 studies were included. As shown in Fig. 2B and 2C, the pooled OR for rs987525 was 1.21 (95% CI: 1.06-1.38, p=0.005), and for rs1530300, the pooled OR was even higher, reaching 1.54 (95% CI: 1.30-1.82, p<0.0001). The SNP rs1801133 in *MTHFR* was evaluated in 3 studies (Gaspar et al., 1999; Gaspar et al., 2004; de Aguiar et al., 2015), with 2 of them finding significant odds for the presence of the T allele in NSCL±P. The pooled OR was 1.20 (95% CI: 1.05-1.37, p=0.007), confirming that the T is a risk allele for NSCL±P (Fig. 2D). The variant C allele in *BMP4* rs17563 was significantly associated with a decreased risk for NSCL±P in the original studies (Araújo et al., 2012; Antunes et al., 2013) and in this meta-analysis (Fig. 2E). The pooled OR for the C allele was 0.68 (95% CI: 0.51-0.90, p=0.008). The meta-analyses did not find significant associations between NSCL±P and the variant alleles of rs7078160 in *VAX1* (Fig. 3A), rs560426 in *ABCA4* (Fig. 3B), rs17085106 in 18q22 (Fig. 3C) and rs13041247 in *MAFB* (Fig. 3D).

The genetic markers with a nominal p value in the allele analysis were further meta-analyzed for the genotype distribution. The pooled results revealed that the rs987525 in 8q24 significantly increased the risk of NSCL±P in the Brazilian population in heterozygous (CA vs. CC, OR: 1.32, 95% CI: 1.08-1.60, p=0.006) and homozygous (AA vs. CC, OR: 1.51, 95% CI: 1.14-2.01, p=0.005) (Fig. 4). Evidences of association between NSCL±P and the genotypes of rs1530300 in 8q24 (CC vs. TT, OR: 2.58, 95% CI: 1.20-5.56, p=0.02), rs1801133 in *MTHFR* (TT vs. CC, OR: 1.43, 95% CI: 1.05-1.94,

p=0.02) and rs17563 in *BMP4* (CC vs. TT, OR: 0.50, 95% CI: 0.39-0.66, p<0.0001) were provided only at the homozygote model (Fig. 4). No significant associations in the genotype analysis was observed for rs642961 in *IRF6* (Fig. 4).

Figure 2. Forest plots for the allele distribution of genetic markers with significant associations with NSCL±P in the pooled analysis. (A) rs642961 in *IRF6*, (B) rs987525 in 8q24, (C) rs1530300 in 8q24, (D) rs1801133 in *MTHFR* and (E) rs17563 in *BMP4*.

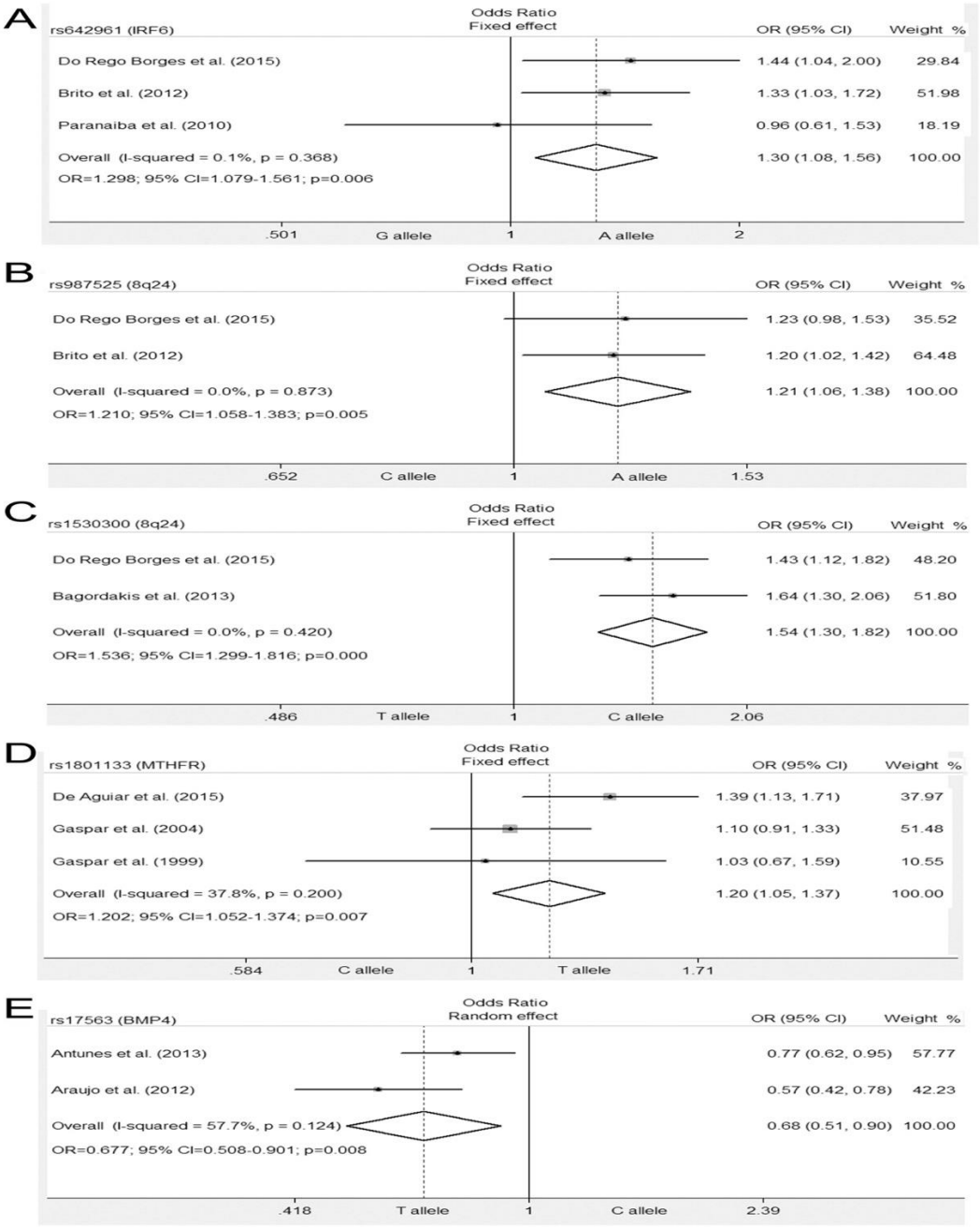


Figure 3. Forest plots for the alleles of (A) rs7078160 in *VAX1*, (B) rs560426 in *ABCA4*, (C) rs17085106 in 18q22 and (D) rs13041247 in *MAFB*.

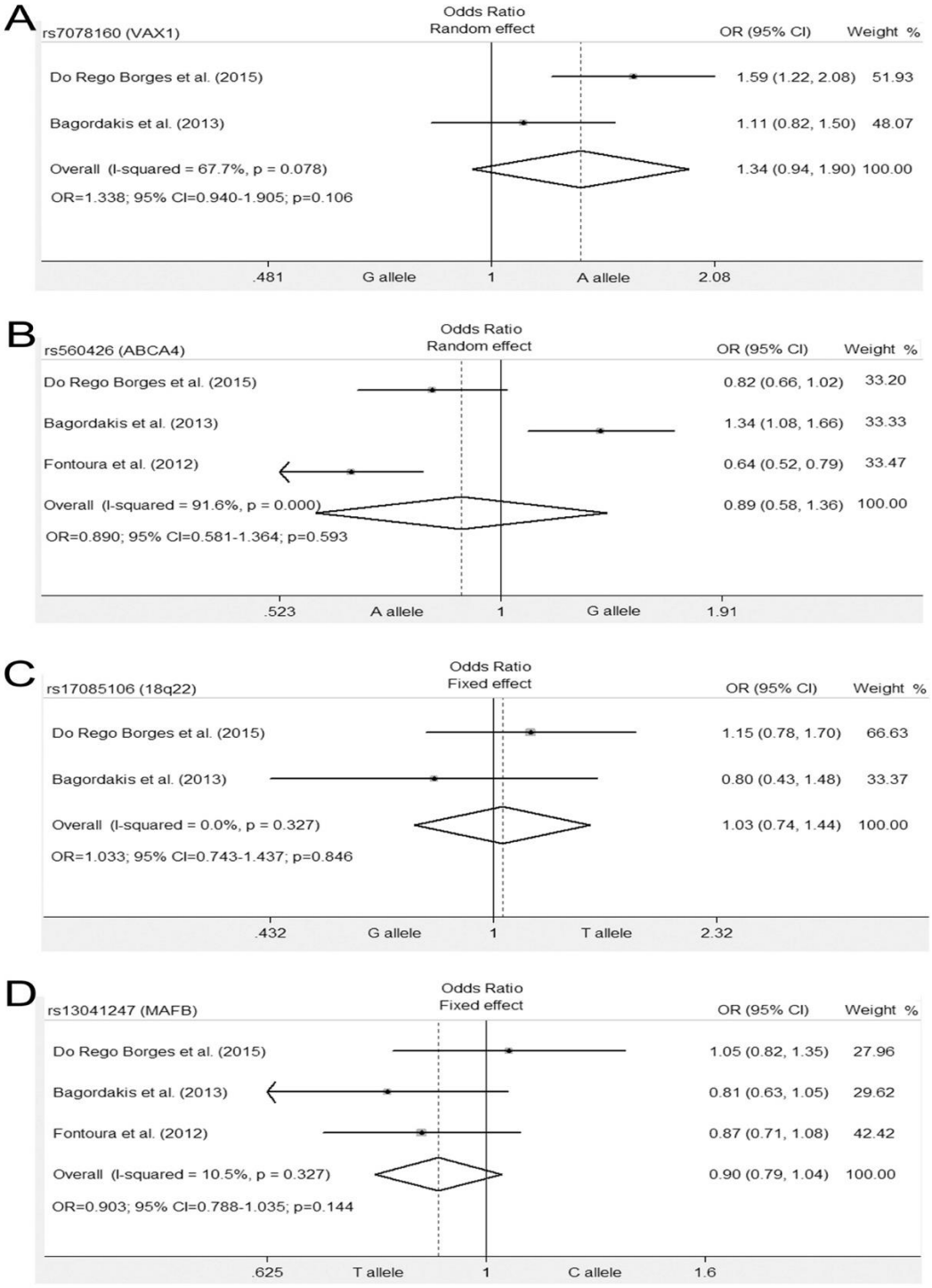
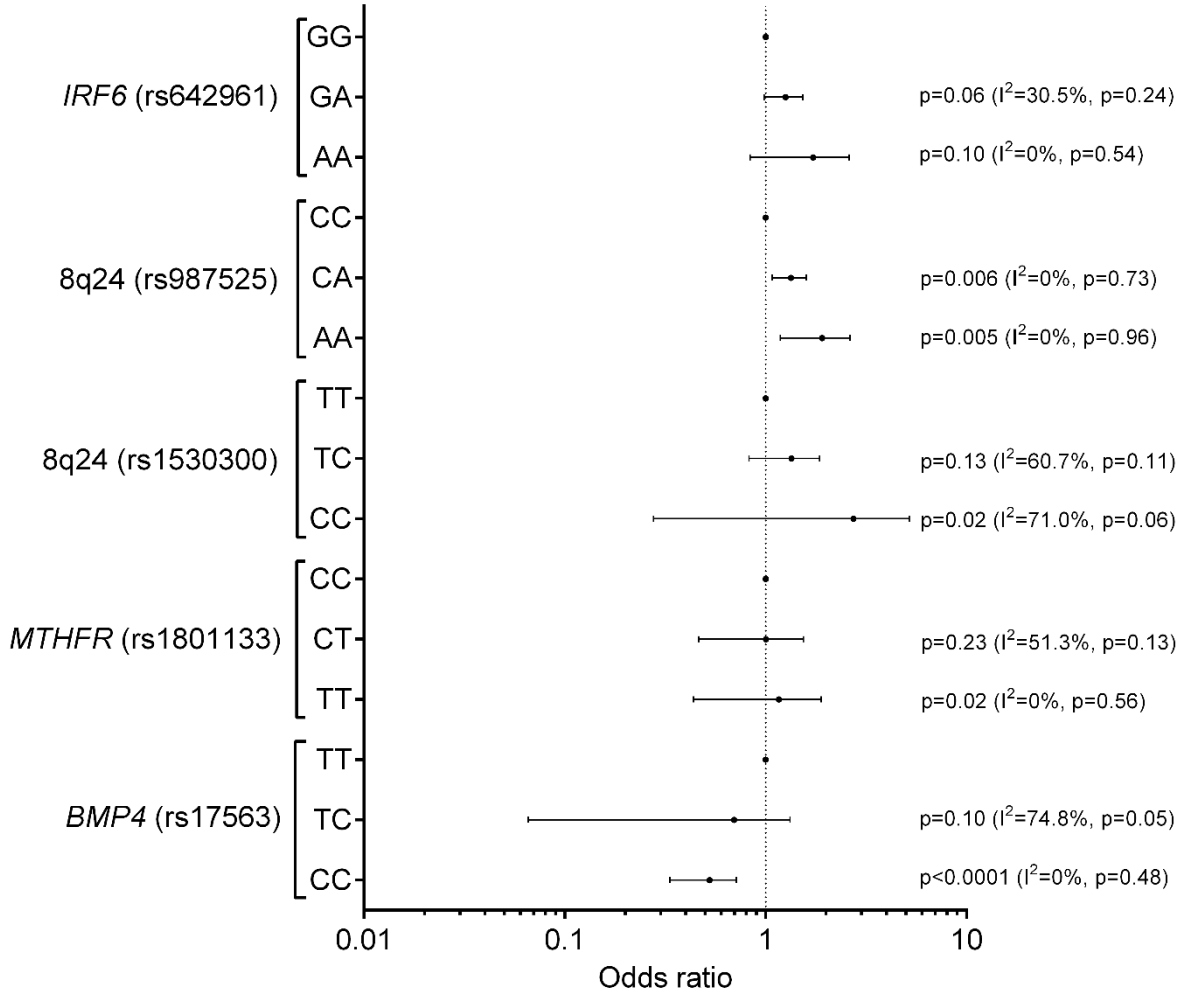


Figure 4. Association of genotypes of *IRF6* (rs642961), 8q24 (rs987525 and rs1530300), *MTHFR* (rs1801133) and *BMP4* (rs17563) with NSCL±P. ORs with 95% CI of NSCL±P were associated genotypes at heterozygote and homozygote model.



QUALITY OF EVIDENCE

Based on GRADE analysis, the quality of the evidence for the majority of the allele analysis was moderate, but 2 analysis showed low quality of evidence (*BMP4* and *ABCA4*), suggesting little confidence in the estimated effect. The moderate risk of bias and the inconsistency in some studies were the main factor responsible for the limited quality of evidence (Supplementary Table 4).

Discussion

Although NOC are amongst the most common and distressing congenital defects, the biological mechanisms associated with this multifactorial disease are partially known. Therefore, genetic and microenvironmental factors are currently the subject of intense research because their knowledge is essential in order to create translational opportunities for prevention and counselling. Different genetic studies, traditionally based in linkage or association analysis, have been applied to identify genomic susceptibility regions for NOC. In the linkage analysis, which requires large families, the investigation searches differences in co-segregation of known and unknown genetic markers transmitted through generations of affected and healthy members within the family, whereas studies of association investigate differences in the distribution of markers between cohorts containing affected and unaffected subjects. Association studies frequently use a case-control approach, but case-parent trio design (one affected child and two living biologic parents) has the advantage of minimizing influence of unknown confounders in the results, mainly differences in the population stratification within and between groups. All 49 studies identified in this revision applied association analyses, with a clear majority using case-control approach (n=32, 65.3%). Three studies (de Aquino et al., 2014b; de Aguiar et al., 2015; Machado et al., 2016b) have applied case-parent trio as a discovery approach followed by validation of the significant signals in a case-control independent sample. This approach is based on the premise that if the same effect of a disease-marker can be obtained from case-parent trio and case-control study, the magnitude of information is strong and true.

Dozens of genes/loci have accumulated enough data to be listed as a putative causal genetic factor for NOC, though only 2 (*IRF6* and 8q24 region) have been validated across multiple populations (Beaty et al., 2016). *IRF6* was initially targeted

for investigation in NOC after mutations were detected in patients with van der Woude syndrome (OMIM 119300), the most common syndrome that has oral cleft in the clinical spectrum (Kondo et al. 2002; Paranaíba et al, 2008), and in patients affected by popliteal pterygium syndrome (OMIN 119500), a rare autosomal dominant disorder with varied clinical expressivity including popliteal pterygia, cleft palate, lower lip pits, syndactyly and genital anomalies (Paranaíba et al., 2011). Zucchero and collaborators (2004) were the first to reveal the association of NSCL±P with genetic variations in *IRF6*, specifically the polymorphism rs2235371 (820G>A) that replaces a valine by an isoleucine at amino acid position 274 (V274I) of the SMIR-binding domain of IRF6. Later, this same group revealed that the association of rs2235371 with the disease is dependent of rs642961, a polymorphism that disrupts the binding site of the transcription factor AP-2a in the *IRF6* promoter, which is in strong linkage disequilibrium with rs2235371 (Rahimov et al., 2008). The association of *IRF6* variants with NOC has been replicated in various populations (Park et al., 2007; Tang et al., 2009; Birnbaum et al., 2009a; Huang et al., 2009; Salahshourifar et al., 2011; Krasone et al., 2014; Mijiti et al., 2015; Tomita et al., 2017; Moreno Uribe et al., 2017) and supported by animal studies (Ingraham et al., 2006; Iwata et al., 2013). Nevertheless, this association has not been identified in studies of African cohorts (Butali et al., 2011; Weatherley-White et al., 2011; Butali et al., 2014; Figueiredo et al., 2014).

Two previous meta-analysis have assessed the contribution of rs642961 for NOC (Wang et al., 2012; Wattanawong et al., 2016). The first meta-analysis pooled 1673 NSCL±P and 3158 controls from 6 studies and showed that the variant A allele of rs642961 was positively associated with increased risk of NSCL±P, and subgroup analyses by ethnicity and type of cleft revealed enhanced odds among Caucasians and Asians and among all cleft types (Wang et al., 2012). The second also found that rs642961 A allele shows risk effects for NSCL±P in both Asian and Caucasian populations (Wattanawong et al., 2016). Notably, both meta-analysis included the study of Paranaíba et al. (2010) and considered the samples of the Brazilian population as Caucasian. In the current study, we identified 9 studies (16 different SNPs and rare variants) that assessed *IRF6* in the Brazilian population, but only 3 of them with the SNP rs642961 showed enough data for meta-analysis. The combined result showed that the A allele is associated NSCL±P in the Brazilian population, yielding a moderate OR of 1.30. Even applying Bonferroni correction for multiple tests, the risk effect of the

A allele remained significant. Although the heterogeneity assumption tested by I^2 metric and publication bias examination suggested robustness of results, we should take into consideration that one study (Brito et al., 2012b) did not comply with Hardy-Weinberg equilibrium, one study (Paranaíba et al. 2010) did not reach significant results in the original analysis and the third study (do Rego Borges et al., 2015) showed only marginal association in a Brazilian population with high African ancestry. Furthermore, we did not observe association between *IRF6* rs642961 and the risk for NSCL±P in the Brazilian population under the heterozygote and homozygote genotype models and the studies excluded from meta-analysis reported divergent results, such as the lack of association of 8 SNPs spanning different regions of *IRF6* with NSCLP (de Araújo et al., 2016), and no association between rs642961 and NSCL±P in a case-parent trio approach but a protective effect of *IRF6* rs2235371 (de Souza et al., 2016), which was originally described in linkage disequilibrium with rs642961 and of risk for the Brazilian population (Zucchero et al., 2004). Very unique associations of rs2235371 and rs2073487 with complete left but not right NSCL±P and rs2013162 with only bilateral complete NSCL±P were also reported in a study with a Brazilian cohort (Letra et al., 2012a). Thus, further large-scale studies, taking into account the ancestry genomic composition of each individual, are warranted to confirm the findings of the current meta-analysis.

The first evidence that 8q24 is associated with NOC originated after the genome-wide association study (GWAS) conducted by Birnbaum et al. (2009b). This locus, mainly represented by rs987525, was extensively validated as a risk marker for NSCL±P by subsequent studies, including other GWAS (Grant et al., 2009; Beaty et al., 2010; Leslie et al., 2016) and in the meta-analyses involving previous GWAS (Ludwig et al., 2012; Leslie et al., 2017). Since this locus is located in an intergenic region (devoid of gene), Uslu et al. (2014) performed a series of in vivo and in vitro analyses and showed that 8q24 region contains cis-acting enhancers that control Myc expression during the development of face. The deletion of 8q24 region resulted in d*ysmorphic facial* features including cleft lip and cleft palate. In support, the multiple genetic variants in 8q24, which are significantly associated with an increased susceptibility to prostate, colorectal and breast cancer, were predicted to alter enhancer elements that physically interact with *Myc* promoter (Pomerantz et al., 2009; Ahmadiyeh et al., 2010; Sotelo et al., 2010). In this meta-analysis, we found a positive

and significant association between the variant alleles of both rs987525 and rs1530300 and NSCL±P in the Brazilian population, with higher yields for rs1530300. Previous studies have revealed a linkage disequilibrium between rs987525 and rs1530300, with the later showing a stronger association with NSCL±P (Grant et al., 2009; Bagordakis et al., 2013). Interestingly, the association between rs987525 and the risk of NSCL±P was evidenced in both heterozygous and homozygous genotypes, and for rs1530300, the association was only at homozygosity (2 copies of the risk allele). The 2 previous meta-analysis on rs987525 (Wang et al., 2012; Wattanawong et al., 2016), involving studies from different populations, corroborated our findings, supporting the participation of 8q24 on pathogenesis of NSCL±P. However, our conclusions are based in only 2 studies for each marker, *warranting further large*-scale study of this locus in the Brazilian population.

Folic acid (folate), acting as a donor and acceptor of carbon units in the methylation of homocysteine to methionine, synthesis of nucleic acids and amino acids and DNA methylation, is an essential nutrient for cell division, gene expression and maintenance of chromosome structure during embryogenesis (Friso et al., 2017). Previous study showed that the ingestion of folic acid or multivitamin complexes containing folic acid in the gestational period reduces the risk of oral cleft (Wehby and Murray et al., 2010), whereas variations on genes related to absorption, transport and metabolism of folate increase the predisposition of NOC (Bufalino et al., 2010; Wang et al., 2016). The most studied folic acid-associated gene in NOC etiology is the methylenetetrahydrofolate reductase (*MTHFR*) gene, which encoded an enzyme that catalyzes the reduction of 5,10-methylenetetrahydrofolate (5,10-CH2-THF) to 5-methyltetrahydrofolate (5-CH3THF). Variations in *MTHFR*, particularly the rs1801133 (also known as C677T) that results in an increased dissociation of FAD cofactor due to the quaternary structure loss of the enzyme and in a thermolabile protein with reduction of 30-70% in its catalytic activity (Frosst et al., 1995; Pejchal et al., 2006), are reported as exacerbating of NOC risk. In face of the high number of studies, there are 4 meta-analyses on rs1801133 and NOC risk. Both Verkleij-Hagoort et al. (2007) and Luo et al. (2012) reported no overall associations between *MTHFR* rs1801133 and NOC, whereas Pan et al. (2012) and Zhao et al. (2014) reported significant associations only in Asian populations. In the current meta-analysis, we identified 3 case-control studies (Gaspar et al., 1999; Gaspar et al., 2004; de Aguiar et al., 2015)

that analyzed rs1801133 in *MTHFR*, and the pooled result indicated an association between this polymorphism and NSCL±P risk in the Brazilian population. Thus, the associated risk of rs1801133 in *MTHFR* seems ancestry-dependent, and the admixed Brazilian population is included in the risk-group. Unfortunately, the studies did not bring information on dietary folate intake or maternal plasma folate concentration, not allowing analysis of gene-environment interactions, which should be important for the underlying mechanisms of *MTHFR* variants.

BMP4 (bone morphogenetic protein 4), a member of the transforming growth factor-β superfamily, plays important roles during facial development, and its loss of function in mice causes a series of craniofacial abnormalities, including cleft lip and palate (Juriloff and Harris, 2008). Corroborating with animal studies, several studies have focused on the effects of *BMP4* genetic variants in NOC risk in humans, with the functional SNP rs17563 being the most frequently investigated. Indeed, 2 meta-analysis have summarized the association of rs17563 with NOC. In the first meta-analysis, published in 2015 by Hu et al., only 6 studies were available and they were all focused in populations of China and Brazil. For the Chinese population, rs17563 increased risk for NSCL±P whereas a protective effect was found in the Brazilian population (Hu et al., 2015). In the second meta-analysis (Li et al., 2017), 5 more case-control studies were included, confirming the risk effect of C variant allele in the Chinese population and the protective effect in Brazilians. This recent meta-analysis also revealed that *BMP4* rs17563 was associated with a higher risk among Caucasians, represented in the meta-analysis by populations from India and Iran (Li et al., 2017). The current meta-analysis, involving 446 patients with NSCL±P and 682 controls, suggests that rs17563 in *BMP4* is a protective factor for developing NSCL±P in the Brazilian population, but the conclusions on *BMP4* were based in only 2 studies, and as revealed by Tau^2 and I^2 values, the effects were heterogeneous, *deserving further* studies to determine the real association of this marker with the susceptibility of NSCL±P in the Brazilian population. Nevertheless, we have applied random-effects model for this marker, which is more conservative than fixed-effects model, which ensures the reliability of the results even though the data come from studies with relative heterogeneity.

Approximately 70% (n=34) of the reviewed studies claimed to have identified at least one significant genetic association, though 79 (69.3%) of the genetic markers

have been analyzed only once, so it is not possible to reach trustworthy conclusions on the basis of such limited evidence. Furthermore, several positive reports were only marginal associations, no resisting to correction for multiple tests, or were associations only at a specific genetic model. Noteworthy, most articles with positive associations have usually evaluated previous known-signals for NOC, mainly those reported in GWAS. Only 20 (40.8%) studies were classified as having low risk of bias after MAStARI guidelines. We particularly noted that guideline 9, which is related to statistics, was not fulfilled in several studies. Several articles did not report the baseline data, limiting the determination of magnitude of effect, the correction for multiple tests was sometimes ignored by the authors, and another important drawback was regarding the composition of groups. Many studies combined the different types of oral clefts in a single group, not respecting the specificity of each one (NSCL±P vs. NSCPO). However, the most important shortcoming was related to the population stratification. Brazil is formed by a remarkably heterogeneous population. The intense miscegenation of the native Indians, Europeans and Africans over 5 centuries generated a high degree of genetic variability, which has been reported to directly affect most genetic polymorphic traits, such as those associated with NOC (de Aquino et al., 2014a; do Rego Borges et al., 2015). Thus, genetic studies involving Brazilian samples should use ancestry informative markers for controlling for population stratifications and to determine the real associations with specific subgroups of Brazilians, as those with high African ancestry or high native American origin. Together, the shortcomings in the susceptibility markers tested limit the possibilities to reach definitive conclusions.

Other limitations also need to be recognized. Only case-control studies were selected for the meta-analysis, which are less powerful for genetic associations than transmission disequilibrium test in case-parent trio approach. Information of environmental interventions or maternal periconceptional behaviors were not available in our study, whereas these factors may influence the associations mainly in *MTHFR* rs1801133 and folic acid intake. Last, we did not consider the genetic ancestry of the samples (groups) in the analysis and did not separate the populations based on ancestry-enrichment, one strategy with interesting results (de Aquino et al., 2014a; do Rego Borges et al., 2015).

In conclusion, this study suggests evidences for an association between increased risk for NSCL±P and *IRF6* (rs642961), 8q24 (rs987525 and rs1530300) and *MTHFR* (rs1801133), and between *BMP4* (rs17563) and a reduced risk for NSCL±P. Additional studies using larger and well-characterized Brazilian groups are needed to further replicate these findings and also those involving other potential genetic markers identified as the susceptibility for NOC. Furthermore, future studies should consider the influence of diversity of the Brazilian population in susceptibility of specific genetic variants in the pathogenesis of NOC.

Acknowledgements

This study was granted by Fundação de Amparo à Pesquisa do Estado de São Paulo-FAPESP, São Paulo, Brazil (2016/02667-0).

Conflict of interest

Authors have declared no conflicts of interest.

References

Ahmadiyeh N, Pomerantz MM, Grisanzio C, Herman P, Jia L, Almendro V, He HH, Brown M, Liu XS, Davis M, Caswell JL, Beckwith CA, Hills A, Macconaill L, Coetzee GA, Regan MM, Freedman ML. 2010. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. Proc Natl Acad Sci USA 107:9742-9746.

Antunes LS, Küchler EC, Tannure PN, Costa MC, Gouvêa CV, Olej B, Granjeiro JM. 2013. BMP4 polymorphism is associated with nonsyndromic oral cleft in a Brazilian population. Cleft Palate Craniofac J 50:633-638.

Araújo TK, Simioni M, Félix TM, de Souza LT, Fontes MÍ, Monlleó IL, Souza J, Fett-Conte AC, Secolin R, Lopes-Cendes I, Maurer-Morelli CV, Gil-da-Silva-Lopes VL. 2012. Preliminary analysis of the nonsynonymous polymorphism rs17563 in BMP4 gene in Brazilian population suggests protection for nonsyndromic cleft lip and palate. Plast Surg Int 2012:247104.

Bagordakis E, Paranaiba LM, Brito LA, de Aquino SN, Messetti AC, Martelli-Junior H, Swerts MS, Graner E, Passos-Bueno MR, Coletta RD. 2013. Polymorphisms at regions 1p22.1 (rs560426) and 8q24 (rs1530300) are risk markers for nonsyndromic cleft lip and/or palate in the Brazilian population. Am J Med Genet A 161A:1177-1180.

Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. 2011. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 64:401-406.

Beaty TH, Marazita ML, Leslie EJ. 2016. Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities. F1000Res 5:2800.

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

Bezerra JF, Oliveira GH, Soares CD, Cardoso ML, Ururahy MA, Neto FP, Lima-Neto LG, Luchessi AD, Silbiger VN, Fajardo CM, Oliveira SR, Almeida Md, Hirata RD, Rezende AA, Hirata MH. 2015. Genetic and non-genetic factors that increase the risk of non-syndromic cleft lip and/or palate development. Oral Dis. 21:393-9.

Birnbaum S, Ludwig KU, Reutter H, Herms S, de Assis NA, Diaz-Lacava A, Barth S, Lauster C, Schmidt G, Scheer M, Saffar M, Martini M, Reich RH, Schiefke F, Hemprich A, Pötzsch S, Pötzsch B, Wienker TF, Hoffmann P, Knapp M, Kramer FJ, Nöthen MM, Mangold E. 2009a. IRF6 gene variants in Central European patients with non-syndromic cleft lip with or without cleft palate. Eur J Oral Sci 117:766-769.

Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Bergé SJ, Reich RH, Schiefke F, Hemprich A, Pötzsch S, Steegers-Theunissen RP, Pötzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nöthen MM, Mangold E. 2009b. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet 41:473-477.

Brandalize AP, Bandinelli E, Borba JB, Félix TM, Roisenberg I, Schüler-Faccini L. 2007. Polymorphisms in genes MTHFR, MTR and MTRR are not risk factors for cleft lip/palate in South Brazil. Braz J Med Biol Res 40:787-91.

Brito LA, Bassi CF, 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. 2012b. IRF6 is a risk factor for nonsyndromic cleft lip in the Brazilian population. Am J Med Genet A 158A:2170-2175.

Brito LA, Paranaiba LM, Bassi CF, Masotti C, Malcher C, Schlesinger D, Rocha KM, Cruz LA, Bárbara LK, Alonso N, Franco D, Bagordakis E, Martelli H Jr, Meyer D, Coletta RD, Passos-Bueno MR. 2012a. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. Birth Defects Res A Clin Mol Teratol 94:464-468.

Brito LA, Yamamoto GL, Melo S, Malcher C, Ferreira SG, Figueiredo J, Alvizi L, Kobayashi GS, Naslavsky MS, Alonso N, Felix TM, Zatz M, Seruca R, Passos-Bueno MR. 2015. Rare Variants in the Epithelial Cadherin Gene Underlying the Genetic Etiology of Nonsyndromic Cleft Lip with or without Cleft Palate. Hum Mutat 36:1029-33.

Bufalino A, Ribeiro Paranaíba LM, Nascimento de Aquino S, Martelli-Júnior H, Oliveira Swerts MS, Coletta RD. 2010. Maternal polymorphisms in folic acid metabolic genes are associated with nonsyndromic cleft lip and/or palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol 88:980-986.

Butali A, Mossey P, Adeyemo W, Eshete M, Gaines L, Braimah R, Aregbesola B, Rigdon J, Emeka C, Olutayo J, Ogunlewe O, Ladeinde A, Abate F, Hailu T, Mohammed I, Gravem P, Deribew M, Gesses M, Adeyemo A, Marazita M, Murray J. 2014. Rare functional variants in genome-wide association identified candidate genes for nonsyndromic clefts in the African population. Am J Med Genet A 164A:2567-2571.

Butali A, Mossey PA, Adeyemo WL, Jezewski PA, Onwuamah CK, Ogunlewe MO, Ugboko VI, Adejuyigbe O, Adigun AI, Abdur-Rahman LO, Onah II, Audu RA, Idigbe EO, Mansilla MA, Dragan EA, Petrin AL, Bullard SA, Uduezue AO, Akpata O, Osaguona AO, Olasoji HO, Ligali TO, Kejeh BM, Iseh KR, Olaitan PB, Adebola AR, Efunkoya E, Adesina OA, Oluwatosin OM, Murray JC; NigeriaCRAN Collaboration. 2011. Genetic studies in the Nigerian population implicate an MSX1 mutation in complex oral facial clefting disorders. Cleft Palate Craniofac J 48:646-653.

Cardoso ML, Bezerra JF, Oliveira GH, Soares CD, Oliveira SR, de Souza KS, da Silva HP, Silbiger VN, Luchessi AD, Fajardo CM, Hirata RD, Almeida MG, Hirata MH,

Rezende AA. 2013. MSX1 gene polymorphisms in non-syndromic cleft lip and/or palate. Oral Dis 19:507-12.

Choi SJ, Marazita ML, Hart PS, Sulima PP, Field LL, McHenry TG, Govil M, Cooper ME, Letra A, Menezes R, Narayanan S, Mansilla MA, Granjeiro JM, Vieira AR, Lidral AC, Murray JC, Hart TC. 2009. The PDGF-C regulatory region SNP rs28999109 decreases promoter transcriptional activity and is associated with CL/P. Eur J Hum Genet 17:774-84.

da Silva AL, Ribeiro LA, Cooper ME, Marazita ML, Moretti-Ferreira D. 2006. Transmission analysis of candidate genes for nonsyndromic oral clefts in Brazilian parent-child triads with recurrence. Genetics and Molecular Biology, 29:439-442.

de Aguiar PK, Coletta RD, de Oliveira AM, Machado RA, Furtado PG, de Oliveira LA, de Aquino SN, Martelli-Junior H, de Almeida Reis SR, Moreira HS, Persuhn DC. 2015. rs1801133C>T polymorphism in MTHFR is a risk factor for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol 103:292-298.

de Aquino SN, Hoshi R, Bagordakis E, Pucciarelli MG, Messetti AC, Moreira H, Bufalino A, Borges A, Rangel AL, Brito LA, Oliveira Swerts MS, Martelli-Junior H, Line SR, Graner E, Reis SR, Passos-Bueno MR, Coletta RD. 2014b. MTHFR rs2274976 polymorphism is a risk marker for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol 100:30-35.

de Aquino SN, Messetti AC, Bagordakis E, Martelli-Júnior H, Swerts MS, Graner E, Coletta RD. 2013. Polymorphisms in FGF12, VCL, CX43 and VAX1 in Brazilian patients with nonsyndromic cleft lip with or without cleft palate. BMC Med Genet. 14:53.

de Aquino SN, Messetti AC, Hoshi R, Borges A, Viena CS, Reis SR, Oliveira Swerts MS, Graner E, Martelli-Júnior H, Coletta RD. 2014a. Analysis of susceptibility polymorphisms for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol 100:36-42.

de Araújo TK, Secolin R, Félix TM, de Souza LT, Fontes MÍ, Monlleó IL, de Souza J, Fett-Conte AC, Ribeiro EM, Xavier AC, de Rezende AA, Simioni M, Ribeiro-dos-Santos ÂK, dos Santos SE, Gil-da-Silva-Lopes VL. 2016. A multicentric association study between 39 genes and nonsyndromic cleft lip and palate in a Brazilian population. J Craniomaxillofac Surg 44:16-20.

de Souza LT, Kowalski TW, Ferrari J, Monlléo IL, Ribeiro EM, de Souza J, Fett-Conte AC, de Araujo TK, Gil-da-Silva-Lopes VL, Ribeiro-Dos-Santos ÂK, dos Santos SE, Félix TM. 2016. Study of IRF6 and 8q24 region in non-syndromic oral clefts in the Brazilian population. Oral Dis 22:241-245.

Dixon MJ, Marazita ML, Beaty TH, Murray JC. 2011. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet 12:167-178.

do Rego Borges A, Sá J, Hoshi R, Viena CS, Mariano LC, de Castro Veiga P, Medrado AP, Machado RA, de Aquino SN, Messetti AC, Spritz RA, Coletta RD, Reis SR. 2015. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry. Am J Med Genet A 167A:2344-2349.

Ehlers Bertoja A, Sampaio Alho C, De França E, Menegotto B, Miriam Robinson W. 2008. TGFA/TAQ I polymorphism in nonsyndromic cleft lip and palate patients from Rio Grande Do Sul, Brazil. Cleft Palate Craniofac J 45:539-44.

Falagan-Lotsch P, Lopes TS, Küchler EC, Tannure PN, Costa Mde C, Amorim LM, Granjeiro JM. 2015. The functional EGF+61 polymorphism and nonsyndromic oral clefts susceptibility in a Brazilian population. J Appl Oral Sci 23:390-6.

Figueiredo JC, Ly S, Raimondi H, Magee K, Baurley JW, Sanchez-Lara PA, Ihenacho U, Yao C, Edlund CK, van den Berg D, Casey G, DeClerk YA, Samet JM, Magee W 3rd. 2014. Genetic risk factors for orofacial clefts in Central Africans and Southeast Asians. Am J Med Genet A 164A:2572-2580.

Filézio MR, Bagordakis E, de Aquino SN, Pereira Messetti AC, Martelli-Júnior H, Swerts MS, Graner E, Coletta RD, Paranaíba LM. 2013. Polymorphisms in GABRB3 and oral clefting in the Brazilian population. DNA Cell Biol 32:125-9.

Fontoura C, Silva RM, Granjeiro JM, Letra A. 2012. Further evidence of association of the ABCA4 gene with cleft lip/palate. Eur J Oral Sci 120:553-7.

Fontoura C, Silva RM, Granjeiro JM, Letra A. 2015. Association of WNT9B Gene Polymorphisms With Nonsyndromic Cleft Lip With or Without Cleft Palate in Brazilian Nuclear Families. Cleft Palate Craniofac J 52:44-8.

Friso S, Udali S, De Santis D, Choi SW. 2017. One-carbon metabolism and epigenetics. Mol Aspects Med 54:28-36.

Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJ, den Heijer M, Kluijtmans LA, van den Heuvel LP, Rozen R. 1995. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet 10:111-113.

Gaspar DA, Matioli SR, de Cássia Pavanello R, Araújo BC, Alonso N, Wyszynski D, Passos-Bueno MR. 2004. Maternal MTHFR interacts with the offspring's BCL3 genotypes, but not with TGFA, in increasing risk to nonsyndromic cleft lip with or without cleft palate. Eur J Hum Genet 12:521-526.

Gaspar DA, Matioli SR, Pavanello RC, Araújo BC, André M, Steman S, Otto PA, Passos-Bueno MR. 2002. Evidence that BCL3 plays a role in the etiology of nonsyndromic oral clefts in Brazilian families. Genet Epidemiol. 23:364-74.

Gaspar DA, Pavanello RC, Zatz M, Passos-Bueno MR, André M, Steman S, Wyszynski DF, Matiolli SR. 1999. Role of the C677T polymorphism at the MTHFR gene on risk to nonsyndromic cleft lip with/without cleft palate: results from a case-control study in Brazil. Am J Med Genet 87:197-199.

Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield JP, Glessner JT, Thomas KA, Garris M, Frackelton EC, Otieno FG, Chiavacci RM, Nah HD, Kirschner RE, Hakonarson H. 2009. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr 155:909-913.

Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.

Hu YY, Qin CQ, Deng MH, Niu YM, Long X. 2015. Association between BMP4 rs17563 polymorphism and NSCL/P risk: a meta-analysis. Dis Markers 2015:763090.

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.

Ingraham CR, Kinoshita A, Kondo S, Yang B, Sajan S, Trout KJ, Malik MI, Dunnwald M, Goudy SL, Lovett M, Murray JC, Schutte BC. 2006. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). Nat Genet 38:1335-1340.

Iwata J, Suzuki A, Pelikan RC, Ho TV, Sanchez-Lara PA, Urata M, Dixon MJ, Chai Y. 2013. Smad4-Irf6 genetic interaction and TGFβ-mediated IRF6 signaling cascade are crucial for palatal fusion in mice. Development 140:1220-1230.

Jehee FS, Burin BA, Rocha KM, Zechi-Ceide R, Bueno DF, Brito L, Souza J, Leal GF, Richieri-Costa A, Alonso N, Otto PA, Passos-Bueno MR. 2009. Novel mutations in

IRF6 in nonsyndromic cleft lip with or without cleft palate: when should IRF6 mutational screening be done? Am J Med Genet A 149A:1319-22.

Juriloff DM, Harris MJ. 2008. Mouse genetic models of cleft lip with or without cleft palate. Birth Defects Res A Clin Mol Teratol 82:63-77.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. 2002. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet 32:285-289.

Krasone K, Lāce B, Akota I, Care R, Deeley K, Küchler EC, Vieira AR. 2014. IRF6 AP-2a binding site promoter polymorphism is associated with oral clefts in Latvia. Stomatologija 16:132-136.

Küchler EC, Sabóia TM, Vieira TC, Lips A, Tannure PN, Deeley K, Reis MF, Ho B, Rey AC, Costa MC, Granjeiro JM, Vieira AR. 2014. Studies of genes involved in craniofacial development and tumorigenesis: FGF3 contributes to isolated oral clefts and may interact with PAX9. Acta Odontol Scand 72:1070-8.

Leslie EJ, Carlson JC, Shaffer JR, Butali A, Buxó CJ, Castilla EE, Christensen K, Deleyiannis FW, Leigh Field L, Hecht JT, Moreno L, Orioli IM, Padilla C, Vieira AR, Wehby GL, Feingold E, Weinberg SM, Murray JC, Beaty TH, Marazita ML. 2017. Genome-wide meta-analyses of nonsyndromic orofacial clefts identify novel associations between FOXE1 and all orofacial clefts, and TP63 and cleft lip with or without cleft palate. Hum Genet 136:275-286.

Leslie EJ, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF, McHenry T, Resick J, Sanchez C, Jacobs J, Emanuele B, Vieira AR, Neiswanger K, Lidral AC, Valencia-Ramirez LC, Lopez-Palacio AM, Valencia DR, Arcos-Burgos M, Czeizel AE, Field LL, Padilla CD, Cutiongco-de la Paz EM, Deleyiannis F, Christensen K, Munger RG, Lie RT, Wilcox A, Romitti PA, Castilla EE, Mereb JC, Poletta FA, Orioli IM, Carvalho FM, Hecht JT, Blanton SH, Buxó CJ, Butali A, Mossey PA, Adeyemo WL, James O, Braimah RO, Aregbesola BS, Eshete MA, Abate F, Koruyucu M, Seymen F, Ma L, de Salamanca JE, Weinberg SM, Moreno L, Murray JC, Marazita ML. 2016. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. Hum Mol Genet 25:2862-2872.

Letra A, da Silva RA, Menezes R, de Souza AP, de Almeida AL, Sogayar MC, Granjeiro JM. 2007b. Studies with MMP9 gene promoter polymorphism and nonsyndromic cleft lip and palate.Am J Med Genet A 143A:89-91.

Letra A, Fakhouri W, Fonseca RF, Menezes R, Kempa I, Prasad JL, McHenry TG, Lidral AC, Moreno L, Murray JC, Daack-Hirsch S, Marazita ML, Castilla EE, Lace B, Orioli IM, Granjeiro JM, Schutte BC, Vieira AR. 2012a. Interaction between IRF6 and TGFA genes contribute to the risk of nonsyndromic cleft lip/palate. PLoS One 7:e45441.

Letra A, Menezes R, Fonseca RF, Govil M, McHenry T, Murphy MJ, Hennebold JD, Granjeiro JM, Castilla EE, Orioli IM, Martin R, Marazita ML, Bjork BC, Vieira AR. 2010. Novel cleft susceptibility genes in chromosome 6q. J Dent Res 89:927-32.

Letra A, Menezes R, Granjeiro JM, Vieira AR. 2009. AXIN2 and CDH1 polymorphisms, tooth agenesis, and oral clefts. Birth Defects Res A Clin Mol Teratol 85:169-73.

Letra A, Silva RA, Menezes R, Astolfi CM, Shinohara A, de Souza AP, Granjeiro JM. 2007a. MMP gene polymorphisms as contributors for cleft lip/palate: association with MMP3 but not MMP1. Arch Oral Biol 52:954-960.

Letra A, Silva RM, Motta LG, Blanton SH, Hecht JT, Granjeirol JM, Vieira AR. 2012b. Association of MMP3 and TIMP2 promoter polymorphisms with nonsyndromic oral clefts. Birth Defects Res A Clin Mol Teratol 94:540-548.

Li YH, Yang J, Zhang JL, Liu JQ, Zheng Z, Hu DH. 2017. BMP4 rs17563 polymorphism and nonsyndromic cleft lip with or without cleft palate: A meta-analysis. Medicine (Baltimore) 96:e7676.

Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, AlChawa T, Nasser E, Böhmer AC, Mattheisen M, Alblas MA, Barth S, Kluck N, Lauster C, Braumann B, Reich RH, Hemprich A, Pötzsch S, Blaumeiser B, Daratsianos N, Kreusch T, Murray JC, Marazita ML, Ruczinski I, Scott AF, Beaty TH, Kramer FJ, Wienker TF, Steegers-Theunissen RP, Rubini M, Mossey PA, Hoffmann P, Lange C, Cichon S, Propping P, Knapp M, Nöthen MM. 2012. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet 44:968-971.

Luo YL, Cheng YL, Ye P, Wang W, Gao XH, Chen Q. 2012. Association between MTHFR polymorphisms and orofacial clefts risk: a meta-analysis. Birth Defects Res A Clin Mol Teratol 94:237-244.

Machado RA, de Freitas EM, de Aquino SN, Martelli DR, Swerts MS, Reis SR, Persuhn DC, Moreira HS, Dias VO, Coletta RD, Martelli-Júnior H. 2017. Clinical relevance of breast and gastric cancer-associated polymorphisms as potential susceptibility markers for oral clefts in the Brazilian population. BMC Med Genet 18:39.

Machado RA, Messetti AC, de Aquino SN, Martelli-Júnior H, Swerts MS, de Almeida Reis SR, Moreira HS, Persuhn DC, Coletta RD. 2016b. Association between genes involved in craniofacial development and nonsyndromic cleft lip and/or palate in the Brazilian population. Cleft Palate Craniofac J 53:550-556.

Machado RA, Moreira HS, de Aquino SN, Martelli-Junior H, de Almeida Reis SR, Persuhn DC, Wu T, Yuan Y, Coletta RD. 2016a. Interactions between RAD51 rs1801321 and maternal cigarette smoking as risk factor for nonsyndromic cleft lip with or without cleft palate. Am J Med Genet A 170A:536-539.

Martelli-Junior H, Porto LV, Martelli DR, Bonan PR, Freitas AB, Della Coletta R. 2007. Prevalence of nonsyndromic oral clefts in a reference hospital in the state of Minas Gerais, Brazil, between 2000-2005. Braz Oral Res 21:314-317.

Menezes R, Letra A, Kim AH, Küchler EC, Day A, Tannure PN, Gomes da Motta L, Paiva KB, Granjeiro JM, Vieira AR. 2010. Studies with Wnt genes and nonsyndromic cleft lip and palate. Birth Defects Res A Clin Mol Teratol 88:995-1000.

Menezes R, Letra A, Ruff J, Granjeiro JM, Vieira AR. 2008. Studies of genes in the FGF signaling pathway and oral clefts with or without dental anomalies. Am J Med Genet A146A:1614-7.

Messetti AC, Machado RA, de Oliveira CE, Martelli-Júnior H, de Almeida Reis SR, Moreira HS, Persuhn DC, Wu T, Coletta RD. 2017. Brazilian multicenter study of association between polymorphisms in CRISPLD2 and JARID2 and non-syndromic oral clefts. J Oral Pathol Med. 46:232-239.

Mijiti A, Ling W, Guli, Moming A. 2015. Association of single-nucleotide polymorphisms in the IRF6 gene with non-syndromic cleft lip with or without cleft palate in the Xinjiang Uyghur population. Br J Oral Maxillofac Surg 53:268-274.

Moher D, Liberati A, Tetzlaff J, Altman, DG, PRISMA Group. 2010. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int. J. Surg 8:336-341.

Moreno Uribe LM, Fomina T, Munger RG, Romitti PA, Jenkins MM, Gjessing HK, Gjerdevik M, Christensen K, Wilcox AJ, Murray JC, Lie RT, Wehby GL. 2017. A

population-based study of effects of genetic loci on orofacial clefts. J Dent Res 96:1322-1329.

Mossey PA, Modell B. 2012. Epidemiology of oral clefts 2012: an international perspective. Front Oral Biol 16:1-18.

Pan Y, Zhang W, Ma J, Du Y, Li D, Cai Q, Jiang H, Wang M, Zhang Z, Wang L. 2012. Infants' MTHFR polymorphisms and nonsyndromic orofacial clefts susceptibility: a meta-analysis based on 17 case-control studies. Am J Med Genet A158A:2162-2169.

Paranaíba LM, Bufalino A, Martelli-Júnior 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.

Paranaíba LM, de Aquino SN, Bufalino A, Martelli-Júnior H, Graner E, Brito LA, Passos-Bueno MR, Coletta RD, Swerts MS. 2013. Contribution of polymorphisms in genes associated with craniofacial development to the risk of nonsyndromic cleft lip and/or palate in the Brazilian population. Med Oral Patol Oral Cir Bucal 18:e414-20.

Paranaíba LM, Martelli-Júnior H, Coletta RD. 2011. What syndrome is this? A newborn with multiple congenital abnormalities of face, genitalia, and extremities. Pediatr Dermatol 28:333-334.

Paranaíba LM, Martelli-Júnior H, Oliveira Swerts MS, Line SR, Coletta RD. 2008. Novel mutations in the IRF6 gene in Brazilian families with Van der Woude syndrome. Int J Mol Med 22:507-511.

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.

Passos-Bueno MR, Gaspar DA, Kamiya T, Tescarollo G, Rabanéa D, Richieri-Costa A, Alonso N, Araújo B. 2004. Transforming growth factor-alpha and nonsyndromic cleft lip with or without palate in Brazilian patients: results of a large case-control study. Cleft Palate Craniofac J. 41:387-91.

Pejchal R, Campbell E, Guenther BD, Lennon BW, Matthews RG, Ludwig ML. 2006. Structural perturbations in the Ala > Val polymorphism of methylenetetrahydrofolate

reductase: how binding of folates may protect against inactivation. Biochemistry 45:4808-4818.

Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, Yao K, Kehoe SM, Lenz HJ, Haiman CA, Yan C, Henderson BE, Frenkel B, Barretina J, Bass A, Tabernero J, Baselga J, Regan MM, Manak JR, Shivdasani R, Coetzee GA, Freedman ML. 2009. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 41:882-884.

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.

Rodrigues K, Sena MF, Roncalli AG, Ferreira MA. 2009. Prevalence of orofacial clefts and social factors in Brazil. Braz Oral Res 23:38-42.

Sabóia TM, Reis MF, Martins ÂM, Romanos HF, Tannure PN, Granjeiro JM, Vieira AR, Antunes LS, Küchler EC, Costa MC. 2015. DLX1 and MMP3 contribute to oral clefts with and without positive family history of cancer. Arch Oral Biol 60:223-8.

Salahshourifar I, Sulaiman WA, Zilfalil BA, Halim AS. 2011. Contribution of variants in and near the IRF6 gene to the risk of nonsyndromic cleft lip with or without cleft palate in a Malay population. Am J Med Genet A 155A:2302-2307.

Sotelo J, Esposito D, Duhagon MA, Banfield K, Mehalko J, Liao H, Stephens RM, Harris TJ, Munroe DJ, Wu X. 2010. Long-range enhancers on 8q24 regulate c-Myc. Proc Natl Acad Sci USA 107:3001-3005.

Souza LT, Kowalski TW, Collares MV, Félix TM. 2013. MSX1 gene and nonsyndromic oral clefts in a Southern Brazilian population. Braz J Med Biol Res 46:555-8.

Souza LT, Kowalski TW, Vanz AP, Giugliani R, Félix TM. 2012. TGFA/Taq I polymorphism and environmental factors in non-syndromic oral clefts in Southern Brazil. Braz Oral Res 26:431-5.

Tang W, Du X, Feng F, Long J, Lin Y, Li P, Liu L, Tian W. 2009. Association analysis between the IRF6 G820A polymorphism and nonsyndromic cleft lip and/or cleft palate in a Chinese population. Cleft Palate Craniofac J 46:89-92.

The JBI. MAStARI: Joanna Briggs Institute Reviewer's Manual 2014, Joanna Briggs Institute: University of Adelaide, 2014.

Tomita D, Yamaguchi T, Nakawaki T, Hikita Y, Adel M, Kim YI, Haga S, Takahashi M, Kawaguchi A, Isa M, Park SB, Ishida H, Maki K, Kimura R. 2017. Interferon regulatory factor 6 variants affect nasolabial morphology in East Asian populations. Arch Oral Biol 85: 142-147.

Uslu VV, Petretich M, Ruf S, Langenfeld K, Fonseca NA, Marioni JC, Spitz F. 2014. Long-range enhancers regulating Myc expression are required for normal facial morphogenesis. Nat Genet 46:753-758.

Verkleij-Hagoort A, Bliek J, Sayed-Tabatabaei F, Ursem N, Steegers E, Steegers-Theunissen R. 2007. Hyperhomocysteinemia and MTHFR polymorphisms in association with orofacial clefts and congenital heart defects: a meta-analysis. Am J Med Genet A 143A:952-960.

Waltrick-Zambuzzi M, Tannure PN, Vieira TC, Antunes LS, Romano FL, Zambuzzi WF, Granjeiro JM, Küchler EC. 2015. Genetic Variants in Folate and Cobalamin Metabolism-Related Genes in Nonsyndromic Cleft Lip and/or Palate. Braz Dent J 26:561-5.

Wang M, Pan Y, Zhang Z, Wang L. 2012. Three polymorphisms in IRF6 and 8q24 are associated with nonsyndromic cleft lip with or without cleft palate: evidence from 20 studies. Am J Med Genet A 158A:3080-3086.

Wang W, Jiao XH, Wang XP, Sun XY, Dong C. 2016. MTR, MTRR, and MTHFR gene polymorphisms and susceptibility to nonsyndromic cleft lip with or without cleft palate. Genet Test Mol Biomarkers 20:297-303.

Wattanawong K, Rattanasiri S, McEvoy M, Attia J, Thakkinstian A. 2016. Association between IRF6 and 8q24 polymorphisms and nonsyndromic cleft lip with or without cleft palate: Systematic review and meta-analysis. Birth Defects Res A Clin Mol Teratol 106:773-788.

Weatherley-White RC, Ben S, Jin Y, Riccardi S, Arnold TD, Spritz RA. 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.

Wehby GL, Murray JC. 2010. Folic acid and orofacial clefts: a review of the evidence. Oral Dis 16:11-19.

Zhao M, Ren Y, Shen L, Zhang Y, Zhou B. 2014. Association between MTHFR C677T and A1298C polymorphisms and NSCL/P risk in Asians: a meta-analysis. PLoS One 9:e88242.

Zucchero 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.

Supplementary Table 1. Search strategy in the databases.

Database	Search
PubMed (September 11, 2017)	#1: "Polymorphism, Genetic"[Mesh] OR Polymorphism OR Polymorphisms OR "single nucleotide polymorphism" OR "SNP" OR "nucleotide variant" OR "single nucleotide variant" OR "SNV" OR "genetic variant" OR "coding variant" OR "genetic marker" OR "Polymorphisms, Genetic" OR "Genetic Polymorphism" OR "Polymorphism (Genetics)" OR "Genetic Polymorphisms" OR "genetic markers" OR "genetic linkage" OR "genetic frequency" OR genes OR genetics OR "fine mapping" OR "gene variants" OR "gene variant" OR "genetic factors" OR "mutational screening" OR "rare variant" #2: "cleft lip"[Mesh] OR "cleft palate"[Mesh] OR cleft OR "oral cleft" OR clefting OR "oral clefting" OR "cleft lip palate" OR "cleft lip" OR "cleft palate" OR "cleft lip only" OR "cleft palate only" OR "CLP" OR "CL/P" OR "CLO" OR "CPO" OR "CL±P" OR "cleft lip and palate" OR "cleft lip and cleft palate" OR "nonsyndromic cleft lip and cleft palate" OR "oral clefts" OR "cleft lip/palate" OR "cleft lip or palate" OR "orofacial clefts" OR "orofacial cleft" #3: Nonsyndromic OR non-syndromic #4: #1 AND #2 AND #3
LILACS (September 11, 2017)	(tw:(Polymorphism OR polimorfismo OR "genetic markers" OR "marcadores genéticos")) AND (tw:("cleft lip" OR "fissura labial" OR "fisura labial" OR "cleft palate" OR "fissura palatina" OR "fisura palatina" OR "cleft lip and palate" OR "fissura labiopalatina" OR "fisura labiopalatina")) AND (tw:(nonsyndromic OR "non syndromic" OR "não sindrômica" OR "no sindrómico"))
LIVIVO (September 11, 2017)	TI=(Polymorphism OR Polymorphisms OR "single nucleotide polymorphism" OR "SNP" OR "nucleotide variant" OR "single nucleotide variant" OR "SNV" OR "genetic variant" OR "coding variant" OR "genetic marker" OR "Polymorphisms, Genetic" OR "Genetic Polymorphism" OR "Polymorphism Genetics" OR "Genetic Polymorphisms" OR "genetic markers" OR "genetic

	linkage" OR "genetic frequency" OR genes OR genetics OR "fine mapping" OR "gene variants" OR "gene variant" OR "genetic factors" OR "mutational screening" OR "rare variant") AND TI=(cleft OR "oral cleft" OR clefting OR "oral clefting" OR "cleft lip palate" OR "cleft lip" OR "cleft palate" OR "cleft lip only" OR "cleft palate only" OR "CLP" OR "CL/P" OR "CLO" OR "CPO" OR "CL±P" OR "cleft lip and palate" OR "cleft lip and cleft palate" OR "nonsyndromic cleft lip and cleft palate" OR "oral clefts" OR "cleft lip/palate" OR "cleft lip or palate" OR "orofacial clefts" OR "orofacial cleft") AND TI=(Nonsyndromic OR non-syndromic)
Scopus (September 11, 2017)	***TITLE-ABS-KEY (polymorphism OR polymorphisms OR "single nucleotide polymorphism" OR "SNP" OR "nucleotide variant" OR "single nucleotide variant" OR "SNV" OR "genetic variant" OR "coding variant" OR "genetic marker" OR "Polymorphisms, Genetic" OR "Genetic Polymorphism" OR "Polymorphism (Genetics)" OR "Genetic Polymorphisms" OR "genetic markers" OR "genetic linkage" OR "genetic frequency" OR genes OR genetics OR "fine mapping" OR "gene variants" OR "gene variant" OR "genetic factors" OR "mutational screening" OR "rare variant") AND TITLE-ABS-KEY (cleft OR "oral cleft" OR clefting OR "oral clefting" OR "cleft lip palate" OR "cleft lip" OR "cleft palate" OR "cleft lip only" OR "cleft palate only" OR "CLP" OR "CL/P" OR "CLO" OR "CPO" OR "CL±P" OR "cleft lip and palate" OR "cleft lip and cleft palate" OR "nonsyndromic cleft lip and cleft palate" OR "oral clefts" OR "cleft lip/palate" OR "cleft lip or palate" OR "orofacial clefts" OR "orofacial cleft") AND TITLE-ABS-KEY (nonsyndromic OR non-syndromic) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "ip"))***
Web of Science (September 11, 2017)	#1: TS=(Polymorphism OR Polymorphisms OR "single nucleotide polymorphism" OR "SNP" OR "nucleotide variant" OR "single nucleotide variant" OR "SNV" OR "genetic variant" OR "coding variant" OR "genetic marker" OR "Polymorphisms, Genetic" OR "Genetic Polymorphism" OR "Polymorphism (Genetics)" OR "Genetic Polymorphisms" OR "genetic markers" OR "genetic linkage" OR "genetic frequency" OR genes OR genetics OR "fine

	mapping" OR "gene variants" OR "gene variant" OR "genetic factors" OR "mutational screening" OR "rare variant") #2: TS=(cleft OR "oral cleft" OR clefting OR "oral clefting" OR "cleft lip palate" OR "cleft lip" OR "cleft palate" OR "cleft lip only" OR "cleft palate only" OR "CLP" OR "CL/P" OR "CLO" OR "CPO" OR "CL±P" OR "cleft lip and palate" OR "cleft lip and cleft palate" OR "nonsyndromic cleft lip and cleft palate" OR "oral clefts" OR "cleft lip/palate" OR "cleft lip or palate" OR "orofacial clefts" OR "orofacial cleft") #3: TS=(Nonsyndromic OR non-syndromic) #4: #1 AND #2 AND #3
Google Scholar (August 23, 2017)	Search 1: tudonotítulo: Polymorphism cleft lip palate Brazilian Search 2: tudonotítulo: Polymorphism cleft lip palate Brazil
ProQuest (August 23, 2017)	TI,AB(Polymorphism OR Polymorphisms OR "single nucleotide polymorphism" OR "SNP" OR "nucleotide variant" OR "single nucleotide variant" OR "SNV" OR "genetic variant" OR "coding variant" OR "genetic marker" OR "Polymorphisms, Genetic" OR "Genetic Polymorphism" OR "Polymorphism (Genetics)" OR "Genetic Polymorphisms" OR "genetic markers" OR "genetic linkage" OR "genetic frequency") AND TI,AB(cleft OR "oral cleft" OR clefting OR "oral clefting" OR "cleft lip palate" OR "cleft lip" OR "cleft palate" OR "cleft lip only" OR "cleft palate only" OR "CLP" OR "CL/P" OR "CLO" OR "CPO" OR "CL±P" OR "cleft lip and palate" OR "cleft lip and cleft palate" OR "nonsyndromic cleft lip and cleft palate") AND TI,AB(Nonsyndromic OR non-syndromic) AND TI,AB(Brazil OR Brazilian OR "Brazilian population" OR "Brazilian people" OR "Brazilian epidemiology" OR "European continental ancestry group" OR "south American" OR "south America")

Supplementary Table 2. Analysis of the risk of bias.

Studies	Questions*									% Yes	Risk of bias#
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9		
Messetti et al (2017)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
Machado et al (2017)	NA	Y	NA	Y	Y	NA	U	U	N	50.0	Moderate
Machado et al (2016)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
de Araujo et al (2016)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
de Souza et al (2016)	NA	Y	NA	Y	Y	NA	U	U	Y	66.6	Moderate
Machado et al (2016)	NA	Y	NA	Y	Y	NA	U	U	Y	66.6	Moderate
Saboia et al (2015)	NA	Y	NA	Y	Y	NA	U	U	N	50.0	Moderate
Brito et al (2015)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
Waltrick-Zambuzzi et al (2015)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Falagan-Lotsch et al (2015)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Fontoura et al (2015)	NA	Y	NA	Y	Y	NA	U	Y	Y	83.3	Low
do Rego Borges et al (2015)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
de Aguiar et al (2015)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
Bezerra et al (2015)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Kuchler et al (2014)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
de Aquino et al (2014)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
de Aquino et al (2014)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
de Aquino et al (2013)	NA	Y	Y	Y	Y	NA	U	Y	Y	85.7	Low
Souza et al (2013)	NA	Y	NA	Y	Y	NA	U	U	Y	66.6	Moderate

Cardoso et al (2013)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Paranaiba et al (2013)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
Bagordakis et al (2013)	NA	Y	Y	Y	Y	NA	U	Y	N	71.4	Low
Antunes et al (2013)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Filezio et al (2013)	NA	Y	Y	Y	Y	NA	U	U	Y	71.4	Low
Letra et al (2012)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Letra et al (2012)	NA	Y	Y	Y	Y	NA	U	Y	N	71.4	Low
Brito et al (2012)	NA	Y	Y	Y	Y	NA	Y	U	Y	85.7	Low
Souza et al (2012)	NA	Y	NA	Y	Y	NA	U	U	N	50.0	Moderate
Araujo et al (2012)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Fontoura et al (2012)	NA	Y	Y	Y	Y	NA	U	Y	N	71.4	Low
Brito et al (2012)	NA	Y	Y	Y	Y	NA	Y	U	Y	85.7	Low
Letra et al (2010)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Paranaiba et al (2010)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Menezes et al (2010)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Bufalino et al (2010)	NA	Y	Y	Y	Y	NA	U	Y	N	71.4	Low
Letra et al (2009)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Jehee et al (2009)	NA	Y	NA	Y	Y	NA	U	U	Y	66.6	Moderate
Choi et al (2009)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Menezes et al (2008)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Ehlers Bertoja et al (2008)	NA	Y	Y	Y	Y	NA	Y	U	N	71.4	Low
Letra et al (2007)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Brandalize et al (2007)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate

Letra et al (2007)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
da Silva et al (2006)	NA	Y	NA	Y	Y	NA	U	U	N	50.0	Moderate
Zucchero et al (2004)	NA	Y	NA	Y	Y	NA	U	U	Y	66.6	Moderate
Passos-Bueno et al (2004)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Gaspar et al (2004)	NA	Y	Y	Y	Y	NA	U	U	N	57.1	Moderate
Gaspar et al (2002)	NA	Y	NA	Y	Y	NA	Y	U	Y	83.3	Low
Gaspar et al (1999)	NA	Y	Y	Y	Y	NA	Y	U	N	71.4	Low

Q1. Is the study based on a random or pseudorandom sample?

Q2. Are the criteria for inclusion in the sample clearly defined?

Q3. Are confounding factors identified and strategies to deal with them stated?

Q4. Are outcomes assessed using objective criteria?

Q5. If comparisons are being made, was there sufficient description of the groups?

Q6. Is follow up carried out over a sufficient time period?

Q7. Are the outcomes of people who withdrew described and included in the analysis?

Q8. Are outcomes measured in a reliable way?

Q9. Is appropriate statistical analysis used?

*Y=Yes, N=No, U=Unclear, NA=Not applicable (which was not considered on the percentage calculation).

Risk of bias was categorized as high when the study reaches up to 49% score “yes”, moderate when the study reached 50% to 69% score “yes”, and low when the study reached more than 70% score “yes”.

Supplementary Table 3. Genes and loci analyzed in only one study.

Gene / Locus	Study
CRISPLD2	Messetti et al (2017)
JARID2	Messetti et al (2017)
TNP1	Machado et al (2016)
TCOF1	Machado et al (2016)
COL2A1	Machado et al (2016)
LHX8	de Araujo et al (2016)
TCEB3	de Araujo et al (2016)
SUMO1	de Araujo et al (2016)
SPRY1	de Araujo et al (2016)
MSX2	de Araujo et al (2016)
TFAP2A	de Araujo et al (2016)
HOXA2	de Araujo et al (2016)
SHH	de Araujo et al (2016)
SOX7	de Araujo et al (2016)
PTCH1	de Araujo et al (2016)
TBX10	de Araujo et al (2016)
SPRY2	de Araujo et al (2016)
GREM1	de Araujo et al (2016)
KIF7	de Araujo et al (2016)
DVL2	de Araujo et al (2016)
RARA	de Araujo et al (2016)
ERBB2	de Araujo et al (2016)
FGF22	de Araujo et al (2016)
TGFB1	de Araujo et al (2016)
APOC2	de Araujo et al (2016)
CLPTM1	de Araujo et al (2016)
TBX22	de Araujo et al (2016)
ADPRT	Machado et al (2016)
OGG1	Machado et al (2016)
MLH1	Machado et al (2016)
APEX1	Machado et al (2016)
XRCC3	Machado et al (2016)
RAD51	Machado et al (2016)
XRCC1	Machado et al (2016)
ERCC2	Machado et al (2016)
BMP2	Saboia et al (2015)
BMP7	Saboia et al (2015)
TCN2	Waltrick-Zambuzzi et al (2015)

EGF	Falagan-Lotsch et al (2015)
RFC1	Bezerra et al (2015)
DLX2	Kuchler et al (2014)
EDAR	Kuchler et al (2014)
GLI2	Kuchler et al (2014)
GLI3	Kuchler et al (2014)
LHX6	Kuchler et al (2014)
PITX2	Kuchler et al (2014)
MTHFD1	de Aquino et al (2014)
PAX7	de Aquino et al (2014)
THADA	de Aquino et al (2014)
3p11.1	de Aquino et al (2014)
8q21.2	de Aquino et al (2014)
3q31.1	de Aquino et al (2014)
15q22.2	de Aquino et al (2014)
17q22	de Aquino et al (2014)
FGF12	de Aquino et al (2013)
VCL	de Aquino et al (2013)
CX43	de Aquino et al (2013)
PVRL1	Paranaiba et al (2013)
MID1	Paranaiba et al (2013)
RUNX2	Paranaiba et al (2013)
TP63	Paranaiba et al (2013)
MYH9	Paranaiba et al (2013)
GABRG3	Filezio et al (2013)
MMP2	Letra et al (2012)
MMP7	Letra et al (2012)
MMP10	Letra et al (2012)
MMP13	Letra et al (2012)
MMP14	Letra et al (2012)
MMP16	Letra et al (2012)
MMP25	Letra et al (2012)
MMP27	Letra et al (2012)
TIMP1	Letra et al (2012)
TIMP2	Letra et al (2012)
TIMP4	Letra et al (2012)
6q14.2-14.3	Letra et al (2010)
PDGF-C	Choi et al (2009)
FGF7	Menezes et al (2008)
FGF18	Menezes et al (2008)

Supplementary Table 4. Quality of the studies assessed with the Grading of Recommendation, Assessment, Development, and Evaluation (GRADE) system.

	Studies (n)	Quality assessment				Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	
rs642961 (*IRF6*)	3 observational studies	Serious[a]	Not serious	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE
rs987525 (*8q24*)	2 observational studies	Not serious	Not serious	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE
rs1530300 (*8q24*)	2 observational studies	Not serious	Not serious	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE
rs1801133 (*MTHFR*)	3 observational studies	Serious[a]	Not serious	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE
rs117563 (*BMP4*)	2 observational studies	Serious[a]	Serious[b]	Serious[c]	Not serious	⊕⊕◯◯ LOW
rs7078160 (*VAX1*)	2 observational studies	Not serious	Serious[b]	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE
rs560426 (*ABCA4*)	3 observational studies	Not serious	Very serious[b]	Serious[c]	Not serious	⊕⊕◯◯ LOW
rs17085106 (*18q22*)	2 observational studies	Not serious	Not serious	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE
rs13041247 (*MAFB*)	3 observational studies	Not serious	Not serious	Serious[c]	Not serious	⊕⊕⊕◯ MODERATE

GRADE Working Group grades of evidence:

High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

[a]Some studies were considered as moderate risk of bias; [b]I^2>50; [c]Most of the studies included in the sample used convenience samples.

2.2 Artigo

2p24.2 (rs7552) is a susceptibility locus for nonsyndromic cleft lip with or without cleft palate in the Brazilian population

Artigo Artigo aceito no periódico *Clinical Genetics, 2018, doi: 10.1111/cge.13246*
(Anexo 2)

R.A. Machado[1] | E.N. Nogueira[1] | H. Martelli-Júnior[2,3] | S.R. Reis[4] | D.C. Persuhn[5] | R.D. Coletta[1]

[1]Department of Oral Diagnosis, School of Dentistry, University of Campinas, Piracicaba, São Paulo, Brazil
[2]Stomatology Clinic, Dental School, State University of Montes Claros, Montes Claros, Minas Gerais, Brazil
[3]Center for Rehabilitation of Craniofacial Anomalies, Dental School, University of José Rosario Vellano, Alfenas, Minas Gerais, Brazil
[4]Department of Basic Science, Bahiana School of Medicine and Public Health, Salvador, Bahia, Brazil
[5]Molecular Biology Department, Federal University of Paraíba, João Pessoa, Paraíba, Brazil

Correspondence
Department of Oral Diagnosis, School of Dentistry, State University of Campinas, CEP 13414-018, Piracicaba, São Paulo, Brazil. Email: coletta@fop.unicamp.br

Funding information
This work was supported by grants from The State of São Paulo Research Foundation-FAPESP, São Paulo, Brazil (2016/02667-0). H Martelli-Junior is supported by The Minas Gerais State Research Foundation-FAPEMIG, Minas Gerais, Brazil.

ABSTRACT

The population of Brazil is highly admixed, with each individual showing variable levels of Amerindian, European and African ancestry, which may interfere in the genetic susceptibility of known risk loci to nonsyndromic cleft lip with or without cleft palate (NSCL±P). Here we investigated 5 reported genome-wide loci for NSCL±P in an ancestry-structured case-control study containing 1,697 Brazilian participants (831 NSCL±P and 866 healthy controls). SNPs rs7552 in 2q24.2, rs8049367 in 16p13.3, rs1880646, rs7406226, rs9891446 in 17p13, rs1588366 in 17q23.2 and rs73039426 in 19q13.11 were genotyped using TaqMan allelic discrimination assays and genomic ancestry was estimated using a panel of 40 biallelic short insertion/deletion polymorphic markers informative of the Brazilian population. Logistic regression analysis of the single-markers revealed rs7552 in 2p24.2 is a susceptibility risk marker for NSCL±P, yielding an OR of 1.71 (95% CI: 1.31-2.24, $p=9x10^{-6}$) in the homozygous state. Several SNP-SNP interactions containing rs7552 reached significance after adjustment for multiple tests (both Bonferroni assumption and 1000 permutation test), with the most significant interaction involving the 3-loci among rs7552, rs9891446 and rs73039426 ($p=6.1x10^{-9}$ and $p_{1000\ permutation}=0.001$). Our study is the first to support the association of rs7552 in 2p24.2 with NSCL±P in the highly admixed Brazilian population.

KEYWORDS

nonsyndromic cleft lip with or without cleft palate, genetic loci, single nucleotide polymorphism, Brazilian population.

1 | INTRODUCTION

Nonsyndromic oral cleft, which include cleft lip with or without cleft palate (NSCL±P) and cleft palate only (NSCPO), is one of the most common birth defects worldwide, showing a global prevalence of 1.43:1000 live births.[1] In Brazil, with approximately 4,000 new cases every year, the prevalence ranges from 1:650 to 1:2700 live births, with considerable geographic variation.[2,3] With the involvement of multiple environmental and genetic risk factors, the etiology of the nonsyndromic oral clefts is recognized as complex and multifactorial.[4]

Previous studies based on genome-wide screening (mainly genome-wide association studies-GWAS) have identified several NSCL±P risk genes and chromosomal regions, but few have shown consistent replication across different populations.[4] Indeed, interferon regulatory factor 6 (*IRF6*) gene and the 8q24 region are the most reliable genetic risk factors for NSCL±P and have been validated in different populations.[5,6] In previous studies we have investigated the association of several GWAS-susceptibility signals in the Brazilian population, confirming some associations, including 8q24 region, and failing to confirm others, such as *IRF6*, most likely due to differential frequencies of the risk alleles which are dependent on ethnicity.[7-13] Two recent well-designed GWAS on NSCL±P have been conducted.[14,15] The 3-stage GWAS with more than 5,000 subjects from China identified a novel risk locus for NSCL±P at 16p13.3, with rs8049367 being the most significant signal.[14] Leslie et al.[15] used a multi-ethnic sample from European, Asian, African and Central and South American countries, and revealed novel susceptibility loci for NSCL±P on 2p24.2, 17q23.2 and 19q13.11, besides supported the involvement of 17p13 locus with the risk to develop NSCL±P. The first evidence of association of 17p13, which encodes *NTN1* (netrin 1), with NSCL±P risk was described by Beaty et al.[16], with later characterization of involvement of this locus with NSCL±P in a subsequent study by the same group.[17]

As validation studies of susceptibility variants in different populations are important to support their roles in development of NSCL±P, we investigated the association of NSCL±P susceptibility loci described in the recent GWAS,[14,15] represented by 2p24.2 (rs7552), 16p13.3 (rs8049367), 17p13 (rs1880646, rs7406226, rs9891446), 17q23.2 (rs1588366) and 19q13.11 (rs73039426), in the Brazilian population using an ancestry-structured case-control approach. We also analyzed the interactions among those SNPs (SNP-SNP interaction) in this cohort.

2 | MATERIAL AND METHODS

2.1 | Study population

A total of random unrelated 1,697 subjects were included in this case-control study. Patients with NSCL±P (n=831) were recruited from 3 different geographic regions of Brazil: the northeast region (University Hospital of Lauro Wanderley-HULW, João Pessoa-PB and Santo Antonio Hospital, Salvador-BA), the southeast region (Center for Rehabilitation of Craniofacial Anomalies, University of José Rosário Vellano, Alfenas-MG), and the southern region (Association of Carrier of Cleft Lip and Palate-APOFILAB, Cascavel-PR). The patients with NSCL±P were carefully examined and screened for the presence of associated anomalies or syndromes by the specialized team of the associated Centers. Among NSCL±P group, 242 were nonsyndromic cleft lip only (NSCLO) and 589 were nonsyndromic cleft lip and palate (NSCLP), and the average age of the patients was 18.49±13.84, favoring the distinction between syndromic and nonsyndromic forms of oral cleft. The control group consisted of healthy individuals, without family history of orofacial clefts and from the same geographic areas. The average age of this group was 25.24±11.79. The study was approved by the ethics review board of each of the centers/hospitals affiliated with this collaborative study. Written informed consent was obtained from the parents or guardians and/or the participants.

2.2 | Genotyping and genomic ancestry assessment

DNA was obtained from oral mucosa cells obtained by mouthwash with a 3% sucrose solution using a salting-out protocol.[18] The SNPs rs7552, rs8049367, rs1880646, rs7406226, rs9891446, rs1588366 and rs73039426 were genotyped using the StepOne Real-Time PCR system with TaqMan 5'-exonuclease allelic discrimination assays (Assay-on-Demand service, Applied Biosystems). For quality control purposes, reactions were randomly repeated in 10% of the samples for each SNP, and the concordance rate was 100%. All samples were successfully genotyped, with a genotype call rate of >99%.

The genomic ancestry of subjects was estimated with a validated panel containing 40 biallelic short insertion-deletion polymorphisms.[19] The genotyping and bioinformatics using the Structure software[20] were previously described.[21]

2.3 | Statistical analysis

The difference between groups was analyzed using Mann-Whitney test (for age), x^2 test (for gender) and Kruskal-Wallis test (for ancestry proportions). Genotype distributions were assessed for derivation from Hardy-Weinberg equilibrium (HWE) in the control group using the x^2 test with 1 degree of freedom. Multiple logistic regression analysis under unrestricted, dominant and recessive genetic models was performed with R software (SNPassoc package), considering gender and genomic ancestry as potential confounders. A Bonferroni-adjusted p-value threshold of $\alpha \leq 0.008$ was considered statistically significant. Linkage disequilibrium (LD) and haplotype analysis of *NTN1* SNPs were estimated using the HaploView software. SNP-SNP interactions were performed using the *model-based multifactor dimensionality reduction* (mbmdr) package in R, using the 1000 permutation test to eliminate the false-positive interactions. The mbmdr is a dimension reduction method for SNP-SNP interaction analysis in case-control studies, allowing the adjustment for confounders.[22] This method merges multilocus genotypes into a one-dimensional construct, categorizing the genotypic interactions into high-risk and low-risk level for the phenotype. Power for detecting a p value ≤0.05 for each SNP was calculated using the Quanto software, assuming a prevalence of NSCL±P in Brazil[2] of 0.00146 and using the most conservative ORs reported in the original studies.[14-16]

3 | RESULTS

As depicted in Supplementary Table 1, there was a significant difference in the gender between groups, but the average ancestry contributions of the groups were quite similar (Supplementary Table 1). The distributions of genotypes in the control group were consistent with predictions under the HWE (Table 1), except for rs7406226 ($p<0.0001$) that demonstrated derivation regarding the frequency of the minor G allele (AG and GG genotypes). This SNP was excluded from further analyses.

A significant association between rs7552 and NSCL±P was observed. The frequency of variant A allele was significantly higher in the NSCL±P group compared with the control group, yielding an OR of 1.35 (95% CI: 1.18-1.55, $p=1x10^{-4}$) (Table 1). Logistic regression analysis adjusted for gender and ancestry showed that the AA genotype is associated with NSCL±P (OR: 1.71, 95% CI: 1.31-2.24) with a $p=9x10^{-6}$, which remained significant after applying the conservative Bonferroni adjustment for multiple testing (Table 2). At recessive genetic model (AA vs. GG+GA), a significant

effect of rs7552 was also observed (Table 2). Significant associations with rs7552 (at homozygote and recessive models) were detected in the subgroup analysis, separating NSCLO (Supplementary Table 2) from NSCLP (Supplementary Table 3). None of the other SNPs was significantly associated with NSCL±P at Bonferroni threshold level ($p \leq 0.008$), however, rs1880646 and rs9891446 in *NTN1* showed, without any adjustment, nominal p values for allele distribution (Table 1). These SNPs were in strong LD ($D' = 0.81$ and $r^2 = 0.49$), and carriers of the A-G haplotype (A allele of rs1880646 and G allele of rs9891446) were found to be more common among NSCL±P patients compared to controls, exhibiting a moderate increased risk to NSCL±P (OR: 1.26, 95% CI: 1.05-1.51, p=0.002) (Supplementary Table 4).

The model-based multifactor dimensionality reduction approach was applied for detecting interactions among SNPs on NSCL±P risk. Table 3 summarized the significant results obtained from SNP-SNP interaction analyses with p values extracted for 1000 permutations. We found several significant interactions at 2, 3, 4 and 5 SNPs model, with most of them involving rs7552, indicating the potential role of this SNP on NSCL±P risk. At 2-SNPs model, rs7552 interacted significantly with all investigated SNPs, and p values of interactions resisted to Bonferroni adjustment for multiple testing and to 1000 permutation test (Table 3). Several 3- and 4-SNPs and one 5-SNPs model showed significance after the permutation test and many of those containing rs7552 as top SNP (Table 3).

Table 1 Characteristics of the single-nucleotide polymorphisms (SNPs), allelic distributions in the control and NSCL±P groups and power calculations.

SNP	Locus	Gene	Alleles	HWE	MAF Control	MAF NSCL±P	OR (95% CI) P value*	Power
rs7552	2p24.2	*FAM49A*	G/**A**	0.71	43.3%	50.8%	1.35 (1.18-1.55) $1x10^{-4}$	94.7%
rs8049367	16p13.3	*CRBBP*	T/**C**	0.82	36.2%	37.4%	1.05 (0.91-1.21) 0.48	99.9%
rs1880646	17p13	*NTN1*	A/**G**	0.85	32.7%	29.7%	0.86 (0.74-1.00) 0.05	86.5%
rs7406226	17p13	*NTN1*	A/**G**	<0.0001	33.9%	28.5%	-	-
rs9891446	17p13	*NTN1*	C/**G**	0.28	27.0%	31.0%	1.21 (1.04-1.41) 0.01	99.9%
rs1588366	17q23.2	*TANC2*	A/**G**	0.57	23.8%	21.5%	0.87 (0.74-1.03) 0.10	66.8%
rs73039426	19q13.11	*RHPN2*	C/**T**	0.97	11.7%	10.8%	0.91 (0.73-1.13) 0.41	60.0%

MAF: Minor allele frequency (minor alleles in bold); NSCL±P: nonsyndromic cleft lip with or without cleft palate sample; NSCLO: nonsyndromic cleft lip only; NSCPO: nonsyndromic cleft palate only.

*P values were calculated with chi-square test.

Table 2 Association of single nucleotide polymorphisms in patients with nonsyndromic cleft lip with or without cleft palate (NSCL±P). P values were adjusted for covariates by logistic regression analysis.

	Control (%)	NSCL±P (%)	OR_{Het} (95% CI)	OR_{Hom} (95% CI)	OR_{Dom} (95% CI)	OR_{Rec} (95% CI)
			p value	*p* value	*p* value	*p* value
rs7552 (GG/GA/AA)	32.4/48.5/19.1	27.3/43.7/29.0	1.03 (0.82-1.29)	1.71 (1.31-2.24)	1.22 (0.99-1.51)	1.68 (1.34-2.11)
			0.78	$9x10^{-6}$	0.06	$8x10^{-7}$
rs8049367 (TT/TC/CC)	40.9/45.8/13.3	39.3/46.6/14.1	1.07 (0.87-1.32)	1.12 (0.83-1.51)	1.08 (0.89-1.32)	1.08 (0.81-1.43)
			0.51	0.38	0.43	0.60
rs1880646 (AA/AG/GG)	45.4/43.8/10.8	50.3/40.1/9.6	0.81 (0.66-0.99)	0.80 (0.57-1.11)	0.81 (0.67-0.98)	0.88 (0.64-1.21)
			0.04	0.19	0.02	0.43
rs9891446 (CC/CG/GG)	52.5/40.9/6.6	47.5/43.0/9.5	1.17 (0.96-1.43)	1.57 (1.08-2.27)	1.23 (1.01-1.49)	1.46 (1.02-2.09)
			0.12	0.01	0.03	0.03
rs1588366 (AA/AG/GG)	57.7/37.0/5.3	62.9/31.2/5.9	0.78 (0.63-0.96)	1.01 (0.66-1.55)	0.81 (0.66-0.98)	1.11 (0.73-1.68)
			0.01	0.93	0.03	0.62
rs73039426 (CC/CT/TT)	77.9/20.7/1.4	81.1/16.1/2.8	0.74 (0.58-0.95)	1.97 (0.97-4.02)	0.82 (0.64-1.04)	2.08 (1.03-4.24)
			0.01	0.05	0.10	0.03

Table 3 SNP-SNP interactions in patients with nonsyndromic cleft lip with or without cleft palate (NSCL±P) assessed by model-based multifactor dimensionality reduction.

	SNP1	SNP2	SNP3	SNP4	SNP5	NH[a]	betaH[b]	NL[c]	betaL[d]	p value[e]	Perm. p value[f]
2 loci	rs7552	rs73039426				2	0.65	2	-0.27	$1.4x10^{-7}$	0.001
	rs7552	rs1588366				1	0.70	2	-0.34	$7.7x10^{-7}$	0.001
	rs7552	rs9891446				3	0.53	2	-0.42	$3.5x10^{-6}$	0.001
	rs7552	rs8049367				3	0.53	1	-0.23	$3.5x10^{-6}$	0.001
	rs7552	rs1880646				1	0.68	2	-0.42	$2.1x10^{-5}$	0.001
	rs73039426	rs9891446				3	0.37	2	-0.30	$1.9x10^{-4}$	0.002
3 loci	rs7552	rs9891446	rs73039426			5	0.59	3	-0.53	$6.1x10^{-9}$	0.001
	rs7552	rs8049367	rs1588366			4	0.79	2	-0.68	$3.6x10^{-8}$	0.001
	rs7552	rs1588366	rs73039426			4	0.67	2	-0.45	$4.6x10^{-8}$	0.001
	rs7552	rs1880646	rs1588366			3	0.73	2	-0.67	$2.3x10^{-7}$	0.001
	rs7552	rs1880646	rs73039426			2	0.83	2	-0.45	$4.0x10^{-7}$	0.001
	rs7552	rs9891446	rs1588366			3	0.70	4	-0.49	$7.7x10^{-7}$	0.001

	rs7552	rs8049367	rs73039426			3	0.60	0	NA	$1.8x10^{-6}$	0.002
	rs7552	rs8049367	rs9891446			4	0.69	2	-0.49	$7.1x10^{-6}$	0.005
	rs73039426	rs1880646	rs1588366			1	0.32	3	-0.77	$7.9x10^{-5}$	0.008
4 loci	rs7552	rs8049367	rs1880646	rs1588366		5	1.02	3	-1.06	$2.0x10^{-9}$	0.001
	rs7552	rs1880646	rs1588366	rs73039426		5	0.50	7	-0.85	$4.9x10^{-7}$	0.001
	rs7552	rs9891446	rs1588366	rs73039426		2	1.00	5	-1.00	$1.0x10^{-6}$	0.004
	rs73039426	rs8049367	rs1880646	rs1588366		4	0.62	5	-0.58	$1.4x10^{-6}$	0.004
5 loci	rs7552	rs8049367	rs1880646	rs1588366	rs73039426	5	1.01	6	-1.33	$4.2x10^{-8}$	0.007

[a]Number of significant high-risk genotypes in the interaction. [b]Regression coefficient in step2 for high-risk exposition. [c]Number of significant low-risk genotypes in the interaction. [d]Regression coefficient in step2 for low-risk exposition. [e]p value for the interaction model adjusted for covariates. [f]Permutation p value for the interaction model.

Supplementary Table 1 Characteristics of patients included in this study.

	Control (n=866)	NSCL±P (n=831)	p value
Gender			
Male	406 (46.88%)	466 (56.07%)	0.0002[a]
Female	460 (53.12%)	365 (43.93%)	
Ancestry			
European	61.4%	59.6%	0.98[b]
African	31.0%	32.4%	
Amerindian	7.6%	8.0%	

NSCL±P: nonsyndromic cleft lip with or without cleft palate sample.

[a]p value calculated with x^2 test and [b]p value calculated was Kruskal-Wallis test.

Supplementary Table 2 Association of single nucleotide polymorphisms in patients with nonsyndromic cleft lip only (NSCLO). P values were adjusted for covariates by logistic regression analysis.

	Control (%)	NSCLO (%)	OR_{Het} (95% CI)	OR_{Hom} (95% CI)	OR_{Dom} (95% CI)	OR_{Rec} (95% CI)
			p value	*p* value	*p* value	*p* value
rs7552 (GG/GA/AA)	32.4/48.5/19.1	25.2/40.9/33.9	1.07 (0.75-1.53)	2.20 (1.49-3.24)	1.39 (1.00-1.92)	2.11 (1.54-2.90)
			0.66	$8x10^{-5}$	0.04	$6x10^{-6}$
rs8049367 (TT/TC/CC)	40.9/45.8/13.3	40.1/48.8/11.1	1.10 (0.81-1.49)	0.90 (0.56-1.46)	1.05 (0.79-1.41)	0.86 (0.55-1.35)
			0.56	0.84	0.72	0.50
rs1880646 (AA/AG/GG)	45.4/43.8/10.8	52.9/38.8/8.3	0.75 (0.55-1.01)	0.65 (0.38-1.10)	0.73 (0.55-0.97)	0.74 (0.45-1.23)
			0.05	0.11	0.03	0.23
rs9891446 (CC/CG/GG)	52.5/40.9/6.6	48.7/39.7/11.6	1.05 (0.77-1.42)	1.81 (1.10-2.98)	1.16 (0.87-1.54)	1.77 (1.09-2.86)
			0.77	0.02	0.32	0.02
rs1588366 (AA/AG/GG)	57.7/37.0/5.3	63.6/31.4/5.0	0.78 (0.57-1.06)	0.82 (0.42-1.60)	0.79 (0.58-1.06)	0.90 (0.47-1.74)
			0.11	0.55	0.10	0.75
rs73039426 (CC/CT/TT)	77.9/20.7/1.4	81.8/15.3/2.9	0.70 (0.48-1.04)	1.81 (0.70-4.70)	0.78 (0.54-1.12)	1.93 (0.75-5.00)
			0.06	0.23	0.17	0.18

Supplementary Table 3 Association of single nucleotide polymorphisms in patients with nonsyndromic cleft lip and palate (NSCLP). P values were adjusted for covariates by logistic regression analysis.

	Control (%)	NSCLP (%)	OR_{Het} (95% CI)	OR_{Hom} (95% CI)	OR_{Dom} (95% CI)	OR_{Rec} (95% CI)
			p value	*p* value	*p* value	*p* value
rs7552 (GG/GA/AA)	32.4/48.5/19.1	28.2/44.8/27.0	1.01 (0.79-1.30)	1.54 (1.15-2.07)	1.16 (0.92-1.47)	1.53 (1.19-1.97)
			0.91	$3x10^{-3}$	0.20	$1x10^{-3}$
rs8049367 (TT/TC/CC)	40.9/45.8/13.3	39.0/45.7/15.3	1.05 (0.84-1.32)	1.20 (0.87-1.67)	1.09 (0.88-1.35)	1.17 (0.86-1.59)
			0.64	0.22	0.45	0.30
rs1880646 (AA/AG/GG)	45.4/43.8/10.8	49.2/40.6/10.2	0.83 (0.66-1.04)	0.85 (0.59-1.22)	0.83 (0.67-1.03)	0.93 (0.66-1.31)
			0.10	0.37	0.08	0.67
rs9891446 (CC/CG/GG)	52.5/40.9/6.6	47.0/44.3/8.7	1.23 (0.98-1.53)	1.49 (0.98-2.25)	1.26 (1.02-1.56)	1.35 (0.91-2.02)
			0.06	0.05	0.03	0.13
rs1588366 (AA/AG/GG)	57.7/37.0/5.3	62.6/31.1/6.3	0.78 (0.62-0.99)	1.09 (0.69-1.73)	0.82 (0.66-1.02)	1.19 (0.76-1.87)
			0.03	0.70	0.07	0.44
rs73039426 (CC/CT/TT)	77.9/20.7/1.4	80.8/16.5/2.7	0.75 (0.57-0.99)	1.99 (0.92-4.27)	0.83 (0.64-1.08)	2.10 (0.98-4.50)
			0.04	0.07	0.15	0.05

Supplementary Table 4 Haplotype analysis of the single nucleotide polymorphisms in *NTN1* in controls and patients with nonsyndromic cleft lip with or without palate (NSCL±P). P value was adjusted for co-variants by logistic regression analysis.

Haplotype	Control	NSCL±P	OR (95%CI)	p value
A-C	43.3%	41.5%	Reference	
A-G	23.9%	28.9%	1.26 (1.05-1.51)	0.002
G-C	29.7%	27.5%	0.96 (0.81-1.14)	0.08
G-G	3.1%	2.1%	0.73 (0.42-1.26)	0.32

Sequence: rs1880646 and rs9891446.

4 | DISCUSSION

Studies have described association of many genes and chromosomal regions with NSCL±P, but a clear majority of these studies was conducted with populations of European and Asian origins,[4] with few studies focusing in other populations such as the Brazilian population, one of the most admixed populations in the world. Replication studies of NSCL±P susceptibility in different populations are quite important to define common and population-specific risk alleles and to understand the underlying genetic architecture of this common malformation.[1] In this study we investigated 7 SNPs previously associated with NSCL±P by recent GWAS using a large ancestry-structured case-control cohort from Brazil. We found a strong association of rs7552 in 2p24.2 with NSCL±P. Noteworthy, the association with rs7552 was detected in both NSCLO and NSCLP subgroups, but a stronger association with NSCLO was observed. This is the first report showing a more pronounced effect of rs7552 among NSCLO than in NSCLP patients. Considering SNP-SNP interactions, several interactions containing rs7552 as lead SNP were significant after Bonferroni correction, and all of them were confirmed in the permutation test. Indeed, rs7552 demonstrated significant interaction effects on risk of NSCL±P with all investigated SNPs in 2-SNPs model, supporting the association of this locus with NSCL±P.

The 2p24.2 locus was initially associated with NSCL±P in a multi-ethnic GWAS, with rs7552 being the top marker.[15] This association was detected in the combined discovery-sample, since the association of rs7552 did not withstand the stratification into continental-ancestry groups (European, Asian and Central/South American). The Central and South American sample was derived from several countries, but samples from Colombia represented more than 60% of this group. Interestingly, this study also validated its top associations in a replication sample (607 NSCL±P and 1,685 controls) of European ancestry, and a strong and significant association between rs7552 and NSCL±P risk was observed, yielding an OR of 1.47 (95% CI: 1.27-1.71, $p=2.68x10^{-7}$), which is contained in the confidence interval of the current report. Two previous genome-wide linkage scans have suggested 2q24.2 region as harboring a susceptibility locus for nonsyndromic oral clefts.[23,24] The first study investigated a Japanese family with soft palate cleft spanning 3 successive generations, yielding a strong linkage score at 2p24.2-24.1.[23] The subsequent study investigated 2 Chinese multiplex families with NSCL±P, giving evidence of linkage for the phenotype at the 2p24-p25 region.[24] The associated variant rs7552 is located in the 3'-untranslated

region (3'-UTR) of *FAM49A*. Beyond its expression in many tissues,[25] with higher levels in brain and thyroid, little is known about the encoded product of *FAM49A* gene. Variants in 3'-UTR may influence polyadenylation, translation efficiency, localization and stability of the mRNA, as well as binding of microRNAs, altering one of the most common mechanisms of translational repression.[26] Furthermore, this SNP may carry out a regulatory function or it may be not the true etiologic variant - it is located in a large haplotype block. Further efforts are needed to clarify the relationship between 2p24.2 and the development of NSCL±P, allowing the elucidation of mechanisms through which 2p24.2 (rs7552 or another variant in LD with it) leads to NSCL±P predisposition, thus translating the findings of association studies into the clinic.

The rs7406226 SNP in 17p13 did not adhere to the HWE and was excluded, and the other 2 SNPs in this gene (rs1880646 and rs9891446) showed only nominally significant associations as single markers, which does not appear to reflect a power limitation. However, both rs1880646 and rs9891446 exhibited LD and carriers of the A-G haplotype were found to be more prevalent among NSCL±P patients as compared to controls, showing a significant risk for oral clefts. The first study associating this locus with nonsyndromic oral clefts was a GWAS performed in Europeans and Asians.[16] In a subsequent study with a case-parent trio cohort of European ancestry, Beaty et al.[17] confirmed the association of *NTN1* in 17p13 with NSCL±P. Moreover, this locus showed genome-wide significance in the genome-wide meta-analysis performed by Ludwig et al.[27] and in the GWAS of the Chinese cases and controls.[14] Together, the results point to an association between 17p13 locus containing *NTN1* and the pathogenesis of NSCL±P, but additional studies are warranted.

The present study failed to replicate the findings of the other GWAS candidate loci for NSCL±P (rs8049367 in 16p13.3, rs1588366 in 17q23.2 and rs73039426 in 19q13.11), and this may have been related to admixed ancestry of our population or due to the sample size of our study. Sufficient power to detect a genetic effect for rs8049367 should have been achieved, however the powers for rs1588366 and rs73039426 were limited (~60%). Of note, we have assumed the most conservative genetic effect sizes of previous reports, whereas the true effect size in the Brazilian population is unknown. Another limitation was that we did not control for environmental risk factors, although the major source of confounding for genetic studies is population stratification, and our results were adjusted for this. Some strengths of this study include the high geographic coverage by including samples from 3 different regions of

Brazil and the assessment of the ancestry contribution of each patient, correcting for specific effects tof population stratification.

To our knowledge, this is the first study regarding the relationship of 2p24.2 locus with NSCL±P in the Brazilian population. However, additional studies are still necessary to unveil the exact mechanism by which 2p24.2 locus contribute to NSCL±P. Furthermore, our findings showed that the diversity of the Brazilian population clearly influences the susceptibility of specific loci in the pathogenesis of NSCL±P, underscoring the importance of detailed genetic studies of NSCL±P in different populations to improve our understanding of the etiology of this complex birth defect.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev Genet.* 2011;12(3):167-178.
2. Martelli-Junior H, Porto LV, Martelli DR, Bonan PR, Freitas AB, Della Coletta R. Prevalence of nonsyndromic oral clefts in a reference hospital in the state of Minas Gerais, Brazil, between 2000-2005. *Braz Oral Res*. 2007;21(4):314-317.
3. Rodrigues K, Sena MF, Roncalli AG, Ferreira MA. Prevalence of orofacial clefts and social factors in Brazil. *Braz Oral Res*. 2009;23(1):38-42.
4. Beaty TH, Marazita ML, Leslie EJ. Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities. *F1000Res*. 2016;5:2800.
5. Wang M, Pan Y, Zhang Z, Wang L. Three polymorphisms in IRF6 and 8q24 are associated with nonsyndromic cleft lip with or without cleft palate: evidence from 20 studies. *Am J Med Genet A*. 2012;158A(12):3080-6.
6. Wattanawong K, Rattanasiri S, McEvoy M, Attia J, Thakkinstian A. Association between IRF6 and 8q24 polymorphisms and nonsyndromic cleft lip with or without cleftpalate: Systematic review and meta-analysis. *Birth Defects Res A Clin Mol Teratol*. 2016;106(9):773-88.
7. Paranaíba LM, Bufalino A, Martelli-Júnior H, de Barros LM, Graner E, Coletta RD. Lack of association between IRF6 polymorphisms (rs2235371 and rs642961) and non-syndromic cleft lip and/or palate in a Brazilian population. *Oral Dis*. 2010;16(2): 193-197.

8. Brito LA, Paranaiba LM, Bassi CF, et al. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. *Birth Defects Res A Clin Mol Teratol*. 2012;94(6):464-468.
9. Bagordakis E, Paranaiba LM, Brito LA, et al. Polymorphisms at regions 1p22.1 (rs560426) and 8q24 (rs1530300) are risk markers for nonsyndromic cleft lip and/or palate in the Brazilian population. *Am J Med Genet A*. 2013;161A(5):1177-1180.
10. de Aquino SN, Messetti AC, Bagordakis E, et al. Polymorphisms in FGF12, VCL, CX43 and VAX1 in Brazilian patients with nonsyndromic cleft lip with or without cleft palate. *BMC Med Genet*. 2013;14:53.
11. de Aquino SN, Messetti AC, Hoshi R, et al. Analysis of susceptibility polymorphisms for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. *Birth Defects Res A Clin Mol Teratol*. 2014;100(1):36-42.
12. de Aguiar PK, Coletta RD, de Oliveira AM, et al. rs1801133C>T polymorphism in MTHFR is a risk factor for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. *Birth Defects Res A Clin Mol Teratol*. 2015;103(4):292-298.
13. do Rego Borges A, Sá J, Hoshi R, et al. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry. *Am J Med Genet A*. 2015;167A(10):2344-2349.
14. Sun Y, Huang Y, Yin A, et al. Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. *Nat Commun*. 2015;6: 6414.
15. Leslie EJ, Carlson JC, Shaffer JR, et al. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. *Hum Mol Genet*. 2016;25(13):2862-2872.
16. Beaty TH, Murray JC, Marazita ML, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet*. 2010;42(6):525-9.
17. Beaty TH, Taub MA, Scott AF, et al. Confirming genes influencing risk to cleft lip with/without cleft palate in a case-parent trio study. *Hum Genet*. 2013;132(7):771-81.
18. Aidar M, Line SR. A simple and cost-effective protocol for DNA isolation from buccal epithelial cells. *Braz Dent J*. 2007;18(2):148-152.

19. Bastos-Rodrigues L, Pimenta JR, Pena SD. The genetic structure of human populations studied through short insertion-deletion polymorphisms. *Ann Hum Ann Hum Genet*. 2006;70:658-65.
20. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945-59.
21. Messetti AC, Machado RA, de Oliveira CE, et al. Brazilian multicenter study of association between polymorphisms in CRISPLD2 and JARID2 and non-syndromic oral clefts. *J Oral Pathol Med*. 2017;46(3):232-239.
22. Calle ML, Urrea V, Vellalta G, Malats N, Steen KV. Improving strategies for detecting genetic patterns of disease susceptibility in association studies. *Stat Med*. 2008; 27(30):6532.
23. Tsuda M, Yamada T, Mikoya T, et al. A type of familial cleft of the soft palate maps to 2p24.2-p24.1 or 2p21-p12. *J Hum Genet*. 2010;55(2):124-6.
24. Wang Y, Li X, Zhu WL, et al. Genome-wide and interaction linkage scan for nonsyndromic cleft lip with or without cleft palate in two multiplex families in Shenyang, China. *Biomed Environ Sci*. 2010;23(5):363-70.
25. Fagerberg L, Hallström BM, Oksvold P, et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13(2):397-406.
26. Króliczewski J, Sobolewska A, Lejnowski D, Collawn JF, Bartoszewski R. microRNA single polynucleotide polymorphism influences on microRNA biogenesis and mRNA target specificity. *Gene.* 2018;640:66-72.
27. Ludwig KU, Mangold E, Herms S, et al. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nat Genet*. 2012;44(9):968-71.

2.3 Artigo

Interactions between superoxide dismutase and paraoxonase polymorphic variants in nonsyndromic cleft lip with or without cleft palate in the Brazilian population

Artigo submetido ao periódico *Journal de Dental Research* (Anexo 3)

Renato Assis Machado[1], Carolina de Oliveira Silva[1], Darlene Camati Persuhn[2], Vanessa de Melo Cavalcanti Dantas[2], Silvia Regina de Almeida Reis[3], Tao Wu[4], Sergio Roberto Line[5], Hercilio Martelli-Junior[6], Ricardo D. Coletta[1*]

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

2 Molecular Biology Department, Federal University of Paraiba, João Pessoa, Paraiba, Brazil

3 Department of Basic Science, Bahiana School of Medicine and Public Health, Salvador, Bahia, Brazil

4 Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China

5 Department of Morphology, School of Dentistry, University of Campinas, Piracicaba, São Paulo, Brazil

6 Stomatology Clinic, Dental School, State University of Montes Claros, Montes Claros, Minas Gerais, Brazil and Center for Rehabilitation of Craniofacial Anomalies, Dental School, University of José Rosario Vellano, Alfenas, Minas Gerais, Brazil

Running title: Interactions between *SOD* and *PON* genetic variants in NSCL±P.

*Corresponding author: Ricardo D. Coletta, Department of Oral Diagnosis, School of Dentistry, State University of Campinas, CEP 13414-018, Piracicaba, São Paulo, Brazil. E-mail: coletta@fop.unicamp.br

Abstract

Scope: During development, oxidative stress is hypothesized to mediate embryotoxicity, which may be intensified by exposition to environmental factors and by genetic variations in the enzymes involved in protecting cells from these damaging effects, including superoxide dismutase (SOD) and paraoxonase (PON). The aim of this study was to evaluate the influence of single-nucleotide polymorphisms (SNP) in genes associated with the neutralization of oxidative stress (*SOD* and *PON* family members) in the risk of nonsyndromic cleft lip with or without cleft palate (NSCL±P) in the Brazilian population, considering gene, gene-gene (GxG) and gene-environmental factor (GxE) interactions.

Methods: This study initially evaluated the association of 28 SNP in *SOD1*, *SOD2*, *SOD3*, *PON1*, *PON2* and *PON3* among 325 NSCL±P trios. Multiple logistic regression analyses were used to explore gene, GxG and GxE, involving factors that induce oxidative stress accumulation during pregnancy such as exposure to agrotoxics, environmental contact with agrotoxics, cigarette smoking, consumption of alcohol and drugs and folic acid supplementation. Signals that resisted to both Bonferroni correction and permutation test were subsequently confirmed in an ancestry-structured case-control analysis with 722 NSCL±P and 866 controls.

Results: In the trio sample, transmission disequilibrium test (allele and haplotype) and GxE analysis showed no significant associations, but multiple pairwise GxG interactions containing 10 SNP in *PON1*, *PON2* and *PON3* were detected and further examined in the case-control sample. The *PON1* rs2237583 and *PON2* rs17166879 yielded significant SNP-SNP interactions after adjustment for multiple tests (both Bonferroni assumption and 10,000 permutation test). The C allele and the CT genotype of *PON1* rs2237583 evoke significant protective effects against NSCL±P, while rs3917490 showed a significant association only in the sample composed of patients displaying high African ancestry.

Conclusion: Our results reveal associations of rs2237583 and rs3917490 in *PON1* and GxG interactions containing rs2237583 and rs17166879 with the susceptibility of NSCL±P in the Brazilian population. Furthermore, the study underlines the recent tendency of taking into account potential GxG interactions to clarify the underlying mechanisms associated with the etiology of this common malformation.

Keywords: nonsyndromic cleft lip with or without cleft palate, single nucleotide polymorphism, gene-gene interaction, *SOD1, SOD2*, *SOD3*, *PON1, PON2* and *PON3*.

Introduction

The facial development involves multiple interactions of several cell populations with a vast range of molecules associated with different signaling pathways. Disruption of those interactions, affecting proliferation, cell fate, apoptosis, cell-to-cell and cell-to-extracellular matrix adhesion or migration, are thought to underlie many of the craniofacial malformations such as the nonsyndromic oral clefts [1, 2]. Oxidation reactions during those interactions are essential for normal cell function, as are important the antioxidants to prevent the damage related to the subproducts of those processes. Indeed, the reactive oxygen species (ROS)-producing oxidative stress have long been considered as a mechanism of teratogenesis [3]. ROS avidly interact with a large number of molecules, including proteins, lipids, carbohydrates and nucleic acids, irreversibly destroying or altering their functions. Consequently, ROS have been increasingly identified as major contributors to damage in biological organisms [4, 5].

The enzymes encoded by the superoxide dismutase gene family, which includes *SOD1*, *SOD2* and *SOD3*, are the most important antioxidant defense system against ROS, particularly against the superoxide anion radicals [6]. All *SOD* family members are metal-dependents, and their loss of functions in teratogen-induced oxidative stress has been associated with dysmorphogenesis of whole embryos [7]. Paraoxonase (*PON*) genes, including *PON1*, *PON2* and *PON3*, possess antioxidant properties and is endowed of capacity of hydrolyze organophosphates and agrotoxic [8]. Although the participation of *SODs* and *PONs* on pathogenesis of nonsyndromic oral clefts is uncertain, polymorphic variants in *SOD1*, *SOD2* and *PON1* were associated with spina bifida [9, 10, 11]. Single-nucleotide polymorphisms (SNP) in *PON1* were also related to defects in hands and feet and cerebral cavernous malformations [12, 13]. Moreover, TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin)-induced cleft palate is associated with oxidative stress, which decreases motility and promotes apoptosis of medial edge epithelia cells, inhibiting palatal fusion [14]. Therefore, controlling of the oxidative stress, which may be intensified by environmental factors, is a critical step for normal development.

Due to the vast extent of the polygenic heritability, studies have applied more powerful strategies to unravel human complex traits, including two-stage analysis, which is based on a preliminary screening of several markers (first stage) followed by

a second (confirmatory) analysis on a subset of markers with a reduced multiple-testing adjustment. In the current study, based on family association tests (first stage), we initially evaluated the relationships of polymorphic variants in *SOD* and *PON* members with the risk of nonsyndromic cleft lip with or without cleft palate (NSCL±P) in the Brazilian population, assessing their roles as gene only, gene-gene (GxG) interactions and gene-environmental factor (GxE) interactions. The environmental factors included were maternal exposure to agrotoxics, environmental contact with agrotoxics, cigarette smoking, consumption of alcohol and drugs, and folic acid supplementation during the first trimester of gestation – all with putative participation on NSCL±P etiology [15, 16]. In the second stage, we analyzed the significant signals in an independent and robust case-control sample, which was stratified by genomic ancestry contribution.

Material and methods

Samples

This study included samples from patients with NSCL±P and healthy controls. We conducted an initial screening based in transmission disequilibrium test (TDT) analysis, followed by validation of the positive signals in a case-control approach. There was no overlapping between the samples used in the TDT and in the case-control study. The samples were collected in 4 different geographic regions of Brazil: the northeast region (University Hospital of Lauro Wanderley-HULW, João Pessoa-PB and Santo Antonio Hospital, Salvador-BA), the northern region (Hospital of Ophir Loyola, Belém-PA), the southeast region (Center for Rehabilitation of Craniofacial Anomalies, University of José Rosário Vellano, Alfenas-MG), and the southern region (Association of Carrier of Cleft Lip and Palate-APOFILAB, Cascavel-PR). All patients with NSCL±P were carefully examined and screened for the presence of associated anomalies or syndromes by the specialized teams of the associated centers or hospitals, and only patients with nonsyndromic form of CL±P, classification taking as reference the incisive foramen [17], were included. The unaffected individuals were healthy subjects without congenital malformations or psychiatric diseases. Individuals of the control group for case-control analysis had also no familial history of orofacial clefts and were born in the same geographic areas.

In the TDT analysis, 325 trios composed of one affected offspring and two healthy parents (227 with nonsyndromic cleft lip and palate [NSCLP] and 98 with

nonsyndromic cleft lip only [NSCLO]) were included. Information on maternal exposures, including exposure to agrotoxics, environmental contact with agrotoxics, cigarette smoking, and consumption of alcohol and drugs during the first trimester of gestation, were collected through direct interview of the mothers. The mothers were also asked whether they used folic acid supplementation before the conception and during the first trimester of gestation (Supplementary Table 1). The average age of patients with NSCL±P was 11.47 ± 8.63 years. For the case-control analysis, 722 patients affected with NSCL±P (515 NSCLP and 207 NSCLO) and 866 healthy individuals were included. The average age of the patients with NSCL±P was 18.49 ± 13.84 years, and the average age of the control group was 25.24 ± 11.79. All volunteers included in this research provided written informed consent, and the study was approved by the ethics review board of each of the centers or hospitals affiliated with the collaborative study.

Supplementary Table 1. Characteristics of patients included in the transmission disequilibrium test (family-based analysis) and frequencies of maternal exposures to environment factors during the first trimester of gestation.

	No. of patients	%
Gender		
Male	205	63.1
Female	120	36.9
Type of cleft		
Nonsyndromic cleft lip only	98	30.1
Nonsyndromic cleft lip and palate	227	69.9
Maternal exposure		
Agrotoxics	14	4.3
Environment contact with agrotoxics	31	9.5
Cigarette smoking	39	12.0
Alcohol consumption	22	6.7
Drug consumption	52	16.0
Folic acid absence	127	39,1

Definitions:

Agrotoxics: defined as mother that had direct contact with agrotoxic by work with those products on day basis, regardless frequency, during the first trimester of gestation.

Environment contact with agrotoxics: defined as mother that did not have direct contact with agrotoxic, but worked with agricultural activities during the first trimester of gestation and it is possible some indirect contact.
Cigarette smoking: mothers that smoked during the first trimester of gestation, regardless frequency.
Alcohol consumption: defined as mother that undertook alcoholic beverages during the first trimester of gestation, regardless amount and type.
Drug consumption: defined as mother that took medicines, such as corticosteroids, antibiotics or anticonvulsants (barbiturates or benzodiazepines), in the first trimester of gestation.
Folic acid absence: defined as mother that not took folic acid before conception and during the first trimester of gestation.

Genotyping and genomic ancestry assessment

The genomic DNA was isolated from oral mucosa cells obtained by mouthwash with a 3% sucrose solution using a salting-out protocol [18]. Twenty-eight SNP (Supplementary Table 2) were genotyped on the StepOne Real-Time PCR system with TaqMan 5'-exonuclease allelic discrimination assays (Assay-on-Demand service, Applied Biosystems). Genotyping analyses were randomly repeated in 10% of the samples for each SNP, and the concordance rate was 100%. All samples were successfully genotyped, with a genotype call rate of >97.7%.

To determinate the genomic ancestry of subjects, samples were genotyped for a panel containing 40 biallelic short insertion-deletion polymorphisms as previously described [19].

Supplementary Table 2. Characteristics of the single-nucleotide polymorphisms (SNP) in the genes *SOD1*, *SOD2*, *SOD3*, *PON1, PON2* and *PON3*.

Gene	SNP	Location	Position	MAF	Alleles
SOD1 (21q22.11)					
	rs1041740	Intron	31,667,849	0.242	C/**T**
	rs4817420	Intron	31,668,058	0.242	C/**T**
SOD2 (6q25.3)					
	rs8031	3' UTR	159,679,608	0.367	A/**T**

	rs2758331	Intron	159,684,038	0.334	C/**A**
	rs2855116	Intron	159,685,093	0.334	A/**C**
	rs4880	5' UTR	159,692,840	0.410	A/**G**
SOD3 (4p15.3)					
	rs699473	5' UTR	24,795,181	0.441	C/**T**
PON1 (7q21.3)					
	rs854552	3' UTR	95,298,612	0.366	T/**C**
	rs3917551	Intron	95,305,143	0.127	G/**A**
	rs3917548	Intron	95,306,593	0.133	A/**G**
	rs662	Exon (Gln192Arg)	95,308,134	0.457	C/**T**
	rs3917527	Intron	95,310,946	0.125	T/**C**
	rs2301711	Intron	95,316,347	0.182	T/**C**
	rs3917490	Intron	95,319,529	0.347	T/**C**
	rs2299262	Intron	95,320,616	0.440	C/**T**
	rs2237583	Intron	95,320,865	0.301	C/**T**
	rs705379	5' UTR	95,324,583	0.349	G/**A**
PON2 (7q21.3)					
	rs7785846	3' UTR	95,404,529	0.282	C/**T**
	rs7786401	3' UTR	95,404,689	0.271	G/**T**
	rs7493	Exon (Ser311Cys)	95,405,463	0.283	G/**C**
	rs2286232	Intron	95,410,133	0.282	C/**T**
	rs2299263	Intron	95,411,099	0.282	C/**T**
	rs17166879	Intron	95,424,103	0.247	A/**G**
	rs2299267	Intron	95,432,609	0.211	A/**G**
PON3 (7q21.3)					
	rs10953143	Intron	95,362,932	0.469	C/**T**
	rs3757708	Intron	95,367,601	0.360	G/**T**
	rs1053275	Exon (Ala99Ala)	95,372,243	0.339	T/**C**
	rs978903	Intron	95,374,855	0.339	G/**A**

Risk alleles in bold. Source: http://www.ncbi.nlm.nih.gov/pubmed/.

Statistical analysis

Genotype deviation from Hardy-Weinberg equilibrium was assessed through x^2 test in the healthy parents of the trio sample, and in the control group of the case-control sample. A p value ≤0.05 was indicative of derivation. The TDT analyses, including allele testing and GxG and GxE interactions, were performed using the TRIO Package in RStudio software (version 3.0.0). The additive conditional logistic models were previously described [20, 21]. The full models containing genotype only (G), GxG and GxE to a null model without containing any terms were compared by the likelihood ratio test (LRT). The 2 degree of freedom (df) test was for G and GxG interactions or G and GxE interactions togethers, and 1df test was for GxG and GxE interactions alone. The haplotype-based analysis was conducted using the HBAT function of the FBAT software [22].

For the case-control analysis, the multiple logistic regression analysis under unrestricted, dominant and recessive genetic models, considering gender and genomic ancestry as potential confounders, was performed with SNPassoc and Haplo.stats packages in RStudio software. The genomic ancestry of each individual of the case-control study was determined with the Structure software version 2.3.4 [23], applying the model of K=3 for parental populations based on the tri-hybrid origin of the Brazilian population. GxG interactions were performed using the model-based multifactor dimensionality reduction (mbmdr) package in RStudio software, using the 10,000 permutation test to eliminate the false-positive interactions.

In TDT analysis a Bonferroni threshold level of α ≤0.001 (28 SNP) was adopted, but for the case-control study, the number of SNP was reduced to 10, then the Bonferroni threshold was adjusted to α≤0.005.

Results

The distributions of genotypes in the healthy parents (control) did not deviate from expectation based on the Hardy-Weinberg equilibrium, with the exception of rs8031 (p=0.04, Table 1). Although no individual SNP were preferentially transmitted from healthy parents to NSCL±P offspring (Table 1) and GxE interaction analysis yielded only nominally significant signals, which did not remain significant after Bonferroni

correction (Supplementary Table 3), signals of GxG interactions were identified (Table 2). Multiple pairwise SNP-SNP interactions were identified, but only 14 of them, which included SNP in *PON1* (rs3917490, rs2299262 and rs2237583), *PON2* (rs2286232, rs17166879 and rs2299267) and *PON3* (rs10953143, rs3757708, rs1053275 and rs978903), showed p values resistant to Bonferroni adjustment for multiple testing and 10,000 permutation tests (Table 2). Haplotype-based analyses revealed no *SOD1*, *SOD2*, *SOD3*, *PON1*, *PON2* and *PON3* haplotype-phenotype association (Supplementary Table 4).

Based on these results that revealed intriguing pairwise GxG interaction signals in 10 SNP located in *PON1*, *PON2* and *PON3* with NSCL±P risk, we performed further analyses in a larger case-control sample. As the genetic basis of susceptibility to NSCL±P varies among different populations [15], and the population of Brazil is highly admixed, with each individual showing variable ancestry proportions of Amerindians, Europeans and sub-Saharan Africans, we determined the proportion of genomic ancestry of each patient of the case-control sample. The European ancestry contribution was the most predominant in both NSCL±P and control groups, followed by African and Amerindian (Supplementary Fig. 1). Regarding gender, there was a significant difference between NSCL±P and controls, with NSCL±P being overrepresented by males (Supplementary Table 5). Therefore, the OR with 95% CI were estimated by multiple logistic regression models after controlling for these confounders in each comparison.

The distributions of genotypes in the control group were consistent with those predicted by the Hardy-Weinberg equilibrium (Table 3). The frequency of the rs2237583 C allele was significantly lower in the NSCL±P patients compared with the control group, yielding an OR of 0.79 (95% CI: 0.67-0.93, p=0.005). The CT genotype (heterozygosis) was more frequent in the control group than in the NSCL±P group (p=0.01), revealing a decreased risk to NSCL±P (OR: 0.77, 95% CI: 0.62-0.95). In the dominant genetic model, rs2237583 SNP also demonstrated a significant association, with an OR of 0.75 (95% CI: 0.61-0.92, p=0.005). The conventional analysis for SNP effect alone also revealed an association between the C allele of rs3917490 and NSCL±P (OR: 0.86, 95% CI: 0.75-0.99, p=0.04), but this association did not withstand the p value of Bonferroni correction for multiple tests.

Although haplotype association analysis showed an association between NSCL±P and T-C-C haplotype formed by rs3917490, rs2299262 and rs2237583 in *PON1* with a nominal p value (p=0.009), only the T-T-T-A haplotype in *PON3* involving the SNP rs10953143, rs3757708, rs1053275 and rs978903 demonstrated an association that remains significant after Bonferroni correction for multiple tests (p=0.004, Supplementary Table 6).

As the proportion of African genomic ancestry of the Brazilian population affects the genetic susceptibility to NSCL±P [19, 24, 25, 26, 27, 28], and the frequency of the risk alleles of some of selected SNP differ according to ethnicity, specially between African and Caucasian populations, we performed similar analyses in the subgroups of samples showing high European ancestry and high African ancestry. In the subgroup with high European ancestry, the average of European ancestry was higher than 90% in both groups (Supplementary Fig. 2), and the subgroup with high African ancestry showed an average African ancestry of 37.2% in the control group and 38.9% in the NSCL±P group (Supplementary Fig. 3). In the group with high African ancestry, rs2237583 showed associations with p values that did not withstand Bonferroni correction for multiple tests, whereas significant associations, which resisted to multiple comparisons, involving rs3917490 were found (Table 4). The frequency of *PON1* rs3917490 C allele was significantly higher in the control group versus the NSCL±P group, with the OR of 0.64 (95% CI: 0.48-0.84, p=0.001). The genotype frequencies of rs3917430 were also associated with a decreased risk of NSCL±P in the subgroup with high African ancestry, reaching the maximum effect in the homozygous state (OR: 0.44, 95% CI: 0.25-0.78, p=0.004). No significant associations in the subgroup with high European ancestry were observed (Table 5).

The mbmdr method was applied for detecting interactions between SNP and disease risk in the case-control sample. This method merges multilocus genotypes into a one-dimensional construct, categorizing the genotypic interactions into high-risk and low-risk level for the phenotype. Table 6 summarized the significant results obtained in 2- and 3-SNP analysis. Among the 2-SNP pairs, rs2237583 (*PON1*) and rs17166879 (*PON2*) yielded the most significant signal (p=0.001) and it was the only one that remains significant after permutation test. At 3-SNP, the interactions containing rs2237583 (*PON1*) and rs17166879 (*PON2*) [rs2237583 - rs17166879 - rs978903

(*PON3*), rs2237583 - rs17166879 - rs1053275 (*PON3*), and rs2237583 - rs17166879 - rs3757708 (*PON3*)] were the only significant after the permutation test.

Table 1. Allelic transmission disequilibrium test of the single-nucleotide polymorphisms (SNP) in the sample containing probands with nonsyndromic cleft lip with or without cleft palate (NSCL±P) and healthy parents (trios).

SNP	HWE (*P* value)	MAF	Number of families	T/NT	x^2	OR (95% CI)	*P* value
rs1041740	0.81	0.270	203	120/83	2.10	0.83 (0.64-1.06)	0.14
rs4817420	0.87	0.286	200	116/84	2.29	0.82 (0.64-1.05)	0.12
rs8031	0.04	0.476	244	182/62	1.38	1.14 (0.91-1.42)	0.24
rs2758331	0.24	0.462	248	175/73	0.44	1.07 (0.86-1.33)	0.50
rs2855116	0.67	0.456	243	169/74	0.38	1.07 (0.85-1.34)	0.53
rs4880	0.06	0.469	251	176/75	3.09	0.82 (0.66-1.02)	0.07
rs699473	0.66	0.468	231	158/73	3.35	0.80 (0.64-1.01)	0.06
rs854552	0.43	0.332	218	141/77	0.09	1.03 (0.81-1.31)	0.76
rs3917551	0.69	0.111	99	44/55	3.47	0.70 (0.48-1.01)	0.06
rs3917548	0.82	0.119	110	52/58	2.22	0.76 (0.54-1.08)	0.13
rs662	0.25	0.405	231	158/73	1.23	0.87 (0.69-1.10)	0.26
rs3917527	0.55	0.110	101	49/52	1.21	0.81 (0.56-1.17)	0.27
rs2301711	0.08	0.156	132	67/65	0.64	0.87 (0.64-1.20)	0.42
rs3917490	0.91	0.463	236	172/64	0.73	1.10 (0.88-1.38)	0.39
rs2299262	0.11	0.356	225	151/74	0.00	1.00 (0.79-1.26)	1.00
rs2237583	0.83	0.240	172	109/63	0.11	0.95 (0.72-1.24)	0.73
rs705379	0.44	0.422	244	162/82	0.44	0.92 (0.74-1.15)	0.50
rs7785846	0.85	0.209	174	99/75	0.52	0.90 (0.70-1.17)	0.46
rs7786401	0.92	0.222	170	95/75	0.55	0.90 (0.69-1.17)	0.45
rs7493	0.98	0.226	169	94/75	1.62	0.84 (0.64-1.09)	0.20
rs2286232	0.73	0.232	177	98/79	0.62	0.90 (0.69-1.16)	0.43
rs2299263	0.82	0.228	173	96/77	1.14	0.86 (0.66-1.12)	0.28
rs17166879	0.79	0.215	169	96/73	0.29	0.92 (0.71-1.21)	0.58
rs2299267	0.71	0.171	148	82/66	0.95	0.86 (0.64-1.16)	0.32
rs10953143	0.83	0.433	227	160/67	0.76	0.90 (0.71-1.13)	0.38
rs3757708	0.80	0.490	230	174/56	0.01	0.98 (0.78-1.23)	0.90

rs1053275	0.58	0.459	229	170/59	0.08	0.96 (0.77-1.21)	0.77
rs978903	0.51	0.471	232	175/57	0.003	1.00 (0.80-1.25)	0.95

HWE: Hardy-Weinberg equilibrium; MAF: minor allele frequency; T/NT: transmission/non-transmission counts.

Table 2. Gene-gene interactions between the single-nucleotide polymorphisms (SNP) in *SOD1, SOD2*, *SOD3*, *PON1, PON2* and *PON3*, using epistatic interactions analysis implemented in TRIO software.

SNP1	SNP2	*P* value	Perm. *P* value
rs17166879 (*PON2*)	rs10953143 (*PON3*)	$1.7x10^{-12}$	$<1x10^{-11}$
rs17166879 (*PON2*)	rs3757708 (*PON3*)	$1.1x10^{-7}$	$<1x10^{-11}$
rs3917490 (*PON1*)	rs1053275 (*PON3*)	$1.7x10^{-7}$	$<1x10^{-11}$
rs2286232 (*PON2*)	rs10953143 (*PON3*)	$1.9x10^{-7}$	$<1x10^{-11}$
rs2286232 (*PON2*)	rs1053275 (*PON3*)	$2.2x10^{-7}$	$<1x10^{-11}$
rs17166879 (*PON2*)	rs978903 (*PON3*)	$4.1x10^{-7}$	$<1x10^{-11}$
rs3917490 (*PON1*)	rs978903 (*PON3*)	$5.7x10^{-7}$	$<1x10^{-11}$
rs2286232 (*PON2*)	rs3757708 (*PON3*)	$1.2x10^{-6}$	$<1x10^{-11}$
rs3917490 (*PON1*)	rs3757708 (*PON3*)	$2.4x10^{-6}$	$<1x10^{-11}$
rs2237583 (*PON1*)	rs2299267 (*PON2*)	$3.9x10^{-6}$	$<1x10^{-11}$
rs2299262 (*PON1*)	rs10953143 (*PON3*)	$3x10^{-5}$	$5x10^{-4}$
rs2299262 (*PON1*)	rs3757708 (*PON3*)	0.001	0.001
rs3917490 (*PON1*)	rs10953143 (*PON3*)	0.001	0.001
rs2299262 (*PON1*)	rs2299267 (*PON2*)	0.004	0.001

Perm. *P* value: p value based on 10,000 permutations.

Table 3. Association of single nucleotide polymorphisms (SNP) in the case-control study. P values were adjusted for confounders by logistic regression analysis.

	HWE *P* value	Control (%)	NSCL±P (%)	OR_{allele} (95% CI) *P* value	OR_{Het} (95% CI) *P* value	OR_{Hom} (95% CI) *P* value	OR_{Dom} (95% CI) *P* value	OR_{Rec} (95% CI) *P* value
rs3917490 (TT/TC/CC)	0.95	28.3/49.9/21.8	33.1/47.4/19.5	0.86 (0.75-0.99) 0.04	0.83 (0.66-1.05) 0.10	0.79 (0.59-1.05) 0.09	0.82 (0.66-1.02) 0.06	0.88 (0.69-1.13) 0.32
rs2299262 (CC/CT/TT)	0.06	40.0/48.8/11.2	42.9/46.6/10.5	0.92 (0.79-1.06) 0.28	0.88 (0.71-1.09) 0.22	0.86 (0.61-1.22) 0.40	0.88 (0.71-1.07) 0.19	0.93 (0.67-1.28) 0.64
rs2237583 (CC/CT/TT)	0.54	54.8/39.0/6.2	61.6/33.7/4.7	0.79 (0.67-0.93) 0.005	0.77 (0.62-0.95) 0.01	0.66 (0.42-1.04) 0.06	0.75 (0.61-0.92) 0.005	0.73 (0.47-1.15) 0.17
rs2286232 (CC/CT/TT)	0.15	59.9/33.9/6.2	61.8/32.8/5.4	0.92 (0.78-1.09) 0.36	0.96 (0.77-1.18) 0.66	0.87 (0.56-1.34) 0.51	0.94 (0.77-1.16) 0.56	0.88 (0.57-1.35) 0.56
rs17166879 (AA/AG/GG)	0.21	62.5/32.2/5.3	66.9/28.0/5.1	0.86 (0.72-1.03) 0.10	0.83 (0.67-1.04) 0.10	0.92 (0.58-1.45) 0.70	0.84 (0.69-1.04) 0.11	0.98 (0.62-1.53) 0.91
rs2299267 (AA/AG/GG)	0.83	70.1/27.1/2.8	69.5/27.6/2.9	1.02 (0.84-1.23) 0.79	1.00 (0.80-1.26) 0.97	1.08 (0.59-1.97) 0.77	1.01 (0.81-1.26) 0.92	1.08 (0.59-1.96) 0.81
rs10953143 (TT/TC/CC)	0.07	35.5/45.6/18.9	38.6/43.1/18.3	0.92 (0.80-1.06) 0.27	0.86 (0.69-1.07) 0.17	0.89 (0.67-1.18) 0.41	0.87 (0.71-1.07) 0.18	0.97 (0.75-1.25) 0.80
rs3757708 (GG/GT/TT)	0.05	28.5/46.6/24.9	29.1/47.2/23.7	0.96 (0.83-1.10) 0.60	0.99 (0.79-1.26) 0.95	0.94 (0.71-1.24) 0.61	0.98 (0.78-1.22) 0.82	0.94 (0.75-1.19) 0.61
rs1053275 (TT/TC/CC)	0.05	30.6/46.3/23.1	32.7/46.3/21.1	0.92 (0.79-1.05) 0.24	0.94 (0.75-1.18) 0.59	0.86 (0.66-1.14) 0.28	0.92 (0.74-1.13) 0.41	0.90 (0.70-1.14) 0.37
rs978903 (GG/GA/AA)	0.07	30.5/46.6/22.9	32.5/47.1/20.4	0.91 (0.79-1.04) 0.19	0.95 (0.76-1.20) 0.65	0.84 (0.63-1.11) 0.20	0.91 (0.74-1.13) 0.41	0.87 (0.68-1.10) 0.24

HWE: Hardy-Weinberg equilibrium; NSCL±P: nonsyndromic cleft lip with or without cleft palate.

Table 4. Association of single-nucleotide polymorphisms in the case-control sample with high African genomic ancestry. P values were adjusted for confounders by logistic regression analysis.

	HWE	Control (%)	NSCL±P (%)	OR_{allele} (95% CI)	OR_{Het} (95% CI)	OR_{Hom} (95% CI)	OR_{Dom} (95% CI)	OR_{Rec} (95% CI)
	P value			*P* value	*P* value	*P* value	*P* value	*P* value
rs3917490 (TT/TC/CC)	0.49	23.8/52.4/23.8	37.4/47.1/15.5	0.64 (0.48-0.84)	0.60 (0.38-0.94)	0.44 (0.25-0.78)	0.55 (0.35-0.84)	0.60 (0.37-0.99)
				0.001	0.02	0.004	0.005	0.04
rs2299262 (CC/CT/TT)	0.10	38.1/51.4/10.5	39.8/49.5/10.7	0.96 (0.72-1.28)	0.90 (0.59-1.36)	0.94 (0.48-1.86)	0.91 (0.61-1.35)	1.00 (0.53-1.89)
				0.82	0.62	0.88	0.62	0.99
rs2237583 (CC/CT/TT)	0.23	55.7/35.7/8.6	63.6/32.5/3.9	0.70 (0.50-0.97)	0.78 (0.51-1.18)	0.38 (0.16-0.92)	0.70 (0.47-1.05)	0.42 (0.18-1.00)
				0.03	0.24	0.02	0.08	0.04
rs2286232 (CC/CT/TT)	0.06	61.4/31.0/7.6	61.2/34.4/4.4	0.91 (0.66-1.27)	1.17 (0.77-1.79)	0.55 (0.23-1.30)	1.04 (0.70-1.55)	0.52 (0.22-1.22)
				0.60	0.47	0.15	0.84	0.12
rs17166879 (AA/AG/GG)	0.45	63.3/31.4/5.3	67.9/27.7/4.4	0.83 (0.59-1.18)	0.87 (0.56-1.34)	0.73 (0.29-1.84)	0.85 (0.56-1.28)	0.76 (0.30-1.91)
				0.31	0.55	0.47	0.42	0.56
rs2299267 (AA/AG/GG)	0.40	69.1/29.0/1.9	70.4/28.2/1.5	0.93 (0.64-1.35)	0.89 (0.57-1.37)	0.85 (0.18-3.94)	0.88 (0.58-1.35)	0.88 (0.19-4.05)
				0.72	0.59	0.83	0.56	0.86
rs10953143 (TT/TC/CC)	0.32	37.6/44.8/17.6	39.3/43.7/17.0	0.95 (0.72-1.25)	0.88 (0.57-1.35)	0.91 (0.52-1.60)	0.89 (0.59-1.32)	0.98 (0.58-1.64)
				0.73	0.52	0.74	0.55	0.92
rs3757708 (GG/GT/TT)	0.17	29.5/45.2/25.2	32.0/48.1/19.9	0.85 (0.64-1.12)	0.94 (0.60-1.49)	0.74 (0.43-1.28)	0.87 (0.57-1.33)	0.77 (0.48-1.23)
				0.25	0.80	0.28	0.53	0.27
rs1053275 (TT/TC/CC)	0.08	31.9/43.8/24.3	37.4/46.1/16.5	0.76 (0.57-1.00)	0.87 (0.56-1.36)	0.59 (0.34-1.03)	0.78 (0.51-1.17)	0.64 (0.39-1.05)
				0.05	0.55	0.06	0.22	0.07
rs978903 (GG/GA/AA)	0.11	31.4/44.3/24.3	35.9/47.1/17.0	0.78 (0.59-1.03)	0.90 (0.58-1.40)	0.62 (0.36-1.08)	0.80 (0.53-1.21)	0.66 (0.41-1.08)
				0.08	0.64	0.09	0.29	0.09

HWE: Hardy-Weinberg equilibrium; NSCL±P: nonsyndromic cleft lip with or without cleft palate.

Table 5. Association of single-nucleotide polymorphisms in the case-control sample with high European origin. P values were adjusted for confounders by logistic regression analysis.

	HWE	Control (%)	NSCL±P (%)	OR_{allele} (95% CI)	OR_{Het} (95% CI)	OR_{Hom} (95% CI)	OR_{Dom} (95% CI)	OR_{Rec} (95% CI)
	P value			*P* value	*P* value	*P* value	*P* value	*P* value
rs3917490 (TT/TC/CC)	0.77	29.7/49.1/21.2	31.4/47.5/21.1	0.96 (0.81-1.13)	0.92 (0.70-1.20)	0.94 (0.68-1.31)	0.93 (0.72-1.19)	0.99 (0.75-1.32)
				0.67	0.52	0.73	0.55	0.97
rs2299262 (CC/CT/TT)	0.20	40.5/48.0/11.5	44.2/45.3/10.5	0.90 (0.76-1.07)	0.87 (0.68-1.11)	0.85 (0.58-1.27)	0.87 (0.69-1.10)	0.92 (0.63-1.33)
				0.24	0.27	0.43	0.23	0.65
rs2237583 (CC/CT/TT)	0.16	54.4/40.1/5.5	60.9/34.1/5.0	0.82 (0.68-1.00)	0.76 (0.60-0.97)	0.85 (0.50-1.44)	0.77 (0.61-0.98)	0.94 (0.56-1.59)
				0.05	0.02	0.53	0.03	0.82
rs2286232 (CC/CT/TT)	0.54	59.4/34.8/5.8	62.0/32.2/5.8	0.92 (0.74-1.13)	0.89 (0.70-1.15)	0.99 (0.60-1.64)	0.91 (0.72-1.15)	1.03 (0.63-1.69)
				0.46	0.37	0.93	0.42	0.90
rs17166879 (AA/AG/GG)	0.30	62.2/32.5/5.3	66.5/28.1/5.4	0.87 (0.71-1.07)	0.82 (0.63-1.05)	0.97 (0.58-1.63)	0.84 (0.66-1.07)	1.03 (0.62-1.73)
				0.21	0.11	0.89	0.14	0.89
rs2299267 (AA/AG/GG)	0.46	70.4/26.5/3.1	69.2/27.3/3.5	1.06 (0.85-1.32)	1.06 (0.81-1.37)	1.19 (0.62-2.28)	1.07 (0.83-1.38)	1.17 (0.61-2.24)
				0.58	0.68	0.61	0.60	0.63
rs10953143 (TT/TC/CC)	0.12	34.7/45.9/19.4	38.4/42.8/18.8	0.91 (0.77-1.08)	0.86 (0.66-1.11)	0.89 (0.64-1.23)	0.87 (0.68-1.10)	0.97 (0.72-1.30)
				0.30	0.24	0.47	0.24	0.82
rs3757708 (GG/GT/TT)	0.12	28.2/47.0/24.8	27.9/46.9/25.2	1.01 (0.86-1.19)	1.00 (0.76-1.32)	1.01 (0.73-1.39)	1.00 (0.77-1.30)	1.01 (0.77-1.32)
				0.87	0.99	0.96	0.98	0.94
rs1053275 (TT/TC/CC)	0.17	30.2/47.1/22.7	30.8/46.3/22.9	0.99 (0.84-1.16)	0.95 (0.73-1.25)	0.97 (0.70-1.34)	0.96 (0.75-1.23)	1.00 (0.76-1.31)
				0.90	0.72	0.86	0.74	0.98
rs978903 (GG/GA/AA)	0.23	30.2/47.4/22.4	31.2/47.1/21.7	0.96 (0.82-1.13)	0.95 (0.73-1.25)	0.92 (0.67-1.27)	0.94 (0.73-1.21)	0.95 (0.72-1.25)
				0.67	0.72	0.61	0.64	0.70

HWE: Hardy-Weinberg equilibrium; NSCL±P: nonsyndromic cleft lip with or without cleft palate.

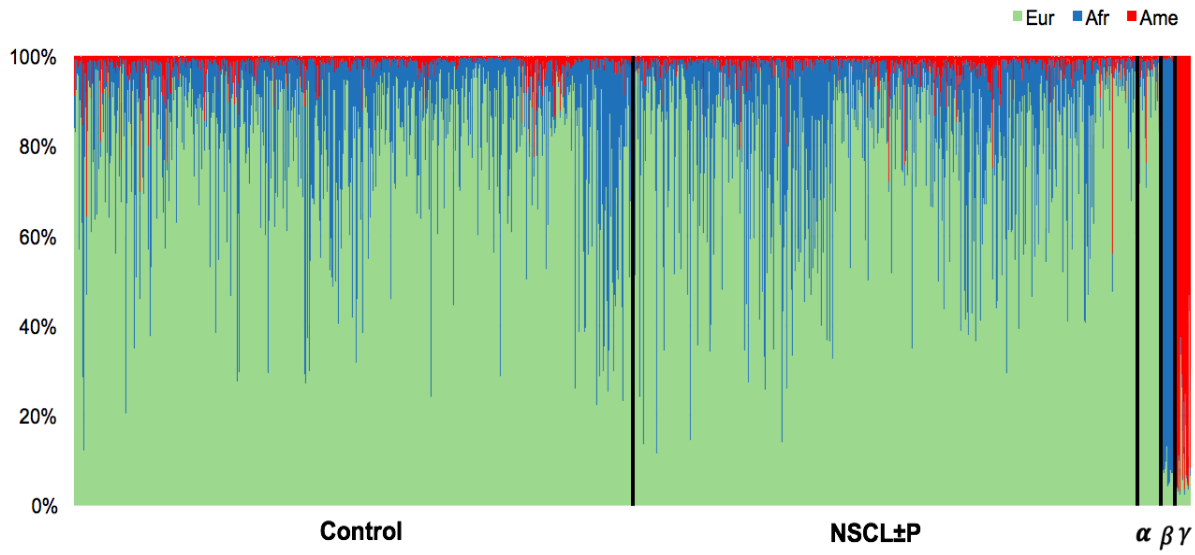
Table 6. Gene-gene interactions in patients with nonsyndromic cleft lip with or without cleft palate (NSCL±P) assessed by model-based multifactor dimensionality reduction method.

	SNP1	SNP2	SNP3	NH[a]	betaH[b]	NL[c]	betaL[d]	*P* value[e]	Perm. *P* value[f]
2-SNP	rs2237583 (*PON1*)	rs17166879 (*PON2*)		1	0.33	0	NA	0.001	0.01
	rs10953143 (*PON3*)	rs3917490 (*PON1*)		1	0.36	2	-0.32	0.003	0.05
	rs2237583 (*PON1*)	rs2286232 (*PON2*)		1	0.29	1	-0.21	0.006	0.06
	rs2286232 (*PON2*)	rs3917490 (*PON1*)		0	NA	1	-0.56	0.007	0.07
	rs17166879 (*PON2*)	rs3917490 (*PON1*)		1	0.21	1	-0.58	0.008	0.10
3-SNP	rs2237583 (*PON1*)	rs17166879 (*PON2*)	rs978903 (*PON3*)	2	0.49	0	NA	1.07×10^{-4}	0.01
	rs2237583 (*PON1*)	rs17166879 (*PON2*)	rs1053275 (*PON3*)	2	0.49	1	-0.36	1.14×10^{-4}	0.01
	rs2237583 (*PON1*)	rs17166879 (*PON2*)	rs3757708 (*PON3*)	2	0.48	1	-0.45	2.63×10^{-4}	0.02
	rs2237583 (*PON1*)	rs2286232 (*PON2*)	rs978903 (*PON3*)	2	0.44	0	NA	9.53×10^{-4}	0.06
	rs2237583 (*PON1*)	rs2286232 (*PON2*)	rs1053275 (*PON3*)	2	0.43	0	NA	0.001	0.07
	rs978903 (*PON3*)	rs17166879 (*PON2*)	rs2299262 (*PON1*)	2	0.43	0	NA	0.001	0.12
	rs1053275 (*PON3*)	rs17166879 (*PON2*)	rs2299262 (*PON1*)	2	0.43	1	-0.30	0.001	0.12
	rs10953143 (*PON3*)	rs2299267 (*PON2*)	rs2299262 (*PON1*)	2	0.86	0	NA	0.001	0.12
	rs10953143 (*PON3*)	rs2299267 (*PON2*)	rs2237583 (*PON1*)	2	0.75	0	NA	0.001	0.10
	rs3757708 (*PON3*)	rs2286232 (*PON2*)	rs2237583 (*PON1*)	2	0.43	0	NA	0.002	0.10
	rs10953143 (*PON3*)	rs2286232 (*PON2*)	rs3917490 (*PON1*)	1	0.54	2	-0.42	0.002	0.17
	rs1053275 (*PON3*)	rs2286232 (*PON2*)	rs2299262 (*PON1*)	1	0.86	0	NA	0.002	0.17
	rs3757708 (*PON3*)	rs17166879 (*PON2*)	rs2299262 (*PON1*)	2	0.42	1	-0.33	0.003	0.19
	rs978903 (*PON3*)	rs2299267 (*PON2*)	rs3917490 (*PON1*)	2	0.44	0	NA	0.003	0.20
	rs3757708 (*PON3*)	rs17166879 (*PON2*)	rs3917490 (*PON1*)	1	0.74	1	-0.77	0.004	0.19
	rs978903 (*PON3*)	rs17166879 (*PON2*)	rs3917490 (*PON1*)	2	0.39	1	-0.75	0.005	0.23

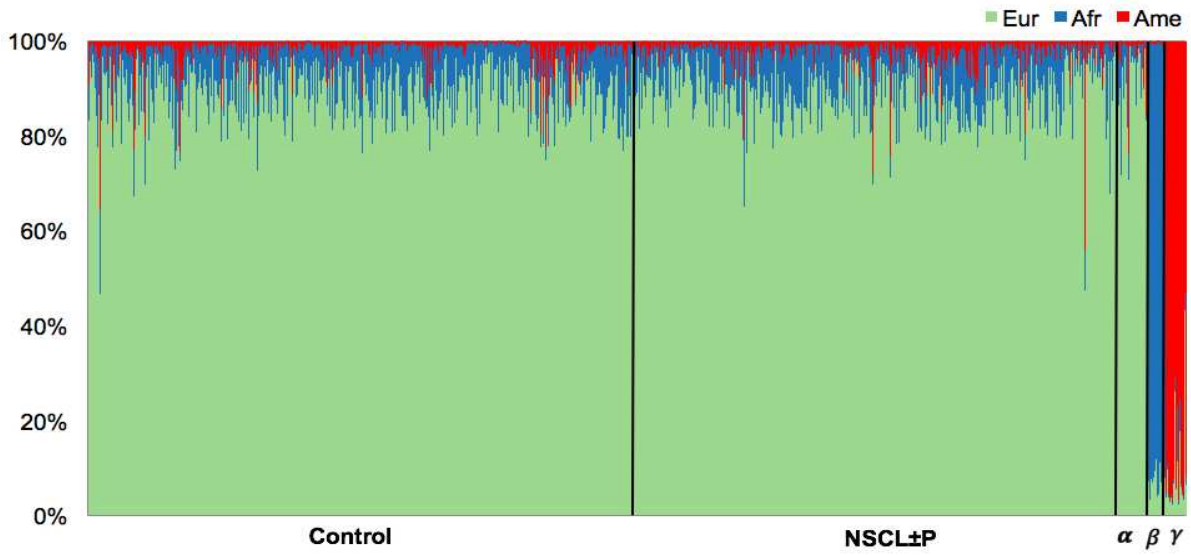
	rs1053275 (*PON3*)	rs17166879 (*PON2*)	rs3917490 (*PON1*)	2	0.42	1	-0.75	0.005	0.24
	rs10953143 (*PON3*)	rs2299267 (*PON2*)	rs3917490 (*PON1*)	1	0.41	2	-0.39	0.007	0.33
	rs978903 (*PON3*)	rs2286232 (*PON2*)	rs2299262 (*PON1*)	1	0.79	0	NA	0.008	0.30

[a]Number of significant high-risk genotypes in the interaction. [b]Regresion coeficient in step2 for high-risk exposition. [c]Number of significant low-risk genotypes in the interaction. [d]Regresion coeficient in step2 for low-risk exposition. [e]*P* value for the interaction model adjusted for covariates. [f]Permutation *P* value for the interaction model.

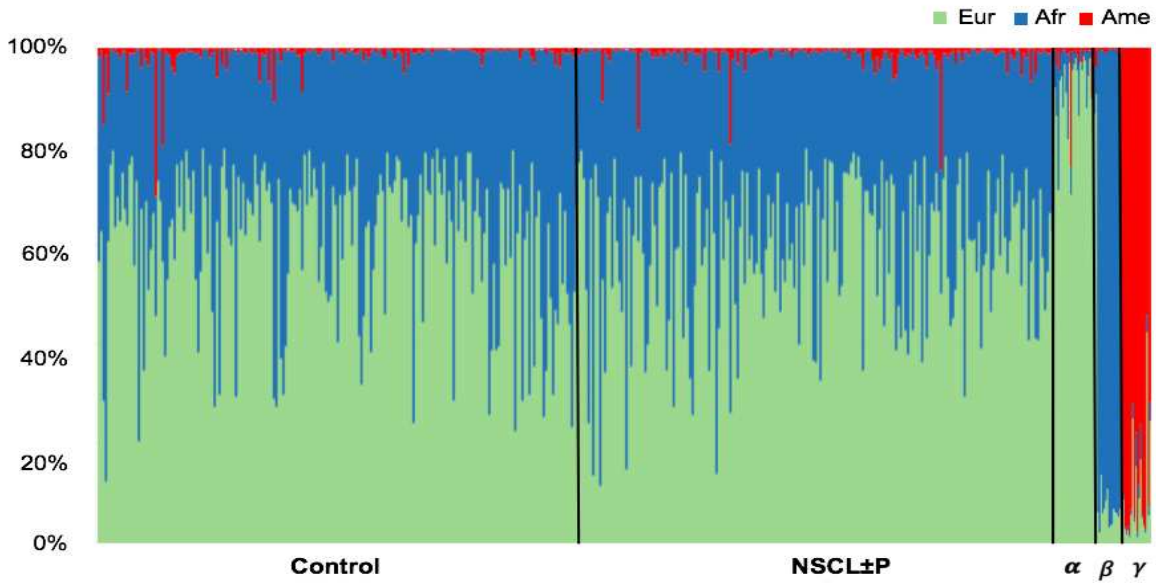
Supplementary Figure 1. Ancestry contributions of individuals of control and NSCL±P groups. Each column is represented by a single individual. European, African and Amerindian parental samples were used as learning samples to assist the Structure software in estimating ancestry of the admixed individuals.



Supplementary Figure 2. Ancestry frequencies of individuals of control and NSCL±P groups displaying high proportion of European ancestry.



Supplementary Figure 3. Ancestry frequencies of individuals of control and NSCL±P groups displaying high proportion of African ancestry.



Supplementary Table 3. Gene-environmental factor interactions between single-nucleotide polymorphisms (SNP) in *SOD1, SOD2*, *SOD3*, *PON1, PON2* and *PON3* and maternal exposures during the first trimester of gestation.

A. Interactions with maternal contact with agrotoxics.

		Agrotoxics			
		Non-Exposure OR (95% CI)	Exposure OR (95% CI)	*P* value (1 df)	*P* value (2 df)
SOD1					
	rs1041740	0.82 (0.64-1.06)	1.00 (0.28-3.45)	0.76	0.33
	rs4817420	0.81 (0.62-1.04)	1.20 (0.36-3.93)	0.52	0.25
SOD2					
	rs8031	1.11 (0.89-1.39)	2.33 (0.60-9.02)	0.27	0.27
	rs2758331	1.05 (0.84-1.31)	2.33 (0.60-9.02)	0.23	0.39
	rs2855116	1.04 (0.83-1.31)	2.33 (0.60-9.02)	0.23	0.40
	rs4880	0.84 (0.67-1.05)	0.37 (0.09-1.41)	0.21	0.09
SOD3					
	rs699473	0.81 (0.64-1.02)	0.66 (0.18-2.36)	0.75	0.17
PON1					
	rs854552	1.03 (0.80-1.31)	1.16 (0.39-3.47)	0.82	0.93
	rs3917551	0.70 (0.48-1.03)	0.66 (0.11-3.99)	0.95	0.17
	rs3917548	0.74 (0.52-1.06)	1.50 (0.25-8.97)	0.44	0.24
	rs662	0.86 (0.68-1.08)	1.40 (0.44-4.41)	0.38	0.41
	rs3917527	0.82 (0.56-1.19)	0.66 (0.11-3.99)	0.82	0.52
	rs2301711	0.87 (0.63-1.20)	1.00 (0.20-4.95)	0.87	0.71
	rs3917490	1.10 (0.87-1.38)	1.16 (0.39-3.47)	0.91	0.68
	rs2299262	0.96 (0.76-1.22)	2.66 (0.70-10.05)	0.11	0.29
	rs2237583	0.91 (0.69-1.20)	5.00 (0.58-42.79)	0.07	0.19
	rs705379	0.90 (0.72-1.13)	1.50 (0.53-4.21)	0.34	0.51
PON2					
	rs7785846	0.94 (0.72-1.23)	0.16 (0.02-1.38)	0.05	0.12
	rs7786401	0.94 (0.72-1.23)	0.16 (0.02-1.38)	0.05	0.12
	rs7493	0.86 (0.66-1.13)	0.20 (0.02-1.71)	0.12	0.13
	rs2286232	0.91 (0.70-1.18)	0.50 (0.09-2.73)	0.47	0.57
	rs2299263	0.89 (0.68-1.16)	0.20 (0.20-1.71)	0.11	0.16
	rs17166879	0.96 (0.73-1.25)	0.20 (0.02-1.71)	0.10	0.22
	rs2299267	0.80 (0.59-1.09)	3.50 (0.72-16.84)	0.04	0.08
PON3					
	rs10953143	0.92 (0.73-1.16)	0.50 (0.15-1.66)	0.31	0.40
	rs3757708	0.96 (0.76-1.20)	1.80 (0.60-5.37)	0.26	0.52
	rs1053275	0.93 (0.74-1.17)	2.00 (0.68-5.85)	0.16	0.36
	rs978903	0.98 (0.78-1.23)	1.80 (0.60-5.37)	0.27	0.55

B. Interactions with maternal environmental contact with agrotoxics.

		Environmental contact with agrotoxics		*P* value (1 df)	*P* value (2 df)
		Non-Exposure OR (95% CI)	Exposure OR (95% CI)		
SOD1					
	rs1041740	0.84 (0.65-1.09)	0.69 (0.29-1.62)	0.65	0.31
	rs4817420	0.83 (0.64-1.07)	0.75 (0.31-1.78)	0.82	0.30
SOD2					
	rs8031	1.21 (0.96-1.53)	0.65 (0.32-1.30)	0.09	0.11
	rs2758331	1.11 (0.88-1.40)	0.77 (0.38-1.56)	0.33	0.50
	rs2855116	1.12 (0.89-1.42)	0.66 (0.32-1.38)	0.17	0.32
	rs4880	0.80 (0.64-1.01)	1.00 (0.48-2.04)	0.57	0.18
SOD3					
	rs699473	0.81 (0.64-1.03)	0.69 (0.29-1.62)	0.70	0.17
PON1					
	rs854552	1.08 (0.84-1.38)	0.64 (0.27-1.48)	0.23	0.47
	rs3917551	0.72 (0.49-1.08)	0.50 (0.15-1.66)	0.55	0.14
	rs3917548	0.81 (0.56-1.17)	0.44 (0.13-1.44)	0.32	0.20
	rs662	0.92 (0.72-1.17)	0.55 (0.25-1.20)	0.21	0.24
	rs3917527	0.81 (0.55-1.19)	0.83 (0.25-2.73)	0.96	0.54
	rs2301711	0.89 (0.64-1.24)	0.75 (0.26-2.16)	0.75	0.69
	rs3917490	1.13 (0.89-1.43)	0.78 (0.35-1.73)	0.37	0.46
	rs2299262	1.00 (0.78-1.28)	0.93 (0.45-1.93)	0.84	0.98
	rs2237583	0.93 (0.70-1.24)	1.10 (0.46-2.59)	0.73	0.88
	rs705379	0.92 (0.74-1.16)	0.92 (0.42-2.02)	0.98	0.80
PON2					
	rs7785846	0.93 (0.71-1.22)	0.71 (0.31-1.60)	0.53	0.63
	rs7786401	0.90 (0.68-1.19)	0.91 (0.40-2.07)	0.97	0.75
	rs7493	0.85 (0.64-1.13)	0.71 (0.31-1.60)	0.67	0.40
	rs2286232	0.93 (0.71-1.22)	0.68 (0.31-1.48)	0.45	0.55
	rs2299263	0.91 (0.69-1.20)	0.56 (0.24-1.27)	0.26	0.30
	rs17166879	0.96 (0.72-1.27)	0.71 (0.31-1.60)	0.49	0.68
	rs2299267	0.83 (0.61-1.14)	1.12 (0.43-2.91)	0.56	0.52
PON3					
	rs10953143	0.87 (0.68-1.11)	1.27 (0.57-2.80)	0.37	0.45
	rs3757708	1.00 (0.79-1.27)	0.78 (0.35-1.73)	0.55	0.83
	rs1053275	0.96 (0.76-1.21)	1.00 (0.46-2.15)	0.93	0.95
	rs978903	1.02 (0.80-1.28)	0.86 (0.41-1.82)	0.67	0.91

C. Interactions with maternal cigarette smoking.

		Cigarette smoking			
		Non-Exposure OR (95% CI)	Exposure OR (95% CI)	*P* value (1 df)	*P* value (2 df)
SOD1					
	rs1041740	0.83 (0.63-1.08)	0.83 (0.42-1.65)	0.99	0.34
	rs4817420	0.83 (0.63-1.08)	0.78 (0.40-1.55)	0.89	0.31
SOD2					
	rs8031	1.11 (0.88-1.40)	1.38 (0.67-2.82)	0.57	0.42
	rs2758331	1.07 (0.85-1.34)	1.13 (0.56-2.26)	0.87	0.79
	rs2855116	1.04 (0.82-1.31)	1.38 (0.67-2.82)	0.45	0.62
	rs4880	0.83 (0.66-1.04)	0.73 (0.36-1.46)	0.73	0.20
SOD3					
	rs699473	0.91 (0.71-1.16)	0.35 (0.17-0.70)	0.007	0.005
PON1					
	rs854552	1.03 (0.80-1.32)	1.06 (0.52-2.15)	0.93	0.95
	rs3917551	0.67 (0.45-0.99)	1.00 (0.32-3.10)	0.51	0.13
	rs3917548	0.74 (0.51-1.07)	1.00 (0.32-3.10)	0.63	0.28
	rs662	0.86 (0.67-1.10)	1.00 (0.48-2.04)	0.70	0.50
	rs3917527	0.79 (0.54-1.16)	1.00 (0.32-3.10)	0.70	0.50
	rs2301711	0.83 (0.59-1.17)	1.22 (0.50-2.94)	0.43	0.53
	rs3917490	1.17 (0.92-1.49)	0.68 (0.35-1.31)	0.12	0.20
	rs2299262	1.00 (0.78-1.29)	0.95 (0.51-1.75)	0.86	0.98
	rs2237583	0.91 (0.68-1.22)	1.25 (0.58-2.67)	0.45	0.71
	rs705379	0.89 (0.70-1.12)	1.23 (0.65-2.34)	0.35	0.51
PON2					
	rs7785846	0.88 (0.67-1.15)	1.25 (0.49-3.16)	0.48	0.59
	rs7786401	0.87 (0.66-1.15)	1.25 (0.49-3.16)	0.47	0.58
	rs7493	0.81 (0.61-1.07)	1.25 (0.49-3.16)	0.38	0.30
	rs2286232	0.87 (0.67-1.14)	1.25 (0.49-3.16)	0.47	0.56
	rs2299263	0.83 (0.63-1.09)	1.42 (0.54-3.75)	0.28	0.32
	rs17166879	0.90 (0.68-1.19)	1.25 (0.49-3.16)	0.51	0.69
	rs2299267	0.80 (0.58-1.11)	1.27 (0.57-2.80)	0.29	0.35
PON3					
	rs10953143	0.87 (0.68-1.11)	1.13 (0.56-2.26)	0.49	0.54
	rs3757708	1.04 (0.82-1.32)	0.65 (0.34-1.24)	0.17	0.40
	rs1053275	1.02 (0.80-1.29)	0.66 (0.35-1.25)	0.21	0.43
	rs978903	1.05 (0.83-1.33)	0.72 (0.38-1.38)	0.28	0.56

D. Interactions with maternal alcohol consumption.

		Alcohol consumption			
		Non-Exposure OR (95% CI)	Exposure OR (95% CI)	*P* value (1 df)	*P* value (2 df)
SOD1					
	rs1041740	0.80 (0.62-1.04)	1.28 (0.47-3.45)	0.36	0.23
	rs4817420	0.78 (0.60-1.01)	1.66 (0.60-4.58)	0.15	0.11
SOD2					
	rs8031	1.09 (0.87-1.37)	2.00 (0.80-4.95)	0.19	0.21
	rs2758331	1.01 (0.81-1.27)	2.50 (0.97-6.44)	0.05	0.13
	rs2855116	1.02 (0.81-1.28)	2.14 (0.87-5.25)	0.10	0.22
	rs4880	0.85 (0.68-1.07)	0.38 (0.13-1.07)	0.11	0.06
SOD3					
	rs699473	0.82 (0.65-1.04)	0.64 (0.27-1.48)	0.57	0.15
PON1					
	rs854552	1.03 (0.80-1.31)	1.11 (0.45-2.73)	0.87	0.94
	rs3917551	0.70 (0.48-1.04)	0.60 (0.14-2.51)	0.82	0.16
	rs3917548	0.75 (0.52-1.08)	1.00 (0.25-3.99)	0.69	0.30
	rs662	0.87 (0.69-1.11)	0.87 (0.31-2.41)	0.99	0.53
	rs3917527	0.80 (0.55-1.17)	1.00 (0.25-3.99)	0.76	0.51
	rs2301711	0.88 (0.63-1.22)	0.80 (0.21-2.97)	0.88	0.71
	rs3917490	1.11 (0.88-1.40)	0.88 (0.34-2.30)	0.64	0.62
	rs2299262	0.96 (0.75-1.22)	1.71 (0.67-4.35)	0.23	0.49
	rs2237583	0.91 (0.69-1.20)	2.00 (0.60-6.64)	0.19	0.41
	rs705379	0.87 (0.69-1.09)	2.14 (0.87-5.25)	0.04	0.11
PON2					
	rs7785846	0.91 (0.69-1.19)	0.87 (0.31-2.41)	0.94	0.76
	rs7786401	0.90 (0.68-1.19)	0.87 (0.31-2.41)	0.94	0.75
	rs7493	0.83 (0.63-1.10)	0.87 (0.31-2.41)	0.93	0.44
	rs2286232	0.90 (0.69-1.17)	0.87 (0.31-2.41)	0.95	0.73
	rs2299263	0.85 (0.65-1.12)	1.00 (0.35-2.85)	0.78	0.54
	rs17166879	0.93 (0.71-1.22)	0.87 (0.31-2.41)	0.90	0.85
	rs2299267	0.78 (0.57-1.06)	4.50 (0.97-20.82)	0.01	0.02
PON3					
	rs10953143	0.86 (0.68-1.09)	2.00 (0.68-5.85)	0.12	0.20
	rs3757708	1.03 (0.82-1.30)	0.50 (0.20-1.23)	0.11	0.29
	rs1053275	1.01 (0.80-1.27)	0.46 (0.17-1.21)	0.10	0.26
	rs978903	1.07 (0.85-1.34)	0.35 (0.12-0.99)	0.02	0.09

E. Interactions with maternal drug consumption.

		Drug consumption			
		Non-Exposure OR (95% CI)	Exposure OR (95% CI)	*P* value (1 df)	*P* value (2 df)
SOD1					
	rs1041740	0.82 (0.63-1.08)	0.85 (0.44-1.62)	0.94	0.34
	rs4817420	0.84 (0.64-1.10)	0.72 (0.38-1.38)	0.67	0.28
SOD2					
	rs8031	1.12 (0.88-1.42)	1.23 (0.69-2.20)	0.76	0.47
	rs2758331	1.12 (0.88-1.42)	0.84 (0.47-1.49)	0.36	0.53
	rs2855116	1.08 (0.85-1.38)	1.00 (0.56-1.78)	0.79	0.79
	rs4880	0.78 (0.62-0.99)	1.04 (0.60-1.80)	0.36	0.14
SOD3					
	rs699473	0.84 (0.65-1.07)	0.65 (0.36-1.16)	0.43	0.13
PON1					
	rs854552	1.10 (0.84-1.42)	0.76 (0.42-1.37)	0.26	0.51
	rs3917551	0.74 (0.49-1.11)	0.50 (0.18-1.33)	0.45	0.12
	rs3917548	0.84 (0.57-1.23)	0.46 (0.19-1.14)	0.22	0.15
	rs662	0.96 (0.74-1.23)	0.54 (0.30-0.99)	0.08	0.11
	rs3917527	0.84 (0.56-1.26)	0.69 (0.29-1.61)	0.67	0.49
	rs2301711	0.93 (0.66-1.32)	0.64 (0.30-1.38)	0.37	0.49
	rs3917490	1.08 (0.84-1.37)	1.22 (0.69-2.15)	0.68	0.63
	rs2299262	0.99 (0.76-1.27)	1.04 (0.58-1.87)	0.87	0.98
	rs2237583	1.04 (0.77-1.41)	0.64 (0.34-1.19)	0.16	0.35
	rs705379	0.93 (0.73-1.17)	0.90 (0.49-1.66)	0.94	0.79
PON2					
	rs7785846	0.79 (0.60-1.05)	2.00 (0.96-4.12)	0.01	0.04
	rs7786401	0.77 (0.57-1.03)	2.44 (1.12-5.30)	0.004	0.01
	rs7493	0.71 (0.53-0.94)	2.44 (1.12-5.30)	0.002	0.004
	rs2286232	0.77 (0.58-1.02)	2.30 (1.09-4.83)	0.005	0.01
	rs2299263	0.72 (0.54-0.97)	2.55 (1.18-5.52)	0.001	0.003
	rs17166879	0.82 (0.61-1.09)	1.90 (0.92-3.95)	0.03	0.08
	rs2299267	0.90 (0.65-1.25)	0.66 (0.32-1.38)	0.44	0.46
PON3					
	rs10953143	0.82 (0.64-1.06)	1.34 (0.78-2.31)	0.10	0.18
	rs3757708	1.09 (0.85-1.40)	0.58 (0.33-1.02)	0.03	0.12
	rs1053275	1.03 (0.81-1.32)	0.65 (0.36-1.16)	0.14	0.33
	rs978903	1.10 (0.86-1.40)	0.60 (0.33-1.07)	0.05	0.15

F. Interactions with maternal folic acid absence.

		Folic acid absence			
		Non-Exposure OR (95% CI)	Exposure OR (95% CI)	*P* value (1 df)	*P* value (2 df)
SOD1					
	rs1041740	0.58 (1.12-0.16)	0.85 (0.58-1.25)	0.82	0.33
	rs4817420	0.78 (0.56-1.09)	0.87 (0.60-1.28)	0.65	0.28
SOD2					
	rs8031	1.10 (0.83-1.46)	1.19 (0.83-1.70)	0.73	0.47
	rs2758331	1.08 (0.82-1.42)	1.06 (0.75-1.51)	0.94	0.80
	rs2855116	1.07 (0.81-1.42)	1.07 (0.74-1.53)	0.98	0.82
	rs4880	0.84 (0.64-1.11)	0.78 (0.55-1.11)	0.74	0.20
SOD3					
	rs699473	0.87 (0.65-1.16)	0.71 (0.49-1.03)	0.41	0.13
PON1					
	rs854552	1.13 (0.83-1.54)	0.91 (0.63-1.32)	0.38	0.65
	rs3917551	0.96 (0.58-1.59)	0.47 (0.26-0.84)	0.06	0.03
	rs3917548	1.00 (0.63-1.57)	0.52 (0.30-0.92)	0.07	0.06
	rs662	0.88 (0.65-1.20)	0.86 (0.60-1.23)	0.91	0.53
	rs3917527	1.00 (0.61-1.63)	0.63 (0.36-1.09)	0.22	0.26
	rs2301711	1.02 (0.68-1.53)	0.70 (0.42-1.16)	0.25	0.37
	rs3917490	1.16 (0.88-1.55)	1.00 (0.69-1.44)	0.51	0.55
	rs2299262	0.89 (0.66-1.19)	1.21 (0.82-1.79)	0.21	0.46
	rs2237583	0.83 (0.59-1.17)	1.20 (0.76-1.90)	0.20	0.42
	rs705379	1.01 (0.78-1.32)	0.77 (0.52-1.12)	0.23	0.39
PON2					
	rs7785846	0.80 (0.57-1.11)	1.12 (0.73-1.73)	0.21	0.35
	rs7786401	0.82 (0.58-1.14)	1.08 (0.69-1.69)	0.33	0.47
	rs7493	0.77 (0.55-1.08)	0.97 (0.62-1.52)	0.42	0.32
	rs2286232	0.78 (0.56-1.08)	1.15 (0.75-1.75)	0.15	0.26
	rs2299263	0.77 (0.55-1.08)	1.05 (0.67-1.62)	0.28	0.31
	rs17166879	0.80 (0.57-1.12)	1.18 (0.76-1.84)	0.16	0.32
	rs2299267	0.82 (0.56-1.19)	0.93 (0.57-1.53)	0.67	0.56
PON3					
	rs10953143	0.83 (0.62-1.11)	1.03 (0.71-1.50)	0.36	0.44
	rs3757708	1.09 (0.82-1.45)	0.82 (0.56-1.19)	0.22	0.47
	rs1053275	1.08 (0.81-1.43)	0.79 (0.54-1.15)	0.18	0.40
	rs978903	1.10 (0.83-1.46)	0.85 (0.58-1.23)	0.26	0.54

Supplementary Table 4**.** Haplotype family-based association tests for *SOD1, SOD2*, *PON1*, *PON2* and *PON3* performed using HBAT.

A. *SOD1*.

Haplotype	Number of families	Frequency	Z score	*P* value
C-C	178	0.706	1.55	0.12
T-T	174	0.279	-1.37	0.16
T-C	10	0.009	-0.56	0.57

Sequence: rs1041740 and rs4817420.

B. *SOD2*.

Haplotype	Number of families	Frequency	Z score	*P* value
A-C-A-G	211	0.461	-1.78	0.07
T-A-C-A	196	0.415	1.29	0.19
A-C-A-A	57	0.058	1.10	0.26
T-C-A-A	16	0.014	0.25	0.80
T-A-A-A	12	0.010	0.003	0.99
T-C-A-G	10	0.008	1.62	0.10
T-A-C-G	22	0.007	-2.15	0.03

Sequence: rs8031, rs2758331, rs2855116 and rs4880.

C. *PON1*.

Haplotype	Number of families	Frequency	Z score	*P* value
C-C-C	217	0.422	1.11	0.26
T-C-C	166	0.217	-0.59	0.55
T-T-T	152	0.198	0.66	0.50
T-T-C	101	0.114	-0.82	0.40
C-T-C	39	0.019	0.75	0.45
C-T-T	35	0.014	-1.03	0.30
C-C-T	23	0.010	-0.85	0.39
T-C-T	25	0.007	-2.18	**0.02**

Sequence rs3917490, rs2299262 and rs2237583.

D. *PON2*.

Haplotype	Number of families	Frequency	Z score	*P* value
C-G-G-C-C-A	147	0.756	0.70	0.48
T-T-C-T-T-G	141	0.193	-0.61	0.53
T-T-C-T-T-A	14	0.011	0.00	1.00

Sequence: rs7785846, rs7786401, rs7493, rs2286232, rs2299263 and rs17166879.

E. *PON3*.

Haplotype	Number of families	Frequency	Z score	*P* value
C-T-C-A	190	0.452	0.05	0.95
T-G-T-G	187	0.416	-0.74	0.45
C-G-T-G	72	0.080	1.25	0.20
C-T-T-G	31	0.028	-0.64	0.51
C-T-T-A	11	0.009	1.87	0.06
T-T-C-A	12	0.003	-1.99	0.04

Sequence: rs10953143, rs3757708, rs1053275 and rs978903.

Supplementary Table 5. Characteristics of patients included in the case-control study.

	Control (n=866)	NSCL±P (n=722)	*P* value
Gender			
Male	406 (46.88%)	409 (56.65%)	0.0001[a]
Female	460 (53.12%)	313 (43.35%)	
Ancestry			
European	83.6%	81.4%	0.99[b]
African	14.6%	16.8%	
Amerindian	1.8%	1.8%	

NSCL±P: nonsyndromic cleft lip with or without cleft palate.

[a]*P* value calculated with x^2 test and [b]*P* value calculated with Kruskal-Wallis test.

Supplementary Table 6. Haplotype analysis of the single-nucleotide polymorphisms in *PON1*, *PON2* and *PON3* in the case-control sample. P values were adjusted for confounders by logistic regression analysis.

A. *PON1* haplotypes.

Haplotype	Control	NSCL±P	*P* value
C-C-C	41.8%	38.8%	Reference
T-C-C	21.6%	26.5%	0.009
T-T-T	22.5%	19.7%	0.42
T-T-C	8.9%	10.4%	0.14
C-T-C	1.9%	2.6%	0.29
C-T-T	2.2%	1.0%	0.08

Sequence: rs3917490, rs2299262 and rs2237583.

B. *PON2* haplotypes.

Haplotype	Control	NSCL±P	*P* value
C-A-A	61.0%	61.9%	Reference
T-G-A	19.9%	17.5%	0.17
C-A-G	15.2%	15.5%	0.91
T-A-A	2.3%	3.1%	0.22
C-G-A	0.4%	0.7%	0.48

Sequence: rs2286232, rs17166879 and rs2299267.

C. *PON3* haplotypes.

Haplotype	Control	NSCL±P	*P* value
T-T-C-A	45.8%	42.8%	Reference
C-G-T-G	41.5%	38.9%	0.94
T-G-T-G	9.9%	12.8%	0.01
T-T-T-G	0.1%	3.3%	0.02
T-T-T-A	0.1%	0.4%	0.004

Sequence: rs10953143, rs3757708, rs1053275 and rs978903.

Discussion

As NSCL±P are among the most common congenital malformations in humans, with high mortality risk in developing countries, and therapeutic options are complex and expensive [15], considerable efforts are invested to identify the environmental and genetic factors associated with NSCL±P susceptibility. The known factors are not able to explain all cases of NSCL±P, and behind the concept of multifactorial disease, the current research has been focused on interacting networks between genes and genes and environmental factors to characterize putative markers with predictive significance for the etiology of disease. In this study, considering two representative and complementary samples of NSCL±P, consisting of more than 1,000 patients, we examined whether SNP in genes encoding enzymes with specific roles in the neutralization of dangerous reactive compounds for DNA damage interact with environmental factors in risk of NSCL±P. In the discovery sample, TDT analysis based on alleles and haplotypes and GxE interactions did not generate any signals achieving the formally adopted threshold of significance, but when GxG interactions were incorporated, several pairwise SNP-SNP interactions between *PON1*, *PON2* and *PON3* attained significant associations with NSCL±P. The case-control structured analysis supported some findings and revealed that GxG interactions involving rs2237583 in *PON1* and rs17166879 in *PON2* are associated with NSCL±P in the Brazilian population. In the case-control sample, our findings also showed that rs2237583 is significantly associated with NSCL±P, when all sample was considered, and rs3917490 in *PON1* reached a significant association with NSCL±P only in patients with a high percentage of African ancestry, suggesting that this polymorphism may have different associations with NSCL±P depending on specific ancestry.

PONs are a family of calcium-dependent hydrolases that are involved in antioxidant defense and the metabolism of various organophosphorus compounds, including insecticides [29]. Indeed, the name paraoxonase is derived of the first known substrate of PON1, paraoxon, the active metabolite of the insecticide parathion [30]. PONs also breakdown endogenous ROS by hydrolyzing oxidized phospholipids and by inactivating lipid hydroperoxides and hydrogen peroxide [31]. Although all three *PONs* seem to be important players in the maintenance of a low oxidative state, they are new identified antioxidant enzymes, warranting new studies to fully elucidate their physiological functions [32]. *PONs* are expressed in nearly all human tissues, and are

implicated in many diseases including cardiovascular disease, Parkinson's disease, Alzheimer's disease and diabetes [8], but little is known about their associations with developmental disorders. However, genetic factors altering expression and activity of PON family members may play a key role in determining susceptibility to oxidative stress associated with environmental exposures during the first trimester of gestation and, subsequently, DNA damage. In support, *PON* expression has been implicated in pathogenesis of neural tube defects, with its activity being essential in protecting of oxidative stress-induced neural tube defects [9, 11]. Moreover, high levels of oxidative stress decrease cell differentiation, resulting in cleft palate in mice [11, 14]. Several *PON* variants are associated with alterations in paraoxonase activity, but the most studied is *PON1* rs662, a variant located in the coding region that promotes reduction in the activity of enzyme [8, 33]. Of all selected SNP of present study, only *PON1* rs662 and *PON2* rs4880 were previously studied in oral clefts, but no associations were found [34, 35]. Similarly, our study failed to identify significant associations between rs662 and rs4880 and NSCL±P, and this does not seem to be related to power, because sufficient power to detect a genetic effect was achieved for these and many other SNP.

In this study the C allele of *PON1* rs2237583 induced a protective effect against NSCL±P. The presence of variant allele (allele T) of this SNP has been associated with a decrease in the enzyme activity in vascular diseases [36, 37]. A study with mothers and children from a region of Northern California identified 12 SNP in *PON1* associated with a reduction in the enzyme activity, including rs2237583 [33]. Thus, the presence of C allele ensures the functionality of enzyme, explaining the protective effect for NSCL±P development found in our study. However, rs2237583 was in strong linkage disequilibrium with rs705379 [33], and rs705379 was considered the causal variant in this locus, because it disrupts the Sp1 factor-binding site in the promoter region, thereby affecting transcription levels [38, 39]. In the current study, rs705379 was not associated with NSCL±P and was not in LD with the rs2237583 (data not shown). As rs2237583 allele frequencies differ according to ethnicity, the contribution of this SNP in different ancestry background populations may also differ.

Indeed, previous studies with *PON* genotypes and phenotypes in different ethnic populations have revealed significant variations in both allele and haplotype frequencies and enzyme activities [40, 41, 42]. Exceeding 200 million people, the Brazilian population is highly heterogeneous and admixed, a fact that has far reaching

implications to NSCL±P susceptibility [19, 24, 25, 26, 27, 28]. The extent of admixture is well documented in a wealth of population genetic studies in Brazilians, and it is evident that most individuals have significant degrees of European, African and Amerindian ancestry. As anticipated, the individual proportions of European, African and Amerindian ancestry varied widely in this study. Applying subgroup analysis, no significant associations were observed in the patients with high European ancestry, though the reduction in the number of samples drove a decreased statistical power. However, even with a small group, *PON1* rs3917490 was significantly associated with NSCL±P in the subgroup of patients with high African ancestry. The frequency of the rs3917490 T allele, which was associated with the protective effect, showed a higher frequency in patients with African ancestry compared with patients with elevated European ancestry. The impact of rs3917490, an intronic SNP, on paraoxonase activity of PON1 has never been assessed. Although genetic association studies often focus on coding SNP, particularly nonsynonymous SNP resulting in amino acid changes because they are likely to be functional, recent studies have showed that SNP in other regions, including introns and 3' UTR regions, may also have regulatory functional consequences. Furthermore, this SNP may be not the true etiologic variant - it can be in LD with some unknown gene or region (the true-associated variant). Further efforts are needed to determine the functionality of rs3917490, allowing the elucidation of mechanisms through this SNP leads to NSCL±P susceptibility in the Brazilian population with high African ancestry. Moreover, these results suggest that some of the population differences in association with NSCL±P can be explained by differences in frequency of alleles and haplotypes on *PON* family members.

Our study has some limitations, including the lack of characterization of impact of SNP on function of the encoded enzymes and the limited power in the analysis of the subgroup-based ancestry proportions. We also do not know whether other environmental factors, such as nutritional deficiency, which is related to oxidative stress induction, could exert important roles under GxE interactions on pathogenesis of this common disease. However, the study has several strengths. First, we pooled NSCL±P samples from several regions of Brazil, giving a high geographic coverage and a representative sample of the Brazilian population. Second, we have applied case-parent trio as a discovery approach followed by validation of the significant signals in a case-control independent sample. This approach is based on the premise

that if the same effect of a disease-marker can be obtained from case-parent trio and case-control study, the magnitude of information is strong and true. Third, applying GxG and GxE interaction analyses, we were able to verify whether the genes directly affect NSCL±P or whether their effects might be exerted through interactions between them or with environmental factors. Fourth, all measurements in the case-control sample were controlled for confounding effects including gender and ancestry proportions. Fifth, by evaluating the genomic structure of each sample, we were able to perform stratification analysis based on proportion of European and African proportions, revealing differential associations for the SNP rs3917490 in *PON1*.

Conclusion

In summary, this study was the first to cover the relationship of *SOD* and *PON* genetic variants and NSCL±P in the Brazilian population. Our results revealed by combining family-based and case-control models that markers in *PON1*, *PON2* and *PON3* may influence the risk of NSCL±P through potential GxG interactions. The presence of C allele in rs2237583 and of T allele in rs3917490, both in *PON1*, were associated with lower odds for NSCL±P, though the effect of rs3917490 was only observed in the sample with high African ancestry. Future studies are needed to confirm the exact mechanism by which *PON* variants may contribute to the development of NSCL±P. Furthermore, our findings showed that the diversity of the Brazilian population clearly influences the susceptibility of specific SNP in *PON1* on NSCL±P pathogenesis.

Funding information

This work was supported by grants from The State of São Paulo Research Foundation-FAPESP, São Paulo, Brazil (2016/02667-0). H Martelli-Junior is supported by The Minas Gerais State Research Foundation-FAPEMIG, Minas Gerais, Brazil.

Author Contributions

Conceptualization: Renato A. Machado, Ricardo D. Coletta.

Performed the experiments: Renato A. Machado, Carolina de Oliveira Silva.

Analyzed the data: Renato A. Machado, Tao Wu, Sergio Roberto Line, Ricardo D. Coletta.

Contributed with samples: Silvia R. A. Reis, Vanessa M. C. Dantas, Darlene C. Persuhn, Hercílio Martelli-Junior.

Writing - original draft: Renato A. Machado, Ricardo D. Coletta.

Writing - review & editing: Renato A. Machado, Silvia R. A. Reis, Vanessa M. C. Dantas, Darlene C. Persuhn, Tao Wu, Sergio Roberto Line, Hercilio Martelli-Junior, Ricardo D. Coletta.

References

1. Crane JF, Trainor PA. Neural crest stem and progenitor cells. Annu Rev Cell Dev Biol. 2006; 22: 267-86.
2. Suga M, Hayashi Y, Furue MK. In vitro models of cranial neural crest development toward toxicity tests: frog, mouse, and human. Oral Dis. 2017; 23(5): 559:565.
3. Larouche G, Hales BF. The impact of human superoxide dismutase 1 expression in a mouse model on the embryotoxicity of hydroxyurea. Birth Defects Res A Clin Mol Teratol. 2009; 85(9): 800-7.
4. Kesarwala AH, Krishna MC, Mitchell JB. Oxidative stress in oral diseases. Oral Dis. 2016; 22: 9-18.
5. Kumar J, Teoh SL, Das S, Mahakknaukrauh P. Oxidative Stress in Oral Diseases: Understanding Its Relation with Other Systemic Diseases. From Physiol. 2017; 8: 693.
6. Fukai T, Folz RJ, Landmesser U, Harrison DG. Extracellular superoxide dismutase and cardiovascular disease. Cardiovasc Res. 2002; 55: 239-249.
7. Kotch LE, Chen SY, Sulik KK. Ethanol-induced teratogenesis: free radical damage as a possible mechanism. Teratology. 1995; 52(3): 128-36.
8. Précourt LP, Amre D, Denis MC, Lavoie JC, Delvin E, Seidman E, et al. The three-gene paraoxonase family: physiologic roles, actions and regulation. Atherosclerosis. 2011; 214(1): 20-36.
9. Gonzalez-Herrera L, Martín Cerda-Flores R, Luna-Rivero M, Canto-Herrera J, Pinto-Escalante D, Perez-Herrera, et al. Paraoxonase 1 polymorphisms and

haplotypes and the risk for having offspring affected with spina bifida in Southeast Mexico. Birth Defects Res A Clin Mol Teratol. 2010; 88: 987-94.

10. Kase BA, Northrup H, Morrison AC, Davidson CM, Goiffon AM, Fletcher JM, et al. Association of copper-zinc superoxide dismutase (SOD1) and manganese superoxide dismutase (SOD2) genes with nonsyndromic myelomeningocele. Birth Defects Res A Clin Mol Teratol. 2012; 94: 762-9.
11. Sakai D, Dixon J, Achilleos A, Dixon M, Trainor PA. Prevention of Treacher Collins syndrome craniofacial anomalies in mouse models via maternal antioxidant supplementation. Nat Commun. 2016; 7: 10328.
12. Scherer SW, Poorkaj P, Allen T, Kim J, Geshuri D, Nunes M, et al. Fine mapping of the autosomal dominant split hand/split foot locus on chromosome 7, band q21.3-q22.1. Am J Hum Genet. 1994; 55(1): 12-20.
13. Rinaldi C, Bramanti P, Famà A, Scimone C, Donato L, Antognelli C, et al. Glyoxalase I A111E, Paraoxonase 1 Q192R and L55M polymorphisms in italian patients with sporadic cerebral cavernous malformations: a pilot study. J Biol Regul Homeost Agents. 2015; 29(2): 493-500.
14. Yuan X, He X, Zhang X, Liu C, Wang C, Qiu L, et al. Comparative Study of Folic Acid and a-Naphthoflavone on Reducing TCDD Induced Cleft Palate in Fetal Mice. Cleft Palate Craniofac J. 2017; 54(2): 216-222.
15. Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011; 12(3): 167-178.
16. Beaty TH, Marazita ML, Leslie EJ. Genetic factors influencing risk to orofacial clefts: Today's challenges and tomorrow's opportunities. F1000Res. 2016; 5: 2800.
17. Spina V, Psillakis JM, Lapa FS, Ferreira MC. Classification of cleft lip and cleft palate. Suggested changes. Rev Hosp Clin Fac Med São Paulo. 1972; 27: 5-6.
18. Aidar M, Line SR. A simple and cost-effective protocol for DNA isolation from buccal epithelial cells. Braz Dent J. 2007; 18: 148-152.
19. Messetti AC, Machado RA, de Oliveira CE, Martelli-Júnior H, de Almeida Reis SR, Moreira HS, et al. Brazilian multicenter study of association between polymorphisms in CRISPLD2 and JARID2 and non-syndromic oral clefts. J Oral Pathol Med. 2017; 46: 232-239.
20. Machado RA, Moreira HS, de Aquino SN, Martelli-Junior H, de Almeida Reis SR, Persuhn DC, et al. Interactions between RAD51 rs1801321 and maternal cigarette

smoking as risk factor for nonsyndromic cleft lip with or without cleft palate. Am J Med Genet A. 2016; 170A(2): 536-9.

21. Wang P, Wu T, Schwender H, Wang H, Shi B, Wang Z, et al. Evidence of interaction between genes in the folate/homocysteine metabolic pathway in controlling risk of nonsyndromic oral cleft. Oral Dis. 2018 (in press).
22. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol. 2000; 19: S36-S42.
23. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155(2): 945-59.
24. Brito LA, Paranaiba LM, Bassi CF, Masotti C, Malcher C, Schlesinger D, et al. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. Birth Defects Res A Clin Mol Teratol. 2012; 94(6): 464-8.
25. Bagordakis E, Paranaiba LM, Brito LA, de Aquino SN, Messetti AC, Martelli-Junior H, et al. Polymorphisms at regions 1p22.1 (rs560426) and 8q24 (rs1530300) are risk markers for nonsyndromic cleft lip and/or palate in the Brazilian population. Am J Med Genet A. 2013; 161A(5): 1177-80.
26. de Aquino SN, Messetti AC, Hoshi R, Borges A, Viena CS, Reis SR, et al. Analysis of susceptibility polymorphisms for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol. 2014; 100(1): 36-42.
27. do Rego Borges A, Sá J, Hoshi R, Viena CS, Mariano LC, de Castro Veiga P, et al. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry. Am J Med Genet A. 2015; 167A(10): 2344-9.
28. Machado RA, de Toledo IP, Martelli-Júnior H, Reis SR, Guerra ENS, Coletta RD. Potential genetic markers for nonsyndromic oral clefts in the Brazilian population: A systematic review and meta-analysis. Birth Defects Res. 2018. (in press).
29. Marsillach J, Mackness B, Mackness M, Riu F, Beltrán R, Joven J, et al. Immunohistochemical analysis of paraoxonases-1, 2, and 3 expression in normal mouse tissues. Free Radic Biol Med. 2008; 45(2): 146-57.
30. Aldridge WN. An enzyme hydrolysing diethyl *p*-Nitrophenyl Phosphate (E600) and its identity with the A-esterase of mammalian sera. Biochem J. 1953; 53: 117-124.

31. Aviram M, Rosenblat M. Paraoxonases 1, 2, and 3, oxidative stress, and macrophage foam cell formation during atherosclerosis development. Free Radic Biol Med 2004; 37: 1304-16.
32. Borovkova EI, Antipova NV, Komeenko TV, Shakhparonov MI, Borovkov IM. Paraoxonase: The Universal Factor of Antioxidant Defense in Human Body. Vestn Ross Akad Med Nauk. 2017; 72(1): 5-10.
33. Huen K, Barcellos L, Beckman K, Rose S, Eskenazi B, Holland N. Effects of PON polymorphisms and haplotypes on molecular phenotype in Mexican-American mothers and children. Environ Mol Mutagen. 2011; 52(2): 105-16.
34. Turgut G, Ozcan A, Çakmak E, Bas L, Tatlidede S, Oguz O, et al. Paraoxonase-1 192 enzyme polymorphism in non-syndromic clefting: In patients and parents. Journal of Cell and Molecular Biology. 2008; 7(1): 67-74.
35. Mostowska A, Hozyasz KK, Lianeri M, Piwowar W, Jagodzinski PP. Polymorphic variants of genes encoding main antioxidant enzymes and the risk of CL/P-affected pregnancies. Clin Biochem. 2007; 40: 416-9.
36. Kim DS, Burt AA, Ranchalis JE, Richter RJ, Marshall JK, Eintracht JF, et al. Additional Common Polymorphisms in the PON Gene Cluster Predict PON1 Activity but Not Vascular Disease. J Lipids. 2012; 2012: 476316.
37. Yoshino S, Cilluffo R, Prasad M, Best PJ, Atkinson EJ, Aoki T, et al. Sex-Specific Genetic Variants are Associated With Coronary Endothelial Dysfunction. J Am Heart Assoc. 2016; 5(4): e002544.
38. Deakin S, Leviev I, Brulhart-Meynet MC, James RW. Paraoxonase-1 promoter haplotypes and serum paraoxonase: a predominant role for polymorphic position -107, implicating the Sp1 transcription factor. Biochem J. 2003; 372(Pt 2): 643-649.
39. Huen K, Yousefi P, Street K, Eskenazi B, Holland N. PON1 as a model for integration of genetic, epigenetic, and expression data on candidate susceptibility genes. Environ Epigenet. 2015;1(1). pii: dvv003.
40. Koda Y, Tachida H, Soejima M, Takenaka O, Kimura H. Population differences in DNA sequence variation and linkage disequilibrium at the PON1 gene. Ann Hum Genet. 2004; 68(Pt 2): 110-9.
41. Rojas-García AE, Solís-Heredia MJ, Piña-Guzmán B, Vega L, López-Carrillo L, Quintanilla-Vega B. Genetic polymorphisms and activity of PON1 in a Mexican population. Toxicol Appl Pharmacol. 2005; 205(3): 282-9.

42. Holland N, Furlong C, Bastaki M, Richter R, Bradman A, Huen K, et al. Paraoxonase polymorphisms, haplotypes, and enzyme activity in Latino mothers and newborns. Environ Health Perspect. 2006; 114(7): 985-91.

2.4 Artigo

Association between *GOLGB1* tag-polymorphisms and nonsyndromic cleft palate only in the Brazilian population

Artigo aceito para publicação no periódico *Annals of Human Genetics, 2018, doi: 10.1111/ahg.12242* (Anexo 4)

Renato Assis Machado[1], Hercílio Martelli-Júnior[2], Silvia Regina de Almeida Reis[3], Darlene Camati Persuhn[4], Ricardo D. Coletta[1]

[1]*Department of Oral Diagnosis, School of Dentistry, University of Campinas, Piracicaba, São Paulo, Brazil*

[2]*Stomatology Clinic, Dental School, State University of Montes Claros, Montes Claros, Minas Gerais, Brazil and Center for Rehabilitation of Craniofacial Anomalies, Dental School, University of José Rosario Vellano, Alfenas, Minas Gerais, Brazil*

[3]*Department of Basic Science, Bahiana School of Medicine and Public Health, Salvador, Bahia, Brazil*

[4]*Molecular Biology Department, Federal University of Paraíba, João Pessoa, Paraíba, Brazil.*

Running title: *GOLGB1* tag-SNPs in nonsyndromic cleft palate.

Corresponding author: Ricardo D. Coletta, Department of Oral Diagnosis, School of Dentistry, State University of Campinas, CEP 13414-018, Piracicaba, São Paulo, Brazil. E-mail: coletta@fop.unicamp.br

Manuscript information: 177 words in the abstract, 1985 words in the text, 15 references, 4 figures/tables, and 2 supplementary files.

Summary

Nonsyndromic oral clefts are common congenital birth defects that exhibit variable prevalence around the world, often influenced by population-dependent genetic predisposition. Few studies have been performed with nonsyndromic cleft palate only (NSCPO), limiting the knowledge of the genetic risk factors related to this type of oral cleft. Genetic variants in golgin subfamily B member 1 (*GOLGB1*), a gene that is essential for normal murine palatogenesis, were analyzed in this study to establish its potential association with NSCPO risk in the Brazilian population. Five tag-single nucleotide polymorphisms (SNPs) of *GOLGB1* (rs1169, rs7153, rs9968051, rs9819530 and rs6794341), which capture the majority of alleles spanning within gene, were genotyped in a case-control study with 270 patients with NSCPO and 284 unrelated healthy controls. The samples were also genotyped for 40 biallelic polymorphic markers to characterize the genetic ancestry. After adjustment for co-variants, the *GOLGB1* tag-SNPs and the haplotypes formed by those SNPs were not significantly associated with NSCPO in this Brazilian case-control cohort. Our results suggest that common polymorphisms of *GOLGB1* are not associated NSCPO susceptibility in the Brazilian population.

Keywords: nonsyndromic cleft palate only, single nucleotide polymorphism, risk factor, *GOLGB1*.

Introduction

Nonsyndromic oral clefts are usually divided into cleft lip with or without cleft palate (NSCL±P) and cleft palate only (NSCPO). NSCPO represents only one third of all nonsyndromic oral clefts, with prevalence around 1-25 per 10,000 live births (Burg et al., 2016). Females are frequently more affected than males (Burg et al., 2016). The etiology of nonsyndromic oral clefts is dependent of genetic and environmental interactions, though the specific underlying factors remain largely unclear. The genetic predisposition is ethnicity-dependent, a situation that is clearly illustrated in the Brazilian population. As result of five centuries of mating between Amerindians, Europeans and sub-Saharan Africans, the Brazilian population displays very high levels of genomic diversity, which has been shown to affect the specific genetic susceptibility of genes and loci previously associated with nonsyndromic oral clefts in other populations (Paranaiba et al., 2010; Bagordakis et al., 2013). Even in specific subsets of the Brazilian population, such as those displaying high African ancestry, a variable participation of different genomic loci in the pathogenesis of nonsyndromic oral clefts is observed (Aquino et al., 2014; do Rego Borges et al., 2016).

A variety of genetic approaches, including genome-wide association studies (GWAS), has been employed to identify genes and loci associated with NSCL±P, but few studies have been performed with NSCPO. Indeed, only 2 independent GWAS for NSCPO were performed, with *GRHL3* (rs41268753) achieving genome-wide significance when considered alone (Leslie et al., 2016), and markers in other few genes approached the genome-wide significance when gene-environment interactions with maternal smoking or multivitamin supplementation were taken into account (Beaty et al., 2011). An alternative has been the traditional studies based on animal models of palatogenesis or specific knockout mice resulting in cleft palate. Those models have yielded insights of participation of specific signaling pathways in the normal palatogenesis, and variations on those genes have been related to NSCPO. For example, Wnt signaling disruption is associated with mouse palatal malformations and single-nucleotide polymorphisms (SNPs) in *WNT3* have been linked to NSCPO in humans (Machado et al., 2016).

A recent study revealed that the loss of function of golgin subfamily B member 1 (*Golgb1) gene* causes cleft palate in mice due to alterations in protein glycosylation

patterns in the embryonic palatal mesenchymal cells (Lan et al., 2016). Furthermore, a 10-bp insertion in *Golgb1* exon 13 in rats, resulting in a truncated protein due to a premature termination codon, causes osteochondrodysplasia, which is characterized by cleft palate, among other phenotypes (Katayama et al., 2011). Owing to the substantial weight of evidence indicating the participation of *Golgb1* in murine palatogenesis and genetic variants in genes involved in palate embryonic development may be associated with NSCPO in humans, we conducted this case-control structured study with 5 tag-SNPs in *GOLGB1*, in which the genetic ancestry variation of each individual was taken into account, to verify the relationship between *GOLGB1* SNPs and NSCPO.

Materials and Methods

Samples

This case-control study included samples from 270 unrelated patients with NSCPO and 284 healthy control individuals. Patients with NSCPO were carefully examined and screened for the presence of associated anomalies or syndromes by the specialized team of the associated Centers for rehabilitation of craniofacial anomalies in Brazil: the Association of Carrier of Cleft Lip and Palate-APOFILAB, Cascavel-PR, the Center for Rehabilitation of Craniofacial Anomalies, University of José Rosário Vellano, Alfenas-MG, the University Hospital of Lauro Wanderley-HULW, João Pessoa-PB and the Santo Antonio Hospital, Salvador-BA. The control group was composed of healthy individuals with no physical illness, psychiatric, birth defects or with a family history of orofacial clefts, living in the same geographic areas (Supplementary Table 1). The study was approved by the ethics review board of each of the centers or hospitals affiliated with the collaborative study. Written informed consent was obtained from the parents or guardians and/or the participants.

Tag-SNP selection

The genomic region containing 5 kb upstream and 2 kb downstream of *GOLGB1* was screened for tag-SNPs using Tagger software (de Bakker et al., 2005). The capture of the tag-SNPs was performed with the 1000 Genomes Project data, following the criteria of minor allele frequency > 0.1, p value in Hardy-Weinberg equilibrium (HWE) > 0.05 and r^2 pairwise correlation > 0.9. Applying this stringent filter, 5 SNPs were selected. The main features of each SNP are described in Table 1.

Table 1 Characteristics of the tag-single nucleotide polymorphisms (SNPs) in *GOLGB1,* located at 3q13.33, and allelic distributions in the control and nonsyndromic cleft palate only (NSCPO) groups.

SNP	Position	Alleles	MAF Control	MAF NSCPO	HWE (p value)	OR (95% CI) / p value*
rs1169	121,663,357	C/t	0.363	0.354	0.73	0.96 (0.75-1.23) / 0.76
rs7153	121,664,450	G/a	0.241	0.237	0.62	0.97 (0.74-1.28) / 0.86
rs9968051	121,665,234	C/g	0.137	0.133	0.02	0.96 (0.68-1.36) / 0.85
rs9819530	121,668,937	T/g	0.375	0.336	0.79	0.84 (0.65-1.08) / 0.17
rs6794341	121,741,762	G/a	0.266	0.311	0.77	1.24 (0.95-1.61) / 0.09

MAF: minor allele frequency. Minor allele in lower case. HWE: *Hardy Weinberg equilibrium*.
*Values of p were adjusted taking into account the structured ancestral distribution.

Genotyping and assessment of genomic ancestry

Genomic DNAs extracted from oral mucosa cells were genotyped by TaqMan 5'-exonuclease allelic discrimination assays (Assay-on-Demand service, Applied Biosystems). For quality control purposes, reactions were randomly repeated in 10% of the samples for each SNP, and the concordance rate was 100%. All samples were successfully genotyped, with a genotype call rate of >99%.

To estimate the genomic ancestry proportions, each sample was also independently genotyped for 40 biallelic short insertion-deletion polymorphisms (Messetti et al., 2017).

Statistical analysis

All genotyped SNPs were tested for the HWE in the control group using the x^2 test. Group comparisons were tested by Student's t-test, x^2 test or Kruskal-Wallis test. Differences in the allele frequency conditioned on population structure were assessed with STRAT (Pritchard et al., 2000). Linkage disequilibrium (LD) was estimated using the HaploView software. Multiple logistic regression analysis was performed with R software, considering age, gender and genomic ancestry as potential co-variants. For

association analysis, the p value was adjusted by Bonferroni correction for multiple testing (p≤0.01). Power for detecting a p value ≤0.05 for each SNP was estimated using the QUANTO software assuming a prevalence of NSCL/P in Brazil of 0.00146 (Martelli-Junior et al., 2007).

Results

The distributions of genotypes in the control group were consistent with those predicted by the HWE (Table 1), except for rs9968051 (p=0.02). LD between each pair of SNPs was assessed by calculating the r^2 value as depicted in Supplementary Figure 1. The analysis revealed none of the tag-SNPs were in LD, confirming that each tag-SNP probably represent a distinct haplotype block.

The average ancestry contributions in the control group were estimated at 75.1% of European, 20.9% of African and 4.0% of Amerindian and those in the NSCPO group were 67.9% of European, 27.8% of African and 4.3% of Amerindian (Supplementary Table 1). There were no significant differences in the ancestry proportions between groups. Taking into account age, gender and genomic ancestry proportions of each sample, none of the tag-SNPs in *GOLGB1* was significantly associated with NSCPO in the additive, dominant or recessive genetic model (Table 2). Indeed, when AG and AA genotypes for rs6794341 were combined and compared against GG genotype (recessive genetic model), a nominal p value was observed (p=0.03). However, this association did not retain statistical significance after adjusting for multiple comparisons. Haplotype association analysis revealed a significant association between NSCPO and C-G-C-G-G haplotype (p=0.02), which did not withstand Bonferroni correction (Table 3).

Since this is the first study analyzing *GOLGB1* SNPs in NSCPO, we performed power analysis assuming our current sample and different effect sizes varying the OR from 1.0 to 2.0 (Figure 1). Assuming the MAFs found in this Brazilian population and the current sample size, the predictive analysis revealed a power of 80% in effects ranging from 1.4 to 1.6. Indeed, 4 out of 5 tag-SNPs showed an OR near 1.4 (1.41 for rs1169, 1.46 for rs7153, 1.41 for rs9819530, and 1.44 for rs6794341).

Table 2 Distribution of genotypes of the tag-single nucleotide polymorphisms (SNPs) in *GOLGB1* in the control and nonsyndromic cleft palate only (NSCPO) groups. Values of p were adjusted for co-variants by logistic regression analysis.

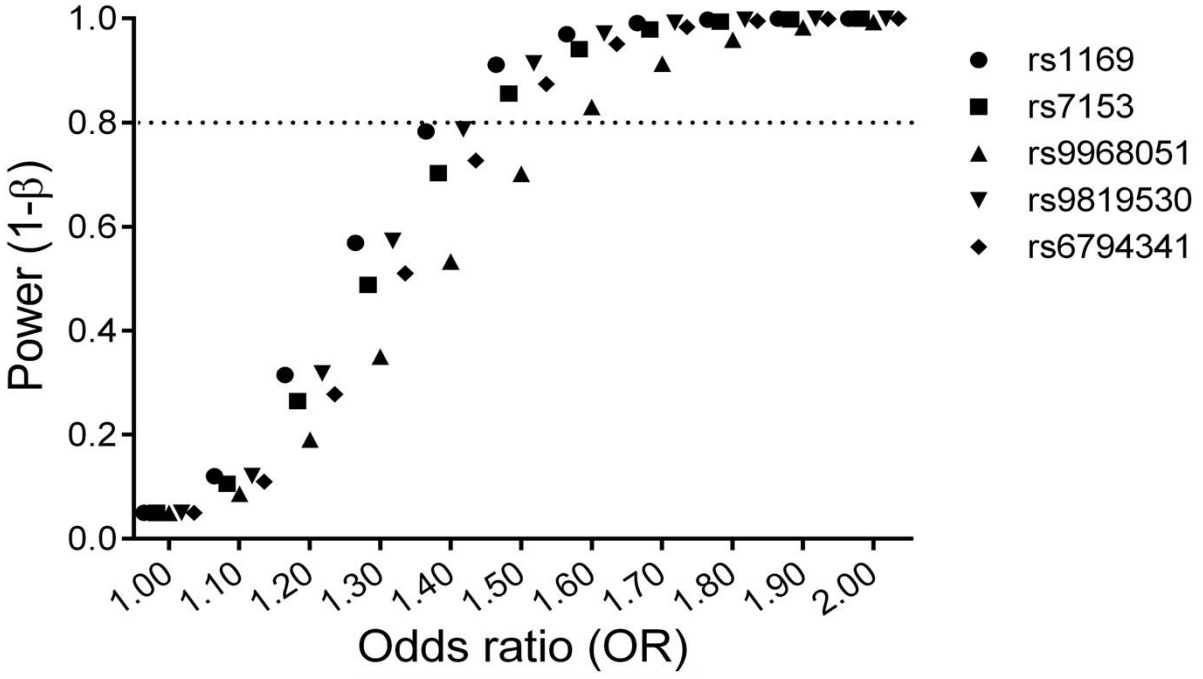
	Control (n=284)	NSCPO (n=270)	OR_{Het} (95% CI) / p value	OR_{Hom} (95 CI) / p value	OR_{Dom} (95% CI) / p value	OR_{Rec} (95% CI) / p value
rs1169 (CC/CT/TT)	40.1%/47.2%/12.7%	45.0%/39.3%/15.7%	0.72 (0.48-1.09) / 0.28	0.93 (0.52-1.66) / 0.28	0.77 (0.52-1.13) / 0.18	1.09 (0.64-1.88) / 0.75
rs7153 (GG/GA/AA)	57.0%/37.7%/5.3%	61.2%/30.2%/8.6%	0.72 (0.48-1.08) / 0.25	1.08 (0.51-2.28) / 0.25	0.77 (0.53-1.14) / 0.19	1.22 (0.59-2.52) /0.59
rs9968051 (CC/CG/GG)	76.1%/20.4%/3.5%	75.6%/22.2%/2.2%	1.25 (0.79-1.99) / 0.47	0.66 (0.19-2.24) / 0.47	1.17 (0.75-1.82) / 0.49	0.63 (0.19-2.12) / 0.45
rs9819530 (TT/TG/GG)	39.5%/46.1%/14.4%	45.1%/42.6%/12.3%	0.78 (0.52-1.17) / 0.37	0.72 (0.40-1.30) / 0.37	0.76 (0.52-1.12) / 0.17	0.81 (0.47-1.42) / 0.46
rs6794341 (GG/GA/AA)	54.2%/38.4%/7.4%	48.7%/40.4%/10.9%	1.42 (0.95-2.13) / 0.07	1.98 (0.98-4.00) / 0.07	1.51 (1.03-2.21) / 0.03	1.70 (0.86-3.35) / 0.12

Table 3 Haplotype analysis of the tag-single nucleotide polymorphisms in *GOLGB1* in controls and patients with nonsyndromic cleft palate only (NSCPO). P value was adjusted for co-variants by logistic regression analysis.

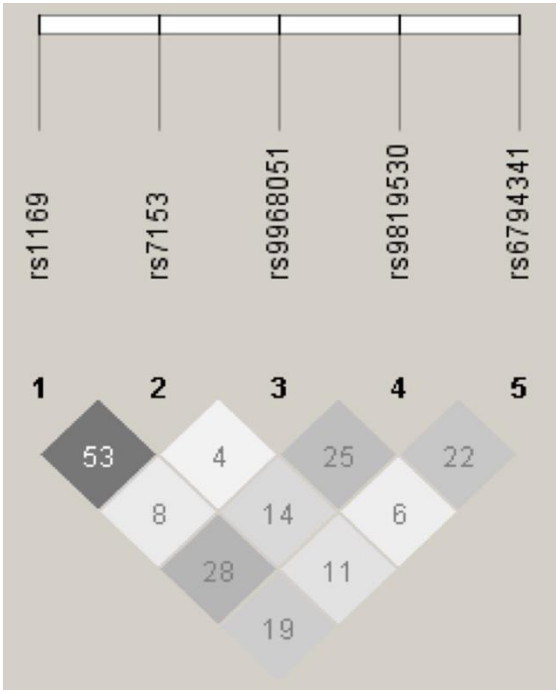
Haplotype	Control	NSCPO	p value
C-G-C-T-A	25.7%	29.5%	
T-A-C-T-G	23.3%	21.8%	0.08
C-G-C-G-G	23.9%	19.8%	0.02
C-G-G-G-G	12.5%	12.7%	0.45
T-G-C-T-G	12.4%	11.5%	0.16
C-G-C-T-G	0.4%	1.7%	0.19

Sequence: rs1169, rs7153, rs9968051, rs9819530 and rs6794341.

Figure 1 Statistical power of the study. If odds of NSCPO is ≥1.4, assuming the current sample, the statistical power of 80% is reached for rs1169, rs7153, rs9819530 and rs6794341.



Supplementary Figure 1 Linkage disequilibrium plot with the tag-single nucleotide polymorphisms (SNPs) within *GOLGB1*. The numbers in the squares indicate the r^2 values of the linkage disequilibrium between pair of SNPs.



Supplementary Table 1 Characteristics of controls and patients with nonsyndromic cleft palate only (NSCPO).

	Control (n=284)	NSCPO (n=270)	p value
Age (years)	24.02 ± 12.15	13.73 ± 11.33	0.001[a]
Gender			
Male	88 (31.0%)	104 (38.5%)	0.06[b]
Female	196 (69.0%)	166 (61.5%)	
Ancestry			
European	75.1%	67.9%	0.51[c]
African	20.9%	27.8%	
Amerindian	4.0%	4.3%	

[a]p value calculated with Student's t-test, [b]p value calculated with x^2 test, and [c]p value calculated with Kruskal-Wallis test.

Discussion

We hypothesized that *GOLGB1* single nucleotide variants would be associated with increased risk for NSCPO. To comprehensively investigate the role of genetic variants in *GOLGB1*, tag-SNP association strategy was performed. This strategy selects the representative SNPs in a specific gene or region, based on the linkage disequilibrium scores, allowing the analysis of contribution of each SNP independently to disease risk. This strategy is particularly suited to cover the entire candidate gene with the smallest number of SNPs. Our study failed to identify significant associations between *GOLGB1* genetic variants and NSCPO susceptibility in the Brazilian population. From a *strategic* perspective, our data is very consistent since we have included samples from 4 different Centers in 4 Brazilian states, covering a high geographic area of Brazil,

and assessed the ancestry contribution of each patient, correcting for specific effects that the population stratification could have.

Giantin, encoded by *GOLGB1* gene, is the largest Golgi complex-associated protein (Munro, 2011). Giantin interacts with p115 and GM130 to tether COPI vesicles with Golgi membranes, but by binding Rab1 and Rab6 proteins, is suggested that giantin may have additional regulatory roles, including the regulation of exocytotic vesicle transport from the endoplasmic reticulum to the Golgi apparatus (Rosing et al., 2007). Besides maintenance of compact Golgi morphology and regulation of vesicle transport, little is known on giantin. The relation between *GOLGB1* variations and human diseases has been reported in very few studies. By using CRISPR/Cas9-inducing *Golgb1* loss-of-function, Lan et al. (2016) demonstrated recently that mouse embryos deficient in *Golgb1* consistently exhibit cleft palate due to failure of palatal shelf elevation. The authors also showed an increased cell density, reduced hyaluronan accumulation and impaired protein glycosylation in the palatal mesenchymal cells. Although *Golgb1* is essential for normal embryonic development of palate in mice, the results of this study did not confirm the presumption that genetic variants in *GOLGB1* are associated with an increased risk of NSCPO in humans.

We estimated that this study has 80% of power to detect a moderate genetic effect (allele relative risk of ~1.4 in the additive risk model) for 4 out of 5 tag-SNPs. As a multifactorial and complex disease, it is well established that multiple genes with modest individual effects (individual odds no higher than 2.0) are capable of disrupting normal lip and palate development under specific circumstances, which may include interactions to environmental risk factors (Beaty et al., 2016). The only SNP with limited power was rs9968051, which showed a low frequency of the minor allele. For this SNP, 80% of power to detect a genetic effect was only obtained in OR ≥1.6 in an additive risk model. The rs9968051 was also the only that did not respect the HWE. Since our assays showed clear discrimination ability, quality control strategy revealed concordance rate of 100% and possible variations in the groups related to population stratification were ruled out by using population structure analysis based on genetic ancestry, rs9968051 was kept in the analyses. As we used tag-SNPs to capture the majority of common variants in the *GOLGB1* gene, it is possible that we might have missed rare variants in this gene associated with the disease. Furthermore, as *GOLGB1* risk allele frequencies differ according to ethnicity, this relationship should

be tested in other populations and functional studies should be performed to clarify the contribution of genetic background to the development of this disease. Thus, further large-scale studies and sequencing of the whole gene are warranted to confirm our findings.

In conclusion, our findings show that *GOLGB1* tag-SNPs are not likely to be associated with NSCPO susceptibility, at least in the ethnically-mixed Brazilian population. Considering the essential role of *GOLGB1* in palatogenesis, further studies with a larger sample sizes, samples from other ancestrally different populations and taking into account also possible interactions with environmental factors are necessaries to increase our understanding and knowledge of the genetic events that contribute to this uncommon form of nonsyndromic oral cleft.

Author Contributions

Conception and design of the study: RDC; Experimental data collection and analysis: RAM, HM-Jr, SAAR, DCP and RDC; Manuscript preparation: RAM and RDC. All authors revised and approved the final manuscript.

Funding Disclosure

This work was supported by grants from The State of São Paulo Research Foundation-FAPESP, São Paulo, Brazil (2016/02667-0). H Martelli-Junior is supported by The Minas Gerais State Research Foundation-FAPEMIG, Minas Gerais, Brazil.

Conflict of Interest

There are no conflicts of interest.

References

Bagordakis, E., Paranaiba, L. M., Brito, L. A., de Aquino, S. N., Messetti, A. C., Martelli-Junior, H., Swerts, M. S., Graner, E., Passos-Bueno, M. R., & Coletta, R. D. (2013) Polymorphisms at regions 1p22.1 (rs560426) and 8q24 (rs1530300) are risk markers for nonsyndromic cleft lip and/or palate in the Brazilian population. *Am J Med Genet A* **161a**: 1177-1180.

Beaty, T. H., Marazita, M. L., Leslie, E. J. (2016) Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities. *F1000Res* **5**:2800.

Beaty, T. H., Ruczinski, I., Murray, J. C., Marazita, M. L., Munger, R. G., Hetmanski, J. B., Murray, T., Redett, R. J., Fallin, M. D., Liang, K. Y., Wu, T., Patel, P. J., Jin, S. C., Zhang, T. X., Schwender, H., Wu-Chou, Y. H., Chen, P. K., Chong, S. S., Cheah, F., Yeow, V., Ye, X., Wang, H., Huang, S., Jabs, E. W., Shi, B., Wilcox, A. J., Lie, R. T., Jee, S. H., Christensen, K., Doheny, K. F., Pugh, E. W., Ling, H., & Scott A. F. (2011) Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. *Genet Epidemiol* **35**: 469-478.

Burg, M. L., Chai, Y., Yao, C. A., Magee, W., & Figueiredo, J. C. (2016) Epidemiology, etiology, and treatment of isolated cleft palate. *Front Physiol* **7**: 67.

de Aquino, S. N., Messetti, A. C., Hoshi, R., Borges, A., Viena, C. S., Reis, S. R., Oliveira Swerts, M. S., Graner. E., Martelli-Júnior, H., & Coletta, R. D. (2014) Analysis of susceptibility polymorphisms for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. *Birth Defects Res A Clin Mol Teratol* **100**: 36-42.

de Bakker, P. I., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J., & Altshuler, D. (2005) Efficiency and power in genetic association studies. *Nat Genet* **37**: 1217-1223.

do Rego Borges, A., Sá, J., Hoshi, R., Viena, C. S., Mariano, L. C., de Castro Veiga, P., Medrado, A. P., Machado, R. A., de Aquino, S. N., Messetti, A. C., Spritz, R. A., Coletta, R. D., & Reis, S. R. (2015) Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry. *Am J Med Genet A* **167A**: 2344-2349.

Katayama, K., Sasaki, T., Goto, S., Ogasawara, K., Maru, H., Suzuki, K., & Suzuki, H. (2011) Insertional mutation in the Golgb1 gene is associated with osteochondrodysplasia and systemic edema in the OCD rat. *Bone* **49**: 1027-1036.

Lan, Y., Zhang, N., Liu, H., Xu, J., & Jiang, R. (2016) Golgb1 regulates protein glycosylation and is crucial for mammalian palate development. *Development* **143**: 2344-2355.

Leslie, E. J., Liu, H., Carlson, J. C., Shaffer, J. R., Feingold, E., Wehby, G., Laurie, C. A., Jain, D., Laurie, C. C., Doheny, K. F., McHenry, T., Resick, J., Sanchez, C., Jacobs, J., Emanuele, B., Vieira, A. R., Neiswanger, K., Standley, J., Czeizel, A. E.,

Deleyiannis, F., Christensen, K., Munger, R. G., Lie, R. T., Wilcox, A., Romitti, P. A., Field, L. L., Padilla, C. D., Cutiongco-de la Paz, E. M., Lidral, A. C., Valencia-Ramirez, L. C., Lopez-Palacio, A. M., Valencia, D. R., Arcos-Burgos, M., Castilla, E. E., Mereb, J. C., Poletta, F. A., Orioli, I. M., Carvalho, F. M., Hecht, J. T., Blanton, S. H., Buxó, C. J., Butali, A., Mossey, P. A., Adeyemo, W. L., James, O., Braimah, R. O., Aregbesola, B. S., Eshete, M. A., Deribew, M., Koruyucu, M., Seymen, F., Ma, L., de Salamanca, J. E., Weinberg, S. M., Moreno, L., Cornell, R. A., Murray, J. C., & Marazita, M. L. (2016) A genome-wide association study of nonsyndromic cleft palate identifies an etiologic missense variant in GRHL3. *Am J Hum Genet* **98**: 744-754.

Machado, R. A., Messetti, A. C., de Aquino, S. N., Martelli-Júnior, H., Swerts, M. S., de Almeida Reis, S. R., Moreira, H. S., Persuhn, D. C., & Coletta, R. D. (2016) Association between genes involved in craniofacial development and nonsyndromic cleft lip and/or palate in the Brazilian population. *Cleft Palate Craniofac J* **53**: 550-556.

Martelli-Junior, H., Porto, L. V., Martelli, D. R., Bonan, P. R., Freitas, A. B., & Della Coletta, R. (2007) Prevalence of nonsyndromic oral clefts in a reference hospital in the state of Minas Gerais, Brazil, between 2000-2005. *Braz Oral Res* **21**: 314-317.

Messetti, A. C., Machado, R. A., de Oliveira, C. E., Martelli-Júnior, H., de Almeida Reis, S. R., Moreira, H. S., Persuhn, D. C., Wu, T., & Coletta, R. D. (2017) Brazilian multicenter study of association between polymorphisms in CRISPLD2 and JARID2 and non-syndromic oral clefts. *J Oral Pathol Med* **46**: 232-239.

Munro, S. (2011) The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harb Perspect Biol* **3**: a005256.

Paranaiba, L. M., Bufalino, A., Martelli-Junior, H., de Barros, L. M., Graner, E., & Coletta, R. D. (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.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

Rosing, M., Ossendorf, E., Rak, A., & Barnekow, A. (2007) Giantin interacts with both the small GTPase Rab6 and Rab1. *Exp Cell Res* **313**: 2318-2325.

3 DISCUSSÃO

Como a discussão foi exaustivamente abordada em cada artigo, usamos esse espaço para contextualizar nossos resultados e discutir as perspectivas futuras para novos estudos.

Nos últimos anos têm ocorrido uma evolução no entendimento dos fatores genéticos relacionados à etiologia das FONS em todo o mundo. No entanto, a complexidade dos mecanismos biológicos envolvidos no desenvolvimento craniofacial e as evidências da heterogeneidade genética entre os sub-fenótipos das FONS e entre as diferentes populações dificultam a identificação dos exatos mecanismos genéticos. Neste sentido, é fundamental estudos de replicação das variantes genéticas identificadas como de susceptibilidade para as FONS em diferentes populações. Para pesquisas desta natureza junto à população brasileira que apresenta uma identidade genética modificada por mais de 500 anos em decorrência da miscigenação principalmente entre europeus, africanos e ameríndios, faz-se necessário inferir variáveis contínuas que estimam a ancestralidade genômica. O agrupamento de indivíduos pela cor de pele ou características físicas que possam indicar sua etnia não garante a separação dos grupos de forma homogênia (Pena *et al*., 2011; Durso *et al*., 2014). Dessa forma, a implicação desta miscigenação sobre os estudos que buscam associação de marcadores genéticos com fenótipos complexos como as FONS pode levar à obtenção de resultados espúrios. Apesar deste aspecto negativo, populações geradas pela miscigenação representam uma possibilidade para a identificação de genes responsáveis pela expressão de determinados fenótipos, quando genes se manifestam de maneira diferenciada nos grupos populacionais ancestrais contribuintes para a formação da população. O manuscrito de revisão sistemática deste estudo identificou diversos estudos com amostras pequenas e de diferentes regiões do país sem adequada avaliação da influência da ancestralidade, indicando que a união entre grupos de pesquisa, bem como a estratificação ancestral dos indivíduos possibilitaria o entendimento concreto da etiologia das FONS na população brasileira.

O resultado da meta-análise indicou uma possível associação do SNP rs642961 no gene *IRF6* e dos marcadores rs987525 e rs1530300 em uma região intergênica na região 8q24 com as FL±PNS na população brasileira, mas esses resultados devem ser observados com cuidado. Mundialmente têm-se observado

essas associações principalmente com populações de ascendência europeia (Birnbaum *et al*., 2009; Grant *et al*., 2009; Nikopensius *et al*., 2009; Beaty *et al*., 2010; Mangold *et al*., 2010; Ludwig *et al*., 2012; Moreno *et al*., 2017), no entanto na população brasileira a participação é discutida. Estudos anteriores do nosso grupo não demonstraram a associação do SNP rs642961 em *IRF6* com populações do sudeste do Brasil, onde há predominância europeia (Paranaíba *et al*., 2010), e em um grupo de amostras do estado da Bahia, onde há predomínio de indivíduos com ancestralidade africana (do Rego-Borges *et al*., 2015). Similarmente, outro estudo baseado em trios de diferentes regiões do Brasil não confirmou a associação de rs642961 em *IRF6* e o risco das FONS (de Souza et al., 2016). Por outro lado, Brito *et al*. (2012a) relataram uma associação significante entre rs642961 em *IRF6* e FLNS. Importante destacar que neste último estudo, a distribuição dos genótipos no grupo controle não respeitaram os preceitos do equilíbrio de Hardy-Weiberg, permitindo a inferência que problemas na genotipagem ocorreram. A associação de 8q24 (ambos rs987525 e rs1530300) com FL±PNS foi encontrada anteriormente em uma população brasileira com ancestralidade européia alta (Brito *et al*., 2012b; Bagordakis *et al*., 2013), mas o estudo de do Rego-Borges *et al*. (2015) demonstrou apenas uma tendência de associação entre rs987525 e rs1530300 e o risco da FL±PNS na população com alta ascendência africana.

A suplementação com ácido fólico é proposta como uma maneira de prevenção dos defeitos associados ao tubo neural e das FONS, visto a sua participação na síntese de nucleotídeos e aminoácidos e na metilação do DNA (Wilcox *et al*., 2007; Arruda *et al*., 2013). No entanto, estudos com suplementação de ácido fólico não encontraram uma diminuição na ocorrência das FONS (Ray *et al*., 2003; Lopez-Camelo *et al*., 2010; Wehby e Murray, 2010), o que sugere o envolvimento de variantes polimórficas em genes como *MTHFR* (methylenetetrahydrofolate reductase), que codifica uma proteína relacionada ao metabolismo do ácido fólico, ao risco aumentado para o surgimento das FONS (Bufalino *et al*., 2010; Bezerra *et al*., 2014; Aquino *et al*., 2014, de Aguiar *et al*., 2015; Wang *et al*., 2016). Embora marcadores polimórficos em *MTHFR* não tenham sido identificados como fatores de risco para as FONS em todos os GWAS realizados e em uma meta-análise usando dois dos maiores GWAS realizados até o momento (Birnbaum *et al*., 2009, Grant *et al*., 2009, Beaty *et al*., 2010, Mangold *et al*., 2010, Ludwig *et al*., 2012, Sun *et al*., 2015, Leslie *et al*., 2016a, Yu *et al*., 2017), há vários trabalhos com abordagem de

genotipagem direta relatando essa associação em diferentes populações (Pan *et al*., 2012; Zhao *et al*., 2014). Vários polimorfismos foram descritos em *MTHFR*, mas o rs1801133 é o mais estudado em FONS devido a redução da atividade enzimática e consequente significado clínico (Schwab *et al*., 2008; Afzal *et al*., 2009; Wang *et al*., 2014; Zhou *et al*., 2014; Hong *et al*., 2017; Wang *et al*., 2017). O rs1801133 (C>T) implica em uma substituição de alanina por uma valina no domínio catalítico, resultando em uma maior dissociação do cofator FAD devido ao deslocamento da estrutura quaternária da enzima (Frosst *et al*., 1995; Yamada *et al*., 2001; Pejchal *et al*., 2006) e em uma proteína termolábil com redução de 30% a 70% em sua atividade catalítica (Kang *et al*., 1991; Frosst *et al*., 1995). Nossa meta-análise confirmou a associação desse polimorfismo como fator de risco para FL±PNS na população brasileira.

Adicionalmente, a meta-análise revelou que o polimorfismo rs17563 no gene *BMP4* apresenta um efeito protetor para as FL±PNS. Esse polimorfismo foi estudado apenas em três estudos no Brasil (Araújo *et al*., 2012; Antunes *et al*., 2013, de Araújo *et al*., 2016), sendo que dois deles são do mesmo grupo de pesquisa (Araújo *et al*., 2012; de Araújo *et al*., 2016) e um não foi incluído na meta-análise devido ao fato de não ter apresentado a frequência dos genótipos entre os grupos FL±PNS e controle (de Araújo *et al*. 2016). O estudo de Araújo e colaboradores (2012) usaram apenas 123 amostras de FL±PNS e 246 controles de 4 regiões diferentes do país sem correção por ancestralidade. Após a inclusão de amostras de mais três estados do país, o mesmo grupo de pesquisa apresentou um estudo com 182 pacientes com FL±PNS e 355 controles e um efeito protetor foi confirmado (de Araújo *et al*., 2016). Antunes *et al*. (2013) usaram 324 amostras FL±PNS e 450 controles do estado do Rio de Janeiro e apenas uma associação com FLNS (71 amostras) foi evidenciada. Esse polimorfismo foi estudado em outras populações e na população de origem asiática, é considerado um fator de risco (Hu *et al*., 2015; Li *et al*., 2017), sugerindo assim que a confirmação do papel deste polimorfismo na população brasileira, com amostras maiores e com estratificação populacional, se faz necessário para o melhor entendimento.

A replicação de marcadores associados as FONS na população brasileira é uma estratégia simples para entender os fatores genéticos associados em nossa população. Com esse princípio, o manuscrito 2 avaliou os SNP rs7552 em 2q24.2, rs8049367 em 16p13.3, rs1880646, rs7406226, rs9891446 em 17p13, rs1588366 em

17q23.2 e rs73039426 em 19q13.11 em 1697 indivíduos brasileiros. Os 7 SNP foram previamente associados com FL±PNS nos recentes GWAS com populações multi-etnicas (Beaty *et al*., 2010; Sun *et al*., 2015; Leslie *et al*., 2016; Yu *et al*., 2017). O SNP rs7552 localizado na região 3' UTR do gene *FAM49A* e suas interações com os outros SNP desse estudo foram associados em nossa população. Embora não se conheça a real função da proteína codificada pelo gene *FAM49A*, a expressão em diversos tecidos foi identificada (Fagerberg *et al*., 2014). Variantes em 3' UTR podem influenciar a poliadenização, a eficiência da tradução, a localização e a estabilidade do mRNA (Króliczewski *et al*., 2018). Além disso, este SNP pode realizar uma função reguladora ou pode estar em desequilíbrio de ligação com outro gene ou região genética e não ser a verdadeira variante etiológica. Dessa forma, estudos adicionais são necessários para esclarecer a relação entre 2p24.2 e o desenvolvimento de FL±PNS, permitindo a elucidação de mecanismos pelos quais a variante rs7552 ou outra variante em desequilíbrio conduz a predisposição FL±PNS, traduzindo assim os achados de estudos de associação na clínica.

O manuscrito 2 também revelou a associação do haplótipo A-G formado pelos SNP rs1880646 e rs9891446 no gene *NTN1* como fator de risco para FL±PNS. *NTN1* codifica uma proteína conhecida como netrin 1, que promove a polaridade das células migratórias durante o desenvolvimento neural, vascular e na morfogênese das glândulas mamárias, pâncreas e pulmões (Ylivinkka *et al*., 2016; Wu *et al*., 2017). Por interagir com a proteína fosfatase 2A (PP2A) e a proteína quinase 1 associada à morte celular (DAPK1), sugere-se que netrin 1 pode desempenhar funções regulatórias, prevenindo a apoptose celular (Lahlali *et al*., 2016). Além disso, netrin 1 foi expressa em níveis elevados no mesêquima, especialmente ao longo das bordas dos processos palatinos, sugerindo seu envolvimento na fusão palatal (Leslie *et al*., 2015). Como estes SNP foram previamente associados apenas com FL±PNS e nenhuma associação com fissuras palatinas isoladas foi descrita em humanos até o momento, trabalhos futuros com esse fenótipo são necessários para entender essa associação.

Obviamente genes e seus produtos não agem isoladamente, eles são componentes de vias nas quais existem várias etapas que precisam ocorrer de maneira correta para que o processo aconteça de modo esperado. Desta forma, cada vez mais tem se estudado como os genes podem interagir entre si e como a combinação de variantes de dois ou mais genes na mesma via pode resultar em um efeito cumulativo. Exemplo disso foi observado no manuscrito 3 que embora nenhuma

interação individual entre os 28 SNPs em *SOD1*, *SOD2*, *SOD3*, *PON1*, *PON2* e *PON3* no núcleo familiar, múltiplas interações GxG entre os SNPs rs3917490, rs2299262 e rs2237583 em *PON1*, rs2286232, rs17166879 e rs2299267 em *PON2* e rs10953143, rs3757708, rs1053275 e rs978903 em *PON3* foram associadas na sucetibilidade para FL±PNS. Os genes *PON1*, *PON2* e *PON3* codificam enzimas importantes na manutenção dos níveis baixos de espécies reativas de oxigênio (ROS) e são expressos em quase todos os tecidos humanos (Précourt *et al*., 2011). Embora a expressão de *PONs* tem sido implicada na patogênese de defeitos do tubo neural, o presente estudo foi o primeiro realizado com FL±PNS.

Em adição, observou-se no estudo 3 que o alelo C do polimorfismo rs2237583 em *PON1* induz um efeito protetor para as FL±PNS e o SNP rs3917490 apresentou uma associação significante apenas em amostras com alta ancestralidade africana. O SNP rs3917490 está no íntron do gene *PON1* e não se conhece a função até o momento. O SNP rs2237583, que é associado com uma diminuição a atividade enzimática, tem sido apontado como um fator de risco para doenças vasculares (Yoshino *et al*., 2016; Yuan *et al*., 2017). Um estudo no norte da Califórnia identificou 12 SNP em *PON1* associados à redução da atividade enzimática, incluindo o SNP rs2237583. No entanto, este SNP estava em forte LD com rs705379 (Huen *et al*., 2011). O SNP rs705379 está fortemente associado a redução dos níveis da enzima e estudos *in vivo* sugerem que pode interromper a sequência de reconhecimento do fator Sp1, afetando a transcrição (Deakin *et al*., 2003; Huen *et al*., 2015). Interessantemente, o alelo C de rs705370 foi associado à expressão aumentada de *PON1* em afro-americanos, mas não em europeus (Chen *et al*., 2003; Holland *et al*., 2006; Rojas-Garcia *et al*., 2005). Em nosso estudo o SNP rs705379 não estava em LD com o rs2237583 e não foi associado com FL±PNS na abordagem de núcleo familiar com amostras de 5 estados diferentes do país. A abordagem em trio é eficaz para reduzir os efeitos da ancestralidade, por que utiliza como controle os pais não afetados da própria família. No entanto, não se consegue analisar a segregação por porcentagem de ascendência dos indivíduos, sendo importante um estudo caso-controle com separação dos pacientes com alta e baixa ancestralidade africana para para um melhor entendimento da participação do SNP rs705370 na etiologia das FONS.

Na revisão de literatura observamos 49 estudos com amostras de pacientes com FONS na população brasileira. Desses estudos, 29 tinham amostras de FPNS,

mas nenhum tinha sido realizado exclusivamente com amostras de FPNS até o momento na população brasileira. Interessante notar que esta tendência, muito em decorrência da menor prevalência deste tipo de fissura oral, é mundial. O estudo 4 foi dedicado exclusivamente as FPNS, na tentativa de verificar a associação de tag-SNP no gene *GOLGB1* com a etiologia das FPNS. No entanto, nossos resultados não detectaram tal associação. Sabendo das diferenças embrionárias e fatores etiológicos envolvidos na formação do palato, a identificação de novas regiões polimórficas em genes candidatos às FPNS é de fundamental importância para o entendimento dessa anomalia. A complexidade e a diversidade dos aspectos clínicos e mecanismos moleculares envolvidos no desenvolvimento craniofacial proporcionam inúmeras oportunidades para investigações futuras da etiologia da FPNS. Dois estudos de GWAS com amostras de pacientes FPNS foram realizados até o momento. Um deles apontou o marcador rs41268753 em *GRHL3* (*grainyhead like transcription factor 3*) como fator de risco para as FPNS (Leslie *et al*., 2016b). O outro demonstrou associações de inúmeros SNP em *MLLT3* (*MLLT3, super elongation complex subunit*) e *SMC2* (*structural maintenance of chromosomes 2*) e mães que consumiram álcool no primeiro trimestre de gravidez no risco de desenvolvimento das FPNS. Também foi encontrado fatores de risco para FPNS nas interações de SNP em *TBK1* (*TANK binding kinase 1*) e *ZNF236* (*zinc finger protein 236*) com o tabagismo materno e com risco diminuído de SNPs em *BAALC* (*BAALC, MAP3K1 and KLF4 binding*) com suplementação multivitamínica (Beaty *et al*., 2011).

De forma geral, os estudos genéticos demonstraram sucesso na caracterização de alguns genes e regiões genéticas com participação na etiologia das FONS, mas a grande maioria destes estudos concentrou-se apenas nos efeitos dos marcadores ou genes de forma individualizada (Birnbaum *et al*., 2009, Grant *et al*., 2009, Beaty *et al*., 2010, Mangold *et al*., 2010, Ludwig *et al*., 2012, Sun *et al*., 2015, Leslie *et al*., 2016a, Yu *et al*., 2017). Dentro do conceito etiológico multifatorial da doença, estudos de interações gene-gene (GxG) e gene-ambiente (GxE) devem contribuir para um melhor entendimento da etiologia das FONS (Cordell, 2009, Beaty *et al*., 2011, Letra *et al*., 2012, Wu *et al*., 2014, Li *et al*., 2015, Machado *et al*., 2016, Wang *et al*., 2018). Exemplos recentes são as interações entre os genes *BHMT*/*BHMT2* (*betaine-homocysteine methyltransferase*) e *DMGDH* (*dimethylglycine dehydrogenase*), que não foram apontados como etiológicos para as FONS quando analisados de maneira individual, mas em conjunto (interações GxG) demonstraram associações de risco

(Wang *et al*., 2018). O tabagismo materno é discutido como uma causa importante para as FONS, mas apenas em interação com o SNP rs1801321 no gene *RAD51* foi apresentado como fator de risco para FL±PNS na população brasileira (Machado *et al*., 2016). Os manuscritos 2 e 3 também revelaram interações entre dos SNP em 2p24.2, 16p13.3, 17p13, 17q23.2 e 19q13.11 e entre os SNP em *PON1*, *PON2* e *PON3* como possiveis fatores de risco para as FL±PNS na população brasileira. Desta forma, estes resultados devem servir de incentivo para novos estudos, principalmente com amostras maiores, para aumentar a compreensão e conhecimento dos eventos genéticos que contribuem para esta complexa alteração do desenvolvimento.

4 CONCLUSÃO

1. O estudo de revisão sistemática e meta-análise evidenciou associações entre rs642961 (*IRF6*), rs987525 e rs1530300 (8q24) e rs1801133 (*MTHFR*) com o risco de FL±PNS. O polimorfismo rs17563 em *BMP4* demonstrou um efeito de proteção para o desenvolvimento das FL±PNS. Estudos futuros com amostras maiores e com populações ancestralmente caracterizadas são importantes para melhor identificar as variantes genéticas com influência na etiologia das FONS na população brasileira.

2. Os genótipos formados pelo alelo A do SNP rs7552 em 2p24.2 e suas interações com rs8049367 (16p13.3), rs1880646 e rs9891446 (17p13), rs1588366 (17q23.2) e rs73039426 (19q13.11) são fatores de risco para desenvolvimento de FL±PNS na população brasileira. Adicionalmente, o haplótipo A-G formado pelos polimorfismos rs1880646 e rs9891446 no gene *NTN1* demonstrou associação significante com FL±PNS.

3. Múltiplas interações GxG foram observadas entre os SNP nos genes *PON1*, *PON2* e *PON3* no risco das FL±PNS, mas as interações contendo rs2237583 em *PON1* e rs17166879 em *PON2* foram as que demonstraram um valor de p resistente a correção para múltiplas comparações e ao teste de permutação. O alelo C do SNP rs2237583 foi associado com um efeito protetor para as FL±PNS, enquanto o SNP rs3917490 foi associado apenas nas amostras com alta ancestralidade africana na população brasileira.

4. Os tag-SNP em *GOLGB1* não foram associados ao risco das FPNS na população brasileira.

REFERÊNCIAS[*]

Afzal S, Jensen SA, Vainer B, Vogel U, Matsen JP, Sorensen JB, et al. MTHFR polymorphismsand 5-FU-based adjuvant chemotherapy in colorectal cancer. AnnOncol. 2009; 20(10):1660-1666.

Antunes LS, Küchler EC, Tannure PN, Costa MC, Gouvêa CV, Olej B, et al. BMP4 polymorphism is associated with nonsyndromic oral cleft in a Brazilian population. Cleft Palate Craniofac J. 2013; 50: 633-638.

Aquino SN, Hoshi R, Bagordakis E, Pucciarelli MG, Messetti AC, Moreira H, et al. MTHFR rs2274976 polymorphism is a risk marker for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol. 2014; 100(1): 30-5.

Araújo TK, Simioni M, Félix TM, de Souza LT, Fontes MÍ, Monlleó IL, et al. Preliminary analysis of the nonsynonymous polymorphism rs17563 in BMP4 gene in Brazilian population suggests protection for nonsyndromic cleft lip and palate. Plast Surg Int. 2012; 2012: 247104.

Arruda IT, Persuhn DC, de Oliveira NF. The MTHFR C677T polymorphism and global DNA methylation in oral epithelial cells. Genet Mol Biol. 2013; 36(4): 490-3.

Bagordakis E, Paranaiba LM, Brito LA, de Aquino SN, Messetti AC, Martelli-Junior H, et al. 2013. Polymorphisms at regions 1p22.1 (rs560426) and 8q24(rs1530300) are risk markers for nonsyndromic cleft lip and/or palatein the Brazilian population. Am J Med Genet Part A. 2013; 161A: 1177-1180.

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 2010; 42(6): 525-9.

[*] De acordo com as normas da UNICAMP/FOP, baseadas na padronização do International Committee of Medical Journal Editors - Vancouver Group. Abreviatura dos periódicos em conformidade com o PubMed.

Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, et al. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. Genet Epidemiol. 2011; 35: 469-478.

Beaty TH, Taub MA, Scott AF, Murray JC, Marazita ML, Schwender H, et al. Confirming genes influencing risk to cleft lip with/without cleft palate in a case-parent trio study. Hum Genet. 2013;132(7): 771-81.

Bezerra J, Oliveira G, Soares C, Cardoso M, Ururahy M, Neto F, et al. Genetic and non-genetic factors that increase the risk of non-syndromic cleft lip and/or palate development. Oral Dis. 2014; 1-7.

Bille C, Winther JF, Bautz A, Murray JC, Olsen J, Christensen K. Cancer risk in persons with oral cleft--a population-based study of 8,093 cases. Am J Epidemiol. 2005; 161(11): 1047-55.

Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, et al. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 2009; 41(4): 473-7.

Brewer C, Holloway S, Zawalnyski P, Schinzel A, Fitz-Patrick D. A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality - and tolerance of segmental aneuploidy - in humans. Am J Hum Genet. 1999; 64(6): 1702-1708.

Brito LA, Bassi CF, Masotti C, Malcher C, Rocha KM, Schlesinger D, et al. IRF6 is a risk factor for nonsyndromic cleft lip in the Brazilian population. Am J Med Genet A. 2012a; 158A(9): 2170-5.

Brito LA, Paranaíba LM, Bassi CF, Masotti C, Malcher C, Schlesinger D, et al. Region 8q24 is a susceptibility locus for nonsyndromic oral clefting in Brazil. Birth Defects Res A Clin Mol Teratol. 2012b; 94(6): 464-8.

Broder HL, Richman LC, Matheson PB. Learning disability, school achievement, and grade retention among children with cleft: a two-center study. Cleft Palate Craniofac J. 1998; 35(2): 127-31.

Bufalino A, Paranaiba LM, de Aquino S, Martelli-Junior H, Swerts MS, Coletta RD. Maternal polymorphisms in folic acid metabolic genes are associated with

nonsyndromic cleft lip and/or palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol. 2010; 88(11): 980-6.

Cardoso EF, Martelli DR, Machado RA, Coletta RD, de Souza JD, Barbosa FT, et al. Nonsyndromic cleft lip and palate, gastric cancer and tooth agenesis. Med Oral Patol Oral Cir Bucal. 2018; 23(1): e44-e48.

Chen J, Kumar M, Chan W, Berkowitz G, Wetmur JG. Increased influence of genetic variation on PON1 activity in neonates. Environ Health Perspect. 2003; 111(11): 1403-1409.

Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. Nat Protoc. 2011; 6(2): 121-133.

Cordell HJ. Detecting gene-gene interactions that underlie human diseases. Nat Rev Genet. 2009; 10(6): 392-404.

De Aguiar PK, Coletta RD, de Oliveira AM, Machado RA, Furtado PG, de Oliveira LA, et al. rs1801133C>T polymorphism in MTHFR is a risk factor for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol. 2015; 103(4): 292-8.

De Souza LT, Kowalski TW, Ferrari J, Monlléo IL, Ribeiro EM, de Souza J, et al. Study of IRF6 and 8q24 region in non-syndromic oral clefts in the Brazilian population. Oral Dis. 2016; 22(3): 241-5.

Deakin S, Leviev I, Brulhart-Meynet MC, James RW. Paraoxonase-1 promoter haplotypes and serum paraoxonase: a predominant role for polymorphic position - 107, implicating the Sp1 transcription factor. Biochem J. 2003; 372(Pt 2): 643-649.

Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011; 12(3): 167-78.

Do Rego Borges A, Sá J, Hoshi R, Viena CS, Mariano LC, de Castro Viega, et al. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry. Am J Med Genet A. 2015; 167A(10):2344-9.

Durso DF, Bydlowski SP, Hutz MH, Suarez-Kurtz G, Magalhães TR, Pena SD. Association of genetic variants with self-assessed color categories in Brazilians. PLoS One. 2014; 9(1): e83926.

Economou AD, Ohazama A, Porntaveetus T, Sharpe PT, Kondo S, Basson MA, et al. Periodic stripe formation by a Turing mechanism operating at growth zones in the mammalian palate. Nat Genet. 2012; 44(3): 348-51.

Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014; 13(2): 397-406.

Feitosa MF, Krieger H. Future of genetic epidemiology in complex traits. Ciência & Saude Coletiva. 2002; 7(1): 73-83.

Fernandez CCA, Pereira CV, Luiz RR, Faraco IM, Marazita ML, Arnaudo M, et al. Third molar agenesis as a potential marker for craniofacial deformities. Arch Oral Biol. 2018; 88: 19-23.

Fogh-Andersen P. Inheritance of harelip and cleft palate, Busck, Copenhagen, 1942.

Foulkes AS. Applied statistical genetics with R: for population-based association studies. Springer; 2009.

Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, et al. A candidate genetic risk factor for vascular disease: a common mutation in methylenete-trahydrofolate reductase. Nat Genet. 1995; 10(1):111-113.

Gaspar DA, Matioli SR, de Cássia Pavanello R, Araújo BC, Alonso N, Wyszynski D, et al. Maternal MTHFR interacts with the offspring's BCL3 genotypes, but not with TGFA, in increasing risk to nonsyndromic cleft lip with or without cleft palate. Eur J Hum Genet. 2004; 12: 521-526.

Gaspar DA, Pavanello RC, Zatz M, Passos-Bueno MR, André M, Steman S, et al. Role of the C677T polymorphism at the MTHFR gene on risk to nonsyndromic cleft lip with/without cleft palate: results from a case-control study in Brazil. Am J Med Genet. 1999; 87: 197-199.

Gonçalves E, Martelli DR, Coletta RD, Vieira AR, Caldeira AP, Martelli H Jr. Risk of leukemia in first degree relatives of patients with nonsyndromic cleft lip and palate. Braz Oral Res. 2014; 28: 1-3.

Gong SG, Gong TW, Shum L. Identification of markers of the midface. J Dent Res. 2005; 84(1): 69-72.

Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, et al. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr. 2009; 155(6): 909-13.

Higgins AW, Alkuraya FS, Bosco AF, Brown KK, Bruns GA, Donovan DJ, et al. Characterization of apparently balanced chromosomal rearrangements from the developmental genome anatomy project. Am J Hum Genet. 2008; 82(3): 712-722.

Holland N, Furlong C, Bastaki M, Richter R, Bradman A, Huen K, et al. Paraoxonase polymorphisms, haplotypes, and enzyme activity in Latino mothers and newborns. Environ Health Perspect. 2006; 114(7): 985-91.

Hong HH, Hu Y, Yu YQ, Zhou L, Lv MQ, Sun Y, et al. Associations of C677T polymorphism in methylenetetrahydrofolate reductase (MTHFR) gene with male infertility risk: A meta-analysis. Eur J Obstet Gynecol Reprod Biol. 2017; 212: 101-109.

Hu YY, Qin CQ, Deng MH, Niu YM, Long X. Association between BMP4 rs17563 polymorphism and NSCL/P risk: a meta-analysis. Dis Markers. 2015; 2015: 763090.

Huen K, Barcellos L, Beckman K, Rose S, Eskenazi B, Holland N. Effects of PON polymorphisms and haplotypes on molecular phenotype in Mexican-American mothers and children. Environ Mol Mutagen. 2011; 52(2): 105-16.

Huen K, Yousefi P, Street K, Eskenazi B, Holland N. PON1 as a model for integration of genetic, epigenetic, and expression data on candidate susceptibility genes. Environ Epigenet. 2015;1(1). pii: dvv003.

Jianyan L, Zeqiang G, Yongjuan C, Kaihong D, Bing D, Rongsheng L. Analysis of interactions between genetic variants of BMP4 and environmental factors with nonsyndromic cleft lip with or without cleft palate susceptibility. Int J Oral Maxillofac Surg. 2010; 39(1): 50-6.

Jindal A, Vieira AR. Family history of cleft lip and palate in subjects diagnosed with leukemia. Am J Med Genet A. 2012; 158A(3): 678-9.

Jones MC. Etiology of facial clefts: prospective evaluation of 428 patients. Cleft Palate J. 1988; 25(1): 16-20.

Jugessur A, Shi M, Gjessing HK, Lie RT, Wilcox AJ, Weinberg CR, et al. Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia. PLoS One. 2009; 4(4): e5385.

Juriloff DM, Harris MJ. Mouse genetic models of cleft lip with or without cleft palate. Birth Defects Res A Clin Mol Teratol. 2008; 82(2): 63-77.

Kang SS, Wong PW, Susmano A, et al. 1991. Thermolabile methylenetetrahydrofolate reductase: an inherited risk factor for coronary artery disease. Am J Hum Genet. 1991; 48(3): 536-545.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, et al. Mutation in IRF6 cause van der Woude and popliteal pterygium syndromes. Nat Genet. 2002; 32(2): 285-289.

Króliczewski J, Sobolewska A, Lejnowski D, Collawn JF, Bartoszewski R. microRNA single polynucleotide polymorphism influences on microRNA biogenesis and mRNA target specificity. Gene. 2018; 640: 66-72.

Lahlali T, Plissonnier ML, Romero-López C, Michelet M, Ducarouge B, Berzal-Herranz A, et al. Netrin-1 Protects Hepatocytes Against Cell Death Through Sustained Translation During the Unfolded Protein Response. Cell Mol Gastroenterol Hepatol. 2016; 2(3): 281-301.e9.

Leslie EJ, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, et al. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. Hum Mol Genet. 2016a; 25(13): 2862-2872.

Leslie EJ, Liu H, Carlson JC, Shaffer JR, Feingold E, Wehby G, et al. A Genome-wide Association Study of Nonsyndromic Cleft Palate Identifies an Etiologic Missense Variant in GRHL3. Am J Hum Genet. 2016b; 98(4): 744-54.

Leslie EJ, Taub MA, Liu H, Steinberg KM, Koboldt DC, Zhang Q, et al. Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. Am J Hum Genet. 2015; 96(3): 397-411.

Letra A, Fakhouri W, Fonseca RF, Menezes R, Kempa I, Prasad JL, et al. Interaction between IRF6 and TGFA genes contribute to the risk of nonsyndromic cleft lip/palate. PLoS One. 2012; 7(9): e45441.

Li Q, Kim Y, Suktitipat B, Hetmanski JB, Marazita ML, Duggal P, et al. Gene-Gene Interaction Among WNT Genes for Oral Cleft in Trios. Genet Epidemiol. 2015; 39(5): 385-394.

Li YH, Yang J, Zhang JL, Liu JQ, Zheng Z, Hu DH. BMP4 rs17563 polymorphism and nonsyndromic cleft lip with or without cleft palate: A meta-analysis. Medicine (Baltimore). 2017; 96: e7676.

Lima LS, Ribeiro GS, de Aquino SN, Volpe FM, Martelli DR, Swerts MS, et al. Prevalence of depressive symptoms in patients with cleft lip and palate. Braz J Otorhinolaryngol. 2015; 81(2): 177-83.

Lopez-Camelo JS, Castilla EE, Orioli IM. Folic acid flour fortification: impact on the frequencies of 52 congenital anomaly types in three South American countries. Am J Med Genet A. 2010; 152a(10): 2444-58.

Ludwig KU, Ahmed ST, Bohmer AC, Sangani NB, Varghese S, Klamt J, et al. Meta-analysis Reveals Genome-Wide Significance at 15q13 for Nonsyndromic Clefting of Both the Lip and the Palate, and Functional Analyses Implicate GREM1 As a Plausible Causative Gene. PLoS Genet. 2016; 12(3):e1005914.

Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, et al. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet. 2012; 44(9): 968-71.

Machado RA, Freitas EM, de Aquino SN, Martelli DR, Swerts MS, Reis SR, et al. Clinical relevance of breast and gastric cancer-associated polymorphisms as potential susceptibility markers for oral clefts in the Brazilian population. BMC Medical Genetics. 2017; 18(1): 39.

Machado RA, Moreira HS, de Aquino SN, Martelli-Júnior H, Reis SR, Persuhn DC, et al. Interactions between RAD51 rs1801321 and maternal cigarette smoking as risk factor for nonsyndromic cleft lip with or without cleft palate. Am J Med Genet. 2016; 179A(2): 536-9.

Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, et al. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet. 2010; 42: 24-6.

Mangold E, Ludwig KU, Nothen MM. Breakthroughs in the genetics of orofacial clefting. Trends Mol Med. 2011; 17(12): 725-33.

Marazita ML, Field LL, Cooper ME, Tobias R, Maher BS, Peanchitlertkajorn S, et al. Nonsyndromic cleft lip with or without cleft palate in China: assessment of candidate regions. Cleft Palate Craniofac J. 2002a; 39(2): 149-56.

Marazita ML, Field LL, Cooper ME, Tobias R, Maher BS, Peanchitlertkajorn S, et al. Genome scan for loci involved in cleft lip with or without cleft palate, in Chinese multiplex families. Am J Hum Genet. 2002b; 71(2): 349-64.

Marazita ML, Lidral AC, Murray JC, Field LL, Maher BS, Goldstein McHenry T, et al. Genome scan, fine-mapping, and candidate gene analysis of non-syndromic cleft lip with or without cleft palate reveals phenotype-specific differences in linkage and association results. Hum Hered. 2009; 68(3): 151-70.

Marazita ML, Mooney MP. Current concepts in the embryology and genetics of cleft lip and cleft palate. Clin Plast Surg. 2004; 31(2): 125-40.

Marazita ML. The evolution of human genetic studies of cleft lip and cleft palate. Annu Rev Genomics Hum Genet. 2012; 13: 263-83.

Martelli DR, Vieira AR, Fonseca AT, Coletta RD, Soares PB, Martelli-Júnior. Risk of nonsyndromic cleft lip and palate in relatives of women with breast cancer. Am J Med Genet A. 2014; 164A(1): 270-1.

Martelli-Junior H, Porto LV, Martelli DR, Bonan PR, Freitas AB, Della Coletta R. Prevalence of nonsyndromic oral clefts in a reference hospital in the state of Minas Gerais, Brazil, between 2000-2005. Braz Oral Res. 2007; 21(4): 314-7.

Melo-Filho MR, Nogueira LA, Martelli DR, Silveira MF, Silva MS, de Barros, et al. Taurodontism in patients with nonsyndromic cleft lip and palate in a Brazilian population: a case control evaluation with panoramic radiographs. Oral Surg Oral Med Oral Pathol Oral Radiol. 2015; 120(6): 744-50.

Montagnoli LC, Barbieri MA, Bettiol H, Marques IL, de Souza L. Growth impairment of children with different types of lip and palate clefts in the first 2 years of life: a cross-sectional study. J Pediatr (Rio J). 2005; 81(6): 461-5.

Moore JH, Williams SM. Epistasis and its implications for personal genetics. Am J Hum Genet. 2009; 85(3): 309-320.

Moore JH, Williams SM. Epistasis: Methods and Protocols. Springer New York, 2015.

Moreno LM, Fomina T, Munger RG, Romitt PA, Jenkin MM, Gjessing HK, et al. A Population-Based Study of Effects of Genetic Loci on Orofacial Clefts. J Dent Res. 2017; 96(11): 1322-1329.

Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, et al. FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. Hum Mol Genet. 2009; 18(24): 4879-96.

Mossey PA, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. Lancet. 2009; 374(9703): 1773-85.

Mossey PA, Little J. Epidemiology of oral clefts: an international perspective. In: Wyszynski DF, ed. Cleft lip and palate: from origins to treatment. New York: Oxford University Press. 2002: 127-158.

Mostowska A, Hozyasz KK, Wojcicki P, Biedziak B, Paradowska P, Jagodzinski PP. 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. 2010; 88(7): 538-45.

Mukhopadhyay P, Greene RM, Zacharias W, Weinrich MC, Singh S, Young WW et al. Developmental gene expression profiling of mammalian, fetal orofacial tissue. Birth Defects Res A Clin Mol Teratol. 2004; 70(12): 912-926.

Murthy J, Bhaskar L. Current concepts in genetics of nonsyndromic clefts. Indian J Plast Surg. 2009; 42(1): 68-81.

Nikopensius T, Ambrozaityte L, Ludwing KU, Birnbaum S, Jagomagi T, Saag M, et al. 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. 2009; 149A(11): 2551-3.

Nopoulos P, Langbehn DR, Canady J, Magnotta V, Richman L. Abnormal brain structure in children with isolated clefts of the lip or palate. Arch Pediatr Adolesc Med. 2007; 161(8): 753-8.

Pan Y, Zhang W, Ma J, Du Y, Li D, Cai Q, et al. Infants' MTHFR polymorphisms and nonsyndromic orofacial clefts susceptibility: a meta-analysis based on 17 case-control studies. Am J Med Genet A. 2012; 158A: 2162-2169.

Paranaíba LM, Bufalino A, Martelli-Júnior H, de Barros LM, Graner E, Coletta RD. Lack of association between IRF6 polymorphisms (rs2235371 and rs642961) and non-syndromic cleft lip and/or palate in a Brazilian population. Oral Dis. 2010; 16(2):193-197.

Paranaiba LM, Coletta RD, Swerts MS, Quintino RP, de Barros LM, Martelli-Junior H. Prevalence of Dental Anomalies in Patients With Nonsyndromic Cleft Lip and/or Palate in a Brazilian Population. Cleft Palate Craniofac J. 2013; 50(4): 400-5.

Pejchal R, Campbell E, Guenther BD, Lennon BW, Matthews RG, Ludwig ML. Structural perturbations in the Ala>Val polymorphism of methylenetetrahydrofolate reductase: how binding of folates may protect against inactivation. Biochemistry. 2006; 45(15): 4808-4818.

Pena SD, Bastos-Rodrigues L, Pimenta JR, Bydlowski SP. DNA tests probe the genomic ancestry of Brazilians. Braz J Med Biol Res. 2009; 42(10): 870-6.

Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy FS, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. PLoS One. 2011; 6(2): e17063.

Pierce BA. Genética: um enfoque conceitual. 3rd. ed. Rio de Janeiro, RJ, Brasil: Guanabara Koogan, 2013.

Précourt LP, Amre D, Denis MC, Lavoie JC, Delvin E, Seidman E, et al. The three-gene paraoxonase family: physiologic roles, actions and regulation. Atherosclerosis. 2011; 214(1): 20-36.

Purcell S, Sham P, Daly MJ. Parental phenotypes in family-based association analysis. Am J Hum Genet. 2005; 76(2): 249-59.

Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, et al. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. Nat Genet. 2008; 40(11): 1341-7.

Ray JG, Meier C, Vermeulen MJ, Wyatt PR, Cole DE. Association between folic acid food fortification and congenital orofacial clefts. J Pediatr. 2003; 143(6): 805-7.

Rice DP. Craniofacial anomalies: from development to molecular pathogenesis. Curr Mol Med. 2005; 5(7): 699-722.

Risch NJ. Searching for genetic determinants in the new millennium. Nature. 2000; 405(6788): 847-56.

Rodrigues K, Sena MF, Roncalli AG, Ferreira MA. Prevalence of orofacial clefts and social factors in Brazil. Braz Oral Res. 2009; 23(1): 38-42.

Rojas-Garcia AE, Solis-Heredia MJ, Pina-Guzman B, Vega L, Lopez-Carrillo L, Quintanilla-Vega B. Genetic polymorphisms and activity of PON1 in a Mexican population. Toxicol Appl Pharmacol. 2005; 205(3): 282-9.

Rojas-Martinez A, Reutter H, Chacon-Camacho O, Leon-Cachon RB, Munoz-Jimenez SG, Nowak S, et al. Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Mesoamerican population: Evidence for IRF6 and variants at 8q24 and 10q25. Birth Defects Res A Clin Mol Teratol. 2010; 88(7): 535-7.

Sá J, Araújo L, Guimarães L, Maranhão S, Lopes G, Medrado A, et al. Dental anomalies inside the cleft region in individuals with nonsyndromic cleft lip with or without cleft palate. Medicina Oral, Patología Oral y Cirugía Bucal. 2016a; 21(1): e48-e52.

Sá J, Mariano LC, Canguçu D, Coutinho TS, Hoshi R, Medrado AP, et al. Dental Anomalies in a Brazilian Cleft Population. Cleft Palate Craniofac J. 2016b; 53(6): 714-719.

Schwab M, Zanger UM, Marx C, Schaeffeler E, Klein K, Dippon J, et al. Role of genetic and nongenetic factors for fluorouracil treatment-related severe toxicity: a prospective clinical trial by the German 5-FU Toxicity Study Group. J Clin Oncol. 2008; 26:2131-2138

Sebro R, Rogus JJ. The power of the Transmission Disequilibrium Test in the presence of population stratification. Eur J Hum Genet. 2010; 18(9): 1032-8.

Shkoukani MA, Chen M, Vong A. Cleft lip - a comprehensive review. Front Pediatr. 2013; 1: 53.

Stanier P, Moore GE. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. Hum Mol Genet. 2004; 13 Spec No 1: R73-81.

Sun Y, Huang Y, Yin A, Pan Y, Wang Y, Wang C, et al., Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. Nat Commun. 2015; 6: 6414.

Taioli E, Ragin C, Robertson L, Linkov F, Thurman NE, Vieira AR. Cleft lip and palate in family members of cancer survivors. Cancer Invest. 2010; 28(9): 958-62.

Trindade L, Gonzales RM, Beck CL, Lautert L. [Workloads in communitarian health agents]. Rev Gaucha Enferm. 2007; 28(4): 473-9.

Vieira AR, Khaliq S, Lace B. Risk of cancer in relatives of children born with isolated cleft lip and palate. Am J Med Genet A. 2012; 158A(6):1503-4.

Vieira AR, McHenry TG, Daack-Hirsch S, Murray JC, Marazita ML. A Genome Wide Linkage Scan for Cleft Lip and Palate and Dental Anomalies. Am J Med Geneti A. 2008; 146A(11): 1406-13.

Wang C, Xie H, Lu D, Ling Q, Jin P, Li H, et al. The MTHFR polymorphism affect the susceptibility of HCC and the prognosis of HCC liver transplantation. Clin Transl Oncol. Em publicação. 2017 Nov 28.

Wang H, Hu C, Xiao SH, Wan B. Association of taggingSNPs in the MTHFR gene with risk of type 2 diabetes mellitusand serum homocysteine levels in a Chinese population. DisMarkers. 2014; 2014:725731.

Wang P, Wu T, Schwender H, Wang H, Shi B, Wang Z, et al. Evidence of interaction between genes in the folate/homocysteine metabolic pathway in controlling risk of nonsyndromic oral cleft. Oral Dis. Em publicação. 2018 Jan 22

Wang W, Jiao XH, Wang XP, Sun YY, Dong C. MTR, MTRR, and MTHFR Gene Polymorphisms and Susceptibility to Nonsyndromic Cleft Lip With or Without Cleft Palate. Genet Test Mol Biomarkers. 2016; 20(6): 297-303.

Wehby GL, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. Oral Dis. 2010; 16(1): 3-10.

Wehby GL, Castilla EE, Goco N, Rittler M, Cosentino V, Javois L, et al. The effect of systematic pediatric care on neonatal mortality and hospitalizations of infants born with oral clefts. BMC Pediatr. 2011; 11: 121.

Wehby GL, Murray JC. Folic acid and orofacial clefts: a review of the evidence. Oral Dis. 2010; 16(1): 11-9.

Weinberg SM, Parsons TE, Fogel MR, Walter CP, Conrad AL, Nopoulos P. Corpus callosum shape is altered in individuals with nonsyndromic cleft lip and palate. Am J Med Genet A. 2013; 161a(5): 1002-7.

Wilcox AJ, Lie RT, Solvoll K, Taylor J, McConnaughey DR, Abyholm F, et al. Folic acid supplements and risk of facial clefts: national population based case-control study. Bmj. 2007; 334(7591): 464.

Wu T, Schwender H, Ruczinski I, Murray JC, Marazita ML, Munger RG, et al. Evidence of gene-environment interaction for two genes on chromosome 4 and environmental tobacco smoke in controlling the risk of nonsyndromic cleft palate. PLoS One. 2014; 9(2): e88088.

Wu W, Lei H, Shen J, Tang L. The role of netrin-1 in angiogenesis and diabetic retinopathy: a promising therapeutic strategy. Discov Med. 2017; 23(128): 315-323.

Xia Y, Hu B, Chen J, Zheng L, Song J. Association between the IRF6 rs2235371 polymorphism and the risk of nonsyndromic cleft lip with or without cleft palate in Chinese Han populations: A meta-analysis. Arch Oral Biol. 2017; 84: 161-168.

Yamada K, Chen Z, Rozen R, Matthews RG. Effects of common polymorphisms on the properties of human methylenetetrahydrofolate reductase. Proc Natl Acad Sci USA. 2001; 98(26): 14853-14858.

Ylivinkka I, Keski-Oja J, Hyytiäinen M. Netrin-1: A regulator of cancer cell motility? Eur J Cell Biol. 2016; 95(11): 513-520.

Yoshino S, Cilluffo R, Prasad M, Best PJ, Atkinson EJ, Aoki T, et al. Sex-Specific Genetic Variants are Associated With Coronary Endothelial Dysfunction. J Am Heart Assoc. 2016; 5(4): e002544.

Yu Y, Zuo X, He M, Gao J, Fu Y, Qin C. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity. Nat Commun. 2017. 8: 14364.

Yuan X, He X, Zhang X, Liu C, Wang C, Qiu L, et al. Comparative Study of Folic Acid and a-Naphthoflavone on Reducing TCDD Induced Cleft Palate in Fetal Mice. Cleft Palate Craniofac J. 2017; 54(2): 216-222.

Zhao M, Ren Y, Shen L, Zhang Y, Zhou B. Association between MTHFR C677T and A1298C polymorphisms and NSCL/P risk in Asians: a meta-analysis. PLoS One. 2014; 9: e88242.

Zhou BS, Bu GY, Li M, Chang BG, Zhou YP. Tagging SNPs inthe MTHFR gene and risk of ischemic stroke in a Chinese population. Int J Mol Sci. 2014; 15(5): 8931-8940.

Zhu JL, Basso O, Hasle H, Winther JF, Olsen JH, Olsen J. Do parents of children with congenital malformations have a higher cancer risk? a nationwide study in Denmark. Br J Cancer. 2002; 87(5): 524-8.

Zucchero T, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, et al Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. Engl J Med. 2004; 351(8): 769-780.

ANEXOS

Anexo 1 – Artigo aprovado no periódico *Birth Defects Research*

Date: 22 January 2018
Manuscript number: BDR-17-0240.R1
Potential genetic markers for nonsyndromic oral clefts in the Brazilian population: a systematic review and meta-analysis

Dear Professor Coletta:

I am pleased to inform you that your revised , "Potential genetic markers for nonsyndromic oral clefts in the Brazilian population: a systematic review and meta-analysis," by Assis Machado, Renato; Toledo, Isabela; Martelli-Junior, Hercilio; Reis, Silvia; Neves, Eliete; Coletta, Ricardo, is acceptable for publication in Birth Defects Research.

Your article cannot be published until the publisher has received the appropriate signed license agreement. Within the next few days the corresponding author will receive an email from Wiley's Author Services system which will ask them to log in and will present them with the appropriate license for completion.

Please note that you will receive email notification when your proofs are available to download and correct. Once all of the articles for a given issue have been proofed and corrected, the contents of the issue will be organized according to the print version and posted as the current issue at http://www3.interscience.wiley.com/cgi-bin/jhome/102526943.

Thank you for submitting your manuscript to Birth Defects Research. I look forward to seeing more of your work in the future.

Sincerely,

Michel Vekemans MD,PhD
EIC
Birth Defects Research

P.S. Bring your research to life by creating a video abstract for your article! Wiley partners with Research Square to offer a service of professionally produced video abstracts. Learn more about video abstracts at http://www.wileyauthors.com/videoabstracts and purchase one for your article at https://www.researchsquare.com/wiley or through your Author Services account. If you have any questions, please direct them to videoabstracts@wiley.com.

Anexo 2 – Artigo aprovado no periódico *Clinical Genetics*

Clinical Genetics <onbehalfof@manuscriptcentral.com> 25 de fev às 03:53

Para: renatoassismachado@yahoo.com.br, evelyn-n@hotmail.com, hmjunior2000@yahoo.com, srareis@uol.com.br, darlenecp@hotmail.com e 1 mais...

Dear Dr. Ricardo D Coletta:

I am very pleased to inform you that your manuscript entitled "2p24.2 (rs7552) is a susceptibility locus for nonsyndromic cleft lip with or without cleft palate in the Brazilian population" (CGE-00033-2018.R1), has been received, accepted, and will be forwarded for publication in Clinical Genetics.

However, your article cannot be published until the publisher has received the appropriate signed license agreement. Within the next few days the corresponding author will receive an email from Wiley's Author Services system which will ask them to log in and will present them with the appropriate license for completion.

If there are any colour figures in your manuscript, please post/courier a colour cost agreement form as soon as possible to

Customer Service (OPI)
John Wiley & Sons Ltd, European Distribution Centre
New Era Estate
Oldlands Way
Bognor Regis
West Sussex
PO22 9NQ

Anexo 3 – Artigo submetido ao periódico *Journal of Dental Research*

Journal of Dental Research <onbehalfof@manuscriptcentral.com> 18 de abr às 08:54
Para: coletta@fop.unicamp.br
Cc: renatoassismachado@yahoo.com.br, carolstz59@hotmail.com, darlenecp@hotmail.com, vanessa_mjp@hotmail.com, srareis@uol.com.br e 5 mais...

18-Apr-2018

Dear Dr. Coletta:

Your manuscript, "Superoxide Dismutase and Paraoxonase Variants in Nonsyndromic Oral Clefts," has been successfully submitted online to the Journal of Dental Research.

Your manuscript ID is JDR-18-0343.

Please mention the above manuscript ID in all future correspondence or when calling the office for questions. If there are any changes in your contact information, please log in to SAGETrack at https://mc.manuscriptcentral.com/jdr and edit your user information as needed. You can also view the status of your manuscript at any time by checking your Author Center after logging into system.

Co-Authors: You are receiving this message because a manuscript has been submitted to Journal of Dental Research listing you as a co-author on JDR-18-0343, "Superoxide Dismutase and Paraoxonase Variants in Nonsyndromic Oral Clefts." If you feel you have received this notification in error, please contact the editorial office at jdr@iadr.org.

As part of our commitment to ensuring an ethical, transparent and fair peer review process SAGE is a supporting member of ORCID, the Open Researcher and Contributor ID (https://orcid.org/). We encourage all authors and co-authors to use ORCID iDs during the peer review process. If you already have an ORCID iD you can link this to your account in ScholarOne just by logging in and editing your account information. If you do not already have an ORCID iD you may login to your ScholarOne account to create your unique identifier and automatically add it to your profile.

Thank you for submitting your manuscript to Journal of Dental Research.

Sincerely,

Editorial Staff
Journal of Dental Research

Anexo 4 – Artigo aprovado no periódico *Annals of Human Genetics*

Annals of Human Genetics <onbehalfof@manuscriptcentral.com> 12 de dez de 2017 às 18:31
Para: coletta@fop.unicamp.br
Cc: renatoassismachado@yahoo.com.br, hmjunior2000@yahoo.com, srareis@uol.com.br, darlenecp@hotmail.com, coletta@fop.unicamp.br

Dear Professor Coletta,

I am pleased to inform you that your manuscript has now been accepted for publication in the Annals of Human Genetics.

*Please note that in order to process your article for publication, we will require the following:

1) An editable version of the text of your manuscript.

If a file of this type was not submitted to ScholarOne Manuscripts for the latest version of your manuscript, please ensure that an editable file (e.g. Word) is emailed as an attachment to the editorial office at AHG.journal@wiley.com.

2) Individual files of each figure in the following format:
---Line art such as charts, graphs and illustrations should be in EPS or PDF format
---Photographic images should be in TIFF format at a resolution of at least 300dpi at final size.

Detailed information on our digital illustration standards is available at:
http://media.wiley.com/assets/7323/92/electronic_artwork_guidelines.pdf.

Supplementary figures for online publication may be in alternative formats.

Files should be sent as email attachments to the the editorial office at AHG.journal@wiley.com. If files are too large to be sent by email, ftp transfer can be made available.