NIH Public Access
Author Manuscript
Eur J Med Genet. Author manuscript; available in PMC 2013 October 01.

Published in final edited form as:
Eur J Med Genet. 2012 October ; 55(10): 510–514. doi:10.1016/j.ejmg.2012.06.005.

GWAS reveals new recessive loci associated with non-syndromic facial clefting

Mauricio Camargo[a], **Dora Rivera**[a], **Lina Moreno**[b,c], **Andrew C. Lidral**[b], **Ursula Harper**[d], **Marypat Jones**[d], **Benjamin D. Solomon**[e], **Erich Roessler**[e], **Jorge I. Vélez**[e], **Ariel F. Martinez**[e], **Settara C. Chandrasekharappa**[d], and **Mauricio Arcos-Burgos**[e,f,*]

[a]Population Genetics and Mutacarcinogenesis Group, University of Antioquia, Medellin, Colombia

[b]Department of Orthodontics, College of Dentistry, University of Iowa, Iowa City, IA, USA

[c]Dows Institute for Dental Research, College of Dentistry, University of Iowa, Iowa City, IA, USA

[d]Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

[e]Neurodevelopment Laboratory, Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

[f]Group of Translational Genomics, Department of Translational Medicine, John Curtin School of Medical Research, The Australian National University, Canberra ACT, Australia

Abstract

We have applied a GWAS to 40 consanguineous families segregating cases of non-syndromic cleft lip with or without cleft palate (NS CL/P) (a total of 160 affected and unaffected individuals) in order to trace potential recessive loci that confer susceptibility to this common facial malformation. Pedigree-based association test (PBAT) analyses reported nominal evidence of association and linkage over SNP markers located at 11q25 (rs4937877, $P = 2.7 \times 10^{-6}$), 19p12 (rs4324267, $P = 1.6 \times 10^{-5}$), 5q14.1 (rs4588572, P-value $= 3.36 \times 10^{-5}$), and 15q21.1 (rs4774497, $P = 1.08 \times 10^{-4}$). Using the Versatile Gene-Based Association Study to complement the PBAT results, we found clusters of markers located at chromosomes 19p12, 11q25, and 8p23.2 overcome the threshold for GWAS significance ($P < 1 \times 10^{-7}$). From this study, new recessive loci implicated in NS CL/P include: B3GAT1, GLB1L2, ZNF431, ZNF714, and CSMD1, even though the functional association with the genesis of NS CL/P remains to be elucidated. These results emphasize the importance of using homogeneous populations, phenotypes, and family structures for GWAS combined with gene-based association analyses, and should encourage. other researchers to evaluate these genes on independent patient samples affected by NS CL/P.

Keywords

Non-syndromic clefting; Facial; GWAS; Recessive loci; B3GAT1; CSMD1

1. Introduction

Orofacial clefts (OFC) are a major public health problem, affecting one every 500–1000 births worldwide[1]. The etiology of OFC is complex, and the genetic contributions are

© 2012 Elsevier Masson SAS. All rights reserved.

*Corresponding author. John Curtin School of Medical Research, The Australian National University, Building 131 Garran Road, Canberra, ACT 0200, Australia. Tel.: +61 2 61259396. Mauricio.Arcos-Burgos@anu.edu.au (M. Arcos-Burgos).

$watermark-text $watermark-text $watermark-text

heterogeneous, likely relating to interacting effects of multiple loci with environmental covariates[2]. Non-syndromic (sometimes termed 'isolated') cleft lip with or without cleft palate (NS CL/P) is the most prevalent type of OFC[3].

Segregation analyses of CL/P have supported models that include genes of major effect[4]. Analyses of recurrence risk patterns estimate that between 3 and 14 genes (possibly interacting) are involved in the etiology of CL/P[5]. Mutation analysis of candidate genes revealed that 2–6% of individuals with NS CL/P are identified as having mutations in several genes including *MSX1, FOXE1, GLI2, JAG2, LHX8, SATB2, RYK1*[6,7]. In addition to analyses of candidate genes/loci, numerous genome-wide linkage screens of NS CL/P have been reported[1,8,9]. Meta-analysis of several of these published genome scans revealed genome-wide significant regions that best fit a recessive model of inheritance[1].

Over the past 20 years, one member of our research group (MA–B) conducted population genetics studies of the isolated Paisa community inhabiting the northwestern region of Colombia [10–12]. This community exhibits high degrees of endogamy, is genetically homogeneous, and the number of sibs is traditionally larger compared to families from other areas of the country [10,11,13]. Furthermore, race composition studies have shown that more than 90% of the genes are of Caucasian ancestry [10]. The study and recruitment of thousands of members from multigenerational and extended families in the Paisa community have been instrumental for the discovery of new gene associations that predispose this population to complex genetic conditions such as idiopathic epilepsy[14–16], Alzheimer's disease and other dementias[17,18], attention deficit/hyperactivity disorder (ADHD) [19–22], and rheumatologic and autoimmune conditions[23–26], among others. In the particular case of NS CL/P, data from more than 600 multigenerational and extended Paisa pedigrees clustering NS CL/P were instrumental in identifying several loci associated with facial malformation. These loci include *IRF6* and *FOXE1*[1,8,27,28] and the replication of loci identified in the GWAS as significantly associated with NS CL/P[29].

The purpose of this study is to contrast the hypothesis that NS CL/P loci following a recessive model of inheritance might be suitable for mapping using the GWAS technology applied to consanguineous extended and multigenerational Paisa families. To achieve this goal, families segregating for NS CL/P were ascertained by identifying affected probands who were recruited at *Clínica Noel*, in Medellin, Colombia. Clinical aspects, ascertainment methodology, and phenotype characterization of probands and their biological relatives have been reported previously[1,8,27–30]. Most of the families involve extended multiplex kindreds, i.e., multigenerational families with several affected individuals.

2. Methods

2.1. Subjects

To date, we have recruited 661 total pedigrees consisting of 2718 individuals (792 affected by clefting and 1926 unaffected), of which 1370 are female and 1291 are male. In 57 cases, mostly from very early generations, an accurate gender determination was not possible. Individuals with dysmorphic syndromes that included CL/P as a characteristic feature (e.g. van der Woude syndrome) were excluded from this study to restrict the analysis to NS CL/P. Because of our *ad hoc* hypothesis of recessive inheritance, we limited our study to pedigrees of known consanguinity or those in which, as a potential indicative of endogamy, spouses shared the same last name. As a result, a total of 40 pedigrees were selected.

2.2. Genotyping and genetic analysis

Blood samples were collected during home visits or during patients' appointments at the clinic. Consent for future genetic and clinical analyses was specifically obtained. DNA was

$watermark-text
$watermark-text
$watermark-text

extracted using the QIAamp DNA Blood Maxi kit (Qiagen, Valencia, California, USA) and stored at −20 °C until use. Genotyping was performed at the NHGRI Genome Technology Branch using 370CNV-Quad SNP-chips from Illumina (www.illumina.com) and the Illumina Infinium assay protocol[31]. In brief, the DNA was whole-genome amplified, fragmented, hybridized, fluorescently tagged, and scanned. Standard quality control was applied.

All genotype data sets underwent the same rigorous quality checks both before and after the affected and unaffected NS CL/P patients were compared. SNPs were excluded from the analysis if they violated Hardy–Weinberg equilibrium ($P < 0.05$ for all genotyped SNPs), had a call rate below 90%, or had a minor allele frequency below 1%. All analyses were carried out with the use of established procedures implemented in the SVS 7.3.1 PBAT module (Golden Helix, Inc. Bozeman, MT, USA. Golden Helix PBAT Software, http://www.goldenhelix.com). PBAT employs a unified approach to calculate the Family Based Association Test (FBAT) statistic, a generalization of the Transmission Disequilibrium Test (TDT) method[32]. The FBAT statistic is based on a linear combination of offspring genotypes and traits, and for this particular set of families is exceptionally well suited as it maximizes the exquisite information provided by these pedigrees. Furthermore, we used FBAT because it is robust against the effects of population stratification and admixture[32]. To deal with the intensive computations needed to analyze these complex pedigrees, we used the PBAT module that identifies clusters of nuclear families in extended pedigrees, which are directly linked (i.e. that share a family member) and analyzes such clusters as extended pedigrees, avoiding the problem of overestimating any parameter and reducing both the complexity and computation time (Golden Helix, Inc. Bozeman, MT, USA. Golden Helix PBAT Software, http://www.goldenhelix.com). We also used the recessive model to test the null hypothesis of no linkage and no association as it maximized the power of the FBAT-statistic, and also because it suits our initial hypothesis of recessive transmission for this trait. Because NS CL/P occurs more often in men than women, ~2:1 ratio, sex was included as a modifier variable because covariates for the selected phenotype are known to substantially increase the power of the FBAT statistic [32,33].

The total set of results from the PBAT analysis (331,352 SNP marker loci that pass quality control) was subjected to a gene-based association test using Versatile Gene-Based Associated Study (VEGAS)[34]. Currently, gene-based tests are popular as an important complementary methodology in GWAS as they combine the markers' implicit linkage disequilibrium information and the structural information on the genes. As a consequence of combining these methods, marginal levels of significance are often confused with random noise may add up to reveal significant signals of association[34]. Specifically, VEGAS performs gene-based association tests that produce a gene-based test statistic and then uses a simulation-based approach to calculate an empirical gene-based P-value. By default, patterns of linkage disequilibrium for each gene were estimated using the HapMap2 CEU population because the Paisa community is mostly Caucasian[10]Since 331,352 SNP marker loci were available for analysis after quality control (threshold P-value for GWAS significance was set at $0.05/331{,}352 = 1.5 \times 10^{-7}$), we performed 10^7 permutations to set up empirical threshold values after using VEGAS.

3. Results

From the 40 extended pedigrees primarily selected for the GWAS, 34 passed quality control and were informative for genetic analyses. A total of 373 individuals, 287 unaffected and 86 affected NS CL/P, 174 females and 199 males, constituted these 34 pedigrees. From the 86 NS CL/P affected individuals, 31 were females and 55 males ($\chi^2 = 5.05$, $P = 0.025$, Odds Ratio = 1.76 males/females, 95% C.I. 1.04–3.00). The family size ranged from 3 to 34

$watermark-text $watermark-text $watermark-text

individuals. A detailed description in terms of size, sex composition as well as status is presented in Table 1 of the supplementary information.

PBAT analysis reported nominal evidence of association and linkage over SNPs located at 5q14.1 (rs4588572, $P = 3.36 \times 10^{-5}$), 11q25 (rs4937877, $P = 2.7 \times 10^{-6}$), 15q21.1 (rs4774497, $P = 1.08 \times 10^{-4}$), and 19p12 (rs4324267, $P = 1.6 \times 10^{-5}$) (Table 1).

After using VEGAS, clusters of significant markers located at chromosomes 8p23.2,11q25, and 19p12 overcame the threshold for GWAS significance ($P < 1 \times 10^{-7}$) (Table 2). In the 8p23.2 region, the CUB and Sushi Multiple Domains 1 (CSMD1) gene is contained; in 11q25 the beta-1,3-glucuronyltransferase 1 (B3GAT1) and beta-galactosidase-1-like protein 2 (GLB1L2) genes, and in 19p12, the Homo Sapiens Zinc Finger Protein 431 (ZNF431) and Homo Sapiens Zinc Finger Protein 714 (ZNF714) genes.

4. Discussion

Most studies of non-syndromic clefts have focused on CL/P rather than isolated cleft palate. This has been biased perhaps by the larger number of cases, easier ascertainment and less confusion from confounding syndromes. To date, there are three published GWAS studies for CL/P using a case-control design[35–37] and one case-parent trio study from an international consortium that is part of GENEVA (the gene-environment association studies consortium) [29,38]. The data from these studies is summarized in Klotz *et al.*[39]. Although a number of important genes showing association to NS CL/P have been reported, disease-causing variants still remain unidentified. In our study, using extended and multigenerational pedigrees from the Paisa community, a genetic isolate in Colombia, South America, we have identified new loci harboring very interesting candidate genes for conferring a risk of susceptibility for NS CL/P.

CSMD1, a complement control-related gene with potential suppressive activity of squamous cell carcinomas, has been associated with the development of head and neck cancers[40,41]. It has been proposed that *CSMD1* may be an important regulator of complement activation and inflammation in the developing central nervous system, and it may play a role in the context of growth cone function [42]. A recent report also associates *CSMD1* with schizophrenia in three independent European populations [43], although the direct relevance of this finding to NS CL/P is unclear. This gene has an intermediate level of expression in the brain, including cerebellum, substantia nigra, hippocampus and fetal brain[44].

B3GAT1 encodes a member of the glucuronyltransferase gene family that functions as the key enzyme in a glucuronyl transfer reaction during the biosynthesis of the carbohydrate epitope HNK-1 (*Human Natural Killer-1*, also known as CD57 and LEU7) [45]. *B3GAT1* has a prominent expression in the brain, and little or no expression in other tissues. Like *CSMD1, B3GAT1* has been previously associated with schizophrenia [46].

ZNF431 is an uncharacterized Krüppel-associated box (KRAB)-containing C_2H_2 zinc finger protein. These transcription factors are involved in the regulation of cell differentiation, proliferation, apoptosis and neoplastic transformation[47–49]. A recent report implicates ZNF431 in controlling both Ptch1 basal expression and cellular response to Hedgehog signaling, which suggests a role for this transcription factor during developmental stages[50].

ZNF714 and *GLB1L2* are of unknown function. Because *ZNF714* is a zinc finger protein belonging to the same subfamily as *ZNF431*, there is a possibility of a developmental role for this gene. *GLB1L2* is highly expressed in neural crest-derived tissues.

$watermark-text $watermark-text $watermark-text

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to extend their deepest gratitude to all the patients and families from Antioquia, Colombia, who took part in our ongoing CL/CP program. This research was supported by COLCIENCIAS (Grant 1115-408-20519), Programa Sostenibilidad Universidad de Antioquia, Colombia, and in part by the Division of Intramural Research, National Human Genome Research Institute, National Institutes of Health, USA.

References

1. Marazita ML, Lidral AC, Murray JC, Field LL, Maher BS, McHenry TG, 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):151–170. [PubMed: 19521098]
2. Moreno LM, Arcos-Burgos M, Marazita ML, Krahn K, Maher BS, Cooper ME, et al. Genetic analysis of candidate loci in non-syndromic cleft lip families from Antioquia-Colombia and Ohio. Am J Med Gen Part A. 2004 Mar 1; 125A(2):135–144.
3. Marazita ML, Mooney MP. Current concepts in the embryology and genetics of cleft lip and cleft palate. Clin Plast Surg. 2004 Apr; 31(2):125–140. [PubMed: 15145658]
4. Marazita ML, Goldstein AM, Smalley SL, Spence MA. Cleft lip with or without cleft palate: reanalysis of a three-generation family study from England. Genet Epidemiol. 1986; 3(5):335–342. [PubMed: 3781238]
5. Schliekelman P, Slatkin M. Multiplex relative risk and estimation of the number of loci underlying an inherited disease. Am J Hum Genet. 2002 Dec; 71(6):1369–1385. [PubMed: 12454800]
6. Jezewski PA, Vieira AR, Nishimura C, Ludwig B, Johnson M, O'Brien SE, et al. Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. J Med Genet. 2003 Jun; 40(6):399–407. [PubMed: 12807959]
7. Watanabe A, Akita S, Tin NT, Natsume N, Nakano Y, Niikawa N, et al. A mutation in RYK is a genetic factor for nonsyndromic cleft lip and palate. Cleft Palate Craniofac J. 2006 May; 43(3):310–316. [PubMed: 16681403]
8. Marazita ML, Field LL, Tuncbilek G, Cooper ME, Goldstein T, Gursu KG. Genome-scan for loci involved in cleft lip with or without cleft palate in consanguineous families from Turkey. Am J Med Gen Part A. 2004 Apr 15; 126A(2):111–122.
9. 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 Gen Part A. 2008 Jun 1; 146A(11):1406–1413.
10. Arcos-Burgos M, Muenke M. Genetics of population isolates. Clin Genet. 2002 Apr; 61(4):233–247. [PubMed: 12030885]
11. Pineda DA, Palacio LG, Puerta IC, Merchan V, Arango CP, Galvis AY, et al. Environmental influences that affect attention deficit/hyperactivity disorder: study of a genetic isolate. Eur Child Adolesc Psychiatry. 2007 Aug; 16(5):337–346. [PubMed: 17487441]
12. Jain M, Velez JI, Acosta MT, Palacio LG, Balog J, Roessler E, et al. A cooperative interaction between LPHN3 and 11q doubles the risk for ADHD. Mol Psychiatry. 2011 May 24.
13. Palacio JD, Castellanos FX, Pineda DA, Lopera F, Arcos-Burgos M, Quiroz YT, et al. Attention-deficit/hyperactivity disorder and comorbidities in 18 Paisa Colombian multigenerational families. J Am Acad Child Adolesc Psychiatry. 2004 Dec; 43(12):1506–1515. [PubMed: 15564820]
14. Jimenez I, Mora O, Jimenez M, Zuluaga L, Isaza R, Sanchez JL, et al. Complex segregation analysis of non-myoclonic idiopathic generalized epilepsy in families ascertained from probands affected with idiopathic epilepsy with tonic-clonic seizures in Antioquia, Colombia. Hum Genet. 1996 Aug; 98(2):214–218. [PubMed: 8698346]
15. Jimenez I, Mora O, Lopez G, Jimenez ME, Zuluga L, Isaza R, et al. Idiopathic epilepsy with generalized tonic clonic seizures in Antioquia, Colombia: is the joint Amerindian and Negroid

$watermark-text $watermark-text $watermark-text

$watermark-text
$watermark-text
$watermark-text

racial admixture the cause of its high prevalence? Biol Res. 1996; 29(3):297–304. [PubMed: 9278700]

16. Palacio LG, Sanchez JL, Jimenez ME, Rivera-Valencia D, Jimenez-Ramirez I, Arcos-Burgos M. Linkage analysis of the 15q25-15q22 region in an extended multigenerational family segregating for idiopathic epilepsy. Rev Neurol. 2004 Dec 1–15; 39(11):1021–1025. [PubMed: 15597263]
17. Arboleda-Velasquez JF, Lopera F, Lopez E, Frosch MP, Sepulveda-Falla D, Gutierrez JE, et al. C455R notch3 mutation in a Colombian CADASIL kindred with early onset of stroke. Neurology. 2002 Jul 23; 59(2):277–279. [PubMed: 12136071]
18. Lopera F, Ardilla A, Martinez A, Madrigal L, Arango-Viana JC, Lemere CA, et al. Clinical features of early-onset Alzheimer disease in a large kindred with an E280A presenilin-1 mutation. JAMA: J Am Med Association. 1997 Mar 12; 277(10):793–799.
19. Arcos-Burgos M, Castellanos FX, Konecki D, Lopera F, Pineda D, Palacio JD, et al. Pedigree disequilibrium test (PDT) replicates association and linkage between DRD4 and ADHD in multigenerational and extended pedigrees from a genetic isolate. Mol Psychiatry. 2004 Mar; 9(3): 252–259. [PubMed: 15094785]
20. Arcos-Burgos M, Jain M, Acosta MT, Shively S, Stanescu H, Wallis D, et al. A common variant of the latrophilin, 3 gene, LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant medication. Mol Psychiatry. 2010 Nov; 15(11):1053–1066. [PubMed: 20157310]
21. Jain M, Palacio LG, Castellanos FX, Palacio JD, Pineda D, Restrepo MI, et al. Attention-deficit/ hyperactivity disorder and comorbid disruptive behavior disorders: evidence of pleiotropy and new susceptibility loci. Biol Psychiatry. 2007 Jun 15; 61(12):1329–1339. [PubMed: 16950213]
22. Domene S, Stanescu H, Wallis D, Tinloy B, Pineda DE, Kleta R, et al. Screening of human LPHN3 for variants with a potential impact on ADHD susceptibility. Am J Med Genet B Neuropsychiatr Genet. 2011 Jan; 156(1):11–18. [PubMed: 21184580]
23. Anaya JM, Correa PA, Mantilla RD, Arcos-Burgos MTAP. HLA-DQB1, and HLA-DRB1 polymorphism in Colombian patients with primary Sjogren's syndrome. Semin Arthritis Rheum. 2002 Jun; 31(6):396–405. [PubMed: 12077712]
24. Anaya JM, Correa PA, Mantilla RD, Arcos-Burgos M. Rheumatoid arthritis association in Colombian population is restricted to HLA-DRB1*04 QRRAA alleles. Genes Immun. 2002 Feb; 3(1):56–58. [PubMed: 11857065]
25. Anaya JM, Rivera D, Palacio LG, Arcos-Burgos M, Correa PA. D6S439 microsatellite identifies a new susceptibility region for primary Sjogren's syndrome. J Rheumatol. 2003 Oct; 30(10):2152–2156. [PubMed: 14528509]
26. Arcos-Burgos M, Parodi E, Salgar M, Bedoya E, Builes J, Jaramillo D, et al. Vitiligo: complex segregation and linkage disequilibrium analyses with respect to microsatellite loci spanning the HLA. Hum Genet. 2002 Apr; 110(4):334–342. [PubMed: 11941482]
27. 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 Dec 15; 18(24):4879–4896. [PubMed: 19779022]
28. Zucchero TM, 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. N Engl J Med. 2004 Aug 19; 351(8):769–780. [PubMed: 15317890]
29. 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 Jun; 42(6):525–529. [PubMed: 20436469]
30. Chiquet BT, Lidral AC, Stal S, Mulliken JB, Moreno LM, Arcos-Burgos M, et al. CRISPLD2: a novel NSCLP candidate gene. Hum Mol Genet. 2007 Sep 15; 16(18):2241–2248. [PubMed: 17616516]
31. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genetics. 2005 May; 37(5):549–554. [PubMed: 15838508]
32. Lange C, Silverman EK, Xu X, Weiss ST, Laird NM. A multivariate family-based association test using generalized estimating equations: FBAT-GEE. Biostatistics. 2003 Apr; 4(2):195–206. [PubMed: 12925516]

33. Bender PL. Genetics of cleft lip and palate. J Pediatr Nurs. 2000 Aug; 15(4):242–249. [PubMed: 10969497]
34. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010 Jul 9; 87(1):139–145. [PubMed: 20598278]
35. 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 Genetics. 2009 Apr; 41(4):473–477. [PubMed: 19270707]
36. 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 Dec; 155(6):909–913. [PubMed: 19656524]
37. 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 Genetics. 2010 Jan; 42(1):24–26. [PubMed: 20023658]
38. Cornelis MC, Agrawal A, Cole JW, Hansel NN, Barnes KC, Beaty TH, et al. The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. Genet Epidemiology. 2010 May; 34(4):364–372. [PubMed: 20091798]
39. Klotz CM, Wang X, Desensi RS, Grubs RE, Costello BJ, Marazita ML. Revisiting the recurrence risk of nonsyndromic cleft lip with or without cleft palate. Am J Med Gen Part A. 2010 Nov; 152A(11):2697–2702.
40. Scholnick SB, Richter TM. The role of CSMD1 in head and neck carcinogenesis. Genes Chromosomes Cancer. 2003 Nov; 38(3):281–283. [PubMed: 14506705]
41. Toomes C, Jackson A, Maguire K, Wood J, Gollin S, Ishwad C, et al. The presence of multiple regions of homozygous deletion at the CSMD1 locus in oral squamous cell carcinoma question the role of CSMD1 in head and neck carcinogenesis. Genes Chromosomes Cancer. 2003 Jun; 37(2): 132–140. [PubMed: 12696061]
42. Kraus DM, Elliott GS, Chute H, Horan T, Pfenninger KH, Sanford SD, et al. CSMD1 is a novel multiple domain complement-regulatory protein highly expressed in the central nervous system and epithelial tissues. J Immunol. 2006 Apr 1; 176(7):4419–4430. [PubMed: 16547280]
43. Havik B, Le Hellard S, Rietschel M, Lybaek H, Djurovic S, Mattheisen M, et al. The complement control-related genes CSMD1 and CSMD2 associate to schizophrenia. Biol Psychiatry. 2011 Jul 1; 70(1):35–42. [PubMed: 21439553]
44. Nagase T, Kikuno R, Ohara O. Prediction of the coding sequences of unidentified human genes. XXI. The complete sequences of 60 new cDNA clones from brain which code for large proteins. DNA Res. 2001 Aug 31; 8(4):179–187. [PubMed: 11572484]
45. Mitsumoto Y, Oka S, Sakuma H, Inazawa J, Kawasaki T. Cloning and chromosomal mapping of human glucuronyltransferase involved in biosynthesis of the HNK-1 carbohydrate epitope. Genomics. 2000 Apr 15; 65(2):166–173. [PubMed: 10783264]
46. Jeffries AR, Mungall AJ, Dawson E, Halls K, Langford CF, Murray RM, et al. Beta-1,3-Glucuronyltransferase-1 gene implicated as a candidate for a schizophrenia-like psychosis through molecular analysis of a balanced translocation. Mol Psychiatry. 2003 Jul; 8(7):654–663. [PubMed: 12874601]
47. Jheon AH, Ganss B, Cheifetz S, Sodek J. Characterization of a novel KRAB/C2H2 zinc finger transcription factor involved in bone development. J Biol Chem. 2001 May 25; 276(21):18282–18289. [PubMed: 11278774]
48. Hennemann H, Vassen L, Geisen C, Eilers M, Moroy T. Identification of a novel Kruppel-associated box domain protein, Krim-1, that interacts with c-Myc and inhibits its oncogenic activity. J Biol Chem. 2003 Aug 1; 278(31):28799–28811. [PubMed: 12748187]
49. Hering TM, Kazmi NH, Huynh TD, Kollar J, Xu L, Hunyady AB, et al. Characterization and chondrocyte differentiation stage-specific expression of KRAB zinc-finger protein gene ZNF470. Exp Cell Res. 2004 Sep 10; 299(1):137–147. [PubMed: 15302581]

$watermark-text

$watermark-text

$watermark-text

50. He Z, Cai J, Lim JW, Kroll K, Ma L. A novel KRAB domain-containing zinc finger transcription factor ZNF431 directly represses Patched1 transcription. J Biol Chem. 2011 Mar 4; 286(9):7279–7289. [PubMed: 21177534]

Appendix A. Supplementary data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.ejmg.2012.06.005.

$watermark-text
$watermark-text
$watermark-text

$watermark-text $watermark-text $watermark-text

Table 1

P-values for the top 30 SNPs reporting nominal evidence of association and linkage.

SNP	Chromosome	Allele (frequency)	avHet	Position (bp)	Closest gene(s)[a]	*P*-value
rs4937877	11	G (0.2542)	0.3565	133,740,601	*GLB1L2*	2.70×10^{-6}
rs4324267	19	A (0.2208)	0.2241	21,116,188	*ZNF431*	1.60×10^{-5}
rs4707479	6	A (0.4083)	0.3776	68,787,829	*BAI3*	3.18×10^{-5}
rs4588572	5	G (0.3621)	0.1951	77,667,389	*SCAMP1*	3.36×10^{-5}
rs787499	1	A (0.375)	0.4571	67,843,969	*SERBP1/GNG12*	4.11×10^{-5}
rs10063742	5	A (0.375)	0.1963	77,795,117	*SCAMP1*	5.01×10^{-5}
rs7700390	5	A (0.3708)	0.2442	77,825,287	*LHFPL2*	5.01×10^{-5}
rs7309401	12	A (0.2417)	0.4444	106,418,919	*BTBD11*	5.33×10^{-5}
rs2133471	10	A (0.3458)	0.4997	54,831,303	*PRKG1/DKK1/PCDH15*	5.75×10^{-5}
rs2115498	19	A (0.3458)	0.3855	21,147,312	*ZNF431*	6.10×10^{-5}
rs6873144	5	G (0.4583)	0.3872	77,777,539	*SCAMP1*	7.41×10^{-5}
rs1159930	5	A (0.4583)	0.4066	77,804,698	*SCAMP1*	7.41×10^{-5}
rs1428864	5	G (0.4542)	0.3927	77,816,018	*SCAP1/LHFPL2*	7.41×10^{-5}
rs239649	21	A (0.3375)	0.4598	27,717,245	BC043580	7.70×10^{-5}
rs2398587	5	G (0.3625)	0.4275	142,600,718	*ARHGAP26*	9.62×10^{-5}
rs3933797	9	A (0.2833)	0.3878	95,225,151	*FAM120A*	9.62×10^{-5}
rs7036984	9	A (0.2417)	0.3477	95,240,360	CR618537	9.62×10^{-5}
rs2058191	19	A (0.3125)	0.3736	33,528,681	*CCNE1*	9.67×10^{-5}
rs1800977	9	G (0.4625)	0.4173	106,730,270	*ABCA1*	9.72×10^{-5}
rs3797390	5	G (0.4667)	0.477	75,942,820	*IQGAP2*	1.02×10^{-4}
rs1035791	11	A (0.2417)	0.4913	124,697,490	*PKNOX2*	1.08×10^{-4}
rs1344542	12	G (0.3833)	0.4871	108,566,536	*MMAB/MVK*	1.08×10^{-4}
rs4774497	15	A (0.35)	0.4216	45,645,546	*SEMA6D*	1.08×10^{-4}
rs1898110	15	C (0.3305)	0.4328	45,677,641	*SEMA6D*	1.08×10^{-4}
rs4796902	18	G (0.2833)	0.431	10,698,597	*FAM38B*	1.08×10^{-4}
rs4239291	18	A (0.2833)	0.4129	10,699,970	*FAM38B*	1.08×10^{-4}
rs6028738	20	A (0.3333)	0.4458	38,012,703	*DHX35*	1.08×10^{-4}

$watermark-text $watermark-text $watermark-text

SNP	Chromosome	Allele (frequency)	avHet	Position (bp)	Closest gene(s)[a]	*P*-value
rs958523	20	A (0.3083)	0.4455	38,015,758	*DHX35*	1.08×10^{-4}
rs928163	20	A (0.2375)	0.4989	55,823,159	*PMEPA1*	1.08×10^{-4}
rs4748264	10	G (0.1875)	0.3945	16,296,342	*PTER*	1.28×10^{-4}

[a]As in the UCSC Genome Browser GRCh37/hg19 assembly.

$watermark-text $watermark-text $watermark-text

Table 2

Results of gene-based association tests using VEGAS with 10^7 permutations.

Chromosome	Gene	*N*	Start (bp)	Stop (bp)	Test statistic	*P*-value
19	*ZNF431*	5	21,116,679	21,160,645	63.955	$<10^{-7}$
19	*ZNF714*	2	21,056,810	21,099,723	34.681	$<10^{-7}$
11	*B3GAT1*	6	133,753,607	133,787,022	62.712	$<10^{-7}$
11	*GLB1L2*	5	133,707,018	133,751,428	57.138	$<10^{-7}$
8	*CSMD1*	19	2,780,281	4,839,736	101.600	$<10^{-7}$
16	*CDH13*	10	81,218,078	82,387,700	53.829	$<10^{-6}$
11	*CNTN5*	19	98,397,080	99,732,683	119.099	$<10^{-6}$
11	*GLB1L3*	5	133,651,484	133,694,668	57.138	$<10^{-6}$
9	*PTPRD*	19	8,304,245	10,602,509	104.050	$<10^{-6}$
2	*CTNNA2*	7	79,593,633	80,729,416	43.953	$<10^{-6}$
20	*CDH4*	10	59,260,953	59,945,694	54.390	2×10^{-6}
1	*CSMD2*	18	33,752,195	34,404,030	132.064	2×10^{-6}
5	*AP3B1*	2	77,333,905	77,626,284	26.100	3×10^{-6}
8	*SULF1*	10	70,541,412	70,735,701	71.135	4×10^{-6}
7	*DNAH11*	9	21,549,357	21,907,982	59.205	4×10^{-6}
5	*ARHGAP26*	6	142,130,475	142,588,765	57.388	4×10^{-6}
12	*TMEM132D*	17	128,122,223	128,954,165	105.118	5×10^{-6}
16	*A2BP1*	12	6,009,132	7,702,500	59.851	7×10^{-6}
10	*CAMK1D*	11	12,431,588	12,911,739	57.470	7×10^{-6}
1	*FMN2*	8	238,321,807	238,705,112	49.759	7×10^{-6}
19	*KANK2*	10	11,135,945	11,167,496	73.331	9×10^{-6}
11	*OPCML*	13	131,790,084	132,907,613	67.945	1.1×10^{-5}
3	*RBMS3*	11	29,297,946	30,021,624	76.673	1.4×10^{-5}
9	*ASTN2*	9	118,227,327	119,217,138	54.474	1.5×10^{-5}
7	*CNTNAP2*	26	145,444,385	147,749,019	160.697	1.6×10^{-5}
5	*LHFPL2*	8	77,816,793	77,980,404	102.079	1.6×10^{-5}
9	*ABCA1*	3	106,583,104	106,730,257	30.986	1.7×10^{-5}

$watermark-text $watermark-text $watermark-text

Chromosome	Gene	*N*	Start (bp)	Stop (bp)	Test statistic	*P*-value
8	*SGCZ*	26	13,991,743	15,140,163	147.671	1.7×10^{-5}
20	*PLCB1*	9	8,061,295	8,813,547	48.709	1.9×10^{-5}
6	*FARS2*	13	5,206,582	5,716,815	87.725	1.9×10^{-5}

VEGAS: Versatile Gene-based Association Study. *N*: Total number of SNPs; bp: Base pair; SNP: Single Nucleotide Polymorphism.