Association of Low-Frequency Genetic Variants in Regulatory Regions with Non-Syndromic Orofacial Clefts

John R. Shaffer1, 2 | Jessica LeClair3, 4 | Jenna C. Carlson1, 2, 5 | Eleanor Feingold2, 5 | Carmen J. Buxó6 | Kaare Christensen7 | Frederic W.B. Deleyiannis8 | L. Leigh Field9 | Jacqueline T. Hecht10 | Lina Moreno11 | Ieda M. Orioli12, 13 | Carmencita Padilla14 | Alexandre R. Vieira1, 2 | George L. Wehby15 | Jeffrey C. Murray16 | Seth M. Weinberg1, 2, 17 | Mary L. Marazita1, 2, 18 | Elizabeth J. Leslie19

1 Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, 15219 USA

2 Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, 15261 USA

 3 formerly of Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, 15261, USA

4 Department of Biostatistics, School of Public Health, Boston University, Boston, MA 02118, USA

5 Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA

6 Dental and Craniofacial Genomics Core, School of Dental Medicine, University of Puerto Rico, San Juan, Puerto Rico, 00936, USA

7 Department of Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, DK-5000, Denmark

8 Department of Surgery, Plastic and Reconstructive Surgery, University of Colorado School of Medicine, Denver, CO, 80045, USA

9 Department of Medical Genetics, University of British Columbia, Vancouver, V6H 3N1, Canada

10 Department of Pediatrics, McGovern Medical School and School of Dentistry UT Health at Houston, Houston, TX, 77030, USA

11 Department of Orthodontics, College of Dentistry, University of Iowa, Iowa City, IA, 52242, USA

12 Department of Genetics, Institute of Biology, Federal University of Rio de Janeiro, Rio de Janeiro, 21941-617, Brazil

13 ECLAMC (Latin American Collaborative Study of Congenital Malformations) at INAGEMP (National Institute of Population Medical Genetics), Rio de Janeiro, 21941-617, Brazil

14 Department of Pediatrics, College of Medicine; and Institute of Human Genetics, National Institutes of Health; University of the Philippines Manila, Manila, The Philippines, 1000; and Philippine Genome Center, University of the Philippines System, Manila, The Philippines 1101

15 Department of Health Management and Policy, College of Public Health, University of Iowa, Iowa City, IA, 52246, USA

16 Department of Pediatrics, Carver College of Medicine, University of Iowa, Iowa City, Iowa, 52242, USA

17 Department of Anthropology, Dietrich School of Arts and Sciences, University of Pittsburgh, Pittsburgh, PA, 15261, USA

18 Clinical and Translational Science, School of Medicine, University of Pittsburgh, Pittsburgh, PA, 15213, USA

19 Department of Human Genetics, Emory University School of Medicine, Emory University, Atlanta, GA, 30322

**Correspondence**

John R. Shaffer, Ph.D., 130 De Soto Street, 3131 Parran Hall, Pittsburgh, PA 15261

Email: john.r.shaffer@pitt.edu

Supplemental Material

**Supplemental Figure S1: Results of genetic association tests of variants with MAF less than 5%:** Quantile-quantile plots depict the observed log10-transformed p-values (y-axis) vs. the expected distribution of p-values (x-axis) under the null hypothesis of no association. Each point represents the evidence of genetic association for a specific craniofacial enhancer. The top panels show results for CP and the bottom panels show results for CL/P. Left panels show results of CMC scans, and right panels show results of SKAT scans. Horizontal dashed lines represent the threshold for suggestive association (p-value < 0.004).

CP

CL/P

CMC

CMC

SKAT

SKAT

**Supplemental Figure S2: SNP genotype cluster plot shows evidence of a null allele.** The putative enhancer located on chromosome 1 at base pair position 209663057-209663962 (hg19) showed suggestive evidence of association in the scan of variants with MAF < 5% (SKAT p-value = 0.0011, CMC p-value = 0.017), and the SNP rs141738759 located in this enhancer region individually shows evidence of association (p-value = 0.0002). However, inspection of the cluster plot for this SNP showing missing genotype calls due to a presumed null allele; therefore, we deem this association as potentially spurious.

