RESEARCH REPORTS
Clinical

Genetic Association of *MPPED2* and *ACTN2* with Dental Caries

B.O.C. Stanley[1], E. Feingold[2,3], M. Cooper[4,5], M.M. Vanyukov[2,6,7], B.S. Maher[8], R.L. Slayton[9], M.C. Willing[10], S.E. Reis[11,12], D.W. McNeil[13], R.J. Crout[14], R.J. Weyant[15], S.M. Levy[16,17], A.R. Vieira[4,5], M.L. Marazita[2,4,5,7,12], and J.R. Shaffer[2]*

[1]Department of Mathematics, Vanderbilt University, Nashville, TN, USA; [2]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA; [3]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA; [4]Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [5]Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [6]Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA, USA; [7]Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [8]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, USA; [9]Department of Pediatric Dentistry, School of Dentistry, University of Washington, Seattle, WA, USA; [10]Division of Genetics and Genomics, Medicine, Department of Pediatrics, School of Medicine, Washington, University at St. Louis, St. Louis, MO, USA; [11]Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [12]Clinical and Translational Science Institute, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [13]Dental Practice and Rural Health, West Virginia University, Morgantown, WV, USA; [14]Department of Periodontics, School of Dentistry, West Virginia University, Morgantown, WV, USA; [15]Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [16]Department of Preventive and Community Dentistry, University of Iowa College of Dentistry, Iowa City, IA, USA; and [17]Department of Epidemiology, University of Iowa College of Public Health, Iowa City, IA, USA; *corresponding author, jrs51@pitt.edu

J Dent Res DOI: 10.1177/0022034514534688

APPENDIX

Participant Recruitment and Data Collection

Center for Oral Health in Appalachia

The Center for Oral Health in Appalachia (COHRA) was initiated as a joint venture between the University of Pittsburgh and West Virginia University to investigate the community-, family-, and individual-level factors influencing oral health disparities in rural Appalachian communities of western Pennsylvania and northern West Virginia (Polk *et al.*, 2008). The first research cohort recruited by COHRA (COHRA1) focused on households, with eligibility criteria stipulating that at least 1 biological parent-child pair be included per household. All other members of eligible households were invited to participant regardless of biological or legal relationships. Participants were enrolled without regard to their oral health status. All participants received complete intraoral examinations by licensed dentists and/or research dental hygienists. After calibration, inter- and intrarater agreement was high (Polk *et al.*, 2008; Wendell *et al.*, 2010). DNA was obtained from a variety of sources, including blood, buccal swab, mouthwash, and saliva samples taken in Oragene kits from DNA Genotek (http://www.dnagenotek.com). Written informed consent was obtained from all adult participants, and verbal assent with written parental consent was obtained for all child participants. The procedures, forms, and protocols for the COHRA1 study were approved by institutional review boards (IRBs) at the University of Pittsburgh and West Virginia University.

Dental Strategies Concentrating on Risk Evaluation

The Dental Strategies Concentrating on Risk Evaluation (SCORE) study was an ancillary project of the Heart SCORE project, a prospective longitudinal cohort study designed to investigate the factors contributing to racial and socioeconomic disparities in cardiovascular risk in adults (Aiyer *et al.*, 2007a,

2007b) recruited from the Pittsburgh area; a subset of Heart SCORE participants also agreed to participate in Dental SCORE. Dental SCORE participants were enrolled without regard to their oral health status, and they received a dental screening by a research dental hygienist following the COHRA study protocols summarized above. DNA was extracted from saliva samples obtained from Oragene kits from DNA Genotek. All participants provided informed consent, and all assessment protocols were approved by the University of Pittsburgh IRB.

Dental Registry and DNA Repository

The Dental Registry and DNA Repository (DRDR) was established as a data warehouse to facilitate interdisciplinary research at the University of Pittsburgh. Patients seeking dental treatment at the University of Pittsburgh School of Dentistry were invited to participate. All patients were offered enrollment without regard to their specific oral health or medical status, and those participating provided a DNA sample via saliva samples taken in Oragene kits from DNA Genotek. DRDR samples are then linked to patient dental records. All participants provided written consent for the future use of their genetic and dental phenotype data in research studies. The DRDR was approved by the IRB at the University of Pittsburgh.

Iowa Head Start Study

The Iowa Head Start Study recruited low-income children aged 3 to 5 yr of age who participated in federally subsidized Iowa Head Start programs (Slayton *et al.*, 2005). Standardized in-field dental examinations were performed to assess dental caries experience. Either buccal or saliva samples taken in Oragene kits from DNA Genotek were used to collect DNA. All participants provided assent with parental or legal guardian consent, and all protocols and procedures were approved by the IRB at the University of Iowa.

Iowa Fluoride Study

The Iowa Fluoride Study is an ongoing study that originally recruited new mothers and newborns from 8 participating Iowa postpartum wards and followed their children longitudinally from birth through adulthood (Wang *et al.*, 2012). The goal of the study is to quantify fluoride exposures from dietary and nondietary sources and to associate fluoride exposure with fluorosis and dental caries. Standardized field dental examinations were conducted by trained dentists for children at aged 4 to 6 yr. DNA was obtained from blood, buccal swab, or saliva samples as part of a related study, the Iowa Bone Development Study. All study questionnaires and protocols were approved by the University of Iowa IRB; all parents provided informed written consent; and all children provided verbal assent.

Center for Education and Drug Abuse Research

The Center for Education and Drug Abuse Research study recruited fathers with and without substance use disorder to prospectively investigate substance use risk factors in their offspring from the ages of 10 to 12 yr through 30 yr (Tarter and Vanyukov, 2001). A subset of the Center for Education and Drug Abuse Research cohort was recruited for genetic studies under a project titled Substance Abuse and the Dopamine System Genes (Vanyukov *et al.*, 2004); study 12 of the NIDA Genetics Consortium: https://zork5.wustl/nida/study_description), and that cohort also had dental examinations conducted at the University of Pittsburgh School of Dental Medicine by calibrated dental hygienists. Blood samples were taken, and lymphoblast cells lines were established at the NIDA Center for Genetic Studies (https://zork5/wustl/nida). For the current study, DNA aliquots were obtained from the NIDA Center for Genetic Studies. Study questionnaires and protocols were approved by the University of Pittsburgh IRB, and all parents provided informed consent, along with assent from children.

Custom Genotyping Panel

Participants were genotyped for a custom panel of single-nucleotide polymorphisms (SNPs) based on the Illumina GoldenGate platform (San Diego, CA, USA) by the Center for Inherited Disease Research at Johns Hopkins University. The content of the custom panel included tagging SNPs from 71 genes of interest, as well as several hundred specific SNPs of interest. Although the genes of interest were chosen for a variety of reasons, most of them were included because they were nominated in 1 or more genome-wide association studies for oral health phenotypes. Nomination of these genes was made by virtue of their physical proximity to and/or linkage disequilibrium with associated SNPs in conjunction with the known biology of the gene, previously reported experimental evidence, or a plausible role in the etiology of the oral health phenotype. The specific SNPs of interest (some of which are located in the aforementioned genes of interest) were chosen because they exhibited genetic associations with oral health phenotypes via genome-wide association studies. The content of the custom panel was chosen with the goal of replicating and fine mapping previously discovered associated loci for a variety of different oral health phenotypes, notably dental caries experience.

The nature of the Illumina GoldenGate platform allowed up to 3,072 SNPs to be simultaneously genotyped. Constraints on genotyping included the "designability score" for each SNP, which is a measure of the predictive genotyping success of a given SNP. In addition, no SNP could be within 60 bases of another SNP on the panel. The approach for choosing SNPs was as follows: 746 specific SNPs of interest, plus another 4,107 SNPs located within genes of interest and having minor allele frequencies ≥ 0.02 and designability scores ≥ 0.8, were considered. This made for a total of 4,853 potential SNPs. With data from HapMap CEU subjects, a sliding window of 50 SNPs was employed to prune out any SNPs representing redundant information as defined by multiple $R^2 > 0.95$. This yielded 3,046 tag SNPs, of which 70 were omitted due to physical proximity to another tag SNP. This in turn yielded 2,976 custom SNPs. Another 96 ancestry informative "barcode" SNPs were included for a total of 3,072 SNPs on the custom panel.

Of the 3,072 SNPs attempted, quality genotypes were released from the Center for Inherited Disease Research for 2,662 (86.7%), which is consistent with the expected success rate for this platform. Mean call rate per SNP was 0.9985 and ranged from 0.9689 to 1.0000. Genotypes were released for 5,354 (of

Appendix Table. Beta Coefficients (*p* Values) of Genetic Association Tests for Select Single-Nucleotide Polymorphisms

Gene	Children								Adults											Children + Adults				
	Whites					Blacks		All	Whites									Blacks	All	Whites			Blacks	All
SNP	COHRA	IHS	IFS	Meta	COHRA	IHS	Meta	Meta	COHRA	Dental SCORE	DRDR	CEDAR	Meta	COHRA	Dental SCORE	DRDR	CEDAR	Meta	Meta	Meta			Meta	Meta
BP, *n*	608	41	136	785	81	23	104	889	994	277	702	173	2146	86	225	173	68	552	2698	2931			656	3587
ZMPSTE24																								
rs10489431	0.96	–3.31	–0.60		–0.39	–0.69			0.80	–0.98	0.53	0.07		–1.87	0.23	–2.60	1.51							
40759064	(.01)*	(.17)	(.25)	(.14)	(.67)	(.79)	(.62)	(.22)	(.16)	(.30)	(.46)	(.94)	(.31)	(.44)	(.85)	(.12)	(.38)	(.45)	(.57)	(.10)			(.37)	(.27)
EDARADD																								
rs10925116	0.52	0.63	–0.10		–0.25	–1.03			–0.39	0.20	–0.34	–0.36		–1.98	–0.45	0.21	1.11							
236562878	(.003)*	(.60)	(.73)	(.01)	(.60)	(.68)	(.51)	(.03)	(.17)	(.65)	(.41)	(.48)	(.15)	(.08)	(.44)	(.79)	(.19)	(.57)	(.12)	(.91)			(.43)	(.82)
rs1999506	0.54	0.68	–0.07		–0.02	–1.03			–0.51	0.04	–0.42	–0.16		–1.69	–0.36	0.32	1.03							
236563287	(.002)*	(.56)	(.81)	(.01)	(.96)	(.68)	(.81)	(.01)	(.09)	(.92)	(.31)	(.76)	(.08)	(.14)	(.52)	(.68)	(.22)	(.74)	(.08)	(.91)			(.69)	(.78)
rs1253626	–0.07	–0.88	–0.58		–0.21	–5.44			0.22	0.05	1.21	0.64		–0.67	–0.84	–0.32	0.05							
236617216	(.87)	(.71)	(.29)	(.50)	(.63)	(.01)	(.08)	(.22)	(.75)	(.96)	(.16)	(.60)	(.24)	(.53)	(.12)	(.66)	(.95)	(.14)	(.69)	(.50)			(.04)	(.79)
ACTN2																								
rs6656267	0.29	3.22	0.90		–0.53	–1.63			0.16	–0.04	–0.14	2.90		–0.04	–0.49	0.19	–1.98							
236861921	(.31)	(.05)	(.13)	(.05)	(.25)	(.43)	(.16)	(.17)	(.73)	(.95)	(.84)	(.003)	(.35)	(.97)	(.41)	(.82)	(.05)	(.27)	(.73)	(.07)			(.12)	(.33)
rs707214	–0.29	5.60	1.24		0.20	–0.91			0.11	1.05	0.96	1.27		0.66	–0.94	–0.28	0.00							
236867394	(.55)	(.01)	(.12)	(.47)	(.64)	(.59)	(.87)	(.46)	(.88)	(.31)	(.40)	(.35)	(.23)	(.54)	(.13)	(.74)	(1.00)	(.36)	(.50)	(.16)			(.44)	(.34)
rs4659704	–0.12	–0.25	0.12		–0.21	6.73			0.23	0.08	–0.20	–0.07		1.95	0.56	–2.31	–0.92							
236876257	(.48)	(.84)	(.66)	(.63)	(.66)	(.005)	(.35)	(.90)	(.43)	(.85)	(.63)	(.89)	(.78)	(.15)	(.38)	(.02)	(.33)	(.58)	(1.00)	(.99)			(.90)	(.95)
rs1366990	0.14	–0.49	0.56		–0.73	–6.38			–0.33	–0.06	–0.28	–0.01		–0.56	0.86	1.89	–0.31							
236881703	(.63)	(.79)	(.31)	(.43)	(.09)	(.01)	(.01)	(.83)	(.45)	(.93)	(.62)	(.99)	(.40)	(.58)	(.15)	(.03)	(.75)	(.07)	(.93)	(.76)			(.56)	(.97)
rs707204	0.20	0.72	–0.38		0.25	0.42			0.50	0.35	0.48	0.80		–2.43	–1.30	–0.47	–0.07							
236897376	(.33)	(.59)	(.29)	(.59)	(.61)	(.83)	(.58)	(.48)	(.15)	(.49)	(.31)	(.23)	(.03)	(.01)	(.02)	(.53)	(.92)	(.004)	(.55)	(.03)			(.01)	(.39)
rs7536539	0.41	–0.39	0.30		0.53	–0.56			–0.29	–0.70	–0.53	3.59		–0.87	0.47	–1.16	–2.01							
236898395	(.29)	(.86)	(.69)	(.29)	(.44)	(.85)	(.56)	(.23)	(.65)	(.40)	(.54)	(.003)	(.91)	(.62)	(.60)	(.54)	(.25)	(.54)	(.71)	(.65)			(.74)	(.78)
rs7534128	–0.12	–1.83	0.02		0.26	–1.78			0.66	1.97	0.55	0.10		–0.39	–0.29	–0.72	1.10							
236922254	(.64)	(.32)	(.96)	(.54)	(.58)	(.35)	(.96)	(.57)	(.11)	(.002)*	(.32)	(.90)	(.01)	(.72)	(.62)	(.40)	(.23)	(.61)	(.02)	(.04)			(.66)	(.09)
rs10082093	–0.19	–1.59	0.02		0.06	–1.34			0.57	1.96	0.35	0.08		–0.10	–0.26	–0.51	0.39							
236929016	(.46)	(.42)	(.96)	(.41)	(.90)	(.46)	(.81)	(.40)	(.16)	(.002)*	(.53)	(.93)	(.02)	(.94)	(.71)	(.63)	(.74)	(.68)	(.05)	(.10)			(.63)	(.20)
MTR																								
rs7556238	0.00	–1.06	0.15		0.30	–2.99			–0.04	–2.15	–0.10	–0.54		–0.05	0.36	1.74	0.83							
236948542	(.99)	(.54)	(.75)	(1.00)	(.52)	(.18)	(.96)	(.99)	(.93)	(.003)	(.87)	(.50)	(.16)	(.97)	(.54)	(.05)	(.33)	(.07)	(.66)	(.23)			(.10)	(.70)
rs4659735	0.56	0.45	0.35		–0.21	–0.17			0.62	–1.06	–0.46	–0.68		–2.39	3.09	–0.84	2.25							
237032925	(.02)	(.83)	(.41)	(.01)	(.80)	(.97)	(.81)	(.02)	(.09)	(.06)	(.38)	(.34)	(.76)	(.34)	(.01)	(.70)	(.47)	(.16)	(.71)	(.30)			(.23)	(.15)
EPHA7																								
rs346462	0.00	–1.24	–0.32		–0.42	5.34			–0.20	0.16	–0.65	0.10		2.29	0.00	–0.53	–1.33							
94004485	(.98)	(.41)	(.30)	(.55)	(.36)	(.001)*	(.44)	(.76)	(.52)	(.75)	(.13)	(.86)	(.25)	(.04)	(1.00)	(.57)	(.10)	(.92)	(.29)	(.20)			(.83)	(.29)
rs7754087	0.58	1.00	0.16		–0.69	–4.56			–0.18	0.07	0.15	–0.23		–0.36	0.27	–1.69	0.16							
94034201	(.004)	(.48)	(.61)	(.003)	(.18)	(.11)	(.05)	(.04)	(.58)	(.89)	(.73)	(.69)	(.81)	(.81)	(.67)	(.06)	(.88)	(.40)	(.55)	(.19)			(.12)	(.60)
MPPED2																								
rs12796693	–0.46	–2.80	–0.56		0.99	–0.03			0.20	0.63	–0.25	–0.10		1.37	–0.35	1.48	–1.15							
30429583	(.02)	(.07)	(.08)	(.002)*	(.35)	(.99)	(.42)	(.01)	(.55)	(.20)	(.58)	(.85)	(.62)	(.52)	(.72)	(.33)	(.46)	(.75)	(.56)	(.23)			(.54)	(.41)
rs11031093	0.69	0.69	0.06		0.10	7.61			–0.01	0.46	0.03	0.21		1.80	–1.30	0.50	–1.95							
30467500	(.0003)*	(.68)	(.87)	(.001)*	(.88)	(.02)	(.20)	(.0003)*	(.96)	(.36)	(.94)	(.70)	(.65)	(.27)	(.14)	(.70)	(.28)	(.51)	(.92)	(.03)			(.92)	(.06)
rs536007	0.71	0.60	0.38		–0.16	4.90			0.02	0.46	0.06	0.32		0.66	–1.58	0.28	–0.25							
30478524	(.0002)*	(.71)	(.27)	(.0001)*	(.79)	(.01)	(.32)	(.0001)*	(.95)	(.36)	(.89)	(.57)	(.54)	(.65)	(.05)	(.81)	(.84)	(.31)	(.92)	(.01)			(.60)	(.04)
rs11031114	0.60	0.61	0.42		–0.48	7.61			0.02	0.36	0.17	–0.05		1.70	–1.68	0.38	–1.17							
30517812	(.002)*	(.70)	(.23)	(.001)*	(.50)	(.02)	(.59)	(.001)*	(.96)	(.46)	(.69)	(.92)	(.61)	(.30)	(.07)	(.78)	(.55)	(.42)	(.93)	(.03)			(.60)	(.09)
rs533880	0.59	9.88	–0.46		2.59	5.41			–0.08	1.09	–0.70	2.18		–4.36	3.24	4.43	NA							
30531305	(.27)	(.002)*	(.69)	(.13)	(.30)	(.33)	(.17)	(.06)	(.92)	(.44)	(.55)	(.20)	(.81)	(.18)	(.20)	(.15)	NA	(.24)	(.47)	(.33)			(.10)	(.12)
rs11031120	0.18	–2.00	0.31		–0.50	3.37			0.10	0.07	–0.66	0.96		1.83	–0.15	–1.33	–0.46							

(continued)

Appendix Table. (continued)

Gene	Children								Adults											Children + Adults		
	Whites					Blacks		All	Whites									Blacks	All	Whites	Blacks	All
SNP	COHRA	IHS	IFS	Meta	COHRA	IHS	Meta	Meta	COHRA	Dental SCORE	DRDR	CEDAR	Meta	COHRA	Dental SCORE	DRDR	CEDAR	Meta	Meta	Meta	Meta	Meta
30537506	(.35)	(.21)	(.35)	(.36)	(.44)	(.34)	(.82)	(.43)	(.77)	(.89)	(.15)	(.13)	(.89)	(.17)	(.84)	(.28)	(.71)	(.76)	(.79)	(.72)	(.71)	(.87)
rs4922835	–0.49	0.91	–0.50		0.37	–1.35			–0.03	0.34	–0.36	–0.15		–0.16	–0.09	0.52	0.79					
30539318	(.02)	(.60)	(.11)	(.01)	(.38)	(.47)	(.66)	(.02)	(.92)	(.48)	(.42)	(.79)	(.72)	(.89)	(.87)	(.54)	(.33)	(.60)	(.94)	(.10)	(.51)	(.23)
LPO																						
rs8178290	–0.02	0.84	–0.25		0.22	1.22			0.49	–1.83	0.28	1.08		0.10	0.06	–0.85	–0.93					
56320158	(.92)	(.61)	(.51)	(.81)	(.62)	(.55)	(.47)	(.99)	(.17)	(.001)*	(.56)	(.08)	(.54)	(.93)	(.92)	(.32)	(.20)	(.36)	(.89)	(.69)	(.59)	(.90)
rs8178307	–0.02	0.79	–0.25		0.22	1.22			0.53	–1.83	0.26	1.08		0.10	0.01	–0.88	–0.84					
56323678	(.95)	(.64)	(.50)	(.82)	(.62)	(.55)	(.47)	(.98)	(.13)	(.001)*	(.59)	(.08)	(.49)	(.93)	(.99)	(.30)	(.24)	(.34)	(.86)	(.64)	(.56)	(.86)

COHRA, Center for Oral Health in Appalachia; IHS, Iowa Head Start; IFS, Iowa Fluoride Study; Dental SCORE, Dental Strategies Concentrating on Risk Evaluation; DRDR, Dental Registry and DNA Repository; CEDAR, Center for Education and Drug Abuse Research; BP, base pair position of SNP.
*Indicates statistical significance after gene-wise adjustment for multiple comparisons.

5502 attempted) samples, as well as 119 (of 126 attempted) blind duplicates and 237 HapMap controls. Genotyping success by DNA source was as follows: 1,499 of 1,512 attempted for blood samples, 138 of 194 attempted for buccal swabs, 214 of 241 attempted for mouthwash samples, and 3,503 of 3,555 attempted for saliva samples. Mean call rate per participant was 0.9985 and ranged from 0.9409 to 1.0000. Estimated Mendelian consistency rate was 99.97%. Genotyping error rate as estimated from blind duplicates was 0.0001.

Multidimensional scaling as implemented in PLINK (Purcell *et al.*, 2007) was used to generate components of ancestry. Self-reported race was consistent with genetically determined ancestry. Self-reported sex was consistent with genetic sex. Self-reported biological relationships among participants were confirmed by genetic relatedness. Together, ancestry-, sex-, and familial relationship-checking procedures were used to verify the identity of the DNA samples.

Statistical Analyses

Genetic association was tested via linear regression separately in age- and race-stratified subsamples for each study. Only non-Hispanic white and non-Hispanic black participants meeting age criteria were included in the present analysis ($N = 3,587$). Caries experience in the primary dentition, as measured by dfs index, was investigated in children aged 3 to 12 yr. Caries experience in the permanent dentition, as measured by DMFS, was investigated in adults 18 yr or older for COHRA, Dental SCORE, and DRDR and in adolescents and young adults 15 yr or older for CEDAR (which reflects the youth of this sample compared to the other adult cohorts). The reason for stratifying by age is that we anticipate that loci may have differential effects on caries experience of the primary and permanent dentitions. The reason for stratifying by race is that population structure (which would result in a mixed sample of participants from different racial groups) can lead to false-positive tests of association. Stratifying analyses by race safeguards against this type of statistical artifact.

Genetic association was tested for each SNP one at a time under the additive genetic model (*i.e.*, genotypes were coded as the number of rare alleles: 0, 1, or 2). The additive model yields high statistical power even for loci that truly behave according to other common genetic modes of inheritance, such as dominance, incomplete dominance, and recessive models. Depending on allele frequency, the additive model may have reduced statistical power for loci that truly exhibit unusual modes of inheritance, such as overdominance and underdominance, which are generally not expected. Sex and age were included as covariates for all analyses. Additionally, the first 4 components of ancestry were included as covariates for analyses of blacks because individuals with African ancestry potentially exhibit population stratification.

Analyses in children were repeated using logistic regression to model a binary caries phenotype (*i.e.*, yes/no; does the child have 1 or more carious teeth?). Covariate adjustments were identical to the linear regression modeling. The binary phenotype was considered because this was the modeling framework for the original genome-wide association study (Shaffer *et al.*, 2011) that nominated the genes that we investigated. This approach may be appropriate for identifying genetic loci that affect whether a participant is susceptible or fully resistant to dental caries, whereas the linear model approach may help identify loci that influence the severity of caries experience. Because of the very small proportion of caries-free adults, logistic regression modeling of the binary phenotype was not conducted for adults.

Stouffer's inverse variance method of meta-analysis was used to combine results across cohorts based on the sample size, direction of effect, and p value of the association test. This method is appropriate even if the scale of effect sizes differs across cohorts, for example, due to differences in age ranges or phenotype distribution (primary *vs.* permanent dentition). Meta-analyses were performed for all white children, all black children, all children, all white adults, all black adults, all adults, and all participants.

Asymptotic p values are presented for all analyses, which, in principle, may be affected by the deviations of our caries phenotypes from normality and the nonindependence among relatives in the COHRA sample. Specifically, the distributions of caries phenotypes are zero inflated, with long right tails. We have explored the impact of this nonnormality on linear models using genome-wide data. First, we have shown that asymptotic p values, such as those presented in this study, are not meaningfully different from true empirical p values generated via permutation. Second, we have applied a severe transformation to create exactly normally distributed transformed phenotypes, and we have shown that results of linear models were not meaningfully different. Given these exploratory analyses (data not shown) and our extensive experience working with dental caries phenotypes, we conclude that the scientific interpretation of our models is likely unchanged despite any potential violations in model assumptions due to phenotype distribution.

We have also thoroughly explored the potential impact of nonindependence among participants on our models. In particular, the COHRA sample, which included households with children from rural communities, includes a variety of relatives. The problem of nonindependence among the sample is partly mitigated by our analysis strategy that separated children from their adult parents. Moreover, we have previously calculated the genomic inflation factor (*i.e.*, a measure of how inflated the p values are across many genetic association tests), which was approximately 1.05 for COHRA children (ranging from 1.03 to 1.07, depending on exact phenotype and model) and approximately 1.00 for COHRA adults (ranging from 0.99 to 1.01). At the candidate gene level, this amount of genomic inflation does not alter the scientific interpretation of results. For example, an asymptotic p value of .001 would still round to .001 after adjustment for genomic inflation factor of 1.05. In addition, we have previously explored the consistency of results from regular linear models that ignore the relatedness among participants to variance components methods (Almasy and Blangero, 1998) that condition on the known relatedness. Such analyses have shown that results were not meaningfully different (results not shown). Therefore, we conclude that the degree of relatedness among our sample does not affect the scientific interpretation of our results. For the present study, the issue of relatedness in the COHRA children sample is further marginalized, given that the benchmarks for interpreting our results are based primarily on replicating genetic associations in additional samples.

APPENDIX REFERENCES

Aiyer AN, Kip KE, Marroquin OC, Mulukutla SR, Edmundowicz D, Reis SE (2007a). Racial differences in coronary artery calcification are not attributed to differences in lipoprotein particle sizes: the Heart Strategies Concentrating on Risk Evaluation (Heart SCORE) Study. *Am Heart J* 153:328-334.

Aiyer AN, Kip KE, Mulukutla SR, Marroquin OC, Hipps L Jr, Reis SE (2007b). Predictors of significant short-term increases in blood pressure in a community-based population. *Am J Med* 120:960-967.

Almasy L, Blangero J (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. American journal of human genetics. 62:1198-1211.

Polk DE, Weyant RJ, Crout RJ, McNeil DW, Tarter RE, Thomas JG, *et al.* (2008). Study protocol of the Center for Oral Health Research in Appalachia (COHRA) etiology study. *BMC Oral Health* 8:18.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* (2007). Plink: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559-575.

Shaffer JR, Wang X, Feingold E, Lee M, Begum F, Weeks DE, *et al.* (2011). Genome-wide association scan for childhood caries implicates novel genes. *J Dent Res* 90:1457-1462.

Slayton RL, Cooper ME, Marazita ML (2005). Tuftelin, mutans streptococci and dental caries susceptibility. *J Dent Res* 84:711-714.

Tarter RE, Vanyukov MM (2001). Theoretical an operational framework for research into the etiology of substance use disorders. *Journal of Child and Adolescent Substance Abuse* 10:1-12.

Vanyukov MM, Maher BS, Devlin B, Tarter RE, Kirillova GP, Yu LM, *et al.* (2004). Haplotypes of the monoamine oxidase genes and the risk for substance use disorders. *Am J Med Genet B Neuropsychiatr Genet* 125B:120-125.

Wang X, Willing MC, Marazita ML, Wendell S, Warren JJ, Broffitt B, *et al*. (2012). Genetic and environmental factors associated with dental caries in children: the Iowa Fluoride Study. *Caries Res* 46:177-184.

Wendell S, Wang XJ, Brown M, Cooper ME, DeSensi RS, Weyant RJ, et al. (2010). Taste genes associated with dental caries. *J Dent Res* 89:1198-1202.