Genome-Wide Association Study of Primary Dentition Pit-and-Fissure and Smooth Surface Caries


Departments of Human Genetics and Biostatistics, Graduate School of Public Health, University of Pittsburgh, and Center for Craniofacial and Dental Genetics, Department of Oral Biology, and Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, Pittsburgh, Pa., Department of Preventive and Community Dentistry, University of Iowa College of Dentistry, and Department of Epidemiology, University of Iowa College of Public Health, Iowa City, Iowa, Department of Periodontics, School of Dentistry, and Departments of Dental Practice and Rural Health, and Psychology, West Virginia University, Morgantown, W.Va., USA

Key Words
Genome-wide association study · Pit-and-fissure surface · Primary dentition · Smooth surface · Susceptibility gene

Abstract
Dental caries continues to be the most common chronic disease in children today. Despite the substantial involvement of genetics in the process of caries development, the specific genes contributing to dental caries remain largely unknown. We performed separate genome-wide association studies of smooth and pit-and-fissure tooth surface caries experience in the primary dentitions of self-reported white children in two samples from Iowa and rural Appalachia. In total, 1,006 children (ages 3–12 years) were included for smooth surface analysis, and 979 children (ages 4–14 years) for pit-and-fissure surface analysis. Associations were tested for more than 1.2 million single nucleotide polymorphisms, either genotyped or imputed. We detected genome-wide significant signals in KPNA4 (p value = 2.0E–9), and suggestive signals in ITGAL (p value = 2.1E–7) and PLUNC family genes (p value = 2.0E–6), thus nominating these novel loci as putative caries susceptibility genes. We also replicated associations observed in previous studies for MPPED2 (p value = 6.9E–6), AJAP1 (p value = 1.6E–6) and RPS6KA2 (p value = 7.3E–6). Replication of these associations in additional samples, as well as experimental studies to determine the biological functions of associated genetic variants, are warranted. Ultimately, efforts such as this may lead to a better understanding of caries etiology, and could eventually facilitate the development of new interventions and preventive measures.

Dental caries (i.e., tooth decay) is the most common chronic disease in children globally, and despite advances in oral healthcare, the incidence of dental caries in young children in the US has increased in recent decades [Beltran-Aguilar et al., 2005; Dye et al., 2007]. The effects of childhood dental caries on quality of life can be profound, including chronic pain, tooth loss, difficulty hearing, eating, and sleeping, and failure to thrive [Acs et al., 1992, 1999], as well as substandard school performance,
GWAS of Childhood Surface-Level Caries

DOI: 10.1159/000356299

等因素，包括口腔微生物，饮食行为，氟摄入和暴露，口腔卫生，唾液成分和流量，以及牙齿的形态特征，以及基因-环境相互作用。此外，牙科疾病的复杂性还表现在表面的异质性，以及影响因素的可能异质性。

Cariogenesis is multifactorial [Hunter, 1988; Anderson, 2002], influenced by sex [Ferraro and Vieira, 2010] and environmental factors, such as bacterial flora, dietary behaviors, fluoride intake and exposures, oral hygiene, salivary composition and flow rate, and tooth positional and morphological features, as well as genetic factors, and gene-by-environment interactions. In addition, the etiology of dental caries is further complicated by the nonuniform risk across tooth surfaces of the primary dentition. Tooth surfaces can be grouped according to the sequence and pattern of decay, which demonstrates a surface hierarchy in susceptibility to dental caries [Batchelor and Sheiham, 2004]. In general, pit-and-fissure surfaces exhibit much greater risk of developing carious lesions than do smooth surfaces [Batchelor and Sheiham, 2004], and progression of decay differs between these surface types.

Just as environmental risk factors exert differential effects between pit-and-fissure and smooth surfaces – e.g., exposure to fluoride better protects smooth surfaces [Jiang et al., 2005], toothbrushing frequency and sugary drinks have greater effects on pit-and-fissure surfaces [Warren et al., 2006] – we hypothesize that genetic factors may also differentially affect pit-and-fissure and smooth surfaces. For example, genetic factors that modulate the differential effects of environmental exposures (e.g., fluoride-sensitivity genes and taste-preference genes), or that are involved in the patterning of tooth morphology during tooth development, may affect pit-and-fissure and smooth surfaces differently. This hypothesis is supported by family studies showing that the heritability of dental caries between pit-and-fissure and smooth surfaces is only partly shared [Shaffer et al., 2012]. Investigating pit-and-fissure and smooth surfaces separately may be beneficial for advancing our understanding of cariogenesis.

Although the importance of genetic factors on dental caries is well established [Townsend et al., 1998], and heritability has been estimated to be 30–50% [Conry et al., 1993; Bretz et al., 2005; Wang et al., 2010; Shaffer et al., 2012], few specific genes influencing disease risk have been discovered. Previous candidate gene studies and a recent genome-wide association scan [Shaffer et al., 2011] have implicated several loci, although few of these findings have been replicated in follow-up samples or corroborated via functional analyses. The complexity of the phenotype and the possible heterogeneity in risk factors presents obstacles for investigating the genetics of caries. Because specific genes could exert stronger effects on one surface type than the other, and because risk and progression of decay differs by surface type, a reasonable approach to studying the genetic causes of dental caries is to investigate pit-and-fissure and smooth surfaces separately. As part of the Gene Environment Association Studies (GENEVA) consortium, we carried out genome-wide association studies (GWAS) to explore the genetic factors affecting dental caries of primary tooth pit-and-fissure and smooth surfaces.

Subjects and Methods

Participants were drawn from two studies: the Center for Oral Health Research in Appalachia (COHRA) [Polk et al., 2008] and the Iowa Fluoride Study (IFS) [Levy et al., 2003]. All participants provided assent with written parental informed consent, and all study procedures were approved by Institutional Review Boards at the pertinent universities.

All surfaces of all primary teeth were individually scored based on visual inspection for each participant by dental experts who were calibrated across study sites annually. Two classes of tooth surfaces were defined by similarities in morphology and risk of developing carious lesions: pit-and-fissure surfaces and smooth surfaces. The traditional primary dentition caries index, decayed and filled surfaces (dfs), was partitioned to generate two phenotypes: dfsPF and dfsSMA, calculated as the summation of surfaces scored as white spot lesion, decayed, or filled, respectively, for pit-and-fissure and smooth surface types.

The COHRA sample was genotyped by the Center for Inherited Disease Research at Johns Hopkins University using the Illumina Human610-Quadv1_B BeadChip (Illumina Inc., San Diego, Calif., USA). Two subsamples of IFS (denoted as IFS1 and IFS2) were genotyped at different labs. IFS1 was genotyped together with the COHRA sample at the Center for Inherited Disease Research. IFS2 was genotyped separately by the Genetic Analysis Platform at the Broad Institute of MIT and Harvard (Broad) using the same Illumina SNPChip. Additional ungenotyped autosomal single nucleotide polymorphisms (SNPs) and sporadic missing data of genotyped autosomal SNPs were imputed for COHRA and IFS1 samples using BEAGLE software [Browning and Browning, 2009] and the HapMap Phase III reference panel. To ensure the appropriateness of the reference panel, only study subjects of European ancestry defined via principal components analysis were included for imputation. Imputation quality scores, defined as the average posterior probability of the most likely genotype across all samples, were high. 531,025 genotyped SNPs passed all the filters and were included for dfsSMA analysis and 531,230 for dfsPF analysis. Including imputed SNPs, the numbers were increased to 1,216,189 and 1,216,074, respectively.

Because specific genes could exert stronger effects on one surface type than the other, and because risk and progression of decay differs by surface type, a reasonable approach to studying the genetic causes of dental caries is to investigate pit-and-fissure and smooth surfaces separately. As part of the Gene Environment Association Studies (GENEVA) consortium, we carried out genome-wide association studies (GWAS) to explore the genetic factors affecting dental caries of primary tooth pit-and-fissure and smooth surfaces.
samples differ in many aspects, including age, social-economic status, and living environment, we combined results using meta-analysis, which eliminates the risk of possible confounding effects introduced by these differences. Specifically, we combined the COHRA and IFS association results using Stouffer’s p value-based method of meta-analysis as implemented in the METAL software [Willer et al., 2010].

All the analyses were limited to self-reported white children with genetically verified European ancestry. Teeth having only smooth surfaces, such as central and lateral incisors, erupt and are exfoliated before teeth having pit-and-fissure surfaces, such as molars. Due to this eruption and exfoliation pattern of the primary dentition, different age (at examination) ranges were considered for analyses of smooth and pit-and-fissure surfaces. Children aged 3–12 years and having at least one smooth surface present were considered for smooth surface analysis. Children aged 4–14 years and having at least one pit-and-fissure surface present were considered for pit-and-fissure surface analysis. A total of 1,006 children were included for smooth surface analysis (among them, 919 also having imputed data), and 979 children were included for pit-and-fissure surface analysis (among them, 895 also having imputed data). Due to the household- and community-based study design, the COHRA sample contains some nuclear and extended biological relatives. To maintain statistical power, related samples were not excluded, though theoretically this source of population substructure among participants could lead to inflated p values. However, by monitoring p values for evidence of genomic inflation, we found that the inclusion of relatives did not adversely impact the results. Moreover, by reanalyzing SNPs of interest using family-based methods that condition on the relatedness among participants in the sample [Almasy and Blangero, 1998], we observed that unadjusted p values were very similar to family-adjusted p values (see online suppl. table; for all online suppl. material, see www.karger.com/doi/10.1159/000356299). Sample sizes for each subset are shown in table 1.

Linear regression was used to test the genetic association between each SNP and dfsSM or dfsPF under additive genetic model. Due to the comparatively wider age range in COHRA, the varying number of exposed surfaces, and the varying durations of exposure to cariogenic processes, we adjusted for age and age^2 in the analyses of the COHRA sample. In contrast, we only adjusted for age in the analyses of the IFS sample, due to the comparatively narrower age range of the IFS sample (from 2.9 to 7.6). A conservative threshold for genome-wide significance was set at \( \alpha = 5.0E^{-8} \). A threshold for suggestive significance was set at \( \alpha = 1.0E^{-5} \). Genome-wide significant associations, as well as suggestive associations near plausible caries genes, were reported. The effect of nonsynonymous SNPs on protein function (tolerated or damaging) was predicted using SIFT [Kumar et al., 2009].

Additional details of the phenotype assessment, genotyping and quality assurance, and statistical analysis are described in the online supplementary methods.

### Results

Previously, we have reported results of GWAS for a dichotomized (yes/no) childhood caries phenotype in the COHRA and IFS samples [Shaffer et al., 2011]. In the present study, we extend this work by analyzing dfs scores in smooth and pit-and-fissure surfaces, separately, ac-

---

**Table 1. Characteristics of COHRA and IFS samples by caries phenotype**

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Sample characteristic</th>
<th>COHRA</th>
<th>IFS1</th>
<th>IFS2</th>
</tr>
</thead>
<tbody>
<tr>
<td>dfsSM</td>
<td>Age at exam, years[^1]</td>
<td>7.31±2.43</td>
<td>5.13±0.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age range</td>
<td>3.0–12.0</td>
<td>2.9–7.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex ratio (males to females)</td>
<td>1.13</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Smooth caries prevalence, %</td>
<td>48.3</td>
<td>18.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dfsSM[^1]</td>
<td>3.00±5.48</td>
<td>0.81±2.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genotyping center</td>
<td>CIDR</td>
<td>CIDR</td>
<td>Broad</td>
</tr>
<tr>
<td></td>
<td>Total sample size</td>
<td>596</td>
<td>397</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Samples genotyped</td>
<td>547</td>
<td>397</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Samples with imputed data</td>
<td>525</td>
<td>394</td>
<td>NA</td>
</tr>
<tr>
<td>dfsPF</td>
<td>Age at exam, years[^1]</td>
<td>7.97±2.44</td>
<td>5.13±0.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age range</td>
<td>4.0–14.0</td>
<td>2.9–7.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex ratio (males to females)</td>
<td>1.10</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pit-and-fissure caries prevalence, %</td>
<td>57.3</td>
<td>31.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dfsPF[^1]</td>
<td>2.51±3.38</td>
<td>1.02±2.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Genotyping center</td>
<td>CIDR</td>
<td>CIDR</td>
<td>Broad</td>
</tr>
<tr>
<td></td>
<td>Total sample size</td>
<td>572</td>
<td>397</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Samples genotyped</td>
<td>520</td>
<td>397</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Samples with imputed data</td>
<td>501</td>
<td>394</td>
<td>NA</td>
</tr>
</tbody>
</table>

[^1] Mean ± SD.
GWAS of Childhood Surface-Level Caries

According to our hypothesis that genes may differentially affect risk of caries by surface morphology.

On average, COHRA children had 3.0 carious smooth surfaces (dfsSM; SD = 5.5) and 2.5 carious pit-and-fissure surfaces (dfsPF; SD = 3.4). In contrast, IFS children had 0.8 carious smooth surfaces (dfsSM; SD = 2.6) and 1.0 carious pit-and-fissure surfaces (dfsPF; SD = 2.1), reflecting the differences in age, demography and environmental factors between the samples. Additional descriptive statistics for the two samples are summarized in Table 1.

Manhattan and quantile-quantile plots of dfsSM and dfsPF GWAS for combined genotyped and imputed SNPs are shown in Figure 1. The genomic inflation factors (λ) are 1.028 for pit-and-fissure surface scan and 1.031 for smooth surface scan. Percentage of p values used to calculate λ was set at 0.95. The two horizontal lines in Manhattan plots represent suggestive significance threshold (p value <1.0E–5) and genome-wide significance threshold (p value <5.0E–8), respectively. Diagonal lines in QQ plots represent the expected distribution of p values under the null hypothesis of no association. Genotyped and imputed SNPs are plotted together.

One genome-wide significant association was observed for dfsPF in the intronic region of KPNA4, encoding a nuclear protein importer (3q26.1, rs17236529, p value = 2.0E–9; Fig. 2a). No function of this gene or adjacent genes is currently known to be related to cariogenesis.

In contrast, several suggestive association signals (p values ≤1.0E–5) near plausible caries genes were observed. Notably, a missense SNP rs1064524 in ITGAL (16p11.2, p value = 2.1E–7) was suggestively associated with dfsSM (Fig. 2b). The T allele causes an amino acid change from arginine to tryptophan, which was predicted to be damaging with high confidence by SIFT. Although associations are of the same direction for the two samples, the effect sizes differ dramatically. On average, each T allele predicted an increase of 0.54 carious surfaces in the COHRA sample and 2.62 carious surfaces in the Iowa sample. ITGAL encodes the integrin alpha L chain of the

Fig. 1. Manhattan plots (left panel) and quantile-quantile (QQ) plots (right panel) for GWAS of dfsPF and dfsSM. a Pit-and-fissure surface scan. b Smooth surface scan. Manhattan plots show −log_{10}-transformed p values (y axis) across the whole genome (x axis). The genomic inflation factors (λ) are 1.028 for pit-and-fissure surface scan and 1.031 for smooth surface scan. Percentage of p values used to calculate λ was set at 0.95. The two horizontal lines in Manhattan plots represent suggestive significance threshold (p value <1.0E–5) and genome-wide significance threshold (p value <5.0E–8), respectively. Diagonal lines in QQ plots represent the expected distribution of p values under the null hypothesis of no association. Genotyped and imputed SNPs are plotted together.
integrin lymphocyte function-associated antigen-1. Lymphocyte function-associated antigen-1 is expressed on all leukocytes and plays a central role in leukocyte intercellular adhesion and also functions in lymphocyte costimulatory signaling. A study that investigated the expression of ITGAL in peripheral blood mononuclear cells found a significantly higher level within the CD4(+) and CD8(+) T cells in chronic periodontitis and aggressive periodontitis patients than in healthy controls, suggesting the participation of ITGAL in the pathogenesis of periodontal lesions [Lima et al., 2011]. It is unknown whether ITGAL may affect dental caries in a similar fashion.

A suggestive association with dfsSM was observed for rs17124372 (p value = 1.99E–6; fig. 2e). This association represents replication of a result from a recent GWAS involving a nonoverlapping subset of adult individuals from the COHRA sample [Shaffer et al., 2013a]. GWAS of novel caries patterns identified by clustering analysis [Shaffer et al., 2013b] in the permanent dentition implicated the region near AJAP1 for caries of the maxillary premolars and canines (p value = 2.4E–8). Although no study directly links the function of AJAP1 to dental caries, its product interacts with basigin [Schreiner et al., 2007], a plasma membrane protein expressed in both human and rat tooth germs that appears to serve as an inducer of matrix metalloproteinase activity during tooth development [Kumamoto and Ooya, 2006; Schwab et al., 2007].

We also observed an SNP 88 kb from the 5′ end of RPS6KA2 suggestively associated with dfsSM (rs3798305, p value = 7.3E–6; fig. 2f). The product of RPS6KA2 is a kinase in p38-dependent mitogen-activated protein kinase signaling important for oral-related diseases including dental caries. SNPs in an intron of this gene have been found suggestively associated with both tooth-level caries and smooth surface caries scores in independent studies of the permanent dentition [Wang et al., 2012; Zeng et al., 2013].

We observed a suggestive association with a region near AJAP1 on 1p36 for dfsSM (rs4654438, p value = 1.6E–6; fig. 2e). This association represents replication of a result from a recent GWAS involving a nonoverlapping subset of adult individuals from the COHRA sample [Shaffer et al., 2013a]. GWAS of novel caries patterns identified by clustering analysis [Shaffer et al., 2013b] in the permanent dentition implicated the region near AJAP1 for caries of the maxillary premolars and canines (p value = 2.4E–8). Although no study directly links the function of AJAP1 to dental caries, its product interacts with basigin [Schreiner et al., 2007], a plasma membrane protein expressed in both human and rat tooth germs that appears to serve as an inducer of matrix metalloproteinase activity during tooth development [Kumamoto and Ooya, 2006; Schwab et al., 2007].

Association results for the top ranking SNP at each locus are shown.

Table 2. Genome-wide significance and suggestive loci

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Chr</th>
<th>SNP</th>
<th>BP</th>
<th>Beta (IFS)</th>
<th>Beta (COHRA)</th>
<th>p value (meta-analysis)</th>
<th>Gene of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>dfsSM</td>
<td>1</td>
<td>rs4654438</td>
<td>4,459,761</td>
<td>–0.30</td>
<td>–1.63</td>
<td>1.60E–6</td>
<td>AJAP1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>rs3798305</td>
<td>167,283,463</td>
<td>2.75</td>
<td>4.83</td>
<td>7.26E–6</td>
<td>RPS6KA2</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>rs1064524</td>
<td>30,400,324</td>
<td>2.62</td>
<td>0.54</td>
<td>2.07E–7</td>
<td>ITGAL</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>rs17124372</td>
<td>31,262,098</td>
<td>–0.89</td>
<td>–5.20</td>
<td>1.99E–6</td>
<td>PLUNC family genes</td>
</tr>
<tr>
<td>dfsPF</td>
<td>3</td>
<td>rs17236529</td>
<td>161,709,001</td>
<td>2.81</td>
<td>1.78</td>
<td>2.03E–9</td>
<td>KPNA4</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>rs7121800</td>
<td>30,521,153</td>
<td>0.49</td>
<td>0.71</td>
<td>6.85E–6</td>
<td>MPPED2</td>
</tr>
</tbody>
</table>

Association results for the top ranking SNP at each locus are shown.

1 Base pair position for reference genome build 36.3 hg18. 2 Imputed SNP.
Fig. 2. LocusZoom plots for regions of interest. 

a. Genome-wide significant locus in KPNA4 for dfsPF. 
b. Suggestive locus in ITGAL for dfsSM. 
c. Suggestive locus near PLUNC family genes for dfsSM. 
d. Suggestive locus in MPPED2 for dfsPF. 
e. Suggestive locus near AJAP1 for dfsSM. 
f. Suggestive locus near RPS6KA2 for dfsSM. 

Negative log_{10}-transformed p values and physical positions for SNPs in the regions are shown. The recombination rate overlay as well as colors indicating linkage disequilibrium between the top SNP (purple) and other SNPs are based on 1000 Genomes CEU data (hg18) (colors refer to the online version only). The rug plot indicates regional SNP density. Gene positions and directions of transcription are annotated. Circles represent genotyped SNPs, and triangles represent imputed SNPs.
attaining the suggestive significance threshold can be found in the online supplementary table, including family-adjusted p values for the COHRA sample.

Discussion

To explore our hypothesis that genes may exert differential effects across tooth surfaces of differing morphology, we performed separate GWAS for smooth and pit-and-fissure surface caries in the primary dentition in children sampled from two populations, a high-risk Appalachian population (COHRA), and comparatively lower risk Iowan population (IFS). We performed meta-analyses to combine results of COHRA and IFS. Novel genes and loci for childhood caries were nominated and putative caries genes previously identified in other studies were replicated.

The association of SNP rs1064524 in ITGAL with dfsSM is of particular interest. Not only was the associated variant a missense SNP, but its associations in the two samples were in the same direction, with the predicted damaging allele (T) associated with more severe caries experience. This suggests that rs1064524 might be a causal mutation. Further functional studies are warranted to investigate its role in cariogenesis.

Suggestive associations within or near AJAPI and RPS6KA2 have been observed for permanent dentition caries phenotypes in other studies involving adult participants from the COHRA sample [Wang et al., 2012; Shaffer et al., 2013a]. We detected suggestive association signals near the two genes (but different associated SNPs) for dfsSM in the current study. These findings made the associations of the two genes more convincing because the samples for primary and permanent caries studies comprise different participants from the same underlying population. The two genes could reflect common genetic risk factors for caries in both the primary and permanent dentition.

Although we nominated some different genes for pit-and-fissure and smooth surface caries, we also observed a number of suggestive loci for caries of both types of surfaces. These findings are consistent with previously reported estimates of genetic correlation between dfsPF and dfsSM based on the COHRA sample [Shaffer et al., 2012], suggesting that approximately 58% of the heritability of the two caries scores was explained by common genetic factors and 42% by surface-specific genetic factors.

Moderate genomic inflation was observed (λ = 1.031 for smooth surface scan and λ = 1.028 for pit-and-fissure surface scan), which is common for GWAS of complex diseases, and may be due to a variety of potential causes. These include population stratification, known or cryptic relatedness, and model misspecification (such as nonnormal phenotype distributions). Additionally, under polygenic inheritance some degree of genomic inflation is expected, depending on sample size, heritability, linkage disequilibrium structure, and the number of true causal variants [Yang et al., 2011]. We speculate that all or some of these causes contribute to the observed genomic inflation in our study. Rather than attempting to identify and adjust for the causes of this inflation, we instead chose to report original p values. This was because the main aim of the current study was to identify new caries genes and generate hypotheses to be tested in additional studies. Therefore, we focused on the rank of SNPs rather than their nominal p values. In light of this moderate inflation, caution should be exerted when interpreting p values from the current GWAS with regard to strict thresholds for statistical significance. Note that, after adjusting for λ, the top SNP for pit-and-fissure surface scan (rs17236529) stays genome-wide significant.

Though the number of participants was modest (for a GWAS), gene-mapping of dental caries is still in its infancy, so pioneering efforts in cohorts such as ours are valuable. Indeed, our sample size of approximately 1,000 was comparable to other GWAS efforts for oral health phenotypes [Schaefer et al., 2010; Shaffer et al., 2011; Divaris et al., 2012; Wang et al., 2012; Shaffer et al., 2013a; Zeng et al., 2013]. We retained related participants in order to maintain statistical power. While not feasible at the genome level, family-based likelihood methods [Almasy and Blangero, 1998], which condition of the biological relationships among the sample, were used to confirm the top hits; p values for unadjusted and family-adjusted associations were not meaningfully different and did not alter interpretation of the results (see online suppl. table).

In summary, we performed GWAS of surface-level caries scores in the primary dentition and nominated KPNA4, ITGAL, and PLUNC family genes as caries susceptibility genes, as well as replicated the associations for MPPED2, AJAPI and RPS6KA2. Replications in additional samples are warranted to confirm the associations of newly nominated genes with dental caries. More studies on the biological functions of these genes are needed to disentangle their roles in cariogenesis. Understanding the genetic nature of caries etiology could ultimately lead to personalized interventions and preventive measures, and could advance the shift in paradigm of dental care from restoration to caries prevention.
Acknowledgments

Foremost, we would like to thank the participants and research staff of the IFS and COHRA studies whose contributions have made this work possible. We would also like to thank the NIH-initiated GENEVA consortium (www.genevastudy.org) for its central role in developing methods and pipelines for the quality assurance and analysis of genome-wide genetic data. This work was funded by NIH grants U01-DE018903, R01-DE014899, R03-DE021425, R01-DE09551, and R01-DE12101. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Author contributions to this work are as follows: R.J.W., R.C., D.W.M., and M.L.M. conceived and designed the IFS study; S.M.L. conceived and designed the COHRA study; S.M.L. conceived and designed the IFS study; B.B. cleaned and processed the IFS caries and demography data; Z.Z., E.F., X.W., D.E.W., M.L., K.T.C., M.L.M., and J.R.S. cleaned, quality-checked, imputed, and processed the genomics data; Z.Z., E.F., X.W., and J.R.S. performed the statistical analysis; Z.Z., E.F., X.W., D.E.W., M.L., K.T.C., S.M.L., M.L.M., and J.R.S. interpreted the results; Z.Z. and J.R.S. wrote the manuscript; all authors read, revised, and approved the manuscript.

Disclosure Statement

All authors have no conflicts of interest.

References


