EXPLORING THE GENETIC BASIS OF EARLY CHILDHOOD CARIES

Jami Lee Ballantine

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillments of the requirements for the degree in Master in Science in the School of Dentistry (Pediatric Dentistry).

Chapel Hill
2017

Approved by:

Kimon Divaris

Michael Roberts

Andrea Zandona

Cary Agler

ABSTRACT

Jami L. Ballantine: Exploring the Genetic Basis of Early Childhood Caries
(Under the direction of Kimon Divaris)

Objectives: We sought to 1) determine the proportion of ECC variance explained by high-density genotyping of the human genome, and 2) identify associated genetic loci.

Methods: The discovery sample comprised 212 children participating in the study (ZOE) in North Carolina. DNA was purified from saliva and genotyped using Illumina HumanOmni2.5-8 arrays. "Prioritized" SNPs with $P<5x10^{-5}$ were examined for replication in 3 independent cohorts of preschool-age children from the COHRA (n=326), IFS (n=348) and IHS (n=247) studies.

Results: Fifty-two percent ($P=0.03$) of ECC variance was explained by all SNPs with minor allele frequency ≥5%. We found no genome-wide significant association ($P<5x10^{-8}$), but 13 loci had $P<5x10^{-5}$. No locus met statistical significance criteria in the replication samples.

Conclusion: These results affirm a heritable component of ECC and demonstrate the feasibility of conducting genomics studies among preschool-age children.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ECC   Early Childhood Caries

ZOE   Zero Out Early Childhood Tooth Decay

SNP   Single nucleotide polymorphism

GWAS   Genome wide association study

COHRA   Center for Oral Health Research in Appalachia

IFS   Iowa Fluoride Study

IHS   Iowa Head Start Study

GCTA   Genome-wide Complex Trait Analysis

MAF   Minor allele frequency

QQ   Quantile-quantile

INTRODUCTION

Early childhood caries (ECC) is the most common chronic childhood disease and an important public health problem. From a pathogenetic standpoint, it is a complex oral disease wherein acid produced from carbohydrate fermentation by a dysbiotic supragingival biofilm results in progressive demineralization of a susceptible tooth surface.[1] Despite decreasing in prevalence in older age groups, dental caries has persisted and in some cases increased among preschool-age children, with recent estimates indicating that about one out of four U.S. children has experienced ECC.[2] Joint efforts of professional, academic, community and policy stakeholders are focused on addressing this important health problem,[3] which tends to disproportionally affect children in families from lower socioeconomic strata.[4]

ECC is specifically defined as one more decayed (noncavitated or cavitated), missing (due to caries) or filled tooth surfaces in a child 71 months or younger.[5] ECC can be aggressive and negatively affects the quality of life of children and their families.[1,6,7] Its many consequences include pain, trouble eating, difficulty sleeping, behavior changes, as well as impacts on school performance and self-esteem. Restorative care often requires sedation or general anesthesia and thus imposes high personal and societal costs.[8] Furthermore, restorative care is not curative and often fails to arrest the disease process.[9]

Despite the latest advances and widespread application of chemotherapeutics (e.g. fluoride and xylitol formulations), oral hygiene instruction and behavioral interventions, caries risk assessment and programs to treat early signs of tooth decay, improvements in the prevention

of ECC at the population level have been negligible. Understanding that similar measures have been effective in adolescents and adults suggests that explanations for the persistent ECC prevalence need to expand beyond traditional behavioral and socioeconomic risk markers. This also suggests that the traditional behavioral, chemotherapeutic, and systematic/environmental approaches alone are insufficient to control ECC at the person-level, and additional approaches are warranted.

It is well-understood that caries is largely determined by behavioral and environmental risk factors, such as diet and fluoride exposure.[1] However, since the late 50s, dental caries has been shown to have a substantial genetic component.[10-13] Reported heritability (the proportion of disease variance explained by genetics) estimates have ranged from 30 to 70%, with higher estimates observed in primary versus permanent dentition caries.[13-16] A logical next-step would be the identification of specific genetic risk factors (e.g., genes, regulatory elements, or others) that underlie ECC. Numerous candidate-gene studies have since been conducted to investigate the postulated role of several hypothesized genes in caries etiology. These studies have largely targeted enamel development and mineralization genes, as well as genes involved in the immune response in early childhood. As reviewed by Vieira et al., these studies have had mixed results and currently no consensus knowledge of the genetic basis of ECC exists.[17]

A major breakthrough in the approaches available to study the genetic basis of human traits (including states of health and disease) has been the advent of genome-wide association studies (GWAS), which allow the 'unbiased', or hypothesis-free examination of trait-associations with the human genome. In GWAS, hundreds of thousands or millions of genomic markers [(mainly single nucleotide polymorphisms (SNPs)] are obtained via high-density genetic analyses (genotyping) for each individual, and they are tested for association with traits or

diseases of interest. GWAS have led to the discovery of genetic associations of several common, multifactorial diseases such as asthma, diabetes, colorectal cancer, cardiovascular and psychiatric conditions.[18-23] To-date, only one GWAS of caries in the primary dentition has been conducted, among a sample of 1,300 European-American (white) children ages 3-12 years old.[24] That study identified several loci (areas of the genome) with 'suggestive' evidence of association that may have plausible biological roles in childhood caries, but did not identify any genome-wide statistically significant associations. The inclusion of a large proportion of children ages 6-12 in that study precluded any inferences regarding ECC; primary teeth begin to exfoliate and therefore evidence of ECC progressively gets lost after the age of 6 unless complete history records are maintained.

Because it has been established that caries experience is to some degree heritable and no GWAS of ECC has been conducted, the logical next step is the conduct of a genome-wide scan for ECC, including a study sample that is exclusively in the age range of ECC (the Shaffer study[24] included older children, beyond the age of 6). Moreover, the inclusion of diverse populations is desirable in this domain, since most studies to date consist largely of Caucasian populations. To address these knowledge gaps and verify the heritability of ECC (*versus* dental caries in the primary dentition, and in general), we sought to conduct a pilot GWAS of ECC in a multi-racial/ethnic sample of preschool children (ages 3-to-5 years old) enrolled in a community-based study of childhood oral health. To further expand our sample size and examine the possible replication and generalizability of our findings, we sought to examine our most promising candidates in external cohorts of preschool children with available clinical and genomic data. Our study's specific aims were to 1) investigate the proportion of ECC variance

(heritability) explained by high-density genotyping of the human genome obtained via a GWAS and 2) to identify genetic risk loci for ECC in the context of a GWAS.

MATERIALS & METHODS

This GWAS was conducted using DNA extracted from saliva samples collected from a multi-ethnic sample of 212 low-income preschool children (ages 3-5) enrolled in the ZOE study previously reported by Barakat et al.[25] Briefly, all children were enrolled in Early Head Start programs or living in nearby 'control' locations in North Carolina and were examined by a single clinical examiner (KD) at the child's preschool or a nearby community location. Ascertainment of tooth surface-level caries lesion diagnoses was based on NIDCR visual criteria at the precavitated level, following a toothbrush prophylaxis, compressed air-drying and artificial light. Very good intra-examiner reliability was achieved for this trait (*kappa*=0.85). Sociodemographic and additional behavioral risk factors were collected via structured, computer-assisted parent interviews that were administered in English or Spanish.

Saliva samples were collected alongside the clinical examinations using the Oragene DNA Genotek OG-575 kit, and sufficient quantity and quality DNA was extracted from these saliva samples by the UNC-Chapel Hill Biospecimen Processing facility. Genotyping was done at the UNC-Chapel Hill Mammalian Genotyping core using the Illumina HumanOmni2.5-8 bead chip (offering ~2.5million markers). Genotyping quality control procedures included HapMap-CEPH trios and duplicates, 7 blind duplicate samples, identification of sex and sample mismatches, and generation of sample call and error rates. Genotypes were called using the Illumina GenomeStudio software. Samples and selected SNPs were evaluated based on multiple

quality values calculated by Genome Studio and were excluded if they failed preset standards. Overall, high-quality DNA and genotypes were obtained, as previously reported.[25]

After quality control and exclusion of SNPs not meeting quality criteria, we used Genome-wide Complex Trait Analysis (GCTA) software provided by the Visscher group to estimate the heritability of ECC explained by the remaining panel of approximately 2.3 million SNPs, as well as subgroups of SNPs with MAF greater or equal than 1%, 5% and 10%.[26] The estimation of trait heritability among unrelated individuals, like in this study's sample, is based upon the creation of a genetic relationship matrix (n by n-1) using the available high-density genotypes and its superimposition on a trait-relationship matrix. To derive h^2 and p-values, GCTA employs a random-effects mixed linear model and restricted maximum likelihood regression adjusting for age, sex and ancestry (10 principal components). Heritability was estimated for the binary ECC case definition and a continuous measure of disease severity, the conventional $d_{1,2-3}$mfs index (the sum of decayed, missing due to caries, and filled/restored primary tooth surfaces) wherein precavitated and cavitated caries lesions are enumerated.

Next, the association of the individual SNPs with the binary ECC case definition was examined using conventional logistic regression models adjusting for age and sex. Additional terms (10 principal components) adjusting for ancestry (obtained via the GCTA procedure, above) were included in all models, according to recommended standards in the field of genetic epidemiology. Traditional risk factors such as fluoride exposure or cariogenic diet are not considered confounders of the measured genetic associations and thus not adjusted for in the presented analyses. The genome-wide statistical significance p-value criterion for individual SNPs in GWAS is $P<5\times10^{-8}$. Nevertheless, a more lenient ($P<10^{-5}$) threshold to identify 'suggestive' evidence of association was considered as a means of highlighting additional

candidate genes. Rare genetic variants [those with minor allele frequency (MAF) less than 5%] were excluded since the study was not sufficiently powered for their interrogation and hundreds of thousands of participants may be required for the formal study of 'rare' alleles. The obtained results were visualized using quantile-quantile (QQ), Manhattan and LocusZoom plots.

Replication of the prioritized ($P<5$x10^{-5}) SNP associations was examined in 3 independent cohorts comprising 921 preschool children from the Center for Oral Health Research in Appalachia study (COHRA, n=326; mean age=35 months; ECC prevalence=25%), Iowa Fluoride Study (IFS) n=348; mean age=60 months; ECC prevalence=35%) and the Genetic, Environment and Health Initiative Research Study, (GEIRS n=247; mean age=48 months; ECC prevalence=25%). Replication was considered at 3 levels of decreasing 'generalization strength', including: 1) directional consistency (same association-direction, e.g. "protective") and multiple testing-corrected statistically significant association in the replication samples; 2) directional consistency and nominal ($P<0.05$) statistically significant association; 3) directional consistency between prioritized SNPs associations in the discovery (ZOE) and the 3 replication samples, as determined by a binomial test ($P<0.05$).

RESULTS

The prevalence of ECC among the 212 participating children (mean age=39 months; range=30-52 months) was 38%. The demographic characteristics of this multi-ethnic/racial sample are provided in Table 1. The heritability of the ECC binary case definition was 52% (P=0.03) when we considered all common (~1.4 million) SNPs with MAF ≥5%. This estimate diminished after adjustment for ancestry (10 PCs): h^2=13% (P=0.4; data not shown on Table). Heritability also reduced to 44% with MAF ≥1% (~1.9 million SNPs). Similarly, heritability was markedly lower, at 14% (P=0.01) for ECC severity ($d_{1-2,3}$mfs index) using the same set of non-rare (MAF ≥5%) SNPs compared to the binary ECC case definition (Table 2).

The GWAS resulted in a high-quality set of association results with no evidence of systematic errors (population stratification) as illustrated by the QQ plot shown in Figure 1. No genome-wide significant associations were found (Figure 2); however, 13 loci had $P<5x10^{-5}$ and thus demonstrated 'suggestive' evidence of association. Of those 13 loci, the 2 most significant ones were marked by SNPs rs4690994 and rs439888. Specifically, the intergenic locus on 4q32 (rs4690994) showed the strongest association with ECC [Figure 3a; $P=2.3x10^{-6}$; odds ratio (OR) = 3.5; 95% confidence interval (CI) = 2.1-5.9], followed by the *CLDN14* locus (Figure 3b; $P=5.3x10^{-6}$; rs439888 intronic to the *CLDN14* gene; OR=3.6; 95% CI=2.1-6.2). None of these SNPs have known functional roles.

All SNPs with $P<5x10^{-5}$ were carried forward to replication examination and their associations with ECC were obtained separately within each of the 3 replication samples. A

Bonferroni multiple testing-corrected p-value threshold for replication was then set as 0.05 / (13 x 3) = 1.3×10^{-3}. Of note, no SNP met even nominal statistical significance association criteria, and only 15 of 39 SNP look-ups (13 SNPs in 3 different replication cohorts) showed directional association concordance. In sum, the replication attempt of the 13 suggestive loci was largely negative.

DISCUSSION

In this study, for the first time, we successfully carried out a GWAS of ECC among a community-based sample of preschool children and obtained new insights in the genetic underpinning of the disease. Our results were based on a small sample of only a couple hundred unrelated preschool-age children, but confirmed that a sizeable heritable component of ECC exists and that larger, future studies are likely to discover specific genetic influences for the disease. Moreover, the study demonstrates that GWAS are feasible (i.e., sufficient amount and adequate quality of DNA can be extracted, and that can be carried forward to high-density genotyping) among populations of young children who are not clinic attending. Dental care attendance, (i.e., selection of a 'clinical' sample) can itself render studied populations non-representative (e.g., due to high health literacy and access to preventive care or, conversely, existence of problems and restorative care seeking).

The absence of statistically-confirmed genome-wide association results and the non-replication of our suggestive findings to 3 independent samples of young children imply that the genetic loci reported here may have no real influence on ECC or, more likely, are reflective of the very low statistical power of this GWAS among ~200 children. Nevertheless, our findings demonstrate the feasibility of the overall approach, from the conduct of clinical examinations under field conditions to the genotyping of millions of genetic markers using saliva samples obtained during those exams. Of note, our approach utilizing commercial saliva sampling kits for DNA extraction enabled this procedure in remote locations without the need for specialized

10

equipment; saliva samples are stable and can be stored in room temperatures for up to 5 years until DNA extraction takes place in a laboratory setting.

Our estimates of heritability are in general agreement with previous reports in the literature.[15] Of note, relatively common SNPs (MAF≥5%) explained the most variance in ECC which is unsurprising, given the small sample size of this GWAS, wherein less frequent SNPs may be adding more 'noise' (i.e., variance) than 'true' association signal. Of note, heritability was markedly lower for the dmfs index compared to the ECC case definition. This finding was somewhat expected, because variations introduced by restorative dental treatment (e.g. placement of a full-coverage, stainless steel crown, coded as a 5-surface restoration versus a pre-existing 2-surface caries lesion) can obscure (actually, inflate) the underlying clinical caries experience. An alternative approach to circumvent this issue may be the interrogation of the d-component of the index alone (indicating diseased-only surfaces); however, this metric may also be confounded by access to care issues, which would affect the ratio of treated: untreated disease. On the other hand, heritability decreased substantially when our models were adjusted for ancestry via the inclusion of 10 principal components; this result should be treated with caution, as such adjustments can produce statistically unstable estimates due to the small sample size. Nevertheless, it is indicative of the impact of race/ethnicity-specific influences, which are at play in a racially-mixed sample like in our study. Finally, limitations of the GCTA approach itself in estimating heritability were recently articulated[27] and alternative or improved approaches may soon be available in this rapidly evolving field.

Our study did not consider traditional risk factors for ECC including socio-economic status, diet, oral hygiene and fluoride exposure. As noted earlier, traditional risk factors, although strongly associated with the trait or disease under study, are not confounders of the genetic

11

associations and adjustments are conventionally not done. Nevertheless, stratification by such factors or examination of gene x environment (e.g., fluoride) interactions can be informative[24] and should be explored in cases where the sample size permits. Interestingly, some biological pathways that are genetically-controlled[28] may be operating via clinical (e.g., saliva and enamel properties) or behavioral risk factors, with the most likely being sweet taste preference, as suggested by relatively recent studies.[29-31]

In sum, the major novelty and strength of this study was the opportunity to do an unbiased scan of the human genome without *a priori* hypotheses for the first time, in a narrow-age range sample, appropriate for the study of ECC. This study also benefits from the uniform clinical examination protocol and the opportunity to replicate or generalize its findings to external samples of almost one thousand preschool children. Lastly, although race- or ethnic-specific results were not examined in these analyses due to the small sample size, the inclusion of under-studied racial/ethnic groups in our investigation is a novel element.

CONCLUSIONS

Based on this study's results, we conclude the following:

1. GWAS of oral health traits, including ECC, are feasible among preschool-age children.

2. A substantial heritable component of ECC exists.

3. Future, large multi-ethnic studies, including pooling of findings across samples and collaborating for cross-validation are likely to identify specific genetic influences for ECC, which can help better understand, prevent and treat this early-onset, aggressive childhood disease.
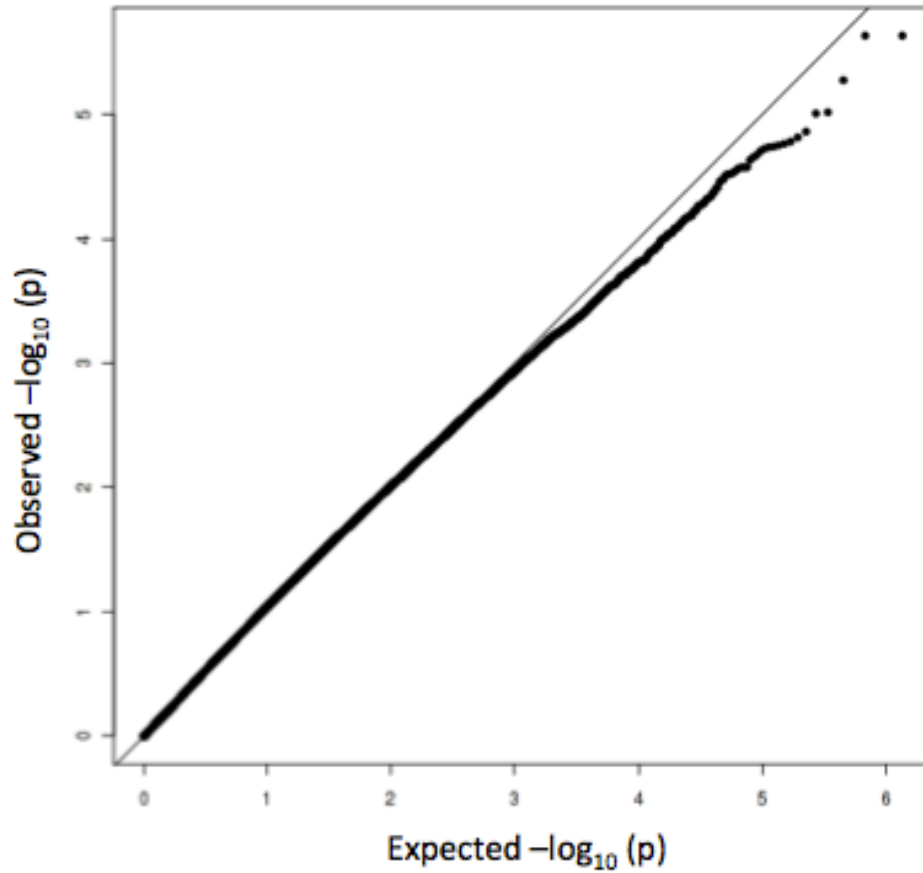
Table 1: Demographic characteristics of the 212 preschool-age children participating in the ZOE GWAS, overall and by early childhood caries (ECC) status.			
	All	ECC	Healthy
	n (col. %)	n (row. %)	n (row %)
Entire Sample	212 (100)	78 (38)	132 (62)
Male	116 (55)	50 (43)	66 (57)
Female	96 (45)	30 (31)	66 (69)
Race/Ethnicity			
African American	67 (32)	26 (39)	41 (61)
Hispanic American/ Latino	74 (35)	34 (46)	40 (54)
European American	49 (23)	13 (27)	36 (73)
Native American	21 (10)	6 (29)	15 (71)
Other	1 (0)	1 (100)	0 (0)

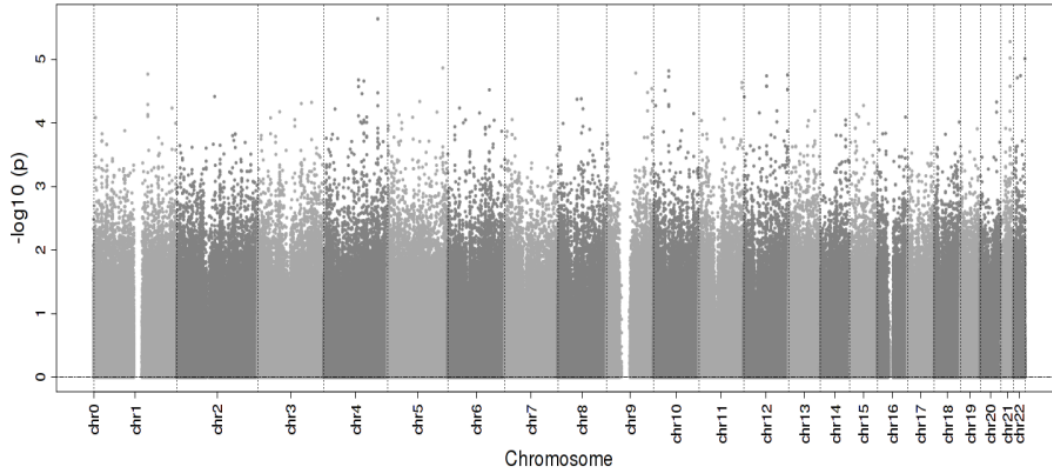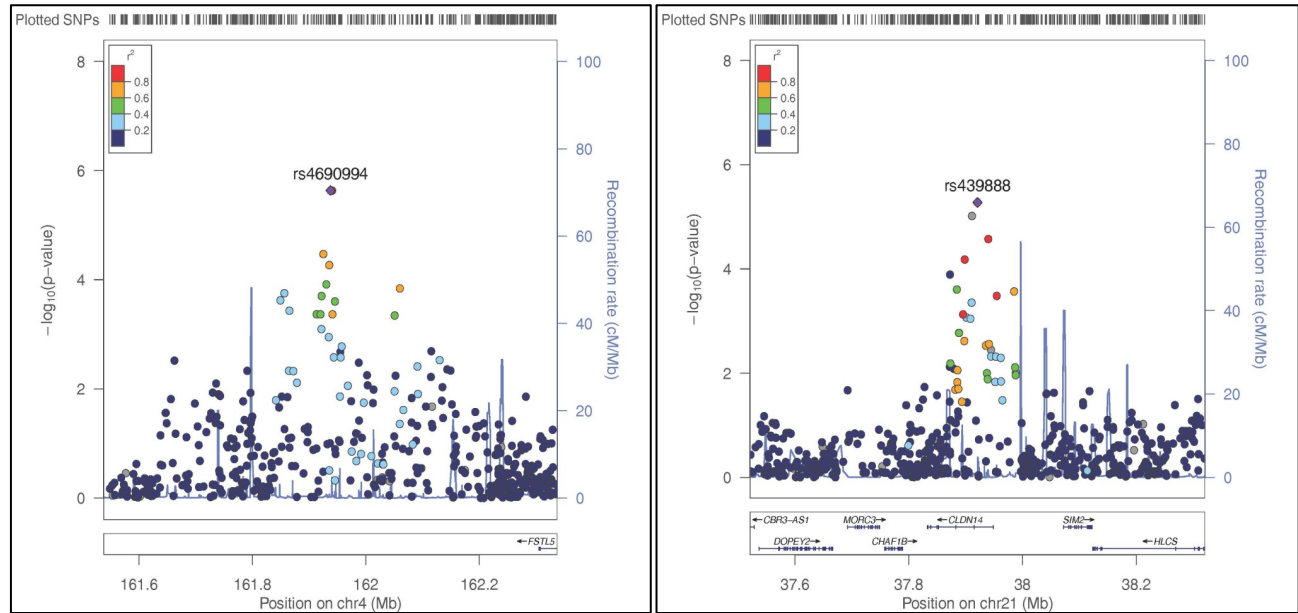|	ECC case status (binary definition)		ECC severity ($d_1d_{2,3}$mfs index)	
Table 2: Phenotypic variance explained for ECC case status and severity ($d_{1\text{-}2,3}$mfs index) among the 212 preschool-age children enrolled in the ZOE GWAS.				
	Variance explained (SE[†])	LR[*] p-value	Variance explained (SE[†])	LR[*] p-value
ALL SNPs (n= 2,331,188)				
Only SNPs considered	0.43 (0.36)	0.043	0.06 (0.08)	0.12
+Age, Sex	0.38 (0.36)	0.07	0.06 (0.08)	0.13
MAF[‡] >0.01 (n= 1,877,037)				
Only SNPs considered	0.44 (0.39)	0.034	0.07 (0.08)	0.08
+Age, Sex	0.39 (0.39)	0.061	0.07 (0.08)	0.08
MAF[‡] ≥0.05 (n= 1,382,931)				
Only SNPs considered	0.52 (0.55)	0.026	0.14 (0.14)	0.01
+Age, Sex	0.43 (0.50)	0.050	0.13 (0.14)	0.02
MAF[‡] ≥0.10 (n= 986,805)				
Only SNPs considered	0.48 (0.53)	0.026	0.21 (0.20)	0.006
+Age, Sex	0.39 (0.48)	0.050	0.19 (0.19)	0.008
*LR, Likelihood Ratio X^2 test; †SE, standard error; ‡MAF, minor allele frequency				

FIGURES

Figure 1: Quantile-Quantile (QQ) plot of GWAS results of ECC among the 212 preschool-age children participating in the ZOE GWAS. This QQ plot illustrates 2.3 million observed (y-axis) versus expected (x-axis) association results [-$\log_{10}$(p-values)] is based on logistic regression genetic models of ECC. The models assumed multiplicative allelic effects and included adjustment for ancestry (10 principal components), age and sex.

Figure 2: Manhattan plot of the ~2.3 million association results [y-axis corresponds to -log$_{10}$(p-value)] of genotyped SNPs with the ECC case definition, arranged by chromosome, among the 212 preschool-age children participating in the ZOE genome-wide association study.

Figure 3. LocusZoom (Pruim et al., 2010) plots of the 2 top loci that emerged as the most significantly associated with ECC among the 212 preschool-age children participating in the ZOE genome-wide association study, left panel (a): rs4690994, P=2.3x10[-6]; odds ratio (OR)=3.5; 95% confidence interval (CI)=2.1-5.9]; right panel (b): the *CLDN14* locus (P=5.3x10[-6]; rs439888 intronic SNP; OR=3.6; 95% CI=2.1-6.2). Position on the x-axis corresponds to genomic coordinates (position), and the position on the left y-axis corresponds to each SNP's –log$_{10}$(p-value). The top, or "lead", SNP is colored purple, while other polymorphisms are color-coded by their r^2, a measure of linkage disequilibrium, with the lead SNP.

REFERENCES

1. Selwitz RH, Ismail AI, Pitts NB. Dental Caries. Lancet. 2007 Jan 6;369(9555):51-59.

2. Dye BA, Tan S, Smith V, Lewis BG, Barker LK, Thornton-Evans G, Eke PI, Beltrán-Aguilar ED, Horowitz AM, Li CH. Trends in oral health status, United States, 1988-1994 and 1999-2004. Vital Health Stat 11. 2007;(248):1-92.

3. Casamassimo PS, Lee JY, Marazita ML, Milgrom P, Chi DL, Divaris K. Improving children's oral health: an interdisciplinary research framework. J Dent Res. 2014 Oct;93(10):938-42.

4. Lee JY, Divaris K. The ethical imperative of addressing oral health disparities: a unifying framework. J Dent Res. 2014 Mar;93(3):224-30.

5. American Academy of Pediatric Dentistry. "Policy on early childhood caries (ECC): Classification, consequences and preventive strategies [Special Edition]." Pediatric Dentistry 31 (2011):40- 43.

6. Cunnion DT, Spiro A 3rd, Jones JA, Rich SE, Papageorgiou CP, Tate A, Casamassimo P, Hayes C, Garcia RI. Pediatric oral health-related quality of life improvement after treatment of early childhood caries: a prospective multisite study. J Dent Child. 2010 Jan-Apr;77(1):4-11.

7. Filstrup SL, Briskie D, da Fonseca M, Lawrence L, Wandera A, Inglehart MR. Early childhood caries and quality of life: child and parent perspectives. Pediatr Dent. 2003 Sep-Oct;25(5):431-40.

8. Casamassimo PS, Thikkurissy S, Edelstein BL, Maiorini E. Beyond the dmft: the human and economic cost of early childhood caries. J Am Dent Assoc. 2009 Jun;140(6):650-7.

9. Berkowitz RJ, Amante A. Kopycka-Kedzierawski DT, Billings RJ, Feng C. Dental caries recurrence following clinical treatment for severe early childhood caries. Pediatr Dent. 2011 Nov-Dec;33(7):510-514.

10. Goodman HO, Luke JE, Rosen S, Hackel E. Heritability in dental caries, certain oral microflora and salivary components. Am J Hum Genet. 1959 Sep;11:263-73.

11. Mansbridge JN. Heredity and dental caries. J Dent Res. 1959 Mar-Apr;38(2):337-47.

12. Boraas JC, Messer LB, Till MJ. A genetic contribution to dental caries, occlusion, and morphology as demonstrated by twins reared apart. J Dent Res. 1988 Sep;67(9):1150—1155.

13. Conry JP, Messer LB, Boraas JC, Aeppli DP, Bouchard TJ Jr. Dental caries and treatment characteristics in human twins reared apart. Arch Oral Biol. 1993 Nov;38(11):937-43.

14. Bretz WA, Corby PM, Hart TC, Costa S, Coelho MQ, Weyant RJ, Robinson M, Schork NK. Dental caries and microbial acid production in twins. Caries Res. 2005 May-Jun;39(3):168-172.

15. Wang X, Shaffer JR, Weyant RJ, Cuenco KT, DeSensi RS, Crout R, McNeil DW, Marazita ML. Genes and their effects on dental caries may differ between primary and permanent dentitions. Caries Res. 2010;44(3):277-284.

16. Shaffer JR, Wang X, Desensi RS, Wendell S, Weyant RJ, Cuenco KT, Crout R, McNeil DW, Marazita ML. Genetic susceptibility to dental caries on pit and fissure and smooth surfaces. Caries Res. 2012;46(1):38-46.

17. Vieira AR, Modesto A, Marazita M. Caries: Review of Human Genetic Research. Caries Res. 2014;48:491-506.

18. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ionnidis JP, Hirschhorn JN. Genome wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet. 2008 May;9(5):356-369.

19. Sordillo JE, Kelly R, Bunyavanich S, McGeachie M, Qiu W, Croteau-Chonka DC, Soto-Quiros M, Avila L, Celedon JC, Brehm JM, Weiss ST, Gold DR, Litonjua AA. Genome-wide expression profiles identify potential targets for gene-environment interactions in asthma severity. J Allergy Clin Immunol. 2015 Apr 24. pii: S0091-6749(15)00341-3.

20. Barret JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Plagnol V, Pociot F, Schuilenburg H, Smyth DJ, Stevens H, Todd JA, Walker NM, Rich SS; Type 1 Diabetes Genetics Consortium. Nat Genet. 2009 Jun;41(6):703-707.

21. Wang H, Burnett T, Kono S, Haiman CA, Iwasaki M, Wilkens LR, Loo LW, Van Den Berg D, Kolonel LN, Henderson BE, Keku TO, Sandler RS, Signorello LB, Blot WJ, Newcomb PA, Pande M, Amos CI, West DW, Bezieau S, Berndt SI, Zanke BW, Hsu L, Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), Lindor NM, Haile RW, Hopper JL, Jenkins MA, Gallinger S, Case G, Colon Cancer Family Registry (CCFR), Stenzel SL, Schumacher FR, Peters U, Gruber SB, Colorectal Transdisciplinary Study (CORECT), Tsugane S, Stram DO, Le Marchad L. Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A. Nat Commun. 2014 Aug 8;5:4613.

22. Smith JG, Newton-Cheh C. Genome-wide association studies of late-onset cardiovascular disease. J Mol Cell Cardiol. 2015 Jun;83:131-41.

23. The Network and Pathway Analysis Subgroup of the Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015 Feb;18(2):199-209.

24. Shaffer JR, Wang X, Feingold E, Lee M, Begum F, Weeks DE, Cuenco KT, Barmada MM, Wendell SK, Crosslin DR, Lauriee CC, Doheny KF, Pugh EW, Zhang Q, Feenstra B, Geller F, Boyd HA, Zhang H, Melbye M, Murray JC, Weyant RJ, Crout R, McNeil DW, Levy SM, Slayton RL, Willing MC, Broffitt B, Vieira AR, Marazita ML. Genome-wide association scan for childhood caries implicates novel genes. J Dent Res. 2011 Dec;90(12):1457-1462.

25. Barakat D, Dejong H, Basta P, Luo J, Rozier RG, North KE, Divaris K. High-density genotyping in preschoolers using saliva obtained during dental examinations. J Dent Res 2015;93 (Spec Iss A): 2117473 (IADR/AADR). [abstract]

26. Yang J, Lee SH, Goddard ME, Visscher PM. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. Methods Mol Biol 2013;1019:215-36.

27. Krishna Kumar S, Feldman MW, Rehkopf DH, Tuljapurkar S. Limitations of GCTA as a solution to the missing heritability problem. Proc Natl Acad Sci U S A. 2016 Jan 5;113(1):E61-70.

28. Divaris K. Predicting Dental Caries Outcomes in Children: A "Risky" Concept. J Dent Res. 2016 Mar;95(3):248-54.

29. Kulkarni GV, Chng T, Eny KM, Nielsen D, Wessman C, El-Sohemy A. Association of GLUT2 and TAS1R2 genotypes with risk for dental caries. Caries Res. 2013;47(3):219-25.
30. Wendell S, Wang X, Brown M, Cooper ME, DeSensi RS, Weyant RJ, Crout R, McNeil DW, Marazita ML. Taste genes associated with dental caries. J Dent Res. 2010 Nov;89(11):1198-202.

31. Shetty V, B L P, Hegde AM. PROP test: prediction of caries risk by genetic taste perception among the visually impaired children. Spec Care Dentist. 2014 Jan-Feb;34(1):34-40.