GENOME-WIDE ASSOCIATION STUDY AND META-ANALYSIS OF MISSING TEETH AND FUNCTIONAL DENTITION

by

Amanda Everman

BS Biology, The Pennsylvania State University, 2016

Submitted to the Graduate Faculty of

Department of Human Genetics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Amanda Everman

It was defended on

March 20, 2018

and approved by

Thesis Advisor:
John R. Shaffer, PhD
Assistant Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Candace Kammerer, PhD
Associate Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Katherine Neiswanger, PhD, MA
Associate Professor
Department of Oral Biology
School of Dental Medicine
University of Pittsburgh

GENOME-WIDE ASSOCIATION STUDY AND META-ANALYSIS OF MISSING TEETH AND FUNCTIONAL DENTITION

Amanda Everman, MS

University of Pittsburgh, 2018

ABSTRACT

Retention of a functional dentition of twenty or more permanent teeth is an important World Health Organization (WHO) goal, as missing teeth adversely affects oral health quality of life, especially masticatory function and aesthetic appearance and satisfaction. Previous studies show that tooth loss is moderately heritable, and genome-wide association studies (GWAS) of periodontitis and dental caries, the two main causes of tooth loss, have successfully identified genetic variants associated with these oral diseases. Thus, this work aimed to identify genetic variants associated with missing teeth by performing genome-wide association scans in five cohorts and a subsequent meta-analyses. Genome-wide association scans using linear and logistic regression for a quantitative trait and functional dentition, respectively, were performed in five cohorts: The Center for Oral Health Research in Appalachia cohort 1 (COHRA; N = 955), Dental Registry and DNA Repository of the University of Pittsburgh School of Dental Medicine (DRDR; N = 227), and cohorts from The Pittsburgh Orofacial Clefts Studies (POFC) project recruited from the United States (POFC-USA; N = 192), Guatemala (POFC-G; N = 272), and the Patagonia region of Argentina (POFC-PA; N = 182). Three p-value based meta-analyses were performed: a white-only meta-analysis (COHRA and DRDR; N = 1182), a Hispanics-only meta-analysis (POFC-G and POFC-PA; N = 454), and a trans-ethnic meta-analysis (COHRA, DRDR, POFC-G, and POFC-PA; N = 1636). Two regions of the genome were associated with missing

teeth at genome-wide significance ($p < 5 \times 10^{-8}$) and were located near genes relevant to dental and oral health (*POSTN*, a critical regulator of periodontal homeostasis, and *MTRR*, which functions in methionine synthesis, a process previously implicated by GWAS of dental caries.) Furthermore, many regions of the genome showed suggestive significance ($p < 1 \times 10^{-5}$) and were located near genes biologically relevant to tooth loss. These discoveries corroborate existing evidence for a genetic contribution to tooth loss, and supports the hypothesis that common genetic variants influence tooth loss. The public health significance of this work is that such findings may ultimately lead to the identification of individuals at risk for tooth loss and the development of novel treatments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This work was funded through NIH grants U01-DE018903, R01-DE014899, R03-DE024264, R56-DE027055, X01-HG00784, R01-DE016148, U01-DE024425, R37-DE008559, R01-DE009886, R21-DE016930, R01-DE012472, and U01-HG004446, and NIH contract HHSN268200782096C. Support for the DRDR cohort was provided by the University of Pittsburgh School of Dental Medicine.

1.0 INTRODUCTION

1.1 SPECIFIC AIMS

Tooth loss negatively impacts oral health quality of life, since it affects aesthetic appearance, satisfaction, and function. It is also a known marker for subsequent cognitive decline and decline in overall health. Loss of a tooth is a failure of oral health, and complete edentulism, the loss of all teeth, is the ultimate endpoint of poor oral health. Missing teeth are known to be associated with environmental and other risk factors, such as age, socioeconomic status, education, and smoking status. Missing teeth are primarily caused by periodontal disease and dental caries, and known genetic factors are involved in the etiology of both of these oral diseases. Furthermore, heritability studies have shown that tooth loss itself is moderately heritable. While some genetic variants influencing loss or retention of teeth have been described previously via candidate gene studies, such approaches are limited and biased in that they rely on assumptions regarding disease pathology to nominate potential genes, and comprehensive knowledge of the functions of all genes is far from complete. Additionally, no published study to date has used a whole-genome approach to systematically identify regions of the genome that are associated with tooth loss. Application of hypothesis-free, genome-wide approaches may lead to the identification of novel genes and biochemical processes relevant to tooth loss.

Since understanding the genetic basis of tooth loss is expected to lead to improvements in oral health and decrease oral health disparities across populations, we will perform genome-wide association analyses of missing teeth and functional dentition, a binary trait defined as retention of at least 20 teeth, in five independent study samples and carry out a subsequent meta-analysis to identify common variants associated with these phenotypes. Given that dental caries and periodontal disease have genetic components, and that tooth loss itself is moderately heritable, we expect that genetic variants may influence tooth loss. Thus, a genome-wide association study, which can successfully identify common variants of modest effect size influencing traits and diseases, is an appropriate methodology for scanning the genome for regions associated with missing teeth and functional dentition.

In order to pursue the following Specific Aims, this project will bring together five existing data sets: the Center for Oral Health Research in Appalachia cohort 1 (COHRA1) recruited from West Virginia and Pennsylvania, samples from the Dental Registry and DNA Repository (DRDR) of the University of Pittsburgh School of Dental Medicine, and unaffected participants from The Pittsburgh Orofacial Clefts Studies (POFC) project recruited from United States (POFC-USA), Guatemala (POFC-G), and Patagonia region of Argentina (POFC-PA.)

Aim 1 is to perform initial genome-wide association analyses in these five independent study samples (COHRA, DRDR, POFC-USA, POFC-G, and POFC-PA) to identify common variants associated with missing teeth and functional dentition and estimate their effect sizes.

Aim 2 is to perform meta-analyses stratified by ethnicity and a subsequent trans-ethnic meta-analysis. To perform replication analyses of top hits identified in the meta-analyses, POFC-USA will be used as a replication sample.

Aim 3 is to perform bioinformatic analyses of significant SNPs and search the literature for nearby genes with biological roles in dental and oral health. The most significantly associated variants identified in individual cohorts and the meta-analyses across cohorts will be interrogated for evidence of functionality and putative roles affecting dental and oral health.

The hypothesis-generating, genome-wide approach that will be utilized in our study will further characterize the genetic factors influencing tooth loss, which may lead to the identification of novel genes and biochemical processes relevant to tooth loss. Associations between common variants and tooth loss will be identified at both the statistical and biological level. Ultimately, understanding the genetic basis of tooth loss may lead to the identification of at-risk individuals as well as the development of novel treatments.

1.2 THE PUBLIC HEALTH IMPORTANCE OF ORAL HEALTH, TOOTH LOSS, AND FUNCTIONAL DENTITION

1.2.1 Social and Individual Impact of Tooth Loss

Poor oral health is a tremendous public health problem worldwide. As of 2010, the global estimate of severe tooth loss, here defined as the retention of 9 or fewer teeth, was 2.3% or 158 million people, with an incidence rate of 205 cases per 100,000 person-years [1]. Economically, the direct and indirect costs of dental diseases are estimated to be $298 billion and $144 billion per year, respectively, amounting to an estimated $442 billion annually. This cost is in line with the economic losses for the most common global causes of death [2]. From the Global Burden of Diseases, Injuries, and Risk Factors (GBD) Study, in 2010, oral diseases (severe tooth loss,

untreated dental caries, and severe periodontitis) were estimated to account for 15 million disability-adjusted life-years (DALYs) globally; in individuals over age 60, the leading cause of DALYs was severe tooth loss. Furthermore, because of population growth, increases in lifespan, and the cumulative nature of oral diseases, the global burden for oral diseases increased from 1990-2010. This is contrary to the overall trend of a decrease in the global burden of other diseases, indicating that oral health is becoming increasingly important from a global health perspective [3]. Due to the costly nature of oral diseases, it is important to study tooth loss, in part, to alleviate the economic burden that it poses on society.

Oral diseases are not only costly at the economic and societal level, but they also have a severe impact on an individual's quality of life, including oral health quality of life. Oral health quality of life is conceptualized as the effect of dental health on an individual's functional capacity, emotional well-being, satisfaction, and sense of self-worth [4]. Because of the serious burden tooth loss has on oral health quality of life, in 1992, the World Health Organization (WHO) made it a goal to retain a functional dentition of 20 or more permanent teeth through life [5]. Tooth loss affects aesthetic appearance and satisfaction [6] and loss of more than 8 teeth impairs mastication and the ability to eat certain foods [7, 8]. Indeed, completely edentate individuals have been found to consume fewer healthy, nutrient-rich foods, like carrots and salads, than individuals with a functional dentition [9]. Accordingly, denture-wearers consume approximately 2-1.5 times less nutrients, and edentate individuals have a poorer, less balanced diet and worse nutritional intake than dentate individuals, particularly in dietary fiber (1.2 times lower), folate (1.3 times lower), and beta-carotene intake (1.7 times lower) [9, 10]. Zhu and Hollis (2014) also found that individuals without a functional dentition had a 20% lower overall energy intake, and consumed 5% less protein, 25% less dietary fiber, 12% less vitamin A, 14%

less vitamin C, and 4% more carbohydrates [11]. The nutritional deficits associated with tooth loss support the WHO goal of retaining a functional dentition through life in addition to the economic losses of tooth loss.

1.2.2 Epidemiologic Correlates of Tooth Loss, Functional Dentition, and Edentulism

Several well-established environmental risk factors have been associated with tooth loss. Perhaps the biggest predictor of tooth loss is age, since tooth loss is the culmination of oral diseases throughout the life course. Multiple studies of different designs that have sampled varying populations have established the association between older age and tooth loss [1, 12-22]. Furthermore, using data from several versions of the National Health and Nutrition Examination Survey (NHANES), Slade *et al*. (2014) observed a cohort effect on edentulism in U.S. adults, whereby edentulism was higher in individuals over 60 years of age and born before 1934 compared to individuals over 60 years old born after 1934. They also found a significant interaction between age and birth year on the prevalence of edentulism; individuals born before 1934 experienced higher rates of edentulism as they aged compared to other cohorts. Such findings may be due to improvements in dental health care services [12]. Other demographic factors associated with tooth loss include race and a rural location, and there are global patterns of tooth loss, with some nations experiencing high rates and some markedly less. African-Americans experience greater tooth loss and higher rates of edentulism than non-Hispanic whites [12-14], but tooth loss is lowest among Asians [12]. Residents of high-poverty rural areas are 1.66 times more likely to be completely edentulous [13]; areas such as Appalachia and the Mississippi Delta have the highest rates of edentulism (8% or more, compared to 4.9% nation-wide) across the United States (U.S.). This is likely explained by high poverty levels and lack of

dental health care services [12, 13]. Globally, the highest rates of edentulism have been reported in North, South, and Central America, Eurasia, and Australia, and edentulism is lowest in sub-Saharan Africa, China, and Oceania [1].

Socioeconomic factors are also highly correlated with tooth loss across populations and age groups. Income is extremely predictive of missing teeth, as loss of functional dentition, edentulism, and any amount of tooth loss is highest in low income individuals[11-14, 18, 20, 21, 23]. Slade et al. (2014) found the difference between the prevalence of edentulism to be as much as 12.6% vs. 0.6% between highest and lowest income quartiles [12]. Education, which is correlated with income, is likewise negatively correlated with missing teeth [12-16, 18-21, 23]. Marital status (being divorced/separated or unmarried vs. married) has been found to be associated with tooth loss as well [13, 20, 23]. Unemployment and being uninsured have been associated with tooth loss [13], though the body of literature supporting these associations is not as substantial as those supporting income and education. Ultimately, these socioeconomic predictors are not independent of each other, and low socioeconomic status is what predisposes to tooth loss more than any one predictor by itself.

In addition to the demographic and socioeconomic environmental factors that predispose to tooth loss, several behavioral risk factors have been associated with edentulism and missing teeth. Unsurprisingly, smoking is strongly predictive of missing teeth, and current and former smokers are approximately 2.3 times more likely to experience tooth loss than never smokers [11, 13, 18, 20-24]. Similarly, high levels of alcohol consumption are associated with tooth loss [20, 23], though some studies have not identified such an association [21]. While body mass index (BMI) has not been associated with tooth loss and/or edentulism, studies have found that physical activity and sedentary time *are* associated with tooth loss [11, 20, 23]. Oral hygiene

behaviors, such as infrequent flossing [25], dental maintenance visits, tooth brushing, and poor personal oral hygiene behaviors [22] are all associated with tooth loss. Taken together, these data suggest that those who do not practice good oral and general health behaviors are more likely to also have oral diseases and thus tooth loss. These results may explain why associations with traits lacking a clear biological link to tooth loss, such as sedentary time, are observed.

Undeniably, environmental and behavioral risk factors play an important role in the etiology of tooth loss across populations, but chronic diseases, including asthma, diabetes, and cardiovascular disease have also been associated with tooth loss and edentulism. Chronic diseases including asthma, diabetes, and cardiovascular disease have been associated with tooth loss [13, 26]. More specifically, cardiovascular comorbidities, such as coronary heart disease, hypertension, angina pectoris, and heart attack [13, 20, 21, 23, 27], and noncardiovascular comorbidities such as lung disease, asthma, and arthritis have all been linked to missing teeth [20]. Irritable bowel syndrome and measures of psychological well-being, such as depressive symptoms and self-reported psychosocial well-being, have also been associated with missing teeth [20, 23, 28]. Medications that alter the oral environment and cause changes to microbial communities, such as asthma inhalers, may explain an association between chronic diseases like asthma and tooth loss. Additionally, researchers have hypothesized that chronic inflammation and/or nutrient deficiencies may explain the association between cardiovascular disease and tooth loss [26].

The role of biological sex in tooth loss is not yet clear; the potential association between sex and tooth loss has not been consistent across several studies. In cross-sectional data, studies of U.S. populations including only individuals over the age of 60 and studies including all adults reported that females were more likely to be fully edentate [13, 20]. Investigators reported that

female adults of all ages and between 20-64 years experience significantly greater tooth loss ($p < 0.001$) as 53.6% of the poor dentition (<=20 teeth) group was female and 53.1% of the full dentition (>=28 teeth) group was male [11]. However, Ylostalo *et al*. (2006) reported that within a cross-sectional cohort of Finnish adults in their mid-30s, males experienced greater tooth loss [23]. Among individuals having a tooth extracted, not restricting age, males and females were equally likely to have a tooth extracted due to caries and periodontitis, though females were more likely to have a tooth removed due to orthodontics, and males due to trauma [29]. Among Brazilian adults ages 20-59, no association was found between tooth loss and sex when tooth loss was categorized as 10 or more teeth in both arches, less than 10 teeth in one arch, or fully edentulous [21]. In rural Ecuadorians over 60 years, no association between sex and tooth loss was found when using a cutoff of 10 or more teeth [16]. Across five cohorts, Slade *et al*. (2014) found no sex differences in the prevalence of edentulism in U.S. adults [12], and sex has not been associated with edentulism in other cross-sectional studies [15]. Total number of missing teeth has also been found to show no sex differences [30]. Furthermore, in our data from COHRA, with ages restricted to 18-60 years, sex was not associated with either total missing teeth or functional dentition. Inconsistencies in defining tooth loss, such as arbitrary cutoffs for severe tooth loss or using complete edentulism only, and inconsistencies in defining study populations based on age groups, likely contribute to the lack of consensus across studies on the effect of sex on tooth loss. It is not yet clear whether observed sex differences in tooth loss are biological in nature or due to differences in exposures to environmental risk factors such as smoking.

1.2.3 Tooth Loss as a Risk Factor for Cognitive and Physical Decline

Not only is loss of permanent teeth a marker of poor oral health, but tooth loss is associated with poor cognitive performance and it is a significant risk factor for dementia. In cross-sectional studies of Swedish and Brazilian populations, loss of natural teeth and edentulism were associated with poorer cognitive performance (OR = 3.3, 95% CI: [1.2-9.3] in the Brazilian population) even after controlling for other factors associated with cognition, such as socioeconomic status (SES), age, social connectivity, education, and smoking [18, 31]. Among identical twins discordant for dementia, the twin with dementia was approximately 4 times more likely to have a history of tooth loss and poor oral health [32]. The association between different measurements of tooth loss and cognition, memory, and dementia has been found in multiple diverse study samples, from rural Ecuadorians and Mexican Americans to urban Chinese populations (OR = 1.56, 95% CI: [1.12-2.18] in Chinese) [16, 30, 33-37]. While cross-sectional studies cannot infer causality from association, the consistency of this result across many distinct populations indicates that tooth loss and poor dentition may be a strong marker of cognition.

Prospective studies have established the temporal relationship between tooth loss and cognition, showing that tooth loss precedes cognitive decline. In a longitudinal study of British adults age 60 and older, edentate individuals had a poorer cognitive (0.88 fewer words recalled) and physical performance (0.09 m/s slower gait speed) than dentate individuals at baseline, and edentulism was additionally associated with subsequent physical and cognitive decline especially in adults age 60-74. This association held even after controlling for multiple covariates such as age, sex, marital status, etc. [20]. Stein *et al*. (2007) found that retention of fewer than 9 teeth was positively associated with both the prevalence at baseline (OR = 4.3, 95% CI[1.16, 15.60]) and subsequent development of dementia (Hazard Ratio = 2.2, 95% CI: [1.1, 4.5]) in a U.S.

population [19]. Other prospective studies have corroborated this evidence that tooth loss precedes and predicts cognitive decline and development of dementia in elderly and middle aged adults representative of both U.S. and global populations [35, 38-40]. Moreover, these associations, in both cross-sectional and longitudinal studies, are robust to the measurement of cognitive function, as some studies have used the Mini–Mental State Examination (MMSE) or the Delayed Word Recall Test, and some have used a binary outcome of dementia cases and controls, and these measurements of cognition and cognitive decline have all been associated with tooth loss.

Tooth loss is not a simple harbinger of cognitive decline, as experimental studies in animal models have demonstrated that the direct loss of teeth leads to cognitive impairment. Rats and mice with bilateral extraction of the maxillary molar teeth have more errors in completing a maze task (average of 3 errors in mice with all molars extracted vs. 0 errors in the control group) [41] and perform worse on passive-avoidance experiments (6 mice with teeth extracted showed impairment vs. 0 mice in the control group showed impairment) [42]. These studies further showed that both experimental and control animals did not show differences in markers of psychological stress [42], eating habits, and motivation several weeks after extraction [41], indicating that the pain and stress of tooth extraction did not confound the results. Upon histological examination, studies also observed that tooth loss affects the structure and function of neurons [43] and results in neuronal cell loss [42]. Additionally, a decrease in levels of tyrosine kinase receptor B mRNA, which enhances nerve transmission and is a marker for elevated synaptic transmission levels, is decreased in the brain tissues of these rats [41]. Together with the data from human epidemiological studies, there is a clear association and plausible causal relationship between tooth loss and cognitive function.

The exact mechanism by which tooth loss leads to cognitive decline is not yet clear, though several theories have been proposed. Inflammation is thought to be involved in Alzheimer's Disease (AD) pathogenesis, and inflammatory molecules such as proinflammatory cytokines may promote the development of AD and cognitive decline [44]. The peripheral inflammation seen in periodontitis may cause an increase in the systemic load of proinflammatory molecules, which may migrate to the central nervous system (CNS), cause inflammation in the brain, and, subsequently, cognitive decline. Additionally, gram-negative bacteria, such as *A. actinomycetemcomitans,* that cause periodontitis may invade tissues other than the gingiva and cause damage, so pathogen damage to brain tissue may contribute to the cognitive decline seen in tooth loss. Hence, the association between tooth loss and cognitive decline may reflect the causal role of periodontitis in both conditions. Indeed, gingival bleeding and attachment loss, both symptoms of periodontitis, are associated with cognitive outcomes independent of tooth loss [37]. Alternatively, because inflammation, particularly inflammatory interleukin molecules, plays a role in the pathogenesis of both cognitive impairment and periodontal disease, this association may reflect a shared genetic predisposition to inflammation-related diseases [45]. Independent of periodontal disease, chewing results in a considerable amount of sensory input to the CNS because periodontal ligament cells are richly innervated, and chewing also increases blood flow to the brain, so tooth loss may decrease CNS activity to some degree via impaired masticatory function [46]. Lastly, low levels of B-vitamins predict cognitive decline [47]; since tooth loss results in poor diet quality [9], deficiencies in essential nutrients in these individuals may be driving the relationship. Altogether, none of these mechanisms are mutually exclusive and several may be acting in concert, contributing to cognitive decline in individuals experiencing tooth loss. Whatever the causal mechanism may be, this robust

association reinforces the WHO goal to retain a functional dentition of 20 or more teeth through life. Only by identifying and characterizing the causes of missing teeth, especially the heretofore understudied genetic component, can this goal be met.

1.3 THE CURRENT STATE OF THE GENETICS OF TOOTH LOSS

1.3.1 Biologic Causes of Tooth Loss

Tooth loss, as assessed by total missing teeth, is a combined phenotype representing teeth missing due to dental caries, periodontal disease, tooth agenesis, and/or dental avulsion [22, 24]. Dental caries occurs when acid-producing bacteria of the oral microbiome cause demineralization of tooth enamel and subsequent tooth decay. Left untreated, carious lesions can result in a tooth falling out or being extracted and thus, missing teeth. Periodontal disease is caused by a dysbiotic shift of the oral microbiome and aberrant growth of pathogenic bacteria, as well as an excessively aggressive immune response to these bacteria. This process leads to inflammation of gum tissues, loss of alveolar bone connecting teeth to the gums, loosening of teeth, and finally tooth loss. Lastly, tooth agenesis is the congenital absence of a tooth resulting from developmental failure, and dental avulsion is the trauma-induced loss of a tooth. The main dentist-reported reasons for extraction are periodontitis and dental caries; whereas caries is more important in younger patients, periodontitis predominates after 50 years of age [29]. To a lesser extent, orthodontics and trauma that did not immediately cause tooth loss are also reasons for extraction, as well as patient preference over other treatment options. It is important to

acknowledge the causes of missing teeth since genetic factors influencing these biological reasons for tooth loss are also likely to contribute to the genetic basis of missing teeth.

1.3.2 Gene Mapping and Heritability Studies for Dental Caries and Periodontitis

At the genome-wide scale, the genetics of dental caries and periodontitis have been studied. As dental caries and periodontitis are the two leading causes of tooth loss, whole-genome gene mapping efforts for these diseases should be informative for tooth loss. Variants associated with dental caries and periodontitis may also be associated with missing teeth. Studies demonstrating the heritability of dental caries and periodontitis are summarized in Tables 1 and 2, and whole-genome gene mapping studies (i.e., genome- wide association studies [GWAS]) of these diseases are summarized in Tables 3 and 4.

Table 1. Heritability studies for dental caries

Author	Year	Sample	N	Phenotype	Heritability Estimate	Reference
Boraas et al.	1988	Twins	44 twin pairs and 3 triplets, reared apart	Dental caries	90%	[48]
Bretz et al.	2004	Brazilian twins	388 twin pairs	SBCPR*	76.3%	[49]
				LSI†	70.6%	
Bretz et al.	2005	Brazilian twins	314 twin pairs	SBCPR	30%	[50]
				LSI	63%	
Bretz et al.	2006	Brazilian twins	115 pairs	SBCPR	65%	[51]
				LSI	62%	
Wang et al.	2010	Whites	2,600	primary tooth caries	54-70%	[52]
				permanent tooth caries	35-55%	
				total caries	29-40%	
Shaffer et al.	2012	Whites	2,600	Smooth surface caries	17-42%	[53]
				Pit and fissure caries	19-53%	

Abbreviations: *lesion severity index (LSI); †surface-based caries prevalence rate (SBCPR)

Table 2. Heritability studies for periodontitis

Author	Year	Sample	N	Phenotype	Heritability Estimate	Reference
Michalowicz et al.	1991	American twins	110 twin pairs	Periodontal disease traits	38% - 82%	[54]
Corey et al	1993	American twins	349 twin pairs	Periodontal disease	MZ pairwise concordance = 0.23 DZ pairwise concordance = 0.08	[55]
Michalowicz et al.	2000	American twins	117 twin pairs	Periodontal disease traits	43-50%	[56]
Mucci et al.	2005	Swedish twins	10,578 twin pairs	Periodontal disease	39% in women 33% in men	[57]
Diehl et al.	2005	White and black Americans	610 related individuals	Periodontal disease traits	14%-30%	[58]

Table 3. Whole-genome gene mapping studies for dental caries

Author	Year	Sample	N	Phenotype	Gene(s) Nominated	Reference
Shaffer et al.	2011	American whites (children)	1,305	Primary dental caries affection status	*ACTN2, MTR, EDARADD, MPPED2, LPO*	[59]
Wang et al.	2012	American whites	7,443	Permanent DMFS*	*RPS6KA2, PTK2B, RHOU, FZD1, ADMTS3, ISL1, TLR2*	[60]
Shaffer et al.	2013	American whites	920	Permanent clustered caries partial DMFS	***LYZL2, AJAP1***, *ABCG2, PKD2*, the dentin/bone *SCPP* sub-family, *TWSG1, EDNRA, NKX2-3, IFT88, TJFBR1, SMAD7, IL17D*	[61]
Zeng et al.	2013	American whites	1,017	Permanent DMFS stratified by PF† and SM‡ surfaces	*BCOR, INHBA, BCORL1, CXCR1/2*	[62]
Zeng et al.	2014	American whites (children)	1,006	Primary DMFS stratified by PF and SM surfaces	***KPNA4***, *ITGAL, PLUNC*	[63]
Morrison et al.	2016	American Hispanics	11,754	DMFT§ and DMFS	***NAMPT, BMP7***, ***IGSF10***, ***MIR5186***, ***MIR548H2***, ***AADACL2***, ***ANK3***, ***CACNA1G***	[64]

Bolded text denotes genes implicated at genome-wide significance; Abbreviations: *decayed, missing, filled surfaces index (DMFS); † pit-and-fissure surfaces (PF); ‡smooth surfaces (SM); §decayed, missing, filled teeth index (DMFT)

Table 4. Whole-genome gene mapping studies for periodontitis

Author	Year	Sample	N	Phenotype	Gene(s) Nominated	Reference
Schaefer et al.	2010	European	1,758	AgP* case/control	***GLT6D1***	[65]
Divaris et al.	2012	European	1,020	Periodontal pathogen colonization	*KCNK1*, *UHRF2*, *FBXO38*, *IL33*, *TRPS1*, *RUNX2*, *CAMTA1*, and *VAMP3*	[66]
Teumer et al.	2013	European	4,032	Periodontal disease traits	Nothing genome-wide significant	[67]
Divaris et al.	2014	European	4,504	CP† cases/controls	*NPY*, *NCR2*, *EMR1*, ***NUAK1*** (smoking interaction), *NIN*, *WNT5A*, 10p15	[68]
Feng et al.	2014	All ethnicities and whites only	866	CP case/control	16q22.3	[69]
Freitag-Wolf et al.	2014	European, stratified by sex	2,183	AgP case/control	*NPY* (increases risk in males only)	[70]
Shaffer et al.	2014	European adults <50 yrs	673	Periodontal pocket probing depth	*LAMA2*, *HAS2*, *CDH2*, *ESR1*, *SOS2* and *NIN*, *OSBPL10*, *HSP90AB2P*, *GVINP1*, *SEL1L*, *FHOD3*	[71]
Shimizu et al.	2015	Japanese	17,918	Periodontitis cases/controls	*KCNQ5*, *GPR141-NME8* (smoking interaction)	[72]
Offenbacher et al.	2016	European	978	Novel derived periodontal traits	***CLEC19A, RBMS3, GGTA2P, TM9SF2, IFI16, TRA, HPVC1, SLC15A4, PKP2, SNRPN***	[73]
Munz et al.	2017	European	7,980	AgP cases/controls	***SIGLEC5, DEFA1A3***, *FCER1G* and *SLC1A3*	[74]
Sanders et al.	2017	Hispanic	10,935	Interproximal clinical attachment level	***TSNAX-DISC1***	[75]

Abbreviations: *aggressive periodontitis (AgP); † chronic periodontitis (CP)

From heritability studies, genetic factors are clearly contributing to the etiology of periodontal disease and dental caries; thus, genetic factors are also contributing to tooth loss and missing teeth. At the same time, these estimates show a wide range, with heritability of dental caries reported to be as high as 90% and as low as 17%. This may be due to differences in the

measurement of dental caries (dental caries cases/controls vs. LSI/SBCPR), calculation of heritability (twin pairs vs. variance component modeling), or study population (Brazilians vs. U.S. whites). The same is true for heritability estimates for periodontitis. Accordingly, GWAS have successfully identified several loci and genes putatively associated with dental caries and periodontitis, and several of these loci have been replicated. For dental caries, *MPPED2* has been associated with different dental caries phenotypes within the same sample [59, 63], and associations for *AJAP1* and *RPS6KA2* have been replicated across GWAS for different dental caries phenotypes in non-overlapping study samples of the primary and permanent dentitions [61, 63]. For periodontal disease, replication of association signals has been problematic. With the exception of *NPY*, associations seen for aggressive periodontitis (AgP) have not generalized to chronic periodontitis (CP), indicating that genetic mechanisms influencing CP are generally not the same as those that influence aggressive periodontitis. This result is consistent with the consensus view that the aggressive and chronic forms of periodontal disease are biologically distinct, which is similar to other complex diseases where the early onset, familial forms are genetically distinct from sporadic, late-onset forms, such as Alzheimer's Disease and breast cancer. Additionally, difficulties in replicating association signals across study populations may be rooted in the differential influence of various environmental factors that are more or less important in different populations [59]. Differences in dental care access, disease prevalence, SES, and drinking water fluoridation, for example, may be driving unmeasured gene-by-environment interactions for both dental caries and periodontitis that are population-specific and thus unable to be replicated across populations. Altogether, GWAS for dental caries and periodontal disease have nominated many genes of potential interest, but there has been limited success in replicating GWAS signals across study populations.

1.3.3 Heritability and Candidate Gene Studies for Missing Teeth

While the heritability of dental caries and periodontal disease has been established, few studies have investigated the heritability of tooth loss. Using the Swedish Twin Registry, Mucci *et al*. (2005) estimated edentulism to be 39% heritable in men and 14% heritable in women, though the 95% confidence intervals of these estimates overlap [57]. Similarly, Kurushima *et al*. (2017) used the Danish Twin Registry to estimate that additive genetic factors explained 36% of the variance in number of teeth (none, 1-9, 10-19, 20-27, all), indicating that tooth loss is moderately heritable. When stratified by sex, tooth loss was 47% and 22% heritable in women and men, respectively, though again, the 95% confidence intervals overlap [76]; interestingly, this result is opposite of that of Mucci *et al.* (2005), where heritability was higher in men than women. Additionally, Kurushima *et al*. estimated heritability of having 20 or more teeth to be 37% and edentulism to be 38%. Heritability estimates from these two studies are harmonious, though that is unsurprising given that both study samples come from Northern European countries with similar environmental factors. To more fully assess the heritability of tooth loss, studies should be undertaken in populations which experience different environmental factors contributing to tooth loss, such as U.S. populations, which differ from European nations in factors like access to dental care. Furthermore, studies should not be limited to whites, since tooth loss is greater in minority populations.

Despite the body of literature supporting the genetic basis of susceptibility to dental caries and periodontitis, the genetic determinants of missing teeth and functional dentition are still emerging areas of research. Various polymorphisms in several genes have already been associated with missing teeth, tooth loss, and edentulism in candidate gene studies. The *OGG1* Ser326Cys polymorphism, a nonsynonymous substitution which results in impaired DNA base

excision repair, has been shown to be associated with tooth loss in the elderly even after adjusting for age, sex, BMI, and heart disease [17]. A repeat polymorphism in the *MPG* gene, which is involved in bone and tooth formation, is associated with retention of a higher number of teeth [77]. In a study of the effect of the *APOE* alleles and edentulism, edentate individuals had a higher frequency of the *APOE* ε4 allele than their dentate counterparts, even when matched on other contributing factors such as age and education, though it is unclear whether this association is attributable to dementia associated with the *APOE* ε4 allele and a subsequent neglect of oral hygiene, or other biologic mechanisms [15]. Variation within the vitamin D receptor gene, *VDR*, is associated with periodontal disease progression and subsequent tooth loss [78], and estrogen-receptor gene polymorphisms are associated with tooth loss in post-menopausal women [79, 80]. While these candidate genes have been reported and studied for over a decade, to date, no published studies have applied unbiased, genome-wide methods to investigate the genetic susceptibility of tooth loss.

A common theme among complex diseases is that genes which cause Mendelian syndromes when disrupted often have regulatory variants that influence common variation in the same traits affected by the syndrome. Therefore, considering syndromes with a missing teeth component may point to more candidates for common variation in missing teeth. Papillon-Lefèvre syndrome, an autosomal recessive Mendelian disorder that results in severe periodontal disease, has been mapped to 11q14–q21 by linkage analysis methods [81]. Altogether, there are several Mendelian diseases that result in periodontal phenotypes relevant to tooth loss, such as inherited neutropenias, which result in a decrease in white blood cells, predisposing to bacterial infections [82]. Amelogenesis imperfecta is a Mendelian disease caused by mutations in one of several genes critical for enamel formation, including *AMELX*, *ENAM*, *KLK4*, and *MMP40*, and

this condition results in excess wear and breakage of teeth [83]. The variety of Mendelian diseases with relevant periodontal manifestations and dental anomalies indicates that genetic factors, including coding variation, may result in generalized or nonsyndromic periodontal disease and tooth loss.

Aside from periodontal disease and dental caries, teeth may also be missing due to tooth agenesis and trauma, and the genetic basis of tooth agenesis is relatively well understood compared to dental caries and periodontitis. Sequencing studies of *WNT10A* have shown that nonsynonymous, nonsense, and missense mutations in *WNT10A* are strongly associated with both agenesis of 1-3 teeth and agenesis of 4 or more teeth, a much rarer outcome [84]. Other genes implicated in the pathogenesis of tooth agenesis include *MSX1*, *PAX9*, *AXIN2*, and *EDA* [85]. While there is a clear genetic component to tooth agenesis, teeth missing due to agenesis only make up a small percentage of total missing teeth; periodontitis and dental caries are the predominant causes of tooth loss. Thus, gene mapping efforts to identify genetic contributors to total missing teeth are unlikely to yield variants implicated in tooth agenesis.

1.4 GENOME-WIDE ASSOCIATION STUDY OF MISSING TEETH AND FUNCTIONAL DENTITION: THE RATIONALE

GWAS is a hypothesis-free, agnostic approach to studying complex traits and diseases. The rationale behind GWAS is the common disease-common variant hypothesis, which proposes that genetic susceptibility to common diseases is largely attributable to common variants, most often single nucleotide polymorphisms (SNPs) or small insertion/deletions (indels), that have moderate effects on disease susceptibility [86]. The assumption behind GWAS is that a systematic scan of

the genome for common variants associated with disease will identify those involved in disease pathogenesis, hence they are ideally suited to studying complex traits with a suspected genetic basis that is not yet fully characterized. GWAS is used for both quantitative traits and binary disease case-control studies. Commonly used statistical tests include linear and logistic regression for quantitative and binary phenotypes, respectively, since the regression framework allows for controlling for covariates (usually age and sex). As discussed previously, GWAS for oral health outcomes have successfully identified associations between SNPs in proximity to biologically relevant genes and replicated these associations across study populations (*AJAP1*, *RPS6KA2*).

There are several advantages to using GWAS to identify genetic associations for complex disease. First and foremost, GWAS are easy to conduct. Many large-scale epidemiologic studies include genotyping as part of their protocol, so well characterized, ready-to-analyze samples of thousands of individuals with genetic and phenotypic information are readily available. The cost of genotyping using SNP microarrays is inexpensive compared to sequencing, which makes GWAS a practical alternative. SNPs not directly genotyped may still be included in GWAS, as well – imputation, a probability-based approach to capture information on un-genotyped variants based on observed genotypes, drastically increases the number of SNPs available for association testing from a million or less to over ten million SNPs. Not only does imputation improve coverage of the genome, but after imputation, it is easy to harmonize genotyping results across SNP microarray platforms, so if different studies have used different microarrays, what was genotyped in one sample is likely to be imputed in the other. While the quality of the data from SNP microarrays is not outstanding for all genotyped (and imputed)

SNPs, enough SNPs are genotyped so that thousands of poor performing SNPs may be excluded from analysis while retaining high quality, extensive coverage of the genome.

Additionally, since GWAS mark all regions of the genome, they are unbiased. This feature is a notable advantage over candidate gene studies, which look for associations in a handful of previously identified or biologically relevant genes and thus can miss the strongest associations, as well as generate false positives [87]. Lastly, GWAS is hypothesis generating. No knowledge of specific genes involved in disease is required to perform a GWAS, though it is advisable that some knowledge of the genetic basis of the disease is known beforehand, such as its heritability or segregation pattern in families. Thus, GWAS is a useful first step for identifying genetic associations for diseases that have not been well characterized but are known to be heritable.

Despite the many advantages of GWAS, this approach has several limitations. Most importantly, GWAS only tests for association with disease, not causation. Most of the top variants in GWAS are typically not causal, and are only highly statistically significant because they are in linkage disequilibrium (LD) with a causal variant, which may or may not be genotyped or imputed. This makes it difficult to disentangle an association signal for a causal variant when ten to a hundred SNPs are all associated with disease because of LD. As GWAS test millions of SNPs, multiple test correction is essential, and a conservative genome-wide significance threshold of $p < 5 \times 10^{-8}$ (the Bonferroni correction for a million tests) has become conventional [87]. Also, GWAS must carefully control for population stratification or risk false positive associations, since variants at a higher frequency in populations with higher disease prevalence will appear to be associated with disease. GWAS do not work well for variants with a low minor allele frequency (MAF) or variants with a small effect size, since the power to detect

these variants is too low [88]. Samples sizes need to be prohibitively large in order to achieve 80% power when testing variants with MAF < ~1% in the population, or with low odds ratios (<1.05). Rare SNPs are also difficult to impute [87]. This means that rare variants and/or those with small effect sizes are not likely to be identified by GWAS despite their contribution to genetic susceptibility to disease.

In addition to these statistical issues raised by GWAS, as the vast majority of SNPs are in noncoding regions, characterization and interpretation of noncoding SNPs remains a challenge since a lot less is known about regulatory elements than coding regions. GWAS does not test structural variants, either. Finally, as GWAS is hypothesis generating, replication of discovery results in an independent study population is ideal, but this becomes challenging when comparable populations with similar phenotypic characterization are not available [61, 89].

Since GWAS works well for common disease and is a good starting point for diseases known to be heritable, missing teeth is an appropriate phenotype for GWAS, since little is known about the genetic mechanisms behind tooth loss. Although associations between periodontitis and dental caries should also generalize to missing teeth, studying only these two oral diseases would not identify genetic mechanisms governing retention of teeth in individuals with periodontitis or extensive dental caries. Furthermore, they do not allow for the identification of heretofore unknown genetic mechanisms behind tooth loss such as wound healing or structural fracture. Altogether, GWAS of missing teeth may identify variants not associated with either periodontitis or dental caries but which still contribute to determining loss or retention of teeth.

In the next chapter (Chapter 2), I will describe the methods used in this project, including descriptions of the datasets and the statistical analyses that were used. In Chapter 3, I will describe the results of the genome-wide association scans, including results from the meta-

analysis and those in the replication sample. In Chapter 4, I will discuss the results further by placing them in the context of current work on oral health and genetics as well as suggest future directions.

2.0 METHODS

2.1 SAMPLE DESCRIPTION AND DATA COLLECTION

2.1.1 COHRA and DRDR

The Center for Oral Health Research in Appalachia (COHRA) cohort was ascertained for the purpose of studying oral health disparities in Appalachia and has been described previously [59]. Participants were recruited from rural Appalachia in West Virginia and Western Pennsylvania using a household based recruitment protocol, which required a minimum of one biological parent-child pair. Oral health status was not considered during recruitment, and all household members were invited to participate regardless of biological/legal relationships. Although COHRA is a larger family-based sample, only unrelated, adult participants of COHRA with missing teeth data (N = 1287) were included in this project. All adults gave written informed consent. A total of 732 households and 740 biological families were ascertained. DNA was collected from blood, mouthwash, buccal swab, or saliva sample. Number of missing permanent teeth was determined by intra-oral examination by licensed dentists and/or dental hygienists. All procedures were approved by the COHRA research committee and both the University of Pittsburgh and West Virginia University Institutional Review Boards.

The Dental Registry and DNA Repository (DRDR) cohort was recruited as a research initiative of the University of Pittsburgh School of Dental Medicine. Description of this cohort has been described previously [60]. DRDR is an ongoing project in which individuals who come to the dental school for treatment are given the opportunity to participate in the registry. DNA was collected from blood, mouthwash, buccal swab, or saliva sample. Number of missing permanent teeth was determined by intra-oral examination by licensed dentists and/or dental hygienists. All procedures were approved by the University of Pittsburgh Institutional Review Board. A subset of the DRDR cohort was included along with COHRA in the Gene, Environment Association Studies (GENEVA) consortium. Missing teeth information was available for 303 adult participants of DRDR.

Genotyping for COHRA and DRDR was performed as part of the GENEVA consortium by the Johns Hopkins University Center for Inherited Disease Research (CIDR) through a National Institutes of Health contract [90]. Genotyping was done using the Illumina Human610-Quadv1_B BeadChip (Illumina, San Diego, CA, USA) and the Illumina Infinium II assay protocol [91]. Genotyping data were released for 99.4% of attempted samples as well as 91 duplicates and 188 HapMap genotyping controls (62 CEU, 78 YRI, 24 JPT, 24 CHB). For each SNP, allele cluster definitions were established using Illumina BeadStudio Genotyping Module version 3.3.7 and the combined intensity data from 98.6% of samples; allele cluster definitions were then used for all samples. Genotypes were not called if quality threshold (Gencall score) was less than 0.15. Genotypes were released by CIDR for 589,735 SNPs (99.53% of attempts). SNPs that did not pass the following filters were excluded: SNP call rate less than 85%, more than 1 HapMap replicate error, call rates with a more than 2% (autosomal) or 10% (X) difference between gender, more than 1.8% male AB frequency (X), or more than a 7% (autosomal) or 5%

(XY) difference in AB frequency. Mitochondrial and Y chromosome SNPs were reviewed manually, and clusters were adjusted and genotypes dropped as appropriate. Mean non-Y SNP call rate and mean sample call rate were each 99.8%, and duplicate reproducibility was 99.99%.

Genotype data cleaning and quality assurance was performed by the GENEVA Consortium Coordinating Center at the University of Washington [90, 92]. Missing call rate and allele frequencies were examined for each 96-well plate of samples. While plates differed significantly in missing call rate, no plates were outliers for either missing call rate or allele frequencies. BeadStudio metrics "BAlleleFreq" and "LogRRatio" [92] were used to identify contamination and large chromosomal aberrations. Two contaminated samples were removed. Biological sex was confirmed by comparing X and Y chromosome SNP intensities. Biological relationship checking was done using identity-by-descent coefficient estimates. Fifty-two samples were removed as sample identity could not be resolved.

Imputation was performed by the GENEVA Coordinating Center with the HapMap Phase III reference panel for 1,387,464 autosomal SNPs using BEAGLE software [93]. Sporadic missing data for 503,167 genotyped SNPs was also imputed (SNP passing pre-imputation filtering of 589,735 genotyped SNPs). Masked SNP analysis (imputation of genotyped SNPs) and Mendelian error checking for imputed SNPs among relatives determined imputation quality to be high. A total of 1,450,678 SNPs were available for analysis, representing 884,297 fully imputed autosomal SNPs, 503,167 genotyped SNPs that were partly imputed to fill in sporadic missingness, 50,154 fully genotyped autosomal SNPs, and 13,060 fully genotyped X-linked SNPs.

2.1.2 POFC

The Pittsburgh Orofacial Clefting Study cohorts from Patagonia (POFC-PA), Guatemala (POFC-G), and the United States (POFC-USA) are part of a larger cohort originally ascertained for the purpose of studying orofacial clefting [94, 95]. Participants in POFC included affected cleft cases, unaffected relatives of cases, and controls. Only unaffected adult (>18 years old) participants were included in this GWAS. POFC comprises a total of 11,727 participants ascertained from 18 sites in 13 countries from North America, Central/South America, Asia, Europe, and Africa. Many sites were part of continuing genetic association studies of the University of Pittsburgh Center for Craniofacial and Dental Genetics and the University of Iowa. Number of missing permanent teeth was assessed by intra-oral examination. Informed consent was obtained from all participants, and Institutional Review Board approval was obtained locally and by the University of Pittsburgh and University of Iowa.

Samples were genotyped for approximately 580,000 SNPs using an Illumina HumanCoreExome array by the CIDR at Johns Hopkins University. Data quality assurance and data cleaning were implemented cooperatively with the CIDR Genetics Coordinating Center (GCC) at the University of Washington [96]. In total, 539,473 genotyped SNPs (96.74% of attempts) passed the GCC recommended quality filters. SNPs with missing call rate >2%, 2 or more discordant calls in 264 study duplicates, 20 or more Mendelian errors in 5,288 parent-offspring trios or dyads, Hardy-Weinberg Equilibrium (HWE) p-value < 0.0001 in participants of genetically confirmed European ancestry, sex differences in allele frequency of 0.2 or greater for autosomes or XY pseudoautosomal region, or sex differences in heterozygosity of 0.3 or greater for autosomes or XY pseudoautosomal region were removed. Samples sizes in each POFC

cohort included in this project before excluding those missing covariate and/or genotype information was as follows: 353 for POFC-G, 245 for POFC-PA, and 227 for POFC-USA.

Prior to imputation, pre-phasing of SNPs passing quality control filters was done using SHAPEIT2 [97]. Imputation was performed using IMPUTE2 [98] and phase 3 of the 1000 Genomes Project (comprising 2504 individuals from 26 populations worldwide) as the reference panel. After imputation, genotype data for 34,985,077 SNPs was available, including fully imputed SNPs and SNPs imputed due to sporadic missingness. Accuracy of imputation was determined by masked variant analysis, which revealed high-quality imputation; mean concordance was 0.960 for SNPs with MAF $\geq$ 0.05 and 0.995 for SNPs with MAF < 0.05. Only SNPs whose "most-likely" genotype probability was > 0.5 were included in statistical analysis. Imputed SNPs out of HWE in European controls were excluded from subsequent analyses.

2.2 GENOME-WIDE ASSOCIATION ANALYSES

2.2.1 COHRA

To describe the distribution of missing teeth and functional dentition and determine which variables to control for in the GWAS, the demographic predictors of missing teeth and functional dentition were explored using the R statistical analysis environment (R Foundation for Statistical Computing, Vienna, Austria). The quantitative trait was defined as a natural log transformation of the total number of missing teeth (ln[Total Missing + 1]). The transformation was used to help normalize the distribution of the data and stabilize the variance. This trait encompasses permanent teeth missing due to decay, missing not due to decay (trauma, orthodontics), and

permanent unerupted teeth. Third molars (wisdom teeth) were not included. Functional dentition was defined as the presence of at least 20 teeth (*i.e.*, missing 8 or fewer teeth.)

Linear regression was used to determine the association between both missing teeth traits and age, participant recruitment site, income, education, race, and ethnicity. Education was defined as 1) none or high school diploma/GED, 2) technical school, associate degree, or some college with no degree, or 3) undergraduate degree or higher. The Pennsylvania participant recruitment sites of Bradford, Burgettstown, and Braddock were combined into one variable denoting Pennsylvania or West Virginia residency. Principal components (PCs) of genetic ancestry were used to control for ancestry and were calculated using PLINK and R software. To further control for population stratification, GWAS was limited to self-identified, non-Hispanic whites; 147 individuals were excluded based these criteria. GWAS analysis was also limited to adults over age 18 and under age 60, as there were too few individuals over age 60 (N = 12) to accurately model an age effect.

Individuals meeting inclusion criteria and for whom genetic data was available were included in the genome-wide association analyses using PLINK software [99]. Linear and logistic regression (--assoc-linear and --assoc-logistic) were used for the quantitative trait and functional dentition, respectively, while including age, sex, site, and the first PC of ancestry as covariates. Education and income, while strongly associated with missing teeth, were not controlled for as 1) this would exclude discovering SNPs possibly influencing both education and missing teeth, and 2) covariate information was not available for all participants. Manhattan and QQ plots were generated using R software. Genomic inflation factor, λ, was calculated for all GWAS as implemented in the GenAbel package [100] for R.

2.2.2 DRDR

Information on covariates such as income and education were unavailable to test for associations between these predictors and the same missing teeth traits as in COHRA. To control for population stratification, principal component analysis was performed on SNPs in low LD using PLINK. Only individuals with self-reported European ancestry and who were confirmed to be of European ancestry by PCA were included in the GWA analyses. That is, individuals who self-reported a race other than white (African-American, Asian, other) or who were outside 2 standard deviations of PC1 or PC2 on the European ancestry region of the PC plots were excluded. PCA was then rerun in this whites-only subset of DRDR

As in COHRA, individuals meeting inclusion criteria and for whom genetic data was available were included in the genome-wide association analyses using PLINK software as previously described. Age, sex, and the first PC of ancestry were included as covariates. Manhattan and QQ plots and genomic inflation factor, λ, were generated as previously described.

2.2.3 POFC-G and POFC-PA

GWA scans for the quantitative (ln[Total Missing + 1]) and binary (missing teeth ≥ 9) traits were performed using variance-component modeling as implemented by EMMAX [101] while adjusting for age, sex, and genetic sharing due to common ancestry and familial relatedness as estimated via identify-by-descent (IBD). Manhattan and QQ plots and genomic inflation factor, λ, were generated as previously described.

2.2.4 POFC-USA

Demographic predictors (recruitment site, sex, age, education, race, ethnicity) were examined for any association with missing teeth and functional dentition. This dataset contained a handful of pairs of related individuals. To control for spurious findings due to genetic relatedness, one individual from a related pair was randomly removed, unless one of the pair did not have a functional dentition, in which case that individual was retained and the other removed. To control for population stratification, principal component analysis was performed on SNPs in low LD using PLINK. As POFC-USA included individuals of both European and Hispanic ancestry, the first 3 PC's were included as covariates in the GWA scan. Both ancestry groups were included in this GWA scan to avoid reducing sample size.

As in COHRA and DRDR, 192 individuals meeting inclusion criteria and for whom genetic data were available were included in the genome-wide association analyses using PLINK software as previously described. Age, sex, and the first three PCs of ancestry were included as covariates. Manhattan and QQ plots and genomic inflation factor, λ, were generated as previously described.

2.3 META-ANALYSIS

Meta-analysis was performed using Stouffer's p-value-based meta-analysis as implemented in METAL software [102]. Six meta-analyses were performed: a whites-only meta-analysis combining COHRA and DRDR for the quantitative trait and functional dentition, respectively, a Hispanics-only meta-analysis combining POFC-G and POFC-PA for the quantitative trait and

functional dentition, respectively, and a trans-ethnic meta-analysis combining COHRA, DRDR, POFC-G, and POFC-PA for the quantitative trait and functional dentition, respectively. POFC-USA was not included in meta-analyses because both ethnicity-stratified GWA scans and GWA scans including both Hispanics and non-Hispanic whites while adjusting for 3 PCs showed over- or under-inflation, thus we were not confident with these results at the genome-wide level. POFC-USA was reserved for replication of specific SNPs only. Results of the trans-ethnic analysis were filtered to include only SNPs that were tested in all four studies. For each meta-analysis, Manhattan and QQ plots and genomic inflation factor, λ, were generated as previously described.

2.4 BIOINFORMATICS

To investigate the functionality of genomic regions showing strong associations with missing teeth traits, regions showing genome-wide significance, that is, SNPs with $p < 5 \times 10^{-8}$, and regions showing suggestive significance ($p < 5 \times 10^{-6}$) were visualized using Regional Association (LocusZoom) plots [103]. The function of nearby genes was investigated using the resources at the National Center for Biotechnology Information (NCBI), specifically the Gene and PubMed databases. PubMed search terms included the name of the gene or corresponding protein, plus terms relevant to dental and oral health, including tooth, dental, oral, periodont*, and inflamm*. Information on gene regulation and LD was obtained from HaploReg [104], and information on functional variants (missense mutations, loss of function mutations, etc.) was obtained from HaploReg and the Exome Aggregation Consortium (ExAC) database [105].

3.0 RESULTS

The analysis plan for all five cohorts is depicted in Figure 1. GWA scans were performed for both the quantitative trait and functional dentition in COHRA, DRDR, POFC-G, POFC-PA, and POFC-USA. Extensive annotation was carried out for GWA scans in COHRA as well as all meta-analyses. Some annotation of suggestively significant results in POFC-G was also performed (checking for nearby genes that were obviously relevant to dental/oral health.) Due to low sample size in DRDR, POFC-PA, and POFC-USA, these GWA scans were not annotated. Results in COHRA are described first, then POFC-G, followed by the white-only meta-analysis (COHRA and DRDR), the Hispanics-only meta-analysis (POFC-G and POFC-PA), and lastly the trans-ethnic meta-analysis (COHRA, DRDR, POFC-G, and POFC-PA). POFC-USA was only used for checking associations of the most significant SNPs in the whites-only and trans-ethnic meta-analyses, and all results regarding POFC-USA are included in the appendix. Descriptions of each GWA scans include results of preliminary data analysis, association testing results, and functional annotation as appropriate. Results of meta-analyses include results of the p-value based meta-analysis and functional annotation of the top hits. Findings in discovery GWA scans and the meta-analyses that were not obviously relevant to dental and oral health are described in the appendix.

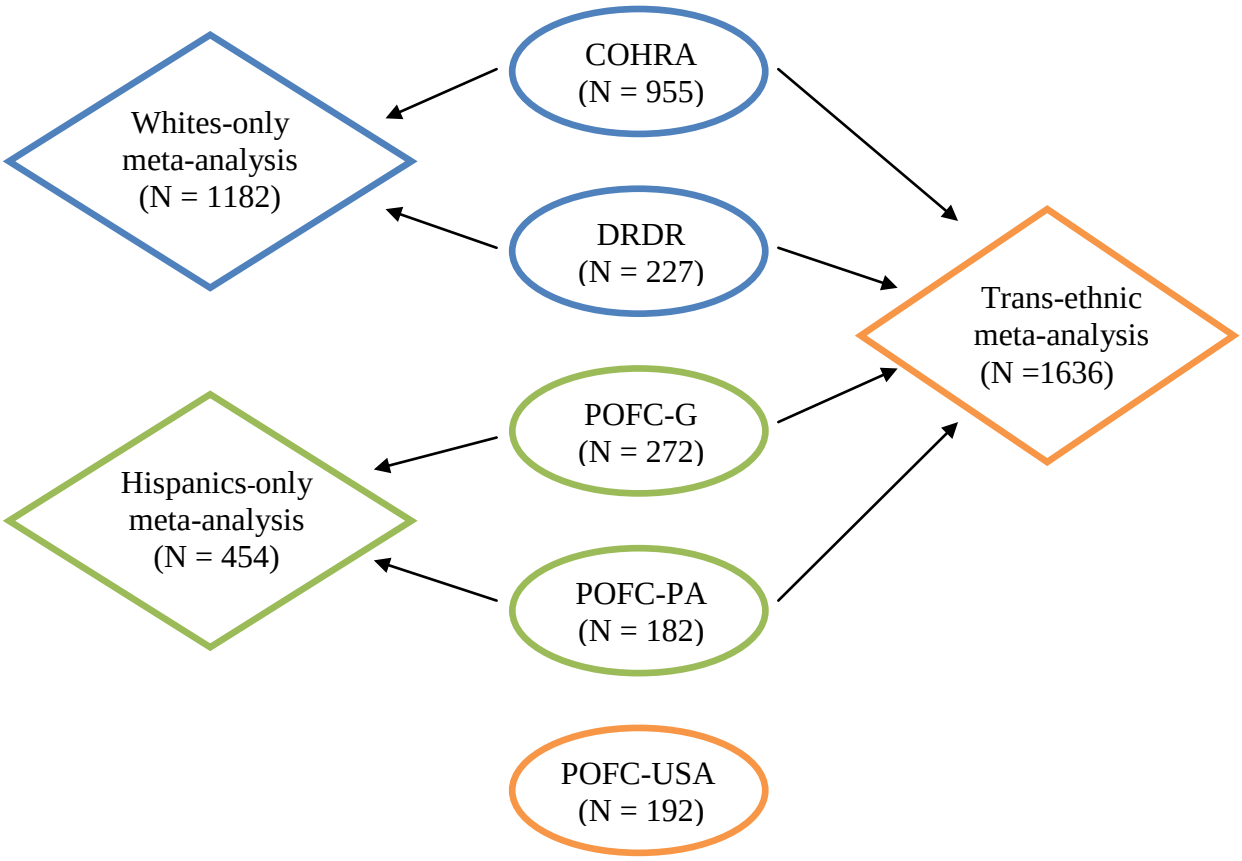


Figure 1. Analysis plan for all five cohorts

Blue coloring signifies whites-only samples, green Hispanics-only samples, and orange mixed samples. Ovals represent cohorts with GWA scans, and diamonds the different combinations of cohorts into meta-analyses. COHRA and DRDR make up the whites-only meta-analysis, and POFC-G and POFC-PA the Hispanics-only meta-analysis. All four of these cohorts were meta-analyzed into a trans-ethnic meta-analysis. POFC-USA was only used for spot-checking significant associations seen in single GWAS scans and meta-analyses.

3.1 COHRA

3.1.1 Trait Development and Covariate Modeling

Distribution of missing teeth was examined in order to define the traits used in the GWAS. Note: the sample size in these results is greater than that of the GWAS, as some individuals did not have genotype information available but were included in initial trait development and

covariate modeling. Figure 2 shows the distribution of the quantitative trait, both before performing a transformation (ln[Total Missing + 1]) and after. Although the distribution is still not normally distributed after the transformation, the transformation helps to stabilize the variance. 37.7% of individuals were not missing any teeth, and 62.2% were missing at least one tooth. Figure 3 shows the distribution of functional dentition (functional dentition status.) Few individuals (11.1%) did not have a functional dentition. Distribution of both traits did not change after removal of individuals not included in the GWAS due to race, ethnicity, or missing genotype information.

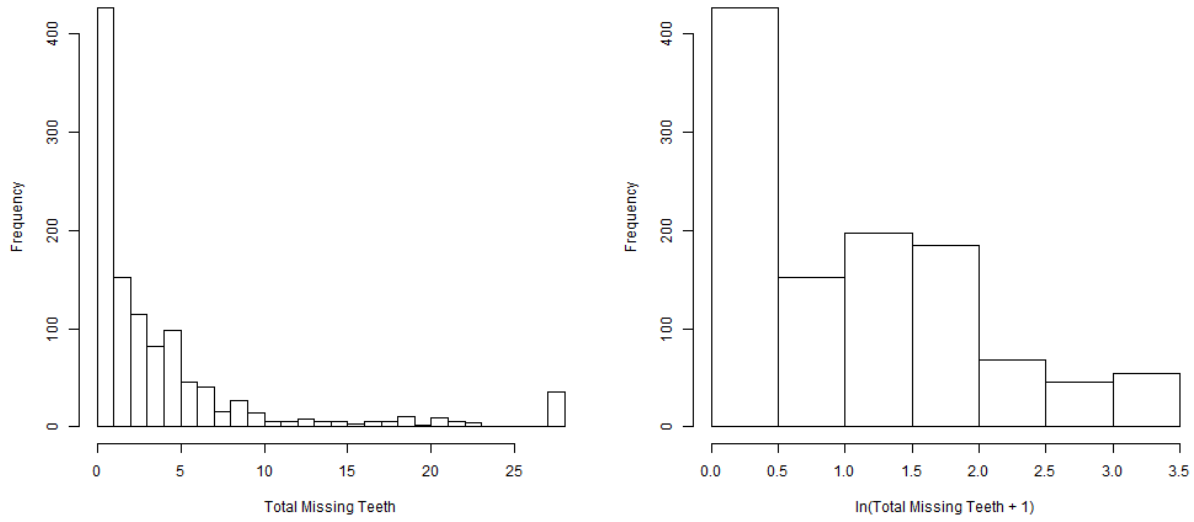


Figure 2. Distribution of the quantitative trait

Distributions of the quantitative trait before (left) and after (right) the transformation

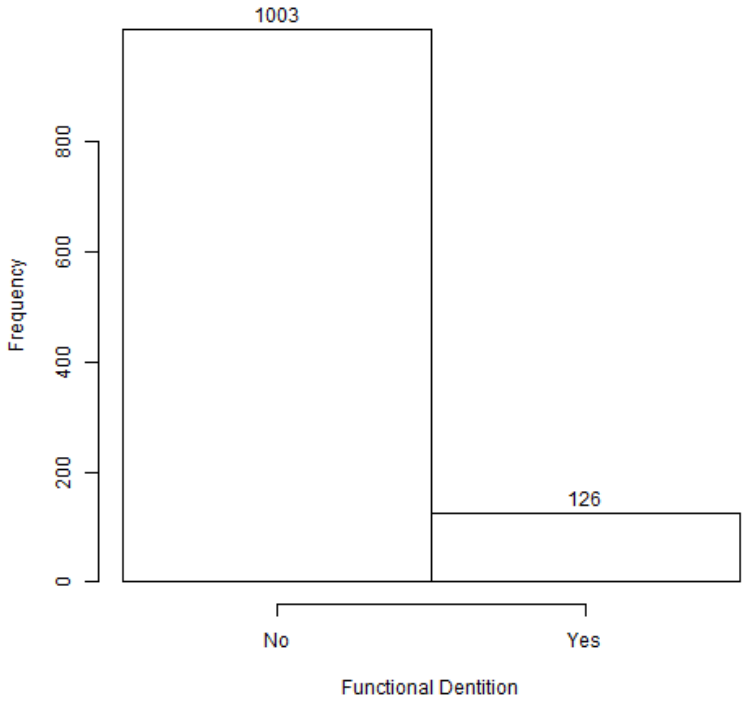


Figure 3. Distribution of functional dentition

For the dataset used in the GWAS, sample characteristics and trait distributions are shown in Table 5, and p-values for the test of association between demographic variables and both missing teeth traits are shown in Table 6. Age, income, education, and site were all associated with missing teeth and functional dentition ($p < 0.005$.) Sex was not associated with either phenotype ($p = 0.43$-0.78.) For functional dentition, 126 individuals (11.2%) of COHRA participants did not have a functional dentition. Mean number of teeth missing was 3.70 (sd = 6.13) before performing the transformation. Figure 4 shows the first two PCs of ancestry. Only the first PC was used as it was sufficient to control for ancestry.

Table 5. Study characteristics and trait distribution in COHRA

Variable	**Value**	**Count (%) or Mean (SD)**
Functional Dentition	Yes	1003 (88.8%)
	No	126 (11.2%)
Total Missing		3.70 (6.13)
Age		34.38 (8.90)
Sex	Male	421 (37.3%)
	Female	708 (62.7%)
Income	Less than 10,000	316 (36.5%)
	10,000 to 14,999	124 (14.3%)
	15,000 to 24,999	140 (16.2%)
	25,000 to 34,999	104 (12%)
	35,000 to 49,999	94 (10.9%)
	50,000 to 74,999	57 (6.6%)
	75,000 to 99,999	21 (2.4%)
	100,000 to 149,999	5 (0.6%)
	150,000 to 199,999	2 (0.2%)
	200,000 or more	2 (0.2%)
Education	High School/GED or none	649 (58.9%)
	Tech School, associate degree, some college	286 (26%)
	Undergrad or higher	167 (15.2%)
Site	PA	313 (27.7%)
	WV	816 (72.3%)

Categorical covariates show counts and percents, and continuous variables show means and standard deviations.

Table 6. Association of Covariates with Missing Teeth Traits in COHRA

Variable	Quantitative Trait		Functional Dentition	
	P	**Beta (SE)**	**P**	**Beta (SE)**
Age	8.61E-27	0.034 (0.003)	1.90E-16	0.09 (0.01)
Sex	0.78	0.02 (0.06)	0.43	-0.15 (0.19)
Income	6.90E-6	-0.08 (0.02)	8.19E-3	-0.17 (0.06)
Education	3.82E-12	-0.27 (0.04)	1.09E-7	-1.05 (0.20)
Site	2.22E-5	0.28 (0.06)	3.85E-3	0.72 (0.25)

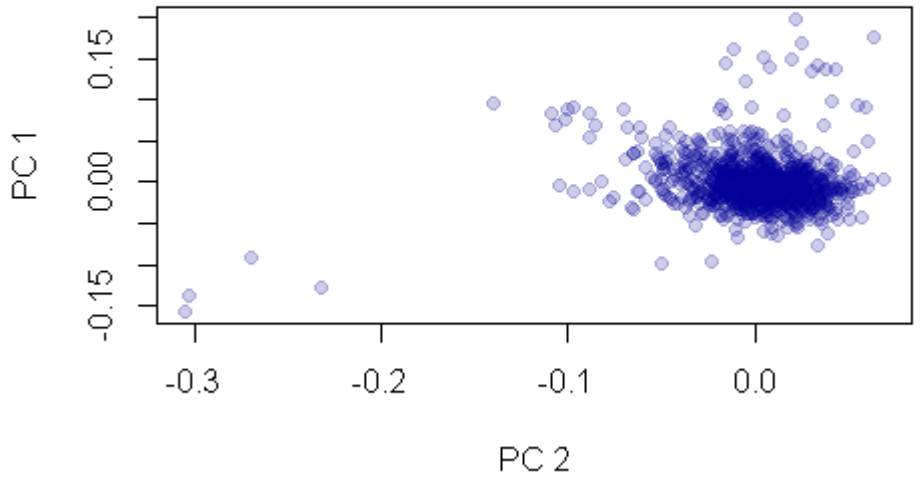


Figure 4. PC1 and PC2 in COHRA

3.1.2 GWAS Results – Quantitative Trait

GWAS was performed using 955 individuals for whom genotype and covariate information was available. Manhattan and QQ-plots for the quantitative trait are shown in Figure 5. Genomic inflation factor was $\lambda = 1.05$ for the quantitative trait, indicating slight inflation. No SNPs reached genome-wide significance ($p < 5 \times 10^{-8}$); however, there were several regions of the genome that showed suggestive significance ($p < 10^{-5}$). Association results for index SNPs are shown in Table 6.

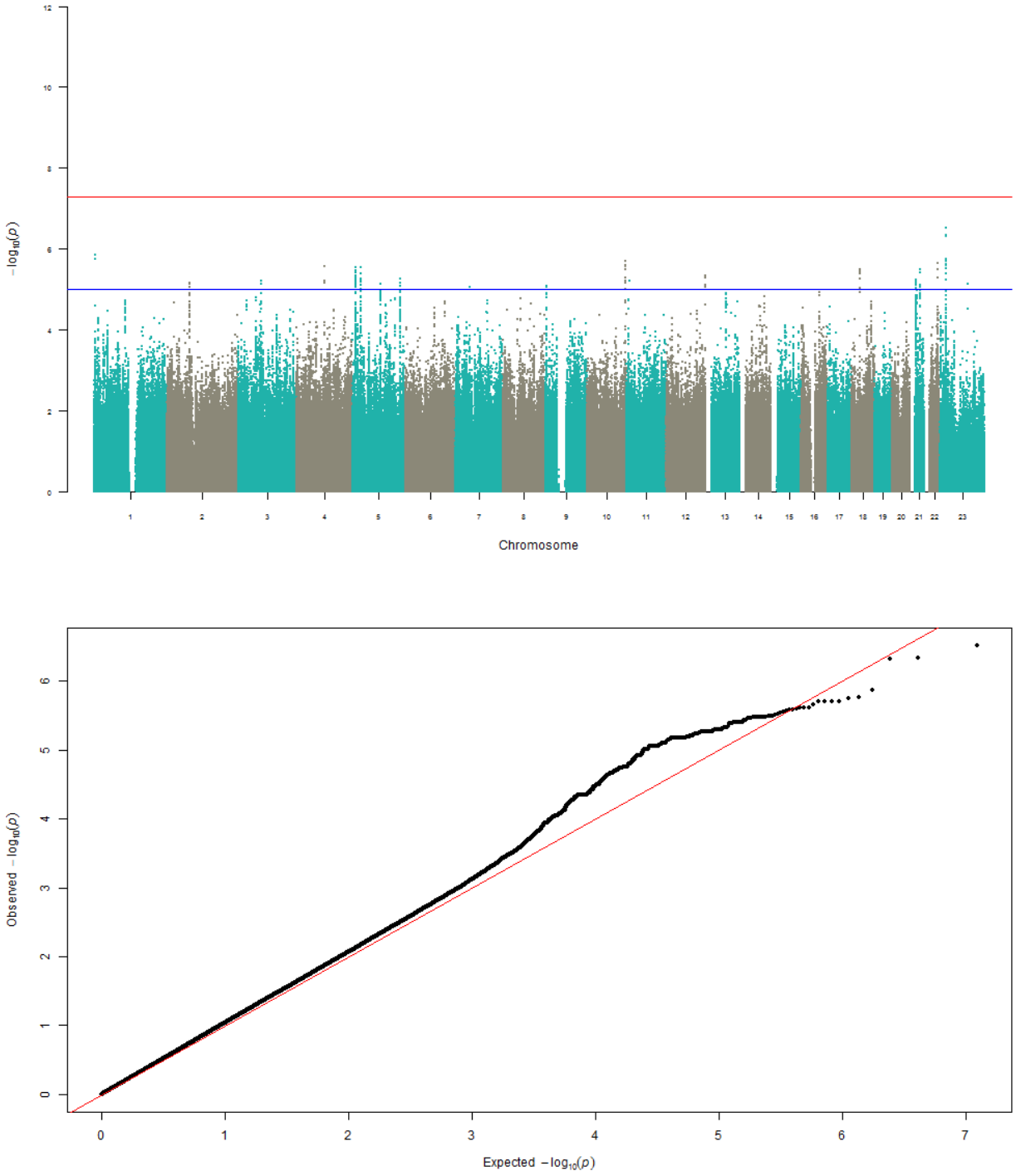


Figure 5. Manhattan and QQ plots for the quantitative trait in COHRA

Table 7. Top GWAS hits for the quantitative trait in COHRA

SNP	CHR	BP	Effect Allele	MAF	N	P	Beta (95% CI)	Type
rs11581023	1	5036549	A	0.14	908	1.37E-06	-0.30 (-0.42, -0.18)	imputed
rs28651854	2	76800193	A	0.14	935	6.84E-06	-0.29 (-0.41, -0.16)	imputed
rs73104494	3	77020811	G	0.23	940	5.98E-06	0.23 (0.13, 0.33)	imputed
rs67374207	4	96181016	G	0.05	950	2.63E-06	0.46 (0.27, 0.66)	imputed
rs2964805	5	10831221	T	0.29	932	2.86E-06	-0.22 (-0.32, -0.13)	imputed
rs77475896	5	29644220	G	0.05	955	2.80E-06	0.43 (0.25, 0.61)	imputed
rs9918187	5	96531684	G	0.35	923	7.26E-06	0.20 (0.11, 0.28)	imputed
rs2860805	5	163265408	C	0.24	953	5.60E-06	-0.22 (-0.32, -0.13)	imputed
rs76362984	7	48491060	A	0.04	864	8.63E-06	0.50 (0.28, 0.72)	imputed
rs10812719	9	2809260	C	0.49	949	8.48E-06	-0.19 (-0.28, -0.11)	imputed
rs75195099	10	129944252	C	0.04	919	2.46E-06	0.54 (0.32, 0.76)	imputed
rs75040946	10	131470783	C	0.10	922	2.01E-06	-0.34 (-0.48, -0.2)	imputed
rs3365	11	8704711	C	0.14	955	6.00E-06	0.29 (0.16, 0.41)	genotyped
rs12297548	12	131431379	T	0.07	954	4.68E-06	0.36 (0.21, 0.51)	genotyped
rs4444411	18	27622307	G	0.09	954	3.19E-06	0.34 (0.2, 0.48)	imputed
rs2825184	21	20202646	A	0.43	876	5.69E-06	0.20 (0.11, 0.28)	imputed
rs73199539	21	32938208	T	0.06	940	3.18E-06	-0.42 (-0.6, -0.25)	imputed
rs4823141	22	44188796	G	0.40	954	2.22E-06	0.21 (0.12, 0.29)	imputed
rs7061889	X	22719993	C	0.15	951	3.03E-07	0.26 (0.16, 0.36)	imputed
rs5966776	X	97002058	T	0.14	892	7.43E-06	0.24 (0.14, 0.35)	imputed

Regions of the genome showing the suggestive or genome wide significance with total missing teeth were further visualized using Regional Association (LocusZoom) plots in order to assess whether the region may have a possible role in dental and oral health.

rs7061889, the top SNP ($p = 3.0E\text{-}7$, $\beta = 0.26$ [0.16, 0.36]), located on the X chromosome, lies within *PTCHD1* antisense RNA (Figure 6). *PTCHD1*, patched domain containing 1, encodes a membrane protein with a patched domain; genomic deletions and loss of function mutations of *PTCHD1* are associated with intellectual disability and autism spectrum disorder, as well as dysmorphic facial features [106, 107], indicating a role of this protein in neurological function and possibly craniofacial development. Functional analyses show that *PTCHD1* protein localizes to the cell membrane, is expressed in the brain [106], and plays a role in the developing mouse brain [108]. It also may function in the sonic hedgehog (Shh) signaling pathway as it shows protein homology to known Shh receptors [106], though this is controversial as experimental knockouts of *PTCHD1* in cells derived from brain tissue showed no change in Shh signaling-dependent cell proliferation [108]. Regardless, *PTCHD1* does not appear to have any known relevant function in dental and oral health. However, if an association between Shh signaling and *PTCHD1* does exist outside of brain tissue, it is plausible that *PTCHD1-AS* and *PTCHD1* may influence oral health and tooth loss as a receptor in Shh signaling. Shh signaling plays a regulatory role in several processes during odontogenesis (tooth development), including periodontal ligament (PDL) stem cell proliferation [109] and tooth root development [110], and the differentiation of cementoblasts [111], the cells that form cementum, which protects the tooth root and anchors PDL attachment. Furthermore, expression of *SHH* is tightly regulated and exhibits specific temporospatial patterns of expression throughout the stages of tooth development in both humans and mice [112]. In addition to this signal's potential effect on

PTCHD1, this broad signal is also located approximately 450kb downstream of *PHEX*, loss of which causes X-linked familial hypophosphatemic rickets, a disorder which includes dental defects such as premature tooth loss, periodontitis, enamel hypoplasia, hypodontia, dental caries, and abscesses [113, 114].

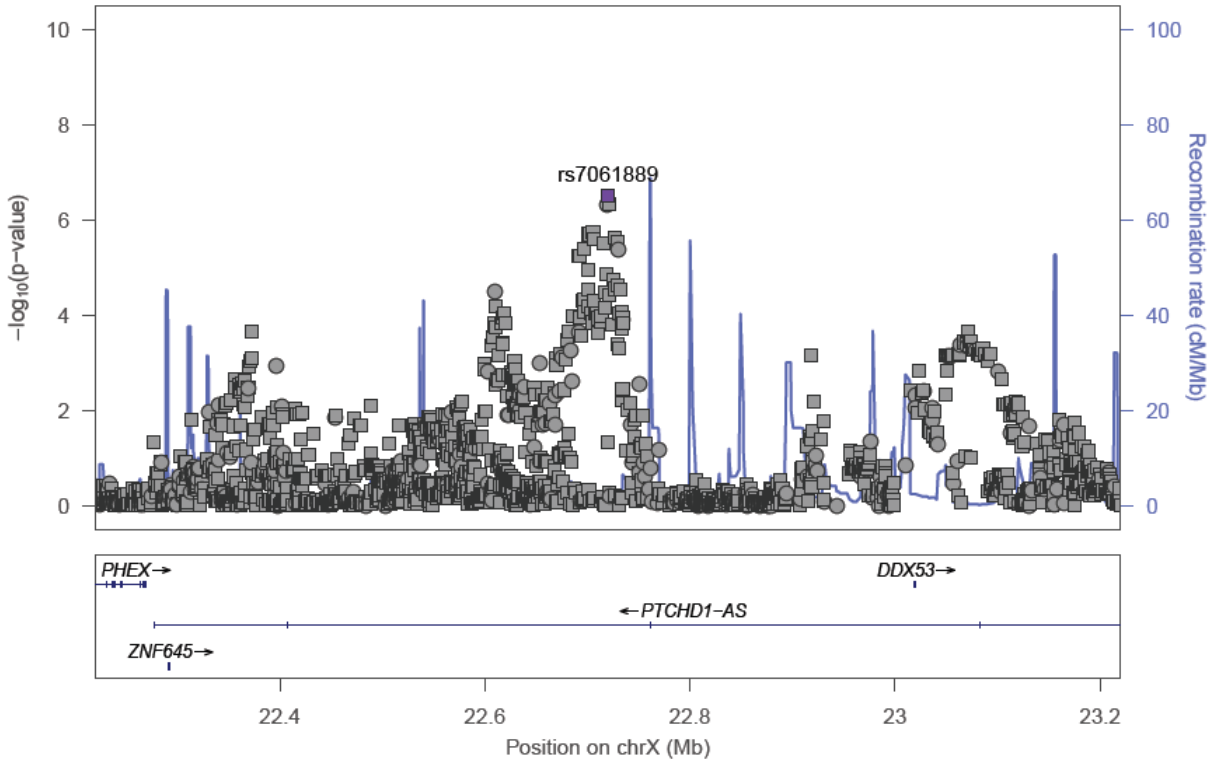


Figure 6. Regional Association Plot of rs7061889

Locations of genes are indicated. The blue overlay represents recombination rate. Squares and circles indicate imputed and genotyped variants, respectively.

rs11581023, the second most significant hit (p = 1.4E-6, β = -0.30 [-0.42, -0.18]), is located approximately 200 kb downstream of *AJAP1* (Figure 7), which was implicated in GWAS of dental caries patterns in the permanent dentition (p = 2.4E-8) [61] and smooth surface caries in the primary dentition (p = 1.6E-6) [63]. The data used in this project is an overlapping sample of adult participants of COHRA, though the region of significance around *AJAP1* differs between the samples (upstream of *AJAP1* for dental caries, downstream of *AJAP1* in this

project); in other words, the association seen here not the same as that for dental caries. *AJAP1* encodes SHREW1, a protein which may mediate matrix metalloproteinase (MMP) activity through an interaction with basigin [61, 115]. Basigin, in turn, is differentially expressed throughout the different stages of tooth development and is involved in its regulation via an interaction with MMPs [116, 117]. Thus, it is plausible that *AJAP1* may play a role in dental health and tooth loss by controlling MMP activity.

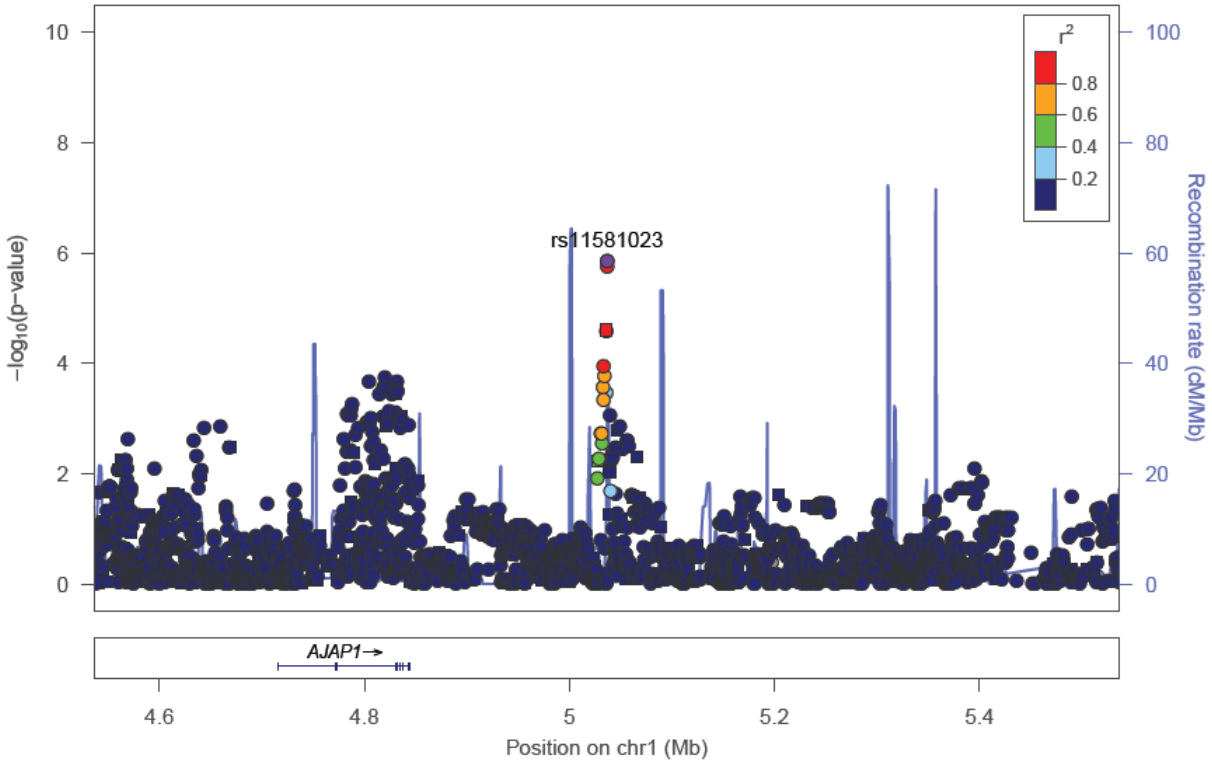


Figure 7. Regional Association Plot of rs11581023

Each point represents a genetic variant and is colored based on its correlation with the top variant. Locations of genes are indicated. The blue overlay represents recombination rate. Circles and squares indicate imputed and genotyped variants, respectively, on all plots unless otherwise indicated.

rs67374207 (p = 2.6E-6, β = 0.46 [0.27, 0.66]) is intronic to *UNC5C* and approximately 100 kb downstream of *BMPR1B* (Figure 8). *UNCSC* is a transmembrane netrin receptor that functions in axon migration during neural development [118] with no known role in dental or

oral health. *BMPR1B* encodes BMP receptor 1B. BMP signaling plays a critical role in proper tooth development and the formation of supporting structures – BMPs are expressed throughout tooth morphogenesis, help form mineralized tooth structures like dentin and enamel as well as the tooth root, and they have distinct expression patterns in periodontal structures [119]. *BMPR1B* may play a role in tooth development as a target of RNA interference mediated regulation, as knockdown of Bmpr-Ib with miR-135a inhibits tooth formation in mice [120]. *BMPRIB* is also part of a signaling cascade involved in ameloblast differentiation and enamel formation [121]. It is plausible that *BMPR1B* expression may influence tooth loss because of its role in proper tooth development, and alterations in its function may lead to defects in enamel or tooth morphology and structure.

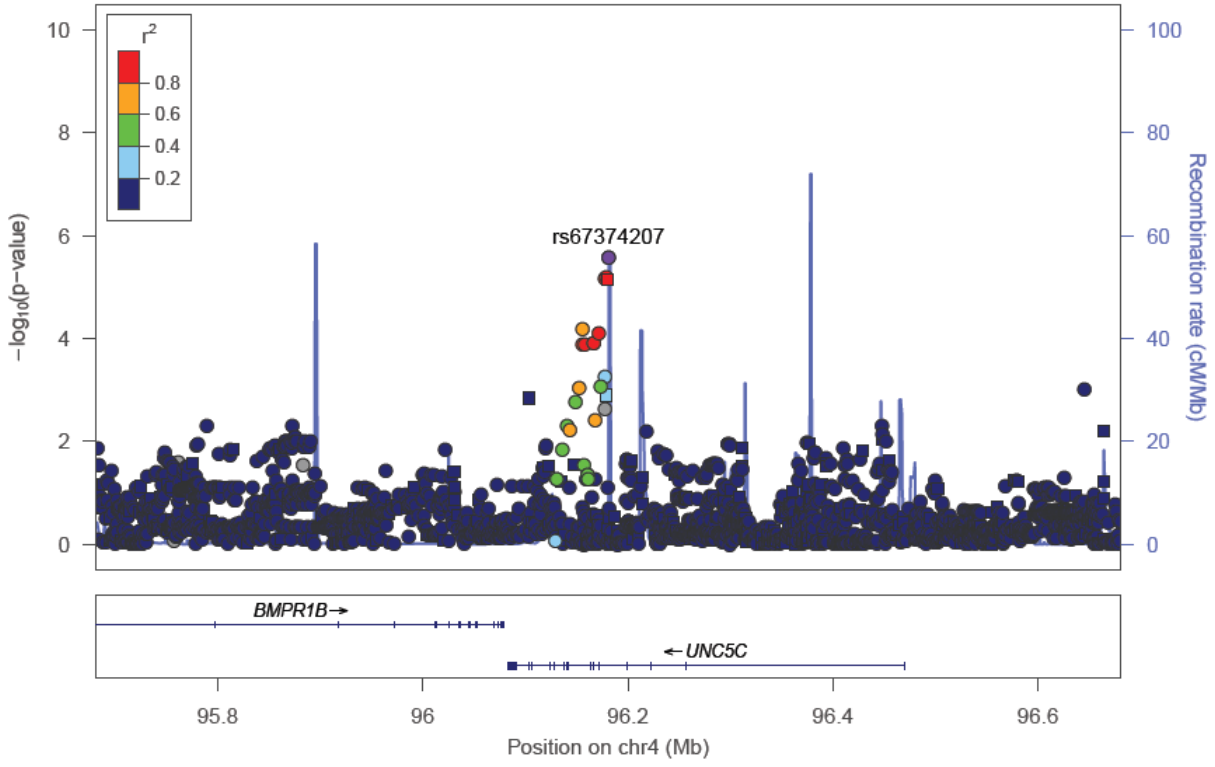


Figure 8. Regional Association Plot of rs67374207

Other SNPs reaching suggestive significance near less interesting genes as well as those not located near genes relevant to dental or oral health are described in the appendix, along with their Regional Association plots.

3.1.3 GWAS Results – Functional Dentition

Manhattan plot and QQ-plots for functional dentition are shown in Figure 9. Genomic inflation factor was $\lambda = 1.00$. One region of the genome reached genome-wide significance (index SNP chr18:8576699, p = 2.65E-8, OR = 5.00, 95% CI: [2.84, 8.82]) and several regions of the genome showed suggestive significance ($p < 10^{-5}$, Table 8.)

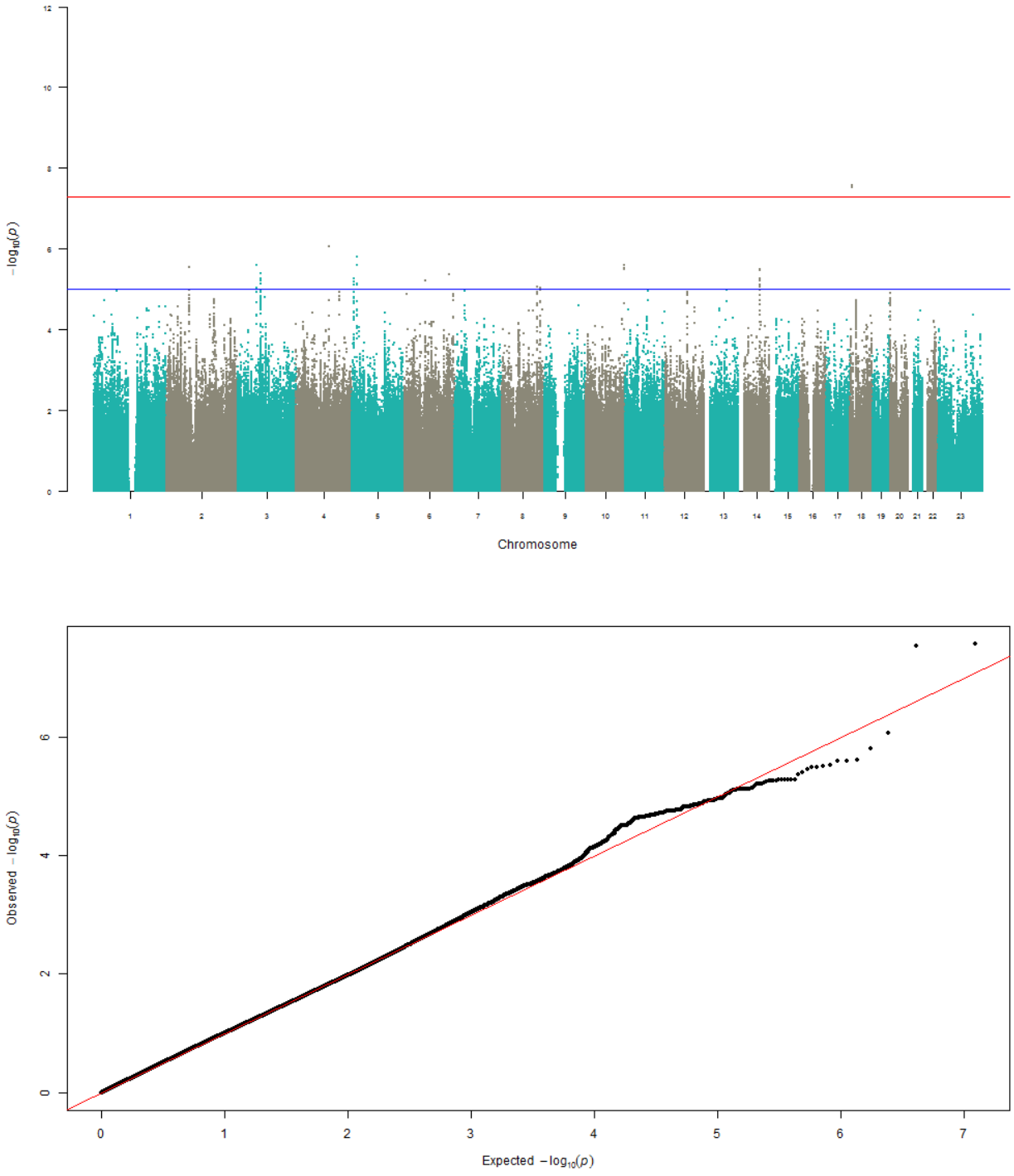


Figure 9. Manhattan and QQ plots for functional dentition in COHRA

Table 8. Top GWAS hits for functional dentition in COHRA

SNP	CHR	BP	Effect Allele	MAF	N	P	OR (95% CI)	Type
rs1487025	2	78001749	A	0.03	947	2.92E-06	5.32 (2.64, 10.71)	imputed
rs7624909	3	64461632	A	0.03	940	2.47E-06	5.29 (2.64, 10.57)	imputed
rs13060599	3	79738599	C	0.44	953	3.97E-06	0.45 (0.32, 0.63)	imputed
rs72673432	4	114066021	A	0.05	914	8.61E-07	4.09 (2.34, 7.18)	imputed
rs10055463	5	7973236	G	0.12	954	5.45E-06	2.56 (1.71, 3.85)	imputed
rs7717485	5	19762485	A	0.02	951	1.58E-06	8.19 (3.47, 19.31)	imputed
rs1349926	6	68983548	G	0.39	954	6.26E-06	0.43 (0.3, 0.62)	imputed
rs629476	6	150673394	T	0.03	948	4.27E-06	5.28 (2.6, 10.73)	imputed
rs1154819	8	120226259	A	0.02	948	8.79E-06	5.85 (2.68, 12.74)	imputed
rs76944100	8	132662517	T	0.04	938	9.11E-06	3.78 (2.10, 6.81)	imputed
rs76798443	10	129951990	T	0.04	938	2.51E-06	4.49 (2.4, 8.38)	imputed
chr14:74029866	14	74029866	C	0.04	912	3.20E-06	4.28 (2.32, 7.88)	imputed
chr18:8576699	18	8576699	A	0.04	903	**2.65E-08**	5.00 (2.84, 8.82)	imputed

Bolded text denoted genome-wide significance.

The genome-wide significant hit at chr18:8576699 contains only two SNPs (Figure 10), both of which are imputed, and lacks a characteristic tower of other SNPs in LD with the index SNP. Furthermore, there are no genes in this region likely to play a role in dental and oral health. Therefore, this finding should be viewed with skepticism. This signal is just upstream of *RAB12*, a GTP binding protein/GTPase involved in vesicle trafficking and regulating autophagy [122, 123] but with no known role in dental health.

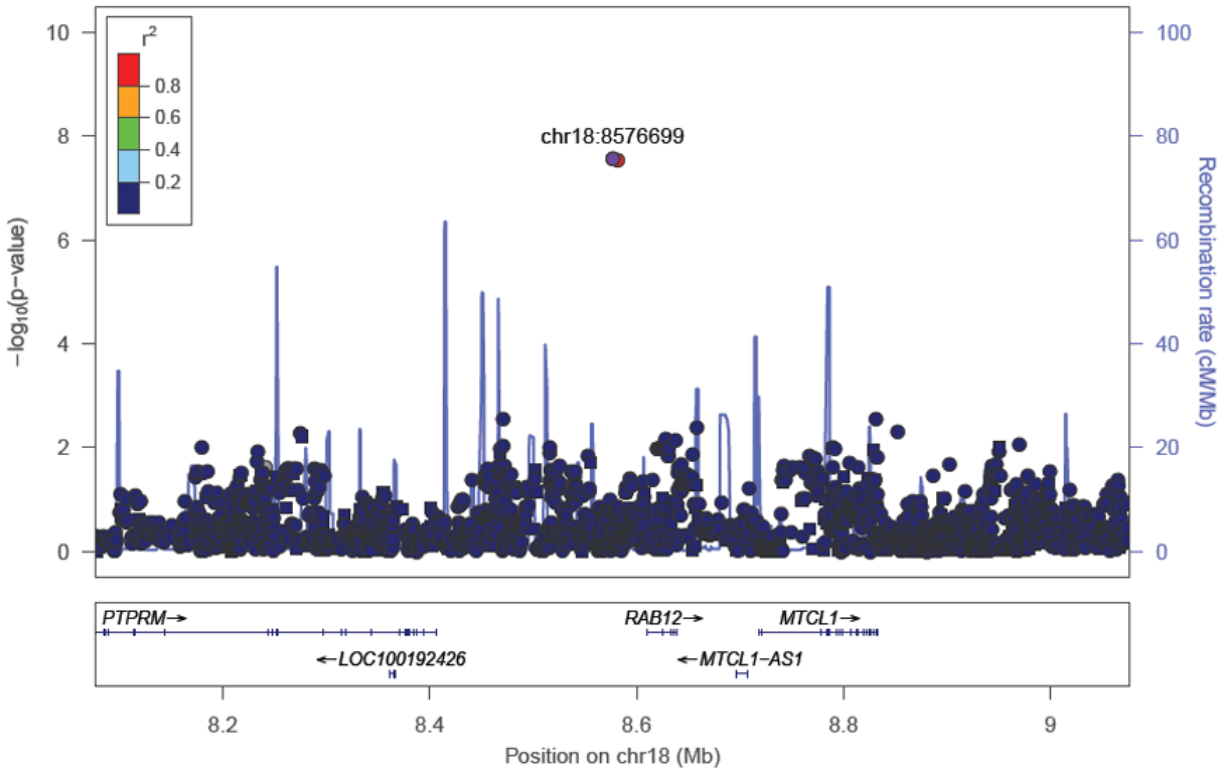


Figure 10. Regional Association Plot of chr18:8576699

rs10055463 (p = 5.45E-6, OR = 2.56, 95% CI: [1.71, 3.85]) is located just downstream of *MTRR* (Figure 11). *MTRR* encodes methionine synthase reductase, an enzyme that restores methionine synthase to its functional state, and is important in folate metabolism [124]. An *MTRR* polymorphism has been associated with dental caries [125], and *MTR*, methionine synthase, was previously implicated in a GWAS of dental caries in the primary dentition [59], and a follow-up study found SNPs within *MTR* to be suggestively associated with dental caries [126].

In addition to *MTRR*, this signal is also approximately 150 kb downstream of *ADCY2*. This gene encodes an adenylyl cyclase that catalyzes the formation of cyclic adenosine monophosphate (cAMP), and the cAMP signaling pathway is an important mediator in cellular

response to mechanical stress [127]. Indeed, cAMP signaling has also been shown to be induced by mechanical force on PDL cells [128, 129], indicating an role of cAMP signaling in maintaining homeostasis analogous to *CTNND2*, as previously discussed. A coding polymorphism in *GPR126*, a G-protein receptor that activates the cAMP/PKA signaling pathway, resulting in decreased cAMP signaling in human PDL cells, has been associated with aggressive periodontitis [130]. Lastly, Du et al. (2016) studied miRNA expression in PDL cells exposed to *P. gingivalis* lipopolysaccharides and found enrichment for miRNAs involved in cAMP signaling among the differentially expressed miRNAs [131], indicating that not only does cAMP signaling play a role in maintaining periodontal homeostasis, but is also potentially involved in the aberrant immune response to oral microbes characteristic of periodontitis.

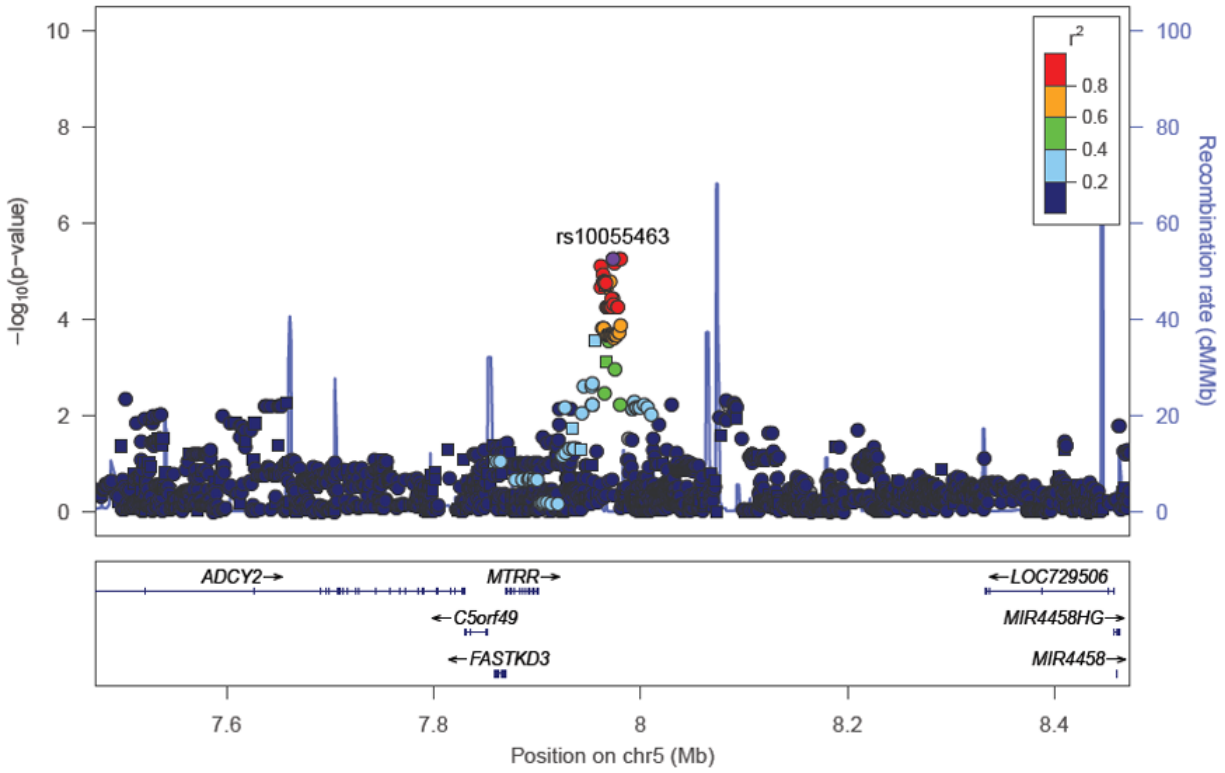


Figure 11. Regional Association Plot of rs10055463

rs1154819 (p = 8.79E-6, OR = 5.85, 95% CI: [2.68, 12.74]) is located approximately 250 kb upstream of *TNFRSF11B* (Figure 12), is also known as *OPG*, which codes for the protein osteoprotegerin. This gene was found to be downregulated in PDL cells in response to mechanical stress [132]. Furthermore, *OPG* regulates the production of osteoclasts, the cells responsible for bone resorption [133]. Indeed, *Opg* knockout mice have severe alveolar bone loss, which is characteristic of periodontitis and periodontal tooth loss [134], and levels of *OPG* are negatively correlated with severity of periodontal disease [135]. Increasing *OPG* levels via gene therapy in rats with experimentally induced periodontitis reduced alveolar bone loss [136]. Also in this region is *NOV*, also known as *CCN3*, which is located approximately 200 kb downstream of the lead SNP. *CCN3* is expressed in the PDL and has been shown to associate with periostin [137], a critical regulator of periodontal homeostasis [138]. *CCN3* is also upregulated in dental pulp stem cells during the dental repair process and promotes dentin formation by regulating Notch and BMP2 signaling [139].

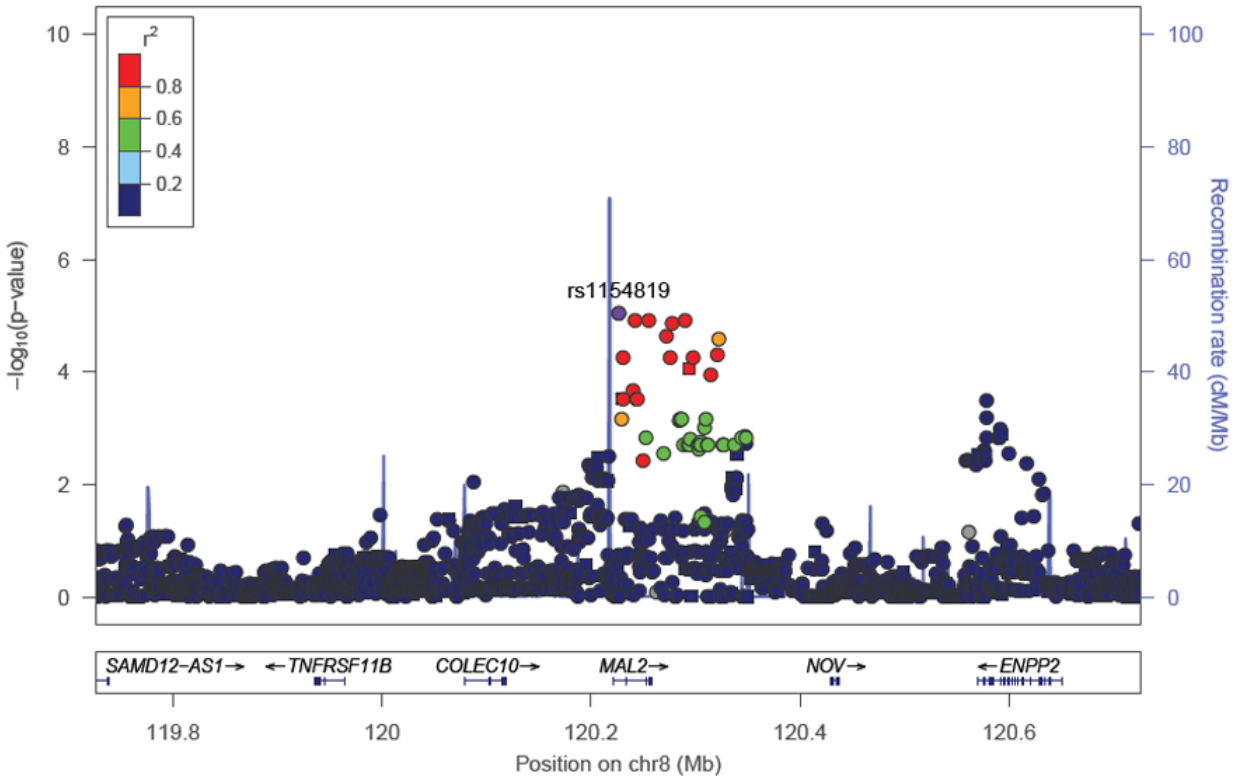


Figure 12. Regional Association Plot of rs1154819

Other SNPs reaching suggestive significance near less interesting genes as well as those not located near genes relevant to dental or oral health are described in the appendix, along with their Regional Association plots.

3.2 DRDR

3.2.1 Principal Components of Ancestry

Figures 13 show PCA plots for the whites-only subset of DRDR. As stated in the methods section, only self-reported and genetically confirmed whites were included in the analysis. As shown in Figure 13, one principal component was sufficient to control for ancestry.

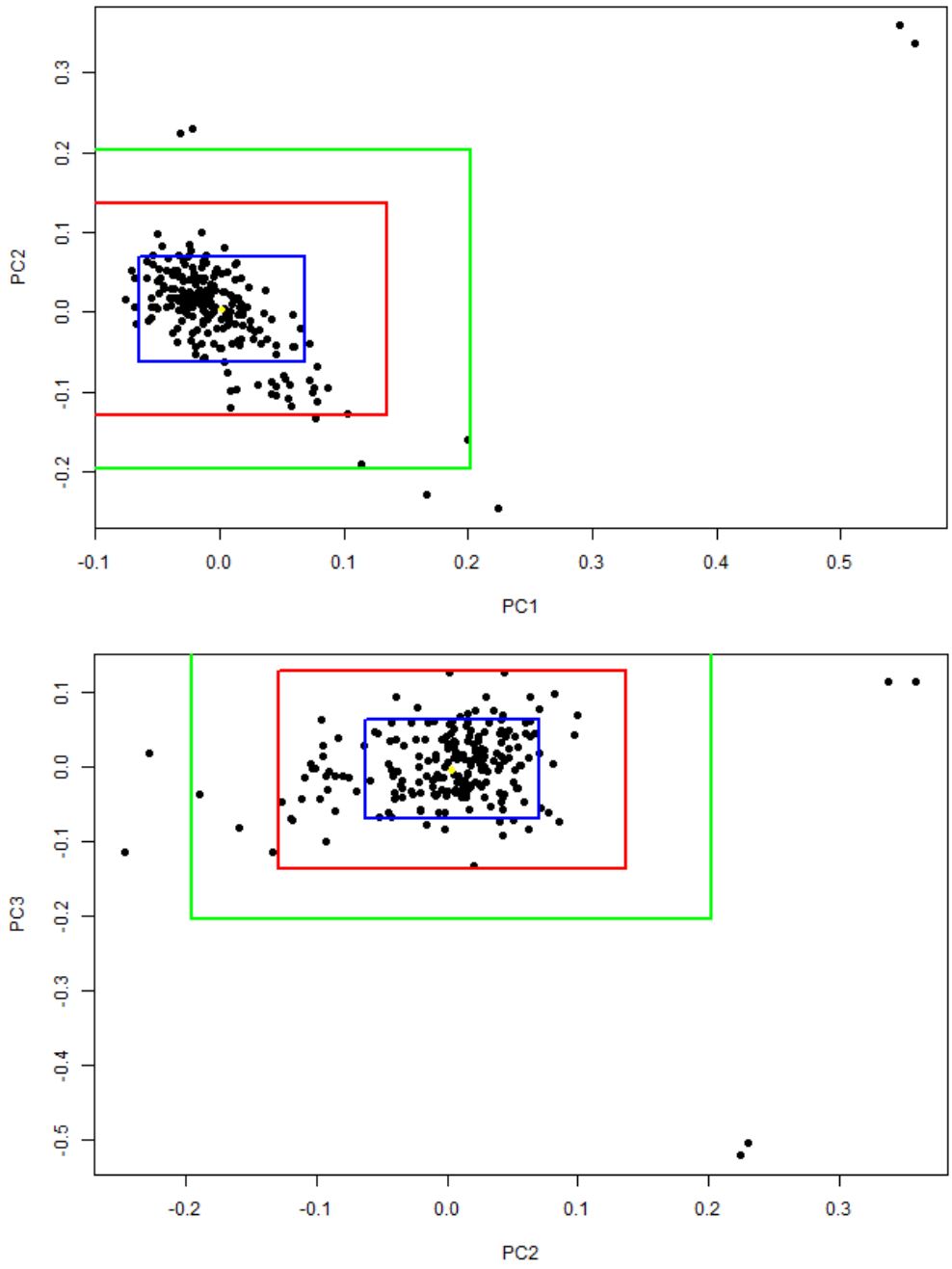


Figure 13. PCA plots of PC1-2 in DRDR

Blue, red, and green boxes indicate one, two, and three standard deviations from the mean on each axis, respectively.

3.2.2 GWAS Results

GWAS was performed using 227 individuals for whom genotype and covariate information was available. Manhattan plot and QQ-plots for the quantitative trait and functional dentition are shown in Figures 14-15. Genomic inflation factor was $\lambda = 1.00$ and $\lambda = 0.97$ for the quantitative trait and functional dentition, respectively, indicating no genomic inflation. Due to the small sample size and lack of genome-wide significant signals, no genomic regions were followed up to see if they contained genes with potential roles in dental or oral health.

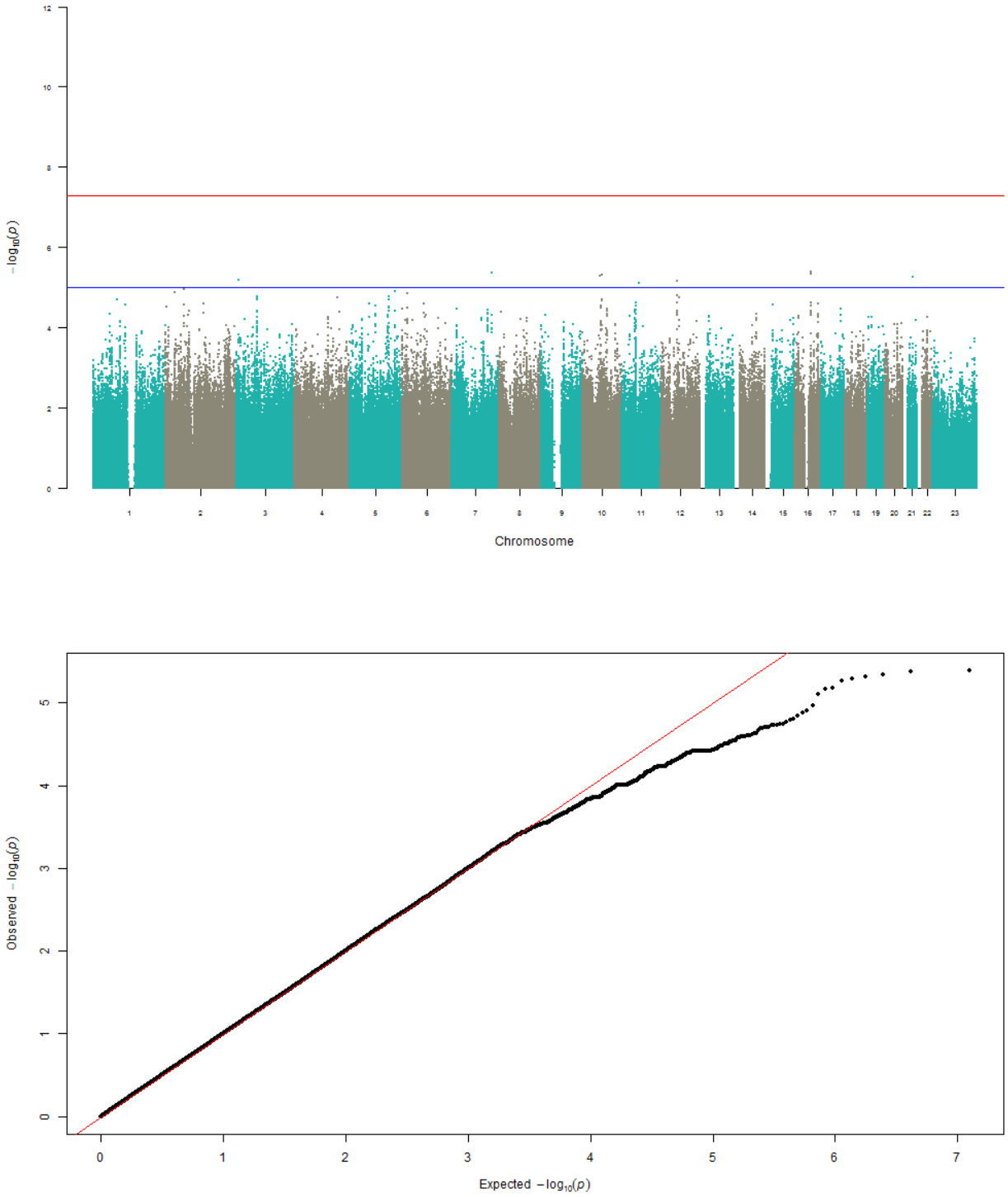


Figure 14. Manhattan and QQ plots for the quantitative trait in DRDR

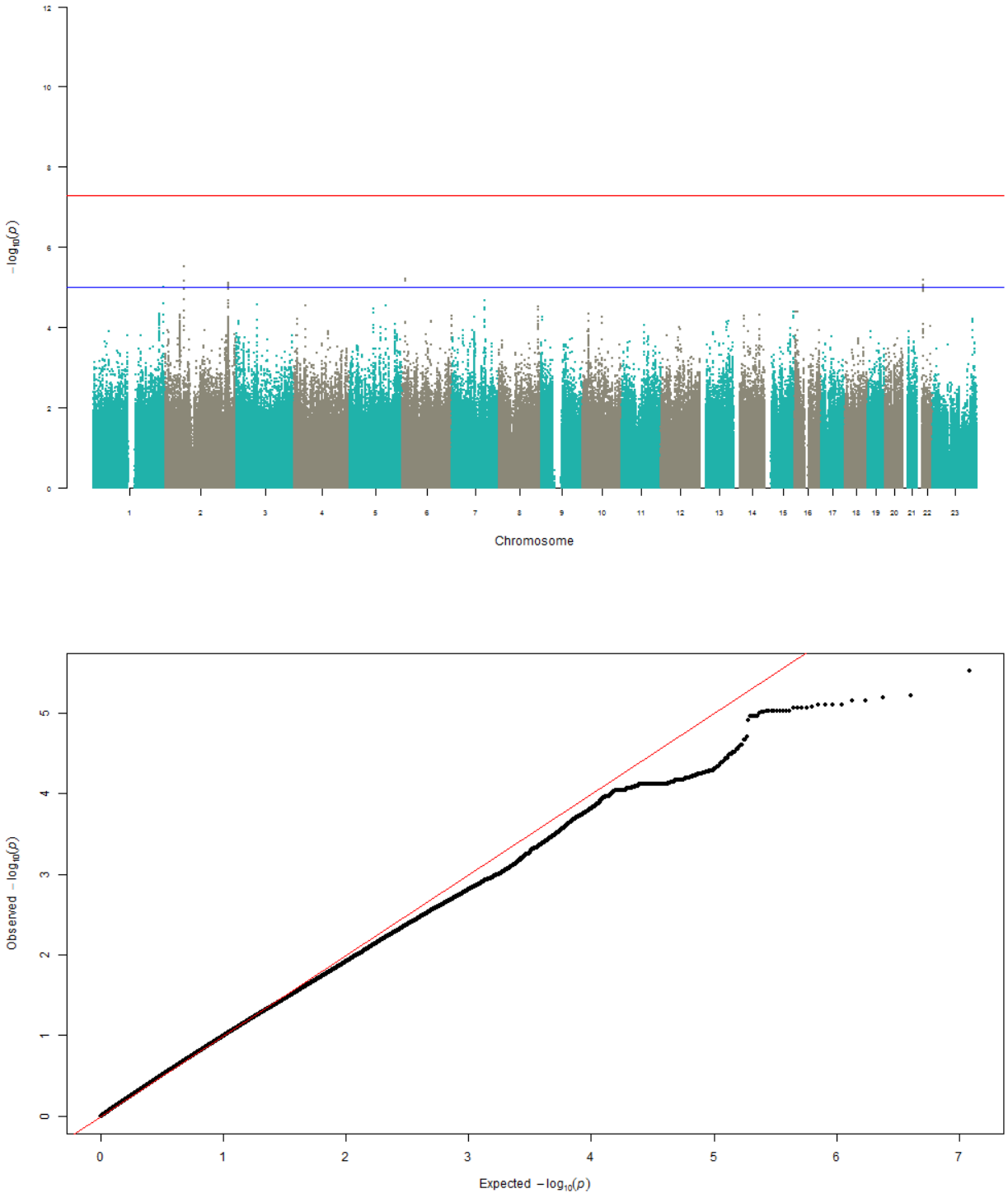


Figure 15. Manhattan and QQ plots for functional dentition in DRDR

3.3 POFC-G AND POFC-PA

3.3.1 Covariate Modeling

Distribution of missing teeth traits, as well as covariates included in the GWA scan, are shown in Table 9. Association between covariates (sex and age) with missing teeth traits are shown in Table 10. As with COHRA, the sample size in these results is greater than that of the GWAS, as some individuals did not have genotype information available but were included in initial trait development and covariate modeling.

Table 9. Study characteristics in POFC-G and POFC-PA

Study	Variable	Value	Count (%)/ Mean (SD)
POFC-G	Sex	Male	109 (31)
		Female	243 (69)
	Age		31.38 (11.32)
	Quantitative Trait		1.14 (1.00)
	Functional Dentition	No	52 (15)
		Yes	301 (85)
POFC-PA	Sex	Male	70 (30)
		Female	170 (70)
	Age		36.11 (12.1)
	Quantitative Trait		1.40 (0.92)
	Functional Dentition	No	46 (19)
		Yes	199 (81)

Categorical covariates show counts and percents, and continuous variables show means and standard deviations.

Table 10. Association of covariates with missing teeth traits in POFC-G and POFC-PA

Study	Variable	Quantitative Trait		Binary Trait	
		P	Beta (SE)	P	Beta (SE)
POFC-G	Sex	0.027	0.27 (0.12)	0.19	-0.47 (0.35)
	Age	9.30E-24	0.044 (0.004)	6.73E-10	-0.083 (0.013)
POFC-PA	Sex	0.74	0.043 (0.13)	0.55	-0.22 (0.37)
	Age	1.17E-21	0.043 (0.004)	1.06E-10	-0.17 (0.018)

3.3.2 GWAS Results

For POFC-G, GWAS was performed in 272 individuals for whom genotype information was available. Manhattan and QQ-plots for the quantitative trait and functional dentition are shown in Figures 16 and 17. Genomic inflation factor was $\lambda = 1.01$ and $\lambda = 0.99$ for the quantitative trait and functional dentition, indicating no inflation due to population stratification. Suggestively significant results for both traits are described in the appendix.

For the quantitative trait, one region of the genome reached genome-wide significance (rs12430287, p = 4.1E-8). This SNP is located approximately 150 kb downstream of *POSTN* (Figure 18), which encodes the protein periostin. Periostin, a matricellular protein expressed in collagenous connective tissues, is critically important in the formation, maintenance, and function of dental tissues [140]. During tooth development, it is widely expressed in dental tissues and regulates tooth development by controlling the composition of the ECM. After development, periostin is expressed in the PDL and alveolar bone, where it is critical in maintaining these structures by ensuring proper formation of the fibers of the ECM. Additionally, periostin mitigates the PDL response to mechanical force and orthodontic tooth movement, likely because of its role in ECM remodeling. Lastly, periostin also functions in

wound healing through pro-fibrogenic processes such as the production of collagen fibers and ECM formation [141]. Periostin null mice present with severe dental defects together with an early-onset periodontal phenotype that recapitulates the destructive tissue loss characteristic of periodontitis [142, 143]. More specifically, enamel defects, malformed teeth, abnormal alveolar bone remodeling, tooth root resorption, widening of the PDL, alveolar bone loss, attachment loss, gingival enlargement, and inflammation were all observed in null mice, indicating that periostin is required for proper functioning of the PDL. Reduction of mechanical force on the periodontium in null mice ameliorated the symptoms of periodontal disease, and mRNA expression of periostin increases in PDL cells in response to mechanical strain [143]. Periostin expression levels are correlated with measures of periodontal disease such as bleeding on probing, and expression patterns in periodontal samples differ between patients with chronic (CP) and aggressive (AgP) periodontitis and normal controls [144]. There is a clear mechanism through which genetic variation around *POSTN* leading to irregular periostin expression may contribute to susceptibility to periodontal disease and tooth loss. In addition to *POSTN*, this signal is also approximately 500 kb upstream of *SMAD9*, which is a transcriptional regulator of BMP proteins and signaling [145], which, as previously discussed, are critical processes in tooth development.

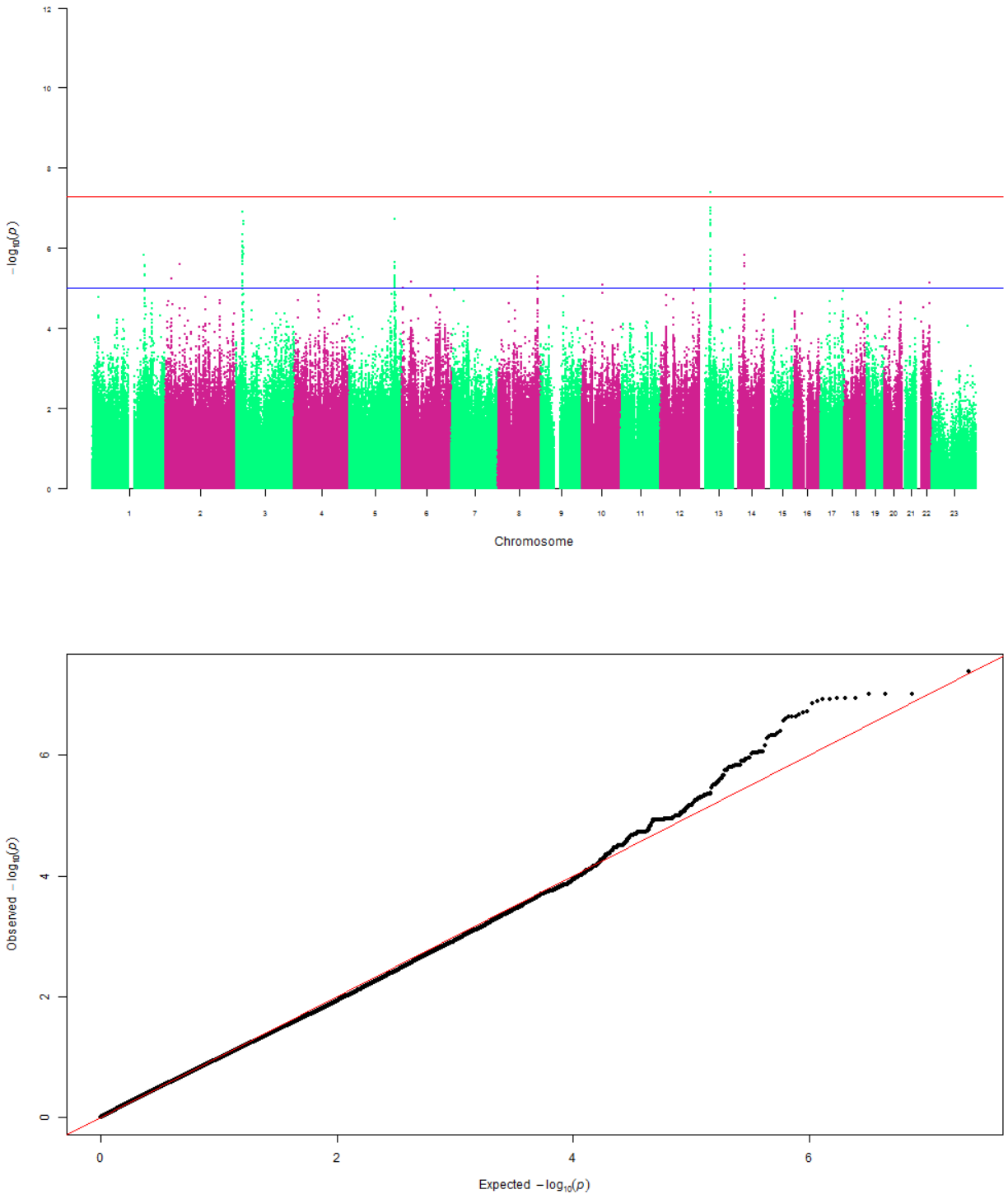


Figure 16. Manhattan and QQ plots for the quantitative trait in POFC-G

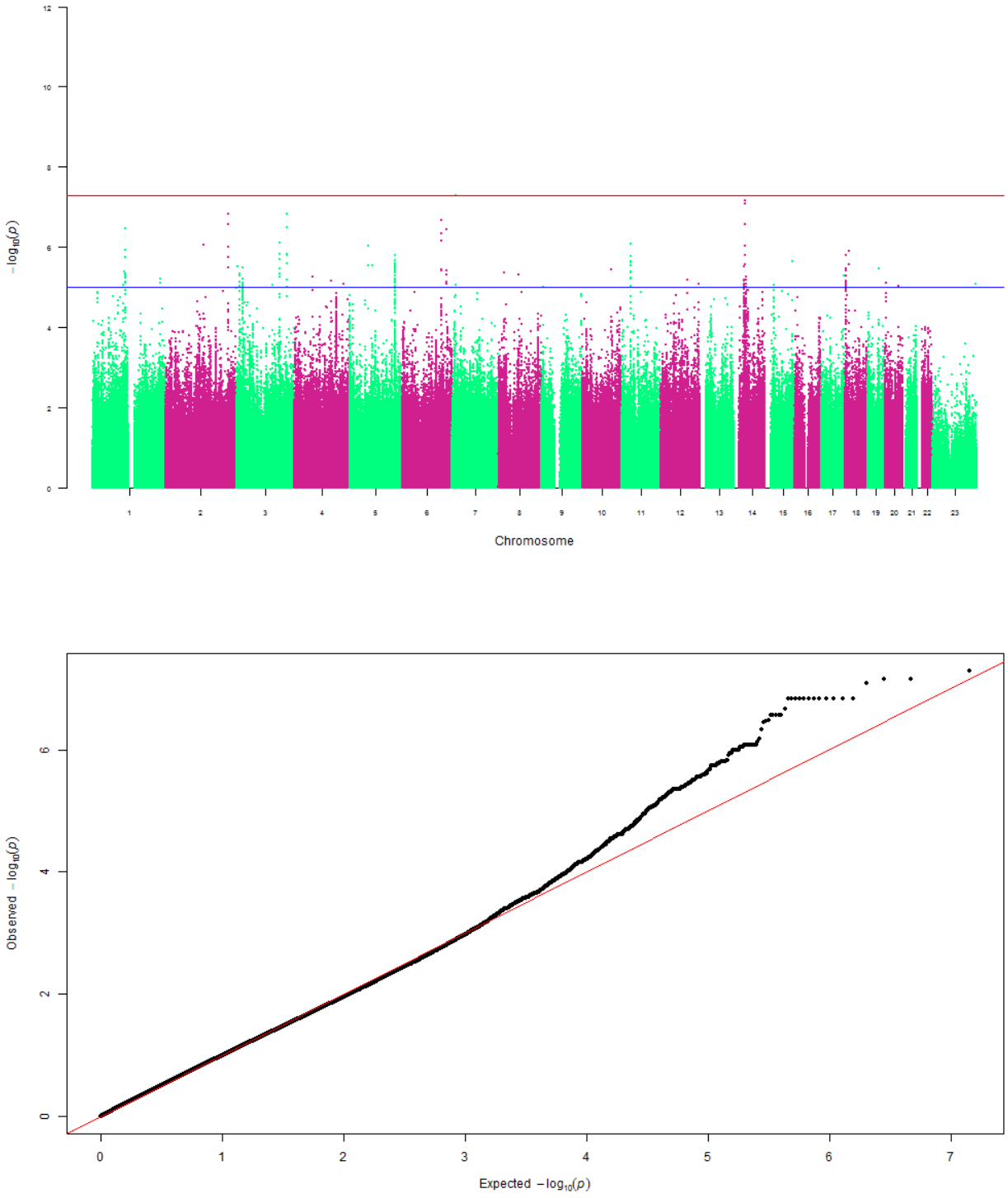


Figure 17. Manhattan and QQ plots for functional dentition in POFC-G

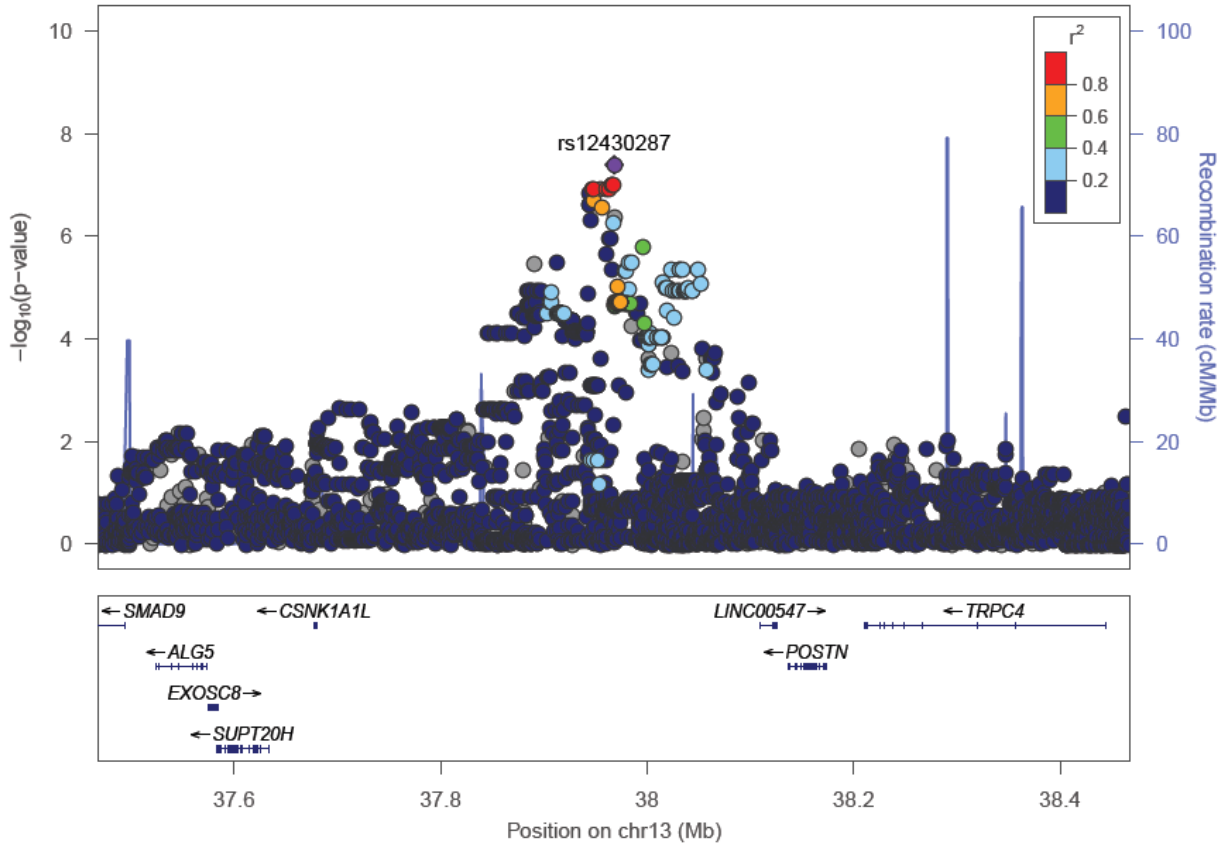


Figure 18. Regional Association Plot of rs12430287

Annotation of imputed vs. genotypes variants was unavailable in POFC-G and POFC-PA.

For POFC-PA, GWAS was performed in 182 individuals for whom genotype information was available. Manhattan plot and QQ-plots for the quantitative trait and functional dentition are shown in Figures 19-20. Genomic inflation factor was $\lambda = 0.92$ and $\lambda = 0.90$ for the quantitative trait and functional dentition, both of which are indicative of deflation. As with DRDR, the results of this GWA scan were not followed up due to the small sample size and lack of genome-wide significant hits.

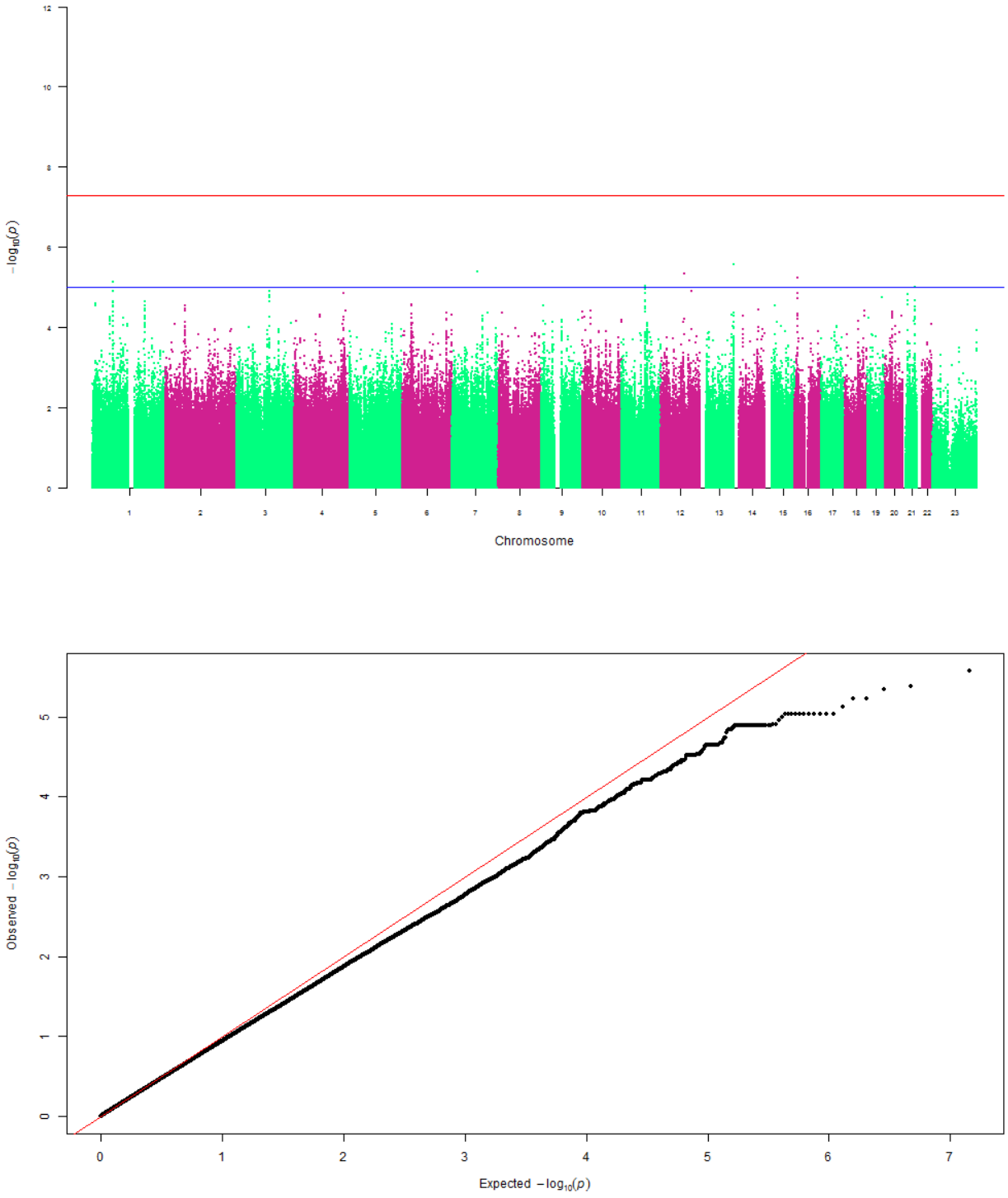


Figure 19. Manhattan and QQ plots for the quantitative trait in POFC-PA

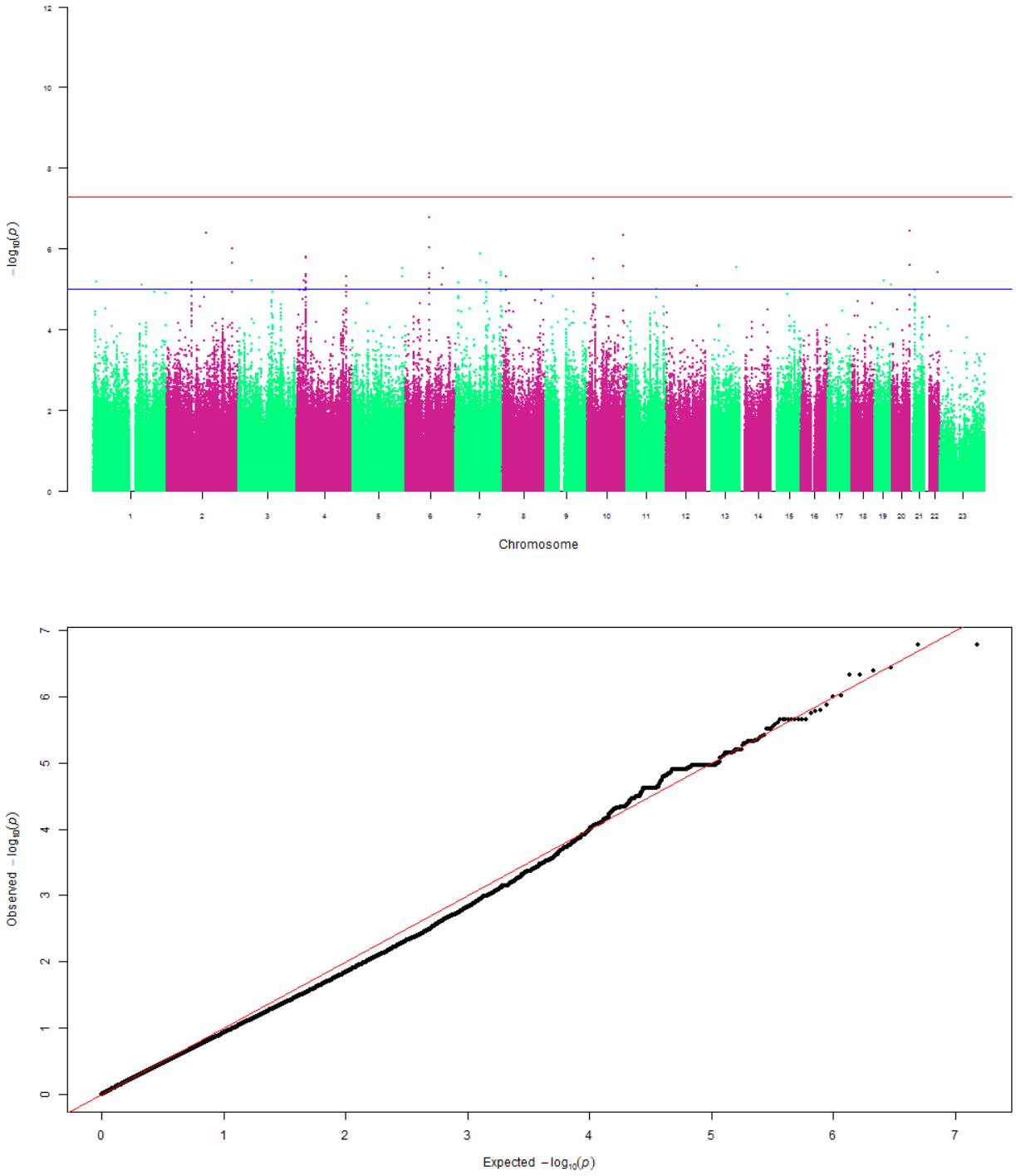


Figure 20. Manhattan and QQ plots for functional dentition in POFC-PA

3.4 META-ANALYSES

3.4.1 Whites-Only Meta-Analysis – Quantitative trait

A p-value based meta-analysis was performed for the quantitative trait to combine GWAS results from COHRA and DRDR, which represent the European ancestry samples. Manhattan and QQ-plots are shown in Figure 21. Genomic inflation factor was $\lambda = 1.05$. While no regions of the genome reached genome-wide significance, several regions showed suggestive significance ($p < 10^{-5}$, Table 11).

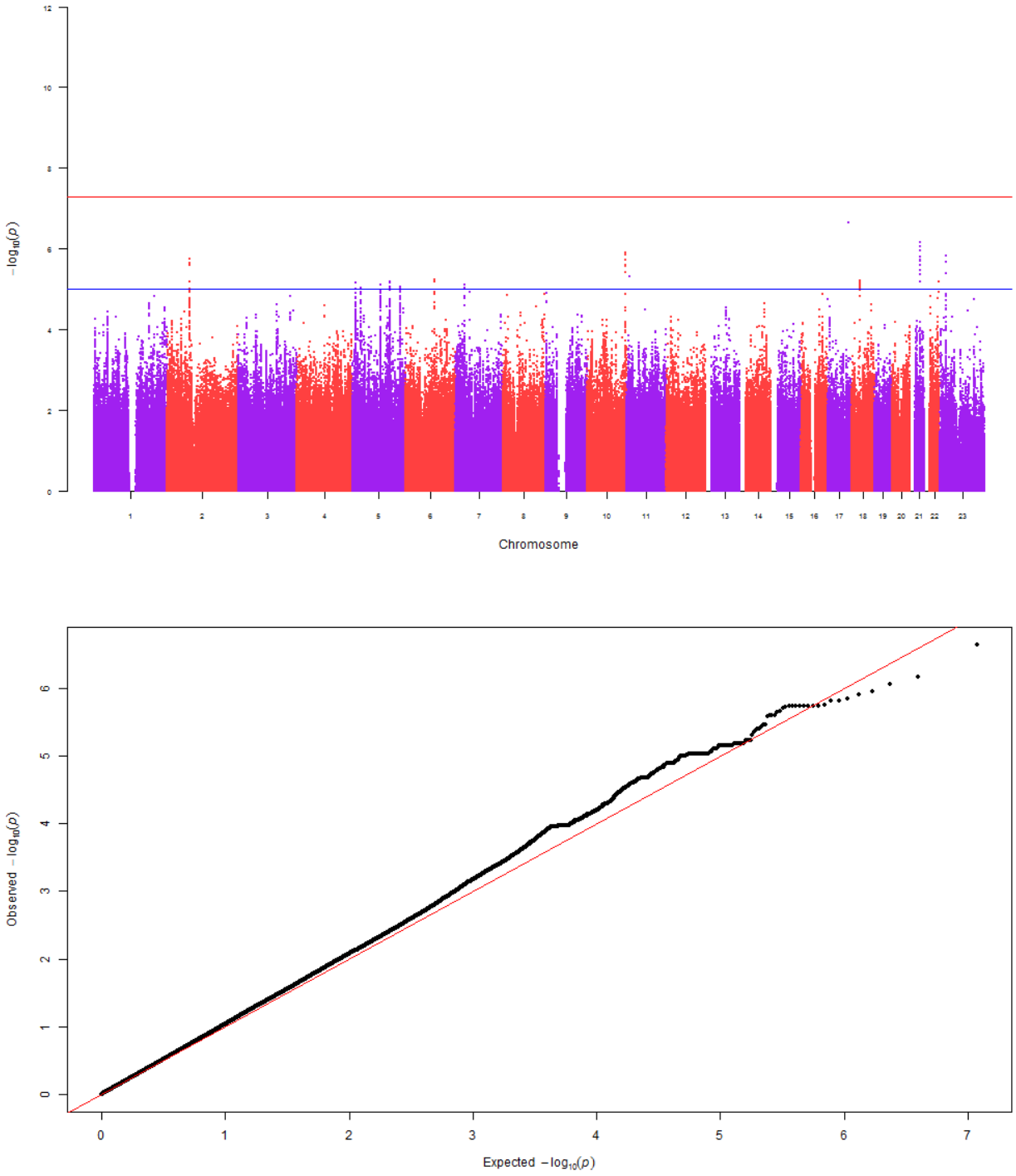


Figure 21. Manhattan and QQ plots for the whites-only meta-analysis of the quantitative trait in COHRA and DRDR

Table 11. Top hits for the whites-only meta-analysis of the quantitative trait in COHRA and DRDR

SNP	CHR	BP	A1/A2	MAF COHRA	MAF DRDR	N	P	Type	DIR*
rs9750906	2	76804252	c/g	0.29	0.31	1113	1.77E-06	imputed	++
rs3797113	5	10695026	c/t	0.06	0.07	1181	7.04E-06	imputed	--
rs79950078	5	29703754	a/c	0.05	0.07	1180	9.19E-06	imputed	++
rs9918187	5	96531684	a/g	0.35	0.33	1148	7.67E-06	imputed	--
rs10062700	5	128180057	a/c	0.36	0.31	1168	6.42E-06	imputed	--
rs2860805	5	163265408	c/g	0.24	0.33	1180	8.78E-06	imputed	--
rs56345510	6	99599021	a/g	0.37	0.34	1150	5.95E-06	imputed	++
rs77332164	7	31048669	c/g	0.03	0.07	1182	7.62E-06	imputed	--
rs78472857	10	131481707	g/t	0.12	0.10	1155	1.25E-06	imputed	++
rs3365	11	8704711	a/c	0.14	0.11	1182	4.82E-06	genotyped	--
rs4969040	17	70947492	a/g	0.09	0.17	1181	2.25E-07	genotyped	--
rs4444411	18	27622307	g/t	0.09	0.12	1181	6.18E-06	imputed	++
rs73199539	21	32938208	c/t	0.06	0.05	1158	6.75E-07	imputed	++
rs11704818	22	48248892	c/t	0.18	0.17	1089	6.68E-06	imputed	--
rs7061889	23	22719993	c/t	0.15	0.16	1177	1.52E-06	imputed	++

*Direction of effect in COHRA and DRDR

rs4969040, the most significant SNP (p = 2.25E-7), is intronic to *SLC39A11*, a member of the solute carrier family of membrane transporters (Figure 22). While this signal lacks a tower of significant SNPs in LD with the index SNP typical of other index SNPs, it was genotyped in these two samples and showed suggestive evidence of association in both COHRA

and DRDR (p = 6.3E-5 and 3.1E-4, respectively.) *SLC39A11* is responsible for cellular transport of zinc [146], which is involved in the mineralization of tissues and helps control dental plaque and calculus accumulation, though the effect of zinc on dental caries is controversial [147]. High zinc levels are correlated with lower rates of enamel demineralization [148], but higher levels of zinc have been reported in adults with dental caries (DMFT > 10) compared to those with no caries (DMFT = 0) [149]. It is unclear exactly how zinc may influence caries susceptibility, and the effect of *SLC39A11* on dental caries has not yet been evaluated, but it is plausible that *SLC39A11* may influence zinc levels, dental caries, and tooth loss. In addition to *SLC39A11*, this signal is also approximately 200 kb upstream of *SSTR2*, somatostatin receptor 2. Somatostatins (SST) are neuropeptides involved in nociception, the sensory response to painful or harmful stimuli, and it has been suggested that this receptor, *SSTR2*, may modulate orofacial pain by inhibiting neuronal activity [150]. SSTs also respond directly to experimentally induced tooth pain in rats, with possible involvement of SST receptor 2 [151, 152]. Excessive pain may lead to tooth extraction [29].

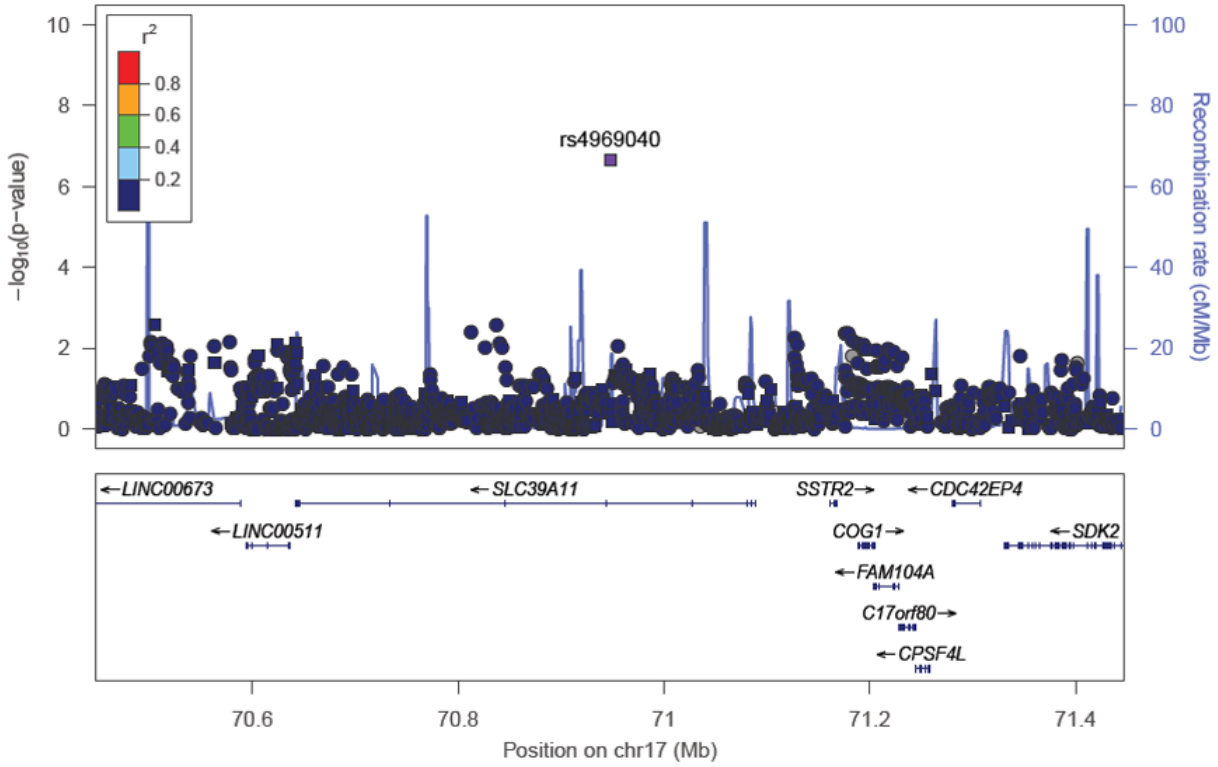


Figure 22. Regional Association Plot of rs4969040

Circles and squares indicate imputed and genotyped variants, respectively.

rs10062700 (p = 6.5E-6) is located about 300 kb upstream of *FBN2*, fibrillin-2 (Figure 23). Fibrillins polymerize to form microfibrils in the ECM of connective tissues, regulate TGF-β and BMP signaling, and are involved in cellular mechanoreception [153, 154]. In periodontal tissues, fibrillin-2 is synthesized by gingival fibroblasts and PDL cells [155-157] and exhibits distinct patterns of expression in the developing tooth and during PDL development [158]. These observations indicate that as a component of microfibrils, fibrillin-2 contributes to the development, structural stability, and homeostasis of the PDL. Its role in both the PDL and in regulating BMP signaling make *FBN2* a strong candidate for future genetic studies of tooth loss.

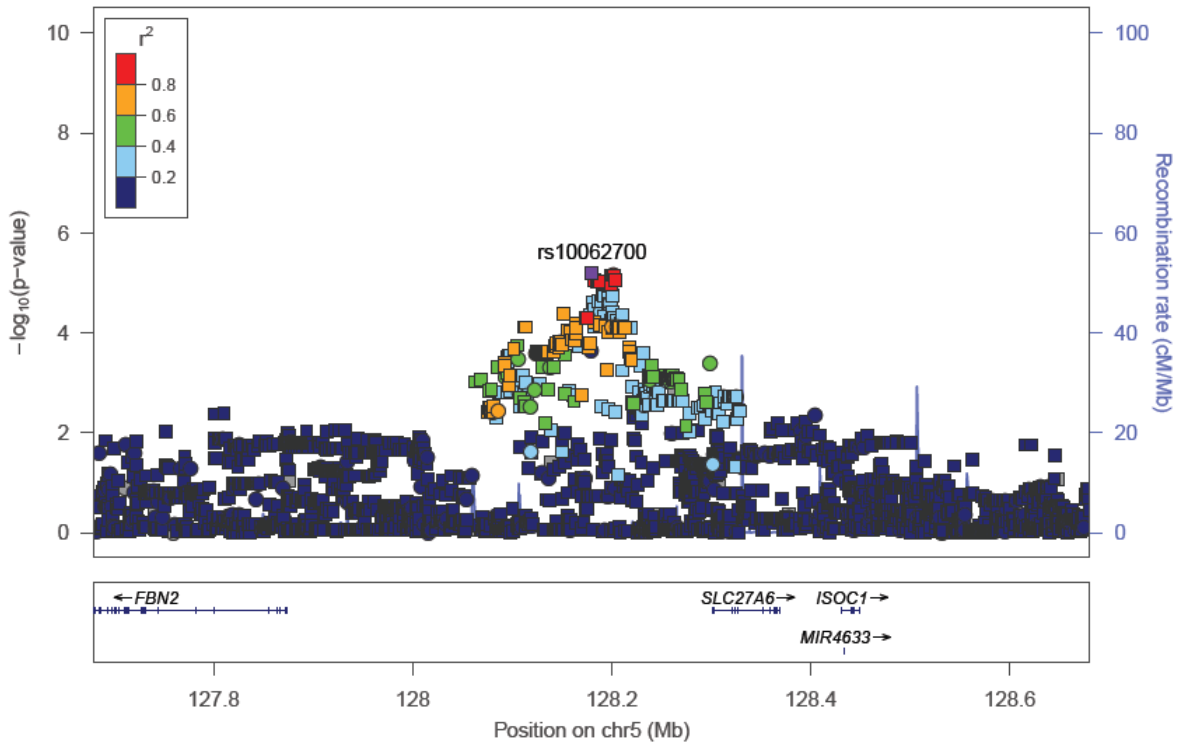


Figure 23. Regional Association Plot of rs10062700

Squares and circles indicate imputed and genotyped variants, respectively.

rs73199539 (p = 6.75E-07) corresponds to the same signal seen in COHRA only, and is located near *SOD1*, which has implications for the role of superoxide dismutase and oxidative damage in dental caries and periodontitis, as previously discussed. This SNP is only nominally significant in DRDR (p = 0.076) and likely remained a top signal in the meta-analysis because of its strong association (p = 3.18E-6) in COHRA, though the direction of effect (negative) is concordant between the two studies. Other SNPs showing strong evidence of association in the meta-analysis because of high statistical significance in COHRA but with weak evidence in DRDR include rs78472857 at *MGMT* (described in the appendix), rs7061189 at *PHEX-PTCHD1-AS*, and rs2860805 near *MAT2B*, as well as rs3365, rs56345510, rs79950078

(COHRA index SNP: rs77475896), rs4444411, and rs9918187 which do not contain relevant genes and were not discussed. Other SNPs reaching suggestive significance near less interesting genes as well as those not located near genes relevant to dental or oral health are described in the appendix, along with their Regional Association plots.

3.4.2 Whites-Only Meta-Analysis – Functional Dentition

A p-value based meta-analysis was performed for functional dentition to combine GWAS results from COHRA and DRDR, which represent the European ancestry samples. Manhattan and QQ-plots are shown in Figure 24. Genomic inflation factor was $\lambda = 1.01$. One region on chromosome 5 reached genome wide significance (index SNP rs6898589, p = 4.10E-8) and several regions of the genome showed suggestive significance ($p < 10^{-5}$, Table 12.)

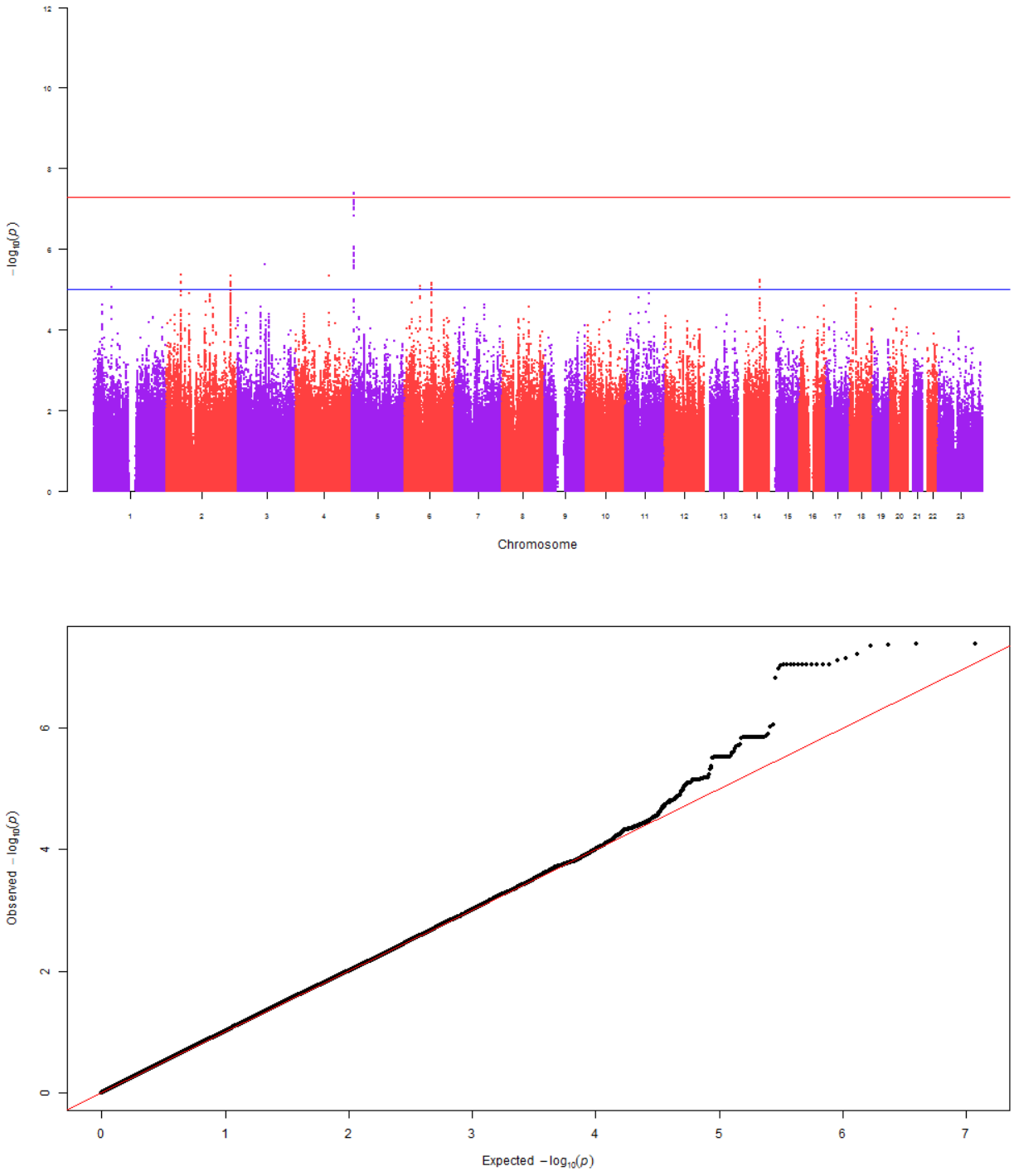


Figure 24. Manhattan and QQ plots for the whites-only meta-analysis of functional dentition in COHRA and DRDR

Table 12. Top hits for the whites-only meta-analysis of functional dentition in COHRA and DRDR

SNP	CHR	BP	A1/A2	MAF COHRA	MAF DRDR	N	P	Type	DIR*
rs600411	1	61355371	a/g	0.04	0.09	1182	8.83E-06	genotyped	--
rs3930612	2	50896945	a/t	0.21	0.21	1152	4.25E-06	imputed	--
rs1455837	2	219912758	a/g	0.35	0.40	1182	4.62E-06	genotyped	++
rs13081582	3	94495036	a/c	0.39	0.42	1181	2.43E-06	genotyped	++
rs72673432	4	114066021	a/g	0.05	0.04	1134	4.69E-06	imputed	++
rs6898589	5	7961941	a/g	0.14	0.18	1171	**4.10E-08**	imputed	++
rs875142	6	52228225	a/g	0.25	0.31	1155	8.49E-06	imputed	--
rs1090071	6	93112404	a/t	0.32	0.30	1178	6.83E-06	imputed	++
rs55705802	14	73995075	c/g	0.04	0.07	1159	5.93E-06	imputed	++

Bolded text denotes genome-wide significance; *Direction of effect in COHRA and DRDR

rs6898589, which corresponds to the same signal near *MTRR* seen in COHRA only, reached genome-wide significance in this whites-only meta-analysis (p = 4.1E-8, p_{COHRA} = 8.0E-6, p_{DRDR} = 7.8E-4; Figure 25). As discussed previously, *MTRR* functions in methionine synthesis and has a putative role in tooth loss as both *MTRR* and methionine synthesis have previously been implicated in dental caries. Also in this region is *ADCY2*, which is involved in cAMP signaling and has implications for tooth loss due to periodontal disease.

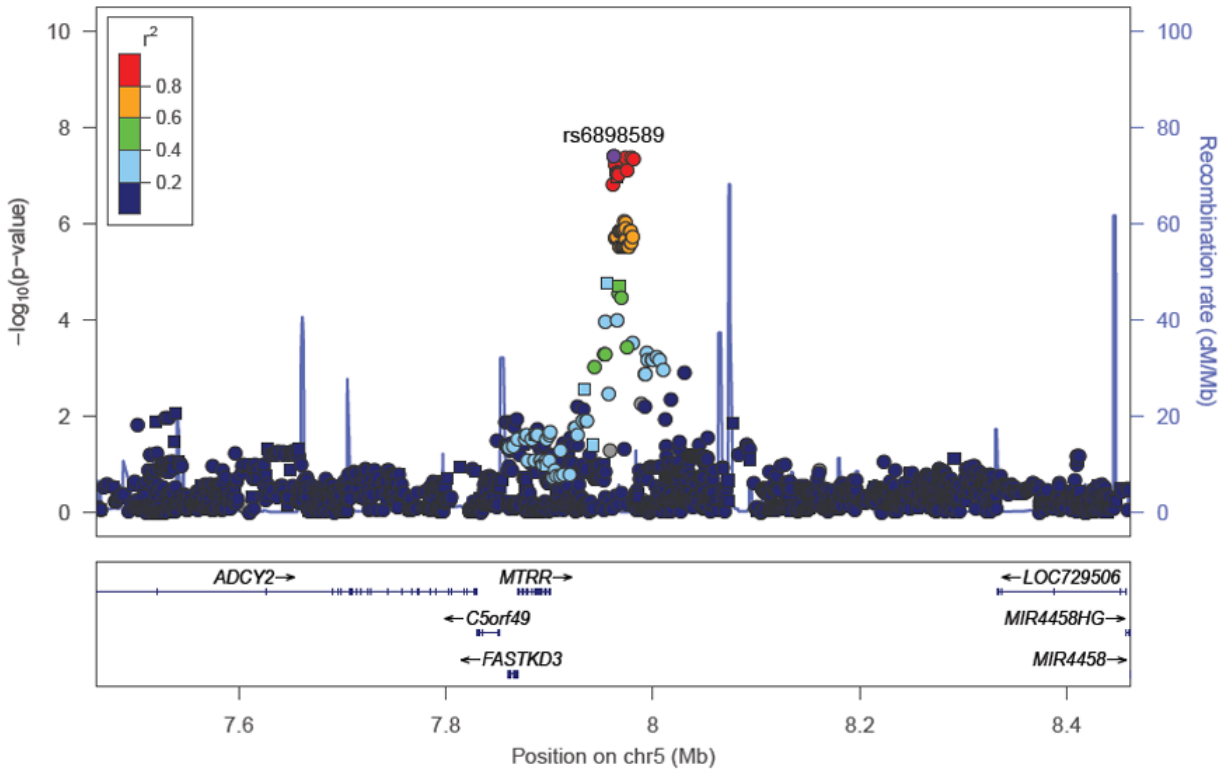


Figure 25. Regional Association Plot of rs6898589

rs3930612 (p = 4.25E-6) is intronic to *NRXN1* (Figure 26). *NRXN1* encodes neurexin 1, a member of the neurexin family of cell surface receptors that function in cell adhesion as well as synapse formation and neurotransmission in the CNS [159]. Deletions of exons of *NRXN1* are associated with autism spectrum disorders and schizophrenia [160]. *NRXN1* was implicated in a GWAS of severe gingival inflammation at suggestive significance (p = 4.0E-6) [161], and is differentially expressed between the dental pulp and PDL [162]; both of these findings indicate a role of *NRXN1* in periodontal health. *NRXN1* was also implicated in GWAS of educational attainment (p = 2E-8 [163] and 5E-6 [164]). The association with educational attainment is interesting as education is a well established predictor of tooth loss, as discussed previously. The

association observed in this meta-analysis may reflect the association between education and tooth loss; alternatively, this may reflect a shared genetic influence of *NRXN1* on both tooth loss and education.

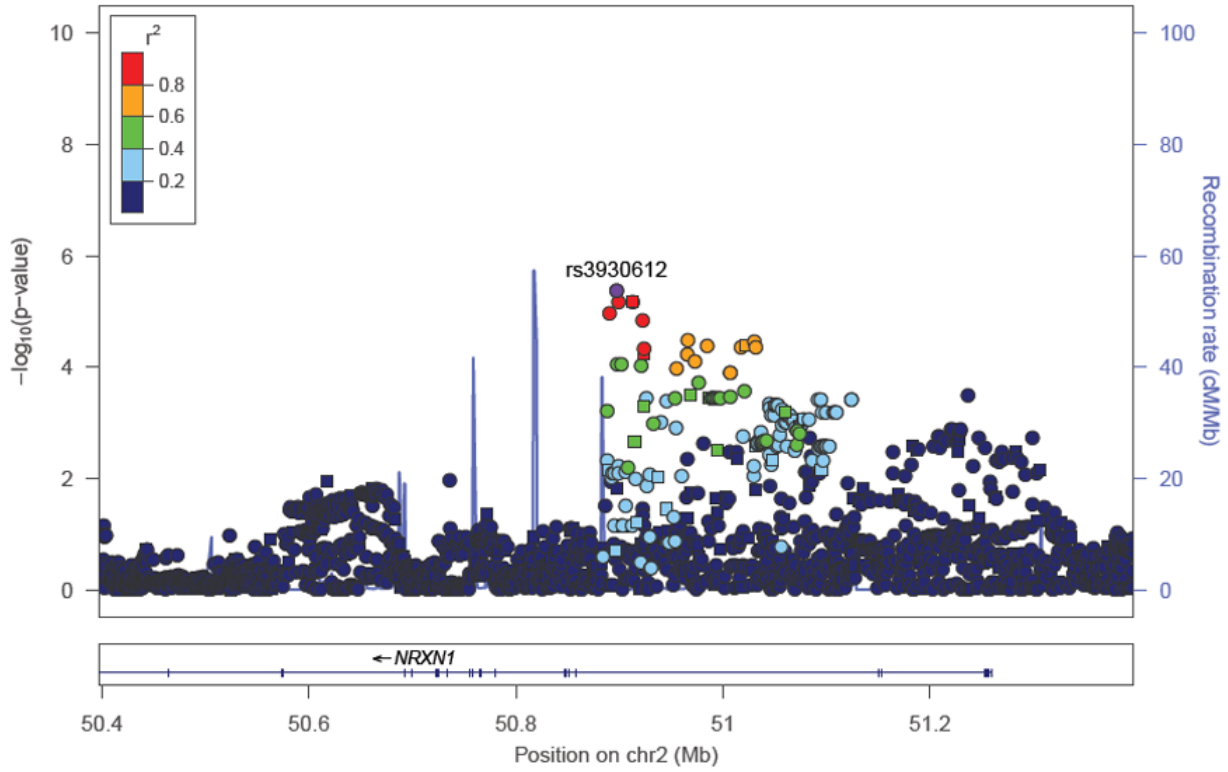


Figure 26. Regional Association Plot of rs3930612

rs1455837 (p = 4.6E-6) is intronic to *IHH*, Indian hedgehog (Figure 27). This SNP is in LD with rs3099 (r^2 = 0.98-99), which lies in 3' UTR and was suggestively significant in this meta-analysis (p = 6.812E-6). Indian hedgehog (Ihh) is a Hedgehog signaling molecule that regulates cell proliferation and differentiation during development and morphogenesis, and it regulates endochondral bone formation and ossification [165]. Missense mutations in the N-terminal active fragment of IHH cause Brachydactyly A-1, and dental anomalies (supernumerary teeth and tooth agenesis) have been reported in some of these patients [166]. Altered Ihh

signaling is associated with temporomandibular joint (TMJ) degeneration, which is partially attributable to occlusal dysfunction that may be caused by missing teeth [167]. In addition to *IHH*, this signal is also 200 kb downstream of *WNT6*, which is differentially expressed throughout tooth development in the developing epithelial layers and enamel structures [168, 169] and promotes differentiation and mineralization of human dental papilla cells [170]. Furthermore, *WNT6* may be involved in dental repair as it has been shown to promote wound healing and induce the migration and differentiation of dental pulp cells [171]. This region also contains *WNT10A*, which, as previously discussed, is one of the major genes implicated in tooth agenesis [84, 85] and is similarly differentially expressed throughout tooth development [169]. Both of these Wnts are clearly candidates for further study for any effect on tooth loss. Also in this region is *CYP27A1*, approximately 250 kb upstream of rs1455837. *CYP27A1* is a vitamin D hydroxylase expressed in gingival fibroblasts and PDL cells, and expression is also induced by inflammatory stimuli (interleukin-1β and *P. gingivalis* lipopolysaccharide) [172]. Both circulating vitamin D and interleukin-1β levels decrease in periodontal tissue after periodontal therapy, indicating a possible role in periodontal inflammation [173]. Thus, *CYP27A1* may influence periodontal tooth loss because of its activation by inflammatory stimuli and role in vitamin D metabolism.

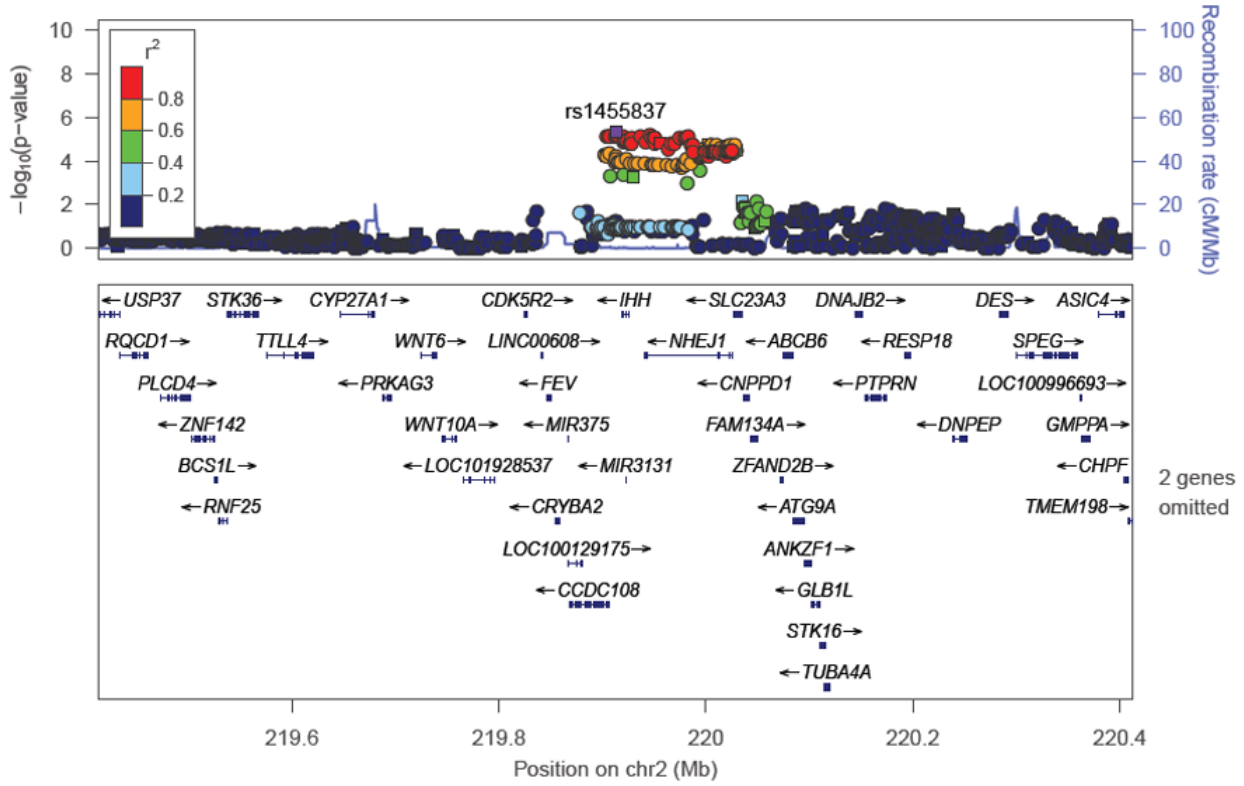


Figure 27. Regional Association Plot of rs1455837

Other SNPs reaching suggestive significance near less interesting genes as well as those not located near genes relevant to dental or oral health are described in the appendix, along with their Regional Association plots.

3.4.3 Hispanics-Only Meta-Analyses

A p-value based meta-analysis was performed for both the quantitative trait and functional dentition to combine GWAS results from POFC-G and POFC-PA, which represent the Hispanic ancestry samples. Manhattan and QQ-plots are shown in Figures 28 and 30. Genomic inflation factor was $\lambda = 0.98$ and $\lambda = 0.95$ for the quantitative trait and functional dentition, respectively, both of which indicate deflation. While no regions of the genome reached genome-wide

significance, several regions showed suggestive significance ($p < 10\text{-}5$). These results are summarized in Tables 13 and 14, including a brief description of nearby genes with potential roles in oral health and tooth loss. Regional Association plots for these results are shown in Figures 29 and 31. Due to the small sample size in this meta-analysis and lack of genome-wide significant findings, these suggestively significant loci should be interpreted with caution.

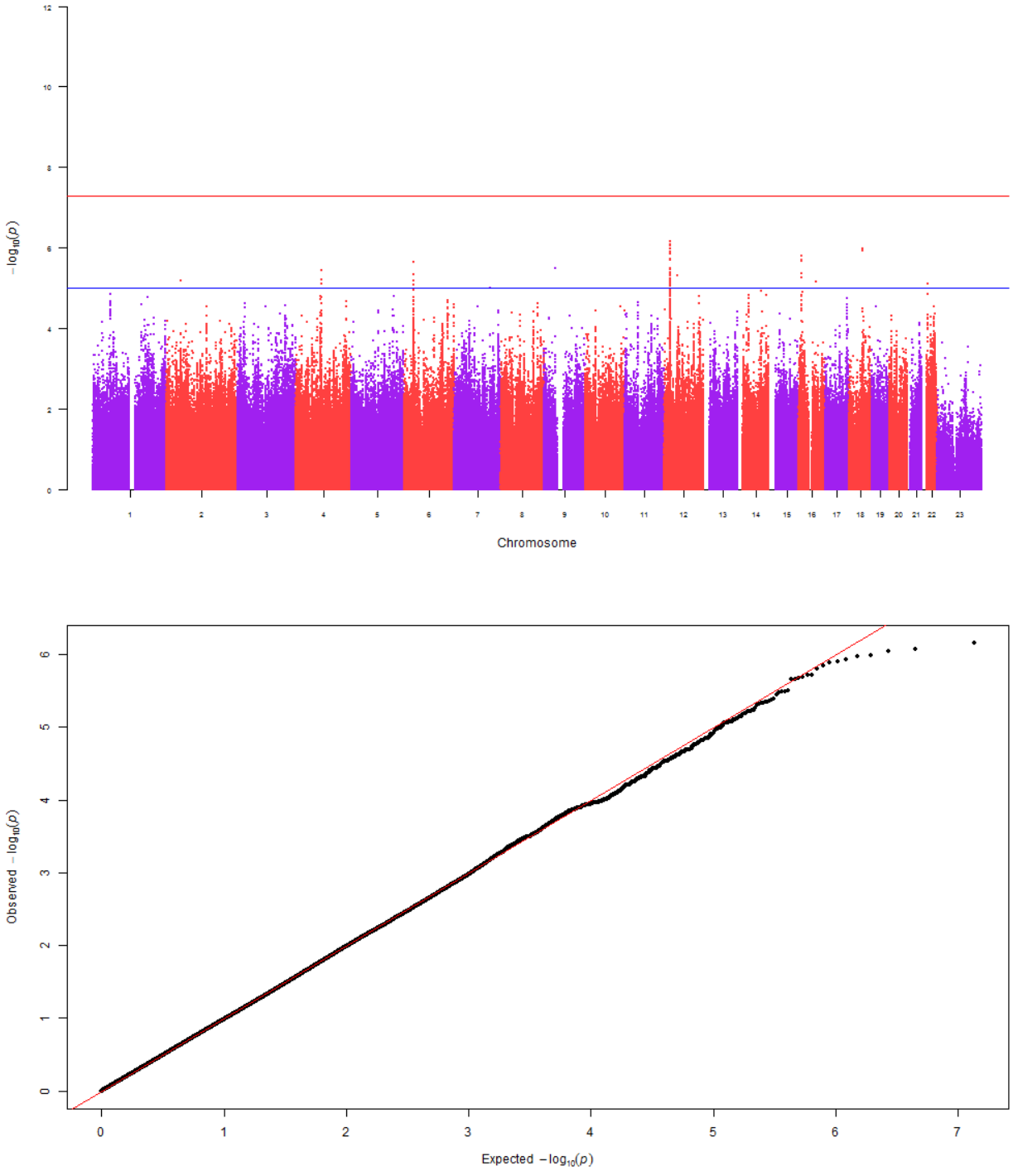


Figure 28. Manhattan and QQ plots for the Hispanics-only meta-analysis of the quantitative trait

Table 13. Top hits for the Hispanics-only meta-analysis of the quantitative trait

SNP	CHR	BP	EA†	MAF POFC-G	MAF POFC-PA	N	P	DIR‡	Gene and Possible Role in Dental/Oral Health
rs3908116	2	51040820	A	0.14	0.18	406	6.45E-06	--	*NRXN1:* Hits in the whites-only and trans-ethnic meta-analyses for functional dentition
rs12506662	4	87861476	A	0.24	0.28	451	3.61E-06	--	*AFF1:* Unknown *MAPK10:* Part of the MAPK signaling pathway, which is involved in immune response and inflammation [174]
NA*	6	32472247	T	0.28	0.12	316	2.20E-06	--	*HLA* genes : HLA proteins have been implicated in periodontal disease [175] and dental caries [176]
rs7795775	7	122540202	C	0.44	0.43	454	9.60E-06	--	*CADPS2:* Unknown *TAR2S16:* Taste receptor for bitterness [177]
rs72719502	9	38701150	C	0.09	0.16	426	3.22E-06	++	*CNTNAP3 :* Member of NCP family of neurexins [178]; neurexin1 was implicated in the meta-analysis for whites
rs67639374	12	18931236	A	0.40	0.21	437	6.95E-07	++	*PLCZ1:* Unknown

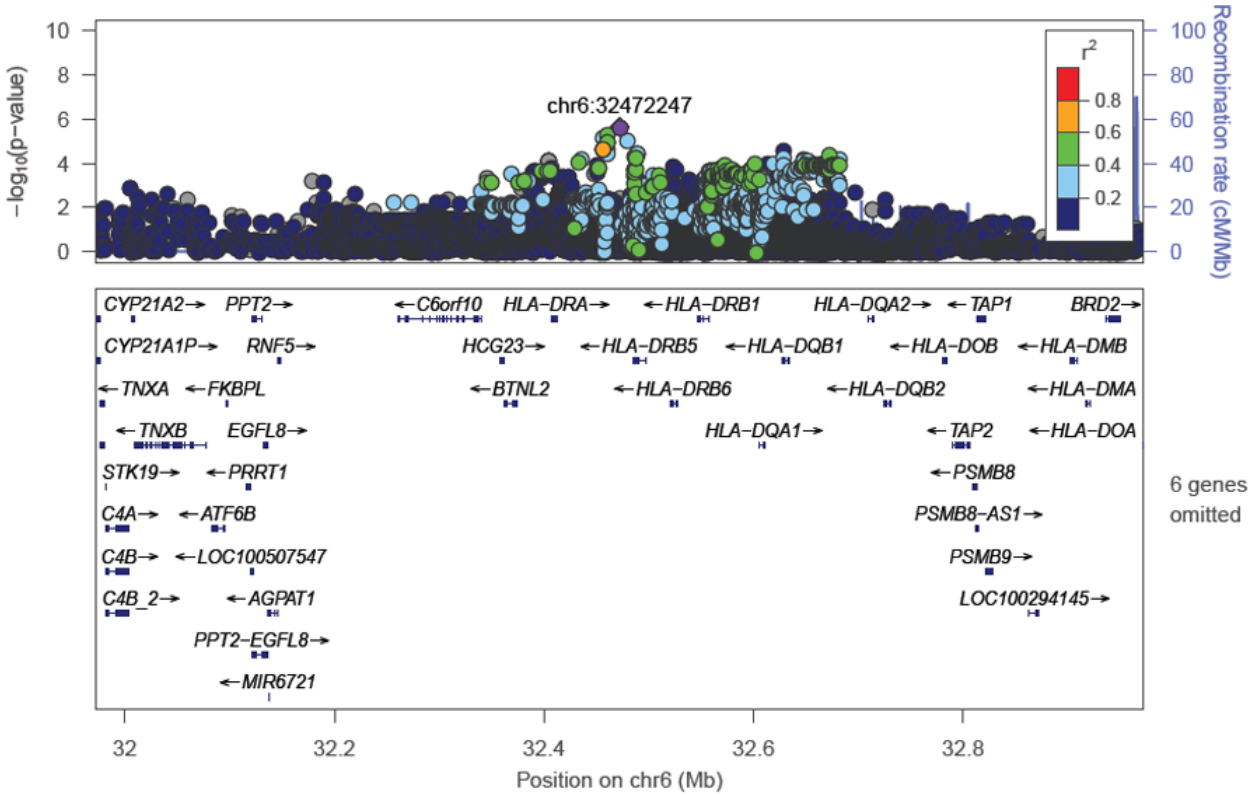
Table 13 Continued

rs138380112	12	44062693	A	0.46	0.45	438	4.91E-06	++	*IRAK4 :* Activates NF-κB to regulate immune and inflammatory response [179]; NF-κB is implicated in periodontal disease [180]
rs34806537*	16	9325529	C	0.15	0.08	435	1.56E-06	--	*CARHSP1:* Regulates TNF-α production by stabilizing TNF-α mRNA [181]; TNF-α is a critical moderator of inflammation, immune response, and periodontal disease progression [182]
rs76720124*	16	57390538	C	0.03	0.03	421	7.05E-06	--	*CCL22:* Cytokine that is increased expression in gingiva of periodontal disease (PD) patients and after induction of PD in mice [183]; loss of *CCL22* in mice lead to increased PD phenotype [184]
									CCL17: Increased expression in gingiva of PD patients [183] and plays a role in periodontal inflammation[185]

Table 13 Continued

rs1441296	18	47956586	C	0.48	0.44	329	1.03E-06	--	*SKA1:* Unknown *MAPK4:* Part of the MAPK signaling pathway [174]
rs405011	22	20164696	C	0.13	0.23	447	7.98E-06	--	*TBX1:* Critical regulator of enamel formation [186-188]

* Regional Association plot included below; all other plots are located in the appendix; †Effect Allele; ‡Direction of effect in POFC-G and POFC-PA



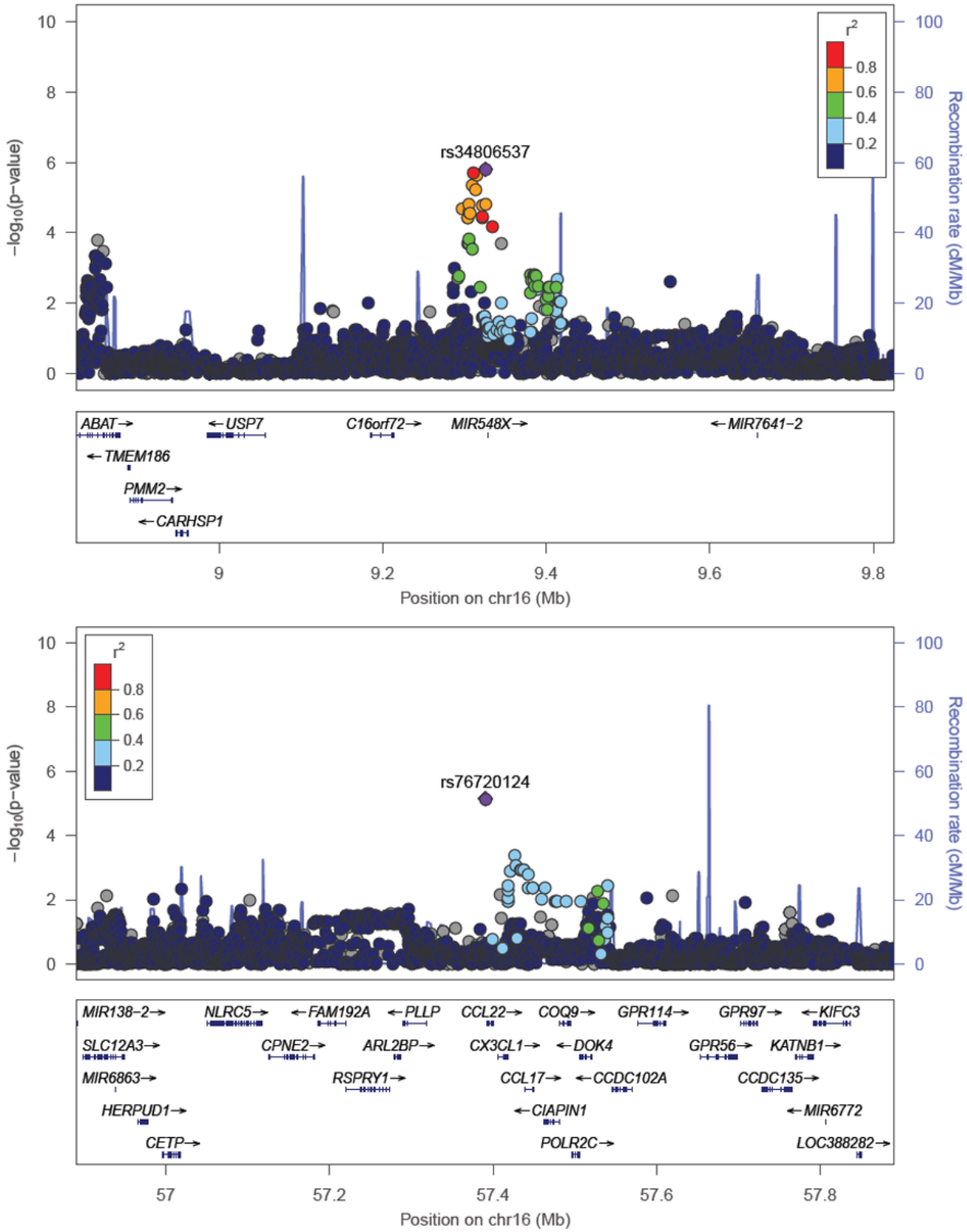


Figure 29. Regional Association Plots for hits in the Hispanics-only meta-analysis of the quantitative trait

LD information was not available for all plots, and annotation of imputed vs. genotypes variants was unavailable.

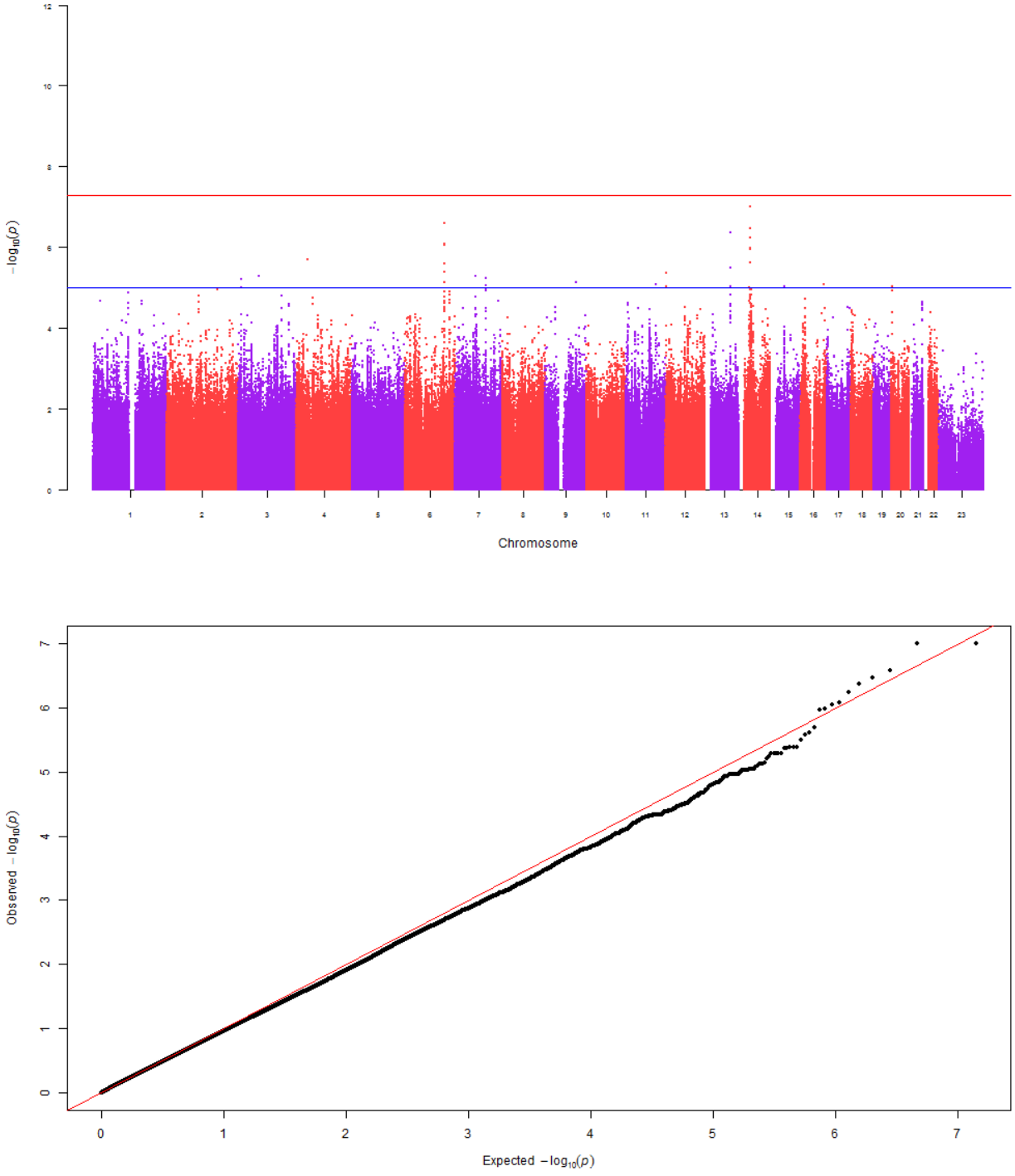


Figure 30. Manhattan and QQ plots for the Hispanics-only meta-analysis of functional dentition

Table 14. Top hits for the Hispanics-only meta-analysis of functional dentition

SNP	CHR	BP	EA†	MAF POFC-G	MAF POFC-PA	N	P	DIR‡	Possible Role in Dental/Oral Health
rs139769216*	3	12649936	C	0.09	0.06	454	6.05E-06	++	*RAF1:* Differentially expressed in developing dental tissues throughout tooth development[189]; Part of Ras/MAPK signaling pathway that regulates cell proliferation/differentiation in tooth development [190]
									TIMP4: Elevated expression in gingival tissue of patients with chronic periodontitis [191]; Induction of TIMP-4 improved periodontitis symptoms in arthritic rats [192]
rs11304053	3	71294110	GA	0.28	0.33	454	5.15E-06	--	*FOXP1:* Also a GWAS hit for lymphocyte count (p = 4 x10-20)[193]
rs59987264	4	39454916	T	0.33	0.39	454	1.99E-06	--	*WDR19:* Mutations are associated with cranioectodermal dysplasia, manifestations of which include dental anomalies (hypodontia, enamel defects, delayed tooth eruption) [194]
rs142286663	6	134788699	A	0.22	0.19	454	2.56E-07	++	*SGK1:* Part of anti-inflammatory pathway that inhibits Toll-like receptor mediated inflammation [195]
rs10085380	7	68661978	C	0.49	0.48	454	5.15E-06	++	*AUTS2 :* Unknown
rs10256805	7	104556251	A	0.37	0.42	454	5.62E-06	++	*LHFPL3-AS2:* Unknown
NA	9	107399986	C	0.04	0.12	454	7.12E-06	--	*ABCA1 :* Increased in expression after oxysterol-induced osteogenic differentiation of PDL stem cells during periodontal regeneration [196]

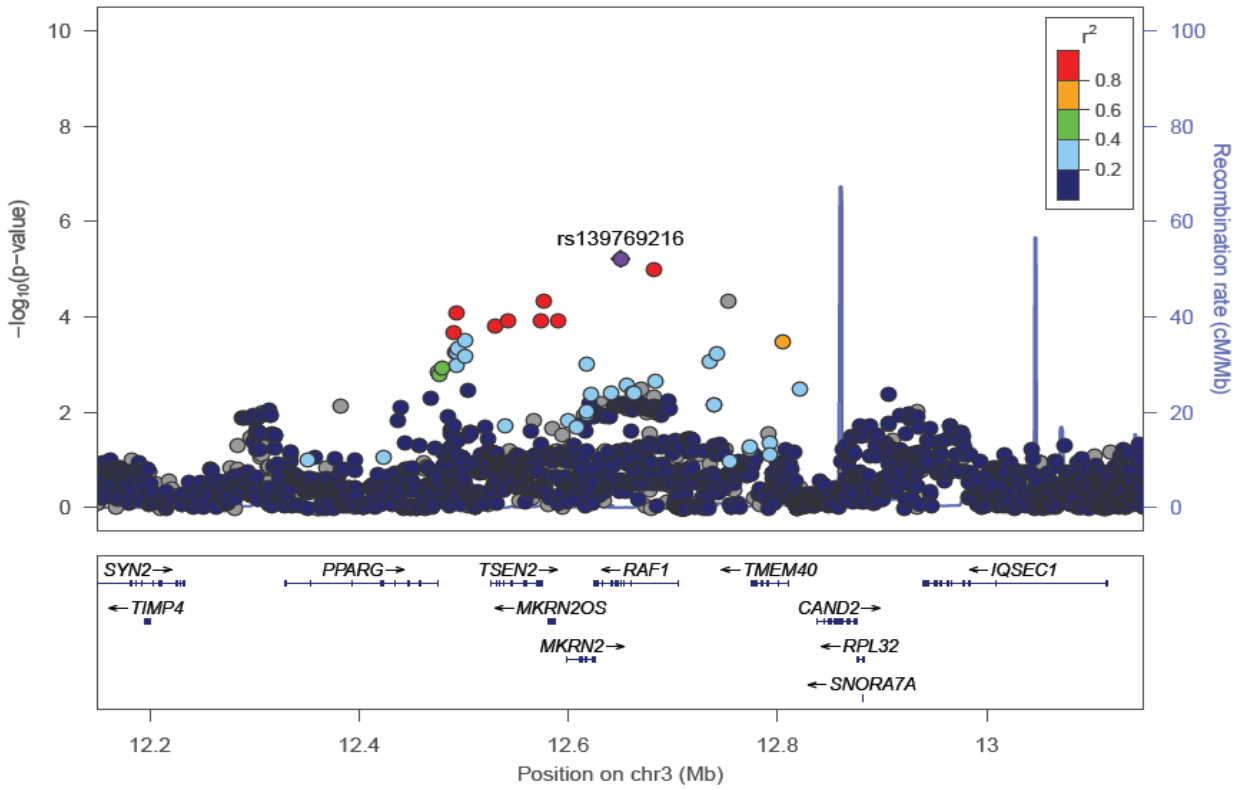
Table 14 Continued

rs72507759	11	101025134	T	0.13	0.07	454	8.03E-06	++	*TRPC6:* Calcium channel protein required for odontogenic differentiation of human dental pulp cells [197]
rs61916630	12	616069	C	0.06	0.04	454	4.22E-06	--	*B4GALNT3 :* Unknown
									12p13.3 deletion syndrome includes dental anomalies, including malocclusion [198]
NA	13	86689246	T	0.36	0.36	454	4.26E-07	--	*SLITRK6:* Expressed in developing tooth tissues [199]
rs8009351*	14	37178427	C	0.18	0.06	454	9.99E-06	++	*PAX9:* Critical for proper tooth development as mutations are implicated in tooth agenesis, including agenesis of >5 teeth [85]
									NKX2-1: Knockout mice show defects in tooth development and morphology [200]
rs139066701	14	39797524	C	0.06	0.05	454	9.76E-08	--	*CTAGE5:* Unknown
									SEC23A : Targeted by TFII-I transcription factors (TFs); the genes encoding these TFs are deleted in Williams Syndrome, features of which include dental and craniofacial anomalies [201]
rs12900666	15	49256375	A	0.17	0.15	454	9.23E-06	--	*FGF7:* Exhibits specific expression patterns in developing dental tissues [202, 203]
									FBN1: FBN2 was implicated in whites-only and trans-ethnic meta-analyses of the quantitative trait; exhibits specific expression patterns in the PDL [158]
rs30387*	16	79612121	C	0.05	0.09	454	8.34E-06	--	*MAF* : Transcription factor expressed in ameloblasts [204]; implicated in GWAS of dental caries in the permanent dentition (p = 5.2E-6) [60]

Table 14 Continued

rs6139511	20	4636334	A	0.14	0.07	454	9.26E-06	++	*PRNP* : Exhibits specific expression patterns during tooth development and may regulate tooth development [205]
									SLC23A2: Intragenic variants may be associated with aggressive periodontitis [206]

* Regional Association plot included below; all other plots are located in the appendix; †Effect Allele; ‡Direction of effect in POFC-G and POFC-PA



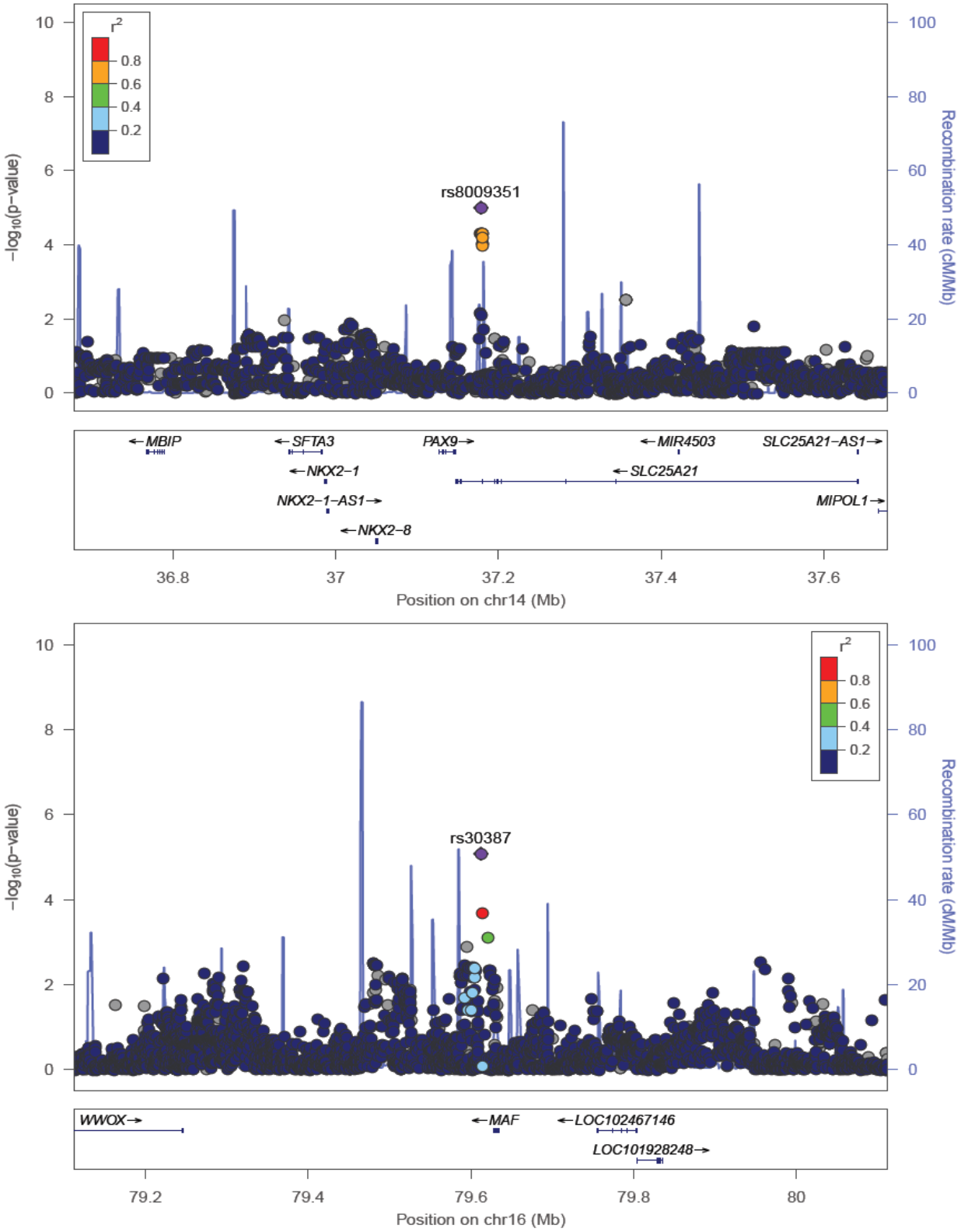


Figure 31. Regional Association Plots for hits in the Hispanics-only meta-analysis of functional dentition

Note: LD information was not available for all plots. Annotation of imputed vs. genotypes variants was unavailable in POFC-G and POFC-PA.

3.4.4 Trans-Ethnic Meta-Analysis – Quantitative trait

A p-value based meta-analysis was performed to combine GWAS results from four study samples (COHRA, DRDR, POFC-G, and POFC-PA) for the quantitative trait. Manhattan and QQ-plots are shown in Figure 32. Genomic inflation factor was $\lambda = 1.03$. While no regions of the genome reached genome-wide significance, several regions showed suggestive significance ($p < 10^{-5}$). Meta-analysis results are shown in Table 15, and allele frequencies in all 4 cohorts are shown in Table 16.

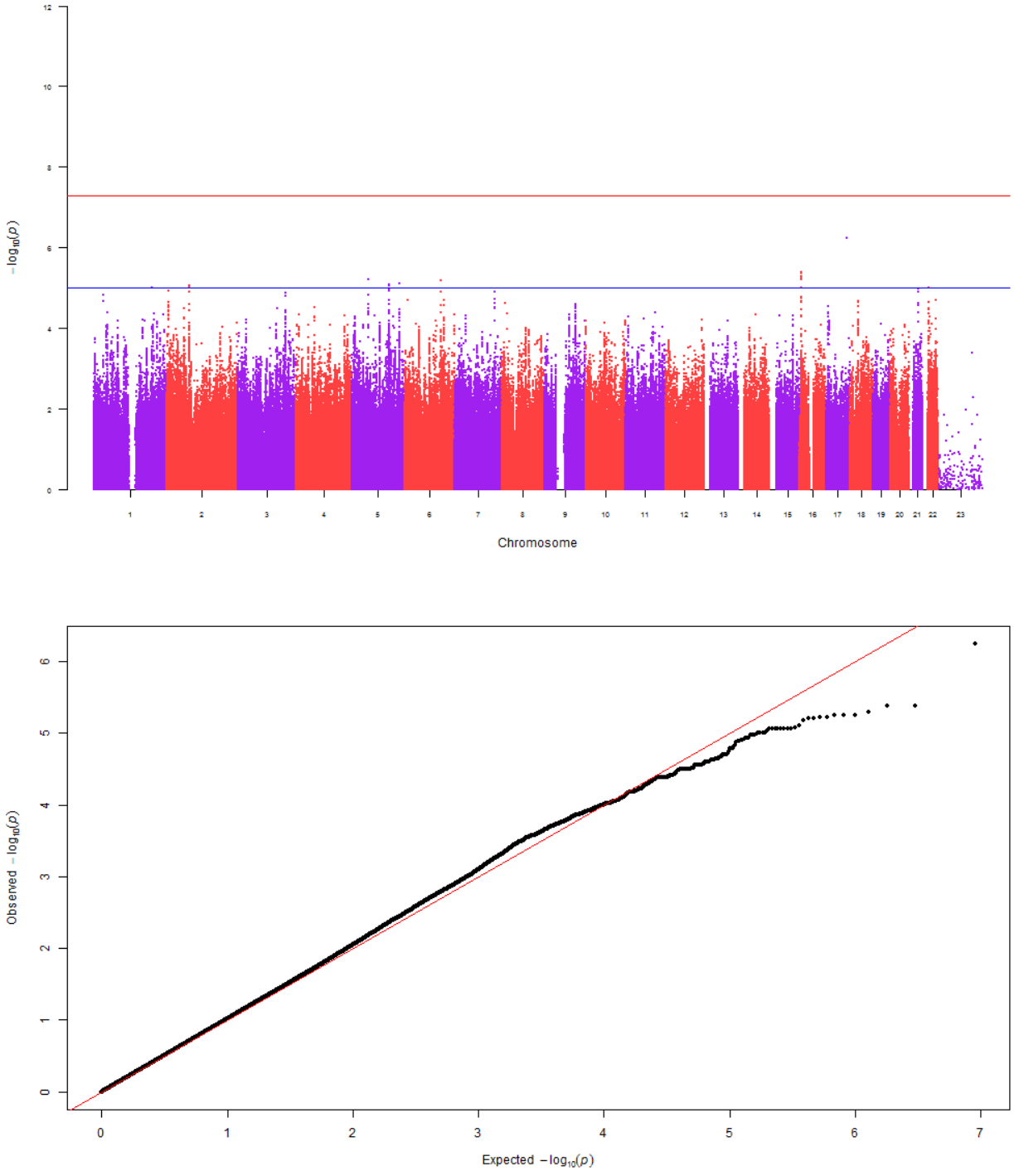


Figure 32. Manhattan and QQ plots for the trans-ethnic meta-analysis of the quantitative trait

Table 15. Top hits for the trans-ethnic meta-analysis of the quantitative trait

SNP	CHR	BP	N	P	Type (in COHRA)	DIR*
rs6663322	1	200649869	1630	9.79E-06	imputed	++++
rs72488321	2	76805700	1574	8.55E-06	imputed	----
rs764629	5	57455043	1608	5.97E-06	imputed	++++
rs10062700	5	128180057	1611	8.34E-06	imputed	---+
rs2860807	5	163265540	1632	7.92E-06	imputed	----
rs4569988	6	123862050	1588	6.68E-06	imputed	----
rs2532011	16	4130204	1584	4.16E-06	imputed	++++
rs4969040	17	70947492	1327	5.66E-07	genotyped	--+-
rs17208994	22	19373861	1631	9.85E-06	imputed	----

*Direction of effect in COHRA, DRDR, POFC-G, and POFC-PA

Table 16. Allele frequencies for the top hits of the trans-ethnic meta-analysis of the quantitative trait

SNP	CHR	BP	A1/A2	MAF COHRA	MAF DRDR	MAF POFC-G	MAF POFC-PA
rs6663322	1	200649869	a/g	0.32	0.30	0.47	0.46
rs72488321	2	76805700	g/t	0.28	0.23	0.25	0.26
rs764629	5	57455043	g/t	0.45	0.47	0.38	0.4
rs10062700	5	128180057	a/c	0.15	0.25	0.36	0.31
rs2860807	5	163265540	a/t	0.29	0.30	0.24	0.34
rs4569988	6	123862050	c/t	0.23	0.39	0.45	0.42
rs2532011	16	4130204	c/t	0.41	0.34	0.41	0.36
rs4969040	17	70947492	a/g	0.08	0.07	0.09	0.17
rs17208994	22	19373861	a/c	0.06	0.05	0.06	0.04

rs4969040 in *SLC39A11* and near *SSTR2* (Figure 33) remained significant in the trans-ethnic analysis (p = 5.7E-7), though this is mostly due to its significance in COHRA and DRDR, as $p_{POFC\text{-}G} = 0.24$ and $p_{POFC\text{-}PA} = 0.05$. *SLC39A11* is a zinc transporter [146], the concentration of which may have important consequences for enamel mineralization [148, 149]. *SSTR2* is a somatostatin receptor that is involved in nociception and orofacial pain [150-152].

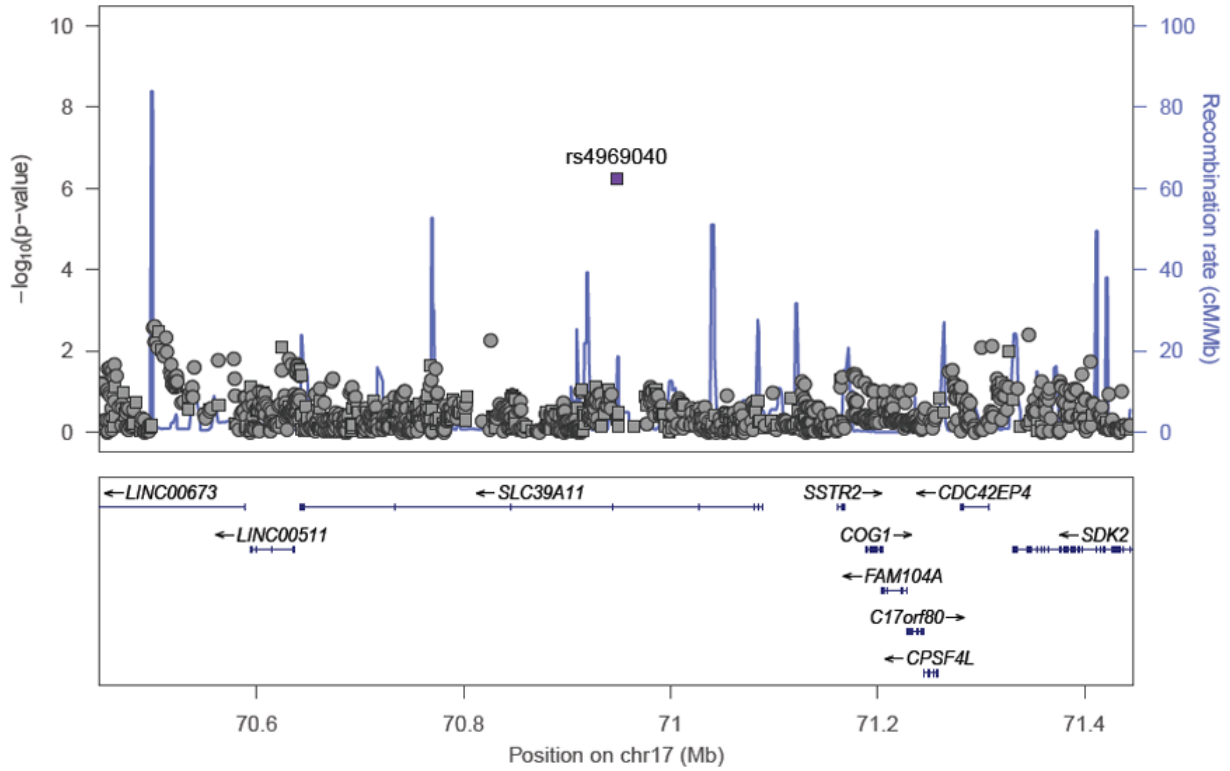


Figure 33. Regional Association Plot of rs4969040

For all plots in the trans-ethic meta-analysis, LD information was not included on the plot as it represents a mixed sample. Annotation of imputed vs. genotypes variants was based on COHRA and DRDR.

rs2532011 (p = 4.2E-6) lies in a region on chromosome 16 with three genes possibly involved in tooth loss (Figure 34). This SNP is intronic to *ADCY9*, another adenylate cyclase; *ADCY2* was implicated in both the COHRA only GWAS and the whites-only meta-analysis of functional dentition. Adenylyl cyclases catalyze the formation of cyclic adenosine monophosphate (cAMP), and the cAMP signaling pathway mediates cellular response to mechanical stress [127]. cAMP signaling is induced in response to mechanical force on PDL cells [128, 129], decreased cAMP signaling in PDL cells has been associated with aggressive periodontitis [130], and miRNAs involved in cAMP signaling are increased in expression in response to *P. gingivalis* lipopolysaccharides [131].

rs2532011 is also located approximately 200 kb upstream of *CREBBP*; mutations in *CREBBP* cause Rubinstein-Taybi syndrome, and manifestations include dental anomalies such as malocclusion, dental caries, and hypodontia [207]. Furthermore, CREBBP binds to cAMP-response element binding protein (CREB), a transcription factor that initiates gene expression in response to cAMP signaling [208]. As this gene is involved in response to cAMP signaling, it may have implications for tooth loss akin to those of *ADCY2* and *ADCY9*. *CREB* is involved in tooth mineralization as it is expressed in the nucleus of odontoblasts, cementoblasts, dental pulp, and PDL fibroblasts, and is phosphorylated, or activated, in the nucleus of molar odontoblasts and cementoblasts [209]. *CREBBP* may influence tooth loss through its interaction with CREB.

Lastly, this region also contains *TRAP1*, approximately 400 kb downstream of rs2532011. *TRAP1* encodes tumor necrosis factor (TNF) receptor associated protein 1. TNF-α is a key mediator of inflammation seen in periodontitis, and it may induce tissue destruction, attachment loss, and bone loss [210]. Levels of TNF-α are associated with periodontal disease status, with the higher levels of expression in gingival crevicular fluid of patients with more severe disease [211, 212]. Expression of TNF-α is associated with infection by *P. intermedia* in gingival tissues of periodontitis patients, which may increase clinical disease measurements and disease progression [182].

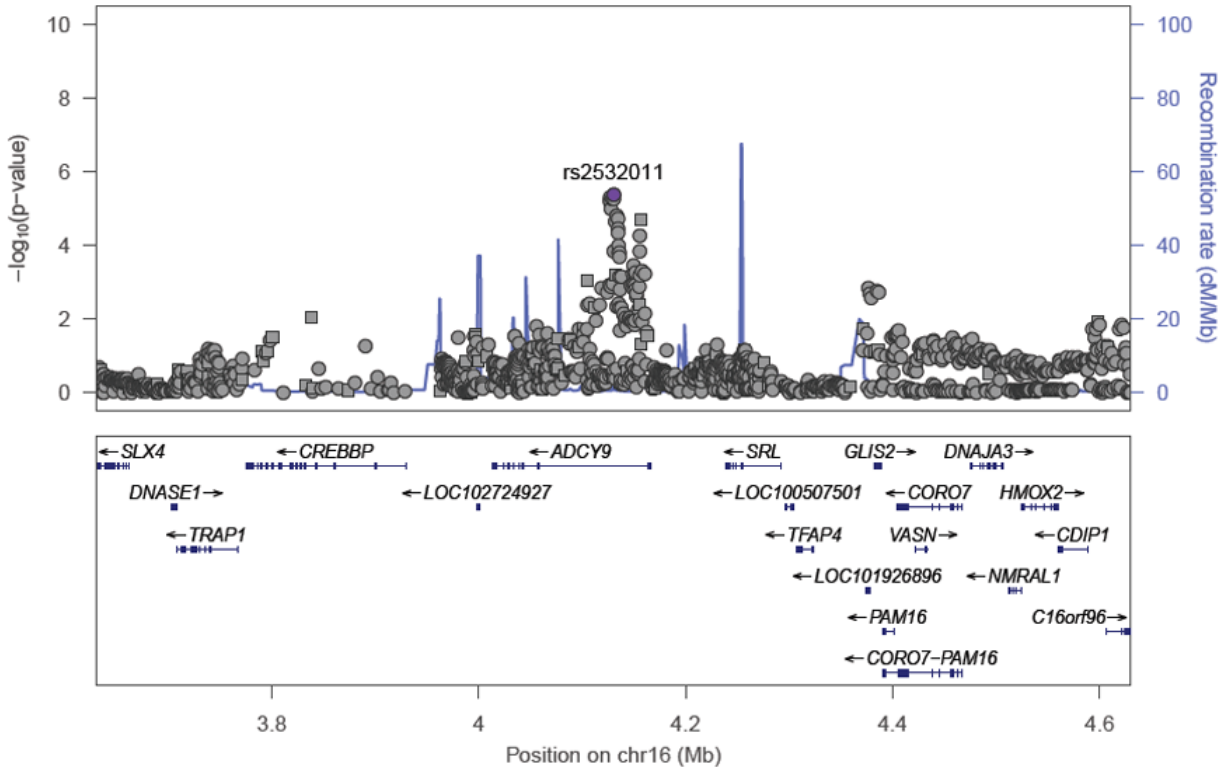


Figure 34. Regional Association Plot of rs2532011

rs4569988 (p = 6.7E-6) is intronic to *TRDN* (Figure 35), triadin, and this index SNP is in moderate LD (r^2 = 0.66) with a missense variant in the third exon of *TRDN* (p.Thr128Ser [c.383C>G]; rs9490809). However, this missense variant is predicted to be benign by ClinVar and is at a relatively high frequency in both European and Amerindian populations (MAF = 0.52 and MAF = 0.38, respectively; 1000 Genomes Phase I.) Triadins are sarcoplasmic reticulum transmembrane proteins that organize microtubules in muscle cells and participate in excitation contraction coupling by regulating cellular Ca^{2+} levels [213, 214]. Deletion of triadin results in loss of muscle strength in mice [214], and loss of function mutations in triadin are responsible for cardiac arrhythmias with sudden death in humans [215]. It is currently unknown how triadin may affect dental and oral health.

This signal is also approximately 200 kb upstream of *NKAIN2*, which was implicated in a GWAS of chronic periodontitis (p = 8E-7)[67] as well of rate of cognitive decline in Alzheimer's disease (p = 6E-7)[216]. *NKAIN2*, Na^+/K^+ transporting ATPase interacting 2, resides in a chromosomal region commonly deleted in cancers, indicating its possible role as a tumor suppressor gene, and it is also involved in nervous system development [217]. Functional studies of *NKAIN2* in prostate cancer cells indicate that it promotes apoptosis to control cell growth and migration, indicating a regulatory role in the cell cycle [218]. It is unclear how *NKAIN2* may affect dental health, but the suggestively significant hit in the GWAS of chronic periodontitis supports the association observed here for missing teeth.

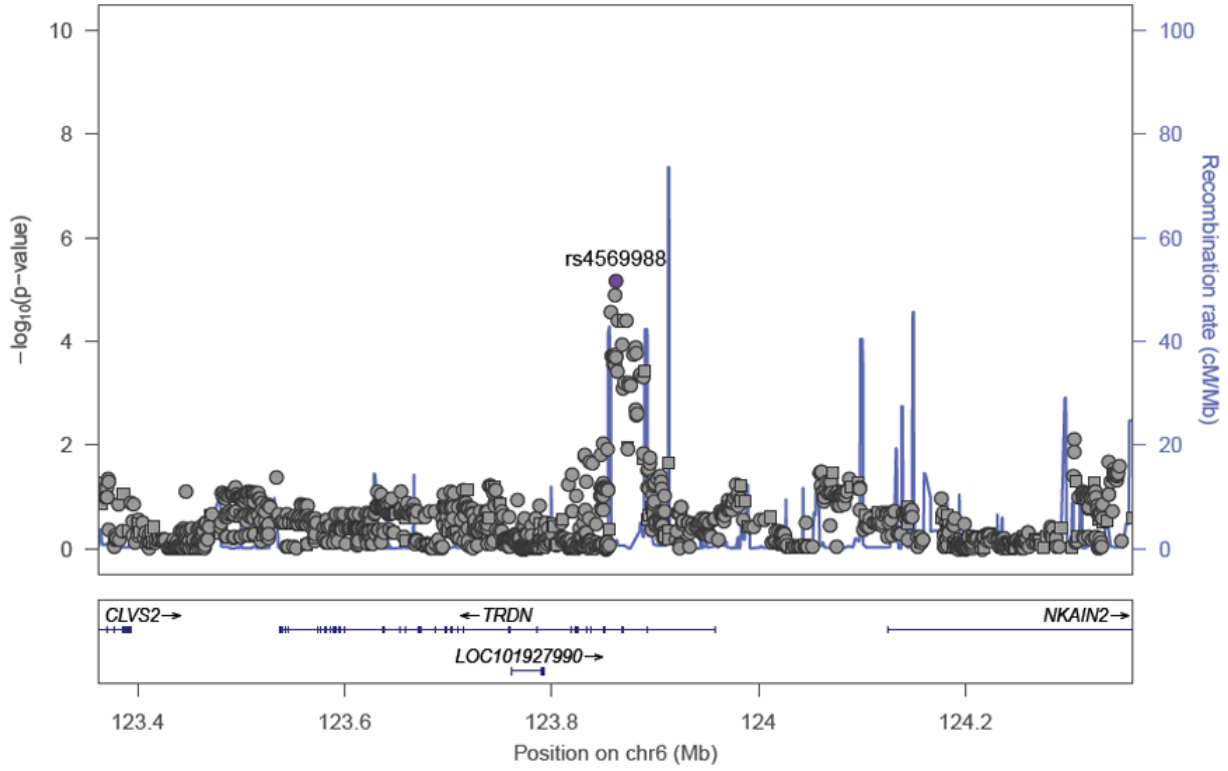


Figure 35. Regional Association Plot of rs4569988

rs17208994 (p = 9.9E-6) is intronic to *HIRA* and is located in the region deleted in 22q11.2 deletion syndrome (Figure 36). Manifestations of 22q11.2 deletion syndrome include palatal abnormalities, characteristic facial features, and dental anomalies, including enamel hypoplasia and hypomineralization, hypodontia, delayed tooth eruption, abnormal tooth shape, gingivitis, and excessive dental caries [219, 220]. rs17208994 also overlaps a region of suggestive significance for Alzheimer's Disease age-of-onset ($p = 5 \times 10^{-7}$) [221], which may reflect the relationship between tooth loss and cognitive decline.

While there is no known role of *HIRA* in dental health, this region of significance is located approximately 400 kb upstream of *TBX1*, which encodes an evolutionarily conserved transcription factor with a T-box DNA binding domain [222]. 22q11.2 deletion syndrome is associated with increased susceptibility to infection, and *TBX1* mutations are associated with low T cell counts, indicating a role of *TBX1* in immunity and immunodeficiency [223], though no relationship between *TBX1* and periodontal disease has been reported. As part of a negative feedback loop including microRNA-96 and PITX2, *Tbx1* plays a critical role in the regulation of dental epithelial cell proliferation and differentiation during tooth development in adult mice, and levels of *Tbx1* expression affect tooth and cusp morphology, ameloblast differentiation, and enamel production [186, 188]. Knockdown of *Tbx1* in mice resulted in decreased production of amelogenin, the major component of the enamel matrix, and subsequent enamel defects [186, 188]. *Tbx1* null mice showed a decrease in tooth mineralization, an absence of ameloblasts during tooth development, decreased amelogenin gene expression, and a lack of enamel on adult incisors [187]. *TBX1* may influence missing teeth as a critical regulator of enamel formation and tooth development.

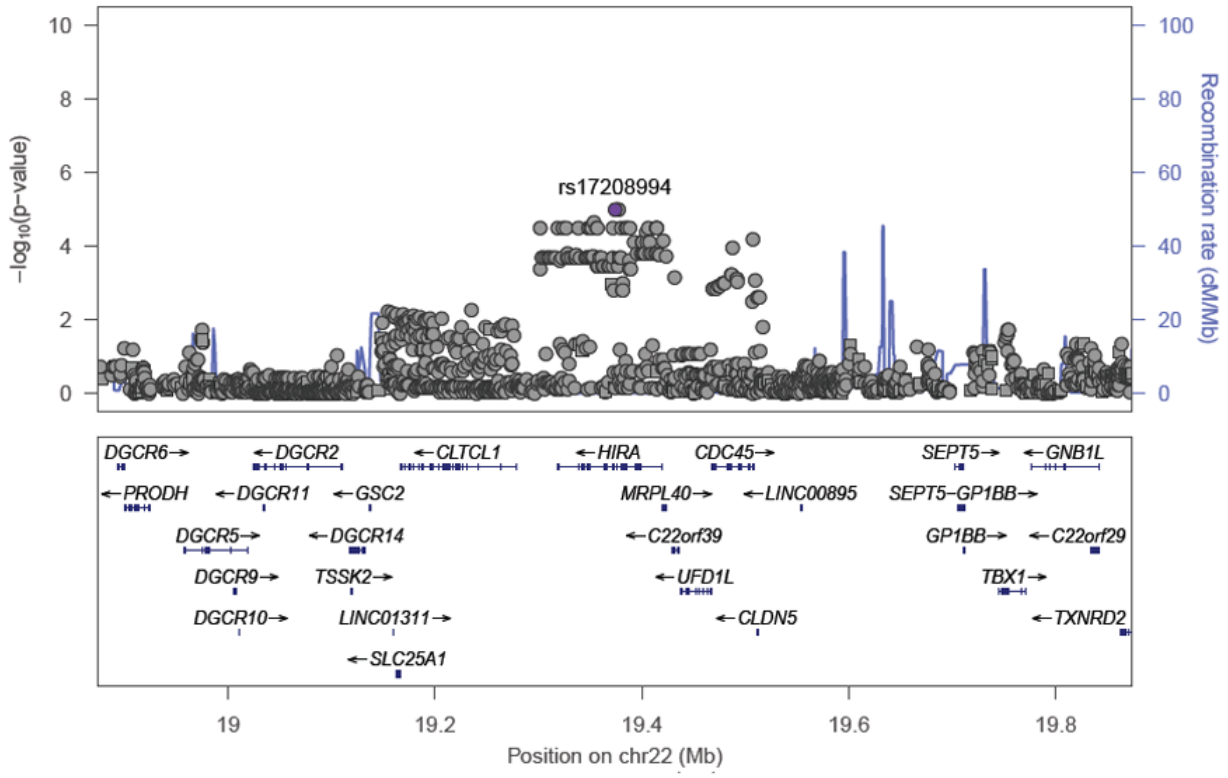


Figure 36. Regional Association Plot of rs17208994

Other SNPs reaching suggestive significance near less interesting genes as well as those not located near genes relevant to dental or oral health are described in the appendix, along with their Regional Association plots.

3.4.5 Trans-Ethnic Meta-Analysis – Functional dentition

A p-value based meta-analysis was performed to combine GWAS results from four study samples (COHRA, DRDR, POFC-G, and POFC-PA) for functional dentition. Manhattan and QQ-plots are shown in Figure 37. Genomic inflation factor was $\lambda = 1.00$. While no regions of the genome reached genome-wide significance, several regions showed suggestive significance ($p < 10\text{-}5$.) Meta-analysis results are shown in Table 17, and allele frequencies in Table 18.

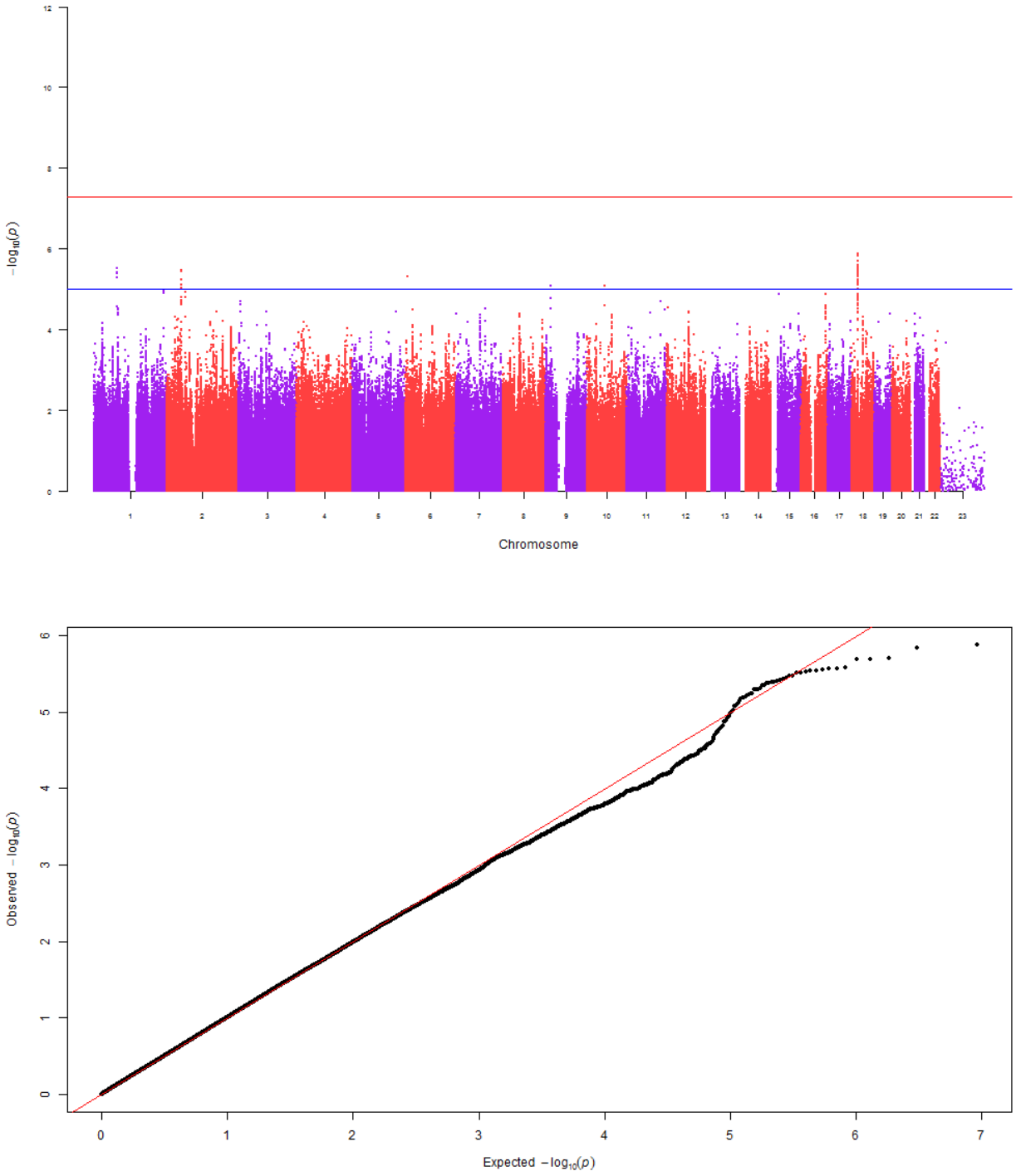


Figure 37. Manhattan and QQ plots for the trans-ethnic meta-analysis of functional dentition

Table 17. Top hits for the trans-ethnic meta-analysis of functional dentition

SNP	CHR	BP	N	P	Type (in COHRA)	DIR*
rs674271	1	78712402	1607	3.01E-06	imputed	++++
rs3930612	2	50896945	1606	3.34E-06	imputed	----
rs2815822	6	6320808	1636	4.90E-06	genotyped	++++
rs10963759	9	18787638	1636	8.32E-06	genotyped	----
rs28734985	10	60084051	1606	8.39E-06	imputed	----
rs4575613	18	22697408	1634	1.31E-06	imputed	++++

*Direction in COHRA, DRDR, POFC-G, and POFC-PA

Table 18. Allele frequencies for the top hits of the trans-ethnic meta-analysis of functional dentition

SNP	CHR	BP	A1/A2	MAF COHRA	MAF DRDR	MAF POFC-G	MAF POFC-PA
rs674271	1	78712402	a/c	0.26	0.27	0.16	0.26
rs3930612	2	50896945	a/t	0.09	0.16	0.21	0.21
rs2815822	6	6320808	g/t	0.12	0.08	0.10	0.13
rs10963759	9	18787638	c/t	0.44	0.43	0.25	0.27
rs28734985	10	60084051	a/g	0.10	0.08	0.09	0.09
rs4575613	18	22697408	c/g	0.42	0.43	0.42	0.48

rs3930612 (p = 3.3E-6) corresponds to the same signal seen in the whites-only meta-analysis that is intronic to *NRXN1* (Figure 38). This gene is important for proper synapse formation and neurotransmission in the CNS [159] and may play a role in periodontal health as it

was implicated in a GWAS of severe gingival inflammation [161] and is differentially expressed in dental tissues [162]. There is increasing statistical evidence that this gene is involved in tooth loss and may function in periodontal health.

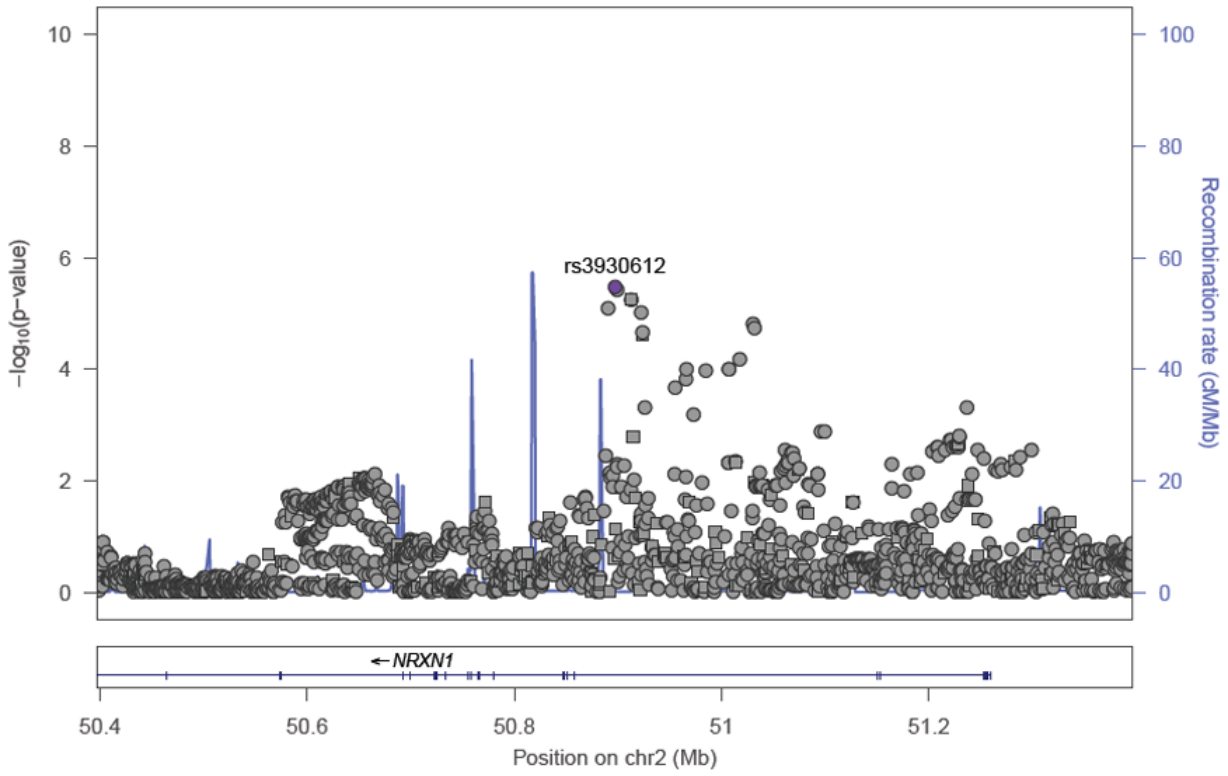


Figure 38. Regional Association Plot of rs3930612

rs2815822 (p = 4.9E-6) is intronic to *F13A1* (Figure 39); this SNP was genotyped in COHRA and DRDR. P-values in each smaller GWA scan were as follows: $p_{COHRA} = 0.05$, $p_{DRDR} = 0.13$, $p_{POFC\text{-}G} = 0.025$, $p_{POFC\text{-}PA} = 0.0599$. Although this plot shows only one SNP, this is in accordance with available LD information in this region, as HaploReg reports LD to be low (no SNPs with $r^2 > 0.5$ in either whites or Hispanics.) *F13A1* encodes the coagulation factor XIII A subunit (FXIIIA), which stabilizes the fibrin clot and prevents fibrinoloysis [224]. Loss of function mutations in either *F13A1* or *F13B* cause factor XIII deficiency (OMIM: 613225), a

rare autosomal recessive disorder resulting in life-long bleeding diathesis and inefficient wound healing [224, 225]. Frequent bleeding in the mouth and gums is a common occurrence reported in patients [224, 226]. This SNP, rs2815822, corresponds to the previously reported Int1(+12)C>A variant, which resides in a regulatory region of intron 1 of *F13A1*; the C>A mutation results in decreased binding affinity of the Sp1 transcription factor, leading to decreased transcription of *F13A1* gene and decreased FXIIIA expression[225]. While the effect of the Int1(+12)A variant on FXIIIA expression was originally ascertained in FXIIIA patients with exon mutations of *F13A1*, this variant is also associated with mild FXIII deficiency in patients with no mutations in either *F13A1* or *F13B* [227]. Decreased FXIIIA levels resulting in oral bleeding could increase risk of infection by periodontal pathogens and periodontal inflammation, and impaired wound healing could affect retention of teeth affected by periodontal disease or trauma.

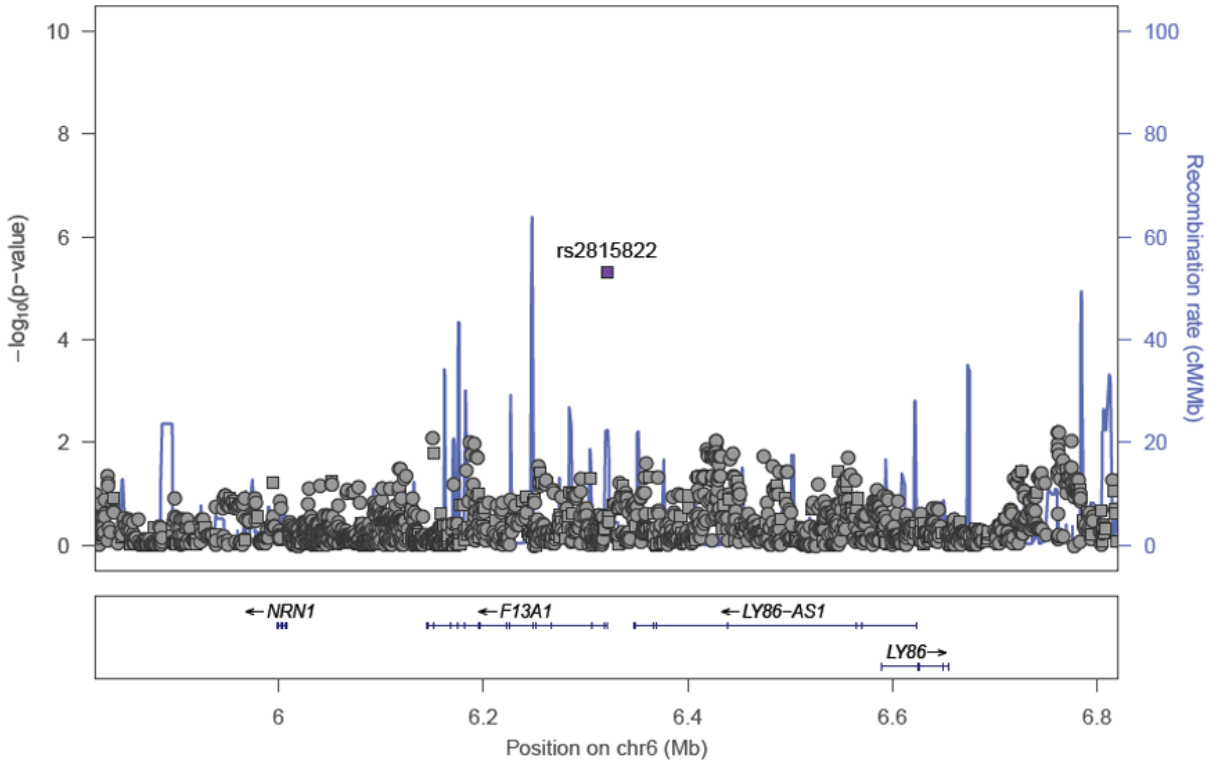


Figure 39. Regional Association Plot of rs2815822

rs10963759 (p = 8.3E-6) is intronic to *ADAMTSL1* (Figure 40), an ADAMTS-like protein showing homology to other members of the ADAMTS family but without the pro-metalloprotease and the disintegrin-like domains typical of ADAMTS proteins [228]. It may function in the ECM and is primarily expressed in skeletal muscle. Hendee et al. (2017) described a multi-generational family segregating a loss of function coding variant in *ADAMTSL1*, and affected family members exhibited dental defects including delayed tooth eruption and early loss of permanent teeth at 20-30 years of age [229]. While it is unclear precisely how *ADAMTSL1* may affect dental health, there is supporting biological evidence that it may influence tooth loss.

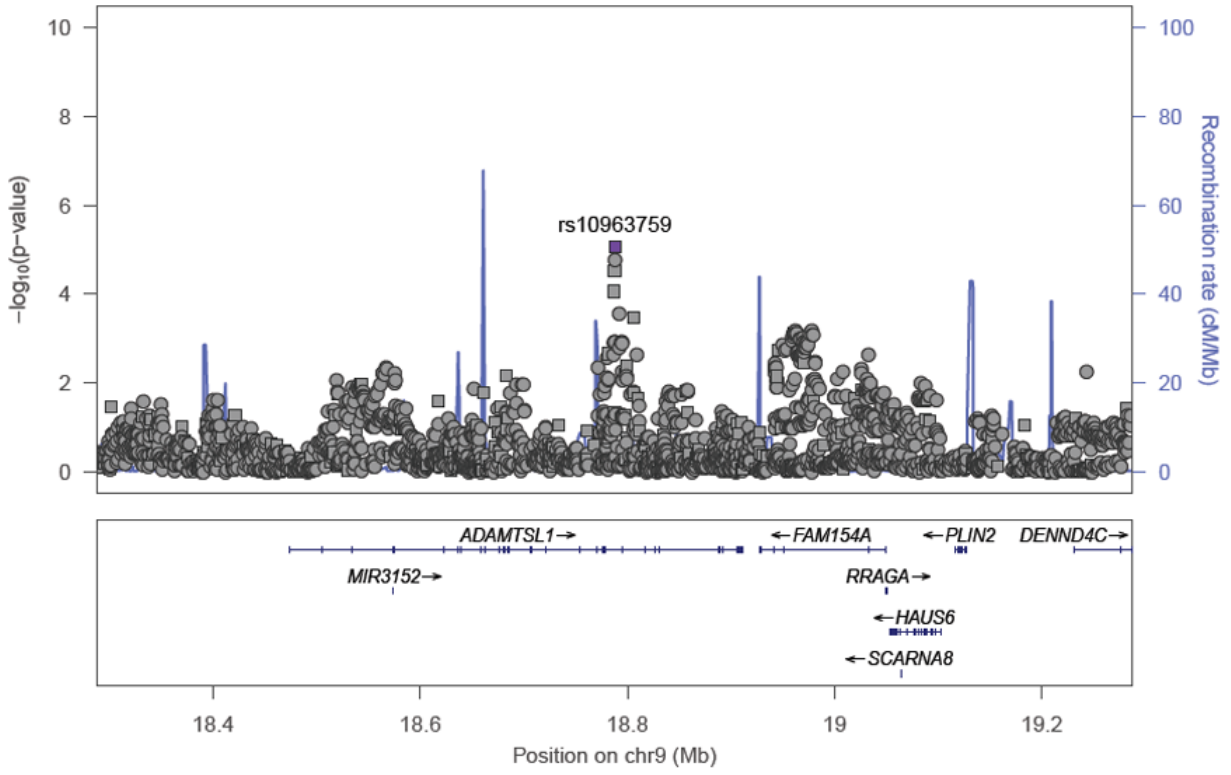


Figure 40. Regional Association Plot of rs10963759

Other SNPs reaching suggestive significance near less interesting genes as well as those not located near genes relevant to dental or oral health are described in the appendix, along with their Regional Association plots.

4.0 DISCUSSION

4.1 NOTABLE FINDINGS

The observation of both suggestive and genome-wide significant hits that were relevant to tooth loss supports a genetic basis to tooth loss and confirms the effect of common variants on susceptibility to tooth loss. As GWAS is hypothesis-generating, the goal of this project not to prove an association between these SNPs and tooth loss but to identify novel genes for further study. This goal was achieved as evidenced by the large number of suggestive findings relevant to tooth loss. Suggestively significant results were annotated to help generate hypotheses and nominate results for further study; we are not advocating that these are true associations nor that the nearby relevant genes are necessarily implicated in dental and oral diseases. Annotation of the results of these analyses relies on known biology of genes and available information from bioinformatic databases, both of which are far from complete. Genes whose known functions are not related to dental and oral health may have other unknown functions that may still be relevant to tooth loss. Similarly, genes of unknown function implicated by proximity to top hits should not be dismissed in light of the known functions of other nearby, biologically plausible genes.

While many SNPs seen at suggestive significance in COHRA remained significant in the meta-analyses, several were only seen in the meta-analyses primarily because of their relatively high significance in COHRA. This result to be expected, as COHRA had the largest weight in

the meta-analysis as it had approximately 3.5 times as many individuals as any other cohort. Additionally, the genome-wide significant hit near *POSTN* (rs12430287, p = 4.1E-8) seen in POFC-G was not significant in POFC-PA (p = 0.37); this SNP was neither imputed nor genotyped in COHRA or DRDR, therefore the association between this variant and tooth loss in whites unable to be tested. Failure of SNPs to achieve significance in the meta-analysis may be due to several reasons. The simplest explanation is that the signal was a false positive and was rightfully not significant in other samples. Alternatively, the most significant SNPs may also be subject to the "winner's curse", where their effect sizes are inflated in the sample in which they were originally discovered. While the effect does exist, the true effect size is smaller than originally estimated, thus there is lower power to replicate the association in additional cohorts [230]. Unmeasured gene-by-environment interactions may also be driving associations seen in only one sample. Differences in drinking water fluoridation govern dental caries experience, and stratification by fluoride levels has previously revealed sample-specific GWAS signals [59]. Such gene-by-environment interactions were not modeled in this project. Additionally, the COHRA sample comprises many individuals from rural areas while DRDR is more metropolitan as participants were ascertained from the University of Pittsburgh dental school. Moreover, POFC-G and POFC-PA represent two very different Hispanic ancestry samples, one coming from Guatemala in Central America and the other from the Patagonia region of Argentina. The two samples likely differ in environmental factors governing tooth loss, such as access to dental care. Altogether, failure to replicate associations across cohorts is unsurprising as there may be variants whose effect is exaggerated or counteracted by differences in environmental exposures associated with oral health and tooth loss.

There was little overlap between association signals seen between the quantitative trait and functional dentition. Only one signal in COHRA was suggestively significant for both traits, that seen near *MKI67* (see Appendix A). This lack of overlap may be due to the fact that functional dentition dichotomizes missing teeth to extreme cases vs. "normal" controls, compared to total missing teeth, which, as a continuous distribution, encompasses common variation in missing teeth. The biology of severe versus common disease may differ, which leads to different variants being more or less important between disease states.

Many of the genes identified in the study share common functions. These themes were generated by the literature, however, which is biased. Gene set enrichment analyses were not performed, though this would be a way to follow up these results in a future study. A large number of these genes play some role in tooth development, either directly or through a putative interaction with a known regulator of tooth development. Genes exhibiting specific temporospatial patterns of expression during tooth development include *WNT6* [168-170], *WNT10A* [169], *PRICKLE2* [231], *ADAMTS9* [232], *MKI67* [233, 234], *NUMB* [235], *POSTN* [140], *AQP1* [236], *TRPC6* [197], *PAX9* [85, 237], *NKX2-1* [200], *FGF7* [202, 203], *PRNP* [205], *TBX1* [186, 188], and *RAF1* [189, 190]. Additionally, *BMPRIB* is a target of RNA interference during tooth development [120] and regulates ameloblast differentiation and enamel formation [121], *PTCHD1* may act as a receptor of Shh signaling [106], which is a critical regulator of tooth development [109-112], and *AJAP1* may influence tooth development though an interaction with basigin [61, 115], which regulates tooth development by regulating MMPs [116, 117]. Genes important during tooth development may be relevant to tooth loss as improper formation of tooth structures may predispose to dental decay, malocclusion and extraction for

orthodontic reasons, dental avulsion, or periodontal disease. Furthermore, these genes may have an additional role in dental health by maintaining tooth structures after tooth development.

Several of the genes implicated at various stages of the analysis are associated with single gene disruption or small deletion disorders that include dental and/or periodontal manifestations. In the whites-only meta-analysis of the quantitative trait, one of the top hits was intronic to *IHH*, mutations of which cause Brachydactyly A-1, and dental anomalies, while not characteristic of the disorder, have been reported in some patients [166]. Mutations in *WDR19*, which was implicated in the Hispanics-only meta-analysis, cause cranioectodermal dysplasia, dental manifestations of which include hypodontia, enamel defects, and delayed tooth eruption [194]. Also implicated in the Hispanics-only meta-analysis is the chromosomal region 12p13.3, deletion of which causes a syndromic condition that includes dental malocclusion [198], and *SEC23A*, a target of TFII-I transcription factors, loss of which cause Williams Syndrome, a disorder that includes dental anomalies [201]. The trans-ethic meta-analyses implicated *CREBBP*, which is linked to Rubinstein-Taybi syndrome, features of which include dental caries, malocclusion, and hypodontia [207]. *DDX59* mutations are associated with oro-facial-digital syndrome, again with dental manifestations [238]. The 22q11.2 deletion syndrome region was implicated in the trans-ethnic meta-analysis of the quantitative trait. Various dental anomalies are common in patients with this disorder, most notably enamel hypoplasia, gingivitis, and excessive dental caries [219, 220]. *TBX1* is the strongest candidate within this region for causing these enamel defects and dental anomalies, as knockout mice recapitulate these features [186-188]. *F13A1* mutations cause factor XIII deficiency, a bleeding disorder that includes excessive bleeding of the mouth and gums [224, 226]. Early loss of permanent teeth has been reported in individuals with a loss of function coding mutation in *ADAMTSL1* [229]. Also notable is *PHEX*,

which was implicated in COHRA only; mutations in *PHEX* cause X-linked familial hypophosphatemic rickets, symptoms of which include dental defects including premature tooth loss [113, 114], though it is unlikely that the associated SNPs affect *PHEX* function or expression as the signal is approximately 450kb downstream. Altogether, these findings lend weight to the theme common in complex trait genetics, that common variants in or near genes causing syndromic disorders when disrupted may contribute to related common traits. In this case, variants causing moderate to severe dental defects when disrupted may contribute to genetic susceptibility to dental caries, periodontal disease, and tooth loss.

A handful of implicated genes are involved in mechanoreception, including *OPG*, *CTNND2*, *ADCY2*, *ADCY9*, *ADCYAP1R1*, *AQP1*, and *POSTN*. Improper response of the periodontal ligament to mechanical forces like those induced by chewing or tooth grinding and failure to maintain periodontal homeostasis are plausible mechanisms through which tooth loss may occur. Additionally, there were numerous associations seen near genes related to immunity and the inflammatory response, including *SOD1*, *FAM195A*, *NRXN1*, *CYP27A1*, *IL17F*, *IL17A*, *MAPK10*, *HLA* genes, *IRAK4*, *CARHSP1*, *CCL17*, *MAPK4*, *SGK1*, *TRAP1*, and *IPMK*. These signals may reflect an underlying genetic susceptibility to periodontitis and periodontal tooth loss. Altogether, there were many positive results located near biologically plausible genes, and it is unlikely that each one of these signals occurred at random.

4.2 SCIENTIFIC CONTEXT

Results of this GWAS and meta-analysis of missing teeth do not support the candidate gene studies of tooth loss and previously reported associations that were discussed in the introduction section. More specifically, the *OGG1*, *ESR1*, *APOE*, and *VDR* variants were not significant after multiple test correction at any stage in the analysis (data not shown), and none of the previously reported candidate genes were implicated by any of top hits. The *OGG1* polymorphism was previously reported to be associated with tooth loss in the elderly [17], however the samples included in the data presented here were not limited to elderly individuals. The effect of this polymorphism may indeed exist in the elderly, but was not associated with tooth loss in this study because of a potential interaction with age. The *APOE* ε4 allele was also not associated with missing teeth or functional dentition at any stage of the analysis. The original study assessed differences in frequency of the ε4 allele between dentate and completely edentulous individuals [15], thus the lack of association in this study between any *APOE* variants and tooth loss may reflect differences in definitions of tooth loss. Similarly, the *VDR* polymorphism was associated with number of teeth lost during a period of follow-up in older individuals [78], and the association between estrogen receptor and *VDR* genotypes reported by Taguchi et al. (2001, 2003) and tooth loss was reported in post-menopausal Japanese women [79, 80]. Association between an *MPG* polymorphism and tooth loss was again observed in elderly Japanese women [77], though *MPG* is involved in bone and tooth formation, and there were many hits for those processes in this study, most notably those involving BMP and hedgehog signaling. The failure of any of these polymorphisms to achieve suggestive significance in this study may reflect the fact that the previously reported associations were false positives, or it may be due to these differences in study populations and definitions of tooth loss.

Because dental caries and periodontitis are the main factors contributing to missing teeth, it is reasonable to assume that GWAS of missing teeth would also implicate genes previously discovered for periodontitis and dental caries. However, with a few exceptions, this was not the case. Only a handful of genes implicated by these analyses were also implicated in dental caries and periodontitis. *NRXN1*, which remained suggestively significant in the whites-only and trans-ethnic meta-analyses was previously implicated in a GWAS of severe gingival inflammation at suggestive significance [161], and *NKAIN2*, also a hit in the trans-ethnic meta-analysis, was implicated in GWAS of chronic periodontitis, again at suggestive significance [67]. *AJAP1*, which was implicated at suggestive significance in both COHRA and the whites meta-analysis, was previously identified in GWAS of dental caries at genome-wide significance [61]. The absence of such signals may be driven by the degree of disease, as tooth loss due to periodontal disease and dental caries are manifestations of severe disease. GWAS of dental caries and periodontitis have generally been done on traits that represent mild to moderate forms of disease; the genetic susceptibility to mild versus severe forms of complex disease may differ.

Despite this relative lack of overlap, there are some similarities between the current study and GWAS for dental caries and periodontitis. Genes implicated in dental caries are frequently involved in tooth development, and those implicated in periodontitis are involved in inflammation and immune response. Genes involved in all of these functions were commonly located near top hits in the current study. Also, there are no highly statistically significant associations seen in multiple study samples or populations for either dental caries or periodontitis, and no such associations were observed in this meta-analysis. This result underscores the multi-factorial nature of oral diseases, as it is the combination of genetic and environmental exposures that ultimately cause disease. Top hits for dental caries and

periodontitis generally have p-values on the order of 10^{-9} or less, which is consistent with findings in this GWAS. P-values of this magnitude are to be expected, as statistical power is low for discovering variant of low effect size, which are theorized to be major contributors to genetic susceptibility to dental caries and periodontitis. Similarly, this may be attributable to relatively small samples sizes in these GWAS; sample sizes for dental caries and periodontitis are typically around 2,000 individuals (see Tables 1 & 2 in Introduction).

4.3 STRENGTHS AND LIMITATIONS

There are several limitations of the current study. First and foremost is the relatively small sample size for GWAS, which limits the statistical power to detect rare (MAF < 0.01) variants and variants with a low effect size. This outcome is primarily reflected in the relative dearth of genome-wide significant results. In addition, a large number of the suggestively significant associations seen in COHRA only and the meta-analyses are likely to be false positives, as they do not exceed the multiple test correction threshold ($p < 5 \times 10^{-8}$) and are expected to be observed by chance alone. Indeed, many of the top associations were not located near genes relevant to oral or dental health. The sample size in the final, trans-ethnic meta-analysis was approximately 1,630 individuals, which is a relatively small sample size compared to GWAS of other complex traits. As sample size is directly related to study power, this may explain the relatively few associations observed at genome-wide significance. Perhaps with a larger sample size, many of the variants observed at suggestive significance levels that were biologically relevant to tooth loss would have surpassed genome-wide significance. Since sample size determines power, it also limits the ability to test rare (MAF < 0.01) variants; variants at MAF <

0.01 would only be seen 15 times in 1,500 individuals, which does not give adequate power when performing tests at the genome-wide scale. Due to the small sample size, only variants with MAF > 0.03 were included in this study at all steps of the analysis (except for POFC-USA, see Appendix C), thus we were unable to examine the association between tooth loss and rare variants. The majority of SNPs are rare variants, and these may potentially have the largest and most significant effects on disease. Further studies with large samples sizes are needed to determine the effect of rare variants on tooth loss.

The effect sizes identified in this GWAS are relatively large for the effect of common genetic variants on common disease. Few examples of common variants with large effect sizes for common disease are known, but the effect of *APOE* alleles on Alzheimer's disease susceptibility is a classic example. Most GWAS only identify variants with odd ratios around 1.2-2.0 [87]. In this study, however, many SNPs had odds ratios of 5 or higher – rs7717485 in COHRA only had an odds ratio of 8.19 and a wide 95% CI (3.47, 19.31.) These large odds ratios may be in part due to the small sample size and correspondingly large standard error; true estimates may be on the lower end of these wide confidence intervals. Alternatively, these variants may indeed have a large contribution to susceptibility to oral diseases and tooth loss, or these variants may have a small effect on oral health, but the cumulative nature of tooth loss leads to a large number of missing teeth and a large effect size.

As the goal of GWAS is gene discovery, only age, sex, genetic ancestry, and study site were controlled for in genetic association analyses. Other factors also associated with missing teeth, such as smoking status and education, were not included as covariates. While Pennsylvania vs. West Virginia residency was controlled for in COHRA, which incorporates income and education, modeling potential gene-by-environment interactions between these two

populations was beyond the scope of this project. Additionally, while the biological role of the top associations in dental and oral health was investigated via bioinformatics, no functional studies were performed for strongly associated SNPs. This study only established statistical associations between genetic variants and missing teeth, not the biological effect of these variants and implicated genes on tooth loss. Further studies are needed to determine the functional effect of these genes on missing teeth. Lastly, as missing teeth is a composite phenotype representing teeth missing due to periodontitis, dental caries, trauma, orthodontics, and tooth agenesis, there may be some "noise" hindering statistical power to detect variants driving specific disease processes. GWAS of dental caries and periodontal disease have used novel derived phenotypes [61, 73] generated through hierarchical clustering and principal component analysis which reflect different aspects of the biology of the diseases. Separating the missing teeth phenotype into teeth missing due to periodontal disease, dental caries, or other reasons may reveal novel associations that are lost in the "noise" of the combined missing teeth phenotype.

Despite these limitations, there were many strengths to this study. As one of the first GWAS of missing teeth [Shaffer, personal communication], this study supports previous evidence for a genetic basis of tooth loss. The genome-wide significant hit at *MTRR* establishes a putative genetic risk factor for tooth loss, and the abundance of suggestively significant hits relevant to dental and oral health, like that of *TBX1*, are strong candidates for further functional studies or replication in other cohorts. Unlike dental caries and periodontal disease, missing teeth is an unambiguous, straightforward phenotype – either the tooth is present or it is not. GWAS of dental caries and periodontal disease have been plagued with issues in defining disease status or parameters for the purpose of GWAS, like periodontal disease status or

quantitative disease measures. This study also used samples with different genetic backgrounds (whites vs. Hispanics) and environmental exposures (POFC-G vs. POFC-PA, COHRA vs. DRDR), which allowed for the examination of the differential effect of genetic variants in these varying populations. Variants important for conferring genetic susceptibility to disease may differ between populations, as evidenced by differences in significant findings in stratified analyses, like *POSTN* in POFC-G and POFC-PA.

4.4 SIGNIFICANCE AND FUTURE DIRECTIONS

This work makes a significant contribution to public health in that it establishes a foundation to build upon for future studies of the genetic contribution to tooth loss. In this study, potential genetic variants associated with missing teeth and functional dentition were identified, albeit mostly at suggestive significance. However, as this is one of the first GWAS of missing teeth, to further assess the effect of these variants, additional studies including larger and more diverse cohorts are warranted. The ultimate goal of studying the genetics of a disorder is to identify at-risk individuals and develop novel treatments. While more work needs to be done to assess the effect of the variants identified in this study before they can be used in screening, this GWAS and meta-analysis is a starting point as it nominates several genes and polymorphisms for further study. Additionally, several biological processes were highlighted by the associations, including tooth development, mechanoreception and periodontal homeostasis, and inflammatory pathways. Targeting these areas, specifically genes that are differentially expressed between disease states, may benefit individuals at risk of tooth loss. The contribution this work makes to understanding

the genetics of missing teeth and tooth loss may help to decrease oral health disparities across communities and alleviate the burden of oral diseases within populations in the future.

APPENDIX A: GWAS RESULTS

Other regions of the genome implicated in the GWA scans in COHRA and POFC-G are described in this section, including their Regional Association Plots.

A.1 COHRA – QUANTITATIVE TRAIT

rs75040946 ($p = 2.0E\text{-}6$, $\beta = -0.34$ [-0.48, -0.20]) on chromosome 10 is intronic to *MGMT* (Figure 41), a DNA repair protein that transfers methyl groups from DNA to protect against carcinogens and mutagenesis [196]. Predictably, alterations in *MGMT* activity are associated with multiple cancer types, including oral cancers [239-242]. Activity of *MGMT* is also upregulated in oral keratinocytes in response to carcinogenic exposure such as areca nut chewing [241], and is underexpressed in oral squamous cell carcinomas from smokers [240, 242]. While there is currently no known role of *MGMT* in dental health, *MGMT* may play a role in tooth loss associated with smoking.

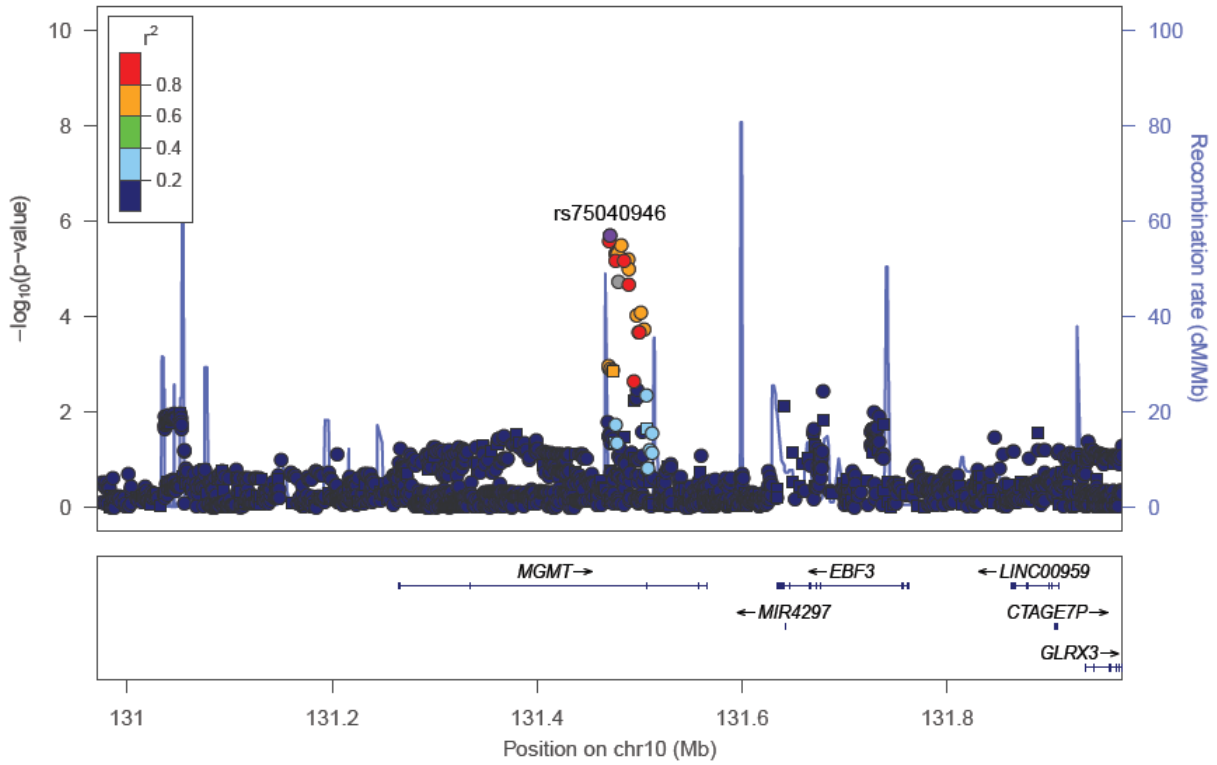


Figure 41. Regional Association Plot of rs75040946

rs4823141 (p = 2.2E-6, β = 0.21 [0.12, 0.29]) is located in a gene-rich region on chromosome 22 (Figure 42). This SNP is intronic to *EFCAB6*, an oncogene which binds the androgen receptor (AR) to regulate AR activity, though *EFCAB6* has reportedly been expressed primarily in the testis [243]. The androgen receptor is expressed in human tooth pulp [244, 245] and may contribute to dentinogenesis and tissue mineralization [245]. Furthermore, the androgen receptor has been shown to be upregulated in gingival fibroblasts in response to drug-induced gingival overgrowth [246]. In addition to the potential role of *EFCAB6* in dental health via the androgen receptor, this SNP is located approximately 400 kb upstream of *SCUBE1*, which was implicated in a GWAS of molar-incisor hypomineralization, a developmental enamel defect [247]. Lastly, this signal is approximately 300 kb downstream of *MPPED1*; *MPPED2* has been

associated with dental caries in previously [59, 63, 126], though it is unknown how *MPPED1* may affect caries susceptibility.

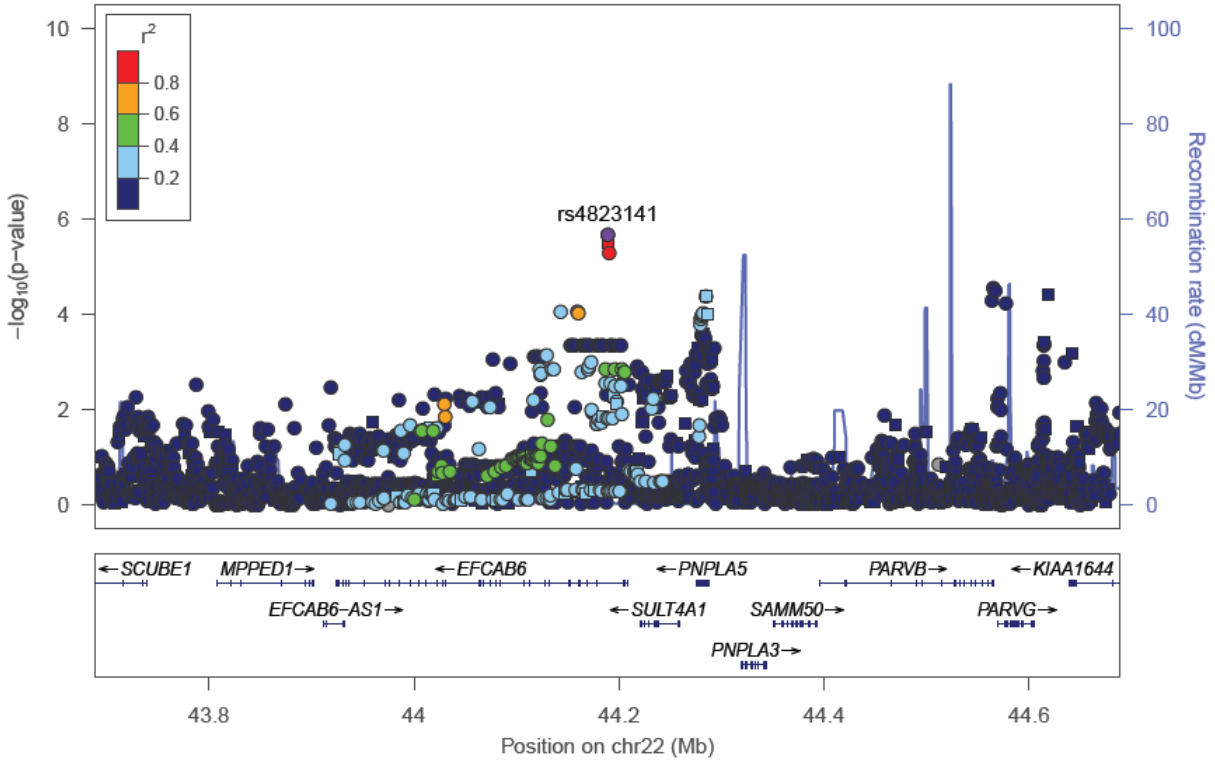


Figure 42. Regional Association Plot of rs4823141

rs75195099 (p = 2.5E-6, β = 0.54 [0.32, 0.76]) is located just upstream of *MKI67* (Figure 43), marker of proliferation 67 (Ki-67). Ki-67, the protein, is widely used as a marker of cellular proliferation in cancer studies and developmental biology. The precise function of Ki-67 in cell proliferation is as yet unclear, though inhibition of Ki-67 inhibits DNA synthesis and it is suggested that Ki-67 plays a role in rRNA synthesis [248]. As Ki-67 is such a widely used marker of cellular proliferation, it has been used in many studies of tooth development. Ki-67 has been shown to be differentially expressed throughout the stages of tooth development in the dental papilla and parts of the developing enamel organ such as ameloblasts (enamel depositing

cells) [233], and decreases in Ki-67 expression and cell proliferation coincide with differentiation of cell types such as ameloblasts and odontoblasts [234]. Follow-up studies would be necessary to determine whether Ki-67 plays a direct role in cellular proliferation and thus tooth development, or if it is merely a marker of proliferation, which is abundant during tooth development.

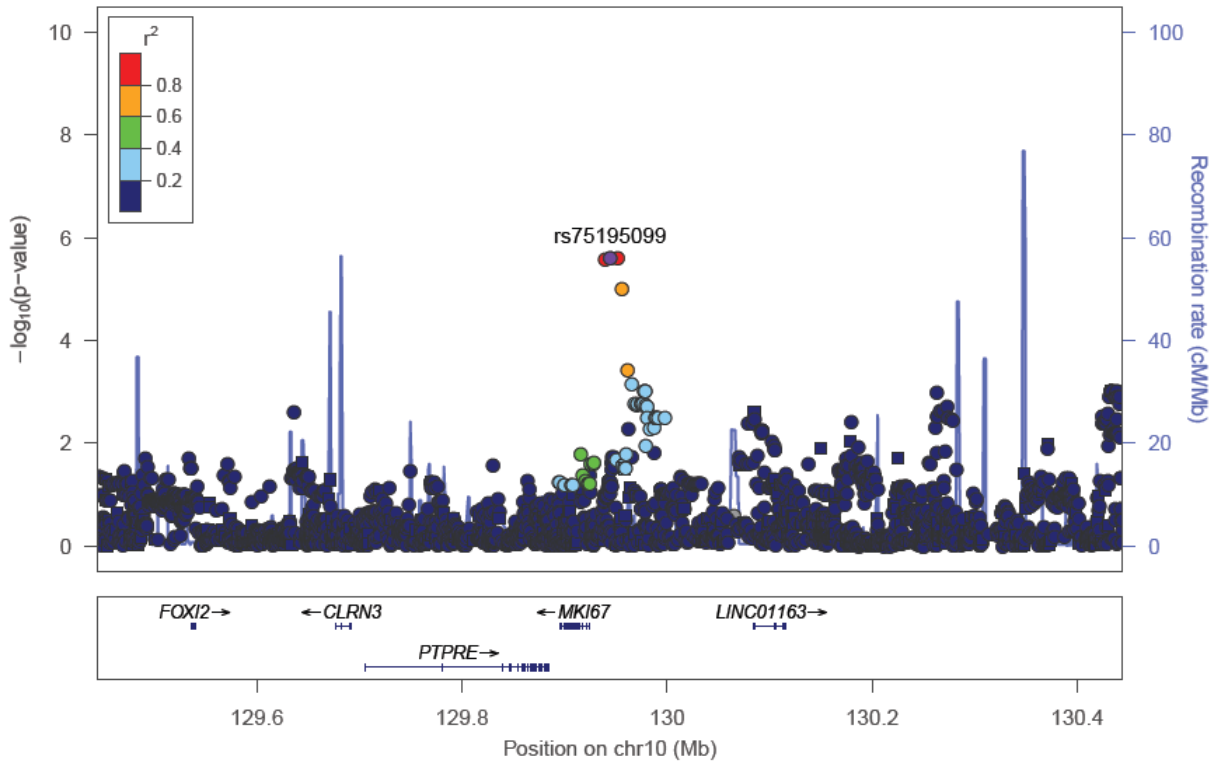


Figure 43. Regional Association Plot of rs75195099

rs2964805 (p = 2.9E-6, β = -0.22 [-0.32, -0.13]) is located approximately 200 kb from *CTNND2* (Figure 44), which encodes an adhesive junction associated protein and plays a role in cell adhesion. *CTNND2* was found to be downregulated 14-fold in (PDL) cells in response to mechanical force loading, indicating a role of this gene in maintaining homeostasis in PDL cells

in response to stretch induced cell realignment, periodontal remodeling, orthodontic tooth movement, and occlusal function [249].

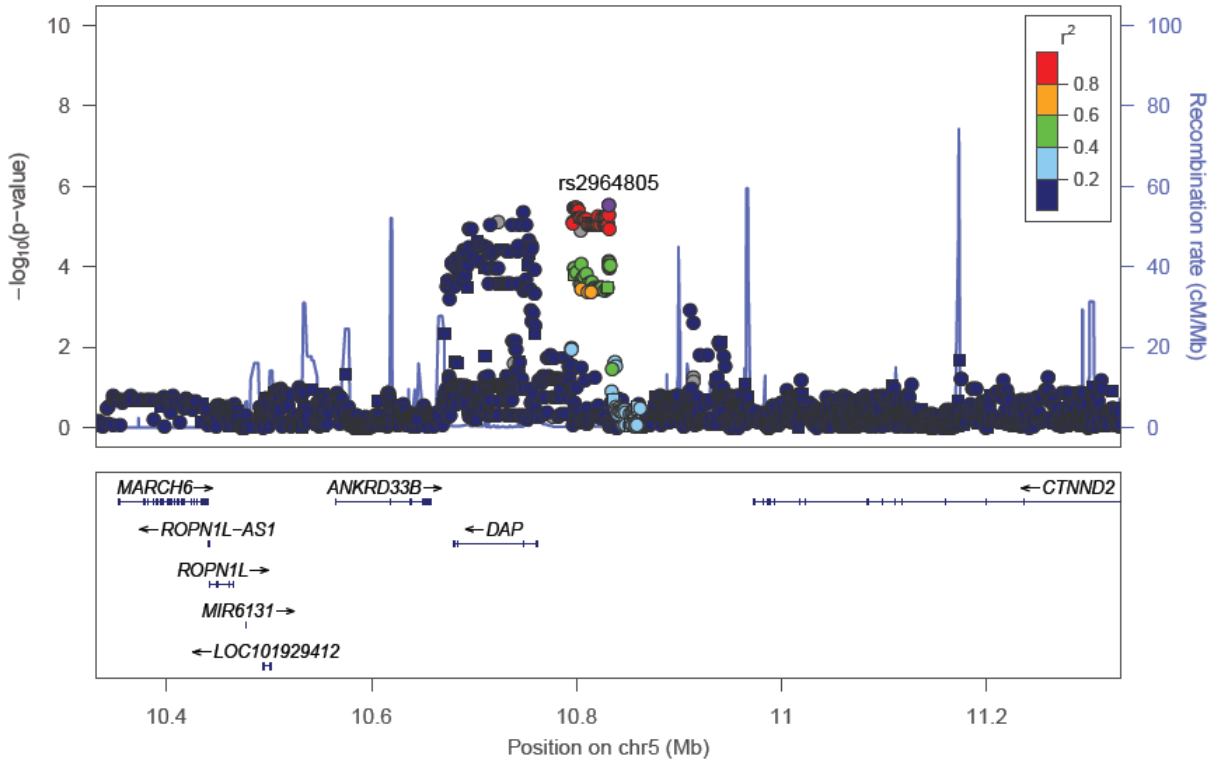


Figure 44. Regional Association Plot of rs2964805

rs73199539 (p = 3.2E-6, β = -0.42 [-0.60, -0.25]) is located approximately 100 kb upstream of *SOD1* (Figure 45); this gene encodes the enzyme superoxide dismutase (SOD), which regulates oxidative damage by destroying damaging superoxide radicals. Lower oxidative damage and increased activity of SOD has been reported in adults with severe caries and children with severe early childhood caries [149, 250], indicating a possible role of *SOD1* in dental caries susceptibility. Oxidative stress generated by reactive oxygen species (ROS) also plays a role in periodontal disease, as ROS are produced by leukocytes during the inflammatory

response and contribute to periodontal tissue destruction [251]. Indeed, increased levels of oxidative damage are associated with periodontal disease [252, 253].

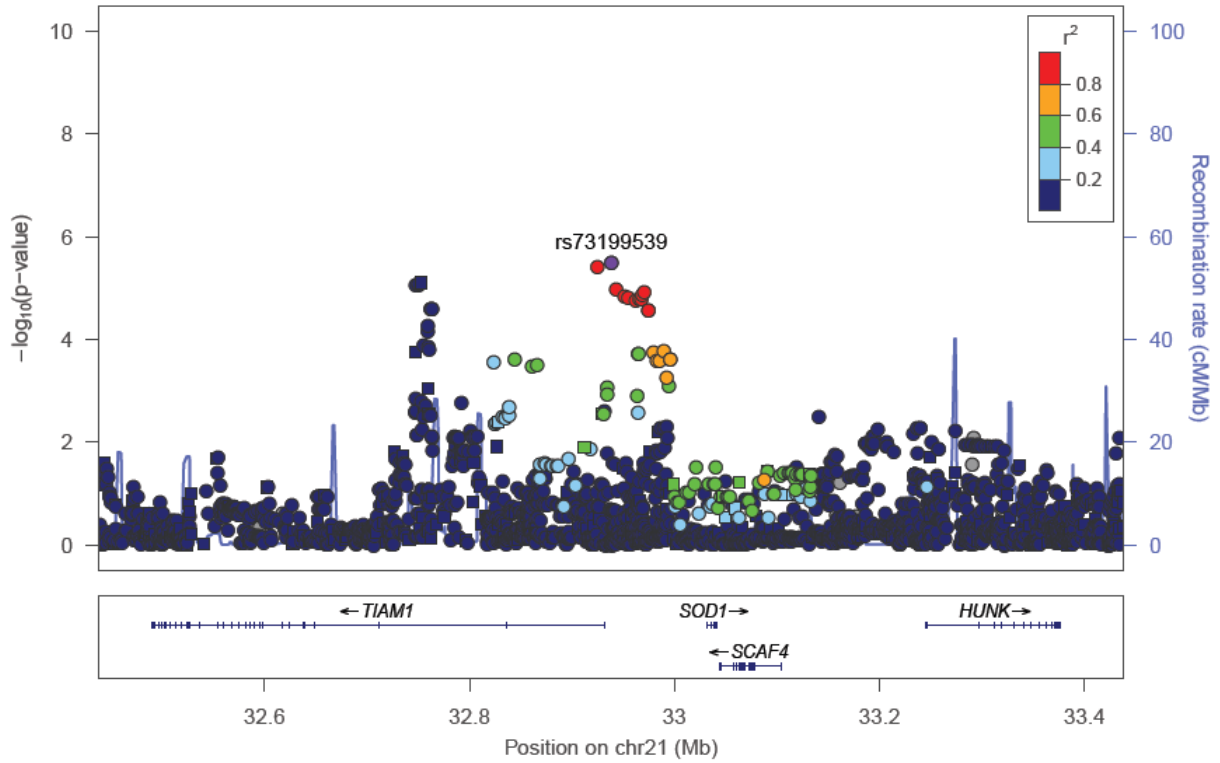


Figure 45. Regional Association Plot of rs73199539

rs2860805 (p = 5.6E-6, β = -0.22 [-0.32, -0.13]) is located near a handful of genes with no known role in dental or oral health (Figure 46). *NUDCD2*, the function of which is unknown, is located approximately 400 kb downstream of this SNP. *HMMR* and *HMMR-AS1* are located approximately 350 upstream of this SNP; *HMMR*, hyaluronan mediated motility receptor, is involved in formation of the mitotic spindle [254] and is implicated in various cancer types, especially breast cancer [255-257]. *MAT2B*, approximately 300 kb upstream of the index SNP, is involved in methionine metabolism, a process implicated by the results in COHRA and the whites-only meta-analysis for functional dentition. *MAT2B* encodes the beta subunit of

methionine adenosyltransferase, which catalyzes the production of S-adenosylmethionine from methionine and ATP [258]. It is unclear how variation within this genomic region may influence dental health and tooth loss.

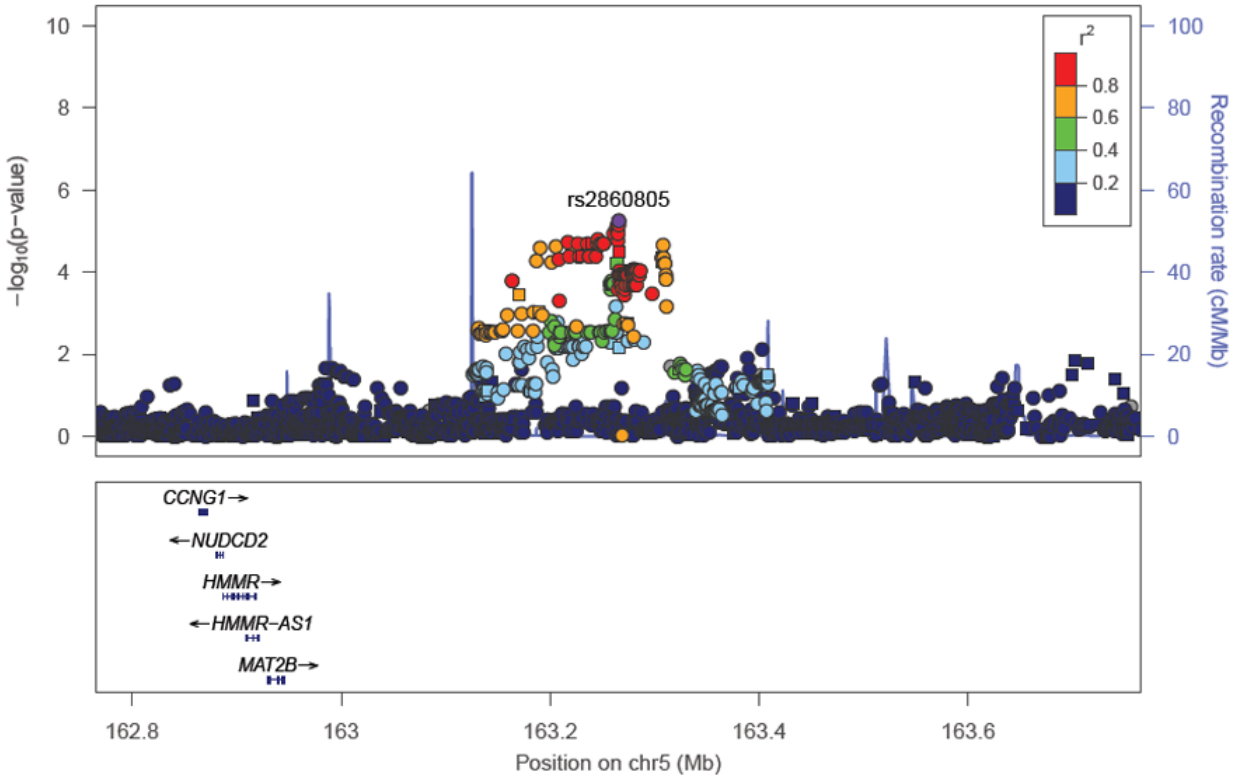


Figure 46. Regional Association Plot of rs2860805

rs12297548 (p = 4.7e−06, β = 0.36 [0.21, 0.51]) is located just upstream of *GPR133* (Figure 47), an adhesion G protein-coupled receptor that functions in the G_s protein/adenylyl cyclase pathway [259]. While the role of this G-protein receptor in dental or oral health is unknown, G-protein signaling, adenylyl cyclases, and cAMP signaling may be important in tooth loss as cAMP signaling may play a role in periodontal homeostasis and immune response, as discussed previously.

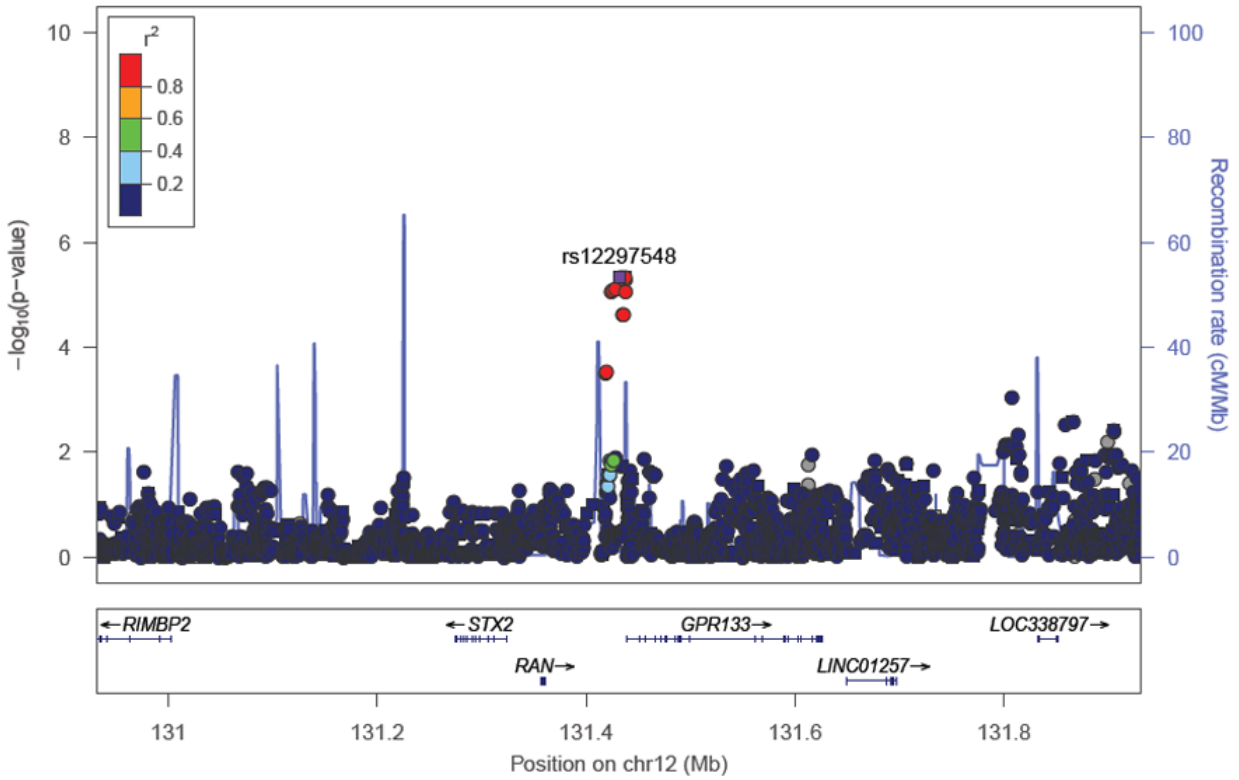


Figure 47. Regional Association Plot of rs12297548

rs73104494 (p = 5.98e−06, β = 0.23 [0.13, 0.33]) is located approximately 100 kb upstream of *ROBO2* (Figure 48), which was implicated in a GWAS of chronic periodontitis, also at suggestive significance (p = 2.64E-6) [67]. *ROBO2* functions in axon guidance during neural development, as well as cell migration [260], though it is unknown how Robo proteins may affect dental or oral health.

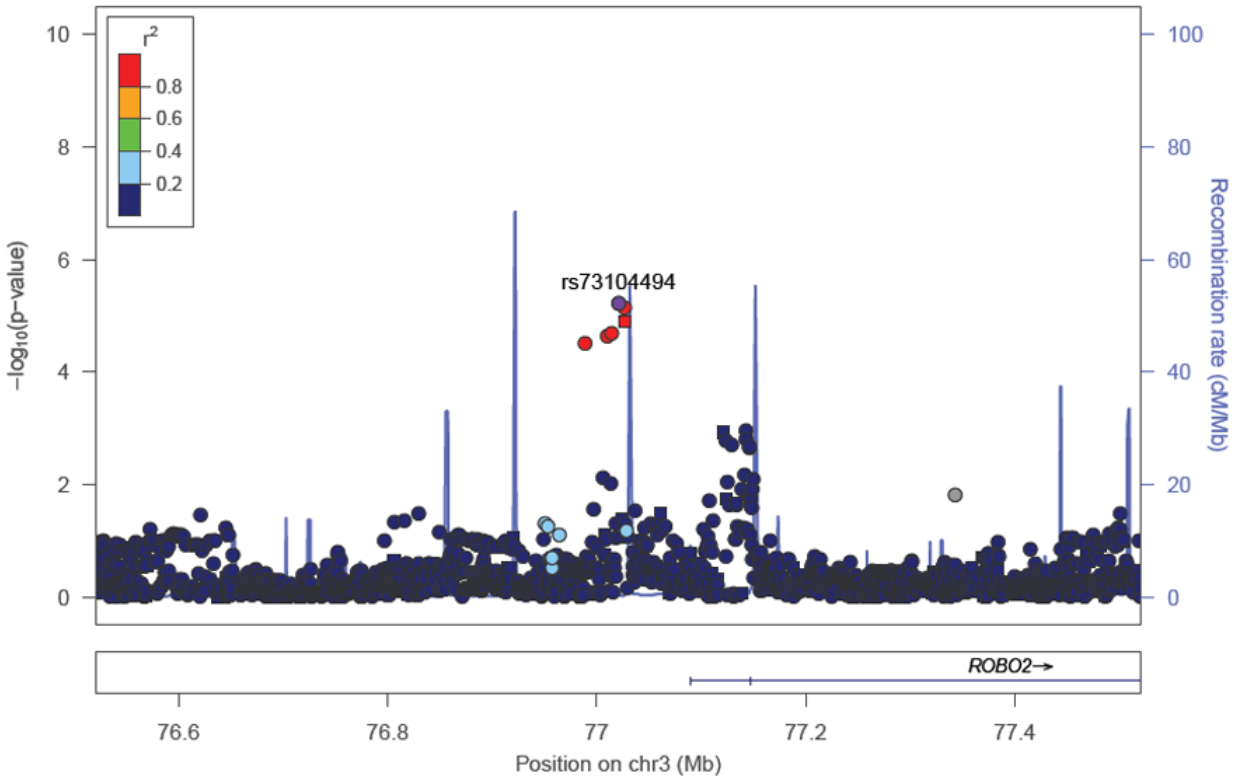


Figure 48. Regional Association Plot of rs73104494

rs10812719 (p = 8.48e−06, β = -0.19 [-0.28, -0.11]) is intronic to *KIAA0020*, also known as *PUM3* (Figure 49). This index SNP is in high LD with rs2270888 (r^2 = 0.96) and rs7036752 (r^2 = 0.95), both of which are missense variants; however, it is unknown how these missense variants affect *PUM3* gene function, or how *PUM3* may affect dental health.

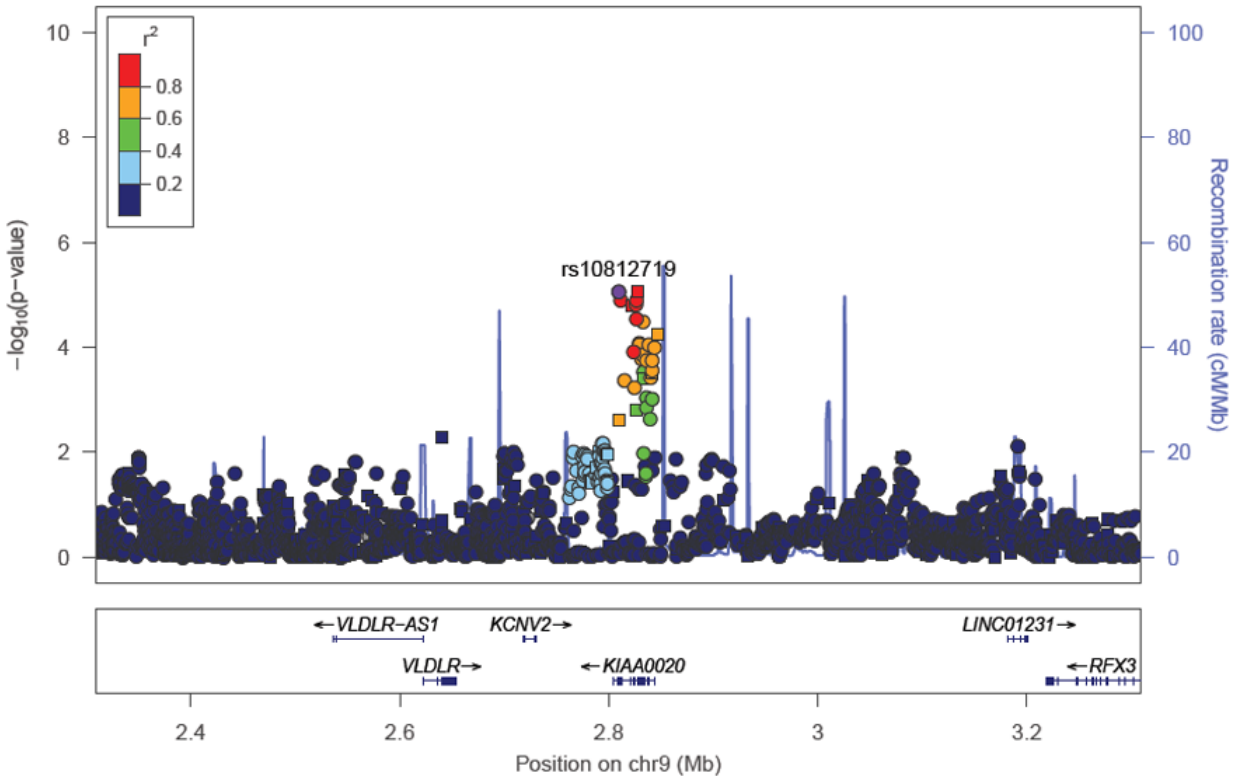
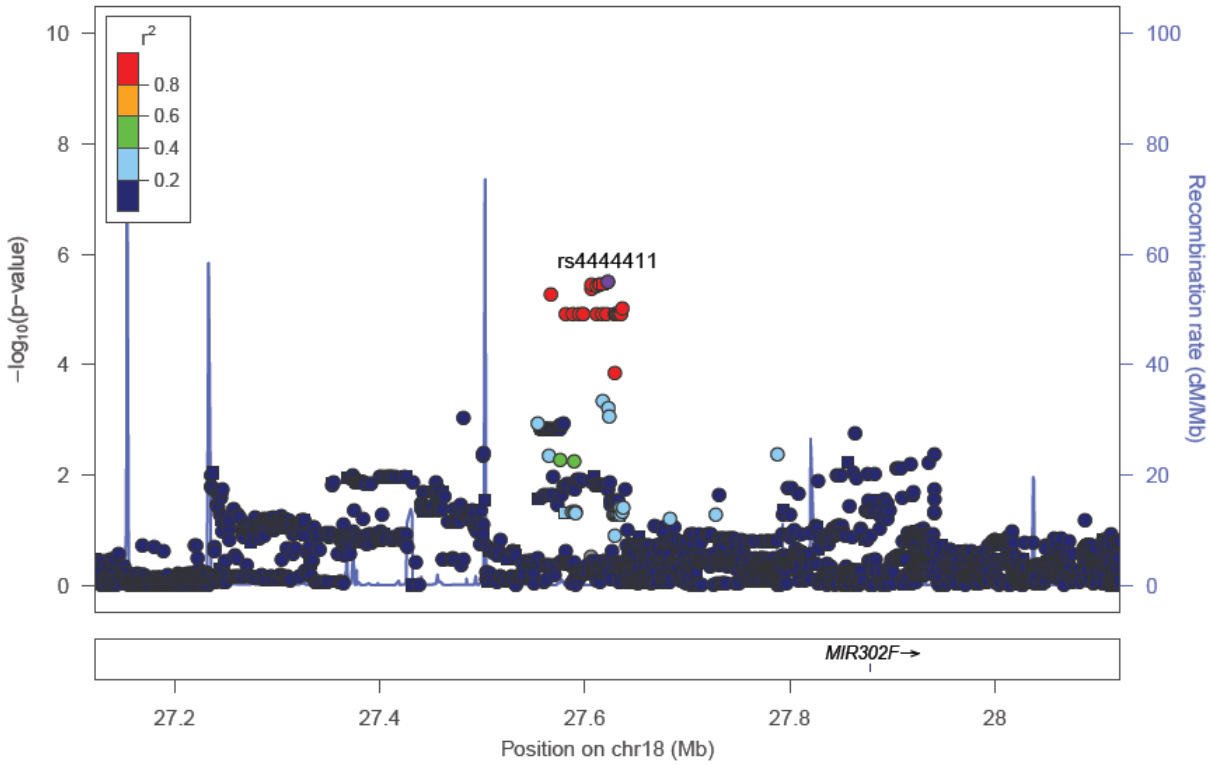
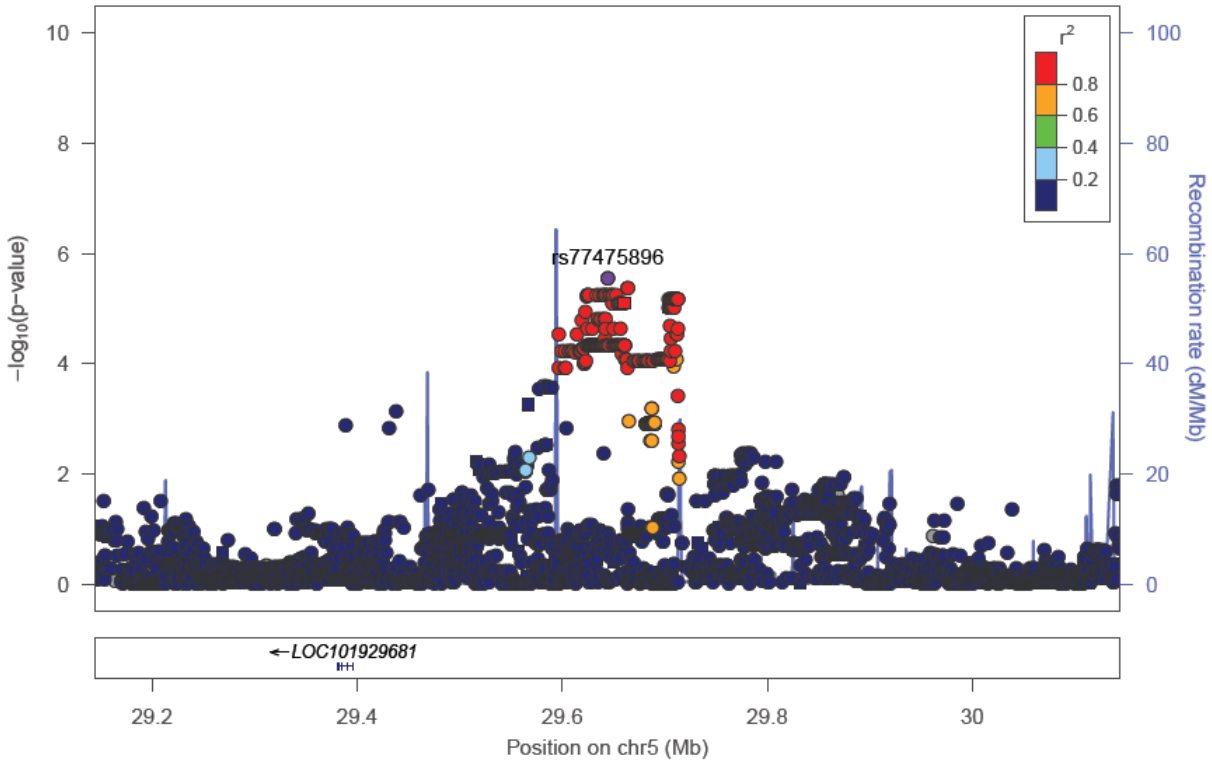
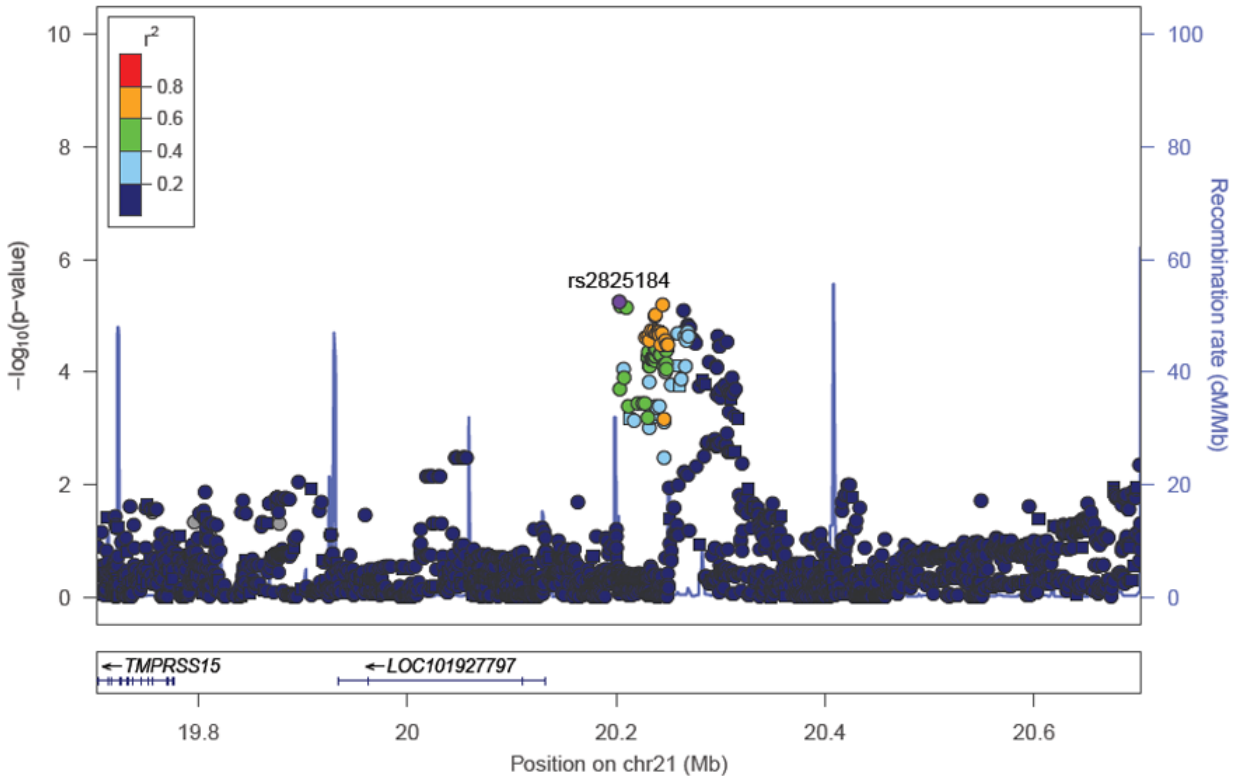


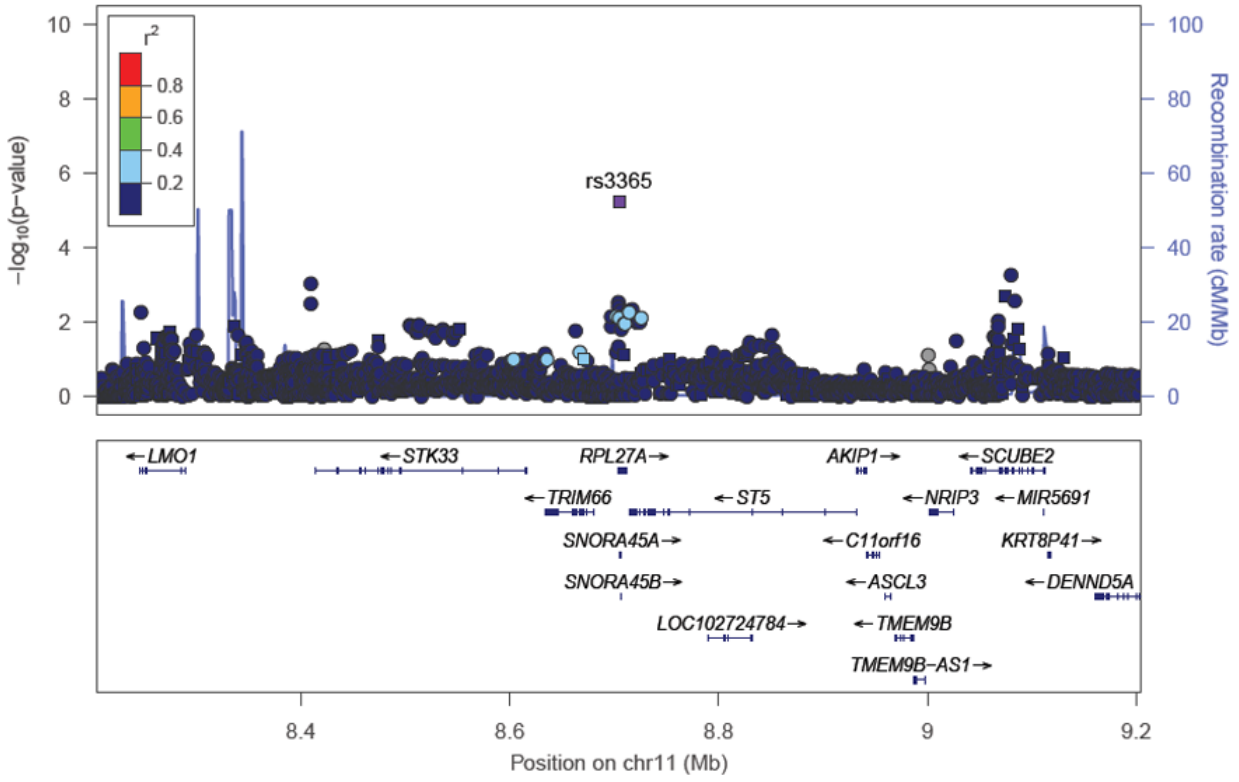
Figure 49. Regional Association Plot of rs10812719

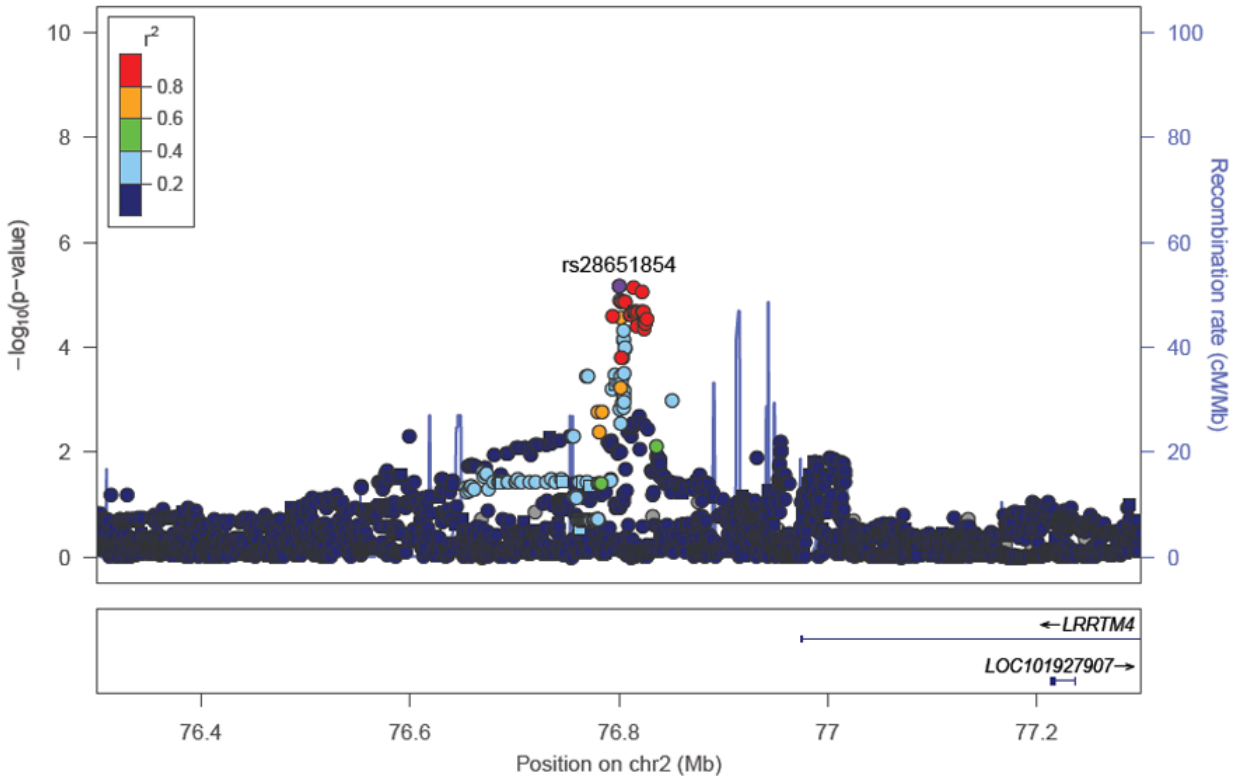
Other SNPs reaching suggestive significance were not located near genes relevant to dental or oral health (Figure 50).

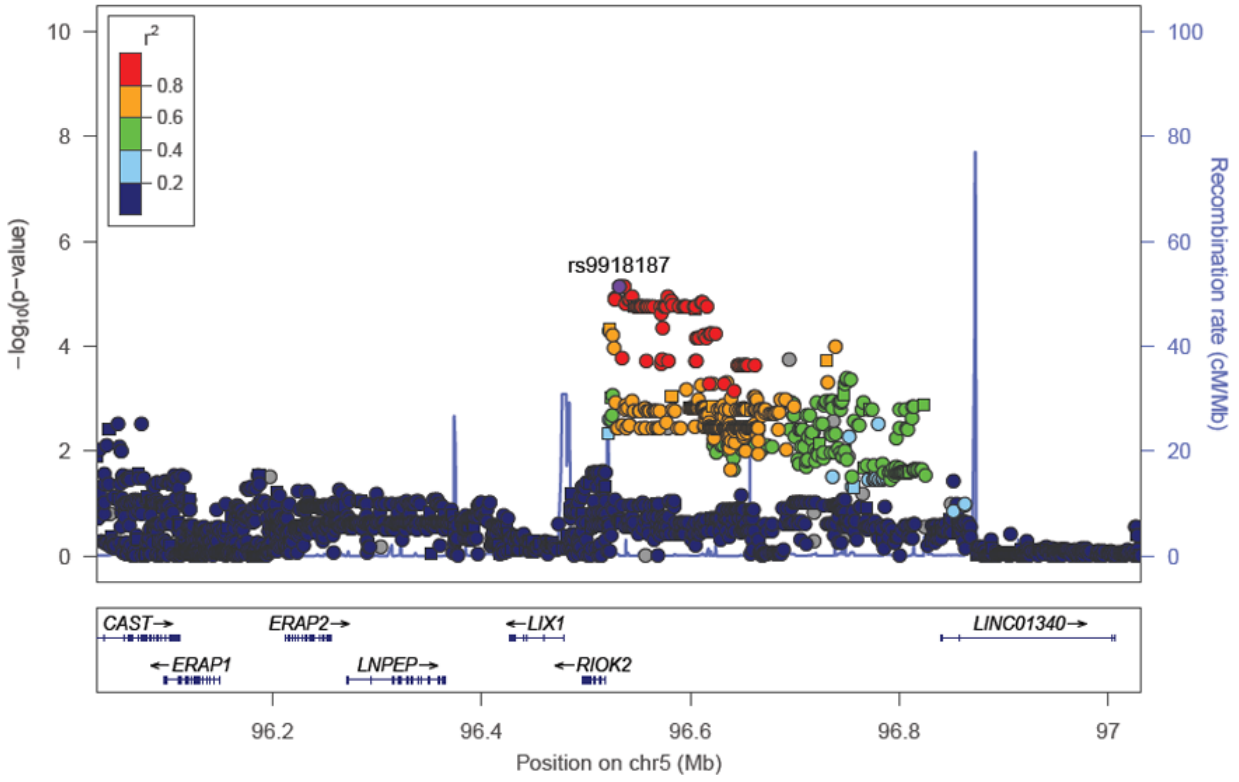

r²
0.8
0.6
0.4
0.2
rs4444411
−log10(p−value)
Recombination rate (cM/Mb)
MIR302F
27.2
27.4
27.6
27.8
28
Position on chr18 (Mb)


r²
0.8
0.6
0.4
0.2
rs77475896
−log10(p−value)
Recombination rate (cM/Mb)
LOC101929681
29.2
29.4
29.6
29.8
30
Position on chr5 (Mb)


r²
0.8
0.6
0.4
0.2
rs2825184
−log10(p−value)
Recombination rate (cM/Mb)
TMPRSS15
LOC101927797
19.8
20
20.2
20.4
20.6
Position on chr21 (Mb)


r²
0.8
0.6
0.4
0.2
rs3365
−log10(p−value)
Recombination rate (cM/Mb)
LMO1
STK33
RPL27A
AKIP1
SCUBE2
TRIM66
ST5
NRIP3
MIR5691
SNORA45A
C11orf16
KRT8P41
SNORA45B
ASCL3
DENND5A
LOC102724784
TMEM9B
TMEM9B−AS1
8.4
8.6
8.8
9
9.2
Position on chr11 (Mb)


r²
0.8
0.6
0.4
0.2
rs28651854
−log10(p−value)
Recombination rate (cM/Mb)
LRRTM4
LOC101927907
76.4
76.6
76.8
77
77.2
Position on chr2 (Mb)


r²
0.8
0.6
0.4
0.2
rs9918187
−log10(p−value)
Recombination rate (cM/Mb)
CAST
ERAP2
LIX1
LINC01340
ERAP1
LNPEP
RIOK2
96.2
96.4
96.6
96.8
97
Position on chr5 (Mb)

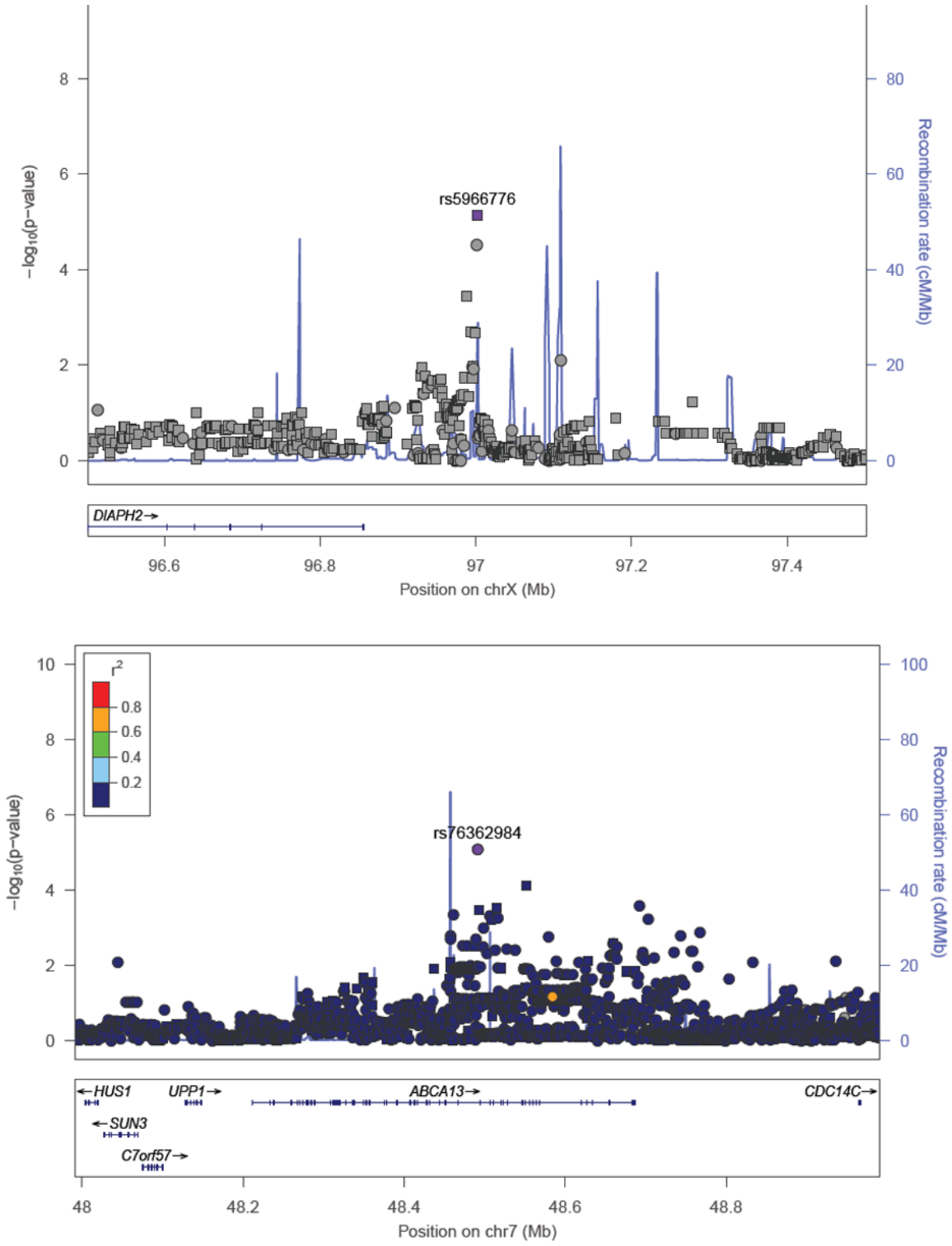


Figure 50. Regional Association Plots for other suggestively significant hits for the quantitative trait in COHRA

LD information was not available for all plots. For rs5966776 annotation of imputed vs. genotyped variants is reversed.

A.2 COHRA – FUNCTIONAL DENTITION

rs72673432 (p = 8.61E-7, OR = 4.09, 95% CI: [2.34, 7.18]) is intronic to *ANK2* (Figure 51), which encodes an ankyrin protein. Ankyrin proteins are known to be involved in ion transport and ion channel function, and mutations in *ANK2* cause ankyrin-B syndrome, which is characterized by a cardiac arrhythmias and sudden death [261]. No known role in dental and oral health has yet been established. This genomic region also contains several mRNAs, though none have been studied in dental and oral health.

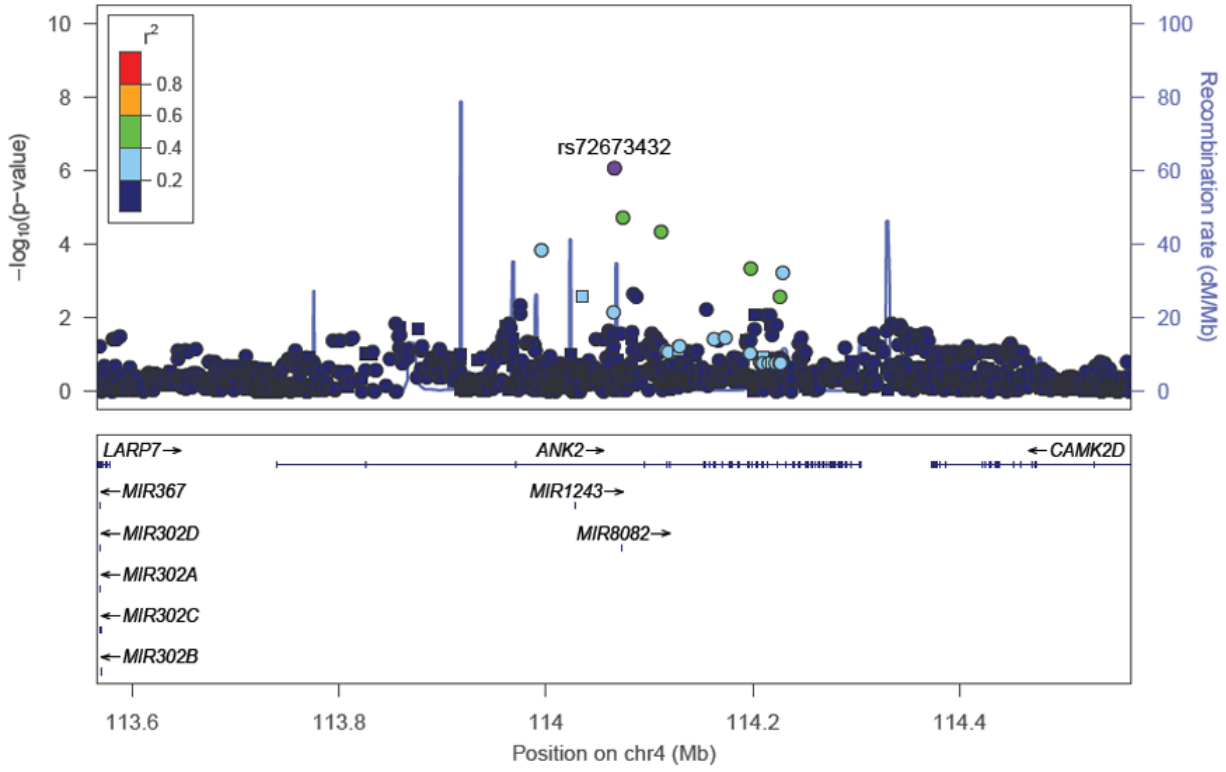


Figure 51. Regional Association Plot of rs72673432

rs7717485 (p = 1.58E-6, OR = 8.19, 95% CI: [3.47, 19.31]) is intronic to CDH14 (Figure 52), which encodes a type II class cadherin molecule expressed in the central nervous system and

is possibly involved in regulating neural development [262], though it has no known role in dental health.

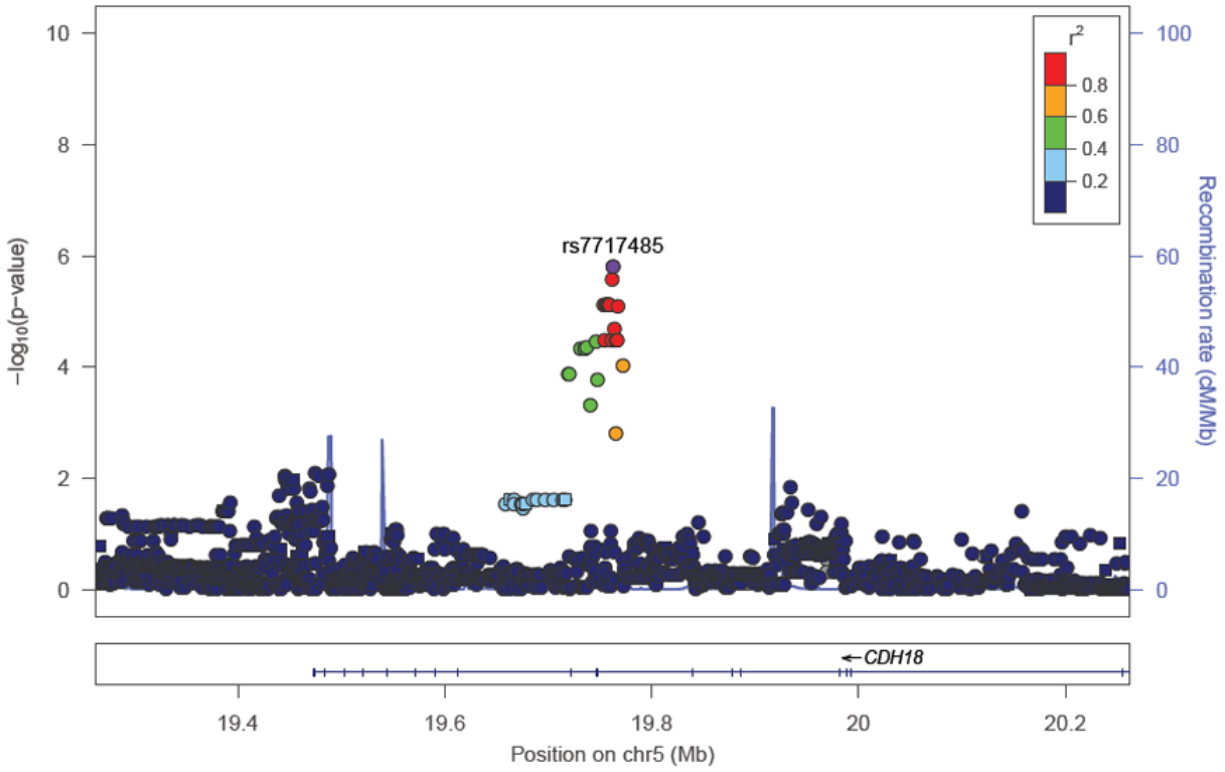


Figure 52. Regional Association Plot of rs7717485

rs7624909 (p = 2.47E-6, OR = 5.29, 95% CI: [2.64, 10.57]) is located just downstream of *ADAMTS9* (Figure 53). The ADAMTS family of proteins, which are expressed in the extracellular matrix (ECM) and interact with ECM proteins, are involved in neural crest formation and craniofacial morphogenesis, and *ADAMTS9* may influence cell proliferation by controlling ECM composition in cells derived from the neural crest [263]. *Adamts9* is expressed in the developing tooth socket of mice as well as the tooth itself, indicating a possible role in tooth development [232]. This signal is also approximately 220 kb upstream of *PRICKLE2*, a planar cell polarity protein with a role in amelogenesis [231]; *PRICKLE2* is expressed in

secretory ameloblasts, and expression was correlated with sites of ameloblast proliferation and differentiation in rats.

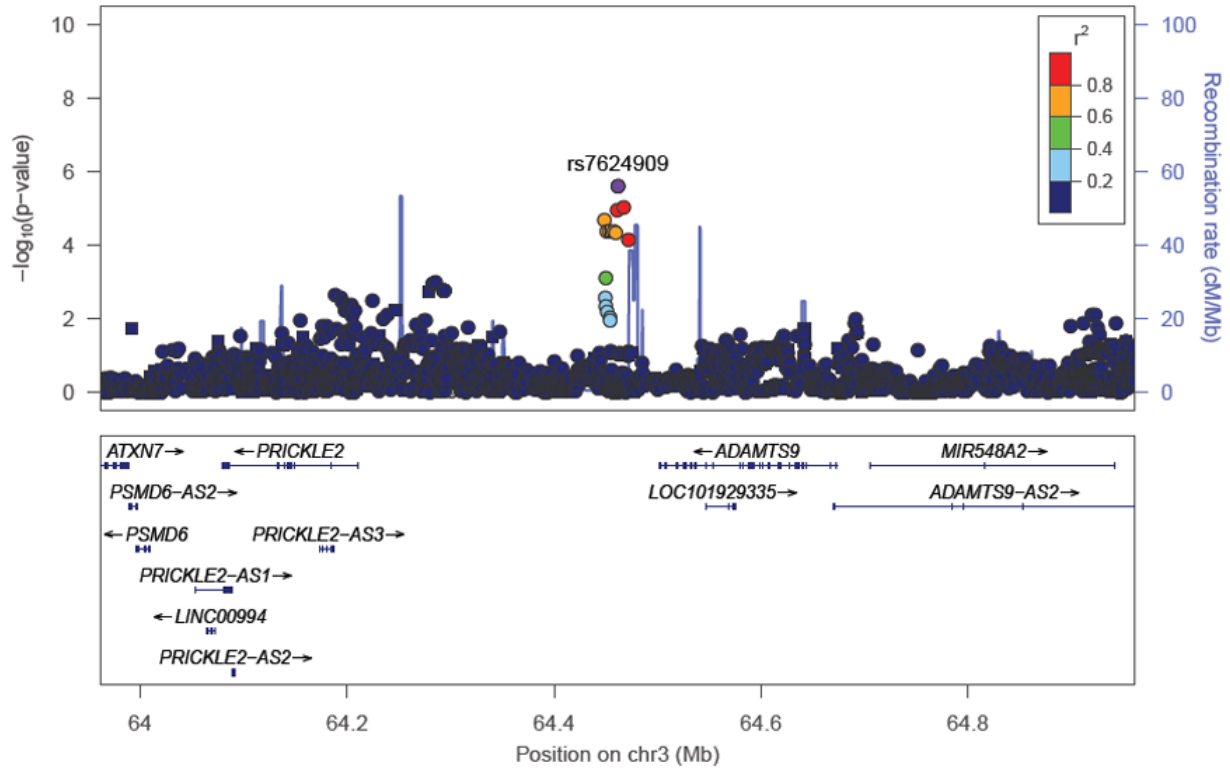


Figure 53. Regional Association Plot of rs7624909

rs76798443 (p = 2.51E-6, OR = 4.49, 95% CI: [2.4, 8.38]; Figure 54) corresponds to the same signal near *MKI67* seen the quantitative trait (Figure 43). Interestingly, this was the only signal in common between both traits at $p < 10^{-5}$.

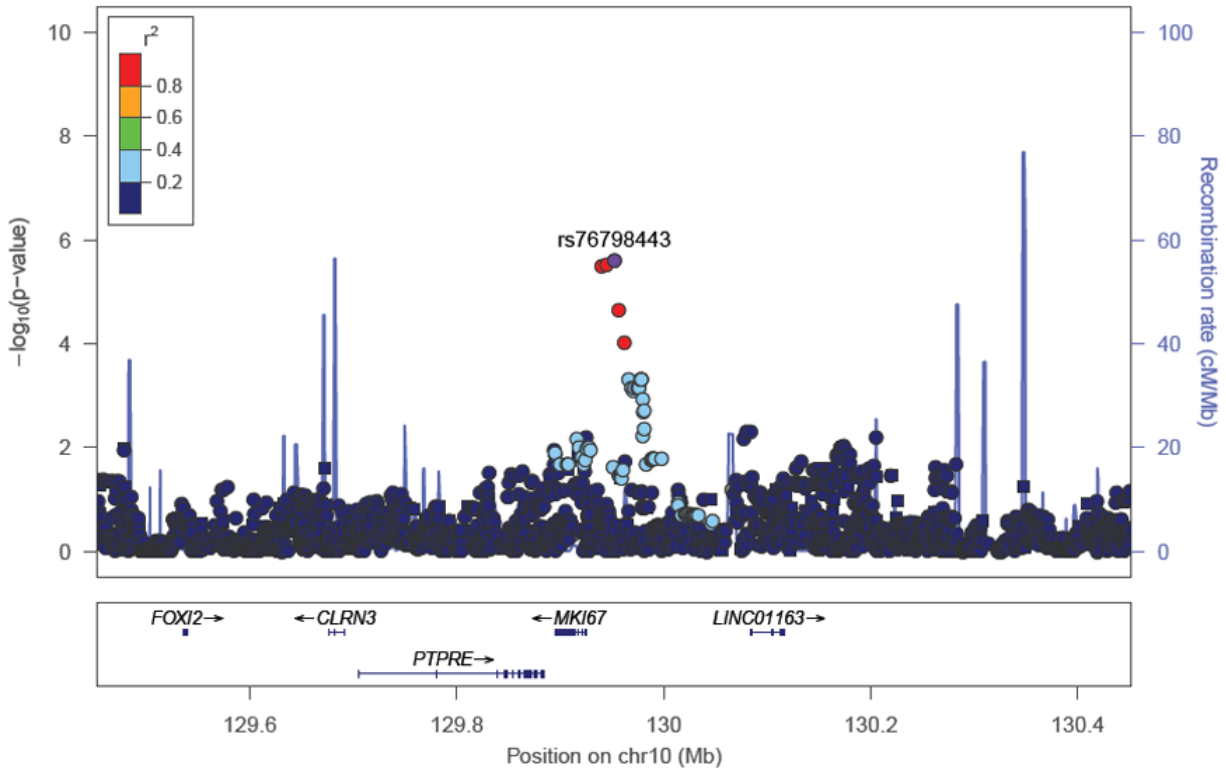


Figure 54. Regional Association Plot of rs76798443

chr14:74029866 (p = 2.51E-6, OR = 4.49, 95% CI: [2.4, 8.38]) is located near a cluster of ACOT genes (Figure 55), though these have no known role in dental health. However, this SNP is located approximately 100 kb upstream of *NUMB*, which is critical in determining cell fate during development and is part of Notch and Hedgehog signaling pathways [264]. NUMB isoforms are differentially expressed in odontogenic cells, and overexpression of NUMB in ameloblasts inhibited Notch and downregulated Hedgehog proteins, indicating that NUMB may regulate ameloblast differentiation through these interactions [235]. In addition to *NUMB*, this signal is also 250 kb downstream of *PAPLN*, an ECM glycoprotein, which was found to be differentially expressed in the developing mouse molar tooth, specifically in ameloblasts [204].

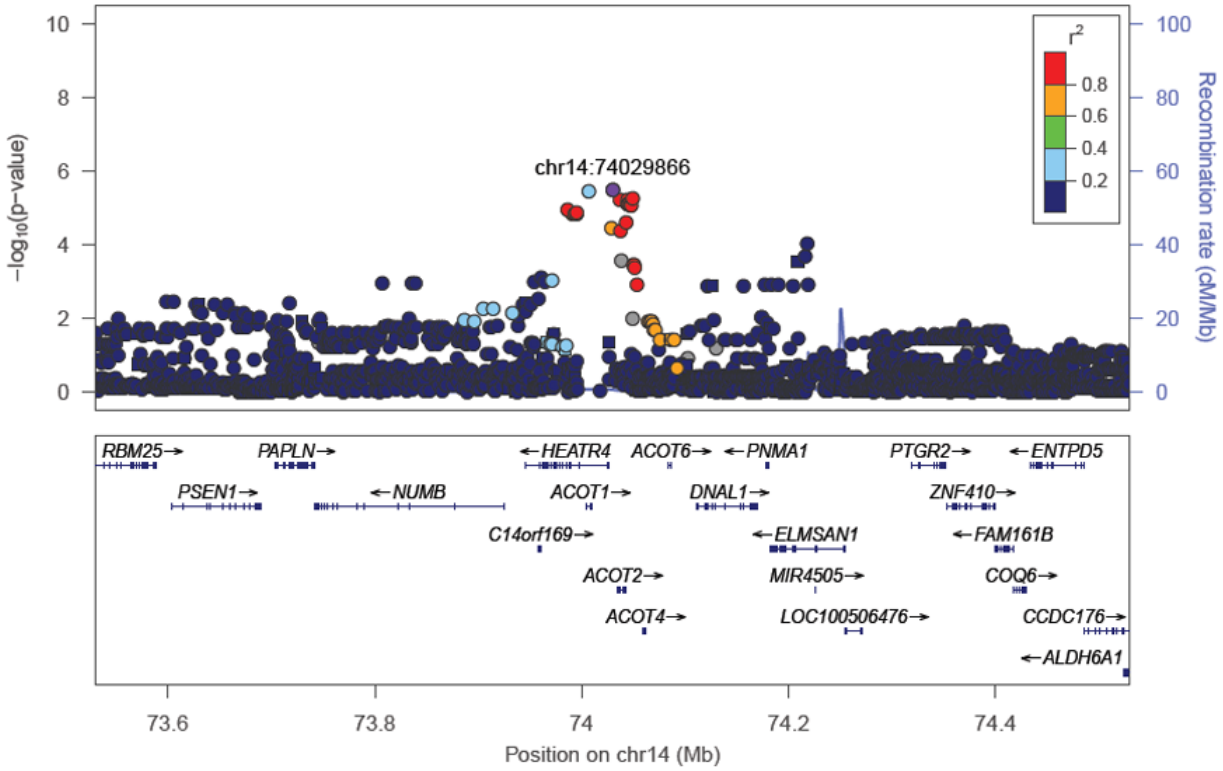


Figure 55. Regional Association Plot of chr14:74029866

rs76944100 (p = 9.11E-6, OR = 3.78, 95% CI: [2.10, 6.81]) is located approximately 250 kb upstream of *EFR3A* and 400 kb downstream of *OC90* (Figure 56), which were implicated in a GWAS of dental caries in the primary dentition at suggestive significance (p = 9.8E-6) [63].

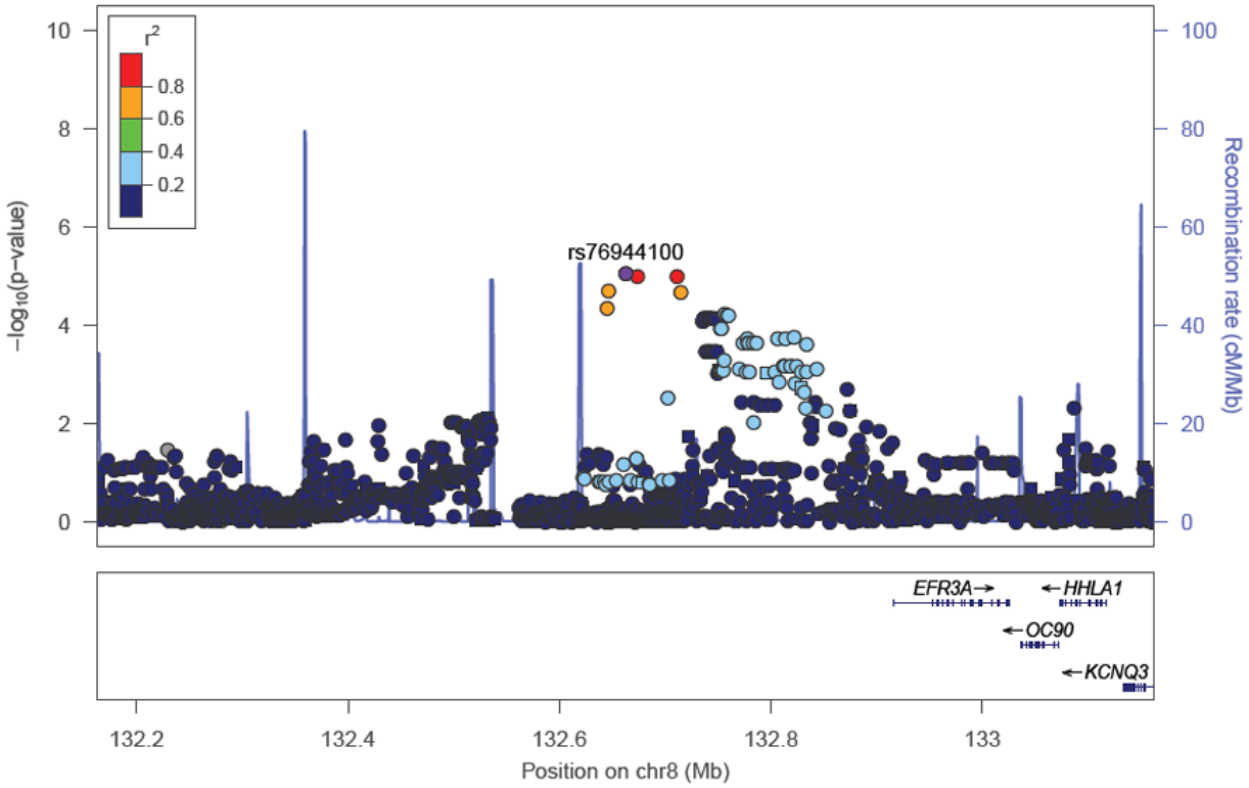
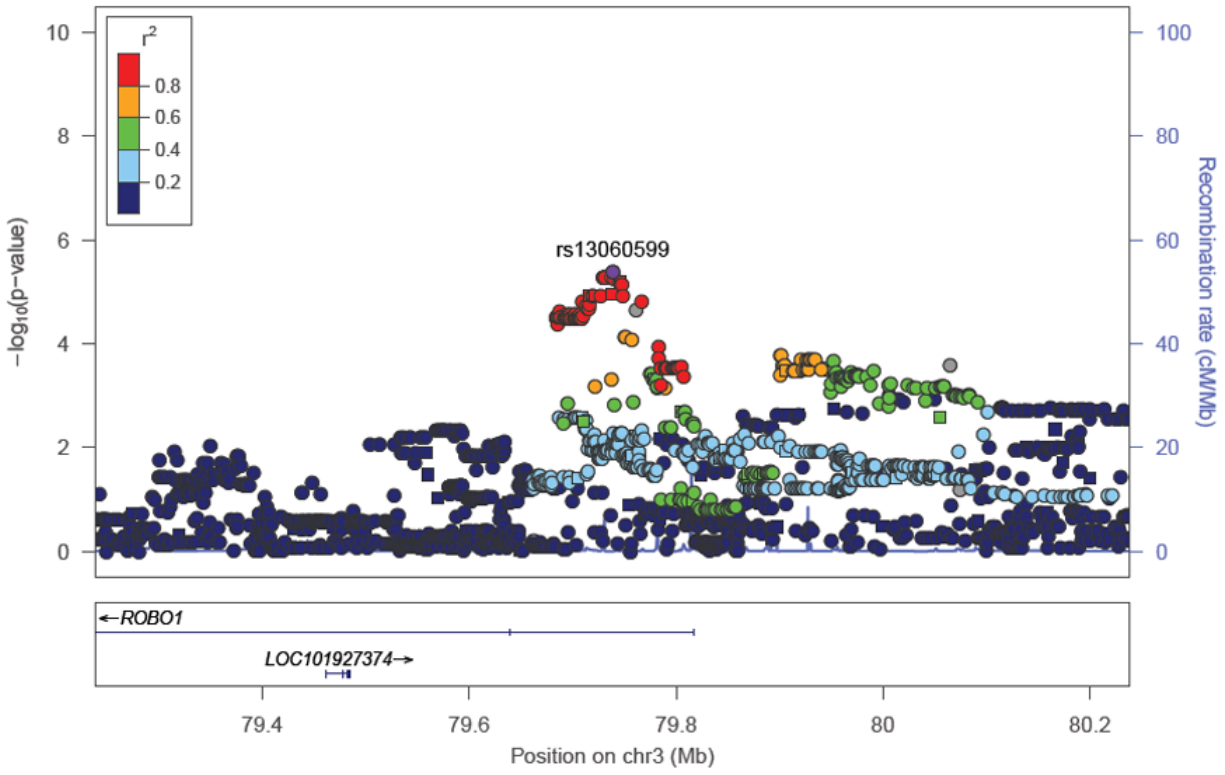
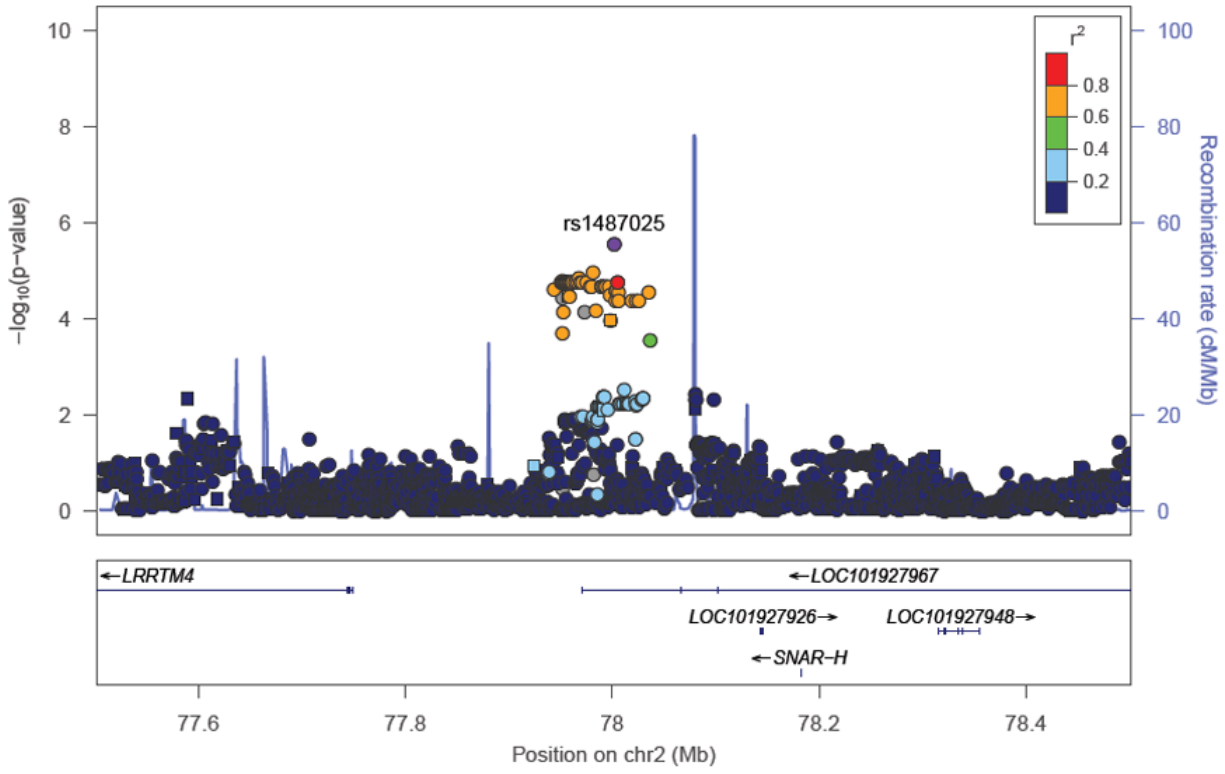
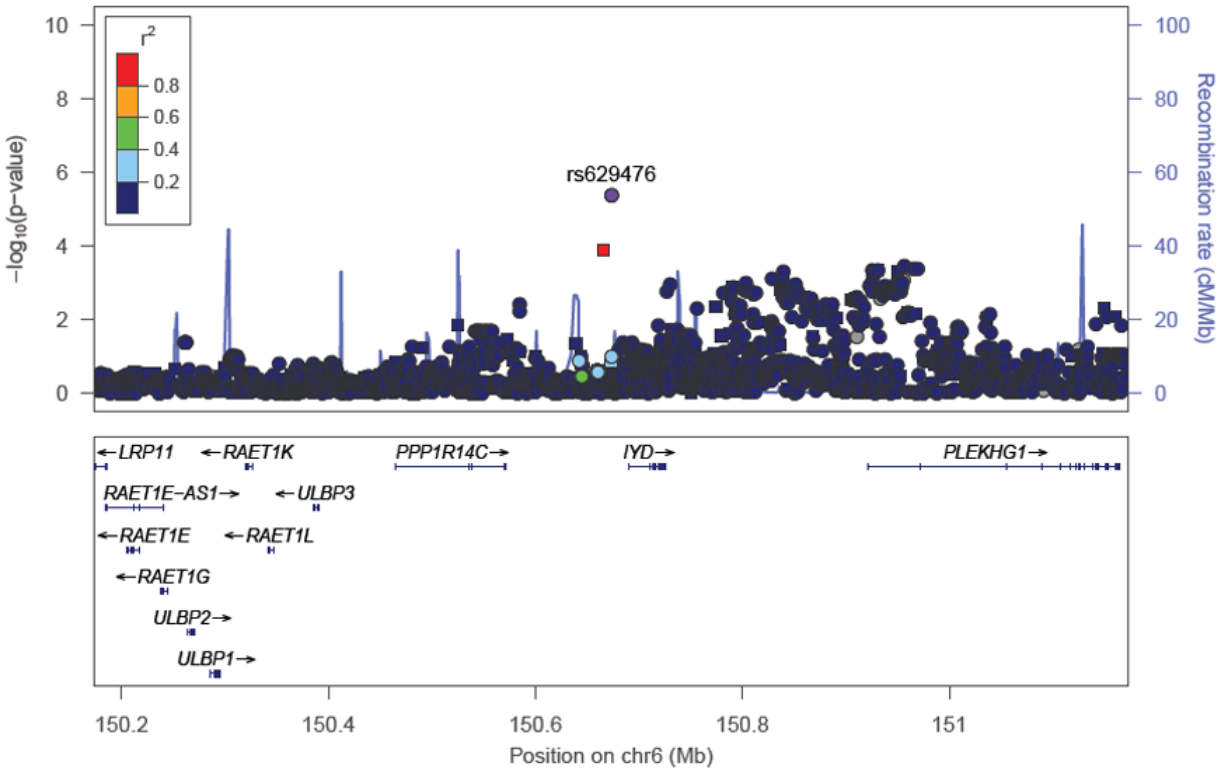


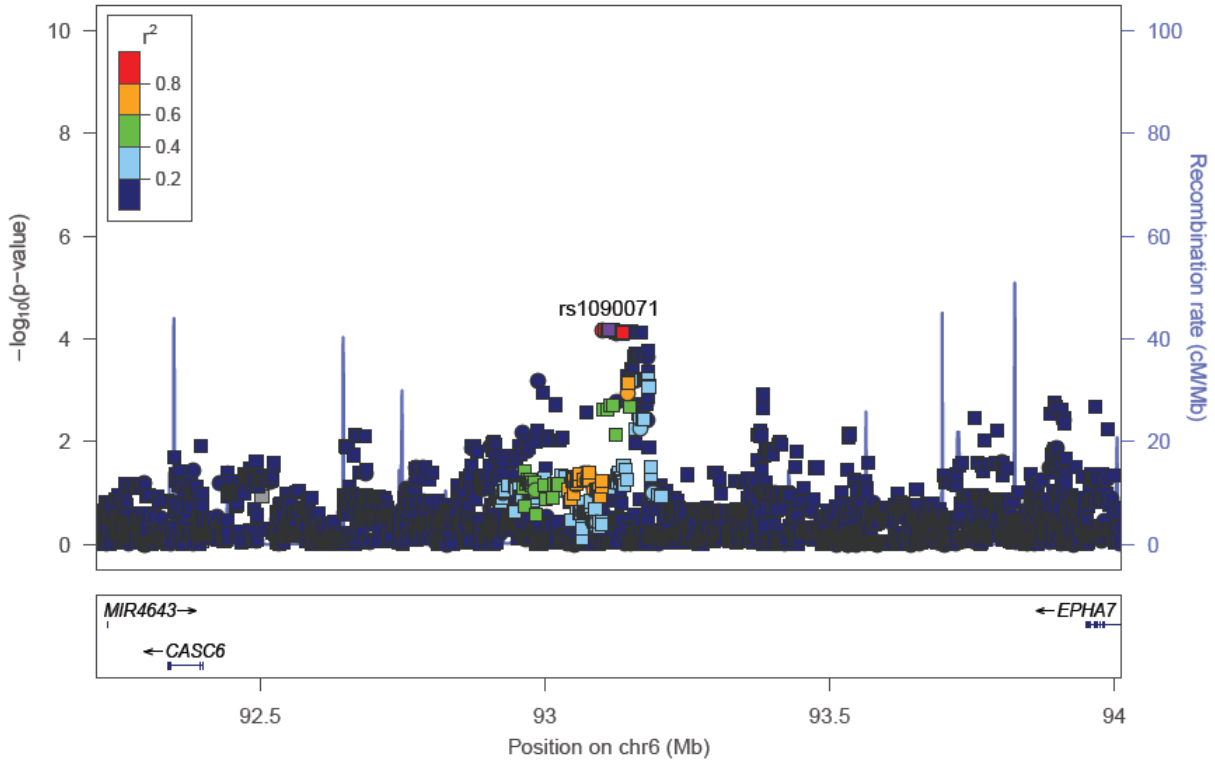
Figure 56. Regional Association Plot of rs76944100

Other associations seen at suggestive significance were not located near genes relevant to dental or oral health (Figure 57).


r²
0.8
0.6
0.4
0.2
rs13060599
−log10(p−value)
Recombination rate (cM/Mb)
←ROBO1
LOC101927374→
79.4
79.6
79.8
80
80.2
Position on chr3 (Mb)


r²
0.8
0.6
0.4
0.2
rs1487025
−log10(p−value)
Recombination rate (cM/Mb)
←LRRTM4
←LOC101927967
LOC101927926→
LOC101927948→
←SNAR−H
77.6
77.8
78
78.2
78.4
Position on chr2 (Mb)


r²
0.8
0.6
0.4
0.2
rs629476
−log10(p−value)
Recombination rate (cM/Mb)
LRP11
RAET1K
PPP1R14C
IYD
PLEKHG1
RAET1E−AS1
ULBP3
RAET1E
RAET1L
RAET1G
ULBP2
ULBP1
150.2
150.4
150.6
150.8
151
Position on chr6 (Mb)


r²
0.8
0.6
0.4
0.2
rs1090071
−log10(p−value)
Recombination rate (cM/Mb)
MIR4643
EPHA7
CASC6
92.5
93
93.5
94
Position on chr6 (Mb)

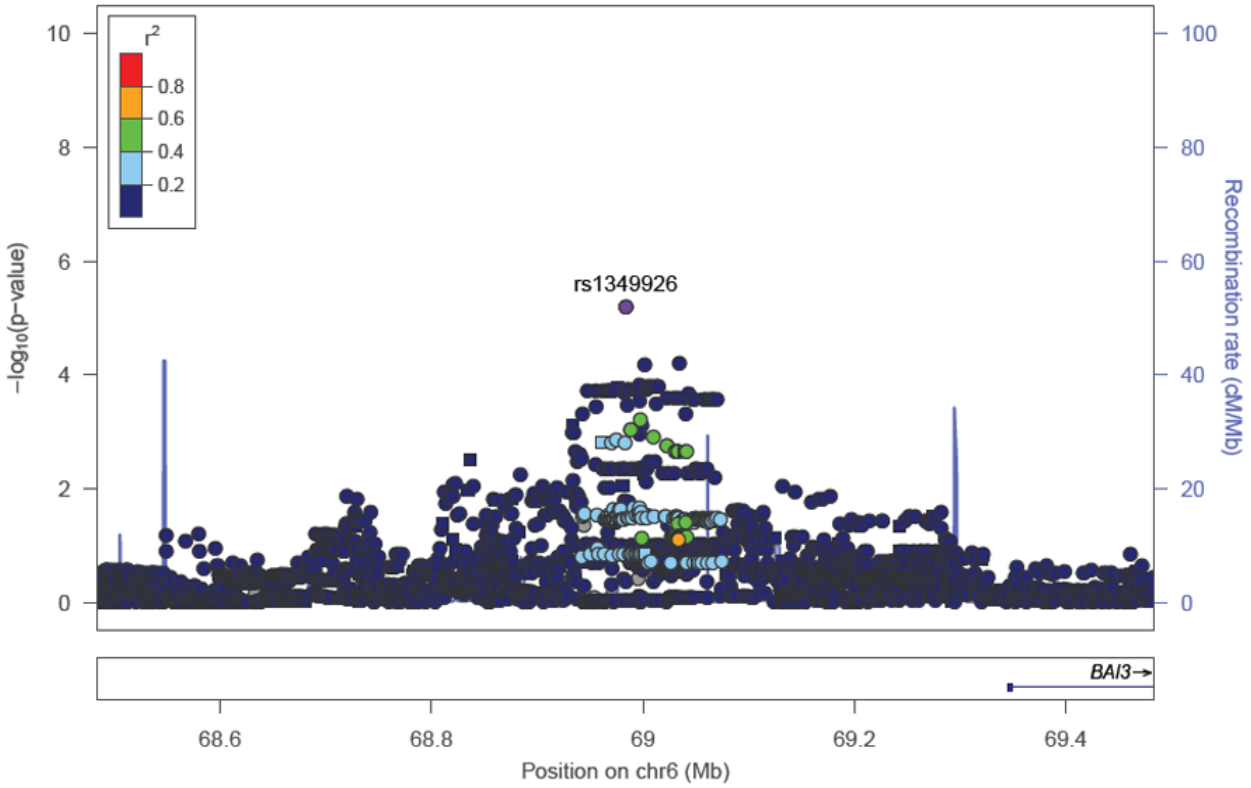


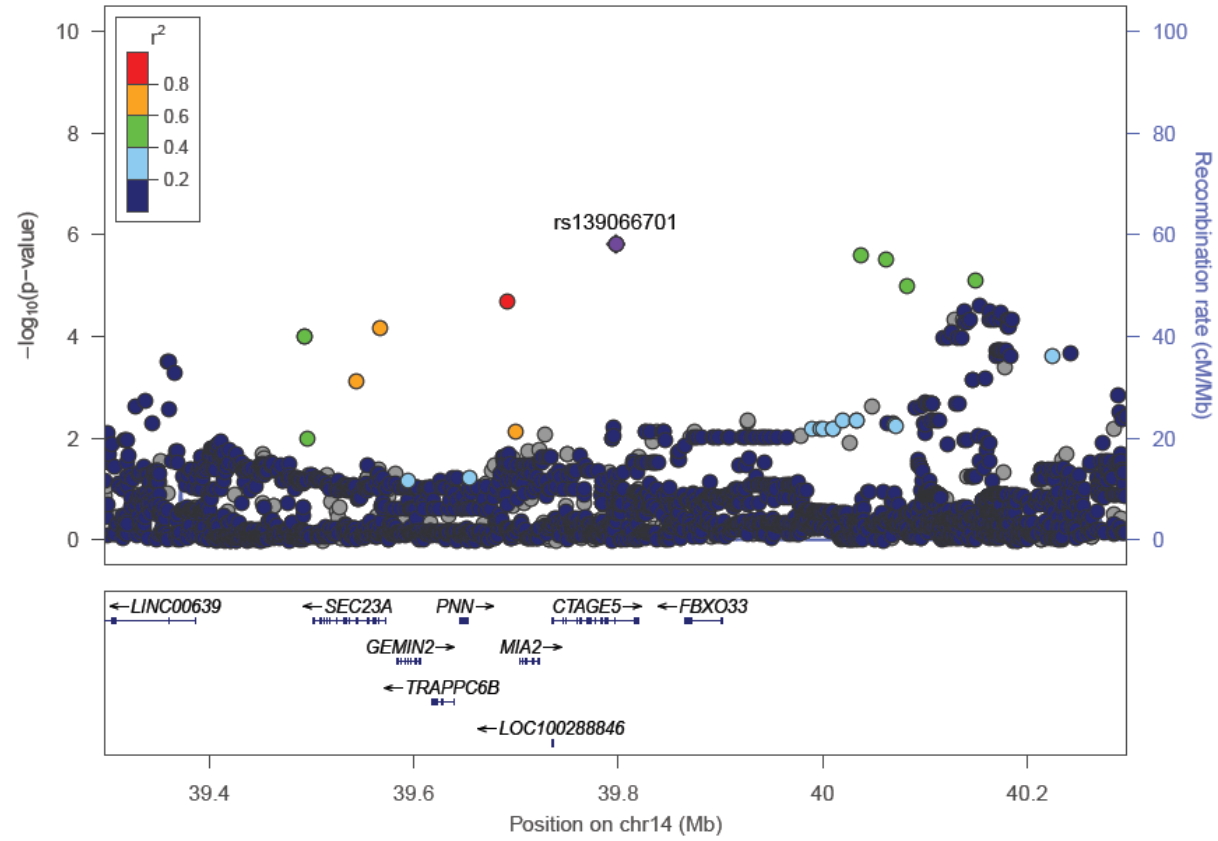
Figure 57. Regional Association Plots for other suggestively significant hits for functional dentition in COHRA

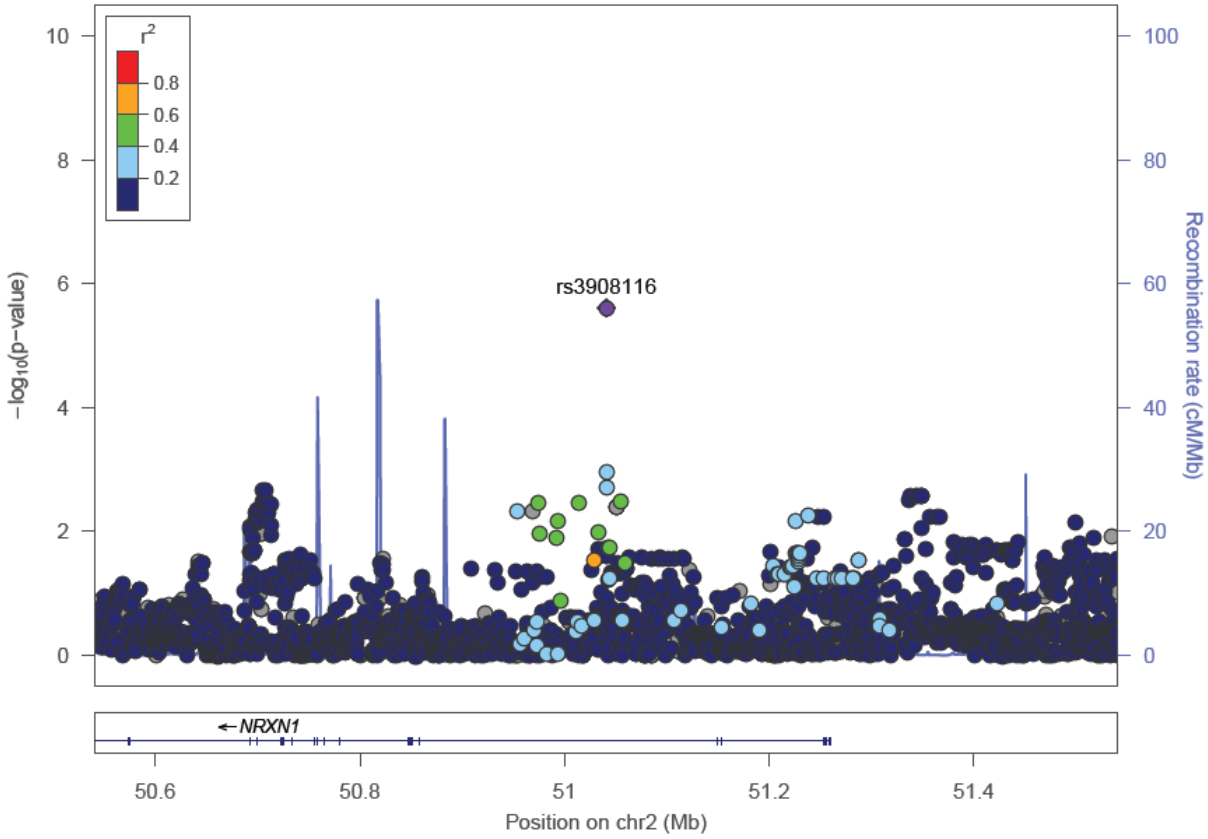
For rs1090071 annotation of imputed vs. genotyped variants is reversed, and the plot covers a larger genomic region due to their being no genes within 500 kb on either side of the index SNP.

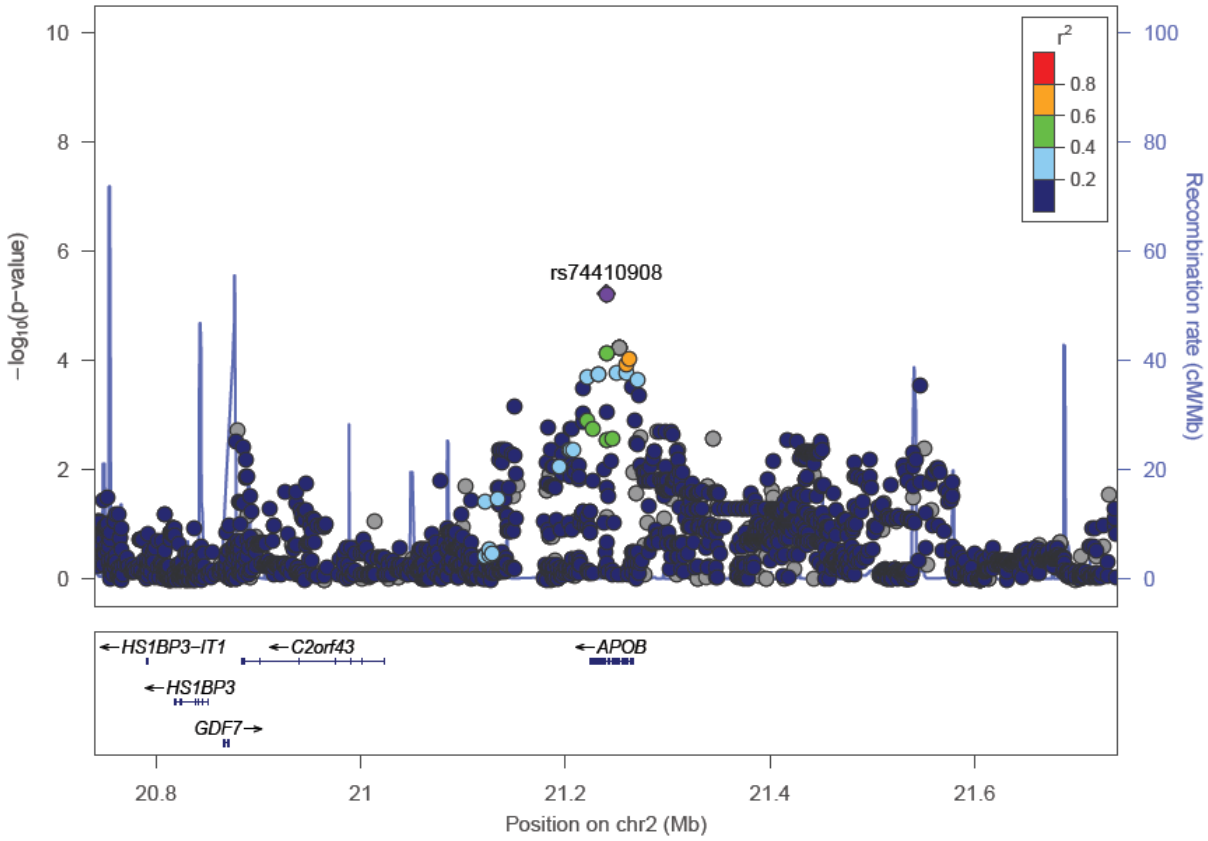
A.3 POFC-G – QUANTITATIVE TRAIT

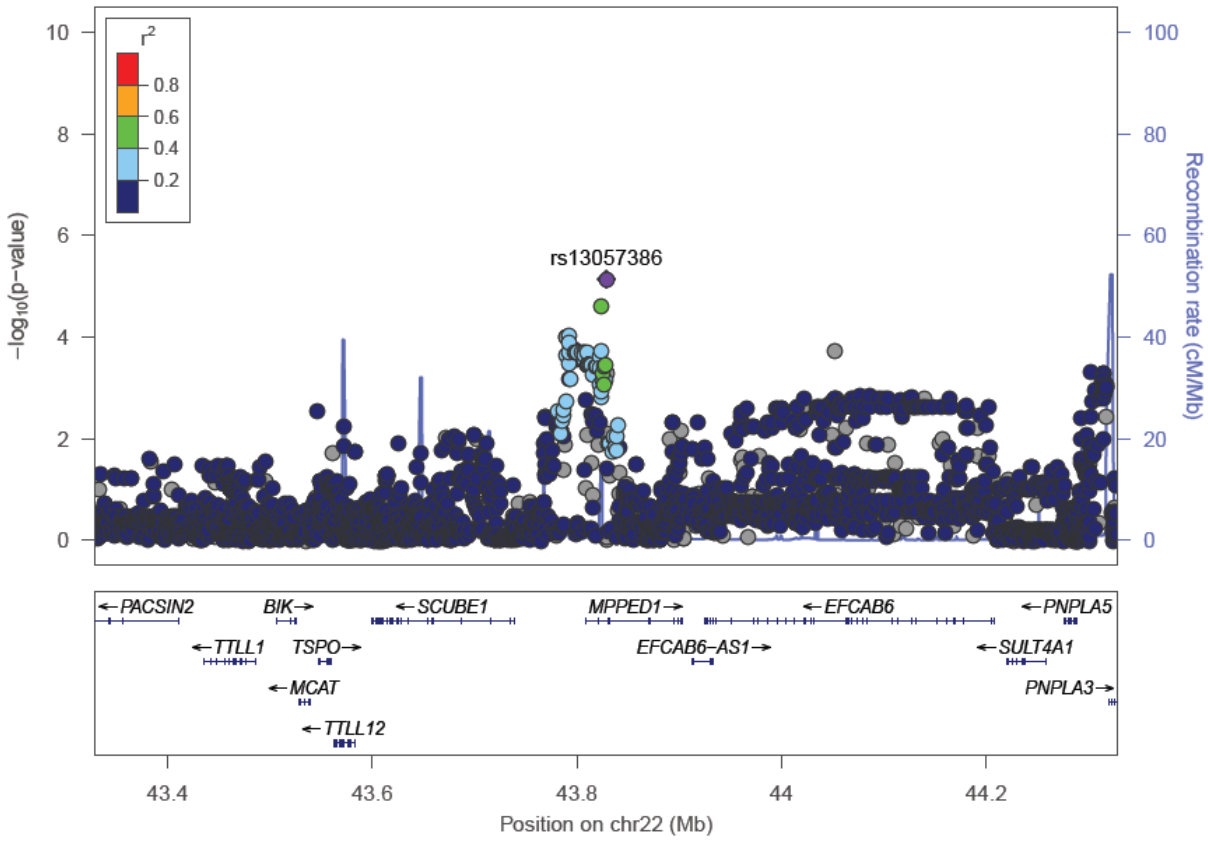
A handful of SNPs associated with the quantitative trait at suggestive significance were located near genes relevant to dental and oral health (Figure 58). rs139066701 (p = 1.5E-6) corresponds to a region also implicated in the Hispanics-only meta-analysis, which is intronic to *CTAGE5* and located near *SEC23A*. While the role of *CTAGE5* in dental health is currently unknown, *SEC23A* is targeted by TFII-I transcription factors, which are deleted in Williams Syndrome, features of which include dental and craniofacial anomalies [201]. rs3908116 (p = 2.6E-6)

corresponds to *NRXN1*, which remained significant in the final trans-ethnic meta-analysis. rs74410908 (p = 5.9E-6) is intronic to *APOB*, and levels of apoB in GCF fluid were higher at diseased sites and decreased after periodontal therapy [265]. rs13057386 (p = 7.24E-6) is intronic to *MPPED1*. This region is approximately 400 kb upstream of rs4823141, which was implicated in COHRA only, located near *EFCAB6*, *MPPED1*, and *SCUBE1*, all with possible roles in dental and oral health. Lastly, rs7100483 (p = 8.5E-6) is located approximately 350 kb upstream of *CDH23*, mutations of which have been implicated in selective mandibular incisor agenesis [237].


r²
0.8
0.6
0.4
0.2
rs139066701
−log10(p−value)
Recombination rate (cM/Mb)
LINC00639
SEC23A
PNN
CTAGE5
FBXO33
GEMIN2
MIA2
TRAPPC6B
LOC100288846
39.4
39.6
39.8
40
40.2
Position on chr14 (Mb)


r²
0.8
0.6
0.4
0.2
rs3908116
−log10(p−value)
Recombination rate (cM/Mb)
NRXN1
50.6
50.8
51
51.2
51.4
Position on chr2 (Mb)


r²
0.8
0.6
0.4
0.2
rs74410908
−log10(p−value)
Recombination rate (cM/Mb)
HS1BP3−IT1
C2orf43
APOB
HS1BP3
GDF7
20.8
21
21.2
21.4
21.6
Position on chr2 (Mb)


r²
0.8
0.6
0.4
0.2
rs13057386
−log10(p−value)
Recombination rate (cM/Mb)
PACSIN2
BIK
SCUBE1
MPPED1
EFCAB6
PNPLA5
TTLL1
TSPO
EFCAB6−AS1
SULT4A1
MCAT
PNPLA3
TTLL12
43.4
43.6
43.8
44
44.2
Position on chr22 (Mb)

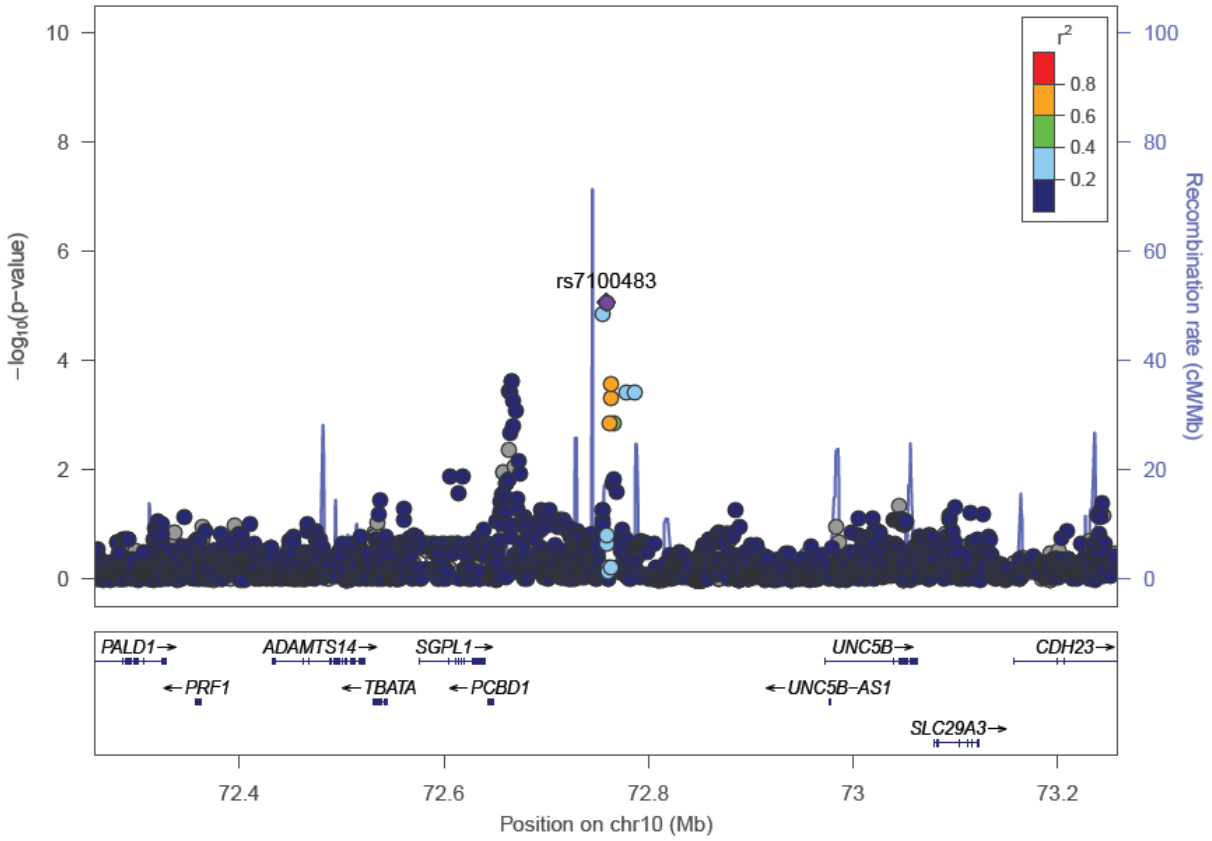


Figure 58. Regional Association Plots for suggestively significant hits in POFC-G for the quantitative trait

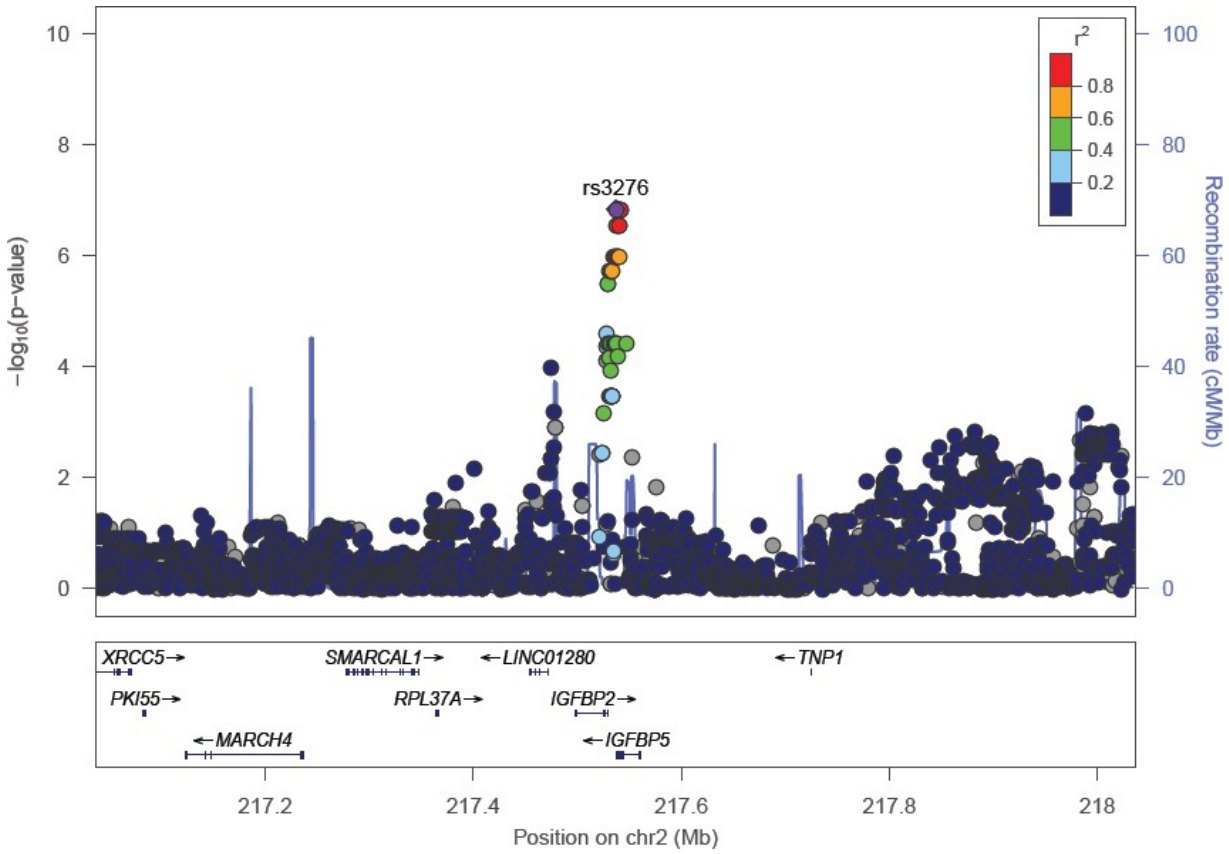
A.4 POFC-G – FUNCTIONAL DENTITION

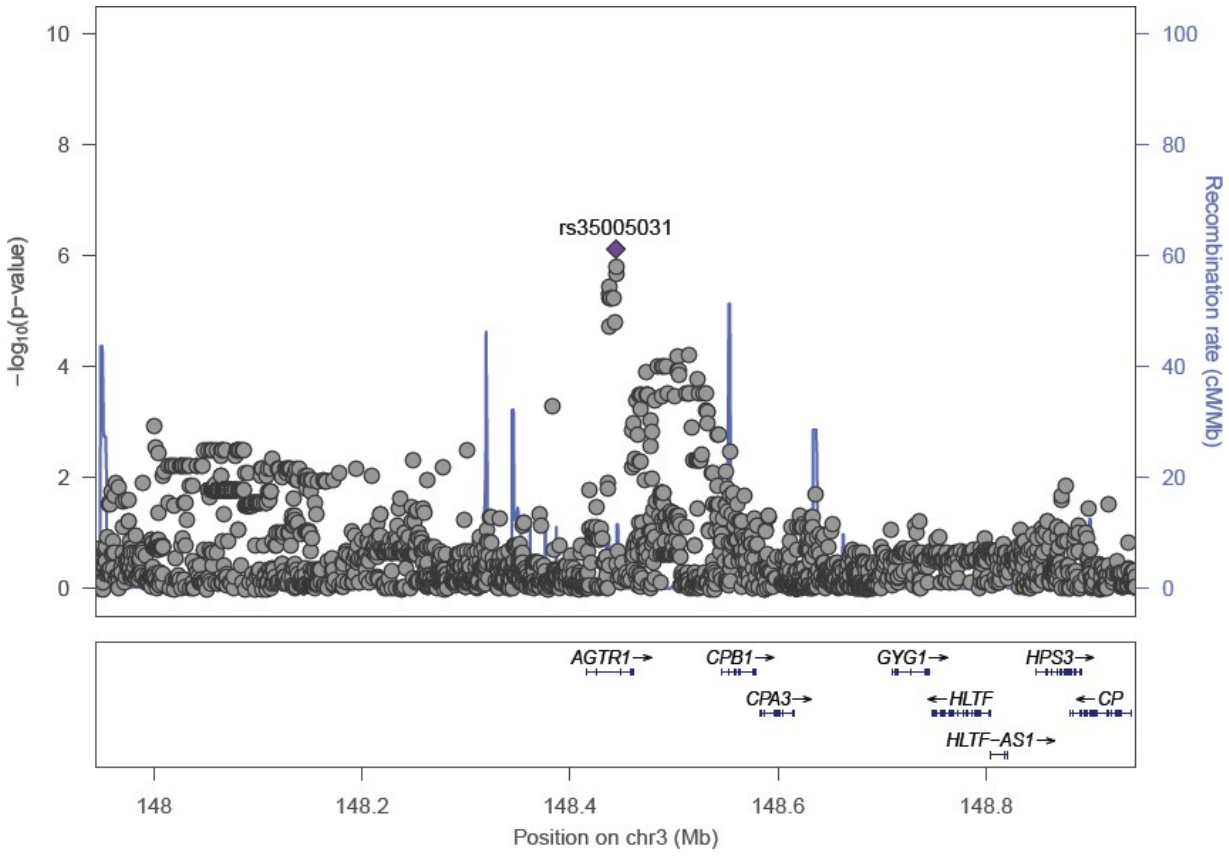
There were several suggestively significant associations relevant to dental and oral health observed in POFC-G for functional dentition that were not seen at other points in the analysis (Figure 59). rs3276 (p = 1.4E-7) is located near several genes with potential roles in dental and oral health. This index SNP is in the 3'UTR of *IGFBP5* and is also just downstream of *IGFBP2*. Both of these genes encode insulin like growth factor binding proteins and are differentially expressed during tooth development [266] and during the differentiation of dental pulp cells [267], and *IGFBP2* may regulate differentiation of dental pulp cells. *IGFBP2* levels in gingival

crevicular fluid are also correlated with periodontal disease parameters [268]. *IGFBP5* levels are decreased in diseased periodontal tissues, and expression improved periodontal tissue regeneration and decreased inflammation in an animal model [269]. *IGFBP5* was also found to be downregulated in PDL cells in response to mechanical force [270]. Also in this region is *TNP1* (approximately 200 kb upstream of rs3276), which was implicated in a GWAS of permanent tooth eruption (p = 2.16E-14) [271]. Lastly, mutations of *SMARCAL1* (approximately 200 kb upstream) cause Schimke Immuno-osseous Dysplasia, features of which include defects of tooth development, such as microdontia, hypodotia, and abnormally shaped molar teeth [272].

rs35005031 (p = 7.6E-7) is intronic to *AGTR1*, which may be involved in periodontal inflammation through the regulation of IL-6 and IL-1β [273]. This SNP is also located approximately 400 kb upstream of *HSP3*; mutations in *HPS3* are a cause of Hermansky-Pudlak Syndrome, manifestations of which include bleeding diathesis, especially gingival bleeding and excessive bleeding after tooth extraction [274]. Also in this region is *CP*, located approximately 450 kb upstream of the index SNP. *CP* encodes the protein ceruloplasmin, which binds copper in plasma; increased levels of ceruloplasmin is associated with periodontitis [275, 276].

rs79982195 (p = 8.3E-7) is intronic to *MPPED2*, has previously been associated with dental caries in whites [59, 63, 126]; this gene may play a role in periodontal pathogen colonization [277]. rs12542122 (p = 4.9E-6) is intronic to *SULF1*, a heparan sulfate endosulfatase that is differentially expressed during tooth development, and *Sulf1/Sulf2* double knockout mice show defects in tooth morphology (dentin hypoplasia and short tooth roots) [278].


rs3276
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
XRCC5
SMARCAL1
LINC01280
TNP1
PKI55
RPL37A
IGFBP2
MARCH4
IGFBP5
217.2
217.4
217.6
217.8
218
Position on chr2 (Mb)


rs35005031
−log10(p−value)
Recombination rate (cM/Mb)
AGTR1
CPB1
GYG1
HPS3
CPA3
HLTF
CP
HLTF−AS1
148
148.2
148.4
148.6
148.8
Position on chr3 (Mb)

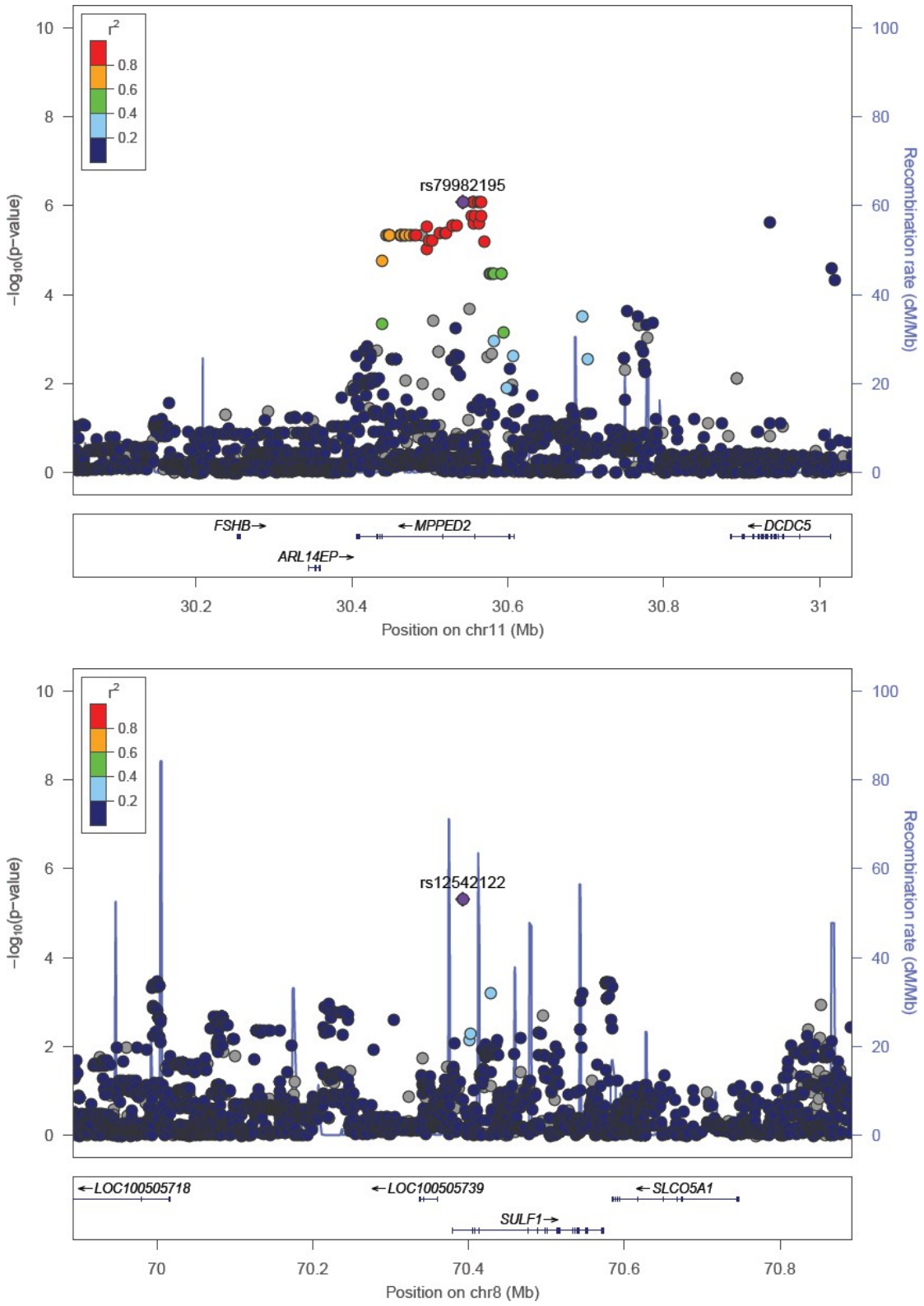
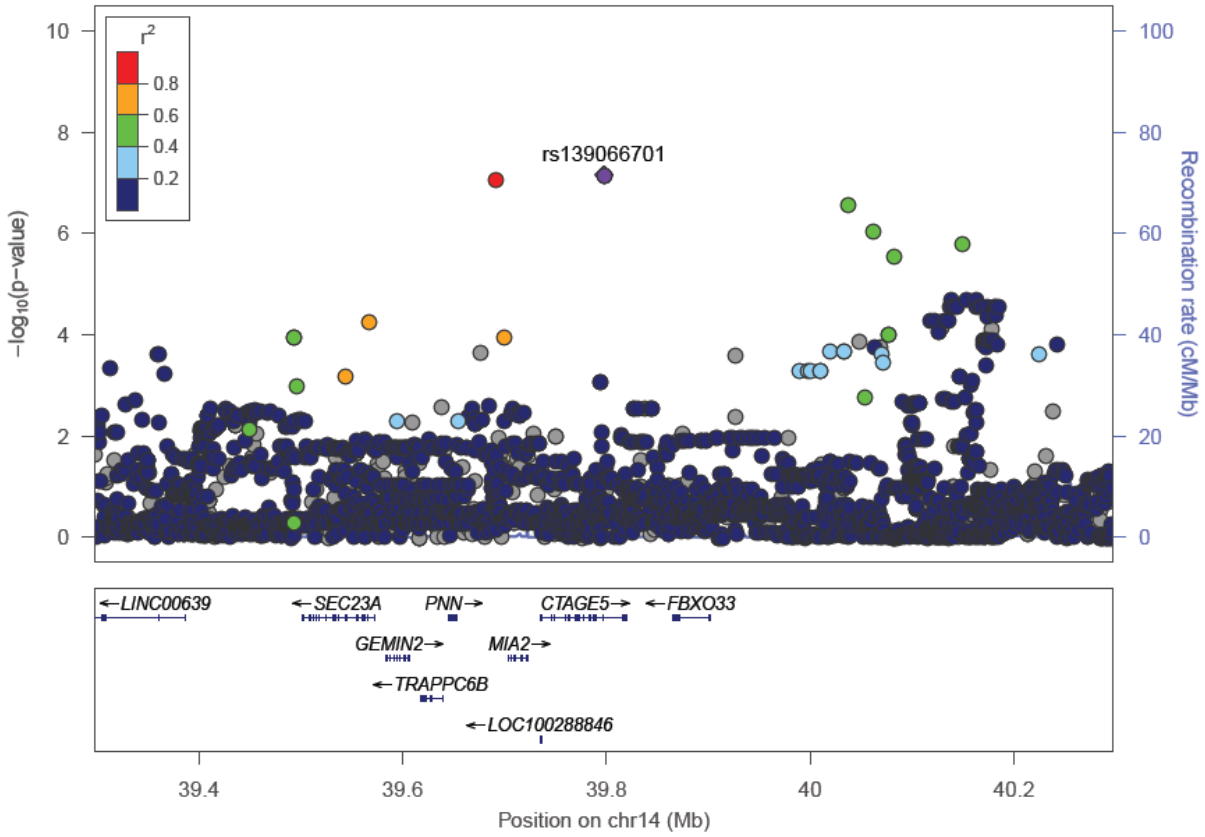
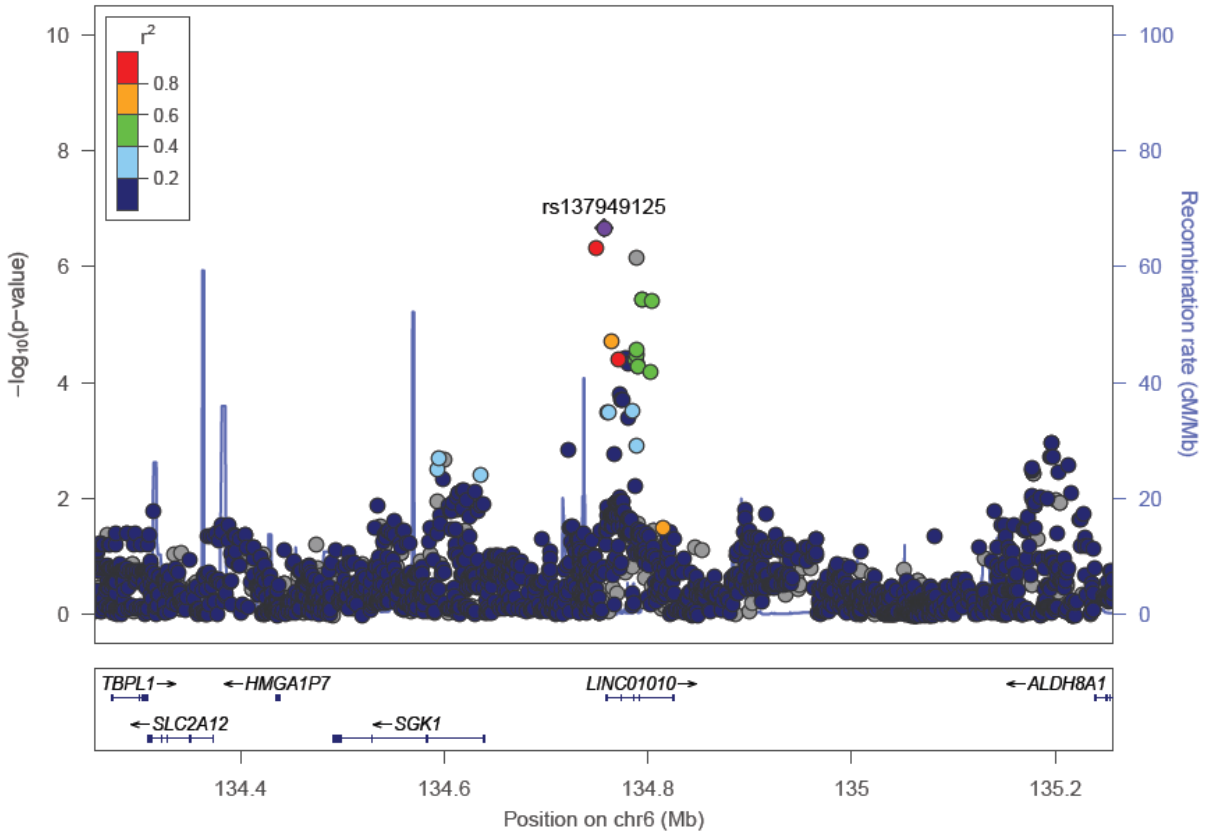


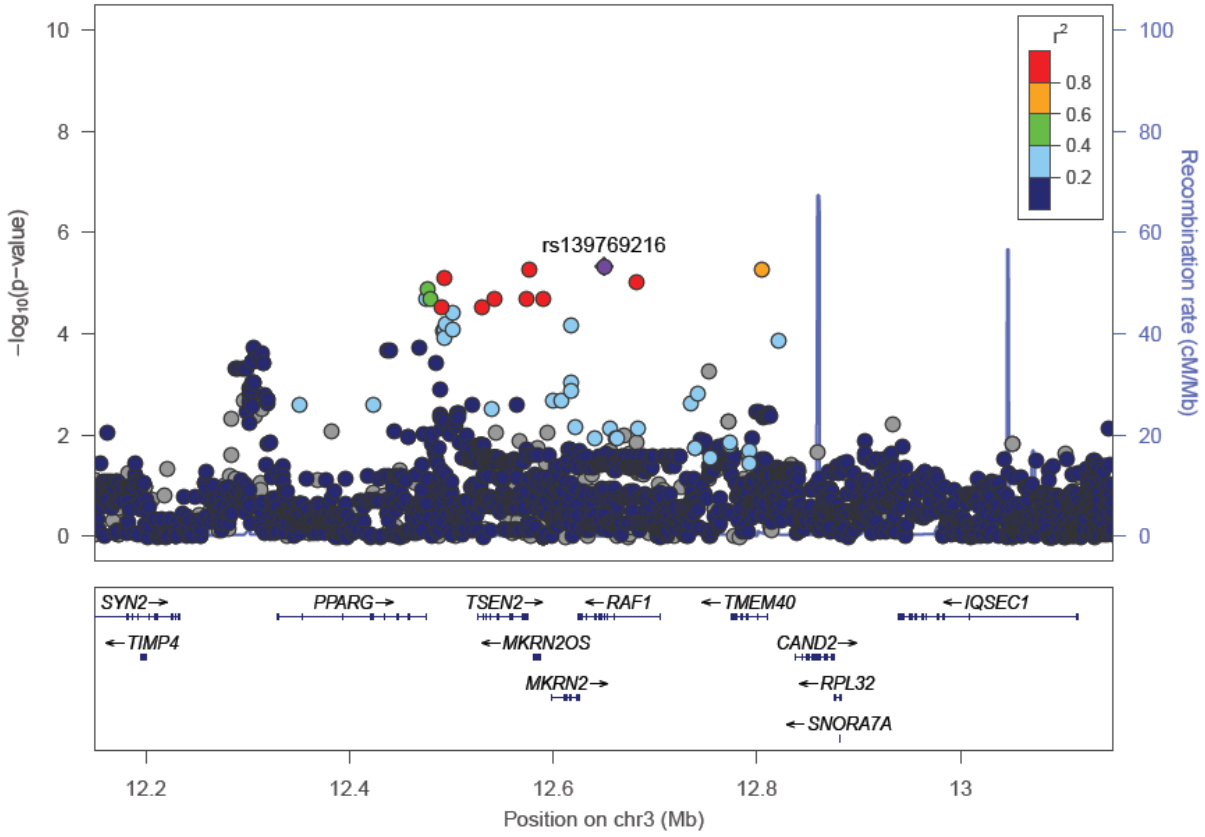
Figure 59. Regional Association Plots for suggestively significant hits in POFC-G for functional dentition

LD information was not available for all plots

Several other suggestively significant signals corresponded to those seen in the Hispanics-only meta-analysis (Figure 60). rs139066701, located near *CTAGE5* and *SEC23A*, was again a significant hit (p = 6.9E-8). rs137949125 (p = 2.1E-7) corresponds to the same signal seen in the Hispanics meta-analysis near *SGK1*, which is part of an anti-inflammatory pathway that inhibits Toll-like receptor mediated inflammation [195]. rs139769216 (p = 4.7E-6) is the same SNP implicated in the Hispanics meta-analysis near *RAF1*, which is important for tooth development [189, 190]. rs8009351 (p = 6.4E-6) was also a hit for the Hispanics-only meta-analysis; this SNP is located near *PAX9*, a major gene implicated in tooth agenesis [85], and *NKX2-1*, which is involved in tooth development as knockout mice show defects in tooth morphology [200].




r²
0.8
0.6
0.4
0.2
rs137949125
−log10(p−value)
Recombination rate (cM/Mb)
TBPL1
HMGA1P7
LINC01010
ALDH8A1
SLC2A12
SGK1
Position on chr6 (Mb)


r²
0.8
0.6
0.4
0.2
rs139769216
−log10(p−value)
Recombination rate (cM/Mb)
SYN2
PPARG
TSEN2
RAF1
TMEM40
IQSEC1
TIMP4
MKRN2OS
CAND2
MKRN2
RPL32
SNORA7A
Position on chr3 (Mb)

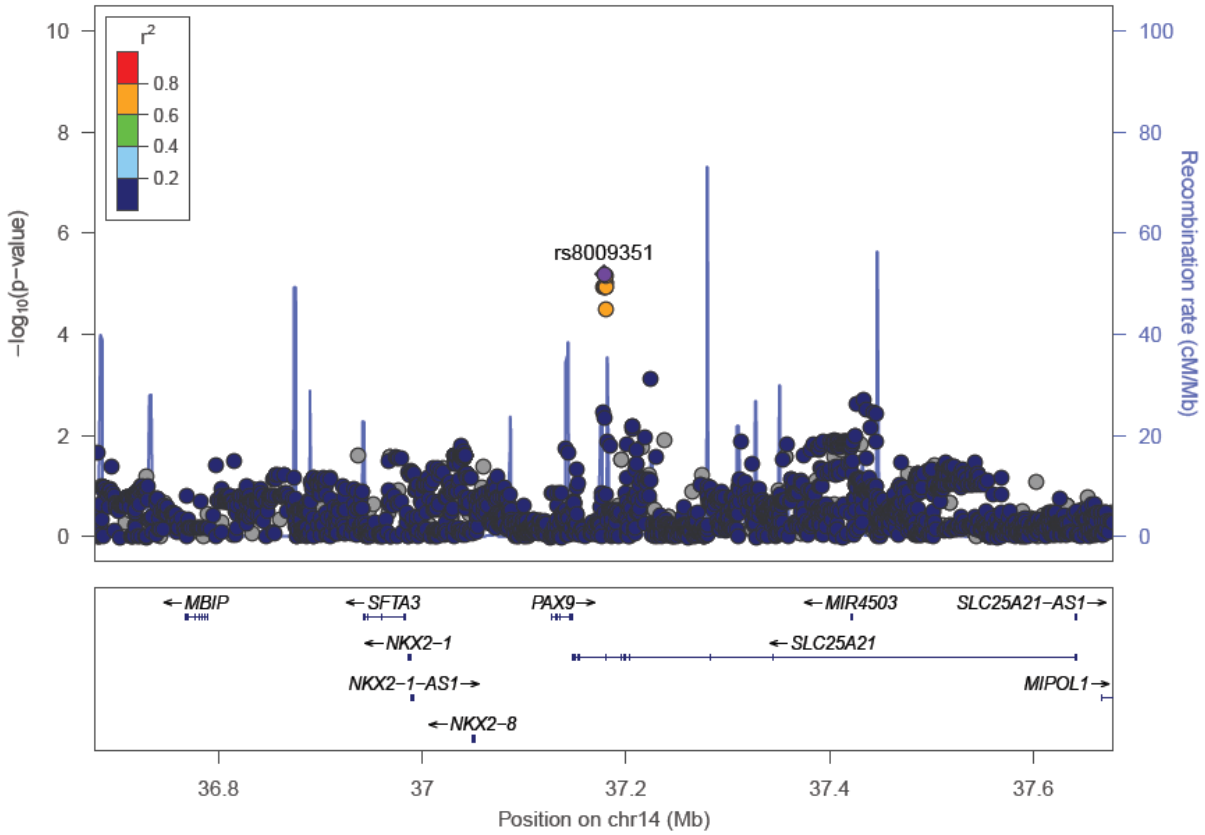


Figure 60. Regional Association Plots for other suggestively significant hits in POFC-G for functional dentition

APPENDIX B: META-ANALYSIS RESULTS

Other regions of the genome implicated in the meta-analyses are described in this section, including their Regional Association Plots.

B.1 WHITES-ONLY META-ANALYSIS RESULTS – QUANTITATIVE TRAIT

rs9750906 (p = 1.77E-06) showed evidence of association in both COHRA and DRDR (p = 4.8E-05 and 0.01, respectively). However, this corresponds to a genomic region containing only an uncharacterized genomic element (*LOC101927907*) and *LRRTM4*, which has no known role in dental health and is located approximately 200 kb downstream of this signal (Figure 61).

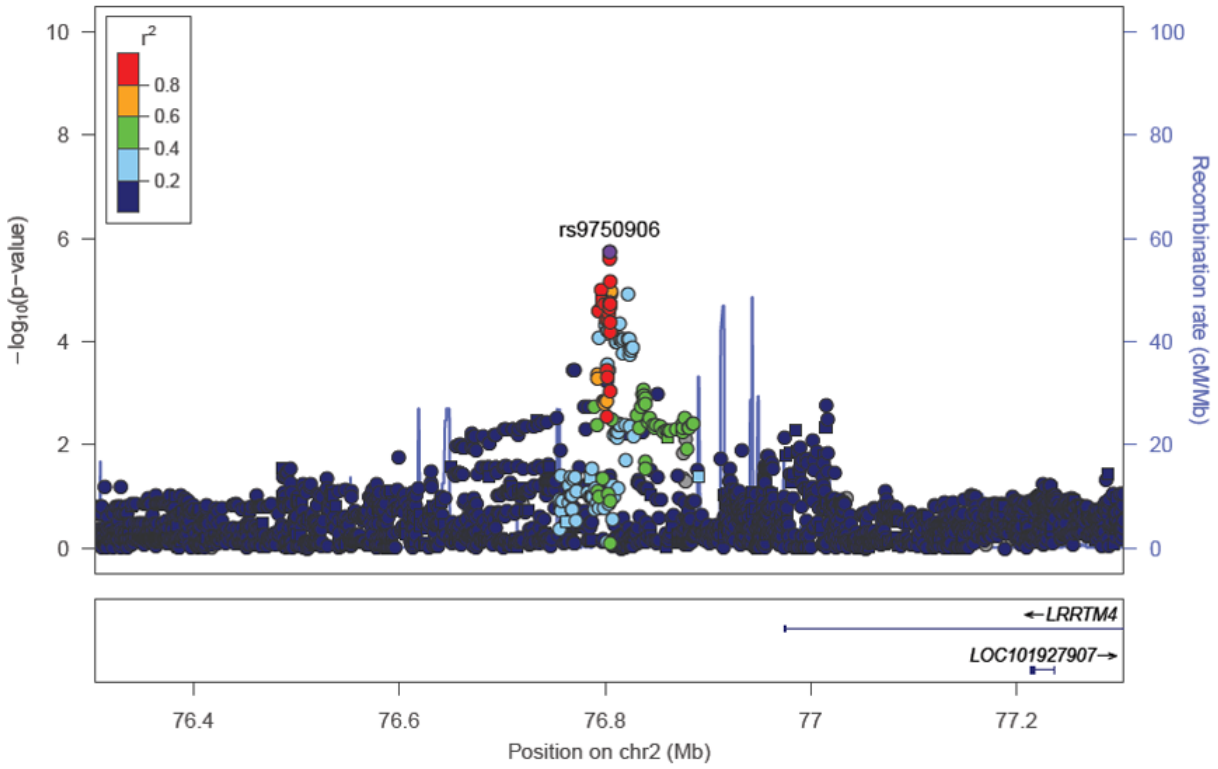


Figure 61. Regional Association Plot of rs9750906

rs11704818 (p = 6.7E-6) approximately 600 kb upstream of *FAM195A* (Figure 62), which may play a role in immune response and inflammation, as it is a chemokine that can activate macrophages and stimulate their migration, though it is supposedly a brain-specific chemokine [279]. Without information on the expression of *FAM195A* in oral tissues it is unclear whether it may play a role in periodontal inflammation and tooth loss.

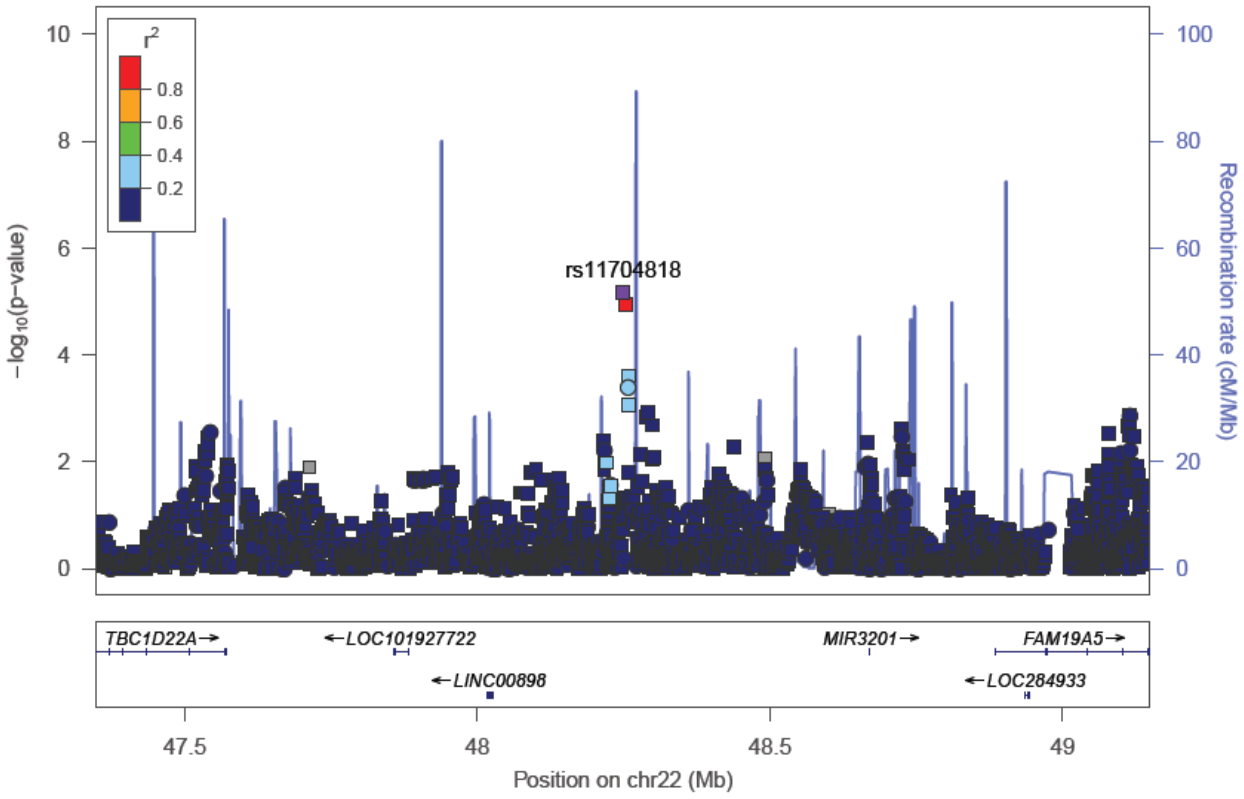


Figure 62. Regional Association Plot of rs11704818

Squares and circles indicate imputed and genotyped variants, respectively.

rs77332164 (p = 7.6E-6) is located just upstream of *ADCYAP1R1* (Figure 63), an adenylate cyclase activating polypeptide receptor. While this gene has not been studied in dental and oral health, it is yet another hit for genes that interact with adenyl cyclases and potentially cAMP signaling, processes that are implicated in periodontal disease. This index SNP also lies just downstream of *AQP1*, which encodes aquaporin 1, a transmembrane water channel protein [280]. During tooth development, *AQP1* is expressed specifically in endothelial cells of capillary vessels of developing dental tissues [236], and increased levels of *AQP1* in gingival tissue are positively correlated with the severity of periodontal disease [281]. *AQP1* may also be involved in periodontal tissue remodeling in response to experimental tooth movement [282] and in mechanoreception in the PDL [283].

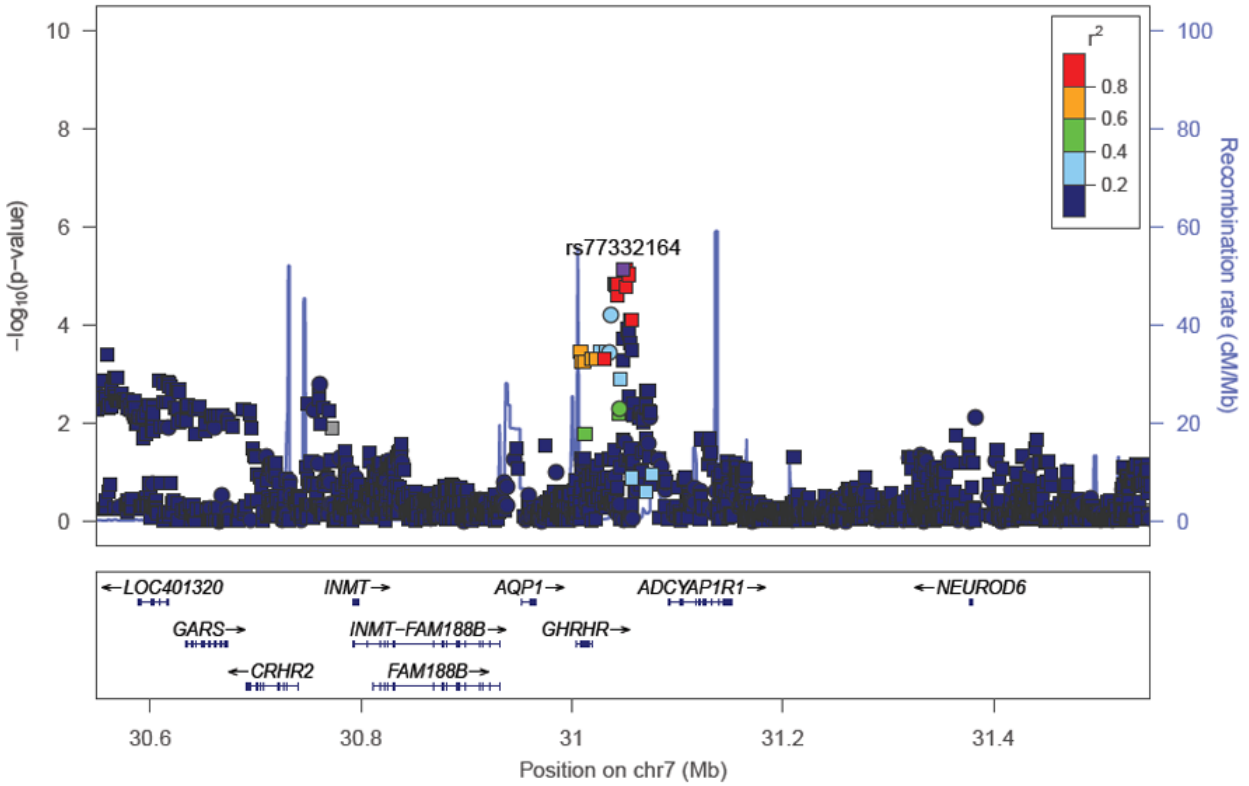


Figure 63. Regional Association Plot of rs77332164

Squares and circles indicate imputed and genotyped variants, respectively.

B.2 WHITES-ONLY META-ANALYSIS RESULTS – FUNCTIONAL DENTITION

rs13081582, the second most significant SNP (p = 2.4E-6), has a broad association signal extending over 500 kb due to extended LD (Figure 64). However, this is also a gene desert, containing only one long intergenic noncoding RNA. There is no supporting evidence that this region is involved in tooth loss.

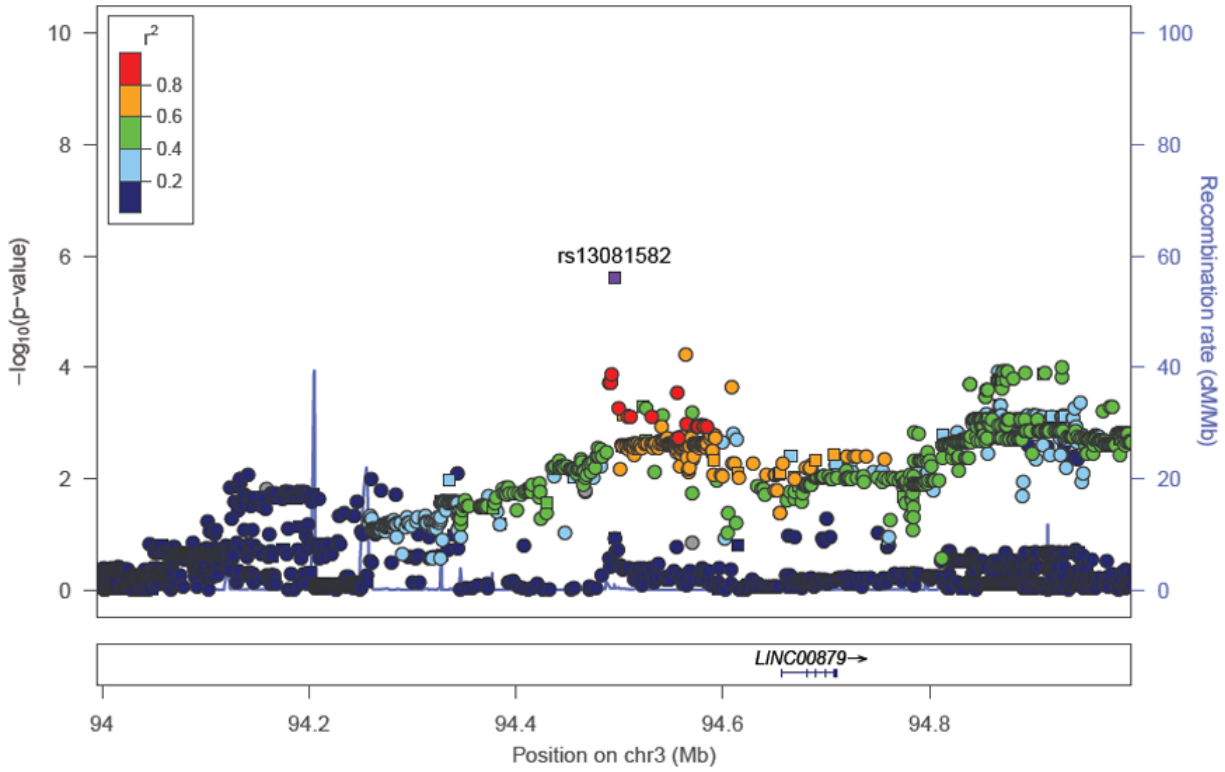


Figure 64. Regional Association Plot of rs13081582

rs1090071 (p = 6.8E-6) is located in a gene desert on chromosome 6 (Figure 65). The nearest gene downstream of this signal is *CASC6*, cancer susceptibility 6, whose function is relatively unknown. Upstream of this signal is *EPHA7*, which was implicated in a GWAS of dental caries in the primary dentition [59]. However, this signal is located approximately 1 million BP downstream of *EPHA7* and is therefore unlikely to affect *EPHA7* expression or function. At this time it is unknown how variation within this region may influence oral health and tooth loss.

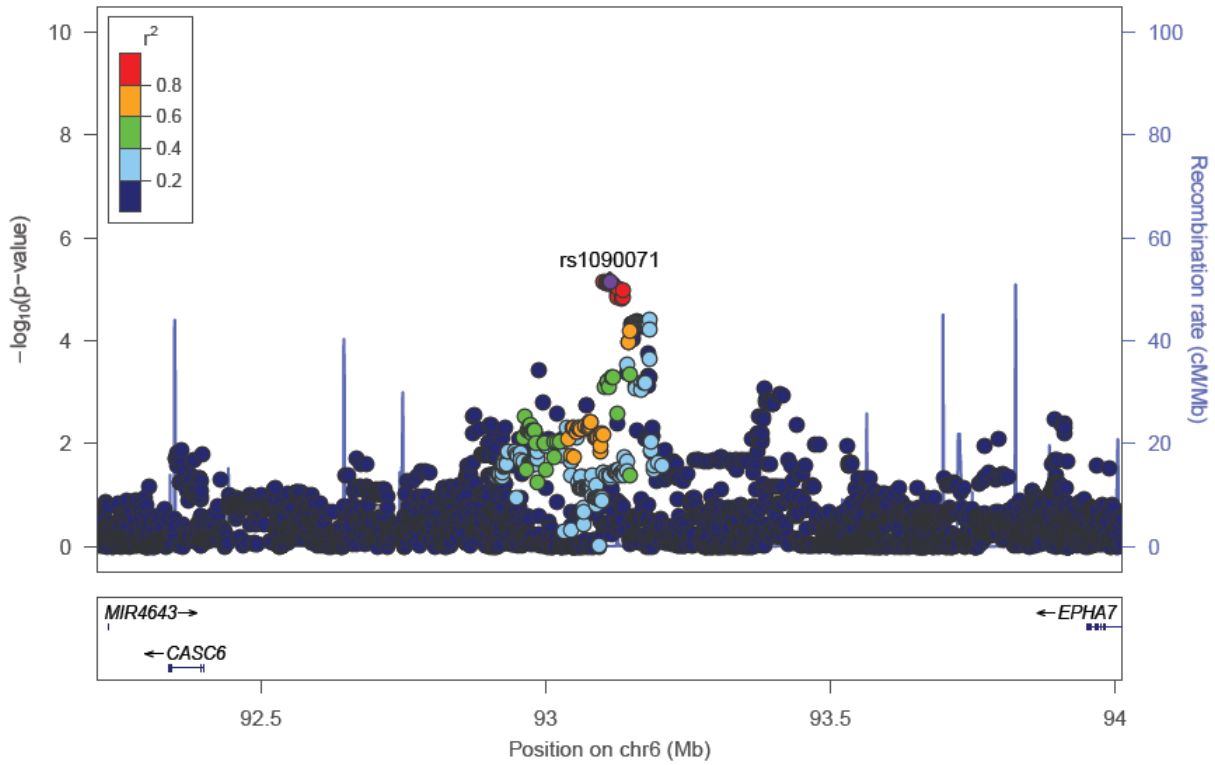


Figure 65. Regional Association Plot of rs1090071

rs875142 (p = 8.5E-6) is intronic to *PAQR8* (Figure 66). There is no known role of *PAQR8* in oral health, but this SNP is also located about 100 kb upstream of *IL17F*, interleukin 17F. IL-17F can promote inflammatory cytokines production through nuclear factor-kappa beta (NF-κB) signaling, and expression of IL-17F mRNA is higher in gingival tissue affected by periodontitis compared to controls [284]. *Il17f* knockout mice show more inflammatory activity and a greater susceptibility to alveolar bone loss following infection by *P. gingivalis*, and IL-17F levels were higher in gingival crevicular fluid of controls compared to periodontitis patients [285]. These findings indicate that proper expression and function of IL-17F may contribute to the pathogenesis of periodontal disease, most likely because of its inflammatory activity. This region also contains *IL17A*, which is similarly implicated in periodontitis [284] and

polymorphisms of *IL17A* have been associated with periodontal disease [286, 287]. However, this region is less likely to have an effect on *IL17A* as it is located upstream of rs875142, compared to *IL17F*, which is just downstream of rs875142.

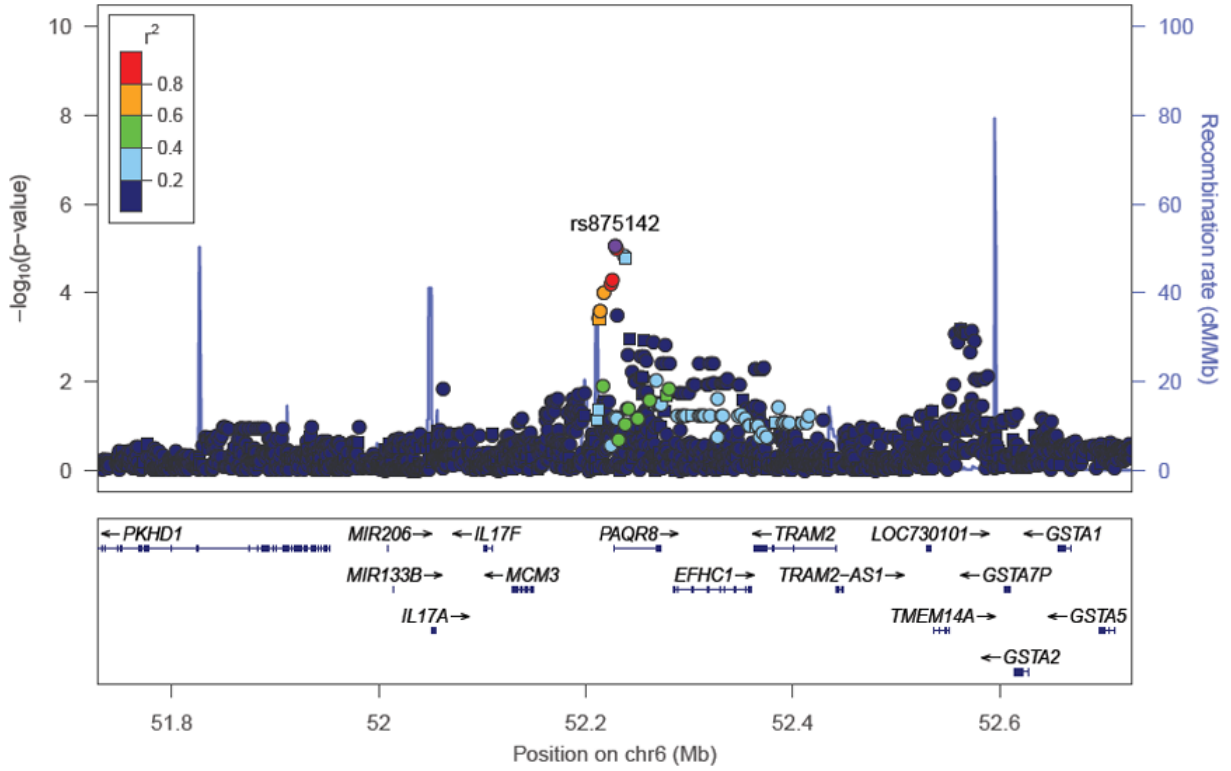
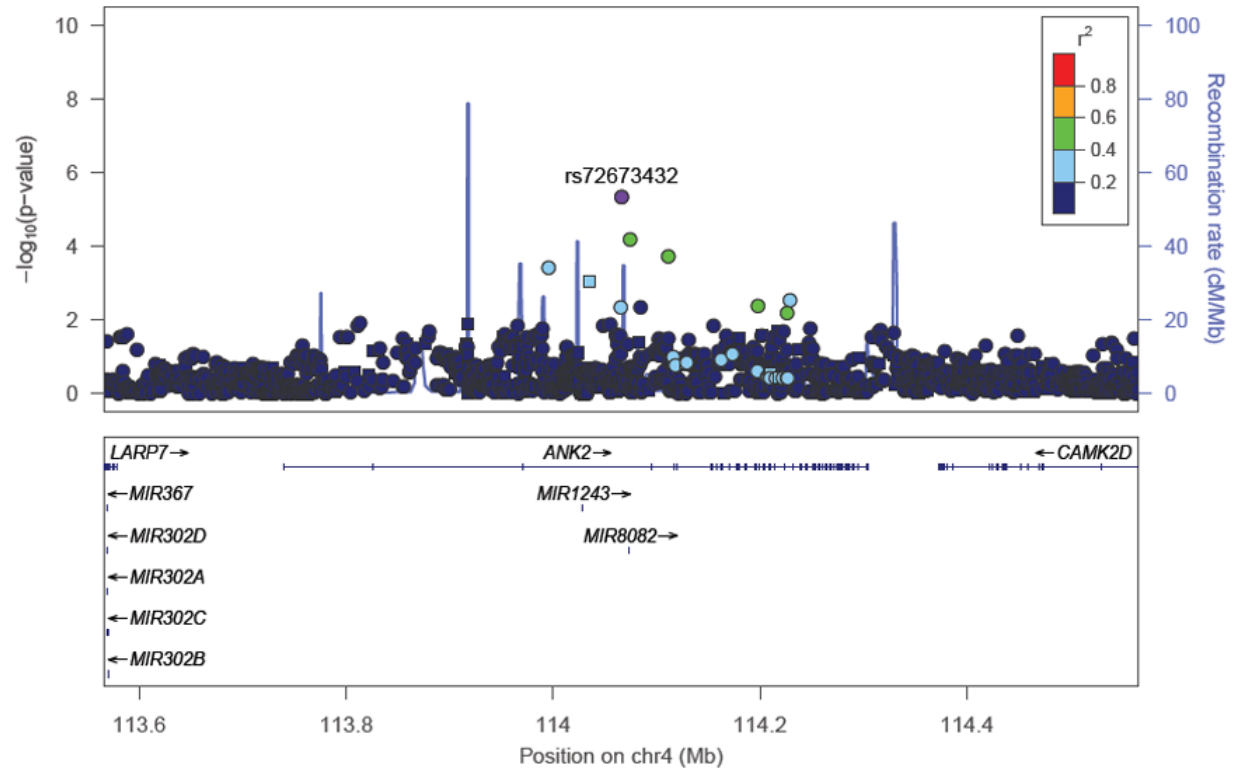
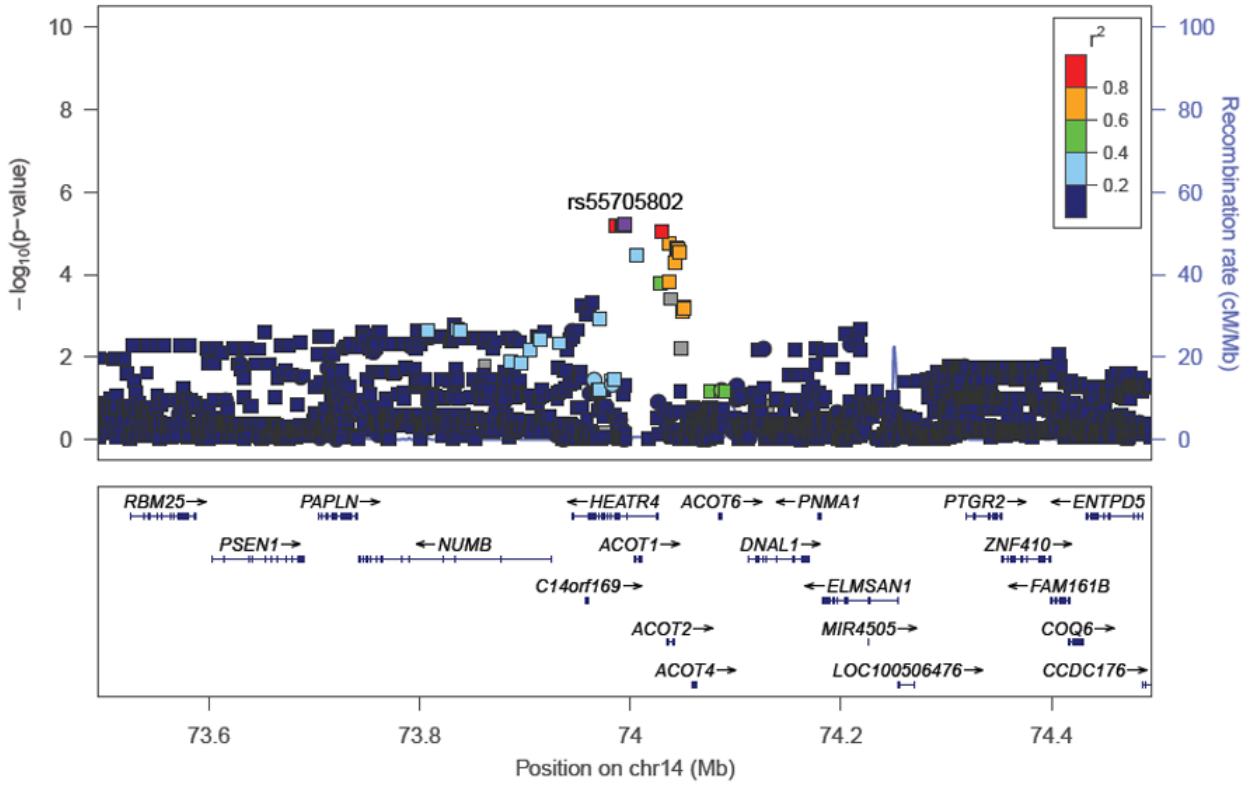


Figure 66. Regional Association Plot of rs875142

As with the quantitative trait, there were several top hits in this meta analysis of functional dentition whose statistical significance were mainly driven by the association in COHRA and were not significant in DRDR (Figure 67). These include rs72673432 near in *ANK2* (p_{meta} = 4.7E-6, p_{COHRA} = 8.6E-7, p_{DRDR} = 0.72), rs55705802 near *NUMB* (p_{meta} = 5.9E-6, p_{COHRA} = 1.4E-5, p_{DRDR} = 0.15), and rs600411 (not discussed) (p_{meta} = 8.8E-6, p_{COHRA} = 4.2-5, p_{DRDR} = 0.08).


rs72673432
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
LARP7→
ANK2→
←CAMK2D
←MIR367
MIR1243→
←MIR302D
MIR8082→
←MIR302A
←MIR302C
←MIR302B
113.6
113.8
114
114.2
114.4
Position on chr4 (Mb)


rs55705802
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
RBM25→
PAPLN→
←HEATR4
ACOT6→
←PNMA1
PTGR2→
←ENTPD5
PSEN1→
←NUMB
ACOT1→
DNAL1→
ZNF410→
C14orf169→
←ELMSAN1
←FAM161B
ACOT2→
MIR4505→
COQ6→
ACOT4→
LOC100506476→
CCDC176→
73.6
73.8
74
74.2
74.4
Position on chr14 (Mb)

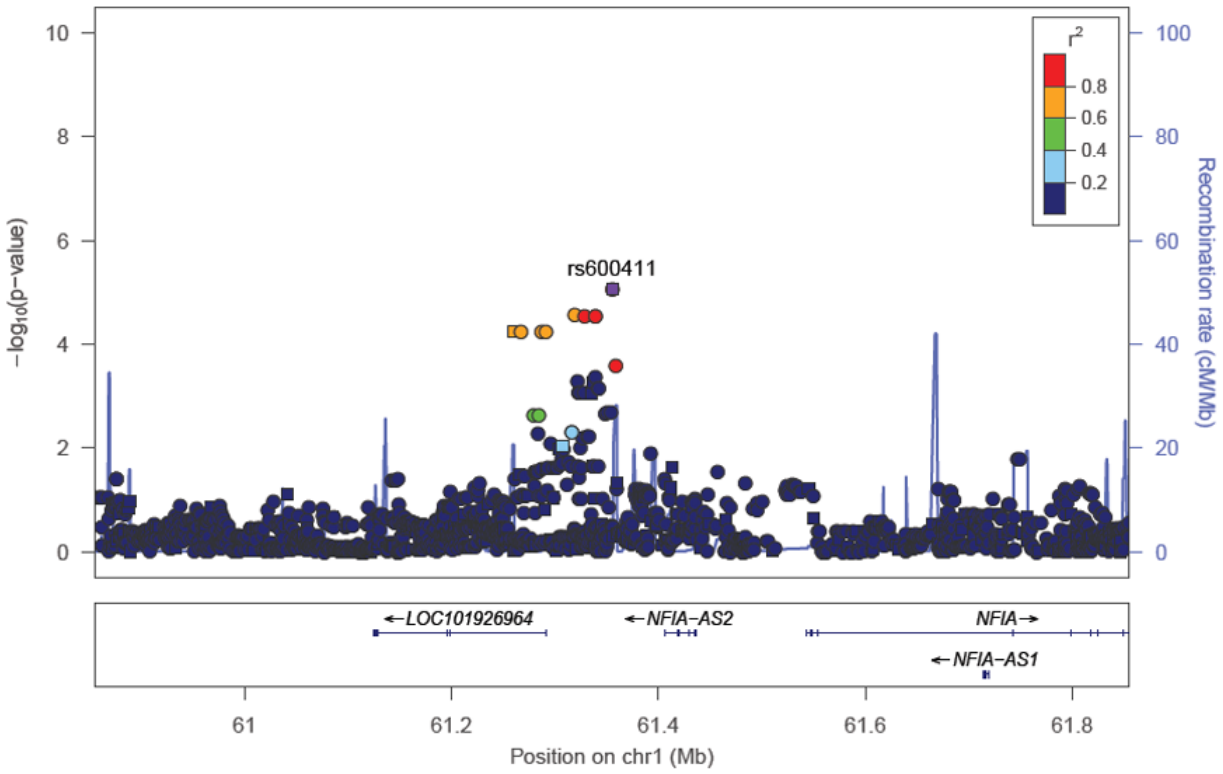
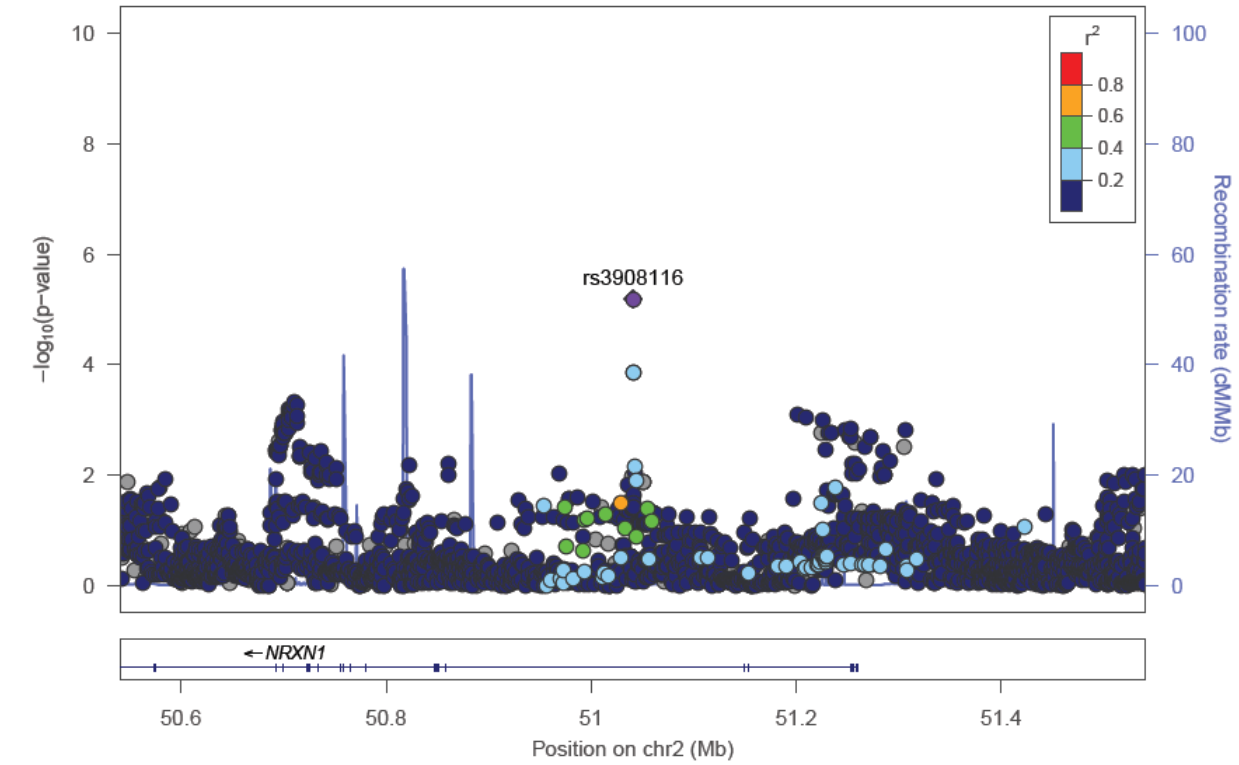


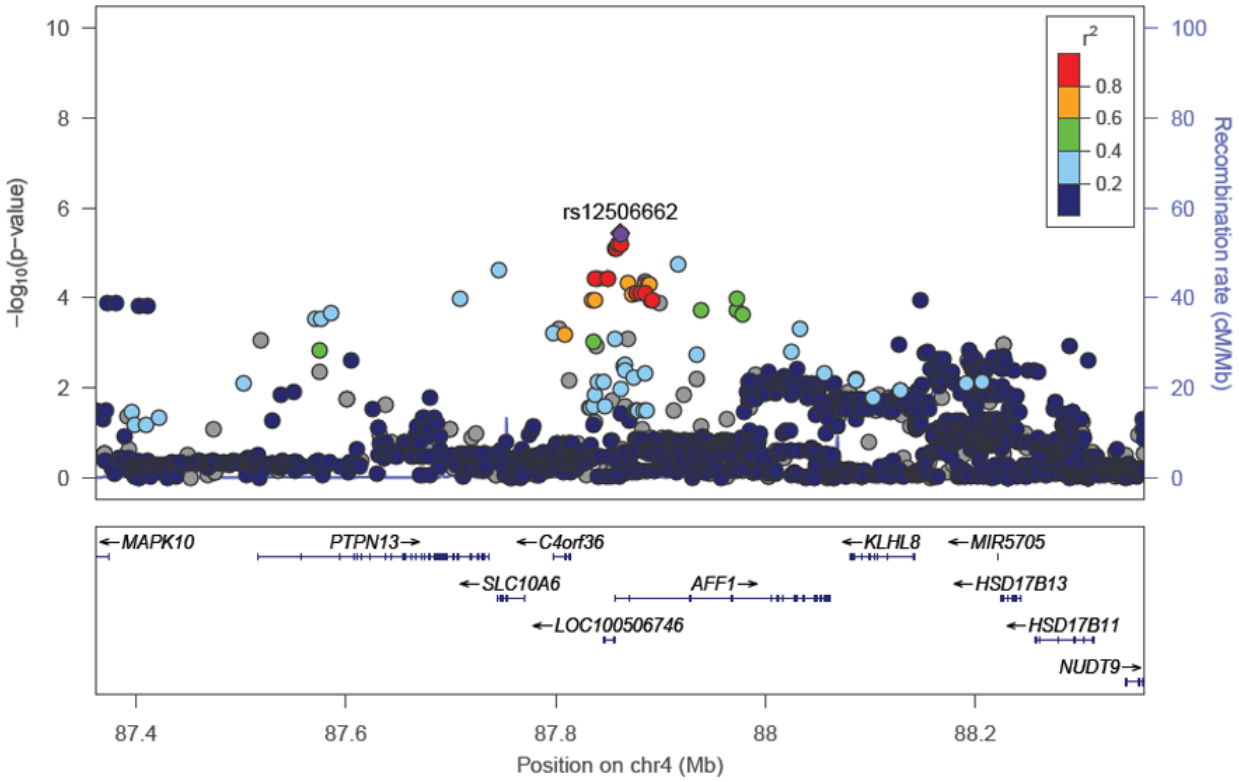
Figure 67. Regional Association Plots of rs72673432, rs55705802, and rs600411

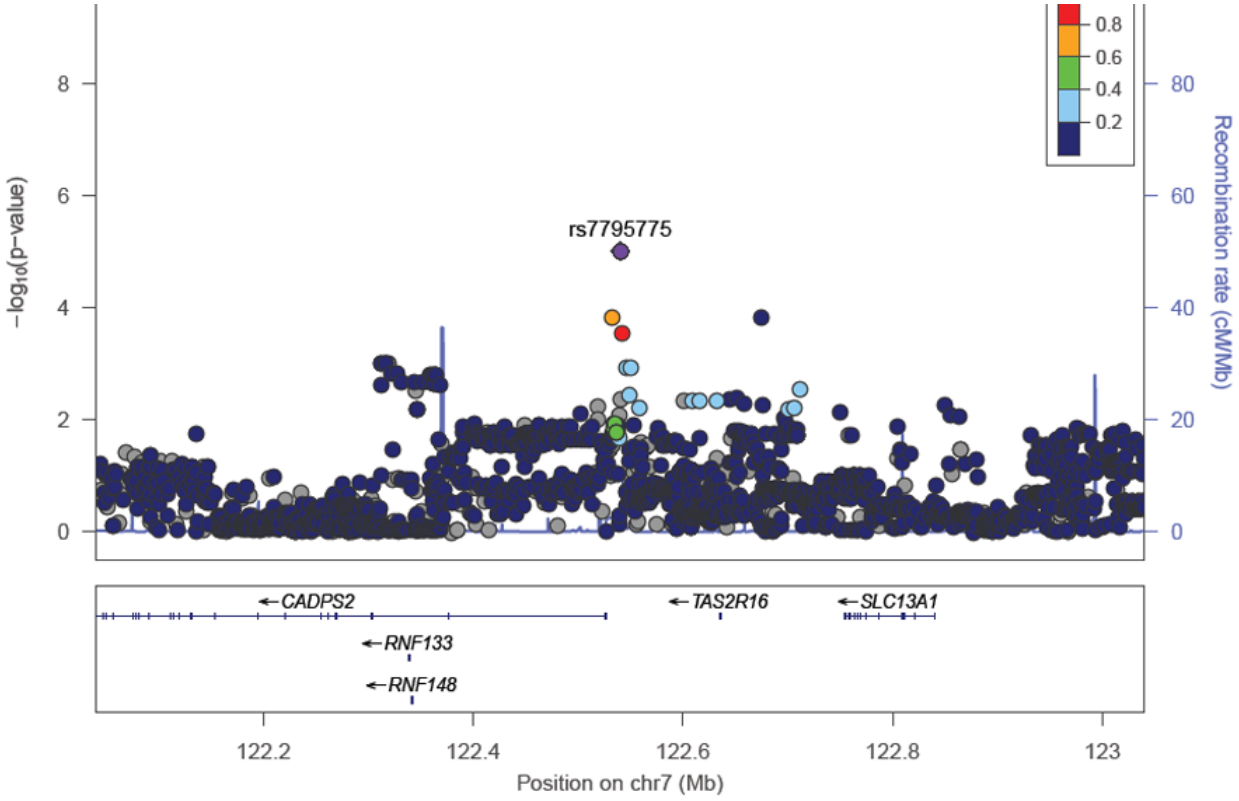
For rs55705802, annotation of imputed vs. genotyped variants is reversed

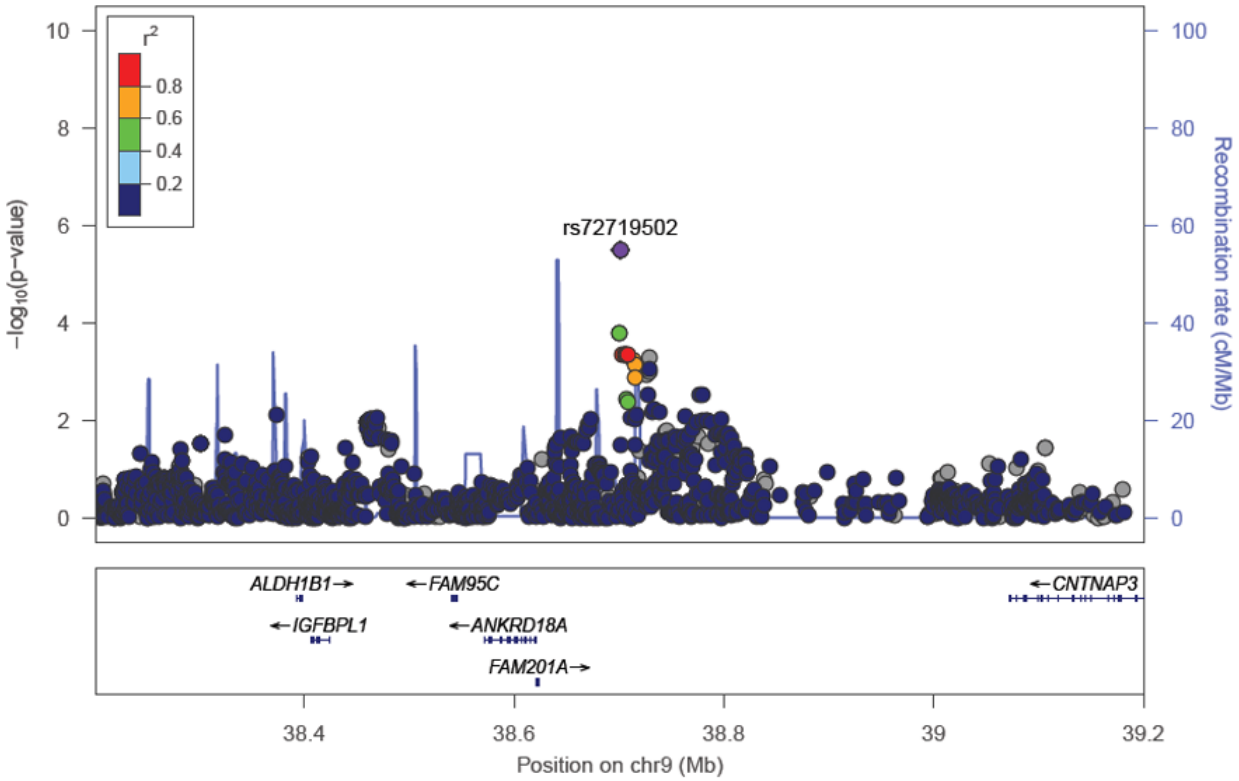
B.3 HISPANICS-ONLY META-ANALYSIS RESULTS – QUANTITATIVE TRAIT

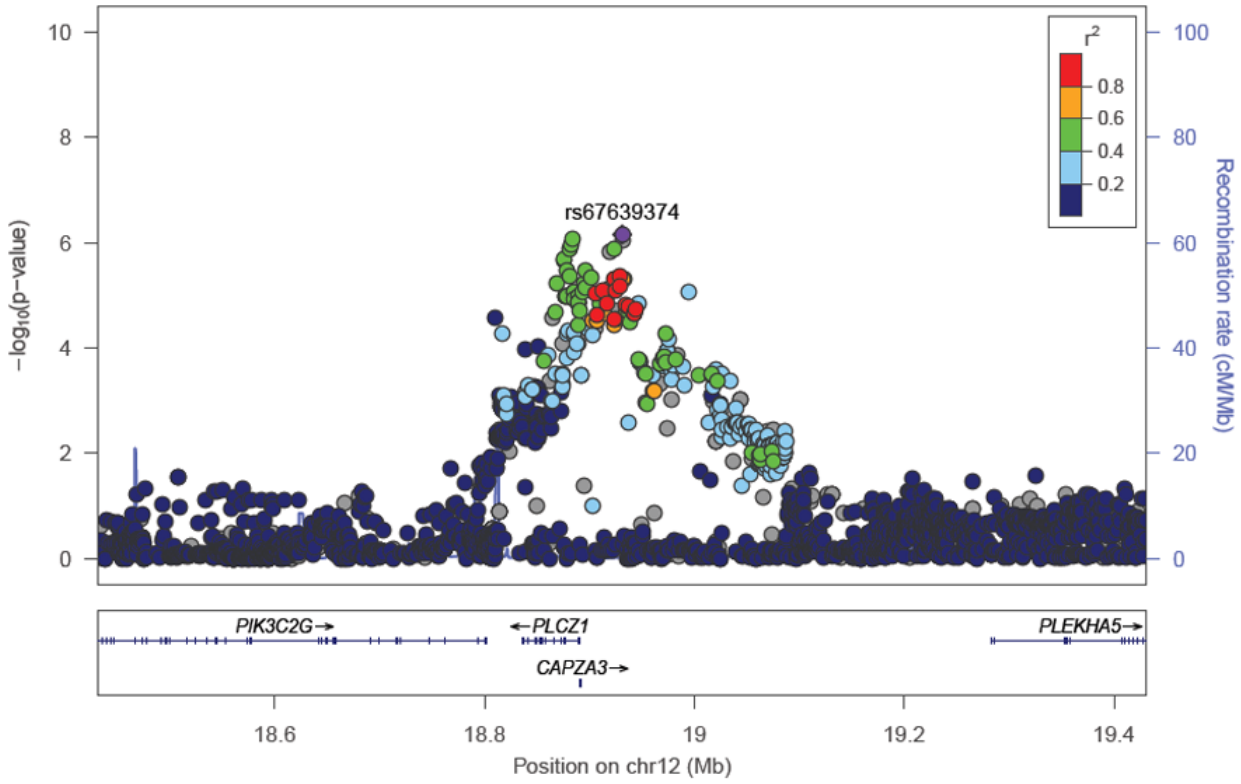
Regional Association plots of regions that do not contain genes with clear biological roles in dental or oral health are included below (Figure 68.) Descriptions of these regions are included in the table of results in the Results section.

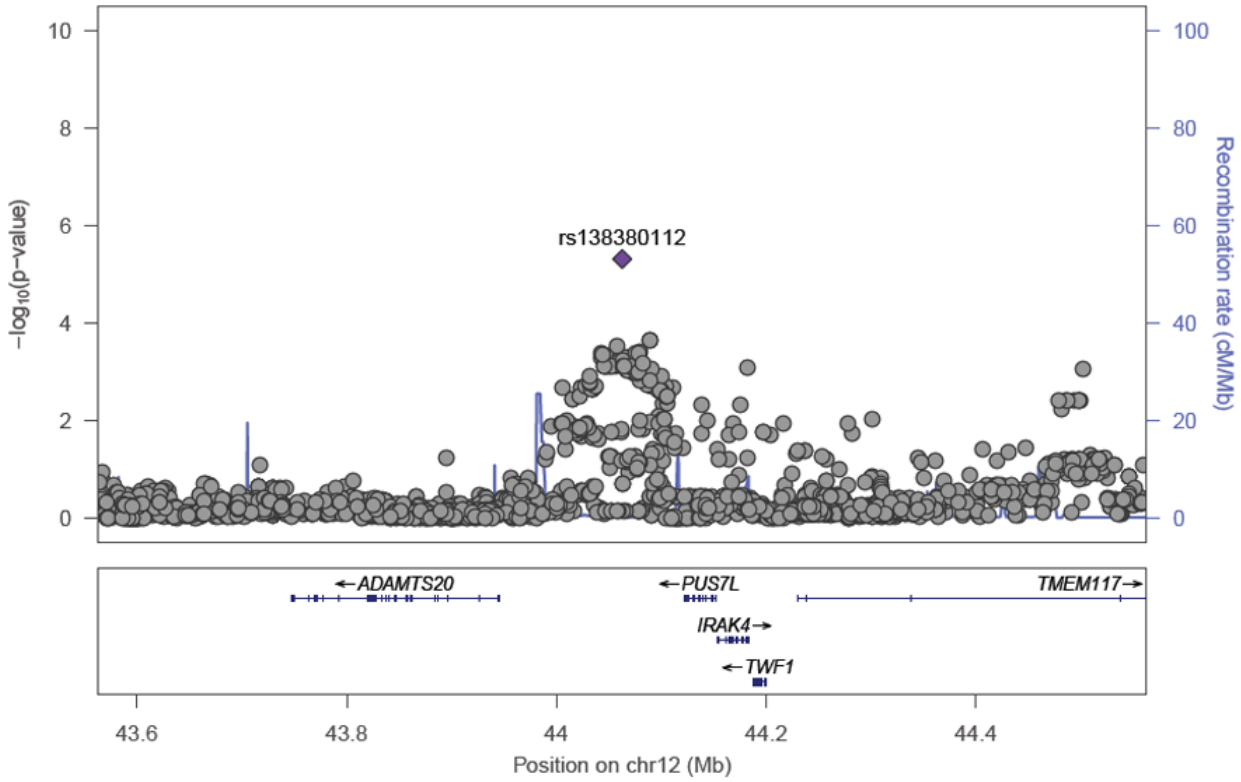

rs3908116
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
NRXN1
50.6
50.8
51
51.2
51.4
Position on chr2 (Mb)


rs12506662
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
MAPK10
PTPN13
C4orf36
KLHL8
MIR5705
SLC10A6
AFF1
HSD17B13
LOC100506746
HSD17B11
NUDT9
87.4
87.6
87.8
88
88.2
Position on chr4 (Mb)


rs7795775
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
CADPS2
TAS2R16
SLC13A1
RNF133
RNF148
Position on chr7 (Mb)


r²
0.8
0.6
0.4
0.2
rs72719502
−log10(p−value)
Recombination rate (cM/Mb)
ALDH1B1
FAM95C
CNTNAP3
IGFBPL1
ANKRD18A
FAM201A
Position on chr9 (Mb)


rs67639374
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
PIK3C2G
PLCZ1
PLEKHA5
CAPZA3
18.6
18.8
19
19.2
19.4
Position on chr12 (Mb)


rs138380112
−log10(p−value)
Recombination rate (cM/Mb)
ADAMTS20
PUS7L
TMEM117
IRAK4
TWF1
43.6
43.8
44
44.2
44.4
Position on chr12 (Mb)

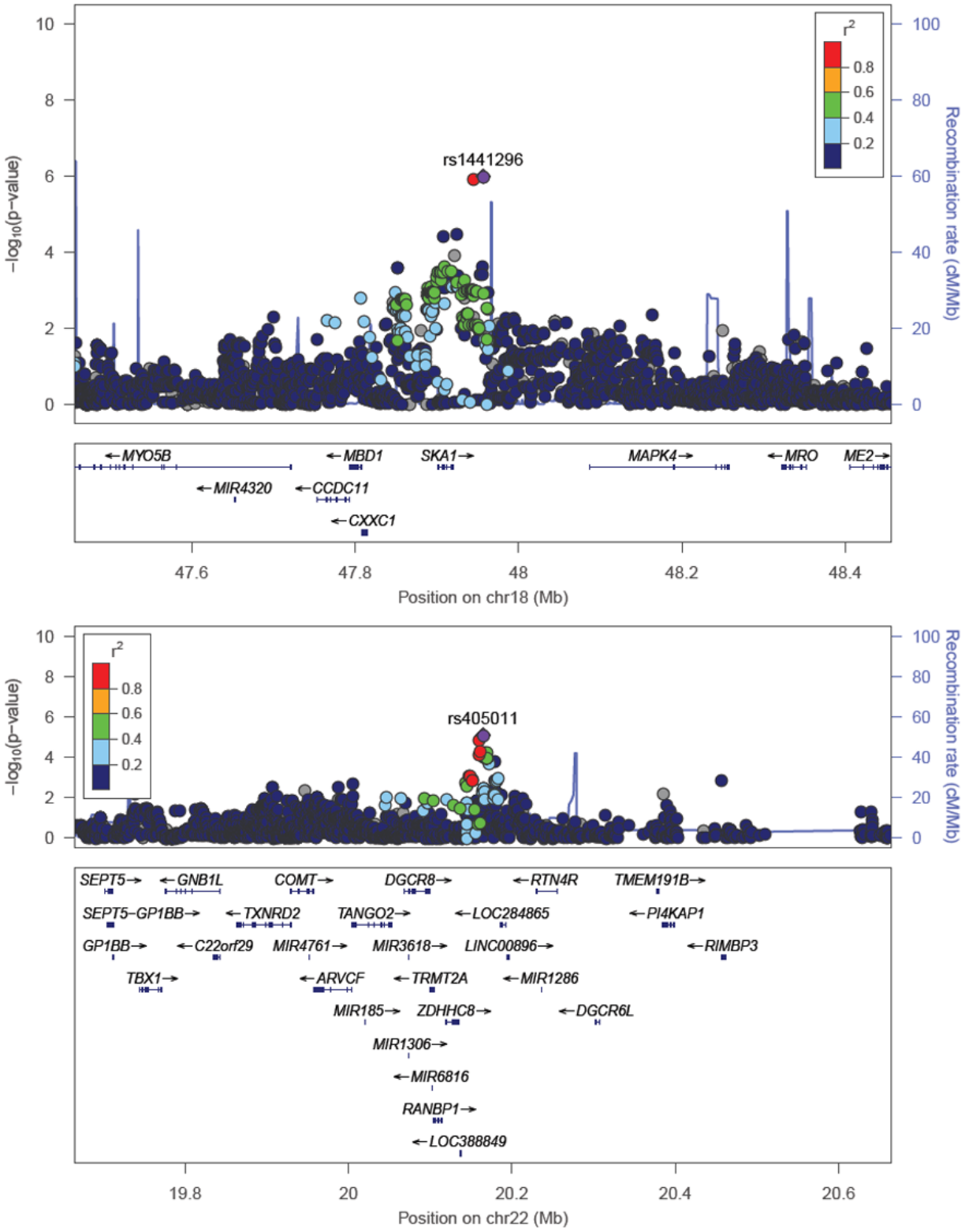
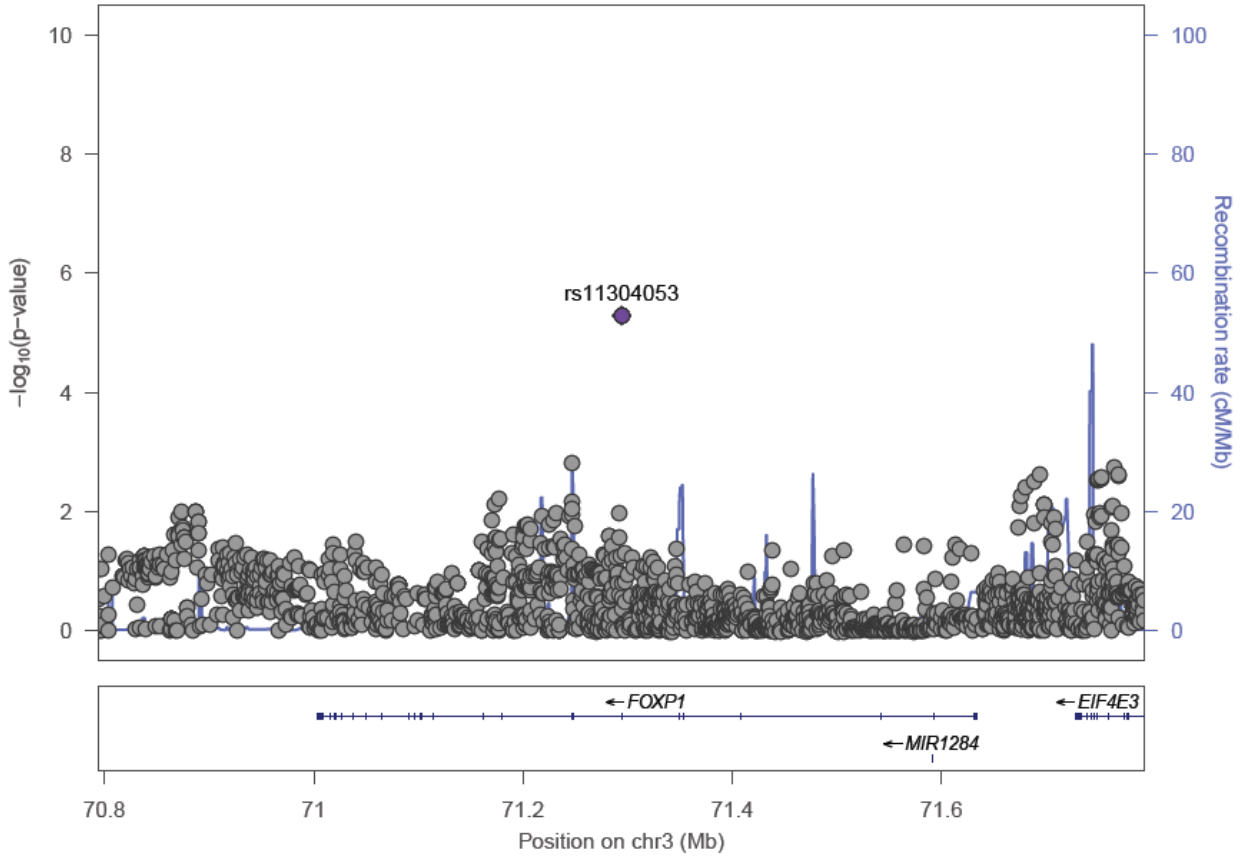
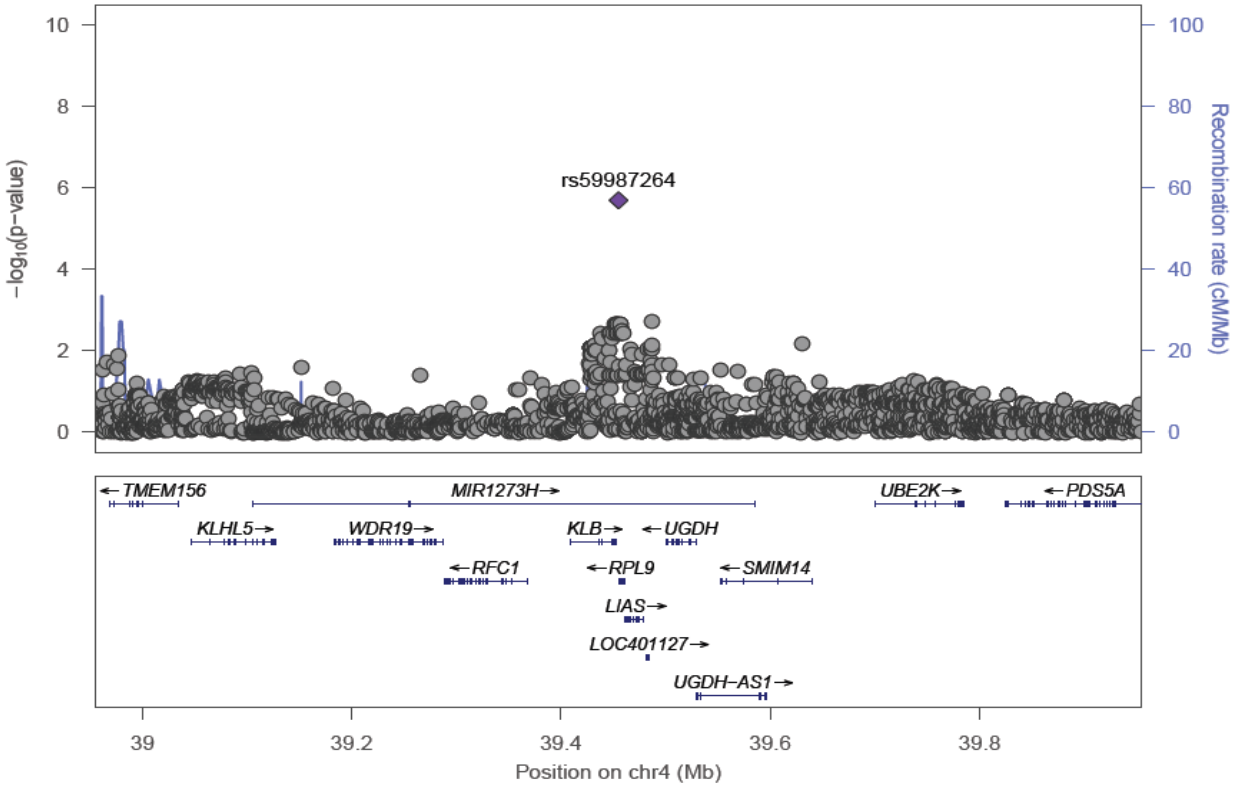


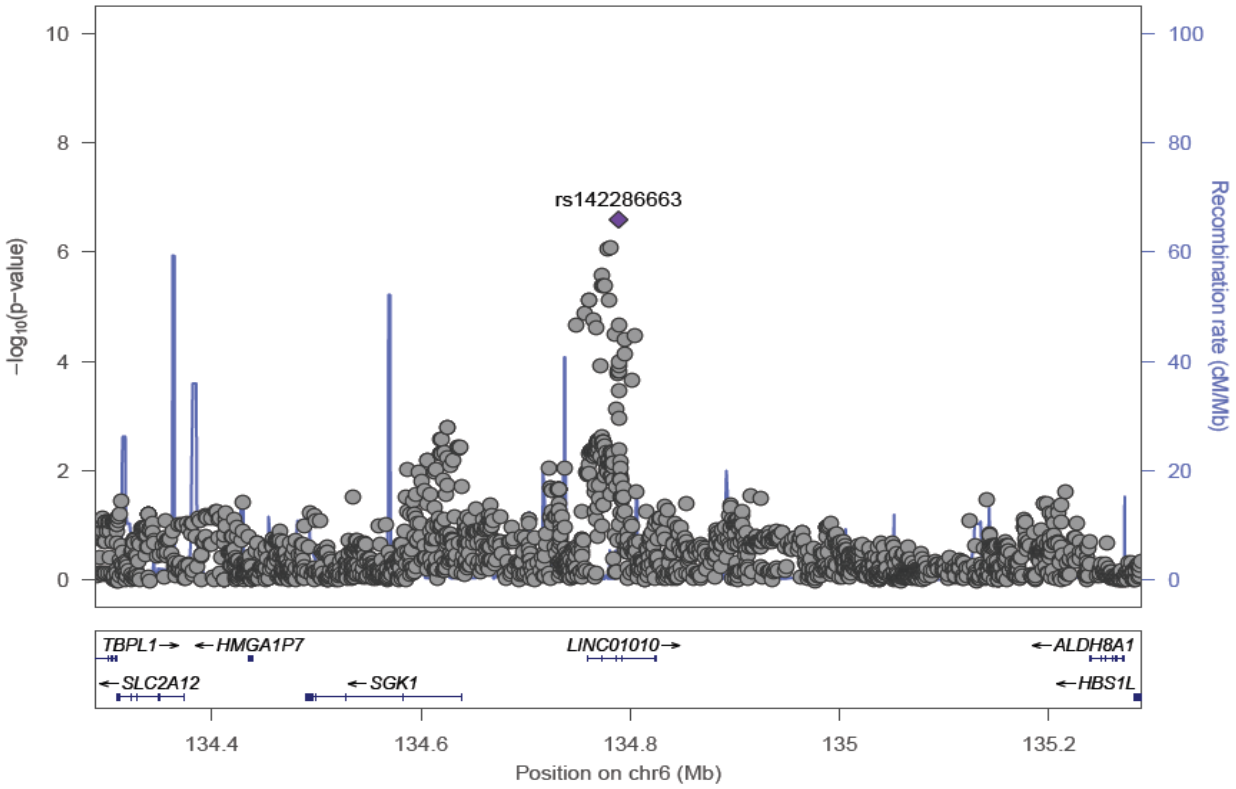
Figure 68. Regional Association Plots of other suggestive significant hits for the quantitative trait in the Hispanics-only meta-analysis

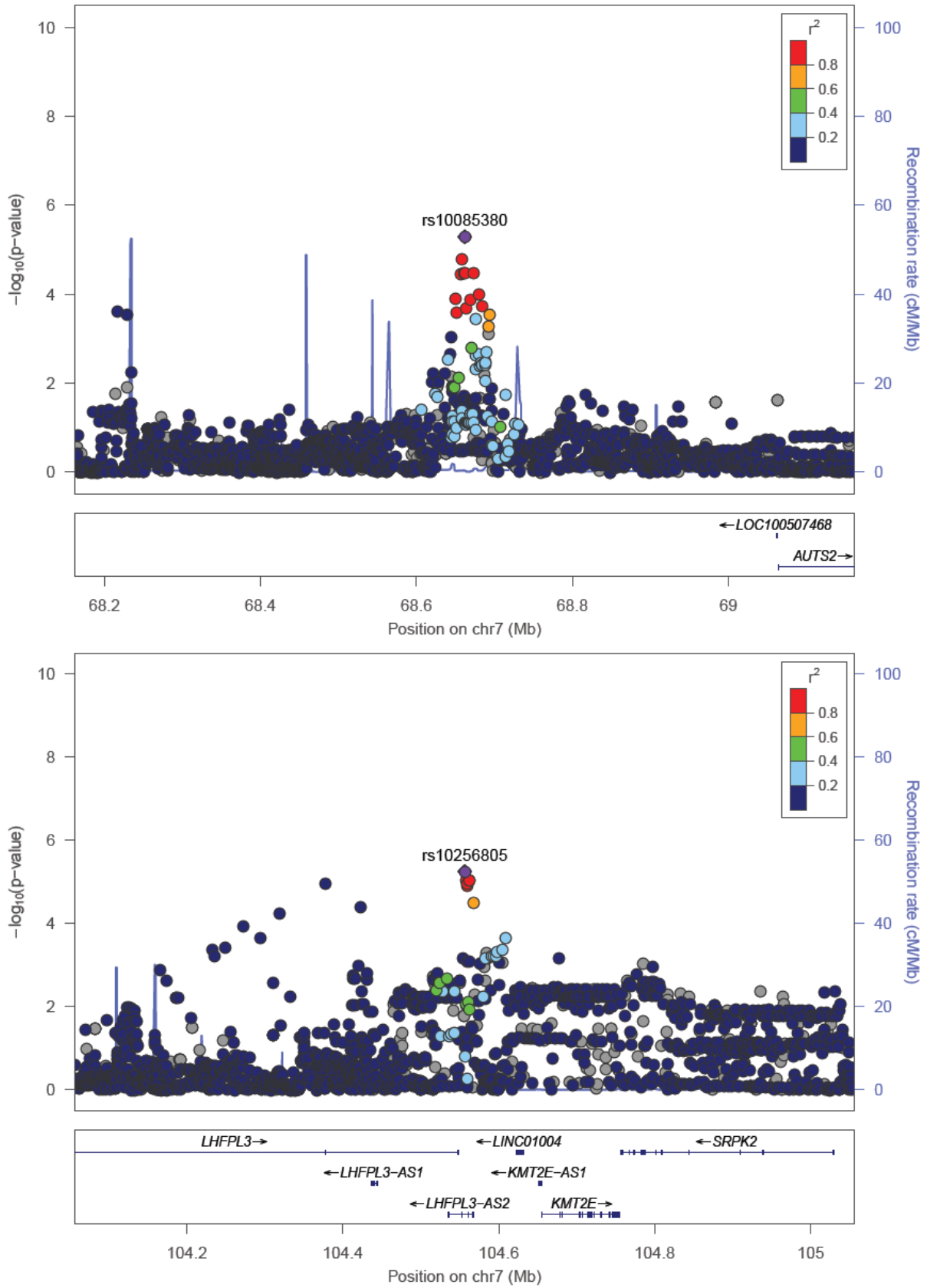
B.4 HISPANICS-ONLY META-ANALYSIS RESULTS – FUNCTIONAL DENTITION

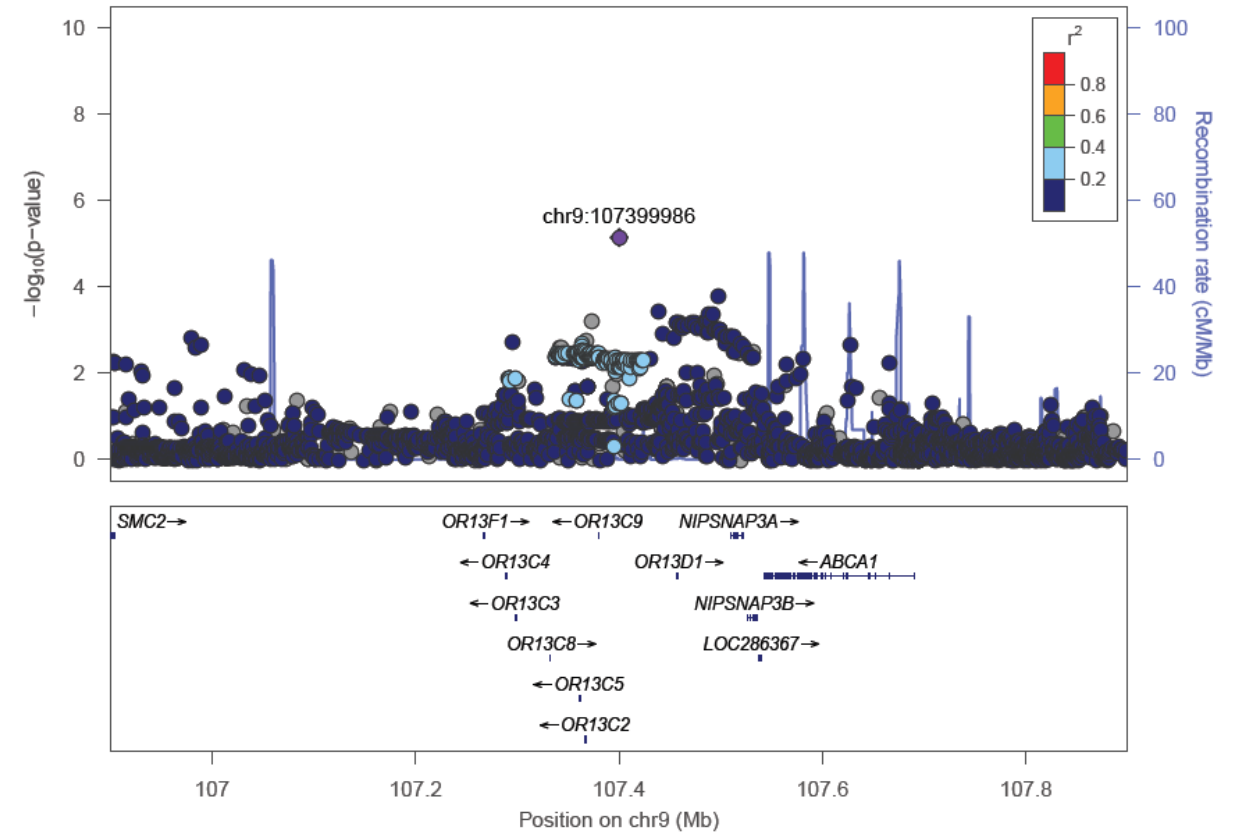
Regional Association plots of regions that do not contain genes with clear biological roles in dental or oral health are included below (Figure 69.) Descriptions of these regions are included in the table of results in the Results section.

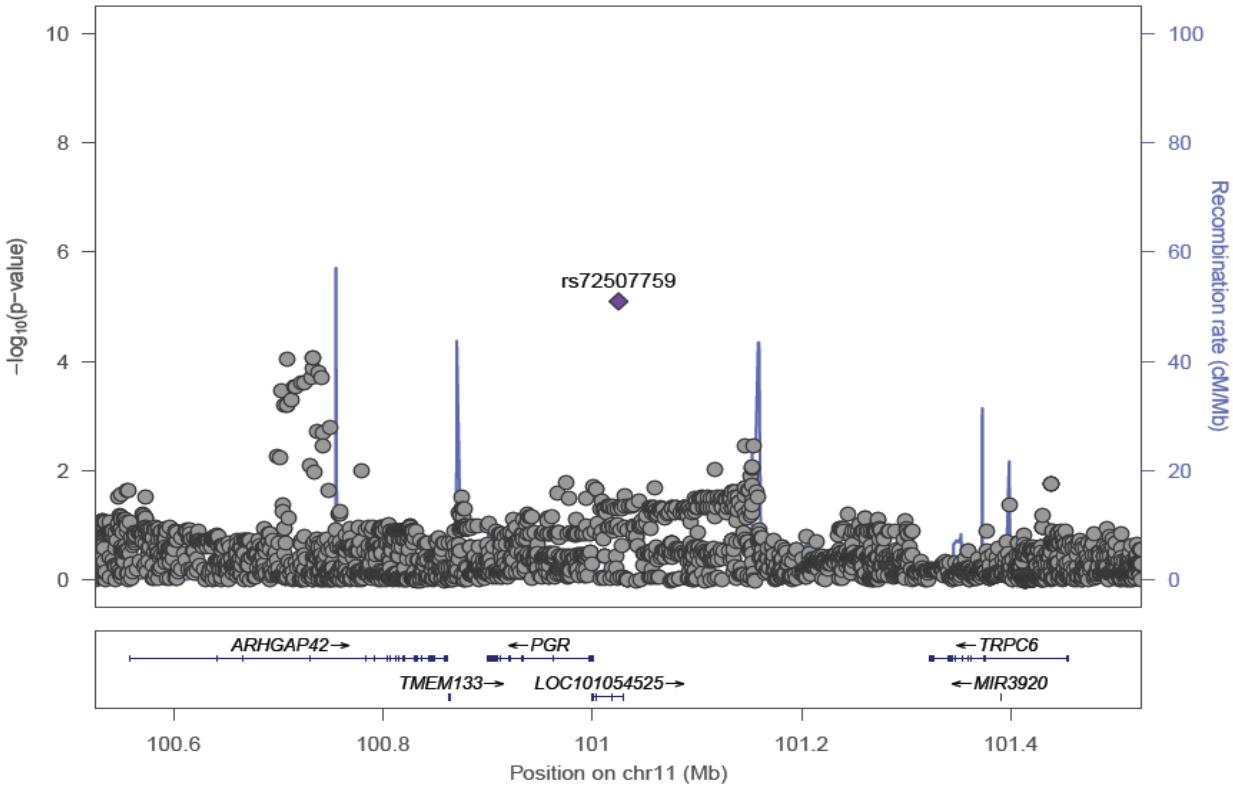


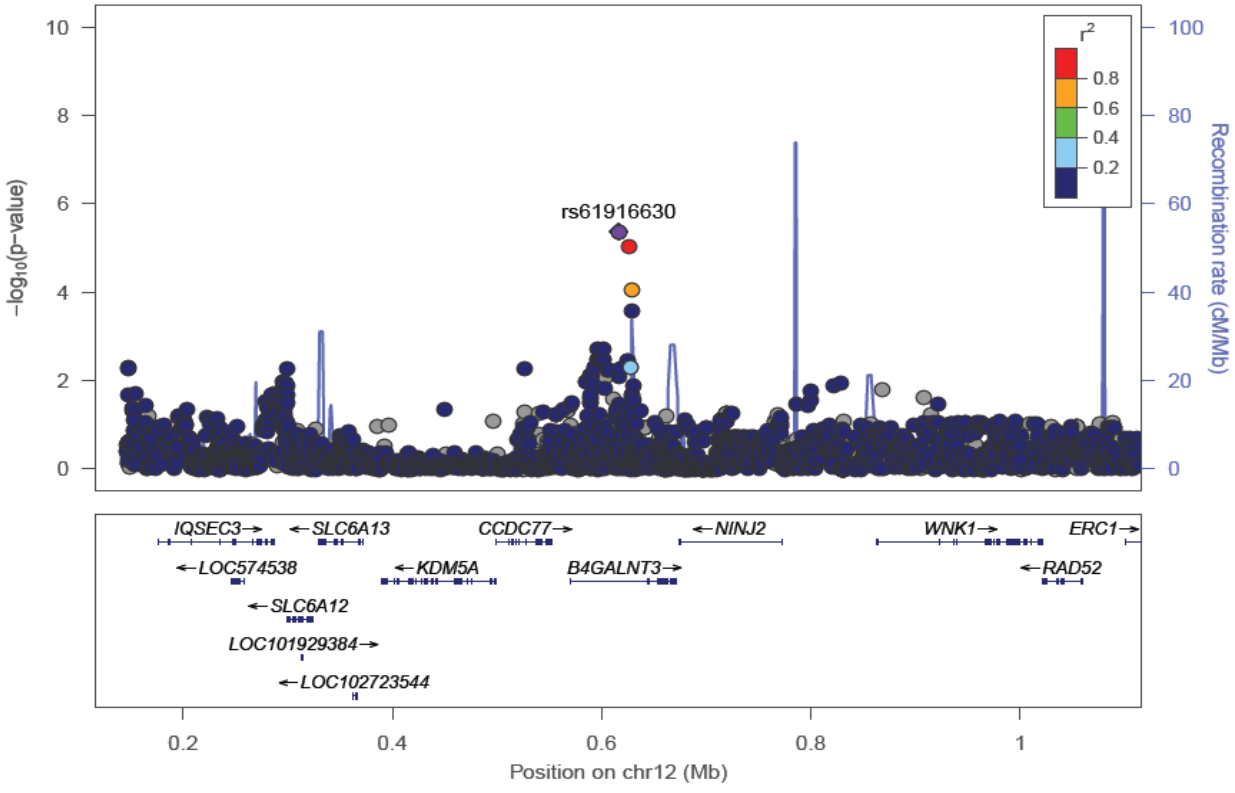

rs59987264
−log10(p−value)
Recombination rate (cM/Mb)
TMEM156
MIR1273H
UBE2K
PDS5A
KLHL5
WDR19
KLB
UGDH
RFC1
RPL9
SMIM14
LIAS
LOC401127
UGDH−AS1
Position on chr4 (Mb)

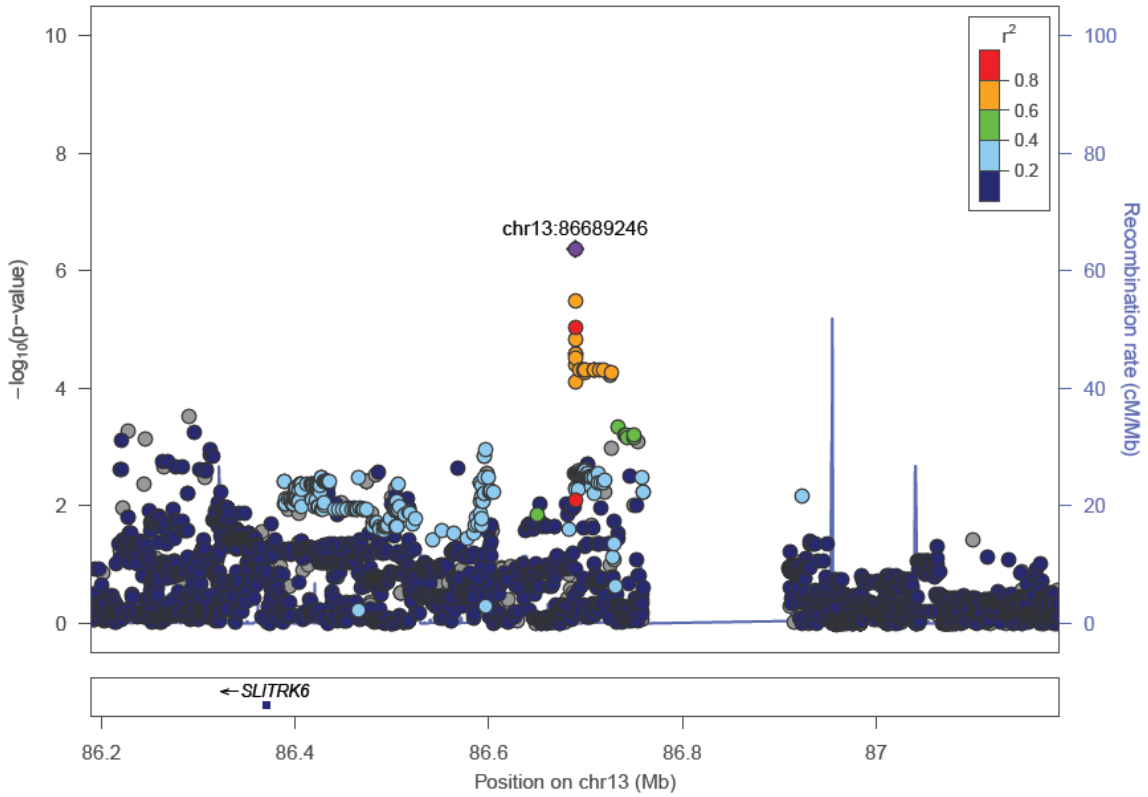

rs142286663
−log10(p−value)
Recombination rate (cM/Mb)
TBPL1
HMGA1P7
LINC01010
ALDH8A1
SLC2A12
SGK1
HBS1L
Position on chr6 (Mb)


rs10085380
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
LOC100507468
AUTS2
Position on chr7 (Mb)
rs10256805
LHFPL3
LINC01004
SRPK2
LHFPL3−AS1
KMT2E−AS1
LHFPL3−AS2
KMT2E
Position on chr7 (Mb)


chr9:107399986
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
SMC2→
OR13F1→
←OR13C9
NIPSNAP3A→
←OR13C4
OR13D1→
←ABCA1
←OR13C3
NIPSNAP3B→
OR13C8→
LOC286367→
←OR13C5
←OR13C2
Position on chr9 (Mb)


rs72507759
−log10(p−value)
Recombination rate (cM/Mb)
ARHGAP42→
←PGR
←TRPC6
TMEM133→
LOC101054525→
←MIR3920
Position on chr11 (Mb)


rs61916630
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
IQSEC3→
←SLC6A13
CCDC77→
←NINJ2
WNK1→
ERC1→
←LOC574538
←KDM5A
B4GALNT3→
←RAD52
←SLC6A12
LOC101929384→
←LOC102723544
Position on chr12 (Mb)


chr13:86689246
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
←SLITRK6
Position on chr13 (Mb)

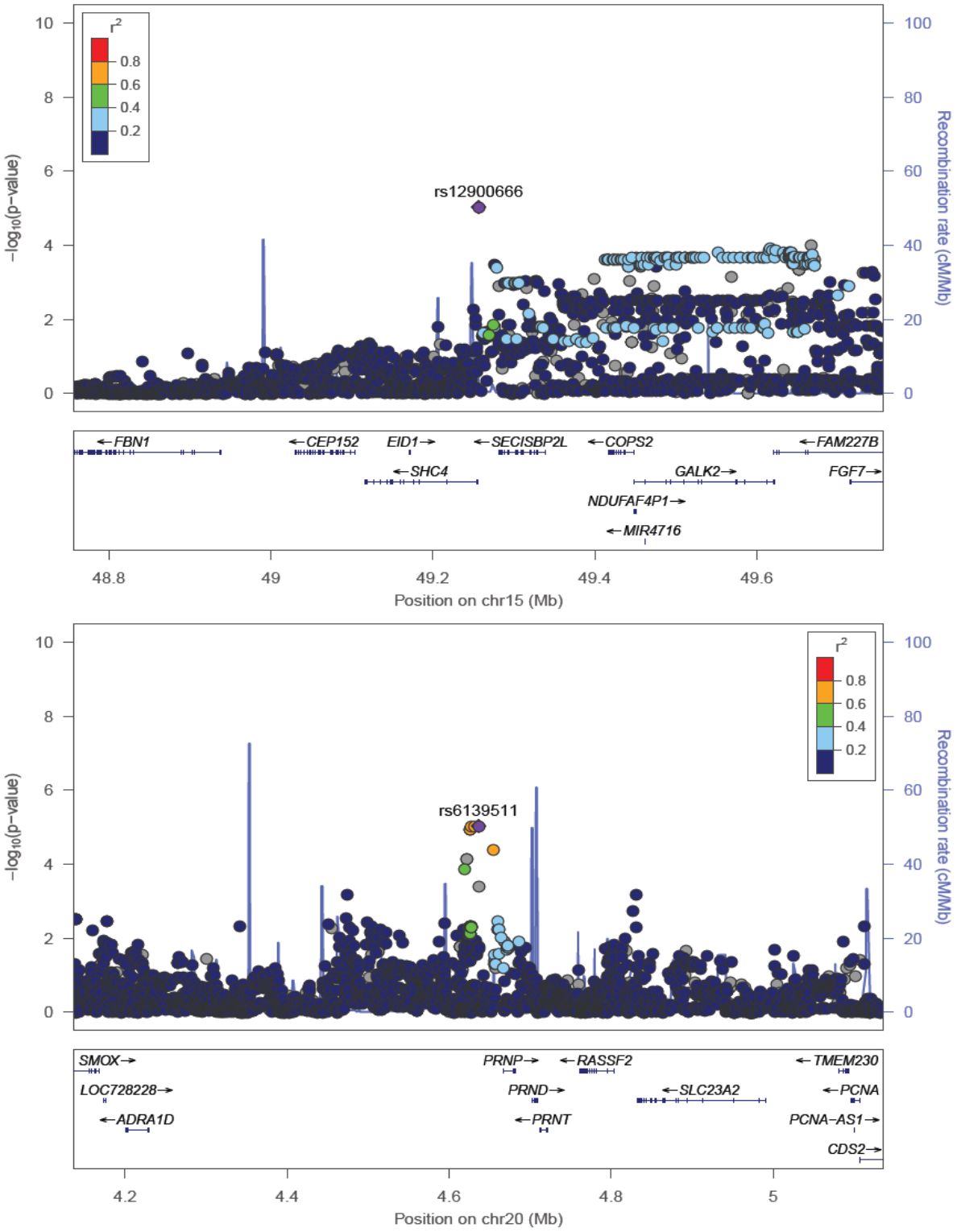


Figure 69. Regional Association Plots of other suggestive significant hits for functional dentition in the Hispanics-only meta-analysis

LD information was not available for all plots.

B.5 TRANS-ETHNIC META-ANALYSIS RESULTS –QUANTITATIVE TRAIT

rs764629 (p = 6.0E-6) is located near several uncharacterized genomic elements and a handful of genes with no known role in dental or oral health (Figure 70). This signal is ~300 kb downstream of the transcription start site of *PLK2*, which is involved in embryonic growth, skeletal development, and cell cycle progression [288]. The nearest gene this signal is upstream of is *GAPT*, GRB2 binding adaptor protein, also approximately 300 kb away. The function of *GAPT* is largely uncharacterized, though GAPT is known to be largely expressed in B cells and may be involved in B cell activation and antibody production [289].

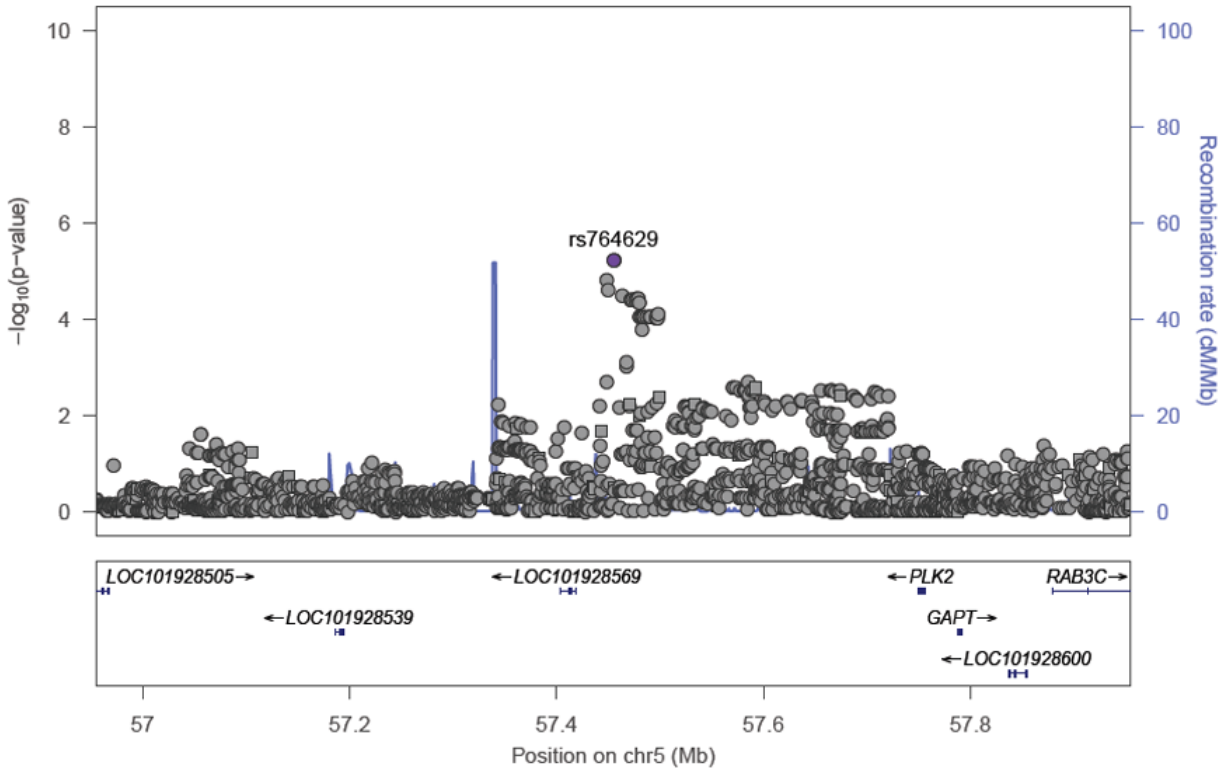


Figure 70. Regional Association Plot of rs764629

For all plots in the trans-ethic meta-analysis, LD information was not included on the plot as it represents a mixed sample, and annotation of imputed vs. genotypes variants was based on COHRA and DRDR.

rs2860807 (p = 7.9E-6; Figure 71) corresponds to the same signal seen in COHRA for the quantitative trait near *NUDCD2*, *HMMR*, *HMMR-AS1*, and *MAT2B*, which is the most interesting gene as it is involved in methionine synthesis, which was implicated in tooth loss by the association seen near *MTRR*.

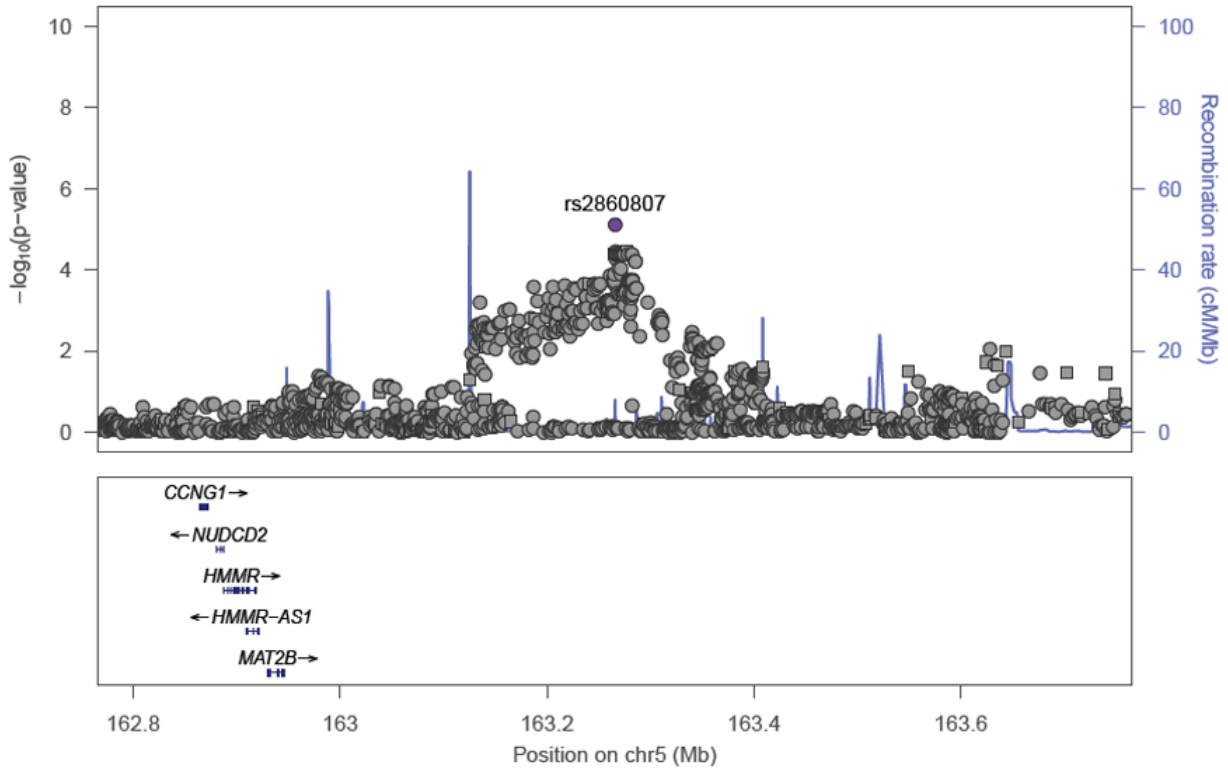


Figure 71. Regional Association Plot of rs2860807

rs10062700 (p = 8.3E-6) is the same signal seen in the whites-only meta-analysis near *FBN2* (Figure 72), which forms microfibrils in connective tissue ECM and is synthesized in the PDL to provide structural stability and maintain homeostasis [67-72]. The function and existing evidence of *FBN2* activity in the PDL corroborates the increasing statistical evidence that this genomic region may influence missing teeth.

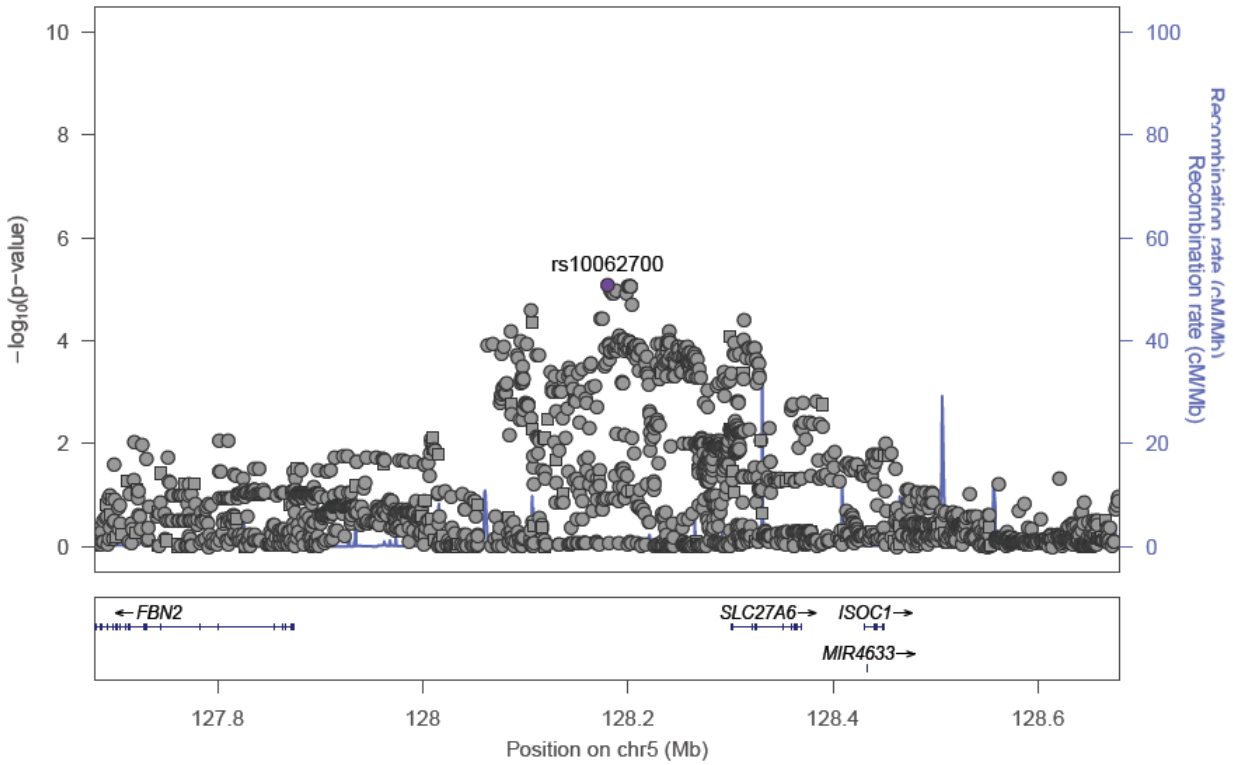


Figure 72. Regional Association Plot of rs10062700

rs72488321 (p = 8.5E-6) lies approximately 200 kb upstream of *LRRTM4* (Figure 73), leucine rich repeat transmembrane neuronal 4, which has no known role in dental or oral health. *LRRTM4* is an excitatory postsynaptic protein, and knockout mice show defects in excitatory synapse formation and function in the brain [290]. There is no existing biological evidence that this genomic region may influence missing teeth.

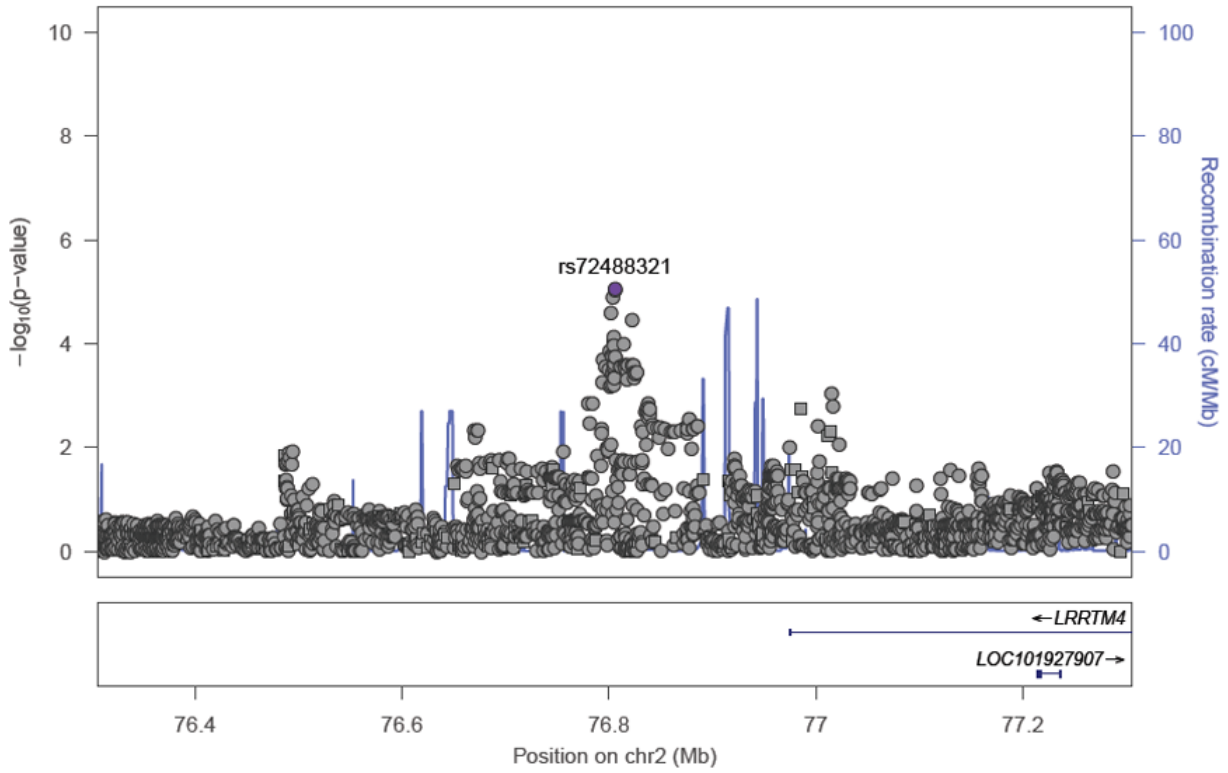


Figure 73. Regional Association Plot of rs72488321

rs6663322 (p = 9.8E-6) is intronic to *DDX59* (Figure 74), DEAD-box helicase 59, which is an RNA helicase involved in RNA metabolism [291]. Mutations in *DDX59* are associated with oral-facial-digital syndrome, manifestations of which include dental anomalies such as hypodontia and extramandibular tooth buds [238]. Additionally, impaired Shh signaling has been implicated in affected families with *DDX59* mutations [291]. Also in this region is *CACNA1S* and *TMEM9*, which were previously identified in a GWAS of permanent tooth eruption [271], albeit in a different genomic position. Timing of permanent tooth eruption may influence missing teeth, as early eruption increases exposure to dental caries and has implications for periodontal disease and malocclusion.

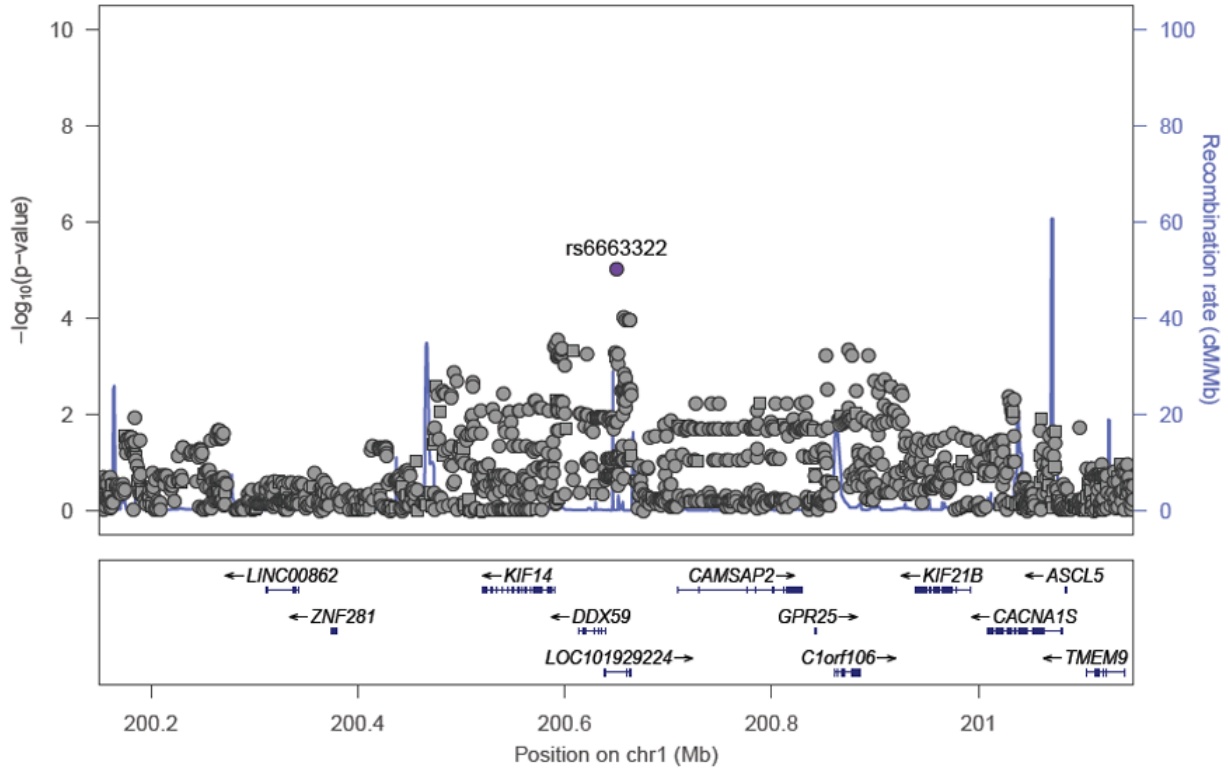


Figure 74. Regional Association Plot of rs6663322

B.6 TRANS-ETHNIC META-ANALYSIS RESULTS – FUNCTIONAL DENTITION

rs4575613 (p = 1.3E-6) is intronic to *ZNF521* (Figure 75), zinc finger protein 521. This association is largely driven by its significance in COHRA (p = 5.5E-7), as it is only nominally significant in POFC-G (p = 0.058) and POFC-PA (p = 0.022) and is not significant in DRDR (p = 0.57.) Zfp521, the protein product of *ZNF521*, regulates lineage commitment of cell types during development and mediates osteoblast over adipocyte differentiation induced by BMP2 in mesenchymal cells [292]. However, Zfp521 has not been studied in tooth development and it is unknown how this protein may affect dental and oral health.

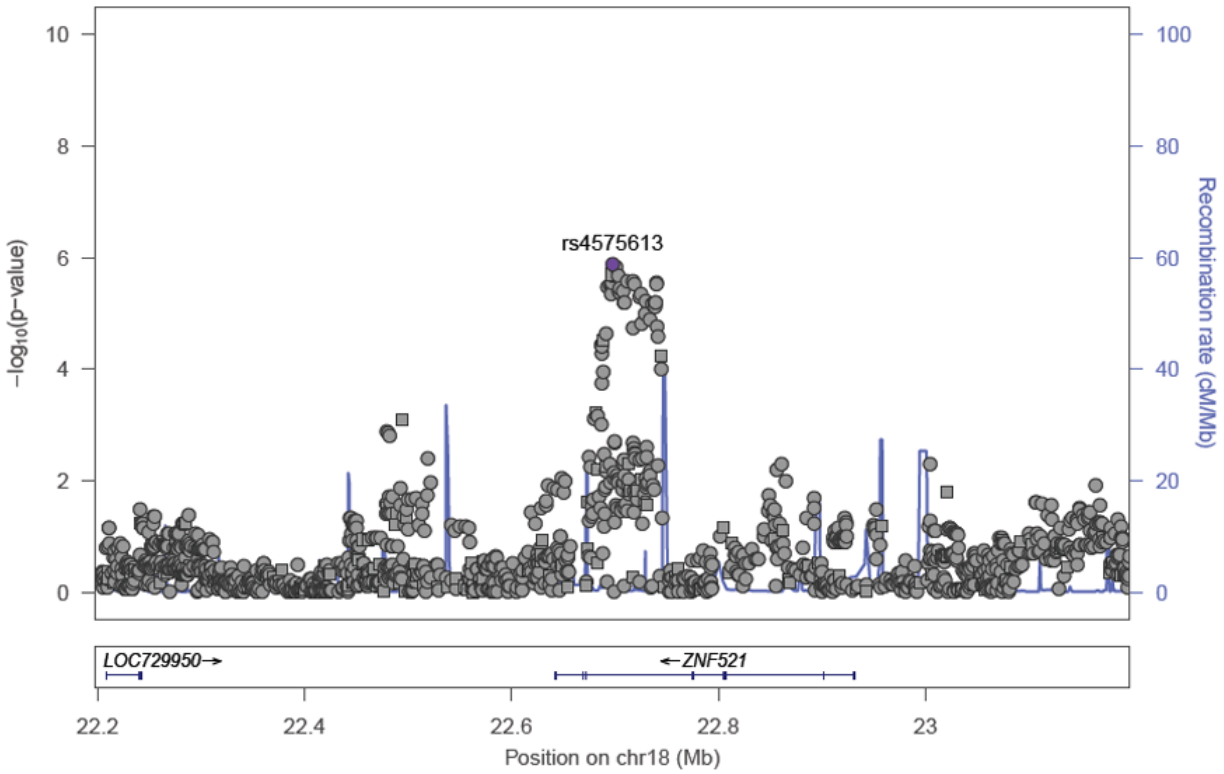


Figure 75. Regional Association Plot of rs4575613

rs674271 (p = 3.0E-6) is intronic to *MGC27382* (Figure 76), an uncharacterized non-coding RNA. It has been found to be upregulated in colorectal cancer stem cells positive for CD133, a biomarker of tumor initiating cells, versus those negative for CD133 [293], and it was implicated in a GWAS of body-mass index in women adjusted for physical activity (p = 6.4E-11) [294]. This region also contains several other genes, those most likely to be affected by the associated SNPs based on proximity being *PTGFR* and *FUBP1*, but none are relevant to dental and oral health; it is unknown how this region of the genome may affect missing teeth.

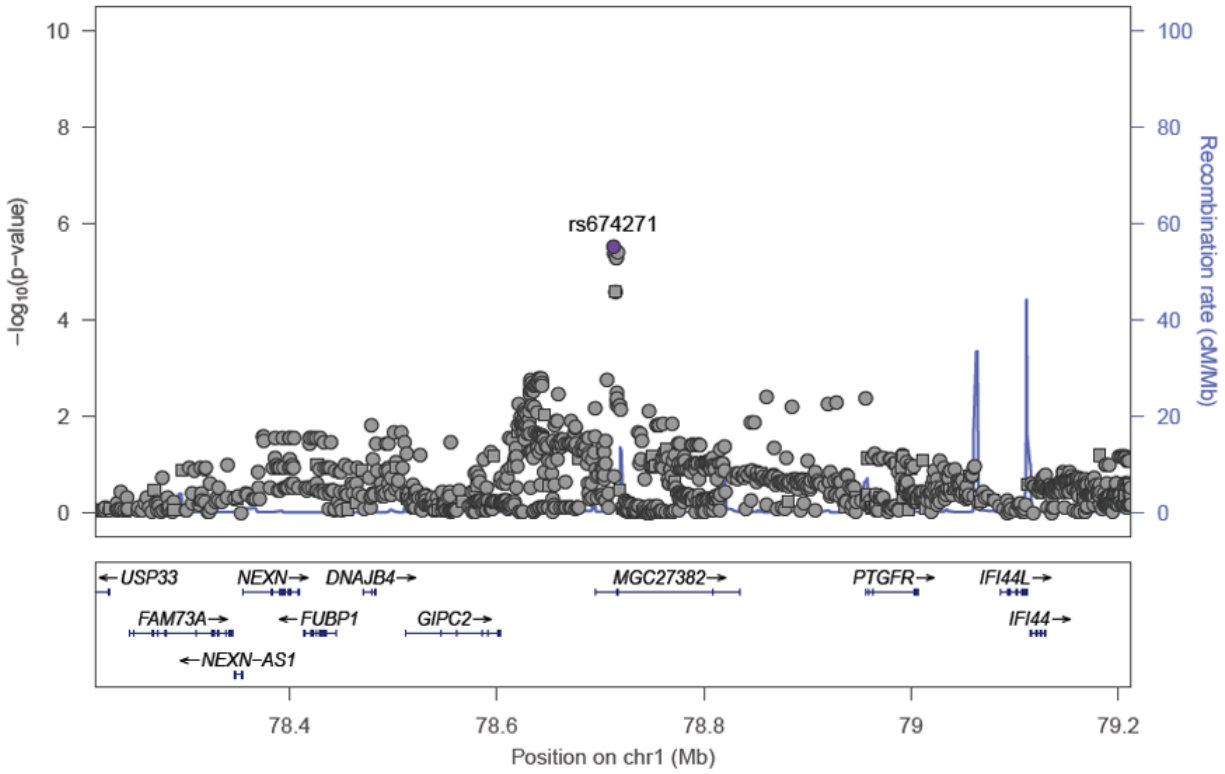


Figure 76. Regional Association Plot of rs674271

rs28734985 (p = 8.4E-6) is located just upstream of *IPMK* (Figure 77), inositol polyphosphate multikinase. It has been shown that IPMK stabilizes TRAF6, tumor necrosis factor receptor-associated factor 6, which is a key mediator of Toll-like receptor (TLR) signaling [295]. Deletion of IPMK in mice macrophages resulted in a decrease in TLR signaling and subsequent inflammatory response, including pro-inflammatory cytokine production; these mice were also resistant to bacterial infection [295]. While IPMK has not been studied in periodontal disease, its regulatory role in TLR signaling makes it a plausible candidate for tooth loss due to periodontitis, as TLR signaling is a key inflammatory pathway implicated in the pathogenesis of periodontal disease [180, 195].

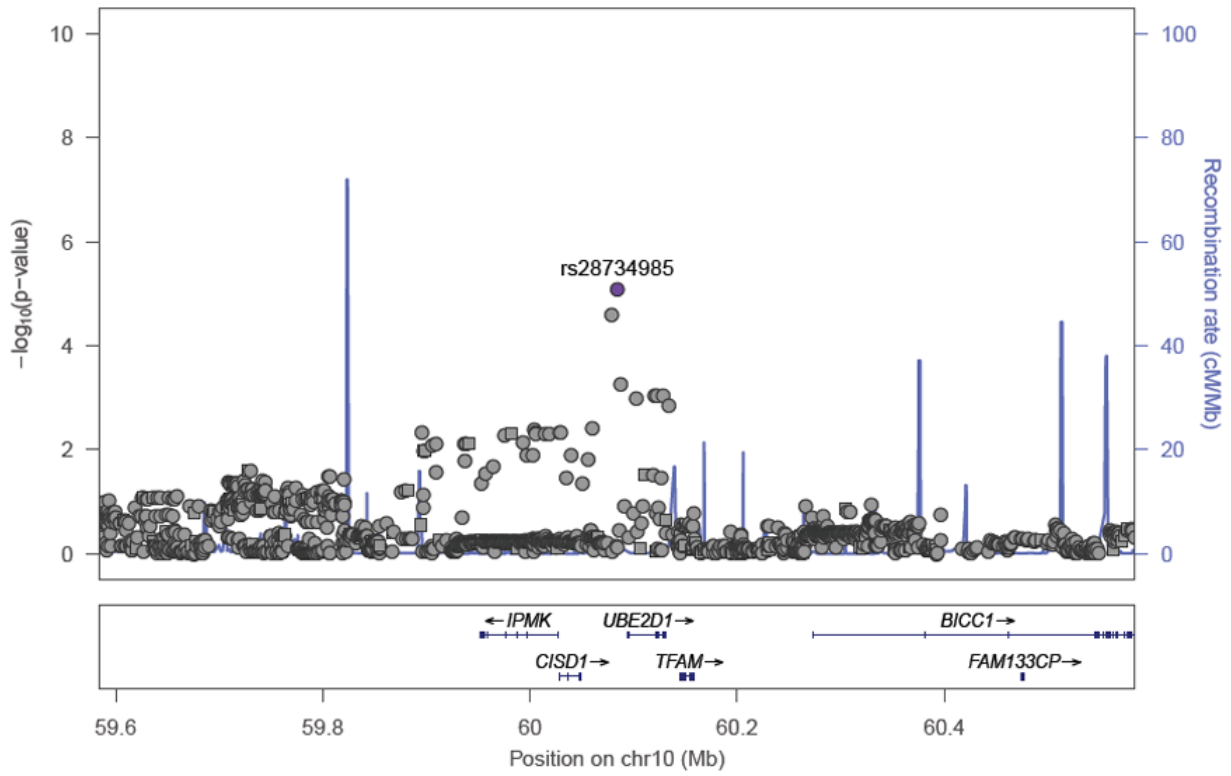


Figure 77. Regional Association Plot of rs28734985

APPENDIX C: RESULTS IN POFC-USA

All results in POFC-USA are included in this appendix section since spot-checking of significant associations did not add to the weight of evidence that discovered variants are associated with missing teeth and functional dentition.

C.1 COVARIATE MODELING

For POFC-USA, sample characteristics and trait distributions are shown in Table 19, and p-values for the test of association between demographic variables and both missing teeth traits are shown in Table 20. As with COHRA, the sample size in these results is greater than that of the GWAS, as some individuals did not have genotype information available but were included in initial trait development and covariate modeling. Only education ($p = 0.016$) and age ($p = 9.19E-8$) were significantly associated with the quantitative trait, and only age ($p = 1.28E-5$) was associated with functional dentition. Plots of the first 4 PCs of genetic ancestry are shown in Figure 78. Three PCs were sufficient to pull out each ancestry group, and the fourth no longer distinguished between them.

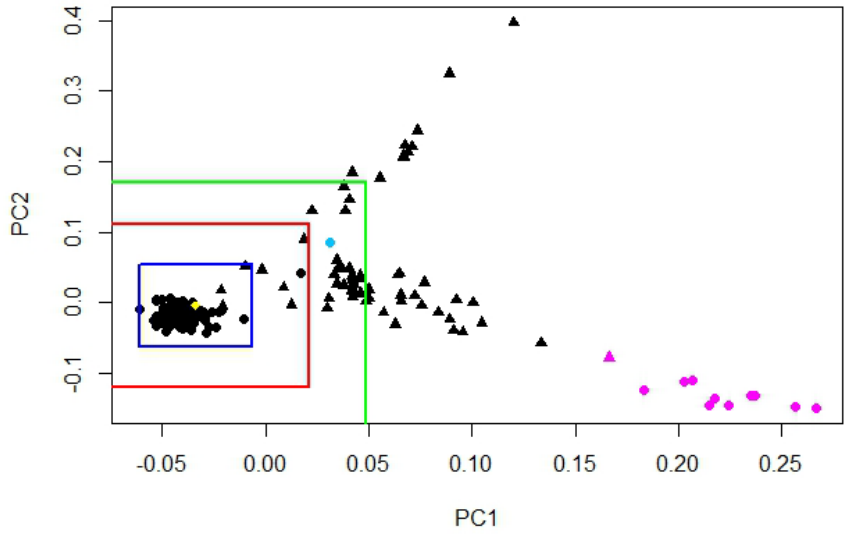
Table 19. Study characteristics and trait distribution in POFC-USA

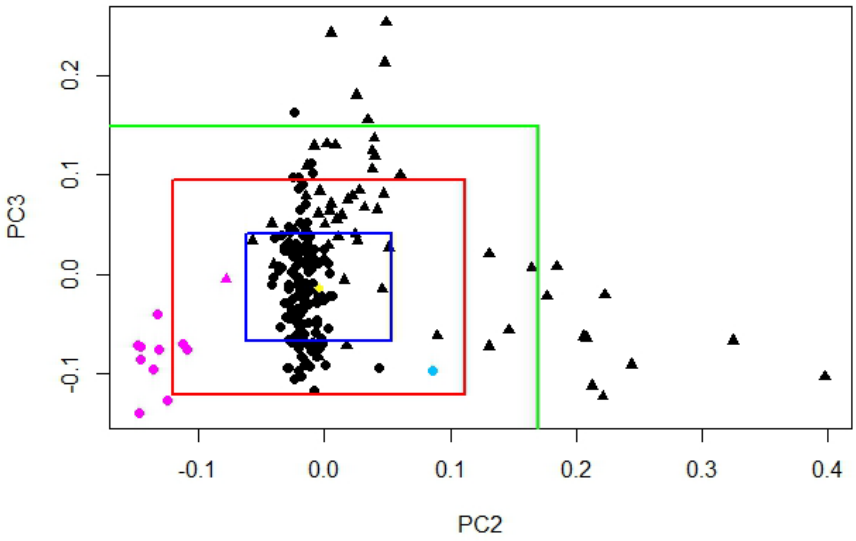
Variable	Value	Count (%)/ Mean (SD)
Functional Dentition	Yes	199 (92.1%)
	No	17 (7.9%)
Total Missing Teeth		0.75 (0.92)
Age		36.2 (11.1)
Sex	Male	80 (36.7%)
	Female	138 (63.3%)
Education*	High School, GED, or lower	82 (42.5%)
	Tech School, Associate degree	35 (18.1%)
	Undergrad or higher	76 (39.4%)
Site	Colorado	27 (12.4%)
	Pittsburgh	136 (62.4%)
	Puerto Rico	47 (21.6%)
	Texas	8 (3.7%)
Race*	White	155 (88.6%)
	Black	11 (6.3%)
	Asian	1 (0.6%)
	Other	7 (4%)
	Caucasian & Native American	1 (0.6%)
Ethnicity	Non-Hispanic	151 (69.6%)
	Hispanic	66 (30.4%)

Categorical covariates show counts and percents, and continuous variables show means and standard deviations.
*Total number of individuals is lower for race and education than other covariates due to missing data.

Table 20. Association of Covariates with Missing Teeth Traits in POFC-USA

Variable	Quantitative Trait		Functional Dentition	
	P	**Beta (SE)**	**P**	**Beta (SE)**
Age	9.19E-8	0.03 (0.005)	1.28E-5	0.087 (0.02)
Sex	0.68	0.05 (0.13)	0.24	0.70 (0.59)
Education	0.016	-0.18 (0.07)	0.13	-0.46 (0.30)
Site	0.46	0.14 (0.19)	0.65	0.36 (0.79)
Race	0.36	0.05 (0.05)	0.14	0.18 (0.12)
Ethnicity	0.29	0.14 (0.14)	0.11	-1.23 (0.77)


PC2
0.4
0.3
0.2
0.1
0.0
-0.1
-0.05
0.00
0.05
0.10
0.15
0.20
0.25
PC1


PC3
0.2
0.1
0.0
-0.1
-0.1
0.0
0.1
0.2
0.3
0.4
PC2

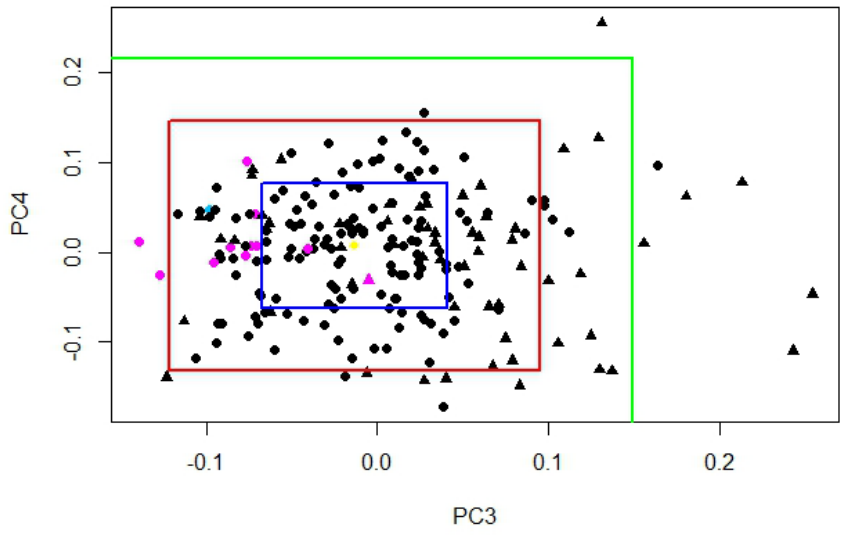


Figure 78. PCA plots of PC1-4 in POFC-USA

Circles and triangles represent non-Hispanic and Hispanic ethnicity, respectively. Black coloring represents white ancestry, pink African American, and blue Asian. Rectangles indicate 1, 2, and 3 standard deviations as calculated from the self-reported whites only, and yellow circle represents the mean.

C.2 GWAS-RESULTS

For POFC-USA, GWAS was performed in 192 individuals for whom genotype information was available. A minor allele frequency threshold of MAF > 0.09 was chosen as GWA scans at MAF > 0.03 did not appear well-behaved. At MAF > 0.03, the quantitative trait showed inflation ($\lambda = 1.02$), especially on the QQ plot, as observed p-values begin to deviate from the expected p-values just after $-\log_{10}(p) = 2$; functional dentition showed deflation ($\lambda = 0.91$) and was exceptionally low powered as only 17 individuals lacked a functional dentition. Manhattan and QQ-plots for MAF > 0.09 are shown in Figures 79-80 for both phenotypes. As with DRDR and POFC-PA, the results of this GWA scan were not annotated due to the small sample size and lack of genome-wide significant hits. At MAF > 0.09, genomic inflation factor was $\lambda = 0.97$ and

$\lambda = 0.92$ for the quantitative trait and functional dentition, respectively, both of which indicate deflation. Additionally, the observed p-values depart from expectations at the lower tail of the distribution. As both GWA scans show these departures from expectations, this indicates that the test statistics are not behaving well at the genome scale. Therefore, POFC-USA was not used for gene discovery and was not included in the meta-analysis; it was only used to check significant associations seen in the meta-analyses.

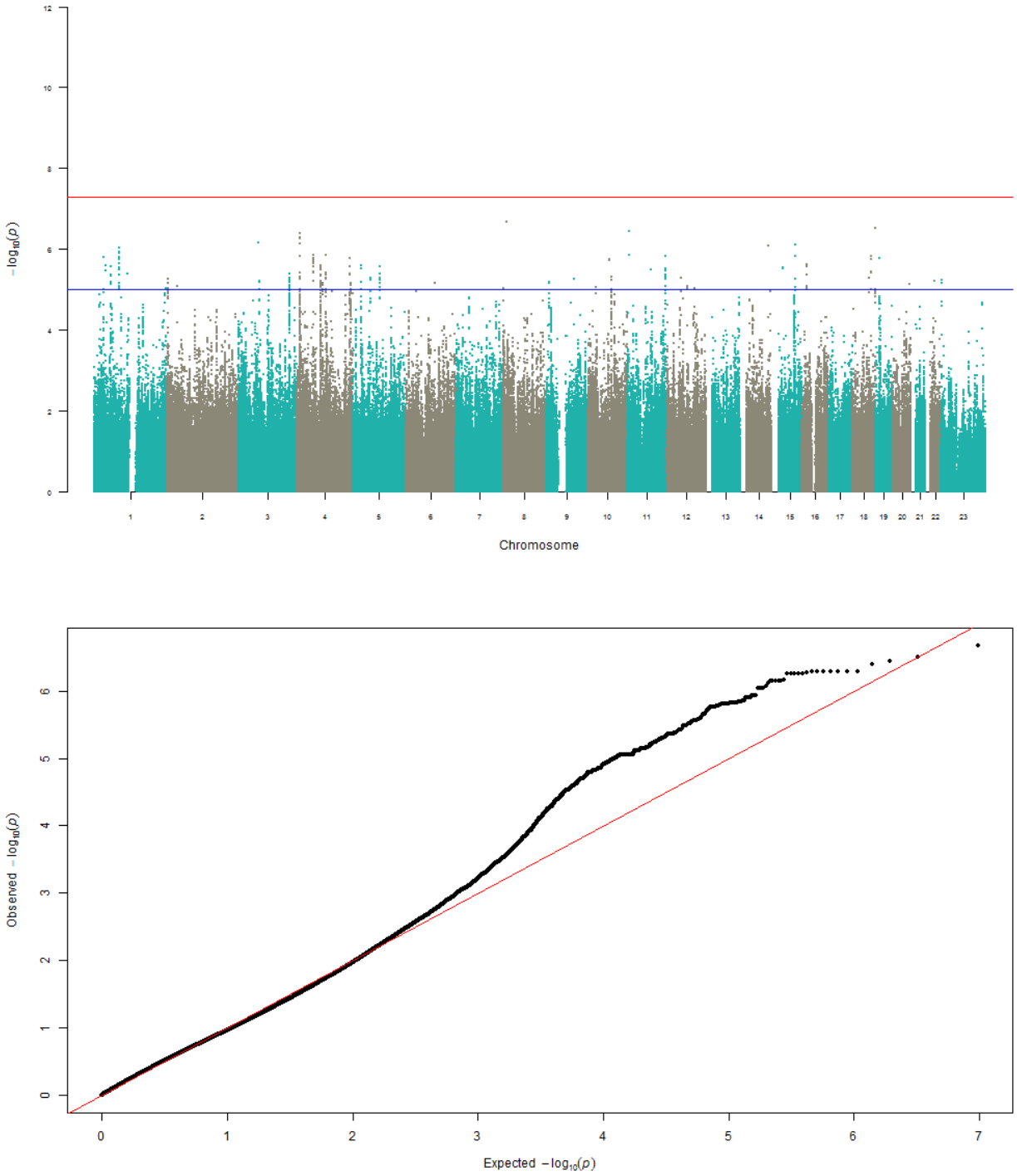


Figure 79. Manhattan and QQ plots for the quantitative trait in POFC-USA at MAF > 0.09

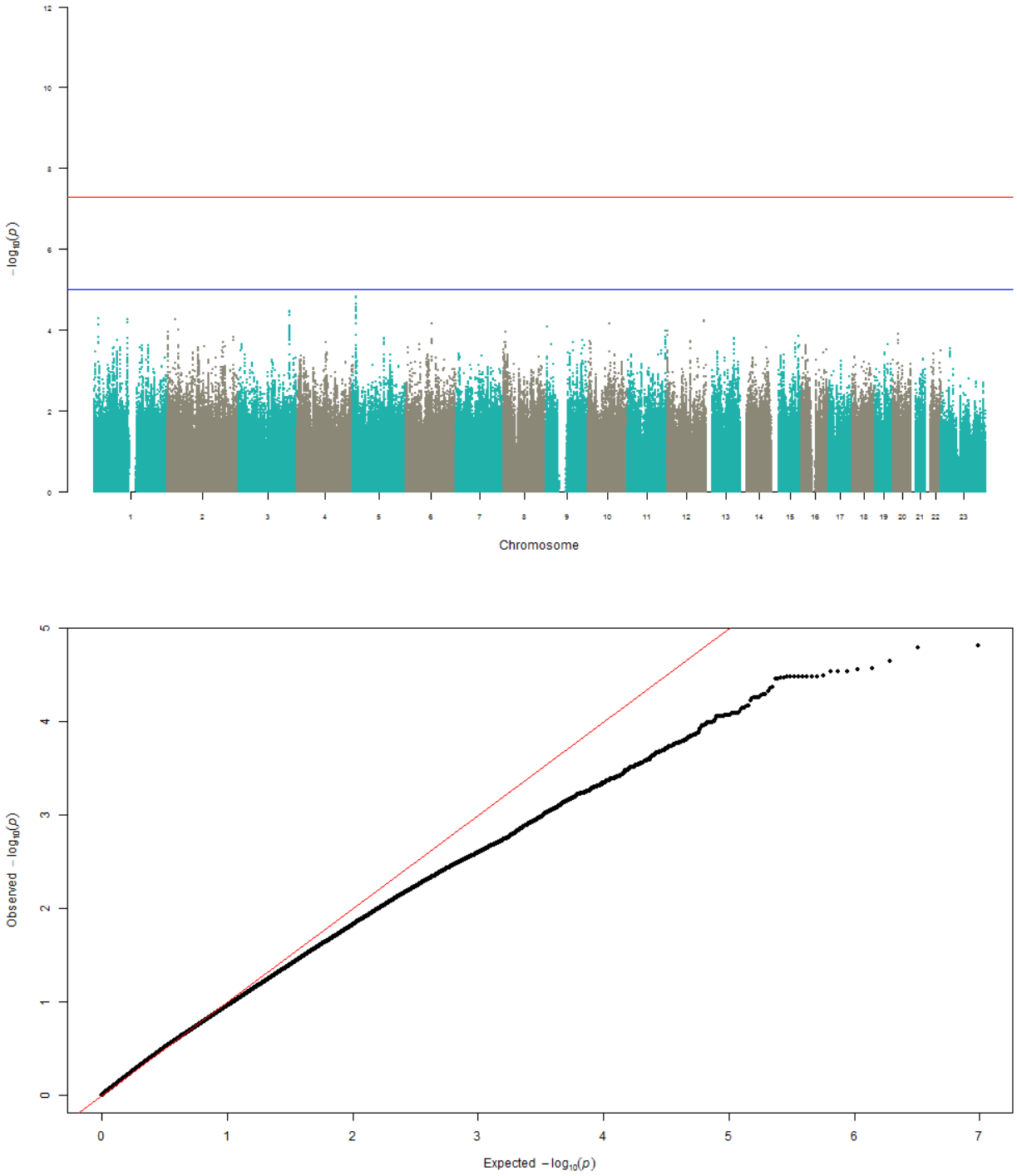


Figure 80. Manhattan and QQ plots for functional dentition in POFC-USA at MAF > 0.09

C.3 SPOT-CHECKING OF ASSOCIATIONS FROM OTHER ANALYSES

No SNPs from the whites-only meta-analysis or the trans-ethnic meta-analysis were significant in POFC-USA at $\alpha = 0.05$ (Tables 21-24). However, two SNPs had $p < 0.10$ and would have lowered the p-value of the variants in the meta-analysis as they were both directionally concordant with the meta-analysis SNPs ($p < 0.10$). This is denoted in the tables.

Table 21. Association of the top hits for the whites-only meta-analysis of the quantitative trait in POFC-USA

SNP	CHR	BP	P_{meta}	$P_{POFC-USA}$	BETA	N	MAF
rs9750906	2	76804252	1.77E-06	0.2174	-0.5023	186	0.23
rs3797113	5	10695026	7.04E-06	0.7326	-0.2181	192	0.08
rs79950078	5	29703754	9.19E-06	0.3692	-0.6881	194	0.05
rs9918187	5	96531684	7.67E-06	0.7838	0.1065	186	0.33
rs10062700	5	128180057	6.42E-06	0.2113	0.464	194	0.34
rs2860805	5	163265408	8.78E-06	0.8489	0.0868	192	0.27
rs56345510	**6**	**99599021**	**5.95E-06**	**0.08327**	**0.6543**	**191**	**0.32**
rs78472857	10	131481707	1.25E-06	0.8539	-0.1127	177	0.12
rs4444411	18	27622307	6.18E-06	0.172	-0.903	193	0.09
rs73199539	21	32938208	6.75E-07	0.7951	-0.2589	186	0.03
rs7061889	23	22719993	1.52E-06	0.3552	-0.3728	181	0.18

Bolded text denoted SNPs with $P_{POFC-USA} < 0.10$

Table 22. Association of the top hits for the whites-only meta-analysis of functional dentition in POFC-USA

SNP	CHR	BP	P_{meta}	$P_{POFC-USA}$	OR	N	MAF
rs600411	1	61355371	8.83E-6	0.6084	0.578	190	0.05
rs1455837	2	219912758	4.62E-6	0.7425	1.15	192	0.38
rs3930612	2	50896945	4.25E-6	0.2465	0.4074	179	0.20
rs13081582	3	94495036	2.43E-6	0.2266	0.543	182	0.42
rs6898589	5	7961941	4.10E-8	0.3186	1.718	188	0.18
rs875142	6	52228225	8.49E-6	0.6351	0.8057	187	0.29
rs1090071	6	93112404	6.83E-6	0.3015	1.548	194	0.30

Table 23. Association of the top hits for the trans-ethnic meta-analysis the quantitative trait in POFC-USA

SNP	CHR	BP	P_{meta}	$P_{POFC-USA}$	BETA	N	MAF
rs6663322	1	200649869	9.79E-06	0.5998	-0.1919	194	0.45
rs72488321	2	76805700	8.55E-06	0.1978	-0.5701	187	0.19
rs764629	5	57455043	5.97E-06	0.9678	-0.01472	193	0.35
rs10062700	5	128180057	8.34E-06	0.2113	0.464	194	0.34
rs2860807	5	163265540	7.92E-06	0.9226	0.04375	191	0.27
rs4569988	6	123862050	6.68E-06	0.492	-0.2444	191	0.48
rs2532011	16	4130204	4.16E-06	0.8799	-0.05504	192	0.34
rs17208994	22	19373861	9.85E-06	0.3576	-0.7355	194	0.05

Table 24. Association of the top hits for the trans-ethnic meta-analysis of functional dentition in POFC-USA

SNP	CHR	BP	P_{meta}	$P_{POFC-USA}$	Effect Direction	N	MAF
rs674271	1	78712402	3.01E-06	0.3204	-	192	0.21
rs3930612	2	50896945	3.34E-06	0.2465	-	179	0.20
rs10963759	9	18787638	8.32E-06	0.8685	-	190	0.27
rs28734985	**10**	**60084051**	**8.39E-06**	**0.06653**	+	**191**	**0.07**
rs4575613	18	22697408	1.31E-06	0.5456	-	191	0.45

Bolded text denoted SNPs with $P_{POFC-USA} < 0.10$

BIBLIOGRAPHY

1. Kassebaum, N.J., et al., *Global Burden of Severe Tooth Loss: A Systematic Review and Meta-analysis.* J Dent Res, 2014. **93**(7 Suppl): p. 20S-28S.

2. Listl, S., et al., *Global Economic Impact of Dental Diseases.* J Dent Res, 2015. **94**(10): p. 1355-61.

3. Marcenes, W., et al., *Global burden of oral conditions in 1990-2010: a systematic analysis.* J Dent Res, 2013. **92**(7): p. 592-7.

4. Sischo, L. and H.L. Broder, *Oral health-related quality of life: what, why, how, and future implications.* J Dent Res, 2011. **90**(11): p. 1264-70.

5. *Recent advances in oral health. Report of a WHO Expert Committee.* World Health Organ Tech Rep Ser, 1992. **826**: p. 1-37.

6. Elias, A.C. and A. Sheiham, *The relationship between satisfaction with mouth and number and position of teeth.* J Oral Rehabil, 1998. **25**(9): p. 649-61.

7. Gotfredsen, K. and A.W. Walls, *What dentition assures oral function?* Clin Oral Implants Res, 2007. **18 Suppl 3**: p. 34-45.

8. Sheiham, A., et al., *The impact of oral health on stated ability to eat certain foods; findings from the National Diet and Nutrition Survey of Older People in Great Britain.* Gerodontology, 1999. **16**(1): p. 11-20.

9. Nowjack-Raymer, R.E. and A. Sheiham, *Association of edentulism and diet and nutrition in US adults.* J Dent Res, 2003. **82**(2): p. 123-6.

10. Sheiham, A., et al., *The relationship among dental status, nutrient intake, and nutritional status in older people.* J Dent Res, 2001. **80**(2): p. 408-13.

11. Zhu, Y. and J.H. Hollis, *Tooth loss and its association with dietary intake and diet quality in American adults.* J Dent, 2014. **42**(11): p. 1428-35.

12. Slade, G.D., A.A. Akinkugbe, and A.E. Sanders, *Projections of U.S. Edentulism prevalence following 5 decades of decline.* J Dent Res, 2014. **93**(10): p. 959-65.

13. Mitchell, J., K. Bennett, and A. Brock-Martin, *Edentulism in high poverty rural counties.* J Rural Health, 2013. **29**(1): p. 30-8.

14. Gilbert, G.H., R.P. Duncan, and B.J. Shelton, *Social determinants of tooth loss.* Health Serv Res, 2003. **38**(6 Pt 2): p. 1843-62.

15. Bergdahl, M., et al., *Difference in apolipoprotein E type 4 allele (APOE epsilon 4) among dentate and edentulous subjects.* Gerodontology, 2008. **25**(3): p. 179-86.

16. Del Brutto, O.H., et al., *Edentulism associates with worse cognitive performance in community-dwelling elders in rural Ecuador: results of the Atahualpa project.* J Community Health, 2014. **39**(6): p. 1097-100.

17. Hasui, Y., et al., *Association of the OGG1 Ser326Cys polymorphism with tooth loss.* J Clin Lab Anal, 2006. **20**(2): p. 47-51.

18. Peres, M.A., et al., *Tooth loss is associated with severe cognitive impairment among older people: findings from a population-based study in Brazil.* Aging Ment Health, 2015. **19**(10): p. 876-84.

19. Stein, P.S., et al., *Tooth loss, dementia and neuropathology in the Nun study.* J Am Dent Assoc, 2007. **138**(10): p. 1314-22; quiz 1381-2.

20. Tsakos, G., et al., *Tooth loss associated with physical and cognitive decline in older adults.* J Am Geriatr Soc, 2015. **63**(1): p. 91-9.

21. Peres, M.A., et al., *Tooth loss is associated with increased blood pressure in adults--a multidisciplinary population-based study.* J Clin Periodontol, 2012. **39**(9): p. 824-33.

22. Al-Shammari, K.F., et al., *Risk indicators for tooth loss due to periodontal disease.* Journal of Periodontology, 2005. **76**(11): p. 1910-1918.

23. Ylostalo, P.V., et al., *Gingivitis, dental caries and tooth loss: risk factors for cardiovascular diseases or indicators of elevated health risks.* J Clin Periodontol, 2006. **33**(2): p. 92-101.

24. Ong, G., *Periodontal disease and tooth loss.* Int Dent J, 1998. **48**(3 Suppl 1): p. 233-8.

25. Tsai, S.J., et al., *Factors associated with having less than 20 natural teeth in rural adults: a cross-sectional study.* BMC Oral Health, 2015. **15**: p. 158.

26. Lowe, G., et al., *Total tooth loss and prevalent cardiovascular disease in men and women: possible roles of citrus fruit consumption, vitamin C, and inflammatory and thrombotic variables.* J Clin Epidemiol, 2003. **56**(7): p. 694-700.

27. Elter, J.R., et al., *Relationship of periodontal disease and tooth loss to prevalence of coronary heart disease.* J Periodontol, 2004. **75**(6): p. 782-90.

28. Esmaillzadeh, A., et al., *Is tooth loss associated with irritable bowel syndrome?* J Oral Rehabil, 2015. **42**(7): p. 503-11.

29. Hull, P.S., et al., *The reasons for tooth extractions in adults and their validation.* J Dent, 1997. **25**(3-4): p. 233-7.

30. Hansson, P., et al., *Relationship between natural teeth and memory in a healthy elderly population.* Eur J Oral Sci, 2013. **121**(4): p. 333-40.

31. Bergdahl, M., et al., *Natural teeth and cognitive function in humans.* Scand J Psychol, 2007. **48**(6): p. 557-65.

32. Gatz, M., et al., *Potentially modifiable risk factors for dementia in identical twins.* Alzheimers Dement, 2006. **2**(2): p. 110-7.

33. Luo, J., et al., *Association between tooth loss and cognitive function among 3063 Chinese older adults: a community-based study.* PLoS One, 2015. **10**(3): p. e0120986.

34. Okamoto, N., et al., *Relationship of tooth loss to mild memory impairment and cognitive impairment: findings from the Fujiwara-kyo study.* Behav Brain Funct, 2010. **6**: p. 77.

35. Reyes-Ortiz, C.A., et al., *Self-reported tooth loss and cognitive function: Data from the Hispanic Established Populations for Epidemiologic Studies of the Elderly (Hispanic EPESE).* Colomb Med (Cali), 2013. **44**(3): p. 139-45.

36. Stewart, R. and V. Hirani, *Dental health and cognitive impairment in an English national survey population.* J Am Geriatr Soc, 2007. **55**(9): p. 1410-4.

37. Stewart, R., et al., *Oral health and cognitive function in the Third National Health and Nutrition Examination Survey (NHANES III).* Psychosom Med, 2008. **70**(8): p. 936-41.

38. Kaye, E.K., et al., *Tooth loss and periodontal disease predict poor cognitive function in older men.* J Am Geriatr Soc, 2010. **58**(4): p. 713-8.

39. Batty, G.D., et al., *Oral disease in relation to future risk of dementia and cognitive decline: prospective cohort study based on the Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified-Release Controlled Evaluation (ADVANCE) trial.* Eur Psychiatry, 2013. **28**(1): p. 49-52.

40. Naorungroj, S., et al., *Cognitive decline and oral health in middle-aged adults in the ARIC study.* J Dent Res, 2013. **92**(9): p. 795-801.

41. Yamazaki, K., et al., *Effect of tooth loss on spatial memory and trkB-mRNA levels in rats.* Hippocampus, 2008. **18**(6): p. 542-7.

42. Oue, H., et al., *Tooth loss induces memory impairment and neuronal cell loss in APP transgenic mice.* Behav Brain Res, 2013. **252**: p. 318-25.

43. Su, S., et al., *Tooth loss inhibits neurogenesis in the dentate gyrus of adult mice.* Neural Regen Res, 2014. **9**(17): p. 1606-9.

44. Kamer, A.R., et al., *Inflammation and Alzheimer's disease: possible role of periodontal diseases.* Alzheimers Dement, 2008. **4**(4): p. 242-50.

45. Su, F., F. Bai, and Z. Zhang, *Inflammatory Cytokines and Alzheimer's Disease: A Review from the Perspective of Genetic Polymorphisms.* Neurosci Bull, 2016. **32**(5): p. 469-80.

46. Onozuka, M., et al., *Age-related changes in brain regional activity during chewing: a functional magnetic resonance imaging study.* J Dent Res, 2003. **82**(8): p. 657-60.

47. Tucker, K.L., et al., *High homocysteine and low B vitamins predict cognitive decline in aging men: the Veterans Affairs Normative Aging Study.* Am J Clin Nutr, 2005. **82**(3): p. 627-35.

48. Boraas, J.C., L.B. Messer, and M.J. Till, *A genetic contribution to dental caries, occlusion, and morphology as demonstrated by twins reared apart.* J Dent Res, 1988. **67**(9): p. 1150-5.

49. Bretz, W.A., et al., *Dental caries and microbial acid production in twins.* Caries Res, 2005. **39**(3): p. 168-72.

50. Bretz, W.A., et al., *Longitudinal analysis of heritability for dental caries traits.* J Dent Res, 2005. **84**(11): p. 1047-51.

51. Bretz, W.A., et al., *Heritability estimates for dental caries and sucrose sweetness preference.* Arch Oral Biol, 2006. **51**(12): p. 1156-60.

52. Wang, X., et al., *Genes and their effects on dental caries may differ between primary and permanent dentitions.* Caries Res, 2010. **44**(3): p. 277-84.

53. Shaffer, J.R., et al., *Genetic susceptibility to dental caries on pit and fissure and smooth surfaces.* Caries Res, 2012. **46**(1): p. 38-46.

54. Michalowicz, B.S., et al., *Periodontal findings in adult twins.* J Periodontol, 1991. **62**(5): p. 293-9.

55. Corey, L.A., et al., *Self-reported periodontal disease in a Virginia twin population.* J Periodontol, 1993. **64**(12): p. 1205-8.

56. Michalowicz, B.S., et al., *Evidence of a substantial genetic basis for risk of adult periodontitis.* J Periodontol, 2000. **71**(11): p. 1699-707.

57. Mucci, L.A., et al., *Environmental and heritable factors in the etiology of oral diseases--a population-based study of Swedish twins.* J Dent Res, 2005. **84**(9): p. 800-5.

58. Diehl, S.R., et al., *Quantitative measures of aggressive periodontitis show substantial heritability and consistency with traditional diagnoses.* J Periodontol, 2005. **76**(2): p. 279-88.

59. Shaffer, J.R., et al., *Genome-wide association scan for childhood caries implicates novel genes.* J Dent Res, 2011. **90**(12): p. 1457-62.

60. Wang, X., et al., *Genome-wide association scan of dental caries in the permanent dentition.* BMC Oral Health, 2012. **12**: p. 57.

61. Shaffer, J.R., et al., *GWAS of dental caries patterns in the permanent dentition.* J Dent Res, 2013. **92**(1): p. 38-44.

62. Zeng, Z., et al., *Genome-wide association studies of pit-and-fissure- and smooth-surface caries in permanent dentition.* J Dent Res, 2013. **92**(5): p. 432-7.

63. Zeng, Z., et al., *Genome-wide association study of primary dentition pit-and-fissure and smooth surface caries.* Caries Res, 2014. **48**(4): p. 330-8.

64. Morrison, J., et al., *Genome-wide association study of dental caries in the Hispanic Communities Health Study/Study of Latinos (HCHS/SOL).* Hum Mol Genet, 2016. **25**(4): p. 807-16.

65. Schaefer, A.S., et al., *A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis.* Hum Mol Genet, 2010. **19**(3): p. 553-62.

66. Divaris, K., et al., *Genome-wide association study of periodontal pathogen colonization.* J Dent Res, 2012. **91**(7 Suppl): p. 21S-28S.

67. Teumer, A., et al., *Genome-wide association study of chronic periodontitis in a general German population.* J Clin Periodontol, 2013. **40**(11): p. 977-85.

68. Divaris, K., et al., *Exploring the genetic basis of chronic periodontitis: a genome-wide association study.* Hum Mol Genet, 2013. **22**(11): p. 2312-24.

69. Feng, P., et al., *Genome wide association scan for chronic periodontitis implicates novel locus.* BMC Oral Health, 2014. **14**: p. 84.

70. Freitag-Wolf, S., et al., *Genome-wide exploration identifies sex-specific genetic effects of alleles upstream NPY to increase the risk of severe periodontitis in men.* J Clin Periodontol, 2014. **41**(12): p. 1115-21.

71. Shaffer, J.R., et al., *Genome-wide association study of periodontal health measured by probing depth in adults ages 18-49 years.* G3 (Bethesda), 2014. **4**(2): p. 307-14.

72. Shimizu, S., et al., *A genome-wide association study of periodontitis in a Japanese population.* J Dent Res, 2015. **94**(4): p. 555-61.

73. Offenbacher, S., et al., *Genome-wide association study of biologically informed periodontal complex traits offers novel insights into the genetic basis of periodontal disease.* Hum Mol Genet, 2016. **25**(10): p. 2113-2129.

74. Munz, M., et al., *A genome-wide association study identifies nucleotide variants at SIGLEC5 and DEFA1A3 as risk loci for periodontitis.* Hum Mol Genet, 2017. **26**(13): p. 2577-2588.

75. Sanders, A.E., et al., *Chronic Periodontitis Genome-wide Association Study in the Hispanic Community Health Study / Study of Latinos.* J Dent Res, 2017. **96**(1): p. 64-72.

76. Kurushima, Y., et al., *Heritability of the Number of Teeth in Middle-Aged and Older Danish Twins.* J Dent Res, 2017: p. 22034517724782.

77. Hirano, H., et al., *Association of natural tooth loss with genetic variation at the human matrix Gla protein locus in elderly women.* J Hum Genet, 2003. **48**(6): p. 288-92.

78. Inagaki, K., et al., *Vitamin D receptor alleles, periodontal disease progression, and tooth loss in the VA dental longitudinal study.* J Periodontol, 2003. **74**(2): p. 161-7.

79. Taguchi, A., et al., *Association of estrogen and vitamin D receptor gene polymorphisms with tooth loss and oral bone loss in Japanese postmenopausal women.* Menopause, 2003. **10**(3): p. 250-7.

80. Taguchi, A., et al., *Relationship between estrogen-receptor genotype and tooth loss in postmenopausal women.* JAMA, 2001. **286**(18): p. 2234-5.

81. Toomes, C., et al., *Loss-of-function mutations in the cathepsin C gene result in periodontal disease and palmoplantar keratosis.* Nat Genet, 1999. **23**(4): p. 421-4.

82. Hart, T.C. and J.C. Atkinson, *Mendelian forms of periodontitis.* Periodontol 2000, 2007. **45**: p. 95-112.

83. Wright, J.T., *The molecular etiologies and associated phenotypes of amelogenesis imperfecta.* Am J Med Genet A, 2006. **140**(23): p. 2547-55.

84. Song, S., et al., *WNT10A variants are associated with non-syndromic tooth agenesis in the general population.* Hum Genet, 2014. **133**(1): p. 117-24.

85. Nieminen, P., *Genetic basis of tooth agenesis.* J Exp Zool B Mol Dev Evol, 2009. **312B**(4): p. 320-42.

86. Kruglyak, L., *The road to genome-wide association studies.* Nat Rev Genet, 2008. **9**(4): p. 314-8.

87. Sham, P.C. and S.M. Purcell, *Statistical power and significance testing in large-scale genetic studies.* Nat Rev Genet, 2014. **15**(5): p. 335-46.

88. Manolio, T.A., *Bringing genome-wide association findings into clinical use.* Nat Rev Genet, 2013. **14**(8): p. 549-58.

89. Jones, A.V., et al., *GWAS of self-reported mosquito bite size, itch intensity and attractiveness to mosquitoes implicates immune-related predisposition loci.* Hum Mol Genet, 2017. **26**(7): p. 1391-1406.

90. Cornelis, M.C., et al., *The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions.* Genet Epidemiol, 2010. **34**(4): p. 364-72.

91. Gunderson, K.L., et al., *Whole-genome genotyping.* Methods Enzymol, 2006. **410**: p. 359-76.

92. Peiffer, D.A., et al., *High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping.* Genome Res, 2006. **16**(9): p. 1136-48.

93. Browning, B.L. and S.R. Browning, *A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals.* Am J Hum Genet, 2009. **84**(2): p. 210-23.

94. Leslie, E.J., et al., *A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13.* Hum Mol Genet, 2016. **25**(13): p. 2862-2872.

95. Leslie, E.J., et al., *A Genome-wide Association Study of Nonsyndromic Cleft Palate Identifies an Etiologic Missense Variant in GRHL3.* Am J Hum Genet, 2016. **98**(4): p. 744-54.

96. Laurie, C.C., et al., *Quality control and quality assurance in genotypic data for genome-wide association studies.* Genet Epidemiol, 2010. **34**(6): p. 591-602.

97. Delaneau, O., J.F. Zagury, and J. Marchini, *Improved whole-chromosome phasing for disease and population genetic studies.* Nat Methods, 2013. **10**(1): p. 5-6.

98. Howie, B.N., P. Donnelly, and J. Marchini, *A flexible and accurate genotype imputation method for the next generation of genome-wide association studies.* PLoS Genet, 2009. **5**(6): p. e1000529.

99. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* Am J Hum Genet, 2007. **81**(3): p. 559-75.

100. Aulchenko, Y.S., et al., *GenABEL: an R library for genome-wide association analysis.* Bioinformatics, 2007. **23**(10): p. 1294-6.

101. Kang, H.M., et al., *Variance component model to account for sample structure in genome-wide association studies.* Nat Genet, 2010. **42**(4): p. 348-54.

102. Willer, C.J., Y. Li, and G.R. Abecasis, *METAL: fast and efficient meta-analysis of genomewide association scans.* Bioinformatics, 2010. **26**(17): p. 2190-1.

103. Pruim, R.J., et al., *LocusZoom: regional visualization of genome-wide association scan results.* Bioinformatics, 2010. **26**(18): p. 2336-7.

104. Ward, L.D. and M. Kellis, *HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants.* Nucleic Acids Res, 2012. **40**(Database issue): p. D930-4.

105. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.

106. Noor, A., et al., *Disruption at the PTCHD1 Locus on Xp22.11 in Autism spectrum disorder and intellectual disability.* Sci Transl Med, 2010. **2**(49): p. 49ra68.

107. Chaudhry, A., et al., *Phenotypic spectrum associated with PTCHD1 deletions and truncating mutations includes intellectual disability and autism spectrum disorder.* Clin Genet, 2015. **88**(3): p. 224-33.

108. Tora, D., et al., *Cellular Functions of the Autism Risk Factor PTCHD1 in Mice.* J Neurosci, 2017. **37**(49): p. 11993-12005.

109. Martinez, C., et al., *Sonic hedgehog stimulates proliferation of human periodontal ligament stem cells.* J Dent Res, 2011. **90**(4): p. 483-8.

110. Liu, Y., et al., *An Nfic-hedgehog signaling cascade regulates tooth root development.* Development, 2015. **142**(19): p. 3374-82.

111. Bae, W.J., et al., *Sonic Hedgehog Promotes Cementoblastic Differentiation via Activating the BMP Pathways.* Calcif Tissue Int, 2016. **99**(4): p. 396-407.

112. Hu, X., et al., *Expression of SHH signaling molecules in the developing human primary dentition.* BMC Dev Biol, 2013. **13**: p. 11.

113. Yuan, L., et al., *Identification of a novel PHEX mutation in a Chinese family with X-linked hypophosphatemic rickets using exome sequencing.* Biol Chem, 2015. **396**(1): p. 27-33.

114. Cheon, C.K., et al., *A novel de novo mutation within PHEX gene in a young girl with hypophosphatemic rickets and review of literature.* Ann Pediatr Endocrinol Metab, 2014. **19**(1): p. 36-41.

115. Schreiner, A., et al., *Junction protein shrew-1 influences cell invasion and interacts with invasion-promoting protein CD147.* Mol Biol Cell, 2007. **18**(4): p. 1272-81.

116. Schwab, W., et al., *Immunocytochemical and biochemical detection of EMMPRIN in the rat tooth germ: differentiation-dependent co-expression with MMPs and co-localization with caveolin-1 in membrane rafts of dental epithelial cells.* Histochem Cell Biol, 2007. **128**(3): p. 195-203.

117. Kumamoto, H. and K. Ooya, *Immunohistochemical detection of MT1-MMP, RECK, and EMMPRIN in ameloblastic tumors.* J Oral Pathol Med, 2006. **35**(6): p. 345-51.

118. Finger, J.H., et al., *The netrin 1 receptors Unc5h3 and Dcc are necessary at multiple choice points for the guidance of corticospinal tract axons.* J Neurosci, 2002. **22**(23): p. 10346-56.

119. Graf, D., et al., *Common mechanisms in development and disease: BMP signaling in craniofacial development.* Cytokine Growth Factor Rev, 2016. **27**: p. 129-39.

120. Kim, E.J., et al., *Failure of Tooth Formation Mediated by miR-135a Overexpression via BMP Signaling.* J Dent Res, 2014. **93**(6): p. 571-5.

121. Lee, H.K., et al., *Odontogenic ameloblasts-associated protein (ODAM), via phosphorylation by bone morphogenetic protein receptor type IB (BMPR-IB), is implicated in ameloblast differentiation.* J Cell Biochem, 2012. **113**(5): p. 1754-65.

122. Matsui, T., T. Itoh, and M. Fukuda, *Small GTPase Rab12 regulates constitutive degradation of transferrin receptor.* Traffic, 2011. **12**(10): p. 1432-43.

123. Xu, J., M. Fotouhi, and P.S. McPherson, *Phosphorylation of the exchange factor DENND3 by ULK in response to starvation activates Rab12 and induces autophagy.* EMBO Rep, 2015. **16**(6): p. 709-18.

124. Nazki, F.H., A.S. Sameer, and B.A. Ganaie, *Folate: metabolism, genes, polymorphisms and the associated diseases.* Gene, 2014. **533**(1): p. 11-20.

125. Antunes, L.A., et al., *A Polymorphism in the MTRR Gene Is Associated with Early Childhood Caries and Underweight.* Caries Res, 2017. **51**(2): p. 102-108.

126. Stanley, B.O., et al., *Genetic Association of MPPED2 and ACTN2 with Dental Caries.* J Dent Res, 2014. **93**(7): p. 626-32.

127. Alenghat, F.J., et al., *Mechanical control of cAMP signaling through integrins is mediated by the heterotrimeric Galphas protein.* J Cell Biochem, 2009. **106**(4): p. 529-38.

128. Santos de Araujo, R.M., Y. Oba, and K. Moriyama, *Role of regulator of G-protein signaling 2 (RGS2) in periodontal ligament cells under mechanical stress.* Cell Biochem Funct, 2007. **25**(6): p. 753-8.

129. Yousefian, J., et al., *A new experimental model for studying the response of periodontal ligament cells to hydrostatic pressure.* Am J Orthod Dentofacial Orthop, 1995. **108**(4): p. 402-9.

130. Kitagaki, J., et al., *A Putative Association of a Single Nucleotide Polymorphism in GPR126 with Aggressive Periodontitis in a Japanese Population.* PLoS One, 2016. **11**(8): p. e0160765.

131. Du, A., et al., *MicroRNA expression profile of human periodontal ligament cells under the influence of Porphyromonas gingivalis LPS.* J Cell Mol Med, 2016. **20**(7): p. 1329-38.

132. Xu, C., et al., *Apoptotic gene expression by human periodontal ligament cells following cyclic stretch.* J Periodontal Res, 2011. **46**(6): p. 742-8.

133. Cochran, D.L., *Inflammation and bone loss in periodontal disease.* J Periodontol, 2008. **79**(8 Suppl): p. 1569-76.

134. Koide, M., et al., *Osteoprotegerin-deficient male mice as a model for severe alveolar bone loss: comparison with RANKL-overexpressing transgenic male mice.* Endocrinology, 2013. **154**(2): p. 773-82.

135. Mogi, M., et al., *Differential expression of RANKL and osteoprotegerin in gingival crevicular fluid of patients with periodontitis.* J Dent Res, 2004. **83**(2): p. 166-9.

136. Tang, H., et al., *In vivo osteoprotegerin gene therapy preventing bone loss induced by periodontitis.* J Periodontal Res, 2015. **50**(4): p. 434-43.

137. Takayama, I., et al., *Periostin is required for matricellular localization of CCN3 in periodontal ligament of mice.* J Cell Commun Signal, 2017. **11**(1): p. 5-13.

138. Romanos, G.E., et al., *PERIOSTIN: role in formation and maintenance of dental tissues.* J Cell Physiol, 2014. **229**(1): p. 1-5.

139. Wang, X., et al., *Promotion of dentin regeneration via CCN3 modulation on Notch and BMP signaling pathways.* Biomaterials, 2014. **35**(9): p. 2720-9.

140. Du, J. and M. Li, *Functions of Periostin in dental tissues and its role in periodontal tissues' regeneration.* Cell Mol Life Sci, 2017. **74**(23): p. 4279-4286.

141. Walker, J.T., et al., *Periostin as a multifunctional modulator of the wound healing response.* Cell Tissue Res, 2016. **365**(3): p. 453-65.

142. Rios, H., et al., *periostin null mice exhibit dwarfism, incisor enamel defects, and an early-onset periodontal disease-like phenotype.* Mol Cell Biol, 2005. **25**(24): p. 11131-44.

143. Rios, H.F., et al., *Periostin is essential for the integrity and function of the periodontal ligament during occlusal loading in mice.* J Periodontol, 2008. **79**(8): p. 1480-90.

144. Aral, C.A., et al., *Gingival Crevicular Fluid and Salivary Periostin Levels in Non-Smoker Subjects With Chronic and Aggressive Periodontitis : Periostin Levels in Chronic and Aggressive Periodontitis.* Inflammation, 2016. **39**(3): p. 986-93.

145. Tsukamoto, S., et al., *Smad9 is a new type of transcriptional regulator in bone morphogenetic protein signaling.* Sci Rep, 2014. **4**: p. 7596.

146. Yu, Y., et al., *Characterization of the GufA subfamily member SLC39A11/Zip11 as a zinc transporter.* J Nutr Biochem, 2013. **24**(10): p. 1697-708.

147. Lynch, R.J., *Zinc in the mouth, its interactions with dental enamel and possible effects on caries; a review of the literature.* Int Dent J, 2011. **61 Suppl 3**: p. 46-54.

148. Mohammed, N.R., R.J. Lynch, and P. Anderson, *Inhibitory Effects of Zinc Ions on Enamel Demineralisation Kinetics in vitro.* Caries Res, 2015. **49**(6): p. 600-5.

149. Hegde, M.N., et al., *Biochemical indicators of dental caries in saliva: an in vivo study.* Caries Res, 2014. **48**(2): p. 170-3.

150. Yin, H., et al., *Inhibitory effects of somatostatin on the substantia gelatinosa neurons of trigeminal subnucleus caudalis via somatostatin type 2 receptors in juvenile mice.* Brain Res, 2009. **1304**: p. 49-56.

151. Takahashi, M., M. Takeda, and S. Matsumoto, *Somatostatin inhibits tooth-pulp-evoked rat cervical dorsal horn neuronal activity.* Exp Brain Res, 2008. **184**(4): p. 617-22.

152. Takahashi, M., M. Takeda, and S. Matsumoto, *Somatostatin enhances tooth-pulp-evoked cervical dorsal horn neuronal activity in the rat via inhibition of GABAergic interneurons.* Brain Res Bull, 2014. **100**: p. 76-83.

153. Sengle, G. and L.Y. Sakai, *The fibrillin microfibril scaffold: A niche for growth factors and mechanosensation?* Matrix Biol, 2015. **47**: p. 3-12.

154. Smaldone, S. and F. Ramirez, *Fibrillin microfibrils in bone physiology.* Matrix Biol, 2016. **52-54**: p. 191-197.

155. Tsuruga, E., et al., *Expression of fibrillins and tropoelastin by human gingival and periodontal ligament fibroblasts in vitro.* J Periodontal Res, 2002. **37**(1): p. 23-8.

156. Tsuruga, E., K. Irie, and T. Yajima, *Gene expression and accumulation of fibrillin-1, fibrillin-2, and tropoelastin in cultured periodontal fibroblasts.* J Dent Res, 2002. **81**(11): p. 771-5.

157. Sugawara, Y., et al., *Immunohistochemical localization of elastin, fibrillins and microfibril-associated glycoprotein-1 in the developing periodontal ligament of the rat molar.* J Periodontal Res, 2010. **45**(1): p. 52-9.

158. Kira-Tatsuoka, M., et al., *Immunohistochemical expression of fibrillin-1 and fibrillin-2 during tooth development.* J Periodontal Res, 2015. **50**(6): p. 714-20.

159. Dean, C. and T. Dresbach, *Neuroligins and neurexins: linking cell adhesion, synapse formation and cognitive function.* Trends Neurosci, 2006. **29**(1): p. 21-9.

160. Reichelt, A.C., R.J. Rodgers, and S.J. Clapcote, *The role of neurexins in schizophrenia and autistic spectrum disorder.* Neuropharmacology, 2012. **62**(3): p. 1519-26.

161. Zhang, S., et al., *The Novel ASIC2 Locus is Associated with Severe Gingival Inflammation.* JDR Clin Trans Res, 2016. **1**(2): p. 163-170.

162. Gong, A.X., et al., *Comparison of gene expression profiles between dental pulp and periodontal ligament tissues in humans.* Int J Mol Med, 2017. **40**(3): p. 647-660.

163. Okbay, A., et al., *Genome-wide association study identifies 74 loci associated with educational attainment.* Nature, 2016. **533**(7604): p. 539-42.

164. Rietveld, C.A., et al., *Common genetic variants associated with cognitive performance identified using the proxy-phenotype method.* Proc Natl Acad Sci U S A, 2014. **111**(38): p. 13790-4.

165. Maeda, Y., et al., *Indian Hedgehog produced by postnatal chondrocytes is essential for maintaining a growth plate and trabecular bone.* Proc Natl Acad Sci U S A, 2007. **104**(15): p. 6382-7.

166. Byrnes, A.M., et al., *Brachydactyly A-1 mutations restricted to the central region of the N-terminal active fragment of Indian Hedgehog.* Eur J Hum Genet, 2009. **17**(9): p. 1112-20.

167. Ishizuka, Y., et al., *TMJ degeneration in SAMP8 mice is accompanied by deranged Ihh signaling.* J Dent Res, 2014. **93**(3): p. 281-7.

168. Sarkar, L. and P.T. Sharpe, *Expression of Wnt signalling pathway genes during tooth development.* Mech Dev, 1999. **85**(1-2): p. 197-200.

169. Suomalainen, M. and I. Thesleff, *Patterns of Wnt pathway activity in the mouse incisor indicate absence of Wnt/beta-catenin signaling in the epithelial stem cells.* Dev Dyn, 2010. **239**(1): p. 364-72.

170. Wang, C., et al., *Effect of Wnt6 on human dental papilla cells in vitro.* J Endod, 2010. **36**(2): p. 238-43.

171. Li, R., et al., *WNT6 promotes the migration and differentiation of human dental pulp cells partly through c-Jun N-terminal kinase signaling pathway.* J Endod, 2014. **40**(7): p. 943-8.

172. Liu, K., H. Meng, and J. Hou, *Activity of 25-hydroxylase in human gingival fibroblasts and periodontal ligament cells.* PLoS One, 2012. **7**(12): p. e52053.

173. Liu, K., et al., *Initial periodontal therapy reduced systemic and local 25-hydroxy vitamin D(3) and interleukin-1beta in patients with aggressive periodontitis.* J Periodontol, 2010. **81**(2): p. 260-6.

174. Kim, E.K. and E.J. Choi, *Compromised MAPK signaling in human diseases: an update.* Arch Toxicol, 2015. **89**(6): p. 867-82.

175. Stein, J., et al., *Are there HLA combinations typical supporting for or making resistant against aggressive and/or chronic periodontitis?* J Periodontal Res, 2003. **38**(5): p. 508-17.

176. Valarini, N., et al., *Association of dental caries with HLA Class II allele in Brazilian adolescents.* Caries Res, 2012. **46**(6): p. 530-5.

177. Bufe, B., et al., *The human TAS2R16 receptor mediates bitter taste in response to beta-glucopyranosides.* Nat Genet, 2002. **32**(3): p. 397-401.

178. Spiegel, I., et al., *Caspr3 and caspr4, two novel members of the caspr family are expressed in the nervous system and interact with PDZ domains.* Mol Cell Neurosci, 2002. **20**(2): p. 283-97.

179. Thomas, J.A., et al., *Impaired cytokine signaling in mice lacking the IL-1 receptor-associated kinase.* J Immunol, 1999. **163**(2): p. 978-84.

180. Hajishengallis, G. and S.E. Sahingur, *Novel inflammatory pathways in periodontitis.* Adv Dent Res, 2014. **26**(1): p. 23-9.

181. Pfeiffer, J.R., et al., *CARHSP1 is required for effective tumor necrosis factor alpha mRNA stabilization and localizes to processing bodies and exosomes.* Mol Cell Biol, 2011. **31**(2): p. 277-86.

182. Monetti, M., et al., *The presence of periodontopathogens associated with the tumour necrosis factor-alpha expression in patients with different periodontal status.* Acta Odontol Latinoam, 2012. **25**(1): p. 82-8.

183. Cardoso, C.R., et al., *Characterization of CD4+CD25+ natural regulatory T cells in the inflammatory infiltrate of human chronic periodontitis.* J Leukoc Biol, 2008. **84**(1): p. 311-8.

184. Araujo-Pires, A.C., et al., *IL-4/CCL22/CCR4 axis controls regulatory T-cell migration that suppresses inflammatory bone loss in murine experimental periodontitis.* J Bone Miner Res, 2015. **30**(3): p. 412-22.

185. Hosokawa, Y., et al., *CC chemokine ligand 17 in periodontal diseases: expression in diseased tissues and production by human gingival fibroblasts.* J Periodontal Res, 2008. **43**(4): p. 471-7.

186. Cao, H., et al., *Tbx1 regulates progenitor cell proliferation in the dental epithelium by modulating Pitx2 activation of p21.* Dev Biol, 2010. **347**(2): p. 289-300.

187. Caton, J., et al., *Enamel-free teeth: Tbx1 deletion affects amelogenesis in rodent incisors.* Dev Biol, 2009. **328**(2): p. 493-505.

188. Gao, S., et al., *TBX1 protein interactions and microRNA-96-5p regulation controls cell proliferation during craniofacial and dental development: implications for 22q11.2 deletion syndrome.* Hum Mol Genet, 2015. **24**(8): p. 2330-48.

189. Sunohara, M., et al., *Expression patterns of Raf-1 suggest multiple roles in tooth development.* Calcif Tissue Int, 1996. **58**(1): p. 60-4.

190. Kumamoto, H., N. Takahashi, and K. Ooya, *K-Ras gene status and expression of Ras/mitogen-activated protein kinase (MAPK) signaling molecules in ameloblastomas.* J Oral Pathol Med, 2004. **33**(6): p. 360-7.

191. Kubota, T., et al., *Altered gene expression levels of matrix metalloproteinases and their inhibitors in periodontitis-affected gingival tissue.* J Periodontol, 2008. **79**(1): p. 166-73.

192. Ramamurthy, N.S., et al., *Experimental arthritis in rats induces biomarkers of periodontitis which are ameliorated by gene therapy with tissue inhibitor of matrix metalloproteinases.* J Periodontol, 2005. **76**(2): p. 229-33.

193. Astle, W.J., et al., *The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease.* Cell, 2016. **167**(5): p. 1415-1429 e19.

194. Arts, H. and N. Knoers, *Cranioectodermal Dysplasia*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 1993: Seattle (WA).

195. Zhou, H., et al., *Inhibition of serum- and glucocorticoid-inducible kinase 1 enhances TLR-mediated inflammation and promotes endotoxin-driven organ failure.* FASEB J, 2015. **29**(9): p. 3737-49.

196. Lee, J.S., et al., *Evaluating the oxysterol combination of 22(S)-hydroxycholesterol and 20(S)-hydroxycholesterol in periodontal regeneration using periodontal ligament stem cells and alveolar bone healing models.* Stem Cell Res Ther, 2017. **8**(1): p. 276.

197. Yang, X., et al., *Role of transient receptor potential channel 6 in the odontogenic differentiation of human dental pulp cells.* Exp Ther Med, 2017. **14**(1): p. 73-78.

198. Guida, V., et al., *A variant in the carboxyl-terminus of connexin 40 alters GAP junctions and increases risk for tetralogy of Fallot.* Eur J Hum Genet, 2013. **21**(1): p. 69-75.

199. Aruga, J., *Slitrk6 expression profile in the mouse embryo and its relationship to that of Nlrr3.* Gene Expr Patterns, 2003. **3**(6): p. 727-33.

200. Biben, C., C.C. Wang, and R.P. Harvey, *NK-2 class homeobox genes and pharyngeal/oral patterning: Nkx2-3 is required for salivary gland and tooth morphogenesis.* Int J Dev Biol, 2002. **46**(4): p. 415-22.

201. Makeyev, A.V. and D. Bayarsaihan, *ChIP-Chip Identifies SEC23A, CFDP1, and NSD1 as TFII-I Target Genes in Human Neural Crest Progenitor Cells.* Cleft Palate Craniofac J, 2013. **50**(3): p. 347-50.

202. Kettunen, P., et al., *Associations of FGF-3 and FGF-10 with signaling networks regulating tooth morphogenesis.* Dev Dyn, 2000. **219**(3): p. 322-32.

203. Lee, H.S., et al., *Comparative gene-expression analysis of the dental follicle and periodontal ligament in humans.* PLoS One, 2013. **8**(12): p. e84201.

204. Pemberton, T.J., et al., *Identification of novel genes expressed during mouse tooth development by microarray gene expression analysis.* Dev Dyn, 2007. **236**(8): p. 2245-57.

205. Zhang, Y., et al., *Multiple effects of the cellular prion protein on tooth development.* Int J Dev Biol, 2011. **55**(10-12): p. 953-60.

206. de Jong, T.M., et al., *SLC23A1 polymorphism rs6596473 in the vitamin C transporter SVCT1 is associated with aggressive periodontitis.* J Clin Periodontol, 2014. **41**(6): p. 531-40.

207. Stevens, C.A., *Rubinstein-Taybi Syndrome*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 1993: Seattle (WA).

208. Johannessen, M., M.P. Delghandi, and U. Moens, *What turns CREB on?* Cell Signal, 2004. **16**(11): p. 1211-27.

209. Klinz, F.J., et al., *Transcription factor CREB is phosphorylated in human molar odontoblasts and cementoblasts in vivo.* Histochem Cell Biol, 2013. **139**(4): p. 615-20.

210. Meikle, M.C., et al., *Gingival fibroblasts degrade type I collagen films when stimulated with tumor necrosis factor and interleukin 1: evidence that breakdown is mediated by metalloproteinases.* J Periodontal Res, 1989. **24**(3): p. 207-13.

211. Ertugrul, A.S., et al., *Comparison of CCL28, interleukin-8, interleukin-1beta and tumor necrosis factor-alpha in subjects with gingivitis, chronic periodontitis and generalized aggressive periodontitis.* J Periodontal Res, 2013. **48**(1): p. 44-51.

212. Escalona, L.A., P. Mastromatteo-Alberga, and M. Correnti, *Cytokine and metalloproteinases in gingival fluid from patients with chronic periodontitis.* Invest Clin, 2016. **57**(2): p. 131-142.

213. Osseni, A., et al., *Triadin and CLIMP-63 form a link between triads and microtubules in muscle cells.* J Cell Sci, 2016. **129**(20): p. 3744-3755.

214. Oddoux, S., et al., *Triadin deletion induces impaired skeletal muscle function.* J Biol Chem, 2009. **284**(50): p. 34918-29.

215. Roux-Buisson, N., et al., *Absence of triadin, a protein of the calcium release complex, is responsible for cardiac arrhythmia with sudden death in human.* Hum Mol Genet, 2012. **21**(12): p. 2759-67.

216. Sherva, R., et al., *Genome-wide association study of the rate of cognitive decline in Alzheimer's disease.* Alzheimers Dement, 2014. **10**(1): p. 45-52.

217. Zhao, S.C., et al., *The structure and function of NKAIN2-a candidate tumor suppressor*. Int J Clin Exp Med, 2015. **8**(10): p. 17072-9.

218. Mao, X., et al., *NKAIN2 functions as a novel tumor suppressor in prostate cancer*. Oncotarget, 2016. **7**(39): p. 63793-63803.

219. McDonald-McGinn, D.M., B.S. Emanuel, and E.H. Zackai, *22q11.2 Deletion Syndrome*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 1993: Seattle (WA).

220. Klingberg, G., et al., *Oral manifestations in 22q11 deletion syndrome.* Int J Paediatr Dent, 2002. **12**(1): p. 14-23.

221. Herold, C., et al., *Family-based association analyses of imputed genotypes reveal genome-wide significant association of Alzheimer's disease with OSBPL6, PTPRG, and PDCL3.* Mol Psychiatry, 2016. **21**(11): p. 1608-1612.

222. Bollag, R.J., et al., *An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus.* Nat Genet, 1994. **7**(3): p. 383-9.

223. Crowley, B., et al., *Variable immune deficiency related to deletion size in chromosome 22q11.2 deletion syndrome.* Am J Med Genet A, 2018.

224. Karimi, M., et al., *Factor XIII Deficiency.* Semin Thromb Hemost, 2009. **35**(4): p. 426-38.

225. Wang, W., et al., *Homozygous intronic mutation leading to inefficient transcription combined with a novel frameshift mutation in F13A1 gene causes FXIII deficiency.* J Hum Genet, 2011. **56**(6): p. 460-3.

226. Lak, M., et al., *Pattern of symptoms in 93 Iranian patients with severe factor XIII deficiency.* J Thromb Haemost, 2003. **1**(8): p. 1852-3.

227. Ivaskevicius, V., et al., *A common F13A1 intron 1 variant IVS1+12(A) is associated with mild FXIII deficiency in Caucasian population.* Ann Hematol, 2013. **92**(7): p. 975-9.

228. Hirohata, S., et al., *Punctin, a novel ADAMTS-like molecule, ADAMTSL-1, in extracellular matrix.* J Biol Chem, 2002. **277**(14): p. 12182-9.

229. Hendee, K., et al., *Identification and functional analysis of an ADAMTSL1 variant associated with a complex phenotype including congenital glaucoma, craniofacial, and other systemic features in a three-generation human pedigree.* Hum Mutat, 2017. **38**(11): p. 1485-1490.

230. Zollner, S. and J.K. Pritchard, *Overcoming the winner's curse: estimating penetrance parameters from case-control data.* Am J Hum Genet, 2007. **80**(4): p. 605-15.

231. Nishikawa, S. and T. Kawamoto, *Localization of Core Planar Cell Polarity Proteins, PRICKLEs, in Ameloblasts of Rat Incisors: Possible Regulation of Enamel Rod Decussation.* Acta Histochem Cytochem, 2015. **48**(2): p. 37-45.

232. Jungers, K.A., et al., *Adamts9 is widely expressed during mouse embryo development.* Gene Expr Patterns, 2005. **5**(5): p. 609-17.

233. Kero, D., et al., *Expression of cytokeratin 8, vimentin, syndecan-1 and Ki-67 during human tooth development.* J Mol Histol, 2014. **45**(6): p. 627-40.

234. Kero, D., et al., *Expression of Ki-67, Oct-4, gamma-tubulin and alpha-tubulin in human tooth development.* Arch Oral Biol, 2014. **59**(11): p. 1119-29.

235. Li, H., et al., *Expression and function of NUMB in odontogenesis.* Biomed Res Int, 2013. **2013**: p. 182965.

236. Felszeghy, S., et al., *Expression of aquaporin isoforms during human and mouse tooth development.* Arch Oral Biol, 2004. **49**(4): p. 247-57.

237. Yamaguchi, T., et al., *Comprehensive genetic exploration of selective tooth agenesis of mandibular incisors by exome sequencing.* Hum Genome Var, 2017. **4**: p. 17005.

238. Faily, S., et al., *Confirmation that mutations in DDX59 cause an autosomal recessive form of oral-facial-digital syndrome: Further delineation of the DDX59 phenotype in two new families.* Eur J Med Genet, 2017. **60**(10): p. 527-532.

239. Don, K.R., et al., *Promoter hypermethylation patterns of P16, DAPK and MGMT in oral squamous cell carcinoma: a systematic review and meta-analysis.* Indian J Dent Res, 2014. **25**(6): p. 797-805.

240. Huang, S.H., et al., *Loss expression of O6-methylguanine DNA methyltransferase by promoter hypermethylation and its relationship to betel quid chewing in oral squamous cell carcinoma.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2010. **109**(6): p. 883-9.

241. Lu, S.Y., et al., *Ripe areca nut extract induces G1 phase arrests and senescence-associated phenotypes in normal human oral keratinocyte.* Carcinogenesis, 2006. **27**(6): p. 1273-84.

242. Lee, S.S., et al., *The expression of O(6) -methylguanine-DNA methyltransferase in human oral keratinocytes stimulated with arecoline.* J Oral Pathol Med, 2013. **42**(8): p. 600-5.

243. Niki, T., et al., *DJBP: a novel DJ-1-binding protein, negatively regulates the androgen receptor by recruiting histone deacetylase complex, and DJ-1 antagonizes this inhibition by abrogation of this complex.* Mol Cancer Res, 2003. **1**(4): p. 247-61.

244. Dale, J.B., et al., *Hormonal regulation of androgen receptor messenger ribonucleic acid expression in human tooth pulp.* J Dent Res, 2002. **81**(5): p. 360-5.

245. Inaba, T., et al., *Expression status of mRNA for sex hormone receptors in human dental pulp cells and the response to sex hormones in the cells.* Arch Oral Biol, 2013. **58**(8): p. 943-50.

246. Huang, W.T., et al., *Immunohistochemical analysis of Th1/Th2 cytokine profiles and androgen receptor expression in the pathogenesis of nifedipine-induced gingival overgrowth.* J Periodontal Res, 2003. **38**(4): p. 422-7.

247. Kuhnisch, J., et al., *Genome-wide association study (GWAS) for molar-incisor hypomineralization (MIH).* Clin Oral Investig, 2014. **18**(2): p. 677-82.

248. Bullwinkel, J., et al., *Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells.* J Cell Physiol, 2006. **206**(3): p. 624-35.

249. Ma, J., et al., *Cyclic stretch induced gene expression of extracellular matrix and adhesion molecules in human periodontal ligament cells.* Arch Oral Biol, 2015. **60**(3): p. 447-55.

250. Silva, P.V.D., et al., *Increased activity of the antioxidants systems modulate the oxidative stress in saliva of toddlers with early childhood caries.* Arch Oral Biol, 2016. **70**: p. 62-66.

251. Waddington, R.J., R. Moseley, and G. Embery, *Reactive oxygen species: a potential role in the pathogenesis of periodontal diseases.* Oral Dis, 2000. **6**(3): p. 138-51.

252. Miricescu, D., et al., *Salivary biomarkers: relationship between oxidative stress and alveolar bone loss in chronic periodontitis.* Acta Odontol Scand, 2014. **72**(1): p. 42-7.

253. Almerich-Silla, J.M., et al., *Oxidative Stress Parameters in Saliva and Its Association with Periodontal Disease and Types of Bacteria.* Dis Markers, 2015. **2015**: p. 653537.

254. Chen, H., et al., *Spatial regulation of Aurora A activity during mitotic spindle assembly requires RHAMM to correctly localize TPX2.* Cell Cycle, 2014. **13**(14): p. 2248-61.

255. Schutze, A., et al., *RHAMM splice variants confer radiosensitivity in human breast cancer cell lines.* Oncotarget, 2016. **7**(16): p. 21428-40.

256. Niedworok, C., et al., *The impact of the receptor of hyaluronan-mediated motility (RHAMM) on human urothelial transitional cell cancer of the bladder.* PLoS One, 2013. **8**(9): p. e75681.

257. Maxwell, C.A., et al., *RHAMM expression and isoform balance predict aggressive disease and poor survival in multiple myeloma.* Blood, 2004. **104**(4): p. 1151-8.

258. Halim, A.B., et al., *Expression and functional interaction of the catalytic and regulatory subunits of human methionine adenosyltransferase in mammalian cells.* J Biol Chem, 1999. **274**(42): p. 29720-5.

259. Bohnekamp, J. and T. Schoneberg, *Cell adhesion receptor GPR133 couples to Gs protein.* J Biol Chem, 2011. **286**(49): p. 41912-6.

260. Blockus, H. and A. Chedotal, *Slit-Robo signaling.* Development, 2016. **143**(17): p. 3037-44.

261. Hashemi, S.M., T.J. Hund, and P.J. Mohler, *Cardiac ankyrins in health and disease.* J Mol Cell Cardiol, 2009. **47**(2): p. 203-9.

262. Shibata, T., et al., *Identification of human cadherin-14, a novel neurally specific type II cadherin, by protein interaction cloning.* J Biol Chem, 1997. **272**(8): p. 5236-40.

263. Christian, L., H. Bahudhanapati, and S. Wei, *Extracellular metalloproteinases in neural crest development and craniofacial morphogenesis.* Crit Rev Biochem Mol Biol, 2013. **48**(6): p. 544-60.

264. Gulino, A., L. Di Marcotullio, and I. Screpanti, *The multiple functions of Numb.* Exp Cell Res, 2010. **316**(6): p. 900-6.

265. Ishizuka, M., et al., *Changes in apolipoprotein B and oxidized low-density lipoprotein levels in gingival crevicular fluids as a result of periodontal tissue conditions.* J Periodontal Res, 2017. **52**(3): p. 594-602.

266. Kim, K.M., et al., *Gene expression profiling of oral epithelium during tooth development.* Arch Oral Biol, 2012. **57**(8): p. 1100-7.

267. Alkharobi, H., et al., *IGFBP-2 and -3 co-ordinately regulate IGF1 induced matrix mineralisation of differentiating human dental pulp cells.* Stem Cell Res, 2016. **17**(3): p. 517-522.

268. Takenouchi, Y., et al., *Insulin-like growth factor-binding protein-2 and -3 in gingival crevicular fluid.* J Periodontal Res, 2010. **45**(6): p. 803-8.

269. Liu, D., et al., *Demethylation of IGFBP5 by Histone Demethylase KDM6B Promotes Mesenchymal Stem Cell-Mediated Periodontal Tissue Regeneration by Enhancing Osteogenic Differentiation and Anti-Inflammation Potentials.* Stem Cells, 2015. **33**(8): p. 2523-36.

270. Rath-Deschner, B., et al., *Regulatory effects of biomechanical strain on the insulin-like growth factor system in human periodontal cells.* J Biomech, 2009. **42**(15): p. 2584-9.

271. Geller, F., et al., *Genome-wide association study identifies four loci associated with eruption of permanent teeth.* PLoS Genet, 2011. **7**(9): p. e1002275.

272. Morimoto, M., et al., *Dental abnormalities in Schimke immuno-osseous dysplasia.* J Dent Res, 2012. **91**(7 Suppl): p. 29S-37S.

273. Nakamura, T., et al., *Involvement of angiotensin II type 1 receptors in interleukin-1beta-induced interleukin-6 production in human gingival fibroblasts.* Eur J Oral Sci, 2011. **119**(5): p. 345-51.

274. Huizing, M., et al., *Hermansky-Pudlak Syndrome*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 1993: Seattle (WA).

275. Harshavardhana, B., S.K. Rath, and M. Mukherjee, *Evaluation of serum ceruloplasmin in aggressive and chronic periodontitis patients.* J Indian Soc Periodontol, 2013. **17**(3): p. 333-7.

276. Iwata, T., et al., *Ceruloplasmin induces polymorphonuclear leukocyte priming in localized aggressive periodontitis.* J Periodontol, 2009. **80**(8): p. 1300-6.

277. Milward, M.R., et al., *Differential activation of NF-kappaB and gene expression in oral epithelial cells by periodontal pathogens.* Clin Exp Immunol, 2007. **148**(2): p. 307-24.

278. Hayano, S., et al., *Roles of heparan sulfate sulfation in dentinogenesis.* J Biol Chem, 2012. **287**(15): p. 12217-29.

279. Park, M.Y., et al., *FAM19A5, a brain-specific chemokine, inhibits RANKL-induced osteoclast formation through formyl peptide receptor 2.* Sci Rep, 2017. **7**(1): p. 15575.

280. King, L.S. and P. Agre, *Pathophysiology of the aquaporin water channels.* Annu Rev Physiol, 1996. **58**: p. 619-48.

281. Buffoli, B., et al., *AQP1 expression in human gingiva and its correlation with periodontal and peri-implant tissue alterations.* Acta Histochem, 2014. **116**(5): p. 898-904.

282. Yoshii, T., et al., *Immunoexpression of aquaporin-1 in the rat periodontal ligament during experimental tooth movement.* Biomed Res, 2012. **33**(4): p. 225-33.

283. Nandasena, B.G., et al., *Immunolocalization of aquaporin-1 in the mechanoreceptive Ruffini endings in the periodontal ligament.* Brain Res, 2007. **1157**: p. 32-40.

284. Luo, Z., et al., *Overexpression and Potential Regulatory Role of IL-17F in Pathogenesis of Chronic Periodontitis.* Inflammation, 2015. **38**(3): p. 978-86.

285. Moretti, S., et al., *Fine-tuning of Th17 Cytokines in Periodontal Disease by IL-10.* J Dent Res, 2015. **94**(9): p. 1267-75.

286. Saraiva, A.M., et al., *Evaluation of IL17A expression and of IL17A, IL17F and IL23R gene polymorphisms in Brazilian individuals with periodontitis.* Hum Immunol, 2013. **74**(2): p. 207-14.

287. Correa, J.D., et al., *Association between polymorphisms in interleukin-17A and -17F genes and chronic periodontal disease.* Mediators Inflamm, 2012. **2012**: p. 846052.

288. Ma, S., J. Charron, and R.L. Erikson, *Role of Plk2 (Snk) in mouse development and cell proliferation.* Mol Cell Biol, 2003. **23**(19): p. 6936-43.

289. Liu, Y. and W. Zhang, *Identification of a new transmembrane adaptor protein that constitutively binds Grb2 in B cells.* J Leukoc Biol, 2008. **84**(3): p. 842-51.

290. Siddiqui, T.J., et al., *An LRRTM4-HSPG complex mediates excitatory synapse development on dentate gyrus granule cells.* Neuron, 2013. **79**(4): p. 680-95.

291. Shamseldin, H.E., et al., *Mutations in DDX59 implicate RNA helicase in the pathogenesis of orofaciodigital syndrome.* Am J Hum Genet, 2013. **93**(3): p. 555-60.

292. Addison, W.N., et al., *Direct transcriptional repression of Zfp423 by Zfp521 mediates a bone morphogenic protein-dependent osteoblast versus adipocyte lineage commitment switch.* Mol Cell Biol, 2014. **34**(16): p. 3076-85.

293. Kim, S.T., et al., *Transcriptome analysis of CD133-positive stem cells and prognostic value of survivin in colorectal cancer.* Cancer Genomics Proteomics, 2014. **11**(5): p. 259-66.

294. Graff, M., et al., *Genome-wide physical activity interactions in adiposity - A meta-analysis of 200,452 adults.* PLoS Genet, 2017. **13**(4): p. e1006528.

295. Kim, E., et al., *Inositol polyphosphate multikinase promotes Toll-like receptor-induced inflammation by stabilizing TRAF6.* Sci Adv, 2017. **3**(4): p. e1602296.