EXPLORING THE GENETIC BASIS OF CHRONIC PERIODONTITIS: A GENOME-WIDE APPROACH

Kimon Divaris

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department Epidemiology of the Gillings School of Global Public Health.

Chapel Hill
2011

Approved by:

Andrew F. Olshan

Ethan M. Lange

Keri L. Monda

Kari E. North

Steven Offenbacher

ABSTRACT

KIMON DIVARIS: Exploring the genetic basis of chronic periodontitis: a genome-wide approach
(Under the direction of Andrew F. Olshan)

Chronic periodontitis (CP) is a common-complex oral disease that affects the majority of the adult population and is a major cause of tooth loss. The disease is characterized by an oral biofilm pathological shift that contributes to cascade of events leading to periodontal destruction. Factors modulating the establishment of a dysbiotic oral microbiome or affecting the host immunity and inflammatory response are promising preventive and therapeutic targets. Although a substantial genetic component of CP is theorized and numerous candidate-gene studies have been completed, to-date no whole-genome association (GWA) studies have been performed.

We performed a GWA analysis of CP in well-defined cohort of 4500 white subjects who were participants of the Atherosclerosis Risk In Communities study. Traits of interest were the three-level disease CDC/AAP classification (healthy/mild, moderate, severe CP) and a continuous extent of disease [proportion of sites exhibiting ≥3 mm attachment loss (EAL)] measure. Additionally, we examined three traits of high bacterial colonization defined as the highest quintile of the distribution of “red” and “orange” complex bacteria, and *Aggregatibacter actinomycetemcomitans* that were quantified using DNA-DNA

checkerboard hybridization in a subset of 1020 white study subjects. Genotyping was performed using the Affymetrix 6.0 platform. Imputation to 2.5million markers was based on HapMap II-CEU and a multiple-test corrected significance threshold was applied ($P<5x10^{-8}$).

We detected no genome-wide significant signals. However, we found suggestive evidence of association ($P<5x10^{-6}$) for CP with ten loci including *NPY, NIN, WNT5A* for severe, *NCR2, EMR1* for moderate CP, and *TBX18*, *ETS1*, *DYNC2H1, TTC26* and *ZC3HAV1* for EAL. Additionally, thirteen loci including *KCNK1, FBXO38, IL33, RUNX2, CAMTA1* and *VAMP3* provided suggestive signals of association ($P<5x10^{-6}$) with the examined "high" bacterial colonization traits. The *NPY* (7p15) locus was replicated in an independent cohort of whites of European descent. These genome-wide scan results from a large well-defined cohort provide information on multiple candidate regions for interrogation in genetic studies of CP. Future investigations providing further replication of these findings may lead to an improved understanding of the complex nature of host-biofilm and -bacteria interactions that characterizes states of health and disease.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS AND SYMBOLS

%	Percent
≥	Greater than or equal to
≤	Less than or equal to
k	Kappa coefficient
λ	Lambda coefficient
μL	Micro liters
Aa	*Aggregatibacter actinomycetemcomitans*
AA	African-American
AAP	American Academy of Periodontology
AgP	Aggressive periodontitis
ARIC	Atherosclerosis Risk In Communities study
b	Beta coefficient
BGI	Biofilm-gingival interface
BMI	Body mass index
BOP	Bleeding upon probing
CAL	Clinical attachment loss
CARe	Candidate-gene Association Resource Consortium
CDC	Centers for Disease Control and Prevention
CEU	Central European
CHD	Coronary heart disease
Chr	Chromosome

cM	Centimorgan ($1/100^{th}$ of a Morgan)
CP	Chronic periodontitis
Cr	*Campylobacter rectus*
CVD	Cardiovascular disease
DM	Diabetes mellitus
DNA	Deoxyribonucleic acid
DST	Allelic sharing distance
DZ	Dizygotic
EAL	Extent of attachment loss
EMM	Effect measure modification
ESI	Extent and severity index
eQTL	Expression quantitative locus
FDR	False discovery rate
Fn	*Fusobacterium nucleatum*
gAgP	Generalized aggressive periodontitis
GWAS	Genome-wide association study
HPFS	Health Professionals Follow-up Study
HWE	Hardy-Weinberg equilibrium
IBD	Identical by descent
IBS	Identical by state
IMT	intima-media thickness
ICC	Intraclass correlation coefficient
IL	Interleukin

Mb	Megabase (1 million base pairs)
MMP	Matrix metalloproteinase
Kb	Kilobase (1 thousand base pairs)
LASSO	Least absolute shrinkage and selection operator
LD	Linkage disequilibrium
MAF	Minor allele frequency
mm	Millimeters
MZ	Monozygotic
N	Number
NHLBI	National Heart, Lung, and Blood Institute
NHANES	National Health and Nutrition Examination Survey
NIDCR	National Institute of Dental and Craniofacial Research
NIH	National Institutes of Health
OHRQoL	oral health-related quality of life
OR	Odds ratio
PC	Principal Component
PD	Probing depth
Pg	*Porphyromonas gingivalis*
Pi	*Prevotella intermedia*
Pn	*Prevotella nigriscens*
PR	Prevalence ratio
Q-Q	Quantile-quantile
QTL	Quantitative trail locus

rho	Spearman’s rank correlation coefficient
SD	Standard deviation
SE	Standard error of the mean
SES	Socio-economic status
SNP	Single nucleotide polymorphism
Td	*Treponema denticola*
Tf	*Tannerella forsythia*
TNF	Tumor necrosis factor
US	United States
VDR	Vitamin D receptor
WHO	World Health Organization

CHAPTER 1

BACKGROUND

A. Introduction

Periodontitis, along with dental caries, is one of the two most common diseases of the oral cavity. It constitutes an inflammatory response to predominantly commensal oral bacteria. This relatively common chronic inflammatory disease is found in about 20% of the adult US population, is characterized by gingival pocket formation and clinical attachment loss, results in gradual destruction of periodontal tissues and tooth-supporting alveolar bone, and is considered the main cause of tooth loss among adults (1,2).

B. Definition and Pathophysiology of chronic periodontitis

1. Dental ecology- the host

The composition and complexity of the oral ecology has recently re-emerged as a focus of research, with the investigation of the oral microbiome's composition and significance gaining increased attention (3). It is commonly assumed that an oral ecosystem in harmonious symbiosis with its host will likely be associated with health (4). Departures from this equilibrium due to pathogenic ecological shifts, up-regulated host inflammatory/immune responses, or both, are characteristics of disease (5). In periodontitis, highly organized oral microbial biofilms (dental plaque) interact with a plethora of host-specific factors

(periodontal anatomy, oral hygiene, inflammatory response, and more) to result in tissue damage. Because tissue destruction appears to be a result of host responses to bacterial challenges rather than of the lytic effects of periodontal pathogens (6), the nature and extent of the microbiome-induced host response is considered an important determinant of the periodontitis phenotype. In fact, recent evidence indicates that autoimmune reactions may play role in the pathogenesis of the aggressive form of periodontitis (7)

2. Dental ecology- the oral microbiome

The "specific plaque" theory was articulated more than three decades ago, and emphasized the role of oral ecological shifts, essentially the bacterial insult in periodontitis (8, 9). In fact, Socransky summarized that in order for a pathogen to cause a disease the following conditions must be satisfied: 1) it must be a virulent type; 2) it must possess the chromosomal and extra-chromosomal genetic factors to initiate disease; 3) the host must be susceptible to the pathogen; 4) the pathogen must be in numbers sufficient to exceed the threshold for the host; 5) it must be located at the right place; 6) other bacterial species must foster, or at least not inhibit, the process; and 7) the local environment must be one which is conducive to the expression of the species' virulence properties (10).

The current state of knowledge with regard to the microbiology in periodontitis has not shifted from these core principles, as they were described in the 1970s and 1980s (11). The importance of the proliferating oral Gram-negative bacteria of the "orange" and the "red complex" is firmly established. Socransky in 1998 (12) defined these complexes as follows: "orange complex": Prevotella *intermedia* (*Pi*), Campylobacter *rectus* (*Cr*), Fusobacterium

nucleatum (*Fn*) and Prevotella *nigriscens* (*Pn*). The "red complex" includes Porphyromonas *gingivalis* (*Pg*), Tannerella *forsythia* (*Tf*) and Treponema *denticola* (*Td*) (Figure 1).

It is commonly agreed that periodontitis manifests as a response to predominantly commensal microbes, but some authors have argued that P*g* and *Aggregatibacter actinomycetemcomitans* (A*a*) which are important periodontal pathogens (13) should be considered exogenous (14, 15) although this is not well established in the literature. Although it is known that the aforementioned bacteria are organized in colonies or biofilms on the tooth surfaces, the dental plaque, the question of whether innate factors may facilitate or prevent such an infection and colonization is intriguing.

Recent findings have added to our understanding of how periodontal pathogens harbor the oral cavity from early age, symbiose and get organized, and interact among themselves and with the host (16, 17). For example, it has been shown that *Fn* possesses the capacity to adhere to and invade oral epithelial cells (18). T*f*, a "red complex" pathogen, has been shown to possess a glycosylated S-layer, which is responsible for evading the bacterium's recognition by the host immune system (19).

Several lines of research have used the presence and colonization level of periodontal pathogens as a "refined" exposure of periodontitis (20, 21). Many investigations have used serum antibody levels (IgG) to periodontal pathogens as a measure of exposure and/or systemic response (17, 22-25). Fewer studies have used direct quantification methods to assess the counts of specific pathogens in the subgingival biofilms (microbial plaque) (21, 26, 27). Interestingly, in a study of the association of periodontal conditions with carotid intima-media thickness (IMT), Desvarieux et al (21) reported that total periodontal microbial

counts and specifically periodontal pathogens were associated with IMT, whereas periodontitis (as classified by CDC) was not. Beck et al (28) suggested that particularly "red complex" bacteria may pose a potential risk for the development of systemic disorders. For example, a recent report suggested that P*g* and A*a* may be associated with the risk for pro-thrombotic state (29), while P*g* possesses properties known to be related to the pathogenesis of atherosclerosis (30). These investigations support the consideration of these specific microbial factors as a distinct exposure in investigations of periodontal, oral and systemic health, consistent with the paradigm of "periodontal medicine" (31, 32).

Interestingly, novel studies have provided insights on the induction by periodontopathogenic bacteria of epigenetic changes such as DNA hypermethylation (33). Epigenetics is one of the most rapidly expanding fields in biology, and the study of epigenetic mechanisms as well as the characterization of the human methylome alterations in health and disease, are priorities in biomedical research (34). Further research on the role of epigenetics in periodontal disease is needed.

3. Modifying factors

Despite important advances in our understanding of periodontal diseases the determinants of the inflammatory host response are not completely understood. While periodontal pathogens are necessary but not sufficient for disease activity to occur (10), several lines of research examine factors affecting or modulating the type and extent of inflammation evoked by the periodontal pathogens (35).

The fact that Diabetes Mellitus (DM) and severe periodontitis co-exist has been documented, and although a two-way relationship has been suggested (36, 37) causality has

not been demonstrated (38-40). While diabetic status modifies the host response to bacterial challenges and constitutes a risk factor for periodontitis, adequate glycemic control may be facilitated by controlling of an existing periodontal infection and inflammation (41). This has been demonstrated by a recent randomized controlled trial (42). It has been shown that although the periodontitis-DM relationship holds for both insulin-dependent and non-insulin dependent DM, it is more pronounced among patients with poorly controlled DM (43). Similar modifying effects have been shown for obesity (44).

Another example of host response-modification is thought to be the case of oral viral infections, which have been gaining attention with regard to CP (45-48). Noteworthy, Contreras et al. (49) showed that herpesviruses infections in combination with subgingival bacterial presence may exert periodontopathogenic potential by modifying the local host response.

Vitamin D has been reported as an important biological parameter with regard to the risk for periodontitis. Recently, Boggess et al. showed that the presence of moderate or severe periodontitis among pregnant women was associated with vitamin D insufficiency (50). Other lines of research indicate that innate antimicrobial factors such as lactoferrin, a salivary defense protein which is also secreted in gingival crevicular fluid during inflammation (51, 52), are important in the modulation of periodontitis (53).

Further elucidation of these biological interactions and pathway dissections can create new preventive and therapeutic opportunities for periodontitis. Moreover, mechanistic insights into the model of periodontitis may be key in understanding pathogenetic pathways relevant to other biologic conditions or chronic diseases, such as diabetes mellitus (DM) (37, 54), cardiovascular disease (CVD) (55, 56) and cancer (57-60).

C. Classification of periodontal diseases

The broad term of “periodontal disease” incorporates conditions other than periodontitis, such as gingivitis (61). The latter constitutes of an entirely reversible gingival inflammation (62) without periodontal destruction, and is typically diagnosed by bleeding upon periodontal probing (BOP). Although gingivitis had long been considered the entry-level stage of periodontitis and integral part of the disease continuum, this is not the case according to the current state of knowledge.

With regard to periodontitis, classification to two main types of chronic (CP) and aggressive periodontitis (AgP) was proposed in 1999 (61). The diagnosis between “aggressive” and “chronic” disease is only possible by knowledge of the rate of periodontal tissue destruction, or indirectly, by the assessment of tissue destruction level disproportional to a patient’s age. This type of sub-type classification may be problematic both in clinical practice (63) and surveillance or population-based studies. Although these two sub-types share several features and may appear close in pathogenetic features (64) distinct classification between AgP and CP is possible and feasible (65). For example, the consideration of an “aggressive” disease type, which likely corresponds to a more genetically penetrant or phenotypically expressed characteristic, may be worthwhile in genetic studies. On the other hand, it has been recognized that AgP cases may often represent manifestations of systemic conditions that interfere with resistance to bacterial infections (61). More recent evidence indicated the presence of auto-antibodies against come collagen types in patients with AgP but not CP (7). This phenomenon could pose a threat to validity due to disease misclassification in studies of periodontitis. At any rate, the classification of periodontitis by

sub-type has been a difficult exercise for clinicians and scientists over the last decades, and has been subject to numerous changes (65). Nonetheless, to aid in the design and interpretation of high-quality population-based studies of periodontitis, an accurate disease classification based on precise case definitions is warranted.

Chronic periodontitis (CP), previously known as “adult periodontitis” (61) is associated with irreversible periodontal tissue damage manifested by gradual periodontal attachment loss (Figure B). In order to establish the presence and quantify the extent and severity of periodontitis a number of clinical signs or measurements have been used: increased periodontal pocket depth (measured as probing depth - PD), loss of clinical attachment (CAL), alveolar bone loss (ABL) evidenced in radiographic examination, and bleeding upon probing (BOP) (65-68). Of those measurements, combined use of both CAL and PD has been recommended as the preferred approach in ascertaining the prevalence or progression of periodontitis in epidemiologic studies (66). Other attempts to quantify the extent and severity of the disease using the maximum amount of clinical information possible resulted in the introduction of the *extent* and *severity* scores, included in the Extent and Severity Index (ESI) by Carlos et al. in 1986 (69). Essentially, the ESI is represents the proportion of sites that exhibit disease expressed as a percent of pre-defined probing sites. Because extent and severity scores quantify the periodontal destruction as evidenced by the intraoral distribution of clinical measurements they can be regarded as an improvement over a categorical nosological model. However, these quantitative measures are also particularly vulnerable to bias due to tooth loss, even more so compared to more “crude” categorical disease models. For this reason, analytical approaches that consider extent scores typically include adjustments for tooth loss (70).

Because various researchers had not been using a universally accepted case definition of periodontitis, collaborative efforts in Europe and in the US were initiated to work towards agreement and convergence in the periodontitis case definition and disease progression standards. A European consensus was reached in the 5th European Workshop in Periodontology. This consensus statement was published in 2005 and recommended the use of a three-level (including health) disease case definition (71). Subsequently, in the US a joint workgroup that was initially commissioned in 2003 by the Centers for Disease Control and Prevention (CDC) and the American Academy of Periodontology (AAP) reviewed existing definitions and suggested another standard (66). Based on that work, an analogous three-level classification of periodontitis was proposed for use in population-based studies (Table 1.2). At the same time, Offenbacher (72) used a combination of clinical, microbial, inflammatory and host-response parameters to propose a five-level disease classification labeled BGI (Biofilm-Gingival Interface) that is defined by a combination of PD and BOP measurements. The BGI classification is considered a major improvement, because it reflects recent advances in our knowledge of the disease characteristics on both biological and clinical level taking into consideration the "biological systems model" (35).

For the purposes of the present investigation we will use the CDC definition and classification of CP. Based on the age structure of the studied population it is assumed that all periodontitis cases are representing chronic types and not AP. Continuous extent scores (e.g. percent of sites exhibiting ≥4mm CAL) will be considered for exploratory analyses.

The "CDC definition" of periodontitis uses the following taxonomy (Table 1.2):

1. **Severe periodontitis**: two or more interproximal sites not on the same tooth with ≥6 mm CAL and one or more interproximal sites with ≥5 mm PD;

2. **Initial** (moderate) **periodontitis**: two or more sites with 4 or 5mm CAL not on the same tooth;

3. **Health/Gingivitis**: individuals not meeting the above criteria.

D. Epidemiology of periodontal diseases

1. Prevalence and methodological issues

Most estimates of periodontitis indicate that advanced adult periodontitis does not exceed a 10-15% prevalence in most populations (1, 73). Estimates from the US are typically derived from the National Health and Nutrition Surveys. In the latest NHANES survey it was reported that 12.8% of the adult population had periodontitis, when the disease was defined as one or more sites with CAL≥ 5 mm. The prevalence was at a substantially higher 43.8% if the disease was defined as one or more sites with CAL≥ 3mm, and this reflects the dramatic impact that classification criteria can have on the disease prevalence estimates (66). Moreover, the NHANES employed a partial-mouth examination protocol, essentially selecting a subset of teeth, and then sites within teeth to examine. The partial-mouth examination protocol, while providing gains in efficiency (maximum examination sites were 42 sites per individual in the 2001-2004 NHANES versus a maximum of 168 theoretically at risk sites per person) was found to substantially underestimate the disease prevalence by as much as 60% relative to the true prevalence, as determined by a full mouth examination (74). The impact of tooth loss on periodontitis estimates is also an important methodological consideration, and was outlined in a previous section. A more detailed investigation and quantification of the impact of tooth loss in population-based studies of periodontitis is warranted.

2. Disease burden

Although the initially reported estimate of approximately 20% prevalence of CP in the adult population in the US is considerable, this figure is likely higher for the methodological reasons outlined above. Moreover, most countries in the world do not have adequate surveillance systems in place to monitor CP (1, 2). Another consideration is that while care-seeking individuals will be diagnosed and subsequently treated for the disease, others with sporadic pattern of or no dental attendance will not receive treatment and may progress to substantial tooth loss at early ages. This phenomenon will increase estimates of edentulism. Moreover, the increasing potential for dentition retention in older ages, combined with an ageing population is expected to contribute to a continuously increasing prevalence of CP in the next decades unless large scale awareness campaigns or interventions are effectively implemented in communities. This population approach is warranted because significant disparities in the prevalence of periodontitis exist between age and racial groups, with individuals of African American and Hispanic ethnicity being disproportionally affected (75).

A rapidly increasing number of reports associate CP with systemic conditions (cardiovascular disease, pregnancy outcomes, diabetes control, oxidative stress, etc), with the microbial load and the resultant inflammation being the likely mediators of these relationships (76-78). Although definitive evidence that links the treatment of CP with improved systemic health outcomes is limited, there are several plausible hypotheses and numerous research reports pointing towards this direction (79).

3. Impact on the quality of life

The impact of periodontitis on quality of life has been well-documented (80-83). Periodontitis is commonly associated with increased impacts on daily activities, even after adjustment for socio-demographic and other clinical conditions, such as caries, tooth loss and prostheses (81). Noteworthy, patients under treatment maintenance were found to report less daily impacts on their oral health-related quality of life (OHRQoL) compared to newly diagnosed patients with CP (83). Of course tooth loss and edentulism, representing the end-stage of CP are disabling conditions that are associated with severe and multi-level impacts (84, 85). Specifically in the US, reports based on NHANES data indicate that oral disease disproportionally impacts disadvantaged groups (86). Another recent report based on NHANES data concluded that periodontitis, and particularly the level of serum antibodies against *Pg*, was associated with cognitive impairment among older US adults (22).

E. Risk factors for chronic periodontitis

Although a plethora of risk factors has been reported to be associated with periodontitis, the importance of each factor remains unknown. In fact person-level factors, such as age and race have been shown to be important determinants of periodontal inflammation (35, 87). Oral hygiene has been the longest considered “traditional” risk factor for periodontitis (1, 88). In fact, while gingivitis parallels the level of oral hygiene on a population-level, this is not the case with CP (75).

Smoking has been found to be the major risk factor for both periodontitis and tooth loss (89-91). This is consistent with the well-documented biological effects of smoking on periodontal health (92-94). Noteworthy, a substantial proportion of over 40% of periodontitis

cases in the US are attributable to smoking based on NHANES III estimates (89); a finding with great public health relevance. Prospective studies have also confirmed the causative role of smoking in the progression of periodontitis, as evidenced by CAL (95). Evidence also exists that smoking modifies the sub-gingival microbial colonization pattern (96) with a mechanism that is consistent with the observation that among smokers there is "higher infectious burden with periodontal pathogens but less inflammation" (97).

With regard to DM, numerous studies have shown that diabetics have a greater prevalence and extent of periodontal pockets (37, 41, 98, 99). Interestingly, it has been reported that inflammatory markers that are evident in periodontitis, were associated with incident type 2 DM (100). These findings emphasize the links between oral and systemic conditions and the need for a comprehensive consideration of such co-morbidities in the diagnosis and treatment of both conditions.

To date, there is no consensus in the literature regarding the role of alcohol consumption as a risk factor for CP (101). A report from the Health Professionals Follow-up Study (HPSF) suggested a weak association (in the range of RR=1.2-1.3) but relied upon self-reports of periodontitis (102). Another investigation suggested that increased periodontitis risk may be conferred by alcohol consumption among individuals with certain polymorphisms of the $ALDH_2$ gene (103). Despite the documented multi-level detrimental effects of alcohol, it is well established that alcohol also exhibits antimicrobial effects that can confer protection from the microbial-plaque diseases. Evidence of such a protective effect has been found in both in vitro (104) and clinical (105) investigations.

Although some evidence suggests that dietary factors such as whole-grain and fiber intake may be associated with decreased risk of periodontitis (106) and nutrients such as

calcium (107) with decreased prevalence of the disease, most evidence indicates that impaired dental condition may be the cause of sub-optimal diet diversity and quality (108, 109).

Stress and other psychosocial factors such as coping skills and personality type have also been reported as risk factors for development or progression of periodontitis (110, 111). For example, in a study by Moss et al. (112), stress and depression traits were important risk factors for being a periodontitis case and manifesting disease progression, respectively.

F. The genetic basis of chronic periodontitis

1. Early reports of familial segregation

The genetic component of periodontitis was supported by early reports in the dental literature. These reports observed a familial aggregation of severe forms of the disease (113-117). Other reports, such as the one by Chung et al. (118) who carried out a path analysis using data from 241 nuclear families, did not detect any important evidence of heritability. These authors instead suggested that common family environment may be a major determinant of periodontal health. Michalowicz (119) later suggested that, although other periodontal diseases such as gingivitis have a substantial behavioral etiologic component, most of the observed heritability in periodontitis is attributable to genetics. The realization of the fact that “genes do not function in a vacuum” and that it is unlikely that a “master gene” for periodontitis will be discovered (120) underlines the necessity to investigate periodontitis risk loci within the full spectrum of common genetic variation.

2. Twin studies

An early report of a periodontitis study among twins was published by Ciancio (121) in 1969; two more reports were published by Michalowicz in the early 1990s (122, 123). In 1993 Corey et al. (124), published their investigation of self-reported periodontitis among twins. For their study they used a mailed survey among the members of a Virginia-based twin registry, and they reported concordance rates of 0.38 and 0.16 between monozygotic (MZ) and dizygotic (DZ) twins, respectively. The authors' conclusion was that, based on their findings, genetic factors make an important contribution to the risk of adult-onset periodontitis. Subsequently, Michalowicz et al. in 2000 (119) improved the current state of knowledge by employing clinical examinations to diagnose periodontitis among 117 MZ and DZ pairs of adult twins. In interpreting their findings the authors estimated that periodontitis has approximately 50% heritability and thus concluded "approximately half of the variance in disease in the population is attributed to genetic variance". A more recent but small investigation that was carried out among 10 pairs of MZ and 8 pairs of DZ twins (125) found a substantial amount of discordance in disease severity among the twins, with that discordance being greater among DZ pairs. The authors suggested that previous estimates of heritability may be exaggerated (125, 126).

3. Candidate-gene studies

A large number of publications have reported on the association of several gene polymorphisms with periodontitis. In these approaches the studied risk variants are thought to be associated with altered immune response to the bacterial insult and include, among other genes, the interleukin-1 (IL-1) (127), tumor necrosis factors (TNF) (128), matrix

metalloproteinases (129), estrogen (130) and vitamin D receptor (50). A summary of these findings is illustrated in Table 1.2. There have been very few candidate-gene investigations of CP among African Americans (131).

Previous candidate gene studies have found evidence for an association between periodontitis risk and interleukin (IL) gene polymorphisms (particularly IL-1), Fc*γ*, TNF, and matrix metalloproteinsases (128, 131, 132). A recent study reported an association between periodontitis and variants in *FAM5C* (133). Other candidate-gene studies did not detect any strong associations between AgP or CP and selected polymorphisms, and suggested the use of haplotype or genome-wide analyses as potentially more fruitful strategies (134). Moreover, the largest and best-powered candidate-gene study that was carried out in a Caucasian population (135) did not detect any important association between the IL-1 cluster and AgP risk. In spite of the high risk of bias from very small studies in the field, most reviews suggest that the balance of published evidence favors a causative role of IL gene cluster polymorphisms with periodontitis (127, 136). Two such meta-analyses summarized the evidence of increased risk for CP conferred by IL-1 (137) and Fc*γ* receptor polymorphisms (138). Nikolopoulos et al. summarized 53 published studies and concluded that there was a statistically significant association of IL-1A and IL-1B polymorphisms with CP risk. Dimou et al. suggested that there is accumulating evidence and supporting biological plausibility linking the FcγRIIIB NA1/NA2 polymorphism with increased risk for both AgP and CP (138). Finally, one recent investigation reported an association of IL-6 polymorphism with colonization with *Aa* but not *Pg*, among Caucasian CP patients (139).

4. Genome-wide association evidence

The opportunity to use whole-genome scans for the exploration of the association with various phenotypes was early recognized, in the era of family-based studies (140). Because current knowledge, technologies, and methods employed in GWAS allow a satisfactory coverage of the common genetic variation (at least among white populations), this approach has gained great popularity (141, 142) for the study of common complex diseases. Although the initial impact of GWAS has been less dramatic than initially postulated (141, 143), they have provided an unprecedented amount of new information regarding the susceptibility for and the pathogenesis of many diseases.

Schaefer et al. (144) were the first to report a genome-wide association hit with periodontal disease. Specifically, these investigators also reported the existence of a shared risk susceptibility locus on 9p21.3 (rs1333048), for both AgP and coronary heart disease (145). This region was mapped to the sequence of the large antisense noncoding RNA *ANRIL*, which partly overlaps regulatory and coding sequences of CDKN2A/CDKN2B. This association was recently replicated in an independent case-control sample with 130 cases of AgP and 339 controls (146), where the authors suggested that the most plausible genetic model underlying the association between the identified SNPs and AgP is the multiplicative one. Further, the same group of investigators in another publication reported that an intronic SNP rs1537415 located in the glycosyltransferase gene GLT6D1 was associated with AgP. In that study, the rare G allele showed 10% enrichment in cases (total ~280 cases). The authors suggested that the rare allele was associated with reduction of the binding affinity of the zinc-finder transcription factor GATA-3, which could be important in the pathogenesis of AgP.

While these GWA investigations provide the first insights into the genome-wide evidence of risk loci for periodontitis, their small sample sizes limited their statistical power. Moreover, both studies examined cases of AgP which represents the most aggressive form of periodontitis, and is found in about 3% or less among various populations (147, 148).

A recent review of published articles up to April 2009 on genetic polymorphisms associated with CP was published by Laine in 2010 (131). The authors noted that for their search strategy they used the keywords: Periodontitis, Periodontal disease in combination with the words: Genes, Mutation, or Polymorphism to identify articles written in the English language, employing a case-control design to study CP or adult periodontitis and reporting genotype distribution. The authors however did not provide an exact search string that could be replicated, information about which databases were searched and with what limits, the exact number of abstract and full-text articles identified from their searches, additional information for seeking additional articles, exclusions including removing duplicates, etc.

G. Tables

Table 1.1 Clinical case definitions of Periodontitis proposed by the CDC working group for use in population-based surveillance of Periodontitis[1] (2)

	Clinical Definition		
Disease Category	Clinical Attachment Level (CAL)		Probing Depth (PD)
Severe periodontitis	≥2 interproximal sites[2] with CAL ≥ **6mm**	and	≥1 interproximal site with PD ≥5mm
Moderate periodontitis	≥2 interproximal sites[2] with CAL ≥ **4mm**	or	≥2 interproximal sites[2] with PD ≥5mm
No or mild periodontitis	Neither "moderate" nor "severe" periodontitis		

1: third molars are excluded

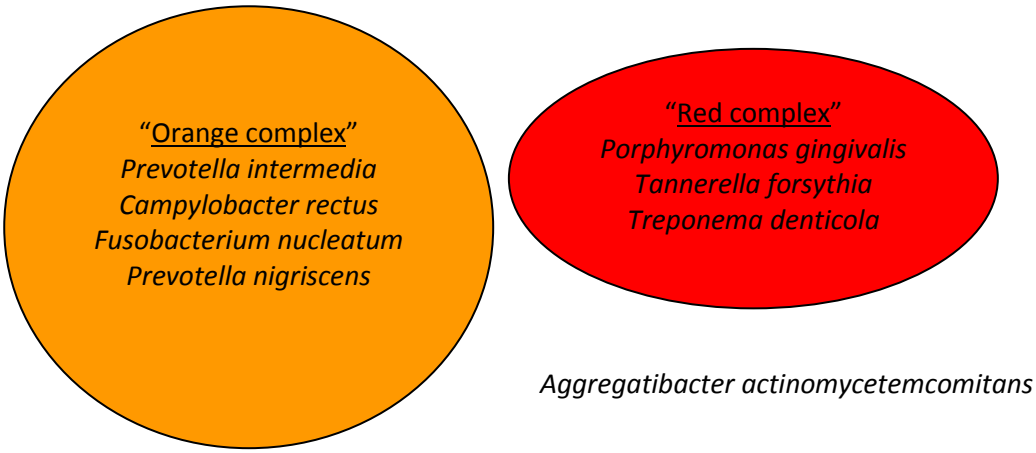
2: not on the same tooth

Table 1.2. Overview of published genetic effects from candidate-gene case control studies of the risk of chronic periodontitis among Caucasians.

Gene	Polymorphism	Race	Number studies found: Association with CP	No association with CP
IL1A	IL1A -889 (+4845) C>T	Caucasian	1	11
IL1B	IL1B +3954 (+3953) C>T	Caucasian	4	12
IL1RN	IL1RN VNTR (+2018) C>T	Caucasian	1	3
TFNA	TFNA (-367) G>A	Caucasian	1	5
IL6	IL6 (-572) C>G	Caucasian	1	2
	IL6 (-1363) G>T	Caucasian	1	1
IL10	IL10 -819 (-824) C>T	Mixed (~80% Caucasians)	2	4
	IL10 -592 (-597) C>A	Mixed & Caucasians	4	4
FcγRIIα	IIα 131 H>R	Caucasian	1 (among smokers only)	4
FcγRIIIb	IIIb NA1> NA2	Japanese	1	4
VDR	VDR Taq1 T>C	Caucasian	2	2
CD14	CD14 -260 C>T	Caucasian	2	6
TLR4	TLR4 Thr399Ile	Japanese	1	5

Figures

Figure 1. Classification of microbial periodontal pathogens according to Socransky 1998 (12)



H. References

1. Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. Bull World Health Organ 2005;83:661-9.

2. Petersen PE, Kandelman D, Arpin S, Ogawa H. Global oral health of older people--call for public health action. Community Dent Health 2010;27(4 Suppl 2):257-67.

3. Avila M, Ojcius DM, Yilmaz O. The oral microbiota: living with a permanent guest. DNA Cell Biol 2009;28:405-11.

4. Handelsman J. Microbial symbiosis: in sickness and in health. DNA Cell Biol 2009;28:359-60.

5. Darveau RP. The oral microbial consortium's interaction with the periodontal innate defense system. DNA Cell Biol 2009;28:389-95.

6. Armitage GC, Robertson PB. The biology, prevention, diagnosis and treatment of periodontal diseases: scientific advances in the United States. J Am Dent Assoc 2009;140 Suppl 1:36S-43S.

7. Hendler A, Mulli TK, Hughes FJ, et al. Involvement of autoimmunity in the pathogenesis of aggressive periodontitis. J Dent Res 2010;89:1389-94.

8. Loesche WJ. Clinical and microbiological aspects of chemotherapeutic agents used according to the specific plaque hypothesis. J Dent Res 1979;58:2404-12.

9. Socransky SS. Relationship of bacteria to the etiology of periodontal disease. J Dent Res 1970;49:203-22.

10. Socransky SS, Haffajee AD. The bacterial etiology of destructive periodontal disease: current concepts. J Periodontol 1992;63(4 Suppl):322-31.

11. Loesche WJ, Grossman NS. Periodontal disease as a specific, albeit chronic, infection: diagnosis and treatment. Clin Microbiol Rev 2001;14:727-52.

12. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL Jr. Microbial complexes in subgingival plaque. J Clin Periodontol 1998;25:134-44.

13. Haffajee AD, Socransky SS. Microbial etiological agents of destructive periodontal diseases. Periodontol 2000 1994;5:78-111.

14. Van Winkelhoff AJ, Winkel EG. Microbiological diagnostics in periodontics: biological significance and clinical validity. Periodontol 2000 2005;39:40-52.

15. Slots J, Ting M. Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis in human periodontal disease: occurrence and treatment. Periodontol 2000 1999;20:82-121.

16. Nakano K, Miyamoto E, Tamura K, et al. Distribution of 10 periodontal bacterial species in children and adolescents over a 7-year period. Oral Dis 2008;14:658-64.

17. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

18. Han YW, Shi W, Huang GT, et al. Interactions between periodontal bacteria and human oral epithelial cells: Fusobacterium nucleatum adheres to and invades epithelial cells. Infect Immun 2000;68:3140-6.

19. Sekot G, Posch G, Messner P, et al. Potential of the Tannerella forsythia S-layer to delay the immune response. J Dent Res 2011;90:109-14.

20. Kebschull M, Demmer RT, Papapanou PN. "Gum bug, leave my heart alone!"--epidemiologic and mechanistic evidence linking periodontal infections and atherosclerosis. J Dent Res 2010;89:879-902.

21. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

22. Noble JM, Borrell LN, Papapanou PN, Elkind MS, Scarmeas N, Wright CB. Periodontitis is associated with cognitive impairment among older adults: analysis of NHANES-III. J Neurol Neurosurg Psychiatry 2009;80:1206-11.

23. Beck JD, Eke P, Heiss G, et al. Periodontal disease and coronary heart disease: a reappraisal of the exposure. Circulation 2005;112:19-24.

24. Madianos PN, Lieff S, Murtha AP, et al. Maternal periodontitis and prematurity. Part II: Maternal infection and fetal exposure. Ann Periodontol 2001;6:175-82.

25. Baker PJ, Wilson ME. Opsonic IgG antibody against Actinobacillus actinomycetemcomitans in localized juvenile periodontitis. Oral Microbiol Immunol 1989;4:98-105.

26. Ximénez-Fyvie LA, Haffajee AD, Socransky SS. Comparison of the microbiota of supra- and subgingival plaque in health and periodontitis. J Clin Periodontol 2000;27:648-57.

27. Haffajee AD, Cugini MA, Dibart S, Smith C, Kent RL Jr, Socransky SS. The effect of SRP on the clinical and microbiological parameters of periodontal diseases. J Clin Periodontol. 1997;24:324-34.

28. Beck JD, Slade G, Offenbacher S. Oral disease, cardiovascular disease and systemic inflammation. Periodontol 2000 2000;23:110-20.

29. Bizzarro S, Nicu EA, van der Velden U, Laine ML, Loos BG. Association of serum immunoglobulin G (IgG) levels against two periodontal pathogens and prothrombotic state: a clinical pilot study. Thromb J 2010;8:16.

30. Nakano K, Wada K, Nomura R, et al. Characterization of aortic aneurysms in cardiovascular disease patients harboring Porphyromonas gingivalis. Oral Dis 2011;17:370-8.

31. Pizzo G, Guiglia R, Lo Russo L, Campisi G. Dentistry and internal medicine: from the focal infection theory to the periodontal medicine concept. Eur J Intern Med 2010;21:496-502.

32. Stamm JW. Periodontal diseases and human health: new directions in periodontal medicine. Ann Periodontol 1998;3:1-2.

33. Bobetsis YA, Barros SP, Lin DM, et al. Bacterial infection promotes DNA hypermethylation. J Dent Res 2007;86:169-74.

34. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol 2010;28:1057-68.

35. Offenbacher S, Barros SP, Beck JD. Rethinking periodontal inflammation. J Periodontol 2008;79(8 Suppl):1577-84.

36. Iacopino AM. Periodontitis and diabetes interrelationships: role of inflammation. Ann Periodontol 2001;6:125-37.

37. Grossi SG, Genco RJ. Periodontal disease and diabetes mellitus: a two-way relationship. Ann Periodontol 1998;3:51-61.

38. Bascones-Martinez A, Matesanz-Perez P, Escribano-Bermejo M, González-Moles MÁ, Bascones-Ilundain J, Meurman JH. Periodontal disease and diabetes-Review of the Literature. Med Oral Patol Oral Cir Bucal 2011;16:e722-9.

39. Ide R, Hoshuyama T, Wilson D, Takahashi K, Higashi T. Periodontal disease and incident diabetes: a seven-year study. J Dent Res 2011;90:41-6.

40. Gurav A, Jadhav V. Periodontitis and risk of diabetes mellitus. J Diabetes 2011;3:21-8.

41. Mealey BL, Oates TW; American Academy of Periodontology. Diabetes mellitus and periodontal diseases. J Periodontol 2006;77:1289-303.

42. Koromantzos PA, Makrilakis K, Dereka X, Katsilambros N, Vrotsos IA, Madianos PN. A randomized, controlled trial on the effect of non-surgical periodontal therapy in patients with type 2 diabetes. Part I: effect on periodontal status and glycaemic control. J Clin Periodontol 2011;38:142-7.

43. Soskolne WA. Epidemiological and clinical aspects of periodontal diseases in diabetics. Ann Periodontol 1998;3:3-12.

44. Borrell LN, Papapanou PN. Analytical epidemiology of periodontitis. J Clin Periodontol 2005;32 Suppl 6:132-58.

45. Pyrc K, Strzyz P, Milewska A, Golda A, Schildgen O, Potempa J. Porphyromonas gingivalis enzymes enhance infection with human metapneumovirus in vitro. J Gen Virol 2011;92(Pt 10):2324-32.

46. Slots J. Herpesvirus periodontitis: infection beyond biofilm. J Calif Dent Assoc 2011;39:393-9.

47. Thomasini RL, Bonon SH, Durante P, Costa SC. Correlation of cytomegalovirus and human herpesvirus 7 with CD3(+) and CD3(+) CD4(+) cells in chronic periodontitis patients. J Periodontal Res 2011 doi: 10.1111/j.1600-0765.2011.01413.x.

48. Slots J. Human viruses in periodontitis. Periodontol 2000 2010;53:89-110.

49. Contreras A, Umeda M, Chen C, Bakker I, Morrison JL, Slots J. Relationship between herpesviruses and adult periodontitis and periodontopathic bacteria. J Periodontol 1999;70:478-84.

50. Boggess KA, Espinola JA, Moss K, Beck J, Offenbacher S, Camargo CA Jr. Vitamin D status and periodontal disease among pregnant women. J Periodontol 2011;82:195-200.

51. Eberhard J, Drosos Z, Tiemann M, Jepsen S, Schröder JM. Immunolocalization of lactoferrin in healthy and inflamed gingival tissues. J Periodontol 2006;77:472-8.

52. Tenovuo J. Clinical applications of antimicrobial host proteins lactoperoxidase, lysozyme and lactoferrin in xerostomia: efficacy and safety. Oral Dis 2002;8:23-9.

53. Wakabayashi H, Kondo I, Kobayashi T, et al. Periodontitis, periodontopathic bacteria and lactoferrin. Biometals 2010;23:419-24.

54. Löe H. Periodontal disease. The sixth complication of diabetes mellitus. Diabetes Care 1993;16:329-34.

55. Beck JD, Offenbacher S. Systemic effects of periodontitis: epidemiology of periodontal disease and cardiovascular disease. J Periodontol 2005;76(11 Suppl):2089-100.

56. Beck J, Garcia R, Heiss G, Vokonas PS, Offenbacher S. Periodontal disease and cardiovascular disease. J Periodontol 1996;67(10 Suppl):1123-37.

57. Divaris K, Olshan AF, Smith J, et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. Cancer Causes Control 2010;21:567-75.

58. Meyer MS, Joshipura K, Giovannucci E, Michaud DS. A review of the relationship between tooth loss, periodontal disease, and cancer. Cancer Causes Control 2008;19:895-907.

59. Michaud DS, Liu Y, Meyer M, Giovannucci E, Joshipura K. Periodontal disease, tooth loss, and cancer risk in male health professionals: a prospective cohort study. Lancet Oncol 2008;9:550-8.

60. Michaud DS, Joshipura K, Giovannucci E, Fuchs CS. A prospective study of periodontal disease and pancreatic cancer in US male health professionals. J Natl Cancer Inst 2007;99:171-5.

61. Armitage GC. Development of a classification system for periodontal diseases and conditions. Ann Periodontol 1999;4:1-6.

62. Offenbacher S, Barros S, Mendoza L, et al. Changes in gingival crevicular fluid inflammatory mediator levels during the induction and resolution of experimental gingivitis in humans. J Clin Periodontol 2010;37:324-33.

63. Picolos DK, Lerche-Sehm J, Abron A, Fine JB, Papapanou PN. Infection patterns in chronic and aggressive periodontitis. J Clin Periodontol 2005;32:1055-61.

64. Armitage GC, Cullinan MP. Comparison of the clinical features of chronic and aggressive periodontitis. Periodontol 2000 2010;53:12-27.

65. Armitage GC. Classifying periodontal diseases--a long-standing dilemma. Periodontol 2000 2002;30:9-23.

66. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78(7 Suppl):1387-99.

67. Armitage GC. Periodontal diagnoses and classification of periodontal diseases. Periodontol 2000 2004;34:9-21.

68. Williams RC. Periodontal disease. N Engl J Med 1990;322:373-82.

69. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.

70. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

71. Tonetti MS, Claffey N; European Workshop in Periodontology group C. Advances in the progression of periodontitis and proposal of definitions of a periodontitis case and disease progression for use in risk factor research. Group C consensus report of the 5th European Workshop in Periodontology. J Clin Periodontol 2005;32 Suppl 6:210-3.

72. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

73. Papapanou PN. Epidemiology of periodontal diseases: an update. J Int Acad Periodontol 1999;1:110-6.

74. Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA. Accuracy of NHANES Periodontal Examination Protocols. J Dent Res 2010;89:1208-13.

75. Albandar JM, Rams TE. Global epidemiology of periodontal diseases: an overview. Periodontol 2000 2002;29:7-10.

76. Hickman MA, Boggess KA, Moss KL, Beck JD, Offenbacher S. Maternal periodontal disease is associated with oxidative stress during pregnancy. Am J Perinatol 2011;28:247-52.

77. Inaba H, Amano A. Roles of oral bacteria in cardiovascular diseases—from molecular mechanisms to clinical cases: Implication of periodontal diseases in development of systemic diseases. J Pharmacol Sci 2010;113:103-9.

78. Noack B, Genco RJ, Trevisan M, Grossi S, Zambon JJ, De Nardin E. Periodontal infections contribute to elevated systemic C-reactive protein level. J Periodontol 2001;72:1221-7.

79. Scannapieco FA, Dasanayake AP, Chhun N. "Does periodontal therapy reduce the risk for systemic diseases?". Dent Clin North Am 2010;54:163-81.

80. Araújo AC, Gusmão ES, Batista JE, Cimões R. Impact of periodontal disease on quality of life. Quintessence Int 2010;41:e111-8.

81. Bernabé E, Marcenes W. Periodontal disease and quality of life in British adults. J Clin Periodontol 2010;37:968-72.

82. Ng SK, Leung WK. Oral health-related quality of life and periodontal status. Community Dent Oral Epidemiol 2006;34:114-22.

83. Needleman I, McGrath C, Floyd P, Biddle A. Impact of oral health on the life quality of periodontal patients. J Clin Periodontol 2004;31:454-7.

84. Divaris K, Ntounis A, Marinis A, Polyzois G, Polychronopoulou A. Loss of natural dentition: multi-level effects among a geriatric population. Gerodontology 2010 doi: 10.1111/j.1741-2358.2010.00440.x.

85. Steele JG, Sanders AE, Slade GD, et al. How do age and tooth loss affect oral health impacts and quality of life? A study comparing two national samples. Community Dent Oral Epidemiol 2004;32:107-14.

86. Sanders AE, Slade GD, Lim S, Reisine ST. Impact of oral disease on quality of life in the US and Australian populations. Community Dent Oral Epidemiol 2009;37:171-81.

87. Salvi GE, Lawrence HP, Offenbacher S, Beck JD. Influence of risk factors on the pathogenesis of periodontitis. Periodontol 2000 1997;14:173-201.

88. Axelsson P, Lindhe J. Effect of controlled oral hygiene procedures on caries and periodontal disease in adults. J Clin Periodontol 1978;5:133-51.

89. Tomar SL, Asma S. Smoking-attributable periodontitis in the United States: findings from NHANES III. National Health and Nutrition Examination Survey. J Periodontol 2000;71:743-51.

90. Ong G. Periodontal disease and tooth loss. Int Dent J 1998;48(3 Suppl 1):233-8.

91. Holm G. Smoking as an additional risk for tooth loss. J Periodontol 1994;65:996-1001.

92. Van Dyke TE, Sheilesh D. Risk factors for periodontitis. J Int Acad Periodontol 2005;7:3-7.

93. Bergström J. Tobacco smoking and chronic destructive periodontal disease. Odontology 2004;92:1-8.

94. Haber J, Wattles J, Crowley M, Mandell R, Joshipura K, Kent RL. Evidence for cigarette smoking as a major risk factor for periodontitis. J Periodontol 1993;64:16-23.

95. Schatzle M, Loe H, Ramseier CA, et al. Clinical course of chronic periodontitis: effect of lifelong light smoking (20 years) on loss of attachment and teeth. Journal of Investigative and Clinical Dentistry 2010;1:8-15.

96. Haffajee AD, Socransky SS. Relationship of cigarette smoking to the subgingival microbiota. J Clin Periodontol 2001;28:377-88.

97. Buduneli N, Larsson L, Biyikoglu B, Renaud DE, Bagaitkar J, Scott DA. Fatty acid profiles in smokers with chronic periodontitis. J Dent Res 2011;90:47-52.

98. Oliver RC, Tervonen T. Periodontitis and tooth loss: comparing diabetics with the general population. J Am Dent Assoc 1993;124:71-6.

99. Tervonen T, Oliver RC. Long-term control of diabetes mellitus and periodontitis. J Clin Periodontol 1993;20:431-5.

100. Pradhan AD, Manson JE, Rifai N, Buring JE, Ridker PM. C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. JAMA 2001;286:327-34.

101. Dye BA. The Relationship Between Periodontitis and Alcohol use is not Clear. J Evid Based Dent Pract 2010;10:225-7.

102. Pitiphat W, Merchant AT, Rimm EB, Joshipura KJ. Alcohol consumption increases periodontitis risk. J Dent Res 2003;82:509-13.

103. Nishida N, Tanaka M, Sekine S, et al. Association of ALDH2 genotypes with periodontitis progression. J Dent Res 2010;89:138-42.

104. Herrera D, Roldán S, Santacruz I, Santos S, Masdevall M, Sanz M. Differences in antimicrobial activity of four commercial 0.12% chlorhexidine mouthrinse formulations: an in vitro contact test and salivary bacterial counts study. J Clin Periodontol 2003;30:307-14.

105. Kongstad J, Hvidtfeldt UA, Grønbaek M, Jontell M, Stoltze K, Holmstrup P. Amount and type of alcohol and periodontitis in the Copenhagen City Heart Study. J Clin Periodontol 2008;35:1032-9.

106. Merchant AT, Pitiphat W, Franz M, Joshipura KJ. Whole-grain and fiber intakes and periodontitis risk in men. Am J Clin Nutr 2006;83:1395-400.

107. Al-Zahrani MS. Increased intake of dairy products is related to lower periodontitis prevalence. J Periodontol 2006;77:289-94.

108. Nowjack-Raymer RE, Sheiham A. Association of edentulism and diet and nutrition in US adults. J Dent Res 2003;82:123-6.

109. Hutton B, Feine J, Morais J. Is there an association between edentulism and nutritional state? J Can Dent Assoc 2002;68:182-7.

110. Genco RJ, Ho AW, Grossi SG, Dunford RG, Tedesco LA. Relationship of stress, distress and inadequate coping behaviors to periodontal disease. J Periodontol 1999;70:711-23.

111. Genco RJ, Ho AW, Kopman J, Grossi SG, Dunford RG, Tedesco LA. Models to evaluate the role of stress in periodontal disease. Ann Periodontol 1998;3:288-302.

112. Moss ME, Beck JD, Kaplan BH, et al. Exploratory case-control analysis of psychosocial factors and adult periodontitis. J Periodontol 1996;67:1060-9.

113. Beaty TH, Boughman JA, Yang P, Astemborski JA, Suzuki JB. Genetic analysis of juvenile periodontitis in families ascertained through an affected proband. Am J Hum Genet 1987;40:443-52.

114. Long JC, Nance WE, Waring P, Burmeister JA, Ranney RR. Early onset periodontitis: a comparison and evaluation of two proposed modes of inheritance. Genet Epidemiol 1987;4:13-24.

115. Page RC, Vandesteen GE, Ebersole JL, Williams BL, Dixon IL, Altman LC. Clinical and laboratory studies of a family with a high prevalence of juvenile periodontitis. J Periodontol 1985;56:602-10.

116. Saxén L, Nevanlinna HR. Autosomal recessive inheritance of juvenile periodontitis: test of a hypothesis. Clin Genet 1984;25:332-5.

117. Saxén L. Heredity of juvenile periodontitis. J Clin Periodontol 1980;7:276-88.

118. Chung CS, Kau MC, Chung SS, Rao DC. A genetic and epidemiologic study of periodontal disease in Hawaii. II. Genetic and environmental influence. Am J Hum Genet 1977;29:76-82.

119. Michalowicz BS, Diehl SR, Gunsolley JC, et al. Evidence of a substantial genetic basis for risk of adult periodontitis. J Periodontol 2000;71:1699-707.

120. Nares S. The genetic relationship to periodontal disease. Periodontol 2000 2003;32:36-49.

121. Ciancio SG, Hazen SP, Cunat JJ. Periodontal observations in twins. J Periodontal Res 1969;4:42-5.

122. Michalowicz BS, Aeppli D, Virag JG, et al. Periodontal findings in adult twins. J Periodontol 1991;62:293-9.

123. Michalowicz BS. Genetic and inheritance considerations in periodontal disease. Curr Opin Periodontol 1993:11-7.

124. Corey LA, Nance WE, Hofstede P, Schenkein HA. Self-reported periodontal disease in a Virginia twin population. J Periodontol 1993;64:1205-8.

125. Torres de Heens GL, Loos BG, van der Velden U. Monozygotic twins are discordant for chronic periodontitis: clinical and bacteriological findings. J Clin Periodontol 2010;37:120-8.

126. Stabholz A, Soskolne WA, Shapira L. Genetic and environmental risk factors for chronic periodontitis and aggressive periodontitis. Periodontol 2000 2010;53:138-53.

127. Grigoriadou ME, Koutayas SO, Madianos PN, Strub JR. Interleukin-1 as a genetic marker for periodontitis: review of the literature. Quintessence Int 2010 ;41:517-25.

128. Soga Y, Nishimura F, Ohyama H, Maeda H, Takashiba S, Murayama Y. Tumor necrosis factor-alpha gene (TNF-alpha) -1031/-863, -857 single-nucleotide polymorphisms (SNPs) are associated with severe adult periodontitis in Japanese. J Clin Periodontol 2003;30:524-31.

129. Repeke CE, Trombone AP, Ferreira SB Jr, et al. Strong and persistent microbial and inflammatory stimuli overcome the genetic predisposition to higher matrix metalloproteinase-1 (MMP-1) expression: a mechanistic explanation for the lack of association of MMP1-1607 single-nucleotide polymorphism genotypes with MMP-1 expression in chronic periodontitis lesions. J Clin Periodontol 2009;36:726-38.

130. Zhang X, Long Y, Li HY, Li XJ, Ding Y. [Estrogen receptors gene polymorphisms in patients with chronic periodontitis]. Hua Xi Kou Qiang Yi Xue Za Zhi 2010;28:139-44.

131. Laine ML, Loos BG, Crielaard W. Gene polymorphisms in chronic periodontitis. Int J Dent 2010;2010:324719.

132. Anovazzi G, Kim YJ, Viana AC, et al. Polymorphisms and haplotypes in the interleukin-4 gene are associated with chronic periodontitis in a Brazilian population. J Periodontol 2010;81:392-402.

133. Carvalho FM, Tinoco EM, Deeley K, et al. FAM5C contributes to aggressive periodontitis. PLoS One 2010;5(4):e10053.

134. Wohlfahrt JC, Wu T, Hodges JS, Hinrichs JE, Michalowicz BS. No association between selected candidate gene polymorphisms and severe chronic periodontitis. J Periodontol 2006;77:426-36.

135. Fiebig A, Jepsen S, Loos BG, et al. Polymorphisms in the interleukin-1 (IL1) gene cluster are not associated with aggressive periodontitis in a large Caucasian population. Genomics 2008;92:309-15.

136. Abu-Saleh T. Interleukin-1 composite polymorphism as a risk factor for chronic periodontitis: a review. SADJ 2010;65:160-2, 164-6.

137. Nikolopoulos GK, Dimou NL, Hamodrakas SJ, Bagos PG. Cytokine gene polymorphisms in periodontal disease: a meta-analysis of 53 studies including 4178 cases and 4590 controls. J Clin Periodontol 2008;35:754-67.

138. Dimou NL, Nikolopoulos GK, Hamodrakas SJ, Bagos PG. Fcgamma receptor polymorphisms and their association with periodontal disease: a meta-analysis. J Clin Periodontol 2010;37:255-65.

139. Nibali L, Tonetti MS, Ready D, et al. Interleukin-6 polymorphisms are associated with pathogenic bacteria in subjects with periodontitis. J Periodontol 2008;79:677-83.

140. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516-7.

141. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.

142. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. Science 2002;298:2345-9.

143. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

144. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19:553-62.

145. Schaefer AS, Richter GM, Groessner-Schreiber B, et al. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet 2009;5:e1000378.

146. Ernst FD, Uhr K, Teumer A, et al. Replication of the association of chromosomal region 9p21.3 with generalized aggressive periodontitis (gAgP) using an independent case-control cohort. BMC Med Genet 2010;11:119.

147. Oliver RC, Brown LJ, Löe H. Periodontal diseases in the United States population. J Periodontol 1998;69:269-78.

148. Löe H, Brown LJ. Early onset periodontitis in the United States of America. J Periodontol 1991;62:608-16.

CHAPTER 2

SPECIFIC RESEARCH AIMS

A. Rationale - the case for a GWAS of chronic periodontitis

Exploring and unraveling the genetic basis of complex diseases has promise in aiding therapy, prevention and care for such diseases, and can ultimately lead to improved health outcomes (1, 2). The opportunity to use whole-genome scans for the exploration of the association with various phenotypes was early recognized, in the era of family-based studies (3). Because current knowledge, technologies, and methods employed in GWAS allow a satisfactory coverage of the common genetic variation among whites (4), this approach has gained great popularity for the study of common complex diseases (5, 6). Although the impact of GWAS has been less dramatic than initially postulated (5, 7), they have generated an unprecedented amount of new information regarding the susceptibility for and the pathogenesis of many common complex diseases. Notably, a large number of nucleotide polymorphisms (SNPs) were found to be associated with a total of 165 traits and were reported in 951 published GWA studies through June 2011 (8).

In the field of dentistry, thus far there have been limited applications of GWAS. Three small GWA investigations have reported genome-wide associations of loci with aggressive periodontitis (9-11) and two have reported on suggestive risk loci for dental caries (12, 13). These investigations were hampered by small sample size and other methodological issues such as clinical examination procedures, disease definition, and more. We propose to

add on to the existing knowledge basis and improve on these previous investigations by carrying out a GWA analysis of chronic periodontitis in the context of a well-defined cohort, using a moderately-sized sample of approximately 4600 white individuals for whom detailed socio-demographic, anthropometric, periodontal, medical, and behavior data have been collected (**Specific Aim 1**). As a supplement to this aim, we will conduct exploratory analyses of genome-wide [gene] x [environment] interactions, by considering SNP interactions with sex, smoking, and diabetes mellitus.

With regard to microbial factors, the pathogenetic role of specific bacterial strains in periodontitis is well-established and has been characterized in detail. While there is a great diversity in the periodontal microbiome, few microorganisms have been directly implicated in the pathogenesis of periodontitis (14). It is well established that colonization with high amounts of "red complex" bacteria (*Pi, Pg, Tf*), as well as *Aa*, is more common in severe or aggressive forms of periodontitis. P*i* causes periodontal tissue destruction by triggering the host immune response, inducing Prostaglandin E2 and increasing the expression of matrix metalloproteinases (15). P*g* and T*f* have been shown to be more common among cases with "refractory" periodontitis versus "treatment-responsive" or periodontally-healthy individuals (16). Tf, in particular, was recently shown to possess an external S-layer that likely is responsible for the attenuated host response to this pathogen (17). Moreover, periodontal pathogens have been associated with systemic morbidities such as CHD, and they have been used as a "refined" exposure in studies of oral-systemic health links over clinical classifications of periodontitis (18, 19). Thus, preventing or controlling oral ecological shifts towards pathogenic biofilms (proliferation of the "orange" or "red" complex) is critical. Moreover, it is intriguing to determine why some individuals harbor more or more

pathogenic periodontal bacteria, and to this end detect whether innate host factors are important. In this domain, one report of a positive association between an IL-6 polymorphism and harboring of *Aa* was based on a small sample of forty Caucasian patients (20). We will address this question by exploring whether there are genome-wide loci that are associated with increased susceptibility for colonization with CP-risk associated bacteria of the "red" and the "orange" complex, as well as *Aa* (**Specific Aim 2**). As an additional exploratory aim, due to the major role of *Pg* in the pathogenesis of CP (21), we will investigate risk loci for *Pg* colonization.

B. SPECIFIC AIMS

By performing the proposed genome-wide association analysis among the white ARIC study participants we aim:

Specific Aim 1: To identify susceptibility loci for chronic periodontitis among a moderate-sized sample of community-dwelling white adults using the CDC disease classification and a continuous measure of disease severity.

As an addition to Specific Aim 1, we will explore for effect measure modification or gene x environment interactions of SNPs with sex, smoking and diabetes mellitus.

Specific Aim 2: To identify susceptibility loci for colonization with periodontal microorganisms of the "orange" and "red" complex, and *Aggregatibacter actinomycetemcominats*.

As an additional exploratory step for Specific Aim 2 we will examine for risk loci for another individual pathogen, *Porphyromonas gingivalis* which is member of the “red” complex.

REFERENCES

1. Lander ES, Linton LM, Birren B, et al.; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860-921.

2. Lander ES. The new genomics: global views of biology. Science 1996;274:536-9.

3. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516-7.

4. Barrett JC, Cardon LR. Evaluating coverage of genome-wide association studies. Nat Genet 2006;38:659-62.

5. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.

6. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. Science 2002;298:2345-9.

7. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

8. Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies. Accessed [November 7 2011]

9. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19:553-62.

10. Schaefer AS, Richter GM, Groessner-Schreiber B, et al Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet 2009;5:e1000378.

11. Ernst FD, Uhr K, Teumer A, et al. Replication of the association of chromosomal region 9p21.3 with generalized aggressive periodontitis (gAgP) using an independent case-control cohort. BMC Med Genet 2010;11:119.

12. Vieira AR, Marazita ML, Goldstein-McHenry T. Genome-wide scan finds suggestive caries loci. J Dent Res 2008;87:435-9.

13. Shaffer JR, Wang X, Feingold E, et al. Genome-wide association Scan for childhood Caries Implicates Novel Genes. J Dent Res 2011 In Press

14. Teles RP, Haffajee AD, Socransky SS. Microbiological goals of periodontal therapy. Periodontol 2000 2006;42:180-218.

15. Eley BM, Cox SW. Proteolytic and hydrolytic enzymes from putative periodontal pathogens: characterization, molecular genetics, effects on host defenses and tissues and detection in gingival crevice fluid. Periodontol 2000 2003;31:105-24.

16. Colombo AP, Boches SK, Cotton SL, et al. Comparisons of subgingival microbial profiles of refractory periodontitis, severe periodontitis, and periodontal health using the human oral microbe identification microarray. J Periodontol 2009 ;80:1421-32.

17. Buduneli N, Larsson L, Biyikoglu B, Renaud DE, Bagaitkar J, Scott DA. Fatty Acid Profiles in Smokers with Chronic Periodontitis. J Dent Res 2011;90:47-52.

18. Kebschull M, Demmer RT, Papapanou PN. "Gum bug, leave my heart alone!"--epidemiologic and mechanistic evidence linking periodontal infections and atherosclerosis. J Dent Res 2010;89:879-902.

19. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

20. Nibali L, Donos N, Farrell S, et al. Association Between IL6 -174 Polymorphism and A.actinomycetemcomitans in Chronic Periodontitis. J Periodontol 2010;81:1814-9.

21. Reddi D, Brown SJ, Belibasakis GN. Porphyromonas gingivalis induces RANKL in bone marrow stromal cells: Involvement of the p38 MAPK. Microb Pathog 2011;51:415-20.

CHAPTER 3

RESEARCH METHODS

A. The Atherosclerosis Risk in Communities (ARIC) study

1. The ARIC study

The ARIC is a National Heart, Lung, and Blood Institute (NHLB)-funded prospective epidemiologic study conducted in four U.S. communities (Jackson, Mississippi; Washington County, Maryland; suburban Minneapolis, Minneapolis; Forsyth County, North Carolina). The study was IRB-approved for every participating site, and all participants had given informed consent. ARIC was designed to investigate the etiology and natural history of atherosclerosis, the etiology of clinical atherosclerotic diseases, and variation in cardiovascular risk factors, medical care and disease by race, gender, location, and date (1). Within the context of ARIC, an NIDCR-funded ancillary dental study was carried out, the Dental ARIC. The Dental ARIC aims were to determine the prevalence, extent, and severity of periodontal conditions in the dentate ARIC population, and to describe the associations between those conditions and prevalent coronary heart disease, coronary artery IMT, presence of carotid artery lesions, and atherosclerosis risk factors (2).

2. Study population

The ARIC Cohort Component began in 1987 and recruited a total of 15,792 community-dwelling participants aged 45-64 sampled from a defined population in their communities (1, 3). These participants were reexamined every three years with the first screen (baseline) occurring in 1987-89, the second in 1990-92, the third in 1993-95, and the fourth and last exam was in 1996-98. The Dental ARIC ancillary study took place during the fourth visit, between 1996 and 1998 among dentate ARIC subjects.

From the initial recruitment of 15,792 participants 11,656 were seen in visit 4. Of those, 6797 underwent the periodontal examination. After exclusions for various reasons a final sample of 6017 subjects formed the Dental ARIC sample (2). There participants had mean age of 62 years (range 52-75), 21% had severe periodontitis when using the study protocol criteria (≥30% of sites with AL ≥3mm), 13% were diabetics and 44% were males.

3. Dental data

The Dental ARIC included a clinical oral examination, collection of gingival crevicular fluid, oral microbial plaque, and serum, as well as in-person interviews. Clinical measurements of PD and CAL were made on six sites on all teeth, and the number of missing teeth was recorded. For these measurements, very good accuracy and reproducibility is to be expected with trained and calibrated examiners (4-6), as was the case in ARIC. Studies have estimates that agreement between examiners for PD and CAL are within 1mm in 90% of the measurements taken (4, 7). In the Dental ARIC weighted kappa statistics ranged between 0.76-0.86 indicating excellent agreement with a standard examiner and intraclass correlation

coefficients (ICC) ranged between 0.76-0.90 indicating excellent to outstanding agreement (2).

Using the PD and CAL measurements and based the CDC/AAP criteria (Table 1.1), subjects were classified as having no disease, moderate or severe CP. Additionally, a continuous measure of disease severity or "extent" based on Carlos et al. (8) was defined as the proportion of sites exhibiting CAL ≥3mm. The distribution characteristics of these variables in the analytical sample overall and across strata of covariates are presented in Tables 3-7. Additional dental (time of and reason for last dental visit, tooth brushing frequency, oral hygiene practices) and anthropometric data (body mass index) were used for exploratory descriptive and bivariate analyses, and are presented in the Appendix. Details about the analytical strategy followed to analyze the CDC and the "extent" traits are presented as part of the analytical approach for Specific Aim 1.

4. Microbiological data

Microbiological data are available for a subset of approximately 1200 white participants (9). Samples of subgingival microbial plaque were obtained during the clinical examination. Subsequently, the levels of eight periodontal pathogens were determined by DNA-DNA checkerboard hybridization analysis, a technique initially described by Gunaratnam et al. (10) and Socransky et al. (11). The selection of these eight microbes was based on their reported and plausible implication in the pathogenesis of periodontitis (10-12). The method has a lowest detection threshold of 10^3-10^4 microbial counts, and has been shown to have good detection properties (10). Therefore, subgingival microbial counts were obtained for organisms of the "red" complex: *Pg*, *Tf*, and *Td*, the "orange" complex: P*i*, P*n*, C*r*, F*n*, as

well as *Aa*. Additional information about the analytical strategy with regard to the microbiological data is presented as part of the approach for Specific Aim 2.

5. Covariates

The ARIC study investigators have collected a comprehensive list of socio-demographic, behavior, anthropometric and biological measurements (clinical, laboratory, and other specialized tests). An extensive list of variables collected can be found on the ARIC website (http://www.cscc.unc.edu/aric/) and in previous publications (1). For the purposes of the GWA analyses we used the following ARIC study variables: examination center (modeled with two indicator variables), age (measured in years and modeled as a continuous variable) and sex (male or female). For exploratory "sensitivity" analyses we also used smoking (modeled as a three-level ordinal categorical variable were 0: never, 1: former, and 2: current smoker), and diabetic status (modeled as a dichotomous variable where 0: healthy and 1: fasting glucose levels of ≥126 mg/dL, non-fasting of ≥200 mg/dL, or pharmacological treatment for diabetes). To correct for population stratification (admixture) we included in all analyses ten principal components derived by the EIGENSTRAT method (see section 3.B.2).

B. Genotyping, quality control and exclusions

1. Genotyping and imputation

In the study population, DNA was extracted from blood samples drawn from an antecubital vein into tubes containing serum separator gel. Blood samples were analyzed at a central ARIC laboratory in Houston, TX. Genotyping was performed with the Affymetrix Genome-Wide Human SNP Array 6.0 chip. The platform offers 906,600 markers for SNPs.

The rigorous quality control procedures included initial blind duplicate genotyping and identification/flagging of SNPs with kappa<0.95 and reconciliation of unintentional duplicate samples (17 duplicates and one triplicate). Imputation to 2.5million markers was performed using 669,450 SNPs and MACH v1.0.16 (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html), based on HapMap Phase II CEU build 36. The SNPs used for imputation were selected from 839,048 autosomal SNPs restricted to those with minor allele frequency (MAF)>0.01 (129,543 excluded), Hardy-Weinberg equilibrium (HWE) $P>10^{-5}$ (12,432 excluded) and call rate >95% (1,693 excluded). We used the following SNP exclusion criteria for further analyses: quality score <0.8 and missing data rate >10% after imputation, and MAF of <5%.

2. Population stratification

To obtain estimates of relatedness and population stratification a subset of 85,947 "high quality" linkage-disequilibrium (LD)-pruned SNPs was selected. These SNPs met the following criteria among self-reported whites: MAF ≥0.1, call rate >99.5%, HWE $P\geq10^{-3}$, autosomal, with annotation in the platform annotation file, not labeled "AFFX" or "chromosome 0", and not monomorphic. Using these SNPs identity-by-state (IBS) allele sharing distance (DST values) were computed using PLINK, as such: DST = IBS distance $(IBS_2 + 0.5*IBS_1)$ / (n SNP pairs). First degree relative status was assigned to pairs of individuals with DST≥0.8 and second degree relatives were considered those with 0.763 ≤DST< 0.8. Among the white ARIC participants there were 380 pairs of first degree and 207 pairs of second degree relatives identified. To minimize exclusions, related pairs were broken by iterative selection of individuals with most relatives using a custom program.

Population stratification was further evaluated with principal component (PC) analysis using the EIGENSTRAT approach (13) and the EIGENSOFT program (14). The above chosen set of LD-pruned SNPs was used for the computation of ten principal components. Genetic outliers were considered those that were further than 8 standard deviations (SD) away from any of the ten PCs over ten runs of PC computation. Based on DST and PC criteria there were 716 subjects flagged from removal from the analysis (206 as genetic outliers based on PCs and 16 based on average DST values ("too little IBS sharing" with the rest of the sample), 351 first degree relatives and 143 second degree relatives. All but ten second degree relatives (whose relatives were excluded as genetic outliers) were re-entered in the dataset and were assigned PCs. After exclusion of 364 individuals (4%) there were 9349 whites who were included in the GWA analysis and of those, 4610 had periodontal phenotype data available as Dental ARIC participants.

C. Analytical strategy

Goldstein (15), Manolio (16), McCarthy (17) and Hirschorn (18) have summarized the key challenges in the analysis and interpretation of genome-wide data. We acknowledge the inherent limitations of GWAS, such as the low power to detect very small effects, and the consideration of common genetic variance for the detection of "single-polymorphism effects". The problem of the "little variance explained" by the discovered SNPs for most diseases has led investigators to look for the "dark matter" of genetic effects in the so-called rare (<5% or <1%) variants (19); the study of the latter however, may be methodologically challenging unless whole-genome typing becomes more accessible, or when multi-stage (20) or novel, more powerful statistical methods (21) are implemented. However, GWAS are a

powerful discovery tool that has the potential to unveil previously unknown genetic risk loci and provide insights on novel mechanisms and pathways. Acknowledging the strengths and limitations of the GWA approach which simultaneously interrogates millions of SNPs across the genome, we embarked on this investigation with a "risk locus discovery" rather than an "effect estimation" approach.

1. Analytical sample

Primary analyses were conducted among the Dental ARIC white subjects with imputed data ('freeze 3' version of genomic dataset). After exclusions specified above (364 whites) and merging of the genetic and clinical datasets, our final analytical sample included 4610 white individuals for whom clinical and genomic data were successfully matched. The analytical sample description for Specific Aim 1 is presented in Tables 3-7, and for Aim 2 in Tables 4-11. The distribution of the phenotype classification (CP status: CDC definition and extent of attachment loss: percent of sites with attachment loss of 3mm or greater) by the study covariates is also presented in Table 3. Participants' mean age is 62 years (range 52-75), 13% are diabetics and 44% are males. When Beck et al. used the CDC classification of periodontitis in a subset of ~5000 Dental ARIC participants, 42% were periodontally healthy, 41% had initial periodontitis and 17% had 'severe' periodontitis (22). In our analytical sample, these proportions are virtually identical, 43 and 17% respectively.

2. Analytical strategy for Specific aim 1

The primary phenotype of interest to address our Specific aim 1 was chronic periodontitis as defined by CDC, a three-level categorical classification. The secondary phenotype was the

"extent" of attachment loss, a continuous measure that is expressed as the proportion of site with attachment loss of 3mm or greater. The following traits were defined and considered for the GWA analyses:

1. Binary outcome: severe periodontitis (coded as 1) versus initial periodontitis or health (coded as 0). Statistical analysis will be based on a logistic regression model assuming multiplicative (log-additive) allelic effects.

2. Binary outcome: moderate periodontitis (coded as 1) versus initial periodontitis or health (coded as 0). As above, statistical analysis will be based on a logistic regression model assuming multiplicative (log-additive) allelic effects.

3. Continuous outcome: "extent" or "severity" of periodontitis (proportion of sites with CAL≥3mm). An appropriate transformation of this measure (z-score, normal curve equivalent transformation of the "crude" proportion of sites) was performed prior to conducting the GWAS. Statistical analysis was based on a linear regression model assuming multiplicative (log-additive) allelic effects.

"Minimally" adjusted models

As noted in the general methods section, all models were adjusted for age, sex and population substructure/stratification. Adjustment for population substructure in the ARIC study has been based on principal component analysis. Although authors have argued that population stratification may not be a serious threat to the validity of GWAS results (23), adjustment for population ancestry has become standard practice (13). Therefore, the "minimal" genetic models include 10 principal components obtained by the statistical package EIGENSOFT (13, 14), as well as for sex and age. Additional terms adjusting for

examination center (two indicator variables) were entered in all models. For implementation in the context of the GWAS we used the probABEL package, which is part of the geneABEL suite (24).

Interpretation of analysis results relied on the evidence that the data offer against the null hypothesis of no SNP (allelic) association with the traits of interest, and therefore was based on the associated P-values. When multiple SNPs emerged below the P-value threshold for prioritization (see following section on hypothesis testing), we presented the "top SNP" per locus, which was the one with the lowest P-value. The additional prioritized SNPs in the same locus were also presented, along with linkage disequilibrium metrics (R^2 with the "top SNP" in the locus) obtained with the SNAP application (25). Moreover, effect estimates (odds ratios for the disease classification and betas for the "extent" trait), as well as 'model-predicted' phenotypic estimates by genotype were also obtained and presented for the "top SNPs" in each locus.

Hypothesis testing

To determine genome-wide "significance" of the tested SNPs' association with the examined phenotypes we considered a multiple-test correction. The Bonferroni has been most frequently used in published GWAS, it is commonly agreed however, that this approach is overly conservative. This issue is further discussed in the limitations part of the discussion section. After applying the multiple-test correction, assuming approximately a million independent tests, a genome-wide significance threshold of $P<5x10^{-8}$ was set. We set another, less stringent threshold of $P<5x10^{-6}$ for prioritizing SNPs for further investigation and locus exploration.

The FDR, as described by Storey and Tibshirani (26), an alternative method proposed to evaluate results from GWAS. In the FDR approach, a *q* value (instead of a *p*) is calculated as a measure of significance in terms of false discovery rate, versus false positive (in the *p* value setting). The interpretation of a q value can be thought as the proportion of significant findings (genome-wide "hits") that turn out to be false positives, after replication/verification (26). Along these lines, Shi et al recently reported a simulation study based on a two-stage FDR approach and use of a least absolute shrinkage and selection operator (LASSO) regression to reduce false positives (27).

Sensitivity analysis - exploratory "fully" adjusted models

Using the accepted definition of a confounder (28), no confounding of the association SNP→phenotype is to be expected by "environmental" or "behavioral" variables. This is supported by the fact that, under standard assumptions, these covariates do not affect the distribution of SNPs in the source population. However, because factors such as smoking and diabetes are risk factors for CP, one may argue that GWAS-identified signals for CP could be in fact, signals marking risk loci for these "intermediate" characteristics. In that scenario, adjusting for smoking and diabetes would "sanitize" the GWAS results from these results. On the other hand, in cases of pleiotropy (implying a true common genetic cause or risk locus of two traits) this result would reflect a valid association. Investigators have proposed methods of adjustment for correlated phenotypes (29), but "non-adjusted" analyses remain the current standard of practice to-date. Another view in favor of "adjusted" genetic modeling supports that the variance reduction that results from such adjustments is favorable

in the GWA "discovery" attempt, because there remains "less variance to be explained" by the genetic effects.

Motivated by the above, we undertook a "sensitivity" analysis as an exploratory step that is presented in the Appendix of Chapter 4. In these supplemental analyses we employed smoking and diabetes-adjusted genetic models for the three traits examined in the first manuscript (severe and moderate CP, and "extent" of attachment loss). We compared 'crude' and 'adjusted' effect estimates for the prioritized SNPs, using an empirical 10% change-in-estimate criterion for 'notable' effect estimate changes, similar to settings of confounding evaluation (30). Moreover, we inspected whether additional SNPs emerged below the threshold of genome-wide significance ($P<5x10^{-8}$) or prioritization ($P<5x10^{-6}$), due to the "available variance reduction" effect that was explained above.

Exploratory assessment of effect measure modification- "gene x environment" interactions

Genome-wide explorations of gene x environment interactions are not common. As noted in the general methods section, our study of approximately 4000 subjects was underpowered to detect interaction effects. However, we conducted exploratory evaluations of EMM by three variables: sex, smoking and diabetes. The rationale for the selection of these three factors is supported by the fact that they represent important and well-documented risk factors for CP: the disease is substantially more common among males, smokers and diabetics. This is evident both in the literature, as well as our Dental ARIC study population. We used the "extent" of attachment loss trait for these exploratory

analyses, because it represents a cumulative and non-reversible measure of periodontal destruction among the teeth present in the oral cavity.

EMM was evaluated on the multiplicative (log-additive) scale in the context of statistical interaction between SNPs and sex (binary variable), smoking (three-level variable) and diabetes (binary variable). Each interaction term was entered in one of three linear regression models that apart from age, sex, examination center and 10 PCs, included the risk factor main effect (i.e. smoking) and its interaction with the SNP (SNP*smoking). We conducted these analyses with the probABEL package, and considered evidence of EMM interaction term P-values of less than 10^{-5}. This relatively high P-value threshold was chosen because interaction evaluations have traditionally low power. An alternative approach of EMM evaluation could have been stratified analyses (by sex, smoking and diabetes) followed by a between-strata homogeneity evaluation, which would also be subject to reduced power. We chose the statistical interaction approach over stratified analyses as a more efficient strategy.

Annotation and visualization of GWA results

Genome-wide significant and 'prioritized' SNPs (based on the P-value criteria that were set in the 'hypothesis testing' section) were annotated using the WGAViewer (31) and Snipper (http://csg.sph.umich.edu/boehnke/snipper/) programs. We explored and reported SNPs locations and their role (i.e. intronic or intergenic, representing a base pair change or a synonymous change, and more), as well as their physical distance (in Kb) from the two or three closest known genes. Additionally, in the Appendix we present associations of the prioritized SNPs with known expression quantitative loci (eQTL) (32), as well as their

interacting genes. This expression-association information was obtained from the Scan (33) database (http://www.scandb.org/newinterface/about.html) via the Snipper annotator. We report eQTL-gene associations found in lymphoblastoid cell lines (LCL) among central European (CEU) population samples and $P \leq 10^{-4}$.

Quantile-quantile (Q-Q) plots were generated to compare the calculated versus expected empirical distributions of the test statistics (-log10 P-values) and detect any substantial deviations that could indicate residual population stratification. For this purpose we evaluated the Lambda (λ) inflation coefficients' proximity to 1. Manhattan plots were also used to display the summary of the genome-wide analysis results (-log10 P-values) by chromosomal location. The Matlab® program was used to display Q-Q plots, λ coefficients and Manhattan plots. To inspect genomic regions that appear associated with the examined phenotypes, we used LocusZoom® version 1.1 (34) and Haploview v.4.2 (35). With these applications we were able to plot selected SNPs on their physical chromosome locations, along with their corresponding –log10 P-values, nearby gene locations, linkage disequilibrium (LD) and recombination rate statistics based on HapMap II-CEU. For missense changes, prediction of the possible impact of amino acid substitutions on protein structure and function was assessed using the PolyPhen-2 application (36). We used the USCF/Encode genome browser (http://genome-preview.ucsc.edu/cgi-bin/hgTracks?db=hg19) to visualize these prioritized loci for genomic areas that may include functionally relevant features, such as open chromatin and DNase I hypersensitivity annotation. Additionally, crude visualizations of functional pathways or interaction networks that identified genes may be implicated were generated with the GeneMANIA application (37). The complete set of

software that we used for annotation, visualization and additional exploration of identified SNPs and loci is presented in Table 3.1.

Power calculations

Several programs are available for power calculations in GWA studies (38). Because of its flexibility in enabling power calculations in the presence of G*E (gene by exposure) interactions, we used the QUANTO v. 1.2.4 program (http://hydra.usc.edu/GxE/).

Typical effect sizes of new SNPs detected by GWAS have been reported to range between 1.1 and 1.3 (15). Using our sample size of ~4,000 individuals and an outcome prevalence of 17 or 60% (depending on the definition) for Specific Aim 1, we had 80% power to detect effects of odds ratio size ≥1.4 for minor allele frequencies (MAF) greater than 10%. As expected, we were less powered to detect interaction effects in the exploratory analyses, unless one assumes relative interaction effects of ≥1.8 magnitude. Illustrations of various iterations of effect sizes, outcome prevalence values, MAF and power are presented in the Appendix.

3. Analytical strategy for Specific aim 2

Counts of the eight periodontal pathogens were derived with a semi-quantitative method, "checkerboard" DNA-DNA hybridization (10, 11). This method was applied on plaque samples to measure the extent of sub-gingival colonization with: *Prevotella intermedia* [American Type Culture Collection (ATCC) 25611], *Campylobacter rectus* (ATCC 33238), *Fusobacterium nucleatum* (ATCC 10953), and *Prevotella nigriscens* (ATCC 33563) (belonging to the "orange" complex); *Porphyromonas gingivalis* (ATCC 33277), *Tannerella*

forsythia (ATCC 43037) and *Treponema denticola* (ATCC 35404) ("red" complex); and *Aggregatibacter actinomycetemcomitans* (ATCC 43718). In this method bacterial levels are expressed as counts relative to established microbial standards.

Three dichotomous traits of "high" colonization with "red" and "orange" complex, and *Aggregatibacter actinomycetemcomitans* were considered for analytical purposes. The two composite phenotypes were derived by the summation of bacterial count for each species belonging to the "red" (n=3) and "orange" (n=4) group, as described above. Because *Porphyromonas gingivalis* is considered the major periodontal pathogen implicated in periodontitis in adult populations, we explored for additional risk loci using its "high" colonization phenotype defined as above, as a separate trait, in analyses that we present in the Appendix.

Various approaches in defining the bacterial colonization profiles have been previously used, including summations of the absolute microbial counts (9), tertile-categorization (39), and five-level categorization of log10-transformed counts (40). To approach Specific Aim 2, we defined a "high" colonization trait as the top quintile (20%) of each trait's distribution. The rationale for the selection of this phenotype is based on the fact that the "checkerboard" semi-quantitative method has a lower detection threshold of ~10^3-10^4 and reduced precision in the lower end of the distribution. For this reason, individuals with "high" bacterial colonization profile may be those with reduced or impaired host and at high risk for periodontal tissue destruction, thus the "high colonization cases" are a trait of interest. Additionally, selection of a smaller subset of individuals (i.e. top 5 or 10%) would reduce our already low power for the GWA analysis among the ~1000 subjects.

Genetic models, annotation and visualization of results, power calculations

Three logistic regression models assuming multiplicative (log-additive) genetic effects were employed to address Specific aim 2, one for each trait of interest: "high red" complex, "high orange" complex, and "high" *Aa* colonization. We used identical model specifications as in Aim 1, and included age, sex, examination center and 10 PCs as covariates. We followed identical post-analysis annotation and visualization procedures. The power analyses are presented in the Appendix. As noted in the Specific Aims section, we developed an additional fourth model to examine for risk loci for *Pg* as an exploratory step. Results of this analysis are presented in the Supplemental material of Chapter 5.

C. Replication plan for GWA findings

Replication of identified SNPs from any GWA study to other samples is a warranted validation step (17, 41). Some investigators have suggested that because replication efforts may fail to confirm even true associations when power considerations exist there is no reason for investigators to place unreasonably high expectations on replication studies (42). In spite of this, the field appears settled into routinely requesting replication of GWA findings.

Candidate replication datasets for the present GWA analysis will ideally include detailed ascertainment of chronic periodontitis based on the CDC taxonomy (or alternatively extent of attachment loss) using full-mouth six-sites per tooth periodontal examinations of an adequately sized sample of white subjects. The investigators' team is actively seeking opportunities for collaboration and eventually replication of our findings. Replication of the microbiological findings (Specific Aim 2) appears less likely, but opportunities may arise in the future, as more investigators explore the host-oral microbiome interactions. Candidate

studies where a replication of our main findings could be performed if genotype data become available include the National Health And Nutrition Examination Survey (NHANES), the Health ABC study, and the Health Professionals Follow-up Study (HPFS). Meta-analysis of replicated results will be performed using the METAL program (43).

D. TABLE

Table 3.1. Software applications used for the genome-wide association analysis, annotation and visualization of results.

Program name (version)	Functions	Website	Developer, year and citation
GeneMANIA	Gene network discovery and visualization	http://www.genemania.org/	Mostafavi, 2008
Haploview (ver.4.2)	Linkage disequilibrium (LD) and haplotype visualization and analysis	http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview	Barrett, 2005
LocusZoom (ver.1.1)	Regional association results plotting	http://csg.sph.umich.edu/locuszoom/	Pruim, 2010
METAL	Meta-analysis of GWAS results	http://genome.sph.umich.edu/wiki/METAL_Program	Willer, 2010
Polyphen (ver.2.1.0)	Prediction of damaging missense mutations	http://genetics.bwh.harvard.edu/pph2/index.shtml	Adzhubei, 2010
PLINK	Genome-wide association analysis package	http://pngu.mgh.harvard.edu/~purcell/plink/	Purcel, 2007
ProbABEL (ver.1.0.3)	Genome-wide association analysis package for imputed data	http://www.genabel.org/packages/ProbABEL	Aulchenko, 2010
QUANTO (ver.1.2.4)	Power and sample size calculation in the presence of gene x environment interactions	http://hydra.usc.edu/GxE/	Gauderman, 2006
SNAP (ver.2.2)	Identification of proxy SNPs and generation of LD plots	http://www.broadinstitute.org/mpg/snap/index.php	Johnson, 2008
Snipper (ver.1.2)	SNPs annotation including nearby genes and expression quantitative loci	http://csg.sph.umich.edu/boehnke/snipper/	Welch, 2010
WGAViewer (ver.1.26l)	GWAS result annotation package	http://people.chgv.lsrc.duke.edu/~dg48/WGAViewer/std.php	Ge, 2008

E. REFERENCES

1. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

2. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

3. White AD, Folsom AR, Chambless LE, et al. Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. J Clin Epidemiol 1996;49:223-33.

4. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78:1387-99.

5. Mombelli A. Clinical parameters: biological validity and clinical utility. Periodontol 2000 2005;39:30-9.

6. Alves Rde V, Machion L, Andia DC, Casati MZ, Sallum AW, Sallum EA. Reproducibility of clinical attachment level and probing depth of a manual probe and a computerized electronic probe. J Int Acad Periodontol 2005;7:27-30.

7. Armitage GC. Manual periodontal probing in supportive periodontal treatment. Periodontol 2000 1996;12:33-9.

8. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.

9. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007 ;78:1911-25.

10. Gunaratnam M, Smith GL, Socransky SS, Smith CM, Haffajee AD. Enumeration of subgingival species on primary isolation plates using colony lifts. Oral Microbiol Immunol 1992;7:14-8.

11. Socransky SS, Smith C, Martin L, Paster BJ, Dewhirst FE, Levin AE. "Checkerboard" DNA-DNA hybridization. Biotechniques 1994;17:788-92.

12. Van Winkelhoff AJ, Winkel EG. Microbiological diagnostics in periodontics: biological significance and clinical validity. Periodontol 2000 2005;39:40-52.

13. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904-9.

14. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet 2006;2:e190.

15. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.

16. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.

17. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

18. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005;6:95-108.

19. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. Am J Hum Genet 2010;86:6-22.

20. Feng T, Zhu X. Genome-wide searching of rare genetic variants in WTCCC data. Hum Genet 2010;128:269-80.

21. Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet 2010;11:773-85.

22. Beck JD, Eke P, Heiss G, et al. Periodontal disease and coronary heart disease: a reappraisal of the exposure. Circulation 2005;112:19-24.

23. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. Cancer Epidemiol Biomarkers Prev 2002;11:513-20.

24. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

25. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24:2938-9.

26. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440-5.

27. Shi G, Boerwinkle E, Morrison AC, Gu CC, Chakravarti A, Rao DC. Mining gold dust under the genome wide significance level: a two-stage approach to analysis of GWAS. Genet Epidemiol 2011;35:111-8.

28. Greenland S, Rothman KJ. Introduction to stratified analysis. In: Rothman KJ, Greenland S, Lash TL. Modern epidemiology. New York: Lippincott, Williams and Wilkins, 2008:258-82.

29. Vansteelandt S, Goetgeluk S, Lutz S, et al. On the adjustment for covariates in genetic association analysis: a novel, simple principle to infer direct causal effects. Genet Epidemiol 2009;33:394-405.

30. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. Am J Epidemiol 1993;138:923-36.

31. Ge D, Zhang K, Need AC, et al. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640-643.

32. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010;6:e1000888.

33. Gamazon ER, Zhang W, Konkashbaev A, et al. SCAN: SNP and copy number annotation. Bioinformatics 2010;26:259-62.

34. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336-7.

35. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

36. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9.

37. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol 2008;9 Suppl 1:S4.

38. Knight J. A survey of current software for genetic power calculations. Hum Genomics 2004;1:225-7.

39. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

40. Papapanou PN, Neiderud AM, Papadimitriou A, Sandros J, Dahlén G. "Checkerboard" assessments of periodontal microbiota and serum antibody responses: a case-control study. J Periodontol 2000;71:885-97.

41. Ioannidis JP. Non-replication and inconsistency in the genome-wide association setting. Hum Hered 2007;64:203-13.

42. Liu YJ, Papasian CJ, Liu JF, Hamilton J, Deng HW. Is replication the gold standard for validating genome-wide association findings? PLoS One 2008;3:e4037.

43. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-1.

CHAPTER 4

A. TITLE AND AUTHORS

MANUSCRIPT #1: EXPLORING THE GENETIC BASIS OF CHRONIC PERIODONTITIS: A GENOME-WIDE APPROACH

K. Divaris, K.L. Monda, K.E. North, A.F. Olshan, E.M. Lange, K. Moss, S.P. Barros, J.D. Beck, and S. Offenbacher

B. MANUSCRIPT #1 SUMMARY

Chronic periodontitis (CP) is a common-complex oral disease that affects the majority of the adult population and is a major cause of tooth loss. Although a substantial genetic component of CP is theorized, to-date, no whole-genome association (GWA) analyses have been performed. We conducted a GWA analysis among 4610 white participants of the Atherosclerosis in Communities Study (mean age of 62 years). Traits of interest were the three-level CDC/AAP periodontitis classification (severe—17%, moderate—43%, healthy—40%), and the continuous trait of "extent" of attachment loss (EAL; proportion of sites exhibiting ≥3mm attachment loss). Genotyping was performed with the Affymetrix 6.0 platform and imputation to 2.5million markers was based on HapMap II-CEU. We used logistic genetic models for the examination of the "severe" and "moderate" CP, and a linear model for the EAL trait using a genome-wide significance threshold of $P<5x10^{-8}$. No genome-wide significant association signals were noted. However, we found suggestive evidence of association ($P<5x10^{-6}$) for seven loci including *NIN, NPY, WNT5A* for severe, *NCR2, EMR1* for moderate, and *TBX18, ETS1, DYNC2H1, TTC26* and *ZC3HAV1* for EAL. These genome-wide association results from a large well-defined cohort provide information on multiple candidate regions for interrogation in future genetic studies of CP.

Keywords: periodontitis, genome-wide association studies; oral health; dentistry; genetics;

C. INTRODUCTION

Chronic periodontitis (CP) is a common-complex disease of the oral cavity that is characterized by an inflammatory response to commensal and pathogenic oral bacteria (1). This relatively common chronic inflammatory disease is found in about 20% of the adult US population, manifests with gingival pocket formation and clinical attachment loss (CAL), and results in gradual destruction of periodontal tissues and tooth-supporting alveolar bone. CP is considered the main cause of tooth loss among most adult populations worldwide (2). Moreover, a growing body of evidence has linked the disease with increased risk for systemic conditions including coronary heart disease (CHD) (3), pregnancy outcomes (4), poor diabetes control (5), and other conditions.

There are more than 450 species identified in the human microbiome (6), and although harboring of periodontal pathogens is virtually universal, only a small proportion of individuals develop the severe form of the disease. Risk factors for CP have been well-studied and include smoking and diabetes mellitus (DM). In addition, age, race and obesity have also been shown to be important risk indicators (7). A genetic component of CP risk was supported by early reports of familial aggregation of severe forms of the disease (8), as well as twin studies (9), but the magnitude of risk conferred by genetics and the role of specific genes has been under debate.

Recent candidate-gene studies for CP have focused on genes related to host immunity and inflammatory response, such as cytokines, cell-surface receptors, chemokines, enzymes and antigen recognition. Most of these studies have examined polymorphisms in the interleukin (IL)-1, IL-6, Fc gamma receptor (FcγR), tumor necrosis factor alpha (TNFα),

human vitamin D receptor (VDR), cluster of differentiation (CD)-14, matrix metalloproteinase (MMP)-1, toll-like receptor (TLR), cyclo-oxygenase-2 (COX-2), and C-reactive protein (CRP) gene coding regions (10).

A recent genome-wide association (GWA) study of generalized aggressive periodontitis (gAgP) among a sample of whites of European descent identified associations with a susceptibility locus on 9p34.3 intronic to the glycosyltransferase 6 domain containing 1 (*GLT6D1*) gene, as well as a shared susceptibility locus on 9p21.3 for both gAgP and CHD (11, 12). However, gAgP is a rare form of periodontitis, found in less than 1% of adults, and is a distinct entity from CP. To-date no GWA exploration has been performed for CP. To add to the knowledge base of the genetic etiology of CP, this study aims to investigate genetic risk loci for CP using a GWA approach in the context of a well-defined cohort.

D. MATERIALS AND METHODS

We conducted a GWA study among 4610 white participants of the Atherosclerosis Risk In Communities (ARIC) longitudinal cohort investigation (13). The ARIC is a longitudinal investigation of atherosclerosis, CVD risk factors and outcomes, which included a complete oral-dental examination between 1996 and 1998. As part of the Dental ARIC ancillary study, participants underwent complete periodontal examinations that recorded the number of missing teeth, probing depth, attachment loss and bleeding upon probing measurements at six sites per tooth, including third molars. Clinical examiners were trained and calibrated against a standard examiner, with corresponding *kappas* indicating excellent to outstanding level of agreement (14).

For the GWA analyses we used two traits, the Centers of Disease Control (CDC) and American Academy of Periodontology (AAP) consensus three-level classification and second, a continuous "extent" score of disease severity. The CDC/AAP classification uses CAL and PD criteria to define three CP categories as, healthy-mild, moderate and severe (15) (Supplemental Table 1). The "extent" of disease score was defined as the proportion of measured sites that exhibited CAL equal or greater than 3mm (16). The rationale for examining the continuous trait is supported by the fact that attachment loss is a non-reversible marker of periodontal destruction, whereas the CDC/AAP taxonomy includes a PD criterion, which is potentially reversible. Additional covariates that we used for descriptive and exploratory 'adjusted' genetic models were smoking (measured as a three-level ordinal categorical variable: never, former and current), and diabetic status (fasting glucose levels of ≥126 mg/dL, non-fasting of ≥200 mg/dL, or pharmacological treatment for diabetes).

Genotyping was based on the Affymetrix Genome-Wide Human SNP Array 6.0 chip which offers 906,600 SNP markers. The platform offers 906,600 markers for SNPs. Following rigorous quality control procedures, imputation to 2.5million markers was performed using 669,450 SNPs and MACH v1.0.16 (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html), based on HapMap Phase II CEU build 36. Comprehensive descriptions of genotyping and imputation, quality control and population stratification procedures, are presented in the Appendix.

Two analytical endpoints were considered for the main effects analysis of the present project: CDC/AAP CP disease classification and "extent" of disease (attachment loss). Two contrasts were considered for the first trait: moderate vs. mild/healthy and severe vs. mild/healthy. The rationale for considering these contrasts is that severe and moderate CP are

considered distinct forms rather than variable expressions or "natural progression" stages of the disease. These analyses relied upon logistic regression models where allele effects were considered multiplicative (log-additive). To investigate the continuous trait of attachment loss, the proportion of sites exhibiting ≥3mm CAL was Z-score transformed into a normal curve equivalent variable, and a linear genetic model was used. All models included age, sex, examination center and ten principal components from the EIGENSTRAT analysis as covariates. A correction for multiple comparisons was employed assuming 1 million independent tests resulting in a threshold of genome-wide statistical significance of $P<5x10^{-8}$. Although variables such as smoking and DM are not likely confounders of the association between genetic polymorphisms and risk of CP, we developed a series of models adjusting for these variables, as a sensitivity analysis. For this step, the results of which are presented in the supplemental material, we used an arbitrary criterion of >10% change-in-estimate which is often used in confounding evaluation in epidemiologic studies (17). We used this criterion to inspect for "important" changes in estimate for the prioritized SNPs, and explored whether additional ones emerged below the $P<5x10^{-6}$ threshold upon the resulting variance reduction.

All genetic analyses were performed with the ProbABEL software (18). Post-analysis procedures included the generation of quantile-quantile (Q-Q) and Manhattan plots. A threshold of $P<5x10^{-6}$ was set for prioritizing SNPs for further investigation. SNPs were annotated using WGAViewer ver.1.26l (19) and Snipper ver. 1.2 (http://csg.sph.umich.edu/boehnke/snipper/), and regions were viewed using LocusZoom ver.1.1 (20) and Haploview ver.4.2 (21). We used additional online resources of the National Center for Biotechnology Information (NCBI- http://www.ncbi.nlm.nih.gov/). Reporting of

genes was based on the "HUGO Gene Nomenclature" naming convention (www.genenames.org).

E. RESULTS

The descriptive characteristics of the Dental ARIC cohort participants that were included in this analysis are presented in Table4.1. Participants had a mean age of 62 years, with a balanced sex distribution. Twelve percent were current smokers, and eleven percent had DM. In terms of CDC/AAP periodontal diagnoses, these were severe—17%, moderate—43%, and healthy—40%. In the main analyses we found no genome-wide significant association signals. Of the total 2178777 examined SNPs, 26 had a $P<5x10^{-6}$ and thus were prioritized for further investigation (Figures 4.1-3). Lambda variance inflation factors for the three traits were low, and ranged between 1.003 for attachment loss to 1.024 for severe CP (Figures 4.10-12). The prioritized SNPs marked three loci as associated with moderate (6p21.1, 19p13.3 and 10p15), three with severe CP (14.21, 7p15, and 3p21), and four loci with the continuous trait of attachment loss (6q15, 11q24, 11q22, and 7q34).

A comprehensive annotation of the prioritized SNPs, as well as corresponding allele frequencies by trait is presented in Table 4.2. Visualizations of the corresponding loci, along with nearby genes and recombination rates are presented in Figure 4.4-9 and Supplemental Figures 4.13-18. For severe CP, the strongest (with respect to P-value) association in the 14q21 locus was produced by rs12883458, intronic to *NIN*. The minor C allele showed a 4% enrichment among severe CP patients, and was associated with an OR=1.89, $P=3.5x10^{-7}$. In the 7p15 locus, the common allele of rs2521634 (47Kb from *NPY*) produced an OR=1.47, $P=1.6x10^{-6}$. Similarly, the common G allele of rs11925054 in the 3p21 locus, adjacent to

WNT5A and *ERC2* produced the strongest signal in the region, showing 4% enrichment among severe CP patients. With regard to moderate CP, the SNP with the lowest P-value in 6p21.1 was rs7762544 (OR=1.41; $P=1.1x10^{-7}$). This variant is 61Kb from *NCR2*, and its minor (risk) allele showed 5% enrichment among moderate CP patients compared to healthy participants. Rs3826782 which is intronic to *EMR1* and 30Kb from *VAV1* provided the strongest signal in that locus with OR=2.00 and $P=4.0x10^{-6}$.

With regard to attachment loss, several SNPs in LD with the top SNP rs17792917 (177Kb from *TBX18*; $P=1.8x10^{-6}$) were found in the 6q15 locus. Individuals with 2 copies of the rare (risk) T allele had more than 10% higher mean extent of attachment loss (22.7% vs. 20.2%) compared to those with CC genotype. The 11q24 locus where rs10790919 ($P=1.9x10^{-6}$) provided the strongest signal was in an intergenic region (671Kb from *ETS1* and 787Kb from *KIRREL3*). Rs7120142 provided the lowest P-value in the 11q22 locus ($P=3.9x10^{-6}$), and was located 67Kb from *DYNC2H1* and 360Kb from *PDGFD*. Carriers of 2 risk allele copies (rare allele T) had a substantially higher mean extent of attachment loss (25.0) compared to those with no risk allele copy (20.4). The 7q34 locus is a gene-rich area, where rs10500130 emerged as the top SNP ($P=4.6x10^{-6}$), with the common C allele being the one associated with higher extent of attachment loss score. Rs10500130 was 3Kb from *TTC26* and 21Kb from *ZC3HAV1*. Rs1537415 which was the "top hit" in a GWA of AgP (11) did not show an important association (P=0.5) with the any of our examined traits.

F. DISCUSSION

This manuscript presents results of the first genome-wide exploration for loci associated with the risk of CP. The study is limited by the sample size for a GWA, of about 4,000

subjects and the absence of a replication sample. The study strengths include a well-defined cohort with a detailed phenotypical characterization of CP, using full mouth periodontal examinations and the latest commonly accepted taxonomy of CP. The CDC/AAP classification has been used in epidemiologic studies and surveillance. Offenbacher and colleagues (22) introduced a refined CP classification characterizing the disease's biological (versus clinical-only) phenotype that may be more useful when exploring genetic effects of a trait with underlying heterogeneity. Nonetheless, the results of the present investigation provide a wealth of new information on potential candidate genes and mechanistic pathway analyses that will require further exploration, replication and validation in future studies. In general, it can be anticipated that interrogations of the genetic etiology of CP will identify markers associated with defense molecules and pathways, as is the case with numerous candidate-gene approaches (10).

Several promising loci and candidate genes were identified in the present analysis. *NIN* is a gene known to encode a protein that plays a role in centrosomal microtubule organization and anchoring, which have recently been recognized as elements of the T cell cytolytic response (23). Interestingly, a recent candidate-gene study by Olson and colleagues reported two of the "top hits" for severe CP in 14q21 of the present study (rs12893300 and rs1004832) as associated with breast cancer risk (24). *NPY* is a gene encoding a neuropeptide that is widely expressed in the central nervous system and has been suggested to function as an anxiolytic peptide that helps explain inter-individual variation in trait anxiety and resiliency to stress (25); a common risk factor for many diseases, including CP. Moreover, experimental evidence supports a role of WNT5A in the activation of MMP-2 and the regulation of inflammatory cytokine genes of macrophages (26). *NCR2* is the gene in the

region with strongest signal for moderate CP. It appears to have an important role in both normal and pathological innate immune responses, and is selectively expressed by Interleukin 2-activated natural killer cells (27). *EMR1, VAV1* and *CELF2* have also been implicated in immune functions such as eosinophilic inflammation, T-cell and B-cell development and activation (28).

With regard to the "extent" of disease trait, apart from two genes are adjacent to the locus 6q15 (*TBX18* at 177kb and *KIAA1009* at 300Kb), the top SNP is 1Kb upstream of the novel processed transcript RP1-90L14.1-001. Other promising loci and potential candidate genes include *PDGFD* in 11q22, *TTC26* and *ZC3HAV1* at 7q34. Wagsater and colleagues (29) reported that *PDGFs* play important roles in atherosclerosis by stimulating matrix metalloproteinase (MMP) activity and influencing monocyte migration. *ZC3HAV1* encodes a zinc finger protein that is thought to prevent infection by retroviruses, possibly by directly interacting with viral RNA (30)

The lack of an overlap of identified loci for the three examined traits is not surprising, and in fact verifies the rationale of examining these three phenotypes separately. Moderate and severe CP are considered largely distinct entities and include in their case definitions a reversible clinical marker (probing depth), while "extent" attachment loss is a cumulative and non-reversible marker of periodontal destruction. From a statistical standpoint, the lack of overlap is not surprising because small stochastic variations can have a big impact on the tails of the test statistic distribution. However, because these traits share a common pathogenetic underpinning, and in our analyses we used the same controls for both contrasts, some overlap in GWA signals should be anticipated. Explorations at lower P-value thresholds may reveal more "good signals" and common risk loci. As an example, in our

exploratory smoking and diabetes-adjusted analyses 4q21.3 emerged as a common risk locus for severe CP and attachment loss and may be a reflection of this overlap (Figure 4.19).

It must be acknowledged that although GWAS have provided invaluable new information on the genetic basis for many disease and health outcomes, a large component of the disease variation remains to-date unexplained (31). The inability of GWAS to detect very small effect sizes and interrogate rare polymorphisms, gene x gene interactions and epigenetics are additional layers of unaddressed complexity. In our study, although no SNP reached the level of genome-wide statistical significance, several loci were nominally associated with substantial effect sizes and, if replicated, may offer promising avenues for further investigation and mechanistic studies.

Further investigations providing replication of these findings and additional examination of specific associations may lead to an improved understanding of the pathogenesis of the disease, as well as novel preventive and therapeutic approaches.

MANUSCRIPT #1- REPORTING OF FUNDING SOURCES

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C; and by a grant from the National Institute of Environmental Health Sciences (P30ES010126). The authors thank the staff and participants of the ARIC

study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The authors declare no potential conflicts of interest with regard to the authorship and publication of this article.

G. TABLES

Table 4.1 Distribution in the total sample and bivariate associations of periodontal diagnosis (CDC/AAP disease classification) and periodontal attachment loss (extent score: percent of sites with ≥3mm attachment loss) with sex, smoking and diabetic status among the white participants of the Dental ARIC study (n=4610).

		CDC/AAP Periodontitis classification			Attachment loss (≥3mm)
	Total (n, column %)	Health-mild (n, row %)	Moderate (n, row %)	Severe (n, row %)	Extent score[1] mean(SD); median[2]
Total sample	4610 (100)	1864 (40)	1961 (43)	785 (17)	21.1 (21.2); 14
Sex[3,4]					
Females	2415 (52)	1197 (50)	939 (39)	279 (12)	16.6 (18.2); 10
Males	2195 (48)	667 (30)	1022 (47)	506 (23)	26.1 (23.2); 19
Smoking status[3,5]					
Never smoker	2104 (47)	1055 (50)	817 (39)	232 (11)	15.0 (15.3); 10
Former smoker	1876 (42)	633 (34)	870 (46)	373 (20)	24.1 (22.5); 17
Current smoker	526 (12)	147 (28)	221 (42)	158 (30)	33.7 (27.7); 25
Diabetic status[3,4]					
Healthy	4077 (89)	1704 (42)	1704 (42)	669 (16)	20.6 (20.9); 14
Diabetes mellitus	527 (11)	158 (30)	254 (48)	115 (22)	25.4 (23.9); 18

1: Third molars were included in the calculation of the extent of attachment loss trait
2: Rounded to the closest integer
3: Chi-square test of equivalence between strata of periodontitis $P<0.05$
4: Median test of stratum-specific attachment loss estimates $P<0.05$
5: Kruskal-Wallis test of stratum-specific attachment loss or covariate estimates $P<0.05$

Table 4.2. Genome-wide association analysis results of the CDC/AAP chronic periodontitis (CP) classification traits (severe CP vs. healthy and moderate CP vs. healthy) and extent of attachment loss trait (EAL-proportion of sites exhibiting attachment loss ≥3mm), among the white participants of the Dental ARIC study (n=4610). Single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF-HapMap II CEU) of ≥5% and associated $P<5x10^{-6}$. The SNP with the lowest P-value per locus is presented; additional prioritized SNPs in each locus are presented in the footnote, along with corresponding R^2 (based on 1000 genomes pilot 1 release) with the top SNPs.

Chr.	SNP	Position Build36	ca[1]	nca[2]	Risk allele frequency (HapMap II-CEU)	Closest gene(s) and position or distance	P value (beta)	Risk allele frequency disease/ healthy	Odds ratio (95% CI[3])	
Severe CP										
14q21[4]	rs12883458	50349129	C	T	[C] 0.104	*NIN* (Intronic)	$3.5x10^{-7}$ (0.64)	0.13/0.09	1.89 (1.48, 2.41)	
7p15[5]	rs2521634	24344565	A	G	[G] 0.754	*NPY* (47Kb)	$1.6x10^{-6}$ (0.39)	0.80/0.74	1.47 (1.25, 1.73)	
3p21[6]	rs11925054	55365926	G	T	[G] 0.865	*WNT5A* (109Kb); *ERC2* (151Kb)	$6.5x10^{-7}$ (0.53)	0.90/0.86	1.69 (1.37, 2.10)	
Moderate CP										
6p21.1[7]	rs7762544	41487293	A	G	[G] 0.184	*NCR2* (61Kb)	$1.1x10^{-7}$ (0.34)	0.21/0.16	1.41 (1.24, 1.60)	
19p13.3[8]	rs3826782	6838736	A	G	[A] 0.070	*EMR1* (Intronic); *VAV1* (30Kb)	$4.0x10^{-6}$ (0.69)	0.05/0.04	2.00 (1.48, 2.70)	
10p15	rs12260727	10378335	A	G	[G] 0.846	*CELF2* (709Kb)	$6.0x10^{-7}$ (0.43)	0.89/0.85	1.54 (1.30, 1.82)	
								Mean EAL[9] (95% CI[10]) by genotype (number of risk allele copies)		
Extent of attachment loss (≥3mm)								0 copy	1 copy	2 copies
6q15[11]	rs17792917	85323684	C	T	[T] 0.242	*TBX18* (177Kb); *KIAA1009* (330Kb)	$1.8x10^{-6}$ (1.92)	20.2 (19.4, 21.0)	22.4 (21.3, 23.4)	22.7 (20.0, 25.4)
11q24[12]	rs10790919	127162281	A	G	[A] 0.788	*ETS1* (671Kb); *KIRREL3* (787Kb)	$1.9x10^{-6}$ (2.04)	17.1 (14.5, 19.5)	19.9 (18.9, 20.9)	22.0 (21.2, 22.8)
11q22[13]	rs7120142	102922991	C	T	[T] 0.197	*DYNC2H1* (67Kb); *PDGFD* (360Kb)	$3.9x10^{-6}$ (1.98)	20.4 (19.7, 21.2)	22.0 (20.9, 23.1)	25.0 (21.3, 28.7)
7q34	rs10500130	138466453	A	C	[C] 0.889	*TTC26* (3Kb); *ZC3HAV1* (21Kb)	$4.6x10^{-6}$ (2.56)	18.8 (12.6, 25.0)	19.0 (17.7, 20.4)	21.6 (20.9, 22.3)

1:coded allele
2: non-coded allele
3: confidence limits
4: Additional SNPs in locus with $P<5x10^{-6}$: rs1004832 (R^2=1.00), rs8009874 (R^2=0.84), rs12893300 (R^2=0.49)
5: Additional SNP in locus with $P<5x10^{-6}$: rs11771124 (R^2=1.00)
6: Additional SNP in locus with $P<5x10^{-6}$: rs503022 (R^2=0.52)
7: Additional SNPs in locus with $P<5x10^{-6}$: rs9357360 (R^2=0.89), rs1853406 (R^2=0.89), rs1535582 (R^2=0.33)
8: Additional SNP in locus with $P<5x10^{-6}$: rs12610529 (R^2=0.79)
9: Assuming a log-additive linear genetic model
10: CL, confidence limits
11: Additional SNPs in locus with $P<5x10^{-6}$: rs9791329 (R^2=0.58), rs11756650 (R^2=0.96), rs7741380 (R^2=0.60), rs4510639 (R^2=0.67)
12: Additional SNPs in locus with $P<5x10^{-6}$: rs6590279 (R^2=0.87), rs10893747 (R^2=0.87)
13: Additional SNP in locus with $P<5x10^{-6}$: rs4440990 (R^2=1.00)

H. FIGURES

Figure 4.1 Manhattan plot of the genome-wide association analysis results for severe chronic periodontitis among the 4610 white dental ARIC participants.

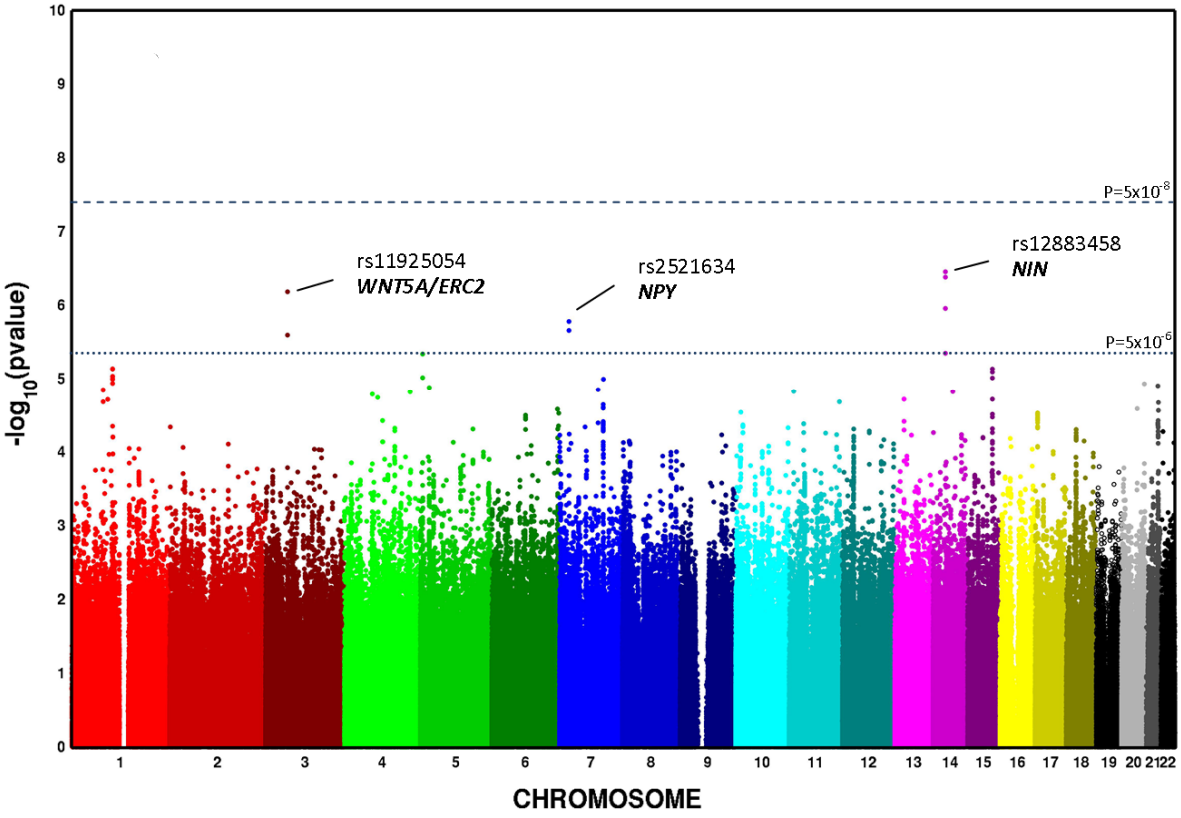


Figure 4.2 Manhattan plot of the genome-wide association analysis results for moderate chronic periodontitis among the 4610 white dental ARIC participants.

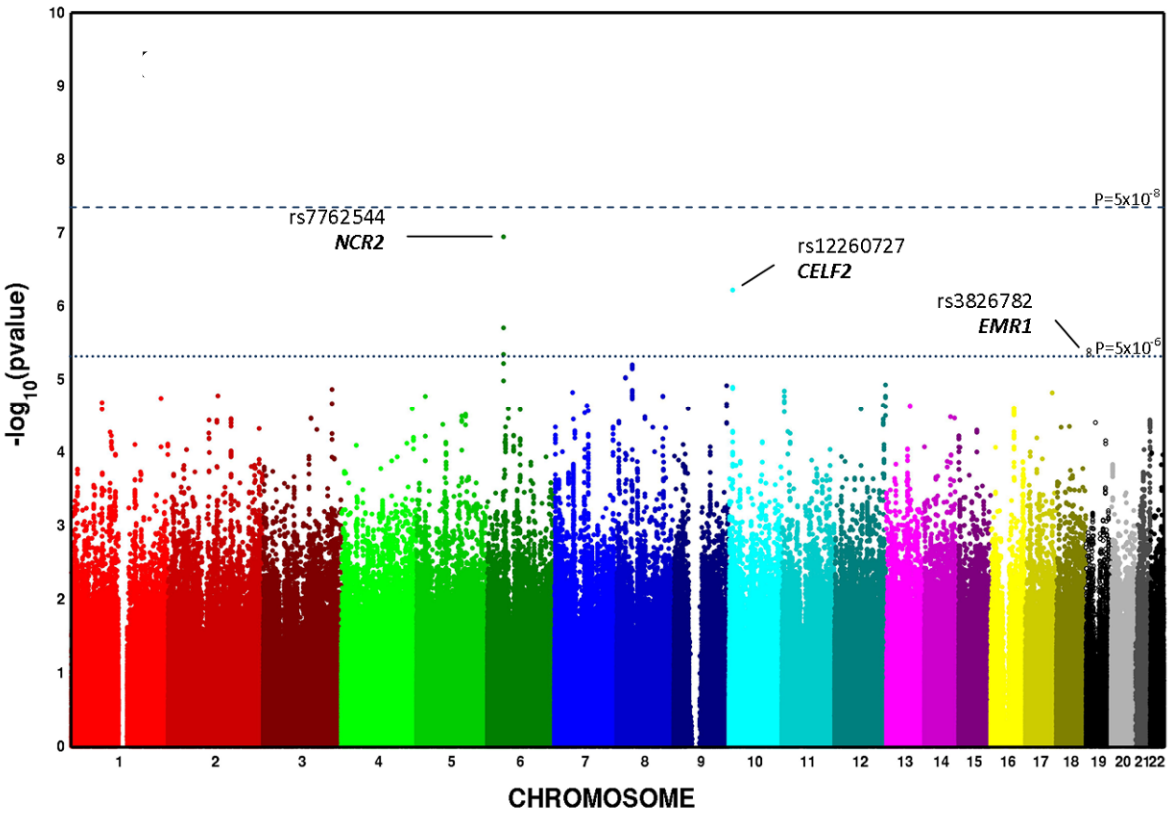


Figure 4.3 Manhattan plot of the genome-wide association analysis results for extent of attachment loss (proportion of sites exhibiting ≥3mm attachment loss) among the 4610 white dental ARIC participants.

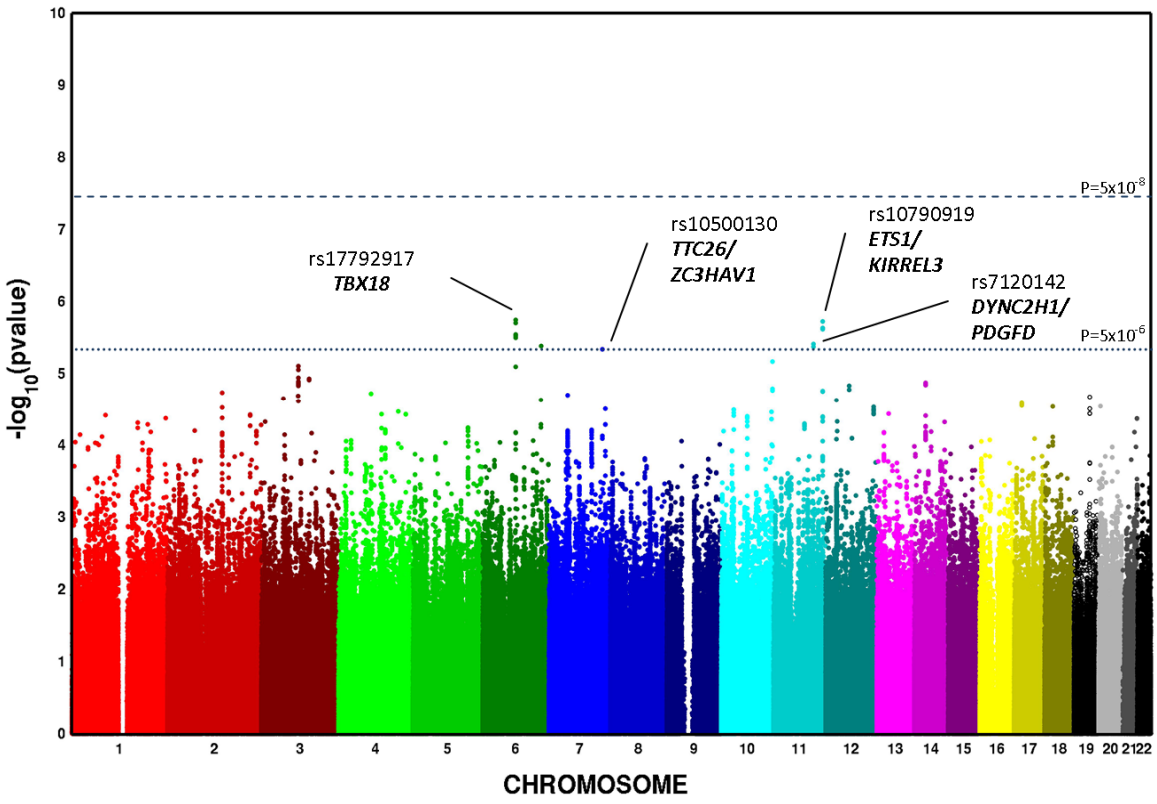


Figure 4.4 Visualization of the 14q21 locus that was marked by rs12883458 ($P=3.5x10^{-7}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated – log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

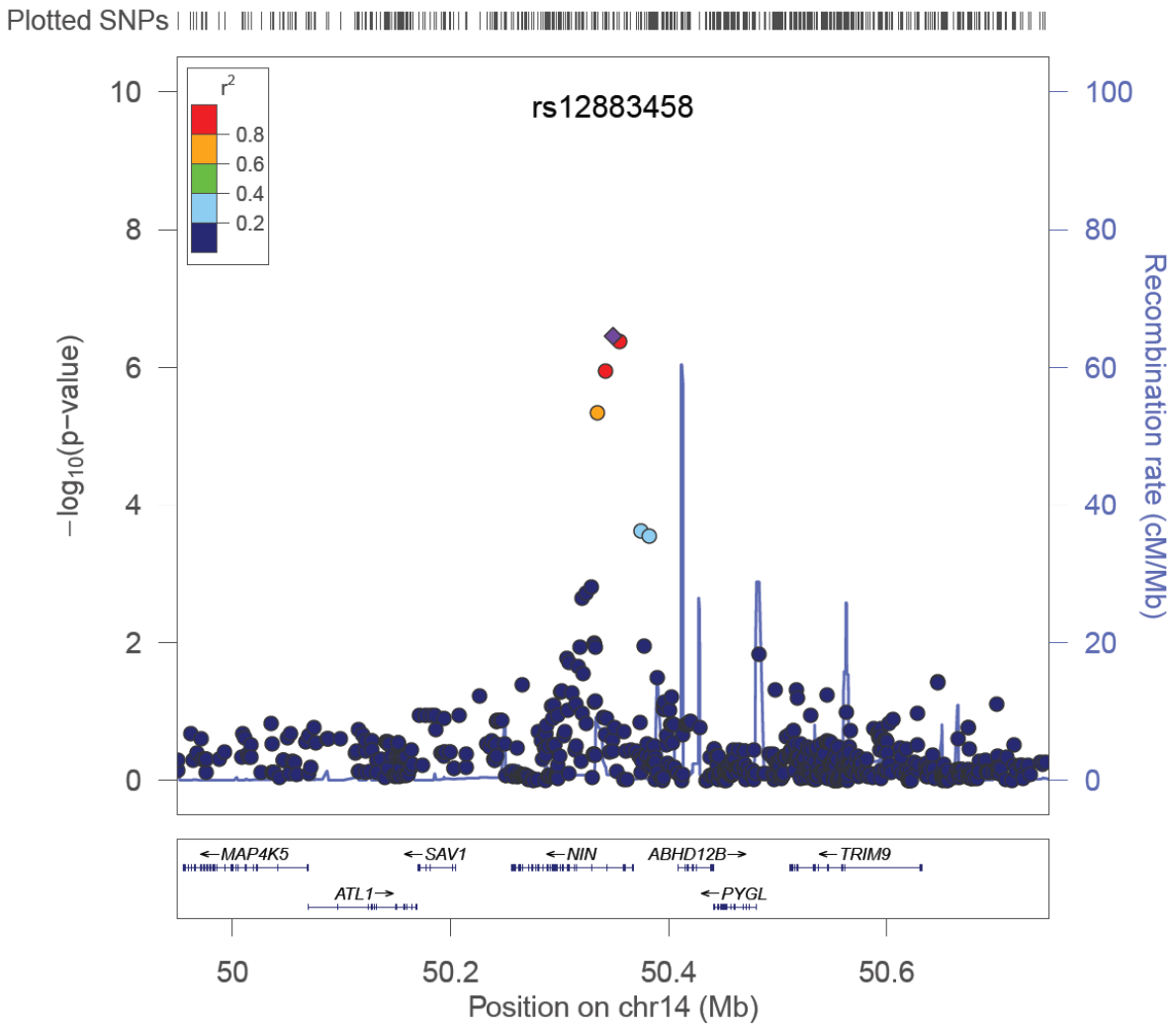


Figure 4.5 Visualization of the 19p13.3 locus that was marked by rs3826782 ($P=4.0x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

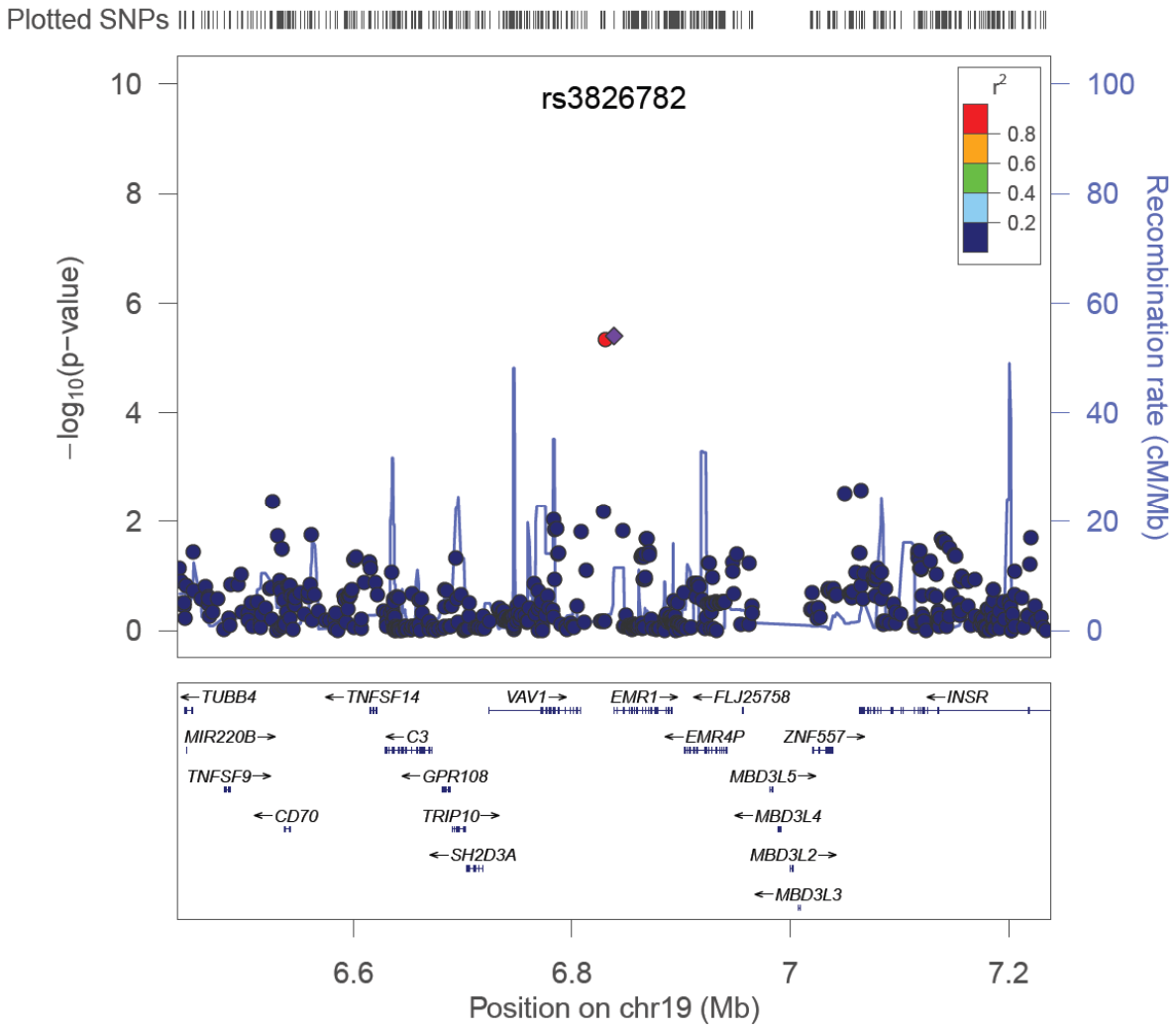


Figure 4.6 Visualization of the 6p21.1 locus that was marked by rs7762544 ($P=1.1x10^{-7}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated – log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

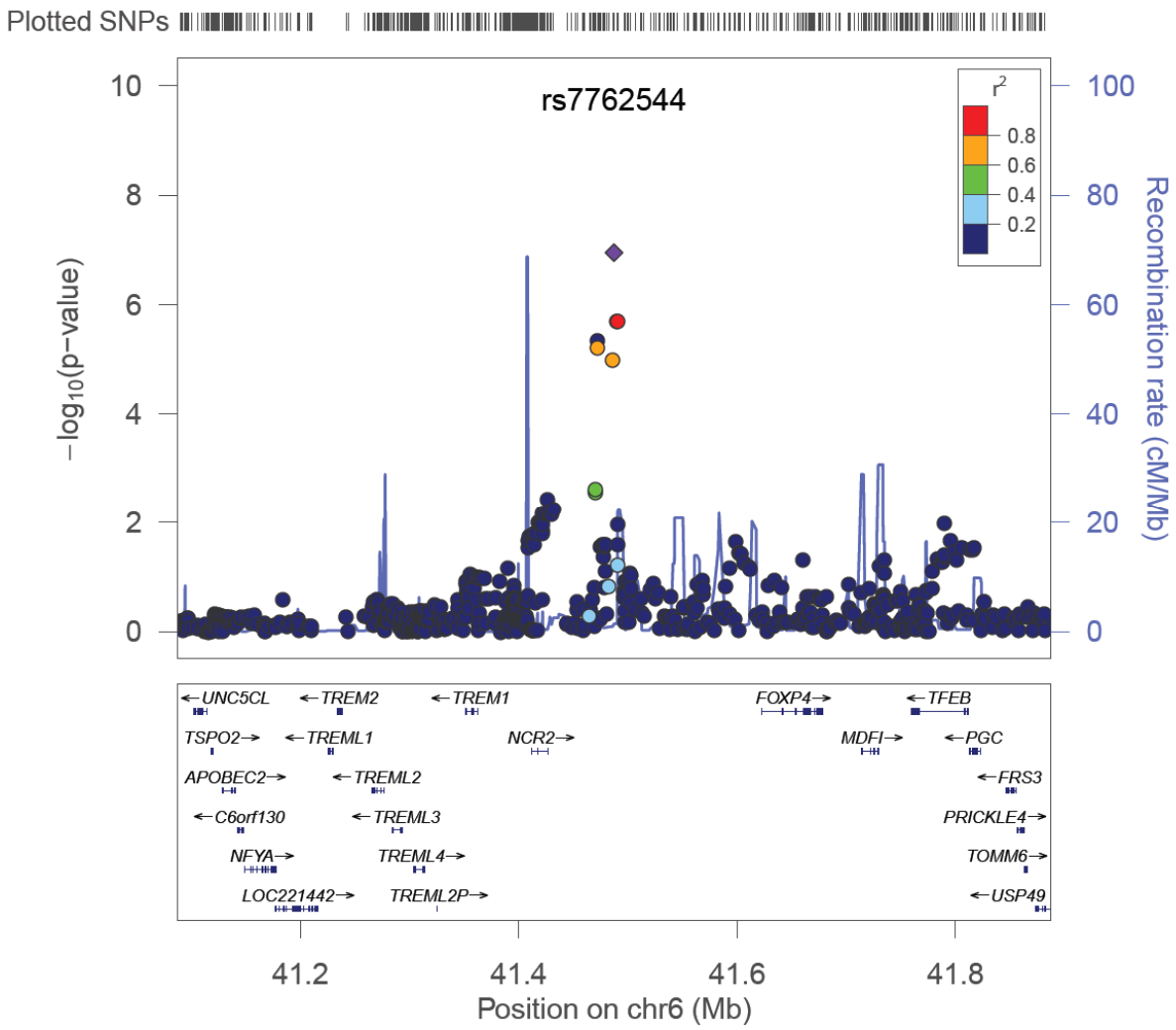


Figure 4.7 Visualization of the 7p15 locus that was marked by rs2521634 ($P=1.6x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated – log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

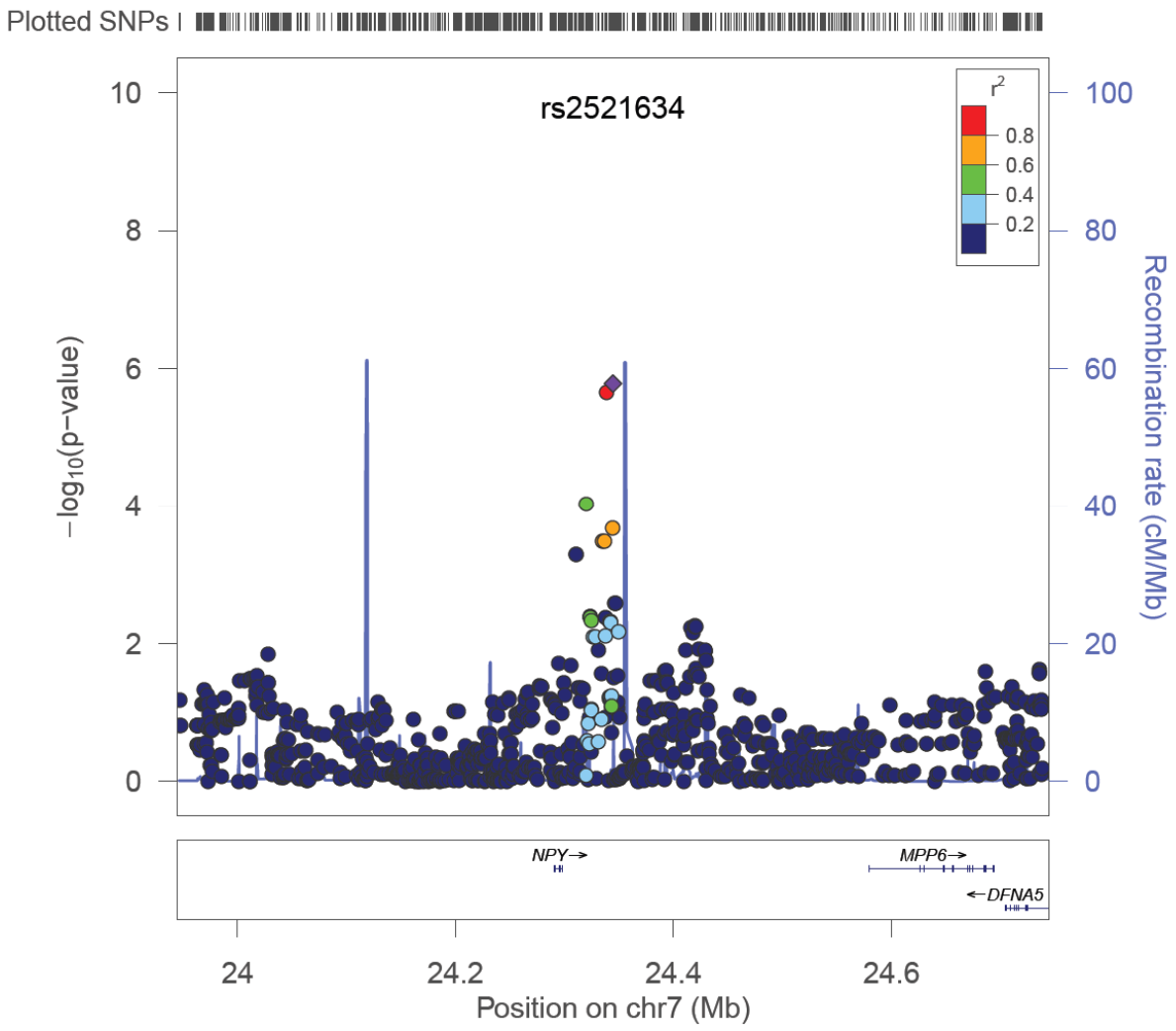


Figure 4.8 Visualization of the 6q15 locus that was marked by rs17792917 ($P=1.8x10^{-6}$) for extent of attachment loss (CAL; sites exhibiting CAL ≥3 mm) among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

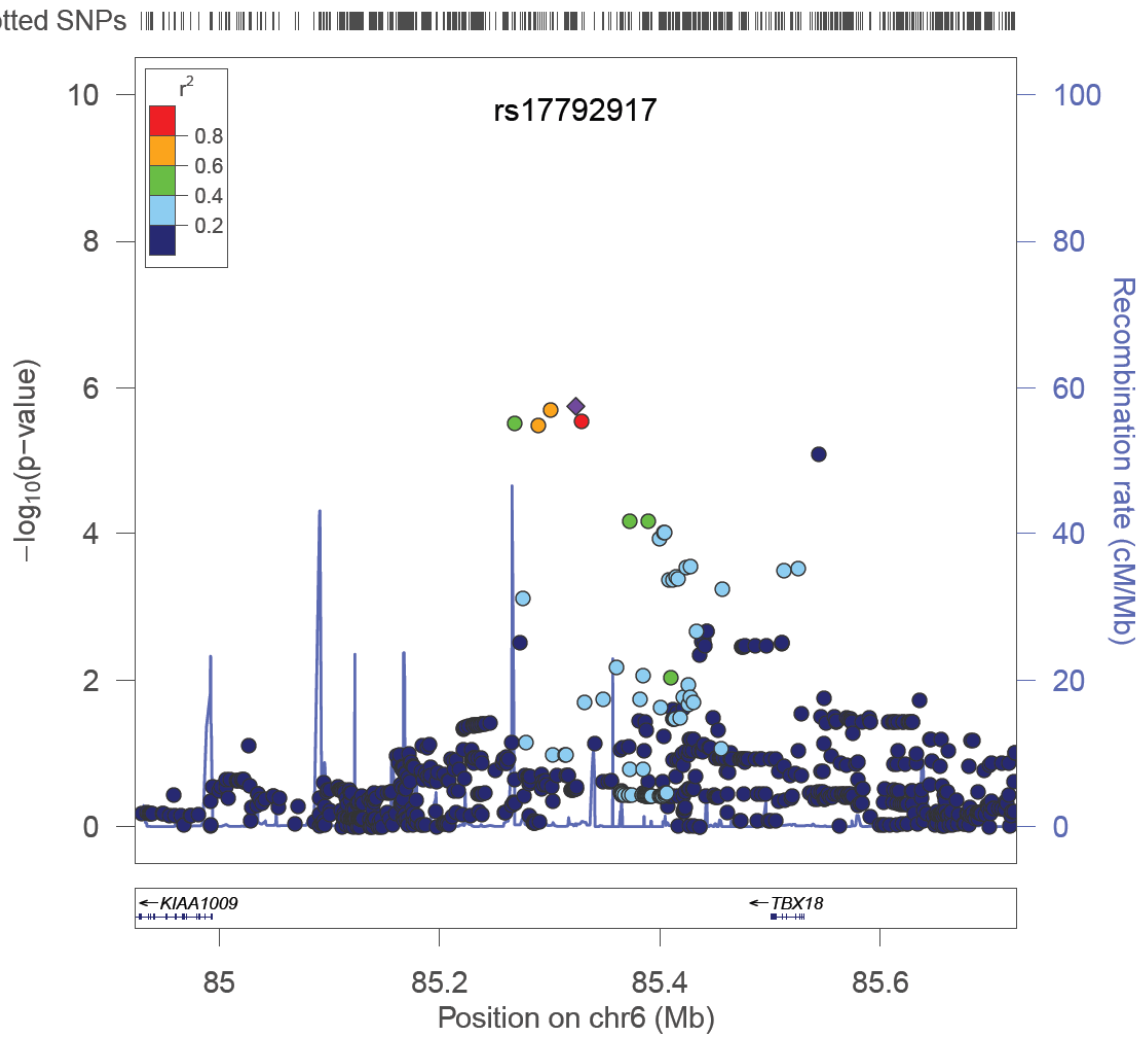
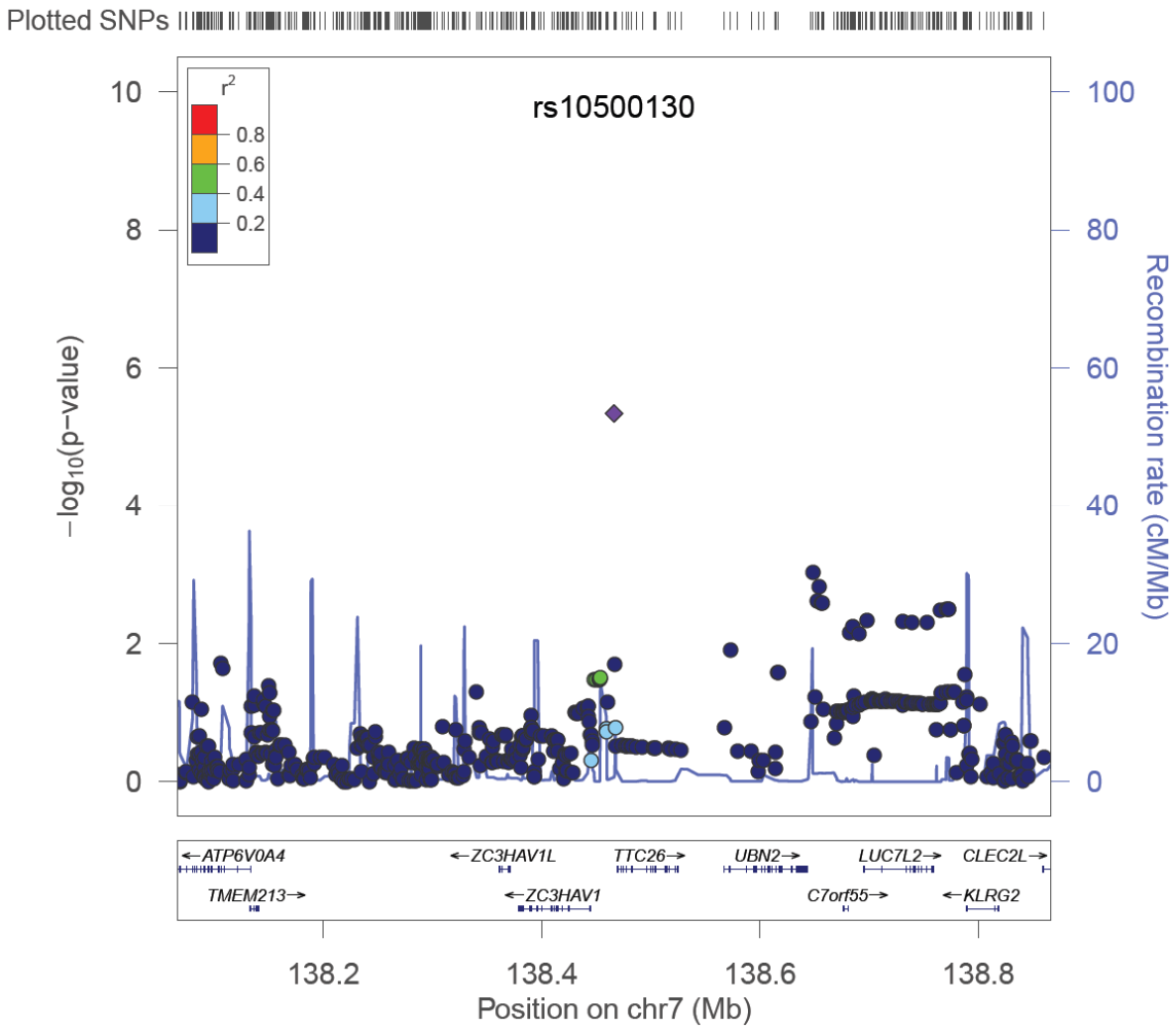


Figure 4.9 Visualization of the 7q34 locus that was marked by rs10500130 ($P=4.6x10^{-6}$) for extent of attachment loss (CAL; sites exhibiting CAL ≥3 mm) among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).



I. REFERENCES

1. Berezow AB, Darveau RP. Microbial shift and periodontitis. Periodontol 2000 2011;55:36-47.

2. Petersen PE, Ogawa H. Strengthening the prevention of periodontal disease: the WHO approach. J Periodontol 2005;76:2187-93.

3. Beck JD, Offenbacher S, Williams R, Gibbs P, Garcia R. Periodontitis: a risk factor for coronary heart disease? Ann Periodontol 1998;3:127-41.

4. Xiong X, Buekens P, Fraser WD, Beck J, Offenbacher S. Periodontal disease and adverse pregnancy outcomes: a systematic review. BJOG 2006;113:135-43.

5. Lalla E, Papapanou PN. Diabetes mellitus and periodontitis: a tale of two common interrelated diseases. Nat Rev Endocrinol 2011. doi: 10.1038/nrendo.2011.106.

6. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

7. Saito T, Shimazaki Y, Sakamoto M. Obesity and periodontitis. N Engl J Med 1998;339:482-3.

8. Saxén L. Heredity of juvenile periodontitis. J Clin Periodontol 1980;7:276-88.

9. Michalowicz BS, Aeppli D, Virag JG, et al. Periodontal findings in adult twins. J Periodontol 1991;62:293-9.

10. Zhang J, Sun X, Xiao L, Xie C, Xuan D, Luo G. Gene polymorphisms and periodontitis. Periodontol 2000 2011;56:102-24.

11. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19:553-62.

12. Schaefer AS, Richter GM, Groessner-Schreiber B, et al. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet2009;5:e1000378.

13. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

14. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

15. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78:1387-99.

16. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.

17. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. Am J Epidemiol 1993;138:923-36.

18. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

19. Ge D, Zhang K, Need AC, et al. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640-43.

20. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336-7.

21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

22. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

23. Jenkins MR, Tsun A, Stinchcombe JC, Griffiths GM. The strength of T cell receptor signal controls the polarization of cytotoxic machinery to the immunological synapse. Immunity 2009;31:621-31.

24. Olson JE, Wang X, Pankratz VS, et al. Centrosome-related genes, genetic variation, and risk of breast cancer. Breast Cancer Res Treat 2011;125:221-8.

25. Zhou Z, Zhu G, Hariri AR, et al. Genetic variation in human NPY expression affects stress response and emotion. Nature 2008;452:997-1001.

26. Kamino M, Kishida M, Kibe T, et al. Wnt-5a signaling is correlated with infiltrative activity in human glioma by inducing cellular migration and MMP-2. Cancer Sci 2011;102:540-8.

27. Cantoni C, Bottino C, Vitale M, et al. NKp44, a triggering receptor involved in tumor cell lysis by activated human natural killer cells, is a novel member of the immunoglobulin superfamily. J Exp Med 1999;189:787-96.

28. Mallory MJ, Jackson J, Weber B, Chi A, Heyd F, Lynch KW. Signal- and development-dependent alternative splicing of LEF1 in T cells is controlled by CELF2. Mol Cell Biol 2011;31:2184-95.

29. Wågsäter D, Zhu C, Björck HM, Eriksson P. Effects of PDGF-C and PDGF-D on monocyte migration and MMP-2 and MMP-9 expression. Atherosclerosis 2009;202:415-23.

30. Jeong MS, Kim EJ, Jang SB. Expression and RNA-binding of human zinc-finger antiviral protein. Biochem Biophys Res Commun 2010;396:696-702.

31. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.

J. MANUSCRIPT #1 - SUPPLEMENTAL MATERIAL

SUPPLEMENTAL MATERIALS AND METHODS

Genotyping and imputation

In the study population, DNA was extracted from blood samples drawn from an antecubital vein into tubes containing serum separator gel. Blood samples were analyzed at a central ARIC laboratory in Houston, TX. Genotyping was performed with the Affymetrix Genome-Wide Human SNP Array 6.0 chip. The platform offers 906,600 markers for SNPs. The rigorous quality control procedures included initial blind duplicate genotyping and identification/flagging of SNPs with kappa<0.95 and reconciliation of unintentional duplicate samples (17 duplicates and one triplicate). Imputation to 2.5million markers was performed using 669,450 SNPs and MACH v1.0.16 (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html), based on HapMap Phase II CEU build 36. The SNPs used for imputation were selected from 839,048 autosomal SNPs restricted to those with minor allele frequency (MAF)>0.01 (129,543 excluded), Hardy-Weinberg equilibrium (HWE) $P>10^{-5}$ (12,432 excluded) and call rate >95% (1,693 excluded). We used the following SNP exclusion criteria for further analyses: quality score <0.8 and missing data rate >10% after imputation, and MAF of <5%.

Population stratification

To obtain estimates of relatedness and population stratification a subset of 85,947 "high quality" SNPs was selected. These SNPs met the following criteria among self-reported whites: MAF ≥0.1, call rate >99.5%, HWE $P\geq10^{-3}$, autosomal, with annotation in the platform annotation file, not labeled "AFFX" or "chromosome 0", and not monomorphic.

Using these SNPs identity-by-state (IBS) allele sharing distance (DST values) were computed using PLINK, as such: DST = IBS distance ($IBS_2 + 0.5*IBS_1$) / (n SNP pairs). First degree relative status was assigned to pairs of individuals with DST≥0.8 and second degree relatives were considered those with 0.763 ≤DST< 0.8. Among the white ARIC participants there were 380 pairs of first degree and 207 pairs of second degree relatives identified. To minimize exclusions, related pairs were broken by iterative selection of individuals with most relatives using a custom program.

Population stratification was further evaluated with principal component (PC) analysis using the EIGENSTRAT method (1). The above chosen set of SNPs was used for the computation of ten principal components. Genetic outliers were considered those that were further than 8 standard deviations (SD) away from any of ten PCs over ten runs of PC computation. Based on DST and PC criteria there were 716 subjects flagged from removal from the analysis (206 as genetic outliers based on PCs and 16 based on average DST values ("too little IBS sharing" with the rest of the sample), 351 first degree relatives and 143 second degree relatives. All but ten second degree relatives (whose relatives were excluded as genetic outliers) were re-entered in the dataset and were assigned PCs. After exclusion of 364 individuals (4%) there were 9349 whites who were included in the GWA analysis and of those, 4610 had periodontal phenotype data available as Dental ARIC participants.

SUPPLEMENTAL RESUTS

Adjusted genetic models

Adjustment for smoking and diabetic status resulted in small (<10%) and predominantly upwards changes in odds ratios of the prioritized SNPs for the two disease classification traits (Supplemental Table S2). Noteworthy, the effect estimate of rs7762544 associated with moderate CP increased by 6% for a genome-wide significant adjusted OR=1.44 ($P=3.8x10^{-8}$). Larger and exclusively downwards changes in effect size estimates were noted for the prioritized SNPs in the domain of attachment loss. These estimate attenuations ranged from 11% for rs10500130 to 23% for rs7120142. Four additional loci emerged below the $P<5x10^{-6}$ threshold in the adjusted analyses, one of those common for severe CP and extent of attachment loss: moderate CP—[rs11615037, OR=1.33; $P_{adj}=4.4x10^{-6}$, in a region between *CPM* (98Kb) and *CPSF6* (178Kb)]; severe CP—(rs10493998, OR=1.68; $P_{adj}=2.2x10^{-6}$, 90Kb from *COL11A)*; extent of CP—(rs8006336, $P_{adj}=1.4x10^{-6}$ adjacent to *RPS29* and *PPIL5*). Rs17006135, was intronic to *SCD5*, approximately 3Kb from the exon boundary, marking the 4q21.3 locus, and was associated with both severe CP (OR=1.82; $P_{adj}=4.4x10^{-6}$) and increased extent of attachment loss ($P_{adj}=8.2x10^{-7}$).

K. SUPPLEMENTAL TABLES

Table 4.3 Clinical case definitions of Periodontitis proposed by the CDC working group for use in population-based surveillance of Periodontitis[1] (2)

	Clinical Definition		
Disease Category	Clinical Attachment Level (CAL)		Probing Depth (PD)
Severe periodontitis	≥2 interproximal sites[2] with CAL ≥ **6mm**	and	≥1 interproximal site with PD ≥5mm
Moderate periodontitis	≥2 interproximal sites[2] with CAL ≥ **4mm**	or	≥2 interproximal sites[2] with PD ≥5mm
No or mild periodontitis	Neither "moderate" nor "severe" periodontitis		

1: third molars are excluded

2: not on the same tooth

Table 4.4 Changes in estimate of top SNPs for the three chronic periodontitis traits after adjustment for smoking (never, former, current) and diabetic status (binary) among the white participants of the Dental ARIC study.

Trait	Chr.	SNP	RAF[1] Disease/ Healthy	Crude estimates		Adjusted[2] estimates		Change in estimate[3] (percent)
				OR/beta[4] (95% CL)	P value	OR/beta[4] (95% CL[5])	P value	
Moderate vs. Healthy	6p21.1	rs7762544	0.21/0.16	1.41 (1.24, 1.60)	$1.1x10^{-7}$	1.44 (1.26, 1.63)	$3.8x10^{-8}$	+6%
	19p13.3	rs3826782	0.05/0.04	2.00 (1.48, 2.70)	$4.0x10^{-6}$	2.09 (1.54, 2.84)	$1.4x10^{-6}$	+7%
	10p15	rs12260727	0.15/0.11	1.54 (1.30, 1.82)	$6.0x10^{-7}$	1.51 (1.27, 1.79)	$3.0x10^{-6}$	-4%
Severe vs. Healthy	14q21	rs12883458	0.13/0.09	1.89 (1.48, 2.41)	$3.5x10^{-7}$	1.88 (1.46, 2.43)	$1.1x10^{-6}$	-0.3%
	7p15	rs2521634	0.80/0.74	1.47 (1.25, 1.73)	$1.6x10^{-6}$	1.43 (1.21, 1.69)	$2.6x10^{-5}$	-7%
	3p21	rs11925054	0.90/0.86	1.69 (1.37, 2.10)	$6.5x10^{-7}$	1.78 (1.42, 2.23)	$2.6x10^{-7}$	+10%
Extent of attachment loss	6q15	rs17792917		1.92 (1.13, 2.71)	$1.8x10^{-6}$	1.61 (0.84, 2.38)	$4.1x10^{-5}$	-16%
	11q24	rs10790919		2.04 (1.20, 2.88)	$1.9x10^{-6}$	1.78 (0.96, 2.59)	$1.9x10^{-5}$	-13%
	11q22	rs7120142		1.98 (1.14, 2.83)	$3.9x10^{-6}$	1.54 (0.71, 2.36)	$2.5x10^{-4}$	-23%
	7q34	rs10500130		2.56 (1.47, 3.66)	$4.6x10^{-6}$	2.28 (1.21, 3.34)	$2.8x10^{-5}$	-11%

1: Risk allele frequency
2: Adjusted for smoking (three-level categorical where 0: never, 1: former, 2: current smoker) and diabetic status (dichotomous variable where 0: healthy and 1: diabetic)
3: Calculated as follows: change-in-estimate= $[(|beta_{adjusted}| - |beta_{crude}|) / |beta_{crude}|] *100$
4: OR (odds ratios) are presented for the two binary periodontitis classification traits and beta coefficients for the continuous trait of attachment loss
5: Confidence limits

Table 4.5 Distribution of genotype (based on number of minor allele copies of SNPs with $P<5x10^{-6}$) by phenotypic classification (three disease traits and mean "extent" of disease) among the Dental ARIC participants.

		Healthy (%)			Moderate chronic periodontitis (%)			Severe chronic periodontitis (%)			Mean extent of attachment loss (≥3mm) and (95% confidence limits)		
Minor allele copies (n)		0 (%)	1 (%)	2 (%)	0 (%)	1 (%)	2 (%)	0 (%)	1 (%)	2 (%)	0 (%)	1 (%)	2 (%)
Chr.	SNP												
6p21.1	rs7762544[1,4]	69.6	27.9	2.5	62.3	33.9	3.9	68.1	29.3	2.6	20.9 (20.1, 21.6)	21.5 (20.4, 22.7)	22.4 (18.6, 26.2)
19p13.3	rs3826782[1,4]	93.0	6.9	0.1	90.2	9.6	0.3	91.7	8.3	0.0	21.1 (20.4, 21.7)	22.5 (20.3, 24.7)	16.1 (9.1, 23.1)
10p15	rs12260727[1]	72.4	25.5	2.1	78.3	20.6	1.1	76.3	22.0	1.7	21.5 (20.8, 22.3)	20.0 (18.7, 21.3)	17.1 (12.6, 21.6)
14q21	rs12883458[2,4]	82.6	17.1	0.3	80.2	18.9	0.9	74.8	23.6	1.7	20.7 (20.0, 21.4)	22.6 (21.2, 24.1)	25.6 (19.1, 32.1)
7p15	rs2521634[2,4]	54.1	39.2	6.8	55.4	38.8	5.8	63.8	32.4	3.8	21.6 (20.8, 22.4)	20.8 (19.8, 21.8)	18.6 (16.2, 21.1)
3p21	rs11925054[2]	74.0	24.0	2.0	74.5	23.8	1.7	81.0	18.0	1.0	21.7 (21.0, 22.4)	19.6 (18.4, 20.8)	16.8 (13.2, 20.5)
6q15	rs17792917[3,4]	61.9	33.0	5.1	56.9	37.6	5.5	53.6	39.1	7.3	20.2 (19.4, 21.0)	22.4 (21.3, 23.4)	22.7 (20.0, 25.4)
11q24	rs10790919[3]	60.2	35.4	4.4	63.0	33.0	3.9	68.7	28.8	2.5	22.0 (21.2, 22.8)	19.9 (18.9, 20.9)	17.1 (14.8, 19.5)
11q22	rs7120142[3,4]	66.7	30.5	2.8	61.6	34.3	4.1	62.9	33.3	3.8	20.4 (19.7, 21.2)	22.0 (20.9, 23.1)	25.0 (21.3, 28.7)
7q34	rs10500130[3]	79.0	19.7	1.3	81.6	17.5	0.9	82.9	15.9	1.2	18.8 (12.6, 25.0)	19.0 (17.7, 20.4)	21.6 (20.9, 22.3)

1: Associated $P<5x10^{-6}$ for moderate CP
2: Associated $P<5x10^{-6}$ for severe CP
3: Associated $P<5x10^{-6}$ for extent of attachment loss
4: The minor allele is the 'risk allele' for the corresponding trait

Table 4.6 Reported genes' symbols [HUGO Gene Nomenclature Committee (HGNC) approved symbols], names and chromosomic locations.

Gene symbol	Gene name	Location
NCR2	Natural cytotoxicity triggering receptor 2	6p21.1
EMR1	Egf-like module containing, mucin-like, hormone receptor-like 1	19p13.3
VAV1	Vav 1 guanine nucleotide exchange factor	19p13.2
CELF2	CUGBP, Elav-like family member 2	10p13
NIN	Ninein (GSK3B interacting protein)	14q21-q22
NPY	Neuropeptide Y	7p15.3
WNT5A	Wingless-type MMTV integration site family, member 5A	3p21-p14
ERC2	ELKS/RAB6-interacting/CAST family member 2	3p14.3
TBX18	T-box 18	6q14.1-q15
KIAA1009	KIAA1009	14.3
ETS1	V-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	11q23.3
KIRREL3	Kin of IRRE like 3 (Drosophila)	11q24
DYNC2H1	Dynein, cytoplasmic 2, heavy chain 1	11q21-q22.1
PDGFD	Platelet derived growth factor D	11q22.3
TTC26	Tetratricopeptide repeat domain 26	7q34
ZC3HAV1	Zinc finger CCCH-type, antiviral 1	7q34
*CPM**	Carboxypeptidase M	12q15
*CPSF6**	Cleavage and polyadenylation specific factor 6, 68kDa	12q15
*COL11A1**	Collagen, type XI, alpha 1	1p21
*RNPC3**	RNA-binding region (RNP1, RRM) containing 3	1p21.1
*SCD5**	Stearoyl-CoA desaturase 5	4q21.3
*SEC31A**	SEC31 homolog A (S. cerevisiae)	4q21
*RPS29**	Ribosomal protein S29	14q21.3
*LRR1**	Leucine rich repeat protein 1 (PPIL5)	14q21.3

*Genes adjacent to loci that emerged in adjusted analyses.

Table 4.7 Reported SNPs address according to the 1000 genomes (1000G) convention and proxy SNPs (R^2>0.8 where available) based on 1000G pilot 1 release. All SNPs were imputed based on HapMap II CEU with the exception of rs2521634 with was directly genotyped on the Affymetrix 6.0 platform.

dbSNP	1000G	Proxy SNPs (R^2>0.8)
rs7762544	6:41379315	rs1853406, rs9357360
rs3826782	19:6887736	rs12610529
rs12260727	10:10338329	rs12258450, rs7919833
rs12883458	14:51279379	rs1004832
rs2521634	7:24378040	rs11771124, rs10487606, rs10487605
rs11925054	3:55390886	rs503022[1]
rs17792917	6:85266965	rs16874800
rs10790919	11:127657071	rs6590279
rs7120142	11:103417781	rs4440990
rs10500130	7:138815913	rs12673905[2], rs12056143[3]

1: R^2=0.40; 2: R^2=0.46; 3: R^2=0.43

Table 4.8 Reported SNPs as expression quantitative loci (eQTLs) and their association with gene expression (lymphoblastoid cell line-based expression) in populations of European descent with $P<10^{-4}$ in the Scan database (http://www.scandb.org).

dbSNP	Gene symbol	Gene name	Locus	P-value
rs3826782	*GPR113*	G protein-coupled receptor 113	2p23.3	$2x10^{-5}$
rs11925054	*TMEM5*	transmembrane protein 5	12q14.2	$3x10^{-5}$
	SLC20A2	solute carrier family 20 (phosphate transporter), member 2	8p12-p11	$8x10^{-5}$
	RABEPK	Rab9 effector protein with kelch motifs	9q33.3	$9x10^{-5}$
rs17792917	*ULK4*	unc-51-like kinase 4	3p22.1	$1x10^{-4}$
	DUSP4	dual specificity phosphatase 4	8p12-p11	$1x10^{-4}$
	CLCF1	cardiotrophin-like cytokine factor 1	11q13.3	$3x10^{-5}$
	CLYBL	citrate lyase beta like	13q32	$2x10^{-5}$
	TSC22D1	TSC22 domain family, member 1	13q14	$9x10^{-6}$
	TRAK1	trafficking protein, kinesin binding 1	3p22.1	$1x10^{-4}$
	DNMBP	dynamin binding protein	10q24.2	$3x10^{-5}$
	TNFRSF19	tumor necrosis factor receptor superfamily, member 19	13q12.11-q12.3	$1x10^{-4}$
	TBC1D8	TBC1 domain family, member 8	2q11.2	$5x10^{-5}$
	MNX1	motor neuron and pancreas homeobox 1	7q36	$3x10^{-5}$
	SLC43A2	solute carrier family 43, member 2	17p13.3	$8x10^{-5}$
	GOLGA8A	golgin A8 family, member A	15q11.2	$7x10^{-5}$
	GOLGA8B	golgin A8 family, member B	15q14	$7x10^{-5}$
	PLCL2	phospholipase C-like 2	3p24.3	$5x10^{-5}$
rs11615037	*CENTA1/* ADAP1	ArfGAP with dual PH domains 1	7p22.3	$1x10^{-4}$

L. SUPPLEMENTAL FIGURES

Figure 4.10 Quantile-quantile plot of genome-wide association analysis results of severe chronic periodontitis among the 4610 white Dental ARIC participants.

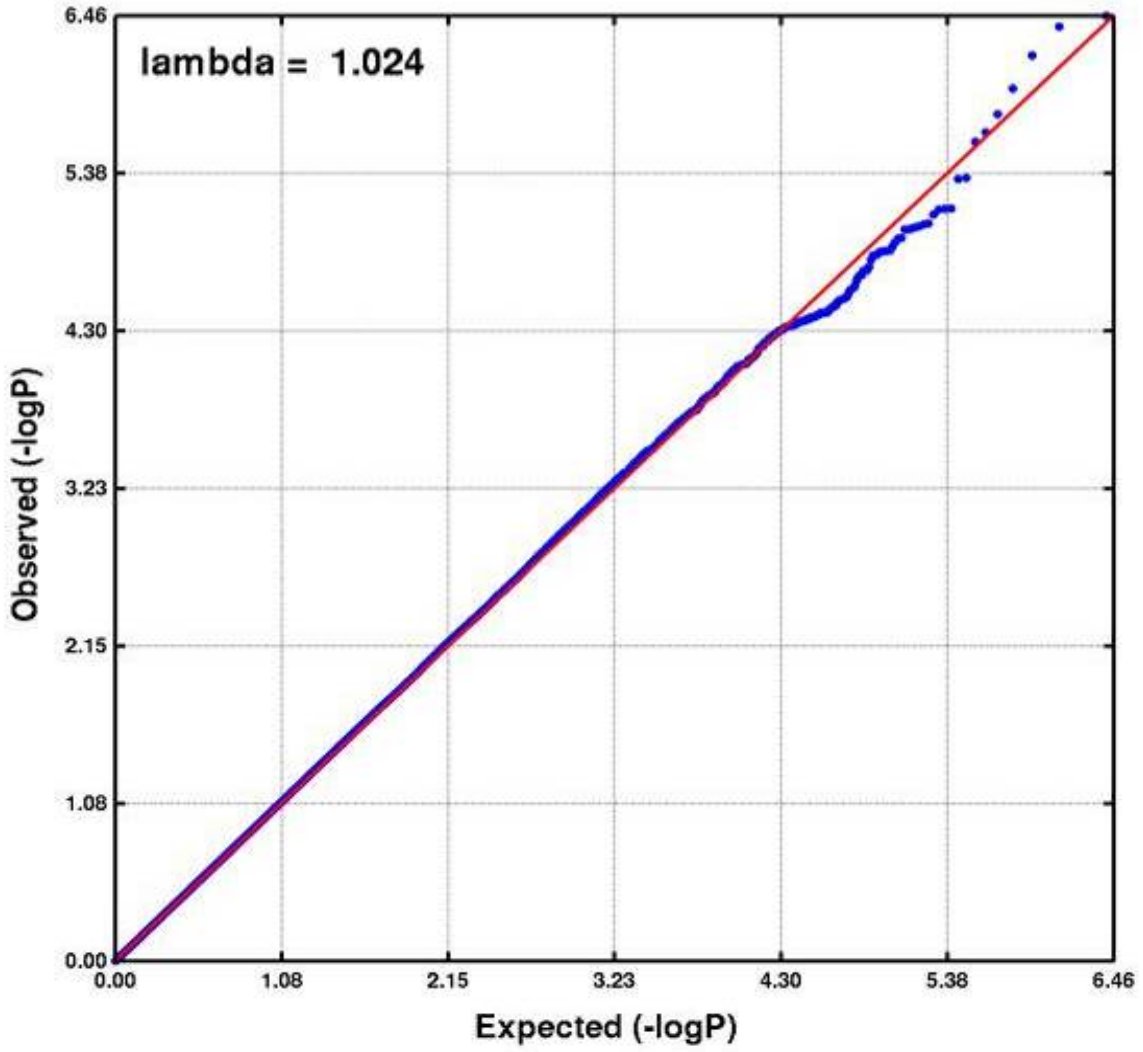


Figure 4.11 Quantile-quantile plot of genome-wide association analysis results of moderate chronic periodontitis among the 4610 white Dental ARIC participants.

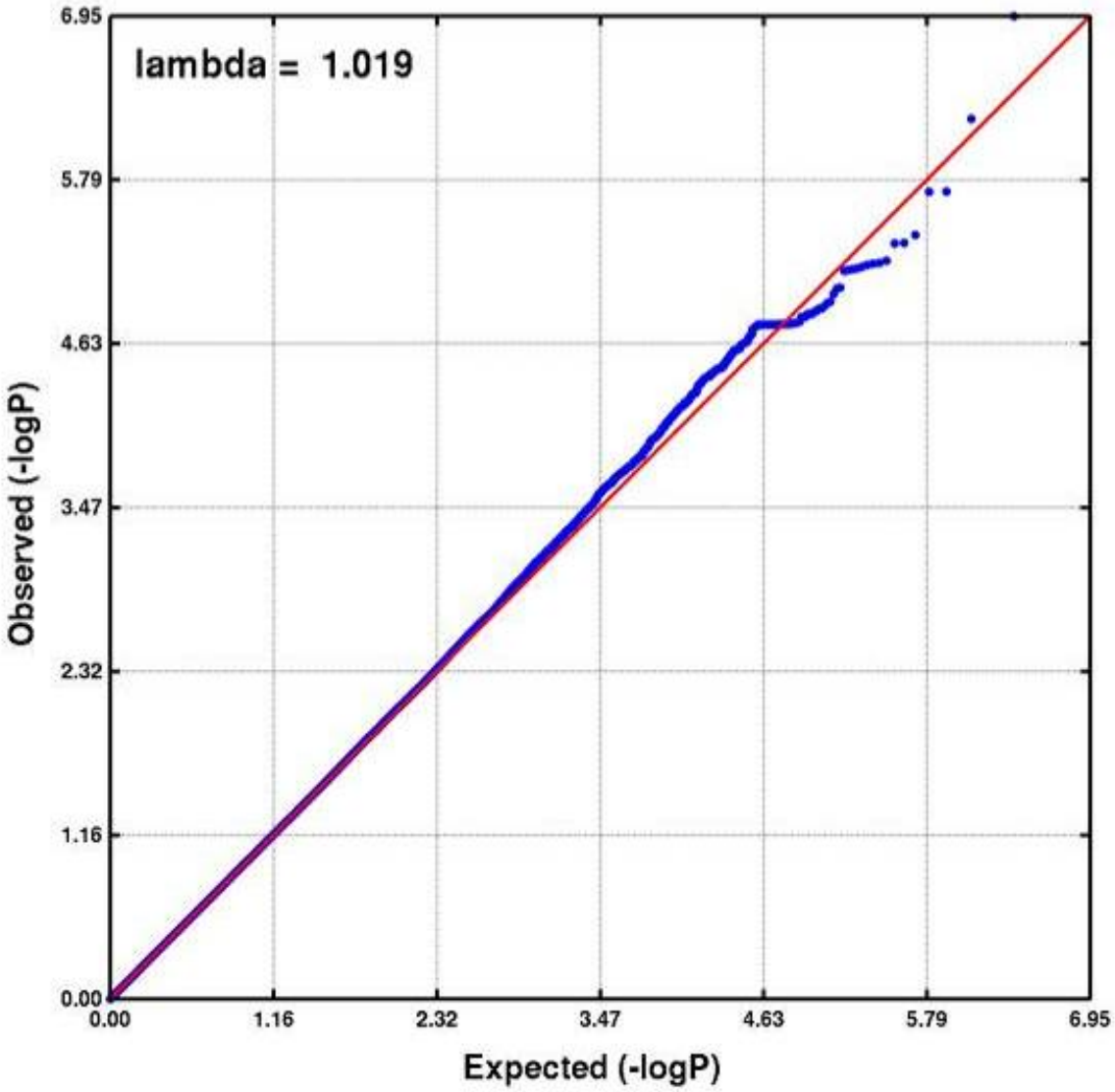


Figure 4.12 Quantile-quantile plot of genome-wide association analysis results of the extent of attachment loss trait (proportion of sites exhibiting ≥3 mm attachment loss) among the 4610 white Dental ARIC participants.

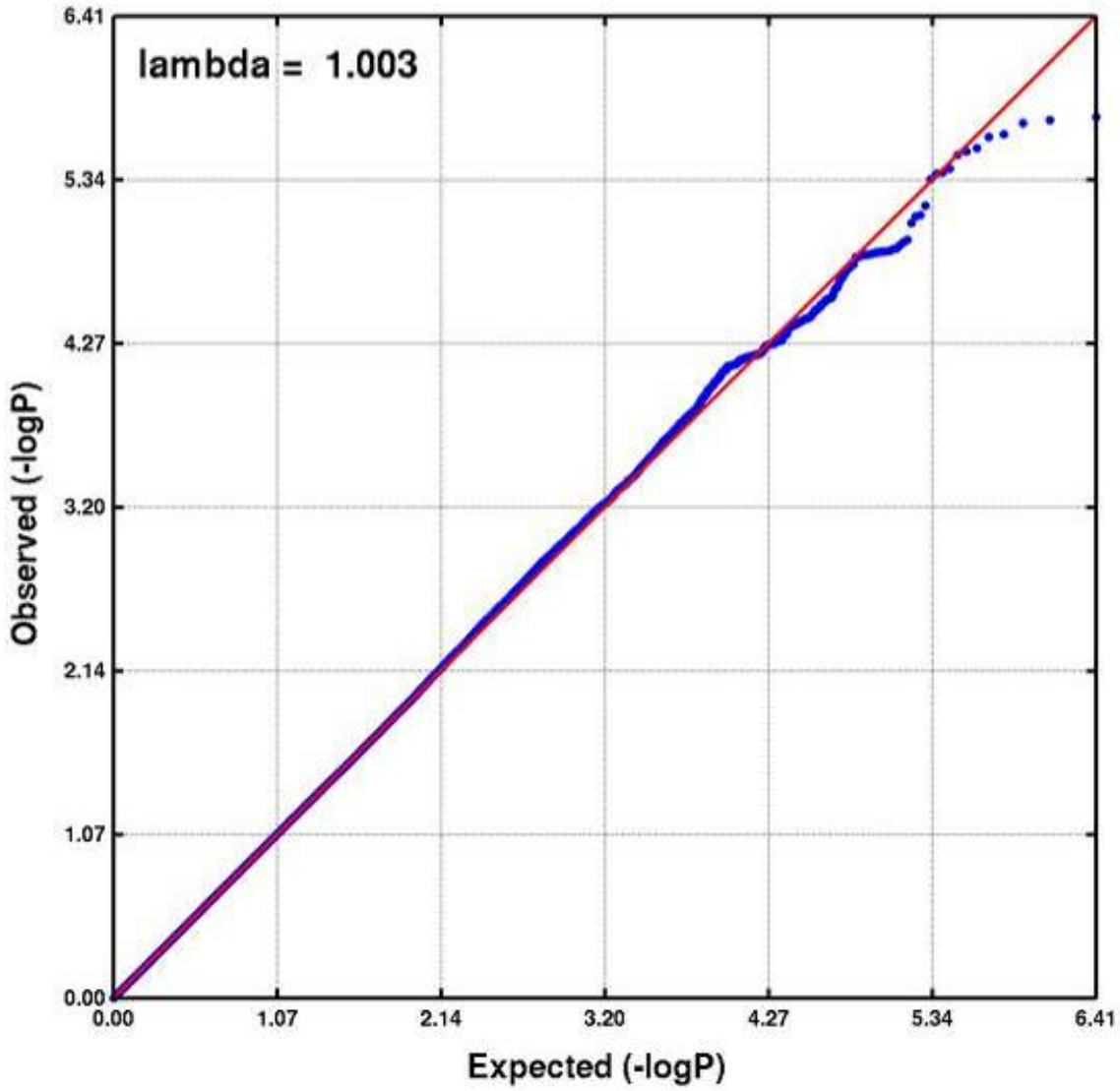


Figure 4.13 Visualization of the 3p21 locus that was marked by rs11925054 ($P=6.5x10^{-7}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

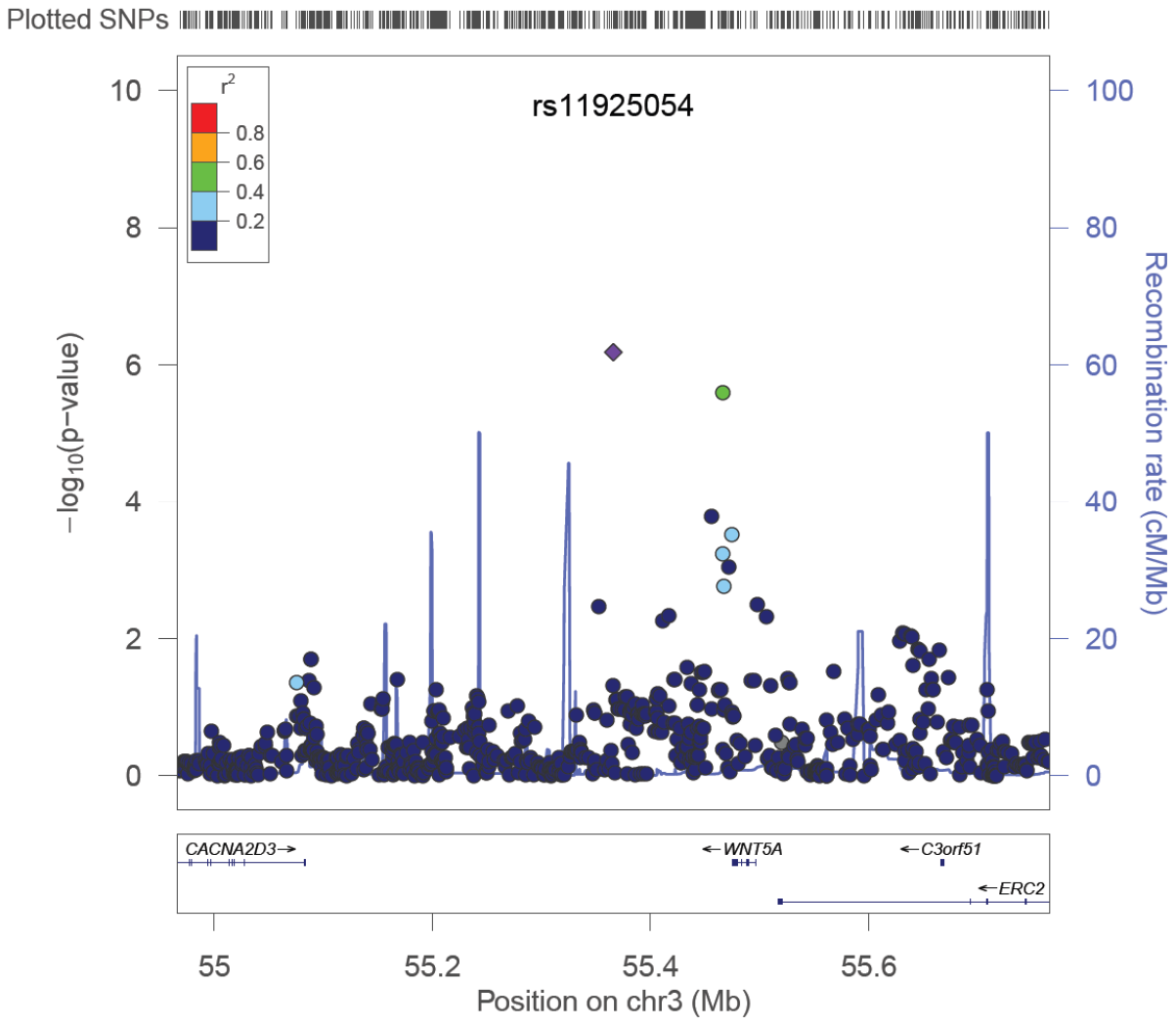


Figure 4.14 Visualization of the 4q21.3 locus that was marked by rs17006135 ($P=1.8x10^{-5}$ and $P=4.4x10^{-6}$ after adjustment for smoking and diabetic status) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated – log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

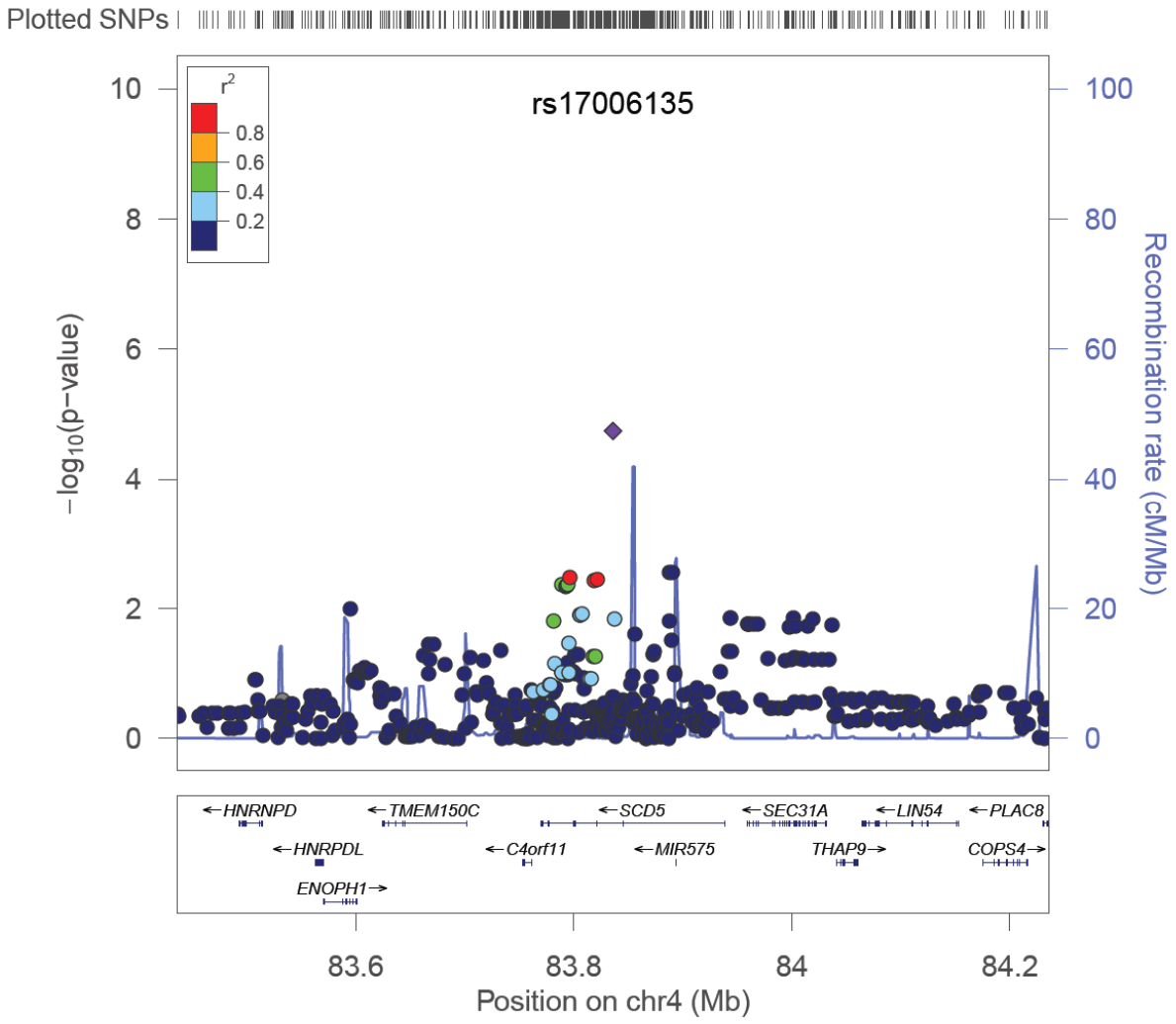


Figure 4.15 Visualization of the 1p21.1 locus that was marked by rs10493998 ($P=9.2x10^{-6}$ and $P=2.2x10^{-6}$ after adjustment for smoking and diabetic status) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated – log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

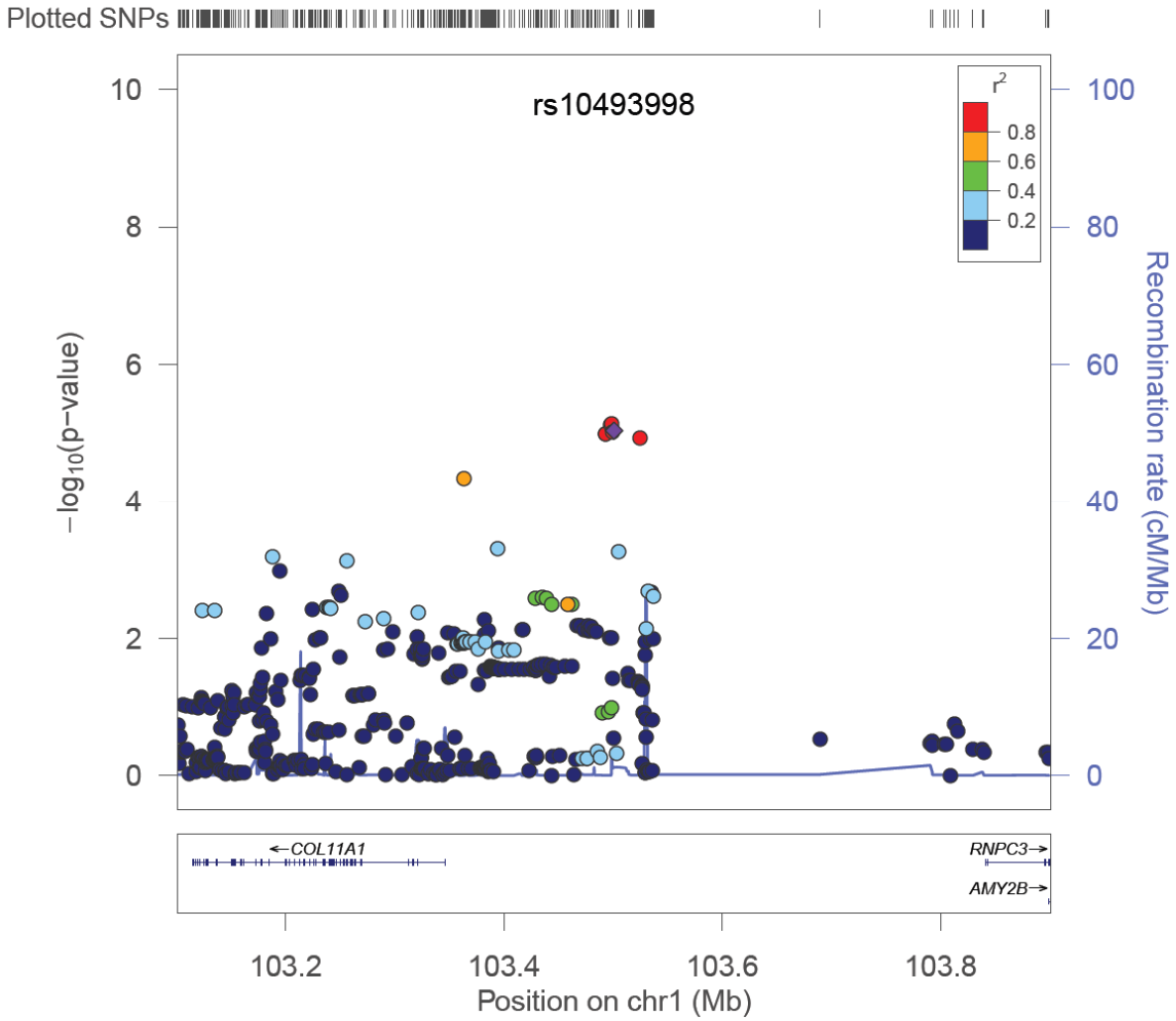


Figure 4.16 Visualization of the 10p15 locus that was marked by rs12260727 ($P=6.0x10^{-7}$) for moderate chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

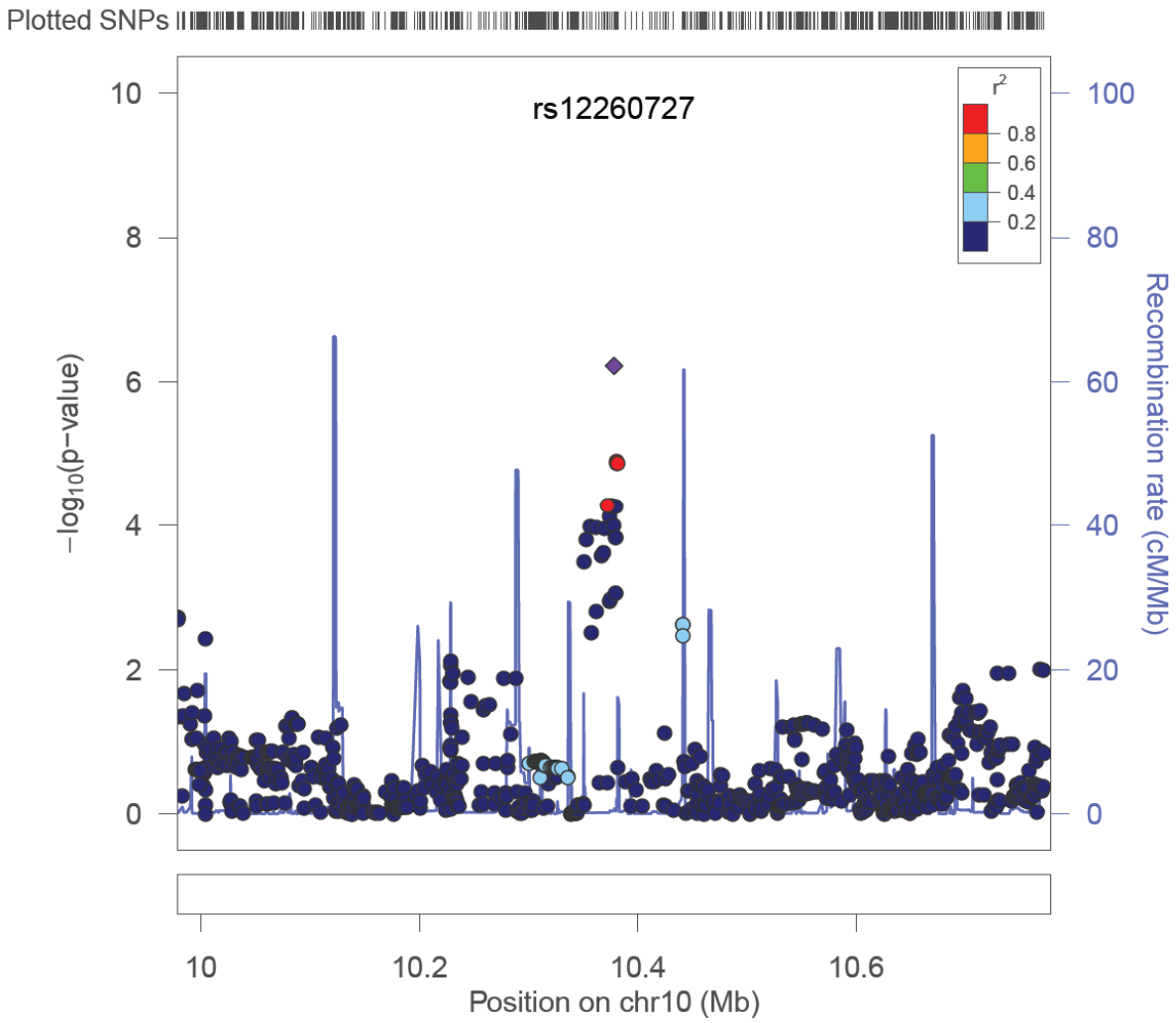


Figure 4.17 Visualization of the 12q15 locus that was marked by rs11615037 ($P=2.6x10^{-5}$ and $P=4.4x10^{-6}$ after adjustment for smoking and diabetic status) for moderate chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

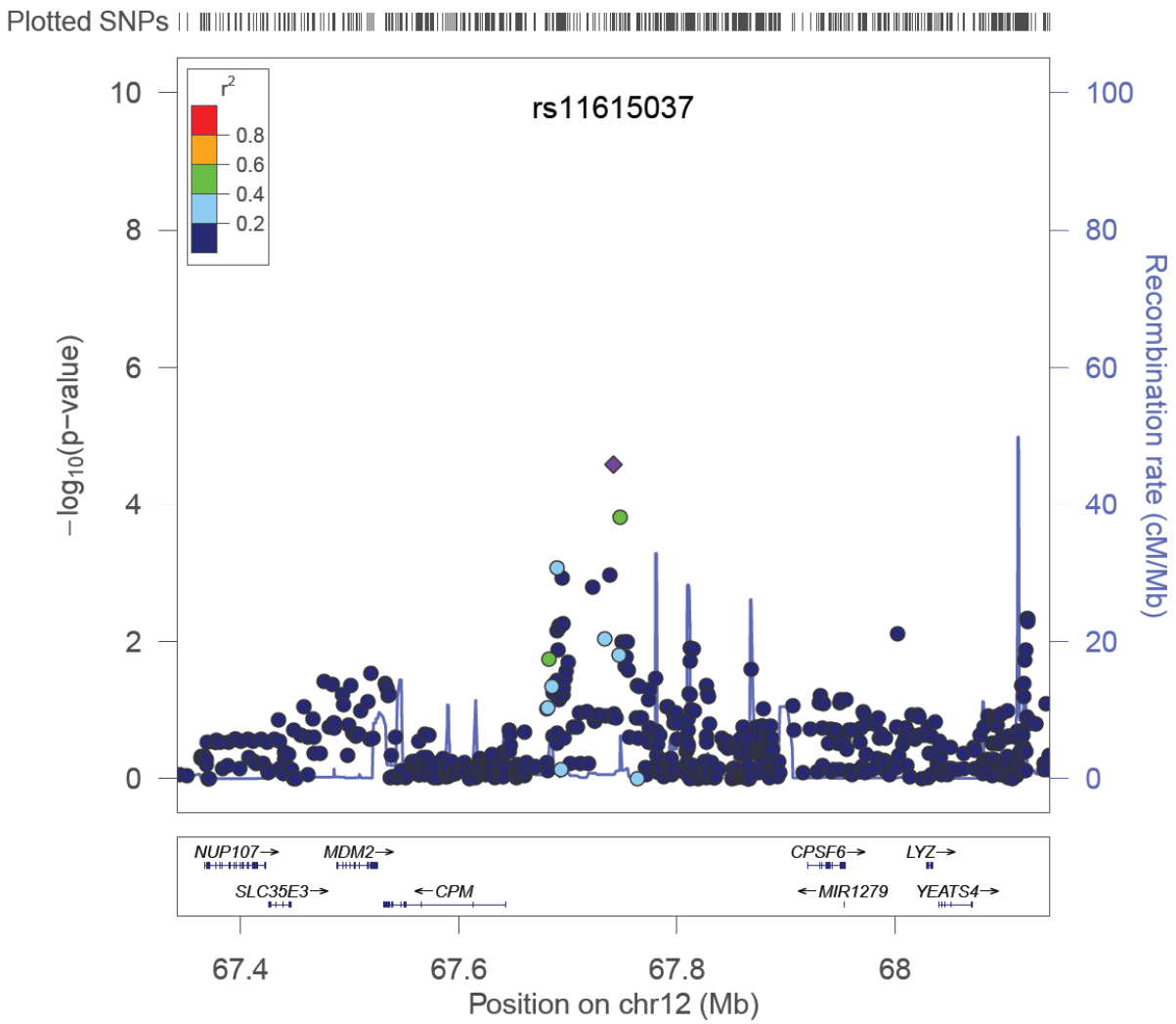


Figure 4.18 Visualization of the 11q22 locus that was marked by rs7120142 ($P=3.9x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

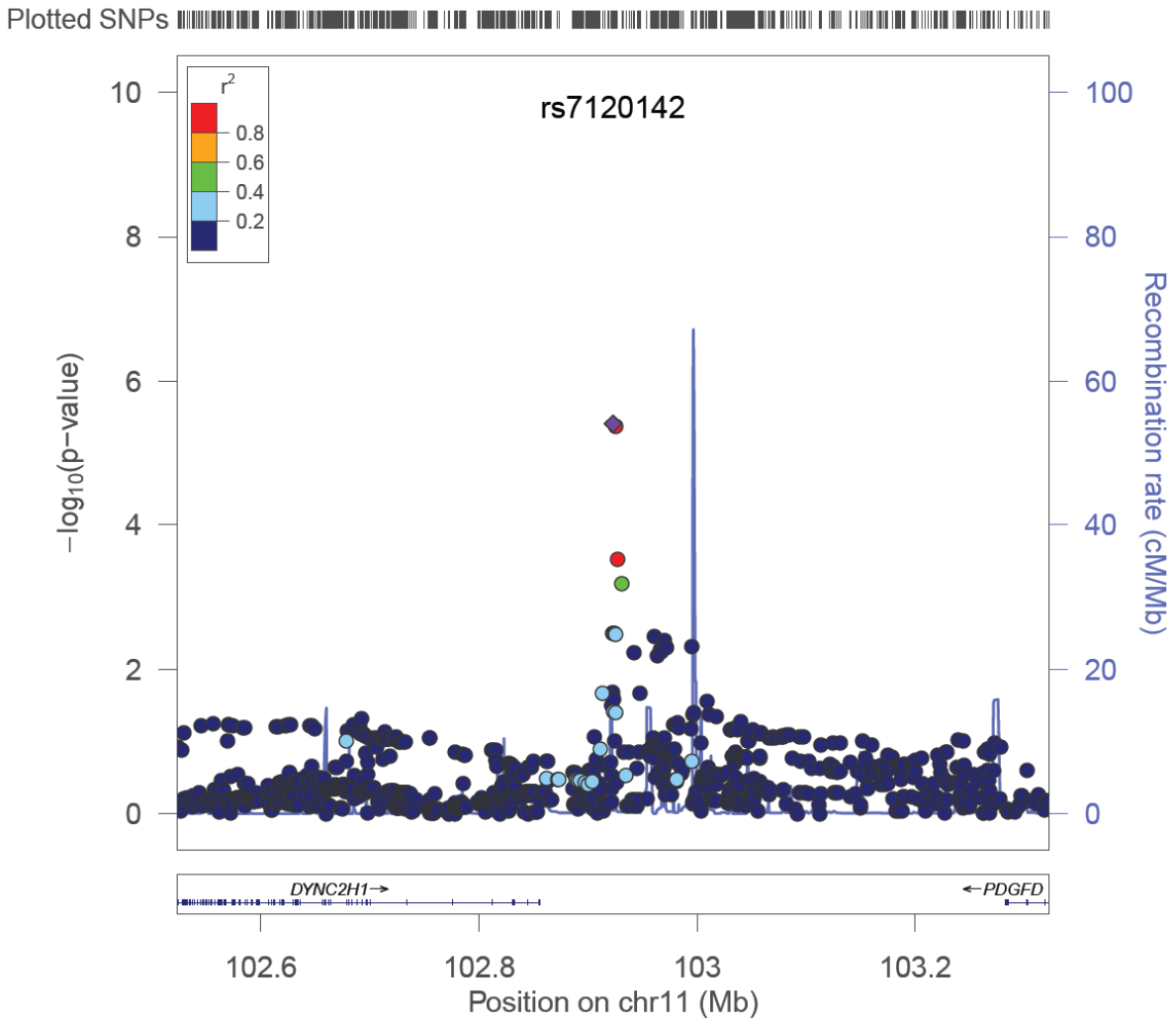
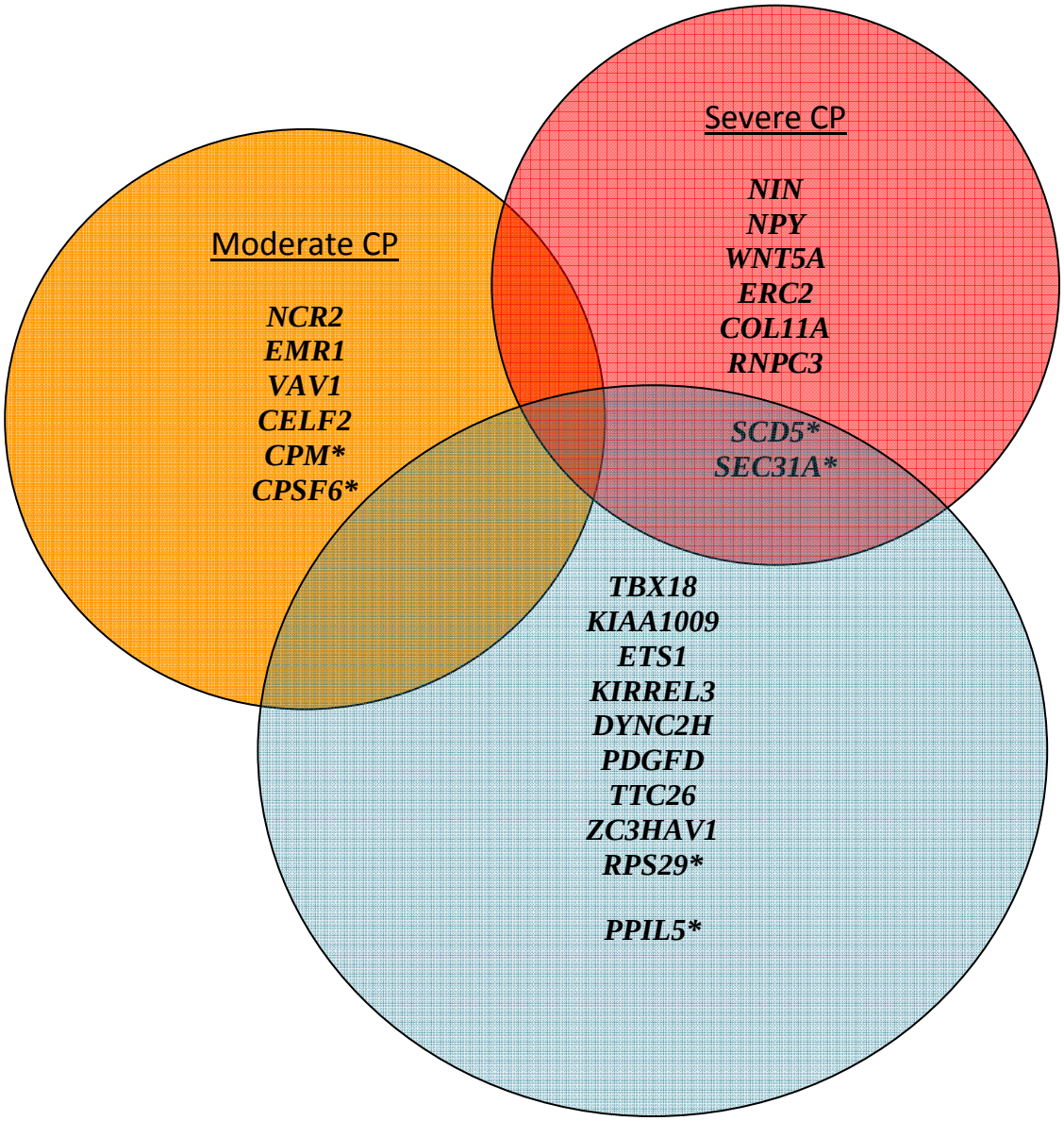


Figure 4.19 Venn diagram representing genes adjacent to identified SNPs for the three chronic periodontitis traits. Genes with asterisk are associated with SNPs that emerged below the $P<5x10^{-6}$ threshold in genetic models adjusted for smoking and diabetic status.



M. MANUSCRIPT #1- SUPPLEMENTAL MATERIAL REFERENCES

1. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904-909.

2. Page RC, Eke PI (2007). Case definitions for use in population-based surveillance of periodontitis. *J Periodontol* 78(7 Suppl):1387-1399.

CHAPTER 5

A. TITLE AND AUTHORS

MANUSCRIPT #2. Genome-wide association study of periodontal pathogen colonization

K. Divaris, K.L. Monda, K.E. North, A.F. Olshan, E.M. Lange, K. Moss, S.P. Barros, J.D. Beck, and S. Offenbacher

B. MANUSCRIPT #2 SUMMARY

Periodontitis is characterized by an oral biofilm pathological shift that contributes to cascade of events leading to periodontal destruction. Host factors modulate the establishment of a dysbiotic oral microbiome, but to-date limited evidence exists with regard to associated genetic risk loci. We conducted a genome-wide association analysis (GWA) among 1020 white participants enrolled in the Atherosclerosis in Communities Study cohort, for whom quantification of eight periodontal pathogens was performed using a "checkerboard" DNA hybridization technique. We examined three traits: "high red" and "high orange" bacterial complex, and "high" *Aggregatibacter actinomycetemcomitans (Aa)* colonization, with "high" being defined as the top quintile of each trait distribution. Genotyping was performed using the Affymetrix 6.0 platform. Imputation to 2.5million markers was based on HapMap II-CEU and a multiple-test correction was applied (threshold of $P<5x10^{-8}$). We detected no genome-wide significant signals. However, thirteen loci including *KCNK1, FBXO38, UHRF2, IL33, RUNX2, TRPS1, CAMTA1* and *VAMP3* provided suggestive evidence ($P<5x10^{-6}$) of association with the measured traits. Further investigations providing replication of these findings and examination of specific associations may lead to an improved understanding of the complex nature of host-biofilm and -bacteria interactions that characterizes states of health and disease.

C. INTRODUCTION

Despite advances in the science and practice of dentistry, periodontal diseases continue to affect adult populations worldwide. The understanding of the pathogenesis of CP has evolved through advances in biology, biochemistry, microbiology, immunology and genetics. Study of the composition and complexity of the oral ecology has recently re-emerged as a focus of research, with investigations of the oral microbiome gaining increased attention (1). To date, over 450 species have been identified in the human microbiome, and its diversity is greater than initially theorized (2, 3). In fact, while harboring of periodontal pathogens is virtually universal, only a small only proportion of individuals develop severe forms of CP.

It is a common ground that an oral ecosystem in harmonious symbiosis with its host will likely be associated with health, whereas departure from this balance is characteristic of disease (4). Specific bacterial species that are implicated in CP have been identified, and in general are commensal and include gram-negative anaerobes. The degree or severity of the host response, which is a determinant of periodontal tissue destruction, has been found to be modulated by several local, systemic, and genetic factors. In this respect, the role of smoking and diabetes, as well as several single nucleotide polymorphisms (SNPs) in immune response-related genes such as interleukin (IL)-1, have been shown to alter the host response and impact the phenotype and clinical course of periodontitis (4,5).

There is a large body of literature in candidate-gene studies investigating the genetics of host inflammatory response, inflammatory mediators and cytokines (5). One recent study found a positive association between an IL-6 polymorphism and harboring of *Aggregatibacter actinomycetemcomitans (Aa),* but that report was based on a small sample of

forty patients (6). No study to our knowledge has carried out a whole-genome investigation for genetic markers of host colonization with periodontal bacteria. To this end, and to add to the knowledge base of the genetic component of periodontitis, the aim of this study was to investigate susceptibility loci for colonization with sub-gingival pathogenic periodontal bacteria using a GWA analysis approach.

D. MATERIALS AND METHODS

Detailed descriptions of the study population, genotyping and imputation, quality control and population stratification procedures are included in the online Appendix. In brief, we conducted a GWA study among 1020 white participants of the Atherosclerosis Risk In Communities (ARIC) longitudinal cohort investigation (7). While ARIC is a study of atherosclerosis, CVD risk factors and outcomes, a complete oral-dental examination took place between 1996 and 1998 during the fourth ARIC visit. As part of the Dental ARIC, apart from a complete clinical examination, sub-gingival microbial plaque samples were collected for a subset of participants (8).

The "checkerboard" DNA-DNA hybridization method (9) was used on plaque samples to measure the extent of sub-gingival colonization with eight periodontal pathogens: *Prevotella intermedia* [American Type Culture Collection (ATCC) 25611], *Campylobacter rectus* (ATCC 33238), *Fusobacterium nucleatum* (ATCC 10953), and *Prevotella nigriscens* (ATCC 33563) (belonging to the "orange" complex); *Porphyromonas gingivalis* (ATCC 33277), *Tannerella forsythia* (ATCC 43037) and *Treponema denticola* (ATCC 35404) ("red" complex); and *Aggregatibacter actinomycetemcomitans* (ATCC 43718). In that method, bacterial levels are expressed as counts relative to established microbial standards. Three dichotomous traits of "high" colonization with "red" complex, "orange" complex, and

Aggregatibacter actinomycetemcomitans were considered for analytical purposes. The two composite phenotypes were derived by the summation of bacterial count for each species belonging to the "red" (n=3) and "orange" (n=4) group, as described above. Because *Porphyromonas gingivalis* is considered the major periodontal pathogen implicated in periodontitis in adult populations, we explored for additional risk loci using its "high" colonization phenotype defined as above, as a separate trait, in exploratory analyses that we present in the Appendix. Various approaches in defining the bacterial colonization profiles have been previously used, including summations of the absolute microbial counts (10), tertile-categorization (11), and five-level categorization of log10-transformed counts (12). For the present investigation, we defined a "high" colonization trait as the top quintile (20%) of each trait's distribution. The rationale for the selection of this phenotype is based on the fact that the "checkerboard" semi-quantitative method has a lower detection threshold of ~10^4 and reduced precision in the lower end of the distribution, whereas a individuals with "high" bacterial colonization profile may be those with reduced or impaired host and at high risk for periodontal tissue destruction.

Participants' DNA was extracted from blood samples drawn from an antecubital vein, and genotyping was performed with the Affymetrix Genome-Wide Human SNP Array 6.0 chip. The platform offers 906,600 markers for SNPs. Following rigorous quality control procedures, imputation to 2.5million markers was performed using 669,450 SNPs and MACH v1.0.16 (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html), based on HapMap Phase II CEU build 36. For analytical purposes, the dichotomous traits of "high colonization" phenotypes were entered in three logistic regression models assuming multiplicative (log-additive) allelic effects. The models included age, sex, examination center

and ten principal components from the population stratification analysis as covariates. A correction for multiple comparisons assuming 1 million independent tests was applied and a threshold of genome-wide statistical significance was set at $P<5x10^{-8}$. An additional arbitrary threshold of $P<5x10^{-6}$ was set to prioritize and investigate loci with suggestive evidence of association. All GWA analyses were performed with the ProbABEL software (13). SNP annotations were performed with WGAViewer ver.1.26l (14), Snipper ver.1.2 (http://csg.sph.umich.edu/boehnke/snipper/), and loci visualizations with LocusZoom ver.1.1 (15) and Haploview ver.4.2 (16). Reporting of genes was based on the "HUGO Gene Nomenclature" naming convention (www.genenames.org).

E. RESULTS

The sample's descriptive information is presented in Table 1. The participants had mean age of 63 years and approximately even gender distribution. "High" bacterial colonization profiles were found in all groups of periodontal diagnosis according to the CDC classification; however, the prevalence of "high red" colonization cases was more than double among participants with severe periodontitis compared to those with mild or no disease. Similar, but less pronounced associations were noted for "orange" complex and *Aa*. The bacterial counts of "high colonization cases" were [median (interquartile range); range)] "red" complex bacteria—[64725 (108516); 21943-3894605), "orange" complex bacteria—[304155 (483373); 111926-$1.5x10^{7}$), and *Aggregatibacter actinomycetemcomitans*—[17309 (20034); 7520-410713].

Of the 2178777 examined SNPs, none had $P<5x10^{-8}$. However, 53 had a $P<5x10^{-6}$ and thus were prioritized for further investigation. Lambda inflation factors for three traits

were: “red”—1.040, “orange”—1.045 and *Aa*—1.032. The corresponding Q-Q plots are presented in Figure S1 in the Supplemental Material. Upon inspection of the prioritized SNPs, there were five loci that emerged for “red”, three loci for “orange” and five loci for *Aa* colonization (Figure 1). Of those, one locus on 1q42 was shared for the “red” complex and *Aa*. Graphical representations of the genomic areas adjacent to six of these loci are presented in Figure 2.

The strongest signal with regard to “red” complex colonization (Table 1) was produced by rs11800854 in the 1q42 locus [$P=2.8x10^{-7}$; OR=12.3 (95% CL=3.7, 41.3); MAF in HapMap-CEU (MAF-CEU): 0.067] in the promoter region (30Kb upstream) of *KCNK1* and adjacent to *KIAA1804*. The common [G] allele showed 3% enrichment among “high colonization cases” for both “red” complex and *Aa*. Another locus in chromosome 1p22 was marked by rs12032672 ($P=9.6x10^{-7}$), ~500Kb upstream of *PKN2*. Rs10043775, in LD with multiple markers in the 5q33 locus, represents a missense change in the *FBXO38* gene (resulting in [Pro]→[Arg] substitution, predicted as ‘benign’ by PolyPhen-2) and provided the strongest signal in that locus ($P=2.4x10^{-6}$), also adjacent to *HTR4*. A high LD area in 9p24 including the *UHRF2*, *GLDC*, *TPD52L3* and *IL33* genes is marked by rs16924631 [intronic to *UHRF2*; $P=3.2x10^{-6;}$ OR=2.29 (95% CL=1.61, 3.24); MAF-CEU: 0.275], of which the [C] risk allele showed almost 10% enrichment among “cases”. An intronic variant (rs10010758) of the *TBC1D1* gene, adjacent (24Kb) to *PTTG2* provided the strongest signal in the 4p14 locus [$P=3.7x10^{-6}$; OR=1.91 (95% CL=1.25, 2.21)]. In the Appendix we present results of the exploratory analysis for *Pg* “high” colonization, which revealed three loci with $P<5x10^{-6}$ including *OTOF*, *C2Orf70*, *CIB4*, *DAB2IP*, *TTLL11* and *AKNRD3*.

The common allele of rs1932040 showed 9% enrichment and provided the strongest association signal with "high orange" bacterial colonization [$P=1.3x10^{-6}$, OR=2.47 (95% CL=1.67, 3.65)], marking an intergenic area between *RUNX2* and *CLIC5* on the 6p21.1 locus. A low recombination area on 8q23, adjacent to *TSPS1* (1.3Mb) and *CSMD3* (672Kb) is marked by multiple alleles, of which rs9942773 provided the strongest signal ($P=1.9x10^{-6}$) and 10% enrichment among "cases". A variant intronic to *CAMTA1* [rs1616122; $P=4.9x10^{-6}$; OR=1.85 (95% CL=1.41, 2.42)] marks the 1p36.2 locus.

The 1q42 locus that was identified for "red" bacteria also provided the third strongest association signal for *Aa* [rs11800854; $P=4.0x10^{-6}$; OR=8.12 (95% CL=2.73, 24.11)]. The common [T] allele of rs11621969 was also nominally associated with high *Aa* colonization ($P=9.4x10^{-7}$) and was adjacent to *FOS* and *JPD2* in 14q24. The rare [G] allele of rs1970525 was more than twice as prevalent among "cases" (0.118 versus 0.054 among "non-cases"), provided the strongest signal in the 10q23 locus [$P=3.8x10^{-6}$; OR=2.89 (95% CL=1.85, 4.52); MAF-CEU: 0.045)] , and represents a nonsense-mediated decay transcript variant in the *GRID1* gene. Rs9287989 is adjacent to *KIAA1715* (30Kb) and *EVX2* (227Kb) and marks the 2q31 locus ($P=4.4x10^{-6}$). An intronic variant of *ODZ2* (rs6885116) provided the strongest signal in the locus 5q35 [$P=1.4x10^{-6}$; OR=2.57 (95% CL=1.76, 3.74); MAF-CEU: 0.084], showing 9% enrichment among "cases". Rs1800795, that was found to be significantly associated with *Aa* colonization in the recent study of Nibali and colleagues (6) did not show any important association in this GWA scan [P=0.34, OR=1.12 (95% CL=0.88, 1.43)].

F. DISCUSSION

This study is the first report of a genome-wide association analysis investigating risk loci for colonization with pathogenic periodontal bacteria. Although limited by the sample size, this investigation explores a novel phenotype and benefits from a comprehensive quantitative phenotypical characterization. Upon replication or validation, these findings have the potential to unveil pathways and mechanisms that direct the host's symbiosis with healthy microflora that if altered may predispose for states of disease. In fact, several of the risk loci identified in this study may offer promising leads for further exploration and mechanistic studies.

The 1q42 locus and rs11800854, which emerged as a common risk marker for both "red" complex and *Aa* colonization is in the promoter region of *KCNK1*, a gene that encodes a potassium channel protein and has been linked to cardiac outcomes (17). The prioritized SNPs at the 9p24 locus marked an area with low recombination rate that includes the *UHRF2, TPD52L3* and *IL33* genes. *IL33*, as other members of the interleukin family, has important roles in immunity and inflammation and has been suggested to function as an "alarmin", alerting the immune system to endogenous trauma such as physical stress or infection (18).

RUNX2, a gene in the 6p21.1 locus that emerged due to the association of rs1932040 with "orange" complex colonization encodes a transcription factor that is essential for osteoblastic differentiation and skeletal morphogenesis. Mutations in this gene have been associated with the cleidocranial disorder syndrome, which has multiple and severe oral manifestations (19). Experimental evidence shows that *RUNX2* is involved in the inhibition of MMP-13

expression, which appears to be involved in periodontitis, as well as osteoblastic and osteoclastic activity (20). Another locus that was identified as associated with “orange” colonization was 1p36.2 with the “top” SNP rs1616122 being intronic to the *CAMTA1* gene. *CAMTA1* has been reported as a candidate gene for type II diabetes risk (21). *VAMP3* is adjacent (310Kb upstream) to the marked locus and may also be plausible gene candidate, as it has been shown to regulate podosome organization in macrophages, and thus mediate their adhesion, spreading and migration (22).

With regard to association results with *Aa* “high” colonization, the strongest signal was produced by the 14q24 locus, where the *FOS* and *JDP2* genes are located. *FOS* was recently reported as part of a novel mechanism of RANKL expression in T cells (23), which may constitute an important link with immune responses, as *FOS* levels have been shown to increase by lippopolysaccharide infusion *in vivo* (24). *JDP2* (85Kb downstream of the top SNP of the 14q24 locus) is a transcription factor that has been associated with the maintenance of Epstein-Barr virus latency (25), which has been implicated in the pathogenesis of chronic periodontitis via inhibition of oral bacteria-induced macrophage activation and phagocytosis (26).

Socransky (27) and other investigators have described how organisms such as *Pg*, *Tannerella forsythia*, *Treponema denticola*, *Aa* and other species found in dental plaque, organize themselves in complex communities collectively called “biofilm”, and interact with each other and with the host to result in different states of health and disease. While the presence of a highly organized biofilm appears a unique feature of dental plaque-induced diseases including caries and periodontitis, several other chronic conditions such as tuberculosis, rheumatic fever, syphilis, gastrointestinal ulcers and Lyme disease share a

bacterial colonization etiologic component (28). Consequently, an increased understanding of the genetic underpinning of interactions between the host and exogenous or symbiosing bacterial communities has the potential to advance the state of knowledge in periodontitis, but also other chronic inflammatory diseases (29).

The findings of the present investigation, if replicated, have the potential to add to the knowledge base of oral microbiome, host-biofilm, host-bacteria interactions, and more. We acknowledge that the study of periodontal pathogen colonization phenotypes is rare and this limits the replication options of our findings. However, the consideration of specific microbial factors as a distinct exposure in investigations of periodontal, oral and systemic health is consistent with the paradigm of "periodontal medicine" and may provide novel insight in the oral-systemic diseases connection. Although the prevention and treatment of periodontitis is an obvious goal, the links of CP with of other systemic conditions and the "common theme" of pathogenic ecological shift in other diseases, provide opportunities for even greater impact.

MANUSCRIPT #2 REPORTING OF FUNDING SOURCES

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HSN268201100012C), R01HL087641, R01HL59367, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C; and by a grant from the National Institute of Environmental

Health Sciences (P30ES010126). The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The authors declare no potential conflicts of interest with regard to the authorship and publication of this article.

G. TABLES

Table 5.1 Distribution of subjects' demographic characteristics and bacterial "high colonization" (defined as the top quintile of the distribution) profile (with "Red", "Orange" complex, and *Aggregatibacter actinomycetemcomitans*) overall, and stratified by periodontal diagnosis (CDC/AAP classification) among the Dental ARIC study participants (n=1020)

	Total n (column %)	Healthy/mild n (column %)	Moderate n (column %)	Severe n (column %)
n (row %)	1020 (100)	416 (41)	415 (41)	189 (19)
Sex				
Females	478 (47)	246 (59)	178 (43)	54 (29)
Males	542 (53)	170 (41)	237 (57)	135 (71)
Age (years; mean, standard deviation)	63.2 (5.7)	62.3 (5.5)	63.6 (5.7)	64.0 (5.9)
"High" bacterial colonization (n, % of column)				
"Red" complex	203 (20)	58 (14)	84 (20)	61 (32)
"Orange" complex	201 (20)	73 (18)	72 (17)	56 (30)
A. *actinomycetemcomitans*	204 (20)	75 (18)	81 (20)	48 (25)

Table 5.2 Genome-wide association analysis results of the high colonization traits (highest quintile of the distribution versus the other four; quantified with DNA-DNA "checkerboard" hybridization) for "Red", "Orange" complex and Aggregatibacter actinomycetemcomitans, among the white participants of the Dental ARIC study (n=1020). Single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF-HapMap II CEU) of ≥5% and associated $P<5x10^{-6}$. The SNP with the lowest P-value per locus is presented; additional prioritized SNPs in each locus are presented in the footnote, along with corresponding R^2 (based on 1000 genomes pilot 1 release) with the top SNPs.

Chr.	SNP	Position Build36	ca[1]	nca[2]	MAF (HapMap II-CEU)	Closest gene(s) and position or distance	Risk allele frequency low/ high colonization	P value	Odds ratio (95% CL)
"Red" complex									
1q42	rs11800854[3]	231786607	A	G	[A] 0.068	*KCNK1* (30Kb)	[G] 0.947/0.978	$2.8x10^{-7}$	12.3 (3.7, 41.3)
1p22	rs12032672	88398224	A	C	[C] 0.350	*PKN2* (524Kb)	[C] 0.332/0.446	$9.6x10^{-7}$	1.99 (1.50, 2.62)
5q33	rs10043775[4]	147785313	C	T	[C] 0.274	*FBXO38* (missense change)[6] /*HTR4* (19Kb)	[T] 0.703/0.791	$2.4x10^{-6}$	2.06 (1.51, 2.83)
9p24	rs16924631[5]	6476308	C	G	[C] 0.142	*UHRF2* (non-coding transcript variant) /*GLDC* (46kb)	[C] 0.122/0.219	$3.2x10^{-6}$	2.29 (1.61, 3.24)
4p14	rs10010758	37614913	C	T	[C] 0.275	*TBC1D1*(intron variant)/ *PTTG2* (24Kb)[7]	[C] 0.291/0.384	$3.7x10^{-6}$	1.91 (1.45, 2.51)
"Orange" complex									
6p21.1	rs1932040[8]	45804766	A	G	[A] 0.142	*CLIC5* (169Kb)/ *RUNX2* (178Kb)	[G] 0.808/0.896	$1.3x10^{-6}$	2.47 (1.67, 3.65)
8q23	rs9942773[9]	115190203	A	C	[C] 0.283	*CSMD3* (672Kb)/ *TRPS1* (1.3Mb)	[A] 0.703/0.803	$1.9x10^{-6}$	2.07 (1.51, 2.82)
1p36.2	rs1616122	7444172	C	T	[C] 0.482	*CAMTA1* (intron variant)/*VAMP3* (310Kb)	[T] 0.506/0.624	$4.9x10^{-6}$	1.85 (1.41, 2.42)
"*Aggregatibacter actinomycetemcomitans*"									
14q24	rs11621969	74883781	C	T	[C] 0.167	*FOS* (65Kb)/ *JDP2* (85Kb)	[T] 0.789/0.885	$9.4x10^{-7}$	2.46 (1.68, 3.62)
10q23	rs1970525[10]	87624904	C	G	[G] 0.045	*GRID1* (NMD[12] transcript variant) /*MI346*/ *WAPAL* (560Kb)	[G] 0.054/0.118	$3.8x10^{-6}$	2.89 (1.85, 4.52)
1q42	rs11800854[11]	231786607	A	G	[A] 0.067	*KCNK1* (30Kb)/ *KIAA1804* (199Kb)	[G] 0.947/0.978	$4.0x10^{-6}$	8.12 (2.73, 24.11)
2q31	rs9287989	176425987	C	T	[T] 0.433	*KIAA1715* (73Kb)/ *EVX2* (227Kb)/ *EXTLP2*[13] (10Kb)	[C] 0.484/0.605	$4.4x10^{-6}$	1.80 (1.39, 2.33)
5q35	rs6885116	167576123	A	G	[G] 0.084	*ODZ2* (intron variant)/ *WWC1* (76Kb)	[G] 0.078/0.169	$1.4x10^{-6}$	2.57 (1.76, 3.74)

1: coded allele
2: non-coded allele
3: Additional SNPs in locus with $P<5x10^{-6}$: rs6682365 (R^2=1.00)
4: Additional SNP in locus with $P<5x10^{-6}$: rs10068216 (R^2=1.00), rs10072051 (R^2=1.00), rs17108251 (R^2=1.00), rs10044061 (R^2=1.00), rs4349707 (R^2=1.00), rs10477376 (R^2=1.00), rs9325095 (R^2=1.00), rs10041283 (R^2=1.00), rs9325097 (R^2=0.87), rs3734120 (R^2=1.00), rs4574533 (R^2=0.87), rs4274967 (R^2=1.00), rs4274968 (R^2=1.00), rs6884076 (R^2=1.00), rs9325098 (R^2=1.00)
5: Additional SNP in locus with $P<5x10^{-6}$: rs11795355 (R^2=1.00), rs7876000 (R^2=1.00), rs10975603 (R^2=0.93), rs16924626 (R^2=0.93), rs16924624 (R^2=1.00), rs10975605 (R^2=0.93), rs10115883 (R^2=0.93), rs10122116 (R^2=0.93)
6: T>C – Ser>Pro, 35b from the exon boundary
7: R^2=0.29 with rs6811863 which is a missense change in *PTTG2*: G>C – [Arg]→[Pro]
8: Additional SNPs in locus with $P<5x10^{-6}$: rs12525547(R^2=0.93), rs9349326(R^2=0.93), rs16873698(R^2=0.93)
9: Additional SNP in locus with $P<5x10^{-6}$: rs10089040 (R^2=1.00), rs9942776 (R^2=1.00), rs10086149 (R^2=1.00), rs7845243 (R^2=0.87), rs10105817 (R^2=1.00), rs7006291 (R^2=1.00), rs11779159 (R^2=1.00), rs11783996 (R^2=1.00), rs10098056 (R^2=1.00), rs7018200 (R^2=0.92)
10: Additional SNPs in locus with $P<5x10^{-6}$: rs4325261 (R^2=1.00)
11: Additional SNP in locus with $P<5x10^{-6}$: rs6682365 (R^2=1.00)
12: nonsense-mediated decay
13: *EXTLP2* is a pseudogene

H. FIGURES

Figure 5.1 Manhattan plot of the GWAS results (-log10 P-values of the ~2.5 million examined SNPs arranged by chromosome) for "high red" complex bacterial colonization among the 1020 white participants of the Dental ARIC Study cohort.

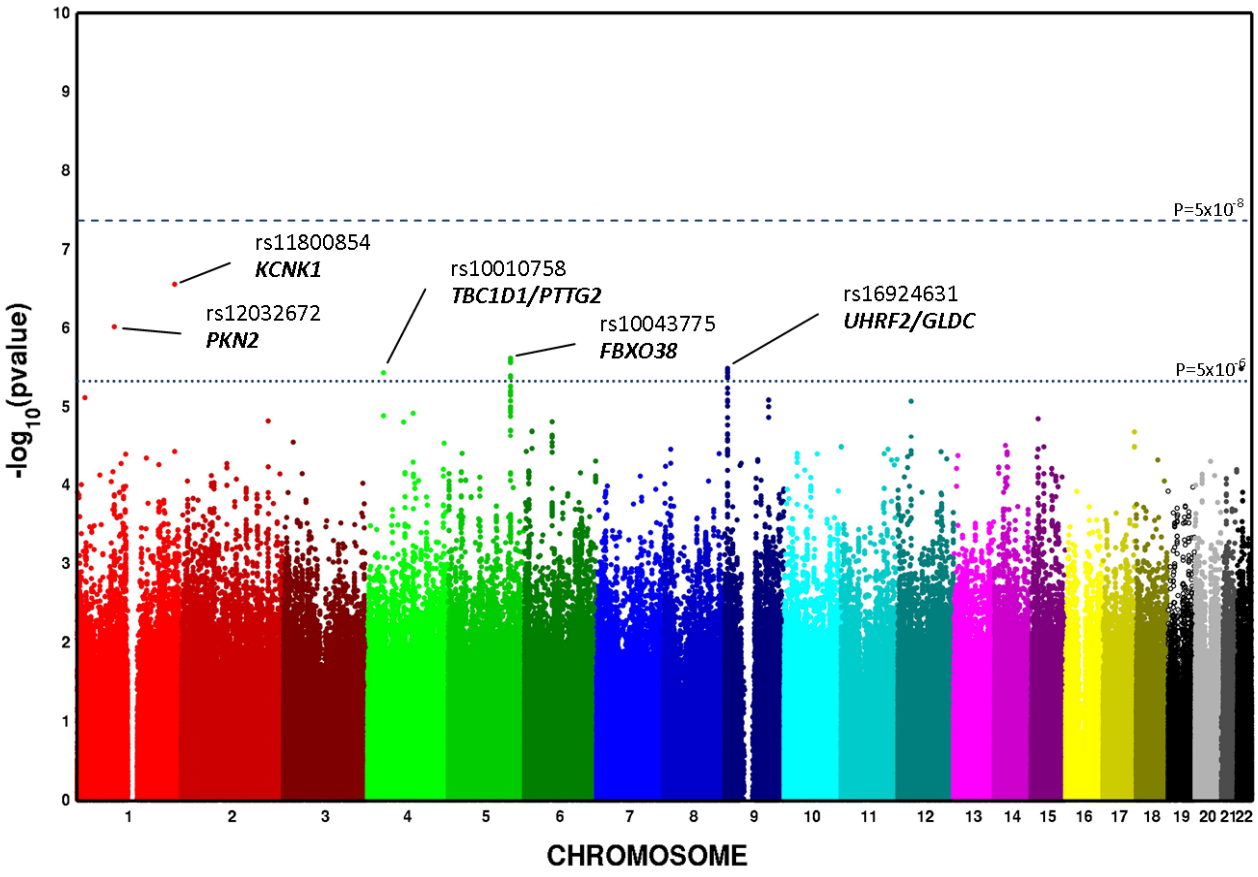


Figure 5.2 Manhattan plot of the GWAS results (-log10 P-values of the ~2.5 million examined SNPs arranged by chromosome) for "high orange" complex bacterial colonization among the 1020 white participants of the Dental ARIC Study cohort.

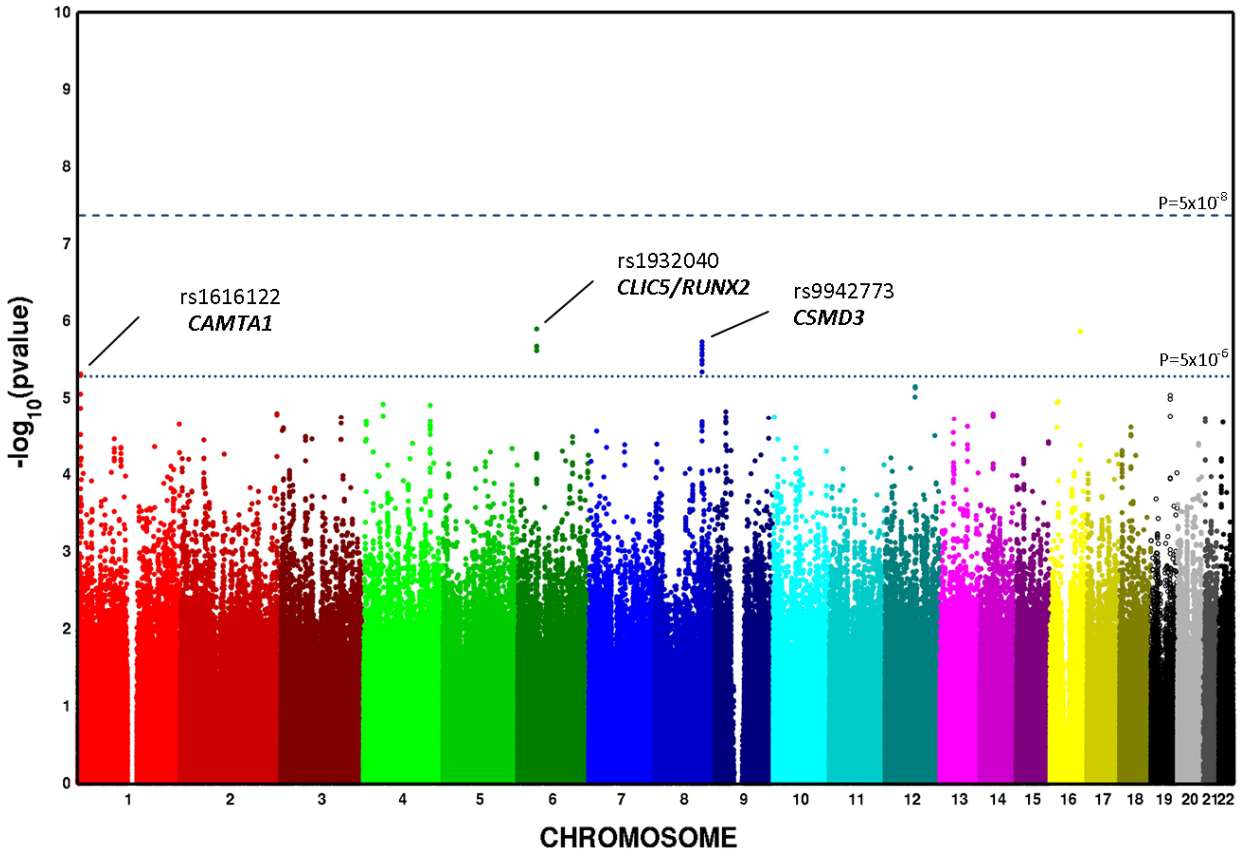


Figure 5.3 Manhattan plot of the GWAS results (-log10 P-values of the ~2.5 million examined SNPs arranged by chromosome) for "high" *Aggregatibacter actinomycetemcomitans* bacterial colonization among the 1020 white participants of the Dental ARIC Study cohort.

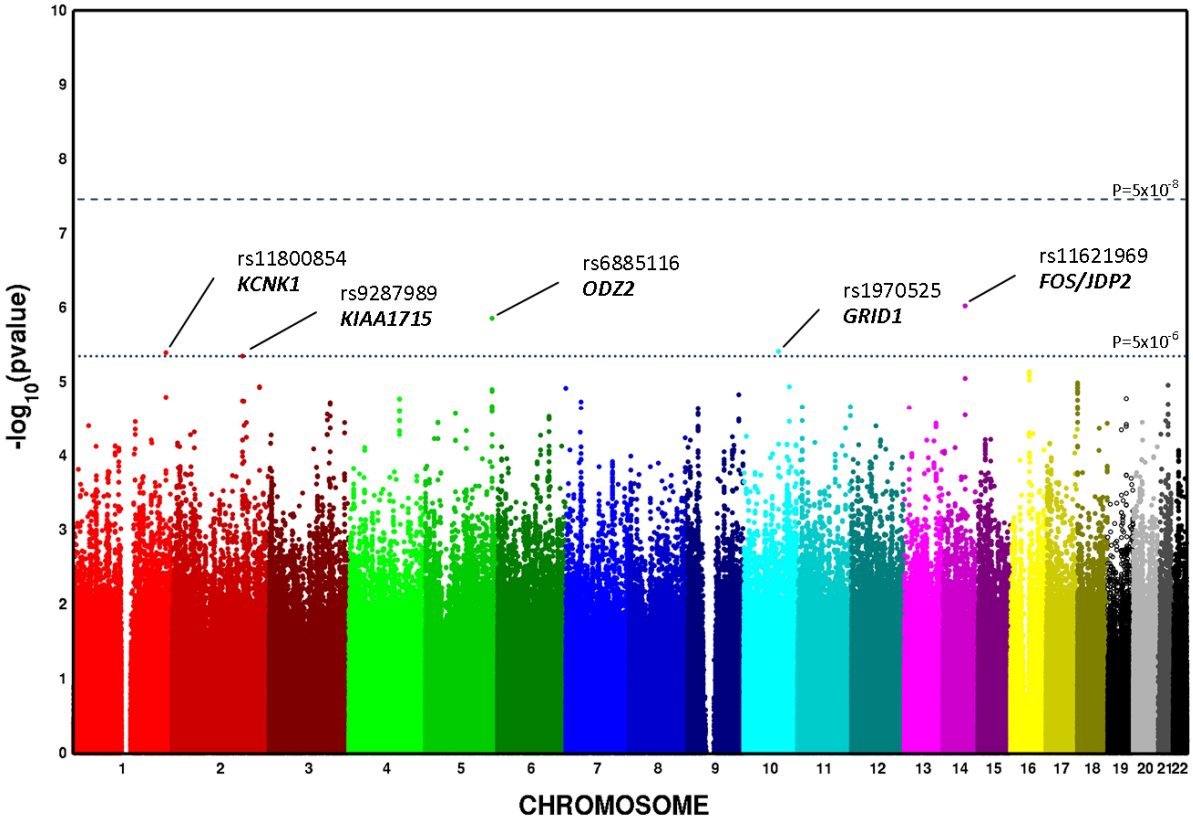


Figure 5.4 Visualization of the 1p22 locus that was marked by rs12032672 ($P=9.6x10^{-7}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

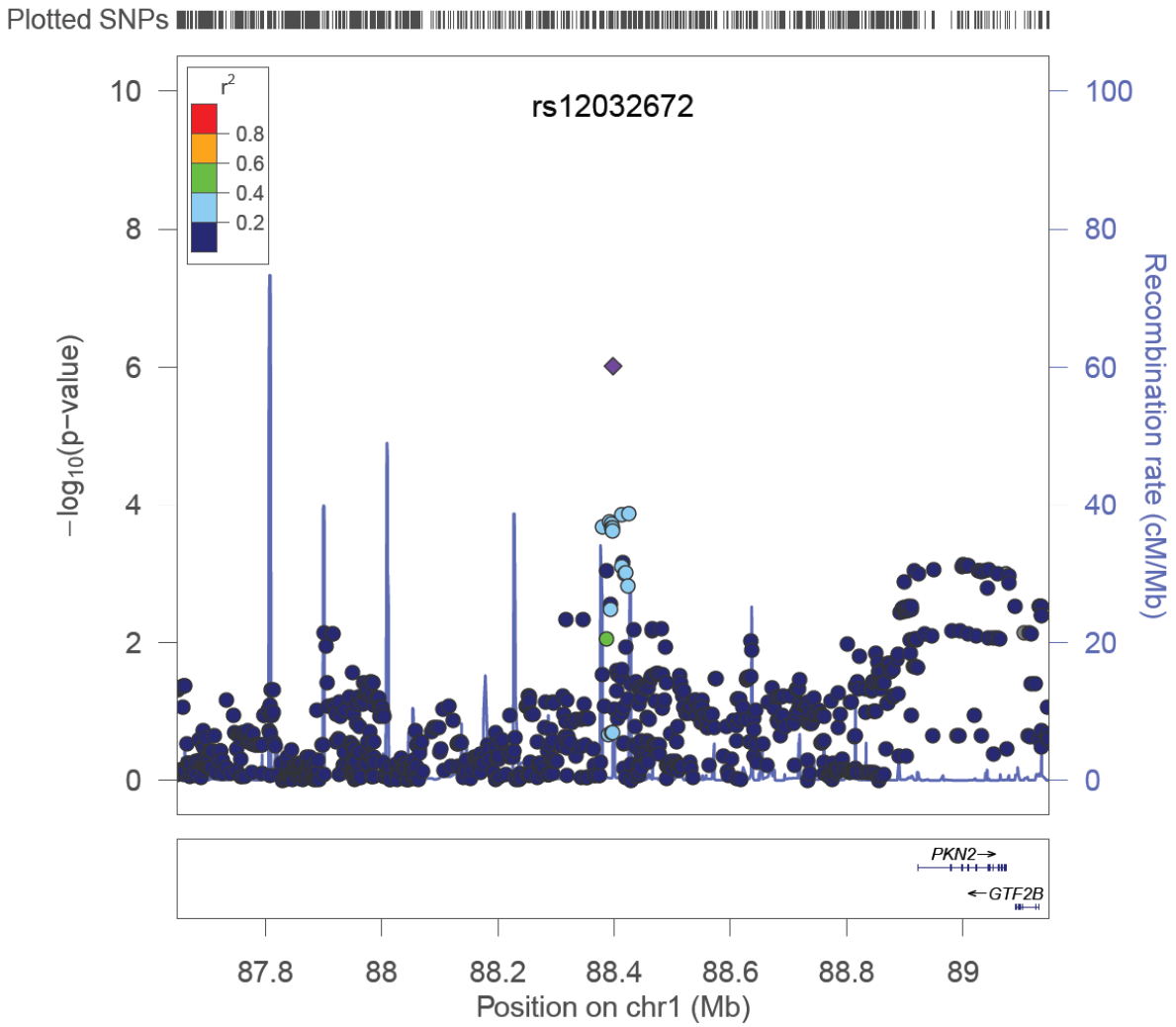


Figure 5.5 Visualization of the 5q33 locus that was marked by rs10043775 ($P=2.4x10^{-6}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

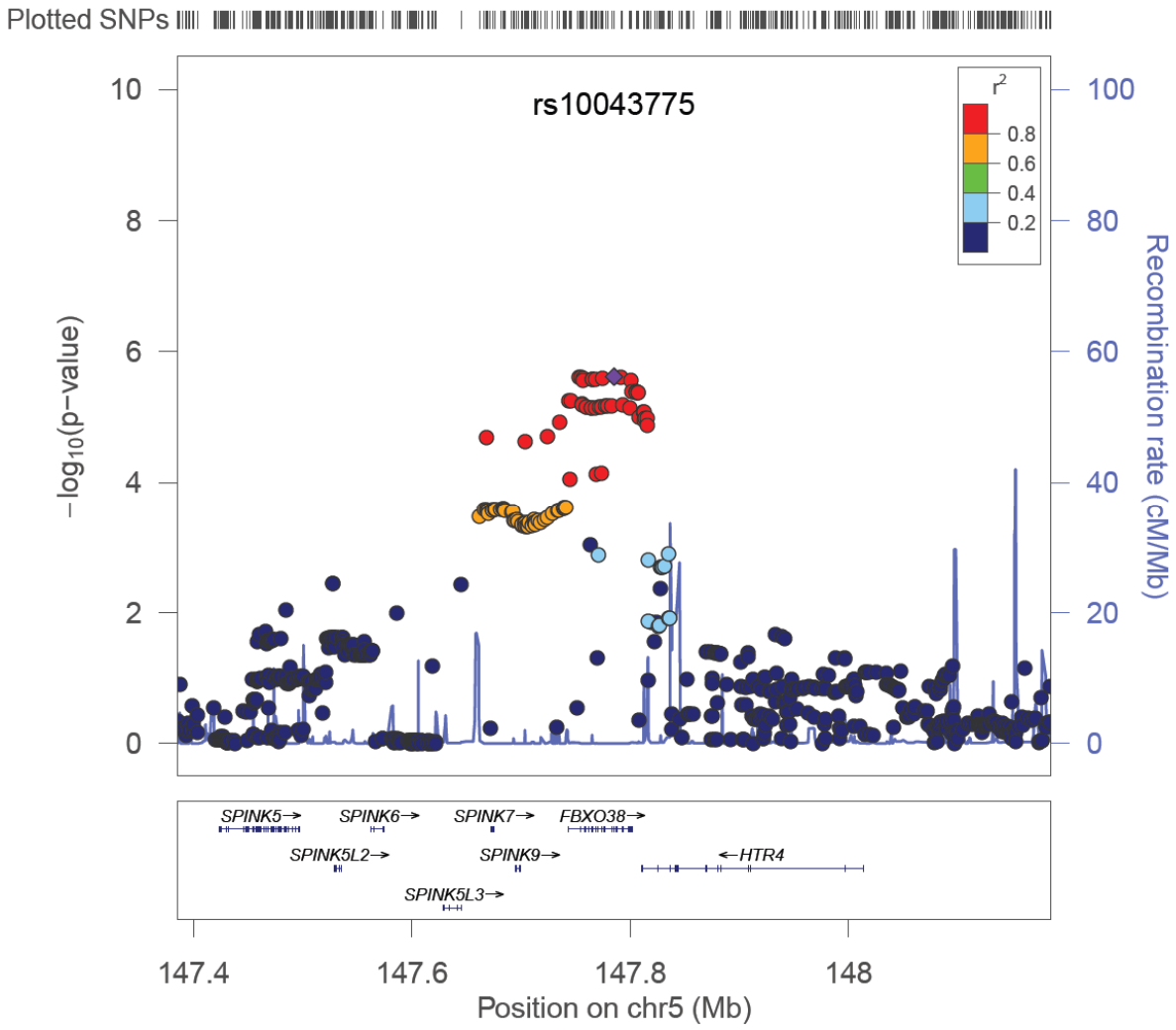


Figure 5.6 Visualization of the 9p24 locus that was marked by rs16924631 ($P=3.2x10^{-6}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

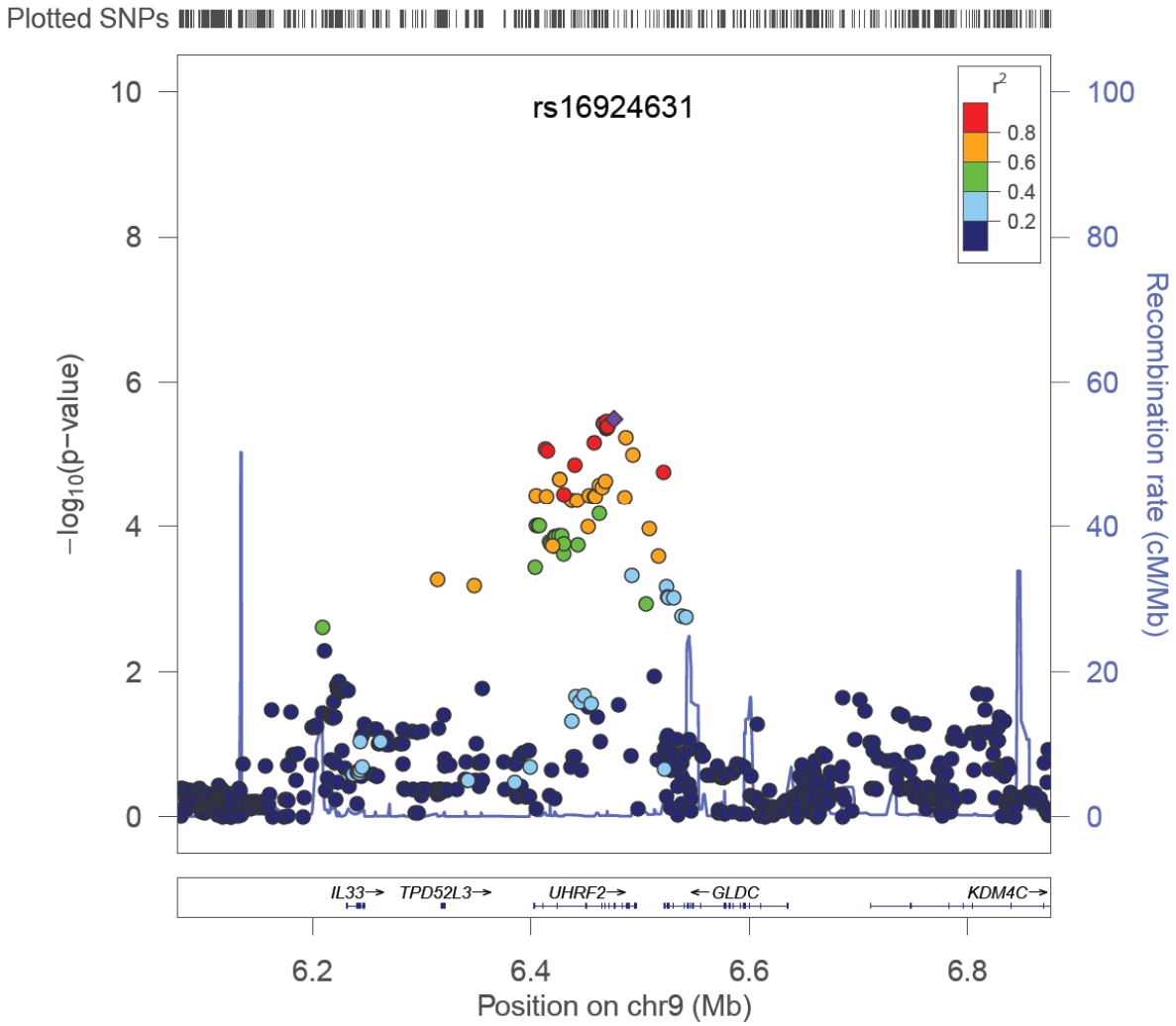


Figure 5.7 Visualization of the 4p14 locus that was marked by rs10010758 ($P=3.7x10^{-6}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

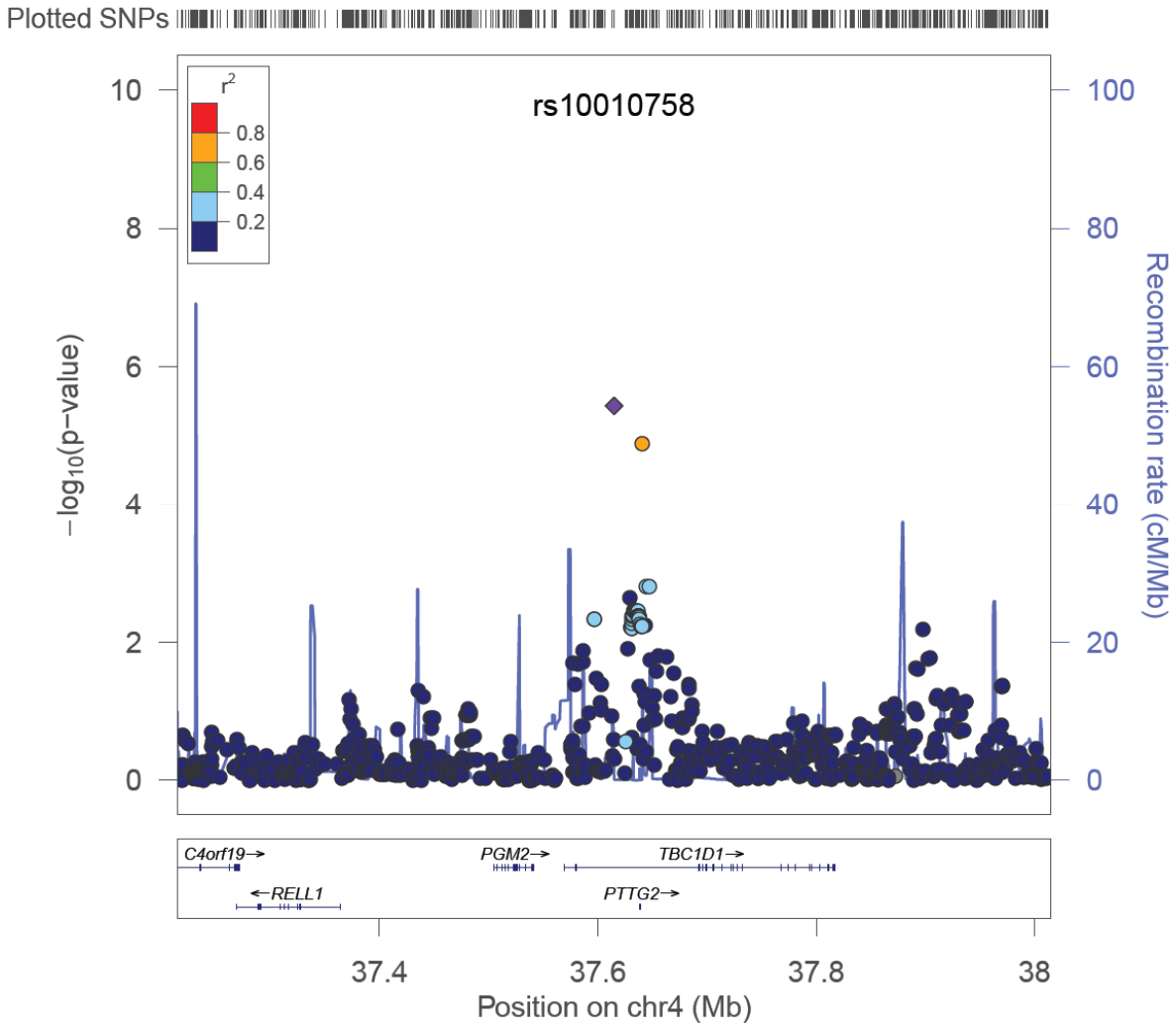


Figure 5.8 Visualization of the 1p36.2 locus that was marked by rs1616122 ($P=4.9x10^{-6}$) for "high orange" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

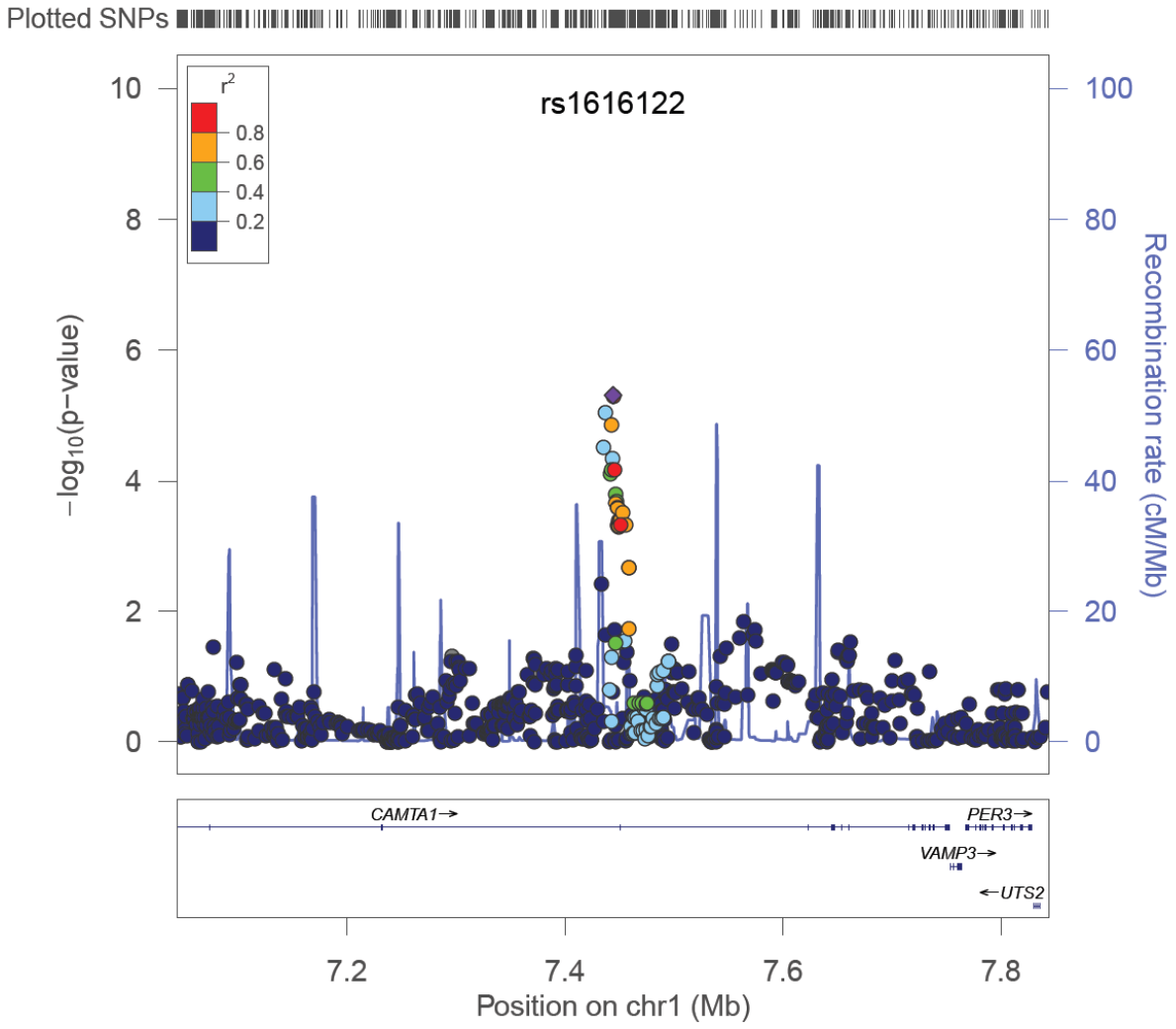
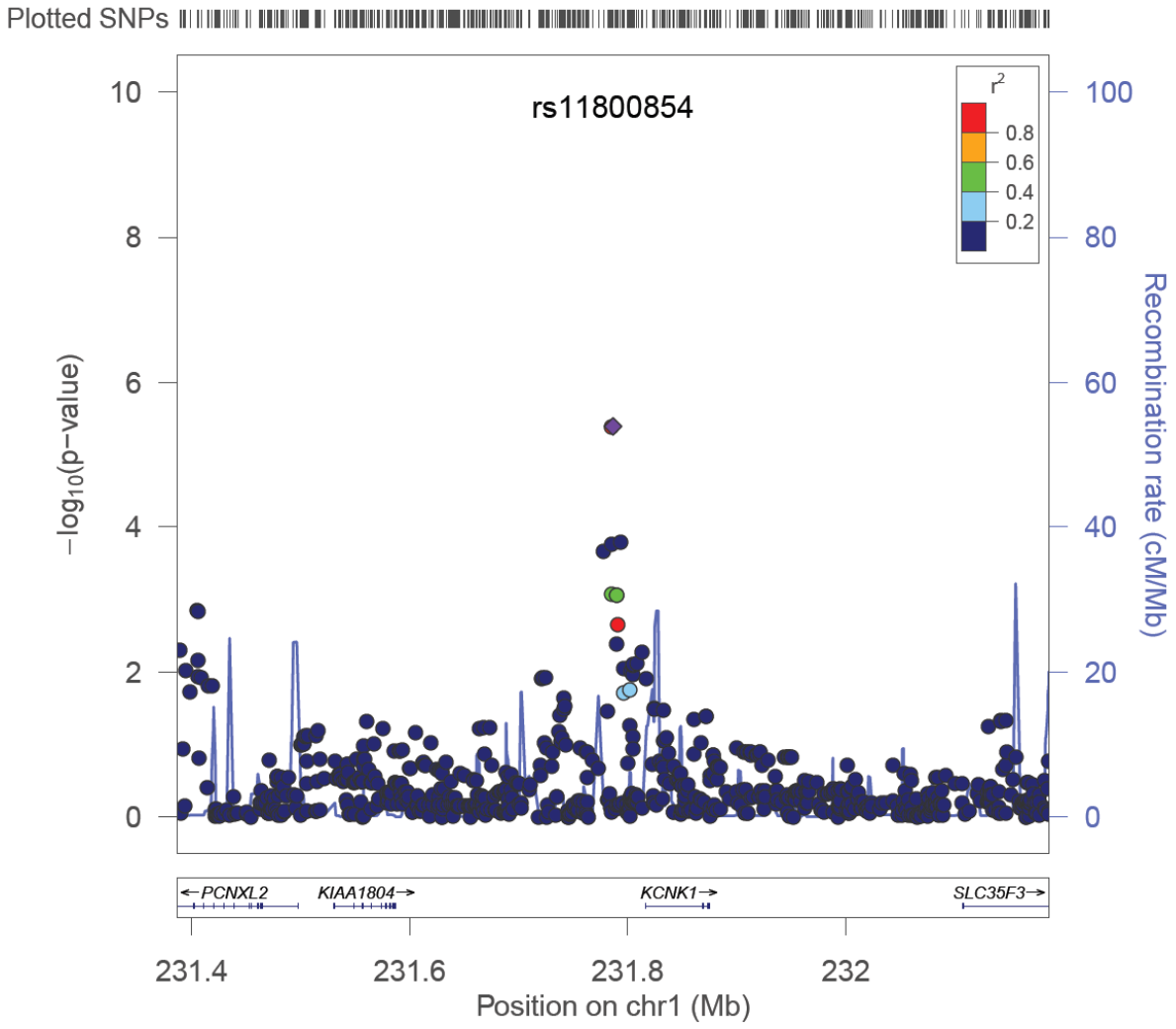


Figure 5.9 Visualization of the 1q42 locus that was marked by rs11800854 ($P=4.0x10^{-6}$) for "high" *Aggregatibacter actinomycetemcomitans* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).



I. REFERENCES

1. Avila M, Ojcius DM, Yilmaz O. The oral microbiota: living with a permanent guest. DNA Cell Biol 2009;28:405-11.

2. Keijser BJ, Zaura E, Huse SM, et al. Pyrosequencing analysis of the oral microflora of healthy adults. J Dent Res 2008;87:1016-20.

3. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

4. Darveau RP. Periodontitis: a polymicrobial disruption of host homeostasis. Nat Rev Microbiol 2010;8:481-90.

5. Zhang J, Sun X, Xiao L, Xie C, Xuan D, Luo G. Gene polymorphisms and periodontitis. Periodontol 2000 2011;56:102-24.

6. Nibali L, Donos N, Farrell S, et al. Association between interleukin-6 -174 polymorphism and Aggregatibacter actinomycetemcomitans in chronic periodontitis. J Periodontol 2010;81:1814-9.

7. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

8. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

9. Socransky SS, Smith C, Martin L, Paster BJ, Dewhirst FE, Levin AE. "Checkerboard" DNA-DNA hybridization. Biotechniques 1994;17:788-92.

10. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

11. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

12. Papapanou PN, Neiderud AM, Papadimitriou A, Sandros J, Dahlén G. "Checkerboard" assessments of periodontal microbiota and serum antibody responses: a case-control study. J Periodontol 2000;71:885-97.

13. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

14. Ge D, Zhang K, Need AC, et al. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640-3.

15. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336-7.

16. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

17. Heidecker B, Lamirault G, Kasper EK, et al. The gene expression profile of patients with new-onset heart failure reveals important gender-specific differences. Eur Heart J 2010;31:1188-96.

18. Kurowska-Stolarska M, Hueber A, Stolarski B, McInnes IB. Interleukin-33: a novel mediator with a role in distinct disease pathologies. J Intern Med 2011;269:29-35.

19. Lin WD, Lin SP, Wang CH, Tsai Y, Chen CP, Tsai FJ. RUNX2 mutations in Taiwanese patients with cleidocranial dysplasia. Genet Mol Biol 2011;34:201-4.

20. Tabuchi S, Sakuta T, Oyama T, et al. Runt-related gene 2 is involved in the inhibition of matrix metalloproteinase-13 expression by roxithromycin in human gingival epithelial cell cultures. J Periodontal Res 2009;44:283-8.

21. Cauchi S, Proença C, Choquet H, et al; D.E.S.I.R. Study Group. Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. J Mol Med (Berl) 2008;86:341-8.

22. Veale KJ, Offenhäuser C, Lei N, Stanley AC, Stow JL, Murray RZ. VAMP3 regulates podosome organisation in macrophages and together with Stx4/SNAP23 mediates adhesion, cell spreading and persistent migration. Exp Cell Res 2011;317:1817-29.

23. Bishop KA, Coy HM, Nerenz RD, Meyer MB, Pike JW. Mouse Rankl expression is regulated in T cells by c-Fos through a cluster of distal regulatory enhancers designated the T cell control region. J Biol Chem 2011;286:20880-91.

24. Yao C, Purwanti N, Karabasil MR, et al. Potential down-regulation of salivary gland AQP5 by LPS via cross-coupling of NF-kappaB and p-c-Jun/c-Fos. Am J Pathol 2010;177:724-34.

25. Murata T, Noda C, Saito S, et al. Involvement of Jun dimerization protein 2 (JDP2) in the maintenance of Epstein-Barr virus latency. J Biol Chem 2011;286:22007-16.

26. Chalabi M, Rezaie F, Moghim S, Mogharehabed A, Rezaei M, Mehraban B. Periodontopathic bacteria and herpesviruses in chronic periodontitis. Mol Oral Microbiol 2010;25:236-40.

27. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL Jr. Microbial complexes in subgingival plaque. J Clin Periodontol 1998;25:134-44.

28. Socransky SS, Haffajee AD. Dental biofilms: difficult therapeutic targets. Periodontol 2000 2002;28:12-55.

29. Badger JH, Ng PC, Venter JC (2010). The human genome, microbiomes, and disease. In: Nelson KE (Ed). Metagenomics of the Human Body. Springer, New York, NY, USA. pp1-14.

J. MANUSCRIPT #2- SUPPLEMENTAL MATERIAL

SUPPLEMENTAL MATERIALS AND METHODS

Study population

The study sample consisted of participants of the Dental ARIC, an ancillary study of the Atherosclerosis Risk In Communities (ARIC) longitudinal cohort investigation (1). ARIC is a study of atherosclerosis, CVD risk factors and outcomes. The Dental ARIC, an ancillary study was undertaken between 1996 and 1998 during the ARIC visit 4. For the purposes of Dental ARIC a complete oral-dental examination was performed among 6979 subjects, whereas collection of gingival crevicular fluid and sub-gingival microbial plaque samples was undertaken for a subset of participants (2). Specifically, of the 6979 Dental ARIC participants, 1450 had microbial plaque samples collected (2-4). For the purposes of the present GWA study, we considered white subjects who had provided microbial plaque samples (n=1236) and were successfully matched with genotype data (n=1020). Additional information that was collected as part of ARIC visit 4 and was used for data presentation and analysis included the following variables: sex, age, CP diagnosis [Centers for Disease Control (CDC) and American Academy of Periodontology (AAP) three-level classification: health/mild, moderate, and severe disease] (5), smoking status (never, former, current), and DM (healthy or DM).

Genotyping and imputation

In the study population, DNA was extracted from blood samples drawn from an antecubital vein into tubes containing serum separator gel. Blood samples were analyzed at a central ARIC laboratory in Houston, TX. Genotyping was performed with the Affymetrix

Genome-Wide Human SNP Array 6.0 chip. The platform offers 906600 markers for SNPs. The rigorous quality control procedures included initial blind duplicate genotyping and identification/flagging of SNPs with kappa<0.95 and reconciliation of unintentional duplicate samples (17 duplicates and one triplicate). Imputation to 2.5 million markers was performed using 669450 SNPs and MACH v1.0.16 (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html), based on HapMap Phase II CEU build 36. The selected SNPs were selected from 839048 autosomal SNPs restricted to those with minor allele frequency (MAF)>0.01 (129543 excluded), Hardy-Weinberg equilibrium (HWE) $P>10^{-5}$ (12432 excluded) and call rate >95% (1693 excluded). After imputation, SNPs with a quality score <0.8 and missing data rate >10% were identified and flagged for removal, while only those with MAF of ≥5% were considered for analytical purposes.

Population stratification

To obtain estimates of relatedness and population stratification a subset of 85,947 "high quality" SNPs was selected. These SNPs met the following criteria among self-reported whites: MAF ≥0.1, call rate >99.5%, HWE $P\geq10^{-3}$, autosomal, with annotation in the platform annotation file, not labeled "AFFX" or "chromosome 0", and not monomorphic. Using these SNPs identity-by-state (IBS) allele sharing distance (DST values) were computed using PLINK, as such: DST = IBS distance ($IBS_2 + 0.5*IBS_1$) / (n SNP pairs). First degree relative status was assigned to pairs of individuals with DST≥0.8 and second degree relatives were considered those with 0.763 ≤DST< 0.8. Among the Caucasian ARIC participants there were 380 pairs of first degree and 207 pairs of second degree relatives identified. To minimize exclusions, related pairs were broken by iterative selection of individuals with most relatives using a custom program.

Population stratification was further evaluated with principal component (PC) analysis using the EIGENSTRAT program (6). The above chosen set of SNPs was used for the computation of ten principal components. Genetic outliers were considered those that were further than 8 standard deviations (SD) away from any of ten PCs over ten runs of PC computation. Based on DST and PC criteria there were 716 subjects flagged from removal from the analysis (206 as genetic outliers based on PCs and 16 based on average DST values ("too little IBS sharing" with the rest of the sample), 351 first degree relatives and 143 second degree relatives. All but ten second degree relatives (whose relatives were excluded as genetic outliers) were re-entered in the dataset and were assigned PCs. After exclusion of 364 individuals (4%) there were 9349 Caucasians who were included in the GWA analysis and of those, 1020 had periodontal microbiological data available as Dental ARIC participants.

SUPPLEMENTAL RESULTS REGARDING *PORPHYROMONAS GINGIVALIS*

We detected no genome-wide significant association signals with regard to *Pg* "high" colonization. However, six SNPs marking three loci emerged below the $P<5x10^{-6}$ threshold and were prioritized for further investigation. Of those, rs10760187 in 9q33.2 provided the lowest P-value [$P=4.5x10^{-7}$; OR=2.07 (95% CL=1.55, 2.76)], with the 'risk' [C] allele (HapMap II CEU MAF [T]: 0.514) showing 12% enrichment among subjects with "high" colonization. This SNP is in an intergenic area between *DAB2IP* (18Kb) and *TTLL11* (18Kb) and in LD with several variants intronic to both genes. Rs1011108 [$P=2.0x10^{-6}$; OR=1.79 (95% CL=1.40, 2.27); MAF-CEU: [T] 0.374] is in a gene-dense locus at 2p23.3, 1Kb from *OTOF*, 3Kb from *C2Orf70* and 18Kb from *CIB4*. In fact, the second prioritized SNP in this locus (rs1275992) is intronic to *CIB4* and in linkage disequilibrium ($R^2=0.58$, D'0.94) with

rs13002673 which represents a missense change in *C2Orf70* resulting to a [Gln]→[His] substitution, predicted as 'benign' by PolyPhen-2. The third locus that was associated with *Pg* "high" colonization was marked by rs1360573 at 10p11.21 [$P=1.5x10^{-6}$; OR=2.75 (95% CL=1.78, 4.26)]. The common "risk" allele [A] showed a 10% enrichment, and the closest gene is *ANKRD30A* (529Kb).

SUPPLEMENTAL DISCUSSION REGARDING *PORPHYROMONAS GINGIVALIS*

The examination of *Pg* colonization as a distinct trait is supported by the major role that this pathogen has in chronic periodontitis (7, 8). These exploratory analysis results, above and beyond those reported for "red" complex bacteria, can be considered as additional candidate regions that may be implicated in increased colonization with *Pg* and thus, higher risk for chronic periodontitis.

With regard to genes in the 2p23.3 locus, *CIB4* (*KIP4*) is known as calcium and integrin binding family member 4. Mutations in the *OTOF* (otoferlin) gene have been associated with non-syndromic types of deafness (9). No information is available on the role of *C2Orf70*. Interestingly, the 9q33.3 locus that was marked by rs10760187 in our GWAS was also identified as associated with abdominal aortic aneurysm in a Northern European GWAS (10). This genomic area has also been reported to harbor important susceptibility variants associated with cleft lip/palate (11). One candidate gene in the 9q33.3 locus, *DAB2IP*, is also known as DAB2-interacting protein or *AIP1*. This gene is member of a family of regulators of extracellular stimuli that serve to maintain the homeostasis of cellular functions (12) and has been shown to be inactivated by methylation in prostate and breast cancer (13). *TTLL11*, another candidate gene in this locus, encodes a member of a large

family of proteins involved in the catalytic ligation of amino acids to tubulins or other substrates (14). Finally, the closest gene to the top SNP in the 21q22.3 locus was *ANKRD30A*, or ankyrin repeat domain-containing protein 30A. Reports have linked this gene and its protein (NY-BR-1) as a differentiation marker of the mammary gland, and thus as a potential diagnostic and immunotherapeutic aid for breast carcinomas (15, 16) with cancer. Moreover, a recent study found an association of a 3'UTR SNP of *ANKRD30A* with the human serum metabolic profile (17).

K. SUPPLEMENTAL TABLE

Table 5.3 *Porphyromonas gingivalis* "high" colonization (quantified with DNA "checkerboard" hybridization) GWA results among the white participants of the Dental ARIC study (n=1020). Single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF-HapMap II CEU) of ≥5% and $P<5x10^{-6}$. The SNP with the lowest P-value per locus is presented; additional prioritized SNPs in each locus are presented in the footnote, along with corresponding R^2 (based on 1000 genomes pilot 1 release) with the top SNPs.

Chr.	SNP	Position Build36	ca[1]	nca[2]	Risk allele frequency (HapMap II- CEU)	Closest gene(s) and position or distance	Risk allele frequency low/ high colonization	P value	Odds ratio (95% CL[3])
2p23.3	rs1011108[4]	26636125	C	T	[T] 0.374	*OTOF* (1Kb); *C2Orf70* (3Kb); *CIB4* (21Kb)	0.341/0.485	$2.0x10^{-6}$	1.79 (1.40, 2.27)
9q33.2	rs10760187[5]	123605641	C	T	[C] 0.514	*DAB2IP* (18Kb); *TTLL11* (18Kb)	0.498/0.618	$4.5x10^{-7}$	2.07 (1.55, 2.76)
10p11.21	rs1360573	36925927	A	G	[A] 0.759	*ANKRD30A* (529 Kb); *FZD8* (956 Kb)	0.752/0.853	$1.5x10^{-6}$	2.75 (1.78, 4.26)

1: coded allele
2: non-coded allele
3: Confidence limits
4: Additional SNP in locus with $P<5x10^{-6}$: rs1275992 (R^2=0.33), intronic to *CIB4* and in linkage disequilibrium (R^2=0.58, D'0.94) with rs13002673 which represents a missense change in *C2Orf70* resulting to a [Gln]→[His] substitution
5: Additional SNPs in locus with $P<5x10^{-6}$: rs7849478 (R^2=0.63), rs10985387 (R^2=0.67)

Table 5.4 Reported SNPs as expression quantitative loci (eQTLs) and their association with gene expression (lymphoblastoid cell line-based expression) in populations of European descent with $P<10^{-4}$ in the Scan database (http://www.scandb.org).

dbSNP	Gene symbol	Gene name	Locus	P-value
rs10043775	KIAA0515 (*PRRC2B)*	proline-rich coiled-coil 2B	9q34.13	$6x10^{-5}$
	CCDC57	coiled-coil domain containing 57	17q25.3	$5x10^{-5}$
rs1932040	MRPL15	mitochondrial ribosomal protein L15	8q11.2-q13	$9x10^{-5}$
	EAF1	ELL associated factor 1	3p25.1	$3x10^{-5}$
	TAF5L	TAF5-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65kDa	1q42.11-q42.3	$8x10^{-5}$
rs16924631	MED15	mediator complex subunit 15	22q11.2	10^{-4}
	PSMB7	proteasome (prosome, macropain) subunit, beta type, 7	9q34.11-q34.12	$2x10^{-5}$
	MVD	mevalonate (diphospho) decarboxylase	16q24.3	10^{-4}
	MAN2A2	mannosidase, alpha, class 2A, member 2	15q25	$9x10^{-6}$
	DNMBP	dynamin binding protein	10q24.31	$6x10^{-5}$
	PQLC1	PQ loop repeat containing 1	18q23	$4x10^{-5}$
	PRKCE	protein kinase C, epsilon	2p21	$3x10^{-6}$
	NEK6	NIMA (never in mitosis gene a)-related kinase 6	9q33.3-q34.11	$2x10^{-5}$
	PREX1	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1	20q13.13	10^{-4}
rs1011108	TMEM2	transmembrane protein 2	9q21.13	$3x10^{-5}$
	LRRK1	leucine-rich repeat kinase 1	15q26.3	10^{-5}
	A4GALT	alpha 1,4-galactosyltransferase	22q13.2	10^{-4}
	MTTP	microsomal triglyceride transfer protein	4q24	$5x10^{-5}$
	MYO6	myosin VI	6q14.1	$2x10^{-5}$
	ASPH	aspartate beta-hydroxylase	8q12.1	$6x10^{-6}$
	CNKSR3	CNKSR family member 3	6q25.2	$9x10^{-5}$
	F13A1	coagulation factor XIII, A1 polypeptide	6p24.2-p23	$4x10^{-5}$
	RASSF6	Ras association (RalGDS/AF-6) domain family member 6	4q21.1	$5x10^{-6}$
	LOC100129069	hypothetical protein LOC100129069	11q24.2	10^{-4}
	HNF4G	hepatocyte nuclear factor 4, gamma	8q21-q22	10^{-4}
	MLSTD1	fatty acyl CoA reductase 2	2p11.23	10^{-4}
	CHL1	cell adhesion molecule with homology to L1CAM (close homolog of L1)	3p26	$6x10^{-5}$
	PIP3-E	interaction protein for cytohesin exchange	6q25.2	10^{-5}

	factors 1		
HNF1B	HNF1 homeobox B	17q12	$3x10^{-5}$
SERPINB1	serpin peptidase inhibitor, clade B (ovalbumin), member 1	6p25	$7x10^{-5}$
ESAM	ribosomal protein L6 pseudogene 25	12q21.31	10^{-4}
IL17RB	interleukin 17 receptor B	3p21.1	10^{-4}
CCDC64	coiled-coil domain containing 64	12q24.23	10^{-4}
C14orf105	chromosome 14 open reading frame 105	14q22.2	$2x10^{-5}$
C1orf161	mab-21-like 3	1p13.1	$2x10^{-5}$
MMAA	methylmalonic aciduria (cobalamin deficiency) cblA type	4q31.1	$5x10^{-5}$
C10orf137	mosome 10 open reading frame 137	10q26.2	$6x10^{-5}$
ZNF503	zinc finger protein 503	10q22.3	$5x10^{-5}$
ZNF44	zinc finger protein 44	19p13.2	10^{-4}
CXorf21	chromosome X open reading frame 21	Xp21.3	$6x10^{-5}$
PIR	pirin (iron-binding nuclear protein)	Xp22.31	$2x10^{-5}$
MGC13057	chromosome 2 open reading frame 88	2q32.2	$4x10^{-5}$
SERPINB9	serpin peptidase inhibitor, clade B (ovalbumin), member 9	6p25	$2x10^{-5}$
CCDC4 (BEND4)	BEN domain containing 4	4p13	$3x10^{-5}$
SYNE2	spectrin repeat containing, nuclear envelope 2	14q22.1-q22.3	$8x10^{-5}$
BHLHB5	basic helix-loop-helix family, member e22	8q12.1	10^{-4}
LRRK1	leucine-rich repeat kinase 1	15q26.3	$4x10^{-6}$
GTF3C6	general transcription factor IIIC, polypeptide 6, alpha 35kDa	6q21	$7x10^{-5}$
PLA1A	phospholipase A1 member A	3q13.13-q13.2	$2x10^{-5}$
WDR91	WD repeat domain 91	7q33	$6x10^{-5}$
LIN7A	lin-7 homolog A	12q21	$9x10^{-6}$
MAGEF1	melanoma antigen family F, 1	3q13	10^{-4}

SUPPLEMENTAL FIGURES

Figure 5.10 Quantile-quantile plot of genome-wide association analysis results of “high red” complex bacterial colonization, among the 1020 Dental ARIC participants.

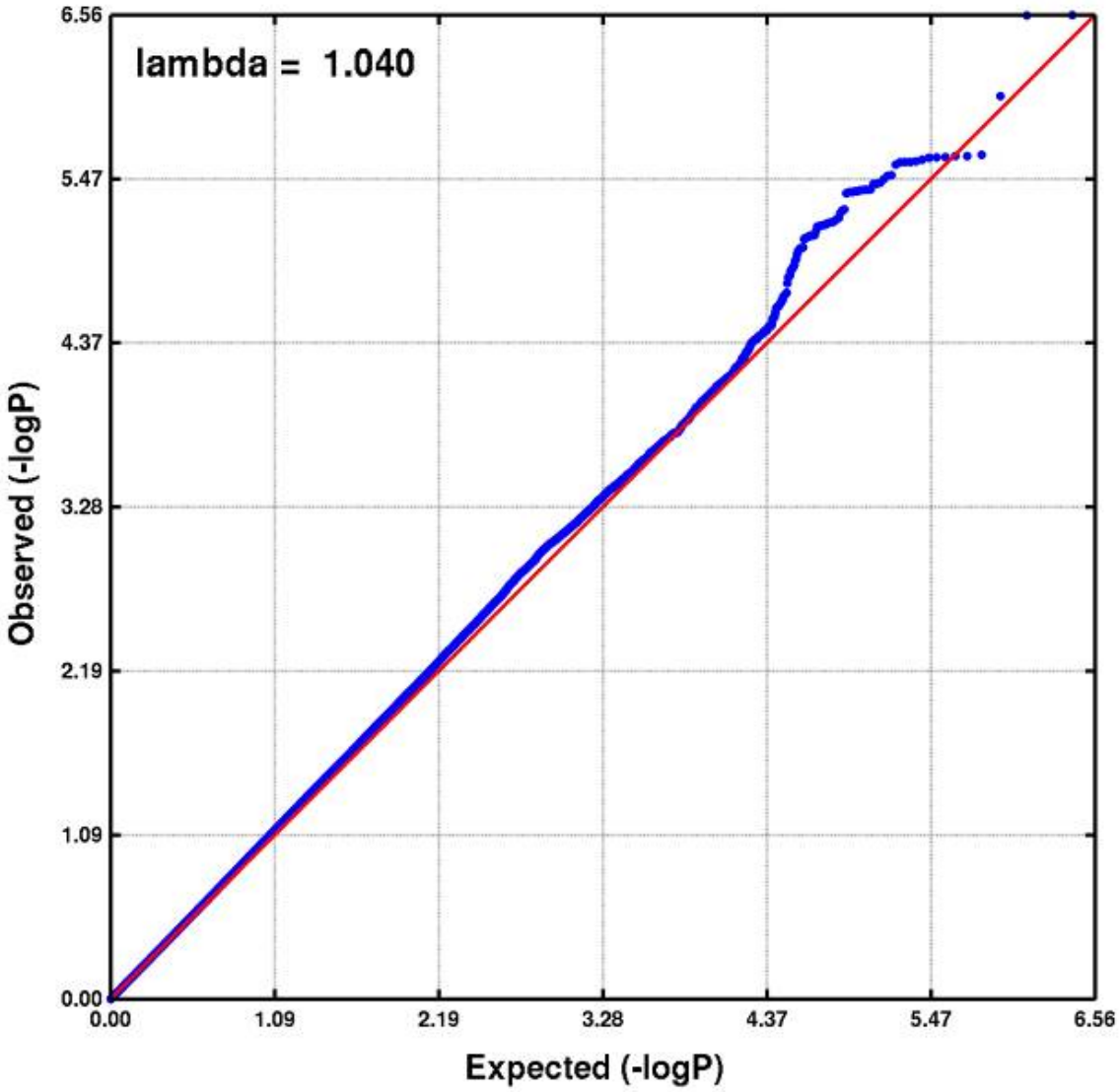


Figure 5.11 Quantile-quantile plot of genome-wide association analysis results of “high orange” complex bacterial colonization, among the 1020 Dental ARIC participants.

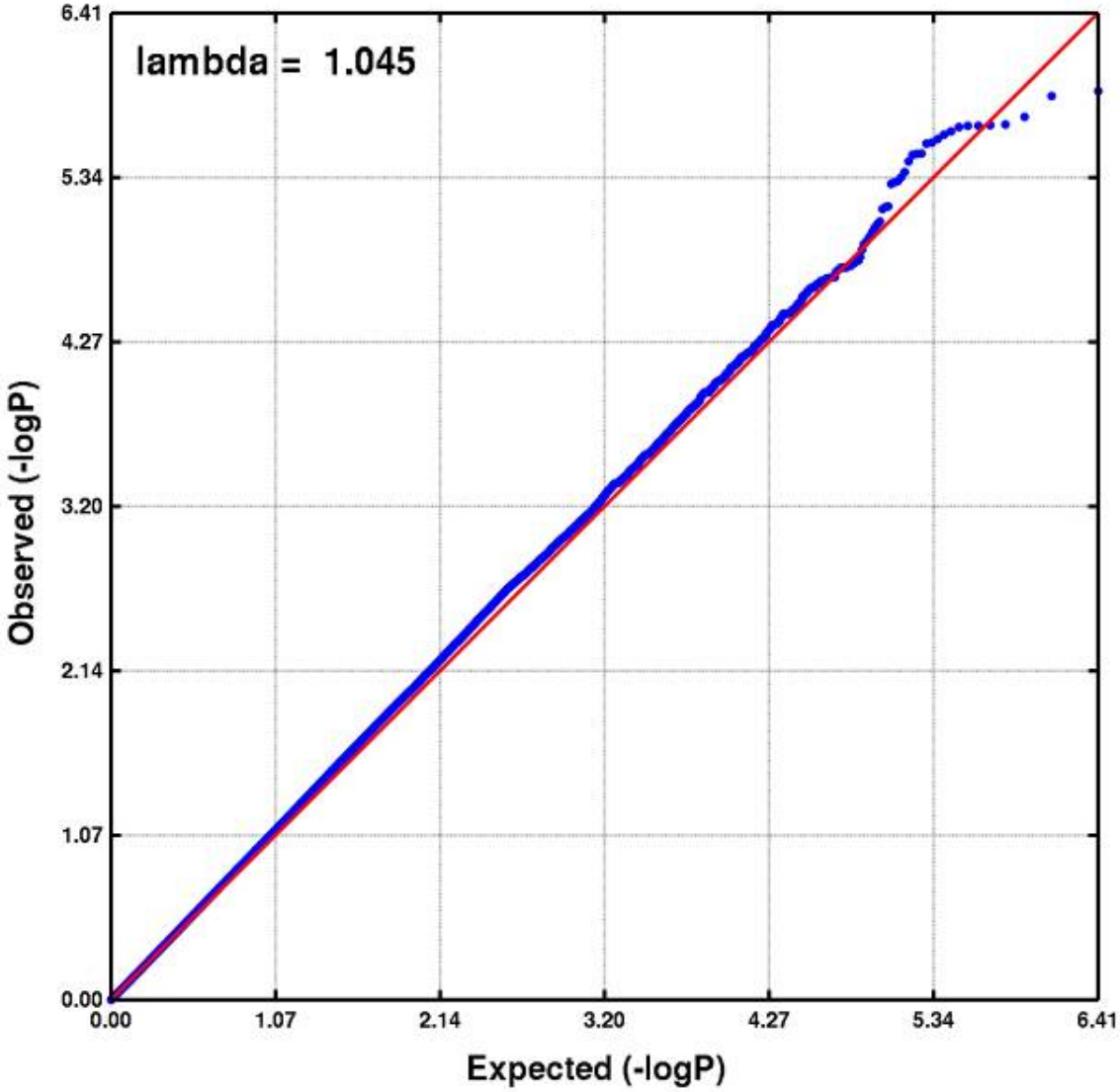


Figure 5.12 Quantile-quantile plot of genome-wide association analysis results of "high" *Aggregatibacter actinomycetemcomitans* colonization, among the 1020 Dental ARIC participants.

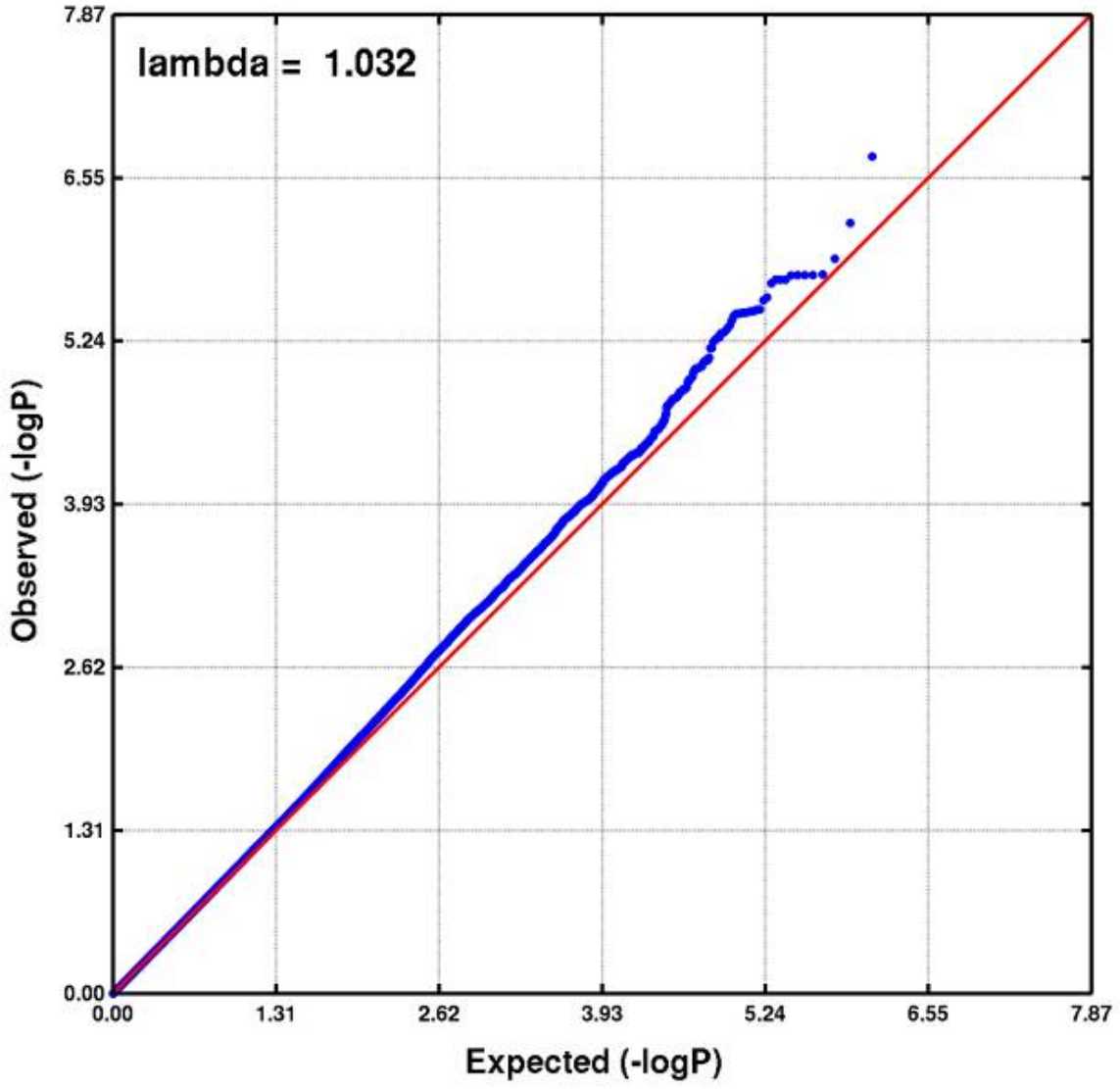


Figure 5.13 Visualization of the 2p23.3 locus that was marked by rs1011108 ($P=2.0x10^{-6}$) for "high" *Porphyromonas gingivalis* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

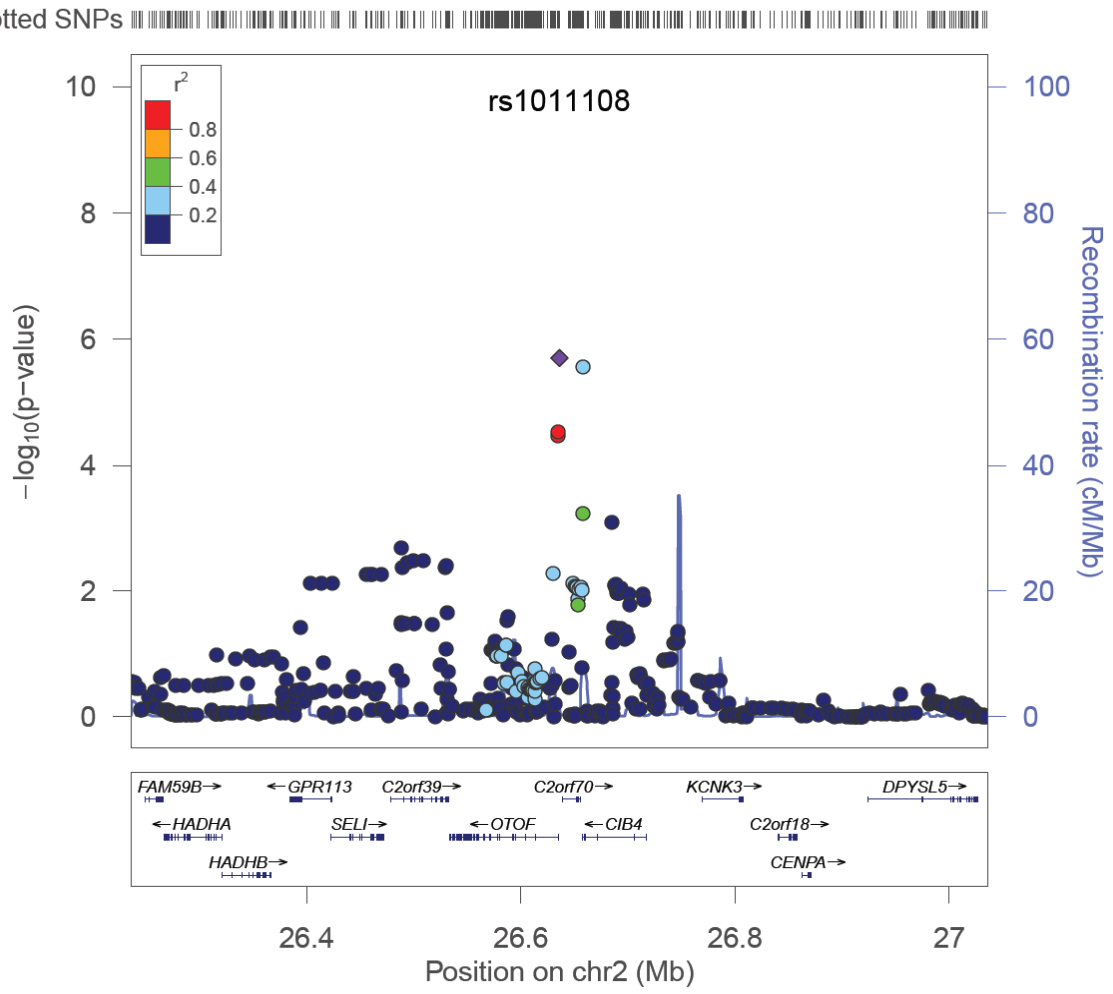


Figure 5.14 Visualization of the 9q33.2 locus that was marked by rs10760187 ($P=4.5x10^{-7}$) for "high" *Porphyromonas gingivalis* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

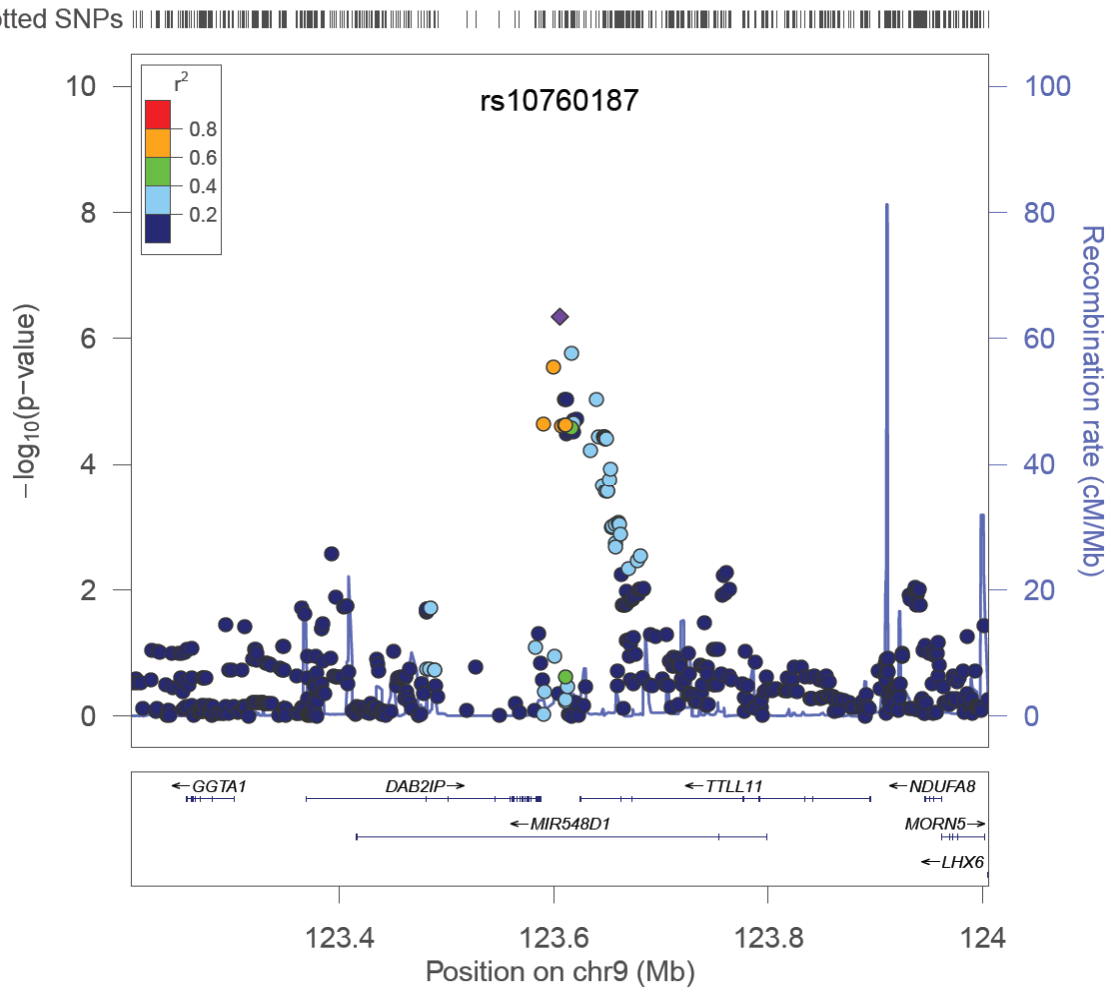
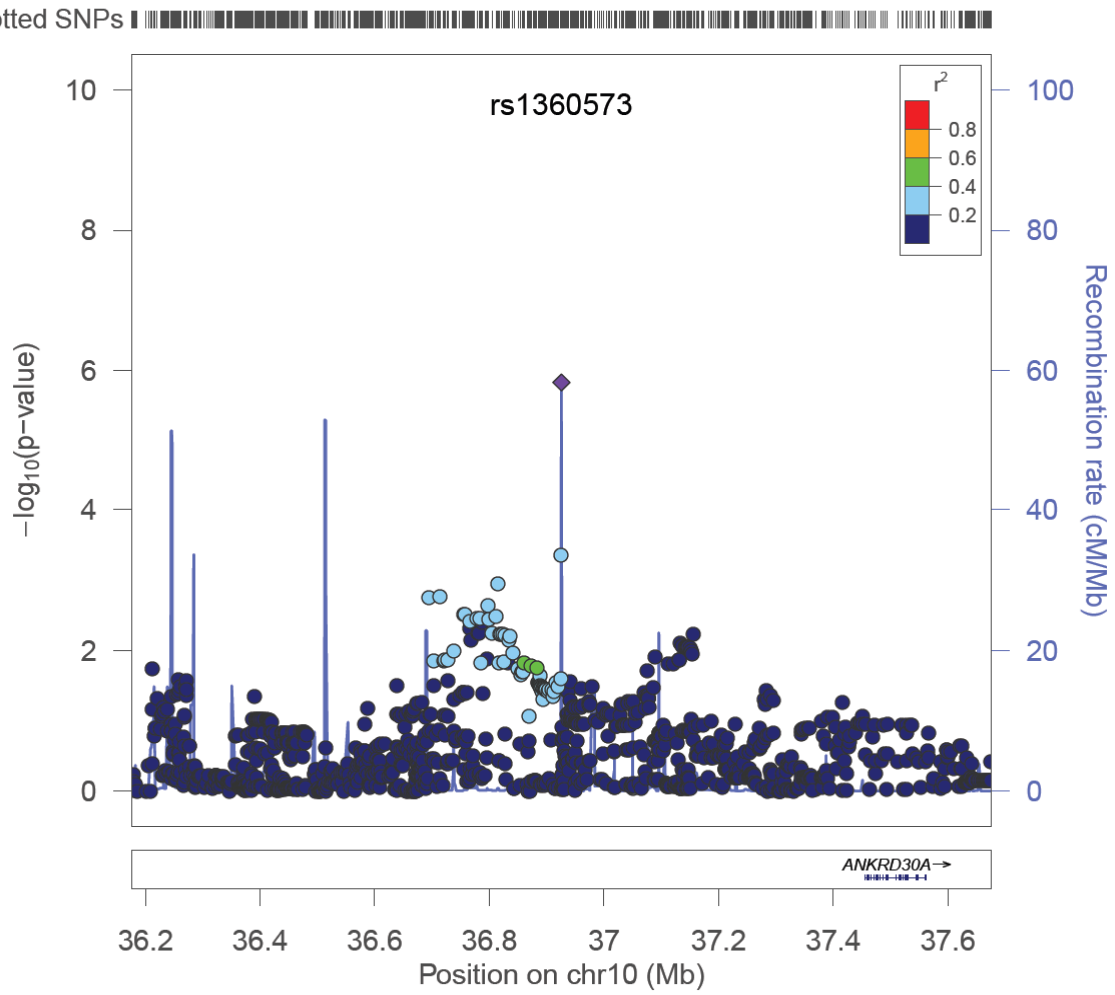


Figure 5.15 Visualization of the 10p11.21 locus that was marked by rs1360573 ($P=1.5x10^{-6}$) for "high" *Porphyromonas gingivalis* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort. The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).



SUPPLEMENTAL REFERENCES

1. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

2. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

3. Elter JR, Offenbacher S, White RP, Beck JD. Third molars associated with periodontal pathology in older Americans. J Oral Maxillofac Surg 2005;63:179-84.

4. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

5. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78:1387-99.

6. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome wide association studies. Nat Genet 2006;38:904-9.

7. Mooney J, Adonogianaki E, Riggio MP, Takahashi K, Haerian A, Kinane DF. Initial serum antibody titer to Porphyromonas gingivalis influences development of antibody avidity and success of therapy for chronic periodontitis. Infect Immun 1995;63:3411-6.

8. van Winkelhoff AJ, Loos BG, van der Reijden WA, van der Velden U. Porphyromonas gingivalis, Bacteroides forsythus and other putative periodontal pathogens in subjects with and without periodontal destruction. J Clin Periodontol 2002;29:1023-8.

9. Yasunaga S, Grati M, Cohen-Salmon M, et al. A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes DFNB9, a nonsyndromic form of deafness. Nat Genet 1999;21:363-9.

10. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat Genet 2010;42:692-7.

11. Letra A, Menezes R, Govil M, et al. Follow-up association studies of chromosome region 9q and nonsyndromic cleft lip/palate. Am J Med Genet A 2010;152A:1701-10.

12. Iwashita S, Song SY. RasGAPs: a crucial regulator of extracellular stimuli for homeostasis of cellular functions. Mol Biosyst 2008;4:213-22.

13. Yano M, Toyooka S, Tsukuda K, et al. Aberrant promoter methylation of human DAB2 interactive protein (hDAB2IP) gene in lung cancers. Int J Cancer 2005;113:59-66.

14. Janke C, Rogowski K, Wloga D, et al. Tubulin polyglutamylase enzymes are members of the TTL domain protein family. Science 2005;308:1758-62.

15. Theurillat JP, Zürrer-Härdi U, Varga Z, et al. Distinct expression patterns of the immunogenic differentiation antigen NY-BR-1 in normal breast, testis and their malignant counterparts. Int J Cancer 2008;122:1585-91.

16. Jäger D, Filonenko V, Gout I, et al. NY-BR-1 is a differentiation antigen of the mammary gland. Appl Immunohistochem Mol Morphol 2007;15:77-83.

17. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 2008;4:e1000282.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

A. Summary of findings

In our genome-wide analysis of chronic periodontitis (CP) in a well-defined cohort of approximately 4600 white subjects we detected no genome-wide significant signals. However, we found suggestive evidence of association ($P<5x10^{-6}$) for CP with markers at ten loci including *NPY, NIN*, and *WNT5A* for severe CP, *NCR2* and *EMR1* for moderate CP, and *TBX18, ETS1, DYNC2H1, TTC26* and *ZC3HAV1* for the "extent" of attachment loss (EAL) trait. Several of these loci include candidate genes that are related to immune function and regulation, as well as neurological function. Noteworthy, four loci marked by rs3826782, rs11925054, rs17792917, rs11615037 have been identified as expression quantitative trait loci (eQTL) on lymphoblastoid cell lines (LCL). Sensitivity analyses including smoking and diabetes-adjusted genetic models did not show any material impact on these results. However, in these adjusted analyses five additional loci including *SCD5, RPS29, PPIL5, COL11A1, CPM* and *CPSF6* emerged below the $P<5x10^{-6}$ threshold. The 4q21.3 (*SCD5*) locus was shared for severe CP and EAL, and the 6p21.1 (*NCR2*) locus reached genome-wide significance for moderate CP (rs7762544; $P_{adjusted}= 3.8x10^{-8}$). The *NPY* (7p15) locus was replicated for severe CP in an independent sample of whites of European descent. We performed inverse-variance weighted meta-analysis for three loci that

revealed no evidence of effect heterogeneity between the two studies. Pooled odds ratios for these three loci were: 7p15 (*NPY*, rs2521634)—OR=1.49 (95% CI=1.28-1.73), P=3.5x10^{-7}; 19p13.3 (*EMR1*, rs3826782)—OR=2.01 (95% CI=1.52-2.65), P=8.2x10^{-7}; 6p21.1 (*NCR2*, rs7762544)—OR=1.40 (95% CI=1.24-1.59), P=7.5x10^{-8}. Our exploratory analyses of SNP interactions with sex, smoking and diabetes detected several interaction effects below an empirical P<10^{-5} threshold: 20 SNPs with sex, 17 with diabetic status, and 12 with the three-level smoking categorization (never, former, current).

In the GWA scan in a subset of approximately 1000 subjects for which periodontal microbiological data were available we detected associations with markers at thirteen loci including *KCNK1*, *FBXO38*, *UHRF2*, *IL33*, *RUNX2*, *TRPS1*, *CAMTA1* and *VAMP3* providing suggestive signals of association (P<5x10^{-6}) with "high" bacterial colonization with "red" complex, "orange" complex, and *Aggregatibacter actinomycetemcomitans* (*Aa*). Exploratory analyses revealed three additional loci including *OTOF*, *C2Orf70*, *CIB4*, *DAB2IP*, *TTLL11* and *AKNRD3* associated with *Porphyromonas gingivalis* (*Pg*) "high" colonization. Four of these loci marked by rs10043775, rs1932040, rs16924631 and rs1011108 were identified to be eQTLs on LCL. Two missense changes that were identified for "high" bacterial colonization traits (rs10043775, 5q33, *FBXO38* gene [Pro]→[Arg] and rs13002673 (in LD with the prioritized rs1011108), 2p23.3, *C2Orf70* gene [Gln]→[His]) were predicted to be 'benign'.

These genome-wide scan results provide new information on multiple candidate regions for interrogation in genetic studies of CP. Further investigations

providing replication of these findings and examination of specific associations may lead to an improved understanding of the pathogenesis of CP as well as the complex nature of host-biofilm and -bacteria interactions that characterizes states of health and disease.

B. Strengths

1. Genome-wide association analysis approach

The opportunity that GWAS provide for investigators to examine the association of currently more than a million DNA markers with a given phenotype without being limited by prior knowledge, known pathways or mechanistic hypotheses in their marker selection, is unique. It has been argued that most associations detected by GWAS represent new discoveries. This “unbiased” or “agnostic” (1) approach in examining the genetic component of various phenotypes facilitates identification of novel loci and thus, potential insights into new pathways or mechanisms of biological processes. The present study is the first report of a GWA analysis of CP to-date, and provides suggestive evidence of association for several genomic loci that may be promising leads for further genetic investigations.

2. Phenotype characterization

The value and quality of the Dental ARIC clinical examinations protocol is truly unique. It is known that clinical-periodontal examination data are accurate and reproducible when they are performed by trained and calibrated dental examiners, as was the case with Dental ARIC study. Recent reports (2) stress that national estimates

of CP prevalence may severely underestimate the true prevalence due to disease misclassification in the NHANES III and IV. The next NHANES oral-dental examinations will be performed for the first time following the "full-mouth six site per tooth" protocol used in the Dental ARIC study. As outlined in the introduction, numerous definitions and quantitative measures exist to capture the extent and severity of periodontitis (3), with each been subject to varying degrees of bias (measurement error, bias due to loss of severely affected teeth, and more). The use of the most widely accepted and consensus CDC/AAP classification of CP, in combination with a clinical examinations protocol of highest standard is a major strength of the proposed investigation. We considered a continuous and non-reversible measure of periodontal destruction as an additional strength.

The investigation of microbiological colonization phenotypes in addition to CP is offering additional insights into the factors facilitating or antagonizing the cascade of pathological oral ecological shifts that are characteristic of the disease. The oral microbial load of Gram-negative and anaerobic periodontal pathogens may be an independent risk factor for, or modifier of several systemic conditions (4-7), and its exploration adds value to this study. While periodontitis is a clinical manifestation of exaggerated immune response to predominantly commensal oral bacteria, investigating genetic factors associated with colonization and proliferation of these causative bacterial agents is a fundamental question.

3. Covariates and sensitivity analysis

Because population stratification ("systematic" differences in ancestry between trait categories) can result in spurious GWA results, appropriate adjustment for population substructure was undertaken in all analyses, using 10 EIGENSTRAT-derived principal components. Inspection of the quantile-quantile plots did not reveal any "early departure" of the observed versus expected P-values. Additionally, genomic inflation (lambda) factors for most analyses were generally low, below 1.030. Our analyses also benefitted from the detailed measurement and consideration in sensitivity analyses of the well-documented risk factors or disease modifiers, smoking and diabetic status. Although no confounding is expected in the association between SNPs and the examined traits, the consideration of smoking history in the sensitivity analysis provides the opportunity to rule out the identification of risk loci for this risk factor, which has been shown to confer epigenetic changes (8, 9). Moreover, the inclusion of the diabetic status as a covariate in the sensitivity analysis is based on the fact that diabetes is associated with altered host response (10, 11) and increased susceptibility to bacterial infections (12). A sensitive definition of diabetes was used in the ARIC study, including fasting and non-fasting serum glucose levels, and pharmacological treatment for the disease.

C. Limitations

1. Phenotype characterization

As noted previously, the use of the CDC/AAP definition of CP represents the most widely accepted standard for classifying the disease. While our clinical data and

case-definition used for CP are of the highest standards possible, the disease itself entails some unique methodological challenges. Although periodontitis is a defined as a disease at the "person- or oral cavity-level", it is actually characterized by multiple-attacks on the periodontal attachment of individual teeth (13). As it progresses, periodontitis leads to gradual periodontal tissue destruction, alveolar bone loss, and eventually tooth loss. For example, in a cross-sectional study of periodontal status in relation to age and tooth type it was found that incisor teeth were the most frequently affected by advanced disease, whereas molar teeth were most frequently missing (14). In "periodontal maintenance" studies, it has similarly been found that molar teeth are the most frequently extracted due to periodontal reasons (15). This constitutes a form of "censoring" of the most affected "diseased units", the teeth. In this respect, CP can be a self-limiting disease on a tooth-level. This is also the case on a patient-level, as (considering the extreme scenario) an edentulous individual is not at risk to develop periodontitis. A consequence of this phenomenon is the observation that the prevalence, as well as extent and severity of CP in epidemiologic studies are likely biased downwards from tooth loss attributable to the disease. It has been recognized that variation in the criteria and case definitions used for the classification of CP has been a source of inconsistencies in the dental literature (16). Also, as outlined in the introduction, more quantitative measures exist to capture the extent and severity of periodontitis compared to "crude" classifications, such as the "extent" of attachment loss, which was employed in this study. It will be worthwhile for future studies to consider juxtaposing the present investigation's findings to additional sensitivity

analyses using extent scores (e.g. ≥3mm CAL) as continuous variables for the GWA analyses, while adjusting for tooth loss.

Another consideration is that periodontitis, when measured by clinical measures of PD and (mainly) CAL, essentially represents the cumulative history of periodontal tissue destruction and not necessarily active disease or infection. Capturing this cumulative impact on periodontal tissues may be beneficial when considering longitudinal effects or progression patterns of periodontitis that may be biologically admixed. While more biologically-relevant disease classifications have been developed, our independent consideration of the association of host genetic variants with the microbial biofilm composition represents an improvement over the study of CP diagnoses alone. Despite being limited by small sample size for a GWAS (approximately 1000 individuals with microbial level counts), we consider this approach as the first exploration of loci associated with the composition of the oral microbiome. However, we acknowledge that this discovery step is limited by the very low power to detect modest effects. Similarly, our GWAS was underpowered to detect interaction effects in this additional exploratory step.

2. Genome-wide analysis approach

Whole-genome sequencing for population-based studies is financially and resource-prohibitive with today's state of technology. For this purpose, imputation methods for non-typed markers have been developed, and imputation has become a standard step in GWA studies (17). Imputing does introduce some uncertainty in the non-typed allele frequencies, but publicly available programs for GWA analyses such as

probABEL (18) allow the use of “dosage files” including imputed SNPs. As denser mapping of the human genome is realized, opportunities to use a 1000 genomes-based (19) imputation of the ARIC genotyped data (up to 10 million markers) will be an important improvement to the current approach.

A frequent criticism of the GWAS concerns the current state of very conservative criteria of (genome-wide) statistical significance that are necessary to reduce the false-positive rate (1). The traditional Bonferroni correction is believed to be ultra-conservative due to the fact that SNPs are to some degree correlated, as the ~2.5 million association tests performed in imputed datasets are not truly independent from each other. This stringency may allow only for relatively strong associations to be discovered, while much “good signal” obtained by SNPs with modest or small effects goes without attention. This is compounded by the fact that many common complex traits are thought to be influenced by a large numbers of low-penetrance, small effect variants. Park et al. recently suggested that if these large numbers of low-penetrance variants is identified from sufficiently powered GWAS, together they could explain 20% or more of the known heritability of such traits (20). On the other hand, another concern with GWAS findings is related to the “winner’s curse” (21) which becomes evident when identified loci are not being replicated in external samples or are not reproduced in subsequent studies. For these reasons, the interpretation of GWAS findings is now proposed to become subject to different criteria (compared to “traditional” epidemiologic evidence) to assess their inferential potential. Based on these proposed “interim guidelines” (22), genome-wide evidence should be weighed in terms of amount of evidence, extent of replication and

protection from bias and thus classified in a three-level categorization as strong, moderate or weak evidence.

Moreover, it is becoming increasingly evident that the variation detected by GWAS is only one dimension of the genetic basis of health and disease (23). For instance, the role of copy number polymorphisms (24), rare alleles (25), epistatic effects (26) and that of the non-coding regions (27, 28) in the human genome are thought to be parts of the missing piece of the "missing heritability" (25). For example, Dickson and colleagues suggested that "synthetic genome-wide associations" can be detected from combinations of rare variants (29). In a recent review, Panagiotou and colleagues summarized the published evidence of genome-wide association for rare (<5%) variants and found that there are still few genome-wide significant rare variants, but suggested that a possible confluence of rare and common variants in the same loci is likely to be depended on the different ancestry references used (30). Bansal and colleagues, in a recent review provided recommendations with regard to analytical approaches that could potentially aid in the discovery of "rare variant" effects (31).

Moreover, it is common ground that biological processes are best represented by "cycles" and the action of biological messengers is understood on "pathways". Therefore, it may be overly simplistic to consider independent effects of SNPs when it is recognized that usually multi-molecular pathways, feedback loops and regulating mechanisms are functioning in tandem. For this reason, any opportunity to jointly evaluate the effect of multiple "functionally relevant" SNPs is appealing and is becoming more feasible with advances in GWA methods (32, 33).

Further, the study of epigenetic phenomena has opened a whole new chapter in research of the pathogenesis and therapeutics, by a joint consideration of genotype and environmental influences (34) that is not routinely considered in GWAS. Feinberg recently suggested that the consideration of epigenomics, the genome-scale study of epigenetics in the function of our genomic make-up will likely be equally important to our understanding of the relation between function and organ anatomy in humans (35). To place this in perspective of studies of CP, it has been shown that bacteria involved in periodontal infection are capable of inducing such epigenetic DNA alterations (36).

3. Potential biases

As explained in detail in the two previous sections, the phenotype characterization and limitations inherent to GWAS are two major potential sources of bias for this investigation. With regard to the CP phenotype, it is possible that the CDC classification is not sensitive or quantitative enough to capture the entire spectrum of disease expression. For example, using the CDC trait may be an inappropriate phenotype if one considers CP equally likely to develop due to a hyper-inflammatory host reaction or an immune-compromised state. Most likely, there are different genetic influences for these two types of host-response profile. To address this limitation, the consideration of biologically-refined classifications of CP such as the BGI may be indicated.

Another phenotype-related source of bias is the definition of the “high” bacterial colonization traits for our Specific Aim 2. It is known that the semi-quantitative method of DNA-DNA “checkerboard” hybridization has a lower count

detection threshold of 10^3 to 10^4. On the other hand, experts agree that CP risk is "frankly elevated" at high count levels, in the range of 10^5 to 10^6. In our study we arbitrarily defined "high colonization cases" those who were in the highest quintile of the distribution for each bacterial trait. Using this approach we were able to retain an "adequate" proportion of 20% of subjects as cases, representing a "pragmatic" approach that allowed an analytically feasible contrast- defining a higher threshold would result in fewer "high colonization cases" and perhaps reduced power to detect GWA signals. Future studies should, however, explore the feasibility of different definitions of the "high colonization" traits. Moreover, although our summation of different bacteria into the "red" and "orange" complex groups is based solid biological and microbiological ground, it could still represent an over-simplification, because different micro-organisms possess different pathogenic and host interaction profiles. Therefore, it will be worthwhile for investigators to examine individual pathogen colonization profiles using a GWA approach.

With regard to the GWAS methodology, in brief, apart from our modest sample size for this type of investigation, there are three inherent biases. First, it is very likely that many "true association" signals are below the stringent threshold set for genome-wide significance or suggestive evidence of association. Second, due the low power of the study, it is also likely that the identified associations are overestimates of the true effect sizes of these loci. Third, our analysis of only common variants restricts our inference to common-only polymorphisms and ignores possibly rare (<5%) but large effect size SNPs. To address these biases future studies

need to be conducted replicating our findings, investigate for novel ones, and explore the role of rare variants.

Tooth loss due to CP is also another important source of bias. An example would be a patient with severe CP that lost all his teeth and in this investigation was classified as edentulous and was excluded from the dental ARIC analytical sample; or another patient who had his "most severely" affected teeth extracted few weeks before the clinical examination, and was subsequently diagnosed with moderate or no CP instead of severe CP. An analytical approach quantifying this source of bias would include sensitivity analyses that either adjust the genetic models for tooth loss, or impute 'likely CP diagnoses' based on certain patterns of tooth loss that are found among CP patients.

D. Potential impact and public health significance

1. Genetic discovery and applications

The significance of any new discovery in health research is measured by its clinical relevance, potential for prevention and public health context (37). The unraveling of the genetic component of human traits and diseases has been challenged in terms of its potential for immediate public health benefits (38). The contrary can be argued, for instance, for pharmaceutical research, were new drugs or vaccines may have immediate and dramatic global public health benefits (39). In this respect, the central role of epidemiology as the "quarterback" of steering translational research into application and public health practice has been increasingly emphasized (40). Other authors stress the fact that still not enough attention or efforts are invested in

translating novel findings in a manner that can have meaningful impacts in public health (41). The findings of this study have the potential to offer insights into the pathogenesis of CP, providing possible therapeutic and preventive avenues, but also benefit the study of other pathological conditions that have a host-microbiome interaction or immunological underpinning.

Despite the accelerated rates of gene discovery for common complex diseases, valid applications of GWAS findings in personalized health care are not yet evident (42, 43). Influences by the emerging "personalized medicine" market may subject individuals to unnecessary costs or even health risks (44); in this regard, the clinicians' and investigators' role in providing accurate and evidence-based counseling about the potential and the limitations of genetic testing, is critical However, regardless of immediate translational potential, discovery of gene variants associated with health or disease can lead to the discovery of previously unknown biological mechanisms and pathways. Intervening on the genetic makeup is not applicable to-date; however, development of drugs that could interact on these pathways is an important opportunity treatment and even prevention (vaccine development) of disease.

2. Therapeutic avenues for periodontitis

Despite the great strides in our understanding of the biology of periodontitis during the last decades, the therapeutic means available for CP have been long unchanged. Efforts in the prevention of CP and periodontal diseases in general have a strong behavioral component (45, 46), because meticulous oral hygiene, oral self-care

and dental attendance are the cornerstones of oral disease prevention. Periodontal therapy however, is needed for individuals who develop the disease, so that the supra- and sub-gingival microbial load can be reduced, the inflammatory response controlled, and the patient can maintain a periodontal health status via preventive efforts and a maintenance-recall protocol. Specifically, the maintenance of a microbial colonization profile that is host-compatible has been proposed as a "microbial treatment goal" for periodontal therapy (47)

The use of mechanical means for the microbial load reduction (conventional or surgical periodontal therapy based on root scaling and planning) remains unchallenged for the treatment of periodontitis (48), while adjunctive antibiotic therapy is becoming more frequent in severe or refractory cases of periodontitis (49, 50). The patterns of colonization by and proliferation of periodontal pathogens has been suggested to have a genetic (host) component (51), and this question was indeed interrogated in our study. Beyond that, the current stage of knowledge allows clinicians and investigators to expect that response to periodontal treatment varies greatly between individuals who harbor the same pathogenic microflora (52). The fact that non-uniform and often unsustainable results are to be expected by CP patients after treatment is a frequent observation (53), which has invited the exploration of individual susceptibility or treatment resistance biological markers.

In medicine, the microbiological component and etiology of disease is well-established and common sense. In dentistry, in spite of advances in oral microbiology and the increasing popularity of antimicrobial applications, the unexplored potential is immense (49). This is supported by a substantial variation in quantitative and

qualitative aspects of periodontal ecology between patients (54), and recent findings that link subgingival bacterial colonization profiles with gingival tissue gene expression (55). The incorporation of oral microbiology applications for the prevention and therapy of periodontitis, consistent with the “periodontal medicine” model, is likely to increase in the future (56-58).

The introduction of antimicrobials (59-61) or anti-inflammatory drugs (61, 62) has showed varying but generally promising results. These reports indicate that combinations of conventional and pharmaceutical treatment regimes are likely to produce the most predictable treatment outcomes in selected cases. However, it would be optimal if individual characteristics of host response could be incorporated in diagnosis and therapy to optimize treatment responses. For example, identification from the present GWA of susceptibility loci and potential functional genetic variants can be used to determine whether adjunctive antimicrobial or anti-inflammatory therapy would be indicated for biological subtypes of periodontitis (63).

Improvements in prevention of periodontitis would have significant and multi-level implications in the oral health and related quality of life of a substantial proportion of the adult population, and may be more important as demographic effects increase the life expectancy and mean age worldwide (64). Prevention of CP will contribute to the reduction of edentulousness rates (65), which constitutes a frequent form of functional and psychosocial impairment. Such advances in prevention could indeed stem from a better understanding of the disease pathogenesis, progression, and treatment or from the development of new preventive strategies, such as the envisioned vaccine against periodontitis (66). Moreover, it will reduce the

disease-related somatic and psychological distress and co-morbidities, as well as the financial and time loss impact to systems and individuals, due to its treatment or sequelae. Additionally, prevention of periodontitis has substantial public health implications when its links with systemic health are considered (48). One major and direct benefit would be, for instance, improved glycaemic control among patients with type 2 diabetes (67).

E. Replication of GWA results and relevance of identified loci

As noted above, replication of results of genome-wide association studies has virtually become a standard requirement for reporting (22, 38). Although successful replication of GWA finding provides some level of additional "confidence" that the results do not represent spurious associations, non-replication does not always imply a "false alarm". The "winner's curse" phenomenon, representing novel findings' failure to replicate in subsequent studies, as well as other methodological features of GWA studies such as stringent significance thresholds may be responsible for overestimation of the true underlying genetic effects (68). For this reason, GWAS are considered essentially "discovery attempts" (69, 70) and their results are subject to very conservative statistical criteria, as well as replication in independent samples.

1. Replication of prioritized SNPs of chronic periodontitis

The Health ABC dental cohort

Replication of the current study's findings was sought via collaboration with investigators of the Health Aging and Body Composition (ABC) study. The Health

ABC is a prospective study of well-functioning adults of ages 70-79 at baseline. White participants were recruited from a random sample of Medicare beneficiaries in the Pittsburgh, Pennsylvania, and Memphis, Tennessee. Eligibility criteria included age between 70 and 79, self-reported no difficulty of walking one-quarter mile or climbing 10 steps without resting, no difficulty performing basic activities of daily living, no reported need of an assistive walking device, and no active treatment for cancer (71). As part of the study year 2 and 3 follow-up clinical visits (1998-2000), a total of 1,133 white and African American participants were eligible and received complete dental and periodontal examinations. About two-thirds of this ancillary dental study's participants were white. Among whites, 21% of men and 14% of women were diabetic. With regard to smoking history, proportions of current/former smokers among men and women were 5%/61%, and 5%/32%, respectively. Replication was based on this subset of approximately 750 Health ABC participants (~400 men and ~350 women) that were subsequently genotyped using the Illumina Human1M Duo platform.

Replication of the dental ARIC GWA findings

Replication was sought for our prioritized SNPs based on the genome-wide association $P<5x10^{-6}$ threshold in 'crude' or 'adjusted' analyses. Fourteen SNPs met this threshold and therefore, effect estimates and P-values based derived from the Health ABC cohort GWAS were obtained. For these SNPs, the same coded and non-coded alleles were genotyped in both GWAS (Table A15). These estimates are presented in Tables A16 and A17. A same direction effect and a conventional $P<0.05$

threshold were set as for evidence of replication. Moreover, we took advantage of the opportunity to perform a meta-analysis of the reported effect estimates, as an additional approach in summarizing the effect of the reported risk loci (72). Meta-analytical pooled odds ratios and P-values were obtained using the METAL program (73), using an inverse-variance weighting method (74). Summary estimates (pooled odds ratio) were obtained for SNPs that did not depart substantially from effect homogeneity between the two cohorts (defined as homogeneity X^2 $P<0.2$).

The summary of the replication findings is presented in the Appendix, in Tables A16-A18. Noteworthy, only one locus associated with severe CP (7p15, rs2521634) met the $P<0.05$ and same effect direction replication criteria. Additionally, one locus associated with moderate CP (19p13.3, rs3826782) neared replication (P=0.06) and another (6p21.1, rs7762544) showed a same direction effect with no evidence of heterogeneity. Therefore, summary estimates based on inverse-variance weighted meta-analysis are presented in Table A18. Pooled odds ratios and associated summary P-values for these three loci were: 7p15 (rs2521634)—OR=1.49 (95% CI=1.28-1.73), $P=3.5x10^{-7}$; 19p13.3 (rs3826782)—OR=2.01 (95% CI=1.52-2.65), $P=8.2x10^{-7}$; 6p21.1 (rs7762544)—OR=1.40 (95% CI=1.24-1.59), $P=7.5x10^{-8}$.

Interestingly, one of these SNPs prioritized for severe CP (rs2521634) provided a replication signal also for moderate CP in Health ABC, whereas one moderate CP-associated SNP in ARIC (rs3826782) provided a replication signal for severe CP in Health ABC. It is possible that these risk loci indeed confer risk for both types of CP. However, this observation warrants more investigation to clarify whether systematic differences between the two study samples may be responsible for this difference (i.e.

tooth loss and retention, or periodontal treatment levels that may result to different phenotypical expression). An alternative explanation may be that these polymorphisms actually confer protection against periodontal pathology, and therefore may be associated with "periodontal health" when studying the two CP traits.

Relevance of identified and replicated loci

Presentation and discussion of all prioritized loci for the CP, EAL and "high" bacterial colonization traits are presented in the discussion sections of and supplemental material of manuscripts 1 and 2 (Sections 4.F, 4.J, 5.F and 5.J). This section offers a more detailed presentation of one replicated locus for severe CP, at 7p15.

The *NPY* (7p15) locus

This locus was marked by rs2521634, which was the top SNP in the region with $P=1.6x10^{-6}$ and located 47Kb downstream of *NPY*. The latter is a gene encoding a neuropeptide that is widely expressed in the central nervous system and has been suggested to function as an anxiolytic peptide that helps explain inter-individual variation in trait anxiety and resiliency to stress (75). Several reports have associated a Leu→Pro substitution in the *NPY* gene with significantly elevated serum triglycerides and lipids among Northern European populations (76, 77). The SNP responsible for this missense change (rs16139, predicted as 'benign' by PolyPhen-2)

was 10Kb upstream of rs2521634 and not in linkage disequilibrium with it (D'=1.00; r^2=0.01).

An early immunohistological Swedish study among by Barr-Agholme and colleagues observed higher neuropeptide Y levels and sensory hyper-innervation in the gingival tissue of Down syndrome (DS) patients, attributing this finding to the gingival inflammatory reaction rather that DS itself (78). A 2004 critical review by Lundy and Linden (79) emphasized the role of the nervous system in the pathophysiology of peripheral inflammation, suggesting a neurogenic inflammatory component for periodontitis. These authors suggested that neurogenic inflammation is in fact a protective mechanism, with NPY having a pivotal role in the cascade of relevant chemical activities.

A recent case-control study by Lundy and colleagues (80) provided experimental evidence consistent with this hypothesis. Elevated NPY levels were found in the gingival crevicular fluid of periodontitis compared to healthy sites. Additional evidence supports a role of the sympathetic nervous system and NPY in the modulation of periodontal inflammation and bone resorption (81, 82). Interestingly, inhibition of nerve growth factor (NGF), which is known to up-regulate neuropeptides in sensory neurons, via systemic anti-NGF was recently shown to be associated with reduced Interleukin 1beta and bone resorption in a rat periodontitis model (83). In sum, combined with our GWAS results, there is accumulating evidence pointing towards the role of the nervous system and neuropeptide messengers (with NPY being an obvious candidate one) across the axis of a "neurological inflammatory reflex" in the cascade of events that lead to periodontitis.

F. Future research directions

1. Improvement over candidate-gene approaches

Candidate-gene studies have been limited in their ability to identify risk loci or genes for common complex diseases. In general, candidate-gene approaches have had better results when studying “highly penetrant” diseases, where a “disease-gene” may exist. Successful candidate-gene studies are typically based on strong prior knowledge or mechanistic theory, as is the case of inflammatory mediator gene markers and cytokines in the context of inflammatory diseases, such as ulcerative colitis (84) and CP (85). In the case of CP in particular, such studies have focused on inflammatory mediator molecules and receptors such as the interleukin gene cluster. Despite the obvious relevance and replication of this finding across many candidate-gene studies, it should be expected that multiple genes are implicated in the pathogenesis of CP, each conferring likely small to modest excess risk. Although previously reported candidate-gene study targets were not marked by strong signals in our GWAS we explored several loci that contain such candidate genes, including *IL1A, IL1RN*, *IL6, TLR4, MMP2*, *GLT6D1, FCGR3A*, and *VDR* with visualizations that are presented in Figures A30-A33. Moreover, we did identify few loci that were contained genes related to previous candidate-gene reports. For example, the 9p24 locus was identified for “high red” complex bacterial colonization and included *IL33*. Additionally, one locus on chromosome 19 that was identified for severe CP contained a tumor necrosis family (*TNFSF14*) and a complement (*C3*) gene, which are well-documented immune- and periodontitis-related genes (86-90). One of the

prioritized loci on chromosome 6 (rs17792917) for the "extent" of attachment loss trait has been reported as an eQTL for another TNF-family gene on chromosome 13 (*TNFRSF19*), which has been reported as a susceptibility locus for lung cancer (91) and vascular dementia (92). Finally, the risk locus on 2p23.3 that emerged for *Pg* "high" colonization is an eQTL for multiple genes including *IL17RB* on chromosome 3. We anticipate that future candidate-gene studies will embark in testing and potentially replicating our GWA findings.

2. Beyond genome-wide association analyses

The analytical methods of GWAS data are constantly evolving, and may allow a future extension of the present analysis via implementation of novel methodologies. For example, despite the initial "agnostic" scan, there is an opportunity to prioritize candidate SNPs for analysis in the GWA context, when for instance epistatic effects are hypothesized or suspected (32, 93). Other examples of novel but not well-established approaches include attempts to link prior knowledge (of validated polymorphisms) with analytical strategies (94), use of pathway-clustering analysis of genetic signals (95), efficient approaches to explore gene-gene and gene-environment interactions (96, 97), and the incorporation of expression quantitative trait loci (eQTL) with GWAS results (98, 99).

It is becoming increasingly understood that major advances in our understanding of the genetic underpinning of human health and disease can be realized by efficient comprehensive, efficient and inexpensive interrogations of genomes, transcriptomes and interactomes (100). Next-generation sequencing studies (101-103) have recently

emerged and are promising in this respect, with exome sequencing investigations providing a high rate of novel findings (104, 105). A particular strength of whole-exome sequencing studies is that they offer opportunities for both disease gene discovery as well as clinical diagnosis, as protein coding genes harbor the vast majority of large-effect disease-related mutations (106).

Finally, several additional explorations of the data presented in this dissertation may offer novel insights in the study of CP. The consideration of gene expression modulation by proximal (*cis*) or distal (*trans*) eQTLs, as noted above (98, 99) can provide information that may be functionally relevant. For example, one of the three replicated loci, 19p13.3 (marked by rs3826782) has been reported as an eQTL for *GPR113* (G-protein coupled receptor 113) on chromosome 2p23.3, which has functions on the neuropeptide signaling pathway as well as signal transducer activity; pathways relevant to the other prioritized locus on 7p15, where neuropeptide Y (*NPY*) is located. Moreover, *GPR113* has been found to be particularly expressed intra-orally, in the taste receptor cells (107).

Moreover, examining larger sets of prioritized loci in 'pathway analysis' networks may reveal patterns of association that are otherwise difficult to detect by 'single-locus' explorations. Examples of such exploratory visualizations of functional and physical interactions between genes that were identified from our analyses are presented in the Appendix Figures A28 and A29. Although these networks arguably contain a lot of "noise", empirical inspection of represented functional pathways and physical interactions can be informative. In this network of interactions (figure A28), some of these well-represented pathways were collagen, protein phosphatase 2A

binding, extracellular matrix, positive regulation of transcription from RNA polymerase II promoter, epithelial to mesenchymal transition, platelet activation, Wnt receptor signaling pathway, cellular response to growth factor stimulus, regulation of fibroblast proliferation and positive regulation of kinase activity.

3. Future studies

As explained in the previous sections, the present study can be considered as a 'fist look' into the genetic component of chronic periodontitis using a genome-wide association approach. Within the limitations of the analyses presented here, this work offers new information about candidate loci associated with CP. There is a large number of follow-up research and different avenues that can and should be followed after the conduct of a GWAS (108, 109). Future studies in this field are needed to replicate and validate these associations, discover new ones, elucidate their functional role and mechanistic pathways, determine their potential for diagnostic and therapeutic applications, and more.

To be more specific, future GWAS examining the three-level CP classification are warranted to replicate the associations with loci that our study reported. Although GWAS of bacterial colonization profiles are harder to realize, replication of the high periodontal pathogen colonization findings will add to our understanding of the contribution of genetic factors to host-microbiome interactions. Because GWAS are generally under-powered to detect most small to modest genetic effects, future studies of equal size may discover new loci that did not meet the suggestive association threshold in our investigation. However, larger GWAS or

pooling of existing cohorts may be necessary to enable the detection of small effects and GxE or GxG interactions. Factors such as sex, smoking and diabetes should be strong candidates as 'environmental exposures' for these interaction analyses. We emphasize the need and importance of conducting GWAS in populations other than whites of European descent, as evidence of the genetic component of CP among populations of African, Asian and American descent are lacking. Such studies in diverse populations can provide insights about loci that were in fact identified in our study population of whites, or discover novel ones.

Identifying the causal variant(s) in the loci that we report is also an important step that future studies should address (110). From our GWAS results it cannot be inferred that the SNPs that were identified tagging the prioritized loci are the causal variants. It is possible that non-genotyped and non-imputed, perhaps rare (<5%) variants are the risk-conferring ones. Fine-mapping represents one strategy that can be implemented in future studies seeking to identify the risk variants and causal alleles in these or other candidate loci for CP. However, some opportunities to interrogate the contribution of rare variants will emerge with the imputation of our existing genotype data using the '1000 genomes' haplotypes, a strategy that will offer approximately 10 million markers for GWAS, including a substantial proportion of rare ones. It is not unrealistic to anticipate that these strategies will eventually be superseded by whole genome sequencing, which will become progressively less costly and time-consuming in the next decade.

Subsequent experimental or association studies linking GWAS data with gene expression information can also provide information about the functional relevance

and candidate mechanistic pathways that the prioritized genes operate on. Experimental studies "knocking-out" the new candidate genes in animal models will be needed to clarify the role of GWAS discoveries. The use of microarrays to interrogate gene expression can also provide tissue-specific insights (111). As more data become available with regard to gene expression regulation by quantitative trait loci on multiple tissues, future studies may be able to annotate our findings to such *cis*- or *trans*-acting eQTLs (112).

Future studies examining traits other than the three-level CDC disease taxonomy that we employed may offer additional insights into the genetic component of CP. Disease indicators or "proxies" are less valid in the ascertainment of cases, but they may be cost-effective and more feasible than gold-standard clinical examinations. It may be possible for indicators such as number of remaining teeth, radiographically measured alveolar bone loss, and self-reported diagnosis of CP to be obtained for population-based samples or existing cohorts with genotype data, thus enabling more GWAS of CP.

Furthermore, genome-wide interrogations of endophenotypes, which represent physiologic indicators, biochemical assays, or other "intermediate" traits relevant to the disease under study (113, 114) may offer additional opportunities to examine the genetic influences for CP. These traits have the benefit of being 'biologically proximal' to the hypothesized genetic effects, whereas the phenotype expression of complex and polygenic diseases is likely to be the result of many genetic and non-genetic, measured, un-measured and unobservable factors. The bacterial colonization profile that we examined as part of our Specific Aim 2 can be considered as an

endophenotype of CP. Others may include *in vivo* quantified inflammatory markers such as gingival crevicular fluid Interleukin levels, serum levels of antibodies to periodontal pathogens, markers of oxidative stress, and more.

Extending the concept of the study of intermediate CP phenotypes, it can be argued that the identified or the still undiscovered risk loci for CP confer excess risk for conditions other than CP. Certain molecular functions and pathways may be common in the pathogenesis of a number of diseases. This is indicative of pleiotropy, which is the phenomenon of a single gene influencing multiple traits (115). Future studies can investigate common risk loci for CP and other, frequently associated conditions, such as diabetes and cardiovascular disease. Moreover, simultaneous examinations of a wide array of phenotypes for a given genotype, in the context of phenome-wide (116, 117) association studies (PheGWAS) are efficient strategies that are gaining popularity and may include CP in the future.

Finally, incorporating the current GWAS findings into epidemiologic and clinical investigations will be an important "research translation" step. One can envisage how novel candidate genes can be considered in investigations of CP risk or treatment response. Genes or promoter regions in the candidate loci that we report, such as *NPY*, could be evaluated as CP risk indicators or prognostic markers CP treatment response, alone or in combination with genetic markers that are already being tested as such (i.e. the interleukin family gene cluster). The identification of a group of genetic markers that have a good combined positive predictive value could in the future be used in the design of genetic tests that would inform clinicians and patients of their CP risk or their likelihood to respond favorably to various therapeutic

options. Such investigations could evaluate for example whether CP patients who are carriers of certain genotype of *NIN*, *EMR1*, *NPY*, *FBXO38*, *TBC1D1*, *CAMTA1* or *GRID1* have better therapeutic results when treated with conventional-surgical, antimicrobial, or anti-inflammatory periodontal treatment, or combinations of these.

References

1. Hunter DJ, Thomas G, Hoover RN, Chanock SJ. Scanning the horizon: what is the future of genome-wide association studies in accelerating discoveries in cancer etiology and prevention? Cancer Causes Control 2007;18:479-84.

2. Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA. Accuracy of NHANES Periodontal Examination Protocols. J Dent Res 2010;89:1208-13.

3. Savage A, Eaton KA, Moles DR, Needleman I. A systematic review of definitions of periodontitis and methods that have been used to identify this disease. J Clin Periodontol 2009;36:458-67.

4. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

5. Beck JD, Slade G, Offenbacher S. Oral disease, cardiovascular disease and systemic inflammation. Periodontol 2000 2000;23:110-20.

6. Bizzarro S, Nicu EA, van der Velden U, Laine ML, Loos BG. Association of serum immunoglobulin G (IgG) levels against two periodontal pathogens and prothrombotic state: a clinical pilot study. Thromb J 2010;8:16.

7. Nakano K, Wada K, Nomura R, et al. Characterization of aortic aneurysms in cardiovascular disease patients harboring Porphyromonas gingivalis. Oral Dis 2011;17:370-8.

8. Launay JM, Del Pino M, Chironi G, et al. Smoking induces long-lasting effects through a monoamine-oxidase epigenetic regulation. PLoS One 2009;4:e7959.

9. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003;33:245-54.

10. Iacopino AM. Periodontitis and diabetes interrelationships: role of inflammation. Ann Periodontol 2001;6:125-37.

11. Lalla E, Papapanou PN. Diabetes mellitus and periodontitis: a tale of two common interrelated diseases. Nat Rev Endocrinol 2011. doi: 10.1038/nrendo.2011.106.

12. Joshi N, Caputo GM, Weitekamp MR, Karchmer AW. Infections in patients with diabetes mellitus. N Engl J Med 1999;341:1906-12.

13. Beck JD, Löe H. Epidemiological principles in studying periodontal diseases. Periodontol 2000 1993;2:34-45.

14. Papapanou PN, Wennström JL, Gröndahl K. Periodontal status in relation to age and tooth type. A cross-sectional radiographic study. J Clin Periodontol 1988;15:469-78.

15. Ong G. Periodontal disease and tooth loss. Int Dent J 1998;48:233-8.

16. Preshaw PM. Definitions of periodontal disease in research. J Clin Periodontol 2009;36:1-2.

17. Nothnagel M, Ellinghaus D, Schreiber S, Krawczak M, Franke A. A comprehensive evaluation of SNP genotype imputation. Hum Genet 2009;125:163-71.

18. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

19. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061-73.

20. Park JH, Wacholder S, Gail MH, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat Genet 2010;42:570-5.

21. Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. Am J Hum Genet 2007;80:605-15.

22. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol 2008;37:120-32.

23. Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 2010;11:446-50.

24. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444:444-54.

25. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.

26. Ritchie MD. Using biological knowledge to uncover the mystery in the search for epistasis in genome-wide association studies. Ann Hum Genet 2011;75:172-82.

27. Alexander RP, Fang G, Rozowsky J, Snyder M, Gerstein MB. Annotating non-coding regions of the genome. Nat Rev Genet 2010;11:559-71.

28. Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223-7.

29. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol 2010;8:e1000294.

30. Panagiotou OA, Evangelou E, Ioannidis JP. Genome-wide significant associations for variants with minor allele frequency of 5% or less--an overview: A HuGE review. Am J Epidemiol 2010;172:869-89.

31. Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet 2010;11:773-85.

32. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. Am J Hum Genet 2010;86:6-22.

33. Luo L, Peng G, Zhu Y, Dong H, Amos CI, Xiong M. Genome-wide gene and pathway analysis. Eur J Hum Genet 2010;18:1045-53.

34. Barros SP, Offenbacher S. Epigenetics: connecting environment and genotype to phenotype and disease. J Dent Res 2009;88:400-8.

35. Feinberg AP. Epigenomics reveals a functional genome anatomy and a new approach to common disease. Nat Biotechnol 2010;28:1049-52.

36. Bobetsis YA, Barros SP, Lin DM, et al. Bacterial infection promotes DNA hypermethylation. J Dent Res 2007;86:169-74.

37. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998;317:465-8.

38. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

39. Drews J. Drug discovery: a historical perspective. Science 2000;287:1960-4.

40. Khoury MJ, Gwinn M, Ioannidis JP. The emergence of translational epidemiology: from scientific discovery to population health impact. Am J Epidemiol 2010;172:517-24.

41. Hiatt RA. Invited commentary: The epicenter of translational science. Am J Epidemiol 2010;172:525-7.

42. VanderWeele TJ. Genetic self knowledge and the future of epidemiologic confounding. Am J Hum Genet 2010;87:168-72.

43. Ioannidis JP. Personalized genetic prediction: too limited, too expensive, or too soon? Ann Intern Med 2009;150:139-41.

44. Khoury MJ, McBride CM, Schully SD, et al; Centers for Disease Control and Prevention. The Scientific Foundation for personal genomics: recommendations from a National Institutes of Health-Centers for Disease Control and Prevention multidisciplinary workshop. Genet Med 2009;11:559-67.

45. Petersen PE, Ogawa H. Strengthening the prevention of periodontal disease: the WHO approach. J Periodontol 2005;76:2187-93.

46. Kallio PJ. Health promotion and behavioral approaches in the prevention of periodontal disease in children and adolescents. Periodontol 2000 2001;26:135-45.

47. Teles RP, Haffajee AD, Socransky SS. Microbiological goals of periodontal therapy. Periodontol 2000 2006;42:180-218.

48. Loesche WJ, Grossman NS. Periodontal disease as a specific, albeit chronic, infection: diagnosis and treatment. Clin Microbiol Rev 2001;14:727-52.

49. Van Winkelhoff AJ, Winkel EG. Microbiological diagnostics in periodontics: biological significance and clinical validity. Periodontol 2000 2005;39:40-52.

50. Herrera D, Sanz M, Jepsen S, Needleman I, Roldán S. A systematic review on the effect of systemic antimicrobials as an adjunct to scaling and root planing in periodontitis patients. J Clin Periodontol 2002;29:136-59.

51. Laine ML, Farré MA, González G, et al. Polymorphisms of the interleukin-1 gene family, oral microbial pathogens, and smoking in adult periodontitis. J Dent Res 2001;80:1695-9.

52. Van Dyke TE, Sheilesh D. Risk factors for periodontitis. J Int Acad Periodontol 2005;7:3-7.

53. Behle JH, Sedaghatfar MH, Demmer RT, et al. Heterogeneity of systemic inflammatory responses to periodontal therapy. J Clin Periodontol 2009;36:287-94.

54. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

55. Papapanou PN, Behle JH, Kebschull M, et al. Subgingival bacterial colonization profiles correlate with gingival tissue gene expression. BMC Microbiol 2009;9:221.

56. Williams RC, Offenbacher S. Periodontal medicine: the emergence of a new branch of periodontology. Periodontol 2000 2000;23:9-12.

57. Pizzo G, Guiglia R, Lo Russo L, Campisi G. Dentistry and internal medicine: from the focal infection theory to the periodontal medicine concept. Eur J Intern Med 2010;21:496-502.

58. Stamm JW. Periodontal diseases and human health: new directions in periodontal medicine. Ann Periodontol 1998;3:1-2.

59. Bland PS, Goodson JM, Gunsolley JC, et al. Association of antimicrobial and clinical efficacy: periodontitis therapy with minocycline microspheres. J Int Acad Periodontol 2010;12:11-9.

60. Sakellari D, Ioannidis I, Antoniadou M, Slini T, Konstantinidis A. Clinical and microbiological effects of adjunctive, locally delivered chlorhexidine on patients with chronic periodontitis. J Int Acad Periodontol 2010;12:20-6.

61. Kirkwood KL, Cirelli JA, Rogers JE, Giannobile WV. Novel host response therapeutic approaches to treat periodontal diseases. Periodontol 2000 2007;43:294-315.

62. Pinho Mde N, Pereira LB, de Souza SL, et al. Short-term effect of COX-2 selective inhibitor as an adjunct for the treatment of periodontal disease: a clinical double-blind study in humans. Braz Dent J 2008;19:323-8.

63. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

64. Mersel A, Call R, Mann J. Demographic trends of aging--application to gerodontology. Gerodontology 1987;6:9-15.

65. Hugoson A, Norderyd O. Has the prevalence of periodontitis changed during the last 30 years? J Clin Periodontol 2008;35:338-45.

66. Page RC. Vaccination and periodontitis: myth or reality. J Int Acad Periodontol 2000;2:31-43.

67. Koromantzos PA, Makrilakis K, Dereka X, Katsilambros N, Vrotsos IA, Madianos PN. A randomized, controlled trial on the effect of non-surgical periodontal therapy in patients with type 2 diabetes. Part I: effect on periodontal status and glycaemic control. J Clin Periodontol 2011;38:142-7.

68. Kraft P. Curses--winner's and otherwise--in genetic epidemiology. Epidemiology 2008;19:649-51.

69. Durick K, Mendlein J, Xanthopoulos KG. Hunting with traps: genome-wide strategies for gene discovery and functional analysis. Genome Res 1999;9:1019-25.

70. Huo D, Olopade OI. Interpretation of genome-wide association study results. Oncology (Williston Park) 2010;24:643, 646.

71. Katancik JA, Kritchevsky S, Weyant RJ, et al. Periodontitis and airway obstruction. J Periodontol. 2005;76:2161-7.

72. Kavvoura FK, Ioannidis JP. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. Hum Genet 2008;123:1-14.

73. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-1.

74. Zhou B, Shi J, Whittemore AS. Optimal methods for meta-analysis of genome-wide association studies. Genet Epidemiol 2011;35:581-91.

75. Zhou Z, Zhu G, Hariri AR, et al. Genetic variation in human NPY expression affects stress response and emotion. Nature 2008;452:997-1001.

76. Uusitupa MI, Karvonen MK, Pesonen U, Koulu M. Neuropeptide Y: a novel link between the neuroendocrine system and cholesterol metabolism. Ann Med 1998;30:508-10.

77. Karvonen MK, Pesonen U, Koulu M, et al. Association of a leucine(7)-to-proline(7) polymorphism in the signal peptide of neuropeptide Y with high serum cholesterol and LDL cholesterol levels. Nat Med 1998;4:1434-7.

78. Barr-Agholme M, Modéer T, Luthman J. Immunohistological study of neuronal markers in inflamed gingiva obtained from children with Down's syndrome. J Clin Periodontol 1991;18:624-33.

79. Lundy FT, Linden GJ. Neruopeptides and neurogenic mechanisms in oral and periodontal inflammation. Crit Rev Oral Biol Med 2004;15:82-98.

80. Lundy FT, El Karim IA, Linden GJ. Neuropeptide Y (NPY) and NPY Y1 receptor in periodontal health and disease. Arch Oral Biol 2009;54:258-62.

81. Haug SR, Heyeraas KJ. Modulation of dental inflammation by the sympathetic nervous system. J Dent Res 2006;85:488-95.

82. Abd El-Aleem SA, Morales-Aza BM, Donaldson LF. Sensory neuropeptide mRNA up-regulation is bilateral in periodontitis in the rat: a possible neurogenic component to symmetrical periodontal disease. Eur J Neurosci 2004;19:650-8.

83. Gaspersic R, Kovacic U, Glisovic S, Cör A, Skaleric U. Anti-NGF treatment reduces bone resorption in periodontitis. J Dent Res 2010;89:515-20.

84. Mansfield JC, Holden H, Tarlow JK, et al. Novel genetic association between ulcerative colitis and the anti-inflammatory cytokine interleukin-1 receptor antagonist. Gastroenterology 1994;106:637-42.

85. Kornman KS, Crane A, Wang HY, et al. The interleukin-1 genotype as a severity factor in adult periodontal disease. J Clin Periodontol 1997;24:72-7.

86. Hosokawa Y, Hosokawa I, Ozaki K, Nakae H, Matsuo T. TNFSF14 coordinately enhances CXCL10 and CXCL11 productions from IFN-gamma-stimulated human gingival fibroblasts. Mol Immunol 2010;47:666-70.

87. Hosokawa Y, Hosokawa I, Ozaki K, Nakanishi T, Nakae H, Matsuo T. Tea polyphenols inhibit IL-6 production in tumor necrosis factor superfamily 14-stimulated human gingival fibroblasts. Mol Nutr Food Res 2010l;54:S151-8.

88. Belstrøm D, Holmstrup P, Damgaard C, et al. The atherogenic bacterium Porphyromonas gingivalis evades circulating phagocytes by adhering to erythrocytes. Infect Immun 2011;79:1559-65.

89. Potempa M, Potempa J, Kantyka T, et al. Interpain A, a cysteine proteinase from Prevotella intermedia, inhibits complement by degrading complement factor C3. PLoS Pathog 2009;5:e1000316.

90. Cutler CW, Arnold RR, Schenkein HA. Inhibition of C3 and IgG proteolysis enhances phagocytosis of Porphyromonas gingivalis. J Immunol 1993;151:7016-29.

91. Hu Z, Wu C, Shi Y, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. Nat Genet 2011;43:792-6.

92. Kong M, Kim Y, Lee C. A strong synergistic epistasis between FAM134B and TNFRSF19 on the susceptibility to vascular dementia. Psychiatr Genet 2011;21:37-41.

93. Lewinger JP, Conti DV, Baurley JW, Triche TJ, Thomas DC. Hierarchical Bayes prioritization of marker associations from a genome-wide association scan for further investigation. Genet Epidemiol 2007;31:871-82.

94. Tintle N, Lantieri F, Lebrec J, Sohns M, Ballard D, Bickeböller H. Inclusion of a priori information in genome-wide association analysis. Genet Epidemiol 2009;33:S74-80.

95. Eleftherohorinou H, Wright V, Hoggart C, et al. Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. PLoS One 2009;4:e8068.

96. Zhang X, Huang S, Zou F, Wang W. TEAM: efficient two-locus epistasis tests in human genome-wide association study. Bioinformatics 2010;26:i217-27.

97. Li L, Yu M, Jason RD, et al. A mixture model approach in gene-gene and gene-environmental interactions for binary phenotypes. J Biopharm Stat 2008;18:1150-77.

98. Nica AC, Montgomery SB, Dimas AS, et al. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet 2010;6:e1000895.

99. Nica AC, Dermitzakis ET. Using gene expression to investigate the genetic basis of complex disorders. Hum Mol Genet 2008;17:R129-34.

100. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 2008;26:1135-45.

101. Schuster SC. Next-generation sequencing transforms today's biology. Nat Methods 2008;5:16-8.

102. Mardis ER. The impact of next-generation sequencing technology on genetics. Trends Genet 2008;24:133-41.

103. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 2008;9:387-402.

104. Biesecker LG. Exome sequencing makes medical genomics a reality. Nat Genet 2010;42:13-4.

105. Biesecker LG, Shianna KV, Mullikin JC. Exome sequencing: the expert view. Genome Biol 2011;12:128.

106. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009;106:19096-101.

107. Khoury MJ, Millikan R, Little J, Gwinn M. The emergence of epidemiology in the genomics age. Int J Epidemiol 2004;33:936-44.

108. Witte JS. Genome-wide association studies and beyond. Annu Rev Public Health 2010;31:9-20.

109. McCarthy MI, Hirschhorn JN. Genome-wide association studies: past, present and future. Hum Mol Genet 2008;17:R100-1.

110. Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H. Interpretation of association signals and identification of causal variants from genome-wide association studies. Am J Hum Genet 2010;86:730-42.

111. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844-7.

112. Petretto E, Mangion J, Dickens NJ, et al. Heritability and tissue specificity of expression quantitative trait loci. PLoS Genet 2006;2:e172.

113. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003;160:636-45.

114. Huang P, Kraja AT, Tang W, et al. Factor relationships of metabolic syndrome and echocardiographic phenotypes in the HyperGEN study. J Hypertens 2008;26:1360-6.

115. He X, Zhang J. Toward a molecular understanding of pleiotropy. Genetics 2006;173:1885-91.

116. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. Bioinformatics 2010;26:1205-10.

117. Pendergrass SA, Brown-Gentry K, Dudek SM, et al. The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. Genet Epidemiol 2011;35:410-22.

APPENDIX

Table A1. Distribution in the total sample and bivariate associations of periodontal diagnosis (CDC classification) and periodontal attachment loss (extent score: percent of sites with ≥3mm attachment loss) with brushing and flossing frequency, dental attendance reason and frequency, and education level among the Caucasian participants of the Dental ARIC study, for whom clinical and genetic data were matched after exclusions (n=4,610)

		CDC Periodontitis classification			Attachment loss (≥3mm)
	Total (n, column %)	Health-mild (n, row %)	Moderate (n, row %)	Severe (n, row %)	Extent score* mean(SD); median†
Total sample	4610 (100)	1864 (40)	1961 (43)	785 (17)	21.1 (21.2); 14
Brushing frequency[1,2] (yesterday)					
One time or none	1350 (29)	482 (36)	592 (44)	276 (20)	24.0(23.6); 17
Two or more	3239 (71)	1374 (42)	1361 (42)	504 (16)	19.9(20.1); 13
Flossed last week[1,2]					
Did not floss	1464 (32)	506 (35)	641 (44)	317 (22)	26.1(24.4); 18
Flossed one or more times	3124 (68)	1349 (43)	1312 (42)	463 (15)	18.8(19.2); 13
Dental visit reason[1,2]					
Regular visits	3809 (83)	1609 (42)	1612 (42)	588 (15)	19.1(19.2); 13
Only when problem	780 (17)	247 (32)	341(44)	192 (25)	31.3(27.3); 23
Last dental visit[1,3]					
<6 months ago	3174 (69)	1312 (41)	1333 (42)	529 (17)	19.8(19.8); 14
6-24 months	1031 (22)	427 (41)	454 (44)	150 (14)	20.8(21.4); 14
More than 2 years ago	381 (8)	117 (31)	163 (43)	101 (27)	32.9(28.0); 25
Education level[1,3]					
Basic (0-11 years)	433 (9)	136 (31)	195 (45)	102 (24)	30.5(27.4); 21
Intermediate (12-15 years)	2102 (56)	838 (40)	904 (43)	360 (17)	21.4(21.3); 14
Advanced (16+ years)	2069 (45)	886 (43)	860 (42)	323 (16)	18.9(19.1); 13

*third molars were included; †rounded to the closest integer; [1]Chi-square test of equivalence between strata of periodontitis P<0.05; [2]Median test of stratum-specific attachment loss estimates P<0.05; [3]Kruskal-Wallis test of stratum-specific attachment loss or covariate estimates P<0.05

Table A2. Distribution in the total sample and bivariate associations of periodontal diagnosis (CDC classification) and periodontal attachment loss (extent score: percent of sites with ≥3mm attachment loss) with quintiles of age and BMI among the Caucasian participants of the Dental ARIC study, for whom clinical and genetic data were matched after exclusions (n=4,610)

		CDC Periodontitis classification			Attachment loss (≥3mm)
	Total (n; median)	Health-mild (n, row %)	Moderate (n, row %)	Severe (n, row %)	Extent score* mean(SD); median†
Total sample	4610	1864 (40)	1961 (43)	785 (17)	21.1 (21.2); 14
Age[1,2] (years, quintiles; mean)					
Q1 (55.5)	1054; 56	504 (48)	403 (38)	147 (14)	20.0 (18.7); 11
Q2 (59.5)	1026; 60	445 (43)	425 (41)	156 (15)	18.8 (19.2); 13
Q3 (63.0)	724; 63	301 (42)	288 (40)	135 (19)	21.6 (21.9); 15
Q4 (66.5)	890; 66	311 (35)	421 (47)	158 (18)	23.6 (23.0); 16
Q5 (70.9)	916; 71	303 (33)	424 (46)	189 (21)	25.6 (22.8); 19
BMI[1,2] (kg/cm^2, quintiles; mean)					
Q1 (22.1)	922; 22.5	445 (48)	346 (38)	131 (14)	18.9 (20.6); 11
Q2 (25.3)	922; 25.3	389 (42)	386 (42)	147 (16)	20.3 (20.7); 14
Q3 (27.5)	919; 27.5	368 (40)	391 (43)	160 (17)	21.5 (21.6); 14
Q4 (30.1)	923; 30.0	324 (35)	421 (46)	178 (19)	22.9 (22.0); 17
Q5 (35.6)	918; 34.7	337 (37)	413 (45)	168 (18)	21.9 (21.1); 15

*third molars were included; †rounded to the closest integer; [1]Chi-square test of equivalence between strata of periodontitis $P<0.05$; [2]Kruskal-Wallis test of stratum-specific attachment loss or covariate estimates $P<0.05$

Table A3. Distribution and bivariate associations of attachment loss (quintile-categorical and three-level classification of extent score) with periodontal diagnosis (CDC classification) among the Caucasian participants of the Dental ARIC study, for whom clinical and genetic data were matched after exclusions (n=4,610)

		CDC Periodontitis classification			Attachment loss (≥3mm)
Attachment loss	Total (n, column %)	Health-mild (n, row %)	Moderate (n, row %)	Severe (n, row %)	Extent score* mean(SD); median†
Total sample	4610	1864 (40)	1961 (43)	785 (17)	21.1 (21.2); 14
Quintile-categorical					
Q1	922	892 (97)	29 (3)	1 (0)	2.3 (1.5); 2
Q2	929	634 (68)	291 (31)	4 (0)	7.6 (1.6); 8
Q3	918	262 (29)	611 (67)	45 (5)	14.3 (2.3); 14
Q4	944	61 (6)	622 (70)	221 (23)	25.1 (4.4); 25
Q5	897	15 (2)	368 (41)	514 (57)	57.3 (19.2); 52
3-level classification					
None-mild (<10%)	1760 (38)	1480 (84)	276 (16)	4 (0)	4.7 (2.9); 5
Moderate (10-<30%)	1769 (38)	365 (21)	1197 (68)	207 (12)	18.0 (5.5); 17
Severe (≥30%)	1081 (23)	19 (2)	488 (45)	574 (53)	52.9 (20.0); 46

*third molars were included; †rounded to the closest integer

Table A4. Median (interquartile range, IQR) of sub-gingival microbial counts of eight periodontal pathogens obtained via DNA checkerboard immunoblotting among a subset of the Dental ARIC participants (n=1020)

	P *gingivalis*	T *forsythia*	T *denticola*	P *intermedia*	P *nigriscens*	C *rectus*	F *nucleatum*	A *actinom.*
Total sample(n)	1020	1020	1017	1018	1020	1018	1012	1020
	945 (3061)	1314 (4368)	2665 (9273)	4270 (15432)	3185 (12758)	2524 (19676)	5605 (32326)	2044 (5592)
Sex								
Females	806 (2716)	1248 (4695)	2788 (9922)	3579 (15930)	3179 (12834)	2260 (10252)	5994 (30347)	1891 (5298)
Males	1162 (3341)	1435 (4188)	2537 (8968)	5074 (15009)	3191 (12683)	2736 (11230)	5355 (32393)	2257 (5762)
Smoking status								
Never	987 (2881)	1184 (4777)	2647 (9895)	4164 (18039)	3146 (14186)	2389 (10985)	5502 (39732)	2129 (6476)
Past- light	1351 (3927)	1558 (4268)	3252 (8968)	5204 (15480)	3514 (13292)	3230 (13005)	7088 (33977)	2138 (5906)
Past- heavy	725 (2470)	1116 (3732)	1869 (8043)	3200 (11721)	2567 (9438)	1827 (06974)	3811 (23385)	1651 (4309)
Current- light	2358 (2935)	2671 (5696)	1718 (13022)	5417 (13724)	3216 (13207)	3375 (8652)	6696 (35865)	2491 (5933)
Current– heavy	1003 (3082)	1318 (6286)	3503 (11592)	6858 (15881)	2726 (13957)	3649 (10491)	6950 (29101)	2250 (4306)
Diabetic status								
Healthy	897 (2996)	1242 (4079)	2574 (9149)	4118 (14895)	2924 (12601)	2499 (10467)	5433 (28535)	2035 (5539)
Diabetes mellitus	1305 (3229)	1849 (6770)	2980 (10907)	6031 (17418)	4595 (15681)	2707 (16178)	6853 (42203)	2398 (6474)
Tooth loss								
0-9 lost teeth	862 (2658)	955 (3280)	2154 (8016)	3466 (13724)	2718 (10288)	1978 (8310)	4621 (26557)	1810 (4818)
10-18 lost teeth	783 (3195)	1445 (5519)	2249 (10392)	3320 (16363)	3407 (14219)	2679 (12632)	5518 (32747)	1997 (5031)
19-32 lost teeth	2247 (5848)	3556 (9088)	6192 (19980)	10486 (32331)	6021 (25088)	4947 (25396)	14486 (48825)	4185 (8360)
No. teeth	-0.20	-0.14	-0.11	-0.10	-0.08	-0.12	-0.10	-0.13
rho (95% CL)	(-0.16, -0.4)	(-0.20, -0.08)	(-0.18, -0.6)	(-0.16, -0.4)	(-0.14, -0.02)	(-0.19, -0.06)	(-0.16, -0.04)	(-0.19, -0.07)
Age	0.01	0.07	0.02	0.01	0.04	0.04	0.03	0.03
rho (95% CL)	(-0.05, 0.07)	(0.01, 0.13)	(-0.03, 0.08)	(-0.06, 0.07)	(-0.02, 0.10)	(-0.02,0.10)	(-0.03,0.10)	(-0.03, 0.09)
BMI	0.01	0.01	0.02	0.00	0.03	0.01	0.00	0.03
rho (95% CL)	(-0.05, 0.07)	(-0.06, 0.07)	(-0.04, 0.08)	(-0.07, 0.06)	(-0.03, 0.10)	(-0.05, 0.07)	(-0.06, 0.07)	(-0.03,0.09)

Table A5. Median (interquartile range, IQR) of sub-gingival microbial counts of eight periodontal pathogens obtained via DNA checkerboard immunoblotting among a subset of the Dental ARIC participants (n=1020)

	P *gingivalis*	T *forsythia*	T *denticola*	P *intermedia*	P *nigriscens*	C *rectus*	F *nucleatum*	A *actinom.*
Total sample(n)	1020	1020	1017	1018	1020	1018	1012	1020
	945 (3061)	1314 (4368)	2665 (9273)	4270 (15432)	3185 (12758)	2524 (19676)	5605 (32326)	2044 (5592)
Brushing frequency (yesterday)								
One time or none	996 (3798)	1506 (5244)	3129 (10393)	4948 (17741)	3075 (18176)	2736 (12907)	7447 (40283)	2584 (6912)
Two or more	904 (2838)	1277 (4109)	2330 (9112)	3977 (14675)	3216 (10557)	2324 (8879)	5048 (26681)	1882 (5004)
Flossed last week								
Did not floss	1386 (3905)	1734 (6439)	3294 (11529)	6185 (21216)	4208 (18039)	3030 (14578)	7252 (42203)	2779 (6603)
Flossed 1+ times	780 (2657)	1094 (3588)	2302 (8314)	3582 (13698)	2711 (10246)	2218 (8797)	4830 (26433)	1750 (5004)
Dental visit reason								
Regular visits	876 (2817)	1143 (3581)	2203 (8226)	3733 (13696)	2827 (10342)	2246 (8377)	4811 (24388)	1876 (4852)
Only when problem	1409 (5391)	2900 (10448)	4624 (19327)	8204 (32715)	6128 (24793)	4328 (26746)	10653 (73484)	3012 (8944)
Last dental visit								
<6 months ago	908 (2880)	1164 (3573)	2183 (7841)	3624 (13697)	2637 (10246)	2231 (8310)	5048 (25670)	1822 (4816)
6-24 months	666 (3086)	1295 (4939)	3355 (11840)	4713 (16502)	4662 (14860)	2618 (13005)	6080 (36723)	2204 (6916)
2+ years ago	1633 (5947)	3497 (12599)	4875 (21366)	10334 (37253)	6654 (26833)	4112 (28367)	10245 (110673)	4208 (10449)
Education level								
Basic (0-11 years)	1719 (5961)	2680 (8607)	3616 (16033)	9248 (34555)	4911 (21757)	5424 (27103)	15088 (77493)	3945 (10016)
Intermediate (12-15 yrs)	859 (2887)	1535 (4914)	3255 (9823)	3980 (14096)	3385 (12601)	2630 (10467)	6272 (38108)	2267 (5525)
Advanced (16+ yrs)	910 (2990)	959 (3383)	1973 (8226)	3948 (13950)	2711 (9889)	2048 (8032)	3966 (21940)	1536 (4710)

Table A6. Median (interquartile range, IQR) values of sub-gingival microbial counts of the "red" and "orange" complex pathogens, and total microbial load obtained via DNA checkerboard immunoblotting, in the total sample, stratified by sex, smoking and diabetic status and their association with age and BMI among the Caucasian participants of the Dental ARIC study, for whom clinical and genetic data were matched after exclusions (n=1020)

	"Red" complex	"Orange" complex	Total microbial load
Total sample	5676 (13629)	22668 (60169)	31640 (76343)
n	1017	1009	1006
Sex			
Females	5684 (13150)	22579 (60294)	31594 (73568)
Males	5676 (14061)	22731 (60252)	32084 (82237)
Smoking status			
Never	5604 (13100)	23437 (73000)	31186 (96592)
Ever- light	6848 (12977)	25544 (69412)	35654 (83731)
Ever- heavy	4591 (10733)	17750 (41055)	26584 (58326)
Current- light	5540 (16016)	23990 (46456)	42370 (48097)
Current– heavy	6500 (19215)	22071 (54349)	32580 (61111)
Diabetic status			
Healthy	5539 (13169)	21728 (56399)	30766 (73756)
Diabetes mellitus	7246 (20397)	26616 (84427)	34893 (107047)
Tooth loss[1]			
0-9 lost teeth	4867 (12089)	18727 (48253)	27334 (63495)
10-18 lost teeth	5770 (16539)	23813 (71346)	32668 (83626)
19-32 lost teeth	11044 (25744)	34946 (143873)	49798 (222836)
Age (rho, 95% CL)	0.05 (-0.01, 0.11)	0.02 (-0.04, 0.08)	0.03 (-0.04, 0.09)
BMI (rho, 95% CL)	0.01 (-0.05, 0.08)	0.01 (-0.05, 0.06)	0.01 (-0.06, 0.07)
Number of natural teeth (rho, 95% CL)	-0.14 (-0.20, -0.08)	-0.13(-0.20, -0.07)	-0.15 (-0.21, -0.09)

[1]Kruskal-Wallis test of stratum-specific bacterial count estimates $P<0.05$

Table A7. Median (interquartile range, IQR) values of sub-gingival microbial counts of the "red" and "orange" complex pathogens, and total microbial load obtained via DNA checkerboard immunoblotting, in the total sample, stratified by brushing and flossing frequency, dental attendance frequency and reason, and education level among the Caucasian participants of the Dental ARIC study, for whom clinical and genetic data were matched after exclusions (n=1020)

	"Red" complex	"Orange" complex	Total microbial load
Total sample	5676 (13629)	22668 (60169)	31640 (76343)
n	1017	1009	1006
Brushing frequency (yesterday)			
One time or none	6352 (17930)	25038 (80044)	33245 (101539)
Two or more	5518 (12998)	20464 (49853)	28904 (69526)
Flossed last week[1(total only)]			
Did not floss	7244 (19794)	25883 (89482)	33900 (127185)
Flossed one or more times	5444 (12521)	20006 (48350)	28182 (64442)
Dental visit reason[1]			
Regular visits	5397 (12535)	20283 (45756)	29129 (61763)
Only when problem	9289 (30032)	33660 (150830)	49084 (233372)
Last dental visit[2]			
<6 months ago	5338 (12387)	20822 (45828)	28888 (62424)
6-24 months	6251 (18396)	21441 (69189)	31233 (131809)
More than 2 years ago	9525 (39649)	40110 (176897)	54380 (367130)
Education level[2]			
Basic (0-11 years)	8709 (29505)	33695 (158940)	49798 (220499)
Intermediate (12-15 years)	6356 (13756)	23504 (65055)	32477 (82767)
Advanced (16+ years)	4935 (12199)	19120 (43498)	27929 (58540)

[1]Median test of stratum-specific bacterial count estimates $P<0.05$; [2]Kruskal-Wallis test of stratum-specific bacterial count estimates $P<0.05$

Table A8. Sample power calculation for SNP main effect among the ARIC cohort study whites. Assumptions are: MAF 0.1-0.5, estimated SNP effects of 1.1-1.6, and outcome prevalence 0.2 (severe periodontitis).

Outcome:	Disease
Design:	Unmatched case-control (1:4)
Hypothesis:	Gene only
Sample size:	800 cases, 4 control(s) per case are required
Significance:	0.050000, 2-sided
Gene	
Mode of inheritance:	Log-additive
Allele frequency:	0.1000 to 0.5000 by 0.1000
Disease model	Summary parameters
*P_0 0.196890	k_P 0.200000
R_G: 1.1000	(*indicates calculated value)

Parameter	Null	Full	Reduced
Gene	$\beta_G=0$	β_G	——

		Power	
Frequency	R_G	Gene	P_0
0.100000	1.1000	0.1797	0.196890
	1.2000	0.5206	0.193948
	1.3000	0.8332	0.191161
	1.4000	0.9664	0.188514
	1.5000	0.9959	0.185998
	1.6000	0.9997	0.183600
0.200000	1.1000	0.2810	0.193833
	1.2000	0.7587	0.188098
	1.3000	0.9713	0.182747
	1.4000	0.9987	0.177741
	1.5000	0.9999	0.173045
	1.6000	0.9999	0.168630
0.300000	1.1000	0.3499	0.190830
	1.2000	0.8570	0.182444
	1.3000	0.9921	0.174743
	1.4000	0.9999	0.167645
	1.5000	0.9999	0.161082
	1.6000	0.9999	0.154995
0.400000	1.1000	0.3883	0.187879
	1.2000	0.8948	0.176981
	1.3000	0.9962	0.167134
	1.4000	0.9999	0.158195
	1.5000	0.9999	0.150049
	1.6000	0.9999	0.142599
0.500000	1.1000	0.3991	0.184980
	1.2000	0.9023	0.171704
	1.3000	0.9967	0.159902
	1.4000	0.9999	0.149355
	1.5000	0.9999	0.139884
	1.6000	0.9999	0.131343

Table A9 Sample power calculation for SNP main effect among the ARIC cohort study whites. Assumptions are: MAF 0.1-0.5, estimated SNP effects of 1.2-1.6, and outcome prevalence 0.2 (highest quintile of bacterial colonization phenotypes).

Outcome:	Disease
Design:	Unmatched case-control (1:4)
Hypothesis:	Gene only
Sample size:	200 cases, 4 control(s) per case are required
Significance:	0.050000, 2-sided
Gene	
Mode of inheritance:	Log-additive
Allele frequency:	0.1000 to 0.5000 by 0.1000
Disease model	Summary parameters
$*P_0$ 0.193948	k_P 0.200000
R_G: 1.2000	(*indicates calculated value)

Parameter	Null	Full	Reduced
Gene	$\beta_G=0$	β_G	——

		Power	
Frequency	R_G	Gene	P_0
0.100000	1.2000	0.1715	0.193948
	1.3000	0.3101	0.191161
	1.4000	0.4741	0.188514
	1.5000	0.6340	0.185998
	1.6000	0.7669	0.183600
0.200000	1.2000	0.2652	0.188098
	1.3000	0.4880	0.182747
	1.4000	0.7022	0.177741
	1.5000	0.8547	0.173045
	1.6000	0.9397	0.168630
0.300000	1.2000	0.3279	0.182444
	1.3000	0.5899	0.174743
	1.4000	0.8034	0.167645
	1.5000	0.9248	0.161082
	1.6000	0.9765	0.154995
0.400000	1.2000	0.3619	0.176981
	1.3000	0.6382	0.167134
	1.4000	0.8431	0.158195
	1.5000	0.9468	0.150049
	1.6000	0.9853	0.142599
0.500000	1.2000	0.3699	0.171704
	1.3000	0.6468	0.159902
	1.4000	0.8482	0.149355
	1.5000	0.9485	0.139884
	1.6000	0.9856	0.131343

Table A10. Sample power calculation for SNP*environment interaction among Caucasians. Assumptions are: MAF 0.1-0.5, estimated SNP effect of 1.4, environmental factor prevalence 0.1 (current heavy smoking), interaction effect 1.2-2.0 and outcome prevalence 0.2 (severe periodontitis).

Outcome:	Disease
Design:	Unmatched case-control (1:4)
Hypothesis:	Gene-environment interaction
Sample size:	800 cases, 4 control(s) per case are required
Significance:	0.050000, 2-sided
Gene	
Mode of inheritance:	Log-additive
Allele frequency:	0.1000 to 0.5000 by 0.1000
Binary environmental factor	
Prevalence:	0.1000
Disease model	Summary parameters

$*P_0$	0.187826	k_P	0.200000
R_G:	1.4000	$*\bar{R}_G$:	1.4269
R_E:	1.0000	$*\bar{R}_E$:	1.0453
R_{GE}:	1.2000	(*indicates calculated value)	

Parameter	Null	Full	Reduced
Interaction	$\beta_{GE}=0$	$\beta_{GE},\beta_G,\beta_E$	β_G,β_E
Gene	$\beta_G=0$	β_G	——
Environment	$\beta_E=0$	β_E	——

		Power					
Frequency	R_{GE}	Interaction	Gene	Environment	P_0	$\bar{R}_G$	$\bar{R}_E$
0.100000	1.2000	0.0989	0.9800	0.0633	0.187826	1.4269	1.0453
	1.4000	0.2277	0.9880	0.0988	0.187202	1.4516	1.0874
	1.6000	0.4020	0.9926	0.1513	0.186634	1.4742	1.1265
	1.8000	0.5782	0.9954	0.2159	0.186117	1.4949	1.1627
	2.0000	0.7247	0.9971	0.2877	0.185645	1.5138	1.1964
0.200000	1.2000	0.1343	0.9995	0.1019	0.176444	1.4271	1.0902
	1.4000	0.3520	0.9998	0.2431	0.175265	1.4521	1.1760
	1.6000	0.6059	0.9999	0.4358	0.174195	1.4751	1.2571
	1.8000	0.8002	0.9999	0.6261	0.173222	1.4959	1.3337
	2.0000	0.9119	0.9999	0.7759	0.172336	1.5149	1.4057
0.300000	1.2000	0.1572	0.9999	0.1647	0.165816	1.4273	1.1347
	1.4000	0.4273	0.9999	0.4585	0.164145	1.4528	1.2660
	1.6000	0.7073	0.9999	0.7523	0.162628	1.4762	1.3929
	1.8000	0.8826	0.9999	0.9173	0.161251	1.4975	1.5147
	2.0000	0.9605	0.9999	0.9783	0.160002	1.5167	1.6310
0.400000	1.2000	0.1686	0.9999	0.2491	0.155900	1.4275	1.1788
	1.4000	0.4640	0.9999	0.6810	0.153795	1.4536	1.3579
	1.6000	0.7517	0.9999	0.9331	0.151881	1.4777	1.5348
	1.8000	0.9129	0.9999	0.9922	0.150145	1.4996	1.7077
	2.0000	0.9749	0.9999	0.9994	0.148573	1.5193	1.8756
0.500000	1.2000	0.1696	0.9999	0.3500	0.146657	1.4278	1.2228
	1.4000	0.4691	0.9999	0.8483	0.144170	1.4545	1.4519
	1.6000	0.7592	0.9999	0.9891	0.141902	1.4794	1.6836
	1.8000	0.9184	0.9999	0.9997	0.139845	1.5021	1.9148
	2.0000	0.9776	0.9999	0.9999	0.137984	1.5226	2.1434

Table A11 Sample power calculation for SNP*environment interaction among Caucasians. Assumptions are: MAF 0.1-0.5, estimated SNP effect of 1.4, environmental factor prevalence 0.5 (ever smoker), interaction effect 1.2-2.0 and outcome prevalence 0.2 (severe periodontitis).

Outcome:	Disease
Design:	Unmatched case-control (1:4)
Hypothesis:	Gene-environment interaction
Sample size:	800 cases, 4 control(s) per case are required
Significance:	0.050000, 2-sided
Gene	
Mode of inheritance:	Log-additive
Allele frequency:	0.1000 to 0.5000 by 0.1000
Binary environmental factor	
Prevalence:	0.5000
Disease model	Summary parameters
$*P_0$ 0.185102	k_P 0.200000
R_G: 1.4000	$*\overline{R}_G$: 1.5369
R_E: 1.0000	$*\overline{R}_E$: 1.0455
R_{GE}: 1.2000	(*indicates calculated value)

Parameter	Null	Full	Reduced
Interaction	$\beta_{GE}=0$	$\beta_{GE},\beta_G,\beta_E$	β_G,β_E
Gene	$\beta_G=0$	β_G	——
Environment	$\beta_E=0$	β_E	——

		Power					
Frequency	R_{GE}	Interaction	Gene	Environment	P_0	$\overline{R}_G$	$\overline{R}_E$
0.100000	1.2000	0.1859	0.9983	0.0869	0.185102	1.5369	1.0455
	1.4000	0.5065	0.9999	0.1869	0.182049	1.6674	1.0879
	1.6000	0.7921	0.9999	0.3300	0.179301	1.7915	1.1277
	1.8000	0.9355	0.9999	0.4881	0.176815	1.9095	1.1650
	2.0000	0.9841	0.9999	0.6346	0.174552	2.0216	1.2000
0.200000	1.2000	0.2801	0.9999	0.1951	0.171372	1.5374	1.0906
	1.4000	0.7203	0.9999	0.5431	0.165745	1.6692	1.1778
	1.6000	0.9438	0.9999	0.8350	0.160741	1.7948	1.2616
	1.8000	0.9931	0.9999	0.9601	0.156263	1.9141	1.3421
	2.0000	0.9994	0.9999	0.9929	0.152230	2.0274	1.4195
0.300000	1.2000	0.3355	0.9999	0.3617	0.158740	1.5381	1.1354
	1.4000	0.8048	0.9999	0.8540	0.150970	1.6715	1.2696
	1.6000	0.9743	0.9999	0.9892	0.144143	1.7994	1.4020
	1.8000	0.9980	0.9999	0.9997	0.138102	1.9215	1.5322
	2.0000	0.9999	0.9999	0.9999	0.132718	2.0376	1.6600
0.400000	1.2000	0.3589	0.9999	0.5522	0.147135	1.5389	1.1799
	1.4000	0.8313	0.9999	0.9744	0.137606	1.6743	1.3634
	1.6000	0.9806	0.9999	0.9998	0.129332	1.8054	1.5491
	1.8000	0.9987	0.9999	0.9999	0.122090	1.9313	1.7360
	2.0000	0.9999	0.9999	0.9999	0.115704	2.0520	1.9235
0.500000	1.2000	0.3559	0.9999	0.7242	0.136486	1.5397	1.2242
	1.4000	0.8235	0.9999	0.9974	0.125537	1.6775	1.4592
	1.6000	0.9774	0.9999	0.9999	0.116141	1.8123	1.7030
	1.8000	0.9981	0.9999	0.9999	0.108007	1.9433	1.9542
	2.0000	0.9999	0.9999	0.9999	0.100909	2.0701	2.2115

Table A12 Sample power calculation for SNP main effect among African Americans. Assumptions are: MAF 0.1-0.5, estimated SNP effects of 1.3-1.7 and outcome prevalence 0.2 (severe periodontitis).

Outcome:	Disease
Design:	Unmatched case-control (1:4)
Hypothesis:	Gene only
Sample size:	200 cases, 4 control(s) per case are required
Significance:	0.050000, 2-sided
Gene	
Mode of inheritance:	Log-additive
Allele frequency:	0.1000 to 0.5000 by 0.1000
Disease model	Summary parameters
*P_0 0.191161	k_P 0.200000
R_G: 1.3000	(*indicates calculated value)

Parameter	Null	Full	Reduced
Gene	$\beta_G=0$	β_G	——

		Power	
Frequency	R_G	Gene	P_0
0.100000	1.3000	0.3101	0.191161
	1.4000	0.4741	0.188514
	1.5000	0.6340	0.185998
	1.6000	0.7669	0.183600
	1.7000	0.8632	0.181314
0.200000	1.3000	0.4880	0.182747
	1.4000	0.7022	0.177741
	1.5000	0.8547	0.173045
	1.6000	0.9397	0.168630
	1.7000	0.9783	0.164471
0.300000	1.3000	0.5899	0.174743
	1.4000	0.8034	0.167645
	1.5000	0.9248	0.161082
	1.6000	0.9765	0.154995
	1.7000	0.9938	0.149335
0.400000	1.3000	0.6382	0.167134
	1.4000	0.8431	0.158195
	1.5000	0.9468	0.150049
	1.6000	0.9853	0.142599
	1.7000	0.9966	0.135762
0.500000	1.3000	0.6468	0.159902
	1.4000	0.8482	0.149355
	1.5000	0.9485	0.139884
	1.6000	0.9856	0.131343
	1.7000	0.9966	0.123611

Table A13. Results of SNP interactions (interaction term $P<10^{-5}$) with sex, diabetes (binary) and smoking status (never, ever, current) in the GWA analysis of chronic periodontitis ("extent" of disease trait: proportion of sites exhibiting attachment loss ≥3 mm) among the white Dental ARIC participants (n=4610).

SNP interaction with Sex					SNP interaction with Diabetes Mellitus					SNP interaction with Smoking				
Chr	SNP	Position Build36	Nearby gene (location)	P value	Chr	SNP	Position Build36	Nearby gene (location)	P value	Chr.	SNP	Position Build36	Nearby gene (location)	P value
7	rs616939	152391164	ACTR3B (208Kb)	$1.4x10^{-6}$	11	rs7107482	rs7107482	NAV2 (intronic)	$1.1x10^{-7}$	19	rs8111486	16732705	NWD1 (intronic)	$2.0x10^{-7}$
1	rs1551110	105334285	LOC100129138 (913Kb)	$1.5x10^{-6}$	3	rs1405597	rs1405597	C3orf58 (603Kb)	$2.4x10^{-7}$	2	rs1533528	199322315	SATB2 (520Kb)	$4.2x10^{-7}$
4	rs10019569	148166054	TTC19 (78Kb)	$2.0x10^{-6}$	10	rs594612	rs594612	NRG3 (intronic)	$3.3x10^{-7}$	12	rs17676308	51310951	KRT73 (12Kb); KRT2 (14Kb)	$1.4x10^{-6}$
8	rs12155819	9610089	TNKS (intronic)	$2.3x10^{-6}$	5	rs2160058	rs2160058	ACTBL2 (287Kb)	$3.8x10^{-7}$	7	rs850380	144771261	TPK1 (607Kb); CNTNAP2 (673Kb)	$2.2x10^{-6}$
4	rs1372486	46864735	GABRB1 (intronic)	$2.5x10^{-6}$	5	rs9327150	rs9327150	PRR16 (intronic)	$1.4x10^{-6}$	2	rs1505839	139512322	NXPH2 (258Kb)	$2.2x10^{-6}$
12	rs4761973	50166536	SLC4A8	$2.6x10^{-6}$	6	rs9461680	rs9461680	HLA-C (3Kb); HCG27 (69Kb)	$1.4x10^{-6}$	12	rs12312201	15003845	ARHGDIB (intronic)	$2.9x10^{-6}$
1	rs4659467	234987803	ACTN2 (intronic)	$3.0x10^{-6}$	16	rs9302669	rs9302669	MMP2 (56Kb); LPCAT2 (86Kb)	$3.4x10^{-6}$	2	rs17288217	175581131	CHN1 (3Kb); ATF2 (66Kb)	$4.5x10^{-6}$
13	rs10851261	110767630	C13orf16 (3Kb); ARHGEF7 (12Kb)	$3.3x10^{-6}$	3	rs795346	rs795346	TGFBR2 (5Kb); GADL1 (27Kb)	$3.4x10^{-6}$	18	rs10871635	64716612	CCDC102B (intronic)	$4.7x10^{-6}$
8	rs2063794	78581773	PEX2 (506Kb)	$3.6x10^{-6}$	2	rs7603311	rs7603311	STON1-GTF2A1L (38Kb); LHCGR (58Kb)	$3.6x10^{-6}$	2	rs16861555	14354512	FAM84A (336Kb)	$5.8x10^{-6}$
12	rs10505960	25165034	CASC1 (intronic)	$4.8x10^{-6}$	15	rs11073534	rs11073534	LOC91948 (470Kb)	$4.7x10^{-6}$	3	rs9817711	102371842	IMPG2 (56Kb)	$6.5x10^{-6}$
6	rs2326476	4234898	ECI2 (154Kb)	$4.8x10^{-6}$	18	rs948649	rs948649	BOD1P (56Kb)	$5.0x10^{-6}$	15	rs2413906	46557872	FBN1 (intronic)	$7.5x10^{-6}$
16	rs11643337	78900950	DYNLRB2 (231Kb)	$5.8x10^{-6}$	3	rs1370045	rs1370045	VGLL3 (25Kb)	$5.1x10^{-6}$	1	rs9326028	54523213	SSBP3 (intronic)	$8.3x10^{-6}$
6	rs1204245	13983271	RNF182 (49Kb)	$6.1x10^{-6}$	9	rs413553	rs413553	PTPRD (intronic)	$5.4x10^{-6}$					
21	rs2834319	34278895	ATP5O (69Kb); SLC5A3 (89Kb)	$6.4x10^{-6}$	17	rs12450764	rs12450764	LOC100499467 (intronic)	$5.6x10^{-6}$					
5	rs7380768	62952975	HTR1A (339Kb)	$6.6x10^{-6}$	6	rs2814124	rs2814124	MCART3P (1Kb); EYS (83Kb)	$6.2x10^{-6}$					
14	rs10140314	21139677	OR10G3 (31Kb)	$7.7x10^{-6}$	9	rs2309932	rs2309932	TMEM2 (195Kb)	$6.5x10^{-6}$					
4	rs12641678	174096485	GALNTL6 (intronic)	$8.4x10^{-6}$	8	rs17348575	rs17348575	CSMD1 (intronic)	$7.2x10^{-6}$					
5	rs323565	40884560	CARD6 (intronic)	$9.0x10^{-6}$										
9	rs7864512	37939638	SHB (intronic)	$9.4x10^{-6}$										
3	rs6767604	176063265	NAALADL2 (intronic)	$9.7x10^{-6}$										

Table A14. Genome-wide association analysis results of the CDC/AAP chronic periodontitis (CP) classification traits (severe CP vs. healthy and moderate CP vs. healthy) and extent of attachment loss trait (EAL-proportion of sites exhibiting attachment loss ≥3mm), among the white participants of the Dental ARIC study (n=4610). Supplemental list of single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF-HapMap II CEU) of ≥5% and that emerged below the $P<5x10^{-6}$ in the "sensitivity" analysis that included logistic regression models adjusted for diabetes (binary variable) and smoking (never, former, current).

Chr.	SNP	Position Build36	ca[1]	nca[2]	Risk allele frequency (HapMap II- CEU)	Closest gene(s) and position or distance	Risk allele frequency "cases"/ healthy	P value (adjusted)[3]	Odds ratio/beta (95% CI[4]; adjusted)[3]
Severe CP									
1p21.1	rs10493998	103500842	C	T	0.120	*COL11A1* (154Kb); *RNPC3* (340Kb)	0.148/0.109	$2.2x10^{-6}$	1.68 (1.36, 2.08)
4q21.3	rs17006135	83836002	C	T	0.071	*SCD5* (36Kb); *SEC31A* (123Kb)	0.096/0.059	$4.4x10^{-6}$	1.82 (1.41, 2.35)
Moderate CP									
12q15	rs11615037	67741708	C	G	0.276	*CPM* (98Kb); *CPSF6* (178Kb)	0.281/0.244	$4.4x10^{-6}$	1.33 (1.18, 1.51)
EAL									
4q21.3	rs17006135	83836002	C	T	0.071	*SCD5* (36Kb); *SEC31A* (123Kb)		$8.2x10^{-7}$	3.15 (1.90, 4.41)
14q21.3	rs8006336	48388426	A	G	0.651	*RPS29* (726Kb); *PPIL5* (747Kb)		$1.4x10^{-6}$	1.68 (1.00, 2.36)

1: coded allele
2: non-coded allele
3: based on logistic regression models that were adjusted for diabetes (binary) and smoking (never, former, current)
4: Confidence interval

Table A15. Coded and non-coded alleles of prioritized SNPs for severe and moderate CP on the Affymetrix 6.0 (ARIC GWAS) and the Illumina Human 1M Duo (ver.3) (Health ABC GWAS) platforms. Pending verification by the Health ABC collaborating group.

SNP	Affymetrix 6.0		Illumina Human1M Duo Beadchip		Different strand (yes/no)
	Coded allele	Non-coded allele	Coded allele	Non-coded allele	
rs12883458	C	T	C	T	No
rs2521634	A	G	A	G	No
rs11925054	G	T	G	T	No
rs10493998	C	T	C	T	No
rs17006135	C	T	C	T	No
rs7762544	A	G	A	G	No
rs3826782	A	G	A	G	No
rs12260727	A	G	A	G	No
rs11615037	C	G	C	G	No

Table A16. Genome-wide association replication results [SNPs prioritized based on a $P<5x10^{-6}$ threshold in the dental ARIC (n=4610)] of severe and moderate CP in the Health ABC cohort. Reported estimates correspond to the ARIC cohort 'risk' allele effect on severe and moderate CP and EAL in the Health ABC cohort.

	Severe CP		Moderate CP		EAL	
dbSNP	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	beta (95% CI)
Severe CP						
rs12883458[1]	**0.25**	**0.65 (0.31, 1.34)**	0.71	0.90 (0.51, 1.58)	0.56	-0.05 (-0.21, 0.11)
rs2521634	**0.05**	**1.65 (1.01, 2.69)**	0.03	1.63 (1.06, 2.51)	0.53	0.04 (-0.08, 0.16)
rs11925054[1]	**0.06**	**0.51 (0.25, 1.04)**	0.05	0.53 (0.29, 0.98)	0.27	-0.08 (-0.22, 0.06)
rs10493998[2]	**0.89**	**1.06 (0.52, 2.15)**	0.34	1.31 (0.76, 1.31)	0.24	0.09 (-0.05, 0.23)
rs17006135[1,2,3]	**0.03**	**0.43 (0.20, 0.90)**	0.32	0.73 (0.40, 1.35)	**0.39**	**-0.08 (-0.26, 0.26)**
Moderate CP						
rs7762544	0.34	1.30 (0.76, 2.20)	**0.28**	**1.32 (0.81, 2.16)**	0.77	0.02 (-0.10, 0.14)
rs3826782	0.01	4.18 (1.66, 10.50)	**0.06**	**2.08 (0.99, 4.37)**	0.10	0.14 (-0.04, 0.32)
rs12260727[1]	0.29	1.39 (0.74, 2.60)	**0.80**	**0.93 (0.55, 1.58)**	0.11	-0.11 (-0.25, 0.03)
rs11615037[1,2]	0.37	0.78 (0.46, 1.32)	**0.77**	**0.93 (0.61, 1.44)**	0.81	0.02 (-0.10, 0.14)

1 opposite direction effect; verification of coded SNP in Health ABC GWAS is pending
2 SNPs that were prioritized based on the $P<5x10^{-6}$ criterion in exploratory smoking- and diabetes-adjusted analysis
3 rs17006135 marked a shared risk locus for severe CP and EAL in the Dental ARIC GWAS adjusted exploratory analysis

Table A17. Replication of GWAS results (P-values and effect estimates) of prioritized SNPs of the ARIC genome-wide association analysis (crude and exploratory smoking and diabetes-adjusted analyses) in the Health ABC cohort. Reported effect estimates correspond to the 'risk' allele in the ARIC cohort.

	ARIC GWAS		ARIC GWAS adjusted		Health ABC	
dbSNP	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
Severe CP						
rs12883458[1]	$3.5x10^{-7}$	1.89 (1.48, 2.41)	$1.1x10^{-6}$	1.88 (1.46, 2.43)	0.25	0.65 (0.31, 1.34)
rs2521634	$1.6x10^{-6}$	1.47 (1.25, 1.73)	$2.6x10^{-5}$	1.43 (1.21, 1.69)	0.05	1.65 (1.01, 2.69)
rs11925054[1]	$6.5x10^{-7}$	1.69 (1.37, 2.10)	$2.6x10^{-7}$	1.78 (1.42, 2.23)	0.06	0.51 (0.25, 1.04)
rs10493998[2]	$9.2x10^{-6}$	1.59 (1.30, 1.95)	$2.2x10^{-6}$	1.68 (1.36, 2.08)	0.89	1.06 (0.52, 2.15)
rs17006135[1,2,3]	$1.8x10^{-5}$	1.70 (1.34, 2.16)	$4.4x10^{-6}$	1.82 (1.41, 2.35)	0.03	0.43 (0.20, 0.90)
Moderate CP						
rs7762544	$1.1x10^{-7}$	1.41 (1.24, 1.60)	$3.8x10^{-8}$	1.44 (1.26, 1.63)	0.28	1.32 (0.81, 2.16)
rs3826782	$4.0x10^{-6}$	2.00 (1.48, 2.70)	$1.4x10^{-6}$	2.09 (1.54, 2.84)	0.06	2.08 (0.99, 4.37)
rs12260727[1]	$6.0x10^{-7}$	1.54 (1.30, 1.82)	$3.0x10^{-6}$	1.51 (1.27, 1.79)	0.80	0.93 (0.55, 1.58)
rs11615037[1,2]	$2.6x10^{-5}$	1.29 (1.15, 1.46)	$4.4x10^{-6}$	1.33 (1.18, 1.51)	0.77	0.93 (0.61, 1.44)

1 opposite direction effect; verification of coded SNP in Health ABC GWAS is pending
2 SNPs that were prioritized based on the $P<5x10^{-6}$ criterion in exploratory smoking- and diabetes-adjusted analysis
3 rs17006135 marked a shared risk locus for severe CP and EAL in the Dental ARIC GWAS adjusted exploratory analysis

Table A18. Meta-analysis of GWAS results (meta-analysis P-values and effect estimates based on inverse variance weighting) of prioritized SNPs for severe and moderate CP in the ARIC and the Health ABC studies. Reported effect estimates correspond to the 'risk' allele in the ARIC cohort. Pooled estimates based on inverse-variance meta-analysis are presented for results that did not demonstrate substantial heterogeneity (homogeneity X^2 P<0.2 criterion).

	ARIC GWAS		Health ABC		Meta-analysis					
					Pooled estimates		Heterogeneity assessment[2]			
dbSNP	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	Diff. direction	I^2	X^2	P<0.2
Severe CP										
rs12883458	$3.5x10^{-7}$	1.89 (1.48, 2.41)	0.25	0.65 (0.31, 1.34)			*	86.6	7.5	*
rs2521634	$1.6x10^{-6}$	1.47 (1.25, 1.73)	0.05	1.65 (1.01, 2.69)	$3.5x10^{-7}$	1.49 (1.28, 1.73)		0	0.2	
rs11925054	$6.5x10^{-7}$	1.69 (1.37, 2.10)	0.06	0.51 (0.25, 1.04)			*	90.1	10.1	*
rs10493998[1]	$9.2x10^{-6}$	1.59 (1.30, 1.95)	0.89	1.06 (0.52, 2.15)				48.7	2.0	*
rs17006135[1]	$1.8x10^{-5}$	1.70 (1.34, 2.16)	0.03	0.43 (0.20, 0.90)			*	91.6	11.9	*
Moderate CP										
rs7762544	$1.1x10^{-7}$	1.41 (1.24, 1.60)	0.28	1.32 (0.81, 2.16)	$7.5x10^{-8}$	1.40 (1.24, 1.59)		0	0.6	
rs3826782	$4.0x10^{-6}$	2.00 (1.48, 2.70)	0.06	2.08 (0.99, 4.37)	$8.2x10^{-7}$	2.01 (1.52, 2.65)		0	0.9	
rs12260727	$6.0x10^{-7}$	1.54 (1.30, 1.82)	0.80	0.43 (0.20, 0.90)			*	67.7	3.1	*
rs11615037[1]	$2.6x10^{-5}$	1.29 (1.15, 1.46)	0.77	0.93 (0.61, 1.44)			*	51.4	2.6	*

1 SNPs that were prioritized based on the $P<5x10^{-6}$ criterion in exploratory smoking- and diabetes-adjusted analysis in the ARIC cohort

2 Asterisks (*) correspond to effects of different direction in the ARIC and Health ABC cohorts, and homogeneity X^2 P-values of less than 0.2

FIGURES

Figure A1. Power calculation for severe CP among whites for MAF: 0.1-0.5.

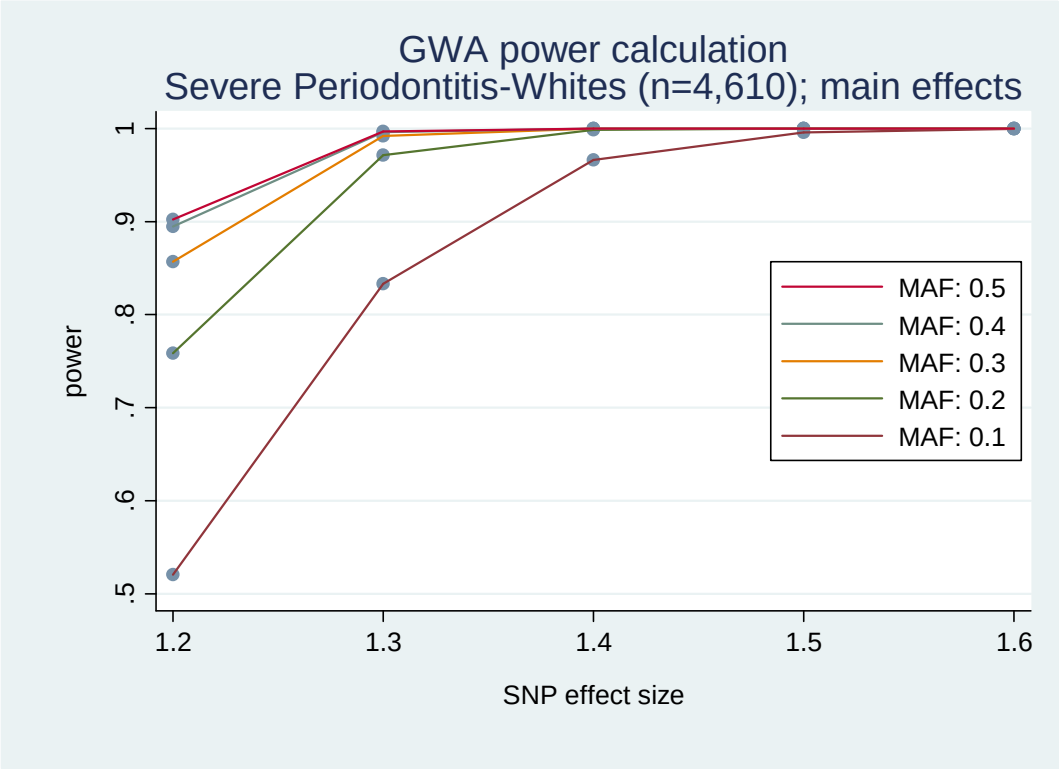


Figure A2. Power calculation for the "high" bacterial colonization trait among whites for MAF: 0.1-0.5.

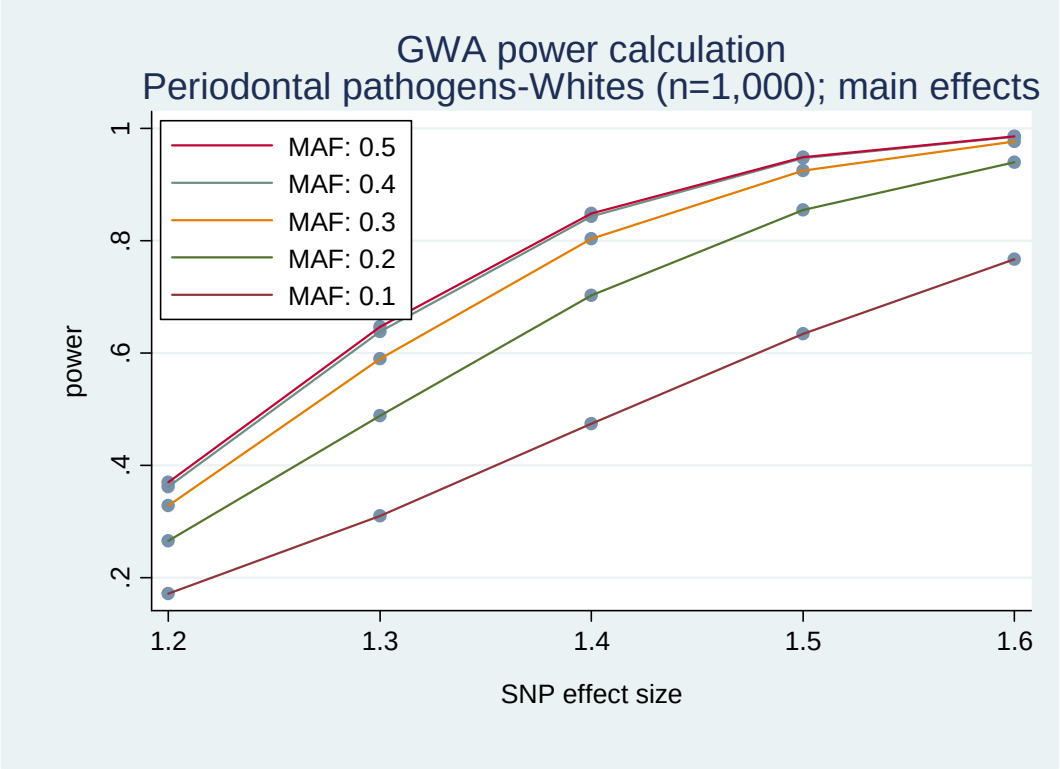


Figure A3. Power calculation for GxE interaction effects for severe CP, assuming 'exposure' prevalence 0.5 and MAF: 0.1-0.5 among the dental ARIC whites.

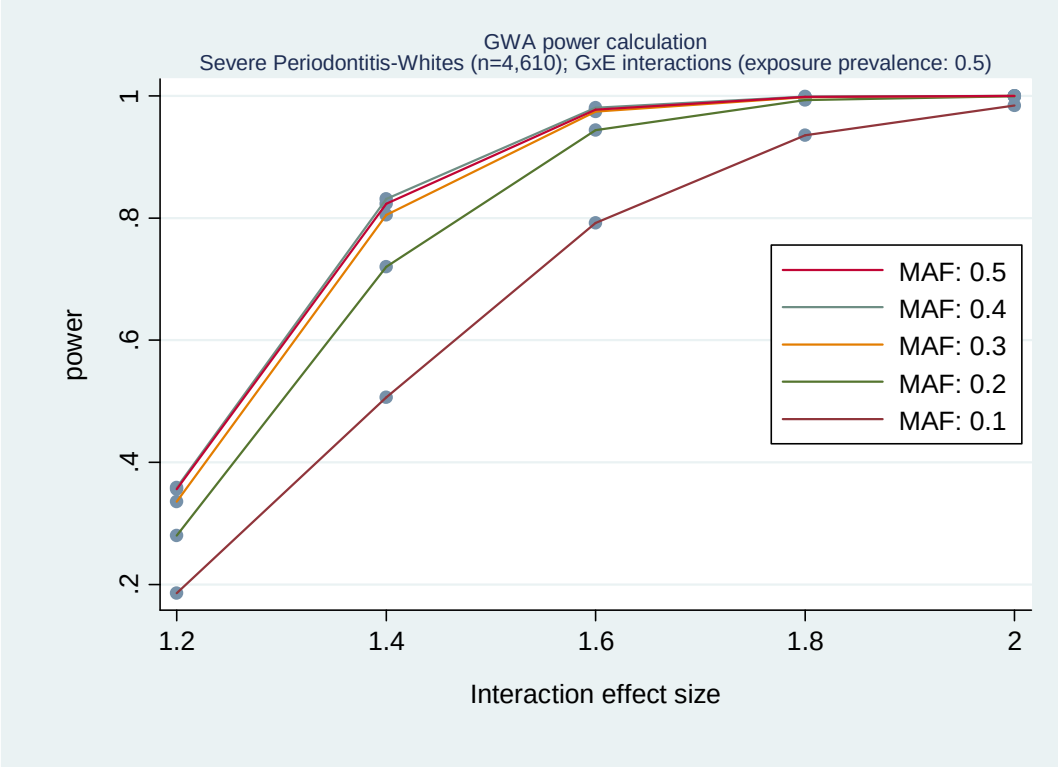


Figure A4. Power calculation for GxE interaction effects for severe CP, assuming 'exposure' prevalence 0.1 and MAF: 0.1-0.5 among the dental ARIC whites.

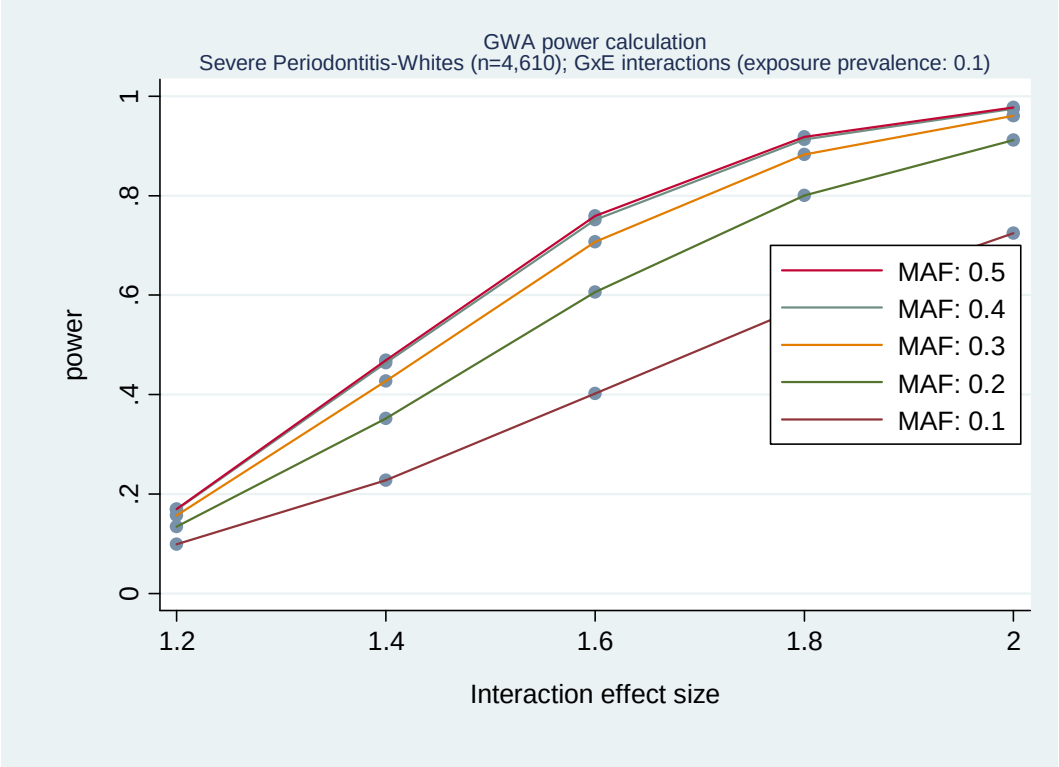


Figure A5. Density plot of average DST among the ARIC cohort study whites zoomed in the regions of DST cut-offs for related individuals (first degree DST ≥0.8, second degree 0.745 ≤ DST < 0.8). Red dots represent those 380 pairs of first degree and 207 pairs of second degree relatives.

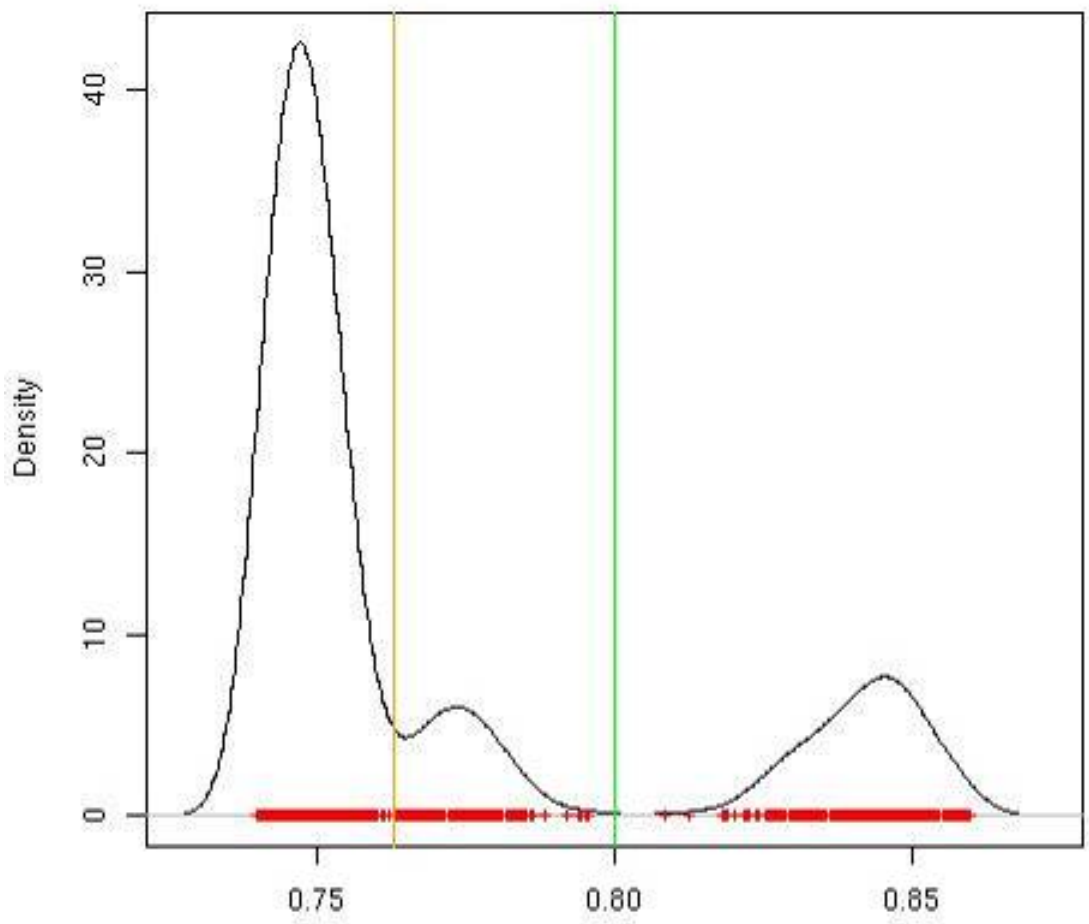


Figure A6. Density plot of average DST among the ARIC cohort study whites. Red dots represent 16 genetic outliers based on average DST values.

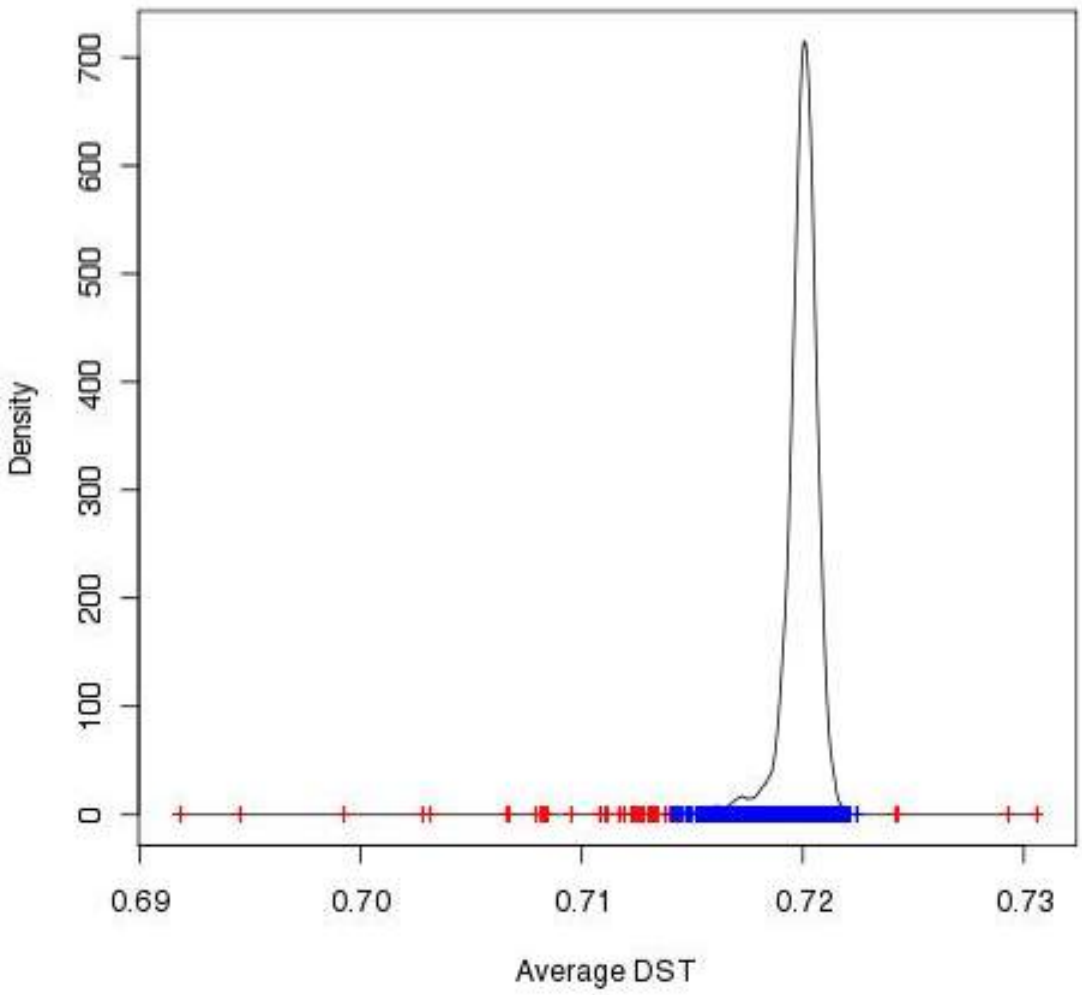


Figure A.7 Scatter plot matrix of the ten principal components generated with the EIGENSTRAT program for the white ARIC cohort study genotyped participants.

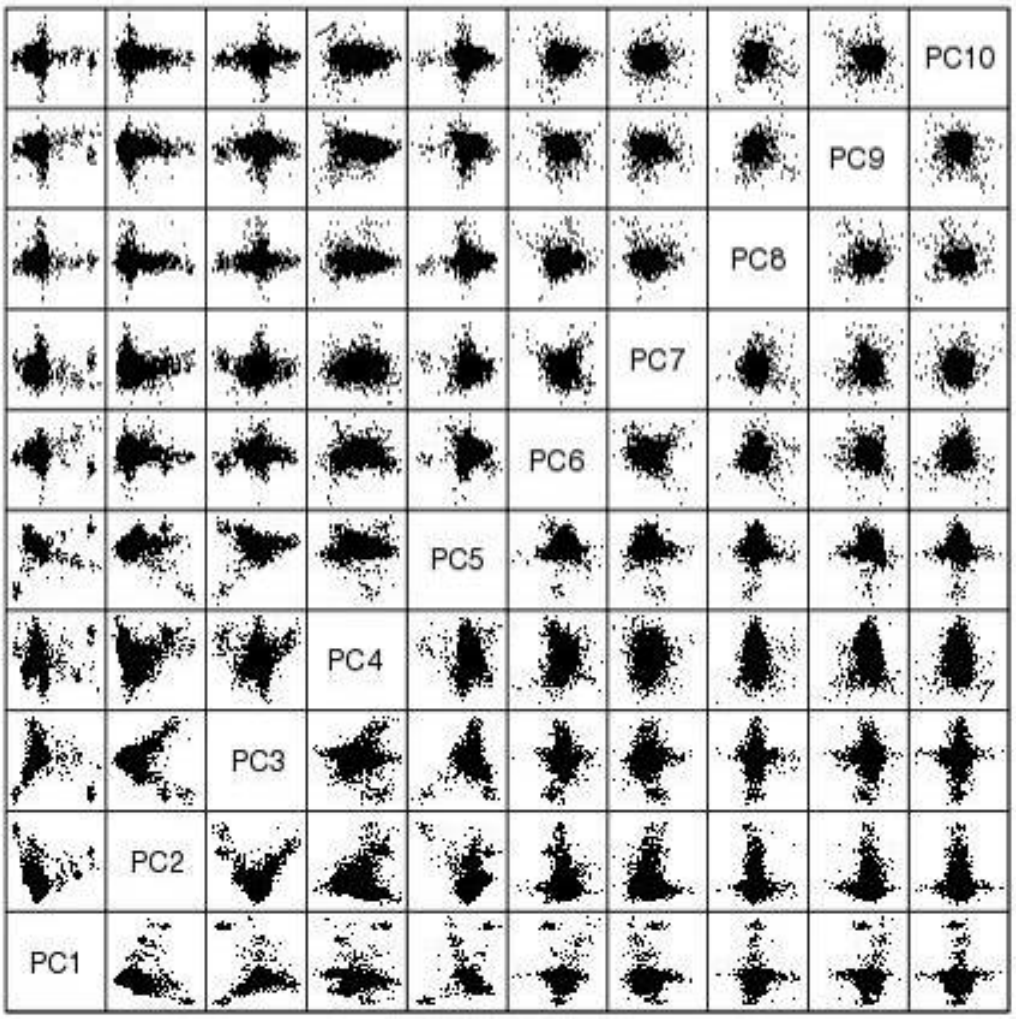


Figure A.8 Catalogue of published GWA studies (n=951) through June 2011.

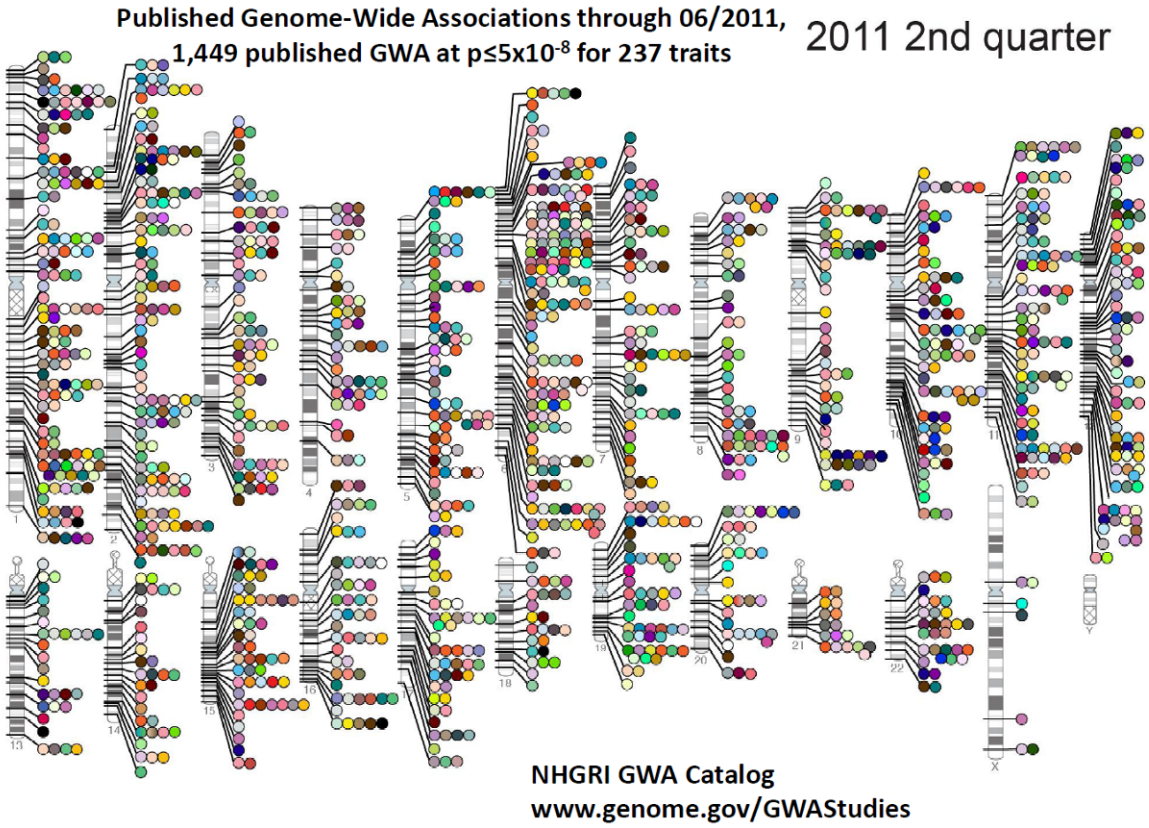


Figure A.9 Visualization of the 14q21 locus that was marked by rs12883458 ($P=3.5x10^{-7}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

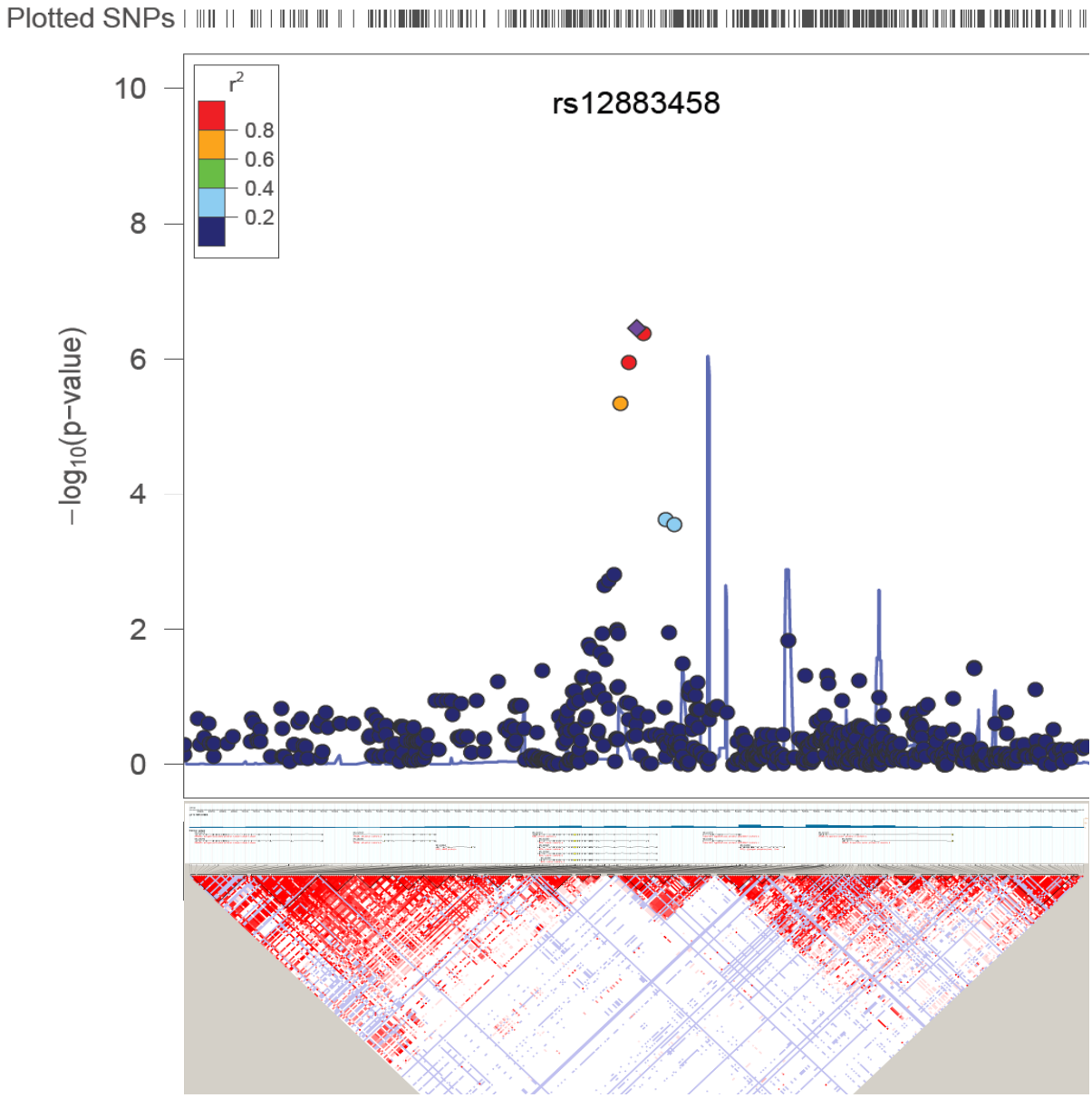


Figure A.10 Visualization of the 7p15 locus that was marked by rs2521634 ($P=1.6x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

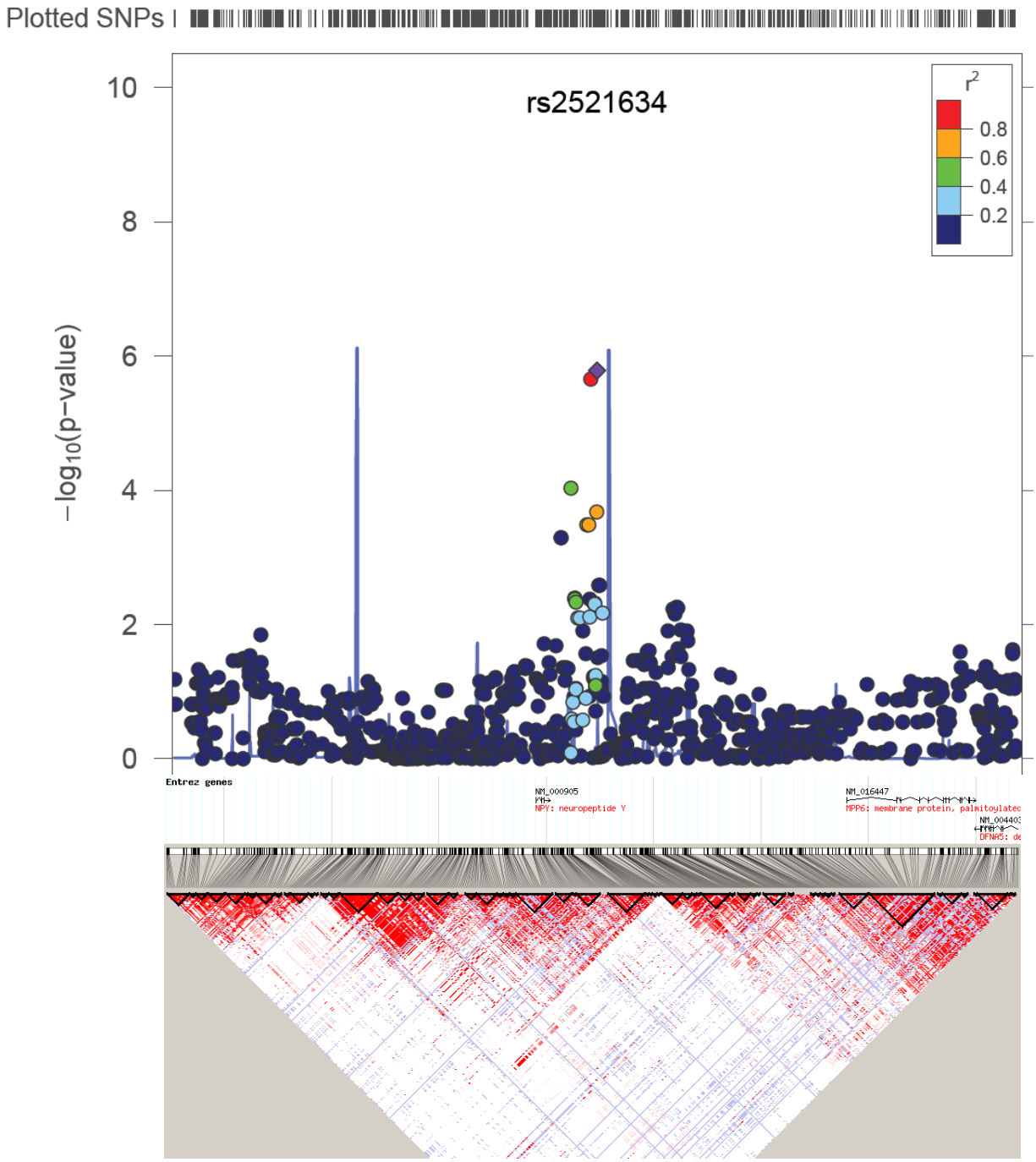


Figure A.11 Visualization of the 3p21 locus that was marked by rs11925054 ($P=6.5x10^{-7}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

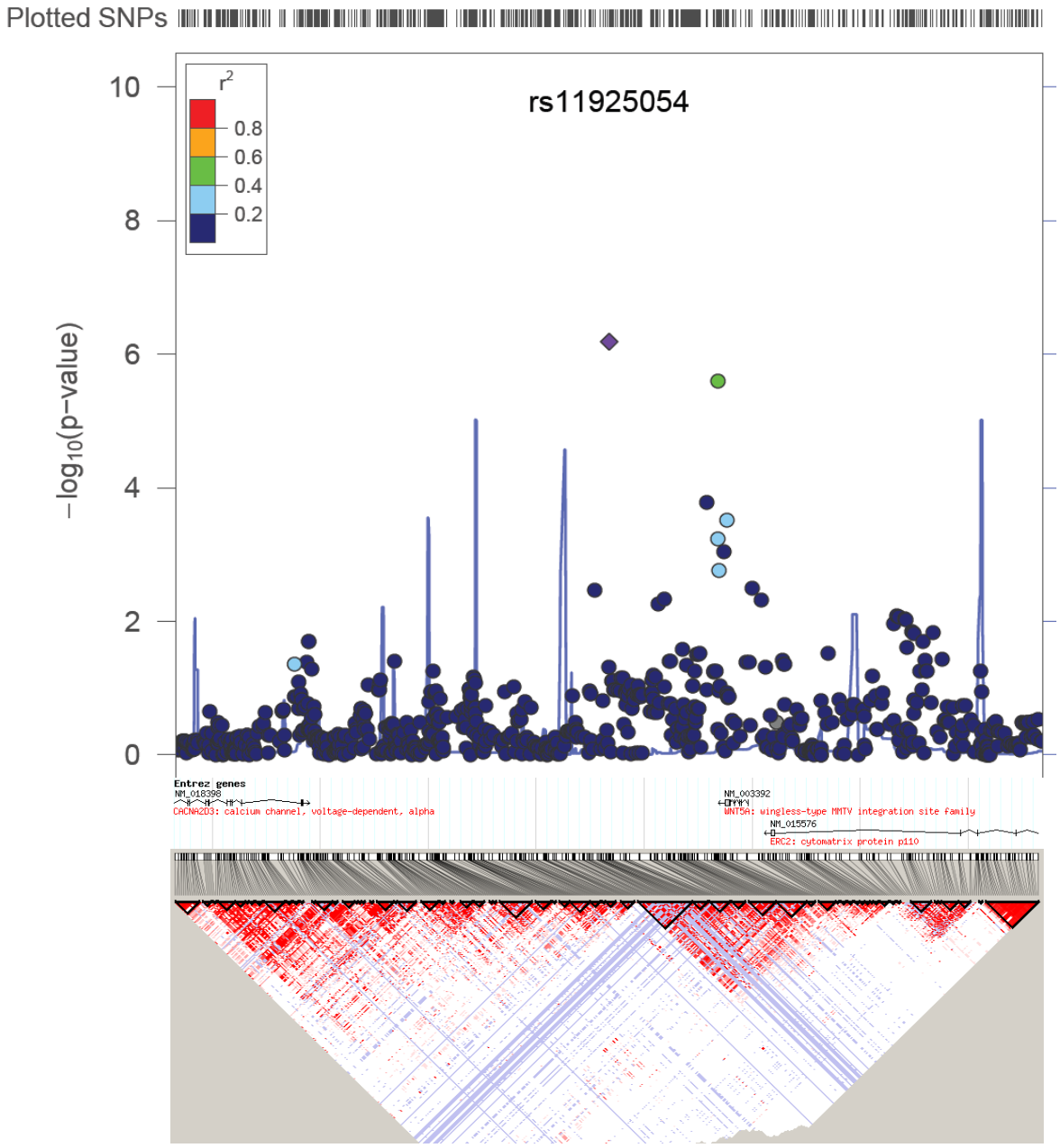


Figure A.12 Visualization of the 6p21.1 locus that was marked by rs7762544 ($P=1.1x10^{-7}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

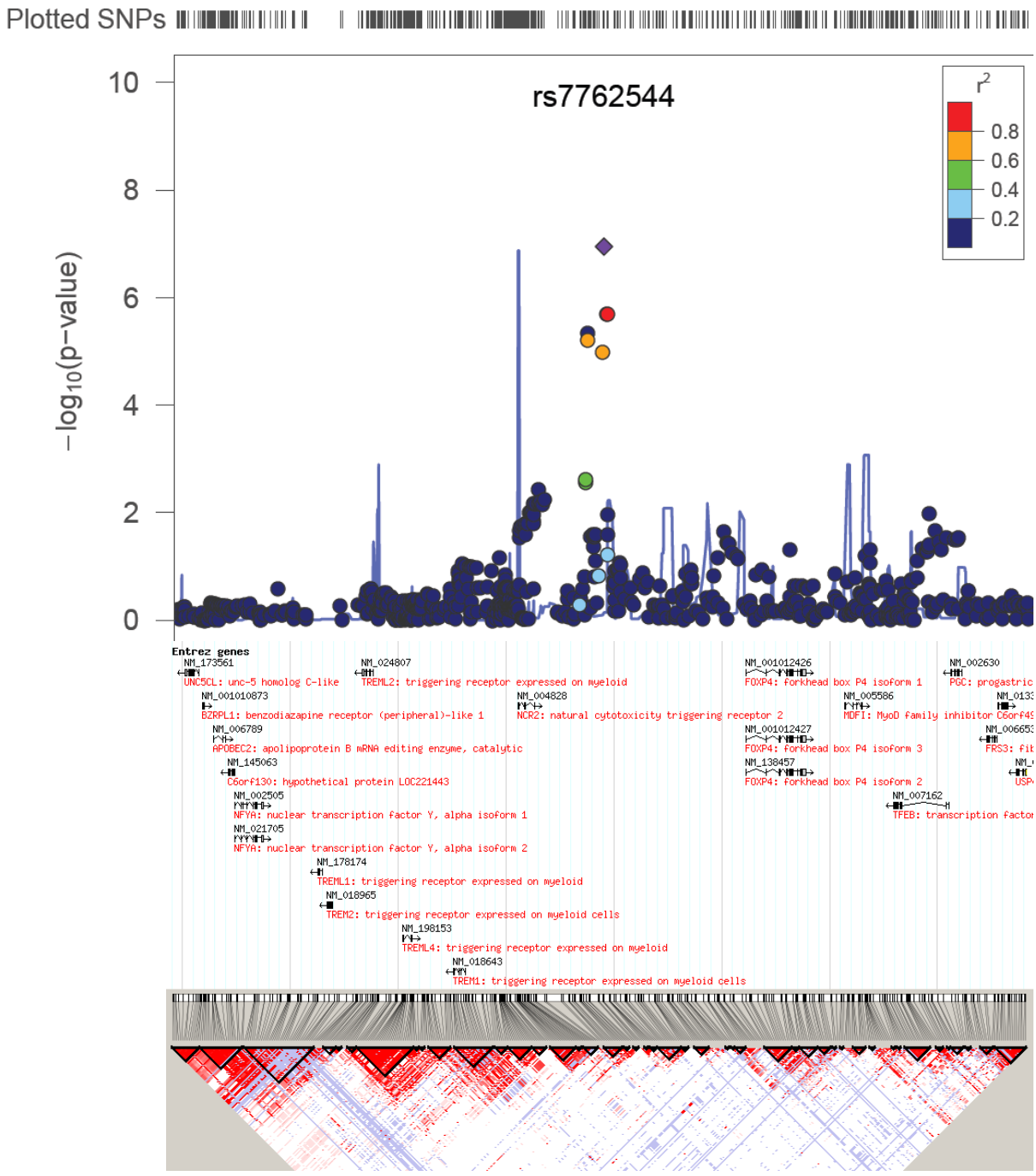


Figure A.13 Visualization of the 19p13.3 locus that was marked by rs3826782 ($P=4.0x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

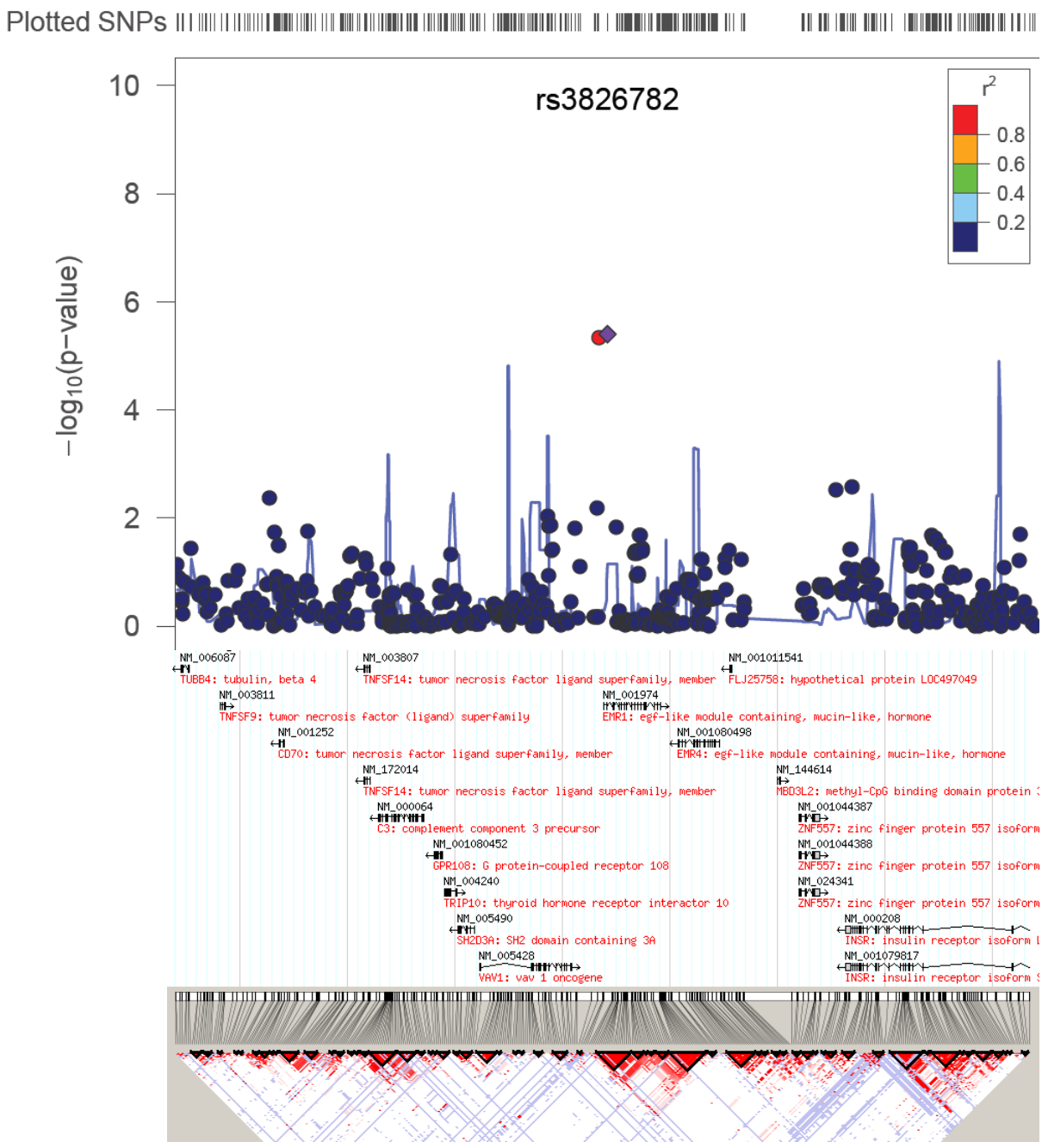


Figure A.14 Visualization of the 10p15 locus that was marked by rs12260727 ($P=6.0x10^{-7}$) for moderate chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

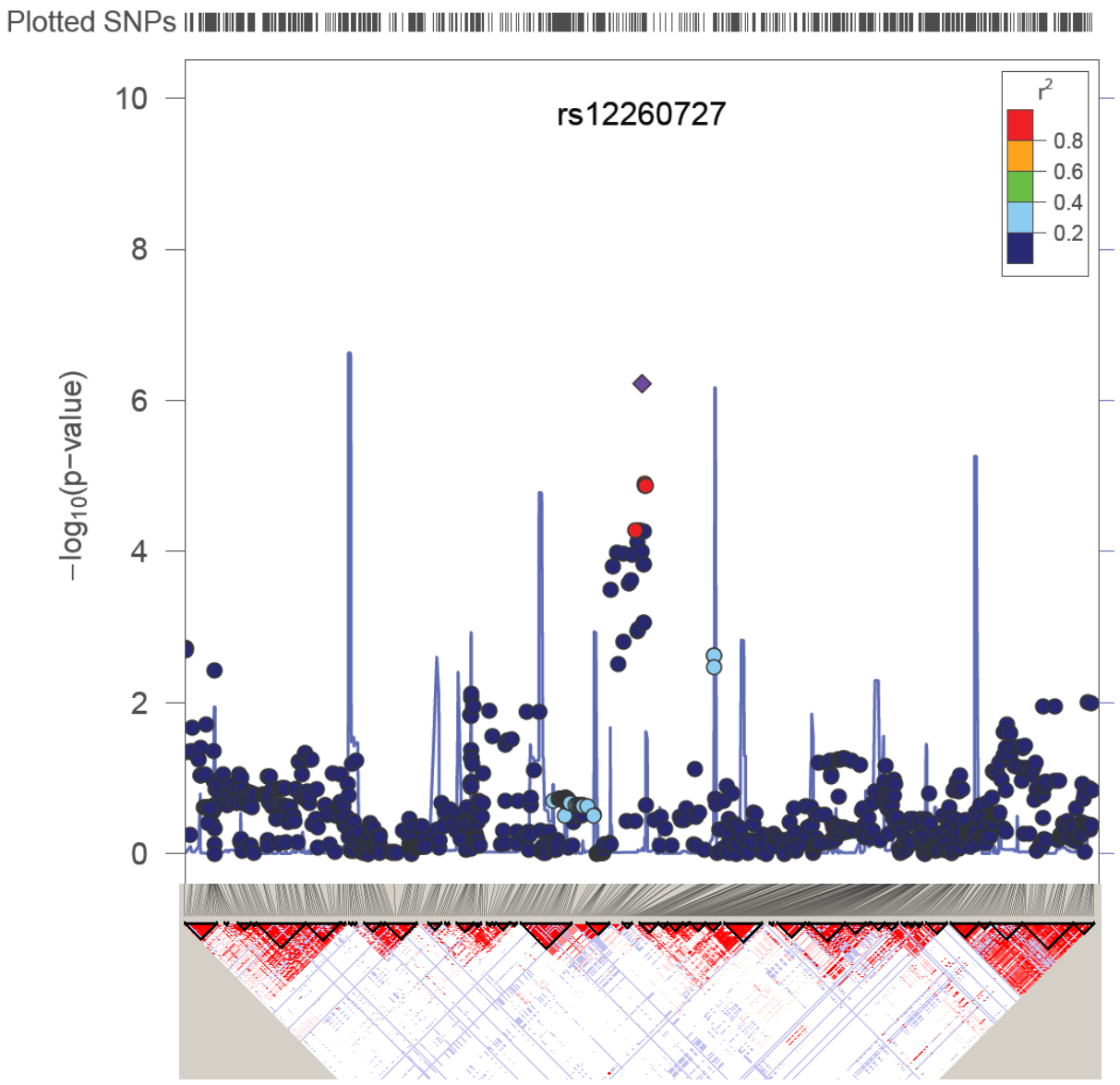


Figure A.15 Visualization of the 6q15 locus that was marked by rs17792917 ($P=1.8x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

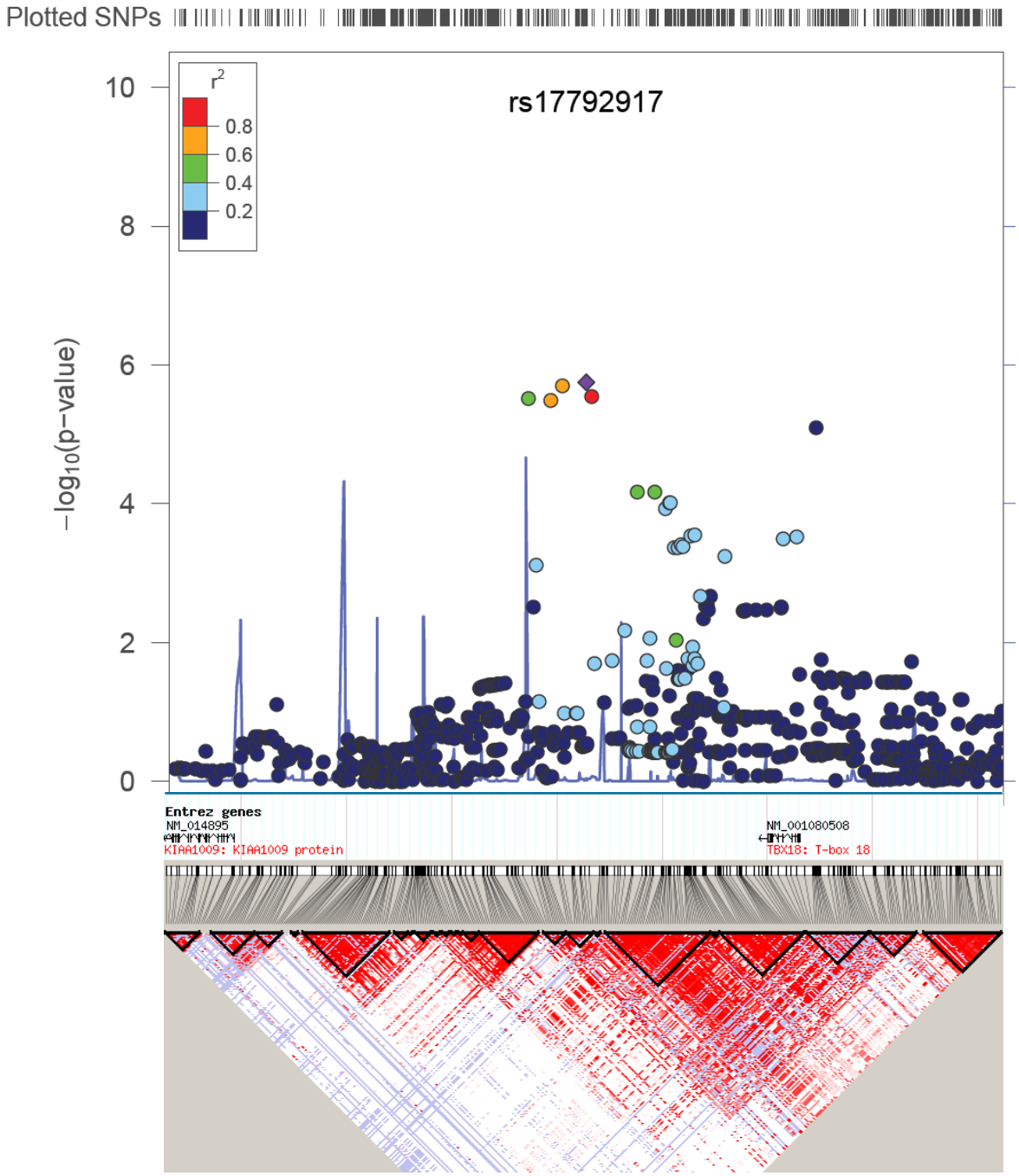


Figure A.16 Visualization of the 4q21.3 locus that was marked by rs17006135 ($P=4.4x10^{-6}$ after adjustment for smoking and diabetic status) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

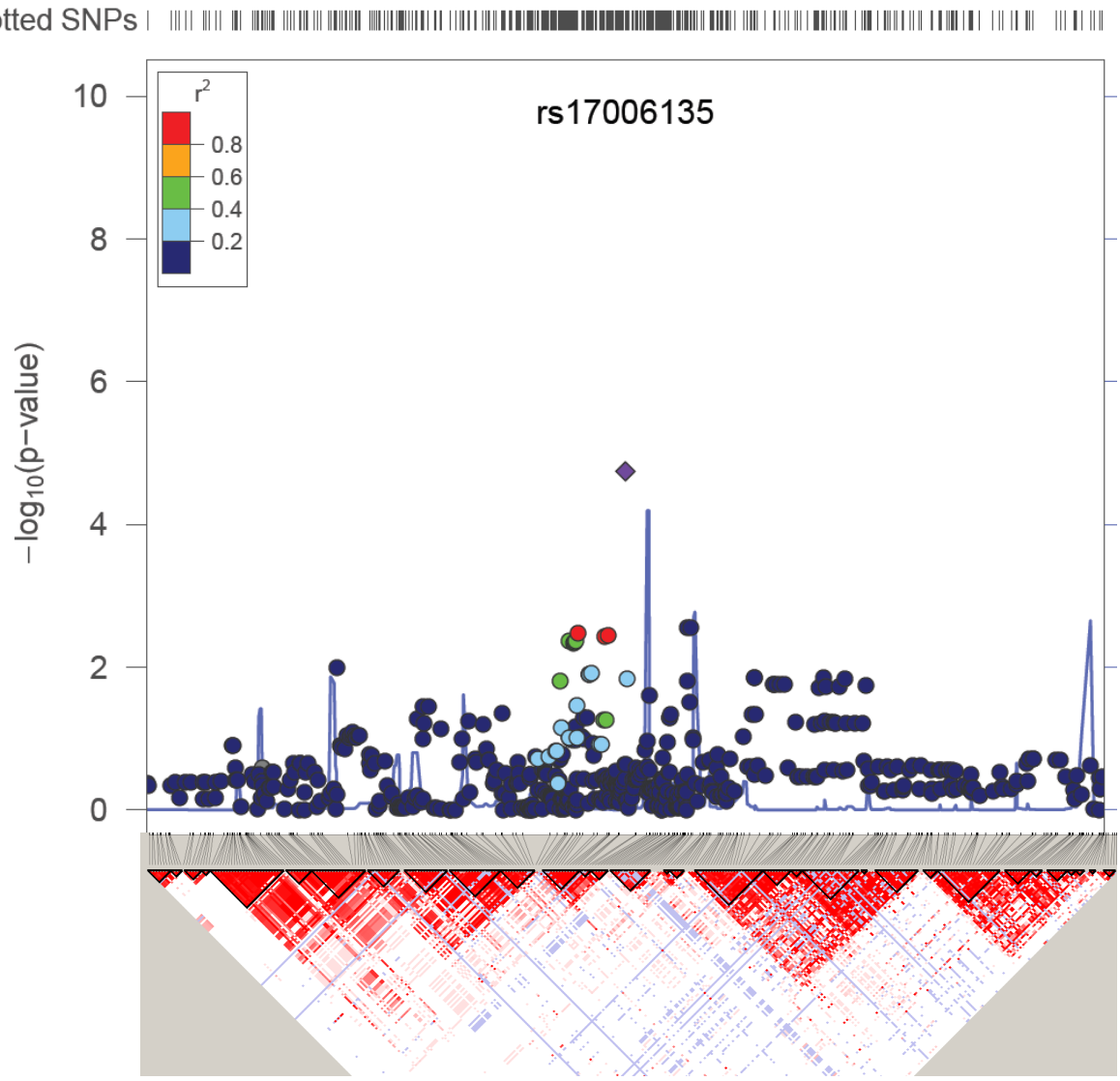


Figure A.17 Visualization of the 11q22 locus that was marked by rs7120142 ($P=3.9x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

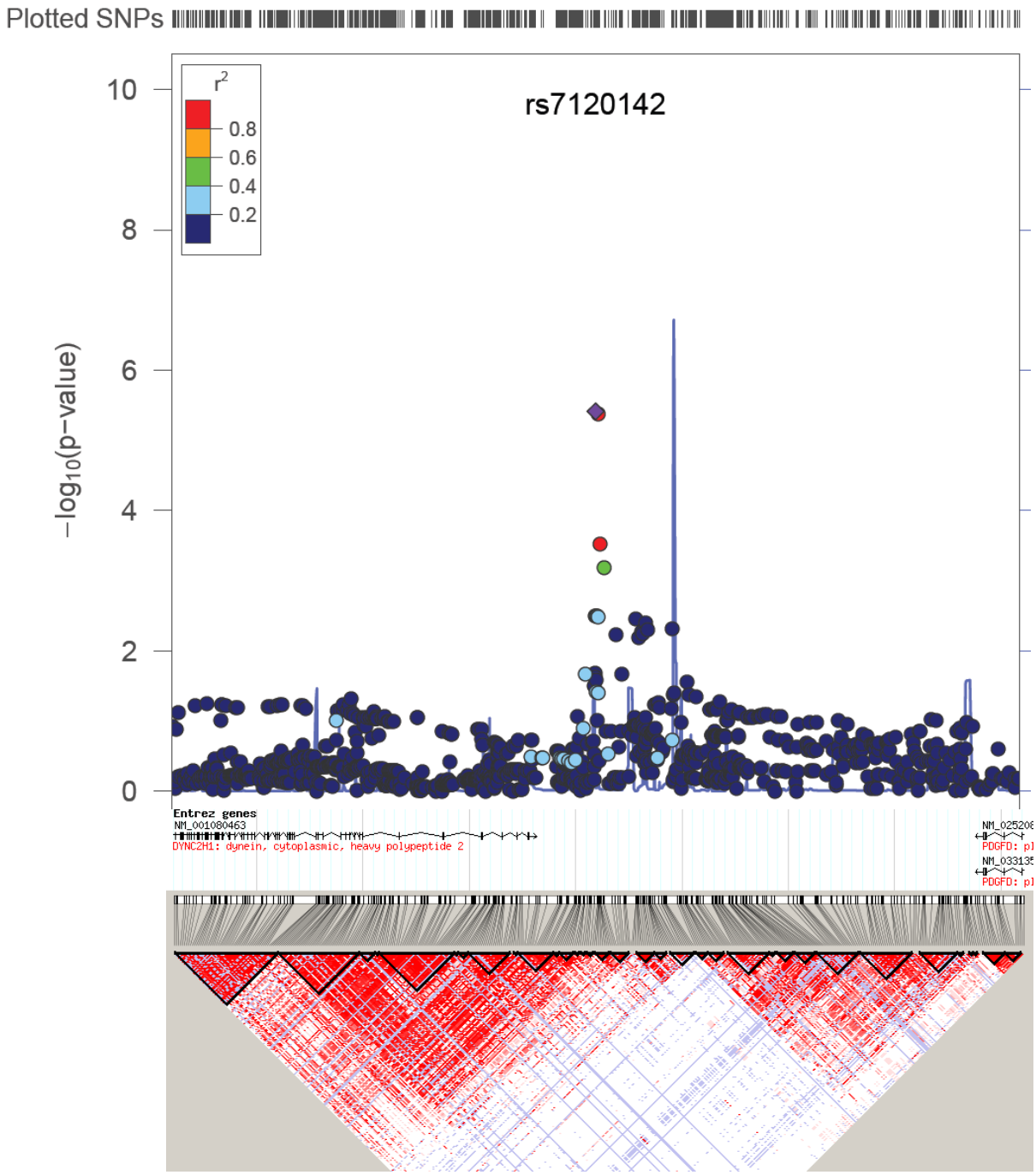


Figure A.18 Visualization of the 7q34 locus that was marked by rs10500130 ($P=4.6x10^{-6}$) for severe chronic periodontitis among the 4610 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

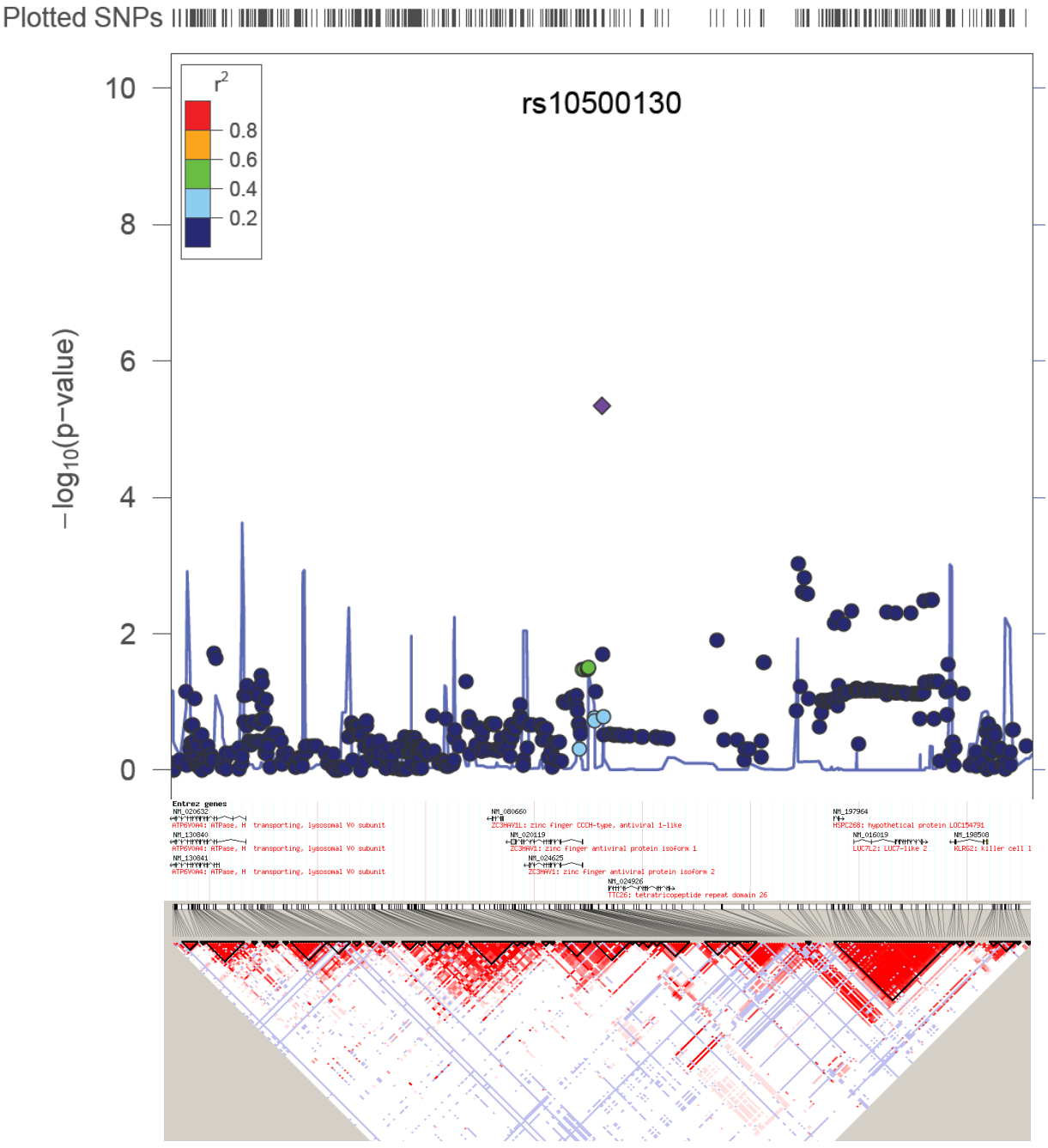


Figure A.19 Visualization of the 1q42 locus that was marked by rs11800854 ($P=4.0x10^{-6}$) for "high" *Aggregatibacter actinomycetemcomitans* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

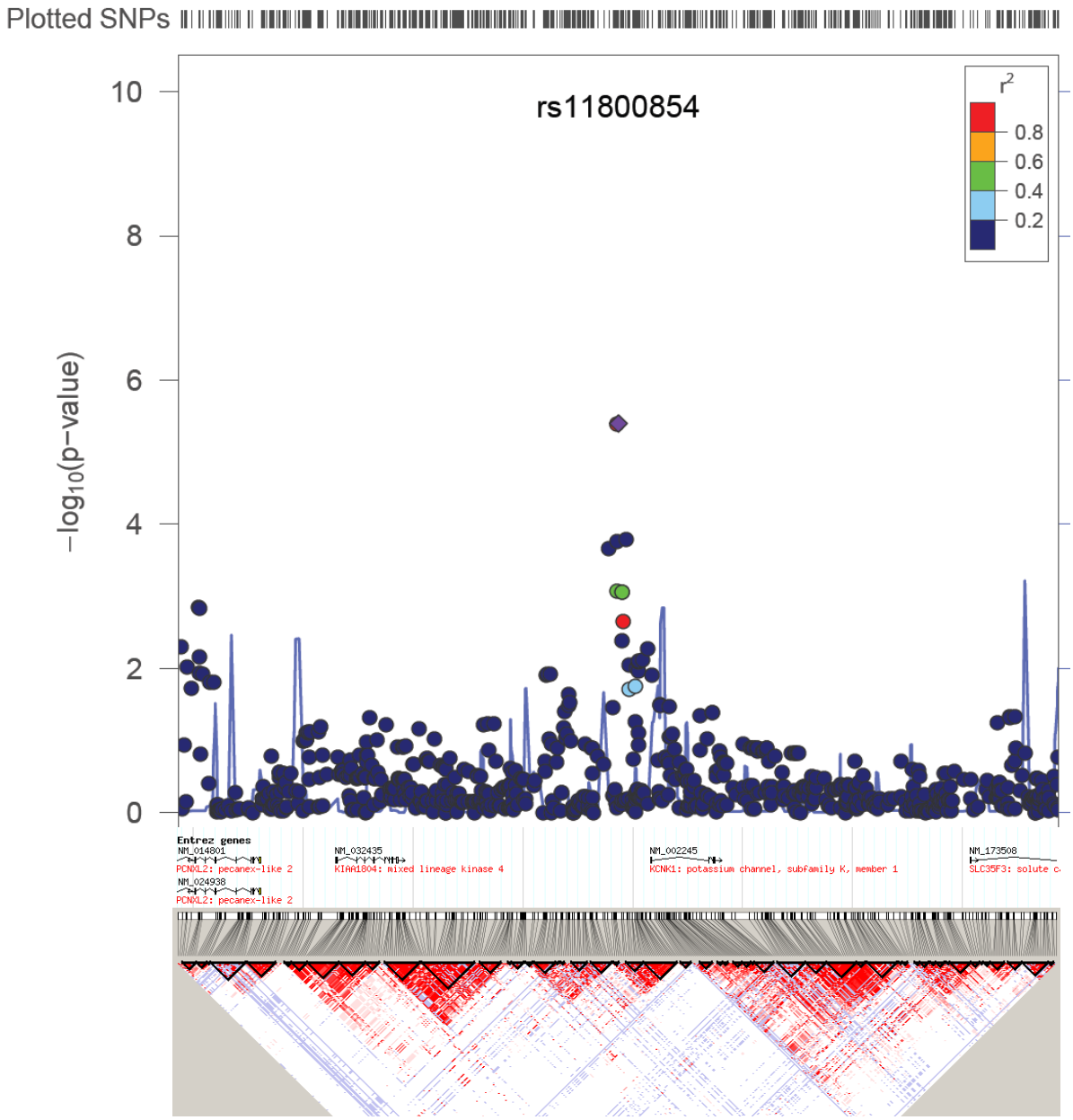


Figure A.20 Visualization of the 1p22 locus that was marked by rs12032672 ($P=9.6x10^{-7}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

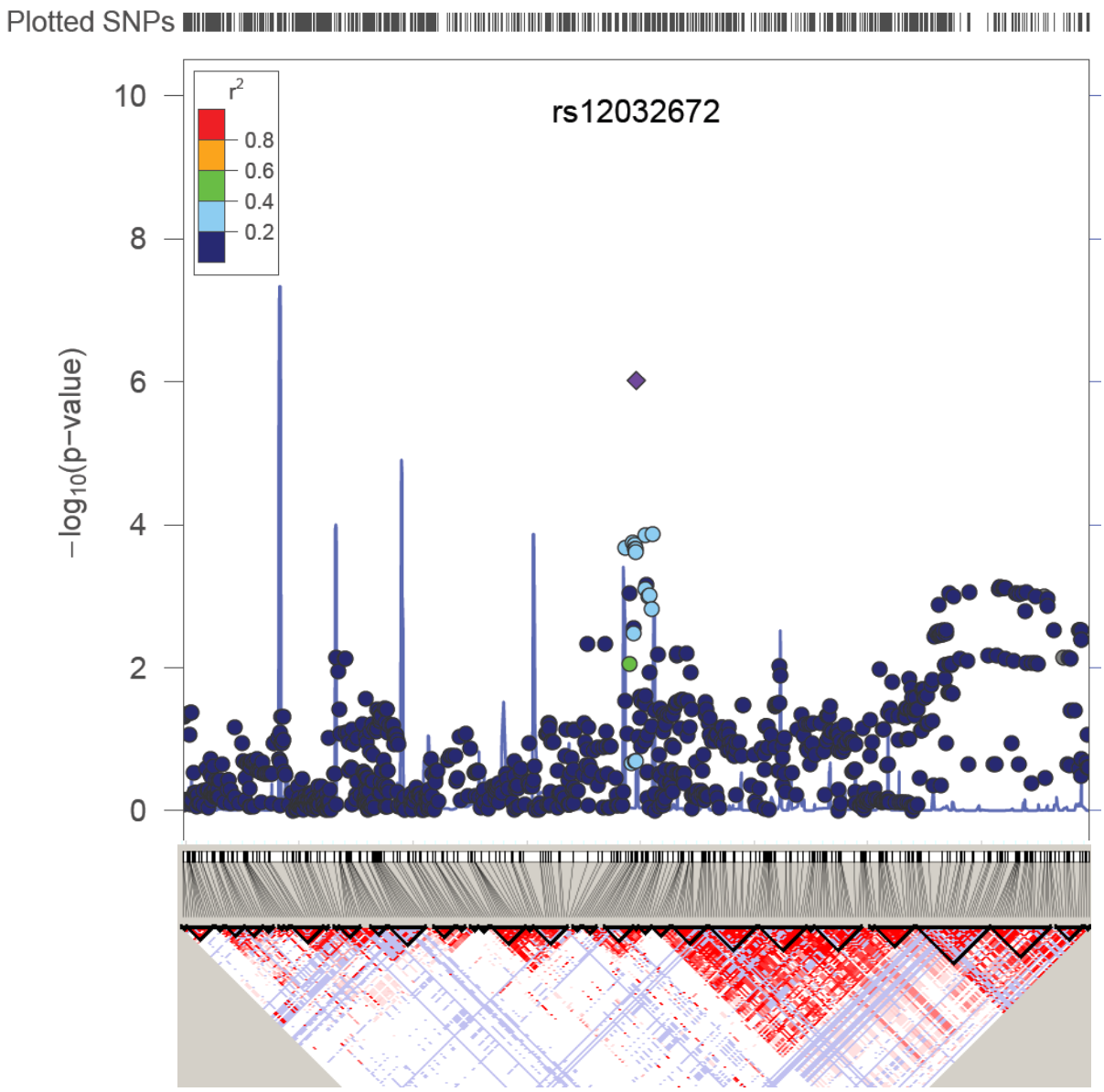


Figure A.21 Visualization of the 5q33 locus that was marked by rs10043775 ($P=2.4x10^{-6}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

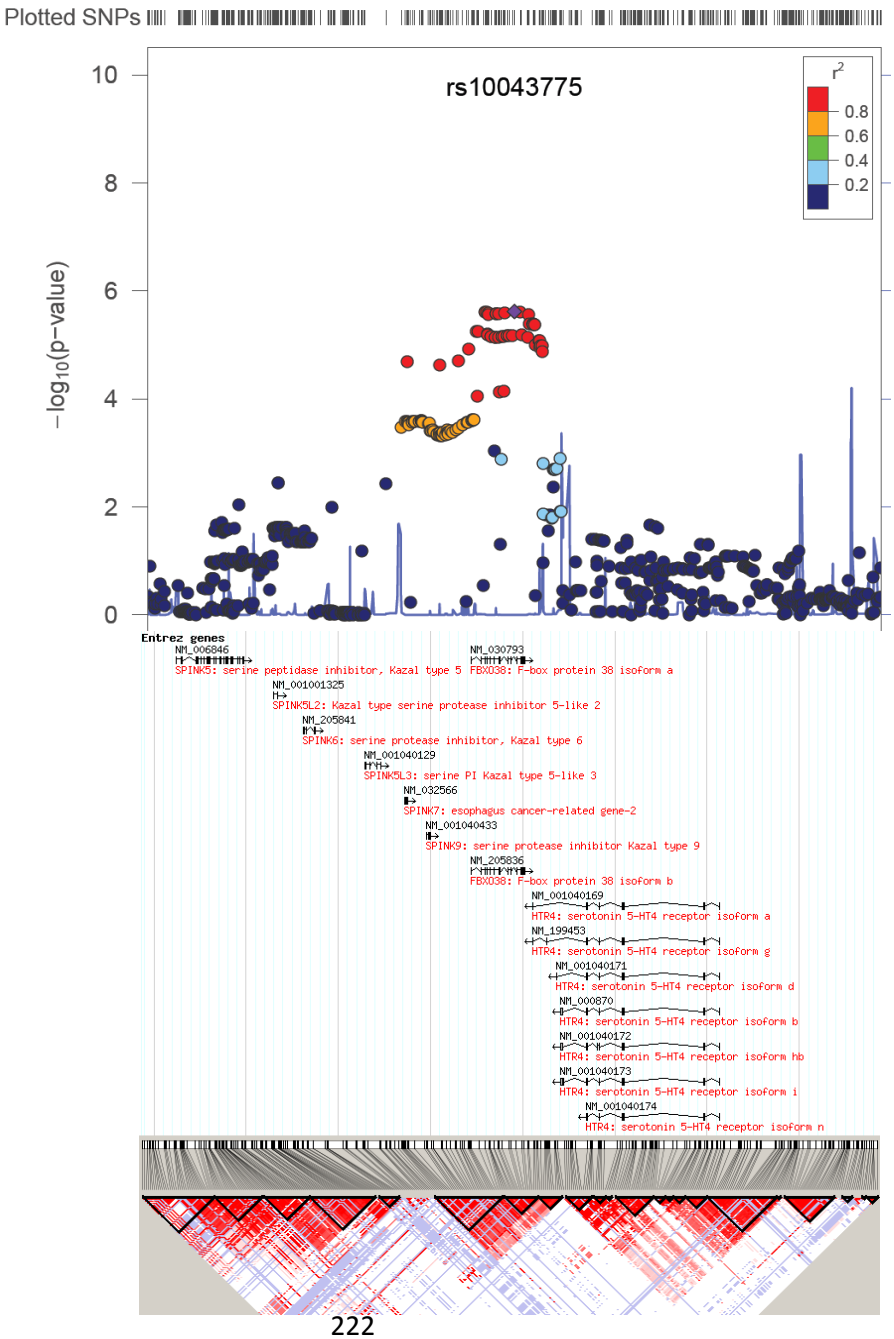


Figure A.22 Visualization of the 9p24 locus that was marked by rs16924631 ($P=3.2x10^{-6}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

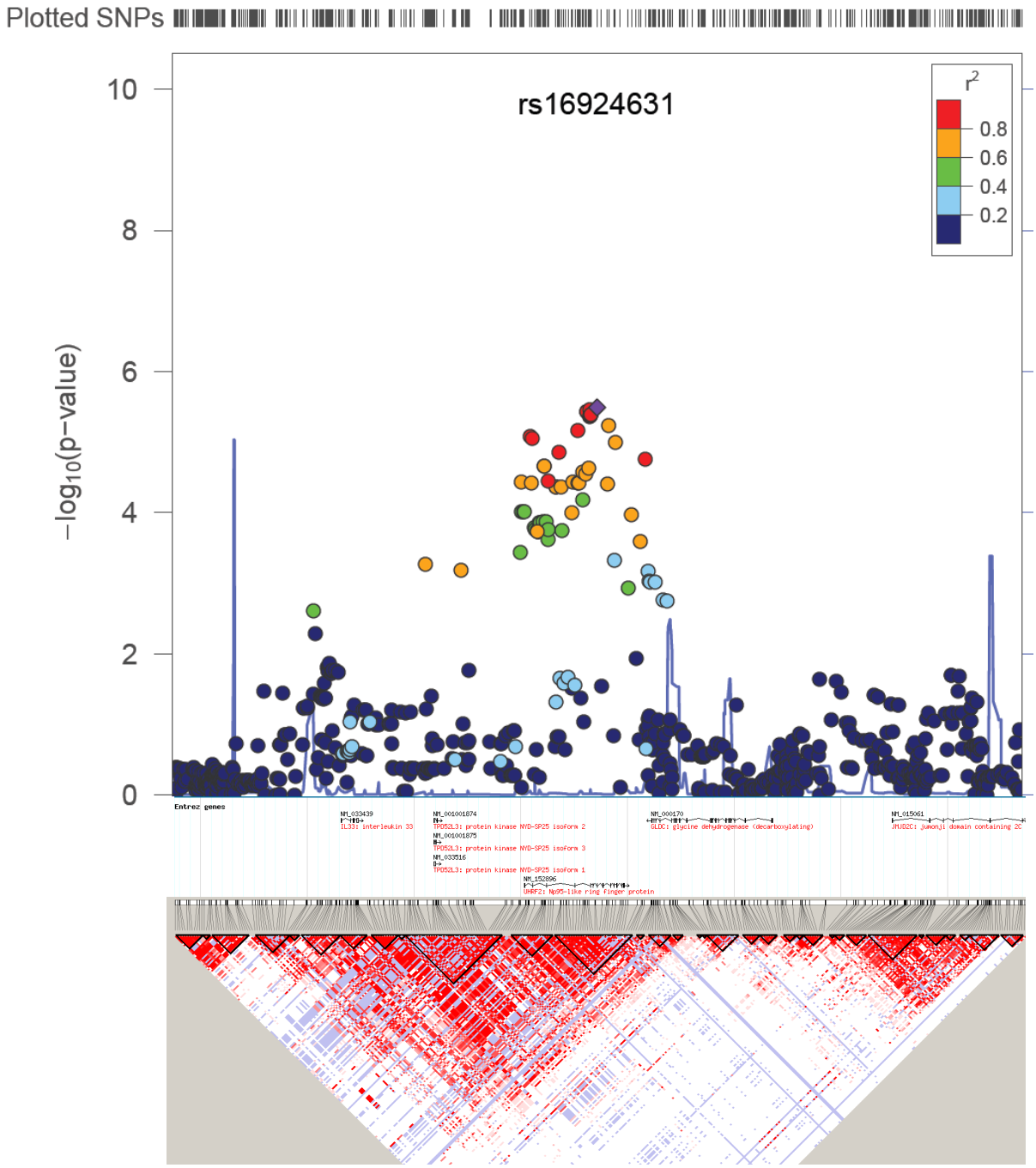


Figure A.23 Visualization of the 4p14 locus that was marked by rs10010758 ($P=3.7x10^{-6}$) for "high red" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

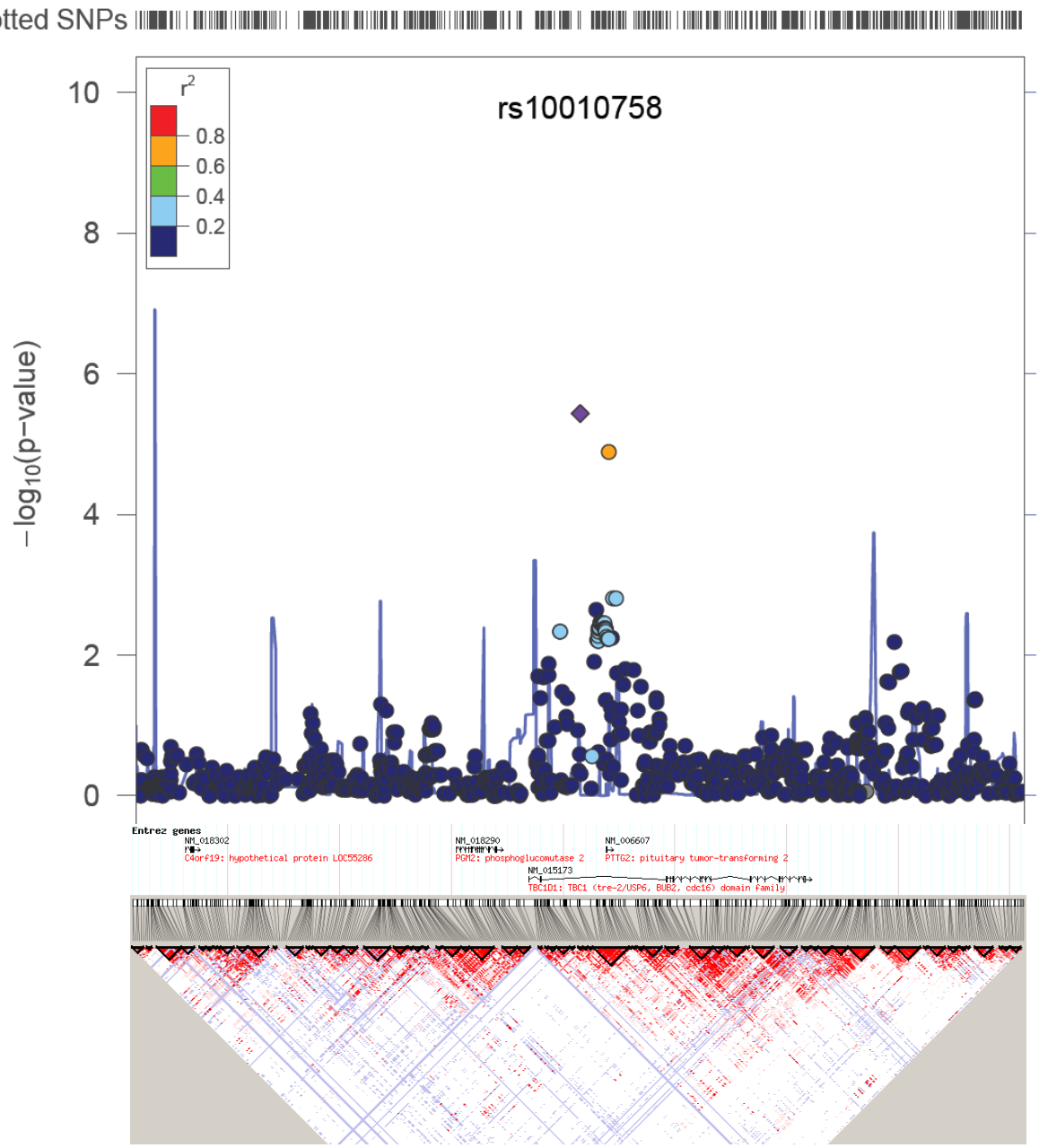


Figure A.24 Visualization of the 1p36.2 locus that was marked by rs1616122 ($P=4.9x10^{-6}$) for "high orange" complex bacterial colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

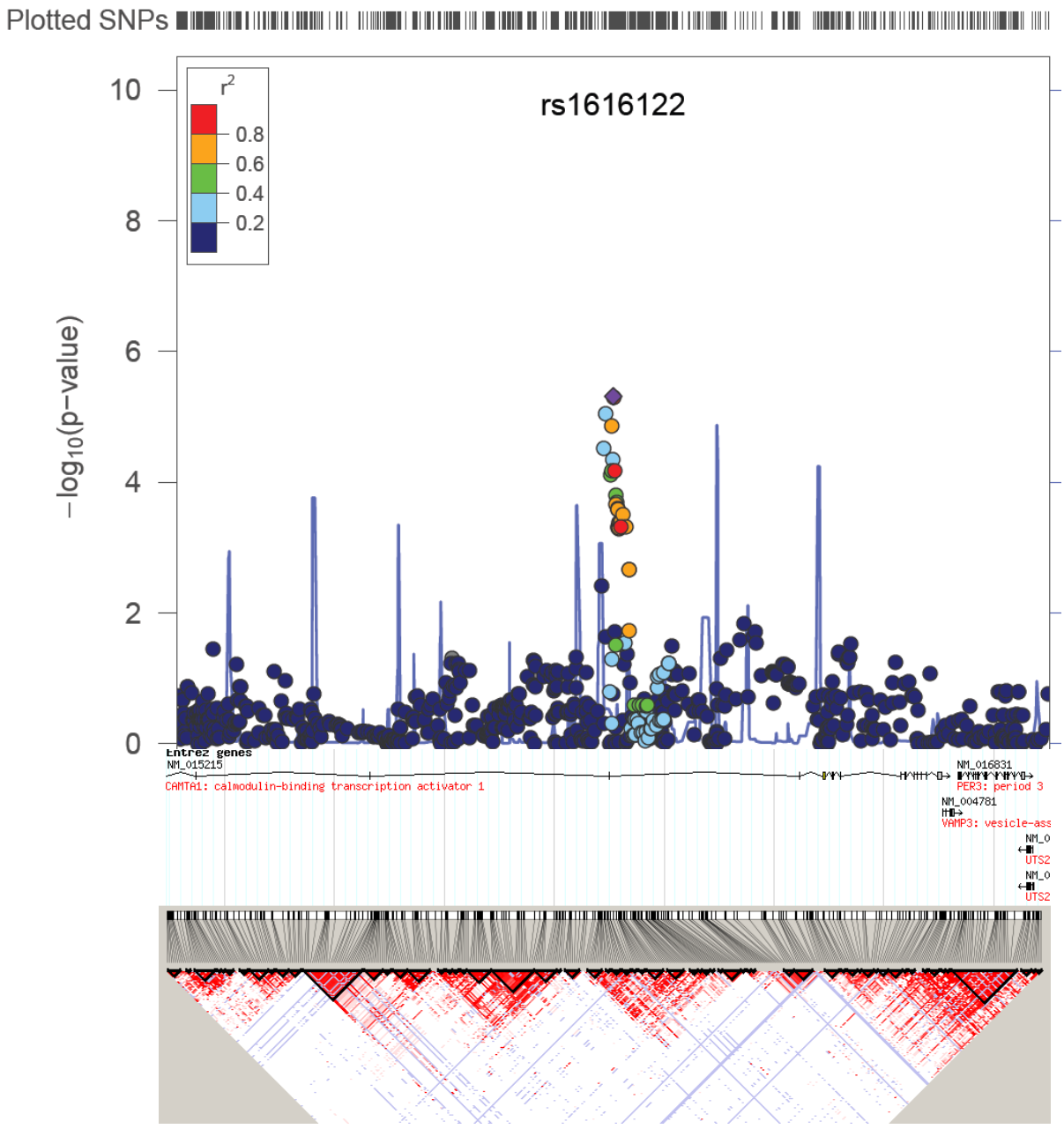


Figure A.25 Visualization of the 1q42 locus that was marked by rs11800854 ($P=4.0x10^{-6}$) for "high" *Aggregatibacter actinomycetemcomitans* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

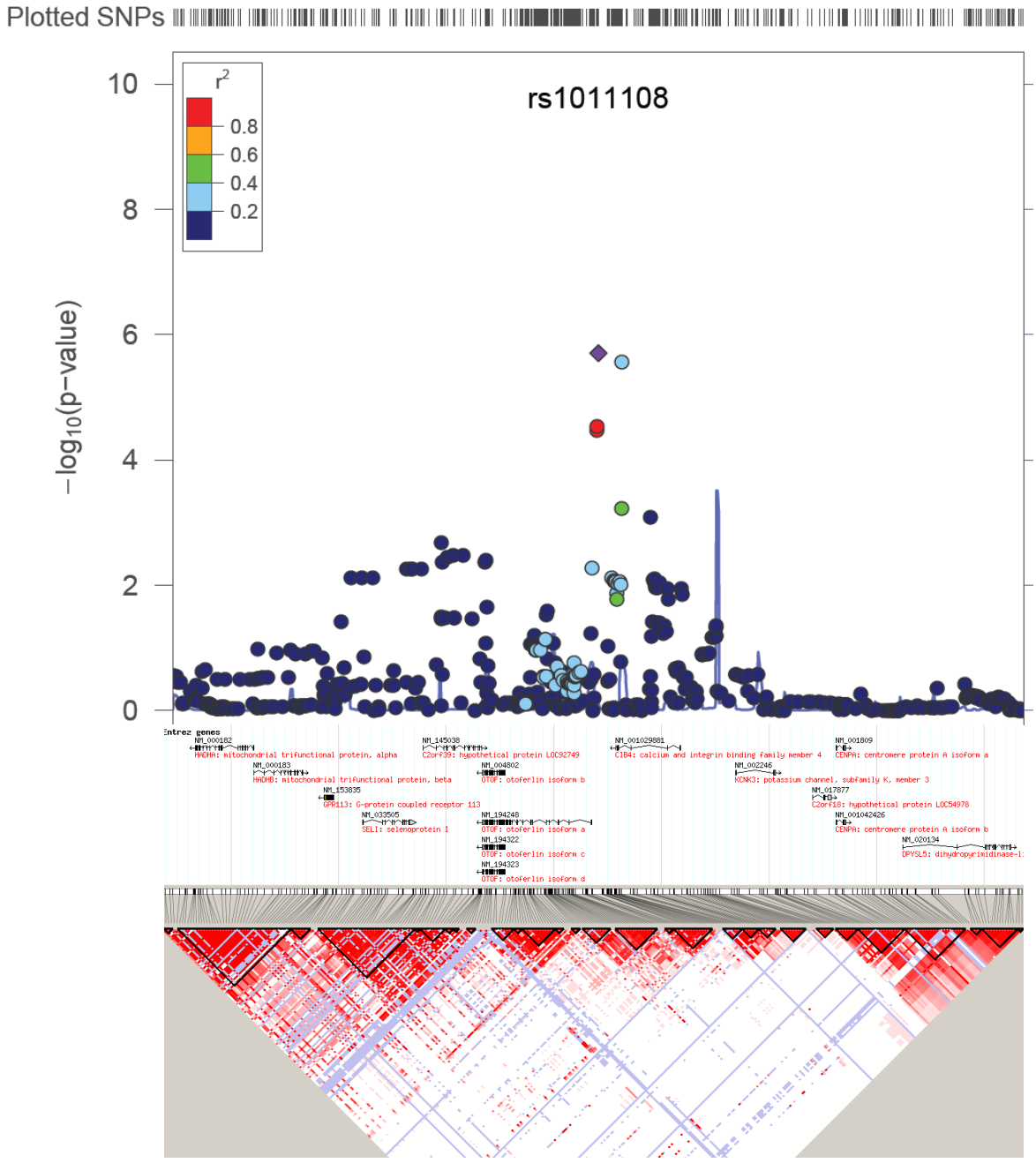


Figure A.26 Visualization of the 9q33.2 locus that was marked by rs10760187 ($P=4.5x10^{-7}$) for "high" *Porphyromonas gingivalis* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

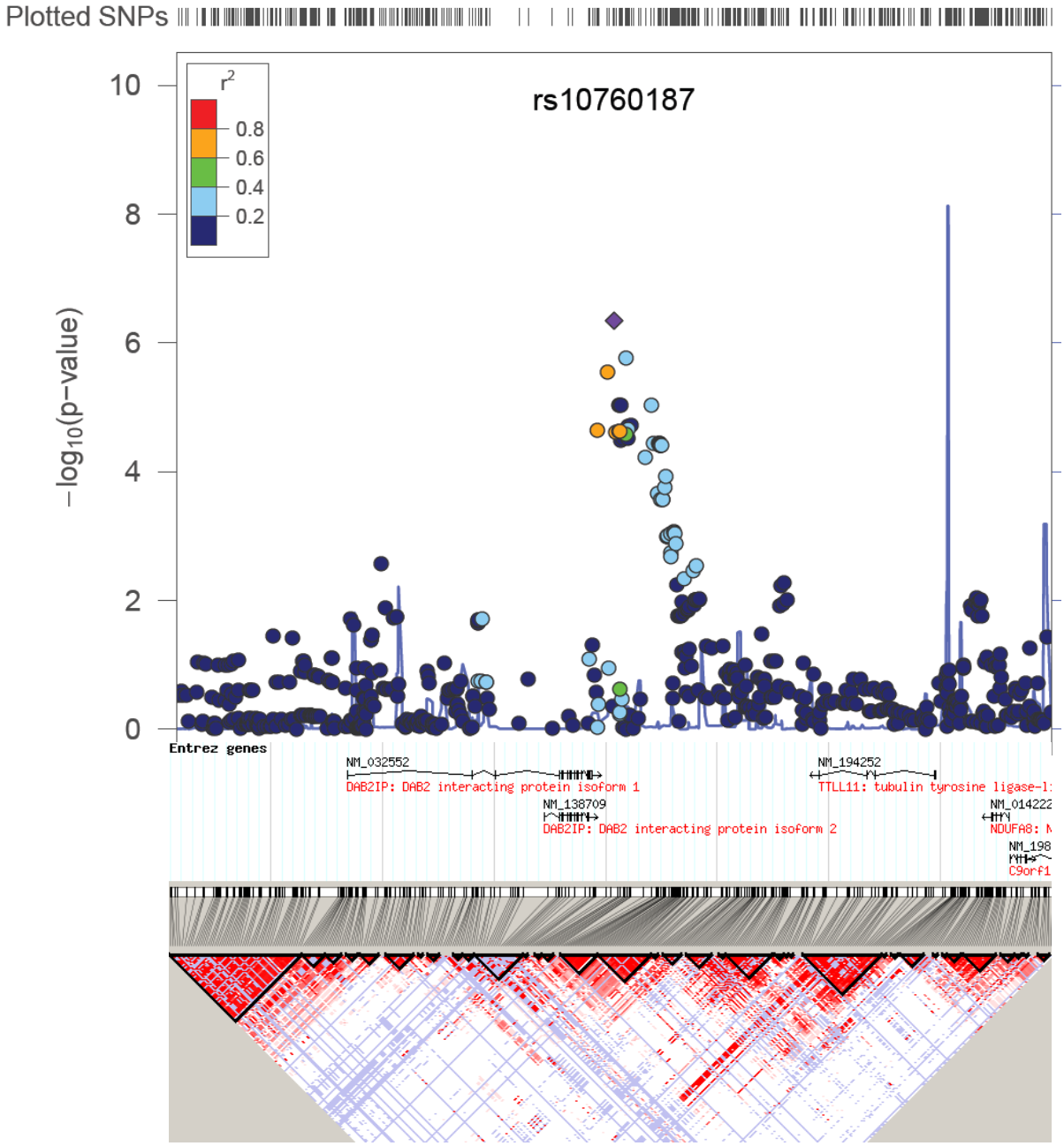


Figure A.27 Visualization of the 10p11.21 locus that was marked by rs1360573 ($P=1.5x10^{-6}$) for "high" *Porphyromonas gingivalis* colonization among the subset of 1020 white participants of the Dental Atherosclerosis in Communities Study cohort combined with a linkage disequilibrium plot generated with Haploview (based on 1000 genomes pilot 1 release haplotype data).

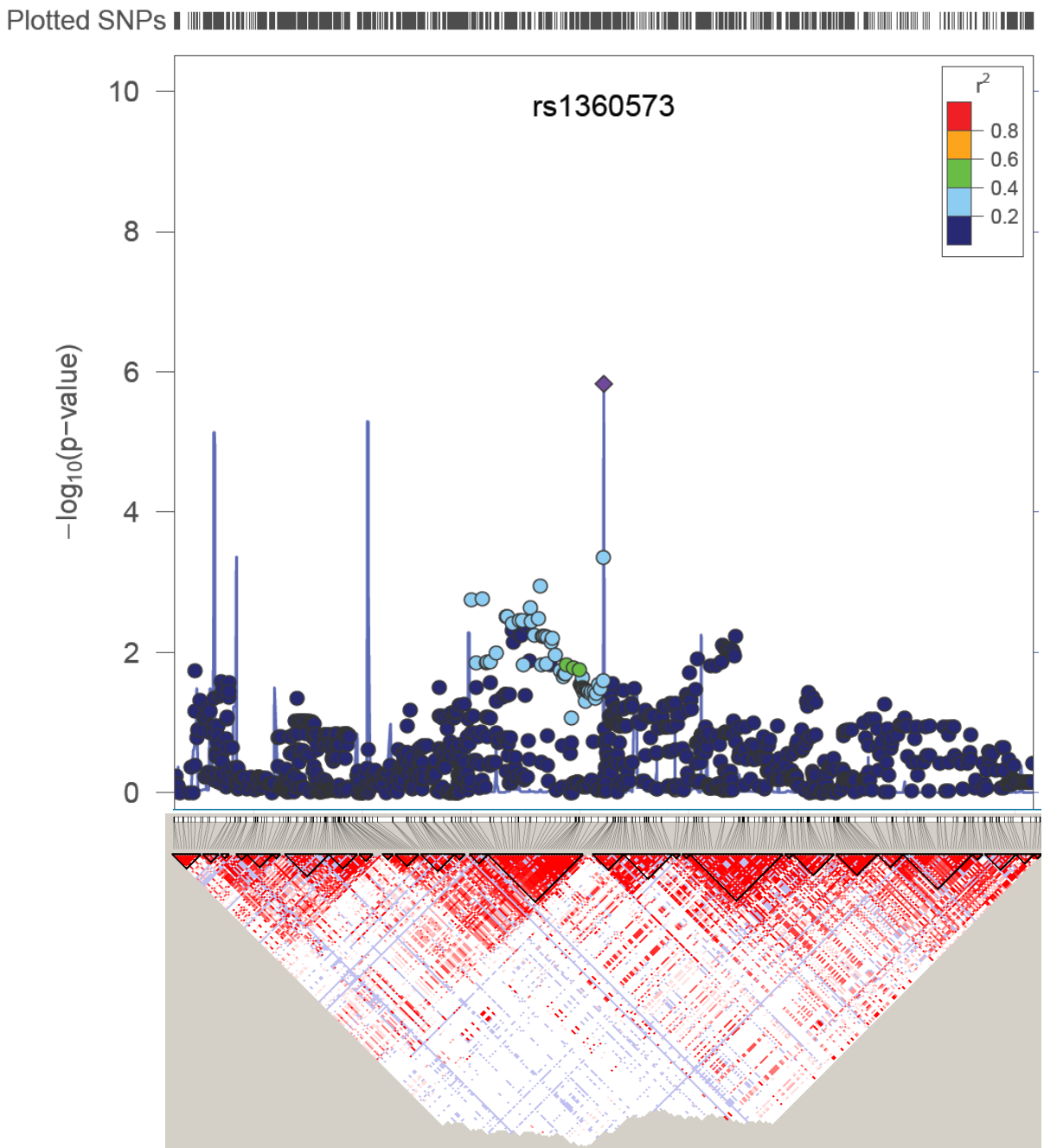


Figure A.28 Predicted network of associations [physical interactions (grey line), co-expression (dark blue line), pathway (light blue line), shared protein domain (green line)] among genes of prioritized loci for the CP and the "high" bacterial traits in the Dental ARIC cohort. Grey lines represent co-expression and blue lines represent physical interactions.

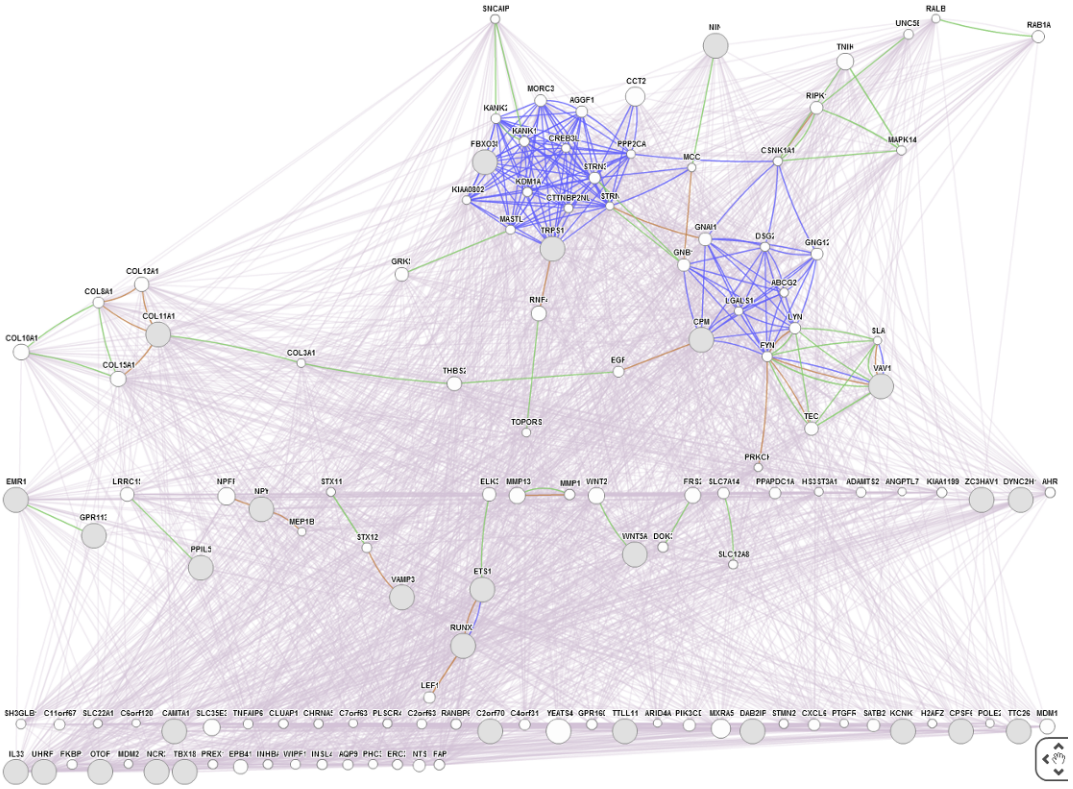


Figure A.29 Predicted network of associations [physical interactions (grey line), co-expression (dark blue line), pathway (light blue line), shared protein domain (green line)], generated with the GeneMANIA module) among genes of prioritized loci for the CP and the "high" bacterial traits in the Dental ARIC cohort, along with genes that have been previously reported to be associated with periodontitis.

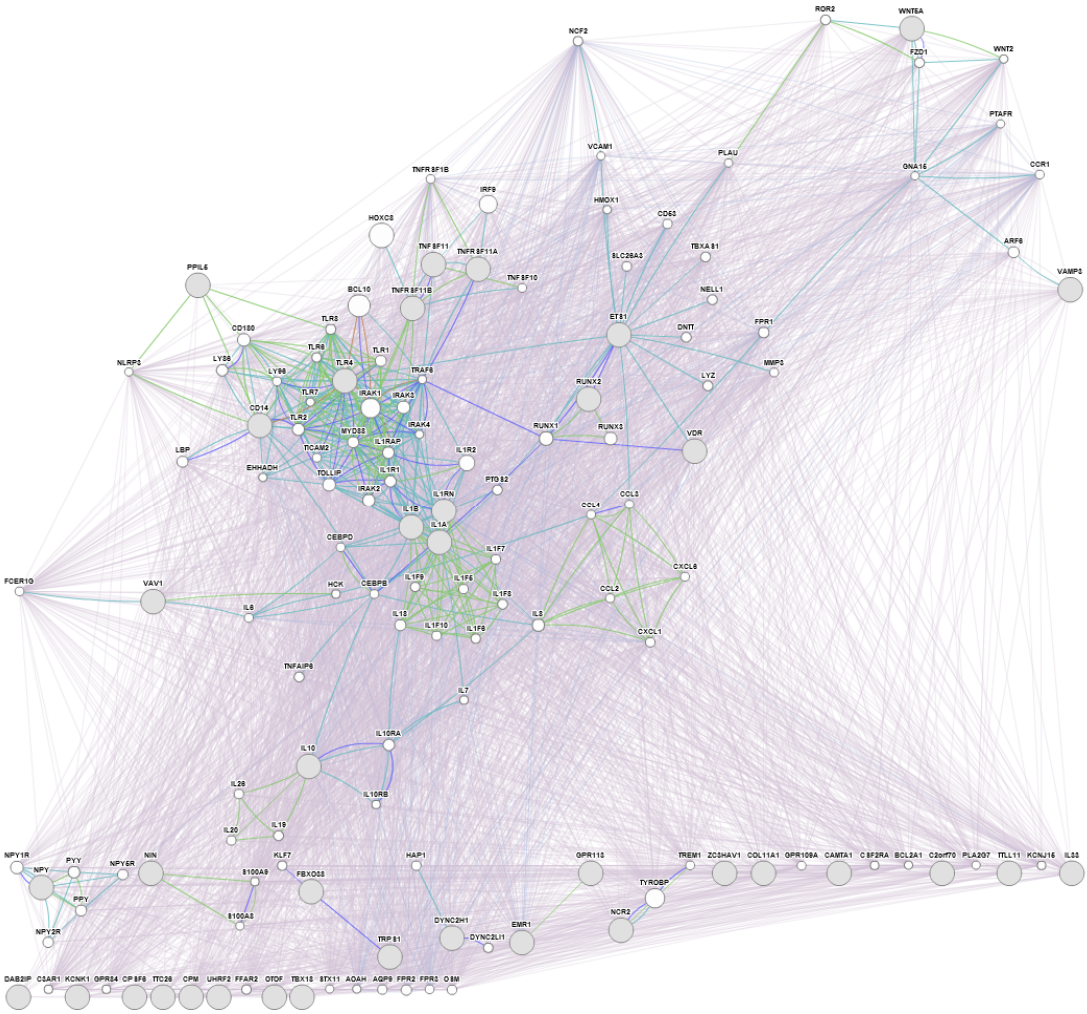


Figure A.30 Visualization of GWA results for severe CP in the ARIC cohort for loci that have been previously associated with periodontitis (*IL1A*, *IL1RN*, *TLR4*, *MMP2*, *GLT6D1*, *GCGR3A*). The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

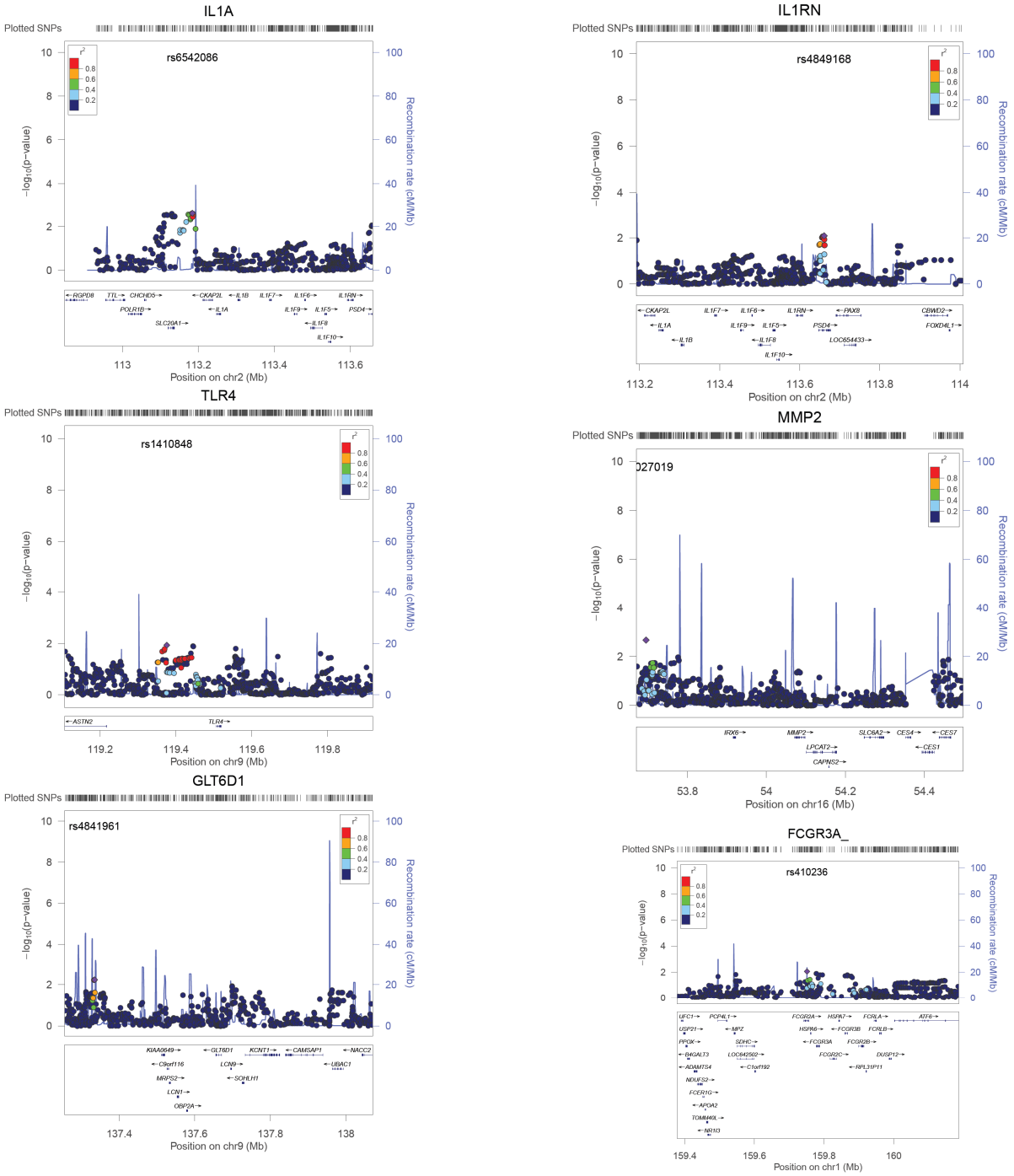


Figure A.31 Visualization of GWA results for moderate CP in the ARIC cohort for loci that have been previously associated with periodontitis (*IL1A*, *IL1RN*, *TLR4*, *MMP2*, *GLT6D1*, *GCGR3A*). The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

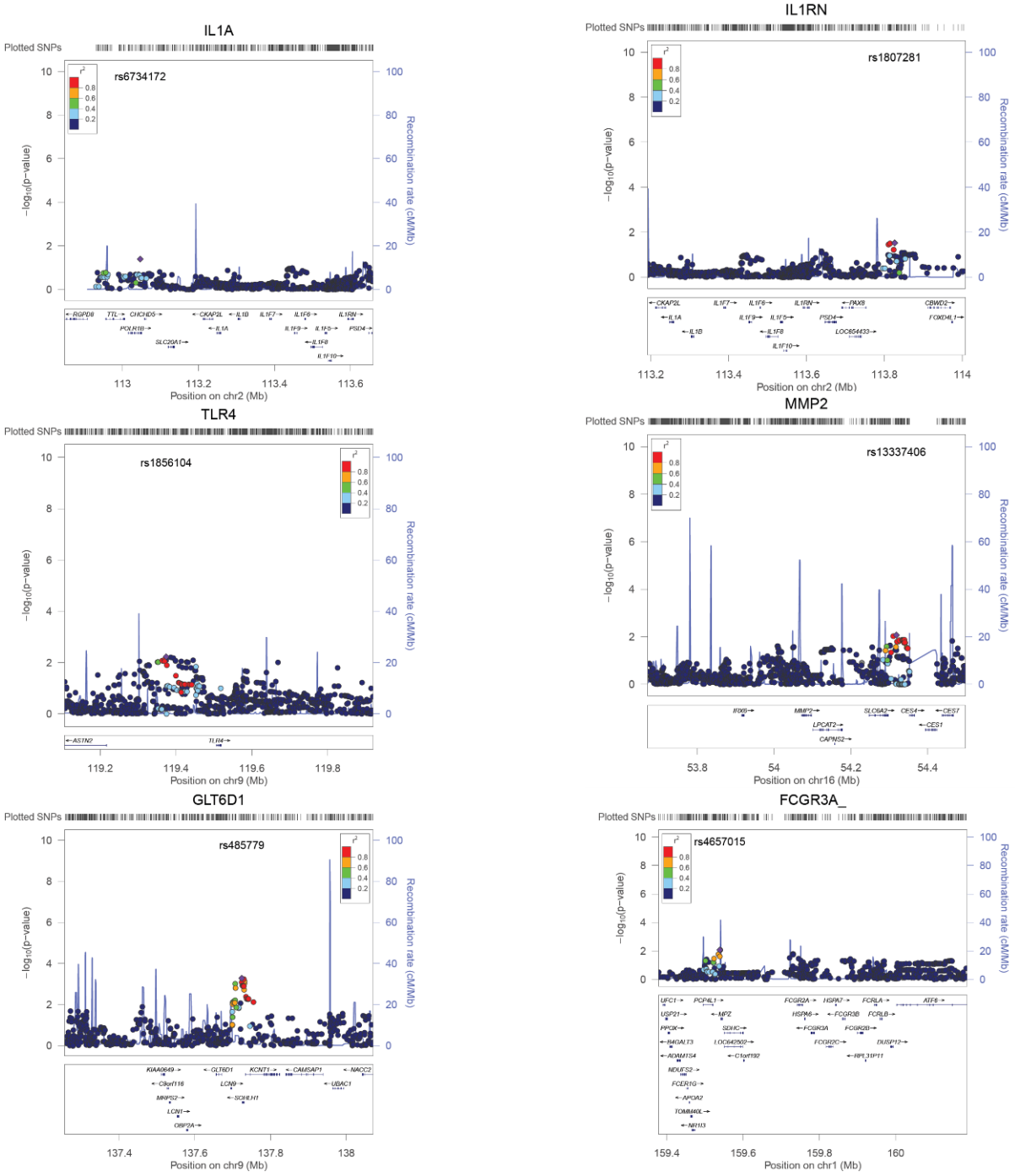


Figure A.32 Visualization of GWA results for the extent of attachment loss trait (proportion of sites exhibiting ≥3mm attachment loss) CP in the ARIC cohort for loci that have been previously associated with periodontitis (*IL1A*, *IL1RN*, *TLR4*, *MMP2*, *GLT6D1*, *GCGR3A*). The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

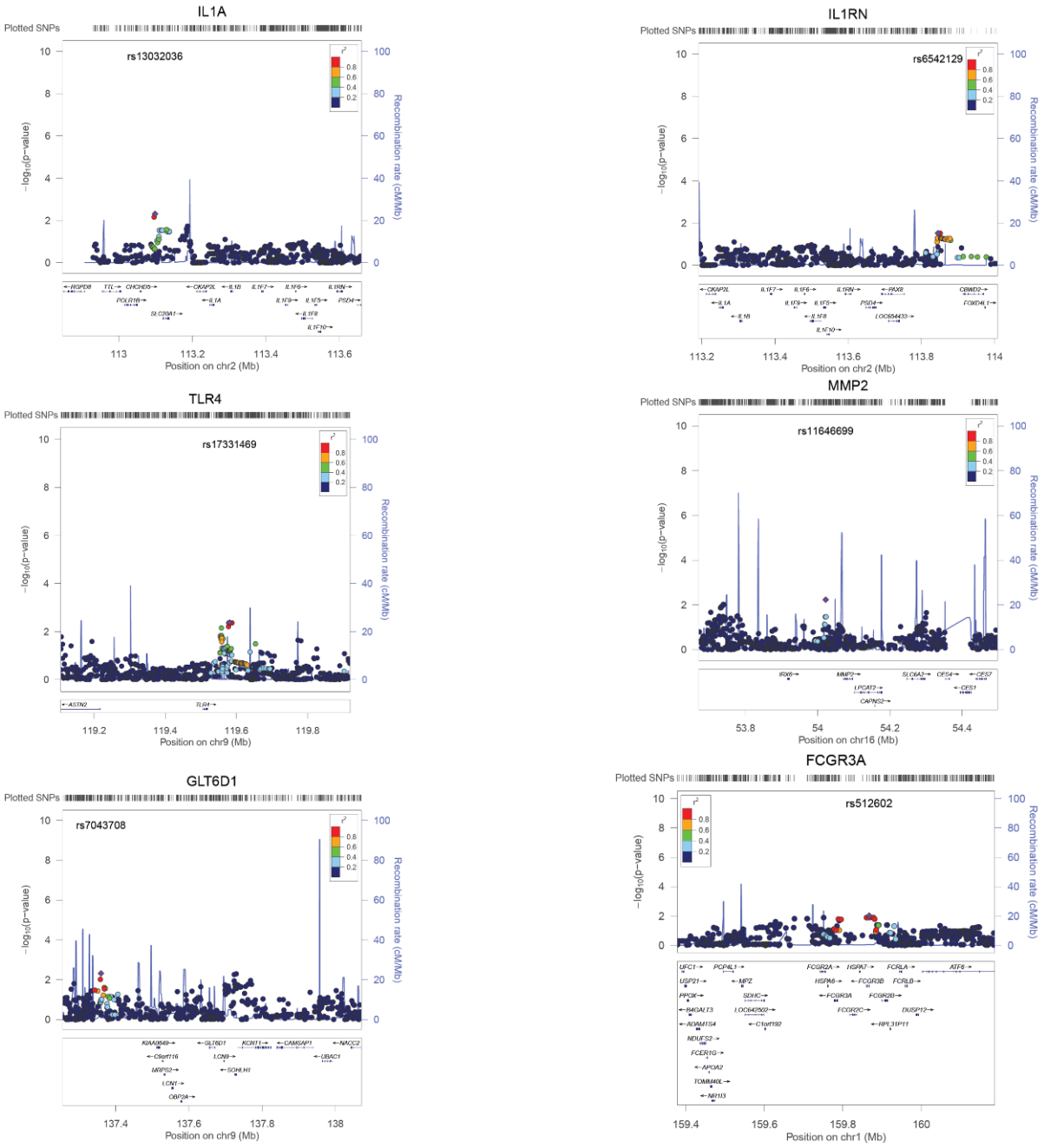


Figure A.33 Visualization of GWAS results in the *IL6* (Interleukin 6) and *VDR* (vitamin D receptor) loci in the ARIC cohort for severe (left column) and moderate CP (right column). The vertical axis corresponds to each marker's associated –log10 P-value. The overlaid recombination rate plot and the color-coded pairwise linkage disequilibrium values with index SNPs were calculated based on HapMap II – CEU (human genome 18, build 36).

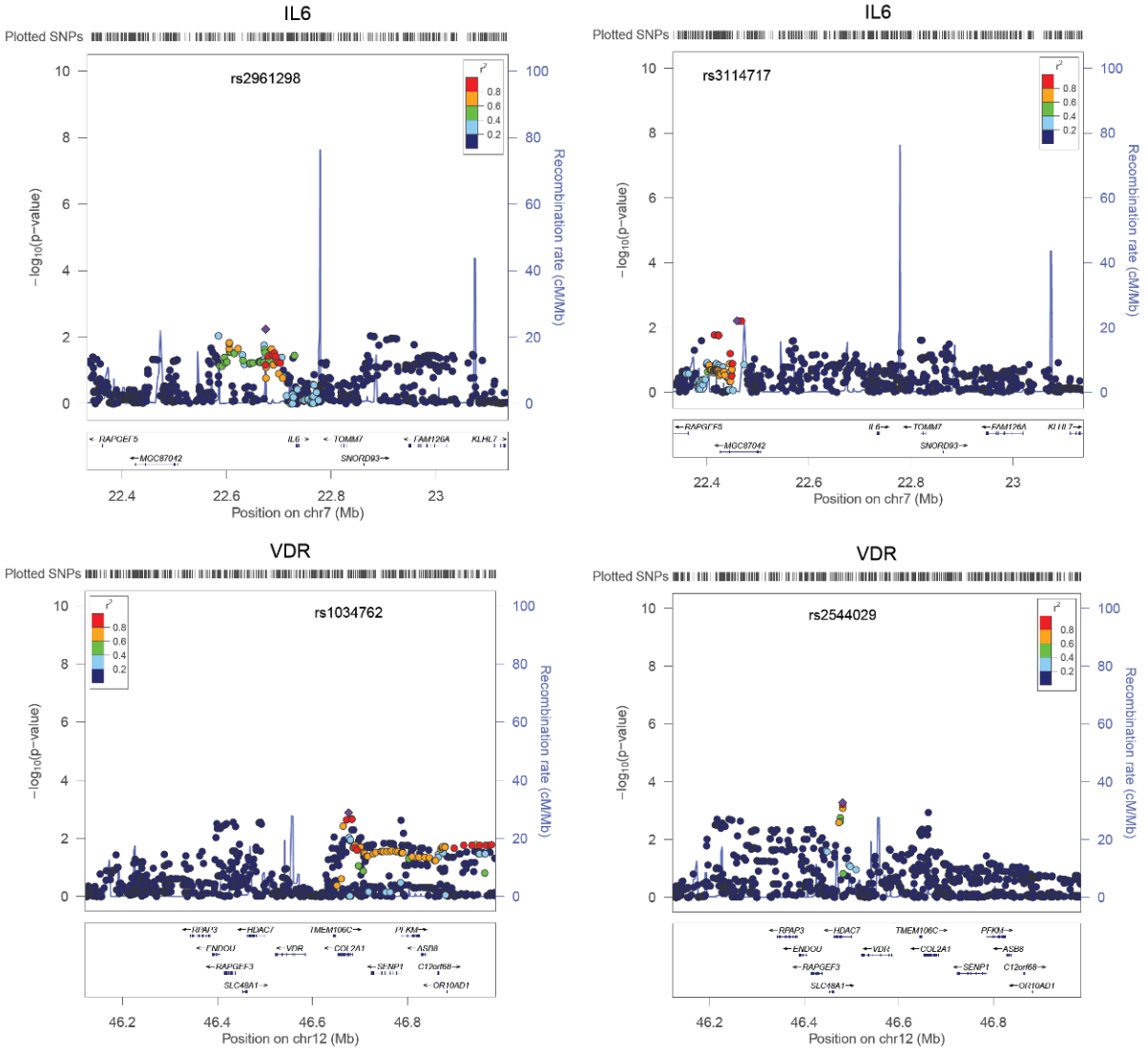


Figure A.34 Visualization of the 6p21.1 locus (chr6:41,279,315-41,479,315; flanking 100Kb of rs7762544, which was prioritized for moderate CP) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

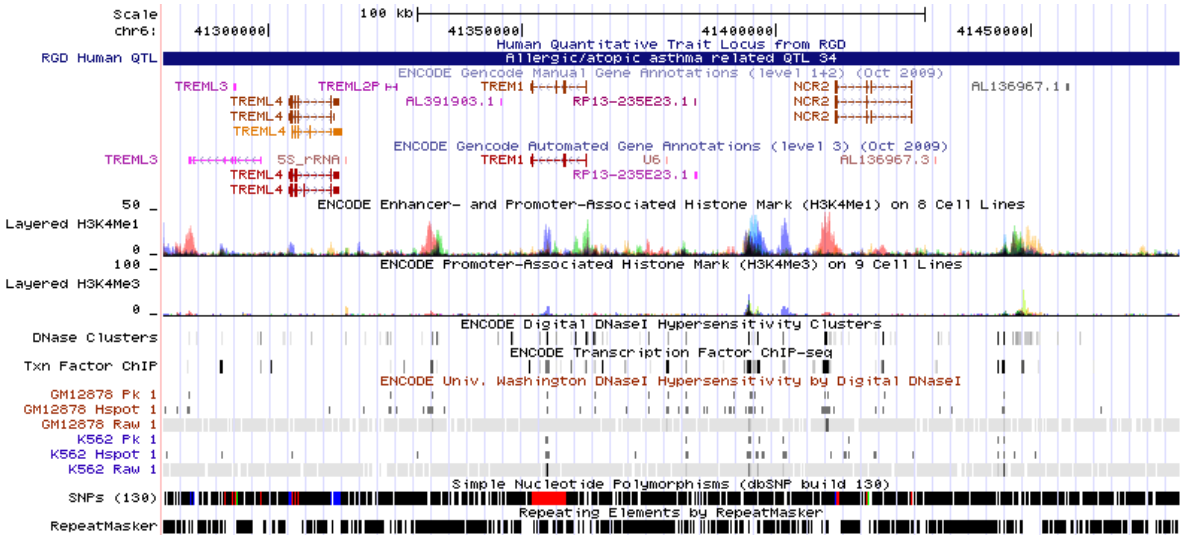


Figure A.35 Visualization of the 19p13.3 locus (chr19:6,787,736-6,987,736; flanking 100Kb of rs3826782, which was prioritized for moderate CP) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

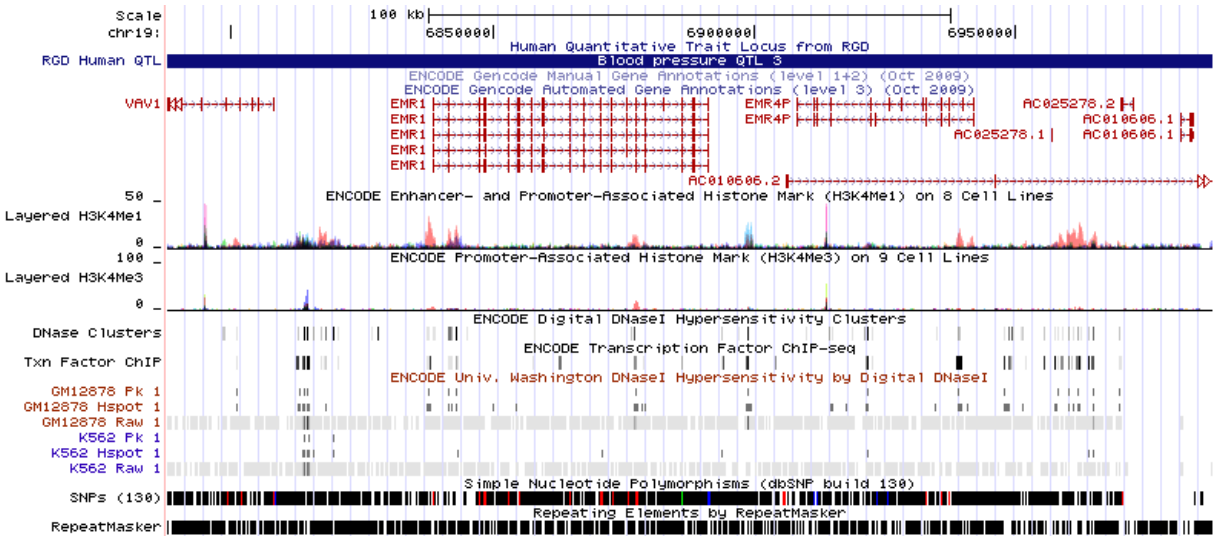


Figure A.36 Visualization of the 10p15locus (chr10:10,238,329-10,438,329; flanking 100Kb of rs12260727, which was prioritized for moderate CP) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

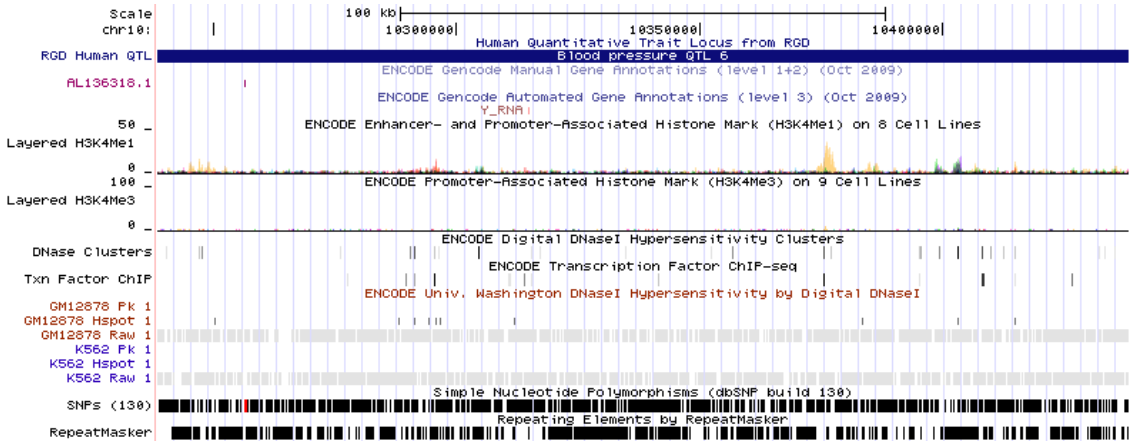


Figure A.37 Visualization of the 14q21 locus (chr14:51,179,379-51,379,379; flanking 100Kb of rs12883458, which was prioritized for severe CP) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

Figure A.38 Visualization of the 7p15 locus (chr7:24,278,040-24,478,040; flanking 100Kb of rs2521634, which was prioritized for severe CP) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

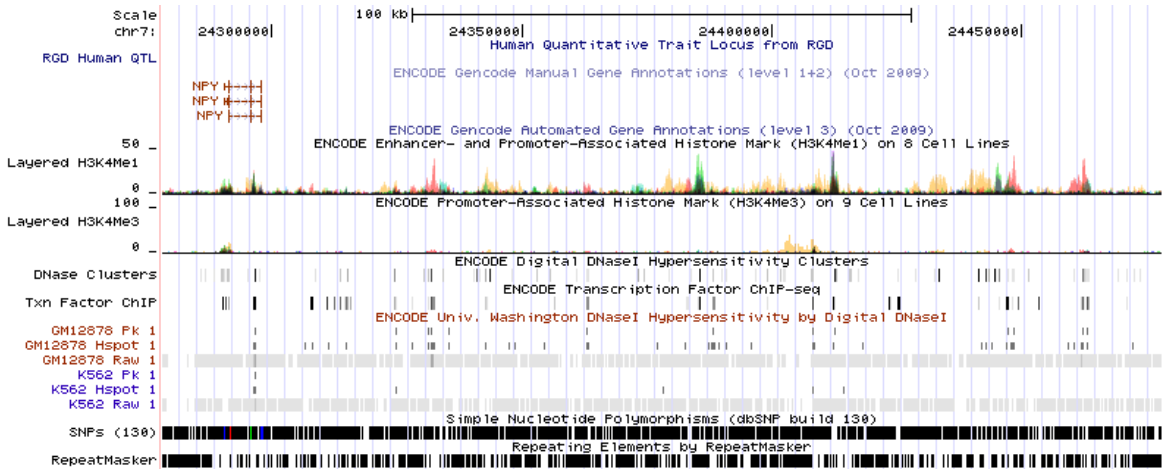


Figure A.39 Visualization of the 3p21 locus (chr3:55,290,886-55,490,886; flanking 100Kb of rs11925054, which was prioritized for severe CP) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

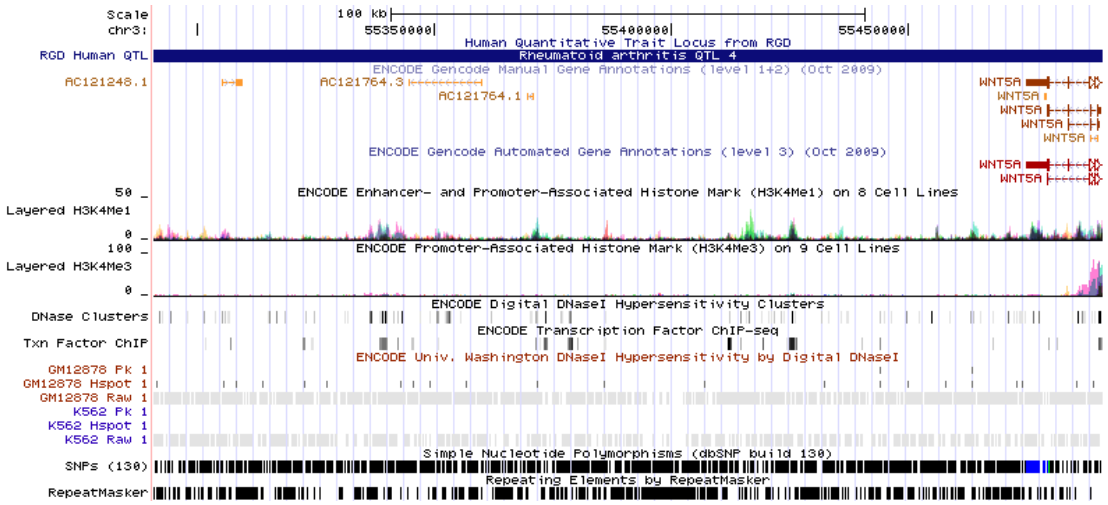


Figure A.40 Visualization of the 6q15 locus (chr6:85,166,965-85,366,965; flanking 100Kb of rs17792917, which was prioritized for the "extent" of attachment loss trait) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

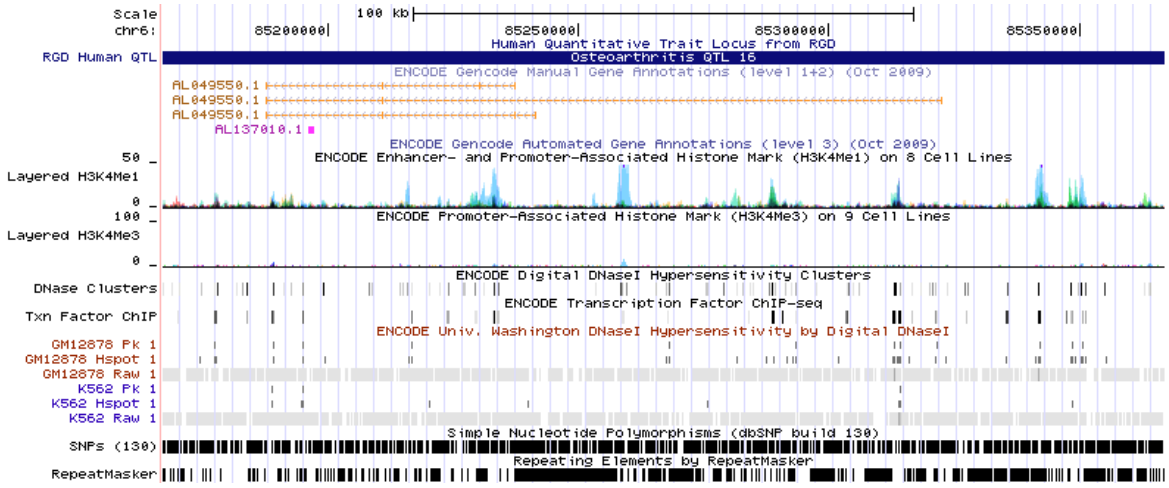


Figure A.41 Visualization of the 11q24 locus (chr11:127,557,071-127,757,071; flanking 100Kb of rs10790919, which was prioritized for the "extent" of attachment loss trait) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

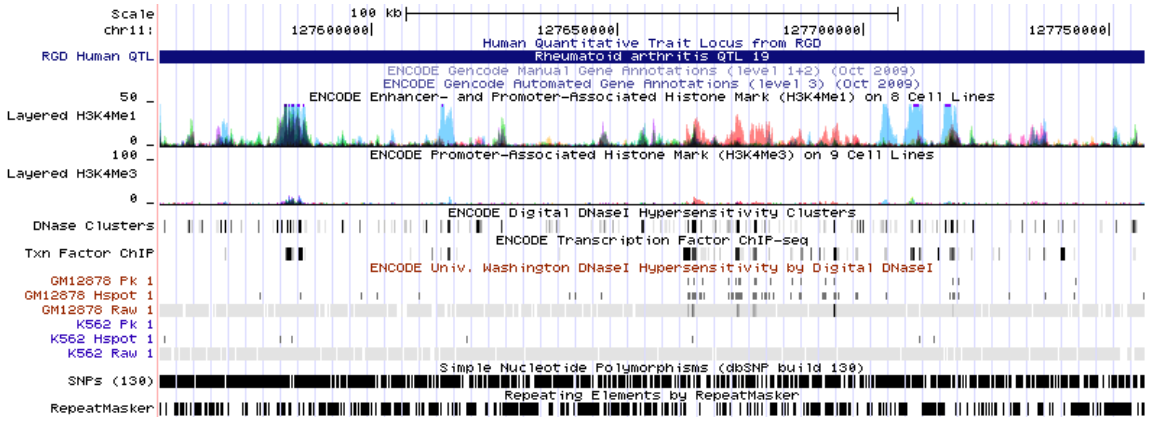


Figure A.42 Visualization of the 11q22 locus (chr11:103,317,781-103,517,781; flanking 100Kb of rs7120142, which was prioritized for the "extent" of attachment loss trait) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.

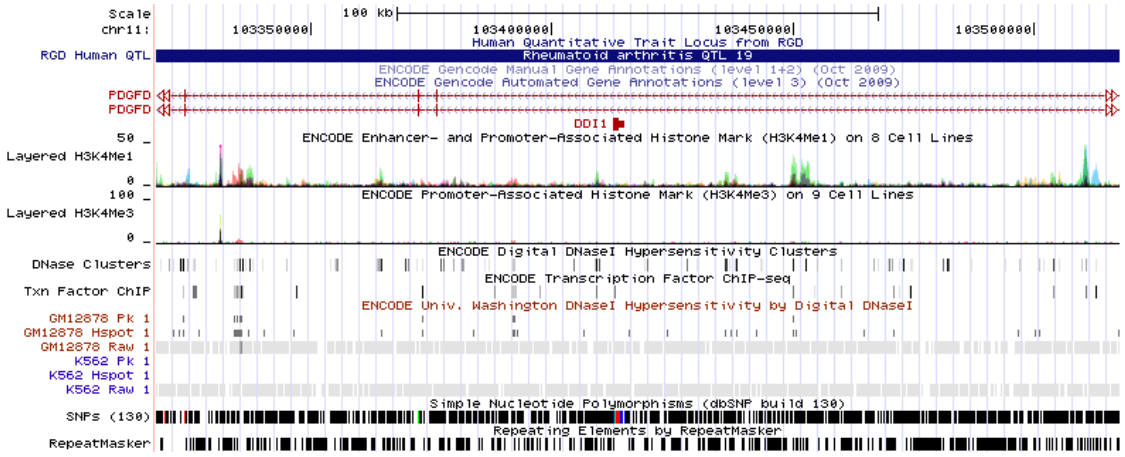
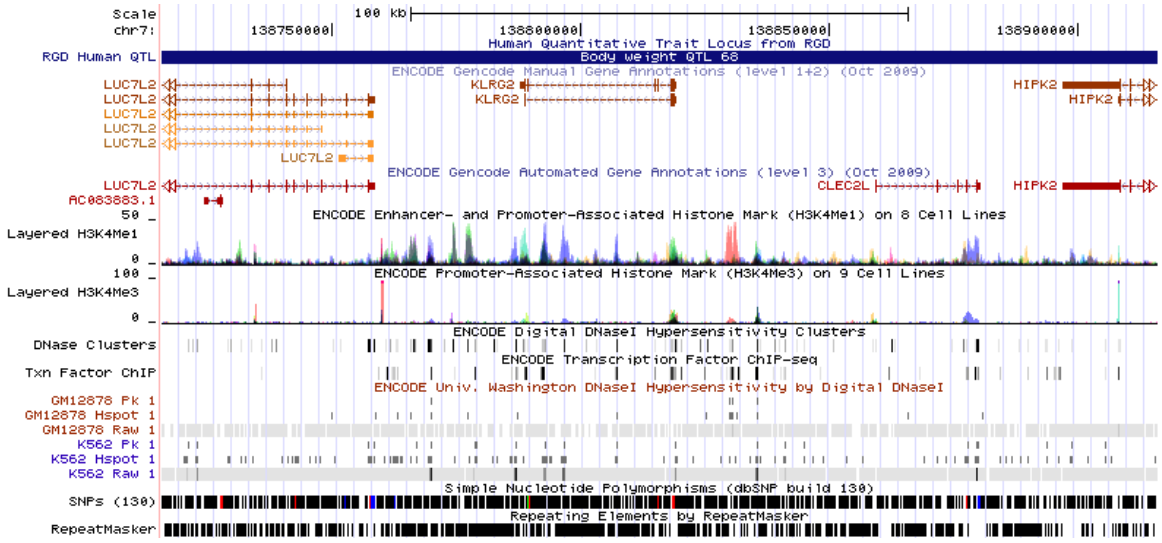


Figure A.43 Visualization of the 7q34 locus (chr7:138,715,913-138,915,913; flanking 100Kb of rs10500130, which was prioritized for the “extent” of attachment loss trait) using the ENCODE/USCF browser and displaying enhancer/promoter associated regions and DNaseI hypersensitivity clusters.



REFERENCES

Chapter 1

1. Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. Bull World Health Organ 2005;83:661-9.

2. Petersen PE, Kandelman D, Arpin S, Ogawa H. Global oral health of older people--call for public health action. Community Dent Health 2010;27(4 Suppl 2):257-67.

3. Avila M, Ojcius DM, Yilmaz O. The oral microbiota: living with a permanent guest. DNA Cell Biol 2009;28:405-11.

4. Handelsman J. Microbial symbiosis: in sickness and in health. DNA Cell Biol 2009;28:359-60.

5. Darveau RP. The oral microbial consortium's interaction with the periodontal innate defense system. DNA Cell Biol 2009;28:389-95.

6. Armitage GC, Robertson PB. The biology, prevention, diagnosis and treatment of periodontal diseases: scientific advances in the United States. J Am Dent Assoc 2009;140 Suppl 1:36S-43S.

7. Hendler A, Mulli TK, Hughes FJ, et al. Involvement of autoimmunity in the pathogenesis of aggressive periodontitis. J Dent Res 2010;89:1389-94.

8. Loesche WJ. Clinical and microbiological aspects of chemotherapeutic agents used according to the specific plaque hypothesis. J Dent Res 1979;58:2404-12.

9. Socransky SS. Relationship of bacteria to the etiology of periodontal disease. J Dent Res 1970;49:203-22.

10. Socransky SS, Haffajee AD. The bacterial etiology of destructive periodontal disease: current concepts. J Periodontol 1992;63(4 Suppl):322-31.

11. Loesche WJ, Grossman NS. Periodontal disease as a specific, albeit chronic, infection: diagnosis and treatment. Clin Microbiol Rev 2001;14:727-52.

12. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL Jr. Microbial complexes in subgingival plaque. J Clin Periodontol 1998;25:134-44.

13. Haffajee AD, Socransky SS. Microbial etiological agents of destructive periodontal diseases. Periodontol 2000 1994;5:78-111.

14. Van Winkelhoff AJ, Winkel EG. Microbiological diagnostics in periodontics: biological significance and clinical validity. Periodontol 2000 2005;39:40-52.

15. Slots J, Ting M. Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis in human periodontal disease: occurrence and treatment. Periodontol 2000 1999;20:82-121.

16. Nakano K, Miyamoto E, Tamura K, et al. Distribution of 10 periodontal bacterial species in children and adolescents over a 7-year period. Oral Dis 2008;14:658-64.

17. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

18. Han YW, Shi W, Huang GT, et al. Interactions between periodontal bacteria and human oral epithelial cells: Fusobacterium nucleatum adheres to and invades epithelial cells. Infect Immun 2000;68:3140-6.

19. Sekot G, Posch G, Messner P, et al. Potential of the Tannerella forsythia S-layer to delay the immune response. J Dent Res 2011;90:109-14.

20. Kebschull M, Demmer RT, Papapanou PN. "Gum bug, leave my heart alone!"--epidemiologic and mechanistic evidence linking periodontal infections and atherosclerosis. J Dent Res 2010;89:879-902.

21. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

22. Noble JM, Borrell LN, Papapanou PN, Elkind MS, Scarmeas N, Wright CB. Periodontitis is associated with cognitive impairment among older adults: analysis of NHANES-III. J Neurol Neurosurg Psychiatry 2009;80:1206-11.

23. Beck JD, Eke P, Heiss G, et al. Periodontal disease and coronary heart disease: a reappraisal of the exposure. Circulation 2005;112:19-24.

24. Madianos PN, Lieff S, Murtha AP, et al. Maternal periodontitis and prematurity. Part II: Maternal infection and fetal exposure. Ann Periodontol 2001;6:175-82.

25. Baker PJ, Wilson ME. Opsonic IgG antibody against Actinobacillus actinomycetemcomitans in localized juvenile periodontitis. Oral Microbiol Immunol 1989;4:98-105.

26. Ximénez-Fyvie LA, Haffajee AD, Socransky SS. Comparison of the microbiota of supra- and subgingival plaque in health and periodontitis. J Clin Periodontol 2000;27:648-57.

27. Haffajee AD, Cugini MA, Dibart S, Smith C, Kent RL Jr, Socransky SS. The effect of SRP on the clinical and microbiological parameters of periodontal diseases. J Clin Periodontol. 1997;24:324-34.

28. Beck JD, Slade G, Offenbacher S. Oral disease, cardiovascular disease and systemic inflammation. Periodontol 2000 2000;23:110-20.

29. Bizzarro S, Nicu EA, van der Velden U, Laine ML, Loos BG. Association of serum immunoglobulin G (IgG) levels against two periodontal pathogens and prothrombotic state: a clinical pilot study. Thromb J 2010;8:16.

30. Nakano K, Wada K, Nomura R, et al. Characterization of aortic aneurysms in cardiovascular disease patients harboring Porphyromonas gingivalis. Oral Dis 2011;17:370-8.

31. Pizzo G, Guiglia R, Lo Russo L, Campisi G. Dentistry and internal medicine: from the focal infection theory to the periodontal medicine concept. Eur J Intern Med 2010;21:496-502.

32. Stamm JW. Periodontal diseases and human health: new directions in periodontal medicine. Ann Periodontol 1998;3:1-2.

33. Bobetsis YA, Barros SP, Lin DM, et al. Bacterial infection promotes DNA hypermethylation. J Dent Res 2007;86:169-74.

34. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol 2010;28:1057-68.

35. Offenbacher S, Barros SP, Beck JD. Rethinking periodontal inflammation. J Periodontol 2008;79(8 Suppl):1577-84.

36. Iacopino AM. Periodontitis and diabetes interrelationships: role of inflammation. Ann Periodontol 2001;6:125-37.

37. Grossi SG, Genco RJ. Periodontal disease and diabetes mellitus: a two-way relationship. Ann Periodontol 1998;3:51-61.

38. Bascones-Martinez A, Matesanz-Perez P, Escribano-Bermejo M, González-Moles MÁ, Bascones-Ilundain J, Meurman JH. Periodontal disease and diabetes-Review of the Literature. Med Oral Patol Oral Cir Bucal 2011;16:e722-9.

39. Ide R, Hoshuyama T, Wilson D, Takahashi K, Higashi T. Periodontal disease and incident diabetes: a seven-year study. J Dent Res 2011;90:41-6.

40. Gurav A, Jadhav V. Periodontitis and risk of diabetes mellitus. J Diabetes 2011;3:21-8.

41. Mealey BL, Oates TW; American Academy of Periodontology. Diabetes mellitus and periodontal diseases. J Periodontol 2006;77:1289-303.

42. Koromantzos PA, Makrilakis K, Dereka X, Katsilambros N, Vrotsos IA, Madianos PN. A randomized, controlled trial on the effect of non-surgical periodontal therapy

in patients with type 2 diabetes. Part I: effect on periodontal status and glycaemic control. J Clin Periodontol 2011;38:142-7.

43. Soskolne WA. Epidemiological and clinical aspects of periodontal diseases in diabetics. Ann Periodontol 1998;3:3-12.

44. Borrell LN, Papapanou PN. Analytical epidemiology of periodontitis. J Clin Periodontol 2005;32 Suppl 6:132-58.

45. Pyrc K, Strzyz P, Milewska A, Golda A, Schildgen O, Potempa J. Porphyromonas gingivalis enzymes enhance infection with human metapneumovirus in vitro. J Gen Virol 2011;92(Pt 10):2324-32.

46. Slots J. Herpesvirus periodontitis: infection beyond biofilm. J Calif Dent Assoc 2011;39:393-9.

47. Thomasini RL, Bonon SH, Durante P, Costa SC. Correlation of cytomegalovirus and human herpesvirus 7 with CD3(+) and CD3(+) CD4(+) cells in chronic periodontitis patients. J Periodontal Res 2011 doi: 10.1111/j.1600-0765.2011.01413.x.

48. Slots J. Human viruses in periodontitis. Periodontol 2000 2010;53:89-110.

49. Contreras A, Umeda M, Chen C, Bakker I, Morrison JL, Slots J. Relationship between herpesviruses and adult periodontitis and periodontopathic bacteria. J Periodontol 1999;70:478-84.

50. Boggess KA, Espinola JA, Moss K, Beck J, Offenbacher S, Camargo CA Jr. Vitamin D status and periodontal disease among pregnant women. J Periodontol 2011;82:195-200.

51. Eberhard J, Drosos Z, Tiemann M, Jepsen S, Schröder JM. Immunolocalization of lactoferrin in healthy and inflamed gingival tissues. J Periodontol 2006;77:472-8.

52. Tenovuo J. Clinical applications of antimicrobial host proteins lactoperoxidase, lysozyme and lactoferrin in xerostomia: efficacy and safety. Oral Dis 2002;8:23-9.

53. Wakabayashi H, Kondo I, Kobayashi T, et al. Periodontitis, periodontopathic bacteria and lactoferrin. Biometals 2010;23:419-24.

54. Löe H. Periodontal disease. The sixth complication of diabetes mellitus. Diabetes Care 1993;16:329-34.

55. Beck JD, Offenbacher S. Systemic effects of periodontitis: epidemiology of periodontal disease and cardiovascular disease. J Periodontol 2005;76(11 Suppl):2089-100.

56. Beck J, Garcia R, Heiss G, Vokonas PS, Offenbacher S. Periodontal disease and cardiovascular disease. J Periodontol 1996;67(10 Suppl):1123-37.

57. Divaris K, Olshan AF, Smith J, et al. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. Cancer Causes Control 2010;21:567-75.

58. Meyer MS, Joshipura K, Giovannucci E, Michaud DS. A review of the relationship between tooth loss, periodontal disease, and cancer. Cancer Causes Control 2008;19:895-907.

59. Michaud DS, Liu Y, Meyer M, Giovannucci E, Joshipura K. Periodontal disease, tooth loss, and cancer risk in male health professionals: a prospective cohort study. Lancet Oncol 2008;9:550-8.

60. Michaud DS, Joshipura K, Giovannucci E, Fuchs CS. A prospective study of periodontal disease and pancreatic cancer in US male health professionals. J Natl Cancer Inst 2007;99:171-5.

61. Armitage GC. Development of a classification system for periodontal diseases and conditions. Ann Periodontol 1999;4:1-6.

62. Offenbacher S, Barros S, Mendoza L, et al. Changes in gingival crevicular fluid inflammatory mediator levels during the induction and resolution of experimental gingivitis in humans. J Clin Periodontol 2010;37:324-33.

63. Picolos DK, Lerche-Sehm J, Abron A, Fine JB, Papapanou PN. Infection patterns in chronic and aggressive periodontitis. J Clin Periodontol 2005;32:1055-61.

64. Armitage GC, Cullinan MP. Comparison of the clinical features of chronic and aggressive periodontitis. Periodontol 2000 2010;53:12-27.

65. Armitage GC. Classifying periodontal diseases--a long-standing dilemma. Periodontol 2000 2002;30:9-23.

66. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78(7 Suppl):1387-99.

67. Armitage GC. Periodontal diagnoses and classification of periodontal diseases. Periodontol 2000 2004;34:9-21.

68. Williams RC. Periodontal disease. N Engl J Med 1990;322:373-82.

69. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.

70. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

71. Tonetti MS, Claffey N; European Workshop in Periodontology group C. Advances in the progression of periodontitis and proposal of definitions of a periodontitis case and disease progression for use in risk factor research. Group C consensus report of the 5th European Workshop in Periodontology. J Clin Periodontol 2005;32 Suppl 6:210-3.

72. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

73. Papapanou PN. Epidemiology of periodontal diseases: an update. J Int Acad Periodontol 1999;1:110-6.

74. Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA. Accuracy of NHANES Periodontal Examination Protocols. J Dent Res 2010;89:1208-13.

75. Albandar JM, Rams TE. Global epidemiology of periodontal diseases: an overview. Periodontol 2000 2002;29:7-10.

76. Hickman MA, Boggess KA, Moss KL, Beck JD, Offenbacher S. Maternal periodontal disease is associated with oxidative stress during pregnancy. Am J Perinatol 2011;28:247-52.

77. Inaba H, Amano A. Roles of oral bacteria in cardiovascular diseases—from molecular mechanisms to clinical cases: Implication of periodontal diseases in development of systemic diseases. J Pharmacol Sci 2010;113:103-9.

78. Noack B, Genco RJ, Trevisan M, Grossi S, Zambon JJ, De Nardin E. Periodontal infections contribute to elevated systemic C-reactive protein level. J Periodontol 2001;72:1221-7.

79. Scannapieco FA, Dasanayake AP, Chhun N. "Does periodontal therapy reduce the risk for systemic diseases?". Dent Clin North Am 2010;54:163-81.

80. Araújo AC, Gusmão ES, Batista JE, Cimões R. Impact of periodontal disease on quality of life. Quintessence Int 2010;41:e111-8.

81. Bernabé E, Marcenes W. Periodontal disease and quality of life in British adults. J Clin Periodontol 2010;37:968-72.

82. Ng SK, Leung WK. Oral health-related quality of life and periodontal status. Community Dent Oral Epidemiol 2006;34:114-22.

83. Needleman I, McGrath C, Floyd P, Biddle A. Impact of oral health on the life quality of periodontal patients. J Clin Periodontol 2004;31:454-7.

84. Divaris K, Ntounis A, Marinis A, Polyzois G, Polychronopoulou A. Loss of natural dentition: multi-level effects among a geriatric population. Gerodontology 2010 doi: 10.1111/j.1741-2358.2010.00440.x.

85. Steele JG, Sanders AE, Slade GD, et al. How do age and tooth loss affect oral health impacts and quality of life? A study comparing two national samples. Community Dent Oral Epidemiol 2004;32:107-14.

86. Sanders AE, Slade GD, Lim S, Reisine ST. Impact of oral disease on quality of life in the US and Australian populations. Community Dent Oral Epidemiol 2009;37:171-81.

87. Salvi GE, Lawrence HP, Offenbacher S, Beck JD. Influence of risk factors on the pathogenesis of periodontitis. Periodontol 2000 1997;14:173-201.

88. Axelsson P, Lindhe J. Effect of controlled oral hygiene procedures on caries and periodontal disease in adults. J Clin Periodontol 1978;5:133-51.

89. Tomar SL, Asma S. Smoking-attributable periodontitis in the United States: findings from NHANES III. National Health and Nutrition Examination Survey. J Periodontol 2000;71:743-51.

90. Ong G. Periodontal disease and tooth loss. Int Dent J 1998;48(3 Suppl 1):233-8.

91. Holm G. Smoking as an additional risk for tooth loss. J Periodontol 1994;65:996-1001.

92. Van Dyke TE, Sheilesh D. Risk factors for periodontitis. J Int Acad Periodontol 2005;7:3-7.

93. Bergström J. Tobacco smoking and chronic destructive periodontal disease. Odontology 2004;92:1-8.

94. Haber J, Wattles J, Crowley M, Mandell R, Joshipura K, Kent RL. Evidence for cigarette smoking as a major risk factor for periodontitis. J Periodontol 1993;64:16-23.

95. Schatzle M, Loe H, Ramseier CA, et al. Clinical course of chronic periodontitis: effect of lifelong light smoking (20 years) on loss of attachment and teeth. Journal of Investigative and Clinical Dentistry 2010;1:8-15.

96. Haffajee AD, Socransky SS. Relationship of cigarette smoking to the subgingival microbiota. J Clin Periodontol 2001;28:377-88.

97. Buduneli N, Larsson L, Biyikoglu B, Renaud DE, Bagaitkar J, Scott DA. Fatty acid profiles in smokers with chronic periodontitis. J Dent Res 2011;90:47-52.

98. Oliver RC, Tervonen T. Periodontitis and tooth loss: comparing diabetics with the general population. J Am Dent Assoc 1993;124:71-6.

99. Tervonen T, Oliver RC. Long-term control of diabetes mellitus and periodontitis. J Clin Periodontol 1993;20:431-5.

100. Pradhan AD, Manson JE, Rifai N, Buring JE, Ridker PM. C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. JAMA 2001;286:327-34.

101. Dye BA. The Relationship Between Periodontitis and Alcohol use is not Clear. J Evid Based Dent Pract 2010;10:225-7.

102. Pitiphat W, Merchant AT, Rimm EB, Joshipura KJ. Alcohol consumption increases periodontitis risk. J Dent Res 2003;82:509-13.

103. Nishida N, Tanaka M, Sekine S, et al. Association of ALDH2 genotypes with periodontitis progression. J Dent Res 2010;89:138-42.

104. Herrera D, Roldán S, Santacruz I, Santos S, Masdevall M, Sanz M. Differences in antimicrobial activity of four commercial 0.12% chlorhexidine mouthrinse formulations: an in vitro contact test and salivary bacterial counts study. J Clin Periodontol 2003;30:307-14.

105. Kongstad J, Hvidtfeldt UA, Grønbaek M, Jontell M, Stoltze K, Holmstrup P. Amount and type of alcohol and periodontitis in the Copenhagen City Heart Study. J Clin Periodontol 2008;35:1032-9.

106. Merchant AT, Pitiphat W, Franz M, Joshipura KJ. Whole-grain and fiber intakes and periodontitis risk in men. Am J Clin Nutr 2006;83:1395-400.

107. Al-Zahrani MS. Increased intake of dairy products is related to lower periodontitis prevalence. J Periodontol 2006;77:289-94.

108. Nowjack-Raymer RE, Sheiham A. Association of edentulism and diet and nutrition in US adults. J Dent Res 2003;82:123-6.

109. Hutton B, Feine J, Morais J. Is there an association between edentulism and nutritional state? J Can Dent Assoc 2002;68:182-7.

110. Genco RJ, Ho AW, Grossi SG, Dunford RG, Tedesco LA. Relationship of stress, distress and inadequate coping behaviors to periodontal disease. J Periodontol 1999;70:711-23.

111. Genco RJ, Ho AW, Kopman J, Grossi SG, Dunford RG, Tedesco LA. Models to evaluate the role of stress in periodontal disease. Ann Periodontol 1998;3:288-302.

112. Moss ME, Beck JD, Kaplan BH, et al. Exploratory case-control analysis of psychosocial factors and adult periodontitis. J Periodontol 1996;67:1060-9.

113. Beaty TH, Boughman JA, Yang P, Astemborski JA, Suzuki JB. Genetic analysis of juvenile periodontitis in families ascertained through an affected proband. Am J Hum Genet 1987;40:443-52.

114. Long JC, Nance WE, Waring P, Burmeister JA, Ranney RR. Early onset periodontitis: a comparison and evaluation of two proposed modes of inheritance. Genet Epidemiol 1987;4:13-24.

115. Page RC, Vandesteen GE, Ebersole JL, Williams BL, Dixon IL, Altman LC. Clinical and laboratory studies of a family with a high prevalence of juvenile periodontitis. J Periodontol 1985;56:602-10.

116. Saxén L, Nevanlinna HR. Autosomal recessive inheritance of juvenile periodontitis: test of a hypothesis. Clin Genet 1984;25:332-5.

117. Saxén L. Heredity of juvenile periodontitis. J Clin Periodontol 1980;7:276-88.

118. Chung CS, Kau MC, Chung SS, Rao DC. A genetic and epidemiologic study of periodontal disease in Hawaii. II. Genetic and environmental influence. Am J Hum Genet 1977;29:76-82.

119. Michalowicz BS, Diehl SR, Gunsolley JC, et al. Evidence of a substantial genetic basis for risk of adult periodontitis. J Periodontol 2000;71:1699-707.

120. Nares S. The genetic relationship to periodontal disease. Periodontol 2000 2003;32:36-49.

121. Ciancio SG, Hazen SP, Cunat JJ. Periodontal observations in twins. J Periodontal Res 1969;4:42-5.

122. Michalowicz BS, Aeppli D, Virag JG, et al. Periodontal findings in adult twins. J Periodontol 1991;62:293-9.

123. Michalowicz BS. Genetic and inheritance considerations in periodontal disease. Curr Opin Periodontol 1993:11-7.

124. Corey LA, Nance WE, Hofstede P, Schenkein HA. Self-reported periodontal disease in a Virginia twin population. J Periodontol 1993;64:1205-8.

125. Torres de Heens GL, Loos BG, van der Velden U. Monozygotic twins are discordant for chronic periodontitis: clinical and bacteriological findings. J Clin Periodontol 2010;37:120-8.

126. Stabholz A, Soskolne WA, Shapira L. Genetic and environmental risk factors for chronic periodontitis and aggressive periodontitis. Periodontol 2000 2010;53:138-53.

127. Grigoriadou ME, Koutayas SO, Madianos PN, Strub JR. Interleukin-1 as a genetic marker for periodontitis: review of the literature. Quintessence Int 2010 ;41:517-25.

128. Soga Y, Nishimura F, Ohyama H, Maeda H, Takashiba S, Murayama Y. Tumor necrosis factor-alpha gene (TNF-alpha) -1031/-863, -857 single-nucleotide polymorphisms (SNPs) are associated with severe adult periodontitis in Japanese. J Clin Periodontol 2003;30:524-31.

129. Repeke CE, Trombone AP, Ferreira SB Jr, et al. Strong and persistent microbial and inflammatory stimuli overcome the genetic predisposition to higher matrix metalloproteinase-1 (MMP-1) expression: a mechanistic explanation for the lack of association of MMP1-1607 single-nucleotide polymorphism genotypes with MMP-1 expression in chronic periodontitis lesions. J Clin Periodontol 2009;36:726-38.

130. Zhang X, Long Y, Li HY, Li XJ, Ding Y. [Estrogen receptors gene polymorphisms in patients with chronic periodontitis]. Hua Xi Kou Qiang Yi Xue Za Zhi 2010;28:139-44.

131. Laine ML, Loos BG, Crielaard W. Gene polymorphisms in chronic periodontitis. Int J Dent 2010;2010:324719.

132. Anovazzi G, Kim YJ, Viana AC, et al. Polymorphisms and haplotypes in the interleukin-4 gene are associated with chronic periodontitis in a Brazilian population. J Periodontol 2010;81:392-402.

133. Carvalho FM, Tinoco EM, Deeley K, et al. FAM5C contributes to aggressive periodontitis. PLoS One 2010;5(4):e10053.

134. Wohlfahrt JC, Wu T, Hodges JS, Hinrichs JE, Michalowicz BS. No association between selected candidate gene polymorphisms and severe chronic periodontitis. J Periodontol 2006;77:426-36.

135. Fiebig A, Jepsen S, Loos BG, et al. Polymorphisms in the interleukin-1 (IL1) gene cluster are not associated with aggressive periodontitis in a large Caucasian population. Genomics 2008;92:309-15.

136. Abu-Saleh T. Interleukin-1 composite polymorphism as a risk factor for chronic periodontitis: a review. SADJ 2010;65:160-2, 164-6.

137. Nikolopoulos GK, Dimou NL, Hamodrakas SJ, Bagos PG. Cytokine gene polymorphisms in periodontal disease: a meta-analysis of 53 studies including 4178 cases and 4590 controls. J Clin Periodontol 2008;35:754-67.

138. Dimou NL, Nikolopoulos GK, Hamodrakas SJ, Bagos PG. Fcgamma receptor polymorphisms and their association with periodontal disease: a meta-analysis. J Clin Periodontol 2010;37:255-65.

139. Nibali L, Tonetti MS, Ready D, et al. Interleukin-6 polymorphisms are associated with pathogenic bacteria in subjects with periodontitis. J Periodontol 2008;79:677-83.

140. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516-7.

141. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.

142. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. Science 2002;298:2345-9.

143. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

144. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19:553-62.

145. Schaefer AS, Richter GM, Groessner-Schreiber B, et al. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet 2009;5:e1000378.

146. Ernst FD, Uhr K, Teumer A, et al. Replication of the association of chromosomal region 9p21.3 with generalized aggressive periodontitis (gAgP) using an independent case-control cohort. BMC Med Genet 2010;11:119.

147. Oliver RC, Brown LJ, Löe H. Periodontal diseases in the United States population. J Periodontol 1998;69:269-78.

148. Löe H, Brown LJ. Early onset periodontitis in the United States of America. J Periodontol 1991;62:608-16.

Chapter 2

149. Lander ES, Linton LM, Birren B, et al.; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860-921.

150. Lander ES. The new genomics: global views of biology. Science 1996;274:536-9.

151. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516-7.

152. Barrett JC, Cardon LR. Evaluating coverage of genome-wide association studies. Nat Genet 2006;38:659-62.

153. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.

154. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. Science 2002;298:2345-9.

155. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

156. Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies. Accessed [November 7 2011]

157. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19:553-62.

158. Schaefer AS, Richter GM, Groessner-Schreiber B, et al Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet 2009;5:e1000378.

159. Ernst FD, Uhr K, Teumer A, et al. Replication of the association of chromosomal region 9p21.3 with generalized aggressive periodontitis (gAgP) using an independent case-control cohort. BMC Med Genet 2010;11:119.

160. Vieira AR, Marazita ML, Goldstein-McHenry T. Genome-wide scan finds suggestive caries loci. J Dent Res 2008;87:435-9.

161. Shaffer JR, Wang X, Feingold E, et al. Genome-wide association Scan for childhood Caries Implicates Novel Genes. J Dent Res 2011 In Press

162. Teles RP, Haffajee AD, Socransky SS. Microbiological goals of periodontal therapy. Periodontol 2000 2006;42:180-218.

163. Eley BM, Cox SW. Proteolytic and hydrolytic enzymes from putative periodontal pathogens: characterization, molecular genetics, effects on host defenses and tissues and detection in gingival crevice fluid. Periodontol 2000 2003;31:105-24.

164. Colombo AP, Boches SK, Cotton SL, et al. Comparisons of subgingival microbial profiles of refractory periodontitis, severe periodontitis, and periodontal health using the human oral microbe identification microarray. J Periodontol 2009 ;80:1421-32.

165. Buduneli N, Larsson L, Biyikoglu B, Renaud DE, Bagaitkar J, Scott DA. Fatty Acid Profiles in Smokers with Chronic Periodontitis. J Dent Res 2011;90:47-52.

166. Kebschull M, Demmer RT, Papapanou PN. "Gum bug, leave my heart alone!"--epidemiologic and mechanistic evidence linking periodontal infections and atherosclerosis. J Dent Res 2010;89:879-902.

167. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

168. Nibali L, Donos N, Farrell S, et al. Association Between IL6 -174 Polymorphism and A.actinomycetemcomitans in Chronic Periodontitis. J Periodontol 2010;81:1814-9.

169. Reddi D, Brown SJ, Belibasakis GN. Porphyromonas gingivalis induces RANKL in bone marrow stromal cells: Involvement of the p38 MAPK. Microb Pathog 2011;51:415-20.

Chapter 3

170. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

171. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

172. White AD, Folsom AR, Chambless LE, et al. Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. J Clin Epidemiol 1996;49:223-33.

173. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78:1387-99.

174. Mombelli A. Clinical parameters: biological validity and clinical utility. Periodontol 2000 2005;39:30-9.

175. Alves Rde V, Machion L, Andia DC, Casati MZ, Sallum AW, Sallum EA. Reproducibility of clinical attachment level and probing depth of a manual probe and a computerized electronic probe. J Int Acad Periodontol 2005;7:27-30.

176. Armitage GC. Manual periodontal probing in supportive periodontal treatment. Periodontol 2000 1996;12:33-9.

177. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.

178. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007 ;78:1911-25.

179. Gunaratnam M, Smith GL, Socransky SS, Smith CM, Haffajee AD. Enumeration of subgingival species on primary isolation plates using colony lifts. Oral Microbiol Immunol 1992;7:14-8.

180. Socransky SS, Smith C, Martin L, Paster BJ, Dewhirst FE, Levin AE. "Checkerboard" DNA-DNA hybridization. Biotechniques 1994;17:788-92.

181. Van Winkelhoff AJ, Winkel EG. Microbiological diagnostics in periodontics: biological significance and clinical validity. Periodontol 2000 2005;39:40-52.

182. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904-9.

183. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet 2006;2:e190.

184. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.

185. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.

186. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

187. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005;6:95-108.

188. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. Am J Hum Genet 2010;86:6-22.

189. Feng T, Zhu X. Genome-wide searching of rare genetic variants in WTCCC data. Hum Genet 2010;128:269-80.

190. Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet 2010;11:773-85.

191. Beck JD, Eke P, Heiss G, et al. Periodontal disease and coronary heart disease: a reappraisal of the exposure. Circulation 2005;112:19-24.

192. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. Cancer Epidemiol Biomarkers Prev 2002;11:513-20.

193. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

194. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24:2938-9.

195. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440-5.

196. Shi G, Boerwinkle E, Morrison AC, Gu CC, Chakravarti A, Rao DC. Mining gold dust under the genome wide significance level: a two-stage approach to analysis of GWAS. Genet Epidemiol 2011;35:111-8.

197. Greenland S, Rothman KJ. Introduction to stratified analysis. In: Rothman KJ, Greenland S, Lash TL. Modern epidemiology. New York: Lippincott, Williams and Wilkins, 2008:258-82.

198. Vansteelandt S, Goetgeluk S, Lutz S, et al. On the adjustment for covariates in genetic association analysis: a novel, simple principle to infer direct causal effects. Genet Epidemiol 2009;33:394-405.

199. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. Am J Epidemiol 1993;138:923-36.

200. Ge D, Zhang K, Need AC, et al. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640-643.

201. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010;6:e1000888.

202. Gamazon ER, Zhang W, Konkashbaev A, et al. SCAN: SNP and copy number annotation. Bioinformatics 2010;26:259-62.

203. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336-7.

204. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

205. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9.

206. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol 2008;9 Suppl 1:S4.

207. Knight J. A survey of current software for genetic power calculations. Hum Genomics 2004;1:225-7.

208. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

209. Papapanou PN, Neiderud AM, Papadimitriou A, Sandros J, Dahlén G. "Checkerboard" assessments of periodontal microbiota and serum antibody responses: a case-control study. J Periodontol 2000;71:885-97.

210. Ioannidis JP. Non-replication and inconsistency in the genome-wide association setting. Hum Hered 2007;64:203-13.

211. Liu YJ, Papasian CJ, Liu JF, Hamilton J, Deng HW. Is replication the gold standard for validating genome-wide association findings? PLoS One 2008;3:e4037.

212. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-1.

Chapter 4

213. Berezow AB, Darveau RP. Microbial shift and periodontitis. Periodontol 2000 2011;55:36-47.

214. Petersen PE, Ogawa H. Strengthening the prevention of periodontal disease: the WHO approach. J Periodontol 2005;76:2187-93.

215. Beck JD, Offenbacher S, Williams R, Gibbs P, Garcia R. Periodontitis: a risk factor for coronary heart disease? Ann Periodontol 1998;3:127-41.

216. Xiong X, Buekens P, Fraser WD, Beck J, Offenbacher S. Periodontal disease and adverse pregnancy outcomes: a systematic review. BJOG 2006;113:135-43.

217. Lalla E, Papapanou PN. Diabetes mellitus and periodontitis: a tale of two common interrelated diseases. Nat Rev Endocrinol 2011. doi: 10.1038/nrendo.2011.106.

218. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

219. Saito T, Shimazaki Y, Sakamoto M. Obesity and periodontitis. N Engl J Med 1998;339:482-3.

220. Saxén L. Heredity of juvenile periodontitis. J Clin Periodontol 1980;7:276-88.

221. Michalowicz BS, Aeppli D, Virag JG, et al. Periodontal findings in adult twins. J Periodontol 1991;62:293-9.

222. Zhang J, Sun X, Xiao L, Xie C, Xuan D, Luo G. Gene polymorphisms and periodontitis. Periodontol 2000 2011;56:102-24.

223. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19:553-62.

224. Schaefer AS, Richter GM, Groessner-Schreiber B, et al. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet2009;5:e1000378.

225. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

226. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

227. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78:1387-99.

228. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.

229. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. Am J Epidemiol 1993;138:923-36.

230. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

231. Ge D, Zhang K, Need AC, et al. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640-43.

232. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336-7.

233. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

234. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

235. Jenkins MR, Tsun A, Stinchcombe JC, Griffiths GM. The strength of T cell receptor signal controls the polarization of cytotoxic machinery to the immunological synapse. Immunity 2009;31:621-31.

236. Olson JE, Wang X, Pankratz VS, et al. Centrosome-related genes, genetic variation, and risk of breast cancer. Breast Cancer Res Treat 2011;125:221-8.

237. Zhou Z, Zhu G, Hariri AR, et al. Genetic variation in human NPY expression affects stress response and emotion. Nature 2008;452:997-1001.

238. Kamino M, Kishida M, Kibe T, et al. Wnt-5a signaling is correlated with infiltrative activity in human glioma by inducing cellular migration and MMP-2. Cancer Sci 2011;102:540-8.

239. Cantoni C, Bottino C, Vitale M, et al. NKp44, a triggering receptor involved in tumor cell lysis by activated human natural killer cells, is a novel member of the immunoglobulin superfamily. J Exp Med 1999;189:787-96.

240. Mallory MJ, Jackson J, Weber B, Chi A, Heyd F, Lynch KW. Signal- and development-dependent alternative splicing of LEF1 in T cells is controlled by CELF2. Mol Cell Biol 2011;31:2184-95.

241. Wågsäter D, Zhu C, Björck HM, Eriksson P. Effects of PDGF-C and PDGF-D on monocyte migration and MMP-2 and MMP-9 expression. Atherosclerosis 2009;202:415-23.

242. Jeong MS, Kim EJ, Jang SB. Expression and RNA-binding of human zinc-finger antiviral protein. Biochem Biophys Res Commun 2010;396:696-702.

243. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.

Chapter 4 appendix

244. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904-909.

245. Page RC, Eke PI (2007). Case definitions for use in population-based surveillance of periodontitis. *J Periodontol* 78(7 Suppl):1387-1399.

Chapter 5

246. Avila M, Ojcius DM, Yilmaz O. The oral microbiota: living with a permanent guest. DNA Cell Biol 2009;28:405-11.

247. Keijser BJ, Zaura E, Huse SM, et al. Pyrosequencing analysis of the oral microflora of healthy adults. J Dent Res 2008;87:1016-20.

248. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

249. Darveau RP. Periodontitis: a polymicrobial disruption of host homeostasis. Nat Rev Microbiol 2010;8:481-90.

250. Zhang J, Sun X, Xiao L, Xie C, Xuan D, Luo G. Gene polymorphisms and periodontitis. Periodontol 2000 2011;56:102-24.

251. Nibali L, Donos N, Farrell S, et al. Association between interleukin-6 -174 polymorphism and Aggregatibacter actinomycetemcomitans in chronic periodontitis. J Periodontol 2010;81:1814-9.

252. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

253. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

254. Socransky SS, Smith C, Martin L, Paster BJ, Dewhirst FE, Levin AE. "Checkerboard" DNA-DNA hybridization. Biotechniques 1994;17:788-92.

255. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

256. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

257. Papapanou PN, Neiderud AM, Papadimitriou A, Sandros J, Dahlén G. "Checkerboard" assessments of periodontal microbiota and serum antibody responses: a case-control study. J Periodontol 2000;71:885-97.

258. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

259. Ge D, Zhang K, Need AC, et al. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640-3.

260. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336-7.

261. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

262. Heidecker B, Lamirault G, Kasper EK, et al. The gene expression profile of patients with new-onset heart failure reveals important gender-specific differences. Eur Heart J 2010;31:1188-96.

263. Kurowska-Stolarska M, Hueber A, Stolarski B, McInnes IB. Interleukin-33: a novel mediator with a role in distinct disease pathologies. J Intern Med 2011;269:29-35.

264. Lin WD, Lin SP, Wang CH, Tsai Y, Chen CP, Tsai FJ. RUNX2 mutations in Taiwanese patients with cleidocranial dysplasia. Genet Mol Biol 2011;34:201-4.

265. Tabuchi S, Sakuta T, Oyama T, et al. Runt-related gene 2 is involved in the inhibition of matrix metalloproteinase-13 expression by roxithromycin in human gingival epithelial cell cultures. J Periodontal Res 2009;44:283-8.

266. Cauchi S, Proença C, Choquet H, et al; D.E.S.I.R. Study Group. Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. J Mol Med (Berl) 2008;86:341-8.

267. Veale KJ, Offenhäuser C, Lei N, Stanley AC, Stow JL, Murray RZ. VAMP3 regulates podosome organisation in macrophages and together with Stx4/SNAP23 mediates adhesion, cell spreading and persistent migration. Exp Cell Res 2011;317:1817-29.

268. Bishop KA, Coy HM, Nerenz RD, Meyer MB, Pike JW. Mouse Rankl expression is regulated in T cells by c-Fos through a cluster of distal regulatory enhancers designated the T cell control region. J Biol Chem 2011;286:20880-91.

269. Yao C, Purwanti N, Karabasil MR, et al. Potential down-regulation of salivary gland AQP5 by LPS via cross-coupling of NF-kappaB and p-c-Jun/c-Fos. Am J Pathol 2010;177:724-34.

270. Murata T, Noda C, Saito S, et al. Involvement of Jun dimerization protein 2 (JDP2) in the maintenance of Epstein-Barr virus latency. J Biol Chem 2011;286:22007-16.

271. Chalabi M, Rezaie F, Moghim S, Mogharehabed A, Rezaei M, Mehraban B. Periodontopathic bacteria and herpesviruses in chronic periodontitis. Mol Oral Microbiol 2010;25:236-40.

272. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL Jr. Microbial complexes in subgingival plaque. J Clin Periodontol 1998;25:134-44.

273. Socransky SS, Haffajee AD. Dental biofilms: difficult therapeutic targets. Periodontol 2000 2002;28:12-55.

274. Badger JH, Ng PC, Venter JC (2010). The human genome, microbiomes, and disease. In: Nelson KE (Ed). Metagenomics of the Human Body. Springer, New York, NY, USA. pp1-14.

Chapter 5 appendix

275. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

276. Beck JD, Elter JR, Heiss G, Couper D, Mauriello SM, Offenbacher S. Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study. Arterioscler Thromb Vasc Biol 2001;21:1816-22.

277. Elter JR, Offenbacher S, White RP, Beck JD. Third molars associated with periodontal pathology in older Americans. J Oral Maxillofac Surg 2005;63:179-84.

278. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

279. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78:1387-99.

280. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome wide association studies. Nat Genet 2006;38:904-9.

281. Mooney J, Adonogianaki E, Riggio MP, Takahashi K, Haerian A, Kinane DF. Initial serum antibody titer to Porphyromonas gingivalis influences development of antibody avidity and success of therapy for chronic periodontitis. Infect Immun 1995;63:3411-6.

282. van Winkelhoff AJ, Loos BG, van der Reijden WA, van der Velden U. Porphyromonas gingivalis, Bacteroides forsythus and other putative periodontal pathogens in subjects with and without periodontal destruction. J Clin Periodontol 2002;29:1023-8.

283. Yasunaga S, Grati M, Cohen-Salmon M, et al. A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes DFNB9, a nonsyndromic form of deafness. Nat Genet 1999;21:363-9.

284. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat Genet 2010;42:692-7.

285. Letra A, Menezes R, Govil M, et al. Follow-up association studies of chromosome region 9q and nonsyndromic cleft lip/palate. Am J Med Genet A 2010;152A:1701-10.

286. Iwashita S, Song SY. RasGAPs: a crucial regulator of extracellular stimuli for homeostasis of cellular functions. Mol Biosyst 2008;4:213-22.

287. Yano M, Toyooka S, Tsukuda K, et al. Aberrant promoter methylation of human DAB2 interactive protein (hDAB2IP) gene in lung cancers. Int J Cancer 2005;113:59-66.

288. Janke C, Rogowski K, Wloga D, et al. Tubulin polyglutamylase enzymes are members of the TTL domain protein family. Science 2005;308:1758-62.

289. Theurillat JP, Zürrer-Härdi U, Varga Z, et al. Distinct expression patterns of the immunogenic differentiation antigen NY-BR-1 in normal breast, testis and their malignant counterparts. Int J Cancer 2008;122:1585-91.

290. Jäger D, Filonenko V, Gout I, et al. NY-BR-1 is a differentiation antigen of the mammary gland. Appl Immunohistochem Mol Morphol 2007;15:77-83.

291. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 2008;4:e1000282.

Chapter 6

292. Hunter DJ, Thomas G, Hoover RN, Chanock SJ. Scanning the horizon: what is the future of genome-wide association studies in accelerating discoveries in cancer etiology and prevention? Cancer Causes Control 2007;18:479-84.

293. Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA. Accuracy of NHANES Periodontal Examination Protocols. J Dent Res 2010;89:1208-13.

294. Savage A, Eaton KA, Moles DR, Needleman I. A systematic review of definitions of periodontitis and methods that have been used to identify this disease. J Clin Periodontol 2009;36:458-67.

295. Desvarieux M, Demmer RT, Rundek T, et al. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). Circulation 2005;111:576-82.

296. Beck JD, Slade G, Offenbacher S. Oral disease, cardiovascular disease and systemic inflammation. Periodontol 2000 2000;23:110-20.

297. Bizzarro S, Nicu EA, van der Velden U, Laine ML, Loos BG. Association of serum immunoglobulin G (IgG) levels against two periodontal pathogens and prothrombotic state: a clinical pilot study. Thromb J 2010;8:16.

298. Nakano K, Wada K, Nomura R, et al. Characterization of aortic aneurysms in cardiovascular disease patients harboring Porphyromonas gingivalis. Oral Dis 2011;17:370-8.

299. Launay JM, Del Pino M, Chironi G, et al. Smoking induces long-lasting effects through a monoamine-oxidase epigenetic regulation. PLoS One 2009;4:e7959.

300. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003;33:245-54.

301. Iacopino AM. Periodontitis and diabetes interrelationships: role of inflammation. Ann Periodontol 2001;6:125-37.

302. Lalla E, Papapanou PN. Diabetes mellitus and periodontitis: a tale of two common interrelated diseases. Nat Rev Endocrinol 2011. doi: 10.1038/nrendo.2011.106.

303. Joshi N, Caputo GM, Weitekamp MR, Karchmer AW. Infections in patients with diabetes mellitus. N Engl J Med 1999;341:1906-12.

304. Beck JD, Löe H. Epidemiological principles in studying periodontal diseases. Periodontol 2000 1993;2:34-45.

305. Papapanou PN, Wennström JL, Gröndahl K. Periodontal status in relation to age and tooth type. A cross-sectional radiographic study. J Clin Periodontol 1988;15:469-78.

306. Ong G. Periodontal disease and tooth loss. Int Dent J 1998;48:233-8.

307. Preshaw PM. Definitions of periodontal disease in research. J Clin Periodontol 2009;36:1-2.

308. Nothnagel M, Ellinghaus D, Schreiber S, Krawczak M, Franke A. A comprehensive evaluation of SNP genotype imputation. Hum Genet 2009;125:163-71.

309. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010;11:134.

310. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061-73.

311. Park JH, Wacholder S, Gail MH, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat Genet 2010;42:570-5.

312. Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. Am J Hum Genet 2007;80:605-15.

313. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol 2008;37:120-32.

314. Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 2010;11:446-50.

315. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444:444-54.

316. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.

317. Ritchie MD. Using biological knowledge to uncover the mystery in the search for epistasis in genome-wide association studies. Ann Hum Genet 2011;75:172-82.

318. Alexander RP, Fang G, Rozowsky J, Snyder M, Gerstein MB. Annotating non-coding regions of the genome. Nat Rev Genet 2010;11:559-71.

319. Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223-7.

320. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol 2010;8:e1000294.

321. Panagiotou OA, Evangelou E, Ioannidis JP. Genome-wide significant associations for variants with minor allele frequency of 5% or less--an overview: A HuGE review. Am J Epidemiol 2010;172:869-89.

322. Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet 2010;11:773-85.

323. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. Am J Hum Genet 2010;86:6-22.

324. Luo L, Peng G, Zhu Y, Dong H, Amos CI, Xiong M. Genome-wide gene and pathway analysis. Eur J Hum Genet 2010;18:1045-53.

325. Barros SP, Offenbacher S. Epigenetics: connecting environment and genotype to phenotype and disease. J Dent Res 2009;88:400-8.

326. Feinberg AP. Epigenomics reveals a functional genome anatomy and a new approach to common disease. Nat Biotechnol 2010;28:1049-52.

327. Bobetsis YA, Barros SP, Lin DM, et al. Bacterial infection promotes DNA hypermethylation. J Dent Res 2007;86:169-74.

328. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998;317:465-8.

329. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9:356-69.

330. Drews J. Drug discovery: a historical perspective. Science 2000;287:1960-4.

331. Khoury MJ, Gwinn M, Ioannidis JP. The emergence of translational epidemiology: from scientific discovery to population health impact. Am J Epidemiol 2010;172:517-24.

332. Hiatt RA. Invited commentary: The epicenter of translational science. Am J Epidemiol 2010;172:525-7.

333. VanderWeele TJ. Genetic self knowledge and the future of epidemiologic confounding. Am J Hum Genet 2010;87:168-72.

334. Ioannidis JP. Personalized genetic prediction: too limited, too expensive, or too soon? Ann Intern Med 2009;150:139-41.

335. Khoury MJ, McBride CM, Schully SD, et al; Centers for Disease Control and Prevention. The Scientific Foundation for personal genomics: recommendations from a National Institutes of Health-Centers for Disease Control and Prevention multidisciplinary workshop. Genet Med 2009;11:559-67.

336. Petersen PE, Ogawa H. Strengthening the prevention of periodontal disease: the WHO approach. J Periodontol 2005;76:2187-93.

337. Kallio PJ. Health promotion and behavioral approaches in the prevention of periodontal disease in children and adolescents. Periodontol 2000 2001;26:135-45.

338. Teles RP, Haffajee AD, Socransky SS. Microbiological goals of periodontal therapy. Periodontol 2000 2006;42:180-218.

339. Loesche WJ, Grossman NS. Periodontal disease as a specific, albeit chronic, infection: diagnosis and treatment. Clin Microbiol Rev 2001;14:727-52.

340. Van Winkelhoff AJ, Winkel EG. Microbiological diagnostics in periodontics: biological significance and clinical validity. Periodontol 2000 2005;39:40-52.

341. Herrera D, Sanz M, Jepsen S, Needleman I, Roldán S. A systematic review on the effect of systemic antimicrobials as an adjunct to scaling and root planing in periodontitis patients. J Clin Periodontol 2002;29:136-59.

342. Laine ML, Farré MA, González G, et al. Polymorphisms of the interleukin-1 gene family, oral microbial pathogens, and smoking in adult periodontitis. J Dent Res 2001;80:1695-9.

343. Van Dyke TE, Sheilesh D. Risk factors for periodontitis. J Int Acad Periodontol 2005;7:3-7.

344. Behle JH, Sedaghatfar MH, Demmer RT, et al. Heterogeneity of systemic inflammatory responses to periodontal therapy. J Clin Periodontol 2009;36:287-94.

345. Paster BJ, Olsen I, Aas JA, Dewhirst FE. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol 2000 2006;42:80-7.

346. Papapanou PN, Behle JH, Kebschull M, et al. Subgingival bacterial colonization profiles correlate with gingival tissue gene expression. BMC Microbiol 2009;9:221.

347. Williams RC, Offenbacher S. Periodontal medicine: the emergence of a new branch of periodontology. Periodontol 2000 2000;23:9-12.

348. Pizzo G, Guiglia R, Lo Russo L, Campisi G. Dentistry and internal medicine: from the focal infection theory to the periodontal medicine concept. Eur J Intern Med 2010;21:496-502.

349. Stamm JW. Periodontal diseases and human health: new directions in periodontal medicine. Ann Periodontol 1998;3:1-2.

350. Bland PS, Goodson JM, Gunsolley JC, et al. Association of antimicrobial and clinical efficacy: periodontitis therapy with minocycline microspheres. J Int Acad Periodontol 2010;12:11-9.

351. Sakellari D, Ioannidis I, Antoniadou M, Slini T, Konstantinidis A. Clinical and microbiological effects of adjunctive, locally delivered chlorhexidine on patients with chronic periodontitis. J Int Acad Periodontol 2010;12:20-6.

352. Kirkwood KL, Cirelli JA, Rogers JE, Giannobile WV. Novel host response therapeutic approaches to treat periodontal diseases. Periodontol 2000 2007;43:294-315.

353. Pinho Mde N, Pereira LB, de Souza SL, et al. Short-term effect of COX-2 selective inhibitor as an adjunct for the treatment of periodontal disease: a clinical double-blind study in humans. Braz Dent J 2008;19:323-8.

354. Offenbacher S, Barros SP, Singer RE, Moss K, Williams RC, Beck JD. Periodontal disease at the biofilm-gingival interface. J Periodontol 2007;78:1911-25.

355. Mersel A, Call R, Mann J. Demographic trends of aging--application to gerodontology. Gerodontology 1987;6:9-15.

356. Hugoson A, Norderyd O. Has the prevalence of periodontitis changed during the last 30 years? J Clin Periodontol 2008;35:338-45.

357. Page RC. Vaccination and periodontitis: myth or reality. J Int Acad Periodontol 2000;2:31-43.

358. Koromantzos PA, Makrilakis K, Dereka X, Katsilambros N, Vrotsos IA, Madianos PN. A randomized, controlled trial on the effect of non-surgical periodontal therapy in patients with type 2 diabetes. Part I: effect on periodontal status and glycaemic control. J Clin Periodontol 2011;38:142-7.

359. Kraft P. Curses--winner's and otherwise--in genetic epidemiology. Epidemiology 2008;19:649-51.

360. Durick K, Mendlein J, Xanthopoulos KG. Hunting with traps: genome-wide strategies for gene discovery and functional analysis. Genome Res 1999;9:1019-25.

361. Huo D, Olopade OI. Interpretation of genome-wide association study results. Oncology (Williston Park) 2010;24:643, 646.

362. Katancik JA, Kritchevsky S, Weyant RJ, et al. Periodontitis and airway obstruction. J Periodontol. 2005;76:2161-7.

363. Kavvoura FK, Ioannidis JP. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. Hum Genet 2008;123:1-14.

364. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-1.

365. Zhou B, Shi J, Whittemore AS. Optimal methods for meta-analysis of genome-wide association studies. Genet Epidemiol 2011;35:581-91.

366. Zhou Z, Zhu G, Hariri AR, et al. Genetic variation in human NPY expression affects stress response and emotion. Nature 2008;452:997-1001.

367. Uusitupa MI, Karvonen MK, Pesonen U, Koulu M. Neuropeptide Y: a novel link between the neuroendocrine system and cholesterol metabolism. Ann Med 1998;30:508-10.

368. Karvonen MK, Pesonen U, Koulu M, et al. Association of a leucine(7)-to-proline(7) polymorphism in the signal peptide of neuropeptide Y with high serum cholesterol and LDL cholesterol levels. Nat Med 1998;4:1434-7.

369. Barr-Agholme M, Modéer T, Luthman J. Immunohistological study of neuronal markers in inflamed gingiva obtained from children with Down's syndrome. J Clin Periodontol 1991;18:624-33.

370. Lundy FT, Linden GJ. Neruopeptides and neurogenic mechanisms in oral and periodontal inflammation. Crit Rev Oral Biol Med 2004;15:82-98.

371. Lundy FT, El Karim IA, Linden GJ. Neuropeptide Y (NPY) and NPY Y1 receptor in periodontal health and disease. Arch Oral Biol 2009;54:258-62.

372. Haug SR, Heyeraas KJ. Modulation of dental inflammation by the sympathetic nervous system. J Dent Res 2006;85:488-95.

373. Abd El-Aleem SA, Morales-Aza BM, Donaldson LF. Sensory neuropeptide mRNA up-regulation is bilateral in periodontitis in the rat: a possible neurogenic component to symmetrical periodontal disease. Eur J Neurosci 2004;19:650-8.

374. Gaspersic R, Kovacic U, Glisovic S, Cör A, Skaleric U. Anti-NGF treatment reduces bone resorption in periodontitis. J Dent Res 2010;89:515-20.

375. Mansfield JC, Holden H, Tarlow JK, et al. Novel genetic association between ulcerative colitis and the anti-inflammatory cytokine interleukin-1 receptor antagonist. Gastroenterology 1994;106:637-42.

376. Kornman KS, Crane A, Wang HY, et al. The interleukin-1 genotype as a severity factor in adult periodontal disease. J Clin Periodontol 1997;24:72-7.

377. Hosokawa Y, Hosokawa I, Ozaki K, Nakae H, Matsuo T. TNFSF14 coordinately enhances CXCL10 and CXCL11 productions from IFN-gamma-stimulated human gingival fibroblasts. Mol Immunol 2010;47:666-70.

378. Hosokawa Y, Hosokawa I, Ozaki K, Nakanishi T, Nakae H, Matsuo T. Tea polyphenols inhibit IL-6 production in tumor necrosis factor superfamily 14-stimulated human gingival fibroblasts. Mol Nutr Food Res 2010l;54:S151-8.

379. Belstrøm D, Holmstrup P, Damgaard C, et al. The atherogenic bacterium Porphyromonas gingivalis evades circulating phagocytes by adhering to erythrocytes. Infect Immun 2011;79:1559-65.

380. Potempa M, Potempa J, Kantyka T, et al. Interpain A, a cysteine proteinase from Prevotella intermedia, inhibits complement by degrading complement factor C3. PLoS Pathog 2009;5:e1000316.

381. Cutler CW, Arnold RR, Schenkein HA. Inhibition of C3 and IgG proteolysis enhances phagocytosis of Porphyromonas gingivalis. J Immunol 1993;151:7016-29.

382. Hu Z, Wu C, Shi Y, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. Nat Genet 2011;43:792-6.

383. Kong M, Kim Y, Lee C. A strong synergistic epistasis between FAM134B and TNFRSF19 on the susceptibility to vascular dementia. Psychiatr Genet 2011;21:37-41.

384. Lewinger JP, Conti DV, Baurley JW, Triche TJ, Thomas DC. Hierarchical Bayes prioritization of marker associations from a genome-wide association scan for further investigation. Genet Epidemiol 2007;31:871-82.

385. Tintle N, Lantieri F, Lebrec J, Sohns M, Ballard D, Bickeböller H. Inclusion of a priori information in genome-wide association analysis. Genet Epidemiol 2009;33:S74-80.

386. Eleftherohorinou H, Wright V, Hoggart C, et al. Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. PLoS One 2009;4:e8068.

387. Zhang X, Huang S, Zou F, Wang W. TEAM: efficient two-locus epistasis tests in human genome-wide association study. Bioinformatics 2010;26:i217-27.

388. Li L, Yu M, Jason RD, et al. A mixture model approach in gene-gene and gene-environmental interactions for binary phenotypes. J Biopharm Stat 2008;18:1150-77.

389. Nica AC, Montgomery SB, Dimas AS, et al. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. PLoS Genet 2010;6:e1000895.

390. Nica AC, Dermitzakis ET. Using gene expression to investigate the genetic basis of complex disorders. Hum Mol Genet 2008;17:R129-34.

391. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 2008;26:1135-45.

392. Schuster SC. Next-generation sequencing transforms today's biology. Nat Methods 2008;5:16-8.

393. Mardis ER. The impact of next-generation sequencing technology on genetics. Trends Genet 2008;24:133-41.

394. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 2008;9:387-402.

395. Biesecker LG. Exome sequencing makes medical genomics a reality. Nat Genet 2010;42:13-4.

396. Biesecker LG, Shianna KV, Mullikin JC. Exome sequencing: the expert view. Genome Biol 2011;12:128.

397. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009;106:19096-101.

398. Khoury MJ, Millikan R, Little J, Gwinn M. The emergence of epidemiology in the genomics age. Int J Epidemiol 2004;33:936-44.

399. Witte JS. Genome-wide association studies and beyond. Annu Rev Public Health 2010;31:9-20.

400. McCarthy MI, Hirschhorn JN. Genome-wide association studies: past, present and future. Hum Mol Genet 2008;17:R100-1.

401. Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H. Interpretation of association signals and identification of causal variants from genome-wide association studies. Am J Hum Genet 2010;86:730-42.

402. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844-7.

403. Petretto E, Mangion J, Dickens NJ, et al. Heritability and tissue specificity of expression quantitative trait loci. PLoS Genet 2006;2:e172.

404. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003;160:636-45.

405. Huang P, Kraja AT, Tang W, et al. Factor relationships of metabolic syndrome and echocardiographic phenotypes in the HyperGEN study. J Hypertens 2008;26:1360-6.

406. He X, Zhang J. Toward a molecular understanding of pleiotropy. Genetics 2006;173:1885-91.

407. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. Bioinformatics 2010;26:1205-10.

408. Pendergrass SA, Brown-Gentry K, Dudek SM, et al. The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. Genet Epidemiol 2011;35:410-22.