Mendelian randomization analysis identified genes potentially pleiotropically associated with periodontitis

Feng Wang¹, Di Liu², Yong Zhuang³, Bowen Feng⁴, Wenjin Lu⁵, Jingyun Yang^{6,7}, Guanghui Zhuang¹

¹Department of Stomatology, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China

²Beijing Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing, China

³Department of Stomatology, Dalian Medical University, Dalian, Liaoning, China

⁴Odette School of Business, University of Windsor, Windsor, ON, Canada

⁵Department of Mathematics, University College London, London, United Kingdom

⁶Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA

⁷Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA

Correspondence to

Guanghui Zhuang

Department of Stomatology

The First Affiliated Hospital of Dalian Medical University

222 Zhongshan Rd, Xigang District

Dalian, Liaoning 116011

China

Tel: 86-18098876609

E-mail: zgh20062009@163.com

or

Jingyun Yang

Rush Alzheimer's Disease Center

Rush University Medical Center

1750 West Harrison Street, Suite 1000

Chicago, IL 60612

USA

Tel: 1-312-563-9035

E-mail: jingyun_yang@rush.edu

Abstract

Objective: To prioritize genes that were pleiotropically or potentially causally associated with periodontitis.

Methods: We applied the summary data-based Mendelian randomization (SMR) method integrating genome-wide association study (GWAS) for periodontitis and expression quantitative trait loci (eQTL) data to identify genes that were pleiotropically associated with periodontitis. We performed separate SMR analysis using CAGE eQTL data and GTEx eQTL data. SMR analysis were done for participants of European and East Asian ancestries, separately.

Results: We identified multiple genes showing pleiotropic association with periodontitis in participants of European ancestry and participants of East Asian ancestry. *PDCD2* (corresponding probe: ILMN_1758915) was the top hit showing pleiotropic association with periodontitis in participants of European ancestry, and *BX093763* (corresponding probe: ILMN_1899903) and *AC104135.3* (corresponding probe: ENSG00000204792.2) were the top hits in participants of East Asian ancestry using CAGE eQTL data and GTEx eQTL data, respectively.

Conclusion: We identified multiple genes that may be involved in the pathogenesis of periodontitis in participants of European ancestry and participants of East Asian ancestry. Our findings provided important leads to a better understanding of the mechanisms underlying periodontitis and revealed potential therapeutic targets for the effective treatment of periodontitis.

Keywords: periodontitis; pleiotropic association; gene expression quantitative trait loci; summary Mendelian randomization

Introduction

Periodontitis is a common disease characterized by an inflammatory response to commensal and pathogenic oral bacteria¹. The primary clinical features of periodontitis include periodontal pocketing, alveolar bone loss (BL), clinical attachment loss (CAL), and gingival inflammation². Based on the 2009–2014 National Health and Nutrition Examination Surveys data, it was estimated that periodontitis affected about 42% of US adults aged 30 to 79 years³. Periodontitis is considered as the main cause of tooth loss in adults. Moreover, it is also associated with various systemic conditions such as coronary heart disease⁴, diabetes⁵ and pre-term birth⁶. Periodontitis not only affects a patient's life, it also brings tremendous economic burden to the society, with an estimated global productivity loss due to untreated severe periodontitis being around \$38.85 billion in 2015⁷.

Periodontitis is a complex, multi-factorial infectious disease with possible contributions from multiple factors, including immunological response⁸, oral bacterial infections⁹, lifestyle factors such as smoking¹⁰ and alcohol consumption¹¹, psychological factors such as stress¹² and depression¹³, and systematic diseases such as diabetes¹⁴. Previous studies also suggested that genetics plays an important role in the pathogenesis of periodontitis. For example, genetically identical monozygotic twins have more than a twofold increased risk of early onset periodontitis, compared with dizygotic twins¹⁵. Another population-based twin study estimated that the heritability of periodontitis was approximately 50%¹⁶. Moreover, many GWAS and candidate gene studies have identified a number of genetic loci associated with the susceptibility of periodontitis¹⁷⁻²². However, the biological

mechanisms of these findings remain largely unclear, and more studies are needed to explore genes that are potentially causally associated with periodontitis to better understand the pathogenesis of periodontitis.

Mendelian randomization (MR) uses genetic variants as the proxy to randomization and is a promising tool to search for pleiotropic/potentially causal effect of an exposure (e.g., gene expression) on the outcome (e.g., periodontitis) without the need of conducting conventional randomized clinical trials (RCTs)²³. Confounding and reverse causation, which are commonly encountered in traditional association studies, can be greatly reduced by using MR. MR has been successful in identifying gene expression probes or DNA methylation loci that are pleiotropically/potentially causally associated with various phenotypes, such as neuropathologies of Alzheimer's disease and severity of COVID-19^{24,25}.

In this study, we applied the summary data-based MR (SMR) method integrating summarized GWAS data for periodontitis and cis- eQTL (expression quantitative trait loci) data to prioritize genes that are pleiotropically/potentially causally associated with periodontitis.

Methods

Data sources

eQTL data

In the SMR analysis, cis-eQTL genetic variants were used as the instrumental variables (IVs) for gene expression. We performed SMR analysis using gene expression data in blood due to the unavailability of eQTL data in the eye. Specifically, we used the CAGE eQTL summarized data²⁶, which included 2,765 participants, and the V7 release of the GTEx eQTL summarized data²⁷, which included 338 participants. The eQTL data can be downloaded at <https://cnsgenomics.com/data/SMR/#eQTLsummarydata>.

GWAS data for periodontitis

The GWAS summarized data were provided by a recent genome-wide association meta-analysis of periodontitis²⁸. The results were based on meta-analyses of 1000 genomes phase 1 version 2/3 imputed GWASs on periodontitis, with a total of nine cohorts from the Gene-Lifestyle Interactions in Dental Endpoints (GLIDE) consortium²⁹. Specifically, the meta-analysis for participants of European ancestry included seven cohorts with a total sample size of 45,563 (17,353 cases and 28,210 controls), and the meta-analysis for participants of East Asian ancestry included two cohorts with a total sample size of 17,350 (1,680 cases and 15,670 controls). All participating studies assumed an additive genetic model, adjusting for age, age-squared and other study-specific covariates. The GWAS summarized data can be downloaded at

<https://data.bris.ac.uk/data/dataset/2j2rqgzdxlq02oqbb4vmcnc2>.

SMR analysis

We conducted the SMR analysis with cis-eQTL as the IV, gene expression as the exposure, and periodontitis as the outcome. The analysis was done using the method as

implemented in the software SMR. Detailed information regarding the SMR method was reported in a previous publication³⁰. In brief, SMR applies the principles of MR to jointly analyze GWAS and eQTL summary statistics in order to test for pleiotropic association between gene expression and a trait due to a shared and potentially causal variant at a locus. We also conducted the heterogeneity in dependent instruments (HEIDI) test to evaluate the existence of linkage in the observed association. A P_{HEIDI} of less than 0.05 indicates that the observed association could be due to two distinct genetic variants in high linkage disequilibrium with each other. We adopted the default settings in SMR (e.g., minor allele frequency [MAF] > 0.01, removing SNPs in very strong linkage disequilibrium [LD, $r^2 > 0.9$] with the top associated eQTL, and removing SNPs in low LD or not in LD [$r^2 < 0.05$] with the top associated eQTL) except relaxing the threshold of eQTL P-value ($P_{\text{eQTL}} < 10^{-4}$) due to the exploratory nature of this study, and used false discovery rate (FDR) to adjust for multiple testing. We performed SMR analysis for participants of European and East Asian ancestries, separately, using CAGE and GTEx eQTL data, respectively, comprising a total of four SMR analyses.

We used Affymetrix exon array S1.0 platforms to annotate the transcripts. We conducted functional enrichment analysis using the functional annotation tool “Metascape” for the top tagged genes to functionally annotate putative transcripts³¹. Gene symbols corresponding to the ten top hit genes were used as the input of the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis.

Data cleaning and statistical/bioinformatical analysis was performed using R version 4.0.3 (<https://www.r-project.org/>), PLINK 1.9 (<https://www.cog-genomics.org/plink/1.9/>) and SMR (<https://cnsgenomics.com/software/smr/>).

Results

Basic information of the summarized data

The number of participants of the CAGE eQTL data is much larger than that of the GTEx eQTL data, so is the number of eligible probes. The sample size of the GWAS data for the European ancestry is much large than that for the East Asian ancestry, so is the number of eligible genetic variants. The detailed information was shown in **Table 1**.

SMR analysis in participants of European ancestry

In participants of European ancestry, we identified two genes showing pleiotropic association with periodontitis after correction for multiple testing using FDR (**Table 2**). Specifically, using the CAGE eQTL data, our SMR analysis identified two genes that were pleiotropically/potentially causally associated with periodontitis, including *PDCD2* (ILMN_1758915; $P_{\text{SMR}}=3.77 \times 10^{-5}$; **Figure 1**) and *D4S234E* (i.e., *NSG1*, ILMN_1772627; $P_{\text{SMR}}=9.08 \times 10^{-4}$; **Figure 2**). GO enrichment analysis of biological process and molecular function showed that the significant genes were involved in two GO terms, including positive regulation of cysteine-type endopeptidase activity (GO:2001056) and positive regulation of defense response (GO:0031349; **Figure S1A**). Concept network analysis of the ten top hit genes also revealed multiple domains related with endopeptidase activity (**Figure S1B**). More information could be found in **Table S1**.

Using the GTEx eQTL data, we did not identify any genes that were pleiotropically/potentially causally associated with periodontitis, after correction for multiple testing using FDR (**Table 2**). However, we found that two genes, *NSG1* (CAGE eQTL: ILMN_1772627, $P_{SMR}=9.08\times 10^{-4}$; GETx eQTL: ENSG00000168824.10, $P_{SMR}=1.35\times 10^{-3}$) and *SI00A12* (CAGE eQTL: ILMN_1748915, $P_{SMR}=1.58\times 10^{-3}$; GETx eQTL: ENSG00000163221.7, $P_{SMR}=1.90\times 10^{-3}$) were among the top hits in both SMR analyses. GO enrichment analysis of biological process and molecular function showed that the ten top hit genes were involved in two MAP kinase-related GO terms (GO:0043405 and GO:0043406; **Figure S1C**). Concept network analysis of the genes revealed multiple domains related with MAP kinase activity and inflammation (**Figure S1D**). More information could be found in **Table S2**.

SMR analysis in participants of East Asian ancestry

In participants of East Asian ancestry, we identified two genes showing significant pleiotropic association with periodontitis after correction for multiple testing using FDR (**Table 3**). Specifically, using the CAGE eQTL data, our SMR analysis identified one gene, *BX093763* (ILMN_1899903, $P_{SMR}=2.33\times 10^{-4}$). GO enrichment analysis of biological process and molecular function showed that the ten top hit genes were involved in one GO terms, axon development (GO:0061564; **Figure S2A**). Concept network analysis of the genes revealed multiple domains related with inflammation (**Figure S2B**). More information could be found in **Table S3**. Using the GTEx eQTL data, our SMR analysis identified one gene, *AC104135.3*, that was pleiotropically/potentially causally associated with periodontitis, after correction for multiple testing using FDR

(ENSG00000204792.2, $P_{\text{SMR}}=7.46 \times 10^{-4}$; **Figure 3**). GO enrichment analysis of biological process and molecular function did not find any significant GO terms. Concept network analysis of the genes revealed multiple domains related with endogenous peptide antigen (**Figure S2C**). More information could be found in **Table S4**.

We found that two genes were among the top hits in both SMR analyses, including *PSD4* (CAGE eQTL: ILMN_2154115, $P_{\text{SMR}}=1.79 \times 10^{-3}$; GTEx: ENSG00000125637.11, $P_{\text{SMR}}=1.84 \times 10^{-3}$) and *GFRA2* (CAGE eQTL: ILMN_1656300, $P_{\text{SMR}}=3.55 \times 10^{-3}$; GTEx: ENSG00000168546.6, $P_{\text{SMR}}=7.07 \times 10^{-3}$). More information could be found in **Table S3-4**.

Discussion

In the present study, we explored putative genes that showed pleiotropic/potentially causal association with periodontitis by integrating GWAS and eQTL data in the SMR analysis. We identified multiple genes, some of which represented novel genes, that might be involved in the pathogenesis of periodontitis in participants of European ancestry and participants of East Asian ancestry. Our findings provided helpful leads to a better understanding of the mechanisms underlying periodontitis and suggested potential therapeutic targets for the treatment of periodontitis.

A recent study investigated molecular biomarker candidates and biological pathways of chronic periodontitis using pooled datasets in the Gene Expression Omnibus (GEO) database, and identified 123 common differently expressed genes (DEGs), including 81 upregulated genes and 42 downregulated genes³². Several of the identified genes were also among the top hits in our SMR analysis. For example, the gene *NSG1* (Neuronal

Vesicle Trafficking Associated 1) was found to be downregulated in persons with chronic periodontitis. It also showed significant pleiotropic association with periodontitis in our study of participants of European ancestry (**Table 2**). *NSG1*, also known as *D4S234E* or *NEEP21*, is located on 4p16.3 in human and is a member of the neuron-specific gene family. It is the most important in regulating receptor recycling and synaptic transmission³³. p53, an important tumor suppressor gene, binds to the promoter region of *NSG1* and regulates its expression in response to DNA damage. Inhibition of *NSG1* expression suppressed apoptosis³⁴. The exact role of *NSG1* in the pathogenesis of periodontitis is unclear and warrants further research.

Another research integrating GWAS and eQTL data identified 10 genes whose expression might influence periodontitis³⁵. Of them, the gene *S100A12* (S calcium-binding protein A12) also appeared among the top hits in participants of European ancestry in the SMR analysis using CAGE and GTEx eQTL data (**Table 2**). *S100A12*, located on 1q21.3, is a member of the S100 family of EF-hand calcium-binding proteins³⁶. Previous studies indicated that it played a prominent role in the regulation of inflammatory processes and immune response³⁷. It was reported that the levels of *S100A12* were higher in participants with high periodontal inflammatory burden and were associated with the percentage of bleeding on probing³⁸. In gingival crevicular fluid and serum, the levels of *S100A12* increased with the inflammation of periodontium³⁹. These findings, together with ours, demonstrated the important role of *S100A12* in influencing periodontitis and highlighted the potential of this gene as a promising target for the prevention and treatment of periodontitis.

Our study was different from the previous study integrating GWAS and eQTL data³⁵. We used the GWAS summarized data for both European and East Asian ancestries, while the previous research only analyzed GWAS data of European ancestry. Similarly, we used both CAGE and GTEx eQTL data, while the previous research used only GTEx data. Moreover, we undertook a SMR analytic framework which focused on exploring genes showing pleiotropic association/potentially causal association with periodontitis while the previous research adopted a Sherlock approach which is a Bayesian statistical framework aiming to identify genes whose expression was associated with periodontitis susceptibility⁴⁰.

Our study was also very different from another MR research on periodontitis²⁹. Although both studies aimed to explore potentially causal factors for periodontitis, the previous research focused examining the causal role of total adiposity in the pathogenesis of periodontitis, while our study attempted to identify genes that were pleiotropically/potentially causally associated with periodontitis. The analytic approaches were also different: the previous research used genetic risk scores as the instrumental variables calculated from three genes (*FTO*, *MC4R* and *TMEM18*) by summing the number of BMI increasing alleles, while in our study, we used all the genetic variants from GWAS summarized data as the potential instrumental variables.

Our study has some limitations. The number of probes used in our SMR analysis was limited, especially in the SMR analysis of participants of East Asian ancestry. As a result,

we may have missed some genes which played important roles in the pathogenesis of periodontitis. The HEIDI test was significant for some of the identified genes (**Table 2-3**). Therefore, we could not rule out the possibility of horizontal pleiotropy, i.e., the identified association might be due to two distinct genetic variants in high linkage disequilibrium with each other. In addition, we only performed SMR analysis for participants of European and East Asian ancestries, and our findings might not be generalized to other populations. More studies are needed to validate our findings in independent populations. Due to the exploratory nature of study, we adopted correction for multiple testing to reduce false positive rate; however, we may have missed important SNPs or genes. Finally, we could not quantify the changes in gene expression in subjects with periodontitis in comparison with the control due to the unavailability of individual eQTL data.

Conclusions

In conclusion, our SMR analysis revealed multiple genes that were potentially pleiotropically associated with periodontitis. More studies are needed to explore the underlying physiological mechanisms in the etiology of periodontitis.

Acknowledgements

The study was supported by NIH/NIA grants P30AG10161, R01AG15819, R01AG17917, R01AG36042, U01AG61356 and 1RF1AG064312-01. Di Liu was supported by China Scholarship Council (CSC 201908110339).

The authors confirmed that all authors have reviewed the contents of the article being submitted, approved its contents, and validated the accuracy of the data.

Author contribution

DL, JY and GZ designed the study. FW, DL, BF and WL analyzed data and performed data interpretation. FW, DL, YZ and JY wrote the initial draft and BF, WL, and GZ contributed to writing subsequent versions of the manuscript. All authors reviewed the study findings and read and approved the final version before submission.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed by the authors.

References

1. Berezow AB, Darveau RP. Microbial shift and periodontitis. *Periodontol* 2000 2011;55:36-47.
2. Flemmig TF. Periodontitis. *Ann Periodontol* 1999;4:32-8.
3. Eke PI, Borgnakke WS, Genco RJ. Recent epidemiologic trends in periodontitis in the USA. *Periodontol* 2000 2020;82:257-67.
4. Humphrey LL, Fu R, Buckley DI, Freeman M, Helfand M. Periodontal disease and coronary heart disease incidence: a systematic review and meta-analysis. *J Gen Intern Med* 2008;23:2079-86.
5. Nascimento GG, Leite FRM, Vestergaard P, Scheutz F, Lopez R. Does diabetes increase the risk of periodontitis? A systematic review and meta-regression analysis of longitudinal prospective studies. *Acta Diabetol* 2018;55:653-67.
6. Walia M, Saini N. Relationship between periodontal diseases and preterm birth: Recent epidemiological and biological data. *Int J Appl Basic Med Res* 2015;5:2-6.
7. Righolt AJ, Jevdjevic M, Marcenes W, Listl S. Global-, Regional-, and Country-Level Economic Impacts of Dental Diseases in 2015. *J Dent Res* 2018;97:501-7.
8. Cekici A, Kantarci A, Hasturk H, Van Dyke TE. Inflammatory and immune pathways in the pathogenesis of periodontal disease. *Periodontol* 2000 2014;64:57-80.
9. Slots J. Periodontitis: facts, fallacies and the future. *Periodontol* 2000 2017;75:7-23.
10. Leite FRM, Nascimento GG, Scheutz F, Lopez R. Effect of Smoking on Periodontitis: A Systematic Review and Meta-regression. *Am J Prev Med* 2018;54:831-41.
11. Wang J, Lv J, Wang W, Jiang X. Alcohol consumption and risk of periodontitis: a meta-analysis. *J Clin Periodontol* 2016;43:572-83.
12. Hilgert JB, Hugo FN, Bandeira DR, Bozzetti MC. Stress, cortisol, and periodontitis in a population aged 50 years and over. *J Dent Res* 2006;85:324-8.
13. Nascimento GG, Gastal MT, Leite FRM, et al. Is there an association between depression and periodontitis? A birth cohort study. *J Clin Periodontol* 2019;46:31-9.
14. Preshaw PM, Bissett SM. Periodontitis and diabetes. *Br Dent J* 2019;227:577-84.
15. Corey LA, Nance WE, Hofstede P, Schenkein HA. Self-reported periodontal disease in a Virginia twin population. *J Periodontol* 1993;64:1205-8.
16. 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.
17. 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.
18. Munz M, Willenborg C, Richter GM, et al. A genome-wide association study identifies nucleotide variants at SIGLEC5 and DEFA1A3 as risk loci for periodontitis. *Hum Mol Genet* 2017;26:2577-88.
19. Teumer A, Holtfreter B, Volker U, et al. Genome-wide association study of chronic periodontitis in a general German population. *J Clin Periodontol* 2013;40:977-85.
20. Divaris K, Monda KL, North KE, et al. Exploring the genetic basis of chronic periodontitis: a genome-wide association study. *Hum Mol Genet* 2013;22:2312-24.
21. Freitag-Wolf S, Dommisch H, Graetz C, et al. Genome-wide exploration identifies sex-specific genetic effects of alleles upstream NPY to increase the risk of severe periodontitis in men. *J Clin Periodontol* 2014;41:1115-21.
22. Laine ML, Loos BG, Crielaard W. Gene polymorphisms in chronic periodontitis. *Int J Dent* 2010;2010:324719.

23. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014;23:R89-98.
24. Liu D, Wang Y, Jing H, Meng Q, Yang J. Mendelian randomization integrating GWAS and mQTL data identified novel pleiotropic DNA methylation loci for neuropathology of Alzheimer's disease. *Neurobiol Aging* 2021;97:18-27.
25. Liu D, Yang J, Feng B, Lu W, Zhao C, Li L. Mendelian randomization analysis identified genes pleiotropically associated with the risk and prognosis of COVID-19. *J Infect* 2020.
26. Lloyd-Jones LR, Holloway A, McRae A, et al. The Genetic Architecture of Gene Expression in Peripheral Blood. *Am J Hum Genet* 2017;100:371.
27. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, et al. Genetic effects on gene expression across human tissues. *Nature* 2017;550:204-13.
28. Shungin D, Haworth S, Divaris K, et al. Genome-wide analysis of dental caries and periodontitis combining clinical and self-reported data. *Nat Commun* 2019;10:2773.
29. Shungin D, Cornelis MC, Divaris K, et al. Using genetics to test the causal relationship of total adiposity and periodontitis: Mendelian randomization analyses in the Gene-Lifestyle Interactions and Dental Endpoints (GLIDE) Consortium. *Int J Epidemiol* 2015;44:638-50.
30. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 2016;48:481-7.
31. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019;10:1523.
32. Suzuki A, Horie T, Numabe Y. Investigation of molecular biomarker candidates for diagnosis and prognosis of chronic periodontitis by bioinformatics analysis of pooled microarray gene expression datasets in Gene Expression Omnibus (GEO). *BMC Oral Health* 2019;19:52.
33. Rengaraj D, Lee BR, Park KJ, et al. The distribution of neuron-specific gene family member 1 in brain and germ cells: Implications for the regulation of germ-line development by brain. *Dev Dyn* 2011;240:850-61.
34. Kudoh T, Kimura J, Lu ZG, Miki Y, Yoshida K. D4S234E, a novel p53-responsive gene, induces apoptosis in response to DNA damage. *Exp Cell Res* 2010;316:2849-58.
35. Li W, Zheng Q, Meng H, Chen D. Integration of genome-wide association study and expression quantitative trait loci data identifies AIM2 as a risk gene of periodontitis. *J Clin Periodontol* 2020;47:583-93.
36. Guignard F, Mauel J, Markert M. Identification and characterization of a novel human neutrophil protein related to the S100 family. *Biochem J* 1995;309 (Pt 2):395-401.
37. Pietzsch J, Hoppmann S. Human S100A12: a novel key player in inflammation? *Amino Acids* 2009;36:381-9.
38. Holmstrom SB, Lira-Junior R, Zwicker S, et al. MMP-12 and S100s in saliva reflect different aspects of periodontal inflammation. *Cytokine* 2019;113:155-61.
39. Pradeep AR, Martande SS, Singh SP, Suke DK, Raju AP, Naik SB. Correlation of human S100A12 (EN-RAGE) and high-sensitivity C-reactive protein as gingival crevicular fluid and serum markers of inflammation in chronic periodontitis and type 2 diabetes. *Inflamm Res* 2014;63:317-23.
40. He X, Fuller CK, Song Y, et al. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *Am J Hum Genet* 2013;92:667-80.

Figure legends

Figure 1. Prioritizing gene around *PDCD2* in association with periodontitis in participants of European ancestry.

Top plot, grey dots represent the $-\log_{10}(P \text{ values})$ for SNPs from the GWAS of periodontitis, and rhombuses represent the $-\log_{10}(P \text{ values})$ for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL results for ILMN_1758915 probe, tagging *PDCD2*. Bottom plot, location of genes tagged by the probes.

GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci

Figure 2. Prioritizing gene around *D4S234E* (i.e., *NSG1*) in association with periodontitis in participants of European ancestry.

Top plot, grey dots represent the $-\log_{10}(P \text{ values})$ for SNPs from the GWAS of periodontitis, and rhombuses represent the $-\log_{10}(P \text{ values})$ for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL results for ILMN_1772627 probe, tagging *D4S234E* (i.e., *NSG1*). Bottom plot, location of genes tagged by the probes.

GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci

Figure 3. Prioritizing gene around *AC104135.3* in association with periodontitis in participants of East Asian ancestry.

Top plot, grey dots represent the $-\log_{10}(P \text{ values})$ for SNPs from the GWAS of periodontitis, and rhombuses represent the $-\log_{10}(P \text{ values})$ for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL results for ENSG00000204792.2 probe, tagging *AC104135.3*. Bottom plot, location of genes tagged by the probes.

GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci

Figure S1. Functional enrichment and gene concept network analysis in participants of European ancestry.

A) Enriched GO terms based on genes identified using CAGE eQTL data; B) Concept network analysis of the genes identified using CAGE eQTL data; C) Enriched GO terms based on genes identified using GTEx eQTL data; and D) Concept network analysis of the genes identified using GTEx eQTL data.

GO, gene ontology; eQTL, expression quantitative trait loci

Figure S2. Functional enrichment and gene concept network analysis in participants of East Asian ancestry.

A) Enriched GO terms based on genes identified using CAGE eQTL data; B) Enriched GO terms based on genes identified using GTEx eQTL data; and C) Concept network analysis of the genes identified using GTEx eQTL data.

GO, gene ontology; eQTL, expression quantitative trait loci

Table 1. Basic information of the GWAS and eQTL data.

Data Source	Total number of participants	Number of eligible genetic variants or probes
European ancestry		
eQTL data		
CAGE	2,765	8,230
GTE_x	338	2,162
GWAS data	45,563	779,1334
East Asian ancestry		
eQTL data		
CAGE	2,765	7,304
GTE_x	338	2,010
GWAS data	17,350	418,8352

GWAS: genome-wide association studies; QTL, quantitative trait loci

Table 2. The top ten probes identified in the SMR analysis in participants of European ancestry.

eQTL data	Probe ID	Gene	CHR	Top SNP	P _{eQTL}	P _{GWAS}	Beta	SE	P _{SMR}	P _{HEIDI}	N _{SNP}
CAGE	ILMN_1758915	<i>PDCD2</i>	6	rs17875294	7.50×10^{-74}	2.34×10^{-5}	0.15	0.04	3.77×10^{-5}	4.52×10^{-1}	20
	ILMN_1772627	<i>NSG1</i>	4	rs6843595	7.10×10^{-291}	8.45×10^{-4}	0.05	0.02	9.08×10^{-4}	1.33×10^{-1}	20
	ILMN_1823130	<i>F01764</i>	4	rs2369111	1.74×10^{-17}	3.62×10^{-4}	0.24	0.07	1.00×10^{-3}	3.31×10^{-2}	20
	ILMN_1808251	<i>C9orf38</i>	9	rs4556138	6.67×10^{-10}	1.94×10^{-4}	0.35	0.11	1.40×10^{-3}	1.51×10^{-2}	20
	ILMN_1748915	<i>S100A12</i>	1	rs3014878	2.17×10^{-177}	1.45×10^{-3}	-0.06	0.02	1.58×10^{-3}	3.77×10^{-1}	20
	ILMN_1748221	<i>PADI6</i>	1	rs1535876	2.45×10^{-50}	1.69×10^{-3}	-0.12	0.04	2.13×10^{-3}	3.50×10^{-1}	20
	ILMN_1659511	<i>LOC645652</i>	1	rs10927894	3.79×10^{-7}	1.38×10^{-4}	0.60	0.20	2.30×10^{-3}	2.88×10^{-2}	20
	ILMN_1710937	<i>IFI16</i>	1	rs12122315	5.38×10^{-19}	1.27×10^{-3}	-0.22	0.07	2.46×10^{-3}	1.37×10^{-1}	20
	ILMN_1729801	<i>S100A8</i>	1	rs58644524	1.93×10^{-19}	1.36×10^{-3}	-0.20	0.06	2.52×10^{-3}	3.72×10^{-1}	20
	ILMN_2096405	<i>WDR37</i>	10	rs12768746	3.40×10^{-7}	2.40×10^{-4}	0.41	0.14	2.91×10^{-3}	4.89×10^{-2}	20
GTE_x	ENSG00000168824.10	<i>NSG1</i>	4	rs6414635	1.94×10^{-41}	9.57×10^{-4}	0.08	0.02	1.35×10^{-3}	2.34×10^{-1}	20
	ENSG00000256049.2	<i>PADI6</i>	1	rs10888031	2.95×10^{-42}	1.32×10^{-3}	-0.06	0.02	1.74×10^{-3}	3.22×10^{-1}	20
	ENSG00000163221.7	<i>S100A12</i>	1	rs57572338	1.56×10^{-15}	7.58×10^{-4}	-0.37	0.12	1.90×10^{-3}	5.90×10^{-1}	9
	ENSG00000184985.12	<i>SORCS2</i>	4	rs62289059	8.07×10^{-15}	8.95×10^{-4}	0.15	0.05	2.23×10^{-3}	4.30×10^{-1}	20
	ENSG00000127952.12	<i>STYXL1</i>	7	rs115332207	3.41×10^{-39}	2.22×10^{-3}	0.12	0.04	2.94×10^{-3}	9.10×10^{-1}	20

	ENSG00000233609.3	<i>RP11-62H7.2</i>	8	rs13259143	5.63×10^{-19}	2.60×10^{-3}	0.16	0.06	4.40×10^{-3}	6.75×10^{-2}	20
	ENSG00000106804.6	<i>C5</i>	9	rs7036980	2.84×10^{-10}	1.63×10^{-3}	0.19	0.07	4.76×10^{-3}	6.56×10^{-1}	20
	ENSG00000213523.5	<i>SRA1</i>	5	rs76128141	7.51×10^{-13}	2.88×10^{-3}	-0.27	0.10	5.91×10^{-3}	4.22×10^{-2}	14
	ENSG00000163421.4	<i>PROK2</i>	3	rs6777956	3.52×10^{-17}	4.01×10^{-3}	-0.21	0.08	6.40×10^{-3}	5.76×10^{-1}	17
	ENSG00000138835.18	<i>RGS3</i>	9	rs41306506	6.49×10^{-20}	5.77×10^{-3}	0.20	0.08	8.22×10^{-3}	9.88×10^{-1}	20

*The GWAS summarized data were provided by the study of Shungin et al. and can be downloaded at <https://data.bris.ac.uk/data/dataset/2j2rqgzexlq02oqbb4vmcnc2>. The CAGE and GTEx eQTL data can be downloaded at <https://cnsgenomics.com/data/SMR/#eQTLsummarydata>.

P_{eQTL} is the P-value of the top associated cis-eQTL in the eQTL analysis, and P_{GWAS} is the P-value for the top associated cis-eQTL in the GWAS analysis. Beta is the estimated effect size in SMR analysis, SE is the corresponding standard error, P_{SMR} is the P-value for SMR analysis, P_{HEIDI} is the P-value for the HEIDI test and N_{snp} is the number of SNPs involved in the HEIDI test.

FDR was calculated at $P=10^{-3}$ threshold.

Bold font means statistical significance after correction for multiple testing using FDR.

CHR, chromosome; HEIDI, heterogeneity in dependent instruments; SNP, single-nucleotide polymorphism; SMR, summary data-based Mendelian randomization; QTL, quantitative trait loci; FDR, false discovery rate; GWAS, genome-wide association studies

Table 3. The top ten probes identified in the SMR analysis in participants of East Asian ancestry.

eQTL data	Probe ID	Gene	CHR	Top SNP	P _{eQTL}	P _{GWAS}	Beta	SE	P _{SMR}	P _{HEIDI}	N _{SNP}
CAGE	ILMN_1899903	<i>BX093763</i>	5	rs984976	4.96×10^{-39}	1.27×10^{-4}	0.42	0.11	2.33×10^{-4}	2.67×10^{-1}	20
	ILMN_1734231	<i>DDOST</i>	1	rs6893	8.13×10^{-40}	1.17×10^{-3}	0.27	0.09	1.62×10^{-3}	9.21×10^{-2}	20
	ILMN_2154115	<i>PSD4</i>	2	rs2241976	1.10×10^{-77}	1.53×10^{-3}	0.26	0.08	1.79×10^{-3}	2.64×10^{-6}	20
	ILMN_2388155	<i>CASP3</i>	4	rs11721363	1.33×10^{-39}	2.07×10^{-3}	-0.88	0.29	2.72×10^{-3}	1.80×10^{-6}	20
	ILMN_1764522	<i>LMBR1</i>	7	rs73167977	3.88×10^{-22}	2.01×10^{-3}	-0.74	0.25	3.26×10^{-3}	9.03×10^{-4}	20
	ILMN_1656300	<i>GFRA2</i>	8	rs1479056	2.14×10^{-62}	3.07×10^{-3}	0.30	0.10	3.55×10^{-3}	5.51×10^{-1}	20
	ILMN_1791211	<i>DOK2</i>	8	rs1479056	1.14×10^{-47}	3.07×10^{-3}	0.34	0.12	3.72×10^{-3}	1.65×10^{-1}	20
	ILMN_1808661	<i>TOMM5</i>	9	rs7018807	9.27×10^{-61}	3.58×10^{-3}	-0.34	0.12	4.09×10^{-3}	4.20×10^{-3}	20
	ILMN_1910292	<i>BX094911</i>	4	rs3111820	3.02×10^{-5}	7.91×10^{-5}	-1.46	0.51	4.13×10^{-3}	2.69×10^{-1}	4
	ILMN_1805590	<i>NAA38</i>	7	rs7799229	1.28×10^{-67}	3.93×10^{-3}	0.24	0.08	4.44×10^{-3}	9.13×10^{-5}	20
GTE_x	ENSG00000204792.2	<i>AC104135.3</i>	2	rs12366	5.06×10^{-108}	6.39×10^{-4}	0.12	0.03	7.46×10^{-4}	3.11×10^{-1}	20
	ENSG00000204469.8	<i>PRRC2A</i>	6	rs2075800	3.28×10^{-8}	6.54×10^{-5}	1.72	0.53	1.22×10^{-3}	8.73×10^{-2}	20
	ENSG00000243753.1	<i>HLA-L</i>	6	rs3094204	1.45×10^{-17}	6.48×10^{-4}	-0.27	0.09	1.54×10^{-3}	1.97×10^{-3}	20
	ENSG00000224769.1	<i>AC069213.1</i>	3	rs6804822	3.13×10^{-14}	5.05×10^{-4}	0.42	0.13	1.56×10^{-3}	1.34×10^{-3}	20
	ENSG00000125637.11	<i>PSD4</i>	2	rs2241976	2.71×10^{-66}	1.53×10^{-3}	0.34	0.11	1.84×10^{-3}	1.96×10^{-5}	20
	ENSG00000144791.5	<i>LIMD1</i>	3	rs34448158	4.35×10^{-11}	5.24×10^{-4}	0.71	0.23	2.16×10^{-3}	5.00×10^{-1}	20

	ENSG00000261490.1	<i>RP11-448G15.3</i>	4	rs3756218	6.00×10^{-12}	1.23×10^{-3}	-0.92	0.31	3.42×10^{-3}	3.45×10^{-3}	20
	ENSG00000164307.8	<i>ERAP1</i>	5	rs26490	3.17×10^{-51}	3.21×10^{-3}	-0.17	0.06	3.85×10^{-3}	8.25×10^{-1}	20
	ENSG00000164039.10	<i>BDH2</i>	4	rs3775972	7.75×10^{-11}	1.71×10^{-3}	0.55	0.20	4.74×10^{-3}	1.11×10^{-1}	20
	ENSG00000168546.6	<i>GFRA2</i>	8	rs1479057	8.81×10^{-11}	3.07×10^{-3}	0.50	0.19	7.07×10^{-3}	5.34×10^{-1}	19

*The GWAS summarized data were provided by the study of Shungin et al. and can be downloaded at <https://data.bris.ac.uk/data/dataset/2j2rqgzexlq02oqbb4vmcnc2>. The CAGE and GTEx eQTL data can be downloaded at <https://cnsgenomics.com/data/SMR/#eQTLsummarydata>.

P_{eQTL} is the P-value of the top associated cis-eQTL in the eQTL analysis, and P_{GWAS} is the P-value for the top associated cis-eQTL in the GWAS analysis. Beta is the estimated effect size in SMR analysis, SE is the corresponding standard error, P_{SMR} is the P-value for SMR analysis, P_{HEIDI} is the P-value for the HEIDI test and N_{snp} is the number of SNPs involved in the HEIDI test.

FDR was calculated at $P=10^{-3}$ threshold.

Bold font means statistical significance after correction for multiple testing using FDR.

CHR, chromosome; HEIDI, heterogeneity in dependent instruments; SNP, single-nucleotide polymorphism; SMR, summary data-based Mendelian randomization; QTL, quantitative trait loci; FDR, false discovery rate; GWAS, genome-wide association studies

A **C** **B** **D**

A **B** **C**