HHS Public Access

Protocols, methods and tools for genome-wide association studies (GWAS) of dental traits

Cary S. Agler[1], **Dmitry Shungin**[2,3], **Andrea G. Ferreira Zandoná**[4], **Paige Schmadeke**[5,6], **Patricia V. Basta**[5,6], **Jason Luo**[7,8], **John Cantrell**[1], **Thomas D. Pahel Jr.**[1], **Beau Meyer**[9], **Jeannie Ginnis**[9], **John R. Shaffer**[10,11,12], **Arne S. Schäefer**[13], **Kari E. North**[7,14], and **Kimon Divaris**[5,9]

[1]Oral and Craniofacial Health Sciences, UNC School of Dentistry, University of North Carolina-Chapel Hill, Chapel Hill, NC, United States

[2]Department of Odontology, Umeå University, Umeå, Sweden

[3]Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, United States

[4]Department of Comprehensive Dentistry, Tufts University School of Dental Medicine, Boston, MA, United States

[5]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, Chapel Hill, NC, United States

[6]Biospecimen core processing facility, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, Chapel Hill, NC, United States

[7]Lineberger Comprehensive Cancer Center, School of Medicine, University of North Carolina, Chapel Hill, NC, United States

[8]Mammalian Genotyping core, University of North Carolina, Chapel Hill, NC, United States

[9]Department of Pediatric Dentistry, UNC School of Dentistry, University of North Carolina-Chapel Hill, Chapel Hill, NC, United States

[10]Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[11]Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[12]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States

[13]Department of Periodontology, Institute of Dental, Oral and Maxillary Medicine, Charité - University Medicine Berlin, Berlin, Germany

Correspondence: Kimon Divaris, DDS, PhD, 228 Brauer Hall, Department of Pediatric Dentistry, UNC School of Dentistry, CB#7450, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-7450, United States, T: (+1) 919-537-3556; F: (+1) 919-573-3950; Kimon_Divaris@unc.edu.

[14]Carolina Center for Genome Sciences, University of North Carolina-Chapel Hill, Chapel Hill, NC, United States

Abstract

Oral health and disease are known to be influenced by complex interactions between environmental (e.g., social and behavioral) factors and innate susceptibility. Although the exact contribution of genomics and other layers of 'omics' to oral health is an area of active research, it is well established that the susceptibility to dental caries, periodontal disease and other oral and craniofacial traits is substantially influenced by the human genome. A comprehensive understanding of these genomic factors is necessary for the realization of precision medicine in the oral health domain. To aid in this direction, the advent and increasing affordability of high-throughput genotyping has enabled the simultaneous interrogation of millions of genetic polymorphisms for association with oral and craniofacial traits. Specifically, genome-wide association studies (GWAS) of dental caries and periodontal disease have provided initial insights into novel loci and biological processes plausibly implicated in these two common, complex, biofilm-mediated diseases. This paper presents a summary of protocols, methods, tools and pipelines for the conduct of GWAS of dental caries, periodontal disease and related traits. The protocol begins with the consideration of different traits for both diseases and outlines procedures for genotyping, quality control, adjustment for population stratification, heritability and association analyses, annotation, reporting and interpretation. Methods and tools available for GWAS are being constantly updated and improved—with this in mind, the presented approaches have been successfully applied in numerous GWAS and meta-analyses among tens of thousands of individuals, including dental traits such as dental caries and periodontal disease. As such, they can serve as a guide or template for future genomics investigations of these and other traits.

Keywords

1. Introduction

Oral health and disease endpoints are the result of complex interactions between innate, behavioral, environmental and social factors. An exhaustive model of risk and protective factors, as well as behavioral and biologic factors at play remains elusive and is, arguably, unattainable. Nevertheless, advances in the oral and craniofacial health sciences and a growing body of evidence have illuminated the major influences on dental caries [1] and periodontal disease [2]—the two main, common-complex oral diseases. From a public health standpoint, both conditions are marked by pronounced social gradients and disparities [3]. From a pathogenetic standpoint, they are both to some degree biofilm- and host immunity-mediated [4–6], and modulated by behavior and host genomics [7].

Meaningful improvements in the prevention and treatment of oral diseases are likely to be achieved when aspects of precision medicine [8,9] are realized in the oral health domain, with genomics being a key part of the puzzle. Although the exact contribution of genomics

and other layers of 'omics' to oral health is an area of active investigation, it is well established that the susceptibility to dental caries and periodontal disease is to some degree determined by the human genome. Estimates of heritability for dental caries and periodontitis reported in the literature vary substantially, but are generally in the range of 20–50% [10–12], with more severe or early-onset forms of disease being more heritable.

The recent advent and increasing affordability of high-throughput genotyping has enabled comprehensive genomics investigations of oral and craniofacial traits; however, the current knowledge base of oral health genomics pales in comparison to what is known for other human traits [13]. While major advances have been made for several common diseases including asthma, diabetes, cancer, cardio-metabolic and psychiatric conditions, genome-wide association studies (GWAS) of dental caries [14–17] and periodontal diseases [18–28] have provided only initial insights into plausible novel loci and biological processes implicated in these two common, complex, biofilm-mediated diseases.

To aid in the conduct and harmonization of genomics investigations in the field and offer an overview of current common ground procedures, this paper presents a summary of protocols, methods, tools and pipelines for the conduct of GWAS of dental caries, periodontal diseases and related traits. First, it describes various traits for both diseases that can be carried forward to GWAS. Further, it outlines the major steps involved in genotyping, imputation, quality control, adjustment for population stratification, heritability and association analyses, annotation, reporting and interpretation. Methods and tools available for GWAS are being constantly updated and improved—with this in mind, the approaches presented in this paper been successfully applied in numerous GWAS and meta-analyses among tens of thousands of individuals, including dental traits such as dental caries and periodontal disease. As such, they can serve as a guide or template for future genomics investigations of these and other traits.

2. Materials

2.1 Sources of phenotype information

- Clinical examinations conducted by trained, ideally calibrated examiners—an examiner is typically consider calibrated upon achievement of pre-set levels of inter-examiner and intra-examiner agreement, usually based on weighted or unweighted *kappa*s, percent agreement or other criteria

- Clinical records (e.g., electronic patient record data)

- Intra-oral photographs, scored by trained and calibrated evaluators or algorithms/ software

- Administrative claims data (e.g., inferred by dental caries- or periodontitis-related treatment procedures)

- Self-reported data (e.g., obtained via written or electronic questionnaires or via telephone)

2.2 Sources of DNA for genotyping

- Blood samples

- Saliva samples [e.g., DNA Genotek Oragene DNA(OG)-500 kit for adults or OG-575 kit for young children or non-spitters, DNA Genotek, Ottawa, Ontario, Canada]

- Buccal swabs (e.g., MAWI technologies iSWAB™ kit, Mawi DNA Technologies LLC, Hayward, CA)

- Newborn blood spots [29]

2.3 DNA extraction, quantitation and quality assessment

- Automated DNA extraction from whole blood with either high-salt extraction methods or automated magnetic-bead extraction methods (i.e., PerkinElmer® Chemagic™ MSM I robotic system) or manual DNA extraction methods

- Automated DNA extraction from saliva, buccal brushes, or blood spots with automated magnetic-bead extraction methods (i.e., PerkinElmer® Chemagic™ MSM I robotic system) or manual DNA extraction methods.

- DNA quantitation using Nanodrop™ spectrophotometry, Quant-iT™ PicoGreen® fluorometry, Qubit™ fluorometry, or human-specific RNaseP assays

- Quality (i.e., sample purity) assessment using spectrophotometric methods (e.g., 260:280 and 260:230 absorbance ratios)

2.4 Genotyping supplies, equipment and software

- High-density genotyping arrays (e.g., Illumina Infinium Omni5Exome-4 BeadChip array, offering ~4.3 million variants and exome content) or targeted genotyping arrays (e.g., Illumina Metabochip or Immunochip)

- Array scanning (e.g., Illumina iScan)

- Variant calling software (e.g., Illumina GenomeStudio)

2.5 Imputation software. As referenced below in §3.3.4, currently the University of Michigan Imputation server is most frequently used for this step.

- Eagle2 [30]

- MiniMac [31]

- IMPUTE2 [32]

- GeneImp [33]

2.6 Genotype storage, transfer and management

- Cloud- or intranet-based storage with secure File Transfer Protocol (FTP) capabilities

- High-performance computing cluster allowing multi-threading and large memory jobs

- Server or workstations with common data management and programming suites (e.g., R, SAS, Stata)

2.7 Software commonly used for genome-wide association analyses

- PLINK [34] - most commonly used

- SUGEN [35] - implements weighted estimators to account for unequal sampling probability in complex survey design settings

- SNPtest [36] – analysis of single SNP frequentist and Bayesian association tests

- GenABEL Project suite [37, 38] - GWAS analyses and statistical 'omics' applications, including mixed model-based GWAS

- GWASTools [39] - an R/Bioconductor package for quality control and association analyses

- EMMAX [40], SOLAR [41, 42], GEMMA [43] - mixed models for association analysis accounting for sample structure (e.g., family-based studies)

- GMMAT [44] implemented via GENESIS [45] – generalized linear mixed model test for GWAS of binary traits accounting for population structure and relatedness, available via the GENESIS R/Bioconductor package

2.8 Software used for supporting analysis routines and visualizations

- GCTA [46] - heritability estimation and generation of eigenvectors for ancestry adjustment

- EIGENSTRAT [47] – principal component analysis-based correction for population stratification

- ADMIXTURE [48] - model-based estimation of ancestry in unrelated individuals

- MAGENTA [49] - gene-centric and pathway/gene-set enrichment analyses

- MAGMA [50] – gene-set analysis of GWAS data

- VEGAS/VEGAS2 [51] – Versatile gene-based association analysis

- LocusZoom [52] - local or web-based creation of regional association plots

2.9 Software used for meta-analysis and subsequent quality control and post-processing of GWAS results

- METAL [53] – meta-analysis of GWAS results

- GWAMA [54] – meta-analysis of GWAS summary statistics

- EasyStrata [55] - evaluation and visualization of stratified GWAS meta-analysis data

- SQC [56] – secure quality control for GWAS meta-analyses

2.10 Resources for genomic context and functional annotations

- ENCODE Project [57]

- Roadmap Epigenomics Project [58]

- UCSC Genome Browser [59]

- Integrative genomics viewer (IGV) [60]

- GTEx [61] - genotype-tissue expression project

- LDlink [62] – exploration of population-specific haplotype structures and links the alleles of interest with possible functional variants

- PolyPhen-2 [63] – prediction of functional effects of human non-synonymous coding single nucleotide polymorphisms (SNPs)

- ScanDB [64] – database including physical- and function-based annotation of SNPs (e.g., association with gene expression)

- Genehopper [65] - multidimensional scoring of gene-gene interactions

3. Methods

Obtaining high-quality phenotypes is the first important step before the actual conduct of a GWAS. Similar to any other type of study, traits carried forward to GWAS are subject to systematic and random errors, which are threats to the validity and precision of the obtained results. Arguably, reliance on very large sample sizes (tens or hundreds of thousands of individuals) may allow for the detection of true genetic signals despite some trait misclassification; however, in principle, higher quality phenotypes (e.g., clinical or biological) provide "better" results compared to lesser quality traits (e.g., self-reported) for a given sample size. Numerous traits are available for the study of dental caries and periodontal disease in the context of a GWAS—Table 1 presents a non-exhaustive list of such traits for dental caries (including early childhood caries or ECC [66]), developmental defects of the enamel (DDE) [67], periodontal disease and edentulism phenotype definitions.

As a second step, it is imperative that investigators have a good understanding of the study design under consideration and the characteristics of the sample—these factors may have an influential impact on the obtained results. For example, case-control studies and small sample sizes are known to produce spurious results due to systematic (e.g., Berksonian [68]) bias and random error. Third, application of strict criteria for determining what is a significant signal or variant is a requirement from a scientific rigor standpoint [69]. Two-stage designs and discovery-and-replication approaches are common strategies used to reduce the unavoidable type I errors in the single-stage or discovery samples and reduce the 'winner's curse' phenomenon [70]. Finally, an additional consideration prior to the conduct of a GWAS is the availability of sufficient quantity and quality of DNA. This is usually not an issue, unless a pediatric population is under study—young children are expected to be

less cooperative than adolescents or adults with research procedures involving venipuncture, as well as produce less saliva. With these in mind, carry out the procedures as follows:

3.1 DNA extraction

1. Extract, quantify and quality-assess DNA from blood, saliva, or buccal swab samples, according to the extraction kit manufacturer's instructions

2. At least 400ng of DNA is required for genotyping with the Infinium Omni5Exome-4 BeadChip referenced in §2.4

3. Plate extracted DNA and ship to genotyping core using a manifest according to each facility's instructions and best practices for safety and sample integrity

3.2 Genotype quality control, variant calling and exclusions.

1. An excellent, comprehensive description of quality control procedures is available in the Supplementary Material accompanying the 2007 Wellcome Trust GWAS [36]

2. Use manufacturer's instructions for processing and scanning of samples

3. Use HapMap-CEPH trios and duplicates and blind duplicate samples for quality control

4. Use GenomeStudio (https://goo.gl/BmNin1) or any other software for genotype calling [e.g., SNPs and copy number variants (CNV)]

5. Generate sample call and error rates

6. Identify sex mismatches and relationship errors (annotated versus genetic), gross chromosomal anomalies, mosaicism, contamination and sample swaps

7. Discard individual samples and markers that fail quality control criteria—these are highly dependent on the study design and sample (e.g., ranges of between 90 and 99% completion for markers and samples, duplicate sample discordances, Mendelian errors)

8. Deposit/upload SNP genotype data for transfer to imputation and association analyses—raw data or intensity files may be used for variant calling with third party software or uploaded for long-term storage

3.3 Imputation of genotypes

1. Use TOPMed, HapMap, 1000 Genomes Project, Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA) or the Haplotype Reference Consortium (HRC) reference panels (including multiple other cohorts) for imputation

2. Pre-phase (estimate haplotypes) using SHAPEIT [71, 72], as needed

3. Use any available imputation software pipeline Eagle2, MiniMac, IMPUTE2 or GeneImp, referenced in §2.5, as needed

4. Use these programs on premise or on the cloud; e.g., the University of Michigan Imputation server [73] provides free genotype imputation services using Minimac3

5. Apply quality filters to imputed SNPs and exclude based upon any of several available criteria [74] [e.g., based upon thresholds of imputation quality score, expected r^2 or 'info' score, Pearson correlation between true and "best-guess" genotypes (R^2), high missing data rate after imputation, or other available metrics]

6. Further exclude rare [minor allele frequency (MAF) <5%, <1%, or any other pre-set threshold] and monomorphic SNPs

3.4 Adjustment for ancestry and population stratification

1. Create eigenvectors or adjust for admixture proportion for population stratification control using programs referenced in §2.8

2. Alternatively used mixed models-based approaches using tools referenced in §2.8 to model putative genetic clustering

3. Exclude genetic outliers or related individuals (e.g., first and/or second-degree individuals) if the study design assumes a sample of unrelated individuals

3.5 GWAS analysis

1. Prepare the phenotype data using appropriate transformations, as necessary and dependent on the trait distribution characteristics

2. Account for population stratification using eigenvectors or admixture proportions developed as described in §3.4, as well as other study and participants' characteristics—typically sex, age (with or without an age-squared term), examination center, as well as complex survey design or familial relatedness (using software [40–45] detailed in §2.7), as necessary.

3. Use appropriate statistical models to test variant-phenotype associations—commonly used methods include linear regression modeling for continuous traits after appropriate transformation if needed (e.g., DMFS index) and logistic regression for binary outcomes (e.g., ECC case status and chronic periodontitis).

4. Consider a P value of less than 5×10^{-8} as the conventional genome-wide statistical significance threshold for GWAS, assuming 1 million independent association tests and a Bonferroni multiple testing correction.

5. One can carry out stratified GWAS analyses, on any variable of interest, depending on the research question (e.g., strata of sex, ancestry, or any strong risk factor such as fluoride for dental caries and smoking or diabetes for periodontal disease) and the study design.

6. Combine results, if applicable, using software and methods detailed in §2.8, using weights proportional to the square root of each study's sample size (i.e., Stouffer's method [75]) or by each study estimates' standard errors.

7. For trans-ethnic meta-analyses, one can use one of several additional approaches for meta-analyses based on random- or fixed-effects (e.g., trans-ethnic meta-analyses), as well as two-stage designs [76–80].

8. Generate QQ plots of association and compute the genomic inflation factor (lambda) to examine for possible systematic departures from expectation due to residual population stratification or other sources.

9. Generate Manhattan plots to visualize genome-wide results and genetic signals.

3.6 Reporting and annotation

1. For each SNP, it is customary to report key information (e.g., chromosome position, reference genome build, strand, reference allele, minor allele frequency, n, beta, standard error, 95% confidence interval, p-value, genotyped or imputed indicator, imputation quality score, as well as meta-analysis statistics (e.g., effect size with corresponding uncertainly estimate and p-values) if applicable.

2. Provide annotation for genomic context, linkage disequilibrium with functional variants, relative position or distance from exons or exon boundaries and gene promoter regions

3. Investigate the predicted effect of non-synonymous coding variants on the protein structure and function using PolyPhen-2 [63] as referenced in §2.10

4. Create regional association plots using LocusZoom [52] as referenced in §2.8

5. Visualize, compare and contrast results with other GWAS findings of the same, similar or potentially related traits using IGV [60] as referenced in §2.10

3.7 Post-GWAS analyses and procedures

1. Conduct gene-centric and pathway or gene-set enrichment analyses using MAGENTA [49], MAGMA [50] or other tools referenced in §2.8.

2. Examine individual marker associations with functional genome elements and gene expression using software or resources detailed in §2.10

3. Typically share or upload GWAS results onto a study-specific or 'community' server, to enable harmonization and meta-analyses or dissemination of results

4. Notes

1. Although here we include DNA extraction, genotyping and associated quality control procedures, most groups collaborate with core facilities to carry out these steps to produce genotype data.

2. Several steps of this protocol can be slow or time-consuming—for instance, genotyping of a large cohort can take months.

3. SNPs identified as genome-wide significant are unlikely to be causal variants but rather highlight genome areas (loci) of interest, where causative variants may lie.

4. Most GWAS signals are located in non-coding areas, where regulatory functional [57] elements of the genome are enriched—variants in regulatory DNA sites may have distant gene targets, and have been shown to systematically affect transcription factor recognition sequences, alter chromatin states and form regulatory networks [81].

5. The interrogation of disease-associated variants' correlation with relevant tissue-specific gene expression is a promising follow-up strategy of GWAS signals.

6. GWAS offer improvements over candidate-gene studies but if sample sizes are small, they are still prone to type I and type II errors—most signals identified in well-powered studies subsequently replicate.

7. Replicate novel signals based on directional consistency, nominal or (ideally) genome-wide significant association in external, independent samples.

8. The combination of multiple cohorts with phenotype and genotype data to interrogate oral and craniofacial health traits (e.g., caries and periodontitis) is warranted—dental consortia have begun forming (e.g., the GLIDE [82]). Despite some unavoidable challenges with achieving phenotype harmonization and addressing differences between populations (e.g., variations in disease prevalence between North American and European studies), single cohort reports are less informative and may be inefficient if realistic opportunities for joint, collaborative analyses exist.

9. Whenever feasible, follow-up promising, replicated loci with experimental (e.g., animal) models and other systems biology approaches.

10. Arguably, biologically-informed traits [20], data-driven clusters [83, 84] or endophenotypes have the potential to be informative targets for interrogation in the GWAS context, especially in cases where clinical endpoints are marked by heterogeneity and are substantially modified by dental treatment.

11. Third molars are frequently excluded from dental caries and periodontitis indices and definitions—this should be made explicit and justified in each cohort/study.

12. Tooth loss influences epidemiologic measures of periodontal disease (e.g., most periodontally affected teeth may be lost). For this reason, explicit consideration of tooth loss patterns [85] can offer advantages when interrogating cross-sectional datasets of periodontitis endpoints, by capturing individuals that may otherwise be misclassified. Similarly interrogation of edentulism [86] and dental caries experience patterns can be a promising strategy for dental caries GWAS, both in the primary [84, 87] and the permanent [83, 88, 89] dentition.

13. Two-step approaches including both genotyping a WGS in a subset of a cohort are considered efficient and cost-effective under certain circumstances to boost power to detect rare variants. For example, selection of a subset (e.g., 10%) of the 'most diseased' participants in the sample for WGS and their subsequent use to enrich the imputation reference panels in integrative analyses, assuming that

they are more likely to harbor rare, causal variants), has been recently proposed as an efficient strategy [90].

Acknowledgement

References

1. Selwitz RH, Ismail AI, Pitts NB. Dental caries. Lancet. 2007 1 6;369(9555):51–9. [PubMed: 17208642]

2. Offenbacher S Periodontal diseases: pathogenesis. Ann Periodontol. 1996 11;1(1):821–78. [PubMed: 9118282]

3. Lee JY, Divaris K. The ethical imperative of addressing oral health disparities: a unifying framework. J Dent Res. 2014 3;93(3):224–30. [PubMed: 24189268]

4. Nyvad B, Crielaard W, Mira A, Takahashi N, Beighton D. Dental caries from a molecular microbiological perspective. Caries Res. 2013;47(2):89–102. [PubMed: 23207320]

5. Genco RJ, Borgnakke WS. Risk factors for periodontal disease. Periodontol 2000. 2013 6;62(1):59–94. [PubMed: 23574464]

6. Hajishengallis G Immunomicrobial pathogenesis of periodontitis: keystones, pathobionts, and host response. Trends Immunol. 2014 1;35(1):3–11. [PubMed: 24269668]

7. Nibali L, Di Iorio A, Tu YK, Vieira AR. Host genetics role in the pathogenesis of periodontal disease and caries. J Clin Periodontol. 2017 3;44 Suppl 18:S52–S78. [PubMed: 27754553]

8. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015 2 26;372(9): 793–5. [PubMed: 25635347]

9. Divaris K Fundamentals of precision medicine. Compend Contin Educ Dent. 2017 9;38(Spec Iss 2): 31–33.

10. Vieira AR, Modesto A, Marazita ML. Caries: review of human genetics research. Caries Res. 2014;48(5):491–506. [PubMed: 24853115]

11. Michalowicz BS, Diehl SR, Gunsolley JC, Sparks BS, Brooks CN, Koertge TE, Califano JV, Burmeister JA, Schenkein HA. Evidence of a substantial genetic basis for risk of adult periodontitis. J Periodontol. 2000 11;71(11):1699–707. [PubMed: 11128917]

12. Loos BG, Papantonopoulos G, Jepsen S, Laine ML. What is the Contribution of Genetics to Periodontal Risk? Dent Clin North Am. 2015 10;59(4):761–80. [PubMed: 26427567]

13. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014 1;42(Database issue):D1001–6.

14. Wang X, Shaffer JR, Zeng Z, Begum F, Vieira AR, Noel J, Anjomshoaa I, Cuenco KT, Lee MK, Beck J, Boerwinkle E, Cornelis MC, Hu FB, Crosslin DR, Laurie CC, Nelson SC, Doheny KF, Pugh EW, Polk DE, Weyant RJ, Crout R, McNeil DW, Weeks DE, Feingold E, Marazita ML. Genome-wide association scan of dental caries in the permanent dentition. BMC Oral Health. 2012 12 21;12:57. [PubMed: 23259602]

15. Sofer T, Shaffer JR, Graff M, Qi Q, Stilp AM, Gogarten SM, North KE, Isasi CR, Laurie CC, Szpiro AA. Meta-Analysis of Genome-Wide Association Studies with Correlated Individuals: Application to the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Genet Epidemiol. 2016 9;40(6):492–501. [PubMed: 27256683]

16. Shaffer JR, Wang X, Feingold E, Lee M, Begum F, Weeks DE, Cuenco KT, Barmada MM, Wendell SK, Crosslin DR, Laurie CC, Doheny KF, Pugh EW, Zhang Q, Feenstra B, Geller F, Boyd HA, Zhang H, Melbye M, Murray JC, Weyant RJ, Crout R, McNeil DW, Levy SM, Slayton RL, Willing MC, Broffitt B, Vieira AR, Marazita ML. Genome-wide association scan for childhood caries implicates novel genes. J Dent Res. 2011 12;90(12):1457–62. [PubMed: 21940522]

17. Ballantine J, Carlson JC, Zandona A, Agler CS, Zeldin LP, Rozier G, Roberts MW, Basta PV, Luo J, Antonio-Obese ME, McNeil D, Weyant R, Crout RJ, Slayton R, Levy S, Shaffer JR, Marazita ML, North KE, Divaris K. Exploring the genomic basis of early childhood caries: a pilot study. Int J Paed Dentistry. doi: 10.1111/ipd.12344

18. Munz M, Willenborg C, Richter GM, Jockel-Schneider Y, Graetz C, Staufenbiel I, Wellmann J, Berger K, Krone B, Hoffmann P, van der Velde N, Uitterlinden AG, de Groot LCPGM, Sawalha AH, Direskeneli H, Saruhan-Direskeneli G, Guzeldemir-Akcakanat E, Keceli G, Laudes M, Noack B, Teumer A, Holtfreter B, Kocher T, Eickholz P, Meyle J, Doerfer C, Bruckmann C, Lieb W, Franke A, Schreiber S, Nohutcu RM, Erdmann J, Loos BG, Jepsen S, Dommisch H, Schaefer AS A genome-wide association study identifies nucleotide variants at SIGLEC5 and DEFA1A3 as risk loci for periodontitis. Hum Mol Genet. 2017 7 1;26(13):2577–2588. [PubMed: 28449029]

19. Sanders AE, Sofer T, Wong Q, Kerr KF, Agler C, Shaffer JR, Beck JD, Offenbacher S, Salazar CR, North KE, Marazita ML, Laurie CC, Singer RH, Cai J, Finlayson TL, Divaris K. Chronic Periodontitis Genome-wide Association Study in the Hispanic Community Health Study / Study of Latinos. J Dent Res. 2017 1;96(1):64–72. [PubMed: 27601451]

20. Offenbacher S, Divaris K, Barros SP, Moss KL, Marchesan JT, Morelli T, Zhang S, Kim S, Sun L, Beck JD, Laudes M, Munz M, Schaefer AS, North KE. Genome-wide association study of biologically informed periodontal complex traits offers novel insights into the genetic basis of periodontal disease. Hum Mol Genet. 2016 5 15;25(10):2113–2129. [PubMed: 26962152]

21. Zhang S, Divaris K, Moss K, Yu N, Barros S, Marchesan J, Morelli T, Agler C, Kim SJ, Wu D, North KE, Beck J, Offenbacher S. The Novel ASIC2 Locus is Associated with Severe Gingival Inflammation. JDR Clin Trans Res. 2016 7;1(2):163–170. [PubMed: 28459102]

22. Shimizu S, Momozawa Y, Takahashi A, Nagasawa T, Ashikawa K, Terada Y, Izumi Y, Kobayashi H, Tsuji M, Kubo M, Furuichi Y. A genome-wide association study of periodontitis in a Japanese population. J Dent Res. 2015 4;94(4):555–61. [PubMed: 25672891]

23. Hong KW, Shin MS, Ahn YB, Lee HJ, Kim HD. Genomewide association study on chronic periodontitis in Korean population: results from the Yangpyeong health cohort. J Clin Periodontol. 2015 7 25. doi: 10.1111/jcpe.12437.

24. Feng P, Wang X, Casado PL, Küchler EC, Deeley K, Noel J, Kimm H, Kim JH, Haas AN, Quinelato V, Bonato LL, Granjeiro JM, Susin C, Vieira AR. Genome wide association scan for chronic periodontitis implicates novel locus. BMC Oral Health. 2014 7 9;14:84. [PubMed: 25008200]

25. Shaffer JR, Polk DE, Wang X, Feingold E, Weeks DE, Lee MK, Cuenco KT, Weyant RJ, Crout RJ, McNeil DW, Marazita ML. Genome-wide association study of periodontal health measured by probing depth in adults ages 18–49 years. G3 (Bethesda). 2014 2 19;4(2):307–14. [PubMed: 24347629]

26. Teumer A, Holtfreter B, Völker U, Petersmann A, Nauck M, Biffar R, Völzke H, Kroemer HK, Meisel P, Homuth G, Kocher T. Genome-wide association study of chronic periodontitis in a general German population. J Clin Periodontol. 2013 11;40(11):977–85. [PubMed: 24024966]

27. Divaris K, Monda KL, North KE, Olshan AF, Reynolds LM, Hsueh WC, Lange EM, Moss K, Barros SP, Weyant RJ, Liu Y, Newman AB, Beck JD, Offenbacher S. Exploring the genetic basis of chronic periodontitis: a genome-wide association study. Hum Mol Genet. 2013 6 1;22(11):2312–24. [PubMed: 23459936]

28. Schaefer AS, Richter GM, Nothnagel M, Manke T, Dommisch H, Jacobs G, Arlt A, Rosenstiel P, Noack B, Groessner-Schreiber B, Jepsen S, Loos BG, Schreiber S. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet. 2010 2 1;19(3):553–62. [PubMed: 19897590]

29. St Julien KR, Jelliffe-Pawlowski LL, Shaw GM, Stevenson DK, O'Brodovich HM, Krasnow MA; Stanford BPD Study Group. High quality genome-wide genotyping from archived dried blood spots without DNA amplification. PLoS One. 2013 5 30;8(5):e64710.

30. Loh PR, Danecek P, Palamara PF, Fuchsberger C, Reshef AY, Finucane HK, Schoenherr S, Forer L, McCarthy S, Abecasis GR, Durbin R, Price LA. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016 11;48(11):1443–1448. [PubMed: 27694958]

31. Browning BL, Browning SR. Genotype Imputation with Millions of Reference Samples. Am J Hum Genet. 2016 1 7;98(1):116–26. [PubMed: 26748515]

32. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012 7 22;44(8): 955–9. [PubMed: 22820512]

33. Spiliopoulou A, Colombo M, Orchard P, Agakov F, McKeigue P. GeneImp: Fast Imputation to Large Reference Panels Using Genotype Likelihoods from Ultralow Coverage Sequencing. Genetics. 2017 5;206(1):91–104. [PubMed: 28348060]

34. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 9;81(3):559–75. [PubMed: 17701901]

35. Lin DY, Tao R, Kalsbeek WD, Zeng D, Gonzalez F 2nd, Fernández-Rhodes L, Graff M, Koch GG, North KE, Heiss G Genetic association analysis under complex survey sampling: the Hispanic Community Health Study/Study of Latinos. Am J Hum Genet. 2014 12 4;95(6):675–88. [PubMed: 25480034]

36. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007 6 7;447(7145):661–78. [PubMed: 17554300]

37. Karssen LC, van Duijn CM, Aulchenko YS. The GenABEL Project for statistical genomics. F1000Res. 2016 5 19;5:914. [PubMed: 27347381]

38. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007 5 15;23(10):1294–6. [PubMed: 17384015]

39. Gogarten SM, Bhangale T, Conomos MP, Laurie CA, McHugh CP, Painter I, Zheng X, Crosslin DR, Levine D, Lumley T, Nelson SC, Rice K, Shen J, Swarnkar R, Weir BS, Laurie CC. GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. Bioinformatics. 2012 12 15;28(24):3329–31. [PubMed: 23052040]

40. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 2010 4;42(4):348–54. [PubMed: 20208533]

41. Fardo DW, Zhang X, Ding L, He H, Kurowski B, Alexander ES, Mersha TB, Pilipenko V, Kottyan L, Nandakumar K, Martin L. On family-based genome-wide association studies with large pedigrees: observations and recommendations. BMC Proc. 2014 6 17;8(Suppl 1 Genetic Analysis Workshop 18Vanessa Olmo):S26.

42. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet. 1998 5;62(5):1198–211. [PubMed: 9545414]

43. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012 6 17;44(7):821–4. [PubMed: 22706312]

44. Chen H, Wang C, Conomos MP, Stilp AM, Li Z, Sofer T, Szpiro AA, Chen W, Brehm JM, Celedón JC, Redline S, Papanicolaou GJ, Thornton TA, Laurie CC, Rice K, Lin X. Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. Am J Hum Genet. 2016 4 7;98(4):653–66. [PubMed: 27018471]

45. Conomos MP, Thornton T and Gogarten SM (2017). GENESIS: GENetic EStimation and Inference in Structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness. R package version 2.6.1.

46. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011 1 7;88(1):76–82. [PubMed: 21167468]

47. 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 8;38(8): 904–9. [PubMed: 16862161]

48. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009 9;19(9):1655–64. [PubMed: 19648217]

49. Segrè AV, Wei N; DIAGRAM Consortium; MAGIC Investigators, Altshuler D, Florez JC Pathways targeted by antidiabetes drugs are enriched for multiple genes associated with type 2 diabetes risk. Diabetes. 2015 4;64(4):1470–83. [PubMed: 25368101]

50. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015 4 17;11(4):e1004219.

51. Mishra A, Macgregor S. VEGAS2: Software for More Flexible Gene-Based Testing. Twin Res Hum Genet. 2015 2;18(1):86–91. [PubMed: 25518859]

52. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010 9 15;26(18):2336–7. [PubMed: 20634204]

53. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010 9 1;26(17):2190–1. [PubMed: 20616382]

54. Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics. 2010 5 28;11:288. [PubMed: 20509871]

55. Winkler TW, Kutalik Z, Gorski M, Lottaz C, Kronenberg F, Heid IM. EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. Bioinformatics. 2015 1 15;31(2):259–61. [PubMed: 25260699]

56. Huang Z, Lin H, Fellay J, Kutalik Z, Hubaux JP. SQC: secure quality control for meta-analysis of genome-wide association studies. Bioinformatics. 2017 8 1;33(15):2273–2280. [PubMed: 28379351]

57. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012 9 6;489(7414):57–74. [PubMed: 22955616]

58. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu YC, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shoresh N, Epstein CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh KH, Feizi S, Karlic R, Kim AR, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJ, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai LH, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M Integrative analysis of 111 reference human epigenomes. Nature. 2015 2 19;518(7539):317–30. [PubMed: 25693563]

59. Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, Fischer CM, Gibson D, Gonzalez JN, Guruvadoo L, Haeussler M, Heitner S, Hinrichs AS, Karolchik D, Lee BT, Lee CM, Nejad P, Raney BJ, Rosenbloom KR, Speir ML, Villarreal C, Vivian J, Zweig AS, Haussler D, Kuhn RM, Kent WJ. The UCSC Genome Browser database: 2017 update. Nucleic Acids Res. 2017 1 4;45(D1):D626–D634. [PubMed: 27899642]

60. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011 1;29(1):24–6. [PubMed: 21221095]

61. Rivas MA, Pirinen M, Conrad DF, Lek M, Tsang EK, Karczewski KJ, Maller JB, Kukurba KR, DeLuca DS, Fromer M, Ferreira PG, Smith KS, Zhang R, Zhao F, Banks E, Poplin R, Ruderfer DM, Purcell SM, Tukiainen T, Minikel EV, Stenson PD, Cooper DN, Huang KH, Sullivan TJ, Nedzel J; GTEx Consortium; Geuvadis Consortium, Bustamante CD, Li JB, Daly MJ, Guigo R, Donnelly P, Ardlie K, Sammeth M, Dermitzakis ET, McCarthy MI, Montgomery SB, Lappalainen T, MacArthur DG Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. Science. 2015 5 8;348(6235):666–9. [PubMed: 25954003]

62. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015 11 1;31(21):3555–7. [PubMed: 26139635]

63. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods. 2010 4;7(4):248–9. [PubMed: 20354512]

64. Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL, Dolan ME, Cox NJ. SCAN: SNP and copy number annotation. Bioinformatics. 2010 1 15;26(2):259–62. [PubMed: 19933162]

65. Munz M, Tönnies S, Balke WT, Simon E. Multidimensional gene search with Genehopper. Nucleic Acids Res. 2015 7 1;43(W1):W98–103. [PubMed: 25990726]

66. Drury TF, Horowitz AM, Ismail AI, Maertens MP, Rozier RG, Selwitz RH. Diagnosing and reporting early childhood caries for research purposes. A report of a workshop sponsored by the National Institute of Dental and Craniofacial Research, the Health Resources and Services Administration, and the Health Care Financing Administration. J Public Health Dent. 1999 Summer;59(3):192–7. [PubMed: 10649591]

67. Clarkson J, O'Mullane D. A modified DDE Index for use in epidemiological studies of enamel defects. J Dent Res. 1989 3;68(3):445–50. [PubMed: 2921385]

68. Westreich D Berkson's bias, selection bias, and missing data. Epidemiology. 2012 1;23(1):159–64. [PubMed: 22081062]

69. Kraft P, Zeggini E, Ioannidis JP. Replication in genome-wide association studies. Stat Sci. 2009 11 1;24(4):561–573. [PubMed: 20454541]

70. Kraft P Curses--winner's and otherwise--in genetic epidemiology. Epidemiology. 2008 9;19(5): 649–51 [PubMed: 18703928]

71. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2011 12 4;9(2):179–81. [PubMed: 22138821]

72. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods. 2013 1;10(1):5–6. [PubMed: 23269371]

73. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, Loh PR, Iacono WG, Swaroop A, Scott LJ, Cucca F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C. Next-generation genotype imputation service and methods. Nat Genet. 2016 10;48(10):1284–1287. [PubMed: 27571263]

74. Hancock DB, Levy JL, Gaddis NC, Bierut LJ, Saccone NL, Page GP, Johnson EO. Assessment of genotype imputation performance using 1000 Genomes in African American studies. PLoS One. 2012;7(11):e50610.

75. Stouffer SA. Adjustment during army life. Princeton University Press; 1949.

76. Morris AP. Transethnic meta-analysis of genomewide association studies. Genet Epidemiol. 2011 12;35(8):809–22. [PubMed: 22125221]

77. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet. 2011 5 13;88(5):586–98. [PubMed: 21565292]

78. Han B, Eskin E. Interpreting meta-analyses of genome-wide association studies. PLoS Genet. 2012;8(3):e1002555.

79. Hong J, Lunetta KL, Cupples LA, Dupuis J, Liu CT. Evaluation of a Two-Stage Approach in Trans-Ethnic Meta-Analysis in Genome-Wide Association Studies. Genet Epidemiol. 2016 5;40(4):284–92. [PubMed: 27061095]

80. Mägi R, Horikoshi M, Sofer T, Mahajan A, Kitajima H, Franceschini N, McCarthy MI; COGENT-Kidney Consortium, T2D-GENES Consortium, Morris AP Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. Hum Mol Genet. 2017 9 15;26(18):3639–3650. [PubMed: 28911207]

81. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutyavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012 9 7;337(6099):1190–5. [PubMed: 22955828]

82. Shungin D, Cornelis MC, Divaris K, Holtfreter B, Shaffer JR, Yu YH, Barros SP, Beck JD, Biffar R, Boerwinkle EA, Crout RJ, Ganna A, Hallmans G, Hindy G, Hu FB, Kraft P, McNeil DW, Melander O, Moss KL, North KE, Orho-Melander M, Pedersen NL, Ridker PM, Rimm EB, Rose LM, Rukh G, Teumer A, Weyant RJ, Chasman DI, Joshipura K, Kocher T, Magnusson PK, Marazita ML, Nilsson P, Offenbacher S, Davey Smith G, Lundberg P, Palmer TM, Timpson NJ, Johansson I, Franks PW. 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 4;44(2):638–50. [PubMed: 26050256]

83. Shaffer JR, Feingold E, Wang X, Tcuenco KT, Weeks DE, DeSensi RS, Polk DE, Wendell S, Weyant RJ, Crout R, McNeil DW, Marazita ML. Heritable patterns of tooth decay in the permanent dentition: principal components and factor analyses. BMC Oral Health. 2012 3 9;12:7. [PubMed: 22405185]

84. Shaffer JR, Feingold E, Wang X, Weeks DE, Weyant RJ, Crout R, McNeil DW, Marazita ML. Clustering tooth surfaces into biologically informative caries outcomes. J Dent Res. 2013 1;92(1): 32–7. [PubMed: 23064960]

85. Morelli T, Moss KL, Beck J, Preisser JS, Wu D, Divaris K, Offenbacher S. Derivation and Validation of the Periodontal and Tooth Profile Classification System for Patient Stratification. J Periodontol. 2017 2;88(2):153–165. [PubMed: 27620653]

86. Recent advances in oral health. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser. 1992;826:1–37. [PubMed: 1462607]

87. Zeng Z, Feingold E, Wang X, Weeks DE, Lee M, Cuenco DT, Broffitt B, Weyant RJ, Crout R, McNeil DW, Levy SM, Marazita ML, Shaffer JR. Genome-wide association study of primary dentition pit-and-fissure and smooth surface caries. Caries Res. 2014;48(4):330–8. [PubMed: 24556642]

88. Shaffer JR, Feingold E, Wang X, Lee M, Tcuenco K, Weeks DE, Weyant RJ, Crout R, McNeil DW, Marazita ML. GWAS of dental caries patterns in the permanent dentition. J Dent Res. 2013 1;92(1):38–44. [PubMed: 23064961]

89. Zeng Z, Shaffer JR, Wang X, Feingold E, Weeks DE, Lee M, Cuenco KT, Wendell SK, Weyant RJ, Crout R, McNeil DW, Marazita ML. Genome-wide association studies of pit-and-fissure- and smooth-surface caries in permanent dentition. J Dent Res. 2013 5;92(5):432–7. [PubMed: 23470693]

90. Hu YJ, Li Y, Auer PL, Lin DY. Integrative analysis of sequencing and array genotype data for discovering disease associations with rare mutations. Proc Natl Acad Sci U S A. 2015 1 27;112(4): 1019–24. [PubMed: 25583502]

91. Divaris K, North KE, Slade GD, Barros SP, Moss K, Beck JD, Offenbacher S. Genome-wide association study of tooth morbidity. J Dent Res 2014;92 (Spec Iss B): 190702.

92. Eke PI, Page RC, Wei L, Thornton-Evans G, Genco RJ. Update of the case definitions for population-based surveillance of periodontitis. J Periodontol. 2012 12;83(12):1449–54. [PubMed: 22420873]

93. Schaefer AS, Richter GM, Groessner-Schreiber B, Noack B, Nothnagel M, El Mokhtari NE, Loos BG, Jepsen S, Schreiber S. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. PLoS Genet. 2009 2;5(2):e1000378.

Table 1.

Example traits for dental caries, developmental defects of the enamel, periodontal disease and edentulism that can be interrogated in the context of GWAS

Condition	Trait	Type	Description
dental caries	DMFS/dmfs	continuous	The sum of decayed, missing due to caries, filled/restored tooth surfaces; 5 surfaces are enumerated for molars (and premolars, in the permanent dentition), 4 for all other teeth [14].
	DMFS/dmfs>0	binary	Case definition for caries experience vs. caries-free; this definition corresponds to 'early childhood caries' or ECC among children <72 months of age [17, 66]. defs and dfs indices may also be considered in the primary dentition
	DFS/S	continuous/ proportion	The proportion of diseased tooth surfaces among the surfaces present [14].
	DM_TFS	continuous	The sum of decayed, missing due to all reasons (presumably mainly caries and periodontal disease) and filled/restored surfaces. A tooth morbidity index [91]
developmental defects of the enamel	DDE>0	binary	The presence of any developmental defect of the enamel, as defined by the Clarkson and O'Mullane modified epidemiologic index [67], of diameter 1mm of greater, as assessed on the facial or labial surface of the entire dentition.
periodontal disease	Chronic periodontitis, CDC/AAP definition (2012)	categorical	Based upon probing depth (PD) and clinical attachment loss (AL), categories of health/no disease, mild, moderate and severe disease are defined [91]. Severe disease is defined as 2 interproximal sites with AL 6 mm (not on same tooth) and 1 interproximal site with PD 5 mm [27, 92].
	Periodontal extent scores	continuous/ proportion	Proportion of all examined sites with probing depth, attachment loss equal or greater to a pre-defined threshold (e.g., 3mm, 4mm, etc.), or bleeding upon probing; e.g., extent of severe gingival bleeding [21].
	Full-mouth summary scores	Continuous	Summary scores or means of periodontal indices; e.g., mean interproximal clinical attachment level [19].
	UNC Periodontal Profile Class (PPC) system	categorical	Latent class analysis-based categories of periodontal disease experience incorporating patterns of tooth loss [85].
	Periodontal Complex Traits (PCTs)	principal component-derived eigenvectors	Complex periodontal traits derived by principal component analysis of clinical, inflammatory (e.g., gingival crevicular fluid IL-1b) and microbial (e.g., periodontal pathogen levels) data [20].
	Aggressive periodontitis (AgP), localized	binary	radiographically-confirmed 50% bone loss at 2–6 teeth, diagnosed at age of 35 or younger [93].

Condition	Trait	Type	Description
	Aggressive periodontitis (AgP), generalized	binary	Radiographically-confirmed 50% bone loss at 7 teeth, diagnosed at age 35 or younger [93].
edentulism	number of remaining natural teeth	continuous	The sum of remaining natural teeth (0–28 or 0–32, depending on the inclusion of third molars).
	"functional dentition"	binary	The World Health Organization definition of reduced dental arches that preserve basic functions—i.e., the retention of a natural, esthetic, functional dentition of no less than 20 teeth throughout life with no need for tooth replacement [86].
	no remaining natural teeth	binary	Complete loss of the natural dentition.