

From the Oral Microbiome to Periodontitis: Bioinformatics Identifies ANK3

Zi-Feng Wei

w2004zff@qq.com

Capital Medical University <https://orcid.org/0009-0006-1885-0064>

Article

Keywords: periodontal disease, oral microbiota, periodontal ligament stem cell, single-cell transcriptomics, spatial transcriptomics, Mendelian randomization, machine learning

Posted Date: September 12th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10879334/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is no conflict of interest

Abstract

Background

Oral microbiome (OM) dysbiosis is associated with periodontitis, but its causal relevance and the biological mechanisms of periodontal tissue destruction remain unclear.

Materials and methods

Bidirectional Mendelian randomization (MR) assessed causal associations of species-level tongue dorsum and salivary microbiome traits with periodontal disease using East Asian genome-wide association study (GWAS) data. Candidate genes mapped from MR instruments were prioritized based on overlap with differentially expressed genes (DEGs), module genes identified by weighted gene co-expression network analysis (WGCNA), and eQTLGen-based MR findings. The candidate gene was further evaluated by MR using East Asian cis-expression quantitative trait locus (cis-eQTL) data and European-ancestry genome-wide association study (GWAS) data for periodontitis subtypes. Single-cell RNA sequencing and spatial transcriptomics examined its cellular role. Phenome-wide association study (PheWAS) and drug–target interaction prediction assessed on-target safety.

Results

Forward MR identified 46 species-level taxa associated with periodontal disease, enriched in *Campylobacter_A*, *Fusobacterium*, and *Streptococcus*; reverse MR identified 47 taxa, mainly in *Prevotella* and *Streptococcus*. Intersection analysis identified *ANK3*, which was downregulated in periodontitis. MR consistently supported a protective causal effect of *ANK3* on periodontitis across cohorts. Single-cell analysis showed predominant *ANK3* expression in periodontal ligament stem cells (PDLSCs), with reduced expression in periodontitis. Spatial transcriptomics showed overlap of *ANK3*-enriched PDLSC regions with periodontitis genetic risk signals. PheWAS and drug–target interaction prediction revealed no major on-target safety concerns.

Conclusions

These findings link causally relevant OM features to *ANK3*–related PDLSC vulnerability in periodontitis and identify a potential host cellular axis for further investigation.

Introduction

The oral cavity is the second largest microbial habitat in the human body, hosting more than 700 species that constitute the OM. Microbial communities of the teeth, tongue dorsum, and saliva differ in composition and exhibit notable spatial heterogeneity. Under healthy conditions, these communities

sustain a stable oral microenvironment through symbiotic and competitive interactions [1]. Once this equilibrium is disrupted, opportunistic pathogens overgrow, dysregulate the host immune-inflammatory response, and ultimately drive irreversible destruction of the periodontal tissues. In recent years, accumulating evidence has linked dysbiosis of the commensal OM to the onset, progression, and refractoriness of periodontitis [2].

Periodontitis is the leading cause of tooth loss in adults worldwide. According to the GBD 2021 analysis, severe periodontitis affected over 1 billion people, with a global age-standardized prevalence of 12.50% that continues to rise [3]. Periodontitis arises from the interplay of OM dysbiosis, host immune dysregulation, and genetic susceptibility. Mechanical debridement combined with local antimicrobial therapy remains the standard of care, yet 20% to 30% of patients still progress to irreversible functional loss [4]. Multicohort studies show that disease progression is accompanied by enrichment of pathogenic taxa in saliva, on the tongue dorsum, and within subgingival plaque, alongside stage-dependent shifts in microbial α -diversity [5]. Beyond serving as a marker of disease activity, OM dysbiosis is increasingly implicated as a driver of disease onset. Whether this relationship is truly causal, and through which targets and mechanisms it operates, however, remains unclear.

To dissect this complex pathogenesis, MR uses genetic variants to investigate causal relationships between exposures and outcomes [6]. Since single nucleotide polymorphisms (SNPs) are randomly allocated during meiosis and allele transmission is independent of disease status, MR effectively mitigates confounding and reverse causation. Integrating MR with bulk, single-cell, and spatial transcriptomic data enables disease-associated signals to be examined at the tissue, cellular, and spatial levels, providing convergent evidence for the identification of key effector genes. This study applies these approaches to assess the causal contribution of the OM to periodontal disease and to identify, validate, and functionally characterize candidate host effector genes and pathways, providing a basis for potential targeted interventions.

Materials and methods

Study Design

The workflow is shown in Fig. 1. Following STROBE-MR guidelines, causal relationships between the OM and periodontal disease were inferred by bidirectional MR, using summary statistics from two East Asian periodontal disease cohorts meta-analyzed with METAL [7]. Instrumental variables (IVs) for the implicated taxa were mapped to genes. Candidate-gene prioritisation integrated bulk transcriptomic evidence, co-expression network evidence, and transcriptome-wide MR evidence. scRNA-seq localized the gene in periodontal tissues, and scDEG, Cellchat, Scissor, and scGSEA analyses characterized its cell-type-specific functions. Spatial transcriptomics resolved its distribution and associated cell types using RCTD, and gsMap localized disease-risk genetic signals within the tissue. Finally, Phenome-wide association analysis assessed potential on-target safety, whereas candidate-drug prediction and molecular docking evaluated target-compound binding.

Data Sources

OM summary statistics were obtained from the 4D-SZ cohort in Shenzhen, China (<https://ftp2.cngb.org/>). Deep shotgun metagenomic data were available for 2,017 tongue-dorsum and 1,915 salivary samples obtained from 2,984 individuals [8].

Periodontal disease GWAS summary statistics were obtained from two Japanese cohorts: bbj-a-142 [9] from OpenGWAS (<https://opengwas.io/>) and GCST90018677 [10] from the GWAS Catalog (<https://www.ebi.ac.uk/>).

Periodontitis GWAS data were retrieved from the GWAS Catalog and FinnGen for periodontitis [11], acute periodontitis [12], and chronic periodontitis [13].

Transcriptomic data were retrieved from GEO (<https://www.ncbi.nlm.nih.gov/geo/>), including bulk datasets GSE16134 [14], the single-cell dataset GSE171213 [15], and the spatial transcriptomic dataset GSE206621 [16].

East Asian cis-eQTL data were derived from the Chinese 1KCP project (<https://yanglab.westlake.edu.cn/1kcp/>), based on 1,116 diploid genome assemblies [17]. European eQTLGen cis-eQTL data came from OpenGWAS [18].

Structures of candidate drugs and the target protein were obtained from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and the Protein Data Bank (<http://www.rcsb.org/>), respectively. The data are summarized in Supplementary Table 1.

Selection of IVs

IVs were selected through the following pipeline. Few SNPs were available at $P < 5 \times 10^{-6}$. SNPs associated with each OM trait at $P < 5 \times 10^{-5}$ were therefore retained to provide sufficient genetic information for MR analyses and the downstream gene-prioritization framework. Independent SNPs were then obtained by LD clumping in PLINK ($r^2 = 0.001$, $kb = 10,000$). SNPs with a minor allele frequency below 0.01 and palindromic SNPs were further excluded. To guard against weak instrument bias, the F -statistic was computed for each SNP, and those with $F < 15$ were discarded. Finally, Steiger filtering was performed to rule out reverse causation.

Bidirectional MR Analysis

Causal associations between OM traits and periodontal disease were estimated using five MR methods: inverse-variance weighted (IVW), MR-Egger, weighted median, weighted mode, and simple mode. The IVW method combined SNP-specific Wald ratio estimates to evaluate the effect of OM traits on periodontal disease. MR-Egger regression was used to assess the robustness of the causal estimates and to test for directional horizontal pleiotropy. The weighted median method was also applied, as it can

provide consistent estimates when at least 50% of the total weight is contributed by valid IVs. IVW served as the primary estimator, with significance set at $P < 0.05$. An association was considered robust when the other four methods produced effect estimates consistent in direction with the IVW estimate.

To assess the robustness of the causal estimates, horizontal pleiotropy was examined using the MR-PRESSO global test and the MR-Egger intercept test, and heterogeneity among IVs was evaluated with Cochran's Q test. $P > 0.05$ was interpreted as no statistical evidence of pleiotropy or heterogeneity. SNPs were annotated using SNPnexus (<https://www.snp-nexus.org/v4/>).

MR Analysis of eQTLGen cis-eQTLs and Periodontitis

Summary-level whole-blood cis-eQTL data for 19,244 genes were obtained from the European-ancestry eQTLGen study via OpenGWAS and used as exposures, with periodontitis (GCST90478260) as the outcome. For each gene, cis-eQTLs with $P < 5 \times 10^{-8}$ were selected as candidate IVs and clumped using PLINK ($r^2 = 0.01$, kb = 100). All subsequent procedures for instrument selection, MR estimation, and sensitivity analyses, were performed as described above.

DEGs Screening

GSE16134 was obtained from GEO and included 69 healthy control samples and 241 periodontitis samples. Gene expression data were generated using the GPL570 platform. Following initial data processing, differential expression analysis was performed using the limma package. Genes with $|\log FC| > 0.25$ and $P < 0.05$ were considered DEGs.

To integrate the MR and differential expression results, genes were matched according to the direction of their effects. Specifically, downregulated DEGs were intersected with protective genes identified by MR, whereas upregulated DEGs were intersected with risk genes identified by MR. The overlapping genes were combined to form a candidate gene set with consistent effect directions.

WGCNA Analysis

WGCNA was applied to GSE16134 to identify co-expression modules associated with periodontitis. After selection of the module most strongly correlated with disease status, the centrality of candidate genes within this module was evaluated using module membership (MM), gene significance (GS), and intramodular connectivity (kWithin). Candidate genes meeting the criteria of $|MM| > 0.80$ and $|GS| > 0.20$ were regarded as strongly module-associated and disease-relevant, supporting their potential hub roles in the periodontitis-associated module. These genes were then intersected with the direction-consistent candidate genes identified through the integrated MR and differential expression analyses.

Cross-ancestry MR Validation

To evaluate genetic evidence for the candidate genes across ancestry groups, two-sample MR analyses were conducted using ancestry-matched cis-eQTL data as exposures. In Europeans, four periodontitis subtype outcomes were examined, comprising acute and chronic periodontitis datasets obtained from the GWAS Catalog and FinnGen. The FinnGen datasets served as the discovery cohorts, whereas the corresponding GWAS Catalog datasets were used as independent replication cohorts. In East Asians, periodontal disease was analyzed as the outcome. IVs were selected at $P < 5 \times 10^{-8}$. The inverse-variance weighted (IVW) method was used as the primary estimator, and sensitivity analyses were performed to assess the robustness of the MR findings.

Single-cell Analysis

The GSE171213 scRNA-seq dataset included nine periodontal tissue samples: four from healthy controls and five from patients with periodontitis. Following quality control in Seurat, cells with 200–4,500 detected genes, 1,000–25,000 total UMI counts, and mitochondrial content below 10% were retained. Harmony was used to correct batch effects across samples, followed by cell clustering at a resolution of 0.20. Initial cell-type annotations were assigned using the tooth-tissue reference in scMayoMap [19], and SingleR was subsequently applied to further annotate the remaining unidentified clusters.

To investigate the biological functions of the candidate genes in specific cell subpopulations, scDEG analysis was used to identify the subpopulations in which these genes were differentially expressed between healthy controls and patients with periodontitis. Scissor then integrated the GSE171213 scRNA-seq data with the GSE16134 bulk transcriptomic dataset using a network-regularized logistic model ($\alpha = 0.60$). Periodontitis-associated Scissor+ cells were identified, with the selected-cell fraction restricted to below 20% [20]. Model reliability was assessed using 10×10 -fold cross-validation with permuted labels. CellChat characterized intercellular signaling pathways, whereas scGSEA evaluated functional pathways associated with the target genes within these subpopulations ($P < 0.05$, adjusted $P < 0.10$, $|\text{NES}| > 1.00$).

Spatial Transcriptomics Analysis

Spatial transcriptomic data from oral mucosal tissues of one healthy control and three patients with periodontitis were obtained from GSE206621. Following quality control, the expression pattern of the candidate gene was mapped across each tissue section. RCTD was then applied using the annotated GSE171213 scRNA-seq dataset as a reference to estimate the cellular composition of each spot [21]. The candidate cell subpopulation, defined as the cluster in which the candidate gene was differentially expressed, was then projected onto the spots to delineate its spatial distribution.

Periodontitis GWAS summary statistics (GCST90478260) were subsequently incorporated, and gsMap was used to map disease-risk signals at the spot level [22]. This integrative framework was used to assess spatial colocalization among candidate gene expression, the candidate cell subpopulation, and genetic risk.

PheWAS Analysis

To characterize the broader phenotypic associations of the target genes, gene-level PheWAS results were queried using the AstraZeneca PheWAS Portal (<https://azphewas.com/>). Each target gene was searched individually, and its associations with binary and continuous phenotypes were retrieved. Associations with $P < 1 \times 10^{-6}$ were retained as suggestive signals and reviewed to evaluate the broader phenotypic relevance and potential safety implications of the target genes.

Molecular Docking Analysis

ANK3-associated compounds were retrieved from the D3 collection of DSigDB (<https://dsigdb.tanlab.org/DSigDBv1.0/>) and screened by reviewing PubMed-indexed studies. Compounds were excluded if no therapeutic relevance to inflammatory conditions, infectious diseases, and oral diseases was reported or if severe toxicity or an unfavorable safety profile had been documented. The retained compounds were subjected to blind protein–ligand docking using CB-Dock2 (<https://cadd.labshare.cn/cb-dock2/index.php>). Predicted binding cavities, docking poses, and Vina scores were evaluated for each protein–ligand complex. A Vina score below -5 kcal/mol was considered indicative of potential binding.

Results

Bidirectional MR Between OM and Periodontal Disease

Forward MR analysis identified 46 species-level oral microbial taxa significantly associated with periodontal disease (Supplementary Table 2). Of these, 23 were protective, comprising 14 tongue-dorsum and 9 salivary taxa, whereas the remaining 23 were risk factors, comprising 6 tongue-dorsum and 17 salivary taxa. These taxa were primarily distributed across the genera *Campylobacter*, *Fusobacterium*, and *Streptococcus* (Fig. 2A). Reverse MR analysis indicated potential causal effects of periodontal disease on 47 taxa (Supplementary Table 3). Among them, 28 taxa, including 11 from the tongue dorsum and 17 from saliva, showed inverse associations, whereas 19 taxa, including 9 from the tongue dorsum and 10 from saliva, showed positive associations. These taxa predominantly belonged to the genera *Prevotella* and *Streptococcus* (Fig. 2C). Sankey diagrams illustrating their taxonomic distributions at the order, family, and genus levels are presented in Fig. 2B, D.

Effect estimates from all five MR methods were statistically significant and directionally concordant. All selected SNPs had F -statistics above 15, supporting adequate instrument strength. The MR-Egger intercept provided no evidence of directional horizontal pleiotropy ($P > 0.05$), and the MR-PRESSO global test indicated no evidence of global horizontal pleiotropy; no outlier variants were detected. Cochran's Q test showed no evidence of heterogeneity ($P > 0.05$), while the Steiger directionality test supported the specified exposure-to-outcome direction.

Periodontitis Gene Set in Periodontal Microenvironment Dysregulation

MR analysis identified 1,898 genes with evidence of causal associations with periodontitis, comprising 944 risk genes and 954 protective genes (Fig. 3A). Sensitivity analyses found no evidence of heterogeneity, horizontal pleiotropy, or reverse causation (Supplementary Table 4). Differential expression analysis of GSE16134 identified 3,887 DEGs, including 2,051 downregulated and 1,836 upregulated genes, as shown in the volcano plot (Fig. 3B). Venn plots identified 184 genes with concordant directions across the MR and differential expression results, including 87 upregulated risk genes and 97 downregulated protective genes (Fig. 3D, E).

WGCNA-Based Validation of the Candidate Gene as a Network Hub

WGCNA identified the black and yellow modules as having the strongest correlations with periodontitis status, with module–trait correlations of -0.65 and 0.65 , respectively (Fig. 3F). Based on the criteria of $|MM| > 0.80$ and $|GS| > 0.20$, 108 genes in the black module and 218 genes in the yellow module were retained (Fig. 3G, H & Supplementary Table 4). The intersection of the IV-mapped gene set, WGCNA-derived gene set, and directionally concordant gene set identified *ANK3* as a shared candidate (Fig. 3I).

In GSE16134, *ANK3* was assigned to the black module and exhibited high intramodular connectivity ($k_{Within} = 5.668$) strong module membership ($MM = 0.873$; $P = 2.919 \times 10^{-98}$), and a significant association with periodontitis status ($GS = -0.58$; $P = 2.942 \times 10^{-29}$). *ANK3* also met the predefined hub-gene criteria, supporting its central position within the periodontitis-associated black module. The expression profile of *ANK3* is shown in Fig. 3J.

Candidate Gene MR Validation

In the European-ancestry analysis, two-sample MR indicated a protective effect of *ANK3* across four outcomes representing acute and chronic periodontitis ($P_{IVW} < 0.05$). In East Asians, MR analysis of periodontal disease also showed a protective association (Fig. 4 & Supplementary Table 6), with evidence of heterogeneity.

Single-cell Analysis and Functional Inference

Following quality control, 11 cell types were annotated: B cells, dental pulp stem cells, endothelial cells, epithelial cells, macrophages, mesenchymal stem cells, monocytes, NK cells, perivascular cells, PDLSCs, and T cells (Fig. 5A, B). *ANK3* was predominantly expressed in PDLSCs. Within this cell type, *ANK3* was downregulated in the periodontitis group relative to the healthy control group (Fig. 5D). Scissor analysis classified epithelial cells as Scissor+ cells, while cell–cell communication analysis showed enhanced communication between epithelial cells and PDLSCs in the disease group (Fig. 5E, F). Within PDLSCs, scGSEA showed that *ANK3* expression was positively correlated with enrichment of TNF- α signaling via NF- κ B but negatively correlated with enrichment of protein secretion, allograft rejection, mTORC1 signaling, and oxidative phosphorylation (Fig. 5G).

Spatial Localization and Genetic Risk Mapping

RCTD-based deconvolution suggested that PDLSCs were enriched in regions adjacent to epithelial cells. Given that the epithelium forms the first line of defense against microbial invasion, this spatial association may reflect a link between epithelial barrier disruption and alterations in the adjacent PDLSC niche. In addition, *ANK3*-enriched regions within PDLSCs overlapped with periodontitis genetic risk regions identified by gsMap, indicating spatial colocalization of gene expression and disease susceptibility signals (Fig. 5H).

PheWAS Results

ANK3 showed no significant associations with adverse clinical phenotypes at the predefined statistical threshold ($P < 1 \times 10^{-6}$). These findings suggest that targeting *ANK3* carries a low risk of unintended side effects in periodontitis treatment.

Molecular Docking Results

DSigDB identified six *ANK3*-upregulating compounds. After excluding compounds without reported therapeutic relevance to inflammatory conditions, infectious diseases, or oral diseases, as well as those with severe toxicity or unfavorable safety profiles, Latomoxef and Metronidazole were retained as candidate drugs. Molecular docking indicated potential binding between Latomoxef and *ANK3*, with a Vina score of -6.7 kcal/mol, further supporting its prioritization (Fig. 5I).

Discussion

This study combined bidirectional MR with multi-omics analyses to examine periodontitis from both microbial and host perspectives. MR identified multiple taxa within OM with causal links to periodontitis, whereas the integrated analyses characterized candidate genes across tissue, cellular, and spatial contexts. These findings provide insight into periodontitis pathogenesis and highlight potential therapeutic targets.

Periodontitis-specific GWAS data remain scarce in East Asian populations; therefore, periodontal disease served as a proxy outcome in the exploratory MR analysis. Although the broader phenotype definition and relaxed instrument-selection threshold require caution when interpreting the MR estimates, candidate genes were prioritized using several complementary lines of evidence. These included cis-eQTL-based MR across ancestries and periodontitis subtypes, together with bulk, single-cell, and spatial transcriptomic analyses that placed the genetic findings in biological and cellular context.

Species-level signals in the forward MR analysis were concentrated in *Campylobacter_A*, *Fusobacterium*, and *Streptococcus*. *Campylobacter_A* was largely risk-associated, *Streptococcus* was protective, and the direction of effect for *Fusobacterium* varied across species. Recent OM studies partly align with these findings. Specific *Campylobacter* species have been reported to increase with periodontitis severity and

to correlate with anaerobic, inflamed periodontal pockets and salivary metabolic disturbances, supporting their role as risk-associated taxa [23, 24]. In contrast, *Streptococcus* is predominantly enriched in periodontal health and help maintain a health-associated microbial community and resist pathogen colonization [25]. *Fusobacterium* has a more complex role. As a bridging organism in dental biofilms, *Fusobacterium* promotes microbial colonization and biofilm maturation [26], whereas under dysbiotic conditions it facilitates red-complex colonization [27], reflecting a context-dependent role consistent with the heterogeneous associations observed across *Fusobacterium* species.

In the reverse MR analysis, species-level signals were mainly enriched in *Streptococcus* and *Prevotella*, with heterogeneous associations within both genera. This pattern suggests that periodontitis reshapes species composition within these genera rather than uniformly increasing or decreasing their constituent species. Under disease conditions, *Streptococcus* species show altered transcriptional activity and a shift toward opportunistic members associated with more severe periodontal states [28]. Similarly, deep periodontal pockets, anaerobic conditions, and tissue degradation products favor inflammation-adapted *Prevotella* species over health-associated core or transitional species [29].

Scissor analysis identified epithelial cells as a major Scissor+ population associated with the periodontitis bulk phenotype, implicating the epithelial compartment in disease processes related to barrier dysfunction. Consistent with this interpretation, disruption of the gingival epithelial barrier facilitates the penetration of bacterial products and promotes periodontal inflammation [30]. Spatial transcriptomics showed that epithelial regions were located near PDLSC-enriched regions, whereas scDEG and gsMap analyses implicated PDLSCs as the principal cellular context associated with *ANK3*. Taken together, these findings raise the possibility that pathogenic cues from the inflamed epithelial compartment may reach the adjacent periodontal regenerative niche. This hypothesis is relevant to the biological distinction between gingivitis and periodontitis: gingivitis involves gingival inflammation without attachment loss, whereas periodontitis entails destruction of the tooth-supporting tissues [31]. Given that bacterial and inflammatory stimuli can impair the osteogenic and regenerative functions of PDLSCs [32], persistent epithelial barrier dysfunction could expose neighboring PDLSCs to pathogenic cues, thereby compromising endogenous periodontal repair and potentially contributing to the progression from reversible gingival inflammation to persistent periodontal attachment loss. However, direct epithelial–PDLSC signaling was not examined in this study, and the proposed relationship requires experimental validation.

Within this framework, *ANK3* expression was lower in PDLSCs from periodontitis samples. scGSEA revealed a positive association between *ANK3* expression and TNF- α signaling via NF- κ B, together with inverse associations with protein secretion, allograft rejection, mTORC1 signaling, and oxidative phosphorylation. These correlations connect *ANK3* to inflammatory, metabolic, and secretory programs relevant to PDLSC homeostasis [33, 34], but do not clarify whether *ANK3* downregulation contributes to regenerative dysfunction or arises as a consequence of altered PDLSC states. Reduced *ANK3* expression is therefore best regarded as a candidate molecular feature of inflammation-associated PDLSC reprogramming rather than direct evidence of stemness loss. Further functional studies are required to

determine the role of *ANK3* in PDLSC self-renewal, osteogenic differentiation, and resistance to inflammatory stress.

This study has several limitations. First, periodontitis-specific GWAS data were unavailable for East Asian populations; periodontal disease was therefore used as a proxy outcome in the exploratory MR analysis. The relaxed IV threshold was used to identify candidate signals rather than to support definitive causal estimates. Second, the transcriptomic, scRNA-seq, and spatial transcriptomic analyses were based on public datasets with limited sample sizes and potential heterogeneity in tissue composition, clinical severity, and sequencing platforms. Third, validation in independent datasets and complementary analyses using WGCNA, multi-ancestry MR, scDEGs, Scissor, and gsMap strengthened the robustness of the findings. However, these computational approaches cannot establish direct molecular mechanisms. The proposed involvement of the epithelial barrier, the spatial proximity between epithelial cells and PDLSCs, and the functional changes associated with *ANK3* therefore require validation in independent cohorts, primary periodontal tissues, and cellular or animal models. Finally, the biological role of *ANK3* in PDLSC homeostasis, inflammatory responses, and periodontal tissue destruction remains to be experimentally defined.

Conclusion

This study provides genetic and multi-omics evidence linking OM dysbiosis to periodontal disease and identifies *ANK3* as a candidate protective hub gene in periodontitis. Integrated single-cell and spatial analyses indicated spatial proximity between PDLSCs and epithelial cells, providing a plausible spatial context in which epithelial barrier disruption could influence the adjacent PDLSC niche through microbial and inflammatory stimuli. Together, these findings highlight a potential role for *ANK3*-associated PDLSC dysfunction in periodontitis pathogenesis and provide a rationale for further evaluating *ANK3* as a therapeutic target.

Declarations

Ethical approval and consent

The original study obtained all necessary ethical approvals and participant consent. Therefore, no additional approval was needed for this study of data.

Conflicts of Interest

Author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

No specific funding was received for this work.

Author Contributions

Zi-Feng Wei: Conceptualization, methodology, data curation, formal analysis, investigation, visualization, writing - original draft, and writing - review and editing. The author gave final approval and agrees to be accountable for all aspects of the work.

Acknowledgments

The author gratefully acknowledges the investigators and participants who contributed to the publicly available datasets used in this study.

Data Availability Statement

This study is based entirely on the secondary analysis of publicly available datasets, and no new primary data were generated. Oral microbiome summary statistics are available from the 4D-SZ cohort (<https://ftp2.cngb.org/>). Periodontal disease and periodontitis GWAS summary statistics are available from the OpenGWAS project (<https://opengwas.io/>; bbj-a-142), the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>; GCST90018677, GCST90436271, GCST90478260, GCST90478266), and FinnGen (<https://www.finnngen.fi/>). Transcriptomic datasets are available from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) under accession numbers GSE16134, GSE171213, and GSE206621. ANK3 cis-eQTL data are available from the 1KCP project (<https://yanglab.westlake.edu.cn/1kcp/>) and OpenGWAS (eqtl-a-ENSG00000151150). Compound and protein structures are available from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) and the Protein Data Bank (<http://www.rcsb.org/>). The analytical methods are described in full in the Materials and Methods.

References

1. Baker JL, Mark Welch JL, Kauffman KM, McLean JS, He X. The oral microbiome: diversity, biogeography and human health. *Nat Rev Microbiol*. 2024;22:89–104. <https://doi.org/10.1038/s41579-023-00963-6>.
2. Rajasekaran JJ, Krishnamurthy HK, Bosco J, Jayaraman V, Krishna K, Wang T, et al. Oral microbiome: a review of its impact on oral and systemic health. *Microorganisms*. 2024;12:1797. <https://doi.org/10.3390/microorganisms12091797>.
3. Nascimento GG, Alves-Costa S, Romandini M. Burden of severe periodontitis and edentulism in 2021, with projections up to 2050: the Global Burden of Disease 2021 study. *J Periodontal Res*. 2024;59:823–867. <https://doi.org/10.1111/jre.13337>.

4. Shaddox LM, Walker CB. Treating chronic periodontitis: current status, challenges, and future directions. *Clin Cosmet Investig Dent*. 2010;2:79–91. <https://doi.org/10.2147/CCIDE.S7712>.
5. Keijser BJB, van den Broek TJ, van der Wurff M, Dulos R, Jagers F, Kool J, et al. Oral bacterial community dynamics during induction of gingival inflammation. *Front Cell Infect Microbiol*. 2025;15:1597690. <https://doi.org/10.3389/fcimb.2025.1597690>.
6. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. *Nat Rev Methods Primers*. 2022;2:6. <https://doi.org/10.1038/s43586-021-00092-5>.
7. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–2191. <https://doi.org/10.1093/bioinformatics/btq340>.
8. Liu X, Tong X, Zhu J, Tian L, Jie Z, Zou Y, et al. Metagenome-genome-wide association studies reveal human genetic impact on the oral microbiome. *Cell Discov*. 2021;7:117. <https://doi.org/10.1038/s41421-021-00356-0>.
9. Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet*. 2020;52:669–679. <https://doi.org/10.1038/s41588-020-0640-3>.
10. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki MI, Koshiha S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*. 2021;53:1415–1424. <https://doi.org/10.1038/s41588-021-00931-x>.
11. Verma A, Huffman JE, Rodriguez A, Conery M, Liu M, Ho YL, et al. Diversity and scale: genetic architecture of 2068 traits in the VA Million Veteran Program. *Science*. 2024;385:eadj1182. <https://doi.org/10.1126/science.adj1182>.
12. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613:508–518. <https://doi.org/10.1038/s41586-022-05473-8>.
13. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50:1335–1341. <https://doi.org/10.1038/s41588-018-0184-y>.
14. Papapanou PN, Behle JH, Kebschull M, Celenti R, Wolf DL, Handfield M, et al. Subgingival bacterial colonization profiles correlate with gingival tissue gene expression. *BMC Microbiol*. 2009;9:221. <https://doi.org/10.1186/1471-2180-9-221>.
15. Chen Y, Wang H, Yang Q, Zhao W, Chen Y, Ni Q, et al. Single-cell RNA landscape of the osteoimmunology microenvironment in periodontitis. *Theranostics*. 2022;12:1074–1096. <https://doi.org/10.7150/thno.65694>.
16. Caetano AJ, Redhead Y, Karim F, Dhami P, Kannambath S, Nuamah R, et al. Spatially resolved transcriptomics reveals pro-inflammatory fibroblast involved in lymphocyte recruitment through CXCL8 and CXCL10. *eLife*. 2023;12:e81525. <https://doi.org/10.7554/eLife.81525>.
17. Wang Y, Duan Z, Chen D, Shi D, Ding Y, Wang Z, et al. The 1000 Chinese Pangenome empowers medical and population genetics. *Nature*. 2026;654:121–130. <https://doi.org/10.1038/s41586-026->

18. Vösa U, Claringbould A, Westra HJ, Bonder MJ, Deelen P, Zeng B, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet.* 2021;53:1300–1310. <https://doi.org/10.1038/s41588-021-00913-z>.
19. Yang L, Ng YE, Sun H, Li Y, Chini LCS, LeBrasseur NK, et al. Single-cell Mayo Map (scMayoMap): an easy-to-use tool for cell type annotation in single-cell RNA-sequencing data analysis. *BMC Biol.* 2023;21:223. <https://doi.org/10.1186/s12915-023-01728-6>.
20. Sun D, Guan X, Moran AE, Wu LY, Qian DZ, Schedin P, et al. Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. *Nat Biotechnol.* 2022;40:527–538. <https://doi.org/10.1038/s41587-021-01091-3>.
21. Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol.* 2022;40:517–526. <https://doi.org/10.1038/s41587-021-00830-w>.
22. Song L, Chen W, Hou J, Guo M, Yang J. Spatially resolved mapping of cells associated with human complex traits. *Nature.* 2025;641:932–941. <https://doi.org/10.1038/s41586-025-08757-x>.
23. Dai W, Ye Y, Mao B, Tang X, Cui S, Zhao J, et al. Microbiological and metabolomic analysis of biomarkers for grades A and B in stage II periodontitis. *Inflammation.* 2025;48:3206–3218. <https://doi.org/10.1007/s10753-025-02260-1>.
24. Seo SH, Lee JW, Oh S, Hong JS, Lee BS, Kwon SJ, et al. Machine learning reveals microbiome differences by periodontitis severity. *PLoS One.* 2026;21:e0349686. <https://doi.org/10.1371/journal.pone.0349686>.
25. Shu Y, Upara C, Ding Q, Zhu M, Zeng E, Banas JA, et al. Spent culture supernatant of *Streptococcus gordonii* mitigates inflammation of human periodontal cells and inhibits proliferation of pathogenic oral microbes. *J Periodontol.* 2023;94:575–585. <https://doi.org/10.1002/JPER.22-0333>.
26. Lenartova M, Tesinska B, Janatova T, Hrebicek O, Mysak J, Janata J, et al. The oral microbiome in periodontal health. *Front Cell Infect Microbiol.* 2021;11:629723. <https://doi.org/10.3389/fcimb.2021.629723>.
27. Song B, Xian W, Sun Y, Gou L, Guo Q, Zhou X, et al. *Akkermansia muciniphila* inhibited the periodontitis caused by *Fusobacterium nucleatum*. *npj Biofilms Microbiomes.* 2023;9:49. <https://doi.org/10.1038/s41522-023-00417-0>.
28. Belstrøm D, Constancias F, Markvart M, Sikora M, Sørensen CE, Givskov M. Transcriptional activity of predominant *Streptococcus* species at multiple oral sites associate with periodontal status. *Front Cell Infect Microbiol.* 2021;11:752664. <https://doi.org/10.3389/fcimb.2021.752664>.
29. Karched M, Bhardwaj RG, Qudeimat M, Al-Khabbaz A, Ellepola A. Proteomic analysis of the periodontal pathogen *Prevotella intermedia* secretomes in biofilm and planktonic lifestyles. *Sci Rep.* 2022;12:5636. <https://doi.org/10.1038/s41598-022-09085-0>.
30. Takeuchi H, Nakamura E, Yamaga S, Amano A. *Porphyromonas gingivalis* infection induces lipopolysaccharide and peptidoglycan penetration through gingival epithelium. *Front Oral Health.*

- 2022;3:845002. <https://doi.org/10.3389/froh.2022.845002>.
31. Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, et al. Periodontitis: consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J Clin Periodontol*. 2018;45(Suppl 20):S162–S170. <https://doi.org/10.1111/jcpe.12946>.
32. Zhao Z, Chen L, Zhu S, Yu H, Chen Y, Song J, et al. Periodontal ligament stem cells in tissue remodeling: from mechanical forces to inflammatory signals. *Stem Cell Res Ther*. 2025;16:653. <https://doi.org/10.1186/s13287-025-04777-6>.
33. Chen Q, Liu X, Wang D, Zheng J, Chen L, Xie Q, et al. Periodontal inflammation-triggered by periodontal ligament stem cell pyroptosis exacerbates periodontitis. *Front Cell Dev Biol*. 2021;9:663037. <https://doi.org/10.3389/fcell.2021.663037>.
34. Zhang W, Jia L, Zhao B, Xiong Y, Wang YN, Liang J, et al. Quercetin reverses TNF- α induced osteogenic damage to human periodontal ligament stem cells by suppressing the NF- κ B/NLRP3 inflammasome pathway. *Int J Mol Med*. 2021;47:39. <https://doi.org/10.3892/ijmm.2021.4872>.

Figures

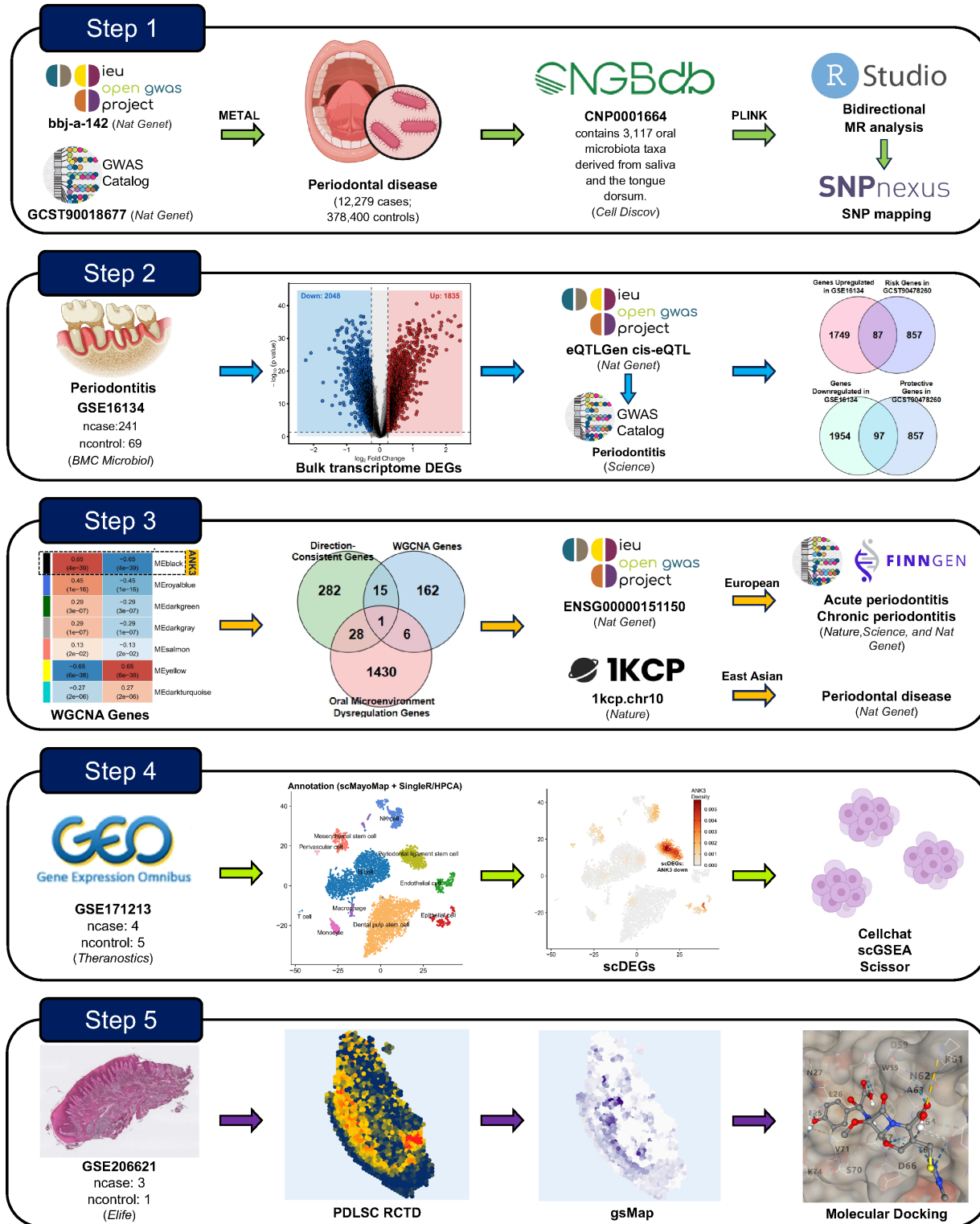


Figure 1

The flowchart of the study.

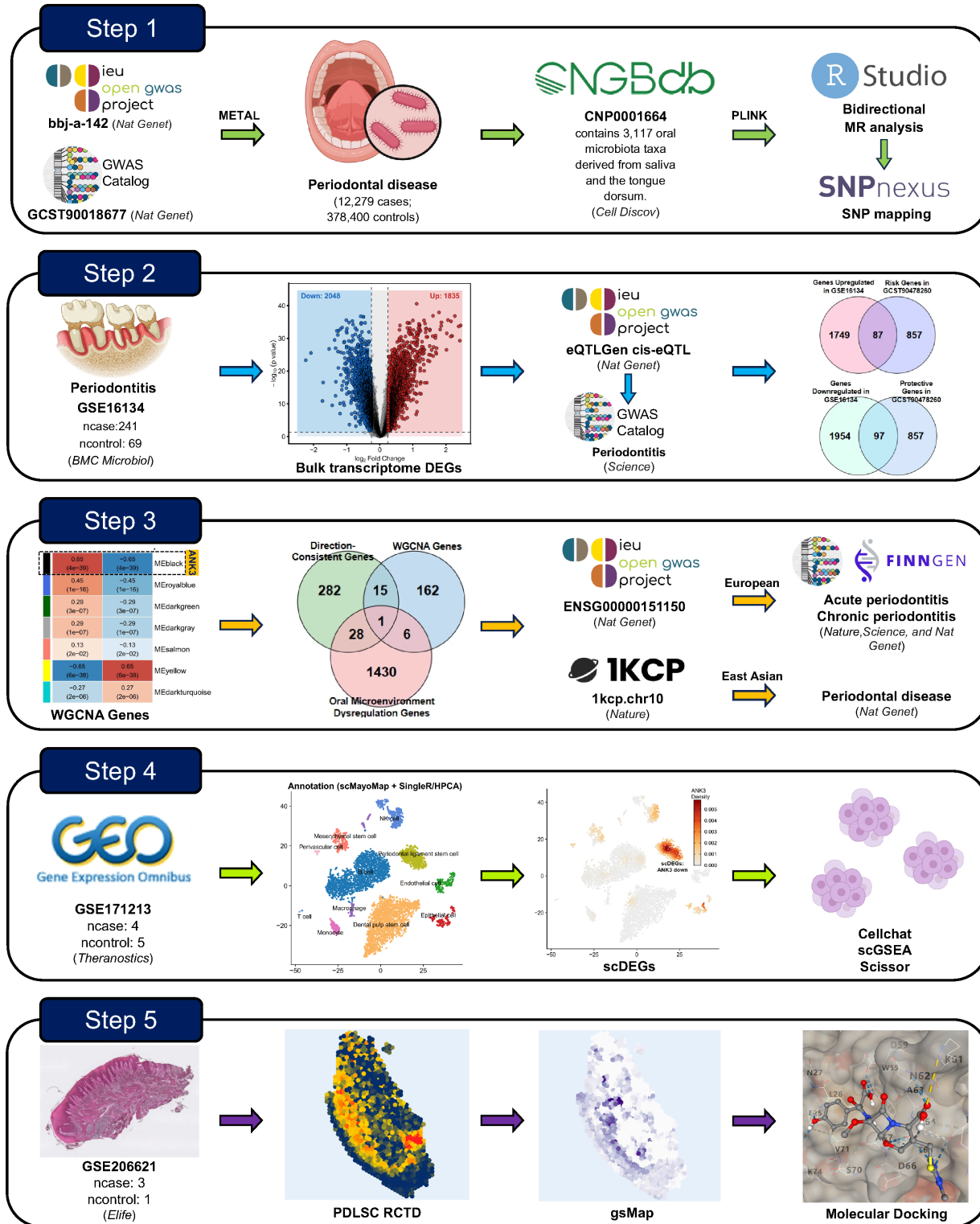


Figure 1

The flowchart of the study.

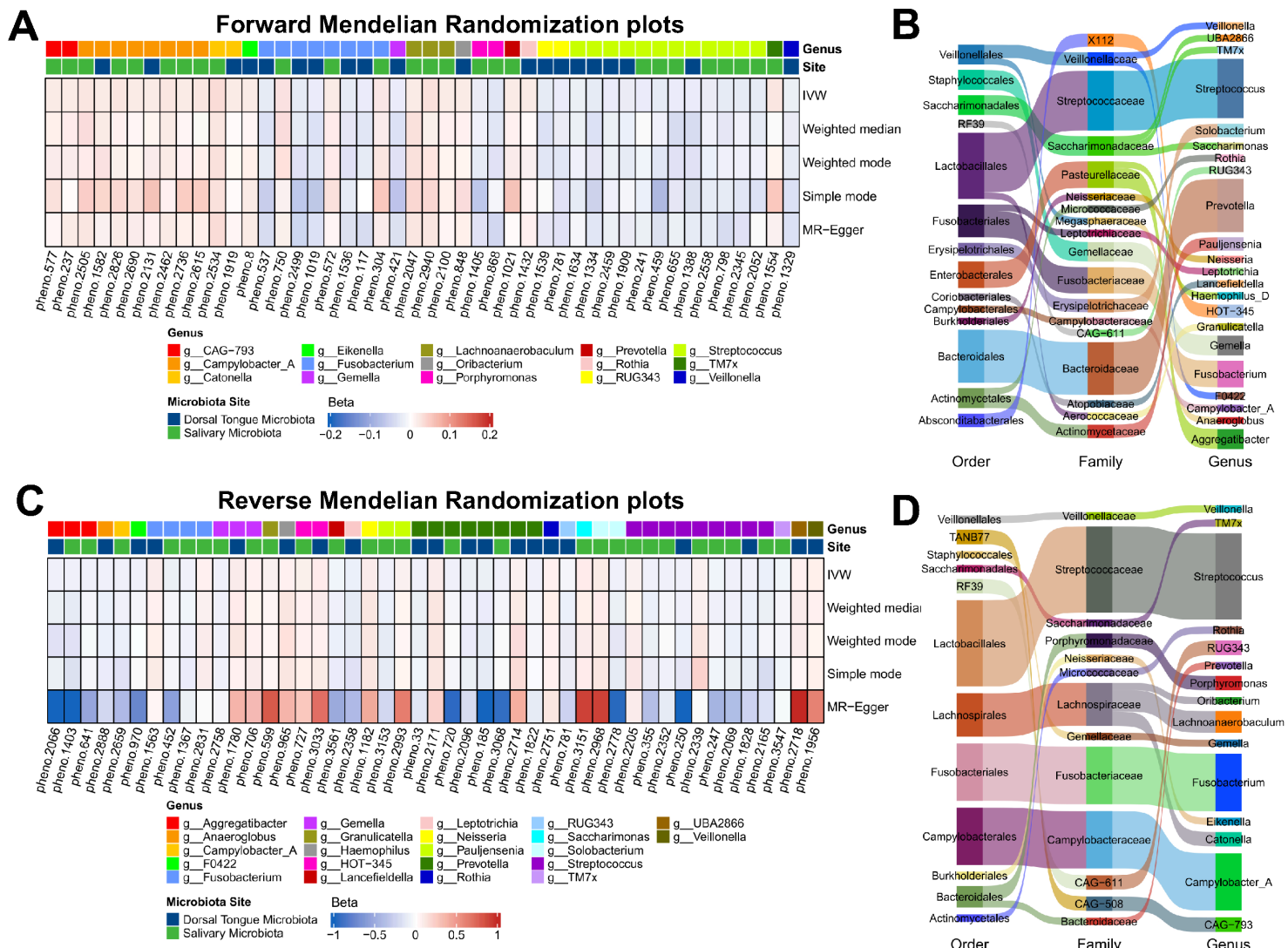


Figure 2

Bidirectional MR analysis of the causal associations between oral microbial taxa and periodontal disease.

(A, C) MR-estimated effects from the forward (A) and reverse (C) analyses, assessing the effects of dorsal tongue and salivary microbial taxa on periodontal disease risk and the effects of periodontal disease on these microbial taxa, respectively. (B, D) Sankey plots showing the taxonomic classification of microbial taxa identified in the forward (B) and reverse (D) MR analyses at the order, family, and genus levels.

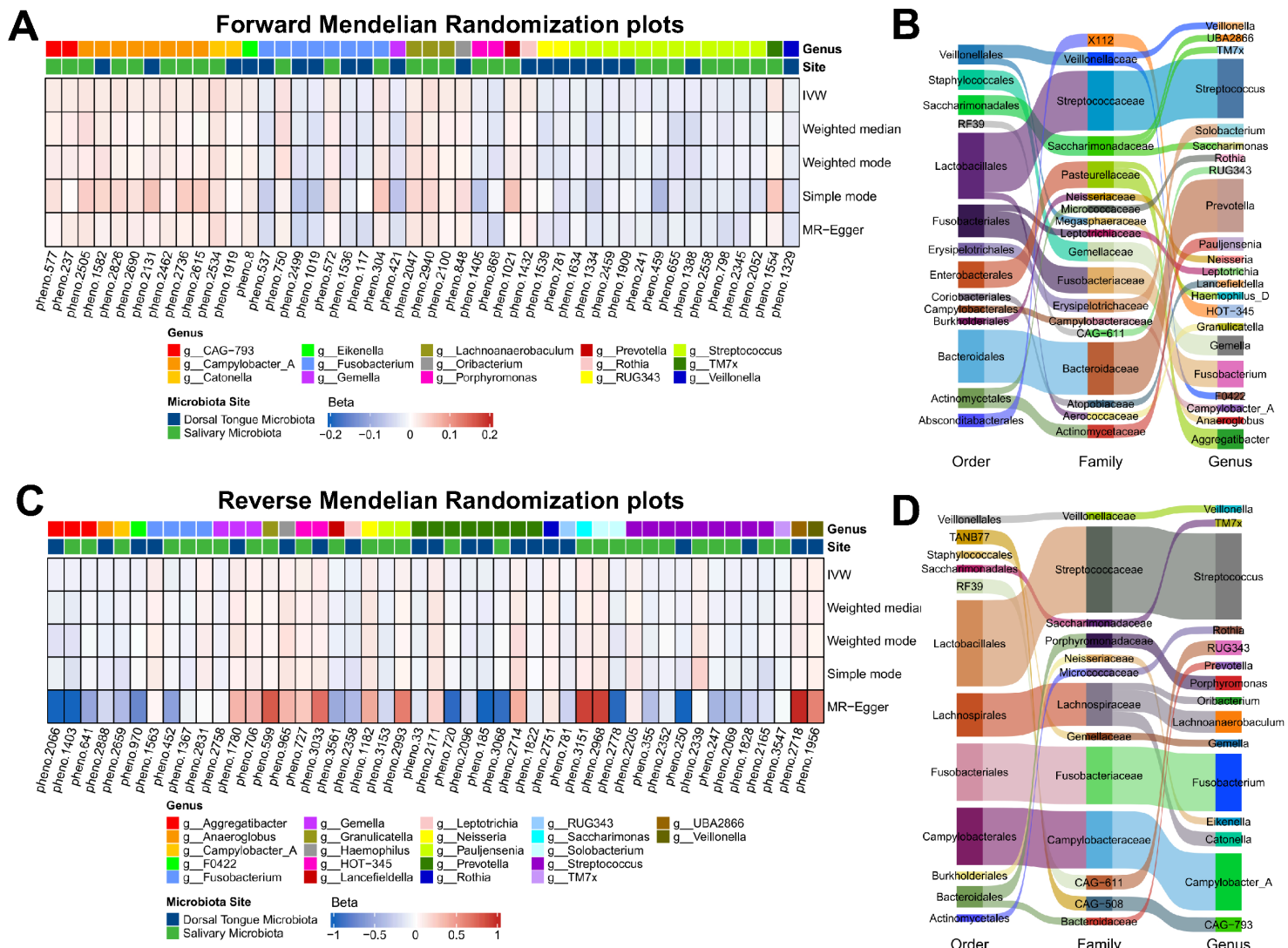


Figure 2

Bidirectional MR analysis of the causal associations between oral microbial taxa and periodontal disease.

(A, C) MR-estimated effects from the forward (A) and reverse (C) analyses, assessing the effects of dorsal tongue and salivary microbial taxa on periodontal disease risk and the effects of periodontal disease on these microbial taxa, respectively. (B, D) Sankey plots showing the taxonomic classification of microbial taxa identified in the forward (B) and reverse (D) MR analyses at the order, family, and genus levels.

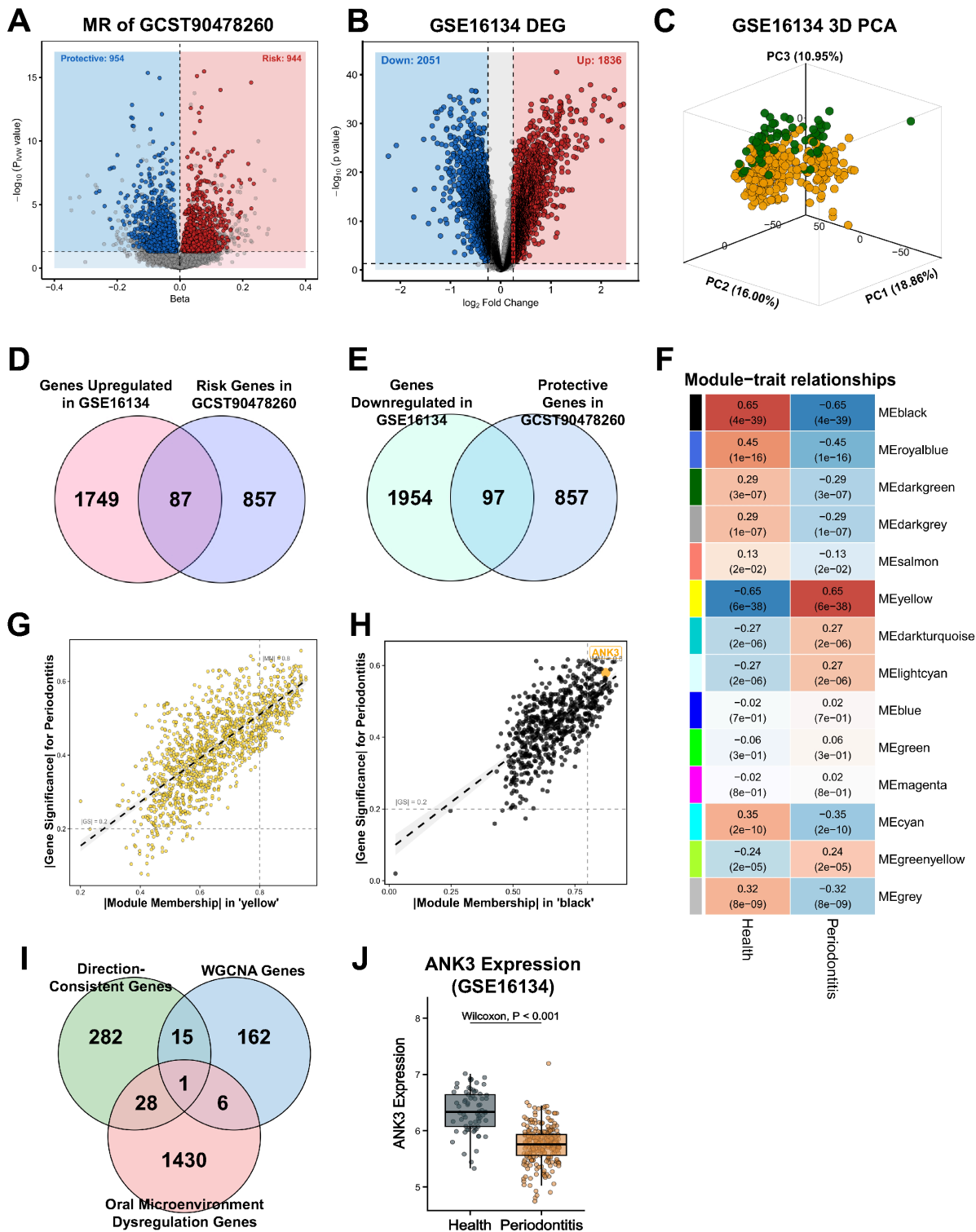


Figure 3

Integrative screening and prioritization of *ANK3* as a candidate gene for periodontitis.

(A) Volcano plot summarizing the MR results obtained using cis-eQTLs from eQTLGen as IVs and periodontitis data from GCST90478260 as the outcome. (B, C) Volcano plot of DEGs in GSE16134 (B) and 3D PCA plot of the same dataset (C). (D, E) Venn diagrams showing directionally concordant genes:

genes upregulated in GSE16134 with MR estimates indicating increased periodontitis risk (D), and genes downregulated in GSE16134 with MR estimates indicating a protective effect against periodontitis (E). (F–H) WGCNA of GSE16134, showing module–trait relationships (F) and GS–MM scatter plots for the black (G) and yellow (H) modules. Candidate genes were selected using the thresholds $|MM| > 0.80$ and $|GS| > 0.20$, with *ANK3* highlighted. (I) Venn diagram integrating WGCNA-identified genes, directionally concordant genes, and genes associated with periodontal microenvironment dysregulation. *ANK3* was identified at the intersection of the three gene sets. (J) Comparison of *ANK3* expression between the HC and periodontitis groups in GSE16134.

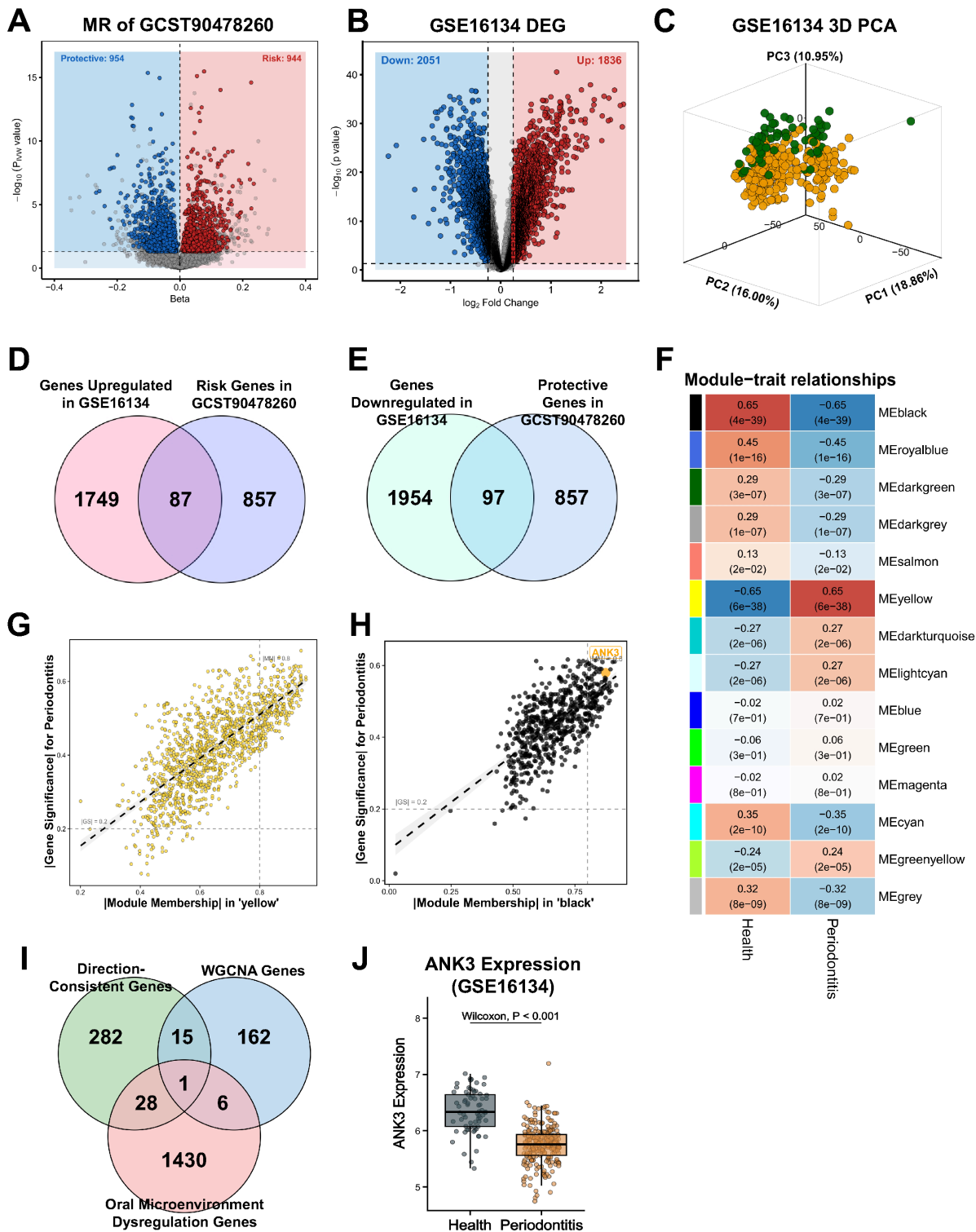


Figure 3

Integrative screening and prioritization of *ANK3* as a candidate gene for periodontitis.

(A) Volcano plot summarizing the MR results obtained using cis-eQTLs from eQTLGen as IVs and periodontitis data from GCST90478260 as the outcome. (B, C) Volcano plot of DEGs in GSE16134 (B) and 3D PCA plot of the same dataset (C). (D, E) Venn diagrams showing directionally concordant genes:

genes upregulated in GSE16134 with MR estimates indicating increased periodontitis risk (D), and genes downregulated in GSE16134 with MR estimates indicating a protective effect against periodontitis (E). (F–H) WGCNA of GSE16134, showing module–trait relationships (F) and GS–MM scatter plots for the black (G) and yellow (H) modules. Candidate genes were selected using the thresholds $|MM| > 0.80$ and $|GS| > 0.20$, with *ANK3* highlighted. (I) Venn diagram integrating WGCNA-identified genes, directionally concordant genes, and genes associated with periodontal microenvironment dysregulation. *ANK3* was identified at the intersection of the three gene sets. (J) Comparison of *ANK3* expression between the HC and periodontitis groups in GSE16134.

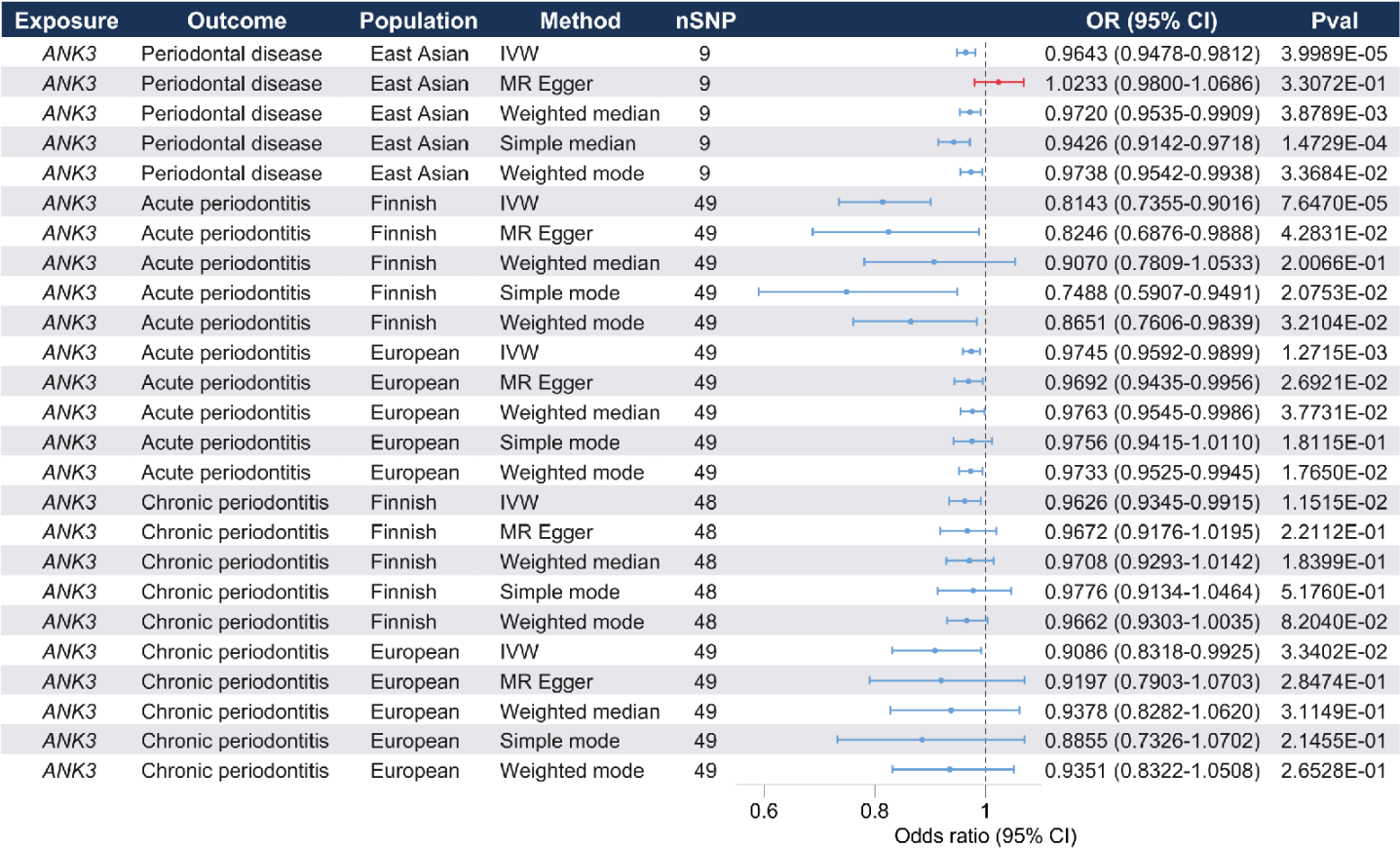


Figure 4

MR analyses of periodontitis-related phenotypes using cis-eQTLs as IVs. The outcomes were periodontal disease from an East Asian–ancestry dataset and acute periodontitis and chronic periodontitis from European-ancestry datasets.

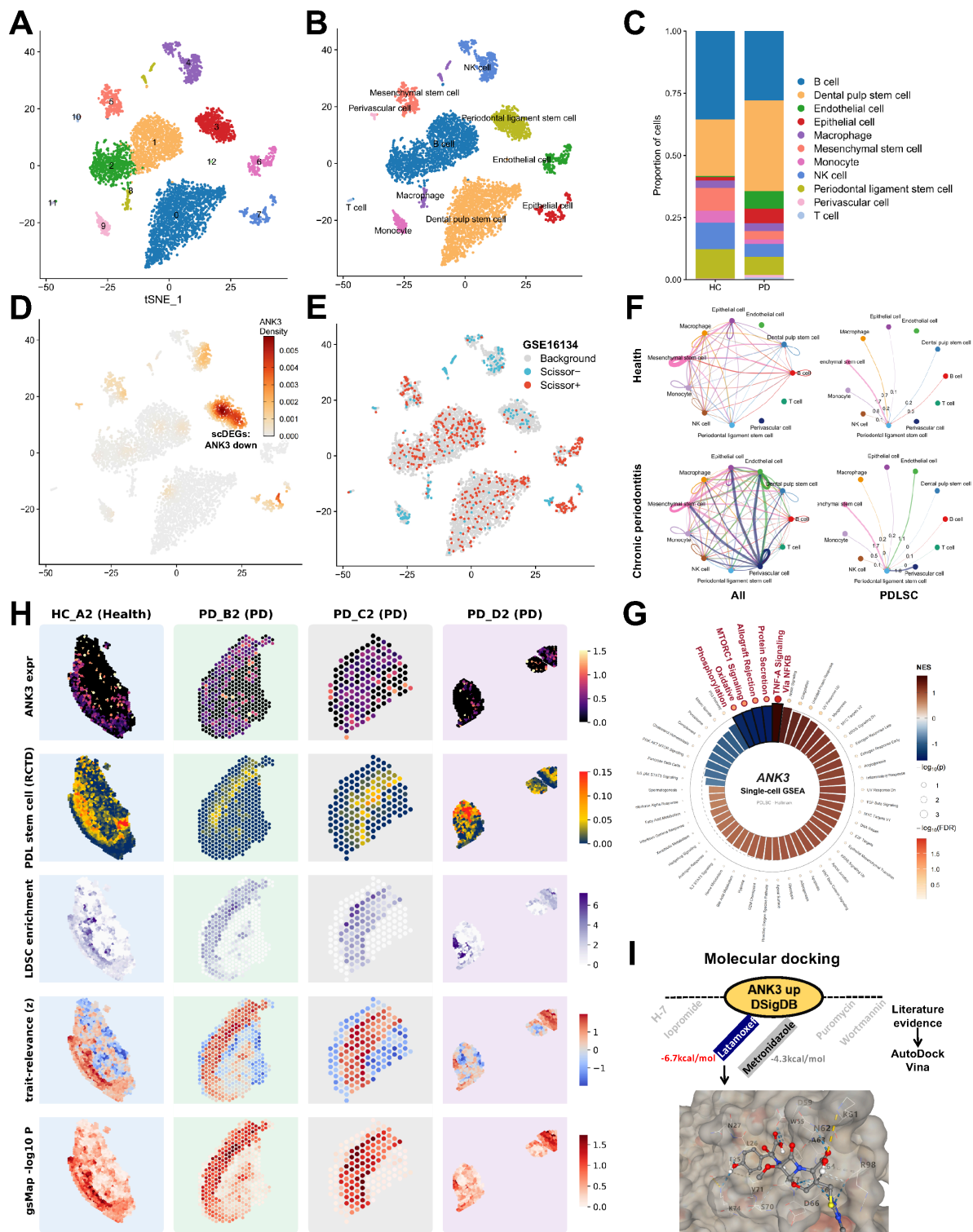


Figure 5

Integrated single-cell, spatial transcriptomic, and molecular docking analyses. (A, B) t-SNE visualization of the integrated single-cell atlas, colored by annotated cluster (A) and cell type (B). (C) t-SNE visualization of *ANK3* expression, with scDEG-identified cells showing *ANK3* downregulation highlighted. (D) Proportions of individual cell types in the healthy control (HC) and periodontitis (PD) groups. (E) Scissor projection of the bulk phenotype onto the single-cell landscape using the GSE16134 dataset. (F)

CellChat-inferred intercellular communication networks in the HC and PD groups, shown across all cell types and for PDLSCs. (G) Circos plot summarizing the scGSEA results for *ANK3* in PDLSCs. (H) Spatial transcriptomic maps of four samples showing *ANK3* expression, the RCTD-inferred spatial distribution of PDLSCs, and gsMap-predicted regions associated with genetic susceptibility. (I) Molecular docking of latamoxef to the *ANK3* protein.

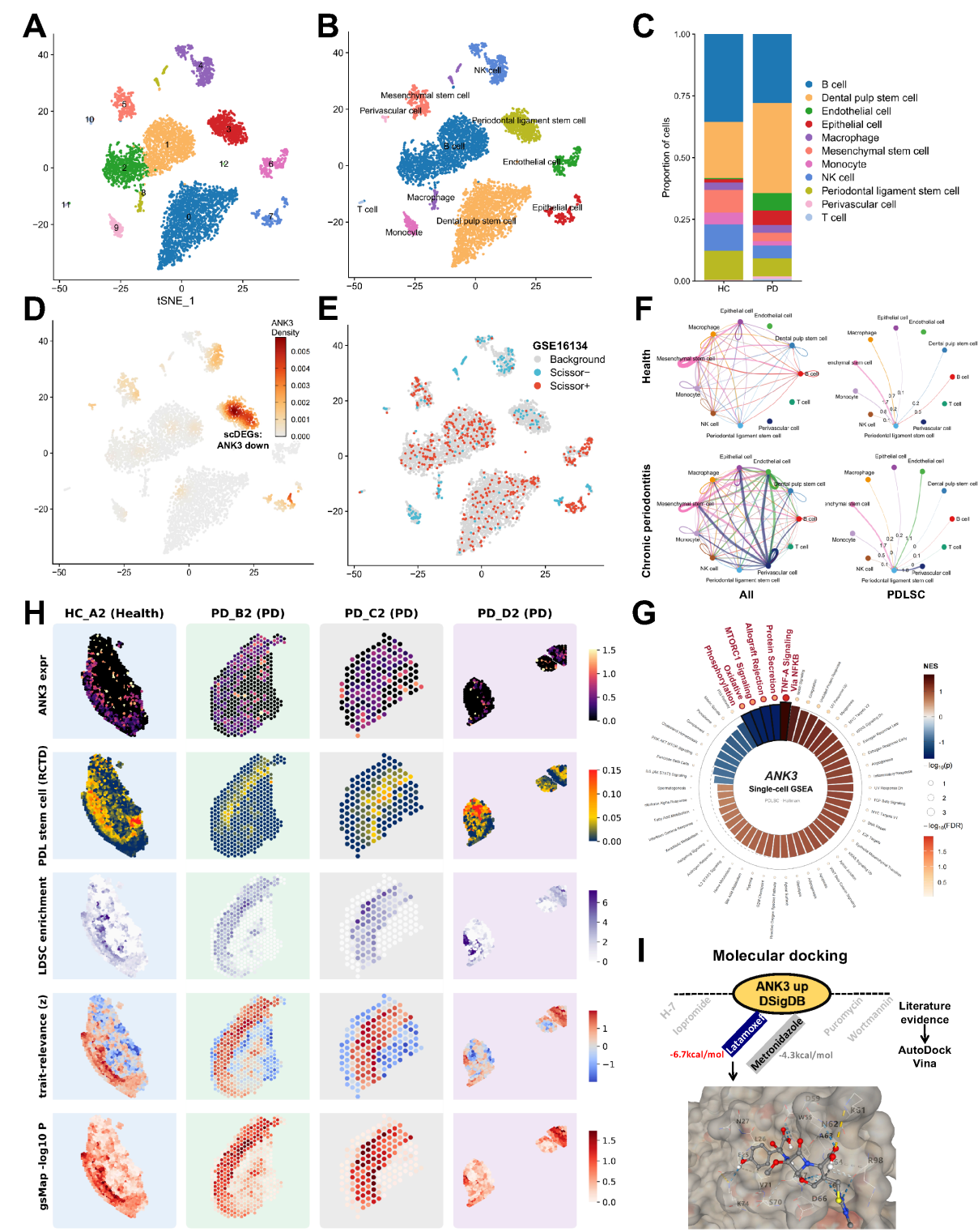


Figure 5

Integrated single-cell, spatial transcriptomic, and molecular docking analyses. (A, B) t-SNE visualization of the integrated single-cell atlas, colored by annotated cluster (A) and cell type (B). (C) t-SNE visualization of *ANK3* expression, with scDEG-identified cells showing *ANK3* downregulation highlighted. (D) Proportions of individual cell types in the healthy control (HC) and periodontitis (PD) groups. (E) Scissor projection of the bulk phenotype onto the single-cell landscape using the GSE16134 dataset. (F) CellChat-inferred intercellular communication networks in the HC and PD groups, shown across all cell types and for PDLSCs. (G) Circos plot summarizing the scGSEA results for *ANK3* in PDLSCs. (H) Spatial transcriptomic maps of four samples showing *ANK3* expression, the RCTD-inferred spatial distribution of PDLSCs, and gsMap-predicted regions associated with genetic susceptibility. (I) Molecular docking of latamoxef to the *ANK3* protein.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable3.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [SupplementaryTable5.xlsx](#)
- [SupplementaryTable6.xlsx](#)
- [SupplementaryTable6.xlsx](#)