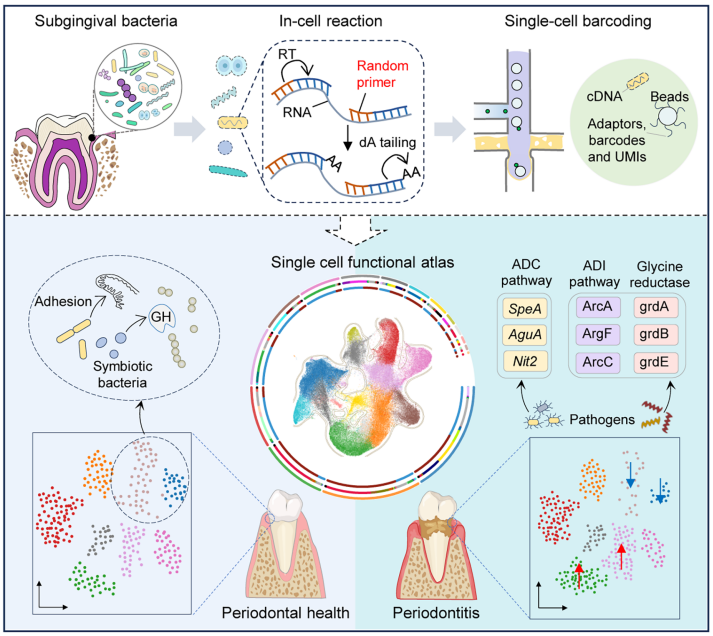
Resource

Cell Host & Microbe

Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs

Graphical abstract



Highlights

- A refined scRNA-seq framework for subgingival bacteria
- A single-cell functional atlas of 133,458 bacterial cells from 285 species
- *N. elongata* maintains adhesion-related transcriptional programs in periodontal health
- *P. intermedia*, *P. gingivalis*, and *T. denticola* show functional heterogeneity and plasticity

Authors

Liguo Ding, Lu Song, Yibing Han, ..., Qianming Chen, Pei-Hui Ding, Yongcheng Wang

Correspondence

phding@zju.edu.cn (P.-H.D.), yongcheng@zju.edu.cn (Y.W.)

In brief

Subgingival microbial communities drive periodontitis, yet their single-cell-specific functions remain poorly defined. Using an optimized mscRNA-seq workflow, Ding et al. map a single-cell transcriptional atlas of subgingival bacteria, revealing loss of health-associated subpopulations and species-specific amino acid-related transcriptional programs in periodontitis.

Ding et al., 2026, Cell Host & Microbe *34*, 1–18
March 11, 2026
https://doi.org/10.1016/j.chom.2026.01.017

CellPress

Cell Host & Microbe

CellPress

Resource

Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs

Liguo Ding,[1,4] Lu Song,[2,4] Yibing Han,[2,4] Jianing Wang,[3,4] Yuexiao Lyu,[1] Mengdi Song,[1] Xiaoyue Li,[1] Wentao He,[2] Qianming Chen,[2] Pei-Hui Ding,[2,*] and Yongcheng Wang[1,5,*]

[1]Department of Laboratory Medicine of the First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, Hangzhou, China

[2]Stomatology Hospital, School of Stomatology, Zhejiang University School of Medicine, Zhejiang Provincial Clinical Research Center for Oral Diseases, Key Laboratory of Oral Biomedical Research of Zhejiang Province, Hangzhou, China

[3]M20 Genomics, Hangzhou, China

[4]These authors contributed equally

[5]Lead contact

*Correspondence: phding@zju.edu.cn (P.-H.D.), yongcheng@zju.edu.cn (Y.W.)

https://doi.org/10.1016/j.chom.2026.01.017

SUMMARY

Periodontitis is a chronic inflammatory disease driven by subgingival dysbiosis. Despite progress in characterizing microbial diversity, functional heterogeneity at the single-cell level remains poorly understood, mainly due to challenges from low microbial biomass and host contamination. Here, we present a single-cell RNA sequencing framework for profiling subgingival bacteria, enabling high-resolution analysis of these communities in health and periodontitis. Using 16 subgingival samples, we generated an atlas spanning 133,458 cells across 285 species, including 57 core active species grouped into eight functional clusters. Health-associated subpopulations specializing in adhesion and polysaccharide degradation decline in periodontitis, coinciding with hypoxia and elevated amino acid availability. The keystone pathogens *T. denticola* and *P. gingivalis* display species-specific tendencies in amino acid metabolism-related transcriptional profiles, while *P. intermedia* harbors a proteolytic subpopulation enriched in periodontitis. These insights deepen our understanding of periodontitis pathogenesis and inform precision diagnostics and therapeutic strategies.

INTRODUCTION

Periodontitis (Pd) is a chronic inflammatory disease affecting 20%–50% of adults globally and contributing significantly to tooth loss and systemic conditions, including diabetes and cardiovascular disease.[1–5] Central to its pathogenesis is the subgingival microbiome, a dense polymicrobial community that orchestrates inflammation and systemic ecological disruption.[6,7] Although microbiological investigations have spanned decades,[8] our understanding of species-specific functions, transcriptional plasticity, and ecological interactions within this community remains constrained, particularly at the single-cell resolution.

Traditional techniques, such as checkerboard hybridization and 16S rRNA sequencing, have been instrumental in identifying hallmark pathogens, including the "red complex," and mapping subgingival community diversity.[9–11] Metagenomic studies extended this by resolving community-wide gene content.[12–14] Landmark metatranscriptomic analyses[15–18] of Pd further revealed community-level transcriptional shifts, including differential expression of virulence factors and metabolic pathways. However, these bulk metatranscriptomics approaches inherently average signals across diverse taxa or higher taxonomic levels, preventing the assignment of functional activities to individual species or subpopulations *in situ*.[19] Functional interpretation is further constrained by gene redundancy, which obscures the ecological roles of specific taxa.

Microbial single-cell RNA sequencing (mscRNA-seq), highlighted by *Nature* as a "Technology to Watch,"[20] enables high-resolution analysis of functional heterogeneity among individual bacterial cells.[21] Initial applications focused on synthetic consortia composed of well-characterized species.[22–25] More recently, mscRNA-seq has been applied to complex natural ecosystems, where it pinpointed keystone subpopulations executing specialized metabolic functions. In the human gut, for instance, mscRNA-seq uncovered intra-species adaptive heterogeneity within *Phascolarctobacterium succinatutens*[26] and identified a *Mitsuokella funiformis* subpopulation responsible for phytate degradation.[27] In the rumen, it delineated metabolic niche trajectories driven by biofilm-associated gene expression in

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

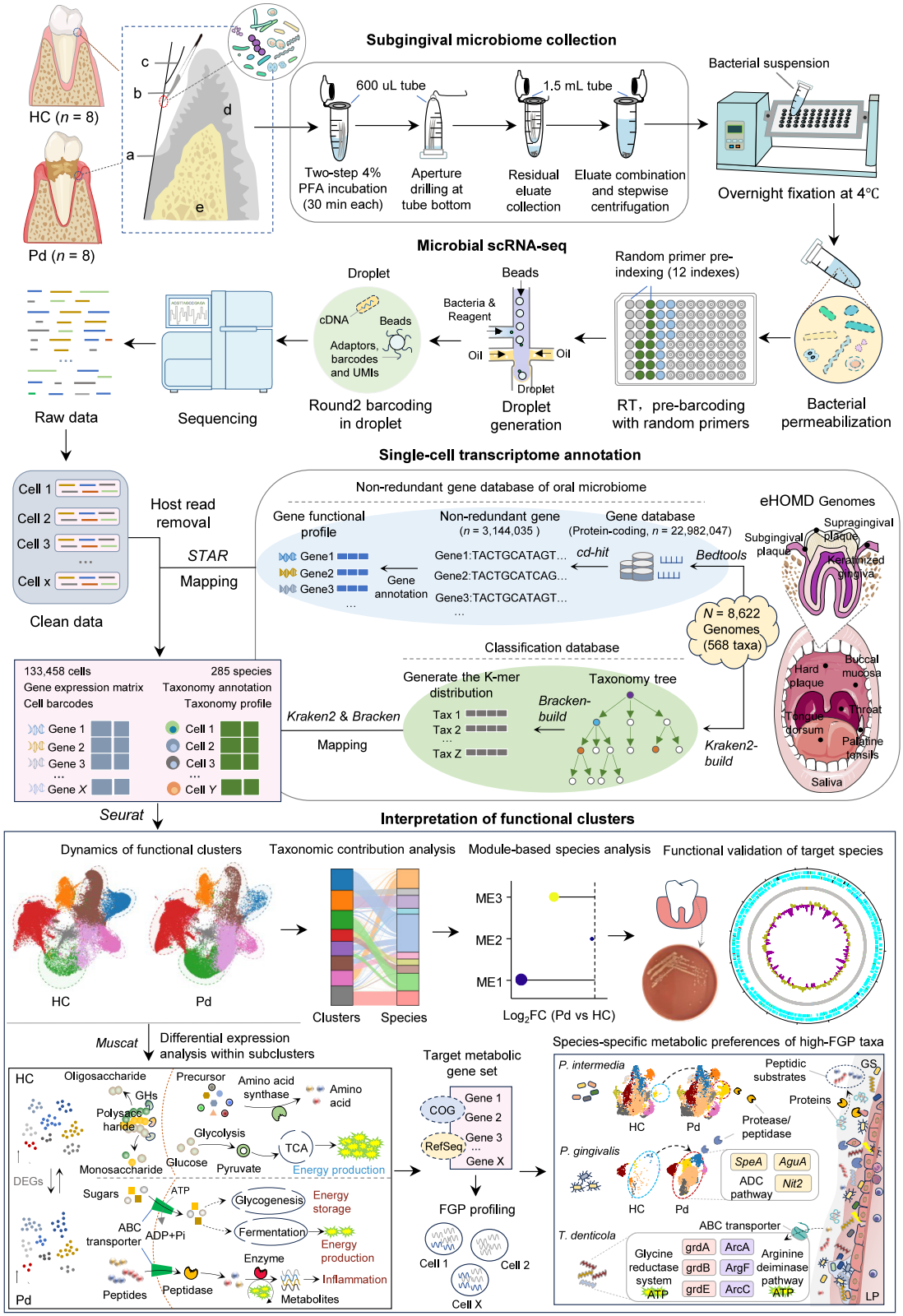


Figure 1. Schematic overview of the mscRNA-seq workflow for subgingival bacteria

The pipeline encompasses subgingival microbiome sampling, mscRNA-seq, single-cell transcriptome annotation, and interpretation of functional clusters. The interpretation of functional clusters involves two key analyses: (1) identifying functions significantly suppressed in Pd, determining the primary species executing

(legend continued on next page)

Bacteroides succiniciproducens.[28] However, current microbial mscRNA-seq workflows demand exceptionally high microbial inputs ($\sim 10^9$ cells) and minimal host contamination—conditions that are rarely met in subgingival plaque samples.[14] Furthermore, conventional plaque sampling often fails to adequately remove host cells and may also introduce bacterial aggregation and cryopreservation artifacts, further compromising data quality and interpretability in single-cell workflows.

To overcome these challenges, we developed an optimized mscRNA-seq pipeline applicable to host-associated environments and applied it to profile the subgingival bacteria at single-cell resolution across both health and disease states. This platform integrates improved host-cell depletion, enhanced bacteria recovery and dispersion, and a re-annotated, non-redundant gene catalog derived from the expanded Human Oral Microbiome Database (HOMD), enabling accurate transcript quantification across diverse taxa. In parallel, we constructed an rRNA-retaining genome reference to resolve taxonomic identity at single-cell level. Applying this framework to 16 subgingival samples, we captured 133,458 individual bacterial cells representing 285 species. Our findings uncover transcriptional heterogeneity and functional plasticity across diverse lineages, establishing a scalable foundation for single-cell bacterial analysis in complex clinical ecosystems.

RESULTS

A high-fidelity scRNA-seq platform for subgingival bacteria

Conventional methods for collecting subgingival plaque, such as paper-strip adsorption, are commonly used to retrieve unattached bacteria from periodontal pockets.[29] However, these approaches often fail to sufficiently remove host-derived contaminants. Moreover, the high-speed centrifugation and snap-freezing steps typically involved may promote bacterial aggregation and RNA degradation, complicating downstream applications of mscRNA-seq that require intact, well-isolated bacterial cells. To address these challenges, we refined the sample-processing pipeline to meet the stringent demands of mscRNA-seq. Subgingival samples were immediately incubated in 4% paraformaldehyde (PFA) to stabilize RNA and prevent cellular damage typically caused by cryopreservation. Gentle pipetting combined with low-speed centrifugation efficiently recovered bacterial cells while minimizing aggregation, achieving yields comparable to conventional protocols (Figure S1A). A stepwise centrifugation with gradually increasing speeds further enhanced sample purity, yielding higher-quality bacterial fractions (Figure S1B). Final overnight fixation in 4% PFA ensured sufficient cross-linking for downstream processing (Figure 1).

The bacterial suspensions were subjected to permeabilization, first with a mild detergent and then with a lysozyme mixture, with repeated pipetting throughout the process to enhance pore formation in the cell wall and membrane and to facilitate the disruption of residual aggregates (Figure S1B). Reverse transcription (RT) with random primers was then performed for pre-barcoding (primer sequences listed in Table S1). Microscopy confirmed that the majority of cells existed as well-dispersed single bacteria prior to droplet generation (Figure S2), allowing their subsequent encapsulation, barcoding, and high-throughput sequencing (Figure 1).

To improve recovery, we extended the centrifugation time during wash steps from the 5 min typically used in gut mscRNA-seq[26] to 15 min, leading to a substantial increase in bacterial yields (Figure S1C). To mitigate aggregation caused by prolonged centrifugation, we replaced vortexing with over 150 strokes of gentle pipetting after each cycle, effectively reducing clumping (Figure S1D). Although ultrasound-based methods have been proposed to reduce aggregation,[30] they were unsuitable for low-abundance microbial samples due to the risk of bacterial rupture (Figure S1D).

High cell viability is critical for mscRNA-seq. To assess the impact of our sampling procedure on cell integrity, subgingival plaque samples were immediately processed in PBS on ice and enriched via stepwise centrifugation. Flow cytometry revealed that ~92% of bacterial cells in both cohorts were PI-negative, suggesting generally high membrane integrity and indicating that our sampling and enrichment procedures largely preserved bacterial viability suitable for single-cell RNA-seq (scRNA-seq) (Figures S1E and S1F).

Together, these methodological refinements markedly improve bacterial recovery, purity, and integrity, providing a robust foundation for high-resolution scRNA-seq of subgingival microbiota.

Construction of a single-cell functional atlas of the subgingival microbiome

Using the optimized protocol, we generated 14.31 billion clean reads from subgingival plaque samples collected from 8 healthy controls (HCs) and 8 Pd patients (demographic and sequencing details in Tables S2 and S3). Alignment to the human genome revealed a median host-read contamination of only 2.69%, representing a significant improvement over the 33.35% observed in our previous metagenomic dataset[31] (Mendeley Data: https://doi.org/10.17632/kdktdhrb6b.1). To enable accurate functional interpretation of single-cell transcriptomes, we constructed a non-redundant gene catalog based on the HOMD,[32] encompassing 568 species, 8,622 genomes, and 22,982,047 protein-coding genes from various oral niches (Figure 1). Deduplication reduced this dataset to 3,144,035 unique protein-coding genes (~13.7%), 74.42% of which were functionally annotated using the clusters of orthologous groups of proteins (COG) database.[33] This represents a substantial improvement over the 30.89% annotation rate in the original catalog, with consistent improvements across all functional categories (Figure S3A). Species- and genome-level coverage analyses showed that our non-redundant gene catalog

these functions, and experimentally validating the findings; (2) analyzing convergent transcriptional responses to Pd by comparing differential gene expression between functional clusters from Pd and control samples. This approach enables us to anchor specific pathways and, subsequently, investigate species- and subpopulation-specific heterogeneity in their expression. a, tooth root; b, paper point; c, tooth crown; d, gingiva; e, alveolar bone; UMIs, unique molecular identifiers.
See also Figures S1 and S2.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

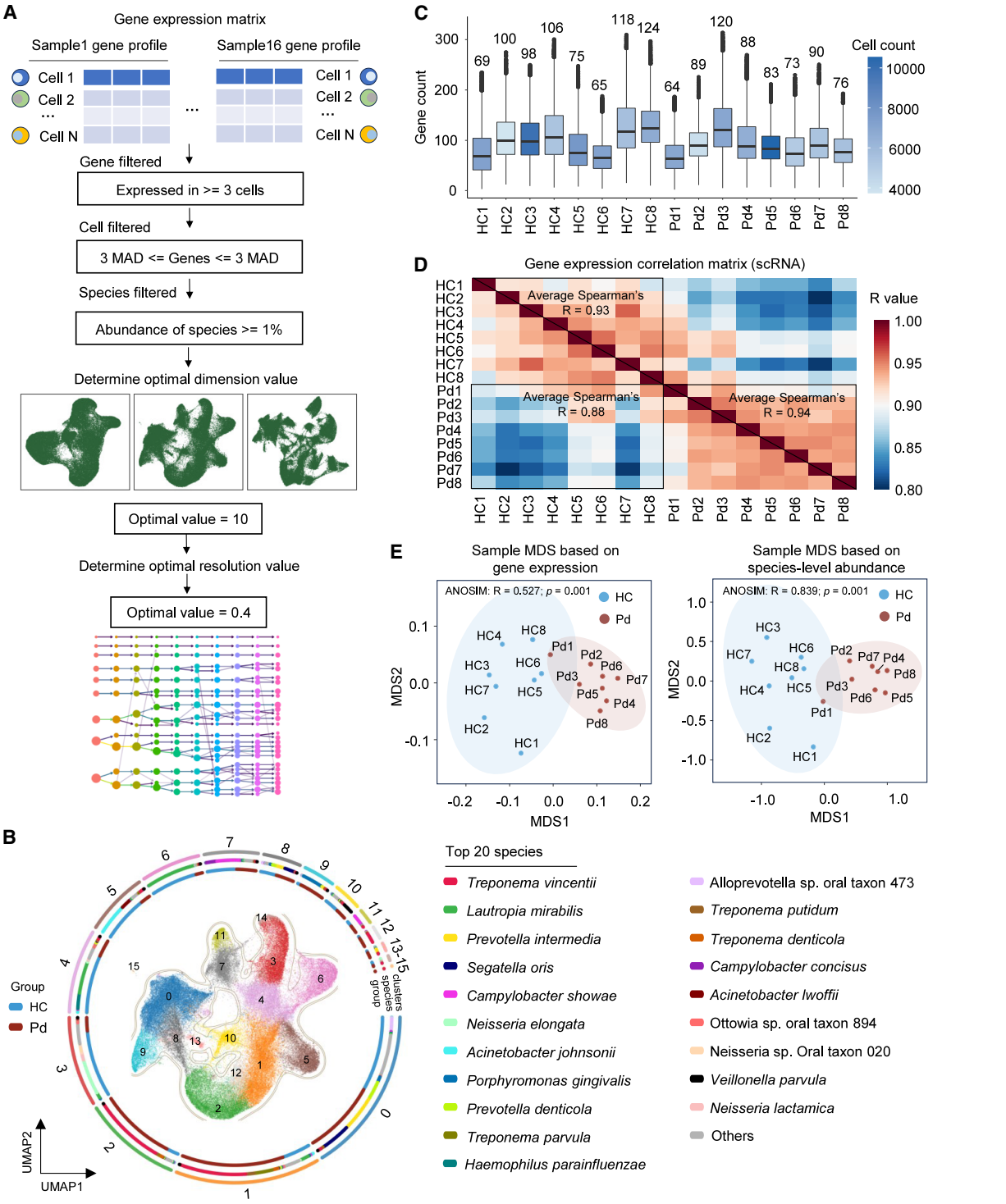


Figure 2. Mapping the functional heterogeneity of the subgingival bacteria at single-cell resolution

(A) Cell and gene filtering steps, along with benchmarking processes.

(B) UMAP visualization of 104,694 high-quality cells from 16 subgingival bacterial samples across HC and Pd groups. Cells are colored by cluster identity, with contour lines delineating cluster boundaries. The circular plot depicts (from outer to inner ring): cluster identity (arc length proportional to cell number), species assignment, and group (HC, blue; Pd, red). Legends for group and species are provided on the left and right, respectively.

(C) Number of annotated unique functional genes identified per sample. Colors correspond to cell counts.

(D) Spearman correlation matrix of pseudobulk gene expression across HC and Pd samples. Correlations were computed from pseudobulk profiles defined as the mean log-normalized gene expression per sample.

(legend continued on next page)

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

CellPress

includes 100% of species and 94.7% of genomes in HOMD. Benchmarking demonstrated improved mapping efficiency, with uniquely mapped reads increasing from 13.8% to 82.6%, while reducing runtime and memory usage (Table S4). This catalog was subsequently used as a reference for mapping and quantifying single-cell transcriptomes.

Given the challenge of residual host contamination, we implemented a host-read depletion step prior to alignment (Figure S3B). To improve taxonomic annotation in mscRNA-seq data, level-by-level annotation was introduced, involving rRNA masking to reduce false-positive assignments[28] (Figure S3C). However, the high functional redundancy within the oral microbiome (~86.3%) suggests that rRNA masking may compromise taxonomic resolution. Therefore, we compared annotations generated using masked versus unmasked rRNA reference databases to evaluate this effect. Notably, unmasked rRNA references showed greater concordance with metagenomic profiles, increasing genus-level overlap from 37% to 47.6% and species-level overlap from 15.2% to 25.3%. Moreover, genus-level correlation coefficients with metagenomic data improved in both HC (from R = 0.51 to R = 0.55) and Pd (from R = 0.41 to R = 0.70) groups (Figure S3D).

Further examination of species assignments using masked references revealed inconsistencies between the two strategies: *Acinetobacter johnsonii* and *Treponema denticola* were identified as the most abundant species in HC and Pd, respectively, whereas BLAST validation revealed that the majority of these reads actually originated from *L. mirabilis* and *T. vincentii* (Figure S3E). Interestingly, most barcoded cells initially annotated as *A. johnsonii* and *T. denticola* under the masked reference were correctly reclassified as *L. mirabilis* and *T. vincentii*, respectively, when using the unmasked reference (Figure S3F). Further analysis confirmed that over 50% of sampled reads accurately mapped to these two species (Figure S3G). Based on these findings, we adopted the unmasked rRNA reference for all downstream taxonomic annotation to ensure maximal accuracy and biological relevance.

To ensure data quality, droplets likely containing cells were identified and filtered using STARsolo's default *soloCellFilter* cutoffs[25,34,35] (Figure S4A), resulting in 133,458 high-confidence cells annotated across 285 species, each containing an average of 2,941 reads and 122 unique functional genes (Table S3). These included key periodontal pathogens from the red complex—*Porphyromonas gingivalis*, *T. denticola*, and *Tannerella forsythia*—a well-established microbial consortium strongly associated with Pd. This complex is typically enriched at active sites, exhibiting deep pockets and bleeding on probing, and directly contributes to tissue destruction. The orange complex, which includes species such as *Fusobacterium nucleatum*, *Prevotella intermedia*, and *Campylobacter rectus*, acts as a critical bridge in biofilm development by modifying the local microenvironment to facilitate establishment of the more pathogenic red complex (Figure S4B). Bacteria within this complex also contribute directly to tissue destruction and host immune activation.[9]

Following quality control, we retained barcodes from core active species (≥1% relative abundance), resulting in a final dataset of 104,694 high-confidence cells. To improve functional clustering, only annotated genes were retained, reducing bias from unannotated sequences. The normalized and batch-corrected dataset, comprising 104,694 cells and capturing an average of 88 unique functional genes per sample, collectively represents 57 species across 27 genera (Figures 2A–2C), with cells from phylogenetically related species tending to cluster together (Figure S4C). Notably, functional gene expression was highly correlated both within (R ≈ 0.93) and between (R ≈ 0.88) groups (Figure 2D), in contrast to much lower correlations based on species composition (within R ≈ 0.45; between R ≈ 0.15; Figure S5A). We then performed multidimensional scaling (MDS) using matrices derived from our mscRNA-seq data, including aggregated gene expression profiles, as well as taxonomic (relative abundance of cells per species or genus) and functional (proportion of cells per Seurat-identified clusters) compositions calculated per sample. These profiles exhibited group-specific separation primarily along the first MDS axis (MDS1; Figure 2E; Figure S5B). Consistently, uniform manifold approximation and projection (UMAP) visualization of sample-level clustering showed a tendency for samples from different groups to occupy distinct regions across clusters (Figure S5C), suggesting that the transcriptional differences observed are broadly representative across samples.

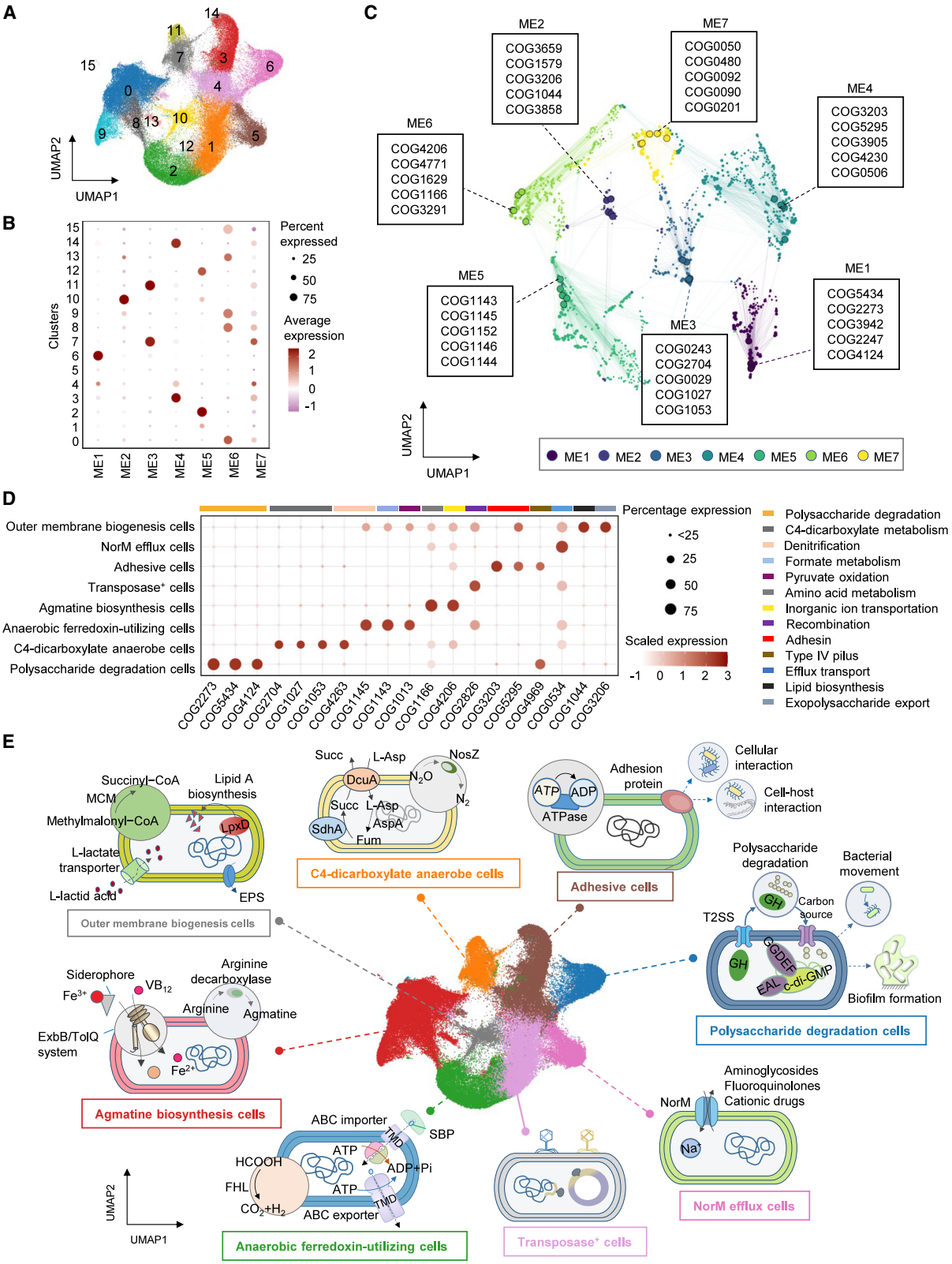
To further assess the representativeness of our data, we performed full-length 16S rRNA-seq on 18 healthy (16S-HC) and 18 Pd (16S-Pd) subgingival plaque samples (demographic details in Table S2). These results, together with a published metagenomic dataset,[31] showed the same trend as our mscRNA-seq data: higher similarity within groups than between groups at the genus or species level (metagenomics: within-group R ≈ 0.47, between-group R ≈ 0.28, Figure S5D; 16S: within-group R ≈ 0.60, between-group R ≈ 0.26, Figure S5E). Moreover, most genera with >1% relative abundance in 16S (~94.7%) were detected by mscRNA-seq, with the shared genera accounting for a median of ~92% of total abundance per sample (Figures S5F and S5G). In one subgingival plaque sample from the 16S dataset, side-by-side microbial mscRNA-seq and 16S rRNA-seq showed a strong correlation for genera >1% abundance ($R = 0.824$, $p = 0.0034$; Figure S5H), indicating that our mscRNA-seq data provide a broadly representative view of the microbial composition.

Defining microbial functional subpopulations via gene co-expression and marker analysis

To elucidate the functional architecture of the subgingival microbiome, we performed high-dimensional weighted gene co-expression network analysis (hdWGCNA)[36] across 16 transcriptionally defined functional clusters, identifying seven core gene co-expression modules (Figures 3A–3C). Module eigengene 1 (ME1), predominantly expressed in cluster 6, was enriched for genes involved in polysaccharide degradation

(E) MDS plots of subgingival samples from HC and Pd groups. The left panel is based on the sample-level pseudobulk gene expression profiles (mean log-normalized expression). The right panel is based on species composition, calculated as the relative abundance of each species per sample.
See also Figures S3–S5.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017



(legend on next page)

(*BglS* [COG2273], *ManB2* [COG4124], and *Pgu1* [COG5434]). ME2, highly associated with cluster 10, was enriched in genes linked to exopolysaccharide export (*GumC* [COG3206]) and lipid biosynthesis (*LpxD* [COG1044]). ME3, overlapping with clusters 7 and 11, showed high expression of genes involved in anaerobic C4-dicarboxylate metabolism (*DcuA* [COG2704], *SdhA* [COG1053], *AspA* [COG1027], and *NadB* [COG0029]). ME4, primarily expressed in clusters 3, 4, and 14, co-expressed genes involved in adhesion (*OmpC* [COG3203] and *Hia* [COG5295]), which are essential for cell-cell interactions.[37] ME5, linked to clusters 2 and 12, was enriched in genes related to anaerobic energy metabolism, particularly formate oxidation (*NuoI* [COG1143]) and ferredoxin-dependent electron transport (*PorD* [COG1144], *PorB* [COG1013], and *NapF* [COG1145]). ME6, spanning clusters 0, 8, 9, 13, and 15, was enriched for genes involved in agmatine biosynthesis (*SpeA* [COG1166]) and inorganic ion uptake (*BtuB* [COG4206], *FepA* [COG4771], and *CirA* [COG1629]). Finally, ME7, broadly expressed across clusters, was enriched in genes related to translation and ribosomal proteins (e.g., *TufA* [COG0050], *FusA* [COG0480], *RpsC* [COG0092], and *RplB* [COG0090]), highlighting a transcriptionally active core module associated with protein synthesis and cellular maintenance in the subgingival microbiome.

Based on co-expression patterns from ME1 to ME6, we defined six distinct microbial functional subpopulations. Clusters 1 and 5, which lacked strong co-expression signals in any module, were classified as two additional subpopulations (Figure 3D). To further characterize the biological functions of these subpopulations, we examined 19 marker genes spanning key metabolic and structural processes, including (1) polysaccharide degradation, (2) C4-dicarboxylate metabolism, (3) denitrification, (4) formate metabolism, (5) pyruvate oxidation, (6) amino acid metabolism, (7) inorganic ion transportation, (8) recombination, (9) adhesin, (10) type IV pilus (Tfp), (11) efflux transport, (12) lipid biosynthesis, and (13) exopolysaccharide export (Figure 3D). By integrating module-based co-expression profiles with marker gene signatures, we assigned representative biological functions to each subpopulation (Figure 3E). For instance, in the agmatine biosynthesis cluster, nine of 18 key arginine metabolism genes were significantly upregulated, including arginine decarboxylase [*COG1166*], peptidylarginine/agmatine deiminase [*COG2957*], and carbon-nitrogen hydrolase [*COG0388*], which are predicted to catalyze consecutive steps from arginine to agmatine and putatively to putrescine (Figure S6A). Arginine decarboxylase was expressed in ~91% of cells, supporting the designation of this subpopulation as "agmatine biosynthesis cells."

Transcriptional profiles of subgingival bacteria in Pd reveal anaerobic energy, carbon, and amino acid metabolic signatures

To investigate the functional reorganization of the subgingival bacteria in Pd, we employed UMAP to visualize bacterial cells based on distinct functional states (Figure 4A). Comparative analysis between healthy and diseased states revealed a significant alteration in the functional composition of the subgingival microbiome. Specifically, in Pd, the transcriptional activity of adhesive cells (ACs) and polysaccharide degradation cells (PDCs) was notably reduced, whereas transposase$^+$ cells (TCs) and anaerobic ferredoxin-utilizing cells (AFCs) expanded markedly (Figure 4B). This was accompanied by downregulation of genes associated with aerobic energy production, including pyruvate dehydrogenase complex components *AceE* [COG2609] and *AceF* [COG0508], and upregulation of fermentative metabolism genes such as alcohol dehydrogenase *AdhP* [COG1064] and *EutG* [COG1454], as well as lactate dehydrogenase *LdhA* [COG1052] (Figure 4C).

To identify convergent transcriptional responses across distinct bacterial states, we examined genes consistently up- or downregulated across multiple functional clusters between health and Pd (Figures 4D and 4E). COG-based functional categorization revealed that energy production, amino acid metabolism, and carbohydrate utilization were among the top five categories with the most differentially expressed genes (DEGs) (Figure 4E). Focusing on energy production and carbohydrate utilization, bacteria in healthy samples exhibited higher expression of glycolytic and tricarboxylic acid (TCA) cycle enzymes, along with polysaccharide-degrading enzymes *BglS*, *ManB2*, and *Pgu1* (Figures 4F–4H). In Pd, upregulated genes included those for simple carbohydrate utilization (*YliI* [COG2133]), starch degradation (*AmyA* [COG0366]), ATP-binding cassette (ABC)-type carbon transporters (*MalK* [COG3839], *RbsB* [COG1879], *UgpA* [COG1175], and *UgpB* [COG1653]), and glycogen utilization (*GlgB* [COG0296] and *GDB1* [COG3408]) (Figure 4H), suggesting that bacteria in Pd may exploit a broader spectrum of carbohydrates, as reported previously.[38]

This transcriptional alteration also involved amino acid metabolism. In health, biosynthetic genes for methionine (*MetH1* [COG0646], *MetH2* [COG1410], and *MetE* [COG0620]), glutamate (*GltB2* [COG0069]), and lysine (*MetL1* [COG0527], *DapA* [COG0329], and *Asd* [COG0136]) were highly expressed (Figure 4I). In Pd, genes encoding cysteine proteases, serine proteases, and metalloproteases (COG4870, COG1572, *AprE* [COG1404], and *TldD* [COG0312]) were upregulated alongside peptide/amino acid transporters (*OppA* [COG4166], *AppF*

Figure 3. Identification and characterization of bacterial functional subpopulations based on gene co-expression and marker analysis

(A) UMAP visualization of single-cell data, with points colored according to their assigned Seurat clusters (0–15).

(B) Dot plot illustrating the association between ME modules and the identified clusters. The dot size corresponds to the percentage of genes expressed in each cluster, while the dot color reflects the average gene expression level within the cluster.

(C) Gene regulatory network of subgingival bacterial cells, color-coded by co-expression modules. Three representative hub genes from each module are highlighted.

(D) Dot plot depicting the representative marker genes for each of the eight functional clusters.

(E) Schematic of the biological processes/structures of the eight functional clusters. SBP, substrate-binding protein; TMD, transmembrane domains; EPS, exopolysaccharide; MCM, methylmalonyl-coenzyme A (CoA) mutase; ABC importer/exporter, ATP-binding cassette importer/exporter; Succ, succinate; DcuA, anaerobic C4-dicarboxylate transporter; L-Asp, L-aspartic acid; AspA, aspartate ammonia-lyase; Fum, fumarate; SdhA, fumarate reductase; FHL, formate hydrogenlyase; NosZ, nitrous oxide reductase; T2SS, type II secretion system.

See also Figure S6.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

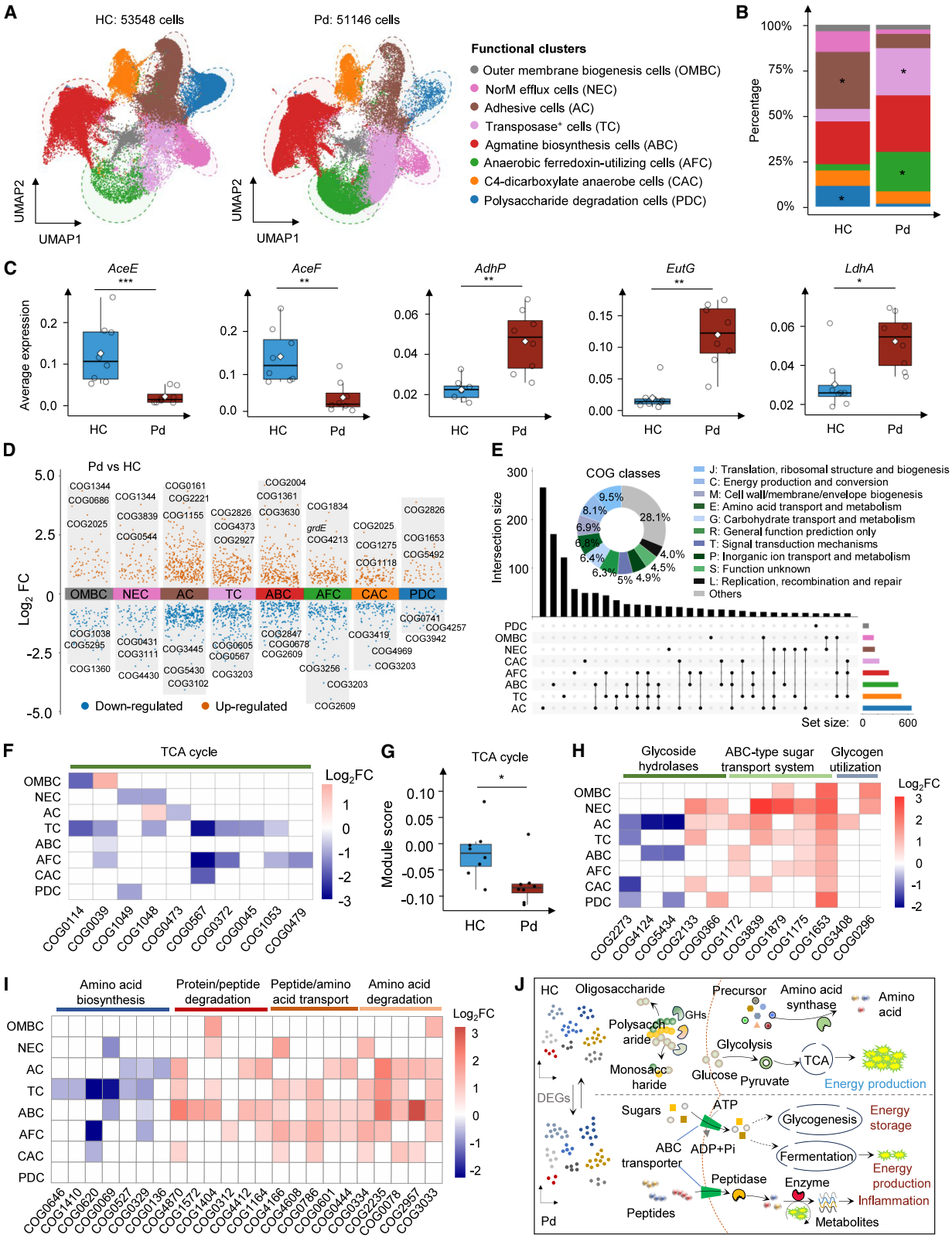


Figure 4. Transcriptional alterations in subgingival bacteria under Pd

(A) UMAP visualization of subgingival bacterial single-cell transcriptomes, split by group (HC, left; Pd, right). Cells are colored by functional cluster, which was defined based on co-expression patterns and marker genes across the entire dataset.

(legend continued on next page)

[COG4608], *GltS* [COG0786], *DppB* [COG0601], and *DppD* [COG0444]). Additionally, genes for the catabolism of glutamate (*GdhA* [COG0334]), arginine (*ArcA* [COG2235], *ArcB* [COG0078], and *AguA* [COG2957]), and tryptophan (*TnaA* [COG3033]) were significantly elevated, suggesting a potential increase in proteolytic and amino acid acquisition activity in Pd (Figure 4I). To complement these transcriptional observations, we performed metabolomics (MB) of amino acid and their derivatives in gingival crevicular fluid (GCF) samples from 18 healthy (MB-HC) and 18 Pd (MB-Pd) subjects (demographic details in Table S2). This analysis revealed a global metabolic alteration, with most detectable amino acids—including L-glycine, L-tryptophan, L-citrulline, and ornithine—elevated in Pd (Figures S6B and S6C), potentially creating an environment that favors the survival and metabolism of amino acid-dependent pathogens.

Together, these findings highlight adaptive transcriptional responses of subgingival bacteria during Pd, including reduced aerobic energy metabolism, enhanced fermentative activity, expanded carbohydrate utilization, and increased amino acid catabolism (Figure 4J).

Diminished translational and adhesive activity impairs *N. elongata* fitness in Pd

To investigate the transcriptional characteristics of microbial populations prevalent in healthy individuals, we focused on two dominant functional clusters: PDC and AC (Figure 4B). In the PDC cluster, *Lautropia mirabilis* dominated in healthy individuals (Figure S7A). Although the number of transcriptionally active *L. mirabilis* cells declined ~9-fold in disease (11,610 to 1,309), their subcluster distribution remained stable (Figures S7B and S7C). Subclustering revealed five *L. mirabilis* subpopulations (Figure S7D): two enriched for glycan-degrading enzymes targeting chitin and pectin, a third with higher per-cell expression of TCA cycle genes (Figures S7D and S7E), a fourth expressing high levels of adhesion and motility genes (e.g., *MotB*, *ClfA*, *OmpA*, and c-di-guanosine monophosphate [GMP] components; Figures S7D and S7F), and a fifth expressing oxidative stress response genes (e.g., *Bcp* [COG1225] and *TrxA* [COG0526]). These data highlight the transcriptional capacity of *L. mirabilis* for diverse carbon utilization, stress tolerance, and adhesion, features consistent with previous reports.[39,40]

In AC, *Neisseria elongata* was the dominant taxon (Figure S7A), a species previously implicated in adhesion and interbacterial signaling.[41] In Pd, the number of transcriptionally active *N. elongata* cells dropped sharply (4,334 to 870; Figure 5A) despite stable taxonomic abundance (1.55% versus 1.36%; Mendeley Data: https://doi.org/10.17632/kdktdhrb6b.1). Single-cell subclustering revealed multiple transcriptionally distinct populations, among which two were notably enriched for adhesins (cell adhesion cluster) and type VI secretion system (T6SS) genes (binding-effector cluster), and a translation-regulation subcluster enriched for ribosomal components (Figure 5B; Figures S7G and S7H). Notably, the relative cell proportions of the binding-effector and translation-regulation subclusters were markedly reduced in Pd (Figure 5C).

hdWGCNA mapped the translation-regulation subcluster primarily to ME3 (Figure 5D), which was the most significantly downregulated module in Pd (Figure 5E). ME3 was predominantly enriched for genes functioning in translation, ribosomal structure, and biogenesis, with approximately 75.9% of its genes (module eigengene-based connectivity [kME] > 0.5; $n = 54$) belonging to this category (Figure 5F). Accordingly, representative genes, including *RplB*, *RpsC*, *SecY*, *InfA*, *NusA*, and *FusA*, were significantly downregulated (Figure 5G), corroborating the module-level expression pattern.

To validate these findings, we isolated *N. elongata* from healthy clinical samples and confirmed its identity via 16S rRNA gene sequencing (Figure S7I). Whole-genome sequencing revealed spatially organized gene clusters encoding adhesion factors, Tfp, T6SS, and nitrate respiration machinery (Figure 5H). Scanning electron microscopy (SEM) confirmed its characteristic rod-shaped morphology—distinct from the coccoid forms typical of other Neisseria species such as *N. lactamica*[42]—and its capacity to adhere to oral epithelial cells (Figure 5I).

Given the reduced oxygen tension in deep periodontal pockets during Pd,[43] we modeled hypoxic stress *in vitro* by culturing *N. elongata* under anaerobic conditions. Although the species remained viable, its growth was markedly impaired (Figure 5J; Figure S7J). Quantitative real-time polymerase chain reaction (real-time qPCR) analysis showed downregulation of genes involved in translation, adhesion, and T6SS, alongside upregulation of nitrate respiration genes under hypoxic conditions (Figure 5K; Figure S7K; primers in Table S5), suggesting that *N. elongata* transcriptionally modulates growth- and respiration-associated programs in response to anaerobic stress.

Transcriptional programs of amino acid metabolism among key subgingival pathogens in Pd

Protein and amino acid metabolism are hallmarks of microbial dysbiosis in the subgingival niche, reflecting how pathogenic species exploit inflammation-induced tissue breakdown as a nutrient source.[44,45] Full-length 16S rRNA-seq revealed a restructured microbial community in Pd, marked by the enrichment of key pathogens such as *P. intermedia*, *P. gingivalis*,

(B) Bar graph illustrating the proportion of cells within each functional cluster, as shown in (A). Statistical significance was assessed using the two-sided Wilcoxon rank-sum test. $^{*}p < 0.05$.
(C) Boxplot illustrating the distribution of expression levels of key aerobic and anaerobic respiration-associated genes across each sample. Statistical significance was determined using the two-sided Wilcoxon rank-sum test. $^{*}p < 0.05$, $^{**}p < 0.01$.
(D) Volcano plot of DEGs across functional clusters with top three genes labeled; analyzed via pseudobulk approach.
(E) UpSet plot depicting the distribution of DEGs across functional clusters. The ring plot illustrates the distribution of all DEGs across COG classes.
(F, H, and I) Heatmaps showing the DEGs related to (F) the TCA cycle, (H) glycoside hydrolases, ABC-type sugar transport system, and glycogen synthesis, and (I) amino acid biosynthesis, protein/peptide degradation, peptide/amino acid transport, and amino acid catabolism across functional clusters, respectively.
(G) Boxplot showing module scores for the TCA cycle gene set (COG-annotated "TCA-cycle" category) in HC and Pd groups, calculated using the AddModuleScore function. Statistical significance was assessed using the two-sided Wilcoxon rank-sum test. $^{*}p < 0.05$.
(J) Schematic diagram summarizing metabolic pathway alterations in the subgingival bacteria between the HC and Pd groups.
See also Figure S6.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

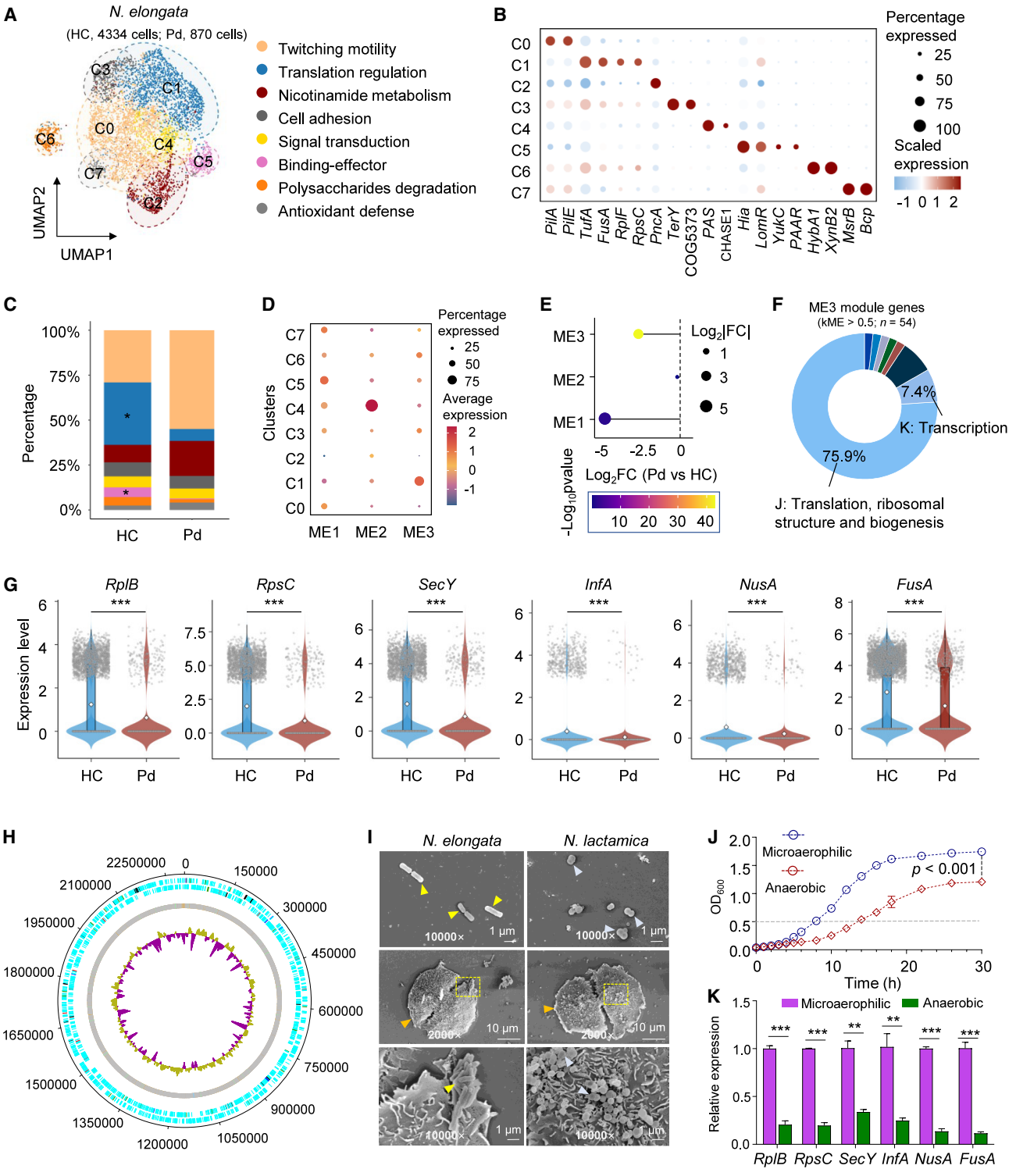


Figure 5. Functional modules and heterogeneity of *N. elongata* in health and Pd

(A) UMAP visualization of *N. elongata* single cells, colored by functional subclusters defined by marker gene expression.

(B) Dot plot displaying the expression patterns of key marker genes across *N. elongata* subclusters.

(C) Proportional distribution of *N. elongata* subclusters in HC ($n = 7$) and Pd ($n = 4$) groups (two-sided Wilcoxon rank-sum test, $^*p < 0.05$).

(D) Dot plot showing the expression patterns of three WGCNA modules (ME1–ME3) across *N. elongata* subpopulations. Dot size indicates the fraction of expressing cells; color intensity indicates mean expression.

(E) Lollipop plot showing differential module eigengene activity between groups, with each module represented by its $\log_2$ fold-change and statistical significance.

(legend continued on next page)

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

and *T. denticola* (Figures S6D and S6E). Building on our observation of elevated amino acids, integrated analysis of 16S rRNA-seq and targeted metabolomics showed that the abundances of these three key pathogenic taxa tended to correlate positively with most amino acids in Pd (Figure S6F). To investigate how these pathogens respond to the amino acid-rich environment, we curated a set of genes involved in amino acid and peptide transport, protease activity, and amino acid degradation from the non-redundant functional dataset (Table S6). We then calculated the functional gene proportion (FGP) for each single cell, defined as the fraction of transcripts assigned to a given functional gene set relative to the cell's total transcriptome, to quantify cellular functional activity. FGP analysis indicated that these three pathogens were among the major contributors to amino acid-related transcriptional activity (Figure S8A). Whereas *P. gingivalis* and *P. intermedia* primarily displayed proteolytic activity (Figure S8B), *T. denticola* showed a tendency toward amino acid uptake and degradation, reflecting differences in transcriptional patterns related to amino acid metabolism (Figure S8C).

Subclustering revealed seven transcriptional states in *P. gingivalis* (Figure 6A; Figure S8D). One subpopulation, annotated as "oxidative stress defense," showed elevated expression of antioxidant-related genes such as thiol peroxidase *Bcp* and thioredoxin *CnoX*. hdWGCNA indicated that this cluster was primarily associated with ME2 (Figure 6B), a module more active in health (Figure 6C), and both *Bcp* and *CnoX* were significantly downregulated in Pd (Figure S8E). By contrast, the "proteolysis" cluster showed higher FGP scores for protease/peptidase functions and aligned with ME5, a module upregulated in Pd (Figures 6B–6D). The "signal transduction" cluster corresponded to ME3, the module showing the largest fold increase and dominated by genes involved in signal transduction mechanisms (Figures 6B, 6C, and 6E). Functional validation under microaerobic and anaerobic conditions indicated that oxidative stress-related genes were strongly repressed under anaerobic conditions, whereas signal transduction genes and gingipains were significantly induced (Figure 6F; primers in Table S5), supporting the gene expression patterns observed in the corresponding transcriptional states.

Compared with *P. gingivalis*, *P. intermedia* exhibited a more narrowly focused proteolytic profile, with significantly higher FGP scores for protease/peptidase activity concentrated in a dedicated subcluster (Figures 6D, 6G, and 6H; Figure S8F), which expanded substantially in Pd (Figure 6I), showing stronger protease and peptidase activity than other subclusters (Figure S8G), reflecting a specialized proteolytic function. qPCR analysis showed that anaerobic conditions promoted the expression of the peptidase-related genes in *P. intermedia* (Figure S8H; primers in Table S5).

To compare metabolic gene expression between *T. denticola* and *P. gingivalis*, we performed differential gene expression analysis (Figure 6J). *T. denticola* preferentially expressed ABC-type oligopeptide transporters (Figure 6K; Figure S9A). In contrast, *P. gingivalis* favored di- and tri-peptide permeases and proton-gradient-driven transporters (Figure 6K), as well as diverse peptidases, including previously reported glycine-liberating proteases (*PepC*)[46] and cysteine proteinases (*rgpA* and *rgpB*) that generate peptidyl-arginine and limited free arginine[47] (Figure S9B). In line with this proteolytic preference, *P. gingivalis* also predominantly expressed genes linked to arginine catabolism, including arginine decarboxylase (*SpeA*), peptidylarginine/agmatine deiminase (*AguA*), carbon-nitrogen hydrolase (*Nit2*), and citrulline utilization hydrolase (*CtlX*) (Figures 6J and 6K), engaging pathways for arginine-to-agmatine conversion, polyamine metabolism, and the ornithine-citrulline branch (Figure S9C). qPCR suggested upregulation of these genes under anaerobic conditions, and metabolomic profiling of *P. gingivalis* culture supernatants showed L-arginine depletion accompanied by increased citrulline and ornithine (Figures S9D–S9F).

In contrast, *T. denticola* preferentially expressed genes of the arginine deiminase pathway (*ArcA*, *ArcB*, and *ArcC*) and the glycine reductase pathway (*grdA*, *grdE*, and *grdB*), both contributing to ATP generation under anaerobic conditions[48–50] (Figures 6J and 6K; Figure S9A). Subclustering indicated that these glycine- and arginine-catabolic genes tended to be preferentially expressed in a specific subpopulation, with similar subcluster-biased patterns observed for histidine and tryptophan degradation (Figures S9G and S9H).

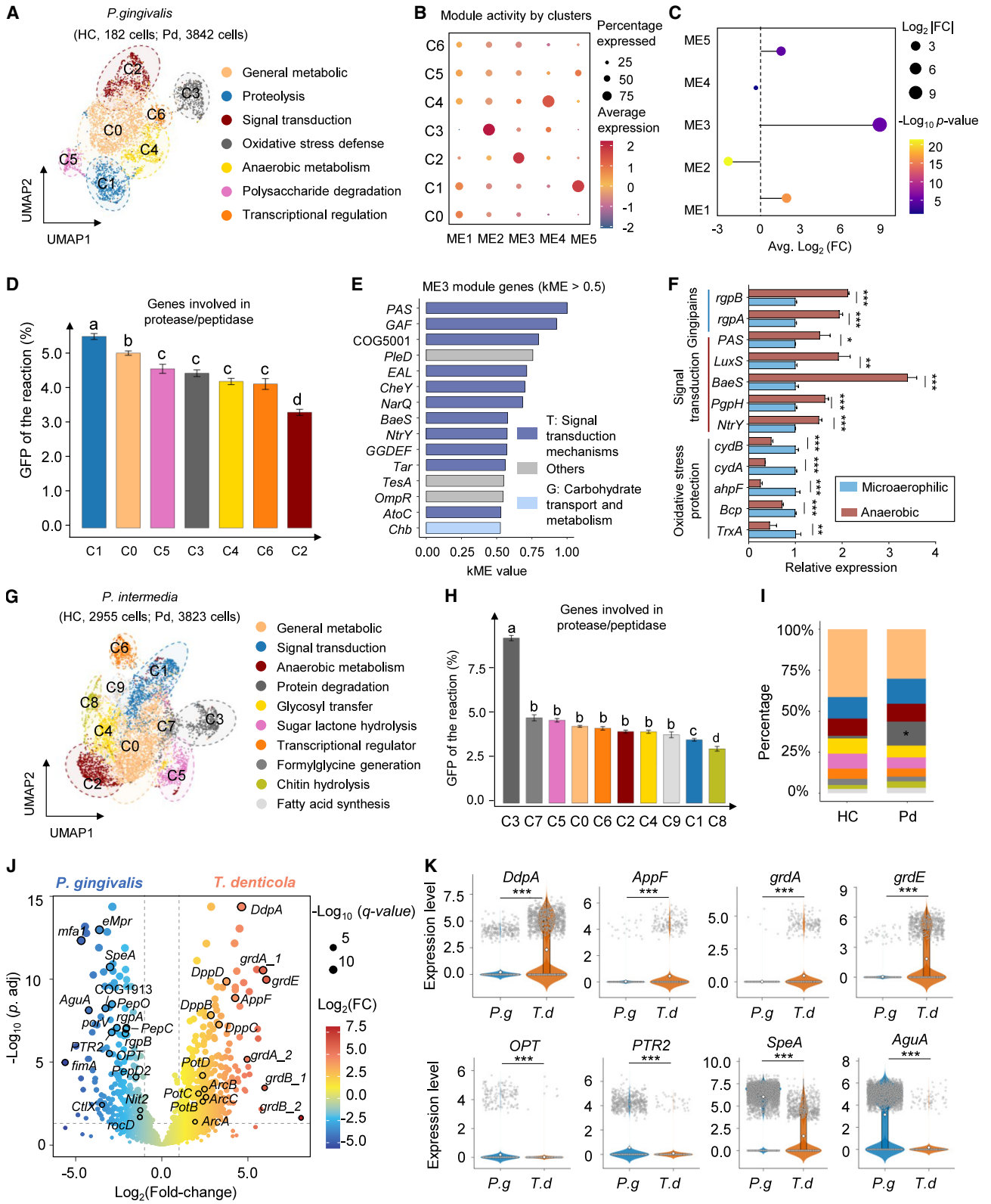
Together, these results reveal species-specific transcriptional patterns in Pd: *P. gingivalis* exhibits elevated expression of proteolytic genes and arginine catabolism pathways, *T. denticola* shows increased expression of genes involved in peptide transport, arginine deiminase activity, and glycine fermentation, and *P. intermedia* displays a focused upregulation of proteolytic genes concentrated in a dedicated subcluster. These patterns should be considered hypothesis-generating, reflecting inferred transcriptional tendencies rather than proven metabolic interactions (Figure S9I).

DISCUSSION

Microbial scRNA-seq holds great promise for resolving cell-specific functions in complex communities.[19,51] However, its extension to host-associated environments has been severely limited by technical constraints, primarily extensive host contamination and low microbial biomass. In oral samples, for

(F) Pie chart representing the COG category distribution of genes within module ME3 (kME > 0.5).
(G) Violin plots showing key protein synthesis genes differing significantly between HC and Pd groups (***$p < 0.001$, two-sided Wilcoxon rank-sum test). White diamonds mark group means.
(H) Circular genome map of *N. elongata* showing forward/reverse coding sequences (CDSs), pseudogenes, T6SS (red), adhesion (dark green), nitrate metabolism (blue), pilus genes (black), tRNAs (light green), and Guanine (G)–Cytosine (C) content (purple/gold).
(I) SEM images of *N. elongata* (left) and *N. lactamica* (right). Yellow triangles indicate *N. elongata*, blue triangles indicate *N. lactamica*, and orange triangles represent oral epithelial cells.
(J) Growth curves under microaerophilic (blue) and anaerobic (red) conditions. Statistical significance was determined by Student's *t* test. Bars show mean ± SEM of three biological replicates. The gray dashed line indicates the optical density 250 (OD_{600}) at which cells were collected for qPCR.
(K) qPCR reveals relative expression of translation-related genes in anaerobic versus microaerophilic groups. Bars show mean ± SEM from three replicates (Student's *t* test, **$p < 0.01$, ***$p < 0.001$).
See also Figure S7.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017



(legend on next page)

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

instance, human-derived reads can constitute up to 74.4% of total sequencing output in dental plaque,[14] with microbial retention dropping to as low as 0.95% after host-read depletion in mucosal swabs.[52] Similar challenges are observed across other low-biomass mucosal environments, such as the nasal and vaginal microbiota, where microbial signals are further obscured by elevated host backgrounds.[53,54] These challenges are compounded by the intrinsically high sequencing depth required to capture single-cell transcriptomes, making mscRNA-seq technically challenging and costly. As a result, its application has remained largely restricted to high-biomass sources such as synthetic consortia and fecal samples, limiting its generalizability across host-associated ecosystems.

To address these barriers, we developed an optimized mscRNA-seq workflow for subgingival bacteria by combining differential gradient centrifugation for host-cell removal.[55] Unlike previous efforts that primarily profiled abundant taxa,[12–14] our workflow resolves transcriptional heterogeneity across and within microbial lineages under native environmental contexts, generating a single-cell transcriptomic atlas of the subgingival microbiome across healthy and Pd-affected sites, encompassing 57 core active species.

Within this framework, *L. mirabilis* emerges as a health-associated commensal, consistent with its enrichment in periodontally healthy individuals.[56–58] Beyond its taxonomic signature, *L. mirabilis* exhibits marked functional heterogeneity: the dominant subcluster expresses genes involved in polysaccharide degradation, consistent with evidence that oligosaccharides produced can modulate host immunity.[59,60] This functional inference aligns with interventional evidence showing that the prebiotic mannan oligosaccharide attenuates alveolar bone loss and modulates immune-inflammatory responses in experimental Pd,[61] underscoring the therapeutic potential of selectively reinforcing health-compatible microbiota.

Similarly, *N. elongata* appears poised to act as a keystone commensal defender. It expresses adhesion machinery and a T6SS implicated in interbacterial antagonism, including inhibition of *P. gingivalis* epithelial invasion[62] and employs immunemodulatory strategies that support stable mucosal colonization.[63] Beyond these defensive roles, *N. elongata* engages in ecological interactions through Tfp-mediated horizontal gene transfer.[41] In Pd, we observed a decline in transcriptional activity—particularly in subpopulations with high translation and T6SS gene expression, although it remains unclear whether these cells are dormant or non-viable. Future studies using image-based approaches, such as bacterial MERFISH,[64] could clarify the physiological significance of this low-transcriptional-activity state.

In Pd, the subgingival niche is characterized by an inflammatory, nutrient-rich milieu, reflected in a substantial increase in host-derived proteins.[65,66] Consistent with this, our metabolomic analysis detected elevated amino acids in GCF, creating an amino acid-enriched milieu that may provide metabolic opportunities for resident microbes.[67] Although host tissue breakdown appears to be a major contributor, bacterial cross-feeding may also shape the amino acid landscape. Future studies are required to disentangle the relative contributions of host- versus microbe-derived substrates and their impact on community function.

In this amino acid-enriched niche, key pathogens appear to adopt species-specific metabolic programs. *P. gingivalis* activates arginine catabolism and proteolytic programs associated with pro-inflammatory and immune evasion,[68–71] while *T. denticola* engages in glycine fermentation, histidine degradation, and tryptophan catabolism—pathways linked to tissue breakdown and malodor.[72–75] Although both species utilize amino acids,[76–78] they adopt distinct nutrient acquisition strategies: *T. denticola* predominantly upregulates ABC-type transporters,[18] whereas *P. gingivalis* favors *Opt* and proton-dependent transport systems.[79] These metabolic divergences underscore the dynamic metabolic reprogramming inherent to pathogenic adaptation and identify peptide transport systems as promising therapeutic targets.[80] Although our findings support species-specific transcriptional tendencies, they do not establish a metabolic division of labor, which will require experimental validation using approaches such as isotopically labeled amino acid tracing.[81]

Beyond mapping the functional landscape of dysbiosis, our single-cell atlas suggests a new therapeutic paradigm for Pd that shifts from broad-spectrum antimicrobials to precision ecological intervention. The metabolic coordination among pathobionts, such as the glycine-liberating proteases (e.g.,

Figure 6. Metabolic signatures of *P. gingivalis*, *P. intermedia*, and *T. denticola* in Pd

(A and G) UMAP visualization of (A) *P. gingivalis* and (G) *P. intermedia* subclusters, with cells colored according to their respective subcluster identities.

(B) Dot plot showing the expression patterns of five WGCNA modules (ME1–ME5) across *P. gingivalis* subpopulations. Dot size indicates the fraction of expressing cells; color intensity indicates mean expression.

(C) Lollipop plot showing differential module eigengene activity between groups, with each module represented by its $\log_2$ fold-change and statistical significance.

(D and H) Bar plots showing the mean FGP for peptidase- and protease-related gene expression across subclusters of (D) *P. gingivalis* and (H) *P. intermedia*. Different letters denote significant differences ($p < 0.05$, Kruskal-Wallis test with Dunn's post hoc test). Data are mean ± SEM.

(E) Bar plot showing the top ME3 hub genes (kME > 0.5), ordered by kME and color-coded by COG categories.

(F) qPCR analysis of *P. gingivalis* genes related to oxidative stress defense, signal transduction, and gingipain proteases under anaerobic versus microaerophilic conditions. Bars represent mean ± SEM of three replicates (Student's *t* test; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$).

(I) Proportional distribution of *P. intermedia* subclusters in HC ($n = 6$) and Pd ($n = 8$) groups. Statistical significance was determined using a two-sided Wilcoxon rank-sum test.

(J) Volcano plot showing differential gene expression between *T. denticola* and *P. gingivalis* pseudobulk samples. DESeq2 was applied with size-factor normalization; genes present in only one species were retained by assigning zero counts to the other species.

(K) Violin plots showing single-cell expression levels of representative genes involved in ABC-type transporters and glycine reductase (top), and proton-gradient-driven peptide transporters and arginine catabolism (bottom) in *P. gingivalis* and *T. denticola*. White diamonds mark group means. Statistical significance was assessed using a two-sided Wilcoxon rank-sum test. $^{***}p < 0.001$. *P.g*, *P. gingivalis*; *T.d*, *T. denticola*.

See also Figures S6, S8, and S9.

PepC) in *P. gingivalis* that facilitate its partnership with *T. denticola*,[46,49] reveals targetable vulnerabilities to destabilize the dysbiotic network. In parallel, metabolic by-products of transcriptionally suppressed commensal populations could identify biomarkers of dysbiosis and early disease, while targeted strategies to reactivate native microbiota may offer therapeutic benefits beyond probiotic supplementation.

In summary, our single-cell transcriptomic atlas provides a high-resolution framework for examining how environmental pressures shape microbial gene expression in health and Pd. By resolving transcriptional heterogeneity within key populations, this work offers new insights into the functional potential of subgingival communities and lays the foundation for future studies linking microbial transcriptional states to ecological interactions and host outcomes.

Limitations of the study

Although this study advances insights into microbial ecology in Pd, several limitations should be considered. First, a subset of genes remained unannotated and were excluded from functional analyses. These likely represent "functional dark matter" of the oral microbiome, potentially encoding unrecognized molecular activities. Second, inherent differences in bacterial cell wall structure, membrane permeability, and intracellular RNA content may introduce species-specific biases during sample processing, potentially affecting taxonomic representation and transcript capture efficiency. Third, most host cells were filtered during sample processing, limiting our ability to assess host-microbe interactions. Developing workflows that preserve both host and microbial transcriptomes *in situ* represents an important future direction to capture the interplay between host and microbial communities in their native subgingival environment. Fourth, the high sequencing cost of single-cell microbiome profiling constrained our sample size, reducing statistical power to evaluate demographic variables such as age, sex, and ethnicity and limiting our ability to detect subtle but biologically important interindividual differences, including subgroup-specific transcriptional patterns related to disease severity, smoking, or host genetics. Future studies that incorporate larger and more diverse cohorts, along with more cost-efficient workflows, will be essential to extend these findings across broader populations and oral ecological settings.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to, and will be fulfilled upon reasonable request by, the lead contact: Yongcheng Wang (yongcheng@zju.edu.cn).

Materials availability

This study did not generate new unique reagents, and all materials used in this study are reported either in the main text or in the supplemental information.

Data and code availability

- Metagenomic raw data from our previous study are available under project ID PRJNA1238390. Whole-genome sequencing of *N. elongata* and mscRNA-seq data reported in this study have been deposited in the Genome Sequence Archive (GSA) under accession numbers GSA: CRA026826 and CRA023798, respectively. 16S rRNA-seq data are available under accession number GSA: CRA030816. Non-redundant gene annotation files, gene expression matrices, cell-level taxonomic assignments, and intermediate analysis files (in RDS format) are available on Figshare: https://figshare.com/s/867dbfdd7997e78d6e71. Raw data from Figures 2, 3, 4, 5, 6, S1, and S3–S9 have been deposited in Mendeley: https://doi.org/10.17632/kdktdhrb6b.1.
- Original code used for the mscRNA-seq analysis has been deposited in Zenodo: https://doi.org/10.5281/zenodo.18344612 and mirrored on GitHub: https://github.com/WangycLab/SGM_code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This project was supported by the Space Medical Experiment Project of CMSP (HYZHXMH03003, Y.W.), the National Key R&D Program of China (2025YFF 0512800, Y.W.), the Pioneer R&D Programs of Zhejiang Province (2024C0 3005, Y.W.), and the Key R&D Program of Zhejiang (2024SSYS0022, Y.W.). Additionally, this work received funding from the National Natural Science Foundation of China (82170953 and 82470973, P.-H.D.), the Science Fund for Distinguished Young Scholars of Zhejiang Province (LR23H140001, P.-H.D.), the Key R&D Program of Zhejiang Province (2022C03088, P.-H.D.), the Fundamental Research Funds for the Central Universities (226-2024-00202, P.-H.D.), and the Postdoctoral Science Foundation of China (2024M762880, L.S.). We thank the core facilities of Zhejiang University and Liangzhu Laboratory for technical support and the Core Facility, Zhejiang Provincial Clinical Research Center for Oral Diseases, for technical support.

AUTHOR CONTRIBUTIONS

Y.W., P.-H.D., L.D., and L.S. conceived this project; L.D., Y.L., M.S., and X.L. performed the single-bacterium RNA-seq experiments; L.S. and L.D. conducted the *in vitro* functional validation experiments; L.S., W.H., and P.-H.D. collected the microbiota samples and conducted oral health evaluations; L.D., J.W., Y.H., and L.S. analyzed the single-bacterium RNA-seq data, 16S rRNA profiles, genome-sequencing data, and both the targeted and untargeted metabolomic datasets; L.D., L.S., Y.H., and J.W. drafted the manuscript; Y.W., P.-H.D., L.D., L.S., and Q.C. revised the manuscript; Y.W. and P.-H.D. supervised this project; all authors have reviewed and approved the final manuscript.

DECLARATION OF INTERESTS

Y.W. is a co-founder of M20 Genomics.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Participant recruitment
  - Inclusion and exclusion criteria
  - Immortalized human oral mucosal epithelial cells
- METHOD DETAILS
  - Sample collection
  - Cell permeabilization
  - In-cell reverse transcription and cDNA 3' poly-A tailing
  - Single-cell droplet generation
  - cDNA enrichment
  - Library preparation
  - Data quality control and filtering
  - Construction of the non-redundant gene catalog
  - Gene abundance calculation and high-quality single-cell screening
  - Preparation of the Kraken2-based gOTUs database
  - Taxonomy determination
  - Filtering and benchmarking before clustering analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2026.01.017.

Received: June 27, 2025
Revised: November 24, 2025
Accepted: January 26, 2026

REFERENCES

1. Kassebaum, N.J., Bernabé, E., Dahiya, M., Bhandari, B., Murray, C.J., and Marcenes, W. (2014). Global burden of severe periodontitis in 1990–2010: a systematic review and meta-regression. J. Dent. Res. *93*, 1045–1053. https://doi.org/10.1177/0022034514552491.
2. Kc, S., Wang, X.Z., and Gallagher, J.E. (2020). Diagnostic sensitivity and specificity of host-derived salivary biomarkers in periodontal disease amongst adults: Systematic review. J. Clin. Periodontol. *47*, 289–308. https://doi.org/10.1111/jcpe.13218.
3. Peres, M.A., Macpherson, L.M.D., Weyant, R.J., Daly, B., Venturelli, R., Mathur, M.R., Listl, S., Celeste, R.K., Guarnizo-Herreño, C.C., Kearns, C., et al. (2019). Oral diseases: a global public health challenge. Lancet *394*, 249–260. https://doi.org/10.1016/S0140-6736(19)31146-8.
4. Pihlstrom, B.L., Michalowicz, B.S., and Johnson, N.W. (2005). Periodontal diseases. Lancet *366*, 1809–1820. https://doi.org/10.1016/S0140-6736(05)67728-8.
5. Villoria, G.E.M., Fischer, R.G., Tinoco, E.M.B., Meyle, J., and Loos, B.G. (2024). Periodontal disease: A systemic condition. Periodontol. 2000 *96*, 7–19. https://doi.org/10.1111/prd.12616.
6. GBD 2021 Oral Disorders Collaborators (2025). Trends in the global, regional, and national burden of oral conditions from 1990 to 2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet *405*, 897–910. https://doi.org/10.1016/S0140-6736(24)02811-3.
7. Kinane, D.F., Stathopoulou, P.G., and Papapanou, P.N. (2017). Periodontal diseases. Nat. Rev. Dis. Primers *3*, 17038. https://doi.org/10.1038/nrdp.2017.38.
8. Jain, N., Dutt, U., Radenkov, I., and Jain, S. (2024). WHO's global oral health status report 2022: Actions, discussion and implementation. Oral Dis. *30*, 73–79. https://doi.org/10.1111/odi.14516.
9. Socransky, S.S., Haffajee, A.D., Cugini, M.A., Smith, C., and Kent, R.L. (1998). Microbial complexes in subgingival plaque. J. Clin. Periodontol. *25*, 134–144. https://doi.org/10.1111/j.1600-051x.1998.tb02419.x.
10. Griffen, A.L., Beall, C.J., Campbell, J.H., Firestone, N.D., Kumar, P.S., Yang, Z.K., Podar, M., and Leys, E.J. (2012). Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. ISME J. *6*, 1176–1185. https://doi.org/10.1038/ismej.2011.191.
11. Li, Y., He, J., He, Z., Zhou, Y., Yuan, M., Xu, X., Sun, F., Liu, C., Li, J., Xie, W., et al. (2014). Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. ISME J. *8*, 1879–1891. https://doi.org/10.1038/ismej.2014.28.
12. Dabdoub, S.M., Ganesan, S.M., and Kumar, P.S. (2016). Comparative metagenomics reveals taxonomically idiosyncratic yet functionally congruent communities in periodontitis. Sci. Rep. *6*, 38993. https://doi.org/10.1038/srep38993.
13. Izawa, K., Okamoto-Shibayama, K., Kita, D., Tomita, S., Saito, A., Ishida, T., Ohue, M., Akiyama, Y., and Ishihara, K. (2021). Taxonomic and Gene Category Analyses of Subgingival Plaques from a Group of Japanese Individuals with and without Periodontitis. Int. J. Mol. Sci. *22*, 5298. https://doi.org/10.3390/ijms22105298.
14. Wang, J., Qi, J., Zhao, H., He, S., Zhang, Y., Wei, S., and Zhao, F. (2013). Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. Sci. Rep. *3*, 1843. https://doi.org/10.1038/srep01843.
15. Duran-Pinedo, A.E., Chen, T., Teles, R., Starr, J.R., Wang, X., Krishnan, K., and Frias-Lopez, J. (2014). Community-wide transcriptome of the oral microbiome in subjects with and without periodontitis. ISME J. *8*, 1659–1672. https://doi.org/10.1038/ismej.2014.23.
16. Jorth, P., Turner, K.H., Gumus, P., Nizam, N., Buduneli, N., and Whiteley, M. (2014). Metatranscriptomics of the human oral microbiome during health and disease. mBio *5*, e01012-14. https://doi.org/10.1128/mBio.01012-14.
17. Szafrański, S.P., Deng, Z.L., Tomasch, J., Jarek, M., Bhuju, S., Meisinger, C., Kühnisch, J., Sztajer, H., and Wagner-Döbler, I. (2015). Functional biomarkers for chronic periodontitis and insights into the roles of Prevotella nigrescens and Fusobacterium nucleatum; a metatranscriptome analysis. NPJ Biofilms Microbiomes *1*, 15017. https://doi.org/10.1038/npjbiofilms.2015.17.
18. Yost, S., Duran-Pinedo, A.E., Teles, R., Krishnan, K., and Frias-Lopez, J. (2015). Functional signatures of oral dysbiosis during periodontitis progression revealed by microbial metatranscriptome analysis. Genome Med. *7*, 27. https://doi.org/10.1186/s13073-015-0153-3.
19. Pountain, A.W., and Yanai, I. (2025). Dissecting microbial communities with single-cell transcriptome analysis. Science *389*, eadp6252. https://doi.org/10.1126/science.adp6252.
20. Eisenstein, M. (2025). Self-driving laboratories, advanced immunotherapies and five more technologies to watch in 2025. Nature *637*, 1008–1011. https://doi.org/10.1038/d41586-025-00075-6.
21. Lloréns-Rico, V., Simcock, J.A., Huys, G.R.B., and Raes, J. (2022). Single-cell approaches in human microbiome research. Cell *185*, 2725–2738. https://doi.org/10.1016/j.cell.2022.06.040.
22. Blattman, S.B., Jiang, W., Oikonomou, P., and Tavazoie, S. (2020). Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. Nat. Microbiol. *5*, 1192–1201. https://doi.org/10.1038/s41564-020-0729-6.
23. Kuchina, A., Brettner, L.M., Paleologu, L., Roco, C.M., Rosenberg, A.B., Carignano, A., Kibler, R., Hirano, M., DePaolo, R.W., and Seelig, G. (2021). Microbial single-cell RNA sequencing by split-pool barcoding. Science *371*, eaba5257. https://doi.org/10.1126/science.aba5257.
24. Ma, P., Amemiya, H.M., He, L.L., Gandhi, S.J., Nicol, R., Bhattacharyya, R.P., Smillie, C.S., and Hung, D.T. (2023). Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. Cell *186*, 877–891.e14. https://doi.org/10.1016/j.cell.2023.01.002.
25. Xu, Z., Wang, Y., Sheng, K., Rosenthal, R., Liu, N., Hua, X., Zhang, T., Chen, J., Song, M., Lv, Y., et al. (2023). Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. Nat. Commun. *14*, 5130. https://doi.org/10.1038/s41467-023-40137-9.
26. Shen, Y., Qian, Q., Ding, L., Qu, W., Zhang, T., Song, M., Huang, Y., Wang, M., Xu, Z., Chen, J., et al. (2025). High-throughput single-microbe

RNA sequencing reveals adaptive state heterogeneity and host-phage activity associations in human gut microbiome. Protein Cell *16*, 211–226. https://doi.org/10.1093/procel/pwae027.

27. Shen, Y., Qu, W., Song, M., Zhang, T., Liu, C., Shi, X., Xu, X., Jiang, J., Ding, L., Mo, F., et al. (2025). Single-microbe RNA sequencing uncovers unexplored specialized metabolic functions of keystone species in the human gut. Imeta *4*, e70035. https://doi.org/10.1002/imt2.70035.
28. Jia, M., Zhu, S., Xue, M.Y., Chen, H., Xu, J., Song, M., Tang, Y., Liu, X., Tao, Y., Zhang, T., et al. (2024). Single-cell transcriptomics across 2,534 microbial species reveals functional heterogeneity in the rumen microbiome. Nat. Microbiol. *9*, 1884–1898. https://doi.org/10.1038/s41564-024-01723-9.
29. Lu, H., Zou, P., Zhang, Y., Zhang, Q., Chen, Z., and Chen, F. (2022). The sampling strategy of oral microbiome. Imeta *1*, e23. https://doi.org/10.1002/imt2.23.
30. Behzadnia, A., Moosavi-Nasab, M., Ojha, S., and Tiwari, B.K. (2020). Exploitation of Ultrasound Technique for Enhancement of Microbial Metabolites Production. Molecules *25*, 5473. https://doi.org/10.3390/molecules25225473.
31. Zhao, Y., Song, L., Li, H.Y., Liu, S.X., Mao, F.F., Li, X.J., and Ding, P.H. (2025). Metagenomic Insights Into the Subgingival Microbiome in Periodontal Health and Different Grades of Periodontitis. J. Periodont. Res. *60*, 788–798. https://doi.org/10.1111/jre.13408.
32. Escapa, I.F., Chen, T., Huang, Y., Gajare, P., Dewhirst, F.E., and Lemon, K.P. (2018). New Insights into Human Nostril Microbiome from the Expanded Human Oral Microbiome Database (eHOMD): a Resource for the Microbiome of the Human Aerodigestive Tract. mSystems *3*, e00187-18. https://doi.org/10.1128/mSystems.00187-18.
33. Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., et al. (2003). The COG database: an updated version includes eukaryotes. BMC Bioinformatics *4*, 41. https://doi.org/10.1186/1471-2105-4-41.
34. Yayon, N., Kedlian, V.R., Boehme, L., Suo, C., Wachter, B.T., Beuschel, R.T., Amsalem, O., Polanski, K., Koplev, S., Tuck, E., et al. (2024). A spatial human thymus cell atlas mapped to a continuous tissue axis. Nature *635*, 708–718. https://doi.org/10.1038/s41586-024-07944-6.
35. Oliver, A.J., Huang, N., Bartolome-Casado, R., Li, R., Koplev, S., Nilsen, H.R., Moy, M., Cakir, B., Polanski, K., Gudiño, V., et al. (2024). Single-cell integration reveals metaplasia in inflammatory gut diseases. Nature *635*, 699–707. https://doi.org/10.1038/s41586-024-07571-1.
36. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E., and Swarup, V. (2023). hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. Cell Rep. Methods *3*, 100498. https://doi.org/10.1016/j.crmeth.2023.100498.
37. Chen, P.Y., Chen, Y.C., Chen, P.P., Lin, K.T., Sargsyan, K., Hsu, C.P., Wang, W.L., Hsia, K.C., and Ting, S.Y. (2024). A whole-cell platform for discovering synthetic cell adhesion molecules in bacteria. Nat. Commun. *15*, 6568. https://doi.org/10.1038/s41467-024-51017-1.
38. Zhang, Y., Zheng, Y., Hu, J., Du, N., and Chen, F. (2014). Functional diversity of the microbial community in healthy subjects and periodontitis patients based on sole carbon source utilization. PLoS One *9*, e91977. https://doi.org/10.1371/journal.pone.0091977.
39. Gerner-Smidt, P., Keiser-Nielsen, H., Dorsch, M., Stackebrandt, E., Ursing, J., Blom, J., Christensen, A.C., Christensen, J.J., Frederiksen, W., and Hoffmann, S. (1994). Lautropia mirabilis gen. nov., sp. nov., a gram-negative motile coccus with unusual morphology isolated from the human mouth. Microbiology (Reading) *140*, 1787–1797. https://doi.org/10.1099/13500872-140-7-1787.
40. Daneshvar, M.I., Douglas, M.P., and Weyant, R.S. (2001). Cellular fatty acid composition of Lautropia mirabilis. J. Clin. Microbiol. *39*, 4160–4162. https://doi.org/10.1128/JCM.39.11.4160-4162.2001.
41. Higashi, D.L., Biais, N., Weyand, N.J., Agellon, A., Sisko, J.L., Brown, L. M., and So, M. (2011). N. elongata produces type IV pili that mediate interspecies gene transfer with N. gonorrhoeae. PLoS One *6*, e21373. https://doi.org/10.1371/journal.pone.0021373.
42. Bovre, K., and Holten, E. (1970). Neisseria elongata sp.nov., a rod-shaped member of the genus Neisseria. Re-evaluation of cell shape as a criterion in classification. J. Gen. Microbiol. *60*, 67–75. https://doi.org/10.1099/00221287-60-1-67.
43. Loesche, W.J., Gusberti, F., Mettraux, G., Higgins, T., and Syed, S. (1983). Relationship between oxygen tension and subgingival bacterial flora in untreated human periodontal pockets. Infect. Immun. *42*, 659–667. https://doi.org/10.1128/iai.42.2.659-667.1983.
44. Ding, J., Zhao, C., and Gao, L. (2023). Metabolism of periodontal pathobionts: Their regulatory roles in the dysbiotic microbiota. Mol. Oral Microbiol. *38*, 181–188. https://doi.org/10.1111/omi.12409.
45. Socransky, S.S., and Haffajee, A.D. (2005). Periodontal microbial ecology. Periodontol. 2000 *38*, 135–187. https://doi.org/10.1111/j.1600-0757.2005.00107.x.
46. Kin, L.X., Butler, C.A., Slakeski, N., Hoffmann, B., Dashper, S.G., and Reynolds, E.C. (2020). Metabolic cooperativity between Porphyromonas gingivalis and Treponema denticola. J. Oral Microbiol. *12*, 1808750. https://doi.org/10.1080/20002297.2020.1808750.
47. Chow, Y.C., Yam, H.C., Gunasekaran, B., Lai, W.Y., Wo, W.Y., Agarwal, T., Ong, Y.Y., Cheong, S.L., and Tan, S.A. (2022). Implications of Porphyromonas gingivalis peptidyl arginine deiminase and gingipain R in human health and diseases. Front. Cell. Infect. Microbiol. *12*, 987683. https://doi.org/10.3389/fcimb.2022.987683.
48. Pols, T., Singh, S., Deelman-Driessen, C., Gaastra, B.F., and Poolman, B. (2021). Enzymology of the pathway for ATP production by arginine breakdown. FEBS J. *288*, 293–309. https://doi.org/10.1111/febs.15337.
49. Tan, K.H., Seers, C.A., Dashper, S.G., Mitchell, H.L., Pyke, J.S., Meuric, V., Slakeski, N., Cleal, S.M., Chambers, J.L., McConville, M.J., et al. (2014). Porphyromonas gingivalis and Treponema denticola exhibit metabolic symbioses. PLoS Pathog. *10*, e1003955. https://doi.org/10.1371/journal.ppat.1003955.
50. Blakemore, R.P., and Canale-Parola, E. (1976). Arginine catabolism by Treponema denticola. J. Bacteriol. *128*, 616–622. https://doi.org/10.1128/jb.128.2.616-622.1976.
51. Xu, Z., Wang, Y., Cai, W., Chen, Y., and Wang, Y. (2025). Single microorganism RNA sequencing of microbiomes using smRandom-Seq. Nat. Protoc. *21*, 160–199. https://doi.org/10.1038/s41596-025-01181-5.
52. Zhi, J., Liang, Y., Zhao, W., Qiao, J., Zheng, Y., Peng, X., Li, L., Wei, X., and Wang, W. (2025). Oral microbiome-derived biomarkers for non-invasive diagnosis of head and neck squamous cell carcinoma. NPJ Biofilms Microbiomes *11*, 74. https://doi.org/10.1038/s41522-025-00708-8.
53. Human; Microbiome; Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214. https://doi.org/10.1038/nature11234.
54. Marotz, C.A., Sanders, J.G., Zuniga, C., Zaramela, L.S., Knight, R., and Zengler, K. (2018). Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome *6*, 42. https://doi.org/10.1186/s40168-018-0426-3.
55. Ding, Y., Fernández-Montero, A., Mani, A., Casadei, E., Miyazawa, R., Zhou, C., Chaumont, L., Posavi, M., Cole, S.D., Shibasaki, Y., et al. (2025). Secretory IgM regulates gut microbiota homeostasis and metabolism. Nat. Microbiol. *10*, 1431–1446. https://doi.org/10.1038/s41564-025-02013-8.
56. Colombo, A.P.V., Boches, S.K., Cotton, S.L., Goodson, J.M., Kent, R., Haffajee, A.D., Socransky, S.S., Hasturk, H., Van Dyke, T.E., Dewhirst, F., et al. (2009). Comparisons of subgingival microbial profiles of refractory periodontitis, severe periodontitis, and periodontal health using the human oral microbe identification microarray. J. Periodontol. *80*, 1421–1432. https://doi.org/10.1902/jop.2009.090185.
57. Shaddox, L.M., Huang, H., Lin, T., Hou, W., Harrison, P.L., Aukhil, I., Walker, C.B., Klepac-Ceraj, V., and Paster, B.J. (2012). Microbiological

characterization in children with aggressive periodontitis. J. Dent. Res. *91*, 927–933. https://doi.org/10.1177/0022034512456039.

58. Abusleme, L., Hoare, A., Hong, B.Y., and Diaz, P.I. (2021). Microbial signatures of health, gingivitis, and periodontitis. Periodontol. 2000 *86*, 57–78. https://doi.org/10.1111/prd.12362.
59. Jana, U.K., Suryawanshi, R.K., Prajapati, B.P., and Kango, N. (2021). Prebiotic mannooligosaccharides: Synthesis, characterization and bioactive properties. Food Chem. *342*, 128328. https://doi.org/10.1016/j.foodchem.2020.128328.
60. Wongkaew, M., Tangjaidee, P., Leksawasdi, N., Jantanasakulwong, K., Rachtanapun, P., Seesuriyachan, P., Phimolsiripol, Y., Chaiyaso, T., Ruksiriwanich, W., Jantrawut, P., et al. (2022). Mango Pectic Oligosaccharides: A Novel Prebiotic for Functional Food. Front. Nutr. *9*, 798543. https://doi.org/10.3389/fnut.2022.798543.
61. Levi, Y.L.A.S., Novais, G.S., Dias, R.B., Andraus, R.A.C., Messora, M.R., Neto, H.B., Ervolino, E., Santinoni, C.S., and Maia, L.P. (2018). Effects of the prebiotic mannan oligosaccharide on the experimental periodontitis in rats. J. Clin. Periodontol. *45*, 1078–1089. https://doi.org/10.1111/jcpe.12987.
62. Fukuda, S., Akatsu, T., Fujii, A., Kawano, S., Minegishi, Y., and Ota, N. (2024). Commensal Neisseria inhibit Porphyromonas gingivalis invasion of gingival epithelial cells. Oral Health Prev. Dent. *22*, 609–616. https://doi.org/10.3290/j.ohpd.b5866430.
63. Zhu, W., Cardenas-Alvarez, M.X., Tomberg, J., Little, M.B., Duncan, J.A., and Nicholas, R.A. (2023). Commensal Neisseria species share immune suppressive mechanisms with Neisseria gonorrhoeae. PLoS One *18*, e0284062. https://doi.org/10.1371/journal.pone.0284062.
64. Sarfatis, A., Wang, Y., Twumasi-Ankrah, N., and Moffitt, J.R. (2025). Highly multiplexed spatial transcriptomics in bacteria. Science *387*, eadr0932. https://doi.org/10.1126/science.adr0932.
65. Barros, S.P., Williams, R., Offenbacher, S., and Morelli, T. (2016). Gingival crevicular fluid as a source of biomarkers for periodontitis. Periodontol. 2000 *70*, 53–64. https://doi.org/10.1111/prd.12107.
66. Bostanci, N., and Belibasakis, G.N. (2018). Gingival crevicular fluid and its immune mediators in the proteomic era. Periodontol. 2000 *76*, 68–84. https://doi.org/10.1111/prd.12154.
67. Santonocito, S., Giudice, A., Polizzi, A., Troiano, G., Merlo, E.M., Sclafani, R., Grosso, G., and Isola, G. (2022). A cross-talk between diet and the oral microbiome: balance of nutrition on inflammation and immune system's response during periodontitis. Nutrients *14*, 2426. https://doi.org/10.3390/nu14122426.
68. Joshi, M.S., Ferguson, T.B., Jr., Johnson, F.K., Johnson, R.A., Parthasarathy, S., and Lancaster, J.R. (2007). Receptor-mediated activation of nitric oxide synthesis by arginine in endothelial cells. Proc. Natl. Acad. Sci. USA *104*, 9982–9987. https://doi.org/10.1073/pnas.0506824104.
69. Ozer, L., Elgun, S., Ozdemir, B., Pervane, B., and Ozmeric, N. (2011). Arginine-nitric oxide-polyamine metabolism in periodontal disease. J. Periodontol. *82*, 320–328. https://doi.org/10.1902/jop.2010.100199.
70. Su, C.H., Liu, I.M., Chung, H.H., and Cheng, J.T. (2009). Activation of I2-imidazoline receptors by agmatine improved insulin sensitivity through two mechanisms in type-2 diabetic rats. Neurosci. Lett. *457*, 125–128. https://doi.org/10.1016/j.neulet.2009.03.093.
71. Gobert, A.P., McGee, D.J., Akhtar, M., Mendz, G.L., Newton, J.C., Cheng, Y., Mobley, H.L., and Wilson, K.T. (2001). Helicobacter pylori arginase inhibits nitric oxide production by eukaryotic cells: a strategy for bacterial survival. Proc. Natl. Acad. Sci. USA *98*, 13844–13849. https://doi.org/10.1073/pnas.241443798.
72. Takahashi, N. (2015). Oral Microbiome Metabolism: From "Who Are They?" to "What Are They Doing?". J. Dent. Res. *94*, 1628–1637. https://doi.org/10.1177/0022034515606045.
73. Zhang, X.X., Ritchie, S.R., and Rainey, P.B. (2014). Urocanate as a potential signaling molecule for bacterial recognition of eukaryotic hosts. Cell. Mol. Life Sci. *71*, 541–547. https://doi.org/10.1007/s00018-013-1527-6.
74. Anumula, L., Ramesh, S., and Kolaparthi, V.S.K. (2024). Matrix metalloproteinases in dentin: Assessing their presence, activity, and inhibitors - a review of current trends. Dent. Mater. *40*, 2051–2073. https://doi.org/10.1016/j.dental.2024.09.011.
75. Basic, A., and Dahlén, G. (2023). Microbial metabolites in the pathogenesis of periodontal diseases: a narrative review. Front. Oral Health *4*, 1210200. https://doi.org/10.3389/froh.2023.1210200.
76. Cai, X.Y., Zheng, C.X., Guo, H., Fan, S.Y., Huang, X.Y., Chen, J., Liu, J.X., Gao, Y.R., Liu, A.Q., Liu, J.N., et al. (2025). Inflammation-triggered Gli1+ stem cells engage with extracellular vesicles to prime aberrant neutrophils to exacerbate periodontal immunopathology. Cell. Mol. Immunol. *22*, 371–389. https://doi.org/10.1038/s41423-025-01271-0.
77. Hespell, R.B., and Canale-Parola, E. (1971). Amino acid and glucose fermentation by Treponema denticola. Arch. Mikrobiol. *78*, 234–251. https://doi.org/10.1007/BF00424897.
78. Seddon, S.V., Shah, H.N., Hardie, J.M., and Robinson, J.P. (1988). Chemically defined and minimal media forBacteroides gingivalis. Curr. Microbiol. *17*, 147–149. https://doi.org/10.1007/BF01573470.
79. Ohara-Nemoto, Y., Sarwar, M.T., Shimoyama, Y., Kobayakawa, T., and Nemoto, T.K. (2020). Preferential dipeptide incorporation of Porphyromonas gingivalis mediated by proton-dependent oligopeptide transporter (Pot). FEMS Microbiol. Lett. *367*, fnaa204. https://doi.org/10.1093/femsle/fnaa204.
80. Nemoto, T.K., and Ohara Nemoto, Y. (2021). Dipeptidyl-peptidases: Key enzymes producing entry forms of extracellular proteins in asaccharolytic periodontopathic bacterium Porphyromonas gingivalis. Mol. Oral Microbiol. *36*, 145–156. https://doi.org/10.1111/omi.12317.
81. Rakoff-Nahoum, S., Foster, K.R., and Comstock, L.E. (2016). The evolution of cooperation within the gut microbiota. Nature *533*, 255–259. https://doi.org/10.1038/nature17626.
82. Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.
83. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
84. Buchfink, B., Reuter, K., and Drost, H.G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods *18*, 366–368. https://doi.org/10.1038/s41592-021-01101-x.
85. Lu, J., Rincon, N., Wood, D.E., Breitwieser, F.P., Pockrandt, C., Langmead, B., Salzberg, S.L., and Steinegger, M. (2022). Metagenome analysis using the Kraken software suite. Nat. Protoc. *17*, 2815–2839. https://doi.org/10.1038/s41596-022-00738-y.
86. Lu, J., Breitwieser, F.P., Thielen, P., and Salzberg, S.L. (2017). Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci. *3*, e104. https://doi.org/10.7717/peerj-cs.104.
87. Zappia, L., and Oshlack, A. (2018). Clustering trees: a visualization for evaluating clusterings at multiple resolutions. GigaScience *7*, giy083. https://doi.org/10.1093/gigascience/giy083.
88. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.
89. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.
90. Crowell, H.L., Soneson, C., Germain, P.L., Calini, D., Collin, L., Raposo, C., Malhotra, D., and Robinson, M.D. (2020). muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. Nat. Commun. *11*, 6077. https://doi.org/10.1038/s41467-020-19894-4.

91. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386. https://doi.org/10.1038/nbt.2859.

92. Xu, S., Hu, E., Cai, Y., Xie, Z., Luo, X., Zhan, L., Tang, W., Wang, Q., Liu, B., Wang, R., et al. (2024). Using clusterProfiler to characterize multiomics data. Nat. Protoc. *19*, 3292–3320. https://doi.org/10.1038/s41596-024-01020-z.

93. Ye, J., McGinnis, S., and Madden, T.L. (2006). BLAST: improvements for better sequence analysis. Nucleic Acids Res. *34*, W6–W9. https://doi.org/10.1093/nar/gkl164.

94. Zhang, J., Guan, J., Wang, M., Li, G., Djordjevic, M., Tai, C., Wang, H., Deng, Z., Chen, Z., and Ou, H.Y. (2023). SecReT6 update: a comprehensive resource of bacterial Type VI Secretion Systems. Sci. China Life Sci. *66*, 626–634. https://doi.org/10.1007/s11427-022-2172-x.

95. Carver, T., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. Bioinformatics *25*, 119–120. https://doi.org/10.1093/bioinformatics/btn578.

96. Silness, J., and Loe, H. (1964). Periodontal Disease in Pregnancy. Ii. Correlation Between Oral Hygiene and Periodontal Condtion. Acta Odontol. Scand. *22*, 121–135. https://doi.org/10.3109/00016356408993968.

97. Papapanou, P.N., Sanz, M., Buduneli, N., Dietrich, T., Feres, M., Fine, D. H., Flemmig, T.F., Garcia, R., Giannobile, W.V., Graziani, F., et al. (2018). 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. *45*, S162–S170. https://doi.org/10.1111/jcpe.12946.

98. Iniesta, M., Chamorro, C., Ambrosio, N., Marín, M.J., Sanz, M., and Herrera, D. (2023). Subgingival microbiome in periodontal health, gingivitis and different stages of periodontitis. J. Clin. Periodontol. *50*, 905–920. https://doi.org/10.1111/jcpe.13793.

99. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

100. Commichaux, S., Shah, N., Ghurye, J., Stoppel, A., Goodheart, J.A., Luque, G.G., Cummings, M.P., and Pop, M. (2021). A critical assessment of gene catalogs for metagenomic analysis. Bioinformatics *37*, 2848–2857. https://doi.org/10.1093/bioinformatics/btab216.

101. Ghanbari Maman, L., Palizban, F., Fallah Atanaki, F., Elmi Ghiasi, N., Ariaeenejad, S., Ghaffari, M.R., Hosseini Salekdeh, G., and Kavousi, K. (2020). Co-abundance analysis reveals hidden players associated with high methane yield phenotype in sheep rumen microbiome. Sci. Rep. *10*, 4995. https://doi.org/10.1038/s41598-020-61942-y.

102. Galperin, M.Y., Wolf, Y.I., Makarova, K.S., Vera Alvarez, R., Landsman, D., and Koonin, E.V. (2021). COG database update: focus on microbial diversity, model organisms, and widespread pathogens. Nucleic Acids Res. *49*, D274–D281. https://doi.org/10.1093/nar/gkaa1018.

103. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. *19*, 455–477. https://doi.org/10.1089/cmb.2012.0021.

104. Bairoch, A., and Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. *28*, 45–48. https://doi.org/10.1093/nar/28.1.45.

105. Liu, C., Zhao, D., Ma, W., Guo, Y., Wang, A., Wang, Q., and Lee, D.J. (2016). Denitrifying sulfide removal process on high-salinity wastewaters in the presence of Halomonas sp. Appl. Microbiol. Biotechnol. *100*, 1421–1426. https://doi.org/10.1007/s00253-015-7039-6.

106. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583. https://doi.org/10.1038/nmeth.3869.

107. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857. https://doi.org/10.1038/s41587-019-0209-9.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

CellPress

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
VITApilote-PFT1200 kit	M20 Genomics	Cat# R20115124
DNA Library Prep Kit	Vazyme	Cat# ND607-03/04
4% paraformaldehyde	Meilunbio	Cat# MA0192-2
AMPure XP beads	Beckman	Cat# A63881
RNAex Pro RNA extraction reagent	Accurate Biology	Cat# AG21101
PrimeScript RT reagent Kit	Takara	Cat# RR037A
Dulbecco's Modified Eagle Medium	Gibco	Cat# 11594416
Fetal bovine serum	Gibco	Cat# 16629525
Chocolate agar	BeNa Culture Collection	Cat# 360798
Brain Heart Infusion	Thermo Fisher Scientific	Cat# CM1135B
Deposited data		
mscRNA-seq clean data	This paper	GSA: CRA023798
Metagenomic raw data	Our previous study[31]	NCBI SRA: PRJNA1238390
16S rRNA sequencing data	This paper	GSA: CRA030816
Whole-genome sequencing of *N. elongata*	This paper	GSA: CRA026826
Intermediate outputs	This paper	Figshare: https://figshare.com/s/867dbfdd7997e78d6e71
Source data	This paper	Mendeley: https://doi.org/10.17632/kdktdhrb6b.1
Experimental models: cell lines		
Immortalized Human Oral Mucosal Epithelial Cells	Creative Bioarray	Cat# CSC-I9117L
Bacterial and Virus Strains		
N. elongata	Isolated from clinical control samples	Chopped Meat Medium
N. lactamica strain ATCC23970	BeNa Culture Collection	Cat# BNCC259832
P.gingivalis strain W83	American Type Culture Collection	Cat# ATCC BAA-308
P. intermedia strain VPI 4197	American Type Culture Collection	Cat# ATCC 25611
Software and algorithms		
CD-HIT-EST	Li and Godzik[82]	https://github.com/weizhongli/cdhit
STAR	Dobin et al.[83]	https://github.com/alexdobin/STAR
Diamond	Buchfink et al.[84]	https://github.com/bbuchfink/diamond
Kraken2	Lu et al.[85]	https://github.com/DerrickWood/kraken2
Bracken	Lu et al.[86]	https://github.com/jenniferlu717/Bracken
Clustree	Zappia and Oshlack[87]	https://github.com/lazappi/clustree
Seurat	Hao et al.[88]	https://github.com/satijalab/seurat
Harmony	Korsunsky et al.[89]	https://github.com/immunogenomics/harmony
hdWGCNA	Morabito et al.[36]	https://github.com/smorabit/hdWGCNA
Muscat	Crowell et al.[90]	https://github.com/HelenaLC/muscat
Monocle 2	Trapnell et al.[91]	https://bioconductor.org/packages/release/bioc/html/monocle.html
ClusterProfiler	Xu et al.[92]	https://github.com/YuLab-SMU/clusterProfiler
NCBI BLAST	Ye et al.[93]	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.13.0/

(*Continued on next page*)

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
t6ss_finder.py	Zhang et al.[94]	https://bioinfo-mml.sjtu.edu.cn/SecReT6/t6ss_prediction.php
DNAplotter	Carver et al.[95]	https://github.com/sanger-pathogens/Artemis
Prism 10	GraphPad software	https://www.graphpad.com/series/prism-10-a-quick-glance
Code for mscRNA-seq analysis	This paper	Zenodo: https://doi.org/10.5281/zenodo.18344612; GitHub: https://github.com/WangycLab/SGM_code

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Participant recruitment

Eligible participants were recruited from the Department of Periodontics at the Stomatology Hospital, Zhejiang University School of Medicine. All participants provided written informed consent prior to enrollment. The study protocol was approved by the Ethics Committee of the Stomatology Hospital, Zhejiang University School of Medicine (Approval No. 2023-128 (R)). Demographic data, including age, gender, smoking status, and medical history, were collected for each participant. Clinical parameters were assessed, including probing depth (PD), bleeding on probing (BOP), clinical attachment loss (CAL), and plaque index (PI).[96]

Inclusion and exclusion criteria

Participants aged 20–65 years with clinically confirmed periodontal health or generalized Stage III/IV, Grade C periodontitis[97] were eligible. Exclusion criteria included pregnancy or lactation, systemic disease (e.g., diabetes, immunodeficiency), recent use (within 3 months) of antibiotics, bisphosphonates, corticosteroids, or NSAIDs, recent periodontal treatment (within 3 months), heavy smoking ($\geq$10 cigarettes/day), ongoing orthodontic therapy, or mental illness affecting compliance.

Immortalized human oral mucosal epithelial cells

Immortalized human oral mucosal epithelial cells derived from gingival mucosa (Creative Bioarray, Cat# CSC-I9117L) were cultured in high-glucose Dulbecco's Modified Eagle Medium (Gibco, Cat# 11594416) supplemented with 10% fetal bovine serum (Gibco, Cat# 16629525) at 37°C with 5% CO_2.

METHOD DETAILS

Sample collection

Microbiological subgingival samples were harvested following a previously describe.[98] In non-periodontitis subjects, subgingival samples were taken from the mesiobuccal and distobuccal sites of the first molars in all four quadrants. In subjects with periodontitis, subgingival samples were taken from the most accessible site, with the deepest PD and BoP, per quadrant. Samples were obtained using two sterile paper points per site.

Before sampling, participants rinsed thoroughly with sterile water, and supragingival plaque was gently removed using a sterile scaler or cotton pellet. Subgingival plaque was collected using sterile paper points inserted apically until light resistance ($\leq$25 N) was met and held in place for 30 s. For 16S rRNA sequencing and targeted metabolomics, samples were obtained from the same individual; paper points were stored separately in 1.5 mL Eppendorf tubes at –80°C until use, with one tube allocated for 16S rRNA analysis and the other for metabolic profiling. For single-bacterium RNA sequencing, paper points were transferred to 0.6 mL tubes preloaded with 200 μL of 4% PFA (Meilunbio, MA0192-2) and fixed at 4°C with gentle agitation for 30 min. The tubes were then incubated at 4°C with gentle agitation for 30 min. To maximize bacterial recovery, air was gently expelled over the paper points for 60 s, and the resulting liquid was transferred to a 1.5 mL sterile tube. Next, a small aperture was created at the center of the bottom of the 0.6 mL tube using a red-hot needle, allowing for rapid extraction by centrifugation at 4000 *g* for 3 min at 4°C. This extraction process was repeated twice. The collected liquid was transferred to a 1.5 mL sterile tube and subjected to three-step differential centrifugation (50 *g*, 100 *g*, 400 *g*, 5 min each, repeated three times) to remove impurities. The purified microbial suspension was then fixed overnight at 4°C using a rotary shaker.

Cell permeabilization

Following fixation, microbial suspensions were centrifuged at 4000 *g* for 15 min at 4°C. The supernatant was removed, and the cells were resuspended in 1 mL of PBS-RI (1× PBS supplemented with 0.1 U/mL NxGen RNase inhibitor). After a second centrifugation

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

under the same conditions, the bacterial count for each sample was adjusted to ~4 × $10^{\wedge 5}$ cells. The pellets were then resuspended in 0.04% Tween-20 in PBS and incubated for 3 min at 4°C. Cells were washed with PBS-RI twice, and then permeabilized in 50 μL of lysozyme mix (100 mM Tris-HCl, 50 mM EDTA, 0.25 U/mL RNase inhibitor, 2.5 mg/mL lysozyme) for 15 min at 37°C. Cells were washed three times with PBS-RI, and resuspended in 30 μL of nuclease-free water. Approximately 100,000 cells of each sample were collected for the next step. All reagents were obtained from the VITApilote-PFT1200 kit (R20115124; M20 Genomics).

In-cell reverse transcription and cDNA 3' poly-A tailing

Twelve random primers containing UMI sequences and pre-index were used for the RT reaction. Cells were evenly divided into 12 wells, each receiving 0.25 μL 100 mM dNTPs, 1 mL 5× reverse transcription buffer, 0.25 μL RNase inhibitor, 0.25 μL reverse transcriptase and 1 μL 10 μM pre-barcoded random primer. The RT reaction was conducted using ten cycles of multiple annealing (8°C for 12 s, 15°C for 45 s, 20°C for 45 s, 30°C for 30 s, 42°C for 2 min, 50°C for 3 min), followed by a 42°C incubation for 30 min in a thermal cycler. Post-RT reaction, bacteria were washed eight times with PBST (0.05% Tween-20 in PBS). The bacteria were then subjected to dA tailing by adding 39 μL cells in PBS, 5 μL buffer T2, 5 μL buffer T1, 0.5 μL TT enzyme, followed by incubation at 37°C for 25 min. After washing five times with PBST, cells were diluted and their concentrations were quantified. The reagents and primers were included in the VITApilote-PFT1200 kit.

Single-cell droplet generation

Single cell suspensions were prepared by dilution in a density gradient solution. Cells, 4× DNA extension reaction mix, and ready-to-use hydrogel barcoded beads were encapsulated into droplets using the microfluidic platform VITAcruizer DP400 system (E20000131; M20 Genomics) and microfluidic chips (E20000131; M20 Genomics). Droplets were incubated under the following thermal cycling conditions: 37°C for 1 h, 50°C for 30 min, 60°C for 30 min, and 75°C for 20 min. All reagents for droplet generation were obtained from the VITApilote-PFT1200 kit.

cDNA enrichment

After extension, droplets were broken with perfluorooctane, and the aqueous phase was purified using AMPure XP beads. qPCR was used to determine the optimal amplification cycle number based on the early exponential phase. Enriched cDNA was purified, eluted in nuclease-free water, quantified with a Qubit 2.0 fluorometer, and assessed using the Qsep100™ DNA Fragment Analyzer. All reagents and primers were from the VITApilote-PFT1200 kit.

Library preparation

Library preparation was performed using the VAHTS Universal DNA Library Prep Kit for Illumina V3 (Vazyme, ND607-03/04). Qualified cDNAs underwent end-repair and adenylation by incubation in a reaction mix containing 50 ng fragmented cDNA, end-repair buffer, end-repair enzymes, and nuclease-free water at 30°C for 30 min. The reaction was then inactivated at 65°C for 30 min. Adapter ligation was performed at 20°C for 15 min, followed by purification of the ligated DNA using AMPure XP beads. Library amplification was conducted, followed by DNA purification and quantification. Library quality was assessed using the Qsep100™ DNA Fragment Analyzer, evaluating fragment size distribution and integrity. High-throughput next-generation sequencing was performed on a NovaSeq 6000 platform using the S4 Reagent Kit.

Data quality control and filtering

The raw sequencing data was processed using VITAbasic (version 1.4; https://share.m20genomics.com/s/eGakzEodeH6Lznc). Adaptor sequences and extra bases generated by the dA-tailing step were trimmed. For each forward read, UMI (8 nts) and cell-specific barcode (20 nts) were extracted. Subsequently, barcodes that could be uniquely assigned to an accepted barcode with a Hamming distance of 2 nts or less were merged. Then, reverse reads shorter than 16 nts were discarded. The remaining clean reads were mapped to the human reference genome (GRCh38) using STAR[83] (version 2.7.10a). Reads aligning to the human genome were removed and excluded from downstream microbial analysis.

Construction of the non-redundant gene catalog

A total of 8,622 high-quality genomes with Prokka annotations were retrieved from HOMD. Protein-coding genes were extracted using BEDTools[99] (version 2.28.0). To reduce redundancy, sequences were clustered with CD-HIT-EST[82] (version 4.8.1; parameters: -c 0.95, -G 0, -T 32, -n 8, -aS 0.9, -M 0) at 95% identity to minimize redundancy while retaining strain-level diversity, a strategy commonly used in microbial gene catalog construction.[100,101] The resulting non-redundant protein sequences (~3.1 million genes) were functionally annotated using DIAMOND[84] (version 2.1.8; parameters: –evalue 10^{-4} –very-sensitive –strand both) against the COG database[102] (the 2020 release), resulting in functional annotations for 74.42% of the genes. This non-redundant gene catalog was used as a reference for mapping and quantifying single-cell transcriptomes in our study.

Gene abundance calculation and high-quality single-cell screening

The host-depleted clean reads were mapped to the non-redundant microbial gene catalog from the HOMD by STAR[83] (version 2.7.10a). Valid cells per sample were identified using the STARsolo *soloCellFilter* default method, which distinguishes high-confidence cell barcodes from background barcodes based on the inflection point of the UMI count distribution.[25,34,35] Only species

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

representing more than 1% of the total composition in each sample for downstream analysis. Genes lacking annotation were excluded, and only genes detected in at least three cells were retained. Following this, cells from all samples were integrated. High-quality cells were defined as those with both nCount and nFeature values within three times the median absolute deviation, while the remaining cells were excluded.

Preparation of the Kraken2-based gOTUs database

For the taxonomic identification of each single cell, a customized Kraken2-based gOTUs database were constructed using Kraken2[85] (version 2.1.3). The kraken2-build module in the Kraken2 software package was used (settings: –no-masking–add-to-library) to create a classification database based on genomes and NCBI taxonomic information. Furthermore, the k-mer distribution file was created with the Bracken –build module[86] (version 2.8; setting: -l 100).

Taxonomy determination

Clean reads with Read2 shorter than 60 nts were ignored, and then classified by Kraken2 against the customized gOTUs database. Cells were classified into seven phylogenetic levels (domain, phylum, class, order, family, genus and species) or unclassified. The adjusted read counts of all taxa were calculated using a custom script adapted from the Bracken (https://ccb.jhu.edu/software/bracken/) est_abundance.py script to attain compatibility with single-cell data. For each cell, we selected the taxon with the highest read count at each taxonomic level, proceeding downward until the species level was reached. If the read count fell below 50% of the parent node's read count at any level, the process halted and the cell was discarded in the further analysis. The final taxon was considered as a reliable classification result for the cell.

Filtering and benchmarking before clustering analysis

The single-cell gene expression matrix was imported into Seurat[88] (version 4.3.0) for subsequent analysis. A series of benchmarks were used to determine the optimal parameters for dimension and resolution in the clustering analysis. The dimension was determined by using the Seurat ElbowPlot function, which shows the standard deviations of top principal components. When the points fall on a plateau (where the standard deviation does not change much) after an inflection point, the corresponding principal component is selected as the best possible dimension value. The resolution was determined using the clustree[87] package (version 0.5.1), which generates clustering trees to interrogate clusters along with resolution increases. When cell clusters start to mix after a certain resolution, that resolution is selected as the best possible resolution value.

Cell clustering and functional cluster identification

After filtering and quality control, dimensionality reduction of all high-quality cells was performed using UMAP. Batch effects across samples were corrected using Harmony[89] (version 0.1). Clusters of cells were identified using the FindClusters function from Seurat (resolution = 0.4). Subsequently, hdWGCNA[36] (version 0.3) was applied to characterize co-expression patterns across the identified clusters, enabling the definition of functional modules and providing biological insights. DEGs were determined using the FindAllMarkers function in Seurat (average $\log_2$ (fold change) > 0.25, p-value < 0.05). Among the DEGs, genes with well-defined functional annotations were selected as marker genes for each cluster. When multiple genes indicated a specific functional category within a cluster, the cluster was named based on that predominant function.

Validation of taxonomic assignments using BLAST

The barcode lists of *A. johnsonii*, *T. denticola*, *T. vincentii*, and *L. mirabilis* in each sample were extracted from the Seurat object in R, and then corresponding clean reads were fetched using a custom script. Next, 50,000 reads were randomly selected from each set of reads using seqtk (version 1.3-r117), and mapped to the NCBI nt database using NCBI BLASTn[93] (version 2.13.0+) (settings: -max_target_seqs 1, -evalue 1e-5).

Re-clustering and pathway functional activity analysis

The cells of the functional cluster from *N. elongata*, *L. mirabilis*, *T. denticola*, *P. gingivalis*, and *P. intermedia* were extracted for re-clustering analysis and the genes that covered fewer than three cells of these two groups were removed. The rest of the clustering and functional cluster identification analyses were the same as described previously.

Pseudobulk differential expression analysis

We performed pseudobulk differential expression analysis using the *muscat* package[90] (version 1.12.1). For the pseudobulk comparison, raw gene counts were aggregated across all cells and normalized by library size. This analysis was conducted separately for each functional cluster comparisons. For interspecies comparisons, single-cell counts were subsetted for *T. denticola* and *P. gingivalis* and aggregated per species and sample to generate pseudobulk counts. The resulting matrix was transposed so that rows corresponded to genes and columns to pseudobulk samples. Differential expression analysis was performed using DESeq2, which normalizes counts by estimating size factors for each pseudobulk sample, accounting for differences in library size and ensuring that observed differences reflect relative expression levels rather than sequencing depth. Genes with $\log_2$(fold change) > 0.25 and $p \leq 0.05$ were considered differentially expressed.

Please cite this article in press as: Ding et al., Single-cell profiling of the subgingival bacteria reveals transcriptional heterogeneity and niche-specific programs, Cell Host & Microbe (2026), https://doi.org/10.1016/j.chom.2026.01.017

CellPress

Functional gene proportion calculation

The functional activity of each cell for a certain function was quantified by the FGP, which was defined as ratio of genes annotated to a specific functional category relative to the total number of annotated genes per cell.[28] specific functional category Functional categories were primarily determined based on COG and RefSeq annotation information.

Flow cytometric assessment of bacterial viability in subgingival plaque samples

The obtained subgingival plaques were processed immediately in PBS on ice. After enrichment of bacterial cells via stepwise centrifugation to remove host cells and debris, the suspension was divided into two equal fractions. One was treated with 75% ethanol for 1 h to generate a killed-cell population, while the other was kept in PBS as a live-cell control. A validation control was prepared by mixing the killed and live suspensions at a 1:1 ratio. Staining was performed using PI (Invitrogen) at a ratio of 1 μL per 10^6 bacteria in 1 mL, followed by incubation on ice for 15 min in the dark. Flow cytometry was performed using a Beckman Coulter CytoFLEX instrument, and data were analyzed using FlowJo software.

Isolation of *N. elongata* from healthy subgingival plaque

Subgingival plaque samples from healthy individuals were stored at 4°C, vortexed, and serially diluted in PBS to 10^{-6}. Aliquots (100 μL) were plated on chocolate agar with 5% defibrinated sheep blood and incubated under microaerophilic conditions (5% O_2, 10% CO_2, 85% N_2) at 37 °C for 3–5 d. Distinct colonies were re-streaked and cultured for 48 h. Genomic DNA was extracted, and the 16S rRNA gene was amplified using primers 27F and 1492R. PCR products were Sanger sequenced (Sangon Biotech), and taxonomic identity was confirmed by comparison to NCBI reference databases. Verified Neisseria elongata isolates were stored in 20% glycerol at −80 °C for downstream analyses.

Growth assay of *N. elongata*, *P. gingivalis*, and *P. intermedia* under microaerophilic and anaerobic conditions

N. elongata isolates were grown overnight on chocolate agar under microaerophilic conditions at 37 °C. Bacterial lawns were scraped and resuspended in Brain Heart Infusion (BHI) broth (Thermo Fisher Scientific, CM1135B) supplemented with 5% heat-inactivated fetal bovine serum (Gibco, 16629525), adjusted to $OD_{600} \approx 0.04$. Cultures were incubated in sterile flasks at 37 °C under both microaerophilic and anaerobic conditions (10% H_2, 10% CO_2, 80% N_2). OD_{600} was measured over 0–30 h using a microplate reader (BioTek Synergy H1). Growth curves were generated from three independent biological replicates, and statistical differences between conditions were evaluated by *Student' s t-test*.

For growth on solid media, lyophilized *N. elongata* was rehydrated and plated (0.1 mL) on chocolate agar (BeNa Culture Collection, 360798) consisting of special casein peptone, starch, NaCl, agar, and 5% defibrinated sheep blood (pH 7.3 ± 0.2). Plates were incubated at 37 °C under both microaerophilic and anaerobic conditions, and colony morphology was recorded.

P. gingivalis W83 was grown in BHI broth supplemented with 5 μg/mL hemin and 0.5 μg/mL vitamin K1. *P. intermedia* ATCC 25611 was cultured in Tryptic Soy Broth (30 g/L) supplemented with 5 g/L yeast extract, 0.5 g/L L-cysteine hydrochloride, 5 μg/mL hemin, and 1.0 μg/mL vitamin K1. Both species were incubated in an anaerobic chamber at 37 °C for 24–48 h.

Scanning electron microscopy

For ultrastructural analysis, *N. elongata* and *N. lactamica* were cultured overnight on sterile glass coverslips. Samples were fixed in 2.5% glutaraldehyde at 4 °C overnight, rinsed with PBS, and post-fixed in 1% osmium tetroxide for 1.5 h at room temperature. After graded ethanol dehydration (50%, 70%, 90%, 100%), specimens were dried using a critical point dryer (Leica CPD300), sputter-coated with platinum-palladium (6 mA, 80 s; Quorum Q150T ES plus), and imaged at 5 kV using a Nova Nano 450 SEM (Thermo FEI) at 2000× and 10,000× magnifications.

For host–microbe interaction imaging, immortalized oral epithelial cells (5×10^4) were seeded onto coverslips for 24 h, followed by co-incubation with either *N. elongata* or *N. lactamica* at a multiplicity of infection (MOI) of 200 for 4 h. Cells were washed, fixed with 2.5% glutaraldehyde, and processed identically as above.

RNA extraction, reverse transcription and qPCR

Total RNA was extracted from *N. elongata* cultured to mid-log phase ($OD_{600} \approx 0.5$) under microaerophilic and anaerobic conditions. Approximately 1×10^9 cells were lysed with lysozyme (0.2 mg/mL, 37 °C, 20 min), homogenized in RNAex Pro reagent (Accurate Biology, AG21101), and purified using Bacterial RNA Mini Columns (AG21023) per manufacturer' s protocol. RNA was eluted in RNase-free water and quantified using a OneDrop spectrophotometer.

For reverse transcription, 500 ng total RNA was converted to cDNA using the PrimeScript RT reagent kit with random hexamers (Takara, RR037A). Quantitative PCR was performed on a QuantStudio™ 7 Flex Real-Time PCR System (Thermo Fisher) with TB Green Premix Ex Taq (Takara). Reactions (10 μL) included 5 μL master mix, 0.4 μL primers (10 μM each), 0.2 μL ROX reference dyeII, 1 μL cDNA, and nuclease-free water. Cycling conditions were: 95 °C for 30 s, then 40 cycles of 95 °C for 5 s and 60 °C for 30 s, followed by melting curve analysis. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method, and visualized in Prism 10.

Genome analysis in *N. elongata*

Raw whole-genome sequencing reads underwent stringent quality control, including removal of adapor-contaminated and low-quality reads (Phred score ≤20, >40%). High-quality reads were de novo assembled using SPAdes[103] (version 3.11.1; -k 107, 117, 127).

Contigs with low coverage (<0.35x mean depth) or shorter than 500 bp were discarded. Coding genes were predicted using Prodigal (version v2.6.2), and then annotated by BLAST alignment (e-value $\leq$ 10-5) against Swiss-Prot[104] and NR databases and DIAMOND[84] alignment (version 2.1.8; –evalue 10^{-4} –very-sensitive) against the COG database[102] (the 2020 release).The genome was screened for T6SS gene clusters using the t6ss_finder.py tool,[94] which employs embedded BLASTp (version 2.11.0+) against the SecReT6 v3 database. Genes associated with adhesion, pilus, and nitrate reductase were identified based on Swiss-Prot, RefSeq and COG annotations. The genomic map was genearted using DNAplotter[95] (Artemis version 18.2.0), highlighting genes of interest through color-coded annotations.

Full-length 16S rRNA gene amplicon sequencing of subgingival samples

Total microbial genomic DNA was extracted from subgingival samples using the FastPure Stool DNA Isolation Kit (MJYH, shanghai, China). DNA concentration and purity were assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA), and integrity was checked by 1% agarose gel electrophoresis. The full-length 16S rRNA gene was amplified from the extracted DNA using the universal bacterial primers 27F (5'-AGRGTTYGATYMTGGCTCAG-3') and 1492R (5'-RGYTACCTTGTTACGACTT-3').[105] Amplified products were purified with AMPure PB Beads (Pacific Biosciences,USA), quantified with a Qubit 4.0 Fluorometer (Thermo Fisher Scientific, USA), and pooled in equimolar ratios. The SMRTbell library was constructed using the SMRTbell Prep Kit 3.0 (Pacifc Biosciences) and sequenced on the PacBio Sequel IIe platform (Pacific Biosciences) at Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China). High-fidelity (HiFi) circular consensus sequences (CCS) were generated using SMRT Link v11.0. HiFi reads were demultiplexed and filtered by length (1,000–1,800 bp). Denoising and Amplicon Sequence Variant (ASV) calling were performed using the DADA2[106] plugin within QIIME2[107] (V2020.2). Taxonomy was assigned using a BLAST consensus classifier in QIIME2 against the NT Taxon core v2024/16S bacteria database. All samples were rarefied to 6,000 sequences for subsequent analysis.

Targeted amino acid and metabolite profiling by UHPLC-MS/MS

Metabolites were extracted from subgingival samples with methanol. For derivatization, a 10 μL aliquot of the supernatant was dried under nitrogen, reconstituted, and reacted with 30 μL of 10 mg/mL dansyl chloride and 40 μL of 0.5 M carbonate-bicarbonate buffer (pH ~9.5) at 60°C for 30 min. Following alkali quenching and a second heating step, the mixture was acidified with formic acid, centrifuged, and the supernatant was collected for analysis.The analysis was performed on an ExionLC AD system coupled to a QTRAP® 6500+ mass spectrometer (Sciex, USA), using a Waters HSS T3 column (2.1 × 150 mm, 1.8 μm). A 15-min gradient of methanol/water (both with 0.1% formic acid) at 0.35 mL/min was used for separation. The ESI source was operated at 350°C and +5000 V, with curtain gas at 35 psi and both ion source gases at 60 psi.Quantification was based on a standard curve of 68 amino acids and derivatives using linear regression. Quality control was ensured by analyzing a pooled QC sample every 10 runs (RSD < 15%). Raw data were processed in Sciex OS software with automated integration and manual verification, and concentrations were calculated against linear calibration curves.

Metabolomic Profiling of *P. gingivalis* Culture Supernatants by LC-MS

P. gingivalis W83 was cultured anaerobically to the mid-exponential phase (OD_{600} ≈ 0.6-0.8). The culture supernatant was collected by centrifugation with 12,000 *g* for 5 min, sequentially filtered using 0.22 μm to remove cells, and stored at -80°C. Metabolites were extracted from the supernatant using cold methanol:acetonitrile (1:1). LC-MS analysis was performed using a UHPLC system (Vanquish, Thermo Fisher) coupled to an Orbitrap Exploris 120 mass spectrometer. Metabolites were separated on a UPLC HSS T3 column (2.1 × 100 mm, 1.8 μm) and detected in information-dependent acquisition (IDA) mode to collect MS/MS spectra. Raw data were processed using R for peak alignment and normalized by internal standard. Metabolites were annotated by matching MS/MS spectra to public databases. Significant metabolites were identified by VIP >1.0 and p <0.05 in orthogonal partial least squares-discriminant analysis (OPLS-DA) models validated by permutation tests.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Comparisons of subgingival microbiota counts obtained using conventional and optimized adsorption methods were assessed with a *paired t-test*. Differences in bacterial retention under varying centrifugation times were evaluated using *one-way ANOVA*. qPCR was performed for *N. elongata*, *P. gingivalis*, and *P. intermedia*, and relative expression levels were calculated using the 2 ^-ΔΔCt method. Differences between experimental conditions were assessed using a two-tailed Student's *t*-test. For microbial single-cell data, pairwise comparisons were conducted using the *two-sided Wilcoxon rank-sum test*, and multi-group comparisons were analyzed using the *Kruskal–Wallis test* followed by *Dunn's post hoc test*. Spearman's rank correlation analysis was employed to assess interrelationships between interested bacterial species and amino acids. All results are presented as mean ± SEM.