[BIOINFO 2025]

Expanding phylogenetic and clinical insights through a comprehensive genome catalog of the human oral microbiome

Jun Hyung Cha[1], Nayeon Kim[1], Junyeong Ma[1], Sungho Lee[1], Geon Koh[1], Sunmo Yang[1], Samuel Beck[2], Iksu Byeon[3], Byunguk Lee[3], and Insuk Lee[1,4]*

[1]Department of Biotechnology, College of Life Science and Biotechnology, Yonsei University, Seoul, Korea

[2]Boston University School of Medicine, Boston, Massachusetts, USA

[3]Korean Bioinformation Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, Korea

[4]DECODE BIOME Co., Ltd. Incheon 21983, Republic of Korea

*Corresponding author: insuklee@yonsei.ac.kr

Understanding the taxonomic and functional landscape of the human oral microbiome requires large, high-quality genomic resources. Here, we present the Human Reference Oral Microbiome (HROM), a catalog of 72,641 high-quality genomes spanning 3,426 species, including 2,019 previously undescribed species, which markedly enhances metagenomic read classification compared with existing references. HROM also reveals 1,137 candidate phyla radiation (CPR) species, establishing Patescibacteria as the dominant oral phylum and uncovering a distinct oral clade separate from environmental counterparts. Within this group, we identify a CPR subclade linked to periodontitis that improves disease prediction when combined with Porphyromonas gingivalis. Comparative analyses with gut microbiome reference genomes further demonstrate substantial taxonomic and functional divergence between oral and gut communities. Moreover, we detect 42 ectopic oral species whose abundance in the gut associates with intestinal, cardiovascular, and liver diseases, underscoring the systemic clinical relevance of the oral microbiota.