**[BIOINFO 2023]**

# Construction of comprehensive human oral microbiome catalog reveals characteristics of underestimated clades

Jun Hyung Cha[1] and Insuk Lee[1,*]

[1] *Department of Biotechnology, Yonsei University*
[2] *Department of Biotechnology, Yonsei University*
**Corresponding author: insuklee@yonsei.ac.kr*

The human oral microbiome is well-known for its strong associations with various oral diseases and a multitude of human health and well-being factors. Therefore, it is imperative to elucidate the hidden dark matter of the oral microbiome and its associations with human health. Despite the existence of reference databases of microbial genomes, such as the NCBI RefSeq database and the Human Oral Microbiome Database(HOMD), which are commonly used for profiling oral microbiomes from shotgun sequencing data, these reference databases predominantly consist of isolate genomes, which underrepresent non-isolate genomes residing in the human oral cavity.

To address this limitation, we constructed the Human Oral Microbiome Catalog (HROM) from various publicly available 7,956 metagenomic samples and isolate genomes originating from the human oral cavity. We applied a recently updated metagenomic binning pipeline to retrieve non-chimeric Metagenome Assembled Genomes (MAGs) to construct a comprehensive catalog of the human oral microbiome. HROM encompasses 145,149 microbial genomes, including 5,113 representative genomes and 76 million nonredundant protein sequences. HROM spans across 21 phyla, including the intriguing phylum Patescibacteria, an episymbiotic CPR(candidate phyla radiation) group. HROM excels in outperforming existing reference genomes in oral microbiome profiling, achieving a higher classification rate. Furthermore, genomic analysis of HROM has revealed definitive compositional differences in the human oral microbiome compared to the gut microbiome. Additionally, a comparative analysis of the gut and oral microbiomes using HROM has unveiled the biological traits of potentially transmissible bacteria from the oral to the gut. Our HROM will provide future researchers with a more comprehensive reference database for analyzing the human oral microbiome.