

[J Microbiol Methods](#). 2009 Dec;79(3):266-71. Epub 2009 Sep 29.

# Metagenomic study of the oral microbiota by Illumina high-throughput sequencing.

[Lazarevic V](#), [Whiteson K](#), [Huse S](#), [Hernandez D](#), [Farinelli L](#), [Osterås M](#), [Schrenzel J](#), [François P](#).

Genomic Research Laboratory, Geneva University Hospitals, Rue Gabrielle-Perret-Gentil 4, CH-1211 Geneva 14, Switzerland. vladimir.lazarevic@genomic.ch

To date, metagenomic studies have relied on the utilization and analysis of reads obtained using 454 pyrosequencing to replace conventional Sanger sequencing. After extensively scanning the 16S ribosomal RNA (rRNA) gene, we identified the V5 hypervariable region as a short region providing reliable identification of bacterial sequences available in public databases such as the Human Oral Microbiome Database. We amplified samples from the oral cavity of three healthy individuals using primers covering an approximately 82-base segment of the V5 loop, and sequenced using the Illumina technology in a single orientation. We identified 135 genera or higher taxonomic ranks from the resulting 1,373,824 sequences. While the abundances of the most common phyla (Firmicutes, Proteobacteria, Actinobacteria, Fusobacteria and TM7) are largely comparable to previous studies, Bacteroidetes were less present. Potential sources for this difference include classification bias in this region of the 16S rRNA gene, human sample variation, sample preparation and primer bias. Using an Illumina sequencing approach, we achieved a much greater depth of coverage than previous oral microbiota studies, allowing us to identify several taxa not yet discovered in these types of samples, and to assess that at least 30,000 additional reads would be required to identify only one additional phylotype. The evolution of high-throughput sequencing technologies, and their subsequent improvements in read length enable the utilization of different platforms for studying communities of complex flora. Access to large amounts of data is already leading to a better representation of sample diversity at a reasonable cost.

PMID: 19796657 [PubMed - in process]