



# SALIVA MICROBIOME CORE: AN ANALYSIS ON MG-RAST DATA

Rafaela Nishiyama<sup>1</sup>, Claudio Trigo<sup>1</sup>, Simone Oliveira<sup>1,2</sup>, Flávio Aguiar<sup>1</sup>, Rodrigo Jardim<sup>3</sup>



<sup>1</sup>Campinas University / Piracicaba Dentistry Faculty, Brazil

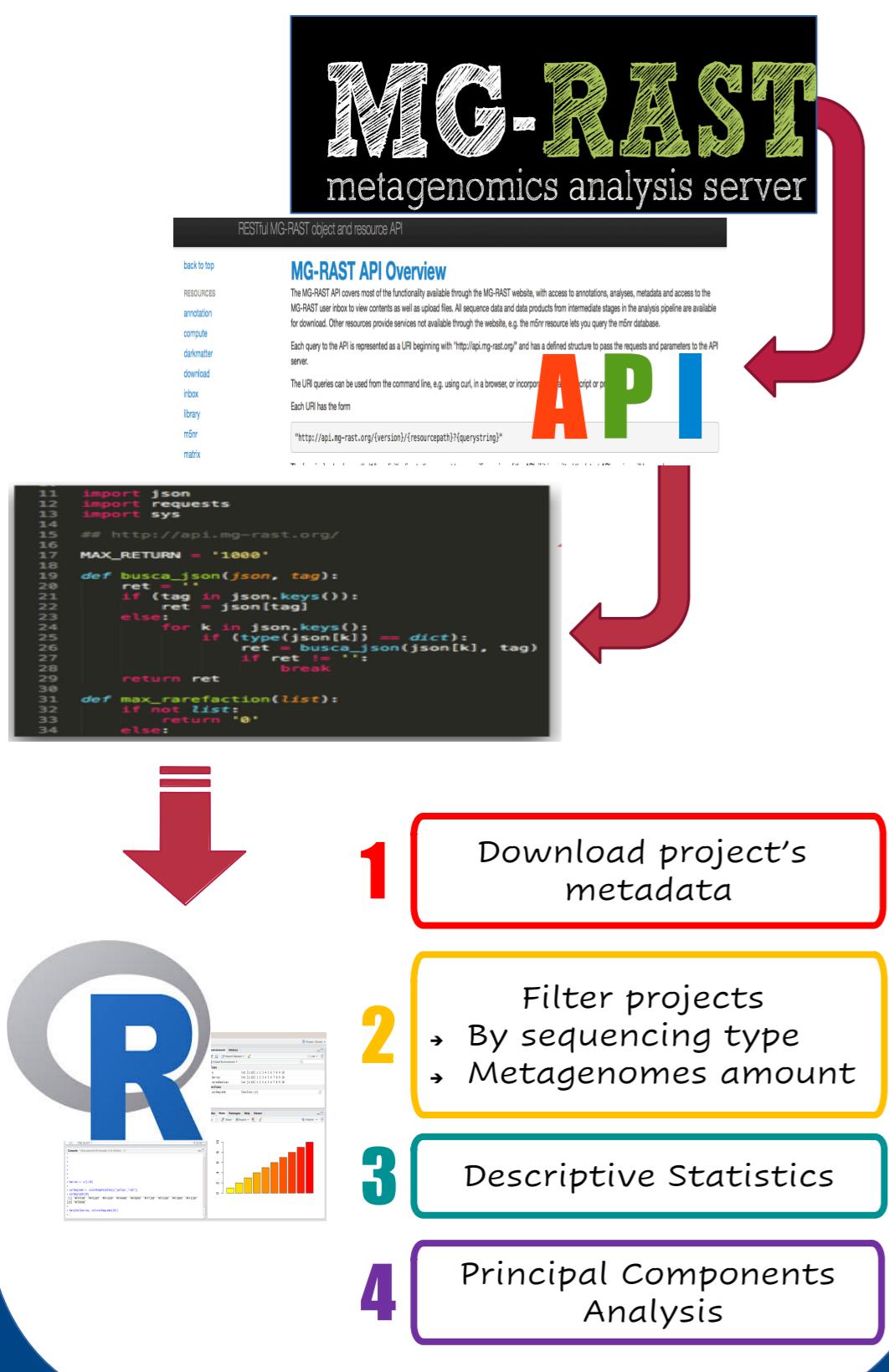
<sup>2</sup>Rio de Janeiro State University / Dentistry Faculty, Brazil

<sup>3</sup>Oswaldo Cruz Foundation / Oswaldo Cruz Institute, Brazil

## Abstract

Oral microbiota is considered the second most complex in the human body and is exposed to more than 700 species of microorganisms. Studies have proven the intrinsic relationship between the environment and its microbiota, showing that the interaction between these is highly synergistic. Identifying the core microbiome is essential for defining how "healthy" the study environment is. It is important to know, study and understand the relationship of this core with the environment and, thus, to predict diseases and changes in environmental behavior. This study aims to find the core of the oral microbiota of saliva samples, regardless of the conditions of the host of data deposited in public databases. Metagenomic samples were downloaded from Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST) using the search term SALIVA. A python script was developed to search for metagenomic data, using the Application Programming Interface provided in order to automate the process since the web interface only allows the download of 1 metagenome at a time. The script is available at GitHub. R software was used for statistical analysis. Descriptive statistical analyzes were performed on the numerical variables grouped by type of sequencing (16s or shotgun) to assess the statistical differences between two microbiome analysis techniques. Principal Component Analysis was carried out in taxonomic studies of metagenomes as a dimension reduction technique in order to find the core microbiome of each sequencing technique. The script was able to recover 474 metagenomics, of which 332 from 16s and 142 from shotgun. After statistical analysis this study found no significant difference between the types of sequencing for the predicted reads ( $W = 25096$ ,  $p\text{-value} = 0.2645$ ). For each type of sequencing it was possible to identify microorganisms belonging to the core of the salivary microbiota.

## Methods



## Results

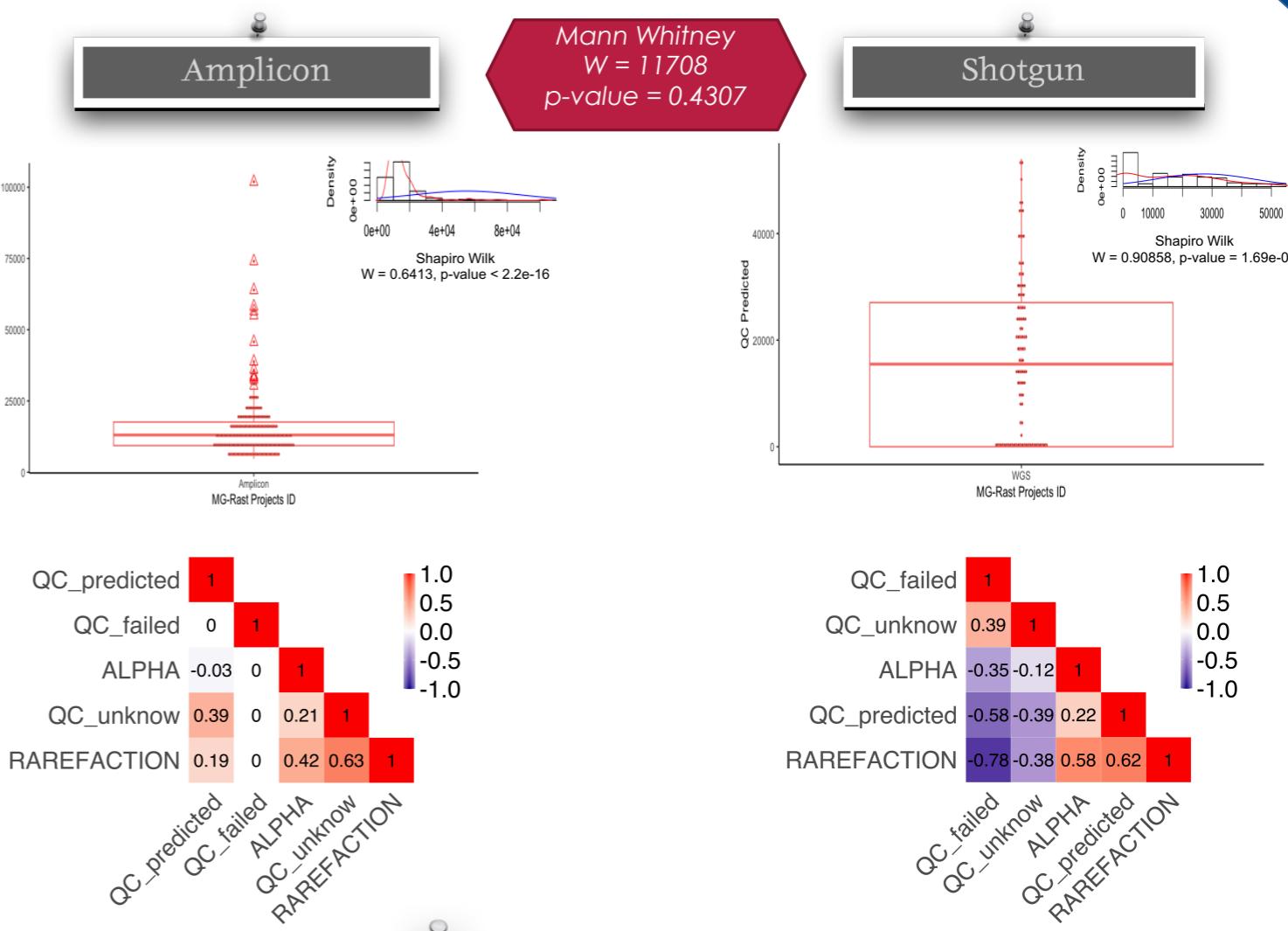


Figure 2: Correlation analysis of variables according to the sequencing method

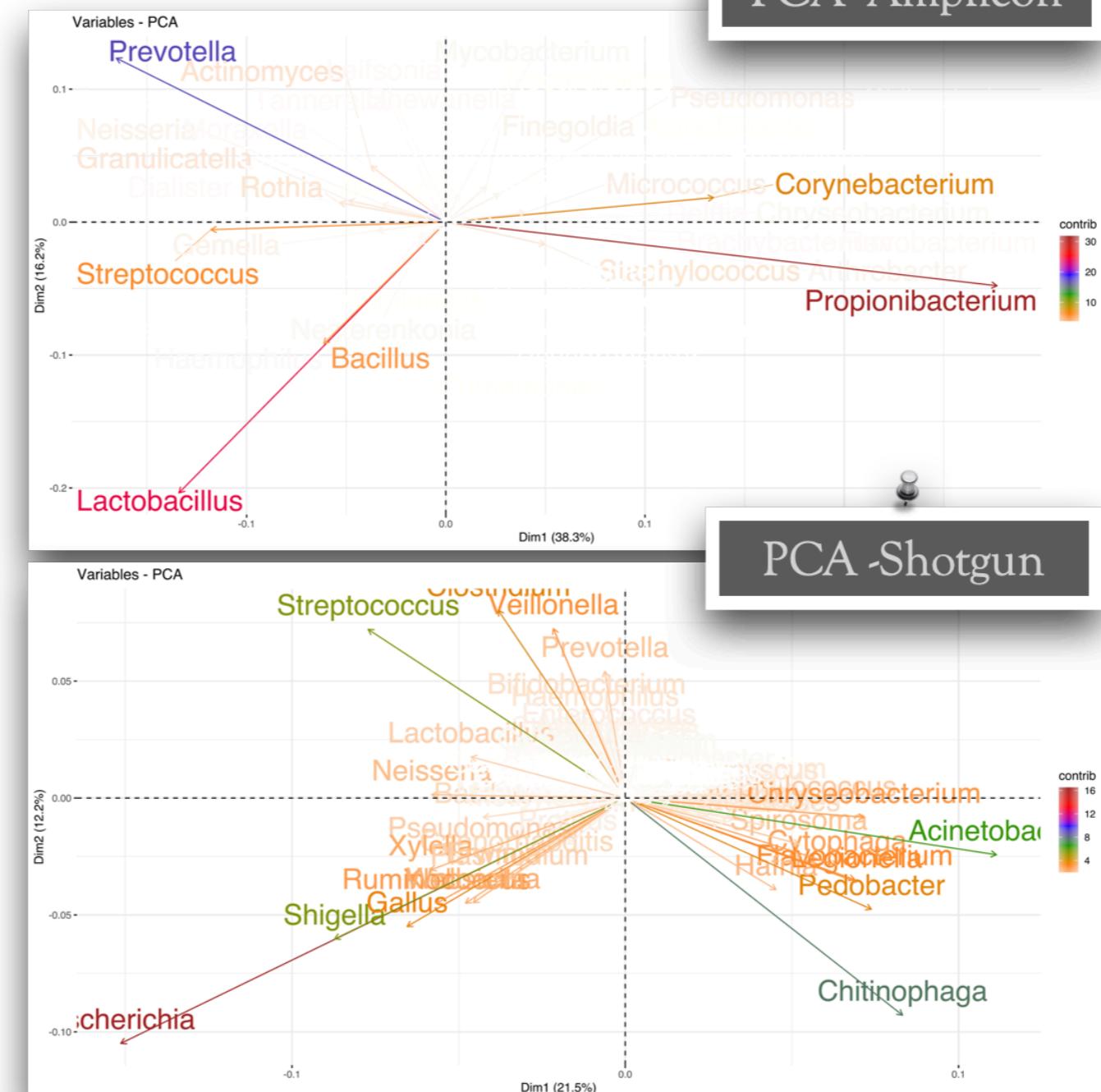
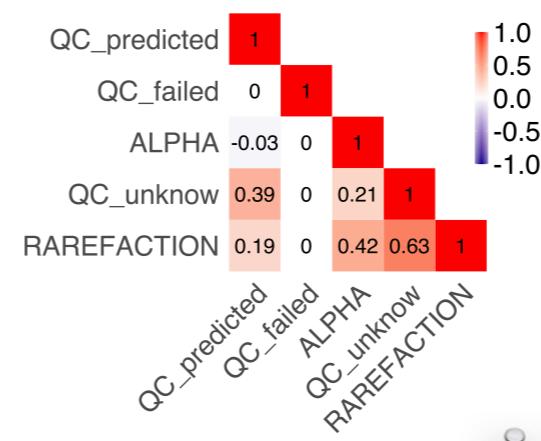


Figure 3: Saliva Microbiome Core (Principal Components Analyses)

## References

- Lamont RJ, Koo H and Hajishengallis G. The oral microbiota: dynamic communities and host interactions. *Nature Reviews*. 2018; 16: 745-759.
- Meyer F et al. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*. 2008; 9(1):386.
- Tierney BT et al. The Landscape of Genetic Content in the Gut and Oral Human Microbiome. *Cell Host and Microbe*. 2019; 26(2):283–295.e8.

## Acknowledgment

