

Title: Comparative Genomics of Bacterial Endophytes and others Functional Groups

Authors: Pablo R. Hardoim¹, Cristiane C. P. Hardoim²

Institutes: ¹ Univ Fed Rio de Janeiro, Inst Bioquim Med Leopoldo de Meis, Lab Biol Mol Plantas, Rio De Janeiro, RJ, Brazil; ² Univ Fed Rio de Janeiro, Inst Microb Paulo de Goés, Lab Bacteriol Mol Marinha, Rio De Janeiro, RJ, Brazil

Summary:

Comparative genomics is a powerful technique to identify functional elements accountable for species competence that enable it to thrive in specific environmental niche, and for species adaptation to implement particular lifestyles. Here, the abundance profile of identified genes was computed for bacterial genome of endophytes (n=40), nodule-forming symbionts (n=42), phytopathogens (n=29), rhizobacteria (n=42) and soil bacteria (n=49). Sequences from protein-encoding genes were assigned to KEGG Ortholog (KO) using the Integrated Microbial Genome comparative analysis system. A feature-by-sample contingency table, where properties with more than 25% abundance and samples within each group with less than 98% functional similarity, was created. The cumulative sum scaling normalization procedure was applied to assigned KO and then a mixture model that implements a zero-inflated Gaussian distribution was computed to detect differentially abundant properties with *metagenomeSeq* package. Protein-coding genes involved in chemotaxis and motility of bacteria suggested that endophytes are more responsive in the use of aspartate/maltose (Tar) and dipeptides (Tap) than rhizobacteria. Genes putatively involved in the regulation of two-component response systems of antibiotic resistance (*evgS/evgA*), redox-responding (*regB/regA*), nitrogen fixation and metabolism (*ntrY/ntrX*), and cell fate control (*pleC/pleD*) are more prominently found among endophytes than among phytopathogens and rhizobacteria. The genes putatively involved in the transcriptional regulation of nitrogen assimilation (*nifA*), reduction of nitric oxide (*norR*), a carbon storage regulator (*sdiA*) among others are detected in a significantly larger proportion among endophytes. Genes encoding glutathione peroxidase (*btuE*), glutathione S-transferase (*gst*), catalase (*katE*) and nitric oxide reductase (*norR*) are also enriched in endophyte genomes. Surprisingly, 28% of the investigated group of prokaryotic endophytes harbored the nitrogenase (*nifH*) gene, indicating that it has an important function to improve plant productivity. Likewise, distinct metabolic functions were enriched for the others plant-associated communities. There was no particular genomic feature that was common to all genomes in each investigated lifestyle, suggesting that multiple, rather than unique, key features are deployed by the symbionts as strategy to interact with the host plant.

Key words: Microbiome, Endosphere, Functional properties, Mutualism, Parasitism

Foment agency: CNPq, CAPES