

**TITLE:** GENOMIC FEATURES OF MUTUALISTIC PLANT BACTERIA

**AUTHORS:** HARDOIM, P. R.<sup>1</sup>; HARDOIM, C. C. P.<sup>2</sup>

**INSTITUTIONS:** <sup>1</sup>Laboratório de Biologia Molecular de Plantas, Instituto de Bioquímica Médica Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Av. Carlos Chagas Filho 373, CCS, Rio de Janeiro, 21941-599 RJ, Brazil [phardoim@gmail.com](mailto:phardoim@gmail.com); <sup>2</sup>Laboratory of Host-Microbe Interactions, Biosciences Institute, Coastal Campus, São Paulo State University (UNESP), Praça Infante Dom Henrique s/nº, 11330-900, São Vicente, São Paulo, Brazil.

**ABSTRACT**

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. It also allows insight into genomic island arising from genomic rearrangements. Here, the abundance profile of identified genes, protein families, metabolic pathways and regulons were computed for endophytes (including nodule-forming plant symbionts), rhizosphere bacteria and phytopathogens. To avoid bacterial species redundancy, of genome sequences with more than 96.5% for gANI and an alignment fraction (AF) more than 0.6 were computed as an intra-species cluster and only one representative of each cluster was further used. For each genome, sequences from protein-encoding genes were assigned to KEGG Ortholog and a feature-by-sample contingency table, where properties with more than 15% abundance in at least one assigned community was created. The assigned KO were normalized with the cumulative sum scaling (CSS) normalization and a mixture model that implements a zero-inflated Gaussian distribution was enumerated and computed to detect differentially abundant properties with *metagenome Seq* package. The lifestyle of endophytes was characterized by significantly overrepresentation of genes encoding for nitrogenase as well as genes involved in the uptake of urea cycle components. The genomes of assigned endophytic bacteria revealed distinct signaling features that differed from those detected among rhizosphere bacteria and phytopathogens. Similar results were also observed for genes encoding proteins involved in transport and secretion systems as well as for transcriptional regulators. Genes involved in chemotaxis receptors are more abundantly represented among phytopathogens than endophytes. Likewise, distinct metabolic functions were enriched for the others plant-associated communities. There was no particular genomic feature that could inhabit 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 statically.

**KEYWORDS:** Bacteria-plant interactions, Endophytes, Rhizobacteria, Phytopathogens, Functional Genomics

**DEVELOPMENT AGENCIES:** CNPq