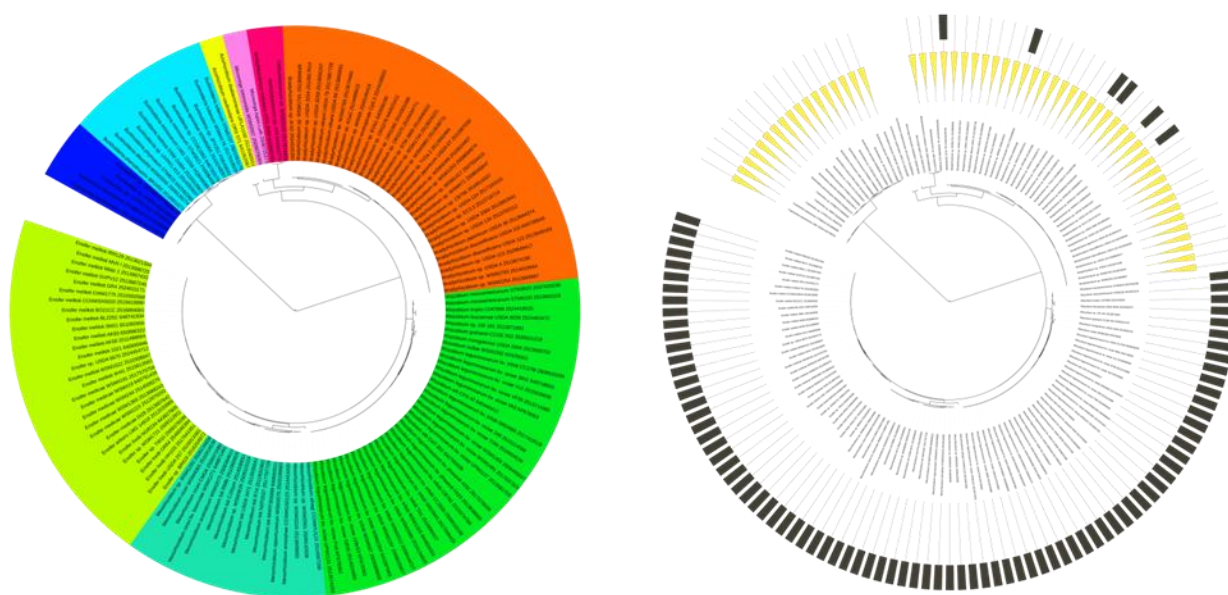


## Comparative Genome Analysis of Diverse Rhizobial Strains Shows Specific Genomic Features Are Associated with Symbiotic Interactions with IRLC Legume Clades

Julie Ardley<sup>1</sup>, Janet Sprent<sup>2</sup>, Wan Adnawani<sup>1</sup>, Rekha Seshadri<sup>3</sup>, Natalia Ivanova<sup>3</sup>, Tanja Woyke<sup>3</sup>, Nikos Kyrpides<sup>3,4</sup>, John Howieson<sup>1</sup> and Wayne Reeve<sup>1</sup>

<sup>1</sup>Centre for Rhizobium Studies, Murdoch University, Murdoch, WA 6150, Australia, <sup>2</sup>Division of Plant Sciences, University of Dundee at JHI, Dundee DD2 5DA, UK, <sup>3</sup>DOE Joint Genome Institute, Walnut Creek, CA, 94598, <sup>4</sup>Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia.

Recent rhizobial genome sequencing projects, including the GEBA-RNB project, have now enabled large-scale comparative genome analyses. This, combined with extensive host metadata, has allowed the identification of genomic markers and characteristics associated with distinct rhizobial phylogenetic groups and patterns of symbiotic association. In this study we screened 5,665 Clusters of Orthologous Groups in 143 phylogenetically diverse sequenced rhizobial strains. A number of molecular functions and biological processes with demonstrated or putative roles in symbiosis were mapped to distinct rhizobial phylogenetic groups, including BacA, proteins involved in hopanoid, vitamin, plant hormone and nitrogenase biosynthesis and key enzymes involved in lysine metabolism and the CBB and TCA pathways. Specific rhizobial functional groups were associated with particular legume symbioses, notably with the Fabaeae and Trifolieae clades within the IRLC legumes. We hypothesise that in these legumes a tighter symbiosis has evolved, in which the plant host has greater control.



**Figure 1.** Left: 16S rRNA NJ phylogenetic tree of 143 rhizobial strains belonging to nine genera. Right: hopanoid biosynthesis (yellow triangle) is found in *Burkholderia*, *Cupriavidus*, *Bradyrhizobium* and *Methylobacterium*, whereas *bacA* (black rectangle) is present in *Ensifer*, *Mesorhizobium*, *Rhizobium* and a few *Bradyrhizobium* strains.

This work was supported by DOE OBER and University of California DE-AC02-05CH11231