

## PS5-4 Genome features of an annual and a perennial clover microsymbiont

Lambert Bräu<sup>1</sup>, Lynn A. Goodwin<sup>2</sup>, Graham O'Hara<sup>1</sup>, Ravi Tiwari<sup>1</sup>, John Howieson<sup>1</sup>, Ron Yates<sup>3</sup>, and Wayne Reeve<sup>1</sup>

<sup>1</sup>Centre for Rhizobium Studies, Murdoch University, Murdoch, WA 6150, Australia; <sup>2</sup>US Department of Energy, Joint Genome Institute, Walnut Creek, CA 94598, USA; <sup>3</sup>Department of Agriculture and Food, Western Australia, South Perth, WA 6151, Australia

*Rhizobium leguminosarum* bv. *trifolii* is one of the most utilized species of root nodule bacteria, required for symbioses with annual (eg. *Trifolium subterraneum*) and perennial (e.g., *T. pratense*, *T. repens*, and *T. polymorphum*) clover species. Most clover rhizobial inoculants only form effective nitrogen-fixing symbioses with either annual or perennial species (and very few with both). This constraint provides a considerable barrier to agricultural productivity, because background populations of *R. leguminosarum* bv. *trifolii* may nodulate with an incompatible host and then fix nitrogen only poorly [1]. Knowledge at the genetic level is essential to develop an understanding of this incompatibility and progress in this pursuit will be greatly enhanced by complete genome sequence information. Thus, the genomes of two *R. leguminosarum* bv. *trifolii* strains were sequenced by the US Joint Genome Institute Community Sequencing Program; the Mediterranean *Trifolium* spp. isolate WSM1325 and the south American *Trifolium polymorphum* isolate WSM2304. Strain WSM1325 is compatible with Mediterranean perennial clovers (i.e., *T. pratense*) but not with American or African perennial clovers, while the reverse is true for strain WSM2304. To sequence the genomes, a shotgun assembly approach was adopted to assemble a draft genome of each organism. Three libraries were constructed for each strain, including one library for each strain in the functional genomics vector pTH1522 [2]. Assembly of the sequence data is in the draft stage and reveals that the genome of WSM2304 is 6.8 Mb in size (27 contigs of 20 reads or greater), contains a G+C content of 61.1%, and encodes 6487 candidate protein-encoding genes. In comparison, the WSM1325 genome is 7.8 Mb in size (143 contigs of 20 reads or greater), contains a G+C content of 60.7%, and encodes 7528 candidate protein-encoding genes. We will discuss comparative studies performed on the genome sequences of these two strains.

[1] Yates *et al.* (2008). *Soil Biol. Biochem.* 40:822-833.

[2] Cowie *et al.* (2006). *Appl. Environ. Microbiol.* 72:7156-7167.