**Title:** An analysis of phylogenesis and phytochemical diversity in *Scutellaria* to identify species with medicinal potential

**Authors:** Bryce Askey, Yeong Hun Song, Yoonkyung Lee, Ru Dai, Sangtae Kim, and Jeongim Kim

*Scutellaria baicalensis* is a medicinal plant whose root extracts have been widely used in Asian medicine for more than 2,000 years. Flavonoids in these extracts possess anti-cancer, antioxidant, antiviral, and neuro-psychologic properties. Because of interest in the medicinal properties of these flavonoids, the reference genome of *S. baicalensis* has been recently completed. Although *S. baicalensis* has been well-studied, the majority of the other over 470 species in the genus have not been analyzed. This study aims to extend the current knowledge of *Scutellaria* andto identify “medicinal hotspots” – clades of species accumulating high concentrations of medicinally relevant flavones. We analzyed flavonoid profiles for a large set of species representing the diversity in the genus. Targeted metabolite analysis of fresh and dried samples was completed with High Performance Liquid Chromatography (HPLC) to quantify 15 flavones. A phylogenetic tree of these species was constructed from chloroplast genome. Based on these results, a subset of species were selected for organ-specific metabolite profile analysis and genome size analyses. A comparison of phytochemical profiles for these species indicates significant diversity in site and identity of flavone accumulation when compared to *S. baicalensis*. Inconsistences in site of accumulation between the proposed flavonoid biosynthesis pathway for *S. baicalensis* and chemical analysis results imply significant differences in the pathway for several species. Flow cytometry results revealed 7 of the 8 species analyzed have similarly sized genomes as *S. baicalensis*. To better understand the cause of these inconsistencies, three of these species and *S. baicalensis* were selected for additional genomic and transcriptomic analysis. After genome sequencing of each of the species, RNA sequencing was used to identify the genes responsible for biosynthetic pathway diversity. Researchers studying *Scutellaria* can use the results of this work to identify previously unstudied species with medicinal potential. Additionally, the comparative metabolite and phylogenetic analysis used will help researchers to better understand the evolutionary history of the *Scutellaria* genus. Finally, identification of genes controlling inter-specific variation in the medicinal pathway can aid in biotechnological development of flavonoid production systems.

**Word count:** 311