

SPINGO: a rapid species-classifier for microbial amplicon sequences

Guy Allard¹, [Feargal J Ryan](mailto:feargalr@gmail.com)¹, Ian B Jeffery¹ and Marcus J Claesson¹

¹ School of Microbiology and Alimentary Pharmabiotic Center, University College Cork, Cork Ireland.
Email: feargalr@gmail.com

Project Website: <https://github.com/GuyAllard/SPINGO>

Source Code: <https://github.com/GuyAllard/SPINGO>

License: GNU General Public License (<https://www.gnu.org/copyleft/gpl.html>)

Taxonomic classification is a corner stone for the characterisation and comparison of microbial communities. Currently, most existing methods are either slow, restricted to specific communities, highly sensitive to taxonomic inconsistencies, or limited to genus level classification. These weaknesses mean that these methods may not uncover crucial microbiota information that can be obtained through a high-resolution analysis of the data. It is therefore imperative to increase taxonomic resolution to species level. In response to this need we developed SPINGO, a flexible and stand-alone software dedicated to high-resolution assignment of sequences to species level using 16S rRNA gene regions from any environment. SPINGO compares favourably to other methods in terms of classification accuracy, and is as fast or faster than available tools that have higher error rates. We also demonstrated its flexibility by successfully applying SPINGO to cpn60 amplicon sequences, demonstrating its ability to identify other types of target genes. SPINGO is an accurate, flexible and fast technique for taxonomic assignment down to the species level. This combination is important for the rapid and accurate processing of ever larger amplicon datasets generated by high-throughput next generation sequencing technologies.