

# Assessing the problem of doing proteomics with unsequenced organisms

LAB ROTATION PROJECT IN BIOINFORMATICS

Róger Bermúdez-Chacón\*  
Supervisor: Katja Bärenfaller†

June 10, 2013

## Project description

When dealing with unannotated or partially annotated proteomes, protein identification is often carried out by searching a database containing protein sequences from other species. Identification of a protein from a different species is often assumed as enough evidence to claim that this protein was identified in a protein sample. This work attempts to systematically evaluate at what extent this holds true, by comparing search results of unannotated proteins extracted from the Cassava (*Manihot esculenta*) root against different annotated databases, including the reference species *Arabidopsis thaliana*, the Viridiplantae database (containing a large number of green plants), and the annotated proteome of Cassava itself.

---

\*D-INFK - Computational Biology and Bioinformatics Master program

†D-BIOL - Plant Biotechnology Group