Understanding the Results

The results can are in CSV format, and can be viewed with any spreadsheet specific software, such as LibreOffice Calc, Microsoft Excel or Google Sheets.

“Main Score” and “Main Prediction” values should be used to evaluate the final priority order of putative secreted proteins in the genome of interest.

**Protein Identification:** The first three columns represent the available id linkouts for the protein.

**Main Score:** This is the output of the main neural network, integrating the results from Phylogenetic Profiling and Protein Sequence features.

**Main Prediction:** Applies an arbitrary threshold in order to discriminate between secreted and not secreted proteins.

**Phylogenetic Profiling Score (sequence independent):** This is the output of a neural network, integrating the results from Phylogenetic Profiling without Protein Sequence features.

**PP Prediction:** Applies an arbitrary threshold in order to discriminate between secreted and not secreted proteins.

**Score:** The Mutual Information between the Phylogenetic Profile of a given protein and The pattern of presence/absence of a specific Secretion System. Intuitively, it measures the dependence of two profiles, higher is better.

**Total Orthologs evaluated:** The number of orthologous protein sequences in the OMA DB for this particular sequence. Intuitively, this represents the support for any given score.

**MI/mean MI:** The ratio between the score of this profile and the mean of every other profile with +/- 10% difference in the number of elements in the profile (Total Orthologs Evaluated). This represents how well a given profile scores when compared to similar protein profiles, higher is better.

**Sequence based ML Score:** All proteins were scored using Secretion System specific(T3SS/T4SS/T6SS) sequence based machine learning methods in order to quantify the similarity to already known secreted proteins.

**GO term:** GO term associated with this protein sequence, the description of this terms can be found at [http://www.geneontology.org/](http://www.geneontology.org/).

**Description:** Description associated with this protein sequence.

For more information, please refer to the original paper: XXXXXXXXXX