# Towards NERVE 2.0: creation and integration of the NERVE database

Mazzocco G, Righetto I, Bertoldo D, Bernante F ,Filippini F

## Motivation

Starting from its milestone work publication in 2000, Reverse Vaccinology (RV) has aroused lively interest and represents one of the best examples of how Bioinformatics can boost Molecular Immunology. Our team developed NERVE (New Enhanced Reverse Vaccinology Environment), the most recent and unique software environment able to automate in silico steps for identifying Vaccine Candidates (VCs) from whole proteomes of bacterial pathogens. NERVE results are achieved integrating the output of several different predictor modules (including e.g. PSORT, SPAAN, etc. ) each considering a particular feature of the protein (e.g. subcellular localization (SCL), number and position of transmembrane (TM) regions, conservation, etc). This approach allows to save time and money by restricting next "wet" analyses to a narrow selection of predicted best VCs to be experimentally validated.

## Methods

NERVE pipeline was built using PERL; the pipeline coordinates serial launch of NERVE modules as well as flowing of data to and from an SQL database. Final output integration then results in ranking. NERVE modules consist of two types of scripts, which either perform specific prediction/analysis tasks or launch optimized versions of well known bionformatic algorithms and then perform results retreval. Among basic tools used to built up the NERVE, BLAST was used to scan for autoimmunity candidate regions and for conservation among strains (autoBLAST module). Also the adhesin prediction software SPAAN and the well known PSORT b SLC predictor for bacteria are used by NERVE as pivotal module elements. The HMMTOP software originally used to predict TM number and position is replaced in the novel NERVE version by TMHMM 2.0; the PSORTb version 2.0 is replaced by 3.0. Finally, script aimed at scanning motifs and profiles is original, written in PERL and it follows PROSITE rules for defining regular expressions (patterns) and position specific matrix-based profiles, and the NERVEdb is a multi-FASTA, flat file.

## Results

In order to enrich the subpopulation of VCs related to already known protective antigens (i.e. proteins the antigenicity of which is already validated) we developed NERVEdb, a manually curated and literature supported knowledgebase containing

MOLBINFO unit, Department of Biology, University of Padua, Padova

roughly 250 protective antigens (PAs) from different prokaryotic pathogen species and strains. Proteomes included in this database are from both Gram+ and Gram- bacteria. The role of NERVEdb within NERVE is to allow for comparison of pathogen proteins and a list of experimentally known PAs. This will provide such proteins with a score depending on similarity to those PAs. A new module based on scanning for typical patterns and profiles characterizing PA classes and pathogen surface proteins was also developed. Improved ranking is obtained by integrating of the outputs from novel modules with those from already existing modules. Last but not least, update of PSORT and replacement of HMMTOP allow for a better prediction of SCL and makes NERVE dependencies 100% open source.

**Contact e-mail**
mazzocco1983@yahoo.it