

Title: Development and application of algorithms of data-mining and classification for the reconstruction of neurodegenerative diseases related networks

Directors: Macha NIKOLSKI (LaBRI) and Klaus PETRY (Inserm U1049)

Keywords: classification, data-mining, peptides, inflammatory processes

Short description:

Neuroinflammation is associated with a large spectrum of trauma or disease related lesions—stroke, traumas, multiple sclerosis, Alzheimer and Parkinson diseases, tumors, etc. Inflammatory regions are characterized by alterations of the molecular phenotype of the luminal side of blood vessels, reflecting functional alterations of the Blood Brain Barrier (BBB).

The study of these alterations for the identification of bio-markers will be helpful for the in vivo monitoring of the apparition and progression of neuroinflammation and a better understanding of the inflammatory processes in a physiological frame.

In vivo screening of phage displayed combinatorial libraries of peptides produce molecular repertoires enriched for molecules targeting the inflammatory lesions, which can be physically isolated by subtractive hybridization techniques or studied in silica by comparison of massively parallel sequencing datasets.

The identification of the proteins mimicked by the selected peptides is a key step for the understanding of the physiology of the inflammatory process. We develop an approach based on peptide/protein sequences homologies to identify mimicked proteins. Prior knowledge of molecular functions, participation to cellular or systemic processes, cellular localization and interaction with other molecules as well as the participation to inflammatory processes, their development or resolution and the expression of these proteins from cell types implicated in such processes (endothelial cells or leucocytes) will help building a more accurate molecular model of the inflammatory process.

A pilot study based on an in vivo screening in Experimental Autoimmune Encephalitis (EAE) affected rats allowed the identification of peptides homing in vivo in experimentally induced inflammatory lesions. One of these was used to show the specificity for inflamed regions within the Central Nervous System and the ability for homing nano-cargos which could be loaded with bioactive substances.

The extension of this work depends on the automation of the tasks necessary for the annotation of the selected peptides aggregating information from a variety of sources: GEO for expression profiles, GO for known functions, Interactome databases to identify molecular partners, KEGG pathways for visualization on signal transduction pathways, PubMed for relevant literature, etc.

Classification, clustering and enrichment algorithms need to be developed to make robust and trustful this process. We aim to define a statistical model that will provide a scoring scheme such that peptides can be chosen according to their scores for further validation at the wet lab as well as the obtention of an overview of the biological processes in play and their representation as annotated graphs, a draft for the biological model, based on already available visualization tools, adapted as needed to the particular problem. In parallel, a knowledge database assisted by surveillance tools able

to feed it newly available information should be setup.
These tools will help fundamental research but also be a substratum for the improvement of design of preclinical tests on animal models of human pathologies. Such studies are of particular interest for the stratification of patients cohorts, allowing the adoption of the best of therapeutic strategies.

Funding: 3 year scholarship from the French Research Ministry (1650 Euros gross income/month) + possibility to be a lecturer at the University
Doctoral school: Mathematics and Computer Science Department, University Bordeaux 1
Starting date: september 2012
Contact: send your resume and cover letter to Macha Nikolski (macha@labri.fr) and Antoine Vekris (avek@mac.com)