

# ERCIM “Alain Bensoussan” Fellowship Scientific Report

Fellow: Ana Capatana  
Visited Location: NTNU  
Duration of Visit: August 27, 2009 - May 27, 2010

## **I - Scientific activity**

### **Academic Activity**

In June Professor Henning Christiansen and me visited NTNU. We had a presentation in the Bioinformatics & Gene Regulation group of Professor Finn Drabløs at Department of Cancer Research and Molecular Medicine, NTNU. The presentation was about my research experience at RUC.

During my stay at NTNU I had 2 more presentations at bioinformatics group meetings. First I presented the method (Markov Clustering) I planed to work with and some indexes of evaluation. The second presentation was concerning a hypothesis on shared evolutionary origin of small nucleolar RNAs and miRNAs.

### **Research**

The initial scientific plan of my research stay was to apply methods from the project at RUC to relevant problems at NTNU. However, this turned out not to be feasible due to the computational complexity. Therefore a revised project plan was made, with focus on detection and identification of non-coding RNAs.

As data set was chosen a set of small RNA, miRNA and intronic RNA sequences. The small RNA and intronic RNA sequences were extracted using the mapped coordinates from University of California at Santa Cruz mm8 assembly of the mouse genome; miRNA annotation was taken from miRBase version 10.0.

RNA secondary structure was generated by RNAfold, and based on RNA secondary structure a distance matrix was generated.

The Markov clustering method was used in order to perform a cluster analysis of the given set of non-coding RNAs. Cluster analysis was made using different peculiarities of different RNA types used in classification. Further it was performed evaluation of the clustering algorithm for non-coding RNAs using different indexes, as biological homogeneity, separation and accuracy of the clusters and different groupings of the sequences. In order to confirm the cluster composition a bootstrapping was performed for all clusters.

Within the analysis different non-coding RNAs properties were used, and different algorithms were implemented.

## **II- Publication(s) during your fellowship**

Publication of relevant results will be considered.

## **III -Attended Seminars, Workshops, and Conferences**

### **Seminars at NTNU**

**The LoSt Project Mini Workshop September 29-31, 2009**