ERCIM "Alain Bensoussan" Fellowship Scientific Report

Fellow: Ana Capatana
Visited Location : DANAIM, Roskilde University
Duration of Visit: October 28, 2008 - July 27, 2009

I - Scientific activity

Academic Activity

I made a presentation to introduce myself in the second week of arrival at Roskilde University. The presentation was about my last years of research experience. The meeting took place in the meeting room of department with about 6 persons attended. The presentation took about 30 min.

During my stay at RUC I have studied Artificial intelligence (emphasizing PROLOG and probabilistic models) and Bioinformatics (role of bioinformatics, implication of different aspects of bioinformatics in gene finding and architecture of gene finders).

Until today, many gene finder programs have been developed. There are two main types of gene finders: based on statistical properties of gene and based on genes' similarity.

The main problems of a genefinders are: ORFs are not equivalent to CDSs, gene prediction programs find new genes that share properties with a given set of genes.

In order to identify all potential genes in raw genome sequences, might be used many tools. Local alignment will compare target DNA/protein sequences with subject DNA/ proteins sequences.

As an alternative option can be used global alignment in order to find regions of local or global similarity between protein or DNA sequences. This way can be revealed functional and evolutionary relationships between sequences and identified members of gene families.

More important is to find the possible function of the studied protein. Sequence analysis methods can put in evidence gene function based on similarity between proteins of unknown function and proteins of known function. The known functions can be achieved from experimental results from molecular biology and genetic studies on model organisms. So, having a certain set of genes with known functions can be established the putative genes by orthology with the similar or the same functions.

There are also many PhD students who are working on gene finding. They are implementing different methods and developing different models.

Research

The overall topic of my research is focused on implementing different methods and aspects of gene structure, and finding some new potential genes.

I have been working on conservation based gene finding, because it helps to identify and reveal conserved information found between related organisms. In order to highlight some potential URF (unassigned reading frame) was developed flow chart data (fig1). There were used different genes properties, characteristics and were implemented algorithms. The conclusion of finding a new gene can be made after experimental tests (proofs).

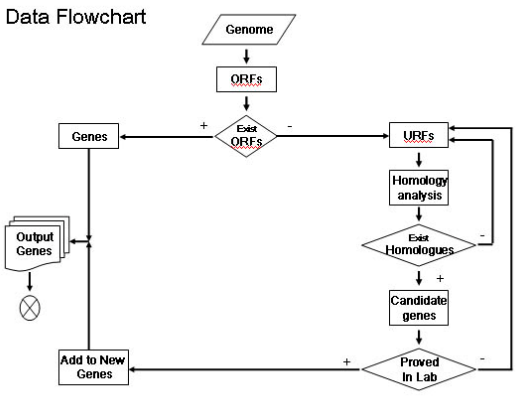

Fig.1 Flowchart data.

II- Publication(s) during your fellowship

A forthcoming paper currently under production (joint with Prof. Henning Christiansen).

III -Attended Seminars, Workshops, and Conferences

Courses
Prof. Henning Christiansen, PROLOG, attended 4 lectures, October 28, 2008-December 20, 2008
Prof. Ole Skovgaard, Bioinformatics, January-April 2009

Seminars in the LOST project group within the period of October 28, 2008 and July 27, 2009

Seminar at NTNU (Henning Christiansen and Ana Capatana **).**

Københavns Universitet Bioinformatics Workshop 2009
Henning Christiansen, Mathieu Petit, Ole Torp Lassen, Ana Capatana, Kobenhaven Univerity, Copenhagen, Denmark, January 26 2009, common project poster for the Roskilde group.

The LoSt Project Mini Workshop March 4, 2009
Henning Christiansen, Ana Capatana , Ole Torp Lassen, Matthieu Petit, Nicos Angelopoulos, Christian Theil Have - presentations
Ana Capatana, A focus on Overlapping Genes: The problem and current investigations. Roskilde University, Roskilde, 4 March 2009, presentation.

System Biology Workshop
Ana Capatana, May 14 and May 15, 2009, Center for Biological Sequences Analysis (CBS) at the Technical University of Denmark offered by the BioSys Innovational Network, Copenhagen, Denmark