JOB OPENING AT UPF BARCELONA

PhD/technical position at the Structural Bioinformatics group of GRIB

The Structural Bioinformatics group (SBI) is integrated in the Research Programme on Biomedical Informatics (GRIB), a joint research programme of the Hospital del Mar Medical Research Institute (IMIM) and the Department of Experimental and Health Sciences of the Universitat Pompeu Fabra, located at the Barcelona Biomedical Research Park (PRBB) in Barcelona (Spain).

Research Outline of SBI group (http://sbi.imim.es)

Knowledge of transcription factor (TF) -binding sites, the locations at which TFs bind to DNA in the genome, is key to understand how genes are regulated. Yet, the binding preferences of most eukaryotic TFs remain unknown. In this scenario, the development of computational tools as a complement to experimental procedures is fundamental. A major difficulty in studying TF-DNA binding specificity and in evaluating models for representing this specificity has been scarcity of structural data. We are involved in the development of a homology modeling-based approach that combines structural information and protein binding microarray (PBM) data to predict the binding preferences of TFs and model TF-DNA interactions.

For more information, visit our web at http://sbi.imim.es

We are seeking for a research student, interested in obtaining the PhD in bioinformatics, with capacity to offer technical support for the improvement of programmes generated in the group and the development of new ideas and services within the scientific subject of this announcement (see further).

DUTIES

- ✓ The successful candidate will be involved in the improvement, development and update of the web-services and tools offered in SBI, with particular attention on the analysis of protein-protein and protein-DNA interactions in network analyses, implying new aspects of research within the topic of this opening (see further)

SKILLS AND EXPERIENCE

- ✓ Knowledge of structural bioinformatics techniques (docking, template-based homology modelling, sequence and structure alignment, etc.)
- ✓ Oral and written communication skills.
- ✓ Fluent spoken and written English.
- ✓ Excellent programming/scripting skills in at least one of each of the following bullet points:
    - Script/Programming languages: Python
        - Specificities: BioPython, Numpy, SciPy, TkInter.
    - Programming languages: C, C++, Java.
    - Web programming knowledge on at least one of the following points:
        - Angular JS, JavaScript, PHP.
        - Django, Bottle.
        - HTML5, CGI.
    - Databases in: SQL, MySQL, SQLite, MongoDB

- ✓ Graduate/engineer in one of the following areas: mathematics, physics, chemistry, biochemistry, informatics, human biology, biomedical engineering.
- ✓ Master or equivalent in bioinformatics or computational biology, or in the process of obtaining it.

WORK CONDITIONS

- ✓ Two-year contract with possible extension of one more year, up to the third year required to defend the thesis.
- ✓ Annual gross salary of 16,500€ with the possibility to add collaborations with SMEs in the field of Bioinformatics.
- ✓ Starting on May-June 2019

SUPPORT

The work is supported by grant: BIO2017-85329-R from the Spanish Ministry of Economy (MINECO) and FEDER sources.

SCIENTIFIC TOPIC

Summary: Interactions between transcription factors (TFs) and their binding sites play important roles in many biological processes. Genomic analyses often involve scanning for potential TF binding sites (named motifs) using models of the specificity of the DNA binding domain (DBD) of the TF. Although a large number of TFs have been identified, it is not well known which motifs they can recognize. A major difficulty in studying TF-DNA binding specificity and, therefore, in evaluating models for representing this specificity has been scarcity of structural data. Several studies have attempted to characterize the motifs of a TF by summarizing all its DNA binding sites as a position weight matrix (PWM) using experimental methods. A large amount of data from these experiments is collected in the database CIS-BP. In this project, we will use homology modelling to span the number of structures of known TF-DNA bindings in CIS-BP. We will use these structures to calculate family-specific statistic potentials for the amino-acid-nucleotide interactions. Then, we will use these potentials to characterize, from a theoretical view, the specificity of the interactions between TFs and their DNA motifs, starting by obtaining theoretical PWMs for all modelled TFs. First, the prediction of the correspondence between TFs and their DNA motifs; second, a potential change of specificity caused by mutations; and third, the protein-engineering of a TF highly specific for a DNA motif, have all ultimate impact on the study of mechanistic insights of diseases and phenotypes and for the use of gene-therapies. We will analyse how the changes of specificity of TF-DNA interactions caused by mutations affect the rewiring of the protein-interaction network. As an example, recent works have evinced that positive selection of mutations in transcription factor binding sites affect the regulation of important cancer cell functions. Likewise, for the study of the cross-species interactome, some proteins can act as transcription factors in another specie, thus affecting the regulatory network. Our goal is to understand the mechanisms of recognition and interaction of TFs with their corresponding DNA binding site, focused in the biomedical applications in networks and systems medicine. With this porpose, we will develop the bioinformatics tools to: 1) predict and model TF-DNA bindings in enhancer and promoter regions; and 2) predict the effect of mutations in both sides of the interface of the interaction. Finally, programmable nucleases for genome editing are synthetically constructed by the fusion of a nuclease and a highly specific DNA binder (such as transcription-activator like effectors, TALE, zinc-finger transcription factors, C2H2-zinc fingers, or RNA-guided using the CRISPR system). Thus, we propose a third application: a bioinformatic tool to obtain the theoretical best sequence of a TF targeting a DNA motif with high specificity, that can be used to construct new programmable nucleases, using the scaffold of other TFs different from zinc-fingers or TALE.

HOW TO APPLY

Applications for this opening should include a CV, a motivation letter and the names of two potential referees. Send them by **e-mail** to baldo.oliva@upf.edu with the subject “SBI position”

Reference: **SBI494.2**

For more information, contact Dr. Baldo Oliva (baldo.oliva@upf.edu) or visit our web site at http://sbi.imim.es