

Postdoctoral Positions Available



Prediction of protein function using evolutionary principles and statistical machine learning

Research Group of Steven Brenner
University of California, Berkeley

Project background

The fierce pace of high-throughput sequencing projects has offered a cornucopia of inferred protein sequences. However, only a minuscule fraction of these have been experimentally characterized, and thus computational methods are widely used for automated annotation. We have developed an algorithm for predicting protein function from sequence, called SIFTER (Statistical Inference of Function Through Evolutionary Relationships). This method employs a statistical graphical model to integrate all known functional information for a protein's family, incorporating the phylogenetic structure of the family. The SIFTER approach yields predictions associated with measures of reliability and traceable evidence.

Our benchmark tests show that SIFTER provides accurate and precise functional predictions, outperforming other available programs. Our most recent results demonstrate that the latest SIFTER version is effective on a wide diversity of protein families and scales to genome-scale function prediction.

Our efforts to improve and deploy SIFTER project are multifaceted, involving fundamental statistical computer science methodology development, practical computational biology for effectively applying SIFTER and assessing its performance, and experimental studies to validate SIFTER predictions and guide future enhancements. Researchers on the SIFTER project work in a collaborative integrated interdisciplinary environment.

Postdoctoral Researcher for Algorithmic Development

Project description

The SIFTER algorithmic core is elegant but presently ignores many features of proteins we know to be important and incorporates unrefined parameters. The successful applicant will further develop the SIFTER inference algorithms, for example incorporate knowledge regarding the multi-domain nature of proteins into SIFTER. The candidate will deploy improved approaches to incorporate information into the prediction process by introducing additional features into the function vector. The SIFTER decision rules will also be improved to provide better effectiveness and more reliable predictions with the possibility of predicting multi-functional proteins. Further, the candidate will work towards incorporating a more fully statistical approach to sampling the alignment and phylogenetic data. This project unites disciplines including statistics, machine learning, applied mathematics, evolutionary and molecular biology.

Position requirements

The candidate should have Ph.D. (preferably in statistics, computer science, mathematics, or a related field) with a strong publication record and strong professional references. The ideal applicant will have outstanding quantitative and computer science skills, and experience with machine learning and statistics. Previous experience with solving biological problems with

computers is an asset. Excellent communication skills are required for effective interaction with the multidisciplinary cohort of researchers in our laboratory and with collaborators.

Postdoctoral Researcher in Computational Biology

Project description

SIFTER currently takes data in a phylogenetic tree inferred based on a Pfam alignment and decorates the tree with molecular function evidence from proteins whose roles are supported by experimental evidence codes from the Gene Ontology (GO) database. In order to develop the SIFTER approach to become more robust on richness and nuances of biological data, the successful candidate will practically expand SIFTER to accept and predict a wider range of information, e.g., terms extracted by natural language processing from the literature; proteins not stored in Pfam; additional GO evidence codes; and more generic levels within the GO hierarchy. Ultimately, SIFTER will be capable of accepting evidence information from a bevy of function prediction methods, such as detailed alignment analysis and context studies; the candidate will integrate these into SIFTER's statistical and phylogenetic formalism allowing SIFTER to operate as a meta-predictor. This project unites disciplines including evolutionary and molecular biology, bioinformatics, and computational biology software development.

Position requirements

Candidate should have Ph.D. (preferably in computational biology, genetics, computer science, molecular biology, biophysics, or a related field) with a strong publication record and strong professional references. The ideal applicant will have strong biological insight and previous experience with solving biological problems with computers. The successful completion of this project requires programming skills, but does not require novel algorithmic development. Excellent communication skills are required for effective interaction with the multidisciplinary cohort of researchers in our laboratory and with collaborators.

The Berkeley academic environment

The Brenner lab is an interdisciplinary research group at the University of California, Berkeley. We are associated with the Department of Plant and Microbial Biology, the Department of Bioengineering, the Department of Molecular and Cell Biology, the Biophysics Graduate Group, and Lawrence Berkeley National Lab. Key collaborators for this project include Michael Jordan from Computer Science and Statistics, and Jack Kirsch from Molecular and Cell Biology

The University of California, Berkeley ranks first nationally in the number of graduate programs in the top 10 in their fields, according to the most recent National Research Council study. Berkeley is committed to diversity in its staff, faculty, and student body, and invites all qualified people to apply, including minorities and women, veterans and individuals with disabilities. The University of California is an Equal Opportunity/Affirmative Action Employer.

Interested applicants should send CV, letters of reference, and statement of interest to jobs@compbio.berkeley.edu

Materials must be submitted no later than August 31, 2009.

Please refer to the University's statement on confidentiality, found at <http://apo.chance.berkeley.edu/evalltr.html>. Minorities and women encouraged to apply.

For more information, see <http://compbio.berkeley.edu/>