Undergraduate Participation in Bioinformatics Training (UPBiT)

Ming-Ying Leung,1,2 Stephen B. Aley,2,3 Vladik Kreinovich,2,4 and Elizabeth Walsh2,3

1Department of Mathematical Sciences, 2Bioinformatics Program, 3Department of Biological Sciences, and 4Department of Computer Science, The University of Texas at El Paso, El Paso, Texas 79968

Abstract

With the advances in biotechnology, there has been an exponential increase of biological data, deposited in databanks around the world. The interdisciplinary field of bioinformatics came into existence because of the demand of well-trained professionals to extract useful information from this enormous amount of data with carefully designed mathematical models, statistical methodologies, computational algorithms, and experimental validation. Building on the existing infrastructure for the graduate degree in bioinformatics, we are developing an undergraduate training program of bioinformatics research at UTEP. This research experience will better prepare biology and mathematics students to pursue graduate studies or careers in fields dependent on mathematical tools for biological data analysis.

Teams of faculty mentors from biological sciences and mathematical sciences will work with groups of undergraduate majors in biology and mathematics on projects stemming from four research themes: Biomolecular Sequence Analysis, Ecology and Evolutionary Biology, Gene Expression and Proteomics Data Analysis, and Protein Structure Prediction. It is also expected that from these interdisciplinary collaborative projects, the interaction among faculty and students from both the biological and mathematical sciences will bring about developments in the undergraduate curricula of both disciplines. We plan to implement modifications in the contents of a number of biology and mathematics courses and to establish new bioinformatics undergraduate courses which will be the foundation for graduate degree studies in bioinformatics.

Program Organization

Student Training Activities

Weekly Workshops to Develop Interdisciplinary Communication Skills
• Writing and presentation assignments in bioinformatics related topics
• Group work to help trainees select research partners
• Discussions on ethical issues to build awareness of social responsibilities

Computer and Wet-Lab Training
Learn to use software like MATLAB, MINITAB, and basic programming in Java and Perl. Shadow graduate students or research staff in wet labs.

Research Rotation
Each trainee will go through a research rotation to work with faculty mentors on several small projects from two to three of the four research themes described below.

Specialized Research Project
After research rotation, trainee will choose a specialized research project from one of these themes.

Theme 1: Biomolecular Sequence Analysis:
Probabilistic models and statistical methods can be developed to analyze words and string patterns on DNA, RNA, and protein sequences. Such analyses can provide good predictions for important functional sites, and have also motivated unexplored mathematical problems

Theme 2: Ecology and Evolutionary Biology:
Research topics include mathematical models for aggregability and speciation, ecoinformatics analysis of rainfall and water resources, and studies of the effects on the biological systems at UTEP’s Indio Mountain Research Station, where linear algebra, tree building, and graph theory will be used.

Theme 3: Gene Expression and Proteomics Data Analysis:
The analysis of high dimensional gene expression and proteomics data from microarrays and mass spectrometry has been a recent topic of interest. Locally generated data from studies of rotifers, Trypanosoma cruzi, and nodaviruses will be analyzed by Bayesian variable selection and spectral analysis techniques.

Theme 4: Protein Structure Prediction:
Understanding the mechanisms for efficient and precise assembly of large macromolecular complexes is an important problem in biochemistry. While crystallography and cryo-EM can determine protein structures, optimization theory and large scale simulations will reduce the cost.

Curriculum Development
• New biology and math course contents
• Undergraduate Bioinformatics Minor
• 4+1 BS-MS Program in Bioinformatics

Long Term Goal

The UPBiT program is a pivotal step towards establishing an interdisciplinary Bioinformatics Research Center at UTEP to bring together talents among our undergraduate and graduate students and the expertise among postdoctoral researchers and faculty to address significant challenges in both the biological and mathematical sciences.

Acknowledgements

UPBiT is supported by the National Science Foundation (NSF) Grants “UBM Institutional: Undergraduate Training in Bioinformatics” (DUE-0926721), “Mathematical Models for RNA” (DMS-0800272), and the following grants:

National Institutes of Health (NIH): S06GM08012-35, 2G12RR008124-11 and 1T36GM078000-01