Developing Computational Tools for RNA Structural Prediction, Proteomics Analysis, and Mammogram Enhancement

> Ming-Ying Leung, PhD Professor, Department of Mathematical Sciences Director, Bioinformatics Program Director, Bioinformatics Computing Core Facility Border Biomedical Research Center (BBRC) The University of Texas at El Paso



BCL at Bell Hall, UTEP

Aims and Activities:

- Interface with other BBRC cores and support BBRC research projects.
- Offer bioinformatics computing services to researchers in the region.
- Provide integrated consultation or collaboration for biomedical researchers in BBRC and other institutions

Bioinformatics and Other BBRC Core Facilities

Bioinformatics

- Biomolecular sequences and structures
- Genomics and proteomics
- Ecoinformatics and Phylogenetics
- Data mining and prediction algorithms
- Grid computing and machine learning
- Mathematical models and theories



Statistical Consulting Lab Statistical Bioinformatics

e.g., cancer informatics

Biomolecule Analysis







Integrate Computational & Experimental Approaches e.g., RNA structure prediction



BCL at Bell Hall (plus satellite sites)

Resource Faculty and Staff

- Biological Sciences: Steve Aley, Igor Almeida, Kyle Johnson, Max Shpak, Elizabeth Walsh, Jianying Zhang
- Chemistry: Ricardo Bernal, Mahesh Narayan
- Computer Science and Engineering: Olac Fuentes, Vladik Kreinovich, Wei Qian
- Mathematical Sciences: Ming-Ying Leung, Naijun Sha, Leticia Velazquez
- Technical staff: Gerardo Cardenas, System Administrator

Ongoing Bioinformatics Projects in 2010

- RNAVLab: Virtual laboratory for RNA sequences analysis and secondary structure prediction
- GPI and GPI-anchored protein structure prediction based on mass spectrometry data
- Bayesian models, coalescent theory, and other statistical methods in biological systems
- Image analysis for cancer detection (e.g., super-resolution and compression algorithms for mammograms)

RNAVLab: A Virtual Laboratory for Studying RNA Secondary Structures Based on Grid Computing Technology



Family Nodaviridae: RNAVLab Results

NoV RNA2 prediction



NoV RNA2 replicons tested



Ability to replicate in transformed yeast cells



Proteomics Data Analysis



GPI-omics

Mammogram Enhancement by Super-Resolution

Automatic Segmentation

(a) Original image. (b) Converting the original image to binary image. (c) Using erosion to remove the labels on the binary image. (d) Converting the binary image to gradient image to find the edge. (e) Using skeletonization to reduce the edge to a single line. (f) Sampling points on the line and make the mesh.

Super-Resolution Results

Sample super-resolution results by a downsample factor of 8: (a) Original high-resolution image. (b) Super-resolution image.(c) Interpolated image. (d) low-resolution image

Bioinformatics Training

- Weekly bioinformatics colloquium
 - Research seminars
 - Hands-on workshops
- Summer training at the Pittsburgh Supercomputing Center (NIH funded partnership with UTEP)
 - Internships
 - Supercomputing workshops
- Research training for
 - Postdoctoral fellows
 - PhD students in Computational Science
 - MS students in Bioinformatics
 - UPBiT undergraduate trainees in Bioinformatics

Bioinformatics Training (cont.)

Biomolecular Sequence Analysis

8

Ecoinformatics and Phylogenetics

Research Rotations

Molecular Structure and 2 Dynamics

Microarray and Proteomics Data Analysis

Undergraduate Participation in Bioinformatics Training (UPBiT)

Development Plan

- Interface with other BBRC cores:
 - Analytical Cytology
 - Cell Culture and High Throughput Screening
- Increase bioinformatics support to the BBRC research units:
 - Infectious Disease and Immunology (sequence alignment, and phylogeny, annotations)
 - Toxicology (toxicological genomics)
 - Neuroscience and Metabolic Disorders
- Provide bioinformatics computing services for the region

Funding Support

- 2007 2011: NIH (SCORE) "Computational Prediction of RNA Viral Genome Structures"
- 2008 2010: Texas State University System Stars Funding "Medical Imaging and Bioinformatics"
- 2008 2011: Texas Higher Education Coordinating Board Advanced Research Program – "Probabilistic Models for Inversions in Viral RNA Sequences"
- 2008 2012: NSF Mathematical Sciences "Mathematical Models for RNA"
- 2008 2013: NIH (MARC) "Enhancement of Quantitative Science in Biology Curricula"
- 2009 2013: NSF Interdisciplinary Training for Undergraduates in Biological and Mathematical Sciences (UBM) – "UBM Institutional: Undergraduate Training in Bioinformatics"
- 2009 2014: NIH/NCRR "Border Biomedical Research Center Bioinformatics Computing Core Facility"