Informatics Expertise to Support Life and Health Sciences Research and Industry

Chi-Ren Shyu, Ph.D.
shyuc@missouri.edu

Director, Informatics Institute
University of Missouri – Columbia

http://MUII.missouri.edu
Informatics at MU

In 2010, MU Informatics Institute (MUII) has 30 core faculty members, 25 doctoral students, $8M ongoing informatics related research and training programs.

60’s—80’s Donald A.B. Lindberg, M.D.

Director of the National Library of Medicine has pioneered in applying computer technology to health care beginning in 1960 at MU
MUII is currently a campus-wide hub for informatics research, education, and services
Bioinformatics

High throughput sequencing, structure biology, cancer research, neural sciences, animal sciences, translational bioinformatics, etc.

Highlights:

MULTICOM methods were ranked among best in six categories during the 8th Critical Assessment of Techniques for Protein Structure Prediction (CASP8), 2008.

ProteinDBS is the first real-time protein 3D structure retrieval engine.
Medical/Health Informatics

EMR/EHR, Imaging Informatics, Healthcare Quality, Texting Mining, Human Computer Interaction, Nursing Informatics, Elder Care, Autism, etc.
Geospatial Informatics

Epidemiology, intelligence applications, agriculture applications, crime analysis, environmental studies, etc.
Scenario – Personalized Medicine/Healthcare

**Personalized Genomics** – Bioinformatics tools to analyze individual genomes from SNP, partial/full genome sequence, etc.

**Structural Bioinformatics** – Rational drug design, host-pathogen interaction, etc.

**Human Factors Studies, Data Mining from Omics Databases and EMR** – Patient safety, Next generation search engine for medical multimedia data, etc.
Development of regulatory complexity through genome duplication


Demonstration of neutral evolution on the surface of mammalian proteins


Identification of partitioned cellular subnetworks created by genome duplication in yeast

Transcription factors with surviving duplicates from a genome duplication in yeast are more central in the regulatory network.
Gene order phylogenetics


Adaptation of yeasts to glucose-rich environments by genome duplication


Effects of gene duplication on cancer-cell metabolism

As the number of tumor samples sharing a duplication increases left to right, the proportion of duplicated metabolic genes present in reactions involved in producing biomass increases (pink line).

Unpublished data

1. Genome duplication led to an increased concentration of glycolytic enzymes
2. Differential scaling of respiration and fermentation means an increased importance of fermentation
3. A polyploid yeast was at a selective advantage because it used glucose rapidly
Rational drug design: A known protein structure may provide a guide to design a drug that modulates the protein's function for curing a disease.
MU Ranked among best in seven categories:

- Template-free modeling
- Template-based modeling
- High-accuracy modeling
- Contact map prediction
- Disorder prediction
- Protein domain prediction
- Protein model quality prediction

8th Critical Assessment of Techniques for Protein Structure Prediction (CASP8), 2008
Bioinformatics pipeline to study host-pathogen interactions (HPIs)

Gather HPI data
Identify HPI key players
Prediction of effector proteins
Characterize HPIs
Study phenomena
Molecular mimicry detection
Apply to specific systems
Our collaborations
Virulent mimicry in host-pathogen interactions

Currently, no computational methods available!
Rational drug design against a potassium channel that is responsible for the long-QT syndrome (a genetic heart problem) and sudden death.

White: Surface of part of the potassium channel;

Red: The region of interest for drug design

Research Highlights:
- 1 copyright (MDock software package);
- 1 patent filed (anticancer therapy);
- Third place in 4\textsuperscript{th} CAPRI (an international competition on structure prediction of protein-protein complexes)
Human Factors in Healthcare Systems

- Application of modeling techniques to identify functional relationships between human constructs (e.g., fatigue, stress) and systems outcomes (e.g., performance, safety)

- Evaluation of medical technology design features to quantify physical and mental work demands for workers, reduce risk of worker injuries, and improve patient safety
Next generation medical multimedia info system

Scanner

Pre-processing

Content Extraction
- Key slices
- Pathology bearing region

Algorithm Library
Computer Vision
Image Processing

Image Archive
All modalities
Cases

Genomics DBS
Sequences
structures

Multi-dimensional Database indexing

Off-line Image Transferring

On-line Signature matching

Current Data Flow

New Data Flow - Onsite

New Data Flow - Remote

Alert Advisor
- Radiology
- Pathology

EMR/EHR

Radiology Reading Station PACS
Next generation search engines
Informatics Expertise – Maximizing Collaborators

- Teaming up with Kansas City area industry and research institutions
- Building a knowledge base that has up-to-date informatics expertise database in Missouri, Kansas, and surrounding states.
- Developing an incubation program to host start-ups, both physically and virtually, for commercialization of informatics software.