# 2007 IEEE International Conference on Bioinformatics and Biomedicine Workshops

## San Jose, CA
## 2-4 November 2007

| | |
|---|---|
| **IEEE Catalog Number:** | **CFP0726D-PRT** |
| **ISBN 10:** | **1-4244-1604-3** |
| **ISBN 13:** | **978-1-4244-1604-2** |

# Table of Contents

## Keynote

An Investigation into the Feasibility of Detecting Microscopic Disease Using Machine Learning ..... 1
*Mary Qu Yang and Jack Y. Yang*

## Workshop 1: Computational Structural Bioinformatics

Predicting Protein-Protein Interactions Based on BP Neural Network..... 3
*Zhiqiang Ma, Chunguang Zhou, Linying Lu, Yanan Ma, Pingping Sun, and Ying Cui*

Point to Face Shortest Paths in Simple Polytopes with Applications in Structural Proteomics..... 8
*Ovidiu Daescu and Yam Ki Cheung*

Refinement of Under-Determined Loops of Human Prion Protein by Database-Derived Distance Constraints..... 14
*Feng Cui, Kriti Mukhopadhyay, Won-Bin Young, Robert L. Jernigan, and Zhijun Wu*

PRTAD: A Database for Protein Residue Torsion Angle Distributions..... 24
*Xiaoyong Sun, Di Wu, Robert Jernigan, and Zhijun Wu*

Non-Quantized Minimum Free Energy in Untranslated Region Exons..... 32
*Keith Knapp, Ahadur Rahaman, and Yi-Ping Phoebe Chen*

Enhancing the Computation of Approximate Solutions of the Protein Structure Determination Problem through Global Constraints for Discrete Crystal Lattices..... 38
*Alessandro Dal Pal, Agostino Dovier, and Enrico Pontelli*

Protein Folding Simulation by All-atom CSAW Method..... 45
*Weitao Sun*

Detecting Constituent Sequences by Means of HP Pattern--Based Grammars to Synthesize Proteins: Inferring Sequence--Structure--Function Relationship..... 53
*Giuseppe Nicosia, Eva Sciacca, and Luca Zammataro*

Structural Characterization of RNA-Binding Sites of Proteins: Preliminary Results..... 60

*Fadi Towfic, David C. Gemperline, Cornelia Caragea, Feihong Wu, Drena Dobbs, and Vasant Honavar*

Incorporating Constraints from Low Resolution Density Map in Ab Initio Structure Prediction Using Rosetta.....67
*Yonggang Lu, Charlie E. M. Strauss, and Jing He*

An Approximate Robotics Algorithm to Assemble a Loop between Two Helices.....74
*Jing He and Kamal Al Nasr*

Flagellar Proteins Prediction after Sequence-Structure Alignments of Coronin and Arp2/3 Complex in Leishmania spp.....80
*Marcília P. Costa, Hálisson L. Ribeiro, Ana C.L. Pacheco, Daniel P. Pinheiro Michel T. Kamimura, Raimundo Araújo-Filho, and Diana Magalhães de Oliveira*

Comparison of Experimental and Computed Protein Anisotropic Temperature Factors.....89
*Lei Yang, Guang Song, and Robert L. Jernigan*

## Workshop 2: Data Mining in Bioinformatics

Are Filter Methods Very Effective in Gene Selection of Microarray Data.....97
*Zhou-Jun Li, Li-Juan Zhang, and Huo-Wang Chen*

Comparative Analysis of Binary Logistic Regression to Artificial Neural Networks in Predicting Precursor Sequence Cleavage.....101
*Allison Tegge, Sandra Rodriguez-Zas, Jonathan V. Sweedler, and Bruce Southey*

Multiresolution Approaches to Representation and Visualization of Large Influenza Virus Sequence Datasets.....109
*Leonid Zaslavsky, Yiming Bao, and Tatiana A. Tatusova*

Using Fuzzy Memberships to Core Patterns to Interpret Connectedness in Gene Expression Clusters.....115
*Noha A. Yousri1, M. S. Kamel, and M. A. Ismail*

Mediated Taxonomy System for Bioinformatics Data Integration.....123
*Feng Ji, Jack Fu, Ramez Elmasri, Nikola Stojanovic, and Geoffrey Grant*

ailSemantic Web-based Data Representation and Reasoning Applied to Disease Mechanism and Pharmacology.....131
*Xiaoyan Angela Qu, Ranga Chandra Gudivada, Anil G Jegga,*

*Eric K. Neumann, and Bruce J Aronow*

Ontology Algebra for Composition of Protein Data Sources.....144
*Amandeep S. Sidhu, Tharam S. Dillon, and Elizabeth Chang*

## Workshop 3: Knowledge Discovery and Management in Hearth Informatics

SHA: A Secure Voice Activated Smart Home for Quadriplegia Patients.....151
*Yusuf Uzunay and Kemal Bicakci*

A Bayesian Approach for the Analysis of Recurrent Events with
Chronic Granulomatous Disease.....159
*Chi-Chang Chang*

Developing an Annotated Corpus of Patient Histories from the
Primary Care Health Record.....165
*Ola Huseth and Thomas Brox Røst*

A Mixture Language Model for Class-Attribute Mining from
Biomedical Literature Digital Library.....174
*Xiaohua Zhou, Xiaohua Hu, Xiaodan Zhang, and Daniel D. Wu*

## Workshop 4: High-throughput Data Analysis for Proteomics and Genomics

V-Lab-Protein: Virtual Collaborative Lab for Protein Sequence Analysis.....183
*Jong Youl Choi, Youngik Yang, Sun Kim, and Dennis Gannon*

INDARE - An Indexed DAG of Regular Expressions for Selecting
Position Frequency Matrices.....191
*Meeyoung Park, Jubin Sanghvi, and Deendayal Dinakarpandian*

Rapid Computation of Large Numbers of LOD Scores in Linkage
Analysis through Polynomial Expression of Genetic Likelihoods.....197
*Hongling Wang, Alberto Maria Segre, Yungui Huang,*
*Jeffrey R. O'Connell, and Veronica J. Vieland*