inferring novel insights by comparing different experiments, often involves the identification of orthologs between species and the conversion of identifiers between various databases formats. HOMECAT (HOmology Mapper for Enrichment and Comparative Analysis with Translation) is a Cytoscape plugin for comparative investigations based on homology. Starting from a list of identifiers or a network, the plugin searches the best consensus orthologs and paralogs from the principal homology databases and conveniently represents them as metanodes, where homologs are satellites of the input nodes. Data provided by the user or collected from various sources, like ArrayExpress or GEO, can be mapped on collapsed metanodes and visualized as pie charts allowing visual inspection of similarities. The plugin is interfaced with EBI PICR and BridgeDB and seamlessly converts user-provided identifiers to query homology databases and to convert results back to user format. Data integration is also facilitated by the possibility for the user to convert identifiers prior to the mapping directly from the plugin. A case study is presented to elucidate the plugin functionalities.

**L11 | Generalized Simulated Annealing Applied to ab-initio Protein Structure Prediction**

Melo M *, Bernardi R, Pascutti P  
* Laboratório de Modelagem e Dinâmica Molecular, Universidade Federal do Rio de Janeiro, 21941-902, Rio de Janeiro - RJ, Brazil

Proteins are the building blocks of cells and the executioners of nearly all cellular functions. Their structure is of paramount importance to understand their dynamics and function, as well as the interactions with other molecules. In this work, we apply the Generalized Simulated Annealing (GSA) to the prediction of protein structures in a new software. The GSA is a stochastic search algorithm employed in energy minimization and used in global optimization problems, such as gravity models, fitting of numerical data and conformation optimization of small molecules. Our software applies the analytical inverse of the probability distribution from GSA, a new method to apply rotations to the phi and psi angles of the peptide bonds and side chains, faster connection with NAMD for potential energy calculation and the possibility of parallel execution, granting a new take on ab-initio protein structure prediction. The new design also allows for an easier inclusion of knowledge derived potentials, based on experimentally determined protein structures. We present results for the 14 amino acid protein mastoparan-X. The chain folds with RMSD of 3.0 angstroms after 500.000 GSA steps. Currently, for this system, the software calculates 5 million GSA steps in under 6 hours using 4 processors. Predicted structures can be refined with molecular dynamics simulations and used to study proteins whose conformation can not be determined with experimental methods.

**L12 | Genome-Wide Annotation Prediction with SVD Truncation based on ROC Analysis**

Chicco D *, Masseroli M  
* Politecnico di Milano, 20133 Milan, Italy

Correct interpretation of many biological experiments is currently based on consistency of biomolecular annotation databases. Such databases are very useful for the scientific community, but, unfortunately, incomplete by definition. To improve their consistence, computational methods able to supply ranked lists of predicted annotations are hence extremely useful. We departed from a previous work on the automatic prediction of Gene Ontology (GO) annotations based on the truncated Singular Value Decomposition (SVD) of the annotation matrix, where every matrix element corresponds to the association of a biomolecular entity to a GO term. Then, we developed a new method where the truncation choice is based on analysis of the Area Under Curve (AUC) of different Receiver Operating Characteristic (ROC) curves for different truncations. To evaluate our method, we used annotations of different organism genes available on July 2009 in an old version of GO Annotation databases. By analyzing Gallus gallus annotations between genes and Biological Process terms, the best truncation parameter, suggested by the algorithm, led to better results than other truncation levels: from all the input annotations, the SVD method with best truncation predicted the highest number of annotations whose presence were confirmed in a more recent GO database version (October 2011). Contrariwise, other truncation levels, related to higher AUC values, led to worst prediction results. To get more correct biomolecular annotation predictions, our SVD best truncation choice method revealed very effective and reliable. Furthermore, since our approach is not limited to specific annotation types, can be applied to any controlled annotation.
List of accepted poster abstracts

Preliminary list as of 3 August 2012; for final list please visit

http://www.eccb12.org/posters

Sections:
A - Applied and Translational Bioinformatics
B - Bioimaging, Spatial-temporal Modeling and Data Visualization
C - Bioinformatics of Health and Disease, Biomarkers and Personalized Medicine
D - Databases, Ontologies, and Text Mining
E - Evolution, Phylogeny, and Comparative Genomics
F - Macromolecular Structure, Dynamics and Function
G - Mutations, Variations, and Population Genomics
H - Protein Interactions, Molecular Networks, and Proteomics
I - Regulation, Pathways, and Systems Biology
K - Sequencing and Sequence Analysis
L - Technology and Software