Click on the menu items (above) for sub-directory links.

Software Demonstrations - *Abstracts*

Industry Demos

Accelrys - From Genes to Protein Structures Prediction: "A Genome Analysis"

Learn advanced in silico high throughput bioinformatics, functional annotation, protein homolog modeling and 3D annotation techniques to study genomes and proteomes. Demo includes using our Discovery Studio software to analyze the West Nile Virus genome. For more information see http://www.accelrys.com/dstudio/.

Apple Computer- *Blending Power with Simplicity in Life Science Clustering: iNquiry and the Apple Workgroup Cluster*

The Apple Workgroup Cluster for Bioinformatics is a turnkey computational cluster designed to enable bench scientists to achieve greater autonomy by controlling their own computational resources. During this session, speakers from Apple and The BioTeam, Inc. will explore the computational challenges that face the modern biologist, and show how the Apple Workgroup Cluster and BioTeam's iNquiry cluster provision tool successfully address them. The session will include live demonstrations of the hardware and software discussed.

Apple Computer- *Xgrid and Mac OS X: A Platform for Distributed Scientific Computing*

The need for computational power in life science computing continues to grow at a rate that, today more than ever, requires creative approaches to resource acquisition and management that enables average scientists to easily harness existing computational power into a logical resource that behaves much like a computational cluster, but without the management complexity and financial burden. Learn about recent announcements regarding Xgrid, and see how you can benefit from its use, even as a non-Macintosh user.

IBM- *Graph Data Management for Bioinformatics*

Many biological objects are best modeled as networks or graphs. We describe a prototype from IBM Research that supports manipulation of biological data directly in the form of graphs, thereby allowing bioinformaticians to manage this data more naturally. We demonstrate its application to clustering genes whose expression levels distinguish between two closely related lymphoma tumor variants.

MathWorks, Inc.- *MATLAB for Bioinformatics*

Bioinformaticians and researchers worldwide rely on MATLAB to accelerate scientific discovery and reduce development time. This software demonstration will demonstrate how to use the MATLAB programming environment for bioinformatics applications such as sequence alignment, microarray data analysis and visualization, phylogenetic tree analysis, and algorithm deployment to stand-alone executables and Excel.

SciTegic - *A Graphical Pipelining Environment for the Integration of Bioinformatics Data and Tools: Pipeline Pilot™*

SciTegic will be introducing the bioinformatics-specific capabilities of its successful data pipelining product, Pipeline Pilot. The graphical user interface will be demonstrated during the composition of analytical pipelines to search, annotate, and align sequences. Further, these new bioinformatics capabilities can be used in pipelines with existing cheminformatics functions for a truly cross-domain analysis environment.

Scottish Enterprise - *The Scottish Bioinformatics Forum - Demonstrating Bioinformatics Excellence in Scotland*

Scotland has a number of centres of excellence in the area of bioinformatics. The Scottish Bioinformatics forum (SBF) was set up to increase multi-disciplinary projects in Scotland, raise the profile of bioinformatics, attract quality people and funding, and provide a stimulating learning environment. In this session Scottish bioinformatics capability, and key research projects will be demonstrated.

Scottish Enterprise demo will contain six following mini-presentations:
1. Interprobe Chemical Services Software
2. Axiope Catalyzer: How do we make it easy to database all the other scientific data?
3. Germinate: A Plant Data Management System
4. SyntenyVista - genome visualisation software
5. PartiGene – from ESTs to gene sequences and beyond
6. Topali - Software for Detecting Recombinants in DNA Multiple Alignments

Sun Microsystems, Inc. - *Managing Knowledge, Throughput, Visualization, Grids & Data Centers*

Sun Microsystems' CTO for Life Sciences will present a collection of demos and information from our commercial and public partners. These include Accelrys, BioTeam, Gene-IT, Geospiza, INCOGEN, TimeLogic and others. APBioNet's "BioBox" collection of public algorithms for Sun and Blueprint's comprehensive BIND database (join us for an exciting launch by Blueprint!). Finally, Sun will also demonstrate our technologies for secure knowledge sharing, image and visualization management, and others that are having high impact on life science research.

SynapseBASE - *Advantages of a Novel Structured Network Database in Prediction of Biologically Significant Patterns in High-throughput Genomics and Proteomics*

SynaBASE is the world's first 2nd generation database for high-throughput genomics and proteomics. By structuring patterns and their relationships by significance, queries on high level increases in speed and scale enable analyses that were previously impossible or impractical; e.g. true comparative genomic comparisons in milliseconds, real-time proteomics analysis and haplotype characterizations.

Student/Non-Profit Demos

Baylor College of Medicine - *Analysis of Intermediate Resolution Structures*

A wealth of information can be mined from electron cryomicroscopy data at intermediate resolutions. To this end, we developed AIRS, the Analysis of Intermediate Resolution Structures toolkit. Airs is capable of identifying secondary structure elements, localize structures and generate ab initio and comparative models. In this demo, we will explore AIRS using UCSF's Chimera.

Bilkent University - *PATIKA - Pathway Analysis Tool for Integration and Knowledge Acquisition*

PATIKA is a software framework for collaboratively constructing pathway models based on an extensive ontology. PATIKA provides users with different facilities such as integration, querying, analysis and visualization, as well as microarray data analysis support.

Bioinformatics Institute - WILDFIRE: Distributed, Grid-enabled Workflow Enactment Engine

We present Wildfire, a graphical user interface for constructing and running workflows. Wildfire borrows user interface features from Jemboss and adds a drag-and-drop interface allowing the user to compose workflows using EMBOSS (and other) programs. For execution, Wildfire uses GEL, the underlying workflow execution engine, which can exploit available parallelism on multiple cpu machines including Beowulf-class clusters and Grids.

Blueprint - *SeqHound: A Bioinformatics Application Programming Platform*

SeqHound is a bioinformatics programming platform offering daily-updated contents of Entrez sequences, 3-D structures, sequence redundancies and neighbours, taxonomy, complete genomes, conserved domains, GO terms and PubMed links. SeqHound is accessible via a remote API (Perl, BioPerl, Java, C/C++). Source code is freely available under GNU public license (http://seqhound.blueprint.org).

Center for Applied Computer Science, University of Cologne and Max Planck Institute for Molecular Genetics, Berlin - GQL: A comprehensive tool for gene expression time course analysis

The GQL is a comprehensive tool for mining gene expression time course data running on all systems. It is written in Python and is based on HMMs. It allows for interactive, robust and highly visualized analyses and supports the use of prior knowledge. The GQL is freely available at http://ghmm.org/gql.

CNRG Genoscope - *YIB: A Tool for Biological Network Motif Analysis*

The Yeast Interaction Browser is a tool for network motif analysis and exploration of cell biochemical interactions networks. It is provided with yeast molecular interactions data: Protein-Protein interactions, Transcriptional Regulation and Metabolic network. The web interface provides an easy way to search for complex heterogeneous motifs, patterns and view statistics.

Cold Spring Harbor Laboratory - *Reactome - A Knowledgebase of Biological Processes*

Reactome located at www.reactome.org, describes life processes in a top-down fashion, that is useful for didactic and data mining purposes. The web interface gives a textbook-like view of biological processes, while the query tools allow bioinformaticists to mine the database. The website and related tools will be demonstrated.

Cornell University - *BIOZON: A System for Unification, Management and analysis of Heterogeneous Biological Data*

Biozon is a unified biological database that integrates heterogeneous data types and the relationships between them into a single extensive schema. This schema allows one to see each data instance in its full biological context. More importantly it allows for complex searches that span multiple data types and for computations on that data.

EMBL - *Protein Disorder Prediction*

DisEMBL (http://dis.embl.de) DisoPred are computational tools for prediction of disordered/unstructured regions within a protein sequence. Avoiding potentially disordered segments in protein expression constructs can increase expression, foldability and stability of the expressed protein. The tools are thus useful for target selection and the design of constructs as needed for e.g. structural biology and structural genomics projects.

European Bioinformatics Institute (EBI) - *ArrayExpress, MIAMExpress, Expression Profiler Tools for Microarray Annotation, Storage and Analysis*

ArrayExpress, MIAMExpress and Expression Profiler are the EBI solution to the problem of microarray data storage, annotation and analysis. MIAMExpress is a standalone Java/SQL application for annotation and generation of MAGE-ML format. ArrayExpress is a MIAME compliant data repository and Expression Profiler is a comprehensive analysis tool.

European Bioinformatics Institute (EBI) - *EMBOSS*

EMBOSS is an open source molecular interaction database and analysis system. It provides manually curated data, network visualisation in the context of GO annotation, and dynamic PSI MI XML download. Software, data, and documentation are available from http://www.ebi.ac.uk/intact

European Bioinformatics Institute (EBI) - *Jemboss - the Graphical User Interface to EMBOSS*

EMBOSS is a free software suite encompassing over 200 applications for computational biological analysis. Jemboss is a java-based point and click interface. We will look at sequence retrieval from specified databases, alignment and other analyses and will include demonstration of the Jemboss alignment and DNA editors.

European Bioinformatics Institute and The Research Collaboratory for Structural Bioinformatics - *MSD/RCSB PDB Access Tools and Databases*

The macromolecular structure database group at the EBI has generated a database of 3D structure from the PDB archive. This database contains proteins that have a biological context plus self consistent set of ligands/Unitprot mappings. Numerous search systems have been developed along with novel visualisation methods for information retrieval.

European Bioinformatics Institute (EBI) - *Taverna - a Bioinformatics Workflow Tool*

Taverna allows users with minimal technical support to access a diverse range of analysis tools and databases across the world, to build process flows using these components and to manage the submission of data into and extraction of results from these flows. It is freely available, open source and platform neutral

Faculty of Science, University of Zagreb - *INCA: Synonymous Codon Usage Analysis Software*

INCA (INteractive Codon usage Analysis) package is an array of features useful in analysis of synonymous codon usage in whole genomes. In addition to computing common codon usage indices, INCA offers numerous options for interactive graphical display, and clusters genes using a neural network algorithm, the self-organizing map (SOM). Available at http://www.bioinfo-hr.org

Fraunhofer FIT- *Biozoom*

We show a new approach for exploration in images and associated data. We demonstrate our Biozoom prototype on a set of marker-tagged fluorescent cell images.We combine visual data mining in symbolic feature data and clustering results together with advanced image inspection features.

Gene Ontology Consortium - *The DAG-Edit Ontology Development System*

DAG Edit is an ontology management system used by the Gene Ontology(GO) and Open Biological Ontologies (OBO) projects. Originally developed as an editor for directed acyclic graphs (DAGs) of controlled terms and metadata, DAG Edit can now supports many of the advanced features of ontology languages such as OWL.

The Hebrew University - *SPRINT – the Side-chain Prediction Inference Toolbox*

SPRINT, the Side-chain PRediction INference Toolbox, represents side-chain prediction problems as graphical models and applies various inference algorithms to it. Specifically, it returns the M lowest energy configurations and characterizes the variability in configurations (e.g. conformational entropy, free energy and rotamers' marginal probabilities). SPRINT main advantage over other side-chain prediction packages is its multi-functionality while using comparable computing resources.

Heriot-Watt University, Edinburgh University - *XSPAN - A Cross-Species Anatomy Network*

XSPAN (www.xspan.org) supports cross-species access to tissue-specific genetic information. It links anatomy ontologies for human and the common model organisms (mouse, Drosophila, Caenorhabditis elegans, zebrafish). Acquisition of tissue mappings is supported by the COBrA tool and a new cell-type ontology. XSPAN can be accessed using web service and web-based GUI interfaces.

Japan Biological Information Research Center, AIST - *H-Invitational Database - A New Human Gene Annotation Database*

H-Invitational Database (H-InvDB) is a human gene database with integrative annotation of 41,118 full-length cDNA clones. H-InvDB describes their gene structures, functions, domains, expression, diversity, and evolution. This is a product of the H-Invitational consortium comprised of 44 research institutes worldwide, organized by Japan Biological Information Research Center (JBIRC) and RIKEN DNA Data Bank of Japan (DDBJ).

Kyushu Institute of Technology - *CADLIVE: A Direct Link of Biochecmial Netoworks to Dynamic Simulations*

We proposed a general method of directly converting concrete biochemical networks to dynamic models, and developed CADLIVE that implements the conversion method. The feasibility of CADLIVE has been demonstrated by using large-scale and complicated biochemical networks, such as nitrogen fixation, heat shock response, circadian clock, and cell cycle systems.

Max Planck Institute for Dynamics of Complex Technical Systems - *Exploring Metabolic Networks Using the FluxAnalyzer*

The FluxAnalyzer is a graphical user interface for MATLAB enabling detailed structural analyses of large-scale metabolic networks. The powerful collection of tools includes those for graph-theoretical calculations, metabolic flux analysis, flux balance analysis, metabolic pathway analysis (relying on elementary modes) and computation of minimal cut sets.

Nanyang Technological University - *MolStudio - A low end system for VR-enhanced protein visualization*

MolStudio enables you to visualize and interact with protein structures with desktop, laptop and tablet computers. You can walk-through and hands-on with proteins to enhance the structure-based learning and research processes using the Virtual Reality Technology.

NESC - *Bridges Project*

The BRIDGES project is developing an advanced Grid based infrastructure for use within the Wellcome Trust Cardiovascular Functional Genomics (CFG) project. Data access and integration, security and usability are all key issues that the BRIDGES project are addressing. This demonstration will highlight some of the tools that the BRIDGES team have developed for the scientists.

Rat Genome Database - *Rat Genome Database*

The Rat Genome Database (RGD) curates and integrates rat genetic and genomic data and provides access to this data to support research using the rat as a genetic model for the study of human disease

San Diego Supercomputer Center - *Encyclopedia of Life*

The Encyclopedia of Life (EOL) project seeks to provide annotation for the vast number of putative protein sequences derived for genome sequencing projects. Data from the EOL genome annotation pipeline (iGAP) is distributed via an innovative web interface. Further information at: http://www.eolproject.info

School of Computer Science, University of Nottingham - *ZigZag for BioInformatics Data*

BioInformatics data stems from a wide variety of sources, and is serialised in a plethora of formats, standards and conventions. Nelson's ZigZag will be demonstrated, with a variety of information types, to illustrate the uses of a single space for storage, viewing, analysis and representation in a deeply interconnected manner.

UniProt - *UniProt (Universal Protein Resource): A Comprehensive Catalog of Information on Proteins*

UniProt (http://www.uniprot.org) is a centralized resource for protein sequences and functional information. Created by unifying the information contained in Swiss-Prot, TrEMBL, and PIR, UniProt provides a richly annotated protein knowledgebase (UniProt), non-redundant reference databases (UniRef) at different levels of redundancy, and an archive database (UniParc) that reflects the history of all protein sequences.

Universitat Autònoma de Barcelona - *PDA: A Pipeline to Explore and Estimate Polymorphism in Large DNA Databases*

PDA is a Web server that automatically can search for polymorphic sequences in large databases and estimate their genetic diversity on different functional regions. It has been applied to create the secondary database DPDB, *Drosophila Polymorphism Database*. Both are publicly available at http://pdap.uab.es and http://dpdb.uab.es.

Université Libre de Bruxelles-SCMBB - *A Snow/aMAZE demo*

Snow is a user-friendly interface for querying and browsing databases of networks. Snow supports iql, a simple language enabling users to perform complex queries knowing only the database conceptual model. Results can be viewed as lists and graphs. For this demonstration, Snow will be interfacing aMAZE, our database for molecular interactions and cellular processes.

University of California Berkeley - *MUSCLE: Faster and More Accurate Multiple Sequence Alignment*

MUSCLE is the most accurate multiple sequence alignment program (up to 5% better than T-Coffee on benchmarks) with speed similar to CLUSTALW. With appropriate options, MUSCLE achieves accuracy equal to T-Coffee and aligns 5,000 sequences of length 350 in 10 minutes. MUSCLE is a free download at http://www.drive5.com/muscle.

University of Georgia - *MAGIC Database: A Tool for Gene Discovery and Expression*

The MAGIC Database and its associated tools provide a Modular Approach to a Genomic, Integrated and Comprehensive resource for gene discovery and expression. It manages and visualizes both DNA sequence and microarray data, and is suitable for both individual research groups and core facilities. Additional information is at http://fungen.org.

University of Leeds - *Databases and Software Available at Leeds University*

A wide range of projects are pursued by the Leeds bioinformatics research group. Those to be demonstrated include: 1, metaSHARK: a database of automated metabolic reconstructions derived from genomic DNA sequence. http://bioinformatics.leeds.ac.uk/shark/ , 2, TmaDB: A repository for tissue microarray data. 3, AtGATA_Base: an Arabidopsis thaliana GATA transcription factor database for discovering the potential regulated targets of GATA factors. 4, A 3D protein-ligand binding site database for structure-function relationship discovery. 5, Software to predict protein-protein binding.

Vanderbilt University Medical Center - *Metaclinic Database System*

The Metaclinic Database System (see poster session) provides a set of web-based tools for rapidly creating clinical research databases. During the demonstration (about 1.5 hrs), I will start from scratch and generate a simple research database, import data from an existing demo, and demonstrate the built-in reporting capabilities of the system.

Weizmann Institute of Science - *The LAMA and CYRCA Protein Profile-to-Profile Alignment and Multiple Profile Alignment Methods*

LAMA and CYRCA search, align, and cluster protein motif alignments (Blocks). Together they provide a sensitive and selective approach to predict protein function, structure and evolution. These tools improve and complement the abilities of other advanced protein analysis methods. The presentation will describe the methods' use, benefits, and limitations.