BioSapiens

A European Virtual Institute for Genome Annotation

www.biosapiens.info

The objective of the BioSapiens Network of Excellence is to provide a large-scale, concerted effort to annotate genome data – with particular focus on the human genome – by leading laboratories distributed around Europe, using both informatics tools and input from experimentalists.

The aim of the European Virtual Institute for Genome Annotation is to help to improve and coordinate European bioinformatics research, and to make our results easily available for all,

- by providing a common infrastructure for the integration of annotations by combining various methods in bioinformatics,
- by organising European meetings and workshops to encourage cooperation, rather than duplication of effort, and
- by supporting the development of novel methods to improve genome annotations.

An important goal is to achieve closer integration between experimentalists and bioinformaticians through a directed programme of genome analysis, focused on specific biological problems.
The annotations generated by the Institute are available in the public domain and easily accessible on the web.

BIOSAPIENS NETWORK

The BioSapiens Virtual Institute for Genome Annotation

The Virtual Institute currently provides a catalogue of genome annotations generated by the members of the Network. This is made available through the Distributed Annotation System (DAS), a standard technology used by genome browsers such as Ensembl. This flexible approach allows the integration of any additional external annotation and permits users to customise their views of a genome to address their specific biological problems. Further technologies are being developed for the improved visualisation of the various annotations.

The BioSapiens Virtual Institute for Genome Annotation provides annotations at the following levels:

Genome annotation

- gene definition (alternative splicing)
- regulators and promoters
- expression
- variation (haplotypes and SNPs)

Proteome annotation

- protein families and orthologue definition
- protein structure and modelling
- membrane proteins and ligands
- post-translational modification and localization

Network annotation

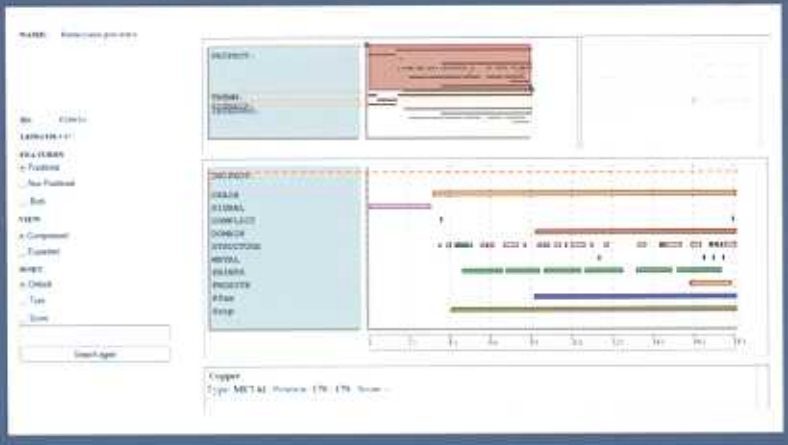
- sequence and structure to function
- protein–protein complexes
- pathways and networks

The BioSapiens Network is focussed on the human genome, but will apply the same technology to specific biological problems in the context of the collaboration with experimental groups. These thematic packages include:

- Down's syndrome genomic region
- Infectious diseases (HIV and HCV)
- Your project – contact us!

BIOSAPIENS NETWORK

Genome Annotation and DAS



The illustration shows annotations on Rusticyanin (a cupredoxin, RUS2_THIFE, P24930) as displayed by the UniProt DAS client (Dasty). The annotations are collected on demand from a number of different DAS servers around Europe. Each feature potentially provides a link to a web page containing further information about that particular annotation.

The Distributed Annotation System (DAS; http://biodas.org/) is a communication protocol that was proposed in 2000 as a way of federating genomic annotation. It was designed as a simple model, built on top of the well known HTTP and XML standards that allowed genomic annotation to be shared, but has since proven to be generic enough to also allow for other types of annotations. The protocol has been semantically extended to serve protein feature annotations and protein structure data.

The BioSapiens European School of Bioinformatics and Outreach Activities

There is a clear need to train and recruit creative and innovative young scientists in bioinformatics, and at the same time help users in experimental labs to keep up with developments in the field.
The BioSapiens Network of Excellence provides extensive training at all levels, from basic bioinformatics courses for experimentalists to more advanced training for experienced computational biologists.

Bringing bioinformatics to European researchers

The Network has established a permanent European School of Bioinformatics, which offers intensive training with a substantial practical component. Its objective is to give a basic overview of the bioinformatics methods and tools available to the community.

The school is held twice a year, in a different country every time, and consists of a six-day entry-level course for inexperienced users and newcomers to the field. It is followed, in the same location, by an advanced workshop on a more specialised topic. Previous workshops have covered molecular interactions (in collaboration with the ESF programme in Functional Genomics), and modeling G-protein-coupled receptors.

A proportion of the BioSapiens training effort is being directed towards sharing expertise in bioinformatics training by the organisation of information on Masters' courses in bioinformatics across Europe.

Advanced training in bioinformatics

The European School of Bioinformatics also organises advanced workshops for bioinformatics experts, including members of the BioSapiens Network of Excellence, in order to continue scientific training for those members working in the areas of science supported by this network. For example, we recently held a workshop on the Distributed Annotation System to help laboratories make their genome annotations publicly available using the BioSapiens DAS portal.

BioSapiens provides sponsorship for young researchers to attend the European Conference on Computational Biology (ECCB), where the network organises specialised workshops on genome annotation, with two very successful events in Glasgow 2004 and Madrid 2005. The Network has also collaborated in the organization of the CASP and CAPRI protein structure, interaction and function community assessment activities.

To access archived material from previous European Schools of Bioinformatics, ECCB annotation workshops or other BioSapiens activities, and to register for upcoming schools, please see http://www.biosapiens.info

About the network

Partners

Steering committee

Janet Thornton, EMBL-European Bioinformatics Institute: Scientific Coordinator (1A)

Søren Brunak, Technical University of Denmark: Work package coordinator (15)

Anna Tramontano, University of Rome "La Sapienza": Training Coordinator (9)

Alfonso Valencia, Consejo Superior de Investigaciones Científicas: Outreach Coordinator (4)

Other partners

European Molecular Biology Laboratory - Peer Bork (1B)

German National Centre for Health and Environment - Dmitrij Frishman (2)

Université Libre de Bruxelles - Jacques van Helden (3)

Institut Municipal d'Assistència Sanitària - Roderic Guigó (5)

The Wellcome Trust Sanger Institute - Richard Durbin (6)

Max-Planck Institute for Informatics - Thomas Lengauer (7)

The Hebrew University of Jerusalem - Michal Linial (8)

Stockholms Universitet - Gunnar von Heijne (10)

University of Oxford - Richard Mott (11)

University College London - Christine Orengo, David Jones (12)

Radboud University Nijmegen - Gert Vriend (13)

Swiss Institute of Bioinformatics - Anne-Lise Veuthey (14)

University of Helsinki - Esko Ukkonen (16)

University of Geneva - Stylianos Antonarakis (17)

Institute of Enzymology (Hungarian Academy of Sciences) - László Patthy (18)

Universität zu Köln - Dietmar Schomburg (19)

Institut Pasteur - Antoine Danchin (20)

BioInfoBank Institute - Leszek Rychlewski (21)

Max-Planck Institute for Molecular Genetics - Martin Vingron (22)

Genoscope – Vincent Schachter (23)

University of Bologna – Rita Casadio (24)

Scientific Advisory Board

Barry Honig, Columbia University

Siv Andersson, University of Uppsala

Janan Eppig, The Jackson Laboratory

Wendy Filsell, Unilever

Ian Harrow, Pfizer

Veronica van Heyningen, MRC Human Genetics Unit

Minoru Kanehisa, Kyoto University

Jonathan K.C. Knowles, Hoffman -La Roche

Trudy Mackay, NC State University

Carlos Martinez, Centro Nacional de Biotecnologia

Iain Mattaj, EMBL

Gert-Jan B. van Ommen, University of Leiden

Kai Simons, Max Planck Institute of Molecular Cell Biology and Genetics

Contacts

Project manager, Kerstin Nyberg, knyberg@ebi.ac.uk

Technical coordinator, Phil Gardner, pgardner@ebi.ac.uk

Coordinated by EMBL-EBI

The BioSapiens project is funded by the European Commission within its FP6 Programme, under the thematic area "Life sciences, genomics and biotechnology for health," contract number LHSG-CT-2003-503265.

Designed and Produced by PickeringHutchins www.pickeringhutchins.com