



For immediate release



European Virtual Institute for Genome Annotation receives 12 million Euro

Hinxton, January 23, 2004 - The Commission of the European Union has awarded 12,000,000 Euro to 24 bioinformatics groups based in 14 countries throughout Europe to create a pan-European BioSapiens Network of Excellence in Bioinformatics. The network aims to address the current fragmentation of European bioinformatics by creating a virtual research institute and by organizing a European school for training in bioinformatics.

Data overload is reaching epidemic proportions among molecular biologists. As genome-sequencing efforts continue apace and are being supplemented by new types of information from microarray, proteomics and structural genomics projects, biologists are literally drowning in a sea of data. Bioinformatics - the science of storing, retrieving and analysing large amounts of biological information - is struggling to keep up and is also contributing to the information overload by generating large numbers of predictions about the biochemical functions of gene products. These predictions need to be tested in the lab, but the infrastructure to "complete the circle" between computational biologists and experimentalists needs to be improved. This will have to change if we are to fulfil the ultimate promise of genomics: better quality of life.

"Europe has excellent bioinformatics environments in many countries, but in order to maximize the overall impact it needs to strengthen and reinforce that excellence by restructuring and coordinating existing research capacities and the way research is carried out," explains Janet Thornton, Director of the European Bioinformatics Institute (EBI) and coordinator of the BioSapiens project. To help realize the goal of a single European research area, which aims to make the best use of Europe's research resources, the Commission of the European Union has devised some new instruments as part of its sixth Framework

Programme (FP6), the EU's main means of funding research in Europe. One of these instruments, the "Network of Excellence" (NoE), is designed to tackle the fragmentation of European research by creating durable structures for future research in certain priority areas, including life sciences, genomics and biotechnology for health.

"The BioSapiens Network of Excellence captures the most important objectives of an NoE," explains Prof. Thornton. "Firstly, it will coordinate and focus excellent research in bioinformatics, by creating a Virtual Institute for Genome Annotation. Annotation is the process by which features of the genes or proteins stored in a database are extracted from other sources, defined and interpreted. Secondly, the Institute will establish a permanent European School of Bioinformatics, to train bioinformaticians and to encourage best practice in the exploitation of genome annotation data for biologists. Thirdly, whilst BioSapiens is primarily a basic research network, it will indirectly benefit the exploitation of biological information to address important social objectives, including improved health-care, better drugs, new vaccines, personalized medicine, and improved understanding of diet and health. By understanding how the normal human organism functions and develops we can improve diet, behaviour and environment to optimize quality of life," she adds.

BioSapiens is coordinated by a steering committee comprising Janet Thornton (chair), Søren Brunak (Technical University of Denmark), Anna Tramontano (University of Rome "La Sapienza") and Alfonso Valencia (Consejo Superior de Investigaciones Científicas, Madrid), and a project manager, Kerstin Nyberg (EBI).

How will this virtual institute work? It will be divided into nodes, each focused on one aspect of genome anno-

tation. The annotations generated will be integrated and made freely accessible to all through a single portal on the web, and will be used as a means of guiding future experimental work. "Experimental validation of a statistically significant subset of computational predictions will be an integral part of the process, leading to an iterative improvement in methods," explains Thornton.

The annotations will be integrated using DAS (Distributed Annotation System), an Open Source system developed by researcher Lincoln Stein and colleagues at Cold Spring Harbor Laboratory (NY, USA) for exchanging annotations on genomic sequence data. "DAS heralds a new era for database structure, where information is distributed by a network rather than a single site," explains Søren Brunak.

Meetings and workshops organized by the institute will encourage cooperation and reduce duplication of effort. They will also be an important medium for fostering closer collaboration between experimentalists and bioinformaticians. Some of these events will be tailored to industry, whose participation the network is keen to encourage. "The development of methods, tools, and servers in close interaction with experimentalists is one feature that distinguishes the network from previous pan-European efforts in bioinformatics," says Thornton, "and although there are 24 formal partners, BioSapiens is not a closed shop: once the infrastructure is established, a primary goal is to make this an open network to promote bioinformatics throughout Europe."

One important aim of the network will be to set up a permanent school of bioinformatics. "There is a clear need to train and recruit creative and innovative young scientists in bioinformatics and at the same time to help users located in experimental labs to keep up with the developments in the field," explains Anna Tramontano, who will coordinate the school's activities. "The network will provide extensive training at all levels, from basic courses for experimentalists to more advanced training for bioinformatics experts," she adds.

As well as providing understanding that will contribute to better health care in the future, a coordinated bioinformatics effort in Europe could have far-reaching economic impact. "The network will stimulate Europe's economic growth by creating new business themes and employment, improving European competitiveness in the bioinformatics and life science industries, and promoting mobility and knowledge sharing," explains Janet Thornton. "BioSapiens will also help to maintain Europe's strong global position in bioinformatics, allowing Europe to compete with the major investments made in this area in the USA, Canada and Japan. When Europeans work together, maximizing collaboration and minimizing duplication, we are better able to meet major challenges such as exploiting the ever-increasing volumes of data and ensuring Europe's full participation in global scientific initiatives," she concludes. ●

About EMBL:

The European Molecular Biology Laboratory is a basic research institute funded by public research monies from 17 member states (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom). Research at EMBL is conducted by approximately 80 independent groups covering the spectrum of molecular biology. The Laboratory has five units: the main Laboratory in Heidelberg, and Outstations in Hinxton (the European Bioinformatics Institute), Grenoble, Hamburg, and Monterotondo near Rome. The cornerstones of EMBL's mission are: to perform basic research in molecular biology; to train scientists, students and visitors at all levels; to offer vital services to scientists in the member states; and to develop new instruments and methods in the life sciences. EMBL's international PhD Programme has a student body of about 170. The Laboratory also sponsors an active Science and Society programme. Visitors from the press and public are welcome.

About the EBI:

The European Bioinformatics Institute (EBI) is part of the European Molecular Biology Laboratory (EMBL) and is located on the Wellcome Trust Genome Campus in Hinxton near Cambridge (UK). The EBI grew out of EMBL's pioneering work in providing public biological databases to the research community. It hosts some of the world's most important collections of biological data, including DNA sequences (EMBL-Bank), protein sequences (UniProt), animal genomes (Ensembl), three-dimensional structures (the Macromolecular Structure Database) and data from microarray experiments (ArrayExpress). The EBI hosts several research groups and its scientists continually develop new tools for the biocomputing community.

Press Contacts:

Cath Brooksbank PhD, EMBL-EBI Scientific Outreach Officer

Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK, Tel: +44 (0)1223 492525,

Fax: +44 (0)1223 494468, www.ebi.ac.uk, cath@ebi.ac.uk

Trista Dawson, EMBL Press Officer

Meyerhofstrasse 1, 69117 Heidelberg, Germany, Tel: +49 6221 387 452, Fax: +49 6221 387 525,

www.embl.de, dawson@embl.de

Policy regarding use

EMBL press releases may be freely reprinted and distributed via print and electronic media. Text, photographs & graphics are copyrighted by EMBL. They may be freely reprinted and distributed in conjunction with this news story, provided that proper attribution to authors, photographers and designers is made. High-resolution copies of the images can be downloaded from the EMBL web site: www.embl.de