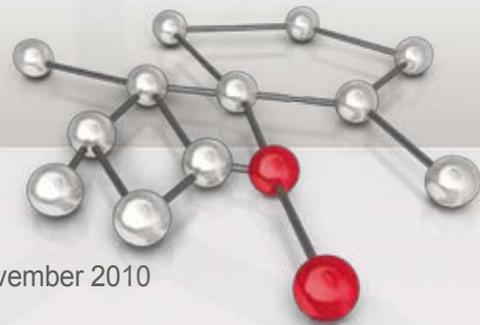


SWISS BIOINFORMATICS

A newsletter published by SIB Swiss Institute of Bioinformatics



Editorial



Researchers do not all study the same disease, virus or insect, yet they all need to store the huge amount of data they collect and to compare their results with up-to-date knowledge produced by their peers. In other words, they need to access state-of-the-art databases and software. SIB Swiss Institute of Bioinformatics supports the life science community in many ways. It has, and continues to develop a wide range of resources, such as SWISS-MODEL, which was created almost 20 years ago and has become essential to researchers worldwide, UniProtKB/Swiss-Prot, STRING, Melanie, EADock or neXtProt, the new platform about to be launched and

dedicated to human proteins. SIB also applies its expertise to collaborative projects aimed at curing diseases such as diabetes. SIB has always known that collaboration is essential and that it must be fostered both internally and externally. By working closely together SIB service and research groups are indeed creating links between their databases and software platforms and this interoperability gives a new dimension to the resources supporting the life science researchers.

To ensure the future of life science research and progress it is, however, not enough to create such resources. It is also crucial to guarantee their reliability based on two criteria: quality and sustainability. What kind of support would a database represent if it disappeared from one day to the next? Years of hard work would simply vanish in the twinkling of an eye. It is therefore essential to create a solid infrastructure to secure the future of bioinformatics expertise and resources at the international level. This is precisely the objective of ELIXIR, a project launched by the European Union. SIB is proud to be part of this European effort and looking forward to its results and impact.

We hope you enjoy reading the new issue of *Swiss Bioinformatics*, which will provide more examples of our activities on the national and international bioinformatics scene.

Ron Appel, Executive Director

● November 2010

● Contents

● Editorial	1
● News in brief	1
● Research	2
● International collaboration	2
● Education and outreach	3
● Link with the industry	3
● About Bioinformatics and SIB	4

News in brief

NIH grant for UniProt

The National Institutes of Health (NIH, USA) awarded a three-year funding to the UniProt consortium, which includes SIB, EBI (European Bioinformatics Institute) and PIR (Protein Information Resource). SIB will thus benefit from a yearly support of more than two million Swiss francs for the maintenance and enhancement of the UniProtKB/Swiss-Prot knowledge base. This grant underlines the importance and international outreach of SIB and its research and service activities.

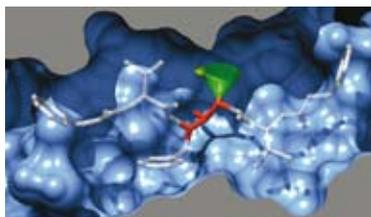
Next step, neXtProt

The first public version of neXtProt, the new knowledge and software platform dedicated to human proteins, will be released in the coming days. neXtProt is expected to become a focal and essential resource for the many biomedical researchers interested in the role of proteins in health and disease. It was developed jointly by SIB group CALIPHO and by GeneBio. Such a collaboration is necessary because it ensures that the platform is designed to meet the needs of both academic and industrial researchers. neXtProt capitalises on the high-quality annotation work carried out in the Swiss-Prot group and it will also synergise with other resources developed by SIB. The first version already makes use of data and analysis encapsulated in Bgee, a database developed by the Evolutionary Bioinformatics Group at the University of Lausanne and SIB.

This edition is only a glimpse into the full scope of data and tools that neXtProt wants to provide for its users, yet it is an important first step that contributes to one of SIB's two core missions, namely to provide quality resources to the life science research community.

The neXtProt server, hosted by the Vital-IT group, is available at beta.nextprot.org.

EADock now online



Malfunctioning proteins cause many diseases, including cancers. Consequently, the role of many anti-cancer drugs consists in modulating their activities. These drugs are small molecules designed to fit into the protein's active site in

order to prevent or permit a given function, similar to finding the right key to open/close a lock.

Predicting such interactions at the atomic scale and in three-dimensional space has been one of the major research themes investigated by SIB's Molecular Modelling Group over the past 10 years. This effort culminated in a computer program called EADock that allows for successful harvesting of many molecules active against various cancer-related target proteins.

Today the know-how accumulated in this software has been made available to the entire scientific community. The software can now be used freely online through the SwissDock website^(*). SIB intends to broaden the use of these methods well beyond just experts in the field, which should lead to many significant advances in molecular oncology and, more generally, molecular medicine in the years to come.

^(*) www.swissdock.ch

Joint efforts against diabetes

To fight the growing diabetes pandemic, academia, biotechnology and the pharmaceutical industry have joined forces in the framework of IMIDIA, the Innovative Medicines Initiative for Diabetes. Professor Bernard Thorens of the University of Lausanne leads



this project that was officially launched on 14 June 2010. It will be looking to develop biomarkers and tools for better disease management and ultimately provide a cure for diabetes.

SIB, along with several research and service groups, will contribute to this challenging adventure both in terms of bioinformatics, computational and data analysis support.

Research

SWISS-MODEL and STRING: services with international outreach

There are about five million proteins in our body produced by 20,000 different genes, so it should not surprise anyone that determining all their roles and interactions will keep the research community busy for many years. One of SIB's missions is to support researchers with core databases and software in the domain of proteins, and some of the services provided by the SIB have become essential tools for scientists worldwide.

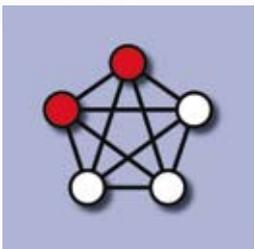


SWISS-MODEL is a server for modelling three-dimensional protein structures. Biochemists and molecular biologists worldwide have been using this tool increasingly, appreciating its free access and easy utilisation. Users simply enter an amino-acid sequence or the UniProt accession code of a protein to build the corresponding 3D model. Besides this

fully automated method for easy modelling cases, SWISS-MODEL offers several operation modes that permit advanced users to fine tune relevant parameters, such as selection of the most appropriate structural template. The SWISS-MODEL Repository provides pre-computed models for human proteins and a selection of relevant model organisms to facilitate data exchange with other services such as Swiss-Prot and STRING. Protein structure models find widespread applications in biomedical research, such as structure-based drug design, rational planning of functional studies, or protein engineering.

Initiated in 1993, SWISS-MODEL has undergone constant improvement and, since 2001, its work has been overseen by Torsten Schwede's Structural Bioinformatics Group at the Biozentrum (University of Basel) and SIB. Every year, more than 200,000 users from all around the world access the server, requesting about half a million models for protein structures.

SWISS-MODEL can be found at: swissmodel.expasy.org.



STRING Identification and structural modelling represent merely the first steps towards understanding a protein's functions. The situation is in fact far more complex because proteins have what could be called a social life, since they interact with each other in myriad ways; this results in functional associations and an increased complexity for functional annotations.

Considerable information is needed to describe the various aspects of a given protein-to-protein association, and the actual annotations and interaction records are scattered over a number of public resources. It is important for those who use such data to have an integration and re-appraisal resource that can be

searched and browsed easily on a single site. The STRING database resource (Search Tool for the Retrieval of Interacting Genes) provides such a service by acting as a "one-stop shop" for all information on functional links between proteins. The main strengths of STRING lie in its comprehensiveness, its confidence scoring and its interactive and highly intuitive interface. It is the only site to cover hundreds (and soon more than 1,100) of organisms ranging from bacteria to humans.

STRING began in 2000 at the EMBL (European Molecular Biology Laboratory) in Heidelberg, and is now being developed and maintained by a small consortium of bioinformatics groups in Europe. SIB plays a crucial role in further developing the resource by providing substantial manpower in Christian von Mering's group (University of Zurich and SIB). STRING is used worldwide and many scientists download and work with its data offline, making it an important cornerstone of research projects internationally.

The STRING resource can be found at: string-db.org.

INCREASED INTEROPERABILITY

With 29 research and service groups based in different Swiss institutions and towns, SIB is a model in terms of inter-institutional and inter-cantonal collaboration. Today, SIB groups work towards the reinforcement of this collaborative aspect, by creating links between the panel of services and resources they create and maintain for the international life science community. Concretely, this means that they aim to give users the possibility to work with several resources at once, for more user friendliness and increased efficiency. STRING and SWISS-MODEL have already taken great strides in this direction. Today, users displaying a 3D protein structure in SWISS-MODEL have for example the possibility to open windows containing information from the STRING database and vice-versa. Our services therefore reach a new dimension and new levels of efficiency, which are highly appreciated by the research community.

International collaboration

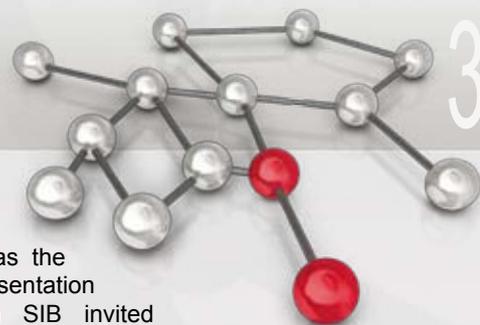
ELIXIR of life

How would life scientists carry out their research without having access to all the knowledge brought together to date in valuable biological databases? How would they, for example, be able to identify a virus if the database containing all relevant information, such as gene sequences, happened to disappear? Likewise, hopeful expectations, such as efficient cancer prevention or treatments, would simply fade away.

The answer to this potentially alarming scenario is ELIXIR – European Life Science Infrastructure for Biological Information, an initiative launched by the European Union that aims to guarantee the future of biological data. Its objective is to secure funding commitments from government agencies, charities, industry and intergovernmental organisations throughout Europe, as well as to strengthen and sustain a world-class infrastructure for managing and integrating information in the life sciences. This goal can be reached by concentrating on two major factors: coordination and the guarantee of sustained funding.

What will ELIXIR provide?

The project involves all stakeholders on the bioinformatics scene in Europe, from data providers to bioinformatics users, to assess all the needs and available resources. This will result in a structured bioinformatics network, which will provide data, tools, coordination, integration, support, standards and training to the whole life science research community. In turn, this will translate into an improved quality of life with better healthcare for ageing populations, sustainable food supply, competitive pharmaceutical and biotechnical industries and increased environmental protection.



How will Switzerland and SIB contribute to this project?

Serving as the Swiss life science infrastructure for biological information, SIB already is at the national level what ELIXIR intends to reach at the European level. Indeed, to date and thanks to the support of the Swiss Confederation, our country is the only one having such a solid bioinformatics infrastructure. And there is no doubt that SIB, as a provider of various renowned resources to the international life sciences community, will play an important role in the ELIXIR project.

For further reading: www.elixir-europe.org.

Ties with India strengthened



In 2009, the Science and Technology Counselor for the Swiss Embassy in New Delhi and SIB decided to gather the Swiss and Indian bioinformatics institutes to discuss possible collaborations. The two bodies hence organised a common symposium at the Indian Institute of Technology (IIT) in New Delhi during the last week of October, where SIB and major Indian bioinformatics institutes presented their activities and projects. The open exchanges in the presence of representatives of the Indian Government are expected to pave the way for a fruitful collaboration.

Education and outreach

SIB provides professional bioinformatics courses, workshops and consulting for life scientists. The demand for such courses is strong, as shown by a high attendance rate and waiting lists. Below are listed several courses and workshops which occurred throughout the current year:

Four one-week courses

- Introduction to statistics for biologists
- Perl and presentation techniques
- Computational analysis of UHT sequencing data
- Advanced statistics: statistical modelling

One-day and two-day courses or workshops

- Perl programming for life sciences
- High-performance computing (HPC) in the life sciences
- Protein identification by mass spectrometry.

All sessions were fully booked and the one- or two-day workshop sessions even had to be doubled to meet the high demand. The quality and the response to current needs explain the success of these courses.



This year's highlight was the session on "Perl and presentation techniques", for which SIB invited Damian Conway, one of the best Perl language specialists in the world, to fly in from Australia. Perl (Practical Extraction and Report Language) is a script programming language which is popular with people who regularly have to write programs for data processing. Monica Aguilera, a post-doctoral researcher in biology at EPFL (Federal Institute of Technology in Lausanne) who attended the session, says: "I found the courses on Perl very interesting. They gave a good overview of the possibilities offered by this programming language. Damian Conway gave us clear explanations and managed to capture our attention during the whole session." As she needs to analyse the high-throughput data she produces during her research, Monica attended other SIB courses and says she hopes there will be more to come.

Based on this success, the education team is already planning the programme for 2011, which will include:

- an introductory course on statistics
- a Unix/Perl course
- an introductory course on microarray data analysis
- an advanced course on statistics
- the renewal of some of this year's workshops and several new ones on diverse bioinformatics projects.

All Bioinformatics courses in one click!

To promote bioinformatics and coordinate bioinformatics education in Switzerland is part of SIB's missions. Already organiser of most bioinformatics courses, SIB now provides a complete list of all courses available in Switzerland on its Website (www.isb-sib.ch/education). With this new service, SIB means to facilitate students' access to the information and thus also their subscription to the courses. Students have the possibility to sign up to a mailing-list to automatically receive the announcements of any new course proposed. About 1,500 people – mainly students from past courses - have subscribed so far and many more are expected to join.

Link with the industry

MELANIE, software that makes the difference

MELANIE is a software package for analysing images from two-dimensional electrophoresis gels. On each gel a mixture of proteins – from a blood or tissue sample, for instance – is separated by size and charge to create a protein map.

The software allows comparison between multiple samples or gels and measures protein expression changes for statistical exploration and analysis. This is required for experiments that study the difference between normal and malignant cells, for example, or to investigate the effect of drugs on an organism.



The goal is to identify specific proteins that differentiate samples, or to study patterns of protein co-expression that offer potential insights into protein regulation and interactions. This information can be used to develop biomarkers for diagnostic testing or to identify likely drug targets.

Melanie started as a PhD project in 1983 and has grown to be currently used by thousands of customers worldwide since its first commercial version in 1994. It is developed and continuously improved by the SIB Proteome Informatics group led by Frédérique Lisacek, under the supervision of Daniel Walther who has been heading the software development team since 1996. The team

collaborates closely with Sonja Voordijk, Melanie's product manager at GeneBio, and with GE Healthcare through which the software is also sold as ImageMaster™ 2D Platinum. This cooperation with GeneBio and GE Healthcare brings in royalties, which allow SIB to finance further developments of the Melanie software.

Researchers from pharmaceutical, biotech and academic groups rely on Melanie for their 2D gel experiments to answer questions such as: "Are there proteins or protein patterns that might be characteristic of a tumour?"

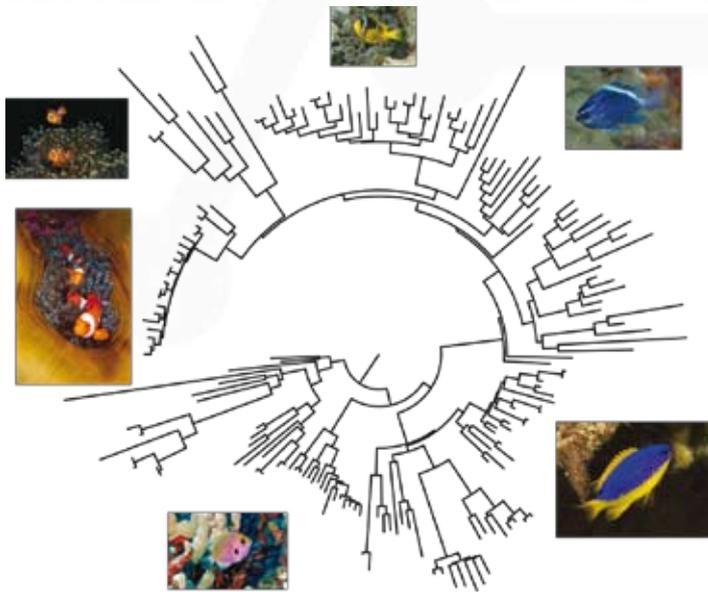
The various analytical methods in Melanie address these questions and guide the choice of proteins for further analysis by mass spectrometry.

(ImageMaster is a trademark of GE Healthcare companies.)

What is bioinformatics?

Over the past 30 years new biological research techniques, along with developments in information technology, have increased both the amount and complexity of biological data. That is why scientists must often apply information technology to biological problems – a science called bioinformatics.

Life scientists use bioinformatics to store, process and analyse large amounts of data, advancing their knowledge and understanding of biological processes. This, in turn, can lead to scientific breakthroughs that enhance our quality of life in fields such as designing better medical treatments or improving crop yields.



About SIB www.isb-sib.ch

The SIB Swiss Institute of Bioinformatics is an academic not-for-profit foundation federating bioinformatics activities throughout Switzerland. Its two-fold mission is to provide world-class core bioinformatics resources to the life sciences research community at both the national and international levels in key fields such as genomics, proteomics and systems biology, as well as to lead and coordinate the field of bioinformatics in Switzerland.

It has a long-standing tradition of producing state-of-the-art software for the life sciences research community, as well as carefully-annotated databases. The SIB includes 29 world-class research and service groups that bring together 400 researchers in the fields of proteomics, transcriptomics, genomics, systems biology, structural bioinformatics, evolutionary bioinformatics, modelling, imaging, biophysics and population genetics in Basel, Bern, Geneva, Lausanne and Zurich. SIB expertise is widely appreciated and its infrastructure and bioinformatics resources are used by life science researchers worldwide.

Institutional members:

- Swiss Federal Institute of Technology Lausanne (EPFL)
- Swiss Federal Institute of Technology Zurich (ETHZ)
- University of Basel
- University of Bern
- University of Geneva
- University of Lausanne
- University of Zurich
- Ludwig Institute for Cancer Research (LICR)
- Friedrich Miescher Institute for Biomedical Research (FMI)
- Geneva Bioinformatics (GeneBio) S.A.
- Hewlett Packard

A full list of SIB research projects can be found at: www.isb-sib.ch/research/projects.

*Phylogenetic tree
© Nicolas Salamin and Glenn Litsios*



Swiss Institute of Bioinformatics

SIB | Swiss Institute of Bioinformatics

Quartier Sorge
Bâtiment Géopode
CH-1015 Lausanne
Switzerland
t +41 21 692 40 50
f +41 21 692 40 55
www.isb-sib.ch

Swiss Bioinformatics is written by SIB Communications with contributions from SIB Members. Design and layout D. Meyer