EMBnet.news

Volume 10 Nr. 4

December 2004

- EMBnet AGM2004
- The MyHits database and web site
- The wEMBOSS interface
- The EBI resources in a Nutshell (part 1)
- Project Management (part 4)

and more...

Editorial

The 3-year EMBCORE FP5 project has now ended. Despite initial financial issues due to a misunderstanding of the procedure, the outcome looks pretty satisfactory (see p8). The most promising outcome is the number of nice new projects emerging from discussions among members of the EMBnet community, some of which are described in this issue. The recently created SRS Federation consortium groups 7 EMBnet nodes to share database management under SRS. This project is currently in an early phase (see p11). The wEMBOSS project involves the Belgian and the Argentinian nodes in the creation and maintenance of a user-friendly web interface to the EMBOSS package and other software. It is already in production, with many nodes already using it (see p13). EMBRACE a FP6 project starting next year with several EMBnet nodes among the participants, (not described in this issue).

This activity reflects a living community of researchers with ideas and the will to improve and ease bioinformatics tools for non-bioinformaticians.

The editorial board: Erik Bongcam-Rudloff, Laurent Falquet, Pedro Fernandes, Oscar Grau, and Gonçalo Guimaraes Pereira.

proteinspotlight

> ONE MONTH, ONE PROTEIN <

Protein Spotlight (ISSN 1424-4721) is a periodical electronic review from the SWISS-PROT group of the Swiss Institute of Bioinformatics (SIB). It is published on a monthly basis and consists of articles focused on particular proteins of interest. Each issue is available, free of charge, in HTML or PDF format at `http://www.expasy.org/spotlight`

We provide the EMBnet community with a printed version of issue 52. Please let us know if you like this inclusion.

Cover picture: Trees covered by snow, winter 2004 [® Erik Bongcam-Rudloff]

Contents

Editorial Board:

Publisher:

EMBnet Administration Office
c/o Jack Leunissen
Dept. of Genome Informatics
Wageningen UR, Dreijenlaan 3
6703 HA Wageningen, The Netherlands
Email: Jack.Leunissen@wur.nl
Tel: +31-317-484074

The MyHits database and web server

Laurent Falquet, Vassilios Ioannidis, Lorenzo Cerutti, Monique Zahn-Zabal & Marco Pagni

Swiss Institute of Bioinformatics
Ch. des Boveresses 155,
1066 Epalinges
Switzerland

Introduction

In the world of the life sciences it often happens that the lab scientists concentrates on a particular set of proteins or particular families of proteins dealing with a pathway or disease. The functional analysis requires comparison with known proteins in the public sequence databases (e.g., using BLAST vs SwissProt) and searching for potential domains or motifs in the public motif databases (e.g., InterPro). Collecting this information and compiling it with private sequences and annotations is tedious due to the different formats of output files, the different information obtained using each tool and the different ways of representing the collection. This task is not easy, especially for occasional bioinformatics users, but requires a lot of parsing even for experienced bioinformaticians.

Based on these facts garnered after many years of helpdesk experience and teaching for the Swiss EMBnet node, we decided to construct a web site and a database to ease the task of the life scientist. We call it **MyHits**.

What is the MyHits database?

A **hit** is a region of a *sequence* matched by a *motif*. Calculating all the hits of a sequence *vs* a motif database (i.e., motif scan) requires heavy computing resources. Since more than 90% of the sequences don't vary from week to week, it is more efficient to calculate the hits once and store them in a database for later use. Moreover sequences and motifs are partially redundant, thus calculating the hits only on the **nr** (non-redundant) versions of sequences and motifs is even more efficient. MyHits automatically calculates and stores all the hits from nr sequence databases *vs* nr motif databases, allowing queries and visualization of hits and annotations.

This task is complex since the motif and sequence databases are updated asynchronously. We use a MySQL relational system to store and manage the smart automatic incremental update of the hit lists.

What is the MyHits web site?

The web site is designed to allow a user to query the hits database and to search **public** and **private** databases with his sequences or motifs. The motifs can be generated with tools provided in a user-friendly interface.

The web site implements the concept of "**hubs**" (sort of UNIX pipes) allowing the output of one program to be redirected to the input of another program.

Two types of users use the web site in different ways. The **anonymous** user has access to all public databases and some of the tools. The **registered** user, can access all the tools, organize and manage his private databases (e.g., mypep, mymsa, mypat…).

Description of the menus

Each page in the MyHits web site displays a menu that helps the user to navigate within the increasing number of tools. The menu bar is divided in five main menus (**Query**, **Hub**, **Result**, **Help**, **Database**) (Figure 1). Each menu contains one or more items or submenus.

The Query menu

PSI-Blast Based on the NCBI PSI-Blast it allows searches in public and private databases with a single sequence or a multiple alignment as query. In addition the output can be clustered according to a user-defined percent identity. It can also be limited to taxonomic species defined by the user.

Pattern search Simple PROSITE pattern search which allows private and public databases to be searched with a user defined pattern following the PROSITE syntax. It also allows the user defined taxonomic species.

Profile search PROSITE profile search starting with a sequence or a multiple alignment. Offers identical features as the PSI-Blast tool.

Motif scan Scans a sequence for the presence of selected motifs, with the option of including the weak matches.

by protein Similar to the motif scan, but instead of recalculating the hits, simply queries the MyHits database of pre-calculated matches (requires the protein to be already in the public databases).

by motif Reverse of the Query by protein, gives the list of proteins containing one or more motifs by querying the MyHits database of pre-calculated matches.

ClustalW Performs a multiple sequence alignment

T-Coffee Performs a multiple sequence alignment, but limited to small numbers of proteins

Translate DNA Translates DNA sequences (with EMBOSS transeq) in one or several reading frames, highlighting the ORF, Start and Stop codons.

The Hub menu

Protein Hub Hub receiving protein ID/AC from the user or other tools, allowing the user to visualize them graphically or using specialized tools (e.g., Dotlet).

MSA Hub Hub receiving multiple sequence alignments (MSA) from the user or other tools, allowing the user to visualize them graphically or using specialized tools (e.g., JalView).

The Result menu

(This menu is not currently implemented.)

The Help menu

Help pages Documentation describing the web site, databases, tools, references, etc…

The Database menu

Show all db Displays the list of current databases and their status

Hit lists status Displays the table of current pairs of motif and sequence databases used to generate the hit lists maintained in MyHits.

Edit my db Edit, Delete, Add sequences, MSA and motifs to the private user database

Edit my info Edit user personal information

What are the available public databases?

Sequences db (code)	Motifs db (code)
SwissProt (sw), TrEMBL(tr), splice variants (*var)	PROSITE pattern (pat)
Trest (te)	PROSITE profile (prf)
Trome (to)	Prefile (pre)
Trgen (tg)	InterPro (ipr)
ENSEMBL (ens)	
Microbial (microb)	

If you are a registered user, you have in addition 3 private databases currently:

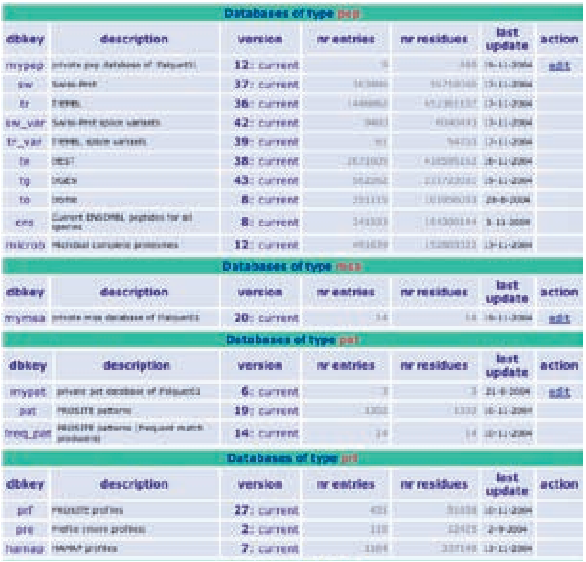
mypep peptide sequences of the user

mypat patterns of the user

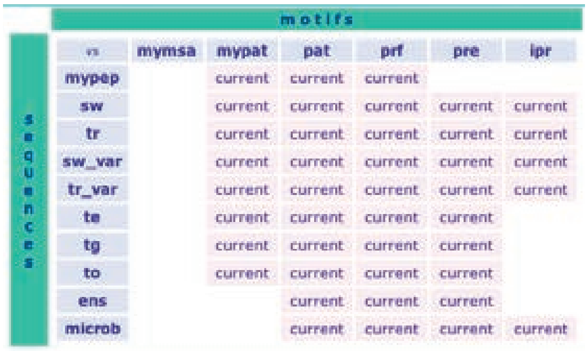
mymsa multiple sequence alignments of the user

We plan to add other private databases in the future (myprf, myhmm, mydna...).

You can visualize the current status of the databases by selecting "**Show all db**" in the database menu (Figure 2).



Only sequence databases connected with a motif database are taken into account for the automatic hit lists update. You can find this information by selecting "**Hit lists status**" from the database menu (Figure 3).



motifs

sequences

vs	mymsa	mypat	pat	prf	pre	ipr
mypep		current	current	current		
sw		current	current	current	current	current
tr		current	current	current	current	current
sw_var		current	current	current	current	current
tr_var		current	current	current	current	current
te		current	current	current	current	
tg		current	current	current	current	
to		current	current	current	current	
ens			current	current	current	
microb			current	current	current	current

How do I register to MyHits?

On the MyHits home page (`http://myhits.isb-sib.ch`) you can click "**a free registration**" to reach the registration page (Figure 4).

Welcome to MyHits registration form

Very Important Note: A few minutes after you click the "register" button you will receive an email with your username and an personal identification number that will be required for your first connection only.
Please keep your password in a safe place, it will not be send to you by email.

Register please

Mandatory please type your personal data, password and press "Register"

Lastname:
Firstname:
Email:
Password:
Password: (confirmation)
Are you an academic user ? ○ Yes ○ No

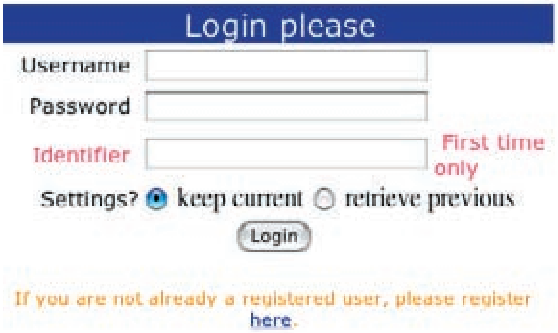
Optional please feel free to give us the following information

Telephone:
Institution:
Address:

Please click only once! The registration may require 15-30 seconds.

MyHits © 2003,2004 - ISB - Swiss Institute of Bioinformatics

Alternatively, on each page you find a "**log in**" button at the top-right side of the window. By clicking this button you reach the login page (Figure 5), where you find a link to the same registration page.



Login please

Username
Password
Identifier
First time only
Settings? ● keep current ○ retrieve previous

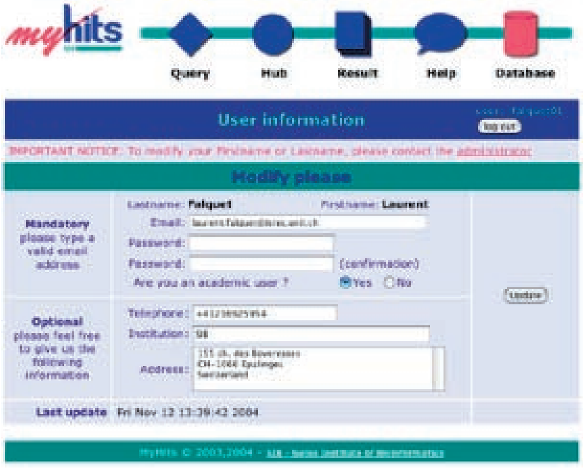
If you are not already a registered user, please register here.

Fill in your personal information: only your lastname, firstname, email, password and affiliation are required.

After clicking on the register button, you should receive an email with your username and an unique identifier. Please login within 48h using this information to validate your account (the unique identifier is required the first time only).

How do I modify my email?

You can always modify your personal information by selecting the "**edit my info**" from the database menu (Figure 6).



User information

User: falquet01

IMPORTANT NOTICE: To modify your Firstname or Lastname, please contact the administrator

Modify please

Mandatory please type a valid email address

Lastname: Falquet
Firstname: Laurent
Email: laurent.falquet@isrec.unil.ch
Password:
Password: (confirmation)
Are you an academic user ? ● Yes ○ No

Optional please feel free to give us the following information

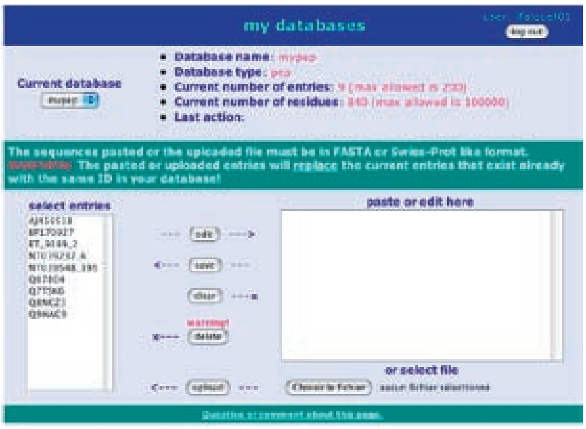
Telephone: +41216925954
Institution: SIB
Address: 155 ch. des Boveresses
CH-1066 Epalinges
Switzerland

Last update Fri Nov 12 13:39:42 2004

MyHits © 2003,2004 - ISB - Swiss Institute of Bioinformatics

How do I access my private databases?

Once you are logged in you can select "**edit my db**" from the database menu, and reach the "my databases" page (Figure 7).



my databases

User: falquet01

Current database: mypep

- **Database name:** mypep
- **Database type:** pro
- **Current number of entries:** 9 (max allowed is 230)
- **Current number of residues:** 840 (max allowed is 100000)
- **Last action:**

The sequences pasted or the uploaded file must be in FASTA or Swiss-Prot like format.
WARNING: The pasted or uploaded entries will replace the current entries that exist already with the same ID in your database!

aucun fichier sélectionné

Question or comment about this page.

Here you can add, edit, delete one or more entries in each of your databases (mypep, mymsa, mypat). FASTA or Pseudo-SwissProt formats are accepted.

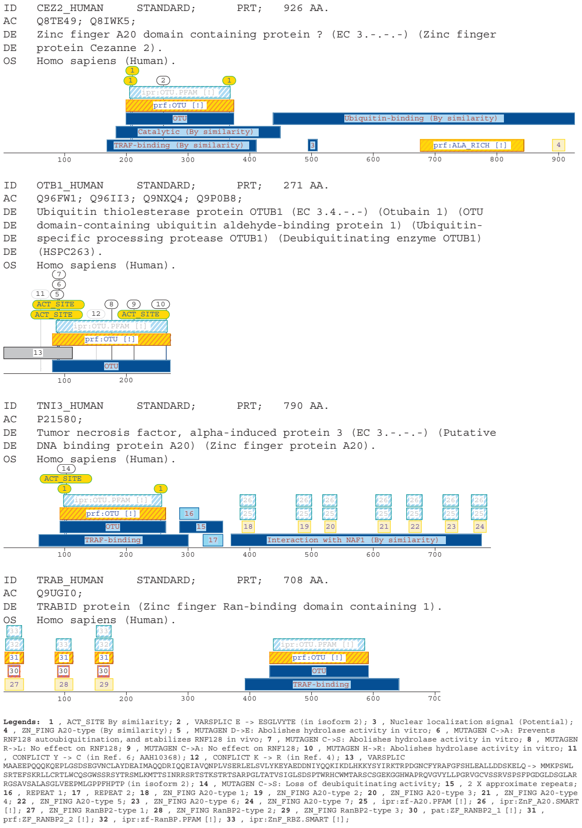
Once a private database has been modified, it takes some time (minutes to hours, depending on the server load), until the new sequences are available and the hit lists calculated. The previous version is still available in the meantime.

I want to define a new protein domain. How do I use MyHits?

A tutorial is available on the web. (`http://myhits.isb-sib.ch/doc/tutorial-psi.html`)

What is the catalogue?

The catalogue is available from both Hubs to graphically display motifs and annotations from a list of proteins. Several options allow the user to select the style of the output (one per line, either shrinked or to scale, or to scale on multiple lines), as well as the format (HTML, PostScript, PDF) (Figure 8).



```
ID   CEZ2_HUMAN     STANDARD;      PRT;   926 AA.
AC   Q8TE49; Q8IWK5;
DE   Zinc finger A20 domain containing protein ? (EC 3.-.-.-) (Zinc finger
DE   protein Cezanne 2).
OS   Homo sapiens (Human).
```

```
ID   OTB1_HUMAN     STANDARD;      PRT;   271 AA.
AC   Q96FW1; Q96II3; Q9NXQ4; Q9P0B8;
DE   Ubiquitin thiolesterase protein OTUB1 (EC 3.4.-.-) (Otubain 1) (OTU
DE   domain-containing ubiquitin aldehyde-binding protein 1) (Ubiquitin-
DE   specific processing protease OTUB1) (Deubiquitinating enzyme OTUB1)
DE   (HSPC263).
OS   Homo sapiens (Human).
```

```
ID   TNI3_HUMAN     STANDARD;      PRT;   790 AA.
AC   P21580;
DE   Tumor necrosis factor, alpha-induced protein 3 (EC 3.-.-.-) (Putative
DE   DNA binding protein A20) (Zinc finger protein A20).
OS   Homo sapiens (Human).
```

```
ID   TRAB_HUMAN     STANDARD;      PRT;   708 AA.
AC   Q9UGI0;
DE   TRABID protein (Zinc finger Ran-binding domain containing 1).
OS   Homo sapiens (Human).
```

Legends: **1** , ACT_SITE By similarity; **2** , VARSPLIC E -> ESGLVYTE (in isoform 2); **3** , Nuclear localization signal (Potential); **4** , ZN_FING A20-type (By similarity); **5** , MUTAGEN D->E: Abolishes hydrolase activity in vitro; **6** , MUTAGEN C->A: Prevents RNF128 autoubiquitination, and stabilizes RNF128 in vivo; **7** , MUTAGEN C->S: Abolishes hydrolase activity in vitro; **8** , MUTAGEN R->L: No effect on RNF128; **9** , MUTAGEN C->A: No effect on RNF128; **10** , MUTAGEN H->R: Abolishes hydrolase activity in vitro; **11** , CONFLICT Y -> C (in Ref. 6; AAH10368); **12** , CONFLICT K -> R (in Ref. 4); **13** , VARSPLIC MAAEEPQQQKQEPLGSDSEGVNCLAYDEAIMAQQDRIQQEIAVQNPLVSERLELSVLYKEYAEDDNIYQQKIKDLHKKYSYIRKTRPDGNCFYRAFGFSHLEALLDDSKELQ-> MMKPSWLSRTEFSKRLLCRTLWCQSGWSSRSYTRSMLKMTTSINRRSRTSTKSTRTSARPGLTATVSIGLSDSPTWRHCWMTARSCSGEKGGHWAPRQVGVYLLPGRVGCVSSRVSPSFPGDGLDSGLARRGSAVSALASGLVEEPMLGPPFHPTP (in isoform 2); **14** , MUTAGEN C->S: Loss of deubiquitinating activity; **15** , 2 X approximate repeats; **16** , REPEAT 1; **17** , REPEAT 2; **18** , ZN_FING A20-type 1; **19** , ZN_FING A20-type 2; **20** , ZN_FING A20-type 3; **21** , ZN_FING A20-type 4; **22** , ZN_FING A20-type 5; **23** , ZN_FING A20-type 6; **24** , ZN_FING A20-type 7; **25** , ipr:zf-A20.PFAM [!]; **26** , ipr:ZnF_A20.SMART [!]; **27** , ZN_FING RanBP2-type 1; **28** , ZN_FING RanBP2-type 2; **29** , ZN_FING RanBP2-type 3; **30** , pat:ZF_RANBP2_1 [!]; **31** , prf:ZF_RANBP2_2 [!]; **32** , ipr:zf-RanBP.PFAM [!]; **33** , ipr:ZnF_RBZ.SMART [!];

Can I print the results?

Yes, most web pages are printable. Moreover you can even generate PDF files from either the Jalview multiple sequence alignment editor (select "send PDF to browser" in the file menu of Jalview), or from the catalogue tool (see above).

Future developments

We plan to improve the MyHits web site and database by consolidating the current interface, by adding new tools (Pftools, HMMer, ESTScan, EST clustering, Coiled-coils, TMPred, SignalP), new databases (DNA, ELM, Tigrfam) and new applets (to annotate MSA and plasmids). We are currently working on a stand-alone package distribution of our automatic incremental update system.
We also wish to integrate MyHits in distributed environments of data and computing power (i.e., GRID).

References

Pagni M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongeneel CV, Falquet L.
MyHits: a new interactive resource for protein annotation and domain identification.
NAR 2004 Jul 1;32:W332-335.

The Belgian EMBnet node announces:

The second edition of the Bioinformatics System Management Workshop (BSMW05) will be organised by the staff of the Belgian EMBnet node, **from April 6 to April 8 2005**. The course will take place in the computer room of the Institut de Biologie et de Médecine Moléculaires of the Université libre de Bruxelles, situated in Gosselies, close to Charleroi.
Last year's course attracted 14 participants from 5 different countries. The aim of the course is to train each participant in installing a complete bioinformatics infrastructure on a personal computer.

The subjects covered are:
- Basic intallation of the operating system on a naked PC
- Installation of network utilities (Apache server, etc.)
- Installation and configuration of the EMBOSS bioinformatics sequence analysis package

- Installation and configuration of a basic SRS 7 server
- Installation under SRS 7 of a nominal set of databanks (part of EMBL, Swissprot, Prosite, etc.)
- Installation of the wEMBOSS interface to the EMBOSS software
- Installation of a few databank similarity search software packages (Blast, Fasta, ...)
- Installation of wrappers for Blast and Fasta under wEMBOSS
- A set of user cases to be solved by the participants

Participation fee is 200€ per person. This fee covers the coffee breaks and lunches. Travel and residence costs will be paid by the participants.
The IBMM building is very close to the Charleroi-Brussels-South airport, the national hub of the RyanAir company providing low cost air transportation in Europe. A few hotels are available within walking distance from the laboratory. The laboratory can be reached from Brussels with a daily bus shuttle (first shuttle from Brussels leaves at 8:00 and last shuttle leaves Gosselies at 18:00).
The number of participants is strictly limited to 15. Prerequisites for participation are :

- a reasonable knowledge of the UNIX operating system
- some familiarity with bioinformatics tools
- a basic knowledge of molecular biology.

Please send an email with your background and motivation to `vledent@dbm.ulb.ac.be` with the mention «**BSMW05**» in the title.

ASEAN-China in Beijing: Bioinformatics Workshop report

The middle of October is the best season of the year in Beijing, with colourful trees along the hillside of the Great Wall. A delegation of twenty-five people from twelve countries were climbing the Wall and enjoying the beautiful scenery on a Friday morning after four days of hard work tapping the keyboard with their fingers. The twenty students of the ASEAN-China Bioinformatics were from the ten South Asian Nations: Brunei, Indonesia, Cambodia, Laos, Myanmar, Malaysia, Philippines, Singapore, Thailand and Vietnam and the five teachers were from Australia, UK, Singapore and China.

It was not an easy job for either the teachers, or the students, since the goal of the Workshop was very ambitious. The workshop started with installing Linux on a brand new PC box and feeding the various bioinformatics tools integrated in the iBiostation CD-ROM, a package developed by Tim Littlejhon and his colleagues at BioLateral. The keyboard looked a bit sticky to several students since they have never typed a Linux command before. However, the enthusiasm of the students impressed the teachers, and the good teaching experience of the teachers made things clear for the students. On the morning of the second day, all the 24 PC Linux boxes were happily hosting the useful tools for sequence analysis. On the morning of the third day, the AP-BioKnoppix Live-CD was burned by Tan Tin Wee and his Singapore colleague for each student to bring back for installation and testing.

The students seemed a bit relax while Lisa Mullan gave a brief introduction to sequence alignment with her lovely example of GREENAPPLE and REDAPPLE, of course in her sweet English. Peter Rice, one of the bosses of the EMBOSS, was proud of showing off his favourite EMBOSS shirt and talking about the big user community of this package, as well as the plans for further development. Lectures by other invited speakers were also arranged during the Workshop. And finally, a panel discussion was organised led by Tin Wee and the other teachers. A great interest was shown by all the students for further bioinformatics training in their own countries and to promote collaboration among the nations. Indeed, the goal of the Workshop is to act as the starting point of long term collaborations, to build a bridge between east and west, to close the gap between south and north, in the field of bioinformatics research, development and education.

Prof. Jingchu Luo

EMBnet Stichting Annual General Meeting 2004

The 18th EMBnet AGM took place this year in Elewijt, a small place north of Brussels. The facilities at the Elewijt Business Centre were particularly adequate as all accommodations and meeting facilities were available in a single place, situated less than 10 km from the airport. The meeting room infrastructure provided permanent Internet connection (copper and WiFi) including two rooms for email reading.

The business meeting started on September 16th with the usual welcome greetings and the roll call. Twenty-six of the 35 members were present. Sadly, the repeated absence of Greece and the apparent absence of activity resulted in the unanimous decision to drop Greece from the group. The resignation of Denmark was also noted with regret. No new membership applications have been received.

This was the last AGM organised during the period covered by the EMBCORE EU project, and this was the right time to look at the results. It was noted that due to various misunderstandings and changes in the persons in charge of the project, both at the EU Commission and at the University of Nijmegen[1], a significant part of the money available could not be spent as foreseen. Given this fact, it was decided to adjourn the meeting for a few hours and take the opportunity offered by the AGM taking place so close to the European Commission to pay a visit to Mr Fred Marcus, EU commissioner for Life Science. A delegation consisting in Erik Bongcam-Rudloff (chairman), Sandor Pongor (treasurer), Martin Grabner (secretary), Jack Leunissen (EMBCORE coordinator) and Robert Herzog (AGM2004 organiser) went to Brussels and met the commissioner. The outcome was that basically there exists no reasonable possibility of extension of the EMBCORE contract. All expenditures will have to end by November 30 this year.

Under the direction of Jack Leunissen, the overall achievements under EMBCORE were summarised, and it was concluded that, even with the reduced spending, a very significant list of deliverables could be reported and notably:

- Many courses were organised worldwide and notably seven in 2004, covering subject ranging from microarray data management to molecular dynamics, introductions to BioPerl and the Staden package
- Four conferences took place with the support of EMBnet over the last few years
- Several specialised workshops were organised
- Seven Quick Guides were produced and printed, three other should appear before the end of the project (see p10).
- The quarterly newsletter "embnet.news" was published on a regular basis
- The web site of EMBnet was turned into a nice portal to access information about all the activities of the group.

All these achievements are clear proof of the willingness of the members of the EMBnet to collaborate and extend their mutual interactions.

The next item on the agenda of the AGM was the presentation of the financial status of the EMBnet Stichting. The present status is about 75K€ available, i.e. enough to cover more than next year's AGM. The treasurer regretted that several members still had not paid the fees for 2004. Also several fees dues for preceding years are still pending. The total amount of outstanding fees reaches about 20K€, which significantly limits actions for the future.

The reports on several EMBnet-associated activities were given, and notably the EMBER project under the leadership of former EB member Terry Attwood and the EMBOSS and EMBRACE projects. Difficulties with the former were noted due to the foreseen closure during the summer 2005 of the HGMP-RC in Hinxton. The creation by the EMBOSS authors of a spin-off company to further develop and support EMBOSS was noted and the assembly wished them success in this enterprise.

---

[1] The institution in charge of the financial coordination of the project which is in contact with Jack Leunissen, the project coordinator.

After the reports from its various committees (Executive Board, Education and Training, Publications and Public Relations, Technical Management and Research Development), the meeting discussed the opportunities for EMBnet members to share resources in order to offer more and better services to national users and try to free more time to work on emerging scientific subjects. The discussion touched such subjects as

> Shall EMBnet redefine its mission ?
> Shall EMBnet open more nodes, new kind of nodes ?
> How do we regild EMBnet's image ?

Erik Bongcam-Rudloff opened the discussion. J.R Valverde asked for further needs of the nodes, which were not included in the presentations. The board agreed that one aspect in Bioinformatics was neglected over the last years by EMBnet: support for Molecular Modelling should be given attention again.

Finally, Erik Bongcam presented a slide with EMBnet's mission statements:

- Assist in biotechnological and bioinformatics related research
- Provide training and education
- Exploit network infrastructure
- Investigate and develop new technologies
- Bridge between commercial and academic sectors

The round of nominations to the committees for next year produced the following:

EB	Erik Bongcam-Rudloff (SE)	chairman
	Sandor Pongor (ICGEB)	treasurer
	Robert Herzog (BE)	secretary
	Martin Grabner (AT)	member
E&T	Isabel Marques (PT)	secretary
	Lisa Mullan (UK)	treasurer
	Valérie Ledent (BE)	member
	Vassilios Ioannidis (CH)	member
P&PR	Gonçalo Guimarães (BR)	chairman
	Pedro Fernandes (PT)	secretary
	Laurent Falquet (CH)	treasurer
	Oscar Grau (AR)	member
TM&RD	José R. Valverde (ES)	chairman
	Nils-Einar Eriksson (SE)	secretary
	Kristian Vlahovicek (ICGEB)	treasurer
	Endre Barta (HU)	member

Further points discussed were possible changes in the fees, none of which were adopted. The discussion also covered possible cooperation with other organisations, like LAC Bionet, APBioNET and the ISMB. The latter will organise its 2006 conference in Brazil, where our member Gonçalo will be a member of the organising committee. This will be a major opportunity to stress the importance of EMBnet's role. The recommendation by Fred Marcus that members of EMBnet try to participate individually in building large projects, as the perspective of funding the consortium as such presents several problems.

The location of the next AGM was discussed. Offers from India and Slovakia were received. A preliminary voting took place, the result of which was in favour of Slovakia, with an expected cost of around 25K€, pending a decision by the board once some financial issues of the Stichting have been cleared. A final decision should be taken on Nov. 15.

The second day of the AGM was concluded with the official dinner at the nearby Kasteel Diependael.

The day before the AGM itself was used for the so-called "EMBnet Collaborative Workshop", where all partners had the opportunity to present some of their activities for the benefit of the other members of the group.

Among the highlights of this meeting, were reports about several 2004 workshops:

Regulatory sequence motifs discovery workshop (Sweden, summer 2004 – 20 students, 10 countries)

Bioinformatics System management workshop (Belgium, April 2004 – 13 students, 5 countries)

The first SRS Federation workshop (Belgium, Sept. 2004 – 22 participants, 17 nodes) and the creation of the initial SRSfed taskforce which consists of: Erik Bongcam (SE), Thure Etzold (Lion), Lubos Klucar(SK) and David Coornaert (BE).

There were also the following presentations:

- Laurent Falquet (CH) gave a nice presentation of the MyHits data management system for protein motifs and patterns. This is a high performance tool available to all academic users
- Guy Bottu (BE) gave a talk about EMBOSS Wrappers – how to install alien software in EMBOSS in a seamless way
- Marc Colet (BE) gave a presentation with a demo of wEMBOSS, a web interface to EMBOSS, its properties and virtues
- J.R.Valverde (SP) introduced the groupware solutions Bioportal Mytikiwiki, a better way to manage communications within a community
- Martin Grabner (AT) presented eGoupware a very powerful but complex way to handle ACL-system with an integrated CMS, which can be configured in a way that communication handled by the groupware modules can be formatted for a web presence directly. LDAP for access control, and webdav as http-extension makes it very complete. A test installation for this workshop is available from `https://www.at.embnet.org/eGroup/`
- Vassilio Ioannidis (CH) presented the management of the helpdesk of the Swiss Institute of Bioinformatics with Request Tracker
- J.R.Valverde (SP) gave a presentation of the current status of the GRID projects in the EU, and how to reap the benefits it promises to offer
- Thure Etzold (LION Bioscience) presented the EU SIMDAT project
- Marc Colet (BE) presented the Belgian bioinformatics BIGRE project

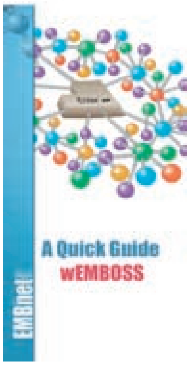
The feeling at the end of this collaborative workshop was that this kind of meeting is very useful and should be pursued, as it stimulates collaborations inside our consortium.

Three new EMBnet QuickGuides available !

We are happy to announce that our famous QuickGuide family has seen the birth of 3 new members.

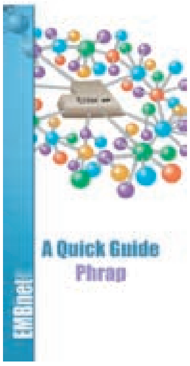
A QuickGuide to wEMBOSS

Designed by Guy Bottu and Robert Herzog of the Belgian EMBnet node, this guide describes the wEMBOSS web interface in detail and provides the user with screenshots to explain the procedures.



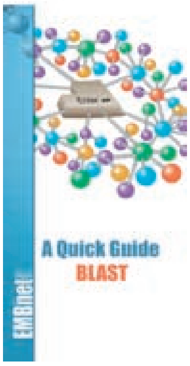
A QuickGuide to Phrap

Designed by Marcos R. Araujo and several members of the Brasilian EMBnet node, this guide describes the Phrap program of the phred/phrap package. Another companion guide is in preparation to describe the Consed software (The Phred QuickGuide was updated to match its Phrap companion guide).



A QuickGuide to NCBI BLAST

Designed by Eduardo Fernandes Formighieri and several members of the Brasilian EMBnet node, this guide describes the NCBI BLAST program and web site.



All these new guides, as well as other guides are freely available from our web site `http://www.embnet.org/download/guides.html`

Report: SRS Federation workshop

Robert Herzog
BEN
ULB Campus de la Plaine,
blv. du Triomphe,
1050 Brussels, Belgium

SRS is a well known and appreciated tool that offers the average biologist a fast and easy way to search biological databanks. Initially conceived at the EMBL laboratories in Heidelberg by Thure Etzold to access biological sequences, this system was soon adopted by many bioinformatics access providers and notably by most EMBnet nodes. However, due to the fast growth of the nucleic acid databanks, many providers dropped out, relying on the access to other partner's SRS servers or to non–European data providers. On the other hand, the efficiency of SRS was recognised by most bioinformatics software developers, and notably the EMBOSS team. Thus having SRS servers around, with all databanks daily updated remains an attractive option, sadly not within the reach of all. Indeed, today the mere size of the most important databanks requires significant disk storage and computer time to build the indices that are needed to answer user queries.

Due to the notable increase in network efficiency, the option of building a kind of federation of SRS servers emerged over the years : each partner would take care of managing in an optimal way sections of the data and provide access for the members of the consortium. In this way, EMBnet members would share resources and at the same time more data would be available within the consortium than on any individual site[1]. Building in the system a reasonable level of redundancy and a clever dispatcher would allow for an unprecedented level of quality, virtually impossible to reach by any isolated system.
Finally, one might start dreaming of a system where local data produced by a group of experimentalists would be seamlessly merged in the consortium and provide efficient links between local unpublished data and those available in the entire currently updated public data sources.

All these ideas prompted several EMBnet node delegates to attend a two-day workshop last September, a meeting linked to the yearly Annual General Meeting of EMBnet. About 20 people originating from 12 different countries attended the workshop, with the special attendance of Thure Etzold himself, accompanied by Peter Maccallum and Dan Staines from the LION Bioscience company[2]. The main issues about building such a system were discussed at length, at times in full sessions, and at times in break-out groups which concentrated on specific issues. The central subjects of the breakout sessions were:

- Contents of databases and tools hosted by the consortium
- Training and support
- How to organise collaboration among members
- QoS, redundancy, infrastructure and technologies
- Branding and communication
- Definition of consortium rules
- Structure and roles
- How to attract funding
- How to maintain national mandates

The final decision of the workshop was to start building a first prototype of an SRS Federation right away as an experiment. Three EMBnet groups came forward and offered to start the project on a purely voluntary basis, provided LION would help with the recent developments of the software (SRS 8, Prisma and WSObjects). A decision was taken to include EMBL, UNIPROT, KEGG, OMIM, INTERPRO and PDB in the initial dataset. Each of these databanks would be kept up-to-date in the best possible way.

---

[1] "L'Union fait la Force" is the basic motto of the Belgian state…

[2] The present owner and developer of SRS

Figure 1. Meeting in Cambridge.
From Left to right, sitting: Marcello Falsarella (BR), Robert Herzog (BE), David Coonaert (BE); standing: Pawel Siedlecki (PL), Volodya Yankovskiy (SE), Erik Bongcam (SE), Lubos Klucar (SK) and Emiliano Barreto (CO)

The Slovak, Belgian and Swedish nodes entered the race around October the 15th, when the special licence agreements with LION had been settled. This resulted in a very complete installation under Linux Slackware in Slovakia by Lubos Klucar, six weeks later. This installation included Prisma to keep an eye on daily updates. The second team to reach the goal was Belgium (mainly David Coornaert) with an installation on specially acquired hardware built around an AMD Opteron and a 2.4 Terabyte file server running Linux SuSE. Here, the databanks were still kept up to date with home-brewed Perl scripts. Sweden followed with a purposely build set of Dual-Opteron machines with the capacity for large data storage. Weekly meetings using the Marratech software helped the partners to exchange their experience with the support of Dan Staines from LION. These easy and cheap to attend meetings helped considerably with the progress of the project.

This was the status of the group when the second SRS federation workshop took place in Cambridge on November 22 to 24 (Figure 1). During those days, a complete hands-on training program was delivered by Dan Staines on Prisma and Peter Maccallum on WSObjects (an SRS specific web service). The third day was devoted to some brainstorming with more members of the SRS development team about future actions within the SRS Federation. Six EMBnet nodes were present at this meeting as delegates from Poland, Columbia and Brazil had expressed the wish to join. It was decided to add PROSITE to the basic datasets and to provide CLUSTALW and PROSITESEARCH as tools installed under SRS. An SRS-connected BLAST server is among the next goals.

The milestones for the group are
- by the end of December 2004, all groups should have a working and permanently up-to-date machine with all initial datasets
- during the next two months, work will begin on building satellite SRS servers receiving from the other partners ready-to-build tarballs containing datasets and the corresponding pre-build indices
- a first prototype of a distributed query interface using a WSObjects approach should become available for experimentation, by the end of February 2005

What will happen next remains to be seen, but in principle, if the experiments proceed as expected, the visibility of the project would be opened to the whole EMBnet community. More EMBnet members could start joining the consortium, provided each one takes the responsibility of a few additional datasets. The concept of redundancy would progressively be implemented in order to provide an increased level of guaranteed quality. This would increase the overall system robustness and prepare it to face hardware problems and network outages. At the time of writing this report no precise date for this phase of the project can be proposed…

Whether the SRS Federation project will fly rests totally on the good will and the dedication of the initial partners, but if it succeeds, the project might well be the start of a new way to consider the management and the access to biological data sources.

wEMBOSS, a featured Web interface to EMBOSS

Guy Bottu
BEN
ULB Campus de la Plaine,
blv. du Triomphe,
1050 Brussels, Belgium

Some background

As the regular readers of EMBnet.news know only too well, a few years ago the EMBnet nodes were compelled to abandon GCG in favor of EMBOSS as their main sequence analysis package. In the course of 2002, the last year for which the Belgian EMBnet Node had a (very expensive !) GCG license, Marc Colet (professor at the U.L.B. and manager of BEN) devoted himself to the development of wEMBOSS. He fortunately did not have to start from scratch, since he could recover a lot of code from his earlier project WWW2GCG, a Web interface for GCG. While Web interfaces tend to be featureless (they keep no record of what the user did before) or otherwise store results of previous program runs in some temporary or custom area, he conceived the idea of storing results in (and taking input data from) a classic UNIX home directory, so that users can work on the same data over the Web as well as in a terminal session. Thus, a user can e.g. in an X-terminal use the Staden package to manage a DNA sequencing project and then use the Web interface to analyze his consensus sequence using GCG or EMBOSS.

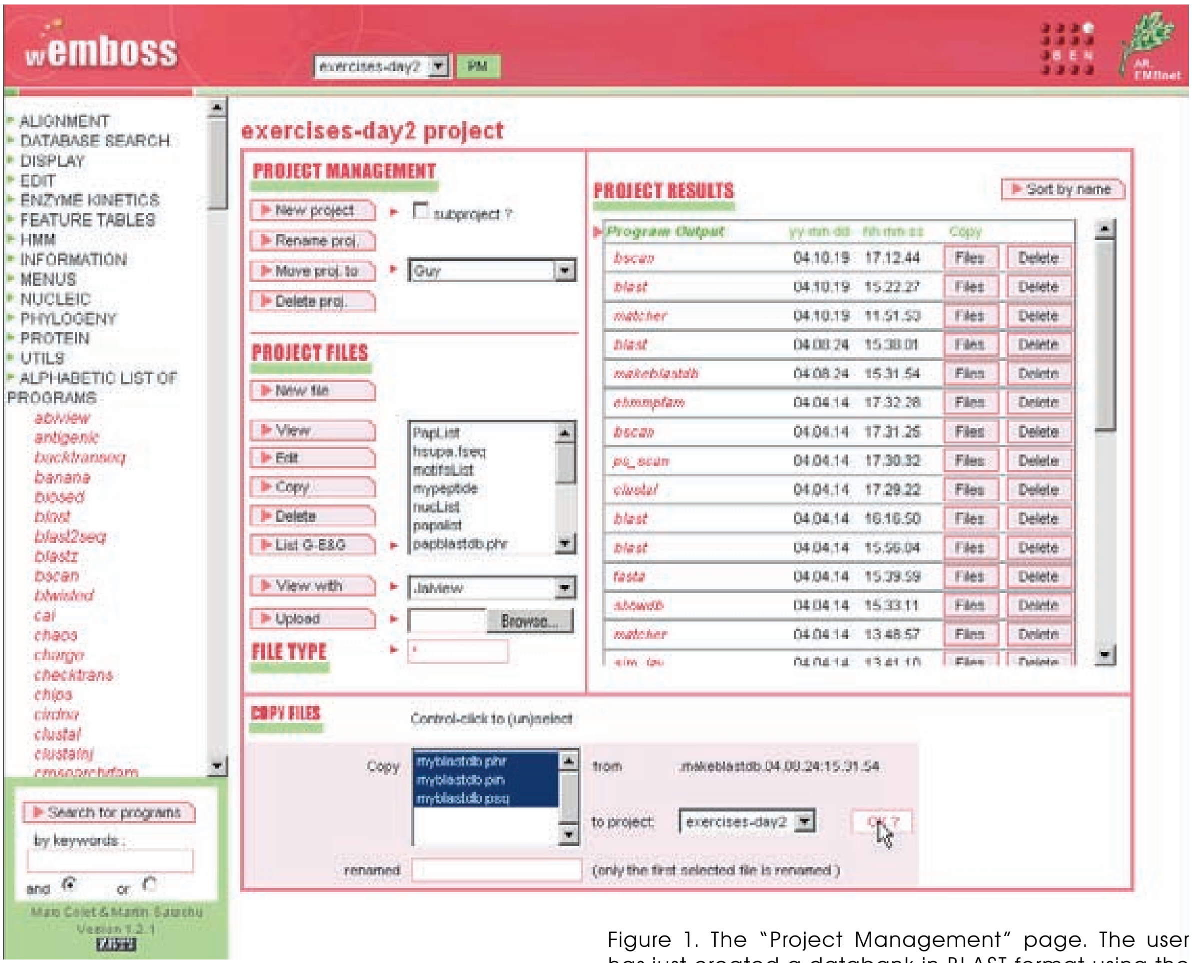
There remained the problem of generating HTML pages with boxes and selectors from the ACD files (each EMBOSS program has an ACD file that defines the input data and other information needed by the program in a precise syntax). For this purpose, code was borrowed from Luke McCarthy's "EMBOSS GUI" (many thanks, Luke !). Note by the way that wEMBOSS parses the ACD files on-the-fly, so that changes in the EMBOSS installation are immediately integrated, without needing to re-run some installation script. JavaScript was added to the pages to make them "alive", letting elements appear/disappear and default values change, so that when the user moves from top to bottom through the page while making his selections, he will be prevented from making inappropriate choices. While developers of GUI's sometimes complain about the complexity of ACD syntax, M. Colet made it a point of honor to make sure that wEMBOSS handles all the formulas correctly.

At the end of 2003, a Web design company was hired to give the interface a look more in agreement with modern standards. In the same period, Martin Sarachu from the Argentinian EMBnet node joined the wEMBOSS development team. A Web site `http://www.wemboss.org` was set up and in May 2004 the first non-field test release was made available for public downloading.

The project management (PM)

To free users of the need to navigate up and down in a classic file system and to make sure that their data are stored tidily and separated from anything not related to their sequence analysis work, wEMBOSS favors the notion of "projects" (which are still UNIX directories, but are presented is a special way). "projects" can be nested to any depth and a selector allows direct navigation from one project to another. Furthermore, it is possible to select not only a data file located in the current working project, but also data in any project higher up in the hierarchy, as input for a program.

When a user starts a wEMBOSS session, the first thing he sees is the "Project Management" page (Figure 1), which contains a series of buttons to create a new project or delete, rename or move an existing project. Data files can be deleted or copied, created or modified using a text area, uploaded from the file system of the local computer or simply viewed.



exercises-day2 project

Move proj. to: Guy

PapList
hsupa.fseq
motifsList
mypeptide
nucList
papolist
papblastdb.phr

View with: Jalview

Program Output	yy mm dd	hh mm ss	Copy	
bscan	04.10.19	17.12.44	Files	Delete
blast	04.10.19	15.22.27	Files	Delete
matcher	04.10.19	11.51.53	Files	Delete
blast	04.08.24	15.38.01	Files	Delete
makeblastdb	04.08.24	15.31.54	Files	Delete
ehmmpfam	04.04.14	17.32.28	Files	Delete
bscan	04.04.14	17.31.25	Files	Delete
ps_scan	04.04.14	17.30.32	Files	Delete
clustal	04.04.14	17.29.22	Files	Delete
blast	04.04.14	16.16.50	Files	Delete
blast	04.04.14	15.56.04	Files	Delete
fasta	04.04.14	15.39.59	Files	Delete
showdb	04.04.14	15.33.11	Files	Delete
matcher	04.04.14	13.48.57	Files	Delete

Copy: myblastdb.phr, myblastdb.pin, myblastdb.psq

from: makeblastdb.04.08.24:15.31.54

to project: exercises-day2

renamed (only the first selected file is renamed)

Figure 1. The "Project Management" page. The user has just created a databank in BLAST format using the program makeblastdb and is going to copy it to the project in order to use it (for BLAST under wEMBOSS, see article about WRAPPERS4EMBOSS in the next issue of EMBnet.news).

Running programs & managing outputs

The user can select a program from an expandable/compressible list. If he cannot readily find the program he needs from the alphabetic list or from the EMBOSS program groups, there is the "Search for program by keywords" function, which searches by relevance first in the one-line description and then in the complete on-line manual. The user can start a program (Figure 2) and wait for the result. A new set of windows will pop up one below the other, with the content of the different output files generated by the program.

If the user wants to go back to see an old result, there is no need to run the same job twice or to save the output locally as it can be called back from a menu at the right side of the PM page. If a user wants to give the output of a program as input to another program (e.g. give a profile made by prophecy to profit or prophet), he can use the "Files" button to copy that file to the project. This introduces an extra step but the developers think that it is a small price to pay for the overall streamlining.

If a program is likely to take much time, there is no need to wait in front of the computer screen and possibly bump on the time-out of the Web; the user can type in his E-mail address in the box at the bottom of the program page, the program will run "in the background" and the user will be notified by E-mail when it has finished. The user can at any time start a new wEMBOSS session and access the program output from the PM page or even, if his Mail program is HTML compliant, open it with a simple click on the link in the E-mail message (Figure 3).

Figure 2. Example of a program page: the dotmatcher.

Handling sequences, graphics and other data

The sequence(s) needed by a program can be provided by typing in an EMBOSS USA (like `swissprot:papa_carpa`, `my.seq` or `list::my.list`), with "public" sequences available on the server or "private" sequences or "Lists" the user must first put them in a file. It is also possible from a program page to upload a sequence file from the user's local computer. If a user wants to run several jobs on the same sequence(s), wEMBOSS has a device that overcomes the need to type or upload these over and over again : when a new "project" is created it already contains files called nucList and protList. It is a small effort to edit these files and type in the USA's, which will appear as if by magic in a selector in the program page.

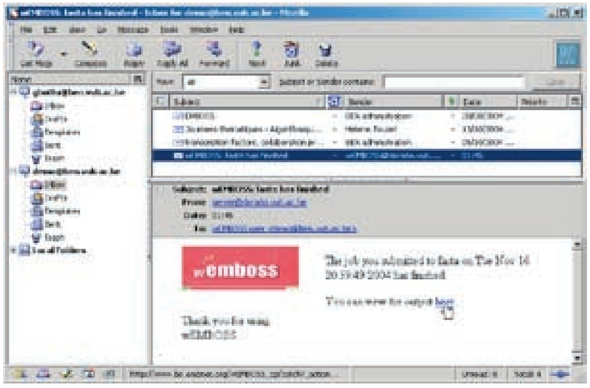


Figure 3. The user has received by E-mail a confirmation that his "background" job has terminated. He intends to view the result by clicking on the link.

If a program needs data like symbol comparison tables or codon usage tables, they can be selected from the EMBOSS distribution, selected from the project files or directly uploaded from the local computer.

As EMBOSS users know, some programs produce a picture as output and it is possible to choose among a series of different graphical formats. wEMBOSS by default proposes PNG, which is displayed directly in the Web page. It proposes PostScript as an alternative, which is not displayed, but can be downloaded.

Plug-ins and applets

Some kinds of output, like multiple sequence alignments or phylogenetic trees, can best be visualized using an appropriate viewer rather than looked at as simple text. This aspect is still under development. When a file is viewed from the PM page or saved from the program output page, it can be sent with a MIME label. The user can configure his local system or Web browser so that the file is opened with an appropriate plug-in or "helper" application. There are currently labels for MSF's, CLUSTAL trees and files in LAV format (produced by SIM for LALNVIEW). A recent addition is the possibility to copy an output file to the "project" and open it with an applet. Currently, the Sanger's Centre Jalview for multiple sequence alignments and Christian Zmasek's ATV for phylogenetic trees are distributed together with wEMBOSS.

References

- Clamp M., Cuff J., Searle S. M. and Barton G. J. (2004).
  «The Jalview Java Alignment Editor»
  Bioinformatics, 12, 426-7.
- Sarachu M. and Colet M. (2004).
  «wEMBOSS: a web interface for EMBOSS»
  Bioinformatics (in press).
- Zmasek C.M. and Eddy S.R. (2001).
  «ATV: display and manipulation of annotated phylogenetic trees»
  Bioinformatics, 17, 383-384.

See also:
The EMBnet QuickGuide to wEMBOSS
`http://www.embnet.org/download/guides.html`

The EBI resources in a Nutshell (part 1)

Lisa Mullan
EMBL Outstation - Hinxton,
European Bioinformatics
Institute, Wellcome
Trust Genome Campus,
Hinxton,
Cambridge, CB10 1SD,
United Kingdom

This practical was designed for the EBI's Small to Medium Sized Industry Programme and aims to offer a practical overview of some of what the EBI can offer in terms of bioinformatics resources in the form of web services.

We will start with a sample gene and investigate the use of various tools to analyse splice sites and SNPs, proteomic information, including interactions with other molecules and finally its tertiary structure. The sequence chosen does not have unlimited information, thus some of the entries may be a little sparse. If this is the case, feel free to experiment with your own data.

Alternative Splicing Database

Access the Alternative Splicing Database at `http://www.ebi.ac.uk/asd/`.

Follow the link on the left hand side to **AEdb** and select the "**simple query form**". Select the correct species from the relevant scroll menu. Use the gene name as a second search term by checking the "Cross-reference or gene symbol" box. Select "Gene Symbol" from the pull down menu and type in the gene name. Hit the Submit button.

☐ Show all the entries.

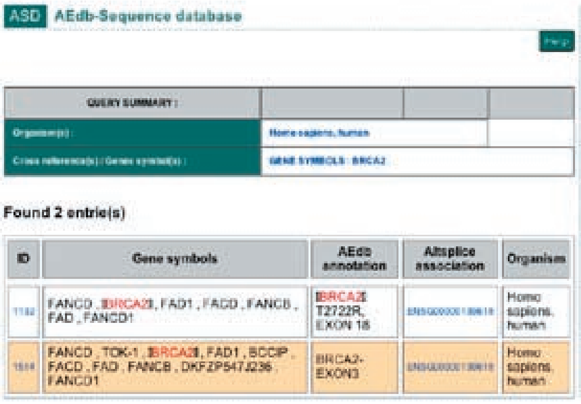
☑ Organism(s) :

☑ Cross reference(s) or gene symbol(s) (e.g. PKM2, ALBZA ALB, S57180 8455946) :

☐ Keywords (you can use wildcard(*) and separate multiple values by & | (e.g. "albumin", G-PROTEIN | albumin, albumin*|*integrin*, ALBZA) :

The Alternative Splicing Database (ASD) Project aims to understand the mechanism of alternative splicing on a genome-wide scale by creating a database of alternatively spliced exons from human, and other model species. It is currently composed of three separate databases. **AltExtron** and **Altsplice** are computational methods for extracting information on alternative splicing. **AEdb** (Alternative Exon Database) is a curated database, with information on alternative splicing extracted from the literature.

Data from all these databases is free to download and store in-house.



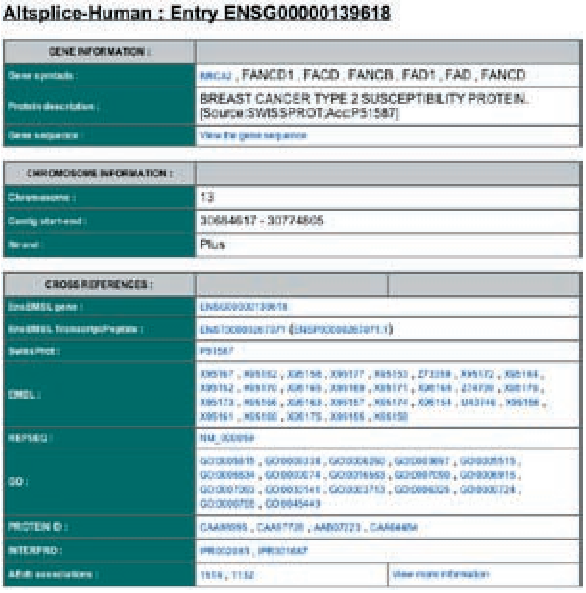
Found 2 entrie(s)

ID	Gene symbols	AEdb annotation	Altsplice association	Organism
1132	FANCD , BRCA2 , FAD1 , FACD , FANCB , FAD , FANCD1	BRCA2 T2722R, EXON 18	[illegible]	Homo sapiens, human
1514	FANCD , TOK-1 , BRCA2 , FAD1 , BCCIP , FACD , FAD , FANCB , DKFZP547J236 , FANCD1	BRCA2-EXON3	[illegible]	Homo sapiens, human

The AEdb returns two splice entries related to the BRCA2 gene, one in Exon 3 and the other in Exon 18.

Follow the link created by the AEdb ID for the splice information on exon18 to see information on the submission and splice type of the BRCA2 gene.

Return to the Gene Information Data and follow the link to the AltSplice entry (under the "**Altsplice Association**" header) for **AEdb 1132**.



Altsplice-Human : Entry ENSG00000139618

GENE INFORMATION :	
Gene symbols :	BRCA2 , FANCD1 , FACD , FANCB , FAD1 , FAD , FANCD
Protein description :	BREAST CANCER TYPE 2 SUSCEPTIBILITY PROTEIN. [Source:SWISSPROT;Acc:P51587]
Gene sequence :	View the gene sequence

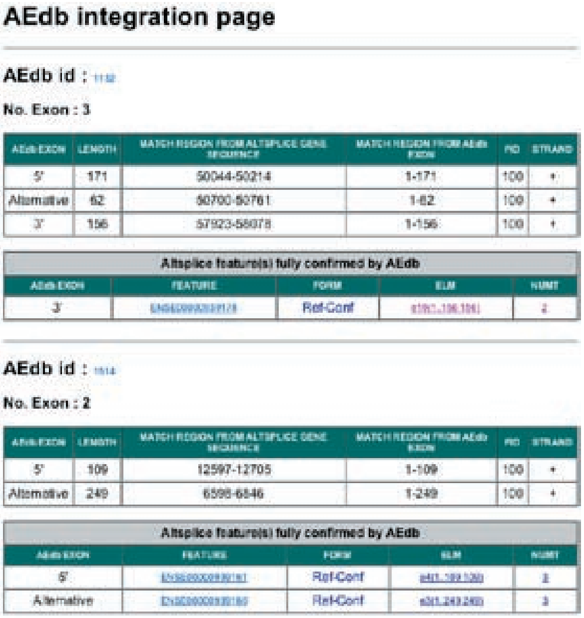
CHROMOSOME INFORMATION :	
Chromosome :	13
Contig start-end :	30684617 - 30774805
Strand :	Plus

CROSS REFERENCES :		
EnsEMBL gene :	[illegible]	
EnsEMBL transcript/Peptide :	[illegible]	
SwissProt :	P51587	
EMBL :	[illegible]	
REFSEQ :	[illegible]	
GO :	[illegible]	
PROTEIN ID :	[illegible]	
INTERPRO :	[illegible]	
AEdb associations :	1514 , 1132	view more information

AltSplice implements a computational pipeline (for detailed detection & characterisation of splice variants) to production standards. Data is taken from the Ensembl database. The AltSplice entry lists features relevant to this.

Scroll down to the cross references and follow the "**view more information**" link relative to the AEdb associations.

This links to the AEdb integration page and information on alternative splicing as matched to the AEdb. The number of exons involved in the feature is given at the top of the page, followed by the AEdb exons and their respective lengths. The match regions are shown as relative positions in the AltSplice (per contig) and AEdb (per exon). The percentage identity (PID) of the match is shown, along with the strand on which the features occur.



AEdb integration page

AEdb id : 1132

No. Exon : 3

AEdb EXON	LENGTH	MATCH REGION FROM ALTSPLICE GENE SEQUENCE	MATCH REGION FROM AEdb EXON	PID	STRAND
5'	171	50044-50214	1-171	100	+
Alternative	62	50700-50761	1-62	100	+
3'	156	57923-58078	1-156	100	+

Altsplice feature(s) fully confirmed by AEdb				
AEdb EXON	FEATURE	FORM	ELM	NUMT
3'	[illegible]	Ref-Conf	[illegible]	2

AEdb id : 1514

No. Exon : 2

AEdb EXON	LENGTH	MATCH REGION FROM ALTSPLICE GENE SEQUENCE	MATCH REGION FROM AEdb EXON	PID	STRAND
5'	109	12597-12705	1-109	100	+
Alternative	249	6598-6846	1-249	100	+

Altsplice feature(s) fully confirmed by AEdb				
AEdb EXON	FEATURE	FORM	ELM	NUMT
5'	[illegible]	Ref-Conf	[illegible]	[illegible]
Alternative	[illegible]	Ref-Conf	[illegible]	[illegible]

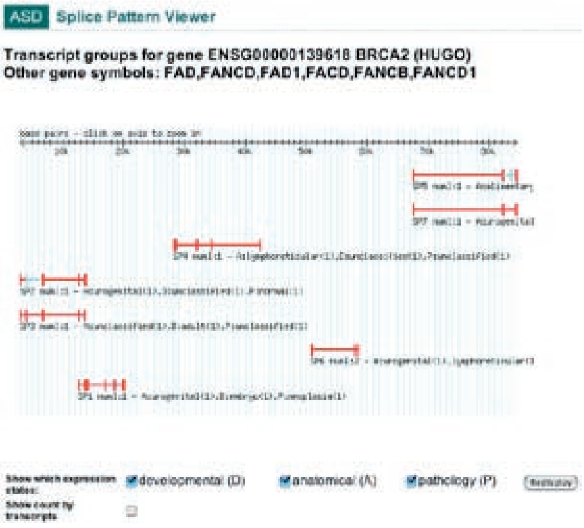
The fully confirmed features are noted in a separate table, with the Feature information linking directly to the relevant Ensembl entry. The "**Form**" column indicates the status of the feature. *Ref-conf* suggests that the reference feature has been confirmed, often with database transcripts. *Ref-PConf* suggests that the reference feature has only been partially confirmed. Indicators in **blue** have been fully confirmed by transcripts, whilst those in **red** have a confirmation transcript unable to confirm all boundaries of the feature. The "**ELM**" column links to information on the sequence of the feature, its start and endpoint and the type of alternate splice site it refers to. The "**numt**" column offers information on the transcripts where this feature has been observed.

The splice information states that the sequence and splice pattern has been confirmed and is located on Ensembl exon 19. The "numt" column indicates that two transcripts confirm this prediction.

Return to the Altsplice entry.

Below the table, a graphical display shows the 7 splice patterns currently annotated along the length of the gene.

Follow the link to "**Click here for additional functionalities**". Select all three expression states and redisplay. The stages where various regions have been experimentally confirmed will appear underneath the splice patterns.



Transcript groups for gene ENSG00000139618 BRCA2 (HUGO)
Other gene symbols: FAD,FANCD,FAD1,FACD,FANCB,FANCD1

Show which expression states: ☑developmental (D) ☑anatomical (A) ☑pathology (P) Redisplay
Show count by transcripts ☐

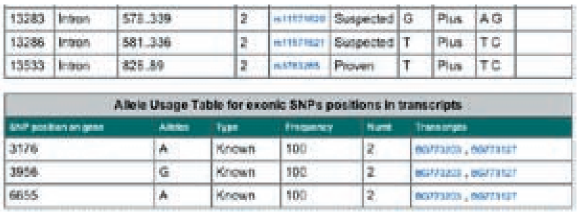
The reference exons are marked in **red** and the reference introns as a connecting **red line**. All insertion exons and introns are in **light blue**. Not all the splice patterns have yet been created, thus some are missing and there are currently only 18 reference exons on view. This compares to the 28 detailed in the Ensembl transcript of this gene.

Zoom in by clicking the sequence ruler at 10K to look at splice patterns SP2 and SP3 in greater detail. The display makes it look as though there is little difference between these two patterns around the start, but the clue is in the colour difference between the second intron.

Return to the AltSplice entry and scroll down to the "**Splice Pattern Table**".

The "**Structure**" column details the exon coverage. The second exon in splice pattern 3 is just 4 bases longer than that in SP2 – thus creating a new intron which shows a 4 base reduction in size. Base counts are taken from the Ensembl genomic sequence size (84.19kb).

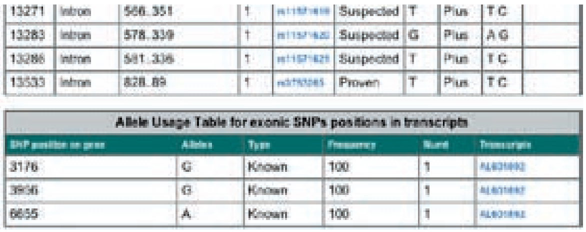
Follow the link to the SNP information (**58**) for splice pattern 2. At the bottom of the page, in the Allele Usage Table, there is information on three alleles of this splice pattern. Note the allele usage and then note the same information for splice pattern 3. Two of them (Gene Positions 3956 and 6655) use an identical allele, but the SNP at gene position 3176 uses an A in SP2 and a G in SP3 – a possible cause of the isoform.



13283	Intron	578..339	2	[illegible]	Suspected	G	Plus	A G	
13286	Intron	581..336	2	[illegible]	Suspected	T	Plus	T C	
13533	Intron	826..89	2	[illegible]	Proven	T	Plus	T C	

Allele Usage Table for exonic SNPs positions in transcripts

SNP position on gene	Alleles	Type	Frequency	[illegible]	Transcripts
3176	A	Known	100	2	[illegible]
3956	G	Known	100	2	[illegible]
6655	A	Known	100	2	[illegible]

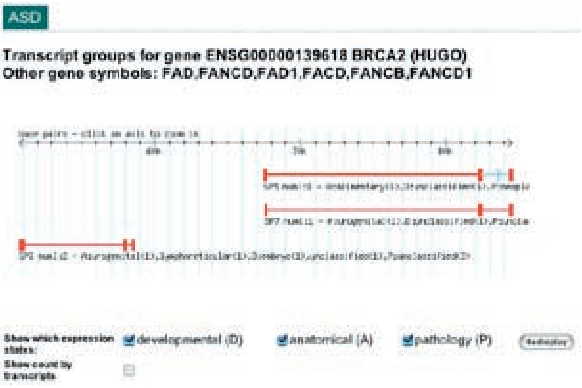


13271	Intron	566..351	1	[illegible]	Suspected	T	Plus	T C	
13283	Intron	578..339	1	[illegible]	Suspected	G	Plus	A G	
13286	Intron	581..336	1	[illegible]	Suspected	T	Plus	T C	
13533	Intron	828..89	1	[illegible]	Proven	T	Plus	T C	

Allele Usage Table for exonic SNPs positions in transcripts

SNP position on gene	Alleles	Type	Frequency	[illegible]	Transcripts
3176	G	Known	100	1	[illegible]
3956	G	Known	100	1	[illegible]
6655	A	Known	100	1	[illegible]

Return to the zoomed out view of the splice pattern viewer and zoom in once again using the browser back button. This time zoom in at 80K to investigate splice patterns SP5 and SP7. There is an additional exon in SP5 caused by a mutation in the intron.



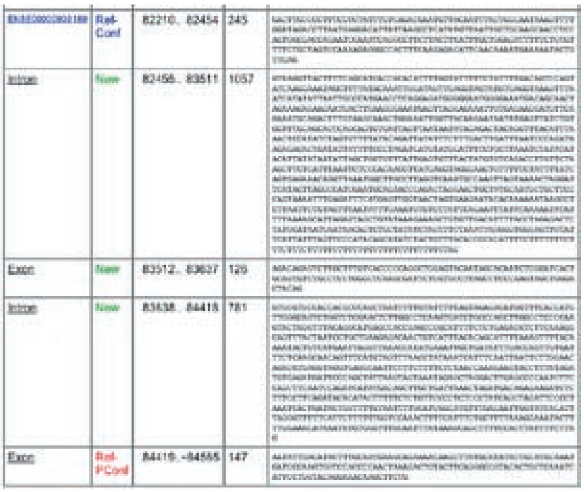
Transcript groups for gene ENSG00000139618 BRCA2 (HUGO)
Other gene symbols: FAD,FANCD,FAD1,FACD,FANCB,FANCD1

Show which expression states: ☑developmental (D) ☑anatomical (A) ☑pathology (P) Redisplay
Show count by transcripts ☐

The Splice Pattern data table in the AltSplice entry information details the sizes of the new exons.

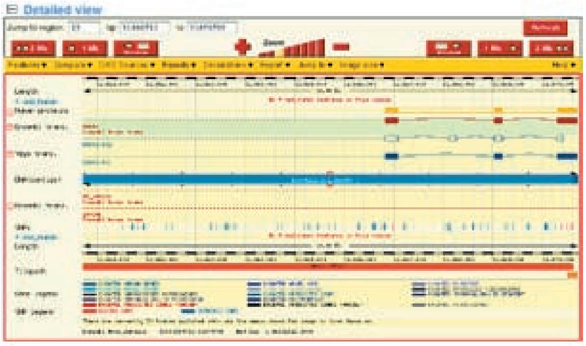
Follow the numbered link from Splice Pattern 5 and open the table detailing the splice information for SP5.

The final reference exon before the new one has a length of 245 bp. From the DAS sources we know the exon is towards the end of the gene, and we can thus identify Ensembl exon 25 as the exon which would be involved in the alternative splice pattern seen in SP5 and SP7.



[illegible]	Ref-Conf	82210..82454	245
Intron	New	82455..83511	1057
Exon	New	83512..83637	126
Intron	New	83638..84418	781
Exon	Ref-PCont	84419..84565	147

Go to `http://www.ensembl.org` and lookup the Human gene BRCA2. Follow the contig link from the GeneView page to the ContigView page and "Jump to Region" 31864711 - 31874709 to zoom in on this splice pattern and the Ensembl and Vega transcripts.



Use the Features ▼ menu to remove all features except Ensembl transcripts; Vega Transcripts, Human proteins and SNPs. Use the DAS Sources ▼ to add Refseq information. The additional exon which is the feature of SP5 can also be seen as a Vega exon, however the unfilled boxes indicate that only one transcript has been found.

The Vertebrate Genome Annotation Database (VEGA) is an initiative of the Sanger Institute and several other collaborators and strives to curate the information available in the automated Ensembl database. The database contains finished sequences only.

This region corresponds to exons 24-28 of the Ensembl transcript and the additional exon appears between Ensembl reference exons 25 and 26.

Look at the SNP information underneath this region and a coding SNP is visible in the display underneath exon 25 (red mark). This corresponds to an SNP detailed in the Alternative Splicing Database for both SP5 and SP7.

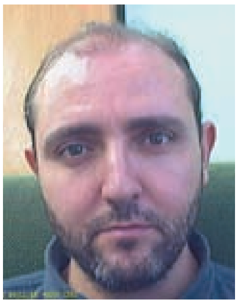
As this SNP is displayed for both splice patterns, it is unlikely that this is the cause of the alternative splicing seen in SP5.

Place your mouse over the coding SNP and follow the link to the SNP properties in the pop up menu.

Follow the link in the resulting SNPView to **HGbase**. This database links to dbSNP as well as defining the exact position of the SNP and detailing the up and downstream sequences. In addition, MeSH terms relating to the SNP are listed.

The Human Genome Variation Database, (http://hgvbase.cgb.ki.se/) is the product of a collaboration between the Karolinska Institute in Sweden and the European Bioinformatics Institute. SNP and other sequence variation information – such as indels and tandem repeats - is manually curated before being stored in this repository. The complete HGVbase data is available in XML, FASTA, MySQL and flatfile format.

Project Management (part 4): A comparison of the tools



José R. Valverde
EMBnet/CNB,
Centro Nacional de
Biotecnología, CSIC
Campus Univ. Autónoma
Cantoblanco, Madrid
28049, Spain

Project Management is a well proven methodology that has been traditionally used in most disciplines. As the relevance of science in society grows, so do the demands put on scientific projects, paralleling those put on other disciplines, often leading to excruciatingly detailed demands by funding agencies. These demands seem to impose an excessive level of housekeeping burden on daily work. And they often certainly do. How are we expected to cope with so much paperwork and still do science?

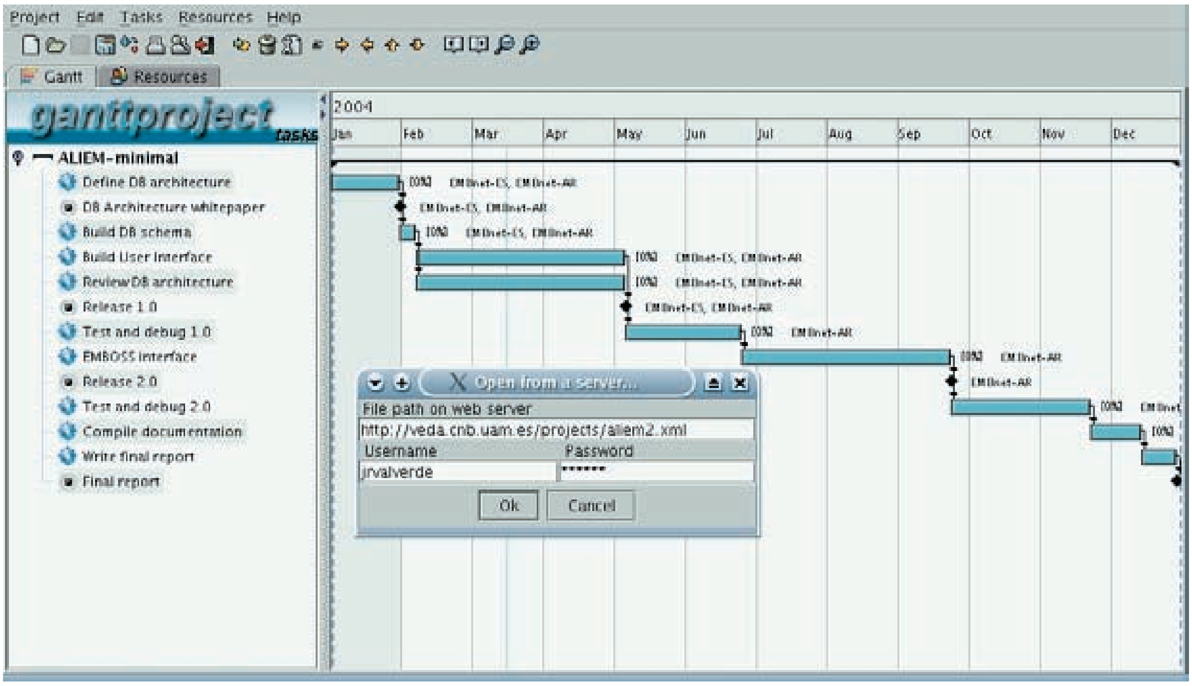
The secret to achieve compliance with these demands is all in the tools.

In this article we use our experience with a wide number of tools to give a broad overview of the most significant and how you may benefit from them. Note that we have installed, and used all of the tools mentioned (and even more). We will concentrate on free tools for two reasons: often science budgets are tighter than their commercial counterparts, and secondly, many of them are reaching a level where they suffice for the needs of the average scientific project.

GanttProject

This is a Java-based tool. As such it will run on any platform with very little effort. You will certainly need to install a Java environment if you don't already have one, but that's all. You can get GanttProject from `http://ganttproject.sourceforge.net/`.

GanttProject offers basic functionality for **planning**, **scheduling** and **tracking** a project, dividing work in tasks, defining task dependencies, assigning resources (people) to tasks, defining milestones and tracking the evolution of task completion. Using it is fairly simple, with a graphical user interface based on a Gantt chart and a resources views which you may manipulate to a great extent,



File path on web server
http://veda.cnb.uam.es/projects/aliem2.xml
Username: jrvalverde
Password: ******

Figure 1. GanttProject example window.

and supported by complete online help. You can generate **reports** for your plan and the progress of your project in HTML and PDF formats as well as a collection of PNG or JPEG images, and of course you can save your planning for future use and modification.

A really interesting feature of GanttProject is that you may save your project over the network to a web server supporting **WevDAV**. This enables you to effectively share the planning with other people and coordinate modifications and access to the plan.

GanttProject is a very nice tool for planning a project and tracking task advances and completion. Its support for task dependencies and dynamic modifications is excellent and its report generation facilities are powerful. It is simple and really easy to use. *GanttProject is probably your best bet for projects that do not require too much overhead and detailed monitoring and where you want to share the Project Management workload with other project members using a centralized coordination repository.*

jxProject

This is another Java-based tool that is supported by advertising (see `http://www.jxproject.com/`). It provides for more detail than GanttProject but has limited reporting and sharing capabilities.

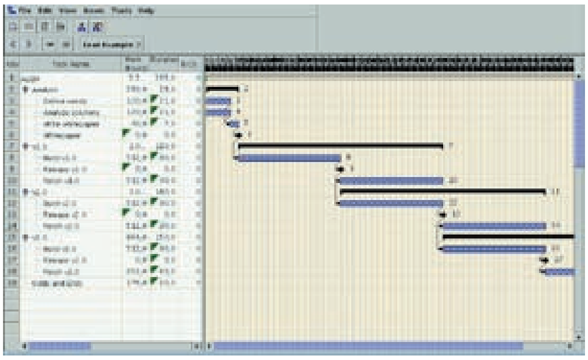
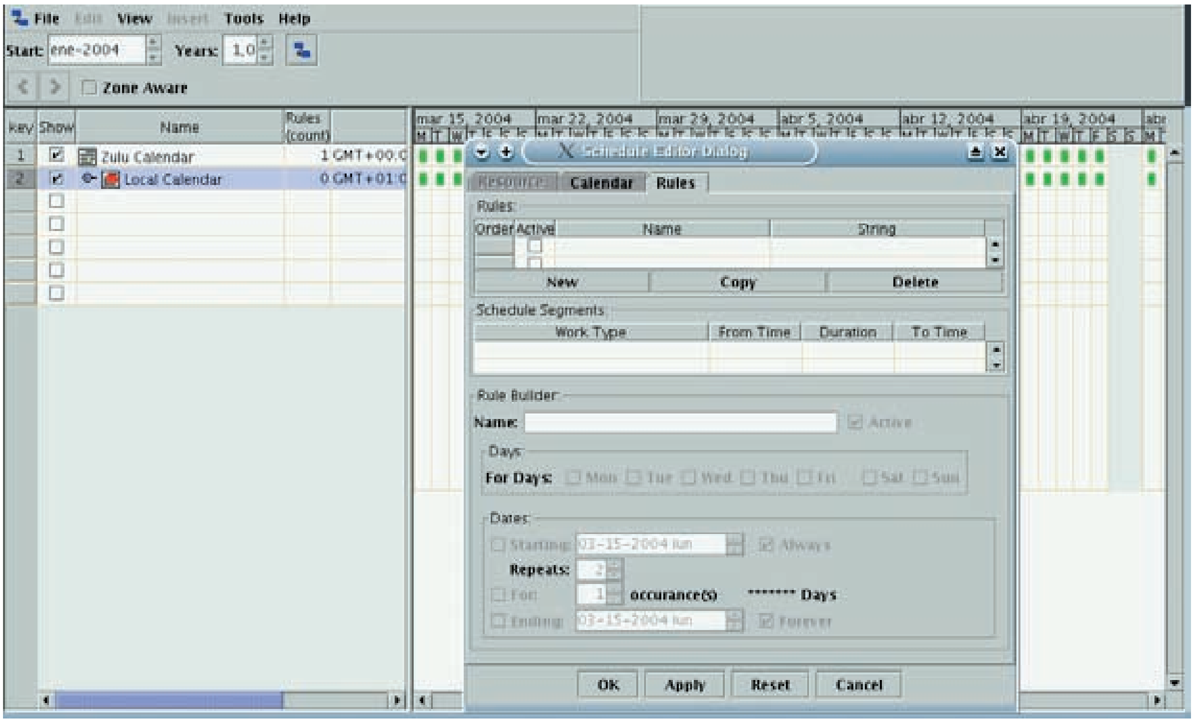


Figure 2. Gantt diagram in jxProject.

JxProject supports task division, dependencies, and resource allocation just like GanttProject does, but it adds some extra tools to better fine tune your organization: in addition to more complete task-based (Gantt diagram) and resource-based views,



Start: ene-2004 Years: 1,0

Key	Show	Name	Rules (count)	
1	☑	Zulu Calendar	1	GMT+00:0
2	☑	Local Calendar	0	GMT+01:0

Rules: Order Active Name String

Schedule Segments: Work Type From Time Duration To Time

Rule Builder: Name: ☑ Active

Days: For Days: ☐ Mon ☐ Tue ☐ Wed ☐ Thu ☐ Fri ☐ Sat ☐ Sun

Dates: ☐ Starting: 03-15-2004 lun ☑ Always

Repeats: 2

☐ For: 1 occurance(s) ******* Days

☐ Ending: 03-15-2004 lun ☑ Forever

Figure 3. Resource window in jxProject.

you get a **calendar definition** system. This tool allows you to define working and non-working days as well as holidays and define different calendars for different people. Since these dates vary widely from one country to another, you will be able to model your project work plan more realistically, accommodating international and personal variations in the availability of the participants in your project.

Finally, jxProject comes with a **resource leveling** engine which can reschedule your project and solve scheduling conflicts according to the task/resource dependencies you have defined. Just as with GanttProject, you may track your project entering information about task progress and completion and make modifications to the plan but, in addition, you can dynamically use the built-in resource leveling solver to detect and solve conflicts and reschedule work as needed.

As you can see, you can achieve finer granularity with jxProject than with GanttProject, but there is a minus side too: jxProject support for reporting is reduced to printing the different views, and it has no support yet for easily sharing your project plans with other people (other than by sharing the project file).

You are better off using jxProject if you need finer granularity beyond GanttProject, want to take into consideration personal and international differences, want to benefit from its resource leveling facility and don't mind doing all of the management work yourself.

Other standalone and Java-based tools

There are other similar tools that you may use, and you are certainly encouraged to give them a try. However, in our opinion, most of them are either too system-specific (i.e. they only run on a given operating system like Linux Planner `http://planner.imendio.org/`), provide more limited functionality (e.g. FUTURe `http://www.wolfkeeper.uklinux.net/FUTURe/`) or are intended as partial tools to be integrated in a more complete framework (e.g. eGantt `http://egantt.com/`). Some of these tools hold great promise in the long run, like Linux Planner (formerly MrProject) and are included by default on most Gnome Linux distributions, but offer little advantages above those described above.

Web based collaboration tools

This is a different area that includes by far the most powerful and useful tools. They are also more demanding to install and require additional effort to master in full, but you get a richer environment with excellent collaborative capabilities.

The basic idea is that all project information will be maintained on a server under a relational database, and access to this information will be done using web interfaces for all project members. The definition of project members is extended to include project managers, engineers, support personnel and, most of all, end users and other stakeholders. These tools normally provide an **enterprise-wide** framework for simultaneously managing multiple projects from multiple groups, teams, departments and sections of an organization, with **access control** on what is visible to each person, their privileges and allowed activities, and with additional tools to enable **collaboration** among all the interested parties.

Normally you will find a variety of tools, which depend on the orientation of the system, but which usually include facilities for managing a shared **address book**, a **calendar** facility for either personal use or shared with other people (e.g. to maintain/coordinate/announce meetings and common events), a **project management** environment for planning and tracking your **project tasks** usually with **logging** capabilities (i. e. annotating the work done like in a lab logbook), capabilities for **sharing information** (files), the ability to create project-specific discussion **forums**, some **problem reporting/issue tracking** facility to allow you to provide support to your users and fix problems as they arise and the ability to generate a variety of **project reports**.

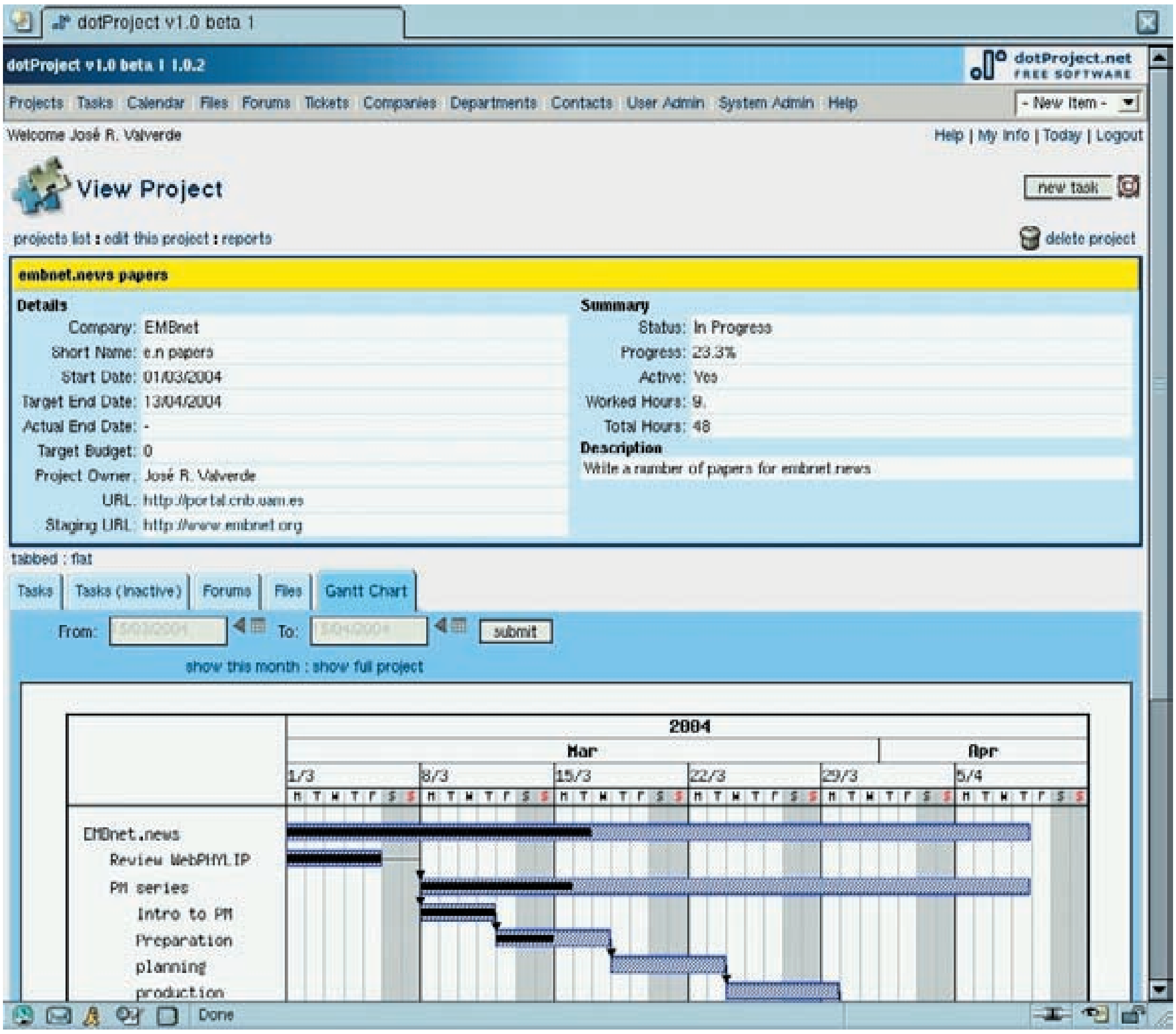
Many of these tools also provide for additional facilities of special interest like **timecards**, a facility to track how much time is devoted to each task: you first check-in before starting work, and then check-out when you stop working on a given task, and the system records automatically the time devoted and collects the appropriate statistics.

You want to have one of these tools when you want to provide a single, common, multiple team tool, manage and coordinate many projects, take into account people involvement in more than one project, give access to users or stakeholders to project information, coordinate daily work (e.g. using a shared calendar, discussion forums, etc..) and have a rich track of project and work progress.

dotProject

This is a generic PM environment that may be obtained from `http://www.dotproject.net/`. It has a simple and pleasant layout, and provides all of the basic facilities described above. It has a **company-based** orientation, and includes facilities for project and task management, address books, calendar, forms, issue tracking and file sharing.

You can **catalogue and make copies of projects** (which often comes handy when you have projects that share many aspects and have a common structure or template) and define as many tasks and task levels as you need, indicating **task dependencies** and automatically computing task duration and



embnet.news papers

Details

Company: EMBnet
Short Name: e.n papers
Start Date: 01/03/2004
Target End Date: 13/04/2004
Actual End Date: -
Target Budget: 0
Project Owner: José R. Valverde
URL: http://portal.cnb.uam.es
Staging URL: http://www.embnet.org

Summary

Status: In Progress
Progress: 23.3%
Active: Yes
Worked Hours: 9.
Total Hours: 48

Description

Write a number of papers for embnet.news

Figure 4. dotProject window.

dates for complex tasks, maintain detailed **task logs** and let the program take care of automatically **recording the time** devoted to each task. The personal assistant facilities are neat and appropriate.

Besides a good tool for drawing charts, dotProject allows a variety of project **reports** to be generated to view task logbooks, time allocation and user performance, which are useful not only for automating the periodic project reporting required by funding agencies but also to help you analyze how good your initial estimates were and make more educated guesses and corrections to future plans. Its minus side is the **access control** and privilege mechanism: it is by far the most detailed and fine-grained of all the tools reviewed, requiring a lot of work to maintain. The authors are working on simpler user-role based mechanisms.

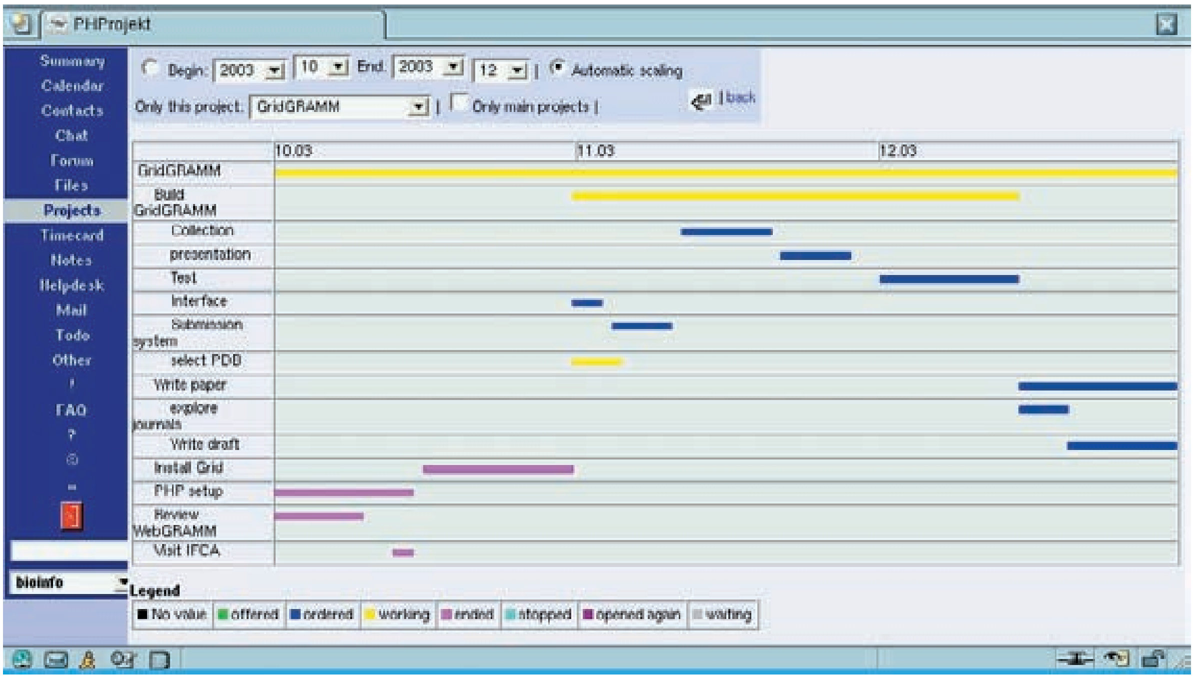
In general, dotProject is good enough for most scientific and research users and has the cleanest and simplest user interface. Unless you have such a broad and dynamic staff that managing access privileges is a problem, this is probably a best-compromise tool.

phProjekt

This tool is more complete, but also has a more airy web interface. You may get it from `http://www.phprojekt.org/` and it has commercial support. In addition to the tools described, it provides an **online chat** system, a **web-mail** environment, **todo lists**, **featured links**, **voting mechanisms** and an integrated **helpdesk**.

phProjekt susbtitutes the concept of a task by that of a **subproject**. Everything is therefore a project, which may be delegated to other people to further refine. This means an effective work and responsibility delegation infrastructure. The work subdivision mechanism may accommodate **complex relationships** among tasks, allowing for a rich variety of task dependencies (although they are not taken into account yet). Time tracking at project level is possible with a complete **timecard** system, and you can generate project activity **reports** and statistics with greater versatility.

User management is easier as well: it is based on roles and privileges: an easy mechanism allows to distribute privileges into roles and



○ Begin: 2003 10 End: 2003 12 | ● Automatic scaling

Only this project: GridGRAMM | ☐ Only main projects |

Figure 5. phProjekt example window.

then to assign a given role (e.g. Project manager, engineer, plain user) to each user or group. Users may be assigned to one or more groups with different roles, thus having different access rights to different group resources.

The **voting** mechanism is another interesting add-on: it allows you to poll project members about relevant issues automatically on log-on and track answers and opinions. **To-do lists** add to the personal information assistant utility and make for a rather complete collaborative work environment. Other niceties include the ability to add **plug-ins** for SMS notifications of events, knowledge-bases, web content management, integration with PostNuke-like portals, etc...

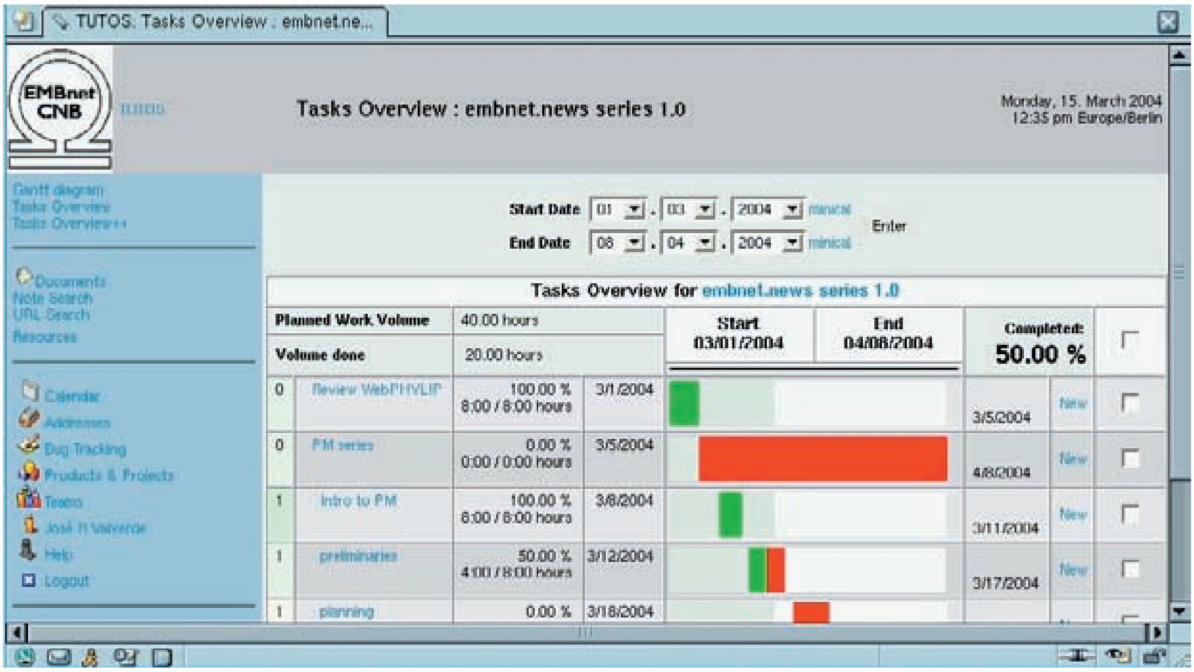
All in all, phProjekt is a very versatile, complete, general-purpose, web-based PM tool. On the minus side, it has a less pleasant look with a user interface that is somewhat wanting.

TUTOS

The Ultimate Team Organization Software (`http://www.tutos.org/`) is a specialized tool which offers a different set of features. It is more oriented towards a company providing **products** or services to external customers.

TUTOS has a more convoluted **user interface**, which changes heavily depending on what the user selects at any time, the results being somewhat confusing for the beginner. On the other hand, it has powerful facilities which may prove really useful if you need them. As the other tools, TUTOS provides a powerful shared **calendar**, a shared **file repository**, a **bug-tracking** system, support for **products and projects**, **time tracking** and a rich **user management** system.

The central concept in TUTOS is that of a product to be delivered to a customer. TUTOS thus provides tools for **cost estimation and invoicing**, and has extended support for the product development life cycle: in addition to classic planning tools (project subdivision into tasks, task assignment, progress tracking) it includes tools for managing **product installations** (deployments of the finished product at different customer sites), and product **support**.



Tasks Overview : embnet.news series 1.0

Monday, 15. March 2004
12:35 pm Europe/Berlin

Start Date 01 . 03 . 2004
End Date 08 . 04 . 2004

Tasks Overview for embnet.news series 1.0

Planned Work Volume: 40.00 hours
Volume done: 20.00 hours
Start 03/01/2004
End 04/08/2004
Completed: 50.00 %

0	Review WebPHYLIP	100.00 % 8:00 / 8:00 hours	3/1/2004	3/5/2004	New
0	PM series	0.00 % 0:00 / 0:00 hours	3/5/2004	4/8/2004	New
1	Intro to PM	100.00 % 8:00 / 8:00 hours	3/8/2004	3/11/2004	New
1	preliminaries	50.00 % 4:00 / 8:00 hours	3/12/2004	3/17/2004	New
1	planning	0.00 %	3/18/2004		

Figure 6. tutos example window.

TUTOS generalizes the concept of **resources** to include everything, allowing for a richer control, reserve and allocation of a project resources (people, computers, reactants, rooms, material, whatever), Users in TUTOS can be assigned priorities according to roles and group membership as in other tools, and there is provision for managing per-user **work calendars** (i.e., internationalize holidays and define working days).

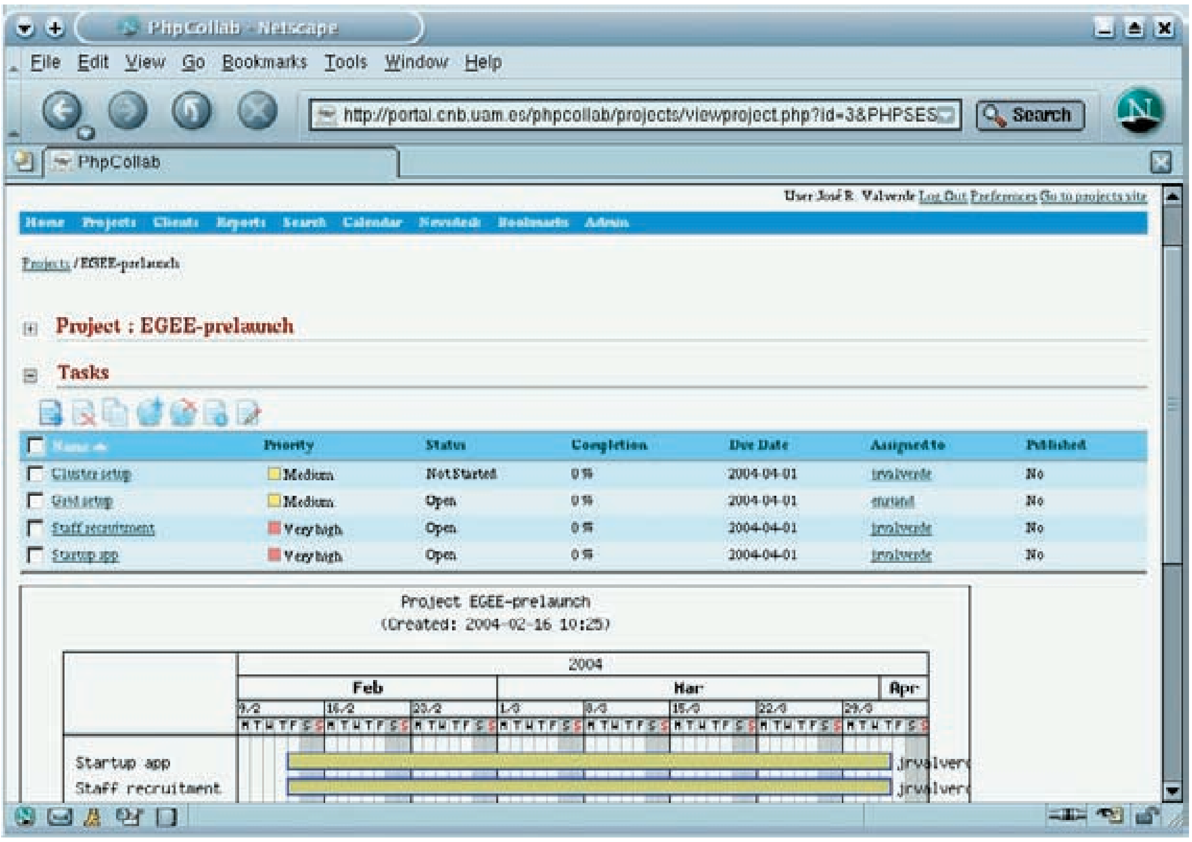
TUTOS is more oriented towards customer-based work, and has some features that make it more professional, but has a steeper learning curve. You will probably prefer it for managing external services, organizing work in your startup or spin-off company or simply want to manage work in an international environment or many projects with advanced computer users.

phpCollab

Available from `http://www.phpcollab.com/` this tool is very pleasant to use and provides interesting features that are of use to facilitate communication among project developers and with stakeholders and users.

While in the other tools, users belong to one or many groups which have associated projects, in phpCollab the **client is the King**: you start with a client, for whom you create one or more projects, and for each project you define a working team out of your pool of employees. Projects may have **discussion forums**, **linked content**, **notes** and **shared data**, and as with the other tools, you get a **shared calendar** tool and a **news desk** which you can use to disseminate news about shared events and relevant changes.

The project model in phpCollab assumes a project is divided in **phases** (which are customisable), and each of these into tasks. Its minus side is that it doesn't provide for tasks to be further sub-divided into smaller



Project : EGEE-prelaunch

Tasks

☐ Name	Priority	Status	Completion	Due Date	Assigned to	Published
☐ Cluster setup	🟨 Medium	Not Started	0 %	2004-04-01	jrvalverde	No
☐ Grid setup	🟨 Medium	Open	0 %	2004-04-01	[illegible]	No
☐ Staff recruitment	🟥 Very high	Open	0 %	2004-04-01	jrvalverde	No
☐ Startup app	🟥 Very high	Open	0 %	2004-04-01	jrvalverde	No

Project EGEE-prelaunch
(Created: 2004-02-16 10:25)

Figure 7. phpCollab example window.

subtasks, hence reducing the levels of abstraction. You have the ability to make **copies of projects** like in dotProject to save work on similar intitiatives.

phpCollab presents **three different interfaces**: one is the **management** interface where you plan everything, a second one is a **developer** interface, with consists of a special web site automatically generated from the management definitions to enable collaborative work by project developers, and the third is a **client** interface, also generated automatically and which consists of a separate web site where users may monitor project development and participate -within the pre-established limits- in the project evolution.

Just like the other tools, phpCollab allows you to generate nice Gantt diagrams and a variety of **reports** to monitor the evolution of your project. Like TUTOS, it allows for the automatic generation of **invoices** based on project costs. The user interfaces (the three of them) are pleasant and intuitive, easy to use, and the program remembers your work environment when you leave so as to bring you back into the same one when you log back in.

phpCollab is therefore a really interesting tool, which is most useful when your work is user-centered, and you want to give a professional, simplified and easy access to users of your project and developers alike.

Other web-based tools

There are other tools around which either provide a generic framework or provide for specific parts of a generic framework. Obviously, there are forum, chat, weblog, bug-tracking and others tools which can be used to do an excellent work.

As for other generic project management tools there are simply not that many additional tools worth mentioning. Some, like webCollab (`http://webcollab.sourceforge.net`) add little to what we have already described, except for user interface aspects. Others are too immature yet for our taste. Regarding specialized task tools, what can we say? They are certainly excellent, but integrating all of them in a single, seamless environment is a lot more work, and unless you need all of their features, you will probably be better off with one of these collaborative tools that provide a basic working framework useful enough for most common projects.

Discussion

We have summarized our experience with a number of tools for project management that will help you better manage and track the progress of projects and better harness the power of teams, much more so if these are -as is often the case nowadays in science- spread over different geographical locations.

To start with you may want to begin by installing the Java-based tools and give them a try to familiarize yourself with the basic PM tasks. As you start to feel that PM is too cumbersome a task (which it usually is), you will want to delegate work to your co-workers. A first step is storing GanttProject files on a WebDAV enabled browser, but if you really want to harness the power of collaborative work, then a centralized, web-based solution is your best choice. dotProject is good enough for most purposes, but if your demands grow beyond it or you want to move to a more customer-oriented system, you will want to give to the other tools a try. And it is very easy: all of them provide **demo sites** where you can try them by yourself. You are hereby encouraged to try each of these tools and find which is the one that best suits your needs.

For further information and discussion you may visit the EMBnet/CNB portal, where you will be able to participate and add your opinions, experiences, pet peeves, etc. and where updated information will be posted as it becomes available.

NATURE'S PHILANDERERS

By Vivienne Baillie Gerritsen

Could chemistry be at the heart of sexual wanderings? Or of sexual devotion for that matter? Well – though the idea certainly lacks romantic appeal – there are signs which point in this direction. The neuropeptide vasopressin is not a newcomer to research revolving around social behavior in animals. However, narrowing one of its roles down to what scientists coolly term 'social cohesion', and Christians would call 'infidelity', is quite a breakthrough. And deserves some thought. It is not so much the quantity of vasopressin but the tissue distribution of its receptor – in males - which seems to have a role in defining flirt or fidelity. In male voles that is.

Monogamy in the animal world is not so common. The examples are few actually. Amongst them the California mouse, the marmoset monkey and certain species of voles. In fact, only five percent of mammals stick to their mates. The ongoing theory is that most males tend to mate with more than one female, not in the pursuit of pleasure but in the pursuit of producing offspring with the same genetic buildup as themselves. Sperm –physiologically speaking – is cheap so they can afford to spread their wares in this manner.

Fig. 1 The furry flirt (*Microtus montanus*)

Where do humans stand one wonders? Well, from a purely biological point of view, humans are no different from the other 95% of non-monogamous mammals. For instance, monogamous species are monomorphic, i.e. the male and female are the same size. Humans are dimorphic: men, on the whole, are larger than women. Monogamy – or polygamy – in human societies today is really set on human beliefs. In fact, monogamy probably became the name of the game in our parts of the world by way of the Roman Catholics in the 11th century. In those days, it was quite common for priests to have more than one wife, or even mistress. Only if they had a son, or more, the latter could inherit church property. So sons were declared illegitimate. The next step was to ban marriage for priests altogether. And, accordingly, celibacy became a standard of spirituality. Monogamy was all right for the spiritually inferior; infidelity was not to be heard of.

The knowledge that vasopressin has a role in male social behavior such as communication, aggression, sexual commitment and paternal care of offspring is not new. The idea had already been emitted in the 1960s following research on rats and mice. Vasopressin was administered into various tissues of the rodents and modulations in their behavior were observed. More recently, the same kind of tests were carried out on various species of voles: the monogamous prairie voles (*Microtus ochrogaster*) and their promiscuous relatives, the non-monogamous meadow vole (*Microtus pennsylvanicus*). When vasopressin antagonists were administered to male prairie voles, they turned into little furry philanderers, mating quite happily with many females. However, when vasopressin was administered to male meadow voles, it had no effect whatsoever on their mating habits.

What made the difference? The number of vasopressin receptors. Prairie voles bear far more vasopressin receptors in their ventral forebrains than their meadow vole homologues. When voles – or any mammal including humans – mate, they release quantities of vasopressin in the forebrain, which bind to their receptors. What is more, this particular region in the brain happens to be very close to what is known as the 'reward center'. What vasopressin ultimately triggers off is the feeling that mating is quite agreeable. But – granted – this is not enough to remain faithful. The vole also has to remember its partner. Besides behavior, vasopressin has a role in social recognition and memory. The vole cue for recognition is probably olfactory. So, the ultimate feeling of something nice coupled to recognizing its partner leads the prairie vole to a form of fidelity.

How about meadow voles then? They also have vasopressin receptors in their forebrains. Yes, but not nearly so many as their prairie counterparts. To prove the fidelity theory, the vasopressin receptor gene was introduced into the forebrains of young meadow voles. And yes, you guessed, the meadow voles lost all wish to flirt. Once they had mated with one female, no other female could lead them astray.

Why is it that prairie voles have more receptors in their forebrain than meadow voles? As a species, they are very close. Their vasopressin receptor sequences are also 99% identical. The flanking sequences on the vasopressin receptor genes differ however; and it is these flanking sequences which determine the tissue distribution of the receptors. The flanking sequences are comparable within one species but vary between species. Hence, differences in the flanking regions could echo differences in social behavior between individuals and between species.

So, does the way we behave depend solely on a bunch of hormones? How many vasopressin receptors do men have in their forebrains some of you may be asking? Can the act of adultery be waved off on the basis of chemistry? Discussing the promiscuousness of men, or women, is not the issue here. Rather, what is interesting is that the distribution of vasopressin receptors could have a role in the variability of human social behavior. For instance, it could explain in part light differences in social behavior between people but also more extreme dysfunctions such as autism where individuals have great difficulty in forming social bonds. Humans are not voles though and an individual's behavior cannot be reduced to a single gene. Social behavior involves highly complex biological pathways which themselves interact with an individual's particular past and present environment. And let us just hope that we have our own say in all this as well.

Cross-references to Swiss-Prot

Vasopressin V1a receptor, *Microtus montanus* (Montane vole) : Q9WTV8
Vasopressin V1a receptor, *Microtus ochrogaster* (Prairie vole) : Q9WTV9

References

1. Lim M.M., Wang Z., Olazabal D.E., Ren X., Terwilliger E.F., Young L.J.
   Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene
   Nature 429:754-757(2004)
   PMID: 15201909

2. Lim M.M., Hammock E.A.D., Young L.J.
   The role of vasopressin in the genetic and neural regulation of monogamy
   J. Neuroendocrinol. 16:325-332(2004)
   PMID: 15089970

3. Balaban E.
   Why voles stick together
   Nature 429:711-712(2004)
   PMID: 15201896

National Nodes

Argentina
Oscar Grau
IBBM, Facultad de Cs. Exactas, Universidad Nacional de La Plata
Email: grau@biol.unlp.edu.ar
Tel: +54-221-4259223 Fax: +54-221-4259223
http://www.ar.embnet.org

Australia
Sonia Cattley
RMC Gunn Building B19, University of Sydney,NSW, 2006
Email: scattley@angis.org.au
Tel: +61-2-9531 2948
http://www.au.embnet.org

Austria
Martin Grabner
Vienna Bio Center, University of Vienna
Email: martin.grabner@univie.ac.at
Tel: +43-1-4277/14141
http://www.at.embnet.org

Belgium
Robert Herzog, Marc Colet
BEN ULB Campus Plaine CP 257
Email: rherzog@ulb.ac.be mcolet@ulb.ac.be
Tel: +32 2 6505146 Fax: +32 2 6505124
http://www.be.embnet.org

Brasil
Gonçalo Guimaraes Pereira
Laboratório de Genômica e Expressão - IB
UNICAMP-CP 6109
13083-970 Campinas-SP, BRASIL
Tel: 0055-19-37886237/6238
Fax: 0055-19-37886235
Email: goncalo@unicamp.br
http://www.br.embnet.org

Canada
Canadian Bioinformatics Resource, National Research Council Canada, Institute for Marine Biosciences,
Email: manager@cbr.nrc.ca
Tel: +1-902-426 7310 Fax: +1-902-426 9413
http://www.ca.embnet.org

Chile
Dr. Ricardo Baeza-Yates
Dept. of Computer Science, Santiago,
Email: rbaeza@dcc.uchile.cl
Tel: N/A
http://www.embnet.cl

China
Jingchu Luo
Centre of Bioinformatics
Peking University
Beijing 100871, China
Tel: 86-10-6275-7281
Fax: 86-10-6275-9001
Email: luojc@pku.edu.cn
http://www.cbi.pku.edu.cn

Colombia
Emiliano Barreto Hernández
Instituto de Biotecnología
Universidad Nacional de Colombia
Edificio Manuel Ancizar
Bogota - Colombia
Tel: +571 3165027 Fax: +571 3165415
Email : ebarreto@ibun.unal.edu.co
http://bioinf.ibun.unal.edu.co

Cuba
Ricardo Bringas
Centro de Ingeniería Genética y Biotecnolgía,
La Habana, Cuba
Email: bringas@cigb.edu.cu
Tel: +53 7 218200
http://www.cu.embnet.org

Finland
Eija Korpelainen
CSC, Espoo
Email: eija.korpelainen@csc.fi
Tel: +358 9 457 2030
http://www.fi.embnet.org

France
Jean-Marc Plaza
INFOBIOGEN, Evry
Email: plaza@infobiogen.fr
Tel: +33 1 60 87 37 11 Fax: +33 1 60 87 37 96
http://www.fr.embnet.org

Germany
Sandor Suhai
EMBnet node at the German Cancer Research Center
Department of Molecular Biophysics (H0200)
Email: genome@dkfz.de
Tel: +49-6221-422 342 Fax: +49-6221-422 333
http://www.de.embnet.org

Hungary
Endre Barta
Agricultural Biotechnology Center
Szent-Gyorgyi A. ut 4. Godollo,
Email: barta@abc.hu
Tel: +36 30-2101795
http://www.hu.embnet.org

India
H.A.Nagarajaram
Laboratory of Computational Biology & Bioinformatics facility, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad
Email: han@www.cdfd.org.in
Tel: +91 40 7155607 / 7151344 ext:1206
Fax : +9140 7155479
http://www.in.embnet.org

Israel
Leon Esterman
INN (Israeli National Node) Weizmann Institute of Science
Department of Biological Services, Biological Computing Unit, Rehovot
Email: Leon.Esterman@weizmann.ac.il
Tel: +972- 8-934 3456
http://www.il.embnet.org

Italy
Cecilia Saccone
CNR - Institute of Biomedical Technologies
Bioinformatics and Genomic Group
Via Amendola 168/5 - 70126 Bari (Italy)
Email: saccone@area.ba.cnr.it
Tel. +39-80-5482100 - Fax. +39-80-5482607
http://www.it.embnet.org

Mexico
Cesar Bonavides
Nodo Nacional EMBnet, Centro de Investigación sobre Fijación de Nitrógeno, Cuernavaca, Morelos
Email: embnetmx@cifn.unam.mx
Tel: +52 (7) 3 132063
http://embnet.cifn.unam.mx

The Netherlands
Jack A.M. Leunissen
Dept. of Genome Informatics
Wageningen UR, Dreijenlaan 3
6703 HA Wageningen, NL
Email: Jack.Leunissen@wur.nl
Tel: +31 317 484074
http://www.nl.embnet.org

Norway
George Magklaras
The Norwegian EMBnet Node
The Biotechnology Centre of Oslo
Email: admin@embnet.uio.no
Tel: +47 22 84 0535
http://www.no.embnet.org

Poland
Piotr Zielenkiwicz
Institute of Biochemistry and Biophysics
Polish Academy of Sciences Warszawa
Email: piotr@pl.embnet.org
Tel: +48-22 86584703
http://www.pl.embnet.org

Portugal
Pedro Fernandes
Instituto Gulbenkian de Ciencia
Unidade de Bioinformatica
2781-901 OEIRAS
Email: pfern@igc.gulbenkian.pt
Tel: +351 214407912 Fax: +351 2144079070
http://www.pt.embnet.org

Russia
Sergei Spirin
Biocomputing Group, Belozersky Institute Moscow
Email: sas@belozersky.msu.ru
Tel: +7-095-9395414
http://www.genebee.msu.ru

Slovakia
Lubos Klucar
Institute of Molecular Biology SAS Bratislava
Email: klucar@embnet.sk
Tel: +421 7 5941 2284
http://www.sk.embnet.org

South Africa
Ruediger Braeuning
SANBI, University of the Western Cape, Bellville
Email: ruediger@sanbi.ac.za
Tel: +27 (0)21 9593645
http://www.za.embnet.org

Spain
José M. Carazo, José R. Valverde
EMBnet/CNB, Centro Nacional de Biotecnología, Madrid
Email: carazo@es.embnet.org,
jrvalverde@es.embnet.org
Tel: +34 915 854 505 Fax: +34 915 854 506
http://www.es.embnet.org

Sweden
Nils-Einar Eriksson, Erik Bongcam-Rudloff
Uppsala Biomedical Centre, Computing Department, Uppsala, Sweden
Email: nils-einar.eriksson@bmc.uu.se
erik.bongcam@bmc.uu.se
Tel: +46-(0)18-4714017, +46-(0)18-4714525
http://www.embnet.se

Switzerland
Laurent Falquet
Swiss Institute of Bioinformatics, CH-1066 Epalinges
Email: Laurent.Falquet@isb-sib.ch
Tel: +41 (21) 692 5954 Fax: +41 (21) 692 5945
http://www.ch.embnet.org

United Kingdom
Alan Bleasby
UK MRC HGMP Resource Centre, Hinxton, Cambridge
Email: ableasby@embnet.org
Tel: +44 (0) 1223 494535
http://www.uk.embnet.org

Specialist Nodes

EBI

Rodrigo López
EBI Embl Outstation, Wellcome trust Genome Campus,
Hinxton Hall, Hinxton, Cambridge, United Kingdom
Email: rls@ebi.ac.uk
Phone: +44 (0)1223 494423
http://www.ebi.ac.uk

ETI

P.O. Box 94766
NL-1090 GT Amsterdam, The Netherlands
Email: wouter@eti.uva.nl
Phone: +31-20-5257239
Fax: +31-20-5257238
http://www.eti.uva.nl

ICGEB

Sándor Pongor
International Centre for Genetic Engineering and Biotechnology
AREA Science Park, Trieste, ITALY
Email: pongor@icgeb.trieste.it
Phone: +39 040 3757300
http://www.icgeb.trieste.it

LION Bioscience

Thure Etzold
LION Bioscience AG, Heidelberg, Germany
Email: Thure.Etzold@uk.lionbioscience.com
Phone: +44 1223 224700
http://www.lionbioscience.com

MIPS

H. Werner Mewes
Email: mewes@mips.embnet.org
Phone: +49-89-8578 2656
Fax: +49-89-8578 2655
http://www.mips.biochem.mpg.de

UMBER

Terri Attwood
School of Biological Sciences, The University of
Manchester, Oxford Road, Manchester M13 9PT, UK
Email: attwood@bioinf.man.ac.uk
Phone: +44 (0)61 275 5766
Fax: +44 (0) 61 275 5082
http://www.bioinf.man.ac.uk/dbbrowser

TECH-MGR

Email: tech-mgr@embnet.org
The team gives support to EMBnet nodes and helps them with maintenance and troubleshooting.
The team is formed of experienced system administrators and programmers who ensure the availability of local services for all EMBnet users.

EMBnet.news

ISSN 1023-4144

Dear reader,

If you have any comments or suggestions regarding this newsletter we would be very glad to hear from you. If you have a tip you feel we can print then please let us know. Please send your contributions to one of the editors. You may also submit material by e-mail.

Past issues of embnet.news are available as PostScript or PDF files (ISSN 1023-4144). You can get them by anonymous ftp from:
the EMBnet organisation Web site
`http://www.embnet.org/download/embnetnews`
the Belgian EMBnet node
`ftp://ftp.be.embnet.org/pub/embnet.news`
the UK EMBnet node
`ftp://ftp.uk.embnet.org/pub/embnet.news`
the EBI EMBnet node
`ftp://ftp.ebi.ac.uk/pub/embnet.news`

Submission deadline for the next issues:

February 28, 2005
May 31, 2005
August 15, 2005
October 31, 2005