EMBnet.news

Volume 11 Nr. 2
June 2005

- Integrating HERV data into ENSEMBL
- The EBI resources in a Nutshell (part 3)
- New QuickGuide

and more...

Editorial

The people of France and The Netherlands recently said NO to the European Constitution, contrasting with the YES of the Spanish people in March. Italy wants to drop the Euro and go back to the Lira. The UK decided to postpone their vote... What a nightmare... Is there a serious crisis in European institutions? Will these events have an influence on European science? The feeling is that the convergence will continue, but at a slower pace. We might have been going too fast and perhaps that a slow-down is necessary to let people digest the changes.

On the Science front things are moving fast. The 6th Framework Program is not yet finished and already the 7th FP is being announced, with double the budget (10 Billion euro / year). It will include similar tools (NoE, IP, STREP, CSA), as well as new tools among which «International Co-operation», which will allow researchers from non-EU members to participate in EU projects. A promising opportunity for non-EU EMBnet nodes!

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 issues 57&58. Please let us know if you like this inclusion.

Cover picture: The Eiffel Tower, Paris, May 2005 [® Laurent Falquet]

Contents

Editorial 2
Integrating HERV data into ENSEMBL 3
New QuickGuide 8
EBI resources in a Nutshell (3) 9
AGM2005 13
Protein spotlight 57 14
Protein spotlight 58 16
Node information 18

Editorial Board:

Publisher:

EMBnet Executive Board
c/o Erik Bongcam-Rudloff
Uppsala Biomedical Centre
The Linnaeus Centre for Bioinformatics, SLU/UU
Box 570 S-751 23 Uppsala, Sweden
Email: erik.bongcam@bmc.uu.se
Tel: +46-18-4716696

Integrating HERV data into Ensembl

Álvaro Martínez Barrio PhD Student
Erik Bongcam-Rudloff Associate Professor
The Linnaeus Centre for Bioinformatics,
BMC, Uppsala University, Box 598
SE-751 24 Uppsala, SWEDEN

Introduction

In the last year I have been working on integrating the display of certain retroviruses, namely Human endogenous retrovirus (HERVs), in Ensembl. We aim to provide a specific tool for investigating the associations between HERV genes and complex diseases, including multiple sclerosis (Perron et al. 1997), schizophrenia (Karlsson et al. 2001) and cancer (Griffiths 2001), for a review see Blomberg et al., 2004. Our ultimate milestone was specifically to set up an Ensembl prototype that would provide experimental serving and integration of new types of data through a Distributed Annotation System (DAS) protocol. As a further step we are planning to enable the same functionality using the web services technology. A wide description of this work can be found in my Master Thesis.

We will describe the entire process of integrating remote data into Ensembl, from the installation of your own Ensembl mirror prototype –as we did in a Mac OS X platform- to the setting and configuration of the DAS server for inhouse data. We developed these services in the same way that Biosapiens effort (`http://www.biosapiens.info/page.php?page=home`) distributes QTL annotation tracks for the mouse genome. Data will be shown on the View pages of Ensembl but served from the LCB. Ensembl provides the means for this purpose.

Mirroring Ensembl

The process for installing a mirror of Ensembl system is well documented at `http://www.ensembl.org/Docs/wiki/html/EnsemblDocs/InstallEnsemblWebsite.html`. The basic steps to be taken are shown but the document does not mention any of the more obscure details of the installation, and how the resulting problems can be solved.

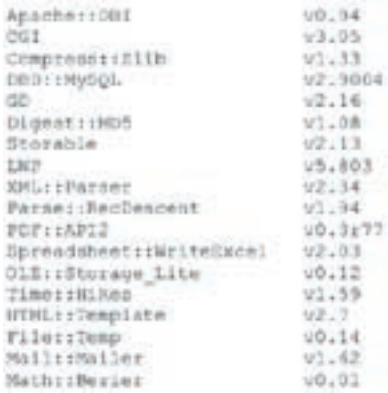
Computer resources for this project were provided by EMBnet Sweden at The Linneaus Centre for Bioinformatics (LCB), Uppsala University. Apache web server, Ensembl installation, database MySQL engine installation, DAS server installation and software development were performed on a double processor workstation (2x1.25GHz PPC Apple G4, 1.75GB DDR SDRAM) with a Mac OS X 10.3 (Panther) operating system (OS) installation.

Due to the outdated version of the OS when the installation was initially performed, we had to patch some problems in different programs with newer installations. As updated versions of Mac OS X were released, these problems were fixed by Apple. From version 10.3.5 on, new installations are no longer required. Installations of such programs as the Apache web server (in the `mod_perl` and `libapreq` libraries), Perl (in `Config.pm`) or gcc (in both versions 2.95 and 3.1) that Apple's OS was equipped with by default are either patched or upgraded, to prevent further incompatibility issues arising.

MySQL should be next installed. We used a binary distribution for Mac OS X.

On Figure 1 the component dependencies are shown. Versions for all the prerequisites installed were:

When installing the various Perl modules, we made extensive use of the Comprehensive Perl Archive Network (CPAN) which has its own module by default installed in Perl and can be executed to fetch and install modules (`$> perl -MCPAN -e shell`). In the new prompt, write help in case you need it (`cpan> help`). The Perl modules (`http://www.cpan.org/modules/by-module/`) installed were:



There were several problems with DBD::mysql and GD. Therefore, we decided to install them manually. We downloaded the modules and fixed the errors at installation because we needed to compile them for our newly fresh Perl installation. But for those users of Mac OS X 10.3.5 and over, this is no longer necessary. Instead you'd rather use Fink (`http://fink.sourceforge.net/`) that ports and distributes different UNIX software that compiles and installs on Mac OS X. By installing the packages of DBD::mysql (named as `dbd-mysql-pm581`) and GD (both `gd` library with its Perl wrapper module `gd-pm581`) corresponding to the Perl version shipped, the Perl modules can be successfully installed.

Dotter is downloadable (`http://www.acedb.org/Software/Downloads/`) as an executable for Mac OS X which dynamically links to libraries placed on "`/sw/lib`", the path where the Fink program places libraries. It requires the Gimp Toolkit and readline, however readline is installed by default. Fink will easily install the Gimp Toolkit (`gtk+2`).

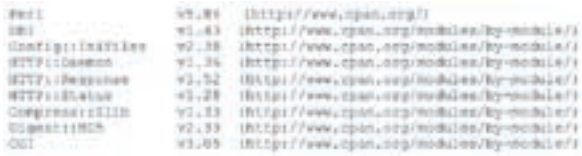
The Ensembl web site software is comprised of three distinct components: Ensembl databases, Ensembl API code and Ensembl web plus the BioPerl modules. Databases schemas and dumps are distributed through an FTP server when new versions are released (see `http://www.ensembl.org/Download/`). Data can be imported directly into MySQL from these files in a tab separated format (command `mysqlimport`). The rest of the components can be downloaded through CVS.

A proper server should start its services (generally represented by OS daemons) on boot time. Therefore, a daemon startup was setup for every service (`http://www.macdevcenter.com/pub/a/mac/2003/10/21/startup.html`; `http://documents.wolfram.com/v4/GettingStarted/OnMacOSX.html`), namely MySQL, Tomcat, Ensembl and the DAS server.

Other systems were created or adapted to cover certain features which are missed in an Ensembl open distribution. These systems were a helpdesk report system, a search system to substitute the Altavista proprietary engine and a submission system to use Apple's Xgrid prerelease 2 for the Ensembl blastview search. These are not reported because they are out of the main focused topic here.

Serving DAS features

Among the available DAS servers, Proserver (`http://www.sanger.ac.uk/Software/analysis/proserver/`), Dazzle (`http://www.biojava.org/dazzle/`) and LDAS (`http://biodas.org/servers/`), we chose ProServer for our prototype. LDAS is very limited by heavy dependencies and file format fields which restrict data sources features. Dazzle is very modular and it is necessary to add small data source plugins to provide access to a range of databases, but Java servlet technology requires installation of BioJava APIs and java programming language skills. However, Proserver is easy to extend with new Perl adaptors for new data sources and it can serve data from virtually anything (files, any kind of database) thanks to the transport modules. It also has few installation requirements. The following are the necessary requirements in Proserver and the versions we installed for them:



The various adaptors and transports might have their own dependencies. In our case, the Perl module that was needed for installation was DBD::mysql to query the MySQL database. It was however installed as part of the Ensembl installation as mentioned above.

In Figure 2, a class schema for the Proserver DAS server is presented. Dependance relations are represented by continuous lines and inheritance relations by dashed lines. The direction of the arrows indicates the class that uses/inherits.

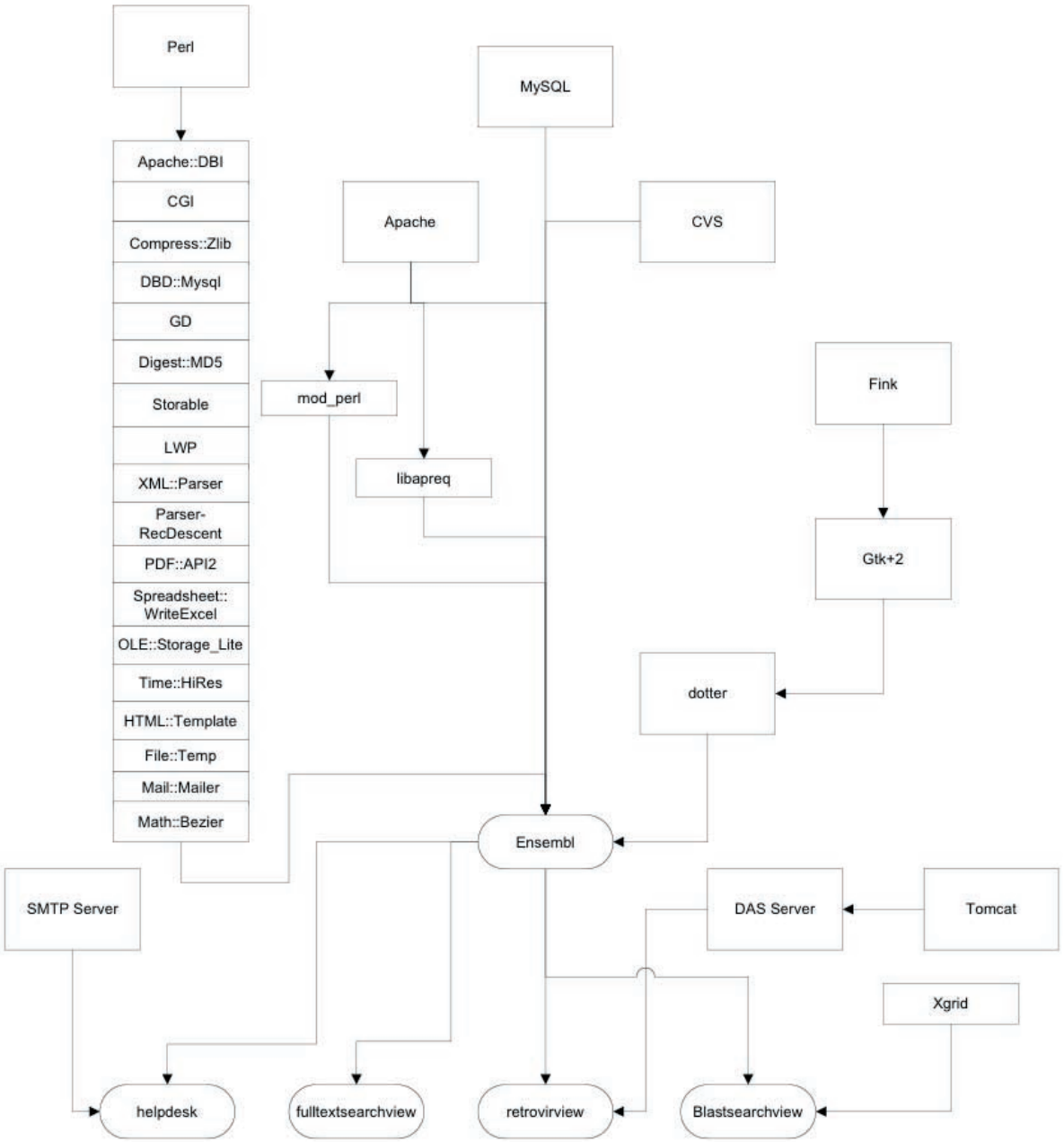


Figure 1: Diagram with software requirements and depencies on installation (indicated by arrows directions).

Figure 2: Proserver class diagram. Rectangular box symbol for the main executable script. Ellipsoid shaded boxes for overwritten classes.

Transport modules encapsulate the transactions between the adaptor and the data source. The dbi and oracle transports will return data, given any SQL "SELECT" statement, in a reference to arrays. Every array will contain a database row, and every array field will contain a row array. This removed the problem of having to connect, prepare, execute and disconnect with the database when querying it.

Adaptors, in order to serve the inhouse annotated features, inherit from Bio::Das:: ProServer::SourceAdaptor and implement a *build_features()* method that returns an array of hash tables.

At `http://www.ensembl.org/das/ensembl-das.pdf` there is a tutorial document placed on how to setup a DAS ProServer to serve data. Follow the simple steps contained in this document to set up a working ProServer. This document also details how to configure the new DAS data source for your Ensembl mirror.

Specifically for our system a configuration script file called `herv.ini` was prepared. When Proserver started, the Config module is loaded with the values of the parameters contained in this file. These parameters define the names of databases, tables, hosts, user and passwords to avoid tying the source name adaptor to rigid built in parametric variables. Ini files are very useful for configuring source adaptors. It can be noted in Figure 2 that these adaptors can easily access the Config module and retrieve

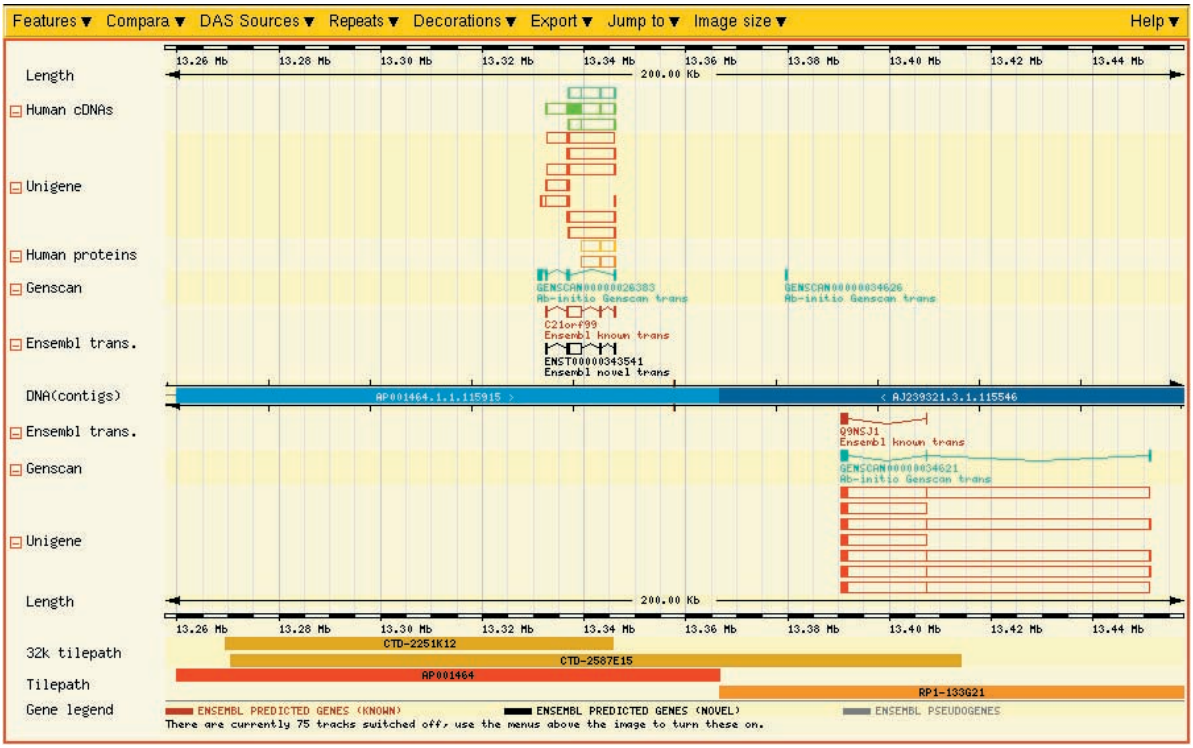


Figure 3: ContigView page with the HERV information deactivated.

the values of parameters into variables for programing purposes.
A source adaptor class for our HERVs database called `herv.pm` was also created. It inherits from SourceAdaptor. It fetchs the data from the database and parses features to be displayed with its *build_features()* method.

A practical example of HERV retrieval within our DAS system is as follows:

In the ContigView page of the local mirror (Figure 3). Use the DAS Sources ▼ menu to add Human HERVs information. A new ContigView image will appear when refreshing the page after closing the menu (Figure 4).

Placing mouse over the HERV and follow the DAS LINK in the pop-up window that leads to the HERView report created to host the information of every HERV (Figure 5).

New features and richer information report in this View page are under development.

At the same time, the Java DAS server, Dazzle, is going to be supported with the recent installation of the Tomcat web servlet engine (notice that `JAVA_HOME` points to `/Library/Java/Home` on Mac OS X) and the ongoing port of the Perl HERV adaptor to Java in order to be compatible and work within Dazzle server. Indications about how to do this can be found at `http://www.ensembl.org/Docs/linked_docs/das_server_v1.2.pdf`.

The development and registry of web services with HERV information will be supported. We are adopting a closer approach to projects as BioMoby (`http://www.biomoby.org/SettingUpPerlServices.html`).

References

•Blomberg, J., Ushameckis, D., Jern, P. 2004. Evolutionary Aspects of Human Endogenous Retroviral Sequences (HERVs) and Disease. *In Retroviral and Primate Genome Evolution*, edited by Eugene D. Sverdlov, Eurekah.com

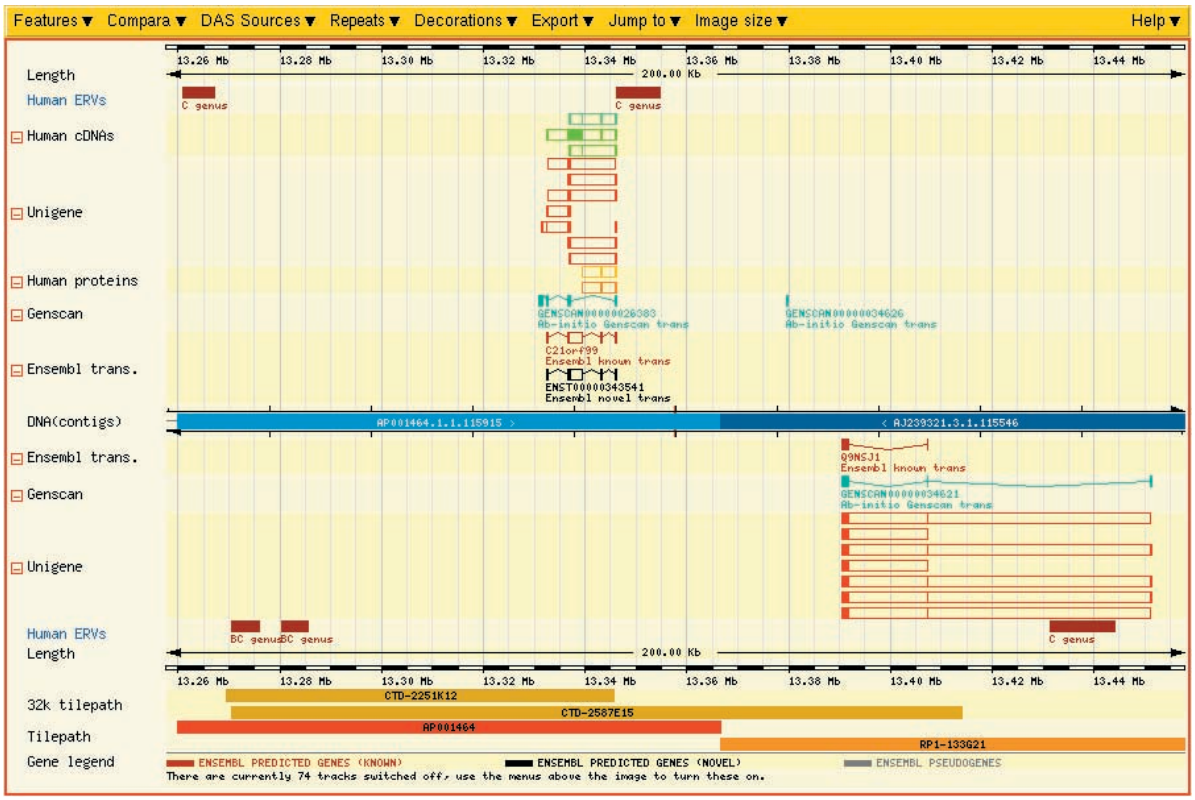


Figure 4: ContigView page with the HERV information activated.

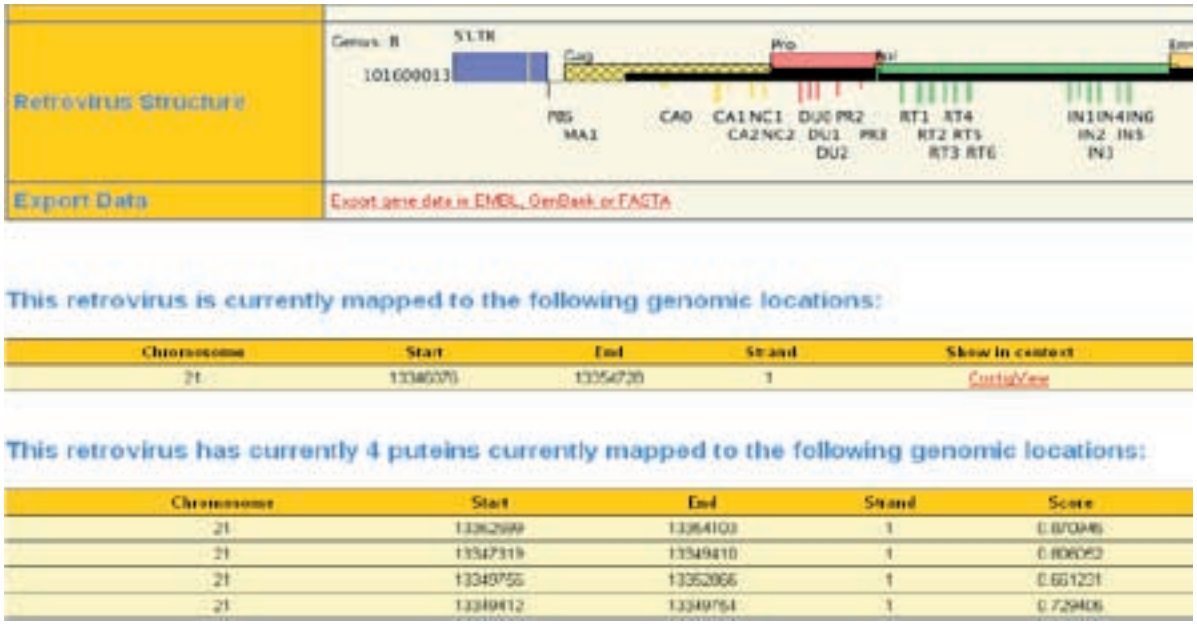


Figure 5: HERView report. Information about putative genes within the displayed viral structure is shown.

•Griffiths, D.J. 2001. Endogenous retroviruses in the human genome sequence. *Genome Biol* 2, 1017.1-1017.3.

•Karlsson, H., Bachmann, X., Silke, B., Johannes, S., Justin, McA., Torrey, E.F., Yolken, R.H. 2001. Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. *PNAS* 98, 4634-4639.

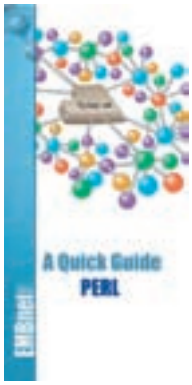
•Perron, H., Garson, J.A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., Mallet, F., Tuke, P.W., Voisset, C., Blond, J.L. 1997. Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. *PNAS* 94, 7583-7588.

Acknowledgements

RetroTector HERV database is a product of Ass. Professor. Göran Sperber, Depts of Neuroscience and Medical Sciences, Uppsala University, Sweden. Information is in silico annotated before storage in a MySQL database.

Two new EMBnet QuickGuide available !

We are happy to announce that our famous QuickGuide series has seen the birth of 2 new members. Designed by Laurent Falquet & Vassilios Ioannidis of the Swiss EMBnet node, these guides describe the PERL programming language syntax and PERL Regular Expressions in detail and provides the user with an extensive list of useful commands and advices. These two guides benefit from being used together.



These new guides and older versions are freely available from our web site `http://www.embnet.org/download/guides.html`

The EBI resources in a Nutshell (part 3)

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 (see Part 1&2 in EMBnet.news issues 10.4 & 11.1).

The UniProt knowledgebase

The Uniprot (Universal Protein Resource) database was implemented in 2003 and combines the protein resources of SwissProt, TrEMBL and PIR. It is accessible by following the protein accession number link from the original table generated in CluSTr.

Alternatively, go to the EBI homepage at `http://www.ebi.ac.uk` and in the search

field at the top of the page enter the SwissProt accession number for the BRCA2_HUMAN protein (P51587). Select Protein sequences from the pull-down menu and hit Go (see next figure).

Beneath the Data Set manager and Query information is the protein entry in Uniprot/Swissprot.

The UniProt Consortium is comprised of the European Bioinformatics Institute, the Swiss

Institute of Bioinformatics and the Protein Information Resource. The UniProt consortium aims to support biological research by maintaining a high quality database that serves as a stable, comprehensive, fully classified, richly and accurately annotated protein sequence knowledgebase, with extensive cross-references and querying interfaces freely accessible to the scientific community.

All data stored in UniProt can be downloaded from the Download Centre at `http://www.ebi.uniprot.org/database/download.shtml`.

Scroll down to the feature information and note that it does not reveal a "varsplic" annotation specifying an alternative splice site. The Uniprot annotators use information of alternative splicing based on cDNA information. In this instance, whilst alternative splicing has been confirmed (see AEdb) it has been done so from EST transcripts.

The IntAct database

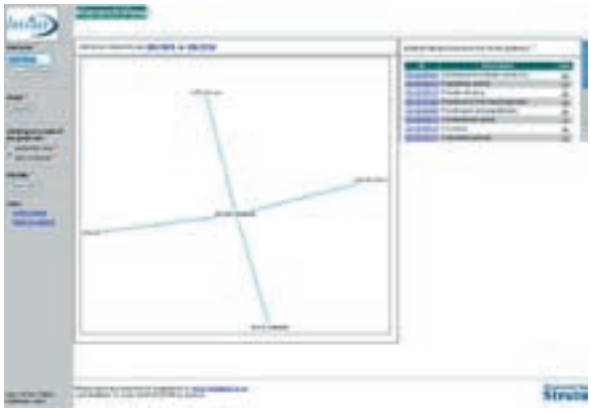
In the UniProt entry, follow the link to the IntAct Database. Every search is done as a live search each time. The resulting table indicates that the protein under investigation has been found to interact DSS1_HUMAN (4 interactions), RAD51_HUMAN (5 interactions), Q8BXW9 (2 interactions).

Click on the links detailing the number of reactions to access more information on each protein.

Go back to the original results page and check the box to the left of the DSS1 _ HUMAN entry. Select the Graph from below the table. Enlarge the resulting window to maximum size and scroll so that the interaction between the BRCA2 _ HUMAN protein and the DSS1 _ HUMAN is displayed.

The length of the line is not currently significant. The dss1 protein is in bold type as that was the original graph query.

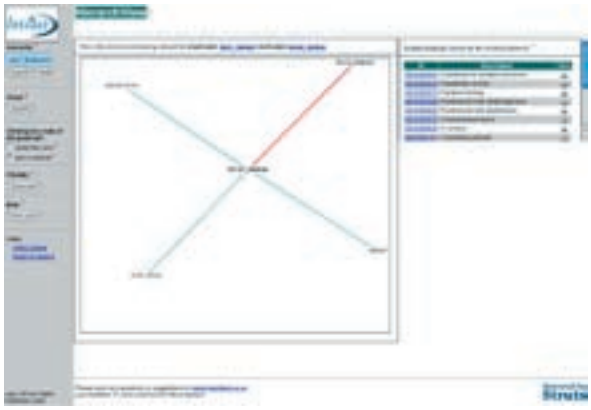
Select "add a network" from the options on the left hand panel and then left click on the BRCA2 _ HUMAN protein in the graph. The graphical view will now display the other interaction that BRCA2 _ HUMAN is associated with RAD51 _ HUMAN.



GO identifiers now appear on the right hand side of the page. Clicking on the various links will display a quick GO ontology.

Try expanding the network of interaction by adding a network to RAD51 _ HUMAN.

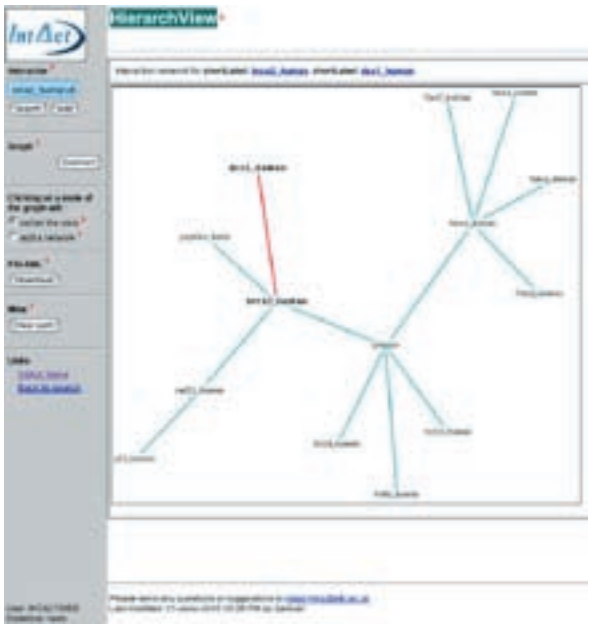
Return to the Database Search page, ensure that and follow the Path button to create a minimal interaction network. This time check the BRCA2 _ HUMAN box at the top of the results table as well as the DSS1 protein



already selected to create an interactive network.

The network can be expanded on either the graph or the path page a maximum of seven times. Once the GO ontology has been accessed, all proteins sharing a selected GO feature will be highlighted in red.

As the network progresses, the IntAct accession numbers relating to that particular interaction are displayed at the top of the graphic. This offers information on the relevant immediate interaction.



Click on the "Use" column on the right hand side of the GO table and note that whilst the BRCA2 _ HUMAN protein is associated with all terms, the majority of other interactions are associated with the nucleus term only.

Note the PSI-XML download option on the left hand side. This allows the XML data relative to the table to be saved and used in other PSI tools.

The Reactome database

Return to the UniProt entry and access the link to the Reactome database.
Reactome – previously the Genome Knowledgebase - is a curated database of biological processes in humans. The basic information is provided by bench biologists who are experts in that domain of biology. The information is then managed and edited by the Reactome staff at CSHL and the EBI, and entered into a relational database. They are then reviewed by other biological researchers for consistency and accuracy. Following peer-review, the information is published to the web.

The data covers biological pathways ranging from the basic processes of metabolism to high-level processes such as hormonal signalling. While Reactome is targeted at human pathways, it also includes many individual biochemical reactions from non-human systems such as rat, mouse, fugu fish and zebra fish. All the information in Reactome is backed up by either a literature citation or an electronic inference based on sequence similarity.

Reactome is an open source project and data can be downloaded from `http://www.reactome.org/download/index.html` or by following the download tab on the Reactome pages.

The area which includes the pathway associated with this protein is highlighted in yellow. Mouse over this region to see the individual pathways appear. Zoom in twice using the radio buttons at the top of the display and clicking over the yellow

highlighted region to display the individual pathways.

The first of the tabular field down indicates references to the protein involved in this particular part of the pathway. This includes databases such as UniProt, Ensembl and LocusLink as well as the hierarchy of the complex within the relevant (in this case DNA repair) pathway.

Follow the link for the complex in the nucleoplasm. This further details the location of the complex and its individual subunits. Database cross reference links are available as icons beside each subunit.

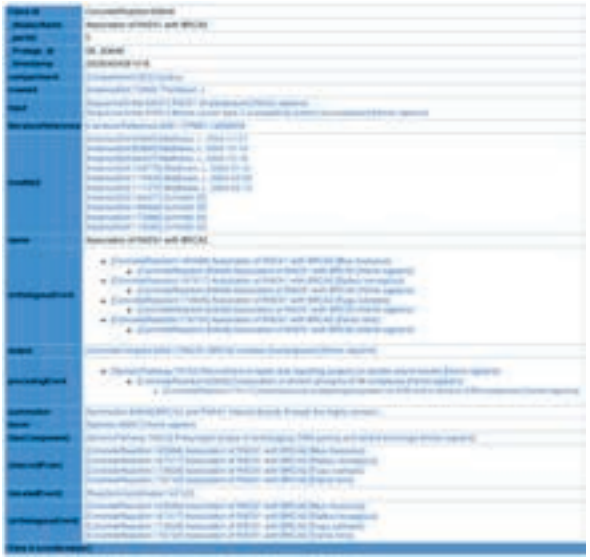
Move down the hierarchy and follow the bottom link to the complex association process. The hierarchy is now displayed on the left hand side, with details of the association on the right.

Follow the first link to the Matthews, L. reference. This links to one of the biological experts who contributed to the information in the Reactome database. Each link may be followed to view detailed information on each pathway process created by this expert.

Details of the input and output of the reaction are displayed together with a note on how this is achieved and the pathway reaction immediately preceding it. Taxonomy is also mentioned as well as the locus within the cell where association takes place.
A list of equivalent events in other organisms is displayed, but these events have only been inferred on the basis of sequence similarity and not experimentally verified.

The reference is split into several links. The PubMed ID is located in parentheses after the journal information.

At the bottom of the page is a link to view the information in the Instance Browser. Previously we had been in the Event Browser. Follow this link.



The Instance Browser lists all the information that Reactome has available for a particular instance in the database. It is representative

of how the data is actually stored. The events browser links it together as a coherent collection of pathway complexes in the context of each event. This is the more user friendly version of the data.

Return to the information page and note the complex requires the association of the RAD51 protein with a repeat region on the BRCA2 protein.

Unfortunately, due to the information in the database, the Breast Cancer Susceptibility protein cannot reveal all that Reactome can offer in terms of event displays. Look to something such as nucleotide metabolism to see a wider range of information.

Announcement

EMBnet Annual General Meeting in Slovakia. Sept.15 - Sept.18, 2005

EMBnet Collaborative Workshop and 19th EMBnet Business Meeting will take place in Smolenice Castle, Slovakia, 15 - 17 September 2005.

Program

We 14 Sept.	arrival
Th 15 Sept.	EMBnet Collaborative Workshop
Fr 16 Sept.	AGM (Part 1)
Sa 17 Sept.	AGM (Part 2)
Su 18 Sept.	departure

More information and the detailed program will be announced later on the web site. `http://www.sk.embnet.org/agm2005/`

The Smolenice Castle

The village of Smolenice with the well-known Smolenice Castle is located at the foot of Malé Karpaty mountains, 60 km northwest of Bratislava. The castle shelters the Congress Centre of the Slovak Academy of Sciences. The representation interior offers interesting views outside. Adjacent English park, nicely integrated into mountain forest edge, offers numerous possibilities for short walks.

First written documents about the existence of Smolenice date back to the 13th century, although its origin is as ancient as the Neo-Feudalism. Several aristocratic families had been the landlords there. In the 15th century the importance of Smolenice increased considerably, because a castle was built there which became the centre and seat of the Smolenice estate. In 1777, Ján Pálffy takes the Smolenice estate as pawn. The Pálffy family did not live at the castle, which had decayed considerably during the life of Krištof III, the last of the Erdödy family – they lacked money for maintenance. The decay was complete during the Napoleon Wars – the main castle building and the tower had burnt down.

Construction of the Smolenice Castle of today had been started early in the 20th century by Jozef Pálffy Jr., the landlord of Smolenice and Dobrá Voda estates. During World War I, the construction was interrupted; provisional adaptation of some rooms was made and archives of the Pálffy family were located there. The construction was not resumed before the end of World War II. In 1945 the Castle became the property of the State; it was taken over by the Slovak National Council who decided to have their summer-house there. The castle was finished and furbished and handed over to the Slovak Academy of Sciences on 26 June 1953 to become a representative place for meetings of scientists from worldwide.

For more information about Congress Centre of Slovak Academy of Sciences check `http://www.kcsmolenice.sav.sk/en/`

Lubos Klucar (klucar@embnet.sk) - organiser

Flower power

Vivienne Baillie Gerritsen

There is not much we have in common with sweet corn yet when one has a closer look at the way an egg cell is fertilized in flowering plants, it is difficult to avoid making comparisons with humans. A long tube – the pollen tube – must make its way into the female gametophyte and release its sperm cells which will ultimately fertilize the egg cell. The great difference however is that a pollen tube has to elongate and travel quite far to perform its business because pollen – unlike sperm – is not mobile. A number of questions arise. What, for instance, guides the pollen tube towards the gametophyte? And how does it know when it has reached it? Scientists have found the beginning of an answer in the form of a small protein: *Zea mays* EGG APPARATUS 1 or ZmEA1, which has a definite role in pollen tube guidance and orientation in the phases preceding egg cell fertilization.

Fig. 1 *Zea mays* (Maize or sweet corn)

Maize has been around for thousands of years. It first appeared in Mexico and became an integrated part of their culture. In fact, many tales relate the origins of maize. According to an Aztec myth, men were first molded out of clay but lacking a favorable outcome these creations became birds and deer. The second time around, men were whittled out of wood but that too proved fruitless so, on their third attempt, the Gods fashioned men from the dough of maize and mixed it with their own blood to create present day humans. In fact, the Nahuatl – the language spoken by the Aztecs – word for maize dough is 'toneuhcayotl' meaning 'our flesh'. From the 1500s onwards, the Spaniards participated in distributing it around the rest of the world, which was not a difficult task since sweet corn grows in almost any climate, and at any altitude.

Let us give a brief reminder of how flowering plants such as maize (*Zea mays*) are fertilized because, unless you are a keen botanist, your memory – like that of the author of this paper – may well need some refreshing. In a nutshell, within the female reproductive organ, or ovary, resides the ovule surrounded by tissues. At one end of the ovule are to be found three cells: two synergid cells and the egg cell itself. It is the egg cell, which once fertilized by pollen, will give rise to an embryo and ultimately a new plant. What happens is that a number of pollen tubes migrate towards the ovule but only one will actually enter – by way of a region termed the micropylar opening – and release its sperm cells which will fertilize the egg cell.

ZmEA1 is a transmembrane protein, barely 94 amino acids long. What scientists have discovered is that its presence is concomitant with the maturing of the egg apparatus and is found both in the synergids and the egg cell itself. It is also found in the region of the micropylar opening just where the pollen tube is supposed to make its appearance. Interestingly, when ZmEA1 is neutralized, pollen tubes do migrate towards the ovule but turn away a little perplexed as though they hadn't found the opening. In wild type maize, a number of pollen tubes head for the ovule but only one changes its orientation markedly as it reaches the

micropylar opening and actually enters. So it does seem apparent that ZmEA1 has a major role in pollen tube guidance.

No one knows how though. In the beginning of the 20th century, scientists had already assumed that pollen tube guidance had something to do with the egg apparatus. Towards the middle of the century, the existence of chemical substances produced by the synergid cells had been suggested. Almost half a century later, fingers are pointing in the direction of a protein: ZmEA1. Though, whether ZmEA1 is a signal itself and binds to receptors, in a secreted form, on the pollen tube tip or whether it somehow triggers off signals in the tissues surrounding the micropylar opening which in turn bind to receptors on the pollen tube…only time will tell. What is sure is that ZmEA1 is largely expressed in the mature egg apparatus during pollen tube guidance and its expression is halted and any secreted form rapidly degraded once fertilization has taken place thus preventing polyspermy.

It's a big job for such a small polypeptide. Here is a little fellow which has its say in a plant's fertility and may well be part of the mechanism which provides Nature with the concept of 'species barrier'. If two species are genetically close, say, but cannot produce offspring, one of the reasons could be a difference in pollen tube guidance. Likewise, plant sterility can be addressed more subtly if molecules such as ZmEA1 and their function are better understood.

To what end do scientists undertake such studies? In an attempt to understand the reproductive process of flowering plants naturally but also in order to acquire a more subtle view on how sterility can arise in the event of crossbreeding. Maize is the staple diet of many populations. Understanding the fine-tuned processes involved in the 'species barrier' concept could provide the knowledge and wherewithal to cross species which otherwise could not be interbred. In turn, interbreeding new species could help in the fight against famine by providing crops – which are not only sturdy and productive but also nourishing – in countries such as Africa where, still today, 33 million children go to bed hungry every night.

Cross-references to Swiss-Prot

Egg apparatus-1 protein (ZmEA1), *Zea mays* (Maize) : Q5G8Z3

References

1. Marton M.L., Cordts S., Broadhvest J., Dresselhaus T.
Micropylar pollen tube guidance
Science 307:573-576(2005)
PMID: 15681383

2. Salvador R.J.
Maize
Iowa State University, Agronomy Department
http://maize.agron.iastate.edu/maizearticle.html

Tintin's blight

Vivienne Baillie Gerritsen

Tintin never grew up. Readers followed his travels around the world for almost fifty years and yet the Belgian journalist showed no signs of aging whatsoever. No grey hair, no wrinkles, no loss of stamina. He never changed style either; he never seemed to tire of his knickerbockers nor of his cranial crest. But that is beside the point… How can a human span a lifetime looking as though he never grew older than the age of fifteen? Hypogonadotropic hypogonadism or HH say some. HH is a condition in which the subject who is inflicted with it never reaches puberty. Typically, in a man, this would mean that he shows no signs of becoming one, i.e. in growing facial hair for example or being the proud owner of a mature reproductive system. Tintin never took his pants down but any of his readers know that he certainly never showed signs of growing a beard. HH in a man is caused when the regulation of the male hormone, testosterone, is deficient. And we now know of one protein which seems to have a key role in such a regulation: the KiSS-1 receptor.

How does a condition like HH occur in the first place? It all has to do with hormones. Naturally. Puberty is triggered off by a flush of hormones, amongst which are those that cause a young boy to grow hair on his chin or a girl to menstruate. If such hormones are hindered in any way, a young boy will never know the joys of a beard, or a razor. At puberty, a part of our central nervous system, the hypothalamus, releases a hormone known as 'gonadotropin-releasing hormone' which itself causes the release of two other hormones – follicle-stimulating hormone (FSH) and luteinizing hormone (LH) – which are both produced in a small gland – the pituitary gland – situated just beneath the hypothalamus. In young boys, FSH and LH ultimately cause the production of testosterone. If the hypothalamus-pituitary gland pathway is hindered in any way, the production of the male hormone will be also and puberty will not take on its proper course.

A number of instances can cause hypogonadotropic hypogonadism. Key hormone receptors involved in the pathway can be deficient or missing due to surgery, head trauma or a congenital disease for instance. It was the study of congenital HH which led to the discovery of the KiSS-1 receptor which proved to have a direct role in the proper development of the male and female gonads. The KiSS-1 receptor is found both in the hypothalamus and the pituitary gland and is a G protein-coupled receptor; that is to say it belongs to the very large family of transmembrane receptors whose role is to transduce a signal. Its ligand is the neuropeptide KiSS-1 and together they are involved in the regulation of gonadotropin secretion. As mentioned above, gonadotropin ultimately stimulates the production of testosterone in males. Take away the KiSS-1 receptor and Tintin becomes Peter Pan.

Fig.1 Tintin in the Land of Black Gold, Hergé, Ed. Casterman

Could it be that the famous reporter was indeed inflicted with HH? Perhaps. Could it have been congenital? No one, save his creator Hergé, could answer that. However, what we do know is that the knickerbockered reporter was beaten about the head a fair number of times during his

travels around the world. A Canadian team, made up of an Associate Professor with the Faculty of Medicine at the University of Sherbrooke in Quebec and his two young sons, scanned all of Tintin's adventures for blows which had been made to the poor boy's head and they discovered that Tintin was subjected to 43 head traumas resulting from such incidents as colliding with tree trunks, falling on ice or receiving a blow from a whisky bottle. The severity of the injuries was judged by the number of candles or stars which revolved above Tintin's head and the number of frames it took him to recover.

Puberty can be arrested following even minor head traumas in children. There can be no loss of consciousness, no important external injuries or subsequent neurological effects, yet if the stalk which leads from the hypothalamus to the pituitary gland is somehow severed, the natural process of puberty cannot take place because hormonal communication has been lost. So, though Tintin visibly suffered from his injuries and most of them involved loss of consciousness, he didn't need all 43 accidents to acquire HH… A mild blow on his head would have been sufficient.

G protein-coupled receptors are diverse and many. And since they are involved in signaling, the pharmaceutical industry has made great use of them by using them as drug targets. And it has been quite successful. Not only is the KiSS-1 receptor involved in neurological processes – such as the onset of puberty for instance – but its ligand, the KISS-1 peptide was previously shown to prevent metastasis in melanoma cells. Such findings suggest that the KiSS-1 receptor and its ligand may be the bearers of essential roles both in the central nervous system and in tumor biology. And, as a consequence, they could provide valuable targets in the future for novel therapies in the fields of oncology as well as neurology.

Cross-references to Swiss-Prot

KiSS-1 receptor, *Homo sapiens* (human) : Q969F8
Metastasis-suppressor KiSS-1, *Homo sapiens* (human) : Q15726

References

1. Cyr A., Cyr L.-O., Cyr C.
Acquired growth hormone deficiency and hypogonadotropic hypogonadism in a subject with repeated head trauma, or Tintin goes to the neurologist
CMAJ 171:1433-1434(2004)
PMID: 15583175

2. Muir A.I., Chamberlain L., Elshourbagy N.A., Michalovich D., Moore D.J., Calamari A., Szekeres P.G., Sarau H.M., Chambers J.K., Murdock P., Steplewski K., Shabon U., Miller J.E., Middleton S.E., Darker J.G., Larminie C.G., Wilson S., Bergsma D.J., Emson P., Faull R., Philpott K.L., Harrison D.C.
AXOR12, a novel human G protein-coupled receptor, activated by the peptide KiSS-1
J. Biol. Chem. 276:28969-28975(2001)
PMID: 11387329

3. de Roux N., Genin E., Carel J.-C., Matsuda F., Chaussain J.-L.
Hypogonadotropic hypogonadism due to loss of function of the KiSS1-derived peptide receptor GPR54
Proc. Natl. Acad. Sci. USA 100:10972-10976(2003)
PMID: 12944565

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

Brazil
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 consists 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 deadlines for the next issues:
August 15, 2005
October 31, 2005