embnet.news

Volume 7 Nr1 (ISSN1023-4144) April 30, 2000

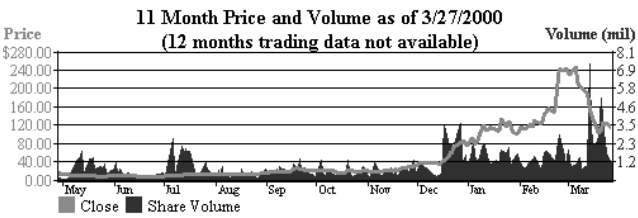
Editorial

What a difference a year makes

In Vol6_1 of EMBnet news the following statement was made :

"In the 90´s it was pop to be in biotech. However when you look at how well these new start-up companies have performed on the stock market, their record is less than impressive."

In the intervening year there has been a remarkable change... biotech stocks suddenly took over from tech stocks as the wonder stocks on the NASDAQ. Witness the meteoric rise of Celera in the past 11 months. It soared from about $20 to over $240 at its peak. It was the proverbial "ten bagger". That was until about a month ago, when at the beginning of March, Tony Blair and Bill Clinton made a joint statement. Now if their joint statement had only been about the fact that they had never inhaled that would not have mattered much, but their statement was about the patenting of the Human Genome, and immediately the share price in Celera dropped off dramatically falling as low as $85.



To make matters worse biotech stocks that had nothing to do with genomics were also caught up in the downward spiral. Even an IT company with "Bio" in its name (but not at all active in biotechnology) saw a drop in share price. It was as though Politics/Science/Economics formed a Bermuda triangle into which profits disappeared, and no-one could explain why. In a recent article Sci/Tech section of the BBC web site Dr David Whitehouse comments :

"It is clear from the reaction of some publicly-funded scientists to the Clinton-Blair statement that they resent the idea that big business is involved in decoding the human genome, the basic instructions for life. They would rather the corporations stay out of it."

Celera on the other hand would like to see its investments pay off. Craig Venter in a recent interview with Bloomberg TV stated that Celera would be publishing the complete drosophila genome within eight days and true to his word you can now get a CD of the Drosophila genome from the Celera website. You can also browse the fruit fly genome at the NCBI, or join the drosophila research community at Flybase in Indiana, who gratefully acknowledge Celera´s release of the Drosophila genome data.

You picks your horse and puts your money down.

He further added that Celera is not a pharmaceutical company but is an information company and intends to make its money by making the data readily accessible. When you look at the software that Celera is using to search and analyse the data, Celera´s Discovery System bears more than a passing resemblance to SRS which has been running at EMBnet nodes for a good few years now.

John Sulston, director of the UK-based Sanger Centre in a recent interview with the BBC states that "Data release is good for patients and is also good for business because it means that all businesses can get at the data and not just a few."

In the April 1st (not a joke) issue of the New Scientist John Sulston gave his opinion in an article called Forever Free.It stated,"Private companies have already filed thousands of patent applications for thousands of human DNA sequences, and in most cases there is little understanding of their biological function. If such patents are granted, researchers in both the public and private sectors may be forced to pay to use this genetic information, if they can get it at all." He further added "Every day, we post our latest data on the Internet. This means that scientists are using the information today without legal restriction or financial obligation."

Contents

Indeed if you go the the Sanger Centre web site you will find press releases for the latest completed genomes. In the past month alone you can find the complete genome of the leprosy bacillus, and the meningitis bacterium

The debate over the ownership of sequence data and its commercial implications has been grabbing the headlines for the past month or more and as the BBC article summarises :

"The genetic revolution and its impact on mankind demands the involvement of everyone, not just scientists. But when politicians become involved, or when scientists try to play politics, the debate can become muddled. Fears and even resentments can become as much a part of the debate as logic and scientific fact."In the recent past both Swiss-Prot and the EBI have had their applications for grants turned down by the EU, because of some misguided political decisions. The subsequent uproar has been covered in both Nature and Science.

The BBC web site have made this into a public debate and they have asked the question : "Has the public sector been too slow? Should the private sector be more open with its data? Does it matter who actually gets to the finishing line first?"

If you think that it is evident that both the public and the private sector are somewhat at the mercy of the politicians, then why not have your say in the debate. For at the moment it seems that the politicians giveth and the politicians taketh away.

Here endeth the first lesson.

The Editorial Board

INTERviewNET

Ewan Birney interviewed by Robert Harper

Q1: Who thought up the name EnsEMBL ?

The Ensembl name (notice the lower caps) was a great invention. Notice that end has EMBL, that you can say it without mangling your mouth, it means "together" and sounds vaguely French. It is an ideal name for an international project on the human genome.

We first talked about a database called "humEMBL" being the reannotation of human EMBL, but started to get worried about the fact that we were going to use it for mouse, fish and other organisms. We wanted a name generically for the software. Over tea at the EBI (I was still at the Sanger at the time) we dreamt up the name Ensembl and everyone has been very happy with that ever since. There are still different capitalisations of the name: but we like to encourage diversity.

Q2: Do you collaborate with other organisations in the production of EnsEMBL?

Ensembl is a joint project between the EBI and the Sanger Centre. I head up the EBI team (we are currently 4, but expecting to grow fast) and Tim Hubbard is the counterpart at Sanger. The original software was designed by Michele Clamp and myself when we were both at the Sanger Centre. She has remained at the Sanger Centre and is one of the key leaders in the project. Collaboration is central part of the culture in Ensembl, partly because of this background.

Ensembl is a completely open project - you can see changes to the code base through our anonymous source code server only hours after we make them. On the mailing lists we have people from over 30 different institutions and many of them actively contribute to the design and the software. For example, people from the Berkeley Drosophila Genome Project helped us write and test the GAME-XML dumper which will appear in the next release of the code.

Anyone is welcome to hang out on the mailing lists and listen into to our design process, as well as use the source code themselves. Check out the resources off http://www.ensembl.org/

Q3: EnsEMBL is for the automatic annotation of eukaryotic genomes. Does it work?

It works. Does it work *well* ? From our perspective it works surprisingly well considering how simple the gene prediction is (it is Genscan followed by BLAST searches of the peptides). However you can always do better. In the next release, due mid April, we will have map integration and ability to view and use long pieces of DNA made from small component stretches of DNA. Looking ahead to May we will be drastically improving our gene prediction when there is a known cDNA sequence.

The list of things one would like to do is almost open-ended. We don't expect to see this finished for a long time, and it means that for potential collaborators there are many opportunities to work with us.

Q4: If you had a million pounds to spend on EnsEMBL what would you spend it on?

That is a hard question. Certainly quality people is the first thing to spend money on. We are building a great team at the moment - we have been very lucky so far to have hired some excellent and dedicated bioinformaticians. Getting good people is the main bottle neck at the moment. The compute requirements are large, but using the Sanger Centre computer resources we have a considerable computer resource to tap into. Of course, the more compute time you spend on solving problems the more accurate you can be. Anyone who has run my Wise2 software will know how compute intensive these algorithms can be.

Q5: What brought you from Oxford to Cambridge?... and don't say the bus.

It certainly wasn't the beer! To be honest, I was attracted by Richard Durbin (currently deputy director at the Sanger Centre). I had two options : to stay at Oxford, write software on my own and not really progress or to join the Sanger Centre and learn from people who were really at the sharp end of the problem. I have never regretted my decision.

Moving from the Sanger Centre to the EBI was also a wonderful opportunity to focus on what I feel is important, namely well written, functional and on-time software. The whole Hinxton campus is the world's best place to do bioinformatics and I think that is one of the reasons behind Ensembl's success.

Q6: Do you still do Bionet or is that only for students?

I used to use Bionet 5 years ago, but I don't now. I find the signal to noise ratio very bad. I find that mailing lists, like the ensembl-dev mailing list or the bioperl mailing lists the most productive. Somehow you get fewer transitory posts and more of a community feeling. In some sense I almost have a larger electronic community I am in touch with than physical research contacts, largely due to these mailing lists.

Q7: Why is there no photo of you on the web... only a tiger?

The Calvin and Hobbes picture was on my first home page back at Oxford (it ran on a VMS system). For me it is sort of "home" on the web and I would hate to lose it. I was shocked however to see it reproduced in the EMBL team leader mug shots for potential students! I guess it will have to go, and a nice picture of me in tie will replace it. That will be a sad day however.

Ewan Birney. birney@ebi.ac.uk

Brazil, a new Mecca for genomics?

Historical perspective

High-throughput sequencing is a highly specialised trade, practised in a very limited number of laboratories in the developed world. It can be estimated that a dozen labs are contributing over half the total sequence data currently being deposited in the public databases, with another 50 or so accounting for the bulk of the rest. All of these labs are located in North America, the larger European countries, Australia and Japan. It may thus come as a surprise that the latest entrant in this select club hails from Brazil, and more specifically the state of São Paulo.

São Paulo has a law stating that 1% of the tax revenue collected by the state has to be given to an independent agency that supports scientific research, known as FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo). As São Paulo is the richest state in Brazil, this amounts to a considerable amount of money (USD 250 Mio in 1998). By law, FAPESP is also forbidden to spend more than 5% of its money on administrative costs. The combination of ample funding and political independence gives the Foundation a lot of freedom to develop innovative scientific programs.

In 1997, FAPESP decided that Brazil should not miss out on the scientific and economic opportunities that can be derived from genome sequencing, and should be able to produce its own data, analyse them, and use the results for local scientific projects. To start off, it was decided that a good target organism should be bacterial, and of interest to the local economy. The agency settled on Xylella fastidiosa, a bacterium that infects orange trees, a major source of income in São Paulo, and causes Citrus Variegated Chlorosis. This choice also brought in additional funding from the citrus growers' association (Fundecitrus).

ONSA

A major goal of this first genome project was to bring sequencing technology to as many laboratories as possible, thus propelling them into the genome age. Therefore, the concept of setting up a single sequencing centre was rejected from the start. Instead, bids were put up for laboratories interested in participating in the project, and those that were selected received equipment (ABI370 sequencers), reagents, and ample technical advice. In total, 30 labs were selected for the Xylella project, dispersed geographically throughout the state of São Paulo. In addition to the sequencing labs, the project steering committee designated a DNA co-ordinator (for the handling and distribution of clones) and a bioinformatics centre. The bioinformatics group, located at the University of Campinas (about 80 km from São Paulo), was made responsible for all of the data handling, from base calling to final assembly verification. The sequencing labs submitted trace files only, and were paid on the basis of the amount of non-vector, high-quality sequences (based on phred scores) that could be extracted from their data. The entire process was automated using Web pages, and enabled the bioinformatics group to keep very close tabs on the daily progress of the project as a whole.

The sequencing consortium that emerged from the Xylella project came to be known as ONSA, the Organisation for Nucleotide Sequencing and Analysis. It is not coincidental that *onça* is the Brazilian name for the jaguar, a slightly smaller but more nimble feline than TIGR or LION. Starting from scratch, i.e. with labs that had never done any sequencing before, ONSA managed to sequence over 90% of the Xylella genome in less than a year. As usual, gap closing and finishing took another year, but the genome is now completed, and was presented at a plant pathogen conference in February 2000. It is notable that this was the first plant pathogen whose genome was sequenced. The consortium is currently sequencing another citrus pathogen, Xanthomonas axonopodis pv. Citri, which causes Citrus Canker.

ORESTES

Once the Xylella project was well under way, and the consortium had proven that it could produce sequence data quickly and efficiently, Andy Simpson of the São Paulo Branch of the Ludwig Institute (the Xylella DNA coordinator) started thinking of more ambitious projects to tackle. He had developed a technique for cDNA cloning using low-stringency PCR and applied it to gene discovery in Schistosoma mansonii. He reasoned that the technique, dubbed ORESTES (for Open Reading frame EST Sequencing), could be applied on a large scale to the generation of novel human ESTs. The novelty of the ORESTES approach lies in its use of defined sequence primers, used at very low stringency, to generate a large number of low-complexity cDNA libraries. Because the probabilities of priming the first and second cDNA strands are distributed randomly on their respective templates, the technique preferentially generates clones coming from the central portions of mRNAs, which are underrepresented in current EST collections.

Andy proposed an ambitious new project, aiming to produce 1 million human EST sequences using the ORESTES technique, and thus to add substantially to our knowledge about the transcriptome. The Human Cancer Genome Project (HCGP) was funded jointly by FAPESP and the Ludwig Institute for Cancer Research, for a total of USD 10 Mio over two years. It uses some of the existing ONSA expertise, but also adds a number of new groups with a stronger interest in human biology and medicine. Technically, it is also different: six megabase high-throughput capillary sequencers have been placed in sequencing centers, each of which is fed by a consortium of geographically clustered labs. Because of the higher complexity of the project, coordinators have been designated for tissue sampling, RNA preparation, and library construction, in addition to the overall project and sequencing coordination.

The HCGP has been a resounding success. The megabase machines were put in production in the summer of 1999, and, after the usual teething pains, have been running at full blast since last October. They have already produced over 200'000 new EST sequences, about 25% of which do not match any sequences in the public EST databases, and about half of which contain novel sequence information. The collection was used to update the annotation of chromosome 22, where it identified about 100 genes for which there was no previous experimental evidence. It is expected that the project will reach the million mark by the end of the current year; it is already producing data at a much higher rate than NCI's CGAP project. It is also sampling a number of cancer types that were not included in CGAP. The sequences are being deposited in the public databases, as the two funding institutions had pledged to do.

The bioinformatics of the HCGP is being handled by a new group at the Ludwig Institute in São Paulo, headed by Sandro de Souza. The group has done an excellent job not only in data management, but also in annotating and databasing the sequences, thus allowing project participants to quickly find sequences based on a number of criteria, including library of origin, annotation class, similarity to sequences in the public databases, etc. They are also integrating the HCGP sequences in contigs with ESTs already in the public databases, and using their data to complete the annotation of emerging human genome sequences. It is very likely that the São Paulo group will become a member of the Ensembl annotation initiative.

Credit

In two years, Brazil (or at least São Paulo state) has gone from essentially nothing to being one of the larger producers of sequence data in the world. It has done so not by investing massively in a large sequencing facility, but by bringing together a large number of individual labs, many of which are already using these new data and know-how in their own research. In this way, the genome projects have already had a major impact on Brazilian science.

The world has not really taken notice yet, but I would bet that within another year or two ONSA and the HCGP will have achieved the same recognition as TIGR and CGAP.

Completed genomes at the European Bioinformatics Institute

Peter Sterk

The first completed genomes from viruses, phages and organelles were deposited into the EMBL Database in the early 1980's. As sequencing technology improved, the number of completed genomes increased, and in the mid 1990's the first bacterial and eukaryotic genomes were completed, marking the start of the genome era. With many sequencing projects in progress, the need has arisen to provide easy and up-to-date access to these genomes. We have set up a genomes web server (http://www.ebi.ac.uk/genomes) from which at the time of this writing nearly 1000 completed genomes can be retrieved. The completed genomes have been divided into a number of groups: viroids, viruses, phages, plasmids, organelles, archae, bacteria and eukaryota. Where many different strains of a species existed - e.g. hiv, we have chosen to represent one genome rather than listing many to keep the lists manageable. We also provide translations of coding regions in FastA format, and in collaboration with SWISS-PROT, we hope to include the proteomes and access to SWISS-PROT entries in due course.

The archaea, bacteria and eukaryota, are represented in the EMBL Database as a series of segments due to the fact that a maximum sequence length of 350 kbase was agreed among the collaborating databases DDBJ, GenBank and EMBL. For each of these genomes, we attempt to provide a CONstructed database entry which contains information on how one could assemble the complete genome from individual segments. For example, the CON entry for the bacterium Rickettsia prowazekii has the following join statement on its CO(nstruct) lines:

join(AJ235270:1..282610,AJ235271:51..312430,
AJ235272:51..279110,AJ235273:51..237523)

The complete genome can be built from four EMBL entries, AJ235270-AJ235273, which have a 50 base overlap. These segments can be retrieved from the web server, as well as the CON entry and the complete genome with annotation.

In March this year we added the complete genome of the fruitfly Drosophila melanogaster. As we are expecting a large number of updates in the coming year, we are not yet able to represent the fruitfly chromosomes as CON entries. However, the segments have been grouped according to chromosomes and can be retrieved from the server. We intend to do the same for other eukaryotic genomes/chromosomes when the completed sequences become available.

Last but not least, we are working hard to improve and expand services in the genome area. Recently, a number of services have been added to the existing EBI search tools, e.g. the Genome and Proteome Fasta3 Searches (http://www2.ebi.ac.uk/fasta3/genomes.html), providing searches for and in complete genomes and proteomes. The Genome Monitoring Table (http://www.ebi.ac.uk/~sterk/genome-MOT/) which monitors the progress of a number of eukaryotic genome projects has had a major overhaul and provides new functionality. The joint EBI-Sanger Centre Ensembl project (http://www.ensembl.org) for automatic annotation of eukaryotic has seen its first release with the next one being due imminently.

UMBER

News from upt' North

In March 1999, the UCL Specialist Node transferred to the University of Manchester. The transition was as smooth as could be expected, given the inevitable hassles associated with moving people and machinery around the country. We endeavoured to keep our Web services running as seamlessly as possible, essentially by effecting the `virtual' move in a stepwise manner. Overall, I think it's fair to say that this last year has been peppered with individual heroics and institutional incompetence. Read all about it....

PRINTS

During the year, Steffen Moeller incorporated PRINTS into his EDITtoTREMBLpackage as part of the EBI's automatic protocol for annotating TrEMBL sequences, and Rodrigo Lopez made the database available for searching on EBI's public servers. The database was released according to its quarterly deadlines and, in addition, an important fifth release was made. This represented a complete automatic update of the database on the current version of SWISS-PROT/TrEMBL.

The change of underlying database (formerly OWL) was necessary for 2 reasons:
i) Leeds has not provided an OWL update since SWISS-PROT partially withdrew from the public domain;
ii) PRINTS joined forces with PROSITE and Pfam in the InterPro project, which required that all partners use the same underlying data source. The conversion was a massive undertaking and, though formally complete, still requires manual attention for fingerprints that failed to update well.

InterPro

As a result of the combined efforts of the SWISS-PROT/

TrEMBL, PROSITE, Pfam, PRINTS and EBI support teams, March 2000 saw the first official release of InterPro. InterPro is a European project, launched in autumn 1998, to integrate the commonly-used pattern databases within a centralised protein family documentation resource. The primary goals of InterPro are to reduce duplication of effort in the time-consuming process of annotation and to provide a resource that will facilitate genome/proteome analysis. Its first major use has been in the comparative genome analysis of D.melanogaster, C.elegans and S.cerevisiae, the results of which were recently published in Science.

CINEMA

CINEMA was finally extricated from the Open Molecule Foundation, and the source code released under a GNU license. After a monumental battle to persuade Manchester to accept funds from EMBnet (which, incredibly, ultimately required Mr.Chairman to visit in person to deal directly with the administrators !), theEMBOSS-CINEMA project was finally able to get off the ground (the day before the EMBnet core grant terminated !). Thus, CINEMA is at last undergoing the redesign promised more than a year ago. The complete old version still runs at UCL, and a version without the 3D viewer is currently running at Manchester.

BioActivity

The BioActivity Web practical saw major revisions in order to link it in with an introductory bioinformatics text book. The link works in two directions : the book provides background material for the concepts, tools and resources used in the practical, setting the scene for one of the later chapters (i.e., a worked example through each of the practical steps) ; in turn, the practical Web pages provide on-line excerpts from the book, supplemented with colour pictures where appropriate. BioActivity was tested at the Gulbenkian Institute on a course hosted by Pedro Fernandes in November. Sadly, we never found out how well the book and practical work together, because the publishers delivered the books the day after the course had finished.

EMBER

And finally, EMBER, the grant submitted to the EC to support the design of a European Molecular Biology Educational Resource. This involved much hard work from a lot of people and suffered various trials and setbacks - not least, the last-minute discovery that the Commission does not support overheads under the initiative to which we applied, and only supports 50% of other costs. The proposal was accordingly "streamlined" but, even in the event of success (which looks doubtful), word is that the budget will be cut by a further 46%. Moving swiftly on (before too many colourful metaphors adorn the pages of EMBnet news), I would like to thank everyone who contributed to EMBER. For better or worse, it was a good effort.

University of Manchester: Bioinformatics Education and Research - UMBER - T.K.Attwood, School of Biological Sciences, The University of Manchester, Oxford Road, Manchester M13 9PT

Bio2000 - a biotech odyssey

Tim Littlejohn, Founder & Chief Scientist, eBioinformatics Head, Australian National Genomic Information Service Email: tim@ebioinformatics.com Web site: http://ebioinformatics.com

I have never been to a conference before I where had to force my way past police to get in - but I suppose there had to be a first time, and this was it.

I've been to hundreds of conferences. They get to be much the same after a while. As a kid I used to sneak in, curious about science. As a young researcher I went to learn about the process of science. As a scientist I went to hear about the latest developments. As a service provider, I go to chair sessions on bioinformatics and man trade displays. This meeting (Bio2000, the international Biotechnology Associations huge annual meeting www.bio.org/events/2000/bio2000.html) was to set the stage for all the ones that follow.

I arrived in hotel in Boston USA late on the evening of Monday 27 March - tired after a 24 hour trek from Sydney Australia. Man that is a dreadful flight. Four anti-inflammatory eye drops later I tuned into the TV and learned that protesters had been barricading the conference and had dumped a truckload of soy-beans (presumably genetically engineered ones) on the entrance steps. Amused, I donned my battle dress (the eBioinformatics team shirt) and ran outside to grab a taxi to the convention. On the way out the reception desk gave me a phone message to call David from eBioinformatics on the Banana, which I dutifully did (the banana is code name for one of our mobile phones).

The taxi dumped me into a scene that could have been from an Arnold Schwarzenegger movie : helicopters hovering overhead, searchlights sweeping the night sky, police and fire trucks on every corner, traffic mayhem. I made my way to the entrance and bumped into two heavily armed policemen. Conference registration is closed ? No entrance at all this evening ? But what about my colleagues locked inside ?

I raced outside and called David on the banana again, this time from a public phone on the street corner opposite. Ten minutes and some social engineering later I was inside wearing a guest pass David had kept tucked up his sleeve for such occasions.

The rest of the meeting was just as exciting. The conference attracted the usual collection of the hundreds of the worlds top names in biotechnology. As the protesters pointed out, biotech is back, and despite some rockiness, NASDAQ proves this. With multiple simultaneous back to back sessions and seemingly a thousand trade displays, it was hard to know what to do first. All my time was occupied by one of three things so I'll talk about those - presenting in sessions, manning the eBioinformatics trade display, and talking to suppliers.

First up I chaired the bioinformatics session entitled Bioinformatics - the brain of Biotech. The session was remarkably well attended - thanks to the quality of the presenters and the hot topic (bioinformatics - the use of computers in the biotechnology discovery process). We must have been doing something right as the session was picked up the next day by biotech news wires such as BioSpace (www.biospace.com/articles/bio_bioinformatics.cfm).

Andrej Sali from Rockefeller University spoke first about "Comparative protein structure modelling of genes and genomes", followed by Tony Kerlavage from Celera Genomics on " Biological Knowledge and the Future of the Life Science Industry". Bruno Gaeta from eBioinformatics spoke on "BioNavigator : an integrated web front-end for bioinformatics analysis" and Jim Ostell from NCBI (National Centre for Biotechnology Information, USA) finished with a talk on "Genomes - Cross-roads for Data". It was a session that spanned molecules to genomes, user interfaces and integration for academics through to big pharmaceutical companies, small and huge corporate entities, medium and large government organisations. In one way it was a whirlwind tour and random sample of where the human genome project is and where it is going. It also highlighted how computers are pivotal to this and other biotech initiatives. "All experiments start out, go through, or end up with bioinformatics", according to the session's chairman.

Next up I was very lucky to be able to participate in the "Bioportals" session. In this session a panel presented their views on what a Bioportal is. Panellists were Neil de Crescenzo (Chemdex, a business to business ecommerce company), Mark Edwards (Recombinant Capital, investors), Bruno Larvol (Cognia, a bioinformatics company), Karen Ferrell (Healtheatre, a video content company), Jerry Williamson (Techex, a technology transfer company), Tim Littlejohn (eBioinformatics, a bioinformatics application service provider).

What struck me most about both these sessions was their popularity - the bioinformatics session was full with three times as many people "denied access" left standing, frustrated, outside. The Bioportals session was the same. The company selling audio tapes of the bioinformatics session sold out completely - and they didn't tape the Bioportals session (I bet they are kicking themselves for that). Fortunately Healtheatre made a video of the session so we can expect to see that streaming down the net soon.

The rest of the meeting was spent on our trade display answering questions about BioNavigator (www.bionavigator.com) or deep in discussions with suppliers and partners. This is a great time to be in a bioinformatics company - biotech and the internet are booming, and eBioinformatics as an internet based bioinformatics application service provider (ASP) is well placed to make the most of this in the biotech arena. Having a ten year pedigree (sprung out of the Australian grown ANGIS bioinformatics service - www.angis.org.au - established in 1991) is one of eBioinformatics' competitive strengths. Indeed listening to the audio tapes from the session Instant Information - how the internet is changing biotechnology indicates that the company's approach is right on track. The need for robust, well integrated, broad, affordable bioinformatics systems has never been greater. The growth of commercial bioinformatics ASPs is a testimony to this, and the history of ANGIS and its sister organisations in academia (many of which are members of the EMBnet consortium - www.embnet.org) shows this concept is here to stay. eBioinformatics is not alone, as other groups such as Lion Bioscience (www.lionbioscience.com) have spun out of EMBnet. eBioinformatics is growing through partnerships with suppliers like Lion and many others - the breadth of software and databases on the BioNavigator site is testimony to this.

The Bio2000 meeting lasted three days. Although I only went to two sessions out of dozens, these few days helped confirm in my mind that biotech is back with a vengeance and that bioinformatics will be a key part of the growth of this, the third, industrial revolution. As I packed up my hotel room on the last night in preparation for my flight back to Sydney I tuned into the TV one last time. There, almost prophetically, was a documentary on the gold rush in the Canadian Klondike. Biotech in the year 2000 must feel exactly like the gold rushes of the previous millennia. I felt like I had just spent three days on the gold fields. It's a privilege to be able to provide the picks and shovels to the miners in this goldrush.

EBI in the year 2000

The EBI has seen a very busy year go by. The new millennium has brought about new challenges and opportunities. In the field of External Services EBI has now a totally dedicated cluster of Silicon Graphics machines which is composed of a single Origin 2000 with 16 R12K and 4 Origin 200's with 4 R12K each.

These machines are exclusively handling the bulk of requests from external users including SRS queries. EBI is handling more than 20K request for searches per day (a figure which is currently doubling up every 4 months !). If we include our own internal/production searches we total more than 150K searches/day. As you can well imagine, our hosts are at 97% utilisation 24/24 hours.

More services have been added in the last few months. These include the new MPsrch which currently runs on Compaq EV6 & EV5-6's and SGI R12K's. There is an increase in the demand for fast searches but also, more sensitive methods. The EBI strives to provide as wide as possible a choice of these in order to satisfy the search requirements of most (if not all !) of our users. Altogether, EBI currently offers over 100 sequence databases in its searches. These range from the EMBL database, Completed Genomes and specialised databases including IMGT, EDG, Parasites, etc. Please refer to http://www.ebi.ac.uk/Databases and http://www.ebi.ac.uk/Tools for further details.

The requirement to provide searches for and in complete genomes and proteomes is being addressed and although we may seem to be slow at this, we are striving to provide a comprehensive set of services. An excellent pointer to the type of ongoing work carried out these days at EBI and its collaborators is EnsEMBL http://www.ensembl.org/ and Completed Genomes http://www.ebi.ac.uk/genomes/

Another exciting activity currently in progress is the clustering of EST's in a project called the EuroGeneIndexes (http://corba.ebi.ac.uk/EST/egi.html). This one has yielded consensus sequences derived from EST belonging to various species for which there are substantial quantities of sequences. These include mouse, rat, cattle, pig, chicken, etc. A blast server to search these sets can be found on http://www.ebi.ac.uk/EGIblast2/. More species are being added to this system and hopefully, very soon now, human consensus sequences will be available !

Another important development is the release of the much awaited InterPro project. This one aims at collating together the most important protein motifs and fingerprint databases (prosite, pfam and prints). This effort has resulted in the most comprehensive database of functional sites and domains which can be searched from EBI either interactively or via email using an array of pattern recognition methods which include FingerPrintScan, ppsearch, pfscan, HMM and SW, the later two on a TimeLogic DeCypher machine. The home page for this project is on : http://interpro.ebi.ac.uk/.

Genome Events

April, 2000
27 April 2000 - 30 April 2000 *Conference*
Bioinformatics 2000 - Organised by Chair: S. Brunak and A. Krogh, Center for Biological Sequence Analysis, Denmark, venue Hotel Marienlyst, Elsinore, Denmark.

May, 2000
11 May 2000 - 13 May 2000 *Conference*
Computational Challenges in the Post-Genomic Age - Organised by Sun Microsystems,San Diego Supercomputing Center,Pacific NW National Laboratory, venue San Francisco, CA.

June, 2000
1 June 2000 - 2 June 2000 *Conference*
Genome Based Gene Structure Determination - Organised by EMBL-European Bioinformatics Institute, Hinxton, near Cambridge, UK, venue Wellcome Trust Genome Campus, Hinxton, near Cambridge, UK.

19 June 2000 - 21 June 2000 *Workshop*
Protein Domain Workshop 2000 - Organised by EMBL-EBI and the Sanger Centre, Cambridge, UK, venue Wellcome Trust Genome Campus, Cambridge, UK.

July, 2000

16 July 2000 - 20 July 2000 *Meeting*
18th International Congress of Biochemistry and Molecular Biology Beyond the Genome - Organised by IUBMB, venue Birmingham, UK.

August, 2000

7 August 2000 - 14 August 2000 *Conference*
The Second International Conference on Bioinformatics of Genome Regulation and Structure - Organised by Institute of Cytology and Genetics, Novosibirsk, Russia, venue Novosibirsk, Russia.

20 August 2000 - 23 August 2000 *Conference*
ISMB 2000 - Organised by Philip Bourne , Chair & Michael Gribskov , Co-chair, venue La Jolla, California.

September, 2000
1 September 2000 - 1 August 2003 *Training*
DIMACS Special Focus on Computational and Molecular

Biology - Organised by DIMACS, Center for Discrete Mathematics and Theor. Comp.Sci, venue DIMACS, Center for Discrete Mathematics and Theor. Comp.Sci.

3 September 2000 - 8 September 2000 *Conference*
BioTechnology 2000 (includes: 2nd European Congress on Applied Genome Research) - Organised by German Society of Chemical Apparatus, Chemical Engineering and Biotechnology, venue Berlin, Germany.

October, 2000

5 October 2000 - 7 October 2000 *Conference*
German Conference on Bioinformatics (GCB2000) - Organised by Erich Bornberg-Bauer, EML, Heidelberg, venue University of Heidelberg, Germany.

January, 2001

20 January 2001 - 25 January 2001 *Conference*
Molecular Helminthology: An Integrated Approach - Organised by Philip T.LoVerde, Christine Li, Rick M.Maizels and Timothy G.Geary, venue Sagebrush Inn, Taos, New Mexico.

25 January 2001 - 30 January 2001 *Conference*
Molecular and Cellular Aspects of Tuberculosis Research in the Post Genome Era - Organised by Gilla Kaplan, Stewart Cole and Patrick J.Brennan, venue Sagebrush Inn, Taos, New Mexico.

February, 2001

7 February 2001 - 13 February 2001 *Conference*
Bacterial Chromosomes - Organised by Susan Gottesman, Nancy Kleckner and John R.Roth, venue Hilton of Santa Fe, Santa Fe, New Mexico.

EMBnet Node News

News from China Node

Jingchu Luo, Centre of Bioinfirmatics, Peking University, luojc@pku.edu.cn

The network infrastructure has been expanding rapidly in China recently. China Education and Research Network (CERNET), one of the major academic networks connecting most of the universities and research institutions has been expanded its bandwith to 155MB between the capital and some big cities such as Shanghai, Wuhan, Guangszhou and Nanjing. The bandwidth of network link between CERNet and ChinaNet, the major commercial network of China, has also been expended to 155MB. Funded by the 985 program which was initiated on the Centenary of Peking University in May 1998, an area subnet domain at the Yan Bei Yuan, the apartment area with hundreds families of university staffs, was setup on 2 Feb 2000. This subnet is connected to the university campus network via fibre cables. Each family now has a unique IP address on their home computers.

In the meantime, scientific network resource has also been increased all over the country. The China Academic Library and Information System (CALIS) consisting of libraries of Peking University and other universities started to operate in 1999. In the field of biological research, the China Human Genome Project is now becoming one of the biggest national research programmes. The North and South Centres of Human Genome Research located in Beijing and Shanghai respectively have been established in 1998. One of the largest sequencing labs has been setup at the Human Genome Centre at the Institute of Genetics in 1999. Joining in the International Human Genome Program in1999, China is now sequencing the 1% of the full human genome.

With the help of EMBnet colleagues, we have gained some significant progress in the past year. More than 60 databases were installed in the database query system SRS. In addition to various mirrors such as GDB and TRASFAC which have been already running, we are installing other mirrors including ExPASy from Switzerland, IMGT from France and Protein Prediction from the US. We have installed two twin CPU SUN servers (1 e3500 and 1 e250) last summer and will upgrade the e3500 machine recently. Supported by the national High-tech programme and the Peking University, we obtained a grant for setting up the Bioinformatics Data Centre this year. We are applying another grant to get a bigger server (10 CPU, 5GB RAM).

As an educational and academic incubator, various seminars, workshops and courses on bioinformatics and biocomputing are being organised. Both local and external experts are invited to give lectures at Peking university. A biweekly lecture series on "Theoretical Biology" organised collaboratively with colleagues from the Institute of Theoretical Physics has been started. A master course of Introduction to Bioinformatics started this semester to some 50 M.S. students. Both Ph.D and M.S programmes of bioinformatics and biocomputing have been started to meet the need of the talents in this fast growing field.

ExPASy Mirror

As one of the collaboration projects between Swiss Institute ofBioinformatics and Peking University, supported by the National ScienceFoundation of Switzerland and the Natural Science Foundation of China, an official mirror of the Expert of Protein Analysis System (ExPASy) has been setup recently at the Centre of Bioinformatics, Peking University: http://expasy.pku.edu.cn/

This mirror contains various bioinformatics resource from sequence database such as SwissProt and TrEMBL, the proteomics database of two-dimensional polyacrylamide gel electrophoresis, to three-dimensional models of protein structure. There are also search engines, free ftp databases and online documentations. The "Amos's WWW link page" created by Dr Amos Bairoch, the boss of SwissProt : http://expasy.pku.edu.cn/alinks.html contains more than 1000 pointer to bioinformatics resources around the world. All the resources on the mirror are free for academic. Commercial users should contact: Dr Roberto Fabbretti: Roberto.Fabbretti@genebio.com - Dr Amos Bairoch: bairoch@cmu.unige.ch for licence scheme.

News from Sanger Center Node

Complete Genome of Leprosy bacillus

The sequence of the entire genome of the bacillus causing leprosy has been determined in a collaborative effort between Stewart Cole's team at Institut Pasteur and the Sanger Centre in the Great Britain. As Stewart Cole will explain at the international meeting 'Genomes 2000' currently being held at Institut Pasteur, comparison of this sequence with that of the genome of the tuberculosis bacillus (entirely sequenced by the same two teams in 1998) will give valuable information about the two diseases. According to the WHO, there are more than 800 000 new cases of leprosy every year worldwide, and more than 2 million people currently suffer severe disabilities because of this disease.

Leprosy is a chronic infectious disease mostly affecting the skin, peripheral nerves, mucosa membranes of the upper respiratory tract and the eyes. The mutilations it causes result in lepers being rejected in many societies. Ninety-two percent of the known cases in the world are in only twelve countries. These most affected countries are: India, Brazil, Indonesia, Myanmar (Burma), Madagascar, Nepal, Ethiopia, Mozambique, Democratic Republic of Congo, Niger, Guinea and Cambodia. In France, about 250 people are treated for leprosy, and there are about 25 new cases a year, mostly in people from countries where the disease is endemic or Antilles, Guiana, Mayotte.

Sequencing the genome of the leprosy bacillus (Mycobacterium leprae) was a priority both for research into the disease and for its control. The collaboration between the Bacterial Molecular Genetics Unit at Institut Pasteur, directed by Stewart Cole, and the Pathogen Genome Sequencing Unit at the Sanger Centre (UK), directed by Bart Barrell, started in 1996. This sequencing project was financed by the Heiser Program for Research in Leprosy and Tuberculosis of the New York Community Trust, and the Association Raoul Follereau.

The entire sequence of the genome of the bacillus Mycobacterium leprae is now available to research scientists. Comparisons with the tuberculosis bacillus, Mycobacterium tuberculosis, have already started. The genetic make-up of these two bacilli are very similar: the similarity of some groups of genes is as high as 93%. However, the M. leprae genome is 3.2 Mb, which is smaller than the 4.4 Mb of M. tuberculosis. The leprosy bacillus seems to have 'got rid of' non-essential genes. This may explain its slow growth, which makes culturing the bacillus in the laboratory very difficult. This in turn greatly hinders research on this bacterium and thus work on the disease it causes. The comparison of the two genomes may lead to the identification of growth factors, absent from M. leprae, which could facilitate research and would be very useful for the production of vaccines. In contrast, other genes in M. leprae are not found in M. tuberculosis: these genes may lead to the development of diagnostic tools for dermatological tests to detect leprosy. Such tools would be valuable as the earlier the disease is detected, the better the treatment. These genes only found in M. leprae may also confer particular properties, for example its neurotropism resulting in the degradation of nerves. The comparative approach being used is expected to identify new therapeutic targets and facilitate the rational development of drugs for treating leprosy.

Note that although current chemotherapy is effective, it is both expensive and extremely impractical as it requires daily administration of a combination of antibiotics for 6 months (for mild cases) to more than 12 months.

For further information contact:

Stewart Cole, Unite de Genetique moleculaire bacterienne, Institut Pasteur -Tel (00 33) 1 45 68 84 46 Email stcole@pasteur.fr http://www.pasteur.fr/recherche/unites/Lgmb/
Don Powell, Press Office, The Sanger Centre, Tel: 01223 494 956, Fax: 01223 494 714, E-mail: don@sanger.ac.uk , http://www.sanger.ac.uk/Info/Press/

News from Finnish node

There are three noteworthy things as for the Finnish EMBnet node:
1. The official name of the company is (has been since the change to CSC !) CSC -Scientific Computing Ltd.
2. Kimmo Mattila will be the node manager from mid-May until April 2001.
3. We will be organizing a workshop on Data mining and microarrays on Sep 11-12 (ten talks). The program is about to be completed, but positive responses include those of Vic-

tor Jongeneel, Thure Etzold and Michael Zhang.

Erja Heikkinen, CSC Science Support/Biosciences, P.O.Box 405, 02101 Espoo, Finland, tel. +358-(0)9-4572433, GSM +358-(0)50-3819503, http://www.fi.embnet.org

News from Indian Node

The Indian EMBnet node have recently placed orders and will soon be upgrading our computing facilities to include another server (SUN e3500), FC-switch and FC-RAID (720Gb), high performance backup DLT 7000, several PENTIUM IIIs (600 Mhz), iMACs and few SGI O2s, Gigabit switch, scanner, CD-writer.

We have started mirroring SCOP recently and installed GCG (10.0) & SeqWeb. Our paper on the Database of Structural Motifs in Proteins (DSMP) is accepted in Bioinformatics, (2000) Vol 16, No.1, pp1-4. Some of my groups other recent publications on the analysis of turns in proteins is referenced on our website (under publications).

The EMBnet India Node was formally inaugurated by Nobel Laureate Prof. Harold E. Varmus on 1st March 2000. We now have 2 independent 64 Kbps internet leased lines but have applied for a 2Mbps line from a private IP in Hyderabad. We may soon upgrade to the 2Mbps perhaps in 8-12 weeks.

Our city, Hyderabad, has recently come to be known as Cyberabad, considering its initiative and our Chief Minister's foresight to develop our city into a major IT centre. This attracted Clinton visit our city.

Dr. Kunchur Guruprasad, Chief, Bioinformatics & EMBnet India Node Manager, CDFD Laboratory, Nacharam, Hyderabad - 500 556. INDIA., http://www.cdfd.org.in, e.mail: guru@www.cdfd.org.in , Tel: +91-40-7155607, Fax: +91-40-7155610/7155479

News from Argentinian node

We are just finishing now a short course in bioinformatics at our node in La Plata Argentina. It was supported by EMBnet education and training committee It was given by Marc Colet from BEN and included the following subjects:

- Databases searching: SRS and ENTREZ
- Introduction to GCG package of sequence analysis tools
- GCG graphical interface WWW2GCG: translation, sequence mapping, Open Reading Frames.
- Pairwise sequence comparison
- Sequence databanks similarity search: FastA, BLAST
- Multiple sequence alignment
- Primers design
- Anonymous and nominal FTP
- Motifs search

It was attended by 17 students from Argentina and Chile.

Dr. Oscar Grau, e-mail: grau@biol.unlp.edu.ar
RELAB Website http://relab.biol.unlp.edu.ar
Tel/Fax: +54 221 425 9223

Node News from HGMP

Briefings in Bioinformatics is edited by Dr Martin Bishop, Head of the Bioinformatics Division at the Human Genome Mapping Project Resource Centre in the UK and was launched in February 2000. The Journal provides an indispensable resource for the experimental practitioner seeking awareness of the disparate sources of data and analytical tools of contemporary genetics and molecular biology. Topics covered by the Journal range in scope and depth, from the introductory level to specific details of experimental methodologies and analyses, encompassing bacterial, plant, animal and human data. Papers should be between 3,000-5000 words in length; to ensure we meet academic standards of integrity, all papers are sent for double-blind peer review.

Notes for contributors for the Journal can be found on the Journal's web-site at : http://www.henrystewart.co.uk/Journals/BiB/ As the Journal has been supported by an educational grant from EMBnet, authors will receive an honorarium of 300 Euros upon publication to help cover expenses.

Martin J.Bishop, HGMP Resource Centre, Hinxton, Cambridge CB10 1SB, UK, Tel +44 1 223 494525 main switchboard +44 1 223 494500, Fax +44 1 223 494512

The EMBnet Nodes

National Nodes

Argentina

Dr Oscar Grau
IBBM Facultad de Ciencas Exactas Universidad Nacional de LaPlata Argentina
Email: grau@biol.unlp.edu.ar
Tel:+54-21-250497 Fax:+54-21-259223

Australia

Dr Tim Littlejohn
ANGIS Electrical Engineering Building J03 University of Sydney
Sydney NSW 2006 Australia
Email: tim@angis.org.au
Tel:+61 2 9351 2948 Fax:+61-2-9351 5694

Austria

Dr Martin Grabner
BioComputing Centre Vienna University
Computing Centre Dr Bohr Gasse 9 Vienna.
Email: martin.grabner@cc.univie.ac.at
Tel: +43-1-4277-14141 Fax: +43-1-7986224

Belgium

Dr Robert Herzog
BEN Université Libre de Bruxelles CP300 Paardenstraat 67 1640 Sint Genesius Rode Belgium
Email rherzog@ulb.ac.be
Tel: +32-2-6509978 Fax:+32-2-6509998

Canada

Dr Christoph Sensen
National Research Council of Canada Institute of Marine Biosciences 1411 Oxford St Halifax Nova Scotia Canada B3H 2Z1
Email: sensencw@niji.imb.nrc.ca
Tel:+1-902-4267310 Fax:+1-902-4269413

China

Professor Jingchu Luo
College of Life Sciences Peking University Beijing 100871 China
Email: luojc@lsc.pku.edu.cn
Tel:+86-10-6275 5206 Fax:+86-10-6275 1843

Cuba

Dr Ricardo Bringas
Centre for Genetic Engineering PO Box 6162 Havana Cuba
Email: bringas@cigb.edu.cu
Tel:+53-7 218200 Fax:+53-7 218070

Denmark

Mr Hans Ullitz-Moller
BioBase - Danish Human Genome Centre Aarhus Universitet Ole Worms Alle 170-171 DK-8000 Aarhus C Denmark
Email: hum@biobase.dk
Tel:+45-86139788 Fax:+45-86131160

Finland

Dr Kimmo Mattila
CSC Center for Scientific Computing PL 405 (Tietotie 6) 02101 Espoo Finland
Email: kimmo.mattila@csc.fi
Tel:+358-9-4572433 Fax:+358-9-4572302

France

Dr Guy Vaysseix
Infobiogen 7 rue Guy Moquet - BP8 94801 Villejuif Cedex France
Email: Guy.Vaysseix@infobiogen.fr
Tel:+33-1-45595241 Fax:+33-1-45595250

Germany

Dr Clemens Suter-Crazzolara
Department of Molecular Biophysics H0200 DKFZ 280 D-69120 Heidelberg Germany
Email: c.suter-crazzolara@dkfz-heidelberg.de
Tel:+49/6221-422342 Fax:+49/6221-423333

Greece

Dr Babis Savakis
FORTH Insitute of Molecular Biology PO Box 1527 711 10 Heraklion Crete Greece
Email: savakis@nefeli.imbb.forth.gr
Tel:+30-81-212647 Fax:+30-81-231308

Hungary

Dr Endre Barta
Agricultural Biotechnology Centre Szent-Gyorgyi u. 4 PO Box 410 2100 Godollo Hungary
Email: barta@abc.hu
Tel:+36-28-430127 Fax:+36-28-420096

India

Dr. Kunchur Guruprasad
CDFD A & H Bldg.,
ECIL Road, Nacharam, Hyderabad-500 076,INDIA
Email: guru@in.embnet.org
Tel: +91-40-717 2241 ext. 748 Fax: +91-40-715 0008

Ireland

Dr Andrew Lloyd
INCBI Dept Genetics Trinity College Dublin 2 Ireland
Email: atlloyd@tcd.ie
Tel:+353-1-608-1969 Fax:+353-1-679-8558

Israel

Dr Leon Esterman
Biological Computing Division Weizmann Institute of Science Rehovot 76100 Israel
Email: lsestern@weizmann.weizmann.ac.il
Tel:+972-8-9343934 Fax:+972-8-9466269

Italy

Dr Marcella Attimonelli
Area di Ricerca CNR-BARI Via Amendoal 166/5 70126 - Bari Italy
Email: marcella@area.ba.cnr.it
Tel:+39-80-5482130 Fax:+39-80-5484467

Netherlands

Dr Jack Leunissen
CMBI University of Nijmegen Toernooiveld 6525 ED Nijmegen Netherlands
Email: jackl@cmbi.kun.nl
Tel:+31 24 365 22 48 Fax:+31 24 365 29 77

Norway

Dr. David Martin
Biotechnology Centre of Oslo University of Oslo
Gaustadalleen 21 0317 Oslo Norway
Email: david.martin@biotek.uio.no
Tel:+47-22958756 Fax:+47-22694130

Poland

Dr Piotr Zielenkiewicz
Institute of Biochemistry and Biophysics Polish Academy of Sciences Pawinskiego 5a 02-106 Warsawa Poland
Email: piotr@ibbrain.ibb.waw.pl
Tel:+48-2-6584703 Fax:+48-39-121623

Portugal

Dr Pedro Fernandes
Instituto Gulbenkian de Ciencia Rua da Quinta Grande Apt. 14 2781 Oeiras Codex Portugal
Email: pfern@pen.gulbenkian.pt
Tel:+351-1-443 1408 Fax:+351-1-443 5625

Russia

Professor Sergei Spirin
Belozersky Institute of PhysicoChemical Biology Moscow State University Laboratory Korpus A - Room 612 119899 Vorobyevy Gory - MOSCOW Russia
Email: sas@brodsky.genebee.msu.su
Tel:+7 (095) 932 8825 Fax:+7 (095) 939 3181

Slovakia

Dr. Lubos Klucar
EMBnet Slovakia
Dubravska cesta 21
842 51 Bratislava - Slovak Republic
Email:klucar@embnet.sk
Tel: +421 7 5941 1284 Fax: +421 7 5477 2316

South Africa

Dr Win Hide
SANBI Private Bag X17 Bellville 7535 University of the Western Cape South Africa
Email: winhide@techno.sanbi.ac.za
Tel:+27 21 959 3645 Fax:+27 21 959 2512

Spain

Dr JoseRamon Valverde
CNB Universidad Autonoma de Madrid Campus de Canto Blanco 28049 Madrid Spain
Email: jvalverde@cnb.uam.es
Tel:+341-5854543 Fax:+341-5854506

Sweden

Mr Nils-Einar Eriksson
Uppsala Biomedical Centre Box 570 S-721 23 Uppsala Sweden
Email: Nils-Einar.Eriksson@bmc.uu.se
Tel:+46-18-471 40 17 Fax:+46-18-55 17 59

Switzerland

Dr Victor Jongeneel
ISREC Bioinformatics Group Chemin des Boveresses 155 CH-1066 Epalinges Switzerland
Email: victor.jongeneel@isrec.unil.ch
Tel:+41-21-692-5994 Fax:+41-21-653-4474

Turkey

Dr Zehra Sayers
National Bioinformatics Node (NBN) MAM GMBAE NBN POB 21 41470 Gebze-Kocaeli Turkey
Email: zehra@bioinfo.rigeb.gov.tr
Tel:+90-262-6412300 ext 4007 Fax:+90-262-6412309

United Kingdom

Dr Martin Bishop
HGMP Resource Centre Hinxton Cambridge CB10 1SB UK
Email: mbishop@hgmp.mrc.ac.uk
Tel:+44 1223 494500Fax: +44 1223 494512

Specialist Nodes

EMBL-EBI

Dr Rodrigo Lopez
EBI Hinxton Hall Hinxton Cambridge CB10 1SD England
Email: rls@ebi.ac.uk
Tel: 1223 494438 Fax:+44 1223 494 468

ETI

Dr Peter Schalk
ETI biodiversity Center Universiteit van Amsterdam Mauritskade 61 1092 AD Amsterdam theNetherlands
Email: pschalk@eti.uva.nl
Tel:+31-20-5257239 Fax:+31-20-5257238

Hoffman-LaRoche

Dr Martin Ebeling
Hoffman-LaRoche PRPZ-B 65/303
CH-4070 Basel Switzerland
Email: martin.ebeling@roche.com
Tel:+41 61 6871398 Fax:+41 61 688 1075

ICGEB

Dr Sandor Pongor
ICGEB Padriciano 99 34012 Trieste Italy
Email: pongor@genes.icgeb.trieste.it
Tel:+39 40 3757300 Fax:+39 40 226555

MIPS

Dr Werner Mewes
MIPS Max Planck Institut fur Biochemie Am Klopferspitz 18a D-82152 Martinsried Germany
Email: mewes@mips.biochem.mpg.de
Tel:+49 89 8578 2656 Fax:+49 89 8578 2655

Pharmacia

Dr Staffan Bergh
Pharmacia-Upjohn AB Strandbergsgatan 49 112 87 Stockholm Sweden
Email: staffan.bergh@eu.pnu.se
Tel:+46-8-6959884 Fax: +46-8-695 4084

Sanger Centre

Mr Peter Rice
The Sanger Centre Wellcome Trust Genome Campus Hinxton Cambridge CB10 1SA England
Email: pmr@sanger.ac.uk
Tel:+44 1223 494967 Fax:+44 1223 494919

UMBER

Dr Terri Attwood
School of Biological Sciences
2.19, Stopford Building The University of Manchester
Oxford Road Manchester M13 9PT
Email: attwood@bioinf.man.ac.uk
Tel.:+44 161 275 5766 Fax: +44161 275 5082

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 in the Tips from the computer room section, please let us know. Submissions for the BITS section are most welcome, but please remember that we cannot extend space beyond two pages per article. Please send your contributions to one of the editors. You may also submit material by Internet E-mail to:

emb-pub@hgmp.mrc.ac.uk

You are invited to contribute to the LETTERS TO THE EDITOR section.

If you had difficulty getting hold of this newsletter, please let us know. We would be only too happy to add your name to our mailing list. This newsletter is also available on-line using any WWW client via the following URLs:

The Online version, (ISSN 1023-4152) :

- ***http://www.be.embnet.org/embnet.news/vol7_1/***
- ***http://www.ebi.ac.uk/embnet.news/vol7_1/***
- ***http://www.ie.embnet.org/embnet.news/vol7_1/***

A Postscript version (ISSN 1023-4144) is available. You can get it by anonymous ftp from:

- ***ftp.uk.embnet.org in the directory pub/embnet.news/***
- ***ftp.be.embnet.org in the directory pub/embnet.news/***
- ***ftp.ebi.ac.uk in the directory pub/embnet.news/***
- ***ftp.ie.embnet.org in the directory pub/embnet.news/***

A pdf version (ISSN 1023-4144) in Acrobat 3 format is also available. You can get it by anonymous ftp from:

- ***ftp.uk.embnet.org in the directory pub/embnet.news/***
- ***ftp.be.embnet.org in the directory pub/embnet.news/***
- ***ftp.ebi.ac.uk in the directory pub/embnet.news/***
- ***ftp.ie.embnet.org in the directory pub/embnet.news/***

Back issues are available at most of these sites.

Publisher:

EMBnet Administration Office.
c/o Jack Leunissen
CMBI
University of Nijmegen
6525 ED Nijmegen
The Netherlands

Editorial Board:

Alan Bleasby, HGMP, UK
(bleasby@hgmp.mrc.ac.uk)
FAX +44 (0)1223 494512
Tel +44 (0)1223 494520

Robert Harper, EBI, Hinxton Hall, UK
(harper@ebi.ac.uk)
FAX +44 (0)1223 494468
Tel +44 (0)1223 494429

Robert Herzog, BEN, Free University Bruxelles, BE
(rherzog@ulb.ac.be)
FAX +32-2-6509998
Tel +32-2-6509978

Andrew Lloyd, INCBI, Trinity College Dublin, IE
(atlloyd@acer.gen.tcd.ie)
FAX +353-1-679-8558
Tel +353-1-608-1969

Rodrigo Lopez, EBI, Hinxton Hall, UK
(Rodrigo.Lopez@ebi.ac.uk)
FAX +44 (0)1223 494468
Tel ++44 (0)1223 494423

Peter Rice, Sanger Centre, Hinxton Hall, UK
(prm@sanger.ac.uk)
FAX +44 (0)1223 494919
Tel +44 (0)1223 494967

embnet.news
Vol.7, No.1
April 30, 2000
ISSN 1023-4144