EMBnet.journal

Volume 17
Supplement A
December 2011

EMBnet AGM 2011 Workshop

Editorial

This EMBnet.journal supplement is dedicated to the 2011 EMBnet Annual General Meeting (AGM) and satellite workshop, which took place at the Instituto Gulbenkian de Ciência in Oeiras, Portugal, between the 23 and 25 May 2011. This, the 23 consecutive EMBnet AGM, was a significant one: it was celebrated during a time when EMBnet has begun to redefine itself in order to continue to contribute effectively to the life sciences' evolving data-driven landscape, for the benefit of the communities it continues to serve.

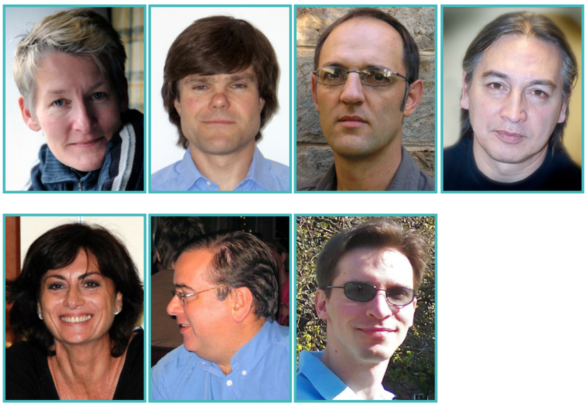
The 2011 workshop and AGM consequently formed an important landmark in EMBnet´s history. The discussions centered around questions on how EMBnet could and should change in order to become a more efficient and agile organisation. Participants in the workshop worked extremely hard during the meeting: this EMBnet.journal supplement highlights some of the important discussions around the work of the EMBnet community towards shaping and re-shaping an effective EMBnet of tomorrow.

EMBnet.journal Editorial Board

Contents

Cover picture: Sunset Surfing, Oeiras, Portugal, May, 2011. [© Erik Bongcam-Rudloff]

Report of the EMBnet AGM 2011 Workshop, Oeiras, Portugal, 23-25 May, 2011



Teresa K Attwood[1], Andreas Gisel[2], Etienne de Villiers[3], Erik Bongcam-Rudloff[4], Domenica D'Elia[2], Pedro Fernandes[5], Georgios Magklaras[6]

[1]Faculty of Life Sciences and School of Computer Science, University of Manchester, Machester, UK,
[2]CNR-Institute for Biomedical Technologies, Bari, Italy
[3]BecA-ILRI, Nairobi, Kenya
[4]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden
[5]Instituto Gulbenkian de Ciência (IGC), Oeiras
[6]The Biotechnology Centre of Oslo, The University of Oslo, Oslo, Norway

The 2011 AGM workshop took place at the Instituto Gulbenkian de Ciência (IGC) in Oeiras, Portugal, from 23-25 May (Figure 1). The goal of the workshop was to build on the demonstrable progress made during the previous year, in particular by helping to deliver on some of the plans outlined during the 2010 AGM. It was also an opportunity to build on our commitment to take EMBnet forward by embracing new partners and new activities. The following pages summarise the workshop content, discussions and conclusions.

Mini Symposium

The 2011 AGM satellite events began with a short scientific session, providing a snapshot of the latest advances in the development of a variety of bioinformatics tools to support 'omics' studies. The mini symposium opened with a talk entitled, *"Bioinformatics of comparative cell biology"*, by José Pereira-Leal, from the Computational Genomics Laboratory of the IGC; his presentation described both RabDB, a comprehensive resource for studying the regulation of membrane traffic pathways mediated by the Rab family of proteins, and mtocDB, a new resource that integrates microscopy data and molecular information to help characterise morphological variation observed in centrioles and centriole-associated structures. Matthias Haimel, from the European Bioinformatics Institute (EBI), UK, presented *"Velvet, Curtain and beyond"*, describing some of the challenges and solutions for *de novo* genome assembly with Velvet and Curtain. Finally, Gert Vriend, director of the CMBI at the Centre for Molecular Life Sciences, Nijmegen, the Netherlands, closed this first part of the symposium with a talk on *"HOPE, a mutational analysis 'machine' for medics"*; HOPE is an easy-to-use webserver workflow, including a diverse set of protein analysis tools and databases for analysing the structural effects of mutations.

To set the scene for the workshop, the programme continued with three further short presentations. Gert Vriend gave a hard-hitting 'perspectives' talk. He examined EMBnet's origins and its current place in the global bioinformatics landscape, reflected on its strengths and weaknesses, and provided a vision of its future opportunities. Thinking about the future, Erik Bongcam-Rudloff, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, described the achievement of SeqAhead, a new COST Action initiative involving a large number of EMBnet Nodes, which, over the next few years, aims to tackle many of the complex informatics issues generated by the ongoing deluge of next-generation sequencing data. Finally, Fran Lewitter, Director of Bioinformatics and Research Computing at Whitehead Institute and Education Committee Chair of the International Society for Computational Biology (ISCB), presented the Society's work in the realms of bioinformatics education and training, and discussed the newly announced alliance with EMBnet.

During the broad-ranging discussions that followed, Erik encouraged Nodes who have an interest in active participation to sign up to the SeqAhead COST Action; Gert pledged his interest in working with an EMBnet-wide/world-wide bioinformatics education activity; and Fran, recapitulating the readiness of the ISCB to take the first steps towards future collaborative bioinformatics education and training initiatives, invited members of EMBnet to participate in a survey on

Figure 1. Collage of events at the 2011 AGM. Left-hand column: invited speakers Fran Lewitter, Gert Vriend and José Pereira-Leal deliver their presentations at the mini-symposium; right-hand column: proceedings of the business meeting; centre column: a hard-hitting perspectives talk from Gert and a relaxing conference dinner afterwards.

bioinformatics curricula. The future looks bright, but much work remains to be done to take such education/training initiatives and this COST Action forward.

Working Group Meetings

Following the symposium, participants got together in working groups to discuss the content and graphics for the new websites being prepared for EMBnet and for *EMBnet.journal*; participants also discussed the status of EMBnet's current e-learning provision and of its Quick Guide portfolio. To assist the work of the groups, tutorials were run by Lubos Klucar and César Bonavides-Martínez on the Open Journal System (OJS) and on the Drupal Content Management System, respectively.

The challenge for the OJS, website-design and e-learning working groups was to analyse the state-of-the-art and to explore possible future improvements within each of these areas in order to allow EMBnet to maintain a relevant and active role in the global bioinformatics community. The OJS working group pointed out the importance of implementing an improved, informative and attractive journal home page. The website-design group, comprising members of the TM PC and P&PR PC, elaborated a new website graphical model, based both on the recommendations and preferences expressed by the website task-force and on comments from those EMBnet members who participated in the meeting. The result of the e-learning working group discussions was a consensus to remove links to the current e-learning portal, and instead to develop a new strategy and a new website in collaboration with the CMBI, the ISCB and, hopefully, the Bioinformatics Training Network in future. The creation of new Quick Guides was also discussed: proposals were made to update some of the older, more popular Guides and to develop several new ones.

Executive Board and Project Committee (PC) Chair Meetings

In parallel with the working group meetings, the Executive Board and PC Chairs held several meetings to discuss proposed new structures for EMBnet's membership and governance. Against a backdrop of flux in the global bioinformatics environment, and spurred on by new European infrastructural initiatives, the time was ripe to consider in what ways EMBnet might need to change in order to remain a relevant, useful and viable bioinformatics organisation, one that could respond in an efficient and agile manner to challenges of the future.

Key ideas for the new structure included: i) phasing out the requirement of member Nodes to obtain government mandates – where Nodes have national status, EMBnet would continue to recognise this, but a mandate to provide national services would no longer be a condition of membership; ii) allowing individuals to become members; and iii) inviting Honorary Members. Taken together, such changes would reduce the barriers to membership, allowing multiple Nodes and individual members per country, each with interests in, or specialising in, the provision of biodatabases, bioinformatics tools and/or bioinformatics training. The overall scheme is illustrated in Figure 2.

To reflect the proposed new structure, changes to EMBnet's management were also discussed. The principal innovation was to instantiate a new 'Operational Board', aiming to achieve a tighter coupling between the work of the Executive Board and of its Committees. Membership of the Operational Board would be derived from the Executive Board, the PC Chairs, and discretionarily, the Chairs of any designated Special Interest Groups (*e.g.*, such as SeqAhead).

Hand-in-hand with these proposals was the recognition that new membership benefits and fees would need to be put in place, and that the Statutes and Byelaws would need to be revised to accommodate the new structure. All changes to the Statutes must be ratified by the full EMBnet Board; to expedite this process, revised Statutes will be made available to members for comment once a complete working draft has been developed.

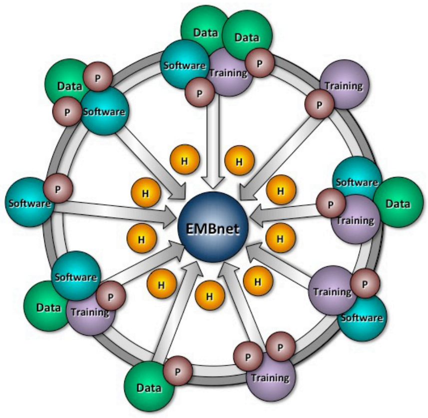


Figure 2. Illustration of the proposed new structure of EMBnet, in which participating countries have multiple Nodes (some, or none, of which may have national status) and individual members (P), with specialisms in data, software and/or training provision. In the new model, Honorary Members (H) may also be invited to join EMBnet.

Conclusion

EMBnet has played, and is poised to continue to play, an important part in the life sciences' evolving data-driven landscape. To remain viable and relevant to today's genomic revolution, however, it must adapt. The 2011 workshop was an important opportunity to discuss how EMBnet could and should change in order to become a more efficient and agile organisation in future. Participants in the workshop worked extremely hard, and we are grateful to them for the energy they invested in shaping tomorrow's EMBnet. We are not there yet, but with the strong collaborative and community spirit that is EMBnet's hallmark, we are in the process of taking significant steps and we expect to get there very soon!

Executive Board: AGM2011 report

Teresa K Attwood

Faculty of Life Sciences and School of Computer Science, University of Manchester, Machester, UK

Commitee members

Chair:	Teresa K Attwood
Secretary:	Andreas Gisel
Treasurer:	Etienne de Villiers
Member:	Erik Bongcam-Rudloff

Since the last Annual General Meeting (AGM), the Executive Board (EB) has continued to meet regularly and to convene monthly video meetings for general discussion of issues relating to the Project Committees (PCs), to *EMBnet.journal*, to the website, *etc.*; the meetings have also allowed us to discuss matters relating to the Stichting accounts, and to prepare the groundwork for this year's AGM.

The monthly video meetings are open to all members of the EMBnet constituency, and those who've attended will appreciate the significant technical issues we've experienced with the platforms used to host them; needless to say, we're striving to find the best solution to these ongoing problems, but meanwhile encourage participation and greater involvement with our activities – despite the technical irritations, there's a lot going on!

This time last year, the EB had been in office for a relatively short time and most of its members were new. The incoming EB therefore faced a steep learning curve, at a time when the global bioinformatics landscape was also clearly in a state of flux. We therefore wanted to use the 2010 AGM as an opportunity to examine those changes, to reflect on EMBnet's position, and to explore its aspirations and motivations moving forward. Recognising the need for both structural and behavioural changes within EMBnet, the idea was to take some practical steps towards trying to make the daily business of EMBnet run more efficiently, to engage more of its members in the activities of the PCs, and to start to improve the EMBnet 'brand'. Specifically, we hoped to be able to establish internal mechanisms by which we might: i) secure strategic alliances with other organisations; ii) compete more strategically for funding for some of our activities; and iii) promote our activities more effectively. Ultimately, our goal was to facilitate the evolution of EMBnet into a respected network of bioinformatics professionals.

Figure 1. EMBnet members during the AGM in Oeiras, Portugal.

Today, our work towards seeding these ambitions is beginning to bear fruit, and the successes we have to report are a testament to our increasing professionalism and the respect this is garnering outside the EMBnet community. A year on, we've managed to i) establish a formal alliance with the ISCB in the field of education [ISCB Press Release[1]]; ii) win funds to underpin a training and visitor exchange programme around the theme of NGS data analysis (this is the SeqAhead[2] COST Action); iii) introduce EMBnet.digest[3], a lightweight monthly flyer for disseminating some of EMBnet's key activities and success stories; and iv) work with the PCs both to develop a professional EMBnet exhibition stand and to upgrade and migrate the www.embnet.org Drupal CMS to the new Linux server in Norway. This year's AGM aims to build on these foundations, to actively engage all members in: developing content and graphics for the new Drupal website; evaluating the e-learning site and outlining strategies for its "rebirth"; updating the content of the *EMBnet.journal* website;

---

[1] http://journal.embnet.org/index.php/embnetjournal/article/view/217/477

[2] http://www.seqahead.eu

[3] http://www.embnet.org/EMBnet.digest

and developing new Quick Guides, *etc.*, as time permits. We also plan to use the AGM as an opportunity to formulate concrete resolutions on EMBnet's internal structure, membership and alliances, and on its policies and mechanisms for sponsorship and fund raising.

Given the transitional period of the last 18 months or so, we considered it important to try to maintain a degree of stability within the EB. Therefore, at the forthcoming AGM, we propose to make no changes, in order to allow the current team both to build on the demonstrable progress we've made during these challenging times, to deliver more on the plans we outlined last year, and to take EMBnet forward by embracing new partners and new activities. There's clearly a lot more to do, and we therefore call on all of you to help us realise our collective ambitions for the most widely distributed and oldest bioinformatics network in the world – a network of which we should rightly be proud!

Education and Training Committee: AGM2011 report

Pedro L Fernandes
Instituto Gulbenkian de Ciência, Oeiras, Portugal

Commitee members

Chair: Pedro Fernandes
Secretary: Emiliano Barreto
Member: Bruno Gaeta,
Member: Shahid Chohan

The activity that was proposed for the 2010-2012 started in the Summer of 2010. Some work had to be put into a better definition of short and long term priorities, an assessment of the difficulties and costs involved, etc.

Mostly there were discussions over the e-mail and on the telephone. We started a series of contacts with prospective authors for the materials that we intend to develop.

Virtual Training

We have realised that there is no sufficient data to be able to make decisions on that. Using the Gulbenkian Training Programme in Bioinformatics (GTPB) programme at the Portuguese node as a possible way of performing comparative experiments, we have been collecting information in the following ways:

- we have made video recordings using commercial equipment for lecture capture on loan (Sonic Foundry MediaSite)[1]. Our intention is to evaluate usage utility with several classes of attendees. Experiments with the materials will start in June 2011;
- we have tried support platforms for running practicals, such as websites and collections of documents (samples available);we have tried remote audiences using collaborative tools such as Adobe Connect. We will also test Webex[2] in an upcoming training course.

---

[1] http://www.sonicfoundry.com/mediasite/

[2] http://www.webex.com/

So far, the conclusion is that only a mixture of methods can cover the specific needs of Bioinformatics training.

We have "Server side" data collection in progress at the moment, building towards a clear definition of requirements.

EMBnet Kits

We have made a series of contacts with potential authors for the materials. Mostly with positive results. We have also conducted a rough evaluation of costs.

As an exercise, we have started a feasibility study on the production of an "Introductory Bioinformatics" book. The conclusion is that it is not an opportune idea, moreover we know of two competitive books coming-up in 2011-2012, backed by commercial publishers.

So, we carry on with the idea of the kits based on materials on electronic support.

We are making the first plan for the "Embnet student kit" and starting the pre-plan for the "Agro" and "Food" kits, in connection with volunteers from universities.

Training for active Bioinformaticians

The format model of the GTPB "Foundations" courses is currently under review. An additional idea is under preparation based on the model of Role Playing Games: a bioinformatician brings previously worked problems to the training room playing the role of a user asking for support. The trainers follow with a series of suggestions that are explored in the training room.

Software for collective exercises

Commercial vendors are not interested so far. We have contacted several commercial vendors, but the general attitude is that they are interested in the undergraduate market and not in cooperation. We are currently exploring open source solutions but all of them require adaptations for which we have no manpower.

Quick Guides

A survey is under preparation to probe for interest in new Quick Guides. The level of response is still low but there is a clear incidence in "R", "Bioconductor" and "Primer design".

We believe that there is room for a new way of thinking about these kits, that reflects the change in the the majority of end users, that has shifted away from "command line" into GUIs, and, in particular, web interfaces. This is up for a discussion in the upcoming Annual General Meeting (AGM). As stated in our business plan, cooperation with the TMPC is deemed necessary.

Cooperation with ISCB

We see the recent joint press release[3] (26th April 2011), as a positive step toward a fruitful cooperation. There is a real chance of defining its contours in the upcoming AGM, that will be attended by Fran Lewitter (Whitehead Institute), who chairs the Education Committee of ISCB[4]. Bruno Gaeta, manager of the Australian EMBnet node New South Wales Systems Biology Initiative (SBI), is a member of that committee. We will explore the possibilities of making our plans converge for the benefit of users and communities of users worldwide.

Bioinformatics Training Network (BTN)

Pedro Fernandes and Terri Atwood are founding members of the BTN[5], that is keen on optimising training methodologies in Bioinformatics. The group has now come-up with a website where resources are being organised. We believe that there are big benefits in cooperating with BTN, such as not duplicating efforts, for example. Again, the contours of such an intention for cooperation are not yet defined, but a proposal can be prepared at the upcoming AGM.

Reflections on Education in Bioinformatics

The ET PC is preparing a set of recommendations that are a mere reflection of our concerns about undergraduate and graduate education in the sense of promoting best practices in the educational universe. It is generally felt that education in Bioinformatics is insufficiently supported by professionals and that more practical work should be introduced, based on a standardised set of problems. The "EMBnet kit for students" - in the initial phase of preparation at the moment - can become the embryo of such an idea. We will evaluate these possibilities more carefully when our cooperation with ISCB is better defined.

---

[3] http://journal.embnet.org/index.php/embnetjournal/article/view/217/477

[4] http://www.iscb.org/iscb-leadership-a-staff-/112

[5] http://www.biotnet.org/

Publicity and Public Relations Project Committee: AGM2011 report

Domenica D'Elia
CNR-Institute for Biomedical Technologies, Bari, Italy

Commitee members

Chair:	Domenica D'Elia
Secretary:	Lubos Klucar
Member:	Martin Norling
Member:	Judit Kumuthini
Member:	Kanchana Senanayake

The main mission of the P&PR PC is to nurture and promote EMBnet's image at large. The P&PR PC is responsible for promoting any type of EMBnet activities, for the advertisement of products and services provided by the EMBnet community, as well as for proposing and developing new strategies aiming to enhance EMBnet's visibility, and to take care of public relationships with EMBnet communities and related networks/societies. In this document, we report proposals, activities and achievements of the committee from June 2010 to May 2011.

Introduction

Since June 2010, the PC's composition and organisation has changed. Namely, Pedro Fernandes, who led the committee since its establishment, moved to guide the ET PC; Judit Kumuthini and Kanchana Senanayache, from the South African and Sri Lankan nodes respectively, joined the committee as new board members, while Goran Nerish, Andreas Gisel and Pedro Fernandes joined the committee as associate members. We would like to thank Pedro for the great contribution he gave in his role of chair of the committee up to June 2010, and for his continuous support as associate member.

Figure 1. The Publicity and Public Relations Project Commitee (P&PR PC) 2010. Left to Right: Judit Kumithini, Kanchana Senanayake, Domenica D'Elia, Lubos Klucar and Martin Norling.

Committee meetings and documents

A first meeting aiming to agree on the new structure and organisation of the P&PR PC was held in June, immediately after the conclusion of the EMBnet AGM 2010 in Ruvo di Puglia (IT). The committee elected as chair Domenica D'Elia, and as secretary, Lubos Klucar. From July 2010 to May 2011, activity plan proposals and working organisation of the committee were discussed and agreed by e-mail and by three Virtual Meetings (VMs) (22 Sep., 2010; 21 Oct., 2010; 18 Jan., 2011). The P&PR PC provided a business plan for the period 2010-2013 on October 2010, and released three activity reports as follow: Nov 2010-Jan 2011, Feb 2011 and March 2011. All these documents are available to EMBnet members from the new archive of the EMBnet website.

P&PR PC visions and goals

1. Assessment of EMBnet community identity and needs

We propose the implementation of periodical surveys, at least on a yearly basis, aiming to get a snapshot of EMBnet and EMBnet communities (i.e., collaborating research groups in the same or different institutions). This activity would help to assess current EMBnet identity (bioinformatics research, capacity-building), developmental potential, as well as the needs of EMBnet-related communities (i.e., training support, development of bioinformatics tools, advice on infrastructure solutions, organisation of thematic workshops, help to initiate bioinformatics activities). Interview results would help to enhance:

- effectiveness of EMBnet activity projects;
- linkage between EMBnet and local research communities;
- collaborative initiatives/projects amongst EMBnet members and related/associated communities;
- improvement of EMBnet website content and of its scientific and social impact.

2. Real-time monitoring of EMBnet's business activity (BAM)

The website should be complemented with a good Web Analytics tool to allow real-time monitoring of EMBnet's target and performance in line with the EMBnet strategic road map. This would allow the EMBnet board to quickly address problems (if any), to make better-informed business decisions, and to take full advantage of emerging opportunities. The first step has already been done, and the whole embnet.org site is being monitored using the Google Analytics system, as of June 2010.

3. Publicity policy for the EMBnet website and *EMBnet.journal*

The proposal is to dedicate some space on the EMBnet website and *EMBnet.journal* portal to publicity banners for commercial companies with interests in bioinformatics and related research fields (*e.g*, IT, Biotechnology and Life Sciences Companies). Income should be invested in tasks of strategic importance for EMBnet, such as the website, *EMBnet.journal*, sponsorship at key conferences, or establishment of special prices or travel grants for young students to EMBnet conferences, workshops, etc.

4. EMBnet sponsorship at key conferences

Sponsorship at key conferences represents a strategic task for enhancing EMBnet visibility. We propose to define terms and conditions for implementing special cooperation agreements with other (even dissimilar) networks and societies for mutual sponsorship of products and activities at key conferences. Steps along these lines have already been taken through EMBnet sponsorship of the 2011 ISCB Africa ASBCB Conference on Bioinformatics[1], in Cape Town. The effectiveness of this policy is demonstrated by special agreements established between ISCB and EMBnet on a collaborative program for education and training in Bioinformatics. A press release[2] on this agreement is available from the EMBnet website.

5. EMBnet Contact Registration

In line with the main objectives of the committee to use new strategies to strengthen the network of EMBnet contacts, the P&PR PC has proposed to establish a registration mechanism through which people interested in the activities of EMBnet can register and hence receive news and posting comments to EMBnet without any type of obligation. A contact registration form[3] has been designed on February 2011 and activated on March 2011. Registrations and statistics are recorded and accessible to privileged users.

Alerts on this initiative have been included in EMBnet publications (brochure and 16.1 issue of *EMBnet.journal*), distributed at the ISCB Africa conference (9-11 March, 2011) and at the workshop organised by EMBnet in conjunction with the conference[4] (7-8 March 2011). Contacts registered up to May 2011 are from Cameroon, Canada, Ghana, Greece, Indonesia, Kenya, Mexico, South Africa, Sweden, Tanzania, Tunisia, United Kingdom, USA, and Vietnam. As expected, the majority (28%) come from South Africa, as a result of dissemination carried out at the 2011 Africa ISCB conference in Cape Town, but other contacts learned about this initiative through the Web. Comments posted by contacts confirm the need of students and young researchers to get training (Master, basic and advanced courses in bioinformatics), a general interest in collaboration on the development of bioinformatics tools, and a desire to participate in our conference. An EMBnet contact mailing list has been implemented by the TM PC on April 2011, but much more is necessary to cope with new needs arising from this initiative. The EMBnet website must have a "Contact" page, and help desks need to be activated and maintained to provide adequate services.

---

[1] http://www.iscb.org/iscbafrica2011

[2] http://journal.embnet.org/index.php/embnetjournal/article/view/217/477

[3] http://www.embnet.org/en/ContactRegistration

[4] http://www.iscb.org/iscbafrica11-program

EMBnet website: proposals and achievements

The P&PR PC has designed and submitted to the EB and TM PC a first proposal for the re-modelling of the EMBnet website on October 2010. A working plan for the coordination and implementation of new developments has been definitively agreed with the TM PC and EB on March 2011. Some of the changes proposed could be implemented only after TM PC has completed the upgrading and updating of the new server, others changes have been made and are:

1. EMBnet documents archive;
2. Public archive of EMBnet conferences and workshops (coll. with TM PC);
3. EMBnet.digest main menu block with hyperlink to the last released digest and to the archive;
4. Internal mailing lists page: this page contains mailing lists maintained by EMBnet, and a guide on how to use the lists (coll. with TM PC);
5. National, Specialist and Associate Nodes lists: the menu block containing information about member nodes has been recently reorganised with the addition of the Associate nodes section. Node information has been updated (where possible) immediately after the AGM in 2010. Some work is still necessary in order to correctly place Specialist and Associate nodes in the relevant lists, as well as to update these lists with new nodes currently missing (coll. with TM PC);
6. Contact Registration Form: public and available from the main menu.

New EMBnet logo

The need to design a new logo for EMBnet was raised by our new global dimension and goals for “21st Century EMBnet” largely discussed and agreed at the AGM in 2010. From 2010, EMBnet decided to adopt a new brand and strapline. The occasion of EMBnet sponsorship at the “2011 ISCB Africa ASBCB Conference on Bioinformatics” prompted the development of a new logo design that could better represent the spirit and essence of EMBnet in its new global dimension. The P&PR PC has proposed that, from 2011, this new logo is officially approved by EMBnet for any use and purpose in place of the old one (where applicable). This proposal was submitted to the EMBnet Board during the VGM on April 2011, and will be voted on at forthcoming AGM (May 2011) in Portugal.

Figure 2. The new EMBnet logo.

Technical Management Project Committee: AGM2011 report

George Vasilios Magklaras

University of Oslo, The Biotechnology Centre of Oslo, Olso, Norway

Commitee members

Chair/Secretary:	George Magklaras (NO)
Member:	Nils-Einar Eriksson (SE)
Member:	César Bonavides-Martínez (MX)
Member:	Romualdo Zayas-Lagunas (MX)
Member:	Nazim Rahman (PK)
Member:	Harald Dahle (NO)

The following is a summary of the activities for 2011 of the Technical Management Project Committee (TMPC):

- **re-organisation of the website**: the TMPC members have worked to migrate all the public websites from the Swedish server to a new Linux server in Norway. In addition, an upgrade of the OJS system was performed, in order to aid the subscribers of the EMBnet.journal with new functionality. A parallel activity was an action plan to upgrade the structure and look of the EMBnet website, following the suggestions of the established website task force. The migration of services was successful, pending the technical issue of replicating the website for redundancy purposes. The re-design and site replication processes are in progress;
- **testing and deployment of a new video conferencing suite**: TMPC members have deployed a test OpenMeetings server. A small number of EMBnet VGMs was held and the server had problems with certain people connecting. EMBnet tried alternative solutions that also experienced similar reliability problems. As a result, TMPC needs to investigate the issue further and provide a solution to the problem;
- **general support of DNS and helpdesk operations**: in the period June 2010 – May 2011, the TMPC ran 38 DNS change requests and 3 HPC/system support issues for various members around the world.

The TMPC would like to thank Harald Dahle for his work. Harald is working with a new UiO institute and he steps down from the TMPC.

Our thanks extend also to Emil Lundberg for his work on various DNS issues in Sweden.

Consequently, the TMPC has two more vacancies to be considered during the AGM.

The Chair/Secretary would like to thank ALL the TMPC members for their work and dedication for this period.

CCSIS specialist EMBnet node: AGM2011 report

Alex Patak

Molecular Biology and Genomics Unit, Institute for Health and Consumer Protection, Ispra, Italy

Central Core Sequence Information System, Molecular Biology and Genomics Unit, Institute for Health and Consumer Protection, Joint Research Centre - European Commission.

The Bioinformatics Competence Groups (BCG) of the Molecular Biology and Genomics Unit (MBG) are dedicated to the collection and organisation of Genetically Modified Organism (GMO) sequence data.

GMO sequence data are not freely available and are provided to the “European Union Reference Laboratory for GM Food and Feed” (EURL-GMFF) and to “European Food Safety Authority” (EFSA), under confidentiality agreement, by applicant Biotech companies as part of the European Union authorisation procedure for the commercialisation of GMO Food and Feed.

The BCG activity is mainly dedicated to give support to the EURL-GMFF in the scientific assessment and validation of detection methods for GM Food and Feed, as part of the EU authorisation procedure, by providing a GMO sequence database integrated with bioinformatics tools (Central Core Sequence Information System, CCSIS).

CCSIS supports policy by implementing the COMMISSION REGULATION (EC) No 641/2004 which foresees following:

- the applicant shall submit the full sequence of the insert(s), together with the base pairs of the host flanking sequences needed to establish an event-specific detection method;
- the CRL shall enter these data in a molecular database;
- by running homology searches, the CRL will thus be in a position to assess the specificity of the proposed method;

Today, the CCSIS is not only used by the EURL-GMFF, but also by the:

- GMO Unit of EFSA for the risk assessment for food and feed safety;
- scientific staff of the MBG Unit in a variety of research projects.

The management of the CCSIS can be divided in two parts: *collection of GMO DNA sequence data* and *system administration of the computing facilities*.

Collection of GMO DNA sequence data

The data stored on the CCSIS comprises a) local copies of public sequence databases, and b) in-house data.

Local copies of public sequence databases

The use of local copies of public databases, like GenBank, protects the confidentiality of the GMO sequence data during the analysis, and improves calculation speeds. Local copies are regularly updated and extended with new datasets when required. The rapidly growing amount of publicly available sequence data is a great challenge for the computing platform, especially for the local storage capacity and network resource needed to copy the data.

In-house data

The most valuable of the BCG activities is the GMO sequence data that have been collected from documents submitted to EURL-GMFF and to EFSA, and complemented with information extracted from GenBank and EMBL. The CCSIS database is a unique source of GMO sequence data world-wide, as this information in not publicly available, but is provided confidentially to the JRC by Biotech companies. The sequences provided to the BCG are verified and manually annotated with all relevant information extracted form the accompanying documents. By following international annotation rules, the sequence records are compatible with most of the available bioinformatics tools.

The new GMO Detection Methods database[1] has been released on the EURL-GMFF website. This is a publicly accessible database that con-

---

[1] http://gmo-crl.jrc.ec.europa.eu/gmomethods/

tains all validated GMO detection methods that the MBG Unit published in the "Compendium of validated GMO Detection Methods" (JRC Reference Report).

System Administration of the computing facilities

The CCSIS runs on a dedicated high performance computing platform. This platform is based on the clustering of several compute nodes using grid computing. Its main use is the alignment of DNA sequences. The grid-computing set-up allows many users simultaneously (high-throughput computing) to perform complex sequence alignments with and acceptable computing speed.

The Apple Workgroup Cluster has 4 Xserve G5 nodes and Xserve RAID with 4 TB of capacity. Upgrade of the hardware platform has been planned and is ongoing. This includes acquisition of new hardware, update to last Operation System versions, and a new version of the iNquiry software have been prepared on a new prototype bioinformatics cluster.

The MRS Sequence Retrieval System will be separated from the cluster and installed on an Ubuntu Linux-based Hewlett-Packard server with enough RAM memory to cope with the continuously growing local copies of public sequence databases.

Chinese EMBnet node: AGM2011 report

Jingchu Luo
Centre of Bioinformatics, Peking University, Beijing, China

It has been 15 years since we joined EMBnet, in 1996, as the national node of China. Great change has been seen in the past couple of years within our organisation. Several national and specialist nodes had to resign from EMBnet for various reasons. With support from several funding sources, we are trying our best to work hard in bioinformatics service, education and development. Our main funding sources are grants from the Ministry of Science and Technology, and Peking University special programs dedicated to hardware purchase. Our computing power has increased dramatically during the past three years.

We have been collaborating for years with EBI colleagues to make our FTP server the central repository for database downloading, and the official redistribution site for the EBI databases, including EMBL releases and the Bio-Mirror project. Owing to changes in the mirroring policy, we have had to stop running the ExPASy mirror. As reported previously, and described in the EMBnet Newsletter, we have been running an Applied Bioinformatics course[1] for graduate students from both Peking University and the Chinese Academy of Agricultural Sciences each semester. More than 200 students are taking the course currently.

We have updated the plant transcription factor (TF) database[2] to version 2.0, which contains predicted TFs from 49 species, classified into 59 families. Several other databases have been constructed or updated, including the leaf senescence database[3] and the Arabidopsis hormone database[4]. We are work-

---

[1] http://abc.cbi.pku.edu.cn/
[2] http://planttfdb.cbi.pku.edu.cn/
[3] http://psd.cbi.pku.edu.cn/
[4] http://ahd.cbi.pku.edu.cn/

ing on a new interface for the Web-based bioinformatics platform WebLab[5], which is becoming a useful analysis platform. We have developed a rice genome browser[6], which was published in BMC genomics. The framework for a general genome browser is being developed, and will be freely available late this year. In addition to service and development, we are also undertaking research in several directions, including plant comparative genomics.

---

[5] http://weblab.cbi.pku.edu.cn/
[6] http://www.ricemap.org/

French EMBnet node: AGM2011 report

Guy Perriere

Pôle Rhône-Alpes de Bioinformatique, Université Claude Bernard, Lyon, France

The Réseau National des plates-formes en Bioinformatique (ReNaBi) is the French EMBnet node since 2008. Present head of ReNaBi is Claudine Médigue (Génoscope, 2 rue Gaston Crémieux, 91057 Evry Cedex)[1] and its delegate for EMBnet is Guy Perrière. This structure is a national network of six regional bioinformatics centres, each centre gathering a variable number of individual platforms (PFs).
URL: http://renabi.genouest.org/

Presently, those centres are:

APLIBIO[2]– Coordinator Ivan Moszer. The APLIBIO regional centre currently brings together eight PFs located in Paris (Institut Curie, GENATLAS, Institut Pasteur, RPBS) and in the Paris area (eBio, MicroScope, MIGALE, URGI) with a staff of 80 people.

PRABI[3] – Coordinator Guy Perrière. The regional platform PRABI currently gathers five PFs located in the Rhône-Alpes region (PRABI-Doua, PRABI-Gerland, PRABI-Lyon Sud, PRABIG and INCa-Synergie) with a staff of 25 people.

ReNaBi-GO[4] – Coordinator Olivier Collin. The ReNaBi-GO regional centre currently gathers three PFs located in the western of France (GenOuest Rennes, ABiMS Roscoff, BiRD Nantes) with a staff of 20 people.

ReNaBi-NE[5] – Coordinator Maud Pupin. ReNaBi-NE currently gathers five PFs located in the North-East of France (CIB in Lille, BIPS in Strasbourg, MBI and SIDR in Vandoeuvre-lès-Nancy, Reims bioinformatics PF), with a staff of 39 people.

---

[1] http://www.genoscope.cns.fr/spip/spip.php?lang=en
[2] http://renabi.genouest.org/platforms/aplibio/
[3] http://www.prabi.fr/
[4] http://renabi.genouest.org/platforms/renabi-grand-ouest/
[5] http://renabi.genouest.org/platforms/renabi-nord-est/

ReNaBi-SO[6] – Coordinator Christine Gaspin. The ReNaBi-SO regional centre brings together 30 people from teams, platforms and a biotechnology company (BIOGEMMA) located in Toulouse and Bordeaux.

ReNaBi-GS[7]– Coordinator Olivier Gascuel. The ReNaBi-GS regional centre currently gathers four PFs located in the South of France (IGS Marseille, CAZy Marseille, IMGT Montpellier, and ATGC Montpellier, Nice and Perpignan) with a staff of 42 people.

There is a long established practice of collaborative work amongst the different PFs that has always been strongly encouraged by ReNaBi, with the support of IBiSA (Infrastructures en Biologie Santé et Agronomie[8]), a French agency in charge of funding PFs for biological research. For instance PFs from the PRABI and ReNaBi-GS regional centres have developed phylogeny-related tools. Pipelines to compute multiple alignments, build trees and view the results have been integrated and made available in the form of a phylogeny service[9]. Another project, RNAspace[10] has brought together three PFs from the regional centres ReNaBi-SO, APLIBIO and RENBI-NE with the aim of developing an integrated tool for annotating small non coding RNAs.

From a more technological standpoint, the BioMAJ project[11]has gathered together the PFs from ReNaBi-GO and APLIBIO for working on the development of a fully automatic management system for biological databanks (*i.e.*, update cycle and supervision of bank repositories mirrored on local servers). Several PFs from different regional centres (APLIBIO, ReNaBi-GO and ReNaBi-GS) are involved in the MobyleNet project[12]that aims at facilitating the interoperability of specialised tools through the provision of a network of web portals.

One important assignment of ReNaBi consists in optimising and rationalising the use of computing facilities on a national scale. The GriSBi initiative (Grid Support for Bioinformatics[13]) has as principal objective to federate the computing resources of the French bioinformatics platforms in a distributed infrastructure for the bioinformatics community at the national level. It involves a number of PFs from ReNaBi-SO, ReNaBi-GO, ReNaBi-NE, PRABI and APLIBIO, and is supported by the national coordination of French grid initiatives (Institut des Grilles[14]).

Collaborative methodological developments also extend to "neighbour" fields. For instance, the ProteHome project[15] groups two bioinformatics PFs (from APLIBIO and PRABI) and three proteomics PFs (located in Rennes, Toulouse and Grenoble). The project purports to provide services and resources to the scientific community involved in proteomics analyses.

Owing to the fact that ReNaBi gathers many PFs that have a broad range of activities, the computing services offered cover the whole spectrum of bioinformatics:

- access to primary or specialised sequence databases;
- alignment and similarity search programs;
- general sequence analysis packages;
- biostatistics packages;
- molecular phylogeny and molecular analysis programs;
- tools for proteomics and transcriptomics data analysis;
- tools for protein structure prediction and modelling;
- Next Generation Sequencing (NGS) specific programs and pipelines.

All the platforms are independent, for this reason it is not possible to give a global financial assessment. The ReNaBi itself receives an yearly funding of € 50,000 from IBiSA in order to support the organisation of workshops, conferences or thematic networks. For the 2011-2015 time frame, the ReNaBi will receive an extra support of € 5,000,000 from the ANR (Agence Nationale de la Recherche[16]), the main national funding agency for research, for NGS data analysis. This funding will be distributed to the six regional centres listed above.

As for machines available accross ReNaBi, because of to its network structure, it is difficult to give the complete list. Standard equipment for a typical ReNaBi regional centre consists usually in a medium size computing cluster with about 300-700 cores, a few disk servers, and a large set (of variable size) of micro-computers.

---

[6] http://renabi.genouest.org/platforms/renabi-sud-ouest/

[7] http://renabi.genouest.org/platforms/renabi-grand-sud/

[8] http://www.ibisa.net/

[9] http://www.phylogeny.fr

[10] http://www.rnaspace.org

[11] http://biomaj.genouest.org/

[12] http://mobylenet.rpbs.univ-paris-diderot.fr

[13] http://www.grisbio.fr

[14] http://www.idgrilles.fr

[15] http://www.grenoble-prabi.fr/protehome

[16] http://www.agence-nationale-recherche.fr/en/

Italian EMBnet node: AGM2011 report

Domenica D'Elia
CNR-Institute for Biomedical Technologies, Bari, Italy

ITB, the Institute of Biomedical Technologies, is an institute of the Italian National Research Council[1] (CNR); it is composed by 4 sections, located in Milano (from where the Institute is directed), Bari, Pisa and Padova. The ITB-Bari (Bioinformatics and Genomics) is the national node of EMBnet in Italy: Domenica D'Elia is the node manager, and Andreas Gisel is a regular member.

URL: http://www.ba.itb.cnr.it

Infrastructure

ITB-Bari hosts the infrastructures for wet and dry laboratories. The wet lab is equiped with instruments for molecular biology studies, cell culture, light microscopy, citofluorimetry and deep sequencing (454 GSFLX Roche). The bioinformatics platform consists on a **Sistem Server HP** Proliant DL 580 (4 CPU Quad-Core Xeon, 32 GB RAM 2.4GHz, storage 6 Tera), **Server Cluster HP** Proliant DL 380 (4 CPU Quad-Core Xeon 2.4GHz, 15 nodes with 30 CPU, Storage 2 Tera), **GSFLX Titanium Cluster** (1 Head node with 2 Intel dual core 2.33GHz, 16GB RAM 4 Node slave with 2 Intel dual core 2.33GHz, 4 GB RAM for each nodes) and **Development Workstation e4** (2 CPU Quad-Core Xeon 2.4GHz, 2 Tesla c1060 Video cards with CUDA, 16 GB Ram, Storage 8 Tera). Further, ITB-Bari is part of the INFN production GRID[2], and has within the Virtual Organisation 'bio', several thousand CPUs at its disposal.

Research Activities

The bioinformatics activities of ITB-Bari are mainly focused on the development of bioinformatics tools for the management and analysis of Next Generation Sequencing data (NGS). In addition, investigations using data-mining tools for genome-wide binding-site analysis, and development of algorithms in gene ontology applications are running. NGS activities focus on the development of algorithms for RNA-seq and analysis of small RNAs, as well as the provision of analysis services on these topics. Further, we are developing reference sequence databases for different themes, such as non-coding RNA to efficiently support NGS data analysis.

Research lines include experimental and bioinformatics studies of transcriptional and post-transcriptional regulatory processes of genes involved in the pathophysiology of cell proliferation, such as the characterisation of new modulators of p53 family member stability and activity; functional analysis and genome-wide investigation of the transcriptome profile for the identification of molecular biomarkers; *in vivo* investigation of the clustering of isoform expression in endometrial tissue. ITB-Bari has also developed a new protocol for the preparation and amplification of representative and strand-specific cDNA libraries for NGS applications (patent number: RM2010A000293). Metagenomics studies are oriented towards evaluation of the taxonomical and functional complexity of environmental microbial communities. The above-mentioned research activities are supported by different national and European projects. The bioinformatics tools and resources of ITB-Bari are available at: http://bioinformatics.ba.itb.cnr.it/.

Networking

ITB-Bari is part of the core organisation for the establishment of an Italian network on NGS data-production research centres[3], an initiative started by agreements established during the 2010 NGS workshop organised in Bari with the support of the CNR-Department of Life Sciences. A Web portal affiliated to COST Action BM1006[4] has also been created: "Italian NGS Network for Data Analysis"[5].

Dissemination

From 2008 to 2010, ITB-Bari was responsible for organising the following EMBnet AGMs, conferences and workshops in Bioinformatics:

---

[1] http://www.cnr.it

[2] http://grid.infn.it

[3] http://ngsnetwork.ba.itb.cnr.it/

[4] http://w3.cost.eu/index.php?id=212&action_number=BM1006

[5] http://www.nextgenerationsequencing.it/

- EMBnet AGM 2008 (Martina Franca) and AGM 2010 (Ruvo di Puglia);
- International Bioinformatics Conference to celebrate the 20th Anniversary of EMBnet (Martina Franca, 2008);
- Seventh Annual Meeting of the Bioinformatics Italian Society, BITS 2010[6];
- NGS data-analysis workshop, 2009[7] (Rome);
- NGS workshop, First Edition, 2009[8] (Bari);
- NGS workshop, Second Edition, 2010[9] (Bari);

The Bari NGS Workshop, in its third edition[10], will be organised in October 2011.

We published many articles in local and national magazines, and released interviews on local TV stations talking about bioinformatics and molecular biology research activities and their social impact on human health and the environment; and EMBnet activities and missions (see links to local and national magazines in: http://www.embnet.org/en/EMBnet20thAnniversary; http://www.artventuno.com/).

Education and Training in Bioinformatics

The ITB-Bari organises courses and tutorials in bioinformatics at local and national levels. Tutorials and hands-on courses were organised as satellite events at conferences and workshops listed above. We also organise Masters courses as part of collaborative agreements with Regional Associations, such as the Master in Bioinformatics and Molecular Epidemiology Techniques Applied to Infectious Diseases (2011-2012), under the Memorandum of Understanding signed with the ABAP – Associazione Biologigi Ambientalisti Pugliesi[11]. ITB-Bari is also active in projects for the dissemination of scientific culture in molecular biology, biotechnology and bioinformatics in middle and high schools, and participates in national and regional Fairs and Festivals of Innovation (December, 2010[12]; September 2011[13]).

---

[6] http://bits2010.ba.itb.cnr.it/

[7] www.nextgenerationsequencing.org

[8] http://mi.caspur.it/workshop_NGS09/

[9] http://mi.caspur.it/workshop_NGS10/

[10] http://mi.caspur.it/workshop_NGS11/

[11] http://www.infoabap.it/ABAP/ABAPeng.html

[12] http://www.ba.itb.cnr.it/wp/category/eventi-2/

[13] http://www.ba.itb.cnr.it/wp/2011/10/%E2%80%9Clight-2011-la-notte-dei-ricercatori%E2%80%9D-porte-aperte-al-futuro/

Portuguese EMBnet node: AGM2011 report

Pedro L Fernandes

Instituto Gulbenkian de Ciência, Oeiras, Portugal

The activity of the Portuguese node of EMBnet is split in two areas: service provision and training.

Service Provision

A re-equipment grant (REEQ/741/BIO/2005 from Fundação para a Ciência e Tecnologia, the national funding agency for science) has allowed the Instituto Gulbenkian de Ciência to move in the direction of High Performance Computing (HPC). The Portuguese node of EMBnet migrated its users into the new computing platform, a 60-node Linux cluster based on IBM Blade Centers in March 2006. Satisfying the needs of about 230 users encompassed two spheres of action: one for users who were not involved in HPC itself but wanted to remain operational in their day-to-day business; and another for 'power users' who wished to implement solutions using the parallel architecture features (via MPI). We continued to provide full support for EMBOSS and other non parallelised code via the Inquiry Portal, and everybody gained availability as the use of non-concurrent Linux machines came to the rescue. The portal offers several flavours of BLAST that use the MPI and run on up to 8 blades in parallel, and integrates locally developed software seamlessly, to the point of automatically generating local web services. The cluster is used to support long-term simulations via ssh, which is mostly used for Population Genetics simulations. Management of the cluster is automated by scripts and by monitoring tasks that require very little attention [Fernandes, 2006]. The cluster usage requires minimal maintenance in terms of replacement of parts, which is eased by the redundancy of the hardware and electrical power. The installation is named Centro Português de Bioinformática – Recursos de Alta Prestação.

Figure 1. A practical, "Proteomics Data Analysis" course in 2010.

Training

The IGC has a permanent Bioinformatics Training Programme (GTPB) that has been running regularly since 1999. This Programme [Fernandes, 2010] presently holds 16 hands-on training courses per year. The attendance is international, and though the travel costs make it more affordable for local trainees, it is more widely accessible, as local lodging is also available at very low rates. Care has been taken to generate expansion plans and testing solutions for handling remote audiences, b-learning, and lecture capture. Future plans include testing telepresence solutions to better approximate face-to-face situations. The GTPB offers an Introductory course once or twice per year, several thematic courses and "Foundations" type courses, aimed at recycling fundamental knowledge like Statistics, Database Design, *etc.* We also offer a "Bioinformatics using Python for Biologists" course, which fosters the development of programming skills using a thematic set of bioinformatics problems. Approximately 300 people now benefit from this programme annually, bringing the total to ~2,000 trainees in the last 12 years.

The activities of the Portuguese node contribute to the integration of users, and communities of users, in global communities and activities; EMBnet-originated software, Quick Guides, newsletter articles, etc. are frequently used. The educational experiments that are under way in GTPB are shared via the planned activity of the ET PC, and it is foreseeable that the resulting actions can benefit the whole EMBnet community.

Pedro Fernandes, the node manager, has served as chair in the P&PR PC (1999-2010) and as chair of the ET PC since June 2010. He is a founding member of the BTN, a thematic group that studies Bioinformatics Training techniques [Schneider *et al.*, 2010].

Figure 2. The new Bioinformatics Training Room at the Instituto Gulbenkian de Ciência.

Publications

Fernandes, P.L. (2006) Single handed node management. EMBnet.news 12, 19-20.

Fernandes, P.L. (2010) The GTPB training programme in Portugal. Brief. Bioinform. 11, 626-634, doi:10.1093/bib/bbq063.

Schneider, M.V., Watson, J., Attwood, T.K., Rother, K., Budd, A., McDowall, J., Via, A., Fernandes, P., Nyronen, T., Blicher, T., Jones, P., Blatter, M-C., De Las Rivas, J., Judge, D.P., van der Gool, W. and Brooksbank, C. (2010) Bioinformatics training: a review of challenges, actions and support requirements. Brief. Bioinform. 11, 544-551, doi:10.1093/bib/bbq021.

Websites

- http://hermes.igc.gulbenkian.pt
- http://gtpb.igc.gulbenkian.pt

Annex

List of GTPB courses in 2010:

- Molecular Evolution, Phylogenetics and Adaptation
- Introductory Bioinformatics (First Course)
- Bioinformatics using Python for Biologists
- Hunting for genes and promoters

- Microarray Data Analysis using GEPAS and Babelomics
- Automatic Functional Annotation and Data Mining
- Macromolecular NMR assignment with CcpNmr Analysis
- Analysis and manipulation of phylogenomic data using ETE
- RNA Bioinformatics
- Microarray Data Analysis using R and Bioconductor
- Proteomics Data Analysis
- Structural Genomics and Drug Design
- Pathway Analysis and Drug Targets
- Integrative Cancer Genomics
- Biostatistical Foundation in Bioinformatics
- Introductory Bioinformatics (Second Course)

List of GTPB courses in 2011:

- Genetic Architecture of Complex Traits
- Massive Data Analysis (using Babelomics)
- Automatic Functional Annotation and Data Mining
- Introductory Bioinformatics
- Model-based Inference of Population Structure
- Molecular Evolution, Phylogenetics and Adaptation
- Bioinformatics using Python for Biologists
- RNA Bioinformatics NGS Data Management
- Hunting for Genes and Promoters
- Building Bioinformatics Web Applications
- Knowledge Discovery and Management in Chemoinformatics
- Microarray Data Analysis using R and Bioconductor
- Integrative Approaches in Genomics
- Biostatistical Foundations in Bioinformatics

Russian EMBnet node: AGM2011 report

Sergei Alexandrovich Spirin
Department of Mathematical Methods in Biology, Belozersky Institute of Moscow State University, Moscow, Russian Federation

The Russian EMBnet node is based on the Belozersky Institute of Physical-Chemical Biology, which is a scientific department of Moscow State University (MSU) and is closely connected with the Faculty of Bioengineering and Bioinformatics of MSU.

People

Department of Mathematical Methods in Biology of Belozersky Institute

- Andrei Alexeevski (head)
- Sergei Spirin
- Olga Zanegina
- Eugeny Aksianov

Computing group of Belozersky Institute (former "GeneBee")

- Vladimir Nikolaev (head)
- Alexander Kravchenko
- Andrei Leontovich
- Lidia Smirenina

Hardware

- Server Prime Power 850 (16 processors SPARC64 IV GP 810 MHz, 2 TB disks, Sun Solaris);
- a number of servers under Linux;
- a number of workstations under Windows XP and Windows 7.

Activity

Education and training

The 95% of the energies were spent to the Faculty of Bioengineering and Bioinformatics of MSU (undergraduate courses: "Practical bioinformatics" for the first and second year of the degree course and "Structural bioinformatics" for the third one). Other educational activity:

a. annual three-day courses in practical bioinformatics for students of MSU specialised in virology;
b. consultations for students of Russian-Armenian (Slavonic) University in Yerevan;
c. annual public lectures (mainly for high school students) in bioinformatics (in frames of so-called "Day of Science");
d. consultations in bioinformatics for "Wet Lab" researchers;
e. consultations and seminars for mathematicians and computer scientists. For example, in Spring of 2010 and of 2011, Sergei Spirin held a special course in bioinformatics for students of Faculty of Computational Mathematics and Cybernetics of MSU.

Software development and support

In 2008-2011 a number of publicly available bioinformatics software was developed.

Among them:

f. Geometrical Core[1] for a family of aligned 3D structures;
g. SheeP[2], an automatic description of beta- and alpha/beta-structural proteins;
h. ConClus[3], conserved hydrophobic clusters in aligned 3D structures of proteins;
i. Malakite:[4] searching for reliable blocks in a given protein sequence alignment;
j. Blocks3D[5]: this service searches for blocks that are supported by 3D superimposition in aligned sequences of PDB proteins;
k. wLake[6]: a tool which allows to identify clusters of a structural water molecules in given superimposed 3D structures.

Also, previously developed software remains supported, in particular:

a. NPIDB[7]: a database of structures of DNA-protein and RNA-protein complexes. A new interface is available at http://npidb.belozersky.msu.ru.
b. AliBee[8]: an original program for multiple alignment;
c. TreeTop[9]: a tool for phylogenetic tree prediction;
d. DotHelix[10]: local similarity map for two biological sequences;
e. SVETKA[11]: simultaneous analysis of a multiple alignment and the phylogenetic tree of the same sequences; search for classifying positions in the alignment;
f. CluD[12]: hydrophobic clusters in 3D structures;
g. GEA[13]: a system for Gene Expression Analysis.

See also http://www.genebee.msu.ru/bioinformatics.html and http://mouse.belozersky.msu.ru/.

Research

Some directions of research in bioinformatics provided in our Institute (those that are not reflected in the above list of new software) are as follows:

a. comparative structural description of several families of proteins[14];
b. drug design;
c. development of algorithms for next generation sequencing data;
d. comparative analysis of performance of phylogeny reconstruction methods.

Funding

The grant of the Russian Foundation of Basic Research No. 10-07-00685 supports software development and installation but not education and training activities. The grant is for 2010–2012, the grant amount for 2010 was 400,000 roubles (~10 000 euro), the grant for 2011 amounts to 350,000 roubles.

Perspective

EMBnet membership

The governing body of the Institute has definitely expressed the intention to maintain the status of Russian EMBnet node. The member fee for 4 years will be paid at the first opportunity.

---

[1] http://mouse.genebee.msu.ru/tools/gc.html

[2] http://mouse.belozersky.msu.ru/~evgeniy/cgi-bin/sheep/sheep.php

[3] http://mouse.genebee.msu.ru/~bennigsen/conclus.html

[4] http://mouse.belozersky.msu.ru/tools/malakite.html

[5] http://kodomo.fbb.msu.ru/blocks3d/

[6] http://mouse.belozersky.msu.ru/~evgeniy/cgi-bin/wLake/wLake.php

[7] http://mouse.genebee.msu.ru/NPIDB/index.html

[8] http://www.belozersky.msu.ru/services/malign_reduced.html

[9] http://www.belozersky.msu.ru/services/phtree_reduced.html

[10] http://www.belozersky.msu.ru/services/dhm/advanced.html

[11] http://monkey.belozersky.msu.ru/~dian/cgi-bin/austerlitz.pl?type=index

[12] http://mouse.belozersky.msu.ru/npidb/cgi-bin/hftri.pl

[13] http://www.genebee.msu.ru/~nikonov/GEA/

[14] http://mouse.genebee.msu.ru/families.html

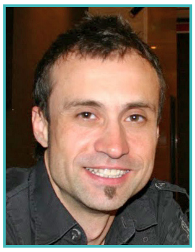
Education

The teaching of undergraduate students will remain the main part of our training activity. We will do our best to increase the number of consultations and possibly to organise bioinformatics courses for "Wet Lab" biologists. Publishing a Russian textbook in bioinformatics is in our plans.

Software development

Hopefully, the intensity of software development (including creation of new public available programs and improvement of old ones) will, at least, not decreased.

Slovak EMBnet node: AGM2011 report



Lubos Klucar

Institute of Molecular Biology, Slovak Academy of Sciences, Slovakia

Since 1999, the Institute of Molecular Biology, Slovak Academy of Sciences, has hosted the Slovak National EMBnet node. Activities of the node are focused mainly on providing access to basic bioinformatics databases and tools, and on training and education in the field of bioinformatics. The node has no direct funding (apart from the EMBnet membership fee). The node actively participates in the management and production of , as well as in the design and development of the embnet.org website. The node is also involved in several international projects, namely: SeqAhead – Next Generation Sequencing Data Analysis Network (COST); and ELIXIR – European Life Sciences Infrastructure For Biological Information (ESFRI).

Staffing

The node possesses two part-time staff members. Lubos Klucar is the national node manager, member of the P&PR PC, and member of the Executive Editorial Board of EMBnet.journal. Matej Stano is a member of the ET PC.

Resources

We utilise 4 PC-based Linux (Slackware) servers, interconnected via SGE, with 8 GB RAM and 3 TB disk space in total. We have 1 Gbps internal network connection, 1 Gbps internal-to-campus network connection, and 10 Gbps campus-to-world connection. We maintain our own DNS domain embnet.sk.

Services

We still maintain the SRS 8.1 database system, with around 50 biological databases installed. This SRS installation is enhanced with EMBOSS, HMMER and ClustalW. Other bioinformatics tools include PHYLIP and the Staden Package. We also maintain two proprietary databases: the phiSITE[1] database of gene regulation in bacteriophages, and the phiBIOTICS[2] database of bacteriophage enzybiotics. Other general services cover mailing lists and a bulletin board system.

Education and Training

We supervise regular one-term undergraduate courses in Bioinformatics, Systems Biology and DNA Arrays at Comenius University, Faculty of Natural Sciences. Other irregular courses are also organised, mainly for PhD students and young scientists. In October 2010, we hosted an EBI SLING Bioinformatics Roadshow, with a focus on Structures, Protein Interactions and Mass Spectrometry-based Proteomics.

---

[1] www.phisite.org

[2] www.phibiotics.org

South African EMBnet node: AGM2011 report

Alan Gilbert Christoffels

South African National Bioinformatics Institute, Cape Town, South Africa

The South African National Bioinformatics Institute (SANBI) is part of the University of the Western Cape[1] situated outside Bellville, near Cape Town. The staff comprises five principal investigators supported by administrative and technical staff. Since 2010, we have seen an increase in graduate student registrations with a total of 18 postgraduate students, representing a range of countries on the African continent including Cameroon (1), Ethiopia (1), Ghana (1), Kenya (4), South Africa (10), Sudan (1).

URL: http://www.nepadst.org/sanbio/

Research

- Analysis and annotation of *Glossina*, the vector for the Tsetse fly;
- molecular epidemiology and dynamics of HIV cohorts in Africa;
- comparison of vector, host and pathogen genomes to deliver unique African knowledge of HIV, malaria, tuberculosis and crops;
- develop and apply understanding of normal and diseased human gene expression to diseases relevant to South Africans.

Funding

Since 2002, SANBI has been a Medical Research Council Unit for bioinformatics capacity development. Recently, we succeeded in a 5-year renewal funding cycle for the Unit. The only bioinformatics research chair in South Africa was awarded to the interim director of SANBI and funds nearly half of the postgraduate research projects.

---

[1] http://www.uwc.ac.za/

EU Collaborative projects

SANBI faculty and post docs have enjoyed funded collaborations through EC-FP projects and currently we serve on the SYSCO programme[2] together with the Institute Pasteur in Tunis and Paris, and the Max Planck Institute in Berlin.

Facilities

The Institute has adequate scientific computing infrastructure, and is the site for a pair of high performance 32 CPU IBM P-690 servers, an 8 CPU Xserve cluster and 4 intel 8-core ubuntu servers which provide a significant proportion of the research computing infrastructure for bioinformatics in Africa. SANBI scientists have workstations and Internet access, as well as backup facilities and disk storage. During August 2010, we relocated to a new Life Sciences Building offering extensive training facilities, a visiting scientist facility and meeting rooms. SANBI is currently a development site for the Galaxy system and provides a development nexus for annotation systems.

Local Training

As part of the undergraduate training programme at the University of the Western Cape, the SANBI staff organises a six-week block of introductory lectures and practicals in bioinformatics to 3rd year undergraduate students in the biological sciences. During 2010 the number of undergraduate students registered for the bioinformatics module increased from 40 to 66. In December 2010, we initiated our first summer vacation internship programme for students registered in computer science, mathematics and statistics. A total of 14 students participated in a four-week research-based summer programme. Two students decided to continue their graduate studies at SANBI during 2011 and one computer science student opted to work part-time in the institute for one year.

National Training

The South African Department of Science and Technology has funded a two-month national training programme in bioinformatics for six years. As of February 2011, this programme has been hosted in the new training facilities at SANBI. The training programme is targeted at new graduated students in South Africa with an annual intake of 25 participants.

During the past two years, we have hosted a popular two-day ENSEMBL workshop in collaboration with the ENSEMBL training team. In 2011, as in 2010, registrations were restricted to 30 individuals. Owing to demand in the previous years, the workshop has been extended to include a more advanced component that deals with access to the ENSEMBL database using the API instead of the browser. The workshop has been mirrored in Johannesburg as part of our attempts to extend the training across South Africa.

Regional Training on the African continent

Over the past four years the International Glossina Genome Initiative (IGGI) consortium, through SANBI, has provided bioinformatics training to African students. In 2009 we established a Glossina Functional Genomics Network. Underpinning these activities were the mentoring of 3 research fellows through short-term research visits. Three research fellows were identified in 2010 based on their research activity on the Trypanosomiasis and on the institutional commitment to support them after the two-month fellowship programme. Namely, Oliver Manangwa (Tsetse and Trypanosomiasis Research Institute, Tanzania), Paul Mireji (Egerton University, Kenya) and Sarah Mwangi (Kenyan studying at SANBI). These candidates were hosted at a Kenya laboratory for population studies, at the Liverpool School of Tropical Medicine and at the Institute of Tropical Medicine in Antwerp (Belgium), respectively. In collaboration with Erik Bongcam-Rudloff, we are providing eBioKit training courses at the Egerton University in Kenya and at the Trypanosomiasis Research Institute in Tanzania, during April 2011.

---

[2] http://www.sysco.com/about-sysco/diversity-internship-program.html

National Nodes

Argentina
IBBM, Facultad de Cs. Exactas, Universidad Nacional de La Plata

Brazil
Lab. Nacional de Computação Científica, Lab. de Bioinformática, Petrópolis, Rio de Janeiro

Chile
Centre for Biochemical Engineering and Biotechnology (CIByB). University of Chile, Santiago

China
Centre of Bioinformatics, Peking University, Beijing

Colombia
Instituto de Biotecnología, Universidad Nacional de Colombia, Edificio Manuel Ancizar, Bogota

Costa Rica
University of Costa Rica (UCR), School of Medicine, Department of Pharmacology and ClinicToxicology, San Jose

Finland
CSC, Espoo

France
ReNaBi, French bioinformatics platforms network

Greece
Biomedical Research Foundation of the Academy of Athens, Athens

Hungary
Agricultural Biotechnology Center, Godollo

Italy
CNR - Institute for Biomedical Technologies, Bioinformatics and Genomic Group, Bari

Mexico
Nodo Nacional de Bioinformática, EMBnet México, Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos

Norway
The Norwegian EMBnet Node, The Biotechnology Centre of Oslo

Pakistan
COMSATS Institute of Information Technology, Chak Shahzaad, Islamabad

Poland
Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa

Portugal
Instituto Gulbenkian de Ciencia, Centro Portugues de Bioinformatica, Oeiras

Russia
Biocomputing Group, Belozersky Institute, Moscow

Slovakia
Institute of Molecular Biology, Slovak Academy of Science, Bratislava

South Africa
SANBI, University of the Western Cape, Bellville

Spain
EMBnet/CNB, Centro Nacional de Biotecnología, Madrid

Sri Lanka
Institute of Biochemistry, Molecular Biology and Biotechnology, University of Colombo, Colombo

Sweden
Uppsala Biomedical Centre, Computing Department, Uppsala

Switzerland
Swiss Institute of Bioinformatics, Lausanne

Specialist- and Assoc. Nodes

CASPUR
Rome, Italy

EBI
EBI Embl Outstation, Hinxton, Cambridge, UK

Nile University
Giza, Egypt

ETI
Amsterdam, The Netherlands

IHCP
Institute of Health and Consumer Protection, Ispra. Italy

ILRI/BECA
International Livestock Research Institute, Nairobi, Kenya

MIPS
Muenchen, Germany

UMBER
Faculty of Life Sciences, The University of Manchester, UK

CPGR
Centre for Proteomic and Genomic Research, Cape Town, South Africa

The New South Wales Systems Biology Initiative
Sydney, Australia

for more information visit our Web site
www.EMBnet.org

EMBnet.journal

ISSN 1023-4144

Dear reader,

If you have any comments or suggestions regarding this journal we would be very glad to hear from you. If you have a tip you feel we can publish then please let us know. Before submitting your contribution read the "Instructions for authors" at http://journal.EMBnet.org/index.php/EMBnetnews/about and send your manuscript and supplementary files using our on-line submission system at http://journal.EMBnet.org/index.php/EMBnetnews/about/submissions#onlineSubmissions.

Past issues are available as PDF files from the Web site:
http://journal.EMBnet.org/index.php/EMBnetnews/issue/archive

Publisher:
EMBnet Stichting
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

Submission deadline for next issue:
15 february 2011