EMBnet.journal

Volume 19 Nr. 1
August 2013

- 2013 EMBnet Annual General Meeting
- 'Next NGS Challenge – Data Processing and Integration' Conference
- Quick Direct-method Controlled (QDC): a simulator of metabolic experiments and more...

Editorial

Dear reader, welcome to a new issue of *EMBnet.journal*, the international, open access, peer-reviewed bioinformatics journal.

In this year, 2013,EMBnet, the Global Bioinformatics Network, celebrates its Silver Anniversary. You can read more about EMBnet's achievements, alliancesand collaborations during this exciting journey, especially in the last three years,in an article written by EMBnet'sChair, Professor Teresa Attwood.

Speaking of collaborations, we report hereon the highly successful conference in Valencia, Spain, dedicated to advances in next-generation sequencing. So popular is this conference that EMBnet, ISCB and EU COST ActionsSeqAhead and Stategranow intend to organise a similar event annually.

In terms of alliances, we report on EMBnet's work in spear-heading a new initiative to create a '*Global Organisation for Bioinformatics Learning, Education & Training*', GOBLET, together with more than 20 international and national societies, networks and institutes from around the world. Also, to build a stronger alliance with the International Society for Biocuration (ISB), weintroduce a new feature in this issue, '*ISB Spotlight*', highlighting some of the work of ISB's biocurators andshowcasing their work more broadly amongst EMBnet's communities.

EMBnet.journal will this year produce one more full issue, 19.2, and two Supplements, marking with this its fourth year of existence as a peer-reviewed journal and successor to *EMBnet.news*,which started in July 1994.

EMBnet.journal Editorial Board

Protein Spotlight (ISSN 1424-4721) is a periodical electronic review from the SWISS-PROT group of the Swiss Institute of Bioinformatics (SIB). It is published
on a monthly basis and consists of articles focused on particular proteins of interest. Each issue is awvailable, free of charge, in HTML or PDF format at http://www.expasy.org/spotlight.
We provide the EMBnet community with a printed version of issue 149. Please let us know if you like this inclusion.

Cover picture: Great Spotted Woodpecker (*Dendrocopos major*) described by Linne in 1758. Rotyxvägen, Uppsala, Sweden. [© Erik Bongcam-Rudloff]

Contents

ISBE – set out for a Systems Biology Infrastructure for Europe



Susanne Hollmann[1], Vitor AP Martins dos Santos[2], Babette Regierer[2,3]

[1]University of Potsdam, Potsdam, Germany
[2]Wageningen University, Wageningen, The Netherlands
[3]LifeGlimmer GmbH, Berlin, Germany

Received 30 June 2013; **Published** 19 August 2013

Systems biology requires the availability, co-ordination and simultaneous interaction of a large number of diverse facilities and activities. These cover an entire spectrum, from mathematical modelling, through biological, biomedical and clinical experiments, to dedicated technology development. The systems biology community needs close cooperation with data-generation groups and bioinformaticians to define a strategy for producing life-science data of sufficiently high quality for model generation. For each medical, biological or biotechnological problem addressed, the optimal combination of facilities and activities is likely to be different. The complexity of biological systems, and the diversity and dynamics of their processes, means that a full analysis is far too complex to be handled by a single entity, industry or country – a variety of specialist expertise and facilities are typically necessary to achieve results suitable for modelling. Systems-level approaches for tackling the complexity of life-science data provide a profound conceptual advance compared to reductionist biological research methods of the past. Rather than focusing on individual laboratories, specialising in a limited number of research technologies, the Infrastructure for Systems Biology in Europe (ISBE) will facilitate the synergistic application of a wide range of research techniques and technologies to problems of major medical and biotechnological importance.

The life sciences are both highly diverse and highly fragmented; initiating and supporting community-building activities to create efficient links and mutual cooperation between them is therefore a major issue. A first cross-disciplinary "community building" event in Amsterdam, in November 2012, initiated by the European Coordination Action "AllBio - Broadening the Bioinformatics Infrastructure to Unicellular, Animal and Plant Science[1]", discussed the current challenges in bioinformatics and future needs in the life sciences. Researchers and coordinators of various life science EU initiatives and COST Actions joined the meeting to identify relevant interfaces and mutual problems. As the participants introduced their projects and activities, it became apparent that the communities they represent are still working quite distantly from each other, and it will need a joint effort to establish sufficient communication and interaction between all parties and relevant stakeholders.

To gain a better understanding of the functioning of cells, tissues, organs and whole organisms, it is clear that the biology of the future will necessarily address questions on a large scale using systems biology approaches. This will require the careful building of efficient pipelines, involving data acquisition, analysis, processing, integration, and finally modelling. Utilisation of the information will be the next step towards acquiring and transferring the knowledge gained, for example to clinics, to animal breeding programmes, to pest control in crop plants, and to numerous other new and existing applications. The mathematical and formal modelling of molecular, physiological and anatomical data derived from various biological, biomedical and clinical experiments still present enormous hurdles, and targeted technology development is still needed to close existing analytical gaps.

Based on opportunities offered by national and European funding programs, numerous excellent initiatives and projects have contributed to the current paradigm shift in modern life sciences, preparing the ground for understanding fundamental aspects of life via systems biology. Now it is time to integrate knowledge and expertise across many disparate scientific areas, to create the efficient, cooperative links necessary to amalgamate them successfully.

1 http://www.allbioinformatics.eu

ISBE is a new initiative to bring such diverse projects together in a single concerted action, to offer their combined wisdom and expertise to Europe's scientific community. Based on the activities of AllBio and other initiatives, ISBE will provide not only a new infrastructure, but also the opportunity to identify links and cross-cutting themes in order both to advance science and to avoid redundancies – it is a unique chance to unite efforts for the benefit of the whole life science community in Europe.

To create a European infrastructure for systems biology, a distributed, but closely interconnected framework will be established that combines data generation, data stewardship (data processing, analysis and curation), and data integration and modelling. Centres across Europe will apply and develop expertise in model-driven knowledge generation, and make this expertise freely available to the community.

The future success of the European economy will depend strongly on our ability to enhance existing and to develop new knowledge-based industries. The ISBE infrastructure will make a considerable impact on European industry, business and society through the systematic study of complex biological processes, incorporating expertise from physics, engineering, biology, chemistry, computer science and mathematics. This will lead to new and important technological applications: for example, in health, agriculture, food production, clean energy, and other areas related to Europe's growing bio-based economy. In particular, the ISBE infrastructure will be highly relevant in transforming our basic knowledge of complex molecular systems into the new arena of predictive, preventive and personalised medicine that reaches across Europe and beyond. The systems biology approaches developed in the ISBE infrastructure will provide new insights and will assist development of tools for designing new medical, biotechnological and environmental applications. ISBE will also transform systems biology into an integrated, pan-European activity by the use of Web-based experimental facilities and on-line mathematical modelling, and make the biological sciences more productive and cost-effective. This is important for European society, industry and economy in genera.l

Acknowledgements

ISBE[2] is a European ESFRI Infrastructure project comprising 23 partners from 11 countries, and is financially supported by the EC under FP7.

ISBE Infrastructure for Systems Biology Europe

2 www.isbe.eu

MIRRI - The Microbial Resource Research Infrastructure: managing resources for the bio-economy

Manuela Schüngel[1], Erko Stackebrandt[1], Chantal Bizet[2], David Smith[3]
[1]MIRRI – Microbial Resource Research Infrastructure, Leibnitz, Germany
[2]MIRRI – Microbial Resource Research Infrastructure Collection de l'Institut Pasteur, France
[3]MIRRI – Microbial Resource Research Infrastructure CAB International, United Kingdom

Received 30 July 2013; **Published** 19 August 2013

Public website: www.mirri.org
MIRRI coordinator: erko@dsmz.de

The term Microbial Resource Centre (MRC) has been introduced to describe the traditional public culture collections (CCs) that are managed under a quality system. Europe has a few of these MRCs that complement the landscape of facilities accessing, monitoring and providing microbial resources. They communicate within ECCO (European Culture Collections' Organisation), and several of them have been networking for more than 30 years with respect to research and operational issues.

Public CCs of microbial material are the key providers of resources for a broad spectrum of users: from researchers in academia to those in bio-industry – all depend on the wealth of microorganisms stored in these collections. Unfortunately, this landscape in Europe is fragmented: smaller public CCs and MRCs do not have harmonised strategies for collection focus, delivery of associated information, bio-security or legal aspects. Although access to holdings is available for some larger collections in Europe (see CABRI project[1]) and worldwide the World Data centre for Microorganisms Global Catalogue of Microorganisms[2] holds data on 60 collections in 29 countries, most of the rest are not connected to a common online portal. This leads to under-utilisation of existing resources, which hampers European research in many fields: it isn't just academic research, but also bio-industry that requires access to authentic well-characterised microbial resources. The microbial raw material, and their genetic and metabolic products, are utilised in many areas: production of healthy and functional food; identification of new antagonists against pathogens; fighting agricultural disease; identifying novel energy sources on the basis of microbial biomass; and screening for new active molecules in the bio-based industries – to mention only a few. Owing to the lack of harmonised approaches, there is duplication of effort, not just for users, who have to perform long and time-consuming investigations before the requested microbial material is found in a database, but also for the CCs, which are often not aware of activities in neighbouring collections.

The G8 Science Ministers' Statement[3], made 12 June 2013, highlighted the need to improve transparency, coherence and coordination of the global scientific research enterprise to address global challenges. They focused on anti-microbial drug resistance as a major health-security challenge, and the need to work together to reduce antimicrobial resistance. They stressed that Research Infrastructures are key elements in research and innovation policies.

MIRRI intends to address these global challenges in many ways, by working with multidisciplinary partners, delivering to them the resources, tools and services needed to facilitate the discovery of solutions. The microbes themselves have most of the answers: after all, they are the great pioneers of our planet, surviving the extremes. Examining organisms from before the introduction of antibiotics and comparing them with resistant strains is the obvious route to improving our understanding, and increasing

1 http://www.cabri.org/
2 http://gmc.wfcc.info/
3 https://www.gov.uk/government/news/g8-science-ministers-statement

Table 1: List of MIRRI partners in the preparatory phase.

Acronym	Full Designation	Country
DSMZ (coordinator)	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH	GE
CABI	CAB International	UK
CSIC-IMEDEA	Agencia Estatal Consejo Superior de Investigaciones Cientificas y University of the Balearic Islands	ES
IAFB-CCIM	Institute of Agricultural and Food Biotechnology , Culture Collection of Industrial Microorganisms	PL
IBPM	All-Russian Collection of Microorganisms, Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences	RU
INRA-CIRM	Institut National de la Recherche Agronomique - Centre International de Ressources Microbiennes	FR
IP	Institut Pasteur	FR
Jacobs Uni	Jacobs University Bremen	GE
KNAW/CBS	Koninklijke Nederlandse Akademie van Wetenschappen - KNAW	NL
MUT	Mycotheca Universitatis Taurinensis	IT
SPP-PS	Service Public Fédéral de Programmation Politique scientifique	BE
UGENT	Universiteit Gent 8.1 UGent - BCCM/LMG 8.2 UGent - Dept. of Applied Mathematics and Computer Science	BE
UGOT	Culture Collection, University of Goteburg	SE
UMinho-MUM	Micoteca do Universidade do Minho	PT
USMI	IRCCS AOU San Martino IST	IT
UVEG-CECT	Universidad de Valencia, Coleccion Española de Cultivos Tipo	ES

our opportunities to discover new possibilities. By understanding the chemistry, and seeking those organisms with the properties needed, be they taxonomic relatives or organisms from particular ecosystems, the coordinated efforts of Microbial Resource Centres can accelerate the process. Well-described microbial resources will play a key role in underpinning the bio-economy and driving economic growth. To do this, there is a need to better utilise microbiological diversity in biotechnology. This is fundamental to the delivery of the bio-economy, and to accelerate the discovery of natural solutions to today's global challenges.

Having recognised the need for improvement, the European Strategy Forum on Research Infrastructure[4] (ESFRI) placed MIRRI on their roadmap in 2010. Having started its preparatory phase at the end of 2012, MIRRI will focus on strategic goals to assure that facilitated uptake of high-quality resources promotes good science and innovation, whilst breaking down the barriers mentioned above. By establishing a user- and quality-driven centre for resources and information, MIRRI will facilitate access to the right microorganisms for the task in hand. Furthermore, by applying coordinated isolation programs, and closely collaborating with similar programs already on-going in the academic environment, existing gaps in the availability of microbial resources will be resolved. Together, these are the prerequisites for enhancing European competiveness in the knowledge-based bio-economy, which gets more and more important as the non-European markets expand.

MIRRI Objectives

In its *preparatory phase*, the MIRRI consortium consists of 16 European partners (see Table 1), as well as 17 European collaborating parties. Based on a history of successful cooperation between most of the participants (*e.g.*, CABRI[5], EBRCN[6],

4 http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri-roadmap

5 http://www.cabri.org/

6 http://www.ebrcn.eu

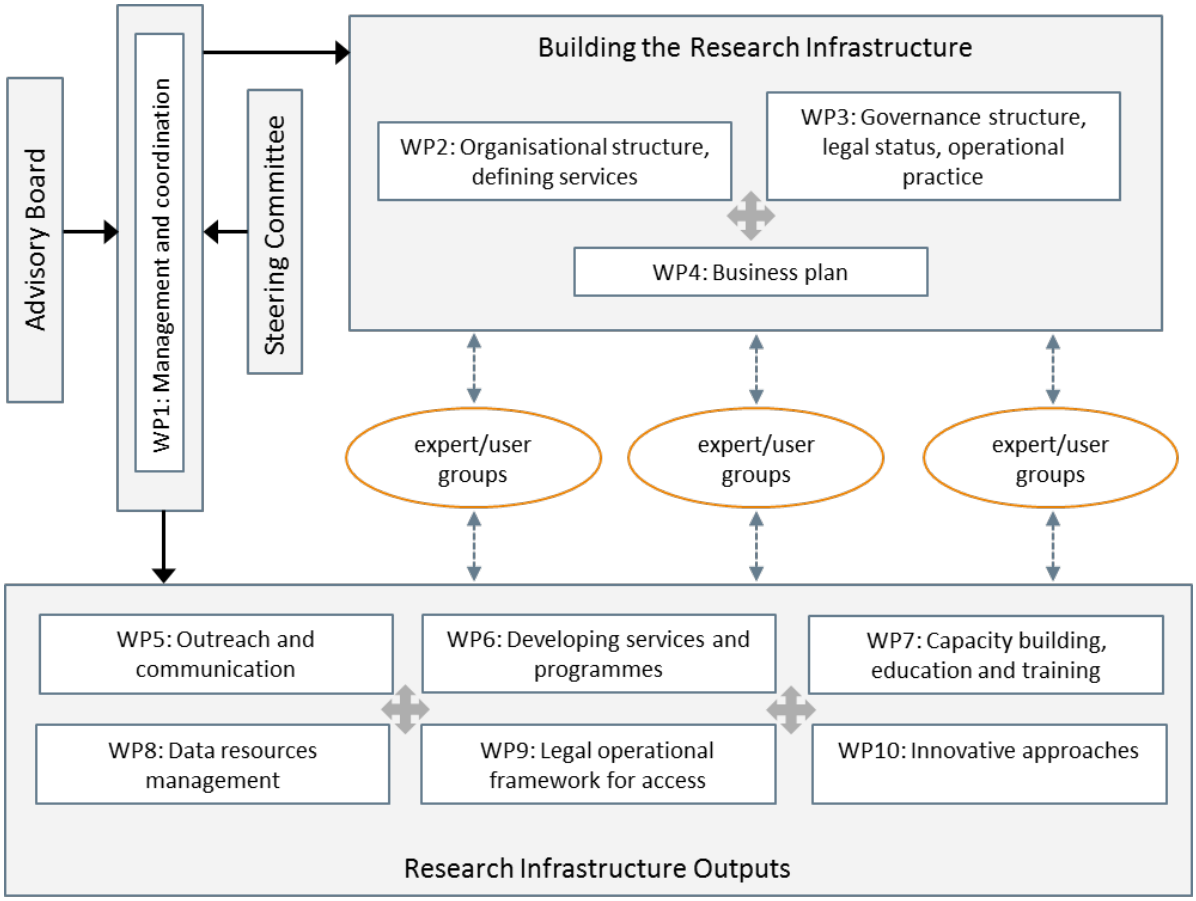


Figure 1. Working scheme of the MIRRI consortium during the preparatory phase.

EMbaRC[7], a well-rehearsed team will address the challenging objectives of the MIRRI project.

The overall objective of MIRRI is to create a quality-driven centre for resources and information to facilitate access to microbial material within Europe. This idea of a European 'onestopshop' to serve users' needs offers not only a significantly broader range of raw material, but also access to their associated data, as well as to expert knowledge in fields of microbiology. A second major goal of MIRRI is to establish a platform for capacity building. This includes both transnational access to facilities and a Masters curriculum for students in microbial systematics and resource management. By offering special training in the field of modern approaches to taxonomy, MIRRI will combat the increasing loss of expertise. Such well-trained, next-generation scientists are necessary to work with bio-industry in Europe, to accelerate discovery processes and innovations.

Strategic Development

During the preparatory phase of MIRRI, strategies will be developed to fulfil the proposed objectives. The workload is distributed across several work packages (see Figure 1), between which close dialogue and cooperation is necessary; more importantly, involvement of different stakeholders is essential (e.g., users/providers of microbial material, bio-industries, policy makers, national authorities) – expert/user groups are a basic requirement for successful strategy development. Several work packages are strongly dependent on stakeholder feedback, especially those dealing with MIRRI's future outputs; they analyse, for example, which resources, expertise and methodological services are requested by users of microbial material (WP6). The field of education and training (WP7) will be strengthened through cooperation with other ESFRI projects, such as

7 http://www.embarc.eu/

ERINHA[8], ELIXIR[9] and EMBRC[10]. Management of data resources (WP8) is one of the most important issues, as access to physical resources is markedly enhanced by the availability of their associated data. This includes the development of tools to extract microbial data from the literature, and development of strategies to allow data interoperability. Bringing together different data-sets will provide an information landscape that will facilitate data-mining and new opportunities for innovation. This leads to open access to an unparalleled wealth of microbial resources within Europe.

MIRRI will create an appropriate legal framework for access and application (bio-risk assessment, as well as bio-security) (WP9) that will ensure user compliance in access and use of biological materials within the plethora of regulations involved. This will include strategies to assure compliance with the Nagoya protocol[11] - *i.e.*, Prior Informed Consent and Access and Benefit Sharing. Further details concerning the work packages can be found on the MIRRI website[12].

The scheduled progress of all work packages is monitored by defined deliverables, to be submitted to the European Commission. Once having entered the *implementation phase* in 2016, the outlined strategies will then be realised.

The proposed future organisational structure of MIRRI is shown in Figure 2. The so-called hub-and-spokes model foresees a central operational hub and several national nodes, each of which is connected to a network of specialised national collections. This model allows a flexible dialogue between hub and spokes. A business plan focusing on a sustainable financing model for the proposed structure is under development, and will be communicated to funders.

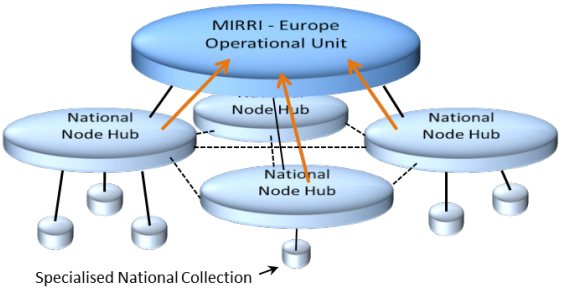


Figure 2. Hub-and-spokes model of the future MIRRI.

On the road towards implementation

Although one of the youngest ESFRI projects, MIRRI has achieved important milestones that will facilitate the step into the *implementation phase*. The management structure chosen for the preparation phase guarantees an open and transparent decision-making process that can be applied to the implementation phase. MIRRI is already well recognised by different stakeholder groups, and their involvement in various aspects of the MIRRI work packages has been achieved. Following the trend of online dissemination, MIRRI established its website and is represented in social media. By using these channels, information can be easily forwarded to all stakeholder groups: young researchers, in particular, will find it easy to communicate with MIRRI via social media. But MIRRI is also using traditional media: the project is presented at conferences and in journals, and, of course, all members have face-to-face meetings with stakeholders. By doing so, the project is moving in the right direction to make an impact on science research and innovation.

As mentioned above, MIRRI's success depends on input and feedback from stakeholders – i.e., users as well as providers of microbial material, bio-industries, policy makers and national authorities. We would like to encourage all of them to contact MIRRI, and to give their valuable insights on their needs and expectations, allowing MIRRI to facilitate discovery for a brighter future.

Acknowledgements

The *preparatory phase* of MIRRI has been financially supported by the European Commission (grant agreement no. 312251).

8 http://www.erinha.eu/
9 http://www.elixir-europe.org/
10 http://www.embrc.eu/
11 http://www.cbd.int/abs/
12 http://www.mirri.org

11th Annual RECOMB Satellite Workshop on Comparative Genomics

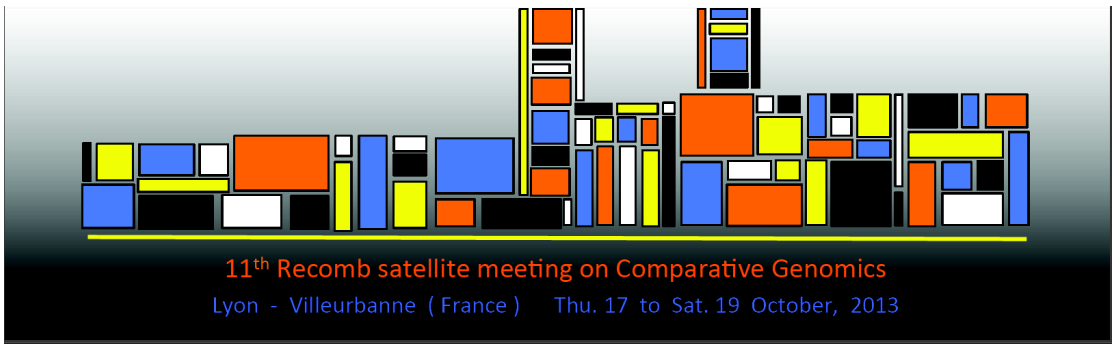
The RECOMB Satellite Workshop on Comparative Genomics aims to provide the premier forum for new computational developments applied to all aspects of comparative genomics, including genome structure and organization, genome function, evolution.

http://rcg2013.sciencesconf.org/

Keynotes speakers:
Tal Dagan, Kiel
France Denoeud, Paris
Laurent Duret, Lyon
Nicolas Galtier, Montpellier
Ludovic Orlando, Copenhagen
Kay Pruefer, Leipzig

Early registration deadline: September 6th

Steering committee:
David Sankoff, Aoife Mclysaght, Jens Lagergren
Program committee chairs:
Macha Nikolski, Yves van de Peer
Organizing committee chairs:
Frédéric Brunet, Eric Tannier



The Global Organisation for Bioinformatics Learning, Education & Training (GOBLET)

GOBLET Consortium

GOBLET Stichting, CMBI Radboud University, Nijmegen Medical Centre, Nijmegen, The Netherlands

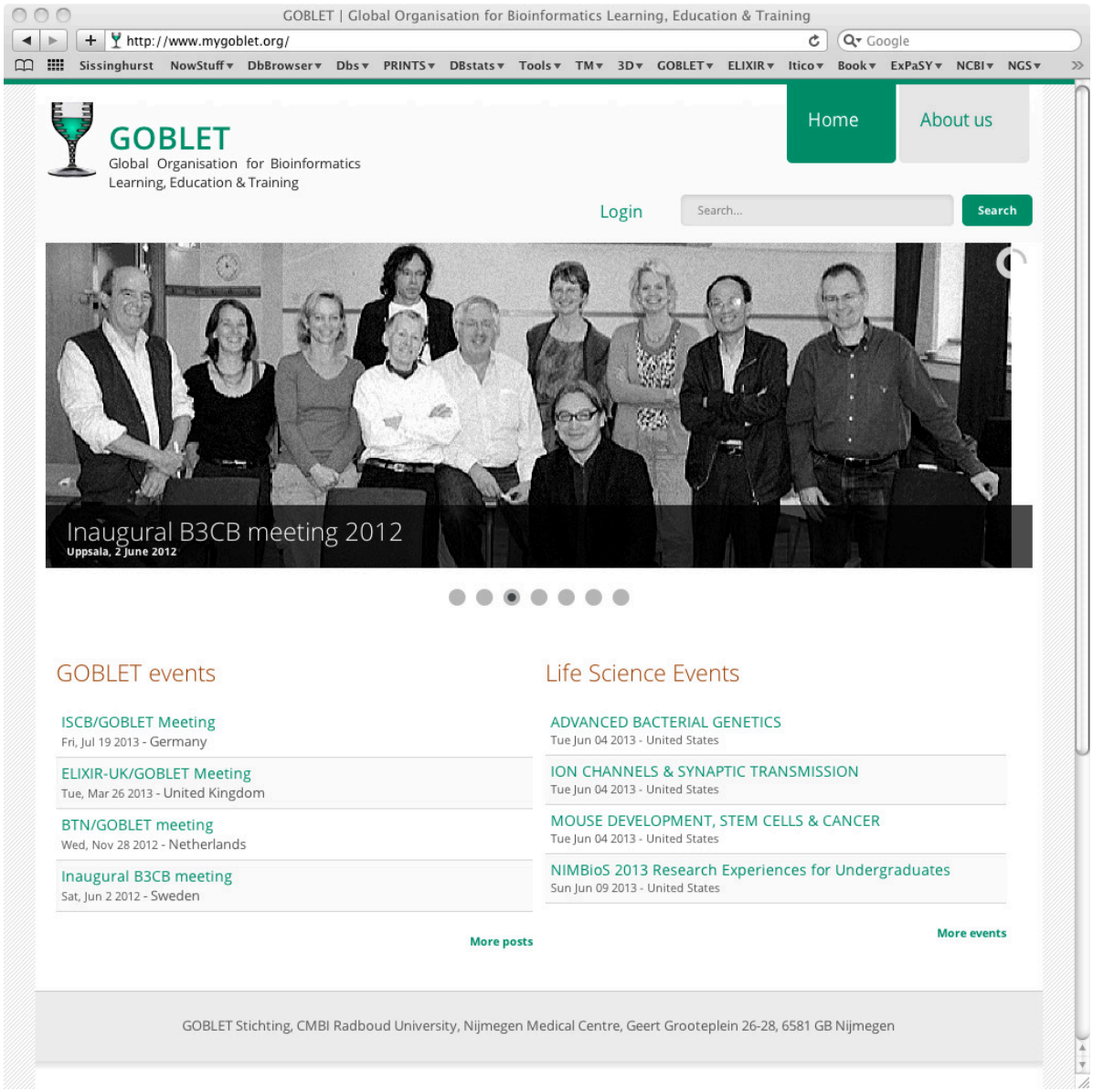
Dispersed across the globe are many organisations that conduct bioinformatics education and training activities as part of their core business. The value of these different educational initiatives within their particular contexts is clear. Nevertheless, we felt that tangible benefits – in terms of efficiency, cost-effectiveness, visibility, etc. – could accrue were such organisations to more openly share their experiences, expertise and resources. Under the auspices of the 24th Annual General Meeting of EMBnet (the Global Bioinformatics Network), leaders and representatives of a variety of international societies, networks and institutes therefore met in Uppsala, Sweden, on 2 June 2012, to discuss how such benefits might be realised in practice.

Invited organisations were: the International Society for Computational Biology (ISCB), the Asia-Pacific Bioinformatics Network (APBioNet), the African Society for Bioinformatics and Computational Biology (ASBCB), the IberoAmerican Society for Bioinformatics (SoIBio), the International Society for Biocuration (ISB), the European Bioinformatics Institute (EBI), the Netherlands Bioinformatics Centre (NBIC), the Next-Generation Sequencing Data Analysis Network (SeqAhead) and the Bioinformatics Training Network (BTN). The rationale for the meeting was that a strategic group like this would be uniquely placed not only to gain an overview of bioinformatics training activities, developments and needs across the globe, but also to begin to address those needs in a concerted way.

To kick-start the meeting, each representative introduced his/her organisation, highlighting its goals and activities in relation to bioinformatics education and training. All were encouraged not only to consider the challenges they face and the needs of the communities they serve, but also to articulate their dreams and visions for the future. From the presentations, several overlaps emerged. In particular, there was a general desire both to harmonise world-wide initiatives in bioinformatics learning, education and training, and to capitalise on increased collaboration across national and international boundaries. The participants agreed that bioinformatics training continues to be critical if we are to fully reap the benefits of the computational tools and resources that underpin the life sciences; moreover, that the need for such training has become increasingly relevant and urgent in light of the emerging challenges of next- and third-generation-sequencing data generation, storage and maintenance.

Overall, there was general consensus that, to help coordinate world-wide bioinformatics training activities, it would be useful to create an umbrella organisation that would allow Bioinformatics, Biotechnology, Biocuration and Computational Biology (B3CB) societies and networks to share, not duplicate, effort; to share, not duplicate, cost; to work together more powerfully towards common solutions and a sustainable future. This 'Global Organisation for Bioinformatics Learning, Education and Training' (GOBLET) was seen as a natural evolution of the BTN (Schneider *et al.*, 2012), providing a more formal, independent structure and allowing the network to expand more effectively.

It was envisaged that GOBLET would act as a forum for cooperation and sharing. In this role, it would, amongst other things, organise international events and acquire funds to stimulate and support relevant activities. To maximise the synergy and value derivable from investments made in local initiatives around the world, GOBLET would also offer centralised services and support for bioinformatics educators, trainers (and trainees) by providing a joint, community-focused, Web-based portal for training information, materials, tools, documents, and so on, open to international communities. Given inevitable institutional ownership issues, inter-institutional competition and internal funding vulnerability, it seemed likely that vesting in a single organisation the responsibility for creating and hosting such a resource was likely to be fraught with difficulties. However, uniting the relevant national and international networks, societies and organisations could allow the creation of an independent entity that was much more focused, much more robust to the volatility of institutional investment strategies, and hence better able to deliver on its aspirations for the benefit of its diverse communities.



To this end, the participants agreed to establish GOBLET as a legally registered Foundation, with a mission to:

1. provide a global, sustainable support structure for bioinformatics educators/ trainers and students/trainees;
2. facilitate capacity development in bioinformatics in all countries;
3. develop standards and guidelines for bioinformatics education and training;
4. act as a hub for fund gathering;
5. reach out to, amongst others, teachers at high schools, to bridge the gap to the next generation of bioinformaticians;
6. foster the international community of B3CB trai-ners;
7. and with an ethos that embraces:
8. inclusivity (welcoming all relevant organisations, networks, societies),
9. sharing (expertise, best practice, materials, tools, compute resources),
10. openness (using Creative Commons Licences),
11. innovation (welcoming imaginative ideas and approaches), and

12. *tolerance* (transcending national, political, cultural and social boundaries).

These discussions were observed by representatives from the Swiss Institute of Bioinformatics, the CSC IT Center for Science (Espoo, Finland), and Peking University (Beijing, China).

Following the meeting, a Memorandum of Understanding (MoU) was drawn up; having acquired its fifth signature, this came into force on 6 July 2012. All organisations invited to the Uppsala 'B3CB' meeting subsequently signed the MoU, indicating their commitment to formally establishing the GOBLET Foundation. The statutes that will form the legal basis of the Foundation, and its Governing Board, will be established during the coming months. The Board will conduct the daily business of GOBLET and will work to garner wider membership. Importantly, as part of its mandate, it will also strive to ensure that the Foundation develops as an open, global bioinformatics learning/education environment that benefits all contributors and helps to enhance the broader bioinformatics skill-level in the life sciences. To this end, GOBLET will also seek to establish synergistic relationships with its end-users and beneficiaries (*e.g.*, by maintaining active dialogues with students and student bodies, for whom education and training are top priorities).

Eleven new partners have since joined this initiative and have signed the MoU: the Swiss Institute of Bioinformatics (SIB); the Canadian Bioinformatics Workshops (bioinformatics.ca); the Instituto Gulbenkian de Ciência (IGC); the Society for Experimental Biology (SEB); the Australian Bioinformatics Network (ABN); the Centre for Proteomic and Genomic Research (CPGR); The Genome Analysis Centre (TGAC); The Sainsbury Laboratory (TSL); Itico; the SLU Global Bioinformatics Centre (SGBC) and the Italian Bioinformatics Society (BITS)[1]. With seed-corn investment from these and future members, it is ultimately hoped to engage a dedicated (part-time) assistant, to orchestrate GOBLET's work, maintain momentum, and assist with strategic planning and scheduling of meetings.

To help drive the initiative forward, a GOBLET kick-off meeting, co-organised by the BTN, was hosted by NBIC in the Netherlands, on 28 November 2012. The MoU and documentation relating to the B3CB and GOBLET meetings are available from the website (www.mygoblet.org). A joint meeting with ELIXIR-UK was later hosted by TGAC in Norwich, UK, on 25-26 March 2013 (www.mygoblet.org/elixiruk _ goblet _ meeting), to explore how GOBLET and ELIXIR could begin to share training experiences, collaborative ideas and best practice across the fields of bioinformatics, computational biology and computing. The next GOBLET meeting will be hosted by the ISCB and is scheduled to take place on 19 July 2013, alongside the programmed events of ISMB2013, in Berlin. We encourage related societies, networks and organisations to participate in this exciting new venture, and welcome expressions of interest via the website's contact form.

1 Since submitting this article for publication, five additional organisations or groups have signed the MoU: the IT Center for Science (CSC); Edinburgh Genomics (EdGe); the Nowgen Centre of Excellence in public engagement, education and professional training in biomedicine; Computational Genomics Analysis and Training (CGAT), and BioSharing.

References

Schneider MV, Walter P, Blatter M-C, Watson J, Brazas MD, Rother K, Budd A, Via A, van Gelder CW, Jacob J, Fernandes P, Nyrönen TH, De Las Rivas J, Blicher T, Jimenez RC, Loveland J, McDowall J, Jones P, Vaughan BW, Lopez R, Attwood TK and Brooksbank C (2012) Bioinformatics Training Network (BTN): a community resource for bioinformatics trainers. Brief. *Bioinform.* **13**, 383-389. doi:10.1093/bib/bbr064

Consortium members

Attwood, T.K.[1], Appel, R.[2,3], Blackford, S.[4], Blatter, M-C.[3], Bongcam-Rudloff, E.[5], Brazas, M.D.[6], Brooksbank, C.[7], Collado-Vides, J.[8], Corpas, M.[9], Cossey, S.[9], De Las Rivas, J.[10], Falquet, L.[3,11], Fernandes, P.L.[12], Gaeta, B.[13], Gaudet, P.[3], Helmer-Citterich, M.[14],Heringa, J.[15], Howard, J.C.[12], Hutchinson, P.16, Jimenez, R.[7], Kok, R.[17], Kumuthini, J.[18], Lewitter, F.[19], Lovell, D.[20], Maclean, D.21, Ouellette, F.[6], Mulder, N.[22], Palagi, P.M.[2,3], Peitsch, M.C.[3,23], Ranganathan, S.[24], Robinson-Rechavi, M.[3,11], Rost, B.[25], Schneider, M.V.[9],Schneider, R.26, Schönbach, C.[27], Trelles, O.[28], van Gelder, C.W.G.[29,30], van Kampen, B. [30], Via, A.[31],Vriend, G.[30]

1. Faculty of Life Sciences and School of Computational Biology, The University of Manchester, Manchester, UK; 2. University of Geneva, Geneva, Switzerland; 3. SIB Swiss Institute of Bioinformatics, Switzerland; 4.The Society for Experimental Biology, Lancaster University, Lancaster, UK; 5. Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala,

Sweden; 6. The Ontario Institute for Cancer Research, Ontario, Canada; 7. European Bioinformatics Institute, Hinxton, UK; 8. Centro de CienciasGenomicas, Universidad NacionalAutonoma de Mexico, Morelos, Mexico; 9. The Genome Analysis Centre, Norwich Research Park, Norwich, UK; 10. Bioinformatics and Functional Genomics Research Group, Cancer Research Center, Salamanca, Spain; 11. University of Lausanne, Lausanne, Switzerland; 12. Instituto Gulbenkian de Ciência, Oeiras, Portugal; 13. School of Computer Science and Engineering, The University of New South Wales, Sydney 2052, Australia; 14. Department of Biology, University of Tor Vergata, Rome, Italy; 15. Center for Integrative Bioinformatics VU, Free University, Amsterdam, The Netherlands; 16. The Society for Experimental Biology, Charles Darwin House, London, UK; 17. NBIC, Nijmegen, The Netherlands; 18.The Centre for Proteomic and Genomic Research, Anzio Road, Observatory, South Africa; 19. Whitehead Institute of Biomedical Research, Cambridge, USA; 20. Australian Bioinformatics Network, CSIRO Mathematics, Informatics and Statistics, Canberra, Australia; 21. The Sainsbury Laboratory, Norwich Research Park, Norwich, UK; 22. Department of Clinical Laboratory Sciences, University of Cape Town Medical School, Cape Town, South Africa; 23. University of Basel, Basel, Switzerland; 24.Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, Australia; 25. Technical University Munich, Garching, Germany; 26. Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg; 27. Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Fukuoka, Japan; 28. Computer Architecture Department, University of Malaga Campus de Teatinos, Spain; 29. Netherlands Bioinformatics Centre, Nijmegen, The Netherlands; 30. Centre forMolecularandBiomolecularInformatics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; 31. Department of Physics, Sapienza University, Rome, Italy.

'Next NGS Challenge – Data Processing and Integration' Conference – Conference report


Ana Conesa[1], Erik Bongcam-Rudloff[2]

[1]Centro de Investigación Príncipe Felipe, Valencia, Spain
[2]Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden

Received 24 March 2013; **Published** 19 August 2013

The '*Next NGS Challenge – data processing and integration*' Conference (held on 15-16 May 2013, in Valencia, Spain), a joint event of EU COST Action BM1006, SeqAhead, the Global Bioinformatics Network, EMBnet, the International Society for Computational Biology, ISCB, and the FP7 Project, STATegra, aims to become a dedicated, annual conference on cutting-edge Next Generation Sequencing (NGS) applications. The goal of the conference was to bring together interested users, computational biologists and bioinformaticians who face new challenges in high-throughput sequencing, and to feature new trends in NGS-based genome research.

The conference consisted of keynote lectures, contributed oral presentations and lively afternoon poster sessions. Submissions were selected on novel NGS applications/discoveries, algorithms for NGS data processing and integration, and efficient solutions for the management of massive volumes of sequence data. Abstracts from the conference are available as an *EMBnet. journal* supplement (Vol.19, Suppl. A[1]).

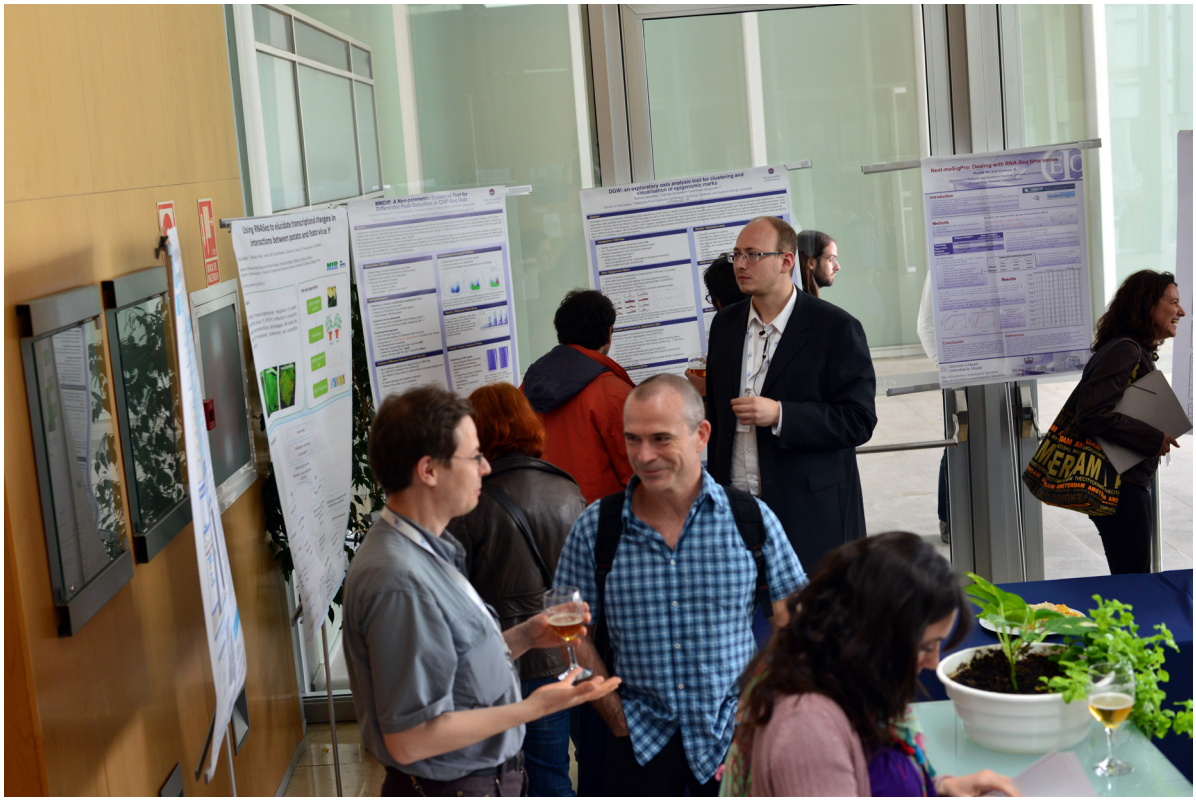
NGS is a highly parallelised approach for quickly and economically sequencing new genomes, re-sequencing large numbers of known genomes, rapidly investigating transcriptomes, analysing communities, etc., under different conditions.

Leif Andersson

Janet Kelso

Producing data on an unprecedented scale, these techniques are now driving the generation of knowledge (especially in biomedicine and molecular life sciences) to new dimensions. The massive data volumes being generated by these new technologies require new data-handling and storage methods. Hence, the life science community urgently needs new and improved

1 http://journal.embnet.org/index.php/embnetjournal/issue/current/showToc



Poster session

approaches to facilitate NGS data management and analysis. A 'moving target', this field requires that bioinformaticians, computer scientists and biomedical scientists join their expertise to bring NGS data management and analysis to new levels of efficiency and integration.

The conference was a true success in terms both of the high number of participants from a wide diversity of countries and backgrounds, and of the quality of the keynote and contributed talks. The conference achieved its main goal of joining together cutting-edge presentations in the newest applications of NGS to genome organisation (from the Neanderthal genome, to genomes of domesticated animals and of economically relevant crops), novel data-mining and integrative bioinformatics approaches, and high-performance computing solutions for the new challenges of massive sequencing. The format of the conference, with extensive breaks and poster sessions, favoured communication and networking among participants. Local logistics were also undertaken professionally, with particular attention to the satisfaction of attendees.

Finally some comments from the participants: Ralph Vogelslang, PacBio, Germany – "Thank you so much already for organising this great conference. It was a real pleasure for me to participate and get the chance to listen to excellent talks. I hope that you will have the chance to set up another conference next year and hopefully we can meet there again latest". Javier Terol, IVIA, Spain – "It has been a great conference, especially the keynote speakers and the second day-morning session. Truly, it was awesome."

More information about the conference: http://www.thenextngschallenge.org.

Acknowledgments

COST Action BM1006, SeqAhead. COST is supported by the EU RTD Framework Programme.

EMBnet - The Global Bioinformatics Network, in 2013: A Silver Anniversary

Teresa K Attwood

University of Manchester, Manchester, United Kingdom

Received 18 March 2013; **Published** 19 August 2013

Summary

EMBnet's 2013 Annual General Meeting (AGM) and associated events were hosted in the TRYP Valencia Oceanic Hotel, Valencia, Spain, from 14 to 18 May. The AGM, the silver anniversary of EMBnet's foundation, provided a timely opportunity both to review our progress since the "*EMBnet Moving Forward: 2010 & beyond*" workshop hosted in Ruvo di Puglia, Italy (Attwood *et al.*, 2010), and to showcase the work of our allied SeqAhead project. EMBnet has undergone significant changes since 2010: it has rebranded and launched a new website; it has formed new alliances; it has incubated new initiatives (notably, GOBLET[1]: (Global Organisation for Bioinformatics Learning, Education & Training) and multidisciplinary research projects (namely, SeqAhead[2] and AllBio[3]); and it has formally introduced structural changes to streamline our activities and broaden our membership.

The events allied to the AGM included a full-day EMBnet hands-on tutorial[4] on "RNA-seq and ChIP-seq data analysis", given by Endre Barta (U. of Debrecen, Hungary) and Eija Korpelainen (CSC- IT Center for Science, Finland), and on "NGS and structural Biology", given by Goran Neshich (EMBRAPA, Brazil), Jose R. Valverde (CNB-CSIC, Spain) and Gert Vriend (CMBI, Holland); the two-day SeqAhead "The Next NGS Challenge: Data Processing and Integration[5]" conference and Management Committee meeting; a full-day EMBnet workshop; and finally, the business meeting itself. Here, we review the motivation for the EMBnet workshop, and present its main conclusions.

Background

EMBnet's 2013 AGM was a significant event, marking the 25th anniversary since its foundation (in 1988), making it the oldest, widest bioinformatics network in the world. However, notwithstanding EMBnet's history and substantial contributions to the field of bioinformatics during the past three decades (*e.g.*, see Attwood *et al.*, 2011), in 2010, the Executive Board[6] set out a vision for a new EMBnet, one that would allow it to evolve more effectively as a global organisation. Amongst its key recommendations, the Executive Board suggested that there was an urgent need to:

1. review, streamline and clarify EMBnet's current membership scheme;
2. review how EMBnet might form strategic alliances with other bioinformatics networks and organisations;
3. consider how it might compete more strategically in global funding calls;
4. review the evolving role and internal structure of *EMBnet.journal*, and consider more tactical publishing strategies; and, in light of these considerations,
5. review and revamp EMBnet's current name, brand and website.

Ultimately, we wished to strengthen EMBnet's foundations, to create a competent, valuable and focused organisation, both to complement existing and emerging bioinformatics institutes, networks, associations and societies worldwide, and to maintain EMBnet's relevance in 2010 and beyond (Attwood *et al.*, 2010). Our goal was to better define EMBnet's role in European and global contexts, and open new avenues via which the organisation could move forward, for the benefit of its global membership.

The 2013 AGM, EMBnet's silver anniversary, was thus an opportune moment to reflect on how far we'd progressed towards meeting the 2010 recommendations.

1 http://www.mygoblet.org/
2 http://www.seqahead.eu/
3 http://www.allbioinformatics.eu/doku.php
4 http://www.thenextngschallenge.org/index.php/program/ngs-workshop
5 http://www.thenextngschallenge.org/
6 http://www.embnet.org/about/executive-board

Update on Achievements

Improving the EMBnet brand & promoting our activities more effectively

A taskforce had been set up to consider how to improve EMBnet's brand and Web presence. Working closely with the Publicity and Public Relations Project Committee[7] (P&PR PC), a new logo was introduced, together with the strapline, "*Bioinformatics without borders*", attempting to encapsulate EMBnet's global mission more effectively. The challenging task of designing a new website was also commenced. After exploring several design prototypes, we engaged Itico (a non-profit Private Limited Company supporting the iAnn project) to implement the ideas – see Figure 1. Although the website is still undergoing revisions, it's gratifying to know that it has already attracted some positive feedback from the community via Twitter.

In parallel with these developments, we had also proceeded to evolve the highly successful EMBnet.news[8] into a peer-reviewed journal – EMBnet.journal[9] (Figure 2, left-hand pane). The additional time taken to produce peer-reviewed issues, however, left a gap in our ability to make timely news announcements. We therefore experimented with a new Web-based publication, so-called *EMBnet.digest* (Figure 2, middle pane), to try to give a flavour of EMBnet's activities and upcoming events on a more regular (monthly) basis. As this appeared to be a popular mechanism for disseminating news, the P&PR PC subsequently overtook its production, and also created a new 'look & feel' (Figure 2, right-hand pane).

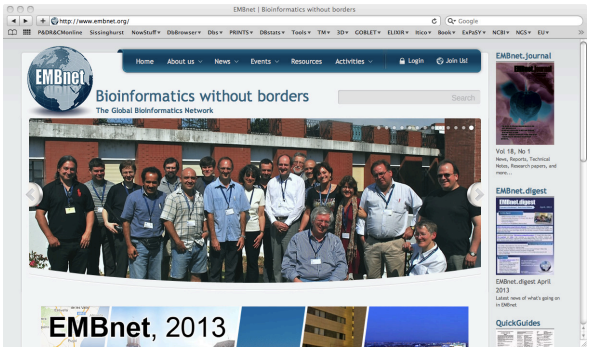


Figure 1. Screen capture showing the new EMBnet website, logo and strapline.

During the last 18 months in particular, the Executive Board also worked closely with the P&PR PC to establish sponsorship/patronage agreements to promote the work of EMBnet more widely, mostly notably at ISCB Latin America 2012[10]

Figure 2. From left to right, screen-captures showing the previous cover of *EMBnet.journal*, the original 'look & feel' of EMBnet.digest, and the current re-designed digest.

7 http://www.embnet.org/about/publicity-public-relations
8 http://journal.embnet.org/index.php/embnetnews/issue/archive
9 http://journal.embnet.org/index.php/embnetjournal/index
10 http://www.iscb.org/iscb-latinamerica2012

Figure 3. The B^{3}CB kick-off meeting that launched GOBLET.

(Chile), BITS 2012[11] (Italy), NETTAB 2012[12] (Italy) and ISCB Africa ASBCB 2013[13].

Securing strategic alliances with other organisations

Two very significant alliances have been formed since 2010, both related to EMBnet's commitment to bioinformatics training. First, having become a formal affiliate of ISCB, a press release was made announcing the intention of the ISCB and EMBnet to collaborate on bioinformatics education and training (ISCB press release[14], 2011). This will ensure that, whenever possible, EMBnet training courses are incorporated into ISCB meetings, especially those in developing regions. Importantly, the collaboration also establishes a framework for fostering future EMBnet and ISCB Student Council activities for the benefit of the global bioinformatics community.

The second alliance was extremely ambitious. A small workshop was held during the 2012 AGM in Uppsala, Sweden, at which representatives from nine other Bioinformatics, Biotechnology, Biocuration and Computational Biology (B3CB) societies, networks and projects came together to discuss the need for a more holistic view of bioinformatics training across the globe – see Figure 3. The participants concluded that, building on the achievements of the Bioinformatics Training Network (Schneider *et al.*, 2012), a Global Organisation for Bioinformatics Learning, Education and Training (GOBLET) was needed to coordinate world-wide bioinformatics training activities: to share, not duplicate, effort; to share, not duplicate, cost; to work together in a mutually respectful way towards common solutions and a sustainable future. To this end, a Memorandum of Understanding (MoU) was drawn up to establish GOBLET as a legally registered foundation, following the successful model of EMBnet. Since the ten founding members (EMBnet, ISCB, ASBCB, ISB, SoIBio, APBioNet, BTN, SeqAhead, NBIC and EBI) signed the MoU, 16 further organisations have pledged their commitment to formally create and support the GOBLET Foundation[15].

Competing for funds for some of our activities

Since 2010, two important research projects have also been incubated and funded, both relating to EMBnet's involvement in NGS activities. First was the Next Generation Sequencing Data Analysis Network (SeqAhead) COST Action (BM1006), the kick-off meeting for which took place in Brussels, Belgium in March 2011. By establishing a strong European network of NGS, data-analysis and informatics centres, SeqAhead aims to facilitate and stimulate the exchange of data, protocols,

11 http://bits2012.dmi.unict.it/conferencevenue.html
12 http://www.nettab.org/2012/
13 http://www.iscb.org/iscbafrica2013
14 http://journal.embnet.org/index.php/embnetjournal/article/view/217/459
15 www.mygoblet.org

Figure 4. From left to right, SeqAhead's first scientific meeting and AllBio's kick-off meeting.

software, experiences, ideas and, crucially, training. Erik Bongcam-Rudloff is the grant-holder and project Chair, and Terri Attwood the project Vice Chair; Andreas Gisel leads Work Group (WG) 2, which is developing an action plan for NGS bioinformatics; Eija Korpelainen leads WG 3, which concerns the design and implementation of software for NGS data analysis; and Gert Vriend leads WG 5, which coordinates the project's dissemination, education and training activities (www.seqahead.eu) – see Figure 4 (left-hand pane).

The second success came in the form of the EU Coordination Action, AllBio, which aims to broaden the bioinformatics infrastructure to unicellular, animal and plant science. The kick-off meeting was also held in Brussels, in November 2011 – Figure 4 (right-hand pane). Once again, Erik Bongcam-Rudloff coordinates the project, and partners include Andreas Gisel, Terri Attwood, Eija Korpelainen, Gert Vriend and Laurent Falquet (www.allbioinformatics.eu).

Structural changes

Arguably, the most significant development since 2010 was the consent, during the 2012 AGM, to radical revisions of EMBnet's statutes and byelaws, in order both to streamline how we conduct aspects of our business (especially our election processes) and, for the first time, to allow admission of individual members.

To this end, we removed the requirement for government mandates, and implemented an online payment system, allowing new members (both individual and organisational) to join more efficiently. Accordingly, EMBnet will no longer have National Nodes, although we will continue to recognise those Nodes that have National status; and there will no longer be a requirement for Node re-elections, as member status will relate strictly to the payment of the annual subscription fee. In addition, the new statutes oblige the Executive Board (EB) to work more closely with Project Committee (PC) Chairs, via a formal Operational Board (OB), and allow the PC Chairs greater flexibility to recruit enthusiastic members to help with their activities. Overall, these changes should allow future AGMs to run more efficiently: *e.g.*, elections will only be held for members of the EB and PC Chairs (committee members will not require election); and there will be less obligatory reporting, and hence more time for discussion and action!

The new statutes were finally ratified in April 2013. Since then, we've welcomed one new organisational member – The Genome Analysis Centre (TGAC), Norwich, UK, represented by Vicky Schneider – and several new individual members.

The 2013 AGM workshop

Taking advantage of the new statutes, and the greater time afforded for discussion, the 2013 AGM was preceded by a workshop (see Figure 5). This was organised into four main sessions, focusing on issues relating to i) membership fees and sponsorship; ii) the website, *EMBnet.digest*, social networking and QuickGuides; iii) *EMBnet.journal* and fund-raising; and iv) video-conferencing and future priorities. The first two of these sessions were run as parallel break-out groups, followed by plenary reportage; the remaining two were run as general group discussions. There were also several 'sticky-note' exercises, to encourage participants to think 'on the hoof' about a range of important topics, including potential future Editorial Board members for *EMBnet.journal*, additional

Figure 5. Participants of the 2013 AGM workshop.

pages, sections and/or functionalities that could be included in the new website; potential sources of funding; and EMBnet's future missions and priorities. The main outcomes of the discussions and exercises are reviewed below.

Membership fees

In previous meetings, the OB had discussed the possibility of changing EMBnet's fee structure: for organisations, the proposal was to drop the fee from €1,000 per annum to €500; for individuals, the proposal was to introduce a fee of €50.

We wished to use this opportunity to explore with the full EMBnet Board whether the proposals were optimal, what might be their pros and cons, and whether there were alternative models that we could adopt. We also wished to consider whether we could offer any new membership benefits, and, where new proposals had been made, what might be their associated benefits.

The discussions on this topic were surprisingly congruent: all groups agreed that a hybrid model of fees would be optimal, in which some organisations would be able to pay an annual fee of €1,000 and others €500, while individuals would pay €50, or €15 if they were students (but the student rate would apply, say, for only three years for any individual, without supporting documentation).

The principal benefit for organisations paying €1,000 would be that their travel to the AGM would be reimbursed (up to a ceiling of €500). In terms of additional benefits of EMBnet membership that could potentially be added to the website, several suggestions were made: these included opportunities to share facilities; to host visiting staff or students; to access a list of trainers; and the possibility to provide sponsorship for students (see next section). Another valuable suggestion was that we could explore the idea of joining student membership with the ISCB.

Sponsorship

In recent years, EMBnet has been approached by several organisations seeking sponsorship, patronage or endorsement of their events. While there is much to be gained by promoting EMBnet's activities in this way, nevertheless, a policy and an appropriate budget would need to be in place in order for us to be able to respond efficiently and consistently to requests of this type.

The questions for the group therefore concerned the types of event we should sponsor, the level of annual budget that could be set aside to support these or other strategic events (in terms of visibility, to prompt new alliances, collaborations, and so on), and what forms of sponsorship EMBnet might be able to provide.

There was general consensus that EMBnet should sponsor a range of events (not just those planned by networks/societies/organisations with whom EMBnet already has special agreements), but preferably those in which mutual benefits are apparent. The kinds of sponsorship offered by EMBnet could include keynote speakers at meetings, or trainers at courses, tutorials and so on; travel bursaries for students; and conference gadgets. It was proposed that we should set aside between 10 and 15% of EMBnet's annual surplus in order to be able to respond to future sponsorship requests.

Website, EMBnet.digest and Social Networking

The new website had been up and running for a while, but the migration of content from the previous site had been an enormous task, and many aspects remained to be resolved. Perhaps more importantly, the burden of maintenance was falling to one or two individuals alone. Help was clearly needed to progress some aspects of the site: these included the addition of EMBnet news and events; updating contents with new activities, new projects, additional resources, *etc.*; adding and updating member details; maintaining the mailing lists, and so on.

During the workshop, a little time was spent reviewing the website, and demonstrating the ease with which all members can login to the system, upload news and events, and generally manage, and take responsibility for, their own content. Overall, participants seemed very happy with

the new site. However, it was recommended to reconsider the rotating banner, which was causing problems in countries where bandwidth is a problem. Another key recommendation was to explore how to adopt a more unified image for promoting EMBnet via member websites. All such feedback will allow us to build on the site and generally improve the image that EMBnet projects to the world.

The 'look and feel', and management of *EMBnet.digest* were also discussed. Here again, most of the work falls to one or two individuals, and help is needed to collate information more efficiently. The principal outcome of the discussions was that we should consider adopting a more exciting 'look and feel', which could be prototyped in the coming months. Moreover, a schedule of 'In focus' contributions was agreed from now until October 2013, commencing with a focus on the AGM in the May issue.

There was also a general discussion around the presence of EMBnet on social networks (via media such as LinkedIn, ResearchGate, Twitter, Facebook, Wikipedia, *etc.*). EMBnet already has a successful LinkedIn site and a Wikipedia entry. Given that EMBnet doesn't have resource to dedicate to all such activities, it was suggested that the website could be made 'social media friendly', to allow news, events, features and so on to be exported to relevant social media, as and when appropriate, more easily.

QuickGuides

QuickGuides[16] have been a long-standing and useful output of EMBnet. However, maintaining or updating them, and writing new ones, requires commitment and effort. As many of the Guides had not been updated and no new Guides had been created for several years, the question was whether we should keep them.

In the discussion that followed, all agreed that we should maintain EMBnet's portfolio of QuickGuides, and that these should be peer-reviewed in order to maintain quality. Two complementary models for managing the Guides were proposed: i) that they should be made available via the GOBLET website, as part of its educational/training content; and ii) that they should be managed via the Open Journal System (OJS). As these models are not mutually exclusive, it would seem sensible to proceed with both!

16 http://www.embnet.org/embnet-quickguides

EMBnet.journal

The transition from the highly successful *EMBnet.news* to the new, peer-reviewed *EMBnet.journal* has been a major challenge. While several supplements containing conference abstracts have been published, the main journal issues have been much more difficult to produce. Therefore, the discussions on *EMBnet.journal* focused primarily on how it could attract more research articles and/or other content, and whether the journal should accept advertisements, in order to bolster its income.

It's clear that the principal problem facing the journal is lack of dedicated manpower. All agreed that hiring a paid, part-time assistant would be helpful, and that a small part of EMBnet's surplus should be set aside annually to achieve this. Several other conclusions were reached: first, that all members of EMBnet should provide a 'good' article, or perhaps encourage a student in their group to submit some of their work; second, that the journal could include a new 'book review' section; third, that the journal should include a section dedicated to the Node 'InFocus' pages of *EMBnet.digest*, highlighting the special events, activities and/or initiatives of EMBnet's Nodes and members; and finally, that we should explore what models other journals adopt with regard to the inclusion of advertisements.

Fund-raising

Fund-raising is an important activity for augmenting EMBnet's finances and for galvanising its work – without dedicated funds, it is much harder always to rely on the good will and dedication of particular individuals. This discussion therefore centred on what types of activity EMBnet should seek to get funded, and what funding opportunities might be available to support them.

There was general agreement that the types of activity that EMBnet would like to be in a position to fund, or offers it would like to be in a position to make, would include travel fellowships for students; support for EMBnet speakers and/or trainers; administrative support for *EMBnet.journal*; EMBnet events (tutorials, workshops, conferences, *etc.*); and joint research projects.

The first, most obvious, mechanism by which to enable such activities and benefits was to increase EMBnet's membership. We are, of course, hopeful that the new, streamlined process for joining EMBnet, with is easy-to-use online payment system, will help to widen our membership;

Figure 6. Left-hand pane, a reward for Domenica D'Elia for sterling work as P&PR PC Chair; right-hand pane, a toast to Vicky Schneider, representative of new organisational member, TGAC.

but clearly, we need to do more. It was recognised that we need to improve our publicity and outreach activities, and that we could and should be doing a lot more both to attract student members and to give them visibility within EMBnet. Other ideas included merchandising (*e.g.*, EMBnet T-shirts!), advertisements in the journal, and possibly attracting donations (which are within our statutes to receive). We should also be proactively consulting development agencies (UNESCO, *etc.*) and seeking joint research projects via national and international funding bodies.

Video conferencing

Maintaining the momentum of EMBnet's activities between AGMs is always a challenge, a challenge that rests with the PC Chairs and EB to address. At one time, it was possible to hold interim Committee meetings, Executive meetings and meetings of the full Board using a (commercial) online conferencing system called Marratech. However, the company was taken over, the software was suddenly no longer available, and we were obliged to find a new solution. After several trials, most monthly meetings of the EB and OB have been conducted with Skype, but this continues to cause problems for many participants. This has become an issue, especially for the bimonthly meetings of the full Board, because a great deal of time is wasted while people try (often unsuccessfully) to connect: this prohibits efficiency, destroys morale, and increasingly deters people from attending.

In consequence, it was agreed that EMBnet will conduct an experiment to test around a dozen video-conferencing packages. Participants in the experiment would need to confirm their commitment after the AGM, a protocol for conducting the experiment would be agreed during the summer, and global testing would commence towards the beginning of the autumn. Ultimately, this should provide the optimal solution for all involved, and the results would be made available as a publication in *EMBnet.journal*.

Future priorities

In thinking about the future of EMBnet, it has been imperative to reflect on the impact of globalisation on our once-European organisation, and to better define EMBnet's role in the global bioinformatics landscape. As EMBnet has evolved, so too has its mission and its priorities. Accordingly, the question was what EMBnet's missions and priorities should now be, and what we should be doing to help get us there.

In terms of missions, there was general consensus that to promote, "*Bioinformatics without borders*" should be our primary goal. EMBnet should disseminate best practice in bioinformatics, promote bioinformatics capacity in all countries, share bioinformatics knowledge, and, crucially, incubate new initiatives (*e.g.*, like GOBLET).

EMBnet's priorities were considered to be to champion education and training; to achieve better PR and outreach; to improve the roll-out and recognition of *EMBnet.journal*; to make better use of our website in order to share more with our communities; to invest strategically in our activities; and, wherever possible, to apply jointly for funds to help realise our research, training and networking objectives.

Conclusion

It's clear that EMBnet has come a long way in the last three years, largely thanks to the enormous efforts of a handful of individuals. Thanks are especially due to Lubos Klucar for his work in running *EMBnet.journal*; to Erik Bongcam-Rudloff for coordinating the SeqAhead and AllBio projects, and to Andreas Gisel, who helps him to keep these projects on track; and finally, to Domenica D'Elia for her indefatigable work as EMBnet's most passionate and committed P&PR PC Chair – see Figure 6.

During the 2013 meeting, we were also fortunate to be able to celebrate the arrival of our latest organisational member, The Genome Analysis Centre[17] (TGAC), represented by Vicky Schneider, who immediately stepped up to take on many new tasks (Figure 6).

The workshop and business meeting were lively affairs, packed to the brim with activities, and suggesting much work that's still to be done. We therefore call on everyone's ongoing support and collaboration to help us evolve EMBnet into a more innovative and ultimately more productive organisation. In particular, we hope that others will take the lead of those who've shouldered so much of EMBnet's work in recent years, and will join us in our efforts to make EMBnet the preeminent bioinformatics network in the world. And perhaps some of us may yet live to see EMBnet celebrate its golden anniversary!

Acknowledgements

We are indebted to Ana Alvarez and Ana Conesa for their considerable efforts with the local organisation, especially for their restaurant recommendations, and overall, for providing such a comfortable and welcoming environment.

References

Attwood TK, Bongcam-Rudloff E, Gisel A, de Villiers E (2010) EMBnet - the European Molecular Biology Network, Moving forward: 2010 & beyond. *EMBnet.journal* **16**(1): 8-14.

Attwood TK, Gisel A, Eriksson N-E, Bongcam-Rudloff E (2011) Concepts, historical milestones and the central place of bioinformatics in modern biology: a European perspective. In: *Bioinformatics - Trends and Methodologies*, Mahdavi MA Ed., Intech Online Publishers, November 2011, ISBN 978-953-307-282-1.

ISCB Press Release (2011) ISCB and EMBnet to collaborate on bioinformatics education and training. *EMBnet.journal* **17**(1): 6.

Schneider MV, Walter P, Blatter MC, Watson J, Brazas MD, Rother K, Budd A, Via A, van Gelder CWG, Jacob J, Fernandes P, Nyronen T, De Las Rivas J, Blicher T, Loveland J, McDowall J, Jones P, Vaughan B, Lopez R, Attwood TK, Brooksbank C (2012) Bioinformatics Training Network (BTN): A community resource for bioinformatics trainers. *Briefings in Bioinformatics* **13**(3): 383-389. doi: 10.1093/bib/bbr064

17 http://www.tgac.ac.uk

2013 Annual General Meeting: Executive Board Report





Teresa K Attwood[1], Andreas Gisel[2], Etienne de Villiers[3], Erik Bongcam-Rudloff[4], Goran Neshich[5]

[1]The University of Manchester, Manchester, United Kingdom
[2]Institute for Biomedical Technologies, CNR, Bari, Italy
[3]Kenya Medical Research Institute (KEMRI), Kenya
[4]Department of Animal Breeding & Genetics, Swedish University of Agricultural Sciences, Uppsala
[5]EMBRAPA, Brazil

Received 18 June 2013; **Published** 19 August 2013

During the past year, the Executive Board (EB) met regularly and held frequent meetings with the Operational Board via Skype. These meetings allowed discussion of a range of issues relating to the Project Committees (PCs), to *EMBnet.journal*, to the website, the Stichting accounts, membership, *etc*.

In alternate months, we also endeavoured to convene Skype meetings open to the full EMBnet constituency; however, for larger numbers of attendees, technical issues continued to cause problems. Attempting to address these issues, we took the first steps towards evaluating the various tools and technologies available for online meetings, by creating a list of existing tools, and a protocol on how to test them. The test will be realised in the same spirit as the 'ping project[1]' of the '80s. The 2013 AGM allowed us to convene a working group to discuss the issues in more detail, and to initiate a common experiment with different software and different Nodes. The ultimate goal is to write a white paper and to publish the results in *EMBnet.journal*.

The last year has been both busy and productive, building substantially on the programme of work we outlined in 2010. In particular, working closely with Itico[2] to improve the EMBnet 'brand', we finally launched the new website, which now includes a new online fee-payment module for individual members. We call on all members to help augment the content of the new site and to help keep it up-to-date.

Since the 2012 AGM in Uppsala, EMBnet's training strategy has been dominated by our leadership of GOBLET[3] (the Global Organisation for Bioinformatics Learning, Education and Training), which has been established as a Stichting, registered in the Netherlands, following the successful model of EMBnet. Working through GOBLET has significantly increased our level of interaction and cooperation with a range of major international societies and networks (including ISCB[4], ASBCB[5], ISB[6], APBioNet[7], SoIBio[8], ABN[9] and so on) – from the original 10 members who signed the Memorandum of Understanding to establish GOBLET, a further 16 organisations and several individuals have committed to join the Foundation.

Another profound advance for EMBnet this year has been the final ratification of the new statutes, which were voted in during the Uppsala 2012 AGM – these statutes became legally binding in April 2013. The most significant result of the change is that we are now able formally to accept individual members (this allows, say, former Node managers to join, or any individual to participate in EMBnet's activities but whose organisation is not a member). Since ratification of the statutes, we have already had several membership applications, and welcomed one new organisation; new memberships will be processed without delay via the online payment system. The new statutes also ushered in changes to the internal structure of the organisation, obliging the PC Chairs and the EB to work together much more closely than they have done in the past; it has also allowed us to streamline the way in which EMBnet's activities

1 en.wikipedia.org/wiki/Ping_%28networking_utility%29

2 www.itico.org
3 http://www.mygoblet.org
4 www.iscb.org
5 www.asbcb.org
6 http://biocurator.org
7 www.apbionet.org
8 www.soibio.org
9 australianbioinformatics.net

are coordinated, no longer requiring Committee members to be elected but rather, giving the PC Chairs greater flexibility to interact with members who most want to contribute to their work.

During the year, we have continued to work closely with other EU-funded projects, such as AllBio and SeqAhead. AllBio[10] had its first AGM in Amsterdam (December 2012), where plans for the project's second year were discussed, including the all-important, 'Bioinformatics Workshop: Evaluation of Test Cases[11]. Authors of selected test cases and specific bioinformatics specialists were invited to this event, hosted in Milan by EMBnet Italy (CNR-ITB[12]). During the three-day workshop, various round-table discussions were held, and several new events were devised to help provide test-case solutions. Two of these 'hackathons' have already happened: one hosted in Amsterdam by EMBnet Netherlands (CMBI[13]) (March 2013); one hosted in Bari by EMBnet Italy (April 2013). In this second year, EBI[14] (EMBnet UK) organised two AllBio-sponsored courses: one on analysis of genome-scale data from plant pathogenic fungi (PhytoPath), the other on data resources and tools for 'plant-omics'. A further key AllBio event took place in Munich (October 2012) – the 'AllBio Ontology Workshop: Using Biomedical Ontologies for Improved Bioinformatics Database Interoperability', organised by EMBnet UK (UMBER[15]) and EMBnet Netherlands. A full list of AllBio events is available from www.allbioinformatics.eu/doku.php?id=public:events.

Within SeqAhead[16], EMBnet Nodes were involved in several events covering different problems in NGS data analysis. In chronological order, EMBnet Sweden (BMC[17]) and EMBnet Netherlands (CMBI) organised the 'Trans-COST Bioinformatics[18], event in Amsterdam; and EMBnet Italy (CNR-ITB) was involved in the COST Workshop 'The next NGS Challenge: High-throughput omics and Data Integration[19], in Barcelona, and organised the COST Workshop 'NGS and non-coding RNA data analysis[20], in Bari. Three further events are worthy of note: in May, the COST conference 'The Next NGS Challenge: Data Processing and Integration[21],in Valencia, involving EMBnet Sweden (BMC), EMBnet Italy (CNR-ITB) and EMBnet Spain[22] (CSIC); in June, a 'Hadoop and NGS data processing' hackathon[23], hosted in Sardinia by EMBnet Finland (CSC[24]), and a COST Workshop on 'Future demands and challenges in ICT and bioinformatics tools for NGS', with EMBnet Finland (CSC), EMBnet Sweden (BMC) and EMBnet Italy (CNR-ITB).

Throughout the year, we have described these and our other activities in our monthly EMBnet.digest[25] and in EMBnet.journal[26]. During 2012, the bulk of the journal work has involved preparation of conference proceedings for the BITS 9th Annual Meeting (see EMBnet.journal 18 Suppl. A[27]) and NETTAB 2012 (see EMBnet.journal 18 Suppl. B[28]).

This year's AGM was an opportunity to strategically re-group in order to build on some of these initiatives, and to prioritise new projects and proposals to attract more funding. This year, one member of the EB (Erik Bongcam-Rudloff) was up for re-election, and was duly re-elected. As always, there's still a lot more work to do. We therefore encourage you all to contribute your energies and visions to EMBnet, both to make this 25th Anniversary year of EMBnet a real celebration and to prepare the way for the next 25 years of the Global Bioinformatics Network!

Chair: T.K. Attwood
Secretary: A.Gisel; **Treasurer:** E. de Villiers; **Members:** E. Bongcam-Rudloff, G. Neshich

10 www.allbioinformatics.eu
11 http://www.allbioinformatics.eu
12 www.embnet.org/nodes/institute-biomedical-technologies-itb
13 www.embnet.org/nodes/centre-molecular-and-biomolecular-informatics-cmbi
14 www.embnet.org/nodes/embl-outstation-european-bioinformatics-institute-embl-ebi
15 www.embnet.org/nodes/university-manchester-umber
16 www.seqahead.eu
17 www.embnet.org/nodes/biomedical-centre-bmc
18 http://seqahead.cs.tu-dortmund.de/meetings:trans-cost
19 www.seqahead.it/cost-bcn-2013
20 http://www.seqahead.it/cost-bari-2013/www.thenextng-schallenge.org
21 www.thenextngschallenge.org
22 www.embnet.org/nodes/centro-nacional-de-biotecnología-csic
23 www.crs4.it/news/-/blogs/hadoop-and-ngs-data-processing-hackathon-iii
24 www.embnet.org/nodes/csc-scientific-computing-ltd
25 www.embnet.org/embnet-digest
26 http://journal.embnet.org
27 journal.embnet.org/index.php/embnetjournal/issue/view/61
28 journal.embnet.org/index.php/embnetjournal/issue/view/73

2013 Annual General Meeting: Education & Training Project Committee Report

Pedro L Fernandes
Instituto Gulbenkian de Ciência, Oeiras, Portugal

Received 1 August 2013; **Published** 19 August 2013

At the end of the 2010-2013 triennium, the Education & Training Project Committee (E&T PC) reports its progress as follows:

Revamping the plans

Our on-going plans were frozen, owing to lack of proper funding. Instead, we started new activities:

- in conjunction with the creation of GOBLET, parallel plans have been prepared to introduce a recognition system for training. Contacts were made with community members, particularly training instructors, to understand their position towards adopting a badges system, a proposal that was uniformly well accepted. The next step will be to form a small group within EMBnet to take the role of endorsing training partners to issue badges. EMBnet can also play a significant role in displaying badges. The Mozilla Open Badges system[1] became fully functional in recent months. A good example can be found at the cloudworks[2] website
- a careful look into MOOCs (Massively Open Online Courses) was started, in order to consider the possibility of partnering with educational resources looking for compatible teaching methods. This is the consequence of adopting peer instruction (Eric Mazur) and Flipped-Classes (Aaron Sams and Jonathan Bergman) in formal education. These methods are much better exercised in the training world, and some EMBnet members can provide knowledge in this area;
- the design of interactive courseware to enable proper e-learning is currently under consideration. This requires familiarisation with techniques first, and then with software that allows such tools to be built. We have started talks with software vendors in order to enable EMBnet to have access to development tools for free, or for a very modest price. So far, we have mainly expressions of interest. Some software vendors are considering an extension of existing deals to larger audiences (*e.g.*, Biomatters);we also participated in talks with industrial parties that can cooperate in bringing their technological expertise to training events, as additional demonstration sessions: *e.g.*, pilot experiments have been successful with Genomatix. This is a growing interest that may lead to mutually beneficial exchanges.

Seeking funding

Implementation of the activities depicted in the 2010-2013 business plan should be assessed fairly, considering the difficulties in proceeding without funds. As funding is crucial to any of the plans we laid-out, the E&T PC explored contacts with European project coordinating bodies and expressed its interest in joining EU-funded initiatives, mainly:

- the Grundtvig programme[3], which focuses on the teaching and study needs of learners taking adult education and 'alternative' education courses, as well as the organisations delivering these services. It aims both to help develop the adult education sector, and to enable more people to undertake learning experiences, especially in other European countries;
- the Leonardo da Vinci Programme[4], which funds many different activities of varying scales. These include 'mobility' initiatives, enabling people to train in another country, co-operation projects to transfer or develop innovative practices, and networks focusing on topical themes in the sector.Entering these programmes is fundamentally different from

1 http://www.openbadges.org
2 http://cloudworks.ac.uk/badge/badge_list
3 http://ec.europa.eu/education/lifelong-learning-programme/grundtvig_en.htm
4 http://ec.europa.eu/education/leonardo-da-vinci/initial_en.htm

applying for a research project. It requires contacting their National Agencies and getting 'invited' to attend their workshops with some regularity. Many of these federated structures run local workshops, termed Lifelong Learning Programmes (LLP), which are often out of our scope – *e.g.*, those on Learning Languages (idioms) or Occupational Geriatrics. However, we must be attentive to areas that are lateral to their actions, such as those on Medical Education, in which we aim at a small share for bioinformatics.

We have found that while these sources have to be considered and negotiations should not stop, funds are pre-allocated to activities, making it difficult to propose new ones.

It is generally felt that EMBnet is in need of a more encompassing funding plan in which the E&T PC can provide cost estimates, data and texts to write grant applications that have an E&T component.

Outreach

EMBnet's track record in training provision throughout its constituency was presented in a talk about networking in science, in which EMBnet's roles were described, at the ABRF conference[5] in Orlando, FL, March 2012. The talk generated a series of contacts that may result in institutional exchanges that lead to training cooperation worldwide, in places where EMBnet's activity is poorly known.

As genomics reaches the clinical arena, it is increasingly evident that the public are largely unprepared for the implications of using personal genome data, in terms of consent. This issue could be less problematic if the educational system helped create a common, informed awareness. In collaboration with Pietro Liò and Lucia Bianchi, such a proposal was made in a paper accepted for *The future of Education*[6] conference in Florence (June 2013) entitled, "*Improving collective awareness and education about the privacy and ethical issues connected with the genome technologies*".

Furthermore, the prospect of using EMBnet's global presence in educational initiatives has been mentioned in talks in Portuguese academic institutions (*e.g.*, Braga, March 2013 and Porto, May 2013) triggering interesting discussions on the need to use networked communities in globalised educational efforts, like MOOCs.

Committee members

Pedro Fernandes (PT)	Chair
Emiliano Barreto (CO)	Secretary
Bruno A. Gaeta (AUS)	Member
Shahid Chohan (PK)	Member

Associate members

J.R Valverde (ES)
Mohamed Abelhouda (EG)
Matej Stano (SK)

5 http://www.abrf.org/index.cfm/page/meetings/ABRF2012/ABRF2012

6 http://conference.pixel-online.net/foe2013

'2013 Annual General Meeting: Publicity & Public Relations Project Committee Report

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

Received 24 July 2013; **Published** 19 August 2013

In May 2013, the Publicity and Public Relations Project Committee (P&PR PC), elected during the 2010 EMBnet AGM, brings to a conclusion its three-year mandate. This article reports on the Committee's activity during the last year (May 2012-2013), and summarises its achievements relative to the objectives set out in its 2010 business plan.

May 2012- 2013 activity report

The main activities and achievements of the P&PR PC during this last year were:

1. finalisation and release of the new EMBnet website;
2. production and release of the monthly EMBnet.digest;
3. promotion and support for the publication of EMBnet.journal Conference Supplements:
   - EMBnet.journal vol. 18, Suppl. A1, BITS 9th Annual Meeting;
   - EMBnet.journal vol. 18, Suppl. B2, NETTAB 2012 workshop;
4. support and production of EMBnet publicity material for sponsored events:
   - ISCB Latin America 2012 Conference3 (Chile);
   - ISCB Africa ASBCB 2013[4] (Morocco);
5. distribution of EMBnet publicity materials at events supported by EMBnet's patronage, such as the BITS 2012 Conference[5] and NETTAB 2012 Workshop[6], and at SeqAhead and AllBio workshops and courses;
6. renewal of EMBnet's ISCB Affiliation;
7. migration and updating of the EMBnet web based Open Conference and Journal Systems (OCS and OJS) (assisting the TM PC);
8. timely updating of the website regarding EMBnet Node information, and announcements of *EMBnet Events and News*, and latest releases of *EMBnet.digest* and *EMBnet.journal*;
9. updating of EMBnet information in Wikipedia;
10. announcement of EMBnet activities on LinkedIn'Special Interest Groups'.

As for the new website, the P&PR PC supported the Executive Board (EB) regarding the layout and content organisation, substantially updated the website contents, contributed to the creation of the membership Web forms, and carried out all the activities at point 8 above.

The P&PR PC also actively participated in monthly virtual meetings of the Operational Board, and contributed in a timely way to EB activities reported in the companion article in this issue (see '2013 Annual General Meeting – Executive Board Report', by Attwood *et al.*).

P&PR PC goals & achievements from 2010 to 2013

As a follow-up to the break-out groups of the 2010 AGM, '*EMBnet moving forward 2010 and beyond* (Attwood *et al.*, 2010), the P&PR PC produced a business plan, including a series of short- and long-term objectives considered crucial for accomplishing 'new' EMBnet's needs and plans for the future. Part of this document, specifically related to the 'P&PR PC visions and goals' was published in 2011 in *EMBnet.journal* (D'Elia, 2011).

The business plan outlined both enhancements to existing activities and proposals for new strategies, aiming to strengthen the network of EMBnet contacts and to broaden its collaborations. Many of the objectives mentioned in the business plan have been successfully accomplished, the major one being the website and actions relating to dissemination and publicity: the P&PR PC designed the new EMBnet logo, pro-

1 journal.embnet.org/index.php/embnetjournal/issue/view/61
2 journal.embnet.org/index.php/embnetjournal/issue/view/73
3 http://www.iscb.org/iscb-latinamerica2012
4 http://www.iscb.org/iscbafrica2013
5 http://bits2012.dmi.unict.it/conferencevenue.html
6 http://www.nettab.org/2012/

Figure 1. Presentation of P&PR PC activity report at EMBnet AGM, 18 May 2013, Valencia (ES).

vided a new design for *EMBnet.digest*, produced a new EMBnet leaflet and managed the dissemination of EMBnet activities at large.

The P&PR PC also worked to enlarge the EMBnet community. In line with this objective, we promoted and managed the establishment of 'EMBnet Contact' registration. This has been a successful initiative, with more than 80 contacts registered, including many young people (*e.g.*, students and early-stage researchers), as well as senior scientists from all over the world. The P&PR PC also promoted and supported the organisation of major conferences and sponsorship activities, which have greatly improved EMBnet's image and reputation during the last three years.

Despite these results, more work still needs to be done. To meet existing and new objectives emerging from the AGM in Valencia (ES) in May 2013 (see the companion article in this issue: '*EMBnet - The Global Bioinformatics Network, in 2013: A Silver Anniversary*', by Attwood TK) will require even greater support from EMBnet. Indeed, as mentioned in the business plan, the feasibility of our proposal was dependent on two key factors: i) the collaborative support of all EMBnet members, and ii) volunteering for specific tasks. The achievements of these last three years came at a substantial cost in terms both of human resources and financial investments made by the EB. Collaborative efforts are now needed to consolidate our achievements and build on them for the future of EMBnet.

As Chair of the Committee, I would like to thank all the Committee members who have supported me in these three last years: namely, the secretary, Lubos Klucar (SK), Martin Norling (SE), Judit Kumuthini (ZA) and Kanchana Senanayake (CL). In particular, on behalf of the P&PR PC, I would also like to thank Pedro Fernandes (Chair of the Education and Training Project Committee) and Andreas Gisel (member of the EB) who have actively collaborated during this period on many of the P&PR PC initiatives as P&PR PC associate members. Last, but not least, my big "Thanks" are for the EB, who have continuously and constructively supported me.

I also would like to thank the EMBnet Board for the trust they demonstrated during the 2013 EMBnet AGM by re-electing me as Chair of the P&PR PC.

With the new statutes, ratified on April 2013, the organisation of the project committees will be under the unique responsibility of the committee Chairs. As for the P&PR PC, in order to properly answer to the most urgent needs raised during the EMBnet workshop in Valencia, two special interest groups have been devised: a website task force, comprising Rafael Jimenez (Itico and iAnn project) and Cesar Bonavides-Martinez (EMBnet Mexico), and a task force that will collaborate on the Committee's dissemination activities. As for this last task, I welcome Vicky Schneider (Node manager of the new organisational member – The Genome Analysis Centre (TGAC), Norwich, UK) – who is willing to help.

References

Attwood TK, Bongcam-Rudloff E, Gisel A, de Villiers E (2010) EMBnet - the European Molecular Biology Network, Moving forward: 2010 & beyond. *EMBnet.journal* **16**(1), 8-14.

D'Elia D (2011) Publicity and Public Relations Project Committee: AGM2011 report. *EMBnet.journal* **17**(Supp.A), 9-11.

2013 Annual General Meeting: Technical Management Project Committee Report

George Vasilios Magklaras

University of Oslo, The Biotechnology Centre of Oslo, Norway

Received 23 July 2013; **Published** 19 August 2013

The Technical Management Project Committee (TM PC) performed the following tasks for the period 2012-2013:

- specified and implemented a cloud virtualised host server, to transfer the main Web and EMBnet node journal operations. The Linode server provider was chosen;
- completed the transfer of the following domains from the Norwegian EMBnet node servers to the Linode hosting provider, minimising the impact on users and maintaining availabi-lity of operations:
  - http://www.embnet.org
  - http://conference.embnet.org
- helped the new website team to resolve a number of system issues in the deployment process of the new EMBnet website;
- set up a newly configured Adobe Connect Video and Audio Meeting software suite, in an attempt to help the EMBnet Operational Board (OB) choose an alternative way to conduct virtual general meetings (VGMs). The server ope-rates at the following URL: https://connect.uninett.no/testembnet

Currently the TM PC continues to manage:

- the EMBnet.journal Web server[1], and is helping the Swedish EMBnet node to transfer its content to a new server in Sweden;
- the EMBnet mailing lists
- the EMBnet DNS operations
- Because the TM PC Chair could not attend the AGM, new committee members willing to collaborate in the activity of the TM PC were hired after the AGM. The Committee is now composed of:George Magklaras – Chair and Secretary (NO), Gang Cheng – Member (NO), RomualdoZayas Lagunas - Member (MX)

The TM PC chair wishes to thank all the previous TM PC members for their contribution and effort.

The following goals of the TMPC are set for the remaining activity time until the next AGM (2014):

a. improve/run the Adobe Connect Video and Audio meeting software for EMBnet;
b. manage the transition of *EMBnet.journal* to a new server, owing to decommissioning of the Norwegian EMBnet node server proteas.uio.no;
c. manage the transitioning of the EMBnet node DNS from Sweden to Linode and
d. secondary DNS servers so that continuity of operations is sustained;
e. manage the transitioning of the EMBnet node mailing list from Sweden to Linode and other servers, to ensure that communications are not disrupted.

The TMPC will meet twice every month, at a predetermined date and time, as called by the TM PC Chair, using the Adobe Connect Video and Audio Meeting software suite service. All members are expected to attend these meetings and contribute towards the specified goals/tasks above.

1 http://journal.embnet.org

ICT needs and challenges for Big Data in the Life Sciences. A workshop report - SeqAhead/ISBE Workshop in Pula, Sardinia, Italy, 6 June 2013

Babette Regierer[1,2], Luca Pireddu[3], Martijn Moné[4], Andreas Gisel[5]

[1]Wageningen University, Wageningen, The Netherlands
[2]LifeGlimmer GmbH, Berlin, Germany
[3] Centro di ricerca, sviluppo e studi superiori in Sardegna (CRS4), Pula, Italy
[4]VU University Amsterdam, Amsterdam, Netherlands
[5]CNR - Institute for Biomedical Technologies, Bari, Italy

Received 15 July 2013; **Published** 19 August 2013

Sequencing has seen major breakthroughs in recent years and has paved the way for developing novel lifescience applications. Consequently, the life sciences are facing a rapidly increasing demand for data-handling capacity; in particular, when going from systems biology approaches to applications in, for example, systems medicine, the amounts of data to store, transfer and process overwhelm present-day capacities. New solutions must be developed by the Information and Communication Technologies (ICT) not only to adequately address the current challenges in the life sciences, but also to prepare for a future with an exponential growth of data during the next 10-20 years.

Now is the right time to address this important issue and bring experts from both fields – life sciences and ICT – together to formulate the necessary goals and prepare for this future. To stimulate this dialogue, two initiatives jointly organised a dedicated workshop with the aim of delineating life sciences' ICT-related requirements for the next ten years. Specifically, Babette Regierer, Luca Pireddu, Martijn Moné and Andreas Gisel joined efforts from the COST Action 'SeqAhead' (www.seqahead.eu) and the European ESFRI initiative 'ISBE – Infrastructure for Systems Biology in Europe[1] to analyse the ICT challenges and discuss the needs along the analytical pipeline, from data generation to data integration and modelling. In total, 20 experts from 12 European member states joined the workshop in Pula (Sardinia, Italy) on 6 June 2013 (Figure 1) to share their experience and perspectives on the Big Data challenge and discuss relevant topics that need new solutions. Emphasis was on: i) data processing – *i.e.*, tools and technologies essential to effectively process the avalanche of data; ii) data integration – *i.e.*, ICT and bioinformatics requirements for intelligently modelling and mining the vast bodies of data generated by NGS and other analytical technologies.

The morning session provided keynote presentations on relevant subjects and on-going activities, in order to stimulate discussions in the afternoon breakouts. Dr. Luca Pireddu (CRS4, IT), the local host and co-organiser of the workshop, and Dr. Andreas Gisel (CNR-ITB, IT) as representative of the SeqAhead COST Action introduced the workshop's goal and structure. The first talk by Dr. Martijn Moné (VU University Amsterdam, NL) on "*An infrastructure for European (systems) biology – ISBE*", gave an introduction to the new ESFRI infrastructure for systems biology in Europe, ISBE. Systems biology attempts to understand the functioning of organisms and of life in general via a process that includes data acquisition, analysis, integration and modelling. Sequencing has become one of the fundamental data-acquisition technologies for systems biology studies, exposing ICT challenges that need to be addressed to enable modelling of complex biological systems.

Dr. Babette Regierer (LifeGlimmer GmbH/ SeqAhead, DE) and Dr. Daniel Jameson (University of Manchester, UK) introduced the

1 www.isbe.eu

Figure 1. The SeqAhead/ISBE workshop participants at CRS4 in Pula (Sardinia) on 6 June 2013. (Source: Valentine Svensson, SciLifeLab)

European initiative on '*IT Future of Medicine (ITFoM) – ICT challenges for a virtual patient*'. This initiative aims to create a virtual patient to help health-care professionals better define personalised therapy and prevention strategies. An important goal is to identify major ICT challenges faced by the project, and to develop a roadmap for their resolution. The ITFoM consortium[2] thus developed a concept and roadmap for ICT challenges for the generation and implementation of a virtual patient. Major issues are expected in the optimisation of hardware and software, and in the efficiency of machine-learning and statistical methods; these are not just important for the virtual patient, but are generally applicable across all of the life sciences.

Dr. Simon Heath (Centre Nacional d'Anàlisi Genòmica (CNAG), ES) presented the operations at CNAG '*Dealing with NGS data - the CNAG experience*', which are an excellent example of efficient NGS data processing. The close cooperation with a computer centre shows the advantages of coupling data-generation and computing facilities. The presentation included information about current practices and expected challenges in scaling up the processes. Dr. Luca Pireddu (CRS4, IT) suggested in his presentation on '*Data-intensive computing in NGS* adoption of technologies developed in data-driven computing activities (Big Data) to help address the ICT challenges faced by modern bioinformatics. Distributed computing frameworks like Hadoop[3] could be a suitable solution to scale processing pipelines for sequencing data. Another principle expected to speed up the processing of high-volume data is the use of distributed, column-oriented databases; adopting these technologies, however, requires the creation of new software tools.

As the last speaker of the morning Dr. Heimo Müller (Medical University Graz, AT) presented the 'BiBBox – Biobanking in a box[4], a system that can store and grant access to patient data. The number of samples (*i.e.*, patients) involved in sequencing studies grows steadily; thus, it is im-

2 http://www.itform.eu

3 http://hadoop.apache.org

4 http://bibbox.org

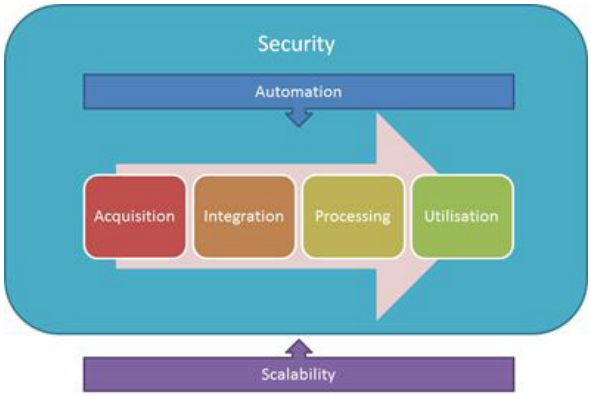


Figure 2. Major steps in the analysis of BIG DATA in the life sciences, including cross-cutting topics, such as scalability, automation and security (Source: D. Jameson).

portant to plan for scalable methods for tracking these samples. BiBBox looks to solve this problem for what could nowadays be considered large studies. On the other hand, the BBMRI initiative[5] looks further into the future, planning for biobanks at the national and European scale. The afternoon session included breakout groups to discuss and catalogue ICT needs and challenges from a range of different expert perspectives, encompassing sequencing technology, (bio) informatics, computer science, systems biology and user communities.

The presentations showed that the topic of the workshop spans a large spectrum of aspects that play key roles in the analysis, handling and management of lifescience data – these include not just steps along the analytical pipeline, like data integration, but also cross-cutting topics such as security (Figure 2).

Summary of the results

The overall process is structured in several different layers: i) the pillars of the principal pipeline, from data generation to modelling, comprise all aspects, from machine to model – and back to the application (*e.g.*, patient); all these areas have certain needs and challenges; ii) cross-cutting issues, like scalability, automation, timeliness, storage, privacy and security, are relevant for all areas; also, organisational aspects, communication, standardisation and metadata are of high relevance, and require development of strategies and agreements on a more general level; iii) time – *i.e.*, while we have to define needs and challenges for the next five years, we also need to prepare now for the next 10-20 years.

In both breakout sessions, the participants identified common themes, including the following:

- data accessibility must be easier and more efficient;
- as many processes as possible must be automated;
- standardisation is a prerequisite to generate automated processes;
- scaling up of data processing and storage needs new approaches (*e.g.*, Hadoop);

CRS4, located in Pula close to Cagliari, was the host institution of the workshop. (Source: CRS4)

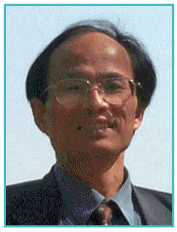
5 http://www.bbmri.eu

- federated approaches might be more successful than a centralised solution;
- metadata and protocols must be standardised and implemented along the whole pipeline (this information must travel with the data);
- standardisation of data, processes, software, and so on, are required;
- intuitive user interfaces are key to allowing the data and tools to be used by non-experts;
- reproducibility and traceability of pipelines.

The participants will initiate a COST Action to address these ICT challenges, to organise the communities involved, to create a strategy for a better communication between the different communities, and to generate a collection of necessary ICT solutions to enable the future of data-intensive life sciences.

Supporting institutions:

Sino-Swiss Workshop on Bioinformatics



Jingchu Luo
Center for Bioinformatics, Peking University, Beijing, China

Received 18 March 2013; **Published** 19 August 2013

As a member of EMBnet, the Global Bioinformatics Network[1], the Center for Bioinformatics[2] (CBI) at Peking University has been working very closely with other member organisations for some 15 years. The collaborations with colleagues from the Swiss Institute of Bioinformatics (SIB)[3] in activities such as exchange visits and Sino-Swiss workshops are among the most memorable and productive.

The first Sino-Swiss Workshop on Bioinformatics (SSWB) was held at CBI in November 1999. Members of the Swiss delegation included Amos Bairoch, the creator of the Swiss-Prot database, and Victor Jongeneel, the then Node manager of EMBnet Switzerland. The late Professor Xiaocheng Gu, founder of the CBI and pioneer of Chinese bioinformatics, hosted the workshop.

A year later, a follow-up SSWB was held in Switzerland, jointly funded by the National Natural Science Foundation of China[4] (NSFC) and the Swiss National Science Foundation[5] (SNSF). The delegation of eight Chinese scientists visited SIB's different partners in Geneva, Lausanne, Basel and Zurich.

Amos Bairoch hosted my presentation in Geneva. His story in running SwissProt, and the high curation standard he brought to the biological database community, are legendary. "It is like a dream!", I started, "I met Amos the first time in London, at the Imperial Cancer Research Fund. We met again last year in Beijing at the first SSWB." As I went on telling the story, "I saw with the amazement that Amos took out a laptop from his rucksack, typed in the plant peptide sequences he'd spotted from one of our posters, and ran a BLAST search against Swiss-Prot, which was installed on his laptop". I was witnessing the creation of the master copy of Swiss-Prot – as the legend has it, the database was all on his laptop. I was left wondering, if there were a match in the BLAST output, what on the poster could be a new piece of annotation to be added to the sequence! The next stop of our visit was Lausanne, where the 2000 EMBnet Annual General Meeting (AGM) was held. The visit was hosted by Ron Appel, the executive director of SIB, and Victor Jongeneel, on behalf of EMBnet Switzerland. Boqin Qiang, a senior geneticist from the Chinese Academy of Medical Sciences[6] (CAMS), gave an introduction about China's involvement in the on-going international Human Genome Project. We then visited the bioinformatics group at Roche in Basel, hosted by Martin Ebeling, the former Node manager of EMBnet Germany.

It was a fruitful visit. With the help of Amos and his colleagues at SIB, we set up an ExPASy mirror at CBI in 2000, making the resource much easier to access for users in China and across Asia.

The collaborations between both sides have continued since then, and I met Amos again several times at various meetings. The photo with Amos and Terri Attwood (Figure 1) was taken at the 2007 EMBnet AGM in Malaga, a southern city in Spain. We sent two young Chinese bioinformaticians to SIB for their post-doctorate training in Lausanne; they came back and developed their

Figure 1. A photo taken during the 2007 EMBnet AGM, in Malaga, Spain. From left to right: Teresa Attwood, Amos Bairoch, Jingchu Luo

1 http://www.embnet.org/
2 http://www.cbi.pku.edu.cn/
3 http://www.isb-sib.ch/
4 http://www.nsfc.gov.cn/
5 http://www.snf.ch/
6 http://www.cams.ac.cn

own labs. Laurent Falquet, the next Node manager of EMBnet Switzerland, visited Beijing in 2008, and gave our students a well-received course.

Soon after the New Year of 2013, a Swiss delegation of ten bioinformaticians, led by Ron Appel, visited Shanghai. Amos joined the delegation as the director of Computer Analysis and Laboratory Investigation of Protein of Human Origin[7], now working on a new protein database NeXtProt[8]. The delegation was also joined by Ioannis Xenarios, who manages SwissProt and ExPASy,

Figure 2. Participants of the Sino-Swiss workshop on Bioinformatics held on 8 January 2013, at Swissnex China in Shanghai.

as well as SIB's HPC infrastructure, VITAL-IT[9]. It was a short visit of only three days, yet a successful workshop was held in the small town of Taicang, north of Shanghai, organised by Swissnex China, Switzerland's outpost for Science, Technology and Culture in China, and the Taicang Institute of Life Sciences Information[10] (TILSI), a new bioinformatics service centre set up by Weimin Zhu, the former head of Database Applications at the EBI. The Chinese delegation was led by the President of CMAS, Xuetao Cao, joined by ten bioinformaticians from organisations at the Chinese Academy of Sciences[11] (CAS), CBI, TILSI and CMAS.

Although time was limited for the talks, participants from both sides did learn a great deal about each others' work in research, development, service and education in bioinformatics, genomics, proteomics, molecular modelling and phylogenetics. Representing the national Node of EMBnet China, my presentation focused on the bioinformatics service and education in which we have been involved during the past 15 years (Figure 3).

Figure 3. Presentation by Jingchu Luo at the Sino-Swiss Workshop, 2013.

It was proposed as early as in 1999 by a Chinese Academician, Bailin Hao, to build a National Bioinformatics Centre of China (NBCC), which would play the same role as the NCBI and EBI, providing bioinformatics resources, services and training. Although many efforts have been made toward this goal, its organisational model and infrastructure have yet to be determined and built. The operational model of the proposed NBCC is one of the key questions. At this workshop, Chinese delegates listened with great interest to the description of the federated model of SIB. A task force, funded by CAS, was recently created, with the goal of conducting a thorough survey to address some of the important questions, and produce recommendations to funding agencies. The task force is in the process of organising a trip to SIB in 2013, to learn more from Swiss colleagues about building and running a successful bioinformatics centre at national level. Trips to NCBI and EBI are also planned.

Acknowledgements

Thanks to WM Zhu and TK Attwood for critical reading of the manuscript.

7 http://www.isbsib.ch/groups/geneva/calipho-bairoch.html
8 http://www.nextprot.org/
9 http://www.vital-it.ch
10 http://www.tilsi.org/
11 http://www.cas.cn/

Summary Report of Centre for Proteomic and Genomic Research - 2012


Reinhard Hiller, Judit Kumuthini
Centre for Proteomic and Genomic Research (CPGR), Cape Town, South Africa

Received 1 August 2013; **Published** 19 August 2013

CPGR background

The CPGR is based on an initiative by the South African Department of Science and Technology (DST), and funded by the Technology Innovation Agency (TIA), to support the development of an internationally competitive bio-economy in South Africa (SA).

Created in 2006, the organisation's vision is to establish a modern, world-class facility that serves the needs of the scientific community in SA by providing state-of-the-art services, technical expertise and collaborative research capabilities in the genomics and proteomics arena.

Based in Cape Town, the CPGR was established as a not-for-profit, contract research organisation to provide support and services to the life science and biotech communities. To this end, it combines state-of-the-art, information-rich genomic and proteomic ('omics') technologies with bio-computational pipelines to create unique solutions for biological problems.

By applying principles of network orchestration, the CPGR combines internal and external resources to create the economies of scale and scope necessary for delivering genomic, proteomic and bioinformatic support in a high-quality, cost-effective fashion[1]. Orchestrating capacity in a networked way allows the CPGR to make agile responses to a wide range of needs and render fit-for-purpose solutions; this approach also facilities the creation and diffusion of knowledge, a pre-requisite for innovation in any sector, and a must in the rapidly evolving 'omics' arena. Ultimately, our aim is to create an eco-system conducive to stimulating life science innovation in Africa!

The CPGR's vision is to be a key driver in SA's effort to become one of the leading bio-economies of the 21st century. Our mission is to be an 'omics' technology platform that provides solutions to innovation gaps and identified opportunities in the development of SA's modern biotech sector!

CPGR offering

Highly skilled laboratory staff maintain and run the key pieces of genomic and proteomic equipment that provide critical service and project support to clients. The organisation can handle samples and isolates from most biological sources, including human, animal, plant, yeast, bacteria and viruses, in a secure Biohazard Class II (BSL II) environment. To create effective solutions in complex biological projects, the CPGR employs a wide range of validated genomic, proteomic and bioinformatic workflows. These include array-based RNA expression profiling and DNA genotyping using Affymetrix cartridge- and GeneTitan-arrays; whole-genome, exome and transcriptome sequencing on a variety of high-throughput sequencing platforms (Illumina, LifeTech, 454); quantitative DNA and RNA detection on digital and qRT-PCR platforms; protein identification and biomarker discovery using a suite of state-of-the-art mass spectrometers (MALDI-ToF/ToF, Q-Exactive, TSQ Vantage, Q-ToF, amongst others); and multiplex biomarker profiling using solid protein arrays and bead-based suspension arrays. Computational workflows for high-throughput analysis of genomic and proteomic data-sets, including standard DNA and next-generation sequencing data analysis complete the portfolio.

Complete Genomics & Proteomics services

The efficient integration of a range of technologies and workflows allows us to render complete genomics and proteomics services. The value we add to scientists is based on the fact that we can choose from a range of options and devise custom solutions that meet diverse requirements, such as throughput, coverage, depth, and costs

[1] For a detailed discussion of the approach, visit http://www.cpgr.org.za/blogspot/?p=367

Figure 1. Jo McBride, Affymetrix platform manager doing the QC which is part of all the Affymetrix workflow implemented at the CPGR.

in genomic and proteomic projects. In response to clients' needs, we can assemble available know-how and resources into fit-for-purpose value propositions.

Through the flexible combination of technologies, and adherence to stage-specific quality management principles, we have the ability to render support and services across the entire genomic and proteomic innovation chain.

Quality is an intrinsic part of all of the organisation's endeavours: it pervades our service processes, product development, and all of our communication. In support of generating good quality outputs, the CPGR employs a modular approach to quality assurance and control. Overall, the organisation is based on the International Organization for Standardization (ISO) approach to quality management, while in projects we adhere to good clinical laboratory practice (G(C)LP) principles. During 2013, the CPGR is preparing for certification according to ISO 9001:2008, building the foundation for the accreditation of individual workflows, as needed. Relevant CPGR staff are trained in G(C)LP and in lean management principles, amongst others.

Project management in support of life science innovation

Technological versatility is relevant in projects aimed at the development of biomarkers for diagnosis and treatment of human diseases. Biomarker development is a lengthy process, unfolding over three essential stages: analytical validity, clinical validity and clinical utility. Often, initially promising discoveries fail to pass later development hurdles, owing to problems in the design, execution or reporting of genomic and proteomic projects. Frequently, 'omics'-driven innovation does not occur for lack of validation of research outputs[2].

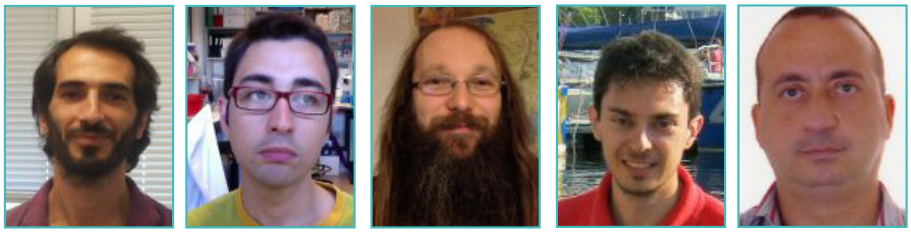
We believe that genuine capacity development in the modern life sciences, aimed at genomics-and proteomics-driven innovation on the African continent, rests on the triple pillars of access to world-class infrastructure, provision of high-quality affordable services, and empowerment through training, in particular in the field of data-analysis and interpretation (bioinformatics).

In order to facilitate a seamless conversion of project ideas into robust project plans, the CPGR has devised an integrated framework for the design and management of genomic and proteomic projects[2]. The framework was developed by taking into account CPGR expertise and inputs from external experts in the field of genomic and proteomic biomarker development. Ultimately, it is a risk-mitigation framework developed to facilitate a more effective migration of early-stage research outputs into downstream development, and translation into products and services.

Figure 2. Jo McBride, Affymetrix platform manager loading arrays into Affymetrix hybridization oven.

[2] For a discussion of the corresponding problems, visit http://www.cpgr.org.za/blogspot/?p=201

Quick Direct-method Controlled (QDC): a simulator of metabolic experiments



Davide Cangelosi[1], Salvatore Fabbiano[2], Claudio Felicioli[3], Luca Freschi[4], Roberto Marangoni[5,*]

[1]Institute Gianna Gaslini, Genova, Italy
[2]University of Salamanca, Spain
[3]Noname Research, Italy
[4]Universite' Laval, Quebec City, Canada
[5]University of Pisa, Italy
* corresponding author (marangon@di.unipi.it)

Received 10 August 2012; **Accepted** 11 Feb-2013; **Published** 19 August 2013
Competing Interests: none

ABSTRACT
Quick Direct-method Controlled (QDC) isa stochastic simulator based on the direct method version of Gillespie's Stochastic Simulation Algorithm(SSA). It has been specifically designed to simulate experiments performed on metabolic networks, when external operators can act on the system, modifying its spontaneous behaviour. Users of QDC can simulate different experimental controls:i.e.,add or remove chemical species at a given time; change the rate of a reaction at a given moment; and describe reactions with complex stoichiometrythat take place once the stoichiometric condition is verified (here called immediate reactions). Moreover, even though QDC is not designed to manage compartments, it can simulate up-take and excretion reactions. QDC represents a useful tool for the specific field of interest thanks to its computational performances and simple input language.

Introduction

Gillespies' Stochastic Simulation Algorithm (SSA) (Gillespie, 1977) is the most widely used approach to simulate the stochastic behaviour of metabolic networks, in particular when the number of molecules involved is relatively small. There are several variants of SSA: the direct-method was the first proposed, followed by many others that address different issues related to the reduction of high computational costs, or dealing with stiff systems, etc. (for a review, see Li *et al.*, 2008). Modern stochastic simulation suites offer wide libraries of different simulation methods, analysis facilities and numerical resources, and can face multi-scale systems, hybrid deterministic and stochastic representations, and many other problems that occur in systems biology. Nevertheless, during some of the studies we carried out on signalling pathways and minimal cells,we needed to extend the description of the target systemby including at least three control actions that would be possible to perform on it: a) to add/rule out a given amount of molecules of a given species at a given time; b) to change the propensity of a reaction,depending on the presence/absence of an external trigger signal; and c) to specify all-or-nothing reactions.

We first tried to describe such extended models using different existing simulation packages, but we encountered severe difficulties; therefore, we developed QDC specifically to match these needs in an easy and efficient way. QDC does

not aim to compete with existing suites, but rather, to complement them for these specific tasks.

QDC

QDC's syntax and input file structure

QDC's core input is represented by an ASCII file written in anin-house format (see Figure 1). Supplementary file 1[1] offers a detailed description of the QDC input language and usage;more detail is presented in the README file that comes with the package download. Here, we give a brief summary.The formal definition of the whole syntax is presented in Supplementary file 2[2].

The input file is structured in sequential blocks, separated by a blank line; each block contains a different category of information: block B1 declares the names of the chemical species present in the system; B2 declares the system volume, measured in litres, that is used for computing stochastic propensities; B3 declares the reactions, using a notation very close to that of standard biochemistry. Each reaction is introduced by the deterministic reaction rate coefficient, which QDC transforms into the stochastic propensity, according to the reaction order, system volume and species concentrations. Among the different reactions, QDC allows users to declare zero-order reactions, useful to simulate the uptake of a chemical species from a generic outside. The left-hand member of the chemical equation is here represented by the operator 'NULL'. First-order reactions with the right-hand part of the equation containing only the operator NULL can be used to simulate excretion reactions. A special reaction class is constituted by the so-called immediate reactions. Theseare introduced by a dash sign instead of a coefficient, and take place immediately after the verification of the condition represented by the left-hand side of the equation. In other words, the stoichiometry noted on the left-hand side of the equation is a logical condition: once it is verified, the immediate reaction takes place, and yields the products indicated on the right-hand side. These reactions allow for complex stoichiometry, where both the reagent species number and the number of molecules per species can be greater than two. We want to remark that *immediate reactions* are not to

```
a, b, c, d

volume, 0.000001

0.1, a > b
$k, c > d

a, 1000
b, 0
c, 1000
d, 0

$k,

0, $k, 0.1
1, 7.0, $k, 0.01
```

Figure 1.

be considered higher-order reactions,asthey do not have a kinetic law, and they happen immediately. Immediate reactions can also be used to simulate the simultaneous excretion of several chemical species, once they have reached a given threshold number of molecules: to do so, the right-hand side of the equation must contain only the 'NULL'term.

Immediate reactions represent the most prominent innovative feature implemented in QDC with respect to other simulation packages. Immediate reactions can be seen as infinite-propensity reactions, as, in fact, they are. Nevertheless, they avoid several implementation problems that a generic infinite-propensity reaction can have. First, immediate reactions avoid the zero-times-infinity problem that will affect a generic infinite-propensity reaction when the reagent species concentrations are nearly zero; second, immediate reactions make it no longer necessary to describe the kinetic law for reaction of order higher than two. Immediate reactions turn out to be very useful in describing all-or-nothing reactions that take place in different biological contexts. For example, the firing of a neuron, which happens when enough synaptic stimulations are received, can be modelled by using immediate reactions.

B4 contains the number of molecules assigned to all the declared chemical species that

1 http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/505/505_supp_1

2 http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/505/505_supp_1

are supplied to the system, along with the time at which they are supplied. When such time is set to zero, the statement denotes the initial number of molecules for the declared species.

B5 and B6(optional) are concerned with the eventual presence of controls exerted on kinetic coefficients that are not constant during the simulation, but change according to some trigger signal: in particular,B5declares the name of all dynamic coefficients used in the simulation (their name begins with a dollar sign); B6 declares the starting instant of the value change and the new value assumed by the variable.

QDC's core

QDC's core, written in C++, implements Gillespie's direct method (Gillespie, 1977) to simulate the time course of a biochemical system. In order to avoid possible violations of the algorithm correctness, QDC does not use any approximation. Given a metabolic system (and eventual actions performed on it) described in an ASCII input file, QDC parses it into C++ source files, compiles and executes them. This procedure has been designed to specialise the source file on the given model, thus allowing one to fully exploit the compiler optimisations. Thanks to this meta-compilation approach, which, to the best of our knowledge,is used for the first time in a metabolic network simulator, QDC offers very good computational performances.

QDC's output

QDC outputs four files: the first contains the time course of the number of molecules of each simulated metabolite; the second contains the counters of each metabolic reaction firing; the third contains the time course of the propensity for each reaction, and the fourth is a log-file of the computation. Files concerning reaction-fire counts and time-course of propensities complement the information contained in the first, helping users to reconstruct more accurate views of the dynamics of the system. For instance, if a given chemical species counts 0 molecules in the reagent file, one cannot conclude that it does not exist: it may be that the genesisreactions are slow with respect to the consuming reactions – then the balance of that species at any sampling point will result in 0, but the species is dynamically existent and gives its contribution to the system evolution. This can be detected by inspecting the reaction-counter file, where one can assess that both genesisreactions and consuming reactions have effectively fired. By examining the propensities time-course file, one can understand which reactionscommandthe highest importance in the system at a given time, therebyrevealingthe evolution of hub-pathways over time.

SBML Import/Export

The QDC package also contains two applications (import_sbml.py and export_sbml.py) that provide the Import(I)/Export(E) from/to SBML[3] level 2, ver. 4, thanks to the libSBML v.4.0[4] libraries. Of course, such I/E is limited to the expressions and the statements that both the languages (of SBML and QDC) support. The SBML I/E can also be managed via the Graphical User Interface (GUI). Supplementary file 3[5] gives a detailed description in SBML of the three main control events managed by QDC.

QDC's GUI

QDC's GUI has been developed to giveusers an immediate visualisation of the simulated system behaviour. The GUI is basic and easy to use: it has been developed in Python v.2.6 and uses the PyQt libraries[6] to manage the interface's elements. This choice confers good portability to-QDC's GUI, as it has been tested on different Linux distributions (Fedora, Ubuntu, *etc.*) and on Mac OS X.

A benchmark test

A very basic benchmark test was run, based on a comparison between QDC and three widely used simulators –StochKit (Kierzek, 2002), Dizzy (Ramsey *et al.*, 2005) and BetaWorkBench (Dematté *et al.*, 2008) – which offer an implementation of Gillespie's direct method. To compare the relative efficiency in implementing Gillespie's algorithm, it is not correct to compare the runtime required by each simulator in simulating the same input model: this kind of measure, in fact, is influenced by other implementation characteristics (disk usage, memory allocation, etc.). Instead, we ran several simulations of the same biochemical model (one strictly similar to that presented in Figure 1, but without control variables)by using the same computer, by varying the simulated time, and measuringthe machine time required to per-

3 www.sbml.org
4 http://sbml.org/Software/libSBML
5 http://journal.embnet.org/index.php/embnetjournal/article/downloadSuppFile/505/505_supp_1
6 http://www.riverbankcomputing.co.uk/software/pyqt

form the task. We built, for each simulator, a curve of dependence of the actual elapsed time vs. the simulated time.There is a region where these curves are quasi-linear (when the simulation requires a number of operations significantly greater than those necessary to launch the program, but not so high as to require the computer to start swapping). We determined, by using a linear regression, the slope of these lines, which represents the coefficient linking the elapsed time to the simulation time. Their average values were as follows: StochKit = 7.25 10^{-3}, Dizzy = 5.8 10^{-3}, BetaWorkBench = 4.09 10^{-3} and QDC= 2.75 10^{-3} (these average values were computed on 1,000 runs; the standard error was less than 10% of the value). We remark that this test does not represent a study on QDC's complexity, nor a proper screening of QDC performance: it represents only an indication that QDC's efficiency can be assumed comparable with that of other freely available simulators.

Availability

QDC is freely available, under the GPL v3 license, through the SourceForge platform[7].

Acknowledgments

The authors thank Davide Chiarugi and Lorenzo Lazzerini for the validation tests they executed,and Roberto Grossi and Pierpaolo Degano for their critical reading of the manuscript, and the formal syntax language in Supplementary file 2.

References

Dematté L, Priami C, Romanel A (2008) The Beta Workbench: a computational tool to study the dynamics of biological systems. *Briefings in Bioinformatics***9**, 437-449.doi: 10.1093/bib/bbn023.

Gillespie DT (1977) Exact Stochastic Simulation of Coupled Chemical Reactions.*J. Phys. Chem.***81**, 2340-2361. doi: 10.1021/j100540a008.

Kierzek AM(2002) STOCKS: Stochastic Kinetic Simulations of biochemical systems with Gillespie algorithm.*Bioinformatics***18**, 470-481. doi: 10.1093/bioinformatics/18.3.470.

Li H, Cao Y, Petzold LR, Gillespie DT(2008) Algorithms and Software for Stochastic Simulation of Biochemical Reacting Systems. *Biotechnol. Prog.***24**, 56-61.doi: 10.1093/bib/bbn050.

Ramsey S, Orrell D, Bolouri H(2005) DIZZY: Stochastic simulation of large-scale genetics regulatory networks.*J. Bioinf. Comp. Biol.***3**, 415-436. doi: 10.1142/S0219720005001132.

7 http://sourceforge.net/projects/gillespie-qdc

News & Views from around the ISB Community

Check if your database or repository is missing

Elsevier is keen to work with researchers and data repositories to ensure that data that is relevant for scientific, technical and medical research can be easily discovered and accessed. One of the ways in which we are doing this is by creating bidirectional links between data repositories and online articles on ScienceDirect. This provides ScienceDirect's readers with one-click access to relevant, trusted data that may help to validate research or drive further investigations. Linking helps to make articles and data better discoverable, attracting more usage. Sharing the data that underpins conclusions is not only good scientific practice, but also increasingly required by funding bodies.

Specific journal instructions for authors depend on the data repository: in some cases data is extracted from the article by curators, while in other cases authors need to upload their data manually.

We need to hear from you with your detailed instructions. Detailed information for already established partnerships is available in the (recently revised and updated) listing of supported databases. Alternatively, copy & paste this link: http://goo.gl/8Jh8k

If you are a data repository manager interested in setting up bidirectional linking with Elsevier publications, please contact us at articleofthefuture@elsevier.com

Submitted by Adriaan Klinkenberg.

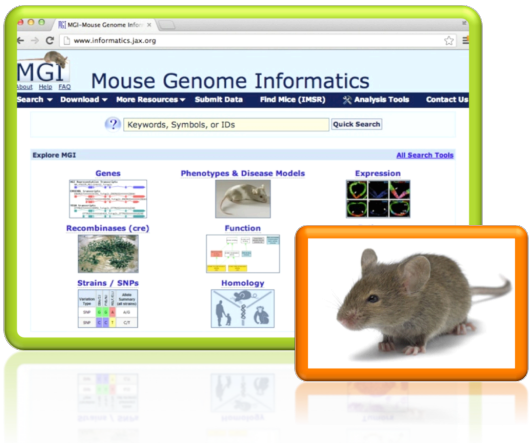
The M:N Project at MGD: Beyond 1:1 Orthology Assertions

The Mouse Genome Database (MGD) curates, integrates, and provides comprehensive genetics, genomic, and phenotypic information for the laboratory mouse, a primary model organism for experimental investigation of human biology and disease. MGD is found at http://www.informatics.jax.org.

A core component of MGD data for over 20 years has been the curated assertion of 1:1 orthology between mouse, human, and rat protein-coding genes. Now, with completely sequenced genomes available for comparative analysis, phylogenetic analysis clearly identifies cases where descent from common ancestor does not always define a 1:1 relationship, but rather that gene duplication following an ancestral speciation event more correctly results in M:N relationship between genes in different species.

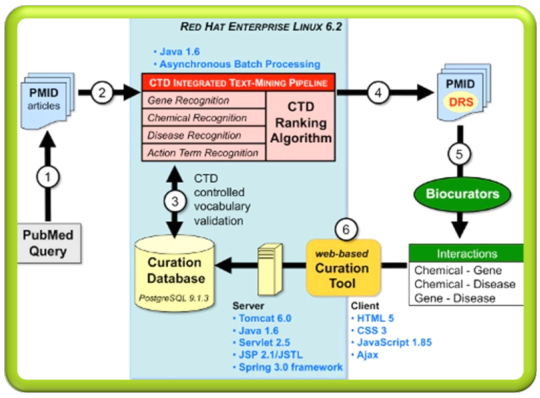
This has implications for the study of human biology in the mouse system and for the presentation of inferential functional and disease associated assertions based on comparative analysis. MGD has recently restructured its database to accommodate such homology classes with concurrent changes in presentation of data related to homology classes and in the representation of human diseases associated with mouse genes by curation of comparative or experimental data. We load data from all mammalian species with completed genome sequences, and will next extend our data to include chicken and Zebrafish protein-coding gene classes. While 1:1 assertions predominate (~80%), we now more clearly represent cases such as the Serpina1 gene class (1 human, 5 mouse, 1 rat), and provide better cross-referencing among related genes, the diseases that have been studied in respect to those genes, and the relationship between genomic features in related genomes. This work is supported by NIH NHGRI grant HG-000330.

Submitted by Judith A Blake, Richard Baldarelli, Mary Dolan, Mark Airey, Jon Beal, Sharon Giannatto, David Miers, Jill Lewis, Carol J. Bult, Janan T. Eppig and James Kadin.



The Comparative Toxicogenomics Database Text-Mining Pipeline

The Comparative Toxicogenomics Database (CTD; http://ctdbase.org/) confirmed the effectiveness of its text-mining pipeline in evaluating and prioritizing scientific literature for the manual curation of chemical-gene-disease information. The results were selected for a special issue of PLoS Text Mining Collection in April (http://goo.gl/f1ubr).



For the study, CTD tested their sophisticated text-mining algorithm by using it to evaluate the text from 15,000 articles and assign a relevancy score to each document. A representative sample of the corpus was sent to their team of biocurators to manually read and evaluate on their own, blind to the computer's score. The biocurators concurred with the algorithm 85% of the time with respect to the highest-scored papers, and there was a clear step-wise progression, wherein the likelihood of an article's true relevancy decreased linearly as the text-mining scores declined.

Ranking papers by text mining allowed biocurators to focus on the most relevant papers and avoid the extraneous ones, increasing productivity by 27% and novel data content by 2-fold. The curated articles were also broad and encompassing with respect to data coverage, finding both shared as well as unique biological processes, pathways, and toxicological end-points, confirming that the ranking system could help identify articles that contribute to a mechanistic understanding of toxicity.

By incorporating similar text mining-based scoring, other databases may also be able to enhance their manual curation by prioritizing more relevant articles, thereby increasing data content, productivity, and efficiency.

Submitted by Allan Peter Davis.

ISB Spotlight provides a snapshot of some of the work and activities of members of the International Society for Biocuration (ISB). The Spotlight features brief descriptions of a range of databases and bio-curation tools, re-published, with permission, from the 'News & Views' section of ISB's monthly newsletter. The newsletter, with the complete 'News & Views from the ISB Community', is freely available from http://biocurator.org/newsletter.shtml.

a case for discomfort

Vivienne Baillie Gerritsen

There is no life without smells. In the wild, smells – and the capacity to sense them – are the basis for survival for plants and animals. They are used to attract, seduce, repel or protect, and are with us night and day; so much so that life would seem very bland without them. On the whole, for any given species, a pleasant perfume implies that all is well, while a bad one suggests that something is wrong. The smells the human body gives off are a combination of who we are, what we eat, and the general state of certain metabolic cycles. When part of a metabolic cycle is deficient, a change in our bodily odours can occur. This is the unfortunate case of what is known as the 'fish malodour syndrome', or trimethylaminuria. People afflicted with trimethylaminuria release a smell of rotting fish. The symptoms were first clinically described in the 1970s and, in the 1990s, scientists discovered the cause: a malfunction of an enzyme known as flavin-containing monooxygenase 3.

"Ish Fish", Sergey Engel (Israel)

Courtesy of the artist

Though the condition was first clearly described in the 1970s, there is little doubt that it had already occurred in a human being. Anecdotal descriptions appear in the ancient literature. The first known depiction goes back almost 3'000 years and recounts the case of an Indian woman who smelt so bad that she had been cast from society. About 750 years ago, it seems that the syndrome was the cause of suicide amongst Thai concubines. Shakespeare himself, obviously came across the condition since he portrays a character who smells of fish in 'The Tempest'; a 'Poor John' he calls him which, in those days, referred to a salted and dried hake. Various other descriptions followed but the first official clinical report was made in the 1970s, on a six year old girl who also suffered from Noonan syndrome*.

Consequently, for a time, not only was it believed that trimethylaminuria was rare but that it was probably linked to other afflictions, such as Noonan syndrome. Over the years, however, scientists have discovered that trimethylaminuria is hereditary and caused by the presence of too much trimethylamine in bodily excretions, i.e. in breath, sweat, saliva and vaginal secretions.

Trimethylamine – or TMA – is the result of bacterial digestion of foods that are rich in choline and carnitine, such as eggs, liver, peas, soybeans and sea fish for example. When everything is working properly, TMA is modified by our liver so that it becomes water soluble. This procedure not only facilitates TMA excretion but also makes the amine odourless. On the whole, for 50 mg of TMA, a human being usually excretes about 1 mg of non-modified 'smelly' TMA, while the rest has been transformed and is odourless.

What is happening inside someone who is suffering from fish malodour syndrome? There are a number of enzymes in our liver that spend their time modifying chemical compounds for further use or excretion. One of these enzymes is a flavin-containing monooxygenase of which

there are five different kinds: FMO1 to 5. Flavin-containing monooxygenase 3 – or FMO3 – is the major catalytic form found in the liver. In normal conditions, FMO3 oxidises incoming TMA, which is further modified to increase water solubility and hence facilitate excretion. The oxidation of TMA by FMO3 acts like a deodorant and wipes away its fishy smell. If FMO3 is faulty, however, the oxidising step is skipped and TMA leaves our body the way it entered it: with a smell.

So far, 150 different mutations have been found in FMO3, which alter the enzyme in such a way that it can no longer work the way it usually does. Despite this abundance of mutations, researchers are still not clear on how the mutations affect TMA oxidation. What they do know, though, is that the mutations do not affect the flavin binding sites. But they may affect the protein's tertiary structure, or directly affect the site where TMA is supposed to bind for its subsequent oxidation. And, if TMA is unable to bind to FMO3, it will not be oxidised.

An intriguing discovery is that fish malodour syndrome seems to go hand in hand with other conditions. Psychological disorders, such as depression and psychoses are frequently stated. Naturally, the simple fact of not smelling nice in society is something extremely difficult to deal with and hardly surprising that it causes individuals to develop antisocial behaviour. Especially as the syndrome can affect children at a very early age. Despite this, researchers are now positive that fish malodour syndrome is frequently associated with neurochemical disorders. This is probably due to the fact that – besides TMA – FMO3 also oxidises other amines which may have a direct effect on depression and psychoses. Foreign compounds that enter our bodies – such as pollutants or environmental chemicals that predispose us to carcinogenesis and birth defects for instance – may also be affected by a deficiency in FMO3.

In the case of fish malodour syndrome, very little can be done. Diets that diminish the amount of TMA entering the body can be followed. Medication to reduce our floral bacterial activity – and hence the amount of TMA released – can also be taken. In the future, we could imagine engineering microorganisms to produce FMO3 in our digestive tract, or the development of suppressants that are able to mask the rotten fish smell. Of course, gene therapy could be used to replace the deficient FMO3 gene by a healthy one... It is always an astonishing thing to realise how such minute changes in our chemical composition are able to affect a person's whole life. How vulnerable each one of us is with respect to our chemical make-up.

* ***See Spotlight issue 121, 'The Matchmaker'.***

Cross-references to UniProt

Dimethylaniline monooxygenase [N-oxide-forming] 3, *Homo sapiens* (Human) : P31513

References

1. Mitchell S.C., Smith R.L.
Trimethylaminuria: The fish malodor syndrome
Drug Metabolism and Disposition 29:517-521(2001)
PMID: 11259343

2. Rehman H.U.
Fish odour syndrome
Postgraduate Medical Journal 75:451-452(1999)
PMID: 10646019

3. Treacy E.P., Akerman B.R., Chow L.M.L., Youil R., Bibeau C., Lin J., Bruce A.G., Knight M., Danks D.M., Cashman J.R., Forrest S.M.
Mutations of the flavin-containing monooxygenase gene (FMO3) cause trimethylaminuria, a defect in detoxication
Human Molecular Genetics 7:839-845(1998)
PMID: 9536088

National Nodes

Argentina
IBBM, Facultad de Cs. Exactas, Universidad de Buenos Aires, Buenos Aires

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

Egypt
Nile University, Giza

Finland
CSC, Espoo

France
ReNaBi, French bioinformatics platforms network, Villeurbanne Cedex

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

Luxembourg
Luxembourg Centre for Systems Biomedicine (LCSB), Luxembourg

The Netherlands
Centre for Molecular and Biomolecular Informatics (CMBI), Nijmegen

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

Norway
The Norwegian EMBnet Node, The Biotechnology Centre of Oslo, 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

United Kingdom
The Genome Analysis Centre (TGAC), Norwich

Specialist- and Assoc. Nodes

CASPUR
Rome, Italy

EBI
EBI Embl Outstation, Hinxton, Cambridge, UK

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 p/a
CMBI Radboud University
Nijmegen Medical Centre
6581 GB Nijmegen
The Netherlands

Email: erik.bongcam@slu.se
Tel: +46-18-4716696