SEB/GOBLET Bioinformatics Workshop
An essential tool for experimental biologists
University of Manchester, 4 July 2013

This short workshop, held as part of the Education and Public Affairs programme of the Society for Experimental Biology(SEB)'s annual meeting, was designed to offer participants an opportunity to discuss some of the 'bioinformatics challenges' they encounter on a routine basis during the course of their work (in terms of choosing the right database or statistical analysis package, how to use freely available software tools, how to interpret the results, *etc.*); it also aimed to allow more general discussion of training approaches that could help to improve their confidence in using bioinformatics tools and resources (whether via focused workshops, technology-enhanced learning, university courses, *etc.*).

The programme, which included brief presentations and small group discussions, was divided into two 90-minute sessions. The first of these built on the results of the survey conducted with SEB members early in 2013, and of the wider survey of GOBLET communities conducted early in 2014, both of which aimed to understand the type and level of bioinformatics training that biologists need to facilitate their work. The second session was included to try to raise awareness of [GOBLET](#) and its training portal, via which life scientists may identify suitable distance- or technology-enhanced learning resources, workshops, courses, summer schools and other training events around the globe.

Sarah Blackford kicked off the event by introducing the workshop, its origins and its aims; she welcomed the participants and presented the workshop facilitators, Angela Davies and Terri Attwood. The number of participants was small (16), so each was asked to say a few words about him- or herself: who they were, where they were from, why they were particularly interested in this workshop, and what they hoped to get out of it. The group included PhD students, postdoctoral researchers and group leaders from a range of countries (including the UK, Portugal, Brazil, Israel and Australia); most indicated that they wanted to learn more about where to find basic training information and how to get help.

Session 1: short background to the survey (Angela Davies)
Angela began by explaining that anecdotal evidence of bioinformatics skills gaps amongst life scientists had prompted GOBLET to work with the SEB to run a brief survey within the SEB community, essentially to try to capture a snap-shot of training needs. The survey, which garnered ~200 responses, included questions on the ways in which bioinformatics training and education had been received; the level of confidence in using bioinformatics databases, software and, especially, command-line tools; how bioinformatics training should be delivered; and what skills, in particular, were needed.

Overall, the results showed some interesting trends. It was therefore agreed to push the survey out more widely through GOBLET partner organisations to try to establish whether the initial findings were general across global communities

of life scientists; this second survey attracted around 500 responses. The results from the two surveys are broadly compared in Table 1.

Table 1: *Comparison of results from the SEB (2013) and GOBLET (2014) surveys.*

Item	Observation	SEB	GOBLET
Respondents (by career stage)	Academic/PI	~25%	**~23%**
	Postdoc/research staff	~35%	**~24%**
	MSc/PhD students	~32%	**~26%**
	Technical staff	~3%	**~5%**
	Biocurator/bioinformatician	~5%	**~9%**
	Manager & other		**~13%**
Respondents (by discipline)	Plant science	~33%	**~6%**
	Biomedical	~15%	**~14%**
	Microbiology/virology	~7%	**~8%**
	Marine/fish	~5%	**~1%**
	Biochemistry/genomics/epigenetics/ molecular biology/bioinformatics…	~40%	**~71%**
Method of learning bioinformatics	Self taught & colleagues	~76%	**~42%**
	Masters/UG courses	~20%	**~21%**
	UG		**~7%**
	Other university courses		**~5%**
	Professional training courses	~30%	**~14%**
	Internet/online courses	~4%	**~8%**
	Other		**~3%**
Level of confidence using current bioinformatics tools *etc.* (where expressed)	5 (most)	~10%	**~19%**
	4	~20%	**~24%**
	3	~34%	**~18%**
	2	~16%	**~21%**
	1 (least)	~7 %	**~18%**
Main bioinformatics skills training required	Analysing & interpreting data	~67%	**~49%**
	Statistics		**~5%**
	Programming	~10%	**~8%**
	Command Line		**~3%**
	Computer Science		**~1%**
	Data management		**~1%**
	Database selection and usage	~16%	**~8%**
	Software tools (using correctly/ adapting)	~35%	**~17%**
	Sequencing analysis (NGS, *etc.*)	~5%	**~4%**
	All		**~4%**
Preferred methods of delivery	Workshop (stand-alone and within own institute/university)	~74%	**~19%**
	UG & PG courses		**~10%**
	e-learning (but with support/access to workshop/Q&As)	~57%	**~25%**
	Other		**~1%**
	Science conference (as a stand-alone workshop/satellite of the conference)	~14%	**~35%**
	Summer schools		**~11%**

The results show that similar proportions of academics and students were surveyed, with a small number of technical staff in each case. For the GOBLET survey, there was a large 'other' category, mostly comprising bioinformaticians, biocurators, statisticians and those who described themselves as managers.

Most respondents to the SEB survey appeared to be life scientists, with one third being involved in plant science and one third in 'other' disciplines, including biochemistry, genomics, epigenetics, molecular biology and bioinformatics. By contrast, as might be expected, the GOBLET survey appears to have reached a rather broader range of bioinformaticians, biochemists/biophysicists, computer scientists, mathematicians and genomic scientists.

Similar trends were seen in both sets of results in terms of how bioinformatics training/education had been received: 'self-taught' and 'via colleagues' were the most reported mechanisms for acquiring bioinformatics skills. Interestingly, Internet/online courses received fewer counts in each case, but this is probably because some of the self-taught training encompassed courses of this type.

By contrast, when asked how responders would like to receive bioinformatics education/training in future, most requested stand-alone workshops (either in the workplace or alongside conferences) and online learning/e-learning, especially with tutor support. Interestingly, undergraduate and graduate programmes seemed rather less popular as routes for bioinformatics training.

Finally, and curiously, confidence with data analysis techniques was reported as relatively high, yet this is clearly still the most sought after skill-set; on the other hand, confidence with command-line software tools was reported as relatively low, yet training in programming was much less sought after.

In addition to the standing questions, the survey also solicited free-text feedback, which helped to 'personalise' the results – a flavour is given here:

- "*It is too late for me, we need to fix the undergraduate syllabus – all biology students should be encouraged to learn programming and statistics*",
- "*User friendly courses for biologists with no programming skills [are needed]*",
- "*Online course ok if includes a good opportunity to discuss with teachers*",
- "*So many learning sites/tutorials/tools – difficult to pick the 'best one'*", and,
- "*As an absolute beginner it is very difficult to even begin to try to use command-line software and all previous courses have been too advanced for someone who has never used Linux OS. It would be helpful if there were opportunities to work through example data in small groups with a tutor.*"

Angela completed her presentation by explaining that GOBLET would like to expand its communities and programmes to try to develop bioinformatics resources to address the kinds of training and education needs highlighted in the two surveys. She then invited the participants to discuss the following questions:

i)   Do the survey trends resonate with you? If not, why not?
ii)   Are there gaps in the survey results? If so, what are they?
iii)   What could universities/other providers do to help (*e.g.*, CPD – if so, in what form – distance, blended, short workshops…?).

Break-out session 1 - Review of bioinformatics training needs

The participants broke into small groups to discuss these three questions, each selecting a rapporteur to convey the outcomes during the reportage session.

In general, the groups felt that the survey results matched their own concerns in terms of their confidence in data analysis and interpretation; some suggested that acquiring knowledge of maths and statistics was more important than programming skills, because the real difficulty was in choosing the right kind of analysis to perform.

The groups also echoed the finding that more training was needed via workshops and short courses, including e-learning resources. It was suggested that access to a range of case studies and establishment of permanent study groups would be helpful; some placed particular emphasis on the need for greater encouragement when using bioinformatics software tools and databases. One suggestion was that there should be ways to test or evaluate the skills of those who have undergone bioinformatics training, perhaps through some kind of accreditation mechanism.

Overall, it was emphasised that students should be repeatedly reminded of the importance of bioinformatics and biostatistics throughout their studies, commencing with basic courses at undergraduate level, in order to introduce the necessary skills at an early stage. Until then, the general feeling was that in order to find the training they needed, "*Google is our best friend*!"

Session 2: short introduction to GOBLET (Terri Attwood)

In this second session, to try to provide more context for the two surveys, Terri explained a little more about what GOBLET is and how it came about, who's in it (and who else should be in it), what members of the Foundation have been doing during the last 18 months and, hopefully, what it can do to help.

About GOBLET

The idea to create GOBLET emerged when representatives of ten international and national networks and societies met in Uppsala, in June 2012, to discuss world-wide bioinformatics training activities and, specifically, how their diverse efforts could be harmonised by working together more efficiently. To help coordinate the spectrum of training activities represented by the different organisations, the group agreed that it would be helpful to create some sort of 'umbrella' body, dedicated not just to bioinformatics, but also encompassing biocuration, biocomputing and computational biology (B^3CB) learning, education and training. The idea, then, was to create a global organisation, to share not duplicate effort, to share not duplicate cost, to work together towards common solutions and a sustainable future. The particular emphasis on sustainability was important, as the group recognised that many of their training activities are *ad hoc,* not specifically funded, and are run by volunteers – so the real issue was how to make the work of all of these organisations somehow more sustainable.

After the meeting, a memorandum of understanding was drawn up, which came into force the following month, as soon as the first five of the original ten

organisations had signed. Four months later, GOBLET was registered as a legal foundation, based in the Netherlands. Eighteen months on, more than 30 international and national networks and societies, national institutes and organisations, research and interest groups, and individuals (including students) have joined the GOBLET Foundation – see Figure 1.

Figure 1: *Partner organisations of GOBLET, as of July 2014.*

Nevertheless, clearly GOBLET is still a very young and growing organisation, with scope for much wider membership. Inevitably, current members are mostly rooted in bioinformatics, but the SEB is different because it provides a more direct connection with life science communities who need bioinformatics training, which is important if GOBLET's work is to remain relevant. So GOBLET also welcomes members from life science disciplines other than bioinformatics.

This fits with GOBLET's ethos, which begins with *inclusivity* – we want to reach out to all relevant biological/biomedical/life science organisations in order to help shape and share what GOBLET does. Hence, our ethos also embraces sharing, openness, innovation and, especially, tolerance – transcending national, political, cultural, social and *disciplinary* boundaries. With this in mind, GOBLET aims to provide a global, sustainable infrastructure to support the provision of bioinformatics training; to facilitate capacity development in bioinformatics across the globe; to help develop standards and guidelines for bioinformatics training; to reach out (amongst others) to teachers at high schools to try to bridge the gap to the next generation of bioinformaticians; overall, to foster the international community of B[3]CB trainers. To this end, GOBLET's work is organised around a number of themes, aiming to help improve the quality of bioinformatics training; exploring methods of training provision, course accreditation and trainer incentivisation; and, as already mentioned, looking at outreach and sustainability.

To provide a framework for GOBLET's work, the first step was to establish a website (www.mygoblet.org) and to put in place a training portal (www.mygoblet.org/training-portal), which includes a registry of courses, of training materials and resources, and of course organisers and trainers, as illustrated in Figure 2.
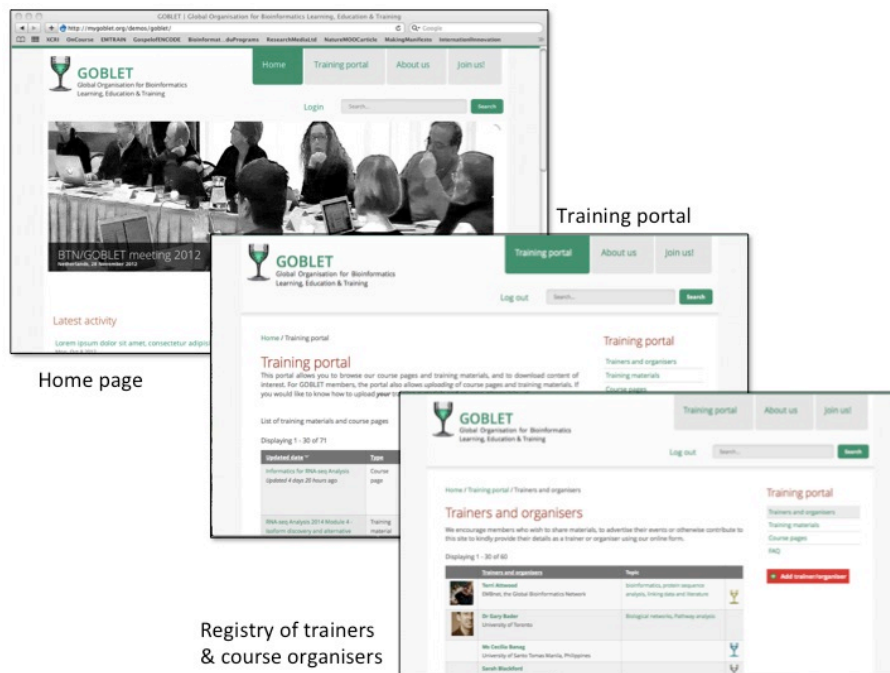
Figure 2: *Screen-grabs of the GOBLET and Training Portal.*

As already mentioned, GOBLET runs community surveys to try to ascertain training needs; the Foundation also publicises its work through conference posters and journal publications, via newsletter articles, and through the creation and dissemination of fact-sheets, as shown in Figure 3.

Figure 3: *Publicising GOBLET's work through posters, fact-sheets & newsletter articles.*

Aside from training events, holding regular meetings and workshops is necessarily part of GOBLET's work, including an Annual General Meeting (AGM),

where GOBLET's formal business (including elections, finances and strategy) is discussed. Last year, highlights of the 2013 AGM were captured in a sketch-note created by Francis Rowland – see Figure 4.

Figure 4: *Sketch-note (by Francis Rowland) highlighting the events of the 2013 AGM.*

Overall, GOBLET's work is orchestrated by an Operational Board, which is led by a small Executive Board and the Chairs of its five committees – its member societies, networks, institutes, individuals/students, *etc.* form the main assembly. This operational structure is illustrated in Figure 5.
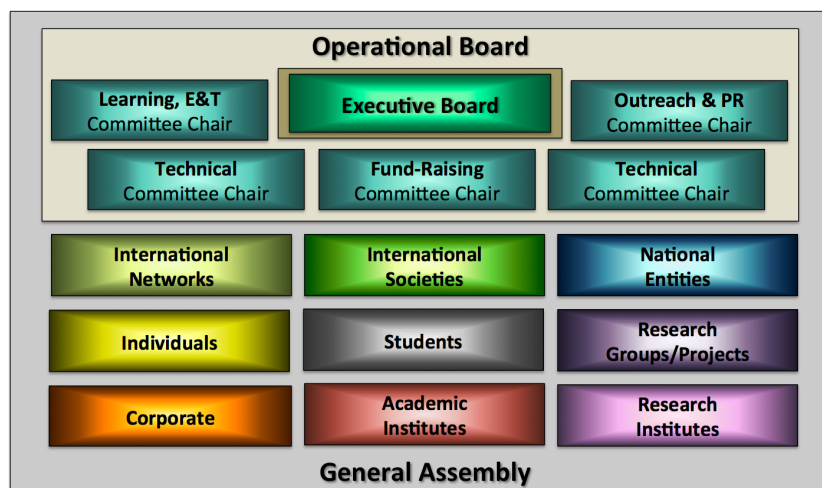

Figure 5: *GOBLET's operational structure.*

The largest of the committees is the Learning, Education and Training Committee, which aims to develop and share best practices in teaching and learning methods, to offer curriculum guidance, and to explore recognition mechanisms and provide support structures and resources for trainers

(including materials, tools, and train-the-trainer programmes). This committee works closely with the Standards Committee, which is currently exploring accreditation mechanisms, and generally aims to try to increase the quality of training materials and programmes.

These two committees, of course, work alongside the Technical Committee, together shaping and providing GOBLET's overarching support structure for trainers and trainees; the Technical Committee is also responsible for maintaining the open Training Portal for all individuals and communities interested in bioinformatics, biocuration, biocomputing and computational biology learning, education and training.

To try to sustain the work of the Foundation in general, and augment the membership subscriptions, the dedicated Fund-raising Committee is aiming to stimulate collaboration and raise funds to support specific projects. Finally, the Outreach and PR Committee is endeavouring to promote GOBLET's mission and work, and to maintain its social networking interactions.

What can GOBLET do to help?

In terms of what can be done to help communities like those represented by the SEB, GOBLET can and does attempt to coordinate world-wide training activities, for example via its portal, openly sharing information and training materials, and making distance/technology-enhanced learning resources, workshops, courses, summer schools, *etc.*, around the globe more accessible and readily discoverable.

Members of GOBLET are committed to promoting and delivering bioinformatics training world-wide, helping to up-skill communities by running courses, and creating and sharing training materials. Working through the Foundation, they can help to disseminate best practices and drive up the quality of *ad hoc* training courses by encouraging the use of standards.

GOBLET can also provide fora for discussing training issues, helping to surface and beginning to address real challenges by organising workshops, meetings, special interest groups, and so on. Ultimately, the Foundation may also be able to lobby funding bodies to try to increase the level of support for training programmes, and perhaps also pressure universities to provide more relevant training. Importantly, GOBLET is about "*empowering people, harnessing communities & networking networks*", creating synergies to help shape the future in collaboration with all kinds of organisations and individuals. Through such interactions, GOBLET can listen to, and is better able to position itself to try to address, the genuine needs of life scientists. To this end, and given the survey trends, Terri invited the workshop participants to consider what more GOBLET could do to help, by discussing the following questions:

i)   What kinds of training/resources could facilitate your work?
ii)   What would make GOBLET useful to you (courses in specific areas, tutorials, workshops, materials, e-learning…)?
iii)   Are there success stories that GOBLET could learn from (*e.g.,* online resources/other organisations that are valuable to you)?

Break-out session 2 – What can GOBLET do for you?

The participants again broke into small groups to discuss these questions, once more selecting a rapporteur to report back.

In general, the groups felt that access to a training portal, with freely-available, open-source materials and online tutorials, searchable by training level and topic, would be valuable; importantly, there was general consensus that such a portal would be much more useful if it included some sort of mechanism(s) for trainee support. The portal, they suggested, should include teaching modules that are standardised and up-to-date, some sort of mechanism for sharing scripts, and an online forum for discussion, sharing knowledge and supporting self-learning. One of the groups felt that GOBLET could usefully prepare materials for schools/high-schools, and suggested that a 'careers day' to promote better understanding of how bioinformatics plays a role in different jobs would be helpful; this group considered Coursera to be an example of a successful model from which GOBLET could learn.

Other specific recommendations were that GOBLET could provide specific workshops aimed at choosing and using the right databases, selecting and using the right statistical packages, and so on; there was also a general feeling that more university-level courses were necessary.

On a different note, one of the groups suggested that access to information about the level of computer power needed for particular applications, the amount of data storage required, and the overall cost of setting up relevant compute infrastructures – essentially, a basic introduction to what sort of computational framework is needed when starting out – would be helpful. This group felt that access to powerful computers via the Cloud would be useful, and thought that GOBLET could provide specific tutorials on free software tools, both online and in person via workshops (with technical staff on hand to help). Once again, the need for accreditation of such tutorials and courses was highlighted.

Wrap-up (Angela Davies & Terri Attwood)

After completing the second discussion session, Angela led a round-up of key points that had emerged from both the morning's break-outs. From the first session, in which bioinformatics training needs were reviewed, the feedback suggested that biologists/life scientists shouldn't be expected to be skilled bioinformaticians – on the contrary, they should be able to consult with appropriate experts, as and when they were needed: on balance, we should be reaching out far more to undergraduates, and really instilling core skills here.

It's clear that real help is needed with the analysis and interpretation of the data generated by certain technologies and some bioinformatics tools, and that standalone workshops and e-learning resources are the preferred routes for acquiring the necessary training. Ideally, universities should both provide access to specific software/databases, but then also take the responsibility to provide training in their use. Most importantly, new users should be given more ***encouragement*** and be nurtured more with some of these tools.

In terms of gaining real confidence with bioinformatics tools, and an appreciation of the level of skills acquired during training courses, short tests could be provided, perhaps using *sample data-sets to test proficiency*. It would be particularly helpful if GOBLET could provide relevant, *discipline-specific examples or case-studies* to illustrate the use of different tools, to better *spark interest and resonate* with novice users; the portal could also usefully hold, for example, *training data-sets* to allow users to gain more confidence in their use.

Some participants felt that maths/stats were more important to learn than programming (many had no stats/bioinformatics support, and so had to seek out their own solutions – often via YouTube!); therefore, some form of *peer support* across different disciplines, including *a forum* to sound-check tools and seek advice would be valuable (although it would undoubtedly require champions to drive this). A final suggestion was that the portal could provide relevant links to Coursera and/or could highlight what GOBLET offers that's different.

From the second session, where we looked more specifically at what GOBLET could do to help, it was recommended that the materials provided by the portal should be open source, at least for students, and that there should be some mechanism (a rating system, perhaps?) for *discerning the quality* of various materials and tools.

Participants wanted to see flexible modules held within a searchable portal, which would also provide the means to share scripts; they wanted to be able to find points of contact for certain specialist areas, and to see greater levels of outreach at conferences, including opportunities to learn more about the range of jobs that bioinformatics encompasses, for example during *careers days*.

The feedback also suggested that GOBLET could provide *guidance/advice on computing infrastructures* and on potential *uses of the Cloud*. Finally, there was consensus on the need for *accreditation* of training resources and courses.

Figure 6: *Participants of the SEB/GOBLET Bioinformatics Workshop, Manchester, UK*

Conclusion

This was a short, sweet, fun but informative meeting. It's gratifying that many of the issues raised are already being addressed by GOBLET and/or are already encompassed in the portal. Nevertheless, many new ideas emerged from the morning's discussions, particularly in terms of priority areas – there's clearly still a lot of work to do!

Although the number of participants was relatively small, the workshop was very useful. We're hence very grateful both to Sarah Blackford and the SEB for their help in organising and promoting the workshop, and to those participants who gave up their time during the SEB conference to share their views – some of the happy band are shown in Figure 6.

After the meeting, Terri agreed to draft a meeting report and post it on the GOBLET website; meanwhile, she encouraged participants to keep in touch by dropping us a line on info@mygoblet.org.