Dr Vicky Schneider shares her thoughts on training, the importance of a solid grounding in bioinformatics and her own experience in the field.

**What is your background as a researcher, and what led you to your current position as Head of Training and Outreach at The Genome Analysis Centre (TGAC)?**

I began a career in research because of people like David Attenborough and Jacques Cousteau; I loved animals, plants, and nature in general from a very early age. Growing up in Brazil, I used to spend hours examining tropical creatures. I therefore decided to study biological sciences at the University of Rome, and from there I went on to do a Master’s in evolutionary biology at the University of Leiden in The Netherlands. It was also there that I completed my PhD. A position within Professor Gianni Cesareni’s group at the University of Rome’s Genetics Department triggered my interest in bioinformatics and the obvious gap between those creating resources and tools, and the end-users. I always had a keen interest in community building, networking and communication, so I then decided to take a big step towards a different role within the research and academic environment: I became the Scientific Training Officer at the European Molecular Biology Laboratory’s Bioinformatics Institute (EMBL-EBI) where I spent six years launching and consolidating the training programme and from there embarked on my current job at TGAC.

**How important is it now for the average biologist to have a holistic understanding of their subject? Is this a relatively recent trend?**

A broad perspective is always important, and I don’t think it is a relatively recent trend; the challenge is presented by the sheer amount of information that is out there. Of course, whatever you do, you gain from knowing your subject from top to bottom – the question here is which aspects of your subject you have to learn about in order to pursue it effectively. Whereas before, around a decade ago, we had sequencing processes that produced a sequence of a certain length and allowed you do lots of things with it quite easily, now researchers have so many sequences that they won’t fit on your laptop, and can’t be moved around. This is before you start to approach the challenge of analysing such data (eg. assemblies).

**A lot of debate surrounds the way pre-university students are taught science. Do you feel there is an urgent need to adapt current curricula to encompass the latest developments in life sciences?**

Yes, this is a crucial issue. I collaborate with lots of people, including many lecturers and university professors, who often feel frustrated when they get young PhD students who might have some knowledge, but lack actual skills. Talking about bioinformatics, it is almost like teaching someone to ride a bicycle: I can peddle a bicycle and tell you how it works, but you can only learn to ride the bicycle by doing it for yourself. Experiments are usually conducted in such a way that bioinformatics could represent a bottleneck at the start of the project, a burden for the lab and actually a misuse of their university time. There is a lot of work to be done in terms of rethinking and prioritising the contents of what students learn, as well as introducing students to bioinformatics and coding earlier.

**Do the different interpretations of bioinformatics across different disciplines and the lack of clear distinction from computational biology ever present problems?**

It presents a problem only if you decide to get caught up in semantics. When I facilitate high-level meetings, I try to remove the assumption that we all have the same definition for these words. I have been in meetings where participants become stuck on one word, for no reason beyond their own idiosyncrasies. In bioinformatics, we tend to remove the semantics and concentrate on problems we want to solve – thereby resolving the issue of interpretation. For TGAC, bioinformatics is a component of computational biology.

**Are there any upcoming events that you would like to draw attention to?**

We have now launched our TGAC Training programme for 2014, consisting of seven core courses. These will be fully immersive and hopefully transformative for those who attend. We will have key speakers that are also part of our TGAC science symposia programme, with a breadth of different disciplines from many different areas represented. We will also contribute to the the Big Bang Science Fair as part of our outreach activity, and will be looking to host more immersive programmes and run key exploratory workshops and focus groups.
The importance of bioinformatics

A group at The Genome Analysis Centre is devoting its attention to training the current and future generations of life scientists in genomics and bioinformatics, while simultaneously teaching the public about the relevance of such subjects to their daily lives.

THE PROCESS OF DNA sequencing has come a long way since the first sequencing techniques were developed in the late 1970s. Since then, the practice has become indispensable to life sciences research and, with the help of ever-advancing technology, much easier, faster and cheaper to perform. The late 1990s brought the advent of next-generation sequencing (NGS) techniques, including pyrosequencing, colony sequencing and massively parallel signature sequencing. These methods have been responsible for an enormous increase in the quantity of genomic data produced.

But while an increased throughput of data is generally advantageous to life science researchers, it does bring its own challenges. Already suffering from a data deluge, thanks to NGS biological research can now easily generate terabytes or even petabytes of data every day. This volume of information necessitates advanced facilities for its storage alone, as well as computationally intensive approaches for its effective analysis. Added to this is the non-trivial issue of making the data accessible to other scientists for collaborative purposes. As a result, the relationship between laboratory biology and bioinformatics has become much closer and the two approaches are now highly dependent on one another.

THE PROBLEM WITH PROGRESS

The issue with this state of affairs is that it has developed so quickly. Within just a couple of decades, skills such as programming, or at least some knowledge of the command line, have become integral to the effective completion of biology projects. This has been a jump that teaching and training programmes have largely failed to cope with, and without a proper remedy, this situation will result in scientists who are unable to effectively pursue their studies. This problem is a global concern, and it is exacerbated by the fact that the necessary skills for effectively handling data are best learnt through practical experience rather than study.

In an effort to redress this issue, one of the UK’s leading research institutes has developed a programme to bring the necessary computational science skills to students and scientists at all levels. The Genome Analysis Centre (TGAC) is located at the Norwich Research Park (NRP) in Norfolk, and its relatively new training and outreach programme is designed to tackle the vital task of disseminating bioinformatics knowledge to both the scientific community and society as a whole. The programme is headed by Dr Vicky Schneider, and its ambitious brief is to enthuse the current and next generation of scientists through advanced skills-based training, make science more accessible to the general public, and enhance interdisciplinary and global collaboration.

UNFETTERED GROWTH

TGAC was established in 2009, and in less than five years has firmly cemented its reputation as an expert partner in high-throughput sequencing. It is home to a team of more than 70 genomics researchers, bioinformaticians and technologists, who between them have worked on projects ranging from the de novo sequencing of a rubber tree genome to research into biomethane production. Collaboration is considered crucial, and the institute has also made knowledge exchange a central part of its strategy. With an array of ongoing international collaborations, as well as many industry links and science seminars hosted on site, TGAC has quickly become a hub for scientific dialogue and exchange.

The researchers have been facilitated in these endeavours by the development of novel sequencing technology, primarily in the form of sequencing machines. The core strategy of TGAC focuses strongly on utilising this technology, as well as their considerable expertise, to strengthen the UK’s bioeconomy. Furthermore, through its Plant and Microbial Genomics Group, TGAC is also bringing advances in sequencing methodologies. By making the most out of the equipment it has, this institute is in some ways setting the pace for many other groups worldwide.

TRAINING AND OUTREACH

The training and outreach programme fulfils a number of important roles within TGAC in an effort to help achieve its ambitious vision. Firstly, it assists in collaboration by providing bioinformatics training to visiting scientists where necessary. This ensures that visitors who may have expertise in a certain field but lack

For more information, the Strategy can be viewed here:
www.tgac.ac.co.uk/uploads/PE%20at%20TGAC.pdf
INTELLIGENCE
THE GENOME ANALYSIS CENTRE’S TRAINING AND OUTREACH PROGRAMME

OBJECTIVES

• To educate current and future scientists about the importance of bioinformatics and genomics, and provide them with skills that they can successfully implement in practice
• To make this area of science more accessible to the general public
• To promote and establish interdisciplinary collaborations and knowledge exchange around the world

PARTNERS

Global Organisation for Bioinformatics Learning, Education and Training (GOBLET) • The Global Bioinformatics Network (EMBnet) • IRSES-DEANN Project

Professor Erik Bongcam-Rudloff, SLU-Global Bioinformatics Centre, Sweden University of Agricultural Sciences, Sweden • Dr Pedro Fernandes, The Gulbenkian Training Programme in Bioinformatics Instituto Gulbenkian de Ciência, Oeiras, Portugal • Dr Ana Conesa, Centro de Investigaciones Príncipe Felipe, Spain • Dr Patricia M Palagi, SIB Swiss Institute of Bioinformatics, Switzerland • Aleksandra Pawlik, The Software Sustainability Institute, University of Manchester, UK • David Judge, Department of Genetics, University of Cambridge, UK • Dr Jacqueline Dreyer, EMBL International Centre for Advanced Training (EICAT), EMBL Heidelberg, Germany • B F Francis Ouellelet, Ontario Institute for Cancer Research, University of Toronto, Canada • Dr Annette McGrath, CSIRO Computational Informatics, Australia • Dr Karthik Ram, University of California, Berkeley and ROPEN Science, USA • Professor Teresa Attwood, University of Manchester, UK

FUNDING

TGAC is a research institute, strategically funded by the Biotechnology and Biological Science Research Council (BBSRC)

CONTACT

Dr Vicky Schneider
Head of Training and Outreach, Senior Management Team
The Genome Analysis Centre (TGAC) Norwich Research Park, Norwich NR4 7UH, UK
T +44 1603 450895
E vicky.sg@tgac.ac.uk
www.tgac.ac.uk/training-outreach

DR VICKY SCHNEIDER leads the Training and Outreach programme at TGAC and is part of the Senior Management team. She is passionate about building communities, working and engaging with others to achieve goals, share experiences and promote best practice. She specialises in bioinformatics training, particularly with the current topics of next generation sequencing, omics and bridging the gap between biologists and computational researchers.

OBJECTIVES

A basics understanding of bioinformatics do not become a burden in research projects. The institute also participates in worldwide information networks – including the Global Organisation for Bioinformatics Learning, Education and Training (GOBLET), of which Schneider is Vice Chair.

Alongside these functions, the main responsibility of the programme is in providing a comprehensive set of education and training activities for students and scientists of all ages. These courses focus on face-to-face, immersive teaching methods, and also incorporate virtual machines and cloud computing in order to ensure that skills learned on the course can be reapplied as easily as possible. Postgraduate training opportunities are also available, and focus on the application of life sciences R&D within academia and industry.

REACHING OUT

The other half of the programme concerns public outreach. Genomics is an area of science that is increasingly relevant to everyday life – and much of the research conducted at the institute is fundamentally linked to important issues such as agriculture, environment and health. Therefore, TGAC is especially concerned with raising public awareness of these issues, as well as highlighting the role that this research can play in bringing about advanced solutions.

The training and outreach programme is regularly involved in science fairs, which allow them to make direct contact with people from every age group and demographic. TGAC4Kids activities introduce children to the inner workings of the TGAC with diagrams, and involve competitive elements to keep them engaged. In the future, the team hopes to adapt this successful initiative into an interactive website. When interacting with the public through any of its activities, the team’s philosophy is that outreach is a cyclical process – to be fruitful, scientists must interact with the public in a two-way dialogue, before reflecting on this dialogue and repeating the process. The audience of public engagement should be both scientists and public.

Across all sections of society, therefore – from the general public to established scientists – there is a need to better understand the role that the life sciences play in everyday life, and the central role bioinformatics plays in the life sciences. In addition to its work as an important facility for genomic research, TGAC is addressing this issue through its training and outreach programme. In doing so, it is not only serving future citizens, but future scientists as well.