Damian Counsell
is a bioinformatics specialist at the Medical Research Council's Human Genome Mapping Project Resource Centre, Cambridge, UK. He teaches the subject at the Genome Campus (and other EMBnet nodes) and lectures on the Master's courses in Bioinformatics and Biomolecular Sciences at Imperial College, London University.

A review of bioinformatics education in the UK

Damian Counsell
Date received (in revised form): 14th October 2002

Abstract

If the completion of the first draft of the human genome represents the coming of age of bioinformatics, then the emergence of bioinformatics as a university degree subject represents its establishment. In this paper bioinformatics as a subject for formal study is discussed, rather than as a subject for research, and a selection of the taught, mainly graduate, courses currently available in the UK are reviewed. Throughout, the author tries to draw parallels between the integration of bioinformatics into biomedical research and teaching today, and that of molecular biology, two decades ago. Others have made this analogy between these two relatively young disciplines.

Although research sources are referenced, the author makes no pretence of objectivity. This article contains his opinions, and those of a number of current bioinformatics course organisers whose comments on the subject were solicited in advance specifically for this paper. The course organisers kindly advised how they planned their curricula, and described the special strengths of their programmes. Comments from present and former students of several bioinformatics degree programmes were also solicited. Except where individuals are directly quoted, any opinions expressed herein should be considered the author's.

Compared with its sister piece [Marion Zatz, in previous issue of *Briefings in Bioinformatics* pp. 353], this paper is less about funding policy – which, in the UK, has lately (if belatedly) been more generous towards bioinformatics teaching – than it is about practice and content; the requirements of the bioinformatics research communities, the corresponding emphases of bioinformatics courses, and the general market for holders of bioinformatics degrees. Individual courses are cited throughout as examples, but the final section contains a full annotated listing with URL addresses.

Based on the author's own experience of practising and teaching bioinformatics, he describes the skills he believes will be most useful to bioinformaticians in the near future and suggests ways to prepare students of bioinformatics for a fall in demand for those abilities.

Keywords: *bioinformatics, education, UK, Master's degrees, Bachelor's degrees, curriculum*

INTRODUCTION

What is bioinformatics?

There is still disagreement over the definition of bioinformatics. At its broadest, the term includes any application of computational methods to biological problems, from statistical genetics to the design of man–machine interfaces. In common usage most people assume 'bioinformatics' to mean 'computational molecular biology'. For the majority of syllabuses in European and US higher education, this narrower definition is preferred. Despite this consensus, some of the most frequently asked questions (FAQ) that I deal with concern the scope of the subject.

I am also often asked: 'Where can I study bioinformatics?' and 'How can I study bioinformatics?' Often correspondents ask more specifically how best to move from the computational to biological sciences or vice versa. My answers to Bioinformatics.Org's visitors are explored in this review.[1] This article should be read in conjunction with the FAQ page on Bioinformatics.Org which is updated frequently.

BIOINFORMATICS EDUCATION TODAY

Currently bioinformatics education has several characteristic features.

Damian Counsell,
MRC Human Genome Mapping Project RC,
Hinxton,
Cambridge CB10 1SB, UK

Tel: +44 (0) 1223 494585
Fax: +44 (0) 1223 494512
E-mail: d.counsell@hgmp.mrc.ac.uk

Graduate level

Bioinformatics is an interdisciplinary field; it is a relatively new area of study; and it is an intense focus of research. For all of these reasons most educational opportunities in the area are for graduates.

There is no standard text, there are few template lecture programmes, and there is no single standard software suite for practical work. Under these circumstances, at least some 'open-ended' guidance is inevitable. This kind of teaching is better suited to graduates, who should be equipped to work independently. By far the majority of the course details submitted to me, from both inside and outside the UK, describe graduate programmes.

Professional development

The need for bioinformatics insights into biomedical research is so acute that few of the 'customers' of bioinformatics programmes feel they have the time to follow a full degree and acquire a broad education in the field. Postdoctoral scientists want to update their computational skills because they need to apply new techniques to their own work (or to improve their attractiveness to employers) immediately. Usually they do so by studying bioinformatics as part of their professional development by taking shorter, part-time or distance-learning courses, such as those offered by the Medical Research Council's Human Genome Mapping Project Resource Centre (HGMP-RC).

Cross-training

It is easier to teach interdisciplinary material to those with a grounding in one of the constituent fields. Bioinformatics students tend to be from the biological rather than the more quantitative sciences, even on the longer, full-time programmes reviewed here. Some courses, such as the Master's degrees in bioinformatics at the universities of Liverpool and Manchester and at Imperial College, London, admit graduates from both computational and biological 'cultures'.

Applications vs. implementations vs. foundations

A less obvious characteristic of bioinformatics teaching today is the separation of the broad categories of bioinformatics activity. The computational component of most courses tends to emphasise training in one or two of the following:

- tool use;
- program development;
- algorithm design and theoretical foundations.

Most short bioinformatics courses for biologists teach the application rather than the *writing* of tools. Most people writing bioinformatics software do not also *apply* a range of existing bioinformatics software programs on a daily basis. In 2000 Gene Myers, then Vice President of Informatics Research of Celera Genomics, described bioinformatics tool users and bioinformatics programme developers respectively as 'miners' and 'engineers'.[2]

New teaching methods

Most modern biology is not theoretical. Most important conceptual developments in biology have come through technical developments, for example: cell culture, X-ray crystallography, cell fusion, monoclonal antibodies, electron and digital microscopy, and the polymerase chain reaction (PCR). Bioinformatics techniques have also helped and will continue to help us to answer fundamental questions of biology and medicine – if only by allowing us to visualise and mine vast quantities of biological data to see new patterns.

Computers are also changing patterns of teaching. Several bioinformatics courses have adopted distance learning, have open-sourced teaching materials (as many bioinformatics developers open source their code), have embraced modern modular teaching formats, and have pooled precious teaching resources from multiple institutions. Examples of these approaches are included in the survey section at the end of this paper.

Constraints on students and teachers in the UK

Students

Four years ago in the journal *Bioinformatics*, Russ Altman wrote an opinion piece[3] on what he felt should be the content of a bioinformatics Master's degree. It struck me at the time that the ambitious number and range of subjects listed in that article were more appropriate to scholars of the calibre of the author himself and the longer duration of the typical US Master's programme than the capabilities of the typical bright graduate student doing an intensive, one-year UK course.

US and UK Master's courses

My personal view is that we should aim to teach fewer, simpler, ideas with broader applicability; if possible we should teach them to a higher standard than would be possible with a wider syllabus. It has been interesting to compare the priorities of course organisers from various institutions as they build their own bioinformatics syllabuses for the more restricted time scales of most UK Master's courses. (UK undergraduate degrees tend to be narrower in content and higher in level than those in the USA.)

Teaching bioinformatics is hard

When time is tight, effective teaching places realism over idealism. I prefer, for example, that students should have a grasp of evolutionary principles for the correct interpretation of sequence and structure analysis programs than 'stamp-collect' a vast range of different techniques for the creation of phylogenetic trees. Similarly, I feel that they should learn how statistical thinking should be used to design and interpret experiments, rather than collect recipes for statistical tests likely to be misapplied later in the laboratory.

Teachers

For both computational and biological students coming to a taught graduate programme in bioinformatics there are many difficult concepts to be acquired from 'the other side'. For most public institutes of higher education, there are precious few teachers with the time, interest and gifts to explain these ideas to members of another intellectual community, and no armies of teaching assistants to share the load. Unfortunately, even when teachers are not *un*interested, they are usually *dis*interested. That is, there is little professional incentive to be a good teacher in UK higher education.

Added to the usual problems of designing a syllabus for a new subject is that of coordinating the efforts of biological and computational departments in that process. Sometimes bioinformatics teaching is assigned to one department (Sheffield University), sometimes it migrates from one to another (Birmingham University), sometimes it is shared (Manchester University).

Within Europe, academics are usually paid less than similarly qualified individuals in the private sector. Additional rewards come for good research, rarely for good teaching. Bioinformatics may continue to be mainly a graduate subject, but teaching it well is hard work. A common complaint of overburdened lecturers when asked to deliver labour-intensive tuition is that students are being 'spoon fed'. The depressing truth is that bioinformatics is difficult and close supervision of students is often necessary.

As joint head of the EBI, and dealing daily with the problems caused by the shortage of trained bioinformaticians, Graham Cameron argued that bioinformatics should be integrated into undergraduate biomedical sciences syllabuses. This would force course organisers to develop formal and accessible teaching programmes and lay the foundations for those teaching at higher levels so that students no longer come to the subject cold. Another advantage of 'assimilation' would be that bioinformatics

'Geeky image'

would be taken for granted to the extent that it would, to use Cameron's words, 'transcend its geeky image'.[4] Although there are only a handful of pure undergraduate courses in bioinformatics in the UK, indeed in the world, there are already a similar number of more conventional courses offering a substantial bioinformatics component: for example Biochemistry with Bioinformatics at Brunel University, Bioinformatics with Computer Science at Royal Holloway College, London University.

Physicists in biology

Teaching quantitative methods to students of the biological sciences is *not* a new problem. For years biologists and medical students have grudgingly sat through statistics lectures. Reluctance on the typical student's part has often been matched by that on the part of the typical tutor, when this is an equally uninterested mathematician or statistician from a discipline taken with elegant proofs, not messy data.

Surveying the literature written and reviewed by graduates of such teaching emphasises the need for us to tackle these general problems. I was recently asked by a group of biologists to perform a completely inappropriate statistical test on data from one of their papers. They had, correctly, not performed the test at submission because the properties of their data did not meet the necessary assumptions, but could not publish the work until this 'analysis' had been wrongly bolted on to their work.

Balancing biology and computing

We must teach quantitative biological from a biological point of view if we are to engage the large minority of students who choose biology because they believe it is a 'science without maths'. Sometimes they become senior scientists. Perhaps one of the best outcomes of the integration of bioinformatics into biological teaching would be to ensure that not only can you not get *by* in biology without mathematical thinking, but that you cannot get *on* without it.

Similarly, it is the job of schools to equip students with the mathematical skills they need to progress, as Mark C. Field, organiser of the Imperial course, put it 'the poor state of mathematics at school level [could] easily exclude people from informatics who may be potentially gifted'.

Teaching biology to quantitative scientists, by contrast, *is* a new problem. At this year's successful O'Reilly Bioinformatics Conference, one of the most popular and well-attended sessions was Nathan Torkington's crash course in (molecular) biology for non-biologists. It is clear that there is a demand for this kind of cross-training amongst computational specialists.

Physicists often have a strong curiosity about biology, a curiosity that played a critical part in the emergence of molecular biology.[5] When I explain biology to students from such backgrounds, they frequently become irritated with the swarms of facts, the hordes of acronyms and the forests of exceptions to biological 'rules' — exceptions that they must learn in order to master even the basics at degree level. Educators must be sensitive to this when they present biology. Once trained, such individuals may later reward us for our efforts with the kinds of insights that only outsiders can offer.

In the private sector it has been argued that it is easier to teach biology to programmers than programming to biologists.[6] While this is often true for bioinformatics software development and implementation work, one problem in the academia is the misapplication of quite brilliant quantitative minds to faulty formulations of biological problems. Clear concepts and sensible experimental goals are as important as rigorous maths. As Andrew Dalby told me when describing the MSc/MRes Bioinformatics he is involved in planning at Exeter University, 'the most important aspect of [this] course is getting the right balance between the biology and computing'

The advantage of full undergraduate programmes

Currently in the UK prospective undergraduate students can choose from only five distinct 'pure' bioinformatics courses. If bioinformatics becomes a more common single honours undergraduate degree there will be one immediate gain for potential bioinformaticians: time.

Time to teach programming

For example, learning to program 'properly' – that is, to acquire both an understanding of programming technique and of the *discipline* of programming – is a slow process. Often, over a one- or two-year course it is difficult to absorb the 'cultural' aspects of good program design.

Whether there is enough time in what is virtually a dual-honours course to cover in depth both computer practice (programming) and theory – for example, the classic algorithms: sorts, compression; and data structures: lists, trees – of a full computer science course is another question. E. James Milner-White feels that 'serious programming' is 'essential' to the Master's degree in bioinformatics at Glasgow University, with Java and Perl being the languages of choice. A special effort is made there to identify students who find this difficult.

Technicians and theoreticians

As bioinformatics matures, the hastily assembled 'cats' cradles' of code used in the early days of the field – before it even had a name – are giving way to documented, engineered and object-oriented software *systems*. The difference is between single coders fixing a local, short-term problem – for example crafting a quick script to calculate molecular weights or isoelectric points – to transnational teams working on 'big' biology – for example, assembling or annotating whole genomes.

The former, more amateur (a word echoed by Milner-White), bioinformatics must give way the more recent, project-based, approach and we must educate a generation of bioinformaticians to 'build' rather than 'tinker'. This professionalisation could extend to a minimum standard technical qualification or a shared common undergraduate core curriculum across Europe. If institutes in the region combine this with the planned standardisation of European higher qualifications,[7] the result could be a workforce of mobile and competent bioinformaticians to speed the advance of life science research in the region.

Time to raise the status of the discipline

While there are so few bioinformatics undergraduate programmes and so many courses in bioinformatics for professional development the subject will continue to be implicitly downgraded to a collection of 'skills' rather than a field of scholarship. 'Doing bioinformatics' will be perceived as a mechanical activity. Falsely believing bioinformatics analysis to be a fully automated step, and unworthy of close attention, biologists may be tempted to do bad bioinformatics, just as many of those who attended tacked-on quantitative courses as undergraduates produce bad statistics as practising biologists. Because of the snobbery that – in the UK especially – places technicians below theoreticians in the pecking order, the subject might not be respected as it should be by those who wish to do serious biology. The irony is that a lot of published, peer-reviewed biology is crying out for quantitative rigour if it is to become serious science.

Integrated syllabuses

Themes and principles

Teaching underlying principles

A well-designed programme of study should be more than just a list of subject areas to be covered. Ideally its various elements should be linked by common themes – natural selection, for example. The Cranfield University MSc in Bioinformatics,[8] in fact, promotes its 'totally integrated approach', emphasising common themes throughout the syllabus, as the most important reason for prospective students to choose its MSc programme over others.

There are two advantages to thematic teaching of a conventional syllabus: firstly, facts in context are much easier for students to understand and retain than facts in isolation; secondly, the presence of an over-arching theoretical framework gives a field coherence both to its practitioners and to the outside world.

The central dogma

In molecular biology, for example, the central dogma of information flow in living things has been illuminating both in the classroom, and in the laboratory, at the least by making the exceptions to rules more obvious: for example, reverse transcription (though this only contravenes the dogma as we commonly know it, not as it was originally framed by Crick).[9]

Teaching at all levels

Full-time courses give students enough time to learn about both tool use and tool development. Those from a programming background can learn much from trying to use other developers' code productively. To biologists, who may never write a full application from scratch – and should not need to because open source programs for most common applications should already be available – perhaps the greatest benefit of learning programming comes from their being able to write scripts to call existing code and automate tedious tasks.

ECONOMICS

Growth in education spending

Funding is growing

All the major funding agencies in the UK have programmes to support increased bioinformatics education.[10] The following summaries are based on the respective organisations' own public policy documents.

The Biotechnology and Biological Sciences Research Council (BBSRC)

Bioinformatics has moved from receiving specific targeted BBSRC funding to being a prominent concern all of its separate subject groups; it has become a so-called 'cross-committee priority'. Every BBSRC research committee will have at least one member with bioinformatics expertise and a bioinformatics network group will be established to for peer review and to influence policy.

This change shows that bioinformatics has entered the mainstream to such an extent that many funded projects depend on bioinformatics techniques and expertise without mentioning the fact. Although this makes it difficult to separate those grants with specific bioinformatics orientation from other, more 'conventional' work – just as it would be almost impossible to dissect molecular biology out from most contemporary biomedical research projects – the nature of taught courses makes the emphasis explicit. Of the 111 Master's places funded by the BBSRC, between 1999 and 2002, 25 were in bioinformatics.

The Engineering and Physical Sciences Research Council (EPSRC)

The EPSRC is involved with other funding agencies in the provision of bioinformatics training, but itself also funds six Master's packages within the UK in bioinformatics or closely related disciplines.

The Medical Research Council (MRC)

The MRC is also convinced that there is a high demand for trained bioinformaticians and has a clearly stated policy to address this: 'The MRC aims to award at least 20 PhD/DPhil research studentships and Masters studentships in different areas of bioinformatics and neuroinformatics every year, and funded 22 of each in 2000/2001.'

As I mentioned earlier, the Medical Research Council also funds the many teaching and outreach activities of the UK HGMP-RC.

The Wellcome Trust

In total the Wellcome Trust spends around £600m per annum on all of its activities. It contributes hugely to research involving bioinformatics in the UK; this is most conspicuous in the form of the Wellcome Trust Sanger Institute, where about a third of the human genome was sequenced. The Wellcome Trust also supports bioinformatics-related short training courses and postgraduate education of scientists and clinicians for professional development, but, since there is no direct funding of bioinformatics Master's programmes by the Trust, it is not possible to give a fixed annual figure for its support of places on taught courses.

Growth in biomedical research spending

This is not an article about research spending, but it is important to note that all UK research councils have made a priority of bioinformatics. Further, in July 2002, the UK government's spending review promised a 10 per cent per annum rise in funding for science.[11]

European funding policy

The European Commission stated in its life sciences strategy document[12] that 'Comprehensive, up-to-date and publicly and freely available bioinformatics data is the basis for advances in biotechnology' and recommended that:

> The Commission . . . and its Member States should also in collaboration with the European Investment Fund develop a competitive bioinformatics infrastructure in support of biotechnology research and focus support for the development in computational biology and bio-medical informatics

over the course of the duration of the sixth framework project (2002–2006). The European Bioinformatics Institute (EBI), an outstation of European Molecular Biology Laboratory (EMBL, itself based in Heidelberg), is located at the Genome Campus near Cambridge in the UK and receives funding from the EMBL member states, the European Commission, and the Wellcome Trust. The annual budget of the EBI is £12–13m.

Cheaper infrastructure than molecular biology

Equipping and running a bioinformatics teaching laboratory costs much less than equipping and running a molecular biology laboratory. Most bioinformatics software is available for free to academics, as are many Internet-based bioinformatics services. The data from the public genome projects is also free to anyone with Internet access. x86-based commodity PCs of the Pentium generation and later are more than capable of tackling the most common bioinformatics tasks in reasonable time scales.

OVERSHOOT

With the costs of bioinformatics lower relative to other biomedical laboratory sciences and funding for research rising, why is it that I believe that there will soon be an end to the shortage of bioinformatics graduates that has marked the past few years?

A personal view

My own story

I graduated with a respectable biological sciences degree from an 'elite' UK university at the beginning of the 1990s, but found it difficult to obtain funding to study a computational subject at graduate level. Instead I worked as a technician and research assistant in good biomedical science research institutes.

A lack of foresight

Even in such well-funded, prominent laboratories it was difficult to persuade senior staff (people leading research council strategy) that e-mail was a preferable alternative to fax for co-writing scientific papers with trans-Atlantic colleagues, let alone that there might be a growing need for computationally trained biologists. The suggestion that the first rough draft of the human genome would be available within a decade was considered absurd.

I eventually acquired formal computational education thanks to EU funding. Now employers pay a premium because of the scarcity of these skills. This expense is a direct result of a predictable (and predicted) shortage of computationally literate biologists. The funding bodies in the UK are now funding computational biology training more eagerly, but, ironically, their earlier error may be reversed: my opinion is that we are in danger of producing more bioinformatics scientists than there are jobs for them in the field. I also fear that if we do not plan carefully we may train them in ways that are inappropriate for the jobs that do exist. The problem may extend far beyond bioinformatics to science and technology training in general,[13] indeed to the expansion of higher education itself.

The next recession

After some false starts, a long-feared economic downturn seems to be beginning.[14] Research, development and information technology are usually part of the second rank of jobs to suffer in a recession (after the likes of advertising). Pharmaceutical firms are among the biggest purchasers of bioformatics skills and their financial health has recently begun to look poorer both in the short and medium term.[15] IT and telecoms have, through larger salaries, previously attracted bioinformatics graduates away from work in the area of their degree discipline. The recent gigantic collapses in the values of public corporations in both markets bode ill for recruitment. Those who would previously have sought employment in the private sector are beginning to compete for positions with not-for-profit organisations.

Miners, engineers and planners

Increasing supply and decreasing demand

Along with the decrease in demand there is now an increase in supply. Bioinformatics has become fashionable with prospective students. Certainly in the UK, if anecdotal evidence is anything to go by, there are more than enough graduates from bioinformatics courses to meet demand from the public and private sectors at this level. (Senior practitioners are still rare.) In the space of a week this laboratory received one offer of free labour from a graduate seeking bioinformatics experience, and a request from a popular science magazine for a bioinformatician interviewee to give advice to young people on how they can break into the field. Bioinformatics has been named as one of the hottest new careers[16] by more than one publication.[17] Speculative applications for jobs are a common constituent of unsolicited email received in bioinformatics research groups.

The de-skilling of certain bioinformatics tasks

At the same time, the tools for general biologists have become more user-friendly (which carries both advantages and drawbacks for the advancement of science) and more and more practising bench scientists are receiving career development training in bioinformatics. Organisations such as EMBnet (see below) are instrumental in this process. Biologists are beginning to use bioinformatics routinely in the laboratory, just as in the 1980s the once-specialised techniques of molecular biology were absorbed into the mainstream of biomedical research and difficult procedures were standardised into kit form by private sector biotechnology firms.

Applications, implementations and foundations revisited

Eugene Myers' division of bioinformaticians into 'miners' and 'engineers' (reported in May 2000)[2] accompanied his view that demand for the miners was much higher than demand for the engineers. In yet another (more recent) editorial on bioinformatics education, William Pearson,[18] author of the widely used FASTA sequence comparison program, argued instead that we should make our priority the training of bioinformatics researchers. Perhaps we can extend Myers' analogy and label these individuals 'planners', by comparison with his miners and engineers:

> the major goal for computational biology and bioinformatics training should be to produce PhD. and post-doctoral fellows who can build independent research careers in Life Sciences Departments. The emphasis should be on research training, rather

> than mastering programming languages, analysis tools, and database administration.

My position is intermediate between these two extremes. Research training seems to me important. However, there is a far smaller demand (in numbers) for individuals to lead independent research into fundamental problems in bioinformatics than there is for those who can do good bioinformatics within a biological research setting. Doing such work is not 'just' a technical activity.

Another glut?

One of the reasons there are so many physics PhDs currently working in bioinformatics is overproduction of pure researchers within that field. This overproduction was motivated by the same kinds of good intentions expressed to policy makers by talented physicists within the physics community. Professors often wish to create students in their own image. We would be wrong to make the same mistake in biomedical research and to encourage bright young people into making the sacrifices necessary to complete a PhD in a supposedly high-demand subject, only for them to find that their specialised training is just not very useful.

I agree bioinformatics should indeed be a subject of basic research, but far higher numbers of bioinformaticians are needed to apply and develop tools than to do basic science in the area, just as few planners, but many miners and engineers, are needed to produce gold. I believe that, by training all bioinformatics graduate students in a pragmatic, but thorough, way we will make better biological scientists of those who go into pure research. The Wellcome Sanger Institute is named after one of the few people to collect two Nobel prizes. The work honoured derived directly from Fred Sanger's own *technical* innovations.[19,20]

Teaching bioinformatics for employability

Transferable skills

One course that advertises training for the world of work as a priority is the MRes in bioinformatics at York University[21] which lists as its three goals, the provision of 'scientific understanding . . . research training . . . [and] a wide range of transferable skills'.

Whatever happens, there will be bioinformatics students who do not go on to do bioinformatics for a living. Bioinformatics courses must be sensitive to the outside world. This relevance should be present in both the coverage of subject matter, and in the range of skills educators provide. From an academic perspective, students need to prepare for a potentially harsh job market by acquiring, at the very least, the following:

- programming discipline;
- a rigorous approach to problem-solving; and
- intellectual 'good taste'.

At York, there is a further emphasis on the learning of languages, management abilities and presentation skills. Equipped with these attributes a bioinformatics graduate would be an asset to any organisation.

Programming discipline

A lot of the earliest examples of 'bioinformatics' were simply cases of computer-literate biologists throwing together a script in a user-friendly higher-level language: BASIC, Hypercard, Perl(-CGI). Just as some of the more arcane and specialised techniques of early molecular biology have become formalised and generalised so bioinformaticians must 'grow up' to share and explain their tools for other biologists. To do this they must write readable, reusable code. A formal bioinformatics course can cultivate these good habits, and this discipline is applicable wherever computers are used.

A rigorous approach to problem solving

One of the attractive aspects of the influx of 'hard' scientists to biology (something that also happened during the early stages of molecular biology) is the arrival of aggressively reductionist approaches. When a biological problem has been well framed, physicists and engineers, for example, often have a talent for 'seeing the wood for the trees'. It is vital in experimental design and analysis to tease out the relevant aspects of a problem and focus on them exclusively. We should cultivate these skills in bioinformatics students. They are widely applicable in the world of work.

Intellectual good taste

Many people suspect that the increase in the number of journals has not been accompanied by an increase in the quality of the science published in them. True or not, honours students should be intelligently sceptical of all publications, even (especially?) the most prominent – where rigour might be sacrificed to newsworthiness.

Learning to think

Scholars learn to discriminate by becoming familiar with their field and becoming familiar with common general fallacies and unsound experimental practices. Both authors and readers can suffer from a lack of critical thinking. Faulty reasoning is even more widespread outside the peer-reviewed literature and can be usefully studied by anyone.

OVERVIEW

The UK has been a leader in the provision of university-level bioinformatics teaching. Below is an annotated list of most full-time, extended bioinformatics courses within the UK, preceded by a description of EMBnet, an important force in bioinformatics education in Europe since the mid-1980s, long before the provision of taught courses in the member states. Table 1 lists the home page and course URLs for bioinformatics courses within the UK.

EMBnet

The grandest example of the 'collected specialists' approach to bioinformatics education in Europe is that of EMBnet.[22] EMBnet is a federation of bioinformatics centres, *nodes*, across Europe and the world. Specialists at each node offer their particular skills to other members of the consortium.

EMBnet is also involved in the collaborative provision of computing resources, technical support and publications (including this one). Importantly, EMBnet is also involved in the development of open source bioinformatics software such as the EMBOSS[23] system.

The official UK EMBnet node is the Medical Research Council's Human Genome Mapping Project Resource Centre,[24] (MRC HGMP-RC) based at the Genome Campus at Hinxton, near Cambridge. There are also two special nodes at the same site: the European Bioinformatics Institute[25] and the Wellcome Trust Sanger Institute[26] (WTSI). UMBER (University of Manchester Bioinformatics Education and Research) is the only UK academic specialist node.

University of Aberdeen

At Aberdeen University the MSc/PgDip Information Technology offers bioinformatics as a specialism – this kind of arrangement may well be the most common route by which taught research courses introduce computational scientists to bioinformatics.

IT in bioscience

University of Abertay, Dundee

The University of Abertay, Dundee, has an MSc/PG Dip in Bioinformatics. The programme defines bioinformatics to be 'the use of IT in the biosciences' and emphasises programming, the use of database technology, and statistical approaches to experimental design and interpretation.

An illustrious pioneer

Birkbeck, University of London

Birkbeck College combined its illustrious history in crystallography, connections with first-class research centres in London, and its tradition of part-time and mature study to create innovative bioinformatics MSc and MRes curricula. Its teaching materials include significant online content. Together with a consortium of research centres in the UK capital it also trains scientists of all levels over short courses. It is disappointing, however, that the material from these courses is not made freely available according to the ideals of open access in academia.

Embracing new bioinformatics

Although a relatively 'old' bioinformatics centre and one generally thought of as arising out of an intellectual heritage of structural biology, Lorenz Wernisch, one of the courses' organisers, is keen to embrace 'new' kinds of bioinformatics. He sees them, not only as embodying technical advances but as changing the nature of the discipline:

> 'Bioinformatics is about to enter a new phase. Traditionally it has been seen as providing services mainly in the form of sequence and structure databases and developing computational methods for comparison and classification. . . . However, with new high-throughput experimental techniques such as microarrays developing very rapidly, bioinformatics will soon be required to interpret and integrate large amounts of data.'

Unlike some other bioinformatics centres Birkbeck has the luxury of being able to teach all of its modules within the school and tailor the teaching around specific bioinformatics problems – another way of maintaining threads or themes over the span of a course.

University of Birmingham

An undergraduate course

Birmingham University's degree in Bioinformatics is one of that handful taught at undergraduate level. Interestingly, this BSc has migrated from its School of Computer Science to the School of Biosciences. Complementing this change, the course answers the challenge of providing a serious grounding in informatics by offering an optional intercalated year within the School of Computer Science.

Brunel University

Brunel in west London also offers a bioinformatics course, but as part of a broader curriculum in its Biochemistry and Bioinformatics BSc. Similarly, the university also offers a postgraduate Medical Genetics and Bioinformatics MSc.

University of Cambridge

Often with interdisciplinary, technical or vocational subjects, the supposed 'top three' science universities in the UK (Oxford, Cambridge and Imperial) have been later than other academic institutions to name or offer full-time, dedicated courses in the area. This has proved true for bioinformatics.

Cambridge's Natural Sciences Tripos system offers extraordinary flexibility to undergraduates in the choice of classes they follow at that level. Although there are recommended patterns of study in named subjects for final year students, these subject headings are broader than might be found at most equivalent institutions, blurring the line between academic programmes.

The university does not offer bioinformatics as a distinct honours subject, but it has for some time taken advantage of the proximity of the Genome Campus in Cambridge (home of the Wellcome Trust Sanger Institute, the EBI and the MRC HGMP-RC) to offer thorough practical training in bioinformatics techniques at the undergraduate and postgraduate levels.

Cranfield University

Cranfield, by contrast with the ancient universities, has specialised in technical and applied research from its founding – it is descended from the Cranfield Institute of Technology, which in turn descended from the College of Aeronautics. Ironically, Cranfield promotes its proximity to the bioinformatics work conducted in all three of Cambridge, Oxford and London as an attraction of its Master's in Bioinformatics at the newly founded

Cranfield Centre for Bioinformatics and IT.

University of Dundee

The University of Dundee offers Biochemistry with Bioinformatics as an undergraduate degree.

University of East Anglia

The University of East Anglia advertises an MSc in Bioinformatics. The course is a collaboration between the Schools of Information Systems and Biological Sciences and takes advantage of the expertise of those working at the neighbouring Norwich Research Park, home of the Institute of Food Research and the John Innes Centre.

University of Edinburgh

Neighbours also offer help to Edinburgh's MSc/Diploma in Quantitative Genetics and Genome Analysis. In this case the contributions come from the Roslin Institute, the Scottish Agricultural College, the Medical Genetics Section of the Department of Medical Sciences, and the MRC Human Genetics Unit. Despite its name, the course is the only one promoted under the university's banner of 'Bioinformatics in Edinburgh'. The prospectus for this particular programme explicitly encourages applicants from both the biological and mathematical and statistical sciences.

University of Exeter

Exeter offers graduate training in bioinformatics under a number of different qualification headings (MSc, MRes, PgDip and PgCert) and is one of several institutions offering such training in a modular distance learning format, especially for international students and those in industry wishing to retrain. Andrew Dalby, one of the organisers, points out that this has been the only way they can cope with the high demand for the course, which is also notable for featuring study of ethical issues prominently in its course outline. Industry links are further emphasised through the course's six-month research projects. Dalby believes that the best way to prepare bioinformaticians is by 'creating expert users with a sound theoretical background and then allowing students to specialise and develop their skills in a specific area'.

Bioinformatics with specialisms

University of Glasgow

James Milner-White, the organiser of the Glasgow MRes Bioinformatics course, feels it is essential for bioinformatics students to acquire 'a solid grounding in programming', both to increase their 'employability and confidence in tackling research projects'. This is helped by the course being jointly run with the larger MSc in IT course at the same institution.

The focus of the course 'is on predicting the protein-coding and other informational potential of genomes, deriving evolutionary models from this, deciphering the function of the proteins encoded by the genomes, analysing and modelling protein structures, and using the information in the diagnosis and cure of disease'. The course is best suited to 'numerate students with a degree in a molecular bioscience who wish to specialise in the computing aspects of biology'.

Imperial College

Imperial College opened its MSc in Computational Genetics and Bioinformatics only last year. Imperial prides itself on the quality of its faculty members, not just in the biological and computational sciences, but also within mathematics. The massive size of the college means that there is a wide range of labs in which students can apply their bioinformatics skills.

University of Leeds

Leeds University has been active in both research and training bioinformatics. Its bioinformatics groups offer shorter courses to both staff and students in addition to a full-time, one-year MRes in Bioinformatics. Presumably to allow students to cover at least one of the many aspects of the wide-ranging discipline in detail, Leeds offers an options system with specialisms in chemical structure and docking, genomics and post-genomics, mathematical biology, biomolecular networks, human linkage analysis, biodiversity and epidemiology. It also offers the possibility of funding for overseas placements. Like the Liverpool course it admits computer scientists and biologists and claims to tailor the content to their specific needs.

Problem solving

Liverpool John Moores University

Subject to approval, Liverpool John Moores University will be offering both three- and four-year undergraduate courses in bioinformatics from 2003.

University of Liverpool

Teaching two cultures

The Liverpool MSc, Postgraduate Diploma and Postgraduate Certificate in Biosystems & Informatics is from that minority of programmes inviting applications from both of the ‘two cultures’ of bioinformatics, that is those with degrees in the life or biomedical sciences *and* those with degrees in computer science, mathematics or engineering. Biologists are preferred to be computer literate and numerate (specifically A-level mathematics is recommended). In keeping with this diversity students admitted to the course are taught, for example, both the fundamentals of molecular biology and Java programming.

University of Manchester

Another pioneer of bioinformatics education is Manchester University. Modules from the Manchester MSc in Bioinformatics are also available in a distance learning format. In fact, bioinformatics education itself has become part of bioinformatics research at Manchester, exemplified by the EMBER Project.

Bioinformatics for mathematics

Terri Attwood, one of the main organisers and co-author of a popular bioinformatics text,[27] feels that, despite the increased funding available for training (and the success of Manchester in obtaining money for their European Multimedia Educational Resource (EMBER), ‘funds are not sufficient still to meet the demands to train more bioinformaticians’. She is further disappointed that other initiatives such as the World Universities Network (WUN) ‘will put money into resources for commercial release rather than for free access’.

Attwood feels that Manchester’s course stands out for its ‘emphasis on research and problem solving, rather than a full programme of lectures’. She also points out that, in the course’s theory and algorithms module, students are exposed to different programming languages.

University of Oxford

Oxford University’s Master’s course begins this year. The ancient university presents the course in a modern, part-time, distance-learning format via its Department for Continuing Education.

Royal Holloway, University of London

Royal Holloway College in the University of London offers an MSc in Computer Science by Research in which a bioinformatics specialism is available. The College also offers undergraduate degrees that combine bioinformatics either with computer science or molecular biology.

Sheffield Hallam University

The MSc/PostGraduate Diploma in Bioinformatics at Sheffield Hallam is aimed primarily at biology graduates and, like the York course, gives consideration to the acquisition of transferable skills.

University of Sheffield

Unlike almost all the other courses outlined here, the Sheffield MSc in Bioinformatics has been offered from this year solely to graduates of mathematical disciplines. Alongside more applied and advanced mathematics the syllabus includes modules in human genetics, proteomics, and genomics.

University of Manchester Institute of Science and Technology (UMIST)

Biological and Computational Science (Bioinformatics) at UMIST can be taken either as a three-year, full-time or a four-year, full-time course including one year (the third) spent working in industry.

A year in industry

University of Warwick

Warwick University offers a BSc in Computational Biology. This is broader in scope than the majority of bioinformatics courses and there are options for students to pursue interests in neurobiology, cell signalling, biological oceanography and other subjects.

University of Westminster

Together with Harrow School of Computer Science, the University of Westminster, a new university in London, is, from this year, offering an MSc in Bioinformatics as both a full- and part-time course. Again this is aimed primarily at graduates of the biological sciences.

University of York

York University was one of the first institutions in the UK to offer any kind of computational biology course, and its

Table 1: Contact details for university bioinformatics courses

Institution	Homepage	Course(s) page
University of Aberdeen	http://www.abdn.ac.uk/	http://www.abdn.ac.uk/prospectus/pgrad/2003/computing.hti#taught
University of Abertay	http://www.abertay.ac.uk/	http://scieng.tay.ac.uk/itb.htm
Birkbeck, University of London	http://www.bbk.ac.uk/	http://www.cryst.bbk.ac.uk/teaching/teaching.htm
University of Birmingham	http://www.bham.ac.uk/	http://www.biosciences.bham.ac.uk/prospective/bioinformatics/
Brunel University	http://www.brunel.ac.uk/	http://www.brunel.ac.uk/courses/ug/B/BIOCHEM+BIOINF.shtml
University of Cambridge	http://www.cam.ac.uk/	http://www.bio.cam.ac.uk/molbio.html
Cranfield University	http://www.cranfield.ac.uk/	http://www.cranfield.ac.uk/ibst/msc_bix/
Dundee University	http://www.dundee.ac.uk/	http://www.dundee.ac.uk/prospectus/undergrad2003/courses/biochem.htm
University of East Anglia	http://www.uea.ac.uk/	http://www.sys.uea.ac.uk/admissions/pgprog/msbio.pdf
University of Edinburgh	http://www.ed.ac.uk/	http://www.sys.uea.ac.uk/admissions/pgprog/msbio.pdf
University of Exeter	http://www.exeter.ac.uk/	http://bioinformatics.ex.ac.uk/
University of Glasgow	http://www.gla.ac.uk/	http://www.gla.ac.uk/Acad/IBLS/GradSchool/mastbio.htm
Imperial College	http://www.ic.ac.uk/	http://www.bioinformatics.ic.ac.uk/bio_mscphd.html
University of Leeds	http://www.leeds.ac.uk/	http://www.bioinf.leeds.ac.uk/mres/index.html
Liverpool John Moores University	http://www.livjm.ac.uk/	TBC
University of Liverpool	http://www.liv.ac.uk/	http://www.csc.liv.ac.uk/~rcp/MSc-BioSystems/MSc-Welcome.html
University of Manchester	http://www.man.ac.uk/	http://www.bioinf.man.ac.uk/education/
University of Oxford	http://www.ox.ac.uk/	http://www.conted.ox.ac.uk/bioinformatics/
Royal Holloway College, University of London	http://www.rhul.ac.uk/	http://www.cs.rhul.ac.uk/msc/mscres/index.shtml
Sheffield Hallam University	http://www.shu.ac.uk/	http://www2.shu.ac.uk/prospectus/op_pglookup1.cfm?id_num=SET001&status=TN
University of Sheffield	http://www.shef.ac.uk/	http://www.shef.ac.uk/mmge/msc-bioinfo/
University of Manchester Institute of Science and Technology (UMIST)	http://www.umist.ac.uk/	http://www.umist.ac.uk/prospectivestudents/undergraduate/courseinfo/fulldetails/00000104.htm
University of Warwick	http://www.warwick.ac.uk/	http://www.warwick.ac.uk/undergrad/biological-sciences/C190
University of Westminster	http://www.wmin.ac.uk/	http://www.wmin.ac.uk/biosciences/oldschool/msc/Bioinformatics.htm
York University	http://www.york.ac.uk/	http://www.york.ac.uk/depts/biol/gsp/masters/bioinf/bioinf.htm

3-year-old MRes in Bioinformatics grew out of its MSc in Biological Computation. The course is aimed at graduates of biology and chemistry. Like Imperial College, York University also offers a Master's in Biomolecular Sciences alongside this course.

As an MRes course, an independent research project forms a substantial part of the programme, and the programme requires one of the research projects undertaken by each student to be an external placing in industry or academia.

References

1. URL: http://bioinformatics.org/faq/
2. Kreeger, K. Y. (2000), 'Hot careers – A look at what's in demand in life sciences', *The Scientist*, Vol. 14 (9), p. 29.
3. Altman, R. B (1998), 'A curriculum for bioinformatics: The time is ripe', *Bioinformatics*, Vol. 14, pp. 549–550.
4. Gavaghan, H. (2000), 'Training: Europe seeks solution to bioinformatics shortfall', *Nature*, Vol. 404, pp. 687–688.
5. Schrödinger, E. (1944), 'What is Life?', Cambridge University Press, Cambridge.
6. Gavaghan, H. (2001), 'Current role suggests the shape of future work opportunities', *Nature*, Vol. 409, pp. 963–964.
7. Koenig, R. (2001), 'Europe seeks to harmonize its degrees', *Science*, Vol. 293, p. 1613.
8. URL: http://www.cranfield.ac.uk/ibst/msc_bix/
9. Crick, F. H. C. (1958), 'On protein synthesis', in 'Symposium for the Society of Experimental Biology XII', Academic Press, New York, p. 153.
10. Brass, A. (2000), 'Bioinformatics education – a UK perspective', *Bioinformatics*, Vol. 16 (2), pp. 77–78.
11. *Research Fortnight* (2002), 'Science scores third spending review boost', *Research Fortnight*, Vol. 8.20, p. 4.
12. European Commission (2002), 'Life sciences and biotechnology – a strategy for Europe', Commission of the European Communities, Brussels.
13. Perutz, M. F. (1999), 'Will biomedicine outgrow support?', *Nature*, Vol. 399, pp. 299–301.
14. *The Economist* (2002), 'The unfinished recession', *The Economist*, Vol. 364, 28 September, supplement.
15. *The Economist* (2002), 'Mercky prospects', *The Economist*, Vol. 364, 11 June, pp. 59–60.
16. URL: http://pubs.acs.org/hotartcl/cenear/jan012001/7901empl.html
17. URL: http://www.newscientistjobs.com/graduate/cac/career.jsp?id=career5
18. Pearson, W. R. (2001), 'Training for bioinformatics and computational biology', *Bioinformatics*, Vol. 17, pp. 761–762.
19. 'Nobel Lectures in Chemistry, Vol. 3, 1942–1962' (1964), Nobel Foundation/Elsevier Science Publishers, Amsterdam.
20. Forsén, S. Ed. (1993), 'Nobel Lectures in Chemistry 1971–1980', World Scientific Publishing.
21. URL: http://www.york.ac.uk/depts/biol/gsp/masters/bioinf/bioinf.htm
22. URL: http://www.embnet.org/
23. URL: http://www.emboss.org/
24. URL: http://www.hgmp.mrc.ac.uk/
25. URL: http://www.ebi.ac.uk/
26. URL: http://www.sanger.ac.uk/
27. Attwood, T. and Parry-Smith, D. (1999), *Introduction to Bioinformatics*, Prentice Hall, Harlow.