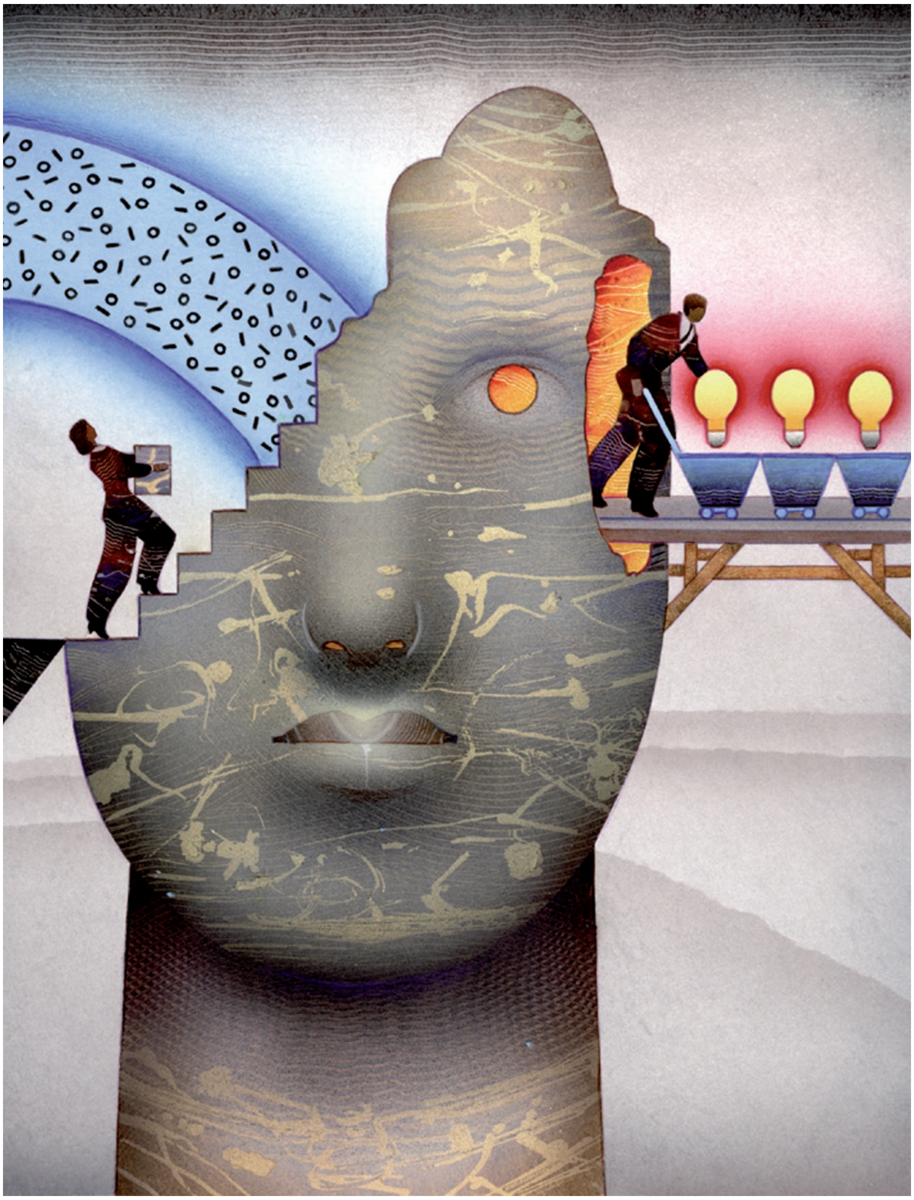
COLUMN Every scientist needs compassion and support in the lab **p.145**

CAREERS BLOG The latest discussions and news on research jobs **go.nature.com/z8g4a7**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**



J. ENDICOTT/IMAGES.COM/CORBIS

EDUCATION

Inspiration for informatics

Trainees in bioinformatics and computational biology should seek depth of knowledge over breadth.

BY VIRGINIA GEWIN

This January, Alexander Sczyrba and his colleagues published what was at the time the largest metagenome ever assembled (M. Hess *et al. Science* **331,** 463–467; 2011). Collecting and collating genetic material from environmental samples is always a challenge; in this case, the metagenome came from parts of a cow's stomach, and contained more than 27,000 biomass-degrading genes and 15 microbe genomes. It totalled 268 gigabases. "We had to develop new algorithms to run analyses on computer clusters, or clouds, as using traditional methods would have taken 80 years on a single computer," says Sczyrba.

Sczyrba wants to focus his career on similar complex, leading-edge analyses. But the path hasn't been straightforward; when he was looking for a postdoc in 2008, it was tough to find institutions that could generate or analyse such large data sets. He landed a post at the US Department of Energy Joint Genome Institute (JGI) in Walnut Creek, California: a large-scale sequencing facility that offered access to data, computing resources and brain power. In 2010 alone, the JGI sequenced 170 metagenomes.

Soon, however, big sequencing centres won't be the only sources of data. "With next-generation sequencing, everybody can produce sequences; it's the analysis that is getting more important," says Sczyrba. Modern biologists need to be able to manage large data sets and explore new computational tools.

FINDING A PATH

Qualified candidates are hard to find, say recruiters in both industry and academia. That may be because, so far, there hasn't been a typical career path for bioinformaticians or computational biologists. "Often we find that it's the people motivated to simply roll up their sleeves and figure out on their own how to work with these data that have the strongest skills," says Jim Bristow, deputy director of programmes at the JGI. As more departments are established, the often circuitous routes once required to attain such skills will probably be replaced by more direct paths. The challenge is finding a training programme that will help researchers to keep pace in a rapidly changing, technology-driven field.

By conventional definitions, bioinformaticians develop new ways to acquire, organize and analyse biological data, whereas computational biologists develop mathematical models or simulation techniques to work out the ▶

data's biological significance. But these lines are blurring, and departments and training programmes are both proliferating and combining the fields.

"The demand for computational-biology training that we have today is way more than was expected a decade ago," says Burkhard Rost, president of the International Society for Computational Biology, which is based in La Jolla, California.

NOT JUST SKIN DEEP

The most obvious training route — pursuing an undergraduate degree in bioinformatics — isn't necessarily the best for a budding researcher. Some undergraduate programmes fail to provide the depth of knowledge sought by employers. "Often these trainees come with great-looking CVs, but when we press them on what they are capable of doing, they tend to be rather weak," says Nick Goldman, research and training coordinator at the European Bioinformatics Institute in Hinxton, UK. Goldman is most impressed by applicants who have actively pursued training in both informatics and the area of research in which they're interested — for example, someone with a computing degree who has done a molecular-biology project (see 'Talent checklist').

Goldman says that students should be wary of learning about only the latest software or genome-mining tool, without gaining a full understanding of the biological topics. Recruiters want savvy scientists who understand technology's ability to address questions. Steve Cleaver, head of quantitative biology at Novartis Institutes for BioMedical Research in Cambridge, Massachusetts, says that the key to a sustainable career in the field is the ability to turn a scientific question into a statistical hypothesis. "But those who can ride the tech waves are well positioned to find career success," he adds. Without a doubt, he adds, the next generation of biologists will be more conversant in bioinformatics. "It's all about cross-training — getting the appropriate training in both analytical science and biology during graduate school to make a meaningful contribution," says Cleaver.

Picking a programme with comprehensive training modules in statistics, computer science and/or biology can be an effective strategy. But Søren Brunak, director of the Center for Biological Sequence Analysis at the Technical University of Denmark in Lyngby, says that researchers should avoid training programmes that focus on just a few data types. With the expansion in high-throughput sequencing of genomes, proteins and metabolites, programmes that focus on a single area, such as genomics, don't adequately prepare students for the job market, says Brunak. "Analyses conducted now are much more reliant on combinations of data types — for example, combining molecular-level data with patient records — than they were before," he notes.

Aspiring principal investigators can go one step further to find the best graduate training for the career they want, by deciding whether to focus on developing tools, such as algorithms to analyse data, or applying those tools to turn data into knowledge. "The most important decision a trainee can make is what kind of research programme they want to build," says Robert Murphy, founding director of a computational-biology PhD programme run jointly between the University of Pittsburgh in Pennsylvania and Carnegie Mellon University, also in Pittsburgh.

"We don't know where we'll be in ten years because the technologies and ideas are moving so fast."
Alexander Sczyrba

D. E. GILBERT

The University of California, Los Angeles (UCLA), has a bioinformatics PhD programme designed to shape the tool developers. It accepts only candidates who demonstrate a core strength in an analytical field such as computer science or maths, or have a dual degree combining one of these fields with biology. Christopher Lee, director of the programme, says that many bioinformatics courses are affiliated with data-rich biology labs on campus, supplying the students needed to tackle a flood of data. They often lack, however, the matrix of expertise necessary to conduct innovative analyses. Lee hopes that the UCLA programme will foster such expertise.

A few graduate training programmes, notably those at the Netherlands Bioinformatics Center in Nijmegen, cater to students with backgrounds in either computer science or biology. "We want to train the tool shapers as well as the people more into applying the tools in a biological setting," says Celia van Gelder, the centre's education project leader. "Over the past 10–20 years, the field of biology has become more computational, with bioinformatics serving as an interdisciplinary field that links researchers who can't otherwise readily talk to one another." The scope of work is widening, she says. As a result, demand for bioinformatics training continues to increase across Europe — with greater emphasis placed on data analysis at all levels. "We produce trainees who have multidisciplinary training in molecular-biology principles as well as algorithms to deal with data," says Jaap Heringa, the centre's scientific director for bioinformatics education. "Things move so fast in bioinformatics, we are constantly innovating our courses," he adds. Murphy agrees; Carnegie Mellon and the University of Pittsburgh offer in-depth training. "We are pretty clear in the application materials that our programme is not for people who want to get enough of a smattering of computational biology to get a job," says Murphy.

EXPANDING OPTIONS

This trend towards creating more comprehensive, interdisciplinary training programmes has gained momentum at biology strongholds in the United States. In July 2010, Dartmouth Medical School in Hanover, New Hampshire, established the Institute for Quantitative Biological Sciences in nearby Lebanon. Its graduate offerings combine modules in bioinformatics, biostatistics and epidemiology. "We have created what we think is a model of the future — training computational-biology students to speak multiple languages beyond bioinformatics," says the centre's director, Jason Moore. He adds that the key is assuming complexity rather than simplicity when approaching a problem.

In August, Moore secured funding to create a US National Institutes of Health (NIH) Center for Biomedical Research Excellence, through which he will mentor five early-career bioinformatics faculty members, to be recruited over the next 3–4 years. After two years of learning how to secure competitive funding, among other things, trainees will be required to submit an application for an R01 grant, the NIH's main funding mechanism. "We really want to provide a well rounded education so that our new recruits can secure funding for — and conduct — well designed studies in computational biology," says Moore.

Other medical schools are also taking the plunge. Duke University School of Medicine in Durham, North Carolina, formed its Department of Biostatistics and Bioinformatics in 2000. This year, it opens its first master's programme, says Elizabeth Delong, chair of the department.

BASIC SKILLS

Talent checklist

- Be at least conversant in the broad range of disciplines contributing to bioinformatics — from statistics to molecular biology to computer science.
- Most work, especially in industry, is done in teams, so communication skills are always in demand.
- Get experience in handling massive data sets. Learn to parse data or run analyses in parallel — using, for example, cloud computing.
- Learn to write programmes in software languages such as Perl or R.
- Cultivate a deep knowledge of at least one area of biology. **V.G.**

And in September, the University of Michigan Medical School in Ann Arbor established a computational-medicine and bioinformatics department to help attract new faculty members and trainees. In June, Emory University School of Medicine in Atlanta, Georgia, launched a biomedical-informatics department with the goal of combining expertise in imaging, computer science and biology to improve patient care. It will recruit four or five researchers over the next few years. "Our particular strength is training computer scientists who want to transition into biomedical informatics, and bringing them together with clinicians to use informatics to treat disease," says department chair Joel Saltz.

Qualified postdocs remain in demand. "It can be very difficult for individual investigators to hire a postdoc in bioinformatics," says Tom Tullius, interim chair of the bioinformatics programme at Boston University in Massachusetts. He attributes the paucity of candidates in part to efforts over the past several years to build large teams at high-powered institutes — such as the Broad Institute in Cambridge, Massachusetts, or the Wellcome Trust Sanger Institute in Cambridge, UK — leaving smaller labs struggling to find talent. The growth of training programmes could ease this.

M. VAN ZWAM

"We want to train the tool shapers as well as the people more into applying the tools."
Celia van Gelder

Now sequencing centres won't be the sole providers of data, individual researchers, particularly at medical centres, will have ample data to fuel research and training. "We've passed out of the period of genome projects where there were amazing public data raining down from the heavens; it's now possible to do exciting work without being associated with data-generating centres," says Lee.

Sczyrba, who begins a junior faculty position in metagenomics at the University of Bielefeld Center for Biotechnology in Germany this autumn, says that unpredictability is what makes the discipline so exciting. "We don't know where we will be in ten years because the technologies and ideas are moving so fast," he says. As Cleaver notes: "Perhaps the best career strategy is to stay flexible and curious." ■

Virginia Gewin *is a freelance writer in Portland, Oregon.*

BEAU LARK/CORBIS

COLUMN

The human touch

A little empathy goes a long way in the competitive confines of a laboratory, argues **Lydia Soraya Murray**.

As almost every scientist knows, a person's first year in research is an emotional minefield. One minute you're flying high. The next, you're banging your head against the wall, resisting the urge to draw in results with a marker pen. Forget the F-word; in science, it's the O-word that generates dread. I sometimes think that 'optimizing' should be spelled 'r.e.p.e.a.t.e.d.f.a.i.l.u.r.e.s'.

After a stimulating yet often soul-destroying start to my PhD, I have decided that coming to terms with the lows is one of the most important things that you can take away from your first year. Never mind the dreaded literature review; this is unquestionably more important.

Because scientists do incredibly specialized and often misunderstood work, it can be hard for people outside our particular fields to empathize with our attachment to our projects. I have a close friend who is a physician. After several weeks of my hard work culminated in what can be described only as 'diddly squat', my friend offered these consoling words: "It's not as if someone has died". To this day, I don't think he realizes how close he came to getting stabbed in the eye with a pipette. Instead of taking bloody revenge, I pointed out that if researchers didn't care so much, he would still be treating head colds with leeches — a less satisfying but more legal response.

Unfortunately, voicing frustrations to colleagues can be just as futile, and prompt the short and not so sweet response: "That's science". To be sure, cultivating a career in the frighteningly competitive world of research leaves no room for hand-holding or molly-coddling, and I truly believe that principal investigators need full-body elephant-hide transplants to achieve the thick skin required for the job. However, we are all human and everyone needs some sort of coping mechanism. Losing this mechanism is always a disaster.

In truth, there is no magic answer for how to deal with a disappointment rate of 90%. Some people build up walls to protect themselves, but this can result in suppression of all emotion. And let's be honest: given the hours that scientists work and the wages we earn, it is mostly our passion that keeps us chained to the lab bench. Dulling the rare moments of true toe-tingling excitement when things work and we discover something for the first time would be far too big a sacrifice. But others might have quite a different attitude and feel the disappointment so acutely that it destroys their confidence and paralyses them. Channelling your emotion into something manageable is truly important. I suggest that anyone new to research should find healthy ways to deal with their frustrations. Some people read or play a sport; others go out dancing. My coping mechanism is a large glass of red wine and a fantastic group of friends who put up with my rants, then shut me up with a good dose of perspective and insight.

Humans are social animals, and sometimes solitude enhances the feeling of ineptitude and makes dealing with a disappointment even harder. Maybe next time someone wanders past you gazing forlornly at their lab book with that oh-so-familiar look of puzzlement and frustration, a wee pat on the back and a bit of camaraderie might help. Yes, 'that's science'. But perhaps they'll see that success is possible despite repeated failures.

Lydia Soraya Murray *is a PhD student in molecular genetics and cell biology at the University of Glasgow, UK.*