Improving quantitative and computational skills in under-graduate bioscience education

Tom Andersen,
Department of Biosciences, University of Oslo
(Currently visiting scholar at Large lakes Observatory, UMN – Duluth)

The Department of Biosciences at the University of Oslo, Norway, has gone through a major revision in its bachelor's and master's programs the last few years. The main principles behind this reform have been to strengthen the basic natural science foundations (chemistry, physics, mathematics), offer a broad introduction to all aspects of biosciences (from molecular to planetary level), and to have computational and quantitative skills as a cross-cutting theme. For the latter we have had the advantage of being affiliated with a national center for computing in science education (https://www.mn.uio.no/ccse/english/). We have also had major inspiration and cooperation with our colleagues at the University of Bergen who run a national center of excellence in biology education (https://bioceed.w.uib.no/). My involvement in this process has been both strategic, as head of education at our department, as well as operational, by developing and performing education in quantitatively oriented skills.

The last 10 years I have been involved in a large, 20 ECTS course (2/3 of a semester workload) on biostatistics and study design (https://www.uio.no/studier/emner/matnat/ibv/BIO2150/index-eng.html). The contents of this course have been a blend of hands-on experiments, theory (statistical and dynamical modelling, experimental design, etc.), practical computing labs (data organization and exploration, data visualization, statistical model building, etc.), and report writing assignments. R (https://cran.r-project.org/) is the computational platform we use throughout the course, and in later years we have also advocated Rmarkdown (https://rmarkdown.rstudio.com/) for lab reports and written assignments. We have tried to use real, messy, biological data from peer-reviewed publications whenever possible in assignments and exams.

The student evaluations have been mixed, to put it mildly – many students feel that this type of course was not what they expected when choosing to study biology and question its relevance for their career plans. Others may even say that they deliberately chose biology because they did not feel comfortable with mathematics and "hard" natural sciences. Still, it never stops to amaze me how fluently many students seem to master quantitative skills at the end of this course (we have about 60 students per year, of which about 90% finish, and hardly any of those fail the final exam). The more disappointing side is how much of this knowledge that seems to be forgotten just a year later when I meet many of the same students in master's courses and as thesis supervisor.

Which is partly why we have chosen a different strategy for our new bachelor's program: instead of having quantitative and computational skills in a single, big course, we now try to embed quantitative activities in many courses across all semesters. We start the first semester with a specially designed scientific computation and modelling course, based on a "just in time" approach. That is, concepts of programming (arrays, loops, conditionals, functions, etc.) and mathematics (difference equations, matrices, eigenvalues, etc.) are introduced through biological examples, and by using only the skills and concepts they have learned so far. We then try to apply the same skills using the same tools (primarily Jupyter notebooks) in computer labs in biochemistry, cell biology, and physiology courses in the same or next semester. We still haven't figured out how to solve the schism of programming languages in the biosciences, so while we start with python in the first 2 semesters, we turn to R in the 3rd semester which the students use in courses in statistics, evolution, biodiversity, and ecology.

We do not yet have a full evaluation of the revised program since the first cohort of students will finish this summer. But we have observed several challenges, especially on the educator side. First, there is always the question of what you must leave out from the curriculum to make room for something new like computational skills. Some of this can be resolved with improving coordination between different courses (e.g. not teaching the Calvin-Benson cycle both in the biochemistry, cell biology, and physiology courses), but definitely not all. Secondly, there is a skepticism among many educators on what is gained from the quantitative skills compared to more classical, factual knowledge. Thirdly, many of our staff lack skills to develop computational course material themselves, which both puts a lot of responsibility on a small number of faculty as well as increasing the tension between different disciplinary traditions. As a long-term solution, we now include a passage on ability to teach computational skills in all our new faculty position announcements. But even as our department is planning to hire 10-15 new staff in the next 5 years, this transformation will obviously take time.

We also try to improve the student's quantitative background, by, starting from autumn 2019, requiring the highest level of high school math for admission to the biosciences program. Since less than half of our students currently have this qualification, this is obviously a risky move that we have not yet seen the results of. But the hope is that, even with an expected reduction in number of applicants for our program, these students will come out better equipped for the biology of the future.

In retrospect, after reading Elrod (2014), I get a feeling that we may have put too much emphasis on just one aspect of quantitative reasoning, namely computational literacy. I will certainly take with me the broader perspective of quantitative reasoning when we start to see the outcomes of the current reform and start planning for further improvements.