

SysMIC: A Blueprint for Interdisciplinary Online Training in the Life Sciences

Gerold Baier^{1*}, Chris Barnes¹, David Crowe², Stephen Gilmore³, Uwe Grimm², Phil Lewis¹, David R. Morse⁴, Clare Sansom⁵, Robert D. C. Saunders⁶, Adrian Shepherd⁵ and Geraint Thomas¹

¹: Department of Cell and Developmental Biology, University College London, London, UK

²: School of Mathematics and Statistics, The Open University, UK

³: School of Informatics, University of Edinburgh, UK

⁴: School of Computing and Communications, The Open University, UK

⁵: Department of Biological Sciences, Birkbeck, University of London, London, UK

⁶: School of Life, Health and Chemical Sciences, The Open University, UK

* Corresponding author: Gerold Baier, Department of Cell and Developmental Biology, University College London, Gower Street, London WC1E 6BT, UK. Email: g.baier@ucl.ac.uk

Abstract

'Big Data' and a surge in quantitative methodologies increasingly urge bioresearchers to train and improve their mathematical and computational skills. In response to this challenge the Biotechnology and Biological Sciences Research Council in the United Kingdom offered funds to develop web-based training in mathematics and computation for life scientists in all fields. The SysMIC consortium won the bid and implemented a new type of online course, which so far has been delivered to more than 1400 scientists, mostly in the UK. Researchers train self-paced or in cohorts in a media rich environment with materials based around intuitive biological examples. We describe the course, its trainee spectrum, and the modes of trainee engagement. We propose that tailored interactive online training is a scalable model for continuing professional development in the interdisciplinary life sciences in the 21st century.

Introduction

Biological research and bio-related industries have undergone a transformation from small-scale laboratory based methodologies to a large-scale computerised and increasingly quantitative interdisciplinary endeavour (Wooley, Lin, and National Research Council (U.S.). Committee on Frontiers at the Interface of Computing and Biology. 2005). Quantitative tools and computer programming are now required for many aspects of the work of life scientists and mathematics, statistics and computation have been identified as “vulnerable skills” in 21st century bioresearch (BBSRC 2015). Researchers in all branches of bioscience can no longer solely rely on colleagues but need to gain the competence in computation to deal with and make sense of their digital data. They face the challenge of integrating their professional development in these skills into tight schedules and making the learning of novel skills part of their work. This training need is, however, not generally addressed by standard university degree programmes. There is therefore an increasing demand for novel solutions to this dilemma.

SysMIC (Systems Approaches in Mathematics, Informatics, Statistics and Computational Biology, sysmic.ac.uk) is a new type of postgraduate course that provides online training for bioresearchers. It is run by a consortium of four UK universities, the University College London (UCL); Birkbeck, University of London; the Open University; and the University of Edinburgh. Development of the course was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) – the UK’s main governmental source of funding for non-medical biological research – to address the recognised and growing need for this kind of interdisciplinary training. The requirement (as defined by the BBSRC) was to develop an on-line distance-learning course suitable for postgraduate researchers with a background in the traditional (non-mathematical) biosciences. It was expected that it could cater for all stages of a scientific career, from PhD student to principal investigator. Now in its fifth year, more than 1400 students have enrolled on the SysMIC course, and it is attracting broader audiences including trainees from biomedical research centres, industry and researchers based outside the UK.

One of the main characteristics of SysMIC is that it occupies a middle ground between traditional face-to-face courses and MOOCs (Massive Open Online Courses) aimed at broad audiences, see (Ding 2014). SysMIC aims to combine the strengths of each mode of delivery, while avoiding some of their weaknesses. Face-to-face courses on the one hand naturally facilitate direct interaction between students and educators, immediate feedback to questions, and rich inter-student interactions within a given cohort; on the other hand, they are typically expensive, and the timing and location of classes is largely inflexible. In contrast, MOOCs are commonly free or cheap, allow (at least in principle) engagement via the Internet without geographical restriction, and the speed, timing and extent of engagement can be student led. However, completion rates are often very low. There are several reasons for this. MOOCs are free to join and the online enrolment process is quick and easy. Actually working through the course requires time and effort and this is often compromised when it conflicts with other obligations and when failure to complete the course has no negative consequences. Students may also decide to stop and therefore not complete a course because they have already achieved their own personal

objectives by working through parts. However, low completion rates can also reflect a substantial degree of dissatisfaction with the quality of the learning experience – perhaps most often due to a lack of individual interaction and support (Hew and Cheung 2014). SysMIC tries to avoid some of these factors by developing online teaching materials that incorporate a rich combination of different media and are complemented by direct interaction between students and instructors as well as more traditional online interactions between students (via forums). Importantly, MOOCs often target beginners and undergraduate audiences but relatively few specifically target professional scientists at the postdoctoral research level like SysMIC.

We give an overview of the course topics and the participant spectrum, highlight the course's modes of operation and interaction and draw some conclusions about the role of web-based training for the future of research-based professional development in the biosciences.

Approach

Mathematics and computing skills are urgently needed in modern biosciences and biomedicine. For the professional, this need is conflicted by the time needed for formal training in those skills. On the other hand, an additional full degree may not be what is actually required. What is needed is to be able to incorporate *tools* from mathematics, computer science and engineering into ongoing biological research. This includes generic skills like formulating a hypothesis in a quantitative mathematical way and specific skills like using a certain programming language to perform an analysis or simulation. The SysMIC course has tried to address this demand by selecting and organising the training material not from the mathematician's or computer scientist's point of view but from the point of view of the life scientist.

Rather than designing a specific course on, for instance “Networks”, “Mathematical Modelling” or “Data Analysis”, we have picked relevant examples from areas of molecular and cellular biology and designed the mathematical concepts and computing around them to address questions that emerge from the examples. This simplifies the entry points for trainees that come with disparate prior knowledge but share an interest in applying these skills to some area of bio research.

Syllabus

The course syllabus was largely determined by an analysis of the skills shortage in the UK bioscience community (BBSRC 2015). There are two modules: the first covers basic mathematical skills and programming for biology, and the second combines a selection of advanced topics from computational modelling and data handling. An outline of the topics covered in the two modules is given in Table 1.

The material in each session is focussed on relevant biological examples. Mathematical concepts and programming are taught in sections related to the context of the respective examples. For instance, the first session in part 1.1 introduces to some aspects of graph theory to aid the understanding of mathematical concepts behind network biology. On the computational side, functions from the Matlab bioinformatics toolbox are demonstrated in order to generate network graphs. The assignment asks trainees to produce a network matrix and create a network graph for the hierarchical organisation of species in a food web. In addition to the worked out examples, elaborated background materials are provided for in-depth study and reinforcement of the relevant mathematics. For the above example, a thorough introduction to graphs and networks is offered, followed by a description of network structures and motifs. Module 1 ends with a mini project which involves solving small biological problems by applying the skills learnt in the module. As an example, one project requires the implementation and subsequent expansion of the FitzHugh-Nagumo model of action potential firing in neurons to account for qualitatively different experimental firing patterns.

Following an introduction to programming, the sessions in Module 1 address three key approaches to biology. The first is network biology, which is based on the assumption that biological function arises from the interaction of molecular and cellular components rather than the working of such components in isolation. This topic requires knowledge of graph theory and the computational handling of multivariate data sets, including high-throughput data from ‘-omics’ technologies. The second approach is dynamical biological systems, based on the idea that biological insight about interaction between components can be formulated as a mathematical model for computational simulations and predictions. And the third is statistical data analysis, which allows rigorous quantitative conclusions from experimental data to be drawn. All three address generic needs in interdisciplinary research and are applicable to many biology, biotechnology and biomedicine.

Module 2 deals with deeper mathematical concepts and more advanced programming skills with an emphasis on dynamical biological systems (Table 1). The sessions comprise a repertoire of modelling and analysis approaches to tackle a variety of research questions in biology. Among these are discrete time models (linear and nonlinear) and continuous time models (deterministic and stochastic) and spatio-temporal models. They are complemented with an introduction to multivariate data analysis and classical and Bayesian parameter fitting approaches.

There is a clear progression from Module 1, which addresses beginners’ needs and starts approximately from the high school mathematics syllabus (A level in the UK), to the more complex research-oriented concepts discussed in Module 2, including for example the stability of solutions of nonlinear ordinary differential equations. Even these, however, are within the level and requirements of mathematics courses given to e.g. first- and second-year undergraduates in physics. SysMIC could perhaps be thought of as a route to these concepts and skills tailored to professional bioscientists at any stage in their careers.

Table 1: Summary of the current SysMIC course syllabus, April 2016

Module 1: Basic quantitative skills for biology				
Section	Biological Examples	Mathematics / Programming	# of sessions	Estimated duration
1.1	Soil biota food webs; Glycolysis; Signalling protein interaction networks	Introduction to MATLAB Graphs & Networks	3	30hrs
1.2	Plasmid replication	Vectors and matrices	1	10hrs
1.3	Bacterial growth; Enzymatic reactions; Gene expression; Gene regulatory networks	Functions and calculus	3	30hrs
1.4	Receptor-agonist interactions, signalling pathways	Model building, differential equations	1	10hrs
1.5	Imaging cells; Microarray analysis	Introduction to R Statistical analysis	3	30hrs
1.6	Predator prey; Allosteric regulation of PFK1; SIR modelling of measles; Neural firing; Lac operon	Mini project		flexible
Module 2: Advanced modelling skills for biology				
2.1	Bacterial ageing; Proteomics of biomarkers	Eigenvalues and eigenvectors; Principal component analysis	2	20hrs
2.2	Simple fluid flow; Enzyme kinetics	Multivariable functions	1	10hrs
2.3	The toggle switch (Gardner, Cantor, and Collins 2000); The repressilator (Elowitz and Leibler 2000)	Systems of differential equations	2	20hrs
2.4	Population growth and chaos; Ion channel gating; Bacterial competition	Discrete systems, cellular automata	2	20hrs

2.5	Pattern formation, Turing patterns and waves	Reaction-diffusion systems	1	10hrs
2.6	Bacterial growth; Dimerisation; Enzymatic reactions	Stochastic systems	1	10hrs
2.7	Optimisation; Genetic algorithm; Bayesian methods	Fitting models	1	10hrs
2.8	Predator-prey; Stem cell transitions; Inferring biochemical parameters	Mini project		1 month

(Table 1, continued)

Software

The programming language Matlab® (the Mathworks Inc., Natick, MA, 2016) is used for parts 1.1 through 1.4 and the mini projects (part 1.6) in Module 1 and throughout Module 2. Upon registration to the course participants are provided with a Matlab license. Matlab serves as the main programming environment because it is widely used within modelling-supported biological research (Hubner, Sahle, and Kummer 2011). It is professionally supported and can be considered an accessible entry point for programming beginners. Nevertheless, an alternative version of the course based on open source software (like Python) might be considered.

The programming language **R** (<http://www.R-project.org>; for statistics) is introduced and used in part 1.5, Statistics. **R** is a free software designed for statistical computing and graphical display. Like Matlab it runs on a large variety of platforms.

One important aspect for the course is to guide non-programmers smoothly to employ programming for their daily needs in research. Participants who complete module 1 will have used two common programming languages and acquired a functioning knowledge of programming that will facilitate their further learning.

Mode of delivery

In its current form, both modules are presented twice each year, beginning in April and October.

In Module 1, the release of materials is split into fortnightly sessions, which corresponds to an average workload of five hours per week. This biweekly release of material applies to eleven sessions in Module 1 and is followed by the flexibly timed miniproject. We hoped that this would allow individuals to comfortably accommodate the workload

within their professional schedules. Upon request, we managed to provide weekly release of materials to individual cohorts.

In Module 2 there is full access to the materials for the ten sessions for the period of 6 months (rather than the regularly paced release of Module 1).

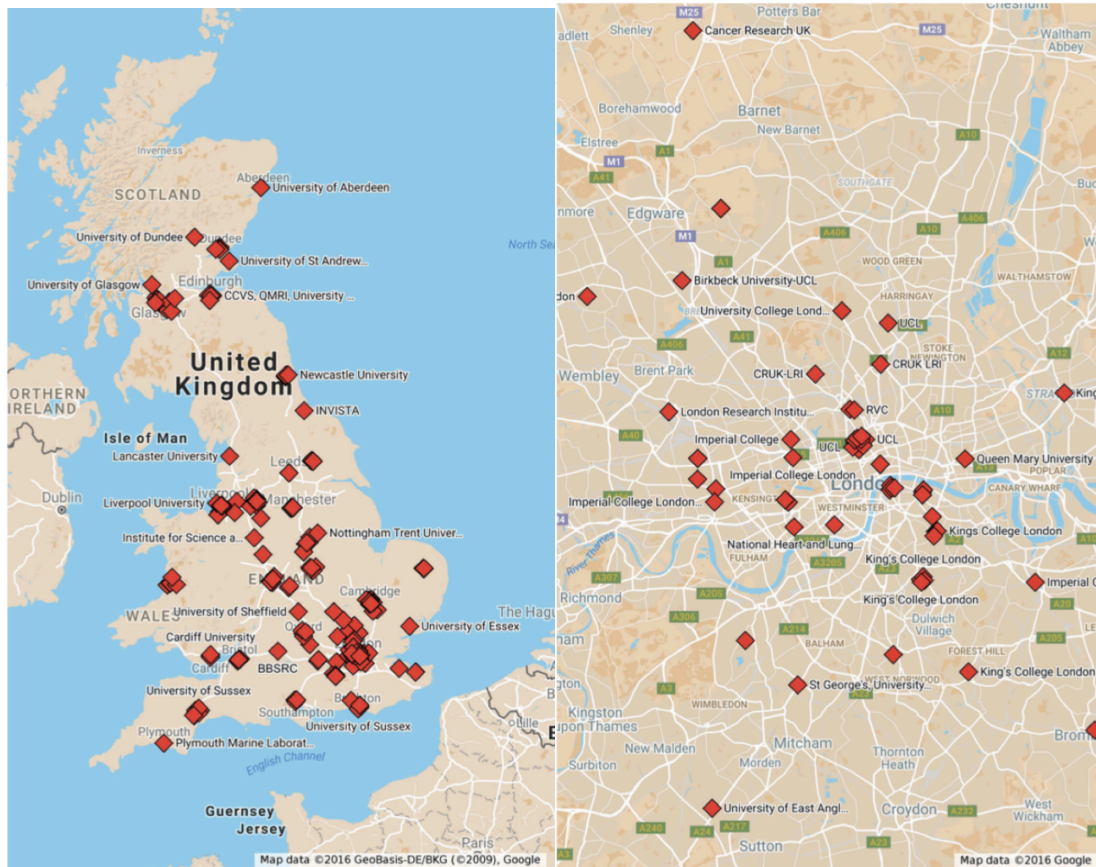


Figure 1: Distribution of SysMIC participants in the UK (left) and zoom into London area (right).

Table 2: SysMIC Student Numbers, April 2016

MODULE 1	Number	Percentage
Total Number of Participants	1233	
Number Completed	618	50.1%
Re-enrolments	96	
MODULE 2	Number	Percentage
Total Number of Participants	241	
Number Completed	82	34.0%

Participants

The course was launched in 2013 and as of April 2016, a total of 1233 trainees had registered for Module 1 (Table 2). Of these 62% were part of a UK-based doctoral training programme (DTP, all together 12 DTPs involving all major research universities in the UK) and the remaining 38% were self-enrolled. See Figure 1 for the distribution of participants' institutions in the UK. The requirement for completion is the submission of at least 10 assignments plus completion of a marked mini project and 50.1% of trainees enrolled in this module completed the course.

An analysis of the demographics of 457 non-DTP trainees (Figure 2) illustrates the broad range of individuals undertaking the course. 80% of these trainees were above the age of 24, with 22% and 9% in the age categories of 35-44 and 45-59 respectively. 42% of individuals were female, which compares favourably with the 28% of registered participants in a bioinformatics MOOC (Ding 2014) and the recently reported 38% in STEM subjects (Arnett 2015; WISE 2015).

The majority of trainees (52%) were at PhD level but 44% were at a later career stage including postdoctoral researchers, research fellows and principal investigators. This spread across career stages reflects our success in addressing the original BBSRC desire to increase mathematical and computing skills at all levels. Many individuals reported being funded by the BBSRC and thus fell within the remit of the original outline of the course. In addition, over time researchers supported by other UK funding bodies were also represented, notably the Medical Research Council (MRC) (Figure 2D). One of the main challenges for implementing the SysMIC course was the lack of formal mathematical training of life scientists in the UK. This is shown in Figure 3A, where 26% of trainees had only the equivalent of the UK's GCSE level (the high school level expected for 16 year olds) mathematics training and only 57% had the equivalent of A-level mathematics in the UK (taught to a fraction of high school students up to the age of 18). In addition, the trainees came from a very wide range of scientific backgrounds;

simple text-based frequency analyses of their self-reported research area and research interests are shown in Figures 3B and 3C respectively.

Despite this wide range of backgrounds and mathematical expertise, our completion rate is around 50% (table 2) for module 1. This a long way above those generally recorded for MOOCs. A meta-analysis of MOOC enrolment and completion rates from 2014 found an average completion rate of only 6.5%, although with far higher student numbers (Jordan 2014). The completion rate also compares well with those of other distance education programmes, for instance about 65% at the Open University for degree courses, which offers significant levels of tutorial support. One reason for the difference might be that unlike the Open University SysMIC currently does not offer formal qualifications. The completion rate for the more advanced Module 2 is 34%.

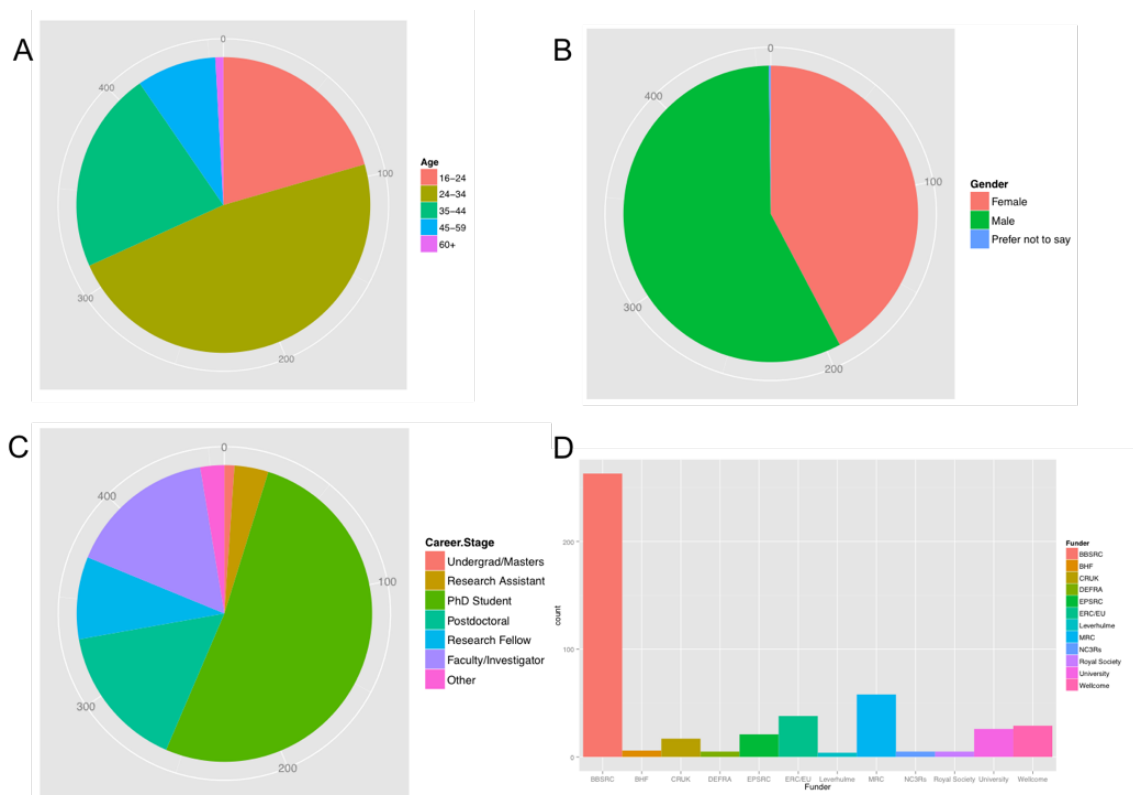


Figure 2: Demographics of 457 individuals who were not part of a doctoral training scheme. A-D) Age, sex, career stage and funding respectively.

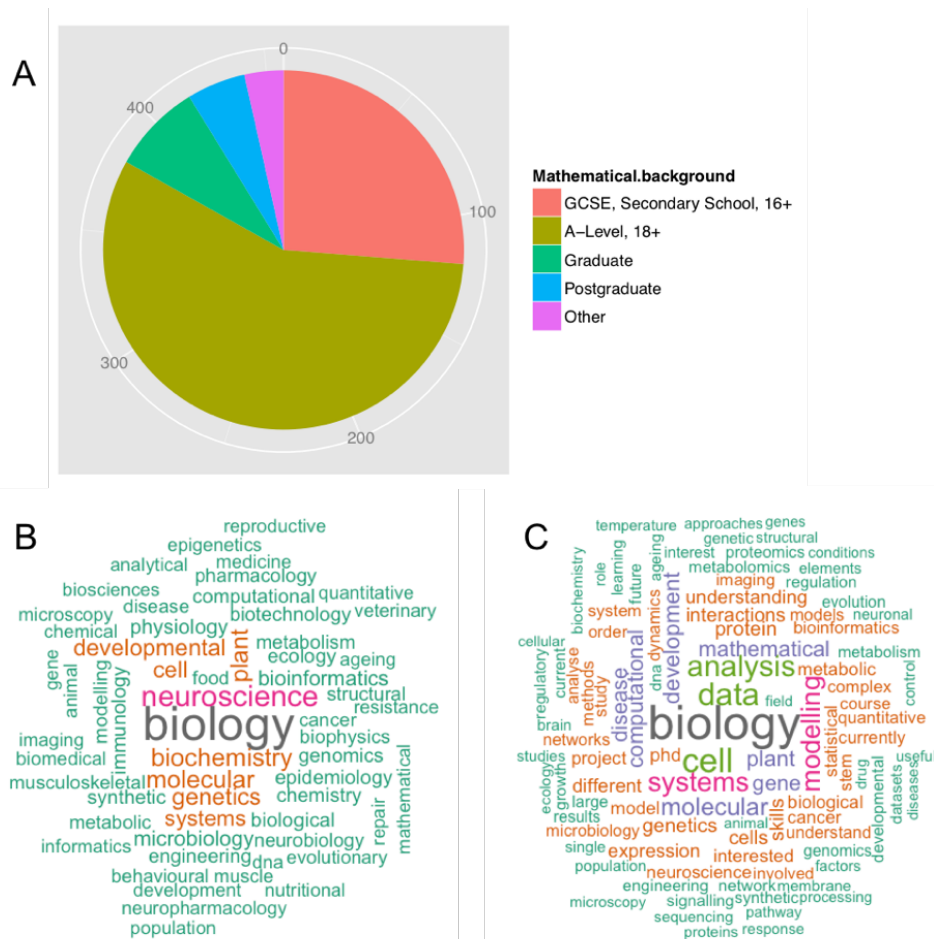


Figure 3: Scientific backgrounds of 457 individuals that were not part of a doctoral training scheme. A) The self-described mathematical level. B and C) Results of basic text mining on research area (B) and research interests (C).

Structure of Materials

Course Material. The course is hosted on the Moodle open source learning platform with access to the course materials via a web portal. Most materials are available as both static documents – a “textbook” of files in PDF format suitable for download and study independent of internet access – and an interactive, hyperlinked online workbook. Maintenance of two delivery modes is made practicable by ensuring the core content is written only once in LaTeX and then converted to both PDF (for the textbook) and HTML (for the workbook). Both modes have a consistent division into the mandatory Biological Examples and complementary Mathematical Background parts. The Biological Examples part elaborates mathematical concepts in a biological context and explains the code used to work out solutions step by step. Program code is contained within text boxes that support syntax highlighting and can be copied conveniently to the trainees’ programming environment. The Mathematical Background parts back up the

concepts taught with detailed derivations and exercises (including model solutions) for practice, refinement and future reference. Introductory screencasts are offered on topics to help with getting started, in particular for programming. Further practice and self-testing is done in quizzes on the mathematical content (Figure 4).

Question 1

Complete

Marked out of 1.00

Look at this graph.

What is the clustering coefficient of node a?

Answer:

Node *a* has 5 neighbours, which are inter-connected by 4 edges (there are $\frac{5 \times 4}{2} = 10$ possibilities). Hence, the clustering coefficient of node *a* is $\frac{4}{10} = 0.4$.

Figure 4. An example of question from an interactive quiz embedded in the SysMIC Module 1 session on biological networks (section 1.1). The concept of clustering coefficient is explained in the materials. The panel at the bottom is revealed after the student submits an answer.

Interactivity is enhanced within the workbook with the addition of embedded applets. These have a mathematical function or equation set up to allow trainees to explore the effects of modifying parameters. Figure 5 shows an example. A model of gene expression regulation is set up and a graph displays simulations of the expression levels of three genes. Trainees can modify the settings and study the model predictions by manipulating sliders that control model parameters. This serves to build up an intuitive understanding of how components of the model contribute to, in this case, the appearance or disappearance of oscillatory gene expression.

Assignments. To complete a session, trainees are required to complete an assignment in which they apply of concepts learned in the session body to a different situation. This serves to consolidate the learning of the concepts and contents with further practice in Matlab or R. Assignments build on the structure of the Biological Examples and typically consist of a programming example with generation of numerical and graphical output.

The results have to be uploaded to the learning environment and are an obligatory part of the course.

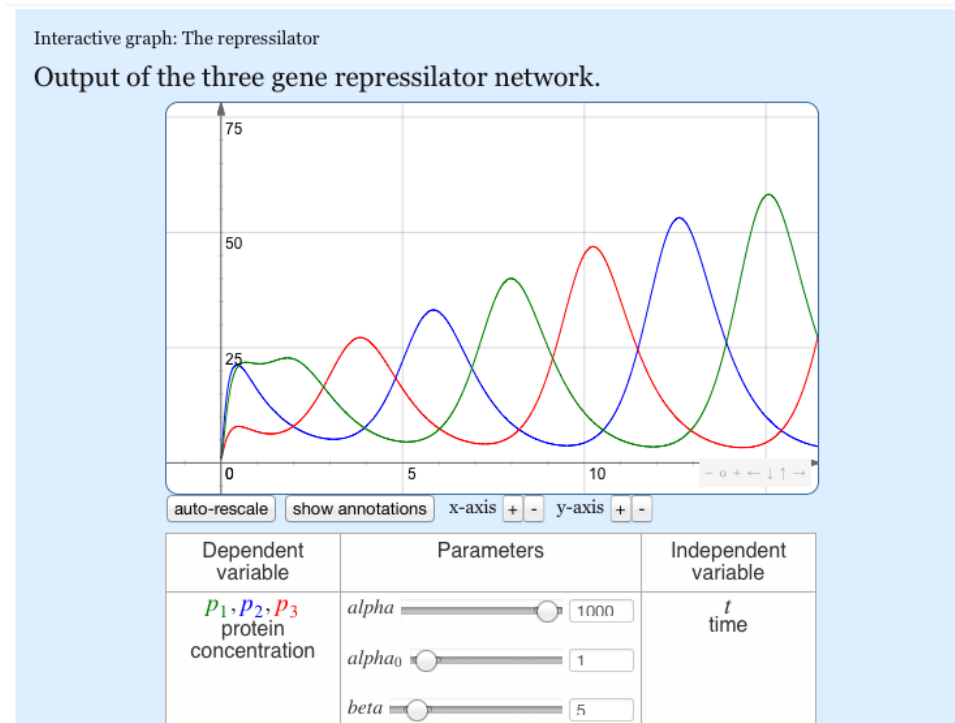


Figure 5. An example of an interactive applet embedded in the SysMIC workbook. It is based on the repressilator model (Elowitz and Leibler 2000). The graph at the top shows varying levels of gene expression and changes dynamically as the sliders are moved. The applet was written using JSXGraph (Gerhauser 2012).

Mini projects. At the end of a module, trainees do a mini project. They are offered a number of biological models with which they work to interpret experimental data and make informed predictions. The result of a mini project is a report containing model description, code and simulations. The report is uploaded to the Moodle site for assessment by SysMIC course tutors. A pass mark for the mini project releases a certificate of the completion of the course.

SysMIC student engagement

In addition to the provision of course materials, SysMIC offers several ways to enhance the quality of student-student and student-educator interactions experienced by course participants.

Forums. An online forum accompanies each session. They contain previously posted questions and answers for immediate help and, similar to those in MOOCs, offer an interactive medium for trainees to post questions and comments, to ask for help and to get in touch. Queries are attended by the core team but at peak times SysMIC also employs dedicated tutors to ensure timely response to queries,. Trainee-driven forums are an important medium for keeping in contact with the SysMIC trainee community.

Monitoring. Our experience suggests that it is important to strike a balance between encouraging participants to stick to a predefined timetable, and being flexible about hand-in deadlines. We therefore monitor individual trainees' levels of individual engagement, collecting information such as timings of assignment submission and last login to the course website, and a student's own declaration of their status (e.g. that they are unable to engage with SysMIC for a defined duration). Based on these observations, we contact directly those students falling behind schedule and offer to provide support in order to allow continuation.

Mini project feedback. Increasingly we have recognized the importance of individualized feedback – trying to understand why a given student is not progressing, and helping them to overcome any obstacles. Given the large number of participants, individual feedback on assignments is not feasible. However, course tutors review the mini project reports and provide detailed feedback on the structure, contents and level of skills as evidenced in the report. For reports considered inadequate, we give detailed instructions on how and where to improve the contents. Thus, in contrast to MOOCs, we are able to offer an individual quality check of our participants' work.

Face-to-face events. A significant percentage of SysMIC participants are part of a *local cohort*, for example a DTP or a group of participants from a particular university or company. Where there is local institutional approval for student participation, regular face-to-face sessions run by local educators have proved highly effective. Where this is not feasible, support sessions via virtual classrooms using a desktop video conferencing system are an alternative. In addition, all SysMIC participants are offered places at dedicated SysMIC workshops. Following a similar model at the Open University, we have run several such workshops on introductory programming and group problem solving within a friendly, informal environment. We find them an important complement to the online provision, addressing individual needs, discussing contents, and building a learning community.

Web seminars and conference calls. As the assignments are time-consuming and challenging, we offer web seminars to support trainees delivered via the professional service Citrix® GoToTraining. These web seminars demonstrate the steps to be taken

during the assignment and explain how to address frequently encountered problems. Video recordings of these web seminars are posted on the course website. Occasionally, the support for individual problems with the course, the assignments or the mini projects via the online forums turns out to be insufficient. We address this by picking such cases and offering conference calls with individuals via Citrix® GoToTraining. Such calls allow detailed questioning and screen sharing and thus are an efficient and quick means to identify the root of a problem (for example programming errors).

Feedback on the course

Summary information about all aspects of the course is collected and aggregated in various ways and this has enabled us to efficiently monitor the enrolment, progression, engagement and completion of participants. In addition, following the upload of an assignment, trainees are surveyed about their opinions about the material and the level of difficulty before they move forward to the next session. And finally, following the successful submission of a mini project participants complete a reflective activity which serves as our “end of course” review. Taken together, these components serve both to provide dedicated feedback to the SysMIC administration and administrators of local cohorts (where we often observe a distinct “local cohort effect”, with participants from some institutions consistently more engaged than others), and also to modify and refine the contents of the course. Modifications introduced in this way include the simplification, splitting and reordering of materials, and the inclusion of a period for catching up with the assignments.

The future of interdisciplinary training

Interdisciplinarity is a feature of most if not all modern bioscience. We may expect increasing adjustment by major research funders to this evolution of the scientific landscape (Rylance 2015). Researchers in the biosciences nowadays find multiple solutions to fund and carry out interdisciplinary work as evidenced by an increasing number of citations from outside disciplines and reference to other disciplines (Van Noorden 2015). However, the associated need for training for interdisciplinary research has so far only been met with piecemeal and uncoordinated efforts, most of them implemented locally by individual higher education or research institutions. With SysMIC we have gone further: we have teamed up as a consortium of four institutions to bring in expertise in the life sciences and in distance education and implemented all materials online for both self- and cohort-based training. It is thus in principle available to anyone who has internet access at any time.

One major aspect for future development of materials is diversification. The current course picks and works through a single biological example for each session. We have seen increasing expressions of interest in this kind of training from pharmaceutical (including industrial) research (MedImmune, an R&D company linked to AstraZeneca)

and we have had various cohorts from a biomedical research institute, the Francis Crick Institute in London. To shift the focus to drug-discovery-related or disease-related research it might be advantageous to include alternative examples taken from the respective areas. This would make the course more modular in the sense that participants might be given a choice of which example to pick. Our experience so far indicates that each area has specific expectations and needs and a modular structure would allow a more tailored composition of materials for e.g. industrial or biomedical needs. We will try to address this by staying in touch with former trainees who come from those areas and who are willing to provide data and background materials that can be transformed into new training examples.

Thinking ahead, there are generic skills and specific topics that are not currently included and which might be increasingly relevant to larger numbers of researchers in the future. Advanced multivariate data analysis to deal with “big data” is an example. As the core team of the SysMIC course cannot be specialists in all potential future fields, one might consider opening up the course to collaborations with other institutions dealing with research-oriented training in other interdisciplinary endeavours. There we could offer expertise in the implementation of innovative training components into a scalable format for either purely web-based training of self-directed individuals or combined web-based / face-to-face training in institutions where course leaders would be able to pick material of their choice.

In the future we envisage that flexible online training environments will become core educational resources for interdisciplinary training in many profession. New demands to deal with training in an area of expertise from outside one’s own background will be a continuous challenge, not an exception. In combination with social media these environments might become part of a new culture of web-based learning and contribute to building global training communities among like-minded individuals.

Based on our experience, we assert that it is perfectly feasible for web-based training in the spirit of the SysMIC course to become a major route along which professionals (in the life sciences and elsewhere) will keep up to date in increasingly interdisciplinary research and research-intensive industries.

References

- Arnett, G. 2015. 'How well are women represented in UK science?', *the Guardian*.
- BBSRC. 2015. 'Review of vulnerable skills and capabilities report', <http://www.bbsrc.ac.uk/about/policies/reviews/consultations/1501-vulnerable-capabilities-report/>.
- Ding, Y.; Wang, M.; He, Y.; Ye, A. Y. X.; Yang, X. X.; Liu, F. L.; Meng, Y. Q.; Gao, G.; Wei, L. P. 2014. "'Bioinformatics: Introduction and Methods," a Bilingual Massive Open Online Course (MOOC) as a New Example for Global Bioinformatics Education', *PLoS Computational Biology*, 10.
- Elowitz, M. B., and S. Leibler. 2000. 'A synthetic oscillatory network of transcriptional regulators', *Nature*, 403: 335-8.
- Gardner, T. S., C. R. Cantor, and J. J. Collins. 2000. 'Construction of a genetic toggle switch in *Escherichia coli*', *Nature*, 403: 339-42.
- Gerhauser, M.; Miller, C.; Valentin, B.; Wassermann, A.; Wilfahrt, P. 2012. 'JSXGraph – Dynamic Mathematics Running on (nearly) Every Device', *Electronic Journal of Mathematics and Technology*, 1: 1933-2823.
- Hew, K. F., and W. S. Cheung. 2014. 'Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges', *Educational Research Review*, 12: 45-58.
- Hubner, K., S. Sahle, and U. Kummer. 2011. 'Applications and trends in systems biology in biochemistry', *FEBS J*, 278: 2767-857.
- Jordan, K. 2014. 'Initial Trends in Enrolment and Completion of Massive Open Online Courses', *International Review of Research in Open and Distance Learning*, 15.
- Rylance, R. 2015. 'Grant giving: Global funders to focus on interdisciplinarity', *Nature*, 525: 313-5.
- Van Noorden, R. 2015. 'Interdisciplinary research by the numbers', *Nature*, 525: 306-7.
- WISE. 2015. 'Women in Science, Technology, Engineering and Mathematics: The Talent Pipeline from Classroom to Boardroom. Statistics 2014', https://www.wisecampaign.org.uk/uploads/wise/files/WISE_UK_Statistics_2014.pdf.
- Wooley, John C., Herbert Lin, and National Research Council (U.S.). Committee on Frontiers at the Interface of Computing and Biology. 2005. *Catalyzing inquiry at the interface of computing and biology* (National Academies Press: Washington, D.C.).