

# Bioinformatics and Systems Biology: bridging the gap between heterogeneous student backgrounds

---

Title: Bioinformatics and Systems Biology: bridging the gap between heterogeneous student backgrounds

Running title: Bridging heterogeneous student backgrounds

Authors:

Sanne Abeln<sup>1,2</sup> [s.abeln@vu.nl], Douwe Molenaar<sup>1</sup>, K. Anton Feenstra<sup>1,2</sup>, Huub C.J. Hoefsloot<sup>3</sup>, Bas Teusink<sup>1,2</sup>, Jaap Heringa<sup>1,2</sup>.

1 Centre for Integrative Bioinformatics (IBIVU) & Amsterdam Institute for Molecules Medicines and Systems (AIMMS), VU University Amsterdam, The Netherlands

2 Netherlands Bioinformatics Centre (NBIC), Nijmegen, The Netherlands

3 University of Amsterdam, Swammerdam Institute for Life Sciences (SILS), Amsterdam, The Netherlands

Abstract:

Teaching students with very diverse backgrounds can be extremely challenging. This paper uses the Bioinformatics and Systems Biology MSc in Amsterdam as a case study to describe how the knowledge gap for students with heterogeneous backgrounds can be bridged. We show that a mix in backgrounds can be turned into an advantage by creating a stimulating learning environment for the students. In the MSc Programme conversion classes help to bridge differences between students, by mending initial knowledge and skill gaps. Mixing students from different backgrounds in a group to solve a complex task, creates an opportunity for the students to reflect on their own abilities. We explain how a truly interdisciplinary approach to teaching helps students of all backgrounds to achieve the MSc end terms. Moreover, transferable skills obtained by the students in such a mixed study environment are invaluable for their later careers.

Keywords: Education, Bioinformatics, Systems Biology

**Author Biographies Sanne Abeln is an assistant professor at the VU University of Amsterdam. Her research focuses on biophysical simulation and medical bioinformatics . She has taught MSc Bioinformatics courses at the University of Oxford and at the VU University in Amsterdam.**

## **Douwe Molenaar**

### *Sanne Abeln*

Douwe Molenaar is a researcher in Systems Biology and Integrative Bioinformatics at the VU University Amsterdam. His research areas of specialisation are evolution and microbiology. He teaches several courses in Systems Biology at Master level.

### *K. Anton Feenstra*

Anton Feenstra is assistant professor in Bioinformatics at the VU University Amsterdam. His research focuses on prediction and analysis of protein-protein interactions and coarse-grained modelling of signalling and regulatory networks. He has been teaching bioinformatics since 2006.

### *Huub C.J. Hoefsloot*

Huub C.J. Hoefsloot holds an MSc in applied mathematics and a PhD in Chemical Engineering. His research activities concern the development of analysis tool for complex dynamical data. He is an associate professor at the University of Amsterdam and has been teaching courses on modelling for more than 20 years.

### *Bas Teusink*

Bas Teusink is professor in Systems Biology and Integrative Bioinformatics at the VU University Amsterdam. His research interest is in evolution and regulation of metabolic networks. He teaches Biochemistry and Systems Biology in several courses at Bachelor and Master level.

### *Jaap Heringa*

Jaap Heringa is full professor in bioinformatics at the VU University Amsterdam and scientific director of the Netherlands Bioinformatics Centre (NBIC), where he additionally directs the educational platform BioWise and NBIC's PhD School. His research focus is on bioinformatics and systems biology and he has longstanding experience in developing sequence analysis tools. He has been teaching bioinformatics for more than 20 years and has initiated the VU master

programme in bioinformatics in 2003.

## **Introduction**

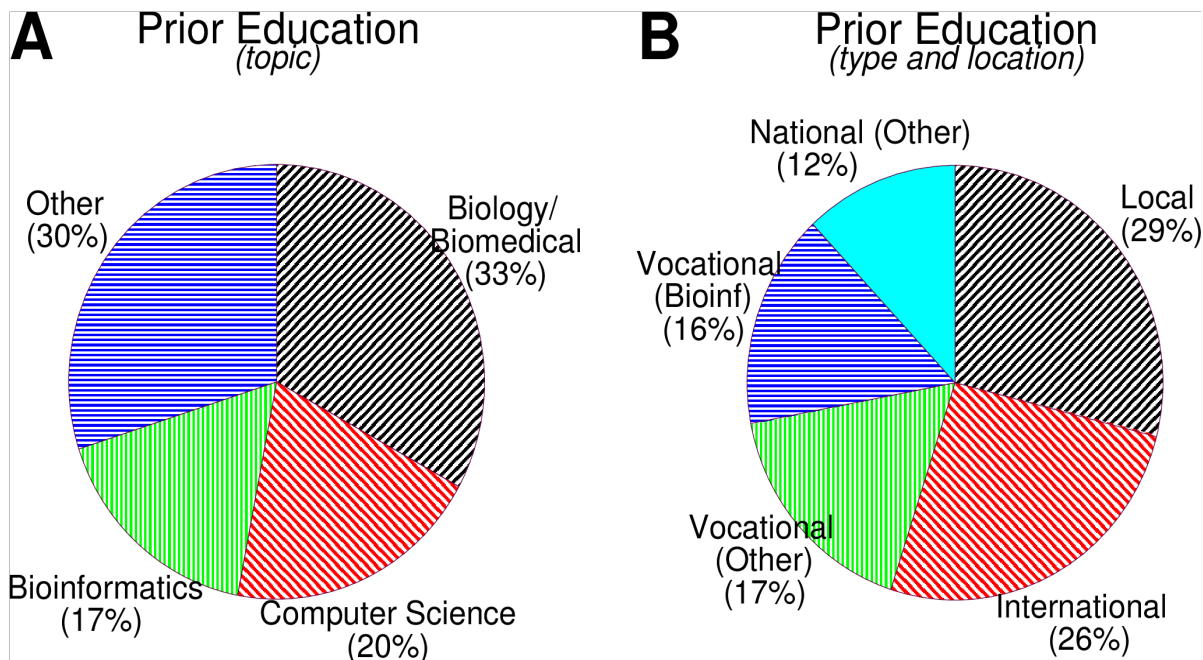
Designing a course or curriculum for students from different backgrounds can be extremely challenging. Typical challenges include creating assignments with enough complexity to reach the right academic level, avoiding the pitfall of teaching a separate course for each background, and motivating students to work on gaps in their knowledge. Below we use the MSc Bioinformatics and Systems Biology in Amsterdam as an example to show how such problems may be overcome.

## **MSc Bioinformatics and Systems Biology in Amsterdam**

The 2-year international MSc programme in Bioinformatics and Systems Biology is a joint programme involving three faculties over two universities: The Faculty of Earth and Life Sciences and the Faculty of Sciences at VU University Amsterdam, and the Faculty of Science at the University of Amsterdam (UvA). The programme has two major profiles: students can major in Bioinformatics or in Systems Biology. Incoming students are recruited separately at these two universities, but all students enroll in the same programme. Each course is taught at one of the two locations, and care is taken for students not having to travel between the two universities on a single day.

The MSc programme in Bioinformatics started in September 2003 with 5 students, and the current programme with the widened focus on Bioinformatics and Systems Biology runs since 2009. In September 2012 a total of 38 students registered for the joint master programme.

## **Heterogeneous student backgrounds**



**Figure 1:** Prior education of students in the MSc Bioinformatics and Systems Biology. **(A)** Type and location separated into academic local (Amsterdam area), vocational (Bioinformatics or other), academic national, and international. **(B)** Topic separated into Computer Science, Bioinformatics, Biology/Biomedical, and other. Statistics were gathered since 2003.

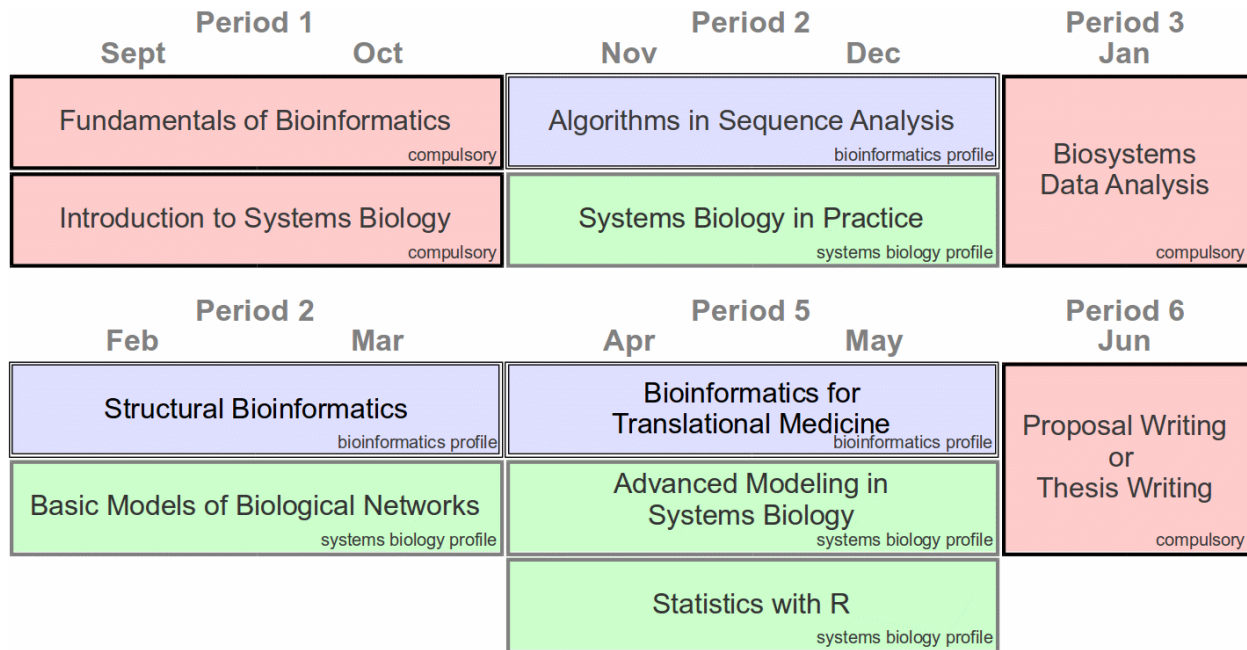
Since setting up the Bioinformatics Programme we have aimed to recruit students with diverse backgrounds, such that both students with a BSc in Computer Science, in Biology, or in a related discipline, can enroll. We believe this mixture provides the best environment for students to develop their Bioinformatics skills. Currently, first degrees (BSc) of our students include Biological Sciences (including Biomedical Science and Biology), Computer Science (including tracks such as Software Engineering and Artificial Intelligence) and Physical Sciences (including Medical Physics and Biochemistry). Note that all students beginning our MSc Bioinformatics and Systems Biology curriculum have some sort of gap in their knowledge - unless they were able to choose very specific optional courses within their BSc. Such knowledge gaps include mathematics, programming, molecular biology, and lab experience.

It is important to note here that in the Netherlands there is no BSc Bioinformatics at University level. Bioinformatics is only given at Universities of Applied Sciences (HBO in The Netherlands) as vocational training. We also accept ambitious students from these vocational backgrounds. For these students mathematical skills and - less obvious to the students themselves - an ability to analyse data independently are the canyons which need to be bridged.

Since 2009 we have merged Bioinformatics and Systems Biology in a single MSc programme. As a result of this combined curriculum, the backgrounds of incoming students have become even more diverse (see Figure 1). We now see a wide variety in the students' fields of interests

for their post-Master careers, ranging from theoretical modelling and simulation, to medical bioinformatics, to combining experiment with computational techniques and development of new computational methods.

## Curriculum outline



**Figure 2:** Overview of the compulsory (red) and profile courses as scheduled in the first year of the MSc programme; Bioinformatics profile compulsory courses are shown in blue, Systems Biology profile compulsory in green. The second year (not shown here) is reserved for a major and minor internship, where the major internship should match the profile chosen. Note also that scheduling is done in such a way that it is attractive (and possible) for students to choose both profiles rather than selecting three optional courses. All courses in the first year are 6 credits in the European Credit Transfer System (ECTS).

The curriculum follows a clear structure (Figure 2), with four compulsory courses (24 ECT) for all students. After two months of parallel compulsory courses (Introduction to Systems Biology and Fundamentals of Bioinformatics) the students can choose either a Bioinformatics or Systems Biology profile. The profile consists of 3 courses (18 ECT) specific to the profile and three optional courses (18 ECT). Good students with few deficiencies are allowed to follow both profiles, by choosing their optional courses strategically. Other students will need to use these optional courses to become more proficient in programming, mathematics or molecular biology.

Most students entering our MSc have a preference for either systems biology or bioinformatics and are typically only familiar with one of these research disciplines. It is interesting to see that

about  $\frac{1}{3}$  of the students decide after the first two months to take both profiles or to take the alternative profile to the one initially chosen. This illustrates the usefulness of having an introductory period of two months in which students can familiarise themselves with both fields. The vast majority of students will choose at least one of their optional courses from the “other” profile, suggesting the combination of bioinformatics and systems biology in a single MSc suits the interests of the students.

Our vision is that systems biology and bioinformatics should be taught as interdisciplinary subjects, and not as a multidisciplinary curriculum (e.g. by giving separate courses in biology, mathematics or programming). Instead, we provide examples of current research in bioinformatics or systems biology to cover, and at the same time amalgamate, necessary knowledge from “classical” disciplines on a need-to-know basis.

Roughly speaking we suggest that bioinformatics is an interdisciplinary field combining computer science with molecular biology, whereas systems biology combines biological experiment with mathematical modelling. Of course in both fields, molecular biology, basic mathematical skills, basic computational skills, statistics and some biophysics are also necessary. Having said this, we do provide our students at the start of our curriculum with short conversion classes to work on specific gaps in their knowledge.

At the end of the two-year MSc programme, students should have a basic understanding of both systems biology and bioinformatics. Moreover, they should have obtained in-depth knowledge in the area of their chosen profile and should be able to do independent research. Below we will outline how we bridge the knowledge gap of our incoming students, such that these end terms can be met.

## **Start of the program (Overcoming initial differences)**

### **Conversion Classes**

The programme has a very intensive start: students have 24 contact hours, with 40 hours of total workload per week, for the first two months. These hours are divided up between lectures, assignment classes and conversion classes. The intensive start sets the pace for the rest of the programme. It also means that students get to know each other well in the first two months; this tends to help them through the rest of the curriculum.

At the start of the course we overcome initial differences in the backgrounds of the students through group projects and conversion classes. Students start with two compulsory courses: Introduction to Systems Biology and Fundamentals of Bioinformatics; in parallel three different conversion classes are provided in Molecular Biology, Mathematics and Programming (Table 1). Students are assigned two out of three conversion classes to mend deficiencies in their background knowledge. Conversion classes count towards 30% of the final grade of the first two compulsory courses, leaving about two weeks' worth of study per class.

<b>Biology Class</b>	<b>Mathematics Class</b>	<b>Programming Class</b>
cell biology	basic calculus	Python scripting
genetics	linear algebra	functions, scope
cellular networks	differential equations	control structures

**Table 1:** An overview of course content for the conversion classes in Biology, Mathematics and Programming.

The aim of these conversion classes is not to teach three years worth of a BSc programme in two weeks. However, due to the higher motivation of students at this level and their previously obtained analytical skills, it is possible to cover much more material than in a similar BSc course. Most importantly, we try to give students the means to start working on their deficiency independently. By the end of the classes, and based on an assessment, they know which concepts and skills they ought to have, which enables the strongest students to work further on their own and advance their skills in the subsequent compulsory profile courses.

Note that not all students are able to pick up the required skills by the end of the conversion classes. Sometimes this can be solved by choosing the profile wisely. For example, while a significant number of students have a good feeling for mathematics, they struggle to get the



procedural and iterative principle of programming, showing that such skills are not equivalent. In such a case, we may advise the student to take the Systems Biology profile. Conversely, students with better programming skills would typically be advised to choose the Bioinformatics profile.

## **Background assessment and mentoring**

On the first day of the course we give a questionnaire in which students have to indicate their proficiency in biology, maths and programming. Based on this, together with the information on previously followed courses and the interests of the student, a mentor assigns two (out of three) conversion classes for the students to follow. Note that the mentors are course teachers of at least assistant professor level. The mentor also helps the student to make a choice between the Systems Biology and Bioinformatics profile. After two months a personal education plan (PEP) is made with the student to select all optional courses, and to discuss possible MSc research projects.

## **After the Conversion Classes**

For some students it is necessary to take further courses in order to follow their desired profile, typically Molecular Biology or Programming (Java/Python) for the Bioinformatics profile, and Mathematics or Molecular Biology for the Systems Biology profile. In agreement with the mentor, one course at BSc level may be taken as an optional course. Our rationale here is that for a student with a BSc background, one programming course at BSc level will be more challenging and more valuable than a typical advanced Molecular Biology course at MSc level.

It is also very important to note that programming and good mathematical skills may, to our observations, indeed be acquired after the BSc level - contrary to popular belief within those disciplines. For example, when students see the relevance of mathematics within research, we noticed that they can overcome problems they had with mathematics at high school. In fact, we see that ambitious students from vocational background will include theoretical work at a high level within their internships, or subsequent jobs. Our observations seem to agree with recent research that indicates motivation is more important in growth of mathematical skills than IQ [1].

In all compulsory (profile) courses, there are assignments that combine molecular biology with either programming, modelling, statistics or lab work, such that the skills learned in the conversion classes can be further developed. Depending on the initial background of the student, they may need to put more effort into one specific aspect of such an assignment. For example, one assignment in the “Algorithms in Sequence Analysis” course is to program out the Needleman-Wunsch algorithm and analyse the subsequent alignments. Students with a biology background will spend more time on the programming whereas students with a computer

science background will find understanding the outcome of the alignment and its biological relevance more challenging. This way the same task can be challenging in its own way for groups of both backgrounds, while keeping the interdisciplinary character of the research topic. Note that we do not take the background of students into account when grading the assignments; instead why try to balance the different skills to complete the tasks, while putting most weight on research and reasoning.

## **Group projects**

In both compulsory courses at the start of the programme there is a large assignment that needs to be solved as a group project. We make sure to mix students with different backgrounds so that they can combine their previously obtained skills to tackle a research project in either Bioinformatics or Systems Biology. In fact the project is set up such that it can only be effectively solved by a combination of background skills and knowledge from molecular biology and computer science. Every year students indicate that these projects work as an eye-opener: they realise how much other students already know about a subject and start valuing these skills. This gives a strong motivation to catch up using the conversion classes. At the same time, students learn to work together with students of different backgrounds, which is an invaluable skill in bioinformatics and systems biology research. Another benefit of these group projects is that, after early mixing, students from different backgrounds tend to support each other throughout the year. This greatly benefits the learning of those involved.

### *Project in Bioinformatics*

The aim of the Bioinformatics project is to benchmark (PSI-)BLAST [2] using the SCOP [3], GO [4] and PFAM [5] databases. To ease this task somewhat, we have selected a set of 100 proteins, with a sufficient number of homologues, on which students can perform the benchmarks. We also provide skeleton scripts in Python to automatically retrieve BLAST results from the web server, parse BLAST, SCOP, GO and PFAM annotation, and generate ROC-curves. Students have to fill in the most crucial parts of the scripts to get them to work, while the I/O is already written for them. This way students can focus on the major learning objectives of the project:

- Understanding the need for automation - Would such a test be possible by manually using the BLAST web server?
- Understanding the fuzziness of biological data – What does a benchmark mean (e.g. function GO, structure SCOP) and how reliable are the “true positives”?
- Relating the research question to the method - How does the performance of BLAST depend on the reference database used (students find this very difficult)?
- Analysing large scale data in a structured way - How to generate roc-plots?
- Interpreting results - What parameter settings for BLAST work best, and why?

Students have to write a report on the project within their group. A draft report can be handed in halfway for formative feedback. Students also compare results between different groups in a final presentation. Typically, this will reveal that seemingly minor changes in methodology and scoring can yield quite different results.

### *Project in Systems Biology*

Students participate in small groups in a literature project in which they analyse two or three original research papers on a topic in systems biology. The topics range from the use of mathematical models to understand gene regulation and metabolism to the statistical analysis of omics data and predictor identification. Each group presents the topic at the end of the course to their fellow students. For this, we organise a small symposium in which also the results from the bioinformatics projects are presented. Apart from this group effort, each student individually writes a summary of the topic. For both the oral presentation and the summary we ask the students to present those in a form that makes the topic understandable to the general scientific audience. We do this to stress the importance of this type of communication, in which the relevance of the topic for a wider field of science and/or society needs to be shown. We thereby enhance the skills of students to think about and also be motivated by the broader implication of particular topics in systems biology.

## **Curriculum of taught courses**

During the taught curriculum our key idea is to show complex and current problems within systems biology and bioinformatics to cover all necessary topics and skills. Typically there will be a couple of larger assignments in which students use the theory provided to tackle such problems. This way all the material can be covered, and skills can be practised, using examples from current research. Hence, within the compulsory profile courses biology, programming and modelling are not separated, but follow an integrated approach. This approach means students can reach the level of performing independent research by the end of the taught courses.

Once they have completed the first two months of compulsory courses (including the conversion classes), students have satisfied the prerequisites for all the profile courses. This means students can choose an individually optimal route through the curriculum. In addition to the profile courses, three optional courses need to be chosen. Students may choose to follow courses on specialist topics such as neuroscience, data mining, synthetic biology and artificial intelligence; alternatively students may choose courses from the 'other' profile. The study route is set out together with the student's mentor.

**[Note to the editor: Section below is a reference to the course content and should be formatted in a different way (for example a smaller font) to make clear it is not part of the main text]**

### **Compulsory Courses**

#### ***Introduction to Systems Biology***

The term "Systems Biology" has different interpretations among scientists, and the main purpose of this introductory course is to illustrate the different aspects of Systems Biology. We show that systems biology research can be roughly divided in "top-down" and "bottom-up" approaches, where top-down refers to the data-driven and bottom-up to the knowledge driven- modelling approach. We show that research in these different areas requires the application of different experimental and analytical techniques. We treat, for example, genome scale metabolic reconstruction and applications of genome scale metabolic models, qualitative reasoning and modelling approaches to understanding gene-regulatory and signalling networks, but also some of the basics of statistical analysis of high-dimensional data. The importance of the top-down approach seems self-evident, given the huge amount of biological data being generated in many biology laboratories. Most students have some knowledge of this type of approach, but the bottom-up approach is new to most students, and demonstrating its relevance requires special attention. Students read Yuri Lazebnik's essay "Can a biologist fix a radio" to understand the necessity of synthetic thinking in biology [6]. By participating in paper-reading and computer practical lessons the students also obtain more detailed and hands-on experience in the type of work of a systems biologist.

### Fundamentals of Bioinformatics

The main aim during the Fundamentals of Bioinformatics is to give students first-hand experience in some of the major open issues in bioinformatics research during the project (as already discussed above). This is supported by the lectures; these have the two-fold aim of giving a general introduction into the subjects (molecular biology and bioinformatics tools and algorithms), and supporting the practical work in the project. The last part of the course are the conversion classes, and the final grade of the course is the average of that for a written exam (testing knowledge from the lectures), the project work, and the conversion classes. Together this course gives students from a diverse array of backgrounds enough of an introduction to follow any of the three core bioinformatics courses in the master programme.

### Biosystems Data Analysis

This course is about the statistical analysis of data obtained from biological experiments. Many data analysis techniques are discussed, e.g. t-tests, ANOVA, Principal Component Analysis (PCA), Partial Least Squares (PLS), Linear Discriminant Analysis (LDA), Partial Least Squares Discriminant Analysis (PLSDA) and Self Organizing Maps (SOMs). After being introduced to different types of data coming from for example metabolomics or transcriptomics experiments, the student is taught when and how to use the different statistical tools. The students have hands on computer tutorials; the answers to these tutorials questions are graded. An important topic is judging the statistical validity of a result. Techniques that are used to do so are permutation testing, cross-validation and multiple testing corrections. This course should give the student a detailed understanding of statistical data analysis and its pitfalls.

## **Bioinformatics Profile**

### Algorithms in Sequence Analysis

This course provides the basic techniques of sequence analysis, including global, semi-global and local alignment by Dynamic Programming; the various BLAST flavours; hidden Markov -based homology search engines; motif searching methods; and phylogenetic methods. These methods are also discussed and embedded in genomics strategies such as comparative genomics, phylogenetic footprinting, and phylogenetic shadowing. The course assignments include a fair amount of programming in the Python language, since students need to design a dynamic programming algorithm for global, semi-global and local alignment, as well as a hidden Markov-model for protein domain prediction from sequence, including the Viterbi, Forward and Forward/Backward (Baum-Welch) algorithms. The course is aimed at providing a thorough understanding and appreciation of algorithmic design, as well as knowledge as to how to apply these algorithms in sequence analysis.

### Structural Bioinformatics

Formerly named 'DNA/Protein structure-function analysis and prediction' introduces students to a number of techniques and algorithms for the analysis and the prediction of molecular structure and structure-function relations, mainly in proteins. It starts with a very short introduction into basics of protein structure, but quickly dives into structure prediction, protein folding, molecular simulation, and even includes some basics of statistical thermodynamics. Like the other courses, a series of practical assignments ensures students acquire practical skills that they may apply in later project work.

### *Bioinformatics for Translational Medicine*

This is a theoretical and practical Bioinformatics course on computational methods for Translational Medicine. We focus on Bioinformatics algorithms that predict the clinical outcome for patients and methods for analysis that obtain a deeper understanding of complex diseases. These methods and algorithms may use data from various high-throughput experiments such as proteomics, microarrays and next-generation sequencing, and may add external information from biological databases.

One large assignment within this course is a competition between students, in which they have to build a classification method that can discern different types of breast cancers based on high throughput biological profiles. Students have to predict how well their methods will perform and write a paper about their classification strategy, validation protocol and results.

For this course we invite several experts in the field to speak about their research. This also gives students a chance to find suitable internship places for their research projects in the second year of the MSc.

## **Systems Biology Profile**

### *Systems Biology in Practice*

In this course students perform hands on experiments on a yeast culture. The physiology of the cells is quantified measuring fluxes as a function of time after a disturbance. Also the proteome and the transcriptome of the cells are measured by the students. All the data of the different students is shared and using the complete data set students make an assessment how the organism is trying to adapt after the disturbance. The students present, compare and discuss their findings in a final session.

### *Basic models of biological networks and Advanced modelling in Systems Biology*

In these two courses concepts of mathematical modelling in biology are explained. Basic Models of biological networks is a compulsory course for the Systems Biology profile. The focus in the basic models course is on biological networks, such as metabolic networks, signal transduction pathways and transcriptional regulation networks. Students learn how to model biological systems using basic mathematical techniques, mostly by solving for the steady state or equilibrium condition of systems of differential equations. In the advanced course model validation (parameter estimation, experimental design), dynamic behaviour (instability, multistability), robustness & sensitivity analysis, metabolic control analysis and stochastic modelling are treated. In both courses we show students how to obtain analytical results first, and teach them how to create and execute a model in Mathematica[7] later. Another important learning objective in both courses is to obtain a feeling for biological dimensions and numbers.

### *Statistics with R*

In this course we teach students how to use the statistical calculation environment R to solve real world problems in biology. Proficiency in a language like R is essential for most systems biologists. Analysis of complex data is applicable in almost all fields of biology nowadays. This course gives the students a firm basis and a head start in their internships, and later industry- or academic career.

## **After the taught courses**

### **Internships**

After a year of taught courses, students go on to do independent research projects in their second year. Typically they do one larger (major) project, and one smaller (minor) project. At this stage students are free to follow their own interest, with the restriction that the major project needs to lie within the field of their chosen profile, e.g. Bioinformatics or Systems Biology. For any bioinformatics project both a programming component and a molecular biology component is required, similarly for systems biology a (mathematical) modelling component and a molecular biology component is required.

Students can be very adventurous in choosing their projects, and many of our students opt to do one or both their projects abroad. Internships may range from biophysical simulations on proteins at the University of Cambridge (UK), to Cancer Profiling at NCBI (US), to neuroinformatics at LONI(US), to miRNA analysis at Riken (Japan).

Other students choose to take up a research project in a company to be able to make a more balanced choice between academia and industry. Projects related to medical and clinical data are also very popular: students work for example on a next generation sequencing project in an academic hospital, or a meta genomics project in the Dentistry department in Amsterdam (ACTA). Usually it is easy for students to find an internship place, especially with the fast developments in high-throughput profiling for medical and biological applications.

### **Research Proposal**

As a compulsory subject during the master, students write a research proposal, applying for one million EURO maximally, which implies the project proposed will involve more than a single researcher. This ambitious upper boundary is set to stimulate 'out-of-the-box' thinking and forcing students, in addition to scientific issues, to really think about how to plan the research, how to budget it, and how to separate tasks over various levels of competence. This course is tutored in a one-to-one fashion and can be taken flexibly over time, allowing students to write their proposal in a period suitable to them. The first step for students is to select a research subject for their proposal, for which they typically need to consult their tutor one to a few times. Then, some suitable literature is handed out, after which students make a first layout of their proposal. After a mandatory check by the tutor, the student then embarks on writing the full proposal, during which time the tutor can be consulted again. The course is concluded by a 10-minute mock defence of the proposal, where a student needs to give a concise overview of the research, which should work as a 'sales pitch' to persuade the mock committee, comprised of tutors and PhD students. During or even before starting to work on their proposal, students have to attend three scientific seminars and write a half-page commentary for each, describing

their personal opinion about these seminars and to what extent they were relevant to their proposal. The main idea is to mimic a scientific setting, where students can select their materials and assess issues in a way that is not typically happening during the other courses. To maximise this effect, students are advised to take this course during or after their internships, which also makes selecting a research topic more natural.

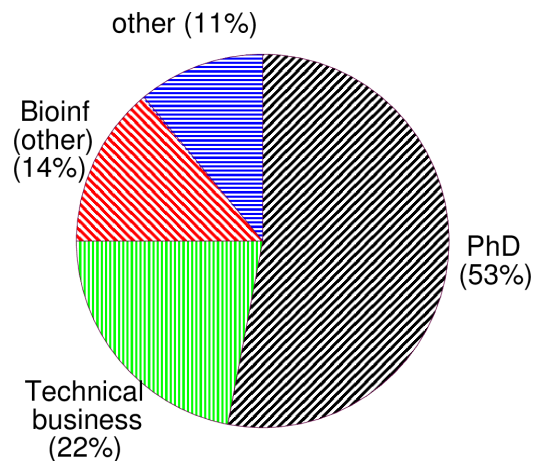
## **iGEM**

The internships of master students is usually limited to doing research only, although the tasks of a scientist in academia or industry often encompasses much more than this. The iGEM (international Genetically Engineered Machines) competition is a perfect opportunity for students to train other essential skills too, like raising money, presenting research to the general audience in presentations, expositions and on the internet, and cooperating in a small group. Since two years we participate in the iGEM competition, and have also organized the European qualifying rounds in Amsterdam. A small group of master students from different disciplines participates in the iGEM competition by working for a couple of months on a subject. To qualify for the competition certain results, strains and sequences have to be submitted to the registry of standard biological parts. The participation in the iGEM competition is acknowledged as a minor internship, and students write an individual report about their contribution to the research part of the work. Because most students are relatively inexperienced in the laboratory, the supervision of a group of students is very intensive. We assign daily supervision to one or two post-doc group members and staff members have advisory roles. Five universities in the Netherlands participate in iGEM, and because of the relatively small travelling distances within this country, iGEM groups also organize meetings to exchange experiences. The high time-investment in this type of education is an important point of consideration, but participation does provide a unique and important experience.

## **Job Prospects**



## Position after MSc



**Figure 3:** Job position of students after completing the master Bioinformatics and Systems Biology divided into PhD (research), non-research Bioinformatics (e.g., technical), business (most often data analysis and/or IT) or other/unknown.

With either strong mathematical or programming skills, and the ability to work with people from a wide variety of backgrounds, our students do not have much trouble finding suitable employment. More than half of the students go on to do a PhD (Figure 3); surprisingly, this fraction holds also for students with a vocational background. Topics of PhD projects have a very broad range, just like the internship projects, including theoretical studies using mathematical modelling, combinations of experiment with data analysis and the development of new algorithms.

Systems biology and bioinformatics are relatively new fields of research. We have observed in the past that highly educated students are in higher demand. E.g. students with a finished PhD in Bioinformatics were in higher demand than those with only an MSc, and in much higher demand than those with vocational BSc education. Currently, we see that the job market is also opening for BSc Bioinformatics education due to the maturation of the field and the higher demand in the medical world for computational data analysis of high-throughput data.

Students who do not go on to do a PhD typically do very well, too. Students may end up in a bioinformatics related job (~ 15%) or work for or set up a technical business company (~ 20%), e.g. finance, pensions, consultancy or IT. Here again the programming, large scale data analysis and ability to work in an interdisciplinary team are highly valued skills.

Interestingly, some students make a complete “switch”: we have seen a student with a Computer Science BSc following a Lab course and becoming so enthusiastic that he is now looking for a job that includes experimental biology. Similarly we have had a couple of students with an MSc in Biology who have become successful in IT jobs with a heavy amount of software development but involving no biology.

Lastly, it is important to point out our aim is not to create a homogeneous set of graduates out of

the heterogeneous set of incoming students. For example, students with a background in computer science will typically remain stronger at method development, while students with a biomedical background will usually have a deeper understanding of the clinical problem that needs to be solved, even after finishing the master. In our experience, such diversity is in fact beneficial - both such graduates are in high demand.

## **Learning points**

In summary, there are several important conclusions for developing a course with heterogeneous student backgrounds.

From a student's perspective, working and studying with peers from different BSc backgrounds is not only an invaluable experience, but also highly regarded in every field of employment. To create such a learning environment, we mix students with different backgrounds in groups early on. In these groups they tackle complicated research tasks and discover gaps in their own knowledge. Due to early mixing, students from different backgrounds tend to support each other throughout the year.

An intensive start to the course is important for setting an exemplary pace and for establishing basic knowledge. In our programme students have 24 contact hours per week in their first two months, with 40 hours of total workload per week. During these first two months we teach Mathematics, Biology and Programming classes to allow students to fill gaps in their knowledge.

After a compulsory start to a programme embracing multiple research fields, students can take a flexible route through the remaining curriculum according to their own interests. Having completed the first two months of our programme, we give students a choice between a Bioinformatics and Systems Biology profile; excellent students have the chance to choose both profiles. Choosing a profile, it is important that students can opt for the discipline that matches their natural talents.

An interdisciplinary subject needs to be taught through research examples in order to obtain high academic standards. During all courses and project work our focus lies on the true interdisciplinarity of bioinformatics and systems biology. This is achieved by setting the same research tasks for all students, which allow them to further practise skills such as programming, mathematics or biological interpretation. This style of teaching gives useful transferable skills to the students, making them fit for the job market.

## **Key Points**

- Mixing students from different background provides a good learning environment at MSc level
- Conversion classes are used to let students tackle their deficiencies,
- In the taught courses we focus on interdisciplinary research within systems biology or bioinformatics, to obtain skills at MSc level.
- An intensive start provides a good learning basis sets expectations for the amount of work necessary in the MSc programme

## **Acknowledgements**

We would like to thank Punto Bawono, Frank Bruggeman, Nicola Bonzanni, Mohammed El-Kebir, Klaas Hellingwerf, Hannes Hettling, Ali May, Erik van Dijk, Bart van Houte, Jens Kleinjung, Remco Kort, Pinar Ozturk, Joost Teixeira De Mattos and Meike Wortel for their part in setting up the curriculum for the Bioinformatics and Systems Biology MSc in Amsterdam. We also would like to thank our education committee, in particular Rob van Spanning, and also the various student members who have been instrumental in shaping up the curriculum of our MSc programme throughout the years. This work is partly financed by the Netherlands Organisation for Scientific Research (NWO) and by the Center for Translational Molecular Medicine (**TraIT**).

## **References**

- 1 Murayama, K., Pekrun, R., Lichtenfeld, S. et al. (2012). Predicting Long-Term Growth in Students' Mathematics Achievement: The Unique Contributions of Motivation and Cognitive Strategies. *Child Development*.
- 2 Altschul, S. F., Madden, T. L., Schäffer, A. A., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402.
- 3 Murzin, A. G., Brenner, S. E., Hubbard, T., et al. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *Journal of molecular biology*, 247(4), 536-540.
- 4 Camon, E., Magrane, M., Barrell, D., et al. (2003). The gene ontology annotation (GOA) project: implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome research*, 13(4), 662-672.
- 5 Finn, R. D., Mistry, J., Tate, J., et al (2010). The Pfam protein families database. *Nucleic acids research*, 38(suppl 1), D211-D222.
- 6 Lazebnik, Y. (2002). Can a biologist fix a radio?--Or, what I learned while studying apoptosis. *Cancer cell*, 2(3), 179
- 7 Wolfram Research, Inc., Mathematica, Version 8.0, Champaign, IL (2010).