The Computational Biology Program

THE HEBREW UNIVERSITY OF JERUSALEM

Report of the Review Committee

5-8 December 2005

Presented to Prof. Haim D. Rabinowitch, Rector

TABLE OF CONTENTS

EXECUTIVE SUMMARY

REVIEW COMMITTEE
The Academic Program in Computational Biology
Executive Summary

Thursday, December 8, 2005

The academic program in Computational Biology here called "the Program," was established in 2001 following an initial phase as a double major beginning 1999. Following the decision of the Rector and the Standing Committee of the Hebrew University, a Review Committee was called to review the Program and assess its success. The members of the Review Committee were: Prof. Ron Shamir (Tel Aviv University, Computer Science); Prof. Jeremy Smith (Heidelberg University, Computational Molecular Biophysics); Prof. Joel Sussman (the Weizmann Institute, Dept. of Structural Biology); Prof. Michael Waterman (Chair) (University of Southern California, Biological Sciences, Math and Computer Science). The Committee convened for four intensive days (5-8 Dec.), read and discussed material prepared by the Program Directors and other faculty involved. They interviewed the Program Directors, the Program's academic committee, the Dean of the Faculty of Science and the chairmen of the relevant departments. The Committee also met separately with several groups of students (undergraduates, graduates and Ph. D) of the Program, and faculty members of CS and LS.

Due to the explosive growth in quantitative biological sciences in recent decades there is a pressing and sustained need for highly-trained people to manage and analyze the large amounts of resulting data. To meet this need the Program Directors created an innovative, highly-successful elite program. The Program, which has particularly stringent entry requirements, attracts some of the brightest students to the Hebrew University. In addition to constructing a strong and coherent interdisciplinary program of courses the Program Directors have introduced an innovative research component as exemplified by the research seminar, retreats and individual projects, some of which have even resulted in high-quality research publications. The Program leads to a full CS degree allowing the graduates to successfully compete in the high tech industry or to go on to a higher degree. An essential component of the Program's success is the close personal student guidance provided by the Directors and Program staff. The selective character and challenging intellectual content has created an impressive esprit de corps among the students.

Further evidence of success is that the majority of the students that completed the undergraduate Program have continued on to graduate studies in various related fields.

While a fraction of the students who begin the Program do not complete it, most of these transfer early and successfully to other programs without significant delay in their academic progress. The Masters Program has attracted only a small number (4) of these students, a fact primarily due to the structure of graduate student support at Hebrew University.

The committee congratulates the Program Directors for their tireless efforts in voluntarily building up this critical elite Program with very few additional or dedicated resources while simultaneously carrying out their normal duties. Hebrew University can take pride in this unique world-class Program which is one of the two

primer undergraduate bioinformatics Programs in Israel. There is for example no such Program on Germany. The Program has now proven its worth and requires some specific measures to be undertaken to ensure its stability and sustainability. The most important of these measures are listed below:

Recommendations:

1) **Teaching credit.** Teaching by faculty members in this Program should be credited fully to their regular teaching load. A faculty member should receive full teaching credit whether the Program course is listed in their own or another department.

2) **Hiring new faculty**. At least three new faculty should be recruited from areas which are complementary to the interests of the present faculty. For example: Computational Genomics, Computational and Mathematical Systems Biology, Structural Bioinformatics, or Computational Genetics.

3) **Teaching Assistants.** The Program requires the allocation of four new Teaching Assistants dedicated to the Program, one to work in the Computational Bioskills Workshop and three to help in supervision of research projects. Allocation of these Teaching Assistants should be independent of the goodwill of the Departments.

4) **Masters Programs.** Recommend integration of the various Masters programs in Quantitative Biology (e.g. this Program (Computational Biology), Bioinformatics and Genomics, Proteomics) as separate tracks under a single umbrella. This will make the Programs easier to administrate, more attractive to students, and likely will reduce costs.

5) **Financial Support.** Secure dedicated fellowships for the top students in the Master level section of the program. Maintaining the student fellowships as well as special activities such as retreats is critical to the continued success of the program.

6) **Program duration.** Given the density of the program students should be given the option to extend the B Sc into the fourth year with minimal additional tuition cost.

7) **Program Content.** Some specific suggestions for streamlining and improving the program content are made in the full report.

THE COMMITTEE'S REPORT

Introduction:

The Review Committee was convened to assess the B.Sc./M.Sc. Academic Program in Computer Sciences and Life Sciences at the Hebrew University of Jerusalem (hereinafter referred to as "The Program"). The Program has been running for six years and has produced its first graduates, and so the review is now timely.

Computational Biology is one of the most rapidly-growing fields in science. From an initial explosion of this field in both academia and industry about 6-7 years ago, the field has now matured and the transformation of the biological and medical sciences towards quantification is now steady and persistent. The Program Directors, who themselves carry out first-class research in quantitative biology/computational biosciences, recognized early on the need to train young people in computational biology so as to meet the need for this transformation.

The initial announcement of the undergraduate Program six years ago was met with an enthusiastic response from the applicants, such that the Program Directors were able to set, and have maintained, particularly high standards for entrance. Thus, the program quickly became "elite" and the articulate, thoughtful and intelligent students that the committee met left an impression of a very high academic level.

The students clearly have to work very hard in the first year, often staying at their computers until the early hours, and this initial workload is something of a shock for some of them. Maybe partly due to this, but more likely due to their own rapid intellectual evolution, some students leave the program early and switch to other, related programs, such as single-major Computer Science or Life Sciences. Few of these students, however, lose much time when they make such a switch. Those remaining in the Program graduate with a valuable, integrated, research-oriented interdisciplinary degree.

Program Support:

This elite Program depends critically on support in several forms. We have been given the impression that the support is not uniform with all the elite interdisciplinary Programs at Hebrew University, and it seems to us this situation should be regularized.

The program benefits from the Stein Donation at $50,000 per year. This allows up to 15 students to receive $2000 tuition and a $1000 stipend. This provides tuition support for students not meeting the University threshold for tuition support, and it is excellence based so that weaker students do not receive the support. Students join the Program for several reasons including the LS and CS themes and the perceived difficulty of the Program (they wish to try the most challenging Program!), but clearly the Fellowships are critical for many of their decisions. The heavy load during the first year is more easily handled with this Fellowship partially freeing some of them from working in addition to taking all the courses. In addition a small portion of the Stein money is used for supporting students to attend conferences and symposia. Were this support to disappear the Program would be quite adversely affected.

One of the important aspects of the program is undergraduate research, and the third year students attend a retreat near the end of the year where each research team presents their work. The best presentations receive an award and are asked to present their work later on the Hebrew University campus. These activities are important to training the students to be active researchers. The activity was supported by a small grant from the Ministry of Science and Technology which is not granted specifically for this purpose.

There is institutional support based in Computer Science for Secretarial and Technical Staff. There are many administrative and organizational details this person is crucial for. Many of the students require or ask for changes in their course of study and after approval by the faculty the secretary assists with the paperwork. The technical staff helps with the hardware and software in the programming and project courses. Clearly this assistance is essential to the program.

Finally we come to a key area of the Program support, that of faculty teaching support. The Program as designed uses s many existing classes which are already being taught but obviously some new classes are necessary to make the Program a success. The non Computer Science faculty receives no regular teaching credits for Program teaching such as Course 67305, Seminar, and Course 76554, Computational Methods in Molecular Biology Laboratory. This issue is becoming even more critical as the Masters Program develops.

A heavy administrative and instructional burden has fallen on three key faculty, Profs. Michal Linial, Nir Friedman and Hana Margalit. It is essential for this important area of science that new faculty is added to diversify the intellectual content and share the load. At least three new faculty should be recruited from areas which are complementary to the interests of the present faculty. For example: Computational Genomics, Computational and Mathematical Systems biology, Structural Bioinformatics, or Computational Genetics.

The Undergraduate Curriculum:

The B.Sc. Program combines a full degree in Computer Science with a substantial fraction of the Life Sciences degree, plus additional courses that focus on Computational Biology with a research emphasis. As a result, the first three semesters are devoted fully to introductory courses (Mathematics, Chemistry, Physics and some CS and LS). The first encounter of the students with Bioinformatics is at the middle of the second year in the workshop in Computational Bioskills and in the Research Seminar. The third year consists of dedicated courses in Bioinformatics, electives in LS and CS, a "wet" lab and a year-long research project.

Overall the teaching program is excellent. The decision to provide the students with a full degree in CS is a good one, as it provides the students with viable employment alternatives in the hi-tech industry upon graduation if they do not opt for graduate school. The need to combine almost two full majors into three years makes the Program extremely intense, especially in the first year, and some hard choices must be made and some good courses are given up. An additional difficulty is that some LS courses are given at a slow pace compared to the abilities of the Program students. Below we discuss suggestions for improvements to the Program courses. Most of our recommendations refer to the introductory courses.

Physics. Currently the Program contains 10 units of Physics on the second year (77163 and 77164) taken jointly with the LS students. Most of the material overlaps the honors Physics high school curriculum. We recommend to exempt those students who passed the 5 units Physics matriculation exam with high grade from these two courses. We also recommend the Program Directors to reconsider the Physics content that is taught, for example to condense and intensify the curriculum.

Probability and Statistics. Currently the students can take either Probability Theory 1 from the Mathematics curriculum (80420) or a two course series Introduction to Probability and Statistics 1 and 2 (52114 and 52115), from the Department of Statistics directed to Mathematics and CS students. Program students who take the second option get only 4 credit units for the 8 credit series. Students who take the first choice are not exposed to Statistics at all. We recommend making the two course series mandatory in order to guarantee that the students get adequate knowledge in fundamentals of Statistics, and giving students full 8 credits for it. Note that this suggestion increases the total official load of the Program.

Chemistry. Students should take General Chemistry (69174), Organic Chemistry (69166) and Physical Chemistry (69123). All are large courses (5 credits each) taught at a relatively slow pace for general LS students. We have heard complaints from Program students that the courses can be taught much more efficiently at a fraction of the time if a dedicated course attuned to the level of the Program students is formed. In fact, on the initiative of the Program Directors, a condensed course replacing General Chemistry was given very successfully in the past for a while, but was closed due to the objection of the Chemistry school. We recommend that such course be generated and given.

Computer architecture. The students are currently required to choose one course from among three courses on computer architecture (67160, 67306, 67925). Since some of our other suggestions may increase the total load of the Program, and since we strongly believe that the total load should NOT be increased, we propose to remove the requirement of an architecture course from the Program. We believe that of the CS courses, this is the least relevant for the Computational Biology emphasis of the Program.

Specialized Biology courses. Some of the students complained about the slow pace of the introductory Biology courses. In fact, the total effort put into LS courses in the first year was described by some students to be one tenth of the effort put into CS courses. Other students found the basic biology courses unchallenging intellectually and taught at a very descriptive and non-analytic level with too much emphasis on details rather than principles. We recommend that the option of creating a revised version of one of the biology courses be considered. Providing such a course to students in several elite interdisciplinary programs may make this economically more viable.

Teaching assistants. Most of the courses in the Program have natural "homes", and teaching assistantships are provided by CS or LS. For some dedicated Program courses both units are reluctant to provide TA-ships. This applies, for example, to the Workshop in Computational Bioskills (76552) and to the Project (76554), a key module in the Program. We view the need to provide dedicated TA-ships for these courses essential to the continued success of the Program. Specifically, we recommend that 4 TA-ships are allocated to the Program, one for the Workshop and three for the year-long Project. If additional ad-hoc courses are initiated, they should be provided their own additional TA-ships.

The first year load. The first year is extremely intense. Students in the second year complained bitterly about the heavy load on the first year, and about the fact that they still did not know where they were going. More mature students on the third year and in graduate school viewed the first year more positively and could appreciate the need for all the introductory courses in retrospect. Some of our recommendations above can reduce the number of hours on the first year and alleviate part of the problem. Other ideas that could help students to survive the first year is to create opportunities for 1^{st} year students to meet 2^{nd} and 3^{rd} year students, and to get exposed to some research in bioinformatics. These together should hopefully reduce some of the first year stress and decrease the number of students who drop out of the program.

The overall load. Students who elect to extend the B.Sc. beyond three years are currently required to pay very high levels of extra tuition. Given the density of the Program, students should be given the option to extend the B. Sc. into the fourth year with minimal additional tuition cost.

Research Component:

The research component is an integral and very strong part of the Undergraduate Program. There are three major parts, i.e. 2nd Year Research Seminars, 3rd year Research Project and Scientific Retreats:

2nd Year Research Seminars: The course "Research Methods in Computational Biology" is a required 2 credit course in the 2nd semester of the 2nd year. The aim of the seminars is to expose students to cutting edge research in computational biology, high-throughput biology and systems biology. The list of papers is selected from fundamental yet technically simple results in top journals. Each meeting is focused on several (2-4) related papers. The meetings include presentations by the students assigned to each paper and a general discussion of the methodology, the analysis, the results and the implications. The emphasis is to provide students with ability to critically read scientific literature, to present it in front of an audience, and to critically evaluate research results.

The students and the faculty both felt that this was a superb course, and the students, in particular, felt it was their first real introduction as to how actual research is done, by forcing them to critically read the scientific literature.

One might consider giving an additional advanced version of this seminar during the 1st semester of the 3rd year. A number of the students felt that this would be very worthwhile.

Research projects: During the 3rd year of the Program the students carry out a large-scale research project under the supervision of a faculty member in the LS or CS. The project is done either individually or in pairs. The students receive 8 credit points for this course ("Computational Methods in Molecular Biology – Lab"). The choice of the project is made from a list of suggested projects drawn up by the Program Directors after consulting with the members of the LS and CS Faculty. The students spend a minimum of 1 day a week doing research and, depending on the lab, some of the students become integral members of the faculty research teams. Examples of the projects are listed in the material given to us.

During the year the students pose a scientific question (often one relating to ongoing research effort at their host lab), devise methodology and tools to address it, apply these tools to data, and analyze the results. It is important that the course staff meet regularly with each team to assess progress and provide guidance.

At the end of the year, the students present their work at a special event open to the faculty members of CS and LS and students, and submit a written report. It should be stressed that a number of scientific papers, published in top journals (like PNAS, JBC, etc.) have come out of these research projects.

The faculty and students both rated this as possibly the most significant course in the Program, as it gave the students a real opportunity to learn and experience how scientific research is actually done.

Annual Scientific Retreat. There is an annual two-day scientific retreat, off-campus, where the students present the results of their work. The meeting is attended by all 3rd-year students, who present their work; as well as 2nd-year students who all attend. In addition, an outside speaker is invited to present some exciting scientific developments in related areas.

The students are asked to vote for the best presentations. The winners are asked to give their talk shortly after the retreat at the HUJI campus. This event is attended by many members of the faculty with all the students in the program presenting posters of their research projects.

These events clearly give the undergraduate students the feel of what it is like to participate in a scientific meeting. One of us (JLS) participated in one such retreat at Sde Boker and was extraordinarily impressed with the quality of research presentations and the integration of the CS and LS students' research. It really felt like attending a high-level scientific meeting.

The Masters Program:

Due to the full Computer Science major component, and their own intelligence, the graduating B. Sc. students are well suited for well-paid jobs in the high-tech industry, and many take this option. The majority, however, remain to pursue a Masters degree. Superficial examination of the statistics of this would suggest that an original objective of the program, to provide a 5-year continuous, integrated B. Sc./M.Sc. education, has not been met, as few of the bachelor graduates have continued on to the Computational Biology M. Sc. However, this is deceptive, and not at all worrisome, as most students do indeed pursue masters study in the field of computational biology, but happen to be registered in other Masters programs.

There exists in the university a plethora of Masters programs in closely-related subjects (e.g. Bioinformatics and Genomics, Computational Biology, Proteomics etc). While the committee does not see this as disadvantageous, as it provides, in principle, flexibility of choice, the choice itself of which masters program to take appears to often have been made for non-academic reasons. That is to say, a student will register for a particular masters program because this is the only way he/she can get a Teaching Assistantship or a fellowship, or because the supervisor happens to be a member of a particular faculty, and so on. This has the detrimental effect that a student can end up taking Masters courses not best corresponding to his/her academic and research interests. We thus recommend that an interfaculty umbrella-structure be set up that administers all Masters programs in fields related to quantitative biology. The faculties of Computer Science/Engineering, Medicine, Chemistry, Physics and Life Sciences should all be associated with this. The umbrella structure would simplify the recruitment, visibility and organization of the students, and would be potentially cost-cutting. Most important, the structure should remove all interfaculty barriers, such that students performing a project with any member of the associated faculties should be able to register for any Masters program. All students in the various Masters programs should be treated equivalently with regards to fellowships/teaching assistantships etc. irrespective as to which faculty they are formally registered in.

Relationship of Program to Physical and Chemical Sciences. Computational Biochemistry and Biophysics are arguably important fields in Computational Biology. However, the optimal training for these fields requires investment in Chemistry and Physics courses well beyond that offered by the present Computational Biology B. Sc. Program. One option for the directors might be to add elective courses in chemistry and physics for students with these interests. However, this is likely to be cumbersome and infeasible, in which case Computational Biochemistry and Biophysics is recommended to be a significant part of the planned Chemistry and Biology program and, failing that or in addition, should be offered as an option in the existing Chemistry with Physics program. More generally, the university would be taking a major risk if it neglects interdisciplinary education bridging the physical and biological sciences.

Prof. Michael Waterman,
University of Southern California _____
(Chair)

Prof. Prof. Ron Shamir,
Tel Aviv University _____

Prof. Prof. Jeremy Smith,
Heidelberg University _____

Prof. Joel Sussman,
the Weizmann Institute _____

APPENDICES

APPENDIX A

RESUMES OF VISITING COMMITTEE MEMBERS

1. Prof. Michael Waterman, University of Southern California (Chair)

2. Prof. Prof. Ron Shamir, Tel Aviv University

3. Prof. Prof. Jeremy Smith, Heidelberg University

4. Prof. Joel Sussman, the Weizmann Institute

APPENDIX B

List of People Who Met With the Committee

The Rector, Prof. Haim Rabinowitch
The Vice-Rector, Prof. Sarah Stroumsa

Head of Academic Review for the Sciences, Prof. Eliahu Friedman

Directors of the Program, Prof. Michal Linial and Prof. Nir Friedman

Dean: Prof. Hermona Soreq, Dean of the Faculty of Sciences

Members of the Academic Committee:
- Prof. Michal Linial (Biological Chemistry, Life Sciences)
- Prof. Nir Friedman (Computer Sciences)
- Prof. Hannah Margalit (Molecular Genetics and Biotechnology)
- Prof. Nati Linial (Computer Sciences)

Chairmen of the Relevant Departments:
- Prof. Michael Ben-Or
  (Head of School of Engineering and Computer Sciences)
- Prof. Noam Nissan (School of Engineering and Computer Sciences)
- Prof. Joseph Hirshberg (Head of Department of Life Sciences)

Junior Faculty:
- Dr. Natali Balaban (Physics)
- Dr. Raz Kuperman (Mathematics)
- Dr. Ariel Darvasi (Ecology, Systematics and Evolution)
- Dr. Ora Furman (Bioinformatics)

Other Faculty:
- Prof. Ruth Sperling (Genetics)
- Dr. Itamar Simon (Molecular Biology)
- Prof. Uzi Motro (Life Sciences & Statistics)

1st, 2nd and 3rd year Students in the Program
Switch-outs of the Program
Graduates of the Program
PhD Students