PROPOSAL TO ESTABLISH
A GRADUATE FIELD AND PH.D. PROGRAM
IN
COMPUTATIONAL BIOLOGY

Contact information:

David Shalloway

Professor, Molecular Biology and Genetics
Chair, Faculty in Computing and Information Science Subcommittee on Computational Biology

e-mail: dis2@cornell.edu
phone: (607)254-4896
267 Biotechnology Building

Contents

1. The Field of Computational Biology

1.1 Field of Study, Subjects, and Concentrations

This is a proposal to create a graduate field of Computational Biology and to offer a Ph.D. degree. This is an interdisciplinary field that should attract a unique group of students desiring interdisciplinary training.

The graduate field of Computational Biology will have seven subjects: computational biology, computational genetics, computational macromolecular biology, computational cell biology, computational organismal biology, computational behavioral biology, and computational ecology. Each subject has a concentration of the same name.

The Graduate School Catalog entry is as follows:

COMPUTATIONAL BIOLOGY [Ph.D.]
Computational biology
Computational biology
Computational genetics
Computational genetics
Computational macromolecular biology
Computational macromolecular biology
Computational cell biology
Computational cell biology
Computational organismal biology
Computational organismal biology
Computational behavioral biology
Computational behavioral biology
Computational ecology
Computational ecology

1.2 Description for the Graduate School Catalog

Computation has become essential to biological research. Genomic databases, protein databanks, MRI images of the human brain, and remote sensing data on landscapes contain unprecedented amounts of detailed information that are transforming almost all of biology. The computational biologist must have skills in mathematics and computation as well as in biology. A key goal in training is to develop the ability to relate biological processes to computational models.

The field provides interdisciplinary training and research opportunities in a range of subareas of computational biology involving topics such as DNA and protein databases, protein structure and function, computational neuroscience, biomechanics, population genetics, and management of natural and agricultural systems.

Students majoring in computational biology are expected to obtain a broad, interdisciplinary knowledge of fundamental principles in biology, computational science, and mathematics. But because the field covers a wide range of areas, it would be unrealistic to expect a student to master each facet in detail. Instead, students choose from specific subareas of study: they are expected to develop competence in at least one specific subdomain of biology (i.e., genetics, macromolecular biology, cellular biology, organismal biology, behavioral biology or ecology) and in relevant subareas of computational science and mathematics.

Students are supervised by field faculty drawn from sixteen departments.

1.3 The Computational Biology Program

Computational Biology (CB) at Cornell is an interdisciplinary program that links the computational and mathematical sciences with the life sciences.
This is particularly appropriate at Cornell given its traditional strengths in the life sciences and the fact that a substantial fraction of its faculty is engaged in life science research. Cornell's computer and information sciences department is recognized among the leaders in the field and it has always been active in forming collaborations across campus.

Quantitative prediction and interpretation are becoming increasingly essential components of biology and other fields. Complex patterns, structures and interactions raise fundamental and fascinating questions that can only be addressed using formal mathematical, statistical, and computational methods. The wealth of data being acquired to address these questions requires new and substantive quantitative approaches to make possible appropriate analysis and interpretation. The unprecedented level of computational power now available to researchers provides the means for increasingly sophisticated analyses of highly complex systems arising in the biological sciences.

Many biologists are faced with the task of analyzing and modeling data collected in the lab, in the field, or *in silico*
that requires sophisticated mathematical and computational analysis. Traditional statistical methods are limited in their scope and are proving inadequate for the complex models and massive data sets that are now becoming commonplace. The solutions to these large-scale problems often lie at the interface of mathematics, computer science and statistics and a new type of scientist and academic who can work and communicate across the inter-disciplinary barriers is needed.

It is presently very difficult for a graduate student to find a unified and coherent program for this type of training in CB at Cornell because of Cornell's unique dispersion of biology and life sciences among several different colleges and many departments.
Research in biomedical and life sciences is being invigorated by the influx of researchers from the physical sciences and mathematics.
A formal program in computational biology to train Ph.D.s would permit students who have strong preparation in mathematics, statistics, or scientific fields with firm quantitative backgrounds, such as physics, chemistry, or engineering, to have a direct entry into CB. Currently, such students must attach themselves to Ph.D. programs in life sciences that may have prerequisites and requirements skew to the students' interests, leading to pedagogical inefficiency.
As will be pointed out below, several highly respected and distinguished universities already have Ph.D. granting programs in CB.

We describe the scope of the field
according to the "NIH Working Definition" of Computational Biology (http://www.bisti.nih.gov/CompuBioDef.pdf), except that social systems are not included. For our purposes "Computational Biology" is the development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological systems

An online search of Peterson's Graduate Programs under "computational biology" yields 471 listings. As is currently the case at Cornell, most of these reflect areas of specialization within standard disciplinary programs (e.g., Computer Science).
However, in the last few years a few dozen specific CB programs have been established in the U.S. Examples include Computational Biology (U Penn), Computational and Genomic Biology (UC Berkeley), Computational Molecular Biology and Bioinformatics (USC), Bioinformatics and Computational Biology (Iowa State), and Bioinformatics and Computational Biology (UCSD). (See also Sec. 2.2, below.) This reflects the importance and rapid growth of the area and the need to provide focused training.

Cornell's CB program will have a distinct academic focus and provide critical centralization and resource deployment at a university where redundancy and lack of centralization in key areas presents problems for graduate students in areas at the interface between traditional disciplines. Such is certainly the case for CB. At the same time, it will provide an academic "umbrella" for a variety of more specialized areas of concentration in computational biology that already have critical mass (e.g., computational functional/structural genomics and computational population genetics). This umbrella will also be critical for the development of other important, yet currently weak, areas (e.g., bioinformatics and computational systems biology).
In the absence of a unifying field, these specialized programs must independently develop *ad hoc* methods of administering students within the existing academic fields. In the long run, the centralizing influence of the field of CB should serve to reduce overall administrative complexity.

1.4 Related Fields and Departments

As can be seen from Appendix 1, the Computational Biology faculty come from 16 departments in a spectrum of endowed and contract colleges, including the Weill Cornell Medical College. As might be expected, these faculty also represent the related fields such as Computer Science, Biophysics, Applied Mathematics, and Biometry.

The general consensus is that the establishment of this new field will provide helpful synergism with the existing fields. No conflicts are anticipated.

Letters of support from the Directors of Graduate Study of these fields are in Appendix 2. In addition, letters of support from the chairs of some of the most relevant departments (Biological Statistics and Computational Biology, Computer Science, and Molecular Biology and Genetics) are included. Existing courses in these (and other) departments will be used to fulfill the academic course requirements of the program.

The intellectual breadth of this field provides a significant opportunity to strengthen the working ties between Cornell (Ithaca) and Cornell Weill.

We have already observed cases where graduate students in the physical sciences at Cornell (Ithaca) have moved to New York City to conduct thesis research in computational biology with Cornell Weill faculty who will be members of the proposed field.

The most important connection may be the Tri-Institutional Training Program in Computational Biology & Medicine, which links Cornell, Cornell Weill, Rockefeller, and the Memorial Sloan-Kettering Cancer Center (http://www.cs.cornell.edu/grad/cbm).

Students enrolled in this tri-institutional program (currently 6 per year) are required to train both in Ithaca and NYC during their first two years (see attached "CBM Timetable of Admissions and Training" or http://www.cs.cornell.edu/grad/cbm/student-guide.pdf). Prof. Harel Weinstein, one of the proposed field members, is the NYC Chairman of this program. Prof. Ron Elber, another proposed field member, is the Ithaca Chairman.

The three other Weill faculty who will be field members are also participants in this program. While some students in this program may choose to do thesis work at Rockefeller of Sloan-Kettering, those who choose Cornell, either in Ithaca or NYC, will be well-served by the proposed field.

1.5 Faculty Membership

Field members are required to teach one general lecture covering their area of computational biology as needed for the academic program (maximum: one lecture every other year). (This may be in addition to normal departmental teaching responsibilities). After the formation of the Field, new membership will require a positive vote by the executive committee.

Faculty who do not serve on a Special Committee of a field student for five years or who do not complete the teaching requirement will be dropped from the field. Faculty may be dropped from the field by two-thirds vote of the executive committee.

2. The Ph.D. in Computational Biology

2.1 The Ph.D. Program

The Ph.D. in Computational Biology is intended for students who are interested in advancing studies at the interface of biological sciences and computer science. The program is designed to allow rigorous pursuit of both disciplines for the purpose of advancement of research that lies directly at the interface. The Cornell program has a strong emphasis on interdisciplinary research that bridges the gap between diverse biological disciplines, from behavior and ecology to molecular biology, to computational fields including algorithm development, numerical analysis and computationally intensive statistics. The focus is on advancement of fundamental research through recognition of common needs and goals of researchers and students working on computationally intensive aspects of biological research.

The program allows for specialization within the biological and computational sciences. The biological subareas are organized by biological size scale: Genetic, Macromolecular, Cellular, Organismal, Behavioral, and Ecological biology. Expertise within one of these is required. The computational subareas are Mathematical science, Computational science, and Programming. Distributed expertise in all three areas is required

A student who is awarded a Ph.D. in Computational Biology will need to achieve three objectives: (a) demonstrated competence in the disciplines that contribute to the field, (b) depth in at least one computational and one biological aspect of the field, (c) original research, on a topic from one or more of the Computational Biology concentrations.

2.2 Student Characteristics

Three primary objectives in organizing the new field are to attract excellent students to Cornell, to attract new faculty who see Computational Biology as their primary discipline, and to advance communication and collaboration among the active and diverse group of computational biologists already at Cornell.
Applications of computer science have expanded to cover nearly all disciplines of the biological sciences, and the growth of computational biology as a discipline has been dramatic. Currently, students interested in computational biology must select among biological disciplines or computer science. As mentioned above, programs in Computational Biology (or Bioinformatics) are becoming available at a large and growing list of universities (see http://anil.chmcc.org/University.html). Some of the schools include: Stanford, UCLA, UCSD, UCSC, Johns Hopkins, Yale, University of Michigan, Indiana University, University of Pittsburgh, Iowa State, and Florida State.
These programs are highly successful at attracting highly motivated students into their Ph.D. programs.

Because of the equal weight being placed on the computational and biological side of this new field, entering students may have a diversity of backgrounds.
There will be undergraduates who have already directly worked in both biological and computer sciences, either as biologists who learned and applied a computer language in undergraduate research projects, or as a computer science student who worked on a biological application. But many entering students will not have a balanced background, and instead will have much more experience in either the biological or the computational side.
The field will be looking for highly able students who can demonstrate strong potential in both computer science and the biological sciences.

The program requires depth in the form of strong analytical skills, and breadth in an appreciation and understanding of a number of the focus areas.
An ideal entering student will have an undergraduate degree in a related area, with solid writing skills, computing experience, and a mathematical foundation that includes probability, statistics, and linear algebra. However, the program is designed so that students have an opportunity to fill gaps in their background at the beginning of their studies.
In practice we expect that most entering students will have a strong undergraduate degree, with a major in a relevant field, including a significant quantitative or technical component.

Students leaving the program will have a very strong knowledge of the intersection between computational and biological sciences. There is great demand for students with such "bridging" skills. As such they will be very well placed to follow professional careers in both the biotech industry and academia. A strong demand for employment in both sectors has already emerged. Two types of academic faculty positions are emerging: (1) Employment in traditional disciplinary departments that are interested in recruiting computational biologists who bring a broader vision and expertise, even at the expense of more limited intradisciplinary training. Computer Science, Molecular Biology and Genetics, Physics, Chemical Engineering, and Theoretical and Applied Mechanics are examples of departments at Cornell that have either made, or have attempted to make, such hires within the past few years. (2) Employment in newly emerging interdisciplinary departments. The new Dept. of Biological Statistics and Computational Biology at Cornell is an example. The recent "Tri-Institutional Program in Computational Biology and Medicine" (http://www.cs.cornell.edu/grad/cbm/) sponsored by Cornell, Rockefeller University, and the Memorial Sloan Kettering Cancer Center, is another example of academic interdisciplinary expansion in this area. Graduates of the proposed field would be excellent candidates for the three faculty positions that were recently recruited under its auspices.

2.3 Academic Requirements

Students must demonstrate competency within both the Biological and Computational foci as described below.

Biological focus:

Within the biological focus area, students must demonstrate competency within one subarea by completing four course-equivalents (see below) for the subarea (Appendix 2). Additional breadth is suggested; at the discretion of the Special Committee, one of the four required courses may be assigned in a second area.

Subareas:

Genetic	(e.g., genomics, population genetics)
Macromolecular	(e.g., protein structure prediction, proteomics, protein dynamics, DNA and RNA structure)
Cellular	(e.g., expression level analysis, signal pathway modeling)
Organismal	(e.g., modeling hand and heart motion, sensory systems)
Behavioral	(e.g., animal behavioral modeling)
Ecological	(e.g., population and ecosystem modeling)

Competency within each subarea will be defined by working faculty subgroups within the guideline that the requirement for a student with no background in the area will consist of 4 (3-credit) courses plus additional background material learned by self-instruction (e.g., organic chemistry). Students with extensive background in a subarea may "advance place" some requirements based on prior course grades and diagnostic oral examination.

Computational focus:

Students must demonstrate competency within each of three areas (see below). Programming competency will usually be satisfied by small projects submitted to the Special Committee. Mathematical and computational competency will be demonstrated by completing a total of four course-equivalents (see below).

The courses will be selected in consultation with the Special Committee from the course list (Appendix 2) to provide the appropriate foundation for the student's anticipated area of study. At least two course-equivalents will be in Mathematical Science and at least one in Computational Science.

- *Mathematical science:* i.e., applied mathematics and statistics (differential equations, linear algebra, analysis, discrete mathematics, engineering mathematics, statistics)
- *Computational science*: e.g., databases, numerical analysis, algorithms
- *Programming:*
  - Standard metalanguage (e.g., Matlab, Mathematica, Stat package) (required)
  - Basic language programming skills (e.g., C++) (required)

Course-equivalents:

Students may "place" course-equivalents based on prior study. It is expected that a typical student will place 2-3 of the course requirements.

At least three of the courses actually taken during the program must be at the graduate level. Placement recommendations will be made following an informal oral examination administered by the Placement Committee[1] as soon as possible after the student enters the field. These recommendations will be added to the student's folder as a guideline for the Special Committee.

The Special Committee will ultimately make the final assessment.

Survey course:

In addition to the course requirements above, it is anticipated that students will take a survey course that will be developed in the future (i.e., once there are enough students in the program).

Minor: A single minor is required.

In consultation with the DGS, students should select this during their first year. The minor will usually be a related biological, computational, mathematical, or physical science field.

Special committee: Students should assemble a Special Committee at the end of the first year. This will consist of the thesis research supervisor (the Chairperson of the Special Committee), a faculty member representing the minor subject, and another faculty member from the field of Computational Biology. The student chooses the members other than the Chairperson with approval by the Director of Graduate Studies. The DGS will serve as Chairperson until the Special Committee has been chosen.

A-exam: The A-exam will be administered by the Special Committee and will consist of: (1) an oral examination to assess completion of the competency requirements in the Biological and Computational foci and in the proposed area of research. (2) a short (e.g., 10 page) written presentation outlining the background, progress and plans for thesis research (submitted to the committee at least 1 week prior to the exam), and (3) a short (1/2 hour) oral presentation of research progress and thesis research plans. The A-exam should be taken before the start of the third year.

Masters degree: The field will only accept students for a Ph.D. program. At the discretion of the Special Committee, students who are not making satisfactory performance may be awarded a terminal Masters degree if they have satisfactorily completed the competency requirements. Students who pass the A-exam may be awarded a Special Masters.

2.4 Ph.D. Thesis

Students are expected to make a thesis proposal by the end of their third year as a part of the A-exam. As part of the thesis proposal, the student will be required to demonstrate depth in at least one concentration, sufficient to carry out fundamental research.
The student's Ph.D. committee will decide how this expertise will be evaluated. The "Papers Option" is acceptable.

In addition to the required group of 600-level courses, the student's committee may require additional courses before approving the thesis proposal.
Each student must satisfy the requirements of one Minor in another graduate field. Primary subject areas of a Ph.D. thesis in Computational
Biology include: computational genetics, computational macromolecular biology, computational cell biology, computational organismal biology, computational behavioral biology, and computational ecology.

2.5 Requirements for Computational Biology Minor

Students must complete the academic requirements of the field. It is expected that a typical student will take four of the courses (from study associated with their major).

2.6 Admissions

Students new to Cornell will be accepted by an admissions committee. (Faculty may bring students of special interest to the attention of the committee.)
The executive committee may function as admissions committee until there are a large number of applications. Students will not be accepted until financial support for the first year has been arranged. If this is not via a fellowship, individual faculty who are interested in the student will be contacted to see if their department can provide a TA.
However, students will not be obliged to work with any specific faculty member.

It is expected that students will join a research group by the end of the first academic year. For a student to join a group, the faculty member and chair of his/her department must sign a letter assuming responsibility for financial support per Graduate School requirements.

The Admissions Committee can accept students transferring into the field from other Graduate fields.

3. Administration

The field will be co-supported by the Faculty of Computing and Information Science and the Department of Biological Statistics and Computational Biology and by the (see letters of support from Dean Constable and Prof. Martin Wells).
This will provide the complementary benefits of a focused departmental home (BSCB) and an overarching university-wide perspective (FCIS).

The Dept. of BSCB has two disciplinary foci: (1) Biological Statistics and (2) Computational and Statistical Genomics. The field of Statistics will play a central role in the first focus, and the new field of Computational Biology will play the central role in the second focus.

The Dept. of BSCB was previously the Dept. of Biometry. Its primary field was Biometry before its name change and broadening of focus. The majority of Biometry graduate students who are studying Biological Statistics have already transferred to the field of Statistics. Moreover, as new faculty are recruited and the area expands, the field of Computational Biology will serve as a more natural home for graduate students studying Computational and Statistical Genomics. Thus, the DGS of Biometry anticipates that the field of Biometry will be phased out in the future as its role is subsumed within the field of Statistics and the new field of Computational Biology. We are working in conjunction with Biometry to arrange a smooth transition for current students and to ensure that graduate teaching assistant needs in this area are satisfied.

The broader university-wide scope of the FCIS will support the new field in its coverage of other areas of computational biology (e.g., computational systems biology, computational neurobiology, etc.) that are not covered by the Dept. of BSCB.

3.1 Executive Committee

The executive committee will be chaired by the DGS and have one representative from each of the six academic subgroups.
If necessary, some members may represent multiple subgroups. Executive committee members will be elected annually by the field membership.

3.2 Administrative Support

Entering graduate students will be co-housed in a location provided by the Dept. of BSCB. After choosing an advisor, students will subsequently migrate to individual research laboratories across campus. The core location will continue to serve as a gathering point for these students for appropriate academic and social (community-building) events.

The new program will only supervise the management of the graduate program and students, and the only administrative need will be a part-time field secretary within the Dept. of BSCB. Additional support as needed in the future will be provided by the Faculty of Computing and Information Science (see letter of support from Dean Constable).

3.3 Student diversity

We are committed to diversity and will work with the existing Cornell programs to support this in our recruiting efforts.

3.4 Student Support

Consistent with the patterns of student support in most of the scientific fields at Cornell, faculty will be expected to support students on research grants or departmental TAs beginning in the second year of study. The field of Computational Biology faculty are, in general, well funded (the average external funding is approximately $350,000/yr), so this should not be a problem.

During the first year students will receive full support from other means. The Chairs of the Departments of Computer Science and Biological Statistics and Computational Biology estimate that their departments will each contribute on average one TA per year. In addition, we expect that a few of the six students per year that are accepted into the Tri-Institutional Training Program in Computational Biology & Medicine (see Sec. 1.4) will enter this field.
(The Tri-Institutional program has guaranteed funding until 2006. It is currently working towards securing continuing funding.)
Furthermore, Computational Biology is a central and essential component of genomics and we expect that one-two of the fifteen Presidential Genomics Fellowships will be awarded to students entering Computational Biology.

This limited funding will permit us to initially matriculate roughly four to five students/year (plus students who obtain independent funding), which will be adequate to establish the Field. However, this is a growing intellectual area and we anticipate that additional funding will be required in the future. Towards this end, the Dean of Computing and Information Science is committed to fund-raising to endow additional support for three students/year (see letter from Dean Constable). Moreover, the University is currently fundraising for the New Life Sciences Initiative, which may “create as many as 100 new graduate fellowships” (http://www.news.cornell.edu/Chronicle/02/5.9.02/NLSI.html). As with the Presidential Genomics Fellowships, it is expected that some of these will support students in Computational Biology. We will also purse additional funding from NIH and NSF Training Grants, targeted endowments.[2] This is partly a “chicken and egg” problem---the field must be established before we can enlist support for these various endeavors.

3.5 Fundraising and the Computational Biology “Umbrella”

While the field will function as an independent unit, a very important auxiliary role will be to serve as an “umbrella” for specialized subprograms within its broad scope (e.g., computational genomics, biological non-linear systems, etc.).
In particular, we anticipate that the field will support these subprograms in individually obtaining external support for training within their area of interest by providing an established academic and administrative framework.

Appendix 1- Faculty

Faculty marked with * will serve on the initial Executive Committee of the Field.

***Carlos Bustamante**
Evolutionary genomics, Bayesian inference, computational statistics, protein structure and evolution

***Andy Clark**
Molecular and theoretical population genetics; molecular evolution; Drosophila evolutionary biology

David Christini (WCMC) Cardiac electrophysiology, dynamics of arrhythmias

***Evan Cooch**
Methods for quantitative analysis and management of population dynamics, theoretical ecology

***Ron Elber** Sequences, structures, dynamics and function of proteins

Steve Ellner Theoretical and computational population biology

Robert Gilmour Dynamics of cardiac electrical activity and arrhythmias

***John Guckenheimer** Analysis of dynamical systems, especially neural systems

***Ron Hoy** Mechanisms of hearing in animals (insects) neuroanatomy, development, behavioral neurobiology

Uri Keich Statistical and algorithmic problems in bioinformatics

Jon Kleinberg
Combinatorial structure of networks and information, with applications to computational biology

***Kelvin Lee** Mathematical frameworks for cellular processes and gene regulation

***Christiane Linster** Neural basis of sensory processing and memory

Hod Lipson Evolutionary robotics and artificial life

Diana Murray (WCMC) Computational analysis of macromolecular interactions in subcellular targeting

Rasmus Nielsen Statistical genomics, molecular evolution and population genetics

Richard Rand Dynamical systems, biological systems

***Kern Reeve**
Evolutionary game theory applied to the analysis of cooperation and conflict in social organisms

Benoit Roux (WCMC) Computational studies of biological membrane ion channels

***David Ruppert** Statistical modeling in biology

Harold Scheraga Physics-based methods to compute protein structure from amino acid sequence

Steve Schwager Multivariate analysis, experimental design, linear models, statistical genetics

Jim Sethna Condensed-matter theory, dynamics of proteins, RNA, and DNA in the cell

***David Shalloway** Global optimization for predicting protein structure and motion

David Shmoys Efficient algorithms for discrete optimization with applications in computational biology

Christine Shoemaker
Population modeling and optimization, including application to bacteria involved in groundwater remediation

Eric Siggia Statistical mechanics and pattern recognition applied to genomic and RNA expression data

***Patrick Sullivan** Spatial and temporal statistical modeling and assessment of populations and communities

Marjolein C H van der Meulen Orthopedic biomechanics of bone adaptation to mechanical stimuli

Jane Wang Biofluid dynamics and scientific modeling and computing

Harel Weinstein (WCMC) Molecular biophysics of signal transduction in function of proteins and DNA

Marty Wells Applications of statistics in biology.

***Golan Yona**
Large scale analysis of protein space and cell-wide processes in search of global principles and self-organizatio

Represented Departments

Biological Statistics and Computational Biology
Biomedical Sciences
Chemical Engineering
Chemistry and Chemical Biology
Computer Science
Ecology and Evolution
Mathematics
Mechanical and Aerospace Engineering
Physiology and Biophysics (WCMC)
Microbiology and Immunology (WCMC)
Molecular Biology and Genetics
Natural Resources
Neurobiology and Behavior
Operations Research and Industrial Engineering
Physics
Theoretical and Applied Mechanics

(WCMC= Weill Cornell Medical College)

Appendix 2

Acceptable Course Recommendations by Subarea

This list is intended as a guideline to Special Committees. It is not exhaustive, and more advanced courses that are not listed are also appropriate. Higher-level courses should be emphasized whenever possible.

Genetics:

BIOGD 481	Population genetics
BIOGD 487	Human Genomics
BTRY 382	Introduction to Statistical Genomics and Bioinformatics
BTRY 482/682	Statistical Genomics
BTRY 652	Computationally Intensive Statistical Inference
BTRY 726	Problems and Perspectives in Computational Mol. Biology
BTRY 408	Probability Theory
BTRY 409	Theory of Statistics
OR&IE 650	Applied Stochastic Processes
OR&IE 651	Probability Theory
OR&IE 670	Statistical Principles
OR&IE 671 I	Intermediate Applied Statistics
OR&IE 678	Bayesian Statistics and Data Analysis
BTRY 694	Stochastic Modeling w/ Applications

Macromolecular Biology:

BEE 760	Nucleic Acid Engineering
BIOBM 330 or (331 and 332) or 333	Principles of Biochemistry: Proteins, Metabolism, and Molecular Biology
BIOBM/ BioGD 439	Molecular Basis of Human Disease
BIOBM 631	Protein Structure and Function
BIOBM 633	Biosynthesis of Macromolecules
BIOBM 638	Macromolecular Interactions and Cell function
BIOBM 639	The Nucleus
BIOBM/ VETPR 730	Protein NMR Spectroscopy
BIOBM/ CHEM 738	Macromolecular Crystallography
BIOPL 462	Plant Biochemistry
BIOPL 652, 653	Plant Molecular Biology I, II
BMEP 302	Cellular Principles of Biomedical Engineering
BMEP 401/ M&AE 466	Biomedical Engineering Analysis of Metabolic and Structural Systems
CHEM 357/358, 359/360	Organic Chemistry for the Life Sciences
CHEM 450	Principles of Chemical Biology
CHEM/ VET MM 672	Kinetics and Regulation of Enzyme Systems
CHEM 686	Physical Chemistry of Proteins

CHEM 780	Chemical Kinetics and Molecular Reaction Dynamics
NS 602	Lipids
VET MM 610	Cellular and Molecular Pharmacology

Cell Biology:

BEE 360	Molecular and Cellular Bioengineering
BEE 484	Metabolic Engineering
BIOBM 432	Survey of Cell Biology
BIOBM/TOX	Eukaryotic Cell Proliferation
BIOBM/ BIOGD 439	Molecular Basis of Human Disease
BIOBM 636	Advanced Cell Biology
BIOBM 639	The Nucleus
BIOG 305/ VETMI 315	Basic Immunology
BIOG/ VETMI 705	Advanced Immunology
BIOMI/ VETMI 408, 409	Viruses and Disease I, II
BIOMI 416	Bacterial Physiology
BIOMI 420	Microbial Genomics
BIOMI/ BIOPL 652	Molecular Plant-Microbe Interactions
BIOMI 690	Prokaryotic Biology
BIONB 425	Molecular Neurophysiology
BIONB 495	Molecular and Genetic Approaches to Neuroscience
BIOPL 422	Plant Development
BIOPL 444	Plant Cell Biology
BIOPL 652, 632	Plant Molecular Biology I, II
BMEP/ CHEME 605/6	Fundamentals of Biomedical Engineering I, II
ENTOM 483	Insect Physiology
VETMI 409	Principles of Virology
VETMI 700	The Biology of Animal and Plant Viruses
VETMM 703	Receptor-Ligand Interactions
VETMM 705	Chemistry of Signal Transduction
VETMM 706	Growth Factor-Coupled Signaling

Organismal Biology:

BEE 454	Physiological Engineering
BIOAP 316	Cellular Physiology
BIOAP 311	Introductory Animal Physiology, Lectures (also Vetph 346)
BIOAP 319	Animal Physiology Experimentation
BIOAP 458	Mammalian Physiology
BIOAP 714	Cardiac Electrophysiology
BIOGD 394	Circadian Rhythms (also Entom 394 and Bionb 394)
BIOPL/ BEE 456	Biomechanics of Plants
NS 341	Human Anatomy and Physiology

Behavioral Biology:

BIONB 326 The Visual System
BIONB/
BMEP/
COGST/
PSYCH 330 Introduction to Computational Neuroscience
BIONB 425 Molecular Neurophysiology
BIONB 491 Principles of Neurophysiology
COGST/
PSYCH 416
and 616 Modeling Perception and Cognition
ECE 577 Feedforward Neural Networks
PSYCH 419 Neural Networks Laboratory

Ecology:

BIOEE/
MATH 362 Dynamic Models in Biology
BIOEE 460 Theoretical Ecology
BIONB 422 Modeling Behavioral Evolution
AN SC 420 Quantitative Animal Genetics
NTRES 305 Applied Population Ecology
NTRES 340 Quantitative Population Analysis
NTRES 459 Wildlife Population Analysis: Techniques & Models
NTRES 670 Spatial Statistics

Mathematical Science:

Differential equations
MATH 420 Differential Equations and Dynamical Systems
MATH 427 Introduction to Ordinary Differential Equations
MATH 428 Introduction to Partial Differential Equations

Linear Algebra
MATH 431 Linear Algebra
MATH 433 Honors Linear Algebra

Analysis
MATH 413 Honors Introduction to Analysis
MATH 422 Applied Complex Analysis
MATH 424 Wavelets and Fourier Series

Discrete Math
MATH 441, 442 Introduction to Combinatorics I, II

Probability and Statistics
BTRY 408 Theory and Probability
BTRY 409 Theory and Statistics
BTRY 603, 604 Statistical Methods III, IV
BTRY 652 Computationally Intensive Statistical Inference
BTRY 682 Statistical Genomics
MATH 471 Basic Probability
MATH 472 Statistics
OR&IE 476 Applied Linear Statistical Models
OR&IE 523 Operations Research II: Introduction to Stochastic Processes I

OR&IE 575 Experimental Design
OR&IE 576 Regression

Engineering Math
T&AM 610,
611, 612,
613, 614 Methods of Applied Mathematics I, II, III, IV, V

Computational Science
Fundamentals
CIS 409 Data Structures and Algorithms for Computational Science
COM S 280 Discrete Structures
COM S 312 Data Structures and Functional Programming

Numerical Analysis
COM S/
BIO BM/
ENGRD 321 Numerical Methods in Computational Molecular Biology
COM S 421 Numerical Analysis
MATH 425 Numerical Analysis and Differential Equations
ORIE 520, 521 Operations Research I: Optimization I, II

Algorithms
COM S 482 Introduction to Analysis of Algorithms

Databases
COM S 432 Introduction to Database Systems

Artificial Intelligence and Machine Learning
COM S 472 Foundations of Artificial Intelligence
COM S 478 Machine Learning
COM S 578 Empirical Methods in Machine Learning and Data Mining

Computational Biology
BioNB 330 Introduction to Computational Neuroscience
BioNB 441 Computers in Neurobiology
COM S 426 Introduction to Computational Biology
COM S 626 Computational Molecular Biology

Programming
Although courses are not required in this area, the following may be of interest:

CIS 401 Introduction to Applied Scientific Computing with MATLAB
CIS 402 Scientific Visualization with MATLAB
CIS 403 Development of Scientific Computing Programs
CIS 404 Survey and Use of Software Libraries for Scientific Computing

Appendix 3-Potential External Reviewers

1. **Professor Jeff Skolnick,**
Director of the Center of Excellence in Bioinformatics, University of Buffalo, 901 Washington Street, Ste. 300, Buffalo, NY 14203-1199, (716) 849-6711, skolnick@buffalo.edu

2. **Professor Simon Levin,**
Moffett Professor of Biology, Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, (609) 258-6880,
slevin@princeton.edu (previous Cornell faculty member)

3. **Professor David Haussler,**
Director, Center for Biomolecular Science and Engineering, Professor of Computer Science, Center for Biomolecular Science & Engineering, 321 Baskin Engineering Bldg, University of California, Santa Cruz, CA 95064, (831)459-2105, haussler@cse.ucsc.edu

4. **Professor Charles Peskin,** Department of Mathematics, NYU, 251 Mercer St., New York, NY 10012, (212) 998-3126, peskin@cims.nyu.edu

5. **Professor John Straub**, Chemistry Department, Boston University, 590 Commonwealth Ave., Boston, MA 02215, (617) 353-6816, straub@bu.edu

6. **Professor Leon Glass**, Isadore Rosenfeld Chair in Cardiology, Professor of Physiology, Centre for Nonlinear Dynamics, Department of Physiology, McGill University, 3655 Prom. Sir William Osler, Montreal, Quebec, Canada H3G lY6, (5l4) 398-4338, glass@cnd.mcgill.ca

7. **Professor Webb Miller**, Department of Computer Science and Engineering, The Pennsylvania State University, 326A Pond Laboratory, University Park, PA 16802-6106, (814) 865-9505 webb@cse.psu.edu

Appendix 4- Letters of Support

1. Dean Robert Constable, Computing and Information Science
2. Prof. Martin Wells, Chair, Dept. of Biological Statistics and Computational Biology and DGS, Field of Biometry
3. Prof. Tom Fox, Chair, Dept. of Molecular Biology and Genetics
4. Prof. Charles VanLoan, Chair, Dept. of Computer Science
5. Prof. Jerry Feigenson, DGS, Field of Biophysics
6. Prof. Terrence Fine, DGS, Field of Applied Mathematics
7. Prof. Joseph Halpern, DGS, Field of Computer Science

---

[1] Initially, the "Placement Committee" will consist of the Executive Committee. This committee will be headed by the DGS and also contain faculty representing expertise in the computational/mathematical and six biological subareas.

[2]

Computational Biology is a field with excellent opportunities for external funding, both through the National Institutes of Health and the National Science Foundation. For example, just in the last year the NIH released a new "Roadmap" for NIH-funded research (http://nihroadmap.nih.gov). This lists "Bioinformatics and Computational Biology" as one of five targeted areas for development.