A.

Proposal for an Interdepartmental Major in

BIOINFORMATICS AND COMPUTATIONAL BIOLOGY (BCB)

Contents:

APPENDICES:

      I.   List of Founding Faculty & Curricula Vitae

      II.   List of Participating Departments & Memos of Support

      III.   Editorial from Bioinformatics:
            "A Curriculum for Bioinformatics: The Time is Ripe"

      IV.   BCB Curriculum and Training Plan

      V.   BCB Governance Document

      VI.   BCB Outcomes Assessment Plan

The text of this proposal is available at:

http://www.public.iastate.edu/~ddobbs/BCBmajor.html

If more information is needed, please contact:

Drena Dobbs
Dept. of Zoology and Genetics
2114 Molecular Biology Building
Iowa State University
Ames, IA 50011

(515)294-1112
ddobbs@iastate.edu

PROPOSAL FOR AN INTERDEPARTMENTAL MAJOR
IN BIOINFORMATICS AND COMPUTATIONAL BIOLOGY

<u>ABSTRACT</u>

Due to the genome sequencing efforts and recent advances in molecular genetics, the next generation of biologists will be able to address questions at new levels of complexity. As a consequence, biological inquiry will increasingly require a multidisciplinary approach that involves not only diverse expertise in the biological sciences - from evolutionary to structural biology - but also expertise from disciplines in which biologists typically have only rudimentary training, including computer science, statistics and mathematics. This document proposes to establish an Interdepartmental Bioinformatics and Computational Biology (BCB) Major at Iowa State University that will address the training needs of modern biology.

The new interdepartmental major will be offered initially by 33 faculty from 12 departments. The programmatic strength of BCB will be rooted in the diverse areas of expertise of these faculty, which span the disciplines of genetics, mathematics, statistics, computer science, physics, and chemical engineering, as well as structural, molecular, integrative and evolutionary biology. The curriculum will include fundamental training in each of these disciplines, with an emphasis on advanced work in molecular biology. The curriculum will also place a heavy emphasis on interdisciplinary and collaborative research interactions. These will, in part, be fostered by interdepartmental seminars, advanced topic lecture courses and annual retreats.

The BCB program will be administered by a Chair and a Supervisory Committee elected by the faculty. Other BCB faculty will serve on committees that will be responsible for student recruitment, student admissions, curriculum and faculty membership. As one of several interdepartmental majors at Iowa State University, BCB will fall under the auspices of the Graduate College.

PREAMBLE

Information is a major theme in biology. The history of modern biology has taught us that understanding the messages used in living organisms is essential to understanding life itself. When specific messages and codes are deciphered, this knowledge can be used to produce powerful tools for discovering still more about life.

The messages and information contained in an organism's genome represent the rules governing the development, behavior and fitness of that organism. These rules determine an organism's structure and function at increasingly higher levels of biological organization. Discovery of how behavior emerges from the lowest levels of organismal organization is the challenge and opportunity that molecular biology, bioinformatics and computational biology could provide to those dedicated and creative enough to seize it.

The last few years have witnessed an explosion in genomic sequence and protein structural data. These data contain a wealth of information about how organisms function and how they came to be. Today there are great opportunities for discovering the nature of this information and for using this knowledge to create tools that will aid biologists, medical practitioners, agronomists, and genetic engineers working in diverse areas. Drug companies and other major corporations are hiring scientists who can combine knowledge of genetics, biochemistry, computer science, mathematics and statistics to exploit this opportunity. Molecular geneticists are witnessing a revolution in the way they do business. With the generation of increasing numbers of complete genomic sequences, genetic experiments now start with a computer analysis and mutations are engineered, not found. The current ability of scientists and practitioners to access, manipulate, and interpret the rapidly accumulating mountain of data is falling badly behind. Bioinformatics is the emerging field which specifically addresses this challenge in the short term, and in the long term will increasingly address deeper questions about how information is utilized by life. The field is interdisciplinary, with successful practitioners bringing together mathematical and statistical theory, knowledge of biology, and state-of-the-art computational approaches. The proposed graduate major in Bioinformatics and Computational Biology (BCB Major) will prepare students for careers at the interface of information science and biology.

Previous activity leading to this proposal

Jim Cornette, Department of Mathematics, worked for many years to stimulate interactions between mathematicians and biologists. In the fall of 1997, he arranged a seminar to explore DNA sequence analysis. This weekly seminar attracted about twenty mathematicians, computer scientists and biological scientists and stimulated extensive dialog among them.

During the 1997-1998 year, there was a remarkable level of excitement and activity based on these interdisciplinary faculty interactions, included multiple research collaborations and grant proposals. The Iowa Computational Biology Laboratory (ICBL) was founded to provide an organized focus and a vehicle for promoting these activities (http://www.math.iastate.edu/danwell/icbl.html). Members of the ICBL initiated student training in computational molecular biology, including workshops, seminars and a new graduate course in the subject. Twenty-two students took the first offering of Math/Gen/ComS 594X "Computational Molecular Biology" in the spring of 1998.

The first step toward a graduate program in bioinformatics and computational biology was made in the spring of 1998, when three existing graduate majors, Computer Science, Genetics, and Mathematics cooperatively agreed

to offer complementary "Areas of Specialization" in Computational Molecular Biology. Graduate minors already provide an alternative mechanism for training in this multidisciplinary field, but a carefully planned graduate major offers the best mechanism for preparing students in this new and exciting area.

In the spring of 1997, an *ad hoc* committee chaired by Tom Peterson submitted a comprehensive report outlining what ISU needed to do to position itself as a premier research institution in the plant sciences. The number one priority identified was the need to hire several computational biologists who would work with biological researchers in the various plant genome projects to enable better access to and use of the rapidly growing body of plant genomic data.

In the spring and summer of 1998, with support from a special appropriation in the state 1999 fiscal year budget, three departments (Zoology & Genetics, Computer Science and Agronomy) pooled resources and hired three new computational biologists. Two structural biologists with computational interests were also hired by the Department of Biochemistry, Biophysics and Molecular Biology. The new hires complement the biology, computer science, physics and mathematics faculty who have initiated active collaborations on campus during the past year. Three faculty (two biologists and one computer scientist) have taken advantage of ISU's on-campus sabbatical opportunities to train in Computational Molecular Biology as a second discipline.

The momentum has continued to build during 1998-99. In February, 1999, the Department of Physics and Astronomy has offered a position to a Computational Structural Biologist and candidates are being interviewed for a joint position in Bioinformatics by the Departments of Statistics and Zoology & Genetics. A Computational Molecular Biology Curriculum Committee, chaired by Steve Willson (Mathematics) and including members of ICBL and other interested faculty, developed a suggested curriculum that, together with the training plan developed by Dan Voytas for an NSF training grant proposal (see below), is the basis for the BCB Curriculum and Training Plan included in this proposal. Bioinformatics faculty have submitted several interdisciplinary grant proposals, including a $2.7 million NSF IGERT proposal (Dan Voytas, P.I.) for a training grant in Computational Molecular Biology. Voytas recently received notification that the training grant has been recommended for funding. The CMB Training Grant will provide research assistantships for 4 new graduate students and 2 postdoctoral fellows each year for a total of 5 years (beginning in July, 1999). An additional source of funding for students majoring in BCB has been obtained through the efforts of Jim Cornette and the ICBL. Pioneer Hi-Bred International has committed up to $500,000 for graduate and postdoctoral fellowships in computational molecular biology over a five year period (beginning in Jan. 1999). In addition, the University has supplied funding for the establishment of the Iowa Genomic Frontier Center (IGF, P. Schnable, Director) which provides support facilities for high-throughput DNA sequencing, microarray RNA expression analysis, high-throughput protein identification and analysis (proteomics), robotics components, and computational support.

Several courses and seminars in bioinformatics are being offered in spring of 1999. These include the second offering of Gen/Math/ComS 594 (taught by new hire, Volker Brendel, Zoology & Genetics), and two new seminar courses, Gen 590 "Topics in Molecular Biology for Computational Scientists" (Drena Dobbs, Zoology & Genetics), and ComS 610 "Intelligent Systems in Molecular Biology" (Vasant Honavar, Computer Science).

At least five new courses are planned for fall of 1999, including two "bridge" courses: Math/ComS/Gen 592X "Computational Mathematics for Biologists" (Dan Ashlock, Mathematics, primary instructor), Gen/ComS/Math 595X "Molecular Biology for Computational Scientists" (Drena Dobbs, Zoology and Genetics, primary instructor) and three new advanced courses: Stat 537 "Statistics for Molecular Genetics (Alicia Carriquiry and

Mike Daniels, Statistics), Gen/ComS 555X "Software Methods for Biologists" (new hire, Hui-Hsien Chou, Zoology & Genetics/Computer Science) and Gen 556X "Computational Genomics and Evolution" (new hire, Xun Gu, Agronomy/Zoology & Genetics).

We propose to establish the Bioinformatics and Computational Biology Major at ISU to build on these initiatives, to provide interdisciplinary graduate training in molecular biology, mathematics and computer science, and to promote innovative research at the boundaries of these disciplines. With the new hires in this area, our faculty now provide a sufficient number of interested and knowledgeable staff to offer a world class graduate major in Bioinformatics and Computational Biology. Few other universities can make this claim.

A. BASIC INFORMATION

1. Name of the proposed interdepartmental major

Bioinformatics and Computational Biology (BCB)

2. Name of the advanced degree

Doctor of Philosophy in Bioinformatics and Computational Biology
Master of Science in Bioinformatics and Computational Biology

3. Names of the departments involved

Animal Science, Botany, Chemical Engineering, Computer Science, Mathematics, Physics and Astronomy, Statistics, Zoology and Genetics.

4. Need for the proposed graduate major

a) National and international need for bioinformaticists/computational biologists

Currently there is an vast shortage of scientists adequately trained in bioinformatics and computational biology. The demand for such scientists is very strong both in academia and industry and is likely to remain high, given the rapid pace at which the genome sequencing and high-throughput gene expression analysis projects are progressing. Employment opportunities for students with advanced degrees in this field are outstanding. (See Appendix III for an editorial that addresses the need for trained Bioinformaticists and Computational Biologists and discusses an appropriate curriculum.)

b) Need for the proposed interdepartmental major in Bioinformatics and Computational Biology at Iowa State

The primary goal of the Bioinformatics and Computational Biology interdepartmental graduate major will be to prepare students for careers in bioinformatics/computational biology research and university teaching. Interest in this field is very high among students at ISU and elsewhere. Computer science courses are taken by about one third of ISU's Genetics graduate students, and several have initiated double majors in Genetics and Computer Science. (And, perhaps due to WEB sites for ICBL and ISU's Bioinformatics job searches, several students from other universities have already written to apply for admission to ISU's program in Bioinformatics.)

5. Objectives of the proposed major

The Interdepartmental Graduate Major in Bioinformatics and Computational Biology (BCB Major) is an interdepartmental and interdisciplinary training program created to provide advanced training and foster research excellence in Bioinformatics and Computational Biology at Iowa State University. Research in biology has been transformed by the richness of information accumulating in genomic DNA sequence databases, rapidly accumulating RNA and protein expression data, and the increasing availability of "organism-specific," "disease-specific," and "protein-specific" databases. These developments have

made bioinformatics a critical component of research in molecular biology, genetics, biochemistry and biotechnology. Bioinformatics addresses a critical challenge in these fields: bridging the gap between the rate at which scientific data are being gathered and stored in digital form and our ability to fully exploit the available data. The proposal for a Bioinformatics and Computational Biology graduate major at ISU reflects a broad recognition both in academia and industry that there are too few scientists adequately trained to meet this challenge.

By incorporating existing strengths in mathematics, computer science, statistics and molecular genetics into an interdepartmental major, the BCB program seeks to fulfill these specific objectives:

- To provide broad and robust graduate student development in Bioinformatics and Computational Biology.

- To foster further intellectual exchange and research collaborations among Iowa State Bioinformatics and Computational Biology faculty, students and staff

- To provide a formal entity for seeking broad-based resources for the support of lecture series, retreats, graduate assistantships, postdoctoral fellowships and various graduate student prizes for excellence in Bioinformatics and Computational Biology Research. This financial support will be sought from the participating Iowa State colleges, private donations and national training grant awards.
- To enhance the national and international reputation of Iowa State University in the field of Bioinformatics and Computational Biology.

A graduate degree in Bioinformatics and Computational Biology from Iowa State University will provide students with skills required to access, manipulate and interpret biological sequence and structure data. The degree program ensures that students receive advanced training in both the biological and computational sciences, but allows students to focus on specific areas, such as gene prediction for functional genomics, or design of algorithms for sequence/structural alignment. BCB program graduates will be prepared for employment in either academic or industrial settings, in jobs ranging from genetic engineering of crop plants or animals in agribusiness, to drug design in the pharmaceutical industry, to research and teaching in a college, university, or biomedical institution. An outline of the BCB Outcomes Assessment Plan is provided as Appendix VI.

6. General description of the program

The Bioinformatics and Computational Biology Program at Iowa State will consist of a diverse group of highly interactive scientists who are enthusiastic about the challenges and opportunities presented by modern biology. These scientists have expertise in four interdependent research areas:

1) Genomics - including ISU faculty involved in Genomics initiatives with plants, animals and microbes. These scientists, along with their colleagues and collaborators from around the world, are producing the prodigious quantities of sequence and expression data that have opened up new avenues of biological inquiry.

2) <u>Genome Evolution</u> - including ISU faculty with expertise in molecular evolution with emphasis on whole genome analyses and on understanding patterns and processes of change that occur among genes and genomes over time

3) <u>Macromolecular Structure and Function</u> - including ISU faculty whose research aims to discern structural and functional meaning from DNA, RNA and protein sequences.

4) <u>Bioinformatics</u> - In our definition, Bioinformatics encompasses aspects each of the other three disciplines and provides the information science techniques necessary to integrate data from research in genomics, molecular evolution and macromolecular structure/function relationships to provide understanding of biological systems at new levels of complexity.

The focus of the BCB major is computational molecular biology. The scope of the major includes research that addresses problems related to storage, retrieval and analysis of information about molecular sequence, structure, function, and evolution. Some specific areas within this scope include: genomics, proteomics, functional genomics, evolutionary genomics, characterization of macromolecular structure/function relationships, prediction of macromolecular structure, design and implementation of biological data warehouses, design and optimization of algorithms for sequence/structure alignment and analysis, development of machine learning/data mining algorithms.

The curriculum of the proposed BCB major will include required core courses, seminars, annual workshops in advanced research topics and an annual BCB research retreat. Students will also receive training in the responsible conduct of research and other relevant issues in bioethics. During the first year, BCB students will do research rotations with three different BCB faculty or research groups. In the second year, students will initiate a multidisciplinary thesis research project under the guidance of two mentors, one biologist and one mathematician/computer scientist. Details of the Curriculum and Training Plan are provided below in Section 8.c. and in Appendix IV.

The BCB program will be administered by a Chair and a Supervisory Committee elected by the faculty. Other BCB faculty will serve on committees that will be responsible for student recruitment, student admissions, curriculum and faculty membership. Details of the administrative structure of the BCB program are described in Appendix V, BCB Governance Document.

As one of several interdepartmental majors at Iowa State University, BCB will fall under the auspices of the Graduate College. Students will be recruited and admitted directly into the interdepartmental program, rather than to individual participating departments. Details of the relationship between the BCB program and participating departments are described in Appendix V.

7. Comparison of the proposed program with

a) Standards, if any, established by the accrediting associations

There are no accrediting associations that establish standards for degrees in this area.

b) Similar programs at other universities

There are no similar programs in the state of Iowa. Only a few degree granting programs are available in the United States, although efforts to establish programs are underway at many universities. Currently, Boston University offers M.S. and Ph.D. degrees in Bioinformatics; George Mason University offers a Ph.D. in Bioinformatics and Computational Biology; Rutgers University and University of Medicine and Dentistry of New Jersey jointly offer a Ph.D. program in Computational Molecular Biology; Washington University in St. Louis offers a Ph.D. in Computational Molecular Biology; Stanford University offers a certificate in Bioinformatics; and two Keck centers (one a collaboration between Carnegie Mellon, the Pittsburgh Supercomputing Center, and the University of Computational Biology at Pittsburgh, the second a collaboration between Baylor College of Medicine, Rice University and the University of Houston) offer training programs in computational biology through other formal programs of study.

These newly established programs are generally similar in curriculum to the proposed BCB program at Iowa State University. Aspects that distinguish the program at ISU from others include: joint mentorship of graduate students by a biologist and a computational scientist (this is also offered at Rutgers), and active participation of scientists from a broad range of departments in both research training and teaching.

8. Program requirements

a) Prerequisites for prospective students

It is expected that entering BCB graduate students will have had background training in one of the foundation disciplines – molecular biology, mathematics, computer science or statistics. During the first two years of graduate school, students will engage in coursework designed to give them breadth in training and to make up any deficiencies (see Section 8.c. below).

b) Language requirements

Language requirements will be determined by the POS committee.

c) Courses and seminars presently available for credit for toward the program

All courses and seminars listed in the following curriculum are currently available or planned for Fall '99 and/or Spring '00.

i. Required background coursework

To participate in the BCB major, students must demonstrate basic competence in mathematics, molecular genetics, computer science, and statistics. This may be demonstrated by taking the indicated ISU courses or by equivalent previous coursework taken at other institutions:

Category I. Calculus and matrix theory

Math 265	Calculus III	4 Cr.

and/or Math 307   Theory of Matrices   3 Cr.

Category II. Molecular genetics

Gen 411	Molecular Genetics	3 Cr.
or Gen 511	Molecular Genetics	3 Cr.
or BB 405	Biochemistry	3 Cr.
or BB 502	Comprehensive Biochem.	4 Cr.

Category III. Analysis of algorithms

Com S 311	Design and Analysis of Algorithms	3 Cr.
or Com S 511	Design and Analysis of Algorithms	3 Cr.

Category IV. Statistics

Stat 401	Statistical Methods for Research Workers	4 Cr.
or Stat 447	Statistical Theory for Research Workers	4 Cr.
or Stat 500	Statistical Methods	4 Cr.
or Stat 543	Theory of Probability and Statistics	4 Cr.

Category V. Discrete structures

Com S 330	Discrete Computational Structures	3 Cr.
or Math 304	Introductory Combinatorics	3 Cr.
or Math 314	Graphs and Networks	3 Cr.

ii. Required core courses

Gen/Math/ComS 594: Introduction to Computational Molecular Biology (3 cr.) (Spring) This core course is required of all BCB students. It provides a state-of-the-art introduction to the subject, with emphasis on concepts and principles, combined with hands-on (keyboard) applications.

iii. Required seminars

BCB graduate students will supplement the basic course requirements by participation in at least four research seminars, two of which must require the student to make a presentation. At least two of the seminars must be interdisciplinary. Seminars that can be used to fulfill this requirement include:

BCB 691 - BCB Faculty Research Seminar. (1 cr.) (Fall) This in-house seminar series will be sponsored by the Iowa Computational Biology Lab (ICBL) and will feature research efforts of participating faculty and industrial collaborators.

BCB 690 - BCB Student Research Seminar. (1 cr.) (Spring) Each spring, students will present an account of their annual research progress to peers and to two faculty mentors, who will promote group discussions of experimental procedures and analysis.

BCB 591 - Annual BCB Symposium. The annual Bioinformatics and Computational Biology Symposium will be a two-week series of lectures by nationally and internationally renowned scientists. Each year the symposium will focus on a specific topic to provide an in-depth experience in a discipline at the forefront of bioinformatics and computational biology. Prior to the Symposium, students will read background materials and papers for each speaker.

Annual BCB Research Retreat. In the beginning of the fall semester, an all-day research retreat will be held at which faculty and advanced students will give short research talks, and all BCB students will make poster presentations of their work.

iv. Advanced group requirements

Students must complete advanced training in two of the three following subject areas: molecular biology, computer science or mathematics/statistics. At least 9 credits of advanced graduate credit must be taken from one subject area and at least 6 credits from another. One of these groups must be molecular biology. The table below provides a list of some of the courses that may be used to fulfill this depth requirement.

Group I. Molecular Biology

BBMB 404	Biochemistry	Cr. 3
BBMB 405	Biochemistry	Cr. 3
BBMB 451	Physical Biochemistry	Cr. 2
BBMB 501	General Biochemistry	Cr. 3
BBMB 502	General Biochemistry	Cr. 3
BBMB 531	Structure and Reactivity	Cr. 1
BBMB 551	Molecular Biophysics	Cr. 3
BBMB 653	Protein Chemistry--Physical Methods	Cr. 1
Gen 411	Molecular Genetics	Cr. 3
Gen 460	Mathematical Genetics	Cr. 2
Gen 462/562	Evolutionary Genetics	Cr. 3
Gen 511	Molecular Genetics	Cr. 3
Gen/Com S 555X	Software Methods for Biologists	Cr. 3
Gen 556X	Computational Genetics & Evolution	Cr. 3

Group II. Computer Science

Com S 311	Design and Analysis of Algorithms	Cr. 3
Com S 511	Design and Analysis of Algorithms	Cr. 3
Com S 461	Introduction to Database Systems	Cr. 3
Com S 561	Principles of Database Systems	Cr. 3
Com S 472/572	Principles of Artificial Intelligence	Cr. 3
Com S 474	Elements of Neural Computation	Cr. 3
Com S 672	Computational Models of Learning	Cr. 3

Group III. Mathematics and Statistics

Math 314	Graphs and Networks	Cr. 3
Math 378	Optimization & Modeling with Artificial Life	Cr. 3
Stat 503	Exploratory Data Analysis	Cr. 3
Stat 536	Genetic Statistics	Cr. 2
Stat 537	Statistics for Molecular Genetics	Cr. 2

d) Proposed new courses or modifications of existing courses

Please see Section 9.b.

e) Thesis and non-thesis options in the master's program

All Masters degrees in Bioinformatics and Computational Biology will require a thesis written on research in bioinformatics and computational biology.

f) Implications for related areas within the university

Participating students, faculty and departments will benefit from new and stimulating research interactions with faculty and students from other cooperating departments.
New course offerings in bioinformatics and computational biology will be of interest to advanced undergraduates and graduate students majoring in other disciplines, e.g., Genetics, Mathematics and Computer Science, all of which offer an Area of Specialization in Computational Molecular Biology.

g) Admissions standards for graduate programs

Admission standards will be above those recommended by the Graduate College. Evidence of strong quantitative skills and a solid background in either the biological or computational sciences will be required for admission.

9. General description of resources currently available and future resource needs

Financial support needed and currently being sought:

Moderate funding from Iowa State University will be necessary to establish salaries for initial BCB students, to provide student recruitment materials, to provide funds for invited speakers for BCB workshops, and to provide support for operation of the program. This funding could be supplied cooperatively by the participating Iowa State Colleges and Departments. The Graduate College will provide administrative housing and other support for BCB.

Strong efforts are currently underway to obtain off-campus funding for BCB graduate student salaries, lecture series and the annual retreat. As mentioned above, a $2.7 million NSF IGERT proposal, "Computational Molecular Biology Training Group" (D. Voytas, P.I.), has been recommended for funding. Pioneer Hi-Bred International has already committed up to $500,000 for graduate and postdoctoral fellowships in computational molecular biology. In addition, requests for Bioinformatics and Computational Biology graduate trainee support are included in several pending extramural and intramural grant proposals, including:

- NIH SBIR (Phase II) proposal: "An Integrated Field Stream Bioinformatics Data System," in collab. with VisualMetrics (V. Brendel, ISU P.I.)
- NSF KDI/KN proposal "Distributed Knowledge Network Tools for Bioinformatics" (V. Honavar, P.I.)
- NSF Plant Genome proposal "Functional Genomics of Plant Gene Regulatory Elements" (T. Peterson, P.I.)
- NSF proposal "High-throughput Mapping Tools for Maize Genomics" (P. Schnable, P.I.)

- ISU Biotechnology Council proposal: "New Bioinformatics Tools for Functional Genomics Research" (D.Dobbs, P.I.)

a) Faculty members, including vitae and publications relating to the program

Curricula vitae (a total of 33) are attached as Appendix I.

Agronomy: Jack Dekkers, Xun Gu, Patrick Schnable (both hold joint appts. in Zoology & Genetics)

Animal Science: Max Rothschild, Christopher Tuggle

Biochemistry, Biophysics and Molecular Biology: Amy Andreotti, Parag Chitnis

Botany: Eve Wurtele

Chemical Engineering: Peter Reilly

Computer Science: David Fernandez-Baca, Vasant Honavar, Leslie Miller

Mathematics: Dan Ashlock, James Cornette, Roger Maddux, Jonathan Smith, Steven Willson

Physics and Astronomy: Kai-Ming Ho, Jamie Morris

Plant Pathology: W. Allen Miller

Statistics: Alicia Carriquiry, Michael Daniels, Edward Pollak (joint appt. in Zoology & Genetics)

Veterinary Microbiology and Preventive Medicine: Susan Carpenter, Chris Minion

Zoology and Genetics: Volker Brendel, Hui-Hsien Chou (joint appt. in Computer Science), Drena Dobbs, Fred Janzen, John Mayfield, Gavin Naylor, Tom Peterson (joint appt. in Agronomy), Daniel Voytas

Founding Faculty of BCB are currently active in training graduate students and in teaching and research in bioinformatics and computational biology. During the last two years, Iowa State has been active and successful in hiring new faculty with expertise in bioinformatics and computational biology. During the 1997-98 academic year, three new faculty were hired by the Departments of Zoology & Genetics, Mathematics, Computer Science and Agronomy. Two structural biologists with computational interests were also hired by the Dept. of Biochemistry, Biophysics, and Molecular Biology. This year, the Department of Physics and Astronomy has offered a position to a Computational Structural Biologist, an additional hire in Bioinformatics is being made jointly by the Statistics and Zoology and Genetics Departments, and several new faculty are being hired by the Department of Computer Science.

b) Effects of the new courses on the work load of the present staff

All of the courses required by students majoring in BCB are already being offered at Iowa State University. These include the core course, Gen/Math/ComS 594, as well as several seminar courses (e.g., Gen/Math/ComS 590). Recent faculty hires will add at least three new advanced courses to the curriculum in Fall '99, including Stat 537 "Statistics for Molecular Genetics," Gen/ComS 555X "Software Methods for Biologists, and Gen 556X "Computational Genomics and Evolution."

c) Research facilities

Each faculty participant listed has the necessary research facilities and computer support required for training students in this area. These facilities have been equipped through federal grant support, industrial support, and university funding. Facilities available include fully-equipped molecular biology research laboratories and ISU's Biotechnology Facilities, including state-of-the-art facilities for DNA sequencing and synthesis, protein sequencing, peptide synthesis, plant transformation, NMR, etc. Additional support facilities for high-throughput DNA sequencing, microarray RNA expression analysis, high-throughput protein identification and analysis (proteomics), including robotics components and computational support facilities, have been obtained through funding for ISU's Iowa Genomic Frontier Center (IGF Center, P. Schnable, Director). Advanced computing facilities associated with the Depts. of Physics and Astronomy, Zoology and Genetics, Computer Science, the ISU Artificial Intelligence Research Laboratory, and Ames Laboratory are available for the proposed research.

In addition, funds for renovation of space in the Molecular Biology Building for a new Computational Biology Laboratory were approved by the Vice Provost of Research in the fall of 1998. This space will provide a research laboratory, classroom and meeting place for BCB students. The laboratory will accommodate computer workstations and servers and a reading room with bioinformatics and computational biology journals.

d) Library facilities (journals, documents, etc.) in the proposed area

The Iowa State University library maintains journals in the basic disciplines relevant to bioinformatics and computational biology , including biochemistry, genetics, molecular biology, statistics, mathematics, physics, and computer science. Additional library resources specific to computational biology will be needed whether or not the program is established. Iowa State currently has more than thirty faculty who work in this area; collections are adequate, but new books and journals will certainly be needed in the future to support this new field.

e) Supplies, field work, student recruitment, etc.

Funds for supplies are available through individual research groups. Funds for recruitment are currently available from the Graduate College and individual departments. Additional recruiting funds have been requested through an NSF IGERT Training Grant proposal.

10. Relationship of the proposed program to the strategic plans of the department, the college, and the university

The new major will bridge the gap between the traditional computational sciences (Computer Science, Mathematics, Physics, and Statistics) and the traditional molecular and structural biological sciences (Genetics, Molecular Biology, Biochemistry, and Biophysics). Developing a vigorous graduate major at this nexus will benefit existing programs by attracting new students and faculty. Graduate students in other majors will benefit by taking new courses, working with faculty in this emerging area, and by participating as a BCB minor, or by participating in journal clubs and seminars. Establishment of a premier program in Bioinformatics is a top university priority, and establishment of a formal degree granting program in Bioinformatics and Computational Biology is a critical component in this effort.

B. REGENTS PROGRAM REVIEW QUESTIONS

Degrees: Masters of Science and Doctor of Philosophy

Major in: Bioinformatics and Computational Biology (BCB)

1. Need

a. How will this proposed program further the education and curriculum needs of the students in this discipline?

There are rapidly increasing intellectual and employment opportunities at the interface between molecular biology and the computational sciences. This program recognizes the need for a formal educational structure for graduate student development in this area and is designed to attract outstanding new students to Iowa State University.

A graduate degree in Bioinformatics and Computational Biology from Iowa State University will provide students with skills required to access, manipulate and interpret biological sequence and structure data. The degree program ensures that students receive advanced training in both the biological and computational sciences, but allows students to focus on specific areas, such as gene prediction for functional genomics, or design of algorithms for sequence/structural alignment. BCB program graduates will be prepared for employment in either academic or industrial settings, in jobs ranging from genetic engineering of crop plants or animals in agribusiness, to drug design in the pharmaceutical industry, to research and teaching in a college, university, or biomedical institution.

b. How does it further the educational and curriculum needs of other units in the college or university?

The new major will bridge the gap between the traditional computational sciences (Computer Science, Mathematics, Physics, and Statistics) and the traditional molecular and structural biological sciences (Genetics, Molecular Biology, Biochemistry, and Biophysics). Developing a vigorous graduate major at this nexus will benefit existing programs by attracting new students and faculty. Graduate students in other majors will benefit by taking new courses, working with faculty in this emerging area, and by participating as a BCB minor, or by participating in journal clubs and seminars.

2.

a. What programs in this field of study are available in other colleges and universities in Iowa?

None. The University of Iowa advertises interdisciplinary biomolecular sciences research opportunities but they do not have a formal program, nor do they advertise a computational dimension in the programs offered.

b. With what representatives of these programs have you consulted in developing this proposal?

Provide a summary of the reactions of each institution consulted as well as the complete text of the responses. (See note 2a.)   None.

c. In what ways is this proposed program similar to those mentioned in 2a? In what ways is it different or does it have a different emphasis? (See note 2c.)   n/a

d. How does the proposed program supplement the current programs available? (See note 2d.) n/a

e. Has the possibility of some kind of interinstitutional program or other cooperative effort been explored? What are the results of this study? (See note 2e.)   n/a

f. List the Iowa institutions in which articulation agreements are being developed for the proposed program. (See note 2f.)   None.

g. Provide the Classification of Instructional Program (CIP) code for the proposed program. (See note 2g.)

There is no code for Bioinformatics. Related general areas are Computer Science (11.0701), Mathematics (27.0101), Statistics (27.0501), Genetics (26.0613), Molecular Biology (26.0402), Biophysics (26.0203).

3. Please estimate the enrollment in this program (for the next five years) as follows:

	1	2	3	4	5
a. Undergraduate					
Majors	0	0	0	0	0
Non-Majors	0	0	0	0	0
b. Graduate					
Majors	5	8	12	16	20
Non-Majors	2	4	4	4	4

c. On what basis were these estimates made?

There are several students in Genetics who would transfer to the new major if it existed. We are getting inquiries for BCB without having made any announcement of a new program. Actual numbers will be determined by resources available resources and student demand. Current internal sources and existing

grants should be able to support 4-5 new Ph.D. students per year. This would lead to a program of 20-25 students in five years.

d. What are the anticipated sources of these students? (See note 3d.)

Students will be drawn from Iowa, the US, and the world.

4. Please provide any available data or information on employment opportunities available to graduates of this program in Iowa and nationally.

We have no numerical data, but industry is hiring everyone with this kind training who will accept a job. Many universities are also beginning to announce hiring in this area. Employment opportunities should be excellent.

5. Are there accreditation standards for this program? If so, please provide a copy of the Accreditation standards. No.

a. What is the accreditation organization? n/a

b. What accreditation timetable is acceptable? n/a

6. Does the proposed program meet minimal national standards for the program, e.g., Council of Graduate Schools or other such bodies? (See note 6.) n/a

7. Please report any reactions of the Iowa Coordinating Council for Post-High School Education.

None as of 2/25/99.

Additional Resource Needs

1. Please estimate the probable marginal increases in expenditure that may be necessary as a result of the adoption of this program for the next three years.

Please note: rather than "marginal increases," the following table shows total program costs for each year.

Estimated Costs* (This table lists total program costs from institutional funds)			
	First Year	**Second Year**	**Third Year**
a. Faculty	0	0	0
b. Graduate Assistants	34,000	51,750	70,000
c. General Expense	10,000	10,000	10,000
d. Equipment	0	0	0

e. Library Resources	0	0	0
f. New Space Needs (est. amount and cost of new space and/or remodeled space	0	0	0
g. Computer Use	0	0	0
h. Other Resources (please explain)	0	0	0
TOTAL:	44,000	61,750	80,000

* Estimated Cost Notes:

a. <u>Faculty</u>: The graduate program will not hire any faculty. However, three new faculty were hired in AY 1998-99 in the Departments of Agronomy, Computer Science, and Zoology and Genetics and two new faculty were hired in the Department of Biochemistry, Biophysics and Molecular Biology. This year three or four new faculty job offers are being made by the Departments of Computer Science, Physics and Astronomy, Statistics, and Zoology and Genetics. Because this field is widely viewed as important and funding opportunities are outstanding, several additional new hires are very likely over the next few years, at least in Computer Science and Mathematics.

b. <u>Graduate Assistantships</u>: University funded assistantships will be focused on first year students. It is expected that upper level students will be funded by extramural grant or departmental funds. This model works well for existing interdepartmental programs at ISU. The University is mounting a major "Informatics Initiative." As part of this effort, resources have been allocated to support three new interdepartmental graduate initiatives. Bioinformatics and Computational Biology is one of the three (the others are "Complex Adaptive Systems" and "Computational Science and Engineering"). Funding for FY99 is $21,000, for FY00 is $136,000, and for FY01 is $200,000. These funds will support general program expenses, part-time office staff, and provide 5 research assistantships in FY00 and 10 research assistantships in FY01. The programs will compete for resources so that those attracting the best applicants will receive the most assistantships. In addition, a request has been made to allocate 4 research assistantships to the BCB major from Plant Science Initiative funds. Also, it is expected that significant support will come from extramural training grant funds. Bioinformatics faculty have submitted several interdisciplinary grant proposals, including a $2.7 million NSF IGERT proposal (Dan Voytas, P.I.) for a training grant in Computational Molecular Biology. Voytas recently received notification that the training grant has been recommended for funding. The CMB Training Grant will provide research assistantships for 4 new graduate students and 2 postdoctoral fellows each year for a total of 5 years (beginning July 1, 1999). In addition, Pioneer Hi-Bred International has committed up to $500,000 for graduate and postdoctoral fellowships in Computational Molecular Biology over a five year period (beginning Jan. 1999). Currently, 3 graduate students are supported by these funds. Several other proposals that include requests for graduate research assistantships in the area of Bioinformatics and Computational Biology are under review. Thus, the budget shown above is very conservative.

c. <u>General Expenses</u>: As explained above, funds have been allocated to establish three new programs. These funds will pay a part-time secretary, allow for the preparation of recruiting materials, and support general program expenses for all three programs.

d. <u>Equipment</u>: Students and faculty in this program will continually require new computer hardware. Various separate initiatives and grant proposals will provide an influx of new computers to Iowa State in the immediate future. Although necessary for the program to succeed, none of these expenditures will be made by the program itself. The program will succeed by cooperating with other units and stimulating other initiatives.

e. <u>Library</u>: Library resources in this area are needed whether or not the program is established. Iowa State currently has more than thirty faculty who work in this area; current collections are adequate, but new books and journals will certainly be needed in the future. This need is independent of whether or not this particular program is approved.

f. <u>New Space</u>. The program will not occupy significant new space devoted exclusively to BCB. Renovation of space in the Molecular Biology Building for a new Computational Biology Laboratory, funded by the Vice Provost of Research, is underway. This space will provide a research laboratory, classroom and meeting place for BCB students and faculty. The laboratory will accommodate computer workstations and servers and a reading room with bioinformatics and computational biology journals.

g. <u>Computer Use</u>: See 'd' above.

2. Describe the nature and justification for the additional resource needs.

Bioinformatics is recognized nationally and internationally as an important new field. Iowa State plans to be a major player in this area and has already invested significant resources. The proposed program builds on existing strengths in the biological and computational sciences. It supports and complements existing programs and other initiatives, including the Plant Sciences Initiative.

3. How is it anticipated that the additional resource needs will be provided? (For programs planning to use external grants, what would be the effect of the grant termination?)

Sources of the minimal necessary resources are described under Notes b and c (above). The Graduate College, various participating departments, and individual research grants will back up successful training grant initiatives.