Systems and Computational Biology Cluster

A hiring proposal from:

Department of Biochemistry and Molecular Biology
Danny Schnell, Head

Department of Biology
Brian O’Connor, Interim Chair

Department of Computer Science
Andrew Barto, Chair

Department of Mathematics and Statistics
George Avrunin, Head

Department of Microbiology
John Lopes, Head
We are experiencing a revolution in life sciences research that aims at integrating the vast array of information on individual molecules, genomes, signaling and metabolic pathways and population dynamics into very large, high-dimensional data sets that give a larger picture of how biological systems work. Interdisciplinary teams of experimental biologists (e.g. molecular and cellular biologists and health scientists), statisticians and computer scientists are combining their expertise in a new field, termed Systems Biology, to generate and analyze large scale data sets that reveal changes or adaptations in cells, whole organisms, and populations. By obtaining a global physiological/ecological picture, systems approaches are having profound impacts on human health, environmental sciences and energy research because they provide the basis for designing novel drug therapies, understanding global climate change, and optimizing agricultural practices. The Departments of Biochemistry and Molecular Biology (BMB), Biology, Computer Science (CS), Mathematics and Statistics, and Microbiology have formed a cluster in Systems and Computational Biology to build on the established excellence of interdisciplinary research in this area on the UMass, Amherst campus. Four of the departments within the cluster (BMB, CS, Math & Stats and Microbiology) are requesting new faculty hires within the Systems and Computational Biology Cluster. Biology is not requesting a hire from the cluster at this time.

This cluster was originally established during AY07-08 with the goal of building and enhancing the systems and computational biology faculty on campus. Four junior positions were requested under the original proposal. Two were funded and filled (BMB and Biology). The positions requested by CS and Math & Stats in the current proposal correspond to two junior positions that were recommended for funding in the future under the original request. The current proposal also broadens the cluster to include a request from Microbiology for a junior faculty position. The fourth position is for a senior faculty member in BMB. This position is supported by a Massachusetts Life Sciences Center (MLSC) New Faculty Startup Grant in Biomedical Systems Biology. To promote the inter-departmental collaborations within the cluster, the participating departments intend to include faculty from the other cluster departments on search committees for each position. The completion of these hires will broadly impact interdisciplinary research and training programs in the life sciences across campus by significantly strengthening the foundation of excellence in systems and computational biology.

Proven Excellence in the Area of Systems and Computational Biology

The Systems and Computational Biology Cluster combines the acknowledged strengths in molecular biology, genetics and biomedical research in the BMB, Biology, and Microbiology Departments with the expertise in analyzing highly complex systems using computational and statistical methods of the Computer Science and Mathematics & Statistics Departments. There is a strong base of excellence in these areas on campus that includes faculty in BMB (Peter Chien, Lila Gierasch, Jennifer Normanly and Scott Garman), Biology (Samuel Hazen, Magdelena Bezanilla and Elsbeth Walker), Chemical Engineering (Michael Henson), Computer Science (Hava Siegelmann, David Kulp and Oliver Brock), Electrical and Computer Engineering (Ramgopal Mettu), Math & Stats (Erin Conlon, Michael Lavine), Microbiology (Jeffrey Blanchard, Derek Lovley), and Public Health (Andrea Foulkes).

The cluster was successful in recruiting two junior faculty with outstanding records of accomplishment in the past two years. Samuel Hazen was recruited into the Biology Department in fall of 2008. His area of expertise is genomics and systems biology of plant biomass optimization as it applies to the bioenergy sector. Dr. Hazen has already secured a $2M Department of Energy grant and is leading a group of plant and microbial scientists in Biology, BMB, Microbiology and Plant Soil and Insect Sciences in developing model systems to study biomass properties. This group has pending proposals for NSF Major Research Instrumentation and DOE ARPA-E funding. Dr. Peter Chien will join the BMB Department in January 2010. His research focuses on complex networks that control progression through the cell cycle. His work is directly relevant to defining normal growth and development at the cellular level and associated biomedically related disease states (e.g. cancer). Dr. Chien will bring a prestigious five-year NIH Pathways to Independence Award (K99/R00) that recognizes his research accomplishments and provides support for his independent research program. The recruitment of these highly successful junior faculty members demonstrates the effectiveness of the cluster in building strength in systems and computational biology on campus.
The requested hires in this cluster are consistent with the current and future hiring plans of the four departments requesting positions. The AQAD report of BMB identified this area as one of exceptional potential for growth and increased success because of the unique combination of world class computer science and life sciences research on the Amherst campus. The report also saw computational and systems biology as two key areas that need to be further developed within the interdisciplinary graduate programs in which BMB participates (Molecular and Cellular Biology, Plant Biology). The AQAD report of Computer Science emphasized the role of computer science in other disciplines: “Increasingly, computer science is ‘outward facing,’ embracing connections with other disciplines by providing a new mode of scientific discovery.” For some years a priority of the UMass Department of Computer Science has been to build strength in Bioinformatics/Computational Biology. The department continues to place a high priority on building in these areas, and the position requested here will be a critical component of maintaining a strong presence of computational biology in CS. The 2005 AQAD review in Math & Stats stressed the need to build the statistics group within the department with a specific emphasis on the area of stochastic modeling in biology. Microbiology also has identified genomics and systems biology as an area of emphasis for hiring new faculty.

**Interdisciplinary/Transdisciplinary Nature of the Cluster**

Rather than focus on a specific area of research, this cluster encompasses faculty with interdisciplinary research interests that use common tools and approaches to address complex problems in the life sciences at the systems level. By definition, systems and computational biology are interdisciplinary as they involve collaborative interactions between life scientists, statisticians and computer scientists. The collaborative interactions of recent hires in these areas demonstrate the interdisciplinary nature of the field and the impact of systems and computational biology on a broad range of life science research. These include, collaborative projects on Protein Folding and Associated Diseases (Brock (CS) and Gierasch (BMB)), Optimization of Biomass Traits for Biofuels (Hazen, Caicedo, Bezanilla, Walker (Biology)), Blanchard, Leschine (Micro), Normanly, Schnell (BMB)), and Building Better Networks for Regulation of Metabolic Pathways (Hazen, Walker, Bezanilla (Biology), Blanchard, Leschine (Micro), Normanly, Schnell (BMB), Henson, and Roberts (Chem Engin)). As such, this cluster represents an emerging, transformative discipline that will impact a broad array of interdisciplinary initiatives on campus, including bioenergy (The Institute of Massachusetts Biofuels Research), cellular engineering (The Institute of Cellular Engineering and the proposed NSF Center for Building a Better Cell), and biophysics (Center for Biological Physics).

These positions will not only enhance research excellence on campus, but they will contribute to the critical mass of faculty in this area that is necessary to provide training in Computational Biology and Bioinformatics. Jeffrey Blanchard, Erin Conlon, and David Kulp developed a highly successful interdisciplinary graduate bioinformatics course, and there is considerable demand for additional courses in these areas. Furthermore, there is tremendous demand for individuals with this expertise in the biotechnology and pharmaceutical industry in the Commonwealth. A recent report from the MLSC and Massachusetts Biotechnology Council identified interdisciplinary training that included information technology as critical to training students to work in the life sciences industry.

**Willingness of Departments and Schools/Colleges to Support the Cluster**

The departments participating in this cluster, the Colleges of Natural Sciences and Mathematics (CNSM) and Natural Resources and the Environment (CNRE), and the campus have a demonstrated commitment to the long term support of this cluster (see attached letters of support). In addition to the commitment of devoting new faculty positions to this area, the BMB and Biology Departments and CNSM have contributed over $1M in start-up funding to support the first two junior faculty positions in this cluster (Hazen and Chien). In addition, CNSM and the campus have pledged the 1:1 match for the $750,000 New Faculty Startup Grant from MLSC that was awarded in the area of Biomedical Systems Biology in 2008. CS recently participated in a proposal to Five Colleges Inc. for a tenure system Five College Joint Faculty Appointee in bioinformatics to be shared between the CS departments of UMass and Mt. Holyoke. The first three years of the appointment will be paid in full by a Mellon grant to Five Colleges. After that, UMass and Mount Holyoke would each fund 1/2 the position. Pending details, this has been approved by the Five College Deans.
Established Leadership

The Systems and Computational Biology Cluster is led by a collaborative group of junior faculty that includes faculty from Computer Science (David Kulp and Oliver Brock) and Microbiology (Jeffrey Blanchard). In 2005, this group established SysBio, a western Massachusetts computational biology special interest group. CS efforts to build in this direction recently suffered setbacks upon learning that Prof. Kulp is resigning for personal reasons and Prof. Brock has left for a prestigious position at the Technical University of Berlin. With the likely loss of David Kulp and Oliver Brock, this cluster will lose two key leaders of the cluster. A major goal of this hiring request is to secure the future leadership in systems and computational biology by hiring additional junior faculty and recruit a highly visible senior faculty member using the existing MLSC startup grant to nucleate this cluster and provide the visibility and leadership to expand and develop both the research and teaching components of the cluster.

Ability to Attract Funding

A strong component of systems and computational biology is now recognized by funding agencies as essential to successful multi-investigator research proposals in genomics, cellular biology, and ecology. These new positions will contribute directly to one or more ongoing interdisciplinary initiatives across campus. In particular, we believe that it is critical to recruit faculty with strong backgrounds in computer science, bioinformatics and statistics to complement the recent hires in the life sciences. These faculty are envisioned to have the highest impact by introducing and developing computational methods that will have a broad impact in many areas of life science research. This will improve the research capacity of the interdisciplinary groups (e.g. TIMBR, ICE/CBBC and Center for Biophysics) and increase competitiveness for large multi-investigator grant opportunities in bioenergy (DOE Energy Biosciences Division, NSF, USDA Competitive Research Initiative), environmental science (NSF, DOE), biomedicine (NIH) and computational biology (NIH, NSF).

Plan for Assessing Effectiveness

Two major criteria will be used to assess the effectiveness of the investment in this cluster in addition to the standard assessment of excellence in research, teaching and service that accompany normal tenure/promotion evaluations. The first will be the submission of multi-investigator grant proposals involving the new hires as principal or co-principal investigators within two years of appointment. The second will be the establishment of a Five College Computational Biology and Bioinformatics program that will offer graduate and advanced undergraduate courses, such as: Introduction to Computational Biology: Biosequence Algorithms and Analysis, Introduction to Bioinformatics: Biologist Tools, Computational Protein Folding, Advanced Computational Biology, and Seminar: Topics in Bioinformatics & Computational Biology.

Description of Positions

Biochemistry and Molecular Biology: BMB has acknowledged strength in research on the structure and function of macromolecular machines within cells. We are proposing to hire a senior level (Associate or Full Professor) in the area of Biomedical Systems Biology. The senior hire is supported by an existing $750,000 New Faculty Startup Grant from the MLSC awarded in 2008. The grant is matched 1:1 by the campus and provides funding to support personnel, equipment, supplies and facilities renovation. BMB has identified laboratory (LGRT 902, 906, 907) and office space (LGRT 912) for this position.

A hire in this area builds on the foundation provided by the isolation and characterization of individual molecules and assemblies with the ability to reveal complex higher-order molecular interactions at the level of cells and organisms. The new faculty member will join a group of BMB faculty that incorporate systems and computational methods into their research programs, including Lila Gierasch (protein folding and associated diseases), Jennifer Normanly (metabolic profiling and control of metabolic pathways), and Peter Chien (regulatory networks controlling the cell cycle). These faculty generate ~$1.5M annually in research activity. We seek to hire someone who brings an integrated view to our Department, whether with a focus on biophysics, computational biology, cell biology, or cellular engineering.

This hire is seen as key to the next stage in development of the cluster. Our goal is to recruit a senior faculty member with an established record of exceptional achievement in systems or computational
biology to provide visibility and leadership in recruiting and retaining junior faculty in systems and computational biology within the five departments. The hire also matches the goals of Governor Patrick’s $1B Life Sciences Initiative as demonstrated by the MLSC award. We anticipate that this individual will contribute directly to the multi-investigator initiatives in the Center for Biophysics or the Institute of Cellular Engineering/Center for Building a Better Cell.

**Computer Science:** In addition to supplying practical tools, Computer Science develops fundamental theory and conceptual tools for describing and understanding complex systems. Computer Science is therefore an essential component of a profoundly interdisciplinary whole-systems approach to the study of complex interactions in biological systems, providing tools that are unique to the computational perspective. It is widely recognized that productive research in a range of biological fields increasingly requires teams of researchers with high levels of competence in mathematics, computer science, and biology. The Department of Computer Science will search for a computer scientist engaged in high-impact research in computational systems biology. This requires core strength in computational/statistical modeling, graphical models, and machine learning, with a track record of applications in systems biology and collaborative research between computational and life sciences. The need is to support activities such as clustering, classification, prediction, information extraction and data mining using graphical models and other techniques incorporating biological knowledge and constraints. A hire in this area as part of this cluster, together with the current strength of the department in machine learning, will make it possible to attract exceptionally strong candidates for this position.

The person filling this position will expand the current strengths of the CS department in machine learning (Barto, Grupen, Learned-Miller, Mahadevan), information extraction (McCallum), and data mining (Jensen) by greatly augmenting our ability to extend these and related methods to the life sciences in support of interdisciplinary projects across the college. He/She will also augment our current strengths in biological dynamic system modeling (Siegelmann) and computational geometry (Streinu, Prof. Smith College and Adjunct Prof. in UMass CS) by providing missing links to machine learning. Additionally, our hire in this area will complement the expertise of the cluster hire in Mathematics and Statistics by sharing interest and expertise in statistical methods while bringing to the mix expertise in the representational and algorithmic techniques being developed by machine learning researchers.

The CS Department has a remarkable record hiring new faculty who are successful in attracting grant funding. 100% of recent hires (since 9/1/02) are lead PIs on sponsored research funding by their third year. Third year funding averages over $200,000 for those hired as assistant professors, and almost $1,000,000 for those hired as associate professors. Fourth and fifth year funding averages $350,000 for assistants, and holds fairly constant at about $1 million for associates. Over the same period of time (FY03 - FY09), average annual funding for tenured faculty members with any sponsored research was $450,000. Average annual funding for research track faculty and research scientists was $300,000. The basic start-up package is approximately $250,000. If additional lab equipment is needed, the cost may rise by $10,000 - $50,000. This is, on average, recouped within five years for assistant professors, and much faster for associates. Hiring CS faculty is in excellent investment. Space is available for this hire in the CS building and the CS space in the low-rise LGRC.

**Mathematics and Statistics:** One of the fundamental challenges in contemporary science is the analysis of enormous and very high-dimensional data sets arising from modern data collection techniques and very large computer simulations. Such new methodologies as micro-array gene expression experiments, large sensor networks, and automated sky surveys, as well as the rapidly increasing use of cheap supercomputers that have made large-scale computer simulation widely available to scientists in all disciplines, present exciting new problems in core areas of statistics and computer science and are part of the impetus leading to a new discipline of informatics.

We would seek a statistician with active research interests in statistical and computational methods for large and complex data such as those arising in the study of complex systems in Computational Biology. The successful candidate would have broad training in modern statistical methods, including Bayesian theory, as well as more traditional theory and methods. This expertise would complement current department strength in bioinformatics (Erin Conlon) and in statistical methods in neuroscience (Michael Lavine), as well as work on systems biology data and on the relations between high-dimensional
molecular and cellular data and measures of disease progression in Public Health (Raji Balasubramanian and Andrea Foulkes).

Rebuilding the statistics group was one of the key recommendations of the 2005 AQAD review of the Department of Mathematics and Statistics and bioinformatics was one of the areas within statistics specifically recommended as part of the life sciences thrust of the College of Natural Sciences and Mathematics. The AQAD review committee also noted that the future of statistics on this campus “lies in serious application and consequential scientific interaction.” This proposal is a major step in that direction.

While statisticians are in high demand, this is particularly so for those with interests in computational biology and complex systems. This cluster, however, should make UMass a very attractive option and the new positions in this cluster would form the beginning of a critical mass in a tremendously exciting area. We would also expect connections to existing computational research in PSE, Astronomy, and other departments across campus. The great potential for interdisciplinary research will certainly contribute to large-scale grant applications in a field with many funding opportunities.

Startup funds in statistics have traditionally been very modest, totaling approximately $30,000. They have been split equally between the department and the college. A statistician interested in large-scale computational analysis would very likely require a considerably higher startup, perhaps on the order of $100,000, but still very much less than typical in a laboratory science. The department would be able to contribute $15,000-20,000 for this purpose. No additional space would be required as a result of this appointment.

**Microbiology:** The Microbiology Department has become a leader in emerging fields of Microbial Systems Biology and Environmental Genomics. Our internationally recognized program competes directly for graduate students with programs at UC-Berkeley, University of Washington, Wisconsin, Michigan State, MIT and Harvard. The Microbiology graduate training program includes courses in Bioinformatics, Bioinformatics Programming Lab, Systems Biology Project Club, and Special Topics in Bioinformatics and summer workshops in Microarray Analysis and Microbial Ecology Bioinformatics.

Microbiology has demonstrated success in developing new faculty in this area. A new faculty mentoring program, initiated by John Lopes, helped get Jeff Blanchard's research program off the ground. His research is now funded by the USDA, DOE and has attracted over $500K in sponsored research agreements from industry. Three of his patent applications were purchased by industry resulting in $150K in revenue to the University and a greater stake in a biofuels company. In just three years, Pat Schloss has been awarded grants from NSF's Biological Databases & Informatics program for Developing Bioinformatics Recipes for Microbial Ecology, the Northeast Sun Grant Initiative for a metagenomics project and a two-year subcontract to modify his "dotur" and "sons" computer programs to accommodate the massive datasets generated by pyrosequencing technologies. Blanchard and Schloss (in collaboration with Sam Hazen and Ana Caicedo) recently put together a $1.5M NSF equipment proposal to bring genomics instrumentation to our campus. Our department's prominence in this area is underscored by Derek Lovley's research laboratory. His laboratory attracts several million dollars in funding per year and incorporates Genomics and Systems Biology methods.

The influence of genomic technologies continues to rapidly expand with the advent of microarray and next generation sequencing technologies and genomics are now a part of all life science-related programs. The scale and reach of genomics has created a need to continue to hire new faculty in this area. We are proposing hiring, at the Assistant Professor level, a new faculty member that develops and/or employs computational methods for emerging genomic technologies. Laboratory and office space in Morrill I N134 or the NSB will be available for the new faculty member. We anticipate that the planned new Genomics, Proteomics, and Metabolomics Core facility in the New Science Building will be a major recruiting tool and that the person could contribute to the further development of the facilities. The new faculty member would benefit from the Institute for Cellular Engineering training grant and the proposed NSF Center for Building a Better Cell.