Accelerating mathematical-biological linkages: Report of a joint NSF-NIH workshop

February 2003

Margaret A. Palmer[1], Peter Arzberger[2], Joel E. Cohen[3], Alan Hastings[4], Robert D. Holt[5], Jennifer L. Morse[1], De Witt Sumners[6], and Zan Luthey-Schulten[7]

[1]University of Maryland, College Park, MD 20742-4454
[2]University of California, San Diego, CA 92903-0043
[3]Rockefeller University and Columbia University, New York, NY 10021-6399
[4]University of California, Davis, CA 95616
[5]University of Florida, Gainesville, FL 32611-8525
[6]Florida State University, Tallahassee, FL 32306-4510
[7]University of Illinois at Urbana-Champaign, Urbana, IL 61801

**Accelerating mathematical-biological linkages:
Report of a joint NSF-NIH workshop**

held 12-13 February 2003

The workshop was supported by grants from
the National Science Foundation
and
the National Institutes of Health.

This document is available in electronic format:

http://www.bisti.nih.gov/mathregistration/report.pdf

Table of Contents

Executive Summary and Recommendations

Complex scientific questions are increasingly being addressed through interdisciplinary approaches. Biology and mathematics have historically benefited from exchange and collaboration, and they lend themselves to further integration and synergy today. New advances in mathematics can grow from addressing the complexities of biological phenomena. In turn, fundamental and emerging biological problems can be better understood and predicted with the aid of new mathematical, computational, and modeling tools and theories. The time is ripe for accelerating progress in understanding and predicting important biological phenomena by using existing mathematical tools, by developing new mathematical approaches, and most importantly by forging new partnerships between mathematicians and biologists to investigate tough problems.

The joint NSF-NIH workshop, "Accelerating mathematical-biological linkages," was held in February 2003 to highlight the opportunities and intellectual challenges present at the mathematics-biology interface, and to challenge the institutional, cultural, and educational barriers to these essential and fruitful partnerships.

The symposium at NIH's Natcher Auditorium featured a plenary address by Dr. Joel E. Cohen and explored cutting-edge mathematical-biological work in three areas: conservation biology, cell structure and function, and bioinformatics and computational problems. This symposium highlighted the wealth of opportunities available at the math-biology interface, the need for mathematics in solving complex biological problems, and the enthusiasm of researchers from many disciplines to engage in this type of work.

Participants in the ensuing half-day workshop joined one of three working groups: Institutional Issues, Education and Training, and Strengthening Ties among the Researchers. Each group was charged with developing and articulating critical actions needed to enhance the math-bio linkage. The rapporteur from each group compiled the emergent ideas of the discussion and guided the development of action items. The three working groups represented areas in which bold initiatives and change could pave the way for new opportunities and results at the math-biology interface. Many ideas were discussed and many recommendations made; some of the recommendations are directed at NSF and NIH while many others are directed at scientists and mathematicians or at academic and research bodies.

Institutional Issues Group Recommendations

<u>Institutional Action Item 1:</u> Two large scale initiatives should be aggressively pursued by both NIH and NSF: 1) large competition(s) to foster new biological breakthroughs through high impact research made possible by mathematical applications and theory; 2) the creation of a *national* research center for Biological Research Interfacing with Mathematics, to review and fund proposals that would bring multidisciplinary groups of scientists from around the world to the center to work together on major unsolved issues in both fundamental and applied arenas.

<u>Institutional Action Item 2:</u> Establish and provide sustained support for a federated network of data resources for biological information from the genomic and proteomic levels through organismal organization and up to environmental information to better enable biomedical and biological multi-scale integrative research. This initiative should build on existing community databases and resources to take advantage of investments and research in cyber-infrastructure, to create an integrated (but distributed) set of resource for the community.

<u>Institutional Action Item 3:</u> Formalize publishing standards for model development to improve the quality and level of access to data and models.

<u>Institutional Action Item 4:</u> Create a federation to coordinate Math/Biology societies, to provide a united voice for major funding initiatives, for enhancing cross-disciplinary interaction, and for interdisciplinary undergraduate and graduate education.

<u>Institutional Action Item 5:</u> Examine existing programs within NSF and NIH, and assess what changes could be readily achieved to improve research and education at the math-bio interface. Develop a subprogram within IGERT, focused on the issue of math-bio linkage training.

Education and Training Group Recommendations

<u>Education Action Item 1:</u> Create new interdisciplinary postdoctoral programs at the interface between mathematics/statistics/computer science and biology/biomedicine.

Education Action Item 2: Convene a high level (e.g., NAS) workshop that attracts department chairs from biology, math, computer science and statistics departments to identify means to promote cross-disciplinary curricula and training appropriate for the kind of biological questions being explored, including support for double degrees in math and biology. Recommended for Fall 2003 (*urgent*).

Education Action Item 3: Create joint faculty positions at the interface between mathematics/statistics and biology/biomedical sciences, with appropriate mentoring and clear expectations for interdisciplinary work.

Education Action Item 4: Develop summer math/bio programs for high school math, computer science, and biology teachers and their students, to introduce them together to the interface between math and biology.

Education Action Item 5: Develop K-12 educational material for the math/biology interface.

Education Action Item 6: Develop a central website and a listserv to disseminate information on research and training opportunities at the math/biology interface.

Strengthening Ties among the Researchers Group Recommendations

Researchers Action Item 1: Compile a list of the top ten most challenging and promising areas in mathematical biology, including modeling of multilevel systems, integrating probabilistic theories, data mining and inference, and computational tools.

Researchers Action Item 2: Organize a national meeting in Washington D.C. on Mathematical Biology, to showcase (over 3 – 5 days) successes where mathematics and computations have helped solve important problems in medicine and to motivate new directions and opportunities on all levels of biology from cells to organs to ecology to the biosphere.

Researchers Action Item 3: Develop a series of advanced 'educational' workshops for mathematicians and biologists to learn about the other discipline, to facilitate communication and interaction.

Researchers Action Item 4: Strengthen the publication ties between mathematics and biology, through merged databases online journal access, and foster the publication of interdisciplinary papers.

Motivation for the Workshop

Complex scientific questions are increasingly being addressed through interdisciplinary approaches. Interdisciplinary collaboration allows scientists to effectively apply a wider array of tools and theories in solving problems that bridge traditional disciplinary boundaries. Each discipline is enriched by these collaborations: new theories, new analytical tools, and new frontiers are developed.

Biology and mathematics are two disciplines that have historically benefited from exchange and collaboration and that lend themselves to further integration and synergy today. New mathematical approaches, aided by new computational and modeling tools, can be applied to fundamental biological problems (Hastings and Palmer 2003).

The time is ripe for accelerating progress in understanding and predicting important biological phenomena by using existing mathematical tools, by developing new mathematical approaches, but most importantly by forging new partnerships between mathematicians and biologists to investigate tough problems. The burgeoning base of theoretical and empirical work, made possible by new methods and technologies, in particular the advances in information science given rise to cyberinfrastructure (Aktins et al 2003; Wooley et al 2003), is providing the impetus to reach robust answers to the major integrative biological challenges faced by our society. Meeting those challenges will require new efforts to train mathematicians and biologists to work at the interface of two disciplines and/or work collaboratively much more frequently (Levin et al. 1992, Levin et al. 1997, and Hastings et al. 2002).

We need a new generation of empiricists with stronger quantitative skills and theoreticians with a detailed understanding of the empirical structure of biological processes. It is very difficult for a single scientist – biologist or mathematician -- to develop the depth needed to address the questions at the frontiers. Instead, we must foster increased research collaboration among biologists and integrative biologists who are trained to understand the language of one another. The mathematical and statistical constructs used in much biological research come mostly from linear theory. In contrast, many biological processes are governed by non-linear dynamics, threshold effects, and more complex relational structures. Both the biological *and* mathematical sciences will benefit from additional collaboration (Hastings et al. 2002).

Workshop Purpose and Organization

The purpose of this workshop was to highlight the opportunities and intellectual challenges present at the math-bio interface, and to challenge the institutional, cultural, and educational barriers to these essential and fruitful partnerships.   The workshop consisted of a day-long symposium (February 12) followed by a half day (February 13) in which small working groups identified key needs to move mathematical-biological linkages forward.   Linkages are defined broadly to include collaborations among mathematicians and biologists, educational and training (at all stages from graduate school to mid-career scientists) opportunities, new research initiatives, society and small group projects as well as other activities.

More than 170 biologists, engineers, mathematicians, computer scientists, physicians, educators, and administrators attended the symposium on "Accelerating Mathematical-Biological Linkages," held on February 12, 2003, at the Natcher Auditorium on the National Institutes of Health campus in Bethesda, Maryland (Appendix 1).   More than 60 attendees participated in the ensuing half-day workshop by joining one of three working groups: Education and Training, Institutional Action, and Collaborations.   Each group brainstormed throughout the morning to develop and articulate critical actions needed to enhance the math-bio linkage.   Reports were then made to all participants and discussion followed.

Day 1: Symposium

The auditorium was filled to capacity – registration closed at 150 but there were an additional 20 – 25 attendees. The symposium was convened by Dr. Margaret Palmer, of the University of Maryland, who opened the meeting by highlighting the urgent need for accelerated linkages between biologists and mathematicians.   She then introduced the rest of the event's organizing committee (Dr. Sam Scheiner and Dr. Mike Steuerwalt of the National Science Foundation; Dr. Jim Cassatt, Dr. Mike Marron, and Dr. John Whitmarsh, of the National Institutes of Health).

Dr. Mary Clutter, Assistant Director of the National Science Foundation, offered her most enthusiastic support of the goals of the conference and discussed several areas in the life sciences in which exciting and fundamentally quantitative opportunities exist. The areas she highlighted included topics from the environmental sciences as well as genomics and

proteomics. Dr. Clutter called for a re-evaluation of cultural, educational, and disciplinary barriers that impede communication and interaction between disciplines.

Dr. Mike Gottesman, Deputy Director of Intramural Research at the National Institutes of Health, followed by underlining the importance of mathematics in the development of ecological and biological theories, and emphasized that in the future, biological research would likely require a quantitative component to be relevant. He expressed the strong support of the National Institutes of Health leadership for enhancing research at the mathematical-biological interface.

Dr. Joel Cohen, of Rockefeller and Columbia Universities, presented his keynote address, "Mathematics Is Biology's Next Microscope, Only Better; Biology is Mathematics' Next Physics, Only Better." Drawing from Descartes, Darwin, Harvey, and many other pivotal figures, Dr. Cohen provided a compelling overview of key historical contributions, both technical and conceptual, by which mathematics can contribute to the advancement of biological knowledge and theory. He also discussed how complex biological problems, from molecular to global-level processes, have provided new topics for mathematical research and development. By summarizing both new opportunities for and potential impediments to these math-bio collaborations, Dr. Cohen generated enthusiasm and support for this burgeoning field of science.

Following the plenary address by Dr. Cohen, the symposium turned to exploring mathematical-biological work in three areas: conservation biology, cell structure and function, and bioinformatics and computational problems. These areas were not meant to be exhaustive of the many biological areas that are ripe for quantitative work, but simply meant to be illustrative.

In the conservation biology session, Dr. Michael Neubert, of Woods Hole Oceanographic Institution, discussed his experience as an applied mathematician collaborating with a plant ecologist, and talked about his scientific findings in applying integro-difference equations to the dispersal of an invasive shrub species. His insight into the cultural and disciplinary differences that arose during the course of the work was particularly valuable. Dr. Leah Gerber, of Arizona State University, spoke about her work as an applied population biologist; she integrates cutting-edge epidemiological models into population viability analysis for management of threatened populations and to determine extinction risk.   Her talk also demonstrated the importance of complex mathematical models in defining

uncertainties in different conservation scenarios, particularly in understanding sea otter populations in California.

The afternoon session began with presentations on cell structure and function. Dr. Alex Mogilner, of the University of California, Davis presented, "Mathematics and a force balance model to understand mitotic spindles," in which he integrated physics approaches and mathematical modeling to explain the role of microtubules in generating key forces that propel the mechanism of cell division. He reported successful results from interdisciplinary training in his Cell and Computational Biology laboratory, which links theory and experimentation, and brings together mathematicians, cell biologists, and biophysicists. Dr. Sharon Lubkin, of North Carolina State University, discussed her work in developmental biology, which uses mathematical models and experiments to understand and predict branching structures in tissues. Her modeling work is also performed in collaboration with experimental biologists.

The final session of the symposium highlighted research in bioinformatics and computational problems. Dr. Chuck Delwiche, of the University of Maryland, spoke about his applications of bioinformatics to the study of gene transfers into plastid genomes in dinoflagellates. His lab uses methods in molecular systematics and genomics, including cDNA screening and selective sequencing of genomic DNA, to understand the phylogenetic history of dinoflagellates. Dr. Monica Hurdal, of Florida State University, rounded out the afternoon by discussing a medical perspective on the math-bio interface in her presentation, "Mapping the Human Brain with Mathematical and Computational Modeling." She is involved in the Human Brain Project and is using MRI (magnetic resonance imaging) scans and EEG (electroencephalography) data to understand human brain functions. Innovative approaches to computer modeling and classical geometric theories are both key aspects of her work.

The day's talks were summed up by rapporteurs for each session: Drs. Bob Holt (University of Florida), Zan Luthey-Schulten (University of Illinois), and De Witt Sumners (Florida State University). A lively open discussion with audience participation was then led by the hosts and rapporteurs. Members of the audience actively exchanged ideas during the discussion, asking questions and voicing opinions on a wide range of issues including curriculum development, national security, collaboration and data sharing, training of graduates and undergraduates, and how to best bring about the cultural and disciplinary changes that stand in the way of this new initiative.

Day 2: Working groups and recommendations

The second day was quite informal, beginning with a charge to participants to develop break-out groups to focus on one of three areas, to generate ideas that could lead to concrete programs or initiatives: Education and Training, Institutional Issues, and Research Collaboration Issues. Each group was led by a rapporteur who compiled the emergent ideas of the discussion and guided the development of action items.

Institutional Action Items
(Rapporteur: Bob Holt)

Institutional Action Item 1: Joint NSF-NIH Initiatives
1a. <u>Concept:</u> Develop a plan for major programs deserving substantial new funding, specifically to target new biological breakthroughs made possible by mathematical applications and theory. This initiative would move well beyond the current NSF/NIGMS Initiative in Mathematical Biology to a much broader programs jointly directed by NIH and NSF.

Rationale: There is an urgent need to solve a proliferating array of major, global scale problems, ranging from ecological and evolutionary dimensions of disease transmission and spread, to analyses of disruptions of neurological mechanisms and repair due to environmental factors, to the management of enormous, burgeoning genomic databases in an efficient fashion. Addressing these problems is difficult (if not impossible) without the application of sophisticated theoretical, mathematical, and computational approaches.

In principle, mathematical, computational, and statistical tools and theory (both existing and created) can be brought to bear in ways never before possible, and major breakthroughs are to be expected in dealing with these crucial problems. But this requires substantial societal support in strengthening ties between the communities of the mathematical sciences and biology, well beyond that provided by existing governmental programs.

Educating a new generation of scientists specifically focused on this interface is essential, and requires substantial support. Joint sponsorship by NSF and NIH of a major initiative funding research and education in this area would send a powerful message about the importance and urgency of this mission. At a pragmatic level, a joint initiative would greatly enhance the synergistic and complementary kinds of science typically funded by NSF and NIH respectively, and help cut across the suture zones currently separating the relevant disciplines.

Proposed Plan: The need for funding for work at the math-bio interface is so urgent that we recommend a special joint NSF-NIH initiative that is extremely large in scope (hundreds of millions of dollars). The planning for such an initiative should from the outset involve both agencies.

If such an initiative were successful, it would specifically broaden the scope of the current NSF/NIGMS Math-Bio study section to include representation/funding from: NIBIB (Bioimaging/Biotech), NHGRI (Human Genome Res. Inst.), NCRR (research resources), NIAID (infectious diseases) as well as many programs at NSF. So rather than an *ad hoc* special study section, one could envisage a permanent chartered study section. This new program should be explicitly charged with using a substantial portion of their funds for high-risk (but potentially high impact) research.

1b. <u>Concept:</u> Create a high visibility, national center specifically devoted to research at the interface between mathematics and biology. It should provide funding and logistical support for the national community of biologists and mathematicians (i.e., this should not be a center that primarily serves a few institutions or topical areas). As a means toward this end, workshops should be held to call for proposals for such a center.

Rationale: Several existing centers provide models of highly effective mechanisms for fostering novel interdisciplinary approaches. Within ecology, the National Center for Ecological Analysis and Synthesis (NCEAS, University of California, Santa Barbara) has provided an exemplary service, bringing together groups of scientists who historically had never worked together, so as to address major unsolved issues in both fundamental and applied arenas. Many of the synthetic activities initiated at NCEAS have continued in ongoing collaborations, grant proposals and the like; in effect, NCEAS has created a large and growing web of interlinked research activity that is much larger in social scope than

was previously possible. In like manner, a center specifically focused on cutting across the interface between mathematics and biology could act as a powerful catalyst for research and educational efforts.

Proposed Plan: National funding by a consortium of agencies (e.g., NSF, NIH) will be required for the creation of a center. However, for a 'think tank' to work, there needs to be a mechanism ensuring the whole-hearted support of the concerned communities. A "bottom-up" approach is to have a series of workshops aimed specifically at honing the mission and structure of a research center. A concrete product of these workshops would be a statement of the content of a call for proposals from universities, existing research institutes, or consortia for housing a Center in Mathematical/Quantitative/ Computational/Theoretical Biology.

Institutional Action Item 2: Databases

Concept: Work to establish integrated data resources for biological information at levels building up from the genomic and proteomic levels, through functional aspects of organismal functioning, to higher levels of organization (including environmental information), building upon new tools and technologies that are creating the cyber-infrastructure. Such federated resources should be developed with guaranteed support that permits it to integrate and make accessible data drawn from scientific efforts around the globe. The multi-agency commitment should be such that investigators can be confident of the long-range durability of the resources.

Rationale: Many of the next generation's advances in biological understanding are likely to involve the integration of vast bodies of information from many levels of biological organization, including information defining the context of environmental drivers of organismal functioning.   The existence of accessible databases can create the opportunity for unexpected, exciting developments in scientific understanding. For instance, gene sequencing databases have permitted systematic evolutionary biologists to revolutionize understanding of the history of life for many organisms.

The development, application, and evolution of sophisticated mathematical and statistical models of biological systems would be greatly facilitated were investigators able to draw on a reliable, carefully structured database spanning the biological sciences. However, creating and managing such a synthetic data base system, covering not only human genomics, but the functional information on all aspects of

organismal functioning (the 'physiome' and 'niche' of each species under study) across the diversity of life is an immense, long-range effort.

One of the great challenges is that the rapid development of technology inexorably makes a database created today into tomorrow's potential electronic white elephant. Broad national commitment would ensure the long-term viability and usability of these resources. Furthermore, models are being built now by communities that take advantage of the new cyber-infrastructure tools and reflect the need to have curation of the data done by experts at local levels. One such example is the Biomedical Informatics Research Network (BIRN), funded by National Centers for Research Resources (see http://www.nbirn.net).

Proposed Plan: There are many agencies currently involved in issues of database creation and management. An initial effort should be made to encourage natural moves to federate these already existing, ongoing activities. This could be achieved by organizing workshops including NLM, NSF, NIH, ODD, DOE, the Smithsonian, NPS, NASA, EPA, and so forth. These workshops should incorporate both agency representatives and working scientists.

There are many crucial issues which could be addressed in these workshops, including metadata standards, national security concerns, the proprietary 'ownership' of data, problems of validation and accessibility, and interagency cooperation. A long-term goal of these workshops and this initiative in general might be to make a proposal to Congress that would provide funding to coordinate the establishment of a sustained federation of strategic biological data in data resources and data bases, to better enable biomedical and biological multi-scale integrative research. This initiative should build on existing community databases and resources to take advantage of investments and research in cyber-infrastructure, to create an integrated resource for the community.

Institutional Action Item 3: Publishing Standards
Concept: Call for the formalization of publishing standards by journals regarding model structures and specifications, and develop mechanisms for the maintenance of specified models on electronic databases. The publishing standards should be complete, including equations, all parameters and units, initial and boundary conditions, and the information sources used to supply the data for parameter specification and validation.

Rationale: It is difficult for investigators at present to rapidly compare and synthesize published models. This slows progress in the field. A systematic, formalized specification of model structures and information sources could greatly facilitate integration among models developed by different investigators, particularly working at different levels of biological organization.

Proposed Plan: The first step would be to have a workshop involving representatives of the major journals and societies in the field, to identify the major issues and plan a course of action. Prior to such a workshop, an electronic call should go out to the community, soliciting input and suggestions. This workshop should be jointly sponsored by NSF and NIH.

Institutional Action Item 4: Federation to Coordinate Mathematical/Biological Societies

Concept: **Investigate the potential for a federation to coordinate and streamline functions of the many societies interested in our subject.**

Rationale: Many societies are already involved in research and education at the math-bio interface. A sampling of such societies includes the Society for Mathematical Biology, the Theoretical Ecology section of the Ecological Society of America, the 5MB Biophysical Society, ESMTB (European Society of Mathematical and Theoretical Biology, ISCB (Inti. Soc.Compo Bioi.) Japan Society, SIAM -Life Sciences, IEEE -EMBS and BMES. A federation of societies could provide a united voice for major funding initiatives, for instance those aimed at enhancing cross-disciplinary dimensions of undergraduate and graduate education. A federation could also help legitimize work in this interdisciplinary field and thus provide support for young researchers.

Proposed Plan: Organize a workshop at a national venue, such as the annual AAAS meeting. This could be coordinated with a symposium, focused on the math-bio interface (comparable in spirit to the recent NSF-NIH workshop in Bethesda).

Institutional Action Item 5: Plan to Foster Support for Interdisciplinary Work at the Institutional Level

Concept: Commission a study of universities where interdisciplinary work thrives. This study would identify "best practices" -structural

and administrative features that promote successful interdisciplinary ventures. This study would necessarily involve sociologists and historians of science and management experts.

Rationale: It has long been recognized that interdisciplinary approaches are required to address many significant conceptual and applied problems, and numerous attempts have been made to foster interdisciplinary studies. So it is likely that there is a great deal of accumulated wisdom about the nature of effective mechanisms (institutional and otherwise) that facilitate – or do not – such activities.   A clear understanding of effective organizational and management practices would greatly facilitate efforts to foster interactions among biologists, mathematicians, statisticians, and computational scientists.

Proposed Plan: Analyzing organizational practices is itself a significant intellectual study. One possibility would be for NSF and NIH to jointly propose to NAS that the NRC carry out a study on interdisciplinary "best practices".   Another option would be to approach foundations such as Hughes and Sloan, which have considerable prior interest in biomedical research and related areas, to carry out such a study.

Institutional Action Item 6:   Specific NSF-NIH Change Initiatives
Concept: Examine existing programs within NSF and NIH, and assess what changes could be readily achieved to improve research and education at the math-bio interface. Develop a subprogram within IGERT, focused on the issue of math-bio linkage training.

Rationale: Both NSF and NIH already have a wide range of existing programs that can be exploited to increase research and education. The relevant community of scientists needs to become savvier in identifying and exploiting these existing programs, and the agencies need to do a better job of clearly advertising the existence of these programs to the community.

Proposed Plan: Program officers should address ways to make the relevant scientific communities more aware of the existence of programs relevant to the math-bio interface. Beyond simply advertising existing programs, the agencies should seriously consider modifications to these programs to foster research and education across the math-bio interface. For instance, at NSF, within the IGERT program, develop an explicit, named program aimed at training that links mathematics and biology. In

like manner, for both NSF and NIH, create then advertise targeted supplements to existing grants (e.g., for mathematicians to be added onto ongoing research projects). This could be named CEMB, for "Collaborative Experiences for Mathematicians and Biologists," provided as 1 semester of salary for release time, as a supplement to an existing grant.

Education and Training Action Items
(Rapporteur: De Witt Sumners)

Education Action Item 1: New Interdisciplinary Postdoctoral Programs
Concept*:* **Outline a program for "group of postdocs" awards (rather than and in addition to individual postdoc awards), aimed at expanding the number of individuals at the interface between mathematics/statistics/computer science and biology/biomedicine. The group concept is meant to create a critical mass of postdoctoral scientists interested in a central question. It is envisioned that this program would augment (not replace) other training activities such as graduate programs in interdisciplinary areas.**

Rationale*:* Postdoctoral fellowships provide a unique opportunity for interdisciplinary training at the interface between the mathematical and biological sciences. Moreover, postdoctoral researchers are valuable in the research enterprise. They have demonstrated their ability to perform independent research at a significantly deep level. At the current time, most support for postdoctoral researchers is either through individual fellowships, associated with a specific research project, or in some cases associated with training programs for graduate students. Some exceptions include institutes such as the mathematics research institutes or the National Center for Ecological Synthesis and Analysis. There is evidence from programs such as the La Jolla Interfaces in Science (LJIS, http://ljis.ucsd.edu) program, DIMACS Special Focus in Mathematical Support for Molecular Biology (http://dimacs.rutgers.edu/SpecialYears/1994_1995/index.html), and the Program in Mathematics and Molecular Biology (http://web.math.fsu.edu/~pmmb/) that postdocs can be provided opportunities to work at interface areas between the mathematical and biological sciences, learn a new discipline, and be productive in a relatively short period of time.

"The most effective way to encourage interactions between mathematicians and computer scientists on the one hand, and biologists on the other, is through direct co-involvement with a particular problem. This applies at all levels from undergraduate through senior scientist" (Levin 1992; Hastings et al. 2002). In issues of training, one must look at the entire pipeline, not just the separate pieces of the educational system. Therefore, we need long-term result horizons and this should include postdoctoral training. At this stage of the development of the interface, many of the resources should be directed towards groups and institutions where there is a critical mass of individuals to be engaged. Resources should be used to build on existing strengths. The critical mass argument applies as well as a rationale for having groups of postdocs working under a coordinated program and on the shared projects, allowing them to learn from each other and provide positive reinforcement through their interactions.

Recommendation 1: Programs should be developed to train postdoctoral students in a multidisciplinary setting in an environment of a critical mass of researchers as well as fellow postdocs. Such groups should be supported at multiple sites.

In particular, we expect that successful activities would:
- Provide the postdoc the opportunity to train in a new area. Explicitly, if the original training were in mathematics, computer science, or statistics, the new training would be in the biological/biomedical sciences, or vice versa.
- Have a critical mass of faculty or senior researchers and postdocs. This critical mass could be in interdisciplinary research centers or institutes, in interdepartmental or inter-institutional partnerships, etc.
- Have opportunities for interactions with graduate students in a project team setting. Have clearly defined projects, in a team setting. Indication of sample projects should be given. These projects could be those that already exist and would serve as the integrative force for the postdoctoral students.
- Provide selection criteria for the postdocs. In particular, some considerations would include the ability of the postdoc to interact with graduate students and form bridges between faculty in different disciplines.

Projects involving postdocs who would work at more than one site during the course of their involvement could be especially inviting. We hope that NIH and NSF would work together to define this program, and attract

Foundations such as Burroughs-Wellcome or Sloan to contribute. Furthermore, we feel that because of the nature of the award, a reduced overhead rate would be appropriate, and consistent with other such programs (e.g. in the range of 8 to 15%).

Recommendation 2: Postdoctoral Group Grants should be made to research groups or other infrastructure centers. Integrative education research training: we are proposing a vehicle that would couple together groups at undergraduate, graduate and postdoctoral levels. We envision that one of the components would be in place, but additional support would be provided to expand an "adjacent" component. We feel that by taking this holistic approach to research and training, each individual at one level will be able to see a possible career trajectory at the next. Furthermore, this could foster a type of mentoring that will last throughout a professional career.

Recommendation 3: Accept proposals that bring together students at different levels of training in a research project. Successful proposals would track the students and assess the impact of the experience on the students' careers.

Education Action Item 2: Mathematical – Biology Department Chairs' Conference

Concept: To hold a meeting of chairs (or their designates) from biology, math, computer science, and statistics departments who are interested in cross-disciplinary education for undergraduate students. The goal would be to identify the means by which existing resources could be utilized most effectively to promote training of undergraduate students in quantitative biology and to begin discussion of other needed curricular changes. Attendees would include chairs from universities, four-year colleges, and junior colleges. Institutions would send cross-disciplinary teams with chairs from biology and from math and/or computer science to the meeting.

Rationale and Proposed Program*:* The conference could be a mixture of workshop and presentations, with a goal of quick-starting the development of programs in quantitative biology and development of model undergraduate curricula that would give students cross-disciplinary "literacy." The meeting would showcase presentations from individuals and groups who have already developed curricula. The MAA, in conjunction with AAAS, ASM and funded by NIH and NSF will be

releasing a report to help with this, based on outcomes of the February 27 - March 1, 2003 NSF/NIH Symposium Meeting the Challenges in Emerging Areas: Education Across the Life, Mathematical, and Computer Sciences.

The goal would be the identification of a list of courses from biology, math, statistics, and computer science appropriate for the kind of biological questions being explored (e.g., environmental biology, neuroscience). Many of these courses already exist at most undergraduate institutions. There would also be discussion of new kinds of coursework required for such curricula.

1. Identification of mechanisms for optimum sharing of existing resources between institutions. Institutions can learn much from math/bio programs already initiated at other institutions and centers. Presentations would be made by faculty already involved in the teaching of innovative biology courses with strong quantitative components. Presentations would involve both a sample of the course and a discussion of the process the faculty used to develop the course.

2. Discussion of mechanisms for faculty development. Training is probably required to prepare faculty both for teaching biology in a more quantitative manner and for initiating cross-disciplinary research. A goal of the meeting might be to prepare some recommendations for teacher training programs appropriate for faculty interested in both classroom teaching and/or research.

Venue: Washington, DC - possibly the National Academy of Sciences. Now that the NRC Board on Mathematical Sciences has cancelled its Fall Mathematics Department Chairs' Colloquium series, this meeting could take its place, and is probably best done as a free-standing meeting because of the number of different departments that could participate.
Time frame: Spring 2004 (urgent)

Education Action Item 3: Mathematics/Biology Joint Faculty Positions

Concept: Create programs for 3-5 years of infrastructure support (startup, postdocs, graduate and undergraduate RA/Fellowships) for beginning Assistant Professors jointly appointed in university tenure track Assistant Professor positions at the interface between mathematics/statistics and biology/biomedical sciences. A university would propose a partnership between its mathematics/statistics department(s) and biology/biomedical

sciences department(s), create a new jointly appointed position between these departments, fund salary, and find office space in each department and lab space for this new appointment. The Departmental partnership would propose a mentoring program for this new faculty member, and make clear in writing the expectations to be met for this person to achieve promotion and tenure.

Rationale: To increase the number of trained professionals in the cross-disciplinary area of quantitative biology, there is a need for new curricula and courses, successful role models, and interdisciplinary research opportunities for faculty and students. Faculty jointly appointed in two departments represent a unique opportunity to efficiently fulfill the needs of interface fields by making them a part of the job description. In the existing departmental structure, it is often difficult to hire and nurture cross-disciplinary faculty since they often represent a new and perhaps unfamiliar area of teaching and research. To the biologists, this person would be a "mathematician" and to the mathematicians, this person would be considered a "biologist."

Considerations for promotion, tenure, graduate student training, etc. for interdisciplinary persons are different from those in a classical "pure" discipline. Consequently, an external funding program such as the one being proposed is needed in order for a new generation of interdisciplinary scientists to overcome existing institutional activation barriers. These interdisciplinary scientists will play an important role in changing the culture of "pure" biology and mathematics departments.

Proposed Program: We propose a program to provide 5 years of infrastructure support (startup, postdocs, graduate and undergraduate research assistantships and fellowships) for Assistant Professors jointly appointed in university tenure track positions at the interface between mathematics/statistics and biology/biomedical sciences. A University will propose a partnership between two of its departments, one from mathematics/statistics/computer science and the other from biology/biomedicine. At least one of the partner departments must be outside of the university's medical school. The departments will work as a team to provide the salary, office and lab space for a new position. The proposal must clearly state the expectations required of the new appointee in the areas of service, teaching and research for promotion and tenure. In addition, the proposal should describe the role the new faculty member will play in developing the teaching and research interface between the two departments and disciplines they represent.

Education Action Item 4: Summer Math/Bio Camp

Concept*:* Summer math/bio programs for high school teachers and/or high school students. The goals would be to introduce students and teachers to the interface between math and biology. Teachers would be introduced to specific topics that could be woven into existing courses in mathematics and biology whereas students would be introduced to ways to pursue these topics in their future studies. NSF/NIH educational resources concerning curricula, course materials, and college and career paths would be advertised and utilized.

An interesting option is to build programs for teachers around the preparation of teaching modules. Educational materials developed through this program would be posted on appropriate websites at NSF and NIH. Students could be given "research" activities and also given assistance in developing and preparing presentations aimed at junior high and elementary school students.

Proposed Program*: These programs could take various forms: a 3-week program for teachers and students together; a 3- or 4-week program for students alone; a 3- or 4-week program for students with teachers in residence for part of the time; a 1- or 2-week program for teachers alone. The teachers in these programs could be exclusively math/computer science teachers, exclusively biology teachers, or a mixed group. Teachers could apply individually to the program or as part of a team of math and biology teachers from the same school. These are not new ideas. In fact, various versions of these ideas have been used in "Young Scholars" programs and "Teacher Enhancement" programs for a long time. What is new is the emphasis on mathematics/ computer science and biology together.

The instructors in these math/bio camps would be university faculty, university students, and high school teachers. Postdoctoral fellows who would be trained in math/bio topics through other programs (see action item 1) could also be involved by giving lecture(s) and/or acting as mentors in the summer camp programs. Time frame: Run a number of such programs starting with one pilot program in summer 2004.

Education Action Item 5: K-12 Interface Educational Materials

Concept*:* A symposium about developing educational material for the math/bio interface is urgently needed.

Proposed Program: The program should bring together a variety of people interested in developing education material at the math/bio interface. Connections made at this symposium could lead to future collaborations/programs to develop educational material. Participants would include K-12 educators, K-12 students, K-12 parents, undergraduate and graduate students, cognitive scientists, learning technologists, large companies or organizations with an innate interest in innovative approaches to education (Disney, Microsoft, Nova, National Geographic, Sea World, NIH, NSF, Biotechnology Institute, National Education Association, for example), advocacy groups such as BIO whose member companies would benefit from having math-bio professionals, and representatives from those companies themselves.

Location: Possibly held in conjunction with FASEB EB 2004 conference, April 17 – 21, 2004 in Washington DC.

Education Action Item 6: Central Math-Bio Website and Listserv

Concept: Create and maintain a central website with information on research and training opportunities at the math/bio interface.

Proposed Program: This website would contain, for example, an archive of power point presentations from the series of NSF/NIH workshops related to math/computer science/physics/biology linkages. Get other websites (professional organizations, etc.) to point to this website. Listserv would periodically be sent updates of interface activities or could serve as a tool to disseminate new information or to host ad hoc discussions.

Education Action Item 7: Double degrees in Mathematics and Biology

Concept*:* While it appears that there now are many more students who take double degrees at the undergraduate level in biology and computer science (although many still claim they are the first to do so in their schools), a similar trend for undergraduate double degrees in math and biology is rare. We propose a program aimed at undergraduates to fund them for an additional time at university (up to 5 semesters) to complete a second undergraduate degree in the complementary field. Such a program may also open the door for

students who did not enter college with sufficient AP credit to make a double degree feasible.

Proposed Program*:* A university will propose a partnership between its math/stats and bio/biomedicine departments to provide up to 5 fellowships for its students to pursue a second undergraduate degree in the complementary field. The students will be chosen as rising juniors, and will be funded for up to 5 semesters (4 academic and one summer) with the goal of obtaining double undergraduate degrees at the end of the funding period.

Strengthening Ties among the Researchers
Action Items
(Rapporteur: Zan Luthey-Schulten)

Researchers Action Item 1: Top Ten list of Problems at the Mathematics-Biology Interface

Concept: Much in the spirit of Hilbert's famous list of unsolved mathematical problems and the less lofty list of the top ten algorithms that had the greatest influence in science and engineering in the last millennium (see Enquist and Schmid, 2001, p. 950), a list of the top ten most challenging and promising areas in mathematical biology should be compiled.

It is hoped that the following list will help focus attention on the mathematical and computational methods needed to answer important and pressing biological questions. Although still in a very preliminary form, it was also suggested that an article on the top ten list would be a useful pedagogical tool for teachers in both mathematics and biology. All of the below would benefit from increased sharing of models and setting of standards for clarity, performance, publication and permanence of software and computational results.   This may require advances in database theory and development, as well as, library science

1. <u>Model multilevel systems: from the cells in people in human communities in physical, chemical and biotic ecologies</u> - Models have been developed at many size scales, from molecular up to global. It is difficult to make models at different scales interact appropriately in a way that is descriptive and accurate. It is also difficult to apply the conclusions

of a model that was developed at one size-level to larger size levels or aggregates of smaller-size-level systems. Areas of mathematics and computational science that might contribute: dynamical systems, physics, statistics, probability theory, real and functional analysis.

2. <u>Network modeling of complex metabolic pathways, cell signaling, and species interactions</u> - Although many of the parts have been elucidated, it is difficult to develop models of control, regulation, and complex interactions. It is also difficult to connect multi-component self-organizing entities. Contributing areas: dynamical systems, physics, computer science, chemistry.

3. <u>Integrate probabilistic theories: understand uncertainty and risk</u> - This requires integration of frequentist, Bayesian, subjective, and other theories of probability. Contributing areas: probability theory, statistics, real and functional analysis, ergodic theory, risk analysis and financial mathematics.

4. <u>Understanding computation: gaining insight and proving theorems from numerical computation and agent-based models</u> - Many agent-based models and many numerical solution techniques rely on computations that do not necessarily reveal the mechanism by which certain outcomes are produced by certain inputs. This indicates a lack of understanding of the system. There is a need for algorithms and theorems that characterize output space and input space and that set up a correspondence between these spaces in a way that gives understanding of the system. Contributing areas: combinatorics, dynamical systems, theory of computation, number theory, computer science, finite automata theory.

5. <u>Data mining and inference</u> - We need a better understanding of data mining and simultaneous inference (beyond Bonferroni). Contributing areas: probability theory, statistics.

6. <u>Linguistic and graph theoretical approaches</u> - There is a need for graphical approaches to represent gene and protein structures as well as super-tree models for evolution that take horizontal gene transfer into account. Contributing areas: graph theory, linguistics, computer science

7. <u>Modeling of brain function</u> - There is a continued need for modeling the mapping of neural function in the brain, particularly in connection with clinical use of functional MRI. Contributing areas: non-linear dynamics, emergent systems.

8. Computational tools for problems with multiple temporal and spatial scales - There is a need for further development of methods that allow modeling of biological processes at varying length and time scales. Improved Hybrid Monte Carlo/Molecular Dynamics Simulations are needed for a dramatic improvement in sampling at all length and time scales in cellular function. Integro-difference equations have allowed large temporal jumps in spatio-temporal models of ecology. Stochastic differential equations have been used to model cellular control and rhythms as well as for noise influenced behavior on larger biological scales. There is a need for new methods to monitor and detect large deviations in living systems. Contributing areas: statistics, statistical physics, stochastic calculus.

9. Ecological forecasting: There is a need for improved modeling of the effects of combined exogenous factors (e.g., climate, land-use change) on large-scale population and community dynamics and our habitat. Linking of multiple sub-models and scaling models up/down is particularly difficult. Contributing areas: environmental science; ecology; dynamical systems; frequentist, Bayesian and other theories of probability.

10. Effects of erroneous data on biological understanding - Erroneous data can be a problem in diverse biological disciplines. In genomics and phylogenetics, gene sequencing systems can give wrong sequences, as a result of a small number of samples processed (which is often only 1 or otherwise low). Conservation biology is another example in which we do not know enough about how faulty data lead to faulty understanding or mistaken decision making. Contributing areas: statistical models with measurement error, measurement error models, cross-validation techniques.

Researchers Action Item 2: National Meeting in Washington D.C. on Mathematical Biology

Concept: The national meeting should showcase successes where mathematics and computations have helped solve important problems in medicine and motivate new directions and opportunities on all levels of biology from cells to the organs to ecology to the biosphere.

Possible hot topics include small interference RNAs, modeling of systems in molecular medicine such as membrane channels like aquaporin, network modeling of complex metabolic pathways and cell signaling, and models of the brain function. It is important that speakers from all aspects

of the math-bio linkage be represented: experimental biologist, biomedical researcher, mathematician, and computational scientist. Most important would be to document a close link between experiment and modeling such as can be done for the aquaporin membrane channels that involved clinical researchers and modelers.

Researchers Action Item 3: Meetings and Workshops
Concept: In general, mathematicians need workshops to learn about the biology and the biologists, and biomedical researchers need workshops to learn about the mathematics and be able to talk with the mathematicians. To accelerate the communication between mathematicians and biologists, interactions on several different time scales and formats will be required:

Mathematics-In-Biology Workshops: 3-5 Days. The workshops are based upon similar successful workshops involving mathematicians and industry. The proposed format is the following: candidate problems will be solicited through advertising or a call for problems. The problems will appear on a web site prior to the meeting. For the initial meeting, a small number of problems would be solicited (say 2-4), but future meetings could present as many as 8-10 projects/problems. On the first day of the workshop, biologists outline the projects and their objectives. Mathematicians will then work collaboratively with the biology participants on the projects for 2-3 days of brainstorming, collaboration and problem solving. Participants are free to work on whichever projects match their interest and expertise. Results and recommendations are presented on the final day.

Each project will be managed by a mathematics and a biology moderator whose role is to coordinate preliminary work including literature searches, keep the problem solving/modeling on track, elicit contributions from all present, provide summaries and assist in the preparation of a final technical report for the biology participants. Longer term goals would be to make this a leading forum for exchange between mathematicians and biologists that would become an annual problem-solving workshop that attracts leading university and government mathematicians to work with biologists.

Workshops on Specialized Topics at the Interface of Biology, Mathematics, and Physics: As many such activities exist already, the idea would be to build upon these activities and provide financial support rather than create new initiatives. Meetings on multi-scale modeling, visualization, information theory and graph theory were repeatedly

suggested. It is important that trainees working at the interface of these fields, both postdoctoral researchers and graduate students, attend the meetings in addition to the seasoned scientists.

Create a Think Tank dedicated to Problems at the Interface of Biology and Mathematics: The community supports the creation of an institution that allows long-term collaborative activities along the lines of the existing NCEAS and recently funded NSF Biophysics center. This new Math-Bio Think Tank could find its home perhaps at Woods Hole or some other existing institute that could support long-term visitors.

Researchers Action Item 4: Strengthen Publication Ties between Mathematics and Biology

Concept: The collaboration between mathematicians and researchers in the biological and biomedical sciences and the use of mathematics in life sciences research would be advanced by providing easy access to the mathematical literature through Pub Med.   Likewise, it would be useful to give mathematicians easy access and guidance to the life sciences literature.

The main database for mathematical research papers and books is Mathematical Reviews and its on-line version is MathSciNet.   Math Reviews is produced by the American Mathematical Society. As the co-chair of the AMS library committee Dr. Laubenbacher will explore the possibility of a linkage between these two databases.

Closing Comments: Time to Act

There appear to be two general roles that mathematics can play in the life and biomedical sciences. The first is to provide off-the-shelf methods to solve problems, e.g., linear algebra methods, numerical solvers for systems of ODEs, etc. The second is to provide fundamentally new approaches to biological problems by formulating them in new ways, using new mathematical methods, with a new set of tools. The first role does not need Ph.D. mathematicians, because the challenge does not lie in the mathematics. To decide whether a particular term should be part of an ODE model, all one needs is three semesters of calculus and a course in

differential equations. Engineers do this all the time. Biologists would benefit greatly from this kind of mathematical training as well.

The second role is most appealing to the mathematicians. At present we are using maybe 20-30% of known mathematics in the life sciences. But there are plenty of examples of how quantum leaps in understanding have been achieved by a fundamentally new framework for a problem. In fact, this is probably the only way of achieving quantum leaps. This is the essential role that mathematicians can play. For this role we need mathematicians that are literate in the life sciences. A great example is David Mumford who works in neuroscience. To most mathematicians he is known as an outstanding algebraic geometer of the most abstract kind. He is now using this formalism to successfully approach some key problems in vision and other areas of brain function. Lectures and workshops directed at the mathematics community at large would help familiarize it with the important biology problems that need to be solved. Then maybe more Mumfords would come out of the woodwork.

To biologists, off-the-shelf mathematics has become an essential part of their work – indeed, competitive research programs in biology must have strong quantitative bases. Increasingly, however theoretical work by biologists and/or work in collaboration with mathematicians is leading to fundamental insights on the part of biologists. Biologists agree that progress in understanding biological problems will depend on fundamental advances in spatial dynamics, stochastic (nonlinear) dynamics, and how models are best fitted to data (Hastings and Palmer, 2003).

This meeting and the related articles referenced below discuss the opportunities at the interface between mathematics and the biological and biomedical sciences. It is important to stress that there is an urgency to move forward. Mathematics provides a unique and necessary lens through which to view the world. In the words of Richard Feynman, "*People who wish to analyze nature without using mathematics must settle for a reduced understanding."* Yet with the problems facing us today in our global society, we can ill afford to plan on a reduced understanding, in the face of SARS (or the next SARS) or in the impact of climate change on our environment. Mathematics has a role to play, but we must begin today since it takes a long time to build up the human capital and the partnerships that are decidedly critical to bringing together these fields.

We see the future, we can obtain it, but it will take an immediate, concerted and sustained effort across the two key agencies that took this first step. Let's begin now!

References

Atkins, D, et al 2003. Revolutionizing Science and Engineering through Cyberinfrastructure. (http://www.communitytechnology.org/nsf_ci_report/).

Enquist, B. and W. Schmid (editors). 2001. Mathematics Unlimited, 2001 and Beyond. Springer Verlag. p. 950.

Hastings, A. et al. 2002. Quantitative biology for the 21st century. Report from a workshop supported by NSF and held at the University of California – San Diego, December 2002. (http://www.sdsc.edu/QEIB).

Hastings, A. and M.A. Palmer. 2003. A bright future for biologists and mathematicians. Science 299: 2003-2004.

Levin, S.A. (editor). 1992. Mathematics and Biology, The Interface: Challenges and Opportunities 1992. Workshop supported by NSF DMS 89 10353. US Department of Energy, DE-AC03-76SF00098.

Levin, S.A., B. Grenfell, A. Hastings, and A.S. Perelson. 1997. Mathematical and computational challenges in population biology and ecosystems science. Science 275: 334-343.

National Research Council. 1995. Report Mathematical Challenges from Theoretical/ Computational Chemistry. National Academies Press

Wooley, J, et.al. BIO Advisory Committee Cyberinfrastructure Report (Interim), Report from a Meeting on 14-15 July.

Appendix 1. Registered Participant list
(note: there were additional un-registered attendees on Day 1)

Cyrus Aidun, Engineering Department, National Science Foundation. caidun@nsf.gov

Mark Alber, Department of Mathematics, Center for the Study of Biocomplexity, University of Notre Dame. malber@nd.edu

Norma Allewell, University of Maryland.

Priyanga Amarasekare, Department of Ecology and Evolution, University of Chicago. amarasek@uchicago.edu

Peter Arzberger, University of California, San Diego. parzberg@ucsd.edu

Craig Aumann, Department of Biology, University of Maryland. caumann@wam.umd.edu

John Baras, Electrical and Computer Engineering Department, University of Maryland. baras@isr.umd.edu

James Bassingthwaighte, Bioengineering Department, University of Washington. jbb@bioeng.washington.edu

Benjamin Bolker, Department of Zoology, University of Florida. bolker@zoo.ufl.edu

Michael Bowers, BIO/ Environmental Biology, National Science Foundation. mbowers@nsf.gov

David Brown, Agronomy and Range Science, University of California, Davis. dhbrown@ucdavis.edu

Margaret Bruce, Pediatrics Department, University of Kentucky. mbruce@pop.uky.edu

Eugene Bruce, Division of Integrative Biology and Neuroscience, National Science Foundation. ebruce@nsf.gov

Denise Caldwell, Physics Division, National Science Foundation. dcaldwel@nsf.gov

Catherine Campbell, National Institute of Neurological Disorders and Stroke, National Institutes of Health. campbelc@ninds.nih.gov

James Cassatt, National Institute of General Medical Sciences, National Institutes of Health. cassattj@nigms.nih.gov

Jean Chin, National Institute of General Medical Sciences, National Institutes of Health. chinj@nigms.nih.gov

David Chopp, Engineering Science and Applied Mathematics, Northwestern University. chopp@northwestern.edu

Ida Chow, Society for Developmental Biology. ichow@faseb.org

Mary Clutter, Assistant Director, Division of Biological Science, National Science Foundation. mclutter@nsf.gov

Joel Cohen, Laboratory of Populations, Rockefeller University. Cohen@rockefeller.edu

Jerry Collins, Department of Biomedical Engineering, Vanderbilt University. jerry.c.collins@vanderbilt.edu

Omar Colon-Reyes, Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University. ocolonre@math.vt.edu

Kathryn Cottingham, Department of Biological Sciences, Dartmouth College. cottingham@dartmouth.edu

Jennifer Couzin, Science Magazine.

Isabel Darcy, Programs in Mathematical Sciences, University of Texas at Dallas. darcy@utdallas.edu

Charles Delwiche, Cell Biology and Molecular Genetics, University of Maryland - College Park. delwiche@umd.edu

Bharati Dhruva, Division of Pathology, Walter Reed Army Institute of Research. paulomi_15@yahoo.com

Emmanuele DiBenedetto, Department of Mathematics, Vanderbilt University. em.diben@vanderbilt.edu

Karin Dorman, Statistics, Genetics, Development and Cell Biology, Iowa State University. kdorman@iastate.edu

Lloyd Douglas, Division of Mathematical Sciences, National Science Foundation. ldouglas@nsf.gov

Clifford Duke, Office of Science Programs, Ecological Society of America. csduke@esa.org

Irene Eckstrand, National Institute of General Medical Sciences, National Institutes of Health. eckstrai@mail.nih.gov

Donald Estep, Department of Mathematics, Colorado State University. estep@math.colostate.edu

Yang Fann, National Institute of Neurological Disorders and Stroke, National Institutes of Health. fann@ninds.nih.gov

Matthew Ferguson, Physics Department, University of Maryland. ferguson@physics.umd.edu

Paula Flicker, National Institute of General Medical Sciences, National Institutes of Health. flickerp@nigms.nih.gov

Christiana Fogg, National Institute of Allergy and Infectious Diseases, National Institutes of Health. cfogg@niaid.nih.gov

Bruce Fuchs, Office of Science Education, National Institutes of Health. Bruce_Fuchs@nih.gov

Dan Gallahan, National Cancer Institute, National Institutes of Health. dg13w@nih.gov

Dylan George, Division of Environmental Biology, National Science Foundation. dylangeorge1@yahoo.com

Leah Gerber, Biology Department, Arizona State University. leah.gerber@asu.edu

Ronald Germain, National Institute of Allergy and Infectious Diseases, National Institutes of Health. rgermain@nih.gov

Avijit Ghosh, Department of Physics, Drexel University.
avijit@physics.drexel.edu

James Glazier, Biocomplexity Institute, Indiana University.
glazier@indiana.edu

Peter Good, National Human Genome Research Institute, National Institutes of Health. goodp@mail.nih.gov

Geoffrey Goodhill, Department of Neuroscience, Georgetown University Medical Center. geoff@cns.georgetown.edu

Michael Gottesman, Office of Intramural Research, National Institutes of Health. mgottesman@nih.gov

Yonatan Grad, Harvard Medical School.

Ananth Grama, Computer Sciences Department, Purdue University.
ayg@cs.purdue.edu

Harvey Greenberg, Mathematics Department, University of Colorado at Denver. Harvey.Greenberg@cudenver.edu

Judith Greenberg, National Institute of General Medical Sciences, National Institutes of Health. greenbej@nigms.nih.gov

Leslie Greengard, MadMax Optics, Inc.

Evan Haskell, Mathematics Department, University of Utah.
haskell@math.utah.edu

Alan Hastings, Department of Environmental Science and Policy, University of California, Davis. amhastings@ucdavis.edu

Neil Hoffman, Molecular and Cellular Biology, National Science Foundation. nhoffman@nsf.gov

Dan Hogan, OCPL, National Institutes of Health. hogand@nigms.nih.gov

Robert Holt, Department of Zoology, University of Florida.
rdholt@zoo.ufl.edu

Monica Hurdal, Department of Mathematics, Florida State University.
mhurdal@math.fsu.edu

Anthony Ives, Zoology Department, University of Wisconsin -Madison.
arives@wisc.edu

Yi Jiang, Theoretical Division, Los Alamos National Laboratory.
jiang@lanl.gov

Sheriff Kafafi

Christine Kelley, National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health. kelleyc@mail.nih.gov

Malgorzata Klosek, National Cancer Institute, National Institutes of Health.
klosek@nih.gov

Mehmet Koyuturk, Department of Computer Sciences, Purdue University.
koyuturk@cs.purdue.edu

Conrad Kufta. cvkufta@aol.com

Reinhard Laubenbacher, Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University. reinhard@vbi.vt.edu

Kevin Lauderdale, National Institute of General Medical Sciences, National Institutes of Health. lauderdk@nigms.nih.gov

Martin Levin. mdlevin@aol.com

Catherine Lewis, National Institute of General Medical Sciences, National Institutes of Health. lewisc@nigms.nih.gov

Frances Li, Office of International Science and Engineering, National Science Foundation. fli@nsf.gov

Soo-siang Lim, Division of Integrative Biology and Neuroscience, National Science Foundation. slim@nsf.gov

Deborah Lockhart, Division of Mathematical Sciences, National Science Foundation. dlockhar@nsf.gov

Wolfgang Losert, Physics Department, University of Maryland. wlosert@glue.umd.edu

Sharon Lubkin, Mathematics and Statistics Department, North Carolina State University. lubkin@eos.ncsu.edu

Gloria Lubkin, Physics Today, American Institute of Physics. glubkin@aip.org

Nadine Lymn, Public Affairs Department, Ecological Society of America. nadine@esa.org

Michael Marron, National Center for Research Resources, National Institutes of Health. marron@nih.gov

Donald Mattison, National Institute of Child Health and Human Development, National Institutes of Health. mattisod@mail.nih.gov

Victor Matveev, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. matveev@nih.gov

Ilya Mazo, Ariadne Genomics Inc. mazo@ariadnegenomics.com

Martin Meier-Schellersheim, National Institute of Allergy and Infectious Diseases, National Institutes of Health. mms@niaid.nih.gov

Raymond Mejia, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. ray@helix.nih.gov

Pedro Mendes, Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University. mendes@vt.edu

Tom Misteli, National Cancer Institute, National Institutes of Health. mistelit@mail.nih.gov

Alex Mogilner, Mathematics and Genetics and Development, University of California, Davis. mogilner@math.ucdavis.edu

Carrie Morjan, Biology Department, Indiana University. cmorjan@bio.indiana.edu

Keith Mott, Biology Department, Utah State University. kmott@biology.usu.edu

Sumeet Muju, National Institute of Neurological Disorders and Stroke, National Institutes of Health. mujus@ninds.nih.gov

Timothy Myers, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Ramesh Nayak, Center for Scientific Review, National Institutes of Health. nayakr@csr.nih.gov

Michael Neubert, Biology Department, Woods Hole Oceanographic Institution. mneubert@whoi.edu

John Nolan, Mathematics and Statistics Departments, American University and NIH. jpnolan@american.edu

Clare O'Connor, Biology Department, Boston College and the National Science Foundation. oconnocn@bc.edu

Mario Ohlberger, CSCAM, University of Maryland. mario@mathematik.uni-freiburg.de

Richard Okita, National Institute of General Medical Sciences, National Institutes of Health. okitar@nigms.nih.gov

Dianne O'Leary, Computer Science Department, University of Maryland. oleary@cs.umd.edu

Janet Oliver, Pathology Department, University of New Mexico Health Sciences Center. joliver@salud.unm.edu

Margaret Palmer, Department of Biology and Entomology, University of Maryland – College Park. mp3@umail.umd.edu

Jack Panossian, National Institute of Neurological Disorders and Stroke, National Institutes of Health. jack_panossian@nih.gov

Richard Pastor, Center for Biologics Evaluation and Rearch, Food and Drug Administration. pastor@cber.fda.gov

Gail Patricelli, Cornell Lab of Ornithology, Cornell University. GailPatricelli@hotmail.com

David Peak, Physics Department, Utah State University. peakd@cc.usu.edu

Michael Pearson, Mathematical Association of America. pearson@maa.org

Peter Preusch, National Institute of General Medical Sciences, National Institutes of Health. preuschp@nigms.nih.gov

Samuel Rankin, American Mathematical Society. smr@ams.org

Martin Reddington, Human Frontier Science Program. m.reddington@hfsp.org

Daniel Reuman, J. Cohen Lab, Rockefeller University. reumand@rockefeller.edu

Fred Roberts, DIMACS Center, Rutgers University. froberts@dimacs.rutgers.edu

Sandra Rodriguez-Zas, Animal Sciences and Statistics Department, University of Illinois at Urbana-Champaign. rodrgzzs@uiuc.edu

Pejman Rohani, Institute of Ecology, University of Georgia. rohani@uga.edu

Peter Roper, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. Pete.Roper@nih.gov

James Rosenberger, Statistics Department, Penn State University. JLR@psu.edu

Joshua Rosenthal, John E. Fogarty International Center, National Institutes of Health. joshua_rosenthal@nih.gov

Andrej Rotter, Mathematical Biosciences Institute, The Ohio State University. rotter.1@osu.edu

Robert Rubin, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. robertr@intra.niddk.nih.gov

Stephanie Ruby, Molecular Genetics and Microbiology, University of New Mexico Health Sciences Center. sruby@unm.edu

Norka Ruiz Bravo, National Institute of General Medical Sciences, National Institutes of Health. ruizbran@nigms.nih.gov

William Rundell, Division of Mathematical Science, National Science Foundation. wrundell@nsf.gov

Estelle Russek-Cohen, Biometrics Program, Dept of Animal and Avian Sciences, University of Maryland. er19@umail.umd.edu

Jaya Satish, National Institute of Dental and Craniofacial Research, National Institutes of Health. satishj@mail.nih.gov

Carl Schaefer, National Cancer Institute, National Institutes of Health. schaefec@mail.nih.gov

Samuel Scheiner, Division of Environmental Biology, National Science Foundation. sscheine@nsf.gov

Tamar Schlick, Chemistry and Courant Institute of Mathematical Sciences (CIMS), New York University. schlick@nyu.edu

Zaida (Zan) Schulten, Department of Chemistry, University of Illinois at Urbana-Champaign (UIUC). zan@uiuc.edu

Jignesh Shah, Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University. jshah@vbi.vt.edu

Bert Shapiro, National Institute of General Medical Sciences, National Institutes of Health. shapirob@nigms.nih.gov

Arthur Sherman, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. asherman@nih.gov

Vladimir Shulaev, Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University. vshulaev@vbi.vt.edu

Anjali Singh, National Institute of Mental Health, National Institutes of Health. anjalis@codon.nih.gov

Satya Singh, National Institute of Allergy and Infectious Diseases, National Institutes of Health. spsingh@niaid.nih.gov

Jennifer Slimowitz, Division of Graduate Education, National Science Foundation. jslimowi@nsf.gov

William Smith, American Statistical Association.
williambsmith@amstat.org

Gregory Smith, Department of Applied Science, College of William and
Mary. greg@as.wm.edu

Peter Sorger, Department of Biology and Biological Engineering,
Massachusetts Institute of Technology. psorger@mit.edu

Julie Stern, National Institute of Child Health and Human Development,
National Institutes of Health. jvstern@helix.nih.gov

Michael Steuerwalt, DMS/Applied Math and Computational Math, National
Science Foundation. msteuerw@nsf.gov

Brandilyn Stigler, Virginia Bioinformatics Institute, Virginia Polytechnic
Institute and State University. bstigler@vbi.vt.edu

Andrea Stith, Office of Legislative and Public Affairs, National Science
Foundation. astith@nsf.gov

Ulrich Strom, National Science Foundation. ustrom@nsf.gov

De Witt Sumners, Department of Mathematics, Florida State University.
sumners@math.fsu.edu

Hong-wei Sun, National Institute of Arthritis and Musculoskeletal and Skin
Diseases, National Institutes of Health. sunh1@mail.nih.gov

Shyam Sundaram, Center for Information Technology, National Institutes
of Health. shyam@mail.nih.gov

Nancy Sung, Burroughs Wellcome Fund. nsung@bwfund.org

Myong-Hee Sung, National Cancer Institute, National Institutes of Health.
sungm@mail.nih.gov

Wojciech Szpankowski, Computer Science Department, Purdue
University. spa@cs.purdue.edu

Paul Tan, National Institute of Diabetes and Digestive and Kidney
Diseases, National Institutes of Health. pault@intra.niddk.nih.gov

Michael Tims, Department of Cell Biology and Molecular Genetics,
University of Maryland. mct@wam.umd.edu

David Vannier, Office of Science Education, National Institutes of Health.
vannieb@od.nih.gov

Venkat Venkatasubramanian, School of Chemical Engineering, Purdue
University. venkat@ecn.purdue.edu

Geert Wenes, Bioinformatics Department, National Center for Genome
Resources. gcw@ncgr.org

John Whitmarsh, National Institute of General Medical Sciences, National
Institutes of Health. whitmarj@nigms.nih.gov

Calvin Williams, Division of Undegraduate Education, National Science
Foundation. cwilliam@nsf.gov

Susan Yang, Neurology Department, Johns Hopkins University.
khyang@jhmi.edu

Jian-Min Yuan, Department of Physics, Drexel University.
yuanjm@drexel.edu
Anton Yuryev, Research Department, Ariadne Genomics, Inc.
info@ariadnegenomics.com

Appendix 2
Article in Science Highlighting the Event

A Bright Future for Biologists and Mathematicians?

Alan Hastings and Margaret A. Palmer

Why, despite vaccination efforts, is Boulder, Colorado, weathering an outbreak of whooping cough (pertussis)—a potentially fatal illness in young children—this winter? The answer to this biological question comes from the classic mathematical analysis of Kermack and McKendrick, whose threshold theorem calculates the minimum level of vaccination required to prevent an outbreak of an infectious disease (1). This example of how mathematics can help biology was just one of many discussed at a recent series of Quantitative Environmental and Integrative Biology workshops (2) and at a recent NIH-NSF workshop that examined forging

A. Hastings is in the Department of Environmental Science and Policy, University of California, Davis, CA 95616, USA. E-mail: amhastings@ucdavis.edu M. A. Palmer is in the Department of Entomology, University of Maryland, College Park, MD 20742, USA. E-mail: mp3@umail.umd.edu

stronger links between mathematicians and biologists (3). A goal of the workshops was to seek answers to the questions: Which biological problems will yield to mathematical analyses, and how should biology and mathematics be integrated to achieve this?

Kermack and McKendrick developed the threshold theorem to determine the conditions under which infectious disease epidemics occur. This theorem has proved crucial for calculating the level of vaccination (less than complete coverage) required to eradicate diseases like polio and smallpox, and for preventing outbreaks of diseases such as pertussis. This theorem relates the occurrence of an epidemic to the number of susceptible individuals, the duration of the infectious period, and the infectivity of the disease. The threshold theorem was initially developed to answer two fundamental biological questions: Why do infectious disease epidemics occur, and

why do they typically die out before all susceptible individuals contract the disease? These questions were answered by using the threshold theorem to develop the SIR (susceptible, infective, removed) model (1), which consists of three differential equations. The SIR model assumes that over the time scale of an epidemic, births and deaths in the host population can be ignored. The model includes the rate of removal (through death or recovery) of infected persons from the group passing on the infection, instead of specifying the more correct but harder to analyze assumption that there is a fixed time period during which an individual can infect others. The threshold theorem was originally illustrated using methods that relied on the graphic display of the number of infective and susceptible individuals during an infectious disease outbreak. The graphic representation of the threshold theorem reveals that the density of susceptible individuals must exceed a certain critical value for an epidemic to occur. This theorem has unquestionable relevance, given heightened concerns about the deliberate introduction of new infectious bioterrorist agents.

Workshop participants agreed that progress in understanding biological problems will depend on mathematical ad-

vances in spatial dynamics (4); stochastic (nonlinear) dynamics, especially as applied to spatial systems (5); and how models are best fit to the data (6). Spatial stochastic systems in biology, such as the population dynamics of species in forest or grassland ecosystems, are motivating development of new mathematical models. A complete description of spatial dynamics is extremely complex (as it would include the dynamics of means, variances, third moments, and higher order moments), and so approximations are needed. For example, moment-closure methods approximate complete dynamics using only a few moments, thus enabling the tracking of, for example, the spatial dynamics of infection rates. This type of mathematical model is particularly valuable for analyzing the dynamics of infectious diseases because the likelihood of a susceptible individual becoming infected does not depend on the overall level of infected individuals in the population, but rather on the severity of infection among those individuals with whom °the susceptible individual is in contact.

Meeting delegates viewed several areas as especially promising candidates for successful application of mathematical and quantitative approaches to solving biological and societal problems. Examples include how natural resources should be managed, forecasting the effects of global climate change, and evaluating the movement of agricultural pests. A good example of how mathematics can benefit biology is the calculation of the size and spatial configuration of marine reserves needed to sustain a fish population that may be overexploited. The basic question is how to calculate the total rate of settlement of new individuals at any point in space, summing up contributions from all other locations. Conditions for the survival and persistence of marine species have been derived from discrete-time and continuous-space models. These models are based on a dispersal kernel model, which gives the probability of offspring from marine organisms being recruited at a given distance along the coast from the point of release from the parent (see the figure) (7). The dispersal kernel model has
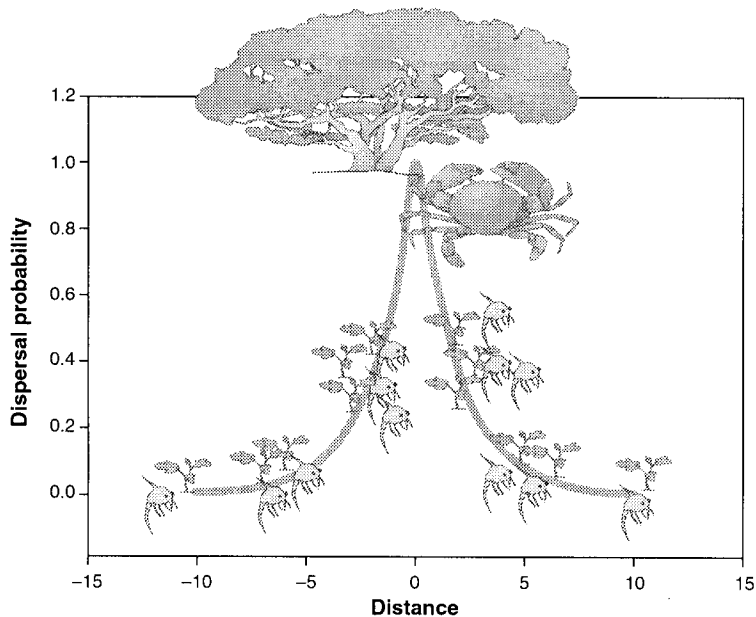
Far away, so close. The dynamics of spatial systems are often at the interface between biology and mathematics. The spatial distribution of offspring around a parent is one example of a spatial process taking place over a discrete time period and can be represented by the dispersal kernel model. The figure shows the dispersal of crustacean larvae along ocean currents away from the parent and, for comparison, the dispersal of seeds from a mature tree. The fact that more offspring land close to the parent rather than far away is described by the double decaying exponential of the dispersal kernel model. This model is valuable not only for predicting the dispersal of offspring but also for addressing biological problems such as the design of marine reserves, the spread of invasive species, and the potential influence of GMOs on natural populations.

spurred the design of a series of interconnected marine reserves off the California coast. The next step is to make sustainability of marine populations apply to more realistic descriptions of oceanographic processes, to integrate economics more fully into calculations of marine reserve management, and to account for the uncertainty in the growth rates of marine populations.

Quantitative approaches can also be used to calculate how spreading of alleles from genetically modified organisms (GMOs) to natural organism populations might affect those populations. Related mathematical analyses examine the best approach for controlled introduction of GMOs that are resistant to insect pests. GMO technology is threatened by the risk that insect pests will evolve resistance to GMOs, and mathematical modeling suggest ways to reduce this risk (8).

Mathematics continues to be essential for understanding the dynamics of infectious disease outbreaks. One dramatic example is the foot-and-mouth epidemic in the United Kingdom in 2001 (9, 10). Tools such as the dispersal kernel model and explicit spatial models allowed comparison of different strategies for controlling the epidemic. These analyses enabled the design of a control strategy based on local culling

of infected and exposed animals that resulted in halting of the epidemic.

Many of the same mathematical themes emerge in cellular and molecular processes. In the cell, chemical energy in the form of ATP is converted into mechanical work by molecular motors—molecules that govern movement in living systems (11). The dynamics of these movements within the cell depend on stochastic forces that lead to discrete conformations of the motors enabling them to operate like molecular ratchets (12). Mathematical modeling opens the door to predicting force-velocity relations and other quantitative characteristics of the motors' actions, which can then be compared to actual measurements (13). Computational approaches make it possible to attack problems that are much more complex than the mere mechanics of single motors and to generate "virtual" structures that can be compared to real data from time-lapse microscopy (14).

Workshop participants agreed that a vital next step will be to promote the training of scientists with expertise in both biology and mathematics. A new generation of empiricists with stronger quantitative skills and of theoreticians with an appreciation for the empirical structure of biological processes will facilitate a bright future for the application of mathematics to solving biological problems.

References
1. W. O. Kermack, A. G. McKendrick, *Proc. R. Soc. London* **115**, 700 (1927).
2. A. Hastings *et al.*, Quantitative biology for the 21st century (2003); www.sdsc.edu/QEIB/
3. M. A. Palmer *et al.*, Accelerating mathematical-biological linkages: Report of a joint NSF-NIH workshop (2003); www.bisti.nih.gov/
4. R. W. van Kirk, M. A. Lewis, *Bull. Math. Biol.* **59**, 107 (1997).
5. B. M. Bolker, S. W. Pacala, *Theor. Pop. Biol.* **52**, 179 (1997).
6. O. N. Bjørnstad, B. T. Grenfell, *Science* **293**, 638 (2001).
7. L. W. Botsford *et al.*, *Ecol. Lett.* **4**, 144 (2001).
8. "Bt corn and European corn borer: Long-term success through resistance management," supplement to NCR-602 (United States Department of Agriculture, 1998).
9. N. M. Ferguson *et al.*, *Science* **292**, 1155 (2001).
10. M. J. Keeling *et al.*, *Science* **294**, 813 (2001).
11. H. Wang, G. Oster *Nature* **396**, 279 (1998).
12. R. D. Vale, R. A. Milligan, *Science* **288**, 88 (2000).
13. A. Mogilner *et al.*, *J. Theor. Biol.* **211**, 143 (2001).
14. E. Cytrynbaum *et al.*, *Biophys. J.* **84**, 757 (2003).