

## Conclusions and Recommendations

### 11.1 DISCIPLINARY PERSPECTIVES

#### 11.1.1 The Biology-Computing Interface

The committee began this study with two key notions. First, it hoped to identify a field of intellectual inquiry associated with the biology-computing interface that drew equally and bilaterally on computing and biology. Second, it hoped to explicate a symmetry between computing and biology in which the impact of computing on biology was increasingly deep and profound and in which biology would have a comparable effect on computing.

Both of these notions proved unfounded in certain important ways. From the standpoint of applications, technology, and practical utility, the committee saw substantial asymmetry. Computing has had a huge transformational impact on biology and will span virtually all areas of life sciences research, but the impact of biology on computing is likely to be much more targeted (i.e., affecting specific problem domains within computing), and large-scale, biology-based technology changes for computing are in the relatively distant future if they occur at all. At the same time, the committee did find that the epistemological and conceptual frameworks of each field may have in the future some substantial influence on the other. The committee believes that an engineering and computational view (as discussed in Chapter 6) will increasingly be recognized as an important way of looking at biological systems. In a parallel though somewhat more speculative vein, the committee also believes that insight into biological mechanisms may have important impact on how certain problems in computing can be approached (as discussed in Chapter 8).

The reason for the deep and transformational impact of computing on biology is that insight into the vast and heterogeneous datasets of 21st century biology will be possible only through the application of computing to analyze and manage those data. (This is not to deny that many quantitative sciences will contribute to biology, although this report has focused primarily on the computing dimensions.) Views among biologists about where best to deploy computing resources will surely differ, but the main contributions of computing to biology will come from new ideas for solving complex biological problems and new models for testing hypotheses; from delivering cyberinfrastructure for biology research, providing ever more computing power, distributed computing and storage, complex software, fault-tolerant computing, and so forth; and from training fearless scientists who can find the right

collaborators for whatever difficulties arise at the frontier. That is, specific computing-to-biology “tech transfer” of intellectual ideas will have some impact, but the greatest impact of computing on biology will come from an overall acceleration of the pace of progress.

To fulfill the promise of 21st century biology, research scientists from both computer and biological science need to work together more extensively, more often, and more closely than ever before. As quantitative methods are increasingly adopted within the biological sciences, it will be possible to answer a new range of scientific questions, not just to accelerate research progress. Uncovering the meaning implicit in the complete sequence of the human genome to deliver on the promises of the project for society is an obvious case.

A revitalized enterprise driven by this newly trained cadre of interdisciplinary scientists and maintained through a balance of individual investigator-initiated and group projects along with continued technology and computational advances, will be able (1) to address fundamental questions in biology such as the relationship of structure to function and the basis for homeostasis; (2) to integrate biological knowledge across the vast scales of time, space, and organizational complexity that characterize biology; (3) to translate basic biology to preventive, predictive, and personalized medicine and to extend biological knowledge to engineering soft materials and other industrial nanobiotechnology contributions; and (4) to uncover how biology can contribute to energy production and environmental restoration. The Committee on the Frontiers at the Interface of Computing and Biology believes that such a vision for 21st century biology is realistic, and that the implementation of its recommendations would ensure decades of exponential progress and a major transformation of our understanding of life.

On the other side of the interface, biological inspiration for new approaches to computing continues to be important, in the sense that biology provides existence proofs that information-processing technology based on biochemistry rather than on silicon electronics is possible. For areas of computing that are generally complex and unwieldy in the associated technologies available so far to address them, or areas lacking in empirical and/or theoretical knowledge, inspiration from whatever source is welcome—and biological inspiration is most likely to be valuable in these areas. (For other areas of computing, whose intellectual terrain is well explored and for which a solid base of empirical and theoretical knowledge is available, biological inspiration is both unnecessary and less interesting, because good and useful solutions are available without any kind of biological connection at all.)

Furthermore, computer scientists tend to be most interested in the general applicability of their work and are often less interested in work that is relevant to only one problem domain. Individuals from this perspective should thus understand the key difference between applications-driven research and applications-specific research. That is, problems in the life sciences can be important drivers of computer science research, and in many cases the knowledge developed in seeking solutions to these problems will be applicable in other domains.

Finally, it is worth noting one possible domain of symmetry between the two fields, although it is a symmetry of ignorance rather than one of knowledge. Both computing and biology provide objects of enormous complexity whose behavior is not well understood—consider the Internet and a cell. It may well turn out that studying each of these objects as systems can yield insights useful in understanding the other—and the same kinds of (yet-to-be-developed) formalism may apply to both—but the jury is still out on this possibility.

### 11.1.2 Other Emerging Fields at the BioComp Interface

Apart from computing-enabled biology and biologically inspired computing, a number of other new areas of inquiry are also emerging at the BioComp interface, although in addition to biology and computing they draw from chemistry, materials science, bioengineering, and biochemistry. Some of these efforts can be characterized loosely as different flavors of biotechnology, and three of the most important are analytical biotechnology, materials biotechnology, and computational biotechnology.

1. Analytical biotechnology describes the application of biotechnological tools for the creation of chemical measurement systems. Examples include the creation of sensors from DNA-binding proteins for the detection of trace amounts of arsenic and lead in ground waters, and the development of nanoscale DNA cascade switches that can be used to identify single molecular events. Significant challenges for analytical biotechnology arise in proteomics, glycomics, and lipidomics.

2. Materials biotechnology entails the use of biotechnological methods for the fabrication of novel materials with unique optical, electronic, rheological, and selective transport properties. Examples include novel polymers created from genetically engineered polypeptide sequences and the formation of nanowires and circuits from metal nanoparticles attached to a DNA backbone.

3. Computational biotechnology focuses on the potential replacement of silicon devices with nanoscale biomolecular-based computational systems. Examples include the creation of DNA switches from hairpin structures and the programmable self-assembly of DNA tiles for the creation of memory circuits.

A common feature of many of the three new biotechnology application areas is that they all require the production of well-characterized, functional biopolymer nanostructures. The molecular precision and specificity of the enzymatic biochemical pathways employed in biotechnology can often surpass what can be accomplished by other chemical or physical methods—a point that is especially relevant to the problem of nanoscale self-assembly. It is this fine control of nanoscale architecture exhibited in proteins, membranes, and nucleic acids that researchers hope to harness with these applied biotechnologies.

An important enabler of the production of such nanostructures, especially on a large scale, is the availability of increasingly standardized and increasingly automatable fabrication techniques. In some ways, the status of fabrication technologies for these nanostructures is similar to the status of integrated circuit fabrication technology several decades ago, which evolved from a laboratory activity with trial-and-error doping of individual devices to a large-scale automated enterprise driven by design automation software over a period of 20 years beginning in the early 1960s.

Although they draw on biology and computing (along with other disciplines), the tools of these parent disciplines are being applied by researchers in these new biotechnological areas to a different and unrelated set of scientific interests and goals, and these areas often attract scientists with no interests in or ties to traditional biology or computing research. Indeed, these researchers are likely to find intellectual homes in areas such as neuroscience, robotics, and space exploration.

These new areas also have obvious relevance to computing. For example, computational biotechnology is relevant to computing in the same way that lithographic silicon fabrication technologies are today—underpinning these latter technologies are understandings of fundamental physics and well-developed electrical engineering techniques and approaches. Similarly, computational biotechnology will draw on materials science and biochemistry as well as biology as it seeks to create highly regular DNA nanoparticles, mate DNA with submicron electronic structures fabricated in silicon, and create networks of interconnecting nanostructures with unique enzyme communication paths. Analytical and materials biotechnologies are also relevant for enabling MEMS—microelectromechanical systems that interact with the physical world (taking in data through various sensors and affecting the world through various actuators).

## 11.2 MOVING FORWARD

The committee believes that the most important barriers today impeding the broader integration of computing and information technology into life sciences research are cultural barriers. Twenty-first century biology will not entail a diminution of the central role that traditional empirical or experimental research plays, but it will call for the whole-hearted embrace of a style of biology that integrates reductionist biology with systems biology research. At the same time, computing and physical science

practitioners must be wary of underestimating the true complexity of biological systems and, in particular, of inappropriately applying their traditional intellectual paradigms for simplicity to biology.

Over the long run, a change in the culture of academic life sciences research is required to sustain the approaches needed for 21st century biology, due to the increased need for disparate skill sets and collaborative approaches, a change that emphasizes interdisciplinary teams that integrate biology and computing expertise. For this reason, the main focus of this chapter's conclusions and recommendations concern actions that can accelerate the required cultural shift. By contrast, reflecting the committee's view that the impact of biological research is likely to be more modest in scope and scale, the conclusions and recommendations place less emphasis on biology's impact on computing. (In this light, Chapters 4-8 of this report should not be seen as laying out a research agenda for computing-enabled biology or for biology-inspired computing, but rather as suggesting some of the areas in which the frontiers of the interface have been pushed—and that still hold considerable intellectual interest.)

### 11.2.1 Building a New Community

The most important target of promoting cultural change is people. Thus, it should be a key objective of science policy makers to create a large, multitalented population of individuals who can act as the intellectual translators and mediators along the frontier, a group that will directly foster interdisciplinary research and technology development. True for any discipline or research area involving disparate skill sets, such an approach is especially critical at the interface between the fields of biology and computing because these areas are enjoying the most rapid growth and intellectual progress. Both junior and senior talent must be cultivated, the former to be the basis of a next generation ready to develop and exploit the technology and conduct the science, and the latter to serve in mentorship and leadership roles.

This message is not a new one—indeed, private programs such as the Burroughs-Wellcome Foundation Interfaces in Sciences have avidly sought the development of community. Nevertheless, it remains true that despite many studies, reports, and proclamations, universities and federal funding agencies have fallen short of the goal of fully facilitating a range of interdisciplinary science and minimizing the birth pains associated with new hypotheses and directions.<sup>1</sup>

An essential aspect of this community is the ability to build on each other's work. Indeed, the most advanced and sophisticated cyberinfrastructure imaginable will be ineffective if different laboratories and researchers are not motivated or are unwilling to work together or to share data and other information. Formal collaborations between individual laboratories or researchers do exist, of course, but these exist entirely on the basis of individually negotiated arrangements between consenting parties. A different, and complementary, model of working together is one in which individuals researchers contribute to and draw from an entire research community. In spirit, this model is the familiar one of publishing research articles and supporting information (data, software) for others to cite and use as appropriate in their own research—and the dominant ethos of the new community should be one of sharing rather than withholding.

This section provides some core principles on how individuals and institutions might help to support and nurture such work. The core principles described here may come across as “motherhood and apple pie,” but it is often the case that such motherhood is not honored as fully as one might think appropriate. The committee does recognize the centrality of providing appropriate incentives for non-

---

<sup>1</sup>For example, a report was prepared by the National Institutes of Health (NIH) and the National Science Foundation (NSF) in August 2001 addressing many of the cultural issues described in Chapter 10. This report on training in bioengineering and bioinformatics, *Assessing Bioengineering and Bioinformatics Research Training, Education, and Career Development*, recommended that measures be taken to (1) increase the number of fellowships and institutional training grants at all career levels that include quantitative, computational biology and integrative systems modeling; (2) include funds to support faculty with complementary expertise (e.g., computer scientists to teach biologists); and (3) support the development of curricula. In the intervening 2 years, the importance of continued efforts in these areas has not diminished.

oring these principles. Both institutions and funding agencies have important roles to play in providing incentives for change when these principles are not honored and for continuity when they are. Thus, the core principles for institutions (Section 11.2.3) and for funding agencies (Section 11.4.1) should be seen partly in this light.

### 11.2.2 Core Principles for Practitioners

The following items are offered as advice to current and prospective researchers at the BioComp interface. These workers include those seeking to retrain themselves to work at the BioComp interface (e.g., a postdoctoral fellow with a computer science background working in a biology laboratory), those facilitating such retraining (e.g., the director of a biology laboratory employing such a postdoc), and those who collaborate as peers with others (e.g., a tenured professor of computer science working with a tenured professor of biology on some interesting problem). Practitioners should:

- *Respect their partners.* Neither the biologist who sees the computer scientist only as a craftsman writing computer programs for data analysis nor the computer scientist who sees the biologist as a provider of dirty and unreliable data shows respect for the other. Scientists with quantitative backgrounds and scientists with biomedical backgrounds must work as peers if their collaborations are to be successful.

- *Have reasonable expectations.* One's intellectual partners in an interdisciplinary endeavor will have differing and often unfamiliar intellectual paradigms. Both vocabulary and epistemology will be different, and a respect for other ways of looking at the world reflects an understanding that paradigms can be different for very sound reasons.

- *Avoid hype.* In the quest for funding and attention, practitioners need to maintain a high degree of questioning to avoid hype, unrealistic expectations, or empty promises.

- *Don't complain.* Complaining to close colleagues about the apparently poor science practiced by other disciplines further reinforces xenophobic arrogance and chauvinism. When the other parties sense such arrogance, the trust needed to achieve scientific collaboration is no longer available.<sup>2</sup>

- *Seek new techniques and intellectual inspiration everywhere.* Both biology and computer science have traditions of applying other disciplines to their problems. For example, Leeuwenhoek's optical microscope led to the discovery of cells, electrical recording devices revealed the voltage-gated channels in neuronal signaling, and knowledge of crystallography uncovered the helical structure and code of DNA. Computer science, originating from a marriage between electrical engineering and mathematics, continues to maintain close intellectual connections to these disciplines.

- *Nurture young talent.* The key to long-term growth of a new field is the ability to sustain and nurture young scientists working in that field. To the extent that attention can be focused on young scientists (e.g., targeting this generation with well-placed, exciting, and novel funding opportunities), problems of competing for funds with senior groups working on classical topics can be reduced.

The committee understands that these principles will have different meaning to researchers at different stages of their careers. For those early in their careers, these recommendations should be taken as a checklist of things to keep in mind as they engage with colleagues and seek support. However, these items are also relevant to senior researchers who serve as role models for their younger colleagues.

---

<sup>2</sup>G. Wiederhold, "Science in Two Domains," unpublished working paper, Department of Computer Science, Stanford University, March 2002, updated February 2003.

### 11.2.3 Core Principles for Research Institutions

The following items are offered as advice to institutions that are supporting work at the BioComp interface. These institutions include academic laboratories, research centers, and departments, as well as business or commercial operations with a research component. Collectively, these items are based on the barriers to collaboration and community discussed in Chapter 10. However, no attempt has been made to specifically align recommendations with barriers because in most cases, the correspondence is many-to-many rather than many-to-one or one-to-many.

Relevant institutions should:

- *Attract and retain professionals with quantitative, computational, and engineering skills to work in biological fields.* As a rule, recruitment and retention will require reasonable career tracks that hold the promise of long-term stability and upward mobility. If good individuals are to be attracted to and retained in any enduring interdisciplinary area, they must have career opportunities that offer the potential for growth. For example, these individuals must be assured that their intellectual work at the interface will be fairly evaluated. Such issues are matters of academic survival for many young faculty, and if processes are not put into place explicitly that ensure an appropriately rigorous but still fair evaluation process, promising faculty may well have strong disincentives to pursue research at the interface. A corollary is that traditional departments often see considerable opportunity cost in supporting (and granting tenure to) individuals who do not fit squarely in their centers of gravity (Section 10.3.3); thus, independent support for researchers with interdisciplinary interests, or support that cannot be converted to individuals with traditional interests, helps to remove the threat that departments may see.

- *Support retraining efforts.* Because much of the computing talent required at the BioComp interface will have to come from individuals with substantial prior experience in computing, retraining will be an essential part of efforts to build the talent base. Individuals considering retraining will be more motivated to do so if funding agencies and tenure and promotion committees wishing to support these faculty members recognize that retooling takes some time to be successful and do not penalize them for lowered productivity during such periods.

- *Develop curricula for interdisciplinary teaching of quantitative, computational, and engineering sciences made relevant to the BioComp interface.* Note the desirability of such curricula being made available in multiple formats—online versus in class, 2-week courses versus semester-length courses, and so on—as well as on multiple topics. Over the long run, it is likely that immersion in these curricula will become a natural part of the educational process for all budding biologists, but today, obtaining this background requires some special effort.

- *Facilitate networking.* Especially for newcomers to a line of work, intellectual connection to others plays an important role in their integration into the new community. An institution can promote informal knowledge exchange and the establishment of social relationships on campus through on-site seminars for like-minded individuals. It can also facilitate off-campus connections by providing support for travel to tutorials, workshops, and seminars.

- *Nurture partnerships.* It is desirable for senior scientists from different intellectual backgrounds to work at the interface and for peer relationships between biologists and computer scientists to develop. Partnerships are best undertaken in close proximity with intense interaction, and even small issues such as office arrangements (e.g., whether or not a computer scientist has an office or a desk in the laboratory of a collaborator or partner) can seriously inhibit the development of close partnerships.<sup>3</sup> Many scenarios could promote partnerships, such as sabbatical visits and the establishment of positions at cen-

<sup>3</sup>For example, a computer scientist developing software to aid in the analysis of biological data would be well advised to spend enough time in the laboratory to understand the actual needs of his or her biologist colleagues. Software delivered “over the transom” is unlikely to be used easily, a point suggesting that there is more to software design than the development of an appropriate algorithm.

ters of excellence. Partnerships with industry could ensure sabbaticals in complementary work environments and stimulate knowledge dissemination to commercial applications.

- *Recognize collaborative work.* A corollary of partnerships is that experts from disparate disciplines will collaborate in publication. Institutions thus have a responsibility to provide fair and appropriate evaluation measures for tenure and promotion cases in which the individuals involved have undertaken large amounts of collaborative work. For example, departments may have to be induced to expand their definitions of tenurable work, or universities may have to establish extradepartmental mechanisms for granting and holding tenure outside of traditional departments.

- *Maintain excellence.* Research at the BioComp interface is inherently interdisciplinary, and evaluation of such research faces all of the problems described above. Nevertheless, problem domains that are at the interface of two disciplines can attract not only highly talented individuals who see interesting and important problems but also individuals of lesser talent who are unable to meet the exacting standards of one discipline and are seeking a home where the standards of acceptance are lower. Individuals in the first category are to be sought and cherished—individuals in the second category ought to be shunned.

- *Provide mentors.* Mentors play a strong role in the success of any retraining effort. However, mentoring individuals who have an established track record of success in another field is different. For example, such individuals may be less able to work autonomously and more likely to flail or drift without an activist mentor than someone with a background in the same field. Shared mentorships may make particular sense in these circumstances, as illustrated by the Burroughs-Wellcome requirement that fellowship awardees have a mentor from outside the department of primary appointment.

- *Reward good behavior.* It has been observed that behavior that is rewarded institutionally is behavior that tends to take hold and to be internalized. The institutions with which individual researchers are associated can play important roles in providing such rewards, especially with respect to the principles described in Section 11.2.2.

### 11.3 THE SPECIAL SIGNIFICANCE OF EDUCATIONAL INNOVATION AT THE BIOCAMP INTERFACE

The pursuit of 21st century biology will require a generation of biologists who can appreciate fundamental statistical approaches, evaluate computational tools and use them appropriately, and know how to choose the best collaborators from the quantitative sciences as a whole. To support the education of this generation, an integrative education, whether formal or informal, will be needed.

Many reports have acknowledged a need for broader training.<sup>4</sup> Increasingly, bioinformatics programs at both the undergraduate and the graduate level do entail study in mathematics, computer science, and the natural sciences.

#### 11.3.1 Content

The committee fully supports these trends and encourages them further, with the strong caveat that an appropriate curriculum to deal with the interface of computing and biology should not simply be the union of course requirements from multiple departments. Courses and other work that deal explicitly with the integrative issues are necessary, and one of the most important skills that such interdisciplinary courses can teach is the ability to communicate among the relevant disciplines. This does not entail simply learning the jargon of each one (though this is, of course, essential), but also interleaving the training in such a way that the student continually sees and explores various parallels between the

---

<sup>4</sup>See, for example, National Research Council, *Bio2010: Undergraduate Education to Prepare Biomedical Research Scientists*, The National Academies Press, Washington, DC, 2003.

different fields of study. Later, in a collaboration, the ability to identify, explain, and exploit these parallels will be valuable.

Cultural barriers should be discussed and addressed specifically. Where it seems easy to dismiss some math or physics as irrelevant to biology, case studies can be assembled to show successes and contrast these with failures or counterproductive avenues. Where it seems easy to dismiss biology as too detail oriented and reductionistic, similar case studies showing the need to understand minute details of the living machinery are also necessary.

It is broadly agreed that an essential element of 21st century biology is the (re)introduction of quantitative science to the biological science curriculum. The committee recognizes, however, that such reintroduction should not be equated with an abstract, theoretical approach devoid of experimentation or phenomenology, and educational programs for 21st century biology must provide sound footing in quantitative science alongside a clear understanding of the intricacies of biology.

In light of the discussions in Chapter 6 regarding the view of biological organisms as engineered entities, the committee believes that students of 21st century biology would benefit greatly from some study of engineering as well. In this view, the committee emphasizes most strongly its support for the recommendations of the BIO2010 report for exposure to engineering principles (discussed in Chapter 10), at the earliest possible time in the training of life scientists. Just as engineers must construct physical systems to operate in the real world, nature also must operate under these same constraints—physical laws—to “design” successful organisms. Despite this fundamental similarity, biology students rarely learn the important analysis, modeling, and design skills common to engineering curricula nor a suite of topics such as engineering thermodynamics, solid and fluid dynamics, control theory, and so forth, that are key to the engineer’s (and nature’s) ability to design physical systems.

The particular area of engineering (electrical, mechanical, computer, and so forth) is probably much less relevant than exposure to essential principles of engineering design: the notion of trade-offs in managing competing objectives, control systems theory, feedback, redundancy, signal processing, interface design, abstraction, and the like (Box 11.1). Ready intellectual access to such notions is likely to enable researchers in this area to search for higher-level order in the data forest. Indeed, as biology continues to examine the system-wide functioning of a large number of interacting components, engineering skills may become necessary for successful biological research.

### 11.3.2 Mechanisms

The committee believes that the availability of individuals with significant computing expertise is an important limiting factor for the rate at which the biological sciences can absorb such expertise.<sup>5</sup> The field, to include both basic and applied life sciences research, is extraordinarily large and dwarfs most other fields outside of engineering itself; thus, influx from other fields is not likely to result in large-scale infusion of computing expertise. Only integrated education of new researchers, along with some re-training of existing researchers, can bring benefits of the computing to a large segment of that world, and previous calls from groups such as Biomedical Information Science and Technology Initiative (BISTI) that a new generation of 21st century researchers must be trained remain compelling, true, and overdue.

Given this perspective, it is appropriate to offer educational opportunities across a broad front. Educational opportunities should span a range in several dimensions, including the following:

- *Time and format.* Monthly lectures or seminars, short-duration workshops (of several weeks), survey courses, undergraduate minors, undergraduate majors, graduate degrees, and postdoctoral

---

<sup>5</sup>This belief is not based on the existence of a “shortage” or “scarcity” in the sense that economists generally recognize. Rather, it is rooted in the premise that most of biology could benefit intellectually from the integration of significant computing expertise, and the observation that such integration is more the exception than the rule when taken across biology writ large.

**Box 11.1**  
**Some Engineering Ideas and Concepts That Biologists May Find Useful**

- Control theory (feedback, optimization, game theory)
- Model design
- Signal processing (gain, signal-to-noise, cross-talk)
- Engineering thermodynamics and energy
- Optimal design theory
- Modularity (and protocols)
- Robustness
- Multiscale and large-scale stochastic simulation
- Network theory or graph theory
- Fluid and solid dynamics or mechanics
- “Collective behavior” from physics
- Reverse engineering
- Computational complexity (decidability, P-NP)
- Information theory, source and channel coding
- Dynamical systems: dynamics, bifurcation, chaos
- Statistical physics: phase transitions, critical phenomena

(re)training focusing on the BioComp interface can serve to motivate interest (when they require little investment or time commitment) or to serve a strong professional interest (when the time commitment required is substantial).<sup>6</sup> Short-term opportunities for cross-disciplinary “pollination” workshops that bring together fields from both sides of the interface and provide a vehicle for tutorials and other educational exchanges are particularly useful in that they have a low cost of entry for participants; thus, those who are dabbling can be enticed more easily.

- *Content.* Although genome informatics is perhaps the most obvious topic, computational techniques and approaches will become increasingly relevant to all aspects of biological research—and educational opportunities should target a wide range of subfields in biology.

- *Target audience.* Given the need for more computing expertise in biology, it is appropriate to provide instruction at multiple levels of sophistication in different fields. Some research biologists have substantial informal computing experience but would benefit greatly from more formal exposure; such

---

<sup>6</sup>In this regard, the model of statistics as a discipline may offer a good example for the way in which bioinformatics might become a discipline of its own. Many universities offer three types of programs in statistics. The first and most formal program is designed for those aspiring to become professional, academic statisticians—that is, those aspiring to become researchers in the field of statistics. This program usually culminates in the Ph.D. degree and establishes an absolutely sound theoretical understanding of the foundations of statistics. The second program is intended for individuals who intend to become professional applied statisticians—that is, those who will work in industry, perhaps in research, and whose primary responsibilities will involve carrying out statistical analyses using established statistical methods. Often, individuals pursuing this degree track will stop with a master’s degree and in some cases even with a bachelor’s degree. The third program involves a set of courses intended for individuals who will be getting a degree in another field, but who have a need for significant understanding of statistical methods so that those methods might be applied in the individual’s home field. In very large universities, sometimes these third-track courses are specialized even further so that we might see courses in business statistics, biological statistics, or even medical statistics. Similarly, it seems reasonably clear that in the field of bioinformatics there will always be a need for researchers, whose primary interest will be in devising new algorithms, new models, and new methods in bioinformatics. There will also be a need for applied bioinformaticians—those whose primary responsibility will be in applying established bioinformatics methods to current projects in biology and biotechnology. It also seems reasonable to suppose that there will be those whose careers in other disciplines will be enhanced by some knowledge of bioinformatics. (These latter two program types may have to provide the biological knowledge to those trained primarily in computation or the computational overview to those trained primarily in biology.)

individuals are in an obviously different situation than those whose only exposure to computing is spreadsheets and word processors.

The development of such educational opportunities generally requires resources, such as release time; assistance in compiling lecture notes, assembling readings, or grading; funding for developing online courses, travel to workshops, and so on. Furthermore, it is desirable to share the outcomes of such development with the academic community (e.g., in the form of online courses, published books, and open commentary about successes and failures). Funding agencies can also provide incentives for such cooperative efforts by giving higher funding priority to research proposals that are put forward in partnerships between or among universities.

#### 11.4 RECOMMENDATIONS FOR RESEARCH FUNDING AGENCIES

The committee believes that it is possible—and feasible—for agencies to support work at the BioComp interface that serves to develop simultaneously (1) fundamental knowledge that enables broad advances in biology; (2) technical innovations that help to improve the quality of life and enhance industrial competitiveness; and (3) the creation and sustenance of a critical mass of talented scientists and engineers intellectually capable and professionally positioned to work creatively at the BioComp interface and to train new generations effectively.

Funding agencies and nongovernmental supporters of research have traditionally been able to influence the course of research through the allocation of resources to particular research fields, and the committee believes that funding at the biology-computing interface is no exception. This support has made important contributions in the past, and the committee urges that such support be continued and expanded.

##### 11.4.1 Core Principles for Funding Agencies

Recognition of the importance in focusing on the BioComp interface amplifies earlier agency-centered studies and reflects its unprecedented richness. Responding to the opportunities, the scientific community, private foundations, and the federal government have taken the first steps in recognizing this enormous intellectual opportunity.

However, no single agency—let alone any individual program, directorate, institute, center, or office—owns the science or the excitement and promise at the interface between computing and biology. Neither can a single agency by itself establish and sustain a process to realize the grand opportunities. In their growing commitment to this frontier science effort, the Defense Advanced Research Projects Agency (DARPA), the National Science Foundation (NSF), the National Institutes of Health (NIH), and the Department of Energy (DOE) each have unique objectives and existing expertise. To exploit the potential fully, the agencies, more than ever before, will have to collaborate and also seek (formal or informal) partnerships with private foundations and industry. Extensive interactions including fully open, joint planning exercises and shared support for technical workshops will be central to true coordination at the agency level.

As is the case for individuals and institutions, a number of core principles provide good desiderata for the funding policies and practices of agencies. Again, these core principles are not particularly new—but remain essential to realizing goals at the BioComp interface. Of course, how these principles are instantiated is key.

To obtain maximum impact, funding agencies and foundations should pay appropriate attention to the following items. Agencies and foundations should:

- *Support awards that can be used for retraining purposes.* While a number of agencies have supported such awards for individuals at early stages of their careers, these programs are fewer in number than in

the past. Also, to the best of the committee's knowledge, there are no programs that explicitly target senior faculty for retraining at the BioComp interface, although, as noted in Section 10.2.2.6, NIH does support a retraining program open to scientists of many backgrounds to undertake biomedical research. To the extent that such programs continue to exist, agencies should seek to publicize them beyond their usual core constituencies.

- *Balance quality and excellence against openness to new ideas in the review process.* Intellectual excellence is central. Yet especially in interdisciplinary work, it is also important to invest in work that challenges existing assumptions about how research in the field "should" be conducted—and the problem is that traditional review mechanisms often have a hard time distinguishing between proposals for such work and proposals for work that simply does not meet any reasonable standard of excellence. This point suggests that agencies wishing to support work at the BioComp interface would be wise to find review mechanisms that can draw on individuals who collectively have the relevant interdisciplinary expertise and, as importantly, an appropriate forward-looking view of the field.

- *Encourage team formation.* It is important not to discriminate against team-researched articles in individual performance evaluations and to provide incentives for universities to reward multiple members of cross-disciplinary teams of investigators. Under today's arrangements, work performed by an individual as part of a team often receives substantially less credit than work performed by an individual working alone or with graduate students.

- *Provide research opportunities for investigators at the interface who are not established enough to obtain funding on the strength of their track record alone.* In these instances, balance must be struck between taking a chance on an unproven track record and shutting down nonfruitful lines of inquiry. One approach is to set time limits (a few years) on grants made to such individuals, requiring them to compete on their own against more established investigators after the initial period. (As in other fields, the duration of "a few years" is established by the fact that it is unreasonable to expect significant results in less time, and norms of regular funding set an upper limit for this encouragement of work outside the boundaries.)

- *Use funding leverage to promote institutional change.* That is, agencies can give priority or differential advantages to proposals that are structured in certain ways or that come from institutions that demonstrate commitments to change. For example, priority or preference could be given to proposals that

- Involve co-principal investigators from different disciplines;

- Originate in institutions that offer grant awardees tenure-track faculty appointments with minimal teaching responsibilities (as illustrated by the Burroughs-Wellcome Career Awards (Section 10.2.2.5.2));

- Have significant and active educational efforts or programs at the BioComp interface; and

- Make data available to the larger biological community in standard forms that facilitate reuse and common interpretation.<sup>7</sup> (This action is predicated on the existence of such standards, and agencies should continue to support efforts to develop these common data standards.)

- *Use publication venues to promote institutional change.* Funding agencies could require as a condition of publication that authors deposit the data associated with a given publication into appropriate community databases in accordance with relevant curation standards. They could also insist that published work describing computational models be accompanied by assurances that detailed code inspection of models is possible under an appropriate nondisclosure agreement.

---

<sup>7</sup>The committee notes without comment that the desire on the part of science agencies to promote wider data sharing and interoperability may conflict with requirements emanating from other parts of the federal government with regard to information management in biomedical research. While science agencies are urging data sharing, other parts of the government can impose restrictions on sharing biomedical data associated with individual human beings in the name of privacy, and these restrictions can have significant impact on the architecture of biomedical information systems. In some cases, these regulatory compliance issues have such impact that biomedical scientists have strong incentives to introduce a *paper* step into their data management processes in order to escape some of the more onerous consequences of these regulations for their information systems.

- *Support cyberinfrastructure for biological research.* Though the National Science Foundation has taken a lead in this area, the issue of supporting cyberinfrastructure for biological research transcends any single agency. Chapter 7 discussed the importance of data repositories and digital libraries in cyberinfrastructure, and it is in these areas that other agencies have important roles to play. Across the board, agencies engaged in supporting biological research will need to support mechanisms for long-term data storage and for continuous curation and annotation of the information resources gathered in publicly supported research for 21st century biology to reach its full potential as a global distributed intellectual enterprise.

- *Recognize quality publicly.* Given the role of peer recognition in the value sets of most scientists (especially in their earlier years), public recognition of innovative work can be a strong motivator. Public recognition can take many forms—though by definition the number of people that can be recognized is necessarily limited. For example, outstanding researchers can be invited to give keynote addresses at important conferences or profiled in reports to Congress or other important public documents.

- *Recognize the costs of providing access to computing and information resources.* Especially at the BioComp interface, collaboration between peers as compared to an investigation conducted by an individual researcher almost always requires larger grants. Researchers need more support for computing and information technology as well as the expertise needed to exploit those capabilities and, in instances that push the computing state of the art, support for high-level expertise as well.

- *Define specific challenge problems that stretch the existing state of the art but are nevertheless amenable to progress in a reasonable time frame.* An agency could pose challenge problems drawn from the problem domains described in Chapter 9. Any number of such challenge problems would be arbitrary, but a selected few goals of broad impact would influence more complete participation by the community and make further funding opportunities by other agencies more likely. Note that when common test sets or other common criteria can be provided or used, clearer metrics for success can be established. A corollary is that agencies should obtain community buy-in with respect to the specifics of such problems. (As one example, the DOE Office of Biological and Environmental Research specified what microbes to tackle for complete genome sequencing through a series of “which bug” workshops to obtain community input on the projects that would be best.)

- *Work with other agencies.* Different agencies bring to the table different types of expertise, and for work at the interface, multiple kinds of expertise are always necessary. Thus, agency partnerships (such as the current collaboration between NIH’s National Institute of General Medical Sciences (NIGMS) and NSF’s Mathematical and Physical Sciences Directorate) may allow proposals at the interface to be evaluated more fairly and ongoing projects to be overseen more effectively.<sup>8</sup>

- *Provide the funding necessary to capitalize on the intellectual potential of 21st century biology.* Chapters 2-7 of this report have sought to demonstrate the broad impact of computing and information technology on biology. However, a necessary condition to realize this impact is a funding stream that is adequate in magnitude and sustained over long enough periods. As noted in Section 10.3.5.2, a benchmark for comparison is that spending in information-intensive fields such as finance is on the order of 5 to 10 percent of overall budgets. A second necessary condition is the use of a peer review process that is broadly sensitive to the perspectives of researchers in the new field and is willing to take chances on new ideas and approaches. As always, the public sector should focus on growing the seed corn for both people and ideas on which the future depends. Finally, although the committee would gladly endorse an increased flow of funding to the furtherance of a truly integrated 21st century biology, it does understand the realities of a budget-constrained environment.

---

<sup>8</sup>This partnership between the NIGMS and the NSF seeks to award 20 grants in mathematical biology and anticipates more than \$24 million in awards over 5 years. NIGMS supports research and training in the basic biomedical sciences. NSF funds mathematical and other quantitative sciences such as physics, computer science, and engineering. See <http://www.nigms.nih.gov/news/releases/biomath.html>.

The following sections are addressed to specific funding agencies.

#### 11.4.2 National Institutes of Health

As the largest funder of life sciences research, the NIH has a special responsibility to support and facilitate the building of bridges between biology and other disciplines, especially including computing. The NIH has already taken a number of commendable steps in seeking to collaborate with other agencies, including the formal NIGMS partnership with NSF for mathematical biology mentioned in the previous section and other less formal partnerships with NSF and DOE for structural biology and the National Center for Research Resources (NCRR)-NSF collaborations in instrumentation. As noted in Chapter 10, a National Research Council report in 2003 called for NIH to increase investment in high-risk, high-potential-payoff life sciences research that would be supported outside the usual NIH peer review system.

Such steps, and others like them, are to be encouraged. At the same time, NIH must address obstacles in a number of other areas that impede the building of bridges between biology and computing. One important issue is that cooperation across organizational boundaries within NIH leaves much to be desired. Translational medicine will not arise from funding mechanisms that isolate narrow slices of human biology, and yet the NIH structure is oriented toward specific diseases and body functions.<sup>9</sup> No component of a human works separately, in isolation. Most diseases are not single-gene defects, most proteins act in macromolecular assemblies, organ systems interact by chemical messengers, the immune system and the circulatory system not only work together but impact all organs of the body, and so on. The NIH structure has been successful for many years, but the fact remains that its organizational structure tends to place similar restraints on cross-institute support for collaborative research (Box 11.2).

A consequence of organization of research fields in biology by subfield (e.g., by disease, or by body function) is that efforts that can benefit the entire community may suffer, even though specialization is necessary to achieve depth of knowledge. The true value of the large-scale deployment of cyberinfrastructure—and especially its data components—is that cyberinfrastructure spans disciplines to integrate findings in one subfield with findings in another subfield—to connect information from one subfield to another subfield, perhaps even via a third subfield. In the absence of explicit direction and coordination, cyberinfrastructure in one subfield is likely to be incompatible in important ways with cyberinfrastructure designed and deployed in another. Achieving coordination is likely to require a level of cooperation across agencies that is substantially greater than has historically been true. It will also require a level of planning and agency involvement in the actual design of the cyberinfrastructure that does not typically happen in the funding of research, in which the role of program officers is primarily to ensure a fair assessment of the science by peer reviewers. In supporting cyberinfrastructure, program officers must act as procurement officers on behalf of the overall scientific community, and not just as impartial brokers of an independent discipline-focused review process.

---

<sup>9</sup>There are 20 institutes within the National Institutes of Health, including the National Cancer Institute (NCI); the National Eye Institute (NEI); the National Heart, Lung, and Blood Institute (NHLBI); the National Human Genome Research Institute (NHGRI); the National Institute on Aging (NIA); the National Institute on Alcohol Abuse and Alcoholism (NIAAA); the National Institute of Allergy and Infectious Diseases (NIAID); the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS); the National Institute of Biomedical Imaging and Bioengineering (NIBIB); the National Institute of Child Health and Human Development (NICHD); the National Institute on Deafness and Other Communication Disorders (NIDCD); the National Institute of Dental and Craniofacial Research (NIDCR); the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK); the National Institute on Drug Abuse (NIDA); the National Institute of Environmental Health Sciences (NIEHS); the National Institute of General Medical Sciences (NIGMS); the National Institute of Mental Health (NIMH); the National Institute of Neurological Disorders and Stroke (NINDS); the National Institute of Nursing Research (NINR); and the National Library of Medicine (NLM).

**Box 11.2**  
**The Shapiro Report on the Structure and Organization of the  
 National Institutes of Health**

A 2003 National Research Council report on the structure and organization of NIH came to conclusions and made recommendations that are consistent with the view of NIH described in this report. Specifically, the earlier report noted:

[T]here is a high payoff potential for carefully selected large- and small-scale strategic projects that require the participation of numerous organizations working in partnership. . . . Well-planned, broad-based, trans-NIH programs will be necessary to meet most effectively scientific or public health needs. . . . Furthermore, there is no formal mandate for NIH to identify, plan, and implement such crosscutting strategic initiatives. [Such crosscutting initiatives are necessary because] scientific mechanisms, risk factors, and social and behavioral influences on health and disease cut across traditional disease categories. Many patients have multiple chronic conditions, so a patient-centered approach to health care and health promotion will sometimes require integration and synergy across [Institutes and Centers]. [Such issues] lend themselves to a strategic coordinated trans-NIH response in which multiple institutes could collaborate on a research plan that cuts across administrative structures in terms of planning, funding, and sharing and disseminating results. . . . Proteomics . . . is [an] example [of such an issue]. . . . [C]oncerted trans-NIH work on the assessment of existing and emerging technology platforms and database formats utilizing reference specimens, could help to advance the whole field and guide NIH-supported studies.

The report went on to recommend that initially 5 percent of the NIH budget and eventually 10 percent should be allocated to the support of such trans-NIH initiatives.

---

SOURCE: National Research Council, *Enhancing the Vitality of the National Institutes of Health: Organizational Change to Meet New Challenges*, The National Academies Press, Washington, DC, 2003, pp. 84-86.

NIH also supports some of the most scientifically sophisticated research environments in the world. As noted in the Botstein-Smarr report,<sup>10</sup> it is in these environments that it makes the most sense to train the leaders of the new generation of biologists with computing expertise. These environments are generally mature enough to support the conduct of interdisciplinary research at the interface, and a widespread geographical diffusion of young scientists with such expertise will help to generate the broad impact sought by NIH.

Perhaps the most important barrier of all is the philosophy that governs much of the current study group approach to proposal review. For historical reasons, the most important and prominent supporters of life sciences research—such as NIH—have focused almost exclusively on hypothesis-testing research—research that investigates well-isolated biological phenomena that can be controlled or manipulated and hypotheses that can be tested in straightforward ways with existing methods. This focus is at the center of reductionist biology and has undeniably been central to much of biology's success in the past several decades.

At the same time, the nearly exclusive focus on hypothesis testing has some important negative consequences. For example, experiments that require breakthrough approaches are unlikely to be directly supported. Just as importantly, advancing technology that could facilitate research is almost always done as a sideline. Thus, investigators must often disguise an attempt to undertake the development of tools or models of great generality by applying them to some (any!) biological system. Subsequent citations of such papers are almost always for the part that explains the new tool or model rather than the phenomenon to which the tool or model was applied.

---

<sup>10</sup>NIH Working Group on Biomedical Computing, *The Biomedical Information Science and Technology Initiative*, June 1999. Available at <http://www.nih.gov/about/director/060399.htm>.

This has had a considerable chilling effect in general on what could have been, but the impact is particularly severe for implementation of computational technologies within the biological sciences. That is, in effect as a cultural aspect of modern biological research, technology development to facilitate research is not considered real research and is not considered a legitimate focus of a standard grant. Thus, even computing research that would have a major impact on the advancement of biological science is simply not done.

The committee believes that 21st century biology will be based on a synergistic mix of reductionist and systems biologies. For systems biology researchers, the committee emphasizes that hypothesis-testing research will continue to be central in providing experimental verification of putative discoveries—and indeed, relevant as much to studies of how components interact as to studies of components themselves. Thus, disparaging rhetoric about the inadequacies and failures of reductionist biology and overheated zeal in promoting systems biology should be avoided. For researchers more oriented toward experimental or empirical work, the committee emphasizes that systems biology will be central in formulating novel, interesting, and in some cases, counterintuitive hypotheses to test. The point suggests that agencies that have traditionally supported hypothesis-testing research would do well to cast a wide “discovery” net that supports the development of alternative hypotheses as well as research that supports traditional hypothesis testing.

#### 11.4.3 National Science Foundation

The primary large-scale initiative of NSF relevant to 21st century biology is its cyberinfrastructure effort. Efforts in this area, including major community databases, collaborative research networks, and interdisciplinary modeling efforts, will require grants that are larger than the Directorate for Biological Sciences (BIO) of NSF has traditionally made, as well as greater continuity and stability. In particular, cyberinfrastructure entails personnel costs (e.g., for programmers, systems administrators, and staff scientists with the necessary computing expertise) that are not associated with the usual BIO-supported grant. As for continuity, windows for support must be consistent with the practical considerations to achieve success. Five-year awards and initial review at that point against specific milestones and deliverables to the community are essential, and only at longer intervals should there be open calls for proposals and competitive processes, save in the case of a resource failing to live up to community expectations.<sup>11</sup>

The professional biological community at large has at least two important roles to play with respect to cyberinfrastructure. First, it must articulate its needs and explicate how it can best exploit the resources that cyberinfrastructure will make available. Second, it must develop a consensus on the expectations that cyberinfrastructure facilities must meet if they are to be continued. Society events (e.g., annual meetings) provide a forum for such discussions to take place.

#### 11.4.4 Department of Energy

The DOE’s Office of Science supports a number of programs in genomic studies and structural biology (as described in Chapter 10). This office has the capacity to provide sufficient funds and a stable environment, but doing so has been a challenge in its overall institutional setting. The committee believes that the payoffs for DOE missions will be extraordinary from the biology supported by the Office of Science, but success requires that priority be given to stable, long-term programs.

---

<sup>11</sup>In principle, review provisions could be analogous to the sunset considerations for NSF-supported Science and Technology Centers and Engineering Research Centers.

#### 11.4.5 Defense Advanced Research Projects Agency

Of all the federal agencies, DARPA appears to be the most heavily involved in exploring the potential of biology for computing. Chapter 8 describes a variety of potential influences of biology on computing (the term “applications of biology for computing” would be promising too much), but in truth, the ultimate value of biology for changing computing paradigms in deep and fundamental ways is as yet unproven. Nevertheless, various biological attributes—robustness, adaptation, damage recovery, and so on—are so desirable from a computing point of view that any intellectual inquiry is valuable if it can contribute to artificial humanly purposive systems with these attributes.

In other words, investigations that consider the impact of biology on computing are—in the vernacular—high-risk, high-payoff studies. They are high risk because biology is not prescriptive in its contributions and success is far from ensured. They are high payoff because computers that possess attributes associated with biological systems would be enormously valuable. It is for this reason that they do logically fall into programs supported by DARPA, which has a long tradition of supporting high-risk, high-payoff work as part of its research portfolio. (As noted in Chapter 10, NSF also sponsors a Small Grants Exploratory Research Program that supports high-risk research on a small scale.)

From the committee’s perspective, the high-level goals articulated by DARPA and other agencies that support work related to biology’s potential contribution to computing seem generally sensible. This is not to say that every proposal supported under the auspices of these agencies’ programs would necessarily have garnered the support of the committee—but that would be true of any research portfolio associated with any program.

One important consequence of supporting high-risk research is that it is unlikely to be successful in the short term. Research—particularly of the high-risk variety—is often more “messy” and takes longer to succeed than managers would like. Managers understandably wish to terminate unproductive lines of inquiry, especially when budgets are constrained. However, short-term success cannot be the only metric of the value of research, and when it is, funding managers invite hyperbole and exaggeration on the part of proposal submitters, and unrealistic expectations begin to characterize the field. Those believing the hyperbole (and those contributing to it as well) thus overstate the importance and centrality of the research to the broader goal of improving computing. When unrealistic expectations are not met (and they will not be met, almost by definition), disillusionment sets in, and the field becomes disfavored from both a funding and an intellectual standpoint.

From this perspective, it is easy to see why support for fields can rise rapidly only to drop precipitously a few years later. Wild budget fluctuations and an unpredictable funding environment that changes goals rapidly can damage the long-term prospects of a field to produce useful and substantive knowledge. Funding levels do matter, but programs that provide steady funding in the context of broadly stated but consistent intellectual goals are more likely to yield useful results than those that do not.

Thus, the committee believes that in the area of biologically inspired computing, funding agencies should have realistic expectations, and these expectations should be relatively modest in the near term. Intellectually, their programs should continue to take a broad view of what “biological inspiration” means. Funding levels in these areas ought to be established on a “level-of-effort” basis (i.e., what DARPA believes is a reasonable level of effort to be expended in this area), taking into account the number of researchers doing and likely to do good work in this area and the potential availability of other avenues to improved computing. Also, programmatic continuity should be the rule, with playing rules and priorities remaining more or less constant in the absence of profound scientific discovery or technology advances in the area.

### 11.5 CONCLUSIONS REGARDING INDUSTRY

Over the past decade, the commercial sector has provided important validation for the proposition that information technology (IT) can have a profound impact on the life sciences. As noted in Chapter

10, there are a host of firms, ranging in size from small start-ups to established multibillion-dollar companies that have significant investments in research efforts and products, that make substantial use of IT in support of medical and pharmaceutical business.

Nevertheless, the committee is aware that some large life science companies (e.g., large pharmaceutical companies) have not found their investments in information technology living up to their expectations. Some such companies have reported investing a great deal of money and time in bioinformatics software and are now looking for and failing to find economic justification for further investment.

The hype of the genome era was as intoxicating to many drug companies, it seems, as the Internet was to mainstream investors, with just as much a comedown. There is a growing realization that the availability of genomic information is not, by itself, sufficient to lead directly to immediately profitable drug breakthroughs, regardless of the IT available to help manage and analyze that information. Indeed, many bottlenecks in drug discovery remain that result from the lack of fundamental biological knowledge about specific expression and pathways. Whereas the initial expectation was that the genome could be mined for likely drug targets, today's approach involves a greater tendency to start with the biology that is known to select likely targets, and then to look to the genome to find genes that interact with those targets.

The committee believes that bioinformatics—and broader uses of information technology—are likely to have a positive effect on drug discovery in the long run, but that those enterprises looking to investments in IT for short-term gain are likely to continue to be disappointed. Commercial advantages to the use of IT will accrue from its integration into the entire process, from gene discovery to clinical trials, benefiting both the entire process and the local situation to which information technology is applied. Also, because of rapidly increasing biological knowledge, the promise of discovering appropriate drug targets in the genome remains, although it is likely to be realized primarily in the long term. Bioinformatics will also enable a more precise genome-based identification of individuals susceptible to a given drug's side effects, possibly providing a basis for excluding them from clinical trials and pharmaceutical applications involving that drug.

## 11.6 CLOSING THOUGHTS

The impact of computing on biology could fairly be considered a paradigm change as biology enters the 21st century. Twenty-five years ago, biology saw the integration of multiple disciplines from the physical and biological sciences and the application of new approaches to understand the mechanisms by which simple bacteria and viruses function. The impact of the early efforts was so significant that a new discipline, molecular biology, emerged, and many biologists, including those working at the level of tissues or systems and whole organisms, came to adopt the approaches and often even the techniques. Molecular biology has had such success that it is no longer a discipline but simply part of bioscience research itself.

Today, the revolution lies in the application of a new set of interdisciplinary tools: computational approaches will provide the underpinning for the integration of broad disciplines in developing a quantitative systems approach, an integrative or synthetic approach to understanding the interplay of biological complexes as biological research moves up in scale. Bioinformatics provides the glue for systems biology, and computational biology provides new insights into key experimental approaches and how to tackle the challenges of nature. In short, computing and information technology applied to biological problems is likely to play a role for 21st century biology that is in many ways analogous to the role that molecular biology has played in biological research across all fields for the last quarter century—and computing and information technology will likely become embedded with biological research itself.

