Remote Biology Labs

Austin J. Che

Massachusetts Institute of Technology

February 19–21, 2005

Abstract

Purpose: The impact of biology in this century will be enormous. As engineers bring the traditional science of biology to consumers, everyone will be capable of tinkering with biological systems. Just as the personal computer allowed ordinary people to apply the physics of electricity and magnetism, molecular biology is entering an era with easily available technology for manipulating living systems. I present a *proposal* for the development of biology engineering education along with a discussion on the *responsible development of e-learning.*

Body: Although biological materials are relatively inexpensive, some necessary equipment is beyond the budget of the home biologist or standard educational program. To borrow from the computer industry, a time-sharing mainframe model in biology providing a centralized remotely-controllable biology lab would allow for distributing costs, more efficiently using resources, increasing scheduling efficiency, and, ultimately, higher work output. The same technology applied to research would bring similar benefits, leading to an expected high demand to drive reductions in cost.

Implications: A programmable biology lab encourages more time spent on designing, planning, and analyzing experiments rather than worrying about inane details such as how to pipette. As the time needed for manual work is transferred to creative time in the brain, an explosion in biological tinkering can be expected. As with the spread of any technology, we need to think about ethical and safety issues. The development of a "bio-hacker" culture is not necessarily good or bad, but should be carefully controlled. Some computer hackers write computer viruses, but the vast majority use their skills to satisfy society's needs. Openly providing necessary resources will allow the controlled growth of biology education. The centralization of resources and the increase in educational possibilities, even as it encourages the spread of knowledge, helps monitor activities and prevent accidents.

1 The Vision of Biology

Alice is 13 years old and has just received a BioConstructor Kit as a gift. Connecting to a website for a biology lab located halfway around the world, she examines the raw materials she has at her disposal. Being her first exposure to molecular biology, she decides to build a simple biological system. Alice follows the tutorial for inserting the green fluorescence protein from jellyfish into bacteria, making the bacteria glow green. She plans out each step deciding what enzymes to use, how long to incubate the solutions, and the website confirms her instructions, scheduling it for execution in the next available time slot. One week later, Alice checks on the status of her experiment and is excited to see her experiment has completed successfully. As she looks at pictures showing her bacteria glowing green, she begins to think about her next experiment.

Bob is a freshman at BTU (Bio-Tech University) eagerly taking his first biology class. The subject is the polymerase chain reaction (PCR) and their assignment is to test how various parameters affect the process. He decides to experiment with varying the temperatures used during PCR. Bob and his classmates connect to the online biology lab and submit their proposed experimental plan. After the instructor has looked over and approved their plans, the remote system automatically carries out the experiments, efficiently scheduling the requests of all students in the class.

Alice and Bob take the technology for granted and do not realize they are in the middle of a hypothetical example in 2005. But, with some planning and work, this can be turned into reality by 2015.

1.1 The Looming Need

Classical biology is a *science*. A scientist discovers what exists and expands the realm of knowledge. Thus, science is inherently unpredictable. The emerging field of *synthetic biology* is being formed by engineers wishing to harness the power of biology in building novel systems (Ball 2004; Ferber 2004). An engineer brings something new into existence using principles that make it likely for success. Whereas scientists make their careers on unpredictability, engineers make their careers on predictability.

Historically, engineering brings the applications driving widespread dissemination and use of scientific knowledge. Electricity and magnetism within the science of physics was and remains beyond the knowledge of most people, but as the engineering discipline of electrical engineering and computers developed, the masses were able to utilize principles learned to build useful applications. Now teens like Alice can build and program a computer, without needing to understand anything about transistors.

As biology matures towards the predictability of engineering, changes in the content and method of education are needed to allow the spread of biological knowledge. My personal experiences with engineering biological systems and teaching molecular biology, and the resulting exasperations, have been the primary driving motivators for this proposal.

1.2 Biology Education

Biology, as used here, will refer primarily to molecular biology, and not to areas such as ecology or taxonomy. Molecular biology forms the foundation of modern biology and includes what would normally be taught in an introductory undergraduate biology course.

I have taught an undergraduate-level biology course for the past two summers. MIT runs a rigorous six-week academic summer program for high school seniors called MITE^{2}S (Minority Introduction to Engineering, Entrepreneurship and Science). Two summers ago, I created a new biology course for this program. During the first summer in 2003, I modeled the class on a typical introductory syllabus, such as used for MIT undergraduate courses. Standard lectures were used with weekly written problem sets. The class feedback I received was overwhelmingly negative. A summer program such as MITE^{2}S has limited funds and must cover all student costs. Most classes including mine did not have any textbooks. Thus, the job of teaching falls heavily and solely on the teacher's performance in the classroom. Perhaps that is partially why the feedback was so negative.

In the summer of 2004, I taught the biology class again. The written problem sets from the previous year were completely eliminated. Several online texts were chosen as primary resources for students, to act as a substitute for textbooks. Interactive online quizzes were used as a substitute for problem sets. Classes were structured around a case-based method where students had to research information about cases online. The difference in feedback between the first and second summers was amazing. It was as if the sign for everything flipped, with negative comments becoming positive. Either I made a miraculous improvement in my teaching or, more likely, the quality of my in-classroom teaching became irrelevant. As the responsibility for learning is subtly shifted away from the teacher on to the student, both students and teachers have an improved experience.

Engineering, in my experience, is much easier to grasp and more interesting for students than abstract science concepts. In addition, at the introductory level, the most effective way for stimulating interest is with direct, hands-on experiences. Notably missing from the revised curriculum were hands-on labs. The one remaining and much more difficult step to pushing learning biology online would have been the ability to run labs online. Although it is *possible* to teach molecular biology without biology labs, it is about as effective as teaching computer programming without computers.

Another kind of biology class focused entirely on engineering has been taught for the past several years at MIT. The synthetic biology design classes, with students ranging from freshmen to professors and with backgrounds ranging from computer science to biology, aim to design and specify new biological systems. Due to time constraints of being taught in one month, there has not been the time to build the designed systems during the class.

During the past summer of 2004, 5 universities across the country held a "competition" where teams tried to design and also build synthetic biological systems. Many of the participants had had no prior biology lab experience. At the end, perhaps the biggest challenge for many teams was worrying about the mundane aspects of lab work. One student who participated in this competition remarked at the conclusion that there was a "disconnect between the complicated designs and the implementation in the lab." The focus in this competition was supposed to be on the engineering and not on debugging lab protocols.

2 Impact of Remote Biology Labs

2.1 Improving Biological Work Efficiency

The total *productivity* of a student or a worker could be defined as the total output produced. Although maximizing productivity seems like a reasonable goal, this is not as useful a metric as the work *efficiency*. If someone has 100 hours to dedicate to a project, but only spends 80 productive hours on the project, then that is 80% efficiency. Loss of perfect efficiency arises from several factors.

Observation 1 *Work efficiency is inversely related to the average amount of time it takes to perform individual tasks.*

This effect is due primarily because of scheduling issues. For example, if a task requires two hours of non-stop work to complete, then a single hour of free time would be wasted, as the task could not be started until a two hour block is freed. Maximum efficiency comes when tasks can be started, stopped, and re-started at any point. As the average time to complete a task increases, the ability to fit the task into a fragmented schedule becomes much more difficult. Time fragmentation is an unavoidable consequence of our busy lives, so as task times increase, in the limit, no work can be accomplished. Biological tasks are inherently time-consuming, requiring hours to run reactions and days to grow cell cultures.

Observation 2 *Work efficiency is inversely related to the number of individual tasks.*

With a project where different tasks take different amounts of time, part of the productive time becomes spent in planning out the use of available resources like time and equipment. This planning time is absolutely necessary, but is overhead that does not contribute any productive output. Biology protocols usually consist of many individual steps that need to be planned out.

Observation 3 *Work motivation fluctuates and efficiency will be greatest when work can be done at the same time one has high motivation.*

Not surprisingly, motivation, determined by many complex factors, greatly impacts the amount of work done. But not only does the amount of motivation matter, *timing* is important. If I only become highly motivated to do work when I am at home and away from the distractions of the lab, it becomes immensely challenging to be efficient performing lab work. In addition, performing repetitive and peripheral lab work such as cleaning lab equipment instead of working on the relevant experiment does not instill high motivation in either students or researchers.

Computer programming or writing papers have efficiencies approaching 100%. They can be started and stopped whenever one wants and done whenever one has the motivation to work. Biology lab work has a low efficiency, due to tasks taking a long time to complete and requirements for a human to be physically present in the lab at specific times.

2.2 Remote Labs Help Manage Time

The above observations show that managing *time* is one of the critical factors for improving efficiency. A system decoupling design from implementation would allow for more efficient use of time for the student or researcher. What is needed is a general, intelligent, programmable, robotic system, providing for common biological operations such as the following:

- Labeling, finding, removing, and storing tubes
- Pipetting and mixing of reagents
- Incubating samples at various temperatures
- Transformation and plating of cells
- Interfacing with equipment such as DNA sequencers

For maximum benefit, the system should be fully controllable over the Internet, and a camera can be used to view the lab in real-time. Many of these common biological operations have already been automated, for example, in the human genome project, but have not been made easily usable and programmable. Also, in many applications, a mechanism for transferring physical materials into and out of the lab would be necessary, probably performed by technicians running the lab.

Remote online labs are not a new idea. Many remote labs built upon LabVIEW (National Instruments) exist, but existing remote labs are mostly related to electrical engineering (Henry 1999; Shen et al. 1999). As can be seen in a review of LabVIEW labs, the number of electrical engineering labs is at least an order of magnitude more than biology related labs (Ertugrul 2000). Even the biology related labs involve controlling standard electrical equipment applied to biological samples.

The MIT iLab project has the goal of creating real laboratories usable over the Internet. It aims to develop a common software architecture, allowing for the deployment of remote labs in diverse areas. Current labs exist in the area of chemical, civil, and electrical engineering. However, no biology related lab is yet available. They are interested in expanding their range of online labs, and I have discussed with them about creating a biology lab. The focus of the iLab framework is aimed at the general tasks needed to manage a lab before and after a session (Harward et al. 2004). Using iLab terminology, a remote biology lab would be most usefully designed as a “batched experiment.” Most biology experiments only require simple decisions that can be easily automated. A user would plan an entire experiment and then the system would execute it without user supervision, with the experiment expected to take a lengthy amount of time.

The most advanced automated biology lab system may be the “Robot Scientist” (King et al. 2004; Morton 2004). The Robot Scientist generates hypotheses, performs experiments to check those hypothesis, and analyzes the data to generate new hypotheses. The Robot Scientist goes further than the proposal here. For the purposes of education, eliminating the intellectual part of designing the experiment is not desirable. However, the Robot Scientist does demonstrate the feasibility of automating a sophisticated biological experiment, making a remote biology lab a feasible goal. The physical implementation of the Robot Scientist uses the Biomek 2000 automated liquid handling workstation. The only human intervention needed was to move microtitre plates manually between the incubator, the plate reader, and Biomek 2000. However, according to the authors, even this would have been trivial to automate, with some added cost. The Biomek 2000 costs $37,900 or less than a quarter of this price on eBay, showing that the necessary equipment is attainable without a massive budget. Although equipment suitable for a remote lab may not exist for every desirable biological procedure, what technology is available is not being used to the fullest extent and developing new technologies for automating common biological tasks has not been a high priority.

2.3 Student and Instructor Benefits

An online remote biology lab benefits students, instructors, and researchers. In classes, students should be focused on obtaining results. Although learning a *process* may be part of the educational experience, the details are usually not. For example, several levels of depth are possible to teach the polymerase chain reaction (PCR). At the shallowest level, PCR can be explained as a method for amplifying DNA (i.e. *what* is PCR). With a bit more depth, the molecular mechanism behind *how* PCR works can be taught. This is a common and easy way to teach PCR and there are many multimedia videos used to show PCR. To go further, the next level of depth would be to *do* PCR. This would include details such as how many microliters of each reagent to add or how to optimize a PCR reaction. These are important real world issues not usually discussed in the traditional discussion of PCR. Currently, implementing learning at this level requires all the resources necessary to run a full lab course. In addition, there will inevitably be the student who contaminates some reagent, messing up the experiments for the entire class. Contamination of PCR reagents is also an important real world issue, but for the purposes of learning PCR, is irrelevant and distracting. Pipetting and learning how to program a thermocycler are also real world issues, but are distracting for learning about PCR. The remote biology lab strikes the perfect balance for depth and hands-on experience without having to worry about issues not important in the learning process.

A real remote biology lab has benefits beyond a computer simulated biology lab. First, our current knowledge does not allow us to simulate most biological operations accurately. The unexpected elements that come from biological experiments is educational in itself.

Also, a simulated lab is limited by the simulator designer's creativity. A non-simulated lab leaves the creativity up to the students, giving them the opportunity to do research with state of the art equipment. In addition, the same remote lab can be used by scientists for their cutting-edge research, where simulation is not sufficient.

The most important limiting issues for a program like MITE^{2}S are time and, secondarily, cost. With only six weeks, each hour spent in class or lab is precious time that cannot be spent elsewhere. An online lab allows people to work on their own schedule. Auto scheduling of equipment and the batched experiment approach makes both people and equipment more efficient. In addition, all biologists have had experiments thrown back weeks due to issues like contaminated water because of careless pipetting or misused equipment. Trying to diagnose and fix these types of problems is time consuming. The goal of biology should be to automate all error-prone and repetitive steps.

Cost is also an important factor for everyone. In a traditional educational setting, equipment usually has an extremely low usage, as TAs and other supervision is often needed. If equipment can be used to capacity at all times, then the cost per experiment will be lowered for everyone. Biology classes may not have the money to invest in the equipment for a couple labs, but may easily be able to afford several dollars per student for the ability to run a real online lab. The marginal cost of biological supplies is small compared with the large fixed costs, making it economically feasible for sharing of the fixed costs. In addition, researchers would certainly pay for a tool that makes it cheaper and more efficient than working in the lab. Expensive equipment like electron microscopes have been remotely connected to the network, allowing researchers around the world access to top-of-the line equipment (Krause 1997). A remote biology lab allows countries or labs with few resources to share time in a lab located in a place with more resources. Furthermore, when misused, equipment may need to be repaired or replaced faster. By controlling what can be done, equipment is guaranteed to be only used in the expected manner, potentially saving a significant amount of cost, time, and training.

3 The Future Storm

Being involved in defining an engineering discipline within the science of biology has shown me a great need in the area of biology education. As a kid, programming my first computer to beep and play games made it a comfortable transition for me to enter the field of computer science. For the biological revolution to occur, kids need to be exposed to the technology long before they enter college and decide what they want to do with their lives. Readily available kits for novices exist in areas such as electrical engineering or mechanical engineering. For example, LEGOs has been successful in many areas, but especially in motivating the integration of designing, building, programming, and testing robotic devices (e.g. Erwin et al.). As biology transitions from science to engineering, much of the technology will become generally obtainable. Introducing Alice at an early age to biology in a safe and fun way, by giving her bio-kits, may lead her to dedicate her life to the field.

A centralized general biology lab allowing for all work to be done remotely, leads to shared costs, more efficient use of resources, and ultimately higher work efficiency. The efficiency benefits from centralization has been demonstrated by the human genome project with its large dedicated sequencing centers replacing the previous small projects done in individual labs. The potential impact and evolution of widespread biological technology can take inspiration from the computer industry. The computer revolution, providing fast and readily available computers to everyone, has spawned an entire generation of hackers. Some use their skills for the benefit of the community, contributing open source software. Others write viruses, causing much monetary damage. But if computers were still the size of large rooms, most existing computer innovations would not exist. Microsoft and Apple happened because the technology was available for tinkering by ordinary people.

Our society may not be ready yet for the technology, as witnessed by the recent paranoid persecution of Steve Kurtz, an artist found with harmless and easily obtainable biological samples in his home. Improved biology education is needed to alleviate the fear of an area unknown to most. With education and gradual societal acceptance, an explosion of knowledge can be expected. Bio-hackers skilled at manipulating biological systems will emerge from the technology, and we have to ponder the implications before it happens. Poliovirus has been synthesized from scratch (Cello et al. 2002) and the ability to make larger genomes is approaching. Computer viruses cost money; biological viruses can cost lives. However, avoiding education cannot be the answer for controlling biological dangers. The current proposal of centralized biological labs would be safer than the current situation, making it easier to monitor and control biological activities.

Predicting future innovations has been notoriously difficult even for those who understand the technology the best (Wulf 1997). In 1943, IBM chairman Thomas Watson predicted a world market for maybe five computers. Having common online biology labs is reminiscent of the time-sharing mainframe model in computer history. But I predict there will be a world market for many more than five remotely accessible biology labs. Perhaps we will move towards having a personal bio-machine in each home, just as personal computers have become ubiquitous. An opportunity exists now to bring biology education to Alice and Bob and everyone else.

Works Cited

Ball, Philip. "Synthetic biology: Starting from scratch." *Nature* 431 (October 2004): 624–626.

Cello, Jeronimo, Aniko V. Paul, and Eckard Wimmer. "Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template." *Science* 297 (August 2002): 1016–1018.

Che, Austin J. "MITES Biology.". `<http://mitesbio.mit.edu/>`.

Ertugrul, Nesimi. "Towards Virtual Laboratories: A Survey of LabView Based Teaching Learning Tools and Future Trends." *International Journal of Engineering Education* 16 (2000): 171–180.

Erwin, Ben, Martha Cyr, and Chris Rogers. "LEGO Engineer and ROBOLAB: Teaching Engineering with LabVIEW from Kindergarten to Graduate School."

International Journal of Engineering Education 16 (2000): 181–192.

Ferber, Dan. “Microbes Made to Order.” *Science* 303 (January 2004): 158–161.

Harward, V. Judson, Jesús A. del Alamo, Vijay S. Choudhary, Kimberley deLong, James L. Hardison, Steven R. Lerman, Jedidiah Northridge, Daniel Talavera, Charuleka Varadharajan, Shaomin Wang, Karim Yehia, and David Zych. “iLab: A Scalable Architecture for Sharing Online Experiments.” *International Conference on Engineering Education*. 2004.

Henry, Jim. “Web Operation of Real, Physical, Engineering Laboratory Experiments – World Wide Resource Center.” *ASEE* (1999).

iLab. “iCampus Projects: iLab: Remote Online Laboratories.”. `<http://icampus.mit.edu/projects/iLab.shtml>`.

King, Ross D., Kenneth E. Whelan, Ffion M. Jones, Philip G. K. Reiser, Christopher H. Bryant, Stephen H. Muggleton, Douglas B. Kell, and Stephen G. Oliver. “Functional genomic hypothesis generation and experimentation by a robot scientist.” *Nature* 427 (January 2004): 247–252.

Krause, Carolyn. “From a Distance: Remote Operation of Research Equipment.” *Oak Ridge National Laboratory Review* 30 (1997).

Minority Introduction to Engineering, Entrepreneurship and Science. “MITE^{2}S Program.”. `<http://web.mit.edu/mites/www/>`.

Morton, Oliver. “A Machine With a Mind of Its Own.” *Wired* (August 2004): 100–103.

National Instruments. “Distance Learning.”. `<http://www.ni.com/academic/distance_learning.htm>`.

Shen, Hong, Zheng Xu, B. Dalager, V. Kristiansen, Ø Strom, Michael S. Shur, Tor A. Fjeldly, Jian-Qiang Lü, and T. Ytterdal. “Conducting laboratory experiments over the Internet.” *IEEE Transactions on Educations* 42 (August 1999): 180–185.

Wulf, Wm. A. “Look in the Spaces for Tomorrow’s Innovations: the science of future technology.” *Communications of the ACM* 40 (1997): 109–111.