PROJECT DESCRIPTION

Introduction

Over the past two decades, astounding technological advances have transformed the discipline of biology. Sequencing the human genome initially cost many millions of dollars and took years to accomplish, however we are now rapidly moving towards a reality of being able to sequence a human genome for a thousand dollars in a few days. Advances in imaging technologies, such as fluorescence microscopy, have revealed and made visible truly amazing events occurring in living cells and organisms. Using recently developed high-throughput techniques, biological researchers now routinely map the locations where proteins bind to DNA genome-wide and illuminate complex signaling networks by perturbing gene expression through siRNA and chemical screens. In listing here just a few of the many exciting examples of recent work in molecular and cell biology, genetics and genomics, proteomics, and other biology subdisciplines, we highlight advances that have potential to capture the imaginations of students who aspire to become scientists.

In this digital age, the fruits of such cutting-edge research are more widely available to students than ever before. Genome browsers, libraries of microscopy images and time-lapse videos, databanks of protein structures and gene expression data, as well as the archives of primary literature that detail the research which produced these data are all widely accessible to students via the internet. This wealth of available resources produced from actual research has the potential to bring the exciting findings of modern research into educational settings, yet educational innovation has not kept pace with scientific discovery. Most current biology curricula still rely heavily on traditional lectures, textbooks, and pencil-and-paper-based homework that fail to capture the experimental processes of conducting research (Knight and Wood, 2005). Across disciplines, numerous studies on student learning have documented the value of active learning for developing deep conceptual understanding (Bransford et al., 1999; Hake, 1998; Kuh et al., 2005; Pascarella and Terenzini, 2005), yet many educators remain apprehensive about incorporating interactive learning methods into biology curricula (Fairweather, 2008; Knight and Wood, 2005). Thus, their students may not experience the benefits of designing and conducting experiments, the interactive problem solving involved in interpreting data and determining next experimental steps, or the joy of discovery gained through these inquiries.

In an effort to overcome this barrier to the emergence of inquiry-based learning in the classroom, we are striving to enrich science education by building freely available educational software and curricular materials that: (i) provide meaningful opportunities for students to design and perform their own experiments through computer-generated simulations, (ii) expose students to the bounty of remarkable research data that is available to them in a manner that empowers their understanding, (iii) can be easily implemented in existing course curricula to enhance student learning, both in class and during independent studies, and (iv) are freely and openly accessible to curious-minded people worldwide. We have already developed and are actively disseminating two such tools, StarBiochem and StarGenetics, described below. With the current proposal, we seek to apply the experience and knowledge gained from these previous initiatives in the design, implementation, evaluation, and dissemination of a new interactive, inquiry-based, virtual experiment simulator: StarCellBio.

We propose to develop this new simulator, StarCellBio, to be a tool designed for use in undergraduate cell biology education that will be freely available online to students, educators, and other interested parties worldwide. To enrich teaching and learning of cell and molecular biology topics, we envision that StarCellBio will bring together the following important design features: (i) visual demonstrations based on actual data derived from cutting-edge research, (ii) interactive inquiry-based experiment simulations, and (iii) significant capacity for instructor-determined customization of StarCellBio simulations.

The genesis of STAR Biology software tools

To provide context for the proposed work, we begin by sharing background information about our working group and our previous educational projects. In 2004, we assembled a group of graduate students, postdoctoral fellows, and faculty engaged in teaching and curriculum design within the MIT Biology Department's Education Group, supported by Dr. Graham C. Walker's HHMI Professor Grant. An early goal of this group of educators was to identify innovative and interactive visualization methods to enrich

students' understanding of the relationship between protein structure and function. To move beyond static two-dimensional renderings and cartoon models (Harris et al., 2009), the group considered using one of the several existing protein visualization tools designed for educational purposes. We found that available tools fell short of meeting our educational targets for various reasons. Some are mere tutorials that do not enable students to manipulate and interact with the protein structure, a feature that is required for visualization tools to be maximally effective for student learning (White et al., 2002). Others could not be practically implemented because of compatibility restrictions regarding computer platforms. The existing more powerful interactive protein visualization tools, designed primarily for research rather than for educational use, required significant technical expertise and substantial training, thus the time required for students to gain facility in using these tools would have diverted them too much from actual biology subject content (Booth et al., 2005). Another roadblock to broad implementation was that many of these existing tools are not freely and easily available, but require a paid-subscription service, a cumbersome registration process, or software installation. Since we were unable to identify a visualization tool that met our needs, we decided to build one that integrated the best components of the existing tools and added innovations of our own.

We initiated a collaboration with Charles Shubert of MIT's Academic Computing Group and MIT Physics Professor John Belcher to design a student-focused, platform-independent, freely available interactive 3-D protein explorer. Shubert and colleagues, later to become the Software Tools for Academics and Researchers (STAR) group within MIT's Office of Educational Innovation and Technology (OEIT), contributed the substantial technical and programming expertise required to realize this project. In 2006, the work culminated in the creation of StarBiochem (Figure 1; web.mit.edu/star/biochem/), an educational tool developed with the specific and primary intent of bridging the gap between existing educational and research-level protein viewers. StarBiochem is now used in introductory biology courses at MIT and Brandeis University (Waltham, MA), high school outreach programs at the Broad Institute (Cambridge, MA) and the MIT Museum, the Boston Public School system, and K-12 teacher training programs in Massachusetts and New Hampshire (Ceraj et al., 2010). The log of visits to StarBiochem's website indicates that the tool has a growing presence outside of MIT. Since 2007, there have been 34,288 visits to the StarBiochem site originating from 1809 cities and 110 countries.

Developing a virtual genetics experiment simulator: StarGenetics

StarBiochem's success generated faculty interest in the development of an interactive inquiry-based tool for genetics education. While StarBiochem is essentially a tool for interactive demonstration of concepts related to protein structure, our goals for the genetics educational tool went further. The study of genetics lends itself readily to laboratory-based learning. The laboratory component of a genetics course offers students opportunities to design, perform, and analyze data collected from their own genetic experiments. Yet, limitations in space, time, and other critical resources pose obstacles to the incorporation of these important experimental exercises in many genetics courses. Many undergraduate institutions, including MIT, do not have an experimental laboratory component that accompanies their introductory genetics courses. Instead, genetics concepts taught in lectures are reinforced, not by experiments that offer students a glimpse into the thought processes that underlie real genetic research, but by text-based genetic problem sets given as homework. While such traditional text-based genetics problems can and do support learning and the development of critical problem-solving skills, the problems themselves usually provide: (i) the rationale for and type of experiment being performed, (ii) the organisms for initial genetic crosses, and (iii) a summary of experimental results that students are asked to interpret. Such exercises do not really recapitulate the logic and processes of the design of actual genetics experiments, in which data acquired from a series of crosses is used to generate additional hypotheses and to guide further experimentation.

Computer simulations are more effective at incorporating inquiry-based learning into course curriculum than traditional methods because they allow students to perform experiments where variables can be easily manipulated and the effects on the simulated phenomenon can be observed (de Jong, 2006; De Jong and Van Joolingen, 1998). To address the limitations of traditional methods for teaching genetics and to promote inquiry-based learning (de Jong, 2006), the MIT OEIT-STAR team collaborated with Professor Chris A. Kaiser of the MIT Biology Department to design and build the interactive educational software tool StarGenetics (Figure 2; web.mit.edu/star/genetics/). StarGenetics enables students to

perform their own genetics experiments through computer-generated simulations. This tool can support all stages of inquiry-based learning: (i) the identification of variables (e.g., organisms for matings) and their relations, (ii) hypothesis generation, (iii) experimentation, (iv) drawing conclusions based upon experimental results, including statistical tests of significance, (v) reflection upon this process, towards the end goal of knowledge acquisition, and (vi) planning and monitoring of the inquiry process (de Jong, 2006). As is also the case with StarBiochem, StarGenetics and the supporting curricular materials that we have designed to accompany this tool are freely available on the internet. Thus, StarGenetics provides an educationally meaningful, financially affordable, time-efficient, and easily implementable substitute for a laboratory component and can enrich student experience in any genetics course.

StarGenetics is quite unique among the existing genetic cross simulators in that it is highly customizable by the instructor. By modifying a simple Excel workbook available on the StarGenetics website, an instructor can generate new exercises with organisms for matings that have custom genotype-phenotype correspondences. Instructors can define sex-linked traits, chromosomal locations and gene linkage, pleiotropy, allele series, and complex gene interactions such as epistasis. Thus, without significant time commitment, faculty can translate their favorite text-based problem sets into interactive genetics exercises. StarGenetics is also unique in that it supports not only fruit fly crosses, but also yeast genetics and related experiments, including mating, sporulation and tetrad analysis, replica plating, and a wide variety of conditions in which phenotypes can be tested (Figure 3). We are now in the process of expanding the educational power of StarGenetics by adding a bacterial genetics experiment simulator, which moves beyond meiosis-based genetics.

StarGenetics was first implemented by Professor Kaiser in MIT's Genetics course (MIT Course 7.03) in Spring 2009. This virtual experiment simulator tool has also been used in courses at Suffolk University (Boston, MA), Tufts University (Medford, MA), Howard University (Washington, D.C.), and in Boston Public Schools. The tool has been used extensively to enrich homework assignments, to provide simulated laboratory activities during class time, and to provide visual demonstrations of genetics concepts during lectures.

We have begun efforts to assess the efficacy of StarGenetics in supporting student learning at some locations of implementation. During Fall semester 2009, we conducted a formal study of StarGenetics at Suffolk University in collaboration with Professor Stacie L. Bumgarner, who incorporated StarGenetics into the curriculum of her undergraduate Genetics course (S.U. Biology Course 274) that semester, and with the assistance of MIT's Teaching and Learning Lab (TLL). This study employed a mixed-method approach to assess the effectiveness of StarGenetics as a tool for teaching genetics concepts. We are currently in the process of analyzing the data collected from this study. Preliminary results from one of the study's quantitative tools, a student-response survey administered at the end of the semester, shows that students' overall experience using StarGenetics was positive (25% reported an "excellent" and 50% reported a "good" overall experience with the software). This survey also provides an early indication that StarGenetics may have met the educational goals that it is intended to address. Students reported that the software enabled them to learn how to design their own experiments (4.43 $\pm$ 0.53 rating using a 5-point Likert scale), to analyze genetic data (4.63 $\pm$ 0.52), and to interpret the results of their own experiments (4.57 $\pm$ 0.53). Furthermore, students judged that StarGenetics had helped them to understand key genetic concepts taught in class (4.38 $\pm$ 0.75) and had offered an effective learning experience (4.43 $\pm$ 0.79). Our analysis of the data collected in this study is ongoing, and we plan to share our results by submitting a report to a peer-reviewed journal. From solicitations for student feedback at other locations where StarGenetics has been implemented, we have received similarly positive reports of student experience using this software tool. Student perceptions of StarGenetics, when fully incorporated into a guest lecture delivered by Professor Bumgarner, at Howard University during Fall 2009 were also overwhelmingly positive. The majority of students (n = 217) reported that in-class demonstrations using StarGenetics enhanced their understanding of the genetic concepts presented in that lecture (6.22 $\pm$ 1.27 rating using a 7-point Likert scale). One student commented that "the software did help with understanding genetics because it showed me more possibilities than the punnett square." Another student observed, "I had difficulty understanding genetics yet viewing the program really helped me get a better understanding."

We are encouraged to learn that there is growing interest in implementation and assessment of genetic experiment computer simulations in undergraduate education (Sved, 2010 and associated response from GSA representatives), and we look forward to working with the Genetics Society of America (GSA) and members of the science education community in continued development and evaluation of StarGenetics. This attention is timely according to our StarGenetics usage statistics. Following its public release on our website in 2009, StarGenetics usage has grown rapidly: 9924 website visits in 1138 cities and 87 countries worldwide. Our Google Analytics data indicate that we have visitors from universities all over the world.

Broader impacts of STAR educational tools and outreach efforts

A particularly important strength of StarBiochem, StarGenetics, and other STAR software is that these tools provide opportunity for teachers and students in resource-limited settings to explore cutting-edge science without needing to overcome potentially insurmountable hurdles, such as obtaining laboratory space and expensive reagents. Any student with a computer and access to the Internet can utilize the full capabilities of these free software tools in their studies. Yet, in order for students and instructors to take advantage of STAR tools, they must first be aware of their existence. Research has shown that, without active outreach, many education initiatives die at the site of local innovation (Fairweather, 2008). Thus, we consider dissemination and outreach efforts to be an intrinsic component of our initiative, critical to expanding the breadth of the impact of this work.

Our dissemination and outreach efforts are led by laboratory-trained Ph.D. research scientists. For such positions, we have been able to recruit high caliber individuals who trained with leading scientists at MIT and who have specific expertise in the research areas that our educational tools address. For instance, the two research scientists spearheading outreach efforts for StarBiochem and StarGenetics are Lourdes Alemán, who completed her Ph.D. with Professor Phillip A. Sharp, and Stacie Bumgarner, who completed her Ph.D. with Professor Gerald R. Fink and has significant experience with teaching undergraduate genetics courses. This approach of dedicating highly trained research scientists to the development and dissemination of STAR educational tools is one that is recommended by an MIT-commissioned iCampus report issued in 2007 by the non-profit organization Teaching, Learning, and Technology (TLT) Group (Ehrmann et al., 2007). This TLT report suggests that sustained educational innovation is critically supported by the work of such highly trained individuals, who can: (i) help faculty members learn about educational development taking place elsewhere, (ii) understand faculty members' ideas and help them to develop their own educational innovations and adaptations, (iii) act as communication links between faculty and technical staff, such as software engineers, and (iv) help faculty assess learning outcomes and disseminate their work to colleagues at MIT and elsewhere. Our STAR research scientists interact closely with MIT Biology faculty, and this interaction has led to encouraging growth in faculty involvement in development, implementation, and assessment of STAR educational initiatives, including participation by Professors Graham C. Walker, Chris A. Kaiser, Gerald R. Fink, Terry L. Orr-Weaver, and Iain Cheeseman.

Our ongoing outreach work was recently recognized with a Davis Educational Foundation grant, which has helped support a multifaceted effort to disseminate StarBiochem and StarGenetics during the past two years. Our STAR research scientists have presented our science education initiatives at national conferences, have initiated collaborations (locally, nationally, and internationally) with universities, colleges, and outreach programs, and have held training workshop at universities to help faculty, instructors, and graduate student teaching fellows become proficient using STAR software tools. We have recently begun using social media (www.epernicus.com/) to support an online community of instructors who use StarBiochem and StarGenetics in their teaching, an approach that appears promising for keeping users connected after participating in our workshops and other outreach events. We are also engaged in ongoing outreach initiatives to support higher education and E-learning in Haiti, Brazil, India, Spain, Puerto Rico, Tanzania, and Ghana. In cases where internet connectivity presents a significant barrier to implementation, we have provided free stand-alone versions of our software tools and curriculum. We are exploring the feasibility of translating our software tools and online curriculum into other languages, especially Spanish, Portuguese, and Haitian Creole.

Beyond work at the university level, our outreach team is collaborating with nonprofit J.F.Y. Networks and teachers in the Massachusetts Public Schools to make our STAR Biology tools more accessible to younger students. To this end, curriculum modules and materials that are appropriate for high school STEM classes will be generated. Our outreach work through the MIT's Center for Environmental Health Sciences (CEHS) Outreach Program and the MIT Museum is significant, involving an integrated permanent interactive StarBiochem-based exhibit and specialized StarBiochem-based interactive lessons designed for high school class field trips to the museum (Ceraj et al., 2010). Our museum collaborators, led by MIT's CEHS Core Director Kathleen Vandiver, recently received funding that will support the export of this interactive lesson to high school classrooms throughout New England.

In addition to outreach efforts with focused target audiences, we also utilize broad advertisement approaches to disseminate information about STAR educational tools. In 2010, these efforts included Google AdWords links and direct mailings to Biology faculty at universities nationwide, which may have contributed to the significant increase in StarBiochem and StarGenetics usage that we have since observed.

Rationale for developing a cell biology experiment simulator: StarCellBio

In summary, StarBiochem is a tool for interactive demonstration of concepts related to protein structure while StarGenetics is a tool for interactive demonstration of concepts related to genetics. The meiosis-simulating "mating engine" at the core of StarGenetics allows students to conduct virtual experiments utilizing instructor-defined "inputs" (i.e., organisms) controlled by a set of algorithms that generate realistic virtual data sets. Prompted by the promising response to StarBiochem and StarGenetics from students and instructors, as well as the ongoing productive collaboration between MIT Biology faculty and the STAR software engineers and laboratory-trained biologists at OEIT, we are encouraged to build upon this basic design to create software tools for use in other areas of biology education.

Introductory Cell Biology and Molecular Biology courses are often conducted without an experimental laboratory component. Often due to limitations in expensive laboratory resources and time, instructors tend to rely heavily on text-based problems that provide descriptions of methods and data as tools to promote student learning. However, such problems fail to recapitulate the logic and processes of conducting actual experiments in these fields. We intend to bring together the best ideas of our faculty, research scientists, and students to develop, implement, and evaluate an interactive educational software tool that will enrich Cell and Molecular Biology curricula with an inquiry-based learning component (de Jong, 2006). Our goal is to build a tool, which we plan to call StarCellBio, designed for use in undergraduate education and freely available online to students, educators, and other interested parties worldwide. We envision that StarCellBio will bring together the important design features of StarBiochem and StarGenetics, including: (i) visual demonstrations based on actual data derived from cutting-edge research, (ii) interactive experimental simulations, and (iii) considerable capacity for instructor-determined customization of exercises.

Our discussions with professors who teach Cell Biology at MIT (MIT Course 7.06), Terry Orr-Weaver, Iain Cheeseman, Angelika Amon, and Harvey Lodish, have helped focus our plans for the development of StarCellBio. A recurring request from faculty is that the tool be able to utilize real cell biology microscopy data in its interactive visualizations and experiment simulations. The immense collection of available image and movie data includes beautiful examples that demonstrate key topics of Cell Biology, including intracellular localization and colocalization of proteins, protein trafficking, cell signaling pathways, the cell cycle, cell motility, cell diversity, and tissue organization. We believe that students' understanding of the experimental basis of these fields will be enriched by incorporating such striking visual data into experimental simulations that can be used to demonstrate concepts in the classroom and, perhaps more importantly, that students can interact with outside of the classroom in their own time.

This project will, in many ways, be more challenging than either of our previous STAR Biology projects. To simulate experiments with capacity to cover the breadth of material covered in cell and molecular biology curricula, the StarCellBio tool will require a set of algorithms that goes beyond the "mating-engine" of StarGenetics.. Systematic application of certain experimental tools can be used to answer diverse biological questions, a process that lends itself well to computer-based simulation and that will take

advantage of the many methodologies and experimental tools now available to cell and molecular biologists. We envision a tool that will combine actual image and time-lapse video data (selected by the instructor from an available library or from the instructor's own data collection) with computer-generated data, such as gel-based (e.g., Western and Southern blots) and fluorescence-activated cell sorting (FACS) analyses, that are determined by instructor "inputs" and realistic "outputs" based on probabilities and calculations controlled by computer algorithms. Given the expertise of MIT's leading Cell Biology faculty and the technical skills of the OEIT staff, our team is uniquely positioned to address this type of pedagogical challenge.

We wish to highlight the fact that the simulation of complex biological systems for educational purposes is not without precedent. The U.S. Medical Licensing Examination (USMLE), a series of exams required for medical licensure in the United States, has recently and successfully incorporated 'case-based simulation' as a major portion of the exam (Dillon et al., 2004; Dillon and Clauser, 2009). The simulations are based on clinical scenarios with complex and broad arrays of possible student-determined inputs (e.g., diagnostic tests and treatments) and algorithm-determined outputs (e.g., resulting clinical improvement or decline, symptomatic change). This interplay between inputs and realistic outputs is precisely the type of scenario that we envision developing for StarCellBio.

Similar to StarGenetics, the StarCellBio tool will support inquiry-based learning, allowing students to consider multiple experimental approaches to answer cell biology questions, conduct experiments, draw conclusions based upon experimental results, conduct further experiments, and reflect upon this process to acquire knowledge. We think that this tool will enrich student learning, improving their problem-solving skills and strengthening their understanding of the principles of experimental biological science. In the context of trying to attract students to scientific fields of study, we hope that offering a tool that models the experience of conducting experimental work will increase students' interest in pursuing careers in science research.

Development of StarCellBio

This project will consist of three major stages: (i) development, (ii) classroom implementation, and (iii) broad dissemination and outreach. In the first stage of the project, our team of faculty, research scientists, and software engineers will build a simulation engine that allows students to perform realistic virtual cell biology experiments. To ensure that we are building a tool that will have utility to the broader community, we will also elicit feedback from colleagues at other institutions who have offered to comment on the fruits of the project at various stages of the development process.

As is the case with StarGenetics, each StarCellBio experiment simulation will require that two distinct types of variables be defined by the instructor. We anticipate that these definitions will be recorded in a modifiable Excel workbook which, together with specified image and/or movie files, will load a custom experiment simulation using the StarCellBio tool. The first type of variables consists of the set of "inputs" needed to set up the exercise: (i) specification of the organism(s) and/or cell type(s) used in the simulation; (ii) specification of available conditions and associated characteristics of the organism(s) under those conditions (e.g., doubling time in various growth conditions, etc.), (iii) specification and definition of wild type and genetic mutants, if applicable, and (iv) specification and definition of experimental treatments available to the student (e.g., specific drugs, siRNAs, etc.). The second type of variables concerns the set of rules that define how and to what extent the model organism's phenotype will change after an experiment is performed. These rules represent the relationships between experimental treatments and experimental outcomes for all possible organisms within the simulation.

The simulation engine that will drive StarCellBio experiment simulations will present two categories of data outputs: (i) real microscopy image and time-lapse movie data, and (ii) computer-generated data, such as gel-based (e.g., Western and Southern blots) and flow cytometry-based data. At the outset of an experiment, background data from one or both of these two categories may be presented, as determined by the instructor who designs the simulation. Students will generate additional data based on their specific choices from among the experimental treatments available. The treatments offered in each simulation will be defined by the instructor, as described above.

Annotated image and movie files for use in simulations will be made accessible to instructors in a visual media library that we will compile and share online. Insofar as possible, this library will be constructed from collections of existing publically available data. We are cognizant of potential issues with copyrighting and will ensure that appropriate permissions are obtained for all material included in the library. The library will be organized and searchable by model organism (e.g., human, mouse, *Drosophila*, yeast), cell type, key concepts, and experimental methodology employed. It will be possible for library files to be called using more than one descriptor (e.g., 'mitotic spindle', 'cell cycle arrest', 'mitosis', 'chromosome segregation mutant', etc.), so that appropriate images can be multi-purposed for various experimental scenarios. In addition to the visual media library, the software will also allow instructors to utilize their own image and movie files in building StarCellBio exercises, thus providing instructors with opportunities to bring their own research into their courses.

The visual image library will contain high-resolution images of individual cells, as well as images containing fields of cells. The field views will potentiate realistic demonstration of the inherent variability observed in experimental results. For example, a one-time treatment of mammalian cells with leptomycin B, a drug that inhibits nuclear export, might affect sixty percent of treated cells, whereas a second treatment using the same experimental parameters might result in ninety percent of cells blocked for nuclear export. As is the case in an actual laboratory setting, in order to draw meaningful conclusions, students will need to repeat any experiment several times, quantify results, and determine statistical averages and deviations. From a pedagogical perspective, this experience should surpass any counterpart text-based cell biology problem by its real-time demonstration of the inherent variability associated with experimental techniques and biological systems. This is an experience that students generally access only in an actual laboratory environment.

The second category of data output in StarCellBio simulations will consist of computer-generated data. We envision that these types of data will include illustrated simulations of gels and blots, as well as simulations of analyses such as flow cytometry, which provide summary information across large populations of cells. Students will interpret these results and perform additional experiments. For instance, they will be able to select specific subpopulations of cells presented in simulated flow cytometry profiles for use in further experimental inquiry.

As with StarGenetics, the experimental outcomes generated by the StarCellBio simulation engine will be based upon a statistical model of a sample group of cells representing the natural distribution of outcomes occurring in a much larger pool of cells. For any given treatment, StarCellBio will return a realistic array of phenotypes based on the probabilities of obtaining the particular phenotypes under definable experimental variables, such as dose-response curves and/or time-dependent progressions.

Initially, we will concentrate on developing functionality that will support topics and concepts covered in the Introductory Biology (MIT 7.01 series) and Cell Biology (MIT Course 7.06) courses at MIT. Throughout development, we will collaborate with MIT Introductory Biology faculty, Professor Hazel Sive, and MIT Cell Biology faculty, Professors Terry Orr-Weaver, Iain Cheeseman, Angelika Amon, and Harvey Lodish. In view of building a tool with broad utility beyond MIT, we will also communicate with educators at other institutions who have offered to provide input and feedback during the StarCellBio development process, including Harvard University Cell and Molecular Biology preceptors, Dr. Tara Mann and Dr. Mary Ellen Wiltrout, respectively, and Dr. Leticia Vega of Barry University, a small minority-serving institution.

Below, we include four examples of cell biology topics, accompanied by descriptions of how a problem about each topic might be converted into a StarCellBio experiment simulation.

***Example 1*: Dissecting a mammalian signaling pathway.** Developing student understanding of intracellular signaling pathways is part of the core curriculum in cell biology courses. StarCellBio could enrich student learning in this domain via an exercise designed to enable students to test the functions of genes involved in the mammalian Ras pathway. For such a simulation, the students might have access to a panel of experimental treatments, including various siRNA sequences targeting genes in the Ras pathway, overexpression constructs, and drugs affecting the activity of proteins in the pathway. Each of these experimental treatments, alone or in combination, would have an instructor-defined dose-response curve to specify the experimental output. For their investigations, students would have access to a

variety of virtual experimental tools (e.g., SDS-page gels/western blots, flow cytometry, fluorescence microscopy, etc.).

Example 2: **Understanding mechanisms of vesicle trafficking.** Another core component of cell biology curriculum is the study of protein trafficking by vesicles in the cell. Investigations of the mechanisms underlying vesicle trafficking often involve visual assays, which lend themselves well to StarCellBio simulations. Students might be presented with time-lapse movie files in which fluorescent probes label two different types of vesicles. Students would have available to them a virtual selection of drugs (e.g., microtubule-inhibiting drugs, such as nocadazole; drugs that disrupt actin microfilament organization, such as Latrunculin A; etc.) and other factors (e.g., non-hydrolysable ATP and GTP analogues; constructs that overexpress microtubule polymerization promoting factor XMAP215 or katanin, etc.) with which they can treat the cells containing labeled vesicles to observe the corresponding effects on vesicle trafficking. The StarCellBio tool would return appropriate image or time-lapse movie files in response to the various treatments, thereby enabling students to characterize the two types of vesicles. Through their investigations, students could make discoveries including (i) whether the actin or microtubule cytoskeleton plays a role in the movement of these vesicles, (ii) how the polarity of microtubules can affect vesicle movement, and (iii) how these vesicles can be subclassified based on the motors that drive their movement.

Example 3: **Investigating the mitotic cell cycle.** The cell cycle is a major topic of all cell biology courses. The StarCellBio tool could enrich studies of the cell cycle in various ways. For example, students might be offered wild type yeast and a selection of temperature-sensitive (t.s.) cell cycle yeast mutants. The initial data presentation might be a flow cytometry profile of cells that were incubated at the permissive temperature and then shifted to the restrictive temperature for at least one cell cycle before being treated with a DNA-staining compound, such as DAPI, used to quantitate DNA content. Using this initial flow cytometry profile, students could make observations about the characteristics of the wild type cell population (e.g., more than half of the cells might be in the G1 phase of the cell cycle with their DNA unreplicated while the remaining cells might be inG2/M phase with a fully replicated DNA and therefore with double the amount of DNA content or in S phase, actively replicating their DNA and therefore with an intermediate DNA content) and begin to characterize the mutant t.s. phenotypes (e.g., Mutant 1 cells are delayed in S phase; Mutant 2 cells are G2/M-arrested, etc.). Students could then have the option to (i) observe these strains using microscopy images for further characterization of phenotypes, (ii) treat cells with drugs with known cellular targets to try to recapitulate the mutant phenotypes towards understanding mechanisms that govern the cell cycle, (iii) use various gene overexpression constructs for rescue experiments, and (iv) use various antibodies from an antibody library to test for changes in phosphorylation states of specific proteins when cells are grown at permissive versus restrictive conditions using gel-based assays.

Example 4: **Understanding cell motility.** Observing mammalian cells in motion is an amazing experience for young and experienced scientists, alike. StarCellBio can enrich the study of cell motility by offering students increased opportunity to interact with informative (and often beautiful) time-lapse microscopy data. With StarCellBio, students might perform their own experiments to identify the targets of compounds that impair cell motility of neutrophils, for example. The initial data presented in the simulation might be a time-lapse movie of actively moving wild type, untreated neutrophils. The interactive component of the simulation might involve students treating the neutrophils with selections from a panel of compounds to see the effects of these compounds, alone or in combination, on cell motility. In response to these treatments, StarCellBio might present students with a combination of appropriate time-lapse movies that enable them to observe resulting phenotypes at high magnification and across fields of cells, the latter for estimations of dose-response relationships. Further investigations using gel-based assays could probe changes in signaling pathway function or in gene expression as a result of treatment with the compounds.

With these four examples, we show how the StarCellBio tool will enable students to: (i) have unprecedented opportunities to interact with visual cell biology data while completing problem sets and studying, (ii) determine the types of experiments needed to answer particular cell biology questions, (iii)

perform those experiments, (iv) analyze the data generated, (v) perform additional experiments, if appropriate, and (vi) draw conclusions based on their own inquiries.

As with StarBiochem and StarGenetics, when a beta-version of the StarCellBio software tool has been developed, we will perform tests (discussed in the evaluation plan below) to assess the tool's ability to simulate realistic cell biology experiments and to provide an intuitive user experience through its graphical user interface. These tests will be conducted in conjunction with MIT Information Services and Technology (IS&T)'s Usability Lab and volunteers from the biology community. Thereafter, we will continue to incorporate additional functionality by building upon the possible experimental treatments, techniques, and model systems, and by expanding the range of concept areas that can be addressed by the tool.

Implementation of StarCellBio in formal educational settings

Our next step will be to collaborate with MIT faculty, undergraduate students, and graduate student teaching fellows to begin incorporating the StarCellBio tool into their course curriculum. Prior to this implementation, we will have established baseline indicators of educational impacts related to MIT's Cell Biology course (MIT Course 7.06) so that we can make comparisons about student learning and engagement before and after implementation of StarCellBio (discussed in the evaluation plan). Using an analogous approach to those employed for StarBiochem and StarGenetics implementation, we will work with Cell Biology faculty Iain Cheeseman and Terry Orr-Weaver to develop a small series of StarCellBio 'curriculum modules' consisting of homework exercises that are appropriate for and easily incorporated into MIT Course 7.06. Implementation of this initial set of curriculum modules will be evaluated using direct and indirect measurements of student learning and responses to the tool (discussed in the evaluation plan). Subsequent steps for StarCellBio software development and course implementation will be revised based upon the results of these early implementation experiments.

After appraising the impacts of local implementation and making any necessary modifications to StarCellBio, we will begin expanding the breadth of our implementation efforts to other institutions. Our experience with StarBiochem and StarGenetics dissemination suggests that the development of easy-to-incorporate 'curriculum modules' will be central to successful implementation of StarCellBio beyond MIT. These curriculum modules, developed by faculty and our research scientists, are designed to be used by students in independent or group learning settings. Ultimately, the set of curriculum modules that will be freely available on our website will span different educational levels (high school STEM-level, undergraduate-level, graduate-level) and cover various biological concepts, thereby making implementation of StarCellBio appropriate in a wide variety of educational settings. The goals of these modules are both educational and practical. The modules should provide meaningful opportunities for students to learn about cell and molecular biology content, while also allowing easy incorporation of StarCellBio into existing course curriculum, thereby lowering instructors' activation energy barrier for implementation. The modules also provide examples for instructors to adapt and customize for their own courses, an activity that we invite, encourage, and support.

Broader impacts of StarCellBio: Dissemination and outreach

Our ongoing dissemination and outreach activities for StarBiochem and StarGenetics, described above, provide a clear blueprint for our planned approach regarding StarCellBio. We are committed to making StarCellBio freely accessible on the STAR website (web.mit.edu/star/index.html). Additionally, the following supporting materials will be developed and made freely available on the STAR website: (i) step-by-step instructions (text and video) that describe how to use StarCellBio, (ii) support for instructors regarding how to set up custom experiment simulations using this software tool, (iii) curriculum modules for use with the software (described above), (iv) video stories that describe case studies of StarCellBio usage in various educational settings, to demonstrate different ways of implementing this software tool, and (v) video stories of seminal cell biology research from scientists at MIT and other universities, to provide context for StarCellBio curriculum modules and to support cell biology education, in general.

We will also conduct training and development workshops for faculty, instructors, postdoctoral fellows, and graduate students at MIT and on-site at other institutions. These workshops will provide support regarding implementation and custom-design of StarCellBio exercises and will explore the capabilities of

the software in various educational settings. Such outreach activities have been instrumental in expanding usage of StarBiochem and StarGenetics. We also plan to reach out to our growing social media community (www.epernicus.com/) to disseminate information about StarCellBio and to encourage the exchange of ideas among our users.

Professor Walker's newly established HHMI Education Group will play a pivotal role in the dissemination efforts for StarCellBio. The participation of Education Group members, composed of graduate students, postdoctoral fellows and, in particular, our dedicated Ph.D.-trained research scientists, will greatly increase our ability to pursue the important dissemination efforts previously described, while at the same time strengthening the Education Group members' teaching skills.

Our recent collaborations with professors at Howard, Suffolk, and Brandeis Universities have provided opportunities to explore how STAR biology tools can be used in diverse educational settings. Feedback from these and other educators will allow us to make further improvements and modifications to our software platforms and teaching tools, including StarCellBio, so that these tools can be made maximally useful across a wide range of educational environments. Finally, as with StarBiochem and StarGenetics, our collaborations with the Broad Institute Outreach Program, the Whitehead Institute Teacher-Partner Program, the MIT Museum, MIT's Center for Environmental Health Sciences (CEHS) Outreach Program, and J.F.Y. Networks will help us make StarCellBio accessible to a broad range of high school teachers and students.

Guiding questions for evaluation and assessment of StarCellBio

The three key stages of this proposed StarCellBio project are (i) software development, (ii) implementation in formal educational settings, and (iii) dissemination and outreach. Throughout the project, a mixed-method approach will be employed to evaluate impacts and outcomes of StarCellBio development and deployment on the learning and educational experiences of students, as well as on instructional practices and teaching experiences of biology educators.

The following questions will be used to guide the assessment of StarCellBio:

(1) To what extent does StarCellBio provide a virtual learning environment that:
  (i) faithfully simulates aspects of actual biology experiments, and
  (ii) is intuitive, easy, and fun to use?

(2) To what extent does StarCellBio improve student learning, with specific focus on:
  (i) their understanding of core cell and molecular biological concepts,
  (ii) their ability to design, conduct, and interpret hypothesis-driven experiments,
  (iii) their critical thinking and problem-solving skills?

(3) To what extent does StarCellBio provide a positive educational experience and promote student interest in the biological sciences?

(4) How effective is StarCellBio in contributing to an online biology education community?

(5) How effective are our StarCellBio outreach activities?

Directed by these questions, the design and application of assessment instruments and subsequent data analyses will be a collaborative effort conducted by MIT Biology faculty, research scientists within the Office of Educational Innovation and Technology (OEIT), and researchers from the Teaching and Learning Lab (TLL) at MIT. Formative and summative data will be collected from faculty and students during each stage of the project.

Evaluation of StarCellBio development

As with StarBiochem and StarGenetics, StarCellBio's development will be directly influenced by Biology faculty, whose integral feedback throughout the software development process will guide next steps and directions. This feedback (formative data) will be documented to record the process of designing this interactive educational software tool, thereby creating a road map for the development of future inquiry-based software tools.

To assess StarCellBio's development process, an established survey instrument (SI-1) will be administered to capture the reactions of cell biology instructors to pre-beta versions of the StarCellBio tool (summative data). SI-1 will provide another source of information about the process of creating a simulated cell biology learning environment. In addition, SI-1 will collect suggestions for improvements to the StarCellBio software, toward our goals of generating a tool that faithfully simulates aspects of real biology experiments and is appropriately simple to use (Question 1). With respect to achieving these goals, a high level of satisfaction from participating faculty will be used as one indicator of success during StarCellBio's development.

During the development process, feedback will also be collected using a focus group (FG) instrument. FG participants will consist of MIT undergraduate students who were previously enrolled in MIT's Cell Biology course (MIT Course 7.06) and graduate student teaching fellows currently or formerly involved in Course 7.06 instruction. Similar to the process for collecting feedback from faculty, the responses of FG participants to pre-beta versions of StarCellBio will be collected throughout the development process. FG participants will be asked to evaluate: (i) the qualitative look-and-feel of StarCellBio's graphical user interface (Question 1); (ii) ease of usability of StarCellBio (Question 1); (iii) StarCellBio's ability to simulate meaningful cell biology experiments (Question 1); and (iv) the tool's ability to support learning of core biological concepts and experimental inquiry (Question 2).

The FG instrument will also be a significant aid towards the development of a maximally useful inquiry-based tool, prior to its broad implementation in formal educational settings. Research shows that students have difficulty with various aspects of inquiry-based learning (de Jong, 2006; De Jong and Van Joolingen, 1998). For example, students have trouble identifying hypotheses to test, knowing which variables are most appropriate to modify, making predictions, choosing reasonable experiments to conduct, and linking experimental data with conclusions that either support or refute a given hypothesis (de Jong, 2006). The most effective computer simulations integrate cognitive tools that support students through the inquiry-based learning process. Such cognitive tools come in various forms, including assignments that "set the simulation in the right state", background and explanatory information, experiment tracking tools, "hypothesis scratchpads", hints, and process coordinators that guide students through simulations (de Jong, 2006). Determining which types of cognitive tools are most appropriate to integrate into StarCellBio and associated curriculum modules (described elsewhere in the proposal) will be an important step in the development of this educational tool. StarBiochem and StarGenetics curriculum modules contain background information, hints, and process simulators to support students through the learning process. Currently, we have only indirect data regarding the effectiveness of these cognitive tools for either software. The FG instrument will assist in evaluating our current strategies involving cognitive tools and will help determine if additional types of cognitive tools are needed for StarCellBio.

Following the development of StarCellBio, we will begin formal usability testing of the software with volunteers from local undergraduate, graduate, postdoctoral, and biology instructor communities. As with StarBiochem and StarGenetics, usability testing (UT) will be conducted in collaboration with MIT's IS&T's Usability Lab. Usability testers will not have been involved earlier in StarCellBio's development process, thus the UT instrument will provide opportunity to evaluate the software with users who have no prior experience with the tool. To assess whether StarCellBio is intuitive and fun to use (Question 1), usability testers will be given a set of specified tasks to complete using StarCellBio. Empirical data will be gathered using think-aloud protocols, moderated by trained staff from the IS&T Usability Lab. In addition, testers will complete a survey instrument (SI-2) following the conclusion of the usability test. UT data will be used to refine and modify StarCellBio to generate improved versions of the software. Repetitions of UT will be conducted with new testers following each major revision of the StarCellBio software. A high level of satisfaction from testers will be used as one indicator of success during StarCellBio's development (Question 1).

Evaluation of StarCellBio implementation

While the beta-version of StarCellBio is being developed, we will collect baseline data concerning learning objectives (Question 2). This initial assessment will be conducted and appraised prior to implementation of StarCellBio in MIT's Cell Biology course (MIT Course 7.06). Baseline indicators will be

collected using a range of direct assessment instruments. Rubric instruments will be used (i) for homework exercises that assess comprehension of concepts to be specifically addressed in StarCellBio curriculum modules (RI-1), and (ii) for exam questions that assess our learning objectives (RI-2). Our direct assessment strategies will utilize a biology concept inventory, called the "Introductory Molecular and Cell Biology Assessment" (ICMA), developed at the University of Colorado (Boulder) by the same team of investigators that generated the genetic concept inventory that we have used in evaluating StarGenetics (Shi et al.). This concept inventory will be employed to evaluate students' understanding of biology concepts with assessment instruments administered before, during, and after their use of StarCellBio in coursework associated with MIT Course 7.06. Upon their completion of the course, student volunteers will be interviewed to provide us with more insight into any educational gains and to probe the accuracy of our other assessment instruments. As has been previously observed, we believe that these interviews will be useful in revealing differences in student learning that can sometimes be missed with other assessment instruments, including homework exercises and exams (Harris et al., 2009). Interview questions will be classified according to Bloom's Taxonomy "competence levels" and will directly address StarCellBio's educational goals (Bloom and Krathwohl, 1956). The quality of student responses will be evaluated by an independent panel of assessors (PhD-trained research scientists) and interview footage will be reviewed by StarCellBio's development team (Harris et al., 2009). All assessment instruments, including RI-1, RI-2, the ICMA, and student interviews, will be administered with Course 7.06 students before and after StarCellBio implementation in effort to obtain comparable results for analysis. Any increase in students' mastery of cell biology concepts and growth in experimental thinking (ability to design, conduct, and interpret hypothesis-driven cell biology experiments) and problem-solving skills, as measured by the described assessment instruments, will be taken as indicators that StarCellBio has successfully met its stated educational goals (Question 2).

Finally, survey instrument (SI-3) will be administered to these same students following their completion of Course 7.06. SI-3 will be an indirect assessment instrument that will evaluate the degree to which students feel that StarCellBio is able to promote its stated learning objectives. SI-3 will also query the extent to which StarCellBio affects students' engagement with and enthusiasm towards the study of biology (Question 3). Any increase in student engagement or enthusiasm will be taken as an indicator that StarCellBio is able to provide a positive learning environment and inspire further inquiry (Question 3). A corresponding survey instrument (SI-4) will be administered to MIT's Cell Biology faculty and to graduate student teaching fellows for Course 7.06 to assess the implementation process and locate any needed improvements for future StarCellBio implementation efforts.

Following implementation in Course 7.06, we will collaborate with additional faculty to implement StarCellBio in MIT's Introductory Biology courses (7.01 series). Since Introductory Biology is a required course at MIT, it presents a highly appropriate educational setting in which to query the extent to which StarCellBio affects students' engagement and enthusiasm towards the study of biology and/or motivates students to consider research science careers (Question 3). The student population enrolled in the Introductory Biology courses covers a diverse demographic range, including freshman and sophomores who have not yet decided upon a major and juniors and seniors already committed to other major fields of study. Thus, a version of the SI-3 instrument will be administered and any results that support StarCellBio's ability to generate enthusiasm and motivation within such a varied student population will be taken as indicators of StarCellBio's ability to provide an effective educational environment and to promote student interest in STEM related careers (Question 3).

Evaluation of StarCellBio dissemination and outreach efforts

Our planned dissemination efforts for StarCellBio can be broadly grouped into two categories: (i) online (remote) outreach efforts and (ii) direct outreach activities.

The success of StarCellBio's online (remote) outreach efforts will be assessed through a combination of usage statistics and user experience surveys. We will collect usage statistics that address the number of visits and the time that individual users spend on the StarCellBio website. This information will provide indicators of StarCellBio's presence beyond MIT and its ability to engage remote users (Questions 4 and 5). In addition, analysis of detailed usage statistics (i.e. number of users per given time accessing StarCellBio from a particular college or university) allows us to estimate the number of organized courses

that are using StarCellBio during a given period of time. This type of usage data can be generated and compared across successive semesters to estimate the growth of StarCellBio usage in college courses over time (Question 5). We are, however, cognizant of the limitations of this type of analysis, as it makes assumptions about what constitutes a "class of students" and will underestimate usage by students who access the StarCellBio website from off-campus locations. Therefore, this type of data will not provide a complete picture of StarCellBio student usage, but does have power to offer a useful window into StarCellBio usage on the campuses of universities worldwide. A 1.5-fold increase in total site visits and/or "course" usage (as described above) from year-one to year-two following StarCellBio's online release will be taken as an indicator of success for our online dissemination efforts.

Surveys will also be designed and offered to assess remote users' experience with StarCellBio more directly. These surveys will evaluate StarCellBio's usability, success at addressing pedagogical challenges, and ability to support our stated learning objectives (Questions 1, 2 and 3). Two different online surveys (SI-5 and SI-6) will be offered, designed specifically for students and for faculty users, respectively. SI-5 and SI-6 will be presented via a dialogue box when a user exits the StarCellBio program. SI-5 and SI-6 will incorporate items from SI-3 and SI-4, respectively and will also include questions that assess: (i) prior StarCellBio usage, (ii) reasons for usage, (iii) ability of StarCellBio to address specified educational objectives and challenges, (iv) quality of StarCellBio curriculum modules, and (v) suggestions for additional educational and technical features.

Direct outreach efforts, such as training and development workshops, will be evaluated via a survey that will assess participant satisfaction with the training (SI-7) and also by follow-up communications to determine whether the training has resulted in successful StarCellBio implementation. Social media (www.epernicus.com) will also be used to support a community of educators who use StarCellBio in their courses. Following participation in training and development workshops, we will follow the communication threads among participants to determine how StarCellBio fosters curriculum growth and development of new teaching methodologies in this community of users (Question 4). In order to gauge how StarCellBio meets the pedagogical challenges and learning objectives characteristic of different educational levels, the SI-5 and SI-6 instruments will also be administered during various direct outreach activities that support high school students or K-12 teachers. These data will provide information regarding the impact of StarCellBio on K-12 education (Question 5).

Figures

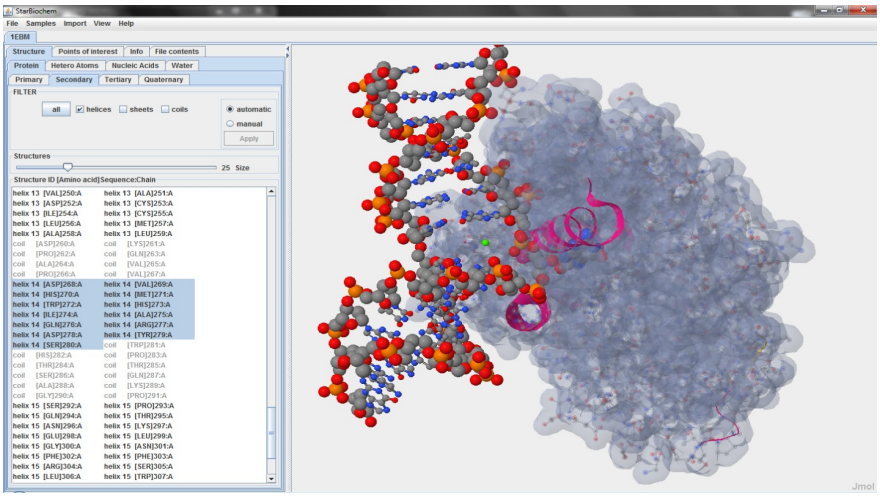

Figure 1. The StarBiochem graphical user interface. StarBiochem is a protein structure explorer designed specifically for student use to enrich their understanding of the biology of proteins through interactive visualization. The tool employs terminology that corresponds to introductory biology textbooks and its interactive controls are entirely encompassed in its graphical interface, so the tool requires no programming knowledge. The software will load any PBD (Protein Data Bank) file. A selection of PDB files that correspond to our curriculum modules are bundled with the software under the Samples menu.

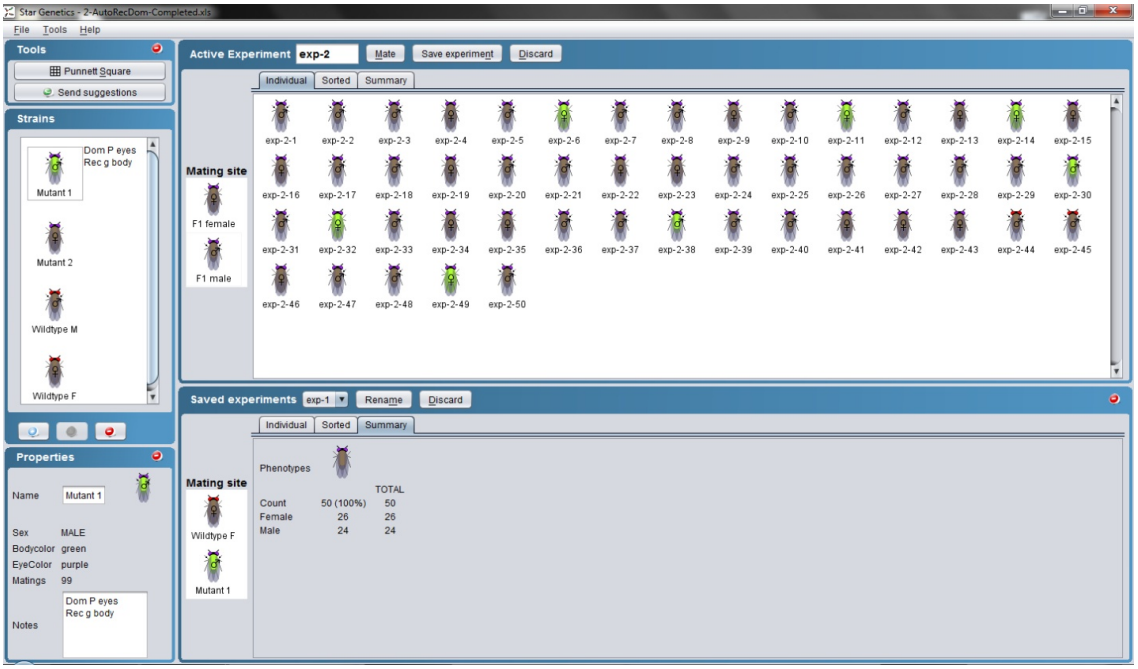

Figure 2. The StarGenetics fly graphical user interface. StarGenetics is a software tool that simulates mating experiments for inquiry-based learning of genetics. Many concepts can be studied with StarGenetics simulations, including sex-linked traits, gene linkage, pleiotropy, allele series, and complex gene interactions such as epistasis. Variation in traits including eye color, body color, wing length, wing vein, antennae length, sterility, and lethality can be simulated in fly experiments. By modifying a simple Excel workbook available on our website, instructors can generate new exercises to load custom experiment simulations. The Excel workbook specifies the organisms, given in the “Strains” window, available for initial crosses. In addition, organisms generated through genetic crosses can be isolated for subsequent mating experiments.

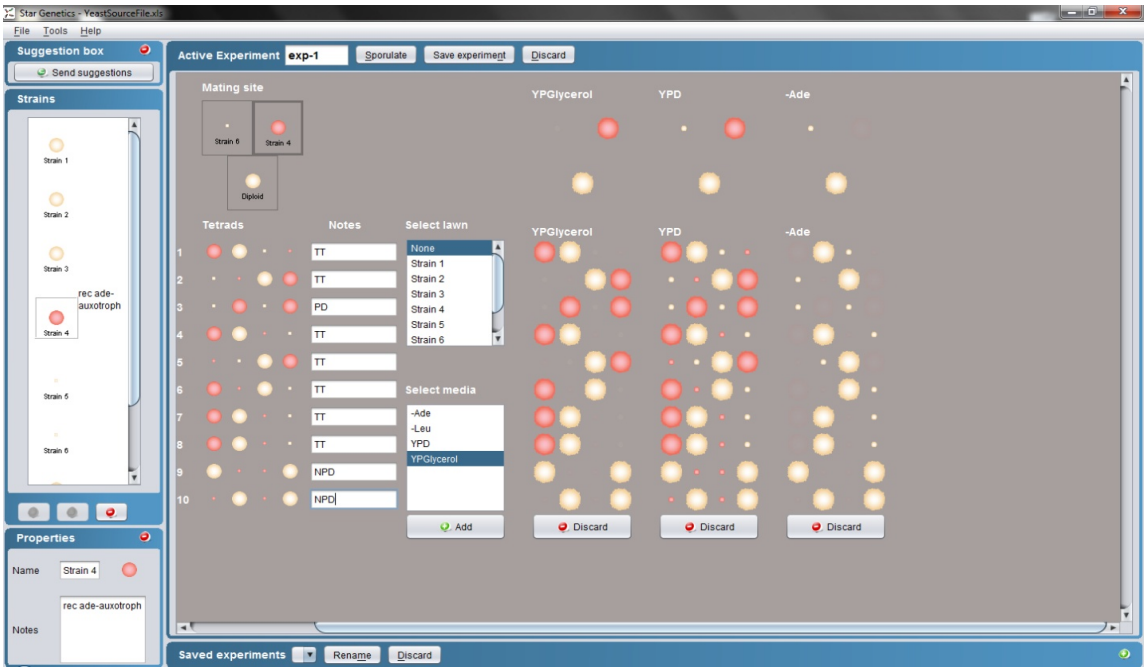

Figure 3. The StarGenetics yeast graphical user interface. The StarGenetics software also supports inquiry-based learning of yeast genetics. Simulations allow mating, tetrad analysis, and replica plating onto various media and lawns of cells. Variation in traits including inability to grow on various media and at defined temperatures, colony color, and colony size can be simulated. Instructors determine genotype-phenotype correspondences and strains available for initial crosses by modifying a simple Excel workbook available on our website. Students can isolate strains generated in their inquiries for subsequent experiments.