Using Physical Models to Complement Computer-Based Bioinformatics Labs: Assessing Student Performance and Reactions

Michelle A. Harris

Biology Core Curriculum Program
University of Wisconsin-Madison
307 Noland Hall, 250 N. Mills St.
Madison, WI 53515
maharris@wisc.edu

Michelle Harris earned a B.S. in the Teaching of Biology from the University of Illinois-Urbana and a M.S. and Ph.D. in Zoology from the University of Wisconsin-Madison. She is an Associate Faculty Associate in the Biology Core Curriculum Program (Biocore), a 4-semester introductory honors biology sequence at the University of Wisconsin-Madison. She teaches sophomore cell biology labs and junior organismal biology labs.

Reprinted From: Harris, M. A. 2004. Using physical models to complement computer-based bioinformatics labs: Assessing student performance and reactions. Pages 358-364, *in* Tested studies for laboratory teaching, Volume 25 (M. A. O'Donnell, Editor). Proceedings of the 25th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 414 pages.

Abstract

We conducted a controlled investigation in March 2003 to assess how well our introductory cell biology lab students achieved the goals of a computer-based bioinformatics lab when they also had physical, hand held molecular models to manipulate. Our students worked in teams of 2-3 to use the web-based Protein Explorer program to investigate regulatory transcription factor motif structure and function. Three of our six lab sections (23-25 students per section) were only allowed to use Protein Explorer to carry out this exercise, but the other 3 lab sections were also given physical models of their assigned motif and the segment of DNA that it normally interacts with. We evaluated student answers and compared the performance of lab sections that have access to physical models and Protein Explorer with those sections that use only Protein Explorer. We also surveyed students about their reactions to using physical models and/or Protein Explorer. Although students felt more confident about answering questions about molecular structure and function after doing our bioinformatics exercise, their performance on these questions improved only slightly. Students

preferred to use both models and Internet tools to investigate these questions, but students with access to both physical models and the Internet did not perform any better than students who used the computer only. We feel that our hypotheses would be better tested if students had more time to use the models and Internet tools.

Introduction

The relatively new field of bioinformatics uses mathematical algorithms to analyze large nucleic acid and protein sequence data sets generated by biologists (Campbell & Heyer, 2003; Krane & Raymer, 2003). Much of the software used to run the algorithms is readily available via the Internet. This has allowed scientists as well as students to have easy access to the information contained in these continuously updated databases. We have developed a "bioinformatics" lab exercise for our honors undergraduate biology students that takes about 2.5 hours for students to complete.

In March 2003 we conducted a study to assess how well these cell biology lab students achieved our learning goals for the bioinformatics exercise. We also wanted to find out whether students' access to physical, hand held molecular models would enhance their understanding of molecular structure and function during this exercise. This paper describes our efforts to empirically assess whether/how our instruction and the use of physical models affected our students' learning.

Some Perspective about Biocore:

The Biology Core Curriculum (or Biocore) Program at the University of Wisconsin-Madison is a 4-semester undergraduate honors biology program. Students apply as freshmen and must complete introductory chemistry and calculus before beginning the Biocore sequence as sophomores. There is a 3-credit lecture course each semester and 2-credit labs accompany the first three lecture courses.

Cell biology (Biocore 304) is the second lab in the Biocore sequence. In spring 2003 there were 142 total students in six lab sections of 23-25 students/lab. Each lab was three hours long and met once each week. Dr. Janet Batzli and I were co-chairs for this course. Graduate teaching assistants ran 50 minute discussion sections that met 1-2 days before lab, and Janet and I were in charge of each 3 hour lab.

This is the 15-week cell biology lab schedule we used in spring 2003: Microscopy (1 wk), Subcellular fractionation (1 wk), Enzyme catalysis (independent project, 3 wks), Photosynthesis (independent project, 2+ wks), Bioinformatics (<1 wk), Biotechnology (2 wks), Muscle motility (1 wk), Signal Transduction (independent project, 4 wks).

The Current Study

We had two main goals for our students as they completed the Bioinformatics exercise: (1) to become aware of and familiar with online bioinformatics resources and physical models as a means of studying molecular structure and function, and (2) to use these tools to learn how transcription factor motif structure influences DNA/protein interactions. We also had three goals for ourselves as instructors: (1) to assess whether our bioinformatics exercise increased students' understanding of transcription factor motif structure/function relationships; (2) to find out if the availability of physical models affects this understanding; and (3) to assess students' reactions to using computer programs and physical models.

We hypothesized that our students' ability to answer questions regarding transcription factor motif structure/function relationships would significantly improve after doing our bioinformatics exercise. We also predicted that students with access to physical models during this bioinformatics

exercise would show a significantly greater amount of improvement on these questions than students who did not use the models. We tested these hypotheses using data from baseline and post-exercise surveys as well as their scores on a follow-up assignment with specific questions about molecular structure and function.

Procedures

One week before the bioinformatics lab exercise, students anonymously filled out a baseline survey with four questions. These questions helped us determine their familiarity with transcription factor structure and function and gave us baseline data for post-exercise comparisons (see Table 1).

Table 1. Baseline Survey Questions

1. How might a transcription factor bind to DNA?
2. How do you think a transcription factor "knows" which part of DNA it should bind to?
3. How might a cell regulate a transcription factor (*i.e.*, turn it on or off)?
4. Do you think that the structure of a transcription factor is important in determining its function within a cell? (rate from 0= not at all important to 5=extremely important) Briefly explain your answer.

We wanted students to become somewhat familiar with the models before lab. One to two days prior to the 2-hour lab exercise, in discussion section, we showed students a 15-minute PowerPoint presentation describing how molecular images are generated from atomic coordinate files. The presentation also briefly described how computers are used to make physical molecular models. We showed students a hemoglobin model at the end of the presentation.

The first hour of lab was used to finish presentations from the previous week's exercise. Afterward we told students to assemble themselves into groups of 2 to 3 and to gather around one of the eight computers available in the lab room. We then randomly assigned each group to study one of three transcription factors: zinc finger, lac repressor, or max protein. We gave each group a specific Protein Data Bank (PDB) accession number for their transcription factor and instructed them to begin working their way through a handout with directions and questions for them to answer. We allowed lab sections 1 through 3 (n=59 students) to use computers with Internet access to view bioinformatics websites and molecular imaging programs. We gave students about 90 minutes to do this part of the exercise. Lab sections 4 through 6 began their exercise using the computer as well, but about 20 minutes into the exercise we showed each group a physical, hand held model of their transcription factor and encouraged them to use the physical models as well as the computer to answer the handout questions. The bioinformatics websites that students viewed are listed in Table 2.

Table 2. Bioinformatics Websites

NCBI (National Center for Biotechnology Information)	http://www.ncbi.nlm.nih.gov/
NCBI's PubMed	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed
Entrez Protein	http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Protein
Entrez Structure	http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Structure
Protein Explorer	http://molvis.sdsc.edu/protexpl/frntdoor.htm

About 30 minutes before the end of the period I led a class discussion in which three randomly chosen teams reported what they had learned about their respective transcription factors. I then gave students more details about their take home assignment. Students used the last 5 to 10 minutes of class to fill out a post survey that contained the same 4 questions as the baseline survey as well as 3 additional questions about their self efficacy beliefs and learning tool preferences (Table 3).

Table 3. Post Survey Questions

1-4. Same questions as on baseline survey
5. Do you believe you can answer the above questions correctly after completing this lab exercise? (rate 1= cannot do at all to 10= certain can do). Please explain response.
6. Rate how the following tools helped facilitate your learning of txn factor motif structure & function (rate 1= of no help to 5= helped a great deal): textbook, pre-lab, intro PPT, lab manual, computer, models, TAs/instructors, peers (Note: the "models" choice was left out of survey for the three labs that did not see the physical models.)
7. (Asked only in the three lab sections that saw the physical models): Imagine that you are asked the following question: *"Summarize how a regulatory txn factor interacts with DNA. What are the secondary protein structures involved in this interaction? What amino acids are interacting directly with the DNA?"* Which materials would you most likely use to answer these questions? (Check only one.)
    - o Protein Explorer computer program
    - o A physical, hand-held model
    - o A combination of Protein Explorer and the model

Data Analysis

Dr. Michael Patrick and I evaluated the responses to baseline and post-survey questions 1-4 on a four point scale (0 = no answer/less than satisfactory answer; 1 = "C" letter grade answer; 2 = "B" answer; 3 = "A" answer). Papers were identified by numbers rather than by student names. The change in mean score between the baseline and post-survey score, for each question, was analyzed using the Mann-Whitney test. We calculated mean responses for questions 5-7.

Results

For questions 1-4, the mean post-survey response grades were significantly higher than baseline survey scores for students in all 6 lab sections (n=119). Post-survey scores, however, remained below a "B" letter grade (see Fig. 1). For questions 1, 3, and 4, students who had access to the physical models and the Internet tools performed equivalently to the students who only had access to the Internet tools. For question 2, however, students who used only the Internet showed a significantly *greater* improvement in their responses from baseline to post-survey than students who used both models & computers (Mann-Whitney U value =1017.5, p=0.000).

The mean score on the take home assignment for the students who had access to the models (86.8%, n=67) was not significantly different from the mean score for students without models (87.2%, n=71; independent samples t-score = 0.414, 2-tailed p=0.68, 130 df).

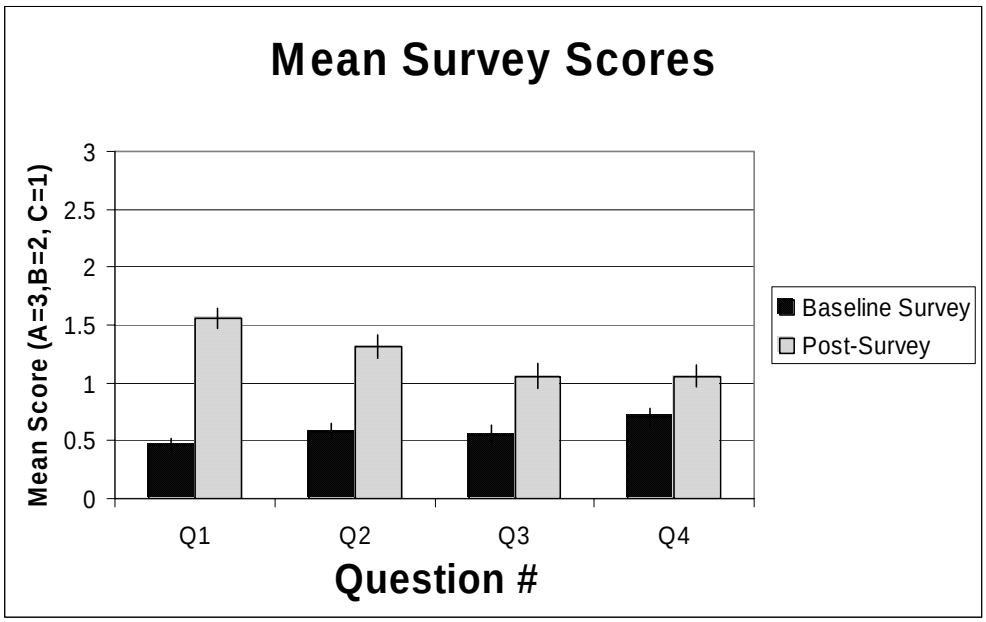

Figure 1. Mean scores for survey questions 1 through 4. These 4 questions were identical on the baseline and post surveys.

Self-efficacy beliefs

When students were asked how confident they felt about answering questions 1-4 on the survey after doing the bioinformatics exercise, the average self-efficacy rating for the entire student population (n=119) was 6.81 (SD.=1.46) on a scale of 1 to 10. The rating selected by most students was a 7 on the 10-point scale.

Preferences

Students in both groups found the Internet tools and the instructors to be the most helpful resources. Students with access to the physical models also rated them relatively high (see Table 4).

Table 4. Learning Tool Preference Mean Ratings (Helped: 1=not at all, 2 =a little, 3 =some, 4 =a good deal, 5 =a great deal)

Learning Tool	Labs without models (SD)	Labs with models (SD)
Computer program	4.4 (.72)	4.4 (.74)
Physical models	NA	4.2 (.88)
TAs/instructors	4.0 (.97)	4.0 (.95)
Peers	3.8 (.86)	3.8 (.77)
Lab manual	3.6 (.90)	3.5 (.93)

When students with access to the models were asked what materials they would most likely use to answer questions about transcription factor structure and function, 63% said they would prefer to use the combination of computer tools and a physical model, 29% chose the computer program alone, and 8% chose the physical models alone.

Discussion

Our first hypothesis, that students' ability to answer questions regarding transcription factor motif structure/function relationships would significantly improve after doing our bioinformatics exercise, was supported. The improvement, however, was very modest (see Fig. 1). This was particularly surprising given that most students reported that they gained some level of confidence in answering questions after doing the bioinformatics exercise.

We had to reject our second hypothesis, which predicted that students with access to physical models during this bioinformatics exercise would show a significantly greater improvement in answering protein structure-function questions than students who do not use the models. Performance on the 3 of the 4 survey questions was not significantly different between the two groups. Regarding survey question 2 (see Table 1), students with models actually did *worse* on one of the questions than students who only had access to the Internet tools. Again, we found a surprising discrepancy between student preferences and their academic performance. Although the majority of students who had access to the models preferred to use them in combination to the computers and also reported them to be just as helpful as the Internet tools and instructors in completing the exercise, their performance on the survey and the take home assignment was not significantly higher than students who had access to the computer tools only.

In retrospect, however, we feel that we very likely did NOT test our hypotheses with this protocol.

We did not give students enough time to become acquainted with either the Internet tools or the physical models. In the future we will introduce students to Internet tools and physical models early in the semester and continue to use them throughout the course. A study by Roberts et al. (in prep) provides good evidence that introductory biochemistry students who repeatedly use these physical molecular models over several weeks do show significant improvement in their ability to answer questions about molecular structure and biochemistry. These students also showed an increased level of sophistication in their answers and used more appropriate vocabulary after they had repeatedly used the models to learn new concepts. Like our students, these biochemistry students rated the models very high in terms of learning.

We feel that investigations of molecular structure-function relationships are probably most relevant to students when they are used to help formulate students' biological rationale in independent research investigations. After we introduce students to Internet tools and molecular models very early next semester we will have clear expectations for their continued use, particularly in doing background research for independent investigations. We will use a variety of physical models over the course of the semester to emphasize key molecules in each unit. Finally, we want to design better ways to evaluate our students' understanding of molecular structure and function, particularly their mastery of 3-dimensional concepts.

Conclusions

Although students felt more confident answering questions about molecular structure and function after doing our bioinformatics exercise, their performance on these questions improved only slightly. We found that students preferred to use physical molecular models in addition to Internet tools to investigate these questions, but that students with access to both physical models and the Internet did not perform any better than students who used the computer only. Because of serious concerns we have about our experimental design, we do not think we truly tested our hypotheses

with this protocol. The current study is still extremely useful though because it will guide improvements to both our class use of Internet and molecular models and in our learning assessment tools.

Acknowledgments

Many thanks to Janet Batzli, UW-Madison Biocore Program Associate Director and my teaching partner, for her invaluable feedback regarding this study. Tim Herman (Milwaukee School of Engineering and Center for Biomolecular Modeling) designed and produced the models and generously allowed us to borrow them for this study. Tim Herman and Mike Patrick (UW-Madison Genetics) provided excellent information about molecular structure and function. Mike Patrick, Paul Dillenburg (UW-Madison LEAD Center), and Eric Hagedorn (UW-Milwaukee) made the data collection and analysis happen.

References

Campbell, A. M., and L. J. Heyer. 2003. Genomics, proteomics, & bioinformatics. Benjamin Cummings, San Francisco, California, 352 pages.

Krane, D. E., and M. L. Raymer. 2003. Fundamental concepts of bioinformatics. Benjamin Cummings, San Francisco, California, 314 pages.

Roberts, J. R., T. Herman, M. Patrick, E. Hagedorn, and P. Dillenburg. In prep. Physical models enhance molecular 3D literacy in an introductory biochemistry course.