CBE—Life Sciences Education
Vol. 5, 52–64, Spring 2006

Article

Incorporating a New Bioinformatics Component into Genetics at a Historically Black College: Outcomes and Lessons

J. David Holtzclaw,* Arri Eisen,† Erika M. Whitney,* Meera Penumetcha,* J. Joseph Hoey,‡ and K. Sean Kimbro§

*School of Medicine and †Department of Biology, Emory University, Atlanta, GA 30322; ‡Georgia Institute of Technology, Atlanta, GA 30332; and §Clark Atlanta University, Atlanta, GA 30314

Submitted April 12, 2005; Accepted September 21, 2005
Monitoring Editor: Elizabeth Vallen

Many students at minority-serving institutions are underexposed to Internet resources such as the human genome project, PubMed, NCBI databases, and other Web-based technologies because of a lack of financial resources. To change this, we designed and implemented a new bioinformatics component to supplement the undergraduate Genetics course at Clark Atlanta University. The outcomes of the Bioinformatics course were assessed. During the first week of the semester, students were assigned the Felder-Soloman's Index of Learning Styles Inventory. The overwhelming majority of students were visual (82.1%) and sequential (75.0%) learners. Furthermore, pre- and postcourse surveys were administered during the first and the last week of the course to assess learning, confidence level, and mental activity. These indicated students increased the number of hours spent using computers and doing homework. Students reported confidence in using computers to study genetics increased, enabling them to better visualize and understand genetics. Furthermore, students were more mentally engaged in a more social learning environment. Although the students appreciated the value of the bioinformatics component, they reported the additional work load was substantial enough to receive additional course credit.

INTRODUCTION

Bioinformatics is the use of computer science, mathematics, and information technology to collect, to organize, and to analyze large volumes of biological data. Biological data come from a large array of subjects including cellular and molecular biology, genetics, biochemistry, evolutionary biology, physiology, and several others. Recent research efforts such as the Human Genome Project and new technology such as DNA microarrays have produced enormous volumes of genetic information waiting to be mined by specialized software. This has produced a growing demand for trained bioinformaticians, making them one of the most sought after and fastest growing sectors in biotechnology. According to a survey of 176 biotechnology companies (Virginia Commonwealth University Center for Public Policy, 2001), most firms plan on hiring two new employees in bioinformatics within the next 12 mo and seven more in the next 5 yr at an average starting annual salary of $45,000 for an entry-level position (M.S. degree or less). Unfortunately, many academic institutions are not prepared to meet this immediate need. Hence, it is estimated that 20,000 jobs in bioinformatics will be left unfilled by 2005 (Eisenberg, 2002).

As previously discussed in *Cell Biology Education,* historically black colleges and universities (HBCUs) are doing their part to help America meet this need despite limited federal support (Suitts, 2003). Unfortunately, many HBCUs lack the resources to implement courses in bioinformatics. Furthermore, faculty at many HBCUs developed their research focus before the evolution of bioinformatics. Hence, most biology students at HBCUs are not exposed to online resources such as the Human Genome Project, PubMed, National Center for Biotechnology Information (NCBI) databases, or other related tools (i.e., BLAST, Cn3D, etc.). To change this, postdoctoral fellows from the Fellowships in Research and Science Teaching (FIRST) Program designed

DOI: 10.1187/cbe.05–04–0071
Address correspondence to: J. David Holtzclaw (david.holtzclaw1@jsc.nasa.gov).

and implemented a new bioinformatics component to supplement the undergraduate Genetics course at Clark Atlanta University (CAU).

FIRST is part of a National Institutes of Health initiative from the Minority Opportunities in Research Division of the National Institute of General Medical Sciences. Grants, known as Institutional Research and Academic Career Development Awards (IRACDA), from this program combine a traditional mentored postdoctoral research experience at a research institution (Emory University) with an experience to develop teaching skills through innovative programs that involve mentored teaching assignments at minority-serving institutions (MSIs; Holtzclaw *et al.*, 2005). The objectives of this initiative are threefold: 1) to enhance research-oriented teaching at MSIs; 2) to increase the research and other skills needed by training scientists to conduct high-quality research in an academic environment; and 3) to promote linkages between research-intensive institutions and MSIs that can lead to further collaborations in research and teaching. Finally, a desired long-term outcome is to increase the number of well-qualified underrepresented minority students entering competitive careers in biomedical research (National Institute of General Medical Sciences, 2002).

The specific goals of the bioinformatics component were threefold: 1) to provide CAU students exposure and introductory training in bioinformatics, demonstrating an alternative career path; 2) to provide CAU students a more interactive, visually oriented, and discovery-based learning approach to genetics; and 3) to allow CAU to assess the need for incorporating bioinformatics at various levels into its curriculum. The Biology Department at CAU was chosen for this initiative because of its computer and distance-learning resources and the existence of a graduate program, which together have the potential to establish CAU as a key segment of the pipeline providing the private and public sectors with well-trained minority bioinformaticians.

Several strategies for incorporating bioinformatics into the undergraduate curriculum have previously been described in *Cell Biology Education* (Campbell, 2003; Honts, 2003). Here we address this issue from the perspective of an HBCU. Second, we examine the potential role of bioinformatics in complementing and enhancing an undergraduate genetics course. Furthermore, we assess student comfort and confidence with the computer, Internet resources, and genetics through pre- and postcourse surveys. We also assess the student's learning styles, level of engagement in Bloom's taxonomy (Bloom and Krathwohl, 1984), effectiveness of the course and course Web site, and how well the course fulfilled the students' career objectives.

METHODS AND ASSESSMENT

Pilot Course

A pilot bioinformatics component for the undergraduate Genetics course (Biology 312) was offered at CAU in Spring 2003. Participation in the bioinformatics component, which occurred during the recitation period, was voluntary with extra credit given for submitted homework projects. The recitation classes were held in the distance-learning laboratory at CAU. This facility holds 28 desktop computers with high-speed Internet access and a classroom projector from which the instructor could project his/her laptop into the screen to guide students through different exercises. In terms of information technology, this was an ideal setup.

The pilot bioinformatics component contained four modules. These modules included a general introduction to PubMed (many of our students had never used PubMed before this class), and tutorials on BLAST, GenBank, and Map Viewer. Fourteen students participated, and postcourse surveys were completed (our unpublished data). On the basis of student feedback, we made several changes to the content and logistics for the Spring 2004 course. It is highly recommended that instructors run a pilot course on a small number of students (5–10) to determine how much time per class to allow for logistical concerns (setting up computers, connecting to the Intranet, accessing databases, etc.).

Course Infrastructure

Because of scheduling complexities of computer facilities, we had to find a new location for the bioinformatics recitation class during the Spring 2004 semester. Available computer facilities were already overburdened and in continual use. Ideally, we wanted to develop an educational environment in which bioinformatics could be taught in any of the four core laboratories and two lecture halls primarily used by the Biology Department at CAU. Conveniently, all these labs and classrooms exist on the same two, concurrent floors of the science building. This facilitated the installation of new information technology resources. With the financial assistance of the FIRST Program, we installed a Cisco 1200 Series wireless local area network (WLAN), 802.11g IOS (Cisco Systems, San Jose, CA) and purchased 22 Dell Inspiron laptops with internal, wireless, mini-PCI cards (1300 WLAN, 802.11g, Dell Computers, Round Rock, TX). These wireless, laptop computers were used in the bioinformatics recitation held in the general biology laboratory space.

Course Description and Content

Genetics was a three-credit course typically taught for 1.5 h twice per week with a voluntary recitation class for 1 h once per week (Genetics course outline and schedule are given in Appendix Table A1). Historically, most of our genetics students are juniors with a few sophomores and seniors. The only course prerequisite was Cell Biology. The current Biology Department curriculum has no computer or calculus course requirements, although they are strongly recommended. Therefore, we could not alter the course requirements by requiring calculus or computer programming, although we would for future courses (see *Discussion* section).

The bioinformatics component was taught once per week during the recitation period of the Genetics course. The complete course syllabus as well as handouts, homework assignments, and supplemental materials are all available on the course Web site (Holtzclaw, 2004). The bioinformatics component consisted of 22% (200/900 points) of the final Genetics grade. We covered introductory bioinformatics such as how to access and interpret information from the publicly available databases (PubMed, nucleotide, protein, and structure databases, etc.). We chose topics that would be of interest to our student population (e.g., sickle cell anemia, diabetes, breast cancer, etc.). Each topic was presented through a case study, termed "module," and case-based teaching pedagogy was used (Herreid, 1994).

Briefly, case-based teaching pedagogy includes the use of a concrete, real-world problem (i.e., diseases, environmental conditions, etc.) to teach scientific theory or knowledge. In the context of a Bioinformatics course, each case was designed to focus on a particular macromolecule, related to a diseased state, and investigated through a database. Students were presented a problem or case (i.e., disease or environmental condition) and were required to use the NCBI databases to address it. During the course of solving the case, students would learn about a particular macromolecule and apply course theories and concepts. For example, a sickle cell anemia case can be used to teach students how to use the NCBI nucleotide database by having them look up hemoglobin A (accession number

NM_000518) and HbS (accession number M25113) and compare the sequences until they find the point mutation for either amino acid or nucleotide sequence (see Appendix for more details on this case).

We then demonstrated, using software and images available in protein-structure databases, how that single point mutation results in a change in the three-dimensional conformation of the molecule. This case was a powerful example of how a single genetic mutation at the nucleotide level can cause conformational changes, resulting in a life-threatening disease. Although students can be exposed to the same material through a traditional lecture format or by reading a textbook, our students learned this information in a rich, hands-on context and can apply these same skills to discover the mechanisms of other genetic diseases.

Modules were organized to focus on a selected database and to build on previously discussed modules. For example, the second module (sickle cell anemia) just described required understanding of the nucleotide database presented in the first module. Another module focused on diabetes and engaged the OMIM (Online Mendelian Inheritance In Man) database and a specific journal article (PubMed) with questions for the students to answer. A third module on breast cancer introduced later in the course required the use of a previously used database (PubMed), as well as introduced new ones (for protein sequences and protein structures).

Importantly, we organized modules to align with content presented in the Genetics course lecture (see Appendix for course schedule). The sickle cell module also required knowledge of gene expression (i.e., transcription and translation), which was covered in lecture during the weeks before the students did that module. Similarly, module 3 on mitochondrial DNA and module 4 on drug resistance corresponded to the related lecture topics of genome analysis, Mendelian genetics, the chromosomal basis of inheritance, and non-Mendelian inheritance. Likewise, the modules on the SNP (single nucleotide polymorphism) database (April 13) and gene therapy (April 20) were presented in sync with the lecture on population genetics (April 20). The Genetics course textbook was *Genetics*, by Peter Russell, 5th edition (Russell, 1997).

Module Design and Format

A typical bioinformatics module consisted of students coming to recitation, checking out a laptop, and downloading the in-class assignment from the course Web site (http://userwww.service.emory.edu/~jholtzc/Courses/Bio312/index.htm). Students in groups of two or three then began working on module exercises (a portion of module 8 is given below as an example).

Module Exercise: Sample from Module 8

1. Create a folder on your desktop labeled structure. Save any structures that you download to this folder.
2. Click on the Structure database from the Entrez homepage, and find the Cn3D tutorial. Download Cn3D if it's not on your computer already. Read the Cn3D "Introduction":
3. Cn3D can show you structures of which of the following (answer all that are correct)?
   A. linear DNA
   B. circular mRNA
   C. a specific chromosome
   D. proteins
   E. all of the above

4. Read the first section, "Retrieving individual structures," and do all of the exercises. http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3dtutP2.shtml

5. What are the results of the PubMed, Protein, and Structure search for Hemoglobin A, HbA, Hbα, and Hb-A? What is the number of results for each search? Give an example of each and the species.

6. Using the following methods, find the corresponding MMDB structure files:
   A. Do an Entrez/PubMed database search to find the crystallographic or NMR structures for PTEN, as in the example, Hemoglobin S (Hb S), and Hemoglobin A (Hb-A).
      i. Which database/query/links did you use for each protein? Did you use any limits? If so, which ones?
      ii. How many results did you get for each protein?
      iii. Give a reference for a structure of each protein. Remember, the left side is highlighted in green or yellow to indicate references that are available online.
      iv. Find 3 pictures/figures from the available references, save them in your structure file. You might have to open the figure in a new window first, before you save the file. Make sure you name the file appropriately.

7. A 3D structure is ideal, but not always available. What if there is no structure file for the protein that you are looking at? To find mutations that have no crystal structure, you can use the reference protein's known structure. Go to the Entrez site. Do an Entrez sequence neighbor search by doing a Protein database/Genpept search for human PTEN (use NP_000305 instead of O00633), Hemoglobin A (HbA) B chain, and HbS beta chain.
      i. Find and save the sequence for each protein in a text file.
      ii. How many 3D domains or chains does each protein (PTEN, Hemoglobin S, and Hemoglobin A) have?
      iii. Do a Blink to find similar structures. List the accession number, gi number, and the protein description of 5 PTEN mutants and 5 alpha and beta chain mutants for both Hemoglobins.
      iv. Describe the mutations for PTEN, Hemoglobin A, and Hemoglobin S.
      v. Look at the 3D alignment of the 5 mutants/protein/chain that have crystal structures in Cn3D. Save to your structure file.

The instructor(s) facilitated this process and then, after 20–30 min, assessed class progress, answered questions, and walked the students through the in-class assignment, providing additional information and examples. Then, the instructor reviewed the homework assignment, which typically was similar to or a continuation of the in-class assignment. The instructor also attended the 2-h, weekly help session held during the evening, which was convenient for some, but not all, of the students.

ILS and Pre- and Postcourse Surveys

The Felder-Soloman's Index of Learning Styles (ILS) is a self-scored, Web-based instrument that assesses learning style preferences on four dimensions: sensing/intuiting, visual/verbal, active/reflective, and sequential/global (Felder and Silverman, 1988). The ILS has been shown to be a suitable psychometric tool for evaluating learning styles of students (Zywno, 2003). During the first week of the semester, the 45 undergraduate students taking genetics (Biology 312) were assigned the ILS. In addition, pre- and postcourse surveys were given during the first and last week of the course. Surveys were anonymous and postcourse surveys were analyzed after grades were submitted. Differences in responses between the precourse (aggregate) and the postcourse (also aggregate) responses were determined by the chi-square test (StatSoft, 2004) using SPSS software (Chicago, IL). Homework credit was given for completion of the ILS as well as the postcourse survey.

RESULTS

Although it is clear that bioinformatics is essential to a contemporary biology curriculum, a major question is how to include it most effectively. We measured the effectiveness of our particular approach by investigating how well it fit

with our students' learning styles and examining a wide range of student learning variables before and after the course.

Learning Styles Assessment

Twenty-eight students completed and submitted the ILS. The ILS scales are bipolar with mutually exclusive answers to each question (either A or B) with an odd number of questions (Zywno, 2003) for each dimension. Students exhibited three predominant learning styles (Figure 1). The overwhelming majority of students were visual (82.1%) and sequential (75.0%) learners, who showed a preference for sensory learning (67%). In other words, these students prefer to visualize the course materials as diagrams, sketches, or schematics. Furthermore, they wanted the syllabus and class material to follow a linear, stepwise, logical path and had a tendency to learn material that had real-life relevance.

On the basis of the ILS assessment shown in Figure 1, we suggest our case-based, module approach to bioinformatics enhanced students' learning of genetics by providing information in the students' preferred learning style in three ways. First, the bioinformatics component, through the use of graphical software such as Cn3D and a computer interface, was highly visual, allowing for our more visually oriented (Figure 1) students to study macromolecules from 360°—a strategy more difficult to integrate within a traditional lecture format. Second, our case-based method provides students with real-world application of genetics through examples such as diabetes and gene therapy. Third, our systematic, stepwise approach to the implementation of different databases provides a logical progression of information that would be beneficial to our sequential learners. Results of the pre- and postcourse surveys strongly support these conclusions. By presenting information in alignment with the students' preferred learning styles, we transfer the effort and energy students exert from formatting information to comprehension and application. Instructors should be careful not to fall into the trap of using only one or two preferred learning styles of the class, but to incorporate as many learning styles as possible to reach every student in the class.

Pre- and Postcourse Surveys

Forty students completed the precourse survey, and 32 students completed the postcourse survey. The goal of these surveys was to provide self-reported measures of student learning, level of mental activity, computer confidence, value of bioinformatics in relation to their educational and career goals, effectiveness of the course Web site, and ratings of instructor attributes. In both pre- and postcourse surveys, students were asked to "estimate your confidence level right now," on a five-point scale where 5 = high and 1 = low, in using various tools to study genetics or molecular biology. Results are given in Figure 2. After the class, 60.6% of students rated their level of confidence as high or good in using the computer to study genetics, a significant increase from the 25% who gave it this rating in the precourse survey ($p \leq 0.005$, Figure 2A). Similarly, 56% of students rated their confidence level as high or good postcourse compared with only 30% precourse in using Internet databases, tools, and software to study genetics ($p \leq 0.04$, Figure 2B).

Because 82% of our students had a visual learning style preference (Figure 1), we assessed whether the visual nature of bioinformatics would aid them in learning genetics. By visual nature of bioinformatics, we mean the computer-based tools such as Cn3D and MapViewer, which present information graphically. When asked about their confidence level in using Internet databases, tools, and software to effectively visualize genetics, 53.1% rated their confidence level as high or good postcourse compared with 32.5% precourse (Figure 2C, $p \leq 0.05$). In the surveys, we did not define what it means to "effectively visualize genetics," leaving it up to the student to define. Clearly, from the postcourse responses, "effectively visualize genetics" was defined by the student as the approach used in the bioinformatics recitation (i.e., using computer technology and databases to learn concepts in genetics). This definition may have been more ambiguous in the precourse survey, potentially leading to the sharp increase in confidence level.

Finally, we assessed the use of case-based learning in our course to increase their understanding of genetics (Figure 2D). Postcourse, 86% of the students rated their confidence level in using problem-based learning to understand genetics as moderate, good, or high as compared with 53% precourse ($p \leq 0.02$). Overall, in Figure 2, one sees a "leftward" shift to high, good, and moderate from moderate, low, and fair in the students' responses postcourse compared with precourse. Therefore, we concluded that the Bioinformatics primer enabled the students to better visualize and understand molecular structures, thus enhancing their learning of genetics.

We also assessed whether the bioinformatics component increased student computer usage. Based on precourse survey results, 67.5% of the students had never used computers in any biology class before (our unpublished data). As a direct result of this course, students spent more time on the computer (Figure 3A, $p \leq 0.001$) and the Internet (Figure 3B). Although we do not know for certain if this extra time was academically related or not, several students did mention to us, both verbally and in postcourse evaluations (our unpub-

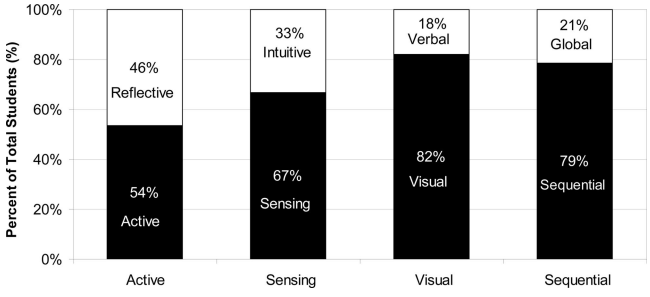

Figure 1. Learning style preferences. The Felder-Soloman's Index of Learning Styles (ILS) is a self-scored, Web-based instrument that assesses preferences on four dimensions: sensing/intuiting, visual/verbal, active/reflective, and sequential/global. The ILS scales are bipolar with mutually exclusive answers to each question (either A or B) with an odd number of questions for each dimension. Twenty-eight students completed and submitted the ILS. The percentage of students that fell into each dimension is given.

lished data), that they did use the NCBI databases for other biology courses. Students also spent more time on homework (Figure 3C), which is probably due to a combination of the increased workload required for this course as well as students taking more core biology courses as they advanced in the degree program. Although students were spending more time on the computer and doing homework, there was little to no change in the amount of time they spent in financially compensated activities (jobs, work study, etc.) or extracurricular activities (student organizations, sports, church-related activities, etc.) between pre- and postcourse evaluations (our unpublished data).

We also assessed the effectiveness of the course Web site by asking a series of questions shown in Table 1. Students were asked "To what extent did utilizing the course Web site. . . . " and were given a five-point scale that ranged from 5 = greatly to 1 = not at all. The Web site promoted greater access to the course materials (84% responded greatly or moderately), and connecting to the NCBI Web site (77% responded greatly or moderately). In addition, the Web site allowed students to schedule time for the course relative to their work and personal responsibilities (66% responded greatly or moderately) while providing background and additional information outside of lecture (71% responded greatly or moderately). Although these results were as expected, we were surprised to find that students credited the course Web site with greatly or moderately increasing interactions among students enrolled in the course (71%) and working collaboratively with other students (79%). Furthermore, students credited the Web site for greatly or moderately increasing their effectiveness to organize or express their comments or questions (66%) or seek answers to their questions (69%). Similar to and in support of results obtained in Figure 2, the course Web site either greatly or moderately enhanced the students' ability to visualize the ideas and concepts taught in the Genetics course (72%). Half the students (48%) responded that the Web site greatly or moderately increased their understanding of genetics.

Although the surveys were anonymous and postcourse surveys were analyzed after grades were submitted, it is possible that the students simply gave us the answers they thought we wanted or just simply filled in bubbles ran-

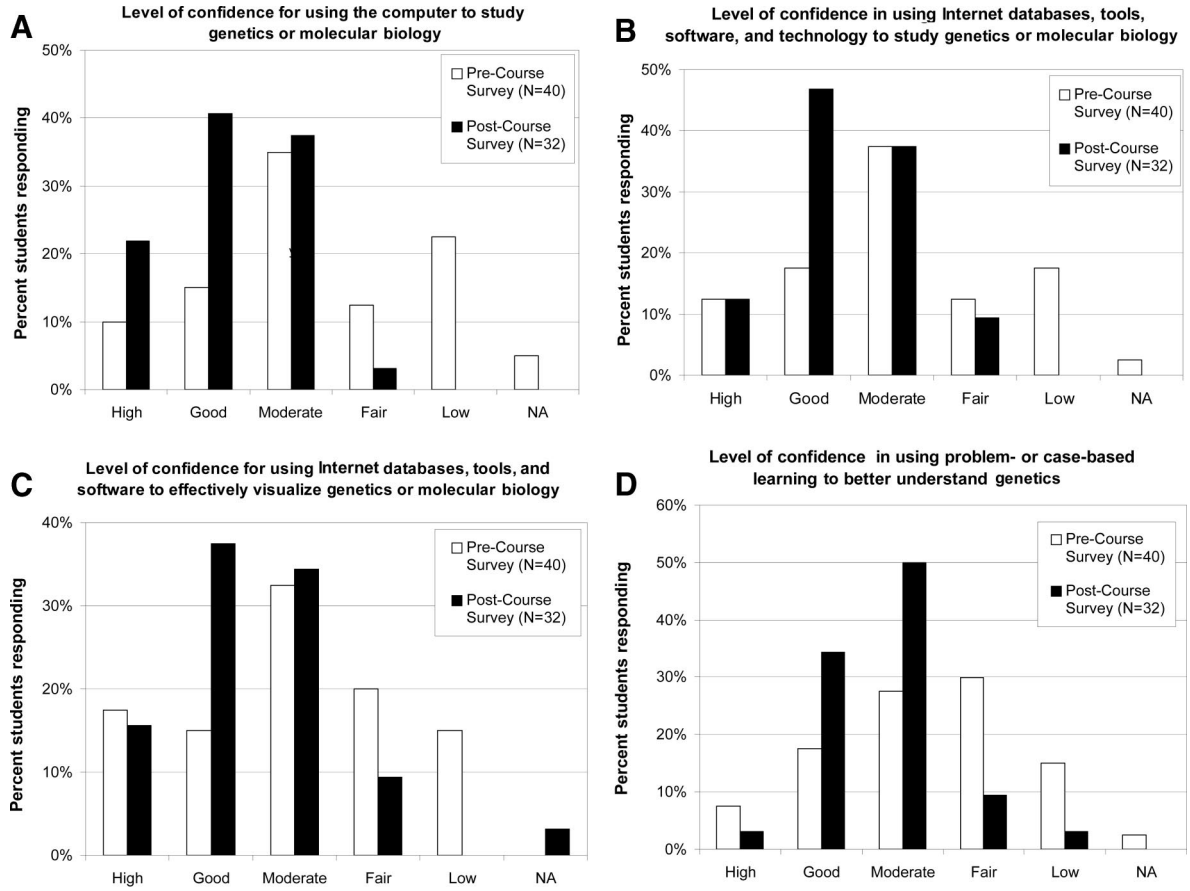

Figure 2. Pre- and postcourse survey results on students' confidence levels. In both pre- and postcourse surveys, students were asked to "estimate your confidence level you have right now in your knowledge and skills in. . . (one answer for each item)": (A) computers to study genetics or molecular biology; (B) Internet databases, tools, software, and technology to study genetics or molecular biology; (C) Internet databases, tools, software, and technology to effectively visualize genetics or molecular biology; (D) problem- or case-based learning to better understand genetics. To respond, students were given a five-level scale that ranged from high to low (NA = no response). Forty students completed the precourse survey, and 32 students completed the postcourse survey.

domly. This is always a possibility in course surveys and is difficult to impossible to completely prevent. For the postcourse survey quality control check, we compared the answers for three concurrent questions: Question A, "How many hours per week did you spend this semester on the computer?"; Question B, "How many hours per week did you spend this semester on the computer for homework or coursework?"; and Question C, "How many hours per week did you spend this semester on the Internet?" If a survey had more hours for Question B or C than Question A, then that survey was eliminated from the database.

Assessment of Mental Activity

We also asked students in the postcourse evaluations to characterize their level of mental activities during the bioinformatics modules in the postcourse evaluations (Table 2). The levels of mental activity were based on Bloom's taxonomy of educational objectives (Bloom and Krathwohl, 1984) and derived from engagement theory (Astin, 1984; Chickering, 1991). Students were asked "Please estimate how often in this course you were engaged in each kind of mental activity given below" and were given a five-level scale that ranged from 5 = very much to 1 = very little. In summary, students felt engaged at several levels of Bloom's Taxonomy. Nearly 20% of students felt they were engaged "very much," analyzing, synthesizing, evaluating, and applying information presented. Approximately 70% of students felt engaged "quite a bit" or "a moderate amount" at knowing, analyzing, synthesizing, or evaluating information presented in class.

Although the students were engaged at several levels of learning, they indicated the bioinformatics modules could have been better organized (Table 2). Only half the students (52% responding "very much" or "quite a bit") felt the bioinformatics modules were well organized, whereas only 32% felt that the modules complemented each other or the genetics lectures (19%) despite our efforts along these lines.

Overall Course Value

Table 3 summarizes feedback from the postcourse student evaluations on the value of the bioinformatics modules in relation to their learning and career goals. Thirty-nine percent of students reported the bioinformatics module was "highly valuable" or "quite valuable" to their career or educational goals. Similarly, 42% of students reported that the Web site informed them of potential graduate or professional opportunities (Table 1). Clearly, we were able to meet our first objective and provide students bioinformatics training and alternative career paths. Furthermore, based on the assessment data presented in Tables 1–3, we provided through the case-based modules more interactive, visually oriented, discovery-based method of instruction. However, we clearly did not present well-organized, teaching sessions nor did we relate modules presented to previously presented modules or genetics lectures.

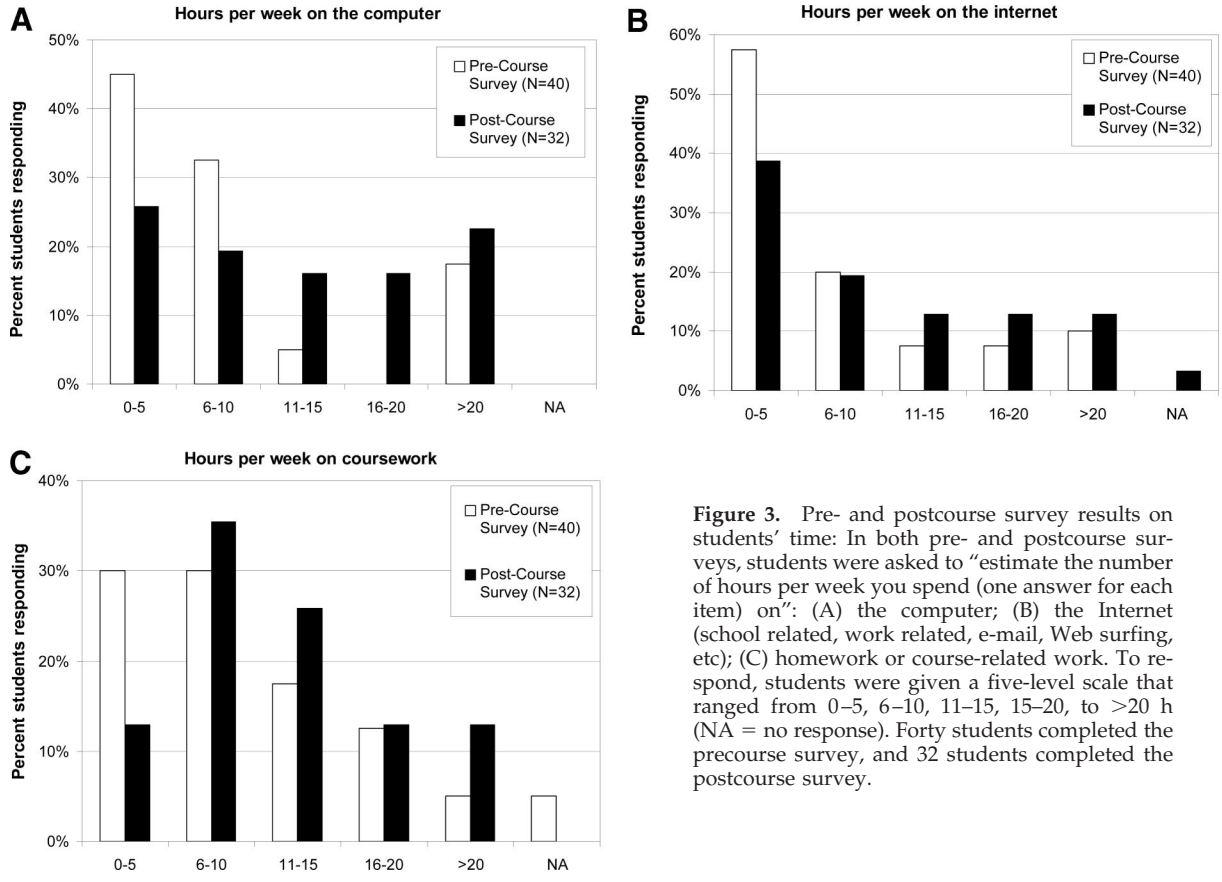

Figure 3. Pre- and postcourse survey results on students' time: In both pre- and postcourse surveys, students were asked to "estimate the number of hours per week you spend (one answer for each item) on": (A) the computer; (B) the Internet (school related, work related, e-mail, Web surfing, etc); (C) homework or course-related work. To respond, students were given a five-level scale that ranged from 0–5, 6–10, 11–15, 15–20, to >20 h (NA = no response). Forty students completed the precourse survey, and 32 students completed the postcourse survey.

The final objective of this pilot course was to assess the need for an undergraduate Bioinformatics course at CAU. All of the students (N = 32) felt that the bioinformatics recitation should be taught as an independent course. This was not because they did not like the course—they clearly gained much from it, and 71% of them said they would recommend it to their peers (Table 4). However, the same percentage felt that the recitation time would have been better spent learning genetics rather than learning bioinformatics. In general, many students struggled in the genetics

Table 1. Evaluation of course Web site

To what extent did utilizing the course Web site:	Greatly	Moderately	Somewhat	Very little	Not at all	NA
Enhance your understanding of the course materials?	6.5	51.6	38.7	3.2	0.0	0.0
Provide greater access to the course materials?	19.4	64.5	16.1	0.0	0.0	0.0
Make the subject matter more relevant to you?	16.1	38.7	29.0	9.7	3.2	3.2
Make connecting to databases easier?	32.3	45.2	22.6	0.0	0.0	0.0
Provide background or additional information outside of lecture?	38.7	32.3	29.0	0.0	0.0	0.0
Increase your understanding of biology or genetics?	19.4	29.0	45.2	6.5	0.0	0.0
Inform you of potential graduate or professional opportunities?	25.8	16.1	19.4	12.9	22.6	3.2
Visualize the ideas and concepts taught in this course?	17.2	55.2	24.1	3.4	0.0	0.0
Be more interactive with other students?	25.8	45.2	22.6	3.2	3.2	0.0
Work collaboratively with other students on assignments and projects?	34.5	44.8	10.3	3.4	6.9	0.0
Organize and express your comments or questions?	10.3	55.2	17.2	10.3	3.4	3.4
Seek answers to your questions?	13.8	55.2	24.1	6.9	0.0	0.0
Schedule time for the course relative to your work and personal responsibilities?	13.8	51.7	27.6	3.4	3.4	0.0

Values are percentages.

Table 2. Self-Assessment of mental activity

Mental activity	Very much	Quite a bit	A moderate amount	Some	Very little	NA
Memorizing facts, ideas, or methods from the lectures, Web sites, and readings so you can repeat them in pretty much the same form	9.7	35.5	35.5	9.7	9.7	0.0
Analyzing the key elements of an idea, event, or theory such as examining a particular case or situation in depth and considering its aspects	22.6	35.5	35.5	6.5	0.0	0.0
Synthesizing and organizing ideas, information, or experiences into new, more complex interpretations and relationships	16.1	35.5	32.3	12.9	3.2	0.0
Making judgments concerning the value of information, arguments, or methods such as investigating how others collected and interpreted data and evaluating the accuracy of their conclusions	19.4	29.0	41.9	6.5	3.2	0.0
Applying theories or concepts to the solution of practical problems or in new situations	19.4	22.6	32.3	12.9	3.2	9.7
Did the bioinformatics recitation modules (lectures) seem well organized?	16.1	35.5	25.8	19.4	3.2	0.0
Did the bioinformatics recitation modules (lectures) complement each other?	9.7	22.6	38.7	19.4	9.7	0.0
Did the bioinformatics recitation modules (lectures) complement your genetics lecture?	3.2	16.1	25.8	29.0	25.8	0.0

Values are percentages.

course, and adding additional work and material did not help the situation.

DISCUSSION

In this course, we used a bioinformatics component to supplement the undergraduate genetics lecture. As far as we are aware, this was the first time bioinformatics has been used to help teach undergraduate genetics, particularly at an HBCU. Based on student responses, the addition of a bioinformatics component improved their computer usage and skills and their understanding of genetics/molecular biology. Although students responded that they enjoyed, were mentally engaged, and appreciated the value of the bioinformatics supplement, the additional work load was perceived as too much without receiving any additional course credit.

Achieving Our Three Objectives

Our objectives for this additional component were threefold: 1) to provide students at CAU exposure and introductory training in bioinformatics; 2) to provide the visual learners with a visual component to the Genetics course; and 3) to allow CAU to assess the need for incorporating bioinformatics at various levels into its curriculum. We achieved each of these goals to varying degrees.

Clearly, the students were mentally engaged (Table 2), increased their confidence level in their own skills (Figure 2), and increased their learning of genetics or molecular biology. From never using PubMed before this course to finding and examining the crystallographic or NMR structures of eIF4A (a cancer therapy target) on the final exam, the level and rate of development the students displayed was quite impressive. However, by the end of the semester, all of the students thought the bioinformatics component should have been a separate class (Table 4), and they reported that adding the bioinformatics component to the recitation period increased the homework burden tremendously. The students thought the recitation period should have been spent directly addressing questions about concepts presented in the genetics lecture. Nonetheless, 70% of the students would recommend the course to a peer, suggesting they would enroll in an independent class in bioinformatics (Table 4) or would conceivably consider a career in bioinformatics or biomedical research. However, <40% thought the bioinformatics component was valuable to their career or educational objectives (Table 3) or informed them of potential graduate or professional opportunities (Table 1). Clearly, we needed to do a better job of communicating to the students the big picture and potential career outcomes of their efforts. Future courses should place great emphasis on this objective.

Over the years, the genetics instructor has observed that the majority of our genetics students were strong visual learners (K. S. Kimbro, personal communication). The results of the students' ILS preferences (Figure 1) support these observations, as do other studies on learning styles in African-American students (Shade, 1992). The large lecture hall used to teach genetics (and most other core biology classes), however, did not possess any LCD projectors or even overhead projectors. At the time this supplement was offered, multimedia presentation equipment was not available in the biology and chemistry lecture halls. In the following academic year (2004–2005), with funds from the FIRST Program, multimedia equipment, LCD projectors, and the previously discussed wireless network were installed in the biology and chemistry lecture halls. These additions will help the faculty at CAU to present information in multiple formats to better accommodate the diverse learning styles of their stu-

Table 3. Postcourse survey—course value

	Highly valuable	Quite valuable	Moderately valuable	Somewhat valuable	Not valuable	Don't know	NA
How valuable was the material you learned in this course to your career goals?	16.1	22.6	9.7	38.7	9.7	0.0	3.2
How valuable was the material you learned in this course to your educational goals?	9.7	29.0	25.8	25.8	6.5	0.0	3.2

Values are percentages.

Table 4. Postcourse survey results on course need

	Yes	No	NA
Should bioinformatics recitation be taught as an independent course?	100	0	0
If there was an independent Bioinformatics course, would you recommend it to someone?	71	29	0
	Genetics	Bioinformatics	NA
Would the recitation time period have been better spent on genetics or bioinformatics?	71.0	25.8	3.2

Values are percentages.

dents. Our data demonstrate that presenting information to students in congruence with their preferred learning style enhances their learning and experience.

The third goal was to assess the feasibility of incorporating bioinformatics into the biology curriculum at CAU. The fact that the student confidence level in using computers to study genetics was doubled after one semester and the fact that they expressed interest in taking an independent Bioinformatics course makes a strong case for incorporating bioinformatics in the undergraduate biology curriculum.

We also demonstrated the capability to teach an interactive, case-based, Web-based course at CAU, a first-time effort at CAU. By installing a WLAN, any course taught within the science building at CAU can now incorporate Web-based technologies into the course. Other faculty within the Biology Department have already incorporated some of the NCBI databases and tools into their courses. Finally, we submitted a comprehensive application in support of incorporating bioinformatics into the biology curriculum. Evaluations obtained from students about the course, course material, and the pedagogical techniques we used were incorporated into the application. The application is currently pending.

Student Exposure to a Variety of Teaching Styles

During the bioinformatics recitations, the teaching approach was more of a "guide on the side" versus a "sage on a stage." Students came into class, checked out a wireless laptop, broke up into small groups, and began to work on the in-class, case-based exercises previously posted on the course Web site. Halfway through class, the instructor would go through the case and cover any additionally relevant materials such as explaining theory (e.g., "how are foods genetically modified?"), equipment (e.g., "what is a flow cytometry?"), or terms (e.g., "what is a microarray?"). Then, the instructor would cover additional topics such as previous homework solutions, test results, questions from the genetics lecture, etc. This "less formal" format was in direct contrast to the predominant lecture style of teaching in the genetics lecture and, in general, in the department. This unfamiliar teaching style took the students several lectures to adapt. By the end of the course, they were quite comfortable with interactive, hands-on, case-based learning methods and actually preferred it to the lecture style format, based on responses to opened-ended questions in the post-course evaluations.

Three FIRST Fellows taught the bioinformatics modules while an assistant professor taught the genetics lecture. Although the FIRST Fellows were well trained in teaching pedagogy (Holtzclaw *et al.*, 2005), they were still "rookies," teaching a new subject area with new equipment, using a different pedagogical approach than the one to which students were accustomed. This would have been a daunting task for any seasoned instructor, let alone a new instructor. Furthermore, each FIRST Fellow had his or her unique teaching style. Future endeavors should consider: 1) only one instructor who teaches both the genetics lecture and the bioinformatics component; or 2) two instructors, both of whom teach sections of the genetics lecture and bioinformatics component. Such approaches would improve course organization and would better integrate the bioinformatics modules and genetics lectures.

Social Interactions and Restrictions on Access at MSIs

Although we were surprised by the high level of student interaction catalyzed by the bioinformatics recitations and course Web site, these findings are supported by previous studies showing African-American students tend to learn best in highly social settings with materials that had a human or social content or in situations guided by a teacher in cooperation with other learners (Shade, 1992; Willis, 1992). These studies help explain the success of our case-based bioinformatics modules using cases of social relevance to our students, such as sickle cell anemia, diabetes, and breast cancer.

Often, instructors at some MSIs have few resources to instruct their students. In particular, information technology (IT) is an area of concern. Most students at MSIs have limited access to computer labs. Although our biology students did have their own computer lab, they had to bring their own paper if they wanted to print anything. Our students also lacked technical support and evening hours, when students most need these facilities. Furthermore, access to scientific journals can be limited, forcing instructors to bring copies of journal articles with them to class that they acquired at neighboring institutions. Given these realities, the level of instruction that occurs at most MSIs is impressive and a testament to the dedication and creativity of the instructors in these often financially strapped environments. By implementing bioinformatics, we had hoped to provide the students an opportunity for greater access to information. Previously, all of their learning had come through lectures and textbooks. By using the NCBI Web site, we exposed students to new, more current sources of information, online databases, and tools, which they were able to access on their own at any time. Another advantage of using bioinformatics to teach genetics is that it allows the interactive viewing of three-dimensional structures and other visual aspects through the use of different programs or Internet tools such as Map Viewer or Cn3D. By demonstrating direct application of the concepts presented, significant increases in learning occurred.

Potential Pitfalls and Suggestions for the Next Course Iteration

The relationship between the Genetics course material and the bioinformatics component is not intuitive. Modules and lectures should be more explicitly linked by chapter and concept, and such links should be outlined in a single syllabus that combines both course elements. Each module should have referenced a specific chapter in the genetics text. The modules should have included the same genes discussed in the genetics lecture, serving as a follow-up or continuation of the class discussions. Likewise, the "database of the week," should have been mentioned or integrated into the genetics lecture, emphasizing the importance of these skills in studying genetics. Also, there are some chapters in the genetics textbook (Russell, 1997) that refer to NCBI databases. When these chapters were covered, we

should have spent some class time examining the Web sites or organisms online.

One of our initial concerns was student access to computers both on campus during the day and at home. However, 90% of students in the precourse survey stated that they had access to computers outside of CAU, and 87.5% said that they had access to a printer (our unpublished data). Students also needed to have access to broadband Internet connections. This is particularly important for institutions, like CAU, whose computer labs and other facilities close at night. Often students, whose only Internet access was through "dial-up" connections or by "dialing into" the university network, complained about slow connections, disconnects, and lost data.

Another concern is the institution's IT support staff. As previously suggested (Honts, 2003), it is important that the instructor be directly involved with the set-up and installation of the computer facilities, which may or may not be in agreement with IT policy. In our case, the university's IT staff installed the WLAN, but would not install necessary software components on individual laptops because the laptops did not belong to CAU (the laptops were bought by funds through the FIRST Program at Emory University, and therefore, technically were the property of Emory University). Hence, it was up to the instructors to install software and set up the correct configurations on each laptop, a very laborious and time-consuming process. Furthermore, the instructors spent a fair amount of time maintaining and troubleshooting the computers. Hence, it is important that the instructors are themselves computer competent.

In the bioinformatics recitations, homework assignments were submitted via e-mail. This became problematic because, according to students, CAU did not have a reliable university-wide e-mail system due to size restrictions on e-mail accounts. So, most CAU students used commercial online e-mail accounts (e.g., Yahoo, Hotmail, etc.) for their personal and academic needs. This led to minor issues such as identifying who submitted the assignments, and the instructors' e-mail accounts receiving viruses, SPAM, and other nuisances from students' accounts. This also led to major issues such as cheating. Because all the homework assignments were electronic (usually Microsoft Word documents) and were submitted via e-mail, it was very easy for students to get homework solutions from their peers. This became evident in the first homework assignment. Question 3e of homework #1 reads "How many genes are on chromosome 1? How many genes are on chromosome Y? (10 points)?" The exact same homework was given during the pilot Bioinformatics course in the spring semester of 2003. During January 2003, the answer was "3044 genes on chromosome 1 and 215 genes on chromosome Y." However, due to the evolving nature of the NCBI databases, which are updated daily, when we gave this question in January 2004, the correct answer was "2475 for chromosome 1 and 247 for chromosome Y." Of the 29 homework assignments submitted in 2004, 19 (65%) had the correct 2003 answers.

This was further investigated by going into Microsoft Word and selecting "options" from the "tools" menu. Then, click on "user information" or "track changes." Depending on which version of Microsoft Word one has and how much effort the students put into reformatting the copied Microsoft Word homework document, one can discover who copied the homework from whom. Another trick is to view the e-mail headers. The e-mail header is a series of lines at the top of an e-mail containing information such as who sent it and from what computer, also called the IP (Internet Protocol) address. Most e-mail programs hide the headers for appearance, but typically viewing the headers is just a matter of clicking on the right key. Therefore, you can discover that student "Tom" (IP Address 255.0.201.556) submitted an e-mail that was originally from "Jerry" (IP Address 255.0.201.742), or Tom got his homework from Jerry. However, if multiple students used the university computer labs, then they may all have the same sets of IP addresses. Also, many Internet service providers (ISP) use "floating IP Addresses." For example, BellSouth may only buy 1000 IP addresses for 10,000 customers because they never have more than 1000 customers online at any given time. Hence, each time you logon, you may get a different IP address. So, a given student may have multiple IP addresses depending on his/her ISP.

We then shared our findings with the class. We demonstrated, in class using two submitted assignments without the students' names, the previously described methods for tracing homework. This was followed by a discussion about the merits of academic honesty. Through this "shock and awe" technique, we believed we decreased the amount of cheating short term. However, it is unclear that cheating was completely eliminated. The issue of academic dishonesty, especially in the digital age, is a concern for future courses.

Although it required significant time and effort to maintain, the course Web site was an important addition and clearly made the course more successful from the student perspective. The inclusion of online discussions, chat rooms, etc., perhaps through an online course management system, would have been even better. Future Bioinformatics course instructors should investigate the use of WebCT, Blackboard, and other such software packages. We would also recommend some prerequisite in calculus or computer science. Most of our students lacked any computer programming experience, which does limit what can be accomplished in student learning from the bioinformatics side. For example, course content regarding the algorithms used to build and search the NCBI databases should be discussed. These algorithms make several assumptions and decisions during queries, which are not always correct. A discussion on how one can safeguard against potential false-positive results is an important course topic.

Finally, we would recommend future instructors carefully consider the additional student workload of incorporating a bioinformatics component into an already tough Genetics course. Studying bioinformatics takes time. First, it requires significant time up front of both the students and the instructor in learning how to navigate the databases, and use various software packages and tools. Second, due to the very nature of bioinformatics, it requires enormous time spent mining, surfing, and modeling data. Students and instructors can spend hours in front of the computer and easily get lost in the sheer volume of information. Assignments need to be clear, specific, and detailed. It could be ideally suited for an honors class. Third, it requires constantly updating assignments in response to constantly updated databases.

CONCLUSIONS

This article details the outcomes and lessons learned from incorporating a bioinformatics component into an undergraduate Genetics course at a historically black college. Our data show that the bioinformatics supplement strengthens the Genetics course as well as exposes students to new information, technologies, and pedagogy. By implementing the simple suggestions previously presented, future endeavors should be even more successful.

Future research on the effectiveness of a Bioinformatics course or series of modules should include student demographic attributes and student performance information on the various assignments. For example, the inclusion of student attributes such as gender would permit an examination of learning style preferences or previous experience with IT and subsequent perceptions of course effectiveness and course performance separately for the men and women enrolled. This, in turn, could enable course instructors to better target their information presentation and case studies to the students and potentially further enhance student learning.

ACKNOWLEDGMENTS

The authors thank Pat Marsteller of Emory University for her advice and feedback in preparing this course. The authors also thank their students for participating in this course experiment. K.S.K. is now at Emory University, School of Medicine. This work was supported by an IRACDA Grant 3K12 GM 00680-03S1 to Emory University School of Medicine.

REFERENCES

Astin, A. W. (1984). Student involvement: a developmental theory for higher education. J. College Student Personnel *25*, 297–308.

Bloom, B. S., and Krathwohl, D. R. (1984). Taxonomy of educational objectives, Handbook 1, Cognitive domain, New York: Addison-Wesley.

Campbell, A. M. (2003). Public access for teaching genomics, proteomics, and bioinformatics. Cell Biol. Educ. *2*, 98–111.

Chickering, A. W., and Gamson, Z. F. (1991). Applying the Seven Principles of Good Practice in Undergraduate Education, San Francisco: Jossey-Bass.

Eisenberg, D. (2002). The coming job boom. Time *159*, 58–61.

Felder, R. M., and Silverman, L. K. (1988). Learning and teaching styles in engineering education. Engineering Educ. *78*, 674–681.

Herreid, C. F. (1994). Case studies in science: a novel method of science education. J. College Sci. Teaching *23*, 221–229.

Holtzclaw, J. D. (2004). Bioinformatics (BIO 312R) Course Web site. http://userwww.service.emory.edu/~jholtzc/Courses/Bio312/index.htm. (accessed 21 February 2005).

Holtzclaw, J. D., *et al.* (2005). FIRST: a new model for training future science faculty. J. College Sci. Teaching *34*, 24–29.

Honts, J. E. (2003). Evolving strategies for the incorporation of bioinformatics within the undergraduate cell biology curriculum. Cell Biol. Educ. *2*, 233–247.

Lipshutz, R. J., Fodor, S. P., Gingeras, T. R., and Lockhart, D. J. (1999). High density synthetic oligonucleotide arrays. Nat. Genet. *21*, 20–24.

National Institute of General Medical Sciences (2002). Institutional Research and Academic Career Development Award. http://grants.nih.gov/grants/guide/pa-files/PAR-02-152.html (accessed 21 February 2005).

Russell, P. (1997). Genetics, Reading, MA: Addison-Wesley.

Shade, B. (1992). Is There an Afro-American Cognitive Style? An Exploratory Study. Newbury Park: Sage Publications.

Suitts, S. (2003). Fueling education reform: historically black colleges are meeting a national science imperative. Cell Biol. Educ. *2*, 205–206.

Virginia Commonwealth University Center for Public Policy (2001). VCU Survey on Bioinformatics and Biotechnology Employment, Richmond, VA: Virginia Commonwealth University Press.

Willis, M. (1992). Learning Styles of African American Children: A Review of the Literature and Interventions, Newbury Park: Sage Publications.

Zywno, M. S. (2003). A contribution to validation of score meaning for Felder-Soloman's index of learning styles. American Society for Engineering Education Annual Conference and Exposition, American Society of Engineering Education.

Appendix

Module Example: Module 2 Sickle Cell Disease

In-Class Assignment:

1. Search the NCBI database for sickle cell anemia. What can you tell me about it?
2. What causes beta-zero-thalassemia? _____
3. Sickle hemoglobin is a "point mutation" meaning only one nucleic acid is exchanged resulting in sickle cell anemia. Which nucleic acid is exchanged? _____ in normal hemoglobin is replaced with _____ is sickle hemoglobin.
4. How does this single nucleic acid exchange cause such a deadly disease?
5. At which base pair position does the sickle hemoglobin mutation occur? _____

Case Project (In-class and for homework)

Case Study: Mixed Signals (50 points).

You and your lab are researching different gene therapy treatments in hopes of finding a cure for sickle cell anemia. To test your therapies, you have been breeding in your laboratory a strain of transgenic mice, which exclusively express human sickle hemoglobin. These homozygous, transgenic mice contain no murine β-globin genes, just human β-globin genes, as shown in the electrophoresis gel in Figure A1. This colony has been carefully controlled and a hemoglobin gel is run on each mouse to ensure that they are in fact homozygotes (i.e., express only human sickle hemoglobin).

In order to determine if your gene therapies are having a negative effect on other genes, you decide to run MURINE cDNA microarrays on each mouse before and after they received a gene therapy treatment. These cDNA microarrays contain probes consisting of 50 bp, oligonucleotide segments from within a gene to help you determine whether or not a

Table A1. Genetics course outline and schedule

January	15	Introduction, Review Chapter 1
	20	Chapter 2, (DNA Genetics Material)
	22	Chapter 2,3 (DNA Replication)
	27	Chapter 5 (Gene Expression: Transcription)
	29	Chapter 5,6 (Gene Expression: Translation)
February	3	Chapter 6 (Gene Expression: Translation)
	5	Chapter 7 (DNA recombinant Technology
	10	Exam 1
	12	Chapter 9 (Applications of Rec. DNA Tech)
	17	Chapter 9 (Genenome Analysis)
	19	Chapter 10 (Mendelian Genetics)
	24	Chapter 11, 12 (Chromosome basis of Inheritence)
	26	Chapter 12, 15 (Non-Mendelian Inheritance)
March	2	Chapter 15
	4	Exam 2
	9	Chapter 13 (Gene mapping in Eukaryotes)
	11	Chapter 14 (Gene mapping in Bacteria)
	16	Chapter 16 (Lac operon; gene reg. In bac.)
	18	Chapter 17 (Gene reg. In euk.)
	23	Exam 3
	25	Chapter 18 (Genetics of Cancer)
	30	Chapter 18 (Genetics of Cancer)
April	1	Chapter 19 (DNA mutation and Repair)
	6	Chapter 19 (DNA mutation and Repair)
	8	Chapter 20 (Transposable elements)
	13	Chapter21 (Chromosomal Mutations
	15	Exam 4
	20	Chapter 22 (Population Genetics)
	22	Chapter 24 (Molecular Phylogeny)
	27	Review
	29	Final

Table A2. Bioinformatics outline and schedule

Date		Topic
January	27	Introduction to Bioinformatics
February	3	Module 1: Introduction NCBI databases
	10	Module 2: Sickle cell disease
	17	Module 3:Mitochrondia DNA
	24	Module 4: Drug resistant bacteria "fast food conundrum"
March	2	Drug resistant bacteria (continued)
	9	Spring break
	16	Module 5: Diabetes
	23	Diabetes (continued)
	30	Module 7: The Human Genome Project
April	6	Module 8: The Protein Data Bank
	13	Module 9: SNP Database
	20	Module 10: Gene Therapy
	27	Review, evaluations

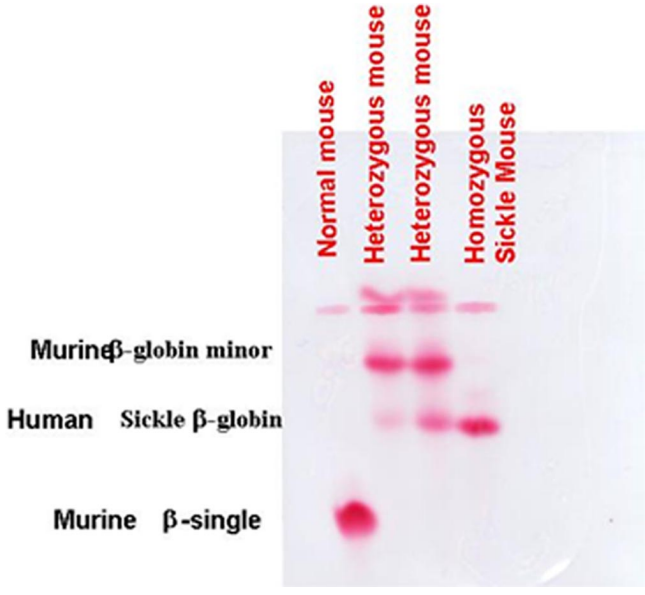

Figure A1. Hemoglobin gel of transgenic mice.

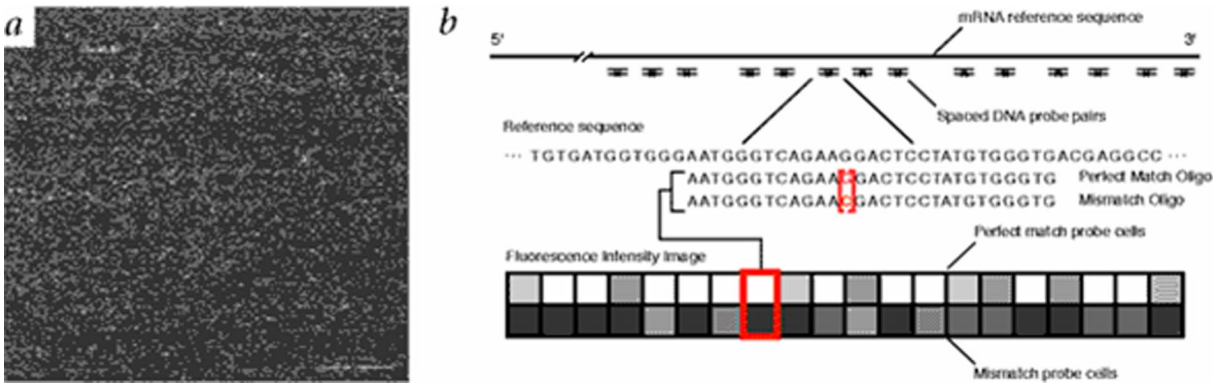

Figure A2. From Lipshutz *et al.* (1999), Figure 2 of publish-ahead-of-print.

particular gene is being expressed (Figure A2). This is repeated for every gene in the murine genome, giving you the ability to determine the effect of the gene therapy treatment on gene expression of every gene.

You have been working on a new gene therapy treatment that you are convinced will revert your transgenic HUMAN sickle, β-globin mice back to normal, MURINE, β-globin mice. One day, one of your students comes into your office looking confused. "When I run the cDNA microarrays on the mice that received the new gene therapy treatment, the cDNA microarrays show that the mice are now expressing murine β-globin. However, when I run a hemoglobin gel, it shows that the mice still express human sickle β-globin. I don't understand what's going on."

A. So does your gene therapy treatment work or not? How can you prove this (i.e., support your answer with reason or data)? (10 points)

B. What caused the mixed signal? (15 points)

C. How do you redesign the experiment so that this does not happen again? (25 points) HINT: Give a specific 50 base pair oligonucleotide segments to include on your microarray that would prevent this from happening again.

Hint: this is a bioinformatics class not a gel electrophoresis class!