Phylogenomic-Based Case Study: An Undergraduate Research Pedagogy at Kentucky State University, the Only Historically Black College and University in the Commonwealth of Kentucky

Narayanan Rajendran, Kentucky State University
Williard Mazhawidza, Kentucky State University

ABSTRACT

Currently the undergraduate biology curriculum at 2 and 4-year colleges especially in Historically Black Colleges and Universities (HBCUs) is undergoing a facelift by implementing a range of recommendations of committees on education. This has produced a demand to educate and train the current and future generation of undergraduates as scientists capable to perform modern biological research that is heavily dependent on genomic training, such as phylogenomics. It is our observation that the HBCUs may not have a strong infrastructure enough to meet this challenge without the development of vigorous research-based learning programs to complement the traditional lecture-only based configuration. In this case study, the first author undertook the objective to facilitate an undergraduate genomic research-pedagogy at the biology department in Kentucky State University, the only HBCU in the Commonwealth of Kentucky. The wet-lab activities in this NSF funded research project include molecular microbiology training in culturing bacteria, isolating DNA, performing PCR, cloning amplicons and analyzing the sequence data. The associated case-study subject was assessed in gaining of wet-lab skills, knowledge, application of protocols, analysis of results and impact on career choice. The results reveal an enhanced student confidence in handling of molecular techniques and positive admiration towards bioinformatics. It suggests that the wet-lab based research pedagogy could play a role in retention of students in biology. Although this case-study subject has appreciated the value of the phylogenomic approaches, a substantial and continual support is required to retain student interest especially institution like this HBCU.

Keywords: Genomic research-pedagogy, Case Study, Kentucky State University

Introduction

According to the "Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st century", (CUBE, 2003), it is recommended to undertake sweeping changes in biology curriculum and education at undergraduate level especially introducing computer based course improvement and curricular adaptation. The exploding advancement in bioinformatics forced many institutions to incorporate computing information across many discipline especially in biology (Campbell, 2003). Research-based phylogenetic learning, which incorporates genomic studies, is more effective (Suchman, et al., 2001) in biology especially in microbiological sciences.

In recent years, genomic information on microbes has tremendously increased since more and more students were involved in the genomic studies (Chan, 2005; Campbell, 2003). Though microbial genomic analyses play a major role in the genomic revolution, many existing hypothetical sequences are not yet annotated. A couple of years ago, it was reported that only 3 % of the sequences are not labeled as hypothetical, which means the rest of those floating sequences are still to be annotated (Brown & Sjolander, 2006). There may be multiple reasons, such as poor bioinformatics training, lack of knowledge or lack of adequate categorization of microbial groups (like medical microbes, agriculture microbes), or untapped bacterial genomics etc. Besides it was reported

earlier that there is lack of sequence identifiers (Brown & Sjolander, 2006). No matter what blocks the flow it is a fact that there are vast hypothetical sequences need immediate attention. There fore, custom designed development of computer-based training activities in phylogenetics is essential and it should start at high school level (Marbach, Rotbain & Stavy, 2008).

Minority students' participation in annotation process particularly in phylogenomics using bioinformatics needs significant attention, especially in 2-and 4-year colleges. Even though as little as 3 % of the annotation was carried out using wet-labs about 97 % of sequences, which were annotated in the public databases, were done based on electronic evidences (Brown & Sjolander, 2006). Training undergraduate students in this direction using this emerging field with wet-lab activities helps in their learning process (Jurkowski, Reid & Labov, 2007) and supports their retention in Biology. However, the need to transform undergraduate learning is challenging (Labov, 2004) especially when considering dry and wet-lab training as it requires an in-depth radical approach both in trainee and trainer. It needs equipments, updated technology, and service-mined human resources. Lack of bioinformatics training at their undergraduate level in many colleges leaves a significant gap in student learning process even to understand the basics of a phylogenetic tree (Meir et al., 2007).

Some Historically Black Colleges and Universities (HBCUs) are lagging behind in this frontier-field due to lack of funding (Holtzclaw, et al., 2006) and dearth of infrastructures. As a result, the students are not well represented in computational fields of sciences especially in graduate studies. In some cases, the undergraduates are lagging behind even in basic biology education (Jurkowski, Reid & Labov, 2007). It is therefore important to train undergraduates, especially minority students, to improve their participation in phylogenomic-based studies.

Undergraduate research experience has been shown to have a positive effect on career choices, development of independent thinking and enhanced participation in the courses that followed after the research experience (Lopatto, 2004). Such research experience boosts their grade point average (GPA) higher and even helps them to pursue PhD degree in science and engineering. With this objective in mind, many attempts were made in the past to introduce research-oriented courses at the 4-year colleges and universities. Induction of bioinformatics into the undergraduate curriculum has been attempted elsewhere successfully as reported earlier (Campbell, 2003; Chan, 2005). However in the absence of such curriculum at Kentucky State University, which is a historically black, liberal arts and 1890 Land-Grant higher learning institute, an attempt was made to induce the research interest in the Phylogenomics with the support of a National Science Foundation grant. In this article, we are exposing the wet-lab methodology of phylogenomic inference of Glutamate Synthase of a new soil isolate *Arthrobacter nicotianae* in an attempt to bring out the research pedagogy by using a case study model towards the phylogenomics field.

Materials and Methods

Conventional Microbiology Activities
The approach was to assign a basic genomic-based research project to be carried out in the wet-lab for a semester period. The case-study subject (CSS) was an undergraduate minority student who has no prior knowledge in this field. The theory and experimental methods to be applied was instructed in the form of written protocol as well as in association with a competent

researcher. In this single-subject case study, newly isolated *Arthrobacter nicotianae* strain PR was used throughout the study. The identification of this soil bacterium had been performed and confirmed using 16sRNA with high resolution automated microbial identification techniques through rRNA gene sequences, hybridization probes and molecular profiles (Accugenix, DE). The CSS grew the bacterium in Nutrient Broth (10mL) at 30°C, by picking a single colony from a Nutrient Agar stock-plate prepared earlier using a streak-plate technique. Genomic DNA was isolated using DNAzol direct kit (Catalog # DN 131, Molecular Research Center, Inc. OH) by mixing 1-10uL or 1-10mg of solid sample with 0.1mL of DNAzol Direct. The lysate was thoroughly mixed and a 2-5µL aliquot transferred directly into 20-50µL of PCR mix.

PCR-based Molecular Microbiology Activities

Previously published probing primers (Horwood, Burgess & Oakey 2004) were synthesized at Integrated DNA Technologies, Coralville, IA, USA. They were employed against the genomic DNA of *Arthrobacter nicotianae* strain PR. The Mastercycler personal thermocycler (Eppendorf) was used to make the polymerase chain reaction (PCR) with a preparation of 2µL of DNA, 10mL of Premix Taq (catalog #RR003, Takara), 2µL of primer (forward and reverse), and 6uL of double distilled water in a 0.5mL microcentrifuge tube. The PCR conditions were based on the protocols that had been standardized earlier (Rajendran, 1999; Rajendran, Rajnarayanan, & Demuth, 2008). The following conditions: 95 ºC (5 min), 95 ºC (1 min), 55 ºC (1 min), and 72 ºC (3 min) for 30 cycles were adopted. The PCR amplicons were analyzed via 1% agarose gel electrophoresis using 100 bp DNA step ladder (catalog # G6951, Promega, Madison WI, USA). Before running electrophoresis, the gel was prepared by mixing 0.5g of Agarose with 50mL of TBE and 1µL of Ethidium bromide. The solution was then poured into the gel casting cassette (Thermo Scientific, USA). Loading dye (6µL) was added to each tube of the PCR products and loaded into the gel slots. The gel was run at 90V for about an hour. By using 'Digidoc-it UVP digital documentation unit', the profile of the amplicons were examined and documented. Individual PCR amplicon was gel extracted using PureLink Quick Gel Extraction kit (Catalog # K2100-12, Invitrogen, Carlsbad CA, USA). In brief, agarose gel with amplicon DNA was dissolved in gel solubilization solution (GS1) at 65ºC. The solution was passed through quick gel extraction column, which retained the DNA. Wash buffer (W9) was also added to the column before eluting with TE buffer.

Subcloning and DNA Sequencing

Amplicons obtained from PCR were subcloned in chemically competent *E. coli,* using manufacturer protocol (Invitrogen USA). PCR products were mixed with pCR4-TOPO vector followed by transformation into chemically competent one shot TOP10 *E. coli*. Positive clones were screened using IPTG/X-Gal (Fermentas, USA) with Ampicillin (Sigma, USA). Positive colonies were grown in Ampicillin/LB culture followed by plasmid isolation using Pure Yield Plasmid Miniprep System (Promega, USA). The clones were confirmed by *Eco*R1 restriction and resolved via 1% agarose gel electrophoresis using 100 bp DNA ladder. Positive clones were sequenced using a T3 or T7 primers at the Center for Genetics and Molecular Medicine (CGeMM) DNA core facility of the University of Louisville, KY, USA.

Sequence Analysis and 3D Analysis

A clean DNA sequence, free from vector sequence, was confirmed using NCBI vector contamination software (Vecscreen). Blast search for homologous nucleotide and/or amino acid sequences were performed using NCBI public database search software at http://www.ncbi.nlm.nih.gov/BLAST/. The determined sequence of Glutamate Synthase was deposited in GenBank and an accession number (FJ979920) was obtained using the BankIT:GenBank at www.ncbi.nlm.nih.gov/BankIT/ submission program. To confirm the specificity signature sequence, a conserved amino acid residue search was performed during multiple sequence alignment. Cn3D 4.1 (NCBI) software was used to determine the domain structure of the clone. The newly cloned sequence was compared to 3D domain model constructed from multiple sequence alignment of many peptides in the same class.

Results and Discussion

Phylogenomics begin with selection of homologs followed by sequence alignment, tree construction and structural annotation (Eisen, 1998). In this study, wet-lab based research pedagogy was designed to train a case-study subject (CSS), undergraduate student, to conduct similar phylogenomic research under the direct supervision of a mentor in association with a post doctoral fellow. Though it is time consuming, mentor-associated research pedagogy helped the CSS to focus on the objective. Based on our assessment of knowledge and skills learned in the laboratory training, the CSS performed the works beyond expectation and was able to demonstrate in-depth understanding of the theory and science behind the protocols that was employed. The CSS assessment, based on the graded activities and the gains from the research project, was expressed in the pie chart (Fig. 1).

Figure 1. Gaining of Wet-lab Skills and Course Work-Knowledge with Career Focus
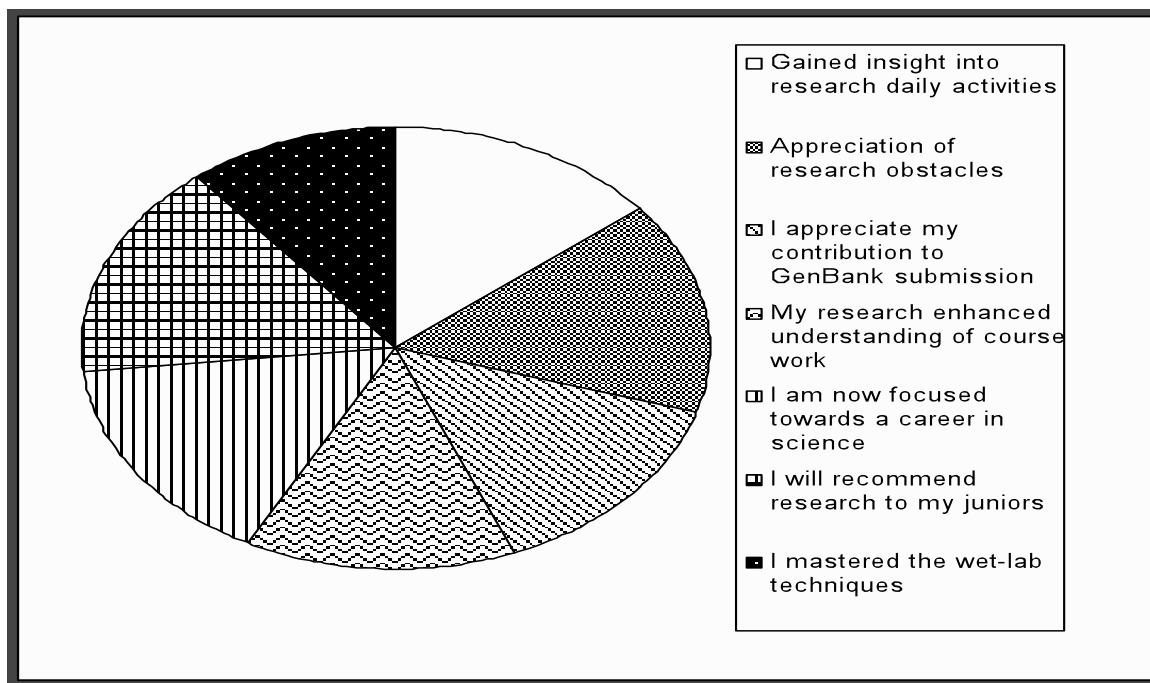
Following the completion of assigned research tasks, the CSS was evaluated with respect to his acquired skill and potential focus to a career in science. The CSS graded the benefit of the research experience on a 0-10 scale. As expressed in the pie chart, the outcome is evenly distributed on all seven assessment approaches except the mastering of the wet-lab technique. This indicates the enthusiastic involvement of the CSS in research activities but required more semester periods to master in the phylogenomic techniques

  At the end of the training, the CSS was asked to take a sequential-approach analysis in order to reveal their mastering of the subject matter. Table 1 represents an unscrambled protocol sequence used in this study. However, a scrambled protocol sequence was given to the CSS and instructed to make the perfect sequential order. Each correct step gets 5 points for a total of 100. As shown in the pie chart and in the table, the training experience of the CSS indeed focused on the core science,

hence it could be a good working model of operation for 2 or 4-year old colleges which only could afford to have a few students to do wet-lab research due to the space constrains and limited research budget. After this study, a few more students were came forward and carried out their research training. Teaching an emerging genomic field especially at HBCU can be challenging, however a brick-at-a time approach (Jurkowski, Reid & Labov. 2007) is ideal to engage undergraduates in this exploration.

  As a part of assessment, following the completion of the research project, the CSS overall performance was assessed in five different categories as shown in Figure 2. This includes theoretical understanding of experimental methods, their applications and wet-lab skills. The grade for each category was scored and measured as a percentage basis. The results were analyzed and presented graphically which is demonstrating high dexterity for bench top skills.

Table 1. Protocol-based Assessment on Wet-lab Learning.

Given an unfamiliar bacterium (*Arthrobacter nicotianae* strain PR) how would you find the sequence of Glutamate synthase?* Please arrange the randomized protocol steps in proper sequential order.

Blind-folded	From the unfamiliar bacterium, how would you find gene of interest?	Arrange them		Total points
1.	Get single colony of the given bacterium using streak-plate method		√	
2.	Grow the single colony in nutrient broth in culture tubes		×	
3.	Transfer in flask, and culture in shaker at 200rpm for overnight.			
4.	Isolate the genomic DNA by lysing the bacteria with 'DNAzol direct'			
5.	Use a PCR ready-made premix plus DNA Taq, and primers (F & R)			
6.	Place PCR tube in a thermocycler and run the PCR with 30 cycles.			
7.	Prepare an electrophoresis gel and load PCR products			

8.	Cut out the gel containing the PCR amplicons, and extract DNA			
9.	Wash with wash buffer and elute DNA with TE buffer.			
10.	Subclone the selected amplicons into a TA vector			
11.	Transform the vector to competent *E. coli* cells			
12.	Screen and grow positive clones using IPTG/X-Gal and Ampicillin			
13.	Isolate plasmid DNA using Pure Yield Plasmid Miniprep (Invitrogen)			
14.	Confirm the clones by E*co*R1 restriction in 1% electrophoresis			
15.	Send positive clones for sequencing using T3 or T7 primers			
16.	Perform DNA sequence analysis on NCBI- blast search			
17.	Search for homologous gene sequences or peptides			
18.	Find the electronic evidence of the gene Glutamate synthase			
19.	Perform multiple sequence alignment to demonstrate homology			
20.	Annotate and submit newly cloned sequence to the GenBank.			
Total score of the wet-lab assessment after completing the protocol steps =				**100**

Figure 2. Significant Gain of Skills and Knowledge Demonstrated by the Student.
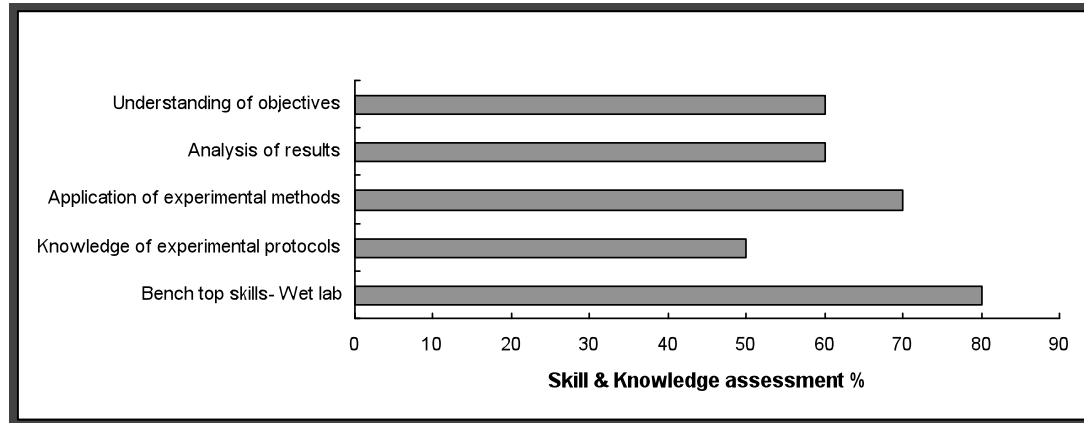
As shown in figure 2, the CSS performed well with wet-lab skills, but showed a limited expansion of the knowledge on the scientific concepts behind the experimental protocols that were applied. As reported earlier (Holtzclaw et al., 2006), phylogenomic component increased the student computer usage in our study also. Despite limited stellar performances in theoretical aspects, it is our assessment that the CSS performed above the average after taking into account with the limited amount of time spent on the project as well as the level of scientific material the CSS was exposed to in other critical area of Biology. This process and the methodology used to bring out the research capability of an undergraduate student towards the phylogenomics were

well received.   As demonstrated earlier at the University of Arizona (Bender, Ward & Wells, 1994), such approach of inclusion of undergraduate students in research laboratories to conduct phylogenomic research has potential to succeed in our HBCU and also had a positive impact both on the student and the instructor.

Since the CSS completed the General Microbiology course prior to begin the research, the level of fundamental knowledge on the theory and the laboratory exercises helped this CSS to learn faster Further training along with a post-doctoral fellow on further analyze of the newly cloned sequence using Cn3D software to determine a putative 3-dimensional structure of the partial sequence enhanced the student learning. As predicted earlier (Jurkowski, Reid & Labov, 2007) the technical advancement with conventional microbiology training and efficient data management helped the CSS to raise his curiosity towards the field and retain the interest until the end. This approach reiterated that research training using advance techniques in an undergraduate laboratory course certainly helps undergraduates to develop an interest towards the cutting edge scientific research and provides an opportunity to stay in research field.

In recent years, serious efforts are being stipulated to address the lack of minority student participation in genomic sciences since many students are not showing much interest to learn the cutting-edge discoveries (Jurkowski, Reid & Labov, 2007). Many agencies including the National Science Foundation and other organizations such as the Genome Consortium for Active Teaching (GCAT) seeks to incorporate the use of modern research methods into undergraduate teaching (Campbell et al., 2007) by providing funding and other supports.

during the wet-lab training session, while growing bacteria in specific media, isolated genomic DNA, performed PCR and resolved PCR products using agarose gel electrophoresis. The plasmids were sequenced, and the sequences were analyzed after cloning the amplicons. It culminated in the revelation of the putative identity of the sequence as Glutamate Synthase by NCBI blast search of microbe genome protein databases and submission of the sequence to the GenBank (accession number FJ979920).

Specifically universities and industries are trying to incorporate students into bioinformatics process of analyzing genetic information using the freely available software and involve them in wet-lab activities. The availability of free software from NCBI and supports from funding agencies such as NSF stimulates the capability of teaching faculty members at small undergraduate universities to steer their students towards careers in this cutting-edge science.

Programs on enriching minority student's learning were existed before (Carline et al., 1998). However, they were not integrated as part of the core learning activities. Hence the effect of their intended purpose has not been achieved or realized in 2 and 4-year colleges and universities. Socioeconomic factors, demographic changes, and technology explosion that challenge public education were blamed for the ineffective integration (Castro et al., 1999; Davidson & Lewis 1997). Institutions of higher learning are subjected to scrutiny since they admit students using race as part of the criteria for admission (Davidson & Lewis, 1997). However, there are reports indicate that the performance of such under-represented minority students who came from those demographic conditions and/or who entered the college with considerable educational disadvantage have the capacity

to perform moderately well if they were challenged with the wet-lab based activities. For example, students who came with significant educational disadvantage into College of Medicine University of Arizona, continued to pass with lower scores on their paper and pencil examinations despite having equal scores on the clinical examinations compared to other students (Campos-Outcalt et al., 1994). These reports reiterated the need to involve the minority students in more learning activities adjunct with wet-lab activities, since it closes the learning gap.

Conclusion

Lack of participation of minority students in phylogenomics is a challenge that institutions of higher learning have to deal with. Despite the persistence of the factors that facilitate the educational and professional gap against minorities in small colleges and universities, vigorous learning activities employing such research-pedagogy should help in closing that inequality. A substantial and continual care to enhance the interest of students to stay is required constantly especially in minority institution such as this HBCU. Research programs that have been established towards the effort to improve academic performance should be enhanced. This case study reiterates the needs for such incorporation of research-based pedagogy in phylogenomic studies at a historically black colleges and universities. It is therefore important to implant the seeds of interest at an early undergraduate level to see an improved outcome of students on the participation in modern genomic studies in near the future.

References

Bender, C., Ward, S., & Wells, M. A. (1994). Improving undergraduate biology education in a large research university. *Molecular Biology of the Cell, 5*, 129-134.

Brown, D., & Sjolander, K (2006). Functional classification using phylogenomic inference. PLoS *Computational Biology, 2,* e77.

Campbell, A. M. (2003). Public access for teaching genomics, proteomics, and bioinformatics. *Cell Biology Education*, *2*, 98-111.

Campbell, A. M., Ledbetter, M. L., Hoopes, L. L., Eckdahl, T. T., Heyer, L. J., Rosenwald, A., Fowlks, E., Tonidandel, S., Bucholtz, B., & Gottfried, G. (2007). Genome Consortium for Active Teaching: meeting the goals of BIO 2010. CBE *Life Science Education, 6,* 109-118.

Campos-Outcalt, D., Rutala P. J., Witzke, D. B., & Fulginiti J. V. (1994). Performances of underrepresented-minority students at the University of Arizona College of Medicine, 1987-1991. *Academic Medicine, 69,* 577-582.

Carline, J. D., Patterson, D. G., Davis, L.A., & Oakes-Borremo P. (1998). Enrichment programs for undergraduate college students intended to increase the representation of minorities in medicine. *Academic Medicine, 73*, 299-312.

Castro, G. A., Bouldin, P. A., Farver, D.W., Maugans, L.A., Sanders, L.C., & Booker, J. (1999). The InterCon network: a program for education partnerships at the University of Texas-Houston Health Science Center. *Academic Medicine, 74***,** 363-365.

Chan, E. Y. (2005). Advances in sequencing technology. *Mutation Research, 573***,** 13-40

CUBE. (2003). Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century, National Research Council. *BIO2010: Transforming Undergraduate Education for Future Research Biologists*. National Academies. Washington, DC.

Davidson, R. C., & Lewis E. L. (1997). Affirmative action and other special consideration admissions at the University of California, Davis, School of Medicine. *Journal of the American Medical Association, 278***,** 1153-1158.

Eisen, J. A. (1998). Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Research, 8***,** 163-167.

Holtzclaw, J. D., Eisen, A., Whitney, E. M., Penumetcha, M., Hoey, J. J., & Kimbro. K.S. (2006). Incorporating a new bioinformatics component into genetics at a historically black college: outcomes and lessons. CBE *Life Science Education***,** *5,* 52-64.

Horwood, P. F., Burgess, G.W., & Oakey. H. J. (2004). Evidence for non-ribosomal peptide synthetase production of cereulide (the emetic toxin) in *Bacillus cereus*. FEMS *Microbiology Letter, 236***,** 319-324

Jurkowski, A., Reid, A. H., & Labov**.** J. B. (2007). Metagenomics: a call for bringing a new science into the classroom (while it's still new). CBE *Life Science Education, 6***,** 260-265**.**

Labov, J. B**.** (2004). From the National Academies: the challenges and opportunities for improving undergraduate science education through introductory courses. *Cell Biology Education, 3,* 212-214.

Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): First findings. *Cell Biology Education* 3**,** 270-277.

Marbach, A. D., Rotbain, G.Y., & Stavy, R. (2008). Using computer animation and illustration activities to improve high school students' achievement in molecular genetics. *Journal of Research Science in Teaching. 45*, 273-292.

Meir, E., Perry, J., Herron, C., & Kingsolver, J. (2007). College students'misconceptions about evolutionary trees. *American Biology Teacher, 69*, 71-76.

Rajendran, N. (1999). Identification and cloning of a gene locus encoding peptide synthetase of *Pseudomonas fluorescens* by two sets of PCR primers. *Zeitschrift fur Naturforschung C, 54***,** 105-109.

Rajendran, N., Rajnarayanan, R. V., & Demuth. D. R. (2008). Molecular phylogenetic analysis of tryptophanyl-tRNA synthetase of *Actinobacillus actinomycetemcomitans*. *Zeitschrift fur Naturforschung C*, *63***,** 418-428.

Suchman, E., Timpson, W., Linch, K., Ahermae, S., & Smith. R. (2001). Student responses to active learning strategies in a large lecture introductory microbiology course *Bioscene: Journal of College Biology Teaching*, *27*, 21-26.

Narayanan Rajendran is Associate Professor of Biology, Kentucky State University, Frankfort, Kentucky

Williard Mazhawidza is Postdoctoral Researcher, Kentucky State University, Frankfort, Kentucky