APPLICATION OF COURSE-BASED RESEARCH EXPERIENCE
IN UNDERGRADUATE PLANT PHYSIOLOGY LAB
AND ITS EFFECT ON LEARNING OUTCOMES

By

JIEYING WANG

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2017
To my respected mentor Bernard Hauser and advisory committee members McDaniel Stuart and Pavlo Antonenko
ACKNOWLEDGMENTS

I would first like to thank my mentor Bernard Hauser of the College of Liberal Arts and Sciences at University of Florida. Dr. Hauser is always there to offer guidance when needed, and he provides profound insights from time to time. He helped me through difficulties both in academics and in adapting to the life abroad as an international student.

I would also like to thank my committee members. Dr. McDaniel Stuart provided plenty of valuable advice on the setup of the experiment and evaluation questions. Dr. Pavlo Antonenko provided suggestions from the educational perspective. Without their dedicated participation and input, the thesis project could not have been conducted as smoothly.

I would also like to acknowledge administrative specialist Susan Spaulding. Her kindness and efficient work helped me greatly in adjusting to the life at the university and completing my degree.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>6</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>7</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 LITERATURE REVIEW</td>
<td>8</td>
</tr>
<tr>
<td>Learning Innovations</td>
<td>8</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>16</td>
</tr>
<tr>
<td>2 MANUSCRIPT</td>
<td>21</td>
</tr>
<tr>
<td>Abstract</td>
<td>25</td>
</tr>
<tr>
<td>Introduction</td>
<td>25</td>
</tr>
<tr>
<td>Results</td>
<td>28</td>
</tr>
<tr>
<td>Discussion</td>
<td>29</td>
</tr>
<tr>
<td>3 CONCLUSION</td>
<td>37</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>39</td>
</tr>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>44</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Flowchart of bioinformatics research project.</td>
<td>29</td>
</tr>
<tr>
<td>2-2</td>
<td>Pre-test and post-test scores in the multiple choice version</td>
<td>30</td>
</tr>
<tr>
<td>2-3</td>
<td>Combined pre-test and post-test scores in the multiple choice version for</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>year 2016 and 2017</td>
<td></td>
</tr>
<tr>
<td>2-4</td>
<td>Pre-test and post-test scores in the short answer version</td>
<td>31</td>
</tr>
<tr>
<td>2-5</td>
<td>Combined pre-test and post-test scores in the short answer version for year</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2016 and 2017</td>
<td></td>
</tr>
<tr>
<td>2-6</td>
<td>Pre-test and post-test scores from year 2017</td>
<td>29</td>
</tr>
</tbody>
</table>
APPLICATION OF COURSE-BASED RESEARCH EXPERIENCE IN UNDERGRADUATE PLANT PHYSIOLOGY LAB AND ITS EFFECT ON LEARNING OUTCOMES

By

Jieying Wang

August 2017

Chair: Bernard Hauser
Major: Botany

Instructional strategies evolve to improve learning outcomes. Active learning has emerged and are commonly believed to yield better learning results than the traditional lectures. More and more forms of active learning are being developed and tried, and course-based research experience (CURE) is one with great potential.

In order to see if CURE can indeed help students learn better, we introduced a research project into our plant physiology lab course. The students accomplished a bioinformatics researched project where they search online database, browsed primary literature and carried out experiments to test hypothesis about gene function under the guidance of the instructors. The lab integrated contents about proteomics, genetics, molecular biology, plant physiology and statistics. A pre/post test was used to evaluate student improvement on content knowledge.

Our results showed that students made significant improvement on most topics during the CURE, and we discussed future revisions for improving the design and evaluation of the CURE to further improve learning outcomes.
CHAPTER 1
LITERATURE REVIEW

What is learning? A knowledgeable person used to be defined as one who knows a lot of information. But with the rise of the internet searches, almost any information can be quickly found. In this setting, a knowledgeable person would be one who knows how to apply the information effectively to solve problems.

According to Bloom’s taxonomy of learning, there are three domains of learning goals, knowledge, skills and attitudes (KSA) and six levels of learning – remember, understand, apply, analyze, evaluate, and create (Bloom, 1956). To remember and to understand are considered lower levels in basic learning, while application, analysis, evaluation and creation are higher levels of advanced learning. In other words, the information age has changed the definition of learnedness from lower level learning to higher level learning.

Learning Innovations

In post-secondary level science courses, instructors traditionally lectured using a blackboard and chalk. Students listened to the teacher passively and recorded notes. (Dufresne, et al., 1996). For over a hundred years, this has been the most convenient and prevalent teaching method. This method of teaching is known as the ‘chalk talk’, and refers to instructors writing on the blackboard and verbally repeating what they just wrote. This is the traditional way to deliver a lecture, as it conveys a great deal of information in a short time and makes the class easier to control (Laronde & MacLeod, 2012). However, this methodology has drawbacks with respect to teaching. To begin with, instructors can easily lose student attention and engagement. In a chalk talk, the voice repeats the content written as an audio duplication of the visual information.
Students can be lulled by a droning voice and lose focus if neither the content taught or the delivery method arouses their interest. When students stop concentrating, only a small amount of the information gets through and it is rarely remembered or understood. Secondly, this method primarily uses uni-directional communication. Most of the time there is only information output from the instructor and no response from the students. Students seldom ask questions or interact between each other to discuss the contents in depth, which discourages higher levels of learning. Moreover, the monotonic teaching mode may not cater each and every student, especially the ones preferring audio-visual materials (Laronde, & MacLeod, 2012). Last but not least, it consumes a great amount of time for the teachers to write things down on the board and for the students to copy them down. The students’ attention may be split between writing and listening, which interferes with the uptake and comprehension of information.

When the age of desktop computers began, a new trend was quickly embraced by the majority of the teachers. Software like powerpoint was created, which facilitated presentation of graphics and videos along with text. These programs have numerous advantages over the chalkboard. First and foremost, the powerpoint file can be distributed out to the students in advance so that they can print it before class. This saves time in taking notes during class. In this way they can better focus on the teachers’ explanation and elaboration about the course contents. Besides, authentic pictures, videos and animations can often intrigue and maintain student interest for a longer period of time (Seth, et al., 2010). Also, the powerpoint allows instructors more time to demonstrate the contents to promote understanding (Apperson, et al., 2006).
Lastly, a class taught with the help of powerpoint tends to be better organized, more precise and systematic (Xingeng, et al., 2012).

Nevertheless, teaching with powerpoint still has disadvantages. A lecture with chalkboard or powerpoint does not make much difference when the instructor simply reads off from the screen, and this greatly reduces student engagement. The convenience of displaying larger amounts of prepared information makes the delivery speed too fast to keep up with, which tends to limit student participation and engagement. The previously designed order of the slides inhibits extemporaneous elaboration on contents by the instructor in response to student questions (Xingeng, et al., 2012). In view of these drawbacks, further innovation in teaching methodology is called for to improve learning effectiveness.

A historic turn happened when Eric Mazur tested a teaching method in 1990. This physics professor realized that the students learned ‘next to nothing’ in his introductory courses. Despite covering a semester of physics concepts, the low retention and comprehension pushed him to make changes to his teaching. He discovered that even though he explained a question to the class for 10 minutes, students still didn’t understand. When the students discussed the question with each other for only 3 minutes, nearly complete comprehension was achieved (Lambert, 2012). He then started experimenting with class discussion by further refining this pedagogy away from lecturing. In order for there to be enough time for the student discussion, textbook information would be assigned to the students as reading materials in advance, which was checked by a five-minute quiz at the beginning of class. Deep learning of the main contents were delivered during the remainder of the class period in
intervals or segments of 10 to 15 minutes. Each section was followed by a conceptual
question to evaluate understanding. The students were not allowed to communicate in
the first trial, but after submitting their answers they were encouraged to discuss the
material and concepts among themselves and try to convince others of their own
answers. After the discussion, answers were collected for the second time, where
students could freely change their choice from their first answer. Mazur reported that
after these discussions the proportion of students getting the right answer significantly
increased (Mazur, 1997).

This innovative style of learning grew into “peer instruction” or “interactive
learning” and gradually became a pedagogical method that spread across countries and
fields (Lambert, 2012). This teaching method works well with physics, since physics is a
subject focusing on problem-solving. Mathematics and chemistry may also use the
same instructional strategy, but in subjects like psychology and biology, it may not work
so well because those subjects are more of the type about finding appropriate approach
to certain questions. In addition, students encounter a large amount of vocabulary in
these subjects, and rote memorization is required, but not enough for understanding the
concepts and applying them to solve problems (Fagen, 2003).

Adaptations to this methodology were made in these subjects accordingly. In
order for the instructors to yield better learning outcomes in concept-oriented subjects, it
is important that they know common misconceptions that the students have in these
courses. This can be done by designing separate assessments for concept application
and comprehension. If the student guessed the correct answer for the concept
application question but failed to correctly answer the corresponding comprehension
question, then the student has not fully understood the concept or had misconceptions about it. Using matched pre- and post-tests can reveal if the misconception is eliminated or if it still persists after instruction. Lecture, accompanied by discussion sections and laboratory experiments enhances student understanding (Fagen, 2003). Peer-graded weekly practice exams with short-answer questions were implemented and improved performance (Haak, 2011). Further improvements such as inclusion of research-based case studies in portions of the textbooks and incorporation of cooperative learning into the discussion sections promote the student achievement in grasping quantitative problem solving, as well as conceptual reasoning (Crouch, et al., 2001).

Apart from peer instruction, many other strategies have appeared under the genre of ‘active learning’. This term was first introduced by R W Revans the English scholar, originally as ‘action learning’. This term referring to the instructional strategy that involves students directly in the learning process and requiring their input on activities (Revans, 1981). Bonwell put forward and refined the definition, excluding listening to the lecture and rather thinking about what the students are doing. Since reported by the Association for the Study of Higher Education (ASHE), active learning became popular in the 1990s (Bonwell, et al., 1991). Over the past two decades, its definition has further developed, and various teaching strategies arose within many fields (Ueckert, et al., 2008). Project Kaleidoscope (PKAL) is launched, with the goal of transforming undergraduate studies in STEM (science, technology, engineering and mathematics) subjects. It attempted to attract and inspire students to persevere in STEM fields by providing better learning experiences through collaborative efforts and innovations in science education.
These teaching innovations produced a range of advancements. Common approaches of active learning include adding authentic field experiences, using case studies and introducing primary literature (Ueckert, et al., 2006). It has by now been well supported that employing these student-centered, active learning techniques greatly enhances learning outcomes in many science-related disciplines (Fagen, et al., 2002; Michael, 2006; Freeman, et al., 2014). For example, Ruhl et al. tried to break up a 45-minute lecture into three parts and let the students discuss in pairs to clarify their notes during the two-minute intervals. He noted improvements in test scores that evaluated both short- and long-term retention of course concepts (Ruhl, et al., 1987). Gokhale used another method. After a task is assigned, he let students discuss it for half an hour. Students then filled out worksheet collaboratively for fifteen minutes. When students are tested on the material they studied, they scored higher than the students learning from regular lectures (Gokhale, et al., 1995).

As to the reason why active learning works well in teaching, it has been proposed that such teaching methods engages students in higher-level learning. During active learning, students are actively participating in tasks that require them to analyze, synthesize, evaluate and create. These are considered as more advanced levels than memorization and understanding in Bloom’s taxonomy, and are believed to enhance learning outcomes.

More recently, a new type of active learning, authentic research experiences (ARE), became implemented at some universities. When compared to traditional instruction methods, authentic research experiences for undergraduate students improves learning outcomes (Kardash, 2000). The National Science Foundation (NSF)
began sponsoring summer research programs for undergraduate students in 1958. The programs included research into math, physics, chemistry, biology, psychology, geology, and computer science. Hundreds of awards have been given to the students participating in the summer research project for their outstanding achievements. Some time ago, such research experiences were limited to a small fraction of undergraduate students in these fields, but more recently universities routinely incorporated research experience into the academic life of undergraduates. There are two general types of undergraduate research experience – course-based research projects and mentored research in academic labs.

Mentored undergrad research experience (MURE) occurs when students volunteer during their spare time to work in a lab that fascinates them and earn credits for doing so. They are mentored by the professor in charge of the lab or the graduate students in that lab. Such one-on-one or multiple-on-one mentoring provides students with abundant opportunity to enhance their understanding of laboratory techniques and analysis skills.

However, unlike MURE, course-based undergraduate research experience (CURE) occurs in a teaching laboratory or classroom setting, where one instructor leads the whole class. In this case, there is less chance for students to be mentored individually. CURE differs from traditional laboratory exercises with prewritten protocols, known conclusions and little to interest the inquiring mind wishing to explore the unknown. CURE engages every student in the class by asking them to address a research question or problem that is of interest to the scientific community (Auchincloss et al., 2014). In CURE learning trials, a wide range of positive outcomes have been
reported. CUREs provide students with chances to build confidence in doing research (Seymour, et al., 2004; Bangera, Brownell, 2014). This is reflected in the fact that research participation experiences increased the fraction of students pursuing PhD degrees (Seymour et al., 2004; Russell et al., 2007; Graham, et al., 2013). The undergraduate research experiences also improve the retention of students in scientific research (Laursen et al., 2010; Eagan et al., 2013). For example, students who participated in a phage discovery and genomics research course in their first year of college showed significantly higher retention rates in this field of inquiry than the students who didn’t take that course (Jordan, et al., 2014). Improvement in the comprehension of course content is reported. In the case of a genomics course, undergraduate students were given the task to annotate assigned genes based on current evidence. In the pre-post analysis on knowledge-based quiz, students made significant progress according to the score (Eckdahl, 2010). Another advantage is that research projects foster critical thinking skills and a more thorough understanding of the content delivered in the course (Burnette, et al., 2013; Handelsman et al. 2004; Thiry, 2011). As a case study showed, students participating in the course-based project where they try to characterize mutant versions of p53 locus in yeast can do equally well when exam difficulty was increased. This outcome shows enhanced ability to analyze and evaluate data and information (Brownell, et al., 2015). Students are also reported to have gained the ability to generate and refine research questions and develop alternative designs (Seymour, et al., 2004). Reports also mentioned improved professional qualities, such as grasp of lab techniques, independence in working, and tolerance to obstacles (Lopatto, 2007).
Even though there are advantages of CUREs in learning outcomes, there are also barriers to implementing CUREs. From the perspective of the instructors, transition from traditional teaching methods to the new approach takes courage and commitment. Not very many instructors are convinced that research-based courses can actually yield better learning outcomes. It also intimidates them to spend much more time preparing for CUREs, because they’re not familiar with it. They think the increased amount of time consumed will hinder their research (DeHaan, 2005). From the perspective of students, they may encounter stress in trying to handle a course with such high demands for the first time. The assessment is greatly different from standard tests, and it also requires much more time (Gülbahar, Tinmaz, 2006). From the perspective of course organization and school academic management, there are also plenty of issues. To begin with, there are few entry-level CURE that can be implemented for freshman and sophomore classes. Also, institutions may not be able to implement CURE because of the demands on space, time and funding. Last but not least, a higher level of proficiency is required of instructors to lead these classes (Spronken-Smith, 2011).

**Bioinformatics**

In order to see how course-based research experiences affect student learning outcomes, we introduced a research project into a student lab. During the semester, the students will use a bioinformatics network to make hypothesis about a gene of unknown function. Based on this hypothesis, students will test the proposed function and make observations about mutant phenotype to verify or refute the hypothesis.

Bioinformatics is a rapidly evolving field where massive amounts of biological data are collected, organized, analyzed and shared online. This interdisciplinary field of science combines statistics, computer science, mathematics and engineering to analyze
and interpret biological data. Bioinformatics evidence from a variety of assays and inferences are uploaded into a database and the association from this information was used to create a gene network. Information in this network can be used for hypothesizing about gene functions.

The data types below were used to create the Arabidopsis gene network (Lee, et al., 2010):

1) Co-citation

Two genes are considered to be semantically related if they’re cited together in the same piece of literature. The higher the frequency of the genes being cited at the same time, the stronger their relationship is. This strength is usually measured as the science citation index (SCI) (Small, 1973). Stapley and Benoit tested the effectiveness of discovering related genes through literature coupling. They recorded the frequency of genes co-occurrences in *Saccharomyces cerevisiae* in biomedical literature, and constructed a matrix for every pair of genes based on their joint and individual occurrence statistics.

This method is proven to be efficient in inferring relationships between genes, as genes found to be grouping together in an example shared up to 32% identity (Stapley, Benoit, 2000).

2) mRNA co-expression

Gene expression profiles are the transcript level of genes in samples. Such data is generated from microarrays that compare mRNA existence and content across treatments, samples or species. Genes are considered to have similar expression patterns if their transcript levels rise and fall together across samples and/or treatments.
Genes with similar expression patterns under different conditions or from different sources may be functionally related (Weirauch, 2011). This theory originated from the fact that, in prokaryotic organisms, genes that are functionally related or are members of the same pathway or protein complex are usually linearly arranged on the chromosome and controlled by the same transcriptional regulatory program called operon. Such system is absent in eukaryotes after millions of years of evolution, but similar patterns of co-expression still are biologically significant. An analysis of human mRNA co-expression showed that co-expression confirmed by multiple data sets and provide credible evidence for a functional relationship for co-expressed genes. Confirmed co-expression links between genes show significantly higher chance of having similarly annotations ($P<10^{-15}$), which means gene microarray data can help form reliable inferences about functionally coherent gene groups (Lee, 2004).

3) Domain co-occurrence

Proteins are the active and functional form of genes or loci. With the amino acids coded by DNA sequence, protein structure can be predicted by computer software. A complete protein usually consists several functional sections called domains. Proteins tend to have similar functions if similar domains exist in them. Marcotte studied genome sequence and protein domains of a vast number of genes, and proteins with similar putative domains were confirmed to be functionally related. There is 18% function overlapping between co-occurring domains, whereas random genes only showed 4% overlap. (Marcotte, et al., 1999).

4) Gene neighbor
As mentioned above, genes that have similar or related functions usually exist next to each other in the operons of prokaryotes. Even though operons no longer exist in eukaryotes, there is still a tendency for functionally related genes to be located closer to each other. This is because the more closely related they are, the less likely they will be separated during evolutionary recombination, since there is strong selective force for them to stay adjacent. The evidence is even stronger when the genes are grouped together in species that are not closely related. It has been reported that functions of uncharacterized genes can be predicted by aligning multiple genomes and looking for gene neighbors. In the research, 4% of the genes studied had significantly clarified functions based on neighbor gene information (Wolf, et al., 2001).

5) Genetic interaction

If a double mutant exhibits the same phenotype as one of the single mutations, there is epistasis between the genes. This may reveal functional relationships between the genes and can be used to construct genetic pathways (Mani, 2008). The study done on fission yeast, *Schizosaccharomyces pombe*, showed that genetic interactions are informative of functional relationships. Interactive gene pairs showed relatively high functional correlation ($r = 0.60$), while universal comparison of gene function correlations showed weak correlation coefficient ($r = 0.14$) (Roguev, et al., 2008).

6) Protein interactions

Proteins physically contact each other so as to carry out certain functions. Such direct interaction is the basis of molecular machines for many biological processes. Physical interaction can be detected by yeast two-hybrid assay, affinity purification
coupled with mass spectrometry and fluorescence resonance energy transfer (FRET).

Interacting proteins are involved in a biological function.

7) Phylogenetic profiles

Phylogenetic profiling detects the joint presence or absence of two traits across various species. Such data can be used to infer biological connection between genes in functional pathways. The assumption is that proteins with related functions have higher chance of correlated evolution and tend to be either preserved or eliminated together in new species (Pellegrini, 1999). A study on microbial genomes reported that gene function linkage was obtained at high level of significance, as phylogenetically linked genes showed 30-fold greater confidence in functional relatedness than control at a level of significance of 0.01 (Wu, 2003).
In our CURE setup, students are guided in accomplishing a project using bioinformatics information to identify the function of an assigned gene by tasking the students to solve incremental steps in the process in an authentic research project setting. This classroom research project attempts to improve student comprehension of bioinformatics models and molecular genetics as they gain hands-on experience in bioinformatics, genetics, plant biology, molecular biology and hypothesis generation and testing.

The approach of hypothesizing the function of a gene based on gene interaction networks is a fast-emerging field that integrates computer science with applications derived from molecular biology (Roos, 2001). The conventional process is to screen gene of interest from expression patterns from different treatments, gather information about the genes with similar expression profiles as the gene of interest and make a hypothesis about the function of the gene in question based on the role of these other genes. The hypothesis can be tested by observation and evaluation of the mutant phenotype, and further supported by molecular analysis of the encoded protein.

Bioinformatics is defined as the application of tools of computation and analysis to the capture and interpretation of biological data (Bayat, 2002). In the last two decades, the amount of biological information that can be found online has exploded exponentially. A combinatorial map of gene network and interactions was created using the whole genome sequences, protein interaction, domain prediction, expression level, co-localization, phylogenetic data and the journal literature. The network integrates biological data from different assays into a single model, enhancing both model
accuracy and coverage. With this network, new candidate genes are proposed for traits based upon network associations with genes previously linked to these traits (Lee, 2010).

There are two traditional ways of studying genes with unknown functions. Forward genetics is the approach where phenotype is observed in a population of randomly mutated species, and the goal is to find out which gene is responsible to the phenotype. Whereas reverse genetics is the approach where a specific gene is mutated, and the goal is to find out what phenotype is related to this gene. In the bioinformatics search, we’re doing reverse genetics, which means trying to find out the consequence of mutating the gene of interest to determine its function. There are various ways of mutating a gene, such as T-DNA insertion, transposon, chemicals, radiation and so on. SALK institute provides a wide range of mutated Arabidopsis seeds, along with a database of the insertion sites and other detailed information about the mutants. The AraNet and GeneMania are the websites with gene interaction networks that scientists can base their hypothesis on and generate functional predictions from (Lee, 2014).

In the course setting, students will form groups of two, and each group will be assigned a mutant to study. They will extract plant DNA and conduct PCR with primers on both sides of the insertion site to determine whether the plant is homozygous mutant, heterozygous mutant or wild type. After genotyping the plants, students are guided in the process of searching the Arabidopsis network maps for other loci that associate with the gene assigned to them. Using these data, students will infer gene function and put forward a hypothesis about gene function. Next, they will search literature to find
common ways of testing their hypothesis. For example, they may be observing total leaf area over the growth of the plant; they may be counting the number of stomata; they may also be dissecting the flowers and look for changed structures. In order to reach a conclusion that is consistent with the scientific method, students are expected to set up control groups and replicates. Wild-type Arabidopsis plants are used as control group to compare mutant plants to. Multiple individuals with the same mutation are observed as replicates. Students will be choosing appropriate statistical tests such as t-test or ANOVA to verify or refute their hypothesis based on the information they collected from their observation.

In this CURE project, we intend to find out whether engaging students in course-based research can improve learning outcomes. Improvement is evaluated by comparison of grades from pre-test and post-test. A pretest-posttest design is a commonly used method where instructors assess the students with the same exam questions before and after implementing a certain pedagogy to see the effects of the teaching method on the learners. Dori and Belcher used pretest and posttest to see the effect of technology-enabled active learning on undergraduate students’ understanding of electromagnetism concepts, and they found that students taught with active learning methods exhibited improved performance in the test of conceptual questions to a significantly higher extent than control group (Dori, Belcher, 2005). In our case, students will take the pre-test at the beginning of the semester and take the post-test and at the end of the semester. The tests consists the exact same questions.

The pre-test and post-test consisted of two formats – short answer questions and multiple-choice questions. The short answer format consisted ten questions, of which
the last is an experimental design/improvement question; the multiple-choice format consisted similar questions as asked in the short answer version, but the questions are more specific and they are accompanied with four detailed choices and ‘none of the above’. We give students the short answer format first, and after it’s been collected, the multiple choice format is handed out. [We switched the order in the second year.] They are allowed ample time to answer the questions.

All the questions in both formats weighed one point except for the experimental design/improvement question, which weighed two points. Short answer questions were graded with partial credit. Quarter points were taken off according to the correctness of the answer with reference to the grading rubric. To ensure consistency of grading criteria, short answer questions of the pre-test and the post-test were graded together at the end of the semester. Multiple choice questions were graded in two ways, with or without penalty for guessing. In the grading criteria without penalty, full point was given to the correct choice and no point was given to the wrong choices. In the grading criteria with penalty, each correct answer is worth 0.2 point, each wrong answer is worth -0.2 points, and leaving the answer blank is given 0 point. The penalty can help evaluate the extent to which students are guessing.

T-test was applied to see if there is significance of difference between the grades of the pre-test and the post-test. Point biserial correlation test was applied to analyze the reliability of the test questions, and changes were made to the questions for the second year after analyzing the feedback from the first year.
CHAPTER 3

Abstract

Instructional strategies evolve to improve learning outcomes. Active learning has emerged and are commonly believed to yield better learning results than the traditional lectures. More and more forms of active learning are being developed and tried, and course-based research experience (CURE) is one with great potential. In order to see if CURE can indeed help students learn better, we introduced a research project into our plant physiology lab course. The students accomplished a bioinformatics research project where they searched online database, browsed primary literature and carried out experiments to test hypothesis about gene function under the guidance of the instructors. The lab integrated contents about proteomics, genetics, molecular biology, plant physiology and statistics. A pre/post test was used to evaluate student improvement on content knowledge. Our results showed that students made significant improvement on most topics during the CURE, and we discussed future revisions for improving the design and evaluation of the CURE to further improve learning outcomes.

Introduction

In our CURE, students are taught to use bioinformatics tools to identify the function of a gene of no known function by tasking the students to solve incremental steps in the process. This classroom research project attempts to improve student comprehension of bioinformatics models and molecular genetics as they gain hands-on experience in bioinformatics, genetics, plant biology, molecular biology and hypothesis generation and hypothesis testing.

The approach of hypothesizing the function of a gene based on gene interaction networks is an emerging field that integrates computer science with applications derived
from molecular biology (Roos, 2001). The conventional process is to screen gene of interest from expression patterns from different treatments, gather information about the genes with similar expression profiles as the gene of interest and make a hypothesis about the function of the gene in question based on the role of these other genes. The hypothesis can be tested by observation and evaluation of the mutant phenotype, and further supported by molecular analysis of the encoded protein.

Bioinformatics is defined as the application of tools of computation and analysis to the capture and interpretation of biological data (Bayat, 2002). In the last two decades, the amount of biological information that can be found online has exploded exponentially. A combinatorial map of gene network and interactions was created using the whole genome sequences, protein interaction, domain prediction, expression level, co-localization, phylogenetic data and journal literatures. The network integrates biological data from different assays into a single model, enhancing both model accuracy and coverage. With this network, new candidate genes are proposed for traits based upon network associations with genes previously linked to these traits (Lee, 2010).

There are two traditional ways of studying genes with unknown functions. Forward genetics is the approach where phenotype is observed in a population of randomly mutated species, and the goal is to find out which gene is responsible to the phenotype. Whereas reverse genetics is the approach where a specific gene is mutated, and the goal is to find out what phenotype is related to this gene. In the bioinformatics search, we’re doing reverse genetics, which means trying to find out the consequence of mutating the gene of interest to determine its function. There are
various ways of mutating a gene, such as T-DNA insertion, transposon insertion, chemicals, mutagens and so on. SALK institute identified more than 200,000 T-DNA mutagenized Arabidopsis seeds and developed a database of the insertion sites and other detailed information about the mutants.

In the course setting, students formed groups of two, and each group is assigned a different mutant to study. They extracted plant DNA and conduct PCR with primers on both sides of the insertion site to determine the genotype at this locus for a population of plants. After genotyping the plants, students were guided in the process of browsing bioinformatics networks for other loci that associate with the locus assigned to them. Using these data, students inferred gene function. Next, they searched literature to find common ways of testing their hypothesis. For example, they may be observing total leaf area over the growth of the plant; they may be counting the number of stomata; they may also be dissecting the flowers and look for changed structures. In order to reach a conclusion that is consistent with the scientific method, students are expected to set up control groups and replicates. Wild-type Arabidopsis plants are used as control group to compare mutant plants to. Multiple individuals with the same mutation are observed as replicates. Students will be choosing appropriate statistical tests such as t-test or ANOVA to verify or refute their hypothesis based on the information they collected from their observation.

In this CURE project, we intend to find out whether engaging students in course-based research can improve learning outcomes. Improvement is evaluated by comparison of grades from pre-test and post-test. A pretest-posttest design is a commonly used method where instructors assess the students with the same exam
questions before and after implementing a certain pedagogy to see the effects of the teaching method on the learners. Dori and Belcher used pretest and posttest to see the effect of technology-enabled active learning on undergraduate students’ understanding of electromagnetism concepts, and they found that students taught with active learning methods exhibited improved performance in the test of conceptual questions to a significantly higher extent than control group (2005).

The pre-test and post-test consisted of two formats: short-answer and multiple-choice questions. The short answer format consisted ten questions, of which the last is an experimental design/improvement question; the multiple-choice format consisted similar questions as asked in the short answer version, but the questions are more specific and they are accompanied with four detailed choices and ‘none of the above’.

Results

This CURE was undertaken in a plant physiology lab course where students accomplished a bioinformatics research project under the guidance of the instructor and teaching assistant. The students learned about making hypothesis and conclusions about gene function using AraNet, an associational bioinformatic network for all Arabidopsis genes. They searched for bioinformatics information about the genes related to the mutant assigned to them, genotyped the mutant plants, browsed literature, planned an experiment to compare mutant phenotype against wild type control to test their hypothesis, and analyzed the data to either refute or accept their hypothesis (Figure 3-1).
In order to see if the research project experience can improve learning outcomes, students were evaluated by a pre-test at the beginning of the semester followed by a post-test with the exact same questions at the end of the course. The pre-test and post-test consisted of two formats: short answer questions and multiple-choice questions. We applied paired t-test to analysis these data from pre-post tests. There were 10 students in the class of year 2016 and 13 students in the class of year 2017.

There were nine questions in the 2016 test. After some adjustments, seven of them are kept for year 2017, and we introduced a few more new questions. Questions 1, 3, 4, 5, 7 are about genetics, with 1 and 3 also touching on bioinformatics. Question 2 is about biochemistry and proteomics. Question 4, 5, 8 was about molecular biology. Questions 3, 5, 9, 10 are about plant physiology. Question 6 and 11 are about statistics.
In the multiple choice test (Figure 3-2), scores generally improved over the semester. There was significant improvement in question 1 and 3 for year 2016 (Figure 3-2a), while for year 2017 there was significant improvement for question 3, 6, 7 and 8 (Figure 3-2B). The variance for each question was smaller in 2017 than 2016. However we do notice that for some questions scores remain unimproved or even dropped after the course.

![Figure 2-2. Pre-test and post-test scores in the multiple choice version. A) Year 2016, B) Year 2017. Error bars show positive and negative standard error. An asterisk indicates significance difference (* for p<0.05, ** for p<0.01, *** for p<0.001).](image)

When we look at the combined results over the two years (Figure 3-3), improvement in performance become more significant. There is only one question where students’ score dropped.
Figure 2-3. Combined pre-test and post-test scores in the multiple choice version for year 2016 and 2017. Error bars show positive and negative standard error. An asterisk indicates significance difference (* for p<0.05, ** for p<0.01, *** for p<0.001).

In the short answer test (Figure 3-4), improvement over the semester is more significant than in the multiple choice test. Only one question in year 2016 showed decrease in score, while all of the scores improved in year 2017. The number of questions with significant improvement increased.

Figure 2-4. Pre-test and post-test scores in the short answer version. A) Year 2016, B) 2017. Error bars show positive and negative standard error. An asterisk indicates significance difference (* for p<0.05, ** for p<0.01, *** for p<0.001).
When we look at the combined result of year 2016 and 2017 (Figure 3-5), there is improvement in every question, with 5 out of 7 questions being significantly improved. The level of significance is also generally higher than in separate years.

Figure 2-5. Combined pre-test and post-test scores in the short answer version for year 2016 and 2017. Error bars show positive and negative standard error. An asterisk indicates significance difference (* for p<0.05, ** for p<0.01, *** for p<0.001).

We also assessed the students’ improvement on various topics covered over the semester by looking at the certain groups of questions (Figure 3-6). Questions testing the same subject are grouped together to reveal differences in comprehension between the two time points. A general trend of improvement was observed in both multiple choice and short answer tests. Short answer questions still showed higher significance than multiple choice questions. Using multiple choice questions, bioinformatics improved the most, followed by genetics and plant physiology. In the short answer test, however, plant physiology showed greatest improvement, followed by genetics, molecular biology and various other topics.
Disc **ussion**

Results from the pre-post assessment tool indicate that this bioinformatics CURE that we implemented in the plant physiology lab generally showed improved learning outcomes. However, there are unforeseen trends, as well as discrepancies between these results and our expectations, and we discuss the possible reasons for this below.

Overall student performance improved more significantly on short answer questions than on multiple choice questions. This may be due to the fact that we assigned partial credits to short answer questions, more accurately assessing incomplete mastery of a subject and rewarding students for the knowledge they had gained. On the contrary multiple choice questions are either ‘completely correct’ or ‘totally wrong’, which is less representative of the amount of content grasped by the students. However, empirical studies on history, English language, economics and biology indicated that multiple choice grades are more representative of student performance (Bridgeman, Lewis, 1994; Walstad, Becker, 1994).
We can also see that the results from year 2017 generally have smaller standard error and higher overall scores. This is because after the trial in year 2016, we reworded some test questions based on student performance and point-biserial analysis. We deleted the questions that were either too difficult or not so relevant to the topics covered in class. We also changed the wording and professional vocabulary in some questions so that the students could better understand what was being asked. Also, to better represent the major topics taught in class, we added some new questions for subjects we wished to assess. This series of adjustments seem to have met our intention of designing pre/post-tests that evaluated learning in this CURE in a variety of subjects related to the research investigation.

When we looked at the results with the years combined, we noticed higher significance of score improvement. This is partially due to the fact that the two years of data are replicates and when we put them together, the number of data points increased, increasing the statistical power.

However, in the data we recorded, we do see that some questions where scores marginally improved or even worsened. A case in point is question four, which asks about how researchers produce null mutation. Student performance dropped in both years in the multiple choice version. This is partially due to the fact that students misunderstood ‘null mutation’ as a mutation which did not affect plant phenotype, instead of one that completely aborted the function of the mutated gene. The term ‘null mutation’ was mentioned for only two or three times without specific emphasis during the semester and may not have left a deep impression on the students. Also, some students may have guessed the correct answer in the pre-test, so that even if they
answered correctly in the post-test, there may not be an overall improvement in average score on this question. Moreover, a few students did not try as hard as they did in the pre-test because answering the same questions twice was tiring and they thought doing badly wouldn’t affect their grade. Such attitude may also explain why some questions didn’t show significant improvement.

Interestingly, students did better in post-test than in pre-test on this same question in the short answer version, and this is because we gave partial credits to answers that touched on how a mutation is produced (even though partial credit was deducted for wrong information in the answers). Plenty of students left this question blank in the pre-test because they either totally didn’t understand the concept or did not know the terminology to answer. In the post-test, students were able to at least say something correct about the question, which led to an increase in scores.

In our assessment of student performance on specific topics, we found that statistics and biochemistry were less familiar to the students, since those topics were not emphasized in the CURE. Even though students are required to apply statistical analysis to lab data in almost every lab report, this subject remained hard for students to grasp. It may be meaningful to systematically introduce statistical knowledge during the course so that the students can better understand it and apply it. The students did not do well in biochemistry questions because this topic is hardly emphasized during the semester, nor were there experiments to let the students have authentic experience about biochemical assays. We may either shift some emphasis onto this topic, or delete this question in the future.
In the years to come, we plan to further improve the CURE setting in several aspects. First, we would like to have a larger sample size so that the conclusions we draw may have greater statistical power. This target can be achieved by continue collecting data throughout the years, increasing student number or class number. Secondly, we would keep revising the questions in the test. One goal is to improve student understanding by adjusting wording and the vocabulary used in the questions, because this is vital for getting representative assessment of authentic student gain of knowledge in the course and also assessment of students entering the course that lacked the needed technical vocabulary. We would also rearrange the questions so that they can better reflect the major topics covered in the course. We would eliminate questions that are too similar and introduce new questions about important concepts that we left out in the current version of tests. Thirdly, we would improve our grading criteria so that the scores can be more representative of student performance and study gain. We tried penalizing guessing wrong answers in the multiple choice test but it did not improve the assessment. We might redesign the choices so that each choice is worth different points depending on how suitable or accurate they are. In this way the multiple choice test can also be assigned partial credits and we may be able to analyze the students’ performance in more detail.
CHAPTER 3
CONCLUSION

In our CURE project, we designed a classroom project for students to participate in scientific research about plant gene function during in a plant physiology lab course. We also developed a set of questions and used them as pre-test and post-test to evaluate student learning outcomes. Our findings showed that engaging students in a course-based research experience can improve their understanding of the course materials, but its effectiveness varied among subjects.

We found that in the multiple choice version of our tests, student scores improved on most questions, and for about half the questions the improvements were statistically significant (p<0.05). These results support the hypothesis that introducing research experiences into traditional lab course enhance student understanding of what was taught in class. Even though there were some questions where students didn’t have improved scores, we concluded that it was because these specific questions are not well-designed or insufficient learning time was devoted to these subject areas. We plan to improve the questions in the near future so that they can better reflect learning outcomes.

We also found that in the short answer version of the tests, the general trend of improvement was similar to the multiple choice version. Nevertheless there are more significant improvements in the short answer version. This revealed that the short answer version more precisely measures student understanding than multiple choice test, since multiple choice questions are cued recall while short answer questions ask more of the students’ grasp on the topic.
We also found varied level of improvement on different course topics. Genetics and plant physiology saw most significant improvement, while statistics saw only marginal improvement. The result was consistent with the extent to which the course was focused on each topic. The topics that did not improve significantly were not emphasized during the semester. In the future these test results will serve as a guide about where we should improve our teaching.

Our research results illustrated the potential of course-based research experience to improve student performance or learning. There are plenty of interesting questions to be further looked into. It would be meaningful to discover the key elements in improving student performance, so that instructors can design better CUREs for better learning results. We also need more appropriate evaluation strategies to assess student performance. Besides students’ academic achievements, we can also observe student attitude on the course and see if CUREs can improve that. We can follow the students and see if students attending CUREs are more likely to participate in that field of research, and if they would retain for more years in that field of study.

Despite the limits of our study, it provided insight into the effects of course-based research experience on student learning outcomes, and also offered opportunities for further research into the potential benefits of this practice. The success of our trial in CURE lays the foundation for applying CUREs in many other subjects and fields.
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

Jieying Wang completed her undergraduate studies at Fudan University in China. She got her bachelor’s degree in life sciences in 2015. She was a Master of Science candidate in biology at University of Florida College of Liberal Arts and Sciences. She majored in botany in the Department of Biology at University of Florida. She worked in the Hauser Lab under the guidance of Dr. Bernard Hauser. She was teaching assistant for integrated biology lab BSC2011L and plant physiology lab BOT3503L.

Jieying Wang also minored in education at University of Florida. Her thesis topic was based on her teaching experience in plant physiology lab and her knowledge in both her major and minor. She worked with Dr. Hauser, Dr. Stuart and Dr. Antonenko on her thesis project about course-based research experience and its effect on learning outcomes.