Evaluating Student Preparedness and Conceptual Change in Introductory Biology Students Studying Gene Expression

Kelly McDonald, Ph.D.
California State University, Sacramento

Joseph Gomes, B.S.
California State University, Sacramento

ABSTRACT

Conceptual difficulties experienced by introductory college biology students studying gene expression are explored in this empirical study. We used an open-ended assessment instrument and a pre-test/post-test design to measure prior knowledge and conceptual change over the course of one semester. Our findings suggest that introductory biology students struggle with the basic terminology necessary to understand complex biological systems at the molecular and genetic level. While conceptual growth from the beginning to the end of the semester was less than expected, learning gains were significant for all concepts examined by our assessment strategy. Qualitative evaluation of pre- and post-tests further highlighted the difficulty students have articulating their knowledge using scientific language. In our discussion, we emphasize the importance of assessing conceptual understanding, developing instructional strategies to promote conceptual change, and the need for closer alignment of curriculum between and within institutions. Ultimately, educational and institutional resources to support faculty development in the area of teaching and learning are critical for the retention and preparation of a diverse student population in the biological sciences.

Introduction

Many students entering Introductory Biology classes designed for majors arrive underprepared for the college curriculum. Few biology departments offer remediation or preparatory courses, as is standard in other disciplines such as math, chemistry and English. Furthermore, most Introductory Biology courses cover a substantial amount of material at a superficial level and rely heavily on traditional lecture and assessment strategies. These practices continue despite mounting evidence that "depth over breadth" and a student-centered learning environment with frequent and formative assessment is more effective (e.g., Freeman et al., 2007; Handelsman et al., 2004; Knight and Wood, 2005; Udovic et al., 2002). Many have linked these short-comings to the high dropout rate, academic failure, and overall dissatisfaction among Science, Technology, Engineering and Math (STEM) majors (Seymour and Hewitt, 1997; Tinto, 1993; Tobias, 1990; Vision and Change, 2007). This is particularly true for under-represented minority groups, which currently represent only 9% of college graduates entering the STEM workforce (NAS, 2010). As faculty teaching introductory biology courses, we can begin to address these problems by 1) understanding what prior knowledge and preconceptions our students bring to the classroom, 2) identifying the concepts that pose the greatest difficulty through frequent and formative assessment, and 3) designing curricula using evidence-based strategies shown to facilitate deep learning and

conceptual change.

Genetics is part of the core curriculum for biology majors, and understanding gene expression, how genes get turned on and off, is fundamental to understanding how all cells, tissues, and organisms function. The Biology-Online.org dictionary defines gene expression as:

"The conversion of the information from the gene into mRNA via transcription and then to protein via translation resulting in the phenotypic manifestation of the gene." In the Introductory Biology course at our institution, an appraisal of exam scores and responses from student surveys indicated that the concept of gene expression routinely posed great difficulty for students. Fundamentally, in order to understand this definition, students must first understand the nature of genes, mRNA, transcription, proteins, translation and phenotypes.

Several validated assessment tools are available to gauge students' conceptual understanding of genetics (Bowling et al., 2008a; Smith et al., 2008); however, like the dictionary's definition above, these diagnostics assume a working knowledge of basic genetics terminology. In this study, we sought to discover whether students at the introductory level had sufficient prior knowledge of the basic concepts needed to understand gene expression, monitor learning gains, and ultimately inform the development of curriculum focused on the conceptual difficulties observed. While we report data on students' knowledge in a specific area of genetics, we believe our results highlight a more pervasive problem related to student preparedness for the college science curriculum.

Theoretical Framework

There is a wealth of knowledge from the fields of cognitive psychology and science education to suggest that conceptual change for students studying science is difficult (reviewed by Zirbel, 2004). The most well-known model of conceptual change in science education espouses confronting students with a conflict between their beliefs and those held by the scientific community, and then supporting their learning as they construct new knowledge (Strike and Posner, 1992). Franke and Bogner (2011) report success using this approach with a constructivist, hands-on

curriculum for high school students studying gene technology. However, even when such strategies are used effectively, research suggests that students have difficulty changing their beliefs, and subsequently retaining these new beliefs (Mazur, 1997). Chinn and Brewer (1993) suggest that "peripheral conceptual change" is most common as students create a new hybrid conception that combines their deeply-rooted beliefs with the new ideas obtained from instruction. Smith and Knight (2012) also report difficulties with conceptual change in a population of college students studying genetics. The authors identified a set of Most Common Incorrect Answers (MCIAs) related to specific concepts on a genetics conceptual assessment. An examination of student response patterns found that if a student selected the MCIA on the pre-test and then missed the same question after instruction, it was likely that the student would select the same MCIA during the follow-up assessment. They concluded that some incorrect ideas are more difficult to correct than others. Experts further agree that conceptual change is slow and iterative, requiring considerable effort by the learner and instructor. Therefore, learning gains made by students exploring complex systems may not be substantial during the course of one unit or even one semester, and this may be particularly true for terms and concepts that are unfamiliar to students prior to instruction (Chi et al., 1994; Ohlsson, 2009).

While multiple perspectives regarding the most effective way to promote and analyze conceptual change have emerged, there is considerable evidence to suggest that traditional approaches are less effective than a variety of student-centered, active learning methods (reviewed by Banet and Ayuso, 2000; Duit and Treagust, 2003). Additionally, the ascertainment of students' prior knowledge and misconceptions through classroom assessment techniques is a crucial part of these strategies (Angelo and Cross, 1993; Sundberg, 2002; Tanner and Allen, 2004). In a literature review on the effect of prior knowledge on learning, Roschelle (1995) claimed that a large body of research concluded that "Learning proceeds primarily from prior knowledge, and only secondarily from the presented materials." This suggests that the most carefully designed instruction may be ineffective for addressing incorrect ideas and supporting conceptual growth

when prior knowledge is not considered, as students will interpret the curriculum through their personal knowledge, beliefs, and attitudes.

Literature Review

A 2005 report indicated that 10% of all students graduating from two and four-year institutions were choosing careers in the life sciences or healthcare fields (National Center for Education Statistics, 2005). Yet, according to an earlier report, only 30% of high school seniors could correctly answer questions related to genetics (National Center for Education Statistics, 2000). A number of genetics misconceptions have been identified for both high school and college students (e.g., Browning and Lehman, 1988; Marback-Ad, 2001; Orcajo and Aznar, 2005; Smith and Knight, 2012). It has been suggested that many of these originate early in middle and high school (AAAS, 1993) as a consequence of conceptual difficulties that K-12 teachers have with the material (Cakir and Crawford, 2001). These difficulties likely relate to the fact that genetics requires a solid foundation of the molecular nature of biological systems, which draws upon an understanding of physics and chemistry. K-12 science teachers rarely master these disciplines during their training (Klymkowsky, 2010). In addition, some of the difficulties students encounter with genetics are not actual misconceptions, but instead may result from incomplete understandings and confusion regarding the relationships between different concepts (Tanner and Allen, 2005; Lewis et al., 2000; Marback-Ad and Stavy, 2000).

While the reasons for the difficulties are varied and complex, it is clear that students at all levels struggle with genetics concepts (Marbach-Ad, 2001; Longden, 1982; Stewart, 1982; Hildebrand, 1991; Lewis and Wood-Robinson, 2000; Smith and Knight, 2012). These findings have serious implications for biology instructors and curriculum developers. Faculty teaching Introductory Biology courses are often faced with highly diverse student populations. These include students from high schools with Advanced Placement (AP) biology training or inadequate science programs, as well as non-traditional students that are returning to college years after high school graduation. Furthermore, with few or no course prerequisites, the

introductory classes enroll students possessing a broad range of interests and aptitudes for the biological sciences. The student diversity in the introductory courses makes assessment of prior knowledge, skills and disposition crucial to planning effective instruction. However, most college level biology courses continue to evaluate student knowledge exclusively through summative assessments that measure achievement for the purpose of assigning grades at the conclusion of a lesson (Tanner and Allen, 2004).

In response to the need for diagnostic instruments, a variety of tools have been developed by the science education community to measure student learning and identify misconceptions across several scientific disciplines, including physics (Hestenes et al., 1992), chemistry (Landis et al., 2001), geology (Libarkin and Anderson, 2005), general biology (Garvin-Doxas et al., 2007) and others (reviewed by Libarkin, 2008, D'Avanzo, 2008). When used in a pre-test/post-test design (Dimitrov and Rumrill, 2003; Sundburg, 2002), these conceptual assessments (or concept inventories) can be effective methods for gauging both prior knowledge and conceptual change. In recent years, several tools have been developed and validated to measure student learning in the field of genetics (Bowling et al., 2008a; Smith et al., 2008). Bowling and colleagues (2008a) developed a Genetics Literacy Assessment (GLA) to evaluate student understanding of 17 concepts central to genetics through 31 multiple-choice questions. While the GLA was designed and validated for non-majors, Smith and colleagues (2008) developed a Genetics Concept Assessment (GCA) consisting of 25 multiple-choice questions covering 9 genetics concepts, intended for both majors and non-majors. This instrument has been employed to distinguish the most and least difficult genetics concepts and identify incorrect ideas that students have the most difficulty changing. While these instruments have been carefully designed and validated, and are simple to administer and score, there are measurable limitations to their use for classroom assessment (Smith and Tanner, 2010). Most concept inventories are comprehensive, but don't allow for the examination of specific topics in depth. The inventories are generally inflexible in their content and structure, and while instructors may choose to select or eliminate specific questions, the validity and

interpretation may be compromised. Furthermore, while the multiple choice answers have been carefully written to include common misconceptions, this format does not allow for the identification of additional misunderstandings. The concept inventories described above were designed as standardized tools for widespread use, but others have designed their own classroom assessments, tailoring them for a particular population or around one or more specific learning outcome (Nazario et al, 2002; Elrod, 2008). For our study, we chose an instrument that afforded us the flexibility to select specific terms related to gene expression and provided rich data from written students' response. Limitations of this instrument included the need to develop a rubric and train scorers, as well as the time-consuming nature of the scoring.

Purpose and Research Questions

The overall purpose of this study was to evaluate both the prior knowledge and conceptual change of introductory level biology students studying concepts related to gene expression. With regard to a select set of terms/concepts, we sought to understand:

1. What prior knowledge do students possess upon entering college Introductory Biology?

2. How much conceptual change can occur over the course of one semester?

3. Which terms present the greatest and least difficulty?

Methods

Participants: A total of 120 introductory biology students from a four-year comprehensive, public university participated in this study during the Spring 2011 semester. The Introductory Biology course, BIO 2: Cells, Molecules and Genes, is a 5-unit class composed of two seventy-five minute lectures, one three hour lab and one two-hour activity per week. It is the second of two lower division courses required for Biological Sciences and related majors (e.g., Biochemistry, Environmental Sciences) as well as students applying to post-baccalaureate health professions programs. The assessment and methodology for this study conformed to, and was approved by, the Institutional Review Board (IRB) of the researchers' institution (Protocol #11-12-116).

Assessment Instrument: The assessment tool used in this study is an open-ended response instrument called the Ten Word Test, developed by Dr. Terry Underwood at California State University, Sacramento, for evaluating a collaborative project between the English department and Center for Community Engagement (personal communication, 2010). We modified the test to investigate students' prior knowledge and conceptual development around the concept of gene expression. The following ten terms were selected for inclusion: DNA, RNA, exon, gene, mutation, transcription, translation, epigenetics, protein and phenotype. The Ten Word Test is comprised of three parts, which examine students' 1) self-assigned confidence rating of each term, 2) ability to define or describe each term, and 3) ability to construct an essay in which they use some or all of the ten terms to explain the main concept.

In Part I, students' rated their confidence or level of comfort with the ten terms on a scale from 1-3, using the following guidelines:

1 = you know little to nothing about the term.

2 = you have "some knowledge" about the term but cannot fully explain it to others.

3 = you know "a lot" about the term, and feel you can define and explain it fully to others.

Parts II (description) and III (essay) were scored on a scale of 1-3, with scorers using the following guidelines:

1 = Answer provides little or no evidence of understanding of the term(s); answer is inaccurate or vague to gauge understanding.

2 = Answer indicates a basic understanding of the concept, but may be lacking in detail, level of complexity or sophistication.

3 = Answer provides a nuanced or complex description of the term. When appropriate, both the structure and function of the term are accurately and thoroughly described.

In addition, the numerical scores in Part II were further characterized with a quality identifier to justify the score. The following quality identifiers were applied: B = Blank (or no evidence of knowledge), I = inaccurate description, V = vague description (lacks clarity). For a score of 2, the quality identifiers are L = limited definition, P = partially correct (this allows for small inaccuracies in details, as long as the primary

description is accurate), N = novice language used (but otherwise, accurate).

Study Design: To validate the Ten Word Test for content and clarity and to develop the rubrics for scoring the tests, we conducted a pilot study on a subset of students enrolled in the BIO 2 introductory course during the Fall 2010 semester. Rubrics were developed in alignment with the course curriculum, and in consultation with two faculty members (other than the lead researcher) from the Department of Biological Sciences. Pilot results from aggregate data on 44 samples indicated minimal prior knowledge and moderate learning gains; however, the pilot tests were anonymous, preventing the comparison of the pre- and post-tests for the same individuals. Findings from the pilot guided the refinement of the Ten Word Test instructions and scoring rubrics used in this study.

For the current study, we used a mixed methods approach employing a pre-test/post-test repeated measures strategy to evaluate students' prior knowledge and learning gains over the course of one semester. Identical pre- and post-Ten Word Tests were administered in the first and last weeks of the 15-week semester. Students took the tests during the regular class period, and were given ample opportunity to complete all three parts. Tests were anonymous, but student-selected numbers were used for the purpose of linking pre- and post assessments. The standard course curriculum was the intervention for the purpose of this study. During the Spring 2011 semester, when these data were collected, instruction for the molecular genetics units that directly related to the terms on the Ten Word Test spanned six weeks, but some of the terms (e.g., DNA, gene, RNA, protein) were used regularly throughout the entire semester.

Data Analysis: All tests (n=240 pre and post) were consolidated into a single batch, blinded, randomized and assigned a unique identification number. The scorers used ten tests for norming and agreement on the use of the rubric. Due to the time-consuming nature of the analysis, a random sample of 60 was chosen for in-depth analysis. Each survey in this sample was scored by at least two trained individuals, and a third was consulted in the case of disparate scores. Inter-rater reliability was determined to be > 80%.

Of the 60 tests analyzed, 32 could be matched to their corresponding pair, and met criteria for inclusion in the statistical analyses. The corresponding pairs that were not included in the set of 60 were pulled from the batch of 240, blinded, randomized and scored, resulting in a final dataset of 32-matched pairs. Allowing students to select their own identifying numbers proved to be a limitation of the design, as we were unable to match some of the pre- and post-test student-selected numbers. Microsoft Excel for Mac 2011, Version14.1.2, was used for data management and to calculate frequency and percent distributions and SPSS, IBM SPSS Statistics, Version 19, was used to compare pre and post-test scores for the 32-paired samples using the Wilcoxon Signed Ranks Test. Statistical significance was assigned to a p value of < 0.05. For the purpose of this study, we focused our data analysis on Part II of the Ten Word Test, first looking quantitatively at general trends of prior knowledge and learning gains and ranking terms in order of their difficulty. We then performed qualitative analysis on student descriptions of the terms in order to gain a deeper appreciation for the alternative conceptions that students hold and the language students use to explain their understandings.

Results

The overall percent distribution of scores assigned to students' descriptions of all ten terms combined is illustrated in Figure 1. For the pre-test, the lowest score (Score = 1) was assigned 87% of the time, indicating that most students were unable to provide adequate descriptions of the terms at the beginning of the semester. A score of 2, representing a basic level of knowledge, was achieved 13% of the time, and a score of 3 was assigned only once on the pre-tests. The aggregate data clearly show improvement from the pre to the post-test, as the percentages of scores of 1 (47%) and 2 (43%) were nearly equal, and there was a significant increase in the number of descriptions assigned a score of 3 (10%). While learning definitely occurred, the prevalence with which students received the lowest score on the post-test remained high, indicating that they were still struggling with basic terminology at the end of the semester.

The quality identifiers (Table 1) provide an explanation for the scores in Figure 1, and are informative in suggesting some general problems

Figure 1
Percent Distribution of Pre- and Post-test Scores.
Data are reported as relative frequencies of scores assigned to student descriptions for the ten terms evaluated in aggregate (n=64 tests, 32 matched pairs with 10 terms per test).
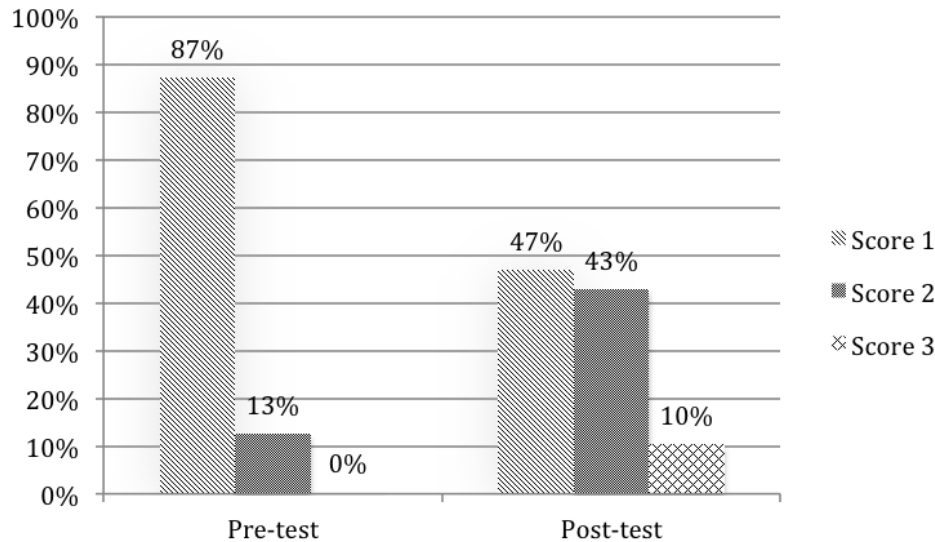
students experienced describing the terms. On the pre-test, it was clear from the number of answers left "Blank" (37%), that students often had no familiarity or were unable to articulate their knowledge of the terms. Some students indicated that they had heard of the term, but could not remember what it meant. The number of "Blanks" decreased to 5% on the post-tests, indicating that most students had enough knowledge to attempt an answer; however, their descriptions were often vague (24%) or inaccurate (21%). The number of descriptions earning a score of 2 that were accurate, but incomplete (limited), increased from 8% on the pre-test to 31% on the post-test. The high percentage of limited answers on the post-test are consistent with the findings of Marbach-Ad (2001), who reported a tendency for high school 12th graders to describe genetics concepts with vague or incomplete explanations on open-ended assessments.

To determine which concepts presented the greatest difficulty for students, we calculated percent distributions of scores for each term individually (Table 2). The difference in the pre- and post-scores was significant for all terms (p-value < 0.05), indicating an improvement in student performance from the beginning to the end of the semester. Of the 32-paired samples evaluated, only 4% of all post-test scores showed a significant decrease compared to their matched pre-test scores, whereas 46% of scores increased from pre- to post-test, and 50% were unchanged, indicating no improvement from pre-test to post-test (data not shown). All of the terms except Epigenetics are found in high school biology textbooks and curricula; however, instruction can vary among courses with regard to the time and depth devoted to each. As predicted, 100% of descriptions for Epigenetics were scored as a 1 on the pre-test, with a large percentage (78%) of students leaving the term blank. While learning gains were observed, only 16% of students scored a 2 (and there were no 3s) for descriptions of Epigenetics on the post-test. This was not surprising, as this concept is complex and only covered briefly in the Introductory Biology curriculum.

In addition to Epigenetics, students showed the greatest difficulty with the terms Exon, Transcription and Translation on the pre-test, often leaving the fields blank. This was not surprising for Exon, as it requires a more nuanced understanding of gene structure; however, we did expect more students to be capable of

providing a basic description of Transcription (i.e., the production of an RNA molecule from a DNA template) and Translation (i.e., the production of a polypeptide or protein from an mRNA molecule). Notably, while these terms were most challenging for students at the beginning of the semester, students demonstrated considerable improvement on the post-tests, with roughly half the number of scores of 1 on the post-test

Table 1

Percent Distribution of Quality Identifiers.

Quality identifiers were assigned to student descriptions scored as a 1 or 2 (n=64 tests, 32 matched pairs with 10 terms per test). There were no quality identifiers for a score of 3. Data are reported as relative frequencies for the ten terms evaluated in aggregate.

Score	Quality Identifier	Pre-Test	Post-Test
1	Blank	37%	5%
1	Vague	29%	24%
1	Inaccurate	21%	21%
2	Novice	4%	7%
2	Limited	8%	31%
2	Partial	1%	4%
3	N/A	0%	8%

Table 2

Percent Distribution of Pre- and Post-Test Scores for Individual Terms.

Relative frequencies of students' descriptions receiving a Score of 1, 2 or 3 are displayed for each of the terms evaluated. Pre- and Post-test scores for 32-paired samples are significantly different (p-value < 0.05) for all ten terms as measured by the Wilcoxon Signed Ranks Test. Due to rounding, percentages of all terms do not add to 100%.

Concept	% Pre Score 1	% Post Score 1	% Pre Score 2	% Post Score 2	% Pre Score 3	% Post Score 3
DNA	72	41	28	47	0	13
Epigenetics	100	84	0	16	0	0
Exon	100	53	0	31	0	16
Gene	78	56	22	44	0	0
Mutation	72	31	28	56	0	13
Phenotype	72	9	28	59	0	31
Protein	91	44	9	44	0	13
RNA	91	50	6	41	3	9
Transcription	100	57	0	41	0	3
Translation	97	44	3	50	0	6

as the pre-test.

For six terms – DNA, Mutation, Phenotype, Protein, RNA, and Translation - students scored a 2 or 3 on the post-test 50% or more of the time, with the highest post-test scores being achieved for Phenotype and the second highest for Mutation (91% and 69% of scores were equal to a 2 or 3, respectively). Mutation and Phenotype were two terms in which students displayed the greatest prior knowledge; however, only 28% of students demonstrated sufficient knowledge to score a 2 on the pre-test for these two terms, and there were no pre-test scores of 3. Students also showed some knowledge of the terms DNA and Gene on pre-tests, yet low scores on the post-test remained relatively high (41% and 56% score=1, respectively).

Calculating the average score for each term on the pre- and post-test provided another means of comparing the relative difficulties of the individual terms and visualizing the learning gains for each (Figure 2). In general, students scored highest at the end of the semester for terms with which they had the greatest prior knowledge. While post-test averages for all terms

fell short of a 2.0, moderate gains were observed for most. The exception was for the term Gene. The narrow spread between pre- and post-test scores for this term was completely unexpected and encouraged us to investigate student understanding of this concept in more detail.

Students' understanding of the concept of a gene was of interest for several reasons. It is a concept covered in the high school curriculum and pre-test scores indicated moderate prior knowledge (22% scored a 2 on pre-tests) relative to the other terms. Furthermore, it is a term that students hear and even use in everyday life, and it is introduced early and used frequently throughout the introductory biology curriculum. Nonetheless, in the relative rankings it proved to be one of the more difficult terms for student to describe at the end of the semester. These data indicate that students may have made less gain in their understanding of a gene compared with other concepts. Only Epigenetics, which is highly complex and covered only briefly, ranked lower. This observation prompted us to wonder how many students were

Figure 2

A Comparison of Average Scores on Pre- and Post-tests for Individual Terms (n=64, 32 matched pairs)

progressing in their understandings, yet not making sufficient advancements to justify an improvement in score based on our rubrics. To answer this question, we examined each of the matched pairs that received a score of 1 for the description of a Gene, on both the pre- and the post-test, looking for changes in language that would indicate student learning. Out of the 15 matched pairs that met this criterion, over half demonstrated some degree of improvement on the post-test, either with regard to detail, clarity or use of scientific language. Three of these were Blank on the pre-test, so for these samples simply making an effort was recognized as improvement.

Several examples of student descriptions are shown in Table 3. Student 1 provides a simple, functional description of a gene on the pre-test; yet, there is no indication that the student understands what the "code" is composed of (chemically) or how the gene is able to influence characteristics at the organismal level. As a result, this description was scored as a 1 due to lack of detail or clarity. On the post-test, the student provides greater detail using more scientific vocabulary in the description. The student demonstrates an understanding of biological complexity, explaining that a protein may be "functional or dysfunctional" and a gene "may or may not be expressed," but overall, the student demonstrates confusion as he or she attempts to relate the concept of the gene to RNA, proteins and operons (which are unique to prokaryotes). The second example also demonstrates positive change, as the

student recognizes that genes "code for specific mRNA, polypeptides," in the post-test description. However, it appears that the student is uncertain of the role that genes play on the X and Y chromosomes, and they do not appreciate that some genes code only for RNA (and not mRNA or polypeptides).

Collectively, the qualitative analyses uncovered several trends. First, students appeared more comfortable with the functional than the structural aspects of a gene, which is consistent with the findings of others (Marbach-Ad, 2001). On the pre-tests, students were most likely to define a gene with regard to its role in determining a trait or phenotype within an organism. Some also described the transmission of genes from parents to offspring, but rarely did students demonstrate knowledge of the structural or compositional characteristics of a gene. Furthermore, a thorough and accurate structural and functional explanation, required to earn a score of 3, was not observed in any of the pre- or post-tests. Student learning outcomes and the course curriculum emphasize the relationship of structure and function in biological systems, and we were looking for students to incorporate both dimensions in their descriptions. In some cases, it was suspected that the students might have possessed greater comprehension than they articulated. In other cases, however, it was unclear as to whether students were simply repeating a term they had heard frequently, or whether they were using it intentionally and truly understood its meaning. For

Table 3

Examples of Student Descriptions of a Gene on Pre- and Post-tests.

	Pre-test description of Gene	Post-test description of Gene
Student 1	A code that determines the characteristics of an organism.	The coding part of RNA. It is translated into a functional or dysfunctional protein that may or may not be expressed by inducible or reversible operons.
Student 2	Forms DNA, building block of life needed for tRNA, rRNA, and mRNA to form other DNA, organelles, and RNA.	Codes for specific mRNA, polypeptides and makes up human genetic make-up if on X and Y chromosome.

example, the term "code" was often used in students' descriptions of genes, but it was used in alternative ways. In the example in Table 3, a gene "is a code," "is the coding part" and "codes for" something else. We have limited our discussion to the challenges students had describing a gene; however, the lack of attention to structure and function (when applicable) and the difficulties with language use were observed for the majority of the terms.

Discussion

Genetics is a rapidly changing field of biology with a growing impact on healthcare, policy and society. In order to adequately prepare students for careers in the field or simply equip them with a degree of literacy required to be knowledge citizens, the concepts and principles of genetics are taught throughout the biology curriculum, from introductory to advanced courses across all sub-disciplines. Our study originated as a classroom assessment strategy to provide Introductory Biology instructors with insight into the nature of the difficulties that students have understanding concepts related to genetics. Our data suggest that many of our students enter Introductory Biology courses having retained minimal knowledge of genetics from prior coursework. We did not expect our students to be capable of detailed explanations of the underlying mechanisms or regulatory patterns of gene expression. However, we did assume them capable of constructing accurate descriptions of some basic genetics terms such as DNA, RNA and gene. This assumption was partly based on our knowledge of the high school state standards, one of which states that "Students know that the central dogma of molecular biology outlines the flow of information from transcription of ribonucleic acid (RNA) in the nucleus to translation of proteins on ribosomes in the cytoplasm." Eight of the ten terms in our study are explicitly cited in the state standards and gene expression appears multiple times throughout. Furthermore, communication with high school teachers and educators in our region confirmed the inclusion of these concepts in the actual curriculum taught to general biology students.

So the question is, why do introductory level college students perform so poorly on a basic assessment of knowledge they allegedly acquired in high school? The explanation is certainly complex, but my queries have pointed to several key factors. First and foremost, K-12 teachers are held accountable for their students' standardized test scores. As a result, most have taken a direct instructional approach in order to cover all of the material in the standards at only a superficial level. For many, this has greatly reduced or eliminated the time spent performing laboratories or other discovery and inquiry-based activities that promote student engagement and deeper long-term learning. In addition, the current science standards are largely fact-based, resulting in instruction delivered as discontinuous facts, rather than complex conceptual ideas and coherent themes. This emphasis on teaching facts, over teaching students how to think like scientists, has been a major criticism of science educators for years, but until teachers are held accountable for the latter, real reform in the K-12 system will gain little momentum. Lastly, a large percentage of students in our state perform below the proficiency level on the state's standardized tests for biology, indicating that they are leaving high school with an inadequate knowledge of the subject. Genetics comprises approximately 20% of the 10th grade biology test and only 43% of students scored proficient or higher in the 2011-2012 school year. The End-of-course scores for the same academic year indicated only 52% of graduating seniors were proficient or advanced in biology. The scores are lower than the state average for many of our urban, high poverty schools (some of which report 0% of students scoring proficient or above in biology), and it is important to note that a significant number of these schools serve as feeders to our institution (California Department of Education, 2013).

The state standardized test scores reflect the fact that urban low-income, culturally and linguistically diverse (LI/CLD) students are at a particular disadvantage when it comes to science education. They often have little or no access to science courses until high school, either because they are remediated to math and English courses at the expense of science, or because their schools have inadequate resources, including properly trained science teachers. The overall number of US STEM graduates is low, at only 6% of 24-year-olds surveyed in 2008. However, the number drops to 2.7% for African-Americans and 2.2% for Latinos, despite

the fact that interest in a STEM field is equal for these populations upon entering college (Mervis, 2010). This disparity continues in the workforce, where Hispanics, African Americans and American Indians made up only 9% of the STEM workforce, while comprising 26% of the US population, according to a 2008 report (National Science Board, 2012). The lack of diversity in the STEM workforce is concerning, as there are well-known benefits to having varied backgrounds and perspectives when it comes to global competitiveness (Ferrini-Mundy, 2013). However, until every child has the opportunity to engage in high quality science throughout their early education, we will continue to see a disproportionate number of LI/CLD students that are unprepared for, and often discouraged from pursuing, advanced STEM training and careers.

The new Common Core State Standards and Next Generation Science Standards present an opportunity for teachers, administrators, and policy makers to affect systemic change and potentially address some of the problems inherent to K-12 science education. With a new emphasis on college and career preparation as well as critical thinking and reasoning (Stage et al., 2013), these standards ensure "fewer, clearer, and higher" standards (Commission on Mathematics and Science Education, 2009). They are promising in their alignment with research on teaching and learning and their intentional coordination between literacy, math and science and engineering (Stage et al., 2013). While there is plenty of optimism surrounding the new standards, successful implementation will require significant investment from districts and administrators in long-term teacher professional development and support. Access to appropriate, high-quality training is often a challenge, and a recent survey from the National Science Teacher's Association (NSTA) indicated that science teachers report less access to discipline-specific compared with general professional development (Luft et al., 2009). These problems can be overcome with the growing number of online and virtual environments that allow teachers to interact and collaborate at a distance. The effectiveness of collaboration has been demonstrated with the "Lesson Study" approach commonly used in the Japanese educational system, and additional data from education research in the US report student learning and testing gains associated with teachers that have the

time and support to work in teams (Rosenholtz, 1989; Jackson and Bruegmann, 2009; Markow and Pieters, 2010). Some other interesting models are emerging, including the advent of the "Teacherpreneurs," a name given to expert teachers that split their time between the classroom and working in leadership roles to influence policy, assist administrators and mentor colleagues (Berry, 2013).

Our observations that learning gains over the course of one semester are relatively small for our topic under study highlight the need for science education reform at the post-secondary level as well. Our modest gains should not be surprising, given that conceptual change by many accounts is both difficult and slow. Others using a pre- and post-test design observed small to modest learning gains over the course of one semester, especially when the concepts evaluated were abstract in nature (Bowling et al., 2008b; Elrod, 2008). Students cannot see or touch genes, as they can a plant or a skeleton, making it difficult for them to visualize or create mental models of the biological processes inherent to genetics. Furthermore, the ability to communicate using scientific language is a relatively advanced skill that takes both time and practice to develop (Elrod, 2008), which may result in an underestimation of student knowledge using an open-response format like the Ten Word Test. The findings described in our study create ongoing problems for students as they progress through the curriculum. Just as Introductory Biology instructors assume students have learned and retained a basic level of genetics knowledge from high school, most faculty teaching upper division courses expect students to have gained significant knowledge of gene expression by the time they arrive in their courses. Yet all too often, faculty teaching these upper division courses find themselves forced to review the basics or forge ahead, leaving a portion of students to fall further behind.

To improve student learning, we, as college faculty, also need to undergo a conceptual change in the way we approach teaching and learning. Curriculum should be designed to help students with abstract and complex concepts, and this generally requires breaking from the lecture and infusing instructional methods in which students actively engage with the material and each other (reviewed by Allen and Tanner, 2005;

Udovic et al., 2002). In addition, faculty can benefit from the use of frequent and formative assessment that is aligned with their learning outcomes and instructional activities to provide an indication of which techniques/curricula are most effective (reviewed by Tanner and Allen, 2004). Information on prior knowledge and conceptual understanding can also be used to carefully select the material that students learn on their own, saving valuable class time for problem-solving and critical thinking exercises.

While additional data collection and analysis on a larger and unrelated population of students is necessary to draw generalizable conclusions, our findings have already prompted several curricular changes within the Introductory Biology course at our institution. First, we have increased the number of weeks devoted to the concept of gene expression in lecture from six to eight, allotting more time for students to build and synthesize the distinct concepts before encountering the more challenging material. In addition, we have re-designed an independent research project in which students apply the molecular and genetics concepts covered in lecture to a single gene disorder that they chose to study. This project now allows additional time for hands-on activities, including some that guide students in the development of conceptual models to explain the molecular basis of the disease they are researching. Instructors are also integrating more active learning exercises in lecture, and formative assessments are being employed to guide curriculum development and revision.

With the proper tools and training, K-12 teachers and college faculty can positively impact student learning at the course-level. However, it is necessary to work beyond the borders of our individual courses because the construction of scientific knowledge is a slow, progressive process. To support conceptual growth and the development of scientific thinking, we must coordinate and align curriculum and teaching practices between high schools, community colleges and university courses. This approach would allow students to construct scientific knowledge with curriculum that builds and reinforces at each level. These curricular and structural reforms are difficult and time-consuming, and the impacts on student learning are not always immediately recognized. These

changes will rely on district and institutional support in the form of professional development, time and resources, and greater incentives for devoting time and energy to teaching and learning. In our opinion, these reforms are critical at all levels of education if we are to prepare students for the exciting changes and scientific challenges of the 21st Century.

Acknowledgements

About the Authors

KELLY K. MCDONALD, Ph.D., is an Assistant Professor in the Department of Biological Sciences at California State University, Sacramento. Her areas of interest include the development and evaluation of instructional practices that facilitate learning and retention in introductory biology courses. She also coordinates several programs, including an NSF Robert Noyce grant, to recruit and train future K-12 science and math teachers.

email: mcdonald@csus.edu

JOSEPH M. GOMES, B S, is a first-year medical student at the American University of the Caribbean School of Medicine. Prior to beginning his career in medicine, he earned his B.S. in Biomedical Sciences at California State University, Sacramento, and subsequently worked as a primary and secondary school substitute teacher. He is looking forward to blending both areas of interest by pursuing an active role as both a future medical doctor and educator.

email: JosephGomes@students.aucmed.edu

REFERENCES

American Association for the Advancement of Science (1993). *Benchmarks for Science Literacy*. Oxford University Press, New York.

American Association for the Advancement of Science (2011). *Vision and change in undergraduate biology education, A call to action*. Retrieved from http://www.visionandchange.org/finalreport

Allen, D., & Tanner, K. (2005). Infusing active learning into the large-enrollment biology class: seven strategies, from the simple to complex. *Cell Biology Education (CBE) – Life Sciences Education, 4(1)*, 262-268.

Angelo, T. & Cross, K. (1993). Classroom assessment techniques: A handbook for college teachers. 2nd ed. San Francisco: Jossey-Bass Publishers.

Banet, E. & Ayuso, E. (2000). Teaching genetics at secondary school: A strategy for teaching about the location of inheritance information. *Science Education, 84(3)*, 313-351.

Berry, B. (2013). Teacherpreneurs: A bold brand of teacher leadership for 21st-Century teaching and learning. *Science, 340(6130)*, 309-310.

Bowling, B. V., Acra, E. E., Wang, L., Myers, M. F., Dean, G. E., Markle, G. C., & Huether, C.A. (2008). Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Genetics, 178(1)*, 15-22.

Bowling, B. V., Huether, C. A., Wang, L., Myers, M. F., Markle, G. C., Dean, G. E., & Jacob, G. A. (2008). Genetic literacy of undergraduate non-science majors and the impact of introductory biology and genetics courses. *BioScience, 58 (7)*, 654-660.

Browning, M. E., & Lehman, J. D. (1988). Identification of student misconceptions in genetic problem solving via computer program. *Journal of Research in Science Teaching, 25*, 747–761.

Cakir, M., & Crawford, B. (2001, January). *Prospective biology teachers' understanding of genetics concepts*. Paper presented at the Annual Meeting of the Association for the Education of Teachers in Science, Costa Mesa, CA. Retrieved from http://www.eric.ed.gov/PDFS/ED463956.pdf

California Department of Education (CDE). (2013, January 18). *Dataquest: 2012 STAR Test Results*. Retrieved May 7, 2013, from http://dq.cde.ca.gov/dataquest/

Chi, M. T. H., J. D. Slotta, & De Leeu, N. (1994). From things to processes: a theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27–43.

Chinn, C. A. & Brewer, W. F. (1993). The role of anomalous data knowledge acquisition: A theoretical framework and implications for science education. *Review of Educational Research, 63*, 1-49.

Commission on Mathematics and Science Education. (2009). *The opportunity equation: Transforming mathematics and science education for citizenship and the global economy*. New York, NY: Carnegie Corporation.

D'Avanzo, C. (2008). Biology concept inventories: overview, status, and next steps. *Bioscience, 58*, 1079–1085.

Dimitrov, D., & Rumrill, P. (2003). Pretest-Posttest designs and measurement of change. *Work, 20(2)*, 159-165.

Duit, R. & Treagust, D. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education, 25(6)*, 671-688.

Elrod, S. (2008, January). *Genetics concept inventory (GenCI) development*. Paper presented at Conceptual Assessment in Biology Conference II, Asilomar, California. Retrieved from http://bioliteracy.net/ CABS.html

Ferrini-Mundy, J. (2013). Driven by Diversity. *Science, 340 (6130)*, 278.

Garvin-Doxas, K., Klymkowsky, M. & Elrod, S. (2007). Building, using, and maximizing the impact of concept inventories in the biological sciences: report on a National Science Foundation sponsored conference on the construction of concept inventories in the biological sciences. *CBE Life Sciences Educucation*, 6, 277–282.

Gene Expression. 2013. Retrieved February 9, 2013, from Biology-Online website, http://www.biology-online.org/dictionary/Gene_Expression

Franke, G. & Bogner, F. (2011). Conceptual change in students' molecular biology education: Tilting at windmills? *The Journal of Educational Research, 104*, 7–18.

Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, C., & Wenderoth. M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE—Life Sciences Education, 6*, 132–139.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., & Wood, W. B. (2004). Scientific teaching. *Science, 304 (5670)*, 521-522.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*, 141-158.

Hildebrand, A. C. (1991). Confusing chromosome number and structure: A common student error. *Journal of Biological Education, 25,* 193-200.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics, 1(4),* 85-108.

Klymkowsky, M. (2010). Thinking about the conceptual foundations of the biological sciences. *Cell Biology Education (CBE) – Life Sciences Education, 9(1)*, 405-407.

Knight, J. K., & Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biology Education, 4 (4)*, 298-310.

Landis, C. R., Ellis, A. B., Lisenky, G. C., Lorenz, J. K., Meeker, K. & Wamser, C. C. (2001). *Chemistry ConcepTests: A pathway to interactive classrooms*. Upper Saddle River (NJ): Prentice Hall.

Lewis, J., J. Leach & Wood-Robinson, C. (2000) All in the genes? Young people's understanding of the nature of genes. *Journal of Biology Education, 34*, 74–79.

Lewis, J., & Wood-Robinson, C. (2000). Genes, chromosomes, cell division and inheritance: do students see any relationship? *International Journal Science Education, 22*, 177–195.

Libarkin, J. (2008). Concept Inventories in Higher Education Science. A manuscript prepared for the National Research Council Promising Practices in Undergraduate STEM Education, Workshop 2. Washington, D.C., October 13-14,

2008. Retrieved from https://www7.nationalacademies. org/bose/Libarkin_CommissionedPaper.pdf

Libarkin, J.C. & Anderson, S.W. (2005). Assessment of learning in entry-level geoscience courses: Results from the geoscience concept inventory. J*ournal of Geoscience Education, 53 (4)*, 394-401.

Longden, B. (1982). Genetics-Are there inherent learning difficulties? *Journal of Biological Education, 16(2)*, 135-140. doi: 10.1080/00219266.1982.9654439

Luft, J., Ortega, I., & Wong, S. (2009). *NSTA's State of Science Education Survey*. NSTA Reports: Arlington, VA.

Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. *Journal of Biological Education, 35(4)*, 183-189.

Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *International Journal of Science Education, 34(4)*, 200-205.

Markow, D., & Pieters, A. (2010). *The MetLife Survey of the American School Teacher: Collaborating for success.* New York: MetLife. Retrieved from www.metlife.com/ assets/cao/contributions/foundation/american-teacher/ MetLife_Teacher_ Survey_2009.pdf

Mazur, E. (1997). Peer instruction: Getting students to think in class. *AIP Conference Proceedings, 399(1)*, 981-988.

Mervis, J. (2010). New answers for increasing minorities in science. *ScienceInsider*. Retrieved from http://news. sciencemag.org/scienceinsider/2010/09/new-answers-for-increasing-minorities.html

National Academy of Sciences (NAS). 2010. *Rising above the gathering storm, revisited: Rapidly approaching Category 5.* Washington, DC: National Academies Press.

National Center for Education Statistics. (2000). *The nation's report card: science 2000*. Retrieved from http://nces. ed.gov/nationsreportcard/itmrls/searchresults.asp.

National Center for Education Statistics. (2005). *Degrees and other formal awards conferred.* Retrieved from http:// nces.ed.gov/programs/digest/d05/tables/dt05_249.asp.

National Science Board. (2012). *Science and engineering indicators 2012*. Arlington, VA: NSB.

Nazario, G., Burrowes, P., & Rodriguez, J. (2002). Persistent misconceptions: Using pre- and post-tests to identify biological misconceptions. *Journal of College Science Teaching, 31(5)*, 292-296.

Ohlsson, S. (2009). Resubsumption: a possible mechanism for conceptual change and belief revision. *Educational Psychololgy, 44*, 20–40.

Orcajo, T. I., & Aznar, M. M. (2005). Solving problems in genetics II: conceptual restructuring. *International Journal of Science Education, 27*, 1495–1519.

Radford, A., & Bird-Stewart, J. (1982). Teaching genetics in schools. *Journal of Biological Education, 16(3)*, 177-180. doi: 10.1080/00219266.1982.9654454

Roschelle, J. (1995). Learning in interactive environments: Prior knowledge and new experience. In J. H. Falk & L. D. Dierking (Eds.), *Public institutions for personal learning: Establishing a research agenda* (pp. 37-51). Washington, DC: American Association of Museums.

Rosenholtz, S. J. (1989). Workplace conditions that affect teacher quality and commitment: Implications for teacher induction programs. *Elementary School Journal, 89(4)*, 421-39.

Seymour, E. & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Smith, J., & Tanner, K., (2010). The problem of revealing how students think: Concept inventories and beyond. *Cell Biology Education (CBE) – Life Sciences Education, 9(1)*, 1-5.

Smith, M. K., & Knight, J. (2012). Using the genetics concept assessment to document persistent conceptual difficulties in undergraduate genetics courses. *Genetics, 191(1)*, 21-32.

Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The genetics concept assessment: A new concept inventory for gauging student understanding of genetics. *Cell Biology Education (CBE) – Life Sciences Education, 7(1)*, 422-430.

Stage, E. K., Asturias, H., Cheuk, T., Daro, P. A. & Hampton, S. B. (2013). Opportunities and challenges in next generation standards. *Science, 340 (6130)*, 276-277.

Strike, K., & Posner, G. (1992). A Revisionist Theory of Conceptual Change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of Science, Cognitive Psychology and Educational Theory and Practice* (pp. 147-164). New York, NY: State University of New York Press.

Stewart, J. H. (1982). Difficulties experienced by high school students when learning basic mendelian genetics. T*he American Biology Teacher, 44*, 80–84, 89.

Sundberg, M. (2002). Assessing student learning. *Cell Biology Education, 1(1)*, 11-15.

Tanner, K. D., & Allen, D. E. (2004). From assays to assessments— on collecting evidence in science teaching. *Cell Biology Education, 3*, 69–74.

Tanner, K. D., & Allen, D. (2005). Approaches to biology teaching and learning: understanding the wrong answers—teaching toward conceptual change. *Cell Biology Education, 4*, 112–117.

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition (2nd ed)*. Chicago: The University of Chicago Press.

Tobias, S. (1990). *They're not dumb, they're different: Stalking the second tier*. Tucson, AZ: Research Corporation.

Udovic, D., Morris, D., Dickman, A., Postlethwait, J., & Wetherwax, P. (2002). Workshop biology: demonstrating the effectiveness of active learning in an introductory biology course. *BioScience, 52*, 272–281.

Zirbel, E. (2004). Framework for conceptual change. *Astronomy Education Review, 3 (1)*, 62-76.