A New Integrative Approach to Evolution Education

PETER J. T. WHITE, MERLE K. HEIDEMANN, AND JAMES J. SMITH

Evolution is a difficult theory for students to understand. Part of the reason for this may be the tendency of instructors to teach evolution in the context of ecological systems, isolated from genetic and cellular mechanisms. To address this, we developed a set of integrative cases that consider the evolution of traits from the genetic scale to the macroecological scale. We implemented two of these cases in a biology course and tested their effectiveness using a pre-and postcourse assessment tool. Students who successfully learned evolution in a case context were better able to explain the molecular basis of mutation, to connect mutation to phenotypic change, and to make mechanistic links between genotypes and phenotypes. These gains were independent of the students' course achievement and precourse understanding of evolution. These findings support the hypothesis that students who acquire a molecular understanding of evolutionary mechanisms will have a better overall understanding of evolution.

Keywords: evolution, evolution education, integrative science, science education, case studies

Evolution is regarded as the most important unifying concept in biology (Dobzhansky 1973), but it presents a particularly difficult set of concepts for university-level science students to fully understand. For example, over the past 40 years, the percentage of undergraduate students who understand the basic tenets of natural selection has remained below 50% (Gregory 2009). Similarly, recent research indicates that fewer than 50% of science graduate students can correctly apply basic evolutionary principles (Alters and Nelson 2002, Gregory and Ellis 2009). The reasons for these deficiencies are complicated. However, at the core, the complexities of the principles underlying the theory of evolution make it challenging to fully comprehend. To clearly understand even the basic idea of natural selection, students must have a firm grasp of biological ideas ranging from the nature of genetic mutations to the population-level consequences of natural selection on novel phenotypes. However, we rarely spend the time, energy, and effort necessary to engage students in exploring the underlying cellular and molecular bases of the genotypes and phenotypes and how these characteristics are fundamental to evolutionary processes (Moore 2008, Smith JJ et al. 2009). In addition, when students learn cellular and molecular biology in their introductory or advanced courses, they rarely do so from an evolutionary biology perspective. This is unfortunate, because descriptions at the molecular level provide some of the most fascinating and definitive evidence supporting evolution.

Typically, university evolution instruction at the introductory level is delegated to organismal, macrobiology courses that include ecology, biodiversity, and Mendelian genetics. Instruction about evolution is focused on variation and natural selection and, perhaps, population genetics. For example, students may learn that beach mice in the genus *Peromyscus* have white coats as a result of natural selection against mice on the beach with dark coats (Vignieri et al. 2010). When teaching this, we might explain that predation is greater on those with dark coats because the predators use visual cues to find prey. However, the biochemistry and cell biology underpinning the variations are rarely discussed. Upper-level evolution courses may include some of the genetic underpinnings (e.g., mutations, alternative alleles, population genetics) but less frequently address the related changes and differences in biochemistry and cell biology.

We are not aware of teaching resources or curricular materials available to instructors that track the evolution of traits from molecular genetics to the array of population phenotypes. We hypothesize that students who do engage with these kinds of resources will develop a better overall understanding of evolution. We propose this hypothesis in a spirit similar to those of studies in which the acceptance of evolution has been shown to be positively correlated with genetic literacy (Miller et al. 2006). Therefore, our work is driven by a primary problem in evolution teaching and learning: a core curriculum that lacks a comprehensive genes-to-selectable-phenotype approach and in which there is little discussion of how biochemistry and cell biology are important for understanding evolution.

BioScience 63: 586–594. ISSN 0006-3568, electronic ISSN 1525-3244. doi:10.1525/bio.2013.63.7.11

Case approach to this problem

We developed a set of cases that connect curricular components and introduce molecular and cellular mechanisms into undergraduate evolution education; two cases are reported here (boxes 1 and 2). This is a radical departure from most—if not all—of the present curricular materials in undergraduate evolution education, which are typically focused on one or a few aspects of the biology of a study system that are involved in evolutionary processes but rarely provide a truly integrated framework or perspective for learning evolution. We would argue that these parallel but separate curricular treatments lead to compartmentalized, disconnected knowledge that prevents the students from developing a deep understanding of evolution and of general biology.

Box 1. The case of seed taste and shape evolution in the garden pea (*Pisum sativum*).

The first case involves the garden pea plant (*Pisum sativum*) and why domesticated populations often have sweet-tasting seeds that are wrinkled in appearance when dried, whereas wild populations have starchy-tasting seeds that are round. The case includes biochemical details of starch synthesis, with a focus on the role of a key protein, the starch branching enzyme (SBE1). Functional SBE1 produces a starchy seed that will stay round when dried. The case provides details of the mutation that occurred in the gene encoding SBE1 (Smith AM 1988, Bhattacharyya et al. 1990), as well as the resulting changes in sugar and starch synthesis (Edwards and ap Rees 1986, Guilfoile 1997). Plants with this mutation produce sweeter seeds with higher water content than those without the mutated *Sbe1* gene; therefore, the seed will wrinkle when it is dried (Hedley et al. 1986). The two gene variants in his case are denoted by the alleles *R*, which encode a functional SBE1 protein, and *r*, which encodes a nonfunctional SBE1 protein. Ancient farmers preferred the sweeter pea, and as a result, the *r* allele became fixed in cultivated pea plant populations (Zohary and Hopf 1973, Ljuština and Mikić 2010, Smýkal et al. 2011). The *R* allele remains more common in wild populations. When it is incorporated into the introductory biology curriculum, this case can facilitate student learning in molecular genetics, cell biology, population genetics, and ecology (artificial selection), which, together, largely explain the evolution of sweeter pea seeds. In addition, this case provides an interesting link to Gregor Mendel's seminal genetic research: The round versus wrinkled phenotype was one of the characteristics he examined experimentally (van der Waerden 1968).

Box 2. The case of fur color evolution in the beach mouse (*Peromyscus polionotus*).

In this case, the reason that some populations of beach mouse (*P. polionotus*) have light fur, whereas others have dark fur, is examined. Fur color is determined by the quantity of the pigment eumelanin produced by melanocytes, cells located at the base of hair follicles (Barsh 1996). Eumelanin production is correlated with the stimulation of the MC1R protein, which, in turn, triggers biosynthesis of this pigment (Barsh 1996). Details of biochemical and cellular processes are presented in the case. Mice that live on light-sand beaches often have a single nucleotide substitution mutation in their *Mc1r* gene that reduces MC1R affinity for binding the melanocyte-stimulating hormone (Hoekstra et al. 2006). This results in less eumelanin production and, ultimately, in a mouse with lighter-color fur. Light fur results in more camouflage on light sand and, therefore, offers protection against visual predators in those habitats. Consequently, the *Mc1r* mutation is selectively advantageous and is, indeed, more commonly found in mouse populations that live on light-sand beaches (Mullen et al. 2009, Steiner et al. 2009). Conversely, the *Mc1r* mutation is not beneficial in dark-soil habitats, where having light fur can result in increased predation, especially from visual predators, and the mutation is rarer in dark-soil habitats (Kaufman 1974, Vignieri et al. 2010). Therefore, there is a relationship between the frequency of occurrence of the mutated allele in various populations and the soil color of the habitat in which each lives (Mullen et al. 2009). When incorporated into the introductory biology curriculum, this case facilitates student learning in molecular genetics, cell biology, ecology, and population genetics, which, together, help explain the evolution of light fur in beach mice.

Our primary criterion for selecting a system to be developed into an instructional case was that the pathway from genes to proteins to selectable phenotypes be well defined and documented in the primary literature. We found surprisingly few such study systems that have been well described (although the number is increasing). In addition, it was important that the Mendelian genetics for the selected cases be well documented with respect to the genetic loci involved, the different alleles segregating at these loci, and their dominance relationships. Thus, each of the cases would illustrate evolutionary processes by beginning with a DNA mutation and ending with the fixation of alternate phenotypes in separated populations. Along the way, the cases allow us to explore the cell biology of the phenotype (e.g., how proteins make some hair cells white and others brown) and the nature of the selective pressure (e.g., artificial selection for sweeter peas) involved in each evolutionary scenario. The details of the cases that we have developed can be found in full at *www.evo-ed.com*. We focus on two of these cases in the present article: In the first, pea seed taste and shape are examined, and in the second, mouse fur color is examined (boxes 1 and 2). The cases themselves represent authentic, factual, and comprehensive instructional activities. They were created using backward-design principles (McTighe and Thomas 2003), with content development guided by clearly articulated learning goals and objectives, which, in turn, were based on the development team's years of teaching experience.

The two cases were implemented in the spring semester of 2012, in an

introductory cell and molecular biology course (LB145) at Lyman Briggs College (LBC), a small residential college for science majors at Michigan State University (MSU; Sweeder et al. 2012, Luckie et al. 2013). LBC is physically part of the MSU campus; students take their introductory courses at LBC, followed by advanced courses in other science, technology, engineering, and mathematics units. Information about the spring 2012 iteration of LB145 can be found at *www.evo-ed.com/lb145s12/index.html*.

Here, we report our assessment of the effects of the case materials on student understanding of evolution. We employed pre- and postcourse assessments to test students' ability to solve problems by applying evolutionary principles, using an assessment tool that was developed through iterative analyses of student feedback. We then applied a regression model to test for a relationship between the students' scores on final exam questions about case material and their achievement in the postcourse evolution assessments.

Our study was conducted in three phases. In the first phase, we designed, tested, and refined an open-ended assessment tool to gauge the students' ability to apply evolutionary principles to solve problems. In the second phase, we implemented the cases in an introductory biology course. In the third phase, we tested how well the students had learned the case material and related that data to postcourse assessment scores. Before we began data collection, our project was reviewed and approved by the MSU Institutional Review Board (IRB# X10-1086). The project was deemed exempt from the protocols put in place for studies on human subjects.

The Assessment Tool for Evaluating Evolution Knowledge. We developed an open-ended assessment tool, the Assessment Tool for Evaluating Evolution Knowledge (ATEEK; box 3), using an iterative design process (Bishop and Anderson 1990) to determine whether the students could solve evolution-oriented questions. First, our project team (the present authors) determined the essential concepts that we considered important for a complete understanding of evolution. Next, we created the open-ended ATEEK, which we field tested in two introductory cell and molecular biology courses (BS161 and LB145) and in an advanced-level evolution course (ZOL445). The project team evaluated the students' answers to the ATEEK questions. When the pattern of answers to any given question differed from the types of answers we expected from introductory- and advanced-level students, we proposed hypotheses to explain how the students may have misunderstood the question. We then revised the ATEEK questions, where that was appropriate, to satisfy the hypothesized misunderstandings and repeated the field test on a cohort of undergraduate biology students. This cycle of testing, evaluation, and revision of our tool occurred three times: once in the spring of 2011 and twice in the fall of 2011. The final fall 2011 ATEEK solicited the types of answers that we expected from introductory- and advanced-level students; no further question modifications were made. The data reported in this article were generated using that final version of the ATEEK.

The students' answers to the ATEEK questions were scored on a three-point scale, on which an answer was scored as 0 if it was wrong or mostly wrong, as 1 if it was partially correct, and as 2 if it was correct or mostly correct in terms of the targeted ideas. A more-detailed explanation of our scoring rubric development and examples of students' answers for each score and for each question are available in supplemental appendix S1, available online at *http://dx.doi.org/10.1525/bio.2013.63.7.11*.

The first question (box 3) probes whether students can describe the molecular connection between genotypes and phenotypes. Specifically, do they know that genes are the information for proteins that operate on a cellular level to produce observed phenotypes? We used a Mendelian framework for this question, because many students' understanding of genetics seems to be rooted in a Mendelian paradigm in which genotypes are connected to phenotypes without a well-defined mechanism. The second question is analogous to the first, except that it prompts students to

Box 3. The questions used in the pre- and postcourse Assessment Tool for Evaluating Evolution Knowledge (ATEEK) test.

Students were given 15 minutes to complete the ATEEK as an independent in-class exercise. For the completion of the precourse and postcourse ATEEKs, 0.94% of the students' final grade was awarded for each (for an approximate total of 1.9%). The ATEEK was administered in class sessions 2 and 30.

1. Jaguars (large predatory cats) can have an orange coat or a black coat. Orange jaguars have either two *G* alleles (genotype *GG*) or one *G* allele and one *g* allele (genotype *Gg*), whereas black jaguars have two *g* alleles (genotype *gg*).
   When a jaguar has the genotype *gg*, what happens so that a black coat is produced?
2. Toxican mushrooms contain a toxin that causes vomiting when ingested. Recently, some Toxican mushrooms were found that did not produce the toxin.
   Describe in detail what might have happened **at the molecular level** so that these mushrooms no longer produce this toxin?
3. The non-poisonous Toxican mushroom has become more frequent in mushroom populations and poisonous Toxican mushrooms have become rare.
   Define Natural Selection and use it to explain this scenario.
4. Considering genetic mutation,
   (a) Describe, at the molecular level, what a mutation is.
   (b) Use your answer from part (a) to describe the *process* whereby a mutation results in a change at the phenotype level.

Note: For a detailed discussion of the questions and scoring rubric, please see appendix S1, available online at *http://dx.doi.org/10.1525/bio.2013.63.7.11*.

Table 1. Time line and description of instructional activities and assessments pertaining to evolution cases in the course LB145, spring semester 2012.

Session	Case	Activities
4	P	Minute paper: "Why are wrinkled peas wrinkled?" Debriefed with a single PowerPoint slide following peer discussion. Minute paper: "What is a protein?" Debriefed following peer discussion.
11	P	Slide set (12 slides) on the relationship of pea shape to taste. Looked at in the contexts of higher sucrose content, osmotic potential, and water content. In-class exercise*: Why do wrinkled peas have a higher water content?* Asked as a before–after instruction question.
21	P, M	Focus of class session: "mutations and their impacts on the cell." Slide set (2 slides): nucleotide differences between the *R* and *r* alleles in peas. Slide set (10 slides) on the effect of C→T mutation on the MC1R protein in mice.
26	P	Take-home questions: Consider Mendel's peas: "When we examine the genetic basis of round versus wrinkled peas, we find that it is a single-locus genetic system governed by two alleles, *R* and *r*. (6 points) Given that, create diagrams that illustrate and explain the difference, at the DNA level, between an *R* allele and an *r* allele. illustrate the chromosomes containing *R* and *r* alleles in individuals who are (a) homozygous *R*, (b) homozygous *r*, and (c) heterozygous at this genetic locus. Show the chromosomes both before (show the homologous pair) and after DNA replication (show the homologous pair as dyads). explain the basis of the round or wrinkled phenotype that would be associated with being (a) homozygous *R*, (b) homozygous *r*, and (c) heterozygous at this genetic locus."
27	M	Slide set (16 slides) on the genetics and cell biology of the MC1R protein in mice. Minute paper: "What is Evolution?" Debriefed with peer–peer and class discussion. Think–pair–share: "What are the possible genotypes of light and dark mice?" In-class exercise: *Examine data on the relationship between genotypes and phenotypes. Do the data confirm or refute the hypothesis that the* Mc1r *gene is responsible for fur-color phenotypes?* In-class exercise: *Examine a pedigree of mice and assign each a genotype based on its phenotype. If the pedigree is representative of population fur colors, what are the light and dark allele frequencies in your population? If a predator ate the darkest (or lightest) mice, what is your new allele frequency?* Clicker question, followed by class discussion: "Can differences in fur color among beach mice be fully explained by different genotypes of the *Mc1r* gene?"
28	M	Slide set (10 slides) on the ecology and population genetics of beach mice. In-class exercise: *Predict the allele frequencies in different beach mouse populations on the basis of their fur colors. Were your predictions correct? If not, theorize why.* Class discussion: *What is the effect of predation on beach mouse populations?*

Note: Activities were associated with the Mendel's peas case (P) or the beach mouse case (M).

Figure 1. Time line of course activities. The numbers indicate the session number. Abbreviations: E, Assessment Tool for Evaluating Evolution Knowledge test; F, final exam; M, instruction involving the beach mouse case; P, instruction involving the Mendelian pea case. See the text for further details.

provide an explanation for the genetic basis for the loss of a trait rather than the genetic basis for two different already-established traits. The third question, related to the second, asks students to use macroscale processes (i.e., natural selection) to describe the change in the allele frequencies within populations. The final question (in two parts) asks students to define the word *mutation* and to explain how it can result in a phenotypic change. Although question 4 is somewhat similar to questions 1 and 2, during the design phase of the ATEEK, we discovered that students would use the word *mutation* to describe how new phenotypes arise (i.e., primarily with regard to the nontoxic mushrooms in question 2). However, they would rarely give any genetic detail to explain what, in their mind, a mutation was. We therefore designed question 4 to assess whether they understood the genetics of mutation and to provide ourselves with a frame of reference to judge the responses to question 2 when the word *mutation* was used.

Implementation of cases. We implemented the *Pisum* seed and *Peromyscus* fur color cases in LB145 in the spring semester of 2012. LB145 is a five-credit course, with a typical enrollment of 60–80 students per semester (primarily sophomores). It is the second in a two-course sequence of introductory biology. The first course, LB144, is an introductory organismal biology course that most students take in the preceding semester. Students in LB145 attend two 80-minute class sessions and two 110-minute lab sessions per week. In the spring of 2012, when the cases were implemented, the course was taught by JJS; PJTW coordinated the lab component of the course and assisted with case implementation in the course lectures. There were 30 class sessions in LB145 over the duration of the semester. The 30 sessions were characterized by a mix of interactive engagement, clicker questions, and class discussions to facilitate learning. The class schedule is available at *www.evo-ed.com/lb145s12/classnotes.html*. The *Pisum* case was taught (by JJS) in class sessions 4, 11, 21, and 26 (table 1, figure 1); the

Peromyscus case was taught by PJTW in class sessions 21, 27, and 28. The *Peromyscus* case was designed to be a capstone case for the LB145 course (table 1).

Evaluating case knowledge: *Peromyscus* fur color. We developed six final-exam questions to assess students' knowledge of the evolution of *Peromyscus* fur color, touching on different subdisciplines of biology (box 4). The questions probed whether the students could explain the nucleotide and amino acid differences associated with different melanocortin-1-receptor (*Mc1r*) alleles, could describe the function of the MC1R protein in eumelanin synthesis, could integrate their knowledge of natural selection with the knowledge that different fur color phenotypes are produced, could translate their knowledge of the MC1R protein into population-level phenomena, could describe genotype to phenotype connections, and could create a plausible (genetic or cellular) scenario in which dark fur alleles would result in light fur. The students earned up to 1.6% of their course grade for answering all of the questions correctly. The questions were graded (by PJTW) on the three-point scale described above, where a 0 was given for an incorrect answer, a 1 was awarded for a partially correct answer, and a 2 was awarded for a correct answer. This resulted in a maximum of 12 for the *Peromyscus* case score.

Evaluating the conceptual impact of the cases. The precourse ATEEK was administered in LB145 in the 2nd class session (out of 30 total class sessions); the postcourse ATEEK was administered in the 30th class session (table 1). Each ATEEK was completed during a normal class session and took approximately 15 minutes of class time. A modest number of grade points were awarded for the completion of each ATEEK (for a total value of 0.94% of the final course grade, based on participation, not on the correctness of the responses). After the students completed the pre- and postcourse ATEEKs in LB145, PJTW scored their answers for each question as was previously described, for a maximum score of 10.

Data analyses. We completed three analyses. First, we assessed the difference in score between identical questions in the pre- and postcourse ATEEKs. Second, we analyzed whether there were any relationships between student improvement on different pairs of questions on the ATEEK. Third, we used multiple regression to relate postcourse ATEEK scores to *Peromyscus* case scores and students' overall achievement in the course (the final course grade).

We analyzed scores on identical questions in the pre- and postcourse ATEEKs to assess gain: If a student's score improved on a question from the pre- to the postcourse ATEEK, it was counted as a *gain* (e.g., from 0 to 1 or from 1 to 2). If a student's scores did not improve, it was counted as *no gain* (therefore, the *no gain* category also included losses). The scores from the two parts of question 4 were pooled to generate a single score for question 4.

We used 2-cell × 2-cell contingency tables to determine whether the students' gains on each question of the ATEEK were independent of the gains on other questions. This resulted in a total of six comparisons. Fisher's exact tests were used for the comparisons, because a preliminary analysis indicated that there were very small expected values for the *no gain* categories in some of the questions; Fisher's exact test gives more reliable *p*-values under such conditions (Upton 1992).

We used paired *t*-tests to analyze the students' scores on identical questions on the pre- and postcourse ATEEKs. Five separate *t*-tests were calculated, one for each question. These analyses allowed us to determine whether the pre- to postcourse ATEEK scores were statistically different, but it did not link any independent variables to the outcome.

Our final analysis was a multiple regression of three independent variables (see below) with respect to postcourse ATEEK scores. A multiple regression is a single test that reveals the relationship between each independent variable and the dependent variable. The coefficient of each independent variable included in the model is calculated so as to maximize that variable's ability to explain the variance of the dependent variable. This is achieved by minimizing the residual sum of squares of the dependent variable for all possible computations of independent variable coefficient values. Independent variable coefficients

Box 4. Test questions on the beach mouse case.

Students were asked six questions on the final exam pertaining to the beach mouse case. Up to 1.6% of a student's final grade could be earned for correct answers. The total score on these questions was the student's *Peromyscus* case score.

1. What is the difference between the R^{67} and C^{67} alleles? Specifically, how do their nucleotide sequences differ, and what is the resulting difference in the amino acid sequences they produce?
2. What is the role of the MC1R protein in eumelanin pigment synthesis in *Peromyscus polionotus* (a) dark fur populations and (b) light fur populations?
3. What role does natural selection play in determining the coat color of *Peromyscus polionotus* populations? Briefly describe some of the studies that have been done to support this.
4. Do populations of *Peromyscus polionotus* with light fur tend to have a high C^{67} allele frequency? Why or why not?
5. What phenotypes do the following genotypes typically code for?

   $R^{67}R^{67}$, $R^{67}C^{67}$, $C^{67}C^{67}$
6. Can the genotype $R^{67}R^{67}$ result in a phenotype different from the one you listed above? Why or why not?

are determined automatically by the computer software (in this case, R; R Development Core Team; *www.r-project.org*), using complex matrix algebra. The R^2 value of a multiple regression model represents the total proportion of dependent variable variance that is explained by the suite of independent variables. Significance was determined for each independent variable on the basis of *t*-values. The first independent variable was the *Peromyscus* case score; we hypothesized that the students who had a better grasp of the details and concepts involved in the *Peromyscus* case would be better able to apply evolutionary concepts to solve the problems presented on the ATEEK. The second independent variable was the precourse ATEEK score; we hypothesized that student achievement on postcourse ATEEK might be positively affected by their knowledge or ability to answer ATEEK questions prior to case instruction. Pre- and postcourse scores on assessment tools like the ATEEK are often integrated into an overall instrument gain score. We chose to include the precourse ATEEK score as an independent variable rather than as part of a gain score, because the former method allowed us to enumerate the statistical effect of the precourse ATEEK scores and to assess the variable's statistical importance. The third independent variable was the students' overall achievement in the course; we hypothesized that students who excelled in learning introductory cell and molecular biology course material would be better able solve problems on the postcourse ATEEK, because much of it requires an application of cellular and molecular biological concepts. Our multiple regression model was computed using the following equation:

postcourse ATEEK = *Peromyscus* case score
+ precourse ATEEK
+ course grade.

Student performance on ATEEK and *Peromyscus* case questions

There were 66 students enrolled in LB145 in the spring of 2012. Of these, 61 completed both the pre- and the postcourse ATEEK (5 did not complete the postcourse ATEEK). A further two students did not complete the case exam questions (a portion of the questions was left blank). Because we were not able to determine whether these omissions were due to a lack of time or to a deficit of knowledge, we also omitted these two students from the analysis.

Of the remaining students ($n = 59$), the average score on the precourse ATEEK was 2.9 out of 10 (standard deviation (SD) = 2.4), and the average score on the postcourse ATEEK was 5.9 out of 10 (SD = 2.6). In total, 48 of the 59 students had postcourse ATEEK scores that were better than their precourse ATEEK scores. Six students had identical pre- and postcourse ATEEK scores, and five students had lower postcourse than precourse ATEEK scores. The average *Peromyscus* case score of these 59 students was 7.7 out of 12 (SD = 2.5). Their average course grade was 84% (SD = 6.8%).

Comparing gains on pairs of ATEEK questions. The Fisher's exact tests did not reveal any significant relationships between the gains that students made from the pre- to the postcourse ATEEK for any pair of questions. However, interpretation of *p*-values should be done with caution, because low sample sizes may reduce the statistical power to detect significant relationships (Upton 1992). If one takes the current *p*-values as suggestive of which relationships are most likely to become significant with larger data sets, the results seem to reflect a conceptual connection. The Fisher's exact test comparison of questions 1 and 2 yielded a *p*-value of .11. This indicates that students may have been more likely have the same score—gain or no gain—for those two questions. As was noted above, these questions are analogous but differ in focus such that question 1 examines an established trait and question 2 examines the loss of a trait. The remaining *p*-values were .76 for the comparison of question 1 and question 3, .99 for that of questions 1 and 4, .99 for that of questions 2 and 3, .45 for that of questions 2 and 4, and .36 for that of questions 3 and 4. Overall, there appeared to be sufficient statistical independence to make multiple questions useful as an assessment tool. Furthermore, even if the statistical independences with a larger sample size were to be found to be weak for some comparisons, it would still be useful to use multiple questions as a check on the quality of learning at each conceptual level.

Learning gains: Pre- versus postcourse ATEEK scores. There were significant differences in the average student's pre- and postcourse scores for four of the five ATEEK questions (figure 2). At the end of the course, the students could more successfully describe the link between genotypes and phenotypes (question 1), they could provide a molecular mechanistic explanation for the disappearance of a trait (question 2), they could accurately describe what a mutation

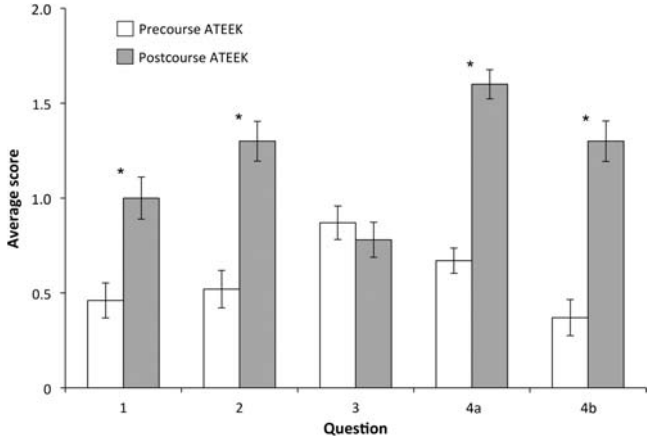

Figure 2. Average scores for each question on the pre- and postcourse Assessment Tool for Evaluating Evolution Knowledge (ATEEK) test. The asterisks denote significance at $p < .001$ (determined by t-test), and the error bars represent the standard error.

was (question 4a), and they could describe how a mutation can affect a phenotype (question 4b). There was a small but not significant drop in ATEEK scores in question 3, which asked the students to use natural selection to solve a problem.

Predicting postcourse ATEEK scores. The precourse ATEEK score and the *Peromyscus* case score were both highly significant positive predictors of the postcourse ATEEK score (table 2). A student's overall course grade was not a significant predictor. Standardized coefficients showed that the *Peromyscus* case score was an important independent variable in the model and was a strong positive predictor of postcourse ATEEK scores. Overall, this model explained 40% ($R^2 = .40$) of the variance in postcourse ATEEK scores.

Conclusions

There was a significant relationship between the students' learning of the last integrative case, demonstrated on the pertinent final exam questions, and their performance on the postcourse ATEEK, which was a deeper assessment of their understanding of evolutionary principles. This relationship was independent of both performance on the precourse ATEEK and overall course achievement (table 2).

Our analysis was based on three measures: (1) a pre- to postcourse ATEEK score comparison for each item (figure 2), (2) the independence of items from one another on the ATEEK (table 2), and (3) a statistically significant explanation of the variance in postcourse ATEEK scores by the students' performance on other measures. These three, taken together, indicate that the cases, as they were presented in the instructional paradigm reported here, helped the students learn about the cellular and molecular underpinnings of evolution. Specifically, the cases helped the students understand that a genetic mutation results in an altered protein that can change phenotype frequencies within a population in ways that affect its selective value. These findings support the hypothesis that students who have an understanding of genetic, molecular, and cellular evolutionary mechanisms will have a better overall understanding of evolution. In analyzing the pre- and postcourse ATEEK scores (figure 2), we showed that the students made gains for most items, which were largely designed to elicit student understanding of molecular connections to phenotypes and behavior. The responses to question 3, which is related to natural selection, showed no change after the instruction and introduction of the cases. The students came in with some knowledge about ecology and natural selection from the previous course in the introductory biology sequence and, apparently, did not change their thinking about them.

When evolution is taught only in relation to natural selection, students may begin to perceive other elements of the evolutionary system as black boxes. They may be asked to "imagine that there is a mutation that gives mice light fur" or to "suppose that pea plant seeds became sweeter over time." Molecular mechanisms are often given short shrift so that instruction can be focused on natural selection and ecological systems. However, learning outcomes can be undermined if a student decides that there is no reasonable mechanism for light-furred mice, for example, to be produced from a population of dark-furred mice. Our expectations are that when students learn about an entire evolutionary system as presented in the cases, they come to see evolution as a synthesis of biological ideas and concepts across the curriculum and that this will strengthen their understanding and retention of each of the constituent ideas (Alters and Nelson 2002, Nelson 2007).

Developing an assessment tool (i.e., the ATEEK) that measured student knowledge of evolution was challenging. We used standard techniques for developing this tool (Bishop and Anderson 1990, Angelo and Cross 1993) and revised it after analyzing feedback from students outside of this study. Notably, the ATEEK does not ask students to tell us what evolution is. Our rationale for this choice was that a given definition of evolution does not necessarily mean that a student understands the cross-disciplinary nature of evolutionary theory. Rather, evolution is a theory that requires a broad understanding of biology. For example, the answer "any change in the frequency of alleles within a gene pool from one generation to the next" (Curtis and Barnes 1989, p. 974) may be correct from a population-genetics perspective but does not show that the student understands the mutational origin of alternate alleles and does not address the role of natural selection in evolutionary processes. In a more recent textbook, Campbell and Reece (2007) defined evolution as "descent with modification; the idea that living species are descendant of ancestral species that were different from the present-day ones; also defined more narrowly as the change in genetic composition of a population from generation to generation" (p. G-14). This is a more

Table 2. Multiple linear regression using precourse Assessment Tool for Evaluating Evolution Knowledge (ATEEK) score, overall course grade, and* Peromyscus *case score to predict postcourse ATEEK score.

Variable	Coefficient	Standardized coefficient	Standard error	*t*	*p*
Precourse ATEEK score	0.41	0.37	0.12	3.4	.0013
Overall course grade	0.044	0.11	0.054	0.82	.42
Peromyscus case score	0.37	0.36	0.13	2.8	.0077

Note: $F(3,55) = 13.76$, adjusted $R^2 = .40$. The R^2 value indicates that this combination successfully explained 40% of the variance in postcourse ATEEK scores. Standardized coefficients show the relative effect that each independent variable has on the calculation of predicted postcourse ATEEK score.

complete definition. However, if a student were to cite this definition verbatim, it would provide little indication whether the student actually understood how evolution *works*. The ATEEK is therefore focused on applying tenets of evolution rather than on definitions of it. Can a student explain how a new allele or genotype comes to be? Can a student connect that genotype to a selectable phenotype? Can he or she explain how natural selection acts on those phenotypes? We argue that if students can make these connections, they then understand a great deal about evolution, regardless of their ability to give a technical definition.

Two subject areas omitted from the ATEEK were population genetics and speciation. A strong case can be made that a complete understanding of evolution requires the knowledge that the result of natural selection is a change in allele frequencies within populations and that cumulative changes occurring independently in two different populations over a long period of time can result in two different species. Population genetics is a significant component of the *Peromyscus polionotus* case, and we will be designing a question to encapsulate this in future iterations of the ATEEK. The two cases discussed here do not directly address the issue of speciation, because we chose to focus on the emergence of novel traits within populations rather than the emergence of new species. Although the latter is often a result of the former, the cases that we implemented were single-species systems. However, we have now developed an additional case that has a strong phylogenetics component—the case of trichromatic vision in Old World primates—in which it may be appropriate to introduce an assessment question asking students how trait evolution is linked to the evolution of new species (see *www.evo-ed.com*).

The cases described here include molecular genetics and cellular biology, in order to help students understand how novel phenotypes arise, starting from the most basic building blocks of life. Some instructors (e.g., Smith JJ et al. 2009) have already come to realize that incorporating molecular aspects into biology curricula might lead to students' better understanding evolution. Andrew Moore, then manager of the Science and Society Program at the European Molecular Biology Organization, called for a stronger infusion of the molecular aspects of evolution into secondary-school curricula in Europe (Moore 2008). Moore argued that "many of the most fascinating and definitive examples supporting evolution—those made in the past four decades using gene-sequencing technology and bioinformatics—are largely absent from European secondary-school curricula." The same can certainly be said for curricula in the United States, and the problem spills into university curricula, as well. The "molecular stuff" is crucial and should ideally be infused throughout the curriculum. We have made our cases freely available on our Web site, *www.evo-ed.com*, where we include descriptions of the cases, slides for teaching, and interactive app-style simulations to engage students and to help them learn different aspects of each case.

Acknowledgments

We thank Miles Loh for his help assembling case materials and online resources; Alex Shingleton, Chuck Elzinga, and Gerald Urquhart for piloting our cases and pilot assessment tool in the courses BS161, ZOL445, and LB144, respectively; and Craig E. Nelson for his ongoing support of our project. Partial support for this work was provided by the US National Science Foundation's Transforming Undergraduate Education in Science, Technology, Engineering, and Mathematics program (award no. 1043876). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the US National Science Foundation.

References cited

Alters BJ, Nelson CE. 2002. Perspective: Teaching evolution in higher education. Evolution 56: 1891–1901.

Angelo TA, Cross KP. 1993. Class Assessment Techniques: A Handbook for College Teachers, 2nd ed. Jossey-Bass.

Barsh GS. 1996. The genetics of pigmentation: From fancy genes to complex traits. Trends in Genetics 12: 299–305.

Bhattacharyya MK, Smith AM, Ellis THN, Hedley C, Martin C. 1990. The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. Cell 60: 115–122.

Bishop BA, Anderson CW. 1990. Student conceptions of natural selection and its role in evolution. Journal of Research in Science Teaching 27: 415–427.

Campbell NA, Reece JB. 2007. Biology, 8th ed. Pearson.

Curtis H, Barnes NS. 1989. Biology, 5th ed. Worth.

Dobzhansky T. 1973. Nothing in biology makes sense except in the light of evolution. American Biology Teacher 35: 125–129.

Edwards J, ap Rees T. 1986. Metabolism of UDP-glucose by developing embryos of round and wrinkled varieties of *Pisum sativum*. Phytochemistry 25: 2033–2039.

Gregory TR. 2009. Understanding natural selection: Essential concepts and common misconceptions. Evolution: Education and Outreach 2: 156–175.

Gregory TR, Ellis CAJ. 2009. Conceptions of evolution among science graduate students. BioScience 59: 792–799.

Guilfoile P. 1997. Wrinkled peas and white-eyed fruit flies: The molecular basis of two classical genetic traits. American Biology Teacher 59: 92–95.

Hedley CL, Smith CM, Ambrose MJ, Cook S, Wang TL. 1986. An analysis of seed development in *Pisum sativum* II: The effect of the *r*-locus on the growth and development of the seed. Annals of Botany 58: 371–379.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 313: 101–104.

Kaufman DW. 1974. Adaptive coloration in *Peromyscus polionotus*: Experimental selection by owls. Journal of Mammalogy 55: 271–283.

Ljuština M, Mikić A. 2010. A brief review on the early distribution of pea (*Pisum sativum* L.) in Europe. Field and Vegetable Crops Research 47: 457–460.

Luckie DL, Smith JJ, Cheruvelil KS, Fata-Hartley C, Murphy CA, Urquhart GR. 2013. The "Anti-cookbook laboratory": Converting "canned" introductory biology laboratories to multi-week independent investigations. Tested Studies for Laboratory Teaching: Proceedings of the Association for Biology Laboratory Education, vol. 34. Forthcoming.

McTighe J, Thomas RS. 2003. Backward design for forward action. Educational Leadership 60: 52–55.

Miller JD, Scott EC, Okamoto S. 2006. Public acceptance of evolution. Science 313: 765–766.

Moore A. 2008. Science teaching must evolve. Nature 453: 31–32.
Mullen LM, Vignieri SN, Gore JA, Hoekstra HE. 2009. Adaptive basis of geographic variation: Genetic, phenotypic and environmental differences among beach mouse populations. Proceedings of the Royal Society B 276: 3809–3818.
Nelson CE. 2007. Teaching evolution effectively: A central dilemma and alternative strategies. McGill Journal of Education 42: 265–283.
Smith AM. 1988. Major differences in isoforms of starch-branching enzyme between developing embryos of round- and wrinkled-seeded peas (*Pisum sativum* L.). Planta 175: 270–279.
Smith JJ, Baum DA, Moore A. 2009. The need for molecular genetic perspectives in evolutionary education (and vice versa). Trends in Genetics 25: 427–429.
Smýkal P, et al. 2011. Phylogeny, phylogeography and genetic diversity of the *Pisum* genus. Plant Genetic Resources 9: 4–18.
Steiner CC, Römpler H, Boettger LM, Schöneberg T, Hoekstra HE. 2009. The genetic basis of phenotypic convergence in beach mice: Similar pigment patterns but different genes. Molecular Biology and Evolution 26: 35–45.
Sweeder RD, Jeffery KA, McCright AM. 2012. Lyman Briggs College: An innovative living–learning community for STEM education. Quality Approaches in Higher Education 3: 7–14.
Upton GJG. 1992. Fisher's exact test. Journal of the Royal Statistical Society 155: 395–402.
Van der Waerden BL. 1968. Mendel's experiments. Centaurus 12: 275–288.
Vignieri SN, Larson JG, Hoekstra HE. 2010. The selective advantage of crypsis in mice. Evolution 64: 2153–2158.
Zohary D, Hopf M. 1973. Domestication of pulses in the Old World. Science 182: 887–894.

Peter J. T. White (pwhite@msu.edu) and James J. Smith are affiliated with Lyman Briggs College, at Michigan State University, in East Lansing. Merle K. Heidemann is affiliated with the Center for Integrative Studies in General Science and the Department of Geological Sciences, also at Michigan State University.