Genetic Literacy of Undergraduate Non–Science Majors and the Impact of Introductory Biology and Genetics Courses

BETHANY VICE BOWLING, CARL A. HUETHER, LIHSHING WANG, MELANIE F. MYERS, GLENN C. MARKLE, GARY E. DEAN, ERIN E. ACRA, FRANCIS P. WRAY, AND GEORGE A. JACOB

With the advancement of genetic information and technologies, there is an increasing need for a genetically literate public. This study looks critically at student learning and at the current instruction of genetics in introductory non–science major biology and genetics courses at the undergraduate level. A new diagnostic tool, the Genetic Literacy Assessment Instrument, was administered pre- and postcourse to more than 300 students in six introductory nonmajor courses that emphasize genetics to varying degrees. Current data from students in these courses show a precourse average score of 43 percent correct on the inventory. Postcourse scores increased only modestly, to an average of 49 percent. In this article, we discuss the impact of teaching methods and course content on scores, as well as student learning in the different content areas of genetics. The results suggest that further studies in genetics education are needed to better understand the effect of teaching methods on achieving genetic literacy.

Keywords: assessment, biology education, genetics education, undergraduate

Over the last several decades, the role of genetic technologies in health and public policy has persistently increased (Miller 1998, Kolsto 2001), and new knowledge in genetics continues to have significant implications for individuals and society (Lanie et al. 2004). Enhancing the general public’s understanding of genetics will improve communication regarding genetic information and technologies, and help to ensure its appropriate use (Haga 2006).

In spite of the increased exposure to genetics issues, recent studies indicate that the general public has a relatively low understanding of genetics concepts (Petty et al. 2000, HGC 2001, Lanie et al. 2004, Bates et al. 2005, Miller et al. 2006), and that genetic information presented informally through different types of media is not always correct (Grinell 1993, Lanie et al. 2004). For these reasons it can be difficult for people without a solid foundation in the basic concepts to distinguish valid genetic information from misinformation (Jennings 2004).

In the United States, the National Science Education Standards provide the basis for state science standards and include standards on genetics for each of the clustered grade levels, K–4, 5–8, and 9–12 (NRC 1996). In grade levels K–4 and 5–8, the basic concepts of inheritance and reproduction are introduced, and in grades 9–12, the molecular basis of heredity and biological evolution are covered. Students graduating from high school should leave with a very basic, but reasonably broad, understanding of genetics. However, despite these standards, the 2000 National Assessment of Educational Progress tested approximately 49,000 US students, and on average only about 30 percent of 12th graders could completely or partially answer genetics questions correctly (NCES 2000).

Bethany Vice Bowling (e-mail: bowlingb2@nku.edu) is an assistant professor in the Department of Biological Sciences at Northern Kentucky University in Highland Heights. Carl A. Huether is a professor in the Department of Biological Sciences, Lihshing Wang is an associate professor in Educational Studies and Leadership, Melanie F. Myers is an assistant professor and Erin E. Acra is a graduate student in the Genetic Counseling Program, Glenn C. Markle is a professor in the Division of Teacher Education, and Gary E. Dean is an associate professor in the College of Medicine, and all at the University of Cincinnati in Ohio. Francis P. Wray is an associate professor of biology at Raymond Walters College, University of Cincinnati, in Blue Ash, Ohio. George A. Jacob is a biology teacher at St. Xavier High School in Cincinnati, Ohio. © 2008 American Institute of Biological Sciences.
Postsecondary education provides further opportunity for educating the public about genetics. More than two million individuals graduate with associate’s or bachelor’s degrees each year in the United States (NCES 2004). Approximately 10 percent of these degrees are in the life sciences and health fields (NCES 2005). The other 90 percent of graduates may receive some genetics instruction through general education courses. A study of institutions indicated that more than 90 percent of those surveyed have general education requirements (Hurtado et al. 1991). However, there has been little exploration or evaluation of genetics knowledge and education for undergraduates, especially non-science majors. Two studies conducted in Israel found that high school and university students had a number of alternative conceptions (or misconceptions) regarding genetics concepts and failed to see the relationships among concepts (Marbach-Ad and Stavy 2000, Marbach-Ad 2001). Also, Fagen (2003) examined the genetics vocabulary of undergraduate students and reported that the students begin introductory classes with significant gaps in their understanding.

Other science disciplines, primarily physics and chemistry, have worked extensively to evaluate students’ preconceptions and knowledge gains over the duration of introductory courses. This has been accomplished through the development of concept inventories, or diagnostic tests, which are multiple-choice tests for evaluating knowledge of a particular set of concepts (Wandersee et al. 1994). Some notable assessments include the Force Concept Inventory (FCI) (Hestenes et al. 1992), the Astronomy Diagnostic Test (Hufnagel 2002), and the Chemistry Concepts Inventory (Mulford and Robinson 2002). Research in undergraduate biology education has more recently moved toward assessment of preconceptions and knowledge gains with the Concept Inventory of Natural Selection developed in 2002 (Anderson et al.), and efforts to produce a biology concept inventory for biology-major courses are under way (Klymkowsky and Garvin-Doxas 2007).

Existing inventories serve as a standardized way to measure student knowledge and changes in knowledge over time. Hake (1998) used the FCI to show how courses using interactive engagement had much greater gains in student knowledge than courses using more traditional teaching methods. Including more than 62 courses and 6000 students, this study provided a convincing argument that reform in undergraduate physics education was necessary. Since then, a number of studies have illustrated that undergraduate science instruction—not just in physics—is significantly more effective when student-centered and interactive teaching strategies are incorporated (NRC 2003). Additionally, students tend to be more confident about their abilities in science when a course uses active learning (ModularCHEM Consortium 2002, Wilke 2003), and self-efficacy has been shown to have a significant impact on student learning (Multon et al. 1991).

In response to the need for a genetically literate society and the evaluation of undergraduate genetics education, the Genetic Literacy Assessment Instrument (GLAI) was developed and evaluated (Bowling et al. 2008). The instrument is based on the central concepts in genetics that an undergraduate non-science major should understand, as determined by a subcommittee of the American Society of Human Genetics (Hott et al. 2002). With the availability of a valid and reliable assessment tool, this study sought to determine the effect of introductory biology and genetics courses on students’ genetics knowledge. The research effort had two goals: (1) to determine the extent to which undergraduate courses in biology and genetics influence the genetic literacy of students completing these courses, and (2) to determine the effect of pedagogy and course content on student learning within the courses.

**Study design**

This study involved the use of a naturalistic setting, meaning that the arrangement of courses, instructors, and students were not manipulated in any way. All of the instruments and procedures were reviewed and approved by the governing institutional review boards. Participants in the study were self-selected from a convenience sample. The biology departments of five colleges and universities provided the names and contact information of individuals teaching introductory biology courses for non-science majors. A total of seven instructors were invited to participate in the study, and six instructors at five different institutions agreed to take part. During the 2006–2007 academic year, a total of 630 students in six classes (five introductory biology courses and one introductory genetics course [Course C]) were invited to participate; 395 completed the precourse inventory, 330 completed the postcourse inventory, and 287 students completed both (46 percent of all students). Table 1 describes the demographic characteristics of each course.

**Genetic Literacy Assessment Instrument**

The GLAI is a 31-item multiple choice assessment tool that addresses 17 concepts identified as central to genetic literacy within six content areas: the nature of the genetic material, transmission, gene expression, gene regulation, evolution, and genetics and society. The development and evaluation of the GLAI have been discussed elsewhere (Bowling et al. 2008). It was found to be a valid and reliable tool for assessing the genetic literacy of undergraduate students.

The online survey tool SurveyMonkey was used to administer the GLAI, which allowed students to complete the pre- and postcourse assessments on their own time. Students were given an incentive of extra credit for completion of both assessments, and were sent an e-mail with a hyperlink that connected to the online assessment. They had one week to complete the precourse instrument. Toward the end of the quarter/semester, the students received an e-mail notice directing them to the postcourse assessment. Again, the students had one week to complete the assessment. There were at least seven weeks between pre- and postcourse administration of the assessments.
Additional data collection: Course data

Observations of the courses were conducted using the Reformed Teaching Observation Protocol (RTOP), a standardized instrument for measuring the degree to which classroom instruction is reformed (MacIsaac and Falconer 2002). Reformed teaching, in contrast to traditional teaching, includes the use of inquiry-based activities, small group learning, class discussion, and other constructivist teaching approaches (MacIsaac and Falconer 2002). Traditional teaching methods are teacher-centered and rely heavily on lectures. The RTOP consists of 25 statements on classroom instruction, with observers rating the class for each statement on a scale from 0 (never occurred) to 4 (very descriptive). The overall scale is from 0 to 100; the more reformed the course is, the higher the score. Observations to collect RTOP data took place throughout the quarter/semester and were unannounced; the dates of observations were scheduled at the convenience of the observers and on days when tests were not being given. Each course was observed twice, once by the primary author and again by the primary author and an additional individual. All observers were trained through an online tutorial to use the RTOP instrument.

Data analysis

The impact of the courses on students’ GLAI scores was analyzed by comparing pre- and postcourse scores for each course using a paired, one-tailed t test. Cohen’s effect size ($d$) was also calculated for each of the courses (Cohen 1988). Normalized gain (NG) scores, defined as the change in score divided by the maximum possible increase ((postcourse score – precourse score) ÷ [100 – precourse score]) (Hake 1998), were calculated for each student, and these scores were averaged for each course (table 1). Hake’s (1998) work with the FCI has shown NG to be an objective measure of student learning in that it is not correlated with the precourse score (Hake 1998, 2002). We used NG as the dependent variable in a regression analysis of the effects of the independent variables of time dedicated to genetics content (“time on genetics”) and pedagogy (RTOP scores) on student learning. This analysis allowed us to determine the predictive power of the two variables (Field 2005).

Results

A total of 287 students completed both the pre- and postcourse GLAI. As seen in table 1, the courses differed in their participation rates. In five of the courses, more than 40 percent of the students completing the course participated, whereas in course A, less than 25 percent did so. In terms of demographics, significantly more females than males participated in the study (table 1), which is consistent with the makeup of the courses as shown. Additionally, the proportion of ethnic minorities who participated in the study was similar to the proportion that completed the courses. These data suggest that the sample of students participating in the study was representative of those students completing the courses.

<table>
<thead>
<tr>
<th>Course</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Originally enrolled</td>
<td>72</td>
<td>18</td>
<td>89</td>
<td>331</td>
<td>104</td>
<td>40</td>
</tr>
<tr>
<td>Completed course</td>
<td>64</td>
<td>18</td>
<td>69</td>
<td>280</td>
<td>92</td>
<td>40</td>
</tr>
<tr>
<td>Participated in study (percentage)$^a$</td>
<td>15 (23.4)</td>
<td>14 (77.8)</td>
<td>31 (44.9)</td>
<td>166 (59.3)</td>
<td>42 (45.7)</td>
<td>19 (47.5)</td>
</tr>
</tbody>
</table>

$^a$ Percentage of students completing the course who participated in the study.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females in class (percentage)</td>
<td>A (100.0)</td>
</tr>
<tr>
<td>Females in study (percentage)</td>
<td>100.0</td>
</tr>
<tr>
<td>Ethnic minorities in class (percentage)</td>
<td>–</td>
</tr>
<tr>
<td>Ethnic minorities in study (percentage)</td>
<td>6.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GLAI mean scores</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precourse</td>
<td>A (39.8)</td>
</tr>
<tr>
<td>Postcourse</td>
<td>44.3</td>
</tr>
<tr>
<td>Normalized gain</td>
<td>0.06</td>
</tr>
</tbody>
</table>
among the areas (table 3). Questions on gene regulation showed the highest percentage correct for the precourse at 56.9 percent, and scores also exhibited the least amount of change from pre- to postcourse. The least average percentage correct for both pre- and postcourse were for evolution, 25.6 and 30.5, respectively. Questions on transmission genetics showed the most change from pre- to postcourse, with an increase of 11 percentage points. The corresponding average numbers of hours spent during class time are also listed in table 3. The most time was dedicated to transmission genetics (an average of 6.3 hours), whereas the least amount was spent on concepts within gene regulation.

**Effects of course content and teaching methods**

The teaching methods and amount of time spent on genetics concepts varied considerably among the courses (table 4). It should be noted that courses C, D, E, and F are at institutions on a quarter system, for which there are generally three instruction time each week for 10 weeks. Also, courses D and E are taught for one quarter in a course that is composed of a three-quarter sequence; and data were obtained for the second quarter, in which virtually all of the genetics content is taught. For course E, the instructor reported spending 30.0 hours of the class on the six main content areas tested by the GLAI (the entire course), whereas hours taught in the other courses varied between 6.5 to 22.0. Teaching methods, as measured by the RTOP, were also quite diverse, but the most reformed course did not exceed a score of 61 out of the possible 100. The interrater reliability between the RTOP reviewers was calculated at $r = 0.69$.

The linear regression analysis shown in table 5 provides insight into the impact of teaching methods and time on genetics on student scores. Step 1 includes only the RTOP as an independent variable; step 2 includes the RTOP and time on genetics in the model. These results indicate a small but significant effect of the RTOP on NG scores ($R^2 = 0.029$, slope = 0.003, $p < 0.005$). Time on genetics appears not to systematically explain the variation.

**The impact of the courses on students’ GLAI scores**

Relatively small gains on the GLAI were seen from pre- to postcourse in all of the courses. Of the six, only two showed a medium effect, according to Cohen’s effect size, on student knowledge gain. Similar studies in physics and chemistry using concept inventories have also found small changes from pre- to postcourse (Hake 1998, Mulford and Robinson 2002, Coletta and Phillips 2005). There is some concern regarding students’ seriousness when responding to concept inventories, as there is frequently no credit given for correct answers. Although they were encouraged to complete the GLAI to the best of their ability, it is possible that some students did not, and their lack of sincerity could have negatively affected the change from pre- to postcourse scores. However, a study by Henderson (2002) showed little difference between graded and ungraded FCI scores.
The percentage of students who answered questions correctly varied among the content areas. An average of 43.5 percent of students correctly answered postcourse questions on the nature of the genetic material. These questions address some very basic concepts: DNA (deoxyribonucleic acid) is the genetic material of virtually all organisms, mutations result in genetic variation, DNA is organized into chromosomes, and so on. Most of these concepts are listed among the high-school science standards. In addition, only 30 percent of the students correctly answered items about concepts in evolution on the postcourse inventory, and little change was seen from pre- to postcourse. The misconceptions that students frequently have regarding evolution present a challenge in nonmajor biology instruction (Anderson et al. 2002). Perhaps that is why we see such little change in the courses observed here. The results shown in table 3 indicate that instructors may be surprised to know their students are possibly leaving their introductory courses lacking an understanding of some very basic concepts, similar to the astonishment expressed by physics instructors upon using the FCI for the first time (Hestenes et al. 1992).

### Table 3. Percentage of students answering GLAI items correctly (pre- and postcourse) and course time spent on each of the six content areas.

<table>
<thead>
<tr>
<th>Content area</th>
<th>Nature of the genetic material</th>
<th>Transmission</th>
<th>Gene expression</th>
<th>Gene regulation</th>
<th>Evolution</th>
<th>Genetics and society</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precourse</td>
<td>36.4</td>
<td>52.6</td>
<td>37.5</td>
<td>56.9</td>
<td>25.6</td>
<td>50.0</td>
</tr>
<tr>
<td>Postcourse</td>
<td>43.5</td>
<td>63.6</td>
<td>46.8</td>
<td>56.5</td>
<td>30.5</td>
<td>52.9</td>
</tr>
<tr>
<td>Change</td>
<td>7.1</td>
<td>11.0</td>
<td>9.3</td>
<td>–0.4</td>
<td>4.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Average time (hours)</td>
<td>2.6</td>
<td>6.3</td>
<td>2.0</td>
<td>0.9</td>
<td>2.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>

The effect of course content and teaching methods

Although there are a number of factors involved in student learning in an undergraduate science course, the teaching methodology used by instructors has been receiving increased attention. The most notable study is Hake’s (1998) research using the FCI, which illustrated that physics courses using interactive engagement produced much greater increases in student knowledge than courses using more traditional teaching methods. Since then, the development of the RTOP has advanced measures for characterizing teaching methods.

Several other studies have been conducted using both the RTOP tool and concept inventories to better identify the impact of teaching methods on student learning. Two studies show a greater correlation than this study does between the RTOP and NG (table 5; \( r = 0.17, p < 0.005 \)). Lawson and colleagues (2002) and MacIsaac and Falconer (2002) report correlations of 0.80 and 0.33, respectively. The range of RTOP scores in our study was restricted to between 33 and 60; this small range may contribute to our finding of a lower correlation. Also, MacIsaac and Falconer’s findings indicate RTOPs around 70 yielded NG scores above 0.40, whereas courses with RTOPs 60 and below had NG scores below 0.20. In our study, the highest RTOP was 61, with an NG score of 0.17. Perhaps there is a threshold for instruction, and the knowledge gain is much greater in significantly reformed courses than in only moderately reformed courses.

### Table 4. RTOP, time on genetics, pre- and postcourse GLAI mean scores, and normalized gains by course.

<table>
<thead>
<tr>
<th>Course</th>
<th>RTOP</th>
<th>Time on genetics (hours)</th>
<th>GLCI mean score</th>
<th>Normalized gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precourse</td>
<td>Postcourse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>33</td>
<td>10.0</td>
<td>39.8</td>
<td>44.3</td>
</tr>
<tr>
<td>B</td>
<td>42</td>
<td>6.5</td>
<td>49.5</td>
<td>62.9</td>
</tr>
<tr>
<td>C</td>
<td>58</td>
<td>16.0</td>
<td>54.1</td>
<td>59.3</td>
</tr>
<tr>
<td>D</td>
<td>52</td>
<td>22.0</td>
<td>39.4</td>
<td>46.5</td>
</tr>
<tr>
<td>E</td>
<td>15</td>
<td>30.0</td>
<td>48.2</td>
<td>48.2</td>
</tr>
<tr>
<td>F</td>
<td>61</td>
<td>12.0</td>
<td>36.3</td>
<td>46.1</td>
</tr>
</tbody>
</table>

### Table 5. Regression analysis of normalized gain scores (N = 287).

<table>
<thead>
<tr>
<th>Step</th>
<th>Parameter</th>
<th>Standard error of B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>–0.032</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>RTOP</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>0.087</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>RTOP</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Time on genetics</td>
<td>–0.004</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Note: \( R^2 = 0.029 \) for step 1; \( ΔR^2 = 0.005 \) for step 2. a. \( p < 0.005 \).
be taken when interpreting data on “time on genetics,” since the instructors’ self-reporting varied enormously (6.5 to 30 hours) among the courses that are classified as introductory biology. Additionally, the topics covered may not have been directly relevant to the six concepts on which the GLAI is based. In future studies, verification of instructors’ self-reports, perhaps in the form of a daily inventory of content discussed, would help validate these data.

There are certainly other factors not explored in this study that affect student learning. For instance, course B, conducted at a private institution with a class of 18 students, showed the highest NG of 0.25, with an RTOP score of 42 and a reported 6.5 hours of time spent on genetics. It is possible that some characteristics of courses have a greater effect on the students’ learning than the instruction observed in the class. Additionally, a study published by Coletta and Phillips (2005) reported that innate “scientific reasoning ability,” as tested by Lawson’s Classroom Test of Scientific Reasoning, may have just as much of an impact on students’ learning in physics as reformed instruction, suggesting that other factors may need to be considered in future research in this area.

Conclusions
There are a number of limitations to this study, some of which have been discussed above. One major limitation concerns the sample, which was a convenience sample and not randomly controlled. Further research conducted in courses using significantly reformed, moderately reformed, and traditional teaching methods would provide more insight into the role of instruction on student knowledge gains in genetics, particularly if genetics content were held more constant among them. Two other limiting factors are the small sample size of six courses and the lack of standardization regarding the instructors’ knowledge of genetics.

This research also exhibits a number of strengths. To our knowledge, it is the first study to consider both teaching methods as measured by the RTOP and the time dedicated to a particular content area when measuring student gains using a standardized assessment. Other studies, primarily in physics education, have considered these two variables independently. This study directly enhances genetics education research by demonstrating the use of the GLAI for assessing genetic literacy in undergraduate students. It also begins to look critically at current genetics education at the undergraduate level and suggests that further studies in this area are needed to better understand the effect of teaching methods on student learning in genetics.

We encourage anyone who wants to obtain the full version of the GLAI, or to collaborate in continued research with the instrument, to contact the authors.

Acknowledgments
We are extremely grateful to the educators who allowed their courses to be used for this study, and to the students who participated.

References cited
doi:10.1641/B580712
Include this information when citing this material.