INTRODUCTION

With the advent of faster and less expensive sequencing technologies (1), more genomic data are being generated every day. At present, there are not enough trained bioinformaticians to deeply analyze the data. The need for people with skills to analyze big data sets will increase, implying that students with training in this area will be desirable in the workforce (2). Moreover, these publicly available data present a good opportunity for faculty and students to get involved in bioinformatics research. Many freely available bioinformatics tools are easy to master (3); thus, anyone with a computer, access to the Internet, and basic training in this field can contribute to novel genomics research.

However, many biology faculty are uncomfortable teaching bioinformatics concepts because they lack training and/or are unaware of available tools and resources (4). To help overcome these challenges, we developed a faculty professional development workshop, the Genome Solver Project (3; https://genomesolver.qubeshub.org) to teach some of the fundamentals of bioinformatics using easily accessible online tools. The workshop contains information on sequence databases and how to navigate them, including the National Center for Biotechnology Information suite of tools (https://ncbi.nlm.nih.gov) and the Integrated Microbial Database at the Joint Genome Institute (https://img.jgi.doe.gov), global and local sequence alignments, pairwise and multiple sequence alignments, and a brief introduction to phylogenetics. The bioinformatics tools we teach include BLAST (5) for pairwise alignments; MUSCLE for multiple sequence alignments (6, 7); and MEGA for phylogenetics (8).

The workshop is broken up into short lectures, with time to practice each of the tools and for conversation among the participants to encourage collaboration and exchange of information. In the early versions of the workshop, we also explicitly discussed an annotation pipeline called RAST (9). The workshop is aimed at biologists and biological research questions; participants need no experience with coding or scripting. We also developed a Community Science Project for faculty and their students to join, and the project was...
explicitly discussed in the later versions of the workshop, beginning in 2015. Although the workshop content has been modified and refined based on changes to the online tools and participant feedback since it was first offered in the summer of 2012, all iterations of the workshop up through the summer of 2017 have served to introduce faculty to the same tools. To date, more than 250 faculty, as well as some graduate students and postdoctoral fellows, at ~125 different institutions have received training through Genome Solver workshops.

With the fifth anniversary of the project, we sought to evaluate the impact of the program on faculty, and, to the extent possible, on student learning. First, we reviewed the data collected from each iteration of the workshop as to whether the participants felt the workshop training and materials were useful to them as they projected future curricular reforms. Second, we conducted a follow-up survey to learn how much bioinformatics content was included in changes in curriculum delivery by workshop participants. In other words, we sought to better understand whether the workshop influenced faculty’s ability to incorporate bioinformatics into their course offerings.

In addition to measuring the extent to which faculty felt enabled to teach bioinformatics, we aimed to assess the effect on student performance relative to bioinformatics content. Therefore, we undertook an evaluation of the outcomes of the project on students. A small group of workshop participants agreed to administer a pre- and post-course quiz to their students to measure performance gains after bioinformatics instruction. To complement the student performance data, we also conducted 30-minute phone interviews with these faculty to gain a better understanding of the individual characteristics of their courses, including how much bioinformatics content was included and what topics were covered. Our results in total suggest that many, but not all, faculty find the workshops useful for encouraging incorporation of bioinformatics content in their courses, and students of faculty who implement bioinformatics content in their courses show increased quiz scores following bioinformatics instruction.

METHODS

To assess the impact of the Genome Solver Project on faculty and students, we gathered data in several different areas. All data collection was deemed “exempt” by Georgetown’s Institutional Review Board.

Workshop recruitment and survey

Workshop participants were initially recruited primarily by word-of-mouth through existing networks, including the Genomics Education Partnership (http://gep.wustl.edu), the Genome Consortium for Active Teaching (http://gcat.davidson.edu), and the American Society for Microbiology Biology Scholars program (http://directory.biologyscholars.org).

Subsequently, we advertised through our former website (http://genomesolver.org) and at national conferences including the American Society for Microbiology Conference on Undergraduate Education (ASMCUE) and the American Society for Cell Biology Annual Meetings, with posters and short talks. We asked the participants at the start of the workshop about their institution. In total, about 20% of participants came from community colleges, 15% from baccalaureate-granting institutions, 20% from master’s-granting institutions, and 35% from doctoral institutions. About 5% were from a variety of other institutions, including medical schools, government institutions, or non-profit institutions, while the remainder could not be identified as associating with a particular institution. About 10% of the participants identified as coming from minority-serving institutions. We developed a survey that workshop participants (n = 277) took directly after the end of the workshop (see Fig. 1). Averages from the five-point Likert scale data were calculated as described (https://higheredassessment.wordpress.com/2012/06/28/weighted-average-standard-deviation/).

Workshop content

The workshop initially consisted of 1.5 days of content, alternating between didactic lectures on bioinformatics content and practice with the tools discussed. In later iterations, the content was trimmed so that the entire workshop could be delivered in a single day. In all cases, participants were given readings about the Human Microbiome Project and information about BLAST (Basic Local Alignment Search Tool) (5) before the workshop. During the workshop, a series of six short lectures was delivered:

• Introduction to the Genome Solver Project
• Introduction to Databases*
• Genome Annotation*
• Comparative Genomics*
• Phylogenetics*
• The Horizontal Gene Transfer Project

Lectures denoted with asterisks also had hands-on exercises for practice with the tools introduced in that lecture. The final lecture served as both a summary of the tools described and an introduction to a community science project we developed in later iterations of the workshop for faculty and their students to join. All current materials are available at QUBES (http://genomesolver.qubeshub.org) and at CURENet (https://serc.carleton.edu/curenet/collection/218072.html).

Faculty survey

A follow-up online survey (Appendix 1) was designed and deployed to all who participated in a workshop between 2012 and 2016 (n = 259) in order to better understand
how the training might have influenced incorporation of bioinformatics into their work with students, both in the classroom and as a means for independent research. The survey link was sent to all previous workshop participants three times beginning November 29, 2016, with reminders on December 5 and December 8, resulting in 47 complete responses (18% response rate). As an incentive, respondents were included in a draw for a $50 Amazon gift card. Descriptive quantitative analysis was performed on the survey using Excel. Coding of the qualitative survey data was conducted manually using an inductive approach, allowing the categories and themes to emerge from the data. One author, McWilliams, conducted the initial qualitative coding, which was then reviewed by Rosenwald, who contributed to naming the categories and to making decisions on coding responses about which there were questions (10).

Pre- and post-course student quiz data

A set of 20 multiple-choice questions that tested basic concepts related to molecular biology and bioinformatics was developed (Appendix 2). This quiz was motivated by the ones developed by Genomics Education Partnership (GEP) discussed in Shaffer et al. (11), which asked questions specifically about bioinformatics and genomics. One important difference between the GEP quiz and ours is that ours focused more on bacterial genomics where the GEP quiz focuses on eukaryotic genomics, reflecting the disparate nature of the two projects. Faculty members at five schools agreed to administer this quiz to their students before and after bioinformatics instruction. Three schools were doctoral institutions and the remaining two were primarily master’s-granting institutions. One of the schools involved was Georgetown University, the home institution of the developers of the Genome Solver workshops. These instructors (Arora and Rosenwald) also developed the multiple-choice quiz. The student performance dataset consists of 640 student responses. Data are reported as number of correct responses pre-course compared with post-course. The proportion of correct responses for each question was determined for the pre- and post-course test answers. A proportions t-test and one-way ANOVA were calculated using JMP v12 software (SAS, Raleigh, NC). For all statistical analyses, p < 0.05 was considered significant. (Mathur conducted the statistical analyses.)

Faculty interviews

We conducted 30-minute phone interviews with each of the faculty who helped us to collect student learning data. The interviews were recorded and transcribed (McWilliams). The interview protocol used to guide the interviews is shown in Appendix 3. Four of the faculty had attended a Genome Solver workshop and, upon return to their home institution, devised ways to utilize the bioinformatics tools and exercises from the workshop. The other two, the instructors at Georgetown University who co-created the original Genome Solver training workshop and materials, used these materials in their courses. The faculty were asked specifically about class size, amount of bioinformatics focus, number of instructors, and class makeup (i.e., number of majors versus nonmajors), as well as details on bioinformatics-related assignments and the extent to which
they were used with students. The courses ranged from large, introductory courses to smaller, upper-level courses and included a sophomore learning community.

RESULTS

Faculty development in bioinformatics

Many life science faculty earned their terminal degrees before the advent of bioinformatics as a major field of endeavor for biologists (4). We therefore developed a workshop to teach some of the fundamental tools used for bioinformatics. As described in more detail in Methods, the workshop contains information about databases, pairwise and multiple sequence alignments, gene annotation, and a brief introduction to phylogenetics, tools that can be downloaded for free or that are web-based.

At the end of each workshop, participants were asked to fill out an online survey to provide feedback for improvements. As part of the survey, participants were asked to rate a number of statements about the utility of the workshop for them and their future plans on a five-point Likert scale. The average response was in the “agree” to “strongly agree” category for each of the prompts, suggesting that the participants had a very positive experience with the workshop overall (average rating of 4.83 out of 5). In addition, participants agreed the workshop contributed to their confidence in applying these techniques to their own research (average 4.3 out of 5). They also thought Genome Solver would help their institutions move ahead with genomics curricula (average 4.27 out of 5). Most planned to use the materials and connections made at the workshop in their genomics teaching and research endeavors (average 4.38 out of 5). Finally, most participants agreed the workshop was important in helping them plan their teaching of genomics (average 4.34 out of 5) (Fig. 1).

Changes in courses and curricula take time, however, and making such changes can be challenging, even when faculty are eager to incorporate new material and tools. To determine the extent to which workshop participants had been able to incorporate bioinformatics into their teaching, we sent an online follow-up survey (Appendix 1) to all faculty who had participated in a Genome Solver workshop from 2012 to 2016. We sought to determine whether the participants had been able to incorporate bioinformatics into their work with students after taking the Genome Solver workshop. To this end, the survey asked:

- whether participants had used any of the tools and materials from the Genome Solver workshop with students in their classes and/or in independent student research
- which tools and materials they used
- what might have prevented them from using Genome Solver materials, and
- their most valuable take-aways from the workshop.

The survey link was emailed to all workshop participants (2012–2016), with two reminders, resulting in 59 initial responses. After culling incomplete responses, there were 47 usable responses in the final dataset. Approximately 57% of respondents had used material from the workshop, used the workshop materials after modification to suit their needs, or used tools discussed in the workshop. Approximately 32% had used this information in classes, 2% with their independent research students, and 23% with both students in class and research students. The most frequently cited tool used was BLAST, followed next by various databases.

When asked about their most valuable take-away from the workshop, responses clustered in three main thematic areas (Table 1):

- improved understanding of tools through hands-on activities
- ideas for integrating bioinformatics content in the undergraduate classroom and student research
- the opportunity to network and learn from other teachers’ experiences.

The first theme—that of “training” faculty in easy-to-use bioinformatics research tools—was an original goal of the Genome Solver project, supported by the website and online community forum. The second theme—that of networking and sharing information—was also an original goal, to boost faculty confidence and skills in order to motivate and prepare them to introduce students to these tools. The third theme—that of enabling faculty to apply what they had learned about bioinformatics research tools in their own classrooms—was the ultimate goal of the project. Genome Solver project leaders hoped that participants would leave armed with tools and skills, but also ideas and inspiration, to try at their home institutions. We found that several participants (n = 11) reported taking ideas from the workshop and modifying them to suit their own curricular needs (Table 2).

Yet many faculty did not implement the tools and ideas from the workshop; approximately 43% of the survey respondents did not use any information from the workshop in their classes (Table 3). Reasons stated included not currently teaching courses where this content fits well; discovering other sources of bioinformatics material that suited their needs better; and lacking the confidence to teach these materials in a class. With respect to this last point, certain comments suggested some faculty needed even more basic training than what the workshop provided or ongoing support beyond the workshop (Table 3). Thus, even though we started at quite a basic level, some individuals nevertheless left the workshop lacking comfort with the material or the confidence to implement it with students. In future efforts, we plan to develop online mentoring using resources on the QUBES hub, the current site of Genome Solver (http://genomesolver.qubeshub.org). We also noted that no respondents reported facing difficulties from their department with
Among the survey respondents who said they had not implemented any of the tools and materials from the workshop in their work with independent research students, there were also a number of interesting findings revealed in the reasons for not doing so (Table 4). The top reason was that although their research includes some bioinformatics, the Genome Solver materials were not appropriate for this research. The next most cited reasons were not doing research, not having a lab, or not having research students. However, we also found that a number of individuals said they lacked confidence in working with Genome Solver materials. Nevertheless, a number of respondents indicated that they still plan to use the materials with undergraduate research students in the future. A final group indicated that bioinformatics research does not align with their existing research. Similar to the responses about not using Genome Solver materials in the classroom, “fit” and confidence emerged as the reasons why implementation was difficult.

Student development in bioinformatics

We partnered with a group of workshop participants to assess student performance gains by administering a pre- and post-course quiz to their students before and after bioinformatics instruction. We designed a set of 20 multiple-choice online questions that tested basic concepts related to molecular biology and bioinformatics (see Appendix 2). The same quiz was administered to the students at the beginning and end of the course.

Analysis of the pre- and post-course quiz demonstrated students had significant increases in performance related to bioinformatics concepts. In a comparison of pre- and post-course scores for all students, results for 15 of the 20 questions showed a significant increase in correct responses for the post-course test scores (Fig. 2). Of the remaining five questions, two were basic biology questions included as controls (Questions 1 and 2, Appendix 2) and, as predicted, do not show a significant increase. Question 10 describes Hidden Markov Models and is an “all of the above” answer, so it is likely students got this question partially right by choosing one of the other answers. Question 17 discusses the concept of synteny or the order of genes on a chromosome; we suspect many faculty may not discuss this concept in detail. Finally, in Question 18, which asks about transversions, we note that scores on the pre-course quiz are already quite high, so little improvement was noted.

We then analyzed the overall scores by school. All five schools showed a significant increase in the post-course mean compared with the pre-course mean (Fig. 3). The
Faculty interviews

Faculty who partnered with us on the quiz also agreed to a 30-minute telephone interview to complement the information collected on student performance. The courses ranged from large, introductory courses to small, upper-level courses, and included a sophomore learning community. The characteristics of these courses, including the class name, size, and percentage of the class focused on bioinformatics, as listed in information gathered from each school, is shown in Table 6. The courses taught at School 5 (Georgetown) had the highest bioinformatics focus (> 80%) and smallest class sizes (15 to 20 students) compared with the other schools (Table 6). All schools showed significant increases in student performance regardless of the amount of focus on bioinformatics teaching, although these data suggest that the amount of time spent on bioinformatics content leads to increased scores on the post-course quiz, as was observed previously (12). Class size has also been shown to affect student performance (13, 14), and the Georgetown classes were among the smallest of the courses taught among the faculty who deployed the performance quiz. The high post-course scores among these students could be explained by the dual roles of the instructors as developers and teachers, although two different instructors are represented in these data. We also note the pre-course test scores at
### TABLE 3.
Reasons preventing faculty from using Genome Solver materials in their classrooms.

| Reason                                                                 | Supporting quotes                                                                                                                                 |
|----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| I’m not [currently] teaching courses where this content fits well. (9 tallies) | “I did not teach the course that would best use the bioinformatics activities this past semester. I will be teaching the applicable course in the upcoming semester.”  
“I was teaching at another institution for the last year and could not use the materials. I plan to use them in my global health and microbiology class.” |
| Although I teach a course that involves bioinformatics, this content doesn’t suit my needs. (0 tallies) | “Better explanation of what the tools are telling us. Why we use each tool. Not just how to use each tool.”  
“A more basic starting point for the workshop.”  
“I needed more hands-on time to figure out what the daily talks were about. While I understood in principle what I was supposed to do, I lacked knowledge on how I’d go about implementing something from scratch.”  
“Acquiring a solid understanding of the subject matter myself so that I am able to teach it.” |
| I am not confident teaching Genome Solver materials to a class. (5 tallies) | “GEP – their ‘ready-to-adopt’ curricular items and student projects made it possible for me to implement an introductory level bioinformatics course as a one-person discipline.”  
“GEP from Washington University St. Louis. Because we use Drosophila in Genetics, it was nice having fly projects.” |
| I prefer to develop and implement my own bioinformatics material. (0 tallies) |                                                                                                                                                  |
| I don’t have support within the department to implement bioinformatics into the curriculum. (0 tallies) |                                                                                                                                                  |

Individuals responding to question 5 in the follow-up survey (n = 20). This question presented a number of discrete responses to the participants (listed on the left) and they could opt for more than one answer. Participants could also enter information in a free-response box. This free-response box is the source of the selected quotes on the right.

### TABLE 4.
Reasons preventing faculty from using Genome Solver materials with their research students.

| Reason                                                                 | Supporting quotes                                                                                                                                 |
|----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| Bioinformatics research does not align with my research. (3 tallies) | “The projects I oversee have not made use of those informatic tools yet.”                                                                                                                                   |
| Although my research includes some bioinformatics, the Genome Solver materials don’t fit well. (12 tallies) | “We use BLAST, Genbank, and NCBI resources in undergraduate research. The workshop has helped at least two undergraduate research scholars to learn about and use bioinformatics resources. Their research is not on the microbiome, but the approach is relevant and useful to their learning.” |
| I am not confident working with Genome Solver material with my research students. (5 tallies) | “I need to spend more time on using the materials before I am confident in passing it off to a student to try.”                                                                                         |
| Write-in response 1: Plan to use materials in the future               | “Hope to use it for some phage work with an undergraduate research student.”  
“I will do it next semester. It is just timing.”  
“Really have not had the opportunity yet.” |
| Write-in response 2: Not doing research                               | “Not currently conducting any research with undergraduate students.”  
“I’m not doing research at this time.”  
“I do not do research.” |

Individuals responding to question 7 of the follow-up survey (n = 34). This question presented a number of discrete responses to the participants (listed on the left) and they could opt for more than one answer. Participants could also enter information in a free-response box. This free-response box is the source of the selected quotes on the right.
this school are not significantly different from the other schools. Nevertheless, when we examined the post-course increases after removing the Georgetown student data, we still observed increased post-course correct responses for 13 of the 20 questions (Appendix 4).

**DISCUSSION**

Bioinformatics, despite its increasing prevalence in research, is not yet a routine part of biology instruction at the undergraduate level. There are a number of issues that prevent incorporation of bioinformatics into biology education. First and foremost, many faculty lack training in this area (4). One of the recommendations of the BIO2010 report is to provide more faculty development opportunities aimed at understanding the integrative nature of biology, mathematics, and the physical sciences, a crucial step toward improving undergraduate biology education (15). Faculty training programs in bioinformatics, such as the GCAT-SEEK (16, 17), GEP (11, 12), SEA-PHAGES (18), and as described here, Genome Solver, aim to create networks of educators who are using genomics and bioinformatics in undergraduate curricula.

Genome Solver's goals for student engagement in genomics and bioinformatics meet those outlined in *Vision and Change for Undergraduate Biology Education* reports (19, 20). Research shows that active-learning techniques employed by teachers improve student performance in STEM subjects (21); classroom-based undergraduate research experiences (CUREs) such as Genome Solver can provide such active learning experiences. Various studies focusing on assessments of CUREs show evidence of increased content knowledge and performance gains for students (12, 22–26). Similar results to those shown here were documented by Shaffer et al. (11), and the fact that our work focuses on microbial genomics rather than annotation of eukaryotic genes demonstrates the broad utility of bioinformatics approaches for undergraduate life science education. Bioinformatics and genomics provide opportunities to apply quantitative approaches to tackle biological problems (27, 28). Finally, bioinformatics provides a cost-effective way for students to gain research experience, which has been suggested to improve retention in STEM, particularly for under-represented minority students (29, 30).

Genome Solver workshop participants provided very strong initial feedback immediately following the workshop about the success of the training in contributing to their
The difference in post-course versus pre-course mean scores by Carnegie classification.

| School Classification          | Pre-Course Mean | Post-Course Mean | Difference | p value |
|-------------------------------|-----------------|------------------|------------|---------|
| Doctoral Universities         | 7.12            | 10.33            | 3.21       | <0.01   |
| Master’s College and Universities | 7.25           | 9.53             | 2.28       | <0.01   |

The pre- and post-course means of the students were analyzed based on the Carnegie Classification of their schools. Means are the average number of correct answers on the 20-question quiz. The difference in learning gains between doctoral and master’s schools was not significant ($p = 0.44$).

Confidence, providing them with helpful materials, and networking with other faculty. The faculty survey data suggest that workshop training was important for three reasons. The first benefit was that faculty learned how to use bioinformatics and genome annotation software, such as BLAST and RAST. The second benefit was going through exercises that served as idea-generators for what they could do with genomics in their own curricula. Finally, workshop participants benefitted from the networking and development of a learning community. One of the faculty partners in collecting student quiz data stated, “I’m more of a traditional microbiologist [so the Genome Solver workshop] helped push me, to help my fundamental understanding so that I can implement it a little bit more into my class.” Qualitative evidence suggests that using CUREs to teach leads to benefits such as connecting teaching and research goals, providing enjoyment and satisfaction in teaching, and providing opportunities to publish research with students (31).

However, as found in the follow-up survey, not all workshop participants gained the necessary confidence to incorporate bioinformatics into their teaching. A recent paper suggests that boot camps and short workshops might not be the optimal way to learn this information (32), that a longer-term approach would be valuable for sustaining gains. Faculty mentoring networks, as developed and deployed by QUBES (Quantitative Undergraduate Biology Education and Synthesis), in which faculty meet with mentors online on a weekly or biweekly basis, with discrete tasks and time for questions, could be a way to address this point in the future (33).

Another issue is the lack of formal, agreed-upon core competencies for bioinformatics, especially with respect to life science students. However, efforts are underway by the Network for Integrating Bioinformatics into Life Sciences Education (NIBLSE). Recently, this group conducted a national survey to develop and refine a list of competencies that could form the basis of national norms (34). Similarly, the International Society for Computational Biology (ISCB) Education Committee has worked to identify the bioinformatics skill sets necessary for three different “personas,” people with different needs, including a bioinformatics user (one who uses existing tools), a bioinformatics scientist (one who employs computational methods to advance understanding of biology), and a bioinformatics engineer (one who creates the computer algorithms used by the other users) (35). The ISCB Committee has recently refined these personas to include more personas and types of jobs that might need to use bioinformatics skill sets (36).

Course-based undergraduate research experiences provide students with opportunities to engage in problem-solving in which the outcome is unknown and which has import for others in the field (37). Bioinformatics CUREs in some ways are easy to adopt—the cost of research is relatively low since all that is needed is a computer and access to the Internet. On the other hand, bioinformatics CUREs require a cadre of faculty with the confidence and ability to guide students in this type of research, which are in short supply. The Genome Solver Project over the last five years has attempted to address this deficit by creating a network of opportunities for faculty and development of a learning community. Although not all faculty who attended a Genome Solver workshop were able to incorporate bioinformatics instruction into their current teaching, for a variety of reasons, the student quiz data suggest that students who are exposed to bioinformatics instruction make significant gains. We continue to promote the integration of bioinformatics into life science curricula as we believe that training undergraduate students to analyze “big data” such as sequence data will better prepare them for graduate study and the workforce.
TABLE 6.
Characteristics of the courses at five different schools.

| School | Course                                      | Class Size | Bioinformatics Focus | Class Makeup           | # of Instructors |
|--------|---------------------------------------------|------------|----------------------|------------------------|-----------------|
| 1 (D)  | The Microbial Universe                     | 50–75      | 15–20%               | Majors                 | 1               |
| 2 (D)  | Foundations of Biology                      | 1,000      | <10%                 | Majors & nonmajors     | 2               |
|        | Honors Foundations of Biology              | 35         | >50%                 | Majors                 | 1               |
| 3 (M)  | Advanced Topics in Biochemistry             | 10–15      | 25–30%               | Majors                 | 2               |
|        | The Human Microbiome and Disease           | 10–15      | 60%                  | Majors & nonmajors     | 1               |
| 4 (M)  | Microbial Physiology                        | 18         | 15%                  | Majors                 | 1               |
| 5 (D)  | Genomics and Bioinformatics                 | 15–20      | 100%                 | Majors                 | 2               |
|        | Microbial Genomics                          | 5–10       | 80%                  | Majors                 | 2               |

This information was obtained as part of the 30-minute phone interviews with faculty based on the questions shown in Appendix 3. M = master’s granting; D = doctoral.

SUPPLEMENTAL MATERIAL

Appendix 1: Follow-up participant survey
Appendix 2: Pre-and post-course quiz questions for students
Appendix 3: Faculty interview protocol
Appendix 4: Supplemental Figure 1. Comparison of pre- and post-course quiz responses after removing school 5 (Georgetown University)

ACKNOWLEDGMENTS

We thank all the members of the Genome Solver community, especially those faculty who implemented the performance quiz in their classrooms and agreed to the faculty interviews. The Genome Solver Project is supported by the National Science Foundation (DUE 1123016 and 1505102). The authors declare that there are no conflicts of interest.

REFERENCES

1. Metzker ML. 2010. Sequencing technologies—the next generation. Nat Rev Genet 11:31–46.
2. Levine AG. 2014. An explosion of bioinformatics careers. Science 344:1303–1304.
3. Rosenwald AG, Russell JS, Arora G. 2012. The genome solver website: a virtual space fostering high impact practices for undergraduate biology. J Microbiol Biol Educ 13:188–190.
4. Williams J, Drew J, Galindo-Gonzalez S, Robic S, Dinsdale E, Morgan W, Triplett E, Burnette J, Donovan S, Elgin S, Fowlks E, Goodman A, Grandgenett N, Goller C, Hauser C, Jungck JR, Newman J, Pearson W, Ryder E, Sayres MW, Sierk M, Smith T, Tosado-Acevedo R, Tapprich W, Tobin T, Toromartinez A, Welch L, Wright R, Ebenbach D, McWilliams M, Rosenwald A, Pauley M. 2017. Barriers to integration of bioinformatics into undergraduate life sciences education. bioRxiv 204420.
5. Altschul SF, Gish W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.
6. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.
7. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797.
8. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874.
9. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206–D214.
10. Thomas DR. 2006. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 27(2):237–246.
11. Shaffer CD, Alvarez C, Bailey C, Barnard D, Bhalla S, Chandrasekaran C, Chandrasekaran V, Chung H-m, Dorer DR, Du C, Eckdahl TT, Poet JL, Frohlich D, Goodman AL, Gossler Y, Hauser C, Hoopes LLM, Johnson D, Jones CJ, Kaehler M, Kokan N, Kopp OR, Kulega GA, Mcneil G, Moss R, Myka JL, Nagengast A, Morris R, Overvoorde Pj, Shoop E, Parrish S, Reed K, Regisford EG, Revie D, Rosenwald AG, Saville K, Schroeder S, Shaw M, Skuse G, Smith C, Smith M, Spana EP, Spratt M, Stamm J, Thompson JS, Wawersik M, Wilson BA, Youngblom J, Leung W, Buhler J, Mardis ER, Lopatto D, Elgin SC. 2010. The genomics education partnership: successful integration of research into laboratory classes at a diverse group of undergraduate institutions. CBE Life Sci Educ 9:55–69.
12. Shaffer CD, Alvarez CJ, Bednarzke AE, Dunbar D, Goodman AL, Reinke C, Rosenwald AG, Wolyniak MJ, Bailey C, Barnard D, Bazinet C, Beach DL, Bedard JE, Bhalla S, Braverman J, Burg M, Chandrasekaran V, Chung HM, Clase K, Dejong RJ, Diangelo JR, Du C, Eckdahl TT, Eisler H, Emerson JA, Frary A, Frohlich D, Gossler Y, Govind S, Haberman A, Hark AT, Hauser C, Hoogewerf A, Hoopes LL, Howell CE, Johnson D, Jones CJ, Kadlec L, Kaehler M, Silver Key SC, Kleinschmit A, Kokan NP, Kopp O, Kuleck G, Leatherman J,
Lopilato J, Mackinnon C, Martinez-Cruzado JC, McNeil G, Mel S, Mistry H, Nagengast A, Overvoorde P, Paetkau DW, Parrish S, Peterson CN, Preuss M, Reed LK, Revie D, Robic S, Roecklein-Canfield J, Rubin MR, Saville K, Schroeder S, Sharif K, Shaw M, Skuse G, Smith CD, Smith MA, Smith ST, Spana E, Spratt M, Sreenivasan A, Stamm J, Szauter P, Thompson JS, Wawersik M, Youngblom J, Zhou L, Mardis ER, Buhler J, Leung W, Lopatto D, Elgin SC. 2014. A course-based research experience: how benefits change with increased investment in instructional time. CBE Life Sci Educ 13:111–130.

13. Bandiera O, Larcinese V, Rasul I. 2010. Heterogeneous class size effects: new evidence from a panel of university students. Econ J 120:1365–1398.

14. Johnson IY. 2010. Class size and student performance at a public research university: a cross-classified model. Res High Educ 51:701–723.

15. National Research Council. 2003. Bio 2010: Transforming undergraduate education for future research biologists. The National Academies Press, Washington, DC.

16. Buonaccorsi VP, Boyle MD, Grove D, Praul C, Sakk E, Stuart A, Tobin T, Hosler J, Carney SL, Engle MJ, Overton BE, Newman JD, Pizzorno M, Powell JR, Trun N. 2011. GCAT-SEEKsequence: genome consortium for active teaching of undergraduates through increased faculty access to next-generation sequencing data. CBE Life Sci Educ 10:342–345.

17. Buonaccorsi V, Peterson M, Lamendella G, Newman J, Trun N, Tobin T, Aguilar A, Hunt A, Praul C, Grove D, Roney J, Roberts W. 2014. Vision and change through the genome consortium for active teaching using next-generation sequencing (GCAT-SEEK). CBE Life Sci Educ 13:1–2.

18. Hatfull GF. 2012. The secret lives of mycobacteriophages. Adv Virus Res 82:179–288.

19. American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education: a call to action: a summary of recommendations made at a national conference organized by the American Association for the Advancement of Science, July 15–17, 2009. Washington, DC.

20. American Association for the Advancement of Science. 2015. Vision and change in undergraduate biology education: chronicling change, inspiring the future. Washington, DC.

21. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderworth MP. 2014. Active learning increases student performance in science, engineering, and mathematics. Proc Nat Acad Sci 111:8410–8415.

22. Brownell SE, Hekmat-Scafe DS, Singla V, Seawell PC, Imam JFC, Eddy SL, Stearns T, Cyert MS. 2015. A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. CBE Life Sci Educ 14(2):ar21.

23. Siritunga D, Montero-Rojas M, Carrero K, Toro G, Velez A, Carrero-Martinez FA. 2011. Culturally relevant inquiry-based laboratory module implementations in upper-division genetics and cell biology teaching laboratories. CBE Life Sci Educ 10:287–297.

24. Rowland SL, Lawrie GA, Behrendorf JBYH, Gillam EMJ. 2012. Is the undergraduate research experience (URE) always best? The power of choice in a bifurcated practical stream for a large introductory biochemistry class. Biochem Mol Biol Educ 40:46–62.

25. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, Dejong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SC, Findley AM, Gissendanner CR, Golebiowska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cressawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. MBio 5:e01051–13.

26. Ward JR, Clarke HD, Horton JL. 2014. Effects of a research-infused botanical curriculum on undergraduates’ content knowledge, STEM competencies, and attitudes toward plant sciences. CBE Life Sci Educ 13:387–396.

27. Altman RB. 1998. A curriculum for bioinformatics: the time is ripe. Bioinformatics 14:549–550.

28. Campbell AM, Ledbetter MLS, Hoopes LLM, Eckdahl TT, Heyer LJ, Rosenwald A, Fowlks E, Tonidandel S, Bucholtz B, Gottfried G. 2007. Genome consortium for active teaching: meeting the goals of BIO2010. CBE Life Sci Educ 6:109–118.

29. Pender M, Marcotte DE, Domingo MRS, Maton KI. 2010. The STEM pipeline: the role of summer research experience in minority students’ Ph.D. aspirations. Educ Policy Ann Arch 18(30):1–36.

30. Eagan MK, Jr., Hurtado S, Chang MJ, Garcia GA, Herrera FA, Garibay JC. 2013. Making a difference in science education: the impact of undergraduate research programs. Am Educ Res J 50:683–713.

31. Shortlidge EE, Bangera G, Brownell SE. 2016. Faculty perspectives on developing and teaching course-based undergraduate research experiences. BioScience 66:54–62.

32. Feldon DF, Jeong S, Peugh J, Roksa J, Maahs-Fladung C, Shenyo A, Oliva M. 2017. Null effects of boot camps and short-format training for PhD students in life sciences. PNAS 114(37):9854–9858.

33. Donovan S, Eaton CD, Gower ST, Jenkins KP, LaMar MD, Poli DB, Sheehy R, Wojdak JM. 2015. QUBES: a community focused on supporting teaching and learning in quantitative biology. Lett Biomath 2(1):46–55.

34. Sayres MAW, Hauser C, Sierk M, Robic S, Rosenwald AG, Smith T, Triplet EW, Williams JJ, Dinsdale E, Morgan W, Burnetze J, Donovan S, Drew JC, Elgin S, Fowlks E, Galindo-Gonzalez S, Grandgenett N, Goller CC, Jungck J, Newman J, Pearson W, Ryderv E, Tosado-Acevedo R, Tapprich W, Tobin T, Toro A, Welch L, Wright R, Ebenbach D, Olney KC, McWilliams M, Pauley MA. Bioinformatics core competencies for undergraduate life science educators. PLOS One 13(6):e0196878.

35. Welch L, Lewitter F, Schwartz R, Brooksbank C, Radiovoc P, Gaeta B, Scheider MV. 2014. Bioinformatics curriculum guidelines: toward a definition of core competencies. PLOS Comput Biol 10(3):e1003496.
36. Welch L, Brooksbank C, Schwartz R, Morgan SL, Gaeta B, Kilpatrick AM, Mietchen D, Moore BL, Mulder N, Pauley M, Pearson W, Radivojac P, Rosenberg N, Rosenwald A, Rustici G, Warnow T. 2016. Applying, evaluating and refining bioinformatics core competencies (an update from the curriculum task force of ISCB’s Education Committee). PLOS Comput Biol 12:e1004943.

37. Dolan EL. 2016. Course-based undergraduate research experiences: current knowledge and future directions. Paper commissioned for the Committee on Strengthening Research Experiences for Undergraduate STEM Students Board on Science Education, Division of Behavioral and Social Sciences and Education Board on Life Sciences, Division of Earth and Life Studies.