Winter 2014

A Central Support System Can Facilitate Implementation and Sustainability of a Classroom-Based Undergraduate Research Experience (CURE) in Genomics

Adam Haberman
University of San Diego, ahaberman@sandiego.edu

A Central Support System Can Facilitate Implementation and Sustainability of a Classroom-Based Undergraduate Research Experience (CURE) in Genomics

Abstract

In their 2012 report, the President's Council of Advisors on Science and Technology advocated “replacing standard science laboratory courses with discovery-based research courses”—a challenging proposition that presents practical and pedagogical difficulties. In this paper, we describe our collective experiences working with the Genomics Education Partnership, a nationwide faculty consortium that aims to provide undergraduates with a research experience in genomics through a scheduled course (a classroom-based undergraduate research experience, or CURE). We examine the common barriers encountered in implementing a CURE, program elements of most value to faculty, ways in which a shared core support system can help, and the incentives for and rewards of establishing a CURE on our diverse campuses. While some of the barriers and rewards are specific to a research project utilizing a genomics approach, other lessons learned should be broadly applicable. We find that a central system that supports a shared investigation can mitigate some shortfalls in campus infrastructure (such as time for new curriculum development, availability of IT services) and provides collegial support for change. Our findings should be useful for designing similar supportive programs to facilitate change in the way we teach science for undergraduates.

CBE—Life Sciences Education
Vol. 13, 711–723, Winter 2014

Article

A Central Support System Can Facilitate Implementation and Sustainability of a Classroom-Based Undergraduate Research Experience (CURE) in Genomics

David Lopatto,[1] Charles Hauser,[2] Christopher J. Jones,[3] Don Paetkau,[4] Vidya Chandrasekaran,[5] David Dunbar,[6] Christy MacKinnon,[7] Joyce Stamm,[8] Consuelo Alvarez,[9] Daron Barnard,[10] James E. J. Bedard,[11]* April E. Bednarski,[12†] Satish Bhalla,[13] John M. Braverman,[14] Martin Burg,[15] Hui-Min Chung,[16] Randall J. DeJong,[17] Justin R. DiAngelo,[18] Chunguang Du,[19] Todd T. Eckdahl,[20] Julia Emerson,[21] Amy Frary,[22] Donald Frohlich,[23] Anya L. Goodman,[24] Yuying Gosser,[25] Shubha Govind,[26] Adam Haberman,[27‡] Amy T. Hark,[28] Arlene Hoogewerf,[17] Diana Johnson,[29] Lisa Kadlec,[30] Marian Kaehler,[31] S. Catherine Silver Key,[32] Nighat P. Kokan,[33] Olga R. Kopp,[34] Gary A. Kuleck,[35] Jane Lopilato,[36] Juan C. Martinez-Cruzado,[37] Gerard McNeil,[38] Stephanie Mel,[39] Alexis Nagengast,[40] Paul J. Overvoorde,[41] Susan Parrish,[42] Mary L. Preuss,[43] Laura D. Reed,[44] E. Gloria Regisford,[45] Dennis Revie,[46] Srebrenka Robic,[47] Jennifer A. Roecklien-Canfield,[48] Anne G. Rosenwald,[49,] Michael R. Rubin,[50] Kenneth Saville,[51] Stephanie Schroeder,[43] Karim A. Sharif,[52§] Mary Shaw,[53] Gary Skuse,[54] Christopher D. Smith,[55] Mary Smith,[56] Sheryl T. Smith,[57] Eric P. Spana,[58] Mary Spratt,[59] Aparna Sreenivasan,[60] Jeffrey S. Thompson,[61] Matthew Wawersik,[62] Michael J. Wolyniak,[63] James Youngblom,[64] Leming Zhou,[65] Jeremy Buhler,[66] Elaine Mardis,[67] Wilson Leung,[68] Christopher D. Shaffer,[68] Jennifer Threlfall,[69] and Sarah C. R. Elgin[68]

[1]Department of Psychology, Grinnell College, Grinnell, IA 50112; [2]Bioinformatics, St. Edward's University, Austin, TX 78704; [3]Department of Biological Sciences, Moravian College, Bethlehem, PA 18018; [4]Department of Biology, Saint Mary's College, Notre Dame, IN 46556; [5]Department of Biology, Saint Mary's College of California, Moraga, CA 94556; [6]Science Department, Cabrini College, Radnor, PA 19087; [7]Biology Department, University of Incarnate Word, San Antonio, TX 78209; [8]Department of Biology, University of Evansville, Evansville, IN 47722; [9]Biological & Environmental Sciences, Longwood University, Farmville, VA 23909; [10]Biology Department, Worcester State University, Worcester, MA 01602; [11]Department of Biology, Adams State University, Alamosa, CO 81101; [12]Chemistry Department, Lindenwood University, St. Charles, MO 63301; [13]Department of Computer Science & Engineering, Johnson C. Smith University, Charlotte, NC 28216; [14]Department of Biology, Saint Joseph's University, Philadelphia, PA 19131; [15]Departments of Biomedical Sciences & Cell and Molecular

DOI: 10.1187/cbe.13-10-0200
Financial disclosure: This work was supported by grant 52007051 from the Howard Hughes Medical Institute to S.C.R.E. under the Professors Program, grant 2U54 HG00307910 from National Human Genome Research Institute (Richard K. Wilson, principal investigator), and by Washington University in St. Louis. None of the above funders had any role in the design or conduct of the study, nor in the collection, analysis, or interpretation of the data, nor in the preparation, review, or approval of the manuscript.

Present addresses: *Department of Biology, University of the Fraser Valley, Abbotsford, BC V2S 7M8, Canada; †Department of Biology, Washington University in St. Louis, St. Louis, MO 63130; ‡Department of Biology, University of San Diego, San Diego, CA 92110; §Biology Department, Massasoit Community College, Brockton, MA 02302.

Address correspondence to: Christopher D. Shaffer (shaffer@biology.wustl.edu).

Biology, Grand Valley State University, Allendale, MI 49401; [16]Department of Biology, University of West Florida, Pensacola, FL 32514; [17]Department of Biology, Calvin College, Grand Rapids, MI 49546; [18]Department of Biology, Hofstra University, Hempstead, NY 11549; [19]Department of Biology & Molecular Biology, Montclair State University, Montclair, NJ 07043; [20]Department of Biology, Missouri Western State University, St. Joseph, MO 64507; [21]Department of Biology, Amherst College, Amherst, MA 01002; [22]Department of Biological Sciences, Mount Holyoke, South Hadley, MA 01075; [23]Biology Department, University of St. Thomas, Houston, TX 77006; [24]Department of Chemistry and Biochemistry, California Polytechnic State University, San Luis Obispo, CA 93405; [25]Grove School of Engineering, City College of New York, New York, NY 10031; [26]Biology Department, City College of New York, New York, NY 10031; [27]Biology Department, Oberlin College, Oberlin, OH 44074; [28]Biology Department, Muhlenberg College, Allentown, PA 18104; [29]Department of Biological Sciences, George Washington University, Washington, DC 20052; [30]Department of Biology, Wilkes University, Wilkes-Barre, PA 18766; [31]Biology Department, Luther College, Decorah, IA 52101; [32]Department of Biology, North Carolina Central University, Durham, NC 27707; [33]Department of Biology, Cardinal Stritch University, Milwaukee, WI 53217; [34]Utah Valley University, Orem, UT 84058; [35]College of Engineering and Science, University of Detroit Mercy, Detroit, MI 48221; [36]Department of Biology, Simmons College, Boston, MA 02115; [37]Department of Biology, University of Puerto Rico at Mayaguez, Mayaguez, PR 00680; [38]Department of Biology, York College, City University of New York, Jamaica, NY 11451; [39]Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093; [40]Departments of Chemistry and Biochemistry, Widener University, Chester, PA 19013; [41]Biology Department, Macalester College, St. Paul, MN 55105; [42]Biology Department, McDaniel College, Westminster, MD 21157; [43]Department of Biological Sciences, Webster University, Webster Groves, MO 63119; [44]Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35401; [45]Department of Biology, Prairie View A&M University, Prairie View, TX 77446; [46]Department of Biology, California Lutheran University, Thousand Oaks, CA 91360; [47]Department of Biology, Agnes Scott College, Decatur, GA 30030; [48]Department of Chemistry, Simmons College, Boston, MA 02115; [49]Department of Biology, Georgetown University, Washington, DC 20057; [50]Department of Biology, University of Puerto Rico at Cayey, Cayey, PR 00736; [51]Biology Department, Albion College, Albion, MI 49224; [52]Department of Natural Sciences, LaGuardia Community College, Long Island City, NY 11101; [53]Department of Biology and Chemistry, New Mexico Highlands University, Las Vegas, NM 87701; [54]Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, NY 14623; [55]Department of Biology, San Francisco State University, San Francisco, CA 94132; [56]Department of Biology, North Carolina A&T State University, Greensboro, NC 27411; [57]Biology Department, Arcadia University, Glenside, PA 19038; [58]Department of Biology, Duke University, Durham, NC 27708; [59]Biology Department, William Woods University, Fulton, MO 65251; [60]Science and Environmental Policy, California State University–Monterey Bay, Seaside, CA 93955; [61]Department of Biology, Denison University, Granville, OH 43023; [62]Biology Department, College of William and Mary, Williamsburg, VA 23185; [63]Biology Department, Hampden–Sydney College, Hampden–Sydney, VA 23943; [64]Department of Biology, California State University–Stanislaus, Turlock, CA 95382; [65]Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA 15213; [66]Department of Computer Science and Engineering and Department of Genetics, [67]Genome Institute, [68]Biology Department, and [69]George Warren Brown School of Social Work, Washington University in St. Louis, St. Louis, MO 63130

Submitted October 30, 2013; Revised September 20, 2014; Accepted September 21, 2014
Monitoring Editor: Erin Dolan

In their 2012 report, the President's Council of Advisors on Science and Technology advocated "replacing standard science laboratory courses with discovery-based research courses"—a challenging proposition that presents practical and pedagogical difficulties. In this paper, we describe our collective experiences working with the Genomics Education Partnership, a nationwide faculty consortium that aims to provide undergraduates with a research experience in genomics through a scheduled course (a classroom-based undergraduate research experience, or CURE). We examine the common barriers encountered in implementing a CURE, program elements of most value to faculty, ways in which a shared core support system can help, and the incentives for and rewards of establishing a CURE on our diverse campuses. While some of the barriers and rewards are specific to a research project utilizing a genomics approach, other lessons learned should be broadly applicable. We find that a central system that supports a shared investigation can mitigate some shortfalls in campus infrastructure (such as time for new curriculum development, availability of IT services) and provides collegial support for change. Our findings should be useful for designing similar supportive programs to facilitate change in the way we teach science for undergraduates.

INTRODUCTION

The *Vision and Change* report from the American Association for the Advancement of Science (AAAS) calls for all undergraduate students to have experience with research to understand the process of science (AAAS, 2011). A recent report to President Obama recommended replacing *all* standard science laboratory courses with discovery-based research courses (President's Council of Advisors on Science and Technology [PCAST], 2012). While there is ample evidence of the importance of research experiences in undergraduate science, technology, engineering, and mathematics (STEM) education (e.g., Hunter *et al*., 2007; Lopatto, 2009; Laursen *et al*., 2010), these recommendations raise questions of feasibility in the minds of faculty and administrators alike (Healey and Jenkins, 2009). Most institutions cannot provide individual (or even small-group) mentored research experiences for all of their STEM students, given limitations in the number of available research mentors (faculty and others), supply budgets, physical facilities, and infrastructure support (Wood, 2003; Desai *et al*., 2008). Implementation of research-centered laboratory courses (classroom-based undergraduate research experiences, or CUREs), an attractive alternative, requires overcoming similar (although less severe) barriers, as well as overcoming entrenched academic practices (Rowlett *et al*., 2012). Deciding to change pedagogical traditions (e.g., scheduling patterns or allocations of class time) is not easy (Henderson and Dancy, 2007; Dancy and Henderson, 2008; Winningham *et al*., 2009). Science by its nature changes constantly, so maintaining up-to-date access to research tools, as well as finding time to develop suitable research curricula, is an intellectual and financial challenge (Spell *et al*., 2014). Our hypothesis is that collaborative, nationwide research projects can help to overcome some of these barriers and support "pioneer" faculty advocating for local change. To examine this hypothesis, we studied faculty reports of the incentives and barriers to successfully implementing a research-based laboratory in genomics. While the data examined here are confined to the Genomics Education Partnership (GEP), we note that national collaborative research projects have a long history of successfully engaging students (and other citizens) in science (e.g., Cornell Lab of Ornithology, www.birds.cornell.edu/page.aspx?pid=1664; NASA Citizen Scientists, http://science.nasa.gov/citizen-scientists). The availability of the Internet now makes it relatively easy to connect members of a nationwide partnership, making this a practical strategy.

The availability of new tools and large data sets in genomics has fostered a number of large-scale collaborative undergraduate research projects in genetics, evolution, and cell/molecular biology. These projects have facilitated access to technical resources, faculty training, scientific and technical expertise, and shared curriculum, as demonstrated by programs such as the Genome Consortium for Active Teaching (GCAT; Campbell *et al*., 2007), the GCAT NextGen Sequencing Group (GCAT-SEEK; Buonaccorsi *et al*., 2014), Phage Hunters Integrating Research and Education (PHIRE; Hatfull *et al*., 2006), and the Howard Hughes Medical Institute (HHMI) Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) course (Jordan *et al*., 2014). Within this arena, using bioinformatics as the major platform for undergraduate research has several obvious advantages: the need for laboratory infrastructure is minimal (only computers and access to the Internet are required); there are no lab safety issues, allowing open access 24/7; there is a large pool of publicly available raw data for students to work with; and the approach lends itself to peer instruction, as many undergraduates quickly acquire the needed technical expertise. There are also some disadvantages: many faculty members lack training in bioinformatics approaches and tools; and some students, anticipating wet-bench labs or fieldwork, fail to see a computer-based project as research (Shaffer *et al*., 2014). However, the growing utilization of large data sets in all areas of biology, as well as growing awareness of the relevance of genomic information in health sciences, is increasing recognition by students and faculty alike of the need to bring bioinformatics into the biology curriculum. Several excellent genomics research projects for undergraduates with an emphasis on bioinformatics have recently been described (Banta *et al*., 2012; Ditty *et al*., 2010; Burnette and Wessler, 2013; Harris and Bellino, 2014).

The GEP is a consortium in which more than 100 colleges and universities (mostly primarily undergraduate institutions, or PUIs) have joined with Washington University in St. Louis (WUSTL) with the goal of providing undergraduates with a research experience in genomics (see http://gep.wustl.edu). The GEP is investigating the evolution of the Muller F element, a region of the *Drosophila* genome that exhibits both heterochromatic and euchromatic properties, and the evolution of the F element genes. Undergraduates are involved in both finishing (improving the quality of draft sequence) and annotating (creating hand-curated gene models based on all available evidence, mapping repeats, and identifying other features) designated regions of the *Drosophila* genome. They work on 40-kb "projects," which, after quality control checks, are reassembled to generate large domains for analysis. GEP materials have been adapted to many different settings, from a short module in a first genetics course to the core of a semester-long laboratory course to an "independent study" research course. A common student assessment is carried out using the central website. Pre/postcourse quizzes demonstrate that GEP students do indeed improve their knowledge of genes and genomes through their research (Shaffer *et al*., 2010, 2014). Postcourse survey results from 2008 and 2010–2012 on science attitudes are consistent and show an overall pattern and numerical scores very similar to those of students in a dedicated summer research program (Lopatto, 2007; Lopatto *et al*., 2008; see especially Shaffer *et al*., 2014). All student projects are completed at least twice independently, and a reconciliation process is carried out by experienced students working at WUSTL during the summer. Student annotations are deposited in GenBank and form the core of our scientific publications, which analyze the reassembled regions as a whole (e.g., Leung *et al*., 2010). A paper based on comparative analysis of the F element of four *Drosophila* species, now in preparation, will have more than 1000 student and faculty coauthors. Thus, by both pedagogical and scientific measures, the GEP appears to have assembled a group of faculty who each have successfully developed a CURE on their campus.

While there have been sweeping calls for the development and use of CUREs (Karukstis, 2008; AAAS, 2011; Kloser, 2011; PCAST, 2012), there has been relatively little study of the practical and pedagogical issues faculty face in developing research-based courses or of the kinds of support

systems/best practices that could help facilitate the change to a CURE-based STEM curriculum. These issues need to be explored if widespread adoption of this strategy is to be successful. Here, we utilize our collective experiences with the GEP to explore these issues, focusing on three questions:

1. What are the barriers for implementing a CURE *in genomics*? An examination of the types of support and resources needed, contrasted with their availability on campus, can potentially identify common barriers to CURE implementation.
2. How does the central core system help faculty and foster *overall success of the GEP*? We examine those features of the central GEP support system that help us overcome local barriers, looking in particular for correlations between faculty needs and the reported value of core resources and determining what program elements faculty members value.
3. What are the incentives for faculty members *to create and sustain a research-based experience in genomics*? We examine the reported incentives and rewards that drive faculty to take on the challenge to create and sustain a research-based laboratory course.

Our collective experience suggests that a collaborative network with core resources can effectively support faculty who want to provide a research experience for their students through their classroom teaching in genomics.

MATERIALS AND METHODS

GEP Members

Faculty members of the GEP were initially recruited through an email invitation using a list of PUI faculty maintained by the Washington University Division of Biology and Biomedical Sciences and the list of GCAT members. Subsequent recruits were attracted by posters that the GEP presented at meetings of the American Society for Cell Biology, the Council on Undergraduate Research, the Drosophila Research Conference, the Association of Biology Laboratory Educators, and other venues; through the GEP website; or through conversation with a colleague. Interested faculty members join the GEP by attending a 3- to 5-d workshop at WUSTL to gain familiarity with the bioinformatics tools and curriculum used by the partnership. They can then claim 40-kb projects posted on the GEP website and submit the results of the finishing and/or annotation completed by their students. Students who complete projects are eligible to be coauthors on the resulting scientific paper that makes use of their analysis and may present their own work at on-campus, local, or national meetings. Faculty members are eligible to be coauthors on the scientific papers, on research papers analyzing educational issues, and on meeting presentations.

Resources Provided by the Core System

The GEP central core support team at WUSTL (staffed by W.L., full-time; C.D.S., 35%; S.C.R.E., 10%; J.B., as needed; E.M. and professional staff of the Genome Institute, as needed) provides a range of IT, bioinformatics, and genomics expertise. The core team organizes and runs the introductory workshop to help new members learn the recommended workflow, gain familiarity with bioinformatics tools, and see firsthand the resources being used in the project. The workshop also provides opportunities for faculty to discuss ways to adapt this research project for their students at their institution with colleagues pursuing similar educational goals. Available training includes both genome annotation (done with J.B.) and sequence improvement (done with E.M. and staff of the Genome Institute). Faculty members may also have a student or colleague come to the same (or a subsequent) workshop, selecting a person who will assist in implementing their GEP course. (This addresses the initial need for a teaching assistant [TA] or peer instructor; it is anticipated that successful and enthusiastic students who took the course in a prior year will serve as TAs/peer instructors in following years.) Group concerns—discussion of ongoing implementation, design of assessment tools, improvements to the curriculum, preparation of joint publications—are addressed at annual alumni workshops, 2 - to 3-d meetings that all GEP faculty are eligible to attend (and approximately 50% do so in a given year). Alumni workshops also provide professional development opportunities in bioinformatics through lectures/labs with WUSTL faculty or guest lecturers. The progress and products of alumni discussions are posted on a private GEP wiki for initial dissemination to the group, comment, and revision. Ultimately, new curriculum is posted on the public website (http://gep.wustl.edu).

Beyond the workshops, the GEP support system is organized through the central website (as described in Shaffer *et al.*, 2010), which provides members with access to the sequence improvement and annotation research projects for students and acts as a pipeline for submitting final project reports to WUSTL for quality control (carried out by experienced undergraduates) and final assembly into the complete domain of interest. Curricular materials, including background lectures on the research question, introductory walkthroughs of basic bioinformatics tools and database procedures, and practice problems, are continually updated by the core staff members and are augmented by new contributions from GEP faculty as well as staff. Video tours of the Genome Institute illustrating various sequencing technologies are also posted. The GEP core staff make all necessary changes in curriculum when the National Center for Biotechnology Information Basic Local Alignment Search Tool interface or University of California–Santa Cruz Genome Browser changes, databases are updated, or new operating systems appear, significantly reducing the load on faculty partners. GEP materials are freely available on the GEP website under a Creative Commons license. The website also contains various Help functions, including the custom software and connections to Web resources required for an efficient workflow for this annotation challenge. The core staff is available to assist individual faculty with issues ranging from software installation to tricky annotations via a website bulletin board, email, Skype, and telephone.

Anonymous Survey

The Faculty Survey (Supplemental Material S1 and S2) was written by subgroups of GEP faculty in attendance at the alumni workshops during the Summer of 2010 and was posted on the GEP private wiki for all GEP members to

review; a revised final version was generated by GEP faculty during the summer of 2011 with input from D.L. and S.C.R.E. GEP faculty coauthorship ensures that the survey design covers all of the major barriers and incentives that the GEP faculty members themselves were able to identify, and the group discussion promoted a common understanding of the questions asked, improving the credibility of the instrument. Administration of the survey was approved by the Human Research Protection Office of Washington University in St. Louis (IRB ID 201104105; approval date 12 April 2012). The purpose of the survey was described in a preamble (Supplemental Material S1), and informed consent was given by moving from the preamble to the survey itself. Anonymity was maintained by stripping out identifiers before transmitting responses to D.L. and J.T. for analysis. Faculty members who had attended GEP workshops between June 2006 and January 2012 were invited by email to participate in the survey. (Six individuals who had previously left the project [deceased, left academia, or reassigned by their chair] were not contacted.) Telephone calls were made to a random subset of faculty to encourage participation in the survey. Items either asked participants to provide free-response comments in a text box or to indicate their feelings/responses on a numerical scale of 1–5.

Common barriers for implementing a CURE were first identified by using the responses to question 21 of the anonymous survey, which asked the faculty members to rank a list of factors that could be incentives or barriers for sustaining GEP curricula (Supplemental Material S2). For assessing the extent to which an item was considered to be a barrier, difference in the means between the participants' rating of the item's "importance" and its "presence on their campus" were calculated. The structure of this new barriers' scale was investigated using an exploratory factor analysis with varimax rotation. SAS 9.2 statistical software was used for this process and subsequent analyses (except where noted below). In this and the following analyses, four respondents who did not complete any part of question 21 were excluded from the analysis.

Other numerical responses used here are presented either as a table displaying the distribution of ratings (generally on a scale of 1–5) or by providing the means; errors are reported as SD. Free responses in the anonymous survey were generally very brief and were not evaluated in detail.

Open Survey

In addition to the anonymous survey, GEP faculty coauthors responded to three questions in an open format; these responses are given verbatim in Supplemental Material S3–S5. These responses were initially analyzed (by D.L.) by simply determining the frequency of keywords and phrases. We built groupings based on the various frequent keywords and phrases indicated. We included all similar phrases, variants, and synonyms thereof, for example, "independence," "independent," and "independently" were grouped together along with synonyms such as "freedom" or "on their own." For a more in-depth analysis, the answers to each question were also analyzed separately by a person with no prior contact with the project (J.T.) using inductive content analysis (Elo and Kyngäs, 2008; Krippendorff, 2013). NVivo 10 software was used for the inductive content analysis. The documents were first open coded to identify themes that emerged from the text (Miles and Huberman, 1994). These initial codes were clustered into categories by constantly comparing data within and between codes (Glaser and Strauss, 1967; Hsieh and Shannon, 2005). Labels and definitions for each category were then developed to produce a codebook; this codebook was then used to code the entire text for each question. Comments contributing to a particular theme are tagged with a common color in Supplemental Material S3–S5.

RESULTS

For evaluating the barriers, impact of core support, and incentives for the implementation of GEP-based student research, an anonymous survey was made available in late Spring 2012 to the 100 GEP faculty partners in the program who had joined before or during academic year 2011–2012. This survey asked respondents to evaluate on-campus conditions affecting the success of their efforts to teach genomics through a research project, the importance of various components of the GEP support system, and their reasons for remaining active in the GEP. GEP faculty members are from a very diverse group of schools from across the country (Supplemental Material S6), and this diversity is represented in the set of faculty completing the survey. Of the 92 faculty members who completed this survey, 64 (70%) remain active in the program, 25 (27%) described themselves as active in genomics education but not using GEP materials at present, and three (3%) described themselves as having left the program. Of the 25 who reported that they no longer use GEP materials, 14 provided reasons: one is retired, one left academia for a different career, nine had their teaching efforts redirected by their institution, and three found alternatives that were a better fit for their curricular needs. All faculty responses given were used in the following analysis; however, not all faculty responded to all items in the survey.

What Are Common Barriers to Implementing a CURE? The Discrepancies between Faculty Needs and Campus Resources

Regarding implementing and sustaining an undergraduate research course (CURE) in genomics, respondents were asked to evaluate the *importance* of 25 items and the *presence* of each item on their campus (Figure 1 and Supplemental Material S7). The largest gaps between the mean importance score and mean presence score occurred for "acceptance of genomics in the curriculum," "availability of teaching assistants," "a reasonable teaching load," "expertise in genome-related topics," "quality of computer resources," "quality of IT support," "acceptance of research in the curriculum," and "availability of computing facilities."

Similar themes emerged from a keyword analysis of faculty responses to one of the questions in the open survey: "What do you perceive as the most significant barrier opposing your efforts to teach genomics by engaging students in research?" (68 comments; Figure 2 and Supplemental Material S3). The most common theme concerned the difficulty of fitting the GEP material into the home institution's schedule. A related challenge was attempting to provide sufficient time within an established course for the genomics work.

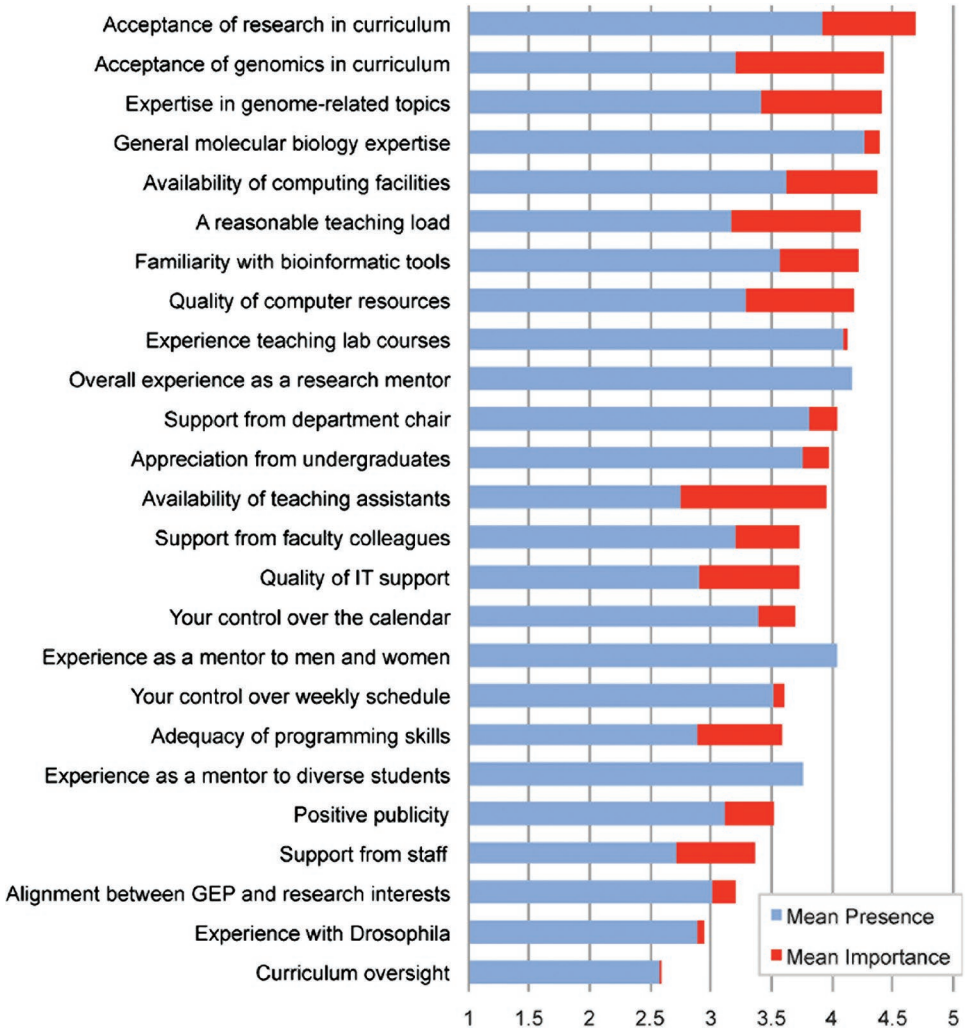

Figure 1. Faculty identification of barriers to implementing and sustaining a research-based lab course in genomics. Mean faculty ratings (on the anonymous survey), scoring both the importance (red bar) and the presence on campus (blue bar) of 25 items, at the time when the respondent attempted to implement genomics research lab activities. Respondents rated importance on a scale of 1 (marginally important) to 5 (very important), and rated presence on a scale of 1 (absent) to 5 (present in abundance). Items are sorted top to bottom by importance (red bar). The mean response for presence (blue bar) was superimposed over the red to highlight the difference; if presence exceeds importance, only the blue bar is visible. The difference between importance (red, what is needed) and presence (blue) suggests barriers to implementation. Numerical data are provided in Supplemental Material S8.

Other commonly cited barriers included the problem of cultivating capable teaching assistants and holding down class size. Less frequently cited barriers included problems with technology, student preparation, keeping up with pedagogy, and lack of colleague support. An independent inductive content analysis of the responses to this question identified the same top two concerns: the "fit" of the course within the wider curriculum and the availability of TA support (Table 1 and Supplemental Material S3). Other factors identified by inductive content analysis mimicked those above, although not in the same order. The consistency in results from three sources (anonymous survey checklist and analyses of the open comments by two independent evaluators [D.L. and J.T.]) supports the credibility of the assessment.

To further explore the faculty data, we attempted to organize the information in the 25 items by using the differences between "importance" and "presence" as an index of a barrier. The differences between the importance and presence responses were taken, and the new variable served as material for an exploratory factor analysis (same data set as used to construct Figure 1, given in numerical form in Supplemental Material S7). The best model from this analysis contains five factors. There were three survey items (appreciation from undergraduates, positive publicity, and experience with *Drosophila*) that did not load strongly on any of the five factors. Only the 22 survey items that did load strongly were used in the subsequent analysis. The five factors identified were conceptualized as follows: 1) items relating to the expertise and experience of the individual faculty member (referred to in the text below as "teaching/mentoring experience"); 2) items related to introducing genomics into the curriculum (familiarity with genomics); 3) administrative support for teaching activities (administrative teaching support); 4) support for computer-based activities (computing

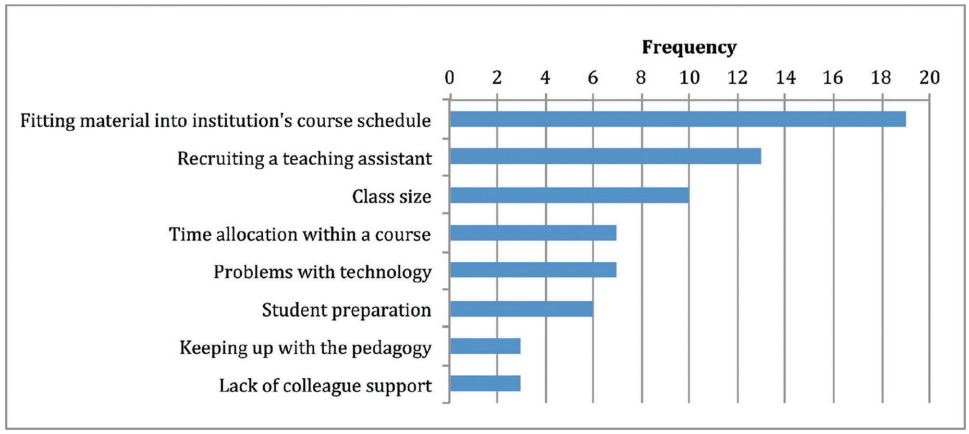

Figure 2. Frequency of the most significant barriers. The results shown are from keyword analysis of responses to the question "What do you perceive as the most significant barrier opposing your efforts to teach genomics by engaging students in research?" Open survey responses; data are presented in Supplemental Material S4.

support); and 5) direct teaching support expressed in terms of faculty teaching load and the presence of TAs (teaching support). (See Supplemental Material S8 for a list of survey items and factor loadings associated with each factor.) Overall, the results of this exploratory factor analysis suggest that the majority of items proposed by the GEP members for the survey did identify underlying variables reflective of five different types of barriers to the implementation and sustainability of GEP activities.

As a further exploration, the items in each subscale were summed, yielding five scores for each faculty respondent. We then used the five subscales as data for exploratory analyses of the possible influence of institutional characteristics (small college vs. research university, etc.; see Supplemental Material S8) but found no pattern of differences. The lack of differences does not mean that there are none, as our analyses lacked effective statistical power. Although we believe that the support provided by the GEP enables faculty to overcome barriers and successfully implement a genomics-based CURE at diverse institutions (based on anecdotal reports), we have no strong statistical evidence to support that belief based on our analysis.

Table 1. Analysis of responses to the question "What do you perceive as the most significant barrier opposing your efforts to teach genomics by engaging students in research?"

Theme	Number of faculty endorsing	Example quote
Fit with wider curriculum	23	"I would love to be able to offer Genomics as a stand-alone course or even as part of a lab course every year, but my teaching load won't allow it. There are too many high enrollment non-majors courses that need to be taught instead."
Finding TA support	19	"20 students in a class is difficult to manage without a TA. The fact that I cannot teach the class often causes potential TAs to graduate before the next offering of the course."
Time intensive	19	"The greatest barrier is simply the time required to instruct students and allow them the opportunity to find comfort in the project."
Student interest	15	"My main problem is in engaging students and getting them interested in the project. They want to do wet lab research in something that has direct practical application."
Technical support	12	"Lack of IT support and the need for upgraded computers have also been problems."
Challenging content	6	"The major barrier is the lack of exposure of students to genomics and bioinformatics in previous courses. It takes time to introduce students to so many tools and then have them use those tools to answer a real research question. Thus students are initially frustrated because the content of the course is so new, and the approach to teaching and learning is a unique experience for them."
Institutional buy-in	5	"Another barrier is resistance to change and innovation by some educators."
Own substantive knowledge	3	"Many GEP members are not actively engaged in genomics research. This makes it difficult for them to provide the expertise needed to teach students how to conduct genomics research in a class that is taught infrequently."

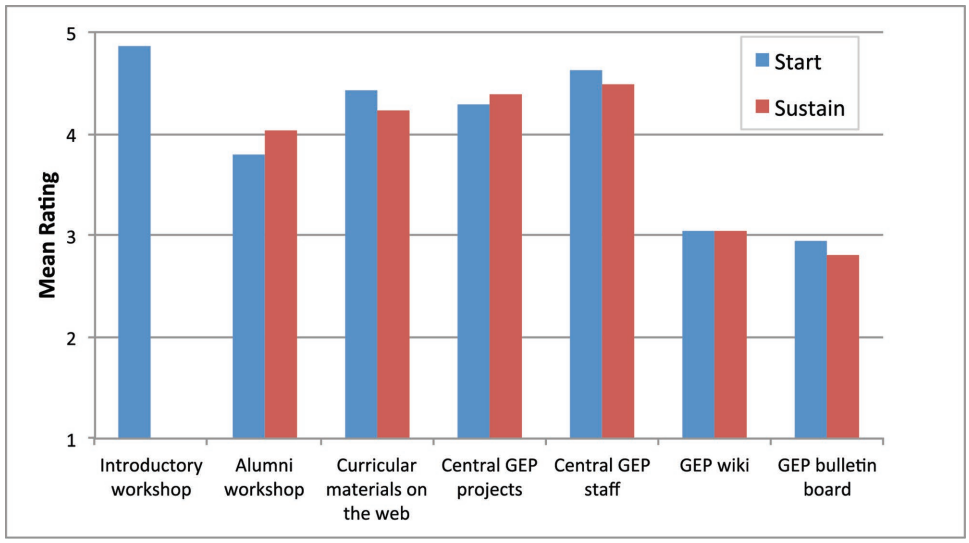

Figure 3. Faculty ratings of GEP assistance for starting and sustaining their research-based genomics lab. Faculty rated the importance of GEP resources/activities for starting (blue bars) or maintaining (red bars) their teaching using genomics research, using a scale of 1 (marginally important) to 5 (very important). Data from the anonymous faculty survey; means are shown.

How Can a Central Core System Help? Program Elements That Support Implementation and Sustainability

Given the barriers to implementation of CURE curriculum, what program elements do faculty members value most? Program features that influenced the faculty members' introduction and maintenance of genomics research in their courses were evaluated using an importance scale of 1–5 (Figure 3 and Tables 2 and 3). Not surprisingly, the introductory workshop was critical for initiating the program. Several centralized features of the program, including setting up the projects, maintaining curriculum, and follow-up staff support, received high marks for importance in both the decision to initiate and the ability to sustain the program, with a modal response of 5 (very important). While alumni workshops were ranked of less importance on average, and some faculty have not participated in this feature, many faculty do see these as important for staying current with the latest bioinformatics tools, ensuring progress toward publication, incorporating new pedagogical innovations, and maintaining a truly collaborative approach to both the research and curriculum (see comments in Supplemental Material S4). Wiki and bulletin board features, on the other hand, were rated very important by fewer than 20% of the respondents, even though there were 45 discussion threads on the bulletin board this past year. This low rating might be due to the willingness of GEP staff to respond to email and telephone queries (~2/d).

Faculty responses to the open survey question "Is a central support system (i.e., a centrally organized research project, shared training curriculum, central IT support) of continuing importance for your teaching genomics?" showed a strong consensus in favor of having a central organization (97%). Sixty-three respondents made 68 comments, many of which used terms such as "crucial" and "essential" to describe the role of the central support system in sustaining their efforts

Table 2. Distribution of responses to the question "What GEP resources or activities helped you to bring genomics research into your courses or curriculum (start up)?"

	Rating of importance (frequency)				
GEP resource or activity	1	2	3	4	5
Introductory workshop	0	0	1	9	75
Alumni workshops	10	3	10	18	31
Curricular materials on the Web	1	1	12	18	52
Central GEP projects	1	2	10	25	40
GEP wiki (Table of Faculty, other)	10	9	24	18	7
GEP bulletin board (frequently asked questions)	13	6	28	9	10
Central GEP staff to help troubleshoot, etc.	0	1	7	14	59

Table 3. Distribution of responses to the question "What GEP resources or activities have helped you maintain genomics research in your courses or curriculum (sustainability)?"

	Rating of importance (frequency)				
GEP Resource or activity	1	2	3	4	5
Alumni workshops	5	4	9	16	34
Curricular materials on the Web	4	3	7	20	43
Central GEP projects	0	1	12	16	42
GEP wiki (Table of Faculty, other)	12	6	18	17	8
GEP bulletin board (frequently asked questions)	16	11	14	9	11
Central GEP staff to help troubleshoot, etc.	0	3	10	10	52

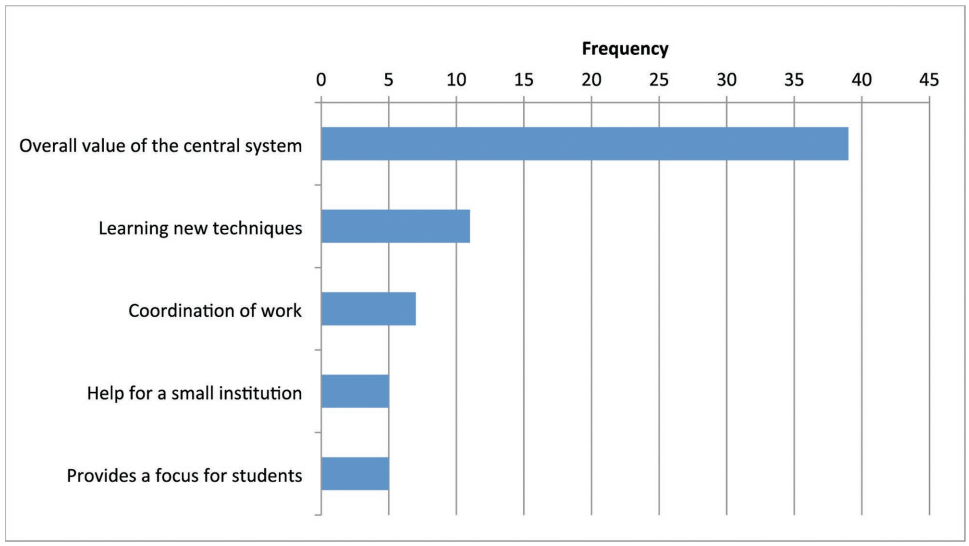

Figure 4. Importance of a central support system. Keyword analysis of responses to the question "Is a central support system (i.e., a centrally organized research project, shared training curriculum, central IT support) of continuing importance for your teaching genomics?" Open survey responses; data are presented in Supplemental Material S5.

(keyword analysis, Figure 4 and Supplemental Material S4). Reasons cited included the support for troubleshooting, the economical use of resources, and the reliance on the central system for community. Several respondents indicated that, without the central support system, they would not be able to sustain the program at their institution. Inductive content analysis of the open responses to this question identified themes of access to significant research, teaching resources, scientific expertise, technical support, and community (Table 4 and Supplemental Material S4).

What Are the Incentives for Faculty? Why Do They Persist in Maintaining a CURE?

What are the incentives that drive faculty to take on the challenge of creating and maintaining a research-based lab course? In response to the question "What do you perceive as the most significant incentive for sustaining your efforts to teach genomics by engaging students in research?" 66 faculty members made 98 comments (Supplemental Material S5). Keyword analysis (Figure 5) indicated that the most frequent comments were about research opportunities for students, particularly involving "high-profile" or "prestigious" research, followed by comments about the value of active learning. Less-frequent but still common responses included the value for both students and faculty of being involved with potential publications (e.g., "The ability to coauthor various publications is a strong incentive"), the value of forming a community of scholars (e.g., "My main incentive has been to connect to a larger community of scientists"), and the value to faculty members of continuing their development as instructors and researchers (e.g., "This forum

Table 4. Analysis of responses to the question "Is a central support system (i.e., a centrally organized research project, shared training curriculum, central IT support) of continuing importance for your teaching genomics?"

Theme	Number of faculty endorsing	Example quote
Access to significant research	36	"As a small institution with limited research resources, we rely on initiatives like the GEP to provide the centralized 'big picture' question to which our students can contribute. We can, of course, devise our own more local projects, but the scale of the GEP's research and the opportunity for collaboration with students from other institutions are large motivating factors for our students to become involved."
Access to teaching resources	29	"Having a community of faculty working on the same pedagogical challenge is essential to our success .... Discussions of our experiences in implementing the curriculum and mutual support during alumni meetings help solve the challenges we encounter."
Access to scientific expertise	15	"Genomics is such a rapidly developing field, it would be difficult for most teachers to keep up while pursuing all the other commitments in teaching, professional development and service. Having central organization that helps us keep on top of new research, computational tools and pedagogical approaches is essential."
Access to technical support	14	"The availability of expert IT help dedicated full time to making the system work … could not be replicated at my home institution."
Access to community	10	"Discussing my successes and challenges with like-minded colleagues has been helpful and motivating."

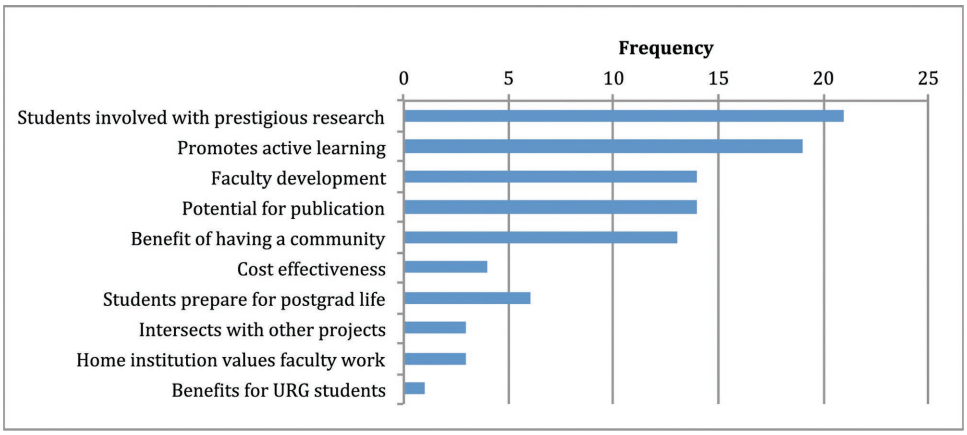

Figure 5. Faculty incentives. Keyword analysis of responses to the question "What do you perceive as the most significant incentive for sustaining your efforts to teach genomics by engaging students in research?" Open survey responses; data are presented in Supplemental Material S6.

provides an opportunity for me to grow professionally and network with other scientists"). Inductive content analysis, done independently (by J.T.), identified the same two themes as most important ("participation in real research" and "increases student learning"), as well as the opportunity to contribute to the field and to participate in a scientific community (Table 5). Interestingly, many faculty members find they are motivated by observing "student learning" that goes beyond the student simply gaining factual knowledge but instead indicates that the students' relationship with knowledge has changed. Examples include "Student enthusiasm and success keep me motivated," "I am primarily motivated by the awareness of the depth of insight students can gain," and "The students in my classes become so engaged in science research, with frustration and then elation upon finding a solution, that I, myself, get excited." Similar rewards (a sense of participating in significant research, gains in understanding from hands-on work) have been cited by GEP students surveyed 1–5 yr after taking a GEP-affiliated course (Shaffer *et al.*, 2014).

Table 5. Analysis of responses to the question "What do you perceive as the most significant incentive for sustaining your efforts to teach genomics by engaging students in research?"

Theme	Number of comments	Example quote
Participation in real research	37	"They are able to apply their learning and creativity to address real scientific questions, despite the inherent frustrations they encounter while doing novel research."
Increases student learning	23	"My main incentive to continuing this effort is the benefit that I have seen the students gain from this experience. It teaches them important content about gene structure, genomics, and bioinformatics in addition to how to solve a real scientific problem using critical thinking skills."
Contribution to field	20	"The most significant incentive for teaching genomics by engaging students in research is the opportunity to produce new knowledge that will result in publications in the primary scientific and education literature, a benefit for both the students and myself."
Scientific community	18	"It does also keep me connected to a larger group of like-minded educators and scientists who want to have more students experience this type of collaborative research project; without the GEP, none of this could be accomplished."
Keeping up with the field	14	"Continued involvement in GEP forces me to stay abreast of the latest developments in the rapidly changing field of genomics, which benefits my students and enriches me professionally as a molecular biologist."
Prepares students for the future	14	"Our students come away with a genuine passion for research that leads them into ideas for their careers that they did not consider before."
Feasibility	12	"A big incentive for me is to be able to involve entire classes in novel research on a very low budget."
Increases student motivation	10	"Research goals give meaning and immediate application to the knowledge and skills students acquire. That's a powerful motivator for students to work hard to succeed and for me to continue supporting them."
Valued by institution	6	"My institution regards my involvement in GEP as contributing to my career advancement through continuing professional scholarship and by being able to offer an innovative lab for students."
Credence	3	"Knowing that significant resources of major scientific institutions (WUSTL and HHMI) are invested in the project gives additional assurance that the investment of my time in this effort is more likely to result in a lasting educational and research resource that will keep up with developments in the field."

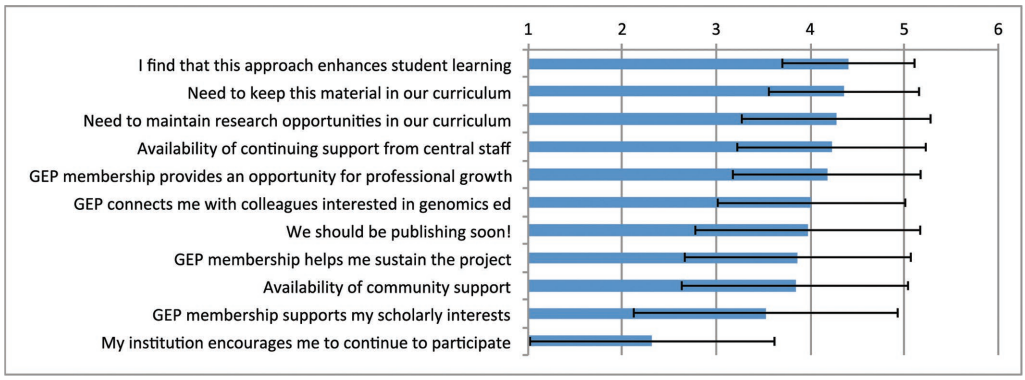

Figure 6. Faculty ratings of reasons for persistence. Faculty members rated their reasons for continuing as a member of the GEP on a scale of 1 (marginally important) to 5 (very important). Data from the anonymous faculty survey. Means and SDs shown.

Faculty members rated their reasons for continuing as a GEP member among a set of options posed on the anonymous survey using a 5-point scale of importance (Figure 6 and Table 6). The most important reason selected for remaining an active member was "I find that this approach enhances student learning," followed by "need to keep this material in our curriculum," and "need to maintain research opportunities in our curriculum." Other reasons, including "availability of community support" and "GEP membership supports my scholarly interests," were rated more modestly, and "My institution encourages me to continue to participate" received a considerably lower rating than all other items. We infer from these ratings that GEP faculty members are motivated by their belief that genomics and active research should be a part of their curricula, but that these goals receive only modest support from their home institutions. However, several faculty members indicated in their open responses that they thought GEP participation had contributed to positive promotion and tenure decisions (e.g., "It has helped to advance my career").

DISCUSSION

Our overall hypothesis is that collaborative research projects can help overcome the present barriers to establishing research-based undergraduate science curricula. The growth of the GEP is an example of the effectiveness of this approach. The GEP joins a growing number of collaborative programs; these include GCAT, GCAT-SEEK, PHIRE, and SEA-PHAGES. We further investigated the barriers and incentives for GEP members by soliciting both anonymous and attributed responses from the GEP faculty cohort. We have identified the most common barriers to implementing a genomics-based classroom research program by identifying the gap between perceived importance and current presence of a number of teaching resources. We found support through both faculty ratings and faculty reports for the central core model that the GEP represents. Finally, we identified the incentives reported by faculty as being instrumental in their continued engagement with the partnership.

Barriers to science education reform, particularly the change to active-learning strategies, have been previously identified (Henderson and Dancy, 2007; Dancy and Henderson, 2008; Spell *et al.*, 2014). Henderson and Dancy (2007), for example, interviewed physics instructors and found reports of many of the difficulties frequently mentioned by GEP faculty members—problems with scheduling, class size, and time structure (Figure 2 and Table 1). While some of these factors are tied to financial constraints (e.g., class size), others have more to do with attitudes that presumably could be addressed at an administrative/community level

Table 6. Distribution of responses to the question "Why have you stayed as an active member of GEP?"

	Rating of importance (frequency)				
GEP resource or activity	1	2	3	4	5
I find that this approach enhances student learning.	0	3	3	33	43
Need to keep this material in our curriculum	0	1	11	26	42
Need to maintain research opportunities in our curriculum	2	4	3	30	39
Availability of continuing support from central staff	3	1	9	24	37
GEP membership provides an opportunity for professional growth.	1	4	12	25	37
GEP connects me with colleagues interested in genomics education.	2	2	17	26	28
We should be publishing soon!	4	6	11	19	23
GEP membership helps me sustain the project.	5	5	12	21	28
Availability of community support.	6	3	15	22	27
GEP membership supports my scholarly interest.	9	7	16	18	23
My institution encourages me to continue to participate.	23	12	16	6	5

on each campus. We (and our curricular committees) need to recognize that research-based laboratories frequently require larger blocks of time and greater flexibility in scheduling than is the norm (Shaffer *et al.*, 2014), and that we will need to adjust our scheduling systems accordingly if we wish to change our pedagogical approach. However, despite the enthusiasm of national groups interested in better STEM education for greater utilization of CUREs (AAAS, 2011; PCAST, 2012; Jordan *et al.*, 2014), it is notable that a lack of "acceptance of research within the curriculum" remains a significant barrier at the grassroots level (Figure 1). For example, one response on the anonymous survey stated "Basically, most faculty are not convinced that students can master content and learn more in a 'research in the classroom' based approach. The criticism I hear is that it may be true at other institutions, but not at our institution with the students we typically attract." Such sentiments are generally not expressed in public, but our findings suggest they are widespread and need to be addressed by continuing research to establish the efficacy of this approach. Most of the data arguing that research experiences enhance learning and retention in the sciences have been obtained from studies of students engaged in individual or small-group mentored research in a faculty lab (Hunter *et al.*, 2007; Lopatto, 2007, 2009; Locks and Gregerman, 2008; Laursen *et al.*, 2010). However, there is increasing evidence and documentation of positive learning outcomes using CUREs (Campbell *et al.*, 2007; Lopatto *et al.*, 2008; Shaffer *et al.*, 2010, 2014; Burnette and Wessler, 2013; Jordan *et al.*, 2014). Nonetheless, there is no doubt that more work is needed to assess CUREs, to understand the process and optimize results for students (Auchincloss *et al.*, 2014).

Another common challenge is that research-based active-learning environments often benefit from having a lower number of students per instructor than is needed for lectures or traditional labs. We find it advantageous to have one knowledgeable person present for every six to seven novices (Shaffer *et al.*, 2010). In the case of GEP, this challenge is addressed by a modest investment in support of peer instructors, as undergraduates who have done well in a GEP course are generally excellent facilitators. The strategy of recruiting top students from a given year to serve as TAs/peer instructors the following year works well at many schools but can collapse if the GEP-affiliated course is offered only every other year; the lack of "availability of teaching assistants" is a frequently cited barrier (Figures 1 and 2, Table 1, and Supplemental Material S3). How far a strategy of peer mentoring could be extended is an important question that remains to be explored. Given a bioinformatics platform, which reduces infrastructure needs and safety concerns, the need for trained TAs may be the limiting factor in our desire to reach larger numbers of students. Our present experience suggests that using a CURE can allow a faculty member to provide research experiences for two- to 10-fold more students than by traditional means (a faculty member who might have two to six students in his/her lab may have four to 40 students in a GEP-based CURE), but whether this project could be managed to provide a research experience for hundreds of students at one site is unknown. Other CUREs have experienced difficulty in scaling up to that level (Brownell *et al.*, 2013).

Other barriers, including technology issues and the need for expertise in the research area, appear to be offset by the support provided by the central core system (Supplemental Material S3 and S5; contrast results in Table 1 with Figure 4). Further, it is possible that the support provided by the GEP community (see Table 5) may offset the lack of administrative support that can occur on campus. Analyses of the reported faculty incentives for sustaining a CURE (Table 5 and Figure 6) indicate strong personal commitment on the part of the faculty to providing research experiences for students, grounded in faculty observations that this approach enhances student learning. The evidence indicates that a central core can help faculty members overcome the many barriers to achieve this outcome. While some of the lessons learned from our survey on the utility of a collaborative research project supported by a core system are specific to managing a research project utilizing a bioinformatics approach, most are broadly applicable.

Using bioinformatics as the platform for undergraduate research simplifies the need for lab infrastructure, as it requires only computers and Internet access. Furthermore, the use of a computer-based bioinformatics approach minimizes supply costs. At a cost of less than $200 per college student (this includes all wet-bench work, all technical support, infrastructure maintenance/development, and travel and expenses for new faculty/TA training and continuing GEP faculty workshops), this project has developed a cost-effective strategy for providing research experiences for more college students. Equally important, students are learning the analytical skills they will need in the coming years as sequencing becomes less expensive and the ability to handle large data sets is deemed essential (Shaffer *et al.*, 2010, 2014). Similar success by other consortia using bioinformatics supports this conclusion (Hatfull *et al.*, 2006; Campbell *et al.*, 2007; Ditty *et al.*, 2010; Banta *et al.*, 2012; Jordan *et al.*, 2014). The GEP's inclusion of undergraduates in hypothesis-driven experiential research directly addresses the recommendations of the PCAST (2012) and is producing novel insights into the regulation of chromatin structure by employing comparative genomics (Leung *et al.*, 2010, unpublished data). The project has brought together faculty whose members share common interests in genetics/genomics research and pedagogy. Given the diverse topics encompassed within a contemporary biology department, our results suggest that national research projects that bring together other groups of faculty members with shared interests in a biology subfield may be one of the most practical means of supporting large-scale pedagogical change to embrace research-based laboratory experiences.

ACKNOWLEDGMENTS

We thank the many students who have participated in GEP-affiliated courses since 2006 and Frances Thuet for setting up the assessment websites and helping with data collection. We also thank the many Washington University undergraduates and staff members of the Genome Institute who have served as TAs in the GEP workshops, as well as staff who have helped to organize and facilitate these meetings.

REFERENCES

American Association for the Advancement of Science (2011). Vision and Change in Undergraduate Biology Education: A Call to Action, Washington, DC: http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf (accessed 4 July 2013).

Auchincloss LC, Laursen SL, Branchaw JL, Eagan K, Graham M, Hanauer DI, Lawrie G, McLinn CM, Pelaez N, Rowland S, *et al.* (2014). Assessment of course-based undergraduate research experiences: a meeting report. CBE Life Sci Educ *13*, 29–40.

Banta LM, Crespi EJ, Nehm RH, Schwarz JA, Singer S, Manduca CA, Bush EC, Collins E, Constance CM, Dean D, *et al.* (2012). Integrating genomics research throughout the undergraduate curriculum: A collection of inquiry-based genomics lab modules. CBE Life Sci Educ *11*, 203–208.

Brownell SE, Kloser MJ, Fukami T, Shavelson RJ (2013). Context matters: volunteer bias, small sample size, and the value of comparison groups in the assessment of research-based undergraduate introductory biology lab courses. J Microbiol Biol Educ *14*, 176–182.

Buonaccorsi V, Peterson M, Lamendella G, Newman J, Trun N, Tobin T, Aguilar A, Hunt A, Praul C, Grove D, *et al.* (2014). Vision and change through the Genome Consortium for Active Teaching using next-generation sequencing (GCAT-SEEK). CBE Life Sci Educ *13*, 1–2.

Burnette JMIII, Wessler SR (2013). Transposing from the laboratory to the classroom to generate authentic research experiences for undergraduates. Genetics *193*, 367–375.

Campbell AM, Ledbetter ML, Hoopes LL, 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.

Dancy M, Henderson C (2008). Barriers and promises in STEM reform. Commissioned paper for the National Academies of Science Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education, held 13–14 October 2008 in Washington, DC.

Desai KV, Gatson SN, Stiles TW, Stewart RH, Laine GA, Quick CM (2008). Integrating research at research-extensive universities with research-intensive communities. Adv Physiol Educ *32*, 136–141.

Ditty JL, Kvaal CA, Goodner B, Freyermuth SK, Bailey C, Britton RA, Gordon SG, Heinhorst S, Reed K, Xu Z, *et al.* (2010). Incorporating genomics and bioinformatics across the life sciences curriculum. PLoS Biol *8*, e1000448.

Elo S, Kyngäs H (2008). The qualitative content analysis process. J Adv Nurs *62*, 107–115.

Glaser BG, Strauss AL (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research, New York: Aldine.

Harris SE, Bellino M (2014). DNA barcoding from NYC to Belize. Science *342*, 1462–1463.

Hatfull GF, Pedulla ML, Jacobs-Sera D, Cichon PM, Foley A, Ford ME, Gonda RM, Houtz JM, Hryckowian AJ, Kelchner VA, *et al.* (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genet *2*, e92.

Healey M., Jenkins A (2009). Developing Undergraduate Research and Inquiry, York, UK: Higher Education Academy.

Henderson C, Dancy M (2007). Barriers to the use of research-based instructional strategies: the influence of both individual and situational characteristics. Phys Rev Spec Top Phys Educ Res *3*, 020102.

Hsieh HF, Shannon SE (2005). Three approaches to qualitative content analysis. Qual Health Res *15*, 1277–1288.

Hunter A-B, Laursen SL, Seymour E (2007). Becoming a scientist: the role of undergraduate research in students' cognitive, personal, and professional development. Sci Educ *91*, 36–74.

Jordan TC, Burnett SH, Carson S, Caruso S, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, *et al.* (2014). A broadly implementable research course for first-year undergraduate students. MBio *5*, e01051.

Karukstis KK (2008). Broadening participation in undergraduate research. J Chem Educ *85*, 1474–1476.

Kloser MJ (2011). Integrating teaching and research in undergraduate biology laboratory education. PLoS Biol *9*, e1001174.

Krippendorff K (2013). Content Analysis: An Introduction to Its Methodology, Thousand Oaks, CA: Sage.

Laursen SL, Hunter A-B, Seymour E, Thiry H, Melton G (2010). Undergraduate Research in the Sciences: Engaging Students in Real Science, San Francisco, CA: Jossey-Bass.

Leung W, Shaffer CD, Cordonnier T, Wong J, Itano MS, Slawson Tempel EE, Kellmann E, Desruisseau DM, Cain C, *et al.* (2010). Evolution of a distinct genomic domain in *Drosophila*: comparative analysis of the dot chromosome in *Drosophila melanogaster* and *Drosophila virilis*. Genetics *185*, 1519–1534.

Locks AM, Gregerman SR (2008). Undergraduate research as an institutional retention strategy: the University of Michigan model. In: Creating Effective Undergraduate Research Programs in Science: The Transformation from Student to Scientist, ed. R Taraban and RL Blanton, New York: Teachers College Press, 11–32.

Lopatto D (2007). Undergraduate research experiences support science career decisions and active learning. CBE Life Sci Educ *6*, 297–306.

Lopatto D (2009). Science in Solution: The Impact of Undergraduate Research on Student Learning, Tucson, AZ: Research Corporation for Science Advancement. http://web.grinnell.edu/sureiii/Science_in_Solution_Lopatto.pdf (accessed 27 January 2014).

Lopatto D, Alvarez C, Barnard D, Chandrasekaran C, Chung HM, Du C, Eckdahl T, Goodman AL, Hauser C, Jones CJ, *et al.* (2008). Undergraduate research: Genomics Education Partnership. Science *322*, 684–685.

Miles MB, Huberman AM (1994). Qualitative Data Analysis: An Expanded Sourcebook, 2nd ed., Thousand Oaks, CA: Sage.

President's Council of Advisors on Science and Technology (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering and Mathematics, Washington, DC: U.S. Government Office of Science and Technology. www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf (accessed 4 July 2013).

Rowlett RS, Blockus L, Larson S (2012). Characteristics of excellence in undergraduate research (COEUR), ed. Nancy Hensel, Washington, DC: Council on Undergraduate Research, 2–19. www.cur.org/assets/1/23/COEUR_final.pdf (accessed 17 November 2014).

Shaffer CD, Alvarez C, Bailey C, Barnard D, Bhalla S, Chandrasekaran C, Chandrasekaran V, Chung HM, Dorer DR, Du C, *et al.* (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.

Shaffer CD, Alvarez CJ, Bednarski AE, Dunbar D, Goodman AL, Reinke C, Rosenwald AG, Wolyniak MJ, Bailey C, Barnard D, *et al.* (2014). A course-based research experience: how benefits change with increased investment in instructional time. CBE Life Sci Educ *13*, 111–130.

Spell RM, Guinan JA, Miller KR, Beck CW (2014). Redefining authentic research experiences in introductory biology laboratories and barriers to their implementation. CBE Life Sci Educ *13*, 102–110.

Winningham RG, Templeton JH, Dutton BE, Scheck SH (2009). A grassroots, faculty-driven initiative to institutionalize undergraduate research: the ins and outs of cultivating administrative support. Counc Undergrad Res Q *30*, 29–34.

Wood WB (2003). Inquiry-based undergraduate teaching in the life sciences at large research universities: a perspective on the Boyer Commission Report. Cell Biol Educ *2*, 112–116.