UNDERGRADUATE RESEARCH

Genomics Education Partnership

The Genomics Education Partnership offers an inclusive model for undergraduate research experiences, with students pooling their work to contribute to international databases.

D. Lopatto,[1]* C. Alvarez,[2] D. Barnard,[3] C. Chandrasekaran,[4] H.-M. Chung,[5] C. Du,[6] T. Eckdahl,[7] A. L. Goodman,[8] C. Hauser,[9] C. J. Jones,[10] O. R. Kopp,[11] G. A. Kuleck,[12] G. McNeil,[13] R. Morris,[14] J. L. Myka,[15] A. Nagengast,[14] P. J. Overvoorde,[16] J. L. Poet,[17] K. Reed,[18] G. Regisford,[19] D. Revie,[20] A. Rosenwald,[21] K. Saville,[22] M. Shaw,[23] G. R. Skuse,[24] C. Smith,[25] M. Smith,[26] M. Spratt,[27] J. Stamm,[28] J. S. Thompson,[29] B. A. Wilson,[30] C. Witkowski,[31] J. Youngblom,[32] W. Leung,[33] C. D. Shaffer,[33] J. Buhler,[33] E. Mardis,[33] S. C. R. Elgin[33]*

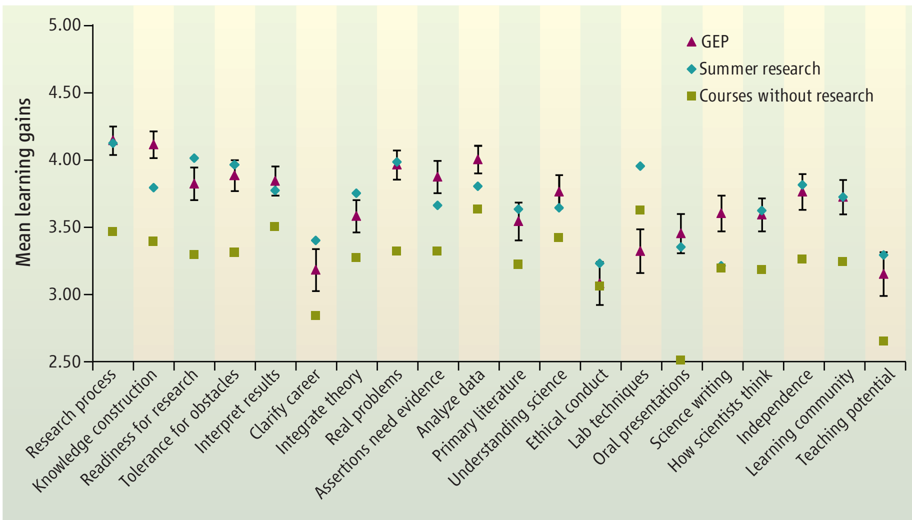
Teaching students to think as scientists do has traditionally been accomplished by immersing undergraduates in a full-time summer research program, an experience that enhances professional development and provides career clarification (*1*, *2*). We believe that conducting research is the best way to learn how knowledge is defined and created in a field. Furthermore, students acquire a deeper understanding of fundamental concepts as they apply what they have learned to accomplish defined research goals. Undergraduate research experiences can sustain student interest in a science career, providing an opportunity to work collaboratively with colleagues while making novel contributions to the community (*1*).

Most undergraduates begin their research with mentoring by a faculty member, graduate student, or postdoctoral student (postdoc) during a summer spent in the laboratory. This approach, while usually successful, excludes many students who do not have the summer free or funded. Other barriers include limited funding and facilities, as well as a lack of experienced mentors (*1*, *3*). Thus, we must incorporate student research experiences into our regular academic-year curriculum to make such experiences more broadly available. Our experiences as founding members of the Genomics Education Partnership (GEP) encourage us to believe that a curriculum in genomics can train students to think like scientists (*4*).

Genomics for Undergraduates

Genomics is an attractive area for student-scientist partnerships. First, genomics represents an exciting advance in the life sciences. With genomics, we can analyze not only one gene, but many genes, and can address questions ranging from patterns of gene expression to evolution of a species. Second, the teaching and research materials can be obtained for little or no cost: DNA sequences and microarray results are freely available on the Web (*5*, *6*). Third, genomics research projects can range from simple to complex, which allows engaging in them to match a variety of academic calendars.

Mindful of these advantages, a number of faculty groups and institutions have launched projects to bring genomics into the undergraduate curriculum. These range from primarily utilizing computer-based resources for annotation and analysis to having a major wet-laboratory focus (*7*, *8*). Two major projects engage students in analyzing genomes from phage and from bacteria or archaea (*9*, *10*) (see supporting online text). GEP was formed in 2006 as a collaboration between faculty from various institutions and the Biology Department and Genome Sequencing Center at Washington University in Saint Louis. The group currently consists of over 40 faculty members who teach undergraduate courses in biology, statistics, mathematics, and computer science at a variety of institutions, including research universities, master's degree–granting comprehensive universities, and private baccalaureate-granting colleges (*4*). The goal of the GEP is to provide opportunities for undergraduate students to participate in original research



Student ratings of learning gains. Mean learning gains for common survey items on the SURE, CURE, and GEP surveys. The SURE survey data represent 1973 undergraduate summer research experiences ("Summer research" data points). The CURE survey data represent 560 student surveys from science courses that did not include a research-like experience "Courses without research" data points). The GEP survey data represent 308 students in GEP courses (SOM). Error bars represent 2 SEM.

[1]Grinnell College, Grinnell, IA. [2]Longwood University, Farmville, VA. [3]Worcester State College, Worcester MA. [4]Texas Wesleyan University, Fort Worth, TX. [5]University of West Florida, Pensacola, FL. [6]Montclair State University, Montclair, NJ. [7]Missouri Western State University, St. Joseph, MO. [8]California Polytechnic State University–San Luis Obispo, CA. [9]St. Edward's University, Austin, TX. [10]Moravian College, Bethlehem, PA. [11]Utah Valley University, Orem UT. [12]Loyola Marymount University, Los Angeles, CA. [13]York College, The City University of New York, Jamaica, NY. [14]Widener University, Chester, PA. [15]Galen College of Nursing, Cincinnati, OH. [16]Macalester College, St. Paul, MN. [17]Missouri Western State University , St. Joseph, MO. [18]Austin College, Sherman, TX. [19]Prairie View A&M University, Prairie View, TX. [20]California Lutheran University, Thousand Oaks, CA. [21]Georgetown University, Washington, DC. [22]Albion College, Albion, MI. [23]New Mexico Highlands University, Las Vegas, NM. [24]Rochester Institute of Technology, Rochester, NY. [25]San Francisco State University, San Francisco, CA. [26]North Carolina Agricultural & Technical State University, Greensboro, NC. [27]William Woods University, Fulton, MO. [28]University of Evansville, Evansville, IN. [29]Denison University, Granville, OH. [30]Jackson State University, Jackson, MS. [31]Missouri State University, Springfield, MO. [32]California State University–Stanislaus, Turlock, CA. [33]Washington University in St. Louis, St. Louis, MO. (For complete addresses, see supporting online material.)

*To whom correspondence should be addressed. E-mail: lopatto@grinnell.edu (D.L.); selgin@biology.wustl.edu (S.C.R.E.).

experiences within the framework of course curricula. The GEP pools student efforts across institutions to improve DNA sequence quality and to generate hand-curated gene models. Developing a coordinated research effort allows us to tackle large projects together and to make work in eukaryotes accessible. The first project undertaken by GEP members was sequence improvement and annotation of the dot chromosomes of several *Drosophila* species. The distal 1.2 megabases of this unusual small chromosome are of interest because the domain appears to be heterochromatic, a packaging state normally associated with gene silencing, but it has the same gene density of more transcriptionally active euchromatic regions (*11*). Because of the high density of repetitious elements, improving the sequence to verify the assembly, close gaps, and raise the quality estimate to 1 error in 1000 bases provides greater credibility to the analysis. Annotation then follows, with students generating the most plausible gene models based on available evidence, and then defending their models (*4*). (See supporting online text for a more detailed description.)

Economies in mentoring can be realized because the students are all using similar tools and strategies as they work independently on a specific region. The amount of sequence assigned to a student can be adjusted according to the available time. During a typical semester laboratory course, one student can improve and annotate 40 kb of *Drosophila* DNA. By organizing the raw data into 40-kb packages that can be "claimed" from a Web site, the GEP can undertake fairly large projects for sequence improvement and annotation.

Participating course instructors integrate this research into their courses using a variety of formats. Results from the individual students are then checked and pooled at Washington University to create the final dot chromosome assembly that is used to generate a variety of statistics documenting the nature of the chromosome and the genes present. Comparing dot chromosomes of several *Drosophila* species allows us to address issues of chromosome evolution.

Assessment and Sustainability

Students know that their work will be used for a scientific publication and deposited in a public database for other scientists to build from. Sequence data are submitted to GenBank, and annotation results to FlyBase. One scientific paper has been published based on student work (*12*); contributing students are coauthors. We anticipate an accumulation of publications and database contributions with time. The significance of this is illustrated by some of the comments from students and teaching assistants (TAs) (table S1).

Students completed a postcourse survey of their perceptions. This survey, a variation of the Classroom Undergraduate Research Experiences survey (CURE) (*4*), parallels one evaluating summer undergraduate research experiences (SURE) (*1*, *13*). GEP students report learning and professional gains similar to those reported by students who have spent a summer in the laboratory (see the figure on page 684 and table S2). High ratings from GEP students on "knowledge construction," "assertions need evidence," "analyze data," and "science writing" are similar to other research-like experiences (*14*, *15*). From the perspective of students, the GEP course experience is more like summer research than like a standard science course without research. As the GEP matures, more assessment protocols are being added. Currently, the faculty are devising content tests for annotation and finishing. The GCAT has reported that students in that program make significant gains in comprehension of topics in functional genomics and show increased interest in research (*16*).

Research opportunities in genomics are likely to increase. There are now 11 sequenced *Drosophila* species in addition to *D. melanogaster*, the commonly used laboratory species (*17*). By choosing regions of biological interest, so that the results contribute to research papers, as well as to the databases, much can be accomplished by a student-scientist partnership that otherwise might be difficult to attain.

We invite scientists to consider what might be accomplished by a distributed community of undergraduate scientists collaborating on a particular set of genes or genomes. At present, the GEP has over 40 members, and membership is growing (*4*). GEP faculty and their (usually undergraduate) TAs come to Washington University for a 1-week workshop to learn the relevant software and to discuss pedagogical issues. As the program develops at an institution, the previous year's students can be recruited as the next year's TAs. This strategy fosters sustainability and allows class size to grow. Efficient implementation favors one experienced person (the TA or faculty member) for every six or seven novices to provide support for students learning to use software and to evaluate data. Although most institutions already maintain computer laboratories, which minimizes start-up costs, training for faculty and TAs is essential. Faculty and project leaders must agree on a common rubric for sequence improvement and annotation goals and on a common reporting mechanism. Good on-site information technology support is important, but the basic strategy is very portable. The opportunity to teach about genes, genomes, and the flow of biological information using this hands-on approach provides entry into a rich domain of data. We have found that involving undergraduates in a genomics research project is a rewarding way for faculty to teach and for undergraduates to learn.

References and Notes

1. D. Lopatto, *CBE Life Sci. Educ.* **6**, 297 (2007).
2. E. Seymour *et al.*, *Sci. Educ.* **88**, 493 (2004).
3. S. H. Russell, in *Creating Effective Undergraduate Research Programs in Science*, R. Taraban, R. L. Blandon, Eds. (Teachers College Press, New York, 2008), pp. 53–80.
4. Genomics Education Partnership (http://gep.wustl.edu). This site includes the list of member institutions and a facsimile of the GEP assessment instrument.
5. Summary of Principles Agreed at the First International Strategy Meeting on Human Genome Sequencing, www.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml#1.
6. *Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility*, report of a meeting organized by the Wellcome Trust, Fort Lauderdale, FL, 14 and 15 January 2003; www.genome.gov/Pages/Research/WellcomeReport0303.pdf.
7. For examples, see this reference and (*8*); *Teaching Big Science at Small Colleges*, http://serc.carleton.edu/genomics/index.html.
8. *Genome Consortium for Active Teaching* (GCAT) introduces students to microarray experiments; www.bio.davidson.edu/projects/gcat/gcat.html.
9. The Science Education Alliance of HHMI works on phage, www.hhmi.org/grants/sea/.
10. The DOE Joint Genome Institute sponsors work on bacteria, www.jgi.doe.gov/education/genomeannotation.html.
11. N. C. Riddle, S. C. R. Elgin, *Chromosome Res.* **14**, 405 (2006).
12. E. E. Slawson *et al.*, *Genome Biol.* **7**, R15 (2006).
13. D. Lopatto, in *Creating Effective Undergraduate Research Programs in Science*, R. Taraban, R. L. Blandon, Eds. (Teachers College Press, New York, 2008), pp. 112–132.
14. L. A. Denofrio *et al.*, *Science* **318**, 1872 (2007).
15. C. Trosset *et al.*, in *Creating Effective Undergraduate Research Programs in Science*, R. Taraban, R. L. Blandon, Eds. (Teachers College Press, New York, 2008), pp. 33–49.
16. A. M. Campbell *et al.*, *CBE-Life Sci. Educ.* **6**, 109 (2007).
17. A. Stark *et al., Nature* 450, 219 (2007).

18. We thank F. Thuet (Washington University) for management of the GEP CURE assessment Web site; L. Crosby and W. Gelbart of FlyBase (Harvard University); the staff of the Washington University Genome Sequencing Center for assistance; and the Howard Hughes Medical Institute for financial support (grant 52005790 to S.C.R.E., HHMI Professor).

Supporting Online Material
www.sciencemag.org/cgi/content/full/322/5902/684/DC1

10.1126/science.1165351