Title
An inclusive Research Education Community (iREC): Impact of the SEA-PHAGES program on research outcomes and student learning

Journal
Proceedings of the National Academy of Sciences of the United States of America, 114(51)

ISSN
0027-8424

Authors
Hanauer, David I
Graham, Mark J
Betancur, Laura
et al.

Publication Date
2017-12-19

DOI
10.1073/pnas.1718188115

Peer reviewed

An inclusive Research Education Community (iREC): Impact of the SEA-PHAGES program on research outcomes and student learning

David I. Hanauer[a], Mark J. Graham[b], SEA-PHAGES[1], Laura Betancur[c], Aiyana Bobrownicki[b], Steven G. Cresawn[d], Rebecca A. Garlena[e], Deborah Jacobs-Sera[e], Nancy Kaufmann[e], Welkin H. Pope[e], Daniel A. Russell[e], William R. Jacobs Jr.[f,2], Viknesh Sivanathan[g], David J. Asai[g,2], and Graham F. Hatfull[e,2]

[a]Department of English, Indiana University of Pennsylvania, Indiana, PA 15705; [b]Center for Teaching and Learning, Yale University, New Haven, CT 06511; [c]Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260; [d]Department of Biology, James Madison University, Harrisonburg, VA 22817; [e]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260; [f]Department of Microbiology and Immunology, Albert Einstein College of Medicine, New York, NY 10461; and [g]Science Education, Howard Hughes Medical Institute, Chevy Chase, MD 20815

Contributed by William R. Jacobs Jr., November 12, 2017 (sent for review October 19, 2017; reviewed by Martin Chalfie and Eric J. Rubin)

Engaging undergraduate students in scientific research promises substantial benefits, but it is not accessible to all students and is rarely implemented early in college education, when it will have the greatest impact. An inclusive Research Education Community (iREC) provides a centralized scientific and administrative infrastructure enabling engagement of large numbers of students at different types of institutions. The Science Education Alliance–Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) is an iREC that promotes engagement and continued involvement in science among beginning undergraduate students. The SEA-PHAGES students show strong gains correlated with persistence relative to those in traditional laboratory courses regardless of academic, ethnic, gender, and socioeconomic profiles. This persistent involvement in science is reflected in key measures, including project ownership, scientific community values, science identity, and scientific networking.

bacteriophage | genomics | science education | evolution | assessment

Engaging undergraduates in scientific research is educationally advantageous, regardless of the students' career aspirations (1–3). Several well-established models, each with benefits and challenges (4), provide this engagement. In apprentice-based research experiences (AREs), students, typically in their later college years, perform research under the direct supervision of an experienced mentor. An ARE can provide a high level of training, but the opportunities are constrained by laboratory space and supervisory capacity, imposing high-stakes selection for a relatively small number of students (5). Course-based research experiences (CREs) represent a second model; in this case, students conduct research as a class. In comparison with AREs, well-designed CREs can engage more students earlier in the curriculum (6), which is expected to have higher impact (7, 8). However, developing authentic research activities suitable for a CRE is challenging. A drawback of both models is that they largely exclude the 40% of US undergraduate students who attend 2-y colleges or 4-y colleges with limited research infrastructures (9).

A third model is the inclusive Research Education Community (iREC), in which a common scientific problem is addressed by students at multiple institutions that are supported by a centralized scientific and programmatic structure. Because of the centralized support, the iREC presents three advantages over other models. (*i*) The iREC is inclusive, because it is designed for students with few prerequisites, thus emphasizing the exploration of a student's potential rather than selection based on past accomplishments. (*ii*) The iREC presents students at all types of institutions with the opportunity to participate in authentic research, including at schools with little or no investigator-driven research. (*iii*) The iREC encourages growth, because the programmatic costs per student decrease as more students participate.

The centralized scientific and programmatic structure of the iREC encourages the development of a collaborative community, in which the students interact with one another both within the same institution and across institutions. The sense of community is strengthened in several ways: all of the schools pursue the same scientific problem, instructors from different institutions regularly come together in training workshops and faculty meetings, and students and faculty are presented with opportunities to share their findings with one another [e.g., the Science Education Alliance–Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) annual symposium]. In these ways, the student's cognitive experience mirrors that of an experienced researcher, and the social community aspects of scientific practice are apparent. Because iRECs require robust centralized programmatic structures that support the study of suitable research topics (10), iRECs are rare (5). Examples include the Genomics Education Partnership (11, 12), Small World Initiative (13, 14), and the SEA-PHAGES program (15).

The special characteristics of the iREC make it a particularly strong candidate for enhancing science education early in a student's career, with the long-term outcome of enhancing engagement and student persistence in the sciences. The iREC educational

Significance

The Science Education Alliance–Phage Hunters Advancing Genomics and Evolutionary Science program is an inclusive Research Education Community with centralized programmatic and scientific support, in which broad student engagement in authentic science is linked to increased accessibility to research experiences for students; increased persistence of these students in science, technology, engineering, and mathematics; and increased scientific productivity for students and faculty alike.

Author contributions: D.I.H., M.J.G., S.G.C., R.A.G., D.J.-S., W.H.P., D.A.R., V.S., D.J.A., and G.F.H. designed research; D.I.H., SEA-PHAGES, L.B., A.B., N.K., and W.H.P. performed research; D.I.H., S.-P., L.B., A.B., and N.K. analyzed data; S.G.C., R.A.G., D.J.-S., W.H.P., D.A.R., D.J.A., and G.F.H. performed program development and support; D.I.H., SEA-PHAGES, L.B., and N.K. collected and analyzed data; M.J.G. and A.B. developed the SEA-PHAGES structure model; S.G.C., R.A.G., D.J.-S., W.H.P., D.A.R., V.S., D.J.A., and G.F.H. provided SEA-PHAGES program development and support; and D.I.H., M.J.G., L.B., A.B., S.G.C., R.A.G., D.J.-S., N.K., W.H.P., D.A.R., W.R.J., V.S., D.J.A., and G.F.H. wrote the paper.

Reviewers: M.C., Columbia University; and E.J.R., Harvard School of Public Health.

The authors declare no conflict of interest.

[1]A complete list of SEA-PHAGES authors can be found in the *Supporting Information*.

[2]To whom correspondence may be addressed. Email: jacobsw@hhmi.org, asaid@hhmi.org, or gfh@pitt.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1718188115/-/DCSupplemental.

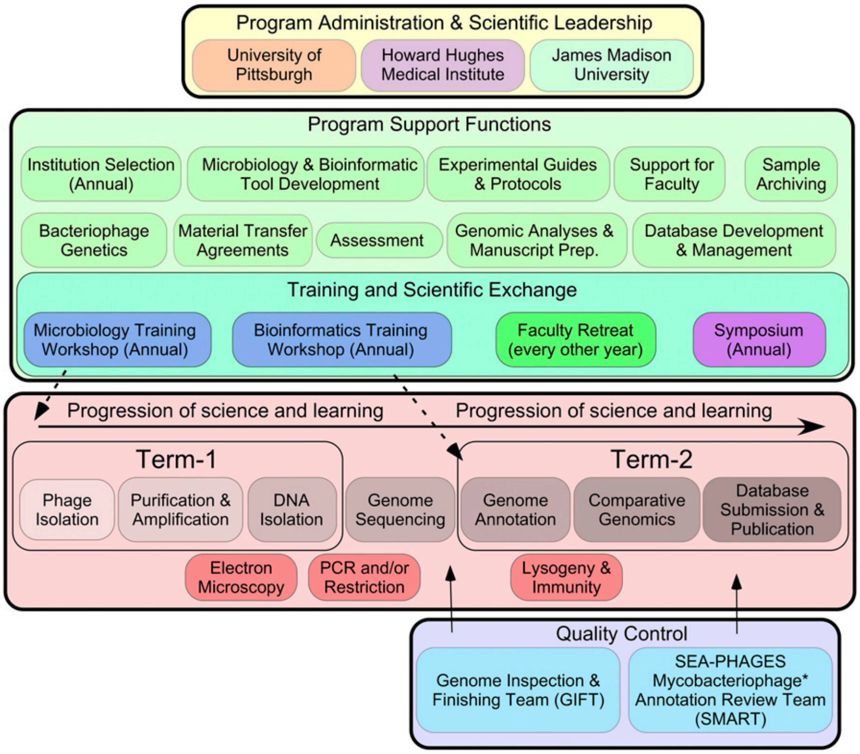

Fig. 1. Organization and structure of the SEA-PHAGES program. The SEA-PHAGES program administrators (yellow box) oversee support components critical to program implementation (green box). Typical two-term course structure (red box) includes phage isolation through comparative genomics; additional characterization includes EM, PCR/restriction analysis, and lysogeny assays (red ovals). Sequence and annotation quality control is shared with SEA-PHAGES faculty teams (purple box).

approach, fully implemented in the SEA-PHAGES program, provides a testing ground to explore the outcomes of this approach in terms of scientific productivity, student engagement, and student persistence in science, technology, engineering, and mathematics (STEM). Here, we report the combined impacts of research productivity and student persistence of the SEA-PHAGES program. The synergy between research authenticity and student engagement suggests that the iREC model could play a transformative role in science education.

Results

SEA-PHAGES Program Infrastructure. The SEA-PHAGES program seeks to understand viral diversity and evolution taught as a two-term laboratory course research experience. The first term is focused on bacteriophage isolation, purification, and DNA purification, and the second term focuses on genome annotation and bioinformatic analyses of the isolated phages (Fig. 1). Because the phage population is vast, dynamic, old, and consequently, enormously diverse (16, 17), the probability that a student will isolate a phage with a new genome or with previously unidentified genes is high (18, 19). When coupled with the technical simplicity of phage isolation, rapid and cheap sequencing capabilities, and powerful bioinformatic tools, SEA-PHAGES presents an accessible and discovery-rich research experience.

Programmatic support and scientific support are critical for success of an iREC. The SEA-PHAGES program elements include the development and publication of detailed experimental protocols, two 1-wk faculty training workshops in (*i*) phage discovery and (*ii*) bioinformatics, curated databases of students' results, archiving of collected bacteriophages, continuous system-wide assessment,

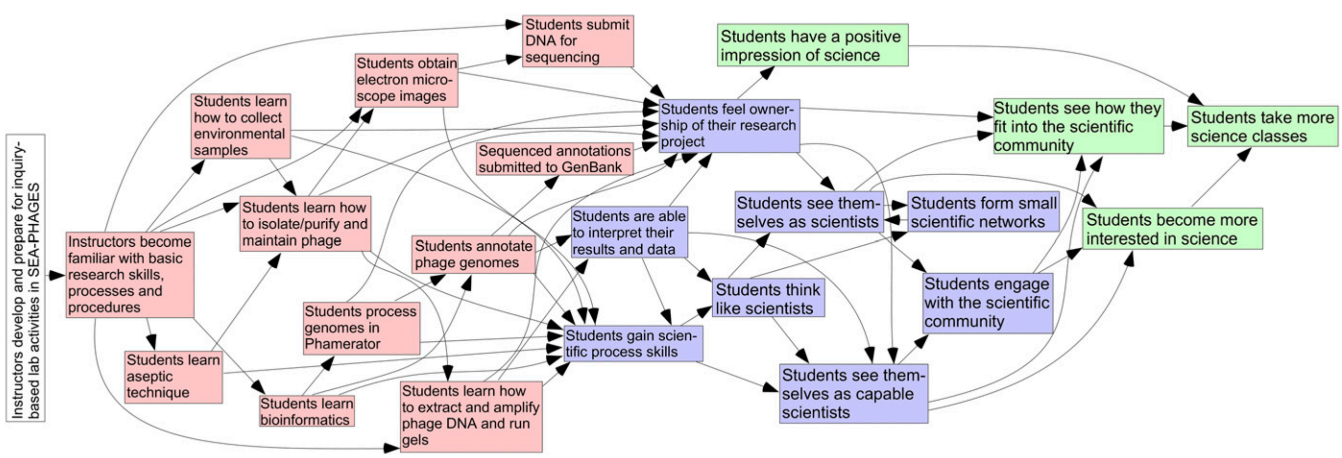

Fig. 2. The SEA-PHAGES systems-level model. Systems-level SEA-PHAGES activities (white box) with short-, medium-, and long-term outcomes (red, blue, and green boxes, respectively). *SI Appendix*, Fig. S1 shows the entire model.

scientific exchange in online forums, and an annual symposium. All of the SEA-PHAGES faculty meet in a biennial faculty retreat, and faculty also participate in advanced genome annotation workshops. In addition, Science Education Alliance faculty teams contribute to quality control of both sequence data and genome annotation (Fig. 1). Two databases facilitate coordination of the scientific and programmatic data (phagesdb.org and https://seaphages.org, respectively).

Because of the potential complexity of SEA-PHAGES, we used systems-level methods (20, 21) to construct a detailed pathway map (Fig. 2 and *SI Appendix*, Fig. S1) that relates program activities to short-, medium- and long-term outcomes in SEA-PHAGES. The full model (*SI Appendix*, Fig. S1) captures all of the program elements and how they connect to outcomes, and a modest subset illustrates the pathways linking course design with student persistence (Fig. 2). This model is helpful for facilitating program development, designing additional iRECs, and providing a framework for assessment strategies.

SEA-PHAGES Program Scale and Costs. The initial investment in iREC administrative and programmatic structure facilitates program growth. The SEA-PHAGES program has grown by addition of 7–25 institutions each year, and over its 9-y development, it now includes over 100 institutions (Fig. 3*A* and *SI Appendix*, Table S1), spanning R1 universities to community colleges (Fig. 3*B* and *SI Appendix*, Table S1). The 104 schools joining in the first 8 y showed a strong propensity to continue for multiple years in the program, and the probabilities for remaining after 3, 4, or 5 y are 97, 89, and 87%, respectively; continuation rates are not significantly different for schools joining in different years. The massively parallel approach enabled inclusion of over 4,000 students in academic year 2016–2017 (16,300 total over 9 y) (Fig. 3*A*), 80% of whom were in their first or second year of study. Although scalability of undergraduate research programs often presents substantial challenges (1), an iREC promotes cost efficiencies, because the program administration expenditures are nearly independent of the number of students involved; thus, as the

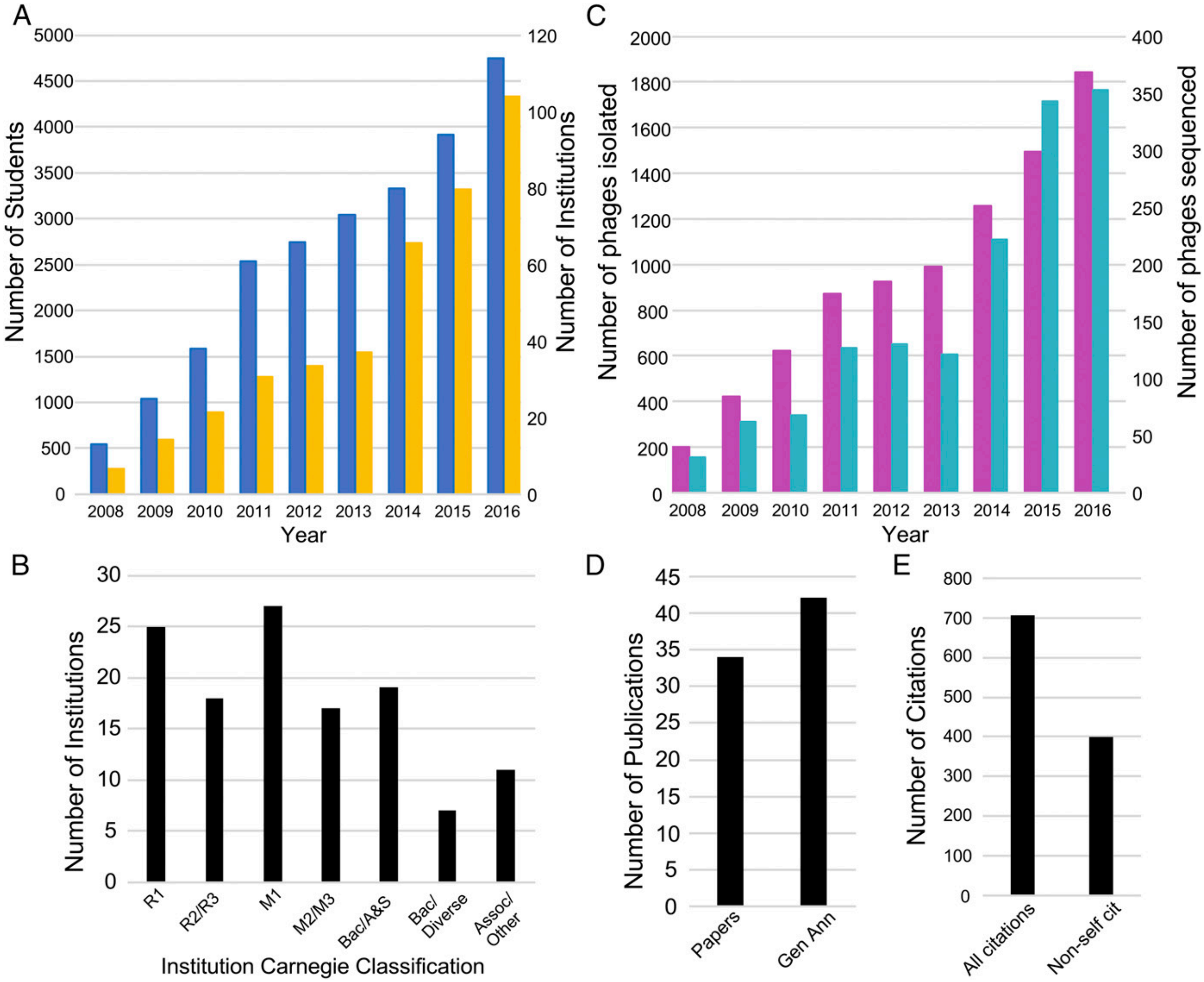

Fig. 3. Program participants and research productivity from the SEA-PHAGES program. (*A*) Numbers of SEA-PHAGES institutions and students (blue and yellow bars, respectively) participating by academic year (fall semester). (*B*) Carnegie Classifications of SEA-PHAGES participating institutions. Assoc/Other, associate's colleges, and others; Bac/A&S, baccalaureate colleges—arts & sciences; Bac/Diverse, baccalaureate colleges—diverse fields; M1–M3, larger, medium, and smaller master's colleges and universities, respectively; R1–R3, doctoral universities with highest, higher, and moderate research activity, respectively. (*C*) Numbers of phages isolated and genomes sequenced (pink and aqua, respectively) by academic year. (*D*) Numbers of peer-reviewed SEA-PHAGES publications as *Genome Announcements* (Gen Ann) and other peer-reviewed papers (Papers) (*SI Appendix*, Table S2). (*E*) Citations of SEA-PHAGES papers, showing all citations and nonself-citations.

number of participating institutions increases, the cost per student decreases. For the SEA-PHAGES program, the current administrative costs per student (∼$500, encompassing all of the support items in Fig. 1) are 33% lower than 2 y previously, and additional program growth will extend the cost-effectiveness. The low per student cost enables the iREC to be delivered to large numbers of students early in their undergraduate careers, thus encouraging students to explore science in a relatively low-risk “gateway” experience. The iREC can introduce the student to research at a better time and at a much lower cost than the more traditional ARE. For those students who find research to be something that they want to explore further, the iREC can provide a stepping stone to subsequent AREs and should facilitate a more productive research experience. We note that the instructional and material costs at SEA-PHAGES participating institutions are greater than for traditional laboratories but are commensurate with other CREs.

SEA-PHAGES Research Productivity. The authenticity of the research conducted in an iREC is critically important, not only for addressing scientific questions but because it also influences the cognitive experiences of student participants (22, 23). In the SEA-PHAGES program, research productivity is reflected in the numbers of phages isolated (∼10,000 in total) (Fig. 3*C*) and sequenced (∼1,400) (Fig. 3*C*), representing substantial proportions of the total numbers of all phages isolated and sequenced to date (24, 25). These findings are reported in over 70 peer-reviewed publications (Fig. 3 *D* and *E* and *SI Appendix*, Table S2) (including 40 short *Genome Announcement* papers), many with student and SEA-PHAGES faculty coauthors. The availability of archived and sequenced phages for experimental manipulation by the scientific community at large provides a valuable resource for gaining insights into bacteriophage biology (24, 25). This research productivity compares favorably with that of one to two NIH R01 grants (26, 27).

Impact of SEA-PHAGES on Student Intention to Persist in STEM. A key iREC educational goal is for students to share the experience of the professional research scientist, including the thrill of discovery, collaboration within a community, and advancing scientific knowledge relevant to the broader community. These psychosocial elements are strongly linked to educational persistence (28–31) and benefit all students, regardless of their intended area of study. Using the psychometric Persistence in the Sciences (PITS) assessment tool (28), we compared 2,850 students taking either SEA-PHAGES or nonresearch traditional laboratory courses at a total of 67 institutions. PITS encompasses five survey components: project ownership (with content and emotion categories), self-efficacy, science identity, scientific community values, and networking, each measuring psychological components that correlate strongly with a student’s intention to continue in science (22, 28). We also collected information on academic performance, socioeconomic status, and other demographics (*SI Appendix*).

To separate the influence of the type of course taken from other variables, including the possibility of student self-selection of

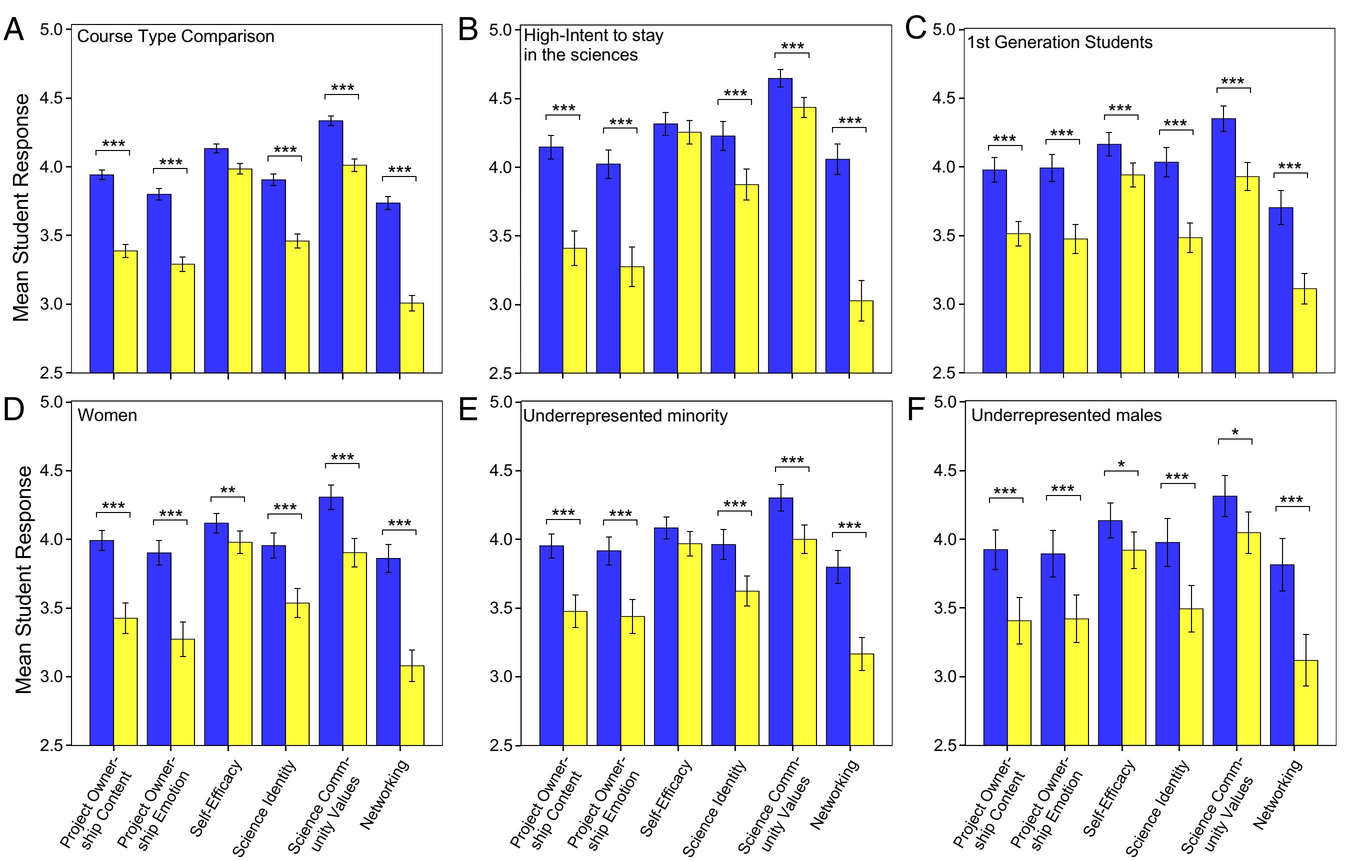

Fig. 4. Comparison of intent to persist in the sciences for students taking SEA-PHAGES and traditional laboratory courses. The PITS survey responses comparing SEA-PHAGES and nonresearch laboratory courses (blue and yellow bars, respectively). (*A*) Propensity score matching balanced all variables, except for course type. (*B*–*F*) Equally sized randomly chosen subsets of students were selected and compared using multivariate ANOVA (MANOVA) (all $P < 0.0001$) and ANOVA, with significant differences indicated. Groups analyzed are those reporting a high (scoring five on a five-point scale) intent to stay in the sciences (*B*), first generation students (*C*), women (*D*), underrepresented minorities (*E*), and underrepresented minority males (*F*). The PITS survey rating scales are from one (strongly disagree) to five (strongly agree) for all measures except for scientific community values, which had a one (not like me at all) to six (very much like me) scale. All scales had full descriptors for each of the levels on the scale. $*P < 0.05$; $**P < 0.01$; $***P < 0.0001$.

SEA-PHAGES or traditional laboratories, we used propensity score matching (32) (Fig. 4*A*). We observed large and significant differences in five of six categories (all except self-efficacy, which assesses students' confidence in their abilities to function as scientists) (Fig. 4*A*), reflecting substantial gains by SEA-PHAGES students. Of the measures used, self-efficacy is the one most closely related to the primary goals of the typical nonresearch traditional laboratory, which are to develop confidence in laboratory procedures and skills. The overall pattern of the PITS measures shows significant increases in multiple aspects of the research experience (project ownership, science identity, science community values, and networking) but little difference in student confidence in laboratory procedures and skills (i.e., self-efficacy). Because the experiments in SEA-PHAGES have greater uncertainty and are directed by the necessities of authentic science, it is reassuring that we did not observe a reduction in self-efficacy compared with traditional laboratories. SEA-PHAGES and traditional laboratories both encourage student development of procedural confidence, but SEA-PHAGES adds an authentic research experience that promotes continued engagement in science.

Because students were not randomly assigned at all 67 institutions, it is plausible that the SEA-PHAGES courses could be disproportionately populated with students interested in pursuing science. To test this, we compared students declaring the highest possible intent to stay in science and observed similarly strong gains by SEA-PHAGES students (Fig. 4*B*). The surprisingly low scores—correlating with poor persistence (28)—from students with high intent to study science who are taking traditional nonresearch laboratory courses resonate with national concerns about science education (9). A simple interpretation is that students keen on pursuing science interests were discouraged by their experiences in traditional laboratory courses.

iREC Inclusion Promotes Broad Student Success. To examine the inclusive nature of the iREC, we compared student cohorts known to have poor science persistence early in college careers (33, 34), particularly first generation college students (Fig. 4*C*), women (Fig. 4*D*), underrepresented minorities (Fig. 4*E*), and underrepresented men (Fig. 4*F*). The broadly shared gains by SEA-PHAGES students strongly support the conclusion that the iREC model provides authentic research experiences (Fig. 4 *C–E*) to all students with similar advantages. We also find that student responses are similar for different types of institutions (Fig. 5*A*)—with small additional project ownership gains at community colleges relative to other schools—and we hypothesize that the supportive iREC programmatic structure (Fig. 1) facilitates success at institutions, such as community colleges, that typically do not have robust investigator-driven research activity. Students with different socioeconomic backgrounds (Fig. 5*B*), academic performance (Fig. 5*C*), gender (Fig. 5*D*), and ethnicity (Fig. 5*E*) also score similarly, reinforcing the inclusive nature of the iREC as exemplified by the SEA-PHAGES program. Finally, to confirm that students reliably self-report their intention to persist in the sciences, we measured the average numbers of science courses taken by subsets of students in each of the three subsequent terms after their introductory laboratory course (Fig. 5*F*). The SEA-PHAGES students enrolled in a consistently higher number of science courses than students taking traditional laboratory courses (Fig. 5*F*).

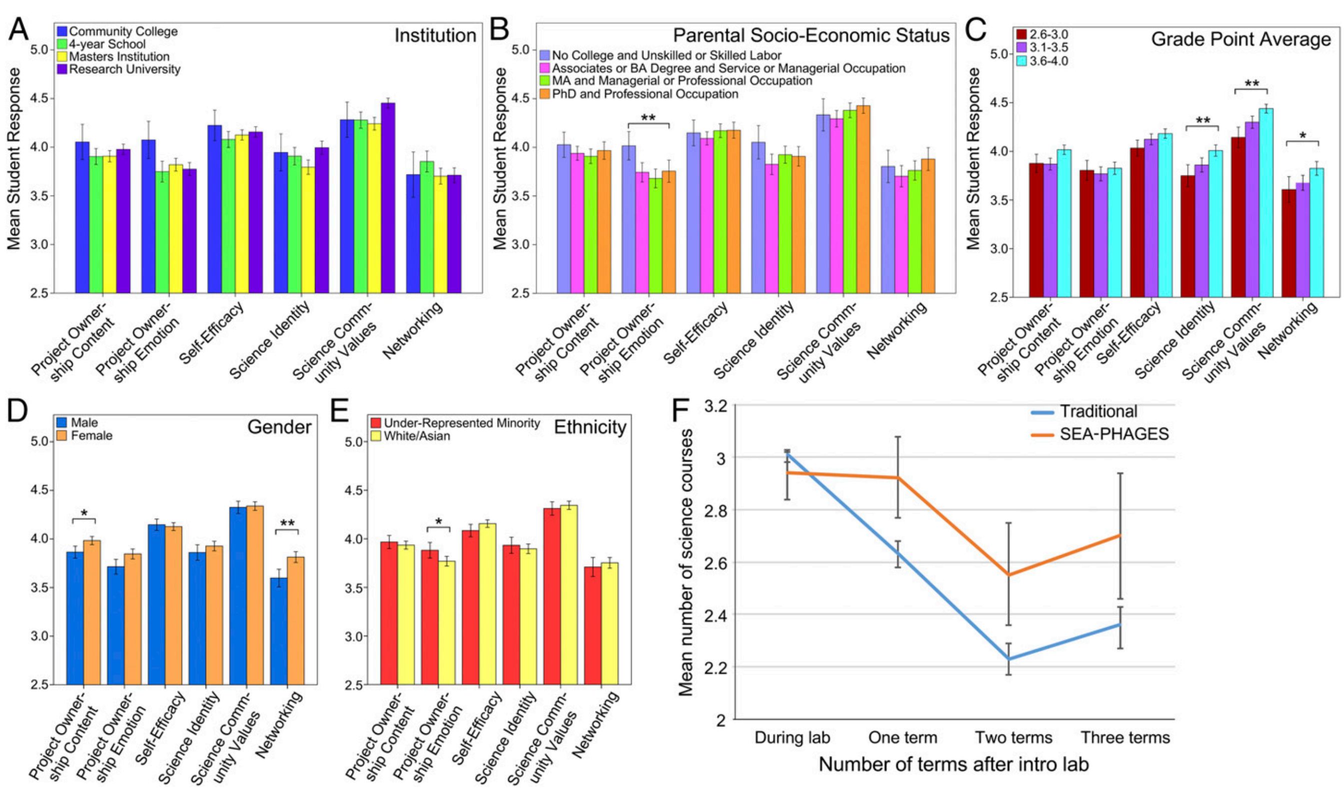

Fig. 5. Comparisons of student subgroups taking the SEA-PHAGES courses on their intent to persist in the sciences. The PITS survey responses for equally sized randomly chosen subsets of students were selected and compared. Groups differed by institutions (*A*), socioeconomic status (*B*), grade point average (*C*), gender (*D*), or ethnicity (*E*). Multivariate ANOVA (MANOVA) showed only small differences for some groups (institution type, $P < 0.049$; grade point average, $P < 0.04$; gender, $P < 0.001$). Significant differences using univariate analyses (ANOVA) are shown. The PITS survey rating scales are from one (strongly disagree) to five (strongly agree) for all measures except for scientific community values, which had a one (not like me at all) to six (very much like me) scale. All scales had full descriptors for each of the levels on the scale. *$P < 0.05$; **$P < 0.01$. (*F*) Average number of science courses taken by students experiencing SEA-PHAGES (red) or a nonresearch laboratory course (blue) in three subsequent terms; 95% confidence intervals are shown.

Discussion

We have described here the iREC model for promoting student persistence in STEM education. The iREC, as illustrated by SEA-PHAGES, focuses on scientific discovery within a community accessible by early career undergraduate students and a centralized administrative structure that supports a broad range of institutions. Furthermore, it enables student development regardless of demographic or academic background. We propose that the iREC concept could have a transformative impact on science education when expanded to include additional research topics. We encourage research institutions to design and implement additional iREC programs. We emphasize that the authenticity of iREC research topics is important, not only for promoting student engagement through project ownership but also for program sustainability and acquiring financial support.

Several important questions arise regarding SEA-PHAGES program implementation and iREC development in general. For example, the SEA-PHAGES program spans experimental approaches, including microbiology, molecular biology, imaging, and computational biology, and the contributions of each of these elements to student persistence are unresolved. Furthermore, as yet, we know little of how the iREC experience influences students' choices in enrolling for other STEM courses and laboratories or in pursuing other research experiences. We also do not know how the SEA-PHAGES experience influences student career choices after graduation. Because early career students succeed in SEA-PHAGES, regardless of background or experience, we predict that the benefit of experiencing the process of discovery—vs. the unfortunately too frequent imposition of exercises for which the "right" answers are already known—will be broadly accrued by all students, including those who sample science via this iREC but who choose to pursue nonscience careers. Layering iREC experiences through different levels of the undergraduate curriculum could multiply their impacts.

Although the initial costs of establishing an iREC administrative structure can be substantial, they can be considerably less so if built on an extant independently funded research program. After it is operational, the program structure can support rapid expansion of the numbers of institutions and student participants, thereby substantially reducing the costs/student. Defining the SEA-PHAGES programmatic structure (Fig. 1), analyzing the relationships among its component elements (Fig. 2), and documenting the research and educational outcomes (Figs. 3–5) provide a path for future iREC development. Widespread use of this model has the potential to drive a major transformation of undergraduate science education.

Materials and Methods

The pathway model was constructed using previously described approaches (20), and detailed methods are described in *SI Appendix*. Program assessment used the PITS survey tool and comprised five existing survey tools covering project ownership, self-efficacy, science identity, scientific community values, and networking, all of which measure different psychological components of a research experience and have individually been used in a range of investigations of educational programs. Before usage in this data collection process, the PITS survey was evaluated for its dimensionality, validity, and internal consistency (28). The tool underwent psychometric evaluation and has been validated for usage in the assessment of research experiences. Details of the survey cohorts, data, and statistical analyses are described in detail in *SI Appendix*. This study was approved and supervised by the Institutional Review Board of the Indiana University of Pennsylvania (14-302) and the University of Pittsburgh Institutional Review Board (PRO14100567 and PRO15030412).

ACKNOWLEDGMENTS. We thank Billy Biederman, Priscilla Kobi, and Crystal Petrone for program assistance and manuscript preparation; Sam Jackendoff for technical expertise and data collection; Tuajuanda Jordan, Lu Barker, Kevin Bradley, and Melvina Lewis for early program development; and SEA-PHAGES students and instructors. We also thank the reviewers for helpful comments on the manuscript. This work was supported by National Science Foundation Grant DUE-1524575 and Howard Hughes Medical Institute Grants 54308198 and 52008197.

1. Gentile J, Brenner K, Stephens A (2017) *Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities* (National Academies, Washington, DC).
2. Lopatto D (2004) Survey of undergraduate research experiences (SURE): First findings. *Cell Biol Educ* 3:270–277.
3. Lopatto D (2007) Undergraduate research experiences support science career decisions and active learning. *CBE Life Sci Educ* 6:297–306.
4. Brewer C, Smith D, eds (2011) *Vision and Change in Undergraduate Biology Education: A Call to Action* (American Association for the Advancement of Science, Washington, DC).
5. Wei CA, Woodin T (2011) Undergraduate research experiences in biology: Alternatives to the apprenticeship model. *CBE Life Sci Educ* 10:123–131.
6. Bangera G, Brownell SE (2014) Course-based undergraduate research experiences can make scientific research more inclusive. *CBE Life Sci Educ* 13:602–606.
7. 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.
8. Linn MC, Palmer E, Baranger A, Gerard E, Stone E (2015) Education. Undergraduate research experiences: Impacts and opportunities. *Science* 347:1261757.
9. 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. Available at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf. Accessed April 7, 2015.
10. Lopatto D, et al. (2014) A central support system can facilitate implementation and sustainability of a classroom-based undergraduate research experience (CURE) in genomics. *CBE Life Sci Educ* 13:711–723.
11. Shaffer CD, 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.
12. Elgin SCR, et al.; Genomics Education Partnership (2017) The GEP: Crowd-sourcing big data analysis with undergraduates. *Trends Genet* 33:81–85.
13. Caruso JP, Israel N, Rowland K, Lovelace MJ, Saunders MJ (2016) Citizen science: The small world initiative improved lecture grades and California critical thinking skills test scores of nonscience major students at Florida Atlantic University. *J Microbiol Biol Educ* 17:156–162.
14. Davis E, et al. (2017) Antibiotic discovery throughout the small world initiative: A molecular strategy to identify biosynthetic gene clusters involved in antagonistic activity. *MicrobiologyOpen* 6.
15. Jordan TC, et al. (2014) A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. *MBio* 5:e01051–e13.
16. Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF (1999) Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc Natl Acad Sci USA* 96:2192–2197.
17. Rohwer F, Youle M, Maughan H, Hisakawa N (2014) *Life in Our Phage World: A Centenial Field Guide to the Earth's Most Diverse Inhabitants* (Wholon, San Diego).
18. Hanauer DI, et al. (2006) Inquiry learning. Teaching scientific inquiry. *Science* 314: 1880–1881.
19. Hatfull GF, et al. (2006) Exploring the mycobacteriophage metaproteome: Phage genomics as an educational platform. *PLoS Genet* 2:e92.
20. Corwin LA, Graham MJ, Dolan EL (2015) Modeling course-based undergraduate research experiences: An agenda for future research and evaluation. *CBE Life Sci Educ* 14:es1.
21. Urban JB, Trochim W (2009) The role of evaluation in research practice integration working toward the "golden spike." *Am J Eval* 30:538–553.
22. Hanauer DI, Hatfull G (2015) Measuring networking as an outcome variable in undergraduate research experiences. *CBE Life Sci Educ* 14:ar38.
23. Brownell SE, et al. (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:ar21.
24. Pope WH, et al.; Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science; Phage Hunters Integrating Research and Education; Mycobacterial Genetics Course (2015) Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *eLife* 4:e06416.
25. Dedrick RM, et al. (2017) Prophage-mediated defence against viral attack and viral counter-defence. *Nat Microbiol* 2:16251.
26. Berg J (2011) Productivity metrics and peer review scores *NIGMS Feedback Loop Blog*. Available at https://loop.nigms.nih.gov/2011/06/productivity-metrics-and-peer-review-scores/. Accessed September 14, 2016.
27. Jacob BA, Lefgren L (2011) The impact of research grant funding on scientific productivity. *J Public Econ* 95:1168–1177.
28. Hanauer DI, Graham MJ, Hatfull GF (2016) A measure of college student persistence in the sciences (PITS). *CBE Life Sci Educ* 15:ar54.
29. Robnett RD, Chemers MM, Zurbriggen EL (2015) Longtidinal associations among undergraduates' research experiences, sefl-efficacy, and identity. *J Res Sci Teach* 52:847–867.
30. Estrada M, Woodcock A, Hernandez PR, Schultz PW (2011) Toward a model of social influence that explains minority student integration into the scientific community. *J Educ Psychol* 103:206–222.
31. Graham MJ, Frederick J, Byars-Winston A, Hunter AB, Handelsman J (2013) Science education. Increasing persistence of college students in STEM. *Science* 341:1455–1456.
32. Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 46:399–424.
33. Asai DJ, Bauerle C (2016) From HHMI: Doubling down on diversity. *CBE Life Sci Educ* 15:fe6.
34. Huang G, Taddese N, Walter E, Peng SS (2000) *Entry and Persistence of Women and Minorities in College Science and Engineering Education* (US Department of Education, National Center for Education Statistics, Washington, DC).