

Innovations in Undergraduate Science Education: Going Viral

Graham F. Hatfull

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania

Bacteriophage discovery and genomics provides a powerful and effective platform for integrating missions in research and education. Implementation of the Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program facilitates a broad impact by including a diverse array of schools, faculty, and students. The program generates new insights into the diversity and evolution of the bacteriophage population and presents a model for introducing first-year undergraduate students to discovery-based research experiences.

In 2012 the President's Council of Advisors on Science and Technology (PCAST) issued their report "Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics," which called for fundamental revisions in undergraduate education, especially in the first 2 years of college or university (1). Noting the economic demand in the United States for additional science, technology, engineering, and mathematics (STEM) graduates over the next decade, they concluded that even a modest improvement in student retention in STEM disciplines would have a substantial impact. They presented five major recommendations, one of which was to replace standard laboratory courses with discovery-based research courses. A second report, "Vision and Change in Undergraduate Biology Education: a Call to Action" from the American Association for the Advancement of Science (AAAS) (2) echoed these sentiments and specifically recommended to "Introduce research experiences as an integral component of biology education for all students, regardless of their major." It is easy to agree with these recommendations, but putting them into action is associated with numerous challenges.

The general idea of engaging large numbers of freshman undergraduate students in authentic research is attractive, as it has the capacity to have a transformative impact on their entire undergraduate careers (3), and early curricular innovations are likely to drive additional changes in upper-division courses. But as attractive as this is, it also exposes the nub of the problem: most authentic research activities require prior knowledge of content, concepts, and techniques not typically characteristic of first-year undergraduate students. This is largely why undergraduate research is typically reserved for apprentice-like experiences in which students work in a research lab under the supervision of a faculty member, graduate student, or postdoctoral researcher. Virology has an important—and possibly unique—role in breaking through this conundrum.

Bacteriophage discovery and genomics presents an attractive solution for introducing novice scientists (e.g., first-year undergraduates) to discovery-based research experiences. Although bacteriophages were discovered 100 years ago (Happy Anniversary!), it is only recently that we have become aware of the general attributes of the bacteriophage population (4). It is vast (the majority of all life forms are phages!), it is highly dynamic (the entire population turns over every few days), it is old (maybe >2 billion years), and not surprisingly, it is highly diverse (5). And yet mapping and defining this diversity requires little prior knowledge or expertise and little more than a scintilla of curiosity. There is no

need for detailed understanding of biological concepts or specific content matter, and the techniques required are relatively straightforward (6). This is not to say that mastery of the context and techniques is not important: it most certainly is. The point is that it is not needed for any student to initiate phage discovery experiments and presents an open door for student participation.

There is nothing new in the basic idea of isolating phages using a specific bacterial host, but rather little can be learned about phage diversity or evolution by only examining phage plaques or virion morphologies by electron microscopy. Understanding diversity and the mechanisms that generate it requires genomic analyses, which is why the combination of phage discovery and genomic analysis is especially important. In 2002, we developed the Phage Hunters Integrating Research and Education (PHIRE) program at the University of Pittsburgh. Within PHIRE we developed the platform of phage discovery and genomics as an introduction to research for undergraduate and high school students (6, 7). In considering this approach as a possible general model, it is helpful to note that there were dual motivations: addressing a core research question and innovating in science education.

The PHIRE program validated the basic platform and set the stage for its broader dissemination, and in 2008 the Howard Hughes Medical Institute (HHMI) established the Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program. There were 12 participating institutions in the first year, and inclusion of additional schools each year has grown the program to 95 active participants in the fall of 2015 (8); in the fall of 2014, over 2,600 undergraduates were involved, 81% of whom were first- or second-year students. Each institution offers a two-term course (~4 h/week), with the first term focusing on microbiology: phage isolation and naming, purification, amplification, electron microscopy, and DNA isolation. The genome of at least one phage per student section is sequenced between the two terms, assessed for quality, and returned to schools for second-term bioinformatics analysis: i.e., genome an-

Accepted manuscript posted online 27 May 2015

Citation Hatfull GF. 2015. Innovations in undergraduate science education: going viral. *J Virol* 89:8111–8113. doi:10.1128/JVI.03003-14.

Editor: F. Goodrum

Address correspondence to gfh@pitt.edu.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JVI.03003-14

notation and comparative genomics. The SEA-PHAGES program has focused on the use of a single host for phage isolation, *Mycobacterium smegmatis* mc²155, and the genetic diversity is sufficiently high that the exact same phage is rarely independently isolated twice (there are only two exceptions among 850 sequenced mycobacteriophages).

The SEA-PHAGES program has had positive impacts on both research and education. The total number of mycobacteriophages isolated is nearly 6,000, and complete genome sequences have been determined for about 850 (9; <http://phagesdb.org>). This is by far the largest collection of sequenced phage genomes for any single bacterial host strain and represents about a third of all sequenced double-stranded DNA phage genomes. Even though all these are in principle in direct genetic contact with each other, the population is highly diverse, with many different types of genomes, albeit with mosaic architectures (9, 10). This wealth of new genetic information has contributed to numerous peer-reviewed publications, underscoring the authenticity of the research advances. Because phage genomes are relatively small and cheap to sequence, the availability of phage isolates becomes a limiting factor, and the SEA-PHAGES program resolves this bottleneck (9). The program also promotes advances in student learning gains and retention in STEM disciplines (8), as well as heightened engagement as seen through networking analyses (20). The breadth of the impact is reflected in the capacity to reach several thousand students, and video testimonials reflect student excitement at participation (<http://seaphages.org>).

The phage discovery and genomics platform is complementary to viral metagenomics (11, 12) as an approach to understanding viral diversity and evolution. Metagenomics yields vast amounts of data but relatively few complete genome sequences. An advantage of the phage discovery and genomics platform is that it populates not only the computer but also the freezer, and individual phages can be genetically, biochemically, and structurally characterized (13). The entire collection thus becomes available for functional analysis of the thousands of novel genes identified or manipulated for use in phage-based diagnostics or phage therapy (14, 15).

The various successes of the phage discovery and genomics platform warrant consideration of the features that contribute toward its implementation, as these may be useful in guiding development of other discovery-based research courses for first-year undergraduates (Table 1). We will briefly discuss 11 such features.

First, there is a single broad research question in defining the genetic diversity of the virosphere, providing coherence to the various individual experiments. Second, there is substantial flexibility in how it can be implemented. We have discussed above its use in the research laboratory (PHIRE) and as a freshman course (SEA-PHAGES), but it has also been successfully implemented at the University of KwaZulu-Natal for the past 7 years as an intensive 2-week workshop (9). Third, it is imbued with discovery both in the isolation of novel phages and in the identification of new genes, and fourth, it is authentic research, as indicated by the associated peer-reviewed publications. Fifth, the platform uses a parallel project structure with a common progression of experimental approaches, simplifying instruction and providing peer-mentoring opportunities. In the SEA-PHAGES program, student alumni are well-prepared and often keen to act as undergraduate teaching assistants for subsequent classes. Individuality arises from the personalized student experiences in isolating and naming their own phages, key aspects of the sixth feature, project own-

TABLE 1 Features of the phage discovery and genomics platform

Feature no.	Feature	Comments
1	Single research question	Mapping the genetic diversity of the virosphere
2	Flexibility	Implemented as research lab, course, and workshop configurations
3	Discovery based	Discovery of new phages and new genes
4	Authentic research	New findings leading to peer-reviewed publications
5	Parallel projects	Common research progression; facilitating peer mentoring
6	Project ownership	Phage isolation and naming, linking virus and student identities
7	Promotes diversity	No requirement for prior content, concept, or technical knowledge
8	Progression	Structured introduction of complexity throughout program
9	Training	Workshops can train nonexpert instructors
10	Program coordination	Databases for data coordination, program implementation; podcast, newsletter, annual symposium
11	Central administration	Coordinated program and scientific leadership promote institutional diversity

ership, which plays an important role in engaging students (16). This is reflected in the keenness of students to discuss their phage research with other students and faculty, as well as their parents and guardians (20).

The seventh feature is that an early-career course-based research experience provides an opportunity for students to recognize an aptitude for research that may go unrealized if students have to compete for access to an apprentice-like research experience that depends on prior accomplishments (such as lecture course performance) as a qualification. Furthermore, the platform is suitable for both STEM and non-STEM majors (17). This feature is thus of particular importance, as it represents an inclusive approach to research opportunities, with a potentially broad impact across undergraduate student communities. The eighth feature is that the platform progresses from simple, concrete approaches at the beginning to increasingly abstract concepts in genome analysis. In essence, it represents a guided tour up Bloom's taxonomy (18).

The final three features concern organizational and administrative aspects (Table 1). An unusual facet of SEA-PHAGES is that the vast majority of faculty instructors do not have prior bacteriophage experience. Faculty preparation is facilitated by two 1-week workshops for the microbiology and bioinformatics parts, respectively. Database components are also critical for coordination of both data and program administration. Students enter all of their phage information into the phagesDB database at <http://phagesdb.org>, from which students can access the collective data. Faculty enter information about their schools and students into the database at <http://seaphages.org>, which collates these data. There is also ongoing development of computational tools for genome annotation and comparative genomics (19) and community-promoting components such as an annual symposium, podcast, and

newsletter. Finally, disseminated programs such as SEA-PHAGES benefit from centralized leadership provided by faculty at the University of Pittsburgh, James Madison University, and the Howard Hughes Medical Institute. This provides scientific guidance, selection of participating institutions, faculty training, and both program and student assessment. These centralized components facilitate inclusion of institutions without robust research infrastructures, and only one-third of SEA-PHAGES schools are in the Carnegie Classification as Research Universities. SEA-PHAGES thus provides a model for bringing authentic research experiences to students and faculty at institutions such as primarily undergraduate institutions and community colleges where such opportunities are relatively rare.

What are the costs associated with such integrated research and educational programs? These are complex but fall into two main categories: costs to participating institutions and costs for program administration. Institutional costs include those for faculty instruction and for research materials and supplies. Faculty costs vary by institution, and supplies are \$100–\$200/student, approximately equivalent to the costs of a textbook. Program administration costs are considerable, but can be considered an offset against research productivity. Total costs per student are but a fraction (<20%) of the costs for undergraduate students doing apprentice-like summer internships, but with similar benefits (8).

Although the number of students potentially benefitting from participation in a phage discovery and genomics experience is large, it is dwarfed by the estimated total of 10^{31} phage particles in the biosphere (5). And with such impressive diversity, we'll run out of students long before we run out of novel phages to isolate! The SEA-PHAGES program is in the process of expanding to include other host strains within the phylum *Actinobacteria*, and the number of available strains to use is large. There is clearly a great deal of juice to be extracted from this approach, promising many new insights into viral biology.

Does the general discipline of virology have other integrated research and education platforms to contribute? Almost certainly so, and the field of virology in general has a long and strong history of educational innovations. The challenge then is to develop specific initiatives that can be broadly implemented early in the undergraduate curriculum. We hope the varied features of the phage discovery and genomics platform will help to guide and inspire such innovations.

ACKNOWLEDGMENTS

I thank David Asai, Deborah Jacobs-Sera, Welkin Pope, and Daniel Russell for comments on the manuscript.

REFERENCES

1. 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. Report to the president. Executive Office of the President, Washington, DC. <http://www.whitehouse.gov/ostp/pcast>.
2. Brewer C, Smith D (ed). 2011. Vision and change in undergraduate biology education: a call to action. American Association for the Advancement of Science, Washington, DC.
3. Harrison M, Dunbar D, Ratmanský L, Boyd K, Lopatto D. 2011. Classroom-based science research at the introductory level: changes in career choices and attitude. *CBE Life Sci Educ* 10:279–286. <http://dx.doi.org/10.1187/cbe.10-12-0151>.
4. Hatfull GF, Hendrix RW. 2011. Bacteriophages and their genomes. *Curr Opin Virol* 1:298–303. <http://dx.doi.org/10.1016/j.coviro.2011.06.009>.
5. Hendrix RW. 2002. Bacteriophages: evolution of the majority. *Theor Popul Biol* 61:471–480. <http://dx.doi.org/10.1006/tpbi.2002.1590>.
6. Hanauer DI, Jacobs-Sera D, Pedulla ML, Cresawn SG, Hendrix RW, Hatfull GF. 2006. Inquiry learning. Teaching scientific inquiry. *Science* 314:1880–1881. <http://dx.doi.org/10.1126/science.1136796>.
7. Hatfull GF, Pedulla ML, Jacobs-Sera D, Cichon PM, Foley A, Ford ME, Gonda RM, Houtz JM, Hryckowian AJ, Kelchner VA, Namburi S, Pajcini KV, Popovich MG, Schleicher DT, Simanek BZ, Smith AL, Zdanowicz GM, Kumar V, Peebles CL, Jacobs WR, Jr, Lawrence JG, Hendrix RW. 2006. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet* 2:e92. <http://dx.doi.org/10.1371/journal.pgen.0020092>.
8. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SC, Findley AM, Gissendanner CR, Golebiewska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cresawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. *mBio* 5(1):e01051–13. <http://dx.doi.org/10.1128/mBio.01051-13>.
9. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Hendrix RW, Lawrence JG, Hatfull GF, 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. <http://dx.doi.org/10.7554/eLife.06416>.
10. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S, Kriakov J, Lawrence JG, Jacobs WR, Hendrix RW, Hatfull GF. 2003. Origins of highly mosaic mycobacteriophage genomes. *Cell* 113:171–182. [http://dx.doi.org/10.1016/S0092-8674\(03\)00233-2](http://dx.doi.org/10.1016/S0092-8674(03)00233-2).
11. Deng L, Ignacio-Espinoza JC, Gregory AC, Poulos BT, Weitz JS, Hugenholtz P, Sullivan MB. 2014. Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature* 513:242–245. <http://dx.doi.org/10.1038/nature13459>.
12. Edwards RA, Rohwer F. 2005. Viral metagenomics. *Nat Rev Microbiol* 3:504–510. <http://dx.doi.org/10.1038/nrmicro1163>.
13. Marinelli LJ, Piuri M, Swigonova Z, Balachandran A, Oldfield LM, van Kessel JC, Hatfull GF. 2008. BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One* 3:e3957. <http://dx.doi.org/10.1371/journal.pone.0003957>.
14. Hatfull GF. 2014. Mycobacteriophages: windows into tuberculosis. *PLoS Pathog* 10:e1003953. <http://dx.doi.org/10.1371/journal.ppat.1003953>.
15. Jacobs WR, Jr, Barletta RG, Udani R, Chan J, Kalkut G, Sosne G, Kieser T, Sarkis GJ, Hatfull GF, Bloom BR. 1993. Rapid assessment of drug susceptibilities of *Mycobacterium tuberculosis* by means of luciferase reporter phages. *Science* 260:819–822. <http://dx.doi.org/10.1126/science.8484123>.
16. Hanauer DI, Dolan EL. 2014. The Project Ownership Survey: Measuring differences in scientific inquiry experiences. *CBE Life Sci Educ* 13:149–158. <http://dx.doi.org/10.1187/cbe.13-06-0123>.
17. Caruso SM, Sandoz J, Kelsey J. 2009. Non-STEM undergraduates become enthusiastic phage-hunters. *CBE Life Sci Educ* 8:278–282. <http://dx.doi.org/10.1187/cbe.09-07-0052>.
18. Bloom BS, Englehart MB, Furst EJ, Hill WH, Krathwohl DR. 1956. Taxonomy of educational objectives: the classification of educational goals. McKay, New York, NY.
19. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12:395. <http://dx.doi.org/10.1186/1471-2105-12-395>.
20. Hanauer DI, Hatfull GF. Measuring networking as an outcome variable in undergraduate research experiences. *CBE Life Sci Educ*, in press.