Understanding genetic diversity of bacteriophages

29 April 2015

Over the last seven years, thousands of undergraduate students have joined the effort to sequence and analyze the genomes of bacteria-infecting viruses as part of the Howard Hughes Medical Institute's Science Education Alliance (SEA). The hundreds of genomes that they have sequenced are the foundation of a new study that broadens scientists' understanding of the vast diversity among these viruses, which are known as bacteriophages.

The study compared the genomes of 627 bacteriophages isolated from a single species of bacteria, and found a continuum of genetic diversity, rather than discrete groups within the population. The paper, published April 28, 2015, in the journal eLife, is authored by 199 faculty and 2,664 students at institutions across the United States and at the University of KwaZulu-Natal in South Africa.

"Getting these data and looking at the phage population with this degree of resolution is a real payoff of having [the SEA] community," said Graham Hatfull, senior author of the eLife paper and an HHMI Professor at the University of Pittsburgh.

Hatfull leads the SEA's Phage Hunters Advancing Genomi… nce (PHAGES) project in conjunction with HHMI's Undergraduate and Graduate Programs group. Students – primarily first-year undergraduates – in the year-long course, which is now being offered at more than 70 campuses, isolate bacteriophages from local soil, extract DNA from their phages, then sequence and analyze the phage genomes. Their data is shared with the research community through online databases.

Last year, the SEA reported on the educational benefits of the SEA-PHAGES program, demonstrating that the course engages students and motivates them to continue learning science. Equally important, Hatfull says, is that the program's community of student researchers is generating new data and making meaningful contributions to scientific knowledge.

Bacteriophages (also called simply phages) flourish in just about every environment on Earth. Hatfull says there are more phage particles on our planet than all other forms of life combined. But despite their abundance, the genetic diversity of bacteriophages has been poorly characterized, and the function of most phage genes remains unknown. This is largely because only a relatively small number of phages have been isolated, which means there are few sequenced genomes available for comparative analyses.

According to Hatfull, a big question has been whether the phages that have so far been found to infect a single type of bacteria represent distinct populations. Previous genetic analyses had sorted Mycobacteria-infecting phages into clusters, with members of each cluster sharing certain genetic features. But these analyses included too few
seven years into the SEA-PHAGES program, students at 81 institutions had sequenced and analyzed the genomes of hundreds of bacteriophages isolated using Mycobacterium smegmatis from soil samples collected around the world. That was enough data to offer some new insight into phage diversity, Hatfull says.

When Hatfull and his colleagues compared the 627 genomes, they found that they could group them into 28 clusters, with some genes or groups of genes shared between clusters. "The phages are related to each other in complex ways, such that their relationships are better explained with a continuum of genetic diversity," Hatfull says.

Still, certain types of phages were more strongly represented than others among the genomes in their analysis. Including hundreds of genomes gave the researchers the chance to classify rarer phage types that might have been overlooked in a smaller study. "We couldn't have gotten [this new] perspective without getting the data the way we did, as a collective consortium," Hatfull says.

Hatfull says he expects the SEA-PHAGE community can learn more by continuing to isolate and analyze additional bacteriophages of M. smegmatis, but in coming years they will also begin to explore the genetic diversity of viruses that infect other bacterial hosts. He and his colleagues have a spate of new questions they are eager to address, and Hatfull is confident that students in the program will continue to contribute to real research advances.


"A Broadly Implementable Research Course in Phage Discovery and Genomics for First-Year Undergraduate Students." DOI: 10.1128/mBio.01051-13

Provided by Howard Hughes Medical Institute