How diversity-generating retroelements promote mutation and adaptation in myriad microbes

Carrie Arnold, *Science Writer*

Adaptation, a cornerstone of evolutionary change, is rarely straightforward. Acquiring a random mutation that promotes survival can take generations. Prokaryotes such as bacteria and Archaea, along with the viruses they harbor, have compact genomes, leaving them with a limited repertoire of DNA to respond to environmental change.

Fifteen years ago, microbiologist Jeffrey Miller at the University of California, Los Angeles and colleagues identified a type of jumping gene known as a retrotransposon in a handful of bacteria and viruses that allowed them to mutate certain surface proteins by using an enzyme called a reverse transcriptase (1). This enzyme can make DNA using a strand of RNA as a template; but unlike the reverse transcriptase found in retroviruses such as HIV, the enzyme, known as a diversity-generating retroelement (DGR), was error prone.

At certain locations in the RNA sequence, the reverse transcriptase inserted random DNA nucleotides. The result was a protein that differed ever so slightly from its progenitor. The process could be repeated over and over, even within the same cell, giving microbes an almost unlimited array of variations. "It's a system that's beautifully constructed to be constantly evolving in a random but targeted fashion," Miller says. "The cell can constantly be optimizing its surface proteins."

For well over a decade, scientists thought that DGRs only cropped up occasionally on the prokaryotic tree of life. A recent study in *Nature Microbiology* reveals that, far from being a rarity, DGRs are commonplace, affecting both bacteria and their Archaeal cousins (2). This Swiss Army Knife of proteins first elucidated by Miller now appears to provide an important and widely coopted means of survival in an ever-changing world. Investigating DGRs

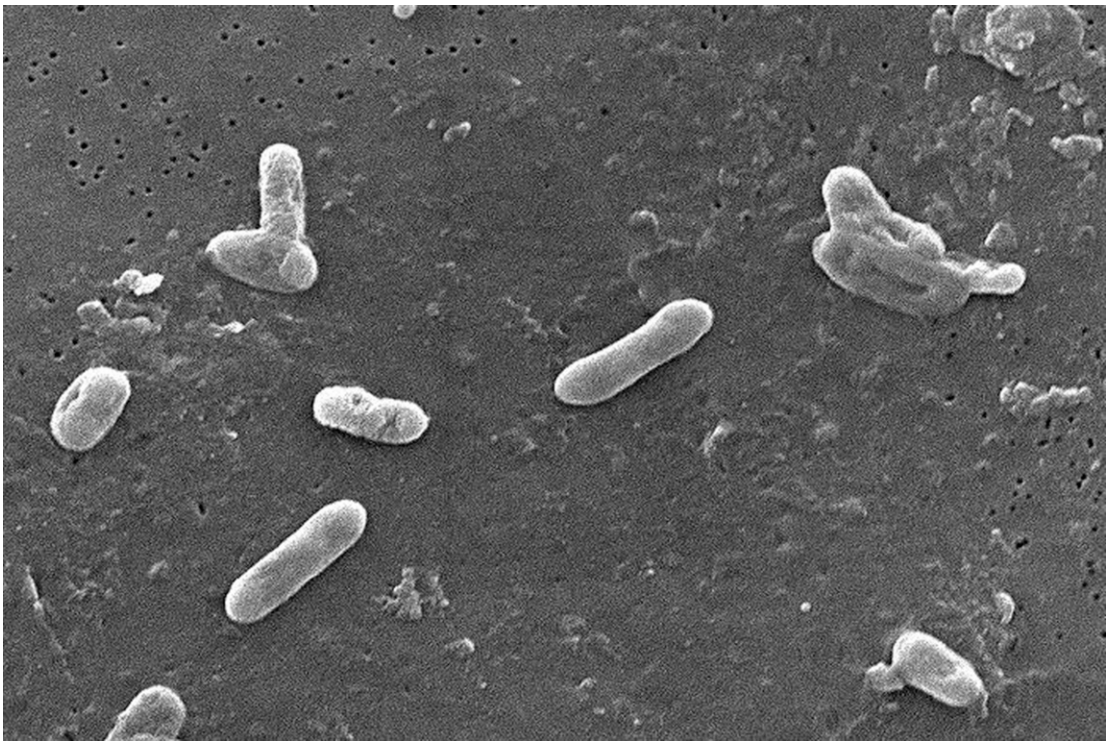

Fig. 1. After infecting the bacterium *Bordetella bronchiseptica* (shown here) with a bacteriophage, researchers stumbled on findings that led them to diversity-generating retroelements, part of a potentially important microbial mechanism for generating a diverse range of proteins. Image courtesy of CDC/Janice Carr.

could lead not only to a better understanding of an important microbial mechanism but also to new strategies for generating a diverse range of proteins and perhaps even the development of more resilient antibiotics.

One in a Million

In 2002, Miller's graduate student Minghsun Liu wanted to find a new bacteriophage virus for a bacterium called *Bordetella bronchiseptica*, closely related to the organism that causes pertussis. *B. brochiseptica* exists in two forms: an infectious form that lives in the respiratory tract, and a free-living form that is suited to life in the droplets of coughs and sneezes. Making a transition between the two forms is as simple as switching different genes on and off.

One of these phages Liu found, BPP-1, only infects the infectious form of *B. bronchiseptica* that invades the respiratory tract. When Liu grew the free-living bacteria as a control and infected them with BPP-1, he expected to see a lawn of healthy, intact bacteria on the Petri dish. Mostly, that's what he found. But on each plate, Liu also found tiny plaques where the virus had infected and killed the bacteria. It was the equivalent of a "one in a million finding," according to Miller.

Viral replication, however, is tremendously error prone, and random mutants crop up all the time. So initially, this discovery didn't seem especially groundbreaking.

> "As he spoke, our jaws just dropped. I had an immediate interest in this system."
>
> —Partho Ghosh

But when Liu and then-undergraduate research assistant Sergei Doulatov, who recently started his own lab at the University of Washington, sequenced the genomes of the various bacteriophages they had isolated in the lab, they found their mutants were far from random.

The mutations were confined to specific viral proteins that hacked surface receptors on the bacteria to inject their DNA. Liu and Doulatov also found that the viral genome contained a reverse transcriptase. "No one had ever seen a reverse transcriptase in a bacteriophage before," Doulatov says. "We knew something weird was going on."

A 2002 *Science* paper outlined their discovery (1). Two additional years of mechanistic experiments worked out the basic mechanism of DGRs, ultimately published in *Nature* in 2004 (3). Proteins targeted by DGRs contain a variable region within the gene that encodes them. A template region sits next to the end of the gene. The reverse transcriptase discovered by Liu and Doulatov copies the mRNA from the template region into DNA, but instead of inserting a thymine to pair with an adenine, the enzyme uses whatever nucleotide happens to be at hand. This error-laden region then replaces the original variable region within the gene. At first, Doulatov thought DGRs only occurred in this group of bacteriophage, but a search of new gene-sequence databases revealed that this strategy was sparsely but widely scattered throughout the bacterial tree of life.

The work was groundbreaking but largely ignored by the larger scientific community. "It was so far out there, no one knew what to do with it," Miller says. "Now, 15 years later, we're finally seeing other groups become interested."

After hearing Miller speak at a Gordon conference shortly after the publication of the *Nature* paper, Partho Ghosh, a biochemist at the University of California, San Diego was one of the only people who immediately saw the promise of DGRs. "As he spoke, our jaws just dropped," Ghosh says. "I had an immediate interest in this system."

Ghosh came aboard and immediately began collaborating with Miller in the early 2000s to solve the structure of the reverse transcriptase and figure out why certain proteins could diversify using DGRs. These proteins all sat on the surface of bacteria or viruses, but Ghosh found they also shared a unique biochemical feature called a C-type lectin fold. Although most well-known for their role in antibodies, C-type lectin folds are large, carbohydrate-binding proteins that Ghosh, Miller, and colleagues found could act as a stable scaffold for DGR tinkering (4).

"This system allows microbes to adapt to a dynamic environment," Ghosh says. He likens it to the human immune system, which can generate countless varieties of antibodies to help us fight off almost any kind of pathogen we might encounter.

DGRs, however, function more broadly than as a sort of immune system. They give prokaryotes and viruses a way to respond to changes in food, predators, and environmental conditions without needing to wait around for a beneficial mutation to crop up. Ghosh's lab is close to figuring out why the DGR reverse transcriptase only mutates DNA at adenines, having submitted his findings for publication—understanding this adenine preference could potentially suggest ways to alter the reverse transcriptase, facilitating alterations to other nucleotides.

As Miller's lab continued its work on DGRs, identifying them in *Legionella* and other species of bacteria, 160 kilometers away, researchers in the lab of earth science and biology professor David Valentine at the University of California, Santa Barbara were making some DGR discoveries of their own (5).

Diverse Elements

One kilometer beneath the azure waves of the Santa Monica Bay, Valentine's postdoctoral researcher Blair Paul guided a submersible called Alvin as it retrieved seafloor samples. Paul hoped the sediment collected contained novel varieties of methane-metabolizing microbes that would provide new insights into Earth's carbon cycle. When he analyzed the DNA sequences of the prokaryotes living in the deep muck, he found his sought-after methanogens, along with other types of Archaea representing a newly identified group called DPANN.

But looking more closely at these sequences Paul found several species with a reverse transcriptase. A more thorough analysis in 2015 showed that DGRs existed in Archaea and were especially prevalent in the Nanoarchaeota, known for their tiny genomes (6). A collaboration with microbiologist Jill Banfield at the University of California, Berkeley enabled Paul and Valentine to search through Banfield's vast collection of microbial genomes. This most recent study, published in *Nature Microbiology* in April 2017, revealed they could be found throughout the microbial world (2). "It's far more widespread that we ever thought," Paul says.

Even as researchers learn more about DGRs' prevalence, the findings thus far haven't helped explain their origin or evolution. Yuzhen Ye, a computational biologist at Indiana University, says that sequence data indicates DGRs only evolved once, most likely from a type of retroelement called Group II introns (7). The reverse transcriptase enzyme found in DGRs also appears to have evolved independently from the reverse transcriptase found in retroviruses. "Even though the genes have different sequences, you can still recognize that it's still the same gene," she says. "It evolved only once. It's an ancient invention, but no one has any idea how far back it goes."

Just as innovative scientists transformed CRISPR from an obscure bacterial immune system into a popular gene-editing technology, Miller hopes that he can do something similar for DGRs. His biotech startup AvidBiotics wants to develop precision antibiotics by inserting DGRs into bacteriophage that target only the disease-causing bacteria. Unlike a CRISPR-generated phage, these DGR-carrying viruses would be able to adapt to the continually evolving genomes of the bacteria they're infecting. Scientists can also use DGRs as a tool to generate a diverse range of proteins to test different drugs and understand how enzymes function by altering their shape.

Just as DGRs can generate a near-endless range of protein variants, Miller thinks that their biotech application might also have a multitude of uses. "Who knows," he asks, "what some creative scientists will think up?"

1 Liu M, et al. (2002) Reverse transcriptase-mediated tropism switching in Bordetella bacteriophage. *Science* 295:2091–2094.

2 Paul BG, et al. (2017) Retroelement-guided protein diversification abounds in vast lineages of Bacteria and Archaea. *Nat Microbiol* 2:17045–17051.

3 Doulatov S, et al. (2004) Tropism switching in Bordetella bacteriophage defines a family of diversity-generating retroelements. *Nature* 431:476–481.

4 McMahon SA, et al. (2005) The C-type lectin fold as an evolutionary solution for massive sequence variation. *Nat Struct Mol Biol* 12:886–892.

5 Arambula D, et al. (2013) Surface display of a massively variable lipoprotein by a Legionella diversity-generating retroelement. *Proc Natl Acad Sci USA* 110:8212–8217.

6 Paul BG, et al. (2015) Targeted diversity generation by intraterrestrial archaea and archaeal viruses. *Nat Commun* 6:6585–6592.

7 Guo H, Arambula D, Ghosh P, Miller JF (2014) Diversity-generating retroelements in phage and bacterial genomes. *Microbiol Spectr* 2:1–16.