Did DNA Come From Viruses?

Research that began with a study of replication enzymes used by bacteria has led to a controversial theory: Viruses may have helped shape all three major domains of life

Scientists who deal in the history of life have never been quite sure what to do with viruses. One measure of their uncertainty is the Tree of Life Web Project, a collective effort to record everything known about the relationships of living and extinct species. The first page of its website—entitled “Life on Earth”—shows the broadest view: From a single root come three branches representing the domains of life (www.tolweb.org). One limb, Eubacteria, includes bacteria such as *Escherichia coli*. Another, Archaea, includes microbes of a different lineage that are less familiar but no less common. The third, Eukaryotes, includes protozoans as well as multicellular organisms such as ourselves. And just below the tree there’s a fourth branch floating off on its own, joined only to a question mark. It is labeled “Viruses.”

A growing number of scientists hope to get rid of that question mark. They recognize that a full account of the evolution of life must include viruses. Not only are they unimaginably abundant—most of the biomass in the ocean is made up of viruses—but they are also extraordinarily diverse genetically, in part because they can acquire genes from their hosts. They can later paste these genes into new hosts, potentially steering their hosts onto new evolutionary paths.

Patrick Forterre, an evolutionary biologist at the University of Paris-Sud in Orsay, France, believes that viruses are at the very heart of evolution. Viruses, Forterre argues, bequeathed DNA to all living things. Trace the ancestry of your genes back far enough, in other words, and you bump into a virus.

Other experts on the early evolution of life see Forterre’s theory as bold and significant. But although they find much to agree with—particularly the importance of viruses to evolution—many also regard Forterre’s ideas as controversial.

“I really applaud the bravery and intellectual power to come up with this picture,” says Eugene Koonin of the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. “But it would be strange if we agreed on all the parts of the picture.”

A new domain

Forterre has been developing and elaborating his theory over many years. He began his scientific career in the early 1970s studying the replication of DNA. He investigated how *E. coli*, the common gut bacteria, use special enzymes to make new copies of their genes without letting the double helix of DNA become tangled.

Exploring the viral world. Patrick Forterre searches for clues to the early evolution of life by investigating exotic viruses.

As Forterre was studying bacteria, another group was developing a complex new view of simple organisms. Carl Woese of the University of Illinois, Urbana-Champaign, demonstrated that some bacteria are not bacteria at all. They belong to a separate branch on the tree of life, which came to be known as Archaea. Archaeans turned out to have a distinct biology. Forterre and his colleagues discovered that they use peculiar enzymes for DNA replication that work differently from those in bacteria and eukaryotes. Meanwhile, other scientists had begun looking at the DNA-replication enzymes used by a virus called T4 bacteriophage, which copy DNA in yet another way.

This discovery made a deep impression on Forterre. At the time, many scientists thought that viruses were merely escaped genetic fragments. After all, viruses are not truly alive: They cannot replicate on their own and have no metabolism. They simply hijack host cells to make new copies of themselves. The best explanation seemed to be that viruses evolved from genes in “true” organisms. Mutations allowed these renegade genes to leave their genomes and become encased in protective protein shells. If that were true, however, the enzymes of viruses should resemble those of their hosts. Yet the DNA-replicating enzymes of T4 bear no relation to the enzymes in the bacteria they infect.

“I thought, maybe T4 is from a fourth domain,” Forterre recalls. But T4 is not unique. Scientists continued to find more DNA-copying enzymes in viruses that have no counterparts in the world of cells, and Forterre’s suspicions deepened. “Maybe each of these viruses is a remnant of a domain that has disappeared,” he wondered.

In 1985, Forterre offered his speculations at a scientific conference. “I suggested that viruses originated from an early cell, perhaps before the origin of the three domains,” he says.

Forterre based his theories on the enzymes in viruses rather than the genes that encode them: genome information was scarce then. And for a long time, the question of viruses’ origin lay dormant. But advances in DNA sequencing have brought a wealth of new information about virus genomes. In Forterre’s view, it reinforces his argument that many viruses are ancient. The genes for the most common proteins in virus shells, for example, turn out to be present in viruses that infect hosts in all three domains, suggesting that these genes originated in a virus that infected an ancestor of all three domains of cellular life. “We find more and more evolutionary connec-
tions between viruses in different domains,” says Forterre. “All these findings have completely destroyed this old idea that viruses are escaped fragments of cells.”

A puzzle that arose from work on microbial genes in the 1990s also gave a new dimension to Forterre’s thinking. As scientists dissected the DNA that codes for enzymes used in replication, they found that the precise sequences in bacteria were radically different from those in archaea and eukaryotes. The discrepancy between the domains of life meant that DNA essential to survival had come from different sources or had been separated by a long period of evolutionary time. Koonin even proposed that DNA replication had evolved twice, once in bacteria and once in the common ancestor of archaea and eukaryotes.

Forterre had a different reaction. “I came back to this idea [of a very early origin for viruses] when genomics showed a clear difference between the DNA replication of bacteria on one side, and the system in archaea on the other,” says Forterre. “Maybe in fact one of these two systems came from a virus.”

Forterre proposed that the genes for DNA replication in bacteria had evolved twice by their viruses. Soon afterward, Luis Villarreal of the University of California, Irvine, pointed out a possible evolutionary connection between DNA-replication genes in certain viruses and those in eukaryotes. He suggested that viruses had replaced the original genes in eukaryotes. “The next step,” Forterre says, “was to say, ‘Why not both systems?’”

For the past several years, Forterre has been expanding his original ideas into a sort of grand unified theory of viruses and cellular life. Forterre proposes that viruses donated more than just their DNA-replication genes to cellular life. He argues that they donated DNA itself. In recent months, he has presented the scenario in a series of papers, the most recent of which appeared last month in the journal *Virus Research*.

Like many scientists, Forterre favors the theory that DNA-based organisms are descended from simpler RNA-based organisms. Experiments on RNA suggest that it could have been versatile enough to support primitive life. Not only can it carry genetic information, but it also has the capacity to act like an enzyme, carrying out chemical reactions. RNA-based life may have been able to absorb nutrients, replicate, and evolve. According to this “RNA world” theory, these organisms later evolved proteins and DNA, which then took over many of RNA’s former tasks.

Forterre proposes that RNA organisms evolved into self-replicating cells that could produce their own proteins. At that point, the first viruses evolved. These RNA viruses parasitized RNA-based organisms, manipulating them to make new copies of themselves. These primordial RNA viruses may have produced lineages that are still with us today, in the form of modern RNA viruses such as influenza, HIV, and the common cold.

Although a great deal of evidence supports the idea of an RNA world, the scenario raises a number of difficult questions. Not the least of these is how RNA-based life might have evolved into DNA-based life. Many scientists have pointed out that DNA is more stable than RNA and less prone to mutations. Once DNA was established, it allowed genes to become longer and more complex. But how the transition came about is difficult to explain, Forterre points out. The stability of DNA provides long-term advantages, not the short-term ones that natural selection can favor. “This is not a Darwinian way of thinking,” says Forterre. “You should explain why the first organisms in which RNA was modified had a selective advantage, not in its descendants.”

Forterre offers a solution. He suggests that viruses were the intermediate agents of change. For viruses, DNA might have offered a very powerful, immediate benefit. It would have allowed them to ward off attacks from their hosts. Cells today use a number of weapons against RNA viruses. They can silence the viral RNA with special RNA molecules of their own. Or they can cut the genome of the virus into fragments.

“RNA viruses have to find a way to avoid these defenses,” says Forterre. They do so by making it difficult for their hosts to grab their RNA. Living RNA viruses chemically modify their genes to thwart their hosts. Forterre proposes that some early RNA viruses altered their genes in a particularly effective way: They combined pairs of single-stranded RNA into double-stranded DNA. The vulnerable nucleotides carrying the virus’s genetic information were now

---

**RNA world**

- RNA
- DNA virus
- Last Universal Common Ancestor (LUCA)

**RNA-to-DNA transition**

- DNA-infected RNA cells
- RNA virus

**DNA world**

- Bacteria
- Archaea
- Eukarya

*From RNA world to DNA world.* Forterre proposes that all living organisms share a common ancestor that stored its genetic information in RNA. Some of its genes evolved into viruses. Later, some of those viruses evolved DNA as a way to defend their genes from attack, and DNA-based viruses became incorporated into hosts. Host genes were then transferred onto viral chromosomes and shared. In the process, the three major domains of DNA-based life emerged.

---

**I really applaud the bravery and intellectual power to come up with this picture.”**

—Eugene Koonin, NCBI
Once roved superior to those of their predecessors.

omains of DNA-based life, the new forms
hey originated," Forterre admits.

p of the features of the eukar
yote cell. “The
aves of viral infections could have built some
W
form nucleuslike shells inside their hosts.

ut that certain viruses, such as poxviruses, can
y eukaryotes. He points
lt to any account of the early evolution of
otes shared a close common ancestor.
R
just so happened that the vir
are to bacterial genes. F
orterre suggests that it
y eukaryotes DNA-replication
that archaean and eukar
iments, Forterre proposes that
one gave rise to one of the three domains of life.

These blendings of genomes could explain the similarities and differences among the three domains, he argues. Take, for example, the fact that archaean and eukaryote DNA-replication enzymes are more similar to each other than they are to bacterial genes. Forterre suggests that it just so happened that the viruses that infected the RNA-based precursors of archaean and eukaryotes shared a close common ancestor.

Eukaryotes in particular pose a special challenge to any account of the early evolution of life. For instance, eukaryotes keep their DNA tucked away in a nucleus, a structure whose origins scientists have debated for years. Forterre suggests that viruses may have played a part in shaping the cells of early eukaryotes. He points out that certain viruses, such as poxviruses, can form nucleuslike shells inside their hosts. Waves of viral infections could have built some of the features of the eukaryote cell. “The eukaryote cell is very strange and very complex, so I don’t have a very clear idea of how they originated,” Forterre admits.

After viruses ushered in these three domains of DNA-based life, the new forms proved superior to those of their predecessors. “Once this occurred, the DNA cells outcom-

Viral diversity. Scientists are uncovering a vast diversity among viruses, such as this giant mimivirus. Their genetics and biochemistry point to an ancient origin.

Only at this point, Forterre argues, could RNA-based life make the transition to DNA. From time to time, genes from the RNA chromosome would be accidentally pasted into the virus’s DNA chromosome. These genes could now enjoy all the benefits of DNA-based replication. They were more stable and less prone to devastating mutations. Natural selection favored organisms carrying important genes in DNA rather than RNA. Over time, the RNA chromosome dwindled while the DNA chromosome grew. Eventually, the organism became completely DNA-based. Forterre proposes that this viral takeover occurred three times, and each one gave rise to one of the three domains of life.

These blendings of genomes could explain the similarities and differences among the three domains, he argues. Take, for example, the fact that archaean and eukaryote DNA-replication enzymes are more similar to each other than they are to bacterial genes. Forterre suggests that it just so happened that the viruses that infected the RNA-based precursors of archaean and eukaryotes shared a close common ancestor.

Eukaryotes in particular pose a special challenge to any account of the early evolution of life. For instance, eukaryotes keep their DNA tucked away in a nucleus, a structure whose origins scientists have debated for years. Forterre suggests that viruses may have played a part in shaping the cells of early eukaryotes. He points out that certain viruses, such as poxviruses, can form nucleuslike shells inside their hosts. Waves of viral infections could have built some of the features of the eukaryote cell. “The eukaryote cell is very strange and very complex, so I don’t have a very clear idea of how they originated,” Forterre admits.

After viruses ushered in these three domains of DNA-based life, the new forms proved superior to those of their predecessors. “Once this occurred, the DNA cells outcom-

headed all the RNA cells,” says Forterre. “And once all the RNA cells were eliminated, you had a limited number of lineages.”

Finding flaws
Forterre’s ideas have been warmly received, even by those who dispute some of them. “Patrick’s ideas need to be taken seriously,” says David Penny of Massey University in Palmerston North, New Zealand. But Penny points out that in Forterre’s scenario, RNA-based organisms are already relatively complex by the time viruses drive them into the DNA world.

“I doubt in the absence of DNA any organism would be very complex at all,” Penny says. He points out that RNA replication suffers a high error rate. Under those conditions, genomes cannot become large without risking catastrophic damage.

Penny and his colleagues dispute the idea that natural selection could not drive RNA organisms to DNA-based replication on their own. They’ve argued that the organisms could have made the transition through a series of evolutionary steps, each of which reduced the error rate during replication. Once life had shifted to DNA, Penny argues, viruses might have played a part in the origin of one or more of the three domains. “I definitely do not want to exclude the possibility of viral takeover,” says Penny.

Koonin, on the other hand, agrees with Forterre that viruses that originated in the RNA world. “That idea makes perfect sense, in my opinion,” he says. He also accepts the possibility that viruses might have invented DNA. But when Forterre argues that viruses gave rise to the three domains, Koonin parts company. “I don’t believe it for a moment,” he declares. Archaean share too many genes in common with eukaryotes for this to be plausible, Koonin argues. Instead, eukaryotes must have evolved from

archaens after they had become DNA-based.

Koonin envisions a different sort of history of viruses and their hosts. He and William Martin of Heinrich Heine University in Düsseldorf, Germany, have proposed that life first evolved in honeycomb-like cavities in rocks around hydrothermal vents. These compartments played the role that cell walls and membranes would later play. Initially, RNA molecules were selected simply for fast replication. Successful molecules could spread from one compartment to another. Over time, groups of RNA molecules were favored by natural selection, working together to reproduce more successfully. It was then that viruslike things emerged—before cells yet existed.

“There would be parasites that only care for their own replication—and here Patrick’s ideas might have their place,” Koonin says. “DNA replication might originally have emerged in such parasitic entities.”

Woese, who has influenced Forterre’s work for more than 20 years, is both enthusiastic and agnostic about the virus theory. “In the most specific form, I don’t know if he’s right, but then it doesn’t make any difference, because he’s going in the right direction,” he says. “I think that’s a significant advance, to be able to fold the viruses into the whole process.”

Woese and others believe that the best way to assess Forterre’s theory would be to find and analyze more viruses. Forterre agrees. “At the moment, we know very few of the viruses in the living world,” says Forterre. “We know some viruses that are human pathogens, and some that infect bacteria that are important to the food industry. But for many groups, we have no idea of their viruses.”

Forterre himself is trying to fill that gap by studying viruses that live in heat-loving archaean. Archaean viruses are proving to be particularly diverse and bizarre—such as lemon-shaped species that don’t finish growing until they have left their host cell. Because viruses have such ancient roots, they preserve a remarkable range of biochemical tricks. “It’s clear now that you have many more genes in the viral world, so there are many interesting new enzymes to be found,” says Forterre. “If we explore the viral world more, I don’t know if we will be able to be sure of one theory or the other. But I am sure we will get many more interesting molecular mechanisms.”

—CARL ZIMMER

Carl Zimmer, a freelance writer, is the author of Parasite Rex and Soul Made Flesh.