



## A New Method to Track Viral Evolution

**I**n 2009, Humboldt County in northwest California experienced an unusual rabies outbreak. Scientists at the California Department of Public Health observed a dramatic increase—up 355 percent—in the number of foxes with the virus. Rabies is a viral disease that affects the central nervous system and can be spread to humans through bites or scratches by rabid animals. Human rabies cases are exceedingly rare in the U.S., in large part because domestic animals are vaccinated against the disease. However, if preventive treatment is not administered to an infected person before symptoms develop, the disease is almost invariably fatal.

The typical rabies virus found in California terrestrial animals is the skunk variant, meaning skunks are the hosts that transmit it to other animals. Skunks with rabies tend not to attack people or other

animals that share their habitat, but rabid foxes do, making the Humboldt County outbreak a concern for health officials.

Sharon Messenger, a virologist at the Department of Public Health, suspected that the 2009 outbreak was being spread through a phenomenon called species or host jumping. She asked a team of Livermore researchers, led by biomedical scientist Monica Borucki in the Physical and Life Sciences Directorate, to examine brain tissue samples collected from infected animals. The team, which includes Livermore scientists Jonathan Allen, Haiyin Chen, Tom Slezak, and Clinton Torres in the Computation Directorate, theorized that the rabies genes mutated as the virus jumped from skunks to foxes. The researchers are now using the Laboratory's computational resources and their expertise in bioinformatics to

In support of the California Department of Public Health, Livermore researchers analyzed samples of rabies genome taken from skunks and foxes using ultradeep sequence analysis.





A map of Humboldt County in California shows the fox encounters and rabies-positive foxes reported from January 1 through August 18, 2009.

better understand the evolution of rabies and other ribonucleic acid (RNA) viruses, an essential step toward improving biological defense against these diseases.

### Going Viral

RNA viruses mutate about 1,000 times faster than bacteria and DNA-based viruses, and many of the new diseases infecting people are RNA viruses that have jumped from animal hosts. These diseases are a homeland security concern, particularly those that have pandemic or epidemic potential, and the Department of Defense considers viral infections a significant threat to soldiers deployed around the world.

The Livermore researchers are developing a more sensitive approach to identify mutations and predict which ones have the potential to jump from host to host. “We’re interested in understanding how a virus population changes as it adapts from one host species to another,” Borucki says. Initially supported by Livermore’s Laboratory Directed Research and Development Program, the team’s work is now funded by the Defense Threat Reduction Agency. Says Allen, “This study is giving us a new appreciation for the complexity and dynamism of viral evolution.”

To understand how viruses mutate, the Livermore team is analyzing 50 rabies samples from the Humboldt outbreak as well as 35 samples of bovine coronavirus—another RNA virus—previously used in experiments to simulate a host-jumping event. In the exploratory stages, the researchers are developing computational tools that can analyze viruses to identify trends

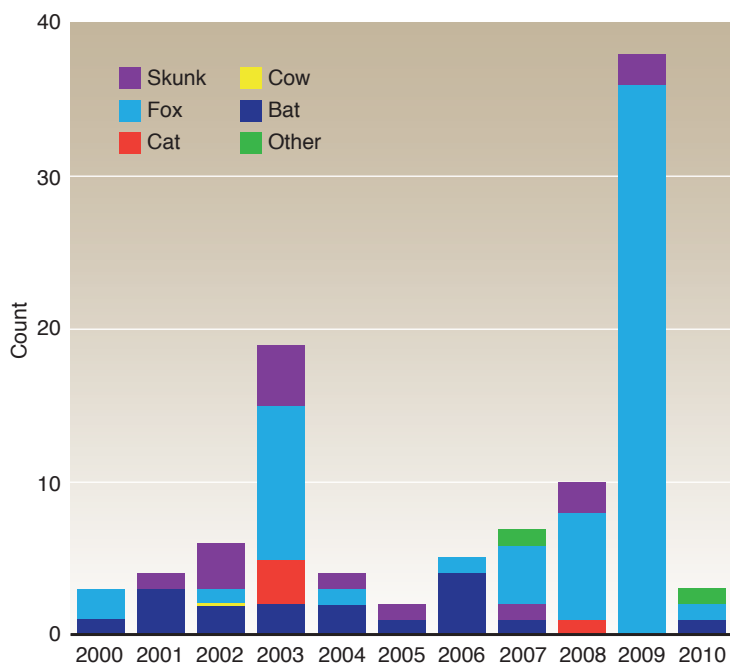
and potentially important mutations—what Allen calls “mutation prospecting.” They will use this information to build a database that, when combined with viral studies at other laboratories, can help them spot rare variants and predict which viruses are most likely to jump from host to host.

Genetic variations found in terrestrial animals that carry rabies indicate each animal’s geographic location. The rabies virus in a California skunk looks genetically different from that of a skunk in Texas. Thus, when a person contracts rabies, a small part of the viral sequence can be analyzed to determine its source, for example, as a raccoon variant from South Carolina.

“Although the rabies virus has jumped species multiple times in the past, it is still a relatively rare occurrence,” says Borucki. “For example, a bat rabies virus in Arizona changed to infect skunks one year. Then nine years later, it changed to infect foxes.” Genetic sequencing indicates that fox-to-fox transmission caused the Humboldt County outbreak. However, the virus originated in a skunk and jumped to a fox. According to Allen, the researchers’ hypothesis is that mutations acquired by the virus allowed it to move through the fox population quickly.

### Capturing Uncommon Variants

In working with the rabies virus, researchers normally sequence only a small part of its genome, about 300 to 400 nucleotides taken from up to 30 copies of the viral RNA. Genome sequencing technology generates copies from random fragments of a genetic sample, and researchers piece those snippets together to look for



This graph compares the number of rabies-positive animals in Humboldt County from 2000 to 2010.

patterns in the variants. The Livermore team uses a technique called ultradeep genome sequencing, which provides more coverage of a given population. In the Humboldt County study, the team analyzed 11,000 of the 12,000 total nucleotides in the rabies genome, using 300,000 copies of the genetic material. Such in-depth analysis provides a more complete picture of the rabies genome, including the uncommon variants in the virus, and helps clarify the diversity of the viral strains.

Some viruses evolve through recombination, a process in which genes swap information. Two viruses might even enter the same cell and take pieces from each other's genomes to create a new disease, such as the coronavirus mutation that led to severe acute respiratory syndrome. "In our work with the bovine coronavirus genome, we found samples with an extra piece of an amino acid sequence that appears to serve a biological function," says Borucki. "This piece commonly doesn't show up in surface analysis. We found the variant at a deeper level when we extended the population analysis." Uncovering such a variant is crucial because it could be the one that supports the mutation needed for a virus to infect another species.

The team is also experimenting with a method that is less deep but includes longer fragments. "Generally, we look at mutations in isolation, but they are most likely correlated with other parts of the genome," says Allen. "Two mutations might be in close proximity

but not picked up on the same fragment. By reading a longer fragment, we could potentially establish that the two mutations link to the same gene."

### Trending Data

The researchers plan to document all of the mutations and differences found in the sample genetic data. Both methods, ultradeep sequencing and long-fragment analysis, will yield huge amounts of data, which will require high-performance computing resources to analyze. Borucki says, "Without Livermore's biocomputational and mathematic capacity, we couldn't do this kind of experiment." Sequencing a single sample generates about 1.5 billion bytes of data in a compressed format. The team has already processed about 70 samples. A 300-sample study would require half a trillion bytes of storage capacity. The data files must also be professionally maintained and backed up, which requires Livermore's information technology expertise.

Allen is working on a software code to organize the data so researchers can discover trends. As sequencing technology gets faster and cheaper, real-time processing will be possible, allowing scientists to predict which viruses are likely to jump from host to host. Tools that can help them understand the potential function of virus changes and the variability present in a virus population could be a boon for forensic studies and vaccine development.

Borucki is also using results from this project to educate people about rabies. She notes that many people do not understand that it doesn't take a bite to be infected with rabies. The disease can be transmitted by a scratch from an infected feral cat, for example, and delays in treatment are dangerous. "When I talked to Livermore high school students, few of them knew that rabies is nearly 100 percent fatal in humans," Borucki says. "By getting this issue on people's radar, we have the potential to influence their behavior."

In the future, the Livermore analysis tools could be used to look for trends in viruses in different environmental situations. With a sufficient pool of candidate mutations—another piece of the data puzzle still needed—public health officials could identify new viral threats to humans in time to produce vaccines against those particular bugs. With advance warning, medical professionals and first responders could then treat the initial infections more efficiently and reduce the need for mass immunizations.

—Kris Fury

**Key Words:** bovine coronavirus, genetic variation, genome, host jumping, mutation, sequencing, rabies, ribonucleic acid (RNA) virus.

**For further information contact Monica Borucki (925) 424-4251 (borucki2@llnl.gov).**