Abstractions

LEAD AUTHOR

In late 2004, Mount St Helens in Washington, the most active of the US Pacific Northwest volcanoes, started a steady eruption that continues today. Earthquakes occur almost constantly as a new dome is pushed into place by rock and magma below. Richard Iverson and colleagues at the US Geological Survey's Cascades Volcano Observatory in Vancouver, Washington, and at the University of Washington in Seattle, have spent the past two years trying to uncover the cause. The data led Iverson to propose the unusually simple model of the volcano's underground mechanics found on page 439.

What has been happening at Mount St Helens in the past two years?

Before 2004, nobody thought that eruptions at Mount St Helens would happen again in our lifetime. Since then, there have been more than one million earthquakes accompanying the eruption. That's not unusual, but what's been extraordinary about these earthquakes is their uniquely periodic nature — so much so we dubbed them 'drum beats'. For months at a time, earthquakes were happening once or twice a minute. Very few things in solid earth sciences are so obviously repetitive — that's what led us to the idea behind this model. The eruption has been a great opportunity to do research because it has been so well behaved.

What's novel about the model detailed here?

The model assumes magma comes in from depth at a steady rate, solidifying and pushing a solid body upwards. The earthquakes result from a lurching motion, called stick-slip behaviour, which accompanies incremental pressure build-up and release as solid rock is pushed out of the ground. Our model demonstrates that something as complicated as a volcano may exhibit surprisingly simple mechanical behaviour.

What led you to a less complex model?

The single most exciting thing about the repetitive earthquakes is that they call for a very simple explanation based on a mechanical process. Whatever is generating the seismicity can happen over and over again without destroying itself. Our data suggest that two solid surfaces are persistently grinding against one another, which can continue if the supply of new rock continues.

How long can Mount St Helens keep erupting?

Most of us expected it to have shut off by now. There is little evidence of systematic changes at depth, implying the magma is being resupplied at depth. Understanding what would shut that off forces us to consider progressively deeper processes, leading us to subduction and melting of tectonic plates. ■

MAKING THE PAPER

Len Pennacchio

A brute-force approach uncovers elusive gene regulatory sequences.

Only 2% of the human genome gives rise to proteins. Stretches of DNA in the remaining 98% of non-coding sequences inform genes when and where to be turned on or off. Len Pennacchio of the Lawrence Berkeley National Laboratory, California, went fishing for enhancers, a particularly elusive type of these regulatory sequences.

Unlike promoters, which typically sit immediately before a gene and signal where transcription should begin, enhancers exist before, after or even within the genes they regulate. In some cases, they function millions of base pairs away from the genes they help control. Scientists don't yet know enough about enhancers to use mathematical algorithms to distinguish them from other sequences. "We haven't identified enough distant enhancers to enable computational scientists to train on this data set," explains Pennacchio. Instead, he and his colleagues came up with a brute-force approach for testing a large number of DNA fragments for enhancer activity. Their work is described on page 499.

First, the team needed to select DNA fragments to test. They reasoned that bits of the genome that do not encode proteins but are extremely well conserved among vertebrates must have a fundamental function and thus be good candidates. So they scoured the human genome for DNA segments that are either conserved with a distantly related species, such as the pufferfish, or ultra-conserved, meaning they are 100% identical with a more closely related species, such as the mouse, for at least 200 base pairs. Their initial list contained 167 such elements, which they tested for function.

Pennacchio's team fused each putative enhancer fragment to a gene engineered to turn blue when expressed, and then injected the fragments into fertilized mouse eggs. They implanted the eggs into female mice and let the eggs grow for 12 days. Then, they examined the resulting embryos for spots of blue in different organs and tissues. "Traditionally, people generate transgenic mice, wait for the mice to grow to maturity and then look for gene expression in the resulting transgenic offspring," says Pennacchio. "We streamlined the process, making it faster and cheaper."

Over about two years, the team identified 75 enhancer elements (45% of the tested fragments), whose sequences and expression patterns are available at http://enhancer.lbl.gov. A scientist wanting to engineer a gene so that it will only be expressed in the heart, for example, can use the public resource to obtain a list of enhancer sequences that drive expression specifically in that tissue. The data set also allows scientists to search for patterns in the sequences of all heart enhancers and develop computational tools to predict where to find other heart enhancers in the genome. In addition, Pennacchio's team is testing potential enhancer elements that other groups have identified. "I receive numerous e-mails per day from people asking us for help," he says.

The group will spend the next five years testing about 2,000 additional elements to add to the resource, with the goal of finding enhancers for all human genes. And they might try other methods of enhancer identification. "For now, comparative genomics alone has worked well for us, but there are surely other ways to further prioritize possible enhancers," says Pennacchio. ■

KEY CONTRIBUTORS

Trolling through 270 people's DNA to identify gains and losses of genetic sequence is a daunting prospect. Five groups identified and mapped these 'copy number variants', which probably play key roles in genetic diseases (see page 444). But making sure the data captured real genetic variance — and not artefacts from *in vitro* samples or automated data analysis — proved the biggest challenge, says Matt Hurles of the Wellcome Trust Sanger Institute in Cambridge, UK. "We had to invent the quality control," he says. "These samples are all derived from cell lines, which sometimes rearrange their genomes."

To avoid artefacts, Charles Lee's group at Brigham and Women's Hospital in Boston, Massachusetts, with Stephen Scherer's team at the Hospital for Sick Children in Toronto, Canada, grew the 270 cell lines, examined them at key stages of division and sorted out discrepancies that seemed a result of cell culturing rather than real copy number variants.

They found obvious problems in 30 cell lines, the most common being signs of three rather than two chromosomes in some cells — a near-impossibility, says Hurles, "you just don't see these particular extra chromosomes in people". Don Conrad of the University of Chicago, Illinois, then used family information to sort out less obvious discrepancies. ■