GENETICS

The sequence of the salamander

Nature **554**, 50–55 (2018)

The genome of the aquatic axolotl salamander, a native of Mexico's lakes, has yielded some surprises, and the technique used could point the way to analysis of other organisms that have complex genomes with large numbers of sequence repeats, such as the lungfish and many species of plants.

The axolotl has attracted much interest as a model organism because salamanders are capable of limb regeneration, a rarity among vertebrates. But they also have large genomes that are riddled with repeating sequences that complicate sequencing attempts. These repeats originate from genetic elements called transposons, which copy themselves over and over again, often sequentially.

Standard sequencing relies on single read lengths of up to about 600 base pairs, which must then be overlaid and stitched back together to read the entire sequence. But long sections of repeats require much

longer read-lengths in order to match up unique sequences on either end of the repeat, and that poses technical challenges with existing sequencing technology. It's a bit like trying to overlay photos to take a panoramic photo in a forest. "Is (the photo) this tree over here, or a stretch of forest over there?" says co-author Michael Hiller, who is a research group leader at the Max Planck Institute in Dresden, Germany.

The researchers used Pacific Biosciences instruments to collect reads of about 14 thousand base pairs in length, an impressive feat that comes with a heftier price tag than other sequencers. They combined these longer read lengths with a software package that keeps the information from long reads intact whenever possible. Other assembly algorithms break the reads into pieces whenever they detect uncertainty, such as stretches with an elevated sequencing error rate.

This new approach preserves the longer reads, allowing them to span the sequence repeats. Finally, the researchers employed a scaffold program from the Bionano Genomics Saphyr system that further stitches the assembled genomic parts together to create even longer sequences.

The technique worked, but it wasn't cheap. "It's very expensive, but the axolotl genome serves as proof of concept" that other giant genomes can be successfully sequenced, says Hiller.

The analysis also turned up a surprise in one of the axolotl's developmental genes. The animal completely lacks the *Pax3* gene, which is present in all other known vertebrates; loss of its function is catastrophic in mice. The gene is a key regulator of expression of other genes during development, and in axolotls, its role appears to have been replaced by the related *Pax7*.

"We don't know yet what this really means, if it's important for regeneration or something else, or simply a coincidence, but the genome allowed us to see that this gene and the surrounding sequences are completely absent. They are neither found in the genome assembly, nor in the individual reads," says Hiller.

Jim Kling

Published online: 26 March 2018
<https://doi.org/10.1038/s41684-018-0031-x>