Advancing wildlife genomics through the development of molecular methods

October 19 2020

Scientists tested the new SIP method for genome sequencing on the koala retrovirus. Credit: David Clode

A team of scientists from the Leibniz Institute for Zoo and Wildlife Research (Leibniz-IZW), the Australian Museum and the Max Delbrück

Center for Molecular Medicine (MDC) report a new method for identifying any genome sequence located next to a known sequence. It is often difficult to precisely determine unknown sequences close to small known fragments. Whole genome sequencing can be a solution, but it's a very cost intensive approach. In order to find a more efficient technique, the scientists developed Sonication Inverse PCR (SIP): First, DNA is cut into random pieces using ultrasound waves. After DNA fragmentation, long-range inverse PCR is performed followed by long-fragment high-throughput sequencing. SIP can be used to characterize any DNA sequence (near a known sequence) and can be applied across genomics applications within a clinical setting as well as molecular evolutionary analyses. The results are reported in the scientific journal *Methods in Ecology and Evolution*.

Many methods have been developed to identify sequences next to a determined sequence of interest. Inverse PCR based methods are among the most common methods and have been used for decades but suffer from bias because of the way DNA is cut apart by enzymes: They need to find specific sequence motifs that are not evenly spread across the DNA. Therefore, many neighboring sequences to a target cannot be characterized without technical difficulty or without the expense and effort of whole genome sequencing. "Sonication Inverse PCR (SIP) circumvents this problem by using high-frequency sound waves to randomly cut the DNA, eliminating the bias resulting from the use of enzymes," Prof Alex Greenwood from Leibniz-IZW explains. "The fragments are then turned into circles and the so-called inverse PCR is applied." With the development of long-fragment sequencing, the authors were able to target 4-6 thousand base long inverse PCR fragments and sequenced them at high-throughput on the PacBio RS II sequencing platform.

The new method was tested on a complex model, the koala retrovirus (KoRV), a high copy retrovirus found in the koala (Phascolarctos

cinereus) genome. Targeting the ends of the integrated virus, the full spectrum of viral integrations in the genome could be determined using a small 'known' piece of viral DNA. Mapping the integrations against reference genomes provided precise genomic locations for each integration at a resolution that would otherwise require a large sequencing effort. "Applying this method allowed us to discover a koala specific defense mechanism against KoRV," says Dr. Ulrike Löber from the MDC,

"SIP is economical and can be simultaneously applied to many samples by including barcodes to the PCR primers, making the method cost efficient," adds Dr. David Alquezar, former member of the Leibniz-IZW team and now manager of the Australian Centre for Wildlife Genomics at the Australian Museum. The authors continue to apply SIP to address different problems, such as how viruses become integrated into genomes and how they cause diseases. In conclusion, SIP provides a new protocol for high-throughput profiling of flanking sequences next to any region of interest coupled with long-range sequencing, allowing scientists to study complex biological systems such as mobile genetic elements.

More information: David E. Alquezar-Planas et al, DNA sonication inverse PCR for genome scale analysis of uncharacterized flanking sequences, *Methods in Ecology and Evolution* (2020). DOI: 10.1111/2041-210X.13497

Provided by Forschungsverbund Berlin e.V. (FVB)