GENOMICS

Targeting ancient DNA

A method that allows precise capture of Neanderthal genome sequences will permit detailed comparison of modern and ancient humans.

Imagine having to isolate DNA fragments that are very short and only make up 0.001% of your total sample. This was the challenge faced by Adrian Briggs and his colleagues in Svante Pääbo's laboratory at the Max Planck Institute for Evolutionary Anthropology in Leipzig when they set out to sequence the mitochondrial genomes of several different Neanderthal individuals.

Briggs is not a novice when it comes to sequencing Neanderthal DNA and is well aware of the difficulty. He is a member of the Neanderthal Genome Project led by Pääbo that uses a shotgun sequencing approach with Roche's 454 high-throughput pyrosequencer to decode the genome. The difficulty, he says, is that "typically what you get out of an ancient bone is more than 99% DNA from bacteria and other microbes. You have to sift through an enormous amount of sequence before you get to the locus you are interested in." Briggs recalls that it took 147 full 454 sequencing runs to obtain one complete mitochondrial sequence; going by current catalog prices that puts the price tag well over a million dollars and makes the sequencing of more mitochondrial genomes by a shotgun approach economically unfeasible.

As an alternative, the researchers wanted a method that would target specific DNA segments before sequencing. The classical technique for targeted DNA enrichment is PCR. Recently a PCR-based strategy has proven its value for obtaining sequence information from a contemporary of the Neanderthal, the cave bear. The study led by Matthias Meyer also at the Max Planck Institute for Evolutionary Anthropology showed that with a combination of standard multiplex PCR, sample barcoding and 454 sequencing, the complete mitochondrial genomes of 31 cave bears could be sequenced (Stiller *et al.*, 2009).

Rib fragment of a Neanderthal infant found in Mezmaiskaya cave, Russia. Image courtesy of A. Briggs.

Could the same strategy be applied to Neanderthal DNA? Briggs lists two main reasons why this would not be possible: first, bear bones are bigger and better preserved than Neanderthals' and thus yield more extract and longer PCR products for sequencing; second, cave-bear studies are not prone to contamination from modern bear. Whereas bear DNA is not common in the laboratory, human DNA is, and Neanderthal DNA is very susceptible to being swamped by modern human DNA. "Trying to avoid single-molecule level of contamination with hundreds of primer pairs is not a feasible way to go," Briggs explains.

In lieu of PCR, the team developed a strategy based on primer extension capture: 5′-biotinylated oligonucleotide primers target specific sequences, and a polymerase extends the primers into the adjacent sequence. As the bones serving as source material are rare and precious, the researchers wanted to use the libraries that had already been made for shotgun sequencing. Neanderthal DNA fragments in these 'immortalized' libraries have an average length of 50–85 base pairs and are flanked by sequencing adapters so they can easily be amplified by PCR with primers hybridizing to the adapter regions.

But successfully targeting a certain region and obtaining the sequence for it was only half the battle. As in all applications using next-generation sequencing, data analysis was difficult. The average read length from a 454 run is 250 base pairs, so the short Neanderthal fragments were easily contained in these reads, but assembling them to a full genome proved tricky. Newbler, the assembler Briggs and colleagues initially used to assemble the mitochondrial genome failed because it relies on overlap between the reads to assemble longer pieces. Unfortunately, ancient DNA is often damaged at fragments' ends, making perfect overlaps between fragments very rare. The researchers had to write their own program that aligned all fragments to the Neanderthal mitochondrial reference genome, and from these alignments they reconstructed five mitochondrial genomes.

These genomes had little diversity, comparable to that of modern Europeans. Considering that Neanderthals had lived in Europe for a total of at least 200,000 years, whereas Europe was colonized by a small number of modern humans only about 40,000 years ago, these data suggest a very low effective population size for Neanderthals.

Another important aspect of this work lies in the method of primer extension capture itself, which according to Briggs "is tailored to libraries of short fragments and particularly complex background." It will allow the targeting of any sequence of interest, not just mitochondrial DNA. Earlier this year the Neanderthal Genome Project completed genome sequencing at 1.5-fold coverage, and the project is now in the analysis phase. It is likely that there will be regions that warrant a closer look, and primer extension capture provides a way to extract these sequences for closer scrutiny.

Homo sapiens and *Homo neanderthalensis* can then be compared gene by gene.

Nicole Rusk

RESEARCH PAPERS

Briggs, A.W. *et al.* Targeted retrieval and analysis of five Neanderthal mtDNA genomes *Science* **325**, 318–321 (2009).

Stiller, M. *et al.* Direct multiplex sequencing (DMPS): a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Res.* advance online publication (27 July 2009).