

Record for decoding the longest DNA sequence is impressive – here's what to expect next

December 28 2018, by Angus Davison



Jeremy the left coiling snail on top of a right coiling snail, Theresa. Credit: Angus Davison, [CC BY-SA](#)

Like other professionals, scientists like to be the best at what they do,

but they also like to have fun in their job. And in 2018, my colleagues managed just that in claiming a record for [decoding the world's longest DNA sequence](#).

For the English scientists involved, perhaps the most important fact is that their DNA read was about twice as long as the previous [record](#), held by their Australian rivals. The glory of gaining the record is the result of an Ashes-style competition to produce ever longer DNA [sequences](#). The record has exchanged hands several times over the past year, but with this new sequence the trophy seems to be safe in the UK – at least for the moment.

But as exciting as it is to win, the most inspiring thing about this record is the science and the future applications that could become available thanks to our ability to decode ever longer sequences.

Jigsaw jumble

The technology that enables scientists to read runs of DNA sequences has come a long way since the millennium-era race to [decode the first human genome](#). There are lots of ways you can now read DNA, but the problem is that many animal and plant genomes are often billions of [base pairs](#) (pairs of DNA building blocks known as A, T, G and C) and so making sense of them is tricky. People have used different methods in the past, but essentially what they do is chop the DNA up into small parts, read each piece and then try to assemble the results back together, a bit like what you would do with a jigsaw puzzle.

Putting the DNA pieces together in the correct order is therefore a major obstacle when it comes to DNA sequencing. This is obviously harder the more pieces you have, especially if they are short and very similar to each other.

Being able to continuously read ever longer pieces – eventually an entire chromosome in one go – would therefore have a huge impact on science and innovation. In my own research, I am interested in finding the genes that determine the left and right sides of animal bodies. And while I can fairly straightforwardly read the genome of [snails like "Jeremy"](#) – which has a shell that coils left instead of right – it is very difficult to make sense of it, because the order is almost completely jumbled.

My colleague, [Matt Loose, also at the University of Nottingham](#), led the team behind the new world record, which [read 2.3m bases of human DNA in one go](#). Putting that in context, in the most common form of DNA sequencing only a few hundred bases are read at once, creating millions of pieces to put together. If a few hundred bases are equivalent to once around a grand prix track, then a 2.3m base pair read is twice around the circumference of the Earth. In comparison, the main rival Australian team at the [Kinghorn Centre for Clinical Genomics](#) is still some way behind. They have still to get once around the world.

Long reads and small holes

The key technology that is pushing these advances is a very small hole, [called a nanopore](#). DNA bases, or letters, are ratcheted through the nanopore, and the order can be read by monitoring disruptions to an electrical current put through it. If the nanopore were scaled up to the size of a thumb and forefinger pinch, then the scientists would have threaded a rope of over seven kilometres in length through the hole, without it becoming tangled or breaking. In comparison, a more typical DNA sequence would be about half a metre in length.

In theory, sequencing a whole chromosome in one go should be possible using this method. This would then avoid the problem of trying to assemble a massive jigsaw. But natural breaks in each chromosome mean that this may not be possible. Whatever the actual limit of read

length, the new methods are already being used to more quickly and cost effectively identify pathogens in disease outbreaks. The same methods are also being used to rapidly and accurately characterise the genome rearrangements that take place as cells progress to become cancerous.

A recent proposal to sequence [the genomes of 1.5m known animal, plant and fungal species](#) will also benefit from these new long-read technologies. In future, the methods will help enable truly personalised medicine – having our individual genomes sequenced. In the UK, [about 85,000 people](#) have already had their entire genetic code read, with an ambition to sequence a million genomes in the next five years. For the moment, most of this is being done using older, short-read technology, which is still cheaper but misses an important layer of structural information.

In my own laboratory, I plan to use the same methods to find the genes that sometimes enable snails to exist in two mirror-image versions of themselves. The same methods may also be used to further unravel the genetics of human diseases, especially those that are due to structural rearrangements and changes in gene copy number.

The scientists behind the record believe that their record might last for a year or so. And the competition is expanding to include other competitors – just in the last month, [a new entrant from the Netherlands](#) came within a whisker of beating the UK record.

But given what's at stake, fierce competition can only be a good thing.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Record for decoding the longest DNA sequence is impressive – here's what to expect next (2018, December 28) retrieved 28 April 2024 from <https://phys.org/news/2018-12-decoding-longest-dna-sequence.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.