Genomics 101 transcript: What is long-read vs short-read sequencing?

**Naimah:** At Genomics England, we read DNA sequences using both long-read and short-read sequencing, but what is the difference between these two methods? Today I will be speaking to Emma McCargow, who is the Programme Lead for Cancer at Genomics England, to find out more.

So, first of all, Emma, can you tell me what is the difference between long-read and short-read sequencing?

**Emma:** So, long-read sequencing is a relatively new methodology for sequencing of DNA. If you can imagine that the existing sequencing that is done at the moment breaks up DNA into lots of small fragments. The fragments are almost 900 base pairs, so they are quite small, and if you imagine there are lots and lots of these base pairs, lots of little fragments, these are then pieced back together again, and then we read that across the genome. For long-read sequencing, we split up the DNA. So, we get these fragments, but they are in much longer pieces. These are 10,000 right the way through to really massive strands, massive fragments of sort of 100,000 base pairs. So, it is very similar to a jigsaw. If you imagine you have two different jigsaws, one is an adult jigsaw with lots of very small pieces, and the other is a toddler’s jigsaw. Long-read sequencing is very much like a toddler's jigsaw, it’s much easier to put back together. It is actually a more efficient and cost-effective way of sequencing nowadays, as well.

**Naimah:** So, how can these two methods of sequencing be applied?

**Emma:** So, one of the things that we are doing in Genomics England is using long-read sequencing in cancer. The reason that we want to look at this across cancer is because there are certain different types of cancers that have specific variations in their genome, such as repeats. So, this is where you might have lots of base pairs that are the same base pairs repeated lots of times. In short-read sequencing, it is really hard to see those repeats. When you piece the genome back together again, sometimes you miss those changes in the genome. With long-read sequencing it is much easier to be able to identify the variations in that genome and the changes that happen in the genome. So certain types of cancers are more suited to long-read sequencing.

**Naimah:** Is there a difference in the time it takes to do long-read sequencing compared with short-read sequencing?

**Emma:** So, the actual time it takes is fairly similar, but actually one of the things that we are looking at in Genomics England is how we apply the technology. So, the operating model. And we are looking, as one of the programmes that we are doing, to see if it is something that we can have the technology and sort of much closer to the patient. So, it is more about the process flow of how we get the sample right through to being sequenced, and then how we generate the data as well.

**Naimah:** And is the long-read sequencing something that is done in high Genomics England?

**Emma:** Yes, so in our R&D lab in Hinxton, we have an amazing team who are researching and doing lots of developmental work using these technologies and they are exploring all sorts of different things with different sequencing methodologies. It is definitely something that our team are looking into, and they are exploring its use case in different ways.

**Naimah:** Is long-read sequencing quite a new concept for Genomics England?

**Emma:** Yes, so it is a relatively new and novel methodology in itself. It is quite new across the board and not just for Genomics England. We have been sort of exploring this for a couple of years now certainly within the R and D team, they have been looking at this a little bit earlier. In the cancer programme, we are looking to work with our partners in NHSE (NHS England) to see if this is something that we could explore for NHSE as it is a sort of diagnostic capability.

**Naimah:** You mentioned that we are using long-read sequencing in the cancer initiative in Genomics England, are we going to look at expanding this into the other gene initiatives at Genomics England?

**Emma:** Yes, so I know that some of my colleagues are also looking at this for the different initiatives, this is definitely something into which we are looking. The reason that we kind of picked cancer first is really because, as I said before, there are certain types of cancers that really would be very suited to long-read sequencing. So, we see that this is where it will be really quite beneficial for patients of the future.

**Naimah:** That was Emma McCargow explaining the difference between long-read and short-read sequencing. If you have any questions, please feel free to email us at info@genomicsengland.co.uk.Thank you for listening.