**What is the Genomics England Research Environment?**

**Genomics 101 transcript**

**Florence: What is the Genomics England Research Environment? My name is Florence Cornish and I'm here with Emily Perry, Research Engagement Manager at Genomics England, to find out more.**

**So Emily, before we dive into the Research Environment, let's set some context. Could you explain what Genomics England is aiming to do as an organization?**

**Emily:** So, Genomics England provides genome sequencing in a healthcare setting for the National Health Service in England. As we sequence genomes for healthcare, the benefit is that we can also put that genomic and clinical data out for research in a controlled manner, and then that can also feed back into healthcare as well. So, it's really, this kind of cyclical process that Genomics England is responsible for.

**Florence: And so, what do we mean when we say Research Environment?**

**Emily:** So, the Research Environment is how our researchers can get access to that clinical and genomic data that we get through healthcare. So, it's a controlled environment, it's completely locked down, so it's kind of like a computer inside a computer. And in there, the researchers can access all of the data that we have and also a lot of tools for working with it in order to do their research.

We refer to the data as the National Genomics Research Library, or the NGRL. The NGRL data is provided inside the Research Environment

**Florence: So you mentioned the National Genomic Research Library. If any listeners want to learn more about this, you can check out our previous Genomics 101 podcast: What is the National Genomic Research Library?**

**And so Emily, could you talk about what kind of data is stored in this library?**

**Emily:** So the library is made up of both genomic data and clinical data, which the researchers use alongside each other. The genomic data includes what we call alignments, which is where we match the reads from sequencing onto a reference sequence, and variants, which is where we identify where those alignments differ from the reference sequence, and this is what we are looking for in genomic research.

The clinical data includes the data that was taken from our participants at recruitment, so details of the rare disease, the cancer, that they have, but also medical history data. So, we work with the NHS and we're able to get full medical history for our participants as well.

This is all fully anonymized, so there's no names, there's no dates of birth, there's no NHS numbers. It's just these identifiers which are used only inside the Research Environment and have no link to the outside world.

**Florence:** **And so how is this clinical and genomic data secured?**

**Emily:** So, as I said there's no names, there's no NHS numbers, there's no dates of birth. And we have very strict criteria for how people can use the data. So researchers, in order to get access to the Research Environment, they have to be a member of a registered institution, they have to submit a project proposal for what it is that they want to study with the data.

There's also restrictions on how they can get the data out, so they do all their research inside, there's no way that they can do things like copy and paste stuff out or move files. The only way that they can get data out of the Research Environment is going through a process called Airlock, which is where they submit the files that they want to export to our committee, who then analyse it, check that it's in accordance with our rules and it protects our participants' safety and that only then would they allow them to export it.

**Florence:** **Who has access to the Research Environment?**

**Emily:** We have researchers working with the Research Environment all over the world. There's 2 kind of major groups. One of them is academia, so this will be researchers working in universities and academic institutions. The other side of it would is industry - so this will be biotech, startups, pharma companies, things like that.

**Florence: And finally, can you tell us about some of the discoveries that have been made using this data?**

**Emily:** There's lots of really cool things that have come out of the Research Environment. A recent story that came out of the Research Environment was the ReNU syndrome, it was initially just one family that they identified this in, and they were able to extend this discovery across and identify huge numbers of individuals who had this same disorder because they had their genomes within the Research Environment.

**Florence: You can hear more about this research in our previous Behind the Genes podcast: How has a groundbreaking genomic discovery impacted thousands worldwide?**

**So, we'll wrap up there. Thank you so much, Emily, for sharing more about what we mean by the Genomics England Research Environment.**

**If you'd like to hear more explain episodes like this, you can find them on our website, at** **www.genomicsengland.co.uk** **or wherever you get your podcasts.**

**Thank you for listening.**