25 Genomes: Mapping Nature's Code

Information sheet

What is the 25 genomes project?

The Wellcome Sanger Institute is one the world's leading genomics research centres. It was the UK home for the Human Genome Project – which revealed the entire DNA code of life. Since then, that research has started a revolution in healthcare from developing new cancer therapies and better diagnostic tests to online genetic testing services. But scientists are now turning their attention to an even bigger challenge. Every life form has a genome, from plants to animals, bacteria to fungi. To celebrate the 25th Anniversary of the Institute, 25 species will have their genome sequenced. This project will be Sanger Institute's first contribution to the global effort taking place to sequence all life on earth. All genomes sequenced will be novel – never sequenced before and they will be the reference genome for that species. A genome that can be used to compare against other species and individuals within that species.

What is DNA?

DNA or deoxyribonucleic acid is a long double stranded molecule that contains our unique genetic code. DNA contains four basic building blocks or 'bases':
adenine(A), cytosine(C),guanine(G) and thymine(T).

The order, or sequence, of these bases form the instructions in the genome.

What is a genome?

A genome is an organism's complete set of genetic instructions. Each genome contains all of the information needed to build that organism and allow it to grow and develop.

Our bodies are made up of trillions of cells (approx 37,000,000,000,000), each with their own complete set of instructions for making us, like a recipe book for the body. This set of instructions is known as our genome and is made up of DNA and is packaged into Chromosomes in the nucleus. Each cell in the body, for example, a skin cell or a liver cell, contains this same set of instructions:

What is genome sequencing?

DNA sequencing is the process of working out the order of the bases, A, C, G and T, in a strand of DNA. The Sanger Institute uses specialised machines that are able to read all of the DNA bases of an organism's genome. This is what we mean when we say sequence a genome.

At present, we can't sequence a genome, or even a single chromosome, from start to finish. We have to break it up into smaller, more manageable chunks, or fragments. The order and number of bases in these fragments of DNA is then identified through techniques that label each base individually with different colours). From this information, scientists are able to work out the sequence of the DNA and find out lots of other interesting things about that species genetic makeup.

Why do you need a reference genome for a species?

A reference genome gives you the complete list of what is required to make an organism. By comparing reference genomes, you see what makes one species unique - what makes us humans, and what makes chimpanzees, chimpanzees. For an infectious organism or pathogen, a reference genome gives you the complete target list for that organism - every drug or vaccine must by definition target a component found within the reference genome

What can a genome tell us about other species?

Studying the genomes of other species can help us understand how the organism functions, how it compares to other species, and how they have evolved from a common ancestor. It can also help us understand if and how some species are responding and adapting to environmental impacts such as climate change.