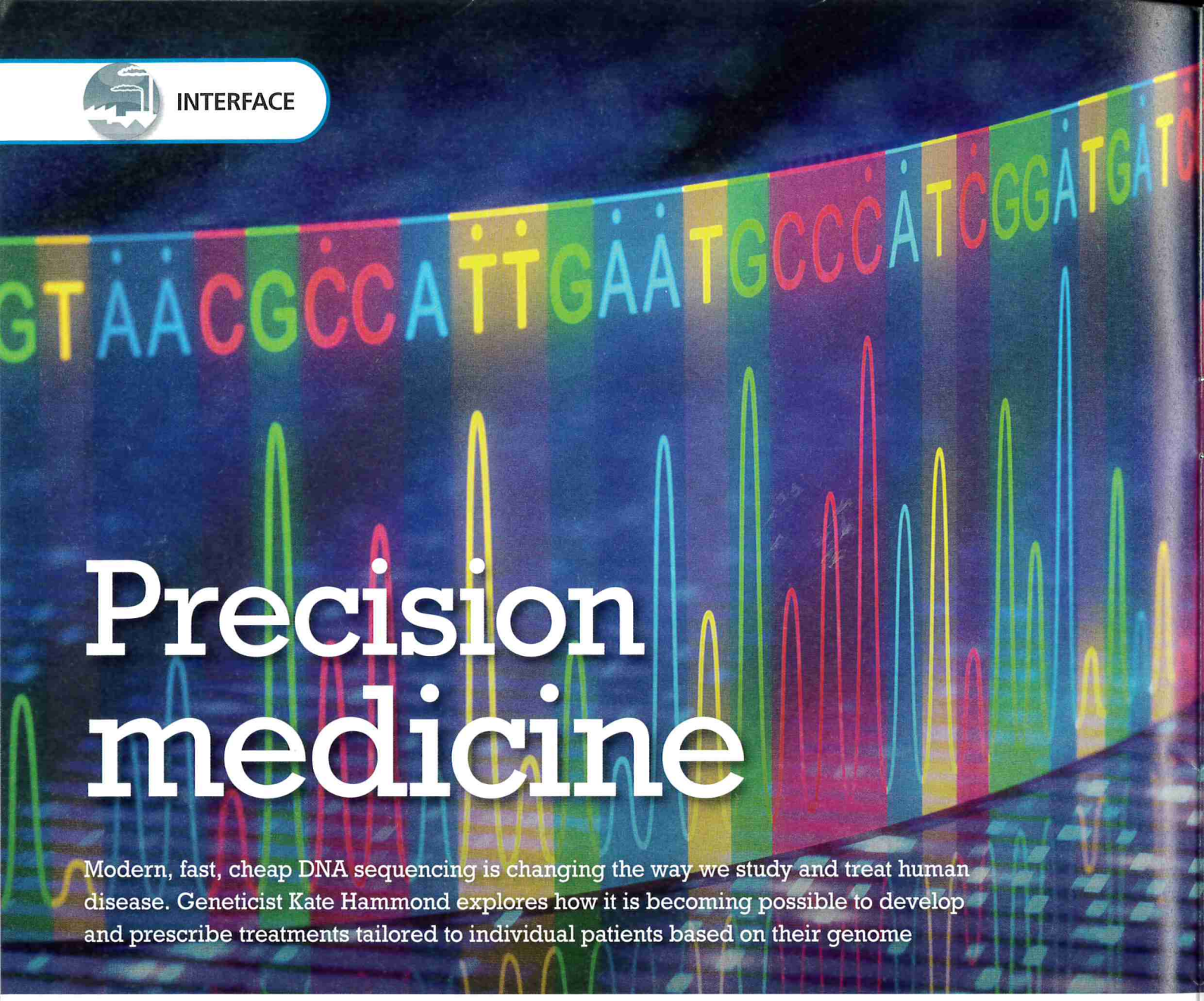


Precision medicine

Modern, fast, cheap DNA sequencing is changing the way we study and treat human disease. Geneticist Kate Hammond explores how it is becoming possible to develop and prescribe treatments tailored to individual patients based on their genome

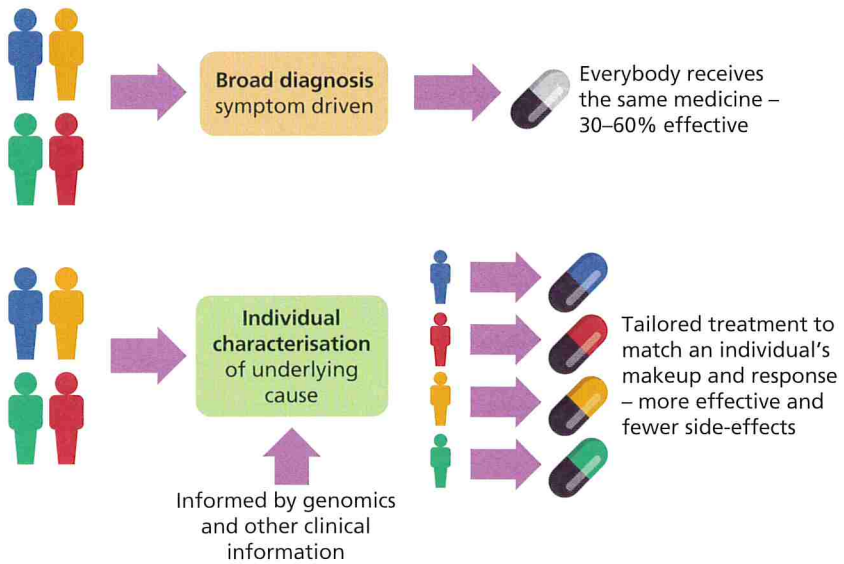
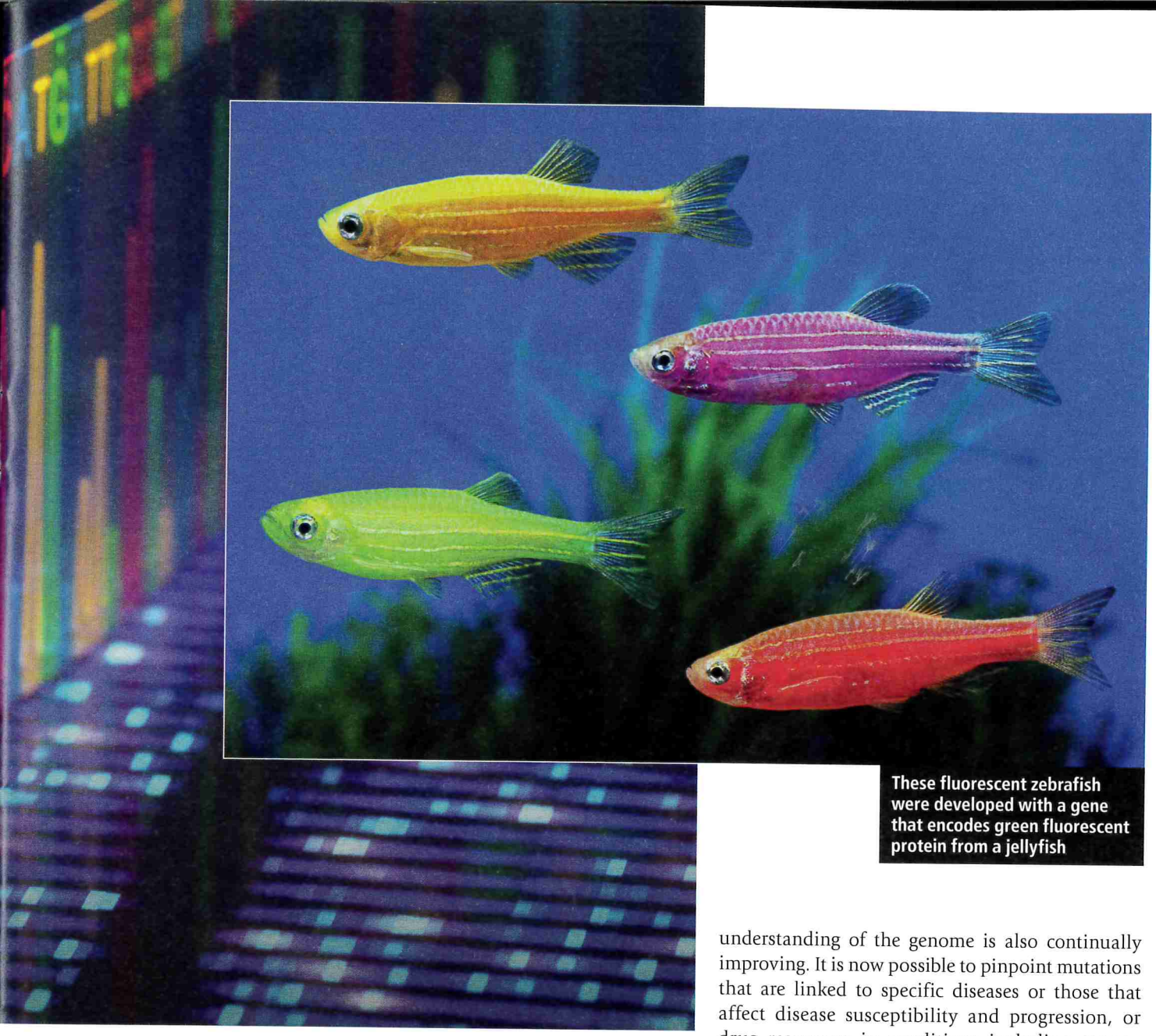


Figure 1 Overview of the concept of precision medicine

In precision medicine, the most appropriate treatment for an individual is selected based partly on the patient's genome, because individual patients can respond very differently to the same treatments. Families and individuals suffering from cancer, neurodegenerative diseases and other rare genetic conditions are set to benefit.

The ideas underlying precision medicine (sometimes called personalised medicine, see Figure 1) are not new. For example, donor tissue for organ transplantation has long been matched as closely as possible to a patient's tissue type. In the past, however, it has been difficult to obtain and make sense of enough information about an individual's genetic makeup for precision medicine to be widely used. Excitingly, this has changed. Modern DNA sequencing methods have become faster and cheaper, giving us access to previously unimaginable amounts of genetic information. Our



These fluorescent zebrafish were developed with a gene that encodes green fluorescent protein from a jellyfish

TERMS EXPLAINED

Codon A sequence of three DNA or RNA nucleotides that code for a specific amino acid or stop signal.

ERK signalling Extracellular receptor that phosphorylates intracellular proteins to stimulate responses from the cell surface membrane.

Exome The protein-coding parts of the genome. Made up of exons.

Exon The parts of a gene that will become part of the mature mRNA and that can encode proteins.

Microarray/chip A collection of short DNA fragments attached to a solid surface.

Oligonucleotide Short strands of nucleic acid (DNA or RNA).

Orthologue A homologous (equivalent) gene found in another species. Evolved from a common ancestral gene during speciation.

Promoter The region of DNA sequence that controls where and when transcription of a gene is initiated. Usually upstream of the gene it controls.

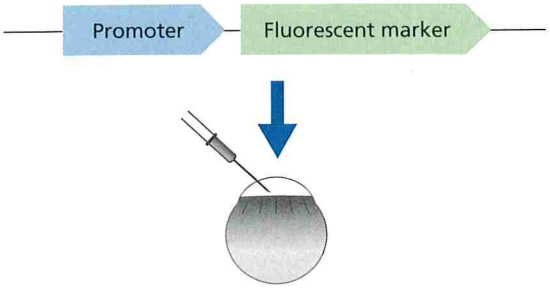
Transgenic DNA sequence(s) from another source have been added to a recipient organism's genome.

understanding of the genome is also continually improving. It is now possible to pinpoint mutations that are linked to specific diseases or those that affect disease susceptibility and progression, or drug responses in conditions including cancer. This information can be used to develop new treatments or guide selection of existing treatments for individual patients.

Precision medicine in the news

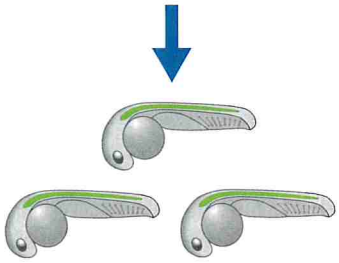
The 100 000 Genomes Project was completed in late 2018 (see Biological Sciences Review, Vol. 29, No. 4, pp. 14–17), sequencing the genomes of 100 000 UK patients and their families. Many of the patients had previously undiagnosed rare diseases. Thanks to the project, in about a quarter of these patients, the mutation underlying their condition has been identified and has already helped identify treatments in many cases.

Several extraordinary demonstrations of the power and potential of precision medicine have been in the news. A beautiful example was reported in 2019 by a team from the Children's Hospital of Philadelphia, USA. This team sequenced the DNA



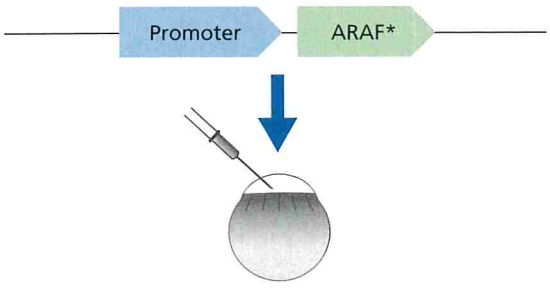
1 DNA coding for a fluorescent marker under the control of the **promoter** for a gene expressed in the lymphatic system (or other tissue of interest) is injected into zebrafish eggs

2 Some of the resulting fish contain the new DNA (a transgene) in the genome of their germ cells (the eggs and sperm). These fish are selected and bred

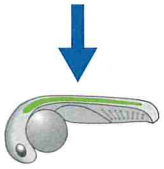


3 The resulting fish (shown as embryos) express the fluorescent marker in the lymphatic system (or other tissue of interest)

Introducing another transgene into the fluorescent line



4 DNA coding for ARAF or mutant ARAF (ARAF*), under the control of a lymphatic system promoter, is injected into eggs from the fluorescent lymphatic system line



5 In some embryos the lymphatic system expresses ARAF* as well as the fluorescent lymphatic system marker. These embryos can be used to investigate the effects of ARAF on the fluorescently marked lymphatic system

Figure 2 Making transgenic fluorescent zebrafish

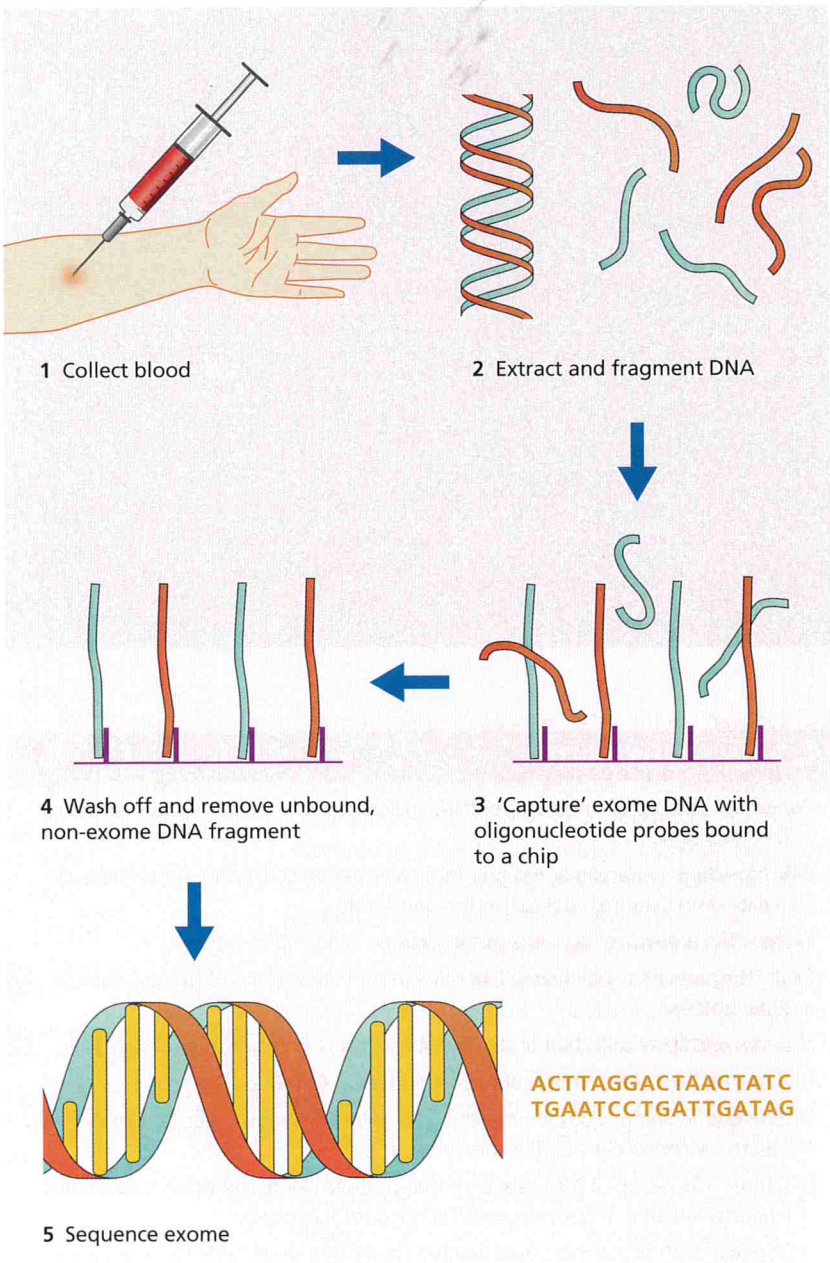


Figure 3 Exome capture

of a 12-year-old patient with a rare, severe lymphatic disease of unknown cause. The lymph glands are part of the immune system, connected by the lymphatic vessels, carrying tissue fluid and cells of the immune system. In this patient, lymph fluid was leaking from the lymphatic system and causing severe swelling in his chest, legs and abdomen, leading to breathing difficulties. DNA sequencing identified a mutation likely to be causing his condition.

Based on what was known about the biology of the protein encoded by the mutated gene, the team designed a potential drug treatment for the boy. Thanks to modern genetic techniques, the team was able to make cell and **transgenic** zebrafish models of the disease, using these to test whether the treatment might be effective, before going on to treat the patient (see Figure 2). A year on, the boy made extraordinary progress, no longer needing supplementary oxygen to support his breathing and having a nearly normal life.

First step: DNA sequencing to genetic diagnosis

Once a patient has been identified for genomic analysis, the first step is to isolate their DNA and sequence it. The cost and speed of sequencing are extremely important if precision medicine is going to be widely used. The first human genome sequence was completed in 2003. It had taken 13 years, cost £3 billion and read approximately 3.3 billion base pairs of DNA. Since then, however, the cost and time taken to sequence a human genome has reduced exponentially and a complete human genome can now be completed for under £1000 in days. This makes individual genome sequencing affordable.

For routine clinical use, to reduce the cost further, it is possible to sequence only the protein-coding regions. These regions come from **exons**, collectively called the **exome**, which makes up just a few percent of the total genome. Nevertheless, it contains the majority of human disease-causing mutations. This was the case for the Philadelphia lymphatic patient.

Exome capture: isolating the protein-coding regions of the genome

To isolate the patient's exome, a DNA sample is fragmented and exposed to short pieces of single-stranded DNA with base sequences that are complementary to each exon in the human exome, termed **oligonucleotide** probes. Fragments of patient DNA containing exons will stick to the corresponding oligonucleotide probes. The oligonucleotide/exons are recovered from the mixture for sequencing, using **microarrays**, or **DNA chips** (see Figure 3).

Sequencing the genome/exome

Whether a whole genome or an exome is to be sequenced, high-throughput modern sequencing technologies are used. These allow many pieces of an individual's DNA to be sequenced at the same time, speeding up the process. The patient's DNA is broken up into short fragments of 100–150 base pairs. These are then sequenced in parallel (see Figure 4). The resulting short sequences are assembled into genome sequences for a patient using the reference human genome as a guide (data obtained originally from 13 individuals). Variations between the patient's DNA sequence and the reference genome and/or control DNA samples from people without the disease can then be identified.

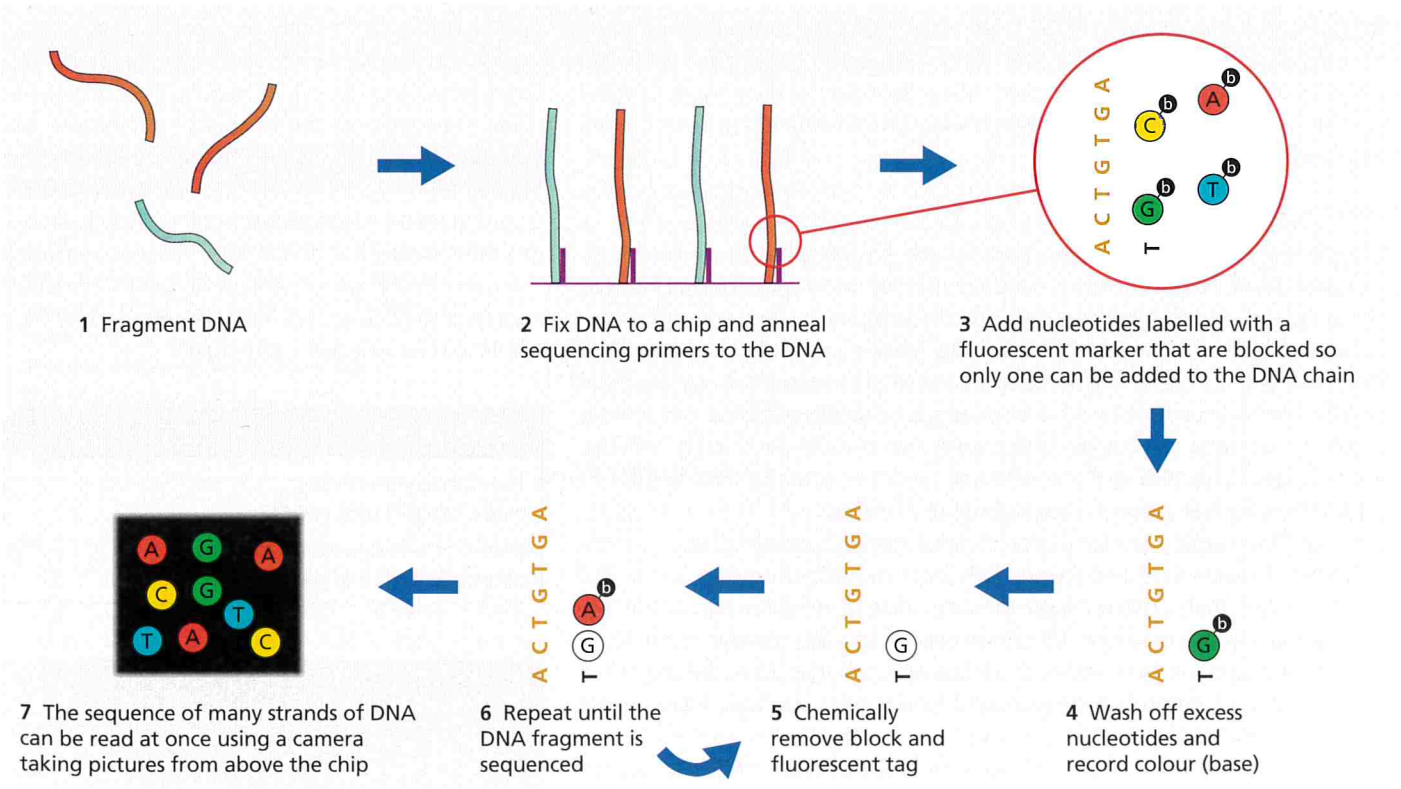
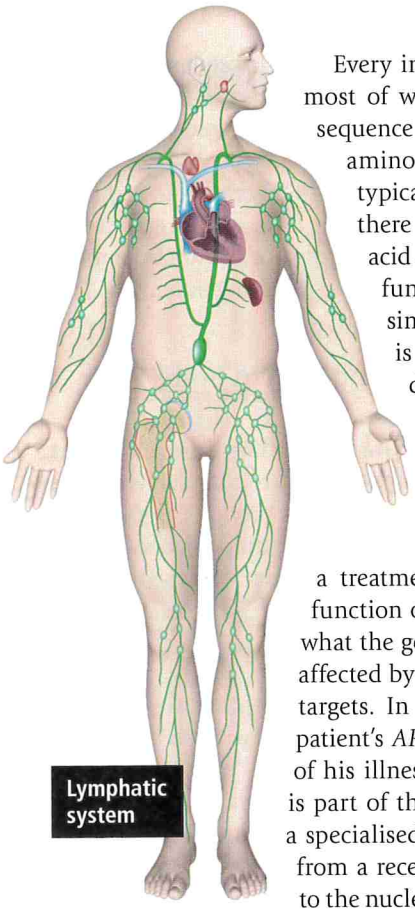


Figure 4 Basics of Illumina sequencing, as used in the 100 000 Genomes Project



Lymphatic system

Every individual will have many genetic variants, most of which are harmless. Variation in the DNA sequence does not always result in variation in the amino acid sequence. First, most amino acids are typically encoded by more than one **codon** – there is redundancy. Second, even if the amino acid sequence is altered, the protein may remain functional if the substituted amino acid has similar physical properties. Careful analysis is therefore required to identify potentially disease-causing mutations.

From mutation to treatment

Once the mutation likely to be causing a patient's condition has been identified, further work is often needed to develop a treatment. In such situations, knowledge of the function of the mutated gene is often vital. Knowing what the gene product normally does and how this is affected by the mutation offers clues for possible drug targets. In the Philadelphia case, a mutation in the patient's *ARAF* gene was identified as the likely cause of his illness. The protein encoded by the *ARAF* gene is part of the **ERK signalling** pathway. This consists of a specialised series of proteins that delivers a stimulus from a receptor on the cell surface membrane of cells to the nucleus. The patient's mutated ARAF protein was found to overactivate the ERK pathway. The medical team hypothesised that inhibiting the ERK pathway would be a good treatment strategy.

Luckily, several ERK pathway inhibitors were already in clinical use. They tested one of these, Trematinib, in a cell culture. The culture contained human lymphatic cells with the mutant *ARAF* gene that was causing over-stimulation. The researchers found that these cells had a reduced amount of cell adhesion molecules between adjacent cells, explaining the cause of the leakage of fluid from the lymph vessels. Encouragingly, Trematinib treatment returned the cell line to normal.

Testing the treatment

Before using Trematinib on their patient, the Philadelphia team needed to confirm its results in an animal model. An animal model was required because cells in culture cannot replicate the complex interactions between cell types that are essential in properly functioning bodies and organ systems. The team chose the zebrafish (*Danio rario*). These small tropical fish are easy and cost-effective to keep and breed, commonly used in biomedicine and similar enough to humans to provide informative pre-clinical data. Like humans, zebrafish are vertebrates and the zebrafish genome contains **orthologues** of about 85% of human genes linked to human diseases.

In the Philadelphia study, the team used transgenic zebrafish in which the lymphatic tissue had been genetically engineered to fluoresce green. The fluorescent lymphatic tissue could then be imaged in detail within living, transparent zebrafish embryos. They were created by adding the gene encoding a fluorescent protein from a jellyfish (*Aquoria victoria*) into the zebrafish genome under the control of the promoter from a gene normally active in the tissue of interest (see Figure 2).

To confirm that the mutated *ARAF* gene was the cause of their patient's illness, the Philadelphia team added a mutated *ARAF* gene into the zebrafish to create double transgenic zebrafish. Zebrafish embryos containing mutant, overactive, ARAF protein developed distended lymphatic ducts, resembling the human condition. The team then treated the zebrafish embryos with an ERK inhibitor. This was straightforward to do because zebrafish embryos develop externally in eggs and are permeable to substances dissolved in the water that they are kept in. The lymphatic abnormalities were reversed and the team now had enough evidence to apply for approval to use Trematinib to treat their patient. This could be done quickly because the drug was already approved for human use for other conditions and it worked extremely well.

Precision medicine: the future

So just how much impact might precision medicine have in the future? For rare diseases it offers the possibility of diagnosis and the development of specific, targeted treatments. For cancer it offers the prospect of tailored treatments specific to both the cancer type and the patient. In all cases, however, precision medicine depends on the infrastructure to perform mass genome sequencing and the ability to analyse a large amount of data being present and cost-effective. In the UK, the newly formed Genomic Medicine Service has been set up to provide this infrastructure. This service will offer genome sequencing and a personalised approach to patients with rare diseases and hard to treat cancers. The aim is to eventually sequence the genomes of 5 million people.

The widespread use of precision medicine also depends on careful consideration of ethical issues and informed consent. What information apart from that directly related to your treatment could your genome reveal about you? Would you want to know? How might this information affect your family? What if it fell into the wrong hands? Could you face discrimination on the basis of your genomic information? Do the benefits outweigh the risks? There are important decisions ahead for scientists and clinicians but also for society as a whole and for affected individuals.

RESOURCES

100 000 Genomes Project:
https://tinyurl.com/yadq7nop

Launch of NHS Genomic Medicine Service:
https://tinyurl.com/ybpkucuw

Dr Kate Hammond is a lecturer in the School of Life Sciences at the University of Liverpool and is programme director for the genetics BSc. Her research interest is in zebrafish genetics.