Few technologies have transformed the field of biology as profoundly as sequencing – the ability to decipher the sequence of base pairs in a fragment of DNA.
The first two human genomes were sequenced over 20 years ago, and since then, the genomes of hundreds of thousands of different individuals have been deciphered. The genes that make up our genomes exist in many variants, called alleles, and each of us has a unique set of such alleles. The point of sequencing more and more human genomes is not just about scientifically explaining the differences between people, the origins of our common traits, or identifying which genes encode, say, our height or intelligence – it’s also about better understanding the genetic basis of many diseases, which may lead to new biomedical applications. Such knowledge can be useful both in diagnosis and in the search for new treatment methods. Genome sequencing also holds the promise of giving rise to personalized medicine, in which treatments are specifically tailored to individual patients, including their individual genomes.

However, sequencing has far more applications than just reading out an individual’s DNA. Apart from the alleles we inherit from our parents and any new mutations, which are present in all our cells, it turns out that new DNA variants arise within our bodies all the time. Firstly, the adaptive immune response that occurs in all vertebrates involves the creation of new antibodies through recombination and selection. As a result, the genomes of cells producing antibodies differ from the genomes of our other cells. Sequencing helps us better understand the mechanisms involved in developing immunity to infections, as well as in autoimmune diseases. Secondly, new, evolutionarily unintended mutations continually occur in the process of cell division in our bodies. These mutations are the cause of cancer and sequencing the genome of cancer cells makes it possible to identify key mutations and select effective therapies.

Gene expression

Another application of sequencing is transcriptomics. Each of our genes undergoes a process of expression, meaning that it is transcribed from DNA to RNA, and if it encodes a protein sequence, it is then translated into a protein. Our cells respond to their environment and communicate with each other largely through the regulation of gene expression: “switching on” or “switching off” the transcription of genes into RNA depending on the context. For example, if the organism has become infected by a virus, a cell may receive a signal stimulating it to produce proteins that inhibit viral infections. Sequencing the transcriptome allows us to detect and characterize the immune response. In recent years, methods that allow us to study the transcriptomes even of individual cells (single-cell RNA-seq, or sc-RNA-seq) have become widespread.

DNA sequencing is one of the fundamental technologies for studying organisms and cells at the molecular level. Unlike other methods, such as those based on antibodies or mass spectrometry, it makes it possible to analyze a vast number of samples in an incredibly short time. Numerous new DNA sequencing methods emerged at the beginning of the twenty-first
The point of sequencing human genomes is not just about scientifically explaining the differences between people, the origins of our common traits, but also better understanding the genetic basis of many diseases, leading to new biomedical applications.
that can be used to monitor changes in the processes occurring in living cells, and so it has been routinely used for well over two decades. It has found a great number of applications in almost all fields of biological and medical science. It allows for precise tracking of changes in gene expression in entire organisms, tissues, selected cell types, and even individual cells. We are now observing another breakthrough in this technology, known as spatial transcriptomics.

Microscopic methods have long made it possible to study the processes underway inside our tissues. By visualizing individual genes and proteins, we can determine where in the tissue or cell the expression of a specific gene occurs. However, current technological advancements have enabled transcriptomics (high-throughput analysis of the expression of many genes simultaneously) to be combined with microscopic visualization of tissue organization. This makes it possible to determine not only which cells regulate gene expression and how they do so, but also where in the examined tissue they are located and how gene expression changes at the boundary between healthy and diseased tissue.

There are roughly two approaches to spatial transcriptomics. The first type of methods is based on the preparation of special glass slides, where each fragment of the surface is specially marked with short DNA sequences (oligonucleotides). A tissue preparation is applied to the slide. When RNA from the cells of the preparation is synthesized into cDNA, oligonucleotides attach to it. Then, cDNA from the entire preparation is harvested and sequenced in the usual way. Labelling makes it possible to later recognize from which place on the slide a particular sequence came and to reconstruct the spatial structure of the preparation, which can then be overlaid onto microscopic images. The advantage of these methods is a relatively large “depth” – the number of different RNA molecules that can be identified through them. The drawback is relatively low resolution.

The second type of method is based on in situ detection, similar to what has been done for many decades, only on a larger scale. The preparation is labeled with fluorescent probes corresponding to different mRNA sequences. This often involves marking short, characteristic DNA sequences with fluorescent markers. Such probes bind specifically to mRNAs, which can then be directly visualized on the preparation. There are many variations in this family of spatial transcriptomic methods, but most of them only allow for the detection of a limited repertoire of genes defined a priori.

Progress in developing new methods in biology and medicine takes years, even decades. Methods that were primarily experimental and rarely appeared in top-tier journals ten years ago are now being refined and routinely used in laboratories worldwide, and their costs, while not trivial, are no longer prohibitive. The methods described in this article – long-read sequencing and spatial transcriptomics – are now entering this second stage. Their practical importance for both fundamental science and biomedicine can be expected to increase significantly in the coming years.