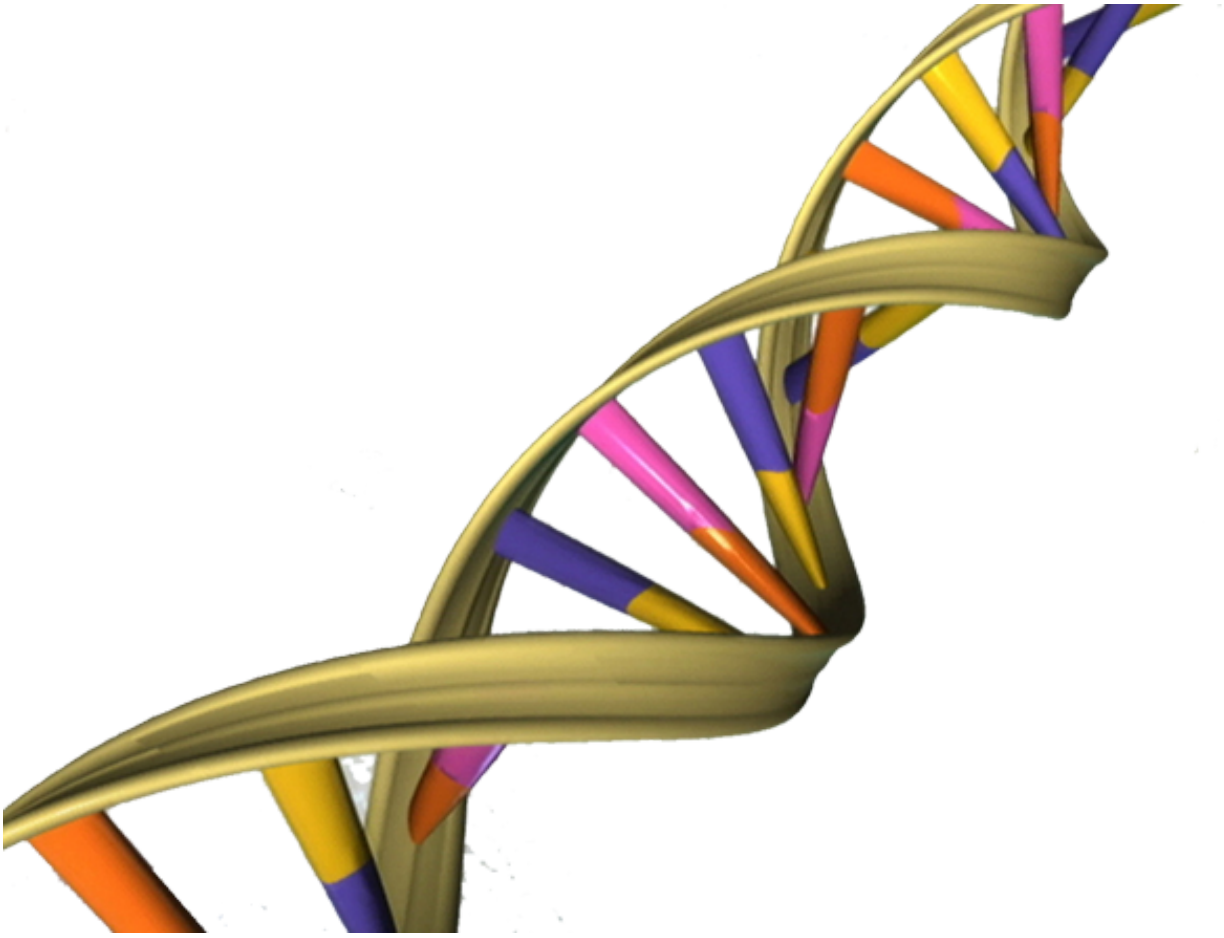
Zipping DNA

November 13 2017



A depiction of the double helical structure of DNA. Its four coding units (A, T, C, G) are color-coded in pink, orange, purple and yellow. Credit: NHGRI

ETH researchers have developed a method that allows large amounts of genetic information to be compressed and then decompressed again in

cells. This could aid in the development of new therapies.

What do you do if you have a large document or a high-resolution image that is too big to send via email? You simply zip it to a more manageable size using a suitable software. "Instead of sending the information 'white-white-white-white-white-...' for every single pixel on a white line, only the message 'white 1,000 times' is transmitted," explains Kobi Benenson, Head of the Synthetic Biology Group at ETH's Department of Biosystems Science and Engineering in Basel. Once received, the information can then be returned to its original size, i.e. unzipped.

Limited transport capacity

This method for digital files inspired Benenson and his colleague Nicolas Lapique to develop an innovative solution for biological systems. They worked out a method that could be used to zip the genetic material DNA: it is compressed for transport into cells and then assembled into functioning genetic information once inside the cell.

This type of solution could be useful for biologists, particularly for synthetic biology or biotechnology, as the scientists are limited when trying to implant large amounts of information into cells in the form of DNA. The problem is that the transport vehicles that are currently used for this purpose can only be loaded with a limited amount of DNA.

Removing repetitions in DNA

The basic principle behind this innovative DNA compression is the same as the principle behind zipping a digital file: "Elements that come up often in the DNA sequence to be implanted are only transmitted once," explains Benenson.

For example, this could apply to promoters - sections of the DNA that regulate if and how the associated gene is read. If the DNA to be transported into the cell contains four different genes that all have the same promoter, it will only be included once.

Tightly packaged and reassembled at the destination

Removing redundancies isn't everything, however. The ETH researchers assemble the DNA to be transported into the cell in accordance with specific rules. Benenson talks about "compressed encoding".

The four genes in our example first receive a joint promoter. The researchers string together the four coding gene sequences compactly on the DNA double strand. They equip it all with individual stop sequences and - importantly - different binding sites for a recombinase, an enzyme that can open, rotate, and reassemble DNA strands.

"The recombinase takes on the role of the decompression software," explains Benenson. It ensures that the components of the compressed DNA are assembled in working order inside the cell. For the four example genes, this means that each will receive their own promoter once reassembled.

Genetic programmes detect tumour cells

Benenson and Lapique were able to demonstrate that this new method actually allows large "genetic programmes" to be implanted in mammalian cells. "These are man-made and carry out specific tasks within the cells," explains Benenson. In other words, they comprise a whole arsenal of biological components such as proteins and RNA that work within the cell in a coordinated manner to achieve a goal defined by the scientists. In biotechnology, this method would allow the creation

of certain complex substances such as active ingredients for medicines.

Benenson's group, however, is working on genetic programmes that will hopefully master much more complicated tasks in the future. One such task is cancer targeting, which means that the programme can detect specific substances, the markers, in a cell. Depending on the concentration, it decides whether the cell is healthy or whether it is a tumour cell- which the programme would then be able to independently eliminate. It would be a kind of all-in-one solution for combating tumours that would cover the examination, diagnosis and even treatment. This approach has been shown to work in cell cultures and the researcher would like to test it in an animal model as well.

More accurate diagnoses thanks to new methods

With currently available DNA delivery vehicles, the accuracy of deciding whether a cell is healthy or cancerous is still not high enough, as not enough different markers can be applied at once due to the limited amount of DNA that can be transmitted.

"A combination of four to six markers would be ideal," explains Benenson. In order to detect all these, however, the corresponding number of sensors is needed to recognise the marker substances. More sensors - this involves proteins, RNA, and DNA components - would also mean more DNA that must be implanted in the cell as the sensors' blueprint. They now hope that a programme can use this new DNA compressing and decompressing method to implement additional sensors and thus increase the accuracy rate.

Borrowing from information technology

It's no coincidence that the genetic programmes developed by Benenson

and Lapique are structured logically and function in a similar way to computer programmes. "Our research is often inspired by computer science and information technology," explains Benenson. He very clearly enjoys thinking outside the box. When it comes to the new DNA transport methods, it's safe to say: it's lucky that email attachments have size restrictions.

The study is published in *Nature Nanotechnology*.

More information: Genetic programs can be compressed and autonomously decompressed in live cells, *Nature Nanotechnology*, doi:10.1038/s41565-017-0004-z , www.nature.com/articles/s41565-017-0004-z

Provided by ETH Zurich