Recording and decoding short digital text messages with DNA Typewriter. a, Base64 binary-to-text was modified to assign 64 NNNGGA barcodes for TAPE-1 to 64 text characters. b, Illustration of the encoding strategy for “WHAT HATH GOD WROUGHT?”, which has 22 characters including whitespaces. The message is grouped into sets of four characters, converted to NNN barcodes according to the TAPE64 encoding table, and plasmids corresponding to each set are mixed at a ratio of 7:5:3:1 for transfection. To encode 22 characters, we sequentially transfected 5 sets of 4 characters and 1 set
of 2 characters 3 days apart into PE2(+) 5×TAPE-1(+) HEK293T cells. c–e, Decoding of three messages based on sequencing of the following 5×TAPE-1 arrays: “WHAT HATH GOD WROUGHT?” (c), “MR. WATSON, COME HERE!” (d) and “BOUND FOREVER, DNA” (e). For each message, the full set of NNNGGA insertions was first identified and cotransfected sets of characters were then identified from the bigram transition matrix (left). Within each set of characters inferred to have been cotransfected, ordering was based on corrected unigram counts (middle), resulting in the final decoded message (right). Misordered characters within each recovered message are colored purple, missing characters are colored red with strikethrough, and unintended characters are colored light blue. Both two-dimensional histogram and corrected read counts were calculated by combining sequencing reads over n = 3 independent transfection experiments. Read counts were corrected using the edit score for each insertion barcode. Credit: Nature (2022). DOI: 10.1038/s41586-022-04922-8

While developing a new system for recording within cells, geneticist Jay Shendure and his team decided to give it a test run by using it to encode text. Since their invention relied on a nearly brand-new recording medium, DNA, they wanted to use messages that evoked a sense of historical significance.

Two choices were obvious: "What hath God wrought?" a Biblical quote used by Samuel Morse in the first long-distance telegraph transmission, and the more mundane, "Mr. Watson, come here!" spoken by Alexander Graham Bell to his assistant in the first telephone call.

Postdoctoral fellow Junhong Choi needed one more phrase for the experiments. Shendure wanted the third quote to encapsulate the excitement of working on this system, which they later named the DNA typewriter, while the COVID-19 pandemic raged around them.
"I wanted Choi to use 'It was the best of times, it was the worst of times,'" says Shendure, a Howard Hughes Medical Institute (HHMI) Investigator at the University of Washington.

Instead of the well-known line from Charles Dickens, Choi, who is Korean, had another idea: A lyric from a Korean pop or K-pop song. So they rounded out their collection with "Bound forever, DNA" a lyric translated from a song by the group BTS.

While the DNA typewriter proved it could tuck these phrases like secret messages into cells, its most powerful potential lies elsewhere. Shendure and his colleagues envision scientists using this technology to document the inner workings of cells as they unfold far from view.

A naturally digital medium

The idea to use DNA to encode information piggybacks on the molecule's natural function. Just as computer software is written in 1s and 0s, the DNA molecule uses a four-letter code for the instructions for living things.

The DNA typewriter isn't the first of its kind, however, it has a combination of attributes that make it particularly promising for illuminating the biology of cells. Namely, it can capture a large number of events while documenting them in chronological order.

"We have accomplished something that's analogous to writing," Shendure says. "We can create thousands of symbols, which we call barcodes, and we can capture them in order."

The researchers have so far adapted the system to accommodate as many as 4,096 barcodes, which are short pieces of DNA. Like the original machine tapping out letters, the DNA typewriter lays down only one
barcode at a time, from left to right.

While such a massive number of symbols isn't necessary for typing out a novel, it better suits the immense complexity of activity within a cell. With this hefty recording capacity, the DNA typewriter promises scientists a new way to examine living cells—an option beyond the typical choice of using a microscope or destroying cells in order to analyze them, Shendure says.

A new window into cells

Leonard Zon, an HHMI Investigator at Boston Children's Hospital and Harvard Medical School, sees plenty of potential. "It's a fascinating technology that could have great applications in the future," he says of the DNA typewriter.

For instance, Zon, who was not involved in developing this system, envisions using a version of it to monitor a cancer patient's response as he or she receives successive doses of chemotherapy over a treatment period, he says. "You would essentially get a history of how the cancer cell responded to the injuries inflicted by the chemotherapy."

Likewise, scientists could use it to examine poorly-understood aspects of development, such as following the events that occur within an early embryo's innermost layer of cells as it transforms into the gut, he says.

From typewriter to flight recorder?

In another experiment, Shendure's team turned their recorder on the cells themselves. By tagging dividing cells with barcodes, they tracked cell divisions. After 25 days, in which one cell gave rise to about 1.2 million, the researchers analyzed the patterns in the cells' barcodes to reconstruct their family tree.
While it tracked the history of the population of cells, this lineage experiment did not capture the processes going on within them, such as the damage or the developmental steps Zon mentions. Other research from Shendure's group delineates this additional step.

In a study published in *Nature*, his team describes a strategy for genetically engineering cells so that the genetic code for a molecule that interests scientists carries with it the instructions for a barcode. When the molecule turns on within a cell, it switches on the barcode too. Then the typewriter goes to work, documenting the molecule's activity.

With some additional engineering, Shendure and his team hope to turn the typewriter into a black box-like system. By tapping into new biological events, improving the system's efficiency, and making other improvements, he wants to build a tool scientists can use to recover a detailed history of a cell, just like investigators looking into an airplane crash.

"There's a lot of potential here, once we take it to the next level," he says.


Wei Chen et al, Multiplex genomic recording of enhancer and signal transduction activity in mammalian cells (2021). DOI: 10.1101/2021.11.05.467434

Provided by Howard Hughes Medical Institute