

[The Scientist](#) » [News & Opinion](#) » [News](#)

DNA-based Data Storage Here to Stay

The second example of storing digital data in DNA affirms its potential as a long-term storage medium.

By Sabrina Richards | January 23, 2013

0 Comments



stock.xchng, schulergd

Researchers have done it again—encoding 5.2 million bits of digital data in strings of DNA and demonstrating the feasibility of using DNA as a long-term, data-dense storage medium for massive amounts of information. In the new study released today (January 23) in *Nature*, researchers encoded one color photograph, 26 seconds of Martin Luther King Jr.'s "I Have a Dream" speech, and all 154 of Shakespeare's known sonnets into DNA.

Though it's not the first example of storing digital data in DNA, "it's important to celebrate the emergence of a field," said [George Church](#), the

Harvard University synthetic biologist whose own group published a similar demonstration of DNA-based data storage last year in *Science*. The new study, he said, "is moving things forward."

Scientists have long recognized DNA's potential as a long-term storage medium. "DNA is a very, very dense piece of information storage," explained study author [Ewan Birney](#) of the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) in the UK. "It's very light, it's very small." Under the correct storage conditions—dry, dark and cold—DNA easily withstands degradation, he said.

Advances in synthesizing defined strings of DNA, and sequencing them to extract information, have finally made DNA-based information storage a real possibility. Last summer, Church's group published the first demonstration of DNA's storage capability, encoding the digital version of Church's book *Regenesis*, which included 11 JPEG images, into DNA, using Gs and Cs to represent 1s of the binary code, and As and Ts to represent 0s.

Now, Birney and his colleagues are looking to reduce the error associated with DNA storage. When a strand of DNA has a run of identical bases, it's difficult for next-generation sequencing technology to correctly read the sequence. Church's work, for example, produced 10 errors out of 5.2 million bits. To prevent these types of errors, Birney and his EMBL-EBI collaborator [Nick Goldman](#) first converted each byte—a string of eight 0s and 1s—into a single "trit" made up of 5 or 6 digits of 0s, 1s, and 2s. Then, when converting these trits into the A, G, T and C bases of DNA, the researchers avoided repeating bases by using a code that took the preceding base into account when determining which base would represent the next digit.

The synthesizing process also introduces error, placing a wrong base for every 500 correct ones. To reduce this type of error, the researchers synthesized overlapping stretches of 117 nucleotides (nt), each of which overlapped with preceding and following strands, such that all data points were encoded four times. This effectively eliminated reading error because the likelihood that all four strings have identical synthesis errors is negligible, explained Birney.

Agilent Technologies in California synthesized more than 1 million copies of each 117-nt stretch of DNA, stored them as dried powder, and shipped it at room temperature from the United States to Germany via the UK. There, researchers took an aliquot of the sample, sequenced it using next-generation sequencing technology, and reconstructed the files.

Birney and Goldman envision DNA replacing other long-term archival methods, such as magnetic tape drives. Unlike other data storage systems, which are vulnerable to technological obsolescence, "methods for writing and reading DNA are going to be around for a long, long time," said molecular biologist [Thomas Bontin](#) of the University of Copenhagen. Bontin, who was not involved in the research, compared DNA information storage to the fleeting heyday of the floppy disk—introduced only a few decades ago and already close to unreadable. And though synthesizing and decoding DNA is currently still expensive, it is cheap to store. So for data that are intended to be stored for hundreds or even thousands of years, Goldman and Birney reckon that DNA could actually be cheaper than tape.

Additionally, there's great potential to scale up from the 739 kilobytes encoded in the current study. The researchers calculate that 1 gram of DNA could hold more than 2 million megabytes of information, though encoding information on this scale will involve reducing the synthesis error rate even further, said bioengineer [Mihri Ozkan](#) at the University of California, Riverside, who did not participate in the research.

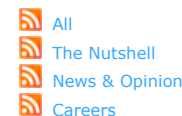
Despite the challenges that lie ahead, however, the current advance is "definitely worth attention," synthetic biologist [Drew Endy](#) at Stanford University, who was not involved in the research, wrote in an email to *The Scientist*. "It should develop into a new option for archival data storage, wherein DNA is not thought of as a biological molecule, but as a straightforward non-living data storage tape."

N. Goldman et al., "Towards practical, high-capacity, low-maintenance information storage in synthesized DNA," *Nature*, doi: 10.1038/nature.11875, 2013.

Follow The Scientist



Subscribe to RSS feed



Popular Posts

1. [DNA Reveals Ancient Looks](#)
2. [Top Science Scandals of 2012](#)
3. [Dung Defeats Drugs](#)
4. [Does Pot Cause Mental Decline?](#)
5. [Universal Flu Vaccines Charge Ahead](#)

Current Issue



[View the January 2013 contents.](#)

SUBSCRIBE
TheScientist

Tags

[next generation sequencing](#), [DNA-based Data Storage](#), [DNA synthesis](#) and [DNA](#)

0 Comments

Add a Comment

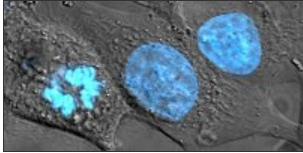


You

[Sign In](#) with your LabX Media Group Passport to leave a comment

Not a member? [Register Now!](#)

Related Articles



Double DNA

By Bob Grant

Scientists provide evidence for the existence of four-stranded human DNA, which has previously only been theorized and synthesized.



Fly Guts Reveal Animal Inventory

By Beth Marie Mole

Stomachs of flesh-eating flies carry the DNA of animals in remote rainforests.



DNA for the Masses

By Edyta Zielinska

An Israeli company plans to synthesize DNA on demand for anyone, including those outside of research labs.

TheScientist

[Home](#) [News & Opinion](#) [The Nutshell](#) [Multimedia](#) [Magazine](#) [Advertise](#)
[About & Contact](#) [Privacy Policy](#) [Job Listings](#) [Subscribe](#) [Archive](#)



© 1986-2013 The Scientist

Now Part of the LabX Media Group: [Lab Manager Magazine](#) | [LabX](#) | [LabWrench](#)