

Why DNA Will Someday Replace the Hard Drive

Researchers who have used the biomolecule to encode MP3s, text files, and JPEGs say it will be a competitive storage medium in just a few decades.

By [Katherine Bourzac](#) on January 23, 2013

DNA could someday store more than just the blueprints for life – it could also house vast collections of documents, music, or video in an impossibly compact format that lasts for thousands of years.

Researchers at the [European Bioinformatics Institute](#) in Hinxton, U.K., have demonstrated a new method for reliably encoding several common computer file formats this way. As the price of sequencing and synthesizing DNA continues to drop, the researchers estimate, this biological storage medium will be competitive within the next few decades.

The information storage density of DNA is at least a thousand times greater than that of existing media, but until recently the cost of DNA synthesis was too high for the technology to be anything more than a curiosity. Conventional methods of storing digital information for prolonged periods continue to pose problems, however. The magnetic tapes typically used for archival storage become brittle and lose their coating after a few decades. And even if the physical medium used to store information remains intact, storage formats are always changing. This means the data has to be transferred to a new format or it may become unreadable.

DNA, in contrast, remains stable over time – and it's one format that's always likely to be useful. "We want to separate the storage medium from the machine that will read it," says project leader [Nick Goldman](#). "We will always have technologies for reading DNA." Goldman notes that intact DNA fragments tens of thousands of years old have been found and that DNA is stable for even longer if it's refrigerated or frozen.

The U.K. researchers encoded DNA with an MP3 of Martin Luther King Jr.'s "I Have a Dream" speech, a PDF of a scientific paper, an ASCII text file of Shakespeare's sonnets, and a JPEG color photograph. The storage density of the DNA files is about 2.2 petabytes per gram.

Others have demonstrated DNA data storage before. This summer, for example, researchers led by Harvard University genetics professor [George Church](#) used the technology to encode a book (see "[An](#)

Entire Book Stored in DNA”).

The difference with the new work, says Goldman, is that the researchers focused on a practical, error-tolerant design. To make the DNA files, the researchers created software that converted the 1s and 0s of the digital realm into the genetic alphabet of DNA bases, labeled A, T, G, and C. The program ensures that there are no repeated bases such as “AA” or “GG,” which lead to higher error rates when synthesizing and sequencing DNA.

The files were divided into segments, each bookended with an index code that contains information about which file it belongs to and where it belongs within that file – analogous to the title and page number on pages of a book.

The encoding software also ensures some redundancy. Each part of a file is represented in four different fragments, so even if several degrade, it should still be possible to reconstruct the data.

Working with genomics company Agilent Technologies of Santa Clara, California, the researchers synthesized the fragments of DNA and then demonstrated that they could sequence them and accurately reconstruct the files. This work is described today in the journal Nature.

Goldman’s group estimates that encoding data in DNA currently costs \$12,400 per megabyte, plus \$220 per megabyte to read that data back. If the price of DNA synthesis comes down by two orders of magnitude, as it is expected to do in the next decade, says Goldman, DNA data storage will soon cost less than archiving data on magnetic tapes.

Victor Zhirnov, program director for memory technologies at the Semiconductor Research Corporation in Durham, North Carolina, says that because the current cost is so high, data-storing DNA will probably find its earliest use in long-term archives that aren’t accessed frequently. Looking ahead, he says, he can envision “a more aggressive technology” to replace flash, the nonvolatile memory technology found in portable electronics, which is already reaching its scaling limits. The key will be developing entire hardware systems that work with DNA, not just sequencers and synthesizers.

Harvard’s Church says he is working on this very problem. “We can keep incrementally improving our ability to read and write DNA, but I want to jump completely out of that box,” he says. Church is currently developing a system for directly encoding analog signals such as video and audio into DNA, eliminating conventional electronics altogether.



Katherine Bourzac Contributor

I’m a freelance journalist based in San Francisco. Before going freelance, I was *MIT Technology Review*’s material science editor; and I graduated from MIT’s Science Writing program in 2004.

[About Katherine »](#)

