

Computer files stored accurately on DNA in new breakthrough

Scientists have recorded data including Shakespearean sonnets and an MP3 file on strands of DNA, in a breakthrough which could see millions of records stored on a handful of molecules rather than computer drives.



The advantage of using DNA over hard drives is that it does not require a constant supply of electricity Photo: Alamy

By Nick Collins, Science Correspondent

6:00PM GMT 23 Jan 2013

By translating computerised files into DNA similar to that found in plants and animals, the researchers claim it is possible to store a billion books' worth of data for thousands of years in just a small test tube.

Although the method is expensive, it could still be much more efficient than hard drives or magnetic tape for long-term storage of large sets of data such as government records, the scientists said.

Within a decade, they expect the technique to have become cheap enough that DNA storage could become cost-efficient for the public to store lifelong keepsakes like wedding videos.

Dr Nick Goldman of the European Bioinformatics Institute, who led the study, said: "We already know that DNA is a robust way to store information because we can extract it from bones of woolly mammoths, which date back tens of thousands of years, and make sense of it.

"It's also incredibly small, dense and does not need any power for storage, so shipping and keeping it is easy."

The advantage of using DNA over hard drives is that it does not require a constant supply of electricity, while "no-power" archiving materials such as magnetic tape degrade within a decade.

Scientists have long been able to read DNA, a code comprising four "letters", but using it to store information has been difficult because it is prone to decoding errors when the same letter is repeated.

In a study published in the *Nature* journal, the researchers demonstrated they could avoid the problem by translating computer files, made up of ones and zeroes, into a form of DNA code which did not allow letters to repeat themselves.

First they converted an audio file of Martin Luther King's "I have a dream" speech, a photograph of their laboratory, a PDF file of an academic paper and a text version of all Shakespeare's sonnets into the DNA code.

The code was sent to a US lab where experts converted it into synthetic strings of DNA which resembled a tiny grain of dust.

The researchers then sequenced the synthetic DNA to retrieve the code, before converting it back into the original computer files with 99.9 per cent accuracy.

Dr Goldman said: "Because it is expensive and one of its big advantages is longevity, the potential applications will initially be in really high value information which you are determined to keep safe but you do not expect to read very often, such as government records or the Doomsday Book.

"As the price starts to come down it will start to become available to people with smaller budgets, so in ten years' time it may be [cost efficient for] something you would look at on a 50 year timescale, such as a wedding video."