MEDICAL SCIENCE MONITOR

EDITORIAL

e-ISSN 1643-3750
© Med Sci Monit, 2018; 24: 1185-1187
DOI: 10.12659/MSM.908313

Received: 2017.11.30
Accepted: 2017.04.12
Published: 2018.02.26

DNA MemoChip: Long-Term and High Capacity Information Storage and Select Retrieval

[1,2] George B. Stefano
[3,4] Fuzhou Wang
[1,2] Richard M. Kream

1 Department of Psychiatry, Charles University, Center for Molecular and Cognitive Neuroscience, Prague, Czech Republic
2 International Scientific Information, Melville, NY, U.S.A.
3 Group of Neuropharmacology and Neurophysiology, Division of Neuroscience, The Bonoi Academy of Science and Education, Chapel Hill, NC, USA
4 Department of Anesthesiology, Obstetrics and Gynecology Hospital, Affiliated to Nanjing Medical University, Nanjing, Jiangsu, P.R. China

Corresponding Author: George B. Stefano, e-mail: gbstefano@yahoo.com
Source of support: In part, ISI Corp.

Over the course of history, human beings have never stopped seeking effective methods for information storage. From rocks to paper, and through the past several decades of using computer disks, USB sticks, and on to the thin silicon "chips" and "cloud" storage of today, it would seem that we have reached an era of efficiency for managing innumerable and ever-expanding data. Astonishingly, when tracing this technological path, one realizes that our ancient methods of informational storage far outlast paper (10,000 *vs.* 1,000 years, respectively), let alone the computer-based memory devices that only last, on average, 5 to 25 years. During this time of fast-paced information generation, it becomes increasingly difficult for current storage methods to retain such massive amounts of data, and to maintain appropriate speeds with which to retrieve it, especially when in demand by a large number of users. Others have proposed that DNA-based information storage provides a way forward for information retention as a result of its temporal stability. It is now evident that DNA represents a potentially economical and sustainable mechanism for storing information, as demonstrated by its decoding from a 700,000 year-old horse genome. The fact that the human genome is present in a cell, containing also the varied mitochondrial genome, indicates DNA's great potential for large data storage in a 'smaller' space.

MeSH Keywords: **Information Storage and Retrieval • Humans • Genome, Human • Genome, Mitochondrial • Computers • Organizational Innovation • DNA**

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/908313

957 — — 8

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Over the course of history, human beings have never stopped seeking effective methods for information storage. From rocks to paper, and through the past several decades of using computer disks, USB sticks, and on to the thin silicon "chips" and "cloud" storage of today [1,2], it would seem that we have reached an era of efficiency for managing innumerable and ever-expanding data. Astonishingly, when tracing this technological path, one realizes that our ancient methods of informational storage far outlast paper (10,000 *vs.* 1,000 years, respectively), let alone the computer-based memory devices that only last, on average, 5 to 25 years. During this time of fast-paced information generation, it becomes increasingly difficult for current storage methods to retain such massive amounts of data, and to maintain appropriate speeds with which to retrieve it, especially when in demand by a large number of users.

In this light, our methodological advancements, which become *de facto* outdated and demand newer technologies at an increasing pace, make for poor informational retention/storage, as well as select retrieval. Moreover, the current level of information storage demands greater capacity as time passes. As Extance notes, "counting everything from astronomical images and journal articles to YouTube videos, the global digital archive will hit an estimated 44 trillion gigabytes by 2020, a tenfold increase over 2013. By 2040, if everything were stored for instant access flash memory chips used in memory sticks, the archive would consume 10–100 times the expected supply of microchip-grade silicon" [3]. As such, there must be another method to stabilize and economize information storage and retrieval, a potential mechanism that may have been staring us in the face all along.

Recently, Goldman and colleagues proposed that DNA-based information storage provides a way forward for information retention as a result of its temporal stability [4]. Indeed, when learning from the past, it is evident that DNA represents a potentially economical and sustainable mechanism for storing information, as demonstrated by its decoding from a 700,000 year-old horse genome [5]. The fact that the human genome is present in a cell, containing also the varied mitochondrial genome, indicates DNA's great potential for large data storage in a 'smaller' space while maintaining a high degree of sophisticated modulation [6]. The accuracy of biogenetics is the basis of determining bio-similarity in its kind, of which DNA sequence is the core of the process. In contrast, the variety of the biogenetic-related DNA code determines the biodiversity (Dobson & Carper, 1993). For these two distinct pathways, it is the remarkable storage capability of DNA that handles this complicated genetic information through its extremely precise control over – and fine-tuning of – the coding and decoding processes. Although the whole coding information that is stored in the genome is kept in a relatively static state, it is ready to use under different physiological environments, coming into effect by cellular specific 'retrieval machinery'. This has recently been demonstrated by encoding a Shakespeare sonnet, including correcting for 'mistakes' [4]. Importantly, this ability of DNA to self-correct, removing its errors, can be incorporated into this new technology, as well [4]. This provides for a predictable and reliable mechanism for information retention by means of DNA's effective, long-term, and correctable characteristics.

Furthermore, binary code has become the fundamental language of modern computing and information storage since it was first noted by Gottfried Leibniz in 1679 [7]. Until 1988, the four base pairs of DNA were transferred into binary code, which appears as though a new information storage method based on DNA would certainly come along [3]. It is clear that as this technology develops further, transforming currently-available and proven binary code processes for their counterparts in, what could be called, "DNA MemoChip technology" (e.g., select information retrieval and DNA strands) will become the basis for stable and efficient long-term information storage. Ensuring that progress is made in this arena is only a matter of substantial initial expenditure [3,8]. Even if this "DNA MemoChip" technology is a promising way forward for managing and storing enormous amounts of information in the future, there still exist a significant number of technological challenges that will need to overcome. One key issue concerns the 'sandwich-like' spiral structure of DNA, which can foster "DNA MemoChip" capabilities in retaining information via binary coding system; however, its secondary and tertiary structures condensing the whole genetic information into a tiny space will be essential to mimic in order to maximize storage volume and, yet, minimize the storage space needed.

Another pivotal aspect of this potential "DNA MemoChip" relates to information retrieval. With respect to how DNA information is – and has been – stored, information retrieval has always been a time relative consuming process, and a case-sensitive process that unfolds in an on-demand manner. Accordingly, how can we retrieve information from these chips within seconds without errors, and present it on our computer monitors precisely? Although the retrieval techniques utilized in "cloud" storage may provide answers, it is hard to be used by individuals flexibly anytime and anywhere. Therefore, it can be argued that using a chip reader is the next, logical step to be explored, including the possibility that it becomes compatible with personal smart devices, such as our phones and tablets. Here, we could be carrying around all the information we may possibly need in portable 'super libraries'; the vehicles through which this information is accessed may change with technology, but the mechanism for storage and retrieval remain the same.

This DNA informational storage and select retrieval technology, which we are calling the "DNA MemoChip", represents a giant

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

step in meeting our technological and informational demands, which simultaneously will transform the field. Indeed, future generations will benefit from a common storage process that can withstand obsolescence, which up until now, has been understood as inevitable.

References

1. Lynch MF: Storage and retrieval of information on chemical structures by computer. Endeavour, 1968; 27(101): 68–73
2. Kowalski GJ, Information Retrieval Systems: Theory and Implementation. 1 ed. The Information Retrieval Series. Zhai CX, de Rijke M (eds.). New York: Springer US, 1997; Vol. 1: 282
3. Extance A: How DNA could store all the world's data. Nature, 2016; 537(7618): 22–24
4. Goldman N, Bertone P, Chen S et al: Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. Nature, 2013; 494(7435): 77–80
5. Zhirnov V, Zadegan RM, Sandhu GS et al: Nucleic acid memory. Nat Mater, 2016; 15(4): 366–70
6. Stefano GB, Kream RM: Artificial Intelligence, DNA mimicry, and human health. Med Sci Monit, 2018; 24: 3923–24
7. Blass U, Honkala I, Litsyn S: On binary codes for identification. J Combin Designs, 2000; 8(2): 151–56
8. Church GM, Gao Y, Kosuri S: Next-generation digital information storage in DNA. Science, 2012; 337(6102): 1628

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]