

A Revolution in Computer Science Technology- How to Store Data in Bacterial-DNA

SAI PRIYA K.S

Student / Department of B.E., Computer Science & Engineering
Velammal Engineering College, India.
Saipriya.ks@gmail.com,

Abstract— The field of information technology has undergone many changes and had caused a lot of revolutions in recent years. There had been many innovating inventions in last few decades, which made human lives easy and reliable. Computers play a vital role in everybody's life. The main purpose of using computer is to actually store and retrieve data for one's own use. The most safest and reliable place to store data nowadays, is found to be in hard-drives, etc., But the life time of a hard-drive is found to be less. Thus, we have some other inventions which helps to overcome this problem. One such invention is the concept of storing data in bacterial DNA. It enables to store the contents of an entire computer hard-drive on a gram's worth of E. coli bacteria. It is found that we can store data like music, images and even videos in the safest storage area, which is found to be the DNA of bacteria. There are methods to store the data and also there are techniques which are undergone to retrieve the data stored in them. Let us discuss them in this paper.

Keywords— *biostorage, Luminometer, Chelex's process, Mutation, Bioluminescence, Intercalating Agent.*

I. INTRODUCTION

Biostorage, the term for storing and encrypting information in organisms, has only existed for close to a decade, but scientists say that the method could soon allow for storing text, images, music or even video to be recorded in E. Coli, according to discovery. The concept of going to storing data in bacteria is that the life time of bacterial enzymes is higher than that of any hardware sources to store data. And also, even very simple bacteria have long strands of DNA with tons of bases available for data encryption, and bacteria are by their nature far resilient to damage than more traditional electronic storage. Bacteria are nature's hardiest survivors, capable of surviving just about any disaster that would finish off a regular hard drive. Hard disks may also be damaged as shown in fig1:



Fig 1: damaged hard disk

II. LITERATURE SURVEY

The study, published in the academic journal *Biotechnology Progress*, describes how DNA encoded with "E=MC2 1905!" was put into the genome of the common soil bacterium *Bacillus subtilis*. While ink may fade and computers may crash, bacterial information lasts as long as a species stays alive — possibly a mind-boggling million years — according to Professor Masaru Tomita, who heads the team of researchers at Keio University.



Fig2: bacterial enzymes

III. DNA

DNA storage is hierarchical and multilayer. We may say it consists of files grouped into folders which are organized into volumes of data. Every animal cell has a nucleus (the central and controlling part of the cell). This contains some fixed amount of chromosomes (depending upon the organism). These chromosomes have many genes. Each gene is made up of millions of DNA. These nucleic acids are just complex organic molecules. As per Watson and Crick (persons who designed the DNA structure at first) it is a helical ribbon structure. There are two ribbons on either side, each contains alternatively placed sugar and protein molecules. These proteins are bonded with Adenine, Guanine, Thiamine and Cytosine.) Any small change in the environment can easily affect this molecule which may sometimes undergo mutation. Mutation is a small or may be even bigger sudden change in the genome sequence. This may even change the entire organism. This mutation helps us in data storage.

The DNA is as stable as its living carrier (bacteria, plant or animal cell). In some cases after the cell's death the DNA data may survive. In suitable conditions the DNA may hold its data for years. On the other hand the DNA is vulnerable to hydrolysis and oxidation. Many factors lead to mutations in the genes. In mutations, letters of the genetic code can be changed and stretches of DNA can be deleted. The mutation rate of DNA replication process of Escherichia coli is near 10^{-9} . Some biological cells are more resistant to environmental conditions than others. For example bacteria such as Deinococcus radiodurans are especially good at surviving extreme conditions. They can tolerate high temperatures, desiccation and ultraviolet and ionizing radiation doses about 1000 times higher than fatal ones to humans. The DNA data protection mechanism has two basic layers. The genetic code is to some extent redundant and the genetic information is massively replicated. Most of the amino acids are encoded by more than one codon (in total 61 codons are used to code 20 amino acids), for example alanine is coded by four different codons. On the other hand there are amino acids (such as methionine) which are coded by single codons. Furthermore, the same DNA information is replicated in every cell of the particular living organism. The DNA strand actually looks like fig3:



Fig3: DNA strand

A mutation occurs when a DNA gene is damaged or changed in such a way as to alter the genetic message carried by that gene. A mutagen is an agent or a substance that can bring about a permanent alteration to the physical composition of a DNA gene such that the genetic message is changed. Once the gene has been damaged or changed, the mRNA transcribed from that gene will now carry an altered message. The polypeptide made by translating the altered mRNA will now contain a different sequence of amino acids. The function of the protein made by folding this polypeptide will probably be changed or lost. No product (red pigment) is produced by the altered protein. In subtle or very obvious ways, the phenotype of the organism carrying the mutation will be changed. In the present study, we make the DNA to mute by radiations, which is induced mutation. Ultra violet rays are ionizing radiations to mute the DNA. We encode the data into the genome. This small change in the DNA can be transmitted over generations so that the data will live along with the molecule. This mutation results in evolution.

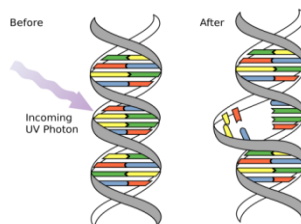


Fig4: mutation of DNA.

When the ultraviolet rays fall on the DNA it absorbs some energy. This energy excites the electrons in the DNA to a higher energy level. As a result it results in mutation. The two nearby thiamine molecules get bonded with each other leaving the adenine molecule. This forms intercalating agent. This instability causes the DNA to get muted. Again when it gets into the normal state the genome sequence is changed and results in mutation. This is induced by ultraviolet rays. The role of ultraviolet rays is to increase the mutation rate. Now here we control the mutation rate by digitally made electrical signals. This forms the basic principle in data storage in DNA.

IV. DATA ENCODING IN DNA

At first a suitable bacterium is chosen which must be chosen in a way that it must tolerate high temperatures, desiccation, ultraviolet light and ionizing radiation doses 1000 times higher than would be fatal to humans. Bacteria like *Escherchia coli*, *Deinococcusra diodurans*, *Serratiamarcescens*, *Bacillus subtilis*, etc are normally used as they can tolerate extreme conditions. The DNA molecule is extracted from it by Chelex's process (A well known biotechnological method to collect bacterial DNA on water on a test tube). Then it is stored in a proper temperature. Normally 20oC is maintained. Then the DNA is cloned (taken synthetic copies of its own by biotechnological method). The synthetic DNA is taken for processing of data storage.

Then the DNA is taken for Bioluminescence. It is a process of taking the DNA to a state by which it is ready to accept electrical signals. The process of Bioluminescence contains the steps that are as follows. The DNA is first heated to 30oC in a biological medium. Then the DNA is brought to luminescence state. In the luminescence state the data will be ready to receive electrical impulses. This is done by using some chemical substances that makes the DNA to a luminescence state. Normally aldehyde n decanal is used as the illuminometer. Culture aliquots were removed at regular intervals and assayed with the aldehyde n-decanal in an inexpensive analog luminometer. This luminometer reduces the density of the DNA contents. After this the density of the DNA in the liquid is very much reduced. The content is kept like this for few hours. As hours increase the effect of luminometer becomes higher. But the bacteria should not be in an extreme luminescence state. It will destroy the DNA. So it can accept only to a limit of luminometer. Then the luminometer is removed from the DNA.

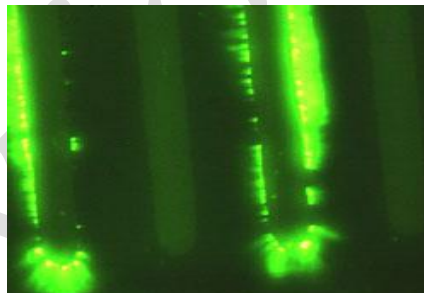


Fig5: gold rods that are inserted in the DNA content.

In the presence of ultraviolet light the liquid medium containing DNA is encoded with digital data. The presence of ultraviolet light is to increase the effect of mutations. By the effect of ultraviolet light the effect of mutation becomes more spontaneous. Now two electrodes are inserted into this liquid medium which contains the DNA. These electrodes are made of gold. They are of one nanometer in length. Four such electrodes are used in this. The digital data is converted into electrical impulses. These impulses are sent through these electrodes and this will encode the data onto the DNA. After encoding the data on the synthetic DNA, the DNA is inserted into living bacteria as a living medium. This data will live along with the bacteria for generations.

By this method data can be stored in the DNA and data will be transmitted over many generations. These data can last even up to 3000 years. Hence, even after 100 human generations we can get the data. Bacteria may be an inexpensive and stable long-term means of data storage. We can store high content of data on a small piece which can just measure a few nanometers. Genetic coding is so massive that information can be stashed away somewhere in the gene without affecting an organism's overall appearance and other traits.

V. RETRIEVAL OF DATA

Data stored on the DNA can be retrieved by decoding it. [2] The reverse process in the encoding is just done. The DNA is extracted from the bacteria at first. The luminometer is again added and the concentration of DNA is fairly reduced. Electrodes are placed into the liquid medium that contains DNA molecules and ultraviolet light and ionizing radiations are given to the content at 30oC. The encoded data in the DNA get transformed into the electrical impulses and the data are retrieved through the electrodes.

There will not be any loss of information due to this process. The original information that was encoded is retrieved from the content. The amount of information that can be stored depends up on the amount of bacteria in the content. If a milliliter of liquid can contain up to billion bacteria, the potential capacity of such a memory system is enormous. Higher the bacteria content, higher the storage capacity.

VI. RETRIEVAL OF DATA : COMPARISION OF DNA ANDHARD – DRIVES

PARAMETER	DIGITAL STORAGE DEVICE	DNA STORAGE
Basic element	Transistor, magnetic or optical Domain	Nucleotide
Basic element Size	1,000/ 10,000 nm ²	1 nm ³
Addressing scheme and Organization	Discrete, hierarchical	Sequential, parallel copying.
Redundancy	Multi layer	Data replication
Capacity	Unlimited	0.1 Mbit/mm
Areal density	iGbit/mm ²	-
Access time	0.01 milli sec	-
Internal data Transfer	10 Mbit/ sec	100 Kbits/ sec
Error rate	10 ⁺⁴	10 ⁻⁹

VII. CONCLUSION

From this discussion we can find that, Bacteria are the future data storing device. Data can be stored in bacterial cells and can be retrieved after 1000 years. Bacteria are cheap and store enormous data. All future technologies may be based on the data storing in biological material including bacterial cells.

References

- [1] Digital and Biological Storage Systems - a Quantitative Comparison, M. Wasizu, O. Kurosawa, I. Arai, S. Suzuki, and N. Shimamoto, IEEE 2005.
- [2] Quantitative Analysis of DNA Orientation in Stationary AC Electric Fields Using Fluorescence Anisotropy Seiichi Suzuki, Member, IEEE, Takeshi Yamanashi, Shinichi Tazawa, Osamu Kurosawa, and Masao Washizu, Senior Member, IEEE 2008
- [3] National Center for Biotechnology Information.
- [4] Boffins store data in bacterial DNA, by Ian Williams
- [5] TRN, the Latest Technology Research News.
- [6] NewScientist.com, Data stored in multiplying bacteria
- [7] PhysOrg.com, DNA For Information Processing and Data Storage
- [8] FORESIGHT foresight.org nanotech institute, Data storage using DNA.
- [9] LiveScience.com New Technique Stores Data in Bacteria by Bill Christensen
- [10] Researchers store data in bacteria DNA. Information lasts as long as a species stays alive — possibly a million years, msnbc.msn.com
- [11] R. Staden, "A mew computer method for the storage and manipulation of DNA gel reading data", MRC Laboratory of Molecular Biology Hills Road, Cambridge CB2 2QH, UK, Oxford Journals, June 16, 1980.



www.ioirp.com

International Journal of Innovative Research in Computer Science & Engineering (IJIRCSE)
ISSN: 2394-6364, Volume – 1, Issue – 4. March 2015

- [12] Be´atrice Godard, Jo´rgSchmidtke, Jean-Jacques Cassiman and Se´gole`neAyme´, "Data storage and DNA banking for biomedical research: informed consent, confidentiality, quality issues, ownership, return of benefits. A professional perspective", INSERM SC11, HopitalBroussais, Paris, France; Institute of Human Genetics, Hannover Medical School, Hannover, Germany; and Center for Human Genetics, Leuven, Belgium. 2003
- [13] Roger Staden, "Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing", Oxford Journals, 1982
- [14] C T Chung, S L Niemela, and R H Miller, "One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution", Proceedings of the National Academy of Science of the United States of America, April 1989

WWW.IOIRP.COM