TRANSACTIONS ON SCIENCE AND TECHNOLOGY

Basic Architecture and Applications of DNA Computing

Nordiana Rajaee[1]*, Azham Zulkharnain[2] & Awang Ahmad Sallehin Awang Hussaini[2]

[1] Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, MALAYSIA.

[2] Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, MALAYSIA

*Corresponding author. E-Mail: rnordiana@unimas.my; Tel: +6082-583277; Fax: +6082-583410.

Received: *20 March 2016*
Revised: *7 April 2016*
Accepted: *24 April 2016*
In press: *25 April 2016*
Online: *30 June 2016*

Keywords:
DNA Computing;
Bio-inspired Computing

Abstract

DNA computing is a new computing paradigm utilizing actual DNA oligonucleotides to do computation by employing biomolecular tools to get the reaction and outputs extraction. In this paper, we introduce basic architecture of DNA computing. Brief explanation on the biomolecular tools employed in DNA computing are also included, and its various applications in many fields are also discussed.

© Transactions on Science and Technology 2016

Introduction to DNA computing

DNA or deoxyribonucleic acid works as a "memory" to store genetic information in cellular organism. Consisting of four bases A (adenine), T (thymine), C (cytosine) and G (guanine), these bases follow a Watson Crick complementary rule whereby A complements T, C complements G and vice versa. One DNA base is called an oligonucleotide and its length denoted in mer. DNA strands are often quoted in 5'-3' order and two single stranded DNA sequences may combine to form a double stranded DNA. The length of a double stranded DNA is denoted as base pairs. Figure 1 shows basic structure of a double helix DNA strand and its bases.

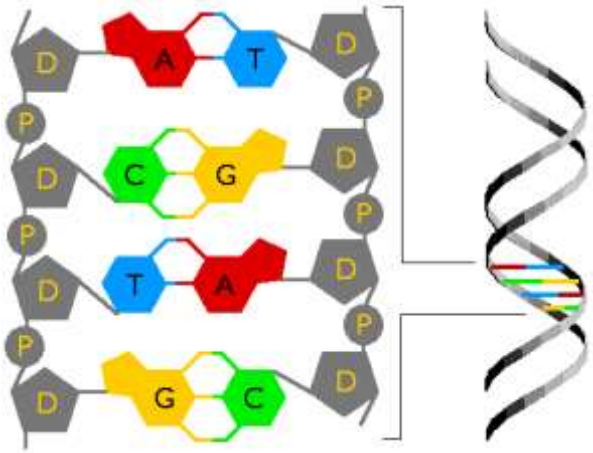

Figure 1. Structure of DNA

DNA computing is a new computing paradigm which encodes the information in a problem into the DNA oligonucleotides and by using biomolecular tools, the constructed DNA sequence strands are extracted to decipher its computation results. Much of DNA computing relies on developing algorithms that solve problems using the encoded information in the sequence of nucleotides that make up DNA double helix and then breaking and making new bonds between them to reach the answer. DNA computing was pioneered by Leonard M Adleman in 1994 when he solved a seven-node Hamiltonian Path Problem (HPP) using actual DNA oligonucleotides. His works later spurred various researches in many fields including biology, chemical, mathematics to information technology. The major advantages of DNA computing are its self-assembly properties, huge information storage capacity, massive parallelism and low dissipation energy (Adleman, 1994).

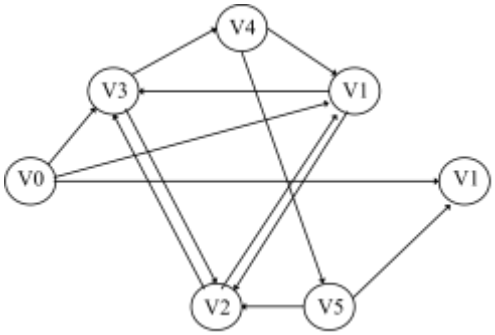

Figure 2. Seven-node Hamiltonian Path Problem (HPP)

Biomolecular tools

DNA computing is a wet-lab process which includes a number of feasible bio-chemical operations (tools) necessary to execute such computations. In general, it consists of two steps: the reaction, involving the interactions of the molecules to produce potential solution molecules, and the extraction, the techniques employed to isolate or otherwise identify any correct solution molecules during the reaction. In this section, we briefly present the basic background of biomolecular tools employed in DNA computing.

Hybridization is the annealing of complementary single stranded molecules to form a double stranded DNA. This is the basis for initial path formation during the reaction step and is subsequently employed during the extraction phase for the isolation of generated path molecules.

Ligation is a process often invoked after single stranded DNA are annealed and concatenated to each other. Many single stranded fragments are connected in series and ligase is used as “glue” to seal the covalent bonds between the adjacent fragments.

Denaturation is a melting process in vitro. Double stranded DNA molecules can be separated without breaking the single strands by applying heat to the solution. The double stranded molecules come apart because the hydrogen bonds between complementary nucleotides are much weaker than the covalent bond between the adjacent nucleotides in the same strands.

Cutting process is carried out by using Restriction Enzymes (RE). Restriction enzymes recognize a specific sequence of DNA known as a restriction site. Any DNA that contains the restriction site within its sequence is cut by the enzyme at that point.

Polymerase Chain Reaction (PCR) is an amplification technique widely used in molecular biology. A pair of DNA sequences known as "primers" is used to signal the starting point and ending point for a specific target DNA sequence for amplification. The PCR process is capable of exponentially amplify a DNA strand into millions of its copies given a site-specific single molecule DNA and the process is usually carried out in three stages of different temperatures.

Gel electrophoresis is a technique used for separation of DNA strands according to their sizes using electric current applied to the gel containing the strands. The size of the DNA strands refers to the weight of the DNA strands which is proportional to the lengths of their sequences. This technique is based on the fact that DNA molecules are negatively charged. Since DNA molecules have the same charge per unit length, they all migrate at the same speed in aqueous solution. However, if electrophoresis is carried out in gel, the migration rate is affected by its size causing less weighted strands to migrate faster. Thus, sorting the strands by their sequence lengths is made possible using this technique. The results of gel electrophoresis process can be viewed by staining gel with fluorescent dye and photographed under UV light (Amos, 2004).

Adleman's architecture

In Adleman's work, he developed an algorithm to encode the HPP into DNA oligonucleotides and utilized hybridization-ligation, cutting process by restriction enzymes, polymerase chain reaction method and gel electrophoresis to extract the outputs of the computation. Figure 3 shows the Adleman's architecture.

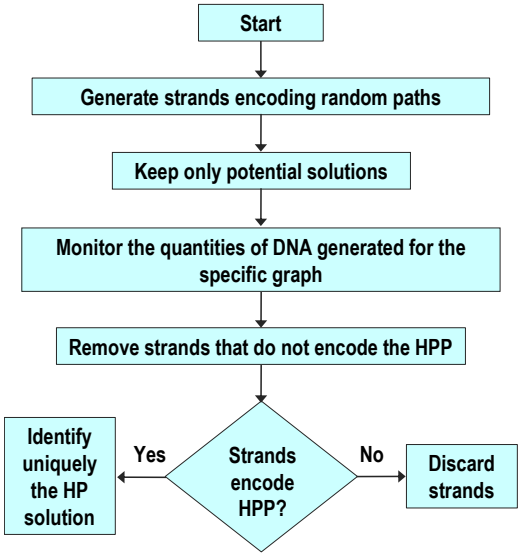

Figure 3. Adleman's Architecture

The Adleman's architecture works by generating random 20 mer-length DNA sequences for represent all the vertices and edges in the seven-node graph. Using hybridization-ligation process, all possible solutions (all paths) are constructed. Polymerase chain reaction is then employed to extract only paths beginning with a and ending with. All paths with various distances from the initial vertex to the terminal vertex are then subjected to gel electrophoresis process to be sorted. From the gel electrophoresis, only 140 base pair length (representing seven nodes) are extracted. The solution path of 140 b.p is then subjected to magnetic bead separation process to check the exact sequence of nodes for the HPP problem (Adleman, 1994).

Applications of DNA computing

Implications to Mathematics

After Adleman's work in solving the HPP, the DNA computing has been proposed to solve various problems in many fields. In the beginning, most application were proposed for solving mathematics, computer science and engineering problems such as Boolean circuits, chained integer arithmetics, addition and substraction operations (Ezziane, 2005). More extended proposals followed by physical implementation of DNA computation were conducted to solve weighted graph problems, job scheduling problems and elevator scheduling problems (Watada, 2008).

Implications to Computer Science and Engineering

More recent works proposed DNA computing intended for solving matching types problem in DSP techniques (Tsaftaris, 2004), applicability of DNA computing algorithm to solve image recognition in intelligent visual mechanics (Tsuboi, 2003) and control systems (Dong, 2009). In the area of information technology and security, DNA computing has also been proposed for encryption of information data (Javheri, 2014), security analysis of images (Gupta, 2015) and to conceal messages in DNA whereby increased privacy, security and big data storage is available (Hakami, 2015).

Implications to Medicine

The big data in clinical settings also benefitted from DNA computing (Cannataro, 2012), especially in diagnosis and treatment of diseases (Hornweder, 2011), creating small scale computers to monitor patients' conditions and analyzing multiple gene expressions simultaneously (Salehahmadi, 2014). Self explanatory DNA authentication chips and bio-sensors are also developed to correctly adjust medication concentration and release of drugs. DNA computing is currently accelerating predicted to revolutionize computing especially in the medical field (Salehi, 2016).

Issues and challenges

DNA computing capitalize on the massive storage information capacity of DNA and its fast simultaneous computation capabilities, which makes them an intriguing prospect for solving hard NP

problems. While the fastest supercomputers right now is operating at 1 billion operations per second, a drop of DNA is estimated to be equivalent to 15,000 trillion computers. However, the progress of the development of DNA computing is still hindered by several limitations such as the short shelf life of DNA which maximum spans of 6 months, and algorithms utilized to perform the computation are slow molecular biological operations.

Future of DNA computing

As recently reported of Moore's Law getting obsolete, the need for new computing paradigms for faster and high performance computation is in search. DNA computing is among the emerging computing paradigms predicted to be the next replacing the conventional silicon computers. Not only capable of solving computationally hard problems, the eminent attractive properties of DNA which proves groundbreaking in the fields of security and encryption also extended its potential across different fields. Currently, the DNA computing is still in its infancy stage even with the latest developments in the last few years as we have yet to find a way to automate a DNA computer without human intervention.

Conclusion

This paper briefly introduces DNA computing, its basic architecture and biomolecular tools utilized in the physical implementation of DNA computing. The area of DNA computing holds great potential to be explored due to its applications in many various fields. However, the DNA computing is still at its early stages and there are limitations which need to be overcome before it can actually replace silicon computers.

Acknowledgements

The authors would like to give thanks to Universiti Malaysia Sarawak for supporting this work under Small Grant Scheme F02(S153)/1154/2014(18).

References

[1] Adleman L. (1994). Molecular computation of solutions to combinatorial problems. *Science*, **266**, 1021-1024
[2] Amos, M. (2004). Theoretical and Experimental DNA Computation, Natural Computing Series, Springer
[3] Boruah, K. & Dutta, J. C. (2015). Twenty years of DNA computing: From complex combinatorial problems to the Boolean circuits, *International Conference on. IEEE Electronic Design, Computer Networks & Automated Verification (EDCAV)*, 29 -30 January, Kolkatta, India, 52-57.
[4] Cannataro, M., Santos, R. W. D, Sundnes, J. & Veltri, P. (2012). Advanced computing solutions for health care and medicine. *Journal of Computational Science*, **3**(5), 250-253
[5] Dong, M, Wang, N. & Tao, J. (2009). DNA computing in control systems: A survey. *Control and Decision Conference.* December 16-18. Guilin. 4942-4946.
[6] Ezziane, Z. (2005). DNA computing: applications and challenges. *Nanotechnology*, **17(**2), R27
[7] Fundamental Concepts of Biology (2016). (http://www.ric.edu/faculty/jmontvilo/109files/109(05) dnaandrna.htm), Accessed on April 25, 2016).
[8] Gupta, S. & Jain, A. (2015). Efficient image encryption algorithm using DNA approach. *2nd International Conference on. IEEE Computing for Sustainable Global Development*, 11-13 March, India, 726-731.

[9] Hakami, H. A., Chaczko, Z. & Kale, A. (2015). Review of Big Data Storage Based on DNA Computing, *Asia-Pacific Conference on. IEEE Computer Aided System Engineering (APCASE)*, 14-16 July, Educador, 113-117

[10] Hornweder, K. S. V. (2011) An Overview of Techniques and Applications of DNA Nanotechnology, *Technical Report* UT-CS-11-682.

[11] Javheri, S. & Kulkarni, R. (2014). Secure Data communication and Cryptography based on DNA based Message Encoding. *International Journal of Computer Applications*, **98**(16), 360-363

[12] Misbahuddin, S. C. S, M. & Hashim, N. P. (2014). DNA for information security: A Survey on DNA computing and a pseudo DNA method based on central dogma of molecular biology. *International Conference on Computer and Communications Technologies*. 11 – 13 Dec, Hyderabad, 1-6.

[13] Salehahmadi, Z. & Khishdoust, H. (2014). DNA Computing Revolutiones Medical World. *Merit Research Journal of Medicine and Medical Sciences*, **2**(9), 179-185.

[14] Salehi, S. A., Jiang, H., Riedel, M. D. & Parhi, K. K. (2016). Molecular Sensing and Computing Systems. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, **1**(3), 249-264

[15] Tsaftaris, S. A., Katsaggelos, A. K. & Papoutsakis, E. T. (2004). How can DNA computing be applied to digital signal processing? *Signal Processing Magazine, IEEE,* **21**(6), 57-61.

[16] Tsuboi, Y. & Ono, O. (2003). Applicability of DNA computing algorithm solving image recognition in intelligent visual mechanics. *Industrial Electronics Society*, **3**, 2800.

[17] Watada, J. (2008). DNA computing and its applications. *Eighth International Conference on Intelligent Systems Design and Applications,* **2**, 288-294.

[18] Yusof, Z. M., Rahim, M. A. A., Nawawi, S. W., Khalil, K., Ibrahim, Z. & Kurniawan, T. B. (2012). DNA Sequence Design for Direct-Proportional Length-Based DNA Computing: Particle Swarm Optimization vs Population based Ant Colony Optimization. *Fourth International Conference on Computational Intelligence, Modelling and Simulation.* 25 – 27 Sept, Pahang, Malaysia. 65-69