DNA Computing.

Changes of the challenges and opportunities in this field over time: a literature review

By Lucas Toner (200307631)

Table of Contents

Introduction

‘DNA computing’ is part of a broader branch of computation known as ‘molecular computing’. It refers to the use of the DNA molecule to perform computations, rather than using silicon and electronic circuits, as is convention.

The DNA molecule has a suite of properties that makes it uniquely suitable for performing computations:
For example, each DNA strand is made up of two complementary strings of proteins in parallel, known as “Watson-Crick Complementarity”. If two complementary strings come into contact, they will “hybridise” and come together to form a molecule of DNA.
DNA also has an extremely high information density[8], in comparison to conventional storage, it is also extremely stable for long periods of time, can be stored 3 dimensionally, and is self repairing.
There a numerous techniques used to manipulate DNA, not limited to the use of only DNA molecules, but using other molecules such as enzymes[6] and CRISPR[20].
It is also tried and and tested, nature has been using DNA to perform computations for a very time! Leonard Adleman is quoted as referring to it as "Nature's Toolbox”[4].

It represents an interesting intersection of disciplines that has already been the source of innovative practical technologies.

Seminal Paper

The seminal paper in DNA Computing describes an experiment performed by the author, Leonard Adleman, the first work of it's kind in 1994[1]. He solved an NP-complete problem, the Hamiltonian path problem, using molecules of DNA as his computation medium.

To briefly summarise the experiment: Adleman's experiment took advantage of the massive parallelisation that DNA computing is capable of to produce every possible solution to the

Hamiltonian path problem for a small graph. These candidate solutions were then filtered via DNA manipulation techniques so that a correct solution could be extracted and read.

In general, this is considered as the beginning of DNA computing as a field.

The State of the Art Then

Much of the early work in DNA computing was focused upon building on the work in the seminal paper [1]. True to the definition of NP-Complete, the algorithm was generalised shortly after Adleman's paper [3], with other researchers producing solutions to other problems, following a similar brute-force technique.

A variety of novel techniques and concepts were pioneered in the early days of the field. Not all of these topics are unique to DNA Computing, however they are all relevant:

- **Surface computing** – Fixing DNA strands to some solid medium. Using two dimensional mediums for DNA computing was more error-resistant, and opened a door towards automation of various processes.[5][6]. It was met with some scepticism, however was considered by some a way forward.
- **DNA Microarrays** – A chip on which DNA molecules can be hybridised, and read from. This was faster and cheaper than direct sequencing and was considered a step towards automation. This can be considered a branch of surface computing, but was regarded with optimism.[6]
- **RNA computing** – Performing computation with RNA rather than (or alongside) DNA to take advantage of it's unique properties. One example[4] takes advantage of an enzyme that digests any RNA molecule that has hybridised with a DNA molecule, using it to filter out specific RNA molecules.
- **Microfluidics** – Controlling the flow of fluids on a micro-level, allowing for fine-grained control of liquids containing DNA strands. This was considered to be very useful in future paths towards automation.[6]
- **DNA Origami** – Not yet referred to as DNA Origami, but as "DNA self-assembly"[5], the construction of structures via DNA that folds itself into shapes. It began to get some real scrutiny in the late 1990s, with extreme optimism about it's potential in the future[7]. Importantly, the relationship between DNA Origami and DNA computing was mainly hypothetical in the early days of both fields.[7]

Early Promises

In the conclusions of his paper [1], Adleman hypothesised what the future of DNA computing may look like.
He envisioned that the massively parallel properties of DNA computing could be exploited to solve complex algorithms that conventional computers are unsuitable for, with greater time and energy efficiency. As a long term prospect, he also imagined a general purpose computer, capable of solving any problem, consisting of a single macromolecule operated upon by enzymes.

Papers published after the field had developed a little maturity have similar promises, but with caveats. In 1997, Gibbons et al.[5] suggests that, should DNA computing become more efficient in terms of the amount of material it requires, then DNA computing may have some potential use in solving problems that are intractable for conventional computers. The concept of using DNA computing for general purpose use is completely disregarded.
In 2000, Ruben And Landweber[6] are a little more optimistic, considering that there are a few approaches that suggest automation is a possibility. They believe that the massive parallelism of DNA provides great potential. However, they lay out several challenges that the field currently faces. For example, like Gibbons et al., they consider that the volume of DNA needed for useful computation is too great, and would need to be reduced for DNA computing to find a place.

The State of The Art Now

While the fundamentals of DNA computing today are still recognisable, the field has evolved significantly in the past 23 years. Contextually, it is important to consider that our understanding of DNA has improved vastly, with a backdrop of advances in Biochemistry and molecular chemistry that DNA computing can take advantage of.

The breadth of DNA computing has also advanced considerably with many avenues of study emerging within the past 23 years, and some losing relevance. Regarding the topics that were of high interest in 2023:

- **Surface Computing –** Surface-based DNA computing still finds use in the niches of the field[22], however is no longer in mainstream use.
- **DNA Microarrays** – While still in use today, they are no longer popular in DNA computing, as the cost of sequencing has fallen dramatically.[8][9]
- **RNA Computing** – Using RNA within molecular computing is no longer a novelty and is used throughout the field for different purposes, from circuit design[10] to environmental responses.[11]
- **Microfluidics** – Microfluidics remains an active and important field across the biochemical spectrum, and remains vital in controlling reactions in DNA computing, for example, a chip using Microfluidics and DNA.[12]
- **DNA Origami** – DNA origami (otherwise known as self assembly) is one of the fundamental areas of DNA computing, used throughout the field in areas such as environmental response[14] and structuring data.[13]

Regarding some of the topics that have emerged over time:

- **Strand Displacement** – A ground-breaking technique that is now one of the foundations of modern DNA computing[8]. It was pioneered in 2000[15] and became mainstream over the next decade.[14]

- **CRISPR –** Another technique that has changed the field is CRISPR, a type of molecule that can be used to precisely detect and modify DNA, the molecule's applications were discovered in 2012[20], and since then CRISPR has seen substantial use in DNA computing.
- **Data Storage –** The idea of taking advantage of DNA's inherent information density is not a new one, it was referenced in Adleman's seminal paper [1], however there is currently substantial theoretical and commercial interest in DNA Data storage. An organisation of note is the DNA Data Storage Alliance, with notable members such as Microsoft and Western Digital.[16]
- **Computation *In Vivo* –** To a greater extent, DNA computing can be seen as a predecessor to Synthetic Biology [17], the engineering of biological systems. A notable example is the use of self-assembled DNA molecules to control molecules inside a living organism[18]. There has also been growing interest in temporal Data storage *In Vivo.*[13]
- **Quantum Applications –** The relationship between DNA and the quantum realm is a relatively new field[17]. The quantum properties of DNA are still under strong research, for example, it was found in 2021 that DNA exhibits quantum properties, such as quantum tunnelling[19].

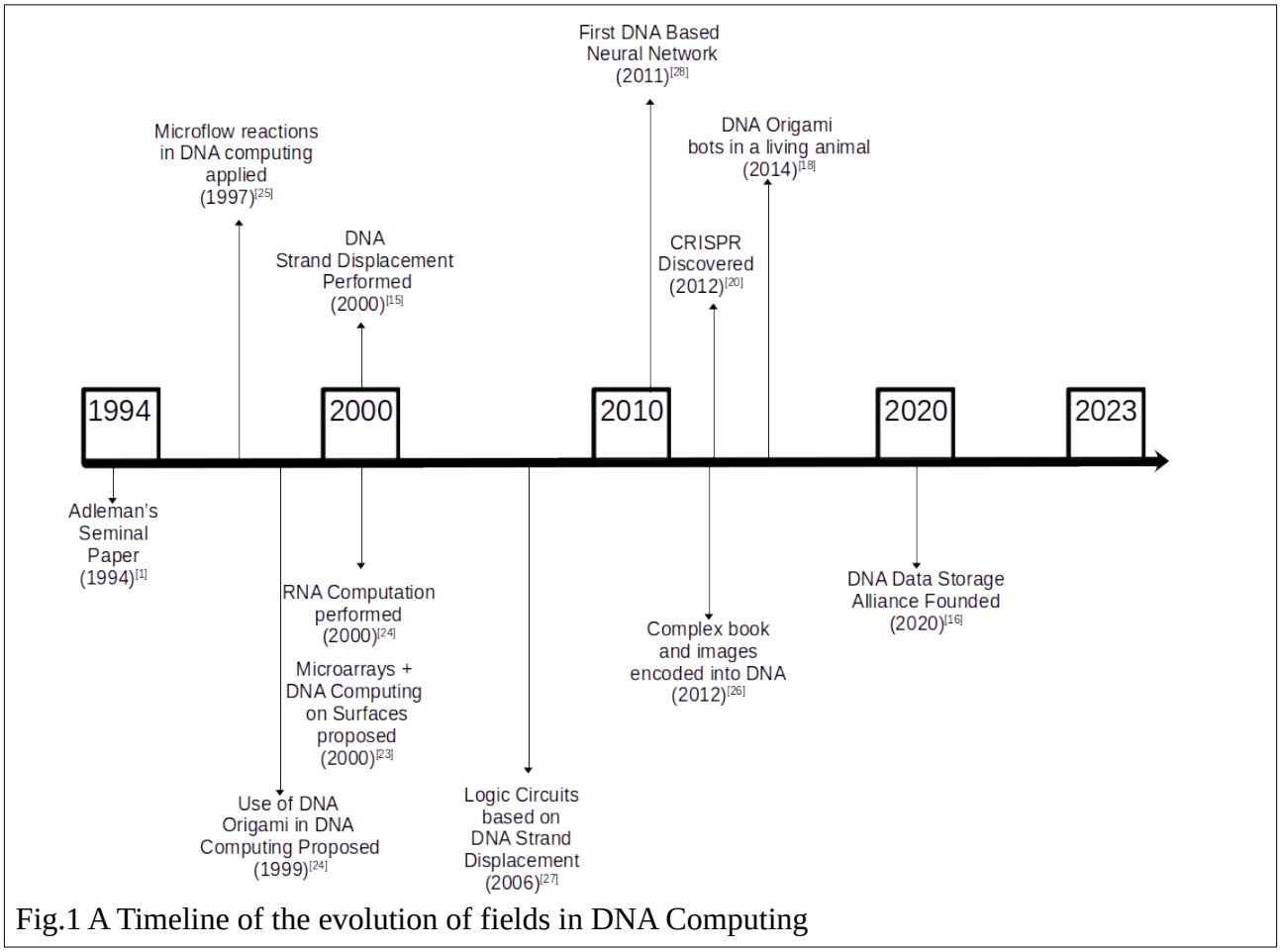

Fig.1 A Timeline of the evolution of fields in DNA Computing

Current Promises and Challenges

Some of the challenges that faced the field in it's infancy are now not as great a concern. The sequencing of DNA, which was once an expensive inconvenience, is now a relatively trivial process[8], with one of the more affordable devices, a *Nanopore MinION* costing less than £900 (*as of December 2023*) for example.
The cost of synthesis however, remains prohibitively high[8]. There has been a focus from the community to reduce the number of materials required in computation[8].

The field of DNA computing in 2000 was focused on algorithms that "isolated solutions, rather than constructed them", in which a brute force technique was used to produce every possible solution, and then a correct one was isolated from that pool[5].
Modern DNA computing takes a more specialised approach, and computations are typically much more specialised. For example, the massively parallel nature of DNA can be exploited to perform mass operations in a database [13], rather than brute force calculations.

The key challenges facing DNA computing today are those on the larger scale, problems of the scalability of technology, reusability and logistics[8]. There are other considerations too, as DNA computing enters the corporate eye, for example in storage, regulation will be needed to support and guide the growth of a new type of technology[13].
It is also important to consider the ethical implications of DNA computing, especially in the context of computation and activity *In Vivo*.

Promises of DNA computing in 1990s/Early 2000s	Promises of DNA computing in 2023
Massive Parallelism will allow us to solve intractable problems in Polynomial time.[1][5][6]	Massive Parallelism will allow us to perform computation on lots of data simultaneously.[13][8]
	Neural Networks will take advantage of Massive parallelism for high speed pattern recognition.[8]
Automation of processes will greatly increase efficiency of computation.[5][6]	Automation of processes will greatly increase efficiency of computation.[8]
Self-Assembly has major potential to produce interesting molecules.[6]	DNA-Origami (Self-Assembly) has a huge range of applications, and yet unrealised potential.[21]
In the long term, general purpose DNA computing machines may be possible.[1]	We can use DNA computing alongside other techniques, such as quantum to take advantage of it's unique capabilities.[13]
	DNA storage will provide us with much more efficient storage, in terms of space and energy use.[13]
	DNA is a tool towards whole cellular computing.[17]
	DNA Computing gives us molecular control for biomedical purposes.[18]

Fig. 2 The Promises of DNA computation in 2000, and today in 2023

Fig. 2 shows a table that demonstrates how the promises of DNA computing have changed over the past 23 years. It's clear that the early promises of the field failed to capture the breadth of specific use cases in which DNA computing would be applicable, with the field moving in directions that was not predicted. Additionally, the promise of massive parallelism solving intractable problems has not been realised, due to the challenge identified in the field's infancy[5][6], of the need of exponential quantities of DNA required to solve larger problems.

Prospects for DNA computing are high, many of the applications of DNA computing today were not even considered 23 years ago, and in those applications in which DNA computing is appropriate, it has great potential. As this technology becomes commercialised for medical and information processing purposes, it is likely to receive more investment and scrutiny then ever before.

References

[1] Adleman, L.M. (1994) 'Molecular computation of solutions to Combinatorial Problems', *Science*, 266(5187), pp. 1021–1024. doi:10.1126/science.7973651.

[2] Song, Y. *et al.* (2018) ‘DNA multi-bit non-volatile memory and bit-shifting operations using addressable electrode arrays and electric field-induced hybridization’, *Nature Communications*, 9(1). doi:10.1038/s41467-017-02705-8.

[3] Lipton, R.J. (1995) ‘DNA solution of hard computational problems’, *Science*, 268(5210), pp. 542–545. doi:10.1126/science.7725098.

[4] Seife, C. (2000) ‘RNA works out Knight Moves’, *Science*, 287(5456), pp. 1182–1183. doi:10.1126/science.287.5456.1182.

[5] Gibbons, A., Amos, M. and Hodgson, D. (1997) ‘DNA computing’, *Current Opinion in Biotechnology*, 8(1), pp. 103–106. doi:10.1016/s0958-1669(97)80164-4.

[6] Ruben, A.J. and Landweber, L.F. (2000) ‘The past, present and future of molecular computing’, *Nature Reviews Molecular Cell Biology*, 1(1), pp. 69–72. doi:10.1038/35036086.

[7] Seeman, N.C. (1999) ‘DNA engineering and its application to nanotechnology’, *Trends in Biotechnology*, 17(11), pp. 437–443. doi:10.1016/s0167-7799(99)01360-8.

[8] Polak, R.E. and Keung, A.J. (2023) ‘A molecular assessment of the practical potential of DNA-based computation’, *Current Opinion in Biotechnology*, 81, p. 102940. doi:10.1016/j.copbio.2023.102940.

[9] *DNA Microarray Technology fact sheet* (2020) *Genome.gov*. Available at: https://www.genome.gov/about-genomics/fact-sheets/DNA-Microarray-Technology (Accessed: 07 December 2023).

[10] Schaffter, S.W. and Strychalski, E.A. (2022) ‘Cotranscriptionally encoded RNA strand displacement circuits’, *Science Advances*, 8(12). doi:10.1126/sciadv.abl4354.

[11] Green, A.A. *et al.* (2017) ‘Complex cellular logic computation using ribocomputing devices’, *Nature*, 548(7665), pp. 117–121. doi:10.1038/nature23271.

[12] Lee, W. *et al.* (2021) ‘Programmable DNA-based boolean logic microfluidic processing unit’, *ACS Nano*, 15(7), pp. 11644–11654. doi:10.1021/acsnano.1c02153.

[13] Song, X. and Reif, J. (2019) ‘Nucleic acid databases and molecular-scale computing’, *ACS Nano*, 13(6), pp. 6256–6268. doi:10.1021/acsnano.9b02562.

[14] Zhang, D.Y. and Seelig, G. (2011) ‘Dynamic DNA nanotechnology using strand-displacement reactions’, *Nature Chemistry*, 3(2), pp. 103–113. doi:10.1038/nchem.957.

[15] Yurke, B. *et al.* (2000) ‘A DNA-fuelled molecular machine made of DNA’, *Nature*, 406(6796), pp. 605–608. doi:10.1038/35020524.

[16] (2020) - *Official site of the DNA Data Storage Alliance*. Available at: https://dnastoragealliance.org (Accessed: 07 December 2023).

[17]Grozinger, L. *et al.* (2019) ‘Pathways to cellular supremacy in Biocomputing’, *Nature Communications*, 10(1). doi:10.1038/s41467-019-13232-z.

[18] Amir, Y. *et al.* (2014) ‘Universal computing by DNA origami robots in a living animal’, *Nature Nanotechnology*, 9(5), pp. 353–357. doi:10.1038/nnano.2014.58.

[19] Slocombe, L., Al-Khalili, J.S. and Sacchi, M. (2021) ‘Quantum and classical effects in DNA point mutations: Watson–Crick tautomerism in at and GC base pairs’, *Physical Chemistry Chemical Physics*, 23(7), pp. 4141–4150. doi:10.1039/d0cp05781a.

[20] Jinek, M. *et al.* (2012) ‘A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity’, *Science*, 337(6096), pp. 816–821. doi:10.1126/science.1225829.

[21] Haydell, M., Ma, Y. (2023). DNA Origami: Recent Progress and Applications. In: Valero, J. (eds) DNA and RNA Origami. Methods in Molecular Biology, vol 2639. Humana, New York, NY. https://doi.org/10.1007/978-1-0716-3028-0_1

[22] Tan, Xi, et al. "Scalable and programmable three-dimensional photonic processor." *Physical Review Applied* 20.4 (2023): 044041.

[23] Liu, Q. *et al.* (2000) ‘DNA computing on surfaces’, *Nature*, 403(6766), pp. 175–179. doi:10.1038/35003155.

[24] Faulhammer, D. *et al.* (2000) ‘Molecular computation: RNA solutions to chess problems’, *Proceedings of the National Academy of Sciences*, 97(4), pp. 1385–1389. doi:10.1073/pnas.97.4.1385.

[25] McCaskill, J.S. (1997) ‘Spatially resolved in vitro molecular ecology’, *Biophysical Chemistry*, 66(2–3), pp. 145–158. doi:10.1016/s0301-4622(97)00073-2.

[26] Church, G.M., Gao, Y. and Kosuri, S. (2012) ‘Next-generation digital information storage in DNA’, *Science*, 337(6102), pp. 1628–1628. doi:10.1126/science.1226355.

[27] Seelig, G. *et al.* (2006) ‘Enzyme-free nucleic acid logic circuits’, *Science*, 314(5805), pp. 1585–1588. doi:10.1126/science.1132493.

[28] Qian, L., Winfree, E. and Bruck, J. (2011) ‘Neural network computation with DNA strand displacement cascades’, *Nature*, 475(7356), pp. 368–372. doi:10.1038/nature10262.