How can digital data stored as DNA be manipulated?

- Data can be encoded as DNA but are difficult to process thereafter.
- A new method enables operations to be performed on DNA-encoded data directly, without having to first translate them into their electronic equivalent.

DNA can be used to reliably store a vast amount of digital data. However, retrieval or manipulation of specific data encoded in these molecules has hitherto been difficult. Now, researchers from the CNRS and the University of Tokyo have pioneered the application of a new method that harnesses enzymes, offering initial clues as to how these technical obstacles may be overcome. Their findings are reported in *Nature*, the 20 October 2022.

Nature has indisputably invented the best solution for storing a massive amount of data: DNA. This understanding has inspired the use of DNA for the storage of digital data, converting binary (0 or 1) numbers into one of the four different DNA "letters" (A, T, C, or G).

But how does one find a specific datum in the library of information stored as DNA? And how can calculations with DNA-encoded data be performed directly, without first converting them back into electronic data? These are the questions that teams from the LIMMS (CNRS / University of Tokyo) and Gulliver (CNRS / ESPCI) research laboratories have sought to answer. They are testing a new approach using enzymes and applying the solutions of artificial neurons and neural networks for direct operations on DNA data.

Specifically, the researchers have made use of the reactions of three enzymes to design chemical "neurons" that reproduce the network architecture and ability for complex calculations exhibited by true neurons. Their chemical neurons can execute calculations with data on DNA strands and express the results as fluorescent signals.

The LIMMS and Gulliver teams have also innovated by assembling two layers of the artificial neurons in order to refine calculations. Precision is further enhanced through microfluidic miniaturization of reactions, allowing tens of thousands to take place.

The fruit of a decade of cooperation between French biochemists and Japanese microfluidics engineers, these breakthroughs may eventually permit better screening for certain diseases as well as the manipulation of gigantic DNA-encoded databases.

When kept away from water, air, and light, DNA can be preserved for hundreds of thousands of years, without any energy input. And stored in a capsule a few centimetres in diameter, it can hold up to 500

terabytes of digital data. By 2025, the total volume of digital data generated by humans is expected to reach 175 zettabytes.[1] Since current storage media are relatively bulky, fragile, and energy-intensive, DNA may provide a viable alternative—able to contain all existing data within the space of a shoebox. Facilitating DNA storage will be the PEPR MoleculArxiv goal, a prority research program provided last May by the CNRS.

Notes

[1] A zettabyte is 10^{21} bytes—i.e., one billion trillion.

Bibliography

Nonlinear decision-making with enzymatic neural networks. S. Okumura, G. Gines, N. Lobato-Dauzier, A. Baccouche, R. Deteix, T. Fujii, Y. Rondelez & A. J. Genot. Nature, 20 october 2022. DOI : 10.1038/s41586-022-05218-7

Contacts

CNRS researcher | Anthony Genot **| T +81 3 5452 6213 |** genot@iis.u-tokyo.ac.jp

CNRS Press Officer | Ouns Hamdi | **T +33 1 44 96 43 90** | ouns.hamdi@cnrs.fr