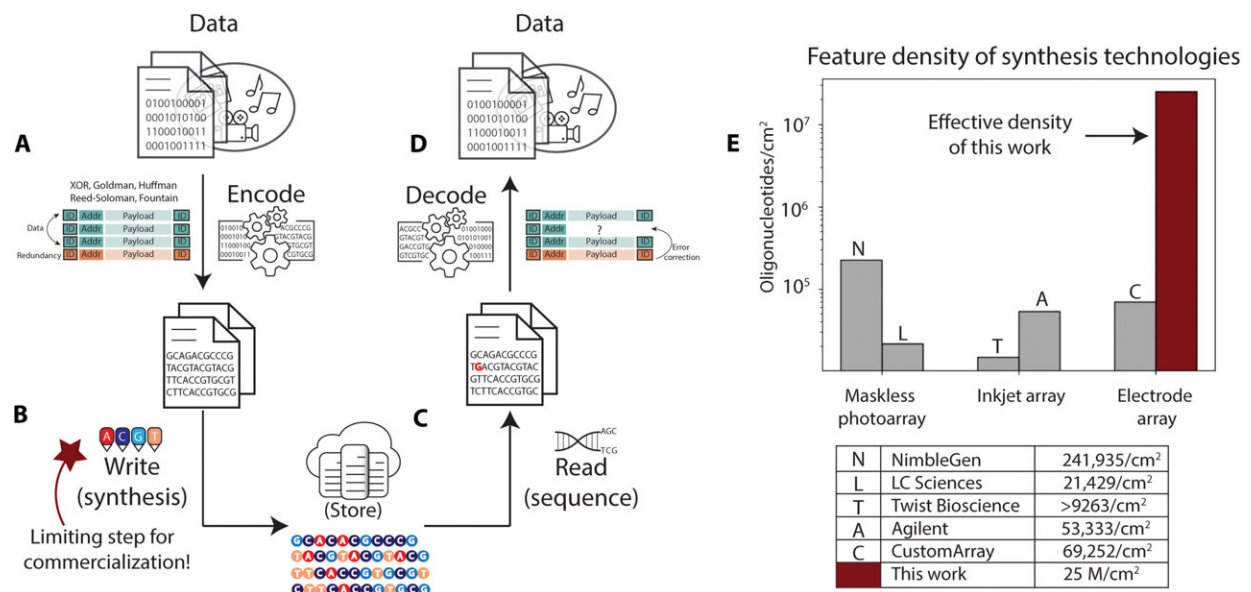


Storing information in DNA: Improving DNA storage with nanoscale electrode wells

December 7 2021, by Thamarasee Jeewandara



DNA data storage requires higher synthesis throughput than is possible with current techniques. (A to D) Overview of the DNA data storage pipeline. (A) Digital data are encoded from their binary representation into sequences of DNA bases, with an identifier that correlates them with a data object, addressing information that is used to reorder the data when reading, and redundant information that is used for error correction. (B) These sequences are synthesized into DNA oligonucleotides and stored. (C) At retrieval time, the DNA molecules are selected and copied via PCR or other methods and sequenced back into electronic representations of the bases in these sequences. (D) The decoding process takes this noisy and sometimes incomplete set of sequencing reads, corrects for errors and missing sequences, and decodes the information to recover the data. (E) Summary of the commercial synthesis processes and corresponding estimated oligonucleotide densities, as reported in

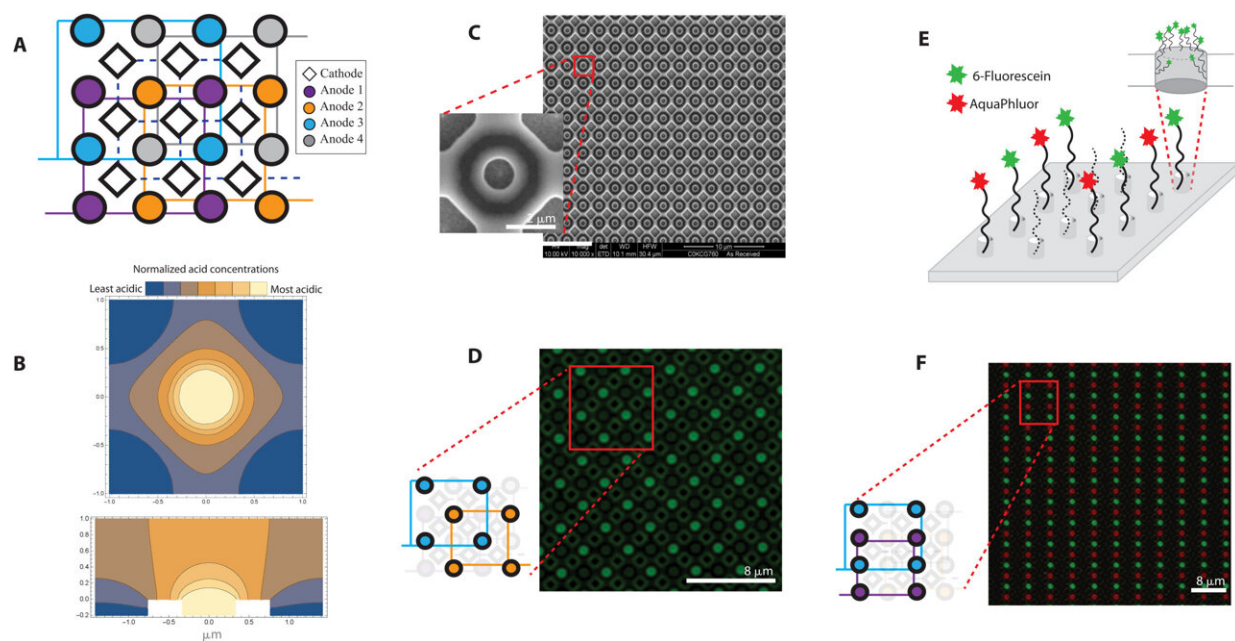
the literature or by the companies themselves. Our electrochemical method density is highlighted in dark red. Credit: *Science Advances*, 10.1126/sciadv.abi6714

Geneticists can store data in synthetic DNA as a medium for long-term storage due to its density, ease of copy, longevity and sustainability. Research in the field had recently advanced with new encoding algorithms, automation, preservation and sequencing. Nevertheless, the most challenging hurdle in DNA storage deployment remains the write throughput, which can limit the data storage capacity. In a new report, Bichlien H. Nguyen, and a team of scientists in Microsoft Research and computer science and engineering at the University of Washington, Seattle, U.S., developed the first nanoscale DNA storage writer. The team intended to scale the DNA write density to 25×10^6 sequences per square centimeter, an improved storage capacity compared to existing DNA synthesis arrays. The scientists successfully wrote and decoded a message in DNA to establish a practical DNA data storage system. The results are now published in *Science Advances*.

Long-term DNA archives

The current pace of data generation exceeds existing [storage](#) capacities, DNA is a promising solution to this problem at an expected practical density of more than [60 petabytes per cubic centimeter](#). The material is [durable under a range of conditions](#), relevant and [easy to copy](#), with promise to be [more sustainable or greener than commercial media](#). During the process, digital data in the form of sequences of bits can be encoded in sequences of the [four natural DNA bases](#)—guanine, adenine, thiamine and cytosine, although additional bases are also possible. The team can next write the sequences into molecular form via de novo DNA oligonucleotide synthesis to create specific molecules based on a set of

repeating chemical steps. The resulting oligonucleotides can be preserved and stored after synthesis. To access the data, the DNA storage can be amplified using [polymerase chain reactions](#) and sequenced to return the DNA base sequences to the digital domain, then the DNA base sequences can be decoded to recover the original sequence of bits.



Overview of 650-nm array pitched 2 μm . (A) Finite element analysis of anodic acid generation and diffusion at a 650-nm-diameter electrode with a 200-nm well is depicted with a cross-sectional view along the $y = x$ plane and (B) top-down view on the $z = 0$ plane. The colors blue and yellow represent regions with relatively low and high acid concentrations, respectively. (C) An overview of the nanoscale DNA synthesis array with scanning electron microscopy images of the 650-nm electrode array and enlarged view of one electrode. (D) A fluorescent image in which the well surrounding each activated anode is patterned with AAA-fluorescein. The cartoon diagram depicts which electrodes in the layout were activated. (E) Illustration of the wells patterned with AAA-fluorescein and AAA-AquaPhluor and (F) corresponding image overlay of the two fluorophores at the end of DNA synthesized on the same 650-nm electrode array. Credit: *Science*

A new method for synthetic DNA data storage

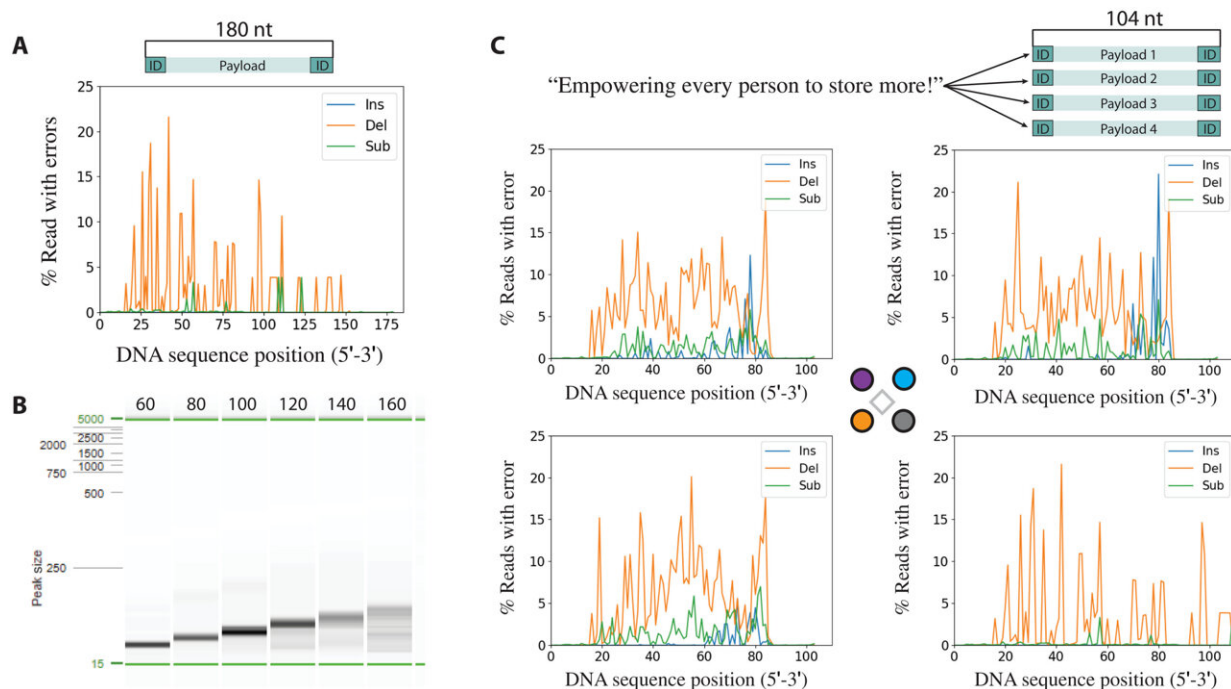
In this study, Nguyen et al. produced an [electrode array](#) which demonstrated independent [electrode](#)-specific control of DNA synthesis with electrode sizes and pitches to establish synthesis density of 25 million oligonucleotides per cm^2 . This value is estimated as the electrode density required to achieve the minimum target of [kilobytes per second of data storage in DNA](#). The team pushed the state-of-the-art in electronic-chemical control and provided experimental evidence to the [write bandwidth](#) necessary for DNA data storage.

The team introduced a proof-of-concept molecular controller in the form of a tiny DNA storage writing mechanism on a chip. The chip could tightly pack DNA synthesis at 3-orders of magnitude higher than before to achieve greater DNA writing throughput. To store information in DNA at the scale necessary for commercial use required two crucial processes. First the team had to translate digital bits (ones and zeros) into strands of synthetic DNA representing bits with encoding software and a DNA synthesizer. Then they must be able to read and decode the information back to its bits to recover that information into digital form again with a DNA sequencer and decoding software.

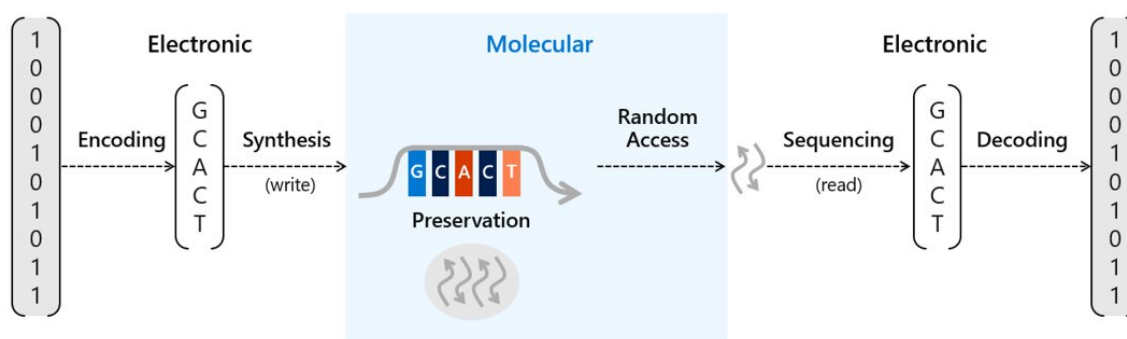
Developing electrochemical arrays for nanoscale features

During the traditional synthesis of DNA chains, scientists use a multistep method known as [phosphoramidite chemistry](#), in which a DNA chain can be grown sequentially by the addition of DNA bases. Each DNA base contains a blocking group to prevent multiple additions of DNA bases to the growing chain. On attachment to a DNA chain, acid can be

delivered in the setup to cleave the blocking group and prime the DNA chain to add the next base. During [electrochemical DNA synthesis](#), each spot in the array contains an electrode and when a voltage is applied, acid is generated at the working electrode (anode) to deblock the growing DNA chains, while an equivalent base is generated at the counter electrode (cathode). The team prevented acid diffusion in the setup by designing an electrode array, where each working electrode around which acid formation occurred during DNA synthesis was sunk in a well, and surrounded by four common counter electrodes, i.e., cathodes that drove base formation, to confine the acid to specific regions. Nguyen et al. verified the effectiveness of the design using [finite element analysis](#). During the experiments, when presented in sufficient concentration, the acid deblocked the surface-bound nucleotides to allow the next nucleotide to couple. Using the setup of chips containing feature spots to confine acids, they developed electrochemical arrays with four individual electrodes to regulate DNA synthesis. The team then performed experiments with two fluorescently labeled bases in green and red. As proof of concept, they showed the device's capacity to write data by synthesizing four unique DNA strands, each 100 bases long with an encoded message, without errors.



Errors stemming from synthesis followed by sequencing. (A) Insertions (Ins), deletions (Del), and substitutions (Sub) per position for a synthesized and PCR-amplified 180-base sequence. (B) Electrophoresis image of synthesis products after PCR amplification. (C) Message encoded into 64 bytes split into four unique sequences of 104 bases (top). Insertions, deletions, and substitutions per locus of each of the four sequences in the multiplex synthesis run. In every error analysis graph, the terminal 20 bases at both 3' and 5' ends come from the primers used in PCR and are not representative of the synthesized errors. Credit: *Science Advances*, 10.1126/sciadv.abi6714



Scaling DNA Data Storage with Nanoscale Electrode Wells. Tiny DNA storage writing mechanism on a chip. Credit: Microsoft Research Blog, *Science Advances*, 10.1126/sciadv.abi6714

Outlook: Synthesizing short oligonucleotides on the electrode array for data storage

Using the setup, Nguyen et al. also demonstrated spatially controlled synthesis of short oligonucleotides on the electrode array to assess the maximum length of DNA that could be formed. The scientists created a single DNA sequence with 180 nucleotides and PCR-amplified various length products from the complete length of the oligonucleotides. As the [amplicon](#) got longer, the expected PCR products appeared fainter and less well defined, while shorter amplicons showed stronger and more well-defined bands indicative of higher [synthesis](#) errors. Based on the results, the researchers selected sequence length accounting to 100 bases for ease of purification to provide a practical demonstration of DNA data storage without further optimization. In this way, the proof-of-concept method demonstrated in this work by Bichlien H. Nguyen and colleagues paved the way forward to generate large-scale and unique DNA sequences [in parallel for data storage](#). The work outpaced [previous](#)

[reports](#) on dense synthetic DNA sequences to provide a first experimental indication to achieve the write bandwidth required for data storage at nanoscale feature sizes. The scientists expect immediate applications of the devices in information technology and foresee their practical applications in materials science, synthetic biology and large-scale molecular biology assays.

More information: Bichlien H. Nguyen et al, Scaling DNA data storage with nanoscale electrode wells, *Science Advances* (2021). [DOI: 10.1126/sciadv.abi6714](https://doi.org/10.1126/sciadv.abi6714)

Goldman et al, Towards practical, high-capacity, low-maintenance information storage in synthesized DNA, *Nature* (2013). [DOI: 10.1038/nature11875](https://doi.org/10.1038/nature11875)

© 2021 Science X Network

Citation: Storing information in DNA: Improving DNA storage with nanoscale electrode wells (2021, December 7) retrieved 19 April 2024 from <https://phys.org/news/2021-12-dna-storage-nanoscale-electrode-wells.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
