MAXIMIZING DATA'S POTENTIAL

Assembly of Data-encoding DNA Fragments via DNAzymes to Reduce the Cost of DNA Data Writing

Gemma Mendonsa

Mengdi Bao, Brett Herdendorf, Sriram Chari, Anil Reddy

June 19, 2025
SC-DNA-2025

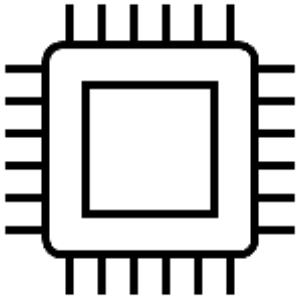
SEAGATE

Why DNA data storage?

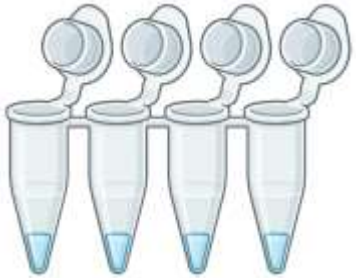
Volumetrically dense

Low power consumption

Stable



Parallel compute capability



Easy to replicate

1.Nadis, S. (2021). 'A New Light on DNA Storage'. *Harvard Magazine*, (March-April). Available at: https://www.harvardmagazine.com/2021/03/right-now-dna-storage.

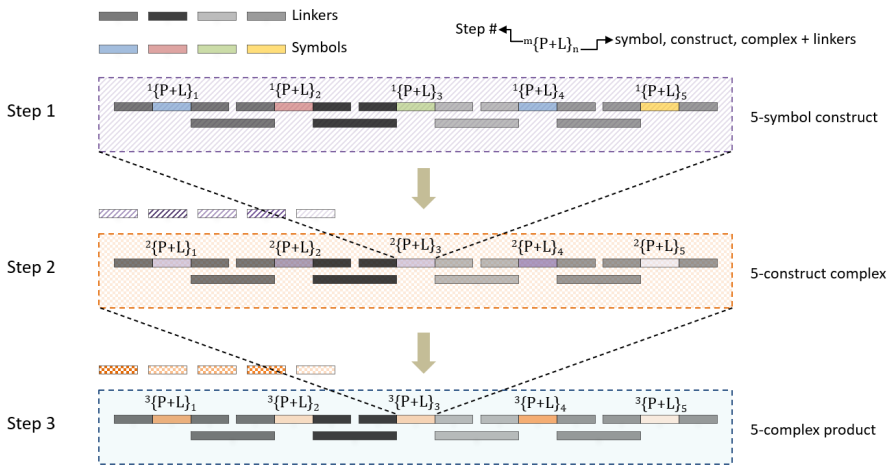
Why not DNA data storage?

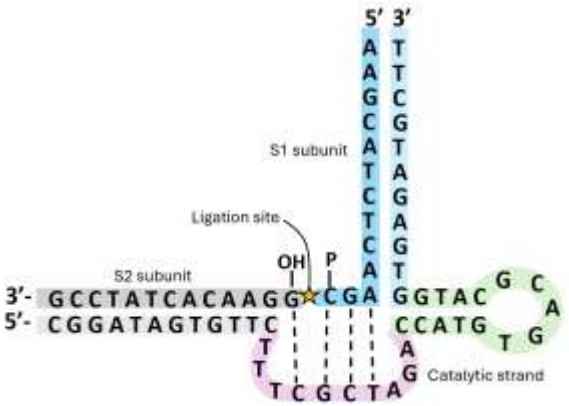
Slow write speeds

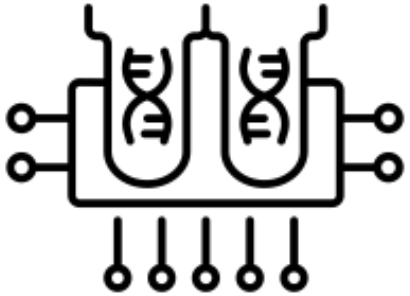
Expensive write

Large volumes of reagents required

Our solutions to the challenges







Parallelization with DNA assembly to increase write speed[1]

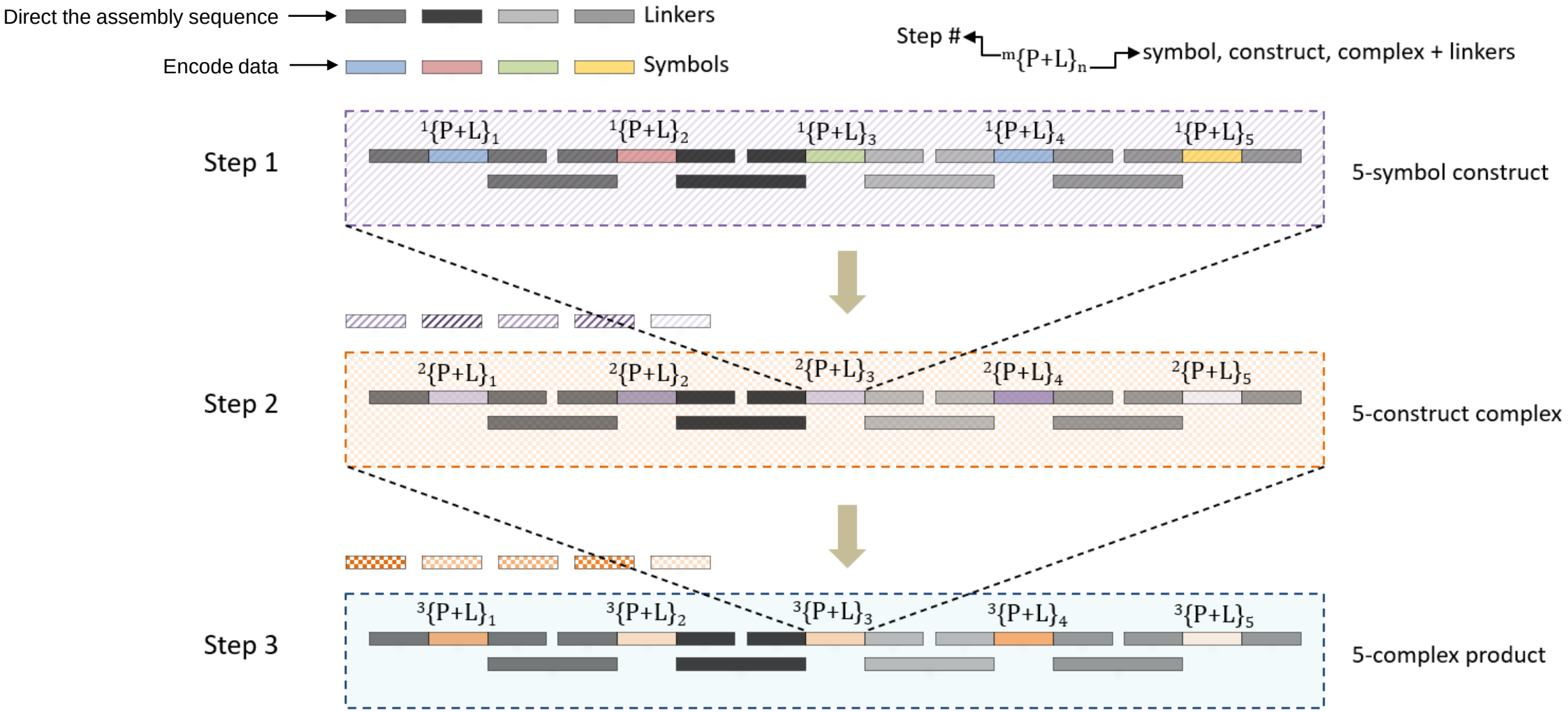
DNAzyme chemistry to reduce write cost[1]

Lab-on-a-chip to scale down reagent volumes[2]

1. Mendonsa, Gemma, Sriram Chari, Mengdi Bao, Brett Herdendorf, and Anil Reddy. "Directed assembly of single-stranded DNA fragments for data storage via protein-free catalytic splint ligation." Nucleic Acids Research (2025): In press. DOI: https://doi.org/10.1093/nar/gkaf582

2. Bao, Mengdi, Brett Herdendorf, Gemma Mendonsa, Sriram Chari, and Anil Reddy. "Low-cost and automated magnetic bead-based DNA data writing via digital microfluidics." Lab on a Chip 25, no. 8 (2025): 2030-2042.

DNA assembly to increase write speed and reduce cost

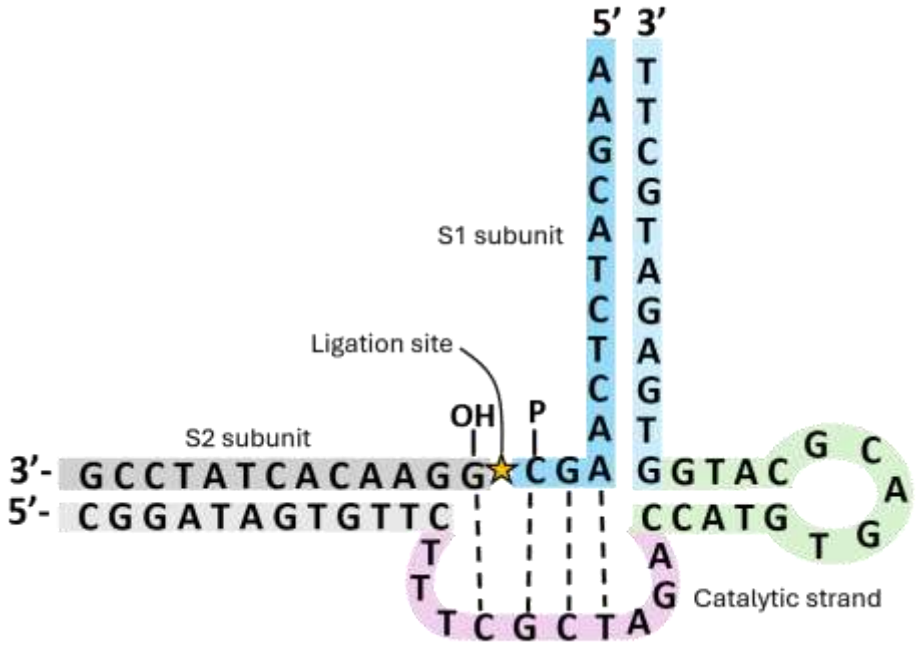


Synthesizing oligos in bulk makes them cheap, but oligo assembly introduces a new cost: proteins used for assembly.

DNA assembly with DNAzymes

DNAzymes are much cheaper than the proteins usually used to assemble DNA.

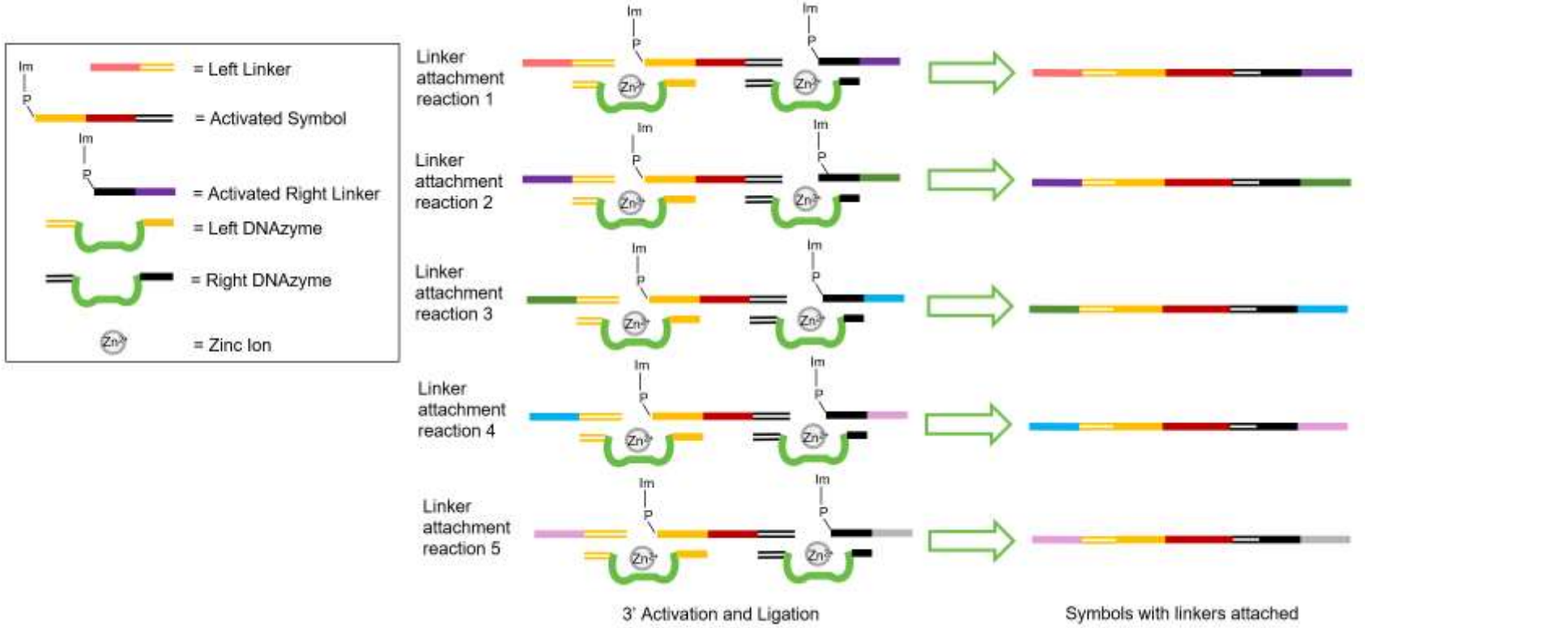
- The E47 DNAzyme joins two short strands of DNA to form one longer strand.
- The S1 and S2 subunit strands (left) are joined/ligated by the catalyst strand in the presence of Zn^{2+} or Cu^{2+}.
- The S1 subunit must be “activated” with imidazole+EDC before ligation can occur.
- The product is a single strand S1+S2.



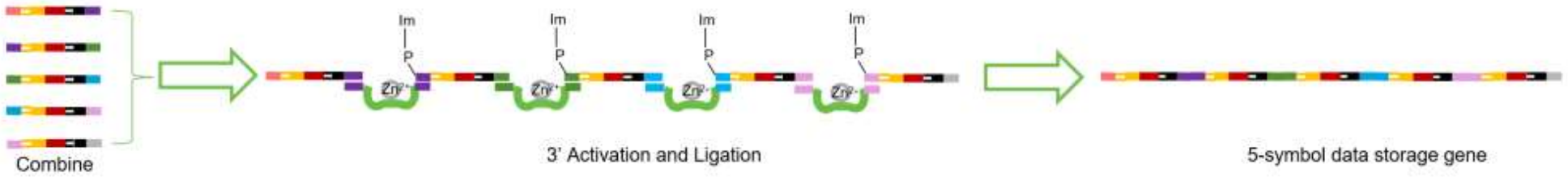
Cuenoud, B. and Szostak, J.W., 1995. A DNA metalloenzyme with DNA ligase activity. *Nature*, *375*(6532), pp.611-614.

Assembly of prefabricated symbols and linkers

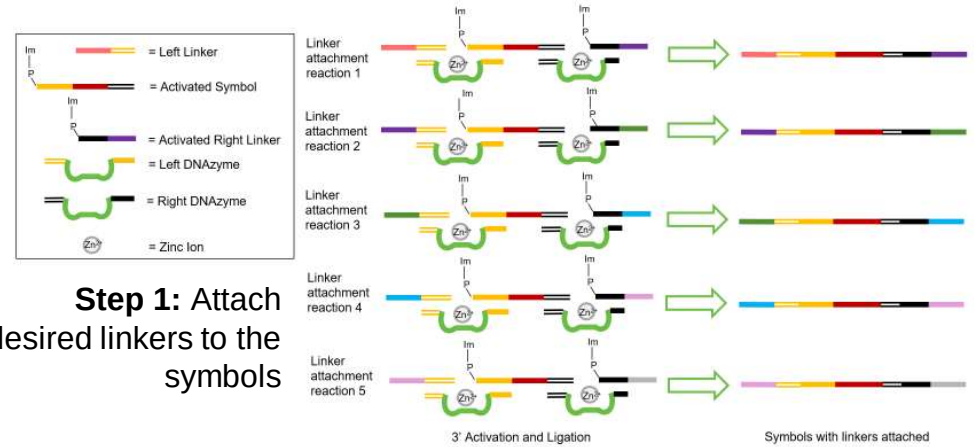
Step 1: Attach desired linkers to the symbols



Step 2: Assemble the long strand via the linker sequences



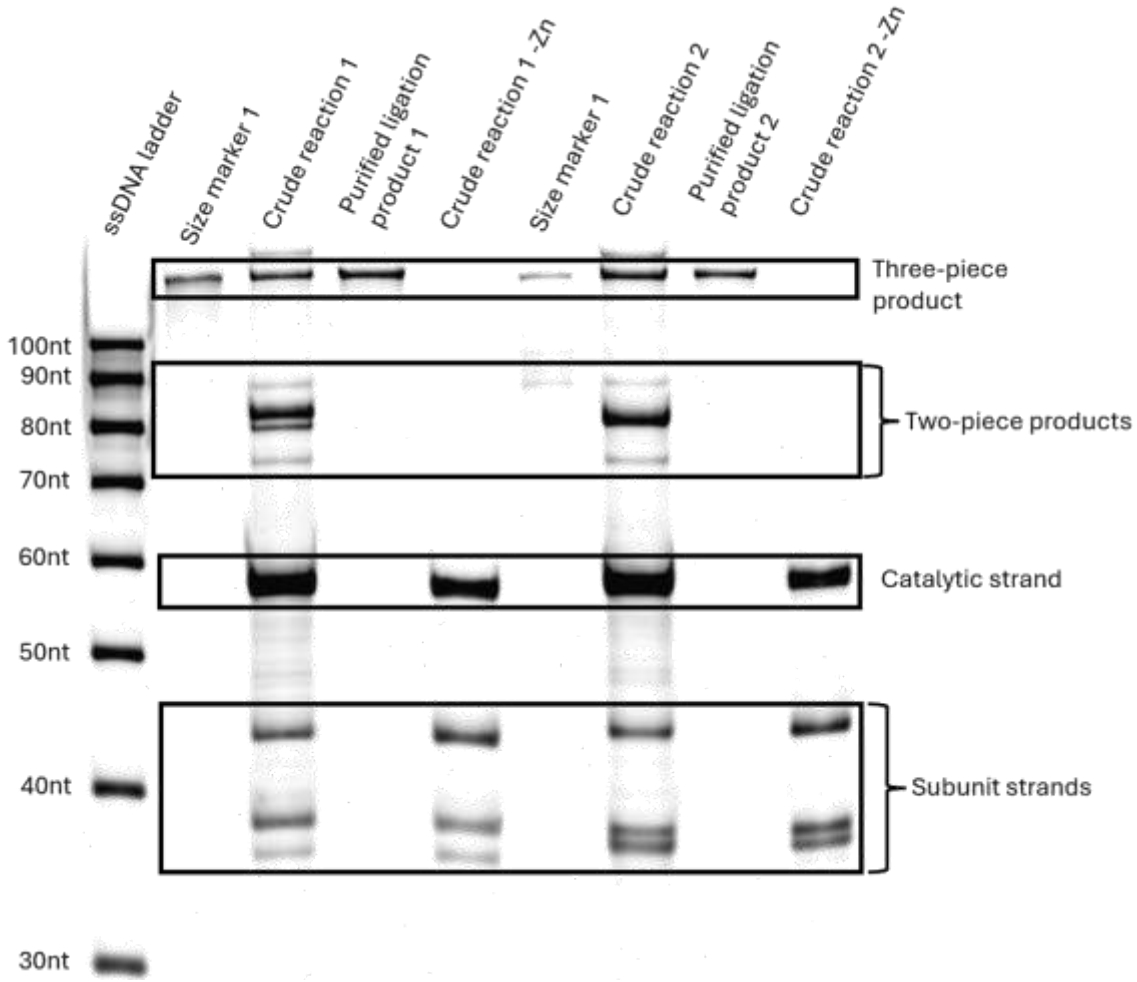
Step 1: Attaching symbols to linkers



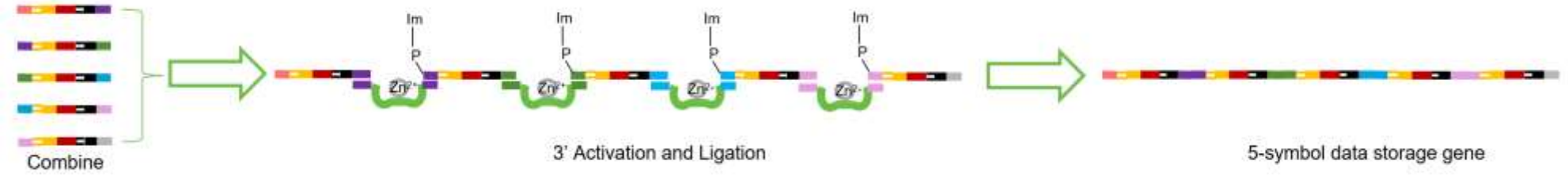
Step 1: Attach desired linkers to the symbols

The order of the data-encoding symbols is determined by the linkers attached to each symbol. In the first step of assembly, two linkers are attached to each symbol.

The symbol+linker complexes were assembled with DNAzymes. The three-piece assembly products could be clearly seen with gel electrophoresis.



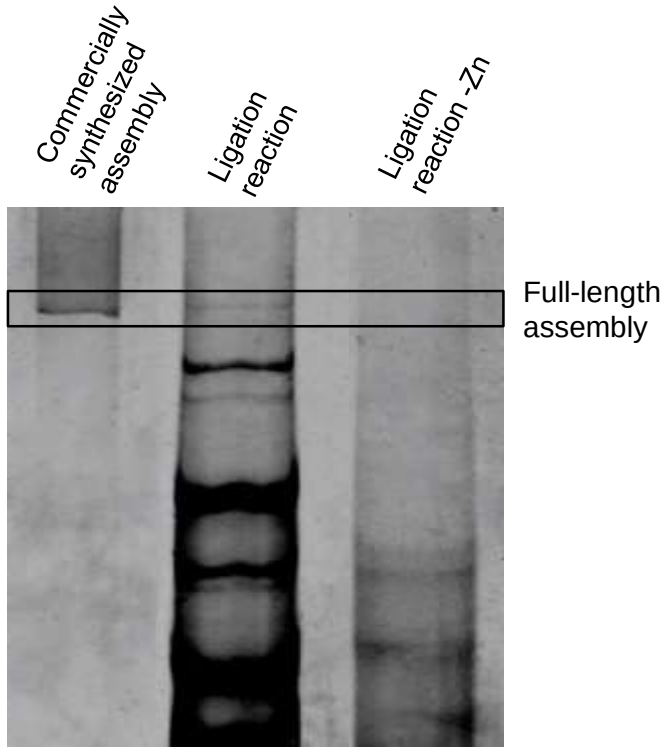
Step 2: Assembling the final gene



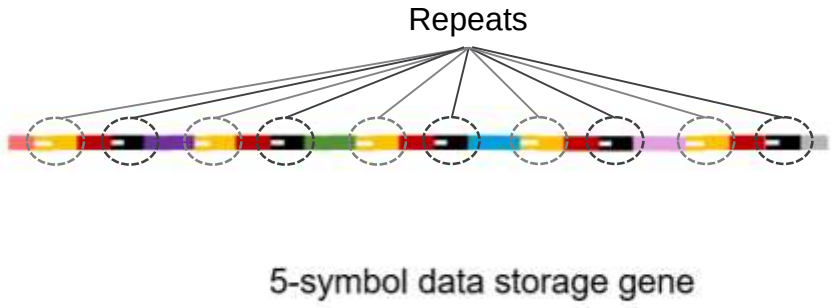
A second DNAzyme assembly reaction was used to join five symbols together via their linker ends. DNAzymes were designed to specifically join the linkers in the correct order.

The yield of the 5-piece assembly was low (~0.08%), but full-length product was present.

Full-length product was purified and PCR amplified.



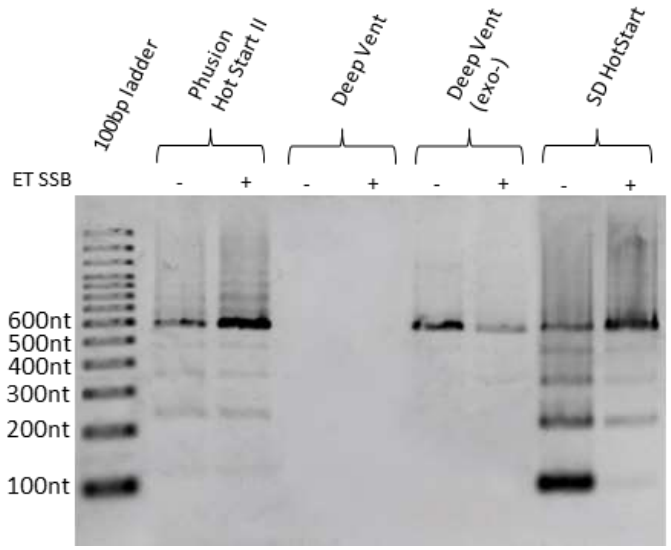
PCR of a DNA strand containing repeats

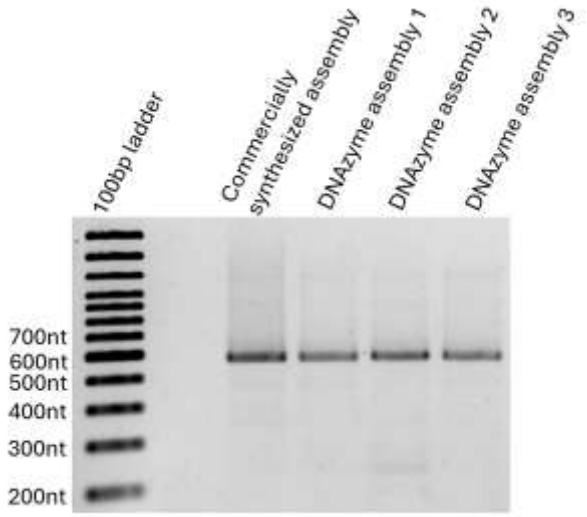


Initial PCR amplifications resulted in many bands, likely due to the repeat regions in the full-length assembly.

Switching to a strand displacing polymerase DeepVent (exo-) reduced the “laddering” effect greatly.

Three different 5-symbol data storage genes were assembled and amplified with Polymerase Chain Reaction (PCR).



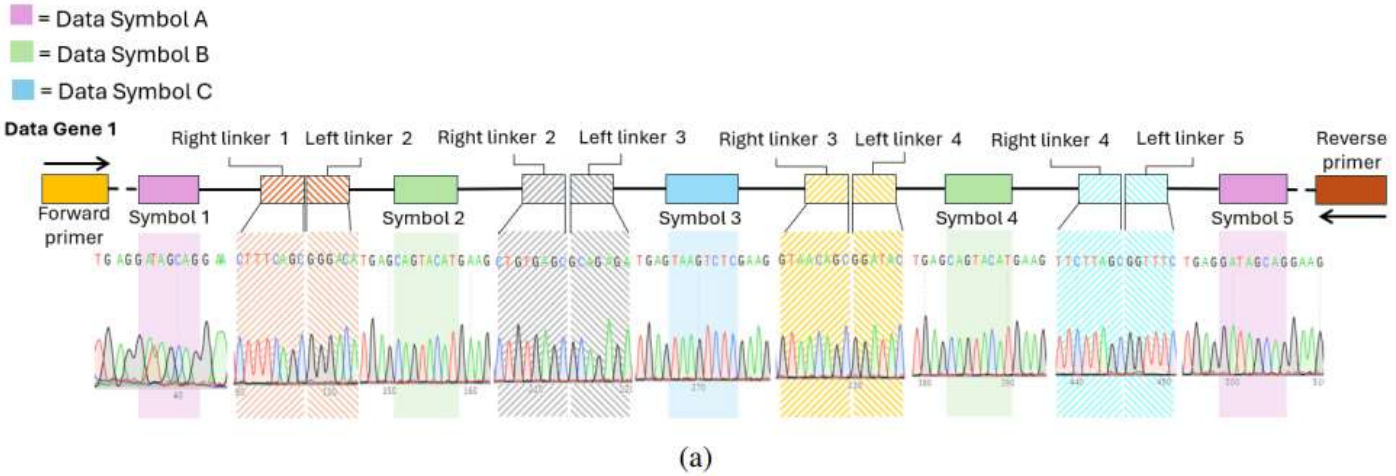


Assemblies confirmed with DNA sequencing

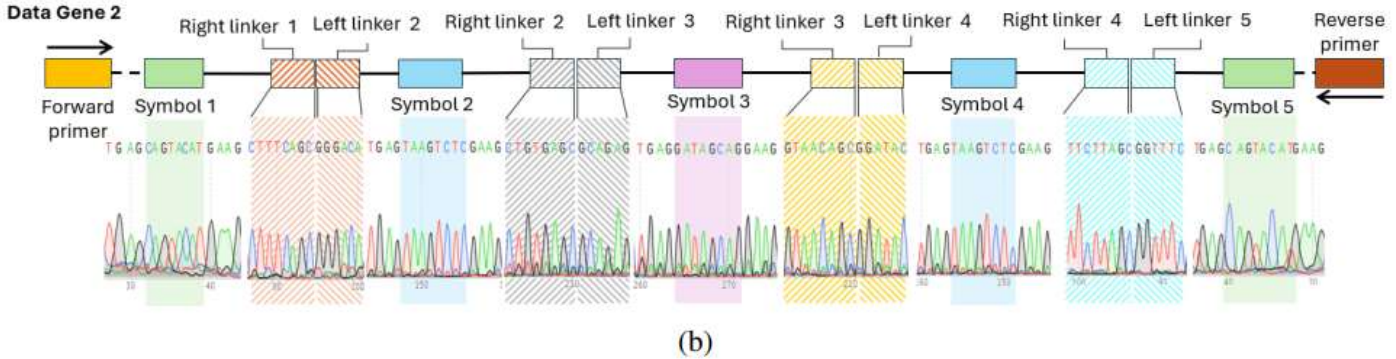
Sanger sequencing was used to confirm the sequences of the assemblies.

Oxford Nanopore sequencing was also used to measure variants/errors.

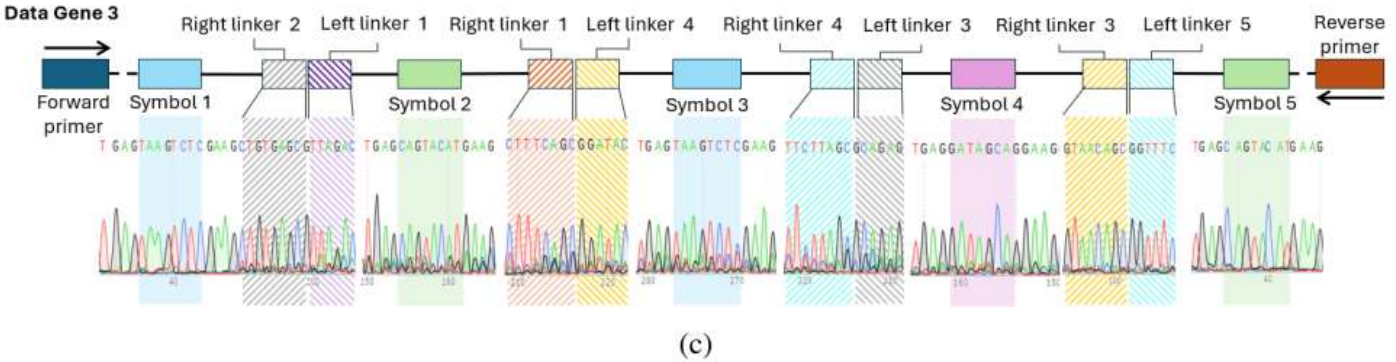
All genes were assembled as expected. No variants (errors) were detected.



(a)



(b)



(c)

Sample alias	Barcode	Reads	Bases	Median read length	Amplicons	Unmapped	Variants (indels)
Assembly_1	barcode01	15.0 k (33%)	7.5 M (33%)	500	1	1 (0%)	0 (0)
Assembly_2	barcode02	15.0 k (33%)	7.5 M (33%)	504	1	57 (0%)	0 (0)
Assembly_3	barcode03	15.0 k (33%)	7.6 M (33%)	492	1	4 (0%)	0 (0)

NAR article coming soon

Article in press at Nucleic Acids Research:

DOI: https://doi.org/10.1093/nar/gkaf582

Contributors: Gemma Mendonsa, Sriram Chari, Mengdi Bao, Brett Herdendorf, Anil Reddy

For more information, contact: gemma.mendonsa@seagate.com

Other work: Automation with digital microfluidic lab-on-a-chip

Lab on a Chip

ROYAL SOCIETY OF CHEMISTRY

PAPER

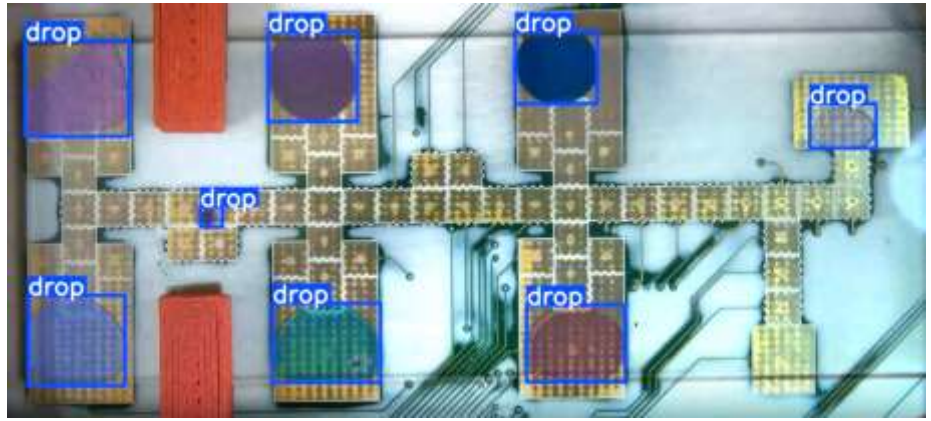
View Article Online
View Journal

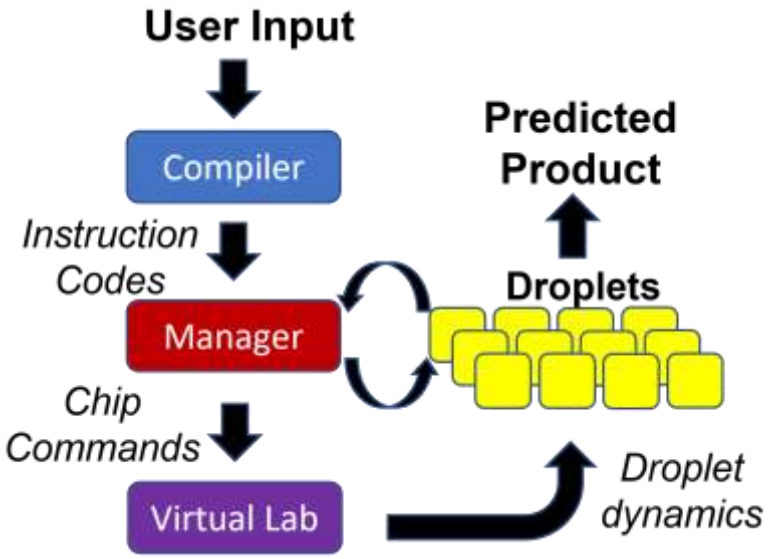
Check for updates

Cite this: DOI: 10.1039/d5lc00106d

Low-cost and automated magnetic bead-based DNA data writing *via* digital microfluidics†

Mengdi Bao, ‡* Brett Herdendorf,‡ Gemma Mendonsa, Sriram Chari and Anil Reddy





DOI: 10.1039/D3DD00083D (Paper) *Digital Discovery*, 2023, **2**, 1436-1451

Automated routing of droplets for DNA storage on a digital microfluidics platform

Ajay Manicka [a], **Andrew Stephan** [a], **Sriram Chari** [b], **Gemma Mendonsa** [b], **Peyton Okubo** [a], **John Stolzberg-Schray** [a], **Anil Reddy** [b] **and Marc Riedel** [a]

[a] *Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota, USA. E-mail: ececomm@umn.edu*

[b] *Seagate Technology, USA*

Future Work

- Improving yield and reaction rate
  - Novel DNAzymes
  - Optimization of DNAzyme sequence and reaction conditions
- Simplifying the steps for automation
  - Simplified purification steps for isolating full-length assemblies
  - Stabilized activated intermediate to enable pre-activation of subunits
- Performing additional tiers of assembly

Thank you!