scientific reports

Check for updates

OPEN

DNA conjugation on functionalized plastic surfaces for sequential, iterative single molecule sequencing

Sharmili Roy, Hanlee P. Ji✉ & Billy T. Lau✉

We developed a method for repeated and sequential retrieval of arbitrary DNA elements. A click chemistry process was used to conjugate the DNA molecules onto a plastic surface within the interior of microcentrifuge tubes. For this study, we utilized synthetic DNA sequences that encode arbitrary data and designed PCR primers for amplification. Specifically, the DNA was tailed with trans-cyclooctene (TCO) and was then conjugated to plastic surfaces functionalized with methyltetrazine (MTz). The covalent DNA attachment to the plastic surface enables repeated and non-destructive polymerase-based copying and amplification of the original source molecules. In summary, we demonstrate a new type of DNA storage media with the property of long-term stability and ability to read different groups or files of DNA data. For this proof-of-concept study, we demonstrate the key features of this technology including: (1) characterization of the DNA conjugation process using control strands, (2) conjugation kinetics, and (3) amplification and sequencing of specific DNA data elements over multiple retrieval operations from the same plastic surface. Specifically, we showed that different DNA "file groups" with information could be accessed multiple times with single molecule sequencing. We also compared the performance metrics of PCR versus recombinase polymerase amplification (RPA). Our results indicate that RPA has better performance metrics in terms of reducing contamination and improved yield of retrieved data from across sequential experiments. Overall, we demonstrated a plastic-based DNA storage system for robust and reliable iterative, targeted DNA retrieval, and sequencing.

DNA is a highly stable molecule and resists degradation over time. Researchers have focused on DNA's molecular properties as a material for encoding any type of arbitrary data with the benefit of long term, highly stable storage[1,2]. Thus, manipulating and sequencing of DNA molecules provides a scalable writing and reading process for storing data, similar to a computer's external storage system[3]. There are multiple approaches for DNA-based encoding of information, synthesis of the DNA data elements, retrieval of the files and optimization of the sequencing technologies[4–6]. However, there is relatively little published work on developing a physical media format that provides long-term storage of DNA data[4,7,8]. One major challenge is that DNA data retrieval operations (i.e., PCR and sequencing) consumes the original DNA source. Hence, the repetitive use of a DNA media format for data storage exhausts the DNA encoding molecules. The ideal material substrate for such a device would be one where DNA elements could be readily and sequentially retrieved in a compact and portable system over the course of many iterations[9,10].

For this proof-of-concept study, we demonstrated the feasibility and robust performance of a DNA storage system using the plastic surface of PCR tubes as a solid phase support. This system involves the chemical conjugation of DNA molecules to the surface of PCR tubes. First, pools of DNA oligonucleotides are synthesized with sequences corresponding to arbitrary desired data. Once the DNA is conjugated, one can retrieve and sequence the individual DNA data elements. Plastic PCR tubes are routinely used in molecular biology applications and thus can be manipulated with standard lab equipment such as micropipettes and thermocyclers. This feature could enable facile adoption of this methodology given the limited equipment requirements. The DNA storage process also allows one to repeatedly read the original DNA molecules encoding the data[11–13]. To generate DNA templates amenable for plastic surface conjugation, DNA in the form of an oligonucleotide pool is enzymatically tailed with trans-cyclooctene functionalized dUTP – this click chemistry modification allows

Division of Oncology, Department of Medicine, Stanford University School of Medicine, CCSR 1115, 269 Campus Drive, Stanford, CA 94305-5151, USA. ✉email: genomics_ji@stanford.edu; billylau@stanford.edu

the DNA fragment to be covalently conjugated to a methyltetrazine-modified (**MTz**) PCR tube surface. The conjugation reaction occurs rapidly, thus providing a quick solution for generating DNA storage media[10].

In this study we demonstrated several innovations. The PCR tube format provided a convenient way of accessing sequence information from the same DNA substrate and the capacity for repeated queries for different DNA data elements. We also tested the use of recombinase polymerase amplification (**RPA**) as an isothermal amplification method for DNA element retrieval and compared it to PCR. Using RPA is easier for amplifying specific target DNA molecules without the need for a thermocycler and changing temperatures[14]. We also combined this method of data retrieval with nanopore DNA sequencing. This feature provides rapid and near real-time sequencing with a more streamlined procedure compared to other next generation sequencing methods[15–18].

Results

Overview of a plastic-based DNA storage system

We developed a new process for iterative retrieval of DNA data from a functionalized plastic surface (Fig. 1). First, the DNA was functionalized by adding *trans-cyclooctene* (**TCO**) modified dUTPs with terminal transferase enzyme. The modified oligonucleotides were added to PCR tubes whose plastic surfaces were grafted with methyltetrazine (**MTz**) functional groups. Click chemistry reactions lead to rapid conjugation of the TCO-tailed DNA to the MTz-functionalized plastic surface in biocompatible buffers and reaction conditions. Stable covalent attachment of DNA fragments allowed iterative DNA polymerase amplifications from the same substrate (Fig. 1B). Full experimental details are provided in the Methods.

Testing of DNA conjugation to plastic tubes

To assess the performance of the DNA conjugation scheme, we used a DNA amplicon from lambda phage ('lambda DNA') as a control analyte. The TCO-conjugated lambda DNA was added into MTz-functionalized PCR tubes. As an initial test, we took one microliter of the supernatant conjugation solution (initially ranging from 0.1ng/ul to 3ng/ul) and measured its DNA concentration with high-sensitivity Qubit fluorimetry (Fig. 2A, Methods). The results from this assay provided the concentration of unconjugated DNA in solution. The concentration of the remaining DNA in solution after 24 h was too low to be reliably detected, suggesting that the conjugation of the DNA molecules onto the plastic surface was complete (Fig. 2B, Supplementary Table 1). There was no measurable DNA in the supernatant, suggesting the starting DNA was completely conjugated to the surface and the capacity for molecular attachment to the MTz-functionalized surface was in excess of the input DNA.

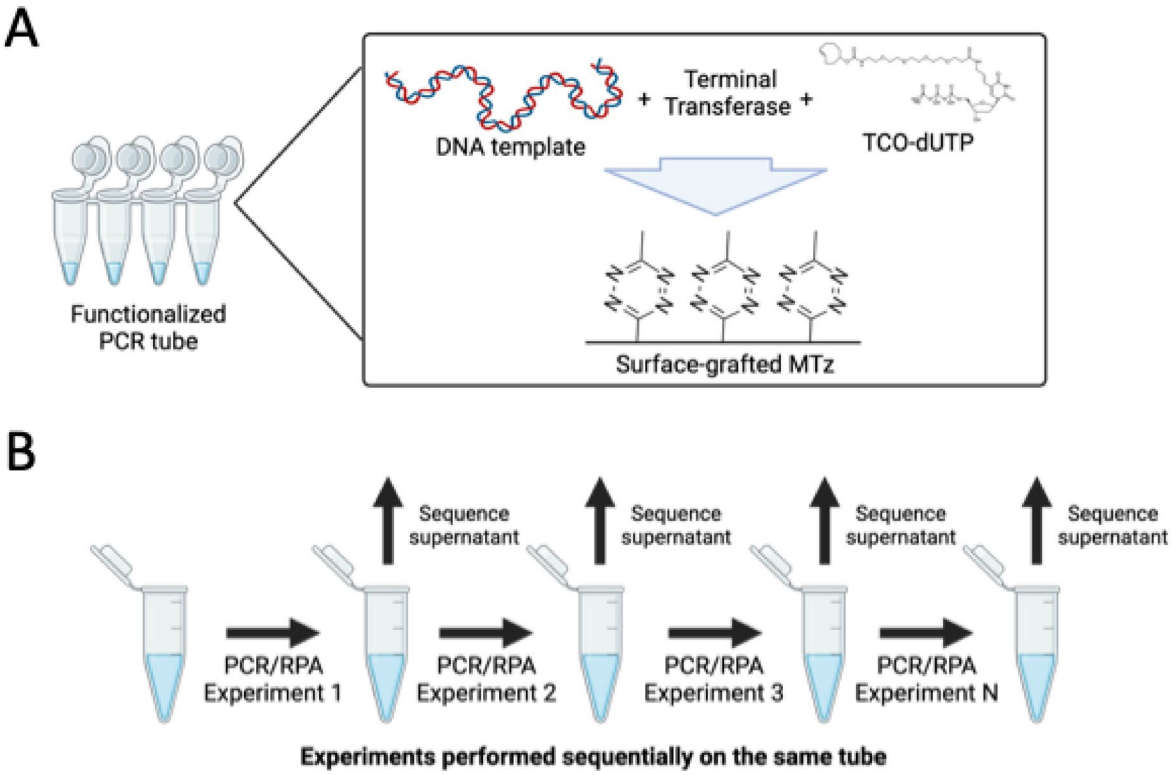

Fig. 1. Overview scheme of DNA conjugation and primer extension from plastic surfaces. (**A**) A TCO-labelled DNA sample library is attached to functionalized onto the surface of a MTz-functionalized plastic PCR tube. (**B**) Polymerase-based reactions such as PCR or RPA can be used to sequentially target DNA molecules of interest.

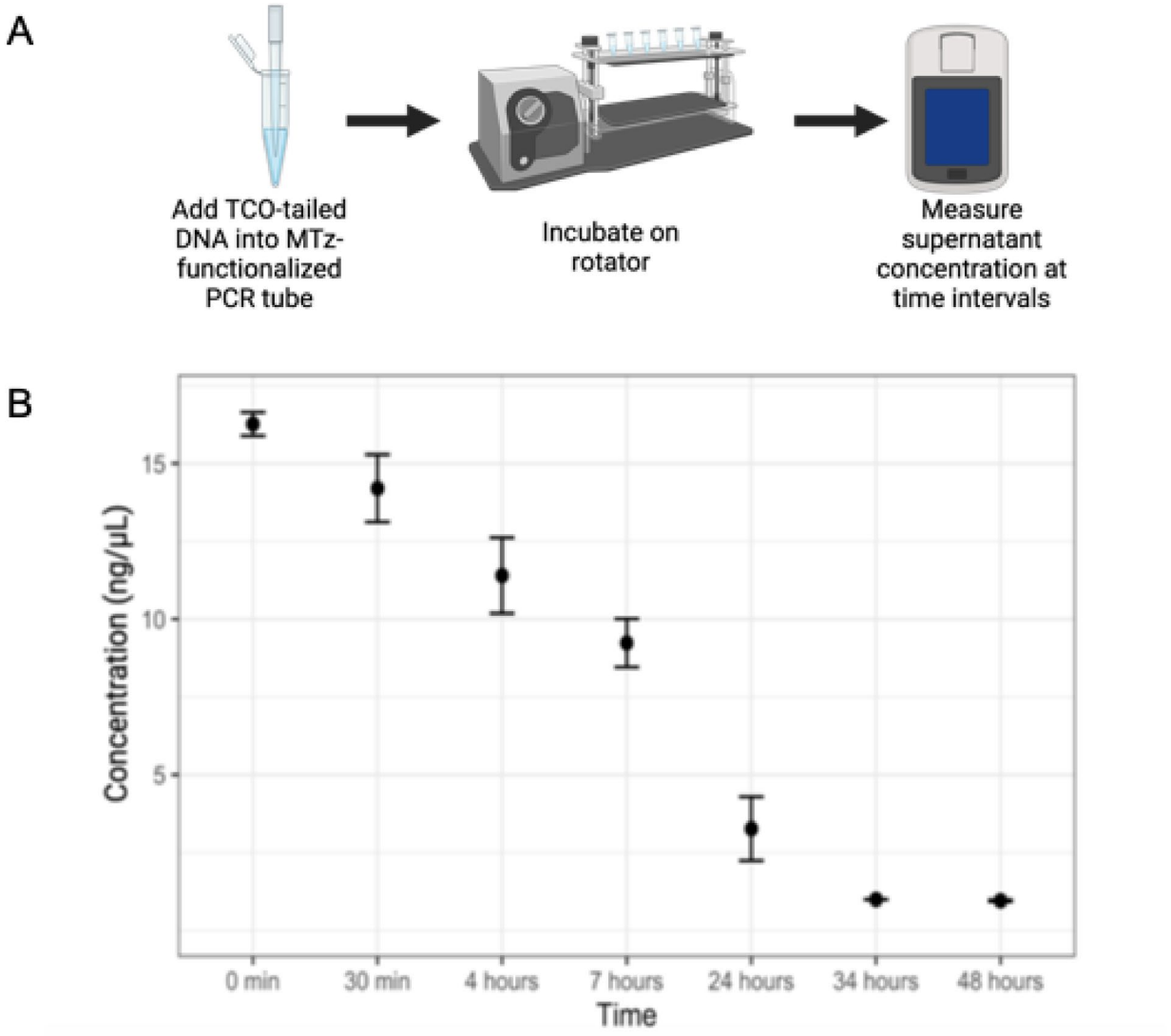

Fig. 2. Development of the APEX chemistry with tube-based system. (**A**) Experimental scheme of conjugation dynamics by measuring remaining amount of DNA in supernatant. (**B**) Concentration of DNA in supernatant versus conjugation time.

For improved resolution into the kinetics of DNA conjugation, we further increased the loading amount of TCO-labelled lambda DNA for conjugation onto MTz-functionalized plastic tubes. Utilizing ~350ng of starting material, we performed conjugation experiments in triplicate and quantified the remaining amount of unconjugated lambda DNA in solution. After 48 h, we observed that nearly 95% of the material was conjugated onto the surface. Considering 350ng of starting material and the lambda DNA fragment size of 3.6 kb, a loading capacity up to ~ 150 femtomoles of DNA template was tested on our platform. This number is subject to a variety of factors that would be further optimized in future studies, such as the total surface area of the plastic tube that is in contact with the DNA solution, number of TCO-dUTPs added by terminal transferase, and the size of DNA fragments to be conjugated.

Conjugation and PCR-based retrieval of DNA-based files

We used an established DNA oligonucleotide pool as the analyte for assessing our DNA conjugation and retrieval strategy[19]. The process of encoding data into DNA sequence was done as described in a prior work – we used these same sequences for the data elements[19]. Briefly, individual data files (i.e., specific text passages, etc.) were encoded as DNA sequences and then synthesized as an oligonucleotide pool. Each oligonucleotide incorporated a unique flanking sequence for polymerase amplification as well as the DNA sequence encoding specific data. For the starting DNA, we array-synthesized 15,000 individual oligonucleotides in a pooled format. Each oligonucleotide was approximately 160 bp average length. Each data file is represented by a number of oligonucleotides called 'file groups'. Each file group contained specific flanking sequences for PCR-based amplification retrieval and the entire set was retrievable using 18 different PCR primer pairs (Supplementary Table 2).

We conducted sequential DNA retrieval operations for each file group using PCR amplification (Fig. 1B, Supplementary Fig. 1). We started with amplification of File Group 1 from the DNA-conjugated PCR tube. After the amplification cycling was completed, the products were retrieved from the tube, a washing step was conducted and then the experiment was repeated in the same tube for File Group 2. Subsequently, this procedure was done for all 18 file groups with a total of 18 consecutive experiments on a single PCR tube with conjugated template DNA. The products from each amplification underwent library preparation and nanopore sequencing. The sequence reads were aligned to a reference containing all the synthesized oligonucleotide sequences to assess DNA retrieval accuracy. Importantly, the sequences of the oligonucleotides were specifically designed to accommodate error profiles specific to nanopore sequencing such that the sequence reads remain uniquely identifiable to their design. We observed that the sequence reads from each file group generally had a high alignment fraction ranging from 79% to 99%. For example, File Group 2 had an alignment fraction of 98% with only a low percentage of read coming from other file groups. This was the case for 16 of the file groups except for File Groups 10 and 13. This result was an indication that these two file groups did not amplify as efficiently or with high specificity (Fig. 3A). Several factors could be contributing to this lower fraction of reads including low specificity from the primers for those two groups.

Exonuclease removal of prior amplification products

We tested an enzymatic digestion method that we previously described for removing carryover amplification products from prior PCR reactions[19]. Removal of carryover amplification products is essential for sequential assays so that sequence reads from separate experiments do not contaminate each other. TCO-functionalized nucleotides are added to the starting DNA at the 3' end by terminal transferase and are thus protected from 3' exonuclease digestion once conjugated to the PCR tube surface. Exonuclease I and exonuclease III digest double-stranded and single-stranded DNA from the 3' end. Consequently, carryover DNA fragments from the previous amplification are degraded while the conjugated DNA remains protected. As a result, the same PCR tube can be used repeatedly to retrieve file groups of interest with reduced carryover production from the prior amplification.

After completing a file retrieval step, PCR tubes were incubated with a combination of exonuclease I and exonuclease III (Methods). We observed minimal carryover of previously amplified DNA through sequential experiments as shown by reads aligning to the previous file number (Fig. 3A, bottom right quadrant of heatmap). The lower part of the heatmap reveals a low presence of carry-over materials across the 18 sequential reactions – at most 1% carryover indicating that most of the off-target reads were due to low primer specificity rather than iterative experimental carryover.

Recombinase polymerase amplification improves retrieval yields

We observed a general trend where there was a decrease in PCR yield with as the number of iterative experiments increased in the same PCR tube (Fig. 3B). Upon querying File Group 18, the PCR-based retrieval yield was very low (approximately 0.1 ng/ul). The observed trend in these iterative PCR reactions suggested a steady rate of depletion throughout the process. This result may be an indicator of conjugated DNA loss from the tube and could also be due to thermal degradation of DNA when repeatedly performing PCR reactions.

As a potential solution to the decreasing yields with iterative PCR-based DNA retrieval, we tested a different file retrieval approach, namely using RPA which uses a recombinase enzyme for annealing primers to target DNA[20,21]. This is an isothermal process and commonly used in rapid diagnostic applications[22,23]. The amplification reaction can be conducted from room temperature up to 42 C and does not require a thermocycler like PCR[16]. This process concludes within 10–20 min, enabling faster file retrieval times[24,25].

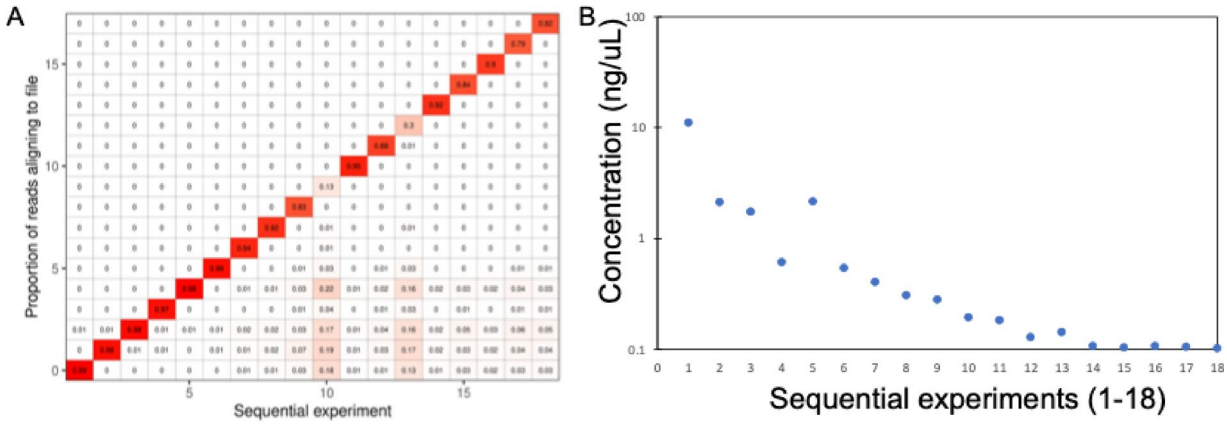

Fig. 3. Sequential amplification of DNA file groups from the same tube using PCR. DNA file groups were sequentially retrieved by targeted 18 different primer pairs. In between each experiment, the tubes were washed, after which another file group was targeted. (**A**) On-target alignment rate is shown for each files experiment on the same tube. (**B**) Amplification yield for each experiment.

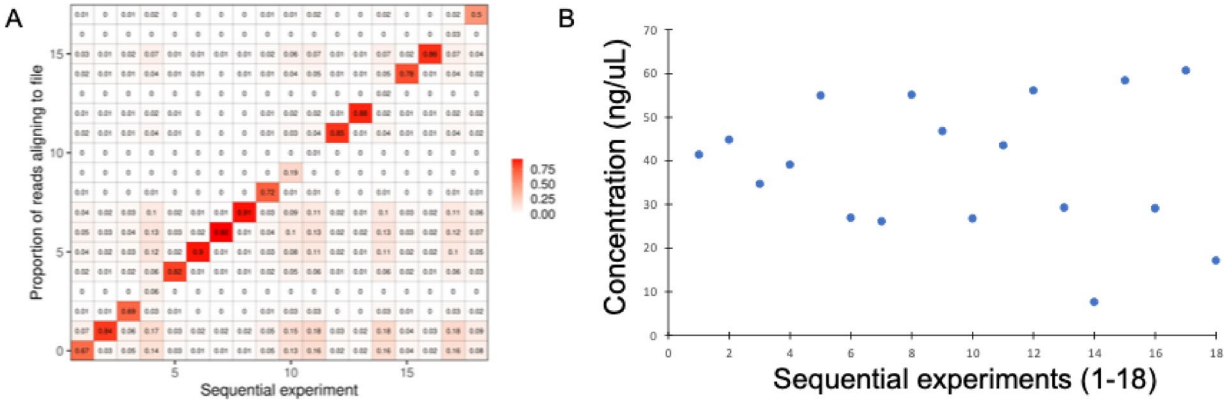

Fig. 4. Sequential amplification of DNA file groups from the same tube using RPA. File groups were retrieved using RPA instead of PCR. (**A**) On-target alignment rate is shown for each files experiment on the same tube. (**B**) Amplification yield for each experiment.

We performed a similar experiment using RPA in a new DNA-conjugated PCR tube. After the conjugation process, we conducted sequential RPA amplification of 18 files from a single tube and prepared library for nanopore sequencing. As with PCR, RPA amplification reaction products accumulate in solution. Therefore, we simply pipetted out the reaction products, purified them, and subjected them to typical nanopore sequencing protocols. In comparison to PCR, the thermal degradation observed in the RPA method was nearly negligible when compared to the PCR approach. The yield of RPA amplicons is significantly higher compared to PCR, reaching approximately 60 ng/μL (Fig. 4B). We also observed minimal experimental carryover across the 18 reactions (Fig. 4A, bottom right quadrant of heatmap). Some primer pair combinations did not amplify the target DNA file group (Fig. 4A). The file groups 4, 10, 11, 14, 17, and 18 did not amplify as efficiently as others. Future studies will focus on improving primer design and the specificity of DNA amplification. These results lead to the conclusion that RPA reactions may be a simpler method for DNA retrieval while avoiding issues with template degradation.

To address whether the surface-conjugated DNA is becoming damaged or is released from the PCR tube surface, we performed a complementary stability assay. We conjugated control lambda DNA to the tube and subjected it to sequential 95 C heating in 5-minute increments, measuring release into solution with a fluorimeter. No detectable DNA release was observed within the assay's sensitivity limits, suggesting that surface conjugates are not simply dissociating into solution. Taking into account that File Group amplification fails after repeated PCR cycles and that we do not observe any control DNA being released into solution, we infer that the covalently attached DNA is irreversibly damaged under thermal stress. Consistent with this interpretation, recombinase polymerase amplification (RPA) reactions at lower temperatures do not show such loss of function, supporting heat as the primary driver of conjugate degradation.

Comparing in solution PCR to RPA amplification on plastic conjugated DNA

We conducted experiments where we compared the differences between PCR and RPA amplification of the unconjugated DNA in solution (i.e., not attached to a solid phase support) versus the plastic-tube conjugated DNA. For these experiments, we performed a total of 36 amplification reactions (18 PCR reactions and 18 RPA reactions to target every file group), which substantially consumed the original DNA material – approximately 30% of the original synthesized oligonucleotide pool on the same tube per amplification reaction. The extent of DNA source material consumption thus motivates our work to enable repeated access.

We observed several key differences among our results. Overall, the PCR from the in-solution, unconjugated DNA had the most on-target amplification products compared to all other protocols (Supplementary Fig. 2A). We did not observe any discernible degradation trend in amplicon yield using PCR (Supplementary Fig. 3A, Supplementary Table 3, Supplementary Table 5). This was expected as each file was independently amplified as opposed to sequential retrieval from the same substrate. The result suggests that the decrease in yield was a result specific to performing many sequential PCR amplification reactions in the tube format.

We compared the RPA method between the unconjugated and conjugated DNA templates. As expected, we did not observe significant differences between in-solution RPA and that performed on conjugated DNA (Supplementary Figs. 2B, 3B, Supplementary Tables 4, 6). This result may be due to the RPA reaction rapidly achieving amplification saturation; therefore, the amplified yields would be expected to be similar.

Evaluation of sequenced oligonucleotide distributions across platforms

We measured the extent of amplification bias between DNA conjugated to the plastic tube versus in solution DNA. To do so, we analyzed the correlation coefficient of aligned read distributions between conjugated and unconjugated DNA contexts when performing file retrieval. We counted the number of sequences reads aligning to specific oligonucleotides with amplification reactions containing unconjugated DNA and measured whether this number was correlated with reactions containing conjugated DNA template. We determined the ratio of correlation coefficients between RPA and PCR for each file experiment. Specifically, we compared the Pearson

correlation coefficients for the read distribution across oligonucleotides in each group (approximately 800 unique oligonucleotide sequences per file group) between conjugated vs. in-solution amplification. Correlation coefficient ratios greater than 1 indicate that RPA from conjugated DNA was better at reproducing read-level distributions of in-solution RPA (Supplementary Fig. 4). Overall, we did not observe a substantial trend differentiating PCR and RPA in terms of sequencing read distributions. For 11 out of 18 file retrieval experiments we determined that RPA better matches the unconjugated in-solution control experiments, suggesting that RPA-based amplification from surface-conjugated DNA templates may be mildly better at mimicking in-solution amplification.

Discussion

We demonstrated a novel process by which DNA can be archived onto the plastic surface of a microcentrifuge tube via click chemistry and accessed by DNA amplification methods such as PCR and RPA. The results showed that the DNA conjugation process can be completed within hours, and that a wide range of input DNA amounts can be effectively used for conjugation. We estimated an upper bound for surface capacity by modeling geometric packing of a surface-tethered DNA bearing a TCO-dUTP tail (effective footprint ~ 50nm^2, yielding ~ 3 pmol per tube for the wetted area (50–100 µl). Accounting for incomplete site conversion, steric crowding from the TCO-dUTP tail, and electrostatic repulsion from the DNA backbone, the expected realized capacity is consistent with our measured 150 fmol that we experimentally loaded (Supplementary Note 1).

MTz-TCO click chemistry creates covalent bonds between DNA templates and the surface of the plastic PCR tube. Conventionally, non-covalent biotin-streptavidin interactions have been used in molecular biology applications. While the biotin-streptavidin interaction is among the strongest non-covalent interactions known, it remains inherently non-covalent which can allow partial dissociation or detachment under thermal stress[26]. Up to 10% of biotinylated molecules bound to streptavidin can dissociate under physiological conditions in less than 12 hours[26], and even more at longer incubation times or higher temperatures. In contrast, MTz-TCO immobilization forms a covalent bond, which is not subject to the equilibrium unbinding kinetics that may compromise biotin-streptavidin under cyclic thermal challenges. Thus, although streptavidin-biotin is highly stable under many conditions and is biocompatible with many assays, our results suggest that covalent immobilization via click chemistry may better preserve DNA functionality under repeated high-temperature cycling by eliminating the dissociation pathway inherent to non-covalent binding.

We also demonstrated scaling up the process for retrieving oligonucleotides by retrieving files – groups of oligonucleotides with common flanking sequences – as part of a synthesized DNA pool mixture. During the analysis of sequential reads from the PCR process, we observed substantial reductions in yield as experiments progressed on the same tube. We speculate that this reduction occurred due to the repeated application of heat from PCR cycling the same tube for 18 retrieval operations. While we could not distinguish between degradation of the functionalized plastic surface versus degradation of the DNA template itself, the extensive heat application reduces the applicability of PCR for DNA retrieval in our system.

In contrast, serial RPA assays on the same tube did not show any degradation patterns. The concentrations of each file retrieval did not follow a specific trend of decreasing yield as would be expected by degradation of the original template molecules. Furthermore, we observed a variability in amplicon yield among the files that is not correlated with sequential experiment number. This suggests that the isothermal reaction conditions of the RPA assays do not sufficiently damage DNA for multiple experiments as compared to PCR.

We demonstrated that RPA on DNA that is click-conjugated on plastic surfaces could be potentially used as a DNA storage platform in low-resource applications. RPA is enzymatically active between room temperature and 42 C, meaning that the reaction setup can be flexible. While the process demonstrates high specificity for many files of interest in our study, there were some off-target amplification issues which warrant future optimization. Lastly, an interesting aspect of this process is the high yields achieved after each RPA experiment, surpassed the yield obtained using traditional PCR methods. This result suggests that DNA files could be robustly targeted without extensive optimizations such as cycle number as in PCR. Overall, we demonstrated in this proof-of-concept study a new approach for repeated interrogation of DNA analytes as conjugated onto plastic surfaces.

Previously, we demonstrated robust attachment of DNA molecules onto crosslinked agarose beads where primer extension products and the original templates could be separated from each other through centrifugal or magnetic force[10,19]. Our current study advances on our previous work in several areas. First, any loss of beads previously led to loss of the stored library. MTz-TCO modifications led to enhanced hydrophobicity and bead clumping that could occur on pipet tips. This was previously addressed by careful pipetting by the user to ensure beads are not left on the pipet tip, but does not extend well to automation. This problem can be completely mitigated by switching to plastic surface conjugation as in this study. Secondly, the use of beads requires the physical transfer of a tube containing the beads to a centrifuge or magnet for manipulation. The use of plastic surfaces reduce the amount of necessary ancillary equipment and enables easy integration into liquid handling platforms.

Several prior studies indicate that immobilizing DNA templates on solid supports or surfaces can reduce PCR amplification efficiency and increase variability compared to using in-solution templates. For instance, membrane-bound templates yield lower and more variable product over repeated rounds, likely due to hindered access or diffusion constraints[27,28]. In solid-phase PCR and bridge amplification[29], high temperature denaturation causes substantial loss of immobilized DNA (~ 40–60%) over cycles, whereas milder denaturation methods (using formamide) preserve DNA better and produce denser clusters and higher retention. These reports align with our observations of decreased primer accuracy and decreased amplification yield. DNA immobilization likely imposes constraints on primer hybridization, extension, and thermal stability of the template, all of which reduce yield and can degrade the template in ways free solution templates do not experience. While primer

optimization to improve on-target accuracy is an ongoing area of study, we have shown that RPA robustly enables repeated interrogation of immobilized DNA elements from a plastic surface.

Experimental section

DNA data files and primers

We used a synthesized oligonucleotide pool consisting of approximately 15,000 oligonucleotides and 160 bp in length. These oligonucleotides were synthesized using a previously described encoding format that is specifically designed for nanopore sequencing[19]. It is used as template for this study to demonstrate file retrieval. The DNA pool was synthesized by Agilent. All sequences are also available at the URL https://github.com/shubhamchandak94/nanopore_dna_storage/tree/bonito/. The lambda DNA control amplicon is a 3.6 kb fragment available from Oxford Nanopore Technologies as the DCS control strand.

DNA sample preparation

DNA samples were prepared and tailed with TCO-functionalized dUTP as previously described[19]. We first denatured the DNA sample into single-stranded DNA by heating it at 95 C for 5 min then immediately cooling to 4 C in a thermocycler in 1X terminal transferase buffer (New England Biolabs, Ipswich, MA) containing 1X CoCl2 additive (5 µL) (New England Biolabs, Ipswich, MA) in a total of 45ul. Tailing reactions were then performed, comprising of 100 µM TCO-PEG4-dUTP (Jena Bioscience, Jena, Germany) and 20U of terminal transferase enzyme (New England Biolabs, Ipswich, MA) in a 50ul reaction. The reactions were incubated at 37 C for 1 h. Post-incubation, the reaction was purified using 1.8X Ampure XP beads (Beckman, Coulter, Brea, CA) and resuspended in elution buffer (10mM Tris-HCl pH 8.0, 0.05% Tween 20).

DNA conjugation process to plastic-based tubes

Methyltetrazine-functionalized surfaces on PCR strip tubes were custom-ordered from PolyAn GmbH (Berlin, Germany). Before use, the functionalized PCR tubes were cleaned with twice with 100ul wash buffer (10mM Tris-HCl pH 8.0, 0.05% Tween-20). Purified TCO-tailed DNA was added directly into the functionalized PCR tube and incubated overnight at room temperature in a rotating mixer. After incubation, the concentration of the supernatant was measured with a Qubit fluorescent assay.

Read access of targeted sequential experimental files with PCR

Individual file groups were retrieved with PCR amplification using independent primer pairs specific for each file group. Primer sequences are listed in Supplementary Table 2. Primer pairs were used at 1 µm concentration in 1X KAPA HiFi HotStart ReadyMix (Roche). We performed PCR reactions that targeted 18 different file groups and sequenced them on the Oxford Nanopore platform. Sequential experiments were directly performed by the PCR master mix solution with the DNA-conjugated onto the MTz-functionalized PCR tube. Amplification conditions were 45 s in 98 C, 18 cycles of 98 C for 15 s, 60 C for 60 s, 72 C for 60 s, and 72 C for 60 s before a final hold at 4 C. The amplification reaction containing the targeted product was removed from the DNA-conjugated PCR tube and was purified using 1.8X volume of Ampure XP beads. To remove residual amplification product, the DNA-conjugated tubes were treated with 1ul each of Exonuclease I and Exonuclease III (New England Biolabs, Ipswich, MA) in NEBuffer 2.1 for 30 min at room temperature. To inactivate the exonuclease reaction, the microcentrifuge tubes were kept at 65 C for 20 min. Then tubes were cleaned with twice with wash buffer (10mM Tris-HCl pH 8.0, 0.05% Tween 20), and then stored at -20 C in storage buffer (10mM Tris-HCl pH 8.0, 0.05% Tween 20, 50% glycerol) for further re-use. The same PCR tube was used for 18 times to sequentially retrieve targeted files.

Read access of targeted sequential 18 files with recombinase polymerase amplification

Individual file groups were retrieved with this amplification enzymes using separate barcodes used as targeted primers at 1 µm concentration in 1X primer rehydration buffer, magnesium acetate and lyophilized TwistAmp Basic reaction mix (TwistDX, United Kingdom) and sequential experiment for each individual files were directly performed with DNA-conjugated plastic tube at 37 C for 20 min followed by hold at 4 C. After the incubation, the reaction mix was pipetted out of the tube and cleaned with 1.8X Ampure XP beads. As with PCR reactions, the DNA-conjugated PCR tube was cleaned with Exonuclease I and Exonuclease III. To prevent spurious amplification that could occur at 4 C, reaction products were pipetted out of the DNA-conjugated PCR tube less than 10 min after reaction completion.

The amplicons were diluted to prepare the sequencing libraries for the ONT platform using LSK 110 kit. The libraries were then sequenced using R9 chemistry on a PromethION flowcell for 72 h while loading 600fmoles of library.

Nanopore sequencing library preparation

From all targeted DNA products, we generated sequencing libraries compatible with the Oxford Nanopore sequencing platform. We first performed end-repair and a-tailing on the DNA amplicons using the Kapa HyperPrep kit (Roche) using the manufacturer's standard protocol. For the ligation reaction, we incubated a barcoded adapter from the Native Barcoding Expansion Kit (Oxford Nanopore Technologies) in the standard Kapa HyperPrep ligation reaction configuration. After ligation of the barcoded adapter, we performed a cleanup step using 0.8X Ampure XP beads. A second ligation step using 1X Quick Ligation Buffer (New England Biolabs, Ipswich, MA), 10ul of ligase from the Kapa HyperPrep kit (Roche), and 5ul of Adapter Mix II (AMII, Oxford Nanopore Technologies, UK) was then performed to incorporate the nanopore-specific motor protein for sequencing. The ligation reaction was purified with 1.5X Ampure XP beads but the beads were washed with SFB buffer (Oxford Nanopore Technologies, UK) instead of 80% ethanol and eluted in EB buffer (Oxford Nanopore

Technologies, UK). The library was then immediately sequenced with a R9 flow cell on the Oxford Nanopore PromethION sequencer for 72 h.

Sequence data processing and alignment

Raw nanopore signal data underwent basecalling to generate sequence reads in FASTQ format, and were demultiplexed into sample-specific files. This was performed using the native basecaller Guppy (Oxford Nanopore Technologies, UK) in high accuracy mode. Sequence alignment was performed using minimap2 v2.17, utilizing the designed oligonucleotide sequences as a reference. To evaluate targeting efficiency, read counts aligning to each oligonucleotide were tabulated using samtools v1.10.

DNA targeting analysis

As above, reads were aligned to a reference genome consisting of the designed pool of oligonucleotides ($N = 14{,}898$ unique oligonucleotide sequences across 18 file groups). The on-target rate was estimated by the number of reads aligning to the intended file group that was amplified versus the total number of reads sequenced for the experiment. Reads per experiment were up to several hundred thousand to over a million reads per experiment.

To estimate the carry-over between iterative experiments, we added the cumulative number of reads belonging to file groups that were targeted in previous experiments to the current targeted file group. For example, in experiments targeting file group 2, the number of reads that aligned to read group 1 would be considered to be carried over from the previous experiment. The cumulative number of reads in this category was then divided by the total number of reads sequenced to derive the carry-over rate. Only primary reads that were mapping to the reference was used. Only reads with a passing base quality of above 7 were used, which is a typical filter for nanopore sequencing. Because the carry-over rate is specific to a single experiment compared the previous one, a single replicate was used. However, an overall carry-over rate can be determined by averaging the per-sample carry-over across the performed experiments ($N = 18$ for both PCR and RPA).

Data availability

Sequence-aligned bam files are available at Zenodo under the DOI 10.5281/zenodo.13179444. It can be accessed via the URL https://doi.org/10.5281/zenodo.13179444.

Received: 22 November 2024; Accepted: 22 January 2026
Published online: 28 January 2026

References

1. Koch, J. et al. A DNA-of-things storage architecture to create materials with embedded memory. *Nat. Biotechnol.* **38**, 39–43. https://doi.org/10.1038/s41587-019-0356-z (2020).
2. Bancroft, C., Bowler, T., Bloom, B. & Clelland, C. T. Long-Term storage of information in DNA. *Science* **293**, 1763–1765. https://doi.org/10.1126/science.293.5536.1763c (2001).
3. Park, S. J. et al. Reducing cost in DNA-based data storage by sequence analysis-aided soft information decoding of variable-length reads. *Bioinformatics* **39**, btad548. https://doi.org/10.1093/bioinformatics/btad548 (2023).
4. Yazdi, S. M. H. T. et al. DNA-Based storage: trends and methods. *IEEE Trans. Mol. Biol. Multi-Scale Commun.* **1**, 230–248. https://doi.org/10.1109/TMBMC.2016.2537305 (2015).
5. Ceze, L., Nivala, J. & Strauss, K. Molecular digital data storage using DNA. *Nat. Rev. Genet.* **20**, 456–466. https://doi.org/10.1038/s41576-019-0125-3 (2019).
6. Church, G. M., Gao, Y. & Kosuri, S. Next-Generation digital information storage in DNA. *Science* **337**, 1628–1628. https://doi.org/10.1126/science.1226355 (2012).
7. Bornholt, J. et al. *Proceedings of the Twenty-First International Conference on Architectural Support for Programming Languages and Operating Systems* 637–649 (Association for Computing Machinery, 2016).
8. Erlich, Y. & Zielinski, D. DNA fountain enables a robust and efficient storage architecture. *Science* **355**, 950–954. https://doi.org/10.1126/science.aaj2038 (2017).
9. Organick, L. et al. Random access in large-scale DNA data storage. *Nat. Biotechnol.* **36**, 242–248. https://doi.org/10.1038/nbt.4079 (2018).
10. Lau, B. T. & Ji, H. P. Covalent click Chemistry-Based attachment of DNA onto solid phase enables iterative molecular analysis. *Anal. Chem.* **91**, 1706–1710. https://doi.org/10.1021/acs.analchem.8b05139 (2019).
11. Chandak, S. et al. *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 8822–8826.
12. Zan, X., Xie, R., Yao, X., Xu, P. & Liu, W. A. Robust and efficient DNA storage architecture based on modulation encoding and decoding. *J. Chem. Inf. Model.* **63**, 3967–3976. https://doi.org/10.1021/acs.jcim.3c00629 (2023).
13. Welzel, M. et al. DNA-Aeon provides flexible arithmetic coding for constraint adherence and error correction in DNA storage. *Nat. Commun.* **14**, 628. https://doi.org/10.1038/s41467-023-36297-3 (2023).
14. Kubo, S., Niimi, H. & Kitajima, I. Improved reverse transcription-recombinase polymerase amplification assay for blood mRNA screening: comparison with one-step RT-qPCR assay. *Forensic Sci. International: Genet.* **63**, 102808. https://doi.org/10.1016/j.fsigen.2022.102808 (2023).
15. Ma, B. et al. A simple and efficient method for potential point-of-care diagnosis of human papillomavirus genotypes: combination of isothermal recombinase polymerase amplification with lateral flow dipstick and reverse Dot blot. *Anal. Bioanal. Chem.* **411**, 7451–7460. https://doi.org/10.1007/s00216-019-02113-5 (2019).
16. Daher, R. K., Stewart, G., Boissinot, M. & Bergeron, M. G. Recombinase polymerase amplification for diagnostic applications. *Clin. Chem.* **62**, 947–958. https://doi.org/10.1373/clinchem.2015.245829 (2016).
17. Algethami, M. et al. Unravelling the clinicopathological and functional significance of replication protein A (RPA) heterotrimeric complex in breast cancers. *Npj Breast Cancer.* **9**, 18. https://doi.org/10.1038/s41523-023-00524-3 (2023).
18. Warmt, C., Broweleit, L. M., Fenzel, C. K. & Henkel, J. An experimental comparison between primer and nucleotide labelling to produce RPA-amplicons used for multiplex detection of antibiotic resistance genes. *Sci. Rep.* **13**, 15734. https://doi.org/10.1038/s41598-023-42830-7 (2023).
19. Lau, B. et al. Magnetic DNA random access memory with nanopore readouts and exponentially-scaled combinatorial addressing. *Sci. Rep.* **13**, 8514. https://doi.org/10.1038/s41598-023-29575-z (2023).

20. Liu, J. et al. A recombinase polymerase amplification (RPA) combined with strip visualization method for RNA-based presumptive tests of saliva and vaginal secretion. *Forensic Sci. International: Genet.* **62**, 102788. https://doi.org/10.1016/j.fsigen.2022.102788 (2023).
21. Srivastava, P. & Prasad, D. Isothermal nucleic acid amplification and its uses in modern diagnostic technologies. *3 Biotech.* **13**, 200. https://doi.org/10.1007/s13205-023-03628-6 (2023).
22. Shang, Y. et al. Development of nucleic acid extraction and real-time recombinase polymerase amplification (RPA) assay integrated microfluidic biosensor for multiplex detection of foodborne bacteria. *Food Control*. **155**, 110047. https://doi.org/10.1016/j.foodcont.2023.110047 (2024).
23. Thoeny, V. et al. Recombinase polymerase amplification in combination with electrochemical readout for sensitive and specific detection of PIK3CA point mutations. *Anal. Chim. Acta*. **1281**, 341922. https://doi.org/10.1016/j.aca.2023.341922 (2023).
24. Ying, J. et al. Development and validation of real-time recombinase polymerase amplification-based assays for detecting HPV16 and HPV18 DNA. *Microbiol. Spectr.* **11**, e01207–01223. https://doi.org/10.1128/spectrum.01207-23 (2023).
25. Kaur, R. et al. Rapid detection of Porcine sapelovirus by reverse transcription recombinase polymerase amplification assay. *Mol. Biol. Rep.* **51**, 178. https://doi.org/10.1007/s11033-023-09123-8 (2024).
26. Chivers, C. E. et al. A Streptavidin variant with slower biotin dissociation and increased mechanostability. *Nat. Methods*. **7**, 391–393. https://doi.org/10.1038/nmeth.1450 (2010).
27. Sheikh, S. N. & Lazarus, P. Re-usable DNA template for the polymerase chain reaction. *Nucleic Acids Res.* **25**, 3537–3542. https://doi.org/10.1093/nar/25.17.3537 (1997).
28. Kadokami, Y. & Lewis, R. V. Membrane bound PCR. *Nucleic Acids Res.* **18**, 3082. https://doi.org/10.1093/nar/18.10.3082 (1990).
29. Huang, J. et al. Low-temperature and high-efficiency solid-phase amplification based on formamide. *Micromachines (Basel)* **15** https://doi.org/10.3390/mi15050565 (2024).

Acknowledgements

The work was supported by the National Science Foundation (NSF) and the Semiconductor Research Council (Award 1807371). Additional support to B.T.L came from NIH grant R35HG011292. Additional support to H.P.J came from NIH grant R33CA247700. Graphical workflows in Figs. 1 and 2 were constructed with BioRender.

Author contributions

B.T.L, S.R. and H.P.J. designed the experiments. S.R. and B.T.L conducted the experiments. B.T.L developed the analysis algorithms. B.T.L and S.R, analyzed the data. B.T.L, S.R. and H.P.J. wrote the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-37575-y.

Correspondence and requests for materials should be addressed to H.P.J. or B.T.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.