Write and read: Harnessing Synthetic DNA Modifications for Nanopore Sequencing.

Uri Bertocchi [a,b], Assaf Grunwald [a], Gal Goldner [a], Eliran Eitan [a], Sigal Avraham [a], Jasline Deek [a], Yael Michaeli [a], Brian Yao [c], Jennifer Listgarten [c], Jared T. Simpson [d,e], Winston Timp [f], Yuval Ebenstein [a,b*]

[a] School of Chemistry, School of Biomedical Engineering, Tel Aviv University, Tel Aviv-Yafo, 6997801, Israel

[b] Sagol School of Neuroscience, Tel Aviv University, Tel Aviv-Yafo, 6997801, Israel

[c] Department of Electrical Engineering & Computer Sciences, University of California, Berkeley, CA 94720, USA

[d] Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada

[e] Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

[f] Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

* E-mail: uv@tauex.tau.ac.il

KEYWORDS Nanopore sequencing, Epigenetics, DNA methylation, 5-hydroxymethylcytosine (5hmC), 5-methylcytosine (5mC), DNA tagging, β-glucosyltransferase (BGT), Methyltransferase (Mtase), Glucosylation, Alkylation,

ABSTRACT: Nanopore sequencing is enabled by detecting nucleotide-specific molecular signatures through changes in ionic current as DNA strands translocate through the pore. Well-trained models allow the detection of naturally occurring DNA modifications such as DNA methylation. Here, we present a versatile "write-and-read" framework, where chemo-enzymatic DNA labeling with unnatural synthetic tags results in predictable electrical fingerprints in nanopore sequencing. We first explore a DNA glucosylation approach that selectively modifies 5-hydroxymethylcytosine (5hmC) with glucose or glucose-azide adducts via T4 β-glucosyltransferase (BGT) and examine the impact on nanopore current signals. We demonstrate that these modifications generate distinct and reproducible electrical shifts, allowing for direct detection of chemically altered nucleotides. We further show that enzymatic alkylation may write and read additional species. Transfer of azide residues to the N6 position of adenines also produces characteristic nanopore signal shifts relative to the native adenine and 6-methyladenine. Beyond direct nucleotide detection, this approach introduces new possibilities for bio-orthogonal DNA labeling, enabling an extended alphabet of sequence-specific detectable moieties. This method expands the capabilities of nanopore-based single-molecule sensing by integrating real-time electrical detection with programmable chemical modifications.

INTRODUCTION

Nanopore sequencing has emerged as a powerful platform in genomics, offering the unique capability to read long nucleic acid molecules at the single-molecule level[1–6]. Protein nanopores function as ultrasensitive molecular sensors, detecting changes in ionic current as individual molecules traverse the pore[6,7]. When an applied voltage drives ionic current through the nanopore, the passage of a molecular species attenuates the flow, producing a measurable change in current[6,7]. This modulation is influenced by the molecule's physical dimensions and its electrostatic and hydrophobic interactions with the pore interior[6,7]enabling signal variations that can be computationally decoded to distinguish nucleotide sequences. These fundamental principles underlie nanopore-based sequencing, where the continuous monitoring of ionic current enables real-time profiling of single-stranded DNA[6], RNA[8], or proteins[9], as they are translocated through the nanopore.

One of the advantages of protein nanopores, such as those offered by Oxford Nanopore Technologies (ONT) for sequencing, is that the electrical signals detected during the process are attenuated by chemical modifications of nucleobases[2]. This sensitivity makes nanopore sequencing especially valuable for epigenetic studies[1,2,10,11]. Current nanopore models focus on calling natural chemical DNA modifications, such as cytosine and adenine methylation[12], providing a limited selection of detectable moieties. Conversely, chemistry may provide a broader spectrum of engineered modifications to read additional genomic information.

Various in-vitro DNA repair, glucosylation, and alkylation reactions have been developed to tag epigenetic and genetic marks with fluorescent tags[13–19]. These reactions were used extensively for optical genome mapping of long-range genetic/epigenetic profiles[20–30]. We hypothesized that the

same approach may be applied to electrical detection by introducing chemical moieties that create electrical contrast and are distinctly detectable via nanopore sequencing. This proof-of-concept study demonstrates selective chemoenzymatic labeling with different synthetic modifications before nanopore sequencing. These modifications influence the measured current in the nanopore, allowing for their specific detection. This approach enables precise control over which DNA targets are chemically marked, integrating three key components: (1) DNA-modifying enzymes that recognize specific sequence motifs or pre-existing modifications[31,32], (2) engineered cofactors carrying detectable chemical groups[33–39], and (3) high-resolution electrical detection of labeled DNA strands via nanopore sequencing. Together, these elements establish a versatile "write-and-read" paradigm, wherein a chosen enzyme installs a tailored chemical marker at specific genomic sites, generating distinct current signatures in the nanopore readout.

We first set out to explore DNA glucosylation as a contrast mechanism. Our study harnesses the natural glucosylation mechanism developed by bacteriophages[40]. We implemented an enzymatic glucosylation strategy using T4 β-glucosyltransferase (BGT) with either glucose or glucose-azide donors to selectively convert 5-hydroxymethylcytosine (5hmC) into either 5-glucosyl hydroxymethyl cytosine (5gmC) or 5-azide-glucosyl hydroxymethyl cytosine ($5N_3$-gmC); see Figure 1A. A similar method was used for 5hmC sequencing through selective pulldown of glucosylated DNA [41,42]. In our case, we tested whether these modifications induce distinct ionic current signatures during nanopore sequencing (see Figure 1B for illustration).

In addition, we demonstrate that other enzyme-substrate systems could be adapted to introduce a broad spectrum of chemical modifications at specific genomic sites. For example, methyltransferases, which catalyze methyl group transfer to a nucleobase within their recognition sequences, could be repurposed for site-specific labeling[43]. By replacing their natural methyl donor, S-adenosylmethionine (AdoMet), with a synthetic cofactor such as 6-azido-S-adenosyl-L-methionine (6-N_3-AdoMet), an azide moiety was directly incorporated, facilitating downstream nanopore detection.

This work establishes a generalizable framework for incorporating non-natural moieties into DNA for labeling and detection using nanopore sequencing. By substituting native cofactors with suitably engineered analogs, our approach retains DNA sequence and natural modification specificity while introducing a new distinct chemical signature. This will enable users to label specific genomic information that is then readable by nanopore sequencing.

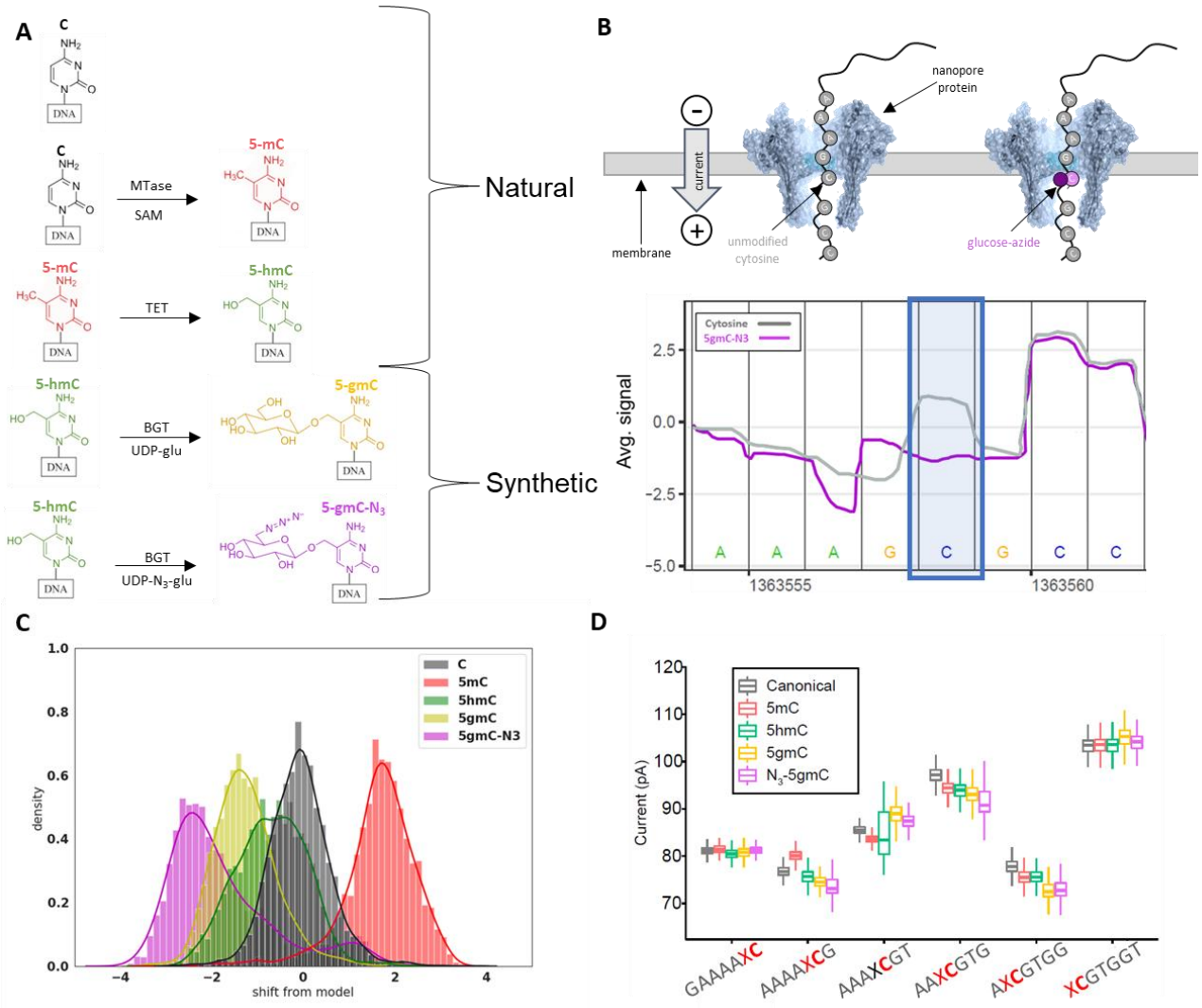


Figure 1. (A) The molecular conversion of cytosine into various natural and synthetic derivatives: 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-glucosyl hydroxymethyl cytosine (5gmC), and 5-glucosyl azide-hydroxymethylcytosine (5gmC-N_3). (B) Illustration of a single-strand DNA with modified cytosine passing through a nanopore. (C) A density plot showing the distribution of signal shifts of a single *k*-mer (GATGCG) demonstrating clear differentiation between modifications. (D) Ionic current distributions produced using nanopolish's eventalign function[2,61] - for canonical cytosine, 5mC, 5hmC, 5gmC, and 5gmC-N_3, demonstrating modification-specific signal variations in nanopore sequencing; C or modified-C position in the 6-mer is colored in red.

RESULTS

DNA glucosylation detection

We selected ~100 sequence compositions (11-mers) for analysis. For each *k*-mer, we prepared four PCR amplicons, each containing a cytosine variant at a predetermined position: C, 5hmC, 5gmC, and 5gmC-N_3. The 100 DNA samples were sequenced using the ONT MinION platform with R9.4.1 flow cells.

Analysis of the ionic current signals showed significant differences for both 5gmC and 5gmC-N_3 when compared to the unmodified 5hmC, specifically at the modification sites and one to two bases upstream of these sites (see Figure 2). For simplicity, we visualize only the 8 positions surrounding signal variation between k-mers. These differences were consistent across all pre-selected *k*-mers, as illustrated in Figure S1. Signal plots and statistics are presented in the supplementary information (Figures S2 and S3). As shown in Figure 1C-D, our chemoenzymatic tagging strategy generates large current shifts for 5gmC and 5gmC-N_3 relative to natural cytosine variants (C, 5mC, 5hmC). As the modified base is at different positions within the pore, the current shift changes; the first position shows no shift, while the 3rd and 4th positions show the strongest shift overall. This is similar to previous findings with natural modifications[2].

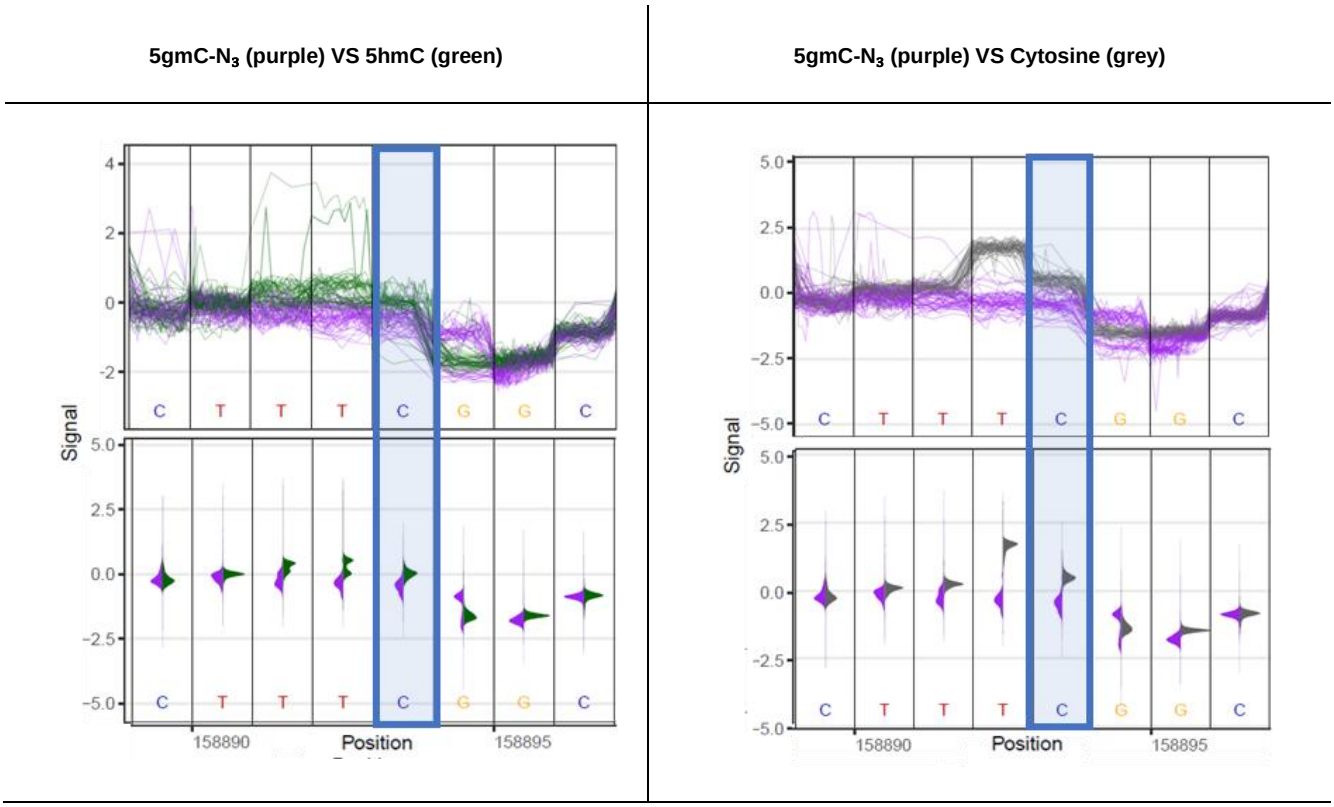


Figure 2. Plots comparing ionic current signals for 5-azide-glucosylhydroxymethylcytosine (5gmC-N_3, purple) versus 5-hydroxymethylcytosine (5hmC, green) (left) and unmodified cytosine (grey) (right). The y-axis represents normalized ionic current (pA), and the x-axis denotes genomic position. Upper panels show raw signal traces, while lower violin plots display signal distributions. The modified cytosine position is highlighted in blue.

Analysis of Non-Assigned Events

To further investigate the impact of our chemoenzymatic labeling on nanopore sequencing, we examined events in the raw signal data that the signal realignment models could not confidently align to the reference genome. These unassigned events can reflect instances where the observed ionic current deviated significantly from the expected patterns defined by the standard model. Since the model is not optimized for glucosylated variants, these deviations further suggest that the synthetic modifications are introducing measurable signal distortions.

Figure S4 illustrates that the frequency of these unassigned events varies with the modification state: both 5gmC and 5gmC-N_3 samples exhibit a higher incidence of such events, particularly at positions 5 and 2 bases upstream of the modified site; 17.15% and 7.4% for 5gmC, 15.39% and 6.15% for 5gmC-N3, respectively). In contrast, native 5hmC shows unassigned events primarily at the 5th base upstream of the modified site (15.4%), while unmodified cytosine is generally recognized across all positions (~2%). These findings correspond to those discussed in Figure 1D, demonstrating that glucosylation increases basecalling errors. This indicates that the synthetic modifications yield signal disruptions distinct from those produced by naturally occurring nucleotides.

DNA alkylation detection

We extended our experiments to DNA alkylation reactions to explore the general utility of our write/read approach. Specifically, we utilized methyltransferases, which naturally alkylate specific sites on DNA by transferring methyl groups from the universal methyl donor S-adenosylmethionine (AdoMet) to particular nucleobases[30,44]. We used native and mutant methyltransferases in combination with synthetic cofactors that enable the transfer of various chemical tags directly to DNA[30,39,43,45]; for an extended list, see Table S4. Such tags may create an electrical contrast for sequencing. We decided to evaluate adenine alkylation, which may serve as a valuable tag for mapping protein-DNA interactions[46,47].

To evaluate the impact of adenine modifications on nanopore sequencing, we generated PCR-amplified DNA products containing the TCGA motif, which serves as the recognition sequence for the M.TaqI methyltransferase. We used M.TaqI along with AdoMet or an azide-functionalized analog, 6-N_3-AdoMet to produce three variants of adenine in the TCGA-motif-containing DNA: unmodified, methylated (m6A), and azide-modified (N3-A) (see Figure 3).

Analysis of current distributions revealed distinct effects for m6A and N3-A. While A and m6A overlap highly in signal in most positions, N3-A exhibited more substantial shifts across multiple pore positions. This suggests that N3-A is more easily detectable, likely due to its more considerable

steric bulk and altered molecular interactions within the pore.

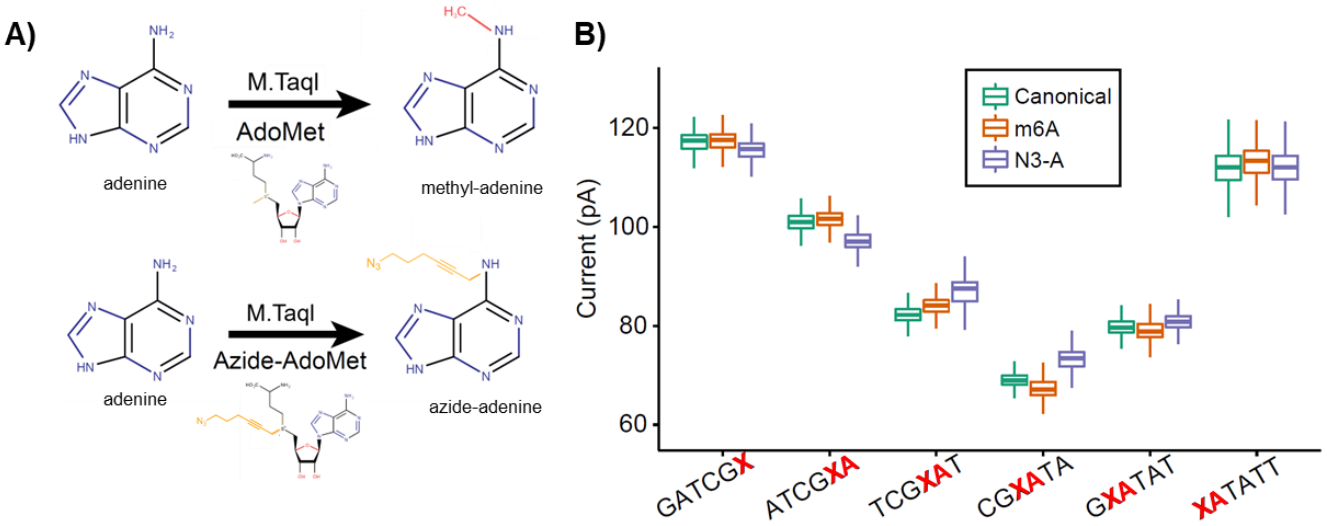


Figure 3. (A) Enzymatic transfer of the azide-methyl group to adenine by M.TaqI methyltransferase using S-adenosylmethionine (AdoMet) co-factor and its azide-modified analog (Azide-AdoMet). (B) Box plots of ionic current (pA) for canonical adenine (N = 6569), N6-methyladenine (m6A, N = 7224), and azide-modified adenine (N_3-A, N = 5191) across selected k-mers, showing distinct current shifts; A or modified-A position in the 5-mer is colored in red

DISCUSSION

This study introduces an exogenous labeling method for nanopore sequencing based on selective enzymatic DNA manipulation with unnatural modifications. We demonstrate that these modified bases produce distinct electrical signatures as they translocate through the pore, broadening the palette of detectable signals. This can enable multiple different pieces of information to be encoded on the same DNA molecule, e.g., chromatin accessibilty[1], protein-DNA interaction[46] and native DNA methylation all at once. By extending the labeling toolbox we previously developed for fluorescent tagging of DNA[15,22,48], we demonstrated the write-and-read concept with chemoenzymatic glucosylation or alkylation with an unnatural moiety. In the first case, we used β-glucosyltransferase (BGT) to selectively modify 5-hydroxymethylcytosine (5hmC) with glucose or engineered glucose-azide cofactors across diverse sequence contexts. Secondly, we showed selective alkylation by incorporating an azide moiety at N6-adenine via a methyltransferase-mediated reaction. These targeted modifications generated distinct ionic current profiles during nanopore sequencing, validating our strategy for encoding chemical information directly into DNA and decoding it at the single-molecule level.

Similarly, our preliminary investigation into adenine modifications revealed that while both natural N6-methyladenine (m6A) and its azide-modified counter-part (N_3-A) produce measurable changes in ionic current, with the azide modification consistently eliciting a more pronounced effect in multiple sequence contexts. Although additional validation in diverse k-mer environments is warranted, these results suggest that nanopore sequencing can accurately distinguish chemically modified adenine residues.

Despite the promising results, this study has several limitations. First, our experiments were primarily conducted on synthetic DNA substrates. While these provide controlled conditions for proof-of-concept validation, they may not fully recapitulate the complexity of native genomic DNA. Additionally, the increased occurrence of unassigned *k*-mer events suggests that the chemical modifications, notably bulky adducts, can interfere with the basecalling accuracy of currently available models, complicating downstream analyses. A dedicated model for these unnatural modifications requires further optimization[49] to ensure robust performance across various sequence contexts and modification densities. Furthermore, this project used the R9 MinION flow cells, which were discontinued in July 2024[50]. While these flow cells provided reliable data for the proof-of-concept, future work will need to re-establish the compatibility and performance of this chemoenzymatic labeling strategy with the newer-generation pore characteristics that may influence signal profiles[51,52].

Protein pore engineering, electronics, and computational algorithms will keep improving our ability to identify native modifications such as DNA methylation. However, the presented approach offers the opportunity for signal engineering beyond the natural repertoire of nucleobases via selective chemical manipulation of DNA. By expanding the molecular alphabet accessible to nanopore sequencing and integrating synthetic modifications into recently reported methods for mapping chromatin structure[1,53] and protein-DNA interactions[46,47,54], our approach may offer a high-resolution view of DNA epigenetics in combination with footprints of the chromatin machinery and transcription factor landscapes.

Our strategy lays the groundwork for integrated multi-omic analyses that capture genetic variation, dynamic epigenetic states, and chromatin organization in a single-pass readout. The addition of new chemistries for selective writing onto DNA may open new avenues for studying the interplay between diverse epigenetic marks—such as methylation, hydroxymethylation, and histone modifications, and their collective impact on gene regulation and cellular identity[55].

METHODS

Glucosylation of Cytosine

In this study, three types of DNA samples containing different cytosine modifications were prepared: (1) 5-hmC, (2) 5-gmC, and (3) 5-gmC-N3. The Escherichia coli (E. coli K-12 substr. MG1655) genome was used as the template for amplification[56]. 25 distinct 11-mers were selected as target sites for modification (Table S1). The forward primer was synthesized with a single 5-hmC residue was utilized to introduce a specific 5-hmC modification into the PCR product. PCR reactions were performed in a final volume of 50 μL, containing 100 ng of E. coli genomic DNA, (extracted with Promega Wizard DNA Extraction kit according to protocol), 0.4 μM primers, 25 μL of the ready mix (Bioline), and ultrapure water. The amplification conditions followed the program detailed in Table S2. After PCR, the DNA was purified using the QIAquick PCR Purification Kit (Qiagen), and fragment sizes were verified via gel electrophoresis.

To selectively label 5-hmC residues, glucosylation reactions were performed using two distinct cofactors: uridine diphosphoglucose (UDP-Glu) for standard glucosylation and UDP-6-N3-Glu for azide-functionalized glucosylation. T4-BGT catalyzed the transfer of glucose moiety from UDP-Glu to 5-hmC, producing 5-gmC. Similarly, using UDP-6-N3-Glu resulted in the transfer of an azide-functionalized glucose moiety to 5-hmC, producing 5-gmC-N3. Glucosylation reactions were carried out in 1× NEBuffer 4 (New England Biolabs). The 5-gmC modification reaction mixture contained 1 μg of synthetic DNA, 20 units of BGT, and 45 μM UDP-Glu in a total volume of 30 μL, incubated overnight at 37 °C. The 5-gmC-N3 modification followed the same procedure but used UDP-6-N3-Glu as the cofactor to introduce the azide-functionalized glucose moiety; RP-HPLC measurements indicate 100% yield for the transfer of the UDP-6-N3-Glu (see Figures S5-S6, and Table S5).

Following glucosylation, the labeled DNA samples were purified using Oligo Clean & Concentrator columns (Zymo Research), according to the manufacturer's recommendations, and stored at −20 °C for further analysis.

Alkylation of Adenine

To chemically modify specific N6-adenine sites by incorporating azide groups, We used the methyltransferases M.TaqI, which catalyze methyl group transfer to adenine within its recognition sequence, TCGA. We performed reactions with both the natural methyl donor S-adenosylmethionine (AdoMet), and the synthetic cofactor 6-azido-S-adenosyl-L-methionine (6-N3-AdoMet), enabling the direct incorporation of an azide moiety at the N6 position of adenine.

DNA labeling reactions were carried out in 1× rCutSmart™ Buffer (New England Biolabs). The methylation and azide modification reaction mixture contained 2 μg of DNA, 30 units of M.TaqI, and 80 μM of either azide cofactor or S-adenosylmethionine (SAM) cofactor in a total volume of 100 μL, incubated at 60°C for 1 hour. The enzyme was inactivated by incubation with 20 μg/mL proteinase K at 45 °C for 2 hours. Following incubation, samples were purified using the Zymo Oligo Clean & Concentrator Kit according to the manufacturer's instructions and stored at −20 °C for further analysis.

Nanopore Sequencing and Data Analysis

We then conducted nanopore sequencing for our samples using the Oxford Nanopore Technologies (ONT) ligation sequencing gDNA SQK-LSK109 kit, equipped with MinION platform R9.4.1 flow cells. After sequencing, the raw data were initially generated as multi-read fast5 files containing multiple reads per file. These files were converted into single-read fast5 format to facilitate downstream analysis using the "multi_to_single_fast5" utility from the "ont-fast5-API" library[57]. This conversion ensured that each fast5 file contained data from a single DNA molecule, as required for the subsequent signal analysis.

Basecalling was performed using ONT's Guppy software (version 6.3.2) in high-accuracy mode (model dna_r9.4.1_450bps_hac), optimized for the R9.4.1 flow cell chemistry. Guppy translated the raw electrical signals into nucleotide sequences, generating FASTQ files (see Table S3 for yield summary).

To integrate the basecalled sequences with the raw ionic current data, we used Tombo's annotate_raw_with_fastqs function[58]. This annotation step linked the FASTQ reads to the corresponding fast5 files, enabling a direct association between the nucleotide sequence and its underlying current profile, as Tombo's 1.5.0 version required.

We then applied Tombo's resquiggle function for signal-level alignment, which involved aligning the raw squiggle data (ionic current traces) to the reference genome. For synthetic samples, we aligned the data to the E.coli reference genome[59]. The resquiggle process corrected baseline drifts and adjusted the observed signals to match the expected current values based on the reference sequence. This allowed us to accurately map the signal shifts associated with 5gmC and 5gmC-N_3 modifications.

The resquiggled data was then analyzed to quantify the distinct current shifts induced by the chemical tagging. To visualize and compare the signal shifts across different modifications, we utilized Tombo's plot command to generate plots showing the ionic current profiles for each modification type—5hmC, 5gmC, and 5gmC-N_3—aligned along genomic positions.

To quantify the observed signal shifts and provide statistical validation of the chemical modifications, we utilized Nanopolish's eventalign function[2,60]. Nanopolish eventalign realigns the reads to the reference genome at the signal level, providing detailed information about each base's timing and magnitude of ionic current events. The output from eventalign included precise event-level data for each read, such as the observed current mean, standard deviation, and dwell time, allowing for a comprehensive analysis of the modifications.

To further assess the impact of synthetic modifications on signal realignment, we examined the occurrence of unassigned k-mers, denoted as "NNNNNN," in the output of Nanopolish's eventalign function. This function provides a position-by-position mapping of ionic current shifts, linking each detected signal to a specific nucleotide sequence. However, when the observed current deviates significantly from the expected signal profile—typically due to unmodeled DNA modifications—Nanopolish masks the k-mer as "NNNNNN", indicating that the raw signal could not be confidently aligned to the reference sequence. Importantly, this does not imply a failure in basecalling, as the ONT basecaller (e.g., Guppy) still assigns standard ACGT sequences, but rather that the signal perturbations introduced by the modifications led to realignment uncertainty in downstream analysis.

AUTHOR INFORMATION

Corresponding Author

* Yuval Ebenstein, School of Chemistry, The Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 6997801, Israel; Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel; Department of Biomedical Engineering, Fleischman Faculty of Engineering, Tel Aviv University, Tel Aviv 6997801, Israel; and Center for the Physics and Chemistry of Living Systems, Tel Aviv University, Tel Aviv 6997801, Israel, E-mail: uv@tauex.tau.ac.il

Funding Sources

WT, JTS and YE are supported by the National Human Genome Research Institute (5R01HG009190). YE acknowledges the European Research Council consolidator grant [817811], the Israel Science Foundation [771/21], and the Koret-UC Berkeley-Tel Aviv University Initiative in Computational Biology and Bioinformatics. The Ontario Institute for Cancer Research supports JTS through funds provided by the Government of Ontario, the Government of Canada through Genome Canada and Ontario Genomics (OGI-136 and OGI-201).

Conflicts

WT has two patents (8,748,091 and 8,394,584) licensed to Oxford Nanopore Technologies. WT has received travel funds to speak at symposia organized by Oxford Nanopore.
YE has a pending patent application concerning synthetic modifications for nanopore sequencing.

REFERENCES

(1) Lee, I.; Razaghi, R.; Gilpatrick, T.; Molnar, M.; Gershman, A.; Sadowski, N.; Sedlazeck, F. J.; Hansen, K. D.; Simpson, J. T.; Timp, W. Simultaneous Profiling of Chromatin Accessibility and Methylation on Human Cell Lines with Nanopore Sequencing. *Nat. Methods* **2020**, *17* (12), 1191–1199. https://doi.org/10.1038/s41592-020-01000-7.

(2) Simpson, J. T.; Workman, R. E.; Zuzarte, P. C.; David, M.; Dursi, L. J.; Timp, W. Detecting DNA Cytosine Methylation Using Nanopore Sequencing. *Nat. Methods* **2017**, *14* (4), 407–410. https://doi.org/10.1038/nmeth.4184.

(3) Kovaka, S.; Fan, Y.; Ni, B.; Timp, W.; Schatz, M. C. Targeted Nanopore Sequencing by Real-Time Mapping of Raw Electrical Signal with UNCALLED. *Nat. Biotechnol.* **2021**, *39* (4), 431–441. https://doi.org/10.1038/s41587-020-0731-9.

(4) Jain, M.; Koren, S.; Miga, K. H.; Quick, J.; Rand, A. C.; Sasani, T. A.; Tyson, J. R.; Beggs, A. D.; Dilthey, A. T.; Fiddes, I. T.; Malla, S.; Marriott, H.; Nieto, T.; O'Grady, J.; Olsen, H. E.; Pedersen, B. S.; Rhie, A.; Richardson, H.; Quinlan, A. R.; Snutch, T. P.; Tee, L.; Paten, B.; Phillippy, A. M.; Simpson, J. T.; Loman, N. J.; Loose, M. Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads. *Nat. Biotechnol.* **2018**, *36* (4), 338–345. https://doi.org/10.1038/nbt.4060.

(5) Howorka, S.; Siwy, Z. Nanopore Analytics: Sensing of Single Molecules. *Chem. Soc. Rev.* **2009**, *38* (8), 2360–2384. https://doi.org/10.1039/B813796J.

(6) Manrao, E. A.; Derrington, I. M.; Laszlo, A. H.; Langford, K. W.; Hopper, M. K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H. Reading DNA at Single-Nucleotide Resolution with a Mutant MspA Nanopore and Phi29 DNA Polymerase. *Nat. Biotechnol.* **2012**, *30* (4), 349–353. https://doi.org/10.1038/nbt.2171.

(7) Howorka, S.; Cheley, S.; Bayley, H. Sequence-Specific Detection of Individual DNA Strands Using Engineered Nanopores. *Nat. Biotechnol.* **2001**, *19* (7), 636–639. https://doi.org/10.1038/90236.

(8) Jain, M.; Abu-Shumays, R.; Olsen, H. E.; Akeson, M. Advances in Nanopore Direct RNA Sequencing. *Nat. Methods* **2022**, *19* (10), 1160–1164. https://doi.org/10.1038/s41592-022-01633-w.

(9) Motone, K.; Kontogiorgos-Heintz, D.; Wee, J.; Kurihara, K.; Yang, S.; Roote, G.; Fox, O. E.; Fang, Y.; Queen, M.; Tolhurst, M.; Cardozo, N.; Jain, M.; Nivala, J. Multi-Pass, Single-Molecule Nanopore Reading of Long Protein Strands. *Nature* **2024**, *633* (8030), 662–669. https://doi.org/10.1038/s41586-024-07935-7.

(10) Stephenson, W.; Razaghi, R.; Busan, S.; Weeks, K. M.; Timp, W.; Smibert, P. Direct Detection of RNA Modifications and Structure Using Single-Molecule Nanopore Sequencing. *Cell Genomics* **2022**, *2* (2), 100097. https://doi.org/10.1016/j.xgen.2022.100097.

(11) Gabrieli, T.; Sharim, H.; Nifker, G.; Jeffet, J.; Shahal, T.; Arielly, R.; Levi-Sakin, M.; Hoch, L.; Arbib, N.; Michaeli, Y.; Ebenstein, Y. Epigenetic Optical Mapping of 5-Hydroxymethylcytosine in Nanochannel Arrays. *ACS Nano* **2018**, *12* (7), 7148–7158. https://doi.org/10.1021/acsnano.8b03023.

(12) Kong, Y.; Zhang, Y.; Mead, E. A.; Chen, H.; Loo, C. E.; Fan, Y.; Ni, M.; Zhang, X.-S.; Kohli, R. M.; Fang, G. Critical Assessment of Nanopore Sequencing for the Detection of Multiple Forms of DNA Modifications. bioRxiv November 19, 2024, p 2024.11.19.624260. https://doi.org/10.1101/2024.11.19.624260.

(13) Zirkin, S.; Fishman, S.; Sharim, H.; Michaeli, Y.; Don, J.; Ebenstein, Y. Lighting Up Individual DNA Damage Sites by In Vitro Repair Synthesis. *J. Am. Chem. Soc.* **2014**, *136* (21), 7771–7776. https://doi.org/10.1021/ja503677n.

(14) Michaeli, Y.; Shahal, T.; Torchinsky, D.; Grunwald, A.; Hoch, R.; Ebenstein, Y. Optical Detection of Epigenetic Marks: Sensitive Quantification and Direct Imaging of Individual Hydroxymethylcytosine Bases. *Chem. Commun.* **2013**, *49* (77), 8599–8601. https://doi.org/10.1039/C3CC42543F.

(15) Nifker, G.; Levy-Sakin, M.; Berkov-Zrihen, Y.; Shahal, T.; Gabrieli, T.; Fridman, M.; Ebenstein, Y. One-Pot Chemoenzymatic Cascade for Labeling of the Epigenetic Marker 5-Hydroxymethylcytosine. *ChemBioChem* **2015**, *16* (13), 1857–1860. https://doi.org/10.1002/cbic.201500329.

(16) Avraham, S.; Schütz, L.; Käver, L.; Dankers, A.; Margalit, S.; Michaeli, Y.; Zirkin, S.; Torchinsky, D.; Gilat, N.; Bahr, O.; Nifker, G.; Koren-Michowitz, M.; Weinhold, E.; Ebenstein, Y. Chemo-Enzymatic Fluorescence Labeling Of Genomic DNA For Simultaneous Detection Of Global 5-Methylcytosine And 5-Hydroxymethylcytosine. *ChemBioChem* **2023**, *24* (20), e202300400. https://doi.org/10.1002/cbic.202300400.

(17) Neely, R. K.; Dedecker, P.; Hotta, J.; Urbanavičiūtė, G.; Klimašauskas, S.; Hofkens, J. DNA Fluorocode: A Single Molecule, Optical Map of DNA with Nanometre Resolution. *Chem. Sci.* **2010**, *1* (4), 453–460. https://doi.org/10.1039/C0SC00277A.

(18) Jin, X.; Hapsari, N. D.; Lee, S.; Jo, K. DNA Binding Fluorescent Proteins as Single-Molecule Probes. *Analyst* **2020**, *145* (12), 4079–4095. https://doi.org/10.1039/D0AN00218F.

(19) McCaffrey, J.; Sibert, J.; Zhang, B.; Zhang, Y.; Hu, W.; Riethman, H.; Xiao, M. CRISPR-CAS9 D10A Nickase Target-Specific Fluorescent Labeling of Double Strand DNA for Whole Genome Mapping and Structural Variation Analysis. *Nucleic Acids Res.* **2016**, *44* (2), e11. https://doi.org/10.1093/nar/gkv878.

(20) Grunwald, A.; Dahan, M.; Giesbertz, A.; Nilsson, A.; Nyberg, L. K.; Weinhold, E.; Ambjörnsson, T.; Westerlund, F.; Ebenstein, Y. Bacteriophage Strain Typing by Rapid Single Molecule Analysis. *Nucleic Acids Res.* **2015**, *43* (18), e117. https://doi.org/10.1093/nar/gkv563.

(21) Gabrieli, T.; Sharim, H.; Nifker, G.; Jeffet, J.; Shahal, T.; Arielly, R.; Levi-Sakin, M.; Hoch, L.; Arbib, N.; Michaeli, Y.; Ebenstein, Y. Epigenetic Optical Mapping of 5-Hydroxymethylcytosine in Nanochannel Arrays. *ACS Nano* **2018**, *12* (7), 7148–7158. https://doi.org/10.1021/acsnano.8b03023.

(22) Gabrieli, T.; Michaeli, Y.; Avraham, S.; Torchinsky, D.; Margalit, S.; Schütz, L.; Juhasz, M.; Coruh, C.; Arbib, N.; Zhou, Z. S.; Law, J. A.; Weinhold, E.; Ebenstein, Y. Chemoenzymatic Labeling of DNA Methylation Patterns for Single-Molecule Epigenetic Mapping. *Nucleic Acids Res.* **2022**, *50* (16), e92. https://doi.org/10.1093/nar/gkac460.

(23) Nifker, G.; Grunwald, A.; Margalit, S.; Tulpova, Z.; Michaeli, Y.; Har-Gil, H.; Maimon, N.; Roichman, E.; Schütz, L.; Weinhold, E.; Ebenstein, Y. Dam Assisted Fluorescent Tagging of Chromatin Accessibility (DAFCA) for Optical Genome Mapping in Nanochannel Arrays. *ACS Nano* **2023**, *17* (10), 9178–9187. https://doi.org/10.1021/acsnano.2c12755.

(24) Detinis Zur, T.; Margalit, S.; Jeffet, J.; Grunwald, A.; Fishman, S.; Tulpová, Z.; Michaeli, Y.; Deek, J.; Ebenstein, Y. Single-Molecule Toxicogenomics: Optical Genome Mapping of DNA-Damage in Nanochannel Arrays. *DNA Repair* **2025**, *146*, 103808. https://doi.org/10.1016/j.dnarep.2025.103808.

(25) Margalit, S.; Tulpová, Z.; Detinis Zur, T.; Michaeli, Y.; Deek, J.; Nifker, G.; Haldar, R.; Gnatek, Y.; Omer, D.; Dekel, B.; Baris Feldman, H.; Grunwald, A.; Ebenstein, Y. Long-Read Structural and Epigenetic Profiling of a Kidney Tumor-Matched Sample with Nanopore Sequencing and Optical Genome Mapping. *NAR Genomics Bioinforma.* **2025**, *7* (1), lqae190. https://doi.org/10.1093/nargab/lqae190.

(26) Margalit, S.; Tulpová, Z.; Detinis Zur, T.; Michaeli, Y.; Deek, J.; Nifker, G.; Haldar, R.; Gnatek, Y.; Omer, D.; Dekel, B.; Baris Feldman, H.; Grunwald, A.; Ebenstein, Y. Long-Read Structural and Epigenetic Profiling of a Kidney Tumor-Matched Sample with Nanopore Sequencing and Optical Genome Mapping. *NAR Genomics Bioinforma.* **2025**, *7* (1), lqae190. https://doi.org/10.1093/nargab/lqae190.

(27) Raseley, K.; Jinwala, Z.; Zhang, D.; Xiao, M. Single-Molecule Telomere Assay via Optical Mapping (SMTA-OM) Can Potentially Define the ALT Positivity of Cancer. *Genes* **2023**, *14* (6), 1278. https://doi.org/10.3390/genes14061278.

(28) Müller, V.; Dvirnas, A.; Andersson, J.; Singh, V.; Kk, S.; Johansson, P.; Ebenstein, Y.; Ambjörnsson, T.; Westerlund, F. Enzyme-Free Optical DNA Mapping of the Human Genome Using Competitive Binding. *Nucleic Acids Res.* **2019**, *47* (15), e89. https://doi.org/10.1093/nar/gkz489.

(29) Margalit, S.; Abramson, Y.; Sharim, H.; Manber, Z.; Bhattacharya, S.; Chen, Y.-W.; Vilain, E.; Barseghyan, H.; Elkon, R.; Sharan, R.; Ebenstein, Y. Long Reads Capture Simultaneous Enhancer–Promoter Methylation Status for Cell-Type Deconvolution. *Bioinformatics* **2021**, *37* (Supplement_1), i327–i333. https://doi.org/10.1093/bioinformatics/btab306.

(30) Sharim, H.; Grunwald, A.; Gabrieli, T.; Michaeli, Y.; Margalit, S.; Torchinsky, D.; Arielly, R.; Nifker, G.; Juhasz, M.; Gularek, F.; Almalvez, M.; Dufault, B.; Chandra, S. S.; Liu, A.; Bhattacharya, S.; Chen, Y.-W.; Vilain, E.; Wagner, K. R.; Pevsner, J.; Reifenberger, J.; Lam, E. T.; Hastie, A. R.; Cao, H.; Barseghyan, H.; Weinhold, E.; Ebenstein, Y. Long-Read Single-Molecule Maps of the Functional Methylome. *Genome Res.* **2019**, *29* (4), 646–656. https://doi.org/10.1101/gr.240739.118.

(31) Kinney, S. M.; Chin, H. G.; Vaisvila, R.; Bitinaite, J.; Zheng, Y.; Estève, P.-O.; Feng, S.; Stroud, H.; Jacobsen, S. E.; Pradhan, S. Tissue-Specific Distribution and Dynamic Changes of 5-Hydroxymethylcytosine in Mammalian Genomes*. *J. Biol. Chem.* **2011**, *286* (28), 24685–24693. https://doi.org/10.1074/jbc.M110.217083.

(32) Klimašauskas, S.; Weinhold, E. A New Tool for Biotechnology: AdoMet-Dependent Methyltransferases. *Trends Biotechnol.* **2007**, *25* (3), 99–104. https://doi.org/10.1016/j.tibtech.2007.01.006.

(33) Lukinavičius, G.; Tomkuvienė, M.; Masevičius, V.; Klimašauskas, S. Enhanced Chemical Stability of Adomet Analogues for Improved Methyltransferase-Directed Labeling of DNA. *ACS Chem. Biol.* **2013**, *8* (6), 1134–1139. https://doi.org/10.1021/cb300669x.

(34) Al Temimi, A. H. K.; Martin, M.; Meng, Q.; Lenstra, D. C.; Qian, P.; Guo, H.; Weinhold, E.; Mecinović, J. Lysine Ethylation by Histone Lysine Methyltransferases. *ChemBioChem* **2020**, *21* (3), 392–400. https://doi.org/10.1002/cbic.201900359.

(35) Willnow, S.; Martin, M.; Lüscher, B.; Weinhold, E. A Selenium-Based Click AdoMet Analogue for Versatile Substrate Labeling with Wild-Type Protein Methyltransferases. *ChemBioChem* **2012**, *13* (8), 1167–1173. https://doi.org/10.1002/cbic.201100781.

(36) Holstein, J. M.; Stummer, D.; Rentmeister, A. Enzymatic Modification of 5′-Capped RNA with a 4-Vinylbenzyl Group Provides a Platform for Photoclick and Inverse Electron-Demand Diels–Alder Reaction. *Chem. Sci.* **2015**, *6* (2), 1362–1369. https://doi.org/10.1039/C4SC03182B.

(37) Muttach, F.; Mäsing, F.; Studer, A.; Rentmeister, A. New AdoMet Analogues as Tools for Enzymatic Transfer of Photo-Cross-Linkers and Capturing RNA–Protein Interactions. *Chem. – Eur. J.* **2017**, *23* (25), 5988–5993. https://doi.org/10.1002/chem.201605663.

(38) Song, C.-X.; Szulwach, K. E.; Fu, Y.; Dai, Q.; Yi, C.; Li, X.; Li, Y.; Chen, C.-H.; Zhang, W.; Jian, X.; Wang, J.; Zhang, L.; Looney, T. J.; Zhang, B.; Godley, L. A.; Hicks, L. M.; Lahn, B. T.; Jin, P.; He, C. Selective Chemical Labeling Reveals the Genome-Wide Distribution of 5-Hydroxymethylcytosine. *Nat. Biotechnol.* **2011**, *29* (1), 68–72. https://doi.org/10.1038/nbt.1732.

(39) Dalhoff, C.; Lukinavicius, G.; Klimasăuskas, S.; Weinhold, E. Direct Transfer of Extended Groups from Synthetic Cofactors by DNA Methyltransferases. *Nat. Chem. Biol.* **2006**, *2* (1), 31–32. https://doi.org/10.1038/nchembio754.

(40) Josse, J.; Kornberg, A. Glucosylation of Deoxyribonucleic Acid: III. α- AND β-GLUCOSYL TRANSFERASES FROM T4-INFECTED ESCHERICHIA COLI. *J. Biol. Chem.* **1962**, *237* (6), 1968–1976. https://doi.org/10.1016/S0021-9258(19)73968-4.

(41) Pastor, W. A.; Huang, Y.; Henderson, H. R.; Agarwal, S.; Rao, A. The GLIB Technique for Genome-Wide Mapping of 5-Hydroxymethylcytosine. *Nat. Protoc.* **2012**, *7* (10), 1909–1917. https://doi.org/10.1038/nprot.2012.104.

(42) Han, D.; Lu, X.; Shih, A. H.; Nie, J.; You, Q.; Xu, M. M.; Melnick, A. M.; Levine, R. L.; He, C. A Highly Sensitive and Robust Method for Genome-Wide 5hmC Profiling of Rare Cell Populations. *Mol. Cell* **2016**, *63* (4), 711–719. https://doi.org/10.1016/j.molcel.2016.06.028.

(43) Klimasauskas, S.; Weinhold, E. A New Tool for Biotechnology: AdoMet-Dependent Methyltransferases. *Trends Biotechnol.* **2007**, *25* (3), 99–104. https://doi.org/10.1016/j.tibtech.2007.01.006.

(44) Kriukienė, E.; Labrie, V.; Khare, T.; Urbanavičiūtė, G.; Lapinaitė, A.; Koncevičius, K.; Li, D.; Wang, T.; Pai, S.; Ptak, C.; Gordevičius, J.; Wang, S.-C.; Petronis, A.; Klimašauskas, S. DNA Unmethylome Profiling by Covalent Capture of CpG Sites. *Nat. Commun.* **2013**, *4* (1), 2190. https://doi.org/10.1038/ncomms3190.

(45) Hanz, G. M.; Jung, B.; Giesbertz, A.; Juhasz, M.; Weinhold, E. Sequence-Specific Labeling of Nucleic Acids and Proteins with Methyltransferases and Cofactor Analogues. *J. Vis. Exp. JoVE* **2014**, No. 93, e52014. https://doi.org/10.3791/52014.

(46) Altemose, N.; Maslan, A.; Smith, O. K.; Sundararajan, K.; Brown, R. R.; Mishra, R.; Detweiler, A. M.; Neff, N.; Miga, K. H.; Straight, A. F.; Streets, A. DiMeLo-Seq: A Long-Read, Single-Molecule Method for Mapping Protein–DNA Interactions Genome Wide. *Nat. Methods* **2022**, *19* (6), 711–723. https://doi.org/10.1038/s41592-022-01475-6.

(47) Weng, Z.; Ruan, F.; Chen, W.; Chen, Z.; Xie, Y.; Luo, M.; Xie, Z.; Zhang, C.; Wang, J.; Sun, Y.; Fang, Y.; Guo, M.; Tan, C.; Chen, W.; Tong, Y.; Li, Y.; Wang, H.; Tang, C. BIND&MODIFY: A Long-Range Method for Single-Molecule Mapping of Chromatin Modifications in Eukaryotes. *Genome Biol.* **2023**, *24* (1), 61. https://doi.org/10.1186/s13059-023-02896-y.

(48) Shahal, T.; Gilat, N.; Michaeli, Y.; Redy-Keisar, O.; Shabat, D.; Ebenstein, Y. Spectroscopic Quantification of 5-Hydroxymethylcytosine in Genomic DNA. *Anal. Chem.* **2014**, *86* (16), 8231–8237. https://doi.org/10.1021/ac501609d.

(49) Yao, B.; Hsu, C.; Goldner, G.; Michaeli, Y.; Ebenstein, Y.; Listgarten, J. Effective Training of Nanopore Callers for Epigenetic Marks with Limited Labelled Data. *Open Biol.* **2024**, *14* (6), 230449. https://doi.org/10.1098/rsob.230449.

(50) *Products scheduled to be discontinued*. Oxford Nanopore Technologies. https://nanoporetech.com/products/prepare/discontinued-kits (accessed 2025-02-06).

(51) Sanderson, N. D.; Kapel, N.; Rodger, G.; Webster, H.; Lipworth, S.; Street, T. L.; Peto, T.; Crook, D.; Stoesser, N. Comparison of R9.4.1/Kit10 and R10/Kit12 Oxford Nanopore Flowcells and Chemistries in Bacterial Genome Reconstruction. *Microb. Genomics* **2023**, *9* (1), 000910. https://doi.org/10.1099/mgen.0.000910.

(52) Ni, Y.; Liu, X.; Simeneh, Z. M.; Yang, M.; Li, R. Benchmarking of Nanopore R10.4 and R9.4.1 Flow Cells in Single-Cell Whole-Genome Amplification and Whole-Genome Shotgun Sequencing. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 2352–2364. https://doi.org/10.1016/j.csbj.2023.03.038.

(53) Yue, X.; Xie, Z.; Li, M.; Wang, K.; Li, X.; Zhang, X.; Yan, J.; Yin, Y. Simultaneous Profiling of Histone Modifications and DNA Methylation via Nanopore Sequencing. *Nat. Commun.* **2022**, *13* (1), 7939. https://doi.org/10.1038/s41467-022-35650-2.

(54) Thatavarty, A.; Sagy, N.; Erdos, M. R.; Lee, I.; Simpson, J. T.; Timp, W.; Collins, F. S.; Bar, D. Z. Detecting Protein-DNA Binding in Single Molecules Using Antibody Guided Methylation. bioRxiv November 20, 2023, p 2023.11.20.567792. https://doi.org/10.1101/2023.11.20.567792.

(55) Janssen, S. M.; Lorincz, M. C. Interplay between Chromatin Marks in Development and Disease. *Nat. Rev. Genet.* **2022**, *23* (3), 137–153. https://doi.org/10.1038/s41576-021-00416-x.

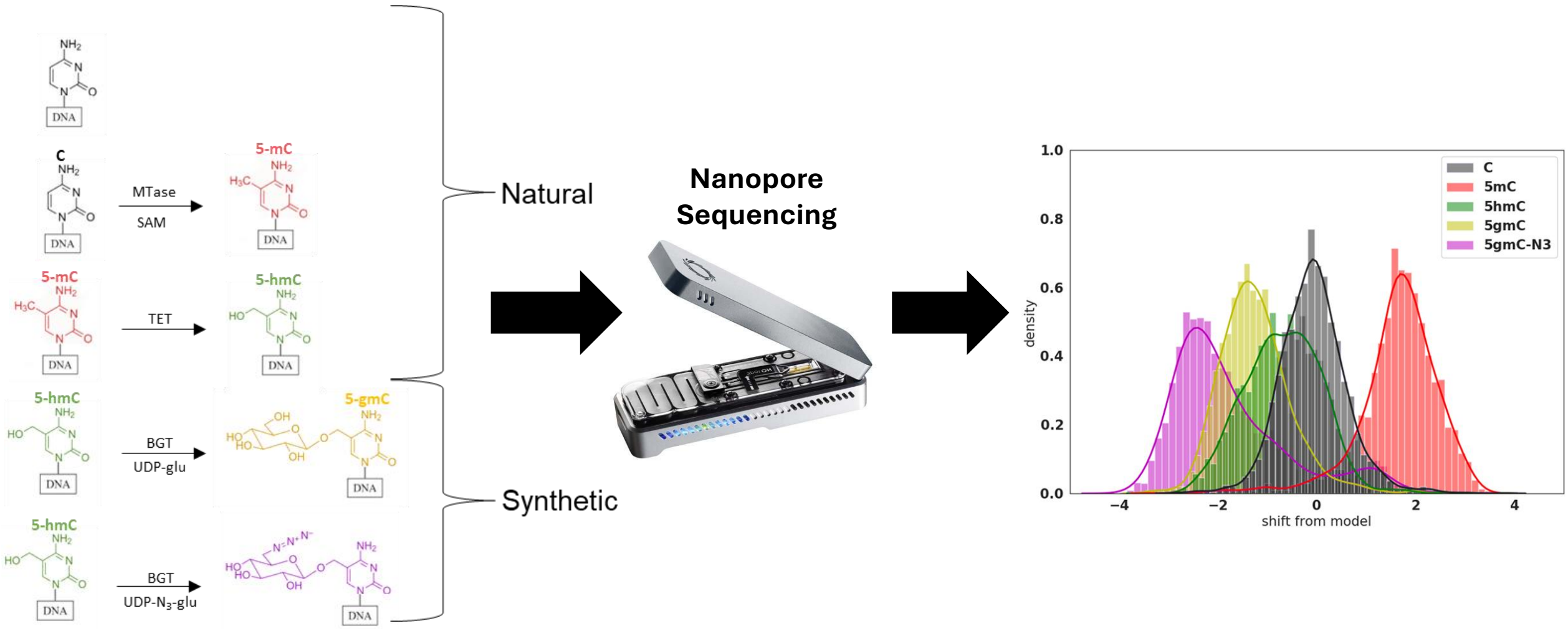
(56) Escherichia Coli Str. K-12 Substr. MG1655, Complete Genome, 2022. http://www.ncbi.nlm.nih.gov/nuccore/NC_000913.3 (accessed 2025-03-10).

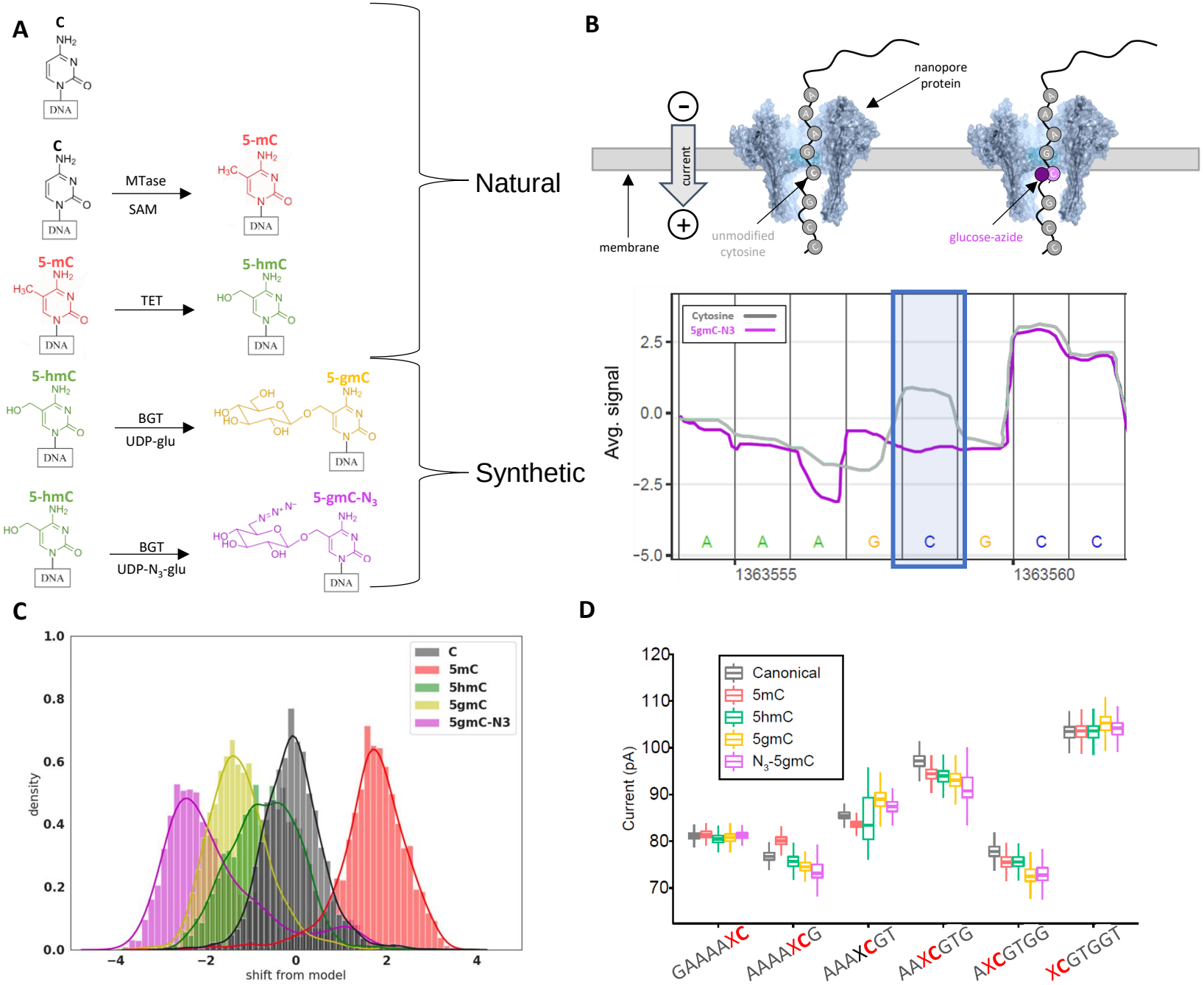
(57) ONT Fast5 API. https://github.com/nanoporetech/ont_fast5_api (accessed 2024-11-14).

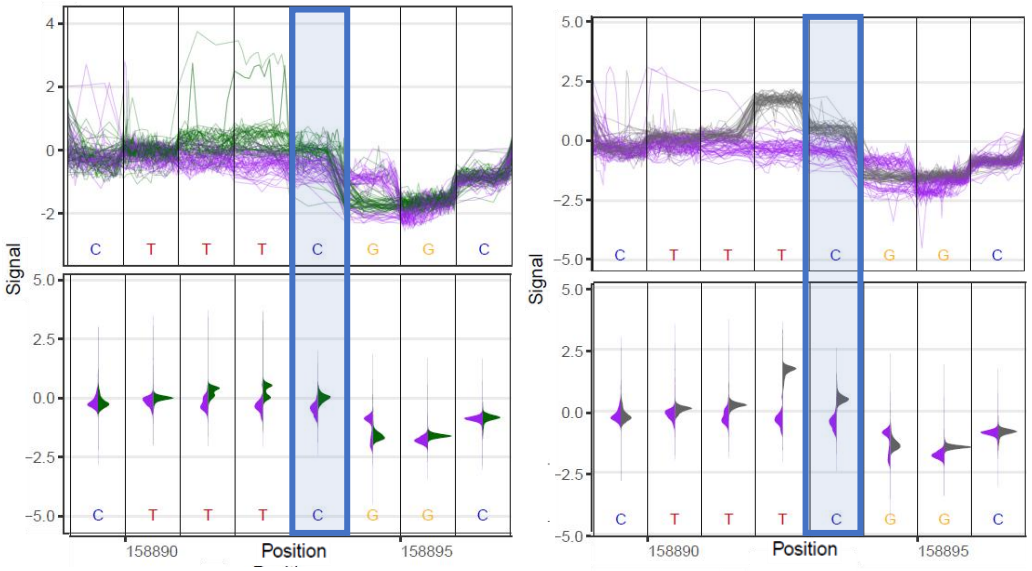
(58) Tombo, 2024. https://github.com/nanoporetech/tombo (accessed 2024-11-14).

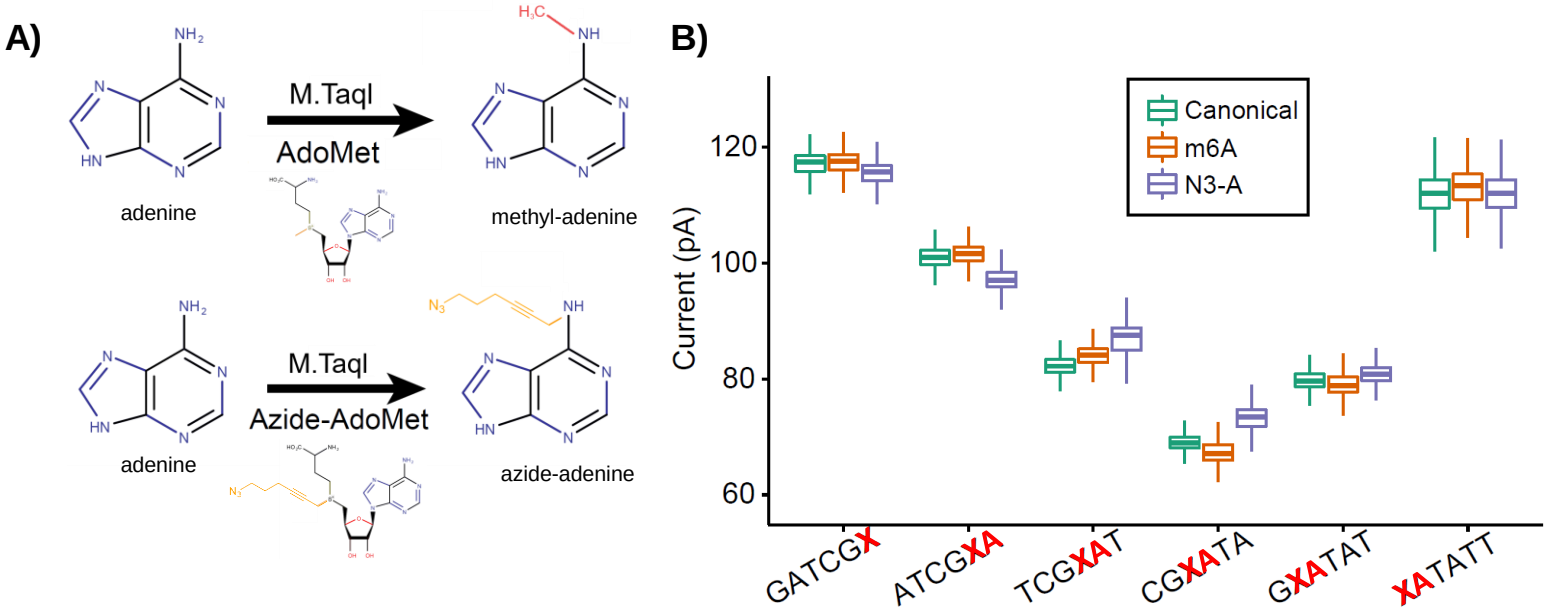
(59) Riley, M.; Abe, T.; Arnaud, M. B.; Berlyn, M. K. B.; Blattner, F. R.; Chaudhuri, R. R.; Glasner, J. D.; Horiuchi, T.; Keseler, I. M.; Kosuge, T.; Mori, H.; Perna, N. T.; Plunkett, G.; Rudd, K. E.; Serres, M. H.; Thomas, G. H.; Thomson, N. R.; Wishart, D.; Wanner, B. L. Escherichia Coli K-12: A Cooperatively Developed Annotation Snapshot--2005. *Nucleic Acids Res.* **2006**, *34* (1), 1–9. https://doi.org/10.1093/nar/gkj405.

(60) Loman, N. J.; Quick, J.; Simpson, J. T. A Complete Bacterial Genome Assembled de Novo Using Only Nanopore Sequencing Data. *Nat. Methods* **2015**, *12* (8), 733–735. https://doi.org/10.1038/nmeth.3444.


C
MTase
SAM
5-mC
5-mC
TET
5-hmC
5-hmC
BGT
UDP-glu
5-gmC
5-hmC
BGT
UDP-N3-glu
DNA
Natural
Synthetic
Nanopore Sequencing
C
5mC
5hmC
5gmC
5gmC-N3
density
shift from model


A
C
DNA
MTase
SAM
5-mC
TET
5-hmC
BGT
UDP-glu
5-gmC
BGT
UDP-N3-glu
5-gmC-N3
Natural
Synthetic
B
nanopore protein
current
membrane
unmodified cytosine
glucose-azide
Cytosine
5gmC-N3
Avg. signal
2.5
0.0
-2.5
-5.0
A A A G C G C C
1363555
1363560
C
1.0
0.8
0.6
0.4
0.2
0.0
density
C
5mC
5hmC
5gmC
5gmC-N3
−4
−2
0
2
4
shift from model
D
120
110
100
90
80
70
Current (pA)
Canonical
5mC
5hmC
5gmC
N3-5gmC
GAAAAXC
AAAAXCG
AAAXCGT
AAXCGTG
AXCGTGG
XCGTGGT


Signal
4
2
0
-2
C T T T C G G C
5.0
2.5
0.0
−2.5
−5.0
C T T T C G G C
158890
Position
158895
Signal
5.0
2.5
0.0
−2.5
−5.0
C T T T C G G C
5.0
2.5
0.0
−2.5
−5.0
C T T T C G G C
158890
Position
158895


A)
M.TaqI
AdoMet
adenine
methyl-adenine
M.TaqI
Azide-AdoMet
adenine
azide-adenine
B)
Canonical
m6A
N3-A
Current (pA)
120
100
80
60
GATCGX
ATCGXA
TCGXAT
CGXATA
GXATAT
XATATT