Supplementary Tables

Supplementary Table 1. Oligonucleotides.

DNA	Sequence	Purpose
254mer GeneBlock	gtcactcagATGTATAGAATGATGAGTTAGGTAGTGTTGATATGGGTTATGAATGAAGTAGTCGATCTTTCATCATATTCTAGATCCCTCTGAAAAAATCTTCCGAGTTTGCTAGGCAGTGATACATAACTCTTTTCCAATAATTGGGGAAGTCATTCAAATCTATAATAGGTTTCAGATTTAATTCTGACTGTAGCTGCTGAAACGTTGCGGAGTGTTAAGGTATATGAGTAGATGATTGATTGGGTATGTTGATAAGTGTAgtcactcag	Generate DNA substrate with homogenously modified cytosines
OTF12	ATGTATAGAATGATGAGTTAGGTAGTGTTGATATGGGTTATGAATGAAGTA	Generate DNA substrate with homogenously modified cytosines
OTR12	TACACTTATCAACATACCCAATCAATCATCTACTCATATACCTTAACACT	Generate DNA substrate with homogenously modified cytosines
OTF2_TruSeq	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGATATGGGTTATGAATGAAGTA	Primers for installing Illumina overhangs
OTR2_TruSeq	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGTGTTAAGGTATATGAGTAGATGA	Primers for installing Illumina overhangs
Taq$^{\alpha}$I-Top-C-FAM	FAM-AGTATGTAGTTGAATTCGATTGATAGGAGG	A3A deamination assay
Taq$^{\alpha}$I-Top-T-FAM	FAM-AGTATGTAGTTGAATTTGATTGATAGGAGG	A3A deamination assay
Taq$^{\alpha}$I-Bottom-5mC	CCTCCTATCAAT/5mC/GAATTCAACTACATACT	A3A deamination assay
Taq$^{\alpha}$I-Bottom-C	CCTCCTATCAATCAAATTCAACTACATACT	A3A deamination assay
HpaII-Top-C-FAM	GTATCTAGTTCAATCCGGTTCATAGCA-FAM	CxMTase opposite strand assay
HpaII-Bottom-5mC	TGCTATGAAC/5mC/GGATTGAACTAGATAC	CxMTase opposite strand assay
HpaII-Bottom-C	TGCTATGAACCGGATTGAACTAGATAC	CxMTase opposite strand assay
Unmethylated Copy Primer (v1)	GACTGGAGTTCAGACGTGT	copy strand synthesis
Methylated Copy Primer (v2)	GA/5mC/TGGAGTT/5mC/AGA/5mC/GTGTG/5mC/T/5mC/TT/5mC/GAT/5mC/	copy strand synthesis

Supplementary Table 2. PCR Methods.

Purpose	Method
Indexing PCRs	98°C for 30s, 10x [98°C for 10s, 65°C for 30s, 68°C for 60s], 4°C hold
254mer generation PCR (OTF12/R12)	95°C for 5 min, 35x [95°C for 30s, 62°C for 30s, 72°C for 30s], 72°C for 5 min, 4°C hold
254mer post deamination PCR	95°C 5min, 25x [95°C 15s, 50°C 30s, 60°C 15s], 60°C 5min, 4°C hold.

Supplementary Table 3. Final Library Statistics

Method	Condition	Indexing PCR cycles	ng / µL final library	total reads	reads after trimming	reads after alignment	reads after deduplication	reads after filtering
BS-Seq	-MTase	17	0.308	16716	16507	16050	14067	
BS-Seq	+M.SssI	17	0.222	4161	3150	2965	2826	
BS-Seq	+M.CviPI	17	0.664	15809	15636	13831	12429	
DM-Seq	-MTase	17	36.0	67400	67333	63973	44799	44377
DM-Seq	+M.SssI	17	30.6	24830	24598	23803	20572	20180
DM-Seq	+M.CviPI	17	23.8	55542	55304	53533	39379	37865
BS-Seq	GBM	14	8.30	223430253	222555471	172961545	89542258	
DM-Seq	GBM	13	23.0	223862027	223764985	165058698	107711545	105968995

Supplementary Table 4. Downsampling Statistics.

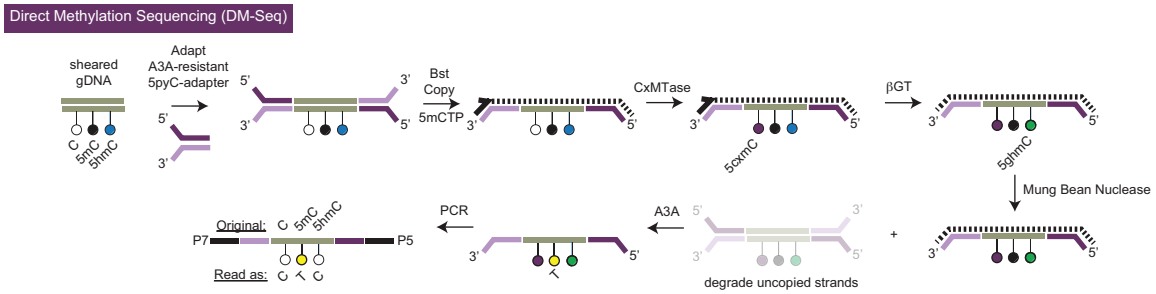
	BS-Seq	DM-Seq
unique CpGs sequenced	25,138,620	30,369,633
% modification of all CpGs sequenced	74.4	75.9
shared CpGs sequenced	16,438,445	
% modification of shared CpGs	76.0	75.4
High 5hmCpGs covered by each data set	2,132	2,538
% modification at high 5hmCpGs	75.0	62.9
High 5hmCpGs covered by both datasets	1,485	
% modification of shared high 5hmCpGs	75.6	61.4
Number of downsamplings	10,000	
Number of CpGs in each simulation	1,485	
Mean modification for all downsamplings	76.0	75.4
Standard deviation for all downsamplings	1.14	1.06

Supplementary Note: Full DM-Seq Protocol

Introduction

Direct Methylation Sequencing (DM-Seq) detects 5-methylcytosine (5mC) at single-base resolution at CpGs using low genomic DNA input. In this method, 5-propynylcytosine (5pyC) containing adapters, which are resistant to deamination by the enzyme APOBEC3A (A3A), are first ligated to sheared genomic DNA. These adapters also serve as a scaffold to prime the creation of a methylated copy strand which is a favorable substrate for DNA carboxymethylation. A DNA carboxymethyltransferase (CxMTase) protects all unmodified CpGs by creating 5-carboxymethylcytosines (5cxmC) and glucosylation with βGT protects all 5-hydroxymethylcytosines (5hmC). The A3A deaminase then only acts on the unprotected 5mC, generating T, and subsequent PCR amplification identifies 5mCs as converted Ts in sequencing. Key advantages of this workflow are that it is 1) direct for sequencing 5mC and 2) non-destructive as compared to chemical methods such as bisulfite sequencing.

A detailed protocol for DM-Seq is provided below.



Materials

Note: while in-house purified A3A enzyme was used for the majority of experiments in this manuscript, we have validated both ACE-Seq and DM-Seq (**Supplementary Figs. 7 and 8**) using commercially-available A3A enzyme (E7125, NEB) and have written this protocol using this more readily available commercial source. We further anticipate that critical reagents such as the DNA CxMTase, CxSAM, and 5pyC-containing adapters will also be available from accessible sources in the future.

- Samples of interest
- Sonicator for shearing DNA (e.g. Covaris M220 Focused Ultrasonicator) with corresponding shearing tubes (e.g. 50 μL microTUBEs, 520166)
- Spike-in controls
    - Unmethylated pUC19 (as a control for DNA carboxymethylation and protection of 5cxmC from deamination)
        - *Unmethylated pUC19 plasmid DNA (dam+/dcm+/cpg-), from standard plasmid preparation.*
    - Methylated M.SssI-treated lambda phage gDNA (as a control for complete deamination of 5mC)
        - *Unmethylated lambda phage DNA (D1521, Promega) was enzymatically methylated at all CpG sites using M.SssI methyltransferase (EM0821, Thermo).*
    - Hydroxymethylated T4-hmC DNA (as a control for βGT protection of 5hmC from deamination)
        - *Genomic DNA from a mutant T4 phage in which inactivating mutations exist in both the alpha- and beta-glucosyltransferase genes, but the machinery to synthesize dhmCTP remains intact. Therefore, every "C" in this genome is 5hmC.*
- NEBNext Ultra II End Prep Mix and Buffer (E7546S)
- NEBNext Ultra II Ligation Master Mix and Ligation Enhancer (E7595S)
- IDT xGEN Duplexed Y-shaped Adapters (15 μM)
    - 5pyC (custom synthesis, IDT, analogous to 10005974)
- Methylated Copy Primer (v2, 10 μM)
- Bst large fragment polymerase (8000 U/mL, NEB – M0275S)
- 10x NEB CutSmart (B7004S)
- Nucleotides
    - dATP, dGTP, dTTP (10 mM, NEB - U120B, U121B, U123C)
    - dmCTP (10 mM, NEB - N0356S)
- Mung Bean Nuclease (NEB - M0250S)
- Proteinase K (NEB - P8107S)
- CxMTase reagents
    - 5x CxMTase Buffer (50 mM Tris-HCl, 250 mM NaCl, 50 mM EDTA, 5 mM DTT, pH 7.9 at 25°C)
    - MBP-M.MpeI-N374K (2.5 μM) – prepped in-house
    - CxSAM (diluted to 1.6 mM) – prepped in-house
- APOBEC3A deamination
    - A3A enzyme (E7125, NEB)
    - Two thermocyclers
    - PCR sample cooler rack stored at -20°C
    - 100% Formamide

    - 10x A3A reaction buffer (E7125, NEB)
- NEBNext Multiplex Oligos for Illumina - Dual Indices (E7600S)
- PCR Master Mixes
    - KAPA qPCR Master Mix (Roche – KK4621)
    - KAPA 2x Hot Start Uracil+ Ready Mix (Roche – KK2802)
- qPCR 384 well plate (4309849) and covers (AB0558)
- Agilent 2100 Bioanalyzer Instrument and High Sensitivity DNA Kit (5067-4626)
- Illumina reagents (e.g. MiSeq Reagent Nano Kit V2, MS-103-1001)
- 0.2M NaOH, diluted from 1M
- Nuclease Free H_2O (AM9937)
- Low EDTA TE (0.22 μm filtered, 10 mM Tris Cl, pH 8.0 at 25°C, 0.1 mM EDTA)
- SPRIselect Beads (Beckman Coulter – B23318) and magnetic rack (S1515S)
- Qubit 3.0 Fluorimeter
    - Qubit dsDNA BR reagent (Q32850, 100 pg/μL - 1000 ng/μL)
    - Qubit dsDNA HS reagent (Q32851, 10 pg/μL - 100 ng/μL)

Procedure

Shear DNA

1. Dilute DNA into a DNA LoBind tube with low EDTA TE to a final volume of 50 μL.
2. Add whole-genome spike-ins at a total ratio of 1:100 spike-ins : gDNA (m:m, 1:300 each). Note that more dilute concentrations can be used if not utilizing shallow (e.g. MiSeq) sequencing as a control.
    a. Unmethylated pUC19
    b. CpG-Methylated M.SssI-treated lambda gDNA
    c. Fully hydroxymethylated T4-hmC phage DNA
3. Use Covaris Sonicator as in typical sonication protocol for appropriate size (50 μL volume)
    a. ~550 bp: 45s duration, 500 cycles per burst, 25 watts, 10% duty factor
4. Perform a left-sided 1.2x SPRI purification and elute in 40 μL low EDTA TE
5. Quantify samples with Qubit HS reagent.

End-Repair DNA

1. Assemble the reactions in a total volume of 60μL using the following chart on a per reaction basis.

	μL
NEBNext Ultra II End Prep Enzyme Mix:	3
NEBNext Ultra II End Prep Reaction Buffer:	6
Fragmented DNA (500 pg - 1 μg total):	51

2. Incubate samples with the following thermocycler method: 20°C for 30 minutes, 65°C for 30 minutes, 4°C hold. Set the heated lid to 75°C
3. Fragmented DNA input should be within the range of 500 pg to 1 μg, and diluted in low EDTA TE per NEB kit instructions, although lower input is possible.

Ligate Illumina Y-Shaped Adapters

1. Anneal the appropriate volume of Y-shaped adapters following the thermocycler method: 95°C 5min, 70°C to 40°C in steps of 5ºC per min, 37°C 5min, 4°C hold.
2. Assemble the reactions in a total volume of 93.5 μL using the following chart on a per reaction basis.

	Vol (μL)
End Prep Reaction Mixture (unpurified)	60
Annealed IDT xGEN Y-shaped Adapter (15 μM, in low EDTA TE)	2.5
NEBNext Ultra II Ligation Master Mix	30
NEBNext Ligation Enhancer	1
Total Volume	93.5

3. Incubate the samples at the thermocycler method: 20°C for 15 minutes, 4°C hold. Turn the heated lid off.
4. Perform a left-sided 1.2x SPRI purification and elute in > 6 μL Nuclease-Free H_2O. Do not vortex samples.

5. Quantify samples using Qubit HS reagent.

Note: Following adapter ligation, the samples may be stored at -20℃ until ready to proceed.

Anneal copy primer and extend 5mC copy strand

Note 1: We chose the Bst large fragment polymerase since it lacks 5'-3' exonuclease activity but has excellent strand displacement activity. We also found that a methylated copy primer increases copy strand efficiency, possibly due to an increase in annealing temperature.

Note 2: We find that the copy strand does not readily sequence based on the CpG and 5hmCpG spike-in phage controls (Figure 4a). If the copy strand were to be amplified after enzymatic deamination, we would anticipate reads where the associated CpGs from the copy stand would read as T rather than C. The accuracy of our detection of unmodified CpG (98.9%) and 5hmCpG (99.7%) with the spike-in control phage DNA indicate that this is not the case.

1. Decide on the amount of ligated input DNA to use. ~1-10 ng was used in our study.
2. Assemble the reactions in a total volume of 9 μL using the following chart on a per reaction basis.

5pyC ligated sample	Up to 9 μL in Nuclease-Free H_2O
10x NEB CutSmart	1 μL
Methylated Copy Primer (v2, 10 μM)	1 μL
10 mM dATP, dGTP, dTTP	1 μL
10 mM dmCTP	1 μL

3. Incubate samples at the thermocycler method: 95°C 5min, 70°C to 40°C in steps of 5ºC per min, 37°C 5min, 4°C hold.
4. Add 1μL Bst large fragment polymerase to each sample.
5. Incubate samples at the thermocycler method: 65°C 30min, 4°C hold.

Protect 5hmC with βGT

1. Assemble the reactions in a total volume of 20 μL using the following chart on a per reaction basis.

Unpurified and copied sample	10 μL
UDP-Glucose	0.4 μL
T4-βGT	1 μL
10x NEB CutSmart buffer	2 μL
nfH_2O	6.6 μL

2. Incubate the samples at the thermocycler method: 37°C 1 hour, 4°C hold

Degrade uncopied strands

1. Adding 1μL of Mung Bean Nuclease to each sample and incubate samples at the thermocycler method: 30°C 30 min, 4°C hold
2. Perform a left-sided 1.2x SPRI purification and elute in 6 μL low EDTA TE.

Note: Make sure sample cools to 4°C before adding Mung Bean Nuclease. This step is an added precaution and is optional.

DNA carboxymethylation

Note: This is the critical enzymatic reaction will convert C to 5cxmC, which is resistant to A3A deamination. 16 hrs was selected for a convenient overnight protocol. Omission of carboxymethylation and copy strand steps sequences 5hmC at base resolution, analogous to ACE-Seq.

1. Assemble the reactions in a total volume of 10 μL using the following chart on a per reaction basis.

Purified DNA	5 μL
5x CxMTase buffer	2 μL
MBP-M.MpeI N374K-His enzyme	2 μL
CxSAM (1.6mM)	1 μL

2. Incubate the samples at the thermocycler method: 37°C 16 hours, 95°C 5 min, 4°C hold

Denaturation and A3A Treatment

1. Treat the remaining post-carboxymethylation DNA with 0.5 μL Proteinase K and incubate at 37°C for 15 minutes.
2. Perform a left-sided 1.2x SPRI purification and elute in 16 μL nuclease-free H_2O. Do not vortex samples.
3. Preheat one thermocycler to 85°C (with heated lid). On the other thermocycler, preheat with the "ramp method," which follows the method:

4°C	10 min
4°C - 50°C	2:15 per degree of ramp (~2 hours total)
50°C	10 min
4°C	Hold

Note 1: The rationale for the ramping conditions is based on balancing the kinetic tendency for reannealing with A3A deamination efficiency. We compared this to incubation at 37°C alone and this ramping method is more favorable for complete deamination. When the ramp starts at a low temperature, the strands should not favor reannealing and A3A deamination will start. As the temperature slowly increases, A3A deamination efficiency will increase, leading to more deamination events; even though the thermal fluctuations may increase and the tendency of the DNA to reanneal will be higher, the fact that some deamination will already have occurred can disrupt base pairing and disfavor reannealing. The protocol thus aims to increase single-stranded-ness and promote more efficient deamination.

Note 2: We have seen slightly more consistent performance using formamide denaturation / heat denaturation relative to DMSO / heat denaturation.

4. Assemble the reactions in a total volume of 20 μL using the following chart on a per reaction basis by pipetting up and down multiple times.

Purified post-CxMTase DNA	16 μL
Formamide	4 μL

5. Incubate the samples at 85°C for 10 minutes.
6. Assemble the below master mix in a total volume of 80 μL using the following chart on a per reaction basis.

nfH_2O	68 μL
NEB APOBEC Reaction Buffer	10 μL
NEB BSA	1 μL
NEB APOBEC	1 μL

7. After DNA is fully denatured for 10 minutes in formamide, swiftly transfer samples to PCR sample cooling rack pre-chilled to -20°C.
8. Add the master mix containing the APOBEC enzyme (80 μL / reaction) to each denatured sample (20 μL) while still on the 20°C block.
9. Spin down on tabletop microfuge.
10. Incubate the samples using the "ramp method."
11. Following deamination, perform a left-sided 1.2x SPRI purification and elute in 10 μL nuclease-free H_2O. Do not vortex samples.

qPCR

Note: The purpose of the qPCR is to determine an approximate number of indexing cycles for each sample, as over-indexing leads to bias in the final sequencing libraries.

1. Assemble the reactions in a total volume of 10 μL using the following chart on a per reaction basis.

nfH_2O	3 μL
2x KAPA qPCR Master Mix (KK4621)	5 μL
Purified, deaminated DNA	1 μL
2μM indexing primers	1 μL

2. Carefully pipette each sample into the 384-well plate and start qPCR method (Applied Biosystems Real-Time PCR instrument): 95°C 3min, 25x (95°C 15s, 63°C 30s, 72°C 30s), 4°C hold.
    a. This method will take 52 minutes to run.
3. For each sample, record the number of cycles required to reach the half-maximum value. The maximum value can be found by approximating where each sample curve plateaus.
    a. Use these values to assign cycles for the indexing PCR.

Indexing PCR

1. Assign unique indices to each sample.

2. Assemble the reactions in a total volume of 50 μL using the following chart on a per reaction basis.
    a. Note: the 50 μL volume should be used for mammalian DNA samples.

nfH_2O	5 μL
2x KAPA Uracil+ PCR Master Mix (KK2802)	25 μL
Purified, deaminated DNA	19 μL
Combined i5/i7 indexing primer	1 μL

3. Incubate samples with the following thermocycler method: 98°C 30s, variable x (98°C 10s, 65°C 30s, 68°C 1 min), 4°C hold.
4. Perform a left-sided 0.8x SPRI purification and elute in 20 μL low EDTA TE.

BioAnalyzer Quality Control

1. Quantify all samples with Qubit HS dsDNA reagent
2. Ensuring 100pg/μL – 10ng/μL concentrations before loading on BioAnalyzer.
3. Use BioAnalyzer to estimate average library size, as well as assessing need for an additional SPRI purification due to adapter dimers.

Illumina Sequencer

1. Load sequencer as in manufacturer instructions utilizing library size and DNA concentration.