nature aging

Technical Report

https://doi.org/10.1038/s43587-023-00555-2

TIME-seq reduces time and cost of DNA methylation measurement for epigenetic clock construction

In the format provided by the authors and unedited

SUPPLEMENTARY INFORMATION

Supplementary Note

Detailed Supplementary Methods

I. DNA extraction, quantification, and normalization

For blood, 100-300 µl of pelleted mouse whole blood (plasma removed) was resuspended in 1mL of red blood cell (RBC) lysis buffer (155 mM NH4Cl, 12 mM NaHCO3, 0.1 mM EDTA, pH 7.3), incubated for 10 minutes on ice, and centrifuged at 2000 RCF for 5 minutes. Pelleted cells were resuspended in 1.5mLs of RBC lysis buffer and spun twice before being lysed in 800 µl TER (50 mM Tris-HCl pH8, 10 mM EDTA, 40 µg/mL RNase A) with 50 µl of 10% SDS added. Lysates were incubated at 37ºC to allow for RNA degradation and then 50 µl of 20 mg/mL proteinase K was added and samples were incubated overnight at 65ºC. To purify DNA, 500 µl of 1:5 diluted (dilution buffer: 20% PEG 8000, 2.5 M NaCl, 10 mM Tris-HCl pH8, 1 mM EDTA, 0.05% Tween-20) SPRI DNA binding beads were added to each sample, and they were incubated with rotation for 30 minutes at room temperature. Tubes were then placed on magnetic racks to capture SPRI beads, and the beads were washed with 1 mL of ice-cold 80% ethanol twice. DNA was eluted in 75 µl of 10mM Tris-HCl (pH 8). Purified DNA was quantified using the Qubit double-stranded DNA broad range kit (Catalog No. Q32850, ThermoFisher) and diluted to 10 ng/µl for TIME-Seq reactions. To check for contaminants, a subset of samples from each extraction were assessed by NanoDrop. For tissues, frozen tissues from mice of various ages were collected from our internal mouse aging tissue bank. Tissues were kept frozen on dry ice while 10-20mg of tissue was cut into a well of a 96-well plate. DNA was extracted using the Chemagic 360 system (Perkin Elmer) and the 10 mg tissue kit (Perkin Elmer, CMG-723) and resuspended in 100 µL of elution buffer.

To achieve consistent DNA normalization from four 96-well DNA plates at a time, we developed a scalable and partially automated protocol to get concentrations of 10 ng/µL with a low-percentage drop-out due to over- or under-dilution. To begin, a standard starting volume of approximately 25 µL of DNA was transferred to a 96-well plate (BioRad, HSP9601). 1µL of unnormalized DNA was removed and transferred to a black 96-well plate (VWR, 33501-812) with 99µL of TE (10mM Tris, 1mM EDTA). 100µL

of 1X picogreen solution (ThermoFisher, P7589) was added to the DNA, after which the plates were sealed, vortexed for 30 seconds, and spun in a plate spinner. Fluorescence was measured by the SpectraMax i3 (Molecular Devices). Standards in triplicate from 1-100ng of total DNA were used to predict the amount of DNA from each sample. Using these DNA concentrations, samples were diluted with 10 mM Tris pH 8 using the Mantis liquid handler (Formulatrix) to a concentration of 40 ng/µL if ≥ 100 ng/µL or 20 ng/µL if < 100 ng/µL. After a second round of DNA quantification, samples were then transferred to a new 96-well plate at a standard volume (typically 20-30 µL) using a BioMek FXp liquid handler (Beckman Coulter). This plate was diluted to a final TIME-Seq starting concentration of 10 ng/µL using the Mantis liquid handler.

Liver DNA from intervention samples was prepared as detailed above using the Chemagic 360 system for DNA extraction and our partially automated DNA normalization protocol. Normalized DNA was prepared with TIME-Seq using mouse clock enrichment baits.

II. Tn5 purification, activation, and activity testing

Tn5 transposase was purified according to a described optimized protocol[49]. 1 mL of overnight culture containing pTXB1-Tn5 (Addgene plasmid #60240) was used to inoculate 1 L of ZYM-505 growth media containing 100 µg/mL ampicillin and 0.001% polypropylene glycol (L14699-AE, Alfa Aesar). After the culture was grown for 4 hours at 37ºC, IPTG (0487-10G, VWR) was added to 0.25 mM, and the culture was grown for an additional 4 hours at 18ºC. The culture was centrifuged at 20,000 RPM (47,800 x g) for 25 minutes, and the pelleted culture was flash-frozen in liquid nitrogen before being stored overnight at -80ºC. The pellet was thawed on ice, resuspended in 10 mL of HEGX buffer (20mM HEPES-KOH pH 7.2, 0.8 M NaCl, 1mM EDTA, 10% glycerol, 0.2% Triton X-100) with a Roche protease inhibitor (SKU11697498001, Millipore Sigma), and 1% w/v of pre-dissolved (50% w/v) octyl-thioglucoside (O-130-5 Gold Bio) was added to help lysis. After a 10-minute incubation on ice, 100 mL of HEG-X was added, and the lysate was transferred to a glass beaker for sonication on the 550 Sonic Dismembrator (ThermoFisher) using 15 cycles (15 seconds on, 30 seconds off) on 70% duty and power 7. The sonicated

lysate was pelleted at 20,000 RPM (47,800 x g) for 30 minutes at 4ºC. The supernatant was transferred to a clean beaker with a stir bar, placed on a magnetic stir plate, and 10% PEI was added dropwise while stirring to remove excess bacterial DNA. After 15 minutes, PEI and precipitated DNA were removed by spinning the mixture at 20,000 RPM (47,800 x g) for another 30 minutes at 4ºC. The lysate was added to 2 chromatography columns (7321010, BioRad) packed with 25 mL each of chitin resin (S6651S, NEB) and equilibrated with 100 mL of HEG-X each. The supernatant was added to the columns in equal proportion and allowed to flow through before the column was washed with 30 column volumes of HEG-X. To elute the purified Tn5, 25 mLs of HEG-X with 100 mM DTT was added to each column and 10 mLs was allowed to flow through before sealing the column and letting it stand for 44 hours. 27 mL of elution was collected, and a Bradford assay (23200, ThermoFisher) was used to quantify protein. The elution was then concentrated to 20 mL using Amplicon Ultra 30 kDA filters (UFC900308, Millipore Sigma) and dialyzed in a Slide-A-Lyzer (66212, ThermoFisher) cassette with 1 L dialysis buffer (50 mM HEPES-KOH, pH 7.2 0.8M NaCl, 0.2 mM EDTA, 2 mM DTT, 20% glycerol). After two rounds of dialysis totaling 24 hours, the eluted protein was removed, aliquoted into 1 mL microcentrifuge tubes, and flash-frozen in liquid nitrogen before being stored at -80ºC. The final concentration of purified protein was 1.5 mg/mL, and we estimate that 1L of culture produced enough Tn5 for approximately 16,000 TIME-Seq reactions with 100 ng of DNA per sample.

TIME-Seq adaptor oligonucleotides (oligos) were ordered from IDT and HPLC purified except for TIME-Seq indexed adaptors and hybridization blocking oligos (see Supplementary Table 5 for all oligonucleotide sequences). To anneal TIME-Seq adaptors, 100 µM TIME-Seq adaptor B containing a 5-bp internal barcode and (separately) 100 µM methylated adaptor A were combined in equal volume with the 100 µM Tn5 reverse ME oligo. Enough methylated A adaptor was annealed to be added in equal proportion with each indexed adaptor B. Oligos were denatured at 95ºC for 2 minutes and then ramped down to 25ºC at 0.1ºC/s. The annealed oligos were then diluted with 85% glycerol to 20 µM, and the methylated A adaptor, indexed TIME-Seq B adaptor, and 50% glycerol were combined in a ratio of 1:1:2. The resulting 10µM adaptors were combined in equal volume with purified Tn5 (1.5 mg/mL), mixed

thoroughly by pipetting 20 times, and incubated at room temperature for 50 minutes. Activated transposase was stored at -20ºC and no loss of activity has been observed up to 8 months.

To test the activity of TIME-Seq transposomes, 100 ng of human genomic DNA (11691112001, Roche) was tagmented in 25 µl reactions by adding 12.5 µl of 2X tagmentation buffer (20mM Tris-HCl pH 7.8, 10 mM dimethylformamide, 10 mM $MgCl_2$) using 1.5 µl of each barcoded transposome. After reactions were incubated at 55ºC for 15 minutes, STOP buffer (100mM MES pH5, 4.125M guanidine thiocyanate 25% isopropanol, 10mM EDTA) was added to denature Tn5 and release DNA. To assess tagmentation, 90% of each reaction was run in separate lanes of a 1% agarose gel at 90 volts (V) for 1 hour. The gel was stained with 1x SYBR gold (S11494, ThermoFisher) in Tris-Acetate-EDTA (TAE) buffer, and DNA fragment size was determined using a ChemiDoc (Bio-Rad) for gel imaging. The remaining DNA was pooled, cleaned up using a DNA Clean & Concentrator-5 (D4013, Zymo) kit, and the DNA was amplified with barcoded TIME-Seq PCR primers. Amplified pools were spiked into sequencing runs to assess relative barcode activity when new transposase was activated.

III. Biotin-RNA bait design and production

Mouse and human discovery baits were designed to enrich for previously described epigenetic clock CpGs[11-13,50] using mm10 and hg19 genomic coordinates. Using bedtools[51] (v2.28.0), regions around each target CpG were first expanded 125 base pairs (bps) up- and downstream (*bedtools slop*) and overlapping loci were merged (*bedtools merge*). Next, 110 bp bait windows were defined every 20 bps in the region (*bedtools window*) and the baits were intersected (*bedtools intersect*) with a file containing RepeatMasker (http://www.repeatmasker.org) annotated regions to identify baits overlapping repetitive DNA regions. Next, the FASTA sequence of each bait was gathered (*bedtools getfasta*), and blat (version 35) was used to get the copy number for each bait with options, -fastMap -maxIntron=50 -stepSize=5 -repMatch=2253 -minScore=40 -minIdentity=0. Using custom R (v. 4.0.2) scripts, information was gathered from the output files from each probe including the percent of each nucleotide, the overlap with repeats, and the bait copy number as determined by blat. Baits were automatically filtered that had overlap with

repeats, however, each locus that had none or very few (< 4) baits after filtering were inspected manually and, if the blat copy number was low (< 10), baits were added back, either to the exact locus or shifted slightly to avoid the annotated repeat. In preliminary biotin-RNA hybridization experiments (data not shown), we noticed that baits with high or low percentage T (less than 8% or greater than 30%) had low coverage, possibly due to stalling of the RNA polymerase while incorporating biotin-UTP. Therefore, when more than half the baits at a target locus had a percent T greater than 30% or lower than 8%, the reverse complement strand was captured instead for all baits at the locus.

To design enrichment baits for mouse and human rDNA, FASTA sequences were prepared from GeneBank accessions BK000964.3 (mouse) and U13369.1 (human) according to previously described methods[31] by moving the last 500 bps of each sequence (rDNA promoter) to be in front of the 5' external transcribed spacer (ETS). From the region comprising the promoter to the 3' ETS (mouse, 13,850bps; human, 13,814bps), 110-nt baits were designed using the *bedtools window* function to create baits every 20bps. Version 1 rDNA baits used in the pilot and targeting the original rDNA clock were designed to specifically enrich rDNA clock CpGs[31] using the same approach described for non-repetitive clock CpGs (i.e., 250bp windows were merged and 110-nt baits were designed to tile the regions).

The sequence of each bait set was appended with a promoter (Sp6 or T7) for *in vitro* transcription (IVT), as well as a promoter-specific reverse priming sequence that contained a BsrDI restriction enzyme motif. Bait sets containing Sp6 and T7 promoters were ordered together in a single-stranded DNA oligo pool (Twist), and pools were resuspended to 10 ng/µl. Bait sets were amplified in reactions containing 12.5 µl of 2X KAPA HiFi HotStart Polymerase Mix (7958927001, Roche), 0.75 µl of each 10 µM primer, and 0.5 µl of the bait pool using the following thermocycler program: initial denaturation, 95ºC for 3 min; 10 cycles amplification, 98ºC for 20 seconds, 61ºC (Sp6) or 58ºC (T7) for 15 seconds, 72ºC for 15 seconds; a final elongation for 1 minute at 72ºC. Amplified DNA was cleaned up using a Clean & Concentrator-5 kit (D4013, Zymo) and then digested with 1 µl BsrDI (R0574S, NEB) at 65ºC for 30 minutes. This reaction was again purified with a Clean & Concentrator-5 kit, and IVT reactions were set up according to the HiScribe™ T7 (E2040S, NEB) or Sp6 (E2070S, NEB) High Yield kits using half of the cleaned DNA

template for each reaction and storing the rest at -80ºC. All ribonucleotides were added to a final concentration of 5mM, including a 1:4 ratio of biotin-16-UTP (BU6105H, Lucigen) to UTP at 5 mM. After the reactions were incubated for 16 hours overnight at 37ºC, 25 µl of nuclease-free water and 2 µl of DnaseI (M0303S, NEB) were added to degrade the DNA template. RNA was purified using the 50 µg Monarch® RNA Cleanup Kit (T2040S, NEB). The concentration of RNA was measured using Qubit RNA BR Assay Kit (Q10210, ThermoFisher) and the size of RNA was measured using an RNA ScreenTape (5067-5576, Agilent) on an Agilent Tapestation.

While the yield from each DNA amplification and IVT reaction varies depending on the size and composition of the bait sets, we estimate that 1,000-10,000 hybridization reactions worth of bait could be produced from just 1 single-stranded oligo pool. For TIME-Seq libraries of 48-64 samples, this could enrich tens to hundreds of thousands of samples. Another advantage of this approach is that baits can be easily shared with other researchers as the ssDNA template, amplified DNA, or prepared biotin-RNA.

IV. Sequencing platforms and parameters

TIME-Seq libraries that were sequenced on an Illumina MiSeq using a 150 cycle MiSeq v3 kit (MS-102-3001, Illumina) or on a NextSeq 500 using a 150 cycle NextSeq High (20024907, Illumina) or Mid (20024904, Illumina) Output v2.5 kit used the following read protocol: read 1, 145-153 cycles; i7 index read, 8 cycles; i5 index read (if needed), 8 cycles; read 2, 5 cycles. This custom read protocol helps maximize sequencing efficiency and decrease cost since the fragment size of amplified TIME-Seq libraries was typically 80-200bps, and sequencing with larger kits results in a large portion of overlapping (unused) cycles. Optimization experiments, smaller deep-sequenced pools, and shallow sequencing data were typically generated by sequencing on a MiSeq using a MiSeq Reagent Micro v2 kit (MS-103-1002, Illumina) using standard paired-end and dual-indexed read cycle numbers. Larger experiments sequenced on a NovaSeq 6000 using a 200 cycles NovaSeq 6000 SP or S1 Reagent Kit v1.5. High GC genome from *Deinococcus radiodurans* was spiked into sequencing runs in different proportions (1-3% on MiSeq, 5-10% on NovaSeq, and 15-20% on NextSeq) to increase base diversity and improve sequencing quality[49].