Flap-enabled next-generation capture (FENGC): precision targeted single-molecule profiling of epigenetic heterogeneity, chromatin dynamics, and genetic variation

Mingqi Zhou[1,2,3], Nancy H. Nabilsi[1,3], Anqi Wang[1,3], Marie-Pierre L. Gauthier[1,3], Kevin O. Murray[4], Hassan Azari[5], William S. Owens[1,3], Jeremy R. B. Newman[2,6], Francisco J. Pardo-Palacios[7], Ana Conesa[7,8], Alberto Riva[9], Thomas L. Clanton[4], Brent A. Reynolds[5], Patrick Concannon[2,6], Jason O. Brant[3,10], Rhonda Bacher[2,3,10] and Michael P. Kladde[1,2,3,*]

[1]Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, FL, USA.

[2]Genetics Institute, University of Florida, Gainesville, FL, USA.

[3]UF Health Cancer Center, University of Florida, Gainesville, FL, USA.

[4]Department of Applied Physiology and Kinesiology, College of Health and Human Performance, University of Florida, Gainesville, FL, USA.

[5]Department of Neurosurgery, University of Florida, Gainesville, FL, USA.

[6]Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, USA.

[7]Institute for Integrative Systems Biology, Spanish National Research Council, Spain.

[8]Department of Microbiology and Cell Science, University of Florida, Gainesville, USA.

[9]Bioinformatics Core, Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL, USA.

[10]Department of Biostatistics, College of Public Health & Health Professions, University of Florida, Gainesville, FL, USA.

*To whom correspondence should be addressed. (Tel: +1-352-273-8142;
Email: kladde@ufl.edu)

Keywords: Targeted sequence enrichment, flap endonuclease, epigenetics, DNA methyltransferase, DNA methylation, nucleosome dynamics, transcription factor binding

Abstract

Targeted sequencing is an increasingly sought technology. Available methods, however, are often costly and yield high proportions of off-target reads. Here, we present FENGC, a scalable, multiplexed method in which target sequences are assembled into 5′ flaps for precise excision with flap endonuclease. Recovery of length-matched sequences, amplification with universal primers, and exonucleolytic removal of non-targeted genomic regions mitigate amplification biases and consistently yield ≥80% on-target sequencing. Furthermore, optimized sequential reagent addition and purifications minimize sample loss and facilitate rapid processing of sub-microgram quantities of DNA for detection of genetic variants and DNA methylation. Treatment of cultured human glioblastoma cells and primary murine monocytes with GC methyltransferase followed by FENGC and high-coverage enzymatic methyl sequencing provides single-molecule, long-read detection of differential endogenous CG methylation, dynamic nucleosome repositioning, and transcription factor binding. FENGC provides a versatile and cost-effective platform for targeted sequence enrichment for analysis of genetic and/or epigenetic heterogeneity.

Multiplexed enrichment of target sequences for next-generation sequencing (NGS) has diverse applications in basic science and medicine[1]. The reduced sample complexity imparts substantial NGS cost savings, as well as higher per base coverages that improve data quality and detection of rare variants. Available methods, however, generally interrogate either genetic or epigenetic signatures, require expensive reagents, have limited flexibility regarding target design, and exhibit high off-target rates. Furthermore, sequence recovery is usually poor due to multiple purifications, necessitating high amounts of input DNA. In addition, most enrichment techniques capture relatively short fragments with wide length distributions from which the shorter fragments are preferentially amplified.

Hybridization-based target enrichment using DNA or RNA oligonucleotide 'baits' is widely used for targeted applications such as DNA methylation analysis[2]. Although tiling oligonucleotides improves coverage uniformity, pools of overlapping, long (50-170 nt), and biotinylated oligonucleotides[3] are cost prohibitive to many laboratories. Hybridization enrichment also suffers from long incubation times, high off-target rates, and inefficient recovery of target sequences[4]. Random fragmentation of input DNA further contributes to off-target sequencing, as well as amplification and sequencing biases. Lastly, hybridization-based capture that yields short reads precludes phasing of epigenetic and genetic information.

Clustered, regularly interspaced short palindromic repeats (CRISPR)-associated (Cas) enzymes, which utilize short RNAs to direct site-specific DNA cleavage[5], are increasingly enlisted for sequence enrichment *in vitro*. However, for multiplexed enrichment of shorter sequences with CRISPR-Cas9 (refs. 6-9), the requirement for a 2-bp protospacer adjacent motif constrains target design and thus length matching, resulting in amplification and sequencing biases. This drawback is in part circumvented by direct sequencing of captured megabase-length fragments with Nanopore technology[10]; however, the requirement for large amounts of sample is problematic for clinical applications. In addition, accurate single-molecule detection of 5-methylcytosine (5mC) as opposed to deriving a population average remains technically challenging. Lastly, off-target cutting[11,12] and the large expense of recombinant Cas enzymes and guide RNAs constitute further barriers to scaling for many laboratories.

Bisulfite patch PCR[13], optionally combined with the methyltransferase accessibility protocol for individual templates (MAPit), termed MAPit-patch[14], are additional methods for targeted genetic and epigenetic surveys. These methods, however, rely on restriction endonucleases, which severely limit target selection design and generate fragments with broad length distributions that introduce substantial amplification and sequencing biases.

Here, we have developed and applied flap-enabled next-generation capture (FENGC), a versatile, rapid, cost-effective, and scalable method for multiplexed sequence enrichment. FENGC enlists flap endonuclease (FEN) to excise target sequences reconstituted into 5′ DNA flaps to achieve high ratios of on- to off-target enrichment. Programmable, nucleotide-specific cleavage impart unparalleled versatility in target design and capture, enabling matching sequence lengths to mitigate PCR bias and further improve on-target sequencing frequency. FENGC of genomic DNA (gDNA) from cultured and primary cells treated with GpC methyltransferase followed by enzymatic methyl sequencing (EM-seq) and high-fidelity, long-read sequencing reproducibly localizes nucleosomes, transcription factor (TF) binding, and endogenous CG methylation on contiguous single molecules. Furthermore, achieving high sequencing coverage facilitates detection of low-level, but significant differences, in chromatin accessibility and DNA methylation, as well as dynamic nucleosome repositioning as a driver of nucleosome-free region (NFR) length. FENGC without deamination provides high-confidence identification of single-nucleotide polymorphisms (SNPs) and insertions/deletions (indels).

Results

FENGC, a novel, cost-effective method for on-target sequence enrichment in genetic and epigenetic studies. To establish FENGC, we retained advantages of multiplexed DNA strand capture and amplification afforded by patch ligation PCR[13,14], but eliminated the requirement for restriction enzymes. To do so, we leveraged the activities of thermostable *Thermococcus* 9° N flap endonuclease 1 (FEN1) and the FEN domain of Taq DNA polymerase I, which tightly bind and efficiently cut double flaps (unpaired 5' flap and 1-nt 3' flap) with exquisite nucleotide-level precision and uniformity[15-19].

In a multiplexed reaction, FENGC enlists a pair of target-specific flap adapters to assemble both ends of each target sequence into a double flap (Fig. 1a; Supplementary Fig. 1). Following specific cleavage by FEN1 or Taq, target sequences are patch ligated to universal primers, one of which protects against subsequent exonuclease digestion. The enriched sequences are purified and amplified by standard PCR for genotyping or, alternatively, subjected to bisulfite- or enzyme-mediated deamination of C to U prior to amplification (methyl-PCR) for base-level detection of 5mC. Shorter library products can be barcoded and sequenced as described previously[14]. Herein, longer PCR and methyl-PCR products were examined by long-read, high-fidelity (Hi-Fi; >Q30) circular consensus sequencing (CCS; Pacific Biosciences, PacBio).

FENGC optimization and validation. We first determined that *Thermococcus* 9° N FEN1 cut an oligo proxy target assembled into a 5' flap more efficiently than Taq (Supplementary Fig. 2; Supplementary Table 1, Sheet 1). Next, we tested if FEN1 activity was affected by DNA methylation and/or changing the base of the 'overlapped' nucleotide (Supplementary Fig. 3). In this experiment, a pair of unmethylated and methylated 80-mers was reconstituted into four double flaps with A, C, G, or T both 5' of the scissile phosphodiester and as a 1-nt 3' flap (Supplementary Fig. 3a; Supplementary Table 1, Sheet 1). Five C or 5mC residues in the two 80-mers reside near the 5' flap cleavage sites, locations contacted by human FEN1 co-crystallized with a double flap (5' flap and 1-nt 3' flap)[19]. Importantly, the use of unmethylated flap adapters parallels FENGC enrichment of methylated targets. Equally efficient cleavage of all four double flaps ± nearby 5mC (Supplementary Fig. 3b,c) suggests FENGC is compatible with DNA methylation analyses, including MAPit.

We first examined the ability of FENGC to enrich single-copy genomic sequences from gDNA of HCT116 colon cancer cells, as well as the method's suitability for bisulfite sequencing. Flap adapters were designed to excise 11 promoters, including 9 from DNA repair genes, *EMP2AIP1* transcribed divergently from *MLH1*, and *HSPA5* as an open chromatin control (Supplementary Table 1, Sheets 2 and 3). The sequences had mean lengths of 298 nt (range 282-315 nt) and 445 nt (range 430-452 nt), anchoring both sets at the 5' end (Supplementary Table 2, Sheets 1 and 2). The ~300-nt and ~450-nt targets were most likely cut and ligated to oligos U1-T and U2 (totaling 47 nt), given amplification of the expected ~350 bp and ~500 bp products, respectively, dependent on Taq and FEN1 (Supplementary Fig. 4a, compare lane 1 with 2 and 3 with 4; Supplementary Fig. 4b, compare lane 1 with 2 and 4 with 5). Superior amplification yield was observed with FEN1 vs. Taq (Supplementary Fig. 4a, compare lane 2 with 4). As expected, FENGC product yields were reduced with longer amplicons and bisulfite (BS) treatment, which hydrolyzes DNA (Supplementary Fig. 4b, compare lane 2 with 5 and lane 3 with 6). Therefore, detection of the ~500 bp BS-PCR product required 4 μg gDNA input for this FENGC reaction that targeted only 11 loci (compare lane 6 with 7).

MAPit-FENGC, a versatile assay for targeted genetic and epigenetic analyses. Given the successful post-capture bisulfite conversion and amplification of target sequences, we combined FENGC with MAPit methylation footprinting, also known as NOMe-seq[20,21]. The

resulting MAPit-FENGC method was tested on biological duplicates of human cell line L0 derived from a GBM tumor of the mesenchymal subtype[22]. Harvested L0 cells were permeabilized and incubated with methyltransferase M.CviPI and methyl donor *S*-adenosyl-*L*-methionine in order to mark accessible GC sites (Methods). FENGC with Taq or FEN1 was used to target the same 11 promoters for sequences of ~300 nt and ~450 nt, followed by standard PCR or BS-PCR, and sequencing using PacBio Hi-Fi long-read CCS (Methods).

In this MAPit-FENGC pilot, ≥1 CCS read aligned to at least 9 of the 11 target promoters (82%) in all 12 reactions (Supplementary Table 3). Zero to few reads were detected for *MSH3* and *MSH6*, possibly due to their high GC content, typical of NGS and patch PCR methods[13,14]. For the remaining 9 targets, FENGC with both Taq and FEN1 using standard PCR achieved ≥25x coverage in the combined duplicates for each condition. For FENGC with FEN1 and BS-PCR, ≥150x coverage was observed in the combined duplicates for the same 6 targets in both the ~300-nt and ~450-nt sequences. For these 6 targets, methylation levels at HCG (H5mCG) and GCH (G5mCH), i.e., excluding GCG overlap, were highly correlated between both biological replicates and between the shared 5' ends of the enriched ~300-nt and ~450-nt sequences (Supplementary Fig. 5; $r \geq 0.92$).

To increase the sensitivity of targeted DNA methylation analysis by FENGC, bisulfite was substituted with EM-seq, a nondestructive, enzymatic means to convert C to U[23,24]. For this analysis, the number of target promoters was scaled from 11 to 119, adding the promoters of 74 genes with Gene Ontology term “DNA repair,” 41 cancer-associated genes, and 3 control genes (Supplementary Table 1, Sheet 3; Supplementary Table 2, Sheet 2). For *FAT4*, oligo 2-T was designed to not cut the 3' end, therefore providing a negative control. This 119-target panel (range 430-452 nt) was used for MAPit-FENGC of different amounts of gDNA from M.CviPI-treated L0 cells. Subsequent EM-seq and amplification (EM-PCR) produced the specific, expected ~500 bp product, including 47 bp universal primers, from as little as 50 ng gDNA (Supplementary Fig. 6). By contrast, bisulfite conversion required 80x more gDNA (4 μg) to detect an amplicon (Supplementary Fig. 4b).

Next, MAPit-FENGC was conducted on two independent cultures of GBM L0, GBM line Nx18-25 (classical subtype), and non-cancerous human fetal neural stem cells (NSC)[22]. GDNA from each of these M.CviPI-treated lines and two 0.1% L0 gDNA:99.9% NSC gDNA mixtures were processed in parallel by FENGC using the 119-promoter panel followed by EM-seq conversion. The ranges of PacBio Hi-Fi CCS reads that aligned to the target references of the hg38 reference genome, off-target but on hg38, and unmapped from the 8 libraries were 77-86%, 6.6-9.9% (excluding one L0:NSC mixture outlier at 14%), and 6.9-13% (Fig. 1b; Supplementary Table 4, Sheet 1). These on-target values compare favorably with single-molecule footprinting enrichment using tiled RNA baits[25]. As expected, zero reads aligned to negative control locus *FAT4* (Supplementary Table 5). MAPit-FENGC detected 105-111 targets (89-94%, excluding *FAT4*) with ≥1 read in the combined biological duplicates for each sample (Supplementary Table 4, Sheet 1). To avoid multiple alignments to gene homologs, reads were further filtered for ≥95% coverage of each reference sequence length. These filtered reads aligned uniquely to 92-105 targets (78-89%; Supplementary Table 4, Sheet 1; Supplementary Table 5), and read number *versus* GC content showed a modest negative fit (Supplementary Fig. 7).

For the fractions of H5mCG (DNA methylation) and G5mCH (chromatin accessibility), all correlation coefficients between the independent biological duplicates of each sequenced library were ≥0.91 (Supplementary Fig. 8). In addition, the proportions of H5mCG and G5mCH were highly correlated between the EM- and BS-PCR data sets, indicating that EM-seq can be substituted for BS-seq in MAPit-FENGC (Supplementary Fig. 9; $r > 0.93$).

MAPit-FENGC with EM-seq informs epigenetic landscapes and efficiently detects differential DNA methylation and accessibility in minor subpopulations of cells. Among the 119 loci profiled by MAPit-FENGC, the *EPM2AIP1* promoter exemplifies marked epigenetic reprogramming in GBM (Fig. 1c-d). To reveal methylation patterns, molecules were hierarchically clustered based on the first principal component of H5mCG and G5mCH levels by methylscaper[26]. In non-cancerous NSC, the promoter was predominantly unmethylated and highly accessible to M.CviPI, except for a +1 nucleosome (blue ellipses; expected ~150- bp protections), which occupied a continuum of positions in different cells across the sample population, thereby lengthening or shortening NFRs (Fig. 1c, right panel). These accessible, variable-length NFRs were also punctuated by a short footprint of fixed location (black rectangle), indicative of strong sequence-specific TF binding. By contrast to NSC, in GBM L0, the *EPM2AIP1* promoter was hypermethylated at HCGs (Fig. 1c,d, compare left panels) and closed, except for relatively short, accessible linkers within randomly positioned nucleosomes (Fig. 1c,d, compare right panels).

To challenge the sensitivity of MAPit-FENGC, duplicate mixtures of 0.1% L0 gDNA:99.9%NSC gDNA were processed. A significant proportion of hypermethylated *EPM2AIP1* molecules was detected, likely originating from the 0.1% L0 spike-in (Fig. 1e, left; red arrowhead; $P = 3.59E{-}09$). The chromatin state of the other molecules resembled those of NSC (Fig. 1d, right). Together, these data show that FENGC maps epigenetic signatures reproducibly, quantitatively, and with high sensitivity, detecting as few as 0.1% hypermethylated epialleles.

The ability of MAPit-FENGC to detect differential epigenetic signatures was tested further by assaying the same 119-target panel from two independent cultures of non-cancerous NSC and GBM Nx18-25. Epigenetic differences between NSC and GBM were assessed for 54 targets with ≥50 total CCS reads in both replicates, but no less than 22 reads in either replicate (Supplementary Table 6). Using a significance criterion of $P < .05$ as well as differential levels of either H5mCG or G5mCH of ≥0.05, ≥0.10, ≥0.20, and ≥0.45, 57% (31/54), 24% (13/54), 13% (7/54), and 7.4% (4/54) of promoters exhibited candidate epigenetic differences, respectively (Supplementary Table 6). The quantitative nature of these results is supported by the highly correlated ($r > 0.91$) H5mCG and G5mCH levels between replicates (Supplementary Fig. 8). In addition, 733 of 735 of *MSH5* promoter copies across all 4 samples had ≥10 G5mCH residues (Extended Data Fig. 1). Such high accessibility rules out incomplete cell permeabilization and limiting M.CviPI activity as possible bases for differential accessibilities either between different loci or molecules from a given locus in the same or different samples.

To illustrate quantitative detection of epigenetic differences between NSC and GBM Nx18-25 by MAPit-FENGC, three representative promoters were chosen and correlated with transcript levels (Fig. 2; Extended Data Fig. 2). *CD44* encodes a transmembrane glycoprotein with functions in cell adhesion, proliferation, and apoptosis[27,28]. *CD44* expression is also a marker of astrocyte-restricted precursors[29]. In particular, *CD44* overexpression in GBM tissue has been linked to the mesenchymal subtype[30,31] and GBM cancer stem cells[32,33]. MAPit-FENGC of NSC showed methylation of two to three HCGs upstream of the *CD44* TSS (Fig. 2a, panel 1). Notably, the major NFR on most of the NSC molecules occupied diverse positions, which coincided with movement of the −1 nucleosome across the entire assayed region (Fig. 2a, panel 2). In striking contrast to NSC, in GBM Nx18-25, reproducible, pronounced, and variable-length expansion of H5mCG was observed (Fig. 2a, panel 1 *vs*. 3; Fig. 2b, left, $P < .0001$). The increased H5mCG in Nx18-25 (compared with NSC) correlated with marked, significant reductions in accessibility (Fig. 2a, panel 2 *vs*. 4; Fig. 2b, right, $P < .0001$), cumulative NFR length (Fig. 2c, $P = 2.2E\text{-}16$), and transcript abundance (Fig. 2d, $P < .01$).

CCN4, also known as WNT1-inducible signaling pathway protein 1 (WISP1), contributes to the tumorigenesis and progression of a wide array of human cancers[34-36]. *CCN4* expression is upregulated in GBM compared with normal tissues[37]. Indeed, *CCN4* transcripts were undetectable in NSC and increased dramatically in GBM Nx18-25 (Fig. 2e; $P < .05$). MAPit-FENGC of the *CCN4* promoter from NSC revealed a majority of molecules with variable levels of H5mCG and disorganized nucleosomes (Fig. 2f, panels 1 and 2, respectively; clusters 2 and 3). Molecules in cluster 1, by contrast, exhibited larger spans of accessibility and contained two footprints (black rectangles); however, nucleosomes (cyan ovals) occluded the TSS. In terms of peaks and valleys, plots of averaged H5mCG and accessibility were qualitatively similar between NSC and GBM Nx18-25 (Fig. 2g). Quantitatively, however, in GBM, the number of molecules bearing 5mCG declined significantly (Fig. 2f, panel 1 *vs*. 3; Fig. 2g, upper; Fig. 2h, left, $P < .0001$), and NFR length per molecule increased, particularly in newly appearing cluster 4 (Fig. 2f, panel 2 *vs*. 4; Fig. 2i, $P = 2.2E-16$). Overall, the quantitative changes in *CCN4* promoter DNA methylation and accessibility are consistent with the observed marked increase in transcript abundance in GBM (Fig. 2e).

HIST1H1B encodes the linker histone protein H1.5, which maintains higher-order chromatin structure, as well as regulates DNA repair and cell proliferation[38,39]. Based on NFR size, nucleosome +1 position, TF binding upstream of the TSS, and DNA hypermethylation, *HIST1H1B* promoter copies from both NSC and GBM Nx18-25 populated seven clusters (Extended Data Fig. 2). Hypermethylated *HIST1H1B* promoter copies increased significantly from 0.6% to 4% in NSC *versus* GBM (Extended Data Fig. 2a, panel 1 *vs*. 3; Extended Data Fig. 2b; Extended Data Fig. 2c, left; Supplementary Table 6, $P < .0001$). In addition, cumulative accessibility, particularly from −80 to +45, decreased significantly in Nx18-25 compared with NSC (Extended Data Fig. 2a, panel 2 *vs*. 4; Extended Data Fig. 2b, bracketed area; Extended Data Fig. 2c, right; Supplementary Table 6, $P < .0001$). However, the modest reduction in accessibility and increase to 4% hypermethylated *HIST1H1B* epialleles in Nx18-25 did not significantly alter bulk transcript abundance (Extended Data Fig. 2d). Taken together, we conclude that MAPit-FENGC efficiently detects prominent as well as low-frequency epigenetic differences, and the profiled epigenetic heterogeneity informs mechanism.

FENGC detects known and novel genetic alterations. For genotyping, gDNA from the same two independent cultures of M.CviPI-treated NSC and GBM Nx18-25 were processed by FENGC with the 119-promoter panel of ~450 nt (Supplementary Table 2, Sheet 2), without deamination prior to PCR. Hi-Fi CCS reads from the 4 sequenced libraries aligned to the 119 hg38 target references, off-target but on hg38, and off hg38 with ranges of 73-89%, 10-20%, and 0.6-8.9%, respectively (Supplementary Table 4, Sheet 2). Among the 97 targets (82%) with ≥1 read in at least one condition, we identified 51 single-nucleotide polymorphisms (SNPs) and 17 indels distributed over 30 promoters (Supplementary Table 7). There were 43 SNPs and 16 indels already recorded in dbSNP (http://www.ncbi.nlm.nih.gov/SNP/). The C-to-A substitution in the *CDH1* promoter at chr16:68737131 (rs16260) was previously identified as a cancer-specific risk factor with clinical significance[40-42]. The SNP A allele was identified within 0% of reads in NSC and 38% of reads in Nx18-25. Eight SNPs and one indel without dbSNP entries were also detected. Reproducible identification of alleles of both known and previously unknown polymorphisms demonstrates the efficacy of FENGC in genotyping.

Bona fide epigenetic perturbations, by definition, require no concurrent DNA mutations. Therefore, we further examined if candidate epigenetic differences in targets (Supplementary Table 6) could be attributable at least in part to genetic variation between NSC and GBM Nx18-25 (Supplementary Table 7). A GBM-specific SNP mapping to the *CHEK1* promoter eliminated a CG site, fully accounting for the loss of H5mCG in Nx18-25 *versus* NSC. Allelic differences between NSC and GBM in nine other promoters could in part underpin the identified

epigenetic variations. Importantly, comparing NSC and GBM, a known indel was represented at similar frequencies in *CCN4* and no allelic variants were detected in *CD44* or *HIST1H1B*, satisfying the criterion for the three showcased loci of no differential underlying mutations contributing to the observed epigenetic alterations.

MAPit-FENGC with ~1-kb reads delineates regulatory relationships between multiple epigenetic modules. Longer reads allow richer examination of epigenetic landscapes, e.g., functional crosstalk between individual regulatory modules at a distance. Therefore, we redesigned 45 of the ~450-nt promoter targets to capture more upstream sequence (range, 930-959 nt; mean, 940 nt; Supplementary Table 1, Sheet 4; and Supplementary Table 2, Sheet 3). These ~940-nt targets were enriched by FENGC of 400 ng and 800 ng from the same two stocks of gDNA from M.CviPI-treated GBM Nx18-25 used to assay the ~450-nt targets. Hi-Fi CCS reads were predominantly the expected ~990 nt (data not shown), of which 77-81% were on target and 38 (84%) targets had ≥1 read in each of the 4 libraries (Supplementary Fig. 10a; Supplementary Table 8). Despite the closely matched target lengths, coverage was variable and only 31% (14/45) of the targets retrieved ≥26 reads in the combined 800-ng samples (Supplementary Table 9). Nevertheless, H5mCG and G5mCH levels of the 14 targets were highly correlated ($r \geq 0.97$) between the two biological replicates for the same gDNA amount (Supplementary Fig. 10b). Correlation coefficients were also excellent between different gDNA amounts (Supplementary Fig. 11a; $r > 0.95$), as well as between the overlapping segments of the ~940-nt and ~450-nt targets (Supplementary Fig. 11b; $r > 0.92$).

To further demonstrate concordance of chromatin architectures between the different lengths of molecules and, moreover, to derive additional regulatory insights, we aligned single-molecule heatmaps of the three *EPM2AIP1-MLH1* amplicons (Fig. 3a-c). The 450-bp *EPM2AIP1* and 438-bp *MLH1* molecules were mostly unmethylated (Fig. 3a) and open (Fig. 3b) in GBM Nx18-25, whereas in GBM L0 they were strongly hypermethylated and closed (Fig. 1d). Moreover, as in NSC (Fig. 1c), a continuum of positions for both +1 nucleosomes and thus NFR lengthening was observed in Nx18-25, but was more constrained for the *MLH1* +1 nucleosome (Fig. 3b, right). In addition, both NFRs harbored a TSS and strong sequence-specific binding of TFs 1 or 2. The 937-bp molecules revealed continuous, variable-length NFRs that encompass *EPM2AIP1* TSS, *MLH1* TSSa, TF footprint 1, and additionally detected TF footprint 3 (Fig. 3c, right). Co-occupancy of the three TFs and two +1 nucleosomes was evident on the long *EPM2AIP1-MLH1* molecules. However, the continuum of *MLH1* TSSa +1 nucleosome repositioning was obscured when clustering of the 937-bp molecules was weighted on the *EPM2AIP1* +1 nucleosome (Fig. 3c, right) and *vice versa* (data not shown). This suggests that the two nucleosomes are mobilized in large part independently of each other, an assertion further supported by weak correlations in accessibility between GCH sites in the two +1 nucleosomes/NFRs (Fig. 3d, bounded by yellow-dashed rectangles). In contrast, considering each +1 nucleosome/NFR separately, higher negative and positive correlations were observed, consistent with nucleosome sliding (Fig. 3d, white-dashed squares). Therefore, the 937-bp *EPM2AIP1-MLH1* molecules netted an additional TF footprint and novel regulatory insight—dynamic, but independent, mobilization of two +1 nucleosomes.

MAPit-FENGC of other ~940-nt targets from GBM Nx18-25 phased additional epigenetic features compared to their shorter counterparts. The divergently transcribed *NPAT* and *ATM* promoters showed a robust TF footprint, a weaker footprint (enclosed by cyan rectangle) at the *NPAT* TSS, and an array of nucleosomes +1 to +3, each progressively less well positioned with increasing distance from the TSS (Extended Data Fig. 3a). The *POLD4* promoter displayed a similar nucleosomal array and two footprints at sequences with strong homology to Sp1 binding sites (Extended Data Fig. 3b, right). Surprisingly, H5mCG localized adjacent to these TF footprints in a sizeable number of molecules (left panel). By contrast to *NPAT-ATM* and *POLD4*

promoters, the *MSH2* 5′ flanking region nucleosomes are more disorganized (Extended Data Fig. 3c), and the NFR of a substantial proportion of molecules contains ~55-bp footprints at the TSS (cyan rectangle), consistent with occupancy by the RNA polymerase II preinitiation complex[43]. Compared to the shorter *CCN4* promoter molecules from Nx18-25 (Fig. 2f), a subset of longer *CCN4* molecules showed NFRs as long as ~400 bp (Extended Data Fig. 3d); however, intriguingly, no NFRs extended beyond a 66-bp A-rich element. Assaying ~800 bp of 5′ flanking sequence from three of the four genes in Extended Data Figure 3 identified clear transitions between areas of 5mCG depletion and hypermethylation.

MAPit-FENGC reproducibly informs epigenetic architectures within primary cells. Having successfully applied FENGC to detect epigenetic and genetic variation from cell lines, we tested the epigenetic arm of the protocol on primary, bone marrow-derived monocytes. Cells isolated from four female mice were separately permeabilized and treated with M.CviPI, and 600 ng gDNA from each sample was processed with flap adapters targeting 78 genes with roles in the cellular inflammatory response (Supplementary Table 1, Sheet 5). For this experiment, primers were designed by FENGC oligonucleotide designer (FOLD; Methods). A computational solution for all 78 targets that avoided repeats and satisfied other command line settings yielded targets ranging from 474-987 nt (mean, 620 nt; Supplementary Table 2, Sheet 4).

Hi-Fi CCS reads from the 4 sequenced libraries were mapped to mm9 build of the mouse genome and the subset of target references: 20-29% did not align to either reference or were removed by filtering, yielding 71-80% on-target reads (Supplementary Table 10). MAPit-FENGC detected 71-75 targets (91-96%) with ≥1 read from each library (Supplementary Tables 10,11).

A mixed effects ANOVA was used to evaluate epigenetic differences between the four monocyte samples (Supplementary Table 12). To provide statistical rigor, testing was restricted to 43 regions with ≥100 filtered CCS reads per sample, of which all were ≤760 nt and ≤70% GC content (Supplementary Fig. 12). For all 43 targets, no single monocyte sample showed significant differences in DNA methylation or accessibility from the other three samples.

The highly consistent chromatin structures recorded by MAPit-FENGC between individual monocyte samples is apparent in plots of both averaged as well as single-molecule H5mCG and G5mCH for eight representative loci, which also illustrate interesting chromatin biology (Fig. 4; Extended Data Fig. 4). The *Hsf1* promoter, transcribed divergently from *Bop1*, is exceptionally open, with 2,333 of 2,334 molecules displaying relatively large NFRs (range 223-445 bp; Extended Data Fig. 4a). As indicated above, such a locus demonstrates efficient probing of chromatin accessibility with M.CviPI. This lends credence to the observed bimodal distribution of *Btk* molecules (Fig. 4a); epialleles bearing 0% H5mCG and TSS-proximal NFRs with a prominently bound TF, consistent with sustaining transcription *versus* those with variable amounts of H5mCG and reduced, disorganized accessibility reminiscent of X inactivation. Notably, H5mCG is preferentially depleted at the TSS of methylated *Btk* copies that presumably reside on the inactive X chromosome.

Among other intriguing loci, the *Pik3r3* promoter harbors a well-positioned –1 nucleosome and progressive repositioning of the +1 nucleosome, widening or narrowing NFRs in individual cells preferentially on the downstream side of the NFR (Fig. 4b). By comparison, incremental movement of both the –1 and +1 nucleosomes produces a ‘tornado-like’ distribution of NFR lengths at the TSSs of *Tlr4* (Extended Data Fig. 4b), *Hsp90ab1* (Extended Data Fig. 4c), *Irf7* (data not shown), and *Cxcr4* (Fig. 4c). Interestingly, the latter promoter displays a 36-bp zone of H5mCG (orange rectangle) in the NFR, located only 12 bp from a strongly footprinted TF.

Compared to the above promoters, *Src* chromatin (–6,420 to –5,805 of RefSeq TSS) is heavily methylated and populated by arrays of randomly positioned nucleosomes with short,

accessible linkers (Extended Data Fig. 4d). Interestingly, in the majority of cells, a sequence with strong homology to the consensus CTCF binding site[44] (map position 462) exhibits a 50-bp area of protection of variable numbers of HCG against methylation. The *Mapk15* gene body (+1,660 to +2,215) is also organized into heavily methylated arrays of randomly positioned nucleosomes, with noticeable depletion of 5mCG in exon 5 (Fig. 4d).

We independently validated MAPit-FENGC of monocytes with single-amplicon MAPit, which reported identical chromatin architectures of the *Btk*, *Cxcr4*, *Hsp90ab1*, and *Tlr4* regions (data not shown). Taken together, our data demonstrate that MAPit-FENGC effectively profiles chromatin architectures of purified primary cells, with striking inter-sample reproducibility.

Discussion

Targeted sequencing offers cost savings, faster turnaround times, and higher sequencing coverages that improve data quality and enable more comprehensive descriptions of genetic and epigenetic heterogeneity than is feasible with whole genome analyses. However, currently available technologies continue to suffer from limitations regarding target sequence design, high off-target sequencing, and/or prohibitive costs. Here, we describe FENGC, a multiplexed, cost-effective, and high-throughput method, wherein FEN activity precisely excises target sequences reconstituted into double flaps (Fig. 1a). Precision targeted cleavage at any desired nucleotide offers outstanding latitude of target sequence design, mitigates off-target sequencing, and enables length-matching that minimizes amplification and sequencing biases.

Several additional features of FENGC make it particularly attractive for targeted sequencing. First, the reaction is scalable and low cost because it employs standard, salt-free synthesis of target-specific oligos, without modifications such as 5mC or biotinylation. Second, high on-target specificities are achieved by several means (Fig. 1a): i) a pair of sequence-specific flap adapters direct double flap formation, target strand cleavage, and 5'-end ligation; ii) a nested, sequence-specific adapter directs 3'-end ligation; iii) neither flap cutting nor ligation tolerate DNA mismatches[15,17,45-47]; and iv) sequence complexity is considerably reduced by exonuclease digestion prior to amplification, as in nested patch, bisulfite patch, and MAPit-patch PCRs[13,14,48]. Third, for every FENGC target, oligo U1-N within flap adapter 2 ligates to the 5' end of each downstream sequence. Therefore, a series of flap adapters annealing successively farther downstream and/or upstream enable chromosome ‘walking.’ Fourth, sample loss is minimized through sequential reagent addition, poising FENGC for robotic automation. Therefore, superior sensitivity and on-target frequencies are reproducibly attained at a fraction of the cost of methods employing CRISPR/Cas or pull downs with biotinylated baits.

FENGC-enriched targets can be amplified for genotyping, after C to U conversion for methylation analysis, or both. Conducting genotyping and epigenetic assays in parallel detects mutations that are potential confounders of *bona fide* epigenetic alterations. To this end, sequencing of 119 FENGC-enriched, but non-deaminated targets, from NSC and GBM Nx18-25 treated with M.CviPI detected 43 SNPs and 16 indels already recorded in dbSNP (Supplementary Table 7). This included rs16260, a known *CDH1* SNP and cancer-specific risk factor that impairs transcription[40-42], as well as nine novel SNPs and one indel.

Non-destructive enzymatic conversion[23,49] of the same promoter sequences from M.CviPI-treated chromatin, termed MAPit-FENGC, provides integrated, single-molecule detection of DNA methylation and chromatin accessibility, with superior sensitivity to bisulfite conversion (Supplementary Figs. 4b,6). Among promoters meeting stringent CCS read filtering criteria in both bioreplicates of both NSC and classical GBM Nx18-25, 24% (13/54) of promoters of genes encoding DNA repair or cancer-associated products exhibited ≥0.1 differential in H5mCG, G5mCH, or both (Supplementary Table 6). Of these 13 promoters, 3 (*GSTA1*, *HOTAIR*, and *HOXB2*) of 4 with the highest epigenetic differences also possessed 1-3 genetic variants with

higher abundances in either non-cancerous NSC or classical GBM Nx18-25. These mutations could contribute to the observed epigenetic reconfigurations, emphasizing the importance of conducting genetic analysis parallel to epigenetic investigations of mutation-prone samples such as cancerous cells.

MAPit-FENGC provides a rich contextual and mechanistic view of epigenetic landscapes by integrating the spatial relationships between endogenous CG methylation, nucleosome positioning/occupancy, and TF binding along contiguous single chromatin fibers. The high coverages provide statistical power for detection of epigenetic differences with a sensitivity as low as 0.1% hypermethylated epialleles (Fig. 1e). Furthermore, hierarchical organization of 100s-1000s of molecules, each reporting a static snapshot from a single cell, captures the range of epigenetic intermediates and therefore DNA methylation and nucleosome dynamics akin to a flipbook animation. For example, within hypermethylated chromatin, MAPit-FENGC revealed regularly spaced arrays of nucleosomes, which were mobile between different cells, i.e., not positioned relative to DNA sequence (Fig. 5d; Extended Data Fig. 4d). Similar observations were made using single-cell micrococcal nuclease sequencing (scMNase-seq)[50], albeit MAPit-FENGC expends far less financial and computational resources.

At transcribed mammalian promoters, MAPit-FENGC efficiently captured heterogeneously sized NFRs, as did scMNase-seq using meta-data that average across many loci and cells[50]. However, the superior coverages of MAPit-FENGC also detected incremental, *promoter-specific* patterns of nucleosome repositioning and hence changes in NFR length, as we previously reported with MAPit of a synthetic yeast locus containing various TF binding sites[51]. Such promoter-specific architectures reflect a net outcome of the competitive interplay between nucleosomes and chromatin-associated proteins, including TFs, recruited chromatin remodelers, etc.[51,52]. A complex regulatory interplay seems likely, for example, at the divergently transcribed *EPM2AIP1-MLH1* locus, where multiple TSSs, TFs, and mobile nucleosomes reside (Fig. 3). Indeed, long-read MAPit-FENGC divulged that the two *EPM2AIP1-MLH1* +1 nucleosomes slide independently and to different extents, challenging the prevailing model that that two genes are coordinately controlled by a shared promoter.

The increased mechanistic granularity of MAPit-FENGC cannot be gleaned from short reads and/or methods that cut accessible DNA and therefore separate nucleosomes and regulatory elements, including DNase-, MNase-, and ATAC-seq. Single-cell versions of these techniques at comparable coverages of specific loci achieved with MAPit-FENGC are presently not feasible due to low numbers of recovered unique fragments per cell. Fiber-seq[53], which uses a non-specific N6-adenine methyltransferase to increase footprinting resolution, also has limitations in that simultaneous detection of N6mA and endogenous 5mC is not presently possible. In sum, FENGC enrichment by itself or in combination with MAPit are expected to have broad applications in basic research, biomarker development, plant breeding, etc.

References

1. Kozarewa, I., Armisen, J., Gardner, A.F., Slatko, B.E. & Hendrickson, C.L. Overview of target enrichment strategies. *Curr Protoc Mol Biol* **112**, 7 21 21-23 (2015).
2. Kacmarczyk, T.J. et al. "Same difference": comprehensive evaluation of four DNA methylation measurement platforms. *Epigenetics Chromatin* **11**, 21 (2018).
3. Gnirke, A. et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* **27**, 182-189 (2009).
4. Samorodnitsky, E. et al. Comparison of custom capture for targeted next-generation DNA sequencing. *J Mol Diagn* **17**, 64-75 (2015).

5. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).

6. Shin, G. et al. CRISPR-Cas9-targeted fragmentation and selective sequencing enable massively parallel microsatellite analysis. *Nat Commun* **8**, 14291 (2017).

7. Nachmanson, D. et al. Targeted genome fragmentation with CRISPR/Cas9 enables fast and efficient enrichment of small genomic regions and ultra-accurate sequencing with low DNA input (CRISPR-DS). *Genome Res* **28**, 1589-1599 (2018).

8. Lee, J. et al. CRISPR-Cap: multiplexed double-stranded DNA enrichment based on the CRISPR system. *Nucleic Acids Res* **47**, e1 (2019).

9. Quan, J. et al. FLASH: a next-generation CRISPR diagnostic for multiplexed detection of antimicrobial resistance sequences. *Nucleic Acids Res* **47**, e83 (2019).

10. Bennett-Baker, P.E. & Mueller, J.L. CRISPR-mediated isolation of specific megabase segments of genomic DNA. *Nucleic Acids Res* **45**, e165 (2017).

11. Lin, Y. et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res* **42**, 7473-7485 (2014).

12. Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S. & Yang, S.H. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nucleic Acids* **4**, e264 (2015).

13. Varley, K.E. & Mitra, R.D. Bisulfite Patch PCR enables multiplexed sequencing of promoter methylation across cancer samples. *Genome Res* **20**, 1279-1287 (2010).

14. Nabilsi, N.H. et al. Multiplex mapping of chromatin accessibility and DNA methylation within targeted single molecules identifies epigenetic heterogeneity in neural stem cells and glioblastoma. *Genome Res* **24**, 329-339 (2014).

15. Lyamichev, V. et al. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat Biotechnol* **17**, 292-296 (1999).

16. Lyamichev, V., Brow, M.A., Varvel, V.E. & Dahlberg, J.E. Comparison of the 5' nuclease activities of Taq DNA polymerase and its isolated nuclease domain. *Proc Natl Acad Sci USA* **96**, 6143-6148 (1999).

17. Kaiser, M.W. et al. A comparison of eubacterial and archaeal structure-specific 5'-exonucleases. *J Biol Chem* **274**, 21387-21394 (1999).

18. Friedrich-Heineken, E., Henneke, G., Ferrari, E. & Hubscher, U. The acetylatable lysines of human Fen1 are important for endo- and exonuclease activities. *J Mol Biol* **328**, 73-84 (2003).

19. Tsutakawa, S.E. et al. Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**, 198-211 (2011).

20. Pardo, C.E., Darst, R.P., Nabilsi, N.H., Delmas, A.L. & Kladde, M.P. Simultaneous single-molecule mapping of protein-DNA interactions and DNA methylation by MAPit. *Curr Protoc Mol Biol* **Chapter 21**, Unit 21 22 (2011).

21. Kelly, T.K. et al. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res* **22**, 2497-2506 (2012).

22. Deleyrolle, L.P. & Reynolds, B.A. Isolation, expansion, and differentiation of adult mammalian neural stem and progenitor cells using the neurosphere assay. *Methods Mol Biol* **549**, 91-101 (2009).

23. Schutsky, E.K. et al. Nondestructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase. *Nat Biotechnol* (2018).

24. Sun, Z. et al. Nondestructive enzymatic deamination enables single-molecule long-read amplicon sequencing for the determination of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Genome Res* (2021).

25. Sönmezer, C. et al. Molecular co-occupancy identifies transcription factor binding cooperativity *in vivo*. *Mol Cell* **81**, 255-267 e256 (2021).

26. Knight, P. et al. Methylscaper: an R/Shiny app for joint visualization of DNA methylation and nucleosome occupancy in single-molecule and single-cell data. *Bioinformatics* **37**, 4857-4859 (2021).

27. Naor, D., Sionov, R.V. & Ish-Shalom, D. CD44: structure, function, and association with the malignant process. *Adv Cancer Res* **71**, 241-319 (1997).

28. Naor, D. Editorial: interaction between hyaluronic acid and its receptors (CD44, RHAMM) regulates the activity of inflammation and cancer. *Front Immunol* **7**, 39 (2016).

29. Liu, Y. et al. CD44 expression identifies astrocyte-restricted precursor cells. *Dev Biol* **276**, 31-46 (2004).

30. Phillips, H.S. et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* **9**, 157-173 (2006).

31. Verhaak, R.G. et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in *PDGFRA*, *IDH1*, *EGFR*, and *NF1*. *Cancer Cell* **17**, 98-110 (2010).

32. Anido, J. et al. TGF-β receptor inhibitors target the $CD44^{high}$/ $Id1^{high}$ glioma-initiating cell population in human glioblastoma. *Cancer Cell* **18**, 655-668 (2010).

33. Fu, J. et al. TGM2 inhibition attenuates ID1 expression in CD44-high glioma-initiating cells. *Neuro Oncol* **15**, 1353-1365 (2013).

34. Gaudreau, P.O. et al. WISP1 is associated to advanced disease, EMT and an inflamed tumor microenvironment in multiple solid tumors. *Oncoimmunology* **8**, e1581545 (2019).

35. Liu, Y. et al. The emerging role of WISP proteins in tumorigenesis and cancer therapy. *J Transl Med* **17**, 28 (2019).

36. Deng, W., Fernandez, A., McLaughlin, S.L. & Klinke, D.J., 2nd WNT1-inducible signaling pathway protein 1 (WISP1/CCN4) stimulates melanoma invasion and metastasis by promoting the epithelial-mesenchymal transition. *J Biol Chem* **294**, 5261-5280 (2019).

37. Jing, D., Zhang, Q., Yu, H., Zhao, Y. & Shen, L. Identification of *WISP1* as a novel oncogene in glioblastoma. *Int J Oncol* **51**, 1261-1270 (2017).

38. Sancho, M., Diani, E., Beato, M. & Jordan, A. Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. *PLoS Genet* **4**, e1000227 (2008).

39. Happel, N. & Doenecke, D. Histone H1 and its isoforms: contribution to chromatin structure and function. *Gene* **431**, 1-12 (2009).

40. Li, L.C. et al. A single nucleotide polymorphism in the E-cadherin gene promoter alters transcriptional activities. *Cancer Res* **60**, 873-876 (2000).

41. Ricketts, C., Zeegers, M.P., Lubinski, J. & Maher, E.R. Analysis of germline variants in *CDH1*, *IGFBP3*, *MMP1*, *MMP3*, *STK15* and *VEGF* in familial and sporadic renal cell carcinoma. *PLoS One* **4**, e6037 (2009).

42. Pittman, A.M. et al. The *CDH1*-160C>A polymorphism is a risk factor for colorectal cancer. *Int J Cancer* **125**, 1622-1625 (2009).

43. Krebs, A.R. et al. Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. *Mol Cell* **67**, 411-422 e414 (2017).

44. Hashimoto, H. et al. Structural basis for the versatile and methylation-dependent binding of CTCF to DNA. *Mol Cell* **66**, 711-720 e713 (2017).

45. Wu, D.Y. & Wallace, R.B. Specificity of the nick-closing activity of bacteriophage T4 DNA ligase. *Gene* **76**, 245-254 (1989).

46. Barany, F. Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc Natl Acad Sci USA* **88**, 189-193 (1991).

47. Hall, J.G. et al. Sensitive detection of DNA polymorphisms by the serial invasive signal amplification reaction. *Proc Natl Acad Sci USA* **97**, 8272-8277 (2000).

48. Varley, K.E. & Mitra, R.D. Nested Patch PCR enables highly multiplexed mutation discovery in candidate genes. *Genome Res* **18**, 1844-1850 (2008).

49. Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935 (2009).

50. Lai, B. et al. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. *Nature* **562**, 281-285 (2018).

51. Chen, H., Kharerin, H., Dhasarathy, A., Kladde, M. & Bai, L. Partitioned usage of chromatin remodelers by nucleosome-displacing factors. *Cell Rep* **40**, 111250 (2022).

52. Ahmad, K., Henikoff, S. & Ramachandran, S. Managing the steady state chromatin landscape by nucleosome dynamics. *Annu Rev Biochem* **91**, 183-195 (2022).

53. Stergachis, A.B., Debo, B.M., Haugen, E., Churchman, L.S. & Stamatoyannopoulos, J.A. Single-molecule regulatory architectures captured by chromatin fiber sequencing. *Science* **368**, 1449-1454 (2020).

Methods

Cell culture. Human fetal telencephalic NSC and human GBM cell lines (L0 and Nx18-25) were cultured in complete NSC medium (basal medium + proliferation supplement at a 9:1 ratio; NeuroCult™ NS-A Proliferation Kit, STEMCELL Technologies, 05751) supplemented with penicillin-streptomycin (1% final concentration; ThermoFisher Scientific Gibco, 15140122), 20 ng/ml human recombinant epidermal growth factor (STEMCELL Technologies, 78006.1), 10 ng/ml human recombinant basic fibroblast growth factor (STEMCELL Technologies, 78003.1), and 0.679 U/ml heparin (Sigma, H3149)[22]. For NSC culture, 10 ng/ml leukemia inhibitory factor (Millipore, LIF1050) was also added. The cells were maintained in a humidified incubator at 37°C and 5% CO_2. A standard protocol was used for passaging the NSC[54] and GBM cells[55], whereby the neurospheres were collected by centrifugation at 110*g* for 5 min every 7-10 days. The pellet was resuspended in 0.05% (w/v) trypsin, 0.53 mM ethylenediaminetetra-acetic acid (EDTA; ThermoFisher Scientific, 25300062), prewarmed to 37 °C. Soybean trypsin inhibitor (ThermoFisher Scientific, 17075029) was then added and gentle pipetting used to dissociate the neurospheres into single cells for replating.

Animal experiments. Animal protocols were approved by the University of Florida Institutional Animal Care and Use Committee (IACUC Protocols 201807422 and 201910745). 'Guide for the Care and Use of Laboratory Animals' was adhered to as prepared by the Committee for the Update of the Guide for the Care and Use of Laboratory Animals of the Institute for Laboratory Animal Research, National Research Council. Mice were closely monitored for signs of dehydration, weight loss, impaired mobility or physiological signs of underlying disorders such as labored breathing or respiratory distress.

C57BL/6J female mice were used in order to evaluate detection of distinct transcriptionally active and inactivated epialleles from the X chromosome. All mice were 8-10 weeks of age upon arrival and individually housed on a 12 h dark/12 h light cycle at 19-22°C and 30-60% humidity, with standard chow diet and water provided *ad libitum*. Prior to cell collection at 4 mo age, anesthetization of mice was induced and maintained with 5.0% and 1.5% isoflurane USP (Patterson Veterinary Supply, NDC 14 043-704-06), respectively, using an Eagle Eye Model 150 anesthesia machine (Jacksonville, FL, USA). The depth of anesthesia was monitored by the absence of pedal withdrawal reflex.

Isolation of monocytes from mouse bone marrow. Bone marrow was collected from spines (below skull to above tail) cleaned of extraneous tissues. All subsequent steps were conducted under sterile conditions. First, the tissue was crushed at room temperature using a ceramic mortar and pestle in a sterile solution of 10 ml of phosphate-buffered saline (PBS), pH 7.2, 2 mM EDTA, 0.5% (w/v) bovine serum albumin (BSA), by mixing MACS BSA Stock Solution (Miltenyi, 130-091-376) and Biotec, autoMACS® Rinsing Solution (Miltenyi, 130-091-222) in a ratio of 1:20. Subsequently, the crushed spine was removed, and the homogenate was filtered through pre-separation filters with 30-µm nylon mesh (Miltenyi Biotec, 130-041-407,) into a 15-ml Falcon® tube. The cellular filtrate was then washed by centrifugation at 300*g* for 10 min, removal of the supernatant, and resuspended in the same PBS/BSA solution for cell counting (Heska, Element HT5 Hematology Analyzer). Monocytes were then enriched using a negative isolation protocol specific to mouse bone marrow (Miltenyi Biotec, 130-100-629), according to the manufacturer's protocol. Briefly, ~40-80 million washed cells were incubated with FcR blocking reagent and cocktail of mouse biotinylated-antibodies, washed with PBS/BSA, resuspended in degassed RPMI 1640 medium (ThermoFisher, 11875101) containing 1% (w/v) penicillin-streptomycin (ThermoFisher Gibco, 15140163), and efficiently depleted of non-target cells by addition of magnetic microbeads conjugated to mouse anti-biotin monoclonal IgG1 antibodies and passage through LS ferromagnetic columns (Miltenyi Biotec, 130-042-401). The flow through, containing highly enriched bone marrow-derived monocytes, was collected into a 1-ml centrifuge tube. To allow recovery from the collection process, a mean of 1.2 million monocytes were plated in a well on a 96-well plate and incubated in a humidified 37°C incubator at 5% CO_2 in RPMI 1640 containing 1% penicillin-streptomycin solution for 3 h before harvesting for MAPit-FENGC.

FEN cleavage of oligonucleotide targets. Cleavage efficiencies of *Thermococcus* 9° N FEN1 and Taq (Supplementary Fig. 1b) were tested on a 129-nt 5' flap assembled by incubating 500 nM each of 200-mer T target, 200 oligo 1-T, and oligo U1 (Supplementary Table 1, Sheet 1). The 200-mer T sequence corresponds to a segment of pGEM-3Z/601b[56] with A 2453 mutated to T. Reactions also contained 1 U APEX Taq (Genesee Scientific), 1 U HotStar Taq (Qiagen; based on polymerization activity), and 32 U FEN1 (New England Biolabs). The buffers used were: 1x PCR buffer (Qiagen; referred to as Taq buffer); 1x ThermoPol reaction buffer (New England Biolabs; referred to as FEN1 buffer); and the mixed buffer reactions contained final concentrations of 1x CutSmart buffer (New England Biolabs) and 1x Taq buffer or FEN1

buffer. The incubation conditions were 3 min at 95°C, followed by 20 min at 65°C and then 9 cycles of 0.5 min at 95°C and 10 min at 65°C. Digestion products were separated on an Agilent 2100 Bioanalyzer and the areas under the peaks of cut and uncut 200-mer T were integrated to determine the cleavage percentages.

To test the effects of different 'overlapping' nucleotides and DNA methylation on FEN1 activity, two 80-mer targets of identical sequence were synthesized with and without five 5mC residues at the same locations (Supplementary Fig. 2a; Supplementary Table 1, Sheet 1). These 5 cytosines occupy the upstream duplex contacted by human FEN1 co-crystallized with a 5' flap and 1-nt 3' flap[19]. The two 80-mers (50 nM) were assembled in separate 20-µl reactions with their corresponding unmethylated flap adapters (500 nM each of oligo U1-A, C, G, or T and 80 oligo-A, C, G, or T; Supplementary Fig. 2b,c; Supplementary Table 1, Sheet 1). This reconstitutes A, C, G, or T 5' of the cleavage site and as the 1-nt 3' flap, as well as recapitulates the scenario of performing FENGC with methylated DNA (Supplementary Fig. 2a). Reactions also contained 1 x FEN1 buffer and 0, 0.01, 0.04, 0.2, or 1 unit(s) of FEN1 and were incubated for 3 min 95°C, followed by 20 min at 65°C. The percentages of cleaved 80-mer at increasing doses of FEN1 were determined by quantitative, real-time PCR, calculated by $[1-2^{-(\text{Ct of FNE1} - \text{Ct of no FEN1})}] \times 100$. The dose *versus* response curves (Supplementary Fig. 2b) and EC_{50} values (Supplementary Fig. 2c) were plotted using GraphPad Prism 5 (www.graphpad.com).

Methylation of accessible GC sites for single-molecule MAPit methylation footprinting. Two million cells were harvested with 0.05% (w/v) trypsin, 0.53 mM EDTA (ThermoFisher Scientific, 25300062) and washed once with ice-cold PBS, pH 7.2, supplemented with 0.015% (w/v) sodium azide. Cells were then centrifuged for 5 min at 1,000*g*, washed once with 500 µl ice-cold cell resuspension buffer (20 mM HEPES, pH 7.5, 70 mM $MgCl_2$, 0.25mM EDTA pH 8.0, 0.5mM EGTA pH 8.0, 0.5% (v/v) glycerol, freshly supplemented 10 mM DTT and 0.25 mM PMSF). After centrifugation for 5 min at 1,000*g*, cells were resuspended in 180 µl cell resuspension buffer with supplemented with 0.05% (w/v) digitonin and incubated on ice for 10 min. Cells were stained with trypan blue to ascertain 100% permeabilization. Cells were then divided into two aliquots and treated with and without 100 U M.CviPI GpC methyltransferase (100 U/million cells; New England Biolabs, M0227B-HI) supplemented with fresh 160 µM *S*-adenosyl-*L*-methionine for 15 min at 37ºC. Reactions were stopped by addition of an equal volume of 1% (w/v) SDS, 100 mM NaCl, 10 mM EDTA and vortexed briefly at medium speed. Nuclei were treated with 100 µg/ml RNase A for 30 min at 37ºC followed by 100 µg/ml proteinase K treatment at 50ºC overnight. Genomic DNA was extracted using phenol-chloroform-isoamyl alcohol (25:24:1, v/v) phase separation, followed by ethanol precipitation, and resuspension in molecular-grade H_2O.

FENGC. Except as noted otherwise, the human gene panels targeted the promoter TSSs of genes coding for proteins with DNA repair and/or cancer-associated functions. Flap oligos 1 and 2 as well as nested oligos 3 were designed for 11 targets of ~300-nt, 119 targets of ~450-nt, and 45 targets of ~940-nt, as catalogued in the respective sheets of Supplementary Tables 1 and 2. The mouse panel targeted the promoters of 78 inflammatory response genes, unless noted otherwise.

To mitigate 5′ flap secondary structure, purified gDNA from M.CviPI-treated cells was fragmented by sonication with a UCD-200 Bioruptor (Diagenode) at high level for 25 sec in 100 µl molecular-grade H_2O. Subsequently, the volume was brought to 20 µl in a SpeedVac vacuum concentrator.

Sequence-specific flap oligos 1 and 2 as well as oligo 3 (4 nmole each) were ordered as 0.3-ml each in 96-well format as salt-free oligos (Eurofins). Lyophilized oligos were resuspended

at 100 μM concentration in molecular-grade water. Unless indicated differently, 500 ng gDNA was processed in each reaction. Multiplexed cleavage of 5′ flaps was performed by adding 1 µl 10 µM oligo U1-T, 2 µl pooled oligos 1, 2 µl pooled oligos 2, 1 x Qiagen PCR Buffer (Qiagen), 3 U APEX Taq (Genesee Scientific), and molecular-grade H_2O to bring the total volume to 35 µl.

Reactions were incubated for 3 min at 95°C, then 20 min 65°C, followed by 14 cycles of 0.5 min at 95°C and 10 min at 65°C. Subsequently, 2 µl pooled oligos 3, 1 µl 10 µM oligo U2, 1 x Ampligase® Reaction Buffer (Lucigen), 10 U Ampligase (Lucigen), and molecular-grade H_2O were added to bring the total volume to 45 µl. Incubation of the reactions at 95°C for 3 min was followed by 30 s at 95°C and 8 min at 65°C for 100 cycles. Unprotected DNA was removed by addition of 20 U exonuclease I, 100 U exonuclease III (both from New England Biolabs), and incubation for 60 min at 37°C and 20 min at 80°C.

The reduced representation products were subjected to either standard PCR, BS-PCR, or EM-PCR. For standard PCR, captured DNA was purified with MinElute PCR purification kit (Qiagen) or AMPure XP beads (1.8x), and PCR was performed with HotStar Taq (Qiagen) as described above. For BS-PCR, captured sequences were converted with EZ DNA Methylation Gold Bisulfite Treatment Kit (Zymo Research) as indicated above, except that 35 cycles of amplification were used. For EM-PCR, captured sequences were purified with 1.8x AMPure XP beads (Beckman Coulter) or 1.8x NEBNext beads (New England Biolabs). Enzymatic conversion of cleaned sequences was done with NEBNext® Enzymatic Methyl-seq Conversion Module (New England Biolabs, E7125), according to the manufacturer's instructions, and converted sequences were eluted in 14 µl molecular-grade H_2O. PCR was performed the same way as BS-PCR. For EM-PCR of ~940-nt targets, sequences were eluted with 25 µl molecular-grade H_2O. PCR was performed with 500 nM of oligo U1 and U2 primer in 2x KAPA HiFi HotStart Uracil+ ReadyMix (Roche) in 50 µl. Amplification conditions for BS- and EM-PCR were 5 min at 95°C, followed by 30 cycles of 20 sec at 98°C, 15 sec at 62°C, 30 sec at 72°C, and a final extension of 1 min at 72°C. Amplified FENGC products were purified with AMPure XP beads (0.65x for ~500-bp products; 0.75x for ~350-bp products).

Pacific Biosciences Sequel and Sequel IIe SMRT sequencing

Purified amplicons passing quality evaluation on the Agilent TapeStation high-sensitivity DNA5000 tape were submitted to the University of Florida Interdisciplinary Center for Biotechnology Research (UF-ICBR, RRID:SCR_019152) for SMRT bell library construction (P/N 101-791-700 Version 07). The manufacturer's protocol was used, except the input amount of each amplicon was increased 3-6x for DNA damage repair, end repair, A-tailing, and barcode ligation. Equimolar amounts of barcoded amplicons were pooled and ligated to SMRTbell adaptors, which were then pooled in equimolar amounts and treated with exonuclease, using the SMRTbell Enzyme Clean-Up Mix (PacBio, 101-932-600). Library construction typically yielded sufficient SMRT bell library (~70 ng) for 2-3 SMRT cells.

Sequencing on the PacBio Sequel or Sequel IIe was performed by UF-ICBR and on the Sequel IIe at the Georgia Genomics and Bioinformatics Core at the University of Georgia. The library pool was loaded at 120 pM, using diffusion loading and 20- to 30-h movies, with HiFi generation and demultiplexing. Sequencing Kit 2.0 (PacBio, 101-389-001) and Instrument Chemistry Bundle Version 11.0 were used. All other steps were performed according the recommended protocol by the PacBio sequencing calculator. One Sequel IIe SMRT cell typically resulted in 3-4 million polymerase reads.

High-fidelity circular consensus sequencing (CCS) reads were generated at ≥3 single polymerase read passes to determine on-target read efficiencies, with other default parameters. To do so, CCS reads were aligned to the human genome (build GRCh38) using the BWA

aligner (bwa-mem)[57]. For reads from deaminated samples, the alignment was performed using bwameth[58], which uses the BWA aligner. Reads that aligned anywhere to the human genome were then extracted and aligned to the target references, again using either BWA or bwa-meth. On-target efficiencies were then calculated as the percent of aligned reads. Both BWA and bwa-meth were used with default parameters.

For epigenetic analyses, CCS reads were generated with default parameters, except for a setting of ≥5 single polymerase read passes. CCS reads were aligned to reference sequences with the python reAminator pipeline[59]. Cut-offs of 95% conversion rate and 95% length of reference sequences alignment were applied. To distinguish endogenous CG methylation and M.CviPI-probed GC methylation unambiguously GCG sites were omitted for calculation of methylation levels, i.e., HCG and GCH, where H is A, C, or T. Single-molecule heatmaps of HCG and GCH methylation were hierarchically clustered (unsupervised) and plotted using methylscaper[26].

For genotyping, CCS reads were obtained with at least 3 passes and minimum predicted accuracy of 0.99. Reads alignment was performed using minimap2 and SNPs and indels calling were carried out using GATK4 pipeline[60]. All variants were filtered with thresholds of GQ>30 and Depth>20, as well as AS_QD<2.0 for SNPs and AS_QD<5.0 for indels.

Reverse transcriptase quantitative PCR (RT-qPCR)

RNA was extracted from one million cells using Direct-zol RNA Miniprep Kit (Zymo Research, 11331). Complementary DNA was synthesized using SuperScript™ First-Strand Synthesis system (Thermo Fisher Scientific, 11904018), then diluted 1:1 with molecular-grade H_2O and used for qPCR measurement. QPCR was done using Taqman assay (Life Technologies 44-449-63) with gene-specific probes *CCN4* (*WISP1*; Thermo Fisher Scientific Hs00180245_m1), *CD44* (Thermo Fisher Scientific Hs01075861_m1), and *HIST1H1B* (Thermo Fisher Scientific Hs01075861_m1).

FENGC Oligonucleotide Designer (FOLD)

For FENGC of the ~620-nt target panel (Supplementary Table S2, Sheet 4) from mouse monocytes, the primers (Supplementary Table S1, Sheet 5) were designed by newly developed program, FOLD (https://github.com/uf-icbr-bioinformatics/FOLD). The programs searches an input file of gene names or genome coordinates for primers that avoid repeats and satisfy criteria of Figure 1, such as locating the 'overlapping' residues that create 1-nt 3′ flaps. Other command-line options include, but are not limited to, increasing the length of default 500-nt sequences and percentage tolerance of departure from this specified, specification of annotated TSS (e.g., RefSeq), and minimum and maximum primer melting temperature (T_m).

FOLD implements a sophisticated optimization method to find the best combination of primers for each input sequence (local optimization), and for all input sequences at the same time (global optimization). For each input sequence, the program identifies all possible locations for each one of the three primers, and finds the optimal triple, taking into account the predicted T_m for each primer and the length of the corresponding products. The deviation from ideal T_m and ideal product length are combined through user-defined weights, and the set of primers with the smallest resulting score is selected.

The goal of global optimization is to ensure that the sizes of the produced fragments for all target sequences are as similar as possible, i.e., exhibit the smallest variance. This is accomplished by picking a random set of primers for each target (among the ones that fulfill the specified design constraints), computing the variance of the resulting fragment sizes, and repeating the process a large number of times, and finally picking the combination that gave the

smallest overall variance. Sub-optimal primer sets are therefore accepted for some targets in order to attain more uniform fragment sizes at the global level.

The program's main output is a table listing the coordinates and size of the resulting PCR product, and information on the three primers (including sequence, length, T_m, and GC%) for each target in input. Optionally, the program can also write the sequences of all PCR products and of all primers to user-specified files in FASTA format.

Statistics and Data Analysis.
To evaluate the 0.1% NSC gDNA:99.9% GBM L0 gDNA mixture (Fig. 1e), a proportions test was conducted in R using the command prop.test(), with a null hypothesis that the observed proportion of 10 hypermethylated out of 1,781 unmethylated epialleles is less than 0.1%. The *P* value of <.0001 indicates that the proportion of hypermethylated epialleles detected was at least or greater than 0.1%.

To identify differentials in H5mCG (methylation) and G5mCH (accessibility) between NSC *versus* GBM Nx18-25, generalized estimating equations were used to model the effect of cell line on the proportion of methylation per molecule. Only targets with ≥50 total CCS reads in the combined replicates (≥22 in any single replicate) were considered (Supplementary Table 5). Errors were modeled as normally distributed, and the correlation structure was assumed exchangeable within each replicate. The geeglm function from the geepack v1.3-2 R package[61] was used in R version 4.1.0. *P* values were corrected for multiple testing using the Bonferroni method to control the family-wise error rate.

To test if MAPit-FENGC detected significant epigenetic differences in at least one of the four analyzed mouse bone marrow-derived monocyte samples, smoothed moving averages (20-bp window) of DNA methylation and accessibility across each gene region were modeled using a mixed effects ANOVA[62]. Testing was restricted to 43 amplicons with ≥100 filtered CCS reads per sample and good diversity, i.e., absence of many duplicates (Supplementary Table 10). The mixed effects ANOVA model was fit using the gls function in the nlme v3.1-152 package in R version 4.1.0. Each sample was treated as a random effect, and correlation along the gene region was modeled as that from an autoregression moving average model (ARMA). Autocorrelation parameters were estimated using the auto.arima function from the forecast v8.15 R package, with a maximum possible value of 1 to avoid overfitting. If the differencing order was estimated as zero, then first-order differences were taken of DNA methylation and accessibility. If the autoregressive and moving average parameters were both estimated as zero, then an autoregressive model of order one was used. The *P* values are of the interaction term of base pair and replicate, which tested for differences between replicates across the gene region. *P* values were corrected for multiple testing using the Bonferroni method to control the false discovery rate (Supplementary Table 10).

The violin density plots were made using the "yarrr" R package[63]. The median (horizontal line), interquartile range of methylation levels (box), and smoothed probability density at different methylation levels (gray area) are plotted. *P* values are corrected for multiple testing using the Bonferroni method (α = 0.05) to control the false discovery rate (Supplementary Table 6).

Additional References

54. Fortin, J.M. et al. Transplantation of defined populations of differentiated human neural stem cell progeny. *Sci Rep-Uk* **6** (2016).

55. Azari, H. et al. Isolation and expansion of human glioblastoma multiforme tumor cells using the neurosphere assay. *Jove-J Vis Exp* (2011).

56. Dechassa, M.L. et al. SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. *Mol Cell* **38**, 590-602 (2010).

57. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

58. Pedersen, B.S., Eyring, K., De, S., Yang, I.V. & Schwartz, D.A. Fast and accurate alignment of long bisulfite-seq reads. *arXiv* **1401**, 1129 (2014).

59. Stees, J.R. et al. High fractional occupancy of a tandem Maf recognition element and its role in long-range β-globin gene regulation. *Mol Cell Biol* **36**, 238-250 (2016).

60. Wenger, A.M. et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol* **37**, 1155-1162 (2019).

61. Højsgaard, S., Halekoh, U. & Yan, J. The R package geepack for generalized estimating equations. *2005* **15**, 11 (2005).

62. Pinheiro, J.C. & Bates, D.M. Mixed-effects models in S and S-Plus. (Springer, 2000).

63. Phillips, N. yarrr: A Companion to the e-Book "YaRrr!: The Pirate's Guide to R. (2017).

Data availability

Raw data are deposited in NCBI BioProject under accession PRJNA752452.

Code availability

Code will be made available at https://github.com/uf-icbr-bioinformatics/FOLD.

Acknowledgements

We are grateful to the Next-generation DNA Sequencing (RRID:SCR_019152) as well as Gene Expression and Bioinformatics Cores RRID:SCR_019145) of the Interdisciplinary Center for Biotechnology Research of University of Florida for their technical support. This work was supported by grants HDTRA1-16-1-0048 awarded by the Defense Threat Reduction Agency to P.C. and R01 CA155390 awarded by The National Institutes of Health to M.P.K.

Author information

Contributions

N.H.N. and M.P.K. conceived and designed the FENGC approach, and N.H.N. conducted initial proof-of-concept experiments with plasmids. M.Z. and M.P.K. conceived the study and designed experiments that M.Z. conducted to develop the mature FENGC protocol. H.Z. and B.A.R. established the novel GBM cell lines. K.O.M. and T.L.C. prepared the primary mouse monocytes and selected the inflammatory response gene targets. A.W. and M.-P.L.G. performed MAPit and qPCR. J.O.B., A.R., J.R.B.N., and W.S.O. provided bioinformatics and computational support, as well as data analyses. A.R. also wrote program FOLD. R.B. conducted biostatistical analyses. F.J.P.-P. and A.C. analyzed the genotyping sequences and prepared the data table. M.Z., P.C., and M.P.K. wrote the manuscript. All authors reviewed, edited, and approved the manuscript.

Ethics declarations

Competing interests

M.Z., N.H.N., and M.P.K. are inventors (assignee, University of Florida) of pending patent application (PCT/US2022/020624, filed March 16, 2022) for the FENGC technology.

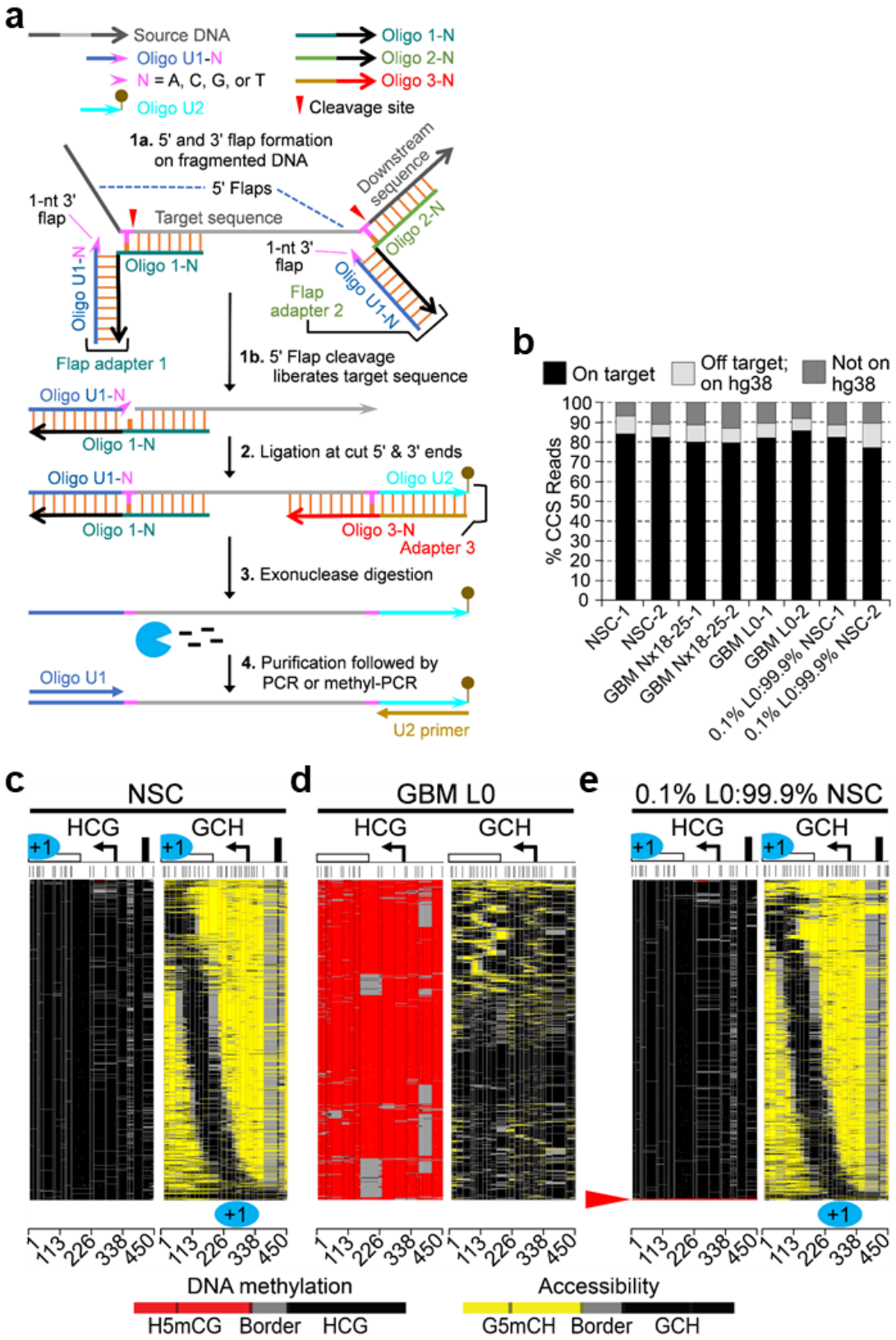


Fig. 1. FEN activity-dependent target sequence enrichment yields high percentages of on-target reads, profiles epigenetic heterogeneity, and detects low-frequency epialleles. a,

FENGC procedure. Mildly fragmented gDNA is hybridized to a pair of locus-specific flap adapters, reconstituting a double flap (5' and 1-nt 3' flap) at both ends of each target sequence (step 1a). To form 1-nt 3' flaps, nucleotides (emphasized in pink) 5' of the scissile phosphodiesters (red arrowheads) 'overlap' and thus displace the 3'-terminal nucleotide of oligo U1-A, C, G, or T (oligo 1-N). Specific cleavage by FEN activity of FEN1 or Taq DNA polymerase I releases single-stranded target strands with ligatable ends, i.e., 5' phosphate downstream of a 1-nt gap and a 3'-hydroxyl (step 1b). The gap is filled by the 1-nt 3' flap, which ligates to the cut 5' end (step 2). A nested 3'-end-specific adapter 3 is added to position oligo 2 for ligation to the cut 3' end. Covalent modification of oligo U2 (lollipop) protects against 3' to 5' exonuclease degradation, dramatically enriching target sequences (step 3). All reagent addition to this point is sequential, avoiding loss of target sequences. Enriched targets are purified and amplified with either standard PCR or methyl-PCR to construct NGS libraries (step 4). **b,** Proportions of Hi-Fi EM-seq CCS reads for MAPit-FENGC of 119 promoter targets mapping to or off the human genome. Suffixes denote the two independent biological replicates for the indicated sequenced libraries. **c-e,** Methylscaper heatmaps of HCG and GCH methylation of single molecules generated from filtered EM-seq CCS reads aligning to *EPM2AIP1* (+307 to **–**143) in the combined replicates of NSC (3,460 reads) (**c**), GBM L0 (3,889 reads) (**d**), and 0.1% L0 gDNA spiked into NSC gDNA (1,781 reads) (**e**). In each pair of plots, the patterns of HCG methylation (left; consecutive H5mCG colored red) and GCH accessibility (right; consecutive G5mCH colored yellow) on each molecule or epiallele shown in the same top-to-bottom order. Two or more consecutively unmethylated sites are connected by black, gray connects the borders between methylated and unmethylated sites, and white indicates missing or unaligned sequences. Arrow, TSS; white rectangle, *EPM2AIP1* protein coding exon; blue partial and full ellipses, footprinted +1 nucleosomes; black rectangle, footprinted TF. Vertical ticks and lines across panels mark respective HCG and GCH motifs. In these and subsequent plots, all features and intervening distances are drawn to scale. A proportions test in R indicated that the proportion of 10 CG-hypermethylated of 1,781 total epialleles detected from the duplicate samples in **e** is at least or greater than 0.1% (P = 3.59E**–**09). Base pair position along each target is indicated at bottom.

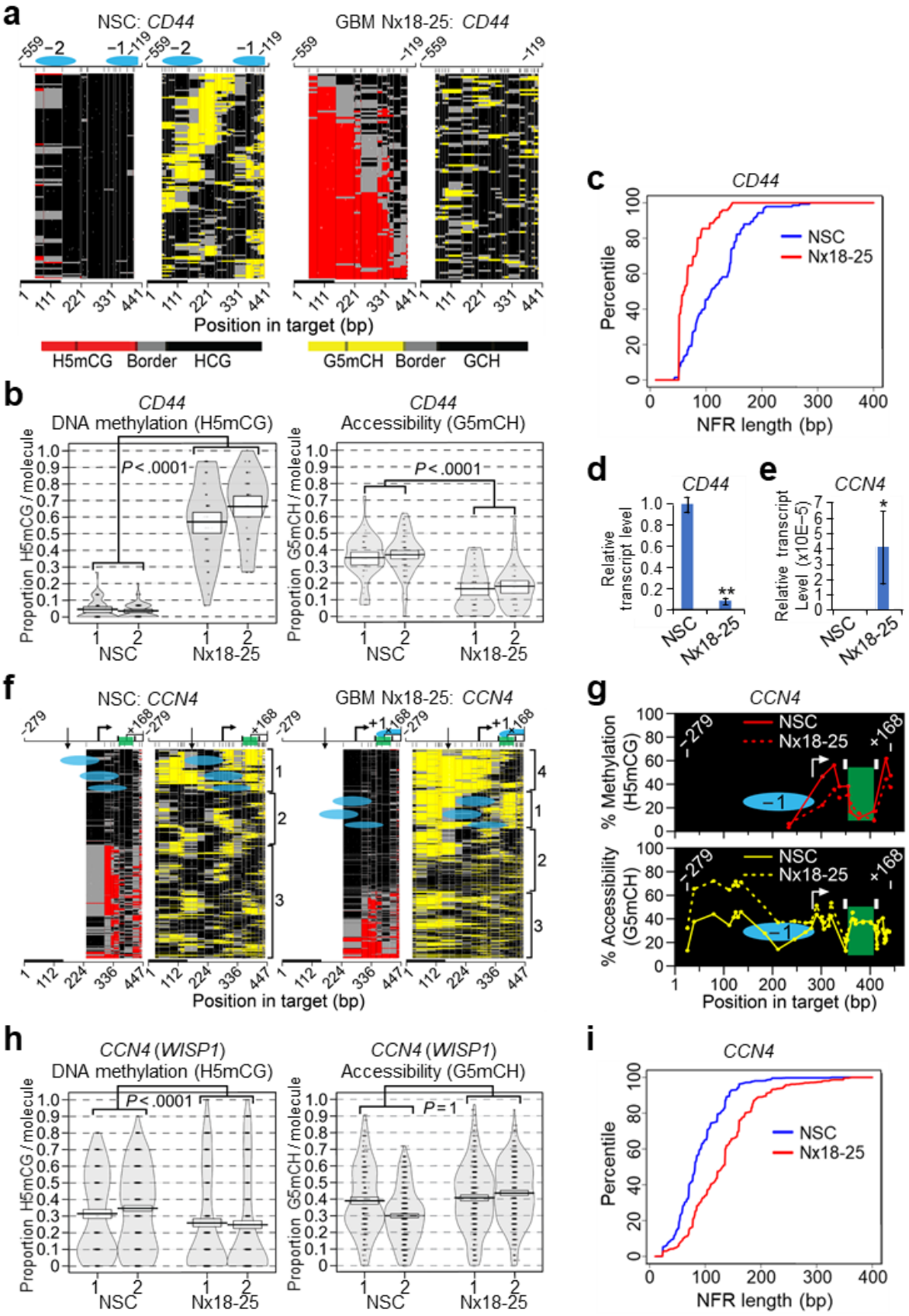


Fig. 2. Targeted MAPit-FENGC detects differential epigenetic configurations that correlate with gene expression. a-c, Filtered EM-seq CCS reads aligning to *CD44* from human NSC (*n* =

2; 141 combined reads) and GBM Nx18-25 (*n* = 2; 96 combined reads) plotted by methylscaper as clustered single-molecule methylation heatmaps (**a**), violin plots of the per molecule proportion of H5mCG (endogenous DNA methylation, left) and G5mCH (accessibility, right) (**b**), and cumulative distribution function of NFR length (mean = 118 bp in NSC and 72.0 bp in Nx18-25, *P* = 2.2E−16) (**c**). **d,e,** Relative expression of *CD44* (*n* = 3, mean ± SD; **, *P* < .01) (**d**) and *CCN4* (*n* = 3, mean ± SD; paired Student's *t* test, two sided, *, *P* < .05) (**e**). **f-i,** Filtered EM-seq CCS reads aligning to *CCN4* from human NSC (*n* = 2; 1,116 combined reads) and GBM Nx18-25 (*n* = 2; 1,143 combined reads; Supplementary Table 5) plotted by methylscaper (**f**), percentages of averaged H5mCG and G5mCH (**g**), violin plots (**h**), and cumulative distribution function of NFR length (mean = 89.4 bp in NSC and 132 bp in Nx18-25, *P* = 2.2E−16) (**h**). Although the difference in H5mCG per molecule between NSC and GBM Nx18-25 in **h** was significant, the difference in accessibility was not, probably due to uncharacteristically high accessibility of molecules bearing H5mCG. Green rectangles in **f,g** demarcate local depletion of H5mCG between two footprinted factors (black rectangles) in cluster 1 molecules. In **f**, a known GA>A indel (dbSNP, rs548251181; straight arrow; plotted in white as it does not align to hg38) present at similar frequencies in both cell types was verified by MAPit-FENGC genotyping with standard PCR amplification (Supplementary Table 7, Indels). White rectangle, coding exon 1 sequence.

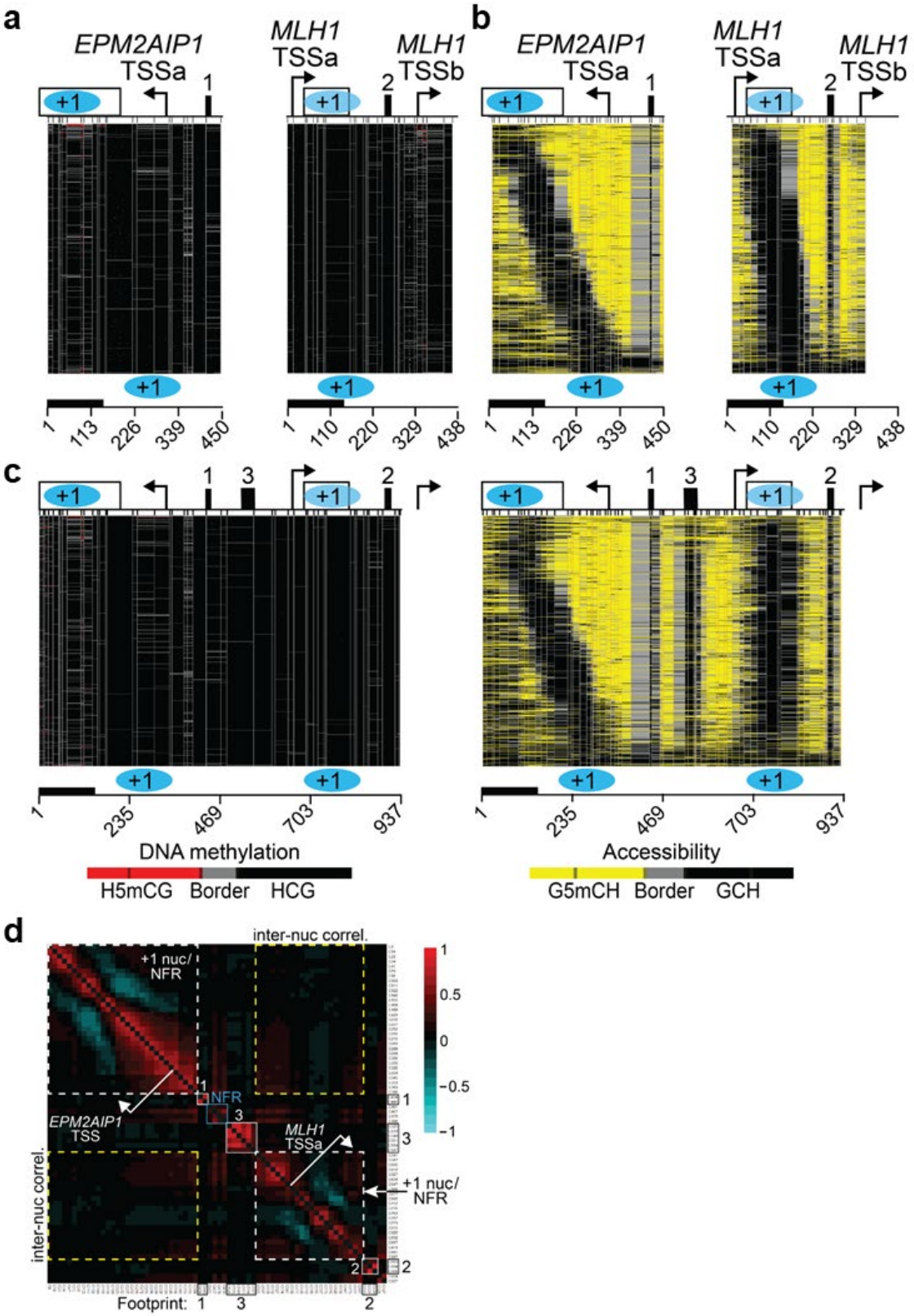


Fig. 3. Independent mobilization of the +1 nucleosomes of the divergently transcribed *EPM2AIP1* and *MLH1* promoters. **a-c,** Methylscaper plots of filtered EM-seq CCS reads aligning to *EPM2AIP1* (+307 to −143) (**a**), *MLH1* (−13 to +425 relative to *MLH1* TSSa) (**b**), and

EPM2AIP1-MLH1 (+328 to −609 relative to *EPM2AIP1* TSS; −650 to +287 relative to *MLH1* TSSa) (**c**). All images contain 1,000 molecules randomly selected from the combined GBM Nx18-25 replicates. The 450-bp *EPM2AIP1* and 438-bp *MLH1* molecules in **a,b** were hierarchically clustered without specifying a subregion and then aligned to the 937-bp *EPM2AIP1-MLH1* molecules in **c**. Symbols as defined above. **d,** Correlation of G5mCH levels (each C indicated at bottom and right) of the same 1,000 molecules plotted in **c**. TF footprints in **c** display expected positive correlations (red) adjacent to the diagonal due to co-protection or co-accessibility of adjacent GCH sites (within squares with solid white lines). Off-diagonal positive correlations within each +1 nucleosome (nuc)/NFR (circumscribed by white-dashed squares) are due to co-protection by histone octamers or co-accessibility in NFRs. Negative correlations (cyan; C70-C155 and C169-C283 in *EPM2AIP1*; C627-C735 and C763-C847 in *MLH1*) arise from inverse protection and accessibility between histones octamers and flanking linkers. Each span of negative correlation corresponds well with the observed range of nucleosome positions at each promoter in **b**. The absence of strong positive or negative inter-nucleosome correlations (inter-nuc correl.; within yellow-dashed rectangles) supports a conclusion that the two +1 nucleosomes are repositioned independently of each other.

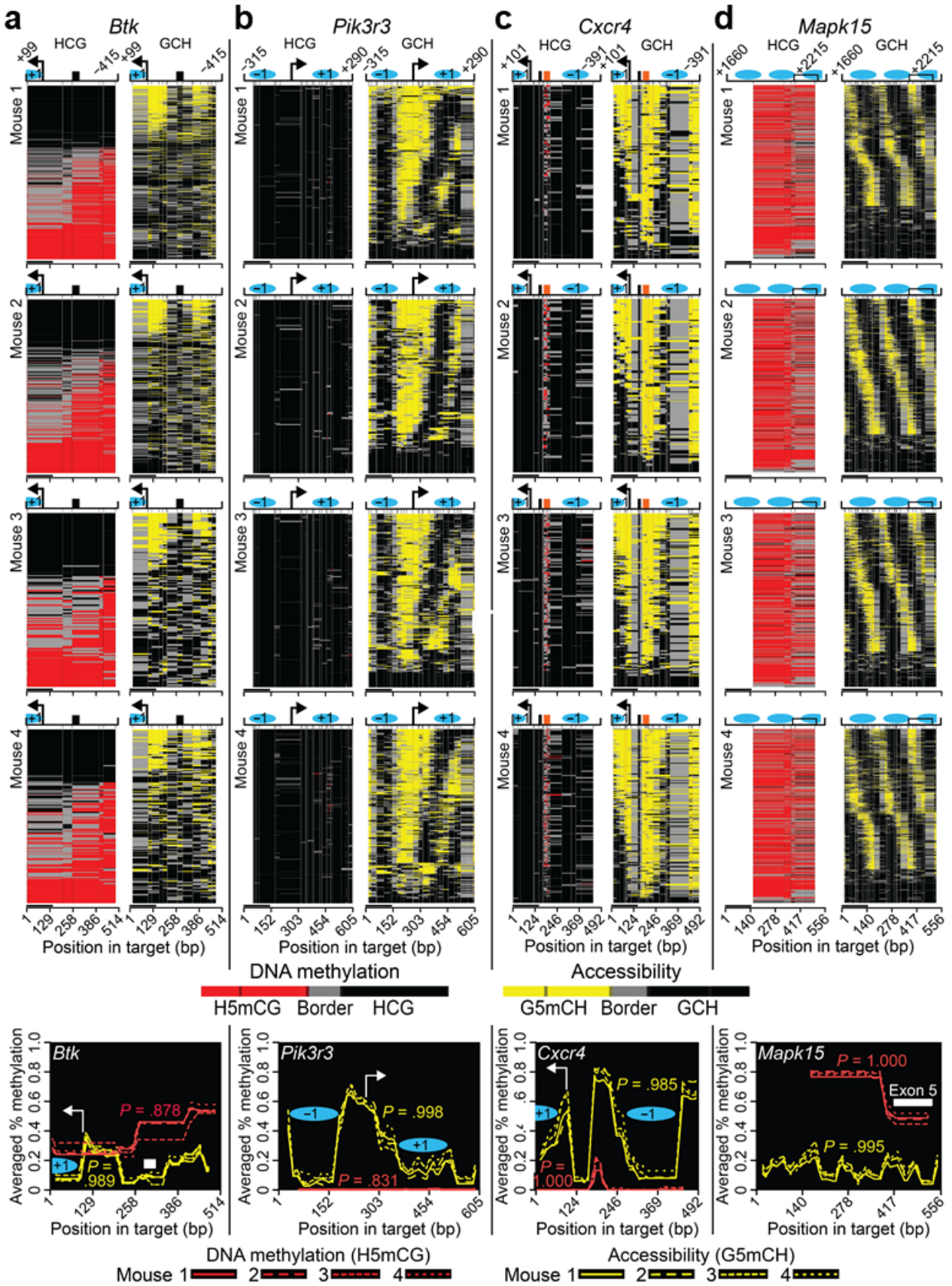
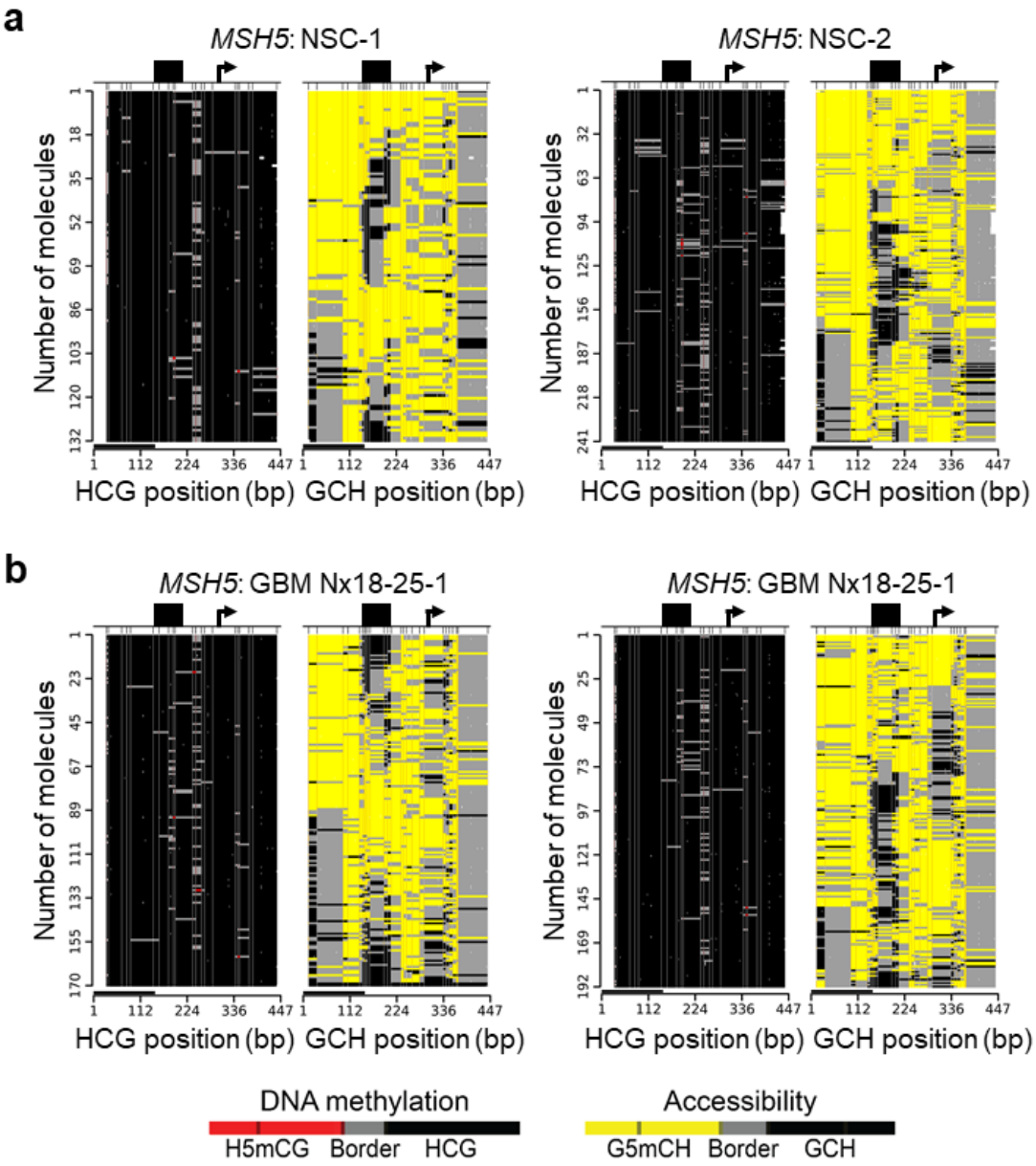


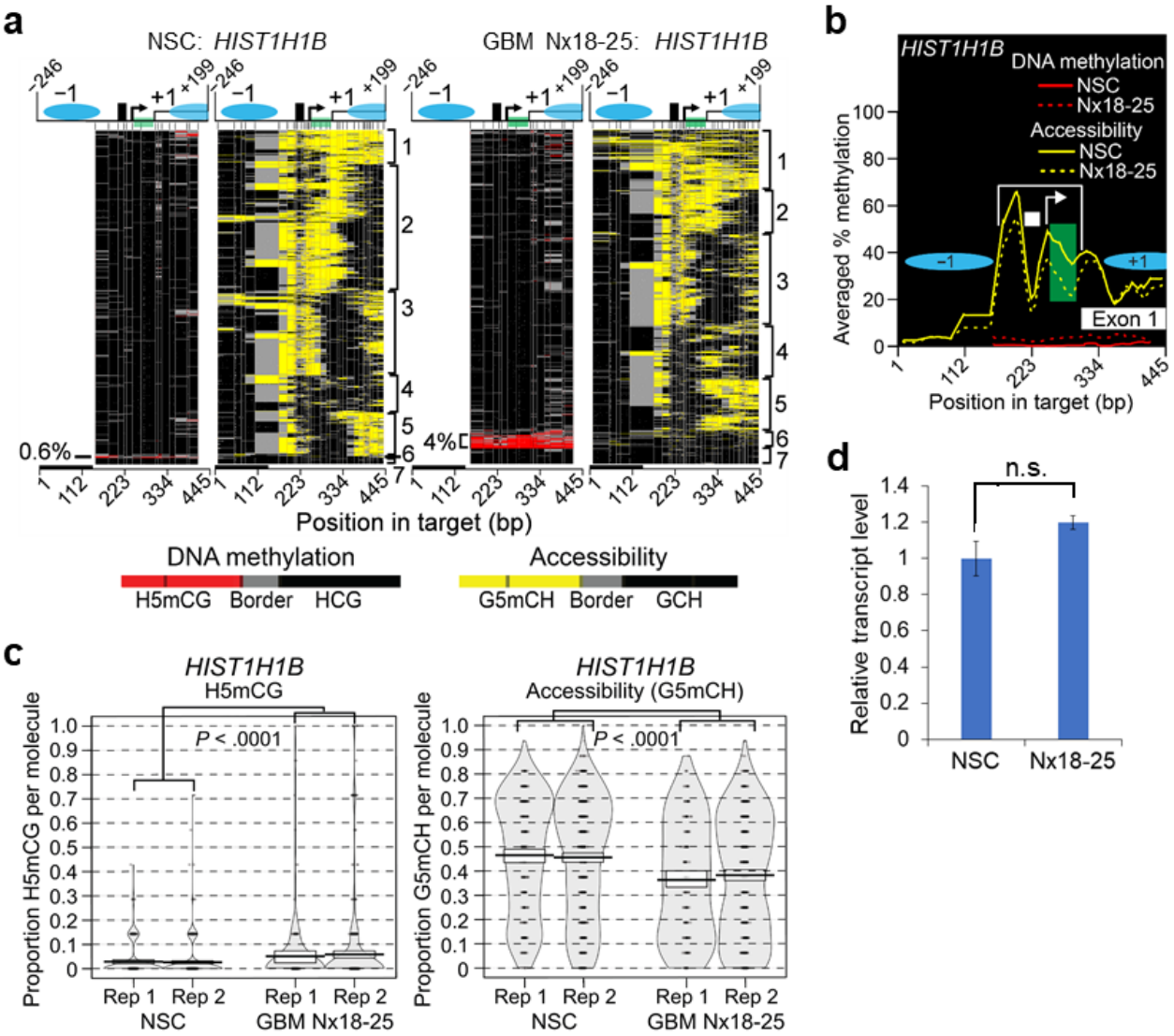
Fig. 4. Reproducible, regiospecific chromatin architectures and dynamics within murine bone marrow-derived monocytes. a-d, Methylscaper single-molecule methylation heatmaps

(top four rows of panels) and averaged levels of H5mCG and G5mCH (bottom row of panels).The plots derive from filtered EM-seq Hi-Fi CCS reads from the sequenced MAPit-FENGC libraries of ~620-nt targets from each of four mouse monocyte samples. Single molecules were aligned to the promoters of *Btk* on the chromosome X (**a**), *Pik3r3* (**b**), *Cxcr4* (**c**), as well as the *Mapk15* gene body (**d**). Black rectangles, strong TF footprints; orange rectangle in **c**, 36-bp zone of H5mCG in NFR of *Cxcr4* that is adjacent to but not within the neighboring TF binding site. Base pair position along each target is indicated at bottom; black bar on left represents 147 bp. Other symbols as defined above. *P* values (Supplementary Table 12) indicate that the levels of H5mCG and G5mCH in any one of four mouse monocyte samples were not significantly different from the other three.

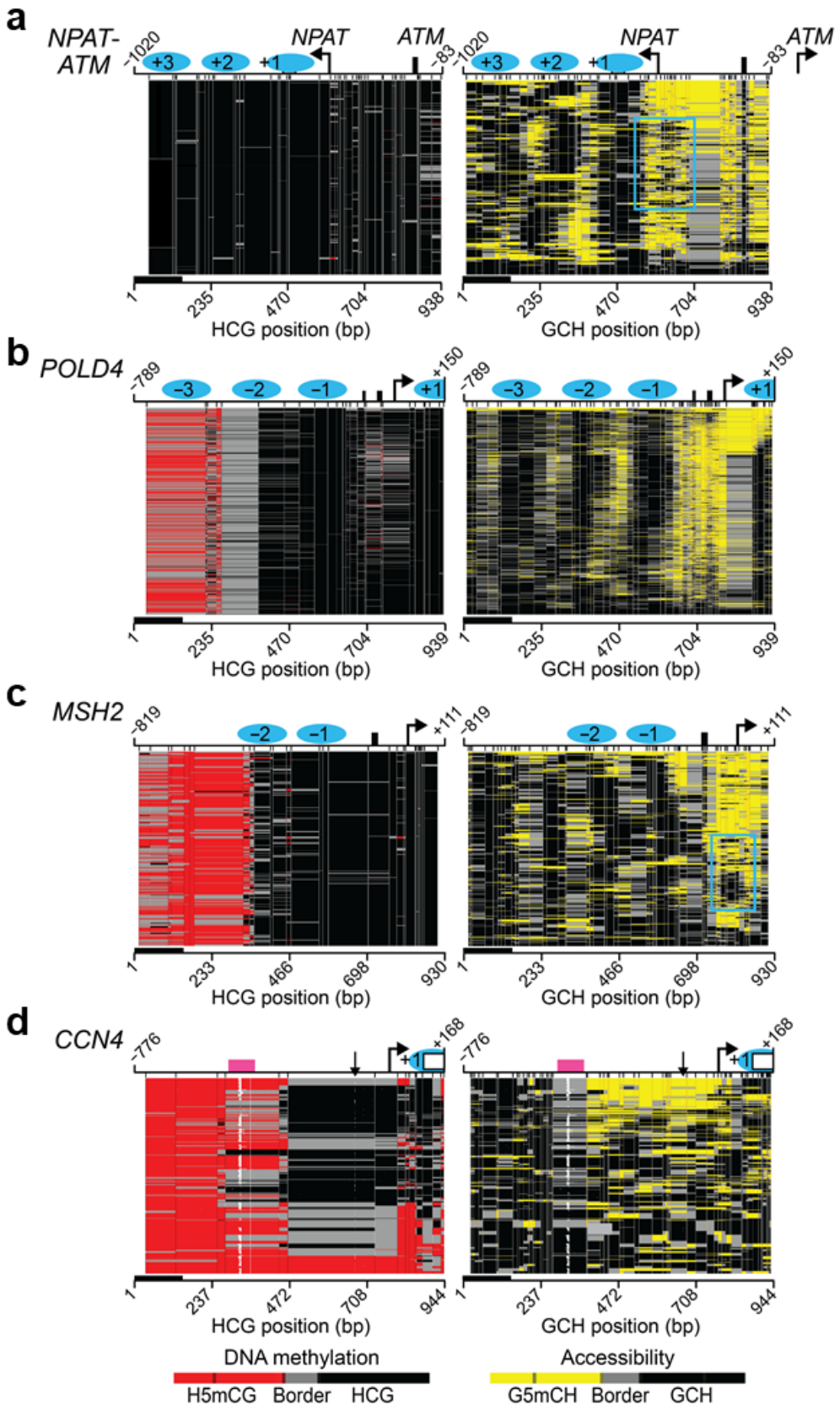
EXTENDED DATA FIGURES



Extended Data Fig. 1. *MSH5* promoter exemplifies a markedly accessible chromatin that controls for inter-sample and -molecule M.CviPI probing efficiency. a,b, Methylscaper plots of filtered CCS reads aligned to the *MSH5* promoter (−302 to +145) from MAPit-FENGC with the 119-target panel of two independent cultures of human non-cancerous NSC (132 and 241 reads; Supplementary Table 5) (**a**) and GBM Nx18-25 (170 and 192 reads) (**b**). Each HCG and GCH position (in base pairs) along the *MSH5* amplicon is indicated at the bottom of its respective panel (black bar represents 147 bp). A heterogeneous-sized footprint is marked by a black rectangle. Almost all *MSH5* promoter molecules across all 4 samples (733 of 735) had ≥10 methylated GCH sites, demonstrating high-level cell permeabilization and M.CviPI activity.

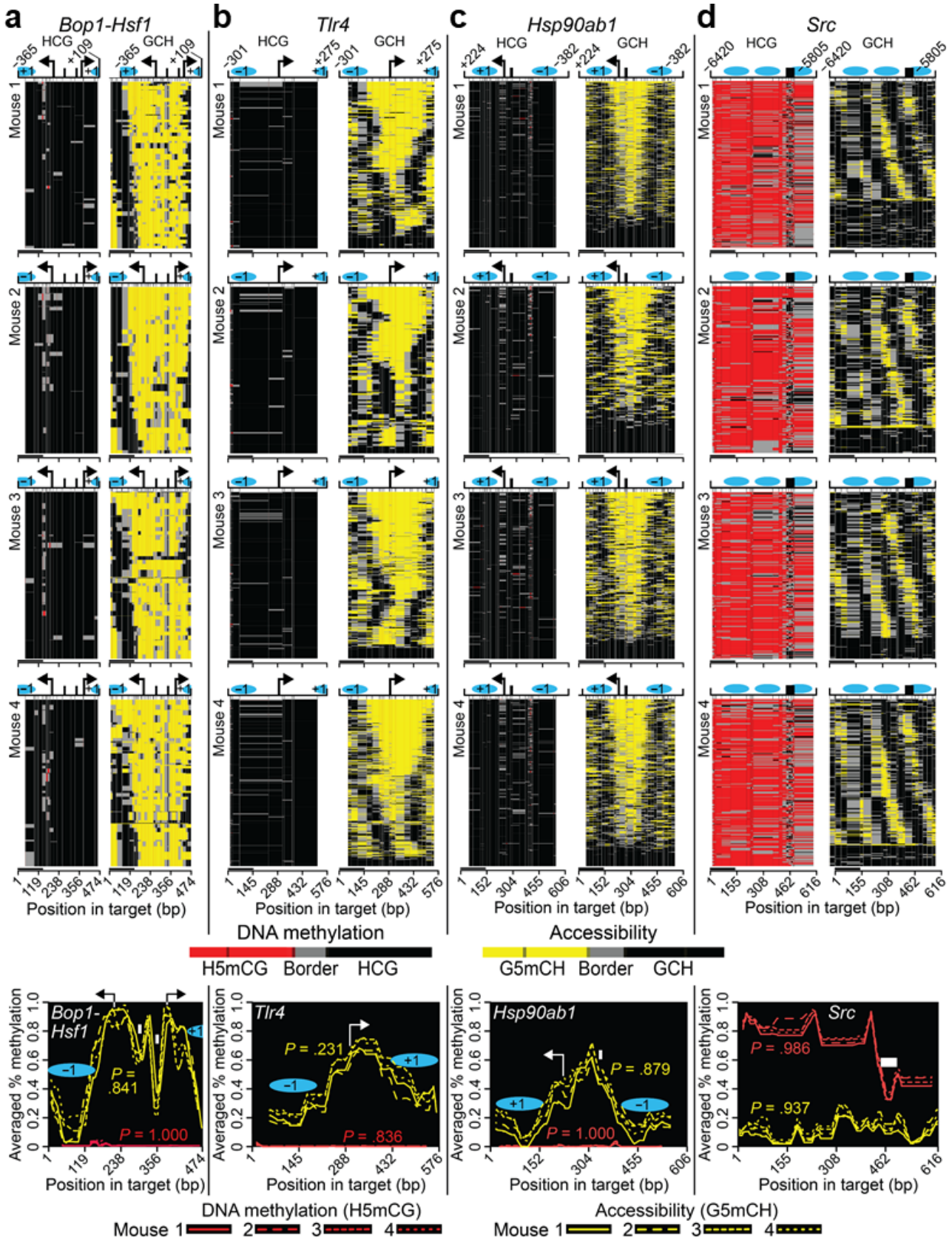


Extended Data Fig. 2. MAPit-FENGC with EM-seq detects minority subpopulations of molecules with differential epigenetic configurations. a-c, Methylscaper heatmaps (**a**), percentages of averaged H5mCG and G5mCH (**b**), and violins plots (**c**) derive from filtered CCS reads aligning to *HIST1H1B* from two human NSC (n = 2; 926 combined reads; Supplementary Table 5) and two GBM Nx18-25 cultures (n = 2; 953 combined reads). The bracketed area from −80 to +45 in **b** shows a significant decrease in accessibility in Nx18-25 relative to NSC that meets our 0.05 threshold (−0.074; Supplementary Table 6; P < 0.0001). **d,** Relative expression of *HIST1H1B* (n = 3, mean ± SD; n.s., not significant). The ~30-bp footprint upstream of the TSS (black rectangle) is consistent with binding of the preinitiation complex to the consensus TATA element[43]. The downstream ~50-bp footprint (green rectangle) may correspond to paused RNA polymerase II[43]. Other symbols as defined above.



Extended Data Fig. 3. Long-read MAPit-FENGC phases multiple epigenetic features. a-d, Methylscaper heatmaps of filtered EM-seq CCS reads aligning to *NPAT-ATM* (−1020 to −83

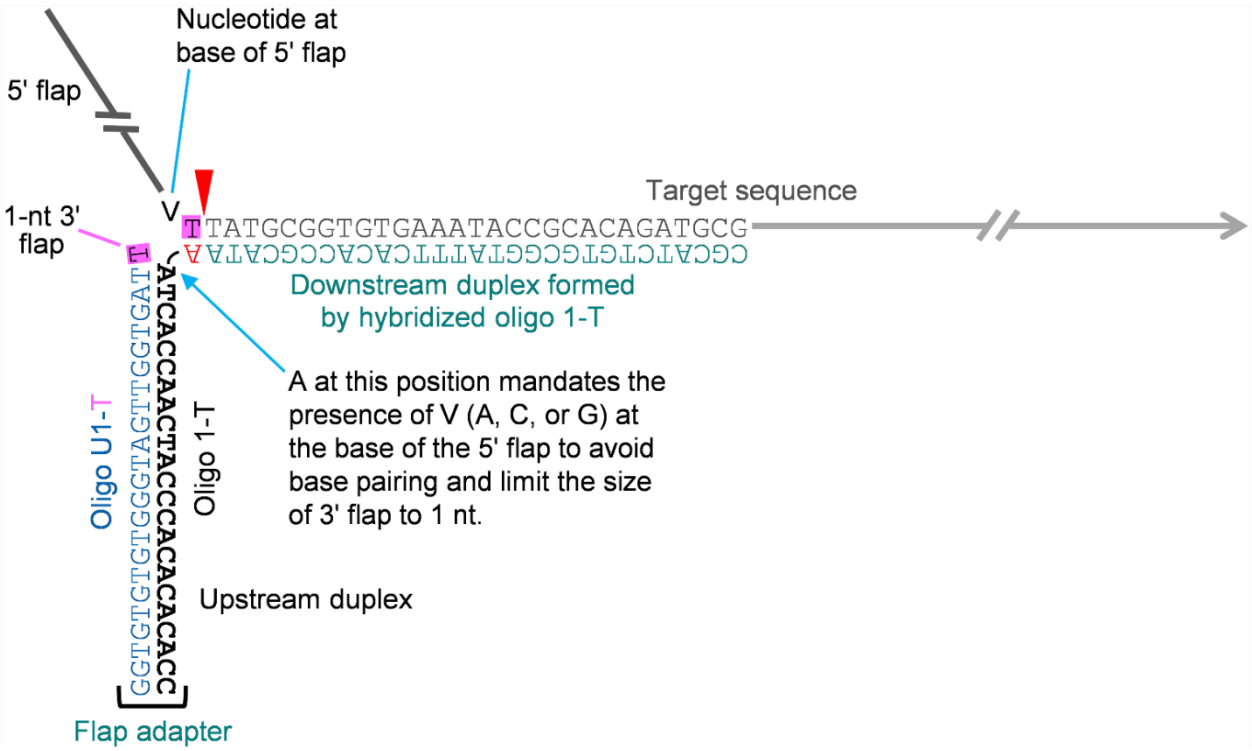
relative to *ATM* TSS; +592 to −346 relative to *NPAT* TSS) (**a**), *POL4D* (**b**), *MSH2* (**c**), and *CCN4* (**d**). All plots contain 1,000 molecules randomly selected from the combined GBM Nx18-25 replicates. Hierarchical clustering was focused on areas containing the longest NFRs. Relatively heterogeneous footprints (cyan rectangles) co-localizing with the *NPAT* TSS in the GCH panels in **a** and **d** are consistent with occupancy by the preinitiation complex[43]. In **d**, a known GA to G indel (dbSNP, rs548251181; Supplementary Table 7) is marked with a vertical arrow and plotted in white due to local non-alignment with the hg38 reference sequence. Similarly, in **d**, variable-length spans of non-alignment in a 66-bp A-rich sequence (88% A; pink rectangle) are plotted in white. Bent arrows, TSSs; cyan ovals, nucleosomes; black rectangles, TF footprints; white rectangles, protein coding sequences.



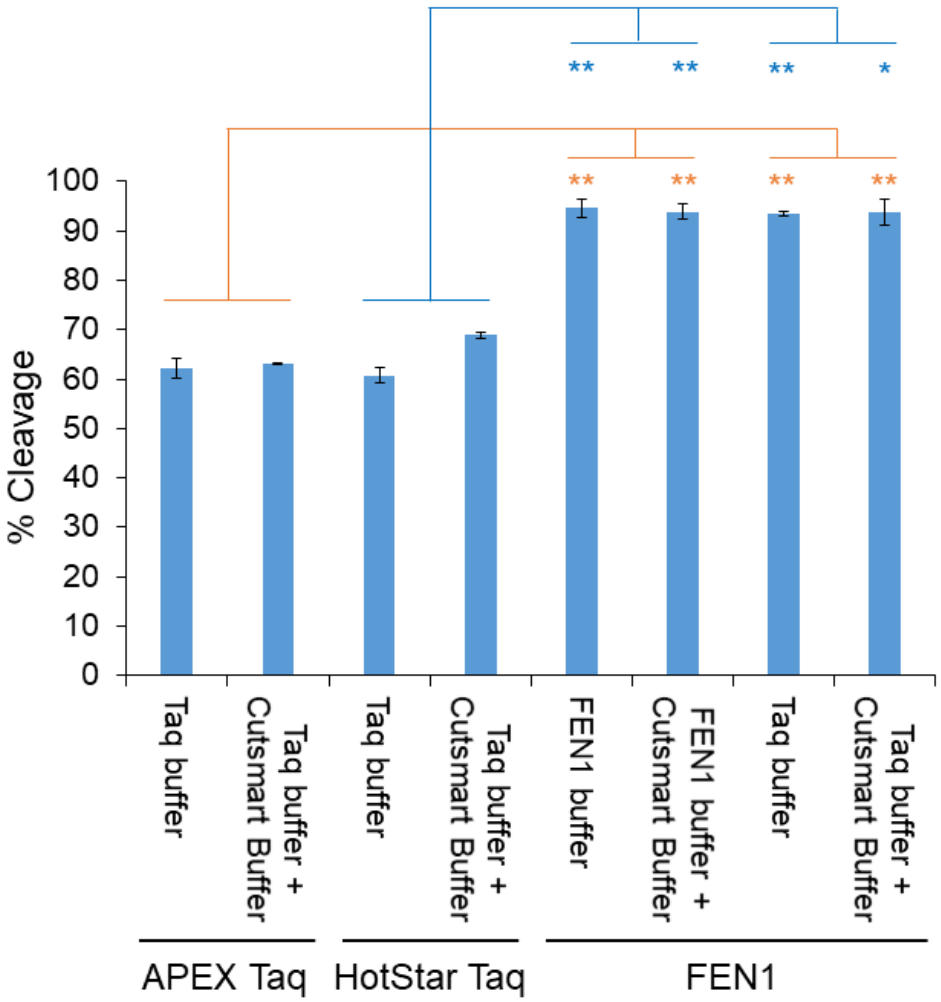
Extended Data Fig. 4. Reproducible, regiospecific chromatin architectures and dynamics revealed by long reads. a-d, Methylscaper single-molecule methylation heatmaps (top four rows

of panels) and averaged levels of H5mCG and G5mCH (bottom row of panels). The plots derive from filtered EM-seq Hi-Fi CCS reads from the sequenced MAPit-FENGC libraries of ~620-nt targets from each of four mouse monocyte samples. Single molecules were aligned to the promoters of divergently transcribed *Bop1* and *Hsf1* (**a**), *Tlr4* (**b**), *Hsp90ab1* (**c**), as well as −6,420 to −5,805 relative to the *Src* RefSeq TSS (**d**). Promoter coordinates in **a** are relative to the *Hsf1* TSS. Black rectangle in **d**, H5mCG footprint at sequence with strong CTCF binding site homology[44]. Base pair position along each target is indicated at bottom; black bar on left, 147 bp. Other symbols as defined above. Note detection of strikingly similar regiospecific architectures from all four mouse samples, despite the granularity of single-molecule heatmaps.

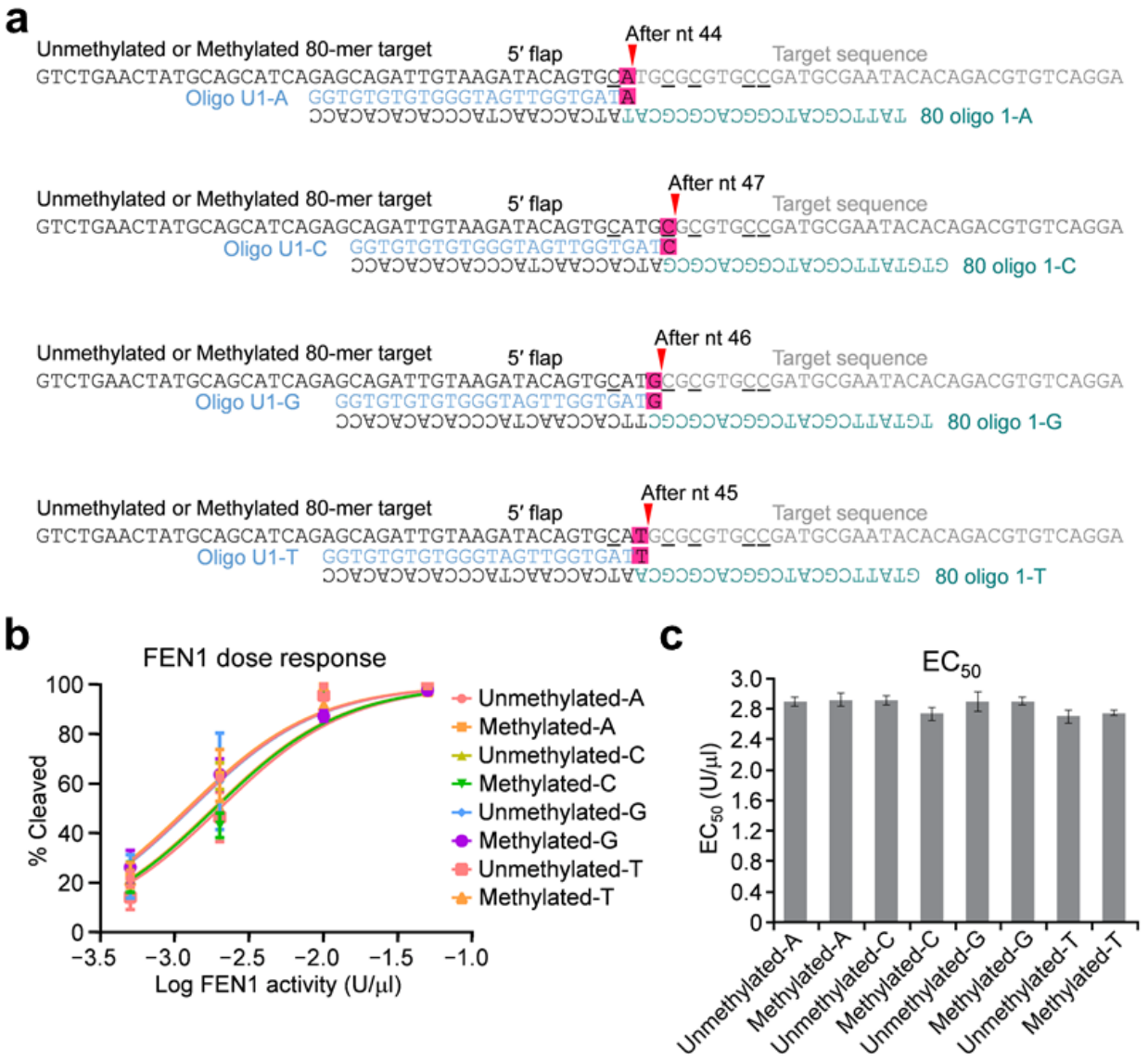
SUPPLEMENTARY FIGURES



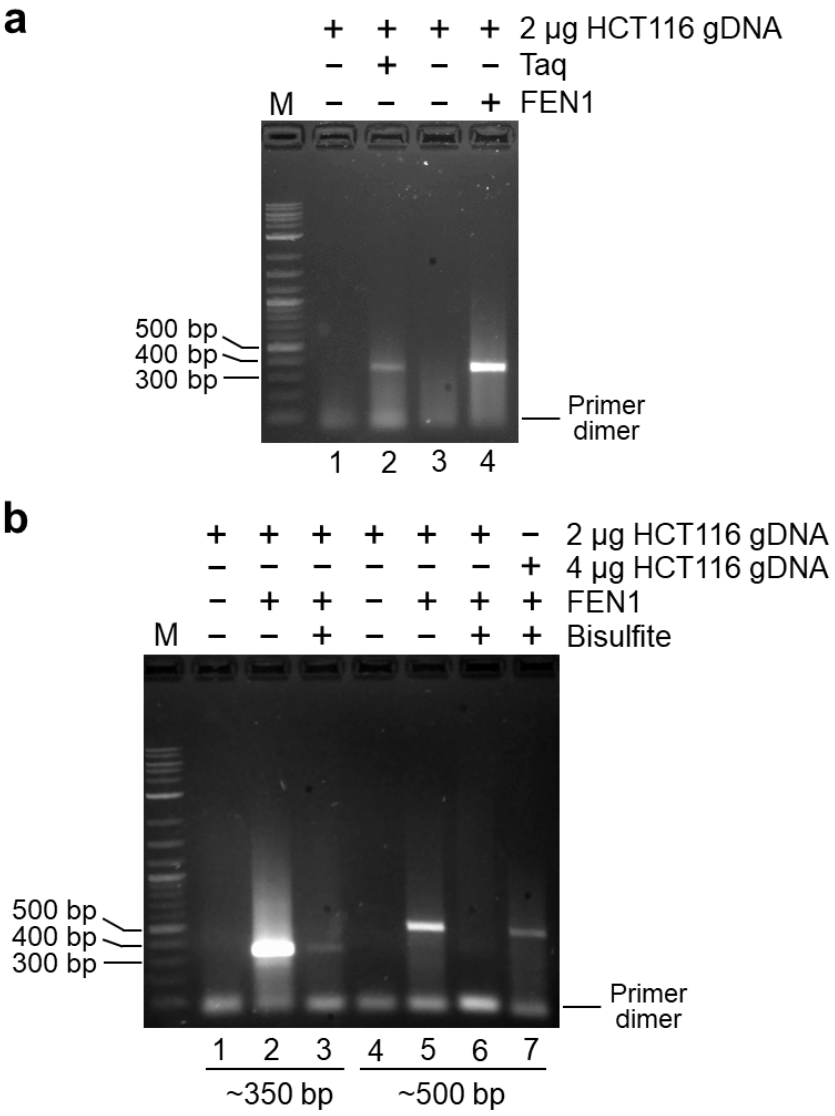
Supplementary Fig. 1. Flap adapter structure to direct efficient, precise 5' flap scission and target strand release. Shown is part of 200 oligo 1-T hybridized to the 5' end of its respective target sequence (200-T oligo, gray), forming a downstream duplex. In multiplexed FENGC, an oligo 2-T also hybridizes to the 3' end of each target sequence (not shown). The constant 3' tail of all oligos 1-T and 2-T binds to oligo U1-T, forming an upstream duplex and reconstituting a plurality of double flaps. The 'overlapped' nucleotides (highlighted pink) of the oligo U1-N 3'-terminus and immediately 5' of the scissile phosphodiester contain the same base. In the shown configuration using oligo U1-T, a T is present immediately 5' end of the cleavage site and base pairs with A (red type). The overlap displaces the 3' T of U1-T, creating an unpaired 3' flap. To ensure that 3' flaps are held at 1 nt in length when oligo U1-T is used, a V (A, C, or G) is placed at the base of the 5' flap, which does not pair with the indicated A oligo 1-T (cyan arrow). Coloration of each sequence matches that used in Fig. 1a. Oligo sequences are listed in Supplementary Table 1, Sheet 1.



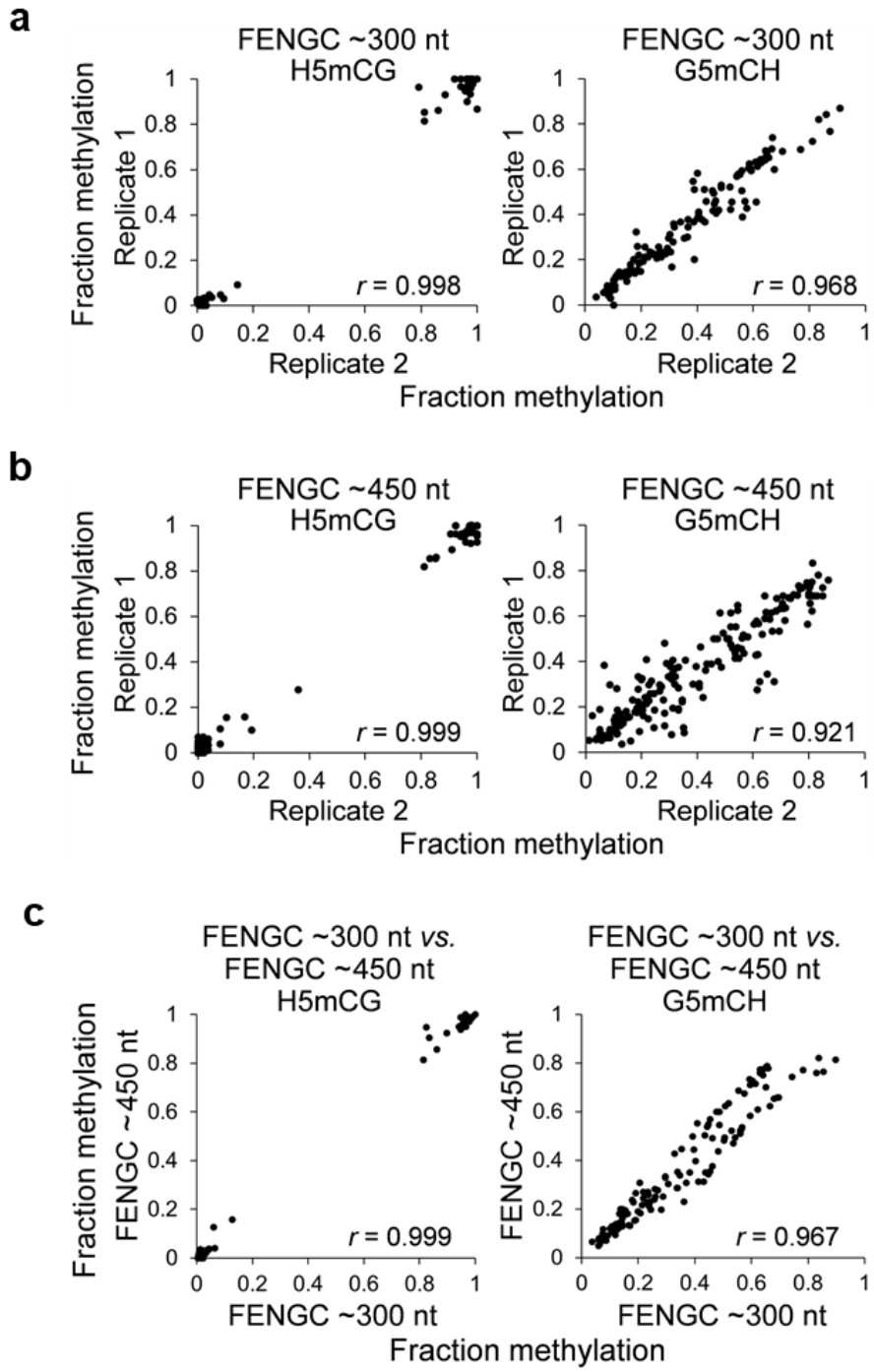
Supplementary Fig. 2. FEN cleavage efficiencies of Taq and FEN1. Percent cleavage of a 200-mer T proxy target assembled into a 129-nt 5' flap by the indicated FENs using the indicated buffers. FEN1 cut the 200-mer with higher efficiency than Taq under all tested conditions (n = 2, mean ± range; paired Student's t test, two sided, *P < .05, **P < .01). Oligos are described in Supplementary Table 1, Sheet 1.



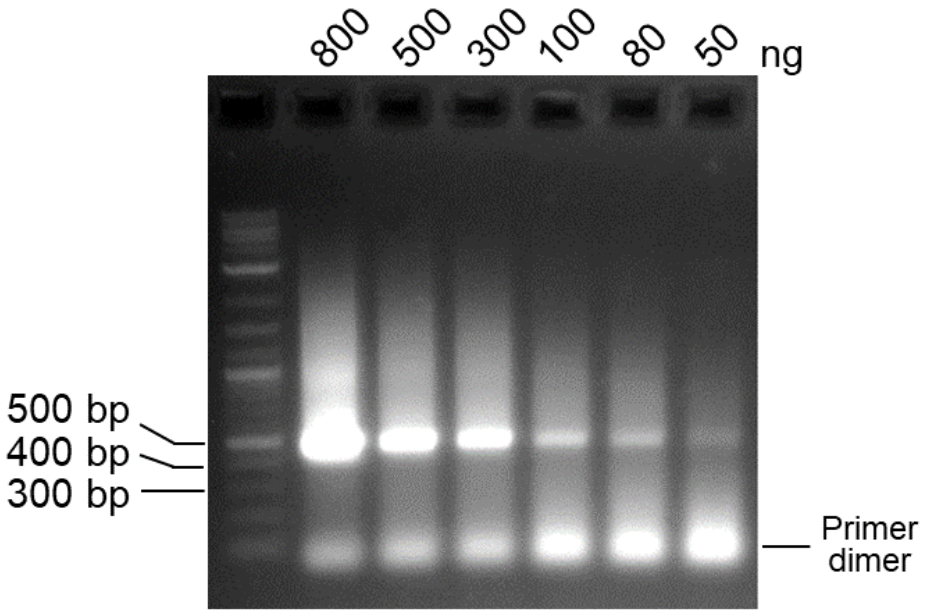
Supplementary Fig. 3. Equally efficient cleavage of double flaps with A, C, G, or T as 'overlap' in the absence and presence of nearby 5mC. a, Unmethylated and methylated double flap substrates used to test FEN1 5' flap cleavage efficiency. Two 80-mers with the same sequence were synthesized without or with 5mC at positions 43, 47, 49, 53, and 54 (underlined). Both 80-mers were reconstituted into four different double flaps by incubation with their respective flap adapter comprised of oligo U1-N and its corresponding 80 oligo 1-N. The 1-nt 'overlap' that displaces the 3'-terminal nucleotide of each oligo U1-N to create an unpaired 1-nt 3' flap is highlighted pink. Red arrowheads, FEN1 cleavage sites. All oligos are described in Supplementary Table 1, Sheet 1. **b,c,** Percentages cleaved of each 80-mer reconstituted into a -A, -C, -G, or -T double flap shown in **a** were quantified by qPCR and plotted as response *vs*. dose (**b**) and EC_{50} (**c**). A one-way ANOVA test of EC_{50} data in **c** showed no significant differences between any pair of conditions (n = 3, mean ± SD). Oligos are described in Supplementary Table 1, Sheet 1.



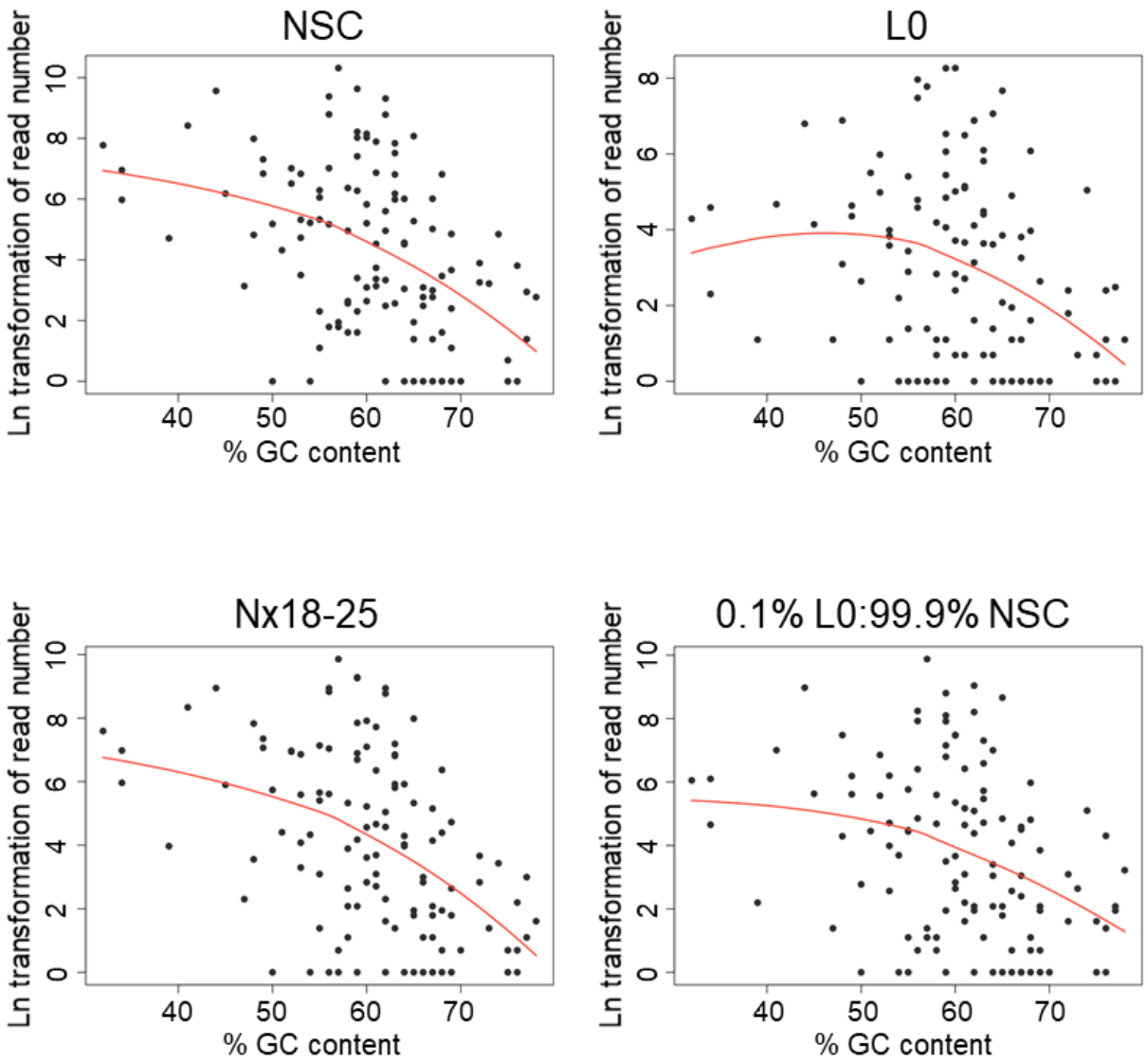
Supplementary Fig. 4. FEN activity-dependent enrichment of single-copy human sequences with and without bisulfite conversion. a,b, Human HCT116 gDNA (2 or 4 µg) was processed by FENGC with a 1-nt 3' flap as described in Figure 1a, using flap adapters 1-T and 2-T specific for 11-promoter panels that excise ~300-nt (**a**) as well as ~300-nt and ~450-nt (**b**) sequences in separate reactions. Both sets of reactions contained the same flap adapters 1-T, whereas the flap adapters 2-T in the ~450-nt reactions were shifted downstream (Supplementary Table 1, Sheets 2 and 3; Supplementary Table 2, Sheets 1 and 2). Reactions also contained either no enzyme, 2 U Taq in 1x Taq buffer, or 32 U FEN1 in 1x FEN1 buffer as indicated. Enriched targets (post-exonuclease treatment and purification) were amplified by standard PCR in **a** or standard PCR (lanes 1, 2, 4, and 5) or BS-PCR (lanes 3, 6, and 7) in **b**. The crude amplification products, expected to be ~350 bp and ~500 bp, were electrophoresed on a 1% agarose gel. The results shown in both gels are representative of two independent experiments using the same HCT116 gDNA.



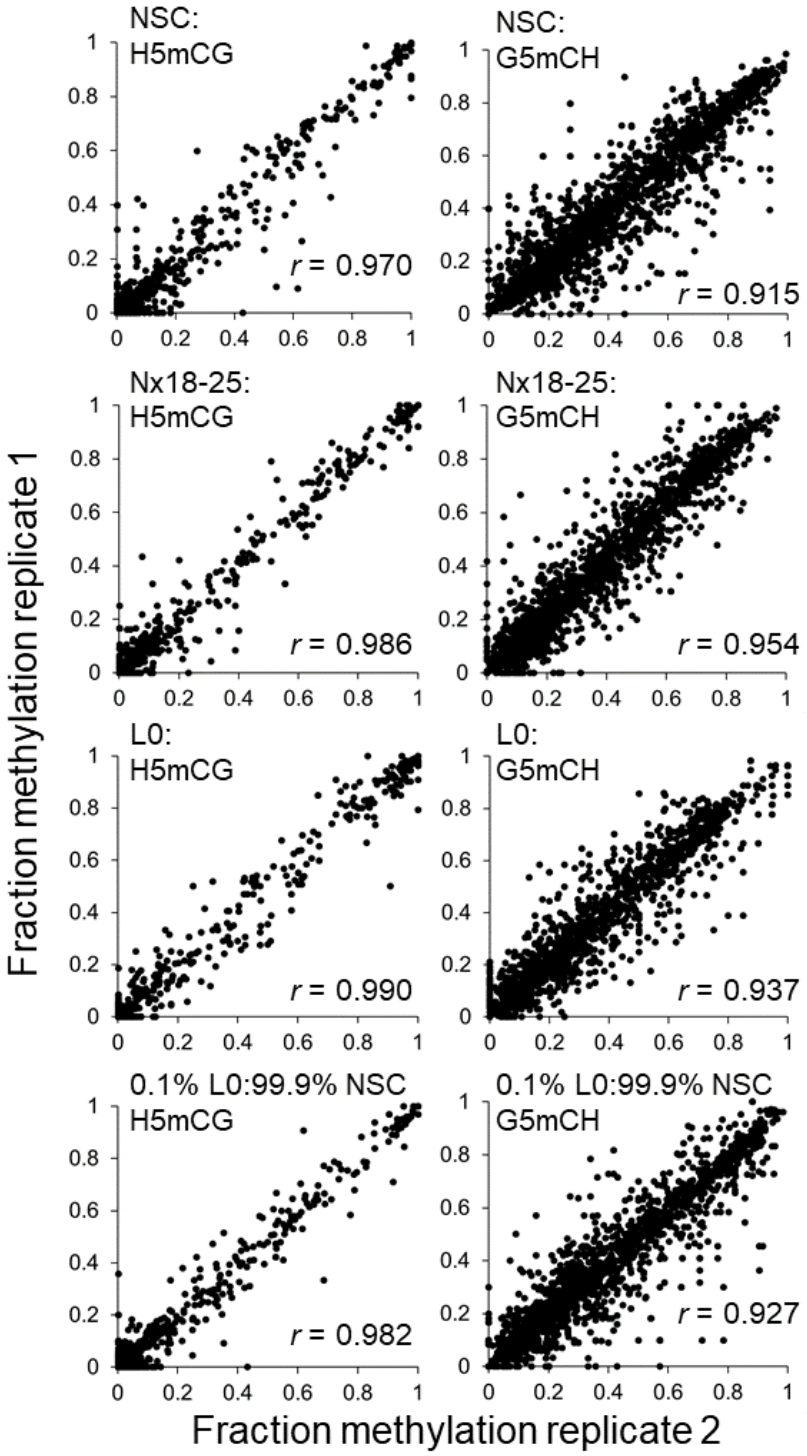
Supplementary Fig. 5. High reproducibility of MAPit-FENGC with BS-seq of different-size targets. a,b, Pearson correlation plots of H5mCG and G5mCH levels between two biological replicates of MAPit-FENGC, processing two independent cultures of M.CviPI-treated human GBM L0 using FEN1 to enrich for ~300-nt (**a**) and ~450 nt targets (**b**) from the 11-promoter panel (Supplementary Table 1, Sheets 2 and 3, respectively). **c,** Correlation plots of H5mCG and G5mCH levels between sequences in common between the ~300-nt and ~450 nt targets. Hi-Fi BS-seq CCS PacBio Sequel reads (≥5 sequencing passes with ≥95% conversion) from the two libraries for each condition were aligned to the reference sequences. Methylation levels are from 6 promoters with >40 aligned reads in each biological replicate (≥150x in combined replicates) in both the ~300-nt and ~450-nt targets (Supplementary Table 3).



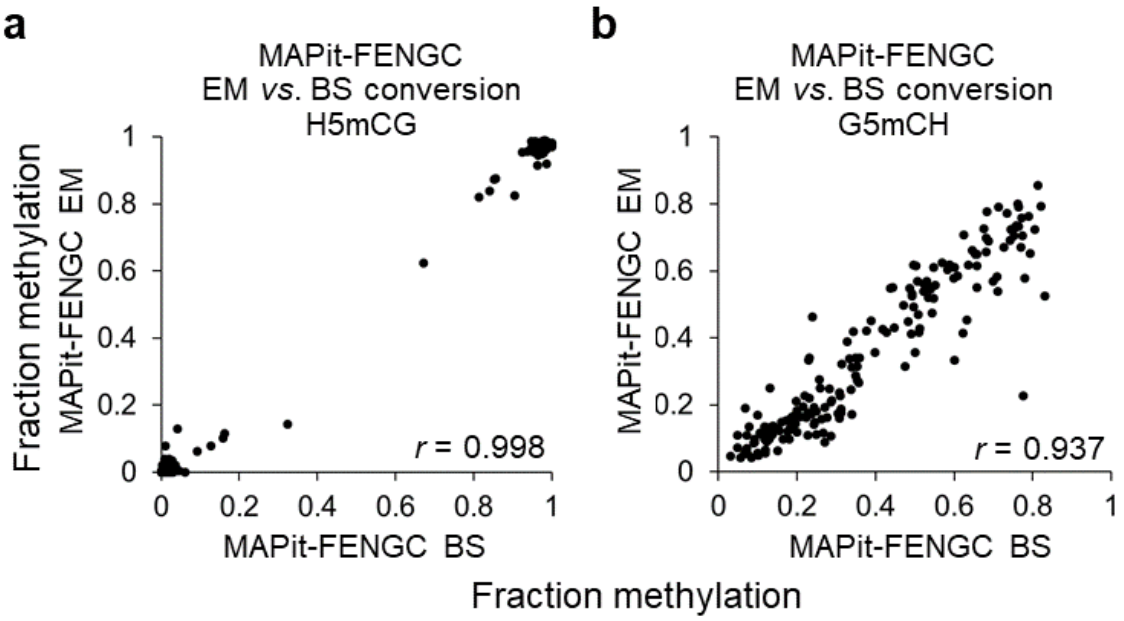
Supplementary Fig. 6. MAPit-FENGC with EM-PCR. Agarose gel electrophoresis of FENGC amplicon yields from the indicated amounts of M.CviPI-treated GBM L0 gDNA following fragmentation with mild sonication. The 5' flaps of 119 promoter sequences of ~450 nt (Supplementary Table 1, Sheet 3) were cleaved by 32 U FEN1 within double flap structures with 3' flaps comprised of one T. Crude EM-PCR amplification products were electrophoresed without purification. For library construction, the primer dimer is readily removed by a single round of size selection with AMPure beads (data not shown). The results shown are representative of two independent experiments.



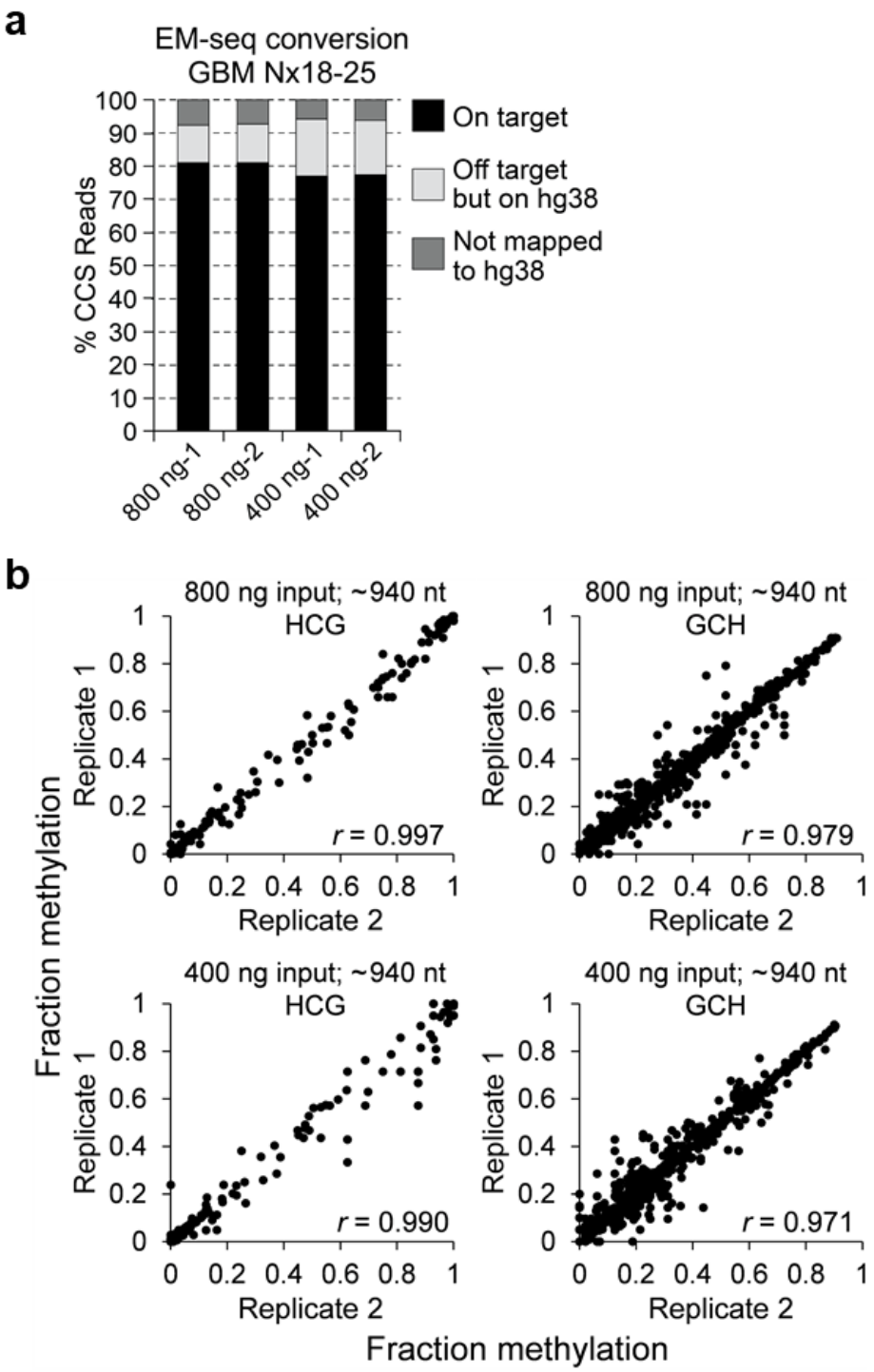
Supplementary Fig. 7. GC content exhibits a moderate, negative correlation with MAPit-FENGC CCS read number. LOESS fit curves of the natural logarithm transformation of the filtered read number in the combined biological replicates plus 1 *versus* GC content for each of the 119 targets of ~450 nt for the indicated samples (Supplementary Table 5). Hi-Fi CCS reads from the 119-target MAPit-FENGC libraries were filtered for ≥95% HCH to HTH conversion and ≥95% coverage of the length of each reference sequence.



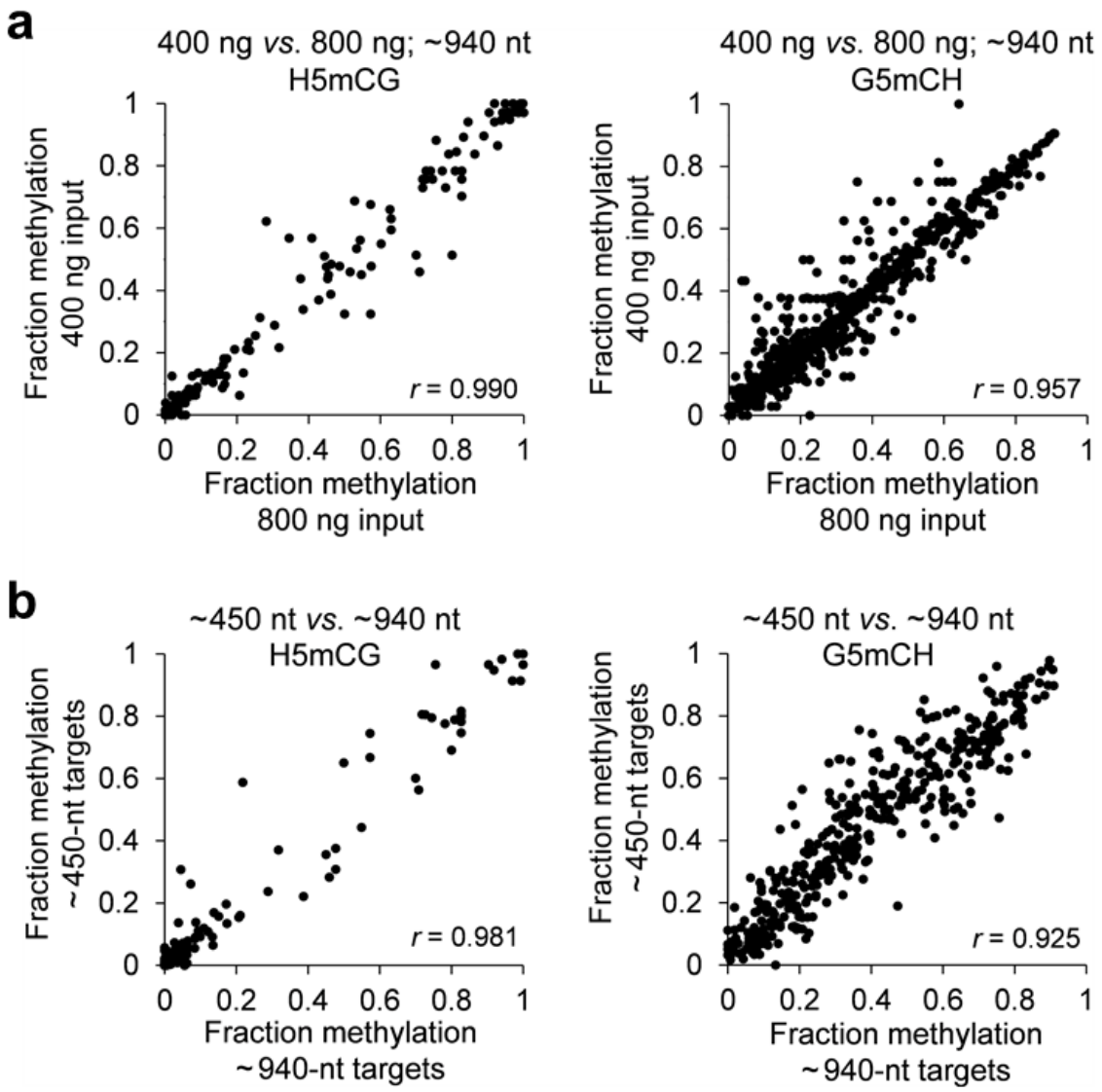
Supplementary Fig. 8. High reproducibility of MAPit-FENGC utilizing EM-PCR of a 119-promoter panel. Correlation plots between the fraction of methylation of each HCG and GCH site in the independent biological replicates of the indicated MAPit-FENGC samples enriched for the 119 human promoter sequences of ~450 nt. Only targets with ≥10 Hi-Fi CCS reads aligned and filtered for ≥95% conversion and ≥95% coverage of reference sequence length in both replicates were included (Supplementary Table 5). This equated to 64 targets for NSC, 52 targets for L0, 64 targets for Nx18-25, and 63 targets for the 0.1% L0 gDNA:99.9% NSC gDNA mixtures.



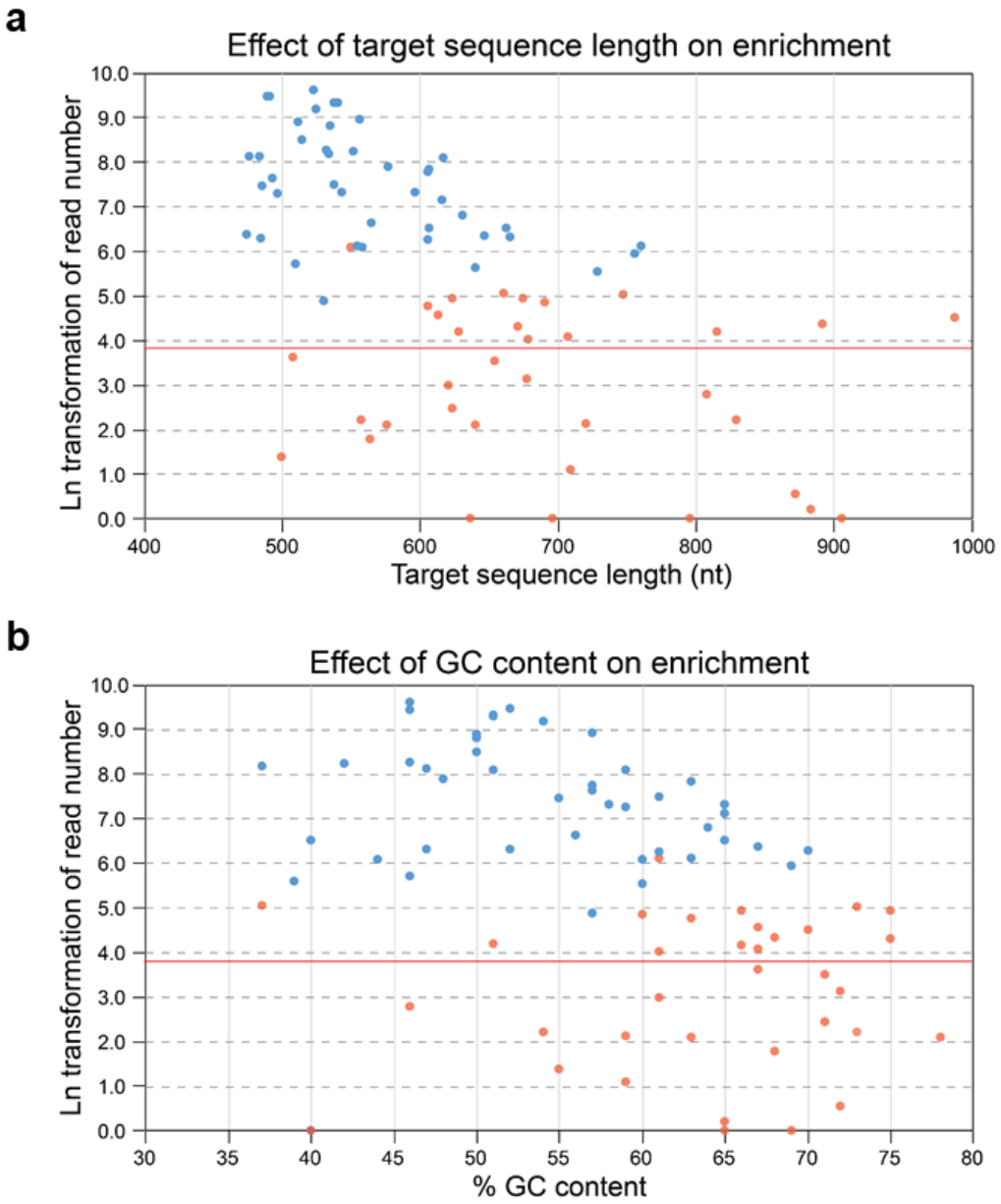
Supplementary Fig. 9. Strong correlation between MAPit-FENGC using EM and bisulfite conversion. a,b, Correlation plots of the fractions of methylation at each HCG (**a**) and GCH site (**b**). Data plotted for the 6 targets of ~450 nt with ≥36 aligned and filtered CCS reads in the combined GBM L0 duplicates for EM conversion (Supplementary Table 5) and bisulfite conversion (Supplementary Table 3).



Supplementary Fig. 10. Performance of MAPit-FENGC for captured ~940-nt targets with two different gDNA amounts. a, Proportions of Hi-Fi CCS reads for MAPit-FENGC of 45 promoter sequences (Supplementary Table 1, Sheet 4; Supplementary Table 2, Sheet 3) mapping to or off the human genome for the indicated sequenced libraries (Supplementary Table 8). **b,** Correlation plots of the fraction of methylation of each HCG and GCH motif in the independent biological replicates of the indicated MAPit-FENGC samples. Only targets with ≥14 Hi-Fi CCS reads aligned and filtered for ≥95% conversion and ≥95% coverage of reference sequence length in both replicates were included (Supplementary Table 9). This equated to 13 and 12 targets for inputs of 800 ng and 400 ng, respectively.



Supplementary Fig. 11. Performance of MAPit-FENGC for captured ~940-nt targets with two different gDNA amounts and comparison to ~450-nt targets. a,b, Correlation plots of the fraction HCG and GCH methylation in the combined biological replicates of ~940-nt targets for 400 *vs*. 800 ng input gDNA (**a**) and for sequences in common between ~940-nt *vs*. ~450-nt targets (**b**) from M.CviPI-treated Nx18-25. After filtering HiFi CCS reads for ≥95% conversion and ≥95% coverage of reference sequence length (Supplementary Table 9), methylation levels of 13 targets with ≥15x coverage in the combined biological replicates were plotted.



Supplementary Fig. 12. Effect of target sequence length and GC content on performance of MAPit-FENGC with EM-seq on primary mouse monocytes. a,b, Natural logarithm transformation of filtered CCS read number + 1 *versus* (**a**) target length and (**b**) % GC content for the 78-target murine gene panel enriched from gDNA isolated from bone marrow-derived monocytes probed with M.CviPI (*n* = 4 female mice). In (**a**) and (**b**), blue and red data points represent 43 targets included and 35 excluded, respectively, from the statistical analysis of Supplementary Table 12. Target loci were excluded because their read number fell below the 100-read per sample threshold (red horizontal line) or presence of too many duplicates, typically observed for targets >760 nt or >70% GC content.

SUPPLEMENTARY TABLES

Supplementary Table 1. FENGC oligonucleotides used in this study.

Supplementary Table 2. Characteristics of FENGC targets used in this study.

Supplementary Table 3. CCS reads aligned to 11 human targets of ~300 nt *vs*. ~450 nt used in FENGC assay development with standard PCR *vs*. BS-PCR.

Supplementary Table 4. CCS reads on- and off-target for 119 ~450-nt human targets in MAPit-FENGC.

Supplementary Table 5. Filtered CCS reads aligned to 119 ~450-nt human targets in MAPit-FENGC using EM-seq.

Supplementary Table 6. Differential H5mCG and G5mCH in NSC and GBM Nx18-25 among ~ 450-nt targets determined by MAPit-FENGC using EM-seq.

Supplementary Table 7. SNPs and indels detected by FENGC using standard PCR of 119 ~450-nt human targets.

Supplementary Table 8. CCS reads on- and off-target for 45 ~940-nt human targets in MAPit-FENGC using EM-seq.

Supplementary Table 9. Filtered CCS reads aligned to 45 ~940-nt human targets in MAPit-FENGC using EM-seq.

Supplementary Table 10. CCS reads on- and off-target for 78 ~620-nt mouse targets in MAPit-FENGC using EM-seq.

Supplementary Table 11. Filtered CCS reads aligned to 78 ~620-nt mouse targets in MAPit-FENGC using EM-seq.

Supplementary Table 12. Differential H5mCG and G5mCH determined by MAPit-FENGC using EM-seq of ~620-nt mouse targets.

Supplementary Table 13. FENGC oligonucleotide concentrations for different numbers of targets.

Supplementary Table 1: FENGC oligonucleotides used in this study; assay development.

Oligo name	Oligo stock name	Oligo sequence	Notes	Figure(s)	Reference
Oligo U1	MZO3782	GGTGTGTGTGGGTAGTTGGTGAT	Oligo does not form a 1-nt 3' flap when base paired with an oligo 1-N or an oligo 2-N annealed to a sequence of interest. Oligo also serves as universal primer.	Figs. 1-4 (as primer only); Extended Data Figs. 1-4 (as primer only); Supplementary Fig. 2 (in flap adapter); Supplementary Figs. 4-12	Varley et al., ref. 13; Nabilsi et al., ref. 14
Oligo U1-A	NNO3458	GGTGTGTGTGGGTAGTTGGTGATA	Nucleotide in red forms the 1-nt 3' flap when base paired with an oligo 1-A or an oligo 2-A annealed to a sequence of interest	Fig. 1a	This study
Oligo U1-C	NNO3453	GGTGTGTGTGGGTAGTTGGTGATC	Nucleotide in red forms the 1-nt 3' flap when base paired with an oligo 1-C or an oligo 2-C annealed to a sequence of interest	Fig. 1a	This study
Oligo U1-G	NNO3451	GGTGTGTGTGGGTAGTTGGTGATG	Nucleotide in red forms the 1-nt 3' flap when base paired with an oligo 1-G or an oligo 2-G annealed to a sequence of interest	Fig. 1a	This study
Oligo U1-T	NNO3449	GGTGTGTGTGGGTAGTTGGTGATT	Nucleotide in red forms the 1-nt 3' flap when base paired with an oligo 1-T or an oligo 2-T annealed to a sequence of interest	Figs. 1-4; Extended Data Figs. 1-4; Supplementary Figs. 4-12	This study
Oligo U2	MZO3794	/5Phos/AGAGAATGAGGAAGGTGG*G*G*T*A*G/3SpC3/	/5Phos/, 5' Phosphate; *, Phosphorothioate Bonds; /3SpC3/, 3' C3 Spacer (C, carbon)	Figs. 1-4; Extended Data Figs. 1-4; Supplementary Figs. 4-12	Varley et al., ref. 13; Nabilsi et al., ref. 14
U2 primer	MZO3783	CTACCCCACCTTCCTCATTCTCT	Used in all reactions using PCR.	Figs. 1-4; Extended Data Figs. 1-4; Supplementary Figs. 4-12	Varley et al., ref. 13; Nabilsi et al., ref. 14
200-mer T target	MZO3885	CGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAG CGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGA GAGTGCACCTTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCA GGCGCCATTCGCCATTCAGG	200-mer proxy target with T residue in red located in "overlap" immediately 5' of FEN cleavage site. Identical sequence as 200-nt target strand in starting pGEM-3Z/601b in which nt 2453 was mutated by site-directed mutagenesis to T.	Supplementary Figs. 1,2	This study
200 oligo 1-T	MZO3878	CGCATCTGTGCGGTATTTCACACCGCATAA**ATCACCAACTACCCACACACACC**	Forms flap adapter with oligo U1-T (NNO3449) that reconstitutes a 5' flap and 1-nt 3' flap upon hybridization of unbolded nucleotides and nucleotide in red to 200-mer T target (MZO3885) and pGEM-3Z/601b-T plasmid; nucleotide in red also hybridizes immediately 5' of FEN cleavage site; bolded nucleotides are complementary to oligo U1-T.	Supplementary Figs. 1,2	This study
Unmethylated 80-mer target	MZO4076	GTCTGAACTATGCAGCATCAGAGCAGATTGTAAGATACAGTGCATGCGCGTGCCGATGCG AATACACAGACGTGTCAGGA	Unmethylated C underlined	Supplementary Fig. 3	This study
Methylated 80-mer target	MZO4077	GTCTGAACTATGCAGCATCAGAGCAGATTGTAAGATACAGTG5mCATG5mCG5mCGTG5m C5mCGATGCGAATACACAGACGTGTCAGGA	5mC (5-methylcytosine) underlined	Supplementary Fig. 3	This study
80 oligo 1-A	MZO4080	TATTCGCATCGGCACGCGCAT**ATCACCAACTACCCACACACACC**	Non-bolded nucleotides are complementary to the 5' end of target sequence; nucleotide in red is immediately 5' of FEN cleavage site; bolded nucleotides are complementary to oligo U1	Supplementary Fig. 3	This study
80 oligo 1-C	MZO4083	GTGTATTCGCATCGGCACGCG**ATCACCAACTACCCACACACACC**	Non-bolded nucleotides are complementary to the 5' end of target sequence; nucleotide in red is immediately 5' of FEN cleavage site; bolded nucleotides are complementary to oligo U1	Supplementary Fig. 3	This study
80 oligo 1-G	MZO4082	TGTATTCGCATCGGCACGCGC**ATCACCAACTACCCACACACACC**	Non-bolded nucleotides are complementary to the 5' end of target sequence; nucleotide in red is immediately 5' of FEN cleavage site; bolded nucleotides are complementary to oligo U1	Supplementary Fig. 3	This study
80 oligo 1-T	MZO4081	GTATTCGCATCGGCACGCGCA**ATCACCAACTACCCACACACACC**	Non-bolded nucleotides are complementary to the 5' end of target sequence; nucleotide in red is immediately 5' of FEN cleavage site; bolded nucleotides are complementary to oligo U1	Supplementary Fig. 3	This study
80 qPCR-F primer	MZO4078	GTCTGAACTATGCAGCATCAGAGC	Forward qPCR primer to assess cutting of unmethylated and methylated 80-nt oligo substrate.	Supplementary Fig. 3	This study
80 qPCR-R primer	MZO4079	TCCTGACACGTCTGTGTATTCG	Reverse qPCR primer to assess cutting of unmethylated and methylated 80-nt oligo substrate.	Supplementary Fig. 3	This study

Supplementary Table 1: FENGC oligonucleotides used in this study; human ~300-nt targets.

Oligo name	Oligo sequence	Notes	
300-nt-EPM2AIP1-oligo 1-T	GCGCCACTACGAGGCTGAGCACGAATA**ATCACCAACTACCCACACACACC**	1. Olgos designed to capture target sequences of ~300 nt in length.	Same as 450-nt-oligo 1-T
300-nt-EXO1-oligo 1-T	AGCCCCAGTTCTGGAGGCTGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-HSPA5-oligo 1-T	GTCGGAGTAGGTCCAGCAGGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-MLH1-oligo 1-T	GGAAACGTCTAGATGCTCAACGGA**ATCACCAACTACCCACACACACC**	2. Non-bolded nucleotides are complementary to the 5' end of target sequence.	Same as 450-nt-oligo 1-T
300-nt-MLH3-oligo 1-T	CTGGCCTATGGTAGGCACCGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-MSH2-oligo 1-T	AAACTAGCCTTCCTCCTCCTCCCAGTA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-MSH3-oligo 1-T	CCGACTCCCGGCGAGAGGA**ATCACCAACTACCCACACACACC**	3. Nucleotide in red is complementary to the nucleotide immediately 5' of cleavage site.	Same as 450-nt-oligo 1-T
300-nt-MSH6-oligo 1-T	GATAACCGGGGGGACCCGAAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-PMS1-oligo 1-T	GCTTTGCCTTCCAGCTGAGATGTAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-PMS2-oligo 1-T	CTCTCCGTGTGGAGCTTCCCA**ATCACCAACTACCCACACACACC**	4. Bolded nucleotides are complementary to oligo U1-T.	Same as 450-nt-oligo 1-T
300-nt-TDG-oligo 1-T	CTGTGTTGCATTATCAGGTAATGAGATGTGAGGGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 1-T
300-nt-EPM2AIP1-oligo 2-T	AACCGGGCTTGTGTGCCTCTGCTGA**ATCACCAACTACCCACACACACC**	5. Used in Supplementary Figs. 4-5; Supplementary Tables 2-3.	
300-nt-EXO1-oligo 2-T	GCCCCGCCCAGAGGATGTGA**ATCACCAACTACCCACACACACC**		
300-nt-HSPA5-oligo 2-T	CGTCGTGGAGGCCGCCGA**ATCACCAACTACCCACACACACC**		
300-nt-MLH1-oligo 2-T	AGGAGAAGGCCTGACTGGCACGTCA**ATCACCAACTACCCACACACACC**		
300-nt-MLH3-oligo 2-T	CAAGGATCCGCATCCCAGGGTTCA**ATCACCAACTACCCACACACACC**		
300-nt-MSH2-oligo 2-T	GTCTGCCGCGGCAACCAATCA**ATCACCAACTACCCACACACACC**		
300-nt-MSH3-oligo 2-T	CGTCTGCGGGCGCGAGCA**ATCACCAACTACCCACACACACC**		
300-nt-MSH6-oligo 2-T	GGGCGGTGCGCGCCA**ATCACCAACTACCCACACACACC**		
300-nt-PMS1-oligo 2-T	GTGACGTGCGGCCGCGAGA**ATCACCAACTACCCACACACACC**		
300-nt-PMS2-oligo 2-T	GGCTGCTTTCGACGTTGTGCTCCA**ATCACCAACTACCCACACACACC**		
300-nt-TDG-oligo 2-T	ACGCGTACTGGCCGTCACGTGA**ATCACCAACTACCCACACACACC**		
300-nt-EPM2AIP1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGTGATCTGGCGCAGAGCGGAGG	1. Oligos designed to capture target sequences of ~300 nt in length.	
300-nt-EXO1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGTCAACGCGTATCCCGCAACCC		
300-nt-HSPA5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTGGTGAAGGCCCTGCTCGTTGG		
300-nt-MLH1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGAACCCGGCGGCCTC	2. Non-bolded nucleotides are complementary to the 3' end of target sequence.	
300-nt-MLH3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTGAACCGGTACCTACAGTCTCAATCCTTATTCC		
300-nt-MSH2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATAAGCAGACGCGGCCCAGC		
300-nt-MSH3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCCGCGACCCTGCGTG	3. Nucleotide in red is complementary to 3' end of target sequence, i.e., immediately 5' of cleavage site.	
300-nt-MSH6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCCCCGCCCCCGG		
300-nt-PMS1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGAAGGGGCGGGTTGGGG		
300-nt-PMS2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCCTTTCCGGGCGGGGC	4. Bolded nucleotides are complementary to oligo U2.	
300-nt-TDG-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGTAGGACAGCCCAATCACGCTCACC		
		5. Used in Supplementary Figs. 4-5; Supplementary Tables 2-3.	

Supplementary Table 1: FENGC oligonucleotides used in this study; human ~450-nt targets.

Oligo name	Oligo sequence	Notes
450-nt-ALKBH1-oligo 1-T	CAGCGTGGTGGCTGAATTAGGAGA**ATCACCAACTACCCACACACACC**	1. Olgos designed to capture target sequences of ~450 nt in length.
450-nt-ALKBH2-oligo 1-T	AGGAGTTCTGCAGTCGGATATCGTGGA**ATCACCAACTACCCACACACACC**	
450-nt-ALKBH3-oligo 1-T	GCATCGGCATCTTTGTGGCAGAAAGA**ATCACCAACTACCCACACACACC**	
450-nt-APEX1-oligo 1-T	GGTCTCCGCTCGCGGGA**ATCACCAACTACCCACACACACC**	2. Non-bolded nucleotides are complementary to the 5' end of target sequence.
450-nt-APEX2-oligo 1-T	GCCACTTTTTAAGCGACGGAACCCTGA**ATCACCAACTACCCACACACACC**	
450-nt-ATM-oligo 1-T	GTGGTTCCTGCTGTGGTTTTGATCA**ATCACCAACTACCCACACACACC**	
450-nt-ATRX-oligo 1-T	CGGCACTACACTAATCCTTTATCCGGGTTATCTCATTTAA**ATCACCAACTACCCACACACACC**	3. Nucleotide in red is complementary to nucleotide immediately 5' of FEN cleavage site.
450-nt-ATXN3-oligo 1-T	CCGCTTCCCACCCATTTCTATGGATTCTTA**ATCACCAACTACCCACACACACC**	
450-nt-AXIN2-oligo 1-T	GTTTCCCTGGGAATTTGGGCTTGGA**ATCACCAACTACCCACACACACC**	
450-nt-BRCA1-oligo 1-T	CCCTAATGGAGGTCTCCAGTTTCGGTAAATATAA**ATCACCAACTACCCACACACACC**	4. Bolded nucleotides are complementary to oligo U1-T.
450-nt-BRCA2-oligo 1-T	GTGGGCGCCCTCGGGA**ATCACCAACTACCCACACACACC**	
450-nt-CCN4-oligo 1-T	CCTGCCTAGGTCTTTCACTACTGTAGGGA**ATCACCAACTACCCACACACACC**	
450-nt-CD44-oligo 1-T	GGGGAGTTGGTGAATCTTCCAAGTGGAAAGA**ATCACCAACTACCCACACACACC**	5. Used in Figs. 1b-e,2a-c,2f-i,3a-b; Extended Data Figs. 1,2a-c; Supplementary Figs. 4b,5b-c,6-9; Supplementary Tables 3-7.
450-nt-CDH13-oligo 1-T	TGGTGCTTCCCTGCCATCTCACAAA**ATCACCAACTACCCACACACACC**	
450-nt-CDH1-oligo 1-T	AGTGCAGTGGTGCGATCACAGCTCA**ATCACCAACTACCCACACACACC**	
450-nt-CDH2-oligo 1-T	CCGCAAGGGGGTGAGGCGA**ATCACCAACTACCCACACACACC**	
450-nt-CDKN2A-oligo 1-T	GGCAGCTCGCCCCCTAGCTA**ATCACCAACTACCCACACACACC**	
450-nt-CDKN2D-oligo 1-T	GCAGGCGGGAAAGAAGCTGGTTTAA**ATCACCAACTACCCACACACACC**	
450-nt-CETN1-oligo 1-T	GACCCCAGTTTTATAGGTTTCTTTAGGGA**ATCACCAACTACCCACACACACC**	
450-nt-CHEK1-oligo 1-T	TGGGGAACATAGGGAGACCCCATCTCTA**ATCACCAACTACCCACACACACC**	
450-nt-CHEK2-oligo 1-T	GCCAGATGTGGCGGGGGAGA**ATCACCAACTACCCACACACACC**	
450-nt-CSDE1-oligo 1-T	GGCCAACAATTCTGATGGCATATTCAGGAGTCA**ATCACCAACTACCCACACACACC**	
450-nt-DDB2-oligo 1-T	CAGCCTCCTAAGTAGATGGGACTACAA**ATCACCAACTACCCACACACACC**	
450-nt-DRD3-oligo 1-T	TGATGTCAATATCCTTACCTGCCTATTTAA**ATCACCAACTACCCACACACACC**	
450-nt-DRD4-oligo 1-T	CGCGACCCGAGAAACCGACAA**ATCACCAACTACCCACACACACC**	
450-nt-EGFR-oligo 1-T	CCCCCTTTCCCTTCTTTTGTTTTACAAAGCAAACTTGTA**ATCACCAACTACCCACACACACC**	
450-nt-EPM2AIP1-oligo 1-T	GCGCCACTACGAGGCTGAGCACGAATA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-EPSTI1-oligo 1-T	CAGGCAACTCTAATTTGAGGTTACCGCTGGA**ATCACCAACTACCCACACACACC**	
450-nt-ERCC1-oligo 1-T	CTCAGACATTGACTTGGCTTCAGTTTCCTCA**ATCACCAACTACCCACACACACC**	
450-nt-ERCC2-oligo 1-T	CCCAGGAGTTCGAGATCAGCTTGAGCAAGA**ATCACCAACTACCCACACACACC**	
450-nt-ERCC8-oligo 1-T	ATCCCGCAGTACAAGAACTGGAGA**ATCACCAACTACCCACACACACC**	
450-nt-EXO1-oligo 1-T	AGCCCCAGTTCTGGAGGCTGA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-FANCC-oligo 1-T	GTAGGGGAGGTAAAATGGAAATTAGTTCTTTTGAA**ATCACCAACTACCCACACACACC**	
450-nt-FAT4-oligo 1-T	GAGTTCCCAGAAAAGTTTCTCAAGCCTCTTATATA**ATCACCAACTACCCACACACACC**	
450-nt-FEN1-oligo 1-T	CGAGAGGCGTGGGGCGAA**ATCACCAACTACCCACACACACC**	
450-nt-FHIT-oligo 1-T	ACGTGTAACGAAAGGCTGTACCGCA**ATCACCAACTACCCACACACACC**	
450-nt-FOXC2-oligo 1-T	TCGGCCCGCGCGAGGA**ATCACCAACTACCCACACACACC**	
450-nt-FZD7-oligo 1-T	CGCACCGCAGCGCGCA**ATCACCAACTACCCACACACACC**	
450-nt-FZD8-oligo 1-T	GGCAGGGCCCTTCCGCA**ATCACCAACTACCCACACACACC**	
450-nt-GPR158-oligo 1-T	CGCTCTCATTGAGGCGCGTTCAGAA**ATCACCAACTACCCACACACACC**	
450-nt-GSTA1-oligo 1-T	CCATGTAAATGTGAATTTCAGGCAAACAATCAA**ATCACCAACTACCCACACACACC**	
450-nt-HIST1H1B-oligo 1-T	CATACTTGTTTGTGTTTTCAAGAATCTTGTGCAATTGA**ATCACCAACTACCCACACACACC**	
450-nt-HMGA2-oligo 1-T	CGGCGGCGGGGTTAAAGGGA**ATCACCAACTACCCACACACACC**	
450-nt-HOTAIR-oligo 1-T	CCGAGCCTTAATGTAGGAGCTAAAGCCAGA**ATCACCAACTACCCACACACACC**	
450-nt-HOXB2-oligo 1-T	GCCTCCGTTTCTCCAGAAAAGCTAGA**ATCACCAACTACCCACACACACC**	
450-nt-HSPA5-oligo 1-T	GTCGGAGTAGGTCCAGCAGGA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-INO80-oligo 1-T	CGCAGGAGAAAATGTCGCGAGAGCTTTCCA**ATCACCAACTACCCACACACACC**	
450-nt-KLF8-oligo 1-T	GGGCTGTCTAAGAGGGAAAGGAGAAAACAAAAAA**ATCACCAACTACCCACACACACC**	
450-nt-LIG1-oligo 1-T	GCAGCTGGACGTGGGGGGA**ATCACCAACTACCCACACACACC**	
450-nt-MBD4-oligo 1-T	GCTTGCTGCCTGGGCCGA**ATCACCAACTACCCACACACACC**	
450-nt-MCM9-oligo 1-T	CCCCAAATTTCCTTCATATTAACCCTGCA**ATCACCAACTACCCACACACACC**	
450-nt-MECP2-oligo 1-T	GCAAGTGAAGGCCTAGGCTTGGGA**ATCACCAACTACCCACACACACC**	
450-nt-MEG3-oligo 1-T	ATTCAGGCACAATGTGTGGGGGGAA**ATCACCAACTACCCACACACACC**	
450-nt-MGMT-oligo 1-T	GATGGCCCTTCGGCCGGTACAA**ATCACCAACTACCCACACACACC**	
450-nt-MLH1-oligo 1-T	GGAAACGTCTAGATGCTCAACGGAA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-MLH3-oligo 1-T	CTGGCCTATGGTAGGCACCGAA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-MNAT1-oligo 1-T	GATTTATGAGCGGTGTACTTACTTTCTTTTGGACCA**ATCACCAACTACCCACACACACC**	
450-nt-MPG-oligo 1-T	ACGTGTTTTCCGGTGGGGTGCCA**ATCACCAACTACCCACACACACC**	
450-nt-MRE11-oligo 1-T	CCGCTTCGTGGTCCCCAGGTA**ATCACCAACTACCCACACACACC**	
450-nt-MSH2-oligo 1-T	AAACTAGCCTTCCTCCTCCTCCCAGTA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-MSH3-oligo 1-T	CCGACTCCCGGCGAGAGGA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-MSH4-oligo 1-T	GCACACTAGCTTATCTTTAAATGTAGAAACCACCCAACA**ATCACCAACTACCCACACACACC**	
450-nt-MSH5-oligo 1-T	GGGTTTTGAGGGGGAGGCCAAACTTAAAGA**ATCACCAACTACCCACACACACC**	
450-nt-MSH6-oligo 1-T	GATAACCGGGGGGACCCGAAA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-MUTYH-oligo 1-T	CAAGCCCTACAAAATGTTTCCTGCAGGA**ATCACCAACTACCCACACACACC**	
450-nt-NEIL2-oligo 1-T	CTGGCGCCCGGCTAATTTTTGTATTTTTAGTAGAGA**ATCACCAACTACCCACACACACC**	
450-nt-NEIL3-oligo 1-T	CCTCTTAATCTGTCCTTGTCTGTGACA**ATCACCAACTACCCACACACACC**	
450-nt-NKX2-5-oligo 1-T	GGGTGGGAAGCGGTGGGGA**ATCACCAACTACCCACACACACC**	
450-nt-OGG1-oligo 1-T	CGTTCTCCCCCGAGCCTTTGCA**ATCACCAACTACCCACACACACC**	
450-nt-OLIG2-oligo 1-T	GGGGATGGAGGGAGGGCTGA**ATCACCAACTACCCACACACACC**	
450-nt-PARP1-oligo 1-T	AGGCGCTCACCACAACTGGCTAA**ATCACCAACTACCCACACACACC**	
450-nt-PCNA-oligo 1-T	AGCCCTGCTCCCGGGAGA**ATCACCAACTACCCACACACACC**	
450-nt-PELO-oligo 1-T	ACCACCCCGACAGGCTCCA**ATCACCAACTACCCACACACACC**	
450-nt-PMS1-oligo 1-T	GCTTTGCCTTCCAGCTGAGATGTAAA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-PMS2-oligo 1-T	CTCTCCGTGTGGAGCTTCCCA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-PMS2P11-oligo 1-T	CCCATCACCATGCCTGGCTAATTTTTGTA**ATCACCAACTACCCACACACACC**	
450-nt-PMS2P1-oligo 1-T	GAGTTCGAGACCAGTCTGGCCAA**ATCACCAACTACCCACACACACC**	
450-nt-PMS2P2-oligo 1-T	CCTTGGTGTCTGGCGGTTTCCAA**ATCACCAACTACCCACACACACC**	
450-nt-PMS2P3-oligo 1-T	CTGGAGCCCAAGTCCTCTGGGTTA**ATCACCAACTACCCACACACACC**	
450-nt-PMS2P5-oligo 1-T	GCGTTGGTGTCTGGCGGTTTCTGA**ATCACCAACTACCCACACACACC**	
450-nt-POLD1-oligo 1-T	TCGGGTGTTGCGCGCGCA**ATCACCAACTACCCACACACACC**	
450-nt-POLD4-oligo 1-T	GCTTCGCTAATCCCCGTGGCCA**ATCACCAACTACCCACACACACC**	
450-nt-POLG-oligo 1-T	GCCGCGGTCGAAGGGAGGTTATGATTAACTTTTA**ATCACCAACTACCCACACACACC**	
450-nt-POLK-oligo 1-T	AACGCAGGGGTCACGGCGA**ATCACCAACTACCCACACACACC**	
450-nt-POLQ-oligo 1-T	AGGTAGATCAGGGAAGTAGCTCTTCTTAA**ATCACCAACTACCCACACACACC**	
450-nt-POLR2B-oligo 1-T	AGGGTCACCACGCCCGGGTAA**ATCACCAACTACCCACACACACC**	
450-nt-POLR2G-oligo 1-T	TGTGCCACCACACCTGACTAATTTTGTA**ATCACCAACTACCCACACACACC**	
450-nt-PRDM2-oligo 1-T	CAGACAAGCTGTGCCAAGAAGTAATTCA**ATCACCAACTACCCACACACACC**	

450-nt-PRKCG-oligo 1-T	GGGCGCTAAGGTTCTCACTTTCTCCGA**ATCACCAACTACCCACACACACC**	
450-nt-PRUNE2-oligo 1-T	GGAGGTGCAGGGGTATGGCGA**ATCACCAACTACCCACACACACC**	
450-nt-PXMP4-oligo 1-T	GGAGGAGCAGGGCCCCA**ATCACCAACTACCCACACACACC**	
450-nt-PYCARD-oligo 1-T	CATGTCGTAGAATCTTGGAATCATGACAGCTAGAA**ATCACCAACTACCCACACACACC**	
450-nt-RNASEH2A-oligo 1-T	CGTAGCTTCGAAGACCCAGCCTA**ATCACCAACTACCCACACACACC**	
450-nt-RNF111-oligo 1-T	CAACATTTAAGCCAAATCTTGACTACATCTTGGTTCA**ATCACCAACTACCCACACACACC**	
450-nt-RPA2-oligo 1-T	GGTCACATTCTCTTCGTGTAATGCCACCA**ATCACCAACTACCCACACACACC**	
450-nt-RPA3-oligo 1-T	GGGGAAATGTATGTGTGTTTTAAAGCTTCCCA**ATCACCAACTACCCACACACACC**	
450-nt-RPA4-oligo 1-T	CTTTGGGAGGCTGAGGCAGGTTGA**ATCACCAACTACCCACACACACC**	
450-nt-SH3TC1-oligo 1-T	CTTCCATTTACTGCACACACACACA**ATCACCAACTACCCACACACACC**	
450-nt-SIM2-oligo 1-T	GCGGCAGGGAGGGCGA**ATCACCAACTACCCACACACACC**	
450-nt-SMAD4-oligo 1-T	GCGTAGAGTGGGCGTCCAGTA**ATCACCAACTACCCACACACACC**	
450-nt-SMUG1-oligo 1-T	CATTCACTATTTTCCCAAAGGGATTGCA**ATCAATCACCAACTACCCACACACACC**	
450-nt-SNAI1-oligo 1-T	GGGCGGGGACTCCTCCGA**ATCACCAACTACCCACACACACC**	
450-nt-SNAI2-oligo 1-T	GCCTGGCTTCCAGATGTGGTGCA**ATCACCAACTACCCACACACACC**	
450-nt-TAF1-oligo 1-T	GGAGCTGCTCGAGTCACGTGGCTTA**ATCACCAACTACCCACACACACC**	
450-nt-TDG-oligo 1-T	CTGTGTTGCATTATCAGGTAATGAGATGTGAGGGA**ATCACCAACTACCCACACACACC**	Same as 300-nt-oligo 1-T
450-nt-TP53-oligo 1-T	CTCTCAGCTGGATCCTTTCTTCTCATCAGTTAAAA**ATCACCAACTACCCACACACACC**	
450-nt-TREX1-oligo 1-T	CATGCCCACCTGTGGGTCCCA**ATCACCAACTACCCACACACACC**	
450-nt-TRIM28-oligo 1-T	TTCGCCGCCGTAGAGTCGTCTCCA**ATCACCAACTACCCACACACACC**	
450-nt-TWIST1-oligo 1-T	GCAGCGCCCCCGAACCCTA**ATCACCAACTACCCACACACACC**	
450-nt-TWIST2-oligo 1-T	AGTCCCAGATGCGCTCCCAGAA**ATCACCAACTACCCACACACACC**	
450-nt-UBB-oligo 1-T	ACCTCCGCCTCCCGGGTTCAA**ATCACCAACTACCCACACACACC**	
450-nt-USP47-oligo 1-T	CTTCTCTAGTAGTAAGCATGGTGCTTTGTGTTATA**ATCACCAACTACCCACACACACC**	
450-nt-WDR48-oligo 1-T	GGCCCAAGCAAAACGAAAGCACTGAAAGAA**ATCACCAACTACCCACACACACC**	
450-nt-WNT1-oligo 1-T	TCCGGAGCCCGTCGAGGACTA**ATCACCAACTACCCACACACACC**	
450-nt-XPC-oligo 1-T	AGGCTGCTTTTTCCAGACTCAGGAAAAA**ATCACCAACTACCCACACACACC**	
450-nt-XRCC2-oligo 1-T	GGCCCCTTTGTCCGTTCTCGGGTTAAAA**ATCACCAACTACCCACACACACC**	
450-nt-XRCC3-oligo 1-T	TCCGCGCGCGGCTGACA**ATCACCAACTACCCACACACACC**	
450-nt-ZEB1-oligo 1-T	GGGCAGGGAGGGATCTGGCTGA**ATCACCAACTACCCACACACACC**	
450-nt-ZNF830-oligo 1-T	GACAGCGGTTGGAGCGGCCA**ATCACCAACTACCCACACACACC**	
450-nt-ALKBH1-oligo 2-T	CTGCGTGGGCCGCCGA**ATCACCAACTACCCACACACACC**	
450-nt-ALKBH2-oligo 2-T	CGATTCAGAAGCGAGACGCAAAGAGGTTAAACA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T-T
450-nt-ALKBH3-oligo 2-T	GGGTGGTACTCGCACCCCCA**ATCACCAACTACCCACACACACC**	
450-nt-APEX1-oligo 2-T	CCTCTTATGAGCAAAAGAGCAACCCCGTA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-APEX2-oligo 2-T	CGGCACAGTTGCTGGGTTCCTGA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-ATM-oligo 2-T	GCCAAGCAGCCGCAGAGCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-ATRX-oligo 2-T	CGTCACTGTAGCTGCTGCTGGAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-ATXN3-oligo 2-T	CCCCTCGCCGGGCGAGA**ATCACCAACTACCCACACACACC**	
450-nt-AXIN2-oligo 2-T	GAGTTGCCAGGACCTTATCAAAGCGCA**ATCACCAACTACCCACACACACC**	
450-nt-BRCA1-oligo 2-T	GGGCCCAGTTATCTGAGAAACCCCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-BRCA2-oligo 2-T	GGTGCAGTTCTTTTTTGGCCGGAGTAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-CCN4-oligo 2-T	AGGCAGGACTCACCGTGGCCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-CD44-oligo 2-T	GGCTGTGACTAAGCAGCCCGCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-CDH13-oligo 2-T	TGGTCATCAGCCTCTACCCAATGCTGCA**ATCACCAACTACCCACACACACC**	
450-nt-CDH1-oligo 2-T	GGCTGGAGCGGGCTGGA**ATCACCAACTACCCACACACACC**	
450-nt-CDH2-oligo 2-T	CCCAGAGATCAAGGAGCTGGGGA**ATCACCAACTACCCACACACACC**	
450-nt-CDKN2A-oligo 2-T	ACCATGTTCTCGCCGCCTCCA**ATCACCAACTACCCACACACACC**	
450-nt-CDKN2D-oligo 2-T	ACGGCGAGGGGCTGGGA**ATCACCAACTACCCACACACACC**	
450-nt-CETN1-oligo 2-T	AAATGCTTCCCGAACTTCTTGCTTCTGA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-CHEK1-oligo 2-T	GAAAGCGAGCAGTTTATGTTGGAGCTTGAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-CHEK2-oligo 2-T	CACACACTCCCCCTTCAGCTCAAAACTACA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-CSDE1-oligo 2-T	GGCTTCTCGTTAGCCTTGGTCTTATTCTCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-DDB2-oligo 2-T	GGGTACTGTGCTCTATGCGAAGATCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-DRD3-oligo 2-T	CCTGAGACTGGATAATTTATAATGAACAGAAA**ATCACCAACTACCCACACACACC**	
450-nt-DRD4-oligo 2-T	CACGAGCGAGTTCCCCGCGA**ATCACCAACTACCCACACACACC**	
450-nt-EGFR-oligo 2-T	CCGCGTCGGGCGCTCA**ATCACCAACTACCCACACACACC**	
450-nt-EPM2AIP1-oligo 2-T	GCTGGGTCTCTTCGTCCCTCCCTGAA**ATCACCAACTACCCACACACACC**	
450-nt-EPSTI1-oligo 2-T	CGAGCCCGGAGTTCACCACTCTA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-ERCC1-oligo 2-T	GGCGCCAATCTCCACCCGCA**ATCACCAACTACCCACACACACC**	
450-nt-ERCC2-oligo 2-T	ACCCAGGCGCGCCCA**ATCACCAACTACCCACACACACC**	
450-nt-ERCC8-oligo 2-T	GGCCTGTTAGCCAAAAAGTAAGGTTTTCTTTA**ATCACCAACTACCCACACACACC**	
450-nt-EXO1-oligo 2-T	CTCCTCCTCCCCTTGAGCGGA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-FANCC-oligo 2-T	TCCTCGCGGGAGCTGCTTCA**ATCACCAACTACCCACACACACC**	
450-nt-FAT4-oligo 2-T	GCTGCTGGAAGAGCAACCTCCA**ATCACCAACTACCCACACACACC**	Does not reconstitute double flap at 3′ end
450-nt-FEN1-oligo 2-T	GCAACCAGTCCCTCCAGGGAGAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-FHIT-oligo 2-T	CTCCCCGCCCCTACCTTCCA**ATCACCAACTACCCACACACACC**	
450-nt-FOXC2-oligo 2-T	TAGCTGCCCGCAGCCCGGTA**ATCACCAACTACCCACACACACC**	
450-nt-FZD7-oligo 2-T	GCCAGCACCAGGGCACAGA**ATCACCAACTACCCACACACACC**	
450-nt-FZD8-oligo 2-T	GGCGAGCACTGGATCTCCACCA**ATCACCAACTACCCACACACACC**	
450-nt-GPR158-oligo 2-T	TCGCGGCTGGCGCCAACA**ATCACCAACTACCCACACACACC**	
450-nt-GSTA1-oligo 2-T	GTCACCGTCCTGGCTCGACAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-HIST1H1B-oligo 2-T	TTAGAAGCAGCCACAGCCTTGGTGA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-HMGA2-oligo 2-T	CAGAGGACAGAGTAGTGGGTGGCA**ATCACCAACTACCCACACACACC**	
450-nt-HOTAIR-oligo 2-T	GGGGCAAGGAGGCATAGGCCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-HOXB2-oligo 2-T	GAGCGACGGCTGGCTGTTTATAAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-HSPA5-oligo 2-T	ACCAGGCGAAGGGCAGGTCTA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-INO80-oligo 2-T	GCGGCTCCTGGGACCCCAA**ATCACCAACTACCCACACACACC**	
450-nt-KLF8-oligo 2-T	CCCCACGCGAGTCAGAACCTAAA**ATCACCAACTACCCACACACACC**	
450-nt-LIG1-oligo 2-T	CACTGGGAGTCGTAGTCTCCCGAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-MBD4-oligo 2-T	GCCATTTCCACGCTCCACCAATCAAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-MCM9-oligo 2-T	GATTTCCATGTTGGTCTCAAACAGAGTCA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-MECP2-oligo 2-T	GCGGCGGCCATTTTCCGGA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T
450-nt-MEG3-oligo 2-T	CAGAGCCCATCCCAGAAGCGGA**ATCACCAACTACCCACACACACC**	
450-nt-MGMT-oligo 2-T	GAGGCTGCCACCGTCCCGA**ATCACCAACTACCCACACACACC**	
450-nt-MLH1-oligo 2-T	GCATACATAATTTTAAATTAAGTGGCTTCCTTA**ATCACCAACTACCCACACACACC**	
450-nt-MLH3-oligo 2-T	GAGTCCCGCTTAACAAATTCTCCGA**ATCACCAACTACCCACACACACC**	
450-nt-MNAT1-oligo 2-T	GGGGTTCCGATATTTGGTGGTCTTA**ATCACCAACTACCCACACACACC**	
450-nt-MPG-oligo 2-T	CCAGCAGCCGCCGACCA**ATCACCAACTACCCACACACACC**	
450-nt-MRE11-oligo 2-T	GAAGCCCCACAATCGCGACCCAA**ATCACCAACTACCCACACACACC**	Same as 950-nt-oligo 2-T

450-nt-MSH2-oligo 2-T	GCACCGCCATGTCGAAACCTCCTCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-MSH3-oligo 2-T	GCTTGCCTCGCAGGGGCTGA**ATCACCAACTACCCACACACACC**		
450-nt-MSH4-oligo 2-T	GGGAAACCGCCGGGGCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-MSH5-oligo 2-T	CTCCCCCAGCTGCCGCCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-MSH6-oligo 2-T	AGGCCCAACCGTTCTGTCGGA**ATCACCAACTACCCACACACACC**		
450-nt-MUTYH-oligo 2-T	GACGGTGAGCGGCTTCCCA**ATCACCAACTACCCACACACACC**		
450-nt-NEIL2-oligo 2-T	ATGCGGCCTCCGTCCCCTTA**ATCACCAACTACCCACACACACC**		
450-nt-NEIL3-oligo 2-T	CGCGCGCGAATCTTCTCTCCA**ATCACCAACTACCCACACACACC**		
450-nt-NKX2-5-oligo 2-T	CCCCTGGCAGCTTCCCTGCA**ATCACCAACTACCCACACACACC**		
450-nt-OGG1-oligo 2-T	TGCCTCGCCCTTTGCGGGTCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-OLIG2-oligo 2-T	GAGGAGAGTCTGTCCCTGGGATCTA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-PARP1-oligo 2-T	CCAGAGCCGCCACCGAACA**ATCACCAACTACCCACACACACC**		
450-nt-PCNA-oligo 2-T	ACCGCTTGGCAAAGCAAAGCAGA**ATCACCAACTACCCACACACACC**		
450-nt-PELO-oligo 2-T	CGCGCTCTTGGTCTCTGAAATCTCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-PMS1-oligo 2-T	TAGGGAACTGCTGGCGGCCTA**ATCACCAACTACCCACACACACC**		
450-nt-PMS2-oligo 2-T	GGGGACACCGGAAGACTGCGA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-PMS2P11-oligo 2-T	CCACATGGAGAAAACATACATTGTATCTCTCAAATTACCAAAA**ATCACCAACTACCCACACACACC**		
450-nt-PMS2P1-oligo 2-T	GTGGCCTGGCGTGGATTCCGA**ATCACCAACTACCCACACACACC**		
450-nt-PMS2P2-oligo 2-T	CACGAGGCGGCCTCCCA**ATCACCAACTACCCACACACACC**		
450-nt-PMS2P3-oligo 2-T	CGCGCGAACCCTGGTCTGCA**ATCACCAACTACCCACACACACC**		
450-nt-PMS2P5-oligo 2-T	CCTCGGCCATGTTCCCCCCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-POLD1-oligo 2-T	GGCCAAAGGGTGCCCTCGCAA**ATCACCAACTACCCACACACACC**		
450-nt-POLD4-oligo 2-T	GAGGCAACTGCGGGGAAGAGGA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-POLG-oligo 2-T	ACCCCAGGCCCACGGCA**ATCACCAACTACCCACACACACC**		
450-nt-POLK-oligo 2-T	GGGATACTCACCCGTTACCTCAGGA**ATCACCAACTACCCACACACACC**		
450-nt-POLQ-oligo 2-T	CGAGAACGAATCTGAGCCTGATTCTGAA**ATCACCAACTACCCACACACACC**		
450-nt-POLR2B-oligo 2-T	AGGATCTGCGCCTGCGCAAATTGTA**ATCACCAACTACCCACACACACC**		
450-nt-POLR2G-oligo 2-T	GGCCTTGCTCTTCCTTTCCACCCA**ATCACCAACTACCCACACACACC**		
450-nt-PRDM2-oligo 2-T	GCCTTCTAACAAGTGGTAACAGAGCTTCA**ATCACCAACTACCCACACACACC**		
450-nt-PRKCG-oligo 2-T	AGCGGGCGAGTGGGGAGGAA**ATCACCAACTACCCACACACACC**		
450-nt-PRUNE2-oligo 2-T	CCTCCTCTGCTGCAACACCTTTTCCTCTCCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-PXMP4-oligo 2-T	CAGTGCGTTGACGACTACGAGCA**ATCACCAACTACCCACACACACC**		
450-nt-PYCARD-oligo 2-T	CCGCTCACCCCGCTGCA**ATCACCAACTACCCACACACACC**		
450-nt-RNASEH2A-oligo 2-T	CCCGCCTCATCGACGCCCA**ATCACCAACTACCCACACACACC**		
450-nt-RNF111-oligo 2-T	GCCTGGACTCGGCCCAGCTAA**ATCACCAACTACCCACACACACC**		
450-nt-RPA2-oligo 2-T	GCCAATCGGTGCTCACGGATGCTA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-RPA3-oligo 2-T	CTTGACTATGTGTGTCTTCAAAATAATACCTTTAAGTTA**ATCACCAACTACCCACACACACC**		
450-nt-RPA4-oligo 2-T	GGCCACCTTCTCACCTGATCGCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-SH3TC1-oligo 2-T	CATACCTGCGGCGTGTTCCTCA**ATCACCAACTACCCACACACACC**		
450-nt-SIM2-oligo 2-T	GGAGCCCAGCGGAGGGGA**ATCACCAACTACCCACACACACC**		
450-nt-SMAD4-oligo 2-T	CTGGGCGGCCGCGGA**ATCACCAACTACCCACACACACC**		
450-nt-SMUG1-oligo 2-T	CTTACCCCATACGGAGGGAGGGA**ATCACCAACTACCCACACACACC**		
450-nt-SNAI1-oligo 2-T	GCAGGCCGCCTCCCTCA**ATCACCAACTACCCACACACACC**		
450-nt-SNAI2-oligo 2-T	CGGATAACGGTCCGGCGGGA**ATCACCAACTACCCACACACACC**		
450-nt-TAF1-oligo 2-T	AATGGGCCGCCTCCAGCGGAA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-TDG-oligo 2-T	CTCACTGCGGTAAGACAACCCCA**ATCACCAACTACCCACACACACC**		
450-nt-TP53-oligo 2-T	GGAAGCGTGTCACCGTCGTGGAAA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-TREX1-oligo 2-T	CCCTCCCCTAGTGCCCCCTTA**ATCACCAACTACCCACACACACC**		
450-nt-TRIM28-oligo 2-T	CCGCCGAAACCCGCACCA**ATCACCAACTACCCACACACACC**		
450-nt-TWIST1-oligo 2-T	GCGGGAGGGGGAGGGGA**ATCACCAACTACCCACACACACC**		
450-nt-TWIST2-oligo 2-T	ACACGCTCTCGGAGTTGCTGGAAA**ATCACCAACTACCCACACACACC**		
450-nt-UBB-oligo 2-T	AGTTTAGGCTTTTCAACTGAGCCCCAAA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-USP47-oligo 2-T	CCTGGACGTCATCCCCTCTTCA**ATCACCAACTACCCACACACACC**		
450-nt-WDR48-oligo 2-T	TGGCGCCGGGCACCTTCTA**ATCACCAACTACCCACACACACC**		
450-nt-WNT1-oligo 2-T	CTCTGTGCTGCAGTTGCGGCGA**ATCACCAACTACCCACACACACC**		
450-nt-XPC-oligo 2-T	GGGCCTTGCTCTTGGCCTTGGA**ATCACCAACTACCCACACACACC**		
450-nt-XRCC2-oligo 2-T	TCCCAACCCGGCGGCTCTCA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-XRCC3-oligo 2-T	CGCCCCCGCCCCTTGA**ATCACCAACTACCCACACACACC**		
450-nt-ZEB1-oligo 2-T	CGGGTTCGCCTGCTTTCTGCGCTTA**ATCACCAACTACCCACACACACC**		
450-nt-ZNF830-oligo 2-T	CGGTTGTACTTCGCGAATGGAGATTCTA**ATCACCAACTACCCACACACACC**		Same as 950-nt-oligo 2-T
450-nt-ALKBH1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAAGTCGATGACCCCTTCCAGGT	1. Oligos designed to capture target sequences of ~450 nt in length.	
450-nt-ALKBH2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGGTAAACCGCACGCAAAATTCTG		Same as 950-nt-oligo 3-T
450-nt-ALKBH3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCCGCGCACCCA		
450-nt-APEX1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCTGCACCGGCCCGATG	2. Non-bolded nucleotides are complementary to the 3' end of target sequence.	Same as 950-nt-oligo 3-T
450-nt-APEX2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTGCCACCCCTTGCAGGGGTCT		Same as 950-nt-oligo 3-T
450-nt-ATM-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGTATGCCCATGCGCATCCAGT		Same as 950-nt-oligo 3-T
450-nt-ATRX-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCCCCACAGCTCAAAGGC	3. Nucleotide in red is complementary to 3' end of target sequence, i.e., immediately 5' of cleavage site.	Same as 950-nt-oligo 3-T
450-nt-ATXN3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGGCATGGGGGCGACTCG		
450-nt-AXIN2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCGGCTCCAGCCGACT		
450-nt-BRCA1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAGCCTGTCCCCCGTCCA		Same as 950-nt-oligo 3-T
450-nt-BRCA2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTGACAAAAACCGCGCCGG	4. Bolded nucleotides are complementary to oligo U2.	Same as 950-nt-oligo 3-T
450-nt-CCN4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGACGGTGCTGGCGGCT		Same as 950-nt-oligo 3-T
450-nt-CD44-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGGTTTTAAGAAGTAGCAGCCCTCCC		Same as 950-nt-oligo 3-T
450-nt-CDH13-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTGGGTAGAGGCTGATGACCAGGACCAATG	5. Used in Figs. 1b-e,2a-c,2f-i,3a-b; Extended Data Figs. 1,2a-c; Supplementary Figs. 4b,5b-c,6-9; Supplementary Tables 3-7.	
450-nt-CDH1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTCTGAACTGACTTCCGCAAGCTCAC		
450-nt-CDH2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGGGCGGGGGAACAGG		
450-nt-CDKN2A-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGCCGAGCTCGGCAGC		
450-nt-CDKN2D-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCGGGCCCAACCCT		
450-nt-CETN1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCTCAGTGAGCTCGGGCTTA		Same as 950-nt-oligo 3-T
450-nt-CHEK1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTCAGGATAACCAGTTTTGCCTTTTTTCTGTG		Same as 950-nt-oligo 3-T
450-nt-CHEK2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGACAAAGCGAAGCTCAGGAGACTCC		Same as 950-nt-oligo 3-T
450-nt-CSDE1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCTTCCCCTGGGTTTCTGCTA		Same as 950-nt-oligo 3-T
450-nt-DDB2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGAGGGAGGATTGAGAGGCCTC		Same as 950-nt-oligo 3-T
450-nt-DRD3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTATTTGGCTCACAGTTCTGGAGTC		
450-nt-DRD4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCACCGCGCCGATGAGCA		
450-nt-EGFR-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACCGTGCGGGGGGC		
450-nt-EPM2AIP1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGGCTACTGCCCGCTACCTAGAA		
450-nt-EPSTI1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTGCGGGTGTTCATGGTTCACAGC		Same as 950-nt-oligo 3-T
450-nt-ERCC1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTCTCCGCCTCCCGCACCT		
450-nt-ERCC2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCGATGACCCCATCTTCAAGACCC		

450-nt-ERCC8-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCCGTGTTGACTCTGCTCT	
450-nt-EXO1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCGATTGCGGTTGGAGAAACC	Same as 950-nt-oligo 3-T
450-nt-FANCC-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCAGTTTGGGCAGTGGCGGAAA	
450-nt-FAT4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTTGGAACACCTGGCGCG	
450-nt-FEN1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAGAGTCTCTCAGCGACCCTG	Same as 950-nt-oligo 3-T
450-nt-FHIT-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGATGTTGACAGCTGGGAATGAAAGGCAG	
450-nt-FOXC2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTAATTCTGCTCGCTCAGGTAGGG	
450-nt-FZD7-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCCCAGGGACGAAAGCGGA	
450-nt-FZD8-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGGCCAGAACTGGTGCACCT	
450-nt-GPR158-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTCCCAATCCCAGCTGAGCAA	
450-nt-GSTA1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGAATTCCAGGTCCTAATGTATTTATAAGC	Same as 950-nt-oligo 3-T
450-nt-HIST1H1B-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCAGCTCTGAGACTGGGGGC	Same as 950-nt-oligo 3-T
450-nt-HMGA2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCGCGCCTCCTAGGGTGG	
450-nt-HOTAIR-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAACACTCACCCCACGGAGCAG	Same as 950-nt-oligo 3-T
450-nt-HOXB2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCCAATCTCCCTCTCAAATTCAAAATTCATGGC	Same as 950-nt-oligo 3-T
450-nt-HSPA5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAAATACAGGCCGCGGCGCT	Same as 950-nt-oligo 3-T
450-nt-INO80-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGATGGCGGCGGCGCTGA	
450-nt-KLF8-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTGCTCTCTGGGCAACCTGAG	
450-nt-LIG1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGGAGGAGGGCGGGAAAAG	Same as 950-nt-oligo 3-T
450-nt-MBD4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCATTCTCGAGGAAGACGCACCG	Same as 950-nt-oligo 3-T
450-nt-MCM9-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGCATTAACCACAACTGGGTAATGA	Same as 950-nt-oligo 3-T
450-nt-MECP2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGCTTTTACCACAGCCCTCTCT	Same as 950-nt-oligo 3-T
450-nt-MEG3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGCCGTCTTCCTTTTGCACATCCTT	
450-nt-MGMT-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGAGAGCTCCGCACTCTTC	
450-nt-MLH1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTTAGTTAACGGTACCATTAAGACAAC	
450-nt-MLH3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACCAACCGCCTCGGACGC	
450-nt-MNAT1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACCGAGGGCAACCCTGATC	
450-nt-MPG-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGAAGCACCCCTAGGACC	
450-nt-MRE11-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCTGTCTTCTTTTCGGGAAGAAAGG	Same as 950-nt-oligo 3-T
450-nt-MSH2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCCTGGTTGAAGAAAATGCGCGACC	Same as 950-nt-oligo 3-T
450-nt-MSH3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTGGAGGCAGCGAGGCC	
450-nt-MSH4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAAGGCGAGGTTGATGAGATCTCA	Same as 950-nt-oligo 3-T
450-nt-MSH5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGCAGCGTTTTCCCGCCTTTTC	Same as 950-nt-oligo 3-T
450-nt-MSH6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGAGCTCCTAAAAGCACCGCAT	
450-nt-MUTYH-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGGTAGCCTTCAAAGCCTCTGCG	
450-nt-NEIL2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGGACGCCCCTGTAGGT	
450-nt-NEIL3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTCAGAGTACAGCCTGGTCCTT	
450-nt-NKX2-5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGTGCCGCCGCCCG	
450-nt-OGG1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCAAAGCCCCGCTTCGTGCCCCAA	Same as 950-nt-oligo 3-T
450-nt-OLIG2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCTGCGGCTGCTGTGG	Same as 950-nt-oligo 3-T
450-nt-PARP1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCCGCACCGGCCA	
450-nt-PCNA-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGAAGCGTGGGTTTCATTTCTCATCATGTG	
450-nt-PELO-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGGCAGCTTCCCAGGCTCTG	Same as 950-nt-oligo 3-T
450-nt-PMS1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGGTCGTCCTGAGAGGGAAA	
450-nt-PMS2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCCGCTCACCTCGAGC	Same as 950-nt-oligo 3-T
450-nt-PMS2P11-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTTTGGCAACAATGGTGTCTTCTTTC	
450-nt-PMS2P1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGTTGGTGTCTGGCGGTTTC	
450-nt-PMS2P2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCGCCTCGGCCATGTTC	
450-nt-PMS2P3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAGCTCCGCGTCCGCCTAG	
450-nt-PMS2P5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTCCAGGGAGGTTGGAAGCCGTG	Same as 950-nt-oligo 3-T
450-nt-POLD1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCGTCTCTCAGATCTATGTGCCTGGCT	
450-nt-POLD4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGACCGGCCAGTGGGC	Same as 950-nt-oligo 3-T
450-nt-POLG-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGCGTCCCCCTTCGC	
450-nt-POLK-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGCGACTACAACTCCCAGGA	
450-nt-POLQ-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCCGCCGTTTCCCACT	
450-nt-POLR2B-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGTCAAGCTTGGGCCGGGGTT	
450-nt-POLR2G-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCTTGCCGCCACC	
450-nt-PRDM2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAATTAAGAATCTAAATTTATCCCAAGAGCCA	
450-nt-PRKCG-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGCCAAAGGCAGGGGCG	
450-nt-PRUNE2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTTTCCATTAGCATACGCGCGCG	Same as 950-nt-oligo 3-T
450-nt-PXMP4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGCCCTTAGCTGCGGC	
450-nt-PYCARD-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCGCCGACCAGGAGGAAGTC	
450-nt-RNASEH2A-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGACGCAAGGCTCCTTGCG	
450-nt-RNF111-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCGGCGTGGGTTCAGAATGCCTAATTT	
450-nt-RPA2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCTTGTTCCTGTCTCCTCTGC	Same as 950-nt-oligo 3-T
450-nt-RPA3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTTGAAGGAAAAAACTATGTATCCAACAAATTCAAAC	
450-nt-RPA4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATAATCCTTCCCTGTGTGTCAGTGGTTGC	Same as 950-nt-oligo 3-T
450-nt-SH3TC1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGGCAGCTCCCAGGCCC	
450-nt-SIM2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGAGTCGCAGGCGGCTG	
450-nt-SMAD4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCTAGGGGAGAGCAGGAAG	
450-nt-SMUG1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCCCAAAATATAGAAATTCTAACCACCCTCTCGGC	
450-nt-SNAI1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGTGCCCCCCACGCAG	
450-nt-SNAI2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGACACGGCGGTCCCTACAGCAT	
450-nt-TAF1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCTTCGTCGCTGTCCGTGTCTGACA	Same as 950-nt-oligo 3-T
450-nt-TDG-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGACCTGGGCCCGATGGCT	
450-nt-TP53-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCACGCTCCCAGCCCGAAC	Same as 950-nt-oligo 3-T
450-nt-TREX1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTCTGGGGTGGGGGAACCCA	
450-nt-TRIM28-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCGCCACAAGCACAAGCACAA	
450-nt-TWIST1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGTGTGGATGGCCCCGA	
450-nt-TWIST2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCTCTGATTTCAAGGCCGGCTTGT	
450-nt-UBB-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTCCTCATAGCCGTAAGAAAGGCTCCT	Same as 950-nt-oligo 3-T
450-nt-USP47-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGATACCCCAGGCCTCGTCCT	
450-nt-WDR48-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTGGAGTGGGCCCGGA	
450-nt-WNT1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTTTGGTTGTTGCCCGCGACGG	
450-nt-XPC-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTCTGGCTGCGCAGTTCGCGT	
450-nt-XRCC2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCGGTCCCAGACTCAGC	Same as 950-nt-oligo 3-T
450-nt-XRCC3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCCTGCCTCACCCCTCACG	
450-nt-ZEB1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACCTGGGGCCATCCGCCAT	
450-nt-ZNF830-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCGTTTCCGACTGGTGCTCAGAC	Same as 950-nt-oligo 3-T

Supplementary Table 1: FENGC oligonucleotides used in this study; human ~940-nt targets.

Oligo name	Oligo sequence	Notes	
940-nt-ALKBH2-oligo 1-T	CTCCCAAATTGTTGGGATTACAGGCGTGA**ATCACCAACTACCCACACACACC**	1. Olgos designed to capture target sequences of ~940 nt in length.	
940-nt-APEX1-oligo 1-T	ACTGAGCGCATTTTATGTGCTAAGTATTCA**ATCACCAACTACCCACACACACC**		
940-nt-APEX2-oligo 1-T	GAGACGTCGAATGAATTCATTTCATTCTCACAGTTCA**ATCACCAACTACCCACACACACC**		
940-nt-ATM-oligo 1-T	GAGAGACCTTCCTGGCGCCCTTCTAATAA**ATCACCAACTACCCACACACACC**	2. Non-bolded nucleotides are complementary to the 5' end of target sequence.	
940-nt-ATRX-oligo1	CAGAGGAAGCTTGGCTGGATGATCTGTAAA**ATCACCAACTACCCACACACACC**		
940-nt-BRCA1-oligo 1-T	TCCGTCGACGCAATCGCCACCA**ATCACCAACTACCCACACACACC**		
940-nt-BRCA2-oligo 1-T	ACCACGCACGGCTAACTTTGCA**ATCACCAACTACCCACACACACC**		
940-nt-CCN4-oligo 1-T	GCATTGCGCCCAACCCTACTTCTGGTTTTCA**ATCACCAACTACCCACACACACC**	3. Nucleotide in red is complementary to the nucleotide immediately 5' of cleavage site.	
940-nt-CD44-oligo 1-T	CCCATCCACCCACATATCTGTTTATTGATATATCCACTCA**ATCACCAACTACCCACACACACC**		
940-nt-CETN1-oligo 1-T	CCATGTGTTCAACTTGACATAAATGGGAAACAGA**ATCACCAACTACCCACACACACC**		
940-nt-CHEK1-oligo 1-T	AGAGGCTGCAGTGAGCAGAGA**ATCACCAACTACCCACACACACC**		
940-nt-CHEK2-oligo 1-T	GAGGCTAGAGGGGAGGGAGATTTGA**ATCACCAACTACCCACACACACC**	4. Bolded nucleotides are complementary to oligo U1-T.	
940-nt-CSDE1-oligo 1-T	CCCAGATTCAATGCTTCTGTTTTCACAGATTCA**ATCACCAACTACCCACACACACC**		
940-nt-DDB2-oligo 1-T	GCAGTGTCTCTCTTCCGTTACCCA**ATCACCAACTACCCACACACACC**		
940-nt-EPM2AIP1-MLH1-oligo 1-T	ACGAATACTACGAGCGGTATGTGGCGGA**ATCACCAACTACCCACACACACC**	5. Used in Fig. 3c,d; Extended Data Fig. 3; Supplementary Figs. 10,11; Supplementary Tables 8,9.	
940-nt-EPSTI1-oligo 1-T	GCTACCTGCATTCTAGGAGTATCCTAGTAGACA**ATCACCAACTACCCACACACACC**		
940-nt-EXO1-oligo	GATAATGCAGACTCAATTCAAGGAATTGAAGGGGAA**ATCACCAACTACCCACACACACC**		
940-nt-FEN1-oligo 1-T	GGCTGACCGGTGACAATGAAGTTAACATTTGTA**ATCACCAACTACCCACACACACC**		
940-nt-GSTA1-oligo 1-T	TGGAACACAGAGGCAGAAATTAGAGTGA**ATCACCAACTACCCACACACACC**		
940-nt-HIST1H1B-oligo 1-T	GAGTGTTTTCTCTTTTTAGTTCCTGGGCAAATCCA**ATCACCAACTACCCACACACACC**		
940-nt-HOTAIR-oligo 1-T	TGGGGTGTCCCGGGTGCAA**ATCACCAACTACCCACACACACC**		
940-nt-HOXB2-oligo 1-T	CCTTATGTACCCCTGCTTTCCTCACTTTCA**ATCACCAACTACCCACACACACC**		
940-nt-HSPA5-oligo 1-T	TTTACTTTCCCCTCCGCTGCCCA**ATCACCAACTACCCACACACACC**		
940-nt-LIG1-oligo 1-T	GGAAGGAGGAGTTCCAGATGAATACGAGAGTTA**ATCACCAACTACCCACACACACC**		
940-nt-MBD4-oligo 1-T	GAGATCTCGTTTCTCTAAAACTGGAGATGGGCA**ATCACCAACTACCCACACACACC**		
940-nt-MCM9-oligo 1-T	GGTTCTAGCTTTTGTTTATATGAACACAGTGTCA**ATCACCAACTACCCACACACACC**		
940-nt-MECP2-oligo 1-T	CTGTTAGCTAACCAGGTGTTCAGGACAGGA**ATCACCAACTACCCACACACACC**		
940-nt-MRE11-oligo 1-T	GTTTCTTAACTGTAAGTCGGGTAGATGCGCCTA**ATCACCAACTACCCACACACACC**		
940-nt-MSH2-oligo 1-T	AGGGCTGAGGCGAAAGGATCACCTGA**ATCACCAACTACCCACACACACC**		
940-nt-MSH4-oligo 1-T	CGGTTTACCCACATGCAGTGACAACA**ATCACCAACTACCCACACACACC**		
940-nt-MSH5-oligo 1-T	GCGGTGAACATAATGAAGTCACTGCTCCATAAAGCTTA**ATCACCAACTACCCACACACACC**		
940-nt-OGG1-oligo 1-T	CTGGGCCATTCTCACCTTTTAGCATAGGGA**ATCACCAACTACCCACACACACC**		
940-nt-OLIG2-oligo 1-T	CTCTCCTTTAGGTTGAGGATGGTGTGGGTA**ATCACCAACTACCCACACACACC**		
940-nt-PELO-oligo 1-T	GCGGTTTTGAGAATTTCAAAGACAGAAAGAATAAGTGGA**ATCACCAACTACCCACACACACC**		
940-nt-PMS2-oligo 1-T	CCATTGCACTCGAGCCTGGGCAA**ATCACCAACTACCCACACACACC**		
940-nt-PMS2P5-oligo 1-T	GGCCAACATGGTGAAACCCCGTGTCTA**ATCACCAACTACCCACACACACC**		
940-nt-POLD4-oligo 1-T	CCATGTGCCAAGTACTGTTCTAAGCACTTTCCCACA**ATCACCAACTACCCACACACACC**		
940-nt-PRUNE2-oligo 1-T	TAGGTGCCCTTATGTCAGCCCTGGGACTAA**ATCACCAACTACCCACACACACC**		
940-nt-RPA2-oligo 1-T	TGTCCAGGCTGCAGTGAGCTGTGA**ATCACCAACTACCCACACACACC**		
940-nt-RPA4-oligo 1-T	CAAGGCTCTGTGGACAGGACAGAAAAGTACAAGA**ATCACCAACTACCCACACACACC**		
940-nt-TAF1-oligo 1-T	CGGAAAATGTGAAAGTTTAGTAAAATGAACAGCTGACCCCTCA**ATCACCAACTACCCACACACACC**		
940-nt-TP53-oligo 1-T	GCTTGATGGGAACGGGAAACCTTCTAACCTTTCA**ATCACCAACTACCCACACACACC**		
940-nt-UBB-oligo 1-T	CTTGAACCTAGGAGGCAGAGATTGCA**ATCACCAACTACCCACACACACC**		
940-nt-XRCC2-oligo 1-T	CCCACCTAGTTTAGTTTTCAAAAACGCTAGGAAAGAGCA**ATCACCAACTACCCACACACACC**		
940-nt-ZNF830-oligo 1-T	GCACCCTTTTCTATAGACAAGAGTGAGGAAAAGAGA**ATCACCAACTACCCACACACACC**		
940-nt-ALKBH2-oligo 2-T	CGATTCAGAAGCGAGACGCAAAGAGGTTAAACA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-APEX1-oligo 2-T	CCTCTTATGAGCAAAAGAGCAACCCCGTA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-APEX2-oligo 2-T	CGGCACAGTTGCTGGGTTCCTGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-ATM-oligo 2-T	GCCAAGCAGCCGCAGAGCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-ATRX-oligo 2-T	CGTCACTGTAGCTGCTGCTGGAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-BRCA1-oligo 2-T	GGGCCCAGTTATCTGAGAAACCCCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-BRCA2-oligo 2-T	GGTGCAGTTCTTTTTTGGCCGGAGTAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-CCN4-oligo 2-T	AGGCAGGACTCACCGTGGCCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-CD44-oligo 2-T	GGCTGTGACTAAGCAGCCCGCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-CETN1-oligo 2-T	AAATGCTTCCCGAACTTCTTGCTTCTGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-CHEK1-oligo 2-T	GAAAGCGAGCAGTTTATGTTGGAGCTTGAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-CHEK2-oligo 2-T	CACACACTCCCCCTTCAGCTCAAAACTACA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-CSDE1-oligo 2-T	GGCTTCTCGTTAGCCTTGGTCTTATTCTCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-DDB2-oligo 2-T	GGGTACTGTGCTCTATGCGAAGATCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-EPM2AIP1-MLH1-oligo 2-T	AACCGGGCTTGTGTGCCTCTGCTGA**ATCACCAACTACCCACACACACC**		Same as 300 bp-MLH1-oligo 2-T
940-nt-EPSTI1-oligo 2-T	CGAGCCCGGAGTTCACCACTCTA**ATCACCAACTACCCACACACAC**C		Same as 450-nt-oligo 2-T
940-nt-EXO1-oligo 2-T	CTCCTCCTCCCCTTGAGCGGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-FEN1-oligo 2-T	GCAACCAGTCCCTCCAGGGAGAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-GSTA1-oligo 2-T	GTCACCGTCCTGGCTCGACAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-HIST1H1B-oligo 2-T	TTAGAAGCAGCCACAGCCTTGGTGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-HOTAIR-oligo 2-T	GGGGCAAGGAGGCATAGGCCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-HOXB2-oligo 2-T	GAGCGACGGCTGGCTGTTTATAAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-HSPA5-oligo 2-T	ACCAGGCGAAGGGCAGGTCTA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-LIG1-oligo 2-T	CACTGGGAGTCGTAGTCTCCCGAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MBD4-oligo 2-T	GCCATTTCCACGCTCCACCAATCAAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MCM9-oligo 2-T	GATTTCCATGTTGGTCTCAAACAGAGTCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MECP2-oligo 2-T	GCGGCGGCCATTTTCCGGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MRE11-oligo 2-T	GAAGCCCCACAATCGCGACCCAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MSH2-oligo 2-T	GCACCGCCATGTCGAAACCTCCTCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MSH4-oligo 2-T	GGGAAACCGCCGGGGCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-MSH5-oligo 2-T	CTCCCCCAGCTGCCGCCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-OGG1-oligo 2-T	TGCCTCGCCCTTTGCGGGTCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-OLIG2-oligo 2-T	GAGGAGAGTCTGTCCCTGGGATCTA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-PELO-oligo 2-T	CGCGCTCTTGGTCTCTGAAATCTCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-PMS2-oligo 2-T	GGGGACACCGGAAGACTGCGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-PMS2P5-oligo 2-T	CCTCGGCCATGTTCCCCCCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-POLD4-oligo 2-T	GAGGCAACTGCGGGGAAGAGGA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-PRUNE2-oligo 2-T	CCTCCTCTGCTGCAACACCTTTTCCTCTCCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-RPA2-oligo 2-T	GCCAATCGGTGCTCACGGATGCTA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-RPA4-oligo 2-T	GGCCACCTTCTCACCTGATCGCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-TAF1-oligo 2-T	AATGGGCCGCCTCCAGCGGAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-TP53-oligo 2-T	GGAAGCGTGTCACCGTCGTGGAAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-UBB-oligo 2-T	AGTTTAGGCTTTTCAACTGAGCCCCAAA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-XRCC2-oligo 2-T	TCCCAACCCGGCGGCTCTCA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-ZNF830-oligo 2-T	CGGTTGTACTTCGCGAATGGAGATTCTA**ATCACCAACTACCCACACACACC**		Same as 450-nt-oligo 2-T
940-nt-ALKBH2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGGTAAACCGCACGCAAAATTCTG	1. All oligos were designed to capture target sequences of ~940 nt in length.	Same as 450-nt-oligo 3-T
940-nt-APEX1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCTGCACCGGCCCGATG		Same as 450-nt-oligo 3-T
940-nt-APEX2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTGCCACCCCTTGCAGGGGTCT		Same as 450-nt-oligo 3-T
940-nt-ATM-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGTATGCCCATGCGCATCCAGT	2. Non-bolded nucleotides are complementary to the 3' end of target	Same as 450-nt-oligo 3-T
940-nt-ATRX-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCCCCACAGCTCAAAGGC		Same as 450-nt-oligo 3-T

940-nt-BRCA1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAGCCTGTCCCCCGTCCA	sequence.	Same as 450-nt-oligo 3-T
940-nt-BRCA2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTGACAAAAACCGCGCCGG		Same as 450-nt-oligo 3-T
940-nt-CCN4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGACGGTGCTGGCGGCT	3. Nucleotide in red is complementary to 3' end of target sequence, i.e., immediately 5' of cleavage site.	Same as 450-nt-oligo 3-T
940-nt-CD44-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGGTTTTAAGAAGTAGCAGCCCTCCC		Same as 450-nt-oligo 3-T
940-nt-CETN1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCTCAGTGAGCTCGGGCTTA		Same as 450-nt-oligo 3-T
940-nt-CHEK1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTCAGGATAACCAGTTTTGCCTTTTTTCTGTG		Same as 450-nt-oligo 3-T
940-nt-CHEK2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGACAAAGCGAAGCTCAGGAGACTCC	4. Bolded nucleotides are complementary to oligo U2.	Same as 450-nt-oligo 3-T
940-nt-CSDE1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCTTCCCCTGGGTTTCTGCTA		Same as 450-nt-oligo 3-T
940-nt-DDB2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGAGGGAGGATTGAGAGGCCTC		Same as 450-nt-oligo 3-T
940-nt-EPM2AIP1-MLH1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGTGATCTGGCGCAGAGCGGAGG	5. Used in Fig. 3c,d; Extended Data Fig. 3; Supplementary Figs. 10,11; Supplementary Tables 8,9.	Same as 300 bp-MLH1-oligo 3-T
940-nt-EPSTI1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTGCGGGTGTTCATGGTTCACAGC		Same as 450-nt-oligo 3-T
940-nt-EXO1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCGATTGCGGTTGGAGAAACC		Same as 450-nt-oligo 3-T
940-nt-FEN1-oligo 2-T	**CTACCCCACCTTCCTCATTCTCT**ACAGAGTCTCTCAGCGACCCTG		Same as 450-nt-oligo 3-T
940-nt-GSTA1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGAATTCCAGGTCCTAATGTATTTATAAGC		Same as 450-nt-oligo 3-T
940-nt-HIST1H1B-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCAGCTCTGAGACTGGGGGC		Same as 450-nt-oligo 3-T
940-nt-HOTAIR-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAACACTCACCCCACGGAGCAG		Same as 450-nt-oligo 3-T
940-nt-HOXB2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCCAATCTCCCTCTCAAATTCAAAATTCATGGC		Same as 450-nt-oligo 3-T
940-nt-HSPA5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAAATACAGGCCGCGGCGCT		Same as 450-nt-oligo 3-T
940-nt-LIG1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGGAGGAGGGCGGGAAAAG		Same as 450-nt-oligo 3-T
940-nt-MBD4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCATTCTCGAGGAAGACGCACCG		Same as 450-nt-oligo 3-T
940-nt-MCM9-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGCATTAACCACAACTGGGTAATGA		Same as 450-nt-oligo 3-T
940-nt-MECP2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGCTTTTACCACAGCCCTCTCT		Same as 450-nt-oligo 3-T
940-nt-MRE11-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCTGTCTTCTTTTCGGGAAGAAAGG		Same as 450-nt-oligo 3-T
940-nt-MSH2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCCTGGTTGAAGAAAATGCGCGACC		Same as 450-nt-oligo 3-T
940-nt-MSH4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAAGGCGAGGTTGATGAGATCTCA		Same as 450-nt-oligo 3-T
940-nt-MSH5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGCAGCGTTTTCCCGCCTTTTC		Same as 450-nt-oligo 3-T
3940-nt-OGG1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCAAAGCCCCGCTTCGTGCCCCAA		Same as 450-nt-oligo 3-T
940-nt-OLIG2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCTGCGGCTGCTGTGG		Same as 450-nt-oligo 3-T
940-nt-PELO-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGGCAGCTTCCCAGGCTCTG		Same as 450-nt-oligo 3-T
940-nt-PMS2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCCGCTCACCTCGAGC		Same as 450-nt-oligo 3-T
940-nt-PMS2P5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTCCAGGGAGGTTGGAAGCCGTG		Same as 450-nt-oligo 3-T
940-nt-POLD4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGACCGGCCAGTGGGC		Same as 450-nt-oligo 3-T
940-nt-PRUNE2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTTTCCATTAGCATACGCGCGCG		Same as 450-nt-oligo 3-T
940-nt-RPA2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGCTTGTTCCTGTCTCCTCTGC		Same as 450-nt-oligo 3-T
940-nt-RPA4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATAATCCTTCCCTGTGTGTCAGTGGTTGC		Same as 450-nt-oligo 3-T
940-nt-TAF1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCTTCGTCGCTGTCCGTGTCTGACA		Same as 450-nt-oligo 3-T
940-nt-TP53-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCACGCTCCCAGCCCGAAC		Same as 450-nt-oligo 3-T
940-nt-UBB-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTCCTCATAGCCGTAAGAAAGGCTCCT		Same as 450-nt-oligo 3-T
940-nt-XRCC2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTCGGTCCCAGACTCAGC		Same as 450-nt-oligo 3-T
940-nt-ZNF830-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCGTTTCCGACTGGTGCTCAGAC		Same as 450-nt-oligo 3-T

Supplementary Table 1: FENGC oligonucleotides used in this study; mouse ~620-nt targets.

Oligo name	Oligo sequence	Notes
620-nt-Ash1l-oligo 1-T	GCTGGCCACGGTGACCGGAATGTTTTGA**AATCACCAACTACCCACACACACC**	1. Oligos designed to capture target sequences of ~620 nt in length.
620-nt-Btk-oligo 1-T	AGGGGAGAATCGGGATCACACAAGTCA**AATCACCAACTACCCACACACACC**	
620-nt-Ccl5-oligo 1-T	GGTCAAGCCATTGGTCTTCACTTGGCAA**AATCACCAACTACCCACACACACC**	
620-nt-Cd180-oligo 1-T	TGGTTAGGAGAAGGGTAGAGCAGGGCAG**AATCACCAACTACCCACACACACC**	2. Non-bolded nucleotides are complementary to the 5' end of target sequence.
620-nt-Cox8a-oligo 1-T	ACTGTTATCCTCTCAGAAACCCGTTTCC**AATCACCAACTACCCACACACACC**	
620-nt-Cxcl15-oligo 1-T	CAGAAGAAATGGAGTCCAGACTGACACA**AATCACCAACTACCCACACACACC**	
620-nt-Cxcr4-oligo 1-T	CCGCCAGGGTTCTATACTCTGAGCTGTC**AATCACCAACTACCCACACACACC**	3. Nucleotide in red is complementary to nucleotide immediately 5' of cleavage site.
620-nt-Dnmt3b-oligo 1-T	CCTGCTTTCACCTAGCTTGGCCTTTGTC**AATCACCAACTACCCACACACACC**	
620-nt-Foxo1-oligo 1-T	CAGCCTCCACTCATCCTGAACCCCAGAG**AATCACCAACTACCCACACACACC**	
620-nt-Gdf5-oligo 1-T	GGGGCCTCCCCATCCCAGAGTAGACTC**AATCACCAACTACCCACACACACC**	4. Bolded nucleotides are complementary to oligo U1-T.
620-nt-Hdac3-oligo 1-T	GCTAGACTGGTGGTCAAAGCCTCTCCTG**AATCACCAACTACCCACACACACC**	
620-nt-Hsf1-oligo 1-T	GGAACCATGCAGGCACTTGGCAAGACC**AATCACCAACTACCCACACACACC**	
620-nt-Hsf2bp-oligo 1-T	GAGCTCCTGGCCCTGAGTCCAAGCTC**AATCACCAACTACCCACACACACC**	5. Used in Fig. 4; Extended Data Fig. 4; Supplementary Fig. 12; Supplementary Tables 10-12.
620-nt-Hsf4-oligo 1-T	GATCCCGATCCCCCACAGTCACAGT**AATCACCAACTACCCACACACACC**	
620-nt-Hsf5-oligo 1-T	AGAGGAACCCACTCCAGTCTTCCCTAAC**AATCACCAACTACCCACACACACC**	
620-nt-Hsfy2-oligo 1-T	ATCGCCCAAAGTCAGGAAAACAAA**AATCACCAACTACCCACACACACC**	
620-nt-Hsp90aa1-oligo 1-T	AGCTGCTGGAGGCCTTCCAATAGGAACA**AATCACCAACTACCCACACACACC**	
620-nt-Hsp90ab1-oligo 1-T	AAGAGTCCAGAACTCAGGGGACTGAAAGA**AATCACCAACTACCCACACACACC**	
620-nt-Hspa1a-oligo 1-T	ACTTAAGGGCTAAAACTTCTCCCCGCCA**AATCACCAACTACCCACACACACC**	
620-nt-Hspa1b-oligo 1-T	TGGTGAGAGTGTGGGACTCAGACCAGGGA**AATCACCAACTACCCACACACACC**	
620-nt-Hspa2-oligo 1-T	CGGTTGCATCGAATACAGAACGGCTGGGA**AATCACCAACTACCCACACACACC**	
620-nt-Ifna1-oligo 1-T	CAAGGTATGGTTGCATGTTTCTCTCA**ATCACCAACTACCCACACACACC**	
620-nt-Ifna2-oligo 1-T	CTAAGGTAGAAGTATTGTTGTGGCCTGGA**ATCACCAACTACCCACACACACC**	
620-nt-Ifng-oligo 1-T	GCCTTTGTTTGTGGTTGGGGACGGCTGA**AATCACCAACTACCCACACACACC**	
620-nt-Ikbkb-oligo 1-T	GAGGCTGGGTATTGTCTTCTCCACCAGGA**AATCACCAACTACCCACACACACC**	
620-nt-Ikbkg-oligo 1-T	AAGTCATAACACCGATCGCTCCCAGCAC**AATCACCAACTACCCACACACACC**	
620-nt-Il10-oligo 1-T	AACGTCCGATATTTCTTGCTGACTCGGA**AATCACCAACTACCCACACACACC**	
620-nt-Il12a-oligo 1-T	ACCATCTCCATAACCACAAGCCTTTGGG**AATCACCAACTACCCACACACACC**	
620-nt-Il12b-oligo 1-T	CCACCTTTCCCCATGTTTTCTTGTTAGC**AATCACCAACTACCCACACACACC**	
620-nt-Il3ra-oligo 1-T	TCCTGTCATCCGTTGCAACTGTACTCGC**AATCACCAACTACCCACACACACC**	
620-nt-Il6-oligo 1-T	AGACACACACTCCCTCTCCCACCCCC**AATCACCAACTACCCACACACACC**	
620-nt-Irak1-oligo 1-T	TGCTTGTGTGCTAGATCTCCATGCCCCC**AATCACCAACTACCCACACACACC**	
620-nt-Irak4-oligo 1-T	CCTGTCTGCCCGCCTCTTG**AATCACCAACTACCCACACACACC**	
620-nt-Irf3-oligo 1-T	CGTTTGCCGCAAAGCATACTAGGATT**AATCACCAACTACCCACACACACC**	
620-nt-Irf5-oligo 1-T	CTTGGGCCACTCTGTCTCCTCAGCCTTA**AATCACCAACTACCCACACACACC**	
620-nt-Irf7-oligo 1-T	CAGGAAAGCACATTGGGACAAGGCAGGG**AATCACCAACTACCCACACACACC**	
620-nt-Jun-oligo 1-T	ACTTCCGGGGGACAGAGCG**AATCACCAACTACCCACACACACC**	
620-nt-Kdm6b-oligo 1-T	CTGCTACCTCCAACCCTAACAAGACACC**AATCACCAACTACCCACACACACC**	
620-nt-Lbp-oligo 1-T	AACCTTGCAATGGGATAGCTGTCTGAAC**AATCACCAACTACCCACACACACC**	
620-nt-Ly86-oligo 1-T	AACAGGAGCTTGTCCTAGGAGCCCACAC**AATCACCAACTACCCACACACACC**	
620-nt-Map3k1-oligo 1-T	AGAGGTTACTCCCCCGTCTTGGAAAGGC**AATCACCAACTACCCACACACACC**	
620-nt-Mapk1-oligo 1-T	GTGGGACAGAAACCAAAAAGCATAACCC**AATCACCAACTACCCACACACACC**	
620-nt-Mapk13-oligo 1-T	AAGGGAGAGGAGAGACTAGAGAGGAGAG**AATCACCAACTACCCACACACACC**	
620-nt-Mapk14-oligo 1-T	CAGGCGCTCCTACATGTTTTCAGGCAGA**AATCACCAACTACCCACACACACC**	
620-nt-Mapk15-oligo 1-T	GCTTTGTGCTCAGGTTCAGGGGAC**AATCACCAACTACCCACACACACC**	
620-nt-Mapk3-oligo 1-T	GCCTGAGCCATCATACAGATTCAGGA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk4-oligo 1-T	AAGTAAAACTCACAGATGGGAACCGAG**AATCACCAACTACCCACACACACC**	
620-nt-Mapk6-oligo 1-T	GAATGAGTGAGGGGTCCTGGGAAAACTG**AATCACCAACTACCCACACACACC**	
620-nt-Mapk8-oligo 1-T	TTCAAGGGCCTACCACCCACCAAGGAGT**AATCACCAACTACCCACACACACC**	
620-nt-Mme-oligo 1-T	AGTGCTGATTGCAGGCTCTCATTCTCTA**AATCACCAACTACCCACACACACC**	
620-nt-Myd88-oligo 1-T	CGGGGTCTGCAAAAAAACACGCGGAACCT**AATCACCAACTACCCACACACACC**	
620-nt-Nfkb1-oligo 1-T	AGCCTGGGGCTAGGTAGCTCCCAAGTG**AATCACCAACTACCCACACACACC**	
620-nt-Nfkb2-oligo 1-T	CAAACTTCCTCTTCTGTGGAAAATTGCC**AATCACCAACTACCCACACACACC**	
620-nt-Nfkbia-oligo 1-T	CGCTTCTCCTAAATTTCTGCGACGACGG**AATCACCAACTACCCACACACACC**	
620-nt-Nfkbib-oligo 1-T	CGCATTTCTCGTACTTTGGCTGCTGAGG**AATCACCAACTACCCACACACACC**	
620-nt-Nr2c2-oligo 1-T	GTGCCCCTCTCCTTCGCGATTTGAGAAAA**AATCACCAACTACCCACACACACC**	
620-nt-Pik3ap1-oligo 1-T	ATAGGGAAGGCCTGCAGGAAGGGGAAGG**AATCACCAACTACCCACACACACC**	
620-nt-Pik3c2a-oligo 1-T	TGAGCGCGGACTGCAATTGGGCTTGTAAA**AATCACCAACTACCCACACACACC**	
620-nt-Pik3c2g-oligo 1-T	TGTGGGAAATCAAAGAGGACTCTTGGAAA**AATCACCAACTACCCACACACACC**	
620-nt-Pik3c3-oligo 1-T	CGGGTTTTGAATCTTCTGAGTGCCTGTAA**AATCACCAACTACCCACACACACC**	
620-nt-Pik3ca-oligo 1-T	CTTTCCGCTCACATCGTGTAAACAAACAA**AATCACCAACTACCCACACACACC**	
620-nt-Pik3r1-oligo 1-T	CTGCCTCTTTCGAATCCACTAGAGCTTC**AATCACCAACTACCCACACACACC**	
620-nt-Pik3r2-oligo 1-T	CCGCCGATAGCTTCCGGAAATTGCCGAA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3r3-oligo 1-T	ACCCCCTAGCCTCGCTCAGAATGGTAC**AATCACCAACTACCCACACACACC**	
620-nt-Pik3r5-oligo 1-T	ACTTCCAGCACTGCTTCTGCAACCTCTG**AATCACCAACTACCCACACACACC**	
620-nt-Pik3r6-oligo 1-T	TGTCCCCCACATTTGCATCTTAGACCAC**AATCACCAACTACCCACACACACC**	
620-nt-Ptpn6-oligo 1-T	AAAGGGTCTAAAGGAGGGGAAAGGCTAG**AATCACCAACTACCCACACACACC**	
620-nt-Rela-oligo 1-T	GAGGACTAGGCAGACACCTGTACCTCTAA**AATCACCAACTACCCACACACACC**	
620-nt-Src-oligo 1-T	ACGTCATCTCATCGTGTACCTCTTGCCA**ATCACCAACTACCCACACACACC**	
620-nt-Stat3-oligo 1-T	TTTGTGAGCTCCTTACCACGTGGGAATCA**ATCACCAACTACCCACACACACC**	
620-nt-Tab1-oligo 1-T	ACCCCAATGCCAAACTAGCTGAAGTTAG**AATCACCAACTACCCACACACACC**	
620-nt-Tab2-oligo 1-T	AGTTTGGCTCGGCCGCTGGA**ATCACCAACTACCCACACACACC**	
620-nt-Tet2-oligo 1-T	GCCAAGCCGCTCCCCATTTCCCCAAA**ATCACCAACTACCCACACACACC**	
620-nt-Tlr2-oligo 1-T	CCTCAGAGGTTGCAACTTGGGCTCAGGA**AATCACCAACTACCCACACACACC**	
620-nt-Tlr4-oligo 1-T	TGTTCATTCGTTCACGGGGGGGATGTAT**AATCACCAACTACCCACACACACC**	
620-nt-Tnf-oligo 1-T	ACTGAAACCCATTTCTTCTCTGTCCTCC**AATCACCAACTACCCACACACACC**	
620-nt-Traf6-oligo 1-T	AGATGCCCAAAAGATTTTCACCCTC**AATCACCAACTACCCACACACACC**	
620-nt-Vav1-oligo 1-T	GTTCAACCTTGCTTTGGGGCCAAAACTC**AATCACCAACTACCCACACACACC**	
620-nt-Ash1l-oligo 2-T	AGAGCGCCGAGTGCAGGCAA**ATCACCAACTACCCACACACACC**	
620-nt-Btk-oligo 2-T	AGTTCTGACTTAATGCAGGTAGCTCCCC**AATCACCAACTACCCACACACACC**	
620-nt-Ccl5-oligo 2-T	GCATCGAGGCAGAGTGGGGCTGTGA**ATCACCAACTACCCACACACACC**	
620-nt-Cd180-oligo 2-T	GAAGAGGGCCACTAAAAAGAAGCAGCTGA**ATCACCAACTACCCACACACACC**	
620-nt-Cox8a-oligo 2-T	TCAGCAGCAGTGGCGTCAGGACAGAC**ATCACCAACTACCCACACACACC**	
620-nt-Cxcl15-oligo 2-T	AGAAAAATTCCCCGTTAGCTTACCGAA**ATCACCAACTACCCACACACACC**	
620-nt-Cxcr4-oligo 2-T	GGATGCAAGTGGACTTACACTCACACTG**AATCACCAACTACCCACACACACC**	
620-nt-Dnmt3b-oligo 2-T	AAAAGTTGCTGCAGGTCTCCTTCCCTCA**AATCACCAACTACCCACACACACC**	
620-nt-Foxo1-oligo 2-T	CGCTGGGTCACCGGTGTCTA**ATCACCAACTACCCACACACACC**	
620-nt-Gdf5-oligo 2-T	GTGCACCGTCTCCAGTCAGCAGCTGAAAA**ATCACCAACTACCCACACACACC**	
620-nt-Hdac3-oligo 2-T	GCCCACATCGGGGTCGTAGAAATACGCC**ATCACCAACTACCCACACACACC**	

620-nt-Hsf1-oligo 2-T	GGCCTTTTAGAAGAGTGGCTGAAGGCTGA**ATCACCAACTACCCACACACACC**	
620-nt-Hsf2bp-oligo 2-T	ATCAAACCAGCCGTAGGCGTTCTGCTGGA**ATCACCAACTACCCACACACACC**	
620-nt-Hsf4-oligo 2-T	GCTGGGTGGACAGGGGTGTTCACGA**ATCACCAACTACCCACACACACC**	
620-nt-Hsf5-oligo 2-T	CTGAGCAGCTCGGCCTCGAAGA**ATCACCAACTACCCACACACACC**	
620-nt-Hsfy2-oligo 2-T	ACTGTCTTCATCTGGCTCTGAGGCACAGA**ATCACCAACTACCCACACACACC**	
620-nt-Hsp90aa1-oligo 2-T	ATCTGAGGTGAACTCACTCTTACCCGCTA**ATCACCAACTACCCACACACACC**	
620-nt-Hsp90ab1-oligo 2-T	GTACAGCCCGAACCTCTAACCTCTAGACA**ATCACCAACTACCCACACACACC**	
620-nt-Hspa1a-oligo 2-T	GCTGGAACACGCCCACGCACGAGTA**ATCACCAACTACCCACACACACC**	
620-nt-Hspa1b-oligo 2-T	GCTGGAACACGCCCACGCACGAGTA**ATCACCAACTACCCACACACACC**	
620-nt-Hspa2-oligo 2-T	AAGCGCAAGCCGGGACCTGA**ATCACCAACTACCCACACACACC**	
620-nt-Ifna1-oligo 2-T	TCAGCTCACTCAGGACAGGGATGGCTTGA**ATCACCAACTACCCACACACACC**	
620-nt-Ifna2-oligo 2-T	TGACCAGTAGCTCATCACTATCAGCATCA**ATCACCAACTACCCACACACACC**	
620-nt-Ifng-oligo 2-T	AGCGTTCATTGTCTCAGAGCTAGGCCGCA**ATCACCAACTACCCACACACACC**	
620-nt-Ikbkb-oligo 2-T	AGGGGAGGCTAGGACACCGGTGTAGTTAA**ATCACCAACTACCCACACACACC**	
620-nt-Ikbkg-oligo 2-T	TTCATCTTGAAGCAGTGACACTGTAGGAA**ATCACCAACTACCCACACACACC**	
620-nt-Il10-oligo 2-T	CCCCTGCTGATCCTCATGCCAGTCAGTAA**ATCACCAACTACCCACACACACC**	
620-nt-Il12a-oligo 2-T	CGGAGCGCGTTGCAAGTTGTGATA**ATCACCAACTACCCACACACACC**	
620-nt-Il12b-oligo 2-T	AGATGCAGGGAGTTAGCGACAGGGAAGA**ATCACCAACTACCCACACACACC**	
620-nt-Il3ra-oligo 2-T	CGCGAGTAAACAGTGTGAGGCAGGGGGA**ATCACCAACTACCCACACACACC**	
620-nt-Il6-oligo 2-T	GTGGTTGTCACCAGCATCAGTCCCAAGAA**ATCACCAACTACCCACACACACC**	
620-nt-Irak1-oligo 2-T	GGCTCCAGGGCGTCCATCACTTTGTAGAA**ATCACCAACTACCCACACACACC**	
620-nt-Irak4-oligo 2-T	GGGACAGAGGACAGCGGAGAGA**ATCACCAACTACCCACACACACC**	
620-nt-Irf3-oligo 2-T	GCTCCCAAACAGTAGTTCTGAGACTCGCA**ATCACCAACTACCCACACACACC**	
620-nt-Irf5-oligo 2-T	AGGCGCAGTGGGGATGTCCA**ATCACCAACTACCCACACACACC**	
620-nt-Irf7-oligo 2-T	GATCACTCACTACTCTGTAGAGGGGACAA**ATCACCAACTACCCACACACACC**	
620-nt-Jun-oligo 2-T	TGTCTCTGCCACACTCAGTGCAACTCTGA**ATCACCAACTACCCACACACACC**	
620-nt-Kdm6b-oligo 2-T	CAAGCTTACCAGATGGAGTTGGAAGTGGA**ATCACCAACTACCCACACACACC**	
620-nt-Lbp-oligo 2-T	CTGTATGGACAGGAACAGCAGCCCCA**ATCACCAACTACCCACACACACC**	
620-nt-Ly86-oligo 2-T	AAGCCAGCCCAATAAAAAGACCTCCGAA**ATCACCAACTACCCACACACACC**	
620-nt-Map3k1-oligo 2-T	CGCCATTTTCTCTCGCGGGCTA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk1-oligo 2-T	GGCGATCGGGAACGAGGAAGGAGGACAA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk13-oligo 2-T	CCCACGAACTCCGAGATCCCTTTTAAAAA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk14-oligo 2-T	CAGATGGTCTTGTTCAGCTCCTGCCGGTA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk15-oligo 2-T	GGGCCAGATTCTCAGGGAAGCCAGGGA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk3-oligo 2-T	GAAGTCACTGCACGATACTTCCCCTGCGA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk4-oligo 2-T	GGCAGCGAGAGAAAGAAAGCGGGCGA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk6-oligo 2-T	ACACACTCACGGGACTTCTCCGACTCA**ATCACCAACTACCCACACACACC**	
620-nt-Mapk8-oligo 2-T	TCACGTCCGCACAAAGAAACCGCCCA**ATCACCAACTACCCACACACACC**	
620-nt-Mme-oligo 2-T	GGTGACTCACCAGTGGGATTGGGGA**ATCACCAACTACCCACACACACC**	
620-nt-Myd88-oligo 2-T	GTTAAGCGCGACCAAGGGTATGGAGAACA**ATCACCAACTACCCACACACACC**	
620-nt-Nfkb1-oligo 2-T	CTGGCTGGAAATTCCCACTGACGTCGAGA**ATCACCAACTACCCACACACACC**	
620-nt-Nfkb2-oligo 2-T	GTGCCCAGAGGTTGGCTATCTCTGAGACA**ATCACCAACTACCCACACACACC**	
620-nt-Nfkbia-oligo 2-T	CACCATTTGCTCGTACTCCTCGTCCTTCA**ATCACCAACTACCCACACACACC**	
620-nt-Nfkbib-oligo 2-T	CGCCCCTGTACTGCACCACAGCCAA**ATCACCAACTACCCACACACACC**	
620-nt-Nr2c2-oligo 2-T	GGGACCGCGCTTTGTTATTGTTTTTGTCA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3ap1-oligo 2-T	CTAAAGAAGTAGGGAGACAGGTCTCAAGA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3c2a-oligo 2-T	GCCCCGAGGAAGGAGTCAGAGA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3c2g-oligo 2-T	TGGAAGGCAGGTCCAGGCCAGCAGA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3c3-oligo 2-T	CAGGAAGAAACCGCCGCTGCATCAAAACA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3ca-oligo 2-T	CCTCCGCTGGTCTTCCATCTACCCACGTA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3r1-oligo 2-T	TCGGCTCGCTTCCTTCCTCTGCCCA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3r2-oligo 2-T	GGTATCGCGGAACCCCCAATGCAAAAGGA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3r3-oligo 2-T	AAAAGCTCCCACCGCCTCCTAATCTTTCA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3r5-oligo 2-T	CCTACCGGCTGCAGTTGGAGCACACTA**ATCACCAACTACCCACACACACC**	
620-nt-Pik3r6-oligo 2-T	CAGCCAATGAGCTCCTGAGGTCCTGTTCA**ATCACCAACTACCCACACACACC**	
620-nt-Ptpn6-oligo 2-T	ATTTCGTTTCTTGGCCACTTCCTCCTCTA**ATCACCAACTACCCACACACACC**	
620-nt-Rela-oligo 2-T	CGGCGCGCTAAAGTAAAGCCATTCGCCA**ATCACCAACTACCCACACACACC**	
620-nt-Src-oligo 2-T	ACCTGAAGGAGAAACTCACCCAGGACA**ATCACCAACTACCCACACACACC**	
620-nt-Stat3-oligo 2-T	CGACACCGAGCGGGGTTGTACCTCA**ATCACCAACTACCCACACACACC**	
620-nt-Tab1-oligo 2-T	CGAGGCGGTCCACACAATGCCAGAGA**ATCACCAACTACCCACACACACC**	
620-nt-Tab2-oligo 2-T	TCCGCATCAAAATACGGCCACA**ATCACCAACTACCCACACACACC**	
620-nt-Tet2-oligo 2-T	CCTTTCTAACAAATCCTACAGGGCAGCCA**ATCACCAACTACCCACACACACC**	
620-nt-Tlr2-oligo 2-T	CACAGGGGTCTGCAAGGGGAACGGA**ATCACCAACTACCCACACACACC**	
620-nt-Tlr4-oligo 2-T	CAGGGAGGCATCATCCTGGCA**ATCACCAACTACCCACACACACC**	
620-nt-Tnf-oligo 2-T	CTGTGCTCATGGTGTCTTTTCTGGAGGGA**ATCACCAACTACCCACACACACC**	
620-nt-Traf6-oligo 2-T	ACTCTAGTTCGAAGCCGCTCACTCACCGA**ATCACCAACTACCCACACACACC**	
620-nt-Vav1-oligo 2-T	TGGCCTGGGGAAGCAGGTTGTTAAGCAA**ATCACCAACTACCCACACACACC**	
620-nt-Ash1l-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTGGAGAGAGAGAGGGTGGCAGCCTG	1. Oligos were designed to capture target sequences of ~620 nt in length.
620-nt-Btk-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGACGTCGTGAGATGCCAGGACTTGGAAG	
620-nt-Ccl5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGTCAGAACCCTCGACTTACATGGTGA	
620-nt-Cd180-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGTCCGGAGCCATCCCAGGTCCAGAG	2. Non-bolded nucleotides are complementary to the 3' end of target sequence.
620-nt-Cox8a-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGACGGCGCGGAACGACCA	
620-nt-Cxcl15-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGATGCCTAGGTTAAGGACAGCCAGCAGG	
620-nt-Cxcr4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGGTTCCATGGCAACACTCGCTCTACA	3. Nucleotide in red is complementary to 3' end of target sequence, i.e., immediately 5' of cleavage site.
620-nt-Dnmt3b-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTCTTCAGAGTGGCGGAGTGCAGCG	
620-nt-Foxo1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGAGCCTCCGAGTTCGTGCCTAGGAC	
620-nt-Gdf5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATAACTCGTTCTTGAAAGGAGAAAGCCGA	
620-nt-Hdac3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGGTCTTGGCCATGGTGCTGG	4. Bolded nucleotides are complementary to oligo U2.
620-nt-Hsf1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTAGTGTCGTCGCTGCCCCT	
620-nt-Hsf2bp-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGAGACCTGACCTCCAACTAACCGCCCT	
620-nt-Hsf4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACATTGGAGGGATCCCGGGGACCACTG	5. Used in Fig. 4; Extended Data Fig. 4; Supplementary Fig. 12; Supplementary Tables 10-12.
620-nt-Hsf5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGGCTGGTCGATGAGCAG	
620-nt-Hsfy2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGTATGCGGTGGCCTCTTTATCAAGAA	
620-nt-Hsp90aa1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACCAGGGAGGGCACACAAGCACC	
620-nt-Hsp90ab1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGTTATCTCCACAGGCTGCCCCGC	
620-nt-Hspa1a-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGTGGTGCCCAGGTCGATGCCGAT	
620-nt-Hspa1b-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGTGGTGCCCAGGTCGATGCCGAT	
620-nt-Hspa2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCAACGGTCGCTTCTCACAACTCTCCA	
620-nt-Ifna1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCTTCTTGATCTGCTGGGCATCCAC	
620-nt-Ifna2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACGAGGAAAGCACAGAGTCTAGCCATTGT	

620-nt-Ifng-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGAGGAGAAGCCCAGAACTTCTGTCTCA	
620-nt-Ikbkb-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAAATAAAAGTTGCGGCGGCCGGGC	
620-nt-Ikbkg-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTAGGCGGCTCAGCTTTAGCACAAGAG	
620-nt-Il10-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGCAGGCAGCATAGCAGTGCT	
620-nt-Il12a-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGGTGGGGATATATCTGGTCATTGCCG	
620-nt-Il12b-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGAGAGAACTTACCTTGCTTTGCTGCGA	
620-nt-Il3ra-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGACCACGAGTTAGTTTCTGAGGCAGGG	
620-nt-Il6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGCAACTGGATGGAAGTCTCCTGCGTGG	
620-nt-Irak1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGGCACATAACCCAGGGTGGCAC	
620-nt-Irak4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGTGGCTTGAGATCAGCCTTCCTCAT	
620-nt-Irf3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCACGCAACCCCCTTCTTTGTCAGC	
620-nt-Irf5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGGGCCCTTGGGACTTTCC	
620-nt-Irf7-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAATTTCCCCTGTGGCACTAGCTCACCT	
620-nt-Jun-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTCCTTATCCAGCCTGAGCTCAACACTT	
620-nt-Kdm6b-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGTCTCATGGAACAAGCACTCCAGCAG	
620-nt-Lbp-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCAATGTAGAGAGCAGAGGTCCAGTCCC	
620-nt-Ly86-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCTCGTCCTCCCCCTGAAGGAGTAGT	
620-nt-Map3k1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACATTCGCTCGTGCCTGC	
620-nt-Mapk1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACAGAGGAGAGAATTAACCGCCGGTAG	
620-nt-Mapk13-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTTACAGCCTTGGAGAACCAGATTCGCC	
620-nt-Mapk14-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAACGTGGGCCTCTCCTGCGACATCTTC	
620-nt-Mapk15-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACCTGTGGGCACAAGGTGGAGCCAAC	
620-nt-Mapk3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACACTCCCTGAGCCCCCCATGT	
620-nt-Mapk4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCAGGGAGCCACTCACCGA	
620-nt-Mapk6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGACAGAGTGCAGGGCCC	
620-nt-Mapk8-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAGTCAAAACACGCTCCCCGTGGGTG	
620-nt-Mme-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAAGTTGGAGGCGACTGAGCAGACACA	
620-nt-Myd88-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGAAGGAGTCCAGGGACCCGGATCCC	
620-nt-Nfkb1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGAGCAGACAGACGGACGGACCCAC	
620-nt-Nfkb2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAGGAAACCGGGTCAGTCCC	
620-nt-Nfkbia-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGGAGTCCAGGCCGCTGTCGT	
620-nt-Nfkbib-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCAGAAAAATGAGGTGGACATTCGCTCG	
620-nt-Nr2c2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGTGAAAGGTCAGAGTTCGTGACCCCGGT	
620-nt-Pik3ap1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTGGGAGGGTGGTGTCAGAGAGCCTCA	
620-nt-Pik3c2a-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGATGGAGTGACCAAGGGGCACGG	
620-nt-Pik3c2g-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTGTGCTGTGGGTGGCAGAGTCAATCC	
620-nt-Pik3c3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATGAAAACTTCCTGTTCCGCCCATCTGCC	
620-nt-Pik3ca-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGCGGAGGCTGACACCGT	
620-nt-Pik3r1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCTGGGCTTCAGCGGCCTCTGAA	
620-nt-Pik3r2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGGGCCTACGATGCACGCATAGTCTGTG	
620-nt-Pik3r3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGGAAGCCACTTTTGTGTCCTTCG	
620-nt-Pik3r5-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACTCCAGAATGGTCCCAACCTGCTCTGGC	
620-nt-Pik3r6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGGAGGCCAGCAACCCCAAGGACCA	
620-nt-Ptpn6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCCAGCCTGTCAAACCAATGGGACTCTG	
620-nt-Rela-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGAGGCGGAAATGCGCG	
620-nt-Src-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCCTTCCTCCAGCCACGTCTGCA	
620-nt-Stat3-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCGATCCGGTTAGGGCTCG	
620-nt-Tab1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATACAGCACGCCTTACACAAGCGGGGG	
620-nt-Tab2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ACAATATGGCGGCGCGG	
620-nt-Tet2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCTTGCTTCCTTCTGCTCCTCTACGAAA	
620-nt-Tlr2-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGCCTGGCATCAGAGCTGGCAGGAG	
620-nt-Tlr4-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTTTCTGCAACCAGTGGAGGCCATGAGA	
620-nt-Tnf-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**AGATGTGGCGCCTTGGGCCAGTGAG	
620-nt-Traf6-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATCGTACACGCTGAGACTTCACAACCAAC	
620-nt-Vav1-oligo 3-T	**CTACCCCACCTTCCTCATTCTCT**ATTGGCACAAGAGGACACCGTCCC	

Supplementary Table 2: Characteristics of designed FENGC target sequences used in this study; human ~300-nt targets.

Gene	Genome coordinates	Captured length (nt)	Captured sequence GC content (%)	Tm			Gene/product name [a,b]	Captured sequence relative to TSS (+1)	Notes
				Oligo 1-T	Oligo 2-T	Oligo 3-T			
EPM2AIP1	chr3:36,992,888-36,993,182	295	62	69	70	71	EPM2A-interacting protein 1	+307 to +13 of EMPAIP1 TSS; -629 to -335 of MLH1 TSSa	Transcribed divergently from *MLH1*
EXO1	chr1:241,847,874-241,848,155	282	66	66	69	69	Exonuclease 1	-317 to -36	
HSPA5	chr9:125,241,405-125,241,709	305	61	64	70	67	Endoplasmic reticulum chaperone BiP	-322 to -18	Control promoter with open chromatin
MLH1	chr3:36,993,504-36,993,803	300	60	61	69	70	DNA mismatch repair protein Mlh1	-609 to -310 of EMPAIP1 TSS; -13 to +287 of MLH1 TSSa	Transcribed divergently from *EMP2AIP1*
MLH3	chr14:75,051,561-75,051,853	293	36	64	66	65	DNA mismatch repair protein Mlh3	-321 to -29	
MSH2	chr2:47,402,729-47,403,020	292	56	65	67	67	DNA mismatch repair protein MSH2	-338 to -47	
MSH3	chr5:80,654,325-80,654,639	315	70	68	71	71	DNA mismatch repair protein MSH3	-323 to -9	
MSH6	chr2:47,782,759-47,783,061	303	78	65	72	74	DNA mismatch repair protein MSH6	-323 to -21	
PMS1	chr2:189,783,763-189,784,054	292	62	62	69	70	PMS1 protein homoloq 1	-322 to -31	
PMS2	chr7:6,009,126-6,009,426	301	61	64	67	71	Mismatch repair endonuclease PMS2	-320 to -20 of PMS2 TSS; -146 to +155 of AIMP2 TSS	Transcribed divergently from *AIMP2*
TDG	chr12:103,965,499-103,965,791	293	56	64	69	67	G/T mismatch-specific thymine DNA glycosylase	-316 to -24	

[a] Gene Ontology term "metabolic process," filtered for "DNA repair"

[b] Control loci with open chromatin

Supplementary Table 2: Characteristics of designed FENGC targets used in this study; human ~450-nt targets.

Gene	Genome coordinates (hg38)	Captured length (nt)	Captured sequence GC content (%)	Tm			Captured sequence relative to TSS (+1)	Gene/product name	Notes
				Oligo 1-T	Oligo 2-T	Oligo 3-T			
ALKBH1	chr14:77,707,866-77,708,314	449	57	64	69	64	+158 to -291 of ALKBH1 TSS; -205 to +244 of SLIRP1 TSS	Nucleic acid dioxygenase ALKBH1	Transcribed divergently from *SLIRP1*
ALKBH2	chr12:109,093,335-109,093,770	436	62	65	65	65	-282 to +154	DNA oxidative demethylase ALKBH2	
ALKBH3	chr11:43,880,520-43,880,969	450	62	64	68	68	-311 to +193	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3	
APEX1	chr14:20,454,833-20,455,278	446	58	67	64	64	-393 to +53	DNA-(apurinic or apyrimidinic site) endonuclease	
APEX2	chrX:55,000,041-55,000,485	445	59	66	67	68	-322 to +123	DNA-(apurinic or apyrimidinic site) endonuclease 2	
ATM	chr11:108,222,537-108,222,984	448	61	62	68	68	+102 to -346 of NPAT TSS; -530 to -83 of ATM TSS	Serine-protein kinase ATM	Transcribed divergently from *NPAT*
ATRX	chrX:77,786,121-77,786,569	449	49	64	64	65	-353 to +96	Transcriptional regulator ATRX	
ATXN3	chr14:92,106,418-92,106,858	441	67	64	70	69	-276 to +165	Ataxin-3	
AXIN2	chr17:65,561,578-65,562,018	441	66	64	66	68	-370 to +71	Axin-2	
BRCA1	chr17:43,125,334-43,125,779	446	52	62	65	67	-415 to +31	Breast cancer type 1 susceptibility protein	
BRCA2	chr13:32,315,183-32,315,631	449	63	68	64	67	-325 to +124	Breast cancer type 2 susceptibility protein	
CCN4	chr8:133,190,760-133,191,206	447	56	65	69	69	-279 to +168	WISP1 isoform 1	
CD44	chr11:35,138,612-35,139,052	441	62	66	68	64	-559 to -119	CD44 antigen	TSS of ENST00000263398.11; same as RefSeq and only isoform expressed at detectable levels in brain (dbGaP accession phs000424.v8.p2 on 09/22/2022)
CDH1	chr16:68,736,898-68,737,344	447	60	67	68	65	-394 to +53	Cadherin-1	
CDH13	chr16:82,626,498-82,626,943	446	50	65	68	67	-471 to -26	Cadherin-13	
CDH2	chr18:28,177,299-28,177,747	449	73	69	66	70	-300 to +149	Cadherin-2	
CDKN2A	chr9:21,994,358-21,994,794	448	75	68	67	70	-402 to +46	Cyclin-dependent kinase inhibitor 2A	
CDKN2D	chr19:10,568,868-10,569,301	434	69	66	69	67	-345 to +89	Cyclin-dependent kinase 4 inhibitor D	
CETN1	chr18:580,059-580,492	434	63	60	62	62	-284 to +150	Centrin-1	
CHEK1	chr11:125,624,871-125,625,316	446	45	66	63	62	-265 to +181	Serine/threonine-protein kinase Chk1	
CHEK2	chr22:28,741,673-28,742,109	437	59	68	66	65	-274 to +163	Serine/threonine-protein kinase Chk2	
CSDE1	chr1:114,757,871-114,758,321	451	56	65	64	67	-271 to +180	Cold shock domain-containing protein E1	
DDB2	chr11:47,214,647-47,215,093	447	63	61	62	65	-327 to +120	DNA damage-binding protein 2	
DRD3	chr3:114,199,258-114,199,704	447	39	57	56	59	-297 to +150	D(3) dopamine receptor	
DRD4	chr11 KI270832v1 alt:166,876-1	448	76	65	68	69	-302 to +146	D(4) dopamine receptor	
EGFR	chr7:55,018,634-55,019,077	444	70	65	69	69	-383 to +61	Epidermal growth factor receptor	
EPM2AIP1	chr3:36,992,888-36,993,337	450	60	69	69	69	+307 to -143 of EMP2AIP1 TSS; -629 to -180 of MLH1 TSSa	EPM2A-interacting protein 1	Transcribed divergently from *MLH1*
EPSTI1	chr13:42,992,151-42,992,580	430	55	66	66	65	-267 to +163	Epithelial Stromal Interaction 1	
ERCC1	chr19:45,423,763-45,424,212	450	59	64	69	69	-292 to +158	DNA excision repair protein ERCC-1	
ERCC2	chr19:45,370,469-45,370,913	445	60	67	68	65	-325 to +120	General transcription and DNA repair factor IIH helicase subunit XPD	
ERCC8	chr5:60,944,930-60,945,380	451	60	62	61	63	+141 to -310 of ERCC8 TSS; -275 to +176 of NDUFAF2 TSS	DNA excision repair protein ERCC-8	Transcribed divergently from *NDUFAF2*
EXO1	chr1:241,847,874-241,848,317	444	64	66	66	65	-317 to +127	Exonuclease 1	
FANCC	chr9:95,317,552-95,317,993	442	63	60	67	67	-283 to +159	Fanconi anemia group C protein	
FAT4	chr4:125,316,119-125,316,562	446	54	61	65	67	-293 to +153	FAT Atypical Cadherin 4	Positive control for open chromatin
FEN1	chr11:61,792,341-61,792,787	447	63	67	66	63	-570 to -124	Flap endonuclease 1; transcribed divergently from TMEM258	
FHIT	chr3:61,251,306-61,251,738	433	58	65	67	66	-276 to +157	Bis(5'-adenosyl)-triphosphatase	
FOXC2	chr16:86,566,961-86,567,408	448	78	70	70	63	-290 to +158	Forkhead box protein C2	
FZD7	chr2:202,034,262-202,034,691	430	76	70	67	69	-325 to +105	Frizzled-7	
FZD8	chr10:35,641,206-35,641,640	435	69	68	67	69	-205 to +230	Frizzled-8	
GPR158	chr10:25,175,048-25,175,483	436	61	66	70	66	-313 to +123	G-protein coupled receptor 158	
GSTA1	chr6:52,803,818-52,804,264	447	41	60	65	58	-449 to -2	Glutathione S-transferase A1	
HIST1H1B	chr6:27,867,390-27,867,834	445	53	61	65	64	-246 to +199	Histone H1.5	
HMGA2	chr12:65,824,153-65,824,600	448	67	69	65	69	-307 to +141	High mobility group protein HMGI-C	
HOTAIR	chr12:53,968,606-53,969,054	449	48	65	68	67	-297 to +152	Homo sapiens HOX transcript antisense RNA, transcript variant 1	
HOXB2	chr17:48,544,881-48,545,329	449	51	63	64	64	-297 to +152	Homeobox protein Hox-B2	
HSPA5	chr9:125,241,260-125,241,709	450	61	64	66	67	-322 to +128	Endoplasmic reticulum chaperone BiP	Open control locus
INO80	chr15:41,116,157-41,116,596	440	68	68	69	69	-349 to +91	Chromatin-remodeling ATPase INO80	
KLF8	chrX:56,232,058-56,232,505	448	57	63	64	64	-298 to +150	Krueppel-like factor 8	
LIG1	chr19:48,170,440-48,170,885	446	72	69	65	65	-281 to +165	DNA ligase 1	
MBD4	chr3:129,440,032-129,440,478	447	65	68	65	65	-298 to +149	Methyl-CpG-binding domain protein 4	
MCM9	chr6:118,931,590-118,932,036	447	32	61	61	60	-297 to +150	DNA helicase MCM9	
MECP2	chrX:154,097,673-154,098,102	430	69	66	68	65	-370 to +60	Methyl-CpG-binding protein 2	
MEG3	chr14:100,825,821-100,826,269	449	60	66	67	63	-287 to +162	Maternally Expressed 3	
MGMT	chr10:129,466,900-129,467,344	445	76	67	69	64	-284 to +161	Methylated-DNA--protein-cysteine methyltransferase	
MLH1	chr3:36,993,504-36,993,941	438	56	61	55	56	-310 to -747 of EMPAIP1 TSS; -13 to +425 of MLH1 TSSa	DNA mismatch repair protein Mlh1	Transcribed divergently from *EPMAIP1*
MLH3	chr14:75,051,407-75,051,853	447	47	64	61	69	-321 to +126	DNA mismatch repair protein Mlh3	
MNAT1	chr14:60,734,448-60,734,889	442	50	63	61	65	-293 to +149	CDK-activating kinase assembly factor MAT1	
MPG	chr16:77,798-78,244	447	75	69	68	65	-372 to +75	DNA-3-methyladenine glycosylase	
MRE11	chr11:94,493,724-94,494,172	449	59	67	68	63	+121 to -328 for MRE11; -260 to +189 of ANKRD48 TSS	Double-strand break repair protein MRE11	Transcribed divergently from *ANKRD48*
MSH2	chr2:47,402,729-47,403,177	449	59	65	68	67	-338 to +111	DNA mismatch repair protein MSH2	
MSH3	chr5:80,654,325-80,654,773	449	72	68	69	69	-323 to +126	DNA mismatch repair protein MSH3	
MSH4	chr1:75,796,574-75,797,021	448	53	64	68	62	-297 to +151	DNA mismatch repair protein MSH4	
MSH5	chr6:31,739,646-31,740,092	447	63	66	70	65	-302 to +145	DNA mismatch repair protein MSH5	
MSH6	chr2:47,782,759-47,783,199	441	77	65	67	66	-323 to +118	DNA mismatch repair protein MSH6	
MUTYH	chr1:45,340,319-45,340,767	449	55	63	66	66	-296 to +153	Adenine DNA glycosylase	
NEIL2	chr8:11,769,235-11,769,674	440	62	65	67	67	-390 to +50	Endonuclease 8-like 2	
NEIL3	chr4:177,309,535-177,309,983	449	55	60	65	60	-302 to +147	Endonuclease 8-like 3	
NKX2-5	chr5:173,235,161-17,323,5610	450	56	69	69	70	-297 to +153	NK2 homeobox 5	
OGG1	chr3:9,749,662-9,750,103	442	64	68	70	70	-282 to +160	N-glycosylase/DNA lyase	
OLIG2	chr21:33,025,611-33,026,059	449	67	67	63	68	-324 to +125	Oligodendrocyte transcription factor 2	
PARP1	chr1:226,408,056-226,408,497	442	66	66	66	69	-396 to +46	Poly [ADP-ribose] polymerase 1	
PCNA	chr20:5,126,395-5,126,840	446	61	68	65	64	-217 to +229	Proliferating cell nuclear antigen	
PELO	chr5:52,787,645-52,788,092	448	60	67	63	67	-295 to +153	Protein pelota homoloq	Transcribed from same promoter as *ITGA1*
PMS1	chr2:189,783,763-189,784,201	439	64	62	66	65	-322 to +117	PMS1 protein homolog 1	
PMS2	chr7:6,008,985-6,009,426	442	62	64	67	69	+122 to -320 of PMS2 TSS; -287 to +155 of AIMP2 TSS	Mismatch repair endonuclease PMS2	Transcribed divergently from *AIMP2*
PMS2P1	chr7:100,336,159-100,336,604	446	67	64	68	66	-296 to +150	Putative postmeiotic segregation increased 2-like protein 1	
PMS2P11	chr7:77,011,149-77,011,595	447	49	63	62	63	-298 to +149	Putative postmeiotic segregation increased 2-like protein 11	
PMS2P2	chr7:75,358,848-75,359,299	452	61	65	67	67	-302 to +150	Putative postmeiotic segregation increased 2-like protein 2	
PMS2P3	chr7:75,527,982-75,528,429	448	61	65	69	68	-305 to +143	Putative postmeiotic segregation increased 2-like protein 3	
PMS2P5	chr7:74,890,474-74,890,918	445	62	67	69	66	+139 to -306 of STAG3L1 TSS; -294 to +151 of PMS2P2 TSS	Postmeiotic segregation increased 2-like protein 5	Transcribed divergently from *STAG3L1*
POLD1	chr19:50,384,048-50,384,479	432	68	70	69	67	-275 to +157	DNA polymerase delta catalytic subunit	
POLD4	chr11:67,353,447-67,353,893	447	65	68	67	67	-297 to +150	DNA polymerase delta subunit 4	
POLG	chr15:89,334,728-89,335,168	441	64	66	70	68	-372 to +69	DNA polymerase subunit gamma-1	
POLK	chr5:75,511,457-75,511,901	445	64	68	64	63	-283 to +152	DNA polymerase kappa	
POLQ	chr3:121,545,842-121,546,288	447	52	60	62	66	-281 to +166	DNA polymerase theta	
POLR2B	chr4:56,978,353-56,978,799	447	48	69	66	69	-287 to +160	DNA-directed RNA polymerase II subunit RPB2	
POLR2G	chr11:62,761,244-62,761,691	448	62	62	67	66	-295 to +153	DNA-directed RNA polymerase II subunit RPB7	
PRDM2	chr1:13,704,558-13,705,004	447	34	61	62	57	-297 to +150	PR domain zinc finger protein 2	
PRKCG	chr19:53,881,833-53,882,277	445	68	67	70	67	-351 to +94	Protein kinase C gamma type	
PRUNE2	chr9:76,905,871-76,906,311	441	68	67	69	67	-223 to +218	Protein prune homolog 2	
PXMP4	chr20:3,372,0179-33,720,629	451	58	67	64	64	-298 to +153	Peroxisomal membrane protein 4	
PYCARD	chr16:31,202,733-31,203,170	449	65	62	68	69	-393 to +45	Apoptosis-associated speck-like protein containing a CARD	
RNASEH2A	chr19:12,806,317-12,806,765	450	66	63	69	66	-298 to +152	Ribonuclease H2 subunit A	
RNF111	chr15:58,987,414-58,987,859	446	55	61	68	68	-252 to +194	E3 ubiquitin-protein ligase Arkadia	
RPA2	chr1:27,914,634-27,915,080	447	56	64	67	64	-282 to +165	Replication protein A 32 kDa subunit	
RPA3	chr7:7,718,466-7,718,915	450	34	62	57	58	-307 to +143	Replication protein A 14 kDa subunit	
RPA4	chrX:96,883,621-96,884,066	446	44	66	67	64	-287 to +159	Replication protein A 30 kDa subunit	
SH3TC1	chr4:8,198,945-8,199,389	445	66	62	65	69	-299 to +146	SH3 domain and tetratricopeptide repeat-containing protein 1	
SIM2	chr21:36,698,824-36,699,273	450	77	68	69	70	-297 to +153	Single-minded homolog 2	
SMAD4	chr18:51,029,915-51,030,359	445	67	64	70	64	-298 to +147	Mothers against decapentaplegic homoloq 4	
SMUG1	chr12:54,188,843-54,189,292	450	53	61	65	66	-297 to +153	Single-strand selective monofunctional uracil DNA glycosylase	
SNAI1	chr20:49,982,773-49,983,210	438	68	69	68	68	-203 to +235	Snail family transcriptional repressor 1	
SNAI2	chr8:48,921,315-48,921,748	434	57	67	68	70	-307 to +127	Zinc finger protein SNAI2	
TAF1	chrX:71,365,956-71,366,401	446	55	68	70	66	-283 to +163	Transcription initiation factor TFIID subunit 1	
TDG	chr12:103,965,499-103,965,940	442	59	64	64	69	-316 to +126	G/T mismatch-specific thymine DNA glycosylase	
TP53	chr17:7,687,409-7,687,855	447	53	62	67	69	-304 to +143	Cellular tumor antigen p53	
TREX1	chr3:48,465,929-48,466,361	448	67	68	66	69	-302 to +146	Three-prime repair exonuclease 1	
TRIM28	chr19:58,544,079-58,544,508	430	69	69	68	68	-390 to +40	Transcription intermediary factor 1-beta	
TWIST1	chr7:19,117,521-19,117,969	449	65	69	70	68	-296 to +153	Twist-related protein 1	
TWIST2	chr2:238,847,667-238,848,112	446	67	66	65	67	-365 to +81	Twist-related protein 2	
UBB	chr17:16,380,497-16,380,942	446	54	69	63	63	-296 to +150	Polyubiquitin-B	
USP47	chr11:11,841,122-11,841,567	446	58	60	64	65	-301 to +145	Ubiquitin carboxyl-terminal hydrolase 47	
WDR48	chr3:39,051,689-39,052,135	447	58	66	68	68	-297 to +150	WD repeat-containing protein 48	
WNT1	chr12:48,978,163-48,978,609	447	74	68	68	68	-290 to +157	Proto-oncogene Wnt-1	
XPC	chr3:14,178,508-14,178,952	445	59	63	68	67	-279 to +166	DNA repair protein complementing XP-C cells	
XRCC2	chr7:152,676,039-152,676,482	444	59	67	70	65	-316 to +128	DNA repair protein XRCC2	
XRCC3	chr14:103,715,325-10,3715,761	437	69	70	69	68	-274 to +163	DNA repair protein XRCC3	
ZEB1	chr10:31,318,813-31,319,255	443	65	69	69	70	-359 to +84	Zinc finger E-box-binding homeobox 1	
ZNF830	chr17:34,961,234-34,961,679	446	57	69	62	66	+227 to -219 of CCT6B TSS; -306 to +140 of ZNF830 TSS;	Zinc finger protein 830	Transcribed divergently from *CCT6B* (two TSSs only 81 bp apart)

[a] Gene Ontology term "metabolic process," filtered further for "DNA repair"

[b] Gene products with known cancer-associated functions

[c] Control loci with open chromatin

Supplementary Table 2: Characteristics of designed FENGC targets used in this study; human ~940-nt targets.

Gene	Genome coordinates	Captured length (nt)	Captured sequence GC content (%)	Tm			Captured sequence relative to TSS (+1)	Notes
				Oligo 1-T	Oligo 2-T	Oligo 3-T		
ALKBH2	chr12:109,093,335-109,094,280	946	54	65	65	65	-791 to +155	
APEX1	chr14:20,454,342-20,455,278	937	56	60	64	64	-789 to +148	
APEX2	chrX:54,999,546-55,000,485	940	55	63	67	68	-777 to +163	
ATM	chr11:108,222,047-108,222,984	938	58	66	68	68	+592 to -346 for NPAT TSS; -1020 to -83 for ATM TSS	Transcribed divergently from *NPAT*
ATRX	chrX:77,786,121-77,787,050	930	43	64	64	65	-780 to +150	
BRCA1	chr17:43,125,334-43,126,270	937	51	69	65	67	-786 to +151	
BRCA2	chr13:32,314,697-32,315,631	935	57	65	64	67	-783 to +152	
CCN4	chr8:133,190,263-133,191,206	944	53	69	69	69	-776 to +168	
CD44	chr11:35,138,111-35,139,052	942	53	64	68	64	-759 to +183	
CETN1	chr18:579,563-580,492	930	54	62	62	62	-780 to +150	
CHEK1	chr11:125,624,364-125,625,316	953	44	63	63	62	-772 to +181	
CHEK2	chr22:28,741,673-28,742,622	950	61	64	66	65	-787 to +163	
CSDE1	chr1:114,757,871-114,758,815	945	47	62	64	67	-764 to +181	
DDB2	chr11:47,214,152-47,215,093	942	58	63	62	65	-790 to +152	
EPSTI1	chr13:42,992,151-42,993,084	934	47	63	66	65	-771 to +163	
EXO1	chr1:241,847,367-241,848,317	951	53	63	66	65	-824 to +127	
FEN1	chr11:61,791,846-61,792,787	942	54	64	66	63	-791 to +151	
GSTA1	chr6:52,803,818-52,804,756	939	39	61	65	58	-788 to +151	
HIST1H1B	chr6:27,867,390-27,868,341	952	45	63	65	64	-759 to +193	
HOTAIR	chr12:53968605-53969564	959	53	68	68	67	-807 to +152	
HOXB2	chr17:48,544,881-48,545,815	935	49	64	62	64	-783 to +152	
HSPA5	chr9:125,241,260-125,242,210	951	59	66	66	67	-823 to +128	Control promoter with open chromatin
LIG1	chr19:48,170,440-48,171,380	941	65	64	65	65	-776 to +165	
MBD4	chr3:129,440,032-129,440,982	951	56	64	65	65	-802 to +149	
MCM9	chr6:118,931,590-118,932,528	939	33	60	61	60	-789 to +150	
MECP2	chrX:154,097,673-154,098,629	957	59	65	68	65	-897 to +60	
MLH1	chr3:36,992,867-36,993,803	937	59	67	69	70	+328 to -609 of EPM2AIP1 TSS; -650 to +287 of MLH1 TSSa	Transcribed divergently from *EPM2AIP1*
MRE11	chr11:94,493,724-94,494,679	956	54	64	68	63	+121 to -835 of MRE11 TSS; -260 to +696 of ANKRD48 TSS	Transcribed divergently from *ANKRD48*
MSH2	chr2:47,402,248-47,403,177	930	52	69	68	67	-819 to +111	
MSH4	chr1:75,796,091-75,797,021	931	46	64	68	62	-780 to +151	
MSH5	chr6:31,739,162-31,740,092	931	53	65	70	65	-786 to +145	
OGG1	chr3:9,749,158-9,750,103	946	61	66	70	70	-786 to +160	
OLIG2	chr21:33,025,108-33,026,059	952	63	65	63	68	-800 to +152	
PELO	chr5:52,787,147-52,788,092	946	50	63	63	67	-793 to +153	Transcribed from same promoter as *ITGA1*
PMS2	chr7:6,008,985-6,009,928	944	63	68	67	69	+122 to -822 of PMS2 TSS; -287 to +657 of AIMP2 TSS	Transcribed divergently from *AIMP2*
PMS2P5	chr7:74,889,984-74,890,918	935	64	68	69	66	+629 to -306 of STAGL3L2 TSS; -784 to +151 of PMS2P5 TSS	Transcribed divergently from *STAG3L2*
POLD4	chr11:67,353,447-67,354,385	939	59	68	67	67	-788 to +151	
PRUNE2	chr9:76,905,871-76,906,815	945	66	68	69	67	-727 to +218	
RPA2	chr1:27,914,634-27,915,584	951	53	68	67	64	-786 to +165	
RPA4	chrX:96,883,116-96,884,066	951	42	66	67	64	-792 to +159	Transcribed from intron 5 of *DIAPH2*
TAF1	chrX:71,365,444-71,366,401	958	53	66	66	70	-795 to +163	
TP53	chr17:7,687,409-7,688,346	938	49	67	67	69	-795 to +143	
UBB	chr17:16,380,000-16,380,942	943	56	63	63	63	-793 to +150	
XRCC2	chr7:152,676,039-152,676,977	939	50	65	70	65	-811 to +128	
ZNF830	chr17:34,960,730-34,961,679	950	53	64	62	66	+731 to -219 of CCT6B TSS; -810 to +140 of ZNF830 TSS;	Transcribed divergently from *CCT6B* (two TSSs only 81 bp apart)

Supplementary Table 2: Characteristics of designed FENGC target sequences used in this study; mouse ~620-nt targets.

Gene	Genome coordinates	Captured length (nt)	Captured sequence GC content (%)	Tm			Captured sequence relative to TSS (+1)	Notes
				Oligo 1-T	Oligo 2-T	Oligo 3-T		
Ash1l	chr3:88,950,129-88,950,957	829	73	65.1	61.4	64.3	-493 to +336	
Btk	chrX:134,583,042-134,583,555	514	50	61	60.8	63.4	+99 to -415	
Ccl5	chr11:83,530,364-83,530,887	524	54	61.8	64.4	60.8	+155 to -369	
Cd180	chr13:102,693,205-102,693,713	509	46	63.2	60.8	64.6	-353 to +156	
Cox8a	chr19:7,217,511-7,218,156	646	52	59.5	65.3	61.5	+106 to -540	Control promoter with open chromatin
Cxcl15	chr5:90,794,068-90,794,625	558	44	58.6	57	63	-466 to +92	
Cxcr4	chr1:128,592,199-128,592,690	492	57	62.6	60.4	63.4	+101 to -391	
Dnmt3b	chr2:153,648,958-153,649,944	987	70	62.7	62	64.4	-492 to +495	
Foxo1	chr3:52,267,705-52,268,609	905	65	64.1	58.5	64.6	-631 to +274	
Gdf5	chr2:155,945,300-155,945,790	491	46	65.2	64.5	58.5	+68 to -423	
Hdac3	chr18:37,954,900-37,955,436	537	61	62.7	65.3	61.6	+117 to -420 of Hdac3 TSS; -250 to +287 of Rell2 TSS	Transcribed divergently from *Rell2*
Hsf1	chr15:76,477,070-76,477,543	474	67	64.9	62.1	61.9	+208 to -266 of Bop1 TSS; -365 to +109 of Hsf1 TSS	Transcribed divergently from *Bop1*
Hsf2bp	chr17:32,036,229-32,036,712	484	70	65	65.2	65.1	+121 to -363 of Hsf2bp TSS; +130 to +613 of Rrp1b	Transcribed divergently from *Rrp1b*
Hsf4	chr8:105,269,428-105,270,050	623	66	61	64.3	64.8	-401 to +222	
Hsf5	chr11:87,616,773-87,617,408	636	69	61.6	60.6	59.2	-391 to +245	
Hsfy2	chr1:56,637,186-56,637,847	662	40	56.1	62.5	61.3	+250 to -412	
Hsp90aa1	chr12:110,702,628-110,703,149	522	46	63.2	60.2	64.7	-6233 to -6754	
Hsp90ab1	chr17:45,573,048-45,573,653	606	63	60.3	60.5	64.7	+224 to -382	
Hspa1a	chr17:34,971,883-34,972,591	709	59	61.6	65.2	65.6	+274 to -435	
Hspa1b	chr17:34,958,965-34,959,604	640	63	64.8	65.2	65.6	+274 to -366	
Hspa2	chr12:76,403,857-76,404,461	605	63	64.3	61.1	64.9	-319 to +286	
Ifna1	chr4:88,849,613-88,850,308	696	40	56.1	64	62.5	-474 to +222	
Ifna2	chr4:88,683,754-88,684,286	533	37	58.1	59	61.4	-374 to +159	
Ifng	chr10:118,440,660-118,441,135	476	47	65.1	65.1	61.8	-386 to +90	
Ikbkb	chr8:22,706,326-22,707,045	720	59	62.6	63.2	64.8	+258 to -462	
Ikbkg	chrX:74,392,804-74,393,357	554	61	63.4	58.2	63.9	+471 to -83 of Fam3a TSS; -487 to +67 of Ikbkg TSS	Transcribed divergently from *Fam3a* (two TSSs only 18 bp apart)
Il10	chr1:131,019,395-131,019,945	551	42	60.3	63	60.5	-450 to +101	
Il12a	chr3:68,690,275-68,690,811	537	51	61.4	60.6	61.5	-369 to +168	
Il12b	chr11:44,399,601-44,400,132	532	46	59.7	62.4	61.8	-462 to +70	
Il3ra	chr14:14,346,110-14,346,724	615	65	63.5	65.4	61.7	-165 to +450	
Il6	chr5:30,012,590-30,013,396	807	46	65.4	62.4	65.3	-524 to +283	
Irak1	chrX:74,023,658-74,024,335	678	61	64.7	63.6	65.1	+261 to -417	
Irak4	chr15:94,543,258-94,543,863	606	65	58.4	59.1	61.4	-387 to +219	
Irf3	chr7:44,997,272-44,997,828	557	54	58.2	61.9	65.1	+308 to -249 of Bcl2l12; -376 to +181 of Irf3 TSS	Transcribed divergently from *Bcl2l12*
Irf5	chr6:29,526,316-29,527,062	747	73	62.9	60.2	58.3	-392 to +355	
Irf7	chr7:141,266,329-141,266,813	485	55	64.2	58.5	62.7	+93 to -392	
Jun	chr4:95,052,198-95,052,827	630	64	59.3	63	61	+25 to -605	
Kdm6b	chr11:69,413,552-69,414,040	489	52	61.4	60.2	62.3	+124 to -365	
Lbp	chr2:158,306,144-158,306,707	564	56	60.4	62.1	62.6	-350 to +214	
Ly86	chr13:37,344,850-37,345,383	534	50	64.2	60.6	63.2	-495 to +39	
Map3k1	chr13:111,808,985-111,809,661	677	72	64	59.6	55.5	+9 to -668	
Mapk1	chr16:16,982,979-16,983,474	496	59	59.9	65	61.3	-403 to +93	
Mapk13	chr17:28,768,866-28,769,660	795	65	59.6	59.1	62.1	-431 to +364	
Mapk14	chr17:28,690,874-28,691,764	891	68	62.7	63.4	65.1	-455 to +436	
Mapk15	chr15:75,995,428-75,995,983	556	57	61.1	64.8	64.8	+1660 to +2215	
Mapk3	chr7:126,759,191-126,759,897	707	67	58.2	63.6	61.2	-435 to +272	
Mapk4	chr18:74,064,823-74,065,577	755	69	57.9	65.1	59.5	+127 to -628	
Mapk6	chr9:75,409,879-75,410,441	563	68	61.4	62.6	58.5	+127 to -436	
Mapk8	chr14:33,446,723-33,447,594	872	72	63.8	64.9	64.8	+436 to -436	
Mme	chr3:63,295,269-63,295,864	596	65	59.8	61.6	63.2	-166 to +430	
Myd88	chr9:119,339,910-119,340,537	628	51	64.9	62.4	64.2	+132 to -496	
Nfkb1	chr3:135,691,529-135,692,148	620	61	65.3	63.6	64.7	+18 to -602	
Nfkb2	chr19:46,303,888-46,304,398	511	50	58.4	62.8	58.1	-432 to +79	
Nfkbia	chr12:55,492,417-55,493,106	690	60	62.6	61.9	62.2	+231 to -459	
Nfkbib	chr7:28,766,590-28,767,119	530	57	62.8	65	59.7	+55 to -475 of Nfkbib TSS; -145 to +385 of Sirt2 TSS	Transcribed divergently from *Sirt2*
Nr2c2	chr6:92,091,014-92,091,520	507	67	62.7	61	64.3	-376 to +131	
Pik3ap1	chr19:41,384,869-41,385,481	613	67	65.2	57	65.2	+228 to -385	
Pik3c2a	chr7:116443199-116444081	883	65	64.8	59.1	63.9	+251 to -632	
Pik3c2g	chr6:139,736,328-139,736,987	660	37	58.3	65.3	62.9	-567 to +93	
Pik3c3	chr18:30,272,232-30,272,871	640	39	59.3	63.9	63	-515 to +125	
Pik3ca	chr3:32,397,129-32,397,798	670	75	58.9	64.1	59.9	-542 to +128	
Pik3r1	chr13:101,768,008-101,768,630	623	71	60.1	65.3	65.8	+210 to -413	
Pik3r2	chr8:70,776,348-70,777,162	815	66	64.9	64.5	65.2	-449 to +366	
Pik3r3	chr4:116,221,346-116,221,950	605	57	63.8	61.7	61.8	-263 to +343	
Pik3r5	chr11:68,431,620-68,432,347	728	60	63.5	63.8	65.1	-501 to +227	
Pik3r6	chr11:68,502,642-68,503,181	540	51	61.6	63.4	65.3	-377 to +163	
Ptpn6	chr6:124,738,642-124,739,124	483	59	60.1	59.7	62.1	+68 to -415	
Rela	chr19:5,637,111-5,637,764	654	71	60	65.6	55.4	-372 to +282	
Src	chr2:157,417,873-157,418,488	616	51	61.3	60.7	65.5	-6420 to -5805	
Stat3	chr11:100,939,384-100,939,932	549	60	61.6	64.1	58.4	+157 to -392	
Tab1	chr15:80,132,766-80,133,525	760	63	59.9	64.4	64.5	-361 to +399	
Tab2	chr10:7,956,254-7,956,828	575	78	61.1	57.2	55.4	-131 to -705	
Tet2	chr3:133,544,278-133,544,951	674	75	65.4	60.4	59.6	+113 to -561	
Tlr2	chr3:83,841,701-83,842,365	665	47	63.7	64.1	64.2	+67 to -598	
Tlr4	chr4:66,827,250-66,827,825	576	48	61.1	58.3	62.1	-301 to +275	
Tnf	chr17:35,201,859-35,202,357	499	55	59.4	61.8	64.9	+149 to -350	
Traf6	chr2:101,678,076-101,678,618	543	58	57	64.2	61	-364 to +179	
Vav1	chr17:57,278,729-57,279,333	605	61	62.1	64	60.9	-371 to +234	

Supplementary Table 3: CCS reads aligned to 11 human targets of ~300 nt *vs*. ~450 nt used in FENGC assay development with standard PCR *vs*. BS-PCR.

	~300-nt targets aligned CCS read number [a]					
Sample [b] / Promoter	~300 nt, Taq, PCR-1	~300 nt, Taq, PCR-2	~300 nt, FEN1, PCR-1	~300 nt, FEN1, PCR-2	~300 nt, FEN1, BS-PCR-1	~300 nt, FEN1, BS-PCR-2
EPM2AIP1	3896	1994	27	34	165	54
EXO1	37	17	3215	2913	2500	3105
HSPA5	21	10	1534	1528	1311	1149
MLH1	5081	2545	456	603	1477	2054
MLH3	141	171	1355	1743	48	34
MSH2	1395	643	810	729	287	159
MSH3	0	0	0	1	0	1
MSH6	0	0	0	0	0	0
PMS1	3140	1759	1610	2500	104	45
PMS2	210	149	716	1059	304	278
TDG	2018	1068	2537	2396	59	33

	~450-nt targets aligned CCS read number [a]					
Sample [b] / Promoter	~450 nt, Taq, PCR-1	~450 nt, Taq, PCR-2	~450 nt, FEN1, PCR-1	~450 nt, FEN1, PCR-2	~450 nt, FEN1, BS-PCR-1	~450 nt, FEN1, BS-PCR-2
EPM2AIP1	4321	3865	236	156	109	41
EXO1	14	21	2678	3371	318	234
HSPA5	103	55	2859	3607	933	501
MLH1	540	277	1682	1362	2325	1527
MLH3	31	33	111	208	5	2
MSH2	454	398	2196	2132	121	50
MSH3	0	0	0	0	6	3
MSH6	0	0	0	0	7	10
PMS1	454	347	817	971	4	10
PMS2	11	14	1150	1431	194	113
TDG	22	28	1148	1259	8	1

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes and ≥95% HCH conversion.

[b] Suffixes -1 and -2 refer to the two independent biological replicates.

[c] BS-PCR control was a PCR product amplified from bisulfite-converted gDNA.

Supplementary Table 4: CCS reads on- and off-target for 119 ~450-nt human targets in MAPit-FENGC; EM-seq. [a,b]

Measurement \ Sample [c]	NSC-1	NSC-2	L0-1	L0-2	0.1%:99.9% L0:NSC-1	0.1%:99.9% L0:NSC-2	Nx18-25-1	Nx18-25-2
Input genomic DNA (ng)	500	500	500	500	500	500	500	500
Total reads	91,751	189,941	15,937	20,601	67,757	53,670	66,882	54,385
Reads aligned on genome	85,408	168,764	14,223	18,884	59,928	48,017	59,286	47,266
% Reads aligned on genome	93.1%	88.9%	89.2%	91.7%	88.4%	89.5%	88.6%	86.9%
Reads aligned on target	77,014	156,295	13,056	17,642	55,750	41,368	53,392	43,269
% Reads off target but on genome	9.8%	7.4%	8.2%	6.6%	7.0%	13.8%	9.9%	8.5%
% Total reads on target	83.9%	82.3%	81.9%	85.6%	82.3%	77.1%	79.8%	79.6%
Number of targets detected [d]	111		105		109		109	
Number of targets detected using reads with ≥95% coverage of each reference sequence and ≥95% conversion	105		92		103		104	

[a] Genomic DNA from M.CviPI-treated cells subject to EM-PCR for detection of DNA methylation.

[b] High-fidelity CCS reads with ≥3 SMRT sequencing passes.

[c] Suffixes -1 and -2 refer to the two independent biological replicates.

[d] At least one read in at least one condition.

Supplementary Table 4: CCS reads on- and off-target for 119 ~450-nt targets in MAPit-FENGC; standard PCR. [a,b]

Measurement \ Sample [c]	NSC-1	NSC-2	Nx18-25-1	Nx18-25-2
Input genomic DNA (ng)	500	500	500	500
Total reads	32,701	27,442	25,062	22,905
Reads aligned on genome	32,512	27,191	22,819	22,652
% Reads aligned on genome	99.4%	99.1%	91.1%	98.9%
Reads aligned on target	28,414	22,963	18,170	20,377
% Reads off target but on genome	12.6%	15.5%	20.4%	10.0%
% Total reads on target	86.9%	83.7%	72.5%	89.0%
Number of targets detected [d]	97			

[a] Genomic DNA from M.CviPI-treated cells subject to standard PCR for genotyping.

[b] High-fidelity CCS reads with ≥3 SMRT sequencing passes.

[c] Suffixes -1 and -2 refer to the two independent biological replicates.

[d] At least one read in at least one condition.

Supplementary Table 5: Filtered CCS reads aligned to 119 ~450-nt human targets in MAPit-FENGC using EM-seq.[a]

	Filtered CCS read number[b]							
Sample[c] / Promoter	NSC-1	NSC-2	L0-1	L0-2	0.1%:99.9% L0:NSC-1	0.1%:99.9% L0:NSC-2	Nx18-25-1	Nx18-25-2
ALKBH1	0	6	0	0	2	0	1	0
ALKBH2	3152	7899	476	503	5595	2842	3381	3072
ALKBH3	5	6	0	0	3	3	2	2
APEX1	159	421	33	32	198	70	124	81
APEX2	701	2995	318	365	1928	813	5865	4908
ATM	807	1866	70	93	418	195	1292	964
ATRX	202	730	40	62	194	79	936	623
ATXN3	0	0	0	0	0	0	0	0
AXIN2	0	0	0	0	0	0	0	0
BRCA1	398	712	217	179	709	236	560	479
BRCA2	129	778	226	218	445	280	785	539
CCN4	333	783	31	66	456	147	566	577
CD44	48	93	7	15	46	33	53	43
CDH1	6	15	0	1	9	4	8	8
CDH13	0	0	0	0	0	0	0	0
CDH2	15	9	0	1	9	4	0	3
CDKN2A	0	0	0	0	0	0	0	0
CDKN2D	2	0	0	0	1	0	0	0
CETN1	280	1555	35	53	133	103	472	434
CHEK1	184	299	22	40	177	101	233	131
CHEK2	3197	12014	1887	1982	4333	2314	5409	5098
CSDE1	63	112	57	62	84	43	145	128
DDB2	69	412	17	20	73	38	182	150
DRD3	48	62	0	2	7	1	26	26
DRD4	0	0	4	6	3	0	0	1
EGFR	0	0	0	0	0	0	1	0
EPM2AIP1	1138	2322	1961	1928	1183	598	687	518
EPSTI1	70	135	13	17	51	33	142	80
ERCC1	249	281	93	137	915	362	422	383
ERCC2	7	6	7	3	10	6	13	23
ERCC8	613	2453	67	82	1295	458	1351	1386
EXO1	4	16	12	24	17	12	30	25
FANCC	1	11	0	1	2	0	1	2
FAT4	0	0	0	0	0	0	0	0
FEN1	705	1828	153	181	986	502	466	480
FHIT	46	95	9	7	71	36	22	26
FOXC2	9	6	2	0	14	10	1	3
FZD7	0	0	0	0	0	0	0	0
FZD8	7	3	0	0	2	4	2	3
GPR158	10	31	6	8	10	11	6	8
GSTA1	1580	2952	54	52	780	320	2468	1718
HIST1H1B	293	633	19	34	371	121	430	523
HMGA2	1	2	1	1	0	0	1	1
HOTAIR	1063	1877	448	528	1175	597	1546	968
HOXB2	31	43	104	140	59	26	57	24
HSPA5	0	28	84	87	3	1	28	11
INO80	3	1	0	0	0	1	0	6
KLF8	0	5	0	0	2	0	0	0
LIG1	14	34	2	3	2	2	21	17
MBD4	50	144	19	27	96	30	112	93
MCM9	655	1721	29	43	267	157	1008	977
MECP2	7	31	0	0	4	3	6	7
MEG3	97	84	8	8	19	19	48	47
MGMT	15	29	0	2	42	31	5	3
MLH1	2872	8950	1339	1532	1741	2054	3558	3287
MLH3	7	15	0	2	2	1	3	6
MNAT1	35	142	5	8	9	6	194	116
MPG	0	1	0	1	4	0	1	0

MRE11	435	1215	43	83	595	299	500	485
MSH2	1117	1939	185	242	2485	801	1531	1036
MSH3	2	23	1	9	9	12	8	8
MSH4	13	99	12	33	21	32	33	25
MSH5	132	241	24	57	198	107	170	192
MSH6	5	13	1	10	0	7	9	10
MUTYH	7	2	0	3	0	2	7	14
NEIL2	0	0	0	0	0	0	0	0
NEIL3	1	1	0	0	0	0	3	0
NKX2-5	0	5	0	0	0	1	0	0
OGG1	131	275	617	549	700	397	189	183
OLIG2	116	291	6	19	46	51	86	86
PARP1	0	11	1	1	0	1	2	0
PCNA	408	552	20	18	66	109	358	216
PELO	103	235	17	23	121	89	89	95
PMS1	27	63	0	1	8	12	33	19
PMS2	56	214	30	30	68	93	82	72
PMS2P1	26	124	16	28	56	33	32	30
PMS2P11	501	991	25	52	340	147	680	489
PMS2P2	5	17	0	1	5	3	12	9
PMS2P3	33	58	336	324	67	36	64	41
PMS2P5	1078	5423	423	547	2232	1425	3784	3841
POLD1	0	0	2	2	2	0	0	1
POLD4	1491	1719	1043	1096	3924	1853	1491	1441
POLG	0	0	0	0	0	0	0	0
POLK	13	82	1	2	5	2	41	31
POLQ	215	456	56	89	195	66	595	469
POLR2B	47	76	9	12	40	32	15	19
POLR2G	3	24	4	0	2	5	5	4
PRDM2	407	641	47	50	310	135	662	413
PRKCG	24	7	226	209	89	33	47	33
PRUNE2	238	673	24	28	247	145	280	303
PXMP4	1	3	0	1	0	2	7	6
PYCARD	0	3	3	4	3	2	3	2
RNASEH2A	11	4	0	6	8	4	10	9
RNF111	130	294	12	5	56	31	159	126
RPA2	1026	5542	890	878	1755	1012	4577	3062
RPA3	187	206	1	8	68	36	234	154
RPA4	3285	10938	423	471	5404	2499	3858	3811
SH3TC1	11	10	59	74	44	14	11	5
SIM2	0	3	0	0	4	2	2	0
SMAD4	12	3	0	0	5	5	1	4
SMUG1	8	24	1	1	10	2	16	10
SNAI1	0	0	0	0	0	0	0	0
SNAI2	1	4	0	3	2	1	1	0
TAF1	168	370	88	134	216	102	711	546
TDG	6	3	0	2	5	1	3	4
TP53	59	144	14	21	72	38	160	107
TREX1	0	3	0	0	0	0	2	0
TRIM28	32	95	7	6	24	22	70	42
TWIST1	4	2	0	0	3	4	5	1
TWIST2	9	10	1	1	16	4	4	3
UBB	66	118	3	5	17	22	43	32
USP47	7	6	0	2	1	0	2	0
WDR48	3	9	0	0	2	0	6	1
WNT1	71	55	82	72	106	57	18	12
XPC	4	0	0	0	0	0	0	0
XRCC2	9	20	28	29	21	11	23	41
XRCC3	0	0	0	0	0	0	0	0
ZEB1	0	0	0	0	0	0	0	0
ZNF830	6095	24059	1156	1234	11769	7706	10358	8753

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.

[b] Filtered for ≥95% coverage of each reference sequence and ≥95% HCH conversion.

[c] Suffixes -1 and -2 refer to the two independent biological replicates.

Supplementary Table 6: Differential H5mCG and G5mCH in NSC and GBM Nx18-25 among ~450-nt targets determined by MAPit-FENGC using EM-seq; no cutoff.[a]

Differential DNA methylation (H5mCG)

Parameter / Promoter	CCS read number [b]	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	43	0.575	0.0327	309.35	<0.0001	<0.0001
HOXB2 [g]	24	0.506	0.0021	57278.74	<0.0001	<0.0001
HOTAIR [g]	968	0.491	0.0032	23831.20	<0.0001	<0.0001
TAF1	168	0.233	0.0168	193.08	<0.0001	<0.0001
APEX2	701	0.145	0.0125	134.27	<0.0001	<0.0001
MCM9	655	0.109	0.0038	799.87	<0.0001	<0.0001
PELO	89	0.067	0.0021	965.23	<0.0001	<0.0001
PRDM2	407	0.051	0.0025	417.96	<0.0001	<0.0001
BRCA2	129	0.044	0.0049	79.29	<0.0001	<0.0001
HIST1H1B	293	0.041	0.0008	2569.17	<0.0001	<0.0001
CETN1	280	0.036	0.011	10.95	0.0009	0.0637
HIST1H1B (TSS) [f]	293	0.033	0.0009	1295.97	<0.0001	<0.0001
RPA2	1026	0.025	0.0015	263.43	<0.0001	<0.0001
MEG3	47	0.009	0.0244	0.12	0.7269	1
EPM2AIP1	518	0.007	0.0014	21.65	<0.0001	0.0002
MNAT1	35	0.006	0.0006	114.13	<0.0001	<0.0001
TP53	59	0.000	0.0019	0.01	0.9103	<0.0001
ERCC1	249	0.000	0.0004	1.11	0.2911	<0.0001
PMS2P1	26	-0.003	0.0024	1.53	0.2167	1
PMS2	56	-0.003	0.0018	3.63	0.0568	1
POLQ	215	-0.004	0.0007	23.70	<0.0001	0.0001
ZNF830	6095	-0.004	0.0007	31.37	<0.0001	<0.0001
OGG1	131	-0.004	0.0019	4.26	0.0391	1
BRCA1	398	-0.004	0.0012	12.94	0.0003	0.0219
CHEK2	3197	-0.005	0.0006	69.66	<0.0001	<0.0001
CSDE1	63	-0.005	0.0037	2.11	0.1466	1
MLH1	2872	-0.006	0.0017	11.79	0.0006	0.0404
MRE11	435	-0.006	0.0014	19.57	<0.0001	0.0007
RNF111	126	-0.006	0.0008	61.54	<0.0001	<0.0001
PMS2P5	1078	-0.007	0.0005	254.04	<0.0001	<0.0001
APEX1	81	-0.008	0.0027	9.30	0.0023	0.1558
PMS2P3	33	-0.010	0.0088	1.19	0.2755	1
ERCC8	613	-0.011	0.0003	1340.62	<0.0001	<0.0001
MSH2	1036	-0.011	0.0036	9.53	0.002	0.1374
FEN1	466	-0.011	0.0013	80.51	<0.0001	<0.0001
DRD3	26	-0.015	0.0282	0.30	0.5838	1
PCNA	216	-0.016	0.001	289.17	<0.0001	<0.0001
FHIT	22	-0.017	0.0108	2.36	0.1246	1
MSH5	148	-0.017	0.0013	159.30	<0.0001	<0.0001
MBD4	50	-0.023	0.0161	2.04	0.1533	1
ATM	807	-0.024	0.0055	18.66	<0.0001	0.0011
POLD4	1491	-0.028	0.0059	22.32	<0.0001	0.0002
TRIM28	32	-0.031	0.0124	6.03	0.0141	0.9576
ATRX	202	-0.035	0.0102	11.70	0.0006	0.0424
OLIG2	86	-0.043	0.0027	256.22	<0.0001	<0.0001
ALKBH2	3072	-0.047	0.0013	1234.02	<0.0001	<0.0001
UBB [g]	32	-0.050	0.0069	53.77	<0.0001	<0.0001
CHEK1 [g]	131	-0.051	0.0019	711.19	<0.0001	<0.0001
PRUNE2	238	-0.065	0.0084	60.30	<0.0001	<0.0001
DDB2	69	-0.068	0.0012	3029.65	<0.0001	<0.0001
CCN4	333	-0.085	0.0098	73.89	<0.0001	<0.0001
RPA4	3285	-0.111	0.0018	3966.60	<0.0001	<0.0001
RPA3	154	-0.231	0.0336	47.50	<0.0001	<0.0001
EPSTI1	70	-0.355	0.0412	74.02	<0.0001	<0.0001
GSTA1 [g]	1580	-0.748	0.0028	73279.51	<0.0001	<0.0001

Differential Accessibility (G5mCH)

Parameter / Promoter	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	-0.201	0.0054	1398.66	<0.0001	<0.0001
HOXB2	0.015	0.0049	9.26	0.0023	0.1597
HOTAIR [g]	-0.067	0.0049	189.45	<0.0001	<0.0001
TAF1	-0.102	0.0051	401.63	<0.0001	<0.0001
APEX2	-0.066	0.0192	11.65	0.0006	0.0437
MCM9	-0.001	0.0109	0.01	0.9182	1
PELO	-0.113	0.0153	53.82	<0.0001	<0.0001
PRDM2	-0.051	0.005	107.73	<0.0001	<0.0001
BRCA2	-0.032	0.0215	2.27	0.1322	1
HIST1H1B	-0.046	0.0067	46.48	<0.0001	<0.0001
CETN1	-0.036	0.007	26.61	<0.0001	0.0001
HIST1H1B (TSS) [f]	-0.074	0.0017	1886.56	<0.0001	<0.0001
RPA2	-0.010	0.0078	1.59	0.2075	1
MEG3	-0.008	0.0157	0.28	0.5952	1
EPM2AIP1	-0.102	0.0191	28.26	<0.0001	<0.0001
MNAT1	-0.032	0.0055	33.64	<0.0001	<0.0001
TP53	-0.046	0.0074	39.66	<0.0001	0.0001
ERCC1	-0.050	0.0142	12.59	0.0004	0.0264
PMS2P1	0.015	0.0265	0.34	0.5601	1
PMS2	-0.086	0.0132	42.10	<0.0001	<0.0001
POLQ	-0.045	0.005	81.52	<0.0001	<0.0001
ZNF830	-0.067	0.0058	133.52	<0.0001	<0.0001
OGG1	-0.084	0.0209	16.06	<0.0001	0.0042
BRCA1	-0.040	0.0132	8.98	0.0027	0.1859
CHEK2	-0.020	0.0047	18.05	<0.0001	0.0015
CSDE1	-0.057	0.0122	21.64	<0.0001	0.0002
MLH1	-0.038	0.0118	10.23	0.0014	0.0942
MRE11	-0.053	0.0076	48.67	<0.0001	<0.0001
RNF111 [g]	-0.061	0.0179	11.56	0.0007	0.0459
PMS2P5	-0.050	0.0098	26.52	<0.0001	<0.0001
APEX1 [g]	-0.068	0.0158	18.22	<0.0001	0.0013
PMS2P3	0.006	0.071	0.01	0.9382	1
ERCC8	-0.074	0.0077	91.65	<0.0001	<0.0001
MSH2	-0.035	0.0099	12.07	0.0005	0.0348
FEN1	-0.019	0.0038	25.16	<0.0001	<0.0001
DRD3	-0.034	0.0118	8.32	0.0039	0.2661
PCNA	-0.037	0.0116	10.36	0.0013	0.0874
FHIT	0.038	0.0172	4.97	0.0258	1
MSH5	-0.046	0.0109	17.54	<0.0001	0.0019
MBD4	-0.006	0.019	0.10	0.7552	1
ATM	0.003	0.0087	0.12	0.7246	1
POLD4	-0.044	0.0252	2.99	0.0837	1
TRIM28	-0.087	0.072	1.45	0.2287	1
ATRX	-0.030	0.0287	1.10	0.295	1
OLIG2	0.067	0.019	12.44	0.0004	0.0285
ALKBH2	-0.059	0.0041	203.23	<0.0001	<0.0001
UBB	-0.017	0.0232	0.52	0.469	1
CHEK1 [g]	-0.072	0.0183	15.63	0.0001	0.0052
PRUNE2	-0.113	0.0108	108.80	<0.0001	<0.0001
DDB2	-0.043	0.0231	3.52	0.0607	1
CCN4	0.079	0.0333	5.56	0.0183	1
RPA4	-0.027	0.0061	20.36	<0.0001	0.0004
RPA3	0.027	0.0099	7.50	0.0062	0.4196
EPSTI1	0.051	0.0278	3.35	0.0673	1
GSTA1 [g]	-0.080	0.0057	192.09	<0.0001	<0.0001

Increase or decrease in proportion H5mCG ≥0.05 that reaches statistical significance (Bonferroni corrected $P < 0.05$).

Increase or decrease in proportion accessibility (G5mCH) ≥0.05 that reaches statistical significance (Bonferroni corrected $P < 0.05$).

[a] For promoters in which ≥50 reads obtained in combined duplicates, but no less than ≥22 reads in any of the 4 samples, i.e., both NSC duplicates and both Nx18-15 duplicates.

[b] Read number of sample with the least amount of reads.

[c] Proportion of difference in H5CG or accessibility (GBM Nx18-25 minus NSC non-cancerous cells).

[d] SE, standard error.

[e] Bonferroni method corrected (alpha = .05).

[f] Values calculated for HCG or GCH sites for −80 to +45 relative to TSS within full-length target from −246 to +199.

[g] Harbors one or more genetic variants (see Supplementary Table 7) that could contribute to the observed epigenetic differences.

Supplementary Table 6: Differential H5mCG and G5mCH in NSC and GBM Nx18-25 among ~450-nt targets determined by MAPit-FENGC using EM-seq; 5% cutoff.[a]

Differential DNA methylation (H5mCG)						
Parameter / Promoter	CCS read number [b]	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	43	0.575	0.0327	309.3459	< 0.0001	< 0.0001
HOXB2 [g]	24	0.506	0.0021	57278.74	< 0.0001	< 0.0001
HOTAIR [g]	968	0.491	0.0032	23831.2	< 0.0001	< 0.0001
TAF1	168	0.233	0.0168	193.0816	< 0.0001	< 0.0001
APEX2	701	0.145	0.0125	134.2705	< 0.0001	< 0.0001
MCM9	655	0.109	0.0038	799.8688	< 0.0001	< 0.0001
PELO	89	0.067	0.0021	965.2265	< 0.0001	< 0.0001
PRDM2	407	0.051	0.0025	417.9615	< 0.0001	< 0.0001
HIST1H1B (TSS) [f]	293	0.033	0.0009	1295.97	< 0.0001	< 0.0001
EPM2AIP1	518	0.007	0.0014	21.6546	< 0.0001	0.0002
ERCC1	249	0.000	0.0004	1.1144	0.2911	< 0.0001
PMS2	56	-0.003	0.0018	3.6271	0.0568	1
ZNF830	6095	-0.004	0.0007	31.3671	< 0.0001	< 0.0001
OGG1	131	-0.004	0.0019	4.2566	0.0391	1
CSDE1	63	-0.005	0.0037	2.1073	0.1466	1
MRE11	435	-0.006	0.0014	19.5723	< 0.0001	0.0007
RNF111	126	-0.006	0.0008	61.5392	< 0.0001	< 0.0001
PMS2P5	1078	-0.007	0.0005	254.038	< 0.0001	< 0.0001
APEX1	81	-0.008	0.0027	9.2998	0.0023	0.1558
ERCC8	613	-0.011	0.0003	1340.616	< 0.0001	< 0.0001
OLIG2	86	-0.043	0.0027	256.2185	< 0.0001	< 0.0001
ALKBH2	3072	-0.047	0.0013	1234.015	< 0.0001	< 0.0001
UBB [g]	32	-0.050	0.0069	53.7674	< 0.0001	< 0.0001
CHEK1 [g]	131	-0.051	0.0019	711.1882	< 0.0001	< 0.0001
PRUNE2	238	-0.065	0.0084	60.2971	< 0.0001	< 0.0001
DDB2	69	-0.068	0.0012	3029.653	< 0.0001	< 0.0001
CCN4	333	-0.085	0.0098	73.8911	< 0.0001	< 0.0001
RPA4	3285	-0.111	0.0018	3966.598	< 0.0001	< 0.0001
RPA3	154	-0.231	0.0336	47.4977	< 0.0001	< 0.0001
EPSTI1	70	-0.355	0.0412	74.0241	< 0.0001	< 0.0001
GSTA1 [g]	1580	-0.748	0.0028	73279.51	< 0.0001	< 0.0001

Differential Accessibility (G5mCH)					
Parameter / Promoter	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	-0.201	0.0054	1398.657	< 0.0001	< 0.0001
HOXB2	0.015	0.0049	9.255	0.0023	0.1597
HOTAIR [g]	-0.067	0.0049	189.452	< 0.0001	< 0.0001
TAF1	-0.102	0.0051	401.6285	< 0.0001	< 0.0001
APEX2	-0.066	0.0192	11.6487	0.0006	0.0437
MCM9	-0.001	0.0109	0.0106	0.9182	1
PELO	-0.113	0.0153	53.8195	< 0.0001	< 0.0001
PRDM2	-0.051	0.005	107.7271	< 0.0001	< 0.0001
HIST1H1B (TSS) [f]	-0.074	0.0017	1886.556	< 0.0001	< 0.0001
EPM2AIP1	-0.102	0.0191	28.259	< 0.0001	< 0.0001
ERCC1	-0.050	0.0142	12.587	0.0004	0.0264
PMS2	-0.086	0.0132	42.1014	< 0.0001	< 0.0001
ZNF830	-0.067	0.0058	133.5214	< 0.0001	< 0.0001
OGG1	-0.084	0.0209	16.0631	< 0.0001	0.0042
CSDE1	-0.057	0.0122	21.6445	< 0.0001	0.0002
MRE11	-0.053	0.0076	48.6682	< 0.0001	< 0.0001
RNF111 [g]	-0.061	0.0179	11.556	0.0007	0.0459
PMS2P5	-0.050	0.0098	26.5186	< 0.0001	< 0.0001
APEX1 [g]	-0.068	0.0158	18.2191	< 0.0001	0.0013
ERCC8	-0.074	0.0077	91.6536	< 0.0001	< 0.0001
OLIG2	0.067	0.019	12.4442	0.0004	0.0285
ALKBH2	-0.059	0.0041	203.2292	< 0.0001	< 0.0001
UBB	-0.017	0.0232	0.5242	0.469	1
CHEK1 [g]	-0.072	0.0183	15.6263	< 0.0001	0.0052
PRUNE2	-0.113	0.0108	108.798	< 0.0001	< 0.0001
DDB2	-0.043	0.0231	3.5194	0.0607	1
CCN4	0.079	0.0333	5.564	0.0183	1
RPA4	-0.027	0.0061	20.3553	< 0.0001	0.0004
RPA3	0.027	0.0099	7.4999	0.0062	0.4196
EPSTI1	0.051	0.0278	3.3488	0.0673	1
GSTA1 [g]	-0.080	0.0057	192.0881	< 0.0001	< 0.0001

Increase or decrease in proportion H5mCG ≥0.050 that reaches statistical significance (Bonferroni corrected $P < 0.05$).

Increase or decrease in proportion accessibility (G5mCH) ≥0.050 that reaches statistical significance (Bonferroni corrected $P < 0.05$).

[a] For promoters with ≥50 reads in the combined duplicates, but ≥22 reads in any of the 4 samples, i.e., both NSC duplicates and both Nx18-15 duplicates.

[b] Read number of sample with the least amount of reads.

[c] Proportion of difference in H5CG or accessibility (GBM Nx18-25 minus NSC non-cancerous cells).

[d] SE, standard error.

[e] Bonferroni method corrected (alpha = .05).

[f] Values calculated for HCG or GCH sites for −80 to +45 relative to TSS within full-length target from −246 to +199.

[g] Harbors one or more genetic variants (see Supplementary Table 7) that could contribute to the observed epigenetic differences.

Supplementary Table 6: Differential H5mCG and G5mCH in NSC and GBM Nx18-25 among ~450-nt targets determined by MAPit-FENGC using EM-seq; 10% cutoff.[a]

Differential DNA methylation (H5mCG)						
Parameter / Promoter	CCS read number [b]	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	43	0.575	0.0327	309.3459	<0.0001	<0.0001
HOXB2 [g]	24	0.506	0.0021	57278.74	<0.0001	<0.0001
HOTAIR [g]	968	0.491	0.0032	23831.2	<0.0001	<0.0001
TAF1	168	0.233	0.0168	193.0816	<0.0001	<0.0001
APEX2	701	0.145	0.0125	134.2705	<0.0001	<0.0001
MCM9	655	0.109	0.0038	799.8688	<0.0001	<0.0001
PELO	89	0.067	0.0021	965.2265	<0.0001	<0.0001
EPM2AIP1	518	0.007	0.0014	21.6546	<0.0001	0.0002
PRUNE2	238	-0.065	0.0084	60.2971	<0.0001	<0.0001
RPA4	3285	-0.111	0.0018	3966.598	<0.0001	<0.0001
RPA3	154	-0.231	0.0336	47.4977	<0.0001	<0.0001
EPSTI1	70	-0.355	0.0412	74.0241	<0.0001	<0.0001
GSTA1 [g]	1580	-0.748	0.0028	73279.51	<0.0001	<0.0001

Differential Accessibility (G5mCH)					
Parameter / Promoter	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	-0.201	0.0054	1398.657	<0.0001	<0.0001
HOXB2	0.015	0.0049	9.255	0.0023	0.1597
HOTAIR	-0.067	0.0049	189.452	<0.0001	<0.0001
TAF1	-0.102	0.0051	401.6285	<0.0001	<0.0001
APEX2	-0.066	0.0192	11.6487	0.0006	0.0437
MCM9	-0.001	0.0109	0.0106	0.9182	1
PELO	-0.113	0.0153	53.8195	<0.0001	<0.0001
EPM2AIP1	-0.102	0.0191	28.259	<0.0001	<0.0001
PRUNE2	-0.113	0.0108	108.798	<0.0001	<0.0001
RPA4	-0.027	0.0061	20.3553	<0.0001	0.0004
RPA3	0.027	0.0099	7.4999	0.0062	0.4196
EPSTI1	0.051	0.0278	3.3488	0.0673	1
GSTA1	-0.080	0.0057	192.0881	<0.0001	<0.0001

Increase or decrease in proportion H5mCG ≥0.10 that reaches statistical significance (Bonferroni corrected *P* value < 0.05).
Increase or decrease in proportion accessibility (G5mCH) ≥0.10 that reaches statistical significance (Bonferroni corrected *P* value < 0.05).

[a] For promoters in which ≥50 reads obtained in combined duplicates, but no less than ≥22 reads in any of the 4 samples, i.e., both NSC duplicates and both Nx18-15 duplicates.
[b] Read number of sample with the least amount of reads.
[c] Proportion of difference in H5CG or accessibility (GBM Nx18-25 minus NSC non-cancerous cells).
[d] SE, standard error.
[e] Bonferroni method corrected (alpha = .05).
[g] Harbors one or more genetic variants (see Supplementary Table 7) that could contribute to the observed epigenetic differences.

Supplementary Table 6: Differential H5mCG and G5mCH in NSC and GBM Nx18-25 among ~ 450-nt targets determined by MAPit-FENGC using EM-seq; 20% cutoff.[a]

Differential DNA methylation (H5mCG)						
Parameter / Promoter	CCS read number [b]	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	43	0.575	0.0327	309.3459	< 0.0001	< 0.0001
HOXB2 [f]	24	0.506	0.0021	57278.74	< 0.0001	< 0.0001
HOTAIR [f]	968	0.491	0.0032	23831.2	< 0.0001	< 0.0001
TAF1	168	0.233	0.0168	193.0816	< 0.0001	< 0.0001
RPA3	154	-0.231	0.0336	47.4977	< 0.0001	< 0.0001
EPSTI1	70	-0.355	0.0412	74.0241	< 0.0001	< 0.0001
GSTA1 [f]	1580	-0.748	0.0028	73279.51	< 0.0001	< 0.0001

Differential Accessibility (G5mCH)					
Parameter / Promoter	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	-0.201	0.0054	1398.657	< 0.0001	< 0.0001
HOXB2	0.015	0.0049	9.255	0.0023	0.1597
HOTAIR	-0.067	0.0049	189.452	< 0.0001	< 0.0001
TAF1	-0.102	0.0051	401.6285	< 0.0001	< 0.0001
RPA3	0.027	0.0099	7.4999	0.0062	0.4196
EPSTI1	0.051	0.0278	3.3488	0.0673	1
GSTA1	-0.080	0.0057	192.0881	< 0.0001	< 0.0001

Increase or decrease in proportion H5mCG ≥ 0.20 that reaches statistical significance (Bonferroni corrected *P* value < 0.05).
Increase or decrease in proportion accessibility (G5mCH) ≥ 0.20 that reaches statistical significance (Bonferroni corrected *P* value < 0.05).

[a] For promoters in which ≥50 reads obtained in combined duplicates, but no less than ≥22 reads in any of the 4 samples, i.e., both NSC duplicates and both Nx18-15 duplicates.
[b] Read number of sample with the least amount of reads.
[c] Proportion of difference in H5CG or accessibility (GBM Nx18-25 minus NSC non-cancerous cells).
[d] SE, standard error.
[e] Bonferroni method corrected (alpha = .05).
[f] Harbors one or more genetic variants (see Supplementary Table 7) that could contribute to the observed epigenetic differences.

Supplementary Table 6: Differential H5mCG and G5mCH in NSC and GBM Nx18-25 among ~ 450-nt targets determined by MAPit-FENGC using EM-seq; 40% cutoff.[a]

Differential DNA methylation (H5mCG)						
Parameter / Promoter	CCS read number [b]	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	43	0.575	0.0327	309.3459	< 0.0001	< 0.0001
HOXB2 [f]	24	0.506	0.0021	57278.74	< 0.0001	< 0.0001
HOTAIR [f]	968	0.491	0.0032	23831.2	< 0.0001	< 0.0001
GSTA1 [f]	1580	-0.748	0.0028	73279.51	< 0.0001	< 0.0001

Differential Accessibility (G5mCH)					
Parameter / Promoter	Methylation change [c]	SE [d]	Wald	*P* value	Adjusted *P* value [e]
CD44	-0.201	0.0054	1398.657	< 0.0001	< 0.0001
HOXB2	0.015	0.0049	9.255	0.0023	0.1597
HOTAIR	-0.067	0.0049	189.452	< 0.0001	< 0.0001
GSTA1	-0.080	0.0057	192.0881	< 0.0001	< 0.0001

Increase or decrease in proportion H5mCG ≥ 0.40 that reaches statistical significance (Bonferroni corrected *P* value < 0.05).

Increase or decrease in proportion accessibility (G5mCH) ≥ 0.40 that reaches statistical significance (Bonferroni corrected *P* value < 0.05).

[a] For promoters in which ≥ 50 reads obtained in combined duplicates, but no less than ≥ 22 reads in any of the 4 samples, i.e., both NSC duplicates and both Nx18-15 duplicates.

[b] Read number of sample with the least amount of reads.

[c] Proportion of difference in H5CG or accessibility (GBM Nx18-25 minus NSC non-cancerous cells).

[d] SE, standard error.

[e] Bonferroni method corrected (alpha = .05).

[f] Harbors one or more genetic variants (see Supplementary Table 7) that could contribute to the observed epigenetic differences.

Supplementary Table 7: SNPs and indels detected by FENGC using standard PCR of 119 ~450-nt human targets; SNPs.[a]

Sample [b] / Promoter	Captured strand (+/-)	Num_SNPs	Location using hg38	dbSNP	Reference base	Variant	Amino acid change	Fraction of reads with variant								Notes
								NSC-1		NSC-2		Nx18-25-1		Nx18-25-2		
ALKBH1	-	0														
ALKBH2	-	2	chr12:109,093,535	rs76163903	G	A	N/A	101/221	45.70%	106/207	51.21%	117/209	55.98%	111/176	63.07%	Eliminates a GCH
			chr12:109,093,559	rs1623845	C	T	N/A	221/221	100.00%	207/207	100.00%	209/209	100.00%	176/177	99.44%	
ALKBH3	+	0														
APEX1 [c]	+	1	chr14:20,454,990	rs1760944	T	G	N/A	53/53	100.00%	47/52	90.38%	0/85	0.00%	0/92	0.00%	NSC-specific SNP
APEX2	+	0														
ATM	+	0														
ATRX	-	0														
ATXN3	-	0														
AXIN2	-	0														
BRCA1	-	0														
BRCA2	+	0														
CCN4	+	0														
CD44	-	0														
CDH1	+	1	chr16:68,737,131	rs16260	C	A	N/A	0/25	0.00%	0/22	0.00%	20/57	35.09%	14/34	41.18%	GBM-specific SNP
CDH13	+	0														
CDH2	-	0														
CDKN2A	-	0														
CDKN2D	-	0														
CETN1	+	0														
CHEK1 [c]	+	3	chr11:125,625,127	rs2298483	C	T	N/A	0/61	0.00%	4/67	5.97%	23/58	39.66%	43/82	52.44%	GBM-specific SNP that eliminates an HCG site at -9 bp (relative to the TSS)
			chr11:125,625,149	rs61612408	G	A	N/A	0/61	0.00%	6/66	9.09%	35/60	58.33%	38/81	46.91%	GBM-specific SNP that eliminates a GCG at +14 bp (relative to the TSS)
			chr11:125,625,183	rs61451117	C	T	N/A	0/60	0.00%	5/58	8.62%	29/57	50.88%	30/79	37.97%	GBM-specific SNP
CHEK2	-	3	chr22:28,741,719		A	G	N/A	45/186	24.19%	0/264	0.00%	45/162	27.78%	39/199	19.60%	
			chr22:28,741,882	rs2236141	C	T	N/A	95/199	47.74%	106/256	41.41%	0/162	0.00%	0/197	0.00%	NSC-specific SNP that eliminates an HCG
			chr22:28,741,956	rs2236142	G	C	N/A	113/202	55.94%	148/255	58.04%	98/198	49.49%	100/206	48.54%	
CSDE1	-	0														
DDB2 [c]	+	4	chr11:47,214,657		T	A	N/A	3/53	5.66%	0/57	0.00%	8/58	13.79%	19/68	27.94%	
			chr11:47,214,743	rs3758666	A	G	N/A	23/50	46.00%	24/48	50.00%	18/50	36.00%	33/75	44.00%	
			chr11:47,214,747	rs3758667	A	G	N/A	23/50	46.00%	24/48	50.00%	18/50	36.00%	36/75	48.00%	
			chr11:47,214,854	rs4237547	C	G	N/A	20/41	48.78%	21/48	43.75%	0/50	0.00%	0/68	0.00%	NSC-specific SNP that eliminates a GCG at -120 bp from TSS
DRD3	-	0														
DRD4	+	0														
EGFR	+	0														
EPM2AIP1	+	0														
EPSTI1	-	2	chr13:42,992,237	rs3825511	A	C	N/A	8/22	36.36%	10/20	50.00%	29/64	45.31%	38/87	43.68%	
			chr13:42,992,240		A	G	N/A	4/22	18.18%	4/20	20.00%	7/65	10.77%	5/87	5.75%	
ERCC1	-	1	chr19:45,423,812		C	T	N/A	4/59	6.78%	9/49	18.37%	13/80	16.25%	14/94	14.89%	Eliminates a GCG
ERCC2	-	0														
ERCC8 [c]	+	2	chr5:60,945,159	rs158922	A	G	N/A	50/114	43.86%	58/129	44.96%	183/187	97.86%	121/122	99.18%	Higher-prevalence SNP in GBM that creates a GCH site
			chr5:60,945,315	rs158921	G	A	no	41/107	38.32%	45/123	36.59%	176/178	98.88%	117/117	100.00%	Higher-prevalence SNP in GBM
EXO1	+	5	chr1:241,847,894		G	A	N/A	0/3	0.00%	2/5	40.00%	5/17	29.41%	13/24	54.17%	Higher-prevalence SNP that eliminates a GCH site
			chr1:241,847,942	rs1776181	A	G	N/A	0/5	0.00%	2/5	40.00%	16/18	88.89%	19/24	79.17%	Higher-prevalence SNP in GBM that creates an HCG and a GCH, i.e., a CGC
			chr1:241,848,042	rs1776180	C	G	N/A	0/5	0.00%	3/6	50.00%	17/18	94.44%	16/21	76.19%	Higher-prevalence SNP in GBM that \eliminates a GCH
			chr1:241,848,104	rs1776179	T	C	N/A	6/6	100.00%	5/5	100.00%	27/31	87.10%	18/22	81.82%	Creates an HCG
			chr1:241,848,232	rs1776178	A	G	N/A	0/6	0.00%	3/4	75.00%	26/31	83.87%	19/29	65.52%	Creates a GCH
FANCC	-	0														
FEN1	+	2	chr11:61,792,350		G	A	N/A	2/37	5.41%	3/61	4.92%	0/68	0.00%	9/77	11.69%	Eliminates an HCG and a GCH
			chr11:61,792,609	rs174538	G	A	N/A	29/57	50.88%	32/64	50.00%	69/72	95.83%	101/102	99.02%	Higher-prevalence SNP in GBM that eliminates a GCH site
FHIT	-	0														
FOXC2	+	0														
FZD7	+	0														
FZD8	-	0														
GPR158	+	0														
GSTA1 [c]	-	4	chr6:52,803,872	rs3957356	T	C	N/A	0/50	0.00%	0/50	0.00%	37/37	100.00%	40/41	97.56%	GBM-specific SNP at -59 bp from TSS
			chr6:52,803,889	rs3957357	A	G	N/A	0/50	0.00%	0/50	0.00%	43/44	97.73%	40/42	95.24%	GBM-specific SNP -74 bp from TSS
			chr6:52,803,986	rs3996998	C	A	N/A	167/167	100.00%	177/177	100.00%	70/70	100.00%	62/63	98.41%	
			chr6:52,804,145	rs9296692	T	C	N/A	154/154	100.00%	175/175	100.00%	75/75	100.00%	63/63	100.00%	
HIST1H1B	-	0														
HMGA2	+	0														
HOTAIR [c]	-	2	chr12:53,968,724		C	T	N/A	24/196	12.24%	27/220	12.27%	0/50	0.00%	0/213	0.00%	Higher-prevalence SNP in NSC that eliminates a GCH
			chr12:53,968,802	rs74613444	T	C	N/A	62/204	30.39%	65/233	27.90%	0/174	0.00%	0/136	0.00%	NSC-specific SNP
HOXB2	-	1	chr17:48,544,929		G	T	N/A	0/74	0.00%	0/73	0.00%	0/113	0.00%	15/77	19.48%	
HSPA5	-	0														
INO80	-	0														
KLF8	+	0														
LIG1	-	0														
MBD4	-	0														
MCM9	-	1	chr6:118,931,610		G	T	Tyr to Stop codon	36/223	16.14%	42/226	18.58%	0/256	0.00%	39/197	19.80%	
MECP2	-	0														
MEG3	+	1	chr14:100,825,826	rs1382379428	C	A	N/A	5/54	9.26%	7/64	10.94%	8/55	14.55%	13/70	16.46%	
MGMT	+	0														
MLH1	+	1	chr3:36,993,455	rs1800734	G	A	N/A	0/91	0.00%	4/68	5.88%	20/59	33.90%	45/70	64.29%	GBM-specific SNP at -62 bp from TSSa
MLH3	-	1	chr14:75,051,627	rs175083	G	A	N/A	15/15	100.00%	21/21	100.00%	45/45	100.00%	38/38	100.00%	Eliminates a GCH
MNAT1	+	0														
MPG	+	0														
MRE11	-	1	chr11:94,493,959	rs11020802	C	A	N/A	31/76	40.79%	27/72	37.50%	44/95	46.32%	49/92	53.26%	Eliminates an HCG and GCH, i.e., a CGC, at -115 bp from MRE11 TSS
MSH2	-	0														
MSH3	+	0														
MSH4	+	3	chr1:75,796,838	rs721204	G	A	N/A	0/30	0.00%	2/31	6.45%	19/39	48.72%	16/37	43.24%	GBM-specific SNP that eliminates a GCH
			chr1:75,797,013	rs777678581	T	C	Ser to Pro	5/21	23.81%	6/27	22.22%	8/34	23.53%	5/31	16.13%	
			chr1:75,797,015	rs369331894	G	A		5/21	23.81%	6/27	22.22%	8/34	23.53%	5/29	17.24%	Eliminates an HCG and GCH, i.e., a CGC, at +134 bp from MSH2 TSS
MSH5	+	0														
MSH6	+	0														
MUTYH	-	1	chr1:45,340,760	rs3219463	C	T	N/A	48/49	97.96%	33/47	70.21%	0/0	0.00%	0/0	0.00%	NSC-specific SNP
NEIL2	+	0														
NKX2-5	-	0														
OGG1	+	0														
OLIG2	+	1	chr21:33,025,870	rs2009130	C	G	N/A	25/51	49.02%	25/66	37.88%	27/49	55.10%	28/46	60.87%	Eliminates an HCG
PARP1	-	0														
PCNA	-	0														
PELO	+	0														
PMS1	+	1	chr2:189,784,035		C	A	N/A	3/15	20.00%	0/21	0.00%	3/28	10.71%	2/31	6.45%	
PMS2	-	2	chr7:6,009,397	rs1367182950	G	A	Gly to Arg	0/25	0.00%	0/25	0.00%	0/26	0.00%	6/40	15.00%	
			chr7:6,009,399	rs1385363023	C	A		0/25	0.00%	0/25	0.00%	0/26	0.00%	6/40	15.00%	Eliminates a GCG
PMS2P1	-	0														
PMS2P11	+	0														
PMS2P2	-	0														
PMS2P3	-	0														
PMS2P5	+	0														
POLD1	+	0														
POLD4	-	0														
POLG	-	0														
POLK	+	0														
POLQ	-	1	chr3:121,546,196	rs13065220	C	T	N/A	68/168	40.48%	94/197	47.72%	119/252	47.22%	112/217	51.61%	
POLR2B	+	1	chr4:56,978,736	rs1713998	A	G	N/A	23/23	100.00%	8/8	100.00%	0/9	0.00%	0/20	0.00%	NSC-specific SNP that creates an HCG
POLR2G	+	0														
PRDM2	+	1	chr1:13,704,576		G	C	N/A	3/59	5.08%	7/57	12.28%	5/66	7.58%	13/81	16.05%	Eliminates a GCH
PRKCG	+	0														
PRUNE2	-	0														
PXMP4	-	0														
PYCARD	-	0														
RNASEH2A	+	0														
RNF111 [c]	+	1	chr15:58,987,672	rs112968854	C	G	N/A	21/52	40.38%	23/50	46.00%	0/88	0.00%	0/89	0.00%	NSC-specific SNP that eliminates a GCH
RPA2	-	0														
RPA3	-	0														
RPA4	+	0														
SH3TC1	+	0														
SIM2	+	0														
SMAD4	+	0														
SMUG1	-	0														
SNAI1	+	0														
SNAI2	-	0														
TAF1	+	0														
TDG	+	0														
TP53	-	1	chr17:7,687,470	rs1406382382	C	A	N/A	23/99	23.23%	22/95	23.16%	18/89	20.22%	18/100	18.00%	Eliminates an HCG
TREX1	+	0														
TRIM28	+	0														
TWIST1	-	0														
TWIST2	+	0														
UBB [c]	+	1	chr17:16,380,612	rs188541510	C	T	N/A	17/51	33.33%	18/47	38.30%	0/62	0.00%	0/66	0.00%	NSC-specific SNP

USP47	+	0														
WDR48	+	1	chr3:39,051,835	rs6599003	G	A	N/A	3/7	42.86%	2/5	40.00%	14/28	50.00%	23/41	56.10%	
WNT1	+	0														
XPC	-	0														
XRCC2	-	0														
XRCC3	-	0														
ZEB1	+	0														
ZNF830 [c]	+	1	chr17:34,961,344	rs9635769	C	T	Arg to Gln	160/160	100.00%	123/128	96.09%	50/113	44.25%	70/138	50.72%	Higher-prevalence SNP in NSC; eliminates an HCG -196 bp from ZNF830 TSS

[a] High-fidelity CCS reads with ≥3 SMRT sequencing passes.

[b] Suffixes -1 and -2 refer to the two independent biological replicates.

[c] Locus with significant epigenetic differences in NSC vs. GBM (Supplementary Table 6) that also exhibits differential variant frequency (detected in at least 5 reads in both replicates).

GBM-specific SNP

NSC-specific SNP

Unannotated SNP detected with Q >30 and depth ≥5 in both two biological replicates.

Supplementary Table 7: SNPs and indels detected by FENGC using standard PCR of 119 ~ 450-nt human targets; indels.[a]

Promoter \ Sample [b]	Reads_NSC-1	Reads_NSC-2	Reads_Nx18-25-1	Reads_Nx18-25-2	Captured strand (+/-)	Num_indels (GQ>30, Depth>15)	Location using hg38	dbSNP	Reference base	Variant	Amino acid change	Fraction of reads with variant				Notes
												NSC-1	NSC-2	Nx18-25-1	Nx18-25-2	
ALKBH1	6	5	24	17	-	0										
ALKBH2	2334	2017	1634	1879	-	0										
ALKBH3	4	6	15	9	+	0										
APEX1	130	159	210	240	+	0										
APEX2	390	321	1251	1293	+	0										
ATM	290	349	702	720	+	0										
ATRX	207	340	551	492	-	0										
ATXN3	0	0	0	0	-	0										
AXIN2	0	0	0	0	-	0										
BRCA1	222	182	254	278	-	0										
BRCA2	73	128	240	274	+	0										
CCN4	249	292	499	497	+	1	chr8:133,190,920-133,190,921	rs548251181	GA	G	N/A	29/59	49.15%	14/54	25.93%	36/94	38.30%	31/93	33.33%	Unmapped sequence plotted in white in Fig. 4f and Extended Data Fig. 20c
CD44	6	12	11	19	-	0										
CDH1	28	23	68	49	+	0										
CDH13	0	2	2	6	+	0										
CDH2	10	9	2	4	-	0										
CDKN2A	0	0	0	0	-	0										
CDKN2D	0	0	0	0	-	0										
CETN1	0	2	3	9	+	0										
CHEK1	651	605	526	608	+	0										
CHEK2	2522	2547	2232	2700	-	0										
CSDE1	21	13	30	32	-	0										
DDB2	199	281	281	284	+	1	chr11:47,214,771-47,214,773	rs781389544	TAA	TA	N/A	16/49	32.65%	12/48	25.00%	6/51	11.76%	27/79	34.18%	
										T	N/A	7/49	14.29%	6/48	12.50%	9/51	17.65%	17/79	21.52%	
DRD3	217	160	241	241	-	0										
DRD4	0	0	0	0	+	0										
EGFR	0	0	3	2	+	0										
EPM2AIP1	1877	1374	1135	1311	+	0										
EPSTI1	29	23	93	111	-	0										
ERCC1	72	67	194	200	-	0										
ERCC2	3	0	3	11	-	0										
ERCC8	622	483	823	702	+	0										
EXO1	6	8	36	38	+	0										
FANCC	2	3	0	0	-	0										
FEN1	447	367	268	339	+	0										
FHIT	97	89	103	100	-	0										
FOXC2	0	0	0	0	+	0										
FZD7	0	0	0	0	+	0										
FZD8	3	1	7	1	-	0										
GPR158	9	16	21	35	+	0										
GSTA1	2770	2549	2850	3147	-	0										
HIST1H1B	560	461	545	531	-	0										
HMGA2	2	10	15	23	+	0										
HOTAIR [c]	2467	2131	1896	2089	-	2	chr12:53,968,800-53,968,801	rs199734232	AT	A	N/A	0/196	0.00%	0/222	0.00%	47/174	27.01%	46/147	31.29%	GBM-specific deletion
							chr12:53,968,801	rs71227279	T	TC	N/A	142/204	69.61%	158/232	68.10%	0/174	0.00%	0/147	0.00%	NSC-specific insertion
HOXB2 [c]	428	421	350	337	-	3	chr17:48,544,973	rs111979968	A	AC	N/A	79/82	96.34%	78/85	91.76%	76/77	98.70%	76/84	90.48%	
							chr17:48,544,986-48,544,987	rs138169880	AC	A	N/A	7/89	7.87%	13/97	13.40%	3/82	3.66%	1/87	1.15%	
										ACC	N/A	3/89	3.37%	2/97	2.06%	23/82	28.05%	32/87	36.78%	GBM-specific insertion
							chr17:48,545,041-48,545,042	rs901811939	CT	C	N/A	31/109	28.44%	41/118	34.75%	40/107	37.38%	35/99	35.35%	
HSPA5	1	9	10	32	-	1	chr9:125,241,509-125,241,510	rs3216733	GC	G	N/A	0/1	0.00%	2/5	40.00%	4/7	57.14%	7/22	31.82%	Eliminates an GCH site
INO80	0	0	0	0	-	0										
KLF8	9	0	18	11	+	0										
LIG1	128	85	109	149	-	0										
MBD4	13	24	71	61	-	0										
MCM9	3836	3105	2793	2917	-	1	chr6:118,931,744-118,931,745	rs768549255	GA	G	N/A	148/356	41.57%	158/377	41.91%	109/309	35.28%	128/278	46.04%	
MECP2	0	0	0	0	-	0										
MEG3	91	81	108	172	+	1	chr14:100,825,832		A	C	N/A	6/60	10.00%	0/69	0.00%	0/60	0.00%	0/87	0.00%	
										ACC	N/A	0/60	0.00%	0/69	0.00%	0/60	0.00%	8/87	9.20%	
MGMT	0	0	0	0	+	0										
MLH1	1872	1369	1122	1300	+	0										
MLH3	21	29	63	56	-	1	chr14:75,051,672-75,051,673	rs200130711	CT	C	N/A	2/15	13.33%	2/21	9.52%	15/45	33.33%	14/38	36.84%	Higher prevalence deletion in GBM
MNAT1	26	30	130	84	+	0										
MPG	0	0	0	0	+	0										
MRE11	185	173	312	279	-	0										
MSH2	881	529	737	762	-	1	chr2:47,402,751-47,402,752	rs544188883	GT	G	N/A	17/57	29.82%	16/53	30.19%	27/72	37.50%	20/72	27.78%	
MSH3	0	0	0	0	+	0										
MSH4	30	39	48	50	+	0										
MSH5	30	11	7	3	+	0										
MSH6	0	0	0	0	+	0										
MUTYH	55	51	240	234	-	0										
NEIL2	0	0	0	0	+	0										
NEIL3	8	15	43	55	+	0										
NKX2-5	10	10	26	22	-	0										
OGG1	58	19	43	63	+	0										
OLIG2	163	194	98	105	+	0										
PARP1	0	0	0	0	-	0										
PCNA	73	69	157	234	-	0										
PELO	62	57	69	66	+	1	chr5:52,788,092		T	TA	N/A	40/40	100.00%	29/29	100.00%	38/38	100.00%	42/43	97.67%	
PMS1	16	23	36	47	+	0										
PMS2	38	67	164	269	-	0										
PMS2P1	289	339	297	352	-	0										
PMS2P11	2616	2081	1356	1479	+	0										
PMS2P2	742	966	1288	1531	-	1	chr7:75,359,192-75,359,194	rs879967661	GTT	G	N/A	6/46	13.04%	6/43	13.95%	0/2	0.00%	0/0	0.00%	
										GTTT	N/A	13/46	36.11%	7/43	16.28%	0/2	0.00%	0/0	0.00%	
PMS2P3	13	15	66	77	-	0										
PMS2P5	737	968	1291	1533	+	1	chr7:74,890,587-74,890,589	rs376727434	GAA	G	N/A	5/45	11.11%	5/50	10.00%	0/9	0.00%	4/9	44.44%	
										GA	N/A	31/45	68.89%	33/50	66.00%	8/9	88.89%	4/9	44.44%	
POLD1	0	0	0	0	+	0										
POLD4	731	379	512	739	-	0										
POLG	0	0	0	2	-	0										
POLK	50	135	166	209	+	0										
POLQ	658	679	1029	987	-	0										
POLR2B	30	8	14	24	+	0										
POLR2G	2	2	11	26	+	1	chr11:62,761,398-62,761,399	rs1222984274	CA	C	N/A	0/1	0.00%	0/2	0.00%	0/9	0.00%	12/22	54.55%	
PRDM2	822	875	1145	1250	+	0										
PRKCG	9	3	9	22	+	0										
PRUNE2	7	8	0	14	-	0										
PXMP4	4	14	66	76	-	0										
PYCARD	0	0	0	2	-	0										
RNASEH2A	1	2	16	7	+	0										
RNF111	132	135	218	204	+	0										
RPA2	309	509	754	739	-	0										
RPA3	1353	1034	1134	1265	-	0										
RPA4	1211	1047	1075	1161	+	0										
SH3TC1	3	2	4	9	+	0										
SIM2	0	0	0	0	+	0										
SMAD4	3	0	10	2	+	0										
SMUG1	32	34	60	103	-	0										
SNAI1	0	0	0	0	+	0										
SNAI2	0	3	27	23	-	0										
TAF1	147	164	443	422	+	0										
TDG	3	6	23	27	+	0										
TP53	330	337	463	485	-	0										
TREX1	0	0	0	2	+	0										
TRIM28	1	2	0	2	+	0										
TWIST1	1	0	2	4	-	0										
TWIST2	0	0	0	0	+	0										
UBB	101	94	83	94	+	1	chr17:16,380,582-16,380,583	rs1213064550	GA	G	N/A	14/51	27.45%	11/48	22.92%	20/51	39.22%	15/50	30.00%	
USP47	4	5	1	9	+	0										
WDR48	7	9	34	52	+	0										
WNT1	13	5	0	0	+	0										
XPC	0	4	0	2	-	0										
XRCC2	36	67	34	44	-	0										
XRCC3	0	0	0	0	-	0										
ZEB1	0	0	0	0	+	0										
ZNF830	2471	1826	1690	1703	+	1	chr17:34,961,414-34,961,415	rs368850951	GA	G	N/A	60/171	35.09%	50/138	36.23%	42/138	30.43%	50/156	32.05%	

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.
[b] Suffixes -1 and -2 refer to the two independent biological replicates.
[c] Locus with significant epigenetic differences in NSC vs. GBM (Supplementary Table 6) that also exhibits differential variant frequency (detected in at least 5 reads in both replicates).
GBM-specific indel
NSC-specific indel
Unannotated indel detected with Q > 30 and depth ≥ 29 in both two biological replicates.

Supplementary Table 8: CCS reads on- and off-target for 45 ~940-nt human targets in MAPit-FENGC using EM-seq. [a]

Sample [b]	800 ng input-1	800 ng input-2	400 ng input-1	400 ng input-2
Total reads	179,458	130,943	54,023	57,231
Reads aligned on genome	165,646	121,108	50,873	53,592
% Reads aligned on genome	92.3%	92.5%	94.2%	93.6%
Reads aligned on target	145,070	105,794	41,467	44,214
% Reads off target but on genome	12.4%	12.6%	18.5%	17.5%
% Total reads on target	80.8%	80.8%	76.8%	77.3%
Number of targets detected	38		38	
Number of targets detected using reads with ≥95% coverage of each reference sequence and ≥95% conversion	25		20	

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.

[b] Suffixes -1 and -2 refer to the two independent biological replicates.

Supplementary Table 9: Filtered CCS reads aligned to 45 ~940-nt human targets in MAPit-FENGC using EM-seq. [a]

Gene	Capture strand length (nt)	%GC	Filtered CCS read number [b]			
			800 ng input-1 [c]	800 ng input-2	400 ng input-1	400 ng input-2
ALKBH2	946	54	0	0	0	0
APEX1	937	56	418	302	109	99
APEX2	940	55	60	50	21	16
ATM	938	58	730	547	155	136
ATRX	930	43	0	0	0	0
BRCA1	937	51	0	0	0	0
BRCA2	935	57	0	0	0	0
CCN4	944	53	255	191	62	49
CD44	942	53	0	0	0	0
CETN1	930	54	0	2	0	0
CHEK1	953	44	0	0	0	0
CHEK2	950	61	0	0	1	0
CSDE1	945	47	1	0	0	0
DDB2	942	58	3	3	1	0
EPM2AIP1-MLH1	937	59	47174	33315	8989	8286
EPSTI1	934	47	1	0	0	0
EXO1	951	53	4	2	0	0
FEN1	942	54	1	1	0	0
GSTA1	939	39	0	0	0	0
HIST1H1B	952	45	0	0	0	0
HOTAIR	959	53	0	0	0	0
HOXB2	935	49	0	0	0	0
HSPA5	951	59	25971	18679	6675	6472
LIG1	941	65	0	0	0	0
MBD4	951	56	9	5	2	1
MCM9	939	33	0	0	0	0
MECP2	957	59	0	0	0	0
MRE11	956	54	3	1	1	2
MSH2	930	52	462	331	108	86
MSH4	931	46	80	55	20	14
MSH5	931	53	4347	3234	1029	965
OGG1	946	61	10	16	2	1
OLIG2	952	63	0	0	0	0
PELO	946	50	0	0	0	0
PMS2	944	63	22554	16530	5198	4918
PMS2P5	935	64	28081	19985	8261	7756
POLD4	939	59	4512	3222	976	923
PRUNE2	945	66	29	24	7	9
RPA2	951	53	7	5	4	2
RPA4	951	42	0	0	0	0
TAF1	958	53	0	0	0	0
TP53	938	49	0	1	0	0
UBB	943	56	1	0	0	1
XRCC2	939	50	0	0	0	0
ZNF830	950	53	0	0	0	0

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.

[b] CCS reads with ≥95% coverage of each reference sequence and ≥95% HCH conversion.

[c] Suffixes -1 and -2 refer to each of the two independent biological replicates.

Supplementary Table 10: CCS reads on- and off-target for 78 ~620-nt mouse targets in MAPit-FENGC using EM-seq.[a]

Sample	Mouse 1	Mouse 2	Mouse 3	Mouse 4
Input gDNA (ng)	600	600	600	600
Total reads	172,876	147,858	168,272	108,306
Reads aligned to other amplicons in library after filtering [b]	19,537	16,271	18,598	4,588
Reads aligned to FENGC targets afer filtering [b]	122,568	101,787	106,040	76,389
Unaligned reads and reads removed by filtering [b,c]	30,771	29,800	43,634	27,329
% On-target reads, ignoring off-target reads from other amplicons	80%	77%	71%	74%
Number of targets detected [c]	73	71	73	75
Number of targets detected in each library using filtered reads [b,d]	70	68	67	71
Number of targets detected in four combined libraries using filtered reads [b,d]	74			

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.

[b] Filtered for ≥95% conversion and ≥95% coverage of each mouse reference sequence (mm9 genome build).

[c] May include off-target reads from other amplicons in library.

[d] At least one read.

Supplementary Table 11: Filtered CCS reads aligned to 78 ~620-nt mouse targets in MAPit-FENGC using EM-seq.[a]

Sample / Promoter	CCS read number [b]					
	Mouse 1	Mouse 2	Mouse 3	Mouse 4	Average CCS read number	Amplicon length
Ash1l	9	14	0	10	8	829
Btk	8412	6930	2687	1761	4948	514
Ccl5	14885	12096	7268	5135	9846	524
Cd180	216	201	290	515	306	509
Cox8a	567	376	739	570	563	646
Cxcl15	178	253	633	701	441	558
Cxcr4	1586	1738	2805	2253	2096	492
Dnmt3b	78	125	92	68	91	987
Foxo1	0	0	0	0	0	905
Gdf5	15135	15154	13733	7529	12888	491
Hdac3	2642	2129	1568	816	1789	537
Hsf1	812	566	649	307	584	474
Hsf2bp	792	475	645	260	543	484
Hsf4	301	120	87	50	140	623
Hsf5	0	0	0	0	0	636
Hsfy2	516	710	784	727	684	662
Hsp90aa1	16707	13095	19155	11218	15044	522
Hsp90ab1	3523	2401	2625	1523	2518	606
Hspa1a	0	4	0	4	2	709
Hspa1b	16	1	2	10	7	640
Hspa2	118	103	166	81	117	605
Ifna1	0	0	0	0	0	696
Ifna2	3174	3855	3031	4350	3603	533
Ifng	3892	2805	3667	3159	3381	476
Ikbkb	8	0	3	19	8	720
Ikbkg	211	279	647	666	451	554
Il10	4907	3722	3439	3110	3795	551
Il12a	14249	14043	8757	7683	11183	537
Il12b	3217	3150	5053	4153	3893	532
Il3ra	1377	1142	1412	1081	1253	615
Il6	18	8	19	16	15	807
Irak1	65	25	78	50	55	678
Irak4	1035	834	451	385	676	606
Irf3	10	0	11	12	8	557
Irf5	219	61	241	83	151	747
Irf7	952	1043	2468	2601	1766	485
Jun	1110	798	966	718	898	630
Kdm6b	14370	13361	15486	9225	13111	489
Lbp	659	850	749	758	754	564
Ly86	7615	5536	8098	5683	6733	534
Map3k1	0	88	0	0	22	677
Mapk1	1815	1389	1589	1004	1449	496
Mapk13	0	0	0	0	0	795
Mapk14	18	202	57	31	77	891
Mapk15	9818	8824	6756	5199	7649	556
Mapk3	85	102	23	21	58	707
Mapk4	507	402	398	232	385	755
Mapk6	6	11	0	3	5	563
Mapk8	1	0	2	0	1	872
Mme	1993	1088	1758	1255	1524	596
Myd88	118	70	28	46	66	628
Nfkb1	19	36	16	4	19	620
Nfkb2	10038	8147	6037	4969	7298	511
Nfkbia	55	118	158	181	128	690
Nfkbib	141	116	105	162	131	530

Nr2c2	54	67	17	7	36	507
Pik3ap1	172	84	73	51	95	613
Pik3c2a	0	0	1	0	0	883
Pik3c2g	88	102	23	408	155	660
Pik3c3	306	232	251	309	275	640
Pik3ca	48	80	90	74	73	670
Pik3r1	24	0	13	6	11	623
Pik3r2	51	179	0	29	65	815
Pik3r3	3304	2329	2313	1533	2370	605
Pik3r5	231	233	289	270	256	728
Pik3r6	12852	10250	13647	8810	11390	540
Ptpn6	2311	1920	5209	3882	3331	483
Rela	50	0	16	66	33	654
Src	4417	3135	3658	1978	3297	616
Stat3	717	320	490	225	438	549
Tab1	721	473	404	224	456	760
Tab2	0	18	0	11	7	575
Tet2	131	82	267	83	141	674
Tlr2	580	644	630	346	550	665
Tlr4	2542	1936	3654	2704	2709	576
Tnf	1	5	0	6	3	499
Traf6	1531	1150	2002	1400	1521	543
Vav1	811	729	280	257	519	605

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.

[b] Filtered for ≥95% conversion of HCH to HTH and ≥95% coverage of the length of each reference sequence.

Supplementary Table 12: Differential H5mCG and G5mCH determined by MAPit-FENGC using EM-seq of ~620-nt mouse targets.[a]

Parameter / Promoter	CCS read number [b]	Mean %H5mCG (endogenous methylation) *P* value [c,d]	Mean %G5mCH (accessibility) *P* value [c,d]
Btk	1761	0.878	0.989
Ccl5	5135	1.000	0.956
Cd180	201	0.300	0.885
Cox8a	376	0.877	0.905
Cxcl15	178	1.000	0.652
Cxcr4	1586	1.000	0.998
Gdf5	7529	1.000	1.000
Hdac3	816	1.000	0.998
Hsf1	307	1.000	0.841
Hsf2bp	260	1.000	0.582
Hsfy2	710	0.903	0.963
Hsp90aa1	11218	0.482	1.000
Hsp90ab1	1523	1.000	0.879
Ifna2	3031	1.000	0.958
Ifng	2805	0.594	0.974
Ikbkg	211	Excluded (duplicates)	
Il10	3110	0.173	0.981
Il12a	7683	0.995	0.995
Il12b	3150	1.000	0.997
Il3ra	1081	0.999	1.000
Irak4	385	1.000	0.951
Irf7	952	0.990	0.928
Jun	718	1.000	0.966
Kdm6b	9225	0.687	0.981
Lbp	659	1.000	0.968
Ly86	5536	0.996	0.618
Mapk1	1004	1.000	0.924
Mapk15	5199	0.998	0.994
Mapk4	232	1.000	0.891
Mme	1088	1.000	0.992
Nfkb2	4969	1.000	0.992
Nfkbib	105	0.997	0.492
Pik3c3	232	1.000	1.000
Pik3r3	1533	0.831	0.998
Pik3r5	231	1.000	0.863
Pik3r6	8810	1.000	0.955
Ptpn6	1920	0.906	0.923
Src	1978	0.986	0.937
Stat3	225	1.000	0.985
Tab1	224	1.000	0.971
Tlr2	346	0.890	0.971
Tlr4	1936	0.836	0.231
Traf6	1150	1.000	0.998
Vav1	257	0.979	0.907

[a] High-fidelity CCS reads with ≥5 SMRT sequencing passes.

[b] Samples with the lowest number of filtered reads from Supplementary Table 9.

[c] Limited to targets with ≥100 reads for each of the four mice with good pattern diversity, i.e., lacking many duplicates.

[d] Not corrected for multiple testing; all Bonferroni method corrected (alpha .05) *P* values were 1.000.

Supplementary Table 13: FENGC oligonucleotide concentrations for different numbers of targets.

Number of targets	Concentration of each of oligos 1-N, 2-N, and 3 (nM) in stock oligo mixtures [a-e]
10	200
20	100
30	67
40	50
50	40
60	33
70	29
80	25
90	22
100	20
110	18
120	17
130	15
140	14
150	13
160	13
170	12
180	11
190	11
200	10
210	10
220	9
230	9
240	8
250	8
260	8
270	7
280	7
290	7
300	7
310	6
320	6
330	6
340	6
350	6
360	6
370	5
380	5
390	5
400	5
410	5
420	5
430	5
440	5
450	4
460	4
470	4
480	4
490	4
500	4

[a] Concentration of separate stocks of U1-N (or U1) and U2 oligos is fixed at empirically determined 10 μM = 10,000 nM (1 μL of each is added).

[b] The total amount of all oligo 1-N and 2-N should be close to but less than that of U1-N (or U1).

[c] Flap oligo 1-N or 2-N concentration (nM) × number of targets × 4 (i.e., total added volume of oligo 1-N and 2-N is 4x of U1-N (or U1)), empirically determined to be 8,000 nM (close to but <10,000 nM).

[d] Concentration of each oligo 3-N should be the same as that of oligo 1-N or oligo 2-N in this separate stock mixture.

[e] Therefore, the concentration of each of oligo 1-N, 2-N, and 3-N is calculated by 8,000 nM/(Number of targets × 4).