

DATA NOTE

Open Access



Targeted native long-read sequencing of DNA methylation alterations following CRISPR-Cas9-induced double-strand breaks in human cells

Yingzi Zhang^{1,2†}, Mengge Wang^{1,2†}, Chongwei Bi¹ and Mo Li^{1,2*}

Abstract

Objectives CRISPR-Cas9 nucleases are widely used to introduce targeted DNA double-strand breaks (DSBs) for genome engineering, but the long-term impact of these lesions on local epigenetic information remains poorly characterized. In a companion research article, we used Cas9-assisted targeted nanopore sequencing (CTS) to reveal that CRISPR-Cas9-induced DSBs can disrupt local epigenetic maintenance across multiple genomic contexts and cell systems. Here, we present a structured description of the raw and minimally processed datasets underlying the study. These datasets provide base-resolution measurements of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) at the differentially methylated regions (DMRs) of several imprinted loci, two heterochromatic regions, a cancer-associated promoter epimutation region, and the *SNRPN* DMR at early/late passages of a clonal line. They enable re-analysis and methodological benchmarking of DSB-associated epigenetic instability.

Data description We provide aligned BAM files and per-CpG methylation calls for multiple genomic contexts under both CRISPR-targeted and non-targeting control conditions. Specifically, the collection includes: (i) imprinted loci in human embryonic stem cells (hESCs), including *small nuclear ribonucleoprotein polypeptide N (SNRPN)*, *paternally expressed 10 (PEG10)*, and *KCNQ1 opposite strand/antisense transcript 1 (KCNQ1OT1)*, (ii) heterochromatic regions in hESCs, including *urothelial cancer associated 1 (UCA1)*, and *cysteine rich C-terminal 1 (CRCT1)*, (iii) the epimutation locus of *MutL homolog 1 (MLH1)* in RKO cells, and (iv) the DMR of *SNRPN* locus in early- and late-passage derivatives of a single hESC clone. For each collection, there is a dataset that includes both the raw aligned Nanopore sequencing reads (BAM) deposited in the NCBI Sequence Read Archive (SRA) and the corresponding processed per-CpG 5mC/5hmC matrices deposited in Zenodo. All higher-level analyses in the research article—such as DMR calling, haplotype-resolved analyses, and structural variant (SV) characterization—are fully reproducible using these deposited data. Additional processed analyses are comprehensively documented in the companion article and are therefore

[†]Yingzi Zhang and Mengge Wang have contributed equally to this work.

*Correspondence:

Mo Li
mo.li@kaust.edu.sa

Full list of author information is available at the end of the article



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

not duplicated here. Together, these datasets offer a rich resource for benchmarking long-read methylation analysis workflows and further investigation of DSB-associated epigenetic instability across diverse genomic contexts.

Keywords CRISPR-Cas9, Epigenetic instability, DNA methylation, Double-strand break, Nanopore sequencing, Long-read sequencing, Genomic imprinting, Human embryonic stem cell, Colorectal cancer cell

Objective

CRISPR-Cas9 has transformed genome engineering and functional genomics, yet accumulating evidence suggests that its activity can have consequences beyond intended sequence changes [1–3], potentially altering epigenetic marks near induced lesions [4]. In particular, how DNA methylation is maintained following CRISPR-Cas9 DSB induction remains incompletely understood.

Long-read sequencing on the Oxford Nanopore Technologies (ONT) platform offers key advantages for resolving complex genetic and epigenetic alterations after CRISPR-Cas9 editing. By generating reads tens to hundreds of kilobases in length, ONT enables haplotype resolution, accurate reconstruction of SVs, and simultaneous detection of native DNA methylations—capabilities not achievable with short-read technologies. Multiple studies have demonstrated the utility of ONT long-read technologies for precisely characterizing genetic instability [5] and CRISPR outcomes [1–3]. Long-read nanopore sequencing has been instrumental in uncovering pervasive unintended SVs at Cas9-induced breaks, highlighting previously unrecognized genotoxic risks of CRISPR editing [1, 3], and has been used as a quantitative, unbiased tool to evaluate genomic repair outcomes [2].

Beyond genetic characterization, ONT provides direct, base-resolution detection of DNA modifications, allowing simultaneous mapping of epigenetic and genetic outcomes on the same reads. Insights from methylation biology further motivate this approach. Yuta et al. [6] showed that targeted perturbations—such as insertion of CpG-free DNA—can induce stable *de novo* CpG island methylation, generate pathogenic epimutations, or correct imprinting defects in human pluripotent stem cells. Additional studies have demonstrated how methylation states contribute to disease traits and may influence epigenetic inheritance [7]. Meanwhile, recent benchmarking efforts have enabled highly accurate ONT-based base-resolution DNA methylation detection [8–12]. Enrichment workflows such as CTS support high-depth profiling of selected genomic regions of interest (ROIs) without PCR amplification [13]. Together, these developments position ONT long-read sequencing as a powerful platform for interrogating how CRISPR-Cas9-induced DSBs influence the maintenance and stability of DNA methylation.

In the companion research article, we used CTS to examine how Cas9-induced DSBs affect DNA methylation at several genomic contexts in human cells, focusing

on (i) canonical imprinted loci with well-defined parental methylation asymmetry, (ii) heterochromatic regions, and (iii) a cancer-relevant epimutant locus in the *MLH1* promoter (Fig. 1). We obtained deep coverage of pre-defined ROIs in H1 hESCs and RKO colorectal cancer cells. CTS datasets were also generated from early- and late-passage derivatives of an *SNRPN*-DMR^{Me/Me} clonal line to assess locus-specific methylation stability during prolonged culture. With read lengths typically spanning tens of kilobases and enabling haplotype resolution, methylation calling, and SV detection, these data allow characterization of the interplay between DSB repair and epigenetic maintenance.

The purpose of this Data Note is to describe the raw and minimally processed datasets needed for future research applications. By depositing raw reads, processed methylation calls, and sample metadata, we aim to support reproducibility of DSB-associated epigenetic analyses and facilitate downstream applications such as benchmarking long-read methylation callers, evaluating DMR detection frameworks, assessing relationships between SVs and local epigenetic perturbations, and generating hypotheses on mechanisms underlying DSB-induced epigenetic instability.

Data description

Experimental design and sample groups

CTS was performed in two human cell lines to investigate how CRISPR-Cas9-induced DSBs affect the maintenance of CpG methylation across diverse genomic contexts (Fig. 1; Table 1). In H1 hESCs, CTS was applied to interrogate three classical imprinted loci (*SNRPN*, *PEG10*, and *KCNQ1OT1*) with strong allelic methylation asymmetry and two densely methylated heterochromatic regions (*UCA1* and *CRCT1*). For each locus, cells were transfected with either a single-guide RNA (sgRNA) targeting an internal site to induce a DSB or with a non-targeting sgRNA serving as a control. This design enabled high-depth profiling of parental-allele-specific methylation states at loci with well-defined imprinting asymmetry, as well as the examination of methylation stability within heterochromatin.

CTS was also performed in the RKO colorectal cancer cell line to interrogate an epimutation in the *MLH1* promoter. RKO cells were subjected to CRISPR-Cas9 editing at the *MLH1* locus or treated with a non-targeting sgRNA, enabling direct comparison of methylation and genetic outcomes in a cancer-relevant context.

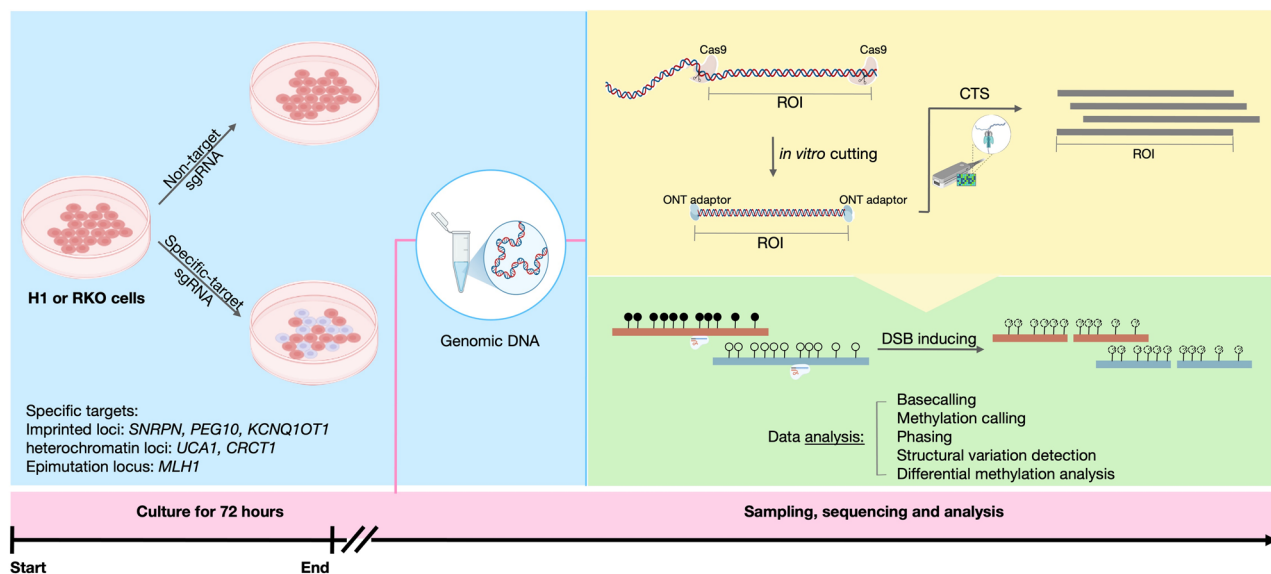


Fig. 1 Schematic overview of the experimental design and data-processing pipeline. H1 hESCs or RKO cells were transfected with either locus-specific or non-targeting control sgRNAs and cultured for 72 h prior to genomic DNA isolation. Target regions—including imprinted loci (*SNRPN*, *PEG10*, *KCNQ1OT1*), heterochromatic loci (*UCA1*, *CRCT1*), and the *MLH1* epimutation locus—were enriched using CTS. Sequenced reads were processed through a unified workflow comprising basecalling, methylation calling, phasing, structural variant detection, and differential methylation analysis

Table 1 Experimental metadata annotation

Dataset	Experimental group	Sample	Cell line	Targeted sequencing region	DSB induction site
1	CTS experiments targeting H1 imprinted loci (<i>SNRPN</i> , <i>PEG10</i> , <i>KCNQ1OT1</i>) and the NonTarget control	Target-SNRPN	H1	<i>SNRPN</i> , <i>PEG10</i> , and <i>KCNQ1OT1</i>	<i>SNRPN</i>
		Target-PEG10	H1	<i>SNRPN</i> , <i>PEG10</i> , and <i>KCNQ1OT1</i>	<i>PEG10</i>
		Target-KCNQ1OT1	H1	<i>SNRPN</i> , <i>PEG10</i> , and <i>KCNQ1OT1</i>	<i>KCNQ1OT1</i>
		Non-Target	H1	<i>SNRPN</i> , <i>PEG10</i> , and <i>KCNQ1OT1</i>	None
2	CTS experiments targeting H1 heterochromatin loci (<i>UCA1</i> , <i>CRCT1</i>) and the NonTarget control	Target-UCA1	H1	<i>UCA1</i> , and <i>CRCT1</i>	<i>UCA1</i>
		Target-CRCT1	H1	<i>UCA1</i> , and <i>CRCT1</i>	<i>CRCT1</i>
		Non-Target	H1	<i>UCA1</i> , and <i>CRCT1</i>	None
3	CTS experiments targeting RKO epimutation locus (<i>MLH1</i>) and the NonTarget control	Target-MLH1	RKO	<i>MLH1</i>	<i>MLH1</i>
		Non-Target	RKO	<i>MLH1</i>	None
4	CTS experiments targeting the <i>SNRPN</i> locus in early- and late-passage SNRPN-DMR ^{Me/Me} cells	SNRPN-DMR ^{Me/Me} clone, early-passage	H1 clone	<i>SNRPN</i>	<i>SNRPN</i>
		SNRPN-DMR ^{Me/Me} clone, early-passage	H1 clone	<i>SNRPN</i>	<i>SNRPN</i>
		non-target control clone, early passage	H1 clone	<i>SNRPN</i>	<i>SNRPN</i>
		non-target control clone, late passage	H1 clone	<i>SNRPN</i>	<i>SNRPN</i>

In addition, to evaluate the stability of methylation alterations during extended culture, for the *SNRPN* locus, CTS was performed on early- and late-passage derivatives of an SNRPN-DMR^{Me/Me} clonal line.

For each sample, high-molecular-weight genomic DNA was prepared without amplification, enriched for the ROIs through CTS using Cas9 ribonucleoprotein complexes assembled with external sgRNAs, and sequenced using one MinION flow cell (R9.4.1 chemistry) per sample. Basecalling, alignment, and methylation calling were performed against the hg38 reference genome (Fig. 1). Across experiments, CTS consistently achieved high on-target enrichment, yielding approximately 200× to over 1,100× coverage per ROI, which facilitated robust

quantification of per-CpG methylation levels, allelic asymmetry, and DSB-associated epigenetic variation.

Raw sequencing data

All raw aligned sequencing data (BAM format with methylation tags) are deposited in the NCBI Sequence Read Archive under accession PRJNA1193628 [14], which includes multiple runs corresponding to different combinations of: cell line (H1 hESC, RKO), target locus (imprinted loci: *SNRPN*, *PEG10*, *KCNQ1OT1*; heterochromatin loci: *UCA1*, *CRCT1*; epimutation locus: *MLH1*), experimental condition (CRISPR-targeted DSB vs. non-targeting control), and culture stage (early- and late-passage SNRPN-DMR^{Me/Me} cells). Individual sample

Table 2 Overview of datasets

Dataset	Experimental dataset	Raw aligned sequencing file format (file extension)	Data repository and identifier (DOI or accession number)	Methylation metrics file format (file extension)	Data repository and identifier (DOI or accession number)
1	CTS experiments targeting H1 imprinted loci (<i>SNRPN</i> , <i>PEG10</i> , <i>KCNQ1OT1</i>) and the NonTarget control	BAM (.bam)	Target-SNRPN (SRR35906854): http://identifiers.org/insdc.sra:SRR35906854 [19]; Target-PEG10 (SRR35906853): http://identifiers.org/insdc.sra:SRR35906853 [20]; Target-KCNQ1OT1 (SRR35797983): http://identifiers.org/insdc.sra:SRR35797983 [21]; NonTarget control (SRR35906856): http://identifiers.org/insdc.sra:SRR35906856 [22]	bedMethyl (.bed)	Zenodo (DOI: https://doi.org/10.5281/zenodo.18759160) [15]
2	CTS experiments targeting H1 heterochromatin loci (<i>UCA1</i> , <i>CRCT1</i>) and the NonTarget control	BAM (.bam)	Target-UCA1 (SRR35171064): http://identifiers.org/insdc.sra:SRR35171064 [23]; Target-CRCT1 (SRR35171063): http://identifiers.org/insdc.sra:SRR35171063 [24]; NonTarget control (SRR35171061): http://identifiers.org/insdc.sra:SRR35171061 [25]	bedMethyl (.bed)	Zenodo (DOI: https://doi.org/10.5281/zenodo.17762285) [16]
3	CTS experiments targeting RKO epimutation locus (<i>MLH1</i>) and the NonTarget control	BAM (.bam)	Target-MLH1 (SRR35171065): http://identifiers.org/insdc.sra:SRR35171065 [26]; NonTarget control (SRR35171062): http://identifiers.org/insdc.sra:SRR35171062 [27]	bedMethyl (.bed)	Zenodo (DOI: https://doi.org/10.5281/zenodo.17762504) [17]
4	CTS experiments targeting the <i>SNRPN</i> locus in early- and late-passage SNRPN-DMR ^{Me/Me} cells	BAM (.bam)	SNRPN-DMR ^{Me/Me} clone, early passage (SRR35171060): http://identifiers.org/insdc.sra:SRR35171060 [28]; SNRPN-DMR ^{Me/Me} clone, late passage (SRR35171059): http://identifiers.org/insdc.sra:SRR35171059 [29]; non-target control clone, early passage (SRR35171058): http://identifiers.org/insdc.sra:SRR35171058 [30]; non-target control clone, late passage (SRR35171057): http://identifiers.org/insdc.sra:SRR35171057 [31]	bedMethyl (.bed)	Zenodo (DOI: https://doi.org/10.5281/zenodo.17762655) [18]

accessions corresponding to each locus and condition are listed in Table 2. Basecalling was done by Dorado (v0.9.6) (<https://github.com/nanoporetech/dorado>). Alignment was done by minimap2 (v2.17). These data support reanalysis of nucleotide variants, phasing, methylation, and SVs using alternative tools; performing DMR calling under different statistical frameworks; integrating with additional annotations; or visualizing SNVs, SVs, and methylation landscapes with independent visualization tools.

Processed methylation data

To maximize reusability while avoiding duplication, we provide the essential processed datasets required for downstream analysis. The processed data include per-CpG 5mC and 5hmC methylation matrices for each sample, which were generated using Modkit (v0.5.1) on Dorado-basecalled, hg38-aligned BAM files [15–18]. Each resulting matrices file, provided in bedMethyl format, reports per-CpG genomic coordinates, modification type, coverage, raw read counts, and methylation proportions. The output contains 18 columns, defined as follows (see the Modkit documentation for further details: <https://github.com/nanoporetech/modkit>): Columns 1–3 denote genomic coordinates in BED

format (chromosome name, 0-based start coordinate, and 0-based exclusive end position), Column 4 indicates the modified base code (m for 5mC or h for 5hmC), Column 5 reports the score, which is equal to the value in Column 10. Column 6 indicates the reference strand. Columns 7–9 correspond to the start position, end position, and RGB color code for visualization compatibility, respectively. Column 10 reports the valid coverage at the site. Column 11 provides the specified modification percentage. Columns 12–18 provide the raw count breakdown used to derive methylation proportions: Column 12 indicates the number of calls classified as a residue with a specified base modification; Column 13 indicates the number of calls classified as the canonical base rather than modified; Column 14 indicates the number of calls classified as other modification types; Column 15 indicates the number of reads with a deletion at this reference position; Column 16 indicates the number of calls where the probability of the call was below the threshold (by default); Column 17 indicates the number of reads with a base other than the canonical base for this modification; and Column 18 indicates the number of reads aligned to this reference position, with the correct canonical base, but lacking a modification call.

These primary processed data allow users to re-perform differential methylation analyses under alternative statistical frameworks, integrate with additional annotations, or visualize methylation landscapes using customized analytical pipelines.

The analytical outputs are not duplicated in this Data Note since they are already publicly available in the companion research article and can be reproduced from the raw data [4].

Limitations

The datasets include raw aligned Nanopore sequencing reads, while raw signal-level files (POD5) are available upon request. Because signal-level data are not contained within BAM files, benchmarking or re-analysis at the raw electrical signal level cannot be performed using the deposited BAM files alone.

The dataset presented here is subject to several limitations inherent to the experimental design. First, CTS provides targeted rather than genome-wide coverage; therefore, the data are restricted to selected imprinted loci, heterochromatic regions, and a cancer-associated promoter, and are not designed to present comprehensive methylome-wide patterns. Second, although we selected two distinct cell lines (H1 hESCs and RKO colorectal cancer cells) and consistently observed that DSBs perturb local methylation patterns, all experiments were nevertheless performed and validated exclusively in *in vitro* cell line systems. As no *in vivo* assays were conducted, the methylation dynamics and SV patterns reported here may not fully recapitulate the chromatin architecture, DNA repair processes, or epigenetic responses that occur under physiological conditions.

Additionally, this study specifically investigates DNA DSBs induced by CRISPR–Cas9. Other forms of chromosomal breakage—such as those induced by ionizing radiation, or chemical genotoxins—were not assessed. Consequently, the extent to which these findings generalize to other sources of DNA damage remains uncertain. Finally, methylation calling and downstream analyses focused specifically on 5mC and 5hmC in the CpG context using the models available at the time of the study. Other base modifications or non-CpG methylation states were not curated, although such signals may be present in the raw nanopore reads and could be explored with future analytical tools.

Abbreviations

5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine
CRCT1	Cysteine rich C-terminal 1
CTS	Cas9-assisted targeted nanopore sequencing
DSB	Double-strand break
DMR	Differentially methylated region
hESC	Human embryonic stem cell

KCNQ1OT1	KCNQ1 opposite strand/antisense transcript 1
MLH1	MutL homolog 1
ONT	Oxford Nanopore Technologies
PEG10	Paternally expressed 10
ROI	Region of interest
SNRPN	Small nuclear ribonucleoprotein polypeptide N
SV	Structural variant
sgRNA	Single-guide RNA
UCA1	Urothelial cancer associated 1

Acknowledgements

We thank all individuals and facilities recognized in the companion Genome Biology research article [4], whose contributions supported the generation of the original experimental datasets described in this Data Note.

Author contributions

M.L., M.W., and Y.Z. jointly designed the experiments. M.W. carried out the CRISPR–Cas9 editing and nanopore sequencing experiments. Data processing and analysis were performed by M.L., M.W., Y.Z., and C.B. Y.Z. and M.W. led the organization and curation of the datasets and drafted this Data Note under the supervision of M.L. All authors reviewed and approved the final version of the manuscript.

Funding

The research of the Li laboratory was supported by the KAUST Office of Sponsored Research (OSR) under award number BAS/1/1080-01-01. This work was also financially supported in part by funding from King Abdullah University of Science and Technology (KAUST) – KAUST Center of Excellence for Smart Health (KCSH), under award number 5932 (M.L.).

Data availability

The data described in this Database article can be freely and openly accessed on NCBI Sequence Read Archive SRR35906854 [19], SRR35906853 [20], SRR35797983 [21], SRR35906856 [22], SRR35171064 [23], SRR35171063 [24], SRR35171061 [25], SRR35171065 [26], SRR35171062 [27], SRR35171060 [28], SRR35171059 [29], SRR35171058 [30], SRR35171057 [31], under accession PRJNA1193628 [14] and Zenodo under DOIs 10.5281/zenodo.18759160 [15], 10.5281/zenodo.17762285 [16], 10.5281/zenodo.17762504 [17], and 10.5281/zenodo.17762655 [18]. Please see Table 2 and references for details and links to the data. Raw signal-level files (POD5) are available upon request.

Declarations

Ethics approval and consent to participate

The work was performed using established human cell lines and does not involve human participants, animal experiments, or data collected from social media platforms. All research procedures involving human pluripotent stem cells (PSCs) were conducted at King Abdullah University of Science and Technology (KAUST), following review and approval by the KAUST Institutional Biosafety and Bioethics Committee (IBEC). KAUST IBEC, as the local ethics committee registered with the National Committee of Bioethics (NCBE) of Saudi Arabia, fulfills the role of an Embryonic Stem Cell Research Oversight (ESCRO) committee. The H1 human embryonic stem cell (hESC) line used in this study was obtained from the WiCell Research Institute under Material Transfer Agreement 19-W0398. Informed consent was obtained from all original donors by the provider institution at the time of derivation. No human embryos or gametes were collected for this study. The human colorectal cancer cell line RKO was purchased from ATCC and cultured according to the supplier's instructions. As a commercially available and anonymized cell line, its use does not require additional ethical approval or informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Bioscience Program, Biomedical Sciences Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

²KAUST Center of Excellence for Smart Health (KCSH), Thuwal 23955, Saudi Arabia

Received: 30 November 2025 / Accepted: 3 March 2026

Published online: 10 March 2026

References

1. Bi C, Wang L, Yuan B, Zhou X, Li Y, Wang S, Pang Y, Gao X, Huang Y, Li M. Long-read individual-molecule sequencing reveals CRISPR-induced genetic heterogeneity in human ESCs. *Genome Biol.* 2020;21(1):213.
2. Yuan B, Bi C, Tian Y, Wang J, Jin Y, Alsayegh K, Tehseen M, Yi G, Zhou X, Shao Y, et al. Modulation of the microhomology-mediated end joining pathway suppresses large deletions and enhances homology-directed repair following CRISPR-Cas9-induced DNA breaks. *BMC Biol.* 2024;22(1):101.
3. Bi C, Yuan B, Zhang Y, Wang M, Tian Y, Li M. Prevalent integration of genomic repetitive and regulatory elements and donor sequences at CRISPR-Cas9-induced breaks. *Commun Biology.* 2025;8(1):94.
4. Wang M, Zhang Y, Bi C, Li M. CRISPR-Cas9-induced double-strand breaks disrupt maintenance of epigenetic information. *Genome Biol.* 2025.
5. Zhang Y, Bi C, Nadeef S, Maddirevula S, Alqahtani M, Alkuraya FS, Li M. NanoRanger enables rapid single-base-pair resolution of genomic disorders. *Med.* 2024;5(10):1307–e13251303.
6. Takahashi Y, Wu J, Suzuki K, Martinez-Redondo P, Li M, Liao H-K, Wu M-Z, Hernández-Benítez R, Hishida T, Shokhirev MN, et al. Integration of CpG-free DNA induces de novo methylation of CpG islands in pluripotent stem cells. *Science.* 2017;356(6337):503–8.
7. Takahashi Y, Morales Valencia M, Yu Y, Ouchi Y, Takahashi K, Shokhirev MN, Lande K, Williams AE, Fresia C, Kurita M, et al. Transgenerational inheritance of acquired epigenetic signatures at CpG islands in mice. *Cell.* 2023;186(4):715–e731719.
8. Laird PW. Principles and challenges of genome-wide DNA methylation analysis. *Nat Rev Genet.* 2010;11(3):191–203.
9. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, Bartlett JJ, Pavlenok M, Gundlach JH. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc Natl Acad Sci.* 2013;110(47):18904–9.
10. Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. Detecting DNA cytosine methylation using nanopore sequencing. *Nat Methods.* 2017;14(4):407–10.
11. Wescoe ZL, Schreiber J, Akeson M. Nanopores Discriminate among Five C5-Cytosine Variants in DNA. *J Am Chem Soc.* 2014;136(47):16582–7.
12. Schreiber J, Wescoe ZL, Abu-Shumays R, Vivian JT, Baatar B, Karplus K, Akeson M. Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc Natl Acad Sci U S A.* 2013;110(47):18910–5.
13. Gigante S, Gouil Q, Lucattini A, Keniry A, Beck T, Tinning M, Gordon L, Woodruff C, Speed TP, Blewitt ME, et al. Using long-read sequencing to detect imprinted DNA methylation. *Nucleic Acids Res.* 2019;47(8):e46–46.
14. NCBI Sequence Read Archive <https://identifiers.org/ncbi/insdc.sra:PRJNA1193628> (2024).
15. Per-CpG 5mC and 5hmC methylation matrices for H1 imprinted loci. Zenodo. <https://doi.org/10.5281/zenodo.18759160>
16. Per-CpG 5mC and 5hmC methylation matrices for H1 heterochromatin loci. Zenodo. <https://doi.org/10.5281/zenodo.17762285>
17. Per-CpG 5mC and 5hmC methylation matrices for RKO epimutation locus. Zenodo. <https://doi.org/10.5281/zenodo.17762504>
18. Per-CpG 5mC and 5hmC methylation matrices for early- and late-passage SNRPN-DMR^{Me/Me} cells. Zenodo. <https://doi.org/10.5281/zenodo.17762655>
19. NCBI Sequence Read Archive. Accession SRR35906854. <http://identifiers.org/insdc.sra:SRR35906854> (2025). NCBI Sequence Read Archive (SRA) dataset.
20. NCBI Sequence Read Archive. Accession SRR35906853. <http://identifiers.org/insdc.sra:SRR35906853> (2026). NCBI Sequence Read Archive (SRA) dataset.
21. NCBI Sequence Read Archive. Accession SRR35797983. <http://identifiers.org/insdc.sra:SRR35797983> (2025). NCBI Sequence Read Archive (SRA) dataset.
22. NCBI Sequence Read Archive. Accession SRR35906856. <http://identifiers.org/insdc.sra:SRR35906856> (2026). NCBI Sequence Read Archive (SRA) dataset.
23. NCBI Sequence Read Archive. Accession SRR35171064. <http://identifiers.org/insdc.sra:SRR35171064> (2025). NCBI Sequence Read Archive (SRA) dataset.
24. NCBI Sequence Read Archive. Accession SRR35171063. <http://identifiers.org/insdc.sra:SRR35171063> (2025). NCBI Sequence Read Archive (SRA) dataset.
25. NCBI Sequence Read Archive. Accession SRR35171061. <http://identifiers.org/insdc.sra:SRR35171061> (2025). NCBI Sequence Read Archive (SRA) dataset.
26. NCBI Sequence Read Archive. Accession SRR35171065. <http://identifiers.org/insdc.sra:SRR35171065> (2025). NCBI Sequence Read Archive (SRA) dataset.
27. NCBI Sequence Read Archive. Accession SRR35171062. <http://identifiers.org/insdc.sra:SRR35171062> (2025). NCBI Sequence Read Archive (SRA) dataset.
28. NCBI Sequence Read Archive. Accession SRR35171060. <http://identifiers.org/insdc.sra:SRR35171060> (2025). NCBI Sequence Read Archive (SRA) dataset.
29. NCBI Sequence Read Archive. Accession SRR35171059. <http://identifiers.org/insdc.sra:SRR35171059> (2025). NCBI Sequence Read Archive (SRA) dataset.
30. NCBI Sequence Read Archive. Accession SRR35171058. <http://identifiers.org/insdc.sra:SRR35171058> (2025). NCBI Sequence Read Archive (SRA) dataset.
31. NCBI Sequence Read Archive. Accession SRR35171057. <http://identifiers.org/insdc.sra:SRR35171057> (2025). NCBI Sequence Read Archive (SRA) dataset.

Publisher's Note

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