PacBio

Application note

Fiber-seq: High-resolution long-read chromatin fiber sequencing in a single multiomic assay

Introduction

A foundational tenet of molecular biology holds that the dynamics of chromatin architecture can exert long-range regulatory effects on gene expression. Most methods that investigate these dynamics are based on short-read sequencing assays that fragment chromatin into smaller pieces. This approach limits the ability to observe these dynamics in the wider genomic context and characterize complex interactions between distal regulatory elements. In addition, most short-read methods do not provide single-molecule level information that masks important heterogeneity and requires separate, specialized assays from those that measure genetic variation.

Fiber-seq is a powerful new long-read whole genome sequencing assay that overcomes the challenges imposed by short-read chromatin sequencing. Fiber-seq simultaneously maps genetic variation, accessible regions, nucleosome position, CpG methylation, and bound transcription factors onto individual chromatin fibers. This rich multiomic data at single-molecule and near base pair resolution would otherwise require 3 or more separate short-read assays (e.g., WGS, WGBS, ATAC-seq). However, Fiber-seq provides this information in a single assay that reveals haplotype-specific differences in gene regulation and individual chromatin fiber dynamics.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.
102-326-654 REV01 SEP2025

PacBio

Fiber-seq highlights with PacBio® HiFi sequencing

- ATAC-seq-like chromatin accessibility, nucleosome positioning, CpG methylation, and transcription factor occupancy at single-molecule resolution
- No additional sequencing beyond standard long-read WGS
- Fiber-seq reagents and optimized protocol available now from EpiCypher®
- Revio® and Vega™ on-instrument **6mA** and **5mC** calling for simple and streamlined analysis
- Community software like fibertools to QC data and infer regulatory elements

What is Fiber-seq?

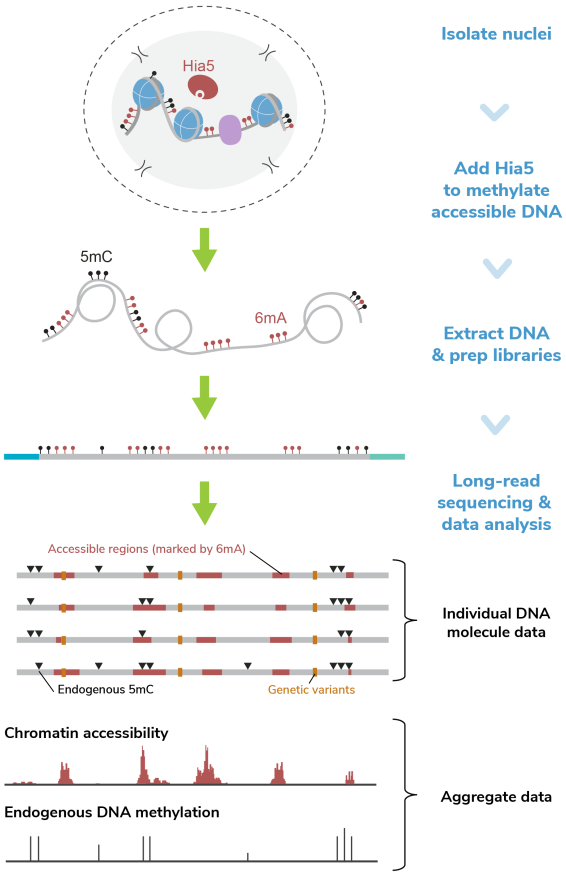

Figure 1. Fiber-seq workflow using EpiCypher CUTANA Hia5. Once 6mA labeled DNA is extracted from nuclei, users prepare libraries following the standard PacBio WGS protocol for HiFi sequencing. Every read represents an individual chromatin fiber with the endogenous 5mC and Hia5 labeled 6mA marks. Data can be aggregated to generate methylation and accessibility maps. Image courtesy of EpiCypher.

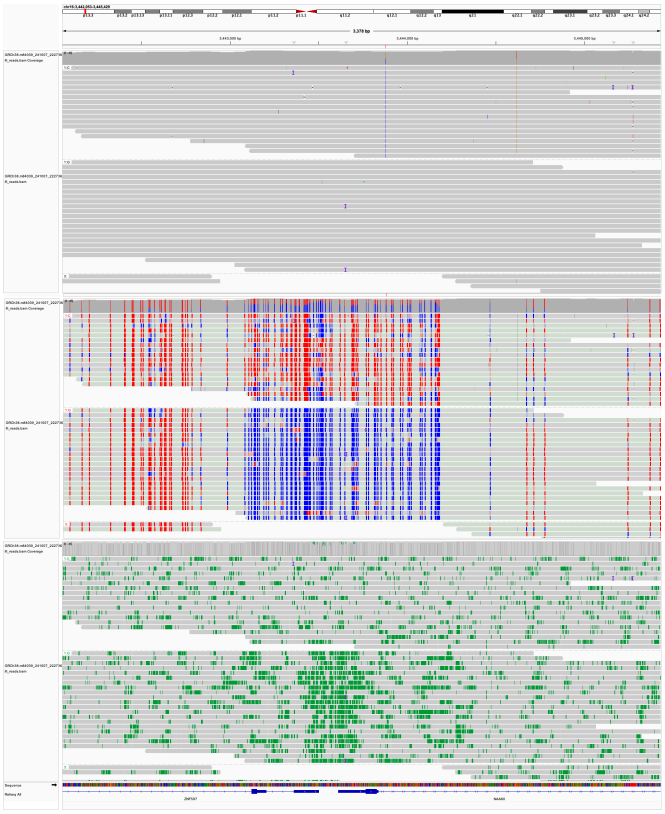

Figure 2. Example of a known imprinted genomic region from Fiber-seq HG002 reads mapped to the GRCh38 reference at Chr16:3,442,053 – 3,445,429 and displayed in IGV. Top panel shows reads sorted by haplotype. The middle panel displays the 5mC calls from the same reads where red indicates a methylated CpG, and blue un-methylated. The top haplotype is methylated at the bi-directional promoters for ZNF597 and NAA60, while the bottom haplotype is un-methylated. Correspondingly, the Hia5 6mA labeling (green; bottom panel) shows that this promoter region is accessible on the bottom haplotype and not the top. A recent study using both Fiber-seq and Kinnex full-length RNA sequencing on GM12878 similarly showed that the top haplotype is inaccessible and results in no transcripts starting at this imprinted promoter (Vollger et al. 2025a).

Fiber-seq is a single-molecule approach that enables nucleotide-resolution readout of chromatin features over multi-kilobase DNA fibers (Stergachis et al. 2020). This method first isolates nuclei from cells, which are then treated with sequence non-specific N^6- adenine methyltransferase, Hia5 (Drozdz et al. 2012). This treatment marks adenines (As) within accessible (unprotected) DNA regions. This 6mA mark is not naturally present in most eukaryotic genomes and therefore is easily distinguished from other natural modifications like 5mC (Kong et al. 2022). Once the nuclei have been treated, DNA is extracted and

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

prepared using PacBio standard whole genome sequencing (WGS) SMRTbell® library preparation. Sequencing is performed on Revio or Vega systems, which detects 6mA methylation along with naturally occurring 5mC base modifications. An example of the different modes of multiomic data is shown in Figure 2 at a known imprinted locus. Fiber-seq not only captures genetic variants for long-range haplotype phasing, but also provides methylation and chromatin accessibility data, simultaneously offering epigenetic insights and a haplotype-resolved view of gene regulation.

Fiber-seq enables novel insights into chromatin biology, epigenomics, and disease mechanisms

The landmark Fiber-seq study from Stergachis et al. (2020) showed it was possible to sequence multi-kilobase individual chromatin fibers using PacBio HiFi reads rather than using bulk averages at discrete loci with short-reads. This innovative approach revealed previously hidden chromatin biology by showing that neighboring regulatory elements can act in an all-or-none fashion, with fibers displaying either fully accessible or fully closed states. In contrast, short-read methods average across molecules, blurring such binary behavior into intermediate levels of accessibility. With Fiber-seq using HiFi sequencing, co-actuation of cis elements can be observed because the open or closed state is "stenciled" onto every read/fiber spanning these regions (Stergachis et al. 2020).

Since this study, Fiber-seq has been used for a variety of applications and employed on many samples from the Human Pangenome Reference Consortium. Some examples of Fiber-seq applications with human samples include:

- **Revealing the chromatin landscape of centromeres (Dubocanin et al. 2025) and "dark" regions of the genome (Peter et al. 2024).** These regions are difficult or virtually impossible to investigate with short reads because of repetitive or homologous sequences, yet they contain some of the most variable regions in the genome (Logsdon et al. 2025).
- **Better understanding disease mechanisms of Mendelian conditions.** For example, Fiber-seq was applied to an undiagnosed network (UDN) participant with a ChrX,13-balanced translocation. The multiomic nature of Fiber-seq data not only helped confirm that the translocation disrupted several genes but also helped disentangle the underlying mechanisms, like the X-chromosome inactivation of autosomal DNA, that explained the disparate phenotypes (Vollger et al. 2025a).
- **Mapping a diploid genome to provide a haplotype-resolved view of human gene regulation.** Vollger et al. (2025b) found that Fiber-seq provides more accurate measures of chromatin accessibility than existing short-read assays like ATAC-seq, especially for elements <200 bp in length. Additionally, this study provided a more complete genome-wide map by including chromatin accessibility in duplicated regions. By leveraging highly accurate HiFi long reads to phase the genome, the authors showed that there are widespread haplotype-specific differences in accessibility largely determined by the underlying sequence variation.
- **Examining chromatinization behavior of transfected plasmid sequences.** Plasmid Fiber-seq has the potential to evaluate the impact of non-coding regulatory variants and improve the design of gene therapy vectors (Mallory et al. 2025).

Moreover, Fiber-seq applications extend beyond humans. Tullius et al. (2024) utilized Fiber-seq in *Drosophila* to visualize individual RNA polymerases on chromatin fibers. The authors identified individual Pol II complexes and demonstrated pause-driven nucleosome destabilization and distance-dependent coupling between transcription events. Fiber-seq has also been used in plants to reveal the regulatory roles of repeats (e.g., LTR retrotransposons) in maize where short reads fail to map but that make up ~80% of the genome (Bubb et al. 2025). These and further studies suggest that Fiber-seq can be a valuable tool for studying epigenetics in many eukaryotic organisms without endogenous adenine methylation aside from humans.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

6mA labeling with EpiCypher CUTANA Hia5 for Fiber-seq

PacBio is partnering with EpiCypher to make Fiber-seq more widely accessible with the commercialization of the CUTANA™ Hia5 enzyme. PacBio independently verified that the Hia5 enzyme and protocol on human cell lines are compatible with HiFi sequencing on Revio and Vega systems with a simple and straight-forward workflow. HiFi sequencing results demonstrated that CUTANA Fiber-seq effectively captures the architecture of single chromatin fibers, enabling the detection of accessible regions, nucleosome position, and regulatory elements (see Figure 2 as an example).

Materials and methods

Metric	Specifications
Protocol	EpiCypher CUTANA Fiber-seq Protocol (v1.0)
Samples	CHM13 cell line
Number of cells	1.8 M
Hia5 concentration range	0.25–8.0X
DNA extraction	Extracting HMW DNA from cultured adherent cells using Nanobind® kits
DNA input	3.5 µg
PacBio library prep protoocol	Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0
Sequencing	Revio with SPRQ™ chemistry

Approximately 1.8 million (1.8 M) CHM13 cells were prepared per treatment and nuclei were isolated following the EpiCypher protocol. Nuclei were then treated with Hia5 enzyme in concentrations ranging from 0.25 to 8X. After Hia5 treatment, genomic DNA was extracted using the PacBio Nanobind PanDNA kit[1] following the protocol for cultured adherent cells. DNA was quality-controlled for size on the Agilent® Femto Pulse® system and prepared using SRE size selection and the PacBio SMRTbell prep kit 3.0. Approximately 3.5 µg of gDNA per sample was used to ensure extra libraries for additional sequencing if necessary, though as little as 500 ng of DNA into library prep is needed for sequencing on the Revio system. After library prep, samples bound with the SPRQ sequencing polymerase were loaded onto SMRT® Cells using a standard 250 pM concentration. Then 5mC and 6mA modifications were called on-instrument using the Revio v13.3 instrument software. For secondary analysis, the fibertools package was employed to QC the datasets and annotate putative nucleosome lengths and positions.

[1] EpiCypher Fiber-seq protocol recommends NEB Monarch Spin gDNA kit. Both kits (Monarch and Nanobind) are compatible with PacBio HiFi sequencing.

Results

All CHM13 Fiber-seq samples yielded between 121 to 148 Gb of HiFi data (~40–49X coverage of the haploid genome) with over 93% of bases at a predicted Q-score of Q30 or higher (Figure 3).

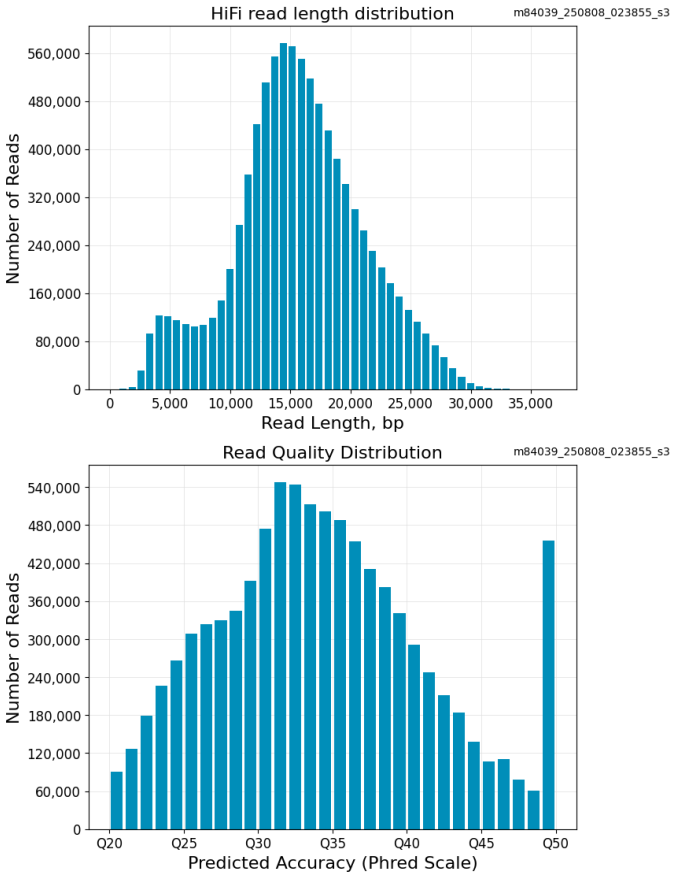

Figure 3. Revio HiFi read length and quality distributions for CHM13-FS03 samples treated with the 2x concentration of Hia5 enzyme (movie id: m84039_250808_023855). High sequencing yields, read length, and data quality demonstrate that Fiber-seq is fully compatible HiFi sequencing.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

Fiber-seq sample	Hia5 Conc.	HiFi Yield (Gb)	HiFi read length N50	Q30+ bases	5mC*	6mA
CHM13-FS01	0.25X	143.8	17,765 bp	93.9%	39.8%	2.3%
CHM13-FS02	1.0X	121.8	17,685 bp	93.9%	39.7%	4.1%
CHM13-FS03	2.0X	142.9	17,155 bp	93.5%	39.2%	5.3%
CHM13-FS04	4.0X	128.2	17,466 bp	93.5%	38.5%	6.4%
CHM13-FS05	8.0X	138.2	17,126 bp	93.3%	38.9%	7.8%

Table 1. Sequencing results of Fiber-seq libraries prepared using EpiCypher CUTANA Hia5 enzyme in the PacBio applications lab, Menlo Park, CA and sequenced on the Revio system with SPRQ chemistry. Note users results may vary depending on the average library insert size. SRE was used on the gDNA prior to SMRTbell library preparation to enrich for large DNA fragments. *Sites where the predicted likelihood is >0.5.

Predicted methylation levels (5mC) for CHM13 were in the expected range (38-39%; Table 1) and consistent with previous reports on the T2T-CHM13 genome being hypomethylated at CpG relative to more differentiated cell lines like HG002 (Gershman et al. 2022). The level of 6mA labeling ranged from 2.3% (0.25X) to 7.8% (8.0X) with increasing methylation at higher concentrations of Hia5, as expected (Table 1 and Figure 4).

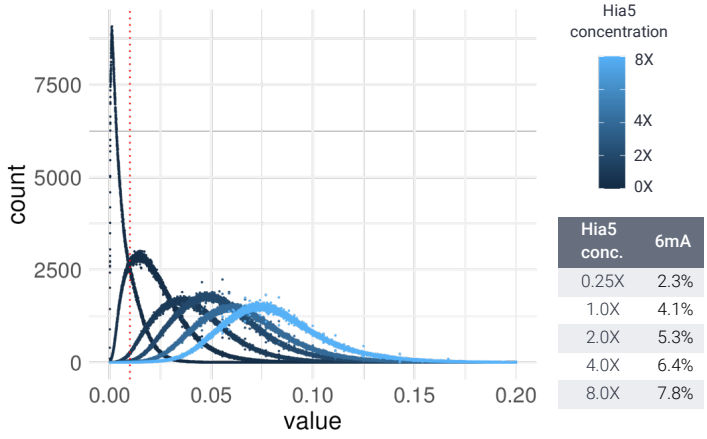

Figure 4. Distribution of reads by %6mA labeling. Lighter shades of blue indicate higher, while darker shades lower concentrations of Hia5 enzyme. Red dashed line approximates the FDR of a sample not treated with Hia5.

Previously-run null controls showed that the false discovery rate (FDR) of 6mA methylation is less than 1% when samples were run through the protocol without Hia5 treatment. For example see the PacBio jasmine GitHub for discussion on 6mA calling accuracy.

The autocorrelation QC of the CHM13 titration data indicates these experiments successfully labeled accessible regions within intact chromatin fibers (Figure 5), consistent with nucleosome length distributions called using the ft add-nucleosomes command with default parameters (Figure 6).

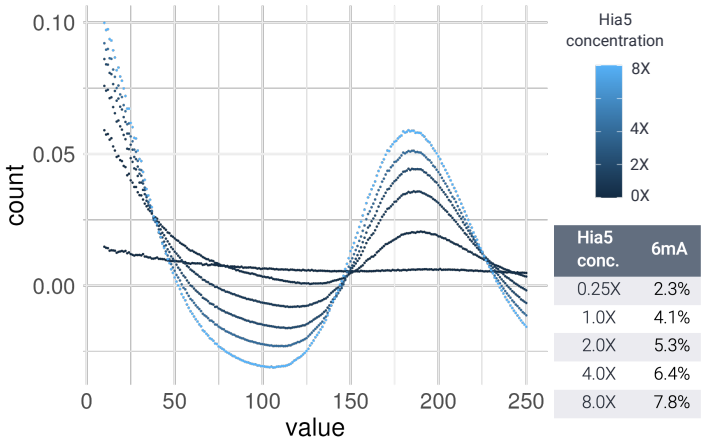

Figure 5. Auto-correlation traces of 6mA signal along single chromatinized DNA molecules reveal the periodic pattern of genome accessibility punctuated by nucleosomes. The gradient from dark to light blue represents increasing concentration of Hia5 during the methylation step. Deeper troughs and higher peaks in the plot are representative of well-ordered nucleosomes.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

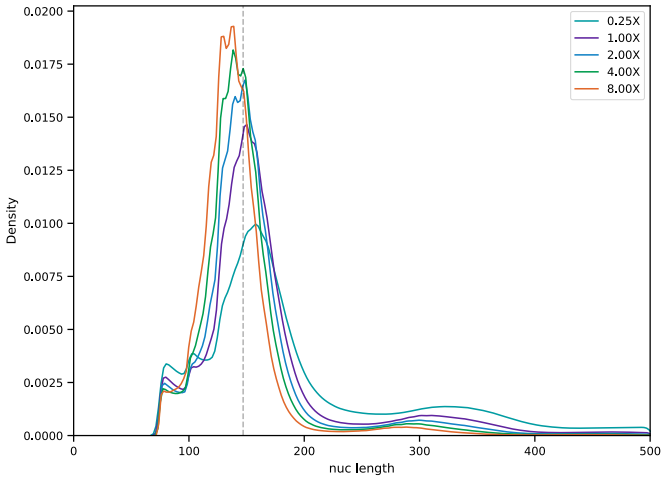

Figure 6. Distribution of Fiber-seq inferred nucleosome lengths from CHM13, chromosome 1, across different Hia5 enzyme concentrations on approximately 1.8 M nuclei. Hia5 0.25X concentration is shown in light blue, 1.0X in purple, 2.0X in dark blue, 4.0X in green, and 8.0X in orange. The vertical dashed line is at 147 bp, the expected average nucleosome length in humans.

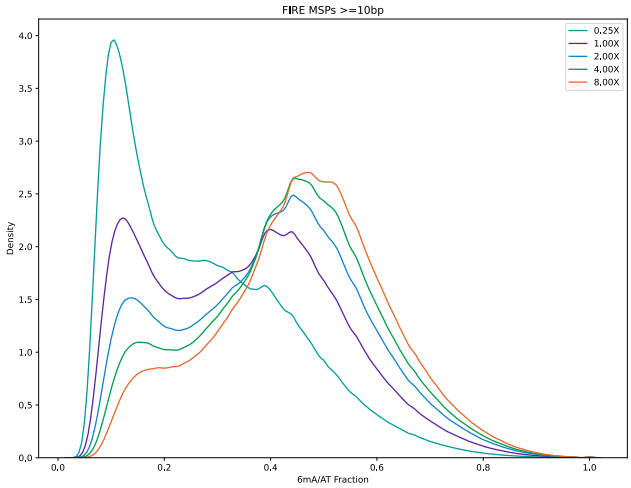

Figure 7. The percent of 6mA modifications per total number of A/T pairs within all putative FIRE regions greater than 10 bp across Hia5 concentrations. Hia5 0.25X concentration is shown in light blue, 1.0X in purple, 2.0X in dark blue, 4.0X in green, and 8.0X in orange. As Hia5 concentration is increased, a higher fraction of the available adenine bases within a putative accessible region are modified (6mA) as expected.

Generally, both plots (Figure 5 and 6) should reflect the known nucleosome length in humans at 147 bp. Experiments with little 6mA labeling will show flatter autocorrelation curves and larger nucleosome length distributions—like the 0.25X and 1.0X treatments in Figure 5 and 6. In contrast, oversaturation will cause shorter nucleosome lengths as 6mA labeling encroaches into the nucleosomes. This is evident in the nucleosome length distributions where the mode shifts to smaller lengths (to the left) as the Hia5 concentration increases. For example, the mode of the 8.0X Hia5 treatment peaks to the left (shorter), while the mode for the 1.0 and 2.0X treatments is at or near the 147 bp length (Figure 6). These metrics, in addition to the total fraction of 6mA, can be used to assess the success and quality of a Fiber-seq experiment.

Fibertools was used to classify putative methylation-sensitive patches (MSPs) and Fiber-seq Inferred Regulatory Elements (FIREs) from each read (Figure 7). These are inferred regions of open chromatin, distinguished from internucleosomal linker regions, and enriched for DNaseI hypersentitivity sites (see Vollger et al. 2025b). Looking at the fraction of 6mA modifications within a FIRE can be used as a metric to indicate the ability to resolve transcription factor footprints.

Here, the mean fraction of 6mA modifications per available adenine base was positively correlated with Hia5 concentrations as expected and suggests that the resolving power to footprint increases with higher fractions of 6mA labeling. For example, the mean rate of 6mA per A/T ranged from 27% (0.25X) to 46% (8.0X) within classified FIREs. Moving from a 4X to 8X concentration provided only a 3% gain in 6mA /AT suggesting minimal benefit in increasing Hia5 concentration further (Table 2 and Figure 7).

Hia5 concentration	%6mA within FIRE (mean)
0.25X	27%
1.0X	35%
2.0X	40%
4.0X	43%
8.0X	46%

Table 2. The mean percent 6mA labeling per 'A/T' within classified FIRE regions across all reads from each Hia5 titration point.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

Discussion

Results on Revio with SPRQ chemistry show that the Fiber-seq protocol using the EpiCypher CUTANA Hia5 enzyme is consistent with standard HiFi sequencing performance. Fiber-seq does not damage DNA or interfere with sequencing. Sequencing on the Vega system produced similarly high-quality results (not shown), confirming that both systems and chemistries are compatible with the protocol. Revio is likely the preferred platform for most Fiber-seq applications because it provides higher sequence coverage on the human genome. Vega may be better suited for applications requiring lower coverage, smaller eukaryotic genomes, or targeted studies (e.g., Bohaczuk et al., 2024)[2]. These results also demonstrate that Fiber-seq is compatible with PacBio Nanobind DNA extraction when substituted for the kit recommended in the EpiCypher protocol.

A critical factor in Fiber-seq experiments is ensuring appropriate 6mA labeling for downstream analysis. Similar to short-read chromatin accessibility profiling, Fiber-seq is sensitive to the ratio of nuclei to enzyme. Optimal 6mA labeling in human samples falls between 5–7%, which captures open chromatin, linker DNA, and nucleosome footprints of ~147 bp (Figure 6). EpiCypher reports that 1 M K562 nuclei with 1X Hia5 yield ~6% labeling and the expected footprint size. Here, a slightly lower rate (4.1%) was observed with 1X Hia5, likely due to the higher input of CHM13 nuclei (~1.8 M). Increasing the Hia5 concentration (2–4X) corrected this, achieving optimal labeling of 5–7% (Table 1) and expected nucleosome size (Figure 6). Furthermore, the frequency of 6mA modifications per adenine within FIRE segments indicates that this labeling range also provides strong resolution for transcription factor footprinting (40–46% 6mA per A/T; Figure 7). The fibertools FIRE pipeline was also trained to account for variation in labeling efficiency, increasing workflow robustness (Vollger et al., 2025b). Nonetheless, shallow sequencing experiments may still be advisable to confirm labeling efficiency before large-scale studies.

Conclusions

Fiber-seq enables the integration of multiple datatypes—genome, methylome, and chromatin accessibility—on the same read and molecule, making it one of the most powerful multiomic sequencing assays available. Unlike analogous short-read assays (e.g., WGS, WGBS, ATAC-seq), Fiber-seq provides a synchronous view across multiple cis-regulatory elements within individual chromatin fibers, yielding insights into chromatin biology and epigenetics that were previously inaccessible. In addition, its ability to phase reads by haplotype and footprint nucleosomes and transcription factors establishes Fiber-seq using PacBio technology as a new gold standard in epigenetics research.

Fiber-seq resources

Publicly available datasets

- PacBio, HG002 Fiber-seq
- PacBio, CHM13 Fiber-seq (EpiCypher CUTANA Hia5)
- EpiCypher, K562 Fiber-seq (EpiCypher CUTANA Hia5)
- HPRC Fiber-seq datasets using PacBio

Protocols

- EpiCypher CUTANA Fiber-seq Protocol
- PacBio WGS protocol with SMRTbell prep kit 3.0

Bioinformatics

- DNA modification detection accuracy using Jasmine
- Fibertools
- Independent cross-technology performance comparison

[2] See https://fiberseq.github.io/analyses/ont.html#overall-conclusions. Both PacBio systems provides more accurate base calling than other long-read technologies.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

References

Bubb, K.L., et al. (2025). The regulatory potential of transposable elements in maize. *Nature Plants*, 11, 1181-1192.

Bohaczuk, S.C., et al. (2024). Resolving the chromatin impact of mosaic variants with targeted Fiber-seq. *Genome Research*, 34(12), 2269-2278.

Dubocanin, D., et al. (2025). Conservation of dichromatin organization along regional centromeres. *Cell Genomics*, 5(4), 100819.

Drozdz, M., Drozdz, M., Piekarowicz, A., Bujnicki, J. M., & Radlinska, M. (2012) Novel non-specific DNA adenine methyltransferases. *Nucleic Acids Research*, 40(5), 2119-2130.

Gershman, A., et al. (2022). Epigenetic patterns in a complete human genome. *Science*, 376, eabj5089.

Kong, Y. et al. (2022). Critical assessment of DNA adenine methylation in eukaryotes using quantitative deconvolution. *Science*, 375, 515-522.

Logsdon, G.A., et al. (2025). Complex genetic variation in nearly complete human genomes. *Nature*, 644, 430-441.

Mallory, B.J. et al. (2025). Principles and functional consequences of plasmid chromatinization in mammalian cells. *bioRxiv*, 2025-05.

Peter, C. J., et al. (2024). Single chromatin fiber profiling and nucleosome position mapping in the human brain. *Cell Reports Methods*, 4(12), 100911.

Stergachis, A.B., et al. (2020). Single-molecule regulatory architectures captured by chromatin fiber sequencing. Science, 368, 1449-1454.

Tullius, T. W., et al. (2024). RNA polymerases reshape chromatin architecture and couple transcription on individual fibers. *Molecular Cell*, 84(17), 3209.

Vollger, M.R., et al. (2025a). Synchronized long-read genome, methylome, epigenome and transcriptome profiling resolve a Mendelian condition. *Nature Genetics*, 57, 469-479.

Vollger, M.R., et al. (2025b). A haplotype-resolved view of human gene regulation. bioRxiv, 2024-06.

Research use only. Not for use in diagnostic procedures. © 2025 Pacific Biosciences of California, Inc. (“PacBio”). All rights reserved. Information in this document is subject to change without notice. PacBio assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of PacBio products and/or third-party products. Refer to the applicable PacBio terms and conditions of sale and to the applicable license terms at pacb.com/license. Pacific Biosciences, the PacBio logo, PacBio, Circulomics, Omniome, SMRT, SMRTbell, Iso-Seq, Sequel, Nanobind, SBB, Revio, Onso, Apton, Kinnex, PureTarget, SPRQ, and Vega are trademarks of PacBio.

© 2025 PacBio. All rights reserved. Research use only. Not for use in diagnostic procedures.

102-326-654 REV01 SEP2025

PacBio