Multi-epigenomic mapping with long read sequencing

James Anderson[1], Bryan J. Venters[1], Paul W. Hook[2], Jamie Moore[2], Vishnu U. Sunitha Kumary[1], Allison Hickman[1], Sabrina Hunt[1], Connor Fraiser[1], Anup Vaidya[1], Ryan Ezell[1], Jonathan M. Burg[1], Zu-Wen Sun[1], Martis W. Cowles[1], Winston Timp[2] & Michael-Christopher Keogh[1]

[1] EpiCypher Inc., Research Triangle Park, Durham NC 27709, USA.
[2] Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA.

EpiCypher®

Long read sequencing (LRS) to interrogate the relationships between chromatin features

Gene transcription is regulated by the complex interplay between histone post-translational modifications (PTMs), chromatin associated proteins (CAPs), and DNA methylation (DNAme). Mapping their genomic locations and examining the relationships between these chromatin elements is a powerful approach to decipher mechanisms of disease, thereby enabling discovery of novel biomarkers and therapeutics. Leading epigenomic mapping technologies (e.g., ChIP-seq, CUT&RUN) rely upon DNA fragmentation to isolate regions of interest for sequencing on short read platforms (e.g., Illumina). This strategy leads to substantial loss of contextual information regarding the surrounding DNA, precluding the identification of multiple co-occurring epigenomic features on a single DNA molecule. By contrast, long-read sequencing (LRS) platforms are capable of sequencing very long reads from a single molecule (typically >10kb), allowing relationships between features on a single molecule to be used to resolve heterogeneity within mixed populations. Here we report a robust multi-omic method that leverages LRS to simultaneously profile histone PTMs (or CAPs), DNAme, and parental haplotype in a single assay.

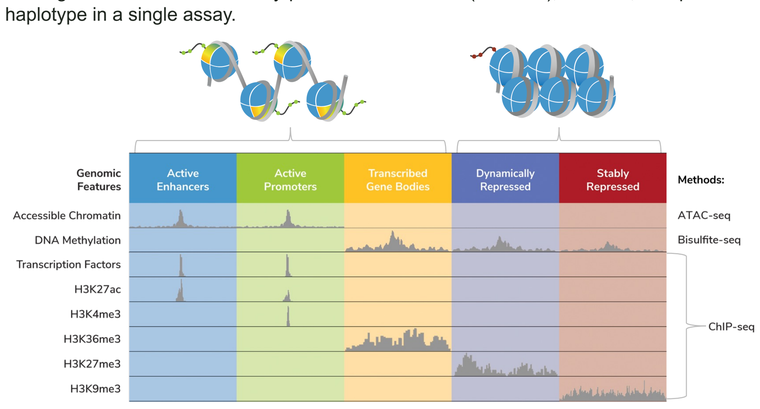
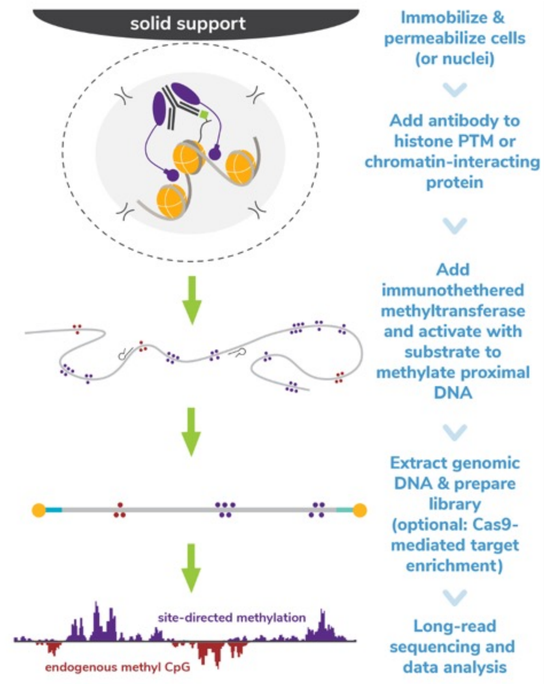


Figure 1: Epigenomic mapping demarks distinct chromatin regulatory features.

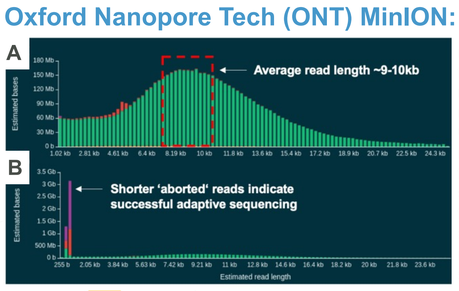
CUTANA-LRS Workflow: optimized immuno-tethering of 6mA writer



Box 1: **CUTANA-LRS.** This nondestructive, epigenomic mapping approach leverages a novel DNA methyltransferase fusion protein (pAG-M.EcoGII) to label DNA underneath antibody-targeted chromatin features, thereby marking sites of interest while preserving DNA molecules intact for LRS.

Inspired by our work with state-of-the-art immunotethering-based approaches (CUT&RUN / CUT&Tag), nuclei are bound to magnetic beads to streamline and automate sample processing. Next, adenosines nearby antibody-targeted chromatin features are methylated with pAG-M.EcoGII, which are then directly read from genomic DNA using Oxford Nanopore Technologies or Pacific Biosciences LRS platforms.

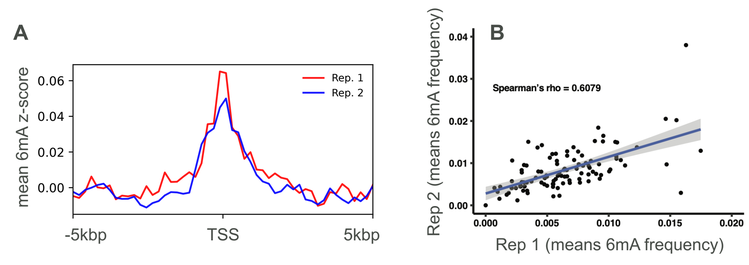
Oxford Nanopore Tech (ONT) MinION: read length analysis



(A) Adaptive sequence reads (selected from chromosome 22) have an modal read length of ~9kb.
(B) Adaptive sequencing allows the ONT sequencer to abort reads that are not from chromosome 22 (thus focusing read-depth to this chromosome). Read length analysis from a representative CUTANA-LRS assay showing reads <1kb comprise many short (aborted) reads, indicating successful selection.

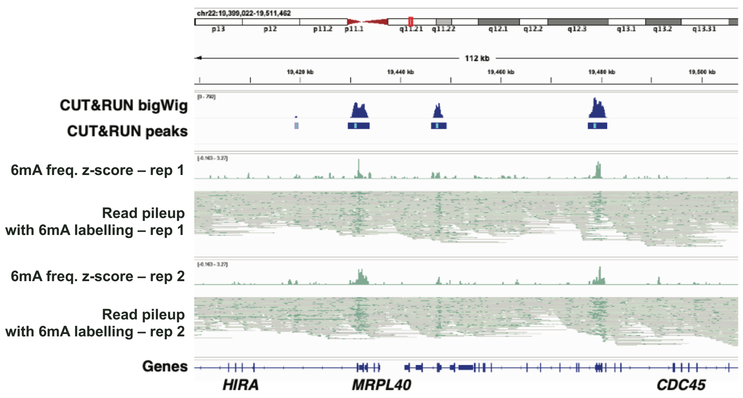
H3K4me3 LRS mapping via 6mA in situ labelling

High signal enrichment and reproducibility for H3K4me3 on chromosome 22 (ONT MinION adaptive sequencing)



(A) Plot showing mean CUTANA-LRS 6mA signal for H3K4me3 from each replicate at TSS from K562 chromosome 22.
(B) Correlation plot of replicates for H3K4me3 6mA signal.

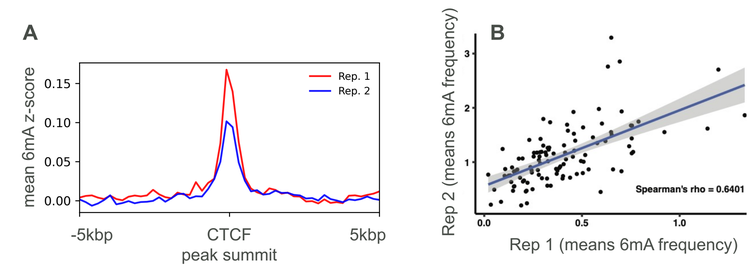
H3K4me3 CUTANA-LRS mapping in K562 cells highly concordant with CUT&RUN SRS



Integrative Genomics Viewer (IGV) showing a 112kb view of chromosome 22. H3K4me3 K562 CUT&RUN signal tracks (blue) and peaks are displayed for benchmarking. Total CUTANA-LRS reads are displayed for each replicate (read pileup), with a composite view (6mA freq. z-score) shown to view labeling enrichment (green).

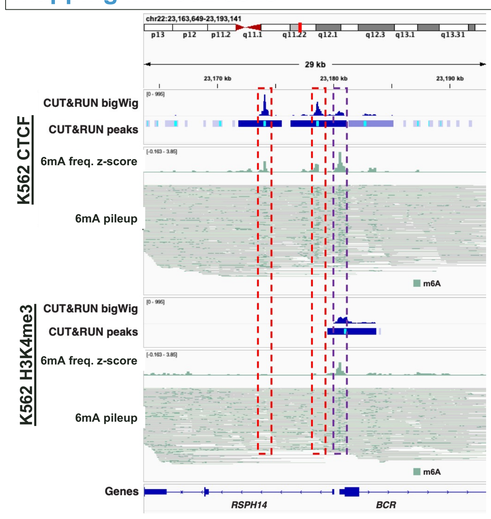
CTCF transcription factor LRS mapping via 6mA in situ labelling

High signal enrichment and reproducibility for CTCF on chromosome 22 (ONT MinION adaptive sequencing)



(A) Plot showing median CUTANA-LRS 6mA signal for CTCF from each replicate at CTCF peak midpoints on chr22 (CUT&RUN) in K562 cells.
(B) Correlation plot of replicates for CTCF CUTANA-LRS 6mA signal.

Distinct patterns of CUTANA-LRS CTCF and H3K4me3 mapping in K562 cells



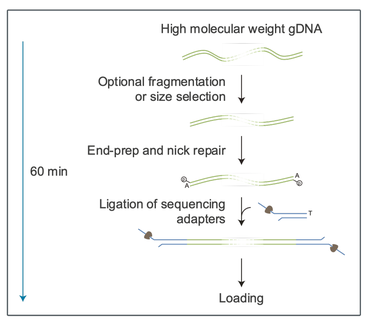
IGV display comparing CUTANA-LRS assay for CTCF and H3K4me3. Purple box = shared peak; red boxes = CTCF-specific peaks. K562 CTCF CUT&RUN tracks and peaks are shown for benchmarking.

Top set of tracks: CTCF CUTANA-LRS

Lower set of tracks: H3K4me3 CUTANA-LRS.

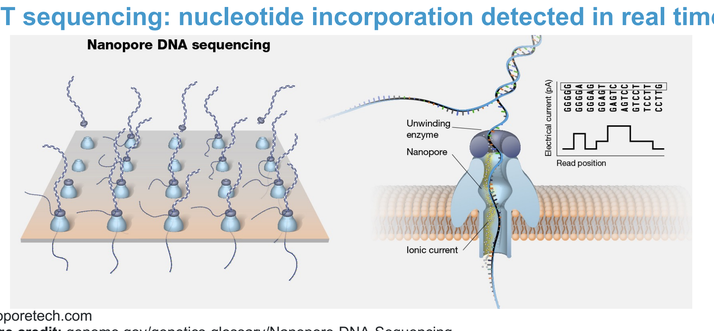
Overview of Oxford Nanopore Technologies

PCR-free ONT library prep: ligation of adapters with helicase



Strengths
- ✓ Real time analysis enables adaptive sequencing
- ✓ Scalable from pocket to benchtop sequencers enables sequencing anywhere
- ✓ Single molecule resolution
- ✓ No PCR enables direct detection of base modifications (eg: 5mC, 6mA)
- ✓ Multi-omics capable

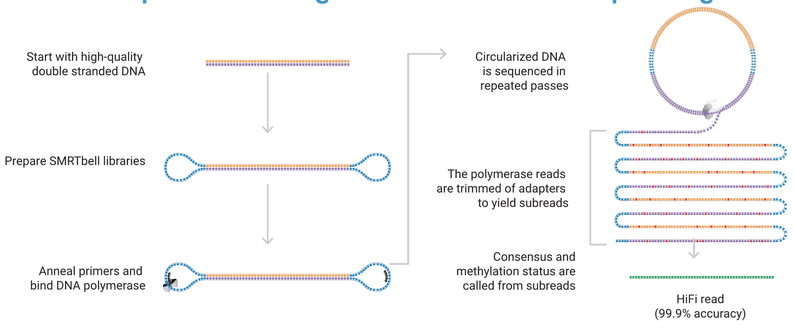
ONT sequencing: nucleotide incorporation detected in real time



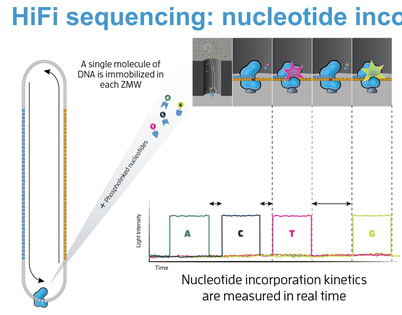
nanoporetech.com
Image credit: genome.gov/genetics-glossary/Nanopore-DNA-Sequencing
Original publication: Howorka and Bayley. (2001) Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnology*.

Overview of Pacific Bio: HiFi LRS

HiFi reads produced using circular consensus sequencing



HiFi sequencing: nucleotide incorporation detected in real time

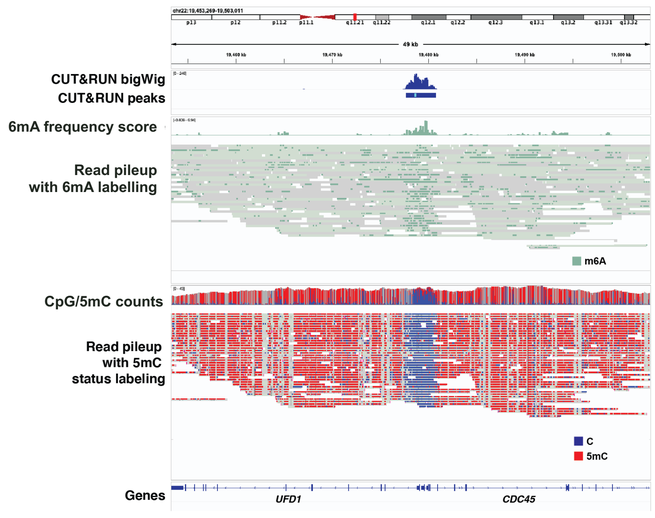


Strengths
- ✓ Long, high-accuracy reads
- ✓ Uniform coverage
- ✓ Single molecule resolution
- ✓ No PCR enables direct detection of base modifications (eg: 5mC, 6mA)
- ✓ Multi-omics capable

Image credit: pacb.com
Original publication: Eid, J., et al. (2009) Real-time DNA sequencing from single polymerase molecules. *Science*.

Simultaneous mapping of 6mA label, native DNA methylation (5mC), and genomic imprinting on a single DNA molecule

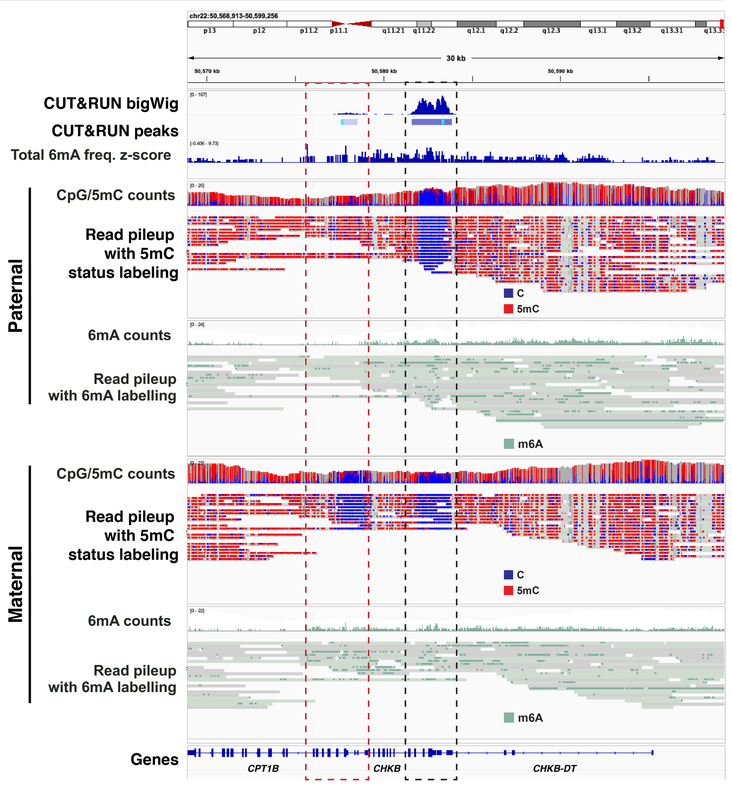
CUTANA-LRS reveals H3K4me3 enrichment flanked by 50 kbp domain of DNA methylation



Measuring both H3K4me3 via 6mA labelling and native DNA methylation (5mC)

IGV display of the CDC45 locus in GM12878 cells.
Top track (blue) H3K4me3 GM12878 CUT&RUN signal and peak calls.
Middle tracks (green): CUTANA-LRS H3K4me3-labeled experiment with 6mA label in green.
Lower tracks (red/blue): Native 5mC DNA methylation from the same reads shown in the middle tracks (green).

CUTANA-LRS resolves epigenomic imprinting



Conclusions

- ➢ CUTANA-LRS is a true multi-omic approach that simultaneously profiles histone PTMs, native DNAme (5mC), and parental single-nucleotide variants from single DNA molecules within a single reaction.
- ➢ Nondestructive and simple workflow preserves chromatin integrity for LRS.
- ➢ Highly reproducible across biological replicates and highly concordant with orthogonal short read assays (e.g., CUT&RUN).

Acknowledgement: This study was supported by the NIH-NHGRI.