WHITEPAPER

A New Era of Single-Molecule Insight in Genome Regulation

PRESENTED BY:

EpiCypher®

PUBLISHED BY:

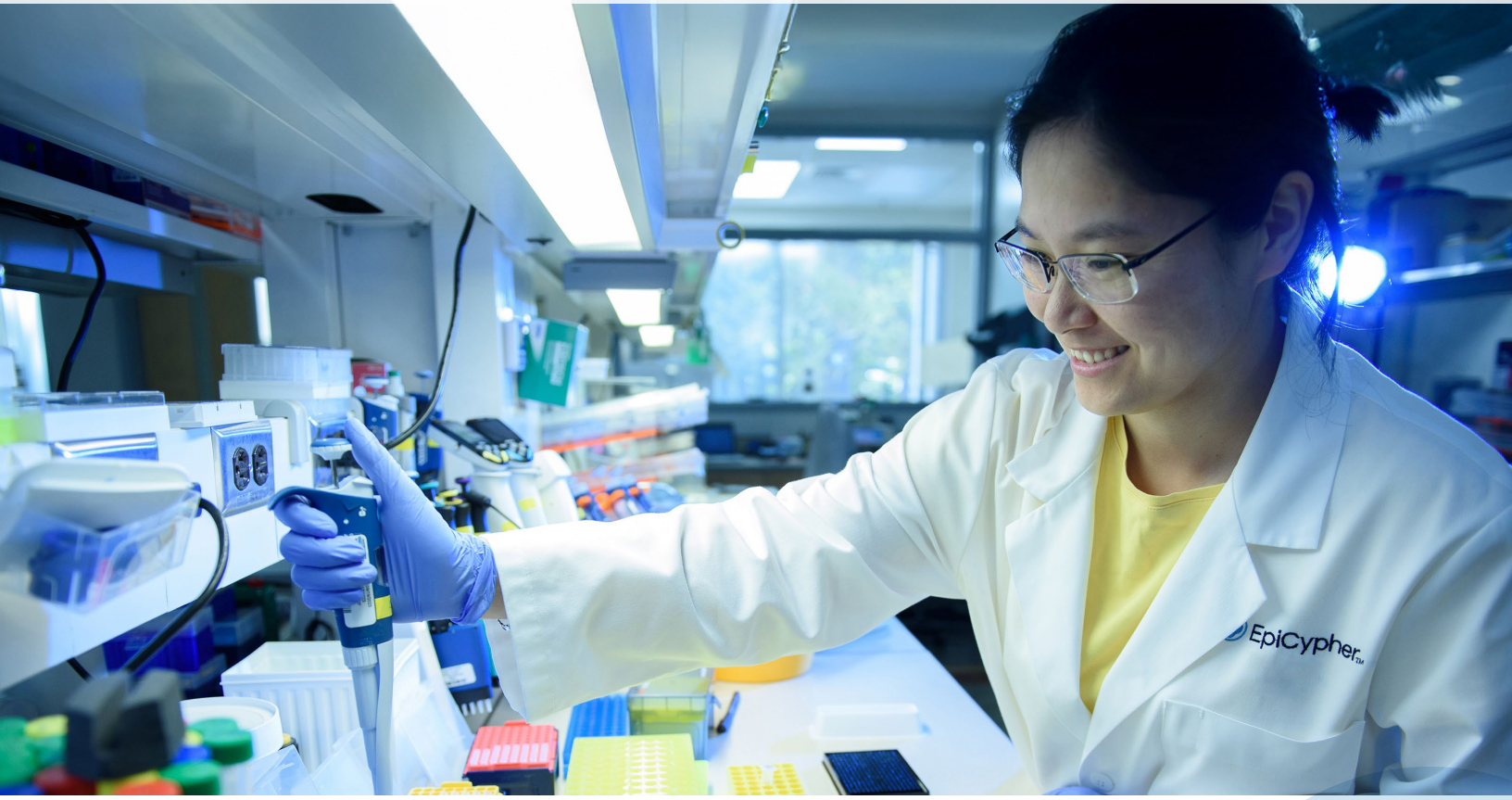
FIERCE Biotech

PRESENTED BY:

EpiCypher®

PUBLISHED BY:

FIERCE Biotech



A New Era of Single-Molecule Insight in Genome Regulation

CONTENTS

Introduction

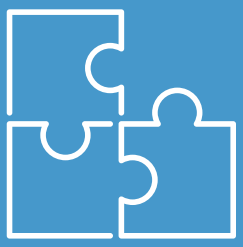
The availability of sequencing options has revolutionized both genomics and epigenomics. Short-read sequencing (SRS), which involves breaking the genome into fragments of less than 1,000 base pairs, is the most widely used approach, but it has many limitations..

Repetitive regions such as centromeres, telomeres and ribosomal DNA constitute about 8% of the human genome. “Short reads are totally blind to these regions of the genome,” says Bryan Venters, PhD, Senior Director of Genomic Technologies at EpiCypher. SRS is also blind to haplotype phasing, meaning it cannot distinguish the different copies of chromosomes inherited from each parent.

To address these shortcomings, researchers turn to long-read sequencing (LRS), which can read long range strands of DNA without the need for fragmenting. Unlike SRS, LRS can shed light on the genome’s repetitive regions. Not only that, but it can provide information on adjacent regions that are tens of kilobases away. LRS can resolve both haplotypes, allowing researchers to more thoroughly investigate imprinted diseases that have a genetic lesion on only one copy, such as Prader-Willi syndrome and Angelman syndrome.

Because of the discoveries it has empowered, LRS is having a heyday. The technology laid the groundwork for the first complete human genome, presented in June 2022 by the Telomere-to-Telomere Consortium. *Nature Methods* named LRS its Method of the Year 2022.[1]

To maximize the discoveries possible with LRS, in 2020 Andrew B. Stergachis and colleagues at Harvard Medical School and John Stamatoyannopoulos at the University of Washington introduced an assay called Fiber-seq. This adds chromatin context to LRS so researchers can see how the DNA is packaged, providing data on chromatin accessibility, DNA methylation and protein footprinting at a per-molecule resolution.



Unlike SRS, LRS can shed light on the genome’s repetitive regions. Not only that, but it can provide information on adjacent regions that are tens of kilobases away.

This whitepaper will explore Fiber-seq's introduction and adoption and explore how researchers are already using it to make discoveries that until recently would have been out of reach.

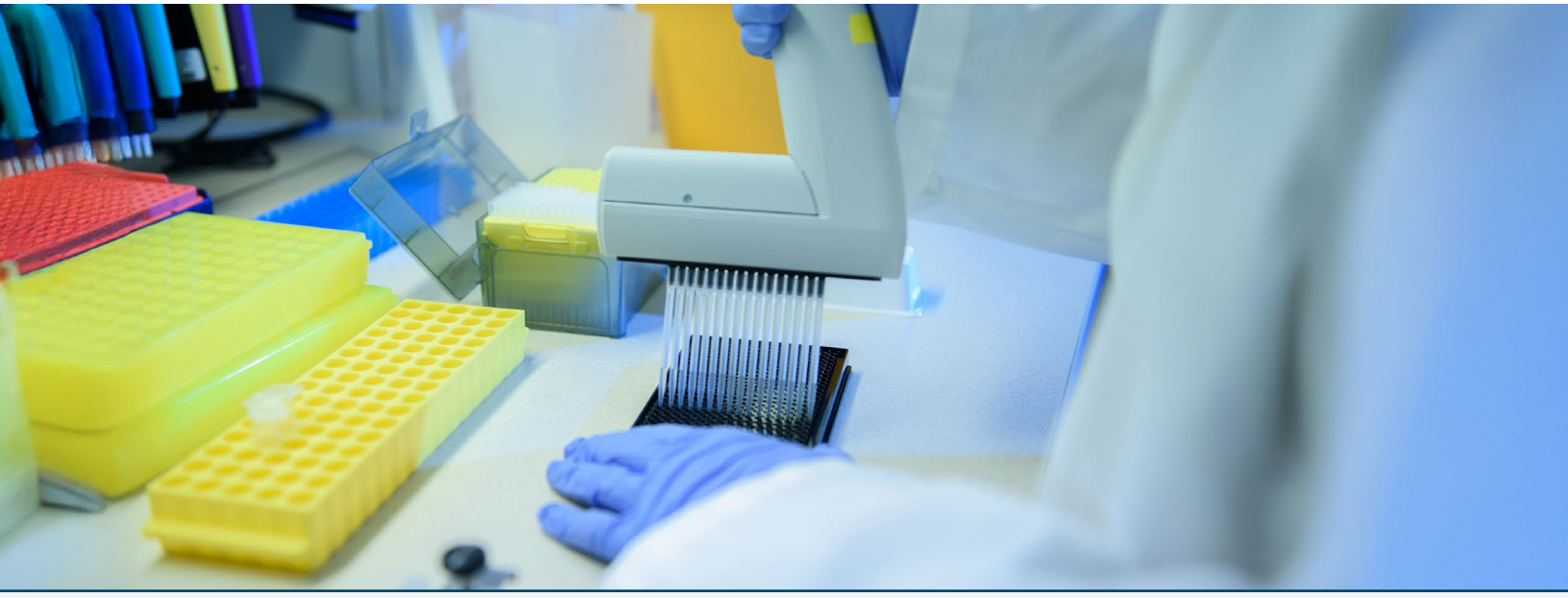
One giant leap for sequencing

Stergachis and colleagues first introduced Fiber-seq in a 2020 paper in *Science*.[2] The technique uses a DNA methyltransferase to mark DNA 6mA (N6-methyladenine) in a way that reveals how chromatin fibers — long stretches of DNA wrapped around proteins — are organized. This technique uses LRS to generate a nucleotide-resolution map of chromatin structure that includes the positions of nucleosomes.

In this initial paper, the authors used Fiber-seq to probe the relationship between regulatory DNA activation, nucleosome position, and DNA variation. This initial publication "was an amazing demonstration of the power of LRS multiomics," Venters says.

When it comes to epigenetics, native DNA 5mC (5-methylcytosine) methylation controls transcriptional patterns in humans and other eukaryotic organisms. This process is important to differentiation, cell identity, and responding to the environment. Fiber-seq overlays LRS results with information on functional chromatin organization, such as methylation patterns, nucleosome positioning, and regulatory protein footprints.

The combination of LRS plus Fiber-seq represents a significant advance over SRS. The latter requires time consuming treatments to distinguish 5mC from unmodified cytosines. LRS, in contrast, can directly and natively read multiple forms of methylation including 5mC; 5hmC (5-hydroxymethylcytosine); involved in embryonic stem cell development and neuronal cell function; and bacterial 4mC (N4-methylcytosine). When LRS is combined with Fiber-seq, the deposition and native detection of 6mA — a mark not typically found in humans — illuminates otherwise invisible chromatin features. "This allows you to see things that you wouldn't otherwise," Venters says.



Fiber-seq also overcomes the typical SRS requirement of multiple parallel assays and enables researchers to peer into structurally complex genomic regions. With Fiber-seq, “in a single 10-minute assay followed by LRS, scientists can get telomere-to-telomere genetic insights such as haplotypes, structural variants, translocations, SNPs, open chromatin, and native 5mC,” Venters explains.

Not only does Fiber-seq record chromatin accessibility, but its key enzyme operates at such a high sensitivity that it footprints nearly all nucleosomes and captures binding sites for many transcription factors, extracting a wealth of information from a single assay. These data are obtained on the same 10 to 20 kilobase DNA molecules, each originating from the same cell. “Not only do you get all those functional and epigenetic insights, but you are able to see cellular heterogeneity across those molecules,” Venters says. “There’s so much biology packed into a single data set that other approaches simply cannot obtain.”

Despite Fiber-seq’s high-profile publications and powerful potential, this promising method has flown under the radar for some time.

Fiber-seq is a sleeper hit

Fiber-seq requires users to make the leap to LRS. Venters believes one reason users have been slow to adopt it is because of misconceptions about the method’s cost. LRS used to be expensive, and that idea has persisted even as the price has come down.

For example, to do typical 30× coverage (or 90 gigabases) for DNA methylation or whole-genome sequencing, “it costs the same to sequence on LRS as it does on SRS - actually, cheaper — actually, cheaper in some cases,” Venters says. “A lot of people haven’t realized just how dramatically the cost of LRS has dropped in recent years.”

LRS with Fiber-seq also yields stronger, more valuable results. “I’ve been in genomics for a while, since the microarray days,” says Venters, a former research professor at Vanderbilt

There’s so much biology packed into a single data set that other approaches simply cannot obtain.

Bryan Venters, PhD,
Senior Director of Genomic Technologies at EpiCypher

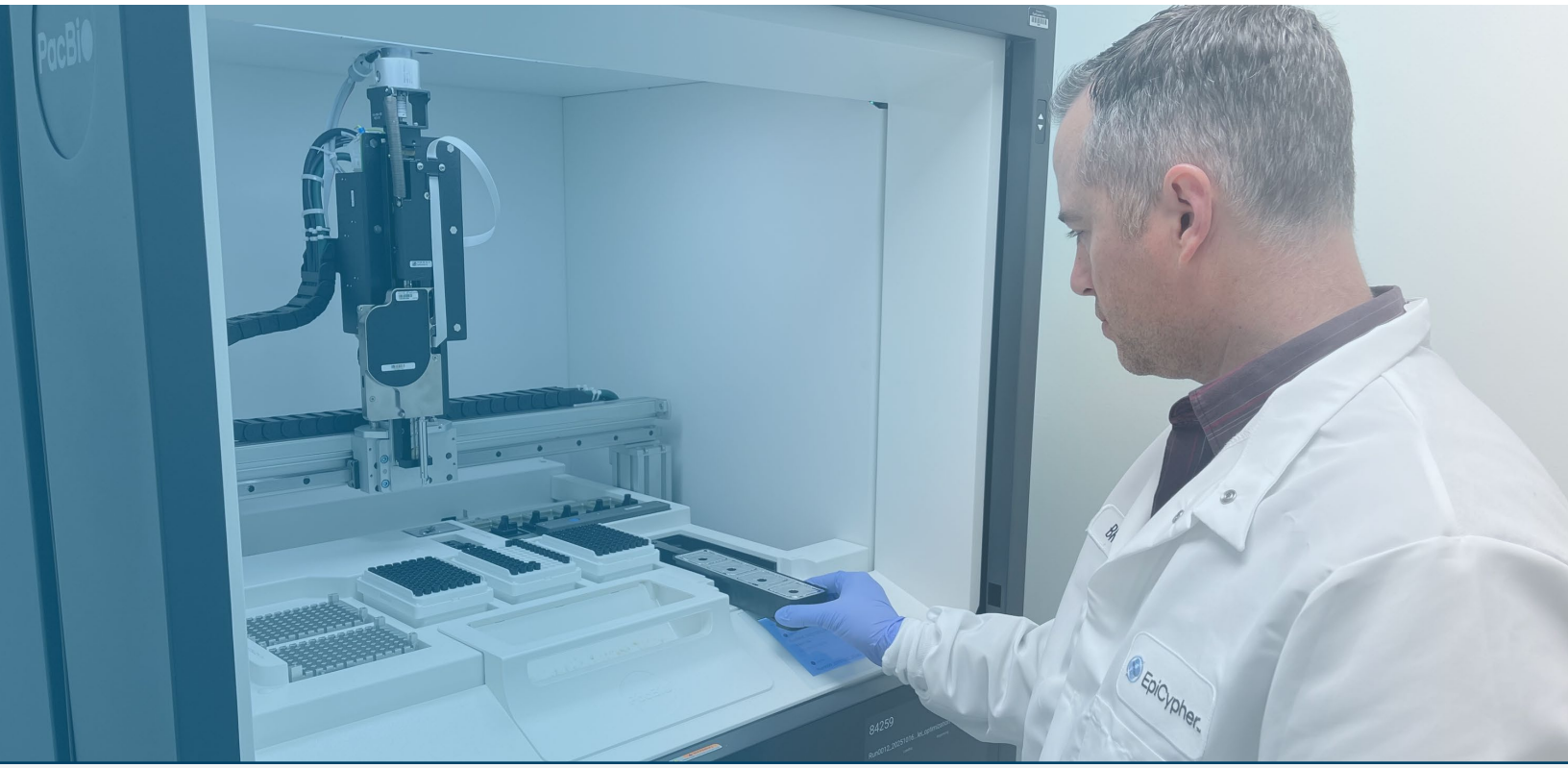
University. A common objection he has encountered is, “How do you get mechanistic insights when you have so many parallel datasets, and then you do correlations?”

In conventional SRS approaches, researchers perform multiple assays, ideally from a single batch of cells: ATAC-seq for chromatin accessibility; CUT&RUN or ChIP-seq for transcription factors and functional binding; MNase-seq for nucleosomes and chromatin structure; bisulfite conversion to identify native 5mC. Then the data sets must be correlated.

“There are some really cool studies that do this,” Venters says. However, correlation is not causation, and “such comparisons lack the strength of single-molecule data,” which provides information on cell heterogeneity with true multiomic power. “Ideally, you want to measure multiple genetic or epigenetic features on the same molecule from the same cell,” he explains.

With SRS, “it actually costs more to do all those assays rather than doing one Fiber-seq reaction, and you’re left with weaker, correlation-driven insights,” Venters says. “But with Fiber-seq, there’s really no comparison. You get all of that in a single assay with a simple 10-minute enzymatic incubation.”

ATAC-seq is one of the most highly published epigenetic assays, generating tens of thousands of research papers per year. “ATAC-seq users will love Fiber-seq,” Venters says. ATAC-seq yields information on open promoters, or stretches of unprotected DNA that are roughly 100 to 120 base pairs. It provides neither nucleosome footprinting nor



transcription factor footprinting because the enzyme that's involved is not as precise. "There have been efforts to get that information, but you have to sequence really deep," he explains.

For ATAC-seq users, "if you're sequencing out to hundreds of millions of short reads anyways, why would you not just do Fiber-seq so you can do the high resolution?" Venters asks. "For just a little bit more sequencing, you can get the open chromatin patterns and so much more — the nucleosomes, the transcription factors, the genetics, the heterogeneity. It's totally worth it."

Fiber-seq relies on the enzyme Hia5, a hyperactive nonspecific adenine methyltransferase first discovered by Monika Radlinska of Warsaw University. The enzyme is derived from a bacterium known as *Haemophilus influenzae*, which causes pinkeye. The enzyme's adenine methyltransferase activity is what allows the assay to record the chromatin architecture onto the DNA.

Until recently, this enzyme was not commercially available, so labs would have to produce their own. But EpiCypher now provides this enzyme commercially. "We fully characterize it, we know its stability, and we'll recommend how to use it," Venters says. EpiCypher has also launched Fiber-seq kits and Fiber-seq assay Services.

Fiber-seq provides insights at single-molecule resolution; integrates seamlessly into existing LRS workflows, and is compatible with the leading LRS manufacturers. "We have talked to many scientists from all over the world," Venters says. "Once they understand what Fiber-seq can do, they are just blown away, and they want to get their hands on it." Interested researchers can start exploring the dozens of published research papers to get a sense of how to integrate Fiber-seq.

> We have talked to many scientists from all over the world. Once they understand what Fiber-seq can do, they are just blown away, and they want to get their hands on it.
>
> **Bryan Venters, PhD,**
> Senior Director of Genomic Technologies at EpiCypher

Insights powered by Fiber-seq

Fiber-seq has been widely applied to many human tissue types, as well as plants, flies, yeast, and some less-common model organisms. Within the past year, Stergachis and colleagues have published on using Fiber-seq to resolve Mendelian conditions and to study the human brain.[3,4] Other labs have used the approach to study centromere biology.[5]

But Venters's favorite example for illustrating the power of Fiber-seq is a 2024 *Genome Research* paper that shed light on two challenging rare diseases: myotonic dystrophy and hereditary persistence of fetal hemoglobin.[6] "These are rare diseases where it has been challenging to understand how the mutation was impacting the disease," he explains.

By using Fiber-seq to resolve the heterogeneity and haplotype phasing, the researchers found that the mutation was present on only one parental haplotype of the regulatory element, and some of the disease phenotype was likely driven by epigenetic changes surrounding a promoter.

"That was one of the coolest examples because it blended the unique genetic insights that LRS can provide with the epigenetic layering that Fiber-seq gives you," Venters says.

Venters has heard from drug-discovery researchers who are interested in using Fiber-seq in their projects. "It's totally capable," he says. "Researchers really like that in a single assay, you are able to see chromatin accessibility, which is a commonly studied epigenetic feature tracked in drug discovery."

Clinicians are also interested in incorporating Fiber-seq into a frontline patient diagnostic regimen. By comparing the Fiber-seq data from drug responders versus non-responders with the same disease, "we would be able to look at their genetic differences, chromatin accessibility differences, DNA methylation patterns, and transcription factor footprints that may be driving the differences in how the patients are reacting to the drugs."

These are just a few of the real and potential discoveries that Fiber-seq enables.

Unlocking Fiber-seq's true potential

How can interested researchers start integrating Fiber-seq? "Start with the literature. There are dozens of papers out there," Venters says. "Fiber-seq will elevate the research impact of any project involving genetics data, methylation, ATAC, or chromatin accessibility." He encourages anyone working in those spaces to take a closer look at the technology.

For SRS users, making the switch to LRS may seem daunting. For example, the bioinformatics are different. But "if you know bioinformatics for SRS, you just need someone who's gone down that path with LRS to help you get started," Venters says. "That's one of the benefits of having a commercial source: we will help people get started on the bioinformatics."

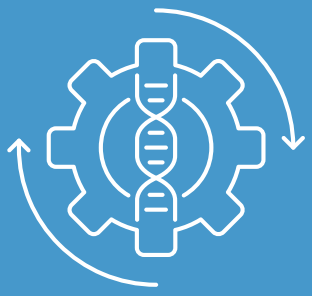
Whatever you do, "don't glide past this technology. Fiber-seq is an amazing multiomic method. These types of technologies don't come along very often," Venters says. Rather than an iteration on an existing technology, LRS with Fiber-seq is a true leap forward. "The number of insights you get per dollar is incomparable to SRS," he adds.

With the interest EpiCypher has seen so far in applying Fiber-seq in the clinic, Venters is optimistic that this innovative assay will be benefitting patients within the next 5 to 10 years. What is he most excited about? "The ability to diagnose disease and to be more informed about how to prescribe drugs to patients who will likely respond to them," he says. "That's really the potential of Fiber-seq."

How to get started with Fiber-seq

If you're interested in Fiber-seq, start exploring resources on the EpiCypher website. The site includes an overview of published literature and technical info to help researchers decide whether it fits with their projects.

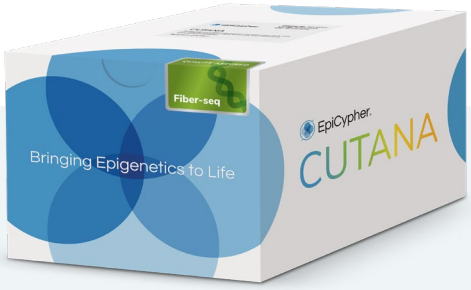
EpiCypher offers the Hia5 enzyme and Fiber-seq kits for labs, as well as fully integrated early-access services such as tissue-processing capabilities, in-house sequencing, and data analysis. The team operates a PacBio Revio on site and maintains high-end compute clusters to handle data processing locally. Reach out to the EpiCypher Services team to get started.



EpiCypher offers the Hia5 enzyme and Fiber-seq kits for labs, as well as fully integrated early-access services such as tissue-processing capabilities, in-house sequencing, and data analysis.

References

1. Method of the Year 2022: Long-read sequencing. Nature. January 12, 2023. https://www.nature.com/collections/eibbdadhga
2. Stergachis AB, Debo BM, Haugen E, Churchman LS, Stamatoyannopoulos JA. Single-molecule regulatory architectures captured by chromatin fiber sequencing. *Science*. Published online June 26, 2020. doi:10.1126/science.aaz1646
3. Vollger MR, Korlach J, Eldred KC, et al. Synchronized long-read genome, methylome, epigenome and transcriptome profiling resolve a Mendelian condition. *Nat Genet*. 2025;57(2):469-479. doi:10.1038/s41588-024-02067-0
4. Peter CJ, Agarwal A, Watanabe R, et al. Single chromatin fiber profiling and nucleosome position mapping in the human brain. *Cell Rep Methods*. 2024;4(12):100911. doi:10.1016/j.crmeth.2024.100911
5. Salinas-Luypaert C, Dubocanin D, Lee RJ, et al. DNA methylation influences human centromere positioning and function. *Nat Genet*. Published online September 4, 2025:1-13. doi:10.1038/s41588-025-02324-w
6. Bohaczuk SC, Amador ZJ, Li C, et al. Resolving the chromatin impact of mosaic variants with targeted Fiber-seq. *Genome Res*. 2024;34(12):2269-2278. doi:10.1101/gr.279747.124



EpiCypher®

EpiCypher was founded in response to the growing demand for high-quality reagents to study chromatin regulation and enable epigenetics-focused drug discovery. The company is at the forefront of chromatin mapping technologies with the CUTANA® platform for ultra-sensitive genomic profiling assays, including CUT&RUN and CUT&Tag, DNA methylation assays, and now, the groundbreaking Fiber-seq platform. EpiCypher also offers the largest collection of defined designer nucleosomes (dNucs) on the market along with complementary high-throughput assays and services. EpiCypher is dedicated to bringing these transformative technologies to market and offers superior products and assay services to researchers worldwide.

EpiCypher Resources