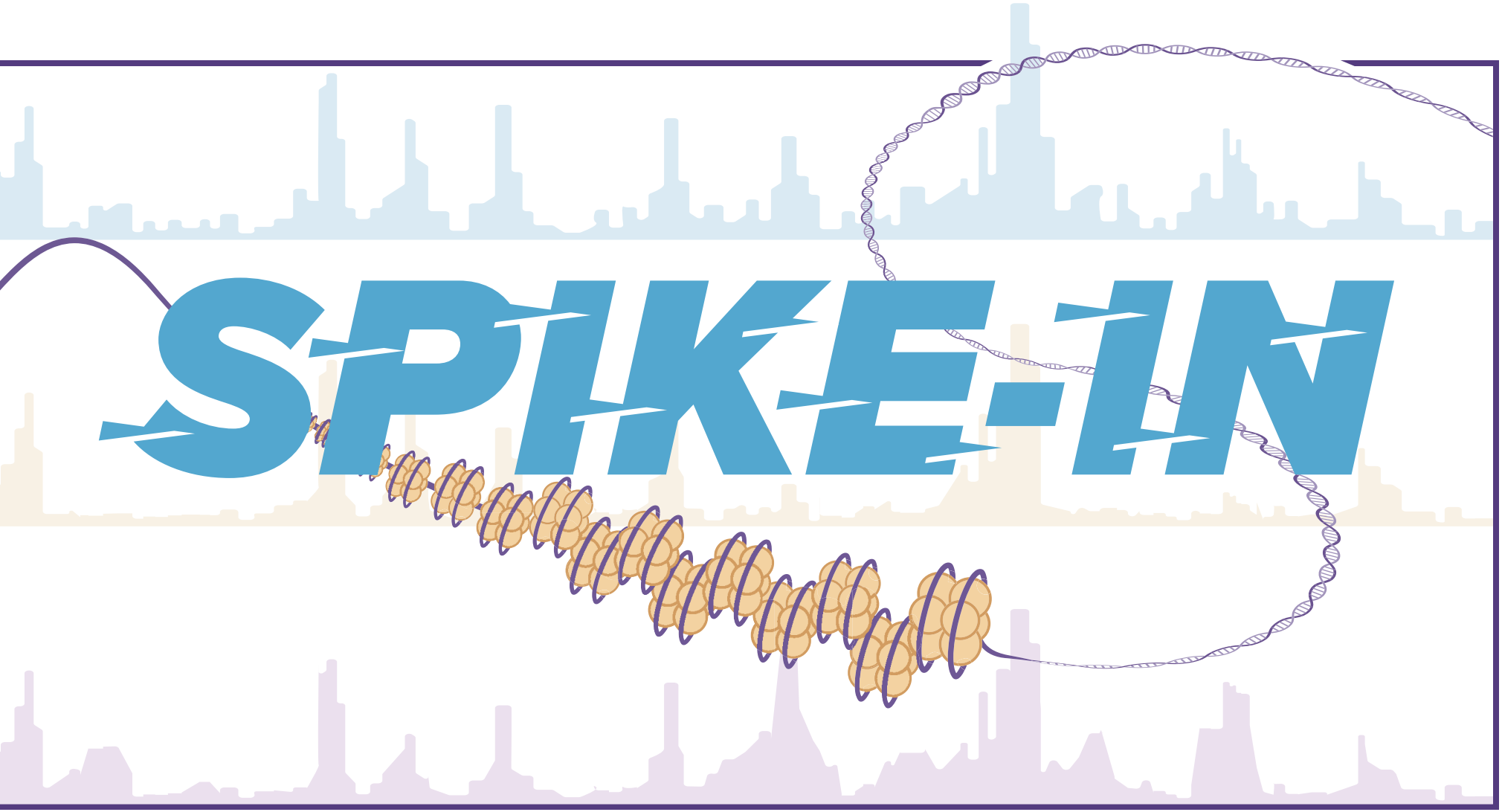
ACTIVE MOTIF®

Enabling Epigenetics Research

Spike-In Normalization: Complete Guide

Overview, Methods, and Resources



Introduction

Identifying differences between epigenetic assay data sets can be challenging when global modification changes occur, such as when studying the effects of chromatin modifying enzyme inhibitors. Additionally, inaccurate quantification of starting material or technical variation during processing results in variation across sample data. Currently available bioinformatic-based normalization methods are not applicable for normalizing across data sets in these instances, and the only reliable way to overcome bias and variation is to add a known standard (Spike-In) into all samples. Specific Spike-In reagents for ATAC-Seq, ChIP-Seq, CUT&RUN, and CUT&Tag are invaluable tools that overcome data variation challenges, allowing researchers to see, compare, and report true biological differences between samples in each of these assays.

With this eBook, you will find a complete resource for normalization of assay data sets – from the background and publication of Spike-In normalization to sample data and comparisons, tools and reagents available, and educational resources for further study.

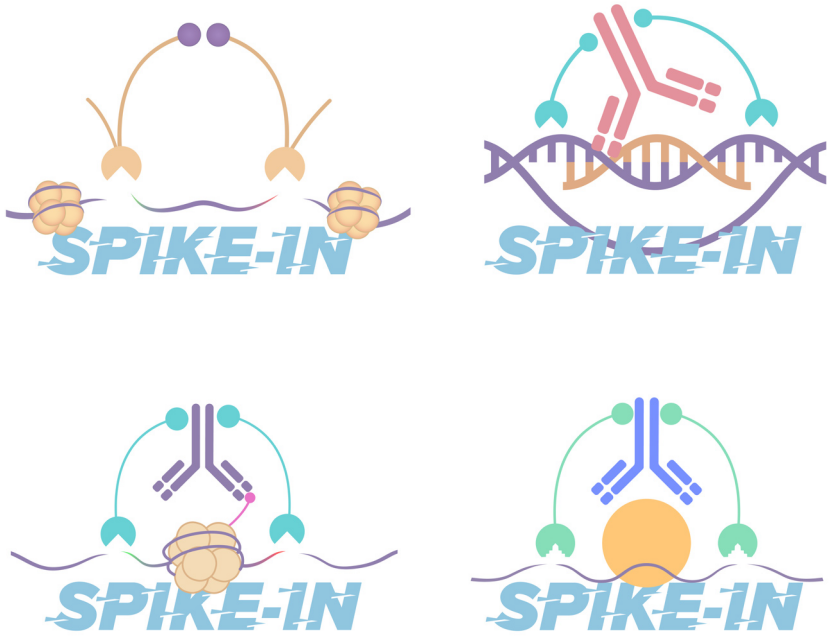


Table of Contents

Overview of Spike-In Normalization

It began with the original publication of this method by Brian Egan *et al.*, An Alternative Approach to ChIP-Seq Normalization Enables Detection of Genome-Wide Changes in Histone H3 Lysine 27 Trimethylation upon EZH2 Inhibition (PLoS One. 2016).[1] This important work focused on ChIP-Seq, which is widely used for the detection of transcription factor binding patterns and histone post-translational modification (PTM) occupancy profiles across the entire genome. ChIP-Seq data from many different cell types and contexts are used to generate genome-wide chromatin modification maps to explore the relationship between transcriptomic and epigenomic landscapes, and cell type comparisons often reveal substantial lineage-related differences in the profiles of specific histone PTMs across genomes. However, when there is manipulation of biological context, e.g. comparisons of knockdown or knockout of individual histone modifying enzymes or their inhibition with small molecules, these conditions can contribute very subtle alterations to the PTM landscape rather than resulting in a completely different, more easily detectable patterns. In an effort to address this, more complex statistical methods, software programs and computational models have been developed in an attempt to adequately compare ChIP-Seq data sets and reliably reveal the differences.

A previous technique, ChIP-Rx, based on the addition of a constant amount of reference cells from a different species, had been published for genome-wide quantitative comparison of histone modifications across different biological samples.[2] The method depends on the ability of the experimental ChIP antibody to recognize the histone modification of interest in both the reference and experimental cell species. Precipitated reference DNA is sequenced along with precipitated experimental DNA, and thus reference sequence reads provide a means to normalize across biological samples. Egan's team developed an alternative normalization approach based on the addition of *Drosophila melanogaster* chromatin and a *D. melanogaster*-specific antibody into standard ChIP reactions. Specifically, the use of an antibody that exclusively recognizes the *D. melanogaster* histone variant H2Av enables precipitation of *D. melanogaster* chromatin as a minor fraction of the total ChIP DNA. The *D. melanogaster* ChIP-Seq tags are used to normalize the human ChIP-Seq data from samples.

One particularly compelling example of the importance of this ChIP-Seq normalization strategy is illustrated in their experiments investigating the genome-wide localization of the repressive histone mark H3K27me3 in the presence of epigenetic inhibitors. H3K27me3 has one identified methyltransferase (EZH2) responsible for catalyzing methyl group addition. However, EZH2 inhibition did not appear to significantly affect H3K27me3 levels in a ChIP-Seq experiment. Following normalization to Spike-In, a substantial reduction in H3K27me3 signal was then observed in ChIP-Seq data from EZH2 inhibitor treated samples, demonstrating how the use of Spike-In normalization can reveal changes in histone modifications that would otherwise be masked by experimental artifacts.

1. Egan B, Yuan CC, Craske ML, Labhart P, Guler GD, Arnott D, Maile TM, Busby J, Henry C, et al. An Alternative Approach to ChIP-Seq Normalization Enables Detection of Genome-Wide Changes in Histone H3 Lysine 27 Trimethylation upon EZH2 Inhibition. PLoS ONE 11(11): e0166438. PMID: 27875550

2. Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER et al. Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. Cell Reports. 2014;9(3):1163–1170. PMID:25437568

ChIP-Seq Spike-In

How Does ChIP-Seq Spike-In Work?

ChIP normalization can be easily implemented simply by integrating Spike-In reagents into your standard ChIP protocol using experimental chromatin and an antibody of interest. Spike-In Chromatin and Spike-In Antibody are also added, or "spiked in", to the ChIP reaction as a minor fraction of the IP reaction. Any variation introduced during the ChIP reaction will also occur with the Spike-In Chromatin. As the amount of spiked-in chromatin is consistent across all samples, a normalization factor can be created based on the Spike-In signal and applied to the experimental signal.

Spike-In chromatin and a Spike-In Antibody can easily be integrated into an existing ChIP-Seq workflow. Spike-In Chromatin and the Spike-In Antibody are added to experimental chromatin and the experimental antibody just prior to immunoprecipitation. The Spike-In antibody recognizes a histone variant that is specific to the species of the Spike-In Chromatin (*Drosophila*), and the experimental antibody specifically recognizes the experimental chromatin. This enables specific detection of the Spike-In Chromatin without any significant increase in background signal. Following sequencing, reads will be mapped to their specific species. Variation introduced during the ChIP procedure will affect the Spike-In Chromatin in the same manner as the experimental chromatin, so a normalization factor can be created from the Spike-In Chromatin and applied to the experimental chromatin to normalize out technical variation and sample bias, or to monitor biological effects.

ChIP Spike-In Normalization Advantages

- Reduce effects of technical variation to detect subtle biological differences not seen with standard ChIP
- Can be applied across different antibodies and samples without bias
- Spike-In Chromatin and Spike-In Antibody can be used with any ChIP kit or protocol
- Strategy works with both qPCR and ChIP-Seq analysis

ChIP-Seq Spike-In Normalization publications

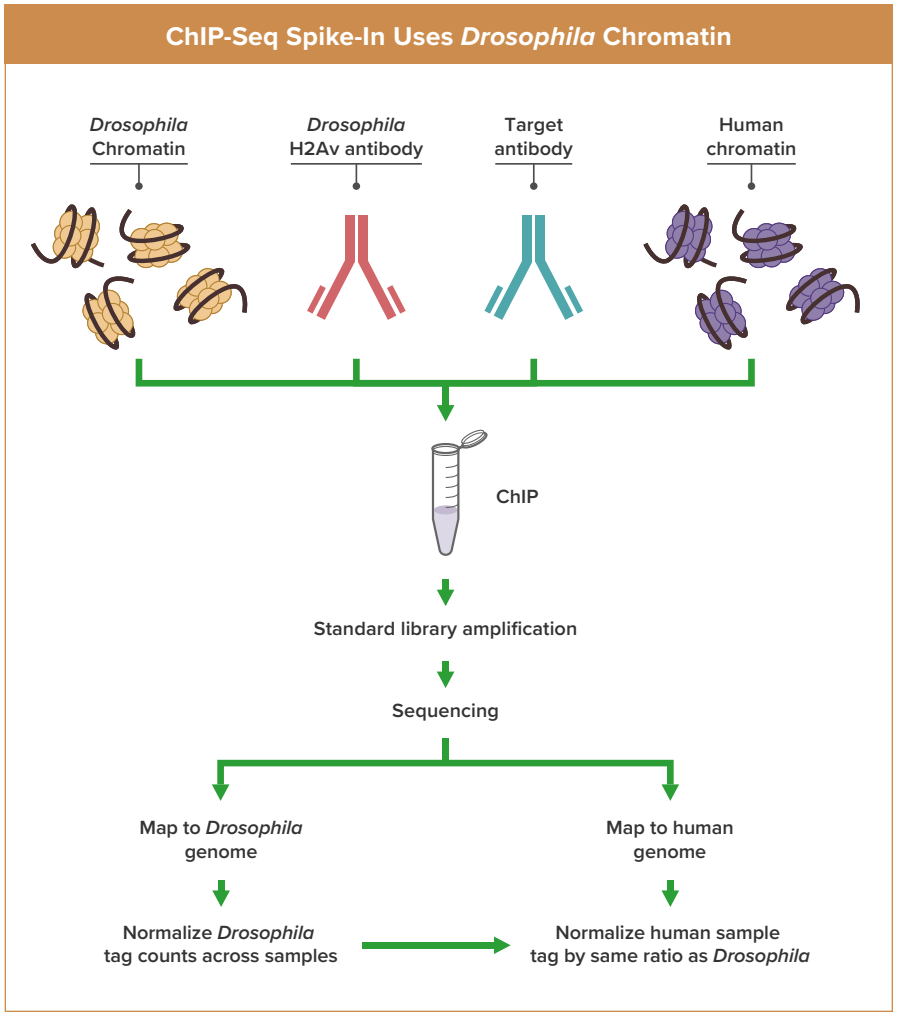


Figure 1. ChIP-Seq Normalization Workflow

A standard ChIP reaction is set up using experimental chromatin (e.g. human) and an antibody of interest. In addition, *Drosophila melanogaster* chromatin is added, or "spiked-in" to each reaction as a minor fraction of the total chromatin. An antibody that recognizes the *Drosophila*-specific histone variant, H2Av, is added to the reaction. The Spike-In Antibody provides a mechanism to reliably pull down a small fraction of the *Drosophila* chromatin that is consistent across all samples. Following ChIP sequencing, the data is mapped to both the *Drosophila* genome and the experimental genome. A normalization factor is created for each sample based on the *Drosophila* tag counts. The experimental tag counts are normalized by the same factor. normalization factor is created for each sample based on the *Drosophila* tag counts. The experimental tag counts are normalized by the same factor.

Specificity of Detection

The Spike-In Chromatin consists of *Drosophila melanogaster* chromatin prepared from Schneider's *Drosophila* Line 2 (S2) cells. The Spike-In antibody recognizes a *Drosophila*-specific histone variant, H2Av. Because of the specificity of the Spike-In Antibody for the Spike-In Chromatin modification, there is no cross-reactivity with mammalian samples leading to reduced background signal.

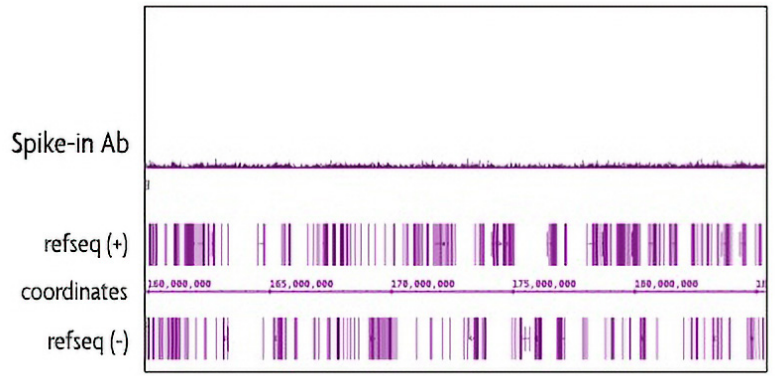
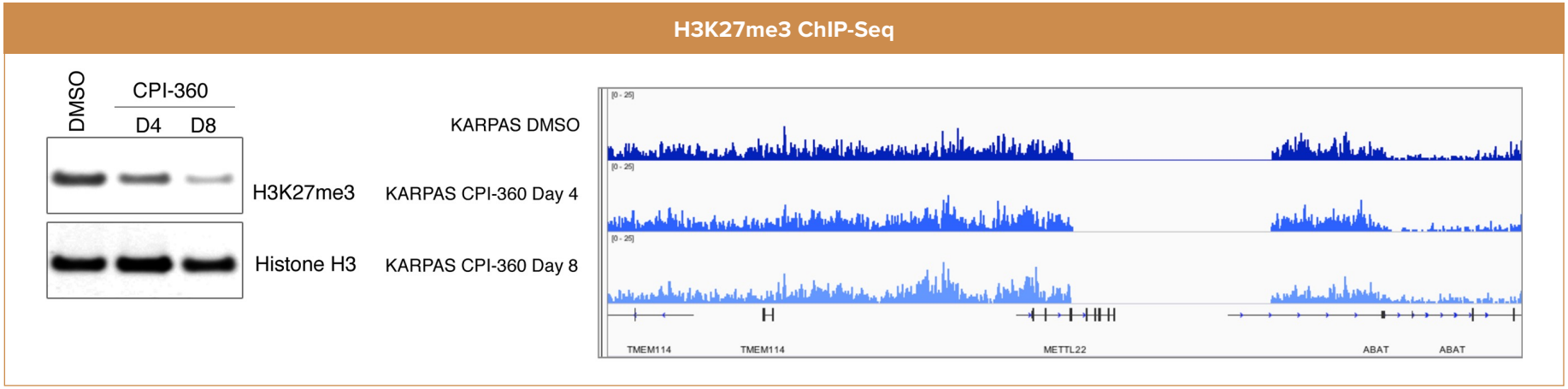


Figure 2: Specificity of the Spike-In Antibody

The Spike-In Antibody shows minimal cross reactivity with mammalian samples. When the Spike-In Antibody was tested in ChIP-Seq with human chromatin, there is little to no signal detected. This demonstrates the specificity of the Spike-In normalization strategy.



How Can Spike-In Normalization Improve ChIP-Seq Analysis?

One particularly compelling example of the importance of ChIP-Seq normalization is illustrated in experiments investigating the genome-wide localization of the repressive histone mark H3K27me3 in the presence of epigenetic inhibitors. H3K27me3 has one identified methyltransferase (EZH2) responsible for catalyzing methyl group addition. However, EZH2 inhibition did not appear to significantly affect H3K27me3 levels in a ChIP-Seq experiment without normalization to Spike-In, in which the significant loss of H3K27me3 becomes apparent. Therefore, Spike-In can reveal changes in histone modifications that would otherwise be masked by experimental artifacts (Figure 3).

Figure 3: Standard ChIP-Seq Methods Do Not Reveal EZH2 Inhibition Without Spike-In Normalization

Left Westerm blot showing reduced global H3K27me3 levels in KARPAS-422 cells treated with 1.5 µM CPI-360 for 4 and 8 days. Whole cell extracts were resolved by SDS-PAGE and immuno-blotted with anti-H3K27me3. Anti-H3 immuno-blots show equal levels of total H3. **Right** Representation of H3K27me3 ChIP-Seq data using IGV. No obvious differences are detected in CPI-360 treated KARPAS-422 cells when compared to KARPAS DMSO controls.

Egan B., et al., An Alternative Approach to ChIP-Seq Normalization Enables Detection of Genome-Wide Changes in Histone H3 Lysine 27 Trimethylation upon EZH2 Inhibition. PLoS One, 2016. 11(11):e0166438

ATAC-Seq Spike-In

Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-Seq) has emerged as a powerful method for investigating open or accessible chromatin across the genome. However, identifying differences between data sets can be challenging when global modification changes occur, such as studying the effects of chromatin modifying enzyme inhibitors. Additionally inaccurate quantification of starting material or technical variation during processing results in variation across sample data. Active Motif's strategy for ATAC-Seq normalization is to Spike-In cryopreserved *Drosophila* cell nuclei into samples prior to nuclei prep and tagmention. During tagmentation both the test cells and the *Drosophila* nuclei are tagged at open chromatin consistently across all samples. A normalization factor is then created based on the *Drosophila* signal and applied to the test genome.

ATAC-Seq Spike-In Control Advantages

- Identify differences between datasets
- Reveal changes in open chromatin that were masked by cell number differences
- Overcome bias and variation with known standard

Spike-In normalizes out differences in starting cell numbers

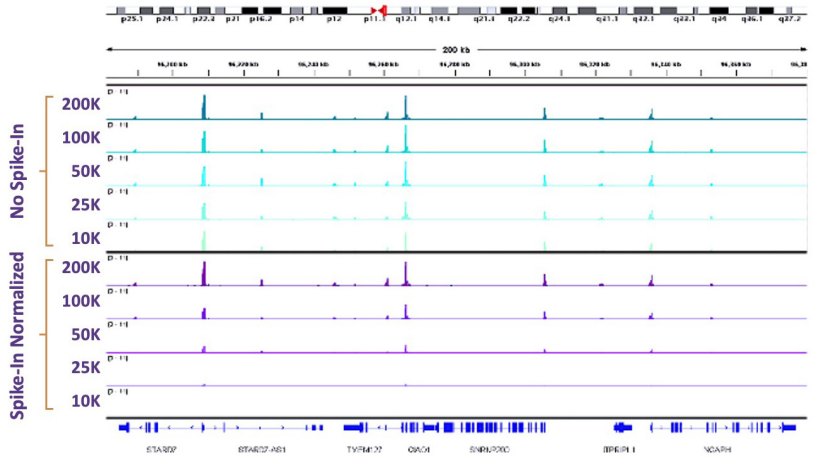


Figure 4: K562 Starting Cell Numbers for ATAC-Seq with and without Spike-In Normalization

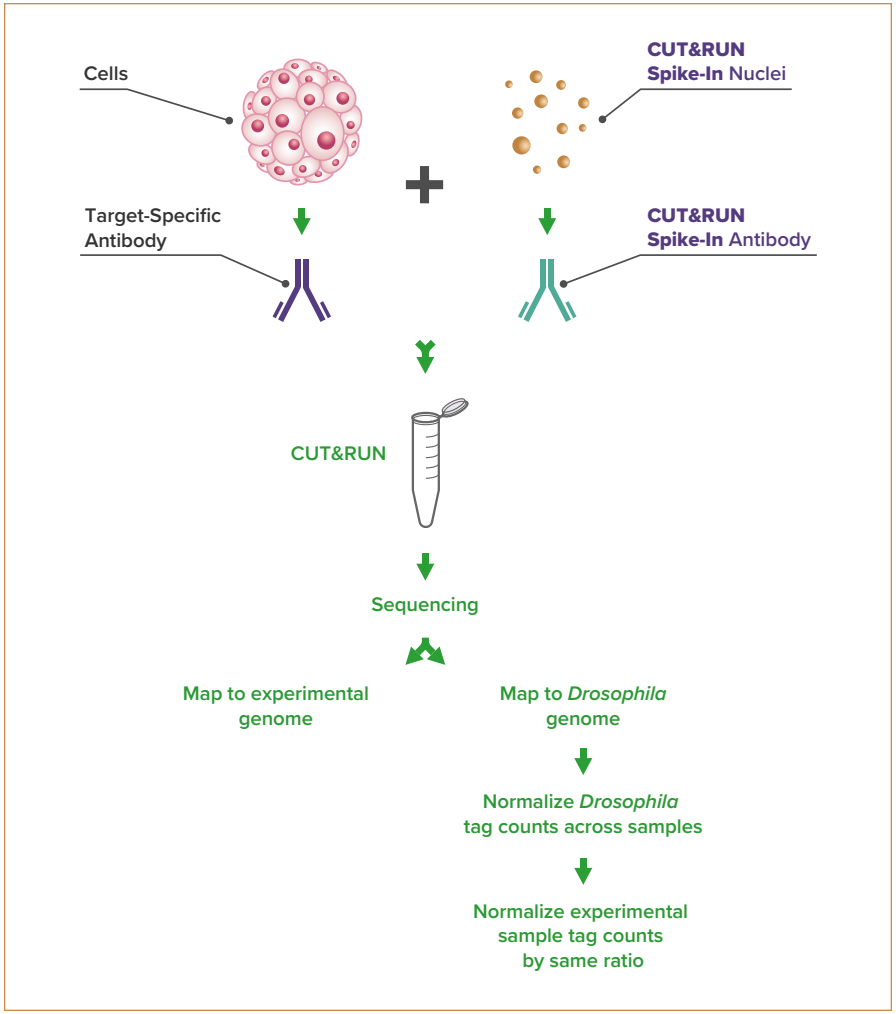
This experiment used 10,000 *Drosophila* nuclei as a spike in for a mock treatment. The mock treatment was a cell titration performed in duplicate from 200,000 cryopreserved K562 cells down to 10,000 cryopreserved K562 cells. The actual cell numbers are 200,000, 100,000, 50,000, 20,000 and 10,000 cells. This figure is both the spike-in normalized and unnormalized tracks for one of the duplicates in the titration.

CUT&RUN Spike-In

Overview

In CUT&RUN, a protein of interest is tagged with an antibody and bound to the chromatin in intact cells. A micrococcal nuclease (MNase) is used to cleave the DNA specifically at the binding sites of the protein of interest. The released fragments are purified, sequenced, and mapped to the reference genome to determine the protein's binding sites. Unlike ChIP, CUT&RUN does not require crosslinking of the protein to the DNA, which can introduce artifacts. CUT&RUN is a valuable tool for studying chromatin-associated proteins because it is sensitive, specific, and requires fewer cells than ChIP, making it ideal for identifying binding patterns of chromatin-associated proteins such as transcription factors or histone modifications genome-wide. Chromatin-associated proteins play critical roles in regulating various cellular processes such as gene expression, DNA replication, DNA repair, and cell differentiation.

Active Motif's strategy for CUT&RUN normalization is to Spike-In cryopreserved *Drosophila* cell nuclei into samples prior to CUT&RUN. Then, during the antibody incubation step, a *Drosophila* H2Av antibody is added in addition to the antibody to the target of interest. This *Drosophila* H2Av antibody provides a mechanism to reliably tag *Drosophila* histones in a consistent way across all samples. A normalization factor is then created based on the *Drosophila* signal and applied to the test genome. This CUT&RUN Spike-In strategy enables normalization of CUT&RUN data independent of the experimental antibody and without bias.



CUT&RUN Spike-In Control Data

Active Motif's CUT&RUN normalization strategy may be applied to any mammalian CUT&RUN assay reaction due to the lack of cross-reactivity of the Spike-In antibody with mammalian samples. The amount of cryopreserved *Drosophila* nuclei and antibody used per CUT&RUN reaction may need to be optimized with the goal of having *Drosophila* reads make up only 5-10% of the total sequencing reads. However, when using robust antibodies against tightly localized histone modifications, such as H3K4me3, we recommend a Spike-In:test sample ratio of 1:20. For antibodies against spreading marks such as H3K27me3, we recommend a Spike-In:test sample ratio of 1:10. For antibodies against transcription factors such as YY1, we recommend a Spike-In:test sample ratio of 1:100.

CUT&RUN Spike-In Control Advantages

- Compare between CUT&RUN datasets between experimental samples
- Reveal differences masked by cell number inputs
- Simply add Spike-In Nuclei to samples and perform CUT&RUN with the CUT&RUN Spike-In Antibody together with experimental target antibody

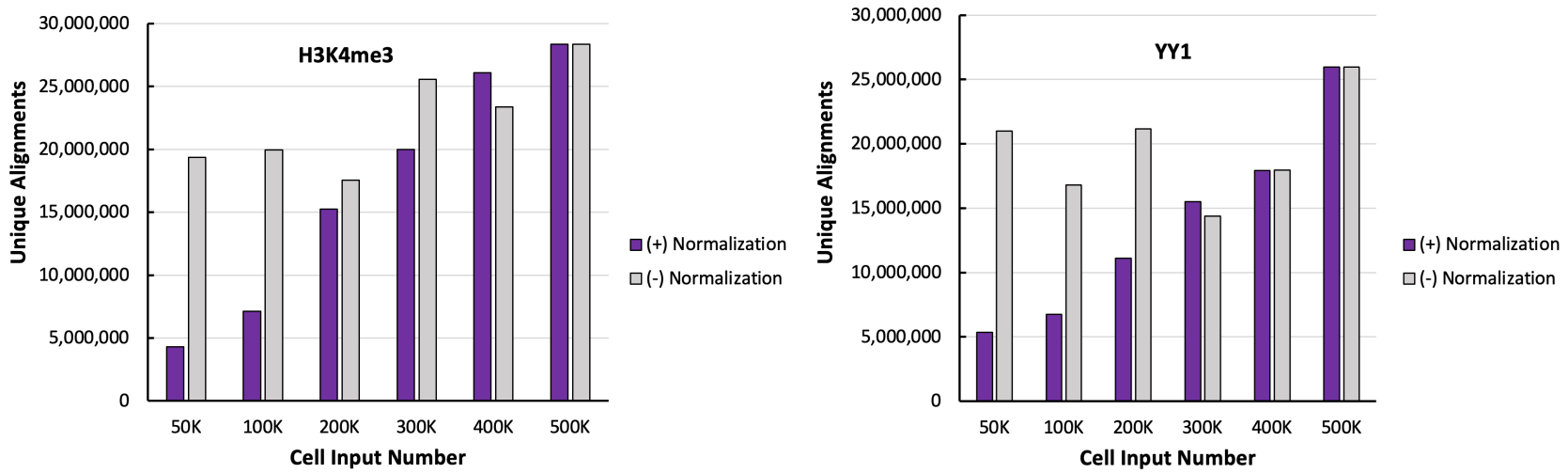


Figure 5. K562 Paired Alignments with and without Normalization

The number of alignments is not reflective of the number of cells added to the experiments targeting H3K4me3 (left panel) and YY1 (right panel), thus normalization is needed.

GSK126 global decrease in EED binding

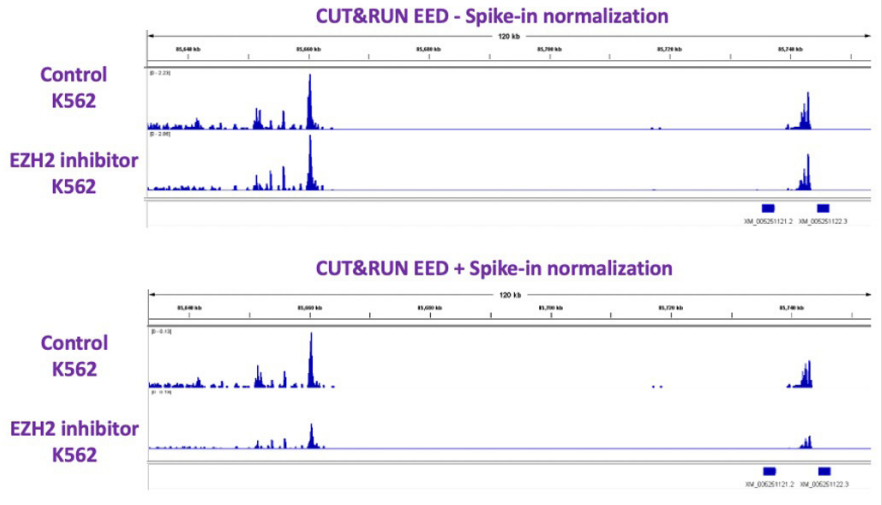


Figure 6. Spike-In Normalization is Necessary to Reveal Global Decrease in EED Binding in the Presence of GSK126 Treatment

500,000 K562 cells (Control) and 500,000 K562 cells treated with GSK126 were assayed in CUT&RUN for EED. Without Spike-In normalization, (top two IGV tracks) no decrease in global EED binding is observed. However, with Spike-In normalization (bottom two IGV tracks), a reduction in EED binding is observed.

CUT&Tag-IT® Spike-In

How Does CUT&Tag-IT Spike-In Control Work?

Active Motif's strategy for CUT&Tag normalization is to spike cryopreserved *Drosophila* cell nuclei into samples prior to CUT&Tag. Then, during the primary antibody incubation step, a *Drosophila* H2Av antibody is added in addition to the antibody targeting the histone mark of interest. This *Drosophila* H2Av antibody does not cross react to other species and provides a mechanism to reliably tag *Drosophila* chromatin in a consistent way across all samples. A normalization factor is then created based on the *Drosophila* signal and applied to the test genome. This CUT&Tag-IT® Spike-In Control strategy enables normalization of CUT&Tag data independent of the experimental antibody and without bias.

CUT&Tag-IT® Spike-In Control Advantages

- Compare CUT&Tag datasets between samples
- Simply add Spike-In Nuclei to samples and perform CUT&Tag with the Spike-In Antibody together with the experimental target antibody
- Obtain normalization factor and reveal true differences between samples

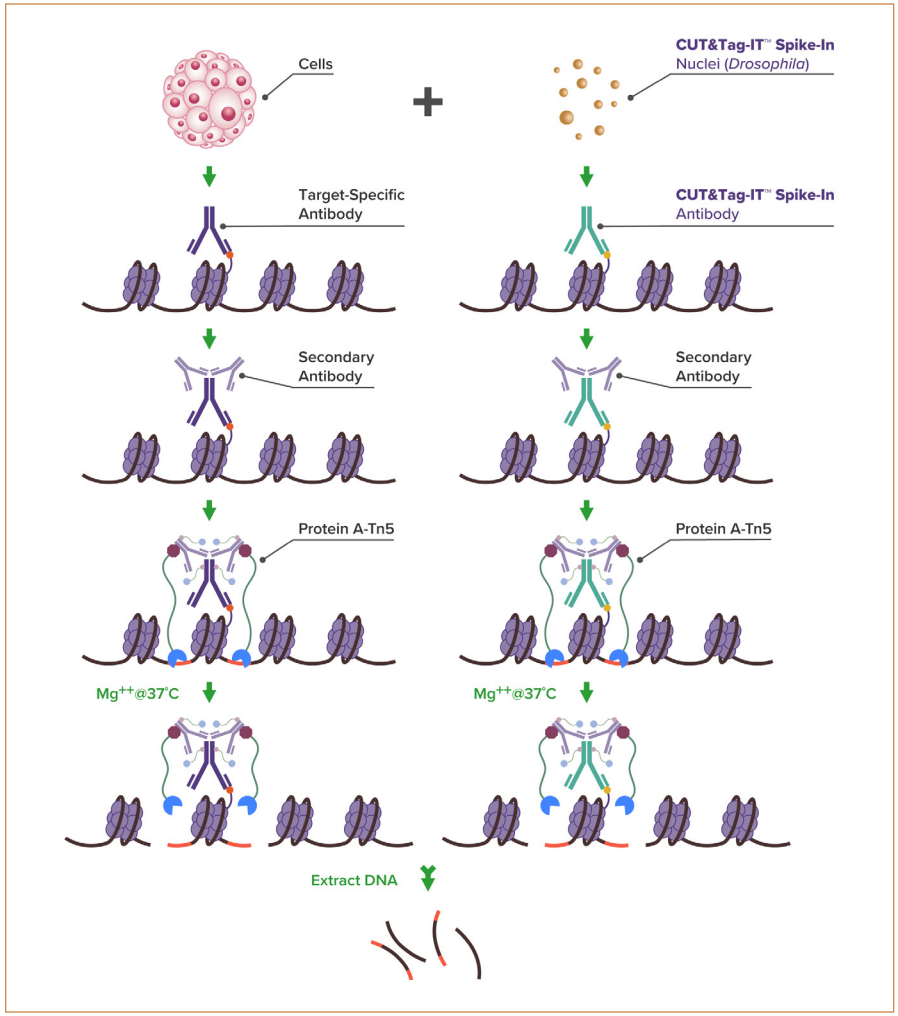


Figure 7: CUT&Tag-IT Spike-In Control Method

Spike-In Nuclei are added to samples and assayed in CUT&Tag with the Spike-In Antibody and experimental target antibody together in the same reaction throughout the complete CUT&Tag assay.

CUT&Tag-IT® Spike-In Control Data

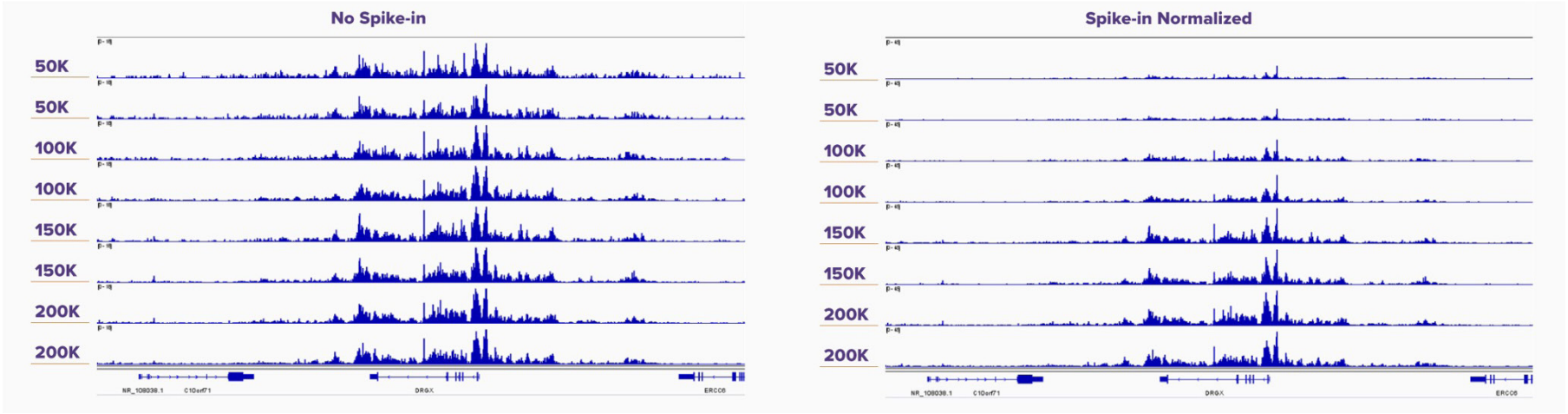


Figure 8: K562 CUT&Tag-IT Assay Results With and Without CUT&Tag-IT Spike-In Control Normalization

IGV browswer tracks are shown for duplicate reactions of 50,000, 100,000, 150,000, and 200,000 K562 cells assayed in CUT&Tag for H3K27me3 without spike-in and with CUT&Tag-IT Spike-In Control normalization. In the Spike-in Normalized panel it is clear that the peak height correlates with increasing cell number, whereas in the No Spike-in control peak height differences are not noticeable from 50,000 to 200,000 cells.

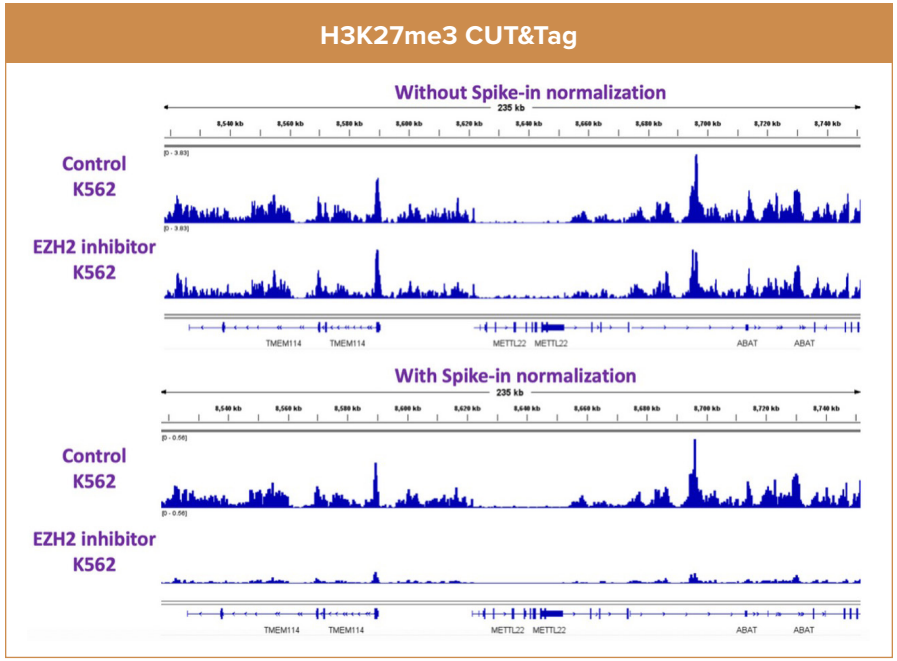


Figure 9. Spike-In Normalization After EZH2 Inhibition Reveals Expected Changes

IGV browser tracks are shown for control samples and samples treated with an EZH2 inhibitor. Data was analyzed both without and with Spike-in normalization. Global decrease in H3K27me3 histone modification are only detectable with Spike-in normalization.

CUT&Tag-IT® R-loop Spike-In

Cleavage Under Targets and Tagmentation (CUT&Tag) has emerged as a powerful method for profiling the localization of histone modifications of interest genome wide. Active Motif applied this approach to study DNA-RNA Hybrid R-loops genome-wide with our CUT&Tag-IT® R-loop Assay Kit. The CUT&Tag-IT® R-loop Assay Kit utilizes a CUT&Tag-based approach to profile R-loops genome-wide using a specifically optimized protocol for cell samples and our DNA-RNA hybrid (Clone S9.6) antibody. Due to variation in cell numbers and NGS sequencing statistics between samples, biological differences can be obscured. By applying Active Motif's CUT&Tag normalization strategy to the CUT&Tag-IT R-loop Assay data, true differences between samples can be revealed.

CUT&Tag-IT® Spike-In Control, R-loop Advantages

- Compare CUT&Tag-IT R-loop Assay Kit datasets between samples
- Simply add Spike-In Nuclei, R-loop to experimental samples and perform CUT&Tag with the Spike-In Antibody, Mouse together with the DNA-RNA Hybrid mAb (Clone S9.6)
- Obtain normalization factor and reveal true differences between samples

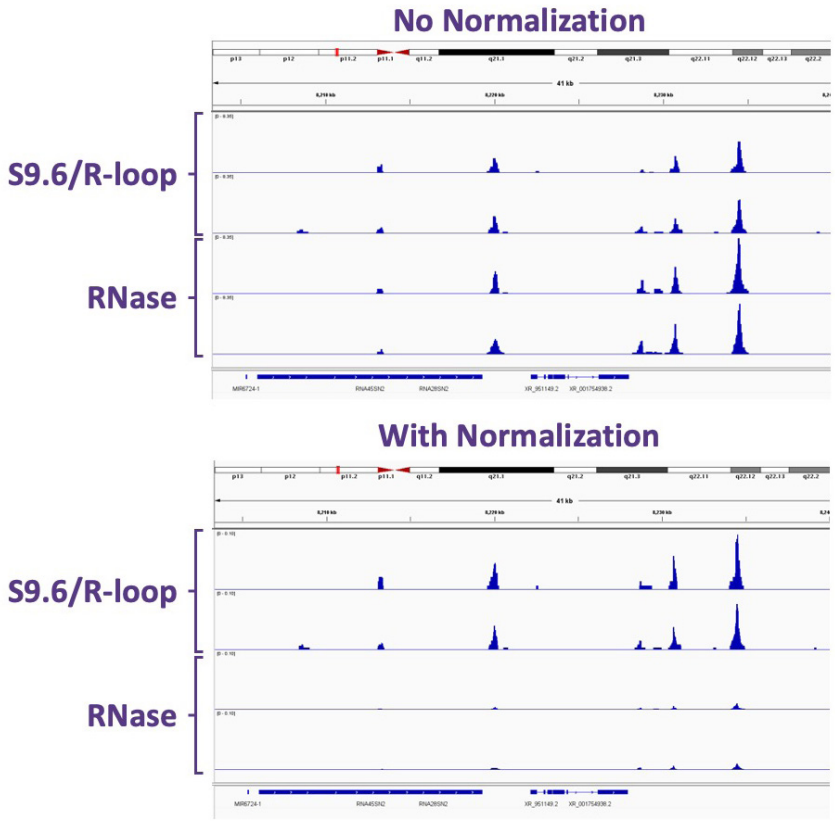


Figure 11. Spike-In is Needed When Using Negative Control RNase Treatment in CUT&Tag-IT R-loop Assay

R-loops are DNA-RNA Hybrids within the genome, and RNase A will cut the hybrid, destroying the R-loop, thus RNase A treatment is used as a negative control in the CUT&Tag-IT R-loop Assay Kit. However, without Spike-In Normalization, (top panel) no differences are observed. With Spike-In Normalization, the effect of the treatment in the negative control is observed.

Resources & Ordering Information

Application Resource Centers

Whether you're just beginning these techniques, or are interested in learning more, access our complete resource centers for each. Get the latest Webinars & Videos, Podcasts, eBooks, TechNotes & Protocols for these and a range of epigenetic assay types:

- ChIP-Seq
- ATAC-Seq
- CUT&RUN
- CUT&Tag
- DNA Methylation

Ordering Information

Name	Format	Cat No.
Spike-in Antibody	50 µg	61686
Spike-in Chromatin	15 rxns	53083
Drosophila Positive Control Primer Set Pbgs	96 rxns	71037
Drosophila Negative Control Primer Set 1	96 rxns	71028
Drosophila Negative Control Primer Set 2	96 rxns	71029
Drosophila Negative Control Primer Set 3	96 rxns	71038
ATAC-Seq Spike-In Control	16 rxns	53154
ATAC-Seq Kit	16 rxns	53150
Fixed Cell ATAC-Seq Kit	16 rxns	53151
CUT&RUN Spike-In Control	24 rxns	53183
ChIC/CUT&RUN Assay Kit	24 rxns	53180
ChIC/CUT&RUN pAG-MNase	50 rxns	53181
CUT&Tag-IT® Spike-In Control, Anti-Rabbit	16 rxns	53168
CUT&Tag-IT® Spike-In Control, Anti-Mouse	16 rxns	53173
CUT&Tag-IT® Assay Kit, Anti-Rabbit	16 rxns	53160
CUT&Tag-IT® Assay Kit, Anti-Mouse	16 rxns	53165
CUT&Tag-IT® Core Assay Kit	16 rxns	53176
CUT&Tag-IT® Spike-In Control, R-loop	16 rxns	53174
CUT&Tag-IT® R-loop Assay Kit	16 rxns	53167

Webinar

To learn more, watch the Spike-In Normalization webinar at activemotif.com/webinars

In this webinar, Brian Egan, Head of R&D at Active Motif discusses why data normalization is often critical for sequencing-based assays like ChIP-Seq, ATAC-Seq, CUT&RUN and CUT&Tag. The concept behind our suite of Spike-In Controls is covered, as well as how and when to use these tools to obtain the best possible results when analyzing experimental data.