EpiCypher® SNAP-CUTANA™ Spike-in User Guide

This User Guide describes EpiCypher's quantitative nucleosome spike-in technology, or SNAP (Sample Normalization and Antibody Profiling) Spike-in Controls for CUTANA™ CUT&RUN and CUT&Tag assays.

Table of Contents

1. CUTANA™ CUT&RUN and CUT&Tag Assay Overview

Cleavage Under Targets & Release Using Nuclease (CUT&RUN) and Cleavage Under Targets & Tagmentation (CUT&Tag) are revolutionary genomic mapping strategies developed by the group of Dr. Steven Henikoff[1,2]. Both assays build on the Chromatin ImmunoCleavage (ChIC) approach from Dr. Ulrich Laemmli[3], in which Protein A/G is used to recruit an enzymatic domain to antibody-bound chromatin *in situ*[3]. An important feature of CUT&RUN/CUT&Tag is the immobilization of cells (or nuclei) to solid support (**Figure 1**), which streamlines assay workflows, improves signal-to-noise (vs. ChIP), and enables low cell inputs and sequencing requirements.

In CUT&RUN, a pAG-Micrococcal Nuclease (pAG-MNase) fusion is used to cleave antibody-labelled chromatin. Fragments are released from cells and purified (**Figure 1A**). In CUT&Tag, pAG is fused with prokaryotic transposase 5 (pAG-Tn5) to cleave and tagment antibody-bound chromatin with sequencing adapters (**Figure 1B**). Both assays are compatible with next-generation sequencing (NGS) to provide high quality profiles of histone post-translational modifications (PTMs) and chromatin-associated proteins. Robust CUT&RUN and CUT&Tag protocols, based on EpiCypher's validated workflows, are available (epicypher.com/protocols).

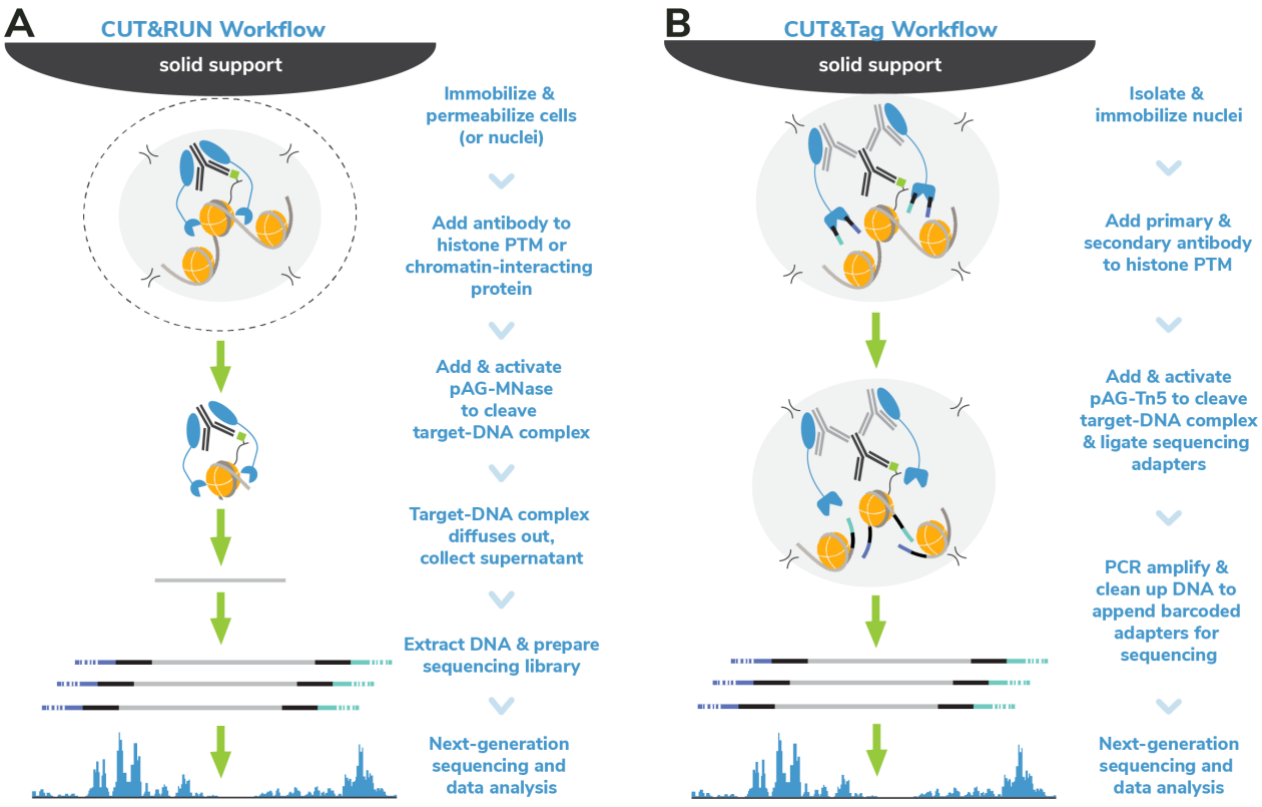

Figure 1: Overview of the CUTANA™ CUT&RUN (**A**) and CUT&Tag (**B**) protocols.

Improved controls for CUT&RUN and CUT&Tag experiments

Compared to Chromatin ImmunoPrecipitation sequencing (ChIP-seq), the leading chromatin mapping assay, CUT&RUN and CUT&Tag consistently generate higher quality data with improved signal-to-noise, while using a fraction of the cellular input and sequencing depth.

But how can users be sure these new assays work as intended? What controls are available?

Typical CUT&RUN/CUT&Tag experiments include positive (e.g. H3K4me3) and negative (e.g. IgG) control antibodies and validated cell lines (e.g. K562 cells). For both assays we also recommend confirming cell viability and bead binding, as well as the size distribution of final NGS libraries (described in the CUTANA™ CUT&RUN and CUT&Tag protocols at epicypher.com/protocols). However, even with such robust controls, questions remain:

- **Is my histone PTM antibody (including H3K4me3 control antibody) specific?** Our previous[4] and ongoing work (chromatinantibodies.com) has shown that PTM antibodies frequently cross-react with related PTM targets and can exhibit application-specific performance. Through our extensive development of CUT&RUN/CUT&Tag assays to various targets, EpiCypher has found that antibodies that work well in ChIP may not always work in CUT&RUN and/or CUT&Tag. In-application testing against a defined panel of on- and off-target nucleosome substrates is the ideal strategy to identify high quality antibodies[4].

- **If my reaction fails or NGS data are of poor quality, how do I troubleshoot my experiment?** CUT&RUN and CUT&Tag methods can be challenging, especially for new users, and it is not always clear which step requires optimization. For instance, poor signal-to-noise in NGS data could be due to failed enzymatic activity, poorly optimized buffer conditions, antibody cross-reactivity, bead clumping, or problems with cell preparation, among others. Defined spike-in controls that replicate physiological chromatin structure (i.e. nucleosomes) can help optimize workflows and guide troubleshooting strategies.

EpiCypher is a leader in the development of semi-synthetic/recombinant nucleosomes and has recently leveraged this technology for the development of SNAP (Sample Normalization and Antibody Profiling) Spike-in Controls (epicypher.com/technologies/snap-spike-in-controls). Our recently launched SNAP-CUTANA™ K-MetStat Panel can be used as a quantitative spike-in control in both CUT&RUN and CUT&Tag reactions targeting histone lysine methylation PTMs.

In this User Guide we describe SNAP-CUTANA™ Spike-ins using the K-MetStat Panel (EpiCypher #19-1002) as our primary example. We review how to incorporate these spike-ins into CUT&RUN/CUT&Tag approaches, data analysis, and interpretations. For additional product information, visit epicypher.com/products/nucleosomes/snap-cutana-spike-in-controls.

2. SNAP-CUTANA™ Spike-in Controls: Overview and Advantages

What are SNAP-CUTANA™ Spike-in Controls, and why should I add these into my CUT&RUN or CUT&Tag workflow?

SNAP-CUTANA™ Spike-in Controls are panels of nucleosomes carrying defined histone PTMs. The spike-ins are distinguished by unique PTM-specific DNA barcodes. Panels are grouped by PTM class, such as that related to histone lysine methylation (K-MetStat Panel; see **Figure 2**) and contain widely studied and disease-relevant modifications.

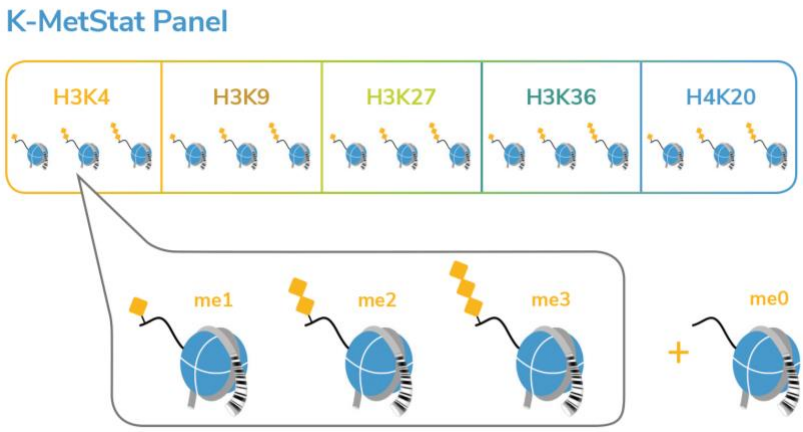

Figure 2: Histone methyl-lysine PTMs in the SNAP-CUTANA™ K-MetStat Panel (EpiCypher 19-1002).

Because the spike-ins represent physiological on- and off-target epitopes and are processed alongside immobilized cells/nuclei in CUT&RUN and CUT&Tag workflows, they provide a powerful, quantitative tool for assay development. This strategy:

- **Provides a direct readout of assay success.** SNAP-CUTANA Spike-ins confirm technical aspects of CUT&RUN and CUT&Tag, including enzymatic activity, DNA purification, and library prep (Section 3.1 and Section 3.2).
- **Guides troubleshooting approaches** by providing a measure of signal-to-noise that can be compared with experimental NGS data to assess experiments (Section 3.2).
- **Enables *in situ* antibody validation** against related PTMs, boosting confidence in NGS data. This is crucial, as many PTM antibodies display cross-reactivity, lot variation, and/or assay-specific performance (Section 3.3).
- **Enables reliable NGS normalization for quantitative cross-sample comparisons.** Still in development; see Section 3.4 for sample applications.

How are the spike-ins detected? What is the "barcode"?

As shown in **Figure 3**, each nucleosome is wrapped with a DNA template containing the Widom 601 nucleosome positioning sequence[5] and a 22 bp barcode unique to each PTM in the panel. Nucleosome assembly DNA is flanked by linker DNA, providing an appropriate substrate for pAG-MNase cleavage or pAG-Tn5 tagmentation in CUT&RUN or CUT&Tag, respectively. The 5' linker DNA also contains a biotin tag (**Figure 3**), allowing spike-ins to be immobilized to streptavidin (SA) magnetic beads for incorporation in CUT&RUN/CUT&Tag workflows alongside cells (or nuclei) immobilized onto ConA magnetic beads.

One important feature of SNAP-CUTANA Spike-ins is that each histone PTM is present in "duplicate," allowing scientists to monitor technical variation of spike-in recovery. This means that for a given panel we prepare each modified nucleosome using two unique DNA-barcoded templates. As an example, the K-MetStat Panel comprises 16 PTM-defined states (**Figure 2**) but contains 32 distinct DNA-barcoded nucleosomes to be monitored by NGS.

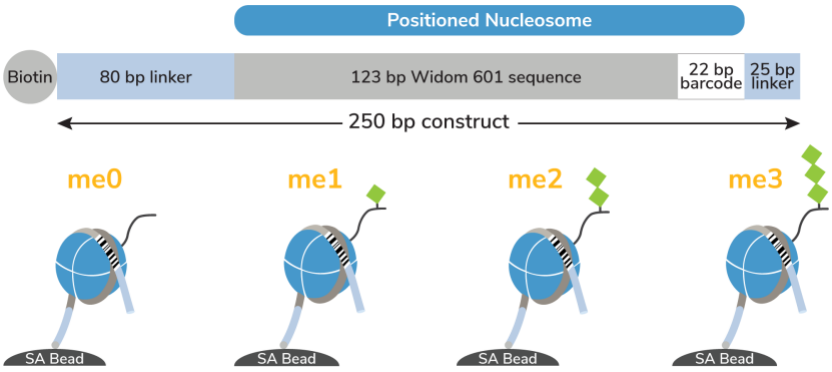

Figure 3: Schematic showing the DNA barcoding and magnetic bead immobilization used in SNAP-CUTANA™ Spike-in Controls.

How are they "spiked" into assays?

SNAP-CUTANA Spike-ins are supplied pre-immobilized to magnetic SA beads (**Figure 3**), allowing them to be processed alongside bead-conjugated cells in CUT&RUN/CUT&Tag. SNAP-CUTANA Spike-ins are added to reactions just prior to antibody addition, in one simple pipetting step (**Figure 4**). The antibody will bind its target epitope in sample cells and the spike-in panel. Addition and activation of pAG-MNase (in CUT&RUN) or pAG-Tn5 (in CUT&Tag) will cleave or tagment the associated spike-in nucleosome, respectively. In both cases, the targeted standard will be processed with sample chromatin through library prep, PCR amplification, and NGS (**Figure 4**). The addition of spike-ins to CUT&RUN/CUT&Tag reactions is described in Section 4.

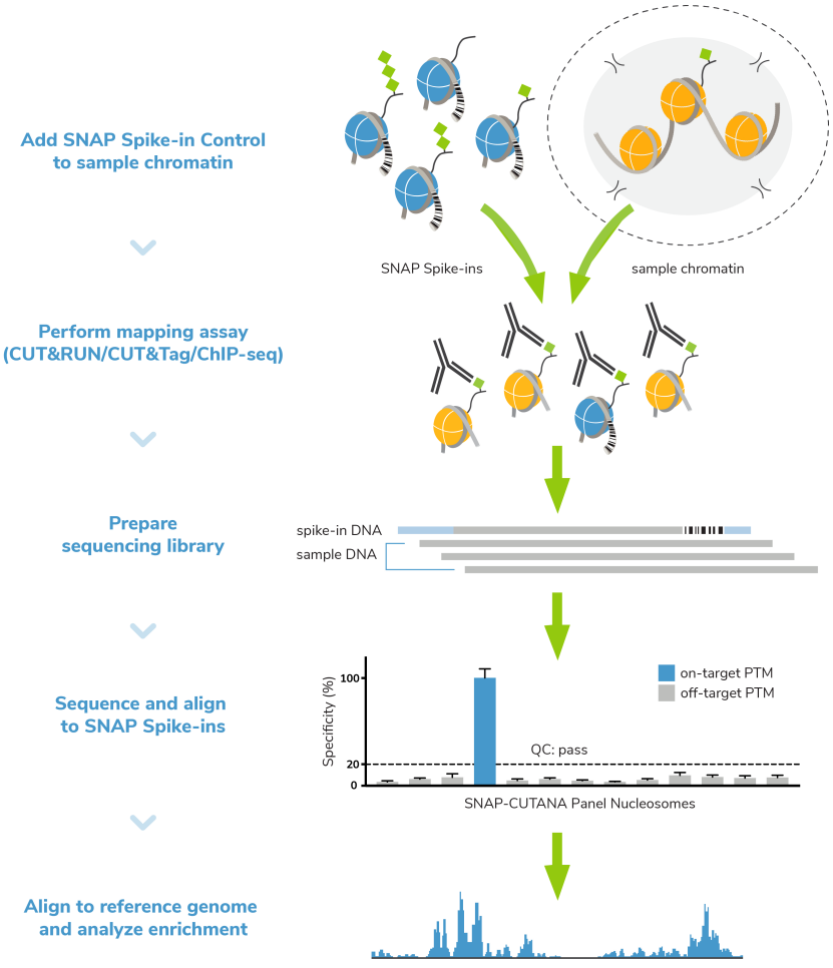

Figure 4: Schematic showing addition of SNAP-CUTANA™ Spike-in Controls during CUT&RUN/CUT&Tag workflows.

3. Applications of SNAP-CUTANA™ Spike-ins for CUT&RUN and CUT&Tag

When should I use SNAP-CUTANA™ Spike-ins for CUT&RUN/CUT&Tag?

Add the K-MetStat Panel to reactions with control antibodies (*e.g.* H3K4me3 and IgG antibodies) in every experiment for:

- Initial protocol optimization and validation for altered workflows (*e.g.* new cell type)
- Continuous monitoring of workflow performance
- Troubleshooting problematic experiments

Add to reactions targeting histone PTMs (*e.g.* in the K-MetStat Panel) to:

- Identify a highly specific antibody
- Validate antibody activity across manufacturing lots

3.1. CUT&RUN/CUT&Tag assay optimization.

When first starting CUT&RUN/CUT&Tag or when trying these approaches under new experimental conditions (new cell types, changing cell number inputs, fixed vs. native preps, etc.) use the SNAP-CUTANA™ K-MetStat Panel with positive (i.e. H3K4me3, epicypher.com/13-0041) and negative (i.e. IgG, epicypher.com/13-0042) control antibodies to validate your CUT&RUN workflow. These antibodies are lot-validated for superior performance and reliability in CUT&RUN and CUT&Tag. EpiCypher includes these positive and negative control antibodies and K-MetStat Panel in every experiment and have found them to be essential controls for initial workflow optimization and daily performance monitoring.

As an example, we used CUT&RUN to map H3K27me3 in K562 cells, with IgG and H3K4me3 antibodies as controls. The K-MetStat Panel (epicypher.com/19-1002) was added to each reaction (**Figure 5**). We first examined relative recovery of each spike-in DNA barcode in NGS data, with recovery of on-target barcode set to 100%. The H3K4me3 control and H3K27me3 antibodies specifically enriched for the spike-in carrying their PTM target (**Figure 5A**; only target PTM is blue). IgG was normalized to the sum total of barcode reads and showed no target preference, confirming multiple technical aspects of the CUT&RUN experiment (e.g. pAG-MNase cleavage conditions). H3K4me3 and IgG antibodies also showed the expected PTM enrichment patterns in NGS (e.g. tight peaks at transcription start sites [TSS], minimal IgG background; **Figure 5B-C**), indicating CUT&RUN conditions were optimized. These robust controls allowed us to have high confidence in our H3K27me3 sequencing results.

Note: Off-target recovery or low signal-to-noise from spike-in data indicates problems with the workflow. Spike-in results can guide troubleshooting experiments (**Table 1** in Section 3.2).

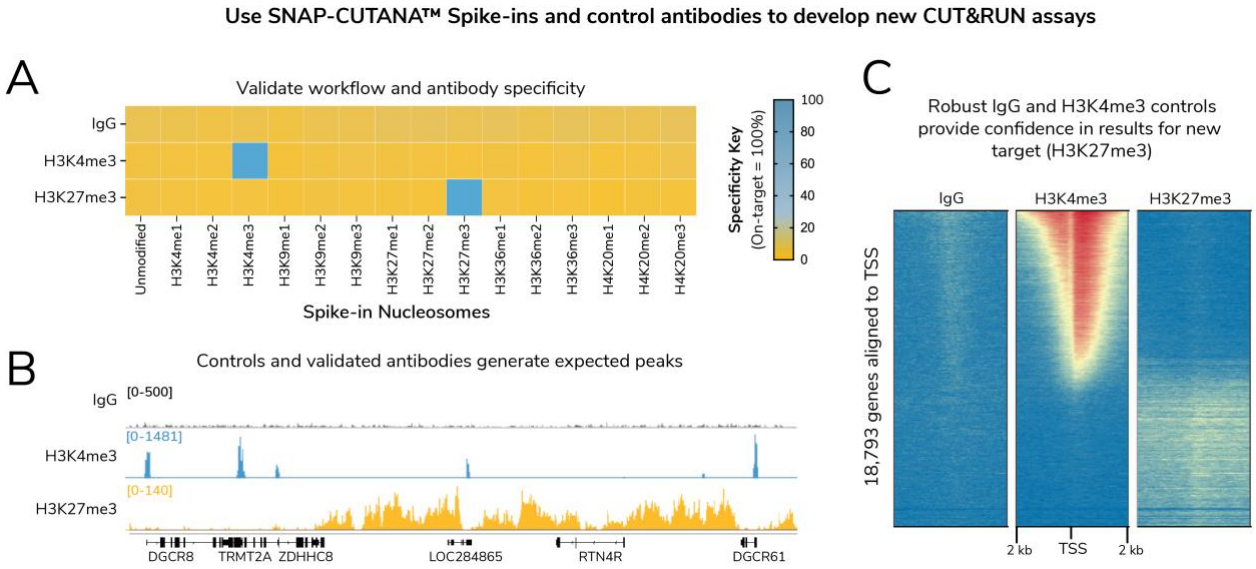

Figure 5: The SNAP-CUTANA™ K-MetStat Panel and control antibodies were used to validate CUT&RUN workflows for K562 cells. (A) PTM antibodies recovered on-target spike-in. Heatmap data normalized to DNA barcodes from on-target PTM; IgG normalized to total reads. (**B,C**) CUT&RUN generated expected H3K4me3 and H3K27me3 maps profiles from K562 sample chromatin. (**B**) RPKM normalized peaks on representative regions. (**C**) Heatmaps display signal intensity +/- 2 kb from transcription start site (TSS). Genes in each heatmap are ordered by signal intensity from H3K4me3 profiles.

3.2. Monitor experimental success.

While EpiCypher CUT&RUN/CUT&Tag protocols are extensively optimized and robust, reactions can occasionally go awry. Adding SNAP-CUTANA™ Spike-in Controls to every reaction identifies problematic experiments and guides troubleshooting. For now, we **minimally** suggest adding the SNAP-CUTANA K-MetStat Panel to reactions designated for positive (*e.g.* H3K4me3) and negative (IgG) control antibodies. We strongly recommend including these controls in every experiment. Here we will discuss how SNAP-CUTANA Spike-ins can be used to identify aberrant CUT&RUN/CUT&Tag reactions (see **Table 1**).

Question: What do successful assays look like using SNAP-CUTANA™ Spike-ins?

In a successful assay, spike-ins reveal specific recovery of the target PTM with minimal background and cross-reactivity, and genomic enrichment patterns appear as expected. In **Figure 6**, CUT&RUN was used to map H3K27me3 in four independently prepared mouse primary B cell samples (optimization experiments with a multi-lab consortium). Control reactions using H3K4me3 and IgG antibodies were included, and the K-MetStat Panel was spiked into each reaction. For Samples 1-3, H3K4me3 positive control and H3K27me3 antibodies displayed high specificity against the K-MetStat Panel (**Figure 6A**), indicated by high recovery of target PTM (blue) and relatively low off-target binding (orange). These results were further reflected in Sample 1-3 sequencing data (**Figure 6B**), which showed expected peak structures with high signal-to-noise (S:N) and low background. NGS results were also consistent across samples.

Question: What do failed assays look like using SNAP-CUTANA™ Spike-ins?

In a failed experiment spike-ins reveal unusual antibody activity, such as cross-reactivity to off-target PTM(s) and/or high background. In contrast to Samples 1-3 (**Figure 6**), Sample 4 displayed uncharacteristic high recovery of multiple off-target PTMs in the K-MetStat panel for both H3K4me3 and H3K27me3 antibodies (**Figure 6A**, red arrows), suggesting high background and low signal-to-noise in these reactions. Sample 4 NGS data exhibited similarly low signal-to-noise in both H3K4me3 and H3K27me3 reactions (**Figure 6B**, red arrows).

Question: How can I use SNAP-CUTANA™ Spike-in results to guide troubleshooting?

When the spike-in controls work as expected, users can trust that their antibody was specific and most CUT&RUN/CUT&Tag experimental conditions are optimal (Samples 1-3). If the spike-in results **and/or** genomic enrichment patterns are not as expected (Sample 4), these data can be used to diagnose and troubleshoot the problem. The K-MetStat Panel provided key insights for assessing failed Sample 4 reactions in **Figure 6**:

- All CUT&RUN reactions were performed in parallel using the same reagents and antibodies, but only Sample 4 displayed problems with S:N and antibody specificity.
- Sample 4 generated poor profiles for multiple targets (H3K4me3 and H3K27me3).
- Low S:N was observed for Sample 4 in both NGS and spike-in data.

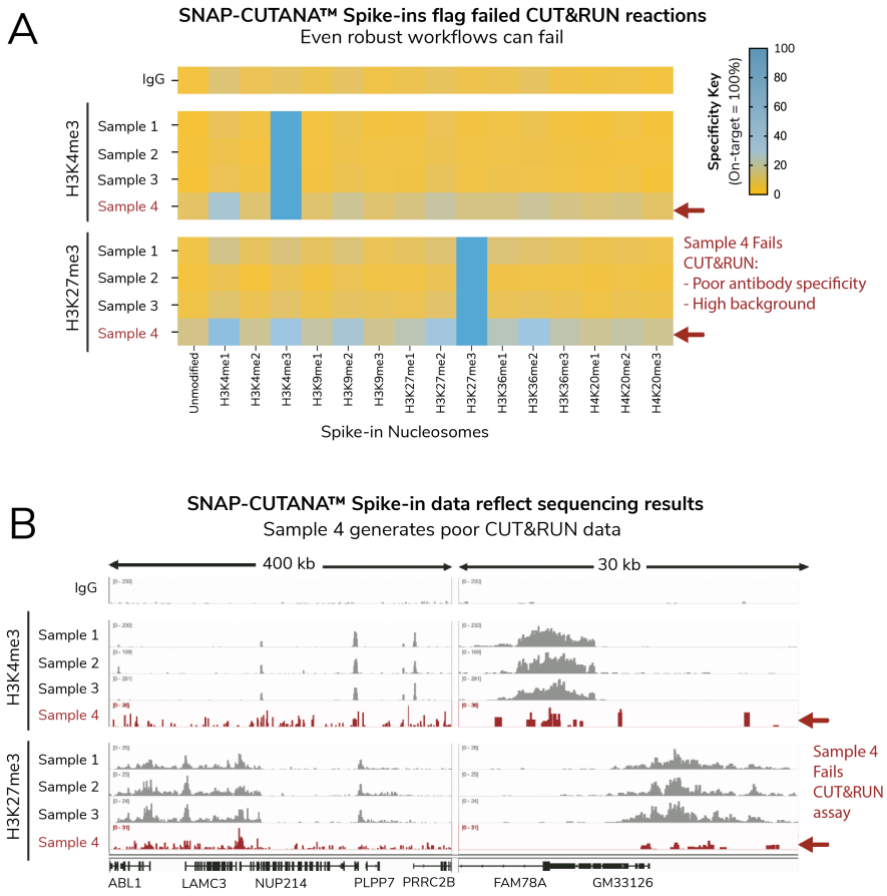

Figure 6: SNAP-CUTANA™ Spike-in Controls identify failed CUT&RUN reactions. CUT&RUN was used to map H3K4me3 and H3K27me3 in four independently prepared mouse primary B cell samples. IgG was included as a negative control. The K-MetStat Panel was spiked into each reaction. **(A)** Heatmaps show recovery of K-MetStat Panel nucleosomes for each reaction, relative to on-target PTM. Sample 4 displays increased cross-reactivity (red arrows). (**B**) RPKM-normalized tracks from sample chromatin shows consistent peaks for Samples 1-3, while Sample 4 displayed low S:N and excessive background (red arrows).

Combined, these results suggest overall problems with the sample prep vs. a CUT&RUN workflow failure. Troubleshooting approaches focused on this aspect (in this case the consortium lab had delivered less than the expected 10,000 cells).

These results illustrate the many advantages of SNAP-CUTANA Spike-ins. It may be unclear from the genome tracks alone that a reaction issue has occurred. **SNAP-CUTANA Spike-ins can flag failed reactions, indicate the cause, and guide troubleshooting**, **improving confidence in experimental results**. **Table 1** (below) outlines how to use the K-MetStat Panel and control antibodies to troubleshoot CUT&RUN and CUT&Tag assays.

Results	Interpretations and troubleshooting approaches
Using H3K4me3 and IgG Control Antibodies: -K-MetStat Panel demonstrates high antibody specificity and high signal-to-noise. -Genomic profiles display low signal-to-noise.	• Enzyme (e.g. pAG-MNase, pAG-Tn5) and wash conditions are optimized. • Perform Quality Control Checks in CUT&RUN/CUT&Tag protocols (epicypher.com/protocols) to confirm cell viability and to avoid cell or bead clumping during assay. Note that low cell viability will increase background. • If using <500,000 cells (CUT&RUN) or <100,000 cells (CUT&Tag), try increasing cell number. • Include control cells (K562 cells) to optimize workflow. • Use native or lightly-crosslinked cells (CUT&RUN and CUT&Tag compatible cross-linking protocol available at epicypher.com/protocols). • Optimize digitonin permeabilization (see CUTANA™ CUT&RUN Kit Manual for detailed instructions; epicypher.com/14-1048).
Using H3K4me3 and IgG Control Antibodies: -Both K-MetStat Panel and genomic profiles display poor signal-to-noise. -The K-MetStat Panel may recover off-target PTMs.	• Indicates a fundamental failure in the workflow. • Carefully re-read the protocol and important notes. • Ensure buffers are prepared FRESH on day of use, and digitonin is added to buffers where indicated. • Perform Quality Control Checks (epicypher.com/protocols) in every experiment
	• Cell integrity may be in question (e.g. Sample 4, **Figure 6**). • Perform Quality Control Checks in CUT&RUN and CUT&Tag protocols (epicypher.com/protocols) to confirm cell viability, number, and binding to beads. • Optimize with 500,000 cells (CUT&RUN) or 100,000 cells (CUT&Tag) before decreasing input.
-H3K4me3 and IgG Control Antibodies generate expected results for the K-MetStat Panel and genomic profiles. -Antibodies to additional targets show no clear enrichment and/or unexpected peaks.	• The workflow is optimized for control antibodies and cell type. • Indicates problem with antibody to experimental target, e.g. poor enrichment efficiency or cross-reactivity. Explore alternative antibodies (shop EpiCypher's validated CUT&RUN antibodies: epicypher.com/cut-and-run-antibodies). • Some targets (e.g. histone acetylation) may benefit from light cross-linking, although yields may be reduced. Cross-linking protocol available at epicypher.com/protocols. • Ensure the target of interest is expressed or localized to chromatin in the study cell/conditions used. Increase cell number or stimulate target expression if needed.

Table 1: Interpreting SNAP-CUTANA™ Spike-in results with genomic NGS results to guide troubleshooting.

Question: Why are SNAP-CUTANA™ Spike-ins better than other methods to identify failed reactions?

Although outliers can be identified using other methods (e.g. separation in Principal Component Analysis [PCA], poor Pearson correlation), SNAP-CUTANA Spike-ins provide an added layer of experimental context. PCA/Pearson are only a proxy for assay stability and cannot determine which part of the experiment is at fault (*e.g.* antibody specificity, cell prep). SNAP-CUTANA Spike-ins provide a direct and quantitative readout of experimental success, providing confidence for researchers to proceed with data analysis and interpretation.

3.3. Identify high quality histone PTM antibodies.

Due to the low background and high sensitivity of CUT&RUN/CUT&Tag approaches, selection of highly specific antibodies is crucial to assay success. EpiCypher is performing extensive antibody testing in epigenomic applications using SNAP Spike-in technology (to date: >1,000 commercial reagents) and has determined that the vast majority (**>70%**) of histone PTM antibodies are nonspecific and/or display poor enrichment[1]. However, high quality antibodies to PTMs for CUT&RUN/CUT&Tag do exist (**Figure 7**). From these studies we have found:

- Antibody capability on modified histone peptide arrays does not transfer to epigenomic assays that consider full nucleosome structure (e.g. CUT&RUN, CUT&Tag, or ChIP-seq).
- Validation of an antibody in one chromatin mapping assay does not guarantee success in another (e.g. ChIP validation does not transfer to CUT&RUN or CUT&Tag).
- Specific PTM recognition under one set of conditions does not guarantee specific PTM recognition in all conditions. Antibodies must always be validated for user-specific conditions with defined nucleosome spike-in controls.

Thus, application-specific testing with PTM-defined, target-representative controls is essential to success. As minimum criteria, EpiCypher recommends selecting antibodies that show <20% cross-reactivity to **each** off-target PTM in the appropriate SNAP-CUTANA™ Panel. These criteria support the accuracy of biological findings based on our prior work[4].

Question: Which histone PTMs can be studied using SNAP-CUTANA™ Spike-ins? Do you have specific antibody recommendations?

The K-MetStat panel has applications to every covered methyl-lysine target (me1-2-3 at H3K4, H3K9, H3K27, H3K36 and H4K20). Test your favorite antibody or inquire at info@epicypher.com for recommendations. EpiCypher also offers CUTANA CUT&RUN Antibodies to Histone PTMs certified using the SNAP-CUTANA K-MetStat Panel (epicypher.com/cut-and-run-antibodies).

Similarly comprehensive panels or targeted reagents are in development for histone lysine acylation (K-AcylStat), ubiquitylation (K-UbStat), and other targets (inquire at info@epicypher.com for specific progress on each).

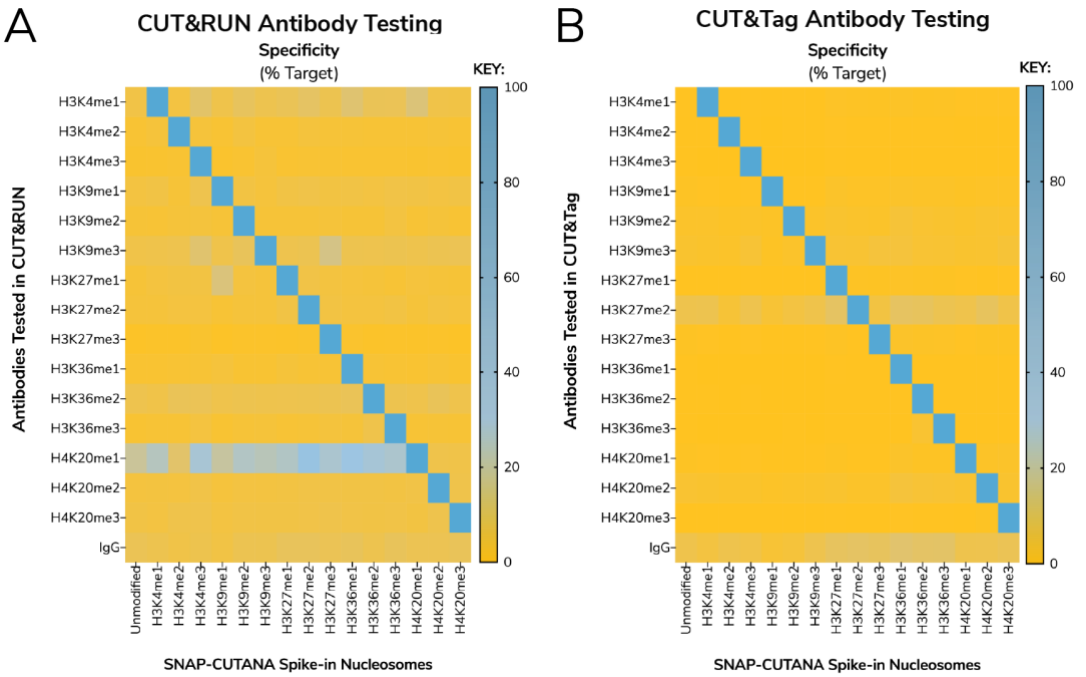

Figure 7: **EpiCypher's best-to-date CUT&RUN (A) and CUT&Tag (B) antibodies to lysine methylation PTMs.** Each row displays SNAP-CUTANA™ K-MetStat Spike-in data for a PTM antibody validated in CUT&RUN (**A**) or CUT&Tag (**B**). Heatmap data for each antibody are normalized to DNA barcodes from the on-target PTM.

Question: How are cross-reactive antibodies detected in SNAP-CUTANA™ Spike-in data?

In **Figure 8** we show specificity profiles for one H3K4me3 and three H3K27me2 antibodies, generated in CUT&RUN reactions spiked with the K-MetStat Panel. Read counts for each DNA-barcoded nucleosome were normalized to the PTM target and used to generate heatmaps (**Figure 8A**). H3K4me3 and H3K27me2 A antibodies showed high specificity for their respective PTM target, with only on-target PTM showing as blue (set to 100% binding) and all other targets showing as yellow in the heatmap (<20% binding relative to target). In contrast, H3K27me2 B and C antibodies displayed substantial off-target binding to H3K4me3 spike-in nucleosome and appear red/purple in the heatmap (>100% binding relative to target). This cross-reactivity was reflected in NGS results (**Figure 8B**). Without the SNAP-CUTANA Spike-ins to flag antibody cross-reactivity, such data could mislead biological conclusions.

Question: Is antibody cross-reactivity from SNAP-CUTANA™ Spike-ins an accurate representation of antibody cross-reactivity in my sample chromatin?

Figure 8B shows heatmaps of CUT&RUN data from K562 cells aligned to gene transcription start sites (TSS, +/- 2 kb). Note the "anti-correlated" enrichment pattern generated from highly specific H3K4me3 and H3K27me2 A antibodies (**Figure 8B**). The red boxes highlight a contaminating, TSS-enriched (i.e. H3K4me3-like) signal using H3K27me2 B and C antibodies that cross-reacted with H3K4me3 (compare with matching spike-in data in **Figure 8A**). These

results demonstrate that SNAP-CUTANA Spike-ins correctly predict antibody specificity and performance in CUT&RUN cell sample data, in agreement with our previous studies using SNAP Spike-ins for ChIP-seq[4]. Thus, including SNAP-CUTANA Spike-ins allows researchers to directly assess the accuracy of NGS data, providing a powerful tool for advanced epigenomics research.

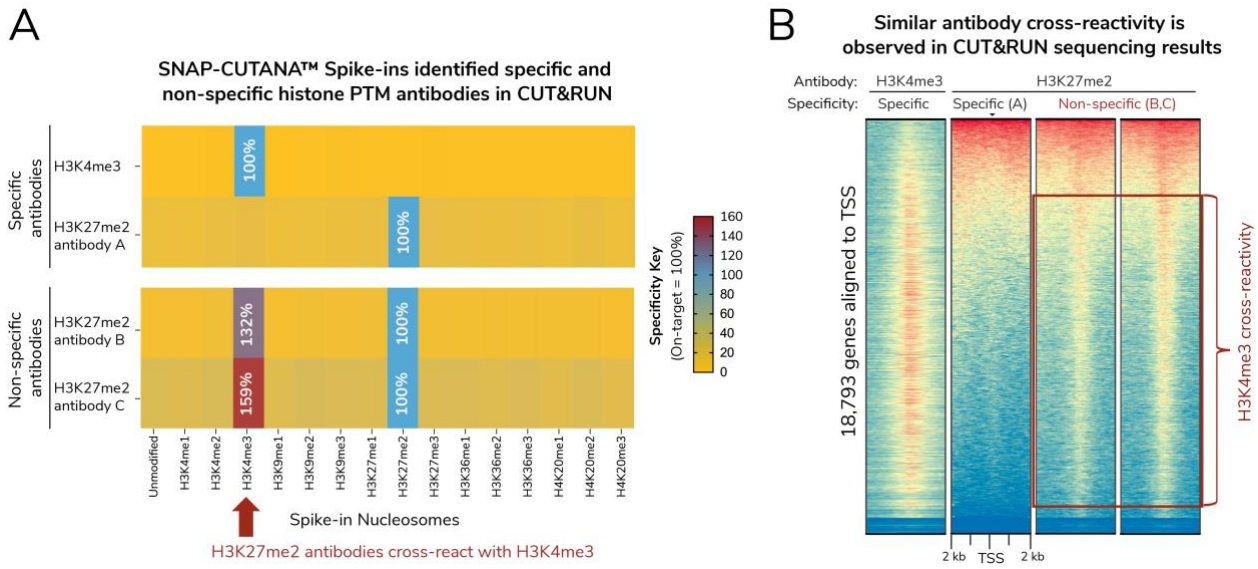

Figure 8: **Antibody cross-reactivity identified by SNAP-CUTANA™ Spike-ins (A) is reflected in CUT&RUN sequencing results (B).** (**A**) SNAP-CUTANA K-MetStat spike-ins were used to characterize H3K4me3 and H3K27me2 antibodies in CUT&RUN workflows using K562 cells. Spike-in data are shown normalized to on-target PTM. (**B**) CUT&RUN data from K562 cells. Heatmaps display signal intensity +/- 2 kb from TSS. Gene rows in each heatmap are linked and ordered by signal intensity sorted by H3K4me3 specific antibody profile.

3.4. Normalize data and quantitatively compare samples.

SNAP-CUTANA™ Spike-ins can also be used to normalize NGS data for reliable, quantitative cross-sample comparisons. EpiCypher is currently investigating multiple normalization methodologies and will share recommendations when available. Below, we describe an emerging application of this exciting technology.

Question: How does normalization to SNAP-CUTANA™ Spike-in Controls compare with other normalization methods?

To demonstrate the application of SNAP-CUTANA Spike-ins for NGS normalization, we show a direct comparison of CUT&Tag data normalization using the SNAP-CUTANA K-MetStat Panel (epicypher.com/19-1002) vs. standard RPKM methods (**Figure 9**). Briefly, CAL27 squamous cell carcinoma cells were treated with either vehicle (0hr) or 3 µM EZH2 methyltransferase inhibitor tazemetostat (24, 48hr). CUT&Tag was performed with an antibody to H3K27me3 (CST #9733) using the EpiCypher CUTANA Direct-to-PCR CUT&Tag Protocol (epicypher.com/protocols). The K-MetStat Panel was added to each sample prior to antibody addition. NGS data were RPKM normalized (top panel) or normalized to the K-MetStat Panel (similar to [6]; bottom panel).

For spike-in normalization, a scale factor was calculated for each sample by dividing the percent of total reads aligned to human genome by the percent of total reads aligned to the spike-in barcodes (Scale Factor = % Human Reads / % Spike-in Reads) and applying this factor to adjust the total sequencing reads of each respective sample.

A very modest difference in H3K27me3 enrichment after EZH2 inhibition is observed in RPKM normalized samples. In contrast, spike-in normalization reveals a pronounced and time-dependent reduction in H3K27me3 enrichment following tazemetostat treatment (**Figure 9**), consistent with immunoblot analysis of total H3K27me3 levels (not shown). Thus, normalization using SNAP-CUTANA Spike-ins can illuminate drug-induced differences in histone PTM enrichment that are not discernable by other approaches.

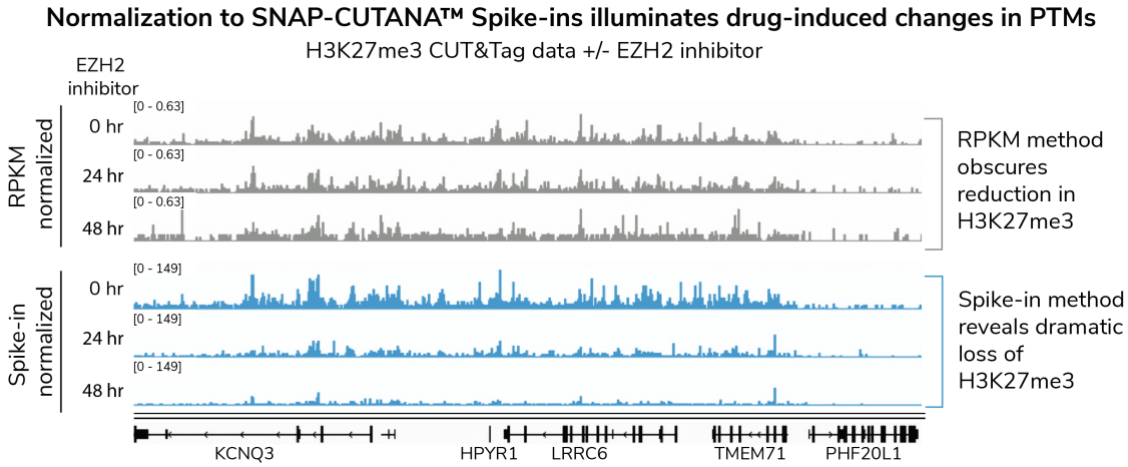

Figure 9: CUT&Tag data normalization using SNAP-CUTANA™ K-MetStat Spike-in Controls illuminates drug-induced differences in histone PTM enrichment. Cancer cells treated with the EZH2 inhibitor tazemetostat for the indicated times show a modest difference in H3K27me3 enrichment compared to vehicle (0hr) when using RPKM normalization. However, spike-in normalization reveals a pronounced and time-dependent reduction in H3K27me3 enrichment following treatment. *Special thanks to Drs. Yinglu Li, Xiao Chen and Chao Lu (Columbia University) for performing the experiment and sharing the data.*

Question: What resources are there for developing normalization strategies using SNAP-CUTANA™ Spike-ins?

Several methods have been reported for experimental normalization using exogenous spike-ins; these approaches can be adopted by calculating a single scalar normalization ratio using total read counts from SNAP-CUTANA Spike-ins.

SNAP-ChIP:

Tay *et al.* Hdac3 is an epigenetic inhibitor of the cytotoxicity program in CD8 T cells. *J. Exp. Med.* 217, e20191453 (2020). (PMID: 32374402)

Lam *et al.* Cell-type-specific genomics reveals histone modification dynamics in mammalian meiosis. *Nat. Commun.* 10, 3821 (2019). (PMID: 31444359)

ChIP-Rx:

Orlando *et al.* Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. *Cell Rep.* 9, 1163-1170 (2014). (PMID: 25437568)

4. Addition of SNAP-CUTANA™ Spike-ins to CUT&RUN and CUT&Tag reactions

Important notes on using SNAP-CUTANA™ Spike-in Panels:

- Store at -20°C. Lower temperatures cause freezing and will permanently damage the magnetic beads.
- Pipette to resuspend beads/nucleosomes before using; DO **NOT** VORTEX.
- The SNAP-CUTANA Panel must be added before addition of primary antibody.
- 50 bp paired-end sequencing is recommended for detection of SNAP-CUTANA barcodes.
- We recommend adding the K-MetStat Panel (epicypher.com/19-1002) to reactions designated for positive (H3K4me3 or H3K27me3) and negative (IgG) control antibodies, in addition to any other reactions targeting a PTM in the K-MetStat Panel (me1, me2, and me3 at H3K4, H3K9, H3K27, H3K36 and H4K20).
- SNAP-CUTANA Spike-ins should comprise ~1% of total sequencing reads. See Section 5 for guidance on how to analyze SNAP-CUTANA Spike-in data from NGS runs.
- The following recommendations are based on our rigorously optimized CUTANA CUT&RUN and CUT&Tag protocols, available at epicypher.com/protocols.

4.1. Adding the K-MetStat Panel to CUT&RUN reactions

Starting from the beginning of Experimental Protocol Section III: Binding of Antibodies.

- For each reaction, resuspend bead-immobilized cells in 50 µL cold Antibody Buffer.
- Add K-MetStat Panel to desired reactions. Amount of panel added is based on the number of cells/nuclei, as outlined in **Table 2**.
  - For reactions using 500,000 cells, add 2 µL K-MetStat Panel directly from the stock.
  - If using less than 500,000 cells, decrease the amount of K-MetStat Panel added linearly by preparing a "working stock" dilution of the panel in Antibody Buffer. General starting recommendations are provided in **Table 2**.
- Proceed with addition of Primary Antibody and overnight incubation.

Starting # Cells in CUT&RUN	Working Stock Dilution in Antibody Buffer [*use FRESH the day of preparation*]	Volume added to reaction	Final dilution in reaction
500,000	Stock	2 µL	1:25
250,000	1:2	2 µL	1:50
100,000	1:5	2 µL	1:125
50,000 or less*	1:10	2 µL	1:250

Table 2: Recommended SNAP-CUTANA™ Spike-in amounts for varying numbers of starting cells in CUT&RUN. ***NOTE: additional dilutions may be used for lower inputs. However, dilution of spike-ins beyond 1:250 increases risk of experimental variation and may require end user optimization.***

4.2. Adding the K-MetStat Panel to CUT&Tag reactions

Starting from the beginning of Experimental Protocol Section III: Binding of Primary and Secondary Antibodies.

- For each reaction, resuspend bead-immobilized nuclei in 50 µL cold Antibody150 Buffer.
- Add K-MetStat Panel to desired reactions. Amount of panel added is based on the number of nuclei, as outlined in **Table 3**.
  - For reactions using 100,000 cells, add 2 µL K-MetStat Panel directly from the stock.
  - If using less than 100,000 cells, decrease the amount of K-MetStat Panel added linearly by preparing a "working stock" dilution of the panel in Antibody150 Buffer. General starting recommendations are provided in **Table 3**.
- Proceed with addition of Primary Antibody and overnight incubation.

Starting # Nuclei In CUT&Tag	Working Stock in Antibody150 Buffer [*use FRESH the day of preparation*]	Volume added to reaction	Final dilution in reaction
100,000	Stock	2 µL	1:25
50,000	1:2	2 µL	1:50
20,000	1:5	2 µL	1:125
10,000 or less	1:10	2 µL	1:250

Table 3: Recommended SNAP-CUTANA™ Spike-in amounts for varying numbers of starting cells in CUT&Tag. ***NOTE: additional dilutions may be used for lower inputs. However, dilution of spike-ins beyond 1:250 increases risk of experimental variation and may require end user optimization.***

5. Analysis of SNAP-CUTANA™ Spike-in Control Sequencing Data

5.1. Count the number of sequencing reads assigned to each nucleosome in the panel

Each spike-in nucleosome contains a unique, PTM-specific barcode that will be detected by NGS. Importantly, each modified nucleosome is assembled using two distinct DNA barcoded templates, providing a technical replicate within each reaction. Thus, for the K-MetStat Spike-in Panel, which contains 16 PTM states, there are 32 DNA barcodes.

These barcode sequences are not contained in mouse, human, fly, or yeast genomes, and thus require an independent alignment procedure. In addition, depending on the directionality of adaptor ligation during library prep, the barcode reads may be contained with R1 reads (adaptor P5) or R2 reads (adaptor P7) from paired-end sequencing. As a result, both R1 and R2 fastq files should be aligned to DNA barcodes.

These steps along with downstream analysis are easily accomplished using the shell script and Excel spreadsheet available on the SNAP-CUTANA Panel product page (e.g. K-MetStat Panel: epicypher.com/19-1002). A shell script is a .sh extension file that can be opened with any basic text editor program (*e.g.* TextEdit on a Mac or any text editing app). It should not be opened or saved as a PDF/Word Doc.

The instructions for aligning sequencing reads to the spike-in DNA barcodes are contained in the shell script, and expanded upon here:

1. Download R1 & R2 paired-end sequencing files (fastq.gz) for reactions that contain spike-ins. Double-click the fastq.gz files to create fastq files and save in a new folder.
2. Navigate to the SNAP-CUTANA Product page at epicypher.com. Scroll down to Documents & Resources and download two files: a SNAP-CUTANA Panel Shell Script (.sh file) and a SNAP-CUTANA Panel Analysis (.xlsx file). Save to your new folder from the previous step.
3. Open the .sh file in TextEdit or any text editing program. Do **NOT** open in Word or in a PDF program. Scroll past the barcode sequences to find the analysis script.
4. The script is a loop that counts the number of reads aligned to each PTM-specific DNA barcode in a reaction.

   You need to customize this script to have one loop per reaction. To customize the script:

   a. Copy the lines **between** # template loop begin ## and # template loop end ## (from the first **echo** to the last **done**)
   b. Paste the loop on the line immediately under the last **done**. Repeat until you have a copy of the loop for each reaction.

   c. In the first loop, replace **sample1_R1.fastq** and **sample1_R2.fastq** with the R1 and R2 files names for **one** reaction. Repeat for each loop/reaction. Save the changes to your .sh file.
6. Open Terminal and change the directory to your new folder from Step 1: Type **cd** and press space once. Drag the folder from your files (i.e. Finder on a Mac) into Terminal to copy the location path. Press return.
7. To execute your shell script in Terminal: type **sh** and press space once. Drag your .sh file into Terminal to copy the file location path. Press Return.
8. Terminal generates read counts from R1 & R2 reads one loop/reaction at a time. For each loop, the script will first generate all barcode read counts from R1 files, and then all barcode read counts from the R2 files. The barcodes will be in the order listed under # Barcode identities in the shell script and results will be annotated by filenames.

Now you can move forward with analysis in Excel.

5.2. Generate a heatmap of the spike-in reads

When the barcode counts generated in Terminal are copied and pasted into the appropriate cells of the Excel spreadsheet, a heatmap of the results (*e.g.* similar to that shown in **Figure 6A**) will automatically be generated. This heatmap will provide information on the success of control reactions and antibody specificity. Instructions on using the Excel template:

9. Open the SNAP-CUTANA Panel Analysis .xlsx file in Excel. The file contains two sheets: the **Panel name** sheet (e.g. K-MetStat), where you paste in data from Terminal, and the **Output Table** which generates the full antibody specificity heatmap.
10. Navigate to the Panel Name sheet, where analysis templates are provided for eight reactions. Copy and paste analysis templates for additional reactions as needed.
11. Type in reaction names and select the On-Target PTM for each using the drop-down menu (Column B). Note that in the K-MetStat Analysis file, the first reaction is pre-set to IgG (negative control), and the second can be set to either H3K4me3 or H3K27me3 (positive control).
12. Copy the R1 barcode read counts from the first loop in Terminal. In Excel, paste into the yellow cells for that reaction in Column C (R1). Copy and paste the R2 barcode read counts from the same loop in Terminal and paste into the adjacent yellow cells in Column D (R2). Repeat for each loop/reaction.
13. The Excel file automatically analyzes spike-in data for each reaction by:
    a. Calculating total read counts (R1 + R2) for each barcode in Column E.
    b. Calculating total barcode read counts (A + B) for each PTM in Column F.
    c. Expressing the total read counts for each PTM as a percentage of on-target PTM read counts (Columns G & J), providing a readout of on- vs. off-target PTM recovery and antibody specificity. Note IgG is auto-normalized to the sum total of barcode reads.

14. Column J populates the Output Table sheet, in which reactions are separated by row and PTM data are sorted into columns. A color gradient is used to visualize recovery of each PTM normalized to on-target PTM, from blue (100%) to orange (below 20%).
15. For each reaction, we also suggest calculating the percentage of unique sequencing reads that have been assigned to the SNAP-CUTANA Panel. Navigate back to the Panel Name Sheet and type the total number of unique reads in the yellow cell in Column B denoted **unique align reads**. The **% total barcode reads** is calculated in the cell immediately below and is added to the Output Table.

5.3. Examine spike-in data

16. Use Output Table to assess antibody binding specificity starting with control reactions. Some guidance on using the controls:
    - EpiCypher IgG control antibody (epicypher.com/13-0042) should not display specific enrichment for any SNAP-CUTANA Spike-in (all boxes orange).
    - K-MetStat Panel: EpiCypher H3K4me3 control antibody (epicypher.com/13-0041) or H3K27me3 antibody (epicypher.com/13-0055) should display strong enrichment for its target (blue), with less than 20% cross-reactivity to off-target PTMs in the panel (orange). Typically our positive control antibodies show less than 10% cross-reactivity.
    - Confirmation of these controls is considered a general readout of workflow success: cells were prepared properly and bound to ConA beads, the enzyme behaved as expected, library prep and sequencing proceeded normally, etc.
17. Once these controls have been confirmed, proceed to analysis of other reactions. Each antibody should display 100% binding to target (blue), and less than 20% binding to off-target PTMs (appears orange).
    - **An antibody with less than 20% binding to all off-target PTMs specific is suitable for downstream data analysis.**
18. Examine the **% total barcode reads** or **spike-in bandwidth** for each reaction. The spike-ins should comprise ~1% of unique sequencing reads for reliable analysis of antibody specificity. However, the range may be higher or lower depending on target abundance and antibody performance. For example:
    - For H3K4me3 (low abundance target in cells), panel barcode reads are typically 1-10% of total sequencing reads.
    - For H3K27me3 (high abundance target in cells), panel barcode reads are typically 0.1-1% of total sequencing reads.
    - For IgG negative control antibody (no target present in sample), panel barcode reads are typically 10-20% of total sequencing reads.

    Outside of this range, consider adjusting the spike-in dilution. The main goal is that thousands of sequencing reads are aligned to the spike-ins for adequate sampling of the panel and reliable use in antibody specificity assessment and data normalization.

6. SNAP-CUTANA™ Spike-in Panel DNA Barcodes

K-MetStat Panel DNA Barcodes are provided in **Table 4**.

Barcodes are also included in the shell script for data analysis on the K-MetStat Panel product page, under the Documents & Resources tab: epicypher.com/19-1002.

SNAP-CUTANA K-MetStat Spike-in Panel: DNA Barcodes		
PTM	**Barcode A (Nuc Replicate 1)**	**Barcode B (Nuc) Replicate 2**
Unmodified	TTCGCGCGTAACGACGTACCGT	CGCGATACGACCGCGTTACGCG
H3K4me1	CGACGTTAACGCGTTTCGTACG	CGCGACTATCGCGCGTAACGCG
H3K4me2	CCGTACGTCGTGTCGAACGACG	CGATACGCGTTGGTACGCGTAA
H3K4me3	TAGTTCGCGACACCGTTCGTCG	TCGACGCGTAAACGGTACGTCG
H3K9me1	TTATCGCGTCGCGACGGACGTA	CGATCGTACGATAGCGTACCGA
H3K9me2	CGCATATCGCGTCGTACGACCG	ACGTTCGACCGCGGTCGTACGA
H3K9me3	ACGATTCGACGATCGTCGACGA	CGATAGTCGCGTCGCACGATCG
H3K27me1	CGCCGATTACGTGTCGCGCGTA	ATCGTACCGCGCGTATCGGTCG
H3K27me2	CGTTCGAACGTTCGTCGACGAT	TCGCGATTACGATGTCGCGCGA
H3K27me3	ACGCGAATCGTCGACGCGTATA	CGCGATATCACTCGACGCGATA
H3K36me1	CGCGAAATTCGTATACGCGTCG	CGCGATCGGTATCGGTACGCGC
H3K36me2	GTGATATCGCGTTAACGTCGCG	TATCGCGCGAAACGACCGTTCG
H3K36me3	CCGCGCGTAATGCGCGACGTTA	CCGCGATACGACTCGTTCGTCG
H4K20me1	GTCGCGAACTATCGTCGATTCG	CCGCGCGTATAGTCCGAGCGTA
H4K20me2	CGATACGCCGATCGATCGTCGG	CCGCGCGATAAGACGCGTAACG
H4K20me3	CGATTCGACGGTCGCGACCGTA	TTTCGACGCGTCGATTCGGCGA

Table 4: SNAP-CUTANA™ K-MetStat Spike-in DNA barcode sequences.

7. References

1 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, doi:10.7554/eLife.21856 (2017).
2 Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, 1930, doi:10.1038/s41467-019-09982-5 (2019).
3 Schmid, M., Durussel, T. & Laemmli, U. K. ChIC and ChEC; genomic mapping of chromatin proteins. *Mol Cell* **16**, 147-157, doi:10.1016/j.molcel.2004.09.007 (2004).
4 Shah, R. N. *et al.* Examining the Roles of H3K4 Methylation States with Systematically Characterized Antibodies. *Mol Cell* **72**, 162-177 e167, doi:10.1016/j.molcel.2018.08.015 (2018).
5 Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**, 19-42, doi:10.1006/jmbi.1997.1494 (1998).
6 Orlando, D. A. *et al.* Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep* **9**, 1163-1170, doi:10.1016/j.celrep.2014.10.018 (2014).