SNAP-CUTANA™ Spike-in User Guide

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

Table of Contents

1. SNAP-CUTANA™ Spike-in Controls

1.1. Background

Typical CUT&RUN and CUT&Tag experiments include positive (*e.g.* H3K4me3 and/or H3K27me3) and negative (*e.g.* IgG) control antibodies. For both assays we also recommend confirming cell viability, bead binding, and size distribution of purified sequencing libraries. However, even with such robust controls, important questions remain: did my reactions work? Are my antibodies specific? How do I troubleshoot experiments?

EpiCypher SNAP (Sample Normalization and Antibody Profiling) Spike-in Controls (epicypher.com/technologies/snap-spike-in-controls) leverages our expertise in the development of semi-synthetic/recombinant nucleosomes. Each SNAP-CUTANA™ Spike-in Panel comprises a pool of DNA-barcoded nucleosomes carrying defined histone PTMs or protein epitopes. The nucleosomes are pre-immobilized on beads, allowing them to be directly added to bead-coupled cells for streamlined CUT&RUN and CUT&Tag workflows (**Figure 1**).

Antibodies are added, which bind their target epitope on cellular chromatin AND in the spike-in panel. Importantly, antibodies distinguish on-target vs. off-target spike-ins; only spike-in nucleosomes containing the target should bind antibody.

Antibody-bound chromatin and spike-in nucleosomes are either released into solution by pAG-MNase (CUT&RUN) or tagmented by pAG-Tn5 (CUT&Tag), followed by next-generation sequencing.

Signal from spike-in nucleosomes is determined from sequencing data via their unique DNA barcode sequences. The spike-in carrying the antibody target should be enriched compared to off-target spike-ins in the panel.

Our SNAP-CUTANA™ K-MetStat Panel can be used as a quantitative spike-in control in CUT&RUN and CUT&Tag reactions targeting histone lysine methylation PTMs. We also offer SNAP Spike-in Controls for CUT&RUN reactions mapping FLAG® and HA-tagged proteins.
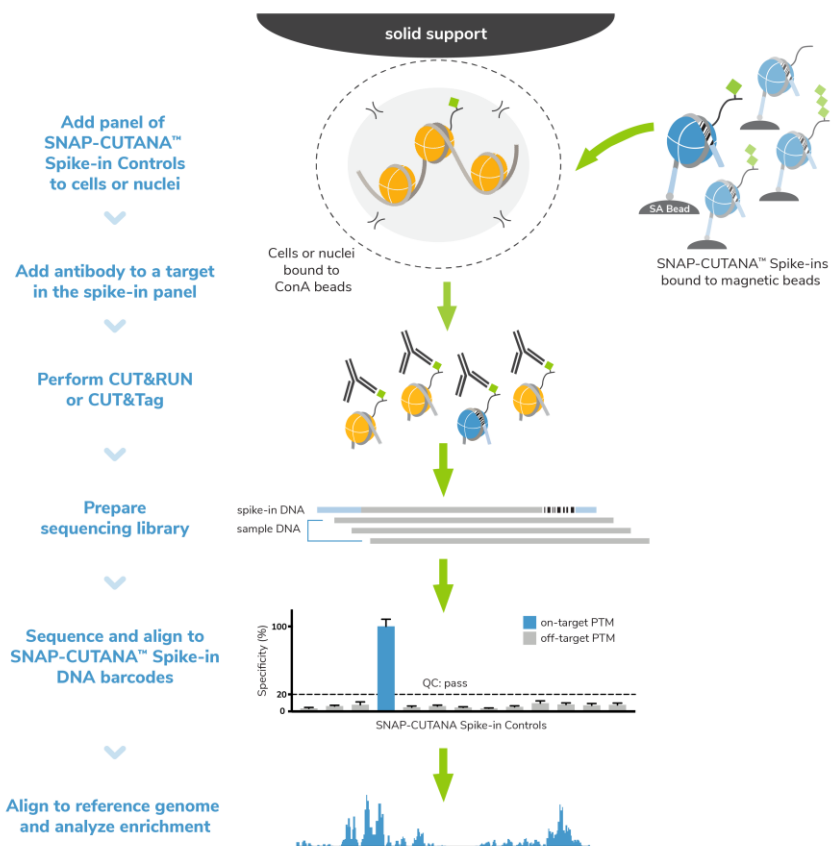

Figure 1: Schematic showing addition of SNAP-CUTANA™ Spike-in Controls during CUT&RUN and CUT&Tag workflows.

FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

1.2. Advantages

- **Fast**: SNAP-CUTANA Spike-ins are added to reactions in one step. The nucleosomes are pooled into rational panels and coupled to magnetic beads for streamlined handling.

- **Easy**: No protocol modifications are necessary. Examine spike-in results in your final sequencing data using our provided analysis tools.

- **Reliable**: SNAP-CUTANA Spike-in data directly report on CUT&RUN and CUT&Tag experimental success. Use results to confirm sample quality and MNase (CUT&RUN) or Tn5 (CUT&Tag) activity, as well as to troubleshoot challenging workflows.

1.3. Applications

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. Address common questions, including:

- **Did my experiment work?** Yields from CUT&RUN and CUT&Tag assays are often quite low (vs. ChIP) due to low cell number requirements and reduced background. In fact, for low-abundance histone PTMs (*e.g.* H3K4me3) or transcription factors, yields are similar to IgG negative control reactions, making it difficult to determine experimental success. For experiments involving drug treatments or transcriptional modulation, target localization may be reduced, which also lowers yields. In each case, SNAP-CUTANA Spike-ins can help confirm CUT&RUN and CUT&Tag assay success.

- **Is my antibody specific for the designated target?** Our previous[1] and ongoing work (chromatinantibodies.com) show that histone PTM antibodies frequently cross-react with related PTM targets and can exhibit application-specific performance, even across different types of chromatin mapping assays. Notably, we found ChIP-validated antibodies are not guaranteed to work in CUT&RUN and/or CUT&Tag. Identification of high-quality antibodies requires assay-specific testing against a defined panel of on- and off-target substrates that mimic physiological chromatin: nucleosomes. SNAP-CUTANA Spike-ins are the **only** control that meet these criteria. SNAP-CUTANA Panels are directly added to CUT&RUN and CUT&Tag reactions and use defined nucleosomes to accurately define antibody specificity to complex chromatin targets[1].

- **How do I troubleshoot failed experiments?** CUT&RUN and CUT&Tag methods can be challenging. For instance, poor signal-to-noise in sequencing data could be due to failed enzymatic activity, poorly optimized buffer conditions, antibody cross-reactivity, bead clumping, or problems with cell preparation, among others. In experiments mapping tagged proteins, loss of genomic signal could be due to protein misfolding or improper localization, all of which are independent of the CUT&RUN and CUT&Tag protocol and antibody performance. SNAP-CUTANA Spike-ins can help optimize workflows and pinpoint areas for troubleshooting.

2. SNAP-CUTANA™ Spike-in Control Protocol

2.1. Overview

- Cells are immobilized on ConA magnetic beads, washed, and aliquoted to 8-strip tubes.

- The SNAP-CUTANA Panel is added to bead-immobilized cells, just prior to addition of primary antibody. The spike-ins are coupled to magnetic beads, allowing them to be processed with your cells, all in the same tube (**Figure 2**).

- Primary antibodies are added, which bind their target in cells AND in the spike-in panel.

- Addition and activation of pAG-MNase (CUT&RUN) or pAG-Tn5 (CUT&Tag) enables selective enrichment of of antibody-bound chromatin AND spike-in nucleosomes.

- Cleaved (CUT&RUN) or tagmented (CUT&Tag) spike-in nucleosomes and cellular chromatin are collected for library preparation, and sequenced.

- SNAP-CUTANA Spike-ins are detected in final sequencing data via their unique DNA barcode sequences. The majority of spike-in barcodes detected should correspond to the antibody target.

- Continue with analysis of sequencing data (*i.e.* align to species reference genome, confirm target enrichment, examine background, etc.).
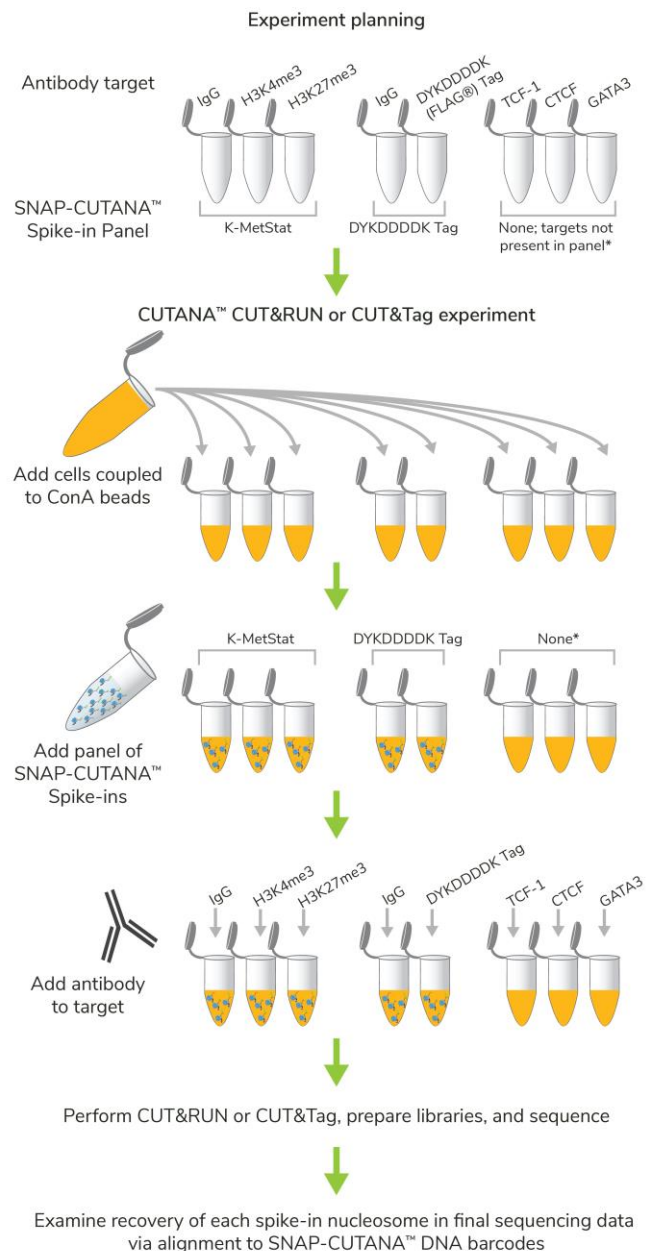

Figure 2: Schematic showing addition of SNAP-CUTANA™ Spike-in Controls during CUT&RUN or CUT&Tag workflows.

*SNAP-CUTANA Controls should only be added when mapping a target included in the spike-in panel. For compatible targets, see Section 4 (K-MetStat Panel) and Section 5 (DYKDDDDK Tag and HA Tag Panels).

FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

2.2. Recommended Products

Available SNAP-CUTANA™ Spike-in Panels

- K-MetStat Panel (epicypher.com/19-1002) for reactions mapping histone lysine methylation PTMs (me1, me2, and me3 at H3K4, H3K9, H3K27, H3K36 and H4K20)

- DYKDDDDK Tag Panel (epicypher.com/19-5001) *only validated for CUT&RUN*

- HA Tag Panel (epicypher.com/19-5002) *only validated for CUT&RUN*

Related Products

- CUTANA™ CUT&RUN Kit (epicypher.com/14-1048)

- CUTANA CUT&Tag Kit (epicypher.com/14-1102)

- HA Tag CUTANA™ CUT&RUN Antibody (epicypher.com/13-2010)

- DYKDDDDK Tag CUTANA™ CUT&RUN Antibody (epicypher.com/13-2031)

- H3K4me3 SNAP-Certified™ Antibody for CUT&RUN (epicypher.com/13-0041)

- H3K27me3 SNAP-Certified™ Antibody for CUT&RUN and CUT&Tag (epicypher.com/13-0055)

- SNAP-Certified™ Antibodies for CUT&RUN (epicypher.com/antibodies-complete-list). If an antibody to your target of interest isn't available, reach out to techsupport@epicypher.com for recommendations.

- CUTANA™ IgG Negative Control Antibody for CUT&RUN and CUT&Tag (epicypher.com/13-0042)

2.3. Important notes – read before using

- Store at -20°C. Lower temperatures can cause freezing and will permanently damage the magnetic beads.

- Do NOT vortex SNAP-CUTANA Spike-ins. The spike-ins are pre-coupled to streptavidin magnetic beads for easy assay incorporation. Vortexing will damage the beads.

- A minimum of 50 bp paired-end reads are required to read the barcoded DNA on SNAP-CUTANA Spike-in Controls.

- SNAP-CUTANA Spike-ins should comprise ~1% of total sequencing reads; adjust spike-in amount as needed. Slightly higher or lower percentages may be acceptable (see Section 3) – the goal is to have sufficient reads for adequate sampling of the spike-in panel without overtaking your sequencing data.

2.4. CUT&RUN spike-in protocol

All SNAP-CUTANA™ Spike-in Panels are compatible with CUT&RUN assays.

SNAP-CUTANA Spike-ins are added to bead-immobilized cells just before addition of primary antibody. Note that the concentration of SNAP-CUTANA Spike-ins is based on the number of cells per reaction and should be adjusted appropriately (**Table 1**).

Steps are outlined below and can also be found in the CUTANA CUT&RUN Kit Manual, Experimental Protocol Section IV: Antibody Binding (epicypher.com/14-1048).

1. Quick spin the SNAP-CUTANA Spike-in Panel to collect liquid in bottom of tube.

2. Gently pipette to resuspend nucleosome/bead mixture - do NOT vortex.

3. Add 2 µL SNAP-CUTANA Spike-in Panel to designated reactions as outlined in **Table 1**. If using fewer than 500,000 cells/reaction, prepare a dilution of the SNAP-CUTANA Panel.

4. Proceed with addition of primary antibody and overnight incubation. No other modifications to the protocol are required.

# Cells	Panel Dilution	Volume per reaction
500,000	Stock	2 µL
250,000	1:2	2 µL
100,000	1:5	2 µL
50,000 or fewer*	1:10	2 µL

Table 1: Scale the amount of SNAP-CUTANA Spike-in Panel to the number of cells. For fewer than 500,000 cells, prepare a working stock dilution of the panel in Antibody Buffer on the day of the experiment.

NOTE: additional dilutions may be used for lower inputs. However, dilution of spike-ins beyond the 1:10 working stock increases risk of experimental variation and may require end user optimization.

2.5. CUT&Tag spike-in protocol

Only the K-MetStat Panel is compatible with CUT&Tag reactions; we have not validated the HA Tag Panel and DYKDDDDK Panel for CUT&Tag. CUT&Tag is generally recommended for histone PTM targets. Chromatin-associated proteins, including those tagged with HA or DYKDDDDK (FLAG®) epitopes, should be mapped using CUT&RUN.

SNAP-CUTANA Spike-ins are added to bead-immobilized nuclei, just before addition of primary antibody. Note that the concentration of SNAP-CUTANA Spike-ins is based on the number of cells per reaction and should be adjusted appropriately (**Table 2**).

Steps are outlined below and can also be found in the CUTANA CUT&Tag Kit Manual, Experimental Protocol Section IV: Antibody Binding (epicypher.com/14-1102).

1. Quick spin the SNAP-CUTANA Spike-in Panel to collect liquid in bottom of tube.

2. Gently pipette to resuspend nucleosome/bead mixture - do NOT vortex.

3. Add 2 µL SNAP-CUTANA Spike-in Panel to designated reactions as outlined in **Table 1**. If using fewer than 100,000 nuclei/reaction, prepare a dilution of the SNAP-CUTANA Panel.

4. Proceed with addition of primary antibody and overnight incubation. No other modifications to the protocol are required.

# Cells	Panel Dilution	Volume per reaction
100,000	Stock	2 µL
50,000	1:2	2 µL
20,000	1:5	2 µL
10,000 or fewer*	1:10	2 µL

Table 2: Scale the amount of SNAP-CUTANA Spike-in Panel to the number of cells. For fewer than 100,000 nuclei, prepare a working stock dilution of the panel in Antibody Buffer on the day of the experiment.

NOTE: additional dilutions may be used for lower inputs. However, dilution of spike-ins beyond a 1:10 working stock increases risk of experimental variation and may require end user optimization.

FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

3. Analysis of SNAP-CUTANA™ Spike-in Controls from sequencing data

3.1. Overview

- SNAP-CUTANA™ Spike-in Panels are pools of modified nucleosomes.

- Each modified nucleosome in a SNAP-CUTANA Panel is prepared with two distinct DNA barcoded sequences, thus providing a technical replicate within each reaction. For an explanation of technical replicates, see Section 6.6.

- The DNA barcode sequences are unique to each epitope in the SNAP-CUTANA Panel. Analysis of barcodes provides a direct readout of spike-in recovery and enrichment.

3.2. Required files

- R1 and R2 paired-end sequencing files (fastq.gz)

- The **shell script** and **Panel Analysis Spreadsheet** available on the SNAP-CUTANA Panel web page, under "Documents and Resources":

    o K-MetStat Panel (epicypher.com/19-1002)

    o DYKDDDDK (FLAG®) Tag Panel (epicypher.com/19-5001)

    o HA Tag Panel (epicypher.com/19-5002)

3.3. Important notes – read before analysis

- The spike-in DNA barcode sequences are unique and will **NOT** align to genomes from mouse, human, fly, yeast, or other model organisms. This protocol is required for detection and analysis of spike-in nucleosomes!

- The percentage of total sequencing reads assigned to spike-in nucleosomes will vary by target abundance. Low abundance targets and negative controls will have a greater percent of total reads from spike-ins (IgG or H3K4me3, ~10-20%), while high abundance targets show lower percent of reads from spike-ins (H3K27me3, 0.1-1%). The goal is that thousands of sequencing reads are aligned to the spike-ins for adequate sampling of the spike-in panel without overtaking your sequencing data.

FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

3.4. Detection of spike-ins from CUT&RUN and CUT&Tag sequencing data

Use the provided shell script to count the number of sequencing reads assigned to each DNA barcode sequence in your SNAP-CUTANA™ Spike-in Panel. Instructions for aligning sequencing reads to the spike-in DNA barcodes are contained in the shell script and outlined below.

1. Download R1 and R2 paired-end sequencing files (fastq.gz) for reactions containing SNAP-CUTANA Spike-ins. Double-click fastq.gz files to create fastq files. Save in a new Folder.

2. Download the accompanying Shell Script (.sh) and Panel Analysis Spreadsheet (.xlsx) from the SNAP-CUTANA Panel web page (see above). Save both files to the Folder in Step 1.

3. Open the Shell Script (.sh) file in TextEdit or any text editing program. Do NOT open in Word or a PDF program. Do NOT save as a PDF.

4. Scroll past instructions and DNA barcode sequences to find the analysis script. The script is a "template loop" that counts the number of reads assigned to each DNA barcode sequence in the SNAP-CUTANA Panel. You need one loop per reaction. To customize:

    a. Copy lines between `# template loop begin ##` and `# template loop end ##` (from the first `echo` to the last `done`).

    b. Under the last `done` paste the loop. You will need one copy per reaction.

    c. In the first loop replace `sample1_R1.fastq` and `sample1_R2.fastq` with R1 & R2 fastq file names for one reaction. Repeat for each loop. Click Save.

5. Open Terminal and set the directory to your Folder from Step 1: Type `cd` and press space. Drag the Folder from your files into Terminal to copy the location. Press return.

6. Run your script in Terminal: Type `sh` and press space. Drag your .sh file from your Folder into Terminal to copy the file location. Press return. Terminal generates barcode read counts from R1 & R2 reads, one loop – or one reaction – at a time. Datasets will be annotated based on the filenames.

3.5. Analysis of spike-in data

When the barcode counts generated in Terminal are copied and pasted into the appropriate cells of the Excel file, a heatmap of the results will automatically be generated. This heatmap will help determine experimental success and antibody specificity.

7. Open the Panel Analysis Spreadsheet (.xlsx) file in Excel and navigate to the first tab.

8. In Column B, fill in reaction names and antibody targets. The first reaction is always set to IgG (negative control); for other reactions, set the target to the appropriate histone PTM or protein epitope.

9. Copy R1 barcode read counts from the first loop in Terminal. In Excel, paste into the corresponding yellow cells for that reaction in Column C. Copy and paste the R2 read counts from the same loop to yellow cells in Column D. Repeat for each loop/reaction.

10. The Excel file automatically analyzes SNAP-CUTANA Spike-in data. For each CUT&RUN or CUT&Tag reaction, it analyzes recovery of each member of the spike-in panel by:

    a. Calculating total barcode read counts for each histone PTM or protein epitope in the Panel (Column F).

    b. Expressing total read counts for each member of the Panel as a percentage of on-target read counts (Columns G & J), providing a readout of on- vs. off-target recovery and antibody specificity. Note that for IgG, binding is normalized to total barcode reads.

11. Column J auto-populates the Output Table in the second tab of the workbook. Reactions are separated by row and spike-in data are sorted into columns. A color gradient is used to visualize the recovery of each spike-in normalized to the target epitope, from blue (100%) to orange (less than 20%).

12. For each reaction, calculate the percent of unique sequencing reads that have been assigned to spike-ins. In Excel, type the total number of unique reads in the yellow cell Uniq align reads (in Column B). The % total barcode reads is calculated in the cell immediately below and is added to the Output Table.

13. View expected results and interpretations for the K-MetStat Panel, HA Tag Panel, and DYKDDDDK Tag Panel.

4. K-MetStat Panel

4.1. Recommendations

- The K-MetStat Panel (epicypher.com/19-1002) can be added to CUT&RUN or CUT&Tag reactions mapping a histone PTM represented in the panel (**Figure 3**).

- In every experiment, include positive control reactions with H3K4me3 (epicypher.com/13-0041) and H3K27me3 (epicypher.com/13-0055) antibodies, spiked with the SNAP-CUTANA K-MetStat Panel (epicypher.com/19-1002).

- Include a negative control reaction in every experiment; use EpiCypher IgG Negative Control antibody (epicypher.com/13-0042).

- Use to screen antibodies to histone methyl-lysine PTMs and to optimize, validate, or troubleshoot workflows. See Section 6.
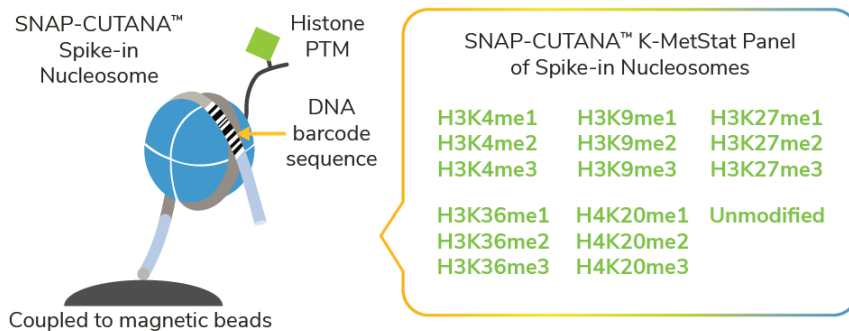

Figure 3: The SNAP-CUTANA™ K-MetStat Panel. The K-MetStat Panel contains 16 different nucleosomes, representing 15 lysine methylation states and an unmodified control. The Panel is coupled to magnetic beads for streamlined integration into CUT&RUN and CUT&Tag workflows and can be added to any reaction mapping a histone PTM in the panel. DNA barcodes correspond to each nucleosome in the panel and are used for detection in sequencing data.

4.2. Expected results

Download the K-MetStat Panel **shell script** and **Panel Analysis Spreadsheet** from epicypher.com/19-1002, under "Documents and Resources." Follow instructions in Section 3.3 to detect and count the number of spike-in reads from sequencing data. Once you have imported spike-in data into the Excel spreadsheet, assess as follows.

1. IgG Negative Control Antibody should show no specific enrichment for any SNAP-CUTANA Spike-in. In the heatmap, signal for each nucleosome is normalized to the total number of spike-in reads from the IgG reaction (all boxes orange; **Figure 4A**).

2. Positive control antibodies to H3K4me3 or H3K27me3 should show strong enrichment for antibody target. In the heatmap, signal for each nucleosome is normalized the on-target PTM (blue; **Figure 4A**); off-target PTMs should show < 20% recovery (orange; **Figure 4A**).

3. Spike-ins should comprise ~1% (0.5-5%) of total sequencing reads. Low abundance targets and negative controls have a greater percentage of total reads from spike-ins (IgG, H3K4me3 ~10-20%), while high abundance targets show a lower percentage of reads from spike-ins (H3K27me3, 0.1-1%). The goal is to have sufficient reads for adequate sampling of the spike-in panel without overtaking your sequencing data.

4. Align sequencing reads to the appropriate reference genome. Examine genomic sequencing tracks and confirm proper target enrichment and low background (**Figure 4B & 4C**). Confirmation of positive and negative control reactions is considered a general readout of workflow success: cells were prepared properly, pAG-MNase or pAG-Tn5 enzyme behaved as expected, library preparation and sequencing proceeded normally, etc.

5. Once controls are confirmed, proceed to analysis of other reactions spiked with the K-MetStat Panel. Each antibody should display 100% binding to the designated on-target PTM (blue), and <20% binding to off-target PTMs (appears orange). Examine sequencing tracks from genomic alignments and assess target enrichment.

What if something goes wrong? Enrichment of off-target spike-in nucleosomes and/or poor signal from genomic analysis can be used to guide troubleshooting, getting you back on the path to experimental success. See Section 6.1 and Section 6.2 for guidance.

> **EpiCypher considers a histone PTM antibody with less than 20% binding to all off-target nucleosomes specific and suitable for downstream data analysis.**
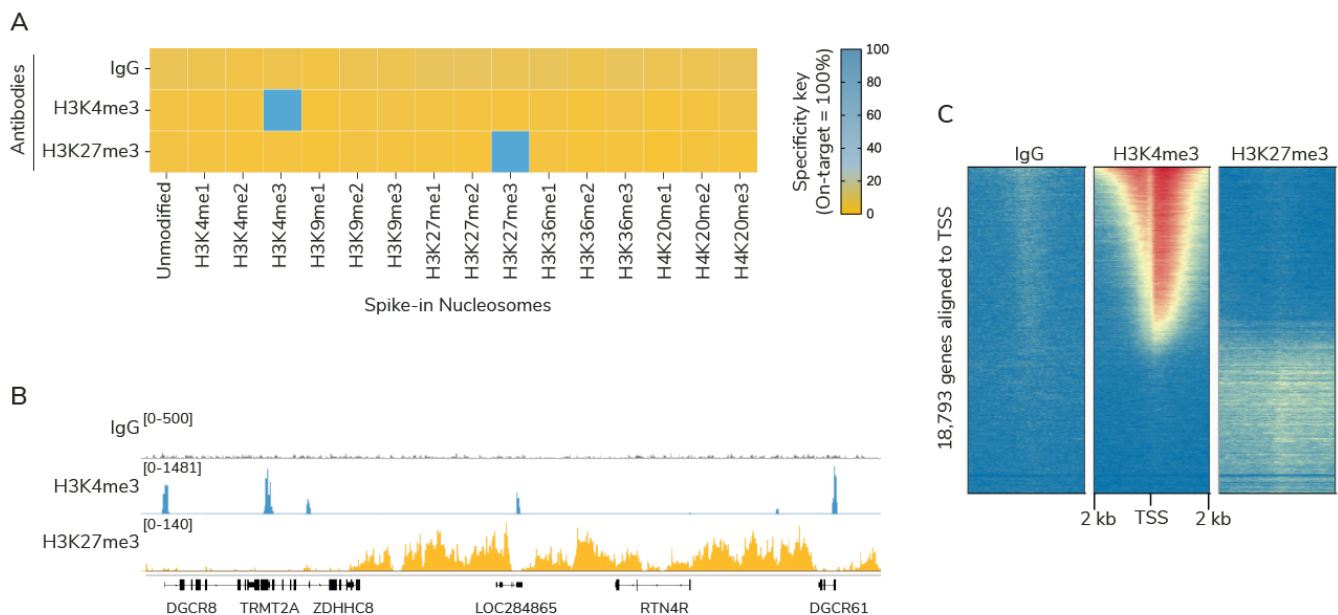

Figure 4: The SNAP-CUTANA™ K-MetStat Panel and control antibodies were used to validate CUT&RUN workflows for K562 cells. (A) PTM antibodies recovered on-target nucleosomes from the K-MetStat Panel. H3K4me3 and H3K27me3 data are normalized to signal from respective on-target PTM; IgG is normalized to total reads from spike-ins. **(B, C)** CUT&RUN generated expected H3K4me3 and H3K27me3 profiles from K562 cells: IgG shows low background, H3K4me3 is localized in sharp peaks at active transcription start sites (TSSs), and H3K27me3 is broadly distributed over repressed regions. **(B)** RPKM normalized peaks at representative regions show high signal over background and the expected peak structures. **(C)** Heatmaps display signal intensity at 18,793 annotated TSSs. Genes in each heatmap are ordered by signal intensity from H3K4me3 profiles, which shows oppositional enrichment to H3K27me3 as expected.

4.3. DNA barcode sequences

PTM	Barcode A (Replicate 1)	Barcode B (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 3: DNA Barcode sequences for nucleosomes in the SNAP-CUTANA™ K-MetStat Panel.

5. DYKDDDDK Tag and HA Tag Panels

5.1. Recommendations

- The DYKDDDDK (FLAG®) Tag Panel (epicypher.com/19-5001) and HA Tag Panel (epicypher.com/19-5002), outlined in **Figure 5**, are **only validated for CUT&RUN.**

  - **Why can't I use these panels in CUT&Tag?** CUT&Tag is generally recommended for histone PTM targets. Chromatin-associated proteins, including those tagged with HA or FLAG® epitopes, should be mapped using CUT&RUN.

- For best results, pair with DYKDDDDK Tag CUTANA™ CUT&RUN Antibody (epicypher.com/13-2031) and HA Tag CUTANA™ CUT&RUN Antibody (epicypher.com/13-2010), respectively. However, these panels can be used with any preferred DYKDDDDK or HA antibody.

- Include a negative control reaction with epitope tag spike-ins in every experiment; use EpiCypher IgG Negative Control antibody (epicypher.com/13-0042).

- For every experiment, EpiCypher recommends positive control reactions with H3K4me3 (epicypher.com/13-0041) and H3K27me3 (epicypher.com/13-0055) antibodies, spiked with the K-MetStat Panel (epicypher.com/19-1002) – see Section 4.

- Use to spike-in results to optimize, validate, or troubleshoot workflows. See Section 6.
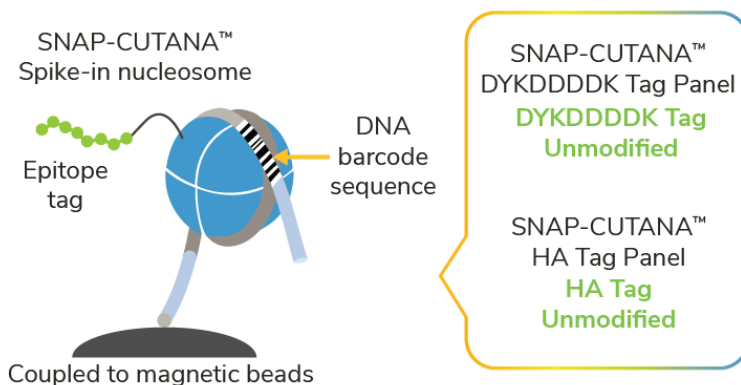

Figure 5: SNAP-CUTANA™ DYKDDDDK Tag and HA Tag Panels for CUT&RUN profiling of epitope-tagged proteins. Each panel contains an epitope-tagged nucleosome (HA or DYKDDDDK) and an unmodified nucleosome. Panels are coupled to magnetic beads for streamlined integration into CUT&RUN workflows. DNA barcodes correspond to nucleosomes in the respective panel and are used to detect spike-ins from sequencing data.

FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

5.2. Expected results

Download the Panel **shell script** and **Panel Analysis Spreadsheet** from epicypher.com/19-5001 (DYKDDDDK Tag Panel) or epicypher.com/19-5002 (HA Tag Panel), under "Documents and Resources." Follow instructions in Section 3.3 to detect and count the number of spike-in reads from sequencing data.

Assess SNAP-CUTANA Spike-in data, starting with IgG negative control reactions, and then moving to reactions mapping epitope-tagged protein. **Figure 6** shows example data from CUT&RUN reactions mapping FLAG-tagged GATA3 in MDA-MB-231 cells[2].

1. IgG Negative Control reactions should show no specific enrichment. In the heatmap, signal for each nucleosome is normalized to the total number of spike-in reads from the IgG reaction. Because there are only two nucleosomes, the normalized signal should be ~50% for unmodified and tagged nucleosomes (light blue in heatmap; **Figure 6A**).

   Note that negative control heatmap data look different for DYKDDDDK and HA Tag Panels compared to the K-MetStat Panel (**Figure 4A** vs. **Figure 6A**). This is due to the number of nucleosomes in each panel. IgG reactions with the K-MetStat Panel divide nonspecific signal across 16 nucleosomes, resulting in < 20% signal for each, while the Tag Panels divide signal across two nucleosomes, resulting in ~50% signal for each.

2. EpiCypher antibodies to HA or FLAG® epitopes should show strong enrichment for HA or FLAG-tagged proteins. In the heatmap, signal for each nucleosome is normalized the on-target PTM (blue; **Figure 6A**); unmodified nucleosomes should show < 20% recovery (orange; **Figure 6A**).

3. Spike-ins should comprise ~1% (0.5-5%) of total sequencing reads, although this may vary by protein abundance. Transcription factors and IgG negative controls often have a greater percentage of total reads from spike-ins (~10-20%). The goal is to have sufficient reads for adequate sampling of the spike-in panel without overtaking your sequencing data.

4. Align sequencing reads to the appropriate reference genome. Examine genomic sequencing tracks and confirm proper target enrichment and low background (**Figure 6B & 6C**). Confirmation of IgG negative control and Tag antibody are considered a general readout of CUT&RUN workflow success: cells were prepared properly, antibody selectively bound the epitope tagged protein, pAG-MNase cleaved antibody-bound nucleosomes, library preparation and sequencing proceeded normally, etc.

What if something goes wrong? Enrichment of unmodified spike-in nucleosomes and/or poor signal from genomic analysis can be used to guide troubleshooting, getting you back on the path to experimental success. See Section 6.1 and Section 6.2 for guidance.

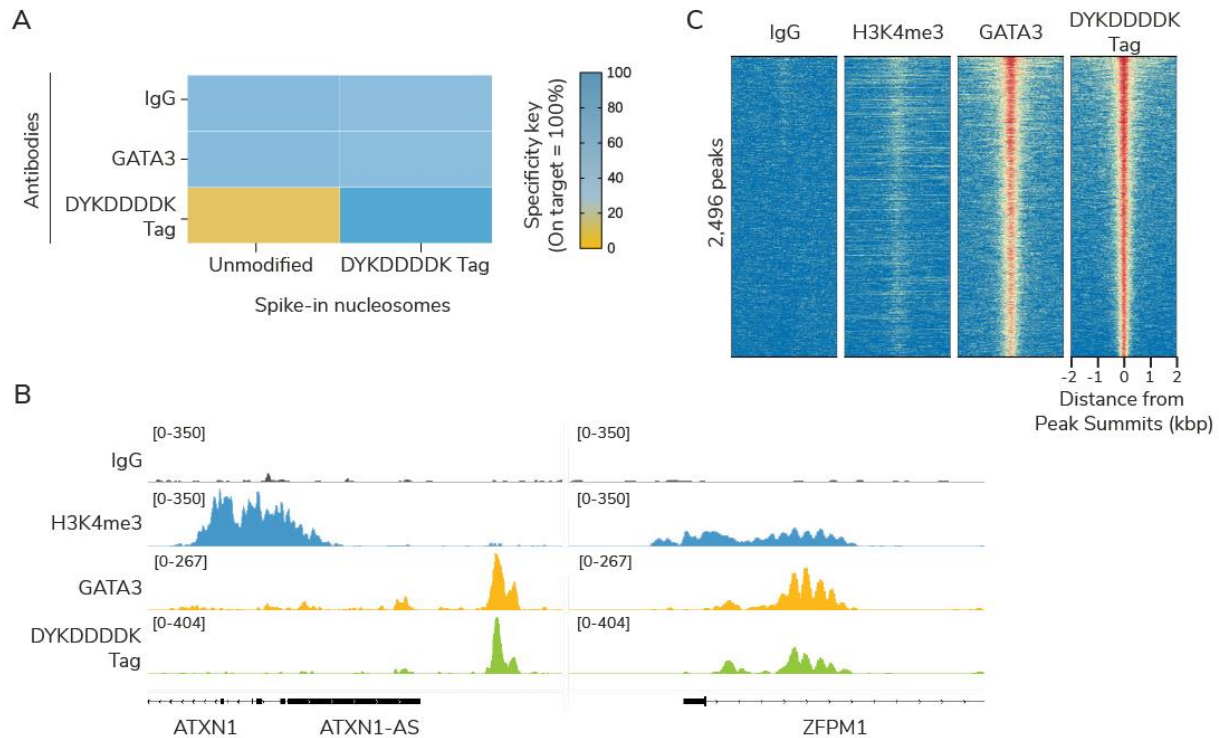FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

Figure 6: The SNAP-CUTANA™ DYKDDDDK (FLAG®) Tag Panel validates CUT&RUN workflows for FLAG-tagged GATA3 in human breast cancer cells. CUT&RUN was performed using MDA-MB-231 cells[2] and antibodies to IgG (EpiCypher 13-0042), GATA3 (CST-5852), DYKDDDDK Tag (EpiCypher 2031), and H3K4me3 (EpiCypher 13-0041). Reactions to IgG, GATA3, and FLAG® were spiked with the DYKDDDDK Tag Panel. **A)** DYKDDDDK antibodies specifically recovered tagged nucleosomes vs. unmodified control. IgG and GATA3 antibodies did not show specific enrichment of spike-ins, as expected. (**B,C**) GATA3 profiles generated with DYKDDDDK Tag antibodies and GATA3 antibodies were nearly identical, validating enrichment of the FLAG-tagged GATA3 in MDA-MB-231 cells. H3K4me3 was included as a positive control. (**B**) RPKM normalized peaks at representative regions. (**C**) Heatmaps display signal intensity +/- 2 kbp from DYKDDDDK Tag peaks. Genes in each heatmap are ordered by signal intensity from DYKDDDDK Tag profiles.

FLAG is a registered trademark of Merck KGaA, Darmstadt, Germany.

5.3. DNA barcode sequences

Epitope	Barcode A (Replicate 1)	Barcode B (Replicate 2)
Unmodified	ACCGCGATACGCGCGCGATATG	ACGACGTACGGCGCGCTATACG
DYKDDDDK	ATCGTACGCCGCGTCGCTCGAA	TAGCGTACCGACGATACGCGTT

Table 4: DNA barcode sequences for nucleosomes in the SNAP-CUTANA™ DYKDDDDK Tag Panel.

Epitope	Barcode A (Replicate 1)	Barcode B (Replicate 2)
Unmodified	ACCGCGATACGCGCGCGATATG	ACGACGTACGGCGCGCTATACG
HA	ACGCGTCGCTACGCGTATCGGT	ACGGTCGTCGACGCGTATTCGG

Table 5: DNA barcode sequences for nucleosomes in the SNAP-CUTANA™ HA Tag Panel.

6. SNAP-CUTANA™ Spike-in Control Applications

6.1. Determining assay success vs. failure

In a successful assay, SNAP-CUTANA Spike-ins reveal specific recovery of the target PTM with minimal background and cross-reactivity and genomic enrichment patterns appear as expected. In a failed experiment, spike-ins reveal unusual antibody activity, such as cross-reactivity to off-target PTM(s) and/or high background. Below, we provide examples of how SNAP-CUTANA Spike-ins can be used to determine assay success and failure and how these results can guide troubleshooting.

In **Figure 7** (next page), 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 7A**), 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 7B**), which showed expected peak structures with high signal-to-noise and low background. Sequencing results were also consistent across samples.

In contrast to Samples 1-3 (**Figure 7**), Sample 4 displayed uncharacteristic high recovery of multiple off-target PTMs in the K-MetStat panel for both H3K4me3 and H3K27me3 antibodies (**Figure 7A**, red arrows), suggesting high background and low signal-to-noise in these reactions. Sample 4 sequencing data exhibited similarly low signal-to-noise in both H3K4me3 and H3K27me3 reactions (**Figure 7B**, red arrows).

6.2. How to troubleshoot failed reactions

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, **Figure 7**). 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 7**:

- All CUT&RUN reactions were performed in parallel using the same reagents and antibodies, but only Sample 4 displayed problems with signal-to-noise and antibody specificity.
- Both H3K4me3 and H3K27me3 antibodies displayed cross-reactivity in Sample 4.
- Low signal-to-noise was observed for Sample 4 in both sequencing and spike-in data.

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

These results illustrate the utility of including SNAP-CUTANA Spike-ins in <u>all experiments</u>. 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 6** (below) outlines how to use the SNAP-CUTANA Panels to troubleshoot CUT&RUN and CUT&Tag assays.
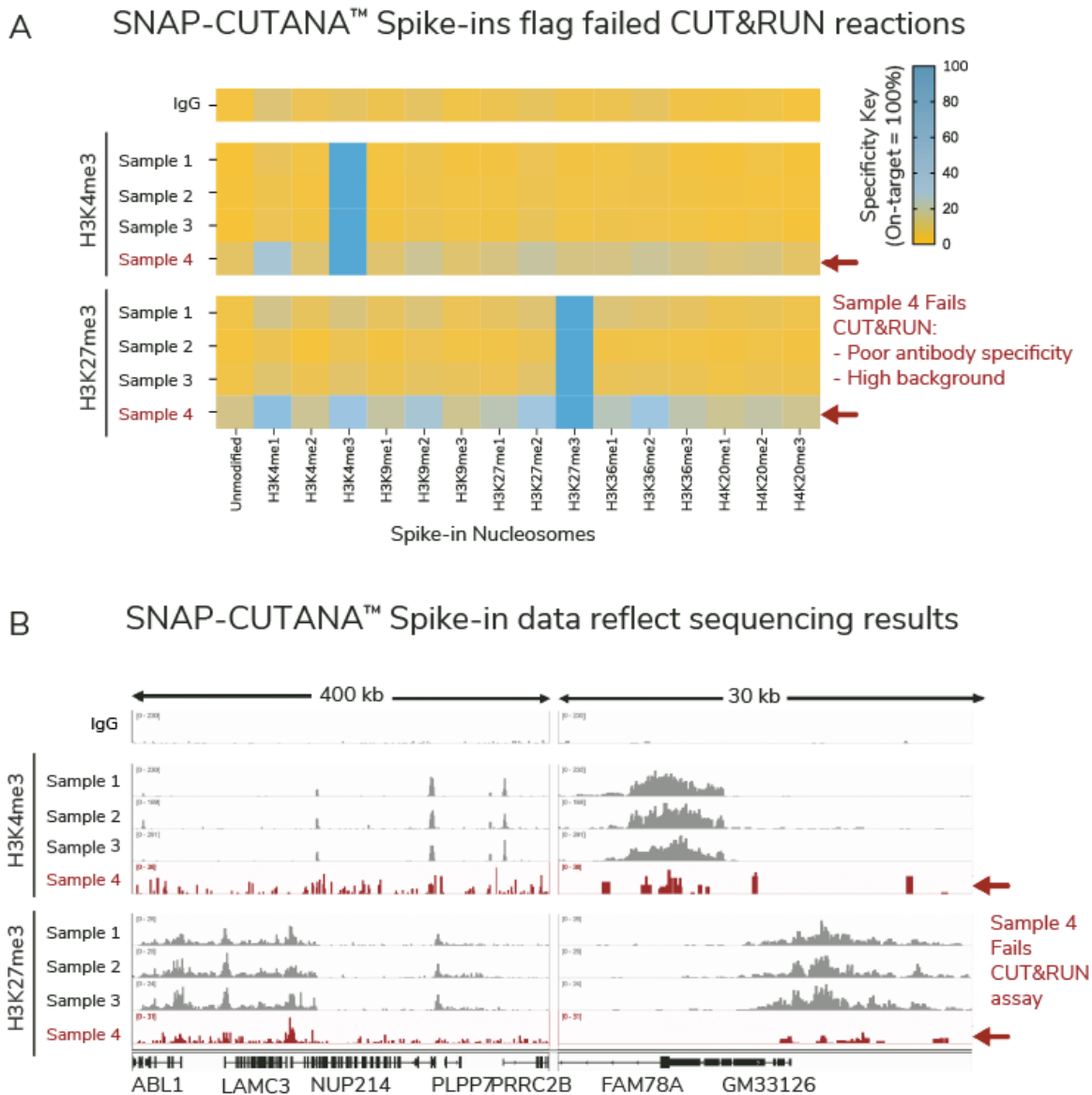

Figure 7: SNAP-CUTANA™ Spike-in Controls identify failed CUT&RUN reactions. CUT&RUN was used to map H3K4me3 and H3K27me3 in four independently prepared mouse primary B cell samples. IgG was included as a negative control. The K-MetStat Panel was spiked into each reaction. **(A)** Heatmaps show recovery 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 signal-to-noise and excessive background (red arrows).

Results	Interpretations and troubleshooting approaches
SNAP-CUTANA Spike-in data shows highly specific recovery of on-target nucleosome. BUT Genomic profiles display low signal-to-noise.	Enzyme (pAG-MNase or pAG-Tn5) and wash conditions are optimized. Antibodies are performing as expected. Problems may include: **Low cell/nuclei numbers** • Optimize with 500,000 cells (CUT&RUN) or 100,000 nuclei (CUT&Tag) before reducing input. • If using nuclei, adherent cells, cross-linked cells, tissues, or frozen samples, see CUT&RUN and CUT&Tag Kit manuals for modifications.* **Poor sample prep** • Perform Quality Control Checks outlined in CUT&RUN and CUT&Tag Kit manuals to confirm cell viability and to avoid cell or bead clumping during assay.* Note that low cell viability will increase background. • Optimize digitonin permeabilization if using whole cells.* • Confirm sample integrity (viability, morphology) and binding to ConA beads.* Note that poor viability and/or lysis increases background. • Avoid ConA bead dry-out during the assay. **Target requires different processing conditions** • Ensure target is present and localized to chromatin. • If possible, use freshly harvested samples. • Test native vs. lightly cross-linked conditions.*
SNAP-CUTANA Spike-in data shows high background and/or nonspecific recovery. AND Genomic profiles show poor signal and/or high background.	**Fundamental failure in 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. • Confirm that ConA beads are not damaged (never frozen, are brown). • Make sure correct parameters are used for indexing PCR. **Low cell number and/or poor sample prep** • Follow instructions above

Table 6: Using SNAP-CUTANA™ Spike-in results to guide troubleshooting.

* CUT&RUN and CUT&Tag Kit manuals available at epicypher.com/protocols.

(Continued on next page)

Results	Interpretations and troubleshooting approaches
SNAP-CUTANA Spike-in data shows high background and/or nonspecific recovery. BUT Genomic profiles show clear peaks and good signal over background	**Indicates problem with antibody to experimental target, e.g. poor enrichment efficiency or cross-reactivity.** • High signal-to-noise in CUT&RUN and CUT&Tag doesn't necessarily indicate a highly-specific antibody! • We suggest screening antibodies from different vendors targeting distinct epitopes.* • Shop antibodies at EpiCypher: epicypher.com/antibodies-complete-list/.
Control antibodies generate expected spike-in data AND genomic profiles. BUT Antibodies to additional targets show no clear enrichment and/or unexpected peaks.	The workflow is optimized for control antibodies and cell type. These results indicate: **Problem with antibody to experimental target**, *e.g.* poor enrichment efficiency or cross-reactivity. • We suggest screening antibodies from different vendors targeting distinct epitopes.* • Shop antibodies at EpiCypher: epicypher.com/antibodies-complete-list/. **Target requires different processing conditions** • Ensure target is present and localized to chromatin. • If possible, use freshly harvested samples. • Test native vs. lightly cross-linked conditions.*

Table 6: Using SNAP-CUTANA™ Spike-in results to guide troubleshooting.

* CUT&RUN and CUT&Tag Kit manuals available at epicypher.com/protocols.

6.3. How can SNAP-CUTANA Spike-ins identify high-quality histone PTM antibodies?

Selection of a highly specific antibody is crucial to CUT&RUN and CUT&Tag success. EpiCypher has used SNAP Spike-in technology to screen >1,000 commercial antibodies in chromatin mapping assays and found that the vast majority (**>70%**) of histone PTM antibodies are non-specific and/or display poor enrichment [1]. However, high quality PTM antibodies for CUT&RUN and CUT&Tag do exist (**Figure 8**). From these studies we have found:

- Modified histone peptide arrays do not reliably predict antibody performance in epigenomics assays (*e.g.* CUT&RUN, CUT&Tag, or ChIP-seq).

- Antibody validation in one assay does not guarantee success in another (*e.g.* ChIP validation does not transfer to CUT&RUN or CUT&Tag).

- Performance can vary by experimental condition. For instance, an antibody that works at 500,000 cells may fail at 10,000 cells, or performance may vary between cell types.

Application-specific testing with defined nucleosome spike-in controls is essential to success. EpiCypher only recommends histone PTM antibodies with <20% cross-reactivity, as determined using SNAP-CUTANA Spike-in Controls. These criteria support the accuracy of biological findings based on our prior work[1].
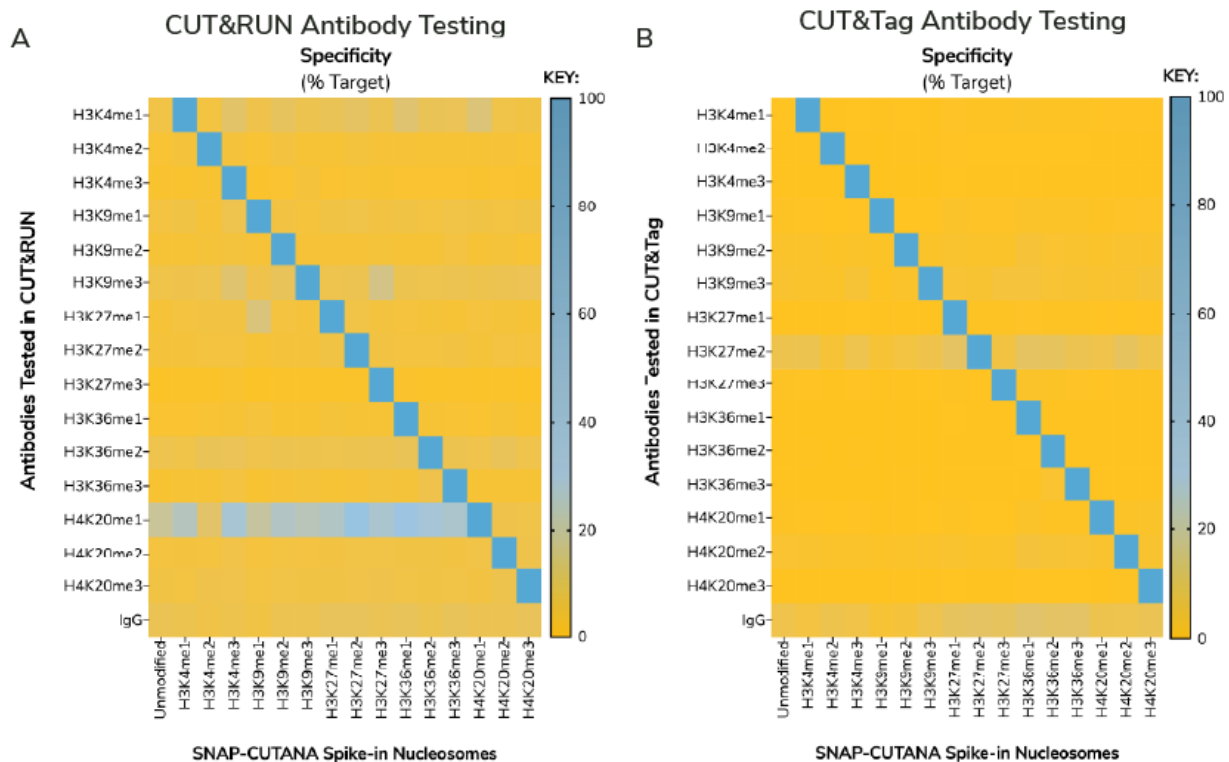

Figure 8: **EpiCypher's best-to-date CUT&RUN (A) and CUT&Tag (B) antibodies to lysine methylation PTMs.** Each row displays SNAP-CUTANA™ K-MetStat Spike-in data for a PTM antibody validated in CUT&RUN (**A**) or CUT&Tag (**B**). Heatmap data for each antibody are normalized to signal from the on-target PTM.

6.4. How accurate are the antibody specificity data?

SNAP-CUTANA Spike-ins correctly predict antibody specificity and performance in genomic profiling experiments, providing a powerful tool for advanced epigenomics research. In **Figure 9** 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 nucleosome were normalized to the PTM target and used to generate heatmaps (**Figure 9A**). H3K4me3 and H3K27me2 A antibodies showed high specificity for their respective PTM 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 sequencing results. **Figure 9B** 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 9B**). 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 9A**). Without the SNAP-CUTANA Spike-ins to flag antibody cross-reactivity, such data could mislead biological conclusions.
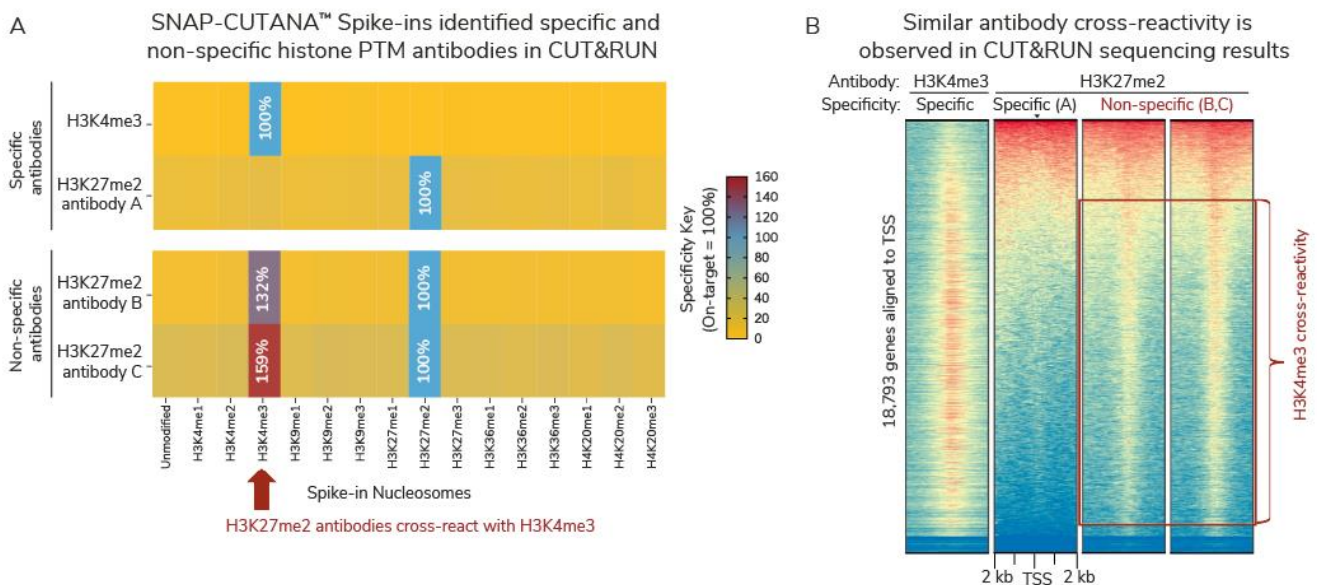

Figure 9: **Antibody cross-reactivity identified by SNAP-CUTANA™ Spike-ins (A) is reflected in CUT&RUN sequencing results (B).** (**A**) SNAP-CUTANA K-MetStat spike-ins were used to characterize H3K4me3 and H3K27me2 antibodies in CUT&RUN workflows using K562 cells. Spike-in data are shown normalized 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.

6.5. How can SNAP-CUTANA™ Spike-in Controls be used for normalization?

Several methods have been reported for experimental normalization using exogenous spike-ins[3-5]. These approaches can be adopted by calculating a single scalar normalization ratio using total read counts from SNAP-CUTANA Spike-ins.

To demonstrate the application of SNAP-CUTANA Spike-ins for sequencing 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 10**). 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. Sequencing data were RPKM normalized (top panel) or normalized to the K-MetStat Panel (similar to [5]; 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 10**), 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.
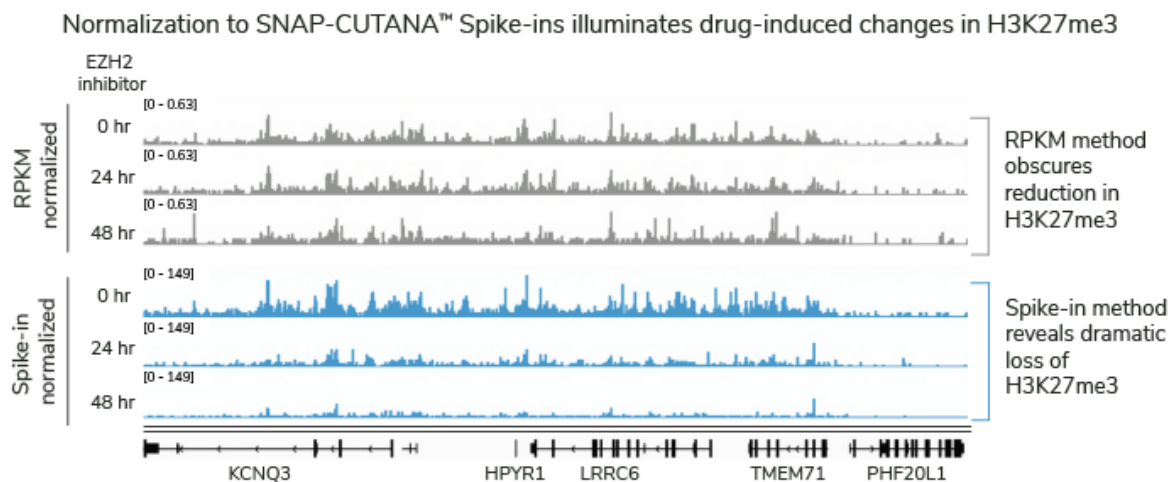

Figure 10: CUT&Tag data normalization using SNAP-CUTANA™ K-MetStat Spike-in Controls illuminates drug-induced differences in histone PTM enrichment. Cancer cells treated with the EZH2 inhibitor tazemetostat for the indicated times show a modest difference in H3K27me3 enrichment compared to vehicle (0hr) 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.*

6.6. Why does each nucleosome have two DNA barcode sequences?

Each modified nucleosome in a SNAP-CUTANA Panel is prepared with two distinct DNA barcode sequences. For each nucleosome, the two DNA barcodes are present at roughly equal amounts - thus providing a technical replicate. To explain, we use the K-MetStat Panel:

- To create the K-MetStat Panel, we assemble nucleosomes carrying 16 defined histone lysine methylation states: me1, me2, and me3 on H3K4, H3K9, H3K27, H3K36, H4K20, and unmodified.

- Nucleosomes are assembled in individual reactions, where the modified histone octamer is wrapped with a DNA template. For each nucleosome, we add **two** DNA templates carrying unique DNA barcode sequences (A and B). The final purified nucleosome is roughly 50% DNA barcode A and 50% DNA barcode B.

- The 16 nucleosomes are pooled together to make the K-MetStat Panel.

- Each nucleosome is represented by two DNA barcode sequences and should be detected similarly in sequencing data.

7. References

1   Shah, R. N. *et al.* Examining the Roles of H3K4 Methylation States with Systematically
     Characterized Antibodies. *Mol Cell* **72**, 162-177 e167 (2018).
     https://doi.org/10.1016/j.molcel.2018.08.015

2   Takaku, M. *et al.* GATA3-dependent cellular reprogramming requires activation-domain
     dependent recruitment of a chromatin remodeler. *Genome Biol* **17**, 36 (2016).
     https://doi.org/10.1186/s13059-016-0897-0

3   Tay, R. E. *et al.* Hdac3 is an epigenetic inhibitor of the cytotoxicity program in CD8 T
     cells. *J Exp Med* **217** (2020). https://doi.org/10.1084/jem.20191453

4   Lam, K. G., Brick, K., Cheng, G., Pratto, F. & Camerini-Otero, R. D. Cell-type-specific
     genomics reveals histone modification dynamics in mammalian meiosis. *Nat Commun*
     10, 3821 (2019). https://doi.org/10.1038/s41467-019-11820-7

5   Orlando, D. A. *et al.* Quantitative ChIP-Seq normalization reveals global modulation of
     the epigenome. *Cell Rep* **9**, 1163-1170 (2014).
     https://doi.org/10.1016/j.celrep.2014.10.018