**CUT&RUN & Bulk ChIC‑seq: A Bioinformatician’s Field Guide**

**1  Core Biological Principle**

CUT&RUN (Cleavage Under Targets & Release Using Nuclease) adapts **chromatin immuno‑cleavage (ChIC)** into a low‑background alternative to ChIP‑seq. An antibody binds its chromatin target *in situ*; a protein A/G–Micrococcal Nuclease (pA/G‑MNase) fusion is then tethered to that antibody. Upon Ca²⁺ activation, MNase cleaves DNA immediately adjacent to the epitope, releasing *native‑sized* fragments that diffuse into solution for library prep ([epicypher.com](https://www.epicypher.com/resources/blogcomplete-stepbystep-guide-to-cutrun/?utm_source=chatgpt.com)).

**Why it matters for informatics:** Because digestion happens only at bound loci, background fragments are orders of magnitude lower than in ChIP‑seq. Consequently, typical CUT&RUN datasets reach saturation at **3–10 M aligned read pairs** for abundant histone marks and ~20 M for transcription factors, versus ~50 M+ for ChIP‑seq.

**2  Wet‑Lab Workflow (explained for dry‑lab folks)**

| **Step** | **Wet‑Lab Action** | **What Bioinformaticians Observe** |
| --- | --- | --- |
| 1. Cell permeabilisation | Digitonin (or saponin) pokes holes in the plasma membrane while leaving nuclei intact. | Fragment size profiles < 150 bp confirm nuclei remain intact (little genomic smear). |
| 2. Antibody incubation | Target‑specific IgG binds chromatin overnight (4 °C). | Successful samples show strong enrichment at canonical loci (e.g. H3K4me3 peaks at TSSs). |
| 3. pA/G‑MNase binding | Fusion protein is added; no cleavage yet. | No sequencing readout at this stage. |
| 4. MNase activation | Ice‑cold Ca²⁺ triggers site‑proximal cutting (∼1–2 min). | Fragment distribution is sharply mononucleosomal (~120 bp) or sub‑nucleosomal (60–90 bp) depending on mark. |
| 5. Stop & release | EDTA + RNase A release fragments; DNA is extracted. | DNA yield 0.6–6 ng is expected for TFs ([cellsignal.com](https://www.cellsignal.com/learn-and-support/frequently-asked-questions/cut-and-run-faqs?srsltid=AfmBOooTZWASuIcemS4wg5JeYIQGOXvvWDDKpyWEXpkEt-Mr6--jvbsi&utm_source=chatgpt.com)). |
| 6. Library prep | End‑repair, A‑tailing, adapter ligation (or Tn5 tagmentation), indexing PCR. | Dual indices enable multiplexing (≥ 5 libraries recommended for colour‑balance) ([epicypher.com](https://www.epicypher.com/content/documents/manuals/14-1001-2-cut-and-run-library-prep-kit-manual.pdf?utm_source=chatgpt.com)). |

**3  Controls in Depth**

**3.1  Positive Control**

A high‑abundance mark such as **H3K4me3** or **CTCF** is processed in parallel. If library complexity, FRiP (fraction of reads in peaks), and peak morphology match published data, the run is technically sound ([epicypher.com](https://www.epicypher.com/resources/blog/epicyphers-h3k4me3-antibody-provides-a-useful-positive-control-for-cutrun/?utm_source=chatgpt.com)).

**3.2  Negative Control**

An isotype IgG or a no‑antibody tube quantifies background cleavage. Use it as the --control BAM in MACS2 or in nf‑core/cutandrun for robust peak calling ([cellsignal.com](https://www.cellsignal.com/learn-and-support/frequently-asked-questions/cut-and-run-faqs?srsltid=AfmBOooTZWASuIcemS4wg5JeYIQGOXvvWDDKpyWEXpkEt-Mr6--jvbsi&utm_source=chatgpt.com)).

**3.3  Spike‑in DNA**

Many kits add 1 % *E. coli* DNA prior to digestion. Reads aligning to the spike‑in genome provide a per‑sample scaling factor that normalises for digestion or recovery efficiency ([nf-co.re](https://nf-co.re/cutandrun/1.0.0/docs/output?utm_source=chatgpt.com)).

**4  Multiplexing & Index Design**

* **Bulk runs**: each antibody/sample combination is processed in a separate tube but libraries are **dual‑indexed (i5/i7)** and pooled for sequencing. Proper index diversity (>25 % at each cycle) prevents colour‑channel bias on Illumina platforms ([epicypher.com](https://www.epicypher.com/content/documents/manuals/14-1001-2-cut-and-run-library-prep-kit-manual.pdf?utm_source=chatgpt.com)).
* **Indexed bulk & single‑cell**: combinatorial barcoding (e.g., CUT&RUN‑Tag or sci‑CUT&RUN) introduces barcodes *before* cleavage, letting multiple reactions share a tube. Reads are demultiplexed *in silico* by barcode.

**5  Alignment & Primary Processing**

1. **Adapter/quality trim** (Trim Galore!/fastp).
2. **Alignment** with Bowtie 2 in very‑sensitive, dovetail‑allowed mode. 4DN’s pipeline aligns to the target genome **and** spike‑in genome, outputting separate BAMs and a normalised BigWig ([data.4dnucleome.org](https://data.4dnucleome.org/resources/data-analysis/cut-and-run-pipeline?utm_source=chatgpt.com)).
3. **Deduplication**: optional for histone marks (few PCR duplicates) but mandatory when UMIs are present.
4. **Peak calling** with MACS2 (--format BAMPE --nomodel --keep-dup all) or SEACR (good for broad marks).
5. **Quality metrics**: FRiP > 0.2 for histones, ENCODE blacklist removal, fragment length enrichment.

**6  Focus on Bulk ChIC‑seq**

Bulk ChIC‑seq is the predecessor to CUT&RUN, using **protein A–MNase without a nuclear permeabilisation step**. The antibody binds in fixed, permeabilised nuclei; MNase cleaves and fragments are released by raising Ca²⁺ *after* nuclei lysis. Key distinctions:

* **Input cells**: ChIC‑seq tolerates 10⁶–10⁷ cells; CUT&RUN works down to 5 × 10³.
* **Background**: ChIC‑seq exhibits higher nonspecific cleavage because intracellular nucleases may activate prematurely.
* **Data characteristics**: fragment sizes are more heterogeneous; hence bulk ChIC‑seq usually requires **deeper sequencing (20–50 M PE reads)** to compensate for background.
* **Analysis**: identical pipeline to CUT&RUN but expect lower FRiP; stringent negative‑control filtering becomes critical.

**7  Other Modalities (Brief Overview)**

| **Method** | **Key Difference** | **Typical Use‑Case** |
| --- | --- | --- |
| **CUT&Tag** | Uses Tn5 transposase tethered to antibody for simultaneous cleavage & adapter insertion. Generates extremely low background; ideal for single cells. |  |
| **Single‑cell CUT&RUN** | Combinatorial indexing or droplet microfluidics attach cell barcodes *in situ*; requires UMI deduplication and cell‑quality filtering (cells > 1 k fragments). |  |
| **CUTAC** | Low Ca²⁺, high Mg²⁺ variant for nucleosome‑sensitive marks; yields open‑chromatin footprints akin to ATAC‑seq. |  |

**8  Best‑Practice Checklist**

1. **Run positive & negative controls *every batch*.**
2. **Aim for 3–10 M aligned read pairs** for abundant histone marks in CUT&RUN; 20 M+ for TFs or bulk ChIC‑seq.
3. **Include spike‑in DNA** and normalise signal tracks accordingly.
4. **Remove blacklist regions** and investigate any IgG‑overlapping peaks.
5. **Document indices & barcodes** to avoid index bleed‑through when multiplexing.
6. **Visual QC**: confirm fragment‑size mode (<150 bp), FRiP, and expected peak patterns before differential analysis.

