DEVELOPMENT OF txci-CAB AS A NEW SINGLE-CELL APPROACH TO STUDY CHROMATIN MODIFICATIONS AND ACCESSIBILITY ACROSS THE GENOME

by

Radhika Rathi

A Thesis Submitted to the Faculty of the

DEPARTMENT OF CELLULAR AND MOLECULAR MEDICINE

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2025

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by Radhika Rathi, titled *Development of txci-CAB as a New Single-cell Approach to Study Chromatin Modifications and Accessibility Across the Genome* and recommend that it be accepted as fulfilling the dissertation requirement for the Master's Degree.

Dr. *Darren Cusanovich* Date: 04/24/2025

Dr. *Casey Romanoski* Date: 04/24/2025

Dr. *Julie Ledford* Date: 04/24/2025

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

Dr. *Darren Cusanovich* Date: 04/24/2025
Master's Thesis Committee Chair
Department of Cellular and Molecular Medicine

Table of Contents

Lists of Figures and Tables

Abstract

Understanding the epigenomic landscape of single cells is critical for uncovering the regulatory mechanisms that govern cell identity, differentiation, and function. Traditional chromatin profiling methods often measure a single modality per assay, limiting the ability to directly capture the interplay between chromatin accessibility and histone modifications. To overcome this limitation, we developed txci-CAB (10x-compatible combinatorial indexing of CUT&Tag and ATAC both) by combining combinatorial indexing with a droplet-based microfluidic system.

In this thesis, we work on systematic optimization of ATAC-seq and CUT&Tag protocols in order to integrate them within the same workflow to simultaneously profile accessible chromatin states as well as specific modifications in individual cells. The txci-CAB protocol incorporates Tn5-based combinatorial barcoding for chromatin accessibility followed by pA-Tn5-guided CUT&Tag for histone modifications.

Although several multimodal single-cell sequencing techniques exist, txci-CAB is designed to enhance scalability, sensitivity and overall convenience of use, both in terms of input reagents and protocol workflow. Our approach also emphasizes increased library complexity while maintaining robustness and reproducibility. Notably, we observed higher estimated complexity in our bulk datasets, demonstrating the method's effectiveness in simultaneously capturing two chromatin modalities at single-cell resolution.

1. Introduction

A wide array of cell types comprises the human body, each with distinct functions, despite sharing nearly identical genomes [1]. This cellular diversity is largely driven by regulation of gene expression, controlled by a complex network [1] of cis-regulatory elements [2] that modulate transcription through changes in chromatin accessibility [3]. The initiation of gene expression is central to enhancers and the transcription factors that bind them among various regulatory elements [2]. Another key mechanism by which chromatin exerts this regulatory influence is through chemical modifications of histone proteins. These histone modifications form a complex that influences chromatin accessibility, gene activation, and silencing, functions that are integral to most biological processes [4].

Therefore, locating these histone modifications and related chromatin accessibility in a comprehensive and multimodal study can help deconvolute cell differentiation trajectories and uncover the regulatory events underlying cell fate decisions prior to the activation of transcriptional programs.

1.1 Chromatin Structure & Function

The nucleus of a eukaryotic cell consists of a majorly conserved and essential structure called chromatin which stores the genetic information and facilitates the important functions of the genome [5]. In Chromatin, DNA is packaged in repeating units called nucleosomes. Each nucleosome is a fundamental structural unit of chromatin which consists of an octamer made up of two copies each of histone proteins H2A, H2B, H3 and H4 and approximately 147 base pairs of DNA wrapped around it [6], [7]. Millions of these nucleosomes compact the DNA into chromatin which is further condensed into chromosomes. This sophisticated packaging strategy enables nearly 2 meters of DNA to fit within the confined space of the nucleus [8].

Beyond its role in packaging DNA, chromatin plays a pivotal role in regulating gene expression for both cellular development and self-maintenance. Chromatin is found to be uncondensed in specific genomic regions (euchromatin) where it is accessible to DNA binding factors or condensed into higher-order conformations (heterochromatin) which was long thought to be uniformly inaccessible but later found be heterogenous and dynamically accessible [9]. This dynamic control to DNA accessibility plays an imperative role in instituting and preserving the cell identity. The chromatin organization of accessible regions reflects a network of enhancers, promoters and chromatin-binding factors which interactively regulate gene expression. External stimuli and differentiation cues are responsible for the change in accessibility landscape. Moreover, new evidence hints towards a competitive relationship between nucleosomes and chromatin-binding factors in regulating the homeostatic maintenance of accessibility [10]. Thus,

profiling these accessible chromatin regions can provide critical insights into the active regulatory landscape of specific cell types or tissues [11].

1.2 Epigenetic Regulation

Cells transition from one state to another due to complex gene regulatory networks which drive dynamic changes in gene expression. Central to these networks are transcription factors (TFs), which not only interpret but can also reshape the chromatin landscape to control gene activity. Chemical modifications to the chromatin affect chromatin accessibility and higher-order chromatin structure, directly impacting gene regulation. [12].

Epigenetics refers to heritable biochemical and structural changes to chromatin that influence gene expression without affecting the underlying DNA sequence [13], [14]. These heritable but reversible epigenetic modifications modulate chromatin accessibility and regulate a wide range of physiological and pathological processes [13], [15], [16]. Chromatin organization and accessibility are shaped by various chemical modifications made to both DNA and histone proteins [12]. These epigenetic modifications govern how the cellular machinery interprets the genome, particularly in the context of transcription [15].

In multicellular organisms, cellular differentiation is driven by the generation of specialized cell types from undifferentiated progenitors or stem cells [17]. This process is orchestrated through epigenetic reprogramming, which establishes lineage specific gene expression patterns. These transcriptional programs are tightly regulated and maintained through various chromatin-based mechanisms including transcription factors binding, post-translational modifications of histones, incorporation of histone variants, DNA methylation, and nucleosome remodeling. Together, these dynamic processes govern gene activation or silencing to ensure proper cell fate decisions during development [18].

Cis-regulatory elements like promoters, enhancers, insulators, and transcription factor (TF) binding sites are epigenetic features largely present in accessible regions of chromatin [16]. These active and accessible chromatin regions are typically bound by regulatory proteins, whereas inactive or repressive chromatin regions are more compacted and less accessible to the transcriptional machinery [12]. Alongside chromatin accessibility, specific histone modifications also play a crucial role in regulating gene expression. Together, these elements form a dynamic and complex regulatory landscape that can influence transcriptional activity without altering the underlying DNA sequence [16], [14].

1.3 Evolution of Sequencing Techniques

Over the years, a wide range of molecular techniques have been developed to enable comprehensive profiling and characterization of epigenome [14]. A major catalyst for this

progress was the release of the first complete sequence of the human genome in 2001 [19]. The successful completion of the Human Genome Project, which spanned 13 years, marked a transformative milestone in genomics. This achievement set the stage for rapid advancements in sequencing technologies, particularly the development of next-generation sequencing (NGS), which made large-scale genome and exome sequencing widely accessible within a decade of the project's completion [20].

Sequencing technologies have undergone remarkable advancements since the introduction of Sanger sequencing, leading to greater data output, improved efficiency and expanded applications. Modern sequencing platforms are now generally classified based on their read length as either second-generation (short-read) or third-generation (long-read) sequencing technologies [21].

Compared to traditional Sanger sequencing, NGS technologies offer dramatically higher throughput at reduced time and cost, enabling large-scale genome analysis with unprecedented efficiency. Furthermore, the integration of diverse sequencing approaches with high-throughput platforms has expanded our ability to study complex biological processes and uncover the genetic basis of disease with greater resolution and depth [22].

To study epigenetic features as described in the previous sections, a variety of DNA sequencing techniques have been developed in order to map histone modifications, DNA modifications, or transcription factor binding sites. Most of these sequencing techniques are antibody-based [12]. Techniques like chromatin immunoprecipitation followed by sequencing (ChIP-seq) are used to study protein-DNA interactions [23] and chemical modifications of histone proteins [24]. ChIP and its derivatives utilize antibodies to enrich DNA fragments bound by a specific protein of interest. The enriched DNA is then sequenced and aligned to a reference genome, where read enrichment at specific regions reflects the presence of the target protein or modification at those loci [25]. ChIP-seq and DNA adenine methyltransferase identification (DamID) [26] are among the two commonly used methods. When coupled with high-throughput technologies such as microarrays or NGS, these methods enable genome-wide profiling of chromatin landscapes [27].

1.4 Chromatin Accessibility Assays

What distinguishes chromatin accessibility assays from other epigenomic approaches like histone ChIP-seq, is their ability to directly evaluate the influence of chromatin structure on gene regulation. While ChIP-seq relies on the detection of histone modifications to deduce transcriptional activity, accessibility assays provide a more straightforward way to measure which regions of the genome are open and potentially active, without the need for antibodies or epitope tagging, thus reducing experimental biases [28].

Several techniques have been developed to measure genome-wide chromatin accessibility in order to understand gene regulation. The majority of them depend on a key principle of evaluating the effectiveness of an enzyme to access and cleave DNA within the chromatin. This ability is shaped by the size of the enzyme, its sequence preferences, and, importantly, the presence of chromatin-bound proteins, all of which influence enzyme access to specific genomic regions [29]. Early efforts to map chromatin accessibility at a genome-wide scale began with DNase-seq [30], which utilizes DNase I digestion, and MNase-seq, which employs Micrococcal Nuclease (MNase) [31]. Later, the development of ATAC-seq (Assay for Transposase Accessible Chromatin) provided a faster and more sensitive method of capturing accessible regions, where chromatin is typically more open and nucleosome occupancy is low. DNase-seq and ATAC-seq are effective in capturing highly accessible or nucleosome-depleted regions and are often used to analyze sub-nucleosomal fragments [28]. These small DNA fragments typically arise from regions of open chromatin where the transposase can readily insert and cut the DNA, while less accessible regions tend to yield longer fragments that are usually excluded during library preparation. ATAC-seq inherently favors the detection of hyper-accessible regions which was proved by the overrepresentation of mitochondrial DNA in early experiments, as unlike nuclear genome, mitochondrial DNA lacks nucleosome packaging. While this bias reflects the method's sensitivity to highly open chromatin, it can be advantageous when the goal is to use minimal sequencing effort to efficiently identify active regulatory elements, such as enhancers and promoters [29].

ATAC-seq exploits a hyperactive Tn5 transposase to fragment DNA while simultaneously inserting sequencing adapters into accessible regions of the genome [32]. Usually performed on 500 to 50,000 unfixed nuclei, the transposition reaction preferentially occurs in nucleosome-depleted regions due to steric hindrance inflicted by chromatin-bound proteins. Following transposition, PCR amplification is used to construct sequencing libraries for next-generation sequencing (NGS). ATAC-seq has proven effective for mapping open chromatin landscapes, nucleosome positioning, and transcription factor footprints across the genome [28].

The widespread adoption of ATAC-seq stems from its greater efficiency and sensitivity compared to earlier techniques and its application in profiling cell-type-specific chromatin landscapes and investigating cellular heterogeneity at single-cell resolution [29], [28]. Cells within heterogeneous tissues work together to carry out complex biological processes, many of which remain incompletely understood [33].

1.5 Single-cell Sequencing Techniques

Recent advances in single-cell sequencing technologies have transformed developmental biology by enabling detailed exploration of tissue complexity, cellular heterogeneity, gene regulatory

networks, and dynamic processes like cell fate transitions and developmental trajectories [33], [34] and provided powerful solutions to overcome challenges associated with limited input material [35].

In addition to enabling genome-wide profiling from small numbers of cells, single-cell sequencing technologies also enable the comprehensive analysis of the genome, transcriptome, and other multi-omic features and provide valuable insights into cellular diversity and evolutionary relationships that are often masked in bulk sequencing approaches, which average signals across large population of cells and obscure cellular heterogeneity [36].

The application of these technologies has profoundly expanded our understanding of diverse biological processes, including gene transcription, embryonic development, and tumorigenesis. Even genetically identical cells can exhibit significant variability at the genomic, transcriptomic, and epigenomic levels. This variability arises from dynamic reprogramming of the genome and epigenome, as well as errors introduced during DNA replication, cell division, and differentiation. Therefore, studying biological processes at single-cell resolution has become essential for uncovering the complexity of cellular states and functions, and has driven the rapid development of innovative sequencing technologies [35].

Although the early adoption of single-cell sequencing was limited by high costs and technical challenges, continued innovation has significantly improved accessibility and reduced costs. As a result, these technologies are now widely applied across various biological and biomedical research fields [36].

Since the first successful application of single-cell transcriptome sequencing in 2009, there has been rapid progress in developing technologies capable of profiling additional layers of cellular regulation at single-cell resolution including the genome, DNA methylation, chromatin architecture, chromatin accessibility, and histone modifications [37]. While single-cell RNA sequencing (scRNA-seq) [38] has been widely used to study gene expression variability across individual cells, profiling the epigenomic landscape provides crucial insights into the regulation of gene expression [33]. And, single-cell epigenomic sequencing presents unique challenges due to the low abundance of epigenetic marks across the genome of a diploid genome, which contains only two copies of DNA [37].

1.6 Tn5-Based Innovations

Efficient single-cell epigenomic profiling thus relies on highly sensitive enzymatic or chemical reactions and minimal DNA loss during processing. Strategies that enable early incorporation of cell-specific barcodes are particularly advantageous, allowing large numbers of cells to be pooled and processed together, thereby improving throughput and scalability. Recent advances in

technologies such as array-based platforms [39], droplet microfluidics [40], and combinatorial indexing approaches [41], have significantly enhanced automation and efficiency of single-cell workflows [37], [33]. Notably, the Tn5 transposase has emerged as a critical tool in this field, offering advantages in sensitivity, simplicity, early barcoding, and scalability, making it central to many recent innovations in single-cell epigenomic technologies [37].

With the help of these improvements it has been made possible to profile chromatin accessibility in thousands of individual cells efficiently and at reduced cost using single-cell Assay for Transposase Accessible Chromatin using sequencing (scATAC-seq) [33].

Single-cell epigenomic profiling methods, such as single-cell ATAC-seq and single-cell DNA methylation sequencing, have provided valuable insights into the epigenetic diversity within tissues. While chromatin accessibility and DNA methylation offer snapshots of active or repressive chromatin states, profiling specific histone modifications can further enhance our understanding of epigenomic regulation [42]. Traditionally, chromatin immunoprecipitation followed by sequencing (ChIP-seq) [25] has been the standard method for studying histone modifications, although it requires large input material and often suffers from low signal-to-noise ratios. Recent low-input methods, including CUT&Run [43] and CUT&Tag, have overcome these limitations, allowing efficient profiling of chromatin features. Building on these advancements, a single-cell CUT&Tag (scCUT&Tag) approach has been developed by integrating CUT&Tag chemistry with droplet-based platforms like 10x Genomics scATAC-seq, enabling large-scale, single-cell resolution mapping of histone modifications in complex tissues [34].

Cleavage Under Targets and Tagmentation (CUT&Tag) is an innovative chromatin profiling method designed to overcome the limitations of traditional techniques like ChIP-seq and CUT&RUN. CUT&Tag utilizes a fusion protein consisting of Protein A tethered to a hyperactive Tn5 transposase (pA-Tn5), which is pre-loaded with sequencing adapters. Upon binding of a target-specific antibody to chromatin in intact cells, the pA-Tn5 complex is recruited to the site of interest. Activation of the transposase enables efficient tagmentation, directly generating sequencing-ready DNA fragments with minimal background noise. CUT&Tag is a rapid, cost-effective method that can be performed in a single tube. Remarkably, it allows high-resolution profiling of various chromatin features, including histone modifications and transcription factors even from low-input samples or single cells [44].

scATAC-seq and scCUT&Tag have greatly improved our ability to study chromatin landscapes at single-cell resolution. Both the methods, utilize Tn5 transposase or its fusion derivatives, allow high-resolution mapping of chromatin accessibility and histone modifications, even in complex or heterogeneous cell populations. They have proven especially valuable for

investigating dynamic biological processes like development and cellular differentiation. Additionally, integrating single-cell epigenomic data with other single-cell omics datasets, such as transcriptomics, has provided deeper insights into the relationship between chromatin state and gene regulation. However, many of these approaches rely on computational correlation across separate datasets, which may not fully capture the complexity of cellular regulation. Therefore, technologies capable of directly measuring multiple molecular modalities within the same cell are essential for achieving a more comprehensive understanding of chromatin function [45].

1.7 txci-CAB Development

Multimodal single-cell sequencing approaches that simultaneously capture different layers of chromatin regulation such as chromatin accessibility and histone modifications provide valuable insight into the coordination of regulatory features that define cell identity and function. Several methods have been developed to enable joint profiling of chromatin modalities in single cells. For instance, nano-CT uses nanobody-tethered Tn5 transposase instead of antibody-Protein A fusion system, to jointly profile two-three chromatin modalities in a single cell including histone modifications and chromatin accessibility [45], [46]. Similarly, multi-CUT&Tag employs a mixture of antibody-tethered transposases to simultaneously map multiple histone modifications within the same cell [47].

While these methods have expanded the capabilities of single-cell epigenomic profiling, challenges remain in terms of assay sensitivity, scalability, and the ability to efficiently capture both accessible chromatin and specific histone modifications in a single, streamlined workflow. Nanobody-based CUT&Tag, while eliminating the need for secondary antibodies and improving resolution, is limited by the availability of validated nanobodies and potential variability in fusion protein expression and activity. Furthermore, reproducibility in chromatin accessibility assays remains a broader issue, partly due to the lack of standardized guidelines for determining replicate numbers. Technical variability, often specific to assay design or biological context, adds another layer of complexity when aiming for robust and reproducible results [28].

To address these limitations, we developed *txci-CAB* (10x compatible combinatorial indexing in CUT&Tag and ATAC both), a multimodal single-cell chromatin profiling method that integrates ATAC-seq and CUT&Tag chemistries to jointly measure chromatin accessibility and chromtin modifications from the same cells.

By leveraging Tn5-based barcoding strategies and antibody-directed profiling within a combinatorial indexing framework, *txci-CAB* enables high-throughput, joint profiling of multiple chromatin modalities while maintaining assay simplicity and flexibility. This approach provides

a two-step strategy for the simultaneous profiling of chromatin accessibility and histone modifications within the same single cell, providing an integrated view of the regulatory landscape.

In txci-CAB, accessible chromatin is first tagged with Tn5-loaded adapters, as in standard ATAC-seq. The same nuclei then undergo CUT&Tag, where an antibody targets a histone mark and a pA-Tn5 fusion inserts adapters at those sites. This combinatorial indexing approach enables simultaneous, single-cell profiling of chromatin accessibility and histone modifications in a streamlined workflow (Fig. 1)

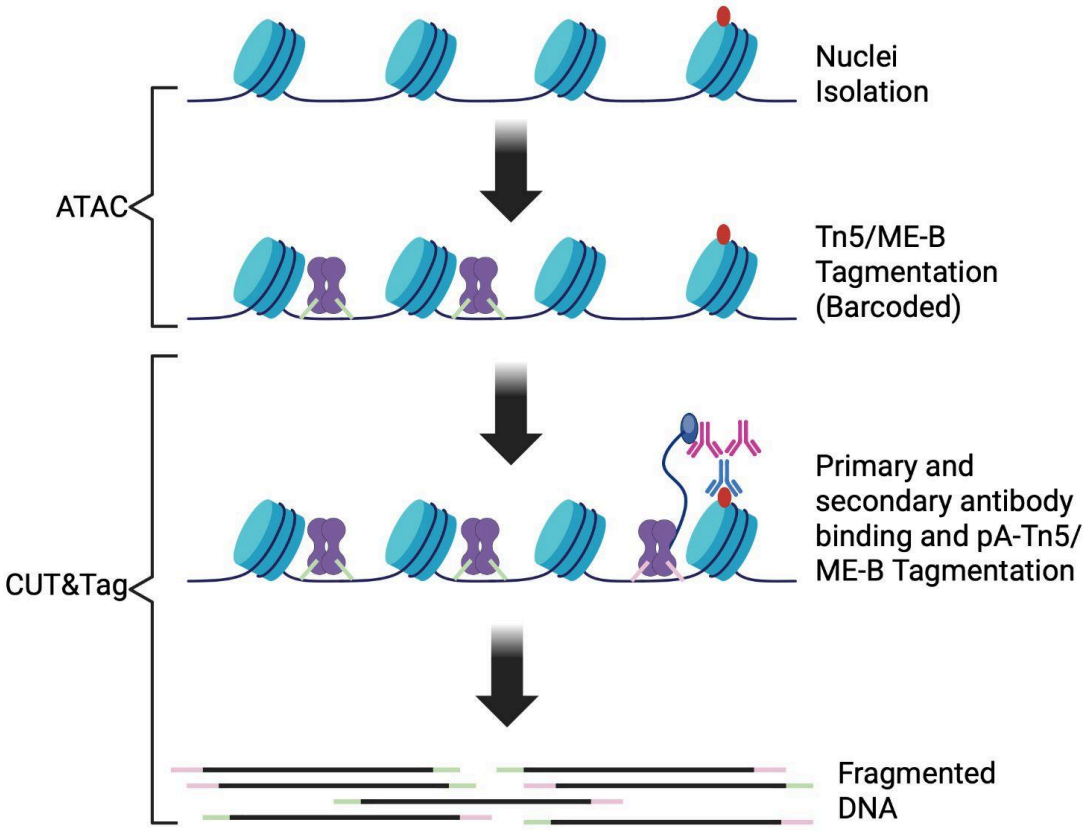


Fig. 1: Schematic of txci-CAB sequencing technique.

2. Results

To address these limitations, we developed **txci-CAB** (10x-compatible combinatorial indexing in CUT&Tag and ATAC both), a high-throughput, multimodal single-cell chromatin profiling method that enables joint measurement of chromatin accessibility and histone modifications from the same cells.

txci-CAB was adapted from the previously established txci-ATAC-seq method, which utilized transposase-based combinatorial indexing to enable high-throughput profiling of chromatin accessibility. txci-ATAC-seq was developed by combining 96-well plate-based Tn5 pre-indexing with 10x Genomics droplet barcoding. This approach allows multiplexing of up to 96 samples and profiling of ~200,000 nuclei in a single reaction without compromising data quality. Unlike previous pre-indexing strategies, txci-ATAC-seq enables accurate deconvolution of nuclei even in overloaded droplets [48].

To optimize the CUT&Tag and ATAC-seq protocols in reference to each other and for single-cell applications, we systematically evaluated key factors that influence experimental efficiency. In our pilot experiment, we examined the efficiency of Tn5 transposase in CUT&Tag buffers. This allowed us to determine how buffer composition influences enzymatic activity, providing crucial insights for protocol optimization.

2.1 Tn5 exhibits enhanced tagmentation efficiency in ATAC-seq buffer in comparison to CUT&Tag buffer

We aimed to integrate Tn5 transposase from the ATAC-seq protocol into the CUT&Tag workflow to enable combinatorial indexing prior to a second round of indexing on the 10x platform. This approach would allow for the simultaneous capture of chromatin accessibility via Tn5 barcoding and the mapping of specific chromatin modifications within the same cell using pA-Tn5. To achieve this dual-layered information, we sought to merge the two protocols by evaluating the efficiency of Tn5 in CUT&Tag buffer. To systematically assess this integration, we performed the standard Omni-ATAC protocol using the CUT&Tag tagmentation buffer to tagment the Tn5 transposase. The “ATAC protocol to evaluate efficiency of Tn5 in CUT&Tag buffers” workflow was followed as described in the methods section.

Our results demonstrated that Tn5 exhibited higher efficiency in the ATAC-seq buffer. This was evident from the qPCR amplification curves, where Tn5 transposase paired with the ATAC tagmentation buffer (Green; Fig. 2B) reached the plateau phase earlier than when used in CUT&Tag conditions (Pink; Fig. 2B).

When amplified DNA libraries from these experiments were run on a 6% PAGE gel, we observed the expected smear from ~100 bp to 1 kb, along with distinct bands near 200 bp and 400 bp (Fig. 2C; 1). These bands correspond to nucleosome-free DNA and mono- and di-nucleosome fragments, respectively, and reflect the underlying chromatin structure [49]. In our experiments, this nucleosomal laddering pattern was more pronounced when Tn5 was used in the ATAC-seq buffer, suggesting more efficient and accurate tagmentation of open and nucleosome-bound regions compared to when Tn5 was used in CUT&Tag conditions. The sharper bands and cleaner smear indicate that the ATAC buffer condition not only enabled better enzyme activity but also preserved nucleosomal architecture more faithfully [49], [50].

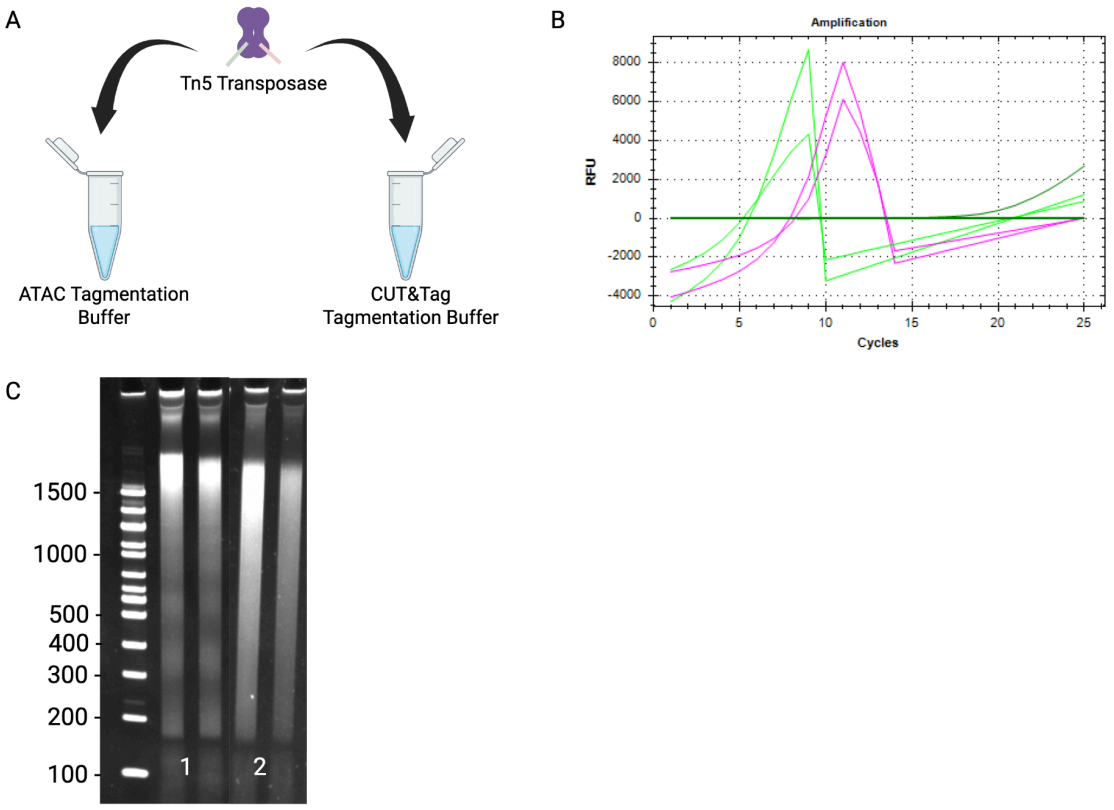


Fig. 2: Schematic and results for Tn5 tagmentation efficiency experiment. A. A schematic of experimental workflow illustrating Tn5 tagmentation using two different buffers to assess compatibility. B. qPCR amplification curves comparing Tn5 activity in ATAC buffer (green) vs. CUT&Tag buffer (pink). RFU- relative fluorescence units. C. Gel electrophoresis showing banding pattern of DNA fragments generated by Tn5 in ATAC buffer (1) and CUT&Tag buffer (2). A 1kb ladder was used as a size reference.

This banding pattern does more than confirm tagmentation, it provides insight into chromatin accessibility and structure. The periodicity of ~200 bp in fragment sizes, which we observed more clearly in ATAC buffer conditions, is indicative of nucleosome spacing and reflects a well-organized chromatin landscape [50]. These patterns suggest that Tn5 transposase in ATAC buffer more effectively captures chromatin features like nucleosome positioning and compaction, whereas the same enzyme in CUT&Tag buffer shows reduced efficiency (Fig. 2C; 2), likely due to suboptimal buffer compatibility. Thus, our results highlight how enzyme-buffer pairing directly impacts both the efficiency and biological interpretability of tagmentation-based chromatin profiling methods.

2.2 Early Tn5 Tagmentation in ATAC Buffer Yields Lower Complexity and Signal Enrichment in CUT&Tag Libraries

To evaluate the efficiency of Tn5-mediated tagmentation within the CUT&Tag framework, we used an anti-H3K4me3 primary antibody (Fisher; Cat. No. NC0014112) to capture accessible regions of chromatin via pA-Tn5 and Tn5 both, which would add sequencing adaptors to the opposite ends of the DNA fragment. The protocol, as described in the methods section under "Testing ME-B-Tn5 with ME-A-pA-Tn5 in CUT&Tag protocol," was used to sequentially incorporate ME-B adaptors via Tn5 followed by ME-A adaptors through pA-Tn5.

qPCR analysis revealed a two-cycle reduction in amplification efficiency compared to the control (Green- Control, Pink- ATAC buffer; Fig. 3B), which at first glance suggested that the library construction process may have avoided over-amplification. However, this modest shift also raised concerns about the overall robustness of the amplification step. While the combined tagmentation strategy did not introduce strong PCR bias, the subtle decrease in efficiency may indicate suboptimal reaction conditions or partial loss of fragment diversity. Sequencing analysis revealed that libraries generated using the ATAC buffer condition exhibited lower estimated library complexity (Fig. 3C). As a proxy for how thoroughly the chromatin landscape is sampled, this reduction points to underlying inefficiencies in the protocol. To address these limitations, we next explored whether modifying the order of Tn5 tagmentation within the CUT&Tag workflow could improve both library complexity and signal enrichment.

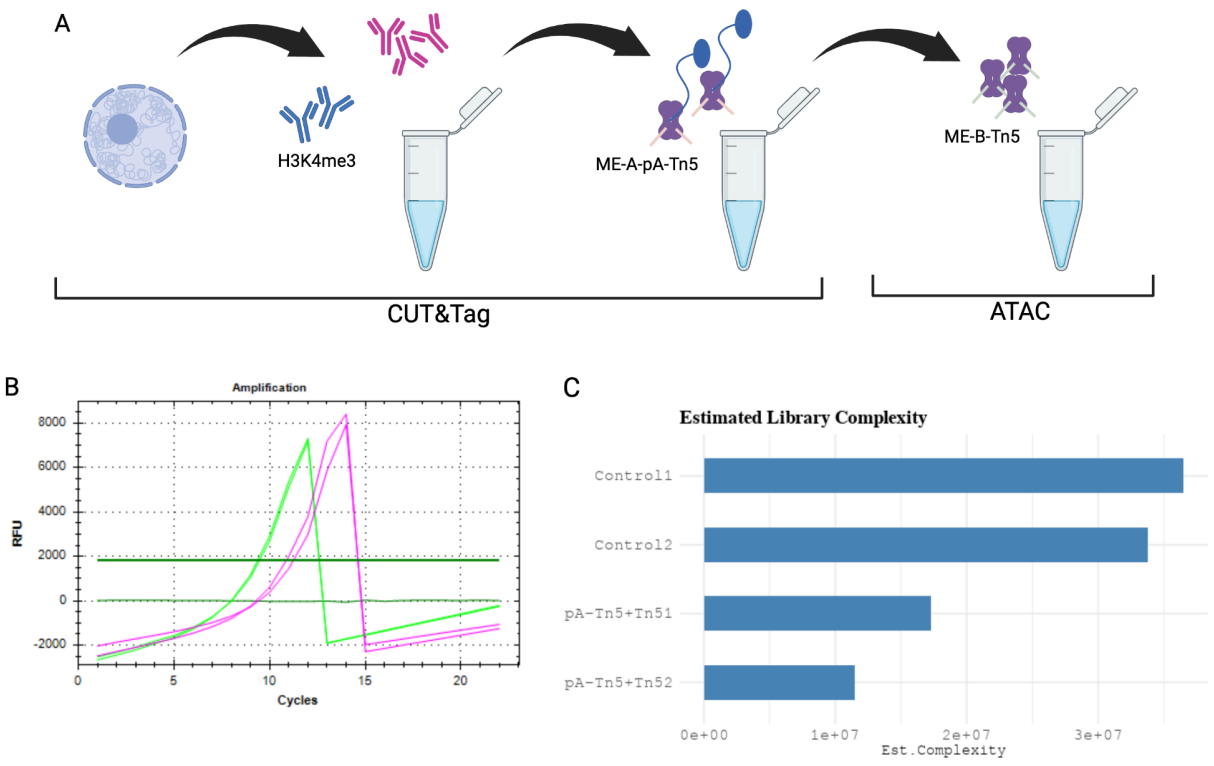


Fig 3: Experimental schematic and QC metrics of Tn5 in CUT&Tag protocol. A. Schematic of experimental workflow illustrating the sequential binding of antibodies, pA-Tn5 (loaded with ME-A adapters), and Tn5 (loaded with Me-B adapters). B. qPCR amplification curves comparing control CUT&Tag libraries (green) to those generated using the dual-enzyme workflow (pink). C. Estimated library complexity of control and experimental samples, including technical replicates.

2.3 Integrating ATAC Pre-Tagmentation into CUT&Tag Workflow Enhances txci-CAB Profiling of Open and Closed Chromatin

Next, we explored integrating ATAC-seq preprocessing into the CUT&Tag workflow, inspired by the Nano-CT [45] protocol. We adapted this approach by incorporating ATAC-seq at the beginning of the workflow to improve library complexity and signal enrichment and allowing for easier indexing in a 96-well plate before pooling all samples and transitioning into the CUT&Tag workflow. To test this adaptation, we conducted an experiment alongside a control. The condition we aimed to evaluate involved applying Tn5 tagmentation during the ATAC workflow before transitioning into CUT&Tag while employing an anti-H3K27me3 antibody. Profiling accessible regions in the same cell along with a repressive mark would provide a clear distinction

between Read1 and Read2 on the same fragment assuming that Read1 reflects the repressive mark and Read2 indicates ATAC-accessible regions on the genome. In this configuration, Tn5 would tag the genome with ME-B sequencing adapters prior to pA-Tn5 activity (Fig. 1).

The qPCR amplification curve showed that when pA-Tn5 was tagmented before Tn5, there was a 3-cycle delay in reaching the amplification plateau compared to the control (Fig. 4B). In contrast, Tn5 tagmentation prior to pA-Tn5 resulted in amplification curves that plateaued at the same cycle as the control (Fig. 4C), suggesting greater library complexity when Tn5 is tagmented earlier. Nucleosomal banding patterns for both samples exhibited ~200 bp periodicity; however, the prior Tn5 tagmentation condition showed sharper bands and a cleaner smear, indicating more efficient enzyme activity while preserving chromatin structure (Fig. 4D).

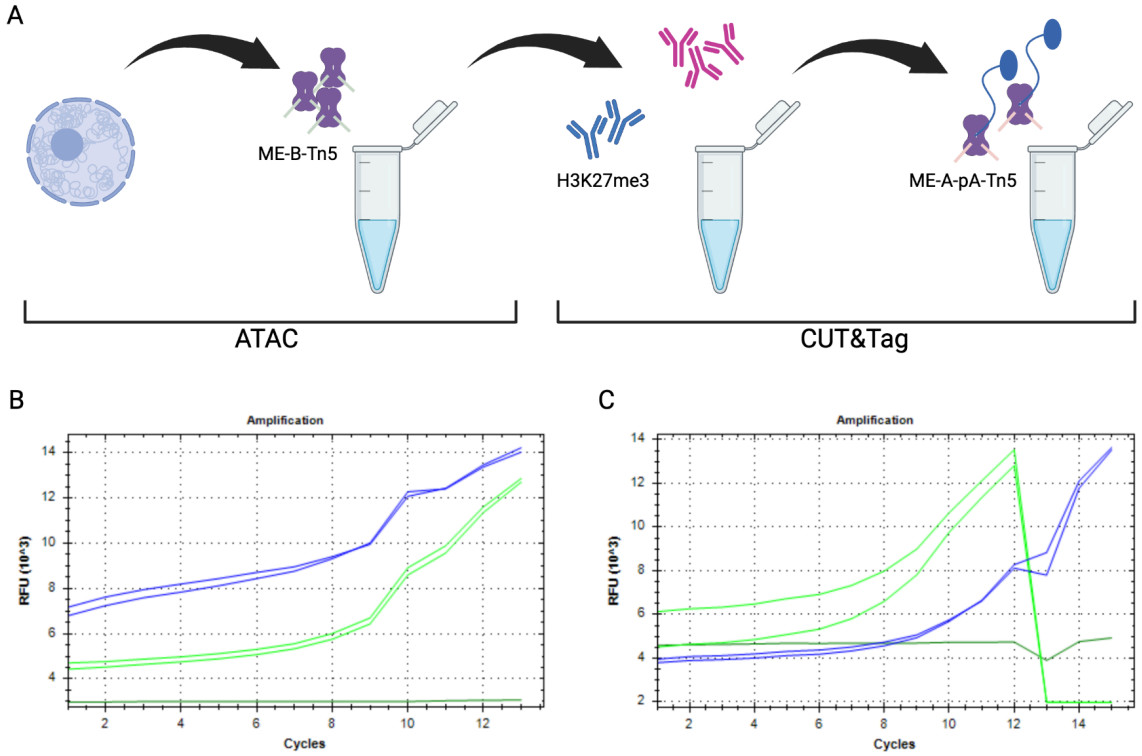
Sequencing analysis revealed that the estimated library complexity for the prior Tn5 tagmentation condition was comparable to the control, while the prior pA-Tn5 condition resulted in significantly reduced complexity (Fig. 4E) as observed previously. The results suggest that chromatin features were profiled more comprehensively and indicated diversity in unique fragments in the case of prior Tn5 tagmentation while prior pA-Tn5 tagmentation indicated less representative library with higher duplication rates which is not ideal for single-cell chromatin profiling.

To evaluate the impact of enzyme order on multimodal chromatin profiling using txci-CAB, we visualized genome browser tracks across a ~1.6 Mb region on chromosome 15, encompassing multiple gene loci including *RSL24D1*, *RAB27A*, *PRTG*, and *RFX7* (Fig. 4F). These tracks compare two enzyme-order configurations, Tn5 → pA-Tn5 and pA-Tn5 → Tn5—alongside a CUT&Tag-only control, and are benchmarked against published single-cell ATAC-seq and CUT&Tag datasets. The Tn5 → pA-Tn5 condition produced the strongest and most continuous signal, with well-defined peaks concentrated in regulatory regions, particularly near *PYGO1* and *NEDD4*. This signal closely mirrors that of the published ATAC-seq dataset [51], suggesting that early ATAC tagmentation effectively captures open chromatin while preserving the accessibility landscape for subsequent antibody-directed pA-Tn5 tagging. This configuration also yielded higher estimated library complexity and greater peak density, indicating enhanced sensitivity and coverage breadth.

In contrast, the pA-Tn5 → Tn5 configuration exhibited markedly weaker and sparser peaks, with reduced signal amplitude (Fig. 4G). This suggests that early histone binding by pA-Tn5 may limit subsequent accessibility or compromise open chromatin integrity, reducing tagmentation efficiency. The CUT&Tag control, which lacks the ATAC tagmentation step, showed expected histone mark (H3K27me3) as compared with the published data [52].

When compared to published CUT&Tag datasets [52], both dual-tagmentation strategies extended beyond canonical repressive domains into intergenic and putative enhancer regions, underscoring the multimodal capacity of txci-CAB. The enhanced peak distribution and coverage in the Tn5-first condition (Fig. 4G) demonstrates that integrating ATAC-seq preprocessing with antibody-guided pA-Tn5 results in superior resolution, complexity, and signal-to-noise ratio, without compromising epitope targeting.

Collectively, these results support the model that initiating chromatin profiling with Tn5 tagmentation enhances accessibility mapping while preserving histone-bound context for downstream CUT&Tag. The optimized txci-CAB workflow allows simultaneous profiling of active and repressive chromatin states within a single cell, offering a streamlined yet comprehensive approach for single-cell epigenomics.



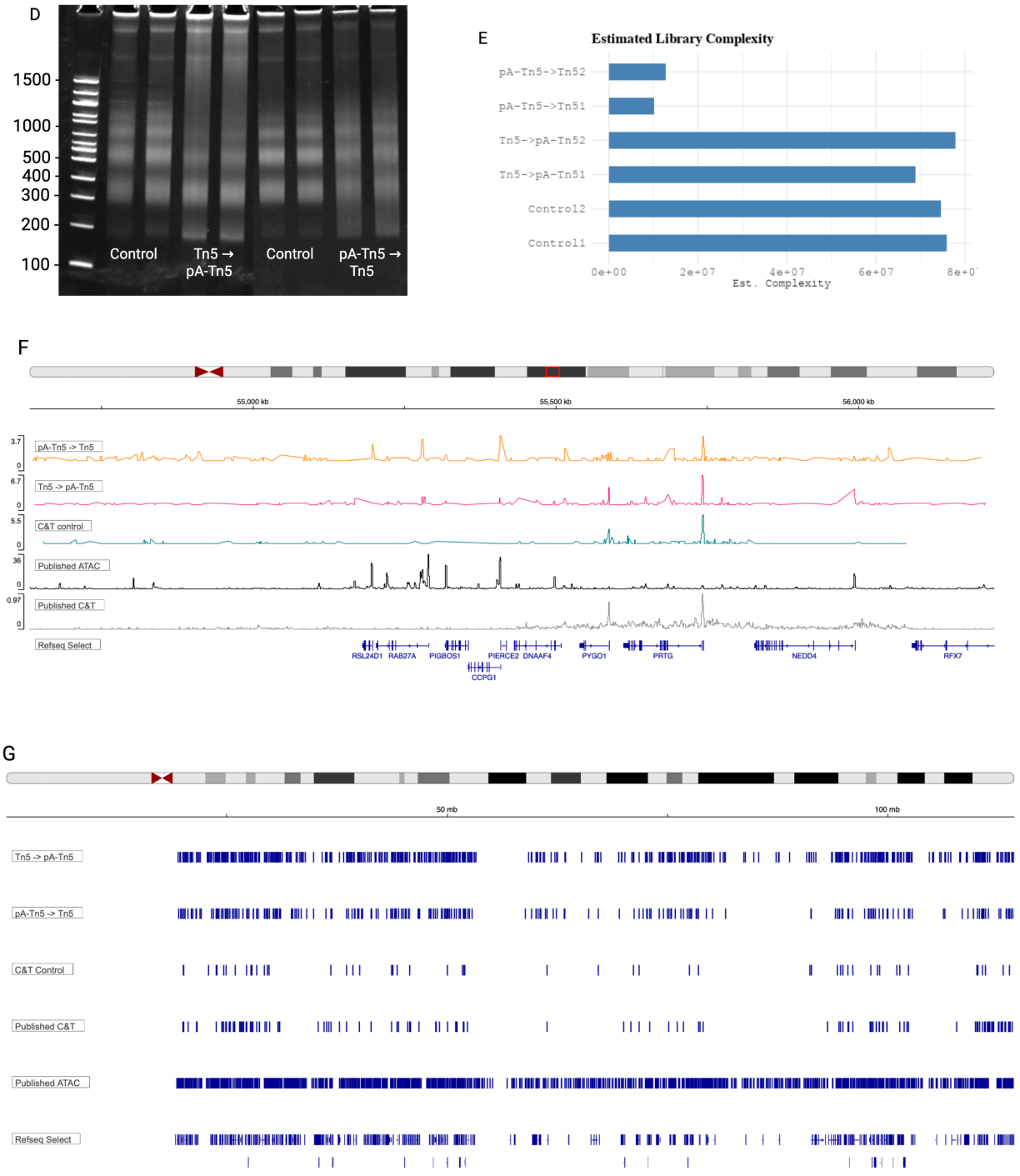


Fig. 4: Experimental schematic, QC and data analysis of Tn5 pre-tagmentation. A. Schematic of experimental workflow illustrating the sequential binding of Tn5 (loaded with Me-B adapters), antibodies, and pA-Tn5 (loaded with ME-A adapters). B. qPCR amplification curves comparing control CUT&Tag libraries (green) to those generated using the dual-enzyme

workflow Tn5 -> pA-Tn5 (blue). C. qPCR amplification curves comparing control CUT&Tag libraries (green) to those generated using pA-Tn5 -> Tn5 workflow from previous experiment (blue). D. Gel electrophoresis showing DNA fragment distributions from the control and experimental conditions (Tn5 → pA-Tn5 and pA-Tn5 → Tn5) alongside a 1 kb DNA ladder. E. Estimated library complexity of control and experimental samples, including technical replicates. F. Genome browser tracks (GBT) generated using IGV, showing chromatin profiling from experimental samples and control alongside reference ATAC-seq and CUT&Tag data (chr15:54,630,498–56,224,111). G. Zoomed-out genome browser view of chromosome 13 highlighting peak distributions in experimental samples and control, compared to published datasets.

2.4 txci-CAB Enables Single-Cell Profiling of Histone Modifications and Chromatin Accessibility Using Indexed Transposition and Microfluidics

We implemented txci-CAB as a single-cell method to simultaneously assess histone modifications and chromatin accessibility, leveraging indexed transposition coupled with droplet-based microfluidics. The methodology followed our standard txci-CAB protocol, incorporating elements from txci-ATAC [48] to enable efficient single-cell resolution. For initial testing, we employed 12 barcodes instead of a full 96-well plate, using barcoded Tn5ME-B pre-indexing in an ATAC-like manner, as exemplied in the Nano-CT protocol. The chromatin profiling was performed using the following antibodies: H3K27me3 for heterochromatin (closed chromatin), H3K27ac for euchromatin (open chromatin), CTCF for chromatin boundary and insulator sites, and IgG as a negative control.

Sequencing analysis validated the robust performance of the 10x Chromium platform, with accurate single-cell barcode recovery and a low doublet rate as confirmed by a species-mixing ("barnyard") experiment and the majority of cells mapping cleanly to either human or mouse, indicating good species separation. Single-cell experiment with collision rate of ~7.5% is in acceptable range as shown in figure 5A (Right panel). Left panel describes read distribution per barcode with number of reads per barcode or cell on x-axis and number of barcodes per cell with a given read count on y-axis. We found that most barcodes have a relatively low number of reads and the range spans from 54 to 4682, which is quite low given the assignment of ~10,000 reads per cell. This distribution suggests that many cells were sequenced at low-depth which could limit downstream analyses.

In order to distinguish between real cell barcodes from background noise or empty droplets, we generated a knee plot (Fig. 5B). In this plot, barcodes are ranked on the x-axis by the number of

associated UMIs or fragments (from highest to lowest), while the y-axis shows the corresponding UMI counts. The plot displays a smooth decline with a clear inflection point often referred to as the "knee", suggesting effective separation between real cells and background noise. This indicates that barcode quality filtering is functioning well and aligns with the low collision rate previously observed, suggesting good 10x barcoding. However, the overall UMI counts per cell appear lower than expected, with the top barcodes reaching only around 1000 UMIs. This reflects the earlier observation of low median reads per cell, pointing to low per-cell complexity. Potential contributing factors include suboptimal tagmentation, over-fragmentation, or insufficient sequencing depth. While the barcoding appears robust, optimizing upstream steps could improve data quality by increasing the number of fragments captured per cell.

We next generated genome browser tracks to discern if the txci-CAB fragment enrichment aligns with the previously published scCUT&Tag dataset [53] (Fig. 5C). Each row in the figure represents a track displaying chromatin signal across a shared genomic region on chromosome 12, with peaks denoting areas of enrichment. The CTCF tracks (top two) include txci-CAB CTCF (hot pink), which shows sharp, narrow peaks indicative of transcription factor binding and notably exhibits stronger signal intensity than the reference dataset (light pink). While both datasets show strong concordance in peak locations, validating txci-CAB's accuracy. Following that, H3K27ac tracks (lavender) show enhancer and promoter-associated activity, where txci-CAB H3K27ac (purple) closely matches broader signal seen in the published dataset (light blue), indicating successful capture of active regulatory elements. The H3K27me3 tracks (orange/brown) present a contrast: txci-CAB H3K27me3 (brown) reveals narrower and less abundant peaks compared to the broader, repressive domains captured in the published dataset. One possible explanation for this difference is variation in chromatin accessibility, txci-CAB may have reduced sensitivity in more compact, heterochromatic regions, which are typically enriched for H3K27me3. Alternatively, the observed discrepancy could reflect biological variability or differences in protocol efficiency when targeting broad histone marks. The published ATAC-seq track (gray, bottom) marks open chromatin regions and serves as an orthogonal reference to confirm overlap between accessible chromatin and regulatory features. Gene annotations along the bottom track link observed peaks to potential regulatory targets.

Collectively, the increased signal strength observed in txci-CAB implies that this method could offer improved resolution or efficiency in capturing transcription factor occupancy compared to existing single-cell CUT&Tag approaches. These results underscore the feasibility of integrating txci-CAB with high-throughput microfluidic platforms to enable multimodal analysis of chromatin accessibility and histone modifications at single-cell resolution.

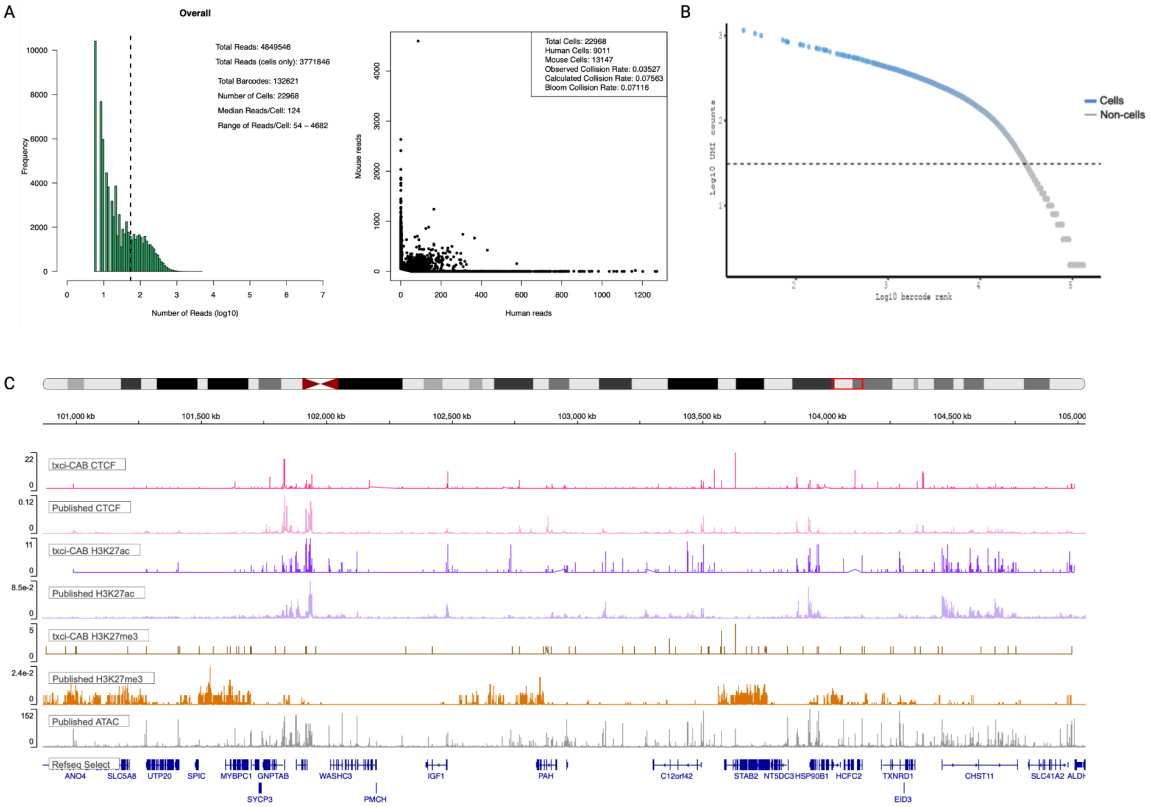


Fig. 5: Data analysis of txci-CAB experiment. A. (Left) Plot showing the distribution of read counts per cell, with frequency on the y-axis and number of reads on the x-axis, representing overall single-cell data quality. (Right) Scatter plot displaying individual cells based on species-specific alignment, with mouse and human reads plotted to assess cross-species contamination. Each dot represents a single cell. B. Knee plot with barcode rank on the x-axis and UMI counts on the y-axis with log10 scale and showing anything above the threshold line as real cells and below it as non-cells. C. Genome browser tracks (GBT) generated using IGV, showing chromatin profiling from experimental samples and control alongside reference ATAC-seq and scCUT&Tag dataset (chr12: ~101 Mb–105 Mb).

3. Discussion

The coordinated regulation of chromatin accessibility and histone modifications underlies the epigenetic control of gene expression, cell identity, and lineage specification. Traditional single-cell epigenomic techniques have provided powerful insights into these layers individually, yet direct and simultaneous measurement of both modalities within the same cell has remained technically challenging. In this thesis, we described txci-CAB, a novel, scalable, and droplet-compatible platform that integrates CUT&Tag and ATAC-seq chemistry through combinatorial indexing, enabling high-throughput, single-cell multimodal profiling of various epigenetic states.

Our results demonstrate that the txci-CAB protocol offers several key advantages over existing multimodal strategies. First, by systematically optimizing buffer compatibility and the sequential order of tagmentation enzymes, we achieved robust and reproducible profiling of both chromatin accessibility and histone modifications. The finding that Tn5 pre-tagmentation followed by pA-Tn5 yields higher library complexity, reduced duplication rates, and more defined nucleosomal laddering underscores the importance of enzyme ordering in preserving the biological signal and minimizing technical artifacts. Furthermore, our qPCR and sequencing data reveal that early Tn5 activity enables more efficient barcoding and enhanced accessibility profiling, as evidenced by broader peak coverage compared to the reverse configuration. Because of Tn5 pre-tagmentation, ATAC-seq signal should not be affected by other histone modalities measured next. Therefore, CUT&Tag signal is bound by this one thing. Because we can only sequence and analyse ~200-300 base pair long sequences, all the CUT&Tag data is thought to be towards ATAC-seq signal in case of the bulk-CAB experiment where we tested anti-H3K27me3. In the results we could see that ATAC and H3K27me3 peaks were co-existing in certain regions (Fig. 4F). Great thing is although there might be a shift in CUT&Tag signal, all the CUT&Tag peaks were still intact and therefore, our approach can profile histone modifications in addition to chromatin accessibility properly.

Importantly, initiating with ATAC-seq does not interfere with subsequent histone profiling, whereas starting with CUT&Tag may constrain the resolution of downstream accessibility detection due to steric hindrance or chromatin compaction effects. In bulk-CAB experiments using H3K27me3 antibodies, we observed co-occurrence of accessibility and histone modification peaks (Fig. 4F), demonstrating that our approach captures both modalities effectively, even in regions marked by repressive histone marks. Although, it would be expected that CUT&Tag signal may exhibit a positional shift when following ATAC tagmentation due to the restrain from one end of the fragment, the underlying peak structure remains intact, supporting the protocol's fidelity in profiling histone modifications alongside open chromatin.

Compared to previously established techniques such as nano-CT or multi-CUT&Tag, txci-CAB retains simplicity and flexibility and enhances throughput. Our approach does not rely on

complex antibody-enzyme fusion engineering or nanobody validation and is compatible with standard 10x Genomics infrastructure. This makes txci-CAB more accessible to a broader community of researchers and lowers the technical barriers to adoption.

Although the single-cell implementation of txci-CAB demonstrated promising results, some limitations were observed—most notably, lower-than-expected sequencing depth, which contributed to reduced per-cell library complexity. txci-CAB libraries showed a high correlation with published reference datasets, validating the specificity and accuracy of the method. When applied to diverse chromatin marks such as H3K27ac, H3K27me3, and CTCF, txci-CAB captured both promoter-enhancer dynamics and repressive domains within the same nuclei, providing an integrated snapshot of regulatory landscapes with increased signal strength in comparison to existing single-cell CUT&Tag approaches.

These observations point to the need for further protocol optimization, particularly around tagmentation efficiency and amplification steps. Due to project time constraints, we were unable to perform a full-scale single-cell experiment, but scaling up to larger datasets, such as 96-well plate formats, will be essential for evaluating txci-CAB's throughput, consistency, and resolution across diverse cellular states. Additionally, our initial validation was conducted in well-characterized cell lines, which are more permissive to nuclei isolation and antibody labeling. Translating txci-CAB to biologically relevant primary tissues will require further refinement, especially in maintaining nuclear integrity and ensuring robust antibody performance under more fragile conditions. Nevertheless, the current results establish txci-CAB as a compelling, integrated approach for simultaneous profiling of transcription factors, histone modifications, and chromatin accessibility at single-cell resolution.

In summary, further optimization is needed, particularly for broad histone marks and primary tissue applications. txci-CAB represents a significant methodological advance in single-cell epigenomics. Its capacity to unify multiple layers of chromatin information within a single assay opens the door to more comprehensive and scalable analyses of gene regulation, with the potential to uncover novel regulatory mechanisms in development, disease, and cellular reprogramming.

4. Conclusion

In this study, we worked on developing and optimizing txci-CAB, a novel multimodal single-cell chromatin profiling method to jointly profile chromatin accessibility and histone modifications. Through systematic testing of enzyme compatibility, reaction conditions, and tagmentation order, we demonstrated that initiating with Tn5-based ATAC-seq followed by pA-Tn5 CUT&Tag yields higher library complexity, better signal quality, and more accurate representation of chromatin states. Our bulk-CAB experiment showed that this dual-tagmentation strategy effectively captures two chromatin modalities in a single-cell and aligns well with published ATAC-seq and CUT&Tag datasets.

txci-CAB maintains scalability and workflow simplicity, and is capable of resolving single-cell barcodes even when the 10x Genomics platform is overloaded with nuclei, demonstrating robust performance in high-throughput settings. We validated txci-CAB using well-characterized cell lines, showing low barcode collision rates in species-mixing experiments. Although limited by time constraints from scaling to a full single-cell experiment, the foundational results highlight txci-CAB's potential as a powerful platform for integrated epigenomic profiling.

5. Methods

Cell Culture conditions

All experiments were performed using the GM12878 lymphoblastoid cell line. Cells were cultured at 37 °C in a humidified incubator with 5% CO_2, using RPMI 1640 medium (Gibco, Cat. No. 11875-093) supplemented with 15% fetal bovine serum (FBS; Gibco, Cat. No. 10437-028) and 1× Penicillin-Streptomycin (Gibco, Cat. No. 15140-122). Cultures were maintained at a density of 300,000 cells/mL, with passaging every other day. Cells were used for experiments 4–5 days post-thaw.

The single-cell txci-CAB experiment was performed with 3T3 cells in addition to GM12878. Cells were cultured at 37 °C in a humidified incubator with 5% CO_2, using high-glucose DMEM with pyruvate (Gibco, Cat. No. 11995-065), supplemented with 10% bovine calf serum (BCS) and 1× Penicillin-Streptomycin. Cultures were passaged at a 1:20 dilution twice a week or when they reached approximately 80% confluency.

Tn5 loading

For Tn5 loading [54], lyophilized linker oligonucleotides were obtained from IDT and resuspended in annealing buffer (50 mM NaCl (Invitrogen, Cat. No. AM9759), 40 mM Tris-HCl (pH 8.0; Invitrogen, Cat. No. 15568-025)) to a stock concentration of 100 µM. Equal volumes of Tn5ME-A or Tn5ME-B were mixed with Tn5MErev to a final working concentration of 50 µM. The oligonucleotide mix was annealed using a thermocycler with the following program: 95 °C for 5 minutes, followed by a slow ramp to 65 °C at 0.1 °C/sec, incubation at 65 °C for 5 minutes, and a second slow ramp down to 4 °C at 0.1 °C/sec. The annealed linkers were aliquoted in 10–20 µL volumes and stored at −20 °C. For double-ended Tn5 loading, 1 µL each of the annealed ME-A and ME-B (50 µM) was added to 20 µL of Tn5 enzyme stock (0.2–0.4 mg/mL), mixed thoroughly, and incubated at 23 °C for 30 minutes with shaking at 350 rpm. The loaded enzyme was stored at −20 °C. For single-ended loading, 2 µL of the ME-B linker was used instead of the two-linker combination, and the incubation conditions remained the same.

pA-Tn5 loading

To prepare pA-Tn5, a final concentration of 8 µM pA-Tn5 was mixed with 4 µM each of annealed Oligo A and Oligo B (1:1 molar ratio) in a single tube. The mixture was gently vortexed or flicked to mix and incubated at room temperature for 50 minutes. The loaded pA-Tn5 complex was then stored at −20 °C. In the case of single-ended pA-Tn5 loading, 8 µM of pA-Tn5 was combined with 8 µM of annealed Oligo A, using the same incubation and storage conditions.

The sequences of Tn5 linker oligos (5’ –> 3’)

Tn5ME-A 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3′

Tn5ME-B 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′

Tn5MErev 5′-[phos]CTGTCTCTTATACACATCT-3′

Bulk CUT&Tag

The bulk CUT&Tag assay was performed following the general protocol established by Kaya-Okur et al. [44], with modifications to accommodate processing in 1.5 mL microcentrifuge tubes. All wash and incubation steps were conducted in 200 µL volumes unless otherwise specified, and centrifugation steps utilized a swinging-bucket centrifuge with appropriate tube adaptors.

To begin, cultured cells (~ 2-3 million) were harvested from the flask and transferred into a 15 mL conical tube, then centrifuged at 500 × g for 5 minutes at 4 °C. The supernatant was aspirated, and the pellet was resuspended in an equal volume of phosphate-buffered saline (PBS) (GIBCO Cat. No. 10-010-023). The sample was centrifuged again under the same conditions, and the supernatant was aspirated. Cells were then resuspended in 500 µL of antibody buffer. Antibody buffer was made by supplementing 2X wash buffer. Wash buffer is 40mM HEPES (pH 7.5; Fisher Cat. No. AAJ60712AK), 300mM NaCl, 1mM spermidine (Sigma, Cat. No. S0266-1G), and EDTA-free protease inhibitor (Sigma Cat. No. 11836170001) prepared in nuclease-free water.. Wash buffer was made on the day of the experiment, stored in 4C and should be used within a week. The antibody buffer was made by adding 8nM EDTA (Invitrogen, Cat. No. AM9260G), 0.05% Digitonin (Invitrogen Cat. No. BN2006), 0.01% IGEPAL (Sigma, cat. no. I3021) and 1% BSA (Fisher, Cat. No. BP1600-100) to wash buffer. Detergent percentages reported are final concentrations. After resuspending the cell pellet with antibody buffer, they were incubated on ice for 3 minutes, and transferred to 1.5 mL low-bind microcentrifuge tubes. At this stage, centrifugation was performed using a fixed-angle rotor at 300 × g for 5 minutes at 4 °C. After aspirating the buffer, nuclei were resuspended in 200 µL of antibody buffer and centrifuged again at 600 × g for 3 minutes at 4 °C, followed by resuspension in 100 µL of antibody buffer. Nuclei were counted using a 50× dilution (2 µL nuclei + 48 µL Omni 2× TD buffer (20 mM Tris HCl (pH 7.5), 10 mM MgCl2 and 20% Dimethyl Formamide) + 50 µL trypan blue (Gibco Cat. No. 15-250-061)), and diluted to 200,000 nuclei per sample. A total volume of 200 µL was prepared by adding 2 µL (1:100 dilution in antibody buffer) of the primary antibody to the nuclei suspension. The tubes were sealed with parafilm and incubated overnight at 4 °C on a rotating wheel, positioned perpendicular to prevent buffer contact with the tube caps. The following day, nuclei were pelleted by centrifugation at 600 × g for 3 minutes at 4 °C, washed once with 200 µL Dig-Wash-BSA buffer. Dig-Wash-BSA buffer was made on the day to be used by adding 0.05% Digitonin, 0.01% IGEPAL and 1% BSA to wash buffer. Once washed, the cells were then resuspended in 200 µL of secondary antibody (1:100 dilution in

Dig-Wash-BSA buffer). Samples were incubated for 1 hour at room temperature on a rotating wheel. After secondary antibody incubation, nuclei were washed three times with 200 µL Dig-300-Wash-BSA buffer. Dig-300-Wash-BSA buffer is made by adding 150mM NaCl to Dig-Wash-BSA buffer. Nuclei are centrifuged each time at 600 × g for 3 minutes at 4 °C and then incubated for 1 hour with 200 µL of pA-Tn5 (1:250 dilution in Dig-300-Wash-BSA buffer), again on a rotating wheel at room temperature. This was followed by three additional washes in Dig-300-Wash-BSA buffer, using centrifugation at 300 × g for 3 minutes at 4 °C. After the final wash, nuclei were resuspended in 200 µL of tagmentation buffer and split into two technical replicates of 100 µL each in PCR tubes. Tagmentation buffer was made by adding 0.05% Digitonin, 0.01% IGEPAL, 150mM NaCl and 10mM $MgCl_2$ (Invitrogen, cat. no. AM9530G) in wash buffer. Samples were incubated at 37 °C for 1 hour in a thermocycler, tapping tubes occasionally to prevent sedimentation. Following tagmentation, DNA was purified using a Zymo DNA Clean and Concentrator-5 columns with 5X binding buffer (Zymo, cat. no. D4004). 5 µL of purified DNA template was amplified in 25 µL PCR reaction containing NEBNext PCR master mix (1X final), 0.5X SYBR Green and 0.4 µM of Ad1 primer and 0.4 µM of N7 primer containing barcodes. Samples were amplified on a Bio-Rad CFX Connect Real-time cycler using a PCR program consisting of an initial extension at 72 °C for 5 minutes, denaturation at 98 °C for 30 seconds, followed by repeated cycles of 98 °C for 10 seconds and 63 °C for 10 seconds, and a final extension at 72 °C for 1 minute. Amplification was monitored in real time, and reactions were stopped when amplification curves plateaued. A 1 µL aliquot of each PCR product was used to assess fragment distribution via gel electrophoresis. AMPure XP bead (Beckman, Cat. No. A63881) cleanup and size selection was used to purify the PCR products. A two-step size selection was performed to remove fragments larger than ~1500 bp (using 0.4× AMPure XP beads) and smaller than ~100 bp (using 1.5× beads). To begin, 25 µl of Elution buffer (Qiagen, Cat. No. 19086) was added to each PCR reaction to bring the volume to 50 µl. Then, 20 µl of AMPure XP beads were added, mixed thoroughly by pipetting, and incubated for 5 minutes at room temperature. The tubes were placed on a magnetic stand, and 68 µl of the supernatant was carefully transferred to a new 1.5 ml tube. For the second size selection, 53 µl of beads were added to the 68 µl supernatant and mixed thoroughly. (This amount was calculated based on estimating 19.4 µl of bead buffer and 48.6 µl of sample in the transferred volume, requiring a total of 72.9 µl beads for 1.5× selection; the remaining 53 µl was added to reach that amount.) After another 5-minute incubation at room temperature, the samples were placed back on the magnet to bind the DNA. The supernatant was removed, and the beads were washed twice with 200 µl of freshly prepared 80% ethanol. After the final ethanol wash, beads were briefly spun down, residual ethanol was removed, and beads were air-dried for 1 minute on the magnet. They were then resuspended in 22 µl of Buffer EB, incubated for 2 minutes at room temperature, and placed on the magnet again. Finally, 20 µl of the eluted DNA was transferred to a new tube. Following this double-sided size selection, CUT&Tag libraries were visualized on a 6% PAGE gel prior to sequencing.

ATAC protocol to evaluate efficiency of Tn5 in CUT&Tag buffers

To assess the efficiency of Tn5 transposase in CUT&Tag-compatible buffers, we performed a modified Omni-ATAC-seq protocol [55]. Cells were first harvested from culture and washed with 4 mL of 1X PBS containing 0.04% BSA (PBSB). The suspension was centrifuged at 500 × g for 5 minutes at 4 °C using a swinging-bucket centrifuge. The pellet was then resuspended in 200 µL of Lysis Buffer. Lysis buffer was prepared by supplementing ATAC Resuspension Buffer (RSB) with detergents as described below. RSB consists of 10 mM Tris-HCl (pH 7.5; Invitrogen, cat. no. 15567027), 10 mM NaCl, and 3 mM $MgCl_2$ (Invitrogen, cat. no. AM9530G), prepared in nuclease-free water. RSB was made in bulk, filtered, and stored at 4 °C for long-term use. On the day of the experiment, ATAC lysis buffer was freshly prepared by adding 0.1% IGEPAL, 0.01% digitonin, and 0.1% Tween-20 (Bio-Rad, cat. no. 1610781) to RSB. After resuspending cell pellets in lysis buffer, they are incubated on ice for 3 minutes. To stop the lysis, 1 mL of Stop Solution was added. Stop solution is made by 1 ml RSB containing 0.1% Tween-20. The nuclei were pelleted by centrifugation at 500 × g for 10 minutes at 4 °C using a fixed-angle centrifuge. Nuclei were subsequently resuspended in 100 µL of 1× PBSB. Nuclear concentration was estimated using a hemocytometer with a 50× dilution factor (2 µL nuclei + 48 µL 2× Omni TD buffer + 50 µL trypan blue). For each reaction, 12.4 µL of CUT&Tag Tagmentation Buffer () was pipetted into a PCR tube, followed by the addition of 6.6 µL containing 20,000 nuclei. Then, 1 µL of Tn5 transposase was added, and the tubes were incubated at 37 °C for 30 minutes in a thermocycler. Following transposition, DNA was purified using the Zymo DNA Clean & Concentrator-5 kit, using five volumes of DNA binding buffer. DNA was eluted in 12.5 µL of the elution buffer provided with the kit. A 5 µL aliquot of purified DNA was used for amplification with a qPCR program consisting of 72°C for 5 min; 98°C for 30s; cycling at 98°C for 10 s, 63°C for 30 s, 72°C for 1 min. The reaction concluded with a final extension at 72 °C for 1 minute. Amplification progress was monitored in real time, and reactions were stopped when the amplification curves plateaued. Finally, 1 µL of the PCR product was analyzed by gel electrophoresis to evaluate the fragment size distribution and confirm tagmentation efficiency.

Testing ME-B-Tn5 with ME-A-pA-Tn5 in CUT&Tag protocol

To evaluate the combinatorial use of ME-B-loaded Tn5 (seTn5) following ME-A-loaded pA-Tn5 in the CUT&Tag workflow, nuclei preparation and antibody incubations were first carried out as described in the bulk CUT&Tag protocol. After incubation with pA-Tn5, tagmentation was performed for 1 hour at 37 °C in a thermocycler. Upon completion of the reaction, the entire sample volume was transferred to a 1.5 mL microcentrifuge tube and centrifuged at 300 × g for 3 minutes at 4 °C. The supernatant was aspirated, and the pellet was resuspended in 12.4 µL of transposition mix (10 µl 2X Omni TD buffer, 0.2 µl of 1% digitonin, 0.2 µl of 10% Tween-20,

and 2 µl of H2O) and 6.6 µL of 1x PBSB. PBSB is 1X PBS supplemented with 0.04% BSA. The sample was then transferred into PCR tubes. To initiate the second round of tagmentation, 1 µL of ME-B-loaded Tn5 (seTn5) was added to each reaction, and the samples were incubated for 30 minutes at 37 °C in a thermocycler. Following this incubation, DNA was purified using the Zymo DNA Clean & Concentrator-5 kit, and samples were eluted according to the manufacturer's instructions. Further steps like qPCR amplification, gel electrophoresis and AMPure XP beads clean-up were all done according to the Bulk CUT&Tag protocol.

Bulk CAB

For bulk CAB experiments, approximately 3×10^6 cells were harvested from culture and washed with 4 mL PBSB. Cells were pelleted by centrifugation at 500 RCF for 5 minutes at 4 °C using a swinging-bucket centrifuge. Following supernatant removal, the pellet was resuspended in 300 µL of lysis Buffer and incubated on ice for 3 minutes. To stop the lysis reaction, 1 mL of Stop Solution was added, and the tube was gently inverted three times. Lysed cells were then pelleted at 500 RCF for 10 minutes at 4 °C using a fixed-angle centrifuge, and the nuclei were resuspended in 100 µL of 1× PBSB. Nuclei were counted using a hemocytometer with a 50× dilution (2 µL nuclei + 48 µL of 2× Omni TD buffer + 50 µL Trypan blue). For each sample, 100,000 nuclei were calculated and diluted to a final volume of 33 µL using PBSB, with one extra sample accounted for to compensate for pipetting loss. To each PCR tube, 62 µL of transposition mix was added, followed by 33 µL of nuclei and 5 µL of seTn5. Tubes were incubated at 37 °C for 30 minutes in a thermocycler. Although improved library complexity has been reported with tagmentation in a thermomixer at 1,000 rpm (with heated lid), both approaches yielded equivalent results in our hands. After Tn5 tagmentation, the contents of five tubes were pooled into either one tube, depending on the experiment. The pooled nuclei were centrifuged for 10 minutes at 500 g at 4 °C. The supernatant was carefully removed while leaving a small volume to avoid disturbing the pellet. The pellet was resuspended in 200 µL of antibody buffer and centrifuged at 600 g for 3 minutes at 4 °C. The supernatant was aspirated, and nuclei were resuspended in 200 µL of primary antibody master mix, mixed thoroughly by pipetting, and incubated overnight at 4 °C on a rotating wheel with the tubes sealed in parafilm. Secondary antibody incubation and pA-Tn5 incubation were performed as described in the bulk CUT&Tag protocol, except that single-ended pA-Tn5 was used for tagmentation in place of the double-ended version. Subsequent steps, including DNA purification using Zymo Clean & Concentrator kits, qPCR amplification, agarose gel electrophoresis, and final clean-up using AMPure XP beads were all conducted as in the bulk CUT&Tag protocol.

Barnyard experiments

To estimate the total rate of barcode collisions during the txci-CAB experiment, two distinct barnyard strategies were implemented. In the first approach, referred to as the "true barnyard," human and mouse cells were combined in equal proportions within the same well prior to barcoded transposition, allowing for assessment of collision events occurring both before and after pooling. In contrast, the "pseudo-barnyard" approach involved performing barcoded transposition separately on human and mouse cells in individual wells, followed by pooling the nuclei. This setup enabled evaluation of collision events arising exclusively after the pooling step.

txci-CAB

Approximately 2–3 million GM12878 and 3T3 cells were harvested from culture, and nuclei isolation was performed following the same procedure described for the bulk CAB protocol. For the initial evaluation of the single-cell protocol, we employed 12 unique barcodes, assigning two wells per barcode, with each well containing approximately 100,000 nuclei. A barnyard strategy was implemented wherein the two wells sharing the same barcode were loaded with the same cell type, allowing for consistent tracking across barcodes (Fig. 6). In total, we used eight wells each for the individual cell types and eight additional wells for the “true barnyard” condition, in which human and mouse cells were mixed prior to processing (Table 1).To prepare each sample, 62 µL of transposition mix was added to each PCR tube, followed by 33 µL of isolated nuclei and 3 µL of Tn5 transposase loaded with indexed ME-B (well barcodes sequences in Table S1). The samples were incubated at 37 °C for 30 minutes in a thermocycler. For each antibody condition in the CUT&Tag portion, six reactions were pooled, combining three barcode groups per antibody, resulting in a final input of approximately 600,000 nuclei per sample. The pooled nuclei were centrifuged at 500 g for 10 minutes at 4 °C. Care was taken to gently remove the supernatant without disturbing the pellet, leaving behind a small residual volume to maintain pellet integrity. Secondary antibody and pA-Tn5 incubations were performed as described in the bulk CUT&Tag protocol, with the modification that single-ended pA-Tn5 was used instead of the double-ended version. Following pA-Tn5 incubation for 1 hour at 37 °C in a thermomixer, tagmentation was stopped by removing the samples from the heat block and adding 200 µL of STOP buffer, mixing thoroughly by pipetting up and down three times. During antibody incubation, tagmentation STOP buffer and 2× Diluted Nuclei Buffer (DNB) were prepared by diluting the 20× stock nuclei buffer (10x scATAC-seq, PN: 2000207). Aliquots of 2× DNB were stored at −20 °C for future use. Nuclei were pelleted by centrifugation at 300 g for 3 minutes, and the supernatant was discarded. The pellet was washed twice by resuspending in 200 µL of 1× DNB supplemented with BSA, followed each time by centrifugation at 300 g for 3 minutes. After the final wash, most of the supernatant was removed, and the nuclei were left in approximately 10 µL of buffer. An additional 10 µL of 1× DNB+BSA was added to reach a final

volume of 20 µL. The volume was confirmed using a P20 pipette. Nuclei concentration was determined by mixing 2 µL of the suspension with 8 µL of trypan blue and counting in duplicate to ensure accuracy. To target the recovery of 55,000 nuclei, 100,000 nuclei were loaded onto the 10x Genomics chip according to the 10x Genomics Chromium Next GEM Single Cell ATAC Reagent Kits v1.1 protocol with some specific modifications in further steps. During the GEM generation and barcoding step (2.1a), a 14 µl nuclei solution containing the desired amount of nuclei in the loading buffer () was then combined with 1 µl of 75 µM short SBS oligo to enable exponential amplification within droplets during in-droplet PCR. After GEM PCR (Step 2.5a), 10 µl (10%) of the reaction was gently transferred to a new PCR tube for separate processing alongside the remaining 90%. Post-GEM incubation cleanup was performed on both fractions, after which the 10% fraction was indexed (Step 4.1) using a modified PCR mix. This mix included 2.5 µl of 25 µM barcoded i7 TruSeq primer (Table S1), replacing the standard Single Index N Set A. Amplification was carried out on a Bio-Rad CFX Connect real-time thermocycler, and stopped once plateauing of the amplification curve was observed for 7 cycles. Each barcode yielded thousands of reads per cell, demonstrating successful amplification.

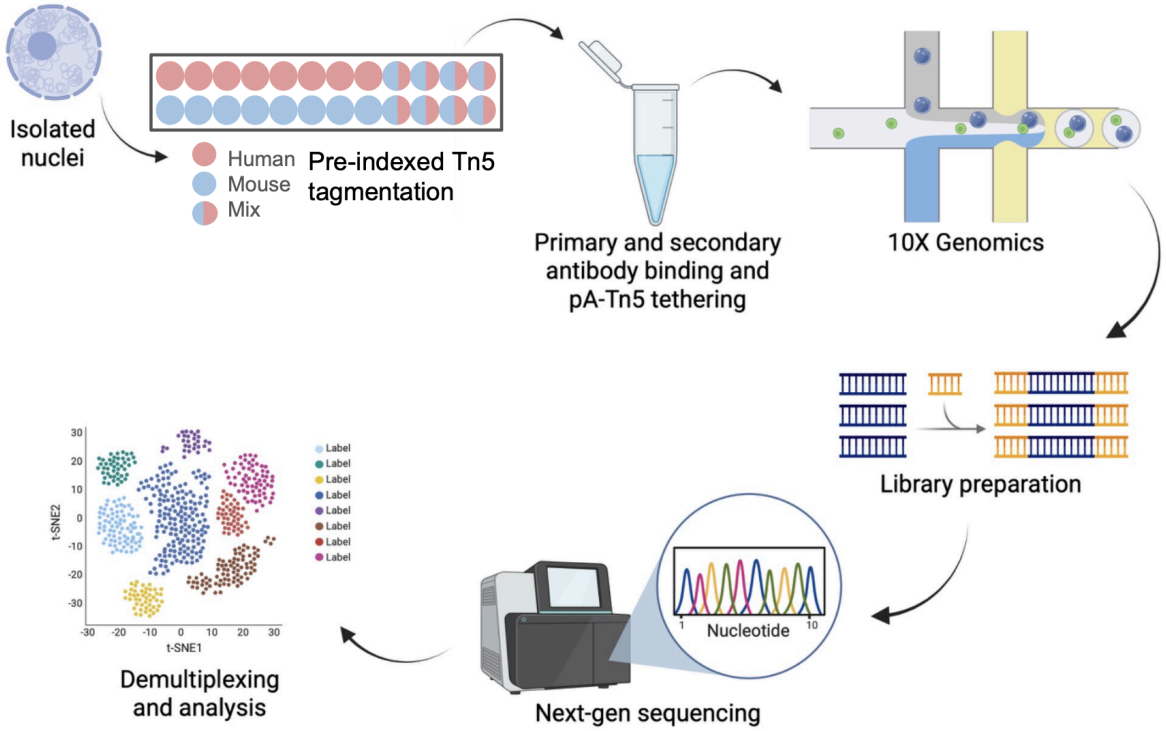


Fig. 6: Experimental design for txci-CAB performed with 12 barcodes (2 wells each).

Well barcode	Cell type	Primary antibody
Tn5_bc1	3T3	H3K27me3
Tn5_bc1	3T3	
Tn5_bc2	GM12878	
Tn5_bc2	GM12878	
Tn5_bc3	3T3+GM12878	
Tn5_bc3	3T3+GM12878	
Tn5_bc4	3T3	H3K27ac
Tn5_bc4	3T3	
Tn5_bc5	GM12878	
Tn5_bc5	GM12878	
Tn5_bc6	3T3+GM12878	
Tn5_bc6	3T3+GM12878	
Tn5_bc7	3T3	CTCF
Tn5_bc7	3T3	
Tn5_bc8	GM12878	
Tn5_bc8	GM12878	
Tn5_bc9	3T3+GM12878	
Tn5_bc9	3T3+GM12878	
Tn5_bc10	3T3	IgG
Tn5_bc10	3T3	
Tn5_bc11	GM12878	
Tn5_bc11	GM12878	
Tn5_bc12	3T3+GM12878	
Tn5_bc12	3T3+GM12878	

Table 1: Experimental design for txci-CAB describing cell types (barnyard) and primary antibody associated with specific well barcodes.

Sequencing

Sequencing of CUT&Tag libraries generated from the initial experiment involving Tn5 and pA-Tn5 in GM12878 cells was performed on the Illumina NextSeq 550 platform using the Mid Output Kit, generating 2 × 76 bp paired-end reads. Libraries from the bulk CAB experiment

were sequenced on the Illumina MiSeq platform with the same read configuration. For txci-CAB libraries derived from GM12878 and 3T3 cells, sequencing was conducted on the NextSeq platform with read lengths of 90 bp for Read 1 and 50 bp for Read 2, using the Mid Output Kit.

Sequencing analysis

The sequencing data from the CUT&Tag and bulk CAB experiments were processed using a pipeline analogous to standard ATAC-seq workflows.

The specific programs used for data analysis were as follows: Trimmomatic v0.36 [56], Bowtie2 v2.4.1 [57], SAMtools v1.10 [58], Picard v2.20.2 [59], MACS2 v2.1.2 [60], bedtools v2.29.2 [61], R v4.2.2 [62], Python v3.13.3 [63].

Initial quality control and adapter trimming were performed with Trimmomatic, applying the parameters: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:10, and MINLEN:20. Cleaned reads were aligned to the appropriate reference genome—either hg19 for human or mm10 for mouse—using Bowtie2. The alignment was performed with -X 2000 to set the maximum fragment length to 2000 bp and -3 1 to trim 1 base from the 3' end of each read, which helped mitigate potential alignment artifacts caused by perfect matches between read pairs.

Post-alignment processing was carried out using SAMtools to retain only properly paired reads (using -f 3 -F 12) with a mapping quality score (MAPQ) of 10 or higher, and reads mapping exclusively to assembled nuclear chromosomes. Duplicate reads were identified and removed using Picard (MarkDuplicates function), which also provided library complexity estimates. Peak calling and further downstream analysis were conducted using MACS2, bedtools, and custom scripts in R.

Data analysis for txci-CAB libraries was performed using a pipeline adapted from the txci-ATAC-seq GitHub repository. Adapter trimming, read alignment, and initial filtering steps were conducted following the same procedures used for bulk libraries. This included the use of Trimmomatic, Bowtie2, and SAMtools for preprocessing and mapping. Subsequent steps specific to single-cell combinatorial indexing such as barcode correction, deduplication, and generation of read count matrices were carried out using custom Python and R scripts tailored for txci-style indexing strategies. These scripts enabled accurate assignment of reads to individual cells and the construction of high-quality chromatin accessibility and histone modification profiles across thousands of single nuclei.

Genome browser tracks for the txci-CAB libraries were generated using the Integrative Genomics Viewer (IGV) [64]. BAM files were first converted to BEDGRAPH format, sorted, and subsequently transformed into BigWig files using standard tools. These BigWig files were then uploaded to the IGV platform for visualization. Tracks were displayed alongside previously published reference datasets to facilitate comparative analysis of chromatin accessibility and histone modification profiles.

Appendix - Supplementary Table

Well barcode	Sequence	Barcode
Tn5_bc1	CGTGTGCTCTTCCGATCTGAACCGCGAGATGTGTATAAGAGACAG	GAACCGCG
Tn5_bc2	CGTGTGCTCTTCCGATCTAGGTTATAAGATGTGTATAAGAGACAG	AGGTTATA
Tn5_bc3	CGTGTGCTCTTCCGATCTTCATCCTTAGATGTGTATAAGAGACAG	TCATCCTT
Tn5_bc4	CGTGTGCTCTTCCGATCTCTGCTTCCAGATGTGTATAAGAGACAG	CTGCTTCC
Tn5_bc5	CGTGTGCTCTTCCGATCTGGTCACGAAGATGTGTATAAGAGACAG	GGTCACGA
Tn5_bc6	CGTGTGCTCTTCCGATCTAACTGTAGAGATGTGTATAAGAGACAG	AACTGTAG
Tn5_bc7	CGTGTGCTCTTCCGATCTGTGAATATAGATGTGTATAAGAGACAG	GTGAATAT
Tn5_bc8	CGTGTGCTCTTCCGATCTACAGGCGCAGATGTGTATAAGAGACAG	ACAGGCGC
Tn5_bc9	CGTGTGCTCTTCCGATCTCATAGAGTAGATGTGTATAAGAGACAG	CATAGAGT
Tn5_bc10	CGTGTGCTCTTCCGATCTTGCGAGACAGATGTGTATAAGAGACAG	TGCGAGAC
Tn5_bc11	CGTGTGCTCTTCCGATCTGACGTCTTAGATGTGTATAAGAGACAG	GACGTCTT
Tn5_bc12	CGTGTGCTCTTCCGATCTAGTACTCCAGATGTGTATAAGAGACAG	AGTACTCC

Table S1: Well barcodes for the txci-CAB experiment with the ME-B oligo sequence.

References

[1] S. Chatterjee and N. Ahituv, “Gene Regulatory Elements, Major Drivers of Human Disease,” *Annu. Rev. Genomics Hum. Genet.*, vol. 18, no. 1, pp. 45–63, Aug. 2017, doi: 10.1146/annurev-genom-091416-035537.

[2] F. Spitz and E. E. M. Furlong, “Transcription factors: from enhancer binding to developmental control,” *Nat. Rev. Genet.*, vol. 13, no. 9, pp. 613–626, Sep. 2012, doi: 10.1038/nrg3207.

[3] S. Ma *et al.*, “Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin,” *Cell*, vol. 183, no. 4, pp. 1103-1116.e20, Nov. 2020, doi: 10.1016/j.cell.2020.09.056.

[4] A. J. Bannister and T. Kouzarides, “Regulation of chromatin by histone modifications,” *Cell Res.*, vol. 21, no. 3, pp. 381–395, Mar. 2011, doi: 10.1038/cr.2011.22.

[5] W. Fischle, Y. Wang, and C David Allis, “Histone and chromatin cross-talk,” *Curr. Opin. Cell Biol.*, vol. 15, no. 2, pp. 172–183, Apr. 2003, doi: 10.1016/S0955-0674(03)00013-9.

[6] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, “Crystal structure of the nucleosome core particle at 2.8 Å resolution,” *Nature*, vol. 389, no. 6648, pp. 251–260, Sep. 1997, doi: 10.1038/38444.

[7] Z. Zhou, R. Yan, W. Jiang, and J. M. K. Irudayaraj, “Chromatin hierarchical branching visualized at the nanoscale by electron microscopy,” *Nanoscale Adv.*, vol. 3, no. 4, pp. 1019–1028, 2021, doi: 10.1039/D0NA00359J.

[8] B. Simpson, C. Tupper, and N. M. Al Aboud, “Genetics, DNA Packaging,” in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2025. Accessed: Apr. 09, 2025. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK534207/

[9] E. I. Campos and D. Reinberg, “Histones: Annotating Chromatin,” *Annu. Rev. Genet.*, vol. 43, no. 1, pp. 559–599, Dec. 2009, doi: 10.1146/annurev.genet.032608.103928.

[10] S. L. Klemm, Z. Shipony, and W. J. Greenleaf, “Chromatin accessibility and the regulatory epigenome,” *Nat. Rev. Genet.*, vol. 20, no. 4, pp. 207–220, Apr. 2019, doi: 10.1038/s41576-018-0089-8.

[11] L. Minnoye *et al.*, “Chromatin accessibility profiling methods,” *Nat. Rev. Methods Primer*, vol. 1, no. 1, p. 10, Jan. 2021, doi: 10.1038/s43586-020-00008-9.

[12] F. C. Grandi, H. Modi, L. Kampman, and M. R. Corces, “Chromatin accessibility profiling by ATAC-seq,” *Nat. Protoc.*, vol. 17, no. 6, pp. 1518–1552, Jun. 2022, doi: 10.1038/s41596-022-00692-9.

[13] Y. Li, “Modern epigenetics methods in biological research,” *Methods*, vol. 187, pp. 104–113, Mar. 2021, doi: 10.1016/j.ymeth.2020.06.022.

[14] M. Mehrmohamadi, M. H. Sepehri, N. Nazer, and M. R. Norouzi, “A Comparative Overview of Epigenomic Profiling Methods,” *Front. Cell Dev. Biol.*, vol. 9, p. 714687, Jul. 2021, doi: 10.3389/fcell.2021.714687.

[15] E. Meaburn and R. Schulz, “Next generation sequencing in epigenetics: Insights and challenges,” *Semin. Cell Dev. Biol.*, vol. 23, no. 2, pp. 192–199, Apr. 2012, doi: 10.1016/j.semcdb.2011.10.010.

[16] A. Chawla, C. Nagy, and G. Turecki, “Chromatin Profiling Techniques: Exploring the Chromatin Environment and Its Contributions to Complex Traits,” *Int. J. Mol. Sci.*, vol. 22, no. 14, p. 7612, Jul. 2021, doi: 10.3390/ijms22147612.

[17] S. A. Ruiz and C. S. Chen, “Emergence of Patterned Stem Cell Differentiation Within Multicellular Structures,” *Stem Cells*, vol. 26, no. 11, pp. 2921–2927, Nov. 2008, doi: 10.1634/stemcells.2008-0432.

[18] R. B. Deal and S. Henikoff, “A Simple Method for Gene Expression and Chromatin Profiling of Individual Cell Types within a Tissue,” *Dev. Cell*, vol. 18, no. 6, pp. 1030–1040,

Jun. 2010, doi: 10.1016/j.devcel.2010.05.013.
[19] Z. Fu, S. Jiang, Y. Sun, S. Zheng, L. Zong, and P. Li, “Cut&tag: a powerful epigenetic tool for chromatin profiling,” *Epigenetics*, vol. 19, no. 1, p. 2293411, Dec. 2024, doi: 10.1080/15592294.2023.2293411.
[20] B. Wilson and S. G. Nicholls, “The Human Genome Project, and recent advances in personalized genomics,” *Risk Manag. Healthc. Policy*, p. 9, Feb. 2015, doi: 10.2147/RMHP.S58728.
[21] T. Hu, N. Chitnis, D. Monos, and A. Dinh, “Next-generation sequencing technologies: An overview,” *Hum. Immunol.*, vol. 82, no. 11, pp. 801–811, Nov. 2021, doi: 10.1016/j.humimm.2021.02.012.
[22] K. R. Kumar, M. J. Cowley, and R. L. Davis, “Next-Generation Sequencing and Emerging Technologies*,” *Semin. Thromb. Hemost.*, vol. 50, no. 07, pp. 1026–1038, Oct. 2024, doi: 10.1055/s-0044-1786397.
[23] E. R. Mardis, “ChIP-seq: welcome to the new frontier,” *Nat. Methods*, vol. 4, no. 8, pp. 613–614, Aug. 2007, doi: 10.1038/nmeth0807-613.
[24] T. S. Furey, “ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions,” *Nat. Rev. Genet.*, vol. 13, no. 12, pp. 840–852, Dec. 2012, doi: 10.1038/nrg3306.
[25] P. J. Park, “ChIP–seq: advantages and challenges of a maturing technology,” *Nat. Rev. Genet.*, vol. 10, no. 10, pp. 669–680, Oct. 2009, doi: 10.1038/nrg2641.
[26] F. Greil, C. Moorman, and B. Van Steensel, “[16] DamID: Mapping of In Vivo Protein–Genome Interactions Using Tethered DNA Adenine Methyltransferase,” in *Methods in Enzymology*, vol. 410, Elsevier, 2006, pp. 342–359. doi: 10.1016/S0076-6879(06)10016-6.
[27] T. D. Southall and A. H. Brand, “Chromatin profiling in model organisms,” *Brief. Funct. Genomic. Proteomic.*, vol. 6, no. 2, pp. 133–140, Sep. 2007, doi: 10.1093/bfgp/elm013.
[28] M. Tsompana and M. J. Buck, “Chromatin accessibility: a window into the genome,” *Epigenetics Chromatin*, vol. 7, no. 1, p. 33, Dec. 2014, doi: 10.1186/1756-8935-7-33.
[29] A. R. Mansisidor and V. I. Risca, “Chromatin accessibility: methods, mechanisms, and biological insights,” *Nucleus*, vol. 13, no. 1, pp. 238–278, Dec. 2022, doi: 10.1080/19491034.2022.2143106.
[30] L. Song and G. E. Crawford, “DNase-seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells,” *Cold Spring Harb. Protoc.*, vol. 2010, no. 2, p. pdb.prot5384, Feb. 2010, doi: 10.1101/pdb.prot5384.
[31] K. Cui and K. Zhao, “Genome-Wide Approaches to Determining Nucleosome Occupancy in Metazoans Using MNase-Seq,” in *Chromatin Remodeling*, vol. 833, R. H. Morse, Ed., in Methods in Molecular Biology, vol. 833. , Totowa, NJ: Humana Press, 2012, pp. 413–419. doi: 10.1007/978-1-61779-477-3_24.
[32] J. D. Buenrostro, B. Wu, H. Y. Chang, and W. J. Greenleaf, “ATAC‐seq: A Method for Assaying Chromatin Accessibility Genome‐Wide,” *Curr. Protoc. Mol. Biol.*, vol. 109, no. 1, Jan. 2015, doi: 10.1002/0471142727.mb2129s109.
[33] H. Chen *et al.*, “Assessment of computational methods for the analysis of single-cell ATAC-seq data,” *Genome Biol.*, vol. 20, no. 1, p. 241, Dec. 2019, doi: 10.1186/s13059-019-1854-5.
[34] M. Bartosovic, M. Kabbe, and G. Castelo-Branco, “Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues,” *Nat. Biotechnol.*, vol. 39, no. 7, pp. 825–835, Jul. 2021, doi: 10.1038/s41587-021-00869-9.
[35] J. Liang, W. Cai, and Z. Sun, “Single-Cell Sequencing Technologies: Current and Future,” *J. Genet. Genomics*, vol. 41, no. 10, pp. 513–528, Oct. 2014, doi: 10.1016/j.jgg.2014.09.005.
[36] X. Tang, Y. Huang, J. Lei, H. Luo, and X. Zhu, “The single-cell sequencing: new

developments and medical applications," *Cell Biosci.*, vol. 9, no. 1, p. 53, Dec. 2019, doi: 10.1186/s13578-019-0314-y.
[37] L. Wen and F. Tang, "Recent advances in single-cell sequencing technologies," *Precis. Clin. Med.*, vol. 5, no. 1, p. pbac002, Feb. 2022, doi: 10.1093/pcmedi/pbac002.
[38] J. Ding *et al.*, "Systematic comparison of single-cell and single-nucleus RNA-sequencing methods," *Nat. Biotechnol.*, vol. 38, no. 6, pp. 737–746, Jun. 2020, doi: 10.1038/s41587-020-0465-8.
[39] M. Dufva, "Introduction to Microarray Technology," in *DNA Microarrays for Biomedical Research*, vol. 529, M. Dufva, Ed., in Methods in Molecular Biology, vol. 529. , Totowa, NJ: Humana Press, 2009, pp. 1–22. doi: 10.1007/978-1-59745-538-1_1.
[40] S.-Y. Teh, R. Lin, L.-H. Hung, and A. P. Lee, "Droplet microfluidics," *Lab. Chip*, vol. 8, no. 2, p. 198, 2008, doi: 10.1039/b715524g.
[41] S. A. Vitak *et al.*, "Sequencing thousands of single-cell genomes with combinatorial indexing," *Nat. Methods*, vol. 14, no. 3, pp. 302–308, Mar. 2017, doi: 10.1038/nmeth.4154.
[42] S. J. Wu *et al.*, "Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression," *Nat. Biotechnol.*, vol. 39, no. 7, pp. 819–824, Jul. 2021, doi: 10.1038/s41587-021-00865-z.
[43] M. P. Meers, T. D. Bryson, J. G. Henikoff, and S. Henikoff, "Improved CUT&RUN chromatin profiling tools," *eLife*, vol. 8, p. e46314, Jun. 2019, doi: 10.7554/eLife.46314.
[44] H. S. Kaya-Okur *et al.*, "CUT&Tag for efficient epigenomic profiling of small samples and single cells," *Nat. Commun.*, vol. 10, no. 1, p. 1930, Apr. 2019, doi: 10.1038/s41467-019-09982-5.
[45] M. Bartosovic and G. Castelo-Branco, "Multimodal chromatin profiling using nanobody-based single-cell CUT&Tag," *Nat. Biotechnol.*, vol. 41, no. 6, pp. 794–805, Jun. 2023, doi: 10.1038/s41587-022-01535-4.
[46] J. R. Bárcenas-Walls *et al.*, "Nano-CUT&Tag for multimodal chromatin profiling at single-cell resolution," *Nat. Protoc.*, vol. 19, no. 3, pp. 791–830, Mar. 2024, doi: 10.1038/s41596-023-00932-6.
[47] S. Gopalan, Y. Wang, N. W. Harper, M. Garber, and T. G. Fazzio, "Simultaneous profiling of multiple chromatin proteins in the same cells," *Mol. Cell*, vol. 81, no. 22, pp. 4736-4746.e5, Nov. 2021, doi: 10.1016/j.molcel.2021.09.019.
[48] H. Zhang *et al.*, "txci-ATAC-seq: a massive-scale single-cell technique to profile chromatin accessibility," *Genome Biol.*, vol. 25, no. 1, p. 78, Mar. 2024, doi: 10.1186/s13059-023-03150-1.
[49] M. S. Magri *et al.*, "Assaying Chromatin Accessibility Using ATAC-Seq in Invertebrate Chordate Embryos," *Front. Cell Dev. Biol.*, vol. 7, p. 372, Jan. 2020, doi: 10.3389/fcell.2019.00372.
[50] J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, and W. J. Greenleaf, "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position," *Nat. Methods*, vol. 10, no. 12, pp. 1213–1218, Dec. 2013, doi: 10.1038/nmeth.2688.
[51] Fuchou Tang, "ATAC-seq reference dataset." GEO accession, NCBI. [Online]. Available: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5825636
[52] xingqi chen, "CUT&Tag reference dataset." GEO accession, NCBI. [Online]. Available: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5530673
[53] Yuqing Guo, "scCUT&Tag reference dataset." doi: 10.1038/s41592-024-02453-w.
[54] S. Picelli, Å. K. Björklund, B. Reinius, S. Sagasser, G. Winberg, and R. Sandberg, "Tn5 transposase and tagmentation procedures for massively scaled sequencing projects," *Genome Res.*, vol. 24, no. 12, pp. 2033–2040, Dec. 2014, doi: 10.1101/gr.177881.114.
[55] M. R. Corces *et al.*, "An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues," *Nat. Methods*, vol. 14, no. 10, pp. 959–962, Oct. 2017, doi:

10.1038/nmeth.4396.
[56] A. M. Bolger, M. Lohse, and B. Usadel, “Trimmomatic: a flexible trimmer for Illumina sequence data,” *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, Aug. 2014, doi: 10.1093/bioinformatics/btu170.
[57] B. Langmead and S. L. Salzberg, “Fast gapped-read alignment with Bowtie 2,” *Nat. Methods*, vol. 9, no. 4, pp. 357–359, Apr. 2012, doi: 10.1038/nmeth.1923.
[58] H. Li *et al.*, “The Sequence Alignment/Map format and SAMtools,” *Bioinformatics*, vol. 25, no. 16, pp. 2078–2079, Aug. 2009, doi: 10.1093/bioinformatics/btp352.
[59] *Picard toolkit*. [Online]. Available: http://broadinstitute.github.io/picard/
[60] Y. Zhang *et al.*, “Model-based Analysis of ChIP-Seq (MACS),” *Genome Biol.*, vol. 9, no. 9, p. R137, Sep. 2008, doi: 10.1186/gb-2008-9-9-r137.
[61] A. R. Quinlan and I. M. Hall, “BEDTools: a flexible suite of utilities for comparing genomic features,” *Bioinformatics*, vol. 26, no. 6, pp. 841–842, Mar. 2010, doi: 10.1093/bioinformatics/btq033.
[62] *R Core Team*. [Online]. Available: https://www.R-project.org/
[63] *Python*. [Online]. Available: https://www.python.org/
[64] H. Thorvaldsdottir, J. T. Robinson, and J. P. Mesirov, “Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration,” *Brief. Bioinform.*, vol. 14, no. 2, pp. 178–192, Mar. 2013, doi: 10.1093/bib/bbs017.