

Video Article

# ATAC-seq Assay with Low Mitochondrial DNA Contamination from Primary Human CD4+ T Lymphocytes

Hannah D. Rickner<sup>1</sup>, Sheng-Yong Niu<sup>1,2</sup>, Christine S. Cheng<sup>1,3</sup>

<sup>1</sup>Department of Biology, Boston University

<sup>2</sup>Department of Computer Science and Engineering, University of California San Diego

<sup>3</sup>Bioinformatics Program, Boston University

Correspondence to: Christine S. Cheng at [chcheng@bu.edu](mailto:chcheng@bu.edu)

URL: <https://www.jove.com/video/59120>

DOI: [doi:10.3791/59120](https://doi.org/10.3791/59120)

Keywords: Epigenomics, ATAC-seq, Tn5, CD4+ lymphocytes, nuclei, mitochondrial DNA contamination

Date Published: 12/13/2018

Citation: Rickner, H.D., Niu, S.Y., Cheng, C.S. ATAC-seq Assay with Low Mitochondrial DNA Contamination from Primary Human CD4+ T Lymphocytes. *J. Vis. Exp.* (), e59120, doi:10.3791/59120 (2018).

## Abstract

ATAC-seq has become a widely used methodology in the study of epigenetics due to its rapid and simple approach to mapping genome-wide accessible chromatin. In this paper we present an improved ATAC-seq protocol that reduces contaminating mitochondrial DNA reads. While previous ATAC-seq protocols have struggled with an average of 50% contaminating mitochondrial DNA reads, the new lysis buffer introduced in this protocol reduces mitochondrial DNA contamination to an average of 3%. This improved ATAC-seq protocol allows for a near 50% reduction in the sequencing cost. We demonstrate how these high-quality ATAC-seq libraries can be prepared from activated CD4+ lymphocytes, providing step-by-step instructions from CD4+ lymphocyte isolation from whole blood through data analysis. This improved ATAC-seq protocol has been validated in a wide range of cell types and will be of immediate use to researchers studying chromatin accessibility.

## Introduction

The assay for transposase-accessible chromatin sequencing (ATAC-seq) has rapidly become the leading method for interrogating chromatin architecture. ATAC-seq can identify regions of accessible chromatin through the process of tagmentation, the fragmenting and tagging of DNA by the same enzyme, to produce libraries with which sequencing can determine chromatin accessibility across an entire genome. This tagmentation process is mediated by the hyperactive Tn5 transposase, which only cuts open regions of chromatin due to nucleosomal steric hindrance. As it cuts, the Tn5 transposase also inserts sequencing adapters that allow for rapid library construction by PCR and next-generation sequencing of genome-wide accessible chromatin<sup>1,2</sup>.

ATAC-seq has become the preferred method to determine regions of chromatin accessibility due to the relatively simple and fast protocol, quality and range of information that can be determined from its results, and small amount of starting material required. Compared to DNase-seq<sup>3</sup> (which also measures genome-wide chromatin accessibility), MNase-seq<sup>4</sup> (which determines nucleosome positions in open genome regions), and the formaldehyde-mediated FAIRE-seq<sup>5</sup>, ATAC-seq is faster, cheaper, and more reproducible<sup>1</sup>. It is also more sensitive, working with starting material of as few as 500 nuclei, compared to the 50 million nuclei required for DNase-seq<sup>3</sup>. ATAC-seq also has the ability to provide more information about chromatin architecture than other methods, including regions of transcription factor binding, nucleosome positioning, and open chromatin regions<sup>1</sup>. Effective, single-cell ATAC-seq protocols have been validated, providing information on chromatin architecture at the single-cell level<sup>6,7</sup>.

ATAC-seq has been used to characterize chromatin architecture across a wide spectrum of research and cells types, including plants<sup>8</sup>, humans<sup>9</sup>, and many other organisms. It has also been critical in identifying epigenetic regulation of disease states<sup>7</sup>. However, the most widely used ATAC-seq protocol includes the major drawback of contamination sequencing reads from mitochondrial DNA. In some data sets, this contamination level can be as high as 60% of sequencing results<sup>1</sup>. Here we present an improved ATAC-seq protocol that reduces the mitochondrial DNA contamination rate to just 3%, allowing reduction of around 50% in sequencing costs<sup>10</sup>. This is made possible by a streamlined process of CD4+ lymphocyte isolation and activation and an improved lysis buffer that is critical in minimizing mitochondrial DNA contamination.

This improved ATAC-seq protocol has been validated with a wide range of primary cells, including human primary peripheral blood mononuclear cells (PBMCs)<sup>10</sup>, human primary monocytes, and mouse dendritic cells (unpublished). It has also been used successfully to interrogate melanoma cell lines in a clustered regularly interspaced short palindromic repeats (CRISPR) screen of non-coding elements<sup>11</sup>. Additionally, the data analysis package described in this protocol and provided on GitHub provides new and experienced researchers with tools to analyze ATAC-seq data. ATAC-seq is the most effective assay to map chromatin accessibility across an entire genome, and modifications to the existing protocol that are introduced here will allow researchers to produce high-quality data with low mitochondrial DNA contamination, reducing sequencing costs and improving ATAC-seq throughput.

## Protocol

This improved protocol provides step-by-step instructions for performing ATAC-seq of CD4<sup>+</sup> lymphocytes, from the starting material of whole blood through data analysis (**Figure 1**).

### 1. Isolation of CD4<sup>+</sup> T cells from whole blood

**NOTE:** The starting material for this protocol is 15 mL of fresh whole blood collected using standard procedures, allowing the source of the starting material to be selected based on research requirements. Scale the protocol as needed. Pre-warm phosphate buffered saline (PBS) + 2% fetal calf serum (FCS) to room temperature (RT) and adjust the centrifuge to RT before starting the CD4<sup>+</sup> T cell enrichment procedure.

1. Add 750  $\mu$ L of human CD4<sup>+</sup> T cell enrichment cocktail to 15 mL of whole blood in a 50 mL conical tube and mix gently by inversion. Incubate at RT for 20 min. When incubation is complete, add 15 mL of PBS + 2% FCS to the tube and mix gently by inversion.
2. Prepare a fresh 50 mL conical tube with 15 mL of density medium. Carefully layer the diluted blood sample onto the top of the density medium, being sure not to disrupt the density medium/blood interface that forms. Centrifuge for 20 min at 1200 x g and RT with acceleration set to 1 and the descending brake OFF.  
**NOTE:** It is critical to set the acceleration to 1 and descending brake OFF on the centrifuge to avoid disruption of the cell layer in the density medium.
3. Collect the enriched CD4<sup>+</sup> T cells from the density medium/plasma interface using a narrow stem transfer pipette. Transfer the collected cells to a fresh 50 mL conical tube. Centrifuge the collected CD4<sup>+</sup> T cells for 8 min at 423 x g and RT.  
**NOTE:** Make sure to return the centrifuge acceleration to 9 and descending brake to ON for step 1.3 and all of the following steps.
4. Discard the supernatant and wash the cell pellet twice with 50 mL of PBS + 2% FCS, centrifuging for 8 min at 423 x g and RT. Discard the final supernatant wash and suspend the washed cell pellet in 2 mL of PBS + 2% FCS.
5. Count cells using a hemocytometer. To continue with ATAC-seq, proceed to step 2.3. To freeze cells for later processing, proceed to step 1.6.
6. Add 1 mL of fresh freezing medium (90% FCS + 10% dimethyl sulfoxide) per 1 million cells. Aliquot 1 mL of cells in freezing medium in cryogenic safe tubes. Place tubes at -80 °C overnight in a slow-freezing container. The next day, transfer tubes to liquid nitrogen for long-term storage.  
**NOTE:** 1 million CD4<sup>+</sup> T cells will be isolated per 2 mL of original volume of whole blood. Freeze 500,000 cells in 1 mL aliquots of freezing medium. Make fresh freezing medium at every use.

### 2. Activate and purify CD4<sup>+</sup> T cells

**NOTE:** This rapid protocol for activating and purifying CD4<sup>+</sup> T cells only requires 48 h and results in 95% viable, activated CD4<sup>+</sup> T cells. Cool the centrifuge to 4 °C before beginning the protocol.

1. Thaw 1 vial (500,000) of CD4<sup>+</sup> T cells in a 37 °C water bath until the ice has just melted. Gently transfer cells to a 15 mL conical tube containing 9 mL of pre-warmed Roswell Park Memorial Institute-1640 (RPMI-1640) media supplemented with 10% FCS, hereafter referred to as complete RPMI.  
**NOTE:** Do not let cells thaw completely to avoid exposure to DMSO.
2. Centrifuge for 6 min at 1500 x g and 4 °C. Discard the supernatant and gently suspend the pellet in 2 mL of complete RPMI. Repeat the spin down, discard the supernatant, and gently suspend cells in 0.5 mL of complete RPMI.
3. Count the cells using a hemocytometer. Adjust the density of the cell suspension with complete RPMI to plate 50,000 cells in 200  $\mu$ L per well of a 96-well round bottom plate.  
**NOTE:** From one frozen tube of 500K CD4<sup>+</sup> T cells, plate 10 wells of 50,000 CD4<sup>+</sup> T cells in 200  $\mu$ L per well.
4. Prepare magnetic beads conjugated with human T-activator anti-CD3 and anti-CD28 antibodies.
  1. Aliquot 12.5  $\mu$ L of human T-activator CD3/CD28 beads per ATAC-seq sample to a 1.5 mL tube. Wash the beads with 1 mL of 1x PBS and place the tube on a magnet for 1 min.
  2. Carefully remove and discard the clear supernatant, remove the tube from magnet, and suspend the beads in 13  $\mu$ L complete RPMI per ATAC-seq sample.  
**NOTE:** Calculate the amount of beads necessary based on the number of ATAC-seq samples being processed. This protocol requires 12.5  $\mu$ L of CD3/CD28 beads per 500,000 cells, which constitutes one ATAC-seq sample.
5. Activate the CD4<sup>+</sup> T cells. Collect 500,000 cells in 2.1 mL of complete RPMI medium in a 15 mL conical tube and add 12.5  $\mu$ L of pre-washed human T-activator beads. Mix gently by inverting the tube.
6. Gently transfer the cells with beads to a sterile reservoir and plate 200  $\mu$ L of cells with beads per well of round bottom 96-well plate using a multi-channel pipette. Incubate for 48 h in a 37 °C, 5% CO<sub>2</sub> incubator.
7. Post incubation, centrifuge the 96-well plate for 8 min at 423 x g and RT. Remove 100  $\mu$ L of medium per well from the 96-well plate, collecting it to a conical tube before discarding to ensure no bead loss. Resuspend the cell pellet in the remaining 100  $\mu$ L and collect all in a 1.5 mL tube.  
**NOTE:** 10 wells of activated T cells will be collected in a 1 mL volume.
8. Perform CD4<sup>+</sup> isolation. Add 50  $\mu$ L of pre-washed CD4 conjugated beads to 500,000 cells in 1 mL of complete RPMI. Combine beads and cells by pipette. Incubate for 20 min at 4 °C in the fridge. Agitate beads and cell mixture every 6 min.  
**NOTE:** This protocol is to be used for one sample of 500,000 cells. Adjust as necessary for two or more samples of 500,000.
9. When the incubation is complete, place the tube on the magnet for 2 min. Remove and discard the supernatant when clear. Wash the bead-bound cells by removing the tube from the magnet, suspending the bead-bound cells in 1 mL of PBS + 2% FCS, replacing the tube on the magnet for 1 min, and discarding the clear supernatant. Repeat this process for a total of 3 washes.  
**NOTE:** Be careful not to discard any beads during washes.

10. After the final wash, place the tube on the magnet for 1 min. Remove and discard the supernatant and place the pellet on ice. Proceed to section 3 (ATAC-seq).

### 3. ATAC-seq

NOTE: In this step, nuclei are isolated from activated CD4<sup>+</sup> T cells for ATAC-seq. The lysis buffer used in this protocol has been improved to be gentler on the nuclei, resulting in more efficient digestion and higher quality results. All centrifugation steps in section 3 are performed with a fixed-angle centrifuge maintained at 4 °C. Cool the centrifuge before beginning protocol.

1. Perform nuclei isolation. Resuspend beads and cells with cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.03% polysorbate 20). Centrifuge immediately at 500 x g for 10 min at 4 °C. Remove and discard the supernatant.
2. Perform the transposase reaction. Suspend the isolated nuclei pellet with 50 µL of Tn5 transposase mix (**Table 1**). Incubate in a thermocycler for 30 min at 37 °C with a 40 °C lid, holding at 4 °C when complete.
3. After thermocycling, immediately perform a quick benchtop centrifuge spin down and place the tube on the magnet for 1 min to remove the beads from the product.
4. Transfer the clear supernatant from the tube on the magnet to a DNA purification column. Wash the column once with 250 µL of Buffer PB and twice with 750 µL of Buffer PE. Elute the sample in 10 µL of elution buffer. Proceed directly to step 3.5.  
NOTE: This step amplifies the DNA fragments with the ligated adaptors, here referred to as the ATAC-seq library. In order to multiplex several ATAC-seq libraries to be run on one NGS lane, use unindexed Primer 1 for all samples and a different indexed Primer 2 for individual samples (**Table 2**). Primers are used at working concentrations of 1.25 µM.
5. Set up the initial PCR amplification reaction in a nuclease-free PCR tube, combining the components in the order and volume specified in **Table 3**. Place the PCR tube into the thermocycler and run the PCR amplification program with cycling conditions as specified in **Table 4**.
6. Monitor the reaction by qPCR. Set up the qPCR reaction in a nuclease-free PCR tube, combining the components in the order and amount specified in **Table 5**. Place in qPCR machine and cycle as specified in **Table 6**.
  1. To determine the optimal number of additional amplification cycles for the remaining 45 µL reaction, create a plot with cycle number on the x-axis and relative fluorescence (RFU) on the y-axis. The optimal number of additional amplification cycles is one-third the number of cycles it takes for the qPCR reaction to reach plateau.  
NOTE: Monitoring the reaction by qPCR allows for determination of the optimal number of PCR cycles desired to obtain optimal library fragment amplification while minimizing GC content and size bias.
7. Complete the final PCR amplification of the remaining 45 µL of PCR reaction. Place the PCR tube with the amplification reaction from step 3.5 back in the thermocycler and run the program as described in **Table 7**.  
NOTE: For samples processed as described in this protocol, the optimal total number of PCR amplification cycles was determined to be 12.
8. After the amplification is complete, purify the libraries using a PCR cleanup kit following the manufacturer's protocol, eluting in 25 µL of the elution buffer.  
NOTE: Amplified libraries can be stored at 4 °C for up to 48 h or frozen at -20 °C for long-term storage.

### 4. ATAC-seq library quality analysis

NOTE: It is important to validate the quality and quantity of ATAC-seq libraries before next-generation sequencing. The quality and quantity of the libraries should be assessed using commercially available kits (see **Table of Materials**).

1. Assess the quality and quantity of the ATAC-seq libraries using a microfluidics-based platform for sizing, quantification and quality control of DNA. See **Figure 2** for representative quality assessment results.  
NOTE: The concentration of the ATAC-seq libraries must be >1 ng/µL to obtain quality sequencing results. On average, 30 nM in 25 µL was obtained.

### 5. Sequencing and data analysis

NOTE: This analysis pipeline allows users to control the quality of reads mapping procedure, adjust coordinates for experimental design, and call peaks for downstream analysis. The following are the commands lines and explanation of execution. The data analysis package is available at ([https://github.com/s18692001/bulk\\_ATAC\\_seq](https://github.com/s18692001/bulk_ATAC_seq)).

1. Sequence the prepared libraries on a next generation sequencer to an average read depth of 42 million reads per sample.
2. Estimate the quality of sequencing reads by checking files generated by FastQC<sup>12</sup> software package using the command:  
fastqc -o <output\_directory> <fastq\_file>
3. Trim the bases of reads by Trimmomatic<sup>13</sup> software, if needed, as determined by the FastQC quality check, using the command (for pair-ended):  
java -jar <path to trimmomatic jar file> PE <input1> <input2> <paired output1> <unpaired output1> <paired output2> <unpaired output2>  
HEADCROP:<crop bases>
4. Align reads to human reference genome (hg38) by Bowtie2<sup>14</sup> software:  
bowtie2 -x <reference genome> -1 <input pair 1> <input pair 2> -S <output SAM file>  
NOTE: The argument -x is for the basename of the index for the reference genome, and -S is for the SAM format output.
5. Check the quality of mapped reads using Samtools<sup>15</sup> flagstat package:  
samtools flagstat <BAM file>
6. Sort the mapped reads file and remove duplicates using Picard<sup>16</sup> tool:  
java -jar picard.jar SortSam INPUT=<input SAM file name> OUTPUT=<output sorted BAM file name> SORT\_ORDER=coordinate and "java -jar picard.jar MarkDuplicates INPUT=<sorted BAM file> OUTPUT=<output BAM file without duplicates> REMOVE\_DUPLICATES=TRUE

7. Index BAM file, trim the unused chromosome reads and shift the coordinates for the experimental reason of ATAC-seq<sup>17</sup>:  

```
java -jar picard.jar BuildBamIndex INPUT=<BAM file>
samtools idxstats <Input BAM file> | cut -f 1 | grep -v chrY | grep -v chrM | grep -v chrUn | xargs samtools view -b <Input BAM file> > <Output trimmed BAM file>
bedtools bamtobed -i <Input trimmed BAM file> > <Output trimmed BED file>
awk 'BEGIN {OFS = "\t"} ; {if ($6 == "+") print $1, $2 + 4, $3 + 4, $4, $5, $6; else print $1, $2 - 5, $3 - 5, $4, $5, $6}' <Input trimmed BED file> > <Trimmed shifted BED file>
```
8. Delete the reads that are shifted to negative coordination:  

```
awk '{if ($2 > 0) print $1 "\t" $2 "\t" $3 "\t" $4 "\t" $5 "\t" $6}' <Input BED file> > <Output non-negative coordination BED file>.
```
9. Convert the BED file to BAM file for following DiffBind<sup>18</sup> analysis:  

```
bedtools bedtobam -i <Input BED file> -g <reference genome> > <Output BAM file>
```
10. Perform peak calling by DiffBind<sup>18</sup> software package:  

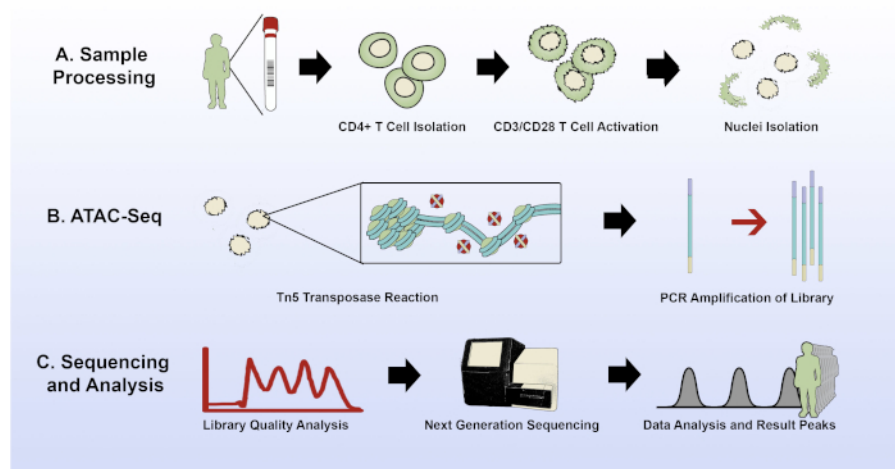
```
macs2 callpeak -t <Input BAM file> -f BAM -g hs -nomodel --nolambda --keep-dup all --call-summits --outdir <Output directory path> -n <Output name> -B -q 0.01 --bdg --shift -100 --extsize 200
```

NOTE: After filtering, expect a median of 37 million reads per sample. Mitochondrial DNA contamination will range from 0.30%–5.39% (1.96% on average). There will be a low rate of multiply mapped reads (6.7%–56%, 19% on average) and a relatively high percentage of usable nuclear reads (60%–92%, 79% on average).

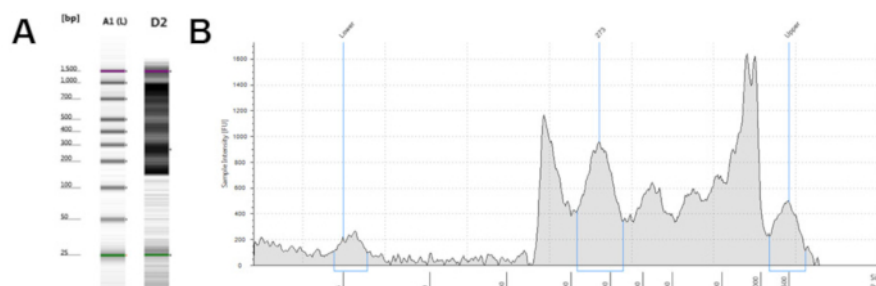
## Representative Results

From 15 mL of fresh whole blood, this protocol generates an average of 1 million CD4+ T cells. These can be frozen for later processing or used immediately. Viability of the CD4+ T cells, fresh or thawed, was consistently >95%. This method of CD4+ T cell isolation allows for flexibility in source material and collection time. This improved ATAC-seq protocol produces a final library of greater than 1 ng/μL for sequencing. Quality control performed using commercially available systems should demonstrate DNA fragments between 200 and 1000 bp (**Figure 2**). Sequencing should only be performed with high quality libraries.

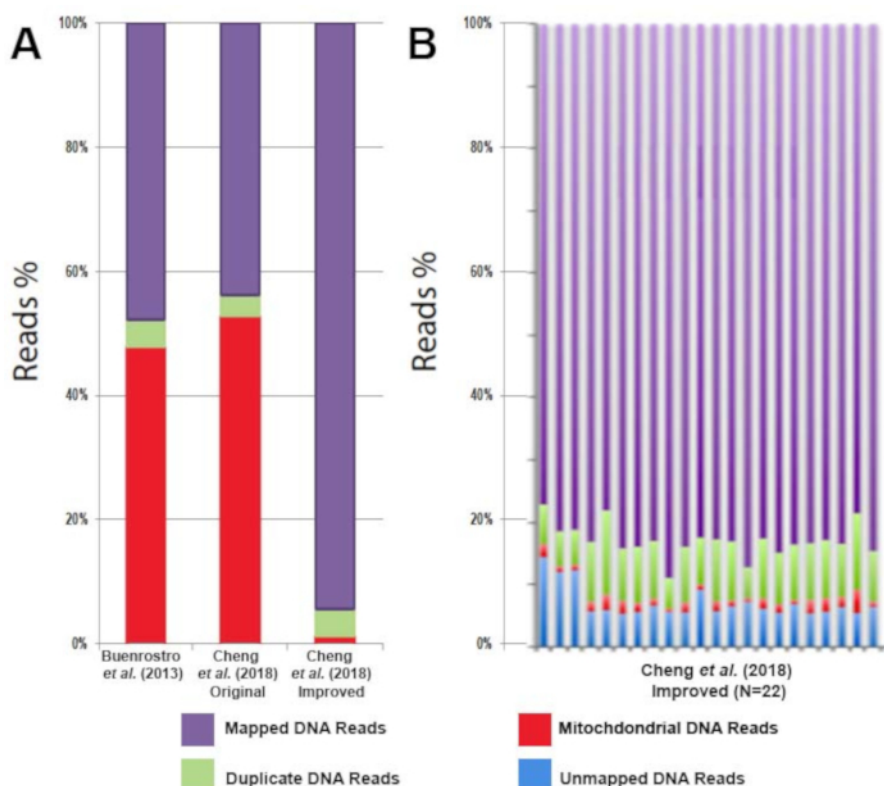
All libraries were sequenced to an average depth of greater than 40 million read per sample. While previous ATAC-seq protocols have been challenged by contaminating sequencing reads from mitochondrial DNA that can range from 50%–60% of the total sequencing reads<sup>1</sup> this improved protocol eliminates the issue. Libraries prepared following this protocol contain on average only 3% mitochondrial reads (**Figure 3A**). The high percentage of usable reads is sufficiently constant across biological replicates (**Figure 3B**). The protocol was able to provide highly reproducible results across technical replicates (**Figure 4A, B**) as well as biological replicates (**Figure 4C, D**). Additionally, the protocol for CD4+ T cell activation presented takes 48 h rather than one week or more and results in consistent and efficient activation, as demonstrated by reproducible sequencing results (**Figure 4A, B**). Predicted ATAC-seq peaks are accurately called by the analysis pipeline (**Figure 4E**). Analysis of sequencing results identified clear changes in chromatin state during human T cell activation. Differentially accessible regions of open chromatin were identified between six samples before and after 48 h of activation (**Figure 5**).



**Figure 1: Experimental overview of the modified ATAC-seq protocol.** (A) Sample acquisition and processing, from 15 mL of patient whole blood through CD4+ T cell isolation, plating and activation of the T cells, and nuclei isolation with the improved lysis buffer. (B) The transposase reaction and PCR amplification of the sequencing library. (C) Quality analysis, sequencing, and data analysis. [Please click here to view a larger version of this figure.](#)

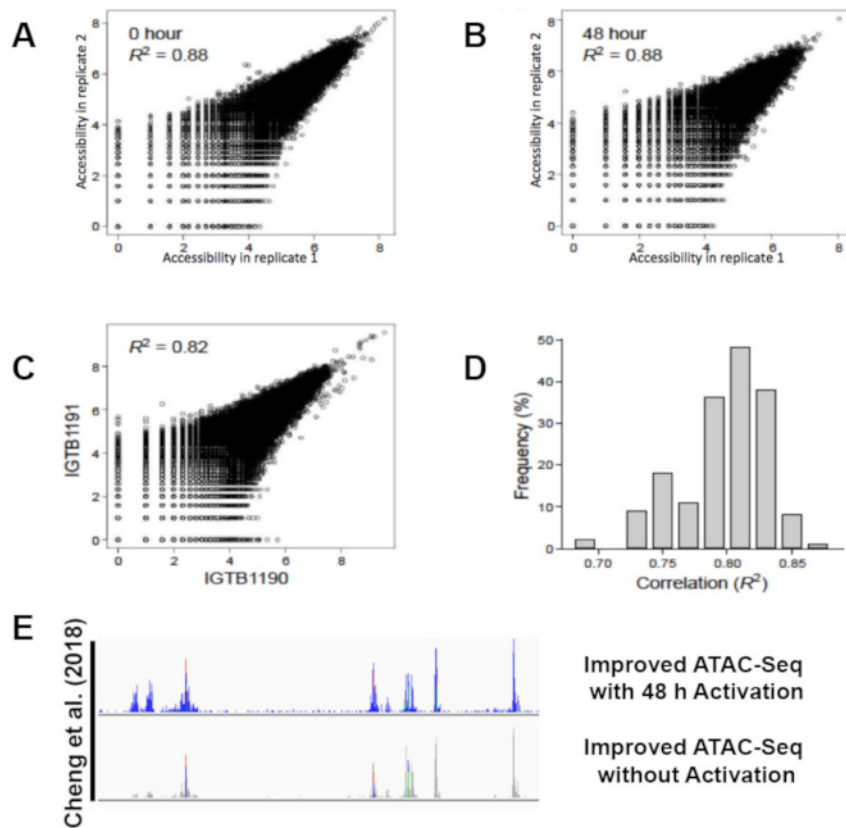


**Figure 2: Representative high quality ATAC-seq libraries from a microfluidics-based platform for sizing, quantification and quality control of DNA.** (A) Electronic gel image of samples B1 and D1 with banding between 200 and 1,000 base pairs. (B and C) Electropherogram trace result of samples B1 (B) and D1 (C), with peaks between 200 and 1,000 base pairs. [Please click here to view a larger version of this figure.](#)

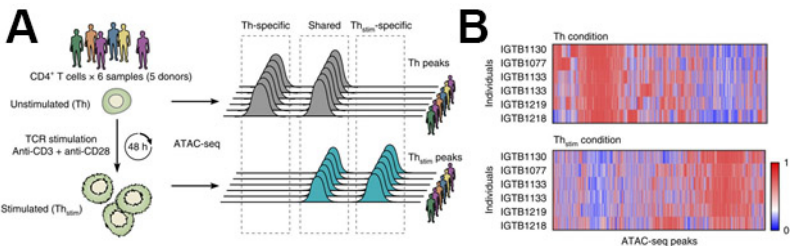


**Figure 3: Decreased mitochondrial DNA contamination with the improved ATAC-seq protocol results in an increase in usable DNA sequencing reads.** (A) Comparison of usable reads (purple), duplicate reads (green), and mitochondrial reads (red) from CD4+ T cell ATAC-seq profiling in the literature. (B) Comparison of usable reads (purple), duplicate reads (green), mitochondrial reads (red), and unmapped reads (blue) of CD4+ T cell ATAC-seq profiling from multiple healthy individuals (n = 22). This figure has been modified from Cheng et al.<sup>10</sup>. [Please click here to view a larger version of this figure.](#)





**Figure 4: Improved ATAC-seq protocol reproducibility and accuracy.** Scatter plots of chromatin accessibility (ATAC-seq signal, x and y-axes) for two replicate experiments of unstimulated (**A**; 36,486 Th peaks) or activated (**B**; 52,154 Thstim peaks) T cells demonstrates technical reproducibility. Chromatin accessibility for activated T cells from individuals IGTB1191 (y-axis) and IGTB1190 (x-axis) (**C**) and histogram of correlations between every pairs of individuals for the 52,154 Thstim peaks (**D**) demonstrates reproducibility between individuals. (**E**) ATAC-seq peaks called with our improved ATAC-seq protocol at chromosome 19 Q13.12. This figure has been modified from Cheng et al.<sup>10</sup>. [Please click here to view a larger version of this figure.](#)



**Figure 5: Representative ATAC-seq results of changes in chromatin state after CD4+ T-cell activation.** (**A**) Experimental overview (left) and nomenclature (right). (**B**) Differentially accessible regions of open chromatin (columns) in six samples (rows) before (top, Th) and after (bottom, Thstim) 48 h activation of primary CD4+ T cells. This figure has been modified from Cheng et al.<sup>10</sup>. [Please click here to view a larger version of this figure.](#)

2x TD Buffer	25 µL
TN5 Enzyme	5 µL
Nuclease Free Water	20 µL
Total Volume	50 µL

**Table 1: Step 3.2 transposase reaction components.**

Ad1_noMX: AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG
Ad2.1_TAAGGCGA CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT
Ad2.2_CGTACTAG CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT
Ad2.3_AGGCAGAA CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT
Ad2.4_TCCTGAGC CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT
Ad2.5_GGACTCCT CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGT
Ad2.6_TAGGCATG CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT
Ad2.7_CTCTCTAC CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT
Ad2.8_CAGAGAGG CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGT
Ad2.9_GCTACGCT CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT
Ad2.10_CGAGGCTG CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT
Ad2.11_AAGAGGCA CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT
Ad2.12_GTAGAGGA CAAGCAGAAGACGGCATACGAGATTCTCTACGTCTCGTGGGCTCGGAGATGT
Ad2.13_GTCGTGAT CAAGCAGAAGACGGCATACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT
Ad2.14_ACCACTGT CAAGCAGAAGACGGCATACGAGATACAGTGGTGTCTCGTGGGCTCGGAGATGT
Ad2.15_TGGATCTG CAAGCAGAAGACGGCATACGAGATCAGATCCAGTCTCGTGGGCTCGGAGATGT
Ad2.16_CCGTTTGT CAAGCAGAAGACGGCATACGAGATACAAACGGGTCTCGTGGGCTCGGAGATGT
Ad2.17_TGCTGGGT CAAGCAGAAGACGGCATACGAGATACCCAGCAGTCTCGTGGGCTCGGAGATGT
Ad2.18_GAGGGGTT CAAGCAGAAGACGGCATACGAGATAACCCCTCGTCTCGTGGGCTCGGAGATGT
Ad2.19_AGGTTGGG CAAGCAGAAGACGGCATACGAGATCCCAACCTGTCTCGTGGGCTCGGAGATGT
Ad2.20_GTGTGGTG CAAGCAGAAGACGGCATACGAGATCACCACACGTCTCGTGGGCTCGGAGATGT
Ad2.21_TGGGTTTC CAAGCAGAAGACGGCATACGAGATGAAACCCAGTCTCGTGGGCTCGGAGATGT
Ad2.22_TGGTCACA CAAGCAGAAGACGGCATACGAGATTGTGACCAGTCTCGTGGGCTCGGAGATGT
Ad2.23_TTGACCCT CAAGCAGAAGACGGCATACGAGATAGGGTCAAGTCTCGTGGGCTCGGAGATGT
Ad2.24_CCACTCCT CAAGCAGAAGACGGCATACGAGATAGGAGTGGGTCTCGTGGGCTCGGAGATGT

**Table 2: ATAC-seq oligos designs used for PCR.**

Nuclease Free Water	11.9 $\mu$ L
100 $\mu$ M Custom Nextera Primer 1 (Table 2)	0.6 $\mu$ L
NEBNext High-Fidelity 2x PCR Master Mix	25 $\mu$ L
ATAC-Seq Library	10 $\mu$ L
25 $\mu$ M Custom Nextera Primer 2 (Table 2)	2.5 $\mu$ L
Total Volume	50 $\mu$ L

**Table 3: Step 3.5 initial PCR reaction mix.**

CYCLE STEP	TEMPERATURE	TIME	CYCLES;
Extension	72 °C	5 min	1
Initial Denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	10
Annealing	63 °C	30 s	
Extension	72 °C	1 min	
Hold	4 °C	Infinity	1

**Table 4: Step 3.5 initial PCR amplification cycling program.**

PCR Reaction Aliquot	5 $\mu$ L
PCR Cocktail from Table 3 with 0.6x Syber Green	10 $\mu$ L

**Table 5: Step 3.6 qPCR reaction mix.**

CYCLE STEP	TEMPERATURE	TIME	CYCLES
Initial Denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	20
Annealing	63 °C	30 s	
Extension	72 °C	1 min	
Hold	4 °C	Infinity	1

**Table 6: Step 3.6 qPCR cycling program.**

CYCLE STEP	TEMPERATURE	TIME	CYCLES
Initial Denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	As Determined
Annealing	63 °C	30 s	
Extension	72 °C	1 min	
Hold	4 °C	Infinity	1

**Table 7: Step 3.7 PCR cycling program for final PCR amplification.**

## Discussion

The modified ATAC-seq protocol presented in this article provides reproducible results with minimal mitochondrial DNA contamination. The protocol has been used to successfully characterize chromatin architecture of human primary PBMCs<sup>10</sup>, human monocytes, mouse dendritic cells (unpublished), and cultured melanoma cell lines<sup>11</sup>. We anticipate this improved lysis condition has the potential to work for other cell types as well. It is also anticipated that this nuclei isolation protocol will be compatible with single nuclei ATAC-seq protocols, minimizing mitochondrial DNA contamination to improve sequencing results.

An additional benefit of this modified protocol is the ability to freeze batches of isolated CD4+ T cells from PBMCs at different times depending on the availability of patient samples. As ATAC-seq can then be performed concurrently on all samples, potential batch effect bias in the transposase reaction and sequencing is minimized<sup>10</sup>. It is critical that freezing medium be made fresh with each use, in order to maintain the high viability that is achieved with this freeze-thaw protocol. Viability of isolated CD4+ T cells should remain above 90% in order to avoid non-specific digestion in the transposase reaction<sup>1</sup>.

Please note that further optimization may be required in order to use this protocol with variable cell types and quantities. Since the Tn5 transposase-to-nuclei ratio has been optimized for 500,000 nuclei in this protocol, if performing ATAC-seq with a different number of nuclei, the amount of Tn5 transposase should be adjusted accordingly. Over-lysis of nuclei due to an excess of Tn5 may lead to high background from closed chromatin and low complexity of sequencing libraries, while under-lysis may not provide a complete PCR amplified library<sup>1</sup>. In order to avoid these complications, it is advised to perform careful nuclei counting and optimize the Tn5 ratio as necessary. To further improve data quality, it is advised to optimize PCR amplification cycles by qPCR monitoring. If the final library undergoes too many amplification cycles, there may be bias introduced in the sequencing data<sup>2</sup>. It is recommended to perform proper quality control of the ATAC-seq libraries prior to next-generation sequencing in order to save time and money.

As we have demonstrated, a new lysis buffer is key to the reduction of mitochondrial DNA contamination and is effective on a wide range of cell types. While other protocols have addressed the issue of mitochondrial contamination through the addition of extra washing steps<sup>19</sup> or the intensive process of a CRISPR-mediated mitochondrial DNA depletion<sup>20</sup>, this ATAC-seq protocol does not add reagents or experimental steps, which makes it a more accessible technique. Together with RNA-seq and single-cell sequencing, ATAC-seq is a powerful tool for exploring epigenetic regulation. This improved ATAC-seq protocol and data analysis package will help decrease sequencing costs and produce higher quality results.

## Disclosures

The authors have nothing to disclose.

## Acknowledgements

We thank Atsede Siba for technical support. C.S.C. is supported by NIH grant 1R61DA047032.



## References

1. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*. **10** (12), 1213-1218 (2013).
2. Buenrostro, J. D., Wu, B., Chang, H. Y., Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology*. **21**, 29.1-29.9 (2015).
3. Song, L., Crawford, G.E. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor Protocols*. (2) (2010).
4. Pajero, A., Muiño, J. M., Angenent, G. C., Kaufmann, K. Profiling Nucleosome Occupancy by MNase-seq: Experimental Protocol and Computational Analysis. *Methods in Molecular Biology*. **1675**, 167-181 (2018).
5. Giresi, P. G., Kim, J., McDaniell, R. M., Iyer, V. R., Lieb, J. D. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Research*. **17** (6), 877-885 (2007).
6. Rosenberg, A. B., et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*. **360** (6385), 176-182 (2018).
7. Corces, M. R. et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature Genetics*. **48** (10), 1193-1203 (2016).
8. Lu, Z., Hofmeister, B. T., Vollmers, C., DuBois, R. M., Schmitz, R. J. Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Research*. **45** (6), e41 (2017).
9. Lara-Astiaso, D. et al. Immunogenetics. Chromatin state dynamics during blood formation. *Science*. **345** (6199), 943-949 (2014).
10. Gate, R. E. et al. Genetic determinants of co-accessible chromatin regions in activated T cells across humans. *Nature Genetics*. (2018).
11. Sanjana, N. E. et al. High-resolution interrogation of functional elements in the noncoding genome. *Science*. **353** (6307), 1545-1549 (2016).
12. Andrews, S. *FastQC A Quality Control tool for High Throughput Sequence Data*. Babraham Bioinformatics. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed September 15, 2018 (2018).
13. Bolger, A. M., Lohse, M., Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. **30** (15), 2114-2120 (2014).
14. Langmead, B., Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. **9** (4), 357-359 (2012).
15. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25** (16), 2078-2079 (2009).
16. Broad Institute. *Picard Tools*. <http://broadinstitute.github.io/picard/>. Accessed September 15, 2018 (2018).
17. Quinlan, A. R., Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26** (6), 841-842 (2010).
18. Stark, R., Brown, G. *DiffBind: Differential Binding Analysis of ChIP-Seq Peak Data*. <http://bioconductor.org/packages/devel/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>. Accessed September 15, 2018 (2018).
19. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nature Methods*. **14** (10), 959-962. (2017).
20. Wu, J. et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature*. **534** (7609), 652-657 (2016).