bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

An optimized protocol for neuronal assay for transposase-accessible chromatin by sequencing (ATAC-seq) library preparation using *Drosophila melanogaster*

Collin B. Merrill[1], Miguel A. Pabon[2], Austin B. Montgomery[2], Aylin R. Rodan[2,3,4] and Adrian Rothenfluh[1,2,4,5,*]

[1]Department of Psychiatry University of Utah, Salt Lake City, UT 84108, USA.
[2]Molecular Medicine Program, University of Utah, Salt Lake City, UT 84112, USA.
[3]Division of Nephrology, Department of Internal Medicine, University of Utah, Salt Lake City, UT 84112, USA.
[4]Department of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA.
[5]Department of Neurobiology, University of Utah, Salt Lake City, UT 84112, USA.

*Corresponding author:
Adrian Rothenfluh, adrian.rothenfluh@hsc.utah.edu

Word count: 5081
Number of figures: 7

ABSTRACT (227 words)

Assay for transposase-accessible chromatin by sequencing (ATAC-seq) is rapidly becoming the assay of choice to investigate chromatin-mediated gene regulation, largely because of low input requirements, a fast workflow, and the ability to interrogate the entire genome in an unbiased manner. Many studies using ATAC-seq use mammalian or human-derived tissues and the established protocols work well in these systems. However, ATAC-seq is not yet widely used in *Drosophila* model systems. Vinegar flies present several key advantages over mammalian systems that make them an excellent model for ATAC-seq studies. For example, there are abundant genetic tools in flies that are not available in mammalian models that allow straightforward targeting, transgene expression, and genetic manipulation. Because current ATAC-seq protocols are not optimized to use fly tissue as input, we developed an optimized workflow that accounts for several complicating factors present in *Drosophila*. We examined several parameters of library construction, including input size, freezing time, and washing, which specifically address the fly cuticle and possible confounds from retinal pigments. Then, we further optimized the enzymatic steps of library construction to account for the smaller genome size in flies. Finally, we used our optimized protocol to generate ATAC-seq libraries that displayed metrics indicating excellent library quality. Our optimized protocol will enable extensive ATAC-seq experiments in *Drosophila*, thereby leveraging the advantages of this powerful model system to understand chromatin-mediated gene regulation.

Keywords: Drosophila, neurons, Tn5, tagmentation, ATAC-seq

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

Assay for transposase-accessible chromatin by sequencing (ATAC-seq) is a relatively new technique that has advanced our understanding of chromatin-mediated gene regulation. The technique offers several advantages over other sequencing techniques, such as using micrococcal nuclease (MNase-seq) or formaldehyde-assisted isolation of regulatory elements (FAIRE-seq), that examine chromatin. These advantages include low input requirements, a fast workflow that does not require harsh chemicals, and genome-wide interrogation in an untargeted manner (Buenrostro et al., 2013). Additionally, ATAC-seq is robust to small sample sizes and has been adapted for single-cell ATAC-seq (Buenrostro et al., 2018; Cusanovich et al., 2015).

The established ATAC-seq protocol was developed using peripheral blood samples (Buenrostro et al., 2013), and many studies using ATAC-seq focus on mammalian or human-derived tissues (Bysani et al., 2019; Davie et al., 2015; Liu et al., 2019). Recent studies have used ATAC-seq to investigate other tissue types, including tissues from plants (Lu et al., 2017), zebrafish (Quillien et al., 2017), and livestock (Halstead et al., 2020). While some ATAC-seq studies using *Drosophila melanogaster* have been published, many of these studies focus on embryogenesis and early developmental stages in larvae (Bozek et al., 2019; Cusanovich et al., 2018; Haines and Eisen, 2018). ATAC-seq using *Drosophila* is a powerful technique to investigate chromatin-mediated gene regulation. Indeed, *Drosophila* model systems provide many advantages over mammalian systems, such as abundant genetic tools, high fecundity, quick life span, and easy husbandry. Additionally, *Drosophila* have approximately 200,000 neurons in the brain, which provides a less-complex system that maintains key functional,

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

physiological, and morphological differences between neuron types. Thus, ATAC-seq is one way to investigate neuronal identity and the genetic differences between neuron populations.

ATAC-seq library preparation, while a relatively quick and straightforward process, has several parameters that can be optimized based on the input tissue. Many of these parameters are addressed in the original ATAC-seq protocol (Buenrostro et al., 2013), and additional protocols have been published further optimizing ATAC-seq library preparation from various tissue types in several storage conditions (Corces et al., 2017; Fujiwara et al., 2019; Ho et al., 2018; Wang et al., 2018). Most studies in *Drosophila* larvae treat the tissue similarly to mammalian tissues, only adding an additional washing step to clean the embryos or larvae before homogenization (Haines and Eisen, 2018; Koenecke et al., 2016). However, ATAC-seq library preparation using adult vinegar flies poses challenges that are not present in other tissues. For example, the insect cuticle is a chitin structure that protects against dehydration and predation. This tough structure is difficult to process by simple homogenization. Another challenge is fly eye color, which ranges from white to red, depending on the fly genotype. Indeed, eye color is commonly used to identify flies with the correct genotype, and many transgenic fly lines contain deliberate insertions a the *mini-white* (w^{+mC}) gene to confer yellow to red eye color to w^{-} flies. Some retinal pigments in vinegar flies emit fluorescence (Franceschini et al., 1981; Miller et al., 1984), which can complicate fluorescently-activated nuclei sorting (FANS). Finally, the *Drosophila* genome is considerably smaller than mammalian genomes (Canapa et al., 2016; Vinogradov, 2004), so established ATAC-seq protocols may not be appropriate for fly tissues. Thus, optimizing each variable parameter for ATAC-seq

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

library preparation for adult flies should be performed prior to sequencing. However, these experiments add considerable sequencing costs, putting the approach out-of-reach for many labs.

To address these challenges, we developed a protocol to generate ATAC-seq libraries from adult *Drosophila* neurons. Our goal was to optimize each variable in the library preparation process, with particular attention to challenges specific to *Drosophila* model systems, such as the insect cuticle and eye color. Our protocol provides guidelines for ATAC-seq library preparation using vinegar flies and outlines the optimal parameters for each library preparation step. This protocol will accelerate studies of chromatin-mediated gene regulation in flies, thereby coupling the powerful genetic tools in *Drosophila* with high-throughput, in-depth examination of chromatin structure.

MATERIALS AND METHODS

Fly strains

The following fly lines were obtained from the Bloomington Drosophila Stock Center (Indiana University, Bloomington, IN, USA) and used in this study: Canton S (BL64349), *w** Berlin (BL2057), nsyb-Gal4 (BL51635), UAS-Stinger (BL84277), cha-Gal4 (BL39691), UAS-GFP-nls (BL4775), and vGAT-Gal4 (BL58980). Flies were reared in bottles containing standard cornmeal agar and grown at 25 °C with 70% relative humidity and a 12-h light/dark cycle. cha-Gal4 and UAS-GFP-nls flies were crossed to produce a stable cha-Gal4>UAS-GFP-nls stock. Male flies from this stock were crossed to w* Berlin and Canton S to generate orange- and red-eyed flies, respectively, for Figs. 4-5. F1 male and female progeny from nsyb-Gal4 x UAS-Stinger and vGAT-Gal4 x UAS-GFP-nls were used for ATAC-seq library preparation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Nuclei isolation

All equipment and buffers were pre-chilled to be ice-cold, and all plastic and glassware used for nuclei isolation were pre-treated with 1% bovine serum albumin (BSA) in PBS. Flies were collected into empty bottles and frozen at -80 °C. Then, the flies were placed in pre-chilled shakers with sieves and agitated for 1 min to separate the heads. The fly heads were added to lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween-20, 0.1% Nonidet P40 substitute, 0.01% digitonin, and 1% bovine serum albumin) in a Dounce homogenizer. The tissue was homogenized with the A pestle until the resistance disappeared. The crude homogenate was passed through a 0.40 μM filter and homogenized with 15 slow strokes with the B pestle. The crude nuclei were washed with wash buffer (10 mM Tris-HCl, 10 mM NaCl, 3 mM $MgCl_2$, 1% BSA, and 0.1% Tween-20) and centrifuged at 500 *g* for 10 minutes per wash. The washed nuclei were resuspended in 1 mL wash buffer with 3 mM DAPI for evaluation and sorting.

Fluorescence-activated nuclei sorting (FANS)

Nuclei extracts were evaluated and sorted with a BD FACS Aria flow cytometer (BD Biosciences, San Jose, CA, USA) operated by the University of Utah Flow Cytometry Core facility. Nuclei collected from *w** Berlin flies (no GFP) were stained with 3 μM DAPI and used as the GFP-negative control to set the gating scheme. GFP-positive nuclei were collected into 500 μL wash buffer and stored on ice until use for ATAC-seq library prep. Nuclei counts were determined using an internal control bead population (Spherotech, ACBP-100-10) at 10^6 beads/mL. Briefly, 20 uL of well-mixed beads were added to 200 uL of nuclei suspension and genetly mixed. Data was recorded for the

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bead and cell mixture until 10,000 singlet beads were collected. Nuclei were defined based on forward scatter signal and DAPI intensity. The number of nuclei was calculated as follows: number of nuclei recorded × 10^5/number of beads recorded = number of nuclei/mL.

DNA tagmentation, amplification, and purification

Purified nuclei were centrifuged for 10 min at 500 *g* at 4 °C. The pellet was mixed with 22.5 µL 2X Tn5 tagmentation buffer (20 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, and 20% dimethyl formamide in sterile water), 16.5 µL 1X PBS, 0.5 µL Tween-20, 5.5 µL water, and Tn5 enzyme (Illumina, Inc., San Diego, CA, USA) and incubated at 37 °C. Tagmented (tagged/fragmented) DNA was purified using a MinElute PCR purification kit (Qiagen, Germantown, MD, USA). The purified DNA was mixed with CD index primers (Illumina) and Phusion High Fidelity PCR MasterMix (New England Biolabs, Ipswich, MA, USA) and amplified for 1 cycle of 72 °C for 5 min and 98 °C for 30 sec, followed by 5 cycles of 98 °C for 10 sec, 63 °C for 30 sec, and 72°C for 1 min in a C100 thermocycler (BioRad, Hercules, CA, USA). 5 µL of this reaction was used for a qPCR side reaction. The aliquot was mixed with the same CD index primers and SsoFast EvaGreen Supermix (BioRad) and amplified with 1 cycle of 98 °C for 30 sec and 40 cycles of 98 °C for 10 sec, 63 °C for 30 sec, and 72 °C for 1 min in an Applied Biosystems 7900HT qPCR instrument (Thermo Fisher Scientific, Waltham, MA, USA). Total fluorescence was calculated from the minimum and maximum fluorescence of each sample, and the cycle number corresponding to a portion of the total fluorescence was determined (~8-10 cycles). Then, the tagmented DNA was amplified for this

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

number of additional cycles. After amplification, the tagmented DNA was purified using AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA) using double-sided size selection (0.5X and 1.1X). The purified DNA was stored at -20 °C until further analysis.

Effect of homogenizer and fly volumes

The effect of homogenizer size and fly volume was assessed using 2 and 7 mL Dounce homogenizers and 0.5 and 2 mL flies, with each fly volume tested in each homogenizer size. Nuclei isolation and sorting were performed as described above.

Effect of freezing time on GFP fluorescence during nuclei isolation

Flies were frozen at -80 °C for 5, 10, 20, and 40 min before isolating heads and nuclei. Nuclei isolation and sorting were performed as described above.

Effect of eye color and washing

To assess the role of eye color on nuclei sorting, we used w*; cha-Gal4 UAS-GFP-nls/+ for orange-eyed flies and w+; cha-Gal4 UAS-GFP-nls/+ males and w+/w*; females for red-eyed flies. Flies were homogenized as described above and washed with wash buffer 1, 3, or 5 times before flow cytometry.

Effect of ATAC-seq library preparation parameters

To assess how changing the tagmentation and amplification parameters affects downstream ATAC-seq library quality, we performed experiments while changing the reaction time, Tn5 concentration, and the number of PCR cycles. For reaction time

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

experiments, purified nuclei were tagmented at 37 °C for 5, 23, or 60 minutes. The rest of the ATAC-seq library was prepared as described above. To assess Tn5 concentration, ATAC-seq libraries were tagmented for 23 minutes at 37 °C using 2.5 µL 10-fold diluted (0.1X), 0.83 µL (0.3X), 2.5 µL (1X), or 7.5 µL (3X) Tn5 enzyme. To identify the optimal number of PCR cycles for ATAC-seq libraries, we amplified the tagmented DNA to 25% (8 cycles), 33% (9 cycles), and 50% (10 cycles) total fluorescence of a qPCR side reaction. Libraries were then purified as described above.

ATAC-seq library quality assessment

ATAC-seq library quality was assessed using an Agilent 2200 TapeStation and High Sensitivity D5000 ScreenTape assays (Agilent Technologies, Inc., Santa Clara, CA, USA) at the Huntsman Cancer Institute High Throughput Genomics core facility or the Genome Technology Access Center of the McDonnell Genome Institute at the Washington University School of Medicine (St. Louis, MO, USA).

Next-generation sequencing and analysis

ATAC-seq libraries constructed using nuclei isolated from GABAergic neurons were sequenced on a Novaseq 6000 instrument (Illumina) using 50-bp paired-end reads. Fastq files were quality checked using FastQC (v 0.11.9). Adapter sequences were removed with CutAdapt 3.4. Then, the sequences were aligned to the dmel_r6.26 genome assembly using Bowtie2 (2.4.2). Aligned sequences were sorted with Samtools (1.12) and deduplicated with Picard (2.23) using the MarkDuplicates command. Insert size distributions were calculated using Picard.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Statistical analysis

The data are represented as means ± standard deviation. Comparisons between homogenizer size and fly volume were analyzed by two-way ANOVA with Bonferroni multiple comparison tests. The effect of eye color was analyzed by Student's t-test, and the effect of freezing on GFP fluorescence was analyzed by one-way ANOVA using Tukey's multiple comparison tests. The effect of multiple washes was analyzed by simple linear regression, and the regression lines were compared by analysis of covariance (ANCOVA). All statistical analyses were performed using GraphPad Prism 9 software (San Diego, CA, USA).

RESULTS

ATAC-seq offers several advantages over previous methods to assay chromatin. For example, the number of nuclei per sample is much lower than for ChIP-seq. Further, the reaction volume can be easily scaled to match the number of nuclei used for each sample. Additionally, ATAC-seq has a much faster workflow, with total hands-on time for library preparation taking 3-4 hours. These advantages are highlighted in studies using mammalian tissues (Buenrostro et al., 2013). Several factors must be considered when using ATAC-seq in *Drosophila* studies. For example, the insect cuticle presents a significant barrier to tissue homogenization and cell lysis. Further, eye color, which is commonly used to identify target genotypes, could influence the successful isolation of GFP-labeled nuclei. Finally, the parameters that are commonly used for ATAC-seq were determined using mammalian systems, which have much larger genomes than

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Drosophila. Optimizing ATAC-seq parameters in *Drosophila* presents a challenge, especially to smaller labs with limited funding, due to the cost of ATAC-seq library preparation and sequencing. Thus, our goal was to identify optimal conditions for nuclei isolation, purification, and ATAC-seq library preparation using *Drosophila*.

We first asked if the number of flies per sample plays a role in the number of extracted nuclei. As expected, using a larger volume of flies increases the number of extracted nuclei ($F(1,8) = 15.61$, $p = 0.0042$). However, the size of the Dounce homogenizer used for extraction does not significantly affect the total extracted nuclei ($F(1,8) = 1.984$, $p = 0.1966$; Fig. 1). Thus, using more flies per sample increases the likelihood of obtaining sufficient nuclei for downstream ATAC-seq library preparation.

One key advantage to using *Drosophila* for ATAC-seq library preparation is the availability of sophisticated genetic tools that allow easy identification of target cells. For example, it is easy to label nuclei of interest with genetically-encoded GFP that can be isolated by flow cytometry. One parameter that greatly influences the false-positive rate and nuclei purity is the gating scheme used for sorting. In all our experiments, we evaluated GFP+ nuclei compared to GFP-/DAPI+ control nuclei using two gating schemes: a "low" gate, which is less stringent, but allows faster sorting, and a "high" gate, which increases specificity at the cost of slower sorting speed (Fig. 2). To determine which is the optimal gating scheme, we evaluated the FANS specificity using the low- and high-gating schemes.

We used tubulin (*tub*)-Gal4 to drive UAS-GFP and analyzed nuclei suspensions. First, control nuclei (GFP-) were identified using forward scatter and DAPI staining (Fig. 2A-B). Then, we used a post-sort purity test to evaluate specificity. In this test, we collected

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

singlet GFP+/DAPI+ nuclei and then evaluated the purity of thecollected nuclei by running them through the flow cytometer a second time. Using the low-stringency gate, we recorded data from 99,941 singlet nuclei. Of these, 79,051 were GFP+ (79%) (Fig. 2C-D). Then, we collected 451,000 nuclei for post-sort purity evaluation. We counted 249,544 nuclei (55% recovery), of which 38,298 were GFP+ (15%) (Fig. 2E-F). Using the high-stringency gate, we evaluated 99,877 nuclei and observed that 13,291 were GFP+ (13%) (Fig. 2G-H). We then collected 137,984 nuclei and determined that 83,046 nuclei (60%) were recovered in the post-sort purity test. 59,668 nuclei in the post-sort purity test were GFP+ (72%) (Fig. 2I-J). These results indicate that sorting nuclei for ATAC-seq using a high-stringency gating scheme provides higher specificity.

In our nuclei isolation procedure, flies are frozen at -80 °C for 5 min, shaken, and passed through a fly sieve to isolate the tissue of interest (bodies vs. heads). Because GFP fluorescence is used to identify target nuclei, we examined whether freezing time influences GFP fluorescence. In nuclei sorted using the low gate, we observed that GFP fluorescence decreased by ~35% and ~30% after freezing for 10 or 20 min at -80 °C, respectively, but these differences were not significantly different than freezing for 5 min. Freezing for 40 min caused a ~44% decrease in GFP fluorescence ($p < 0.05$). Using the high gate, we observed significantly decreased GFP fluorescence after freezing for 10 min ($p < 0.05$; Fig. 3), but GFP fluorescence was not significantly decreased after freezing for 20 or 40 min. Therefore, we conclude that flies should be frozen at -80 °C for 5 min or less when used for nuclei isolation and ATAC-seq library preparation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Retinal pigment is another potential confounding factor. In flies, eye color is routinely used to identify genotypes of interest. Because our nuclei isolation procedure uses fly heads as the starting tissue, rather than dissected brains, we reasoned that the presence of eye pigment may decrease sorting specificity. Thus, we used flies with intense, wild-type red eye color (w^+ or w^+/w^-) and flies with orange eyes (w^- with two transgenes marked with w^{+mC}) for nuclei isolation. When nuclei were evaluated with the low gate, we observed that the GFP/DAPI ratio was significantly higher in the red-eyed flies than in the orange-eyed flies (p = 0.02; Fig. 4). However, using the high gate, we saw no significant difference in the GFP/DAPI ratio. These results indicate that red-eyed flies lead to a higher false-positive rate when using a less stringent gating scheme for sorting. Thus, using flies with more intense eye color requires sorting with increased stringency. Based on these results, we hypothesized that increasing the number of washes during nuclei isolation would mitigate the false positive rate caused by increased red retinal pigment. Washing the crude nuclei extracts one, three, or five times (Fig. 5A, C) dose-dependently reduced the GFP/DAPI ratio in $w^+;w^{+mC}$ and $w^-;w^{+mC}$ flies (Fig. 5B, D). To determine the optimal number of washes for each eye color, we performed linear regression analysis of the GFP/DAPI ratio in nuclei isolated from $w^+;w^{+mC}$ and $w^-;w^{+mC}$ flies evaluated using low- and high-stringency gating schemes. We observed that the regression lines were significantly different between $w^+;w^{+mC}$ and $w^-;w^{+mC}$ samples when evaluated with each gating scheme (p = 0.0008, low gate; p = 0.02, high gate; ANCOVA). These results indicate that when sorting using a less-stringent gating scheme, at least 5 washes are needed, especially for $w^+;w^{+mC}$ flies. For FANS with high-stringency gating, at least three washes are necessary.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The published ATAC-seq guidelines are all based on mammalian genomes, which are approximately 10-fold larger than the *Drosophila* genome (Canapa et al., 2016; Vinogradov, 2004). Therefore, the parameters used for tagmentation with Tn5 (reaction time and enzyme concentration) in mammalian cells may not be appropriate for flies. For these experiments, we evaluated the library quality using a Bioanalyzer, as recommended by Buenrostro et al. (2015). In an excellent ATAC-seq library, a periodic banding pattern should be present, indicating the presence of mono-, di-, and tri-nucleosome fragments in addition to nucleosome-free fragments. We reasoned that examining the expected banding pattern of each library would indicate the library quality, and thus could be used to determine the optimal reaction condition for ATAC-seq library prep in flies. First, we examined how changing the reaction time altered the fragment distribution within the library by performing the tagmentation reaction using 1X Tn5 enzyme and incubating the samples for 5, 23, and 60 min (Fig. 6A-C). Low tagmentation time increased the proportion of 600-1000 bp fragments in the library, suggesting that tri- and multi-nucleosome fragments are over-represented (Fig. 6A). Incubation for 23 min provided a fragment distribution with nucleosome-free and mono-nucleosome fragments, with decreasing amounts of higher-nucleosome fragments, as expected (Fig. 6B). Increasing the tagmentation time biased the library toward smaller fragment sizes, indicating that extended incubation causes DNA over-digestion (Fig. 6C). Thus, for library preparation using *Drosophila* nuclei, a 23-minute tagmentation step is optimal.

Another factor affecting library preparation is the enzyme concentration used for tagmentation. We examined the effect of Tn5 concentration by preparing libraries with

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

0.1X, 0.3X, 1X, and 3X Tn5 (Fig 6D-G). The reaction time was 23 minutes. We observed the major peak at approximately 200 bp in each sample, which corresponds to nucleosome-free regions. Increasing the Tn5 concentration increased the proportion of 600-1000 bp fragments and decreased the width of the major peak.

Next, we examined whether the number of PCR cycles after tagmentation affects library quality. Like other sequencing library types, ATAC-seq libraries must be amplified to obtain sufficient DNA concentrations for sequencing. Buenrostro et al. (2013, 2015) suggested that ATAC-seq libraries should be amplified to 33% of total fluorescence in a qPCR side reaction to ensure that sufficient material is present for sequencing, but does not introduce GC and size bias to the libraries. We hypothesized that the smaller genome size in *Drosophila* would require fewer PCR cycles than libraries derived from mammalian nuclei. We amplified tagmented DNA to 25%, 33%, and 50% of total fluorescence in a qPCR side reaction to determine the number of additional PCR cycles needed for library amplification. In the 25%-amplified library, the fragment distribution did not show a PCR bubble and the main peak was wide with a regular slope (Fig. 6H), indicating efficient amplification that did not add bias to the library. In the 33%-amplified library, we observed a large proportion of fragments at 1000-1200 bp (Fig. 6I), suggesting the presence of a PCR bubble that arose from depleted PCR reagents (most likely primers)(Kanagawa, 2003). Libraries amplified by 50% showed a narrow major peak and no PCR bubble (Fig. 6J), suggesting that the fragments in the PCR bubble were likely resolved by the additional PCR cycles. These results indicate that ATAC-seq libraries derived from *Drosophila* nuclei should be amplified to 25% of total qPCR

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fluorescence, though 33%-amplified libraries may also be suitable for sequencing after reconditioning PCR (Thompson et al., 2002).

Finally, we combined our optimized parameters to generate ATAC-seq libraries. We used vGAT-Gal4;UAS-eGFP-nls flies, which express nuclear GFP specifically within GABAergic neurons. We isolated the nuclei after freezing approximately 1 mL (300-400) flies at -80 °C for 5 min. The crude nuclei extracts were washed three times with wash buffer and sorted by FANS using a high-stringency gating scheme. This returned 65,000 nuclei per sample, which were tagmented for 23 min using 1X Tn5 (from an Illumina Nextera kit). The resulting libraries were amplified to 25% qPCR fluorescence (8 cycles) before sequencing. We observed the expected nucleosomal pattern when quality checked using a Tapestation instrument (Figure 7A), which was also present in the sequencing data (Figure 7B). These results indicate that our protocol for generating ATAC-seq libraries from *Drosophila* tissues provides excellent library quality and sequencing data.

DISCUSSION

ATAC-seq is a powerful way to interrogate the chromatin state within tissues of interest. Critically, the technique provides key advantages over previous methods, such as ChIP-seq, MNase-seq, or FAIRE-seq, to assay the chromatin state. For example, ATAC-seq has low input requirements (50,000-60,000 nuclei vs. 1-10 million nuclei for MNase- or FAIRE-seq) (Kidder et al., 2011; McKay and Lieb, 2013; Simon et al., 2012), has less hands-on time, does not use harsh chemicals such as paraformaldehyde, and does not

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

require antibody optimization and complicated pull-down methods. Thus, ATAC-seq is quickly becoming the assay of choice for chromatin investigations.

Most published studies using ATAC-seq focus on either tumor tissue or white blood cells, though recent studies have investigated the brain and other tissues in mice and humans (Liu et al., 2019; Rocks et al., 2021). Indeed, the initial study describing ATAC-seq was performed in mammalian tissue (Buenrostro et al., 2013), and published protocols have been optimized for use in these tissues (Buenrostro et al., 2018; Corces et al., 2017). *Drosophila melanogaster* is a powerful model organism that is routinely used in studies of development and human disease models. Importantly, *Drosophila* model systems enable numerous genetic tools that are not available in mammalian systems. For example, the Gal4-UAS system allows transgenes to be expressed in specific cell types using cell-specific drivers (Duffy, 2002). The Gal4 system can be used to express RNAi, CRISPR/Cas9, and temperature-sensitive transgenes, thereby allowing specific temporal and spatial manipulation. The Gal4-UAS system has also been adapted to use combinatorial approaches that allow transgene expression within only a few cells (Dionne et al., 2018). Thus, the ability to examine the chromatin state in this powerful model system will allow detailed investigations of chromatin-mediated gene regulation. Our protocol was designed specifically to harness *Drosophila* model systems, thus taking advantage of the powerful genetic tools available in flies. We used Gal4 to drive GFP coupled to a nuclear localization signal within nuclei of interest, which were then purified by FANS. While we used neuron-specific drivers to express GFP, this approach can be used to isolate nuclei from any other tissue in the fly. Our results

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

indicate that our nuclei isolation procedure is an effective way to collect genetically-tagged *Drosophila* nuclei for downstream ATAC-seq library preparation.

Using *Drosophila* for ATAC-seq experiments presents several complicating factors. First, the number of cells available in a single fly is several orders of magnitude lower than the number of cells in mammals. This complication is mitigated by high fecundity and a rapid life cycle. The ability to rapidly produce large numbers of flies in a relatively quick period makes it easy to achieve the necessary nuclei numbers for ATAC-seq, even for rare cell populations.

Another complicating factor in using *Drosophila* for ATAC-seq studies is the fly cuticle. The chitin exoskeleton prevents easy dissection of target tissues or enzyme-mediated tissue digestion. To overcome this limitation, we homogenized flies with a Dounce homogenizer, with a filtration step between homogenization with the "A" and "B" pestles. Dounce homogenization with the "A" pestle disrupts the fly cuticle while maintaining nuclear integrity. Indeed, when homogenizing isolated fly heads, we noticed intact, empty fly heads in the crude homogenate, indicating that the tissue inside the head is pushed out of the cuticle during homogenization. Passing the crude homogenate through a cell strainer eliminates much of the cuticle debris. Subsequent Dounce homogenization with the "B" pestle and ATAC-seq lysis buffer, which was specifically designed to disrupt the plasma membrane without lysing the nuclear membrane (Corces et al., 2017), provides an efficient way to isolate healthy nuclei for FANS.

Eye color is routinely used to identify flies with the correct genotype. Retinal pigments present a complicating factor because several *Drosophila* eye pigments are fluorescent (Franceschini et al., 1981; Miller et al., 1984), which may cause false-positive results

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

during nuclei isolation via FANS. False-positive nuclei will decrease the number of real nuclei per sample and increase the likelihood of ATAC-seq library failure. Washing the crude nuclei extract several times with wash buffer removes retinal pigments, cuticle fragments, and cellular debris, thereby decreasing the false-positive rate during FANS. Importantly, these washes do not increase nuclei loss. To further minimize nuclei loss due to sample transfer, we suggest performing all wash steps in the same tube. This tube can even be used to resuspend the washed nuclei, which can then be used for FANS.

The *Drosophila* genome is ~10-fold smaller than mammalian genomes, so enzyme-mediated steps during ATAC-seq library preparation may be affected by the amount of available DNA. Tn5 tagmentation in mammalian samples is generally performed for 30 min at 37 °C. Because the amount of accessible chromatin in *Drosophila* nuclei may be lower than in mammalian nuclei, we hypothesized that extended tagmentation time would bias the library toward smaller fragment sizes. Indeed, increasing the tagmentation time decreases library complexity, specifically increasing the number of nucleosome-free fragments. Tagmentation for only 5 min increased the number of mid-length fragments, while tagmentation for 23 min produced a library that most closely resembled the expected nucleosome periodicity.

Established ATAC-seq protocols in mammalian tissues use 2.5 µL Tn5 transposase included in Illumina Nextera kits or made in-house. Previous studies showed that decreasing the Tn5 concentration in mouse embryonic stem cells only decreases tagmentation efficiency when diluted to 10 nM or lower (Corces et al., 2017). Notably, this experiment used homemade Tn5 (Picelli et al., 2014), which may have different

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

kinetics or concentrations than the Tn5 included in the Illumina Nextera kit. Thus, we examined whether varying the Tn5 concentration affects ATAC-seq library quality in *Drosophila* samples. We generated libraries using dilute or concentrated Tn5 and observed few differences in library quality using 0.1X, 0.3X, or 1X Tn5. Using 3X Tn5 produced a library with an increased proportion of 600-1500 bp fragments. Thus, increasing the Tn5 concentration may prevent enzyme binding to smaller open chromatin regions such as mono-and di-nucleosome fragments, possibly via steric hindrance, thereby reducing Tn5 digestion and misrepresenting the chromatin structure.

The last parameter we examined was the number of PCR cycles needed to amplify ATAC-seq libraries before sequencing. Protocols for generating ATAC-seq libraries from mammalian tissues call for amplification to one-third the total fluorescence in a qPCR side reaction (Buenrostro et al., 2015). Because of the smaller genome size in *Drosophila*, amplifying the libraries to one-third of total qPCR fluorescence may introduce GC and size bias to the libraries. Indeed, amplifying our libraries to one-third qPCR fluorescence introduced a PCR bubble, indicating that one or more reagents in the PCR reaction (most likely primers) were depleted during amplification (Kanagawa, 2003). This causes further amplification to occur using annealed adaptor sequences, which may increase the number of mid-size fragments in the library. These libraries may still be adequate for sequencing after one cycle of reconditioning PCR (Thompson et al., 2002), but these libraries should be quality-checked again to verify that the PCR bubble was resolved before sequencing. Amplifying the libraries for 25% of qPCR fluorescence prevents the development of a PCR bubble while preserving a good fragment size distribution. In our experiments, 25% qPCR fluorescence generally corresponds to 8-9

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cycles and 33% qPCR fluorescence corresponds to 10-11 cycles. Based on these results, we recommend amplifying ATAC-seq libraries to 25% qPCR fluorescence, which preserves library complexity and does not introduce GC or size bias.

We used our optimized protocol to generate ATAC-seq libraries from nuclei isolated from *Drosophila* GABA neurons. These libraries were of excellent quality, with the expected nucleosomal banding patterns and fragment size distributions. Additionally, these libraries had high alignment rates (92.4-96.5%), further confirming their quality. Therefore, we propose that ATAC-seq libraries from *Drosophila* tissues should use the following parameters: freezing at -80 °C for no more than 5 min, at least three washes, 23-min tagmentation time with 1X Tn5, and amplification to 25% qPCR fluorescence. Our protocol harnesses the powerful genetic tools available in *Drosophila* to achieve excellent cell-type specificity. Then, ATAC-seq is used to interrogate the chromatin landscape with unprecedented resolution. Thus, our protocol will enable detailed investigations of the role of chromatin-mediated gene regulation in normal and pathological states.

ACKNOWLEDGMENTS

This work was also supported by the University of Utah Flow Cytometry Facility, the University of Utah Genomics Core Facility, the High Throughput Sequencing Core at the Huntsman Cancer Institute, and the National Cancer Institute through Award Number 5P30CA042014-24. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FUNDING

This study was supported by grants from the National Institute on Drug Abuse (Grant #5R21DA049635-02, to A.R.) and the National Institute on Alcohol Abuse and Alcoholism (Grant #5R01AA019526-10, to A.R.).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

Bozek, M., Cortini, R., Storti, A.E., Unnerstall, U., Gaul, U., and Gompel, N. (2019). ATAC-seq reveals regional differences in enhancer accessibility during the establishment of spatial coordinates in the Drosophila blastoderm. Genome Res. *29*, 771–783.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. In Current Protocols in Molecular Biology, (Hoboken, NJ, USA: John Wiley & Sons, Inc.), pp. 21.29.1-21.29.9.

Buenrostro, J.D., Corces, M.R., Lareau, C.A., Wu, B., Schep, A.N., Aryee, M.J., Majeti, R., Chang, H.Y., and Greenleaf, W.J. (2018). Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell *173*,

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1535-1548.e16.

Bysani, M., Agren, R., Davegårdh, C., Volkov, P., Rönn, T., Unneberg, P., Bacos, K., and Ling, C. (2019). ATAC-seq reveals alterations in open chromatin in pancreatic islets from subjects with type 2 diabetes. Sci. Rep. *9*.

Canapa, A., Barucca, M., Biscotti, M.A., Forconi, M., and Olmo, E. (2016). Transposons, genome size, and evolutionary insights in animals. Cytogenet. Genome Res. *147*, 217–239.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962.

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914.

Cusanovich, D.A., Reddington, J.P., Garfield, D.A., Daza, R.M., Aghamirzaie, D., Marco-Ferreres, R., Pliner, H.A., Christiansen, L., Qiu, X., Steemers, F.J., et al. (2018). The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature *555*, 538–542.

Davie, K., Jacobs, J., Atkins, M., Potier, D., Christiaens, V., Halder, G., and Aerts, S. (2015). Discovery of Transcription Factors and Regulatory Regions Driving In Vivo Tumor Development by ATAC-seq and FAIRE-seq Open Chromatin Profiling. PLoS Genet. *11*, 1–24.

Dionne, H., Hibbard, K.L., Cavallaro, A., Kao, J.C., and Rubin, G.M. (2018). Genetic

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

reagents for making split-GAL4 lines in Drosophila. Genetics *209*, 31–35.

Duffy, J.B. (2002). GAL4 system in Drosophila: A fly geneticist's Swiss army knife. Genesis *34*, 1–15.

Franceschini, N., Kirschfeld, K., and Minke, B. (1981). Fluorescence of photoreceptor cells observed in vivo. Science (80-.). *213*, 1264–1267.

Fujiwara, S., Baek, S., Varticovski, L., Kim, S., and Hager, G.L. (2019). High Quality ATAC-Seq Data Recovered from Cryopreserved Breast Cell Lines and Tissue. Sci. Rep. *9*, 1–11.

Haines, J.E., and Eisen, M.B. (2018). Patterns of chromatin accessibility along the anterior-posterior axis in the early Drosophila embryo. PLoS Genet. *14*.

Halstead, M.M., Kern, C., Saelao, P., Chanthavixay, G., Wang, Y., Delany, M.E., Zhou, H., and Ross, P.J. (2020). Systematic alteration of ATAC-seq for profiling open chromatin in cryopreserved nuclei preparations from livestock tissues. Sci. Rep. *10*.

Ho, Y.T., Shimbo, T., Wijaya, E., Ouchi, Y., Takaki, E., Yamamoto, R., Kikuchi, Y., Kaneda, Y., and Tamai, K. (2018). Chromatin accessibility identifies diversity in mesenchymal stem cells from different tissue origins. Sci. Rep. *8*, 1–11.

Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). J. Biosci. Bioeng. *96*, 317–323.

Kidder, B.L., Hu, G., and Zhao, K. (2011). ChIP-Seq: Technical considerations for obtaining high-quality data. Nat. Immunol. *12*, 918–922.

Koenecke, N., Johnston, J., Gaertner, B., Natarajan, M., and Zeitlinger, J. (2016). Genome-wide identification of Drosophila dorso-ventral enhancers by differential histone acetylation analysis. Genome Biol. *17*, 1–19.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Liu, C., Wang, M., Wei, X., Wu, L., Xu, J., Dai, X., Xia, J., Cheng, M., Yuan, Y., Zhang, P., et al. (2019). An ATAC-seq atlas of chromatin accessibility in mouse tissues. Sci. Data *6*.

Lu, Z., Hofmeister, B.T., Vollmers, C., DuBois, R.M., and Schmitz, R.J. (2017). Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Res. *45*.

McKay, D.J., and Lieb, J.D. (2013). A Common Set of DNA Regulatory Elements Shapes Drosophila Appendages. Dev. Cell *27*, 306–318.

Miller, G. V, Itoku, K.A., Fleischer, A.B., and Stark, W.S. (1984). Studies of fluorescence in Drosophila compound eyes: changes induced by intense light and vitamin A deprivation (Springer-Verlag).

Picelli, S., Björklund, A.K., Reinius, B., Sagasser, S., Winberg, G., and Sandberg, R. (2014). Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res. *24*, 2033–2040.

Quillien, A., Abdalla, M., Yu, J., Ou, J., Zhu, L.J., and Lawson, N.D. (2017). Robust Identification of Developmentally Active Endothelial Enhancers in Zebrafish Using FANS-Assisted ATAC-Seq. Cell Rep. *20*, 709–720.

Rocks, D., Jaric, I., Tesfa, L., Greally, J.M., Suzuki, M., and Kundakovic, M. (2021). Cell type-specific chromatin accessibility analysis in the mouse and human brain. Epigenetics 1–18.

Simon, J.M., Giresi, P.G., Davis, I.J., and Lieb, J.D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat. Protoc. *7*, 256–267.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Thompson, J.R., Marcelino, L.A., and Polz, M.F. (2002). Heteroduplexes in mixed-template amplifications: Formation, consequence and elimination by “reconditioning PCR.” Nucleic Acids Res. *30*, 2083–2088.

Vinogradov, A.E. (2004). Evolution of genome size: Multilevel selection, mutation bias or dynamical chaos? Curr. Opin. Genet. Dev. *14*, 620–626.

Wang, J., Zibetti, C., Shang, P., Sripathi, S.R., Zhang, P., Cano, M., Hoang, T., Xia, S., Ji, H., Merbs, S.L., et al. (2018). ATAC-Seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. Nat. Commun. *9*, 1–13.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

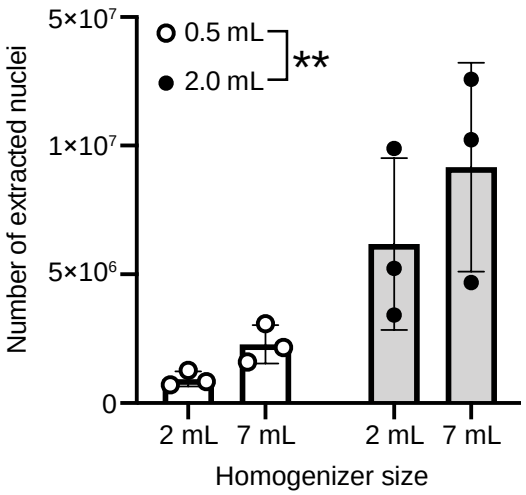

Figure 1. Fly volume increases the nuclei yield. The number of flies used for nuclei isolation increases the number of extracted nuclei, but the homogenizer size does not. Data represent the mean ± SD of three independent experiments. $F(1,8) = 15.61$, $^{**}p = 0.0042$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

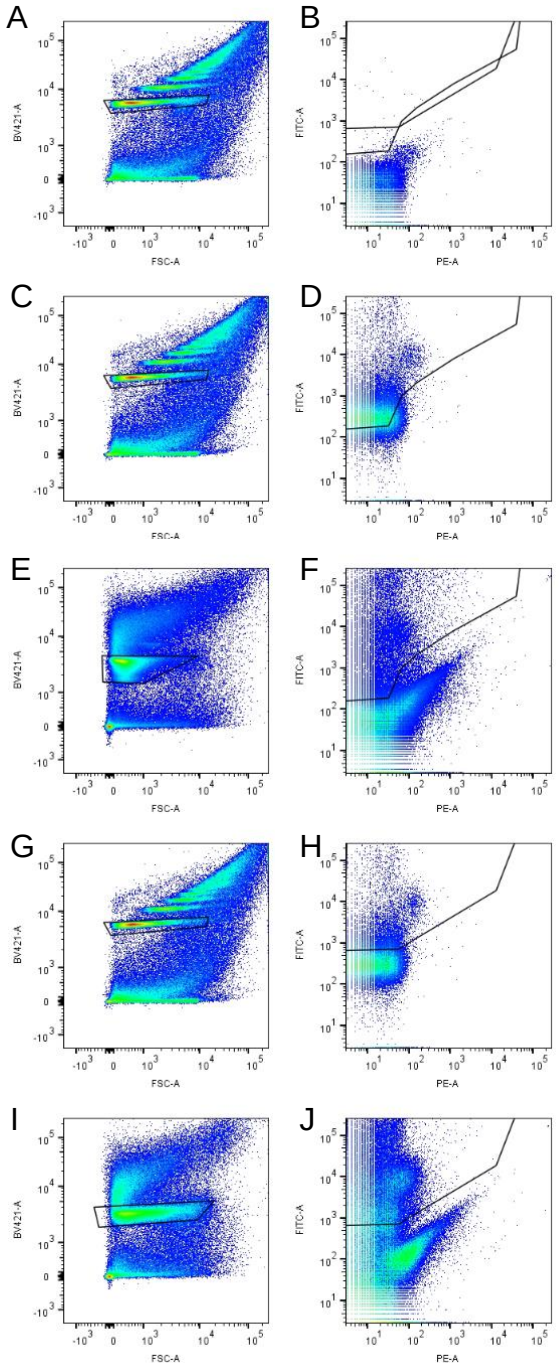

Figure 2. Sorting nuclei using a high-stringency gating scheme increases specificity. Representative dotplots of DAPI+ nuclei (A-B), nuclei sorted with a low-stringency gate (C-F), and nuclei sorted with a high-stringency gate (G-J). Foward scatter vs. BV421 nm fluorescence (DAPI) (left panels). FITC-A (GFP) vs. PE-A (right

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

panels). Black lines represent various sorting gates. Singlet nuclei (enclosed population in left panels) were selected for sorting using the presence of GFP (enclosed population in right panels).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

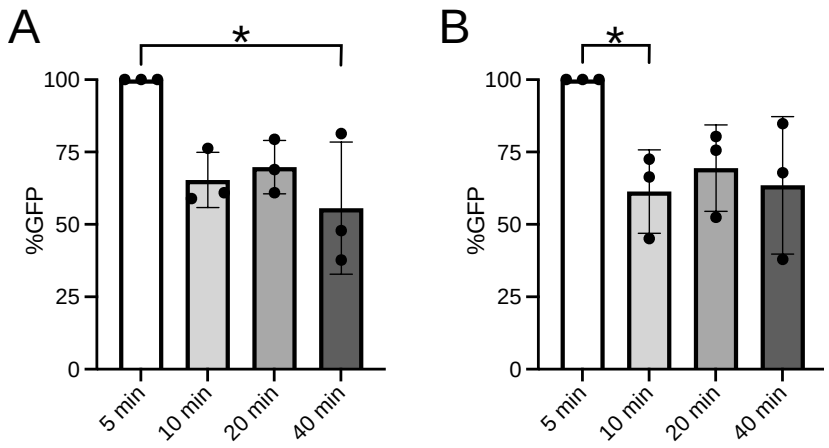

Figure 3. Freezing at -80 °C reduces GFP fluorescence. Flies were frozen at -80 °C and the heads isolated for nuclei isolation by flow cytometry. A) Sorting using a less-stringent (low gate) or B) more-stringent (high gate) gating scheme. Freezing reduces the detection of GFP-positive nuclei by approximately 30% when sorting using the low or high gate scheme. Data represent mean ± SD. Differences were analyzed by Kruskal-Wallis test with one-way ANOVA with post hoc Dunn's multiple comparisons test. $H(4) = 6.787$, $p < 0.05$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

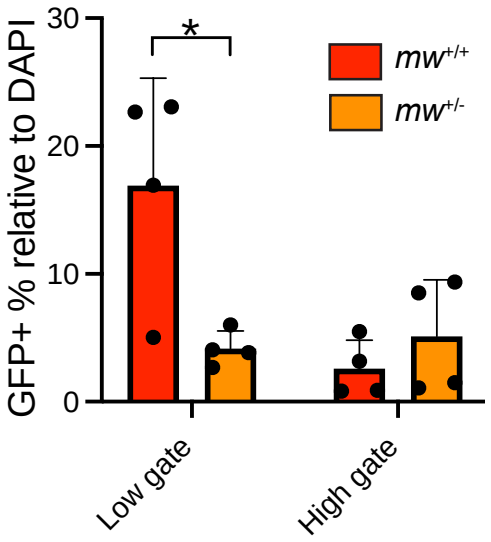

Figure 4. Drosophiia eye color increases the false positive rate for fluorescent activated nuclei sorting. Cha-eGFP-nls flies (mini white+ were crossed to Canton S (mini white+/+) or Berlin (mini white-/-) flies to produce progeny with red (mw+/+) and orange (mw+/-) eyes, respectively. The number of GFP+ nuclei was significantly higher in red-eyed flies when the nuclei were sorted using the low gate. Data shown are the mean ± SD of three independent experiments. Student's t-test, *$p < 0.05$.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

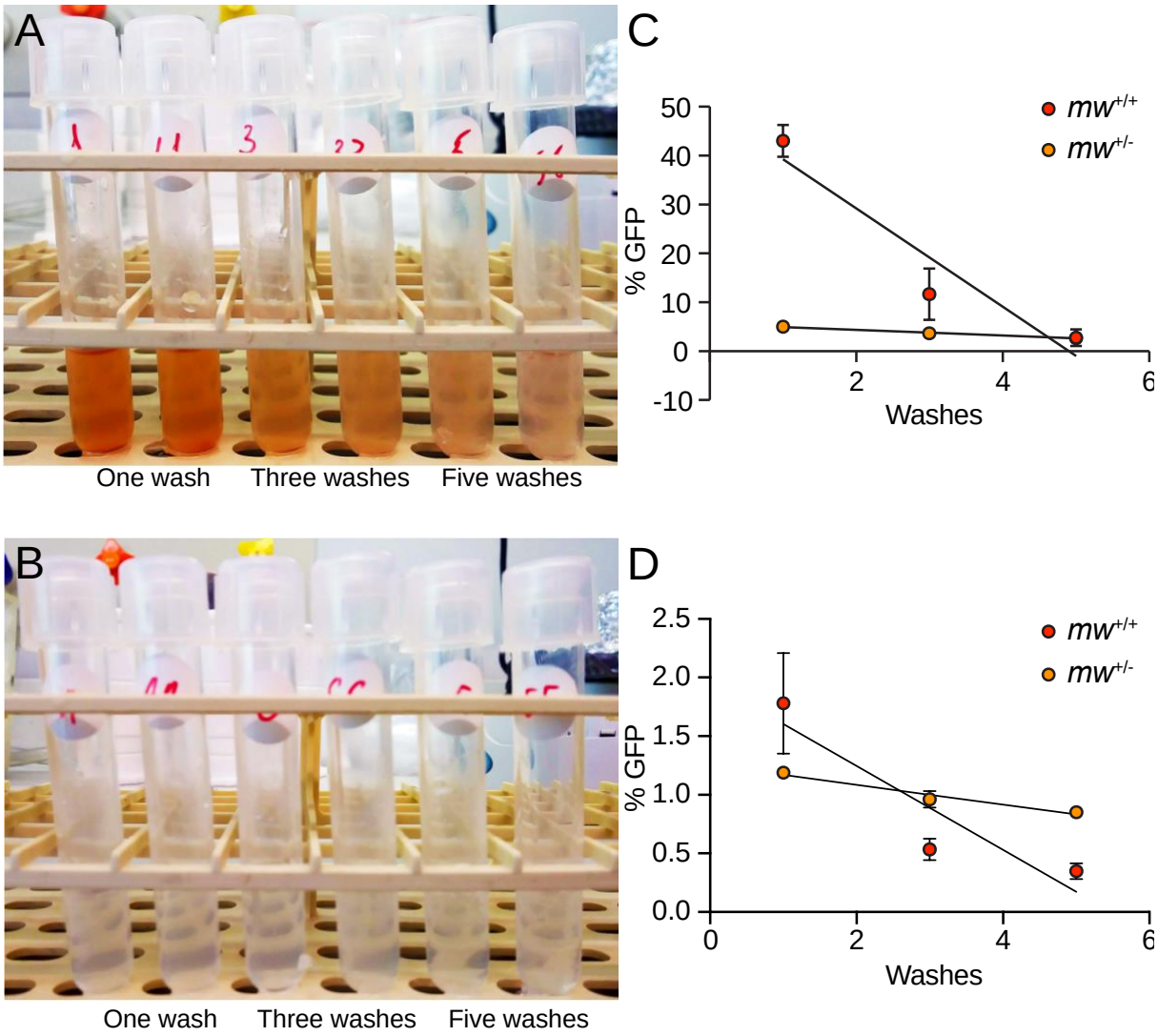

Figure 5. Increasing washes reduces the false positive rate while sorting. Cha-Gal4 UAS-GFP-nls (w^{+mC}) flies were crossed to w^{+} Canton S or w* Berlin ($mw^{-/-}$) flies to produce progeny with red and orange eyes, respectively. Crude homogenates from A) Canton S and B) Berlin were washed 1, 3, or 5 times before sorting. The number of false-positive GFP+ nuclei decreases with more washes, but not the total number of nuclei in red- and orange-eyed flies. Percent GFP+ to DAPI+ nuclei from red- and orange-eyed flies sorted with C) low and D) high stringency gating. Data represent the mean ± SD of two replicates. The regression line slopes were compared by analysis of

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

covariance (ANCOVA). The slopes between $mw^{+/+}$ and $mw^{+/-}$ flies were significantly different for nuclei sorted using low (p = 0.0008) and high (p = 0.02) stringency gating schemes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

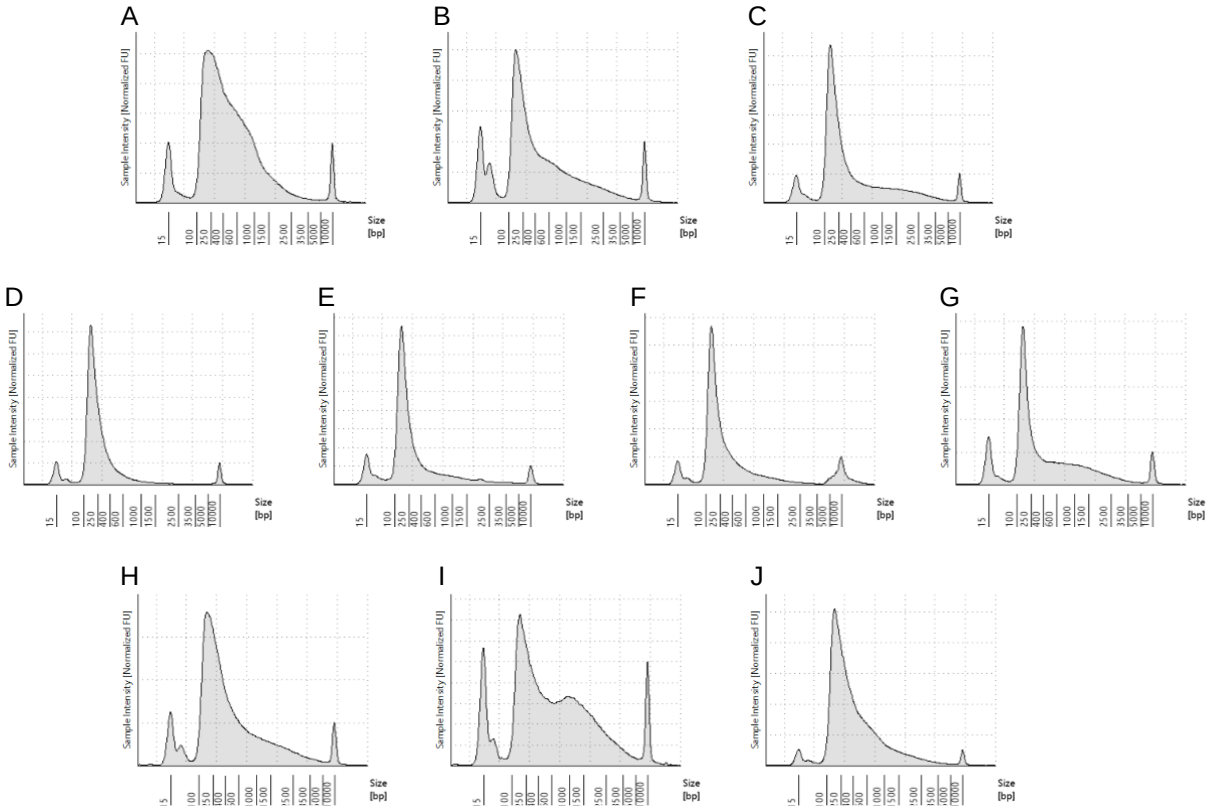

Figure 6. Reaction conditions during tagmentation and amplification affect ATAC-seq library quality. Enzymatic reaction parameters were varied to identify the optimal conditions for ATAC-seq library preparation from Drosophila nuclei. A representative Tapestation trace is shown for each condition. A-C) tagmentation time. A) 5 min, B), 23 min, C) 60 min. D-G) Tn5 concentration. D) 0.1X, E) 0.3X, F) 1X, and G) 3X Tn5. H-J) PCR cycles. H) Number of cycles corresponding to 25% total fluorescence from a qPCR side reaction, I) 33% total fluorescence, and J) 50% total flourescence.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.05.442838; this version posted May 5, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

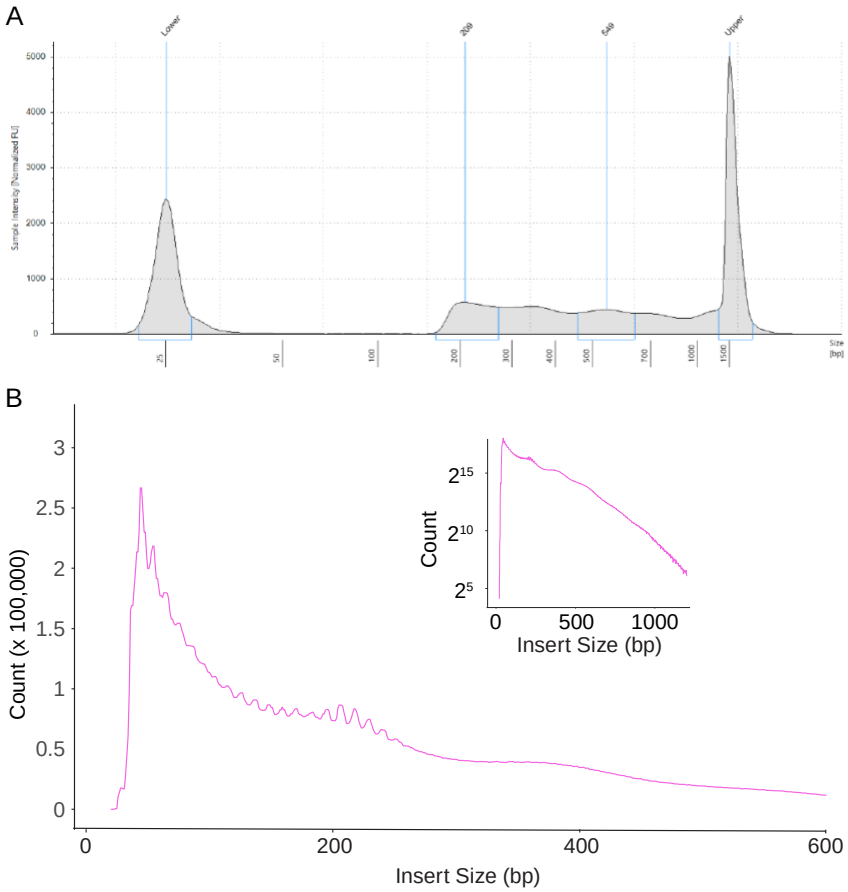

Figure 7. Optimized ATAC-seq parameters generate consistent libraries. Representative traces from an ATAC-seq library prepared from GABAergic neurons isolated from *Drosophila*. The libraries were prepared with vGAT-Gal4 x UAS-GFP-nls flies (orange eyes) after freezing at -80 C for 5 min. Tagmentation was performed for 23 minutes with 1X Tn5. The libraries were amplified for an additional 8 cycles, which corresponded to 25% total fluorescence from the qPCR side reaction. A) Tapestation trace. B) Insert size distribution showing the expected nucleosomal periodicity. Inset: Log-scaled insert size distribution.