UNIVERSITY OF BIRMINGHAM

University of Birmingham Research at Birmingham

Protocol for assay of transposase accessible chromatin sequencing in non-model species

Kissane, Stephen; Dhandapani, Vignesh; Orsini, Luisa

Link to publication on Research at Birmingham portal

Protocol

Protocol for assay of transposase accessible chromatin sequencing in non-model species

Stephen Kissane,
Vignesh
Dhandapani, Luisa
Orsini

l.orsini@bham.ac.uk

HIGHLIGHTS

ATAC-seq modified
protocol for
applications in non-
model species

Transposase titration
identifies thresholds
for optimal
transposition at lower
costs

Minimal number of
cells/tissue is
identified for low
input reactions

Proof of concept
ATAC-seq analysis for
the waterflea *Daphnia
magna*

The assay for transposase accessible chromatin (ATAC-seq) is a method for mapping genome-wide chromatin accessibility. Coupled with high-throughput sequencing, it enables integrative epigenomics analyses. ATAC-seq requires direct access to cell nuclei, a major challenge in non-model species such as small invertebrates, whose soft tissue is surrounded by a protective exoskeleton. Here, we present modifications of the ATAC-seq protocol for applications in small crustaceans, extending applications to non-model species.

STAR Protocols

Protocol

Protocol for assay of transposase accessible chromatin sequencing in non-model species

Stephen Kissane,[1,2] Vignesh Dhandapani,[1] and Luisa Orsini[1,3,*]

[1]School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK

[2]Technical contact

[3]Lead contact

*Correspondence: l.orsini@bham.ac.uk
https://doi.org/10.1016/j.xpro.2021.100341

SUMMARY

The assay for transposase accessible chromatin (ATAC-seq) is a method for mapping genome-wide chromatin accessibility. Coupled with high-throughput sequencing, it enables integrative epigenomics analyses. ATAC-seq requires direct access to cell nuclei, a major challenge in non-model species such as small invertebrates, whose soft tissue is surrounded by a protective exoskeleton. Here, we present modifications of the ATAC-seq protocol for applications in small crustaceans, extending applications to non-model species.
For complete information on the use and execution of this protocol, please refer to Buenrostro et al. (2013).

BEFORE YOU BEGIN

The assay for transposase accessible chromatin (ATAC-seq) has been optimized for mammalian cells to map genome-wide chromatin changes, allowing simultaneous interrogation of nucleosome positions in regulatory sites and genome-wide chromatin accessibility (Buenrostro et al., 2013; Buenrostro et al., 2015). More recently, the protocol has been applied to other model species, such as the fruitfly to discover transcriptional stages of aging in the brain (Davie et al., 2018), and mice to identify tissue-specific transcription factors (Liu et al., 2019). However, to investigate the chromatin rearrangement within cells, a direct *in vitro* transposition of sequencing adapters into native chromatin is needed. This requires between 500 and 5,000 free nuclei from cell suspensions obtained from blood or soft tissue; the sensitivity of detection diminishes with smaller input and quality of the input material, limiting the application of ATAC-seq in non-model species (Buenrostro et al., 2013).

Here, we present a modified ATAC-seq protocol for applications in small crustaceans, in which a chitin exoskeleton obstructs access to cell nuclei. The protocol is optimized for low input material, relevant for applications in non-model species in which starting material may be of limited quantity. The current costs of ATAC-seq are prohibitive for population-level applications because of the costs of the transposase enzyme. We perform a titration experiment demonstrating that the amount of the enzyme transposase required to obtain high-quality *in vitro* transposition can be significantly reduced without affecting data quality. We demonstrate the application of the modified protocol on the biomedical and ecological model species *Daphnia*. Like the model species *Drosophila*, *Daphnia* enjoys many technical advantages over vertebrate models: they are easy and inexpensive to culture in the laboratory, have a short life cycle, and produce large numbers of externally laid embryos. In addition, *Daphnia* has a parthenogenetic life cycle, allowing the rearing of populations of isogenic individuals (clones) from a single genotype. These properties enabled us to test different sample inputs and to quantify the quality of output without confounding factors associated with genetic variation among genotypes.

Experimental design considerations

The first step of this protocol requires the collection of sufficient tissue from *Daphnia* genotypes. Depending on the experimental design, multiple genotypes may be used. As epigenomic patterns may be influenced by environmental factors, special care should be adopted in controlling for environmental and developmental variation among genotypes and biological replicates. Before samples collection, *Daphnia* genotypes should be maintained in common garden conditions for at least two generations to reduce interference from maternal effect. When multiple genotypes are used, they should be synchronized and sampled at the same developmental stages. To prevent variation introduced by sex-specific expression, only females should be collected, following direct inspection of the transparent *Daphnia* body under a stereomicroscope. Populations of identical clones from the same genotype at the same developmental stage can be used to achieve sufficient starting material for nuclei extraction.

Material preparation

⊘ Timing: 0.5 h

Before starting the protocol, ensure that the following reagents and equipment are available:

1. 80% ethanol solution corresponding to 1 mL per sample.

   Note: It is recommended that a fresh ETOH solution is prepared daily to reduce alteration of ETOH final concentration due to evaporation.

2. Lysis tubes for tissue homogenization with Geno/Grinder or other suitable tissue homogenizer. These are prepared by adding 25–30 ceramic beads of 1.4 mm diameter to sterile tubes. Alternatively, purchase sterile tubes prefilled with ceramic beads of 1.4 mm diameter.

   Note: To avoid sample contamination, tubes are filled with ceramic beads under a laminar flow hood and stored in sealed sterile bags.

3. Ensure that two heat blocks are available and set at 65°C and 37°C, respectively.
4. Ensure that a centrifuge suitable for Eppendorf tubes is at 4°C for nuclei purification (step 2 below).
5. Ensure that ATAC sequencing primers are available (Table 1). Primers are used at a final concentration of 25 μM per sample.

   Note: For convenience, primers at a final concentration of 25 μM are pre-aliquoted in a 96-well plate to enable automated library preparation with a liquid handling robot or manual library preparation using a multichannel pipette.

   Note: To reduce index hopping between samples, a unique combination of paired end primers per sample should be preferred (Table 1). Index hopping has impacted NGS technologies from the time sample multiplexing was developed. It causes the incorrect assignment of libraries from the expected index to a different index in the multiplexed pool. Using unique dual indexes over combinatorial dual indexes alleviates this problem (Kircher et al., 2012)

6. Prepare the following solutions using sterile Milli-Q water and store up to 3 months at 18°C–20°C

Table 1. ATAC primers

Primer	Sequence F primer 5′-3′	Sequence R primer 5′-3′
ATAC-1	CAAGCAGAAGACGGCATACGAGATGACACAGTGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACGACACAGTTCGTCGGCAGCGTCAGATGTG
ATAC-2	CAAGCAGAAGACGGCATACGAGATGCATAACGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACGCATAACGTCGTCGGCAGCGTCAGATGTG
ATAC-3	CAAGCAGAAGACGGCATACGAGATACAGAGGTGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACACAGAGGTTCGTCGGCAGCGTCAGATGTG
ATAC-4	CAAGCAGAAGACGGCATACGAGATCCACTAAGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCCACTAAGTCGTCGGCAGCGTCAGATGTG
ATAC-5	CAAGCAGAAGACGGCATACGAGATTGTTCCGTGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTGTTCCGTTCGTCGGCAGCGTCAGATGTG
ATAC-6	CAAGCAGAAGACGGCATACGAGATGATACCTGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACGATACCTGTCGTCGGCAGCGTCAGATGTG
ATAC-7	CAAGCAGAAGACGGCATACGAGATAGCCGTAAGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACAGCCGTAATCGTCGGCAGCGTCAGATGTG
ATAC-8	CAAGCAGAAGACGGCATACGAGATCTCCTGAAGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCTCCTGAATCGTCGGCAGCGTCAGATGTG
ATAC-9	CAAGCAGAAGACGGCATACGAGATACGAATCCGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACACGAATCCTCGTCGGCAGCGTCAGATGTG
ATAC-10	CAAGCAGAAGACGGCATACGAGATAATGGTCGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACAATGGTCGTCGTCGGCAGCGTCAGATGTG
ATAC-11	CAAGCAGAAGACGGCATACGAGATCGCTACATGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCGCTACATTCGTCGGCAGCGTCAGATGTG
ATAC-12	CAAGCAGAAGACGGCATACGAGATCCTAAGTCGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCCTAAGTCTCGTCGGCAGCGTCAGATGTG
ATAC-13	CAAGCAGAAGACGGCATACGAGATTTGCTTGGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTTGCTTGGTCGTCGGCAGCGTCAGATGTG
ATAC-14	CAAGCAGAAGACGGCATACGAGATCCTGTCAAGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCCTGTCAATCGTCGGCAGCGTCAGATGTG
ATAC-15	CAAGCAGAAGACGGCATACGAGATAGCCTATCGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACAGCCTATCTCGTCGGCAGCGTCAGATGTG
ATAC-16	CAAGCAGAAGACGGCATACGAGATTGATCACGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTGATCACGTCGTCGGCAGCGTCAGATGTG
ATAC-17	CAAGCAGAAGACGGCATACGAGATCCACATTGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCCACATTGTCGTCGGCAGCGTCAGATGTG
ATAC-18	CAAGCAGAAGACGGCATACGAGATTCGAGAGTGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTCGAGAGTTCGTCGGCAGCGTCAGATGTG
ATAC-19	CAAGCAGAAGACGGCATACGAGATGGTCGTATGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACGGTCGTATTCGTCGGCAGCGTCAGATGTG
ATAC-20	CAAGCAGAAGACGGCATACGAGATACAGGCATGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACACAGGCATTCGTCGGCAGCGTCAGATGTG
ATAC-21	CAAGCAGAAGACGGCATACGAGATGTGATCCAGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACGTGATCCATCGTCGGCAGCGTCAGATGTG
ATAC-22	CAAGCAGAAGACGGCATACGAGATTTCGTACGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTTCGTACGTCGTCGGCAGCGTCAGATGTG
ATAC-23	CAAGCAGAAGACGGCATACGAGATATGACAGGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACATGACAGGTCGTCGGCAGCGTCAGATGTG
ATAC-24	CAAGCAGAAGACGGCATACGAGATCGACCTAAGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCGACCTAATCGTCGGCAGCGTCAGATGTG
ATAC-25	CAAGCAGAAGACGGCATACGAGATTATGGCACGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTATGGCACTCGTCGGCAGCGTCAGATGTG
ATAC-26	CAAGCAGAAGACGGCATACGAGATATAACGCCGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACATAACGCCTCGTCGGCAGCGTCAGATGTG
ATAC-27	CAAGCAGAAGACGGCATACGAGATGTAGTACCGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACGTAGTACCTCGTCGGCAGCGTCAGATGTG
ATAC-28	CAAGCAGAAGACGGCATACGAGATCGCGTATTGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACCGCGTATTTCGTCGGCAGCGTCAGATGTG
ATAC-29	CAAGCAGAAGACGGCATACGAGATATCCACGAGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACATCCACGATCGTCGGCAGCGTCAGATGTG
ATAC-30	CAAGCAGAAGACGGCATACGAGATTAACGTCGGTCTCGTGGGCTCGGAGATGT	AATGATACGGCGACCACCGAGATCTACACTAACGTCGTCGTCGGCAGCGTCAGATGTG

List of forward (F) and reverse (R) primer sequences used in the sequencing of ATAC-seq data on an Illumina platform. Primers used at 25μM final concentration are shown in 5′-3′ orientation.

Tris-HCl solution

Reagent	Final concentration (mM)	Volume (mL)/weight (g)
Tris-HCl	1,000	7.88 g
Nuclease-free H$_2$O	n/a	50 mL
Total	n/a	50 mL

NaCl solution

Reagent	Final concentration (mM)	Volume (mL)/weight (g)
NaCl	5,000	14.61 g
Nuclease-free H$_2$O	n/a	50 mL
Total	n/a	50 mL

MgCl$_2$ solution

Reagent	Final concentration (mM)	Volume (mL)/weight (g)
MgCl$_2$	300	1.43 g
Nuclease-free H$_2$O	n/a	50 mL
Total	n/a	50 mL

RSB buffer

Reagent	Final concentration (mM)	Volume (mL)
Tris-HCl	1,000	0.25
NaCl	5,000	0.05
MgCl$_2$	300	0.25
Nuclease-free H$_2$O	n/a	24.45
Total	n/a	25

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals		
Buffer EB	Qiagen	CAT# 19086
Ceramic beads (1.4 mm diameter; 325 g)	Qiagen	CAT# 13113-325
*Ethanol absolute, molecular biology grade (200 proof)	Fisher Scientific	CAT# 10517694
High sensitivity D1000 reagents	Agilent	CAT# 5067- 5585
High sensitivity D1000 ScreenTape	Agilent	CAT# 5067- 5584
Magbio HighPrep PCR beads	Auto Q Biosciences	CAT# AQ 60500
Illumina Tagment DNA TDE1 enzyme and buffer kits	Illumina	CAT# 20034198
Liquid Proteinase K	Fisher Scientific	CAT# NC1442588
*MasterPure Complete DNA and RNA Purification Kit	Lucigen	CAT# MC85200
*MgCl$_2$	Sigma-Aldrich	CAT# M8266-100G
*NaCl	Sigma-Aldrich	CAT# S7653-250G
NEBNext High-Fidelity 2× PCR Master Mix	New England Biolabs	CAT# M0541L
Qubit 1× dsDNA HS Assay Kit	Invitrogen	CAT# Q33231
*RNase A	Qiagen	CAT# 19101

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Screw cap tubes (2 mL)	VWR International	CAT# 211-0432
iTaq Universal SYBR Green Supermix	Bio-Rad	CAT# 172-5121
Tris-HCl	Sigma-Aldrich	CAT# 93363-500G
Tween 20	Sigma-Aldrich	CAT# P9416-100ML
Water DEPC treated	National Diagnostics	EC-625
Experimental models: organisms/strains		
Daphnia magna LRV0_1	University of Birmingham	NA
Daphnia magna LRV2_1	University of Birmingham	NA
Daphnia magna LRV3.5_15	University of Birmingham	NA
Daphnia magna LRV13_3	University of Birmingham	NA
Software and algorithms		
fastqc (V0.11.9)	(Andrews, 2010)	https://www.bioinformatics.babraham.ac.uk/projects/fastqc
Trimmomatic (v0.32)	(Bolger et al., 2014)	=http://www.usadellab.org/cms/?pagetrimmomatic
Galaxy pipeline	(Batut et al., 2018)	https://training.galaxyproject.org/
multiqc	(Ewels et al., 2016)	https://github.com/ewels/MultiQC/
ataqv	(Orchard et al., 2020)	https://github.com/ParkerLab/ataqv
Bowtie2 (v2.2.6)	(Langmead et al., 2019)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools	(Li et al., 2009)	http://www.htslib.org/doc/samtools.html
Picard tools	Online resource	https://broadinstitute.github.io/picard/
Qualimap 2	(Okonechnikov et al., 2016)	http://qualimap.conesalab.org/
deeptools alignmentSieve	(Ramirez et al., 2018)	https://deeptools.readthedocs.io/en/develop/content/tools/alignmentSieve.html
MACS2 callpeak	(Gaspar, 2018)	https://hbctraining.github.io/Intro-to-ChIPseq/lessons/05_peak_calling_macs.html
bedtools	(Quinlan, 2014)	https://bedtools.readthedocs.io/en/latest/
Hypergeometric Optimization of Motif EnRichment (HOMER)	(Duttke et al., 2019)	http://homer.ucsd.edu/homer/motif/

⚠ CRITICAL: An asterisk (*) marks chemicals or reagents that are harmful or toxic. In the following hazard and storage requirements are identified:

Ethanol: highly flammable, cause possible fire flashpoint, potential carcinogen. Necessary precautions should be adopted when handling; storage of large volumes should be in fire-proof cabinets.

$MgCl_2$: Irritant and corrosive. Protective gloves and lab coat should be wore at all times.

NaCl: Irritant. Protective gloves and lab coat should be wore at all times.

RNase A: May cause allergy or asthma if breathed. Avoid inhalation.

Masterpure kit: Danger, causes skin irritation, serious eye damage, harmful to aquatic life. Protective gloves and lab coat should be wore at all times. Disposal in clinical waste

MATERIALS AND EQUIPMENT

Equipment required is listed:

	Source	Identifier
Equipment		
Fine Balance	VWR	LA2541
Centrifuge compatible for Eppendorf tubes (5424R)	Eppendorf	5404000332
2020 Geno/Grinder	Spex Sample Prep	n/a
NanoDrop 8000	Labtech	ND-8000-GL
Nexus Gradient	Eppendorf	6331000041
Qubit 3.0 Fluorometer	Invitrogen	Q33216
Tapestation 2200	Agilent	G2964AA
Thermomixer C	Eppendorf	5382000031
Optional equipment		
Biomek FxP[a]	Beckman Coulter	A31844
Bioruptor Pico-sonication device	Diagenode	B01060010

aThe ATAC-seq protocol was optimized manually on a small number of samples. After this optimization step, a liquid handling Biomek FxP workstation was used to prepare large sample numbers. Any liquid handling robotic workstation can be used.

STEP-BY-STEP METHOD DETAILS

The modified ATAC protocol has the following main steps:

Cells extraction from tissue

⊙ **Timing: 1.5 h**

Step 1 enables the extraction of cells from tissues by shearing tissue and creating a cells suspension.

1. Flash-freeze 7 last instar female *Daphnia* juveniles in 2 mL VWR Screw Cap tubes containing 25–30 ceramic beads of 1.4 mm diameter. Store at –80°C.

   ⯀⯀ **Pause point:** The frozen tissue can be stored at −80° before further processing.

2. Thaw samples on ice.
3. Add 300 μL MasterPure Tissue and Cell Lysis Solution with 2.8 μL of proteinase K (50 U/mL).
4. Shear samples using an automated tissue homogenizer (e.g., Geno/Grinder) at 1,750 rpm for 30 s.
5. Place tubes in 65°C for 15 min, inverting 3 times during the incubation.
6. Cool samples at 37°C for 5 min.
7. Add 4 μL of 1 mg/mL RNase A.
8. Incubate at 37°C for 30 min.
9. Place on ice for 5 min.

   △ CRITICAL: The amount of tissue used to obtain sufficient nuclei is determined experimentally. A titration experiment using different amounts of tissue is recommended for this optimization step. For the protocol optimization, we used pools of last instar Daphnia juveniles ranging from 7 to 60 individuals. Please refer to the Expected outcomes section below for a description of results of the titration experiment.

Nuclei purification

⏱ Timing: 0.5 h

Step 10 enables the separation of cell material and proteins from the nuclei

10. Add 175 μL MPC protein precipitation reagent (MasterPure Kit) and mix well by gentle pipetting
11. Centrifuge at 10,000 × *g* for 10 min in a refrigerated centrifuge at 4°C

Note: If chromatin deconvolution is not efficient following the steps above, a pico-sonication step for DNA shearing (e.g., Bioruptor, Covaris) can be introduced. For the optimization of this protocol, two settings were used on the Bioraptor: 600 bp and 150 bp. No appreciable differences were observed in the DNA libraries and in the sequencing results between these two settings (see Expected outcomes).

12. Transfer supernatant to a clean tube and store at –20°C

▐▐Pause point: The purified nuclei can be stored at −20° until further processing.

13. Thaw samples on ice and transfer 1/3 of supernatant into tubes containing 1 mL of resuspension buffer (RSB; MasterPure Kit) and 0.1% Tween 20.

Note: All supernatant can be transferred into the resuspension buffer. It is possible to determine whether the whole resuspension is needed by testing DNA library quality when a proportion of the supernatant is used. Using only a fraction of the supernatant may provide sufficient material for technical replicates. .

14. Centrifuge at 500 × *g* for 10 min in a fixed-angle pre-chilled 4°C centrifuge to pellet the nuclei.
15. Discard Supernatant.

▐▐Pause point: The pelleted nuclei can be stored at −20° until further processing.

Transposition reaction optimization

⏱ Timing: 1.5 h

This step is only used in the optimization stage to reduce the amount of transposase enzyme without affecting data quality. A titration of the manufacturer recommended transposase enzyme concentration were used to identify the minimum amount required for a high-quality ATAC library. The results of this titration experiment are in Expected outcomes.

16. Immediately following the nuclei preparation, resuspend the pellet in different dilutions of the transposase reaction mix [25 μL 2× TD buffer; 0.63 μL transposase (100 nM final concentration); 22.5 μL of nuclease-free water].
17. Test different transposase concentrations. Here we tested four concentrations of the transposase enzyme as follows:
    a. 1/1 of the volume recommended: 25 μL 2× TD Buffer; 2.5 μL transposase; 22.5 μL of nuclease-free water.
    b. 1/2 of the volume recommended: 25 μL 2× TD Buffer; 1.25 μL transposase; 23.75 μL of nuclease-free water.
    c. 1/4 of the volume recommended: 25 μL 2× TD Buffer; 0.63 μL transposase; 24.37 μL of nuclease-free water.

d. 1/8 of the volume recommended: 25 μL 2× TD Buffer; 0.32 μL transposase; 24.68 μL of nuclease-free water.

18. Incubate the transposition reaction at 37°C for 30 min in a thermomixer using 1,000 rpm mixing speed

Note: Longer incubation times may be tested to improve the transposition efficiency. In the optimization phase, incubation of 16–20 h was tested. The Tapestation traces did not show significant difference between the two incubation times (see Expected outcomes).

19. Store samples at −20°C until further processing

⯈⯈⯈**Pause point:** The transposition reaction can be stored overnight or up to few days at −20° before progressing with the purification step.

20. Purify samples using a 1:1 volume Magbio HighPrep PCR beads (Auto Q Biosciences) as follows:
    a. Place Magbio HighPrep PCR beads at 18°C–20°C for 30 min
    b. Add nuclease-free water to the ligation reaction bringing the reaction volume to 100 μL.

    △ CRITICAL: Ensure that the final volume is 100 μL before adding Magbio HighPrep PCR beads.
    c. Add 100 μL (1.0×) of resuspended Magbio HighPrep PCR beads and mix well using a vortex mixer or a pipette (pipet gently at least 10 times).
    d. Incubate for 5 min at 18°C–20°C.
    e. Spin tubes in a microcentrifuge and place on a magnetic rack to separate beads from the supernatant. After the solution clears (about 5 min), discard the supernatant that contains unwanted fragments.

    △ CRITICAL: Do not discard the beads, which bind to the target fragments.
    f. Add 200 μL of freshly prepared 80% ethanol while the tubes are on the magnetic rack. Incubate at 18°C–20°C for 30 s.
    g. Carefully remove the supernatant without disturbing the beads and discard.
    h. Repeat washing step f.
    i. Spin briefly and place tubes onto the magnetic rack.
    j. Remove residual ETOH without disturbing the beads and air dry for 5 min.

    △ CRITICAL: Do not exceed 5 min; over drying the beads may result in lower recovery of target DNA.
    k. Remove tubes from the rack and elute DNA in 23 μL Buffer EB. Mix well on a vortex mixer or by pipetting, and incubate for 2 min at 18°C–20°C.
    l. Place tubes onto the magnetic rack until the solution appear clear.
    m. Transfer 20 μL of supernatant to a clean tube and discard the beads.

    ⯈⯈⯈**Pause point:** The purified samples can be stored overnight and up to few days at −20° before the transposition reaction.

Transposition reaction and purification

⏱ Timing: 1.5 h

Following the optimization in steps 16–20, in this step the transposition reaction is completed using the optimal transposase enzyme concentration, followed by a purification step.

21. Immediately following the nuclei preparation, resuspend the pellet in the transposase reaction mix optimized in step 3 [25 μL 2× TD buffer; 0.63 μL transposase (100 nM final concentration); 22.5 μL of nuclease free water].

> ⚠ CRITICAL: The pellet has to be mixed gently at least six times.

> *Note:* The transposase concentration used here is optimized for tissue obtained from 7 last instar Daphnia juveniles.

22. Incubate the transposition reaction at 37°C for 30 min in a thermomixer using 1,000 rpm mixing speed.
23. Store samples at −20°C until further processing

> ⚠⚠Pause point: The transposition reaction can be stored at −20° until further processing.

24. Purify samples using a 1:1 volume Magbio HighPrep PCR beads (Auto Q Biosciences) as follows:
    a. Place Magbio HighPrep PCR beads at 18°C–20°C for 30 min
    b. Add nuclease-free water to the ligation reaction bringing the reaction volume to 100 μL.

> ⚠ CRITICAL: Ensure that the final volume is 100 μL before adding Magbio HighPrep PCR beads.

    c. Add 100 μL (1.0×) of resuspended Magbio HighPrep PCR beads and mix well using a vortex mixer or a pipette (pipet gently at least 10 times).
    d. Incubate for 5 min at 18°C–20°C.
    e. Spin tubes in a microcentrifuge and place on a magnetic rack to separate beads from the supernatant. After the solution clears (about 5 min), discard the supernatant that contains unwanted fragments.

> ⚠ CRITICAL: Do not discard the beads, which bind to the target fragments.

    f. Add 200 μL of freshly prepared 80% ethanol while the tubes are on the magnetic rack. Incubate at room temperature for 30 s.
    g. Carefully remove the supernatant without disturbing the beads and discard.
    h. Repeat washing step f.
    i. Spin briefly and place tubes onto the magnetic rack.
    j. Remove residual ETOH without disturbing the beads and air dry for 5 min.

> ⚠ CRITICAL: Do not exceed 5 min; over drying the beads may result in lower recovery of target DNA.

    k. Remove tubes from the rack and elute DNA in 23 μL Buffer EB. Mix well on a vortex mixer or by pipetting, and incubate for 2 min at 18°C–20°C.
    l. Place tubes onto the magnetic rack until the solution appear clear.
    m. Transfer 20 μL of supernatant to a clean tube and discard the beads.

> ⚠⚠Pause point: The purified samples can be stored at −20°C before the amplification step.

Library amplification optimization

> ⏱ Timing: 2.5 h

An optimization of the PCR cycling of transposed DNA is advised as suboptimal amplification or overamplification can introduce biases in the DNA fragments distribution within the libraries. For example, in case of over amplification small DNA fragments are preferentially amplified. Consequently, the DNA libraries suffer from reduced complexity.

Complete 5 PCR cycles following the cycling conditions below and using NEBNext High-Fidelity PCR master mix and 25 μM Nextera PCR primers: 25 μL NEBNext High-Fidelity 2× PCR Master Mix; 10 μL transposed DNA; 2.5 μL (25 μM) Forward Primer (Table 1); 2.5 μL (25 μM) Reverse Primer; 10 μL nuclease-free H$_2$O.

25. Optimization step 1: PCR cycling
    a. 1 cycle
       i. 30 s, 98°C
    b. 5 cycles
       i.   10 s, 98°C
       ii. 30 s, 63°C
       iii. 60 s, 72°C
       iv. Hold at 4°C until optimization step 2 is complete.

26. Optimization step 2: qPCR to determine the optimal number of PCR cycles
    Perform a qPCR (5 μL NEBNext High-Fidelity 2× PCR Master Mix; 5 μL of amplified DNA from optimization step 1; 0.25 μL (25 μM) Nextera Forward PCR Primer; 0.25 μL (25 μM) Nextera Reverse PCR Primer; 0.09 μL 100 x SYBR Green I; 4.41 μL nuclease-free H$_2$O) with the following cycling program:
    a. 1 cycle
       i. 30 s, 98°C
    b. 20 cycles
       i.   10 s, 98°C
       ii. 30 s, 63°C
       iii. 60 s, 72°C

27. Calculate the total number of PCR cycles and compete step 25. To determine the optimal number of PCR cycles, plot the qPCR fluorescent signal, identify the value corresponding to 1/3 of the plateau fluorescent value (Figure 1, red line). The number of PCR cycles to complete the optimization step 25 above is the number of cycles that intersects the threshold (red line). If this value falls between two cycles, the lower cycle should be preferred.

28. Purify amplified libraries using a 1:1 volume Magbio HighPrep PCR beads (Auto Q Biosciences) as follows:
    a. Place Magbio HighPrep PCR beads at 18°C–20°C for 30 min
    b. Add nuclease-free water to the ligation reaction bringing the reaction volume to 100 μL.

    ⚠ CRITICAL: Ensure that the final volume is 100 μL before adding Magbio HighPrep PCR beads.
    c. Add 100 μL (1.0×) of resuspended Magbio HighPrep PCR beads and mix well using a vortex mixer or a pipette (pipet gently at least 10 times).
    d. Incubate for 5 min at 18°C–20°C.
    e. Spin tubes in a microcentrifuge and place on a magnetic rack to separate beads from the supernatant. After the solution clears (about 5 min), discard the supernatant that contains unwanted fragments.

    ⚠ CRITICAL: Do not discard the beads, which bind to the target fragments.
    f. Add 200 μL of freshly prepared 80% ethanol while the tubes are on the magnetic rack. Incubate at 18°C–20°C for 30 s.
    g. Carefully remove the supernatant without disturbing the beads and discard.
    h. Repeat washing step f.
    i. Spin briefly and place tubes onto the magnetic rack.
    j. Remove residual ETOH without disturbing the beads and air dry for 5 min.
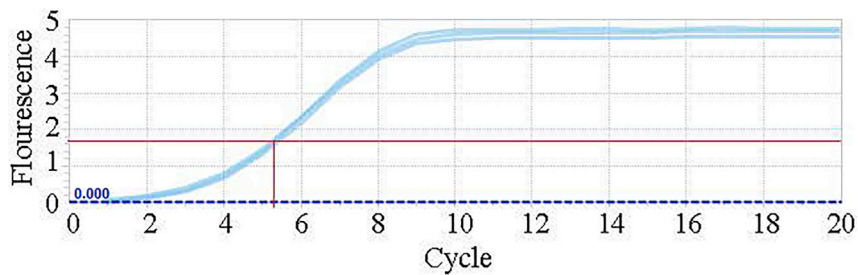
Figure 1. Library amplification optimization via qPCR

Fluorescence plot used in point 5 of the ATAC protocol to optimize DNA library amplification steps and prevent over or underamplification, which can causes biases in the DNA fragment representation within libraries. To determine the optimal number of PCR cycles, we plot the qPCR fluorescent signal, and identify the value corresponding to 1/3 of the plateau fluorescent value (red line). The number of PCR cycles to achieve an optimal amplification is the number of cycles that intersects the threshold (red line). If this value falls between two cycles, the lower cycle should be preferred. The vertical red line is the number of cycles that intersects the threshold.

⚠ CRITICAL: Do not exceed 5 min; over drying the beads may result in lower recovery of target DNA.

k. Remove tubes from the rack and elute DNA in 23 μL Buffer EB. Mix well on a vortex mixer or by pipetting, and incubate for 2 min at 18°C–20°C.
l. Place tubes onto the magnetic rack until the solution appear clear.
m. Transfer 20 μL of supernatant to a clean tube and discard the beads.

⏸⏸Pause point: the purified libraries can be stored at -20°C before the amplification step

Library amplification

⏱ Timing: 1.5 h

In this step, DNA libraries are amplified to generate sufficient material for next generation sequencing (NGS). This step should follow an optimization of the PCR conditions as explained in point 5.

29. Following DNA purification, amplify DNA libraries using NEBNext High-Fidelity PCR master mix and 25 μM Nextera PCR primers: 25 μL NEBNext High-Fidelity 2× PCR Master Mix; 10 μL transposed DNA; 2.5 μL 25 μM Forward PCR Primer; 2.5 μL 25 μM Reverse PCR Primer; 10 μL nuclease-free H_2O (Table 1), using the following PCR cycling:
    a. 1 cycle
        i. 5 min, 72°C
        ii. 30 s, 98°C
    b. 12 cycles
        i. 10 s, 98°C
        ii. 30 s, 63°C
        iii. 60 s, 72°C

⚠ CRITICAL: The optimal number of cycles for non-saturated libraries obtained from 7 last instar Daphnia individuals is 12. The number of cycles may change with the input material and should be optimized using step 29 above.

30. Purify amplified libraries using a 1:1 volume Magbio HighPrep PCR beads (Auto Q Biosciences) as follows:
    a. Place Magbio HighPrep PCR beads at 18°C–20°C for 30 min

b. Add nuclease-free water to the ligation reaction bringing the reaction volume to 100 μL.

⚠ CRITICAL: Ensure that the final volume is 100 μL before adding Magbio HighPrep PCR beads.

c. Add 100 μL (1.0×) of resuspended Magbio HighPrep PCR beads and mix well using a vortex mixer or a pipette (pipet gently at least 10 times).
d. Incubate for 5 min at 18°C–20°C.
e. Spin tubes in a microcentrifuge and place on a magnetic rack to separate beads from the supernatant. After the solution clears (about 5 min), discard the supernatant that contains unwanted fragments.

⚠ CRITICAL: Do not discard the beads, which bind to the target fragments.

f. Add 200 μL of freshly prepared 80% ethanol while the tubes are on the magnetic rack. Incubate at 18°C–20°C for 30 s.
g. Carefully remove the supernatant without disturbing the beads and discard.
h. Repeat washing step f.
i. Spin briefly and place tubes onto the magnetic rack.
j. Remove residual ETOH without disturbing the beads and air dry for 5 min.

⚠ CRITICAL: Do not exceed 5 min; over drying the beads and may result in lower recovery of target DNA.

k. Remove tubes from the rack and elute DNA in 23 μL Buffer EB. Mix well on a vortex mixer or by pipetting, and incubate for 2 min at 18°C–20°C.
l. Place tubes onto the magnetic rack until the solution appear clear.
m. Transfer 20 μL of supernatant to a clean tube and discard the beads.

⏸⏸Pause point: The purified libraries can be stored at −20°C before next generation sequencing.

Next generation sequencing

In this step, DNA libraries are prepared for next generation sequencing using Illumina or Illumina-compatible platforms.

31. Assess the quality and size of the ATAC libraries using a High Sensitivity D1000 Screentape (Agilent - 5067- 5584) and High Sensitivity D1000 Reagents (Agilent - 5067- 5585) on a Tapestation 2200 (Agilent, G2964AA).
32. Quantitate libraries on a Qubit 3.0 Fluorometer (Invitrogen, Q33216) using Qubit 1× dsDNA HS Assay Kit (Invitrogen, Q33231). Use the average library size from the Tapestation QC test to determine the final molar concentration. Other quantitation approaches can be used (e.g., qPCR). Here, we used Qubit combined with Tapestation traces because it allows a higher throughput and is more cost-effective. Information from the Tapestation traces allow to assess the fragment distribution within libraries. A bioanalyzer can be used in alternative to the Tapestation.

Note: The Tapestation traces appear smooth with a peak at 200–400 bp, depending on the insert size used in the library preparation (see below). The same libraries run on a bioanalyzer tend to show a more laddered profile (data not shown). The profile can vary among samples with some samples showing a wider peak; however, these difference do not affect the quality of the libraries.

33. Prepare equimolar pools of ATAC libraries ensuring that different sets of unique paired indexes are used for different libraries.

34. Assess the quality and size of the ATAC library pool on Tapestation 2200 (Agilent, G2964AA) using High Sensitivity D1000 Screentape (Agilent - 5067- 5584) and High Sensitivity D1000 Reagents (Agilent - 5067- 5585).

35. Quantitate ATAC library pool using Qubit 3.0 Fluorometer (Invitrogen, Q33216) and Qubit 1× dsDNA HS Assay Kit (Invitrogen, Q33231).

36. Optional step. If necessary, pools can be concentrated using a 1:1 HighPrep bead purification described above (step 32).

⸏⸏Pause point: The pooled libraries can be stored at −20°C until further processing.

37. Run prepared 2 × 250 bp ATAC libraries on an Illumina or other compatible platform to generate 40 million reads per sample. A HiSeqXten was used here.

   Note: A test run on an Illumina MiSeq platform is advised to validate the protocol before processing large sample numbers. The libraries constructed using different concentrations of the transposase enzyme (see Expected outcomes) were sequenced at low coverage. The normalized ATAC-seq traces from these low coverage libraries were compared with a high coverage sample (40 million reads) at four scaffolds (see Expected outcomes).

Bioinformatics analysis: ATAC-seq standards and processing pipeline

This step describes the processing pipeline used to perform quality checks on the ATAC sequences and to generate summary metrics, for which we followed the ENCODE guidelines (https://www.encodeproject.org/atac-seq/), where possible.

   Note: Metrics that require a chromosomal-level assembly of the reference genome could not be calculated as the reference genome of D. magna 2.4 is not resolved at chromosomal level.

38. Perform an initial sequence quality check using fastqc (V0.11.9) for each fastq files (Andrews, 2010).

39. Use Trimmomatic (v0.32) (Bolger et al., 2014) to remove Illumina adapters and low quality reads, as well as reads shorter than 50 bp using the following parameters: "ILLUMINACLIP: List_of_Adapaters.fa:2:30:10 LEADING:30 TRAILING:30 MINLEN:50 HEADCROP:10."

40. Validate libraries, estimate sample replicate bias and quantify read quality following the gold standard of ENCODE (https://www.encodeproject.org/atac-seq), where possible, and the ATAC-seq data analysis recommendations by galaxy projects (Batut et al., 2018) using multiqc (Ewels et al., 2016) and ataqv (Orchard et al., 2020).

41. After indexing your reference genome, use Bowtie2 (v2.2.6) (Langmead et al., 2019) ultra-fast aligner to map reads that passed the quality threshold to the reference genome using very sensitive search mode settings as follows: "-D 20 -R 3 -N 0 -L 20 -i S,1,0.50" .

42. Apply Samtools (V1.4) (Li et al., 2009) to convert Bowtie2 mapped SAM files to compressed binary format BAM; sort and index for further analysis.

43. Remove PCR duplicates from the sorted BAM files using the Broad Institute Picard tools (2.10.5) (http://broadinstitute.github.io/picard/).

44. Filter out r that map onto the mitochondrial DNA (NCBI: NC_026914.1) before calling ATAC peaks.

45. Quantify mapping quality statistics and visualize results using Qualimap 2 (Okonechnikov et al., 2016) (http://qualimap.conesalab.org/).

46. Shift aligned reads "4 , -5" and "5, -4" for positive and negative strand reads, respectively, using deeptools alignmentSieve (Ramirez et al., 2018).

47. Use MACS2 callpeak (Gaspar, 2018) to call the peaks from each replicate BAM file and sample; QC matrices are calculated on unmerged technical or biological replicates. For some downstream analyses (e.g., comparative peak analysis) replicates can be merged.

48. Filter, combine (where relevant) and convert peaks for visualization using bigwig in bedtools (Quinlan, 2014) and/or custom R scripts.
49. Quantify replicates concordance by calculating the Irreproducible Discovery Rate (IDR), and following the ENCODE processing pipeline for replicated or unreplicated samples as necessary (https://github.com/kundajelab/idr).
50. Visualize high-quality genome browser tracks on the focal species reference genome using Deeptools pyGenomeTracks (Ramirez et al., 2018).
51. Identify and extract Transcription Start Site (TSS) from the reference gene annotation file into deeptool compatible bed format. Scale peak traces across samples using deeptools computed matrix and visualize TSS heatmaps using plotHeatmap (Ramirez et al., 2018).
52. Search for enriched known motifs at transcription start sites (TSS) using HOMER discovery algorithm for regulatory elements (Hypergeometric Optimization of Motif EnRichment; Duttke et al., 2019). A search window spanning 200 bp upstream from the peak center with up to 12 bp motif length can be used following (Heinz et al., 2010).

EXPECTED OUTCOMES

Expected outcomes at critical steps of the protocol and optimization steps are described.

Optimizing DNA input for ATAC-seq in *Daphnia* (related to step 1)

The amount of starting material for ATAC-seq libraries was a critical step in our protocol optimization. As compared to mammalian or vertebrate models, the crustacean *Daphnia* has limited starting material. The whole adult organism measures less than 4 mm. *Daphnia* provides many of the felicities of invertebrate model species: they are easy and inexpensive to culture in the laboratory, have a short life cycle, and produce large numbers of externally laid embryos. *Daphnia* has a parthenogenetic life cycle, in which sexual and asexual reproduction alternate (Ebert, 2005). Sexual recombination results in early stage embryos that arrest their development and enter dormancy (Kerfoot and Weider, 2004). Dormant embryos can be revived and propagated clonally under standard laboratory conditions, allowing the rearing of populations of isogenic individuals (clones) from a single genotype (Cambronero Cuenca and Orsini, 2018). To determine the optimal amount of tissue for high-quality ATAC libraries, we took advantage of *Daphnia* clonality. We performed a titration experiment using 7, 15, 30 and 60 last instar female *Daphnia* juveniles from the same genotype (Figure 2). The quality of ATAC libraries was assessed for: i) animal tissue manually homogenized using a pestle; ii) tissue treated with a tissue homogenizer (Geno/Grinder); and iii) tissue treated with automated tissue homogenizer followed by DNA shearing A successful ATAC library is expected to result in a clear peak between 200–300 bp. A larger size peak can be visible, representing unfragmented genomic DNA (>1,000 bp).

Libraries obtained from tissue treated with an automated tissue homogenizer (Geno/Grinder) were more replicable in quality and quantity of material obtained than libraries obtained from tissue homogenized manually (Figure 2). The combined use of a tissue homogenizer and a picosonicator did not visibly improve the quality of the libraries when 7 and 15 individual *Daphnia* were used (Figure 2), whereas it improved the quality of the ATAC libraries when more than 30 individuals were used (Figure 2). Even when tissue homogenizer combined with DNA shearing was used, the ATAC libraries were of lower quality when more than 30 individuals were used as compared to <15 individual, likely because of suboptimal homogenization of tissue and suboptimal access to free nuclei (Figure 2; 60 individuals).

Optimizing transposase concentration (related to step 3)

The prohibitive costs of the enzyme transposase impose serious limitations to the application of open chromatin analysis in population epigenomics. We perform a titration experiment to determine the minimal amount of transposase enzyme needed to generate good quality ATAC-seq libraries. Dilutions corresponding to 100%, 50%, 25%, and 12% of the manufacturer suggested concentrations were used. These dilutions were tested on libraries prepared with a tissue homogenizer
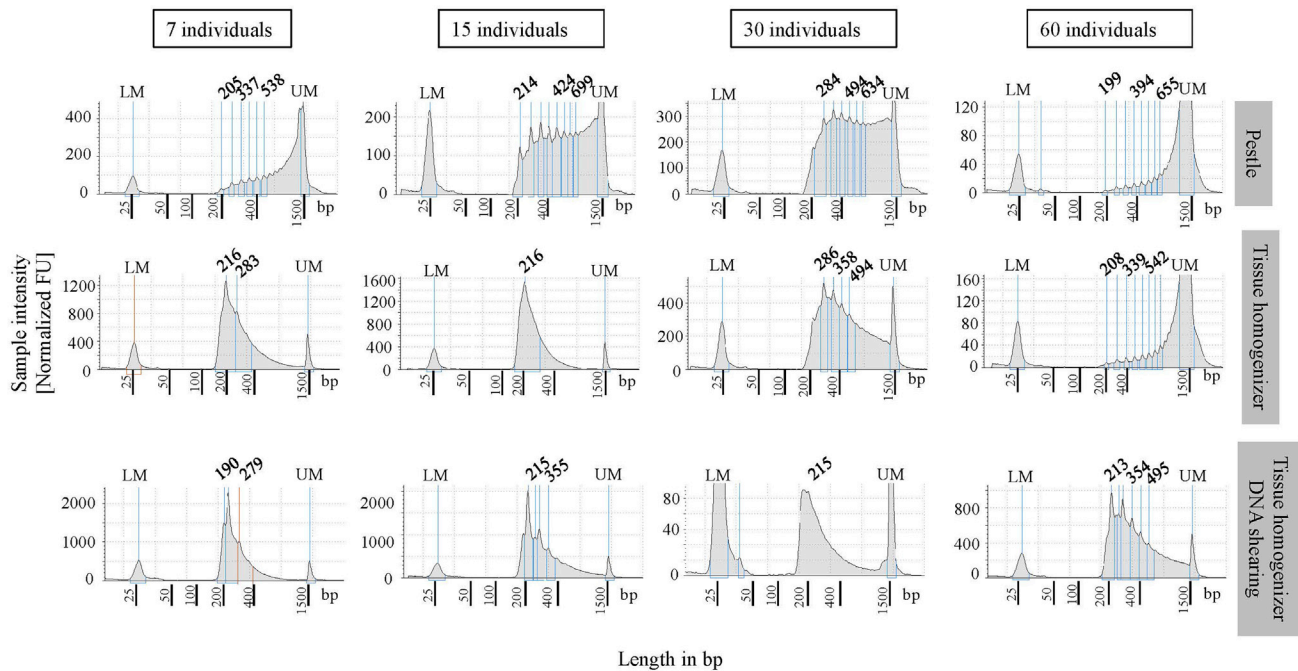
Figure 2. Optimizing DNA input for ATAC-seq

A titration experiment was completed to identify the optimal input material (tissue) to generate a robust ATAC-seq library. The results of this titration experiment are shown as tape station traces of ATAC-seq libraries obtained from different starting material: 7, 15, 30, and 60 last instar *Daphnia* female juveniles, from a single genotype. These different starting materials were tested with three tissue homogenization techniques: manual homogenization with a pestle, automated tissue homogenizer (Geno/Grinder) and a combination of automated tissue homogenizer (Geno/Grinder) and DNA sharing using a pico-sonicator (Bioruptor). The scale on the y axis may differ. The scale on the x axis shows size in base pairs (bp).

step and a 16–20 h (overnight) transposase reaction (Figures 3A–3D) as well as on libraries prepared using a tissue homogenizer and a picosonicator (Figures 3E–3H). There were no appreciable differences among 100%, 50%, and 25% dilutions and between the two methods used (Figure 3). The amount of libraries obtained was also comparable (Figure 3). However, whereas the 12% dilution did not produce an acceptable quality library with transposase alone, it produce an acceptable quality library in the transposase combined with a picosonicator step (Figures 3D and 3H). For the 100% and 25% dilutions, the transposase incubation time was reduced to 30 min as per manufacturer guidelines (Figures 3A1 and 3C1). The quality of these libraries was comparable to the ones obtained with16–20 h (overnight) incubation (Figures 3A–3D), even if they resulted in a comparatively lower yield.

Validation of ATAC-seq modified protocol (related to step 7)

ATAC-seq libraries obtained from the titration experiments in point b (Figure 3) were sequenced at low coverage to assess whether they produced qualitatively comparable results, namely replicable peak traces. The same genotype sequenced at high depth of coverage (40 million reads) was used as reference in this analysis (Figure 4, R). We then compared chromatin profiles in the reference and in the low coverage samples. ATAC fragments were consistently observed across the low coverage and the reference samples, indicating that the tested range of DNA input and the dilutions of transposase tested did not significantly alter the quality of ATAC libraries and ATAC-seq peak traces (Figure 4). We did not do any quantitative analyses on these data because of the low depth of sequencing.

To validate the modified ATAC-seq protocol for downstream applications, we constructed ATAC-seq libraries using 25% dilution of transposase and 7 last instar individuals from 3 *D. magna* genotypes. Taking advantage of the properties of *Daphnia*, namely the opportunity to generate clonal copies of each genotype, we sequenced 6 ATAC-seq libraries with high depth of coverage
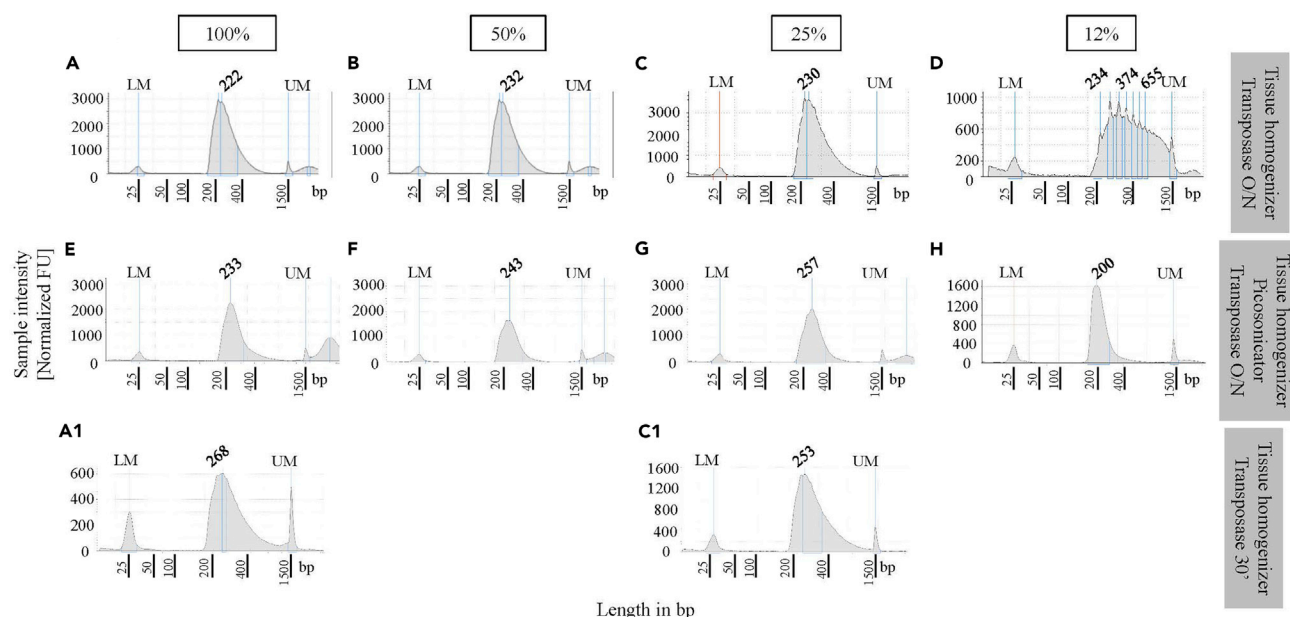
Figure 3. Optimizing transposase dilution in ATAC-seq libraries
Population epigenomics is largely unexplored due to the prohibitive costs of the transposase enzyme. Here we show the outcome of a titration experiment for different dilution of the transposase enzyme suggested by the manufacturer. ATAC-seq tape station traces are shown for libraries obtained with different dilutions of transposase: 100%, 50%, 25%, and 12% of the manufacturer suggested concentration. (A–D) ATAC libraries obtained with an over/night (O/N) incubation of the transposition reaction preceded by an automated tissue homogenization step; (E–H) libraries obtained with an over/night (O/N) incubation of the transposition reaction preceded by an automated tissue homogenization step combined with pico-sonication (DNA shearing); A1 and C1: ATAC libraries obtained for 100% and 25% dilutions of the transposase with an incubation of the transposition reaction of 30 min.

(40 million reads per sample) from an exposure experiment in which clonal replicates of three geno-types (LRV2_1; LRV3.5_15 and LRV13_3) were either maintained in control conditions (not exposed), or exposed to 8 μg/L of a carbamate insecticide. This insecticide has been previously shown to have an adverse effect on *Daphnia* fitness within a single generation (Cambronero et al., 2018; Jansen et al., 2011). The six libraries were first used to generate data quality metrics, and then for peak and motif analyses.

Metrics. We assessed the quality of our ATAC sequences using the golden standard metrics of the ENCODE project (https://www.encodeproject.org/atac-seq/) and the ATAC-seq data analysis rec-ommendations by galaxy projects, where possible. ENCODE metrics that require a chromosomal-level genome assembly could not be used. The metrics calculated following the ENCODE standard are shown in the following. Additional sequence metrics are in Table 2. All metrics support that the modified ATAC-seq protocol optimized here can be used for population epigenomics of invertebrates.

| ENCODE standard	Current paper
Experiments should have two or more biological replicates or at least two technical replicates	Three biological replicated and two technical replicates per samples
Each replicate should have 25 million on-duplicate, non-mitochondrial aligned reads	Reads per replica ranged between 14 and 50 million reads (Average read per replica: 26 million reads)
Percentage of mapped reads should be greater than 95%, though values >80% may be acceptable	Percentage of mapped reads ranged between 83.2% and 93.8%
Irreproducible Discovery Rate (IDR) should be less than 2	IDR <1.6
The number of peaks within a replicated peak file should be >150,000, though values >100,000 may be acceptable	>50,000 (corrected for the genome size of *D. magna*, 130 Mb; this value is over the quality threshold indicated)
The number of peaks within an IDR peak file should be >70,000, though values >50,000 may be acceptable	> 11,700 (corrected for the genome size of *D. magna*, 130 Mb; this value is over the quality threshold indicated)
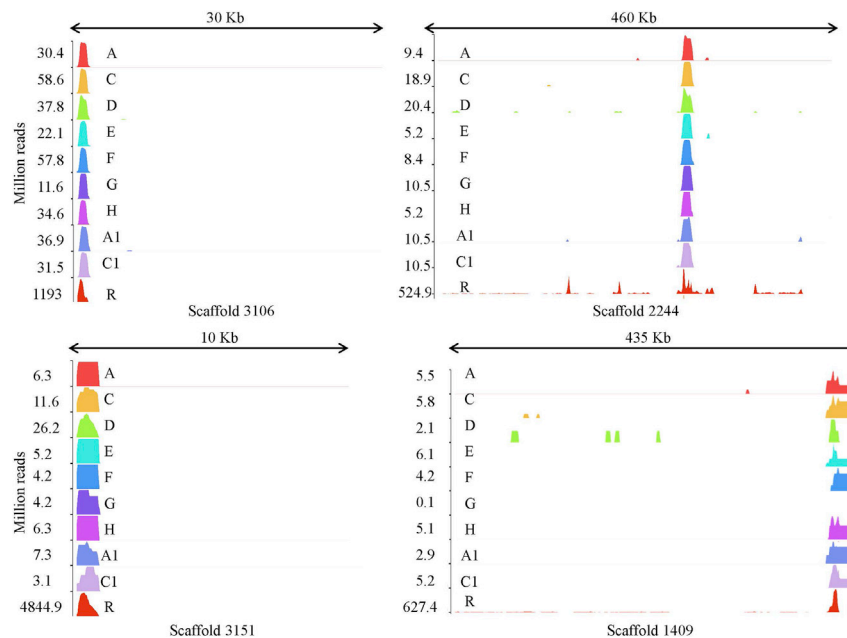
Figure 4. Normalized ATAC-seq traces for four scaffolds on the *Daphnia magna* genome
Normalized ATAC-seq traces shown for 4 scaffolds on the *Daphnia magna* genome (v2.4) obtained from the ATAC-seq libraries in Figure 3. These traces are obtained from low coverage libraries and compared to a high coverage library obtained from the same *Daphnia* genotype R, used as reference.

Peak and motif analysis. We used the six high coverage samples to identify divergence in chromatin accessibility driven by exposure to a carbamate insecticide (carbaryl). To visualize changes in chromatin signatures, we compare the normalized ATAC-seq traces of exposed and unexposed genotypes (Figure 5). As *Daphnia* has a parthenogenetic life cycle, genotypes are genetically distinct but can be propagated clonally, therefore clonal copies of the same genotype can be exposed to different environments. Genome-wide counts of chromatin motifs were quantified in exposed and non-exposed genotypes and summarized in a principal component analysis (Figure 6). Treated and control animals were clearly distinguishable in the PCA plot (Figure 6A). Treatment-specific

Table 2. Sequence metrics

Sample name	No. of raw paired end reads	No. of QC paired end reads	Insert size median	Coverage mean	Mapped reads %	No. of reads mapped to the mitochondria	No. of peaks
LRV13_3R1_C	18,635,973 x 2	16,135,365 x 2	151	24,5574	91.33%	17,974	61,285
LRV13_3-R2C	29,340,610 x 2	25,005,090 x 2	153	37,6782	91.84%	19,888	65,281
LRV13_3R1_I	16,668,003 x 2	12,773,098 x 2	131	16,3577	89.55%	24,572	59,735
LRV13_3R2_I	13,709,892 x 2	10,100,948 x 2	123	12,4835	88.91%	21,810	60,661
LRV2_1R1_C	12,140,926 x 2	10,137,006 x 2	197	16,1945	91.36%	3,280	28,993
LRV2_1R2_C	14,302,926 x 2	11,238,267 x 2	168	15,8192	90.47%	7,696	43,685
LRV2_1R1_I	12,388,306 x 2	7,082,046 x 2	111	7,4666	83.23%	21,434	61,512
LRV2_1R2_I	13,813,597 x 2	9,790,744 x 2	129	11,1734	83.20%	21,059	58,803
LRV3_5_15R1_C	18,483,584 x 2	13,824,157 x 2	155	18,9622	93.69%	8,435	51,683
LRV3_5_15R2_C	25,080,326 x 2	20,644,243 x 2	174	30,8094	93.78%	7,970	48,057
LRV3_5_15R1_I	14,942,965 x 2	10,405,452 x 2	129	13,1903	92.63%	20,251	55,355
LRV3_5_15R2_I	13,019,186 x 2	8,165,030 x 2	129	10,0822	92.68%	17,601	53,166

QC metrics for NGS data generated from the high coverage samples (LRV2_1; LRV3.5_15 and LRV13_3) obtained from an exposure experiment in which clonal replicates were either exposed to the insecticide Carabaryl (8μg/L) or maintained in non-stress (control) conditions. C = control; I = insecticide carbaryl; R1 = technical replicate 1; R2 = technical replicate 2.
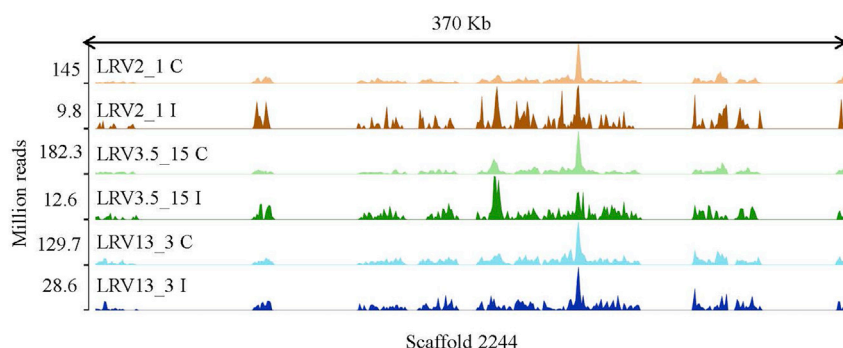
Figure 5. Normalized ATAC-seq traces of three *Daphnia* genotypes exposed to the insecticide carbaryl
Normalized ATAC traces are shown for three genotypes of *D. magna* (LRV2_1; LRV3.5_15 and LRV13_3) for a representative scaffold of ca. 400 kb (Scaffold 2244). For each genotype a clonal replicate exposed to 8 μg/L of the insecticide carbaryl (I) and a non-exposed clonal replicate (C) are shown. These traces are obtained from high coverage libraries (40 million reads).

ATAC traces were identified in a Venn diagram including all genotypes (control and treatment) (Figure 6B).

A heatmap of the genome-wide TSS motifs, scaled using deeptools and visualized with plotHeatmap, show a distinct pattern for exposed and non-exposed copies of the same genotypes, identifying a clear regulatory response to the experimental perturbation imposed by the insecticide carbaryl (Figure 7). A list of known motifs searched against the HOMER discovery algorithm for regulatory elements (Hypergeometric Optimization of Motif EnRichment; (Duttke et al., 2019)) are listed in Table 3 for control and exposed samples (LRV2_1; LRV3.5_15 and LRV13_3).

Overall, our results show that the modified ATAC protocol presented here is compatible with applications in population epigenomics of *D. magna*, with possible application in other non-model invertebrates.

LIMITATIONS

Many steps in the optimized protocol are temperature sensitive (e.g., enzymatic reactions). If incubation temperatures are not respected and if room temperature exceeds 30°C, the protocol may be
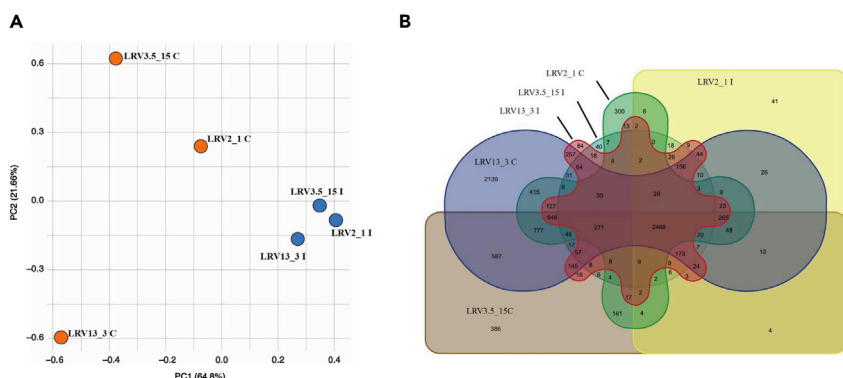
Figure 6. Chromatin accessibility patterns in *Daphnia magna*
(A and B) (A) Principal component analysis of genome-wide accessibility variation at three *D. magna* genotypes (LRV2_1; LRV3.5_15 and LRV13_3). Clonal replicates of the three genotypes exposed to the insecticide carbaryl (I; 8 μg/L) as well as non-exposed clonal replicates (C) are shown; (B) Venn diagram showing unique ATAC normalized traces identified in *D. magna* genotypes exposed to the insecticide carbaryl (I), as well as traces shared between treated (I) and untreated (C) genotypes.
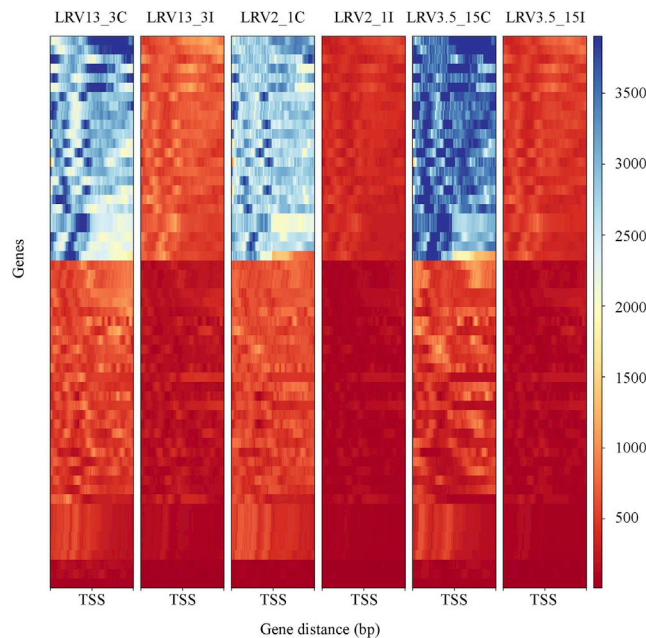
Figure 7. Heatmaps of transcription start sites (TSS)
Distribution of genome-wide DNA motifs shared among three *Daphnia magna* genotypes exposed to the insecticide carbaryl (8 μg/L) (LRV3.5_15I, LRV2_1I and LRV13_3I) and maintained in non-stressful conditions (LRV3.5_15C, LRV2_1C, and LRV13_3C). Peak traces across samples were scaled using deeptools and visualized with plotHeatmap. TSS were scaled to include 1 kb up and downstream of the TSS. Similarity decreases from red to blue.

unreliable. Furthermore, tissue storage conditions may be a cause of failure of the protocol due to degradation of the input material. Animal tissue should be used fresh or flash frozen immediately after collection for later use. Improperly stored tissue may lead to unreliable outputs. The use of DNA exacted from human cell lines is advised as positive control to assess reliability of critical steps in the protocol.

The protocol has been optimized for the crustacean *Daphnia magna*. Although transferability to other crustaceans is largely expected, it is advised that transposase enzyme, staring material and quality of ATAC libraries is assessed for other organisms before the essay is applied to large sample numbers.

TROUBLESHOOTING
Problem 1: inefficient DNA fragmentation (related to step 2)
The transposition reaction efficiency may be diminished or have reduced replicability because of inefficient DNA fragmentation. This may occur in species with very compact genomes.

Potential solution: mechanic fragmentation
A mechanic DNA fragmentation step (e.g., pico-sonication) may be introduced prior to the transposition reaction to improve evenness of fragmentation and accessibility of the tagmentase, enhancing replicability and efficiency of the transposition reaction. An example of how mechanic fragmentation may improve DNA shearing is in Figure 2. Using a tissue homogenizer improves the quality of the ATAC-seq libraries; the combination of a tissue homogenizer and DNA shearing using a picosonicator improves the quality of the libraries when excess tissue may limit accessibility to the nuclei.

Problem 2: low transposition efficiency (related to step 3)
Excess transposition or low efficiency of the transposition reaction may affect the identification of chromatin motifs.

Table 3. Known motifs enrichment results

Motif name	Consensus	Log p value	q value (Benjamini)	# of target sequences with motif (of 10,325)	% of target sequences with motif	# of background sequences with motif (of 32,685)	% of background sequences with motif
Control
NFIL3(bZIP)/HepG2-NFIL3-ChIP-Seq(Encode)/Homer	VTTACGTAAYNNNNN	−7.14E+02	0	2,868	27.77%	4,448.9	13.61%
Foxh1(Forkhead)/hESC-FOXH1-ChIP-Seq(GSE29422)/Homer	NNTGTGGATTSS	−1.28E+03	0	3,435	33.27%	4,471.3	13.68%
AT2G20400(G2like)/colamp-AT2G20400-DAP-Seq(GSE60143)/Homer	DNVGAATATTCBNHN	−6.69E+02	0	1,164	11.27%	1,050.5	3.21%
AT1G76880(Trihelix)/col-AT1G76880-DAP-Seq(GSE60143)/Homer	ACGGTAAAAW	−6.13E+02	0	2,177	21.08%	3,130.6	9.58%
At2g03500(G2like)/col-At2g03500-DAP-Seq(GSE60143)/Homer	WWAGAATATTCT	−5.68E+02	0	1,315	12.73%	1,457.9	4.46%
At5g29000(G2like)/col-At5g29000-DAP-Seq(GSE60143)/Homer	RGAATATTCYHH	−5.35E+02	0	1370	13.27%	1,618.3	4.95%
JGL(C2H2)/col-JGL-DAP-Seq(GSE60143)/Homer	ACYTTCAGTT	−4.94E+02	0	3,956	38.31%	7,971.9	24.39%
Foxa3(Forkhead)/Liver-Foxa3-ChIP-Seq(GSE77670)/Homer	BSNTGTTTACWYWGN	−3.54E+02	0	1,701	16.47%	2,746.3	8.40%
FOXP1(Forkhead)/H9-FOXP1-ChIP-Seq(GSE31006)/Homer	NYYTGTTTACHN	−3.12E+02	0	2,146	20.78%	3,964.4	12.13%
At3g04030(G2like)/col-At3g04030-DAP-Seq(GSE60143)/Homer	DRGAATCT	−2.84E+02	0	1,903	18.43%	3,465.8	10.60%
At1g25550(G2like)/colamp-At1g25550-DAP-Seq(GSE60143)/Homer	NAGATTCY	−2.79E+02	0	1,921	18.60%	3,530.5	10.80%
FOXK2(Forkhead)/U2OS-FOXK2-ChIP-Seq(E-MTAB-2204)/Homer	SCHTGTTTACAT	−2.58E+02	0	2,290	22.18%	4,558.8	13.95%
FUS3(ABI3VP1)/col-FUS3-DAP-Seq(GSE60143)/Homer	DNNWTNTGCATGKNN	−2.36E+02	0	1,323	12.81%	2,240.2	6.85%
At1g68670(G2like)/colamp-At1g68670-DAP-Seq(GSE60143)/Homer	WNWWHNRAAGATTCT	−2.30E+02	0	1,419	13.74%	2,489.4	7.62%
Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq(Fan_et_al.)/Homer	CTGTTTAC	−2.13E+02	0	3,355	32.49%	7,721.5	23.63%
GTL1(Trihelix)/colamp-GTL1-DAP-Seq(GSE60143)/Homer	WWTTTACCKY	−1.89E+02	0	3,023	29.28%	6,935.8	21.22%
Foxo3(Forkhead)/U2OS-Foxo3-ChIP-Seq(E-MTAB-2701)/Homer	DGTAAACA	−1.86E+02	0	2,570	24.89%	5,689.8	17.41%
FoxD3(forkhead)/ZebrafishEmbryo-Foxd3.biotin-ChIP-seq(GSE106676)/Homer	TGTTTAYTTAGC	−1.81E+02	0	2,115	20.48%	4,483.8	13.72%
OCT:OCT-short(POU,Homeobox)/NPC-OCT6-ChIP-Seq(GSE43916)/Homer	ATGCATWATGCATRW	−1.65E+02	0	2,091	20.25%	4,511	13.80%
At5g04390(C2H2)/col200-At5g04390-DAP-Seq(GSE60143)/Homer	AGTGANDN	−1.55E+02	0	6473	62.69%	17,722.3	54.23%
Foxa2(Forkhead)/Liver-Foxa2-ChIP-Seq(GSE25694)/Homer	CYTGTTTACWYW	−1.48E+02	0	2,418	23.42%	5,514.6	16.87%
At3g24120(G2like)/col-At3g24120-DAP-Seq(GSE60143)/Homer	NWWAGMATMW	−1.21E+02	0	5,036	48.77%	13,498.4	41.30%
AT5G60130(ABI3VP1)/col-AT5G60130-DAP-Seq(GSE60143)/Homer	WTTYTAAGVAAA	−1.20E+02	0	3,896	37.73%	10,024.7	30.67%
At2g01060(G2like)/colamp-At2g01060-DAP-Seq(GSE60143)/Homer	AGATKCBNWW	−1.17E+02	0	5,572	53.96%	15,213.2	46.55%

Table 3. *Continued*

Motif name	Consensus	Log p value	q value (Benjamini)	# of target sequences with motif (of 10,325)	% of target sequences with motif	# of background sequences with motif (of 32,685)	% of background sequences with motif
RBFox2(?)/Heart-RBFox2-CLIP-Seq(GSE57926)/Homer	TGCATGCA	−1.10E+02	0	2,618	25.35%	6,366.5	19.48%
AT5G47660(Trihelix)/colamp-AT5G47660-DAP-Seq(GSE60143)/Homer	AWTTTTACCG	−1.10E+02	0	3,348	32.42%	8,493.7	25.99%
Fox:Ebox(Forkhead,bHLH)/Panc1-Foxa2-ChIP-Seq(GSE47459)/Homer	NNNVCTGWGYAAACASN	−9.71E+01	0	1,766	17.10%	4,066.8	12.44%
STZ(C2H2)/colamp-STZ-DAP-Seq(GSE60143)/Homer	HNBTCACT	−9.54E+01	0	6,447	62.43%	18,252.9	55.85%
At3g12730(G2like)/colamp-At3g12730-DAP-Seq(GSE60143)/Homer	AAGATTCT	−8.99E+01	0	2,458	23.80%	6,079.2	18.60%
AT5G45580(G2like)/colamp-AT5G45580-DAP-Seq(GSE60143)/Homer	ADRGAATCTH	−7.41E+01	0	3,038	29.42%	7,938.1	24.29%
GT2(Trihelix)/colamp-GT2-DAP-Seq(GSE60143)/Homer	AMGGTAAAWWWN	−7.10E+01	0	2,944	28.51%	7694.4	23.54%
FOXK1(Forkhead)/HEK293-FOXK1-ChIP-Seq(GSE51673)/Homer	NVWTGTTTAC	−6.86E+01	0	2,808	27.19%	7,319.6	22.40%
FOXM1(Forkhead)/MCF7-FOXM1-ChIP-Seq(GSE72977)/Homer	TRTTTACTTW	−6.43E+01	0	2,428	23.51%	6,249.7	19.12%
AT2G38300(G2like)/col-AT2G38300-DAP-Seq(GSE60143)/Homer	ADRGAATGTT	−6.01E+01	0	2,429	23.52%	6,299.9	19.28%
FoxL2(Forkhead)/Ovary-FoxL2-ChIP-Seq(GSE60858)/Homer	WWTRTAAACAVG	−4.43E+01	0	2,395	23.19%	6,396.6	19.57%
AP-2gamma(AP2)/MCF7-TFAP2C-ChIP-Seq(GSE21234)/Homer	SCCTSAGGSCAW	−4.38E+01	0	205	1.99%	322.3	0.99%
AP-2alpha(AP2)/Hela-AP2alpha-ChIP-Seq(GSE31477)/Homer	ATGCCCTGAGGC	−4.31E+01	0	150	1.45%	207.7	0.64%
KAN2(G2like)/colamp-KAN2-DAP-Seq(GSE60143)/Homer	ATATTCTY	−4.10E+01	0	2,289	22.17%	6,126.4	18.75%
FOXA1(Forkhead)/MCF7-FOXA1-ChIP-Seq(GSE26831)/Homer	WAAGTAAACA	−2.48E+01	0	2,438	23.61%	6,834.6	20.91%
Foxf1(Forkhead)/Lung-Foxf1-ChIP-Seq(GSE77951)/Homer	WWATRTAAACAN	−1.78E+01	0	2,621	25.38%	7,539.7	23.07%
KANADI1(Myb)/Seedling-KAN1-ChIP-Seq(GSE48081)/Homer	ARGAATAWWN	−1.77E+01	0	2,421	23.45%	6,932.4	21.21%
EBF(EBF)/proBcell-EBF-ChIP-Seq(GSE21978)/Homer	DGTCCCYRGGGA	−1.45E+01	0	29	0.28%	31.1	0.10%
EBF2(EBF)/BrownAdipose-EBF2-ChIP-Seq(GSE97114)/Homer	NABTCCCWDGGGAVH	−1.41E+01	0	178	1.72%	384.3	1.18%
Dorsal(RHD)/Embryo-dl-ChIP-Seq(GSE65441)/Homer	GGGAAAAMCCCG	−1.39E+01	0	130	1.26%	262.8	0.80%
MYB73(MYB)/col-MYB73-DAP-Seq(GSE60143)/Homer	NNNNHAACNGHHDHN	−1.38E+01	0	4,006	38.80%	11,938.3	36.53%
E2F4(E2F)/K562-E2F4-ChIP-Seq(GSE31477)/Homer	GGCGGGAAAH	−1.35E+01	0	421	4.08%	1,055	3.23%
Tlx?(NR)/NPC-H3K4me1-ChIP-Seq(GSE16256)/Homer	CTGGCAGSCTGCCA	−1.23E+01	0.0001	117	1.13%	238.7	0.73%
HDG7(HB)/col-HDG7-DAP-Seq(GSE60143)/Homer	WGCATTTAATGC	−1.20E+01	0.0001	509	4.93%	1,324.4	4.05%
AT2G40260(G2like)/colamp-AT2G40260-DAP-Seq(GSE60143)/Homer	WAAAYATTCTTT	−1.19E+01	0.0001	2,664	25.80%	7,829.9	23.96%
LBD2(LOBAS2)/colamp-LBD2-DAP-Seq(GSE60143)/Homer	TCCGAWTTTTCGGN	−1.14E+01	0.0002	547	5.30%	1,442.8	4.41%
RKD2(RWPRK)/colamp-RKD2-DAP-Seq(GSE60143)/Homer	GACKTTTCRDCTTCC	−1.07E+01	0.0004	741	7.18%	2,020.9	6.18%

Table 3. Continued

Motif name	Consensus	Log p value	q value (Benjamini)	# of target sequences with motif (of 10,325)	% of target sequences with motif	# of background sequences with motif (of 32,685)	% of background sequences with motif
ZNF528(Zf)/HEK293-ZNF528.GFP-ChIP-Seq(GSE58341)/Homer	AGAAATGACTTCCCT	−9.87E+00	0.001	6	0.06%	2.4	0.01%
ZNF143|STAF(Zf)/CUTLL-ZNF143-ChIP-Seq(GSE29600)/Homer	ATTTCCCAGVAKSCY	−9.78E+00	0.0011	77	0.75%	151.6	0.46%
LRF(Zf)/Erythroblasts-ZBTB7A-ChIP-Seq(GSE74977)/Homer	AAGACCCYYN	−8.93E+00	0.0025	317	3.07%	813	2.49%
FOXA1(Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	WAAGTAAACA	−7.56E+00	0.0095	2,801	27.13%	8,400.2	25.70%
Pax8(Paired,Homeobox)/Thyroid-Pax8-ChIP-Seq(GSE26938)/Homer	GTCATGCHTGRCTGS	−6.85E+00	0.019	98	0.95%	223.7	0.68%
THRa(NR)/C17.2-THRa-ChIP-Seq(GSE38347)/Homer	GGTCANYTGAGGWCA	−6.79E+00	0.0199	106	1.03%	246	0.75%
EBF1(EBF)/Near-E2A-ChIP-Seq(GSE21512)/Homer	GTCCCCWGGGGA	−6.50E+00	0.0261	122	1.18%	291.4	0.89%
LOB(LOBAS2)/col-LOB-DAP-Seq(GSE60143)/Homer	CGCCGKAWWTTHCGS	−6.18E+00	0.0353	188	1.82%	479.4	1.47%
At3g60580(C2H2)/col-At3g60580-DAP-Seq(GSE60143)/Homer	WTTYTACT	−6.16E+00	0.0354	5,763	55.81%	17,780	54.40%
GEI-11(Myb?)/cElegans-L4-GEI11-ChIP-Seq(modEncode)/Homer	CCGACAYYTYACGGG	−6.02E+00	0.0402	47	0.46%	95.1	0.29%

Motif Name	Consensus	Log p value	q value (Benjamini)	# of target sequences with motif (of 6,508)	% of target sequences with motif	# of background sequences with motif (of 22,227)	% of background sequences with motif
Exposed to carbaryl							
Foxh1(Forkhead)/hESC-FOXH1-ChIP-Seq(GSE29422)/Homer	NNTGTGGATTSS	−1.58E+03	0	3,274	50.31%	4,276.4	19.24%
NFIL3(bZIP)/HepG2-NFIL3-ChIP-Seq(Encode)/Homer	VTTACGTAAYNNNNN	−1.16E+03	0	2,668	41.00%	3,562.3	16.03%
AT2G20400(G2like)/colamp-AT2G20400-DAP-Seq(GSE60143)/Homer	DNVGAATATTCBNHN	−8.36E+02	0	1,138	17.49%	923.1	4.15%
At5g29000(G2like)/col-At5g29000-DAP-Seq(GSE60143)/Homer	RGAATATTCYHH	−7.31E+02	0	1,316	20.22%	1,361	6.12%
At2g03500(G2like)/col-At2g03500-DAP-Seq(GSE60143)/Homer	WWAGAATATTCT	−7.29E+02	0	1,264	19.42%	1,267.8	5.71%
AT1G76880(Trihelix)/col-AT1G76880-DAP-Seq(GSE60143)/Homer	ACGGTAAAAW	−6.95E+02	0	2,061	31.67%	3,039.7	13.68%
JGL(C2H2)/col-JGL-DAP-Seq(GSE60143)/Homer	ACYTTCAGTT	−6.35E+02	0	3,297	50.66%	6,576.1	29.59%
At3g04030(G2like)/col-At3g04030-DAP-Seq(GSE60143)/Homer	DRGAATCT	−4.18E+02	0	1,548	23.79%	2,473.8	11.13%
At1g25550(G2like)/colamp-At1g25550-DAP-Seq(GSE60143)/Homer	NAGATTCY	−4.04E+02	0	1,546	23.76%	2,505	11.27%
FUS3(ABI3VP1)/col-FUS3-DAP-Seq(GSE60143)/Homer	DNNWTNTGCATGKNN	−3.83E+02	0	1,145	17.59%	1,616.2	7.27%
Foxa3(Forkhead)/Liver-Foxa3-ChIP-Seq(GSE77670)/Homer	BSNTGTTTACWYWGN	−3.44E+02	0	1,643	25.25%	2,928.1	13.18%
At5g04390(C2H2)/col200-At5g04390-DAP-Seq(GSE60143)/Homer	AGTGANDN	−3.06E+02	0	4,722	72.56%	12,862.7	57.88%
At1g68670(G2like)/colamp-At1g68670-DAP-Seq(GSE60143)/Homer	WNWWHNRAAGATTCT	−3.06E+02	0	1,248	19.18%	2,052.4	9.24%

Table 3. *Continued*

Motif Name	Consensus	Log p value	q value (Benjamini)	# of target sequences with motif (of 6,508)	% of target sequences with motif	# of background sequences with motif (of 22,227)	% of background sequences with motif
OCT:OCT-short(POU,Homeobox)/ NPC-OCT6-ChIP-Seq(GSE43916)/ Homer	ATGCATWA TGCATRW	−2.76E+02	0	1,888	29.01%	3,839.3	17.28%
FOXP1(Forkhead)/H9-FOXP1-ChIP-Seq(GSE31006)/Homer	NYYTGTTT ACHN	−2.64E+02	0	2,013	30.93%	4,242.2	19.09%
FOXK2(Forkhead)/U2OS-FOXK2-ChIP-Seq(E-MTAB-2204)/Homer	SCHTGTTT ACAT	−2.47E+02	0	2,081	31.98%	4,514.6	20.32%
At2g01060(G2like)/colamp-At2g01060-DAP-Seq(GSE60143)/ Homer	AGATKCB NWW	−2.25E+02	0	4,320	66.38%	11,893	53.52%
FoxD3(forkhead)/ZebrafishEmbryo-Foxd3.biotin-ChIP-seq(GSE106676)/ Homer	TGTTTAYT TAGC	−2.22E+02	0	1,821	27.98%	3,889.1	17.50%
GTL1(Trihelix)/colamp-GTL1-DAP-Seq(GSE60143)/Homer	WWTTTA CCKY	−2.10E+02	0	2,489	38.25%	5,930.3	26.69%
Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq(Fan_et_al.)/Homer	CTGTTTAC	−2.02E+02	0	2,670	41.03%	6,536.3	29.41%
Foxo3(Forkhead)/U2OS-Foxo3-ChIP-Seq(E-MTAB-2701)/Homer	DGTAAACA	−1.81E+02	0	2,309	35.48%	5,549.6	24.97%
STZ(C2H2)/colamp-STZ-DAP-Seq(GSE60143)/Homer	HNBTCACT	−1.75E+02	0	4,561	70.08%	13,110.5	59.00%
At3g24120(G2like)/col-At3g24120-DAP-Seq(GSE60143)/Homer	NWWAGM ATMW	−1.70E+02	0	3,705	56.93%	10,141.5	45.64%
AT5G60130(ABI3VP1)/col-AT5G60130-DAP-Seq(GSE60143)/ Homer	WTTYTAA GVAAA	−1.60E+02	0	3,003	46.14%	7,879.9	35.46%
RBFox2(?)/Heart-RBFox2-CLIP-Seq(GSE57926)/Homer	TGCATGCA	−1.56E+02	0	2,151	33.05%	5,222.6	23.50%
At3g12730(G2like)/colamp-At3g12730-DAP-Seq(GSE60143)/ Homer	AAGATTCT	−1.37E+02	0	1,807	27.77%	4,304.9	19.37%
Foxa2(Forkhead)/Liver-Foxa2-ChIP-Seq(GSE25694)/Homer	CYTGTTTA CWYW	−1.32E+02	0	2,170	33.34%	5,447.6	24.52%
AT5G47660(Trihelix)/colamp-AT5G47660-DAP-Seq(GSE60143)/ Homer	AWTTTTA CCG	−1.18E+02	0	2,630	40.41%	7,002	31.51%
Fox:Ebox(Forkhead,bHLH)/Panc1-Foxa2-ChIP-Seq(GSE47459)/Homer	NNNVCTGW GYAAACASN	−1.08E+02	0	1,541	23.68%	3,701.6	16.66%
AT5G45580(G2like)/colamp-AT5G45580-DAP-Seq(GSE60143)/ Homer	ADRGAATCTH	−9.23E+01	0	2,378	36.54%	6,416	28.87%
FOXM1(Forkhead)/MCF7-FOXM1-ChIP-Seq(GSE72977)/Homer	TRTTTACTTW	−8.64E+01	0	2,118	32.54%	5,644.6	25.40%
GT2(Trihelix)/colamp-GT2-DAP-Seq(GSE60143)/Homer	AMGGTAAA WWWN	−8.22E+01	0	2,462	37.83%	6,782.8	30.52%
KAN2(G2like)/colamp-KAN2-DAP-Seq(GSE60143)/Homer	ATATTCTY	−7.27E+01	0	1,751	26.91%	4,618.3	20.78%
FOXK1(Forkhead)/HEK293-FOXK1-ChIP-Seq(GSE51673)/Homer	NVWTGTTTAC	−5.79E+01	0	2,358	36.23%	6,705.1	30.17%
AT2G38300(G2like)/col-AT2G38300-DAP-Seq(GSE60143)/Homer	ADRGAATGTT	−5.36E+01	0	1,736	26.67%	4,761.6	21.43%
FoxL2(Forkhead)/Ovary-FoxL2-ChIP-Seq(GSE60858)/Homer	WWTRTAA ACAVG	−3.55E+01	0	2,059	31.64%	6,026.5	27.12%
FOXA1(Forkhead)/MCF7-FOXA1-ChIP-Seq(GSE26831)/Homer	WAAGTAA ACA	−3.52E+01	0	2,147	32.99%	6,318.9	28.44%
KANADI1(Myb)/Seedling-KAN1-ChIP-Seq(GSE48081)/Homer	ARGAATA WWN	−3.16E+01	0	1,833	28.17%	5,351	24.08%
MYB73(MYB)/col-MYB73-DAP-Seq(GSE60143)/Homer	NNNNHAAC NGHHDHN	−1.30E+01	0.0001	2,749	42.24%	8,766.4	39.45%

Table 3. *Continued*

Motif Name	Consensus	Log p value	q value (Benjamini)	# of target sequences with motif (of 6,508)	% of target sequences with motif	# of background sequences with motif (of 22,227)	% of background sequences with motif
AP-2gamma(AP2)/MCF7-TFAP2C-ChIP-Seq(GSE21234)/Homer	SCCTSAG GSCAW	−1.29E+01	0.0001	64	0.98%	116.8	0.53%
FOXA1(Forkhead)/LNCAP-FOXA1-ChIP-Seq(GSE27824)/Homer	WAAGTA AACA	−1.13E+01	0.0003	2,454	37.71%	7,822	35.20%
AP-2alpha(AP2)/Hela-AP2alpha-ChIP-Seq(GSE31477)/Homer	ATGCCCTG AGGC	−1.01E+01	0.001	46	0.71%	83	0.37%
EBF1(EBF)/Near-E2A-ChIP-Seq(GSE21512)/Homer	GTCCCCW GGGGA	−8.99E+00	0.0029	55	0.85%	109.4	0.49%
Foxf1(Forkhead)/Lung-Foxf1-ChIP-Seq(GSE77951)/Homer	WWATRTA AACAN	−8.66E+00	0.004	2,300	35.34%	7,385.2	33.23%
At3g60580(C2H2)/col-At3g60580-DAP-Seq(GSE60143)/Homer	WTTYTACT	−8.25E+00	0.0058	3,791	58.25%	12,469.4	56.11%
SeqBias: GCW-triplet	GCWGCWG CWGCW	−8.06E+00	0.0069	6,440	98.96%	21,878	98.46%
STOP1(C2H2)/colamp-STOP1-DAP-Seq(GSE60143)/Homer	DTATCTGG KGRAGGT	−7.26E+00	0.015	152	2.34%	396.6	1.78%
At2g33710(AP2EREBP)/colamp-At2g33710-DAP-Seq(GSE60143)/Homer	WTKGCGG CKR	−6.01E+00	0.0516	597	9.17%	1,821.2	8.20%
LOB(LOBAS2)/col-LOB-DAP-Seq(GSE60143)/Homer	CGCCGKAW WTTHCGS	−5.92E+00	0.0551	81	1.24%	199.5	0.90%

ATAC motifs found in the control (non-exposed) and in carabaryl exposed samples. The motif name; the consensus motif sequence; the log p value; the corrected Benjamini q value; the Number of Target Sequences (including their percentage); and the number of Background Sequences with Motifs (including their percentage) are shown.

Potential solutions

Each suggested solution below or a combination thereof can be used to improve the transposition efficiency.

Different amounts of transpositions enzyme may be used. In the optimization of this protocol, different dilutions of the transposase suggested concentration were used: 100%, 50%, 25%, and 12% (Figure 3). A concentration corresponding to 25% of the manufacturer suggested concentration resulted in good quality ATAC-seq Daphnia libraries. Different concentrations may be used for other organisms.

Different incubation times may be tested for the transposition reaction. In the optimization of this protocol we used 16–20 h long incubation as well as 4h incubation times for the transposase reaction. Whereas no appreciable differences were observed in the Daphnia libraries, this may differ in other species.

Access to chromatin may be improved by mechanic fragmentation. Access to chromatin within nuclei can be enhanced by a mechanic fragmentation of the animal tissue. Including a mechanic fragmentation improves chromatic accessibility (Figure 2; tissue homogenizer) resulting in better quality libraries as compared to libraries obtained from manual fragmentation (Figure 2; pestle).

It is advisable to perform a protocol optimization on a reduced number of samples prior to processing large sample numbers.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Luisa Orsini (l.orsini@bham.ac.uk).

STAR Protocols
Protocol

CellPress
OPEN ACCESS

AUTHOR CONTRIBUTIONS

S.K. performed the protocol optimization with input from L.O. V.D. performed data analysis. L.O. conceived and coordinated the experiments. All authors contributed to the manuscript writing.

DECLARATION OF INTERESTS

The authors declare no competing interests

REFERENCES

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Batut, B., Hiltemann, S., Bagnacani, A., Baker, D., Bhardwaj, V., Blank, C., Bretaudeau, A., Brillet-Gueguen, L., Cech, M., Chilton, J., et al. (2018). Community-driven data analysis training for biology. Cell Syst. *6*, 752–758.e1.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. 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. Curr. Protoc. Mol. Biol. *109*, 21 29 21–21 29 29.

Cambronero, C.M., Marshall, H., De Meester, L., Davidson, A., Beckerman, A.P., and Orsini, L. (2018). Predictability of the impact of multiple stressors on the keystone species *Daphnia*. Sci. Rep. *8*, 17572.

Cambronero Cuenca, M., and Orsini, L. (2018). Resurrection of dormant *Daphnia magna*: protocol and applications. J. Vis. Exp. *131*, e56637.

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, L., Aibar, S., Makhzami, S., Christiaens, V., Bravo Gonzalez-Blas, C., et al. (2018). A single-cell

transcriptome atlas of the aging *Drosophila* brain. Cell *174*, 982–998.e20.

Duttke, S.H., Chang, M.W., Heinz, S., and Benner, C. (2019). Identification and dynamic quantification of regulatory elements using total RNA. Genome Res. *29*, 1836–1846.

Ebert, D. (2005). Ecology, Epidemiology, and Evolution of Parasitism in Daphnia (National Library of Medicine (US), National Center for Biotechnology).

Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048.

Gaspar, J.M. (2018). Improved peak-calling with MACS2. bioRxiv. https://doi.org/10.1101/496521.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Jansen, M., De Meester, L., Cielen, A., Buser, C.C., and Stoks, R. (2011). The interplay of past and current stress exposure on the water flea *Daphnia*. Funct. Ecol. *25*, 974–982.

Kerfoot, W.C., and Weider, L.J. (2004). Experimental paleoecology (resurrection ecology): Chasing Van Valen's Red Queen hypothesis. Limnol. Oceanogr. *49*, 1300–1316.

Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in

multiplex sequencing on the Illumina platform. Nucleic Acids Res. *40*, 2513–2524.

Langmead, B., Wilks, C., Antonescu, V., and Charles, R. (2019). Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics *35*, 421–432.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

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*, 65.

Okonechnikov, K., Conesa, A., and Garcia-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics *32*, 292–294.

Orchard, P., Kyono, Y., Hensley, J., Kitzman, J.O., and Parker, S.C.J. (2020). Quantification, dynamic visualization, and validation of bias in ATAC-seq data with ataqv. Cell Syst *10*, 298–306 e294.

Quinlan, A.R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. Curr Protoc Bioinformatics *47*, 11.12.1–11.12.34.

Ramirez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Gruning, B.A., Villaveces, J., Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat. Commun. *9*, 189.