METHOD ARTICLE

A simple ATAC-seq protocol for population epigenetics

[version 1; peer review: 2 approved with reservations]

Ronaldo de Carvalho Augusto¹ ² ³, Aki Minoda³, Christoph Grunau¹

¹Univ. Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Univ. Montpellier, Perpignan, 66860, France
²LBMC, Laboratoire de Biologie et Modélisation de la Cellule Univ Lyon, ENS de Lyon, Université Claude Bernard Lyon 1, CNRS, UMR 5239, INSERM, U1210, Lyon, 69007, France
³RIKEN Center for Integrative Medical Sciences, Epigenome Technology Exploration Unit, Tsurumi, Kanagawa, 230-0045, Japan

First published: 09 Jun 2020, 5:121
https://doi.org/10.12688/wellcomeopenres.15552.1
Latest published: 07 Jan 2021, 5:121
https://doi.org/10.12688/wellcomeopenres.15552.2

Abstract
We describe here a protocol for the generation of sequence-ready libraries for population epigenomics studies. The protocol is a streamlined version of the Assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq) that provides a positive display of accessible, presumably euchromatic regions. The protocol is straightforward and can be used with small individuals such as daphnia and schistosome worms, and probably many other biological samples of comparable size, and it requires little molecular biology handling expertise.

Keywords
epigenetics, epigenomics, ATAC-seq, Daphnia pulex, Schistosoma mansoni

Open Peer Review

Reviewer Status
Invited Reviewers
1 2

version 2
(revision)
07 Jan 2021

version 1
09 Jun 2020

report

Any reports and responses or comments on the article can be found at the end of the article.
Corresponding author: Christoph Grunau (christoph.grunau@univ-perp.fr)

Author roles: Augusto RdC: Conceptualization, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Minoda A: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Grunau C: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: R.A. and C.G. are supported by Wellcome (Strategic award 107475). This work was performed with the support of LabEx CeMEB, an ANR « Investissements d’avenir » program (ANR-10-LABX-04-01) and the RTP Epigenetics in Ecology and Evolution (RTP3E). C.G. and A.M. received support from JSPS grant S17087.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Augusto RdC et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Augusto RdC, Minoda A and Grunau C. A simple ATAC-seq protocol for population epigenetics [version 1; peer review: 2 approved with reservations] Wellcome Open Research 2020, 5:121 https://doi.org/10.12688/wellcomeopenres.15552.1

First published: 09 Jun 2020, 5:121 https://doi.org/10.12688/wellcomeopenres.15552.1
Introduction
Understanding the dynamic cross-talk between epigenetic mechanisms and environmental cues of animal populations is of fundamental importance for ecologists and evolutionary biologists (Shi et al., 2019). The dynamics of chromatin has long been of interest as a source of phenotypic variance within and among animal populations (Hu & Barrett, 2017; Zhang et al., 2018) and can affect their ecological performance (Augusto et al., 2019; Hawes et al., 2018). In eukaryotic cells, chromatin is a dynamic structure that provides epigenetic information to regulate cell function and gene expression (Chen & Dent, 2014). The physical organization of the chromatin landscape modulates accessibility of genomic regions and dynamically response to both external and internal stimuli. In general, accessible genomic regions are enriched in regulatory elements during gene activity while inaccessible regions restrict binding of transcriptional regulators resulting in gene silencing (Stergachis et al., 2013). Assay for transposase accessible chromatin with high throughput sequencing (ATAC-Seq) is a technique used to assess genome-wide chromatin accessibility. ATAC-seq uses the Tn5 transposase, which fragments DNA and adds adapters simultaneously in open chromatin regions (that are accessible to Tn5 transposase). Deep sequencing of the PCR amplified Tn5 accessible regions provides a high-resolution map of accessible chromatin regions in the genome. We reasoned that this technique can not only be used to establish functional links between chromatin structure and gene function, but also to quantify epigenetic diversity in populations. This would require generation of ATAC-seq chromatin maps in single individuals. In addition, the technique should be sufficiently robust to be used by scientists who are experts in the field of population (epi)genetics and ecology, but having potentially received little training in molecular biology.

Here, we describe a streamlined and robust method for ATAC-Seq of individuals of the crustacea Daphnia pulex and also for the trematode Schistosoma mansoni. Our procedure is based on the protocol from Buenrostro et al. (2015); Corces et al. (2016) and Nextera DNA Library Preparation Kit (2017). Besides their ecological and epidemiological importance, both abovementioned organisms show high phenotypic plasticity in response to environmental cues (e.g. the presence of predator for daphnia) or during their development (schistosoma). This phenotypic plasticity is in line with epigenetic plasticity and makes them suitable models for monitoring chromatin accessibility changes. To validate the robustness of the method, we ask 13 experimenters with different levels of expertise in molecular biology to run the experimental procedure independently using D. pulex. This validation was performed during the summer school ‘Epigénétique en Ecologie et Evolution’ organized by the RTP3E in June 2018. ATAC-Seq on S.mansoni was done in our laboratory. The procedure provides robust results with individual D. pulex and adult worms of S. mansoni specimens, but other organisms of similar cell number can probably be used. Controls must be done without organisms as input.

Materials and methods
Animal sampling and transposase mixture
Fresh adults S. mansoni worms were collected from 4 female 4-week-old Swiss OF1 mice (weight mean 18g) supplied by Charles River, L’arbresle, France Mice had been infected by peritoneal injection with 150 mixed sexes cercariae. Hepatic perfusions were performed with lethal injection of 1mg per kg body weight of sodium pertobartial solution (Dolehal, Vetoquinol, Lure, France) after 65 days post infection. Water and food were given ad libitum, 12h light/dark cycle, 25°C. Housing, feeding and animal care followed the national ethical standards established in the wrt of 1 February 2013 (NOR: AGRG1238753A) setting the conditions for approval, planning and operation of establishments, breeders and suppliers of animals used for scientific purposes and controls. The French Ministère de l’Agriculture et de la Pêche and French Ministère de l’Éducation Nationale de la Recherche et de la Technologie provided permit A66040 to our laboratory for experiments on animals and certificate for animal experimentation (authorization 007083, decree 87–848) for the experimenters.

Living Daphnia pulex were sampled by pipetting through a 1 mL automatic pipette with enlarged openings of the pipetting tips. To avoid experimenter bias, 13 different persons sampled at least one individual. Finally, each specimen was then individually transferred to a 1.5 mL Eppendorf tube and was immediately processed for ATAC-seq library preparation as follows.

The necessary material is listed in Table 1 and Table 2 and must be prepared in advance. In addition, nuclease-free water, high fidelity DNA polymerase for PCR and corresponding buffers, freshly prepared 80% ethanol, refrigerated centrifuge, 0.2 ml PCR tubes, 1.5 ml tubes, ThermoMixer with agitation, PCR thermal cycler, qPCR instrument, magnetic rack, 1 mL pipette, 100 μL pipette, and 10 μL pipette are needed.

An Eppendorf ThermoMixer was then set with agitation to 37°C and the following steps performed.

Table 1. Externally sourced materials.

<table>
<thead>
<tr>
<th>Item name</th>
<th>Vendor</th>
<th>Catalogue ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1% Molecular biology-grade IGEPAL CA-630</td>
<td>Sigma-Aldrich</td>
<td>I8896</td>
</tr>
<tr>
<td>2xTD (Tagment DNA buffer from Nextera kit)</td>
<td>Illumina</td>
<td>FC-121-1030</td>
</tr>
<tr>
<td>TDE1 (Tagment DNA Enzyme from Nextera kit)</td>
<td>Illumina</td>
<td>FC-121-1030</td>
</tr>
<tr>
<td>QIAquick PCR Purification Kit</td>
<td>Qiagen</td>
<td>28104</td>
</tr>
<tr>
<td>AMPure XP beads</td>
<td>Agencourt</td>
<td>A63880</td>
</tr>
<tr>
<td>Bioanalyzer High-Sensitivity DNA Analysis kit</td>
<td>Agilent</td>
<td>5067-4627</td>
</tr>
<tr>
<td>10.000X SYBR I</td>
<td>Invitrogen</td>
<td>S-7563</td>
</tr>
</tbody>
</table>

Table 2. Reagents produced in the laboratory.

<table>
<thead>
<tr>
<th>Phosphate buffered saline (PBS)</th>
<th>137mM NaCl; 27mM KCl; 100mM NaHPO₄; 18mM KH₂PO₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagmentation buffer (TD buffer)</td>
<td>(Wang et al., 2013) 20 mM Tris(hydroxymethyl)aminomethane; 10 mM MgCl₂; 20% (vol/vol) dimethylformamide</td>
</tr>
</tbody>
</table>
• For *D. pulex* remove all water by pipetting with 100 μL tip;
  or
  Perfuse *S. mansoni* worms and take single worm as dry as possible with forceps
• Wash once with 50 μL of cold 1x PBS buffer and remove all supernatant by pipetting, being careful not to remove your sample;
• Add to each sample
  o 25 μl 2× TD buffer *(Wang et al., 2013)*
  o 2.5 μl TDE1
  o 0.5 μl 1% IGEPAL
  o 22 μl Nuclease-free water

This gives 50 μL of transposase mixture for each sample. The samples are pipetted up and down 10 times to disrupt cells.

**Chromatin tagmentation**

This step uses the Nextera Tn5 transposome to ‘tagment’ the chromatin, which is a process that fragments the chromatin and tags the DNA with adapter sequences in a single step.

• Tagmentation reactions are incubated at 37°C for 30 min in an Eppendorf ThermoMixer with agitation at 300 rpm.
• Tagmented chromatin is immediately purified using a QIAGEN MinElute Reaction Cleanup kit or a QIAquick PCR Purification Kit, and purified DNA is eluted into 20 μl of elution buffer (10 mM Tris-HCl, pH 8).

**Library amplification**

This step amplifies the tagmented DNA using a limited-cycle PCR program. PCR is carried out with a universal index Ad1 and an index (barcode) primer Ad2, as described in Table 3 *(Buenrostro et al., 2015)*. Two library amplification methods were tested and validated in our hands as follows:

---

**Table 3. PCR Primer ID and sequence.**

<table>
<thead>
<tr>
<th>Index ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad1_noMX:</td>
<td>AATGATACGCGACCATCCAGCTCAGATGTCAGATGT</td>
</tr>
<tr>
<td>Ad1.1_TAAGGCAGA</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.2.CGACTAG</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.3_AGCCAGA</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.4_TCTCTAC</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.5_GGATCTCTC</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.6_TAGGCATG</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.7_CTCTCTAC</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.8_CAGAGAGG</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.9_GCTACGCT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.10.CGAGGCTG</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.11_AAGAGGCA</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.12_GTGAGGA</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.13_GTCGTGAT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.14_ACCACTGT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.15_TGGATCTG</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.16_CCGTGGT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.17_TGCTGGT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.18_GAGGGGT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.19_AGGTGGG</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.20_GTGGGGT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.21_TGGTGGTC</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.22_TGGTGGTC</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.23_TGGTGGTC</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
<tr>
<td>Ad1.24.CGACTCCT</td>
<td>CAAAGCAAGCGCGATACAGAGATCCCTAGTCTCGAGGAGATGT</td>
</tr>
</tbody>
</table>
Option 1 (for Promega GoTag G2). Combine the following in a PCR tube for each sample: 9.5 μl Nuclease-free MilliQ water; 20 μl Purified transposed DNA; 10 μl 5x GoTag Q2 buffer; 4 μl MgCl₂; 2.5 μl Universal Ad1_noMX primer (25 μM); 2.5 μl Specific Index primer Ad2.*, different for each sample (25 μM); 1 μl dNTPs (10 mM); 0.5 μl GoTaQ G2.

Or

Option 2 (for NEB mix, more convenient but more expensive). Combine the following in a PCR tube for each sample: 20 μl purified transposed DNA; 2.5 μl Universal Ad1_noMX primer (25 μM); 2.5 μl Specific Index primer Ad2.*, different for each sample (25 μM); 25 μl NEBNext High-Fidelity 2X PCR Master Mix

In both options the final volume is 50 μl. The samples are pre-amplified using a PCR machine with the program described in Table 4.

In order to reduce GC and size bias in PCR, the PCR reaction is monitored using qPCR to stop amplification prior to saturation. To run a qPCR side reaction, we combined the following depending on the option that had been chosen previously:

Option 1: 5 μl PCR product of the initial pre-amplification reaction (keep the remaining 45 μl at 4°C); 2.5 μl 5x GoTag Q2 buffer; 0.1 μl GoTaQ 2; 3.14 μl Nuclease-free MilliQ water; 0.25 μl Universal Ad1_noMX primer (25 μM); 0.25 μl Ad2.* indexing primer (25 μM); 1 μl MgCl₂; 0.25 μl dNTPs; 0.1 μl 100X SYBR I

or

Option 2: 5 μl PCR product of the initial pre-amplification reaction; (keep the remaining 45 μl at 4°C); 4.41 μl Nuclease-free MilliQ water; 0.25 μl Ad1_noMX primer (25 μM); 0.25 μl Ad2.* indexing primer (25 μM); 0.25 μl dNTPs; 0.25 μl 20X SYBR I; 5 μl NEBNext High-Fidelity 2X PCR MasterMix

The samples are amplified in a qPCR machine with the program set out in Table 4.

To calculate the optimal additional number of cycles needed for the remaining 45 μl PCR, relative fluorescence is plotted against cycle number and the cycle number that corresponds to one-third of the maximum fluorescent intensity is determined (Figure 1). In our experience, the total number of amplification cycles should not exceed 21 (Augusto et al., 2019).

The remaining 45 μl PCR reaction is run with the additional number of cycles and purified with a QIAGEN MinElute Reaction Cleanup kit or a QIAquick PCR Purification Kit or similar, and eluted with 20 μl of elution buffer (10 mM Tris-HCl, pH 8).

Fragments are separated by electrophoresis through a 1.5% agarose gel or on a Bioanalyzer chip. A ladder that corresponds to the nucleosome-free region and multiple nucleosome-size fragments should be seen (one nucleosome = about 150 bp). A single band at around 150 bp indicates sample degradation or over-fragmentation. Ideally, five bands should be obtained, three bands are acceptable (Figure 2).

AMPure XP beads double-side purification
This step enriches for the nucleosome-free (~300 bp). Removing small fragments (primer dimers) is important for optimal sequencing. First transfer 45 μl to an Eppendorf tube (or use PCR tube directly), add 22.5 μl (0.5X original volume, to remove large fragments) AMPure XP beads, pipet up and down 10 times to mix thoroughly. Incubate at room temperature for 10 minutes and place tubes in magnetic rack for 5 minutes. Transfer supernatant to new tube and add 58.5 μl (1.3X original volume, to remove small fragments) AMPure XP beads, pipet up and down 10 times to mix thoroughly. Incubate at room temperature for 10 minutes, place tubes in magnetic rack for 5 minutes and discard supernatant. Wash beads with 200 μl 80% ethanol (freshly made), pipet ethanol over beads 10 times, then discard ethanol. Ensure all ethanol is removed. Leave tube on magnetic rack with cap open for 3 to maximum 10 minutes depending on ambient humidity. The beads should be ‘glowing’ but not wet. Be careful not to over-dry them, which will decrease elution efficiency. Resuspend beads in 20 μl nuclease-free water, pipet up and down 10 times to mix thoroughly, place tube in magnetic rack for 1–5 minutes and transfer supernatant to new tube. This step can be replaced by Diagenode IP-Star, size selection 320 bp.

We have not systematically investigated if different purification procedures influence on the result. Purified libraries should be stored at -20°C.

| Table 4. PCR program for library pre-amplification. |
| --- | --- | --- | --- |
| Step | Temp | Duration | Cycles |
| Pre-Warming | 72°C | 5 min | 1 |
| Initial denaturation | 98°C | 30 sec | 1 |
| Denaturation | 98°C | 10 sec | 1 |
| Annealing | 63°C | 30 sec | 5 |
| Extension | 72°C | 1 min | 1 |
| HOLD | 12°C | ∞ | 1 |

| Table 5. PCR program for library amplification. |
| --- | --- | --- | --- |
| Step | Temp | Duration | Cycles |
| Initial denaturation | 98°C | 30 sec | 1 |
| Denaturation | 98°C | 10 sec | 1 |
| Annealing | 63°C | 30 sec | 20 |
| Extension | 72°C | 1 min | 1 |
| HOLD | 12°C | ∞ | 1 |
Figure 1. Example amplification profile for four samples. (X-axis) Number of PCR cycles. (Y-axis) Fluorescence intensity. An optimal number of additional cycles to perform for four ATAC-seq libraries are indicated.

Figure 2. Examples of fragmentation profiles. (X-axis) Base pairs. (Y-axis) Fluorescence intensity. Peaks correspond to nucleosome-free region, mono- to tetra-nucleosome fractions. Bottom lane: too strong fragmentation, thus Tn5 incubation time needs to be decreased.

Libraries check
Size profiling can be performed using an Agilent Bioanalyzer High Sensitivity DNA Assay. Expected profiles are shown in Figure 2. Bioanalyzer profiles or KAPA library quantification kit are used to quantify libraries and proceed to sequencing.


Results
The method can be used by scientists with low expert level in molecular biology
The protocol described in the methods section was tested by 13 volunteers with molecular biology expertise level ranging from untrained, some with many years of experience and others who had retired from active wet-bench work several years ago. Each volunteer processed one experiment using *Daphnia*. In only two cases ATAC-seq library production did not succeed.

The method can be used with small organisms
After having had firmly established that the method delivers robust results for *Daphnia*, we used it on individual adult worms of *Schistosoma mansoni* and the protocol delivered expected results in terms of fragmentation, number of PCR cycles and library insert size distribution (Figure 3). Data underlying these results are available at Zenodo (Augusto *et al.*, 2020).
Discussion

Phenotypically, plasticity plays an important role in development and evolution. The relative contribution of genetic and epigenetic components to heritable plasticity is a matter of lively scientific debate (Hu & Barrett, 2017; Roquis et al., 2018). One of the caveats of analyzing epigenetic information is that it is stored in several, very different bearers of information (e.g. DNA methylation, modification of histones, non-coding RNA and topological position in the interphase nucleus). Nevertheless, these types of information converge towards a change in chromatin structure which can be approximated by DNA accessibility. We reasoned that a straightforward ATAC-seq method to map the chromatin accessibility status in populations with high phenotypic plasticity would facilitate further investigations of the role of epigenetics in plasticity. This study field is also of particular importance to field ecologists. We therefore set-out to establish a robust, easy to use protocol that can be used with little molecular biology training. Our protocol was successfully used in the framework of a summer school ‘Epigénétique en Ecologie et Evolution’ by participants with different levels of expertise in molecular biology using D. pulex. We also used single adult S. mansoni worms as biological material. We believe that our protocol is suitable for fast epigenotyping of other organisms as well. From our experience, the only parameter that it might be necessary to optimize is tagmentation reaction time if over- or under-fragmentation occurs. A potential issue is contamination with microorganisms whose DNA might be present in the libraries.

Data availability


This project contains the following underlying data:

- Agarose picture (TIF). (Example of electrophoresis fragment separation.)
- Agarose profile (PNG) (Example of fragment separation on a BioAnalyzer chip.)
- BioAnalyzer, BioAnalyzer 2–4 (PDF). (BioAnalyzer profiles generated in this study.)
- qPCR cycles (XLSX). (Quantification of qPCR cycles for each indicated organism.)
- qPCR plot (JPG). (qPCR amplification cycles plot.)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgments

We are grateful to the 13 volunteers who tested the procedure.

References


Nextera XT DNA Library Prep Kit - Reference Guide. Reference Source


Open Peer Review

José María Santos-Pereira
Andalusian Center for Development Biology (CABD), Spanish National Research Council, Pablo de Olavide University, Seville, Spain

In this protocol article, Augusto and colleagues describe a version of the popular ATAC-seq method for chromatin accessibility profiling in the invertebrate organisms *Daphnia pulex* and *Schistosoma mansoni*. Their main claim is that this protocol can be used by researchers with little molecular biology training. However, only the wet-lab component of the ATAC-seq experiment is covered by the manuscript, while the computational analysis of the obtained sequences is omitted. I think that this is a major limitation that has to be addressed before indexing (please, see below).

Major points:
- Using Tn5 transposase with the appropriate number of cells is critical for this protocol. Authors mention that TAGmentation time should be reduced in case that libraries are over-TAGmented. However, they do not mention the cell numbers used with these organisms, nor the number of *Schistosoma* individuals. This is important to be clearly explained in the manuscript.

- The authors claim that this protocol produces robust results. However, the results are not shown at all. A complete ATAC-seq protocol should cover from the sample preparation in the lab to the analysis of the generated results. I think this is a major limitation of this article and that the authors should explain, at least, the primary computational analysis of the data, the performed quality controls and examples of data visualization. An example of basic analysis of the data, such as peak calling and motif enrichment analysis in the called peaks would be much more helpful to understand how this technique may help to study epigenetic diversity at the population level. Of course, these data have to be available for the community, for example by a GEO accession code.

Minor points:
- The origin of the *Daphnia* samples is not explained.
- While I assume that the use of 1% Igepal in the TAGmentation reaction is what lyses the
samples, this has to be stated more clearly, since usually cell lysis is performed before TAGmentation in other ATAC-seq protocols.

- The Promega GoTaq G2 polymerase is not a high-fidelity enzyme and therefore I would not recommend its use for this purpose, since this could result in mutations in the amplified molecules and decreased alignment efficiency to the reference genome. Have the authors noted this when analyzing their data in comparison with NEB Next?
- How was the quality of the results obtained by the 13 volunteers? If this protocol can be performed by researchers with little molecular biology training, then the authors should show that the experiments performed by their volunteers were of enough quality to be used.
- The agarose gel showed in the supporting material has no lane labels, so it cannot be known what is shown. It would be helpful to show pictures of the agarose gels underlying Fig. 2 data in the main figure.

Is the rationale for developing the new method (or application) clearly explained? Yes

Is the description of the method technically sound? Partly

Are sufficient details provided to allow replication of the method development and its use by others? Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility? No

Are the conclusions about the method and its performance adequately supported by the findings presented in the article? No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Chromatin structure, developmental biology, functional genomics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 18 Nov 2020

Christoph Grunau, Univ. Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Univ. Montpellier, Perpignan, France

1- In this protocol article, Augusto and colleagues describe a version of the popular ATAC-seq
method for chromatin accessibility profiling in the invertebrate organisms *Daphnia pulex* and *Schistosoma mansoni*. Their main claim is that this protocol can be used by researchers with little molecular biology training. However, only the wet-lab component of the ATAC-seq experiment is covered by the manuscript, while the computational analysis of the obtained sequences is omitted. I think that this is a major limitation that has to be addressed before indexing (please, see below).

**It was our initial intention to show the robustness of the wet-bench part. The computational analysis of the obtained sequences is now also provided.**

**Major points:**

2- Using Tn5 transposase with the appropriate number of cells is critical for this protocol. Authors mention that TAGmentation time should be reduced in case that libraries are over-TAGmented. However, they do not mention the cell numbers used with these organisms, nor the number of *Schistosoma* individuals. This is important to be clearly explained in the manuscript.

*A sentence describing the number of Schistosoma individuals was added at the animal sampling and transposase mixture section. Number of nuclei were counted and added to the results section, paragraph “ATAC-Seq can be used on individual Daphnia and individual Schistosoma adults”*

3- The authors claim that this protocol produces robust results. However, the results are not shown at all. A complete ATAC-seq protocol should cover from the sample preparation in the lab to the analysis of the generated results. I think this is a major limitation of this article and that the authors should explain, at least, the primary computational analysis of the data, the performed quality controls and examples of data visualization. An example of basic analysis of the data, such as peak calling and motif enrichment analysis in the called peaks would be much more helpful to understand how this technique may help to study epigenetic diversity at the population level. Of course, these data have to be available for the community, for example by a GEO accession code **We have added a data analysis section and two examples of population epigenetics study to show the feasibility.**

The following sentence was added to the Data availability section: This article contains supporting information online on NCBI SRA (BioProject [PRJNA587385](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA587385)).

**Minor points**

4-The origin of the *Daphnia* samples is not explained **This is now in the Animal sampling part.**

5- While I assume that the use of 1% Igepal in the TAGmentation reaction is what lyases the samples, this has to be stated more clearly, since usually cell lysis is performed before TAGmentation in other ATAC-seq protocols. **We understand the reviewer's concern. When AM, RCA and CG optimized the method we realized that adding IGEPAL directly into the tagmentaion reaction did not alter**
the outcome and that it streamlined the protocol and made it more straightforward.

6- The Promega GoTaq G2 polymerase is not a high-fidelity enzyme and therefore I would not recommend its use for this purpose, since this could result in mutations in the amplified molecules and decreased alignment efficiency to the reference genome. Have the authors noted this when analyzing their data in comparison with NEB Next? **We did not observe differences between both polymerases following our pipeline.** GoTaq is cheaper and that could be an advantage when large number of libraries must be prepared.

7- How was the quality of the results obtained by the 13 volunteers? If this protocol can be performed by researchers with little molecular biology training, then the authors should show that the experiments performed by their volunteers were of enough quality to be used.

As indicated “The protocol described in the methods section was tested by 13 experimenters with molecular biology expert level ranging from untrained to over several 10 years of experience, or some who had retired from active wet-bench work several years ago. In only two cases ATAC-seq library production did not succeed. “ We also added the Figure 7 which shows the clustering of individual *daphnia* based on their ATAC-Seq profiles.

8- The agarose gel showed in the supporting material has no lane labels, so it cannot be known what is shown. It would be helpful to show pictures of the agarose gels underlying Fig. 2 data in the main figure.

**It is our understanding that the supporting data show the raw files. In the Zenode description was added: “In "Agarose picture.Tif“ the left lane shows the 100 bp size marker, first 10 bands from down to top: 100bp, 200bp, 300bp, 400bp, 500bp, 600bp, 700bp, 800bp, 900bp and 1kbp.”**

**Competing Interests:** No competing interest.

Reviewer Report 21 July 2020

[Link](https://doi.org/10.21956/wellcomeopenres.17031.r38998)

© 2020 Bantignies F. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Frédéric Bantignies

Institute of Human Genetics, CNRS, University of Montpellier, Montpellier, France

The authors describe here a simple ATAC-seq protocol for its use in population epigenetics. ATAC-
seq is now a common molecular biology technique used to assess chromatin accessibility in nuclei of culture cells or tissues. The authors have adapted this technique to small non-model organisms, which could be a way to determine epigenetic variation/signatures possibly associated with phenotypic variance in animal populations. Although this protocol seems really straightforward and useful for the community (even for non-molecular biologist), several points need to be clarified before its indexing.

**Major points:**
- Would be nice to have an estimation of the size and cell number in both *S.mansoni* and *D.pulex* specimens. This could be important to further adapt this protocol with other organisms of similar size, smaller or larger.
- There is no mention of nuclei isolation or tissue treatment before Transposase treatment. Maybe worth to mention if such steps are needed or not with these organisms, and why.
- There is no mention at all of the sequencing part per se of this ATAC-seq protocol. The protocol closes with “library check”. A paragraph should be added to present the sequencing part of this ATAC-seq protocol. In particular, mention if *D.pulex* and *S.mansoni* genomes are well annotated and/or mention if alternative procedure can be used for sequence data analysis in these cases.
- What would have been very interesting is to provide an illustration that this technique is indeed suitable to quantify epigenetic diversity in populations. For instance, it would have been nice to have examples of a genome browser display emanating from their multiple ATAC-seq libraries in *D.pulex* and/or *S.mansoni*. Maybe, the authors could also refer to some published or ongoing work.

**Minor points:**
- Page 2/Introduction: when mentioning “ATAC-seq uses the Tn5 transposase, which fragments DNA and adds adapters simultaneously in open chromatin regions (that are accessible to Tn5 transposase)”, the author should further precise where the transposase is cutting the DNA, i.e. in open chromatin, on both sides of nucleosomes and in larger nucleosome free-regions... This would also help to interpret Figure 2 and 3.
- Page 2/Introduction: when mentioning “Besides their ecological and epidemiological importance, both abovementioned organisms show high phenotypic plasticity in response to environmental cues (e.g. the presence of predator for daphnia) or during their development (schistosoma)”. For non-ecology specialists, the authors should elaborate a bit more on that, or at least cite references illustrating this intriguing abovementioned phenotypic plasticity in line with associated epigenetic changes.
- Page 2/Introduction/end of the paragraph: I would move “Controls must be done without organisms as input” in the Mat&Met/transposase mixture section.
- Page 2/ Mat&Met: I would separate “Animal sampling” and “transposase mixture” sections.
- Page 2/Table2/Reagents produced in the laboratory: I would recommend using commercially available PBS1X, and not sure why another Tagmentation (TD) buffer is indicated here. Usually, the TD buffer is directly provided with the enzyme (as indicated in
Table 1). If so, Table 2 is not needed.

○ Page 4/Figure 1 could be a bit more explicit. At least indicate on the graph the one-third of the maximum fluorescent intensity to determine the additional PCR cycles.

○ Page 4/AMPure XP beads double-side purification: “This step enriches for the nucleosome-free (~300 bp) ... “regions as well as di and tri-nucleosome fragments”. Maybe more correct like this. This specific point could be also mentioned in the Figure 3 legend.

○ Also mentioned: “First transfer 45 ul to an Eppendorf tube (or use PCR tube directly)”. After the additional PCR cycles and purification, DNA should be in 20 ul elution buffer. Need to be corrected or clarified for consistency in the following of the protocol.

○ Page 5/Results: Most data are also presented via a zenodo interface. However, the legends and labels of these figures are not always explicit, and the correspondence with the actual figures of the protocol is not always intuitive, something that could be readily improved.

○ Page 6/Figure 3: the figure is showing fragment sizing. All profiles are rather homogeneous, except one, D. Pulex/bottom-left, with two clear peaks, one of small fragments in the range of 35 bp (probably the primers), one of large fragments in the range of 10 kb. As they did for Figure 2, the authors should comment on that and propose an alternative solution when this particular situation is encountered, i.e. needs further purification or it is acceptable to process.

Is the rationale for developing the new method (or application) clearly explained?
Partly

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Genetic and epigenetic, chromatin structure and organization.

I confirm that I have read this submission and believe that I have an appropriate level of
expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Author Response 18 Nov 2020**

**Christoph Grunau**, Univ. Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Univ. Montpellier, Perpignan, France

**Major points:**

1- Would be nice to have an estimation of the size and cell number in both *S.mansoni* and *D.pulex* specimens. This could be important to further adapt this protocol with other organisms of similar size, smaller or larger

**The number of nuclei were counted in daphnia and schistosoma worms.**

2- There is no mention of nuclei isolation or tissue treatment before Transposase treatment. Maybe worth to mention if such steps are needed or not with these organisms, and why

We appreciate the reviewer’s concern. This was one of the streamlining steps and in our hands no nuclei isolation was necessary. We tested both approaches. Maybe the small size of the organisms play a role.

3- There is no mention at all of the sequencing part per se of this ATAC-seq protocol. The protocol closes with “library check”. A paragraph should be added to present the sequencing part of this ATAC-seq protocol. In particular, mention if *D.pulex* and *S.mansoni* genomes are well annotated and/or mention if alternative procedure can be used for sequence data analysis in these cases.

**The sequencing and analysis parts were added.**

4- What would have been very interesting is to provide an illustration that this technique is indeed suitable to quantify epigenetic diversity in populations. For instance, it would have been nice to have examples of a genome browser display emanating from their multiple ATAC-seq libraries in *D.pulex* and/or *S.mansoni*. Maybe, the authors could also refer to some published or ongoing work.

We have added the description of the population epigenetics part. We agree that IGV visualisation is definitely useful to get an impression of the data especially in relation to genome annotations. However, we feel that the display is not so useful for the population studies. It was not our intention to use this method to infer any functional relationships between chromatin structure and e.g. gene expression.

**Minor points:**

5- Page 2/Introduction: when mentioning “ATAC-seq uses the Tn5 transposase, which fragments DNA and adds adapters simultaneously in open chromatin regions (that are accessible to Tn5 transposase)”, the author should further precise where the transposase is cutting the DNA, i.e. in open chromatin, on both sides of nucleosomes and in larger nucleosome free regions...

This would also help to interpret Figure 2 and 3

We modified the sentence as follow: “ATAC-seq works similarly as DNase-seq (DNase I...
hypersensitive sites with high-throughput sequencing) (Song & Crawford 2010), and determines which genomic regions are accessible to Tn5 transposase (i.e. open chromatin regions), presumably the regulatory regions. Tn5 transposase inserts Illumina adapter sequences upon accessing the chromatin, which removes the need for additional steps to make the sequencing libraries later. This simple and efficient protocol reduces the starting material required, compared to DNase-seq. It also avoids many other steps such as the interaction with antibodies (e.g. ChIP-seq) or chemical treatment (e.g. FAIRE-seq, WGBS) that might introduce bias.”

6- Page 2/Introduction: when mentioning “Besides their ecological and epidemiological importance, both abovementioned organisms show high phenotypic plasticity in response to environmental cues (e.g. the presence of predator for daphnia) or during their development (schistosoma)”. For non-ecology specialists, the authors should elaborate a bit more on that, or at least cite references illustrating this intriguing abovementioned phenotypic plasticity in line with associated epigenetic changes.

We added the following sentence: “Classical studies have paved the way to a subsequent rich literature that has documented the amazing property of Daphnia to modify their phenotypes at the morphological, physiological, behavioral and more recently at the molecular levels in response to a large panel of environmental stressors including diet, pollution, heavy metals, and predator kairomones (reviewed in (Riessen 2011, Harris et al. 2012)). Furthermore, schistosomes deal with a multitude of signals from the water environment as well as cues that come from their hosts, shaping morphology, metabolism, and infection success in the short-term and also their full development later in life (Roquis et al. 2017, 2018, Augusto et al 2017, 2019). Our group characterized epigenetic mechanisms behind the phenotypic plasticity of schistosomes and their cross-talk with environmental cues, however Daphnia phenotype plasticity stills poorly understood.”

7- Page 2/Introduction/end of the paragraph: I would move “Controls must be done without organisms as input” in the Mat&Met/transposase mixture section.

This was done.

8- Page 2/ Mat&Met: I would separate “Animal sampling” and “transposase mixture” sections.

This was done.

9- Page 2/Table2/Reagents produced in the laboratory: I would recommend using commercially available PBS1X, and not sure why another Tagmentation (TD) buffer is indicated here. Usually, the TD buffer is directly provided with the enzyme (as indicated in Table 1). If so, Table 2 is not needed.

We understand the reviewer concern, however we would like to keep this information on the paper. Our group is running several experiments using ATAC-seq and this is a useful information in the case of missing PBS 1X and/or TD buffer. When we use the Illumina kit in the lab, we can often dilute very much the Tn5 enzyme and what gets limiting is the 2xTD buffer. The information about the composition of the buffer is actually not so easy to find and we wish to keep it in the article. We did not see any difference between commercial 2xTD buffer and “home-made” 2xTD. Also, Tn5 can
now be purchased separately and even with custom adaptors, which has the potential of even more streamline the protocol.

10- Page 4/Figure 1 could be a bit more explicit. At least indicate on the graph the one-third of the maximum fluorescent intensity to determine the additional PCR cycles. The figure was changed (now figure 2).

11- Page 4/AMPure XP beads double-side purification: “This step enriches for the nucleosome-free (~300 bp)” ... “regions as well as di and tri-nucleosome fragments”. Maybe more correct like this. This specific point could be also mentioned in the Figure 3 legend. We agree with reviewer suggestion and added the following sentence: This step enriches for the nucleosome-free (~300 bp) as well as di and tri-nucleosome fragments.

12- Also mentioned: “First transfer 45 ul to an Eppendorf tube (or use PCR tube directly)”. After the additional PCR cycles and purification, DNA should be in 20 ul elution buffer. Need to be corrected or clarified for consistency in the following of the protocol. Thank you to have found this! It was a “vestige” of an earlier version of the protocol. Elution is done into 45 µL now and not 20 µL.

13- Page 5/Results: Most data are also presented via a zenodo interface. However, the legends and labels of these figures are not always explicit, and the correspondence with the actual figures of the protocol is not always intuitive, something that could be readily improved. We understand the comment. We had provided an Excel file with the description of the supporting raw data. However, this file is not directly accessible. We would need guideline from the editors. Should we provide a detailed description in the manuscript or on the Zenodo depository?

14- Page 6/Figure 3: the figure is showing fragment sizing. All profiles are rather homogeneous, except one, D. Pulex/bottom-left, with two clear peaks, one of small fragments in the range of 35 bp (probably the primers), one of large fragments in the range of 10 kb. As they did for Figure 2, the authors should comment on that and propose an alternative solution when this particular situation is encountered, i.e. needs further purification or it is acceptable to process. Peaks at 35 bp and 10380 bp are spiked-in marker peaks for the BioAnalyser. They do not come from the samples. This was added to the figure legend.

**Competing Interests:** No competing interest.