Chapter 1

Genome-Wide Mapping of Active Regulatory Elements Using ATAC-seq

Georgi K. Marinov, Zohar Shipony, Anshul Kundaje, and William J. Greenleaf

Abstract

Active *cis*-regulatory elements (cREs) in eukaryotes are characterized by nucleosomal depletion and, accordingly, higher accessibility. This property has turned out to be immensely useful for identifying cREs genome-wide and tracking their dynamics across different cellular states and is the basis of numerous methods taking advantage of the preferential enzymatic cleavage/labeling of accessible DNA. ATAC-seq (**A**ssay for **T**ransposase-**A**ccessible **C**hromatin using **seq**uencing) has emerged as the most versatile and widely adaptable method and has been widely adopted as the standard tool for mapping open chromatin regions. Here, we discuss the current optimal practices and important considerations for carrying out ATAC-seq experiments, primarily in the context of mammalian systems.

Key words Enhancers, Promoters, Chromatin accessibility, ATAC-seq, High-throughput sequencing

1 Introduction

Eukaryotic chromatin is generally packaged by nucleosomes, octamer particles comprised of the four core nucleosomal histones H3, H4, H2A, and H2B [1]. Nucleosomal packaging has an inhibitory effect on transcriptional activity and prevents the binding of most transcription factors and other regulatory proteins. Active promoter and enhancer elements differ from the rest of the genome in that they usually exist in a depleted of nucleosomes, open chromatin state. This property is highly useful in practice because just as regulatory factors can access active cREs so can various enzymes, whose action is otherwise inhibited by nucleosome particles. That enhancers and promoters exhibit this property was already appreciated nearly four decades ago, when their hypersensitivity to cleavage by DNase enzymes was first reported [2–4].

Georgi K. Marinov and Zohar Shipony authors contributed equally to this work.

DNase remained the main tool for mapping active cREs into the genomic era, initially coupled to microarray readouts [5–7] and eventually adapted to a high-throughput sequencing format [8–10]. In parallel to these developments as well as more recently, a wide variety of alternative methods taking advantage of the preferential enzymatic/chemical cleavage/modification of accessible DNA were also developed, employing methyltransferases [11–15], restriction enzymes [16], nicking enzymes [17], small molecules [18], viral integration [19], and others.

ATAC-seq, which is based on the preferential insertion into unprotected DNA by a hyperactive mutant version of the Tn5 transposase [20] (Fig. 1), has emerged as the most convenient, widely adaptable and straightforward to execute method for profiling open chromatin. Treatment of chromatin with Tn5 results in the insertion into accessible DNA of adapters that then enable the direct amplification of open chromatin fragments. This eliminates much of complex series of enzymatic steps that are unavoidable features of previous methods such as DNase-seq, allows for the protocol to be completed in just a few hours, and also dramatically lowers the input requirements, down to a few tens of thousands of cells in bulk reactions as well as enabling single cell (scATAC) assays [21, 22].

In this chapter, we describe the most important considerations for carrying out successful ATAC-seq experiments in the context of the Omni-ATAC protocol, an optimized version of the ATAC-seq assay that produces high-quality ATAC libraries for most mammalian cell lines and cell types, as well as for a number of other eukaryotes.

2   Materials

Prepare a master stock of the ATAC-RSB buffer without detergents in a large volume (e.g., 50 mL) and store it 4 °C.

Prepare master stocks of 2× TD buffer (e.g., in 2-mL tubes) and keep those at -20°C

2.1   Transposition Buffers and Reagents

Prepare the ATAC-RSB-Lysis and ATAC-RSB-Wash buffers immediately before use by adding the necessary detergents; keep on ice.

1. IGEPAL CA-630 detergent (Sigma Cat# 11332465001; supplied as a 10% solution)

2. Tween-20 detergent (Sigma Cat# 11332465001, supplied as a 10% solution; store at 4 °C)

3. Digitonin detergent (Promega Cat# G9441, supplied as a 2% solution in DMSO; store at -20°C))
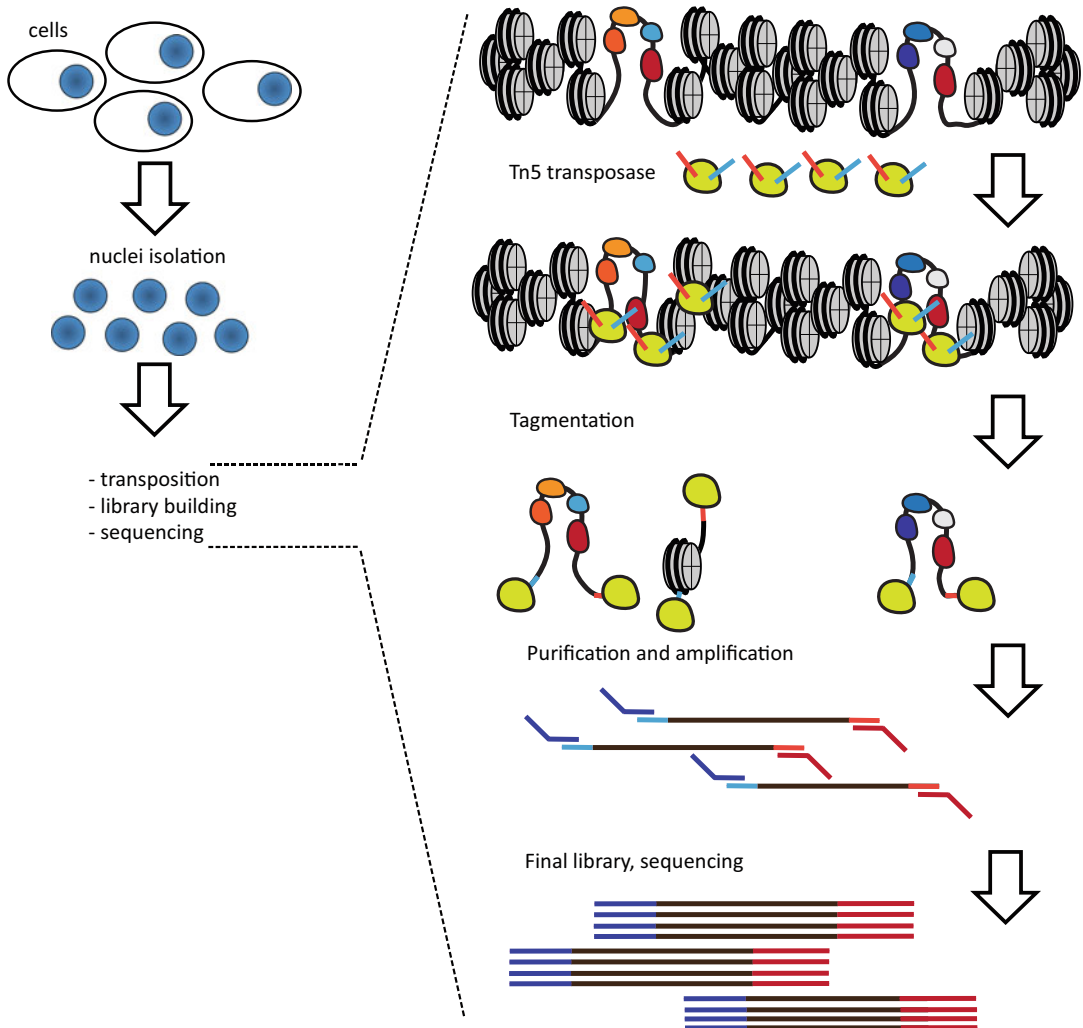
Fig. 1 Outline of the ATAC-seq assay. Nuclei are isolated from cells and chromatin is incubated with an active Tn5 transposase carrying PCR amplification adapter sequences. Tn5 preferentially inserts into accessible chromatin, such as that found at active regulatory elements. After transposition, DNA is purified and PCR amplification is carried out from the primer landing sites deposited by Tn5

4. ATAC-RSB buffer (master stock)

    10 mM Tris-HCl pH 7.4

    10 mM NaCl

    3 mM $MgCl_2$

5. ATAC-RSB-Lysis buffer

    10 mM Tris-HCl pH 7.4

    10 mM NaCl

    3 mM $MgCl_2$

    0.1% IGEPAL CA-630

        0.1% Tween-20

        0.01% Digitonin

    6. Lysis Wash Buffer (ATAC-RSB-wash)

        10 mM Tris-HCl pH 7.4

        10 mM NaCl

        3 mM $MgCl_2$

        0.1% Tween-20

    7. 2× TD buffer

        20 mM Tris-HCl pH 7.6

        10 mM $MgCl_2$

        20% Dimethyl Formamide

    8. Tn5 transposase (*see* **Note 1**)

2.2 Library Building, Sequencing, and Quality Evaluation

1. 200-µL PCR tubes
2. Sequencing primers/adapters (*see* **Note 2**)
3. NEBNext High-Fidelity 2× PCR Master Mix (NEB, Cat# M0541S)
4. Qubit fluorometer or equivalent
5. QuBit tubes
6. QuBit dsDNA HS Assay Kit
7. TapeStation (Agilent) or equivalent, e.g., BioAnalyzer (Agilent).
8. TapeStation D1000 tape and reagents (Agilent)
9. 10 mM dNTP mix
10. 25× SYBR Green (Thermo Fisher Cat# S7563. Supplied as 10,000X)
11. Phusion High-Fidelity DNA Polymerase (NEB, Cat# M0530L)

2.3 General Materials and Equipment

1. 1.5-mL microcentrifuge tubes, preferably low protein and DNA binding (*see* **Note 4**)
2. 2-mL, 15-mL, and 50-mL tubes
3. Incubator (37 °C), or a Thermomixer.
4. Tabletop centrifuge
5. Thermal cycler
6. MinElute PCR Purification Kit (Qiagen Cat# 28004/28006), Zymo DNA Clean and Concentrator Kit (Zymo Cat# D4013/D4014), or equivalent (*see* **Note 5**)
7. Nuclease-free H_2O

8. 1× PBS buffer solution

9. qPCR machine (Step One or equivalent)

3 Methods

The general outline of the ATAC-seq assay is shown in Fig. 1. Nuclei are isolated from cells, then a transposition reaction is carried out, DNA purified, and sequencing libraries prepared. Here we discuss the Omni-ATAC protocol [23] in its most widely applicable version, as it derives the optimal results in terms of reduced mitochondrial contamination (*see* **Note 6**) compared to other versions of the assay [20, 24]. The Omni-ATAC protocol works as described for the great majority of mammalian and insect cell lines, as well as for many other eukaryotic cells without cell walls. Different protocols need to be applied for nuclei isolation from other sources, such as tissues (*see* **Note 7**), plant cells (*see* **Note 8**), various small metazoan animals (*see* **Note 9**), yeast (*see* **Note 10**), and others.

It is also in principle possible to carry out ATAC-seq on cross-linked material but this generally produces suboptimal results and we advise against it (*see* **Note 11**).

3.1 Removal of Non-viable Cells (Optional)

The presence of non-viable cells can negatively affect the quality of final ATAC-seq libraries as dead cells generate a general background of dechromatinized DNA, decreasing the enrichment for open chromatin regions. Two strategies are usually used to address this problem:

1. If the fraction of dead cells is not too high (i.e., ∼5–15%), cells are treated with DNAse (200 U/mL) in culture media, usually for 30 min at 37°C. Cells are then washed thoroughly with 1×PBS to remove DNAse.

2. If the fraction of dead cells is higher, live cells can be separated from dead cells using a Ficoll gradient (Sigma Cat# GE17-1440-02), with the exact conditions varying depending on the cell type.

3.2 Preparation of Nuclei

Once the quality of the input cells has been ensured, the next step is to prepare nuclei and transpose them. The empirically determined optimum input number of cells for a species with a mammalian-sized genome is 50,000 diploid cells. Scale appropriately according to expected genome size and ploidy, and also change other parameters, such as centrifugation speeds, if necessary.

1. Centrifuge 50,000 viable cells at 500 g for 5 min at 4°C

2. Carefully aspirate the supernatant avoiding the pellet.

3. Add 50 μL of cold ATAC-RSB-Lysis Buffer and pipette up and down several times.

4. Incubate on ice for 3 min

5. Add 1 mL cold ATAC-RSB-Wash Buffer, and invert several times to mix well.

6. Centrifuge at 500 g for 5 min at 4°C

7. Carefully aspirate the supernatant as fully as possible while avoiding the pellet.

3.3 Transposition

Carry out transposition as follows:

1. Immediately resuspend the pellet in the transposase reaction mix:

   25 μL TD buffer

   2.5 μL Tn5

   22.5 μL nuclease-free H_20

2. Incubate at 37°C for 30 min in a Thermomixer at 1000 RPM.

3.4 DNA Purification

1. Immediately stop the reaction using 250 μL (i.e., 5×) of PB buffer (if using MinElute) or DNA Binding Buffer (if using Zymo; also *see* **Note 12**).

2. Purify samples following the kit instructions.

3. Elute with 10 μL of Elution Buffer.

3.5 PCR Amplification and Library Generation

Typically, a dual-indexing approach is used when amplifying ATAC-seq libraries. The general structure of an ATAC-seq library as well as the relevant adapter and primer sequences are shown in Fig. 2. *See* **Note 2** for further discussion.

1. Set up a PCR reaction as follows:

   10 μL eluted transposition reaction

   10 μL Nuclease-free H_2O

   2.5 μL of Adapter 1

   2.5 μL of Adapter 2

   25 μL NEBNext High-Fidelity 2× PCR Master Mix (*see* **Note 3**)

2. Optimization of PCR conditions, pre-amplification. Amplify DNA for 5 cycles as follows:

   72°C for 3 min

   98°C for 30 s
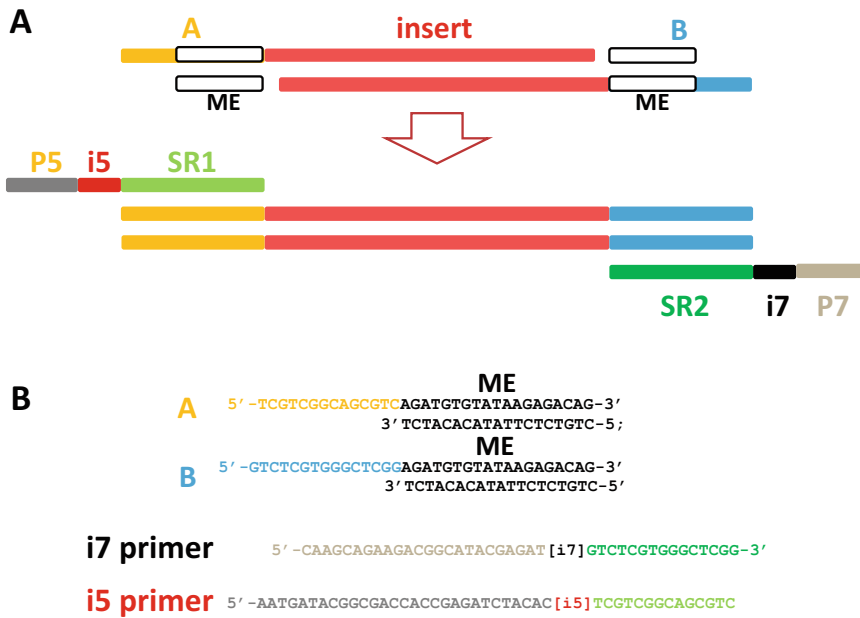
Fig. 2 Structure of an ATAC-seq library. (**a**) After transposition, an original DNA fragment is flanked by two Tn5 molecules with their adapter. Note that all three possible configurations—A-A, B-B, and A-B/B-A (where "A" and "B" indicate the two different adapters that Tn5 molecules used for transposition carry; these sequences have a common "ME" segment)—are produced, but only the A-B ones can be subsequently amplified and sequenced under conventional protocols. The A and B are used as landing sites for the PCR primers that add the i5 and i7 barcodes and the P5 and P7 sequences needed for Illumina sequencing. (**b**) Typical sequences of A and B adapters and of i5 and i7 PCR primers. The [i7] and [i5] sequences are typically 8-bp long and should be chosen appropriately so as to maximize the sequence distance between each pair of indexes

5 cycles of:

98 °C for 10 s

63 °C for 30 s

72 °C for 30 s

Hold at 4 °C

3. Determining additional cycles using qPCR. Use 5 μL of the pre-amplified reaction in a total qPCR reaction of 15 μL as follows:

3.76 μL nuclease-free H_2O

0.5 μL of Adapter 1

0.5 μL of Adapter 2

0.24 μL 25× SYBR Green (in DMSO)

5 μL NEBNext High-Fidelity 2× PCR Master Mix

5 μL pre-amplified sample

4. Determining additional cycles using qPCR. Run the qPCR reaction with the following settings in a qPCR machine:

   98°C for 30 s

   20 cycles of:

   98°C for 10 s

   63°C for 30 s

   72°C for 30 s

   Hold at 4°C

5. Assess the amplification profiles and determine the required number of additional cycles to amplify. Typical results are shown in Fig. 3.

6. Carry out final amplification by placing the remaining 45 μL in a thermocycler and running the following program:

   N_{add} cycles of:

   98°C for 10 s

   63°C for 30 s

   72°C for 30 s

   Hold at 4°C

   Where N_{add} is the number of additional cycles.

   In practice, we have found that 8–10 cycles are usually sufficient to amplify a standard mammalian ATAC library, and, if a very large number of samples are being processed at a time, the following reaction can be run:

7. Single-step PCR.

   72°C for 3 min

   98°C for 30 s

   8–10 cycles of:

   98°C for 10 s

   63°C for 30 s

   72°C for 30 s

   Hold at 4°C

8. Purify the amplified library following the same procedure used for purified the ATAC reaction.

a

b

Phix - 200pM - 100pM - 50pM - 25pM - 12.5pM - 6.25pM - 3.125pM - 1.56pM
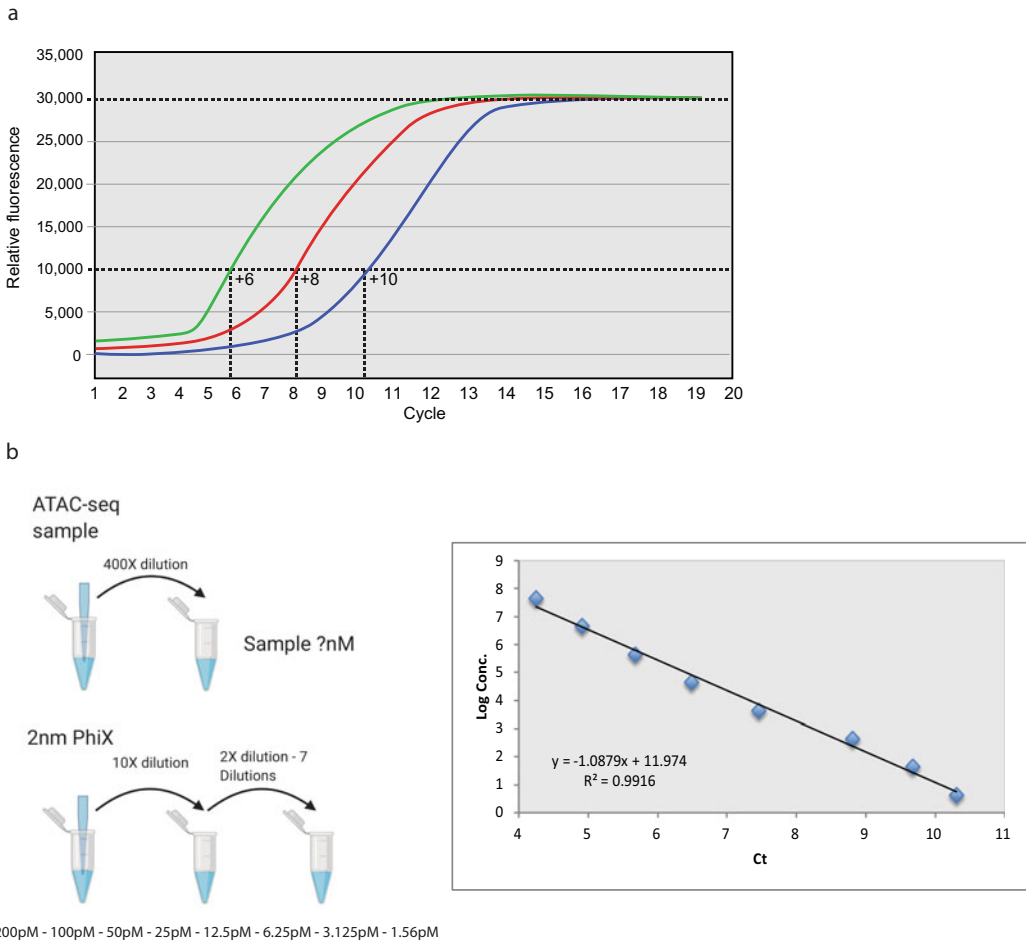
Fig. 3 Determination of additional PCR cycles (post pre-amplification) and library quantification using qPCR. (**a**) Determination of additional PCR cycles; qPCR is performed to determine the number of extra cycles to perform on the pre-amplified ATAC material without reaching saturation. To determine the number of extra cycles, find the number of cycles needed to reach 1/3 of the maximum relative fluorescence, and then carry out this number of additional PCR cycles. (**b**) Quantification of libraries; qPCR qualification is performed on diluted ATAC-seq libraries (400×) against a serial dilution of PhiX (200–1.56 pM). A standard curve is generated based on the PhiX dilutions and used to calculate the molarity of the ATAC-seq library

3.6 Library Quantification and Evaluation of Library Quality

Before libraries can be sequenced, they need to be properly quantified and their quality evaluated. There are two components to this process—first, evaluation of the insert distribution, and second, quantification.

1. Examination of library size distribution. This step can be carried out using a variety of instruments that are now available for this purpose, such as a TapeStation or a BioAnalyzer. In our practice we prefer to use a TapeStation (with the D1000 or HS D1000 kits) due to its ease of use, flexibility, and rapid turnaround time. Typical results are shown in Fig. 4. A successful
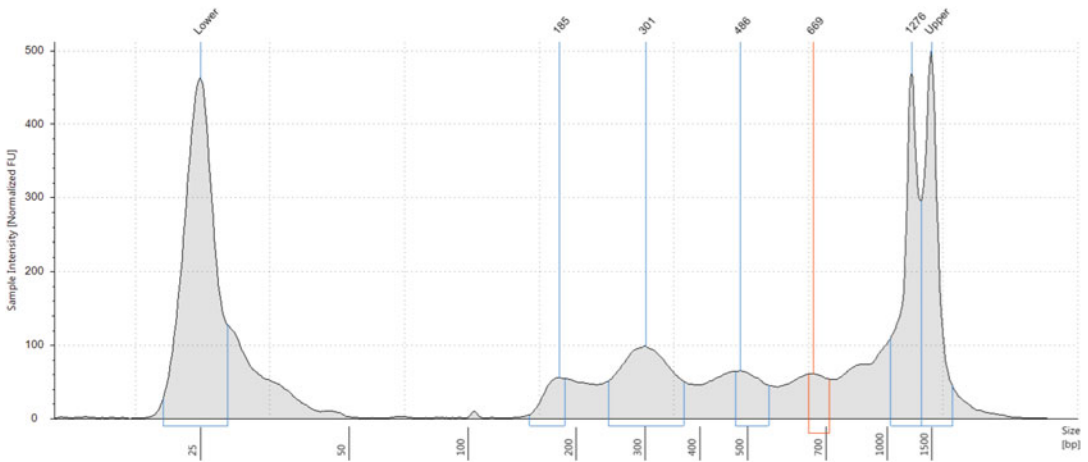
Fig. 4 Evaluation of ATAC-seq library size distribution. Shown is the fragment length distribution as evaluated using a TapeStation instrument and a D1000 TapeStation kit for an ATAC-seq library for the human GM12878 cell line. When a clear nucleosomal signature is observed, as in the example shown here, the library is most likely of high quality. Note that the nucleosomal signature can in some cases be obscured by the presence of high levels of mitochondrial contamination or some other source of highly accessible DNA (*see* **Note 13** for further discussion)

mammalian ATAC-seq library usually exhibits a clear nucleosomal signature (though the reverse is not always true; *see* **Note 13** for further discussion).

2. Quantification of library concentration. For most high-throughput sequencing applications, this step is standardly carried out using a Qubit fluorometer. This works well for most libraries as they exhibit a unimodal fragment length distribution, and the Qubit generally returns highly accurate and reliable measurements.

    However, ATAC libraries do not exhibit a unimodal fragment distribution and in fact often contain fragments of length higher than what can be sequenced on standard Illumina instruments. Thus the effective library concentration often differs from the apparent library concentration measured using Qubit (though Qubit measurements can still be used, with that caveat in mind, if no other information can be obtained)

3. Estimation of effective library concentration using qPCR.

    A standard curve is generated using Illumina PhiX standard (10nM) by first making a 50× dilution to 200 pM, from which additional seven serial 2× dilutions are made (to 100 pM, 50 pM, 25 pM, 12.5 pM, 6.25 pM, 3.125 pM, and 1.56 pM).

    Set up a 20 μL qPCR reactions as follows:

7.9 μL nuclease-free H_2O

5 μL ATAC-seq 400× diluted library or PhiX standards

4 μL Phusion HF Buffer

1 μL 25 μM i7 primer

1 μL 25 μM i5 primer

0.4 μL 10mM dNTP mix

0.5 μL 25× SYBR Green (in DMSO)

0.2 μL NEB Phusion HF

Run the qPCR reaction with the following settings in a qPCR machine:

98°C for 30 s

20 cycles of:

98°C for 10 s

63°C for 30 s

72°C for 30 s

Hold at 4°C

Create a standard curve based on the PhiX dilutions and estimate the true molarity of the qPCR library based on it.

Commercial kits such as NEBNext Library Quant Kit for Illumina or KAPA Library Quantification Kits can also be used, in a similar manner.

3.7 Sequencing

The protocol described here generates libraries designed to be sequenced on Illumina sequencers. A decision usually needs to be made regarding the format to be used when sequencing.

We strongly advise against sequencing ATAC-seq libraries in a single-end format, for two reasons. First, analysis of the fragment length distribution is an important part of the quality evaluation of ATAC-seq datasets, and this is only truly possible in paired-end format. Second, many analyses of ATAC-seq data (e.g., transcription factor footprinting) operate at the level of examining insertion points rather than read coverage; paired-end reads produce twice as many such insertion points for the same cost.

In practice, we have observed that ATAC-seq insert length distributions peaks around 50–60 bp (Fig. 5). Therefore it is most cost effective to sequence ATAC libraries in 2×36 bp or 2 × 50 bp formats (depending on the exact sequencer and kits available), as sequencing kits with more cycles are usually priced significantly higher.

However, for some applications (for example, if aiming to study the effects of sequence variation on chromatin accessibility), longer reads can provide important additional information.

Thus how exactly sequencing is to be executed is to be determined depending on the specific needs of the study being carried out.

A

Fragment length distribution

B

C
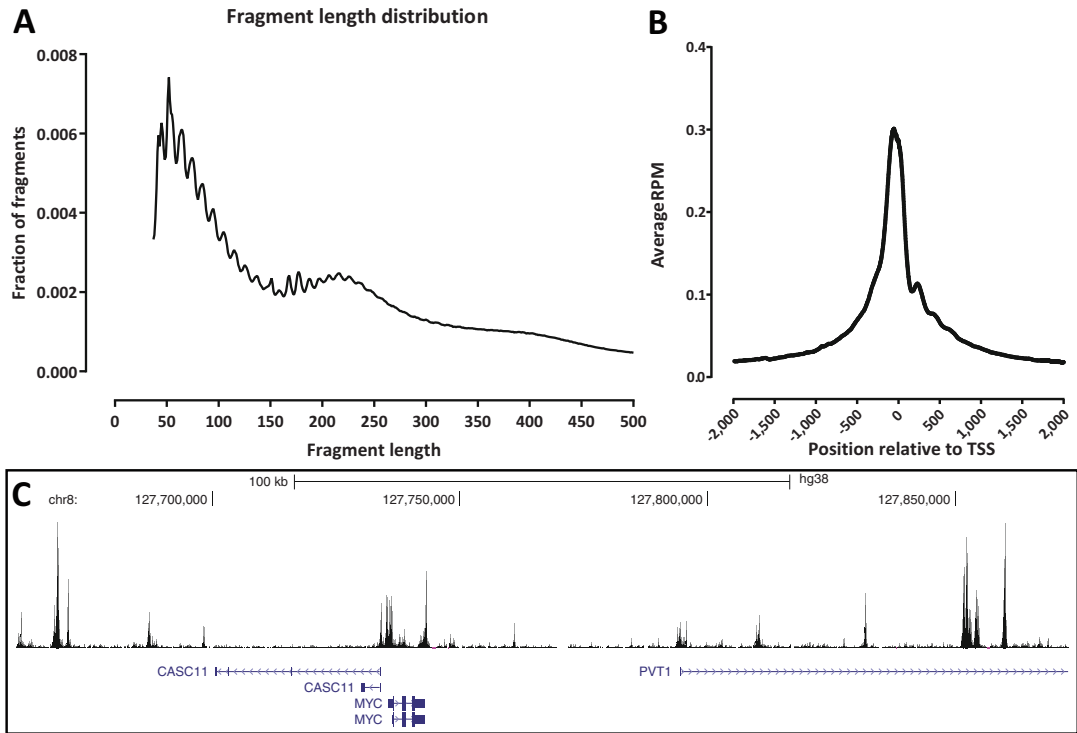
Fig. 5 Expected results from a successful ATAC-seq experiment. (**a**) Shown is the insert length distribution of a typical sequenced mammalian ATAC-seq library, showing a prominent subnucleosomal peak, as well as a mononucleosomal and a less pronounced dinuleosomal peak. (**b**) Aggregate ATAC-seq signal profile around transcription start sites (TSSs). (**c**) ATAC-seq profile in a 212-kb neighborhood around the human *MYC* gene. The ENCODE Consortium [34] keratinocyte dataset with accession ID ENCSR798IJQ was used for this example

4   Expected Results

After sequencing, reads mapped to the reference genome, and several quality evaluation metrics are considered before proceeding with downstream analysis. A typical ATAC-seq library exhibits a nucleosomal signature as shown in Fig. 5a. Enhancer, promote and insulator regions should be strongly enriched relative to the rest of the genome (Fig. 5c shows the accessibility profile in the neighborhood of the *MYC* gene, which highlights a number of candidate distal regulatory elements). A quick way to evaluate the degree of enrichment is to examine aggregate plots of ATAC-seq signal around annotated TSSs, as shown in Fig. 5b, which can also be further formalized as a TSS ratio score, which is calculated by dividing the average number of fragments within ±100 bp of the TSS to the sum of the average number of fragments within the two 100-bp windows at the points $+2$ and -2 kb away from the TSS. The advantage of this metric is that it is independent of peak calling

or sequencing depth; its disadvantage is that it is annotation-dependent and well calibrated only for the human and mouse genomes. In the latter two cases, good ATAC-seq libraries usually exhibit TSS rations ≥8.

5   Notes

1. The Tn5 transposase can be obtained as part of the various Nextera DNA Library Prep kits offered by Illumina, and also from several other commercial vendors. It can also be made by individual laboratories following previously published protocols [25]. The latter approach is, although laborious, the most cost effective, especially for large-scale projects. If homemade Tn5 is used, its activity should ideally be well characterized relative to standard enzymatic formulations.

2. PCR and indexing primers/adapters supplied with the Nextera DNA Library Prep kits offered by Illumina can be used. Alternatively, or if a larger number of indexing sequences is needed, custom-designed, and synthesized oligos can also be used with equivalent success. The structure of an ATAC-seq library with the relevant sequences is shown in Fig. 2. The i7 primer sequence is:

   ```
   5'-CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG-3'
   ```

   The i5 sequence is:

   ```
   5'AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC-3'
   ```

   Where [i7] and [i5] are the index sequences (typically 8-bp long). Dissolve and dilute to 25 μM.

3. The initial extension is very important when amplifying transposed DNA as it is needed to fill in the gap left from the transposition itself (*see* Fig. 2) and allow PCR primers to land in subsequent amplification cycles. For this reason, it is not recommended to use hot-start polymerase mixes, in which the polymerase is only activated by exposing to denaturation temperatures.

4. Low-binding tubes are preferable, though not absolutely required, as a low number of cells (only ∼50,000) is usually used as input to an ATAC reaction.

5. The MinElute kit can be replaced with other DNA purification kits; for example, we have also had equivalent success using the DNA Clean & Concentrator from Zymo. The important

variables regarding the DNA isolation procedure after transposition are the efficiency of recovery and the lower size limit of the recovered fragments. The insert length distribution of most ATAC-seq libraries peaks around 50–60 bp, i.e., even including Tn5 adapter sequences, many of the informative fragments are shorter than 90–100 bp, and should ideally be preserved during the DNA purification procedure.

6. Early versions of the ATAC-seq protocol [20] exhibited very high proportions of reads originating from the mitochondrial genome, often exceeding 80% of the total. This is due to the fact that the mitochondrial genome is not packaged by nucleosomes, and is therefore highly accessible to transposase insertion. Decreasing the fraction of mitochondria has been a key part of the improvement of the ATAC-seq protocol in its the currently used variants relative to the original version, and has been achieved thanks to the addition of the combination of digitonin, Tween-20 and IgePAL detergents during the cell lysis and nuclei preparation step. As a result ATAC-seq libraries generated using modern protocols frequently contain as little as ≤5% of reads for many cell types.

   We do note, however, that there are cells that simply contain an extremely high number of mitochondria due to very high levels of metabolic activity (e.g., some cancer cell lines), and even with the optimized Omni-ATAC protocol mitochondrial fractions are still quite high for them. These are special cases though.

   We also note that high levels of mitochondrial contamination do not necessarily correspond to poor-quality ATAC-seq datasets in terms of signal-to-noise ratios in the nuclear genome. We have found no inverse correlation between the fraction of reads mapping to the mitochondrial genome and the levels of enrichment for open chromatin regions. The key benefit of eliminating mitochondrial reads is to reduce sequencing costs as fewer overall reads need to be sequenced to achieve the necessary coverage over the nuclear genome.

7. Nuclei isolation from tissues, especially when frozen, is a multistep procedure involving tissue homogenization by douncing followed by density gradient centrifugation. The reader is referred to [23] for more details.

8. Plant cells are a challenging system to isolate nuclei from because of their thick cellulose cell walls. Nuclei are isolated by grinding tissue material in liquid nitrogen [26–28] and then sorting nuclei by sucrose sedimentation or by FACS, if the cell type of interest has been labeled accordingly, e.g., using the INTACT approach [29].

9. It is also often necessary to carry out homogenization followed by nuclei isolation when working with whole animals, e.g., *C. elegans* [30], with the exact protocol optimized according to the specifics of the organism being studied.

10. Yeast (and fungal cells in general) have thick cell walls comprised of polysaccharides, lipids, and chitin in various proportions. They present a barrier to the access of Tn5 to the nucleus, thus ATAC-seq protocols tailored to such cells involve treatment with zymolyase or chitinase enzymes [31], with the exact details varying depending on the species studied.

11. It is in principle possible to carry out ATAC-seq experiments and obtain enriched libraries from crosslinked sources. This has in fact been how a number of scATAC-seq studies have been executed in recent years [32, 33]. However, ATAC-seq libraries generated from crosslinked cells are generally suboptimal, with lower signal-to-noise ratio than standard ATAC-seq datasets, and they also tend to display a pronounced loss of subnucleosomal fragments compared to the standard protocol. We thus advice against using fixed material for ATAC-seq except for special circumstances where this is the only available option.

12. This is also a possible stopping point if necessary. DNA can be stored in PB buffer at -20°C before proceeding with subsequent clean up steps at a later time.

13. If a clear nucleosomal signature is observed in a TapeStation profile that in almost all cases indicates a successful ATAC-seq experiment. The inverse is not always true, as the fragment distribution can be dominated by the presence of large amounts of strongly accessible DNA in the original sample. For example, libraries with high levels of mitochondrial contamination often exhibit an obscured nucleosomal signature, and so do nearly all yeast libraries (in the latter case it is because yeast genomes contain a large number of ribosomal DNA copies, which are nearly nucleosome-free when being actively transcribed, and often comprise half or even more of yeast ATAC-seq libraries [13]). As discussed in **Note 6**, high levels of mitochondrial contamination are undesirable in terms of the efficient utilization of sequencing resources but do not necessarily result in poor-quality datasets in the nuclear genome.

Acknowledgements

References

1. Luger K, Mäder AW, Richmond RK et al. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389: 251–260

2. Wu C (1980) The 5′ ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. Nature 286(5776):854–860

3. Keene MA, Corces V, Lowenhaupt K et al. (1981) DNase I hypersensitive sites in Drosophila chromatin occur at the 5′ ends of regions of transcription. Proc Natl Acad Sci U S A 78:143–146

4. McGhee JD, Wood WI, Dolan M et al. (1981) A 200 base pair region at the 5′ end of the chicken adult β-globin gene is accessible to nuclease digestion. Cell 27:45–55

5. Dorschner MO, Hawrylycz M, Humbert R et al. (2004) High-throughput localization of functional elements by quantitative chromatin profiling. Nat Methods 1:219–225

6. Sabo PJ, Humbert R, Hawrylycz M et al. (2004) Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. Proc Natl Acad Sci U S A 101:4537–4542

7. Sabo PJ, Kuehn MS, Thurman R et al. (2006) Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. Nat Methods 3:511–518

8. Crawford GE, Holt IE, Whittle J et al. (2006) Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). Genome Res 16:123–131

9. Boyle AP, Davis S, Shulha HP et al. (2008) High-resolution mapping and characterization of open chromatin across the genome. Cell 132:311–322

10. Thurman RE, Rynes E, Humbert R et al. (2012) The accessible chromatin landscape of the human genome. Nature 489:75–82

11. Kelly TK, Liu Y, Lay FD et al. (2012) Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res 22:2497–2506

12. Krebs AR, Imanci D, Hoerner L, Gaidatzis D et al. (2017) Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. Mol Cell 67: 411–422.e4

13. Shipony Z, Marinov GK, Swaffer MP et al. (2018) Long-range single-molecule mapping of chromatin accessibility in eukaryotes. bioRxiv 504662

14. Wang Y, Wang A, Liu Z et al. (2019) Single-molecule long-read sequencing reveals the chromatin basis of gene expression. Genome Res 29:1329–1342

15. Aughey GN, Estacio Gomez A, Thomson J et al. (2018) CATaDa reveals global remodelling of chromatin accessibility during stem cell differentiation in vivo. Elife 7:pii: e32341

16. Chereji RV, Eriksson PR, Ocampo J, Clark DJ (2019) DNA accessibility is not the primary determinant of chromatin-mediated gene regulation bioRxiv 639971

17. Ponnaluri VKC, Zhang G, Estéve PO et al. (2017) NicE-seq: high resolution open chromatin profiling. Genome Biol 18(1):122

18. Umeyama T, Ito T (2017) DMS-seq for in vivo genome-wide mapping of protein-DNA interactions and nucleosome centers. Cell Rep 21: 289–300

19. Timms RT, Tchasovnikarova IA, Lehner PJ (2019) Differential viral accessibility (DIVA) identifies alterations in chromatin architecture through large-scale mapping of lentiviral integration sites. Nat Protoc 14:153–170

20. Buenrostro JD, Giresi PG, Zaba LC et al. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10:1213–1218

21. Buenrostro JD, Wu B, Litzenburger UM et al. (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523:486–490

22. Cusanovich DA, Daza R, Adey A et al. (2015) Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science 348:910–914

23. Corces MR, Trevino AE, Hamilton EG et al. (2017) An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14:959–962

24. Corces MR, Buenrostro JD, Wu B et al. (2016) Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet 48:1193–1203

25. Picelli S, Björklund AK, Reinius B et al. (2014) Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res 24:2033–2040

26. Lu Z, Hofmeister BT, Vollmers C et al. (2017) Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Res 45:e41

27. Maher KA, Bajic M, Kajala K et al. (2018) Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. Plant Cell 30:15–36

28. Bajic M, Maher KA, Deal RB (2018) Identification of open chromatin regions in plant genomes using ATAC-seq. Methods Mol Biol 1675:183–201

29. Deal RB, Henikoff S (2010) A simple method for gene expression and chromatin profiling of individual cell types within a tissue. Dev Cell 18:1030–1040

30. Daugherty AC, Yeo RW, Buenrostro JD et al. (2017) Chromatin accessibility dynamics reveal novel functional enhancers in C. elegans. Genome Res 27:2096–2107

31. Schep AN, Buenrostro JD, Denny SK et al. (2015) Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Res 25:1757–1770

32. Cusanovich DA, Reddington JP, Garfield DA et al. (2018) The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature 555:538–542

33. Cao J, Cusanovich DA, Ramani V et al. (2018) Joint profiling of chromatin accessibility and gene expression in thousands of single cells. Science 361:1380–1385

34. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74