ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide

Jason D. Buenrostro,[1,2] Beijing Wu,[1] Howard Y. Chang,[2] and William J. Greenleaf[1]

[1]Department of Genetics, Stanford University School of Medicine, Stanford, California
[2]Program in Epithelial Biology and the Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California

This unit describes Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq), a method for mapping chromatin accessibility genome-wide. This method probes DNA accessibility with hyperactive Tn5 transposase, which inserts sequencing adapters into accessible regions of chromatin. Sequencing reads can then be used to infer regions of increased accessibility, as well as to map regions of transcription-factor binding and nucleosome position. The method is a fast and sensitive alternative to DNase-seq for assaying chromatin accessibility genome-wide, or to MNase-seq for assaying nucleosome positions in accessible regions of the genome.

Keywords: ATAC-seq • transposase • chromatin accessibility

How to cite this article:
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.* 109:21.29.1-21.29.9. doi: 10.1002/0471142727.mb2129s109

BASIC PROTOCOL

In a human cell, approximately two meters of DNA are packed within a five-micron nucleus—a spectacular topological challenge solved by the cell through the hierarchical folding of DNA around histone proteins to form nucleosomes, and the compaction of nucleosomes into chromatin (Kornberg, 1974). This hierarchical packaging sequesters inactive genomic regions and leaves biologically active regions—be they promoters, enhancers, or other regulatory elements—accessible to transcription machinery (Gross and Garrard, 1988; Bell et al., 2011). Atop this landscape of physical compaction operates a dynamic epigenetic code that includes DNA methylation, nucleosome positioning, histone composition, and modification, as well as transcription factors, chromatin remodelers, and non-coding RNAs (Kouzarides, 2007; Rinn and Chang, 2012). Cellular phenotypes are substantially governed by epigenetic mechanisms that manipulate the composition, compaction, and nucleoprotein structure of chromatin (Chen and Dent, 2014).

Methods such as MNase-seq (*UNIT 21.1*), ChIP-seq (Landt et al., 2012), and DNase-seq (Song and Crawford, 2010) in particular have proven to be information-rich, genome-wide analysis methods for understanding this epigenetic structure, providing information on transcription factor binding, the positions of modified and canonical nucleosomes, and chromatin accessibility at regulatory elements such as promoters, enhancers, and insulators (Valouev et al., 2011; Consortium, 2012; Thurman et al., 2012). However, current methods for assaying chromatin structure and composition often require tens to hundreds of millions of cells as input material, averaging out heterogeneity

Current Protocols in Molecular Biology 21.29.1-21.29.9, January 2015
Published online January 2015 in Wiley Online Library (wileyonlinelibrary.com).
doi: 10.1002/0471142727.mb2129s109

in cellular populations. In many cases, rare and important cellular subtypes cannot be acquired in amounts sufficient for genome-wide chromatin analyses. The assay of transposase-accessible chromatin (ATAC-seq; Buenrostro et al., 2013) uses hyperactive Tn5 transposase (Goryshin and Reznikoff, 1998; Adey et al., 2010) to simultaneously cut and ligate adapters for high-throughput sequencing at regions of increased accessibility. Genome-wide mapping of insertion ends by high-throughput sequencing allows for multidimensional assays of the regulatory landscape of chromatin with a relatively simple protocol that can be carried out in hours for a standard sample size of 50,000 cells.

ATAC-seq data sets can also be bioinformatically separated into reads that are shorter than the canonical length generally protected by a nucleosome and reads consistent with the approximate length of DNA protected by a nucleosome. These different populations of reads provide information about the positions of nucleosomes, as well as nucleosome-free regions. In this way, paired-end ATAC-seq data provides MNase-seq-like data sets for nucleosomes at relatively accessible regions of the genome.

ATAC-seq is compatible with a number of methods for cell separation and isolation, including cell sorting as well as disruption of intact tissues to cellular suspensions. The method has also worked well across many cell types and species. However, the following protocol has been optimized for human lymphoblastoid cells, and provides guidance on modifications to the method that might allow adaptation to different cell types. In general, the method is separated into three components: cell lysis, transposition, and amplification. Cross-linking generally reduces library creation efficiency, and therefore we recommend starting with fresh, unfixed cells for maximum sensitivity of the methodology.

Materials

50,000 cells in a single-cell suspension
Phosphate buffered saline (PBS; *APPENDIX 2*)
Lysis buffer (see recipe)
Molecular biology-grade IGEPAL CA-630 (Sigma-Aldrich, cat. no. I8896)
TD (2× reaction buffer from Nextera kit; Illumina, cat. no. FC-121-1030)
TDE1 (Nextera Tn5 Transposase from Nextera kit; Illumina, cat. no. FC-121-1030)
Nuclease-free H_2O (available from various molecular biology suppliers)
Qiagen MinElute PCR Purification Kit
25 μM PCR Primer 1 [custom-synthesized by Integrated DNA Technologies (IDT); sequences provided in Buenrostro et al. (2013)]
25 μM Barcoded PCR Primer 2 [custom-synthesized by Integrated DNA Technologies (IDT); sequences provided in Buenrostro et al. (2013)]
NEBNext High-Fidelity 2× PCR Master Mix (New England Biolabs, cat. no. M0541)
100× SYBR Green I (Invitrogen, cat. no. S-7563)
5% TBE polyacrylamide gel (see *UNIT 10.2*; optional)
100-bp DNA ladder (New England Biolabs; optional)

Refrigerated centrifuge
0.2-ml PCR tubes
PCR thermal cycler
qPCR instrument (Applied Biosystems StepOnePlus Real-Time PCR System; cat. no. 4376600)
Typhoon TRIO Variable Mode Imager (Amersham Biosciences; optional)
Bioanalyzer High-Sensitivity DNA Analysis kit (Agilent; optional)

Additional reagents and equipment for counting cells (*APPENDIX 3F*), PCR (*UNIT 15.1* and other units in Chapter 15), and polyacrylamide gel electrophoresis (*UNIT 10.2*; optional)

Cell preparation

1. Harvest and count cells (counting protocol to be defined by the user; see *APPENDIX 3F*).

   Cells should be intact and in a homogenous, single-cell suspension.

2. Centrifuge 50,000 cells 5 min at 500 × *g*, 4°C.

   The number of cells at this step is crucial, as the transposase-to-cell ratio determines the distribution of DNA fragments generated. See Critical Parameters.

3. Remove and discard supernatant. Wash cells once with 50 µl of cold PBS buffer. Centrifuge 5 min at 500 × *g*, 4°C.

4. Remove and discard supernatant. Gently pipet up and down to resuspend the cell pellet in 50 µl of cold lysis buffer. Centrifuge immediately for 10 min at 500 × *g*, 4°C.

   This step affords lysis of cells with nonionic detergent and generates a crude nuclei preparation.

5. Discard the supernatant, and immediately continue to transposition reaction.

Transposition reaction and purification

6. Make sure the cell pellet is placed on ice.

7. To make the transposition reaction mix, combine the following:

   25 µl TD (2× reaction buffer from Nextera kit)
   2.5 µl TDE1 (Nextera Tn5 Transposase from Nextera kit)
   22.5 µl nuclease-free H_2O.

8. Resuspend nuclei pellet (from step 5) in the transposition reaction mix.

9. Incubate the transposition reaction at 37°C for 30 min.

   Gentle mixing may increase fragment yield.

10. Immediately following transposition, purify using a Qiagen MinElute PCR Purification Kit.

11. Elute transposed DNA in 10 µl elution buffer (Buffer EB from the MinElute kit consisting of 10 mM Tris·Cl, pH 8).

12. Store purified DNA at −20°C if necessary.

    This is a convenient stopping point. Please note that these DNA fragments are not PCR amplifiable if melted at this point.

PCR amplification

13. To amplify transposed DNA fragments, combine the following in a 0.2 ml PCR tube:

    10 µl transposed DNA
    10 µl nuclease-free H_2O
    2.5 µl 25 µM PCR Primer 1
    2.5 µl 25 µM Barcoded PCR Primer 2
    25 µl NEBNext High-Fidelity 2× PCR Master Mix.

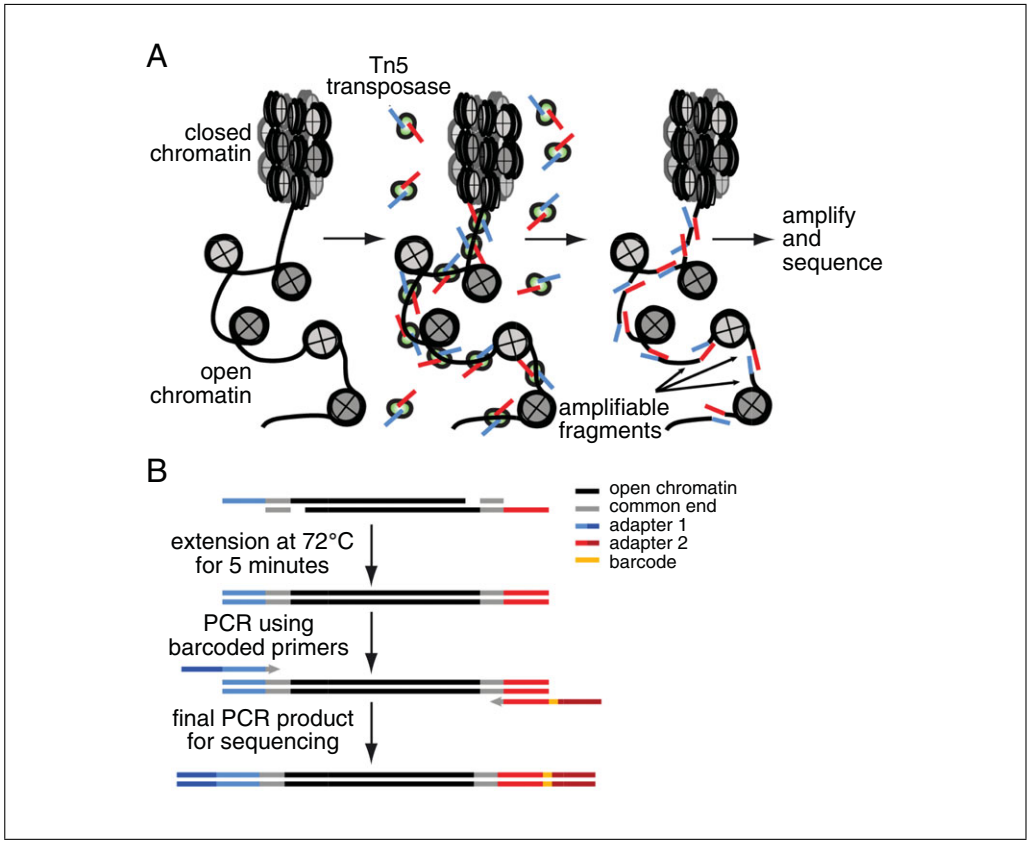

Figure 21.29.1 (**A**) Library preparation schematic. (**B**) Transposition results in fragmented DNA. Prior to amplification, adapters have to be completed with a 72°C extension step. During the subsequent PCR, additional sequence is incorporated into the adapters, which include common sequencing ends and a sequencing barcode.

Primers and PCR conditions are optimized for amplifying large-molecular-weight fragments from low-input material, therefore, PCR reagents provided by Illumina are not recommended. Primers are synthesized by Integrated DNA Technologies (IDT) with no additional modifications. A complete list of primers is available in Buenrostro et al. (2013). Care should be taken to ensure that samples are barcoded appropriately for subsequent pooling and sequencing.

14. Thermal cycle as follows:

1 cycle:	5 min	72°C
	30 sec	98°C
5 cycles:	10 sec	98°C
	30 sec	63°C
	1 min	72°C.

This first 5-min extension at 72°C is critical to allow extension of both ends of the primer after transposition, thereby generating amplifiable fragments (see Fig. 21.29.1). This short pre-amplification step ensures that downstream quantitative PCR (qPCR) quantification will not change the complexity of the original library.

See UNIT 15.1 *for additional detail on PCR.*

Perform qPCR

To reduce GC and size bias in PCR, the appropriate number of PCR cycles (*N*) is determined using qPCR, allowing us to stop amplification prior to saturation.

15. To run a qPCR side reaction, combine the following in qPCR compatible consumables:

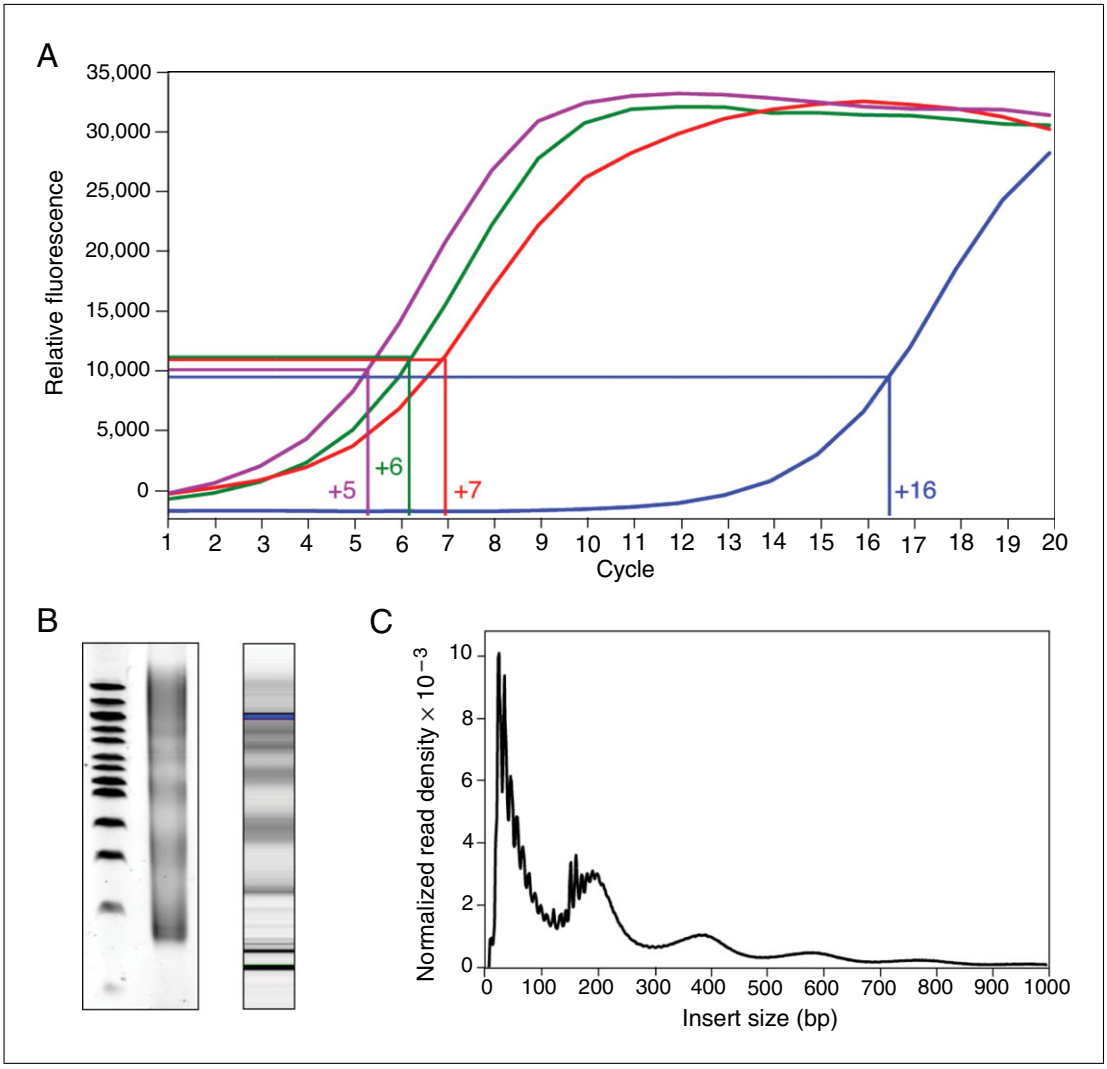

Figure 21.29.2 (**A**) Representative amplification plot demonstrating the correct number of additional cycles to perform for four ATAC-seq libraries. (**B**) Fragment sizes for amplified ATAC-seq libraries, determined by gel electrophoresis (left) and Bioanalyzer (right). Bioanalyzer image contrast has been enhanced from its original image for clarity. (**C**) Insert sizes determined by high-throughput sequencing. Adapters are an additional 124 bp and are not included when measuring fragment size in this panel.

5 µl of previously PCR-amplified DNA (from step 14)
4.41 µl nuclease-free H_2O
0.25 µl 25 µM Custom Nextera PCR Primer 1
0.25 µl 25 µM Custom Nextera PCR Primer 2
0.09 µl 100× SYBR Green I
5 µl NEBNext High-Fidelity 2× PCR Master Mix.

16. Using a qPCR instrument, cycle as follows:

1 cycle:	30 sec	98°C
20 cycles:	10 sec	98°C
	30 sec	63°C
	1 min	72°C.

17. To calculate the additional number of cycles needed, plot linear Rn versus cycle and determine the cycle number that corresponds to `one-third` of the maximum fluorescent intensity (see Fig. 21.29.2A).

The purpose of this qPCR step is to generate libraries that are minimally PCR amplified. Most PCR bias comes from later PCR cycles that occur during limited reagent concentrations. This determination of the optimal number of cycles to amplify the library reduces artifacts associated with saturation PCR of complex libraries.

18. Run the remaining 45 μl PCR reaction to the cycle number determined by qPCR. Cycle as follows:

1 cycle:	30 sec	98°C
N cycles:	10 sec	98°C
	30 sec	63°C
	1 min	72°C.

Cycle for an additional N cycles, where N is determined using qPCR.

Optional library quality control using gel electrophoresis

For assessing quality, amplified libraries can be visualized using gel electrophoresis prior to PCR purification. The low concentration of the amplified materials requires a 5% TBE polyacrylamide gel optimized for sensitivity. We find that adding 0.6× SYBR Green I to libraries provides excellent signal-to-noise without the need for post-staining. We routinely load 15 ng of 100-bp DNA ladder from NEB with 0.6× SYBR Green I. Although in principle any instrument containing a blue-light source or imaging system equipped with a laser that emits at 488 nm can be used to visualize DNA stained with SYBR Green I dye, we use the Typhoon TRIO Variable Mode Imager from Amersham Biosciences for improved sensitivity. Images are best obtained by digitizing at 100-μm pixel size resolution with a 520-nm band-pass emission filter to screen out reflected and scattered excitation light and background fluorescence. Figure 21.29.2B (left) is a representative example of a high-quality library.

19a. Purify amplified library using Qiagen MinElute PCR Purification Kit. Elute the purified library in 20 μl elution buffer (Buffer EB from the MinElute kit consisting of 10 mM Tris·Cl, pH 8). Ensure that the column is dry prior to adding elution buffer to avoid ethanol contamination in the final library.

The concentration of DNA eluted from the column ought to be approximately 30 nM; however, 5-fold variation is possible and not detrimental.

Alternative to gel electrophoresis: library quality control with Bioanalyzer

19b. Assess the quality of purified libraries using a Bioanalyzer High-Sensitivity DNA Analysis kit (Agilent).

See Figure 21.29.2B for an example. We have seen that a preponderance of high-molecular-weight DNA (>1000 bp) can create misleading and inaccurate Bioanalyzer results.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Lysis buffer

10 mM Tris·Cl, pH 7.4 (*APPENDIX 2A*)
10 mM NaCl
3 mM $MgCl_2$
0.1% (v/v) Igepal CA-630
Store up to 1 week at 4°C

COMMENTARY

Background Information

The ATAC-seq methodology relies on library construction using the hyperactive transposase Tn5. Tn5 is a prokaryotic transposase, which endogenously functions through the "cut and paste" mechanism, requiring sequence-specific excision of a locus flanked by 19-base-pair inverted repeats. The Nextera DNA Sample Preparation kit provides the Tn5 transposase reagent loaded with sequencing adapters, creating an active dimeric transposome complex (Adey et al., 2010). The Nextera version of the transposase harbors specific point mutants to the Tn5 backbone, which significantly increases activity (Goryshin and Reznikoff, 1998; Reznikoff, 2008). This transposase preferentially inserts sequencing adapters into unprotected regions of DNA, thereby acting as a probe for measuring chromatin accessibility genome-wide.

Critical Parameters

Cell number

ATAC-seq is often robust to relatively minor variations in cell number (roughly 25,000 to 75,000, for example). In general, using too few cells causes over-digestion of chromatin and appears to create a larger fraction of reads that map to inaccessible regions of the genome (i.e., noise); using too many cells causes under-digestion and creates high-molecular-weight fragments, which may be difficult to sequence. However, the total number of cells also sets the fundamental diversity of the sequencing library (i.e., the number of unique DNA fragments). To obtain more complex libraries, larger reaction volumes or serial reactions with the same reaction volume can be carried out.

Cell collection

Methods for cell collection vary and may need to be optimized for ATAC-seq. We find that fixatives can reduce transposition frequency and are not recommended. We also note that methods involving mechanical shearing can significantly reduce signal-to-noise. Methods that produce intact cells in a homogenous single-cell suspension tend to produce the best data sets.

PCR and fragmentation distribution

The PCR methods described are designed to reduce the effect of size and GC bias from the library construction process. These methods are particularly useful for ATAC-seq because of the highly diverse fragment sizes. We find that samples containing an excess of large fragments (>1 kb) are relatively hard to quantify and result in reduced clustering efficiencies when sequencing. If a library is enriched for long fragments, library fragmentation can be optimized using more or fewer cells per reaction. Alternatively, size selection prior to sequencing might be carried out to eliminate these potentially confounding long fragments. Although it is common to size-select a narrow interval for sequencing, we recommend excising a large fragment size window of 100 to 800 bp to maintain high library complexity and enable the richness of the inferences that can be extracted from the full fragment size distribution. The described PCR method has the additional benefit of serving as an estimate for library complexity. We find that if >6 additional cycles are needed (>11 total cycles), library complexity becomes a concern. Library complexity can be improved by optimizing the input cell number or by making libraries of technical replicates.

Library quantitation

We use qPCR-based methods to quantify our ATAC-seq libraries. We have found that other methods, such as Qubit, can potentially give misleading and inaccurate results due to variation within the fragment size distribution. We recommend quantifying libraries using the KAPA Library Quant Kit for Illumina Sequencing Platforms (KAPA Biosystems). Alternatively, integrated Bioanalyzer traces can be used to approximate library concentration.

Sequencing

ATAC-seq and the Nextera workflow are designed for sequencing using Illumina high-throughput sequencing instruments. When sequencing ATAC-seq libraries, Nextera-based sequencing primers and reagents must be used. Depending on the sequencing instrument, this may require modification of the standard sequencing workflow. For nucleosome mapping, paired-end sequencing is preferred. Paired-end 50-cycle reads generally provide accurate alignments with reasonable costs. For inferring differences in open chromatin within human samples, we generally use >50,000,000 mapped reads, and for transcription factor foot-printing, we use >200,000,000 mapped reads (Neph et al., 2012). Data yield can sometimes be impacted by a large fraction of mitochondrial reads. We note that the mitochondrial read fraction can vary between 10% and

>50%, depending on the cell type. For troubleshooting cluster densities and sequencing quality, refer to Illumina's technical support.

Troubleshooting

Assuming that the cells of interest are healthy and intact prior to beginning the procedure, the biggest source of failure comes from variations in cell number. In general, the addition of too many cells leads to "under transposition" (with a majority of large fragments), while the addition of too few cells leads to "over transposition" and a preponderance of short fragments on the gel, with the possible elimination of any banding pattern. When applying this protocol to nonhuman cells, we see large differences in the number of cells required, depending on the species; however, variation in outcome can also come from different cell types. If results using the standard protocol are not ideal, an efficient method for optimization is to scale the reaction down 10× and to optimize the lysis methods as well as the cell number, starting from 5000 cells per reaction. The nucleosomal banding pattern shown in Figure 21.29.2 is correlated with high-quality libraries.

Anticipated Results

After sequencing human libraries, we often find that approximately half of reads are of sub-nucleosomal length (less than approximately 150 bp) and approximately half of the reads are longer than this length (Fig. 21.29.2C). In general, we find excellent agreement with DNase-seq methods and enrichment for regions of accessible chromatin, with ∼20% of reads concentrated in ∼2% of the genome. We estimate that a single 50-μl reaction volume can generate approximately 50,000,000 to 100,000,000 unique amplifiable fragments.

Time Considerations

ATAC-seq was designed with speed in mind; we have optimized the protocol to require a total time of 3 hr. The above protocol for nuclei extraction requires 30 min. Transposition and purification require 45 min. PCR and purification require 1 hr and 45 min. Methods for QC and library quantitation can vary. A natural stopping point exists after transposition and cleanup, at which point DNA fragments can be frozen indefinitely prior to the PCR amplification.

Acknowledgements

J.D.B. acknowledges support from a National Science Foundation graduate research fellowship and NIH training grant T32HG000044. H.Y.C. acknowledges support as an Early Career Scientist of the Howard Hughes Medical Institute. This work is supported by NIH (H.Y.C., W.J.G., and J.D.B.), including P50-HG007735, RC4NS073015, U01DK089532, and U19AI057229, Scleroderma Research Foundation (H.Y.C.), and California Institute for Regenerative Medicine (H.Y.C.). W.J.G acknowledges support as a Rita Allen Foundation Young Scholar.

Competing Financial Interests

Stanford University has filed a provisional patent application on the methods described, and J.D.B., H.Y.C., and W.J.G. are named as inventors.

Literature Cited

Adey, A.A., Morrison, H.G., Asan, Xun, X., Kitzman, J.O., Turner, E.H., Stackhouse, B., MacKenzie, A.P., Caruccio, N.C., Zhang, X., and Shendure, J. 2010. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* 11:R119.

Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. 2011. Determinants and dynamics of genome accessibility. *Nat. Rev. Genet.* 12:554-564.

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

Chen, T. and Dent, S.Y.R. 2014. Chromatin modifiers and remodellers: Regulators of cellular differentiation. *Nat. Rev. Genet.* 15:93-106.

Consortium, T.E.P. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57-74.

Goryshin, I.Y. and Reznikoff, W.S. 1998. Tn5 in vitro transposition. *J. Biol. Chem.* 273:7367-7374.

Gross, D.S. and Garrard, W.T. 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57:159-197.

Kornberg, R.D. 1974. Chromatin structure: A repeating unit of histones and DNA. *Science* 184:868-871.

Kouzarides, T. 2007. Chromatin modifications and their function. *Cell* 128:693-705.

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C., Fisher-Aylor, K.I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A.J., Hoffman, M.M., Iyer, V.R., Jung, Y.L., Karmakar, S., Kellis, M., Kharchenko, P.V., Li, Q., Liu, T., Liu, X.S., Ma, L., Milosavljevic, A., Myers, R.M., Park, P.J., Pazin, M.J., Perry, M.D., Raha, D., Reddy, T.E., Rozowsky, J.,

Shoresh, N., Sidow, A., Slattery, M., Stamatoyannopoulos, J.A., Tolstorukov, M.Y., White, K.P., Xi, S., Farnham, P.J., Lieb, J.D., Wold, B.J., and Snyder, M. 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* 22:1813-1831.

Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., Maurano, M.T., Humbert, R., Rynes, E., Wang, H., Vong, S., Lee, K., Bates, D., Diegel, M., Roach, V., Dunn, D., Neri, J., Schafer, A., Hansen, R.S., Kutyavin, T., Giste, E., Weaver, M., Canfield, T., Sabo, P., Zhang, M., Balasundaram, G., Byron, R., MacCoss, M.J., Akey, J.M., Bender, M.A., Groudine, M., Kaul, R., and Stamatoyannopoulos, J.A. 2012. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489:83-90.

Reznikoff, W.S. 2008. Transposon Tn5. *Annu. Rev. Genet.* 42:269-286.

Rinn, J.L. and Chang, H.Y. 2012. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81:145-166.

Song, L. and Crawford, G.E. 2010. DNase-seq: A high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor Protoc.* 2010:pdb.prot5384.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., Boatman, L., Canfield, T.K., Diegel, M., Dunn, D., Ebersol, A.K., Frum, T., Giste, E., Johnson, A.K., Johnson, E.M., Kutyavin, T., Lajoie, B., Lee, B.K., Lee, K., London, D., Lotakis, D., Neph, S., Neri, F., Nguyen, E.D., Qu, H., Reynolds, A.P., Roach, V., Safi, A., Sanchez, M.E., Sanyal, A., Shafer, A., Simon, J.M., Song, L., Vong, S., Weaver, M., Yan, Y., Zhang, Z., Zhang, Z., Lenhard, B., Tewari, M., Dorschner, M.O., Hansen, R.S., Navas, P.A., Stamatoyannopoulos, G., Iyer, V.R., Lieb, J.D., Sunyaev, S.R., Akey, J.M., Sabo, P.J., Kaul, R., Furey, T.S., Dekker, J., Crawford, G.E., and Stamatoyannopoulos, J.A. 2012. The accessible chromatin landscape of the human genome. *Nature* 489:75-82.

Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. 2011. Determinants of nucleosome organization in primary human cells. *Nature* 474:516-520.