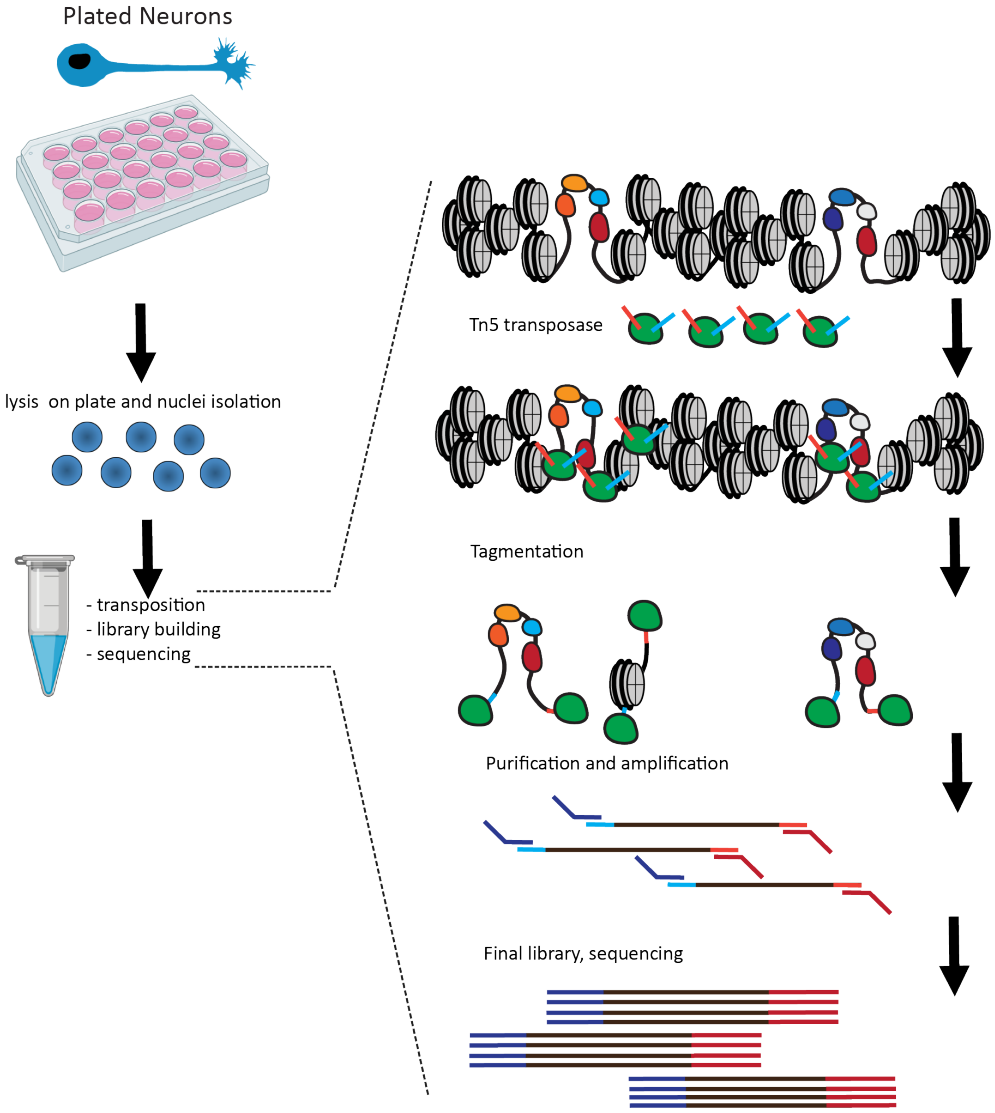
Mapping active regulatory elements genome-wide in cultured primary neurons using ATAC-seq

Graphical Abstract



ATAC-seq (**A**ssay for **T**ransposase-**A**ccessible **C**hromatin using **seq**uencing; based on the preferential insertion of the Tn5 transposase into physically accessible DNA) enables the rapid and straightforward genome-wide profiling of open chromatin regions in the genome, and thus allows researchers to map the identity and track the activity of regulatory elements across cell types and cellular states. However, applying the assay to cultured neurons is not straightforward as their dissociation causes rapid cell death, which interferes with ATAC-seq results. Here we describe a version of the ATAC-seq protocol adapted to the problem of profiling accessible chromatin in cultured primary neurons.

Mapping active regulatory elements genome-wide in cultured primary neurons using ATAC-seq

Maya Maor-Nof[1], Zohar Shipony[1], Georgi K. Marinov[1], William J. Greenleaf[1,2,3,4], and Aaron D. Gitler[1]

[1]*Department of Genetics, Stanford University, Stanford, CA 94305, USA*
[2]*Center for Personal Dynamic Regulomes, Stanford University, Stanford, California 94305, USA*
[3]*Department of Applied Physics, Stanford University, Stanford, California 94305, USA*
[4]*Chan Zuckerberg Biohub, San Francisco, California, USA*

Summary

A key feature of active *cis*-regulatory elements (cREs) in eukaryotes is their nucleosomal depletion, which in turn translates into elevated physical accessibility. Methods for identifying cREs genome-wide and tracking their dynamics across cell types and cellular states rely on this property, taking advantage of preferential enzymatic cleavage or labeling of accessible DNA. ATAC-seq has become established in recent years as a versatile, adaptable and widely adopted method for accomplishing the task of mapping open chromatin regions. However, some biological systems present unique challenges to its application. Primary neurons are one such example – conventional ATAC-seq would require their dissociation, but dissociating them leads to rapid cell death and major changes in cell state, affecting ATAC-seq results. We describe the ATAC-seq protocol that we have developed in order to address this challenge for cultured primary neurons.
For complete details on the use and execution of this protocol, please refer to Maor-Nof et al. (2021)

Before You Begin

There should be some stuff about handling neurons in here; I am not sure what and how to write

Key resources table

Table 1: Key resources table.

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Tn5[1]	Illumina	FC-131-1024
Sequencing primers/adapters[2]	Illumina	FC-131-1024
NEBNext High-Fidelity 2× PCR Master Mix	NEB	M0541S
IGEPAL CA-630 detergent[3]	Sigma	11332465001
Tween-20 detergent[4]	Sigma	11332465001

Continued on next page

[1]Tn5 is the key reagent in the ATAC-seq protocol; it can be obtained from Illumina as listed here, but it can also be prepared in-house, following the protocol described previously by Picelli et al. (2014)
[2]PCR primers for amplifying ATAC-seq libraries can also be ordered directly from other sources; 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)
[3]Supplied as a 10% solution
[4]Supplied as a 10% solution; store at 4 °C

Table 1 – *Continued from previous page*

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Digitonin detergent[5]	Promega	G9441
1M Tris-HCl pH 7.5	Thermo Fisher	15567027
5M NaCl	Thermo Fisher	AM9759
1M $MgCl_2$	Thermo Fisher	AM9530G
Dimethyl Formamide	Sigma	D4551
Deoxyribonuclease I	Worthington	LS006331
10mM dNTP mix	Thermo Fisher	18427013
25× SYBR Green	Thermo Fisher	S7563
Phusion High-Fidelity DNA Polymerase	NEB	M0530L
MinElute PCR Purification Kit	Qiagen	28004/28006
Zymo DNA Clean and Concentrator Kit	Zymo	D4013/D4014
Nuclease-free H_2O	Thermo Fisher	AM9916
1× PBS buffer solution	Thermo Fisher	10010023
qPCR machine (StepOne or equivalent)		
200-μL PCR tubes		
1.5-mL microcentrifuge tubes[6]	Eppendorf	022431021
Thermomixer	Eppendorf	5382000023
Tabletop centrifuge		
Thermal cycler		
Qubit fluorometer	Thermo Fisher	Q33238
QuBit tubes	Thermo Fisher	Q32856
QuBit dsDNA HS Assay Kit	Thermo Fisher	Q32854
TapeStation	Agilent	
TapeStation D1000 tape	Agilent	5067-5582
TapeStation D1000 reagents	Agilent	5067-5583

Materials and equipment

Table 2: ATAC-RSB buffer (master stock, 50 mL) .

Reagent	Final concentration	Amount per sample
1M Tris-HCl pH 7.4	10 mM	500 μL
5M NaCl	10 mM	100 μL
1M $MgCl_2$	3 mM	150 μL
H_2O		49.25 mL

Table 3: ATAC-RSB buffer (1 mL) .

Reagent	Final concentration	Amount per sample
10% IGEPAL CA-630	0.1%	10 μL
10% Tween-20	0.1%	10 μL
2% Digitonin	0.01%	5 μL
ATAC-RSB		970 μL

[5]Supplied as a 2% solution in DMSO; store at -20 °C)
[6]Tubes should be preferably low protein- and DNA-binding

Table 4: ATAC-RSB buffer (10 mL) .

Reagent	Final concentration	Amount per sample
10% Tween-20	0.1%	100 μL
ATAC-RSB		9.9 mL

Table 5: 2× TD buffer (10 mL) .

Reagent	Final concentration	Amount per sample
Tris-HCl pH 7.5	20 mM	200 μL
1M $MgCl_2$	10 mM	100 μL
Dimethyl Formamide	20%	2 mL
H_2O		9.78 mL

Note: store the ATAC-RSB master buffer at 4 °C, and only prepare the ATAC-RSB-Lysis and ATAC-RSB-Wash buffers prior to use. Stored the 2× TD buffer and the Tn5 enzyme at -20 °C

Step-by-step method details

Note: This protocol has been adapted from the omniATAC version of the ATAC-seq assay, previously described in Corces et al. (2017)

Note: We advise to perform at least two independent replicates for each condition assayed.

Preparation of cells

XXX SOMETHING ABOUT NEURONS IN CULTURE, OR MAYBE NOTHING AND WE DON'T HAVE THIS SECTION XXX

DNAse treatment of cells

Timing: ∼40 minutes.

Note: DNAse treatment helps improve signal to noise by removing free-floating DNA and digesting DNA from dead cells.

1. Add DNAse at a concentration of 200 U/mL to the media in the plate with the neurons
2. Incubate at 37 °C for 30 minutes
3. Remove the media with the DNAse from the plate
4. Add cold 1× PBS
5. Remove the PBS
6. Add cold 1× PBS for a second time
7. Remove the PBS
8. Add cold 1× PBS for a third time
9. Remove the PBS
10. Add cold 1× PBS for a fourth time
11. Remove the PBS

Preparation of nuclei

Timing: ~30 minutes.

Note: To avoid cell death caused by trypsinization, neurons are directly lysed on the plate, and the nuclei are prepared from the lysate.

12. Add 1 mL cold ATAC-RSB-Lysis Buffer to the neurons
13. Incubate on ice for 10 minutes
14. Collect nuclei from the plate and count them
15. Centrifuge ~50,000 nuclei at 500 g for 5 minutes in a pre-chilled 4 °C fixed-angle centrifuge
16. Carefully aspirate the supernatant in two steps, by first removing most of it, then using the P200 pipette to remove the last ~100 μL
17. Resuspend the pellet in 1 mL of ATAC-RSB-Wash Buffer
18. Centrifuge for 10 minutes at 500 g ar 4 °C
19. Carefully aspirate the supernatant in two steps as described above.

Transposition

Timing: ~35 minutes.

Carry out transposition as follows:

20. Immediately resuspend the pellet in the transposase reaction mix (prepare a master mix for multiple samples in the same proportions):
    25 μL TD buffer
    2.5 μL Tn5
    5 μL nuclease-free H_20
    16.5 μL 1× PBS
    0.5 μL 1% digitonin
    0.5 μL 10% Tween-20
21. Incubate at 37 °C for 30 min in a Thermomixer with shaking at 1000 RPM.

DNA purification

Timing: ~20 minutes.

Note: Reactions can be cleaned up either with the Zymo DNA Clean and Concentrator or the Qiagen MinElute Cleanup kits, with equivalent results.

22. Immediately stop the reaction using 250 μL (i.e 5×) of PB buffer (if using MinElute) or DNA Binding Buffer (if using Zymo).
23. Purify samples following the kit instructions.
24. Elute with 10 μL of Elution Buffer.

PCR amplification and library generation

Timing: ~1 hour.

Note: When amplifying tranposed DNA, the initial extension is needed to fill in the gap left from the transposition itself and allow PCR primers to land in subsequent amplification cycles. Hot-start polymerase mixes, in which the polymerase is only activated by exposion to denaturation temperatures, are therefore not recommended for amplifying ATAC-seq libraries.

25. Set up the following PCR reaction:
    10 μL transposition elutate
    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
26. Optimize PCR conditions, pre-amplification. Amplify DNA for 5 cycles as follows:
    72 °C for 3 minutes
    98 °C for 30 seconds
    5 cycles of:
        98 °C for 10 seconds
        63 °C for 30 seconds
        72 °C for 30 seconds
    Hold at 4 °C
27. Determine 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
28. Determine additional cycles using qPCR. Run the qPCR reaction with the following settings in a qPCR machine:
    98 °C for 30 seconds
    20 cycles of:
        98 °C for 10 seconds
        63 °C for 30 seconds
        72 °C for 30 seconds
    Hold at 4 °C
29. Assess the amplification profiles and determine the required number of additional cycles to amplify.
30. 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 seconds
        63 °C for 30 seconds
        72 °C for 30 seconds
    Hold at 4 °C

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

    In practice, 8-10 cycles are usually sufficient to amplify a standard ATAC library thus if a large number of samples are being processed at the same time, the following reaction can be run:
31. Single-step PCR.
    72 °C for 3 minutes
    98 °C for 30 seconds
    8-10 cycles of:
        98 °C for 10 seconds
        63 °C for 30 seconds
        72 °C for 30 seconds
    Hold at 4 °C
32. Purify the amplified library as described abode for the purified ATAC reaction. Elute in 20 μL Elution Buffer.

Library size distribution profiling

Timing: ~10 minutes.

There are multiple options for carrying out this step, e.g. the TapeStation and BioAnalyzer instruments. We prefer to use a TapeStation with the D1000 or HS D1000 kits due to its ease of use, flexibility and rapid turnaround time. Follow the manufacturer's instructions depending on the exact instrument and kit used.

Library quantification

Timing: ~50 minutes.

Note: This step is typically carried out using a Qubit fluorometer for most high-throughput sequencing libraries that exhibit a unimodal fragment length distribution. However, ATAC-seq fragment distribution is usually not unimodal and ATAC-seq libraries often include fragments longer than what can be sequenced on standard Illumina instruments. Effective library concentrations therefore often differ from apparent library concentrations measured using Qubit. The best way to estimate effective library concentration is thus qPCR. Commercial kits such as the NEBNext Library Quant Kit for Illumina or KAPA Library Quantification Kits can also be used, in a similar manner.

33. Generate a standard curve using Illumina PhiX standard (10nM) by first making a 50× dilution to 200 pM, then making additional serial 2× dilutions to 100 pM, 50 pM, 25 pM, 12.5 pM, 6.25 pM, 3.125 pM, and 1.56 pM.
34. 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 Short Oligo C/Adapter 1
    1 μL Short Oligo D/Adapter 2
    0.4 μL 10mM dNTP mix
    0.5 μL 25× SYBR Green (in DMSO)
    0.2 μL NEB Phusion HF

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

    98 °C for 30 seconds
    20 cycles of:
        98 °C for 10 seconds
        63 °C for 30 seconds
        72 °C for 30 seconds
    Hold at 4 °C

36. Create a standard curve based on the PhiX dilutions and estimate the library's true molarity based on it.

Sequencing

This protocol generates libraries intended to be sequenced on Illumina sequencers.

Once libraries have been made, the optimal sequencing format needs to be decided on, as there are multiple different Illumina kis, which differ in their output, read length, and cost.

ATAC-seq libraries should not be sequenced in a single-end format, as the analysis of fragment lengths is important for the quality evaluation of ATAC-seq datasets and for a number of downstream analyses, which are only possible in a paired-end format. In addition, some analytical tasks, (such as transcription factor footprinting) focus on Tn5 insertions rather than read coverage, and paired-end reads produce twice as many such data points for the same cost.

We also note that the post-sequencing ATAC-seq insert length distribution peaks between 50 and 100 bp (Figure 2A). It is accordingly most cost-effective to sequence ATAC libraries in 2×36 bp or 2 × 50 bp formats (depending on whether using a NextSeq or some of the higher-throughput Illumina instruments). However, some applications (e.g. analyzing the effects of sequence variation on chromatin accessibility) can benefit from longer reads, and this should be kept in mind depending on the goals of the particular study.

Expected Outcomes

Figure 1 shows examples of typical ATAC-seq library size profiles. A clear nucleosomal pattern is expected, with peaks corresponding to subnucleosomal, mononucleosomal, dinucleosomal, and so on fragments, due to the inhibitory effect of nucleosomes on Tn5 insertion. Occasionally, it is possible to see a "flattened" profile, in which the nucleosomal peaks stand out less than usual. This is often due to the presence of a large amount of mitochondrial genome-derived fragments and does not necessarily affect enrichment for open chromatin in the nuclear genome.

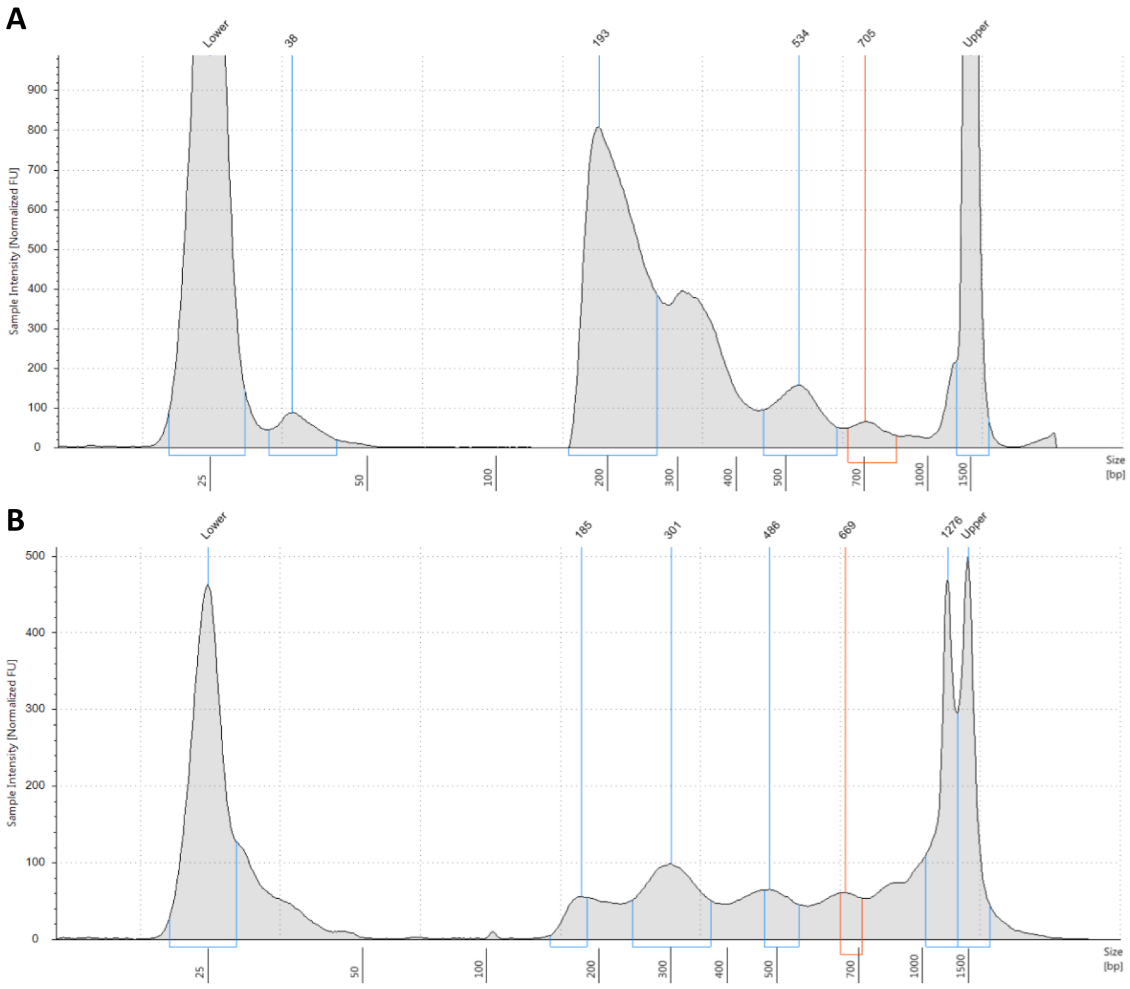


Figure 1: Typical TapeStation profile (D1000 TapeStation in this case) of an ATAC-seq library. ATAC-seq libraries tend to display a nucleosomal pattern with dominant peaks corresponding to subnucleosomal fragments (the ∼180-∼250 range; note that the length of adapters is included in these values), mononucleosomes, dinucleosomes, and so on. The relative height of peaks can occasionally vary between different libraries (A) and (B).

Quality characterization and evaluation after sequencing is based on the following criteria:

1. Evaluation of the fragment length distribution (Figure 2A). It is possible to have a very prominent subnucleosomal peak without a strong mononucleosomal one and still have quite high enrichment for open chromatin, but high-quality ATAC libraries in eukaryotes typically display the characteristic nucleosomal signature in their fragment length distribution
2. Evaluation of open chromatin enrichment (Figure 2B-C). To this end, the average profile around the transcription start sites (TSS) of protein coding genes is a very useful, intuitive and independent of ad hoc parameters such as peak calling thresholds measure. It can also be distilled to a single "TSS score" number, which is calculated as the ratio of signal over the region immediately (e.g. ±100 bp around the TSS) and the regions of equal size located ±2 kb on the flanks either side of the TSS (Marinov & Shipony 2021). High-quality ATAC libraries tend to have TSS

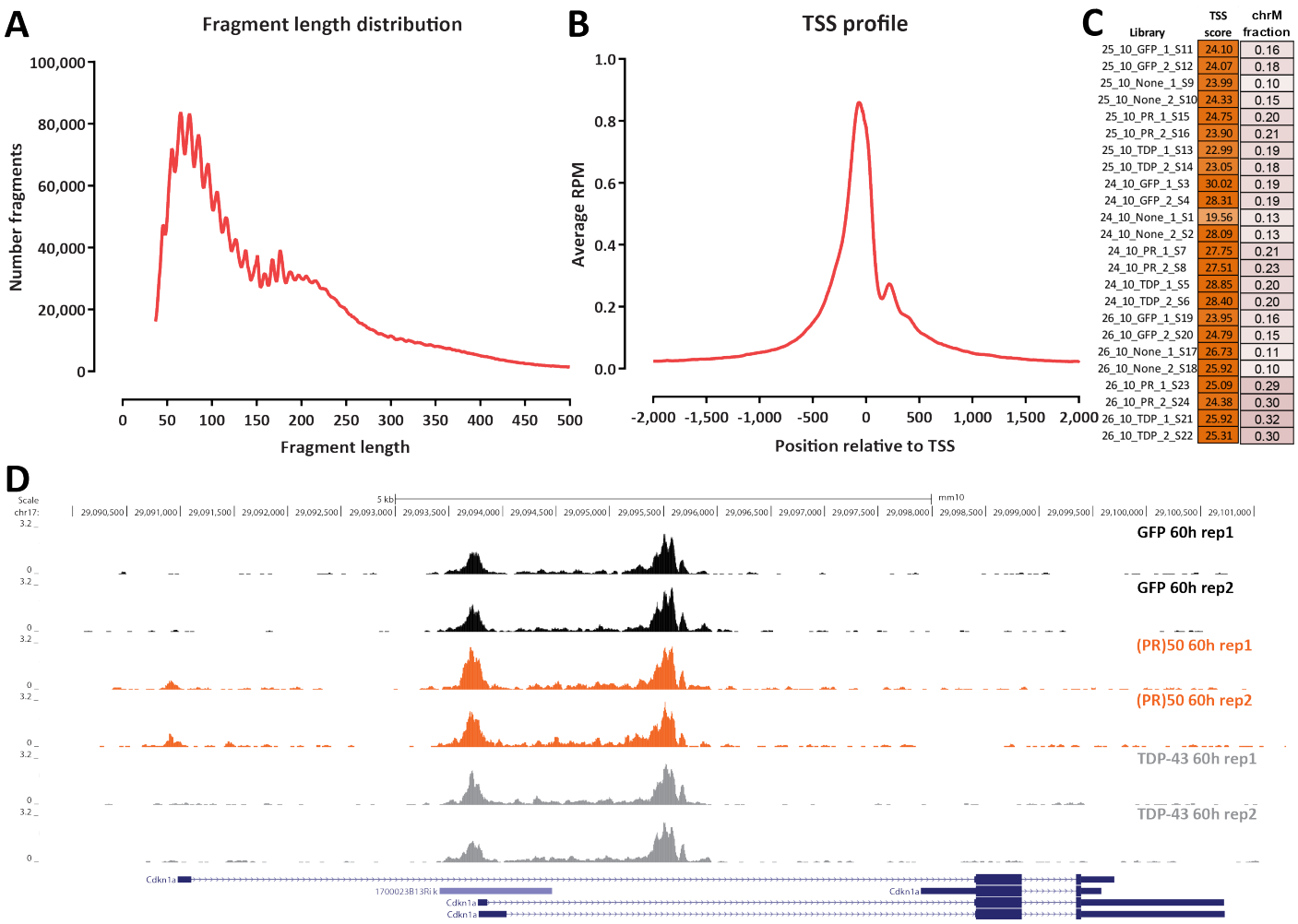


Figure 2: Typical ATAC-seq results after processing of sequencing data. (A) Length distribution for mapped fragments (shown is the dataset corresponding to SRA accession SRR13120289) (B) TSS profile (for the same sample) (C) TSS scores for the library from Maor-Nof et al. (2021). (D) ATAC-seq shows activation of one of the promoters of the *Cdkn1a* gene in $(PR)_{50}$ cells relative to TDP-43 and control GFP cells.

scores $\geq$10 for mammalian genomes. TSS scores for the libraries from Maor-Nof et al. (2021) are shown in Figure 2C.

3. Evaluation of the extent of mitochondrial contamination. As mitochondria do not have nucleosomes and their DNA is highly accessible, they are preferentially transposed by Tn5. High levels of mitochondrial reads are not necessarily associated with poor open chromatin enrichment in the nuclear genome, but they result in having to sequence much deeper to obtain the same effective sequencing coverage, and are a sign of a need to optimize the protocol. While in early versions of the ATAC-seq protocol (Buenrostro et al. 2013), most of ATAC libraries consisted of mitochondria-derived fragments, since then the omniATAC protocol (Corces et al. 2017), on which the plated neuronal ATAC protocol is based, has greatly reduced the level of mitochondrial contamination. The fraction of chrM reads for the libraries from Maor-Nof et al. (2021) is shown in Figure 2C.
4. The molecular complexity of libraries – high-quality libraries should contain a large number of distinct fragments.
5. The effective sequencing depth – in general we aim for ~20-30 million reads after deduplicating fragments mapping to the nuclear genome.

ATAC-seq libraries should also display visible enrichment over promoters and other regulatory elements when examined using a genome browser, as shown in Figure 2D.

Limitations

XXX I actually cannot think of anything right now XXX

Troubleshooting

Problem 1

Libraries exhibit low TSS enrichment, i.e. TSS scores substantially below 10.

Potential Solution

ATAC-seq is quite robust and usually works well in terms of producing good enrichment for open chromatin. When issues with poor enrichment are encountered, this is typically due to problems with the input material, such as the presence of many dead and nonviable cells (which contain significant quantities of dechromatinized DNA) or free floating DNA. The DNAse pretreatment step is designed to address these issues; also make sure that the neurons are in optimal condition when starting the protocol

Problem 2

Libraries contain a high fraction of mitochondrial reads.

Potential Solution

The omniATAC protocol and its derivatives are usually quite successful at minimizing the extent of mitochondrial contamination. A typical reason for very high levels of chrM-mapping reads is failure to aspirate all the supernatant (which contains the mitochondria) during the nuclei preparation procedure prior to transposition. Make sure to remove all of it while being careful not to disturb the pellet.

Problem 3

Final ATAC-seq libraries contain few distinct fragments and are of generally low molecular complexity.

Potential Solution

This issue could be due to cell loss during the nuclei preparation procedure. As ATAC-seq works on relatively small number of cells/nuclei – only ∼50,000 – cell and nuclei pellets are often quite small and barely visible in tubes. It can thus be easy to inadvertently disturb them while aspirating supernatants, leading to cell loss. Be careful to avoid pellets by using the usual methods of spinning tubes consistently with one side pointing outwards within the centrifuge then carefully pipetting out liquids on the opposite side of the tube.

Resource availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Maya Maor-Nof (maormaya@stanford.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The raw and analyzed sequence data from our original paper carrying out ATAC-seq in primary cultured neurons (Maor-Nof et al., 2021) can be found on NCBI GEO (GSE162048).

Acknowledgments

This work was supported by NIH grants R35NS097263(10) (A.D.G.), P50HG007735 (W.J.G.), R01HG008140 (W.J.G.), U19AI057266 (W.J.G.), UM1HG009442 (W.J.G.), and 1UM1HG009436 (W.J.G.), the Brain Rejuvenation Project of the Wu Tsai Neurosciences Institute (A.D.G.). W.J.G. is a Chan Zuckerberg Biohub investigator. Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. This work used the Genome Sequencing Service Center by Stanford Center for Genomics and Personalized Medicine Sequencing Center, supported by NIH grant S10OD020141. Some of the figures were created with BioRender.com.

Author Contributions

Conceptualization, M.M.-N., Z.S. and A.D.G.; methodology, M.M.-N., and Z.S.; writing, M.M.-N., G.K.M.; supervision, W.J.G., A.D.G.

Declaration of Interests

A.D.G. has served as a consultant for Aquinnah Pharmaceuticals, Prevail Therapeutics, and Third Rock Ventures and is a scientific founder of Maze Therapeutics. W.J.G. has affiliations with 10x Genomics (consultant), Guardant Health (consultant), and Protillion Biosciences (co-founder and consultant).

References

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.G. 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.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., Kathiria, A., Cho, S.W., Mumbach, M.R., Carter, A.C., Kasowski, M., Orloff, L.A., Risca, V.I., Kundaje, A., Khavari, P.A., Montine, T.J., Greenleaf, W.J., Chang, H.Y. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* 14, 959–962.

Maor-Nof, M., Shipony, Z., Lopez-Gonzalez, R., Nakayama, L., Zhang, Y.J., Couthouis, J., Blum, J.A., Castruita, P.A., Linares, G.R., Ruan, K., Ramaswami, G., Simon, D.J., Nof, A., Santana, M., Han, K., Sinnott-Armstrong, N., Bassik, M.C., Geschwind, D.H., Tessier-Lavigne, M., Attardi, L.D., Lloyd, T.E., Ichida, J.K., Gao, F.B., Greenleaf, W.J., Yokoyama, J.S., Petrucelli, L., Gitler, A.D.. 2021. p53 is a central regulator driving neurodegeneration caused by C9orf72 poly(PR). *Cell* 184, 689–708.e20

Marinov, G.K., Shipony, Z. 2021. Interrogating the Accessible Chromatin Landscape of Eukaryote Genomes Using ATAC-seq. Methods. Mol. Biol. 2243, 183–226

Picelli, S., Björklund, A.K., Reinius, B., Sagasser, S., Winberg, G., Sandberg, R. 20214. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome. Res.* 24, 2033–2040.