How to tackle challenging ChIP-seq, with long-range cross-linking, using ATRX as an example.

Julia Truch[1], Jelena Telenius[2], Douglas R Higgs[1] and Richard J Gibbons[1]

Affiliation:

[1]Medical Research Council Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, UK.

[2]Computational Biology Research Group, MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, UK.

Corresponding authors:

Richard J Gibbons

richard.gibbons@ndcls.ox.ac.uk

Douglas R Higgs

doug.higgs@imm.ox.ac.uk

Running Head: ATRX ChIP-seq

Abstract

Chromatin immunoprecipitation coupled with high-throughput, next generation DNA sequencing (ChIP-seq) has enabled researchers to establish the genome-wide patterns of chromatin modifications and binding of chromatin-associated proteins. Well-established protocols produce robust ChIP-seq data for many proteins by sequencing the DNA obtained following immunoprecipitation of fragmented chromatin using a wide range of specific antibodies. In general, the quality of these data mainly depends on the specificity and avidity of the antibody used. However, even using optimal antibodies, ChIP-seq can become more challenging when the protein associates with chromatin via protein-protein interactions rather than directly binding DNA. An example of such a protein is the alpha-thalassaemia mental retardation X-linked (ATRX) protein; a chromatin remodeller that associates with the histone chaperone DAXX, in the deposition of the replication-independent histone variant H3.3 and plays an important role in maintaining chromatin integrity [1]. Inherited mutations of ATRX cause syndromal mental retardation (ATR-X Syndrome) [2] whereas acquired mutations are associated with myelodysplasia, acute myeloid leukaemia (ATMDS syndrome) [3] and a range of solid tumours. Therefore, high quality ChIP-seq data have been needed to analyse the genome-wide distribution of ATRX, to advance our understanding of its normal role and to comprehend how mutations contribute to human disease. Here we describe an optimised ChIP-seq protocol for ATRX which can also be used to produce high quality data sets for other challenging proteins which are indirectly associated with DNA and complement the ChIP-seq toolkit for genome-wide analyses of histone chaperon complexes and associated chromatin remodellers. Although not a focus of this chapter, we will also provide some insight for the

analysis of the large dataset generated by ChIP-seq. Even though this protocol has been fully optimised for ATRX, it should also provide guidance for efficient ChIP-seq analysis, using the appropriate antibodies, for other proteins interacting indirectly with DNA.

Keywords

Chromatin immunoprecipitation, High-throughput DNA sequencing, ChIP-seq, Alpha-thalassaemia mental retardation X-linked, ATRX, histone variant H3.3, Library preparation, NGS, chromatin remodeller, double cross-linking, EGS.

1 Introduction

Chromatin immunoprecipitation (ChIP) is now a key technique for analysing protein-DNA interactions *in vivo* [4,5]. The power of this technique comes from the ability of crosslinking agents, such as formaldehyde, to covalently join the protein-DNA complexes in their native chromatin environment in living cells [6,7]. The cross-linked chromatin is then sheared (by sonication or nuclease digestion) and immunoprecipitated using a specific antibody against a protein of interest. After reverse cross-linking of the ChIP sample, the target DNA fragments, pulled down with the protein of interest and enriched in the immunoprecipitate, are purified. This DNA enrichment can then simply be assessed, in a locus-specific manner, by quantitative Real Time PCR (qPCR): so called ChIP-qPCR [8]. To appraise the ChIP data in defined segments of the genome and genome-wide, DNA fragments in the immunoprecipitate were, in the past, analysed by labelling and hybridising to a microarray, a technique known as ChIP-Chip [9,10]. With the improvement of the next generation sequencing (NGS) technoligies, this method has now been supplanted

and, today, DNA enrichment in ChIP experiments is assessed by high-throughput DNA sequencing (ChIP-seq) [11]. Using this technique it is now possible to study the genome-wide distribution of DNA-protein interactions at an economical cost. ChIP-seq has paved the way to many studies focusing on DNA binding proteins, such as transcription factors, and histone modifications, as illustrated by the ENCODE project [12].

However, standard single cross-link ChIP experiments using a short spacer arm cross-linker agent such as formaldehyde (with an interaction range of 2Å) have limitations and are most appropriate for proteins which directly bind DNA (nucleosomes and transcription factors) [6,13,14]. By contrast, assessing the genome-wide profiling of proteins indirectly bound to DNA and/or displaying hyperdynamic interactions can be more challenging and often requires additional cross-linking agents such as ethylene glycol-bis(succinimidylsuccinate) (EGS) with a spacer arm of approximately 16 Å [15-17].

This is the case of chromatin remodeller ATRX. This member of the SNF2 family of helicase/ATPases [2] is involved, in association with the histone chaperone DAXX, in the deposition of the replication-independent histone variant H3.3 in banks of repetitive sequences including endogenous retroviral elements, telomeric DNA, pericentric DNA, and ribosomal DNA [18-20]. Furthermore, ATRX co-localises with variable tandem repeats which include GC-rich sequences predicted to form non-B DNA structures known as G-quadruplexes [21,22]. ATRX has also been shown to associate with chromatin through its interactions with several proteins such as H3K9me3, HP1 [23] and MeCp2 [24].

Mutations in the *ATRX* gene were first identified as the cause of the ATR-X syndrome, characterized by severe intellectual disabilitiy, morphological

abnormalities and alpha-thalassaemia (a blood disorder due to defective production of alpha-globin) [25]. Since then, ATRX mutations have also been associated with a subset of cancers using the telomerase-independent Alternative Lengthening of Telomeres pathway to maintain their telomere length [26,27]. Multiple lines of evidence have demonstrated the role of ATRX in expression of a subset of genes (e.g. the alpha-globin genes) [21,28]. In addition, the loss of ATRX function affects diverse nuclear processes including DNA methylation, replication, genome stability, mitosis and meiosis [29-34].

These observations emphasise the need to produce robust ATRX ChIP-seq data to identify the targets of this chromatin remodeller and help to define its role in health and disease. However, producing high quality ATRX-ChIP has been challenging for various reasons. 1) A large number of cells (10^8) is necessary. 2) A two step cross-linking using EGS and formaldehyde is required, making the fixed chromatin more resistant to sonication. The use of stringent sonication parameters may have significant effects on the quality of the ChIP-seq data and exacerbate the artefacts such as the preferential fragmentation of DNA within regions of open chromatin, producing the artifactual appearance of a nuclease accessibility like-profile [35-37]. As a result, such data may produce a "noisy" background impeding their interpretation. Therefore, it is crucial to establish the optimal sonication. 3) The specificity and avidity of the antibody is also crucial and may be batch dependent. 4) The degree of DNA enrichment produced per ChIP is limited. 5) The small quantity of starting material available for the library preparation as well as the proportion of GC-rich sequences lead to potentially increases in PCR amplification bias. 6) The repetitive nature of many ATRX targets increases the complexity of the data analysis (e.g. issues with unique mapping reads).

In this chapter, we describe the detailed methodology to carry out ChIP experiments requiring a double cross-linking and to generate the subsequent libraries for high-throughput sequencing using the Illumina NextSeq500. Furthermore, we also discuss the quality control steps and provide some insights into the analysis of such data.
It should be noted that while this protocol has been optimised to generate high quality ATRX ChIP-seq data, it is potentially versatile enough to be adapted for other protein targets indirectly associated with DNA and/or with hyperdynamic interactions and requiring a double-crosslinking such as DAXX [38],the LKB1 and cofactor FOG-1 [17], HIRA, UBN1 and AFS1a [39].

2 Material

2.1 Crosslinking Reagent

1. Phosphate-buffered saline (PBS)
2. 38 % formaldehyde solution
3. EGS stock solution: make 500 mM EGS in Dimethyl sulfoxide (DMSO) stock solution immediately prior use. Dissolve 0.046 g EGS in 200 µL of DMSO. Mix well until the powder is fully dissolved and the solution is clear. Use immediately (see 1).
4. Glycine stock solution: 1M in ultrapure water. Dissolve 3.75 g of glycine powder in ultrapure water to a final volume of 50 mL. Mix until the powder is fully dissolved and the solution is clear. Store at 4°C

2.2 Bead preparation

1. Dynabeads protein A and/or G

2. Magnetic stand
3. Primary antibody (see)
4. BSA solution: Make fresh on the day bovine serum albumin (BSA) solution 5 mg/mL in PBS. Dissolve 250 mg of BSA in PBS to a final volume of 50 mL. Mix until complete dissolution of the powder. Chill on ice.
5. PBS

2.3 Lysis buffers

1. Protease inhibitor cocktail tablets
2. Lysis buffer 1 (100 mL) (100mM Hepes-KOH, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40 and 0.25% Triton X-100). In 76.25 mL of ultrapure water, add 10 mL of 1M Hepes-KOH, pH 7.5, 2.8 mL of 5M NaCl, 0.2 mL of 0.5M Ethylenediaminetetraacetic acid (EDTA), 10 mL glycerol, 0.5 mL of 100% NP-40 and 0.25 mL of Triton X-100. Store at 4°C. Protease inhibitor cocktail should be added to 1X final concentration on the day of use. Do not store the left over solution once the Protease inhibitor cocktail has been added.
3. Lysis buffer 2 (100 mL) (200mM NaCl, 1mM EDTA, 0.5mM EGTA and 10mM Tris-HCl). In 93.4 mL of ultrapure water add 4 mL of 5M NaCl, 0.2 mL of 0.5M EDTA, 0.4 mL of 125mM Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA) and 2 mL of 0.5M Tris pH 8. Store at 4°C. (see)
4. Lysis buffer 3 (100 mL) (1mM EDTA, 0.5mM EGTA, 10mM Tris-HCl at pH 8, 100mM NaCl, 0.1% Na-Deoxycholate and 0.5% N-lauroyl sarcosine). In 50 mL of ultrapure water add 0.2 mL of 0.5M

EDTA, 0.4 mL of 125 mM EGTA, 2 mL of 500 mM Tris-HCl at pH 8, 2 mL of 5M NaCl, 0.1g of Na-Deoxycholate and 500 mg of N-lauroyl sarcosine. Then adjust the volume to 100 mL with ultrapure water. Store at 4°C.

2.4 Sonication

1. Covaris S-series
2. Covaris milliTUBE 1mL AFA Fiber
3. 10% Triton X-100

2.5 Washes, elution and reverse cross-linking

1. Wash buffer (100 mL) (50mM Hepes at pH 7.6, 1mM EDTA, 0.7% Na-Deoxycholate, 1% NP-40 and 0.5M LiCl). In 50 mL of ultrapure water add 10 mL of 500 mM Hepes at pH 7.6, 200 µL of 0.5M EDTA, 0.7g of Na-Deoxycholate, 1 mL of NP-40 and 6.25 mL 8M LiCl (or 2.12g powder). Then adjust the volume to 100 mL with ultrapure water. Store at 4°C.
2. Elution buffer (5 mL) (0.5% SDS and 100mM NaHCO3). Make fresh on the day elution buffer by adding 125µL of 20% SDS and 500 µl of 1M NaHCO3 in 4.375 mL of ultrapure water. Keep at room temperature. Scale up if needed depending on the number of ChIP samples.
3. PBS
4. 5M NaCl

2.6 DNA purification

1. Proteinase K 20 mg/mL

2. 1M Tris, pH 6.5
3. 0.5 M EDTA
4. Phenol-chloroform-isoamylalcohol (25:24:1, v/v/v)
5. Ethanol 100% and 70%
6. 3M Sodium Acetate
7. Glycogen

2.7 DNA quantification

1. Qubit dsDNA HS Assay Kit

2.8 Sonication validation

2. 2% agarose gel
3. Loading dye
4. SYBR safe DNA gel stain
5. DNA Ladder

2.9 ChIP validation by qPCR

1. Fast SYBR Green Master Mix
2. Primers for positive and negative control regions (see)

2.10 Library preparation and quantification and sequencing

1. NEBNext Ultra DNA Library Prep Kit for Illumina (E7370S/L)
2. NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) (E7335S/L)
3. AMPure XP Beads
4. 10 mM Tris-HCl, pH 7.5–8.0
5. 0.1X TE, pH 8.0
6. Ethanol 80% (made fresh on the day)

7. High Sensitivity D1000 ScreenTape
8. High Sensitivity D1000 Reagents
9. Agilent 2200 TapeStation
10. KAPA Library Quantification Kit
11. NextSeq 500/550 High Output Kit v2 (150 cycles) or Mid Output Kit v2 (150 cycles) (if using NextSeq500 as Illumina sequencing platform)
12. NextSeq500 (or other Illumina Sequencing platform)

3 Method

3.1 ATRX ChIP

The ATRX ChIP protocol as been optimised based on Law *et al.* (2010). The experimental workflow is shown in Figure 1.

3.1.1 Cell cross-linking

Both room temperature and cold PBS are required for the crosslinking step.

1. Use 10^8 live cells per ChIP (upon harvest, the cell viability should be ≥ 90 %)
2. Wash twice with room temperature PBS.
3. Resuspend the cells in a 50 mL falcon tube in 20 mL room temperature PBS. Make sure that the cells are in single cell suspension. Add 80 µL of freshly made 500 mM EGS stock solution per 20mL of PBS (final concentration of 2 mM) (see).
4. Incubate at room temperature for 45 minutes on a roller with constant mixing (see).

5. Add 526 µL of 38% formaldehyde solution per 20 mL of cell suspension (1% final) (see).
6. Incubate at room temperature for 20 minutes on a roller with constant mixing.
7. Add 2.56 mL of 1M glycine stock solution per 20.5 mL cross-linked cell suspension (125 mM final) to quench the cross-linking reagents (formaldehyde and EGS). Mix well and incubate for 2 minutes on the roller at room temperature.
8. Spin sample at 1,000 rpm for 5 minutes at 4°C.
9. Discard the supernatant and gently wash the pellet with 10 mL of 4°C PBS on ice.
10. Spin sample at 1,000 rpm for 5 minutes at 4°C.
11. Discard the supernatant and repeat the wash by gently resuspending the pellet with 10 mL of 4°C PBS on ice.
12. Spin sample at 1,000 rpm for 5 minutes at 4°C.
13. Discard the supernatant (see).
14. Place the washed pellet on ice.

3.1.2 Dynabead preparation

This section will require 5 mg/mL BSA solution and 4°C PBS. Work on ice.

Each sample requires three 1.5 mL DNA LoBind microcentrifuge tubes: one for the preclearing during section 3.1.2 Chromatin Immunoprecipitation step 1, one for the ChIP of the targeted protein (in this case, ATRX) and the last one for the IgG control (both the for chromatin immunoprecipitation during section 3.1.4 Chromatin Immunoprecipitation step 4).

1. To one 1.5 mL DNA LoBind microcentrifuge tube for the pre-clearing, add 100 µL of magnetic beads (Dynabeads Protein A or G depending on the antibodies used for the ChIP) (see).
2. To two 1.5 mL DNA LoBind microcentrifuge tubes, add 50 µL of magnetic beads (see). One of these will be for the ATRX immunoprecipitation (IP) and the other for the IgG control.
3. For each tube, wash the beads with 1 mL of 5 mg/mL BSA solution, quick spin.
4. Place the tube on the magnetic stand, wait for 3 minutes until the solution is clear and remove the supernatant.
5. Repeat the previous step three times (4 washes in total) (see).
6. Resuspend the 100 µL of washed beads for preclearing in 300 µL of 5 mg/mL BSA solution. Keep them at 4°C under constant agitation until needed.
7. For the ATRX IP: resuspend one of the 50 µL of washed beads in 300 µL of cold PBS containing 10 µg of ATRX antibody (see). Incubate for 3-4 hours on a roller at 4°C.
8. For IgG control: resuspend the other 50 µL of washed beads in 300 µL of cold PBS containing 10 µg of IgG antibody (see). Incubate for 3-4 hours on a roller at 4°C.
9. Before adding to the precleared cell lysate (see section 3.1.4 Chromatin Immunoprecipitation step 4), wash the beads incubated with the ATRX and IgG antibodies, 4 times in 5mg/mL BSA solution and resuspend the beads in 100 µL of PBS at 4°C.

3.1.3 Cell sonication

The cell sonication step is done at 4°C unless specified otherwise. Work on ice.

This is an essential step for a successful ATRX ChIP. It will require optimisation as the sonication settings are cell-type dependent (see) [40].

1. Add protease inhibitor cocktail (1X final) to all lysis buffers before use.
2. Chill centrifuge to 4°C.
3. If the cross-linked cell pellets were frozen at the step 13 of the section 3.1.1 Cell cross-linking, thaw the pellets on ice (see).
4. Resuspend each cell pellet in 10 mL of 4°C lysis buffer 1 containing 1X protease inhibitor cocktail.
5. Rock samples at 4°C for 10 minutes.
6. Centrifuge the samples at 2,000 rpm for 2 minutes at 4°C.
7. Discard supernatant and resuspend cells in 10 mL of 4°C lysis buffer 2 containing 1X protease inhibitor cocktail.
8. Centrifuge the samples at 2,500 rpm for 2 minutes at 4°C.
9. Discard supernatant and resuspend the pellet in 3 mL of 4°C lysis buffer 3 containing 1X protease inhibitor cocktail.
10. Syringe at least 10 times each sample with a G27 needle (see).
11. Divide each sample in three milliTUBE 1 mL tubes with AFA fiber for the Covaris (see and Figure 1).
12. Sonicate the samples to get fragments between 100 and 250 bp (Figure 2) (see and section 3.1.6 DNA purification step 12).
13. After sonication, transfer the sample from each 1 mL tube with AFA fiber into a 1.5 mL DNA LoBind microcentrifuge tube.

14. Add 1/10 volume (100 µL to each 1 mL of sonicated sample) of 10% Triton X-100.
15. Spin at 14,000 rpm for 10 minutes at 4°C.
16. Collect the supernatant and for each sample (divided during step 11) (Figure 1), combine the three 1 mL aliquots together into a graduated 14 mL falcon tube. As volume loss may occur during the sonication step, top up to 3 mL with 4°C lysis buffer 3 containing protease inhibitors if necessary.
17. Take 50 µL of sonicated chromatin from each sample as input and store this aliquot at -20°C (see).

3.1.4 Chromatin immunoprecipitation

The chromatin immunoprecipitation step is done at 4°C unless specified otherwise. Work on ice.

1. Preclear the lysate by adding 300 µL of the washed preclearing beads (section 3.1.2 Dynabead Preparation, step 5) to the 3 mL of cell lysate and incubating the samples on roller at 4°C for 1 hour.
2. Split the 3 mL precleared samples into two 1.5 mL DNA LoBind microcentrifuge tubes (Figure 1).
3. Precipitate the beads from the chromatin using a magnetic stand and collect one 1.5 mL fraction of supernatant for the ATRX IP and the second 1.5 mL fraction of supernatant for the IgG control, each into a new 5 mL falcon tube (Figure 1).
4. Add the washed beads conjugated with the ATRX antibody (from the section 3.1.2 Dynabead Preparation, step 8) to the tube for the ATRX IP containing the first 1.5 mL precleared lysate and add the washed beads conjugated with

the IgG (from the section 3.1.2 Dynabead Preparation, step 8) to the tube for the IgG control containing the second 1.5 mL precleared lysate (Figure 1).

5. Incubate overnight at 4°C on a roller.

3.1.5 Washing, eluting and reverse cross-linking

The washing steps are done at 4°C unless specified otherwise. Work on ice. It is recommended to perform the washing steps in a 4°C cold room.

1. Add protease inhibitor cocktail to the wash buffer to a final concentration of 1X.
2. Transfer each sample (lysates and beads) into a new 1.5 mL DNA LoBind microcentrifuge tube (the targeted chromatin is now bound to the beads).
3. Precipitate the beads using the magnetic stand and discard the supernatant.
4. Wash the beads by adding 1 mL of wash buffer containing 1X protease inhibitor cocktail at 4°C. Incubate for 5 minutes on the roller at 4°C. Spin down shortly. Precipitate the beads on the magnetic stand and discard the supernatant (see).
5. Repeat the above wash step four times (five washes in total) (see).
6. Wash the beads with 1 mL of PBS containing 1X protease inhibitor cocktail at 4°C. Precipitate the beads on the magnetic stand and discard the supernatant (see).
7. Elute the beads at room temperature by adding 250 µL of freshly made elution buffer. Incubate at room temperature for 15 minutes on the roller.
8. Precipitate the beads on the magnetic stand and collect the eluate in a new 1.5 mL DNA LoBind microcentrifuge tube.

9. Repeat the elution by resuspending the beads with another 250 µL of freshly made elution buffer. Using a ThermoMixer, incubate at 65°C for 20 minutes with mixing cycles of 40 seconds ON, 20 seconds OFF at 1,400 rpm.
10. Precipitate the beads on the magnetic stand, collect the eluate and combine it with the previous round of elution from step 8 (500 µL in total).
11. Reverse cross-link by adding 20 µL of 5 M NaCl to each 500 µL elute and incubate at 65°C for 4 hours (see).
12. Defrost the 50 µL input for each sample (from section 3.1.3 Cell sonication, step 17). Add 450 µL of freshly made elution buffer at room temperature. Reverse cross-link by adding 20 µL of 5 M NaCl to each 500 µL eluate and incubate at 65°C for 4 hours (see).

3.1.6 DNA purification

1. To each tube (ATRX IP, IgG control and input) add 20 µL of 1M Tris pH6.5, 10 µL of 0.5M EDTA and 2uL of 20 mg/mL Proteinase K and incubate for 1 hour at 45°C.
2. Add 500 µL of phenol-chloroform-isoamylalcohol (25:24:1, v/v/v), mix well by vortexing and spin at 13,000 rpm for 5 minutes at room temperature (see).
3. Collect the top phase containing the DNA in a new tube (see).
4. Add 1 mL of 100% EtOH, 50 µL of 3M Sodium Acetate and 1µL of Glycogen (20 µg). Incubate at -20°C overnight (see).
5. Spin down at 13,000 rpm for 1 hour at 4°C.
6. A tiny pellet should be visible. Discard the supernatant carefully without disturbing the pellet.

7. Wash the pellet by adding 1 mL of 70% EtOH, mix and spin at 13,000 rpm for 10 minutes at room temperature.
8. Carefully discard as much supernatant as possible and allow the pellet to dry for about 15 minutes (see).
9. For the ATRX ChIP and IgG control samples: resuspend the pellets in 20 µL of ultrapure water.
10. For the input sample: resuspend the pellet in 50 µL of ultrapure water.
11. Incubate at room temperature for 30 minutes to allow the DNA to go into solution. Mix and spin down.
12. On a 2% agarose gel, run 12 µL of input to check the sonication (Figure 2) (see).
13. Quantify the samples using the Qubit dsDNA HS Assay Kit with 1 µL of each sample (see).
14. Store the samples at -20°C.

3.2 ChIP validation by qPCR

In order to assess the ChIP quality, the ATRX ChIP enrichment over the IgG negative control has to be checked both at a locus enriched for ATRX (positive control) and a locus depleted for ATRX (negative control) (see). At all loci, the IgG ChIP enrichment should remain low, whereas the ATRX enrichment should display a significant increase at the positive locus and drop close to the IgG level at the negative locus (see and Figure 3). It is recommended to analyse the data as a percentage of input.

1. For each sample (ATRX, IgG and Input) take a 4 µL aliquot and dilute it in 36 µL of Sigma water in a new tube (1 in 10 dilution) (see). For each qPCR reaction, add 5 µL of diluted sample with 6.5 µL of PCR grade water, 0.5 µL of 10µM Forward primer, 0.5 µL of 10µM Reverse primer and 12.5 µL of Fast SYBR Green Master Mix (25 µL total).
2. Using a Real-Time PCR Detection System, first incubate the sample at 90°C for 20 seconds (holding stage), then for 40 cycles incubate at 95°C for 3 seconds followed by 30 seconds at 60°C for annealing/extension. Performing a melting curve step is also recommended for quality control.
3. If the samples passed all the validation steps (Figure 3), they can be used for ChIP-qPCR or sequencing.

3.3 Library preparation

Do not size select your sample for the library preparation. Only the ATRX ChIP and the input samples will be used for the following step. The IgG ChIP is mainly useful for the quality check by qPCR or for any further qPCR analysis if desired.

The library preparation is done using the NEBNext Ultra DNA Library Prep Kit for Illumina (E7370S/L) and the NEBNext Multiplex Oligos for Illumina (E7335S/L) following New England Biolabs's instructions, adapted to the ATRX ChIP as described below. This protocol will produce library to be sequenced on Illumina platform (here, we use the NextSeq500 as an example).

The manufacturer's instructions require at least 5 ng of ChIP purified DNA per sample. A single ATRX ChIP sample, however, is not expected to produce a sufficient yield to reach this threshold. Nevertheless, the 15 µL sample (after qPCR validation) should result in at least 2 ng of fragmented DNA. If it is the case, the sample can be processed as if a 5 ng sample and produce an acceptable library. Alternatively, up to three ATRX ChIP replicates (45 µL total) can be pooled and processed as a single 5 ng fragmented DNA sample.

3.3.1 DNA end preparation: DNA end repair, 5' phosphorylation and dA-Tailling

1. Dilute the 15 µL of ATRX ChIP sample (or more if combining replicates) with ultrapure water to give a final volume of 55.5 µL (see).
2. Dilute 5 ng of input sample (based on the measurement obtained at the section 3.1.6 DNA purification step 13) in a sufficient volume of ultrapure water to a final volume of 55.5 µL of diluted fragmented DNA.
3. To each 55.5 µL diluted fragmented DNA sample, add 3 µL of End Prep Enzyme Mix and 6.5 µL of End Repair Reaction Buffer (10X) for a final volume of 65 µl.
4. Mix by pipetting and quickly spin the samples.
5. Using a thermocycler, heat the samples for 30 minutes at 20°C and then for 30 minutes at 65°C. Hold the samples at 4°C until the next step.

3.3.2 Adaptor ligation

1. Near to the end of the 30 minutes incubation at 65°C, make a fresh 10-fold dilution of the NEBNext Adaptor for Illumina (15 µM stock solution) in 10 mM Tris-HCl to a final concentration of 1.5µM. Use immediately and discard any surplus.
2. To each 65 µL end prep sample, add 15 µL of Blunt/TA Ligase Master Mix, 2.5 µL of 1.5 µM NEBNext Adaptor for Illumina and 1 µL of ligation enhancer for a final volume of 83.5 µL.
3. Mix by pipetting and quickly spin the samples.
4. Using a thermocycler, incubate the samples for 15 minutes at 20°C.
5. Immediately add 3 µL of USER Enzyme to each 83.5 µL adapter ligated reaction.
6. Mix by pipetting and quickly spin the samples.
7. Incubate the samples for 15 minutes at 37°C (see).

3.3.3 Cleanup of adaptor-ligated DNA without size selection

1. Make a fresh solution of 80% ethanol by mixing 40 ml of 100% EtOH with 10 ml of ultrapure water.
2. Transfer the 86.5 µL of ligation reaction in a new 1.5 mL DNA LoBind microcentrifuge tube.
3. Add 86.5 µL of resuspended AMPure XP Beads (see).
4. Mix by pipetting 15 times and incubate at room temperature for 5 minutes.
5. Briefly spin the samples and precipitate the beads using a magnetic stand (see).
6. Remove the supernatant (see).

7. Wash the beads by adding 200 µL of fresh 80% ethanol. Incubate 30 seconds at room temperature and carefully discard the supernatant without disturbing the beads (during this step, do not remove the samples from the magnetic stand).
8. Repeat the previous step (two washes in total).
9. Quickly spin the tube and precipitate the beads on the magnetic stand and remove any residual ethanol (without removing the tubes from the magnet).
10. Allow the beads to air dry for 5 minutes (cap opened) (see).
11. Remove the tube from the magnet and add 17 µL of 10 mM Tris-HCl to elute the DNA target. Mix by pipetting 15 times and incubate at room temperature for 2 minutes.
12. Briefly spin the tube and precipitate the beads on the magnetic stand (see).
13. Carefully transfer 15 µL of the supernatant to a new PCR tube (without disturbing the beads) for amplification.

3.3.4 PCR enrichment of adaptor ligated DNA

1. For each 15 µL of cleaned adapted ligated DNA reaction, add 25 µL of NEBNext Q5 Hot Start HiFi PCR Master Mix (see), and 5 µL of 10 µM Universal PCR Primer.
2. Add 5 µL of the appropriate 10 µM Index Primer for a final volume of 50 µL (see).
3. Using a thermocycler, incubate the samples for 30 seconds at 98 °C (initial denaturation) followed by twelve cycles of denaturation at 98°C for 10 seconds and annealing/extension at 65°C for 75 seconds. After twelve cycles

of denaturation and annealing/extension, incubate at 65°C for 5 minutes. Hold the reaction at 4°C.

3.3.5 Cleanup of PCR amplification

1. Transfer the 50 µL of PCR amplified reaction to a new 1.5 mL DNA LoBind microcentrifuge tube.
2. Add 45 µL of resuspend AMPure XP Beads (see).
3. Mix by pipetting 15 times and incubate at room temperature for 5 minutes.
4. Briefly spin the samples and precipitate the beads using a magnetic stand (see).
5. Remove the supernatant (see).
6. Wash the beads by adding 200 µL of fresh 80% ethanol. Incubate for 30 seconds at room temperature and carefully discard the supernatant without disturbing the beads (by keeping the samples on the magnetic stand).
7. Repeat the previous step (two washes in total).
8. Briefly spin the tube and precipitate the beads on the magnetic stand and remove any left over of ethanol (without removing the tubes from the magnet).
9. Allow the beads to air dry for 5 minutes (cap opened) (see).
10. Remove the tube from the magnet and add 33 µL of 0.1X TE to elute the DNA target.
11. Mix by pipetting 15 times and incubate at room temperature for 2 minutes.
12. Briefly spin the tube and precipitate the beads on the magnetic stand (see).
13. Carefully transfer 28 µL of supernatant (library) to a new 1.5 mL DNA LoBind microcentrifuge tube without disturbing the beads.

14. Dilute 2 µL of the library in 2 µL of 0.1X TE and check the size distribution of the library on the TapeStation using a High Sensitivity D1000 ScreenTape (Figure 4).
15. Store the library at -20°C.

3.4 Library quantification and sequencing

An accurate quantification of the libraries is essential for a successful sequencing.

1. Quantify the library using the Qubit dsDNA HS Assay Kit.
2. Based on the Qubit quantification and the average fragment size of the library (obtained in section 3.3.5 Cleanup of PCR amplification, step 14), make 15 µL of 12 nM library from the stock library and store the rest at -20°C (see).
3. Quantify the 12 nM library using the KAPA library Quantification kit by making a 1:1,000 and a 1:2,000 dilution in 10 mM Tris-HCl (pH 8.0), 0.05% Tween 20.
4. Based on the KAPA library Quantification, make 5 µL of 4 nM library and store the rest at -20°C.
5. If several libraries are sequenced together, pool the 4 nM libraries with different indexes together (see).
6. Use a NextSeq 500/550 High Output Kit v2, 150 cycles (or Mid Output Kit v2, 150 cycles depending on the number of samples) and sequence using the NextSeq500 from Illumina. It is recommended to have a sequencing depth of at least 20 millions uniquely mapped reads per sample with an input coverage similar to or higher than its related ATRX ChIP-seq sample [37] (see). A

successful ATRX ChIP-seq should display a significant enrichment above background and a very low input background (Figure 5).

3.5 Analysis of the ChIP-seq data

Each ChIP-seq experiment generates a large amount of data that is not trivial to analyse and bioinformatic expertise is essential for a comprehensive analysis. Here, we attempt to provide some basic guidance for this process. It should be mentioned, however, that there is no single way to perform ChIP-seq data analysis and the tools mentioned below are not exhaustive.

When processing ChIP-seq data, few essential steps have to be considered:

Analysis of the raw sequencing data

1. Mapping the sequences to the reference genome: A widely used tool for mapping is Bowtie (1 and 2) [41,42] but other mapper such as BWA [43] could be considered. Several versions of a given reference genome may be available (for instance, the human genome - hg18 and hg19) and it is essential to map all the replicates to the same version of the genome. Additional parameters, such sequence variations (e.g. SNP) with the reference genome should be considered [36]. In the context of the ATRX ChIP-seq data analysis, in which a large number of sequences are known to be repetitive, subsequent analysis of enrichment at repetitive elements for individual repeat types [44] or using the REPEATOME [45] (if applicable) should be considered.
2. Trimming of the adapter sequences: As mentioned during library preparation, Illumina adapters were added to the sample prior to sequencing (see section

3.3.2 Adaptor ligation) and, subsequently, during the data analysis, residual adapter sequences could impair the mapping analysis. Therefore, depending on the mapper used, the trimming of the sequences should be considered using tools such as trim-galore [46].

3. Assess the data quality at each stage of the process: the generation of quality control reports such as fastQC reports [47] is a good practice for each step of the data processing and mapping analysis. Further quality controls such as cross-contamination check using tools such as FastQ Screen (http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/. Accessed 18 July 2017) could be relevant.
4. Remove the PCR duplicates and retain only the unique reads mapping to a single location: PCR duplicates are artefacts generated during the library amplification step, and may create artificial enrichments. They should not be considered in subsequent analysis such as enrichment analysis and should be removed using tools such as Samtools [48] and Picard (http://broadinstitute.github.io/picard/. Accessed 18 July 2017)
5. Remove the signal artefact blacklisted regions [49]: these regions produce artefact signals with high signal/read counts due to sequence mappability shortcoming and should not be considered as biologically relevant.
6. Visualise the data on a genome browser such as UCSC [50] (Figure 5).

Analysis of the ChIP-seq data per se.

1. Assess the quality of the ChIP-seq using the ATRX ChIP and Input samples (e.g. assess the coverage, the enrichment of the ChIP over input) using packages as Deeptools [51].

2. Data normalisation: The data normalisation such as correction for the sequencing depth or GC bias may be critical for downstream analysis. Versatile packages such as Deeptools [51] and Homer [52] could be useful. Depending on the types of samples (e.g. different genders), it could be relevant to take mitochondrial and/or sex chromosomes mapping reads out of the normalisation of the ChIP signals.
3. Peak calling: Central goal of a ChIP-seq experiment, the enrichment analysis characterising the genomic regions targeted by the protein of interest can be supported by a vast numbers of packages including MACS [53] and HOMER [52]. In the context of ATRX ChIP-seq, the peak calling is relatively challenging due to the complexity of the ATRX enrichment (diversity of peak shapes and density across the genome) (Figure 5). Fine-tuning of the peak caller parameters and the use of an input are required to achieve a high quality peak calling. It is essential to visually assess the peak calling using tools like MIG [54].
4. Assess for reproducibility: for a robust analysis, it is essential to produce biological ChIP-seq replicates and assess their reproducibility (Figure 5).

4 Notes

1. As EGS is water-insoluble, it has to be first dissolved in DMSO before being added to aqueous solution. It is essential to prepare the 500 mM EGS stock solution just before using it and discard any leftover as DMSO promotes the hydrolysation of the EGS NHS ester moiety which becomes non-reactive (see manufacturer's instructions for Thermo Scientific EGS, cat. 21565).

2. Successful ATRX ChIP-seqs have been performed using the ab97508 anti-ATRX antibody (Rabbit polyclonal) from Abcam or the H300 anti-ATRX antibody (Rabbit polyclonal) from Santa Cruz Biotechnology. These antibodies could be substituted by antibodies targeting proteins requiring a long range cross-linking. It cannot be guaranteed, however, that the substitute antibodies could produce a successful ChIP-seq following this protocol and further optimisation may be required. Furthermore, it is important to pay attention to the protein isoform and select the most appropriate antibody based on the experimental design. In the case of ATRX, the selective targeting of the full-length protein requires an antibody raised against the C-terminal end. Always take note of the antibody batch, as the efficiency can be batch dependant (especially with polyclonal antibodies). Use the IgG antibody from the same species as the one used to raise the antibody for the protein of interest. Use rabbit IgG in the case of the ATRX ChIP if the antibody used is ab97508 anti-ATRX antibody (Rabbit polyclonal) from Abcam or the H300 anti-ATRX antibody (Rabbit polyclonal) from Santa Cruz Biotechnology.

3. To dissolve EGTA powder, the solution needs to be adjusted at pH 8 using NaOH.

4. These loci may be cell type dependent. However, rDNA is known as an ATRX target in several human cell types and may be a good positive control to test [21].

Potential primer pairs for the qPCR validation of ChIPs performed on human derived samples:

ATRX ChIP positive control for human-derived sample (rDNA)

Positive-hg-F: TTCAAAGCCCCATTCGTATGC

Positive-hg-R: AGTTTTCAGCCCCAACACACC

ATRX ChIP negative control for human-derived sample (RHBDF1)

Negative-hg-F: GAGATGCTGGAGTCAGGACCAT

Negative-hg-R: AGGAGTCAGGAGCAGCAGTCA

5. It is critical that the PBS temperature is above 20°C (up to 30°C) as cold PBS could affect the cross-linking efficiency. For an efficient cross-linking, it is essential that the samples are in single cell suspension prior to the addition of the crosslinking agent. Upon addition of the 500 mM EGS stock solution, the solution will become cloudy and clear again very quickly. After the EGS dispersion, no EGS precipitate should be visible. If some EGS precipitates once the solution is clear, check that the PBS was at room temperature.

6. Pre-cool the centrifuge at 4°C for step 8 of the section 3.1.1 Cell cross-linking.

7. Use the laminar flow cabinet to add the formaldehyde. Do not store an open bottle of 38% formaldehyde solution for a long period of time (usually no more than 6 months to a year).

8. This is a safe stopping point: samples can be snap frozen on dry ice and ethanol and stored at -80°C for up to three days. However, do not store the cross-linked pellets for a longer period of time as it may affect the sonication efficiency (see Figure 2) [55].

9. The choice between the Dynabeads protein A and G is dependent on the affinity of the protein A/G for the IgG types from different species. Choose the Dynabeads protein A if using the ab97508 anti-ATRX antibody (Rabbit polyclonal) from Abcam or the H300 anti-ATRX antibody (Rabbit polyclonal) from Santa Cruz Biotechnology. Alternatively, a mix of Dynabeads protein A and G can be used. If this protocol is adapted to another target protein requiring long range crosslinking, check the species in with the antibody has been raised as well as the IgG type to select

the most appropriated Dynabeads protein type (for further information, refer to: NEB, Tools & Resources, Affinity of Protein A/G for IgG Types from Different Species: https://www.neb.com/tools-and-resources/selection-charts/affinity-of-protein-ag-for-igg-types-from-different-species. Accessed 9 Jun 2017).

10. Do not let the beads dry in between washes.

11. As the chromatin undergoes a double cross-linking (first long range cross-linking using EGS followed by a second cross-linking using formaldehyde), essential for an efficient and reliable ATRX ChIP, it became more resistant to sonication. More stringent sonication parameters are therefore required to achieve similar results than with a single cross-linked sample using only 1% formaldehyde for 10 minutes (as used in standard single cross-link protocol [13,14]. It is strongly recommended to optimise the sonication for each cell type. Be careful to not over-sonicate or under-sonicate the DNA (see Figure 2) as this may affect the quality of the ChIP-seq. Possible settings to try as a starting point: Duty = 5% and Intensity = 5 at 4°C for 18 to 32 min, Duty = 10% and Intensity = 5 at 4°C for 18 to 32 min, Duty = 12.5% and Intensity = 5.2 at 4°C for 18 to 32 min. For each setting, assessing the sonication after 18, 22, 25, 28 and 32 min should be informative.

12. Syringing the sample helps to improve the sonication efficiency.

13. The Covaris has been chosen for the sonication step as it offers a significant flexibility of settings to design an optimal programme for each cell type as well as a good reproducibility. Alternatively, the sonication could also be carried out using a Bioruptor.

14. The input is a very important control. It is used to check the sonication of the sample, correct for the percentage input during the ChIP-qPCR experiments and show the background in the ChIP-seq data.

15. Alternatively, the washes can be done using the wash buffers from the Chromatin Immunoprecipitation Assay Kit from Merck (cat. 17-295) after adding the protease inhibitor cocktails to 1X final concentration. In that case, the first wash is done with 1 mL of Low Salt Immune Complex wash buffer; the second wash is done with 1 mL of High Salt Immune Complex wash buffer; the third wash is done with 1 mL of LiCl Immune Complex wash buffer and the two last washes are done with 1 mL of TE buffer. After the second wash with the TE buffer, skip the steps 4, 5 and 6 of the section 3.1.5 Washing, eluting and reverse cross-linking and resume at the step 7.

16. Alternatively, the sample can be incubated overnight.

17. Take extra care to avoid contact with phenol and use the laminar flow cabinet. Alternatively, the Qiagen MinElute PCR Purification kit (cat. 28004) may be used. However, note that DNA purification following this method will not retain the fragments smaller than 70 bp and may alter some of the qPCR analysis. This alternative is only recommended if the sequencing input displays a strong background due to a failure to generate an optimal sonication (presence of a significant fraction of oversonicated fragments) (Figure 2). In these circumstances, using the Qiagen minElute purification kit will allow removing the over sonicated fragments and decrease the background generated during the library preparation and sequencing.

18. Be careful not to contaminate the top phase by touching the middle interphase.

19. At this stage, the samples can stay at -20°C for up to 48h. Alternatively, quick precipitation can also be carried out at -80°C for at least 4 hours.

20. Once dry the pellet may become transparent and almost invisible. Care must be taken to not overdry the pellet, as it will be difficult to dissolve.

21. A successful sonication should appear as a smear from 100 to 250 bp without significant over or under sonication (Figure 2). Further quality checks can be done by running 1 µL of input on the TapeStation using a D1000 tape.

22. ATRX ChIP sample should be quantifiable and give at least 140 ng/mL, IgG ChIP sample may not be quantifiable (reading: "too low < 0.5 ng/mL"), the input should be > 50,000 ng/mL).

23. If not the case, there is a high probability that the ChIP failed (Figure 3). One common issue may come from the batch of antibody. In that case, repeat the ChIP experiment with a new batch of antibody. In addition, ensure that the sonication is optimal and that the cell viability was > 90%. If the issue persists, use a fresh solution of formaldehyde and make fresh buffers. Alternative positive/negative loci can be tested.

24. This is enough to run each qPCR experiment in triplicate.

25. If several ATRX ChIP samples are combined, the volume of ultrapure water to add should be adjusted to reach a final volume of 55.5 µL of fragmented DNA.

26. A potential presence of a precipitate may be observed in the NEBNext Q5 Hot Start HiFi PCR Master Mix. Equilibrate the NEBNext Q5 Hot Start HiFi PCR Master Mix at room temperature and mix gently by inverting the tube to ensure an optimal performance.

27. Be sure to thoroughly resuspend the AMPure XP Beads by vortexing before use.

28. Incubate the sample for 5 minutes on the magnetic stand until the solution is clear and the beads are fully precipitated.

29. Be cautious and do not disturb the beads. If some beads were aspirated along with the supernatant, pipet down the sample and wait another 3 min to allow the beads to precipitate, then remove the supernatant carefully (with the tubes on the magnetic stand).

30. Be careful and do not over dry the beads as it may decrease yield of recover DNA target.

31. Each index number will give a different barcode. Using a different index number for each individual sample allows sequencing these samples together as a pool and extract the data for each individual sample after sequencing using the unique barcode associated to each DNA fragment. It is important to think in advance which samples will be pooled together to judiciously select the appropriate indexes. For best results, it is recommended to use indexes 6 and 12 when only two samples are pooled. Use indexes 4, 6 and 12 for a pool of three samples and the indexes 2,4,5,6,7 and 12 for a pool of six samples.

32. Making a pre-dilution to 12nM on a fraction of the stock library before the KAPA quantification allows:

- A better comparison between different libraries
- A more accurate dilution to a 4 nM library as the dilution factor will be decreased compared with the stock library

- Ensuring to be in the range of the KAPA standard by doing 1:1,000 and 1:2,000 dilutions for the KAPA quantification

The Qubit quantification is expressed as ng/µL. To obtain the molarity of the library, the following calculation is needed:

Library molarity (M) = Library concentration (g/L) / (Average fragment size of the library (bp) * Average molecular weight of 1 bp)

Then, dilute an aliquot of the stock library to obtain 15 µl of 12 nM dilute library:

Volume of stock library aliquot (µL) = 15 x 12 / stock library molarity (nM)

Add 0.1X TE to the stock library aliquot to obtain a final volume of 15 µL.

33. As ATRX binds to repeats, generating longer reads (compared to the 75-cycle kit) increases the probability to map single reads overlapping with repeats. The libraries prepared with NEBNext Ultra DNA Library Prep Kit for Illumina (E7370S/L) and the NEBNext Multiplex Oligos for Illumina are compatible with the Illumina sequencing platform. In this protocol we use the example of the NextSeq500 but other Illumina platforms can be used. The sequencing coverage is an essential factor that can influence peak calling. Sufficient sequencing depths for ChIP sample and its related input are required for good quality peak calling. The sequencing depth of the input should be similar to or higher than its related ChIP-seq sample [37].

5 Acknowledgements

We would like to thanks David Clynes and Thomas Kent for a critical reading of this manuscript. This work was supported by the Medical Research Council [grant number MC_UU_12025/ unit programme MC_UU_12009/3].

6 References

1. Clynes D, Higgs DR, Gibbons RJ (2013) The chromatin remodeller ATRX: a repeat offender in human disease. Trends in biochemical sciences 38 (9):461-466
2. Picketts DJ, Higgs DR, Bachoo S et al. (1996) ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. Human molecular genetics 5 (12):1899-1907
3. Steensma DP, Higgs DR, Fisher CA et al. (2004) Acquired somatic ATRX mutations in myelodysplastic syndrome associated with α thalassemia (ATMDS) convey a more severe hematologic phenotype than germline ATRX mutations. Blood 103 (6):2019-2026
4. Kuo M-H, Allis CD (1999) In vivo cross-linking and immunoprecipitation for studying dynamic protein: DNA associations in a chromatin environment. Methods 19 (3):425-433
5. Landt SG, Marinov GK, Kundaje A et al. (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome research 22 (9):1813-1831
6. Orlando V, Strutt H, Paro R (1997) Analysis of Chromatin Structure byin VivoFormaldehyde Cross-Linking. Methods 11 (2):205-214
7. Park PJ (2009) ChIP–seq: advantages and challenges of a maturing technology. Nature Reviews Genetics 10 (10):669-680
8. Collas P, Dahl JA (2008) Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. Front Biosci 13 (17):929-943
9. Ren B, Robert F, Wyrick JJ et al. (2000) Genome-wide location and function of DNA binding proteins. Science 290 (5500):2306-2309
10. Lee TI, Johnstone SE, Young RA (2006) Chromatin immunoprecipitation and microarray-based analysis of protein location. Nature protocols 1 (2):729
11. Furey TS (2012) ChIP–seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. Nature Reviews Genetics 13 (12):840-852
12. Consortium EP (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489 (7414):57-74
13. O'Geen H, Echipare L, Farnham PJ (2011) Using ChIP-seq technology to generate high-resolution profiles of histone modifications. Epigenetics Protocols:265-286
14. O'Geen H, Frietze S, Farnham PJ (2010) Using ChIP-seq technology to identify targets of zinc finger transcription factors. Engineered Zinc Finger Proteins: Methods and Protocols:437-455
15. Gavrilov A, Razin SV, Cavalli G (2014) In vivo formaldehyde cross-linking: it is time for black box analysis. Briefings in functional genomics:elu037
16. Tian B, Yang J, Brasier AR (2012) Two-step cross-linking for analysis of protein–chromatin interactions. Transcriptional Regulation: Methods and Protocols:105-120
17. Zeng P, Vakoc CR, Chen Z et al. (2006) In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. Biotechniques 41 (6):694

18. Elsässer SJ, Noh K-M, Diaz N et al. (2015) Histone H3. 3 is required for endogenous retroviral element silencing in embryonic stem cells. Nature 522 (7555):240-244
19. Goldberg AD, Banaszynski LA, Noh K-M et al. (2010) Distinct factors control histone variant H3. 3 localization at specific genomic regions. Cell 140 (5):678-691
20. Lewis PW, Elsaesser SJ, Noh K-M et al. (2010) Daxx is an H3. 3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proceedings of the National Academy of Sciences 107 (32):14075-14080
21. Law MJ, Lower KM, Voon HPJ et al. (2010) ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner. Cell 143 (3):367-378. doi:http://dx.doi.org/10.1016/j.cell.2010.09.023
22. Nguyen DT, Voon HPJ, Xella B et al. (2017) The chromatin remodelling factor ATRX suppresses R‐loops in transcribed telomeric repeats. EMBO reports 18 (6):914-928
23. Eustermann S, Yang J-C, Law MJ et al. (2011) Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. Nature structural & molecular biology 18 (7):777-782
24. Nan X, Hou J, Maclean A et al. (2007) Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. Proceedings of the National Academy of Sciences of the United States of America 104 (8):2709-2714. doi:10.1073/pnas.0608056104
25. Gibbons RJ, Picketts DJ, Villard L et al. (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α-thalassemia (ATR-X syndrome). Cell 80 (6):837-845
26. Lovejoy CA, Li W, Reisenweber S et al. (2012) Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. PLoS genetics 8 (7):e1002772
27. Schwartzentruber J, Korshunov A, Liu X-Y et al. (2012) Driver mutations in histone H3. 3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482 (7384):226-231
28. Voon HP, Hughes JR, Rode C et al. (2015) ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci and Imprinted Genes. Cell reports 11 (3):405-418
29. Gibbons RJ, McDowell TL, Raman S et al. (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nature genetics 24 (4):368-371
30. Clynes D, Jelinska C, Xella B et al. (2014) ATRX dysfunction induces replication defects in primary mouse cells. PloS one 9 (3):e92915
31. De La Fuente R, Baumann C, Viveiros MM (2011) Role of ATRX in chromatin structure and function: implications for chromosome instability and human disease. Reproduction (Cambridge, England) 142 (2):221-234. doi:10.1530/rep-10-0380
32. Huh MS, O'Dea TP, Ouazia D et al. (2012) Compromised genomic integrity impedes muscle growth after Atrx inactivation. The Journal of clinical investigation 122 (12):4412
33. Ritchie K, Seah C, Moulin J et al. (2008) Loss of ATRX leads to chromosome cohesion and congression defects. The Journal of cell biology 180 (2):315-324

34. De La Fuente R, Viveiros MM, Wigglesworth K et al. (2004) ATRX, a member of the SNF2 family of helicase/ATPases, is required for chromosome alignment and meiotic spindle organization in metaphase II stage mouse oocytes. Developmental biology 272 (1):1-14
35. Teytelman L, Özaydın B, Zill O et al. (2009) Impact of chromatin structures on DNA processing for genomic analyses. PloS one 4 (8):e6700
36. Meyer CA, Liu XS (2014) Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. Nature Reviews Genetics 15 (11):709-721
37. Chen Y, Negre N, Li Q et al. (2012) Systematic evaluation of factors influencing ChIP-seq fidelity. Nature methods 9 (6):609-614
38. Drané P, Ouararhni K, Depaux A et al. (2010) The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3. 3. Genes & development 24 (12):1253-1265
39. Pchelintsev NA, McBryan T, Rai TS et al. (2013) Placing the HIRA histone chaperone complex in the chromatin landscape. Cell reports 3 (4):1012-1019
40. Kidder BL, Hu G, Zhao K (2011) ChIP-Seq: technical considerations for obtaining high-quality data. Nature immunology 12 (10):918-922
41. Langmead B, Trapnell C, Pop M et al. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10 (3):R25
42. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature methods 9 (4):357-359
43. Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26 (5):589-595
44. Day DS, Luquette LJ, Park PJ et al. (2010) Estimating enrichment of repetitive elements from high-throughput sequence data. Genome biology 11 (6):R69
45. Woo T-H, Hong T-H, Kim S-S et al. (2007) REPEATOME: A Database for Repeat Element Comparative Analysis in Human and Chimpanzee. Genomics & Informatics 5 (4):179-187
46. Krueger F (2015) Trim Galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files
47. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data.
48. Li H, Handsaker B, Wysoker A et al. (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25 (16):2078-2079
49. Kundaje A (2013) A comprehensive collection of signal artifact blacklist regions in the human genome. ENCODE[hg19-blacklist-README doc-EBI] Available online at: https://sites google com/site/anshulkundaje/projects/blacklists
50. Kent WJ, Sugnet CW, Furey TS et al. (2002) The human genome browser at UCSC. Genome research 12 (6):996-1006
51. Ramírez F, Dündar F, Diehl S et al. (2014) deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research 42 (W1):W187-W191
52. Heinz S, Benner C, Spann N et al. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell 38 (4):576-589
53. Zhang Y, Liu T, Meyer CA et al. (2008) Model-based analysis of ChIP-Seq (MACS). Genome biology 9 (9):R137

54. McGowan SJ, Hughes JR, Han Z-P et al. (2013) MIG: Multi-Image Genome viewer. Bioinformatics 29 (19):2477-2478
55. Bortz PDS, Wamhoff BR (2011) Chromatin immunoprecipitation (ChIP): revisiting the efficacy of sample preparation, sonication, quantification of sheared DNA, and analysis via PCR. PloS one 6 (10):e26015

Figure 1: Flowchart highlighting the main steps of the ATRX ChIP protocol. The beads prepared during the section 3.1.2 Dynabead preparation step are highlighted by (*).

Figure 2: Schematic example of a sonication control on a 2% agarose gel. Lane 1: Efficient sonication done on a cross-linked sample that has been frozen for less than 48 hours. The majority of the DNA fragments are comprised between 100 and 300 bp. Lane 2: Example of an over-sonicated sample requiring less stringent sonication settings (e.g. shorter sonication time). Lane 3: Example of an under-sonicated sample requiring more stringent sonication conditions (e.g. increase sonication time and/or intensity). Lane 4: Artefact example often observed with sonicated cross-linked sample that has been frozen for weeks before sonication. Despite the formation of a fraction of properly sonicated fragments (between 100 and 300 bp), note the presence of both over-sonicated and under-sonicated chromatin.

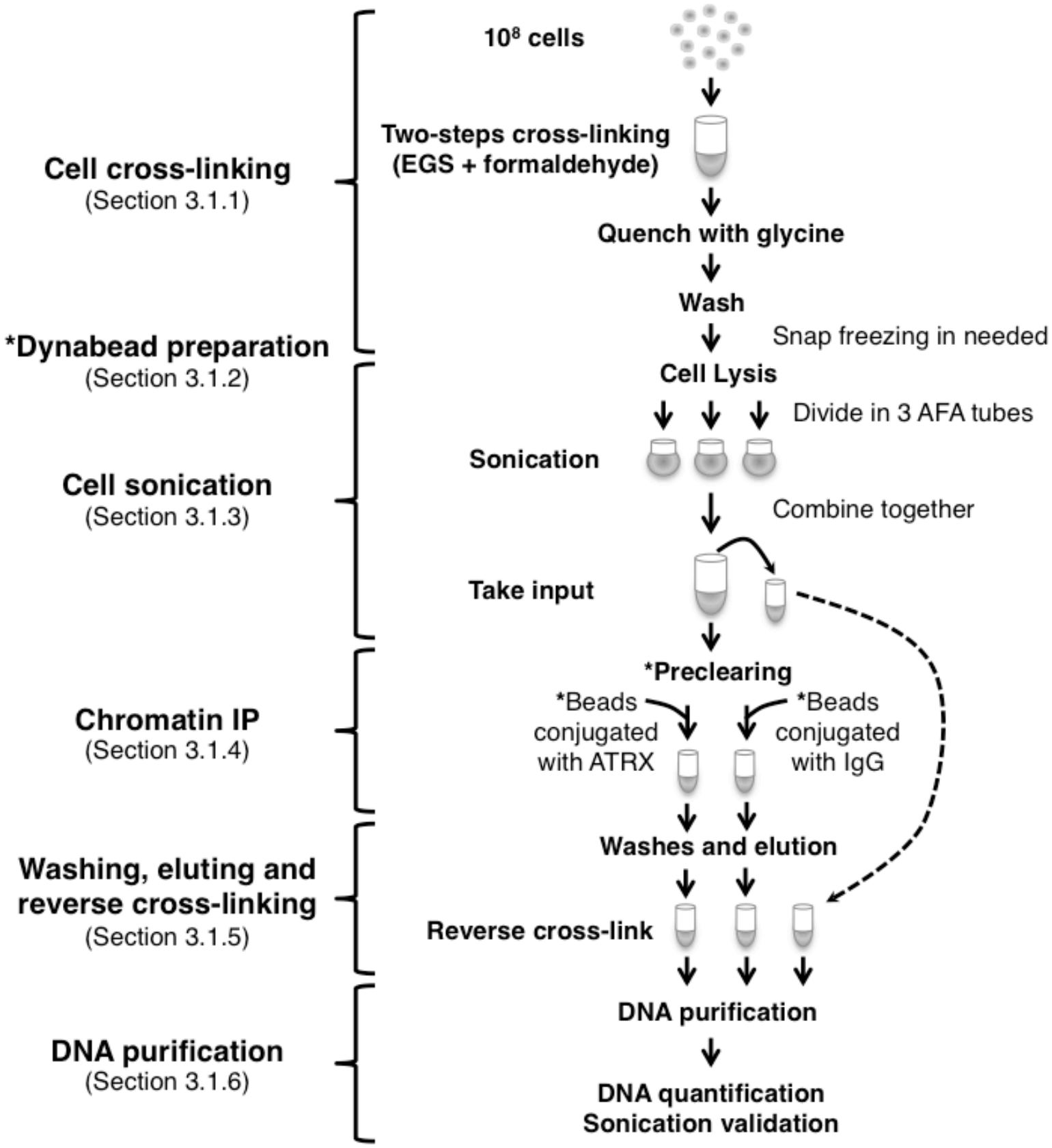
Figure 3: Examples of ChIP validation by qPCR. a) and b) Successful ChIP experiments showing a strong ATRX enrichment at the positive control and a basal level at the negative control. Note that the ChIP efficiency (%Input) can vary between experiments (usually between 0.02 and 0.06 % input for a successful ChIP). The IgG control, however, remains very low at both loci. c) Example of a high IgG

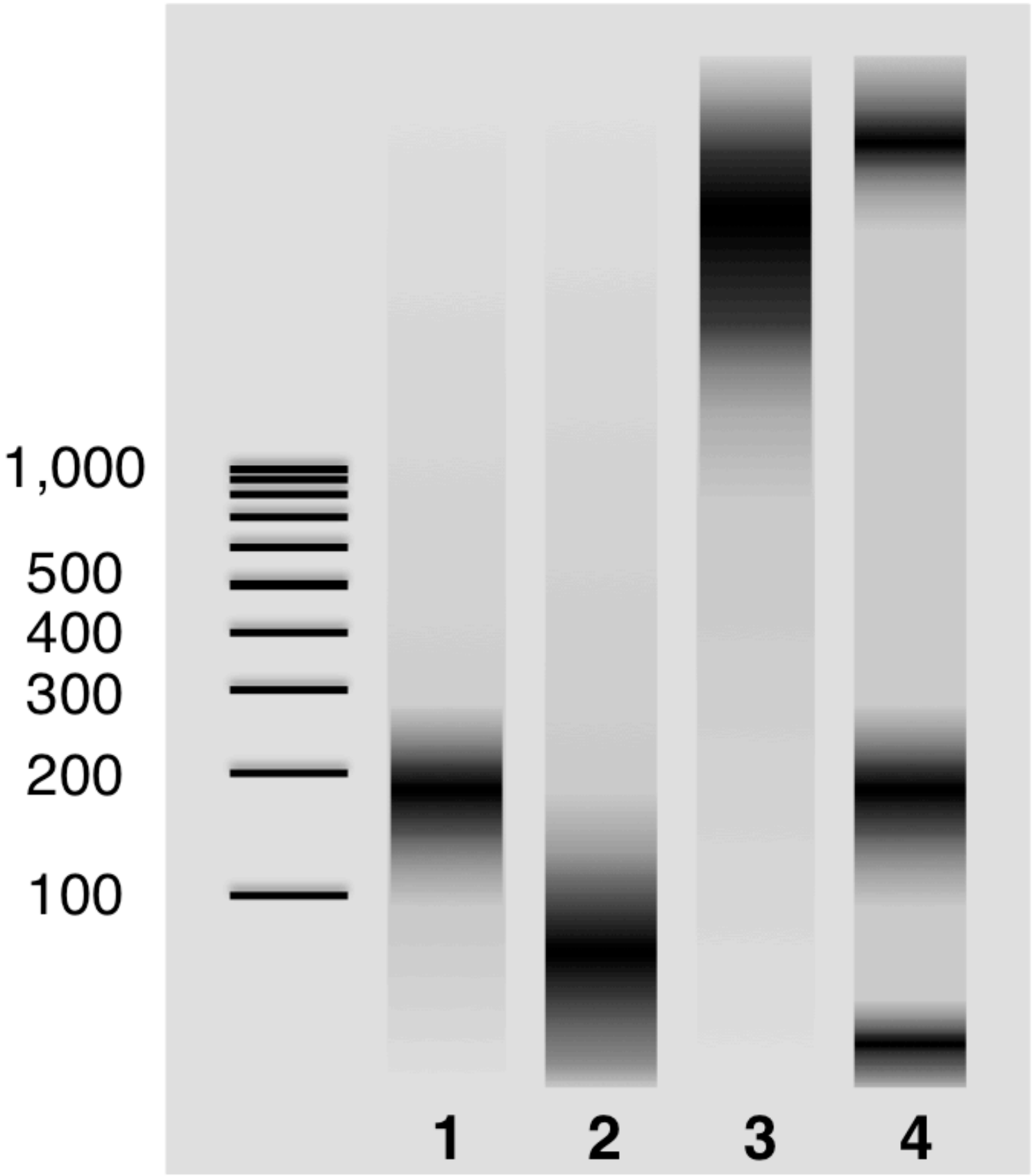
background, which may result in a very high background in the ChIP-seq data. The ChIP efficiency is very low. The batch of antibody used may not display a high avidity. The ChIP yield was too low (< 0.5 ng/mL) to perform a library preparation (see). d) Example of a batch of antibody displaying a low specificity for the protein of interest. The ChIP-seq data may have a high background and may not be interpretable (as shown in Figure 5d). In both c) and d) examples, it is strongly advised to re-start the ChIP instead of pursuing the library preparation. All ChIP experiments were done using lymphoblastoid cell lines and different batches of antibody. Positive: rDNA, Negative: RHBDF1 (see).

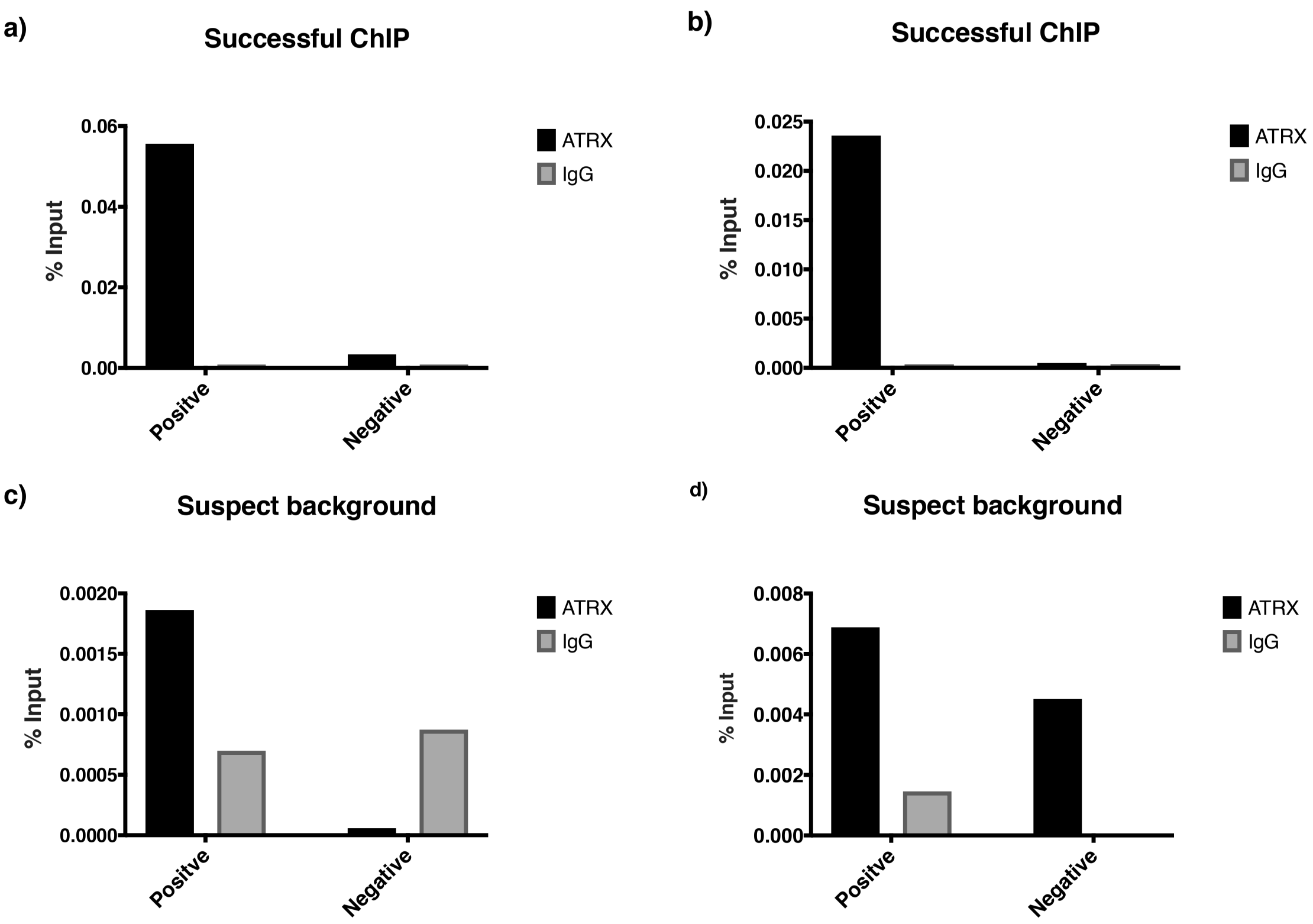
Figure 4: Example of ATRX library run on the TapeStation using a High Sensitivity D1000 ScreenTape. The library was made from 2.5 ng of ATRX ChIP material from a lymphoblastoid cell line.

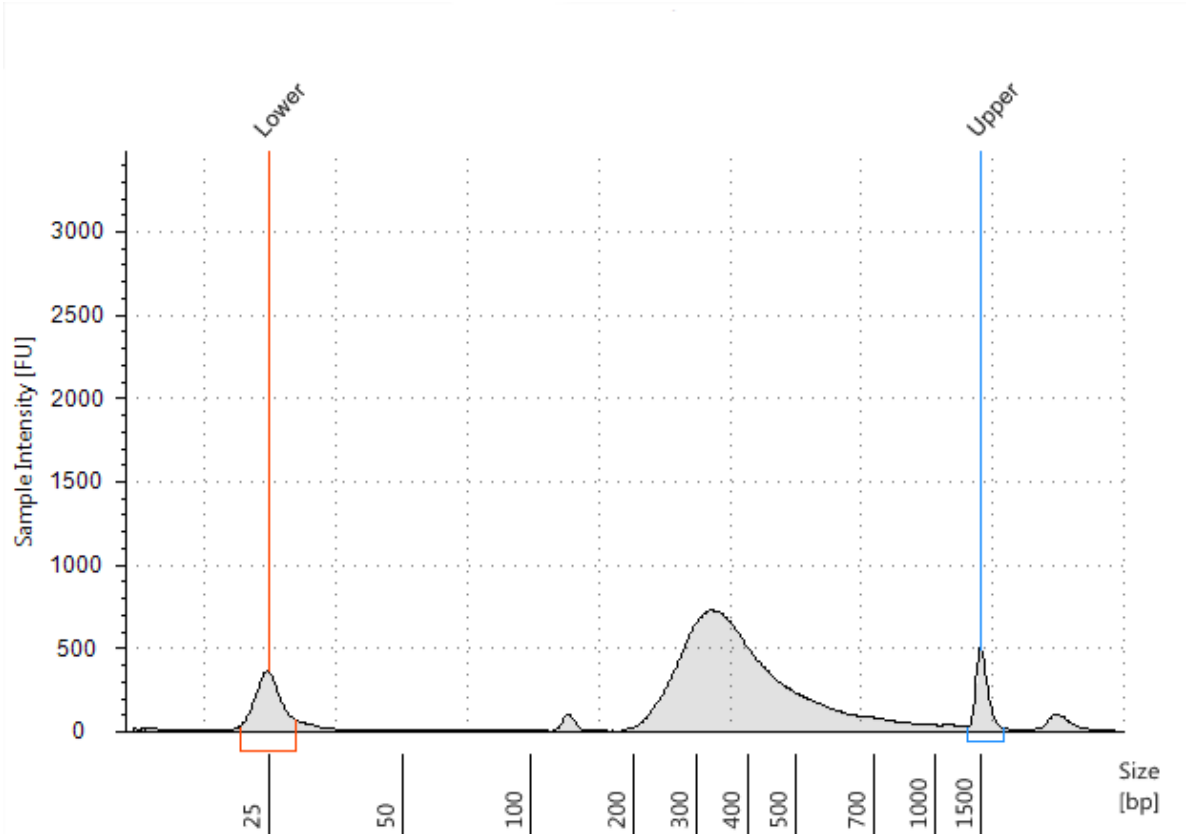
Figure 5: Examples of ATRX ChIP-seq data visualised on the UCSC genome browser. ATRX ChIP-seqs were carried out using lymphoblastoid cell lines and mapped on hg19 using Bowtie2, only the unique reads mapping to a single location were retained. a) and b) Examples of ATRX enrichment over input highlighting successful ChIP-seqs. Notice the variability in peak shapes. c) Example of a ChIP-seq artefact. The enrichment in the ATRX track is probably not real based on the background observed in the input track. d) Example of antibody batch effect on the quality of the ChIP-seq experiment. The two ATRX ChIP-seq data done on the same sample, following the same protocol using two different batches of ATRX antibody. The top track was produced with a poor batch of ATRX antibody as observed in Figure 3d. The enrichment is poor and the background is high. e) Example of an

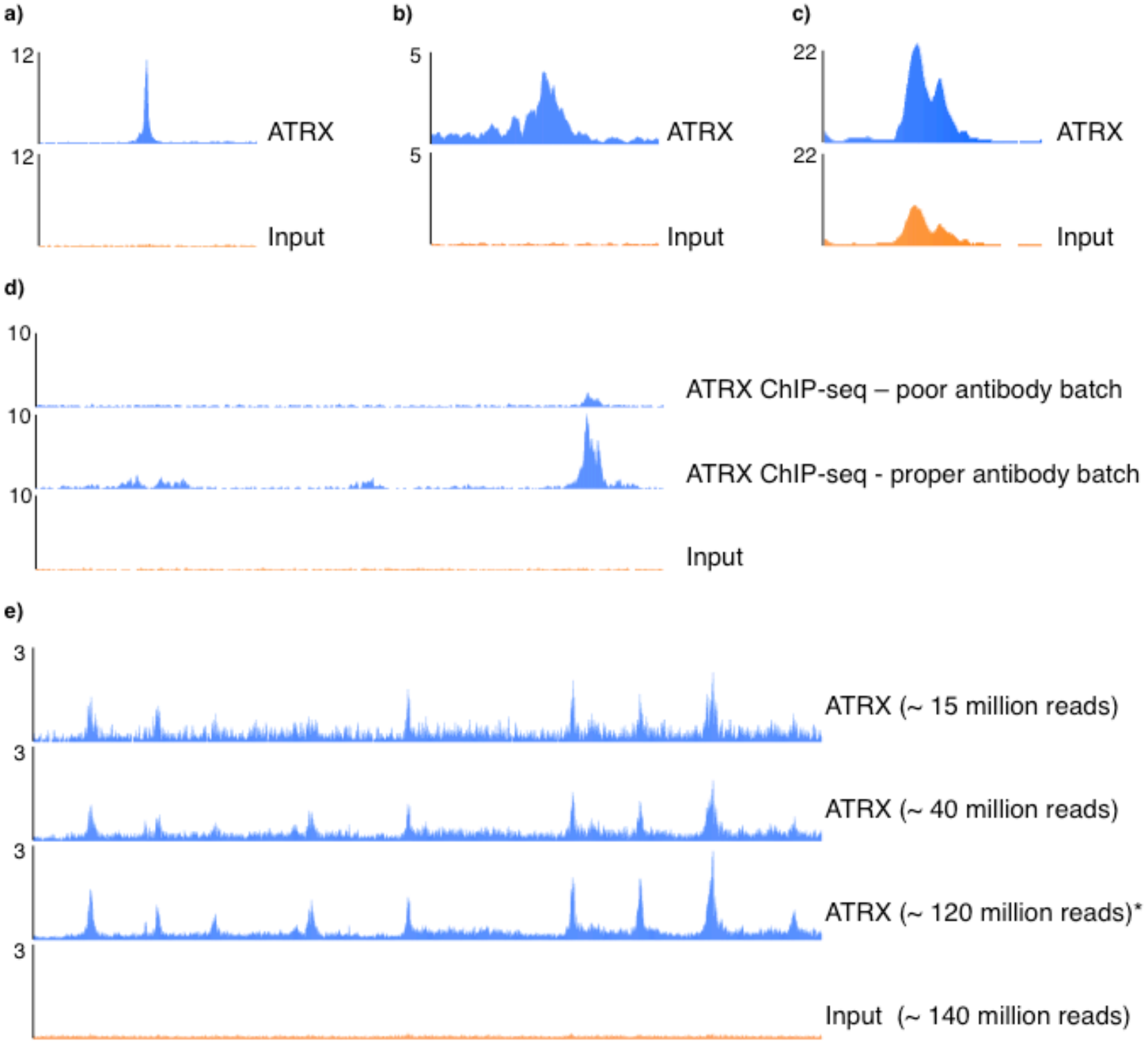
ATRX binding rich region illustrating the importance of the sequencing depth. Deeper sequencing can improve the resolution, providing more information for further analysis. (*) Pool of three independent biological replicates highlighting the reproducibility of the data.


Cell cross-linking
(Section 3.1.1)
*Dynabead preparation
(Section 3.1.2)
Cell sonication
(Section 3.1.3)
Chromatin IP
(Section 3.1.4)
Washing, eluting and reverse cross-linking
(Section 3.1.5)
DNA purification
(Section 3.1.6)
10^8 cells
Two-steps cross-linking (EGS + formaldehyde)
Quench with glycine
Wash
Snap freezing in needed
Cell Lysis
Divide in 3 AFA tubes
Sonication
Combine together
Take input
*Preclearing
*Beads conjugated with ATRX
*Beads conjugated with IgG
Washes and elution
Reverse cross-link
DNA purification
DNA quantification
Sonication validation


1,000
500
400
300
200
100
1
2
3
4


a)
Successful ChIP
0.06
0.04
0.02
0.00
% Input
Positive
Negative
ATRX
IgG
b)
Successful ChIP
0.025
0.020
0.015
0.010
0.005
0.000
% Input
Positive
Negative
ATRX
IgG
c)
Suspect background
0.0020
0.0015
0.0010
0.0005
0.0000
% Input
Positive
Negative
ATRX
IgG
d)
Suspect background
0.008
0.006
0.004
0.002
0.000
% Input
Positive
Negative
ATRX
IgG


Lower
Upper
Sample Intensity [FU]
3000
2500
2000
1500
1000
500
0
25
50
100
200
300
400
500
700
1000
1500
Size
[bp]


a)
12
ATRX
12
Input
b)
5
ATRX
5
Input
c)
22
ATRX
22
Input
d)
10
ATRX ChIP-seq – poor antibody batch
10
ATRX ChIP-seq - proper antibody batch
10
Input
e)
3
ATRX (~ 15 million reads)
3
ATRX (~ 40 million reads)
3
ATRX (~ 120 million reads)*
3
Input (~ 140 million reads)