Expanding the Epi-discovery tool box

Systematic and direct screening for chromatin regulators and chromatin binders using DNA barcoding

Poramba-Liyanage, D.W.

Publication date
2020

Document Version
Other version

License
Other

CHAPTER 2

Epi-ID: systematic and direct screening for chromatin regulators in yeast by Barcode-ChIP-Seq

2

Epi-ID: systematic and direct screening for chromatin regulators in yeast by Barcode-ChIP-Seq

Deepani W Poramba-Liyanage[#], Tessy Korthout[#], Fred van Leeuwen

Division of Gene Regulation, Netherlands Cancer Institute, Amsterdam, The Netherlands.

# these authors contributed equally to this work

Published in Methods in Molecular Biology, 2019

Abstract

The assembly and regulation of chromatin requires coordinated activity of multiple mechanisms. Many factors feed into signaling networks that control the epigenome of a cell. It is this complexity that makes understanding the layers of epigenetic regulation a challenge. Genetic screens have been indispensable for studying chromatin processes. However, they can be laborious and the readout for chromatin changes is often indirect. Epi-ID is a screening strategy in yeast that enables the direct assessment of chromatin status in thousands of gene mutants in parallel. Epi-ID takes advantage of DNA sequences called DNA barcodes that are introduced into a library of yeast knock-out mutants at a common chromosomal location in the genome. Chromatin immunoprecipitation on pools of barcoded mutant strains followed by barcode counting by high throughput sequencing will report on the abundance of the chromatin mark of interest in each mutant strain. Epi-ID is applicable to a wide range of chromatin proteins and modifications that are present and can be immunoprecipitated at or around the barcoded region.

2.1 Introduction

All processes taking place on DNA, whether its transcription, replication or repair, involve chromatin. The assembly of DNA into chromatin by histone proteins can be further adjusted by post-translational modifications of histones and interactions with other proteins that bind to DNA and (un)modified histones. Thus, the composition and structure of chromatin is subject to many modifications and they are under the influence of a range of cellular signals and metabolic activities, leading to dynamic chromatin states that vary across the genome, and between cell stages and cell types. Genetic screens have proven to be indispensable for studying chromatin processes and have led to the discovery of many factors involved in the different layers of epigenetic regulation (Bonini & Berger, 2017; Fuchs & Quasem, 2014). Genetic screens can be laborious, however, and often involve functional readouts or reporter assays that can be influenced by indirect effects. To overcome these issues, we developed Epi-ID, a screening method in yeast in which a chromatin state of a reporter locus is directly assessed in a collection of mutants. Genetic screens in budding yeast have been spearheaded by the availability of systematic deletion collections of the ~5000 non-essential genes and the development of Synthetic Genetic Array (SGA) technologies to perform systematic crosses (Boone, Bussey, & Andrews, 2007; Duina, Miller, & Keeney, 2014; Giaever & Nislow, 2014; Tong & Boone, 2006). In addition, barcode technologies have been developed in which short DNA sequences or DNA barcodes serve as unique and quantifiable identifiers, allowing for parallel analysis of fitness in pools of cells (Giaever et al., 2002; Smith et al., 2009; Yan et al., 2008). In Epi-ID these two concepts are combined in a library in which each clone contains a known gene knock-out and a known pair of unique barcodes.

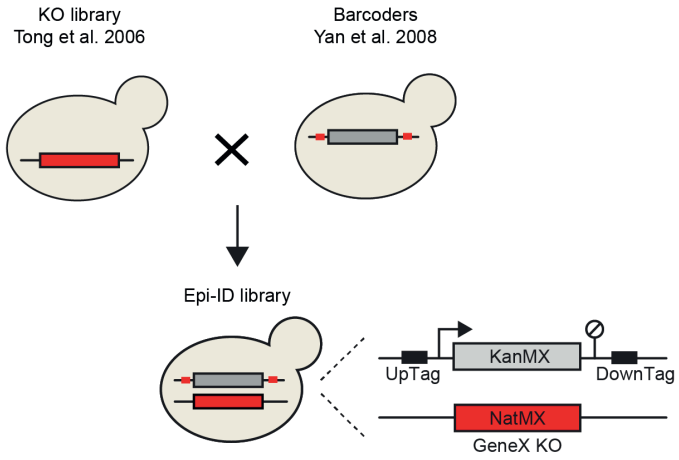

Figure 1

The Epi-ID library was created by crossing a NatMX knock-out library (Tong & Boone, 2006) with a KanMX Barcoder collection (Yan et al., 2008) by using SGA methods. The two barcodes are located in distinct genomic contexts: the UpTag is located in close proximity to the promoter region and the DownTag to the terminator region of the KanMX cassette at the HO locus.

As depicted in Figure 1, a genome-wide Epi-ID library was created (Vlaming et al., 2016) by crossing a knock-out library (Tong & Boone, 2006) with a Barcoder library (~1100 unique pairs of barcodes), which contains in each strain a barcoded KanMX reporter gene containing a G418-resistance marker at a common chromosomal safe-harbor locus, the *HO* gene (Yan et al., 2008). The cross resulted in the generation of approximately 4300 strains divided over 5 plates with unique barcodes (Vlaming et al., 2016). The gene knock-outs contain a nourseothricin (CloNat) resistance marker (NatMX). The KanMX cassette at the HO locus is flanked by two unique DNA barcodes: the UpTag and DownTag. The UpTag is located at the promoter region of the HO/KanMX gene and the DownTag in the HO/KanMX terminator region. Therefore, the two

barcodes are in the same region but in different functional contexts. The Epi-ID library differs from classical yeast knock-out libraries (Giaever et al., 2002) in that the barcodes are placed in a common position in the genome, rather than at the position of the deleted gene. This allows for the assessment of the chromatin feature at a common location in each strain, avoiding possible position effects that are known to influence the chromatin landscape (M. Chen, Licon, Otsuka, Pillus, & Ideker, 2013; X. Chen & Zhang, 2016).

With this Epi-ID library, one can systematically analyze how a knock-out affects the chromatin feature of interest around the DNA barcodes. Briefly, Chromatin Immunoprecipitation (ChIP) is performed on pooled Epi-ID libraries for the chromatin mark of interest, e.g. RNA polymerase or a histone post-translational modification. The DNA barcodes in the input DNA and immunoprecipitated DNA are amplified by a single PCR reaction. During the amplification, the sequence elements necessary for high-throughput sequencing are added and each sample receives a short index sequence to allow multiplexing of samples. Following barcode counting in each sample, barcode enrichment in each ChIP (compared to input) is then used as a measure for the relative abundance of the chromatin mark in each knock-out. For example, a high barcode count of a certain mutant indicates that more of the assessed chromatin feature occurred in that mutant at the barcoded locus (Figure 2).

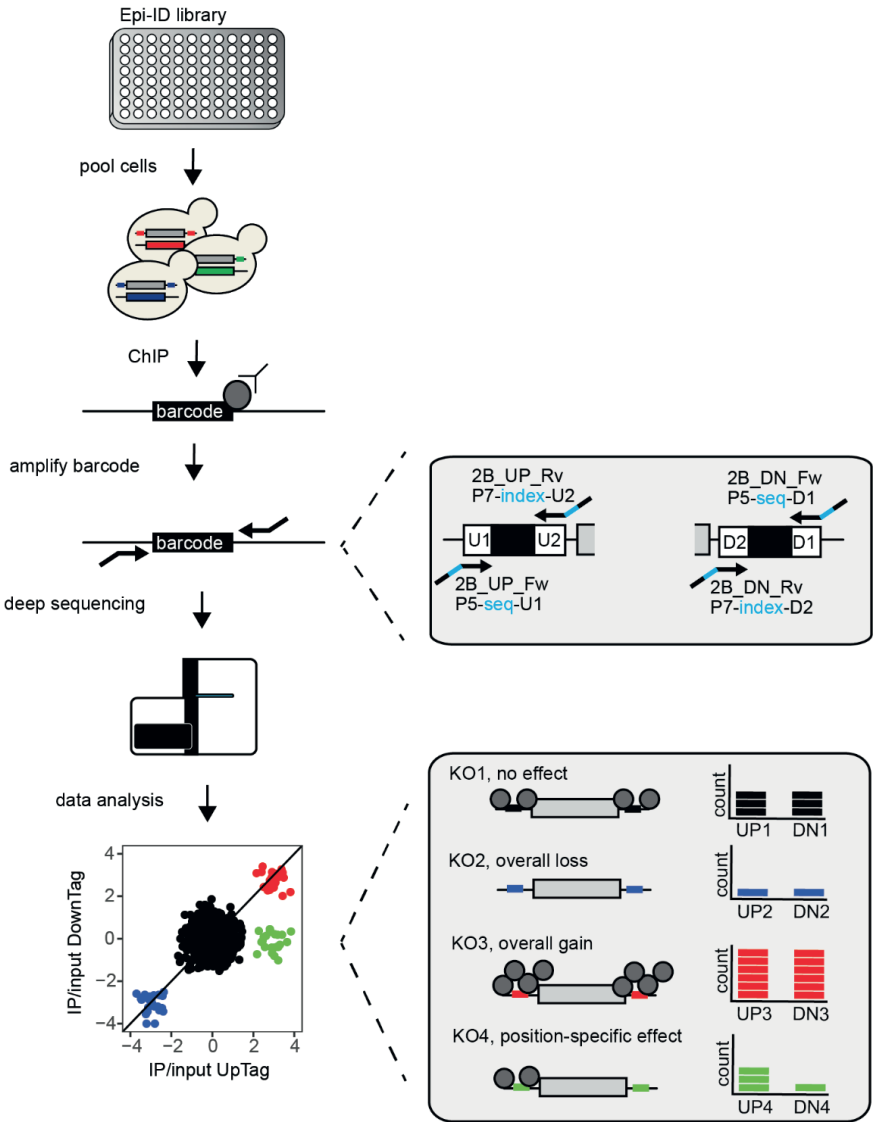

Figure 2

Schematic overview of the Epi-ID screen. The cells from each plate are pooled and a ChIP is performed for the chromatin mark of interest. The UpTag and DownTag are amplified separately with specific primers that also add adaptors for deep sequencing. The reverse primer also

introduces an index to identify each sample and allow for multiplexing. The barcodes are counted by deep sequencing. The majority of the knock-outs do not have an effect on the chromatin mark, so their median-normalized barcode count will be around 0. Increased or decreased counts identify knock-outs with altered chromatin state at the barcode. This can occur at both barcodes or specifically at the UpTag or DownTag.

A key requirement for Epi-ID is that the ChIP against the chromatin feature can successfully immunoprecipitate the barcode DNA. If no antibodies are available for the protein of interest, a gene allele coding for a tagged version of the protein can be introduced in the Epi-ID library using SGA methods. We successfully applied Epi-ID to screen for regulators of H3K79 methylation by Dot1 (Vlaming et al., 2016) and in a small-scale pilot study we combined Epi-ID with the Recombination-Induced Tag Exchange (RITE) assay in a small set of candidate KOs and identified regulators of histone turnover (Verzijlbergen et al., 2011). Finally, the development of advanced genome engineering strategies such as CRISPR-Cas9 will greatly facilitate the development of custom barcoder libraries at chromatin sites other than the HO locus. Such strategies will allow for the determination of the regulatory networks of many other chromatin modifications in the future.

2.2 Materials

Yeast strains and plasmids

The genome-wide Epi-ID library that we recently created consists of five subsets, covering 4921 NatMX gene knock-outs with KanMX barcoders. The strain background is: MAT**a** *can1*Δ*::STE2pr-Sp_his5 lyp1*Δ *his3*Δ *leu2*Δ *ura3*Δ*0 met15*Δ *ho::barcoded-KanMX gene::NatMX*. Each library subset is split into 3*384-well plates (with the exception of the 5th subset which is split into 2*384-well plates). The strains are stored and available upon request as 384-well glycerol freezer stock plates (Vlaming et al., 2016). The plates can be handled efficiently by robotics systems such as a RoToR (Singer Instruments). Alternatively, the library subsets can be stored and maintained as liquid pools, which may be preferred when robotics is a limitation in carrying out a screen.

Yeast media and drugs

1. G418 (Geneticin): Dissolve in water at 250 mg/ml, filter-sterilize and store in aliquots at 4°C.
2. CloNat (Nourseothricin): Dissolve in water at 100 mg/ml, filter-sterilize, and store in aliquots at 4°C.
3. YEPD plates: Dissolve 10 g yeast extract, 20 g bacto peptone, 20 g bacto agar in 860 ml water in a 1 l flask. After autoclaving, add 100 ml of 20% glucose solution, 40 ml 1 mg/ml uracil, 10 mg/ml Tryptophan. Mix thoroughly, cool to approximately 65°C, and pour plates.
4. YEPD + G418 + CloNat plates: Cool YEPD medium to approximately 65°C, add 0.8 ml of G418 (final concentration 250 mg/ml), and 1 ml of CloNat (final concentration 100 mg/L) stock solutions, mix thoroughly, and pour plates.
5. YEPD medium: Dissolve 10 g yeast extract, 20 g bacto peptone in a 1 l flask. After autoclaving, add 100 ml of a 20% glucose solution and mix thoroughly.

Yeast plates and accessories

1. Plus Plates for RoToR (PLU-001, Singer Instruments)
2. Pins for RoToR (Singer instruments)
3. Aluminum sealing tape (Nunc, cat. no. 276014) is used for resealing the 384-well plates that contain the frozen stocks of the Epi-ID library.

Buffers

1. TBS: 20 mM Tris-HCL pH 7.9, 150 mM NaCl
2. PBS: 140 mM NaCl, 2.5 mM KCl, 8.1 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.5
3. BSA: 100 mg/ml in H_2O
4. PBSB: PBS + 5 mg/ml BSA
5. Fix Solution: 50 mM Hepes-KOH, pH 7.5, 100 mM NaCl 5 M, 1 mM EDTA, 11% Formaldehyde
6. Breaking Buffer: 100 mM Tris pH 7.9, 20% glycerol, Protease Inhibitors Cocktail tablets
7. FA buffer: 50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA 1% TritonX-100, 0.1% Na-deoxycholate, H_2O
8. FA-HS: 50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA 1% TritonX-100, 0.1% Na-deoxycholate, H_2O
9. FA+0.8% SDS: similar to FA, for 50 ml, replace 2 ml H_2O with 2 ml 20% SDS
10. FA+0.8% SDS: add 8 ml FA+0.8% SDS to 40 ml FA
11. RIPA buffer: 10 mM Tris pH 8, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate 1mM EDTA, H_2O
12. Elution buffer: 50 mM Tris pH 8, 10 mM EDTA, 1% SDS
13. TE: 10 mM Tris pH 8,1 mM EDTA

Equipment

1. RoToR HDA from Singer Instruments (Watchet, UK)
2. Sorvall fixed angle centrifuge
3. Screw-cap tube for bead beater
4. Bead beater (BioSpec)
5. Silica/zirconia beads (0.5 mm)
6. Bioruptor Pico (Diagenode)
7. 15 ml Bioruptor Tubes and sonication beads Pico (Diagenode)
8. Protein A/G Dynabeads (Thermo Fisher Scientific)
9. Gel imaging system
10. Microcentrifuge.
11. Medium bench swing-out centrifuge.

2.3 Methods

Cell culture and preparation of cell pellets

The number of cells needed to obtain enough barcode copies for preparing a sequencing library depends on the IP efficiency of the antibody. Therefore, the IP efficiency needs to be determined at both the UpTag and DownTag for each new antibody. This can be calculated as percent input by ChIP-qPCR with primers listed in Table 1.

1. Calculate the number of cells needed (see Note 1). Here we assume a culture of 450 ml.
2. Thaw the 384-well library plates and pin directly to 1536 (3*384) YEPD agar plates with the antibiotics (see Note 2).
3. Keep at 30 °C overnight and pin to 1536 plates (YEPD+antibiotics) the following day. Repeat 2 times and prepare for the harvesting of cells on day 3.

4. On day 3, pipette 10 ml YEPD onto the pate and scrape off the cells with a cell scraper. Transfer the medium with cells into a 50 ml Falcon tube. Repeat 3 times to collect the cells in a total of 30 ml YEPD.
5. Dilute the cell suspension 10 times to measure the OD660. Make 450 ml YEPD culture with OD ± 0.150. Avoid cross contamination between the 5 library subsets.
6. Grow for approximately 5 h until an OD660 between 0.500 and 0.700 is reached (an OD660 of 0.700 should yield ~$1x10^7$ cells).
7. Crosslink the cells by adding 1/10th of the volume fresh Fix solution (45 ml Fix solution for 450 ml culture volume) and slow shaking for 20 min at room temperature. Work in a fume hood, keep the lids closed and use fresh formaldehyde (see Note 3).
8. Add 1/20 volume 2.5 M glycine to get a final concentration of 0.125 M (22.5 ml for a 450 ml culture volume) and shake 5 more minutes at room temperature (see Note 4). Place cells at 4 °C in the subsequent steps.
9. Harvest the cells using GSA buckets in a Sorvall fixed angle centrifuge. Spin at 4000 rpm (2500 x g) for 10 min at 4 °C. If more than one bucket is required for spinning the volume of cells, spin down in multiple steps, each time pouring off supernatant and adding more culture.
10. Resuspend each pellet in 30 ml cold TBS and transfer to 50 ml Falcon tubes. Spin down at 4000 rpm (3300 x g) for 5 min at 4°C.
11. Resuspend in 2 ml TBS + 1:500 PMSF and divide over 4* 2 ml screw-cap tubes for bead beating (see Note 5). Spin down 15 seconds at max speed and remove supernatant.
12. Store cell pellets at -80 °C (at least overnight).

Coupling beads and antibody

Select protein A or G Dynabeads according to the antibody used (see Note 6). The optimal ratio of beads and antibody varies per antibody. If a new antibody is used, it is recommended to titrate the antibody for optimal beads:antibody ratio. Here, we assume a beads:antibody ratio of 20:5.

1. 20 µl of Dynabeads per IP in a 1.5 ml tube (use wide tip or cut the tip of a p200 off at ~2 mm from the tip!).
2. Concentrate the beads on a magnetic particle concentrator (MPC), remove supernatant. Do not leave beads on MPC for more than 2 min.
3. Wash the beads 3 times with 1 ml cold PBS+5 mg/ml BSA (PBSB, should be freshly prepared).
4. Add 100 µl per IP of PBSB, resuspend beads well and add antibody.
5. Rotate 2 hours to overnight at 4°C to bind (see Note 6).
6. Wash the beads 2 times with PBSB, then 1 time with cold FA.
7. Resuspend beads in 100 µl FA per IP.
8. Keep on ice until used.

Chromatin preparation

Chromatin preparation is done in bead beater tubes and 15 ml Bioruptor tubes to prepare for subsequent sonication using the Bioruptor Pico (see Note 7).

<u>Preparing 15 ml Bioruptor tubes for sonication</u>

1. Add 800 mg beads to the tube (this is equivalent to the 3rd line of the 15 ml Bioruptor tube graduation scale).
2. Wash the beads 3 times with PBS (add PBS, quick vortex, 5 min spin at 2000 rpm (835 x g)).
3. Remove all buffer by decanting.

4. Store tubes on ice until used.

Preparing chromatin

All reagents and tubes should be kept at 4°C (on ice). Pre-cool the buffers.

1. Resuspend the pellet in bead beater tube in 200 µl cold breaking buffer (add fresh protease inhibitors to the buffer).
2. Add ~500 µl silica/zirconia beads (0.5 mm).
3. Bead beat in bead beater in a cold block for 4 min. Change the cold block after 2 min (see Note 8).
4. Combine the samples from two bead-beater tubes into one Bioruptor tube already prepared with beads. Add 500 µl of cold FA buffer (+ PIC) to bead beater tubes, invert a couple of times and let beads sink to bottom of the tube and transfer the samples with broken cells to the prepared 15 mL Pico tube with sonication beads (see Subheading 3.3.1). Repeat 2 times with another 500 µl of FA buffer. For smaller culture volumes the reagents can be scaled down (see note 9).
5. Spin 5 min at 4000rpm (3300 x g) at 4°C to pellet chromatin and debris. Discard supernatant.
6. Wash pellet again with 1 ml of cold FA (+PIC), spin 5 min at 4000 rpm (3300 x g) at 4 °C and discard supernatant. The performance of some antibodies can be improved by adding SDS (see Note 10).
7. Add 1 ml of cold FA (+PIC) to the pellet and sonicate using the “Bioruptor Pico” for 10 min at 30 second intervals at high power. Depending on cell counts the DNA should be ~200-2000 bp, but mostly ~600 bp.
8. Add 1 ml cold FA (+PIC) to the suspension to dilute chromatin.
9. Spin 5 min at 4000 rpm (3300 x g) at 4°C to pellet debris and transfer soluble chromatin to a new 2 ml tube.

10. Take an input sample for the chromatin prep (10% of IP volume) and keep at 4 °C.
11. Snap freeze chromatin in liquid nitrogen and store at -80 °C.

Check the chromatin size

The size of the chromatin can be checked by running the purified DNA of the input samples on a gel. It is recommended to confirm that the chromatin was sheared to fragments of ~600 bp before proceeding with the ChIP. Alternatively, the ChIP can be performed immediately and the input sample de-crosslinked simultaneously with the IP samples and analyzed afterwards.

1. Prepare the input sample of the sonicated chromatin. To 20 µl chromatin, add 80 µl elution buffer and 70 µl TE (make volume equal to IP sample). Then add 0.5 µl RNase A (10 mg/ml) and 10 µl ProtK (10 mg/ml).
2. Incubate the samples for 1 h at 50 °C and then overnight (or at least 6 hours) at 65 °C to reverse crosslinks.
3. Clean up DNA using QIAquick PCR Purification Kit (Qiagen). Briefly, to each eluted ChIP sample, add 5 times volume binding buffer and apply ~650 µl to a QIAquick spin column. Spin at 13000 rpm (16,000 x g) for 1 min. Discard the flow through, add the rest of the sample and spin again. Wash with 750 µl wash buffer and spin again to dry. Elute in 30 µl elution buffer at 50 °C.
4. To check input DNA on gel, treat the column purified DNA with RNase A before loading on gel. Add 1 µl of a 1:10 dilution of RNase (10 mg/ml) to 15 µl DNA. Incubate 20 min at room temperature and load fragments on a 1% TAE gel.
5. Depending on the cell counts the DNA should be between 200 and 2000 bp, but mostly ~600 bp (Figure 3).

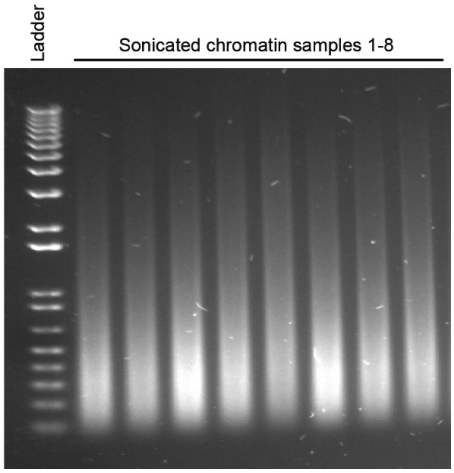

Figure 3

Sonicated chromatin from 1x10^9 cells in 1 ml of cold FA sonicated using the "Bioruptor Pico" (ten minutes with cycles of 30 s intervals on and 30 s off at high power), run on a 1% agarose gel. Depending on cell counts the DNA should be ~200-2000 bp, but mostly ~600 bp.

Chromatin immunoprecipitation

All reagents and tubes should be at 4°C (on ice). Pre-cool the buffers.

1. Use the amount of chromatin calculated previously (see Note 1).
2. Add chromatin to Dynabeads coupled to antibody.
3. Rotate at 4 °C overnight.
4. Concentrate on magnetic particle concentrator (MPC). After 30 seconds invert twice and wait 30 seconds and aspirate off the buffer (e.g. use vacuum pump with clean yellow tip).
5. Wash with 1 ml FA buffer and rotate for 5 min at room temperature (or 4 °C).
6. Repeat wash steps for total of:
    a. 2x with FA
    b. 2x with FA-HS

c. 2x with RIPA
d. 1x with TE (when using TE, the beads do not stick well to the wall of the tube on the MPC. Do not aspirate with vacuum pump, but instead use a pipette.)

7. Add 100 µl elution buffer, incubate 10 min at 65°C in a heating block and shake at 650 rpm to prevent the beads from precipitating.
8. Concentrate on MPC. Take the supernatant (~100 µl), add 70 µl TE, 0.5 µl RNase A (10 mg/ml) and 10 µl ProtK (10 mg/ml).
9. Incubate the eluted samples for 1 h at 50 °C and then overnight (or at least 6 hours) at 65 °C to reverse crosslinks. Clean up DNA using QIAquick PCR Purification Kit (Qiagen). Briefly, to each eluted ChIP sample, add 5 times volume binding buffer and apply ~650 µl to a QIAquick spin column. Spin at 13000 rpm (16,000 x g) for 1 min. Discard the flow through, add the rest of the sample and spin again. Wash with 750 µl wash buffer and spin again to dry. Elute in 30 µl elution buffer at 50 °C.

Library preparation

The UpTag and DownTag are amplified separately with primers that anneal to a common sequence immediately flanking the barcode (Figure 2). The forward primer introduces the Illumina P5 sequence and extra nucleotides for annealing of the 5' end of custom sequencing primers. The reverse primer introduces the Illumina P7 sequence and a 6-base-pair unique index used as a molecular identifier for each sample to allow for multiplexing. The introduction of this unique index allows for identification of each sample+plate combination. Therefore, the ChIPs and inputs from the 5 subset Epi-ID libraries can pooled after this point. The unique index allows for extensive pooling of up to at least 150 samples in one Illumina HiSeq lane. As a quality control for the ChIP and the primers, a test PCR is advised before using all IP material for the library preparation.

Test PCR analysis for ChIP efficiency

1. Dilute the ChIP samples for a test PCR to check the approximate efficiency for the ChIP and specificity. Use 2 µl of the ChIP sample and make a 1:10 dilution.
2. Use 10 µl of this in a 50 µl PCR using the 2B_UP_Fw/2B_DN_Fw and one of the unique reverse primers (see Table 1).
3. Use the following cycle program:
    a. 98°C for 30 sec;
    b. 10 cycles of 98°C for 15 sec, 55°C for 20 sec and 72°C for 15 sec
    c. 13 cycles of 98°C for 15 sec, 70°C for 20 sec and 72°C for 20 sec
    d. final extension of 72°C for 10 min
4. Run 15 µl on a 1.5% agarose gel to check if the PCR worked. A faint band is a good indication that the ChIP efficiency was sufficient.

Library preparation

1. Repeat the PCR described in the test, but now use as much as possible of the ChIP material (13 µl of ChIP for UpTag and 13 µl DownTag). For input, 3 µl from the undiluted sample is more than sufficient.
2. Run 10 µl of each reaction on 1.5% agarose gel and use the Gel Doc to generate an image for quantification using Image Lab software. The amplicons from each reaction (UpTag input and UpTag ChIP samples) should be run on one gel for quantification.
3. For quantifying the signal intensity of DNA bands, use the scanner-generated image and analyze using Image Lab software using volume tools. Make sure that the image is not over exposed. Any software that allows for comparing the band intensities can be used for this (see Note 11).
4. Open the image in Image Lab, select volume tools and define the area of interest using a rectangle. Then encircle the rest of the bands using the same rectangle (Figure 4). Click on the local background subtraction method, units=none and regression method=linear.

Click on analysis table to get a list of background adjusted intensity values (Adj.Vol.(int). This table can be exported into excel and the Adj.Vol.(int) value can then be used to compare the intensities of bands across the gel.

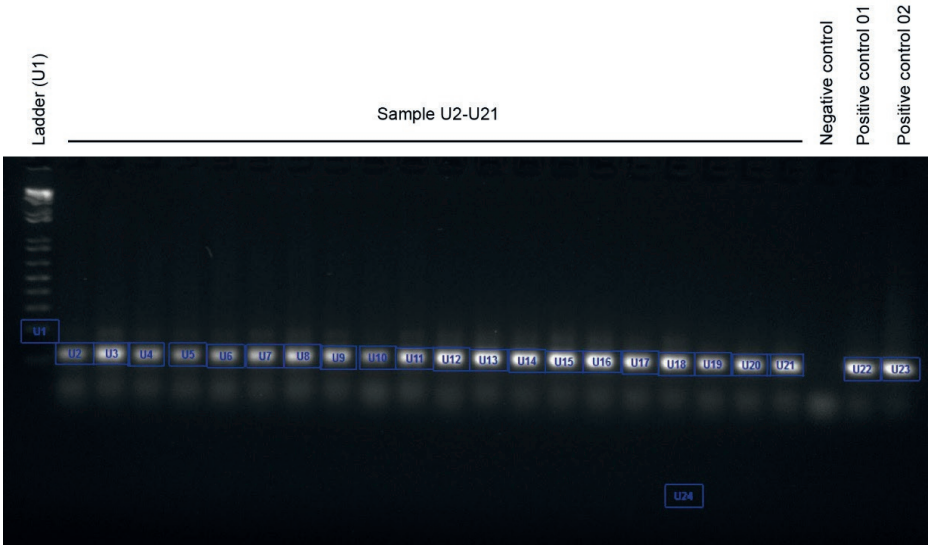

Figure 4

Image Lab quantification of amplified barcodes. 10 µl of barcode PCR reaction is run on a 1.5% agarose gel. Gel doc is used to generate an image and quantification is done with the Image Lab software (volume tools). U2-U21 are UpTag amplified PCR reactions. U1 represents the 200 bp marker from 10 µl of DNA ladder, which can be used to compare the intensity of bands between gels. U22 and U23 are bands from positive control reactions. U24 gives an idea of background intensity.

5. Once the quantification is done, mix the UpTags in equimolar amounts and repeat the steps for the DownTags.
6. Next, run the pooled UpTags and DownTags on one agarose gel. Quantify and mix the UpTags and DownTags in equimolar amounts.

7. Once mixed, run the final pool on a 1.3% agarose gel with large wells. Excise the band (100-150 bp size selection) from the gel and isolate DNA with a QIAquick gel extraction kit (Qiagen): Excise the DNA fragment with a clean scalpel and weigh the gel slice in the tube. Add 3 volumes of binding buffer and incubate in a shaking heat block at 50 °C until the gel has dissolved. Add 1 gel volume of isopropanol to the sample and mix. Place in a QIAquick spin column and centrifuge at 13000 rpm (16,000 x g) to bind DNA. Following the washing steps, elute the DNA with elution buffer to end up with approximately 40 µl purified UpTag and DownTag mix.
8. Run 3 µl on gel to check if there is enough DNA. Even a faint band is sufficient for a good sequencing run.
9. The library is now ready for sequencing with custom UpTag and DownTag sequencing primers (Table 1, 50 µM each in TE).

Sequencing and data analysis

The amplified and pooled barcodes are sequenced using a mix of custom sequencing primers for the UpTag and DownTag. These primers consist of part of the standard Illumina sequencing primers, adjusted to optimize the melting temperature and to include the U1 or D1 sequence present in the UpTag and DownTag amplicon, respectively (Table 1).

Sequencing is done on a HiSeq2500 platform (Illumina, San Diego, CA), single read, >50 bp) with High Output Run Mode. Smaller scale Epi-ID experiments can also be run on a MiSeq (Illumina). The primers described here are compatible with either platform.

To extract the barcodes and indices from each read, we developed the Perl script eXtracting Counting And LInking to Barcode References (xcalibr). This script locates the constant region in each read (U2 or D2, Figure 2) and assigns the flanking sequences to the barcode and index,

respectively. The output of this script is a counts table for each index-barcode combination. The xcalibr source code is available at https://github.com/NKI-GCF/xcalibr.

The counts table is further processed by removing counts below ten and dividing by the median per index to normalize for differences in library size of each pooled sample. At this point, several quality control steps could be performed (see Note 12). Finally, barcode-index combinations are converted to ORF names. Further analysis can involve dividing IP values over each other or normalizing for input of reference IPs.

2.4 Notes

1. Pull-down efficiency and calculation of culture quantities

We determined that at least 250 copies per barcode (500 copies for the UpTag and DownTag together) are required to be present in a PCR reaction to minimize jackpot effects. The following formula can be used to calculate the culture quantities required for the respective antibody.

- Required number of copies per barcodes for successful Epi-ID = 250
- Number of barcodes in pool = 1100 (barcode complexity)
- Total number of barcodes in PCR = 250*1100 (based on the two values above)
- ChIP efficiency at UpTag of your antibody = x (depends on the volume of chromatin, antibody and beads used in the ChIP)
- ChIP efficiency at DownTag of your antibody = y (depends on the volume of chromatin, antibody and beads used in the ChIP)
- Input yield (amount of ChIPable DNA that can be extracted from cells) is estimated to be 50%
- Number of cells required for UpTag = Total number of barcode loci in PCR/ (ChIP efficiency UpTag * Input yield)

- Number of cells required for DownTag = Total number of barcode loci in PCR/(ChIP efficiency DownTag * Input yield)
- Number of cells in total ChIP = Number of cells required for UpTag + Number of cells required for DownTag

  At OD660=0.5, the number of cells/ml of culture = ~ $7x10^6$
- Culture volume required = Number of cells needed in total ChIP/ $7x10^6$

2. Cell culture

The Epi-ID library contains a NatMX deletion cassette and the HO locus is replaced by a barcoded KanMX cassette. Thus, the library plates are grown on YEPD plates with CloNat (10 mg/ml) and G418 (250 mg/ml). Once the 384 glycerol plates are thawed, they are immediately arrayed in 1536 format (where 3*384 glycerol plates make one 1536-format plate). When using a library in pelleted cell format, thaw cells in 30 ml rich media and move directly to step 4 in section 3.3.2 Preparing chromatin.

3. Fresh formaldehyde in Fix solution

Fix solution should be prepared freshly with formaldehyde before use as formaldehyde is unstable and will deteriorate with time.

4. 2.5M Glycine

Glycine is added to quench the formaldehyde and is important to prevent the antibody from being crosslinked in the following steps.

5. Cells per bead beater tube

Place a maximum of $1.5x10^7$ cells in one bead beater tube and divide pellet equally into an even number of tubes to make it easier for subsequent steps. Adding an excess of cells to a tube will hinder the bead beating process and will result in lower extraction of cell material.

6. Preparing beads and antibody

Coupling of antibody to Dynabeads is done for approximately 2 hours to overnight and should be performed immediately before the IP. For best results, couple beads for ~6 hours and use for IP on the same day. Select protein A or G beads according to the antibody used. For rabbit antibodies use Protein A or Protein G and for mouse and goat antibodies use Protein G beads. Protein G binds better at low pH. Make sure the pH is not higher than pH8 in subsequent steps.

7. Sonication using Bioruptor Pico

Sonication of larger volumes is more efficient when using the 15 ml sonication tubes and beads with the sonicator Pico. Optionally, a Diagenode Bioruptor sonicator can be used, but sonicating time must be adjusted accordingly to ensure that fragments are approximately ~600 bp.

8. Bead beating large cell volumes

Bead beating is done in bead beater tubes using a bead beater. If more cells/tube are used, check under microscope to verify that the cells have lysed. Cells should be almost 100% broken. If not, bead beat a bit longer.

9. Preparing chromatin for smaller culture volumes

For cultures of 150 ml with OD660 of 0.5-0.7, collect pellets into one bead beater tube. Following lysis, continue all steps (step 5 to step 9) with half the volume of reagents.

10. Using SDS in buffers

Some antibodies work best on proteins that are denatured by SDS. For such antibodies, steps 6 to 8 of section 3.3.2 Preparing chromatin should be modified as described below.

6. Wash pellet again with 1 ml of cold FA (+PIC+0.13%SDS), spin 5' 4K at 4°C discard supernatant.

7. Add 1 ml of cold FA (+PIC+0.8%SDS) to pellet and sonicate using the “Bioruptor Pico” 10' at 30 second intervals at high power. Following sonication add 1 ml cold FA (+PIC) to the suspension.
8. Add 1 ml cold FA (+PIC) to the suspension to dilute chromatin. At this point, the SDS concentration is ~0.4%. For ChIP, the SDS concentration needs to be lowered to ~0.12% by adding FA buffer to prevent antibody denaturation.

11. Quantifying DNA with a Bioanalyzer

The protocol described here uses the Gel Doc imaging system to quantify the relative amount of DNA in the ChIP and input library samples. We note here that it is also possible to quantify the DNA samples by using a Bioanalyzer.

12. Quality control by using plate signatures

The five different plates of the genome-wide Epi-ID library use the same barcoder set but each has a specific signature of empty spots corresponding to missing barcode pairs. Visualization of this signature can be done by plotting the values of each column (representing the counts of one indexed sample) of the counts table. This plot allows for an extra check to verify that the correct plate number was used and reveals potential mix-ups. In addition, the barcodes of the empty spots should be completely absent. A low barcode count of empty spots can be an indication of cross-contamination between libraries or samples.

2.5 Acknowledgements

The authors thank the RHPC facility of the Netherlands Cancer Institute for providing computational resources and Hanneke Vlaming and Kitty Verzijlbergen for developing the Epi-ID technology. This work was supported by the Dutch Cancer Society (KWF2009-4511 and NKI2014-7232) and the Netherlands Organisation for Scientific Research (NWO-VICI-016.130.627). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Table 1: List of the primers used to perform an Epi-ID screen.

Name	Sequence
HOpromQfw	GAAGCTTGTTGAAGCATGATGAA
HOpromQrev	TTGCTGCTTATGAGGATATGGATTT
HOtermQfw	GAGTAGAAATACGCCATCTCAAGATACA
HOtermQrev	GGAAAGTTGATCAAGACCCAATAATAA
2B_UP_Fw	AATGATACGGCGACCACCGAGATCTACACCTTCCGATCTAGATGTCCACGAGGTCTCT
2B_DN_Fw	AATGATACGGCGACCACCGAGATCTACACTCTTCCGATCTACGGTGTCGGTCTCGTAG
2B_UP_Rv_all	CAAGCAGAAGACGGCATACGAGATNNNNNNGTCGACCTGCAGCGTACG
2B_DN_Rv_all	CAAGCAGAAGACGGCATACGAGATNNNNNNAACGAGCTCGAATTCATCGA
2B_UP_Seq	ACACCTTCCGATCTAGATGTCCACGAGGTCTCT
2B_DN_Seq	ACACTCTTCCGATCTACGGTGTCGGTCTCGTAG

The qPCR primers amplify the HO promoter and terminator. These can be used on the ChIP material to determine the IP efficiency of each antibody. 2B_UP_Fw and 2B_UP_Rv_all amplify the UpTag, and 2B_DN_Fw and 2B_DN_Rv_all the DownTag. 2B_UP_Fw and 2B_DN_Fw introduce the Illumina P5 sequence, a custom sequencing region to anneal sequencing primers (underlined region) and the U1/D1 common region next to the barcodes. 2B_UP_Rv_all and 2B_DN_Rv_all introduce the Illumina P7 sequence, a unique 6-bp index to identify each sample (indicated by N's) and the U2/D2 common region next to the barcodes. A list of the 6-bp indices that we used is available upon request. 2B_UP_Seq and 2B_DN_Seq are custom sequencing primers that consist of the custom sequencing region (underlined region) and the U1/D1 region.