Phosphorylation of histone H3.3 at serine 31 promotes p300 activity and enhancer acetylation

Sara Martire, Aishwarya A. Gogate, Amanda Whitmill, Amanuel Tafessu, Jennifer Nguyen, Yu-Ching Teng, Melodi Tastemel and Laura A. Banaszynski*

The histone variant H3.3 is enriched at enhancers and active genes, as well as repeat regions such as telomeres and retroelements, in mouse embryonic stem cells (mESCs)\textsuperscript{9–16}. Although recent studies demonstrate a role for H3.3 and its chaperones in establishing heterochromatin at repeat regions\textsuperscript{4–8}, the function of H3.3 in transcription regulation has been less clear\textsuperscript{9–16}. Here, we find that H3.3-specific phosphorylation\textsuperscript{17–19} stimulates activity of the acetyltransferase p300 in trans, suggesting that H3.3 acts as a nucleosomal cofactor for p300. Depletion of H3.3 from mESCs reduces acetylation on histone H3 at lysine 27 (H3K27ac) at enhancers. Compared with wild-type cells, those lacking H3.3 demonstrate reduced capacity to acetylate enhancers that are activated upon differentiation, along with reduced ability to reprogram cell fate. Our study demonstrates that a single amino acid in a histone variant can integrate signaling information and impact genome regulation globally, which may help to better understand how mutations in these proteins contribute to human cancers\textsuperscript{20–22}.

The histone variant H3.3 was first identified as a component of active chromatin, and genome-wide studies suggest that at least 75% of H3.3 deposition occurs at regions that are associated with active genes in mESCs\textsuperscript{9}. Although it is known that H3.3 is enriched at active enhancers and genes, it is debatable whether its deposition at these regions is functionally required or simply a byproduct of the high levels of nucleosome turnover that these regions experience\textsuperscript{9–16}. As H3.3 is an integral nucleosomal component of enhancers (Fig. 1a and Supplementary Fig. 1a), we wanted to understand how its loss might affect the chromatin signature at these regions. Intriguingly, we found that mESCs that lack H3.3 (H3.3 KO)\textsuperscript{1} show a global reduction of H3K27ac without any appreciable reduction in H3K4me1 (Fig. 1b and Supplementary Fig. 1b–d). Mass spectrometry confirmed that the overall level of H3.1K27ac in H3.3 KO mESCs was approximately 60% of that in wild-type mESCs (Fig. 1c). Reduced acetylation was accompanied by an increase in global H3.1K27me2 and H3.1K27me3 (Supplementary Fig. 1e)\textsuperscript{22}.

To explore the relationship between H3.3 and chromatin signatures at regulatory elements genome-wide, we performed chromatin immunoprecipitation followed by sequencing (ChIP–seq) of several histone modifications in wild-type and H3.3 KO mESCs. Inspection of genes encoding three pluripotency transcription factors, Sox2, Pdcd5fi, and Nanog, showed that the regulatory elements controlling these genes (evidenced by binding of Oct4 (ref. 23), Nanog\textsuperscript{21}, and p300) were marked by reduced levels of H3K27ac in the absence of H3.3 (Fig. 1d). Genome-wide, we found that the majority (13,135 out of 17,589, ~75%) of promoter-distal enhancers (greater than ±3 kb from the transcription start site) marked by both H3K27ac and p300 showed reduced H3K27ac enrichment in the absence of H3.3 (Fig. 1e and Supplementary Fig. 1f–h). We confirmed these results in an independent ChIP coupled with quantitative PCR (ChIP–qPCR) experiment at 16 individual enhancers in two H3.3 KO clonal mESCs (Supplementary Fig. 1i,j). Reduced acetylation was more pronounced at promoter-distal enhancers than at promoters themselves (Supplementary Fig. 1k). Reduced acetylation was not limited to H3K27ac, as H3K18ac, H3K64ac, and H3K122ac\textsuperscript{24} also showed reduced levels at enhancers in H3.3 KO mESCs (Supplementary Fig. 1l–n). We did not observe an appreciable reduction in H3K4me1 at enhancers (Supplementary Fig. 1o,p).

In line with the model of enhancer priming\textsuperscript{25–28}, these results suggest that H3K4me1 is upstream of enhancer acetylation.

We next considered the two independent chaperone systems that deposit H3.3 into distinct regions of chromatin. The HIRA complex is responsible for the majority of genic deposition (that is, promoters, gene bodies, and enhancers), whereas the ATRX–DAXX complex deposits H3.3 at repeat regions such as telomeres and interstitial heterochromatin (Supplementary Fig. 2a,b)\textsuperscript{2}. We therefore predicted that loss of HIRA, but not ATRX or DAXX, would result in similar loss of enhancer acetylation as observed upon loss of H3.3. In agreement with this, we observed both a global reduction and enhancer-specific loss of H3K27ac in HIRA KO mESCs (Supplementary Fig. 2c,d). The vast majority of these regions (7,950 out of 8,476, 94%) also showed reduced H3K27ac in H3.3 KO mESCs (Supplementary Fig. 2e). By contrast, ATRX KO did not alter H3K27ac enrichment and DAXX KO resulted in increased H3K27ac at enhancers in mESCs (Supplementary Fig. 2f). We then generated an independent HIRA KO mESC line that was isogenic with the ATRX and DAXX KO mESCs used in our study (Supplementary Fig. 2g). The reduction of H3K27ac at enhancers observed for these two independent HIRA KO mESCs was strikingly similar (Supplementary Fig. 2h,i) and consistent across independent experiments (Supplementary Fig. 2j). Overall, these data suggest that HIRA-dependent deposition of H3.3 at enhancers has a role in maintaining histone acetylation of these regions.

We then wanted to understand how loss of H3.3 leads to changes in H3K27ac at regulatory elements. We did not detect changes in acetyl-CoA levels or increased histone deacetylase (HDAC) activity in H3.3 KO mESCs (Supplementary Fig. 3a,b). Furthermore, we did not observe an appreciable difference in expression of the p300 or CBP histone acetyltransferases (Supplementary Fig. 3c). As our data and that of others suggest that p300 is responsible for the majority of H3K27ac in mESCs (Supplementary Fig. 3d)\textsuperscript{22}, we focused
Fig. 1 | Histone H3.3 facilitates histone acetylation at enhancers. a, Heatmap showing enrichment of H3 and H3.3 at promoter-distal enhancers (±3 kb from TSS) in mESCs. Each row represents a single enhancer (n = 17,589). b, Immunoblot with the indicated antibodies from wild-type (WT) and H3.3 KO mESC whole-cell lysates loaded with increased protein concentration (1x, 2x). Blot is representative of more than five independent experiments. See also Supplementary Fig. 9a. mAb, monoclonal antibody; pAb, polyclonal antibody. c, Quantification of H3K27ac levels in WT and H3.3 KO mESCs by tandem mass spectrometry (MS/MS, Mod Spec, Active Motif). Data are represented as percent of the total H3.1(27–41) peptide observed, ±3 ± s.d., two-tailed t test. ** P = 0.0071. d, Genome browser representations of Oct4, Nanog, p300, H3.3, and H3 ChIP–seq in WT mESCs and of H3K27ac ChIP–seq in WT and H3.3 KO mESCs. The y axis represents read density in reads per kilobase per million mapped reads (RPKM) normalized to an external standard. Gray boxes indicate regulatory elements. Data are representative of three independent experiments using two different antibodies. e, ChIP–seq average profiles in WT and H3.3 KO mESCs of H3K27ac. Black line, WT; red line, H3.3 KO. c, Heatmap showing acetylation enrichment of H3 and H3.3. Enhancers region ±3 kb. e, Heatmap showing enrichment of H3 and H3.3 at promoter-distal enhancers (±3 kb from TSS). Enhancers region ±3 kb. Enhancer boxes indicate regulatory elements. Data are representative of three independent experiments using two different antibodies. e, ChIP–seq average profiles in WT and H3.3 KO mESCs of H3K27ac. Black line, WT; red line, H3.3 KO.
nucleosomes resulted in an increase in p300 activity, as compared to the unphosphorylated substrate (Supplementary Fig. 5a,b). As our cellular data suggest that this stimulation occurs in trans (see Figs. 1c and 2b), we tested whether phosphorylated nucleosomes containing H3.3 could stimulate p300 activity on unphosphorylated canonical nucleosomes (Fig. 3b). Using mass spectrometry, we found that the presence of phosphorylated H3.3 nucleosomes resulted in increased acetylation on canonical H3 (Fig. 3c). Overall, these data support a role for H3.3 phosphorylation in stimulating p300 activity in trans (Fig. 3d).

As histone acetylation is correlated with gene expression, we compared transcription levels (RNA seq) from equal numbers of wild-type and H3.3 KO mESCs using synthetic spike-in standards. Despite global reduction of H3K27ac (Fig. 1), we did not observe global reduction of transcription in H3.3 KO compared to wild-type mESCs (Supplementary Fig. 6a). We next compared up- and down-regulated gene sets to the nearest neighboring genes of enhancers affected by H3.3 loss in mESCs (Supplementary Fig. 6b). Although genes with increased nearby acetylation generally showed increased transcription (480 out of 1,051, 46%), only 12% (344 out of 2,796) of genes with decreased nearby acetylation showed significantly reduced transcription in H3.3 KO mESCs. We observed a similar relationship between enhancer acetylation and steady state transcription in HIRA KO mESCs (Supplementary Fig. 6c–f). Overall,
found that H3.3 KO mESCs displayed delayed or defective differentiation, as nearest neighboring genes showed reduced transcription (Supplementary Fig. 7a). These regions did not show notable increases in H3K27ac that mimicking H3.3 phosphorylation allows p300 to function enzymatically to a greater extent. According to these findings coupled with our in vitro observations, we propose that phosphorylation of H3.3S31, in addition to other potentially phosphorylated residues that are common to both canonical and variant H3, provides a ‘phosphothreshold’ for p300 stimulation. In this scenario, H3.3 represents a nucleosomal cofactor for p300 that, through its deposition and phosphorylation, acts to integrate signaling information into chromatin to promote robust enhancer acetylation (Fig. 3d). This model is complementary to other proposed mechanisms of p300 activation, including autoacetylation and signal-mediated transcription factor dimerization.

Collectively, we show that H3.3 is functionally linked to enhancer acetylation in both mESCs and differentiating cells (Supplementary Note). We were surprised that loss of H3.3 in mESCs did not affect p300 recruitment but impacted p300 enzymatic activity, and furthermore, in trans. Although our data indicate that H3.3 deposition alone may stimulate p300, we observe that mimicking H3.3 phosphorylation allows p300 to function enzymatically to a greater extent. According to these findings coupled with our in vitro observations, we propose that phosphorylation of H3.3S31, in addition to other potentially phosphorylated residues that are common to both canonical and variant H3, provides a ‘phosphothreshold’ for p300 stimulation. In this scenario, H3.3 represents a nucleosomal cofactor for p300 that, through its deposition and phosphorylation, acts to integrate signaling information into chromatin to promote robust enhancer acetylation (Fig. 3d). This model is complementary to other proposed mechanisms of p300 activation, including autoacetylation and signal-mediated transcription factor dimerization.

H3.3S31ph has been reported previously during mitosis and was known to be enriched at repeat regions such as telomeres. 

As many latent enhancers are activated during differentiation, we wanted to address whether H3.3 is needed to establish acetylation during enhancer activation. We performed ChIP-seq of H3K27ac in wild-type and H3.3 KO cells that had been differentiated for 4 d into embryoid bodies (EBs). Genome-wide, we identified 2,054 enhancer sites (‘EB enhancers’) that had significantly higher H3K27ac in wild-type differentiated cells than in mESCs and were not pre-bound by p300 in mESCs (Fig. 4a and Supplementary Fig. 6g–i). These regions did not show notable increases in H3K27ac in differentiating cells lacking H3.3 (Fig. 4a). Furthermore, these enhancers remain ‘closed’ in the absence of H3.3, suggesting an early and fundamental role for H3.3 in activating latent enhancers (Fig. 4b). Unlike mESCs, differentiating cells were affected greatly by loss of H3.3, as nearest neighboring genes showed reduced transcription in H3.3-deleted EBs compared to wild-type EBs (Fig. 4c). In agreement with reduced enhancer activation and associated transcription, and in line with a role for H3.3 during development, we found that H3.3 KO mESCs displayed delayed or defective differentiation, evidenced both by morphology and an inability to switch from a pluripotency transcriptional program to those active upon differentiation (Supplementary Fig. 7b–g).

Last, to test the role of H3.3 phosphorylation during enhancer activation, we performed EB differentiation using our exogenous addback system. Expression of the phosphomimic H3.3S31E in H3.3 KO cells resulted in higher H3K27ac at EB enhancers than any other exogenous protein, including H3.3S31A (Fig. 4d and Supplementary Fig. 8a). Accordingly, H3.3S31E expression resulted in increased transcription from differentiation-specific genes and improved EB morphology, as compared to H3.3S31A expression (Fig. 4e and Supplementary Fig. 8b). Interestingly, expression of the constitutive phosphomimic H3.3S31E during differentiation also perpetuated H3K27ac at mESC-specific enhancers (Supplementary Fig. 8c), supporting our hypothesis that H3.3 phosphorylation promotes enhancer acetylation (Supplementary Fig. 8d).

Collectively, we show that H3.3 is functionally linked to enhancer acetylation in both mESCs and differentiating cells (Supplementary Note). We were surprised that loss of H3.3 in mESCs did not affect p300 recruitment but impacted p300 enzymatic activity, and furthermore, in trans. Although our data indicate that H3.3 deposition alone may stimulate p300, we observe that mimicking H3.3 phosphorylation allows p300 to function enzymatically to a greater extent. According to these findings coupled with our in vitro observations, we propose that phosphorylation of H3.3S31, in addition to other potentially phosphorylated residues that are common to both canonical and variant H3, provides a ‘phosphothreshold’ for p300 stimulation. In this scenario, H3.3 represents a nucleosomal cofactor for p300 that, through its deposition and phosphorylation, acts to integrate signaling information into chromatin to promote robust enhancer acetylation (Fig. 3d). This model is complementary to other proposed mechanisms of p300 activation, including autoacetylation and signal-mediated transcription factor dimerization.

H3.3S31ph has been reported previously during mitosis and was known to be enriched at repeat regions such as telomeres.

these results suggest that reduced enhancer acetylation is well tolerated in mESCs (Supplementary Fig. 6g–i).
Fig. 4 | H3.3 phosphorylation promotes the establishment of new enhancers during differentiation. a, Box plot showing H3K27ac enrichment at enhancers acetylated after 4 d of differentiation (n=2,054) in WT and H3.3 KO day 4 EBs. P<2.2×10^{-6} for all comparisons by Wilcoxon rank sum two-side test. The bottom and top of the boxes correspond to the 25th and 75th percentiles, and the internal band is the 50th percentile (median). The plot whiskers correspond to 1.5×interquartile range and outliers are excluded. D0, mESCs; D4, EB at day 4. b, Box plot showing ATAC-seq at enhancers acetylated after 4 d of differentiation (n=2,054) in WT and H3.3 KO day 4 EBs. P<2.2×10^{-6} for all comparisons by Wilcoxon rank sum two-side test. Box plots are displayed as in part a, c. Ratio (log2) of RNA expression from nearby genes of differentiation-specific enhancers (n=1,840) in day 4 EBs versus mESCs for WT and H3.3 KO cells. x axis values >0 indicate increased gene expression upon differentiation. P<2.2×10^{-6} between WT and H3.3 KO curves by Kolmogorov-Smirnov test. d, Heatmap of H3K27ac enrichment in WT and H3.3 KO cells expressing exogenous histone mutants at day 4 at differentiation-specific enhancers. For each analysis, 3 kb around the center of enhancers are shown. Each row represents a single enhancer (n=2,054). e, RT-qPCR analysis of differentiation-specific genes of day 4 EBs in H3.3 KO cells expressing exogenous histone mutants. Data represent mean±s.d (n=3).

Notably, our model is supported by our observation of H3.3S31ph enrichment at mESC enhancers and by a reported role for H3.3S31ph during gene activation35. In addition, Chk1, the reported H3.3S31 kinase35, has been implicated in gene regulation through histone phosphorylation, supporting a role for this cell-cycle kinase in mediating gene activation35. Moreover, binding of an H3.3 lysine 36 reader and elongation factor has been shown to be affected directly by serine 31 phosphorylation26,27, which suggests an important role for H3.3S31ph in guiding transcription. Taken together, our results highlight a functional role for H3.3 in gene regulation and offer important insights towards understanding how point mutations to this histone variant influence human health26,27.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0428-5.

Received: 4 March 2019; Accepted: 25 April 2019; Published online: 31 May 2019

References
1. Shi, L., Wen, H. & Shi, X. The histone variant H3.3 in transcriptional regulation and human disease. J. Mol. Biol. 429, 1934–1945 (2017).
2. Goldberg, A. D. et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell 140, 678–691 (2010).
3. Wong, L. H. et al. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. Genome Res. 20, 351–360 (2010).
4. Banaszynski, L. A. et al. Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. Cell 155, 107–120 (2013).
5. Elässer, S. J., Koh, K.-M., Diaz, N., Allis, C. D. & Banaszynski, L. A. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. Nature 522, 240–244 (2015).
6. Sadic, D. et al. Atrx promotes heterochromatin formation at retrotransposons. EMBO Rep. 16, 836–850 (2015).
7. He, Q. et al. The Daxx/Atrx complex protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation. Cell Stem Cell 17, 273–286 (2015).
8. Voon, H. P. J. et al. ATRX plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted genes. Cell Rep. 11, 405–418 (2015).
9. Jang, C.-W., Shibata, Y., Starmer, J., Yee, D. & Magnuson, T. Histone H3.3 maintains genome integrity during mammalian development. Genes Dev. 29, 1377–1392 (2015).
10. Hödl, M. & Basler, K. Transcription in the absence of histone H3.3. Genes Dev. 29, 1221–1226 (2015).
11. Sakai, A., Schwartz, B. E., Goldstein, S. & Ahmad, K. Transcriptional and developmental functions of the H3.3 histone variant in Drosophila. Curr. Biol. 19, 1816–1820 (2009).
12. Kim, H., Heo, K., Choi, J., Kim, K. & An, W. Histone variant H3.3 stimulates HSF70 transcription through cooperation with HP1γ. Nucleic Acids Res. 39, 8329–8341 (2011).
13. Tamura, T. et al. Inducible deposition of the histone variant H3.3 in interferon-stimulated genes. J. Biol. Chem. 284, 12217–12225 (2009).
14. Zhang, H. et al. RPA interacts with HIRA and regulates H3.3 deposition at gene regulatory elements in mammalian cells. Mol. Cell 65, 272–284 (2017).
15. Nashun, B. et al. Continuous histone replacement by hira is essential for normal transcriptional regulation and de novo DNA methylation during mouse oogenesis. Mol. Cell 60, 611–625 (2015).
16. Kong, Q. et al. Histone variant H3.3-mediated chromatin remodeling is essential for paternal genome activation in mouse preimplantation embryos. J. Biol. Chem. 293, 3829–3838 (2018).
31. Cheung, P. et al. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol. Cell* **34**, 347–361 (2016).

32. Li, M., Dong, Q. & Zhu, B. Aurora kinase B phosphorylates histone H3.3 at serine 31 during mitosis in mammalian cells. *J. Mol. Biol.* **429**, 2042–2045 (2017).

33. Devecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E. & Panne, D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat. Struct. Mol. Biol.* **20**, 1040–1046 (2013).

34. Ortega, E. et al. Transcription factor dimerization activates the p300 acetyltransferase. *Nature* **562**, 538–544 (2018).

35. Thorne, J. L., Ouboussad, L. & Lefevre, P. F. Heterochromatin protein 1 gamma and IκB kinase alpha interdependence during tumour necrosis factor gene transcription elongation in activated macrophages. *Nucleic Acids Res.* **40**, 7676–7689 (2012).

36. Shimada, M. et al. Chk1 is a histone H3 threonine 11 kinase that regulates DNA damage-induced transcriptional repression. *Cell* **132**, 221–232 (2008).

37. Guo, R. et al. B669/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol. Cell* **56**, 298–310 (2014).

38. Wen, H. et al. ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature* **508**, 263–268 (2014).

Acknowledgements

We thank members of the Banaszynski and Hon laboratories for helpful discussions; E. Duncan, H. Yu, I. Kraus, H. Zhit, and S. Morrison for critical comments on this manuscript; A. Shoshnev for the illustrations used as templates in this manuscript; UTSW BioHPC for computational infrastructure; UTSW McDermott Center and the Children's Research Institute for providing next-generation sequencing services; UTSW Flow Cytometry Core for single cell isolation; UTSW Proteomics Core for mass spectrometry analysis. L.A.B. is a Virginia Murchison Linnicuim Scholar in Medical Research (UTSW Endowed Scholars Program) and a Peterson Investigator of the Neuroendocrine Research Foundation (NETRF). This work was supported in part by CPRIT RR140042, The Welch Foundation F-1892, DoD KCP K170230, and NIH R35 GM124958 (L.A.B.), the American-Italian Cancer Foundation (S.M.), the Taiwan Postdoctoral Research Abroad Fellowship (Y.-C.T), and the Green Center for Reproductive Biology Sciences.

Author contributions

S.M. and L.A.B. conceived and designed the study; S.M. performed experiments with help from A.T., J.N., Y.-C.T., and M.T.; A.A.G. performed computational analyses; A.W. wrote the manuscript with input from all co-authors.
Methods

mESC culture. mESCs were cultured on gelatin-coated plates under standard serum and leukemia inhibitory factor (LIF) conditions at 37°C with 5% CO2 (KO-DMEM, 2 mM Glutamax, 15% ES grade fetal bovine serum, 0.1 mM 2-mercaptoethanol, 1× pen/strep (Pen/Strep), 1× NEAA, and LIF). During thawing and early passages, cells were maintained on an irradiated feeder layer. Cells were passaged at least twice to transfer them from the feeder to gelatin-coated plates. Generation of H3.3 KO, ATRX KO, DAXX KO, and HIRA KO mESCs have been described previously, and HIRA KO22 was generated in our laboratory using CRISPR–Cas9 (clustered regularly interspaced short palindromic repeats (CRISPR))–CRISPR-associated protein 9 (Cas9) genome editing. mESCs were tested routinely for mycoplasma.

Generation of HIRA KO using CRISPR–Cas9. HIRA KO22 was generating using a lentivirus CRISPR v2 plasmid (Addgene no. 52961) containing HIRA guide RNAs (gRNAs) (Supplementary Table 1). lentiviral CRISPR v2 vector (5 μg), psPAX2 (5 μg), and VSVG plasmids (0.5 μg) were transfected in serum-free media into 3 × 105 293T cells in a 10 cm2 tissue culture dish using Lipofectamine 3000. Lentivirus-containing supernatants were collected 48 and 72 h post transfection, pooled, and concentrated tenfold with Lenti-X (Clontech, 63123). mESCs (2×104) were transduced with 0.2 μl concentrated virus and 8 μg/ml polybrene for 24 h followed by selection with 1 μg/ml puromycin for 3 d. Single cells were isolated by cell sorting. Genotyping of individual clones confirmed deletion and immunoblot confirmed loss of protein.

Antibodies. H3 general (ab1791, Abcam, no. GR177884-2), H3.3 (9B-838, Millipore, no. 2578126), H3.3 (09-838, Millipore, no. 2578126), H3K4me1 (ab8895, Abcam, no. GR193882-2), H3K27ac (50253, Active Motif, no. 8308001), H3K27me3 (sc-5966, Santa Cruz, no. E30160), H3K9me3 (2118, Cell Signaling, no. 10), Lamin A (ab26300, Abcam, no. GR155781-1), β-Tubulin (T5201, Sigma, no. 088K4832), H3S10ph (39253, Active Motif, no. 8308001), H3S12ph (H3S12ph, Abcam, no. GR421512), H3S14ph (ab15906, Abcam, no. GR67227-8), Oct4 (sc-5279, Santa Cruz, no. C1509), anti-mouse IgG-HRP (NA937V, GE, no. 9732381), anti-rabbit IgG-HRP (170-6515, Bio-Rad, no. 350003248).

Chromatin immunoprecipitation. Native and crosslinking (X) ChIP were performed with 5 × 105 and 6 × 105 cells, respectively, as described previously.23 A spike-in normalization strategy was used to normalize all ChIP–seq data to reduce the effects of technical variation and sample processing bias. Spike-in chromatin was generated according to the manufacturer’s instructions.

Native ChIPs. ChIPs were trypsinized, washed and subjected to hypotonic lysis (50 mM Tris HCl, pH 7.4, 1 mM CaCl2, 0.2% Triton X-100, 10 mM Na-butyrate, 50 mM NaCl). Chromatin was sheared with a Bioruptor to an average size of 300 bp. ChIP DNA was purified and treated with RNase A and proteinase K. ChIP DNA was purified and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher). Libraries with unique adaptor barcodes were multiplexed and sequenced on an Illumina NextSeq 500 (paired-end, 33 base pair reads). Typical sequencing depth was at least 20 million reads per sample.

ChIP–seq data quality control, alignment and spike-in normalization. Quality of ChIP–seq data sets was assessed using the FastQC tool (v.0.10.1). ChIP–seq raw reads were aligned separately to the mouse reference genome (mm10) and the spike-in Drosophila reference genome (dm3) using Bowtie2 (v.2.3.2)24. Only one alignment is reported for each read (either the single best alignment or, if more than one equivalent best alignment was found, one of those matches selected randomly). Duplicate reads were filtered using the rmdup function of SAMTools (v.1.6.2).25 Uniquely mapped Drosophila reads were counted in the sample containing the least Drosophila reads and used to generate a normalization factor for random down sampling. Reads were converted into bigWig files using BEDTools (v.2.25.0)26 for visualization in Integrative Genomics Viewer (v.2.3).27

Peak calling. Peak calling was performed using MACCS2 software (v.2.1.2)28 using cut–value settings –vpvalue 1e–2 –mfold 10,50 and input as a control in each replicated cell. Peaks were filtered by peak height to reduce false positives and to retain only robust peaks for further analyses.

Average profiles. BigWig files were used to generate average ChIP–seq profiles using deepTools (v.3.0.2).29

Box plots. Box plot representations were used to quantitatively assess the read distribution in a fixed window. Box plots are defined by the median, box limits at upper and lower quartiles of 75% and 25%, and whiskers at 90% and 10%. The read distribution surrounding the peak center was calculated and plotted using custom R scripts. Wilcoxon rank sum tests were performed to determine the statistical significance of all comparisons.

Heatmaps. The read densities surrounding 6kb (±3kb) of the peak center of 300 peaks (rank ordered by highest to lowest H1K27ac enrichment in wild-type mESCs) were determined and visualized as heatmaps using deepTools (v.3.0.2).30

Density plots. Density plots representing fold-change differences between samples were generated using custom R scripts. Kolmogorov–Smirnov tests were performed to determine the statistical significance of all comparisons.

ATAC-seq. ATAC-seq was performed as described previously31 with minor changes. For each sample, 100,000 cells were collected, washed and lysed with ATAC buffer (Tris 10 mM, pH 7.4, 10 mM NaCl, 3 mM MgCl2, NP-40 0.1%). Nuclei were collected and subject to tagmentation at 37°C for 30 minutes. Reaction was stopped with 0.2% SDS and DNA was collected using Qiaquick PCR purification columns and eluted in 10 μl 10 mM Tris, pH 8. Eluted DNA was amplified using a KAPA non-hot-start PCR kit and purified using AMPure XP beads (positive and negative selection). Samples were pooled for multiplexing and sequencing using paired-end sequencing on the Illumina NextSeq 500.

ATAC-seq data quality control, alignment and normalization. Quality of the ATAC-seq data sets was assessed using the FastQc tool (v.0.10.1). The ATAC-seq reads were then aligned to the mouse reference genome (mm10) using BWA (v.0.7.5).32 For unique alignments, duplicate reads were filtered out. The resulting uniquely mapped reads were normalized to the same read depth across all samples and converted into bigWig files using BEDTools (v.2.25.0)33 for visualization in Integrative Genomics Viewer (v.2.3).34 Heatmaps were generated using deepTools (v.3.0.2).

Addback experiments. pCDH-EF1a-H3-Flag-HA-IRE6Neo constructs containing histone mutations were made using site-directed mutagenesis followed by Sanger sequencing to confirm the mutations (H3.2 and H3.3 sequences are listed in Supplementary Table 1). Lentivirus was packaged in 293T cells as described above. Wild-type and H3.3 KO mESCs (2 × 104) were incubated with 0.2 ml concentrated lentivirus and polybrene (8 μg/ml). The next day, the medium was replaced with complete mESCs culture medium containing 1 μg/ml puromycin. After 4 d of selection, mESCs were used for downstream analysis.

shRNA transduction. For p300 and CBP knockdown, (Dharmacon) plasmids were packaged as described above. Wild-type and H3.3 KO mESCs (2 × 104) were incubated with 0.2 ml concentrated lentivirus and polybrene (8 μg/ml). The next day, the medium was replaced with complete mESCs culture medium containing 1 μg/ml puromycin. After 4 d of selection, mESCs were used for western blot analysis.

In vitro phosphorylation and HAT assay. Full-length p300 was expressed in Sf9 cells and purified with 3× FLAG beads. Both recombinant human CHK1 protein

ChIP–seq. ChIP–seq library preparation. ChIP–seq libraries were prepared from 5–10ng ChIP DNA following the Illumina TruSeq protocol. The quality of the libraries was assessed using a D1000 ScreenTape on a 2200 TapeStation (Agilent) and quantified using a Qubit dsDNA HS Assay Kit (Thermo Fisher). Libraries with unique adaptor barcodes were multiplexed and sequenced on an Illumina NextSeq 500 (paired-end, 33 base pair reads). Typical sequencing depth was at least 20 million reads per sample.
and recombinant mononucleosomes were purchased from Active Motif (31163, 81070). First, H3.3 mononucleosomes were incubated with CHK1 (50 ng) in kinase buffer (50 mM Tris HCl, pH 7.4, 10 mM MgCl2, 1 mM DTT, 9% glycerol, protease inhibitor cocktail) in the presence or absence of 0.2 mM ATP at 37 °C for 5 min. The reaction was terminated by the addition of 5 μM CHK1 inhibitor SB218078 (Tocris, no. 2560). Immediately after kinase inhibition, the HAT (histone acetyltransferase) assay was performed by adding 25 μM acetyl-CoA (H acetyl-CoA in the case of HOT reaction), 100 ng of full-length p300 (plus H3.1 mononucleosomes in case of the trans assay) for 30 min at 30 °C. The reaction was stopped by adding SDS sample buffer and the samples were subjected to downstream analysis (immunoblot analysis, autoradiography, mass spectrometry).

Colony formation assay. Wild-type and H3.3 KO mESCs were seeded at a density of 100 cells per well in six-well plates. Medium was changed every 2–3 d. For all experiments, 7 d after initial seeding, cells were fixed and stained using the Alkaline Phosphatase Staining Kit (STEMGEN, 00-0055). Colonies were scored manually in a blinded fashion as undifferentiated, differentiated, or mixed.

Quantitative PCR with reverse transcription and mRNA-seq. mRNA was isolated using QIAGEN RNasy. Total RNA (500 ng) was reverse transcribed using random hexamers and MultiScribe reverse transcriptase. qPCR primer sequences used in this study are listed in Supplementary Table 1. For RNA-seq, total mRNA from equal cell numbers was mixed with synthetic RNA standards (ERCC RNA Spike-In Mix, Thermo Fisher)1. Libraries were prepared according to the Illumina TruSeq protocol and sequenced on an Illumina NextSeq 500 (paired-end, 33 base pair reads).

Analysis of RNA-seq data. Data quality control, alignment and normalization. Quality of the RNA-seq raw reads was assessed using the FastQC tool (v0.10.1). The reads were then aligned to the mouse reference genome (mm10) and the spike-in control ERCC92 using TopHat (v.2.0.9)45. Reads mapping to ERCC92 were counted using htseq-count46 and used to normalize the counts to genes. After normalization, the reads were converted into bigWig files using BEDTools (v.2.25.0) for visualization in Integrative Genomics Viewer (v.2.3) or the UCSC genome browser47.

Differential expression analysis. Gene expression levels measured as FPKM (fragments per kilobase of transcript per million mapped reads) was determined by the maximum likelihood estimation method implemented in the Cufflinks software package (v.2.2.1) with annotated transcripts as references. Differential expression was analyzed using the Student’s t test in the program Cufflinks (fragments per kilobase of transcript per million mapped reads) was determined by the maximum likelihood estimation method implemented in the Cufflinks software package (v.2.2.1) with annotated transcripts as references. Differential expression was analyzed using the Student’s t test in the program Cufflinks. The reads were then aligned to the mouse reference genome (mm10) and the spike-in control ERCC92 using TopHat (v.2.0.9)45. Reads mapping to ERCC92 were counted using htseq-count46 and used to normalize the counts to genes. After normalization, the reads were converted into bigWig files using BEDTools (v.2.25.0) for visualization in Integrative Genomics Viewer (v.2.3) or the UCSC genome browser47.

Venn diagrams. To check the extent of overlap of the up- and downregulated genes from RNA-seq with the nearest neighbor genes of the ChIP–seq peaks (obtained from GREAT software), Venn diagrams were generated using custom & scripts.

MA plots. MA plots were used to graphically represent genes that were up- or downregulated by more than twofold (Supplementary Fig. 6a,c). The log2 fold values corrected for multiple testing.

Quantification and statistical analysis. To check the significance of all comparisons, the Wilcoxon rank sum test was used to calculate P values for data used to generate box plots. The two-sample Kolmogorov–Smirnov test was used to calculate P values to show significant changes between two density curves. Analysis of variance (ANOVA) was used for two or more independent comparison groups.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data sets are deposited in the NCBI Gene Expression Omnibus under the following accession numbers: SuperSeries GSE114551, ATAC-seq GSE114547, ChIP–seq GSE114548, and RNA-seq GSE114549.

Code availability

Code to generate figures is available at https://github.com/utsw-medical-center-banaszynski-lab/Martire-et-al-2019-Nature-Genetics.git

References

39. Roberts, C. et al. Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. Mol. Cell. Biol. 22, 2318–2328 (2002).
40. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25 (2009).
41. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).
42. Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26 (2011).
43. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
44. Jiang, L. et al. Synthetic spike-in standards for RNA-seq experiments. Genome Res. 21, 1543–1551 (2011).
45. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111 (2009).
46. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

- LightCycler® 480 System Performance Data Software
- BMG Labtech Control Software
- MARS Data Analysis Software

Data analysis

- FastQC (v. 0.10.1), BOWTIE2 (v. 2.3.2), BWA (v. 0.7.5), SAMtools (v. 1.6), BEDtools (2.25.0), Integrative Genomics Viewer (IGV v. 2.3), UCSC genome browser, MACS (v. 2.1.2), TopHat (v. 2.0.9), Cufflinks (v. 2.2.1), Cuffdiff, R (v. 3.2.1), Prism (v. 8), Excel (v. 16.18), Deeptools (v. 3.0.2)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Datasets are generated during the current study are deposited in the NCBI Gene Expression Omnibus using series accession number GSE114551 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114551). Subseries are GSE 114567 (ATAC-seq), GSE114548 (ChIP-seq), GSE114549 (RNA-seq).
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [x] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

| RNA-seq studies were performed at least in duplicate for use with the DESeq2 pipeline. ATAC-seq studies were performed in duplicate. The majority of the Chip-Seq experiments were performed at least in duplicate, and for HIRA and H3.3, with two independent clonal KO mESCs. Biochemical figures were generated with a variety of replicates as reported in the figure legends of our manuscript. |

Data exclusions

No data were excluded from the analysis.

Replication

Critical ChIP-seq and RNA-seq data sets were obtained in duplicate. Many ChIP-seq data sets were validated with independent ChIP-qPCR assays and/or experiments with independent KO mESCs. All western blots and related experiments were performed in duplicate or more. Results obtained were consistent.

Randomization

Random allocation is not relevant to our study, since we are performing molecular and genomic analysis of known KO mESCs.

Blinding

Blinding was used to score differentiation state in our mESC colony formation assay. For the other experiments blinding was not relevant to our study; our techniques were all molecular biological techniques where the experimenter designs and executes the experimental conditions so blinding is not possible.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- [x] Antibodies
- [x] Eukaryotic cell lines
- [x] Palaeontology
- [ ] Animals and other organisms
- [ ] Human research participants
- [x] Clinical data

Methods

- [x] ChIP-seq
- [ ] Flow cytometry
- [ ] MRI-based neuroimaging

Antibodies

| Antibodies used | H3 general (ab1791, Abcam, Lot # GR177884-2), H3.3 (09-838, Millipore, Lot # 2578126), H3K4me1 (ab8895, Abcam, Lot # GR193882-2), H3K27ac (39133, Active Motif; Lot # 31814008), H3K27ac (39685, Active Motif, Lot # 14517014), H3K64ac (ABE1057, Millipore, Lot # 2826017), H3K122ac (ab33309, Abcam, Lot # GR284790-3), H3K18ac (ab1191, Abcam, Lot # GR3211480-1), H3.3/2 (ABE154, Millipore), p300 (sc-584, Santa Cruz, Lot # F3016), Gapdh (2118, Cell Signaling, Lot # 10), Lamin A (ab26300, Abcam, Lot # GR155781-1), b-Tubulin (T5201, Sigma, Lot # 088K4832), H3S10ph (39253, Active Motif, Lot # 8308001), H3S28ph (07-145, Millipore, Lot # 27707), H3.3S31ph (ab33319, Abcam, Lot # GR267227-8), Oct4 (sc-5279, Santa Cruz, Lot # C3109), anti-mouse IgG-HRP (NA93V, GE, Lot # 9773218), anti-rabbit IgG-HRP (170-6515, Biorad, Lot # 35003248), Spike-In antibody (61686, Active Motif, Lot # 00419007) |

Validation

Many histone modification antibodies, including several used in our study, have been validated using Brian Strahl’s histone tail peptide array (http://histoneantibodies.com). We validated the H3.3, HIRA, ATRX, and DAXX antibodies using our KO mESCs. We validated acetyl and phospho antibodies using peptide dot blots.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s) | H3.3 WT and KO mESCs - Banaszynski et al., Cell 2013; WT and ATRX KO and DAXX KO mESCs- Sadic et al., EMBO Rep 2015;
Cell line source(s)

WT and HIRA KO mESCs: Roberts et al., Mol Cell Biol 2002; WT and HIRA KO mESCs - this study
293T were purchased from ATCC
MFEI ("feeder") were derived from mice and subjected to radiation to mitotically inactivate the cells.

Authentication
KO cell lines were validated by sequencing of genomic DNA and by immunoblot of the protein product.

Mycoplasma contamination
All cell lines tested were negative for mycoplasma. Cells were tested monthly.

Commonly misidentified lines
None

Palaeontology

Specimen provenance
Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition
Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods
If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

ChIP-seq

Data deposition
☑ Confirm that both raw and final processed data have been deposited in a public database such as GEO.
☑ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

To review GEO accession GSE114551:
Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114551
Enter token qxenasoitnqpxyr into the box

Files in database submission

RNA-seq
ESC_H33WT_RNA-seq_rep1
ESC_H33WT_RNA-seq_rep2
ESC_H33KO_RNA-seq_rep1
ESC_H33KO_RNA-seq_rep2
EB_H33WT_RNA-seq_rep1
EB_H33WT_RNA-seq_rep2
ESC_H33KO_RNA-seq_rep1
ESC_H33KO_RNA-seq_rep2
ESC_HIRAWT1_RNA-seq_rep1
ESC_HIRAWT1_RNA-seq_rep2
ESC_HIRAKO1_RNA-seq_rep1
ESC_HIRAKO1_RNA-seq_rep2
ESC_HIRAWT2_RNA-seq_rep1
ESC_HIRAWT2_RNA-seq_rep2
ESC_HIRAKO2_RNA-seq_rep1
ESC_HIRAKO2_RNA-seq_rep2
ESC_H33WT_ATAC
ESC_H33KO_ATAC
ESC_HIRAWT_ATAC
ESC_HIRAKO1_ATAC
EB_H33WT_ATAC
EB_H33KO_ATAC
ChIP-seq
ESC_H33WT_H3K27ac_pAb
ESC_H33KO_H3K27ac_pAb
ESC_H33WT_H3K4me1_rep1
ESC_H33KO_H3K4me1_rep1
ESC_H33WT_H3K4me1_rep2
ESC_H33KO_H3K4me1_rep2
ESC_H33WT_H3K64ac
ESC_H33KO_H3K64ac
Methodology

Replicates

The following experiments were performed in replicate and are in high agreement:

WT/H3.3 KO mESC H3K4me1 ChIP-seq
WT/H3.3 KO mESC H3K27ac ChIP-seq
WT/H3.3 KO EB H3K27ac ChIP-seq

In another sense, the H3.3 KO and HIRA KO samples serve as biological replicate demonstrating the pathway responsible for our observations.

Sequencing depth

Sample / Total Reads / Unique Reads / Read length / PE or SE RNA-seq
ESC_H33WT_RNA-seq_rep1 39,152,592 36,983,355 75 bp PE
ESC_H33WT_RNA-seq_rep2 37,095,065 35,092,925 75 bp PE
ESC_H33KO_RNA-seq_rep1 38,675,988 36,453,961 75 bp PE
ESC_H33KO_RNA-seq_rep2 38,259,977 36,042,510 75 bp PE
ESC_H33WT_RNA-seq_rep1 32,802,966 30,875,934 75 bp PE
ESC_H33WT_RNA-seq_rep2 37,095,745 35,103,958 75 bp PE
ESC_H33WT_RNA-seq_rep1 37,394,096 35,290,777 75 bp PE
ESC_H33WT_RNA-seq_rep2 37,281,876 35,113,024 75 bp PE
ESC_HIRAWT1_RNA-seq_rep1 59,081,868 55,403,722 75 bp PE
ESC_HIRAWT1_RNA-seq_rep2 44,774,879 42,751,379 75 bp PE
ESC_HIRAKO1_RNA-seq_rep1 40,676,497 38,829,512 75 bp PE
ESC_HIRAKO1_RNA-seq_rep2 49,642,924 47,352,169 75 bp PE
ESC_HIRAWT2_RNA-seq_rep1 40,836,408 39,056,154 75 bp PE

http://genome.ucsc.edu/cgi-bin/hgTracks?db=mm10&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr17%3A35506032%2D35510777&hsid=668907157_UstPhYeQZkjoaCD2CF2oED4caLw
ESC_HIRAWT2_RNA-seq_rep2 50,653,012 48,185,546 75 bp PE
ESC_HIRAKO2_RNA-seq_rep1 49,334,622 47,082,156 75 bp PE
ESC_HIRAKO2_RNA-seq_rep2 50,099,462 48,762,189 75 bp PE
ESC_HIRAWT_ATAC 56,067,263 36,567,115 75 bp PE
ESC_H3KO_ATAC 59,541,728 38,433,171 75 bp PE
ESC_HIRAKO1_ATAC 4,750,640 4,424,468 75 bp PE
ESC_HIRAWT_ATAC 23,125,315 19,554,683 75 bp PE
ESC_H3KO_ATAC 32,599,756 30,186,597 75 bp PE

**Antibodies**

H3.3 (09-838, Millipore, Lot # 2578126), H3K4me1 (ab8895, Abcam, Lot # GR193882-2), H3K27ac (39133, Active Motif, Lot #31814008), H3K27ac (39685, Active Motif, Lot # 14517014), H3K64ac (ABE1057, Millipore, Lot # 2826017), H3K122ac (ab33309, Abcam, Lot # GR284790-3), p300 (sc-584, Santa Cruz, Lot # F3016).
Peak calling parameters

macs2 callpeak -t $experiment -c $control -f BED -n $out_dir/$prefix $genomesize -p 1e-2 --mfold 10,50 --nomodel --shift 0 --extsize $fraglen --keep-dup all --SPMR

Data quality

FastQC tool was used to perform a quality check on all sequencing data to assess the quality of the raw reads and identify possible sequencing errors or biases to ensure that the raw data looks good and can be used for further downstream analyses. Peak calling by MACS2 reported over 70% of peaks that had above 5-fold enrichment. Enrichment metrics of pvalue 1e-2 and mfold 10,50 were used for calling peaks to ensure high-confidence enrichment against background.

Software

FastQC, BOWTIE, BWA, SAMtools, BEDtools, MACS2, R, Deeptools, Treeview