Histone H3 globular domain acetylation identifies a new class of enhancers

Madapura M Pradeepa1,2, Graeme R Grimes1, Yatendra Kumar1, Gabrielle Olley1, Gillian C A Taylor1, Robert Schneider3,4 & Wendy A Bickmore1

Histone acetylation is generally associated with active chromatin, but most studies have focused on the acetylation of histone tails. Various histone H3 and H4 tail acetylations mark the promoters of active genes1. These modifications include acetylation of histone H3 at lysine 27 (H3K27ac), which blocks Polycomb-mediated trimethylation of H3K27 (H3K27me3)2. H3K27ac is also widely used to identify active enhancers3,4, and the assumption has been that profiling H3K27ac is a comprehensive way of cataloguing the set of active enhancers in mammalian cell types. Here we show that acetylation of lysine residues in the globular domain of histone H3 (lysine 64 (H3K64ac) and lysine 122 (H3K122ac)) marks active gene promoters and also a subset of active enhancers. Moreover, we find a new class of active functional enhancers that is marked by H3K122ac but lacks H3K27ac. This work suggests that, to identify enhancers, a more comprehensive analysis of histone acetylation is required than has previously been considered.

Covalent modifications in the globular domains of the core histones have been implicated in a variety of chromatin functions5. Post-translational modifications located on the lateral (outer) surface of the histone octamer can alter contacts between histones and nucleosomal DNA and directly affect chromatin structure5. Acetylation of histone H3 at lysine 56 (H3K56ac) is associated with DNA unwrapping from the nucleosome and has been implicated in chromatin assembly and genome stability6. H3K64ac, a modification at the lysine residue at the start of the first α helix in the histone fold domain (HFD), destabilizes nucleosomes and facilitates nucleosome dynamics in vitro7. Metylation of H3K64 is implicated in heterochromatin establishment8. Histone–DNA interactions reach their maximum strength in the nucleosome dyad, and, unlike acetylation on histone tails, H3K122ac is sufficient to stimulate transcription in vitro from chromatinized templates9 and promote nucleosome disassembly10.

Metagene analysis of H3K122ac and H3K64ac chromatin immunoprecipitation and sequencing (ChIP-seq) reads from mouse embryonic stem cells (mESCs) showed that these marks correlate with the magnitude of gene expression (Fig. 1a). Surprisingly, given the link between histone acetylation and active chromatin, we found H3K122ac marks over a subset of inactive or poised genes that are repressed by Polycomb complexes in mESCs (Fig. 1b,c). Sequential ChIP–qPCR analysis confirmed the presence of H3K122ac on bivalently marked nucleosomes (with H3K27me3 and trimethylation of histone H3 at lysine 4 (H3K4me3)) (Fig. 1d).

Pearson correlation analysis across multiple histone modifications in mESCs indicated that H3K64ac and H3K122ac cluster with each other and with monomethylation of H3K4 (H3K4me1) (Fig. 1e), a marker for enhancers11. H3K122ac and H3K64ac reads were also enriched at active promoters and strong enhancers across hidden Markov model (HMM)–based chromatin states (ChromHMM)12,13 (Supplementary Fig. 1). Given these findings, we aligned H3K64ac, H3K122ac and H3K27ac ChIP-seq data with the midpoints of enhancers in mESCs, as defined by H3K4me1 peaks 2 kb away from RefSeq transcription start sites (TSSs)11 (Fig. 2a).

The data clustered into three groups (Supplementary Data 1) on the basis of overlap of H3K4me1 peaks with peaks for H3K27ac and H3K122ac. Group 1 enhancers (n = 23,153) were H3K27ac+ and were, for the most part, also marked by significantly high levels of H3K122ac and H3K64ac (P < 0.01, Wilcoxon rank-sum test; Supplementary Table 1). This group of enhancers would be classified as active on the basis of H3K27ac status3,4. At the other extreme of the distribution, group 3 enhancers (n = 5,265) were negative for all three acetylation marks and would be classified as inactive enhancers. Group 2 enhancers (n = 9,340) were H3K27ac− but were marked by significantly high levels of H3K122ac and H3K64ac (P < 0.01, Wilcoxon rank-sum test; Supplementary Table 1). This group of enhancers would be classified as inactive or poised. H3K122ac (which co-occupies promoters with acetylation of the histone variant H2A.Z (H2A.Zac)9 and can induce transcription14) and H2A.Zac were comparably enriched in group 1 and group 2 enhancers (Fig. 2b). Group 2 enhancers also had high levels of binding for EP300, which acetylates H3K64, H3K122 and H3K27 (refs. 7,9,15).

We found that group 1 enhancers had high levels of H3K122ac and H3K64ac (Fig. 2b). A subset of the clustered enhancers

1MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh, UK. 2School of Biological Sciences, University of Essex, Colchester, UK. 3Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U964, Université de Strasbourg, Illkirch, France. 4Helmholtz Zentrum München, Institute of Functional Epigenetics, Neuherberg, Germany. Correspondence should be addressed to M.M.P. (pmadap@essex.ac.uk) or W.A.B. (wendy.bickmore@igmm.ed.ac.uk).
associated with highly expressed genes in ESCs—termed ‘super-enhancers’ (ref. 16)—was also heavily enriched for H3K64ac (Fig. 2a,d) and Supplementary Fig. 3a). Our data suggest that there are a class of putative regulatory elements (group 2 in Fig. 2) in mESCs that is marked by H3K122ac and/or H3K64ac but lacks the H3K27ac marks that are usually used as a predictor of active enhancers. Gene Ontology (GO) analysis of the subclases indicated that both the H3K27ac+ and H3K27ac−H3K122ac+ groups of enhancers are associated with terms such as ‘stem cell maintenance’. But H3K27ac+ enhancers are also significantly enriched (P < 0.01) for terms associated with cell adhesion, which are lacking for the H3K27ac−H3K122ac+ set. Instead, terms for hindbrain morphogenesis, placental development and germ layer formation were prominent in this set (Supplementary Fig. 3a). A subclass of group 2 enhancers, which were H3K27me3+, were enriched for terms associated with negative regulation of transcription, differentiation and development (Supplementary Fig. 3b).

Analysis of transcription factor motif enrichment found that binding sites for SP1, SP2, SP4, KLF5, EGR1, TFAP2a, TFAP2b and TFAP2c, which we note generally have a high GC content, were enriched in group 2 enhancers (Supplementary Fig. 4a). In comparison to group 1 and group 3 enhancers, group 2 enhancers also had higher levels of H3K27me3 and H2A.Z (Fig. 2b)—both markers for poised promoters and enhancers17,18. A subset of group 2 enhancers with H3K27me3 peaks were enriched for unmethylated CpG islands (CGIs) (Supplementary Fig. 4b), which are located at promoters and enhancers19. Bidirectional transcription of enhancers correlates with enhancer activity20; however, the transcripts are degraded by the exosome complex, making them difficult to detect. Analysis of exosome-sensitive RNAs (eRNAs; identified by comparing RNA-seq reads from mESCs null for Exosc3 (encoding exosome component 3; Exosc3−/−) with wild-type mESCs)21 showed that group 2 enhancers transcribe high levels of eRNAs (Fig. 2c).

We tested the enhancer activity of these elements using luciferase reporter assays in mESCs; a well-characterized Nanog enhancer22 (Fig. 2d) served as a positive control. Group 2 genomic regions (H3K27ac−) with enrichment for H3K122ac (Fig. 3a,b) exhibited 4- to 120-fold higher activity than the negative control (empty vector) and were equally, or more, active than the Nanog enhancer. Similarly, enhancer assays performed in a human breast adenocarcinoma cell line (MCF-7) showed that H3K27ac−H3K122ac+ enhancers3 had greater reporter activity than H3K27ac+ enhancers (Fig. 3c,d).

To demonstrate the in vivo functional importance of group 2 enhancers, we used the CRISPR/Cas9 approach23 to delete them from the mESC genome (Fig. 4a,b). As positive controls, we also deleted one allele of the super-enhancers located near Nanog and Klf4 (Fig. 2c), leading to a significant reduction in Nanog and Klf4 expression, respectively (Fig. 4c). In mESCs with deletion of the Nanog super-enhancer, there was no difference in the expression of Dppa3, located 80 kb upstream of Nanog. Expression of Rad23b, located 180 kb downstream of Klf4, was somewhat affected by deletion of the intervening enhancer. Homozygous deletion of
the group 2 putative enhancer 42 kb downstream of *Lif* (*Lif* 42k en−/) led to reduced expression of *Lif* but not of the flanking gene *Hormad2* (Fig. 4c). Similarly, deletion of one allele of the putative enhancer 30 kb upstream of *Tbx3* (*Tbx3* 30k en−/) led to down-regulation of *Tbx3*.

To examine whether histone acetylation is important for the function of these newly identified regulatory elements, we used dCas9 to recruit the SiDhx repressor complex to them (Fig. 4d). In positive-control experiments, recruitment of dCas9–SiDhx to the *Nanog* enhancer and to the super-enchancers of *Nanog*, *Klf4* and *Sox2* led to significant reduction in expression of the respective target genes but not other nearby genes (Fig. 4e).

For the group 2 enhancers analyzed, ChIP–qPCR analysis showed that recruitment of dCas9–Sid4x effectively reduced the levels of H3K122ac at the target *Tbx3* 30k en enhancer, with no effect on the off-target control (Sox2 super- enhancer) (Fig. 4d). RT–qPCR analysis showed reduced expression of putative target genes upon Sid4x recruitment to the *Foxd3* 57k en, *Tbx3* 30k en, *Sox2* 40k en and *Sox2* 60k en enhancers but not of control genes (Fig. 4e).

To investigate H3K122ac as an enhancer mark in more detail, we performed ChIP-seq for H3K122ac, H3K27ac and H3K4me1 in a human erythroleukemic (K562) cell line. As in mESCs, H3K122ac was enriched in chromatin with active promoter, strong enhancer and poised promoter states (ChromHMM) in K562 cells (Fig. 5a,b). H3K122ac was also enriched at super-enchancers and H3K27ac+ enhancers (Fig. 5c–e). As in mESCs, a subset of H3K27ac+ enhancers was marked with H3K122ac (Fig. 5c–e), had DNase 1–hypersensitive sites (DHSs) and was bound by transcription factors (Fig. 5e and Supplementary Data 2). Analysis of enrichment for transcription factor binding using Encyclopedia of DNA Elements (ENCODE) ChIP-seq data showed that group 2 enhancers were enriched for CTCF, ZNF143, SMC3, RAD21, EZH2 and USF1 in comparison to group 1 enhancers (Supplementary Fig. 4c).

Rather than a simple definition of active enhancers as regions marked by H3K4me1 and H3K27ac, a more complex picture of different histone acetylation marks at enhancers is emerging25. Our data suggest that using H3K27ac alone gives an incomplete catalog of the active enhancer repertoire and that acetylation of histone H3 on the lateral surface of the histone octamer can be used to identify a new class of active enhancers that has no significant enrichment for H3K27ac.

Lysine acetyltransferases (KATs) generally have relaxed substrate specificity, with the exception of KAT8, which acetylates lysine 16 of histone H4 (H4K16; refs. 25,26) and is critical for the maintenance of mESC pluripotency and differentiation27,28. H4K16 acetylation (H4K16ac) marks active enhancers in mESCs, including some that lack H3K27ac25. Like the globular domain acetylations of histone H3, acetylation at this position is required for gene activation and can promote strong enhancer activity24.
The role of most histone acetylation marks at enhancers suggests that opening of local chromatin structure might be an important facet of enhancer function and may be important to the physical properties of the nucleosome.

Figure 3  In vitro enhancer assays. (a) Representative H3K27ac− group 2 putative enhancers from mESCs marked with H3K122ac (PE2, PE5) and a negative-control group 3 putative enhancer lacking all histone acetylation marks tested (C2) are shown as in Figure 2d. The regions used for cloning into the enhancer reporter vector (pGL4.26) are indicated by gray boxes and are detailed in Supplementary Table 3. (b) Luciferase reporter assays were performed in mESCs for the genetically defined enhancer of Nanog (Nanog en) and H3K27ac− putative active enhancers randomly chosen on the basis of the presence of H3K122ac (PE1–PE5, Sox2−40k en), H3K64ac (PE6, PE7), and both H3K122ac and H3K64ac (PE8). Additionally, regions with H3K4me1 but no acetylation (C1, C2) were assayed, and empty vector (pGL4.26) served as negative control. Mean log2-transformed fold change relative to empty vector in luciferase activity is plotted; error bars, s.e.m. from two biological replicates each with two technical replicates (n = 4). (c) Putative enhancers from MCF-7 cells9 shown as in a; transcription factor (TF) ChIP peaks from ENCODE are shown below. Genome coordinates are from the GRCh37/hg19 assembly of the human genome. (d) Luciferase reporter assays performed in MCF-7 cells for randomly chosen H3K27ac− enhancers (G1E1, G1E2) and H3K27ac−H3K122ac− putative human enhancers (PE1–PE6) are shown as in b. Nanog enhancer (Nanog en, as in b) and vector alone served as controls (Supplementary Table 3). Mean log2-transformed fold change in luciferase activity is plotted; error bars, s.e.m. from two biological replicates each with two technical replicates (n = 4).
Contre La Cancer (Equipe Labellise) and INSERM Plan Cancer (Epigénétique et Cancer).

AUTHOR CONTRIBUTIONS
M.M.P., Y.K., G.O. and G.C.A.T. performed the experiments. M.M.P. and G.R.G. conceived the project, designed experiments and wrote the manuscript. All authors contributed to writing, read the manuscript and provided feedback.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Wang, Z. et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. 40, 897–903 (2008).
2. Kim, T.W. et al. Ctbp2 modulates NuRD-mediated deacetylation of H3K27 and facilitates PRC2-mediated H3K27me3 in active embryonic stem cell genes during exit from pluripotency. Stem Cells 33, 2442–2455 (2015).
3. Creyghton, M.P. et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA 107, 21931–21936 (2010).
4. Rada-Iglesias, A. et al. A unique chromatin signature uncovers early developmental enhancers in humans. Nature 470, 279–283 (2011).
5. Trogerber, P. & Schneider, R. Scratching the (lateral) surface of chromatin regulation by histone modifications. Nat. Struct. Mol. Biol. 20, 657–661 (2013).
6. Neumann, H. et al. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. Mol. Cell 36, 153–163 (2009).
7. Di Cerbo, V. et al. Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. elife 3, e01632 (2014).
8. Daujat, S. et al. H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. Nat. Struct. Mol. Biol. 16, 777–781 (2009).
9. Trogerber, P. et al. Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. Cell 152, 859–872 (2013).
10. Simon, M. et al. Histone fold modifications control nucleosome unwrapping and disassembly. Proc. Natl. Acad. Sci. USA 108, 12711–12716 (2011).
11. Heintzman, N.D. et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. 39, 311–318 (2007).
12. Ernst, J. et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473, 43–49 (2011).
13. Bogu, G.K. et al. Chromatin and RNA maps reveal regulatory long noncoding RNAs in mouse. Mol. Cell. Biol. 36, 809–819 (2015).
14. Hailey, J.E., Kaplan, T., Wang, A.Y., Kober, M.S. & Rine, J. Roles for H2A.Z and its acetylation in GAL1 transcription and gene induction, but not GAL1-transcriptional memory. PLoS Biol. 8, e1000401 (2010).
15. Tie, F. et al. Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. Development 141, 1129–1139 (2014).
16. Whyte, W.A. et al. Master transcription factors and Mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319 (2013).
17. Zentner, G.E., Tesar, P.J. & Scacheri, P.C. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. 21, 1273–1283 (2011).
Figure 5 H3K122ac marks at K562 enhancers. (a) Enrichment values for H3K122ac, H3K27ac, H3K4me1, H3K27me3 and H3K4me3 ChIP and input reads from K562 cells across ChromHMM segmentations(12). Ten, transcription; CNV, copy number variation. (b) Box plots showing log-transformed ChIP-seq RPM distributions (median, line inside the box). The IQR shows 50% of the data, and the whiskers extend to 1.5 times the IQR. (c,d) Box plots (c) and heat maps (d) showing enrichment (RPM) of H3K122ac, H3K27ac and H3K4me1 in K562 cells across five groups of enhancers, grouped on the basis of acetylation patterns: super-enhancers; enhancers marked with H3K27ac and H3K122ac (group I); enhancers lacking H3K27ac but marked with H3K122ac (group II); enhancers with H3K27ac but not H3K122ac (group III); and enhancers lacking both H3K27ac and H3K122ac (group IV). (e) UCSC Genome Browser tracks (RP10M) showing H3K27ac, H3K122ac and H3K4me1 ChIP and input reads from K562 cells for super-enhancers and group I and II enhancers. Transcription factor ChIP data, DHS clusters, and K562 ChromHMM + Segway tracks are shown below. The genomic coordinates of K562 cell enhancers are listed in Supplementary Data 2.

18. Voigt, P., Tee, W.W. & Reinberg, D. A double take on bivalent promoters. Genes Dev. 27, 1318–1338 (2013).
19. Long, H.K., et al. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. eLife 2, e00348 (2013).
20. Andersson, R., et al. An atlas of active enhancers across human cell types and tissues. Nature 507, 455–461 (2014).
21. Pefanis, E., et al. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. Cell 161, 774–789 (2015).
22. Jiang, J., et al. A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat. Cell Biol. 10, 353–360 (2008).
23. Mali, P., et al. RNA-guided human genome engineering via Cas9. Science 339, 823–826 (2013).
24. Konermann, S., et al. Optical control of mammalian endogenous transcription and epigenetic states. Nature 500, 472–476 (2013).
25. Taylor, G.C., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M.M. & Bickmore, W.A. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. Genome Res. 23, 2053–2065 (2013).
26. Smith, E.R., et al. A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. Mol. Cell. Biol. 25, 9175–9188 (2005).
27. Li, X., et al. The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. Cell Stem Cell 11, 163–178 (2012).
28. Ravens, S., et al. Mof-associated complexes have overlapping and unique roles in regulating pluripotency in embryonic stem cells and during differentiation. eLife 3, 1–23 (2014).
29. Shogren-Knaak, M., et al. Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science 311, 844–847 (2006).
30. Lovén, J., et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320–334 (2013).
Sequential chromatin immunoprecipitation. Antibodies recognizing H3K122ac and H3K64ac were previously described. mESCs were cross-linked in 1% formaldehyde for 10 min, and reactions were quenched by the addition of glycine to a final concentration of 0.125 M. Chromatin was sheared using a Bioruptor (Diagenode) to an average fragment length of ~100–200 bp. Sequential ChIP was performed as described previously. Briefly, 5 μg of antibodies against H3K4me3 (07–473, Millipore) and H3K27me3 (07–449, Millipore) was covalently coupled to Dynabeads using the Invitrogen Dynabeads Antibody Coupling kit (14311D) according to the manufacturer’s instructions. The first ChIP was performed using antibody to either H3K4me3 or H4K27me3, and the immunoprecipitated chromatin was then eluted with 10 mM DTt and diluted 30-fold with RIPA buffer (1× PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and Protease Inhibitor Cocktail (Roche)). The second ChIP was performed with 1 μg of antibody to H3K122ac. Purified chromatin was quantified by qPCR using the standard-curve method, and quantitation was expressed as the percentage of input bound. Primer details are given in Supplementary Table 2.

Native chromatin immunoprecipitation. mESCs and K562 cells (10 × 10^6) were centrifuged at 500 g for 3 min, washed twice in PBS and resuspended in 200 μl of NBA buffer (85 mM NaCl, 5.5% sucrose, 10 mM Tris-HCl, pH 7.5, 0.2 mM EDTA, 0.2 mM PMSF, 1 mM DTt and 1× protease inhibitors (Calbiochem, 539134-1SET)). Cells were lysed by the addition of an equal volume of 0.1% NP-40 in NBA buffer and incubated on ice for 3 min. Nuclei were pelleted at 2,000 g for 3 min at 4 °C, washed with NBR buffer (85 mM NaCl, 5.5% sucrose, 10 mM Tris-HCl, pH 7.5, 3 mM MgCl2, 1.5 mM CaCl2, 0.2 mM PMSF and 1 mM DTt) and pelleted at 2,000 g for 3 min at 4 °C. Nuclei were resuspended (10 × 10^6 nuclei/ml) in NBR buffer supplemented with RNase A (20 μg/ml) and incubated at 20 °C for 5 min. Chromatin was fragmented for 30 min at 20 °C using 0.133 U/μl microccocal nuclease (MNase; Boehringer- units; Sigma-Aldrich, N3755-5000U; titrated to give predominantly mono-nucleosomes). Digestion was stopped with the addition of an equal volume of STOP buffer (215 mM NaCl, 10 mM Tris-HCl, pH 8.0, 20 mM EDTA, 5.5% sucrose, 2% Triton X-100, 0.2 mM PMSF, 1 mM DTt and 2× protease inhibitors), and digested nuclei were left on ice overnight to release soluble, fragmented chromatin. Chromatin was precleared by centrifugation at 12,000 g for 10 min at 4 °C, and the soluble chromatin (supernatant) was transferred to a fresh tube. Five percent of the released chromatin was retained as input, and the remainder was incubated for 4 h at 4 °C on a rotating wheel with ~5 μg of antibody (against H3K122ac, H3K64ac, H3K4me1 (Abcam, ab8895, lot GR251663-1) or H3K27ac (Abcam, ab4729, lot GR254707-1)) pre-coupled to protein A Dynabeads (Life Technologies, 10002D) in PBS containing 5 mg/ml BSA and 0.1 mM PMSF. Immune complexes bound to beads were washed with five washes with buffer 1 (150 mM NaCl, 10 mM Tris-HCl, pH 8.0, 2 mM EDTA, 1% NP-40 and 1% sodium deoxycholate) on a rotating wheel for 5 min each wash and once in room-temperature TE buffer for 1 min. Chromatin was released from the beads by incubation with 0.1 M NaHCO3 in 1% SDS for 30 min at 37 °C followed by the addition of protease K (100 μg/ml) and Tris, pH 6.8 (100 μM) and incubation at 55 °C overnight. For both native and cross-linked ChiP, Dynabeads were removed using a magnetic rack and chromatin was purified using QiAquick PCR Purification columns (Qiagen) according to the manufacturer’s instructions.

ChiP-seq library preparation and deep sequencing. Libraries were prepared as previously described with the following modifications. No purification was performed between the polyadenylation and ligation reactions. After the polyadenylation reaction, enzymes were inactivated by incubation at 75 °C for 20 min, and the ligation reaction was supplemented with ligation reagents (400 U of T4 DNA ligase (New England BioLabs), 1× Buffer 2 (New England BioLabs), 7.5% PEG-6000, 1 mM ATP and 13.3 mM annealed Illumina adaptors (AU)) and incubated at 16 °C overnight. Size selection following the ligation and PCR steps was performed with 1× and 0.8× reaction volumes, respectively, on Agencourt AMPure XP beads (Beckman Coulter, A68800).

Replicate 1 of the H3K122ac and H3K64ac ChiP analyses was sequenced at the Danish National High-Throughput DNA Sequencing Center (Copenhagen; 42-base single-end reads). Replicate 2 of the H3K122ac and H3K64ac ChiP analyses, two replicates of H3K27ac ChiP analysis, and all ChiP and input samples prepared from K562 cells were sequenced at Edinburgh Genomics (University of Edinburgh; 50-base single-end reads).

Read mapping. Fastq files were aligned to the reference genome using Bowtie (version 0.12.8) with parameters set to retain uniquely mapped reads with a maximum of two mismatches (Bowtie options: -e 40 -m 1 -v 2). For mapping, mm9 and hg18 Bowtie indexes were used for the mouse (mESC) and human (K562 and MCF-7) data sets, respectively. Mapped reads from two biological replicates of H3K27ac, H3K122ac and H3K64ac ChiP were merged for further analysis.

Peak calling. Peaks were called using SICER. For mESCs, MNase-digested ChiP input DNA (GEO, GSM1156619) was used as a background control for H3K27ac, H3K64ac and H3K122ac. For H3K4me1 in mESCs (E14TG2a, GEO, GSM103750), input (GEO, GSM1003746) was used as a background control. mESC biological replicates were merged using SAMtools (v0.1.19) before peak calling with SICER (v1.1). SICER parameters were as follows: window size, 200 bp for H3K122ac, H3K64ac, H3K27ac and H3K4me1 fragment size, 150 bp; false discovery rate, 0.01; and gap size, 600 bp.

Generation of BedGraphs for visualization on the UCSC Genome Browser. BedGraphs for each histone mark were generated from aligned read files using the HOMER software suite (v4.7) at a resolution of 10 bp and with a normalized tag count of 10 million. Mapped reads from two biological replicates for H3K122ac, H3K64ac and H3K27ac ChiP in mESCs were combined in generating the BedGraphs for Figures 1–4. UCSC Genome Browser tracks for individual replicates covering representative loci are shown in Supplementary Figure 2. Similarly, data from single experiments for MCF-7 and K562 ChiP-seq analyses were processed to generate BedGraphs for visualization in the UCSC Genome Browser.

Heat maps and average profiles. Heat maps and average profile for RefSeq gene TSSs (±2 kb), RefSeq gene TESs (±2 kb) and enhancer midpoints (±2 kb) and for the entire length of super-enhancers (all scaled to an equivalent length of 4 kb) were generated using ngsplot (version 2.61) (ref. 37).

For Figure 1a, gene expression quartiles from high (Q4) to low (Q1) were obtained from our previous study and used to generate average profile plots for H3K122ac and H3K64ac across TSSs and TESs as detailed above. The heat map in Figure 1b was generated for the TSSs (±2 kb) of genes that have been shown to be repressed by Polycomb complexes.

The average profile plots (Fig. 2a) for enrichment of strand-specific RNA-seq reads in Exonic+/− relative to wild-type mice (Sequence Read Archive (SRA), SRP042355) for the three enhancer groups were generated using ngsplot v2.61 (ref. 37).

Genome-wide correlation analysis of histone marks. Pearson’s correlation coefficients were calculated between data sets using the bamCorrelate tool (version 1.5.9; removing duplicate reads and using a resolution of 10 kb). The correlation matrix was hierarchically clustered and visualized using the Bioconductor package pheatmap.

ChromHMM analysis. To calculate the distribution of histone marks across different chromatin states, the bamCorrelate tool was used to count the reads within ChromHMM segments for K562 cells and mESCs. Data sets were normalized to RPM values.

Enhancer analysis. Enhancers were defined as having H3K4me1 peaks, with gene TSS (RefSeq TSS ± 2 kb) and genome blacklist regions removed. Active enhancer regions (group 1) were defined as genomic intervals overlapping

DOI: 10.1038/ng.3550
both H3K4me1 and H3K27ac peaks. Inactive enhancers, defined as genomic regions with H3K4me1 peaks with no associated H3K27ac peaks, were stratified into groups 2 and 3, representing enhancers with and without associated H3K122ac peaks, respectively (Supplementary Data 1 and 2). Peak intersections were performed using the BEDTools \(^41\) (v2.23.0) intersect function. Super-enhancer coordinates for K562 cells and mESCs were obtained from the Super-Enhancer Archive (SEA). H3K27me3 peak regions were called using MACS2 (ref. 42; v2.1.0; broadpeak with no-input control).

Transcription factor motif enrichment analysis. Transcription factor motif enrichment analysis was performed using the Regulatory Sequence Analysis Tools (RSAT) server. Nucleotide sequences from group 2 enhancers (H3K122ac \(^+\) in mESCs) were used as input for transcription factor motif enrichment analysis with group 1 enhancer coordinates as the background.

Gene ontology enrichment analysis. GO (Biological Process) enrichment analysis was performed using GREAT \(^43\). Bed files from group 1 and group 2 enhancers intersecting with H3K27me3 peaks (H3K122ac \(^+\)H3K27me3 \(^+\)) and group 2 enhancers lacking H3K27me3 (H3K27me3 \(^-\)H3K122ac \(^+\)) were used as input and the whole genome was used as background to select significantly enriched GO terms for nearby genes.

Enrichment analysis of DNase I–hypersensitive sites and unmethylated CpG islands. To determine the enrichment of DHSs (GEO, GSM1014154) and CGIs at subgroups of enhancers, a Fisher’s exact test was performed using DNase I–hypersensitive sites and unmethylated CpG islands. To determine the enrichment of DHSs (GEO, GSM1014154) and CGIs at subgroups of enhancers, a Fisher’s exact test was performed using DNase I–hypersensitive sites and unmethylated CpG islands. Transcription factor motif enrichment analysis with group 1 enhancer coordinates as the background.

Transcription factor motif enrichment analysis. Transcription factor motif enrichment analysis was performed using the Regulatory Sequence Analysis Tools (RSAT) server. Nucleotide sequences from group 2 enhancers (H3K122ac \(^+\) in mESCs) were used as input for transcription factor motif enrichment analysis with group 1 enhancer coordinates as the background.

Gene ontology enrichment analysis. GO (Biological Process) enrichment analysis was performed using GREAT \(^43\). Bed files from group 1 and group 2 enhancers intersecting with H3K27me3 peaks (H3K122ac \(^+\)H3K27me3 \(^+\)) and group 2 enhancers lacking H3K27me3 (H3K27me3 \(^-\)H3K122ac \(^+\)) were used as input and the whole genome was used as background to select significantly enriched GO terms for nearby genes.

Enrichment analysis of DNase I–hypersensitive sites and unmethylated CpG islands. To determine the enrichment of DHSs (GEO, GSM1014154) and CGIs at subgroups of enhancers, a Fisher’s exact test was performed using DNase I–hypersensitive sites and unmethylated CpG islands. Transcription factor motif enrichment analysis with group 1 enhancer coordinates as the background.

Transcription factor motif enrichment analysis. Transcription factor motif enrichment analysis was performed using the Regulatory Sequence Analysis Tools (RSAT) server. Nucleotide sequences from group 2 enhancers (H3K122ac \(^+\) in mESCs) were used as input for transcription factor motif enrichment analysis with group 1 enhancer coordinates as the background.

Gene ontology enrichment analysis. GO (Biological Process) enrichment analysis was performed using GREAT \(^43\). Bed files from group 1 and group 2 enhancers intersecting with H3K27me3 peaks (H3K122ac \(^+\)H3K27me3 \(^+\)) and group 2 enhancers lacking H3K27me3 (H3K27me3 \(^-\)H3K122ac \(^+\)) were used as input and the whole genome was used as background to select significantly enriched GO terms for nearby genes.