ZFP462 safeguards neural lineage specification by targeting G9A/GLP-mediated heterochromatin to silence enhancers

Ramesh Yelagandula [1,2,7] ✉, **Karin Stecher**[1,3,7], **Maria Novatchkova**[4], **Luca Michetti** [1,6], **Georg Michlits**[1], **Jingkui Wang**[4], **Pablo Hofbauer**[1], **Gintautas Vainorius**[1], **Carina Pribitzer** [4], **Luke Isbel** [5], **Sasha Mendjan**[1], **Dirk Schübeler** [5], **Ulrich Elling** [1], **Julius Brennecke** [1] **& Oliver Bell** [1,2] ✉

ZNF462 haploinsufficiency is linked to Weiss–Kruszka syndrome, a genetic disorder characterized by neurodevelopmental defects, including autism. Though conserved in vertebrates and essential for embryonic development, the molecular functions of *ZNF462* remain unclear. We identified its murine homologue ZFP462 in a screen for mediators of epigenetic gene silencing. Here we show that ZFP462 safeguards neural lineage specification of mouse embryonic stem cells (ESCs) by targeting the H3K9-specific histone methyltransferase complex G9A/GLP to silence meso-endodermal genes. ZFP462 binds to transposable elements that are potential enhancers harbouring pluripotency and meso-endoderm transcription factor binding sites. Recruiting G9A/GLP, ZFP462 seeds heterochromatin, restricting transcription factor binding. Loss of ZFP462 in ESCs results in increased chromatin accessibility at target sites and ectopic expression of meso-endodermal genes. Taken together, ZFP462 confers lineage and locus specificity to the broadly expressed epigenetic regulator G9A/GLP. Our results suggest that aberrant activation of lineage non-specific genes in the neuronal lineage underlies ZNF462-associated neurodevelopmental pathology.

A totipotent zygote develops into a multicellular organism via cell specialization[1–3]. Cell fate specification is primarily controlled by transcription factors (TFs) that activate lineage-specific genes[4], but the silencing of lineage non-specific genes is also important. Silencing often involves histone methyltransferases (HMTases) that establish repressive chromatin modifications, including histone H3 lysine 9 di- and trimethylation (H3K9me2/3) (ref. 5). H3K9me2/3 binds heterochromatin protein 1 isoforms (HP1α, β and γ) (ref. 6), and chromatin-bound HP1 can recruit additional heterochromatin modifiers, thus promoting chromatin compaction and transcriptional silencing[7,8]. Epigenetic feedback mechanisms promote stable inheritance of heterochromatin, gene silencing and cell identity[5,9–13].

H3K9me2 is catalysed by an HMTase complex comprising G9A (also known as euchromatic histone methyltransferase 2, EHMT2) and G9-like protein (GLP, also known as euchromatic histone methyltransferase 1, EHMT1) (ref. [14]), which is essential for embryogenesis[15]. G9A/GLP-dependent heterochromatin modifications at promoters and enhancers have been proposed to block TF binding, thereby preventing aberrant expression of lineage non-specific genes[16–19]. Indeed, G9A/GLP-dependent heterochromatin silences the pluripotency-linked gene *Oct3/4* (also known as *Pou5f1*) during embryonic stem cell (ESC) differentiation, preventing its expression in somatic tissues[20,21]. Inactivation of G9A facilitates *Oct3/4* induction during induced pluripotent stem cell reprogramming of somatic cells[22]. G9A/GLP is proposed to regulate neurodevelopment[23] by interacting with the RE1-silencing TF (REST also known as NRSF) to target and repress neuronal genes in non-neuronal cells[24–26]. G9A/GLP also has important and distinct functions in silencing both non-neuronal and early neuron progenitor genes in mature neurons[27,28]. Notably, de novo mutations in genes encoding either GLP or proteins that cooperate in facultative heterochromatin formation cause the neurodevelopmental disorder Kleefstra syndrome[29]. Although the biochemical functions and developmental importance of G9A/GLP are well established, the mechanisms underlying its precise spatiotemporal silencing of lineage non-specific genes remain largely unclear[14,30,31]. G9A/GLP are broadly expressed and lack sequence specificity; therefore, heterochromatin mediated by G9A/GLP requires yet-unknown cell-type-specific interactions with TFs.

In this Article, we identified zinc finger protein 462 (ZFP462) in a clustered regularly interspaced short palindromic repeats (CRISPR) genetic screen for effectors of heterochromatin-mediated, heritable silencing of *Oct3/4*. ZFP462 is a vertebrate-specific, putative TF of unknown function. Notably, its human orthologue, *ZNF462*, was recently identified as a high-confidence risk gene for a neurodevelopmental disorder (OMIM# 618619) (refs. [32–34]). We demonstrate that ZFP462 is required to silence inappropriate meso-endodermal gene expression in mouse ESCs and neural progenitor cells (NPCs). ZFP462 silences lineage non-specific genes by recruiting G9A/GLP to deposit heterochromatin modifications at enhancers, thereby limiting their accessibility to pluripotency TFs. Thus, ZFP462 precisely targets formation of G9A/GLP-dependent heterochromatin to prevent unscheduled activation of meso-endodermal genes and maintain cell fate identity. Our findings illuminate the developmental regulation of G9A/GLP-dependent gene silencing and suggest that neural phenotypes associated with *ZNF462* haploinsufficiency arise from defects in cell fate specification during embryogenesis.

Results

Zfp462 encodes a regulator of heterochromatic gene silencing

To dissect the mechanisms underlying heterochromatin-mediated silencing of a lineage non-specific gene, we focused on *Oct3/4* (refs. [20,22,35]). *Oct3/4* is highly expressed in ESCs encoding OCT4 TF essential for pluripotency and self-renewal[36], but is inactivated and acquires H3K9me2 upon differentiation. Subsequent HP1 binding and DNA methylation irreversibly silence *Oct3/4* (refs. [21,22,35]).

Tethering the chromoshadow domain of HP1α (HP1) to a genetically modified *Oct3/4* allele in mouse ESCs is sufficient to recruit H3K9-specific HMTases and recapitulate heterochromatin-mediated gene silencing without affecting the wild-type allele[13]. To enable efficient, reversible HP1 tethering via the Tet-OFF system we established an ESC clone with Tet operator sites (TetO) and a puromycin-BFP reporter gene downstream of the *Oct4-GFP* reporter allele (Fig. 1a) that expresses HP1 fused to a FLAG-Tet repressor domain (TetR-FLAG-HP1) (Extended Data Fig. 1a). This ESC clone also expressed hCas9 to enable CRISPR-based screening in ESCs.

TetR-FLAG-HP1 binds to TetO sites with high affinity, but doxycycline (Dox) addition abolishes this interaction (Fig. 1b). TetR-FLAG-HP1, but not TetR-FLAG alone, resulted in transcriptional silencing of the *Oct4-GFP* reporter (Fig. 1c and Extended Data Fig. 1c), concomitant with loss of the active mark H3K4me3, acquisition of the inactive marks H3K9me2 and H3K9me3 (Fig. 1d), and enrichment of endogenous HP1γ (Extended Data Fig. 1d)[13]. Importantly, transcriptional silencing and inactive histone modifications were maintained even after the addition of Dox (Fig. 1c,d and Extended Data Fig. 1d). In agreement with previous results, using small-molecule inhibition of DNA methyltransferases[13], we found that genetic *Dnmt1* deletion and loss of global DNA methylation (Extended Data Fig. 1e) did not affect the initiation of heterochromatin-dependent reporter gene silencing but compromised its maintenance upon Dox addition (Fig. 1c). We infer that ectopic heterochromatin recapitulates G9A/GLP-induced silencing of the endogenous *Oct3/4* gene in somatic cells[20], including the stable inheritance of heterochromatin.

To identify genes required for heterochromatin-dependent silencing of *Oct3/4*, we performed a pooled CRISPR screen with unique molecular identifiers (UMIs), which enables single-cell analysis of mutant phenotypes[37]. The UMI CRISPR library contained approximately 27,000 single guide RNAs (sgRNAs) targeting all annotated mouse nuclear protein-coding genes with four sgRNAs per gene. Each sgRNA was paired with thousands of barcodes representing UMIs improving signal-to-noise ratio and hit calling. We transduced with the pooled library into the hCas9-expressing CiA reporter cells, selected with G418 for 5 days, treated with Dox for 2 days, and then isolated GFP-positive cells by fluorescence-activated cell sorting (FACS) (Fig. 1e). The unsorted population served as background control. Relative enrichment of sgRNAs was determined by sequencing UMIs in both populations followed by statistical analysis using MAGeCK[38]. A total of 130 genes were significantly enriched in the GFP-positive cell population (P < 0.001) (Fig. 1e and Supplementary Table 1), including *Dnmt1*, *Uhrf1*, *Setdb1*, *G9a*, *GLP*, *Atf7ip*, *Daxx*, *Atrx*, *KMT5C* and genes encoding members of the HUSH complex. These factors have all been linked to DNA methylation or H3K9me2/3-dependent gene silencing[14,39,40], supporting the known crosstalk between these pathways in heterochromatin[13,39]

The uncharacterized *Zfp462* gene scored very highly in the screen but has not been associated with heterochromatin regulation. We validated loss of heritable *Oct4-GFP* silencing in *Zfp462* mutant CiA ESCs by independent CRISPR–Cas9 editing (Fig. 1f and Extended Data Fig. 1f).

ZFP462 represses through interaction with G9A/GLP and HP1γ

To further investigate ZFP462, we engineered ESCs with an Avi-GFP-FLAG tag inserted into the endogenous *Zfp462* gene (*Avi-Zfp462*) (Fig. 2a) and identified ZFP462-interacting proteins in ESCs. Several well-known co-repressors affinity purified with ZFP462 (Fig. 2b and Supplementary Table 2), including HP1γ, G9A, GLP and WIZ, a known G9A/GLP interacting protein[41]. Similarly, ZFP462, HP1γ, G9A and WIZ affinity purified with Avi-fused endogenous GLP (Avi-GLP) (Fig. 2c and Supplementary Table 3)[41]. Thus, ZFP462 interacts with co-repressor proteins, suggesting it is involved in heterochromatin regulation.

ZFP462 is a highly conserved, vertebrate-specific C_2H_2-type zinc finger protein (Fig. 2d and Extended Data Fig. 2a) C_2H_2-type zinc fingers are typically associated with DNA binding activity but may also contribute to protein–protein interactions[41,42]. To determine whether ZFP462 recruitment is sufficient to silence the *Oct3/4* reporter, we transduced CiA ESCs with TetR-FLAG fusions of full-length ZF462 (FL), or of N-terminal (NT), mid (Mid), C-terminal truncation (NT+Mid) and C-terminal (CT) fragments (Fig. 2d and Extended Data Fig. 2b). TetR-FLAG-ZFP462-FL targeting led to loss of GFP expression in over 80% of CiA reporter cells (Fig. 2e and Extended Data Fig. 2c). TetR-FLAG-ZFP462-CT also resulted in substantial GFP silencing and interacted with G9A, GLP, WIZ and HP1γ by co-immunoprecipitation (co-IP) (Fig. 2f and Supplementary Table 4). In contrast, GFP silencing was not observed in CiA ESCs expressing TetR-FLAG fusions of ZFP462-NT, ZFP462-Mid or ZFP462-NT+Mid. Thus, ZFP462 associates
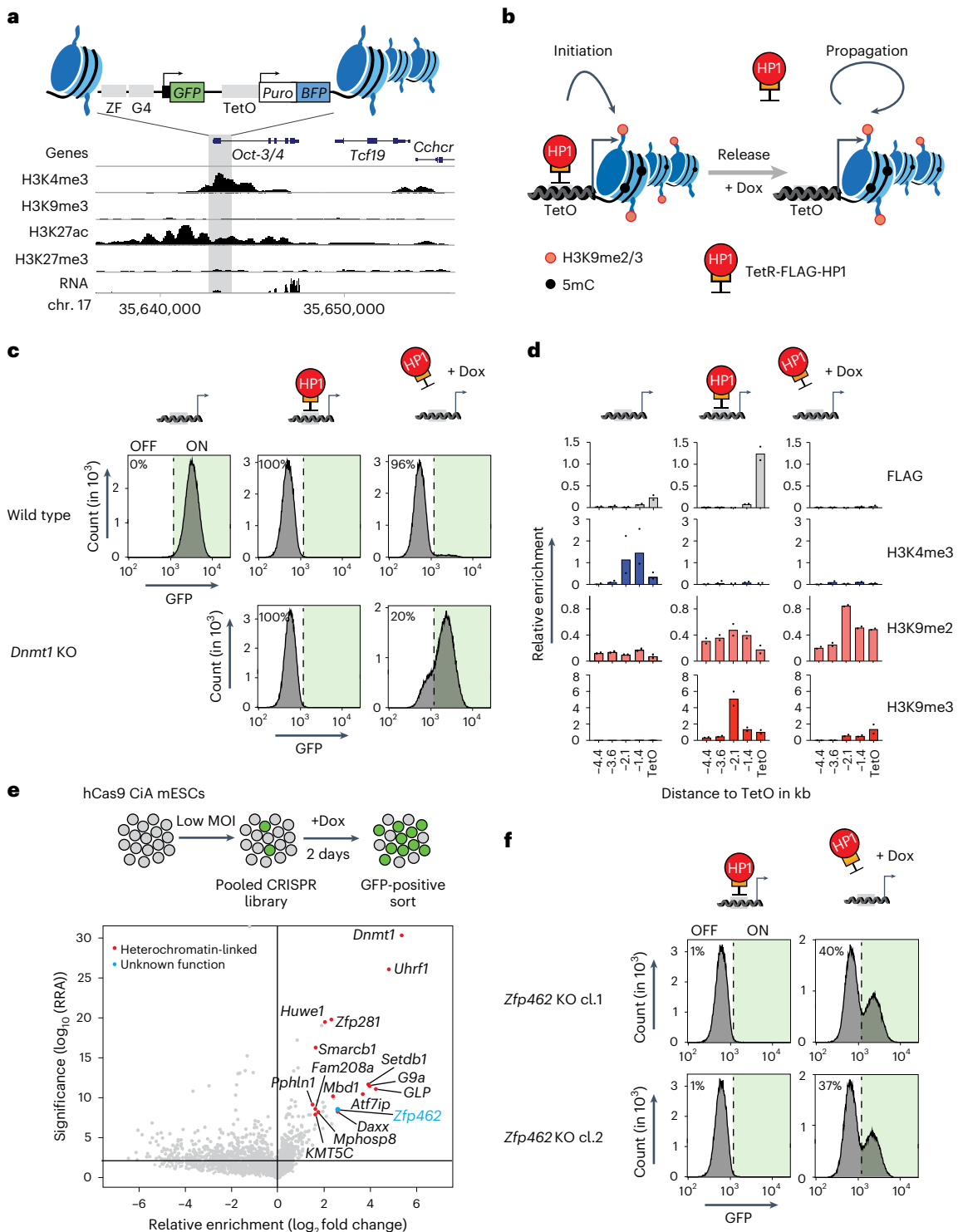
Fig. 1 | CRISPR screen identifies heterochromatin regulators required for heritable *Oct3/4* gene silencing. a, Design of the CiA *Oct4* dual reporter locus in ESCs. One of the *Oct3/4* alleles was modified in ESCs by inserting seven Tet operator sites (TetO) flanked by a GFP and a BFP reporter gene on either side. GFP expression is under control of the *Oct3/4* promoter, whereas a PGK promoter drives BFP expression. The genomic screen shot (below) shows histone modifications and RNA expression at the *Oct3/4* locus in wild-type ESCs. ZF, 12 binding sites for ZFHD1; G4, five binding sites for GAL4. **b**, Scheme of the experimental design. TetR facilitates reversible HP1 tethering to TetO binding sites to establish heterochromatin and silence both GFP and BFP reporters. Dox addition releases TetR binding to distinguish heritable maintenance of chromatin modifications and gene silencing in the absence of the sequence-specific stimulus. **c**, Flow cytometry histograms of wild-type and *Dnmt1* KO CiA *Oct4* dual reporter ESCs show GFP expression before

TetR-FLAG-HP1 tethering, in the presence of TetR-FLAG-HP1 and after 4 days of Dox-dependent release of TetR-FLAG-HP1. Percentages indicate fraction of GFP-negative cells. **d**, ChIP–qPCR shows relative enrichment of TetR-HP1 (FLAG) and histone modifications surrounding TetO before TetR-FLAG-HP1 tethering, in the presence of TetR-FLAG-HP1 and after 4 days of Dox-dependent release of TetR-FLAG-HP1 (*n* = 2 independent biological replicates). **e**, Scheme of CRISPR screen design. MOI, multiplicity of infection. Volcano plot shows enrichment (log fold change GFP-positive sorted versus unsorted cells) and corresponding significance (−log$_{10}$ MAGeCK significance score) of genes in CRISPR screen (*n* represents mean of three independent experiments). **f**, Flow cytometry histograms show GFP expression of two independent *Zfp462* knockout (KO) CiA *Oct4* dual reporter cell lines in the presence of TetR-FLAG-HP1 and after 4 days of Dox-dependent release of TetR-FLAG-HP1. Percentages indicate fraction of GFP-positive cells.
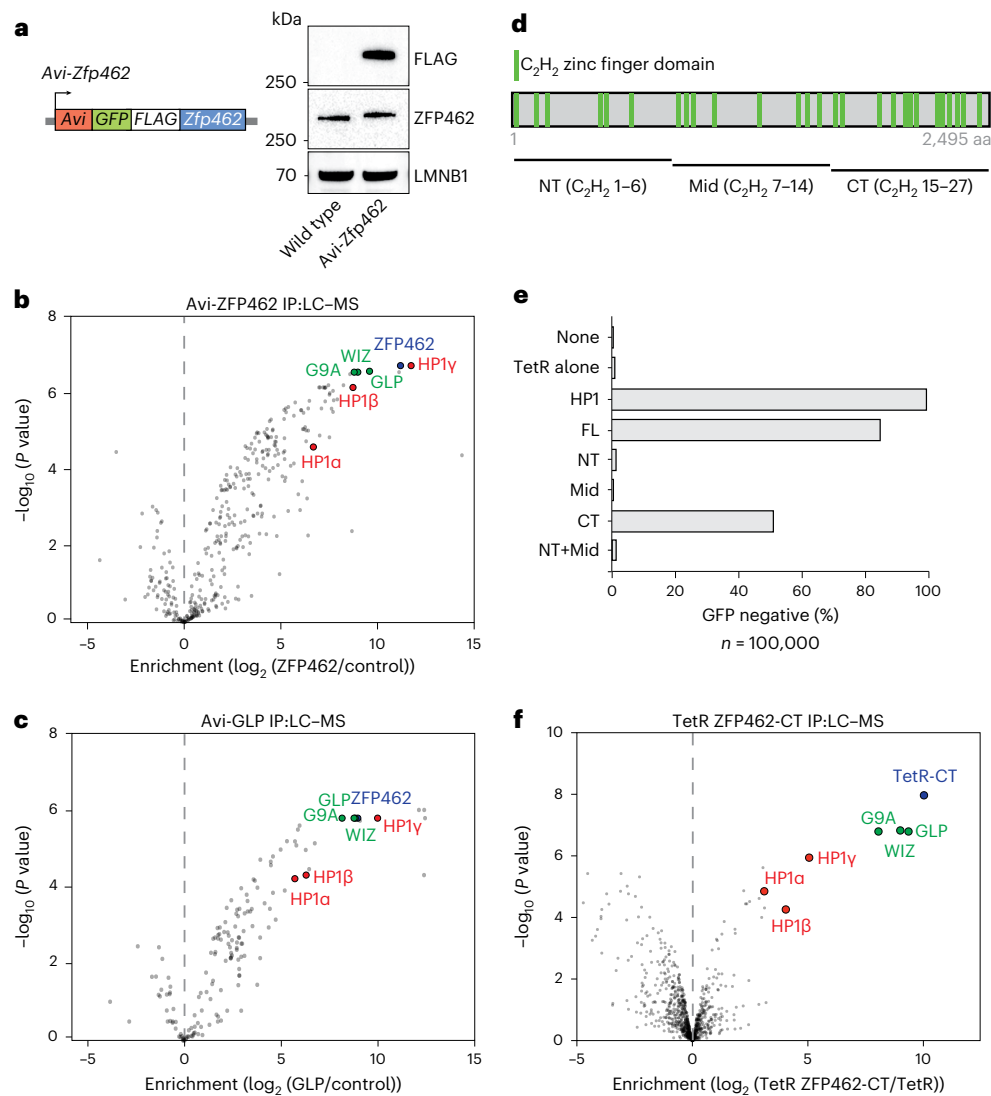
Fig. 2 | ZFP462 elicits silencing function through interaction with G9A/GLP and HP1γ. a, Design of Avi-*Zfp462* ESCs and western blot validation. Mouse ESCs expressing biotin ligase (BirA) were used to modify the endogenous *Zfp462* gene by inserting the Avi-GFP-3XFLAG tag downstream of the translation start codon. Western blot with FLAG and ZFP462 antibodies confirms ZFP462 tagging. **b,c**, LC–MS analysis of Avi-tagged ZFP462 (**b**) and Avi-tagged GLP (**c**) ESCs. Volcano plots show enrichment and corresponding significance of co-purified proteins (*n* = 3 replicates). **d**, Schematic of ZFP462 protein depicts locations of 27 C_2H_2 zinc finger domains (green bars). Fragments used to generate TetR fusions for tethering in CiA *Oct4* dual reporter assay are indicated below. **e**, Bar plot shows percentage of GFP-negative CiA *Oct4* ESCs measured by flow cytometry in response to ectopic TetR fusion protein expression (y axis). **f**, Volcano plot of LC–MS analysis compares enrichment and corresponding significance of co-purified proteins between TetR-FLAG-ZFP462-CT and TetR-FLAG (*n* = 3 replicates).

with G9A/GLP and HP1γ via its C-terminal domain and is sufficient to initiate transcriptional gene silencing.

ZFP462 represses primitive endoderm differentiation

To investigate the role of ZFP462 during mouse embryogenesis, we first examined available single-cell RNA sequencing (RNA-seq) data of three developmental stages representing exit from pluripotency and primary germ-layer specification[43]. Consistent with its role in regulating heterochromatin in ESCs, *Zfp462* was initially expressed in epiblast stem cells (embryonic day (E)4.5). At gastrulation onset, *Zfp462* displayed almost exclusive expression in pluripotent and neural cells (E6.5–7, Extended Data Fig. 3a).

Next, we generated homozygous *Zfp462* mutant ESCs using CRISPR-Cas9 (*Zfp462*[−/−]). Additionally, we engineered premature stop codons into one allele of *Zfp462* using CRISPR–Cas9-assisted homology-dependent repair (*Zfp462*[+/Y1195*] and *Zfp462*[+/R1257*]). *Zfp462*[+/R1257*] mutant ESCs mimic human ZNF462 haploinsufficiency

associated with Weiss–Kruzska syndrome (*ZNF462*[+/R1263*]) (ref. [32]). We isolated two independent *Zfp462* homozygous mutant clones and one clone of each *Zfp462* heterozygous mutation and confirmed reduction and loss of ZFP462 protein (Fig. 3a) and heterozygous point mutation (Extended Data Fig. 3b). Compared with wild-type ESCs, heterozygous and homozygous *Zfp462* mutant ESCs displayed morphological changes with dispersed, refractile cells spreading out of characteristically densely packed ESC colonies (Fig. 3b and Extended Data Fig. 3c). Further, *Zfp462* mutant ESC colonies showed reduced expression of the pluripotency marker alkaline phosphatase, suggesting a function for ZFP462 in ESC self-renewal. RNA-seq revealed differential expression of ~1,800 genes in homozygous *Zfp462* mutant ESCs and ~1,400 genes in heterozygous *Zfp462* mutant ESCs (cut-off: log₂ fold change (LFC) ≥1, adjusted *P* value ≤0.01) (Fig. 3c and Extended Data Fig. 3d, e). Gene Ontology (GO) terms of differentially expressed genes were related to developmental processes, consistent with the spontaneous ESC differentiation observed upon *Zfp462* deletion (Fig. 3d and Extended
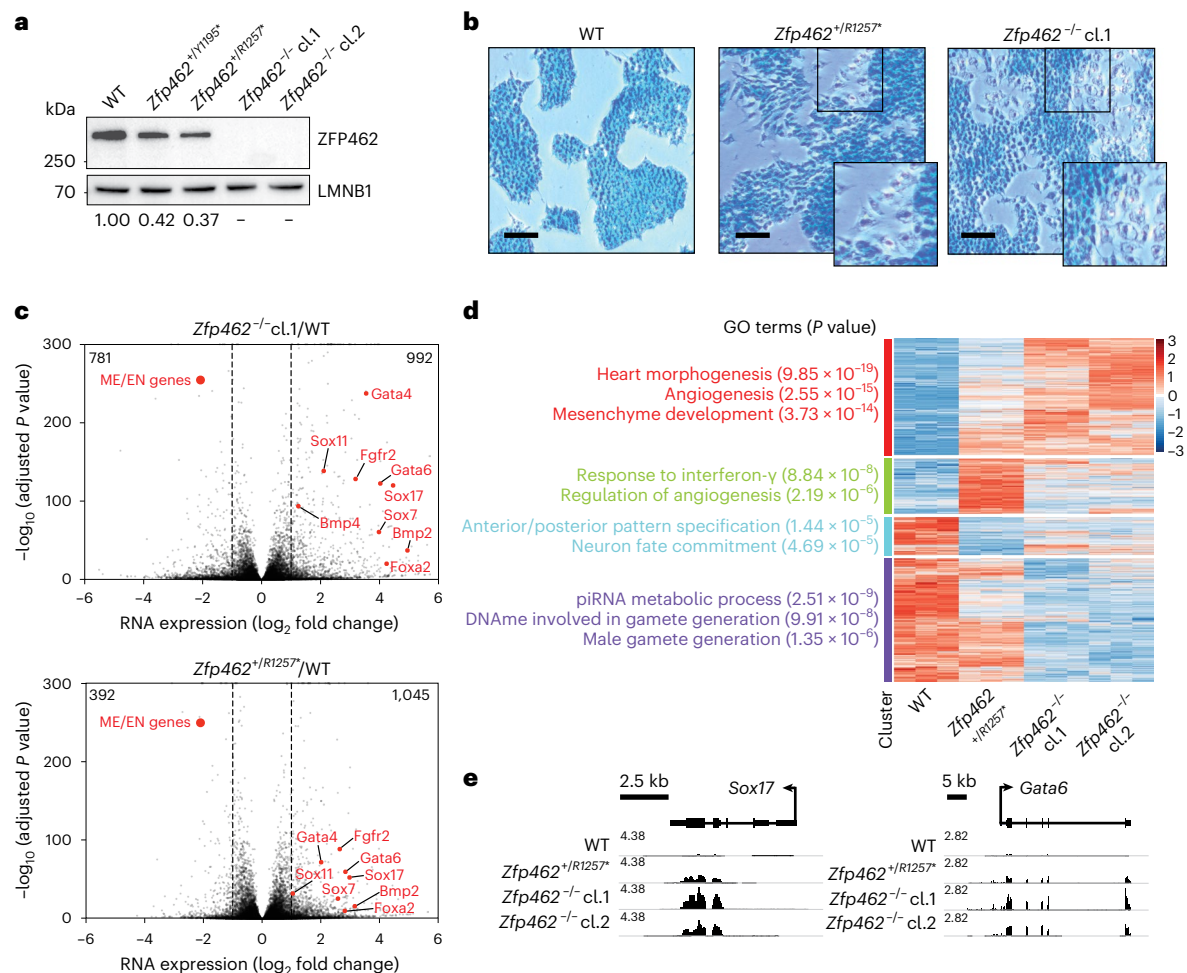
Fig. 3 | Depletion of *Zfp462* leads to aberrant expression of lineage-specifying genes. a, Western blot shows ZFP462 protein expression in wild-type (WT), two heterozygous and two homozygous *Zfp462* mutant ESC lines. LMNB1 serves as loading control and reference for relative ZFP462 quantification (numbers below). **b**, Alkaline phosphatase staining of WT, heterozygous and homozygous *Zfp462* mutant ESCs. Enlarged region is marked as square in the image. Scale bar, 100 μm. **c**, Volcano plots show gene expression changes in homozygous (top) and heterozygous (bottom) *Zfp462* mutant ESCs compared with WT ESCs (*n* = 3

replicates). Indicated are the numbers of significantly up- or downregulated genes. (adjusted *P* value ≤0.05; LFC ≥0.5). **d**, Heat map shows cluster analysis of differentially expressed genes (adjusted *P* value ≤0.05; LFC ≥1) in WT and *Zfp462* mutant ESCs. Top GO terms and corresponding significance are indicated for each cluster (left). piRNA, Piwi-interacting RNA; DNAme, DNA methylation. **e**, Genomic screen shots of *Sox17* and *Gata6* show mRNA expression levels in WT and *Zfp462* mutant ESCs. All RNA-seq profiles are normalized for library size.

Data Fig. 3f). Several key developmental TFs were upregulated in both heterozygous and homozygous *Zfp462* mutant ESCs, including *Bmp4*, *Gata6*, *Gata4* and *Sox17*, required for cell fate specification towards meso- and endodermal lineages (Fig. 3c,e). *Gata6* and *Sox17* expression is restricted to primitive endoderm in early embryos, and overexpression of these TFs in ESCs promotes their differentiation into primitive/extra-embryonic endoderm-like cells[44–46] with similar phenotypes to heterozygous and homozygous mutant *Zfp462* ESCs[44]. Thus, *Zfp462* is required to maintain ESC pluripotency and *Zfp462* haploinsufficiency upregulates key developmental genes including TFs that promote meso-endodermal lineage specification.

ZFP462 represses endodermal genes in neural differentiation

On the basis of the link between *Zfp462/ZNF462* haploinsufficiency and defects in mouse and human neurodevelopment[32,33,47], we investigated the role of *Zfp462* in neural differentiation. To reduce variability before differentiation, we tested whether enforcing the naïve ground state of ESCs using inhibitors against MAPK and GSK3 (termed 2i) could restore self-renewal of *Zfp462* mutant ESCs. Culturing *Zfp462* mutant ESCs under 2i conditions (2i/S/L) suppressed their spontaneous

differentiation (Fig. 4a and Extended Data Fig. 4a,b) and ESCs grown in 2i/S/L medium displayed lower ZFP462 expression (Extended Data Fig. 4c,d). We used an established protocol to induce neuronal differentiation starting from naïve ground-state ESCs[48]. Sequential withdrawal of 2i and leukaemia inhibitory factor (LIF) initiated ESC differentiation into embryoid bodies (EBs) containing progenitors of ecto-, meso- and endodermal lineages. Subsequent addition of retinoic acid (RA) stimulated ectoderm expansion followed by enrichment of NPCs. ZFP462 expression levels increased during neural differentiation with highest levels in NPCs, consistent with previous reports (Extended Data Fig. 4c,d)[47]. Importantly, heterozygous and homozygous *Zfp462* mutants gave rise to EBs and NPCs with reduced size compared with wild-type ESCs (Fig. 4a and Extended Data Fig. 4b). Quantitative reverse transcription polymerase chain reaction (RT–qPCR) analysis revealed that *Nanog* and *Oct4* were downregulated and *Pax6* and *Ngn2* were upregulated, indicating successful exit from pluripotency and induction of neural differentiation in wild-type and mutant cells (Fig. 4b). However, endodermal markers *Gata6* and *FoxA2* were strongly induced in *Zfp462*[−/−] ESCs upon 2i withdrawal and remained high even in NPCs (Fig. 4b). Although heterozygous *Zfp462* mutant cells showed similar

marker gene expression levels to wild-type cells at the onset of differentiation, they failed to downregulate endoderm-specific genes in NPCs similar to knockout (KO) cells.

Next, we profiled global gene expression changes throughout the differentiation time course (Fig. 4c). Cluster analysis of differentially regulated genes yielded three separate groups based on distinct expression kinetics. Genes in cluster 1 comprising ESC-specific markers were highly expressed in ESCs under 2i/S/L and S/L conditions and were downregulated in EBs and NPCs. Cluster 2 contained ~1,300 genes including meso- and endodermal lineage markers. These genes were induced in EBs and gradually decreased in NPCs derived from wild-type ESCs. In contrast, in heterozygous and homozygous *Zfp462* mutant cells, cluster 2 genes were highly induced in EBs and remained overexpressed in NPCs (Fig. 4c,d and Extended Data Fig. 4e). Addition of RA induced upregulation of cluster 3 genes including key neural TFs. Compared with wild-type cells, upregulation of neural genes was lower in *Zfp462* mutant NPCs suggesting a delay in neural differentiation and/or maturation (Fig. 4c,d and Extended Data Fig. 4e). Although deregulation of meso- and endodermal genes was less severe, neural gene expression was similarly impaired in heterozygous and homozygous mutant cells. This is consistent with significant neurodevelopmental defects seen in patients with *ZNF462* haploinsufficiency[32,33]. Finally, immunohistochemistry in NPCs revealed that the deregulation of genes extended to the protein level (Fig. 4e,f). Most wild-type NPCs were SOX1-positive, whereas FOXA2 expression was barely detected, reflecting efficient neural differentiation[49]. In contrast, enrichment and distribution of FOXA2-positive cells was strongly increased in *Zfp462*⁻/⁻ NPCs. Thus, *Zfp462* deletion results in misspecification towards meso-endodermal lineages under conditions that normally induce neural cell identity. We infer that ZFP462 is required to silence expression of lineage non-specific genes and promotes specification into neural cell lineages, consistent with developmental delay and neural defects of heterozygous *Zfp462* KO mice[47].

ZFP462 mediates sequence-specific G9A/GLP recruitment

To investigate how ZFP462 contributes to lineage specification, we used chromatin immunoprecipitation followed by sequencing (ChIP–seq) to profile binding of Avi-tagged ZFP462 genome-wide, which revealed 16,264 significantly enriched sites (Fig. 5a and Extended Data Fig. 5a). Most ZFP462 peaks (86%) were found in intergenic and intronic regions, distal to transcription start sites. As the related family of Krüppel-associated box (KRAB) domain-containing zinc finger proteins (KZFP) target heterochromatin modifiers to silence transposable elements (TE)[50], we analysed ZFP462 binding at repeat DNA families and uncovered a large overlap (Fig. 5b). Among the various repeat DNA classes, ZFP462 binding was strongly enriched at mouse-specific long terminal repeat (LTR) families of endogenous retroviruses (58.11%). Specifically, ZFP462 frequently occupied RLTR9a, RLTR9d, RLTR9d2, RLTR9a2 and RLTR9e families, suggesting a role in repressing these TE subfamilies via G9A/GLP recruitment (Fig. 5c).

To determine its role in targeting the G9A/GLP complex to chromatin, we used CRISPR-mediated deletion to remove ZFP462 in Avi-GLP expressing ESCs and profiled GLP, WIZ and H3K9me2 (Extended Data

Fig. 5b–d). In wild-type ESCs, ZFP462 targets were co-occupied by GLP and WIZ, similar to REST binding sites, and displayed modest levels of H3K9me2 (Fig. 5d–f). HMTase complex subunits and histone methylation were markedly reduced at ZFP462 targets in ESCs lacking ZFP462. By comparison, REST TF binding sites displayed increased GLP and WIZ ChIP–seq signals upon *Zfp462* deletion. As global G9A levels remained unchanged in the mutant (Extended Data Fig. 4c), we conclude that the gain at REST binding sites reflects genomic redistribution of the H3K9-specific HMTase complex in the absence of ZFP462 interaction. Next, we used CRISPR to delete ZFP462 in Avi-HP1γ expressing ESCs and tested its role in HP1γ chromatin targeting (Extended Data Fig. 5c). Compared with known binding sites such as ADNP targets[51], Avi-HP1γ showed only modest enrichment at ZFP462 peaks (Extended Data Fig. 5e). Moreover, the low HP1γ signal remained largely unchanged by *Zfp462* deletion arguing against a primary role of ZFP462 in directly recruiting HP1γ to its target loci. Thus, ZFP462 is primarily required for sequence-specific targeting of the G9A/GLP complex to induce heterochromatin modifications at its targets.

G9A/GLP restricts DNA accessibility at ZFP462 peaks

Transcriptional silencing by G9A/GLP is associated with chromatin compaction[5]. We investigated whether failure to target G9A/GLP and H3K9me2 leads to increased DNA accessibility in *Zfp462* KO ESCs using assay for transposase-accessible chromatin using sequencing (ATAC-seq). Despite G9A/GLP occupancy and H3K9me2, most ZFP462 peaks were already accessible in wild-type ESCs (Fig. 6a,b and Extended Data Fig. 6a). Nonetheless, DNA accessibility was further increased at target loci upon loss of G9A/GLP and H3K9me2 in *Zfp462* KO ESCs. In contrast, REST binding sites remained unchanged (Fig. 6a–c).

To examine whether G9A/GLP-dependent chromatin modifications directly contribute to restricting DNA accessibility at ZFP462 peaks, we analysed available ATAC-seq data of *G9a*/*Glp* double KO (*G9a*/*Glp* dKO) ESCs[52]. As expected, G9A/GLP loss led to increased DNA accessibility at REST targets. DNA accessibility was also increased at ZFP462 peaks, suggesting that G9A/GLP also acts as transcriptional co-repressor at these targets (Fig. 6d,e).

To determine whether G9A/GLP regulation of DNA accessibility is required for ZFP462-dependent silencing of lineage non-specific genes, we compared differential gene expression between *G9a*/*Glp* dKO ESCs and *Zfp462* KO ESCs, which revealed an overall positive, but moderate, correlation (Extended Data Fig. 6b). We speculate that the discrepancy is likely because in ESCs G9A/GLP interacts with multiple TFs including REST, enforcing repression of diverse sets of target genes.

To address whether ZFP462-mediated recruitment of the G9A/GLP complex is required for meso-endodermal gene silencing, we used CRISPR–Cas9 to introduce sequences encoding either full-length ZFP462 (ZFP462-FL) or CT-truncated ZFP462 (ZFP462-NT+Mid) at the endogenous locus in *Zfp462* KO ESCs, isolating two independent ESC clones for each *Zfp462* rescue construct (Extended Data Fig. 6c–e). Unlike ZFP462-NT+Mid, expression of ZFP462-FL rescued ESC morphology (Extended Data Fig. 6f). Furthermore, ATAC-seq analysis showed that ZFP462-FL, but not ZFP462-NT+Mid, mostly restored wild-type levels of DNA accessibility at ZFP462 peaks (Extended Data Fig. 6g).

Fig. 4 | *Zfp462* mutant cells show abnormal cell fate specification during neuronal differentiation. a, Scheme shows design of neural differentiation experiment from ESCs to NPCs (top). Stepwise withdrawal of 2i inhibitors (2i) and LIF leads to formation of cellular aggregates called EBs. Subsequent treatment with RA induces enrichment of NPCs. Representative bright-field images show WT and *Zfp462* mutant cells at corresponding stages of neural differentiation. Scale bar, 100 μm. **b**, Line plots shows RT–qPCR analysis of lineage marker expression during neural differentiation (*n* = 2 replicates). Expression levels are shown relative to ESCs (2i/S/L). **c**, Heat map shows cluster analysis of differentially expressed genes (adjusted *P* value ≤0.05; LFC ≥1) in WT and *Zfp462* mutant cells during neural differentiation (*n* = 2 replicates). Selected

lineage marker genes are indicated for each cluster (left). **d**, Heat maps show differential expression of selected marker genes specific for endodermal and neural lineages in heterozygous and homozygous *Zfp462* mutant cells during neural differentiation. Prim. and def. (EN), Primitive and definitive endoderm. **e**, Immunohistochemistry analysis shows SOX1 and FOXA2 expression in WT and homozygous *Zfp462* mutant at day 8 of neural differentiation. Cell aggregates are counterstained with DAPI. Scale bar, 50 μm. **f**, Bar plots show quantification of SOX1 and FOXA2 immunofluorescence in WT (*n* = 9 cell aggregates over two independent experiments) and *Zfp462*⁻/⁻ cell (*n* = 12 cell aggregates over three independent experiments). Data are presented as mean ± standard deviation. *P* values derived from a two-tailed *t*-test are indicated.

Finally, RT−qPCR showed that ZFP462-FL but not ZFP462-NT+Mid restored silencing in *Zfp462* KO ESCs (Fig. 6f). ZFP462-FL expression also restored endodermal marker gene silencing during NPC differentiation, whereas *Gata6* and *FoxA2* expression remained high upon rescue with ZFP462-NT+Mid (Extended Data Fig. 6h). We infer that ZFP462 recruits G9A/GLP to restrict DNA accessibility and enforce

silencing of lineage non-specific genes, potentially by affecting TF occupancy at *cis*-regulatory sequences.

ZFP462-dependent heterochromatin restricts TF binding

TEs contribute up to 25% of binding sites for pluripotency TFs and play an integral role in the core regulatory network of ESCs[53,54]. Indeed, the
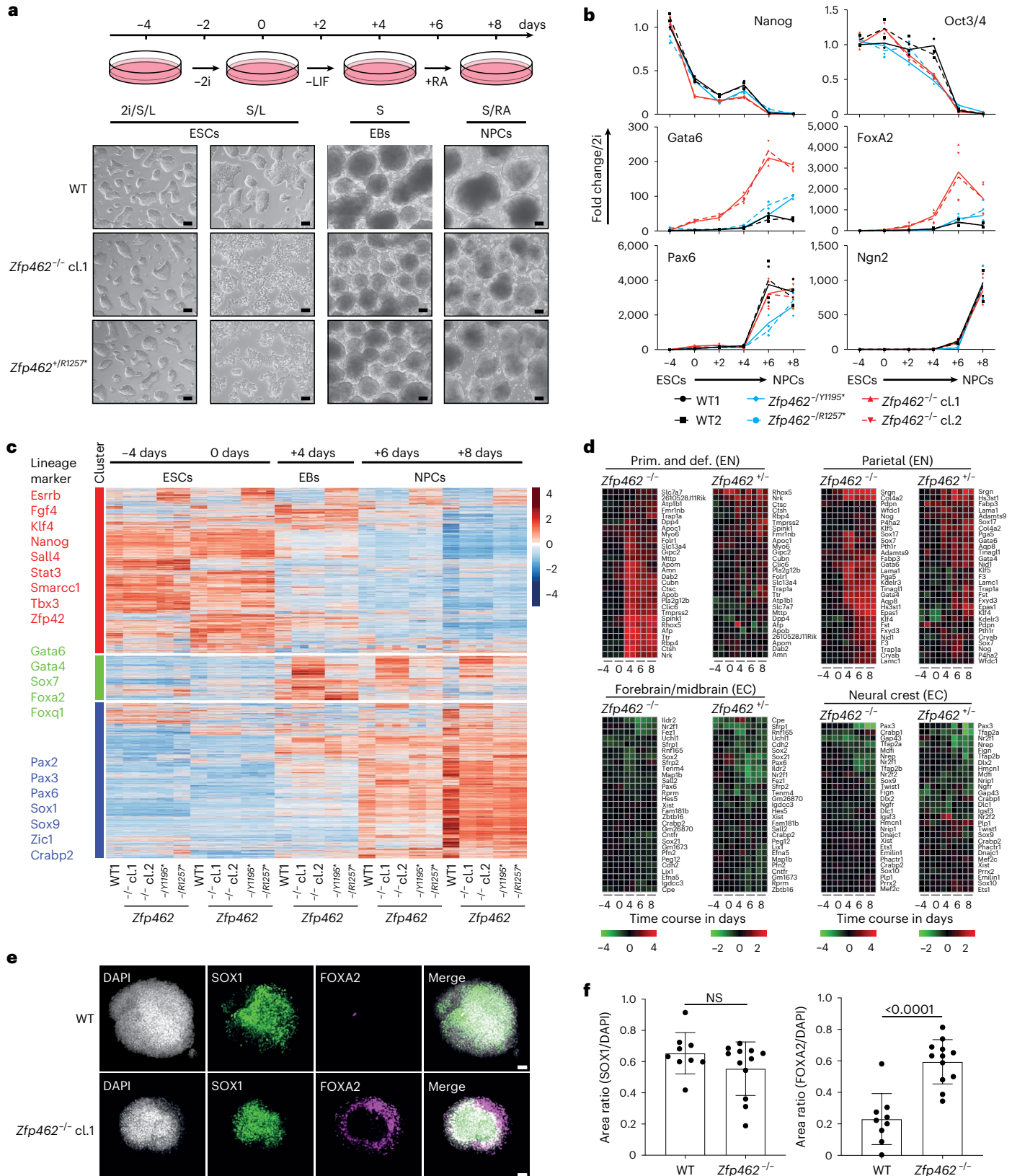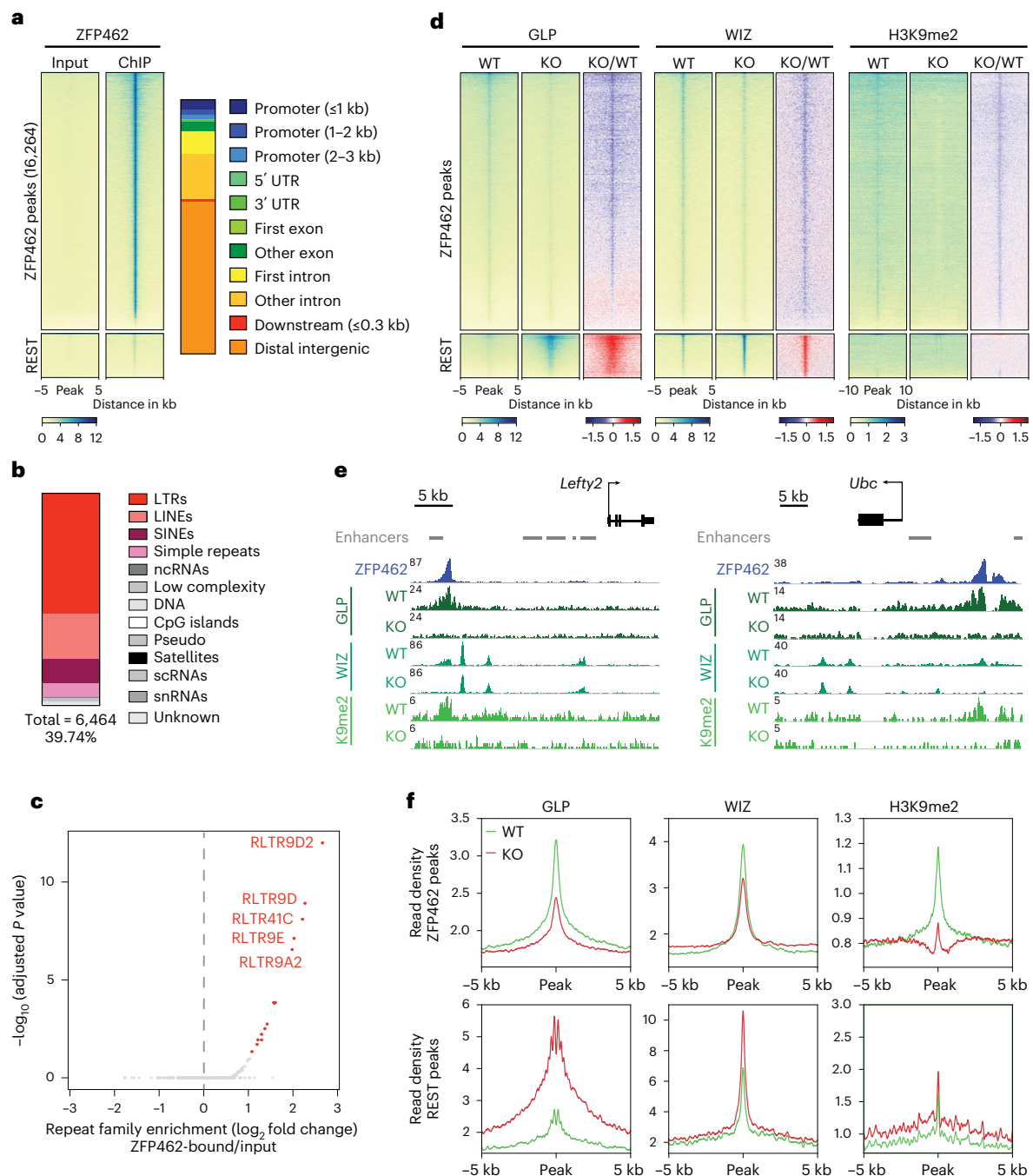
Fig. 5 | ZFP462 establishes H3K9me2 containing heterochromatin by recruiting GLP and WIZ proteins. a, Heat maps of ZFP462 ChIP–seq enrichment at significant ZFP462 peaks (16,264) (top cluster) and REST peaks (bottom cluster) in ESCs. Each row represents a 10 kb window centred on peak midpoint, sorted by ZFP462 ChIP signal. Input signals for the same windows are shown on the left. Below are scale bars (*n* indicates average distribution of two replicates). Bar plot, on the right, shows percentage of genomic features overlapping with ZFP462 peaks. **b**, Bar plot shows fraction of repetitive DNA types including short and long interspersed retrotransposable elements (SINEs, LINEs), non-coding RNAs (ncRNAs), small conditional RNAs (scRNAs), small nuclear RNAs (snRNAs) overlapping with ZFP462 peaks. A total of 39.74% of ZFP462 peaks are associated with repeat elements. **c**, Volcano plot shows enrichment and corresponding significance of TE families overlapping with ZFP462 peaks. **d**, Heat maps of GLP,

WIZ and H3K9me2 ChIP–seq enrichment at ZFP462 and REST peaks in WT and *Zfp462* KO ESCs. Blue-to-red-scaled heat map represents corresponding ChIP–seq enrichment ratio between *Zfp462* KO versus WT (KO/WT). All heat maps are sorted by GLP enrichment signal in WT. For GLP and WIZ, each row represents a 10 kb window centred on ZFP462 peak midpoints. For H3K9me2, each row represents a 20 kb window (*n* indicates average distribution of two replicates). **e**, Screen shots of two selected genomic regions with ZFP462 peaks display GLP, WIZ and H3K9me2 ChIP–seq signals in WT and *Zfp462* KO ESCs. Grey bars indicate ENCODE enhancer annotations. ChIP–seq profiles are normalized for library size. **f**, Metaplots show average GLP, WIZ and H3K9me2 signal at ZFP462 and REST peaks in WT and *Zfp462* KO ESCs. For each plot, read density is plotted at 10 kb window centred on peak midpoints.

mouse-specific RLTR9 subfamilies contain sequence motifs for pluripotency TFs including OCT4 and SOX2 and promote gene expression in ESCs[55–57]. Intriguingly, consensus sequences for several pluripotency

TFs were enriched at ZFP462 peaks (Fig. 6g). ChIP–seq profiling of OCT4, SOX2, NANOG, NR5A2 and ESRRB confirmed robust binding at ZFP462 peaks overlapping with H3K9me2 in wild-type ESCs (Extended

Data Fig. 7a). Notably, ZFP462 also co-localized with OCT4, SOX2 and NANOG at upstream regulatory sequences of the *Oct3/4* gene, suggesting that it limits gene expression by targeting G9A/GLP (Extended Data Fig. 7b).

The precise regulation of pluripotency TFs is essential to maintain ESC identity[58]; a less than two-fold increase in OCT4 expression is sufficient to trigger spontaneous differentiation of ESCs into primitive endoderm and mesoderm[36,59–62]. To investigate whether heterochromatin mediated by G9A/GLP affects TF interactions, we analysed TF binding changes at ZFP462 targets (Fig. 6h and Extended Data Fig. 7a). OCT4, SOX2 and NR5A2 displayed increased enrichment at ZFP462 targets in KO ESCs, concomitant with loss of H3K9me2. Notably, binding of these TFs was unaffected at OSN peaks that do not overlap with ZFP462, indicating specificity. Consistent with spontaneous differentiation into primitive endoderm, NANOG and ESRRB were downregulated and had reduced occupancy at OSN peaks (Fig. 6h and Extended Data Fig. 7a). In contrast, NANOG and ESRRB binding remained higher at ZFP462 peaks in KO ESCs, consistent with a relative increase in DNA accessibility. OCT4 binding was also increased at ZFP462 target site upstream of the *Oct3/4* gene (Extended Data Fig. 7b). We infer that ZFP462-dependent recruitment of G9A/GLP restricts TF binding at its target sites.

ZFP462 targets heterochromatin to repress enhancers in ESCs
To examine the role of ZFP462 at *cis*-regulatory sequences, we used the ChromHMM annotation of ESC chromatin states[63,64]. Nearly half of ZFP462 peaks overlapped with genomic regions categorized as active enhancers (44.6%, Fig. 7a). ZFP462-bound TEs represented 35.37% of ESC enhancers (Extended Data Fig. 7c). In contrast, less than 13% of ZFP462 peaks were associated 'repressed chromatin' and 'constitutive heterochromatin' categories. ChIP–seq profiling of euchromatic histone modifications, H3K4me1, H3K4me3 and H3K27ac, confirmed that many ZFP462 peaks localized at active enhancers in ESCs (Fig. 7b). Notably, G9A/GLP levels were comparable at all ZFP462 peaks, but H3K9me2 abundance was low at active ESC enhancers and accumulated mostly at ZFP462 peaks devoid of euchromatic chromatin modifications. Loss of G9A/GLP and H3K9me2 in *Zfp462* KO ESCs led to increased H3K27ac at active ESC enhancers and at a subset of intergenic ZFP462 peaks marked primarily by H3K4me1, potentially revealing aberrant activation of silent enhancers (Fig. 4b).

To determine whether ZFP462-dependent heterochromatin is required to repress mesoderm- and endoderm-specific enhancers in ESCs, we identified ZFP462 peaks overlapping with enhancers in endoderm (EN), mesoderm (ME) and ectoderm (EC). We used ChomHMM annotation based on epigenomic profiling of mouse foetal tissues of E14 (ref. 65) and identified 4,272 ZFP462 peaks (26.3%) overlapping with enhancers that are active in at least one of these three primary germ layers. From this set we selected enhancers specific for one germ layer and inactive in ESCs to analyse H3K9me2 and DNA accessibility changes in

wild-type and *Zfp462* KO ESCs (Fig. 7c). EN- and ME-specific enhancers displayed ZFP462-dependent H3K9me2 enrichment and reduced DNA accessibility in ESCs (Fig. 7d), in contrast with EC-specific enhancers. To determine whether loss of ZFP462-dependent heterochromatin led to increased TF binding at EN- and ME-specific enhancers, we first performed a motif analysis to identify the consensus sequences with the highest enrichment at ZFP462 peaks overlapping ME/EN-specific enhancers. We uncovered NR5A2 and ESRRB motifs (Extended Data Fig. 7d), whereas OCT4-SOX2-NANOG motifs were not enriched. Our TF ChIP–seq data showed that NR5A2 occupancy was markedly increased at ME/EN-specific enhancers, but unchanged at EC-specific enhancers, in *Zfp462* KO relative to wild-type ESCs (Extended Data Fig. 7e). Consistent with its global downregulation at ZFP462 peaks and OSN targets, ESRRB binding decreased at all germ-layer-specific enhancers. We conclude that inactivation of a set of ME/EN-specific enhancers in ESCs relies on ZFP462-dependent heterochromatin to limit DNA accessibility and aberrant TF binding.

Finally, we examined whether elevated TF binding and aberrant activation of EN- and ME-specific enhancers at ZFP462 peaks impacts transcriptional regulation. Genes proximal to ZFP462 peaks overlapping with activated EN- and ME-specific enhancers were more frequently upregulated, arguing that enhanced DNA accessibility and TF binding promotes aberrant gene expression (Extended Data Fig. 7f). To test whether proximal gene regulation is linked to spontaneous meso-endodermal differentiation, we performed GO term analysis. Notably, the most enriched GO terms were related to developmental processes, in particular cardiovascular system development (Fig. 7e). These results suggest that ZFP462 binds to enhancers and controls their regulatory activity in ESCs, particularly preventing unscheduled activation of EN- and ME-specific enhancers. In the absence of ZFP462, these enhancers have increased DNA accessibility and TF binding, leading to aberrant expression of proximal genes involved in meso-endoderm differentiation.

## ZFP462 represses ME/EN enhancers in neuronal cells
Our data argue that ZFP462 is required for normal neuro-differentiation and development by regulating enhancers in neural cell lineages. To determine whether ZFP462 occupies enhancers later in development, we differentiated Avi-ZFP462-expressing ESCs and mapped its genome-wide distribution in NPCs. ZFP462 peaks in ESCs were largely conserved in NPCs (Fig. 7f). To investigate DNA accessibility at ZFP462 targets in neural cells, we performed ATAC-seq in neural stem cells (NSCs) isolated from the subventricular zone of adult mice. ZFP462 was abundant in NSCs (Extended Data Fig. 7g), and we found that ZFP462-bound enhancers were mostly inaccessible in wild-type NSCs by ATAC-seq (Fig. 7g). These data suggest that G9A/GLP-dependent heterochromatin efficiently restricts inappropriate TF binding in NSCs (Fig. 7g).

Fig. 6 | ZFP462-targeted heterochromatin restricts DNA accessibility and TF binding. a, Heat maps of GLP, WIZ and H3K9me2 ChIP–seq enrichment ratios between *Zfp462* KO versus WT (KO/WT) ESCs at ZFP462 and REST peaks. On the right, heat maps of ATAC-seq signal at ZFP462 and REST peaks in WT and *Zfp462* KO ESCs. Blue-to-red-scaled heat map represent corresponding enrichment ratios between *Zfp462* KO versus WT (KO/WT) ESCs. GLP, WIZ and ATAC-seq heat maps represent a 10 kb window, while H3K9me2 heat map represents a 20 kb window centred on peak midpoints, sorted by H3K9me2 KO/WT enrichment ratio (*n* indicates average distribution of two ChIP–seq replicates and average distribution of three ATAC-seq replicates). **b**, Metaplots show average profiles of ATAC-seq signal in WT (green) and *Zfp462* KO (red) ESCs at ZFP462 peaks (left) and REST peaks (right). **c**, Volcano plot shows DNA accessibility changes between WT and *Zfp462* KO ESCs. *X* axis represents fold change in accessibility and corresponding significance on *Y* axis. Differentially accessible sites bound by ZFP462 are highlighted in red. Indicated in grey is the number of loci with increased DNA accessibility and in red the number of ZFP462-bound loci with

increased DNA accessibility with significance <0.05. **d**, Heat maps of ATAC-seq signal at ZFP462 and REST peaks in WT and *G9a/Glp* dKO ESCs. Blue-to-red-scaled heat map represent corresponding enrichment ratios between *G9a/Glp* dKO versus WT (KO/WT) ESCs. Each row represents a 10 kb window centred on peak midpoints, sorted by dKO/WT enrichment ratio (*n* indicates average distribution of two ATAC-seq replicates). **e**, Metaplots show average profiles of ATAC-seq signal in WT (green) and *G9a/Glp* dKO (red) ESCs at ZFP462 peaks (top) and REST peaks (bottom). **f**, RT–qPCR expression analysis of meso-endodermal marker genes in WT, *Zfp462* KO and *Zfp462* KO ESCs expressing ZFP462-FL or ZFP462-NT+Mid (*n* = 2 independent biological replicates). **g**, HOMER analysis of known DNA sequence motifs enriched at significant ZFP462 peaks. Top-ranked DNA sequence motifs and respective significance values are shown in the table. **h**, Metaplots show average pluripotency TF ChIP–seq signal in WT and *Zfp462* KO ESCs at ZFP462 peaks with LFC ≥1 in KO/WT H3K9me2 ChIP signal loss and at shared OCT4-SOX2-NANOG (OSN) peaks that are not bound by ZFP462 (*n* indicates average distribution of two replicates).

Finally, we interrogated public ATAC-seq data to compare DNA accessibility in meso-endoderm progenitors that lack ZFP462 expression (Extended Data Fig. 3a)[66]. In contrast to NSCs, meso-endoderm

progenitors displayed substantial DNA accessibility at ZFP462 targets, indicative of lineage-specific enhancer activity (Fig. 7g). We infer that ZFP462 controls meso-endoderm enhancers by recruiting G9A/
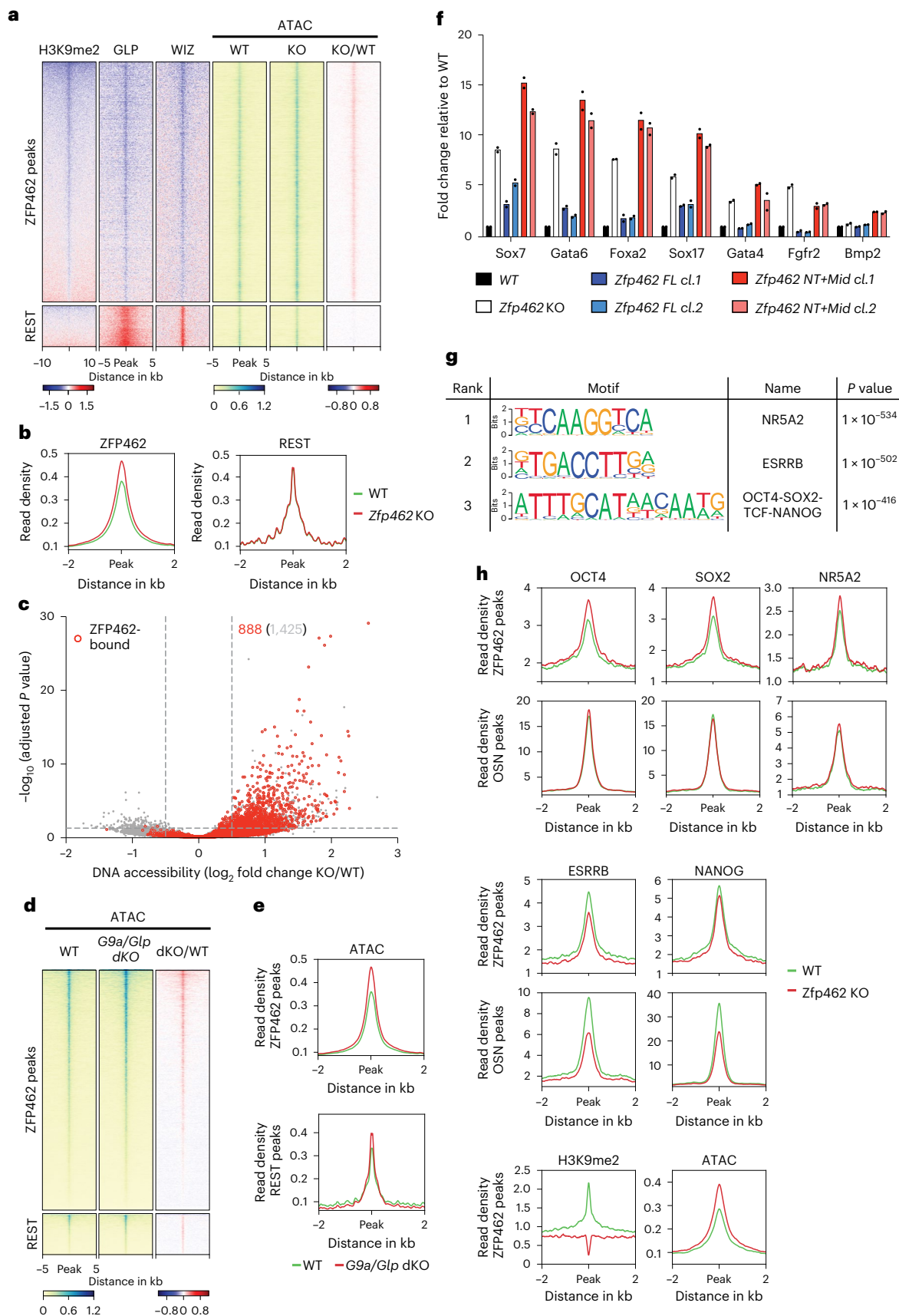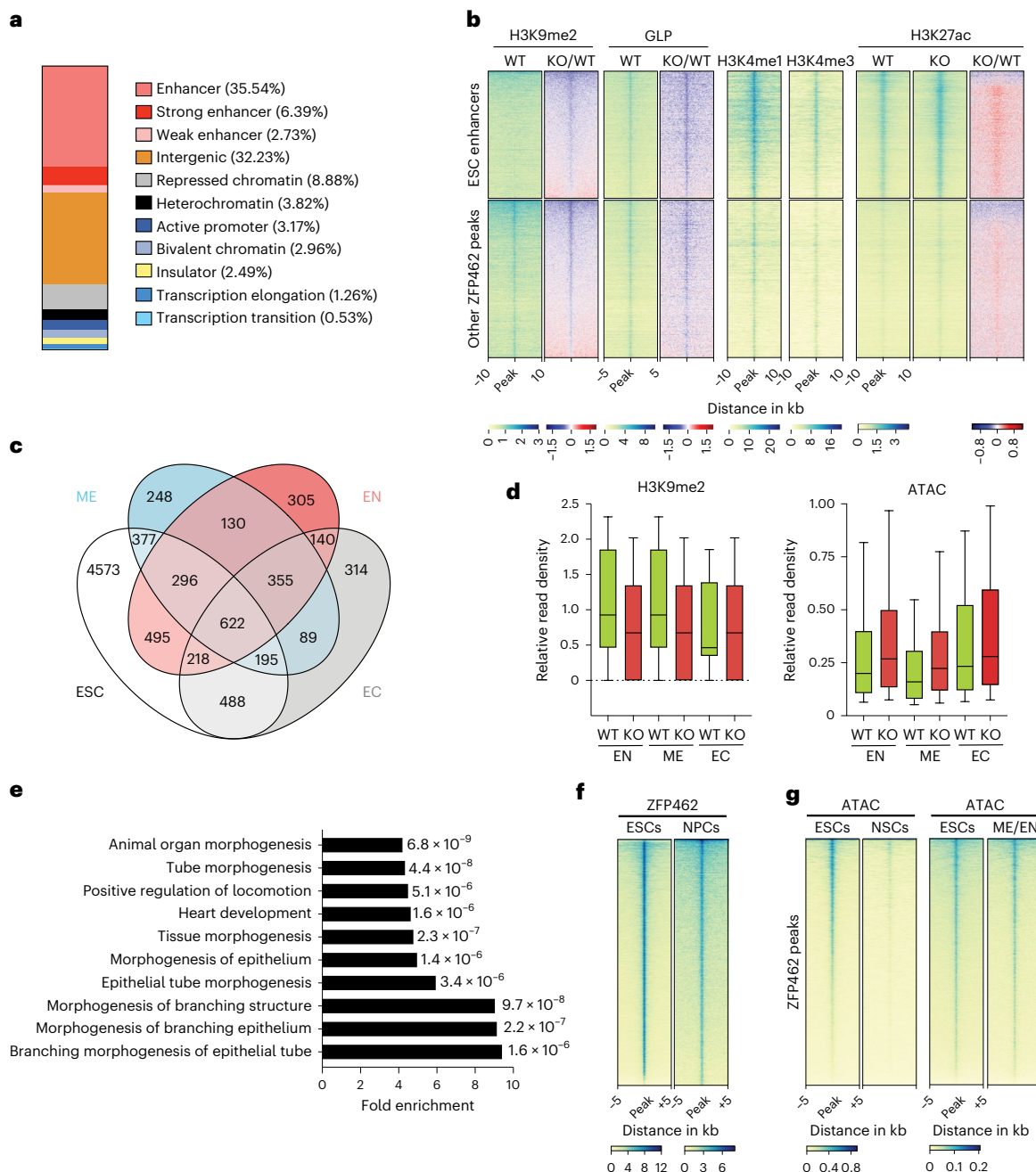
Fig. 7 | ZFP462 represses meso-endodermal enhancers in ESCs. a, Bar plot shows percentage of ZFP462 ChIP–seq peaks overlapping with ChromHMM states in ESCs. **b**, Heat maps of GLP, H3K9me2, H3K4me1, H3K4me3 and H3K27ac ChIP–seq signals at ZFP462 peaks separated into two clusters: ESC enhancers and other ZFP462 peaks. Blue-to-red-scaled heat map represents corresponding enrichment ratios between *Zfp462* KO versus WT (KO/WT). GLP heat map represents 10 kb and histone modifications heat maps represent a 20 kb window centred on peak midpoints, sorted by H3K9me2 KO/WT enrichment ratio (*n* indicates average distribution of two ChIP–seq replicates). **c**, Venn diagram shows overlap of ZFP462 peaks with ChromHMM-annotated enhancers in ESC, endoderm (EN), mesoderm (ME) and ectoderm (EC). **d**, Box plots show enrichment of H3K9me2 ChIP signal and ATAC-seq signal in WT and *Zfp462* KO ESCs at ZFP462 peaks overlapping with EN-, ME- and EC-specific enhancers.

n = 305 (EN), *n* = 248 (ME), *n* = 314 (EC). Shown are median (horizontal line), 25th to 75th percentiles (boxes) and 90% percentiles (whiskers). Outliers are not shown. **e**, Bar plot shows GO term analysis of biological processes of genes with significant differential expression (KO/WT, LFC ≥1 and adjusted *P* value ≤0.05) located proximal to ZFP462 peaks annotated as meso-endoderm-specific enhancers[77]. **f**, Heat map of ZFP462 ChIP–seq enrichment at ZFP462 peaks in ESCs and NPCs. ChIP–seq rows represent 10 kb window centred on ZFP462 peak midpoints, sorted by ZFP462 ChIP–seq signal intensity (*n* indicates average distribution of two replicates). **g**, Heat maps of ATAC-seq signal at ZFP462 peaks in ESCs and NSCs (left). Heat maps of ATAC-seq signal at ZFP462 peaks in ESCs and meso-endoderm (ME/EN) cells[66]. ATAC-seq rows represent 10 kb window centred on ZFP462 peak midpoints, sorted by ESCs ATAC-seq signal intensity.

GLP and establishing heterochromatin in ESCs and cells of the neural lineage. Our data suggest that ZFP462 promotes robust differentiation along the neural lineage by preventing lineage non-specific gene expression.

Discussion

Our CRISPR genetic screen for regulators of heritable H3K9me2/3-dependent gene silencing in mouse ESCs uncovered *Zfp462*, an essential gene encoding a C_2H_2 zinc finger TF, previously

of unknown function. Heterozygous mutations of the human orthologue, *ZNF462*, cause the neurodevelopmental disorder Weiss–Kruszka syndrome[33], but the underlying molecular mechanism is unknown. We demonstrate that ZFP462 regulates cell identity by preventing aberrant expression of endoderm-specific genes in ESCs and during neural differentiation.

ZFP462 binds to enhancer sequences that contain pluripotency and meso-endodermal TF binding motifs and represses these loci by recruiting G9A/GLP, which catalyses H3K9me2. Establishing heterochromatin at ZFP462 target sites restricts DNA accessibility and thereby controls TF binding and enhancer activity. Failure to target G9A/GLP-dependent heterochromatin in heterozygous and homozygous *Zfp462* deletion mutants results in mis-expression of meso-endoderm-specifying TF, which impairs self-renewal and interferes with timely establishment of neural-specific gene expression during NPC differentiation. Neural differentiation is similarly impaired in heterozygous and homozygous Zfp462 mutant cells, consistent with the neurodevelopmental defects observed in patients with *ZNF462* haploinsufficiency.

ZFP462 maintains closed chromatin at a subset of ME/EN-specific enhancers, blocking NR5A2 binding which otherwise may promote meso-endoderm cell fates[67–69]. In contrast, most ZFP462 targets are co-occupied by pluripotency TFs and show hallmarks of enhancer activity in ESCs despite the presence of G9A/GLP-dependent heterochromatin modifications. The capacity of OCT4 and SOX2 to engage with sequence motifs in the context of heterochromatin is a unique feature of pioneer TFs and a critical first step in induced pluripotent stem cell reprogramming[70,71]. Nevertheless, OCT4 and SOX2 binding is limited by positioning the binding motif on the nucleosome surface and chromatin accessibility[72,73]. For example, to engage with its target sites in closed chromatin, OCT4 requires the chromatin remodelling factor Brg1 to mobilize nucleosomes[74]. Here we show that ZFP462-dependent heterochromatin also contributes to limiting DNA accessibility for OCT4 binding at enhancers preventing aberrant gene expression. Precise levels of OCT4 are critical for ESC self-renewal and govern cell type specification during early embryogenesis[36]. Thus, by modulating OCT4 binding at ESC-specific enhancers, ZF462 may contribute to the quantitative readout of OCT4 and define one of the earliest steps of cell fate specification during embryogenesis.

Although ZFP462 is conserved across vertebrates, it predominantly bound to evolutionarily young, species-specific TEs[55]. It is possible that ZFP462-bound TEs contain DNA sequence motifs that are much 'older' that the rest of the TE sequence. Indeed, mouse-specific RLTR9 subfamilies harbour conserved binding sites for key pluripotency TFs including OCT4 and SOX2 (refs. [55,56]). Adaptation of pluripotency TF binding sites bestows TEs with enhancer activity specific for germ cells and preimplantation embryos optimizing the likelihood to persist by vertical transmission[54,75,76]. We surmise that ZFP462/ZNF462 emerged in vertebrates to repress TEs by recognizing a DNA sequence that overlaps with pluripotency TF binding motifs. Therefore, TE-derived enhancers were co-opted to contribute not only to the ESC-gene regulatory network via pluripotency TFs, but also to cell specification via ZFP462.

On the basis of our findings, we propose that the human orthologue ZNF462 regulates G9A/GLP and targets non-neural genes for repression during neurogenesis. Failure to silence lineage non-specific genes due to a ZNF462 heterozygous LOF mutation might impair neural cell differentiation and underlie neurodevelopmental pathology. However, given the species specificity of ZFP462-regulated TEs in mouse, we predict that ZNF462 will have different human targets controlling a distinct gene regulatory network compared to mice. Nevertheless, our findings predict that *ZNF462*-linked pathology involves loss of heterochromatin silencing at enhancers and interference in neural differentiation by lineage non-specific gene expression. Our results argue that Weiss–Kruszka syndrome arises from deregulation of neural lineage-specific gene expression early in embryogenesis, consistent with the broad defects affecting cells of neuronal and neural crest origin.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-022-01051-2.

References

1. Eckersley-Maslin, M. A., Alda-Catalinas, C. & Reik, W. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19**, 436–450 (2018).
2. Burton, A. & Torres-Padilla, M.-E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–735 (2014).
3. Gökbuget, D. & Blelloch, R. Epigenetic control of transcriptional regulation in pluripotency and early differentiation. *Development* **146**, dev164772 (2019).
4. Iwafuchi-Doi, M. & Zaret, K. S. Cell fate control by pioneer transcription factors. *Development* **143**, 1833–1837 (2016).
5. Padeken, J., Methot, S. P. & Gasser, S. M. Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat. Rev. Mol. Cell Biol.* https://doi.org/10.1038/s41580-022-00483-w (2022).
6. Nishibuchi, G. & Nakayama, J. Biochemical and structural properties of heterochromatin protein 1: understanding its role in chromatin assembly. *J. Biochem.* **156**, 11–20 (2014).
7. Larson, A. G. & Narlikar, G. J. The role of phase separation in heterochromatin formation, function, and regulation. *Biochemistry* **57**, 2540–2548 (2018).
8. Kumar, A. & Kono, H. Heterochromatin protein 1 (HP1): interactions with itself and chromatin components. *Biophys. Rev.* **12**, 387–400 (2020).
9. Nicetto, D. & Zaret, K. S. Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Curr. Opin. Genet. Dev.* **55**, 1–10 (2019).
10. Chen, J. et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* **45**, 34–42 (2013).
11. Ragunathan, K., Jih, G. & Moazed, D. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348**, 1258699 (2015).
12. Audergon, P. N. C. B. et al. Restricted epigenetic inheritance of H3K9 methylation. *Science* **348**, 132 LP–132135 (2015).
13. Hathaway, N. A. et al. Dynamics and memory of heterochromatin in living cells. *Cell* **149**, 1447–1460 (2012).
14. Shinkai, Y. & Tachibana, M. H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev.* **25**, 781–788 (2011).
15. Tachibana, M. et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16**, 1779–1791 (2002).
16. Lienert, F. et al. Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet.* **7**, e1002090 (2011).
17. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* **41**, 246–250 (2009).
18. Liu, N. et al. Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability. *Genes Dev.* **29**, 379–393 (2015).
19. Zylicz, J. J. et al. G9a regulates temporal preimplantation developmental program and lineage segregation in blastocyst. *eLife* **7**, e33361 (2018).

20. Feldman, N. et al. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.* **8**, 188–194 (2006).

21. Kellner, S. & Kikyo, N. Transcriptional regulation of the Oct4 gene, a master gene for pluripotency. *Histol. Histopathol.* **25**, 405–412 (2010).

22. Epsztejn-Litman, S. et al. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15**, 1176–1183 (2008).

23. Benevento, M., van de Molengraft, M., van Westen, R., van Bokhoven, H. & Nadif Kasri, N. The role of chromatin repressive marks in cognition and disease: a focus on the repressive complex GLP/G9a. *Neurobiol. Learn. Mem.* **124**, 88–96 (2015).

24. Roopra, A., Qazi, R., Schoenike, B., Daley, T. J. & Morrison, J. F. Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. *Mol. Cell* **14**, 727–738 (2004).

25. Mozzetta, C. et al. The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol. Cell* **53**, 277–289 (2014).

26. Mulligan, P. et al. CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Mol. Cell* **32**, 718–726 (2008).

27. Schaefer, A. et al. Control of cognition and adaptive behavior by the GLP/G9a epigenetic suppressor complex. *Neuron* **64**, 678–691 (2009).

28. Deimling, S. J., Olsen, J. B. & Tropepe, V. The expanding role of the Ehmt2/G9a complex in neurodevelopment. *Neurogenesis* **4**, e1316888 (2017).

29. Kleefstra, T. et al. Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am. J. Hum. Genet.* **79**, 370–377 (2006).

30. Kramer, J. M. Regulation of cell differentiation and function by the euchromatin histone methyltranserfases G9a and GLP. *Biochem. Cell Biol.* **94**, 26–32 (2015).

31. Scheer, S. & Zaph, C. The lysine methyltransferase G9a in immune cell differentiation and function. *Front. Immunol.* **8**, 429 (2017).

32. Weiss, K. et al. Haploinsufficiency of ZNF462 is associated with craniofacial anomalies, corpus callosum dysgenesis, ptosis, and developmental delay. *Eur. J. Hum. Genet.* **25**, 946–951 (2017).

33. Kruszka, P. et al. Phenotype delineation of ZNF462 related syndrome. *Am. J. Med. Genet. A* **179**, 2075–2082 (2019).

34. González-Tarancón, R. et al. A novel mutation in the ZNF462 gene c.3306dup; p.(Gln1103Thrfs*10) is associated to Weiss–Kruszka syndrome. A case report. *Acta Clin. Belg.* https://doi.org/10.1080/17843286.2020.1780391 (2020).

35. Athanasiadou, R. et al. Targeting of de novo DNA methylation throughout the Oct-4 gene regulatory region in differentiating embryonic stem cells. *PLoS ONE* **5**, e9937 (2010).

36. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).

37. Michlits, G. et al. CRISPR-UMI: single-cell lineage tracing of pooled CRISPR–Cas9 screens. *Nat. Methods* **14**, 1191–1197 (2017).

38. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).

39. Rose, N. R. & Klose, R. J. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim. Biophys. Acta Gene Regul. Mech.* **1839**, 1362–1372 (2014).

40. Ren, W. et al. DNMT1 reads heterochromatic H4K20me3 to reinforce LINE-1 DNA methylation. *Nat. Commun.* **12**, 2490 (2021).

41. Ueda, J., Tachibana, M., Ikura, T. & Shinkai, Y. Zinc finger protein Wiz links G9a/GLP histone methyltransferases to the co-repressor molecule CtBP*. *J. Biol. Chem.* **281**, 20120–20128 (2006).

42. Olsen, J. B. et al. G9a and ZNF644 physically associate to suppress progenitor gene expression during neurogenesis. *Stem Cell Rep.* **7**, 454–470 (2016).

43. Argelaguet, R. et al. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**, 487–491 (2019).

44. McDonald, A. C. H., Biechele, S., Rossant, J. & Stanford, W. L. Sox17-mediated XEN cell conversion identifies dynamic networks controlling cell-fate decisions in embryo-derived stem cells. *Cell Rep.* **9**, 780–793 (2014).

45. Wamaitha, S. E. et al. Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes Dev.* **29**, 1239–1255 (2015).

46. Schrode, N., Saiz, N., Di Talia, S. & Hadjantonakis, A.-K. GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Dev. Cell* **29**, 454–467 (2014).

47. Wang, B. et al. Zfp462 deficiency causes anxiety-like behaviors with excessive self-grooming in mice. *Genes Brain Behav.* **16**, 296–307 (2017).

48. Bibel, M., Richter, J., Lacroix, E. & Barde, Y.-A. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat. Protoc.* **2**, 1034–1043 (2007).

49. Venere, M. et al. Sox1 marks an activated neural stem/progenitor cell in the hippocampus. *Development* **139**, 3938–3949 (2012).

50. Ecco, G., Imbeault, M. & Trono, D. KRAB zinc finger proteins. *Development* **144**, 2719–2729 (2017).

51. Ostapcuk, V. et al. Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* **557**, 739–743 (2018).

52. Jiang, Q. et al. G9a plays distinct roles in maintaining DNA methylation, retrotransposon silencing, and chromatin looping. *Cell Rep.* **33**, 108315 (2020).

53. Kunarso, G. et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* **42**, 631–634 (2010).

54. Sundaram, V. & Wysocka, J. Transposable elements as a potent source of diverse *cis*-regulatory sequences in mammalian genomes. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190347 (2020).

55. Sundaram, V. et al. Functional cis-regulatory modules encoded by mouse-specific endogenous retrovirus. *Nat. Commun.* **8**, 14550 (2017).

56. Todd, C. D., Deniz, Ö., Taylor, D. & Branco, M. R. Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells. *eLife* **8**, e44344 (2019).

57. Bourque, G. et al. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res.* **18**, 1752–1762 (2008).

58. Niwa, H. How is pluripotency determined and maintained? *Development* **134**, 635–646 (2007).

59. Thomson, M. et al. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**, 875–889 (2011).

60. Wang, Z., Oron, E., Nelson, B., Razis, S. & Ivanova, N. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* **10**, 440–454 (2012).

61. Le Bin, G. C. et al. Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development* **141**, 1001–1010 (2014).

62. Frum, T. et al. Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Dev. Cell* **25**, 610–622 (2013).

63. Pintacuda, G. et al. hnRNPK recruits PCGF3/5-PRC1 to the Xist RNA B-repeat to establish polycomb-mediated chromosomal silencing. *Mol. Cell* **68**, 955–969.e10 (2017).

64. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).

65. Gorkin, D. U. et al. An atlas of dynamic chromatin landscapes in mouse fetal development. *Nature* **583**, 744–751 (2020).

66. Cernilogar, F. M. et al. Pre-marked chromatin and transcription factor co-binding shape the pioneering activity of Foxa2. *Nucleic Acids Res.* **47**, 9069–9086 (2019).

67. Nissim, S. et al. Iterative use of nuclear receptor Nr5a2 regulates multiple stages of liver and pancreas development. *Dev. Biol.* **418**, 108–123 (2016).

68. Seitz, C. et al. The orphan nuclear receptor LRH-1/NR5a2 critically regulates T cell functions. *Sci. Adv.* **5**, eaav9732 (2022).

69. Olivieri, D. et al. Cooperation between HDAC3 and DAX1 mediates lineage restriction of embryonic stem cells. *EMBO J.* **40**, e106818 (2021).

70. Soufi, A. et al. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).

71. Morgunova, E. & Taipale, J. Structural insights into the interaction between transcription factors and the nucleosome. *Curr. Opin. Struct. Biol.* **71**, 171–179 (2021).

72. Michael, A. K. et al. Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science* **368**, 1460–1465 (2020).

73. Isbel, L., Grand, R. S. & Schübeler, D. Generating specificity in genome regulation through transcription factor sensitivity to chromatin. *Nat. Rev. Genet.* https://doi.org/10.1038/s41576-022-00512-6 (2022).

74. King, H. W. & Klose, R. J. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife* **6**, e22631 (2017).

75. Frank, J. A. & Feschotte, C. Co-option of endogenous viral sequences for host cell function. *Curr. Opin. Virol.* **25**, 81–89 (2017).

76. Bourque, G. Transposable elements in gene regulation and in the evolution of vertebrate genomes. *Curr. Opin. Genet. Dev.* **19**, 607–612 (2009).

77. Carbon, S. et al. AmiGO: online access to ontology and annotation data. *Bioinformatics* **25**, 288–289 (2009).

Methods

The experiments in this manuscript are in compliance with relevant guidelines and ethical regulations.

Cell culture

Mouse ESCs (TC1/RRID:CVCL_M350) were cultured without feeders on gelatin-coated dishes in Dulbecco's modified Eagle medium supplemented with 15% foetal bovine serum (F7524/Sigma), 1× non-essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), 2 mM L-glutamine (Gibco), 0.1 mM 2-mercaptoethanol (Sigma), 50 mg ml^{-1} penicillin, 80 mg ml^{-1} streptomycin and homemade LIF, at 37 °C in 5% CO_2. Cells were regularly tested for mycoplasma. NSCs were isolated from subventricular zone of 7–8-week-old mouse brain and cultured as previously described[78].

Generation of TetO *Oct3/4* reporter ESCs and delivery of TetR-FLAG-fusion constructs

TetO *Oct3/4* reporter ESCs were derived from the original *Oct4* reporter (CiA:Oct4) ESCs[13] by excising the floxed neomycin gene cassette with Cre-mCherry. After confirming successful excision by genotyping PCR, CRISPR-assisted homologous recombination was employed using electroporation of a sgRNA/Cas9 ribonucleoprotein (RNP) complex and the donor plasmid YR9 (Supplementary Table 5) to introduce a TetO-BFP reporter cassette (7xTetO-PGK-Puro-BFP-pA) downstream of the Oct4-GFP reporter. Finally, a TetO Oct4 reporter ESCs clone expressing GFP and BFP was isolated by FACS analysis and genotyping. constructs encoding TetR-3xFLAG-GOI-P2A-mCherry were transduced into TetO Oct4 reporter ESCs using lentivirus.

CRISPR genetic screen

A TetO *Oct4* reporter ESC clone with stable Cas9 expression was generated for CRISPR genetic screening. Next, the Cas9-expressing ESC clonal line was transduced with lentivirus encoding TetR-3XFLAG-HP1-P2A-mCherry, and mCherry-positive ESCs were isolated by FACS for clonal expansion. ESC clone #H5 was isolated on the basis of robust initiation and maintenance *Oct4* reporter gene silencing. A retroviral sgRNA library targeting nuclear factors was used for CRISPR screening as previously described[37] with minor modifications. After 5 days and completion of G418 selection, cells were cultured in ESC medium containing Dox for 2 days. The GFP-positive cell populations of Dox-treated cells were sorted on a FACSAria II cell sorter (BD Biosciences). Unsorted mutant populations were served as background controls. Following genomic DNA isolation from GFP-positive sorted and unsorted cells, sgRNA cassettes were amplified by PCR and subjected to next-generation sequencing on an Illumina HiSeq 2500. Data were analysed as previously described[37]. Gene enrichment was determined using MAGeCK[38].

Generation of *Zfp462* endogenously tagged ESC line

For endogenous *Zfp462* tagging, $5 × 10^6$ Rosa26:BirA-V5-expressing cells[79] were electroporated with sgRNA/Cas9 RNP complex (sgRNA pre-incubated with Cas9 in cleavage buffer) mixed with 15 µg of donor plasmid containing the Tag (Avi-GFP-3XFLAG) sequence flanked by ~500 bp of homology arms (Supplementary Table 5) using Mouse Embryonic Stem Cell Nucleofector Kit (Lonza). GFP-positive ESCs were isolated by FACS and seeded for clonal expansion after 48 h. Individual ESC clones were screened by genotyping and western blot.

Editing of endogenous *Zfp462* to generate mutant and rescue ESCs

Zfp462$^{-/-}$ ESCs were generated with CRISPR–Cas9 using two different sgRNAs (Supplementary Table 5). The sgRNA sequences were cloned into SpCas9-Thy1.1 plasmid that was electroporated using Mouse Embryonic Stem Cell Nucleofector Kit (Lonza). After 36 h, Thy1.1-positive ESCs were isolated by FACS and seeded for clonal expansion. Individual clones were screened by genotyping and western blot.

Heterozygous mutant ESCs were generated as previously described[80] with few modifications. ESCs were electroporated with sgRNA/CAS9 RNP complex and a mix of 1:1 ratio WT:mutant donor single-stranded oligodeoxynucleotides carrying PAM mutation (Supplementary Table 5) using Mouse Embryonic Stem Cell Nucleofector Kit (Lonza). Twenty-four hours after transfection, ESCs were seeded for clonal expansion. Individual ESC clones were screened by PCR followed by restriction fragment length polymorphism. In case of *Zfp462*$^{+/Y1195*}$, WT allele is resistant to MseI and mutant allele is sensitive to MseI. In case of *Zfp462*$^{+/R1257*}$, WT allele is sensitive to TaqI and mutant allele is resistant to TaqI. Final ESC clones were also confirmed by Sanger sequencing and western blot.

Zfp462-FL and *Zfp462*-NT+Mid rescue ESCs were generated by transfecting sgRNA1/Cas9 and sgRNA6/Cas9 RNP complex (1:1 ratio mix) with a repair template carrying *Zfp462*-FL-GFP-3xFLAG-*Zfp462* 3'UTR-pA or *Zfp462*-NT+Mid-GFP-3xFLAG-*Zfp462* 3'UTR-pA rescue construct into *Zfp462*$^{-/-}$ ESCs using Mouse Embryonic Stem Cell Nucleofector Kit (Lonza) (Extended Data Fig. 6c and Supplementary Table 5). GFP-positive ESCs were isolated by FACS after 36 h and seeded for clonal expansion. Individual ESC clones were screened by genotyping PCR and western blot.

Western blot analysis

Cells were grown to confluency on 10 cm plates, collected in PBS, pelleted by 3 min centrifugation at 300*g*, and cell pellets were then resuspended in 5 ml of buffer 1 (10 mM Tris−HCl at pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$ and Complete Protease Inhibitor (Roche)), incubated for 20 min at 4 °C followed by a centrifugation step. The cell pellet was resuspended in buffer 2 (10 mM Tris−HCl at pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$, 0.5% IGEPAL CA-630, 10% glycerol and Roche Complete Protease Inhibitor), incubated for 10 min at 4 °C followed by centrifugation. Isolated nuclei were lysed in buffer 3 (50 mM HEPES−KOH, pH 7.3, 200 mM KCl, 3.2 mM $MgCl_2$, 0.25% Triton X-100, 0.25% NP-40, 0.1% Na-deoxycholate and 1 mM dithiothreitol (DTT) and Complete Protease Inhibitor (Roche)), 2 µl of Benzonase (E1014, Millipore) was added and incubated for 1 h at 4 °C. The lysate was cleared by centrifugation at 16,000*g* for 10 min at 4 °C, and the protein concentration in the nuclear extract was determined using the Broadford protein assay. For western blotting, 30 µg of protein was resolved on NuPAGE-Bis-Tris 4–12% mini protein gels (Invitrogen), which were transferred on to polyvinylidene fluoride membrane, blocked for 30 min in 5% non-fat dry milk in TBS with 0.1% Tween20 (TBST) and stained with primary antibodies at 4 °C overnight. The primary antibodies used for western blotting were mouse anti-Flag (1:1,000, Sigma clone M2), rabbit anti-ZNF462 (1:500, PA5-54585, Invitrogen), rabbit anti-Lamin B1 (1:5,000, ab16048, abcam), rabbit anti-G9a (1:1,000, CST# 68851, Cell Signaling) and mouse anti-H3K9me2 (1:1,000, ab1220, Abcam), Signal was detected with corresponding horseradish-peroxidase-conjugated secondary antibodies and Clarity Western ECL substrate (170-5061, Bio-Rad).

Differentiation of ESCs to NPCs and immunostaining

Differentiation was performed as previously described[48] with few modifications. ESCs are cultured on gelatin-coated plates. In addition, cells are cultured in ESC medium containing 2i inhibitors (3 µM glycogen synthase kinase inhibitor and 10 µM MEK inhibitor) for five passages before starting the differentiation experiment. For immunostaining, 4-day-RA-treated EBs were fixed with 4% paraformaldehyde (Thermo Scientific, #28906) for 30 min, washed three times with PBS and stored at 4 °C until further processing. EBs were incubated in blocking solution, which consisted of PBS (Gibco, #14190094), 4% goat/donkey serum (Bio-Rad Laboratories, #C07SA/Bio-Rad Laboratories, #C06SB) and 0.2% Triton X-100 (Sigma-Aldrich, #T8787) for at least 15 min before application of primary antibody in the same blocking solution. The primary antibody was incubated with the EBs for 2 days at 4 °C while shaking. Used primary antibodies were SOX1 (R&D Systems, #AF3369,

1:200) and FOXA2 (Cell Signaling Technologies, #8186T, 1:200). Subsequently, EBs were washed twice for 15 min each using a PBS/Tween20 solution (0.1% Tween20, Sigma-Aldrich, #P1379). A secondary antibody was then applied 1:500 in blocking solution for another 2 days at 4 °C, while shaking. Secondary antibodies used were: donkey anti-goat IgG (H + L) highly cross-adsorbed secondary antibody, Alexa Fluor Plus 488 (Invitrogen, #A32814) and goat anti-rabbit IgG (H + L) highly cross-adsorbed secondary antibody, Alexa Fluor Plus 647 (Invitrogen, #A32733). DAPI (Sigma-Aldrich, #9542) was added to the secondary antibody staining solution. Afterward, EBs were washed once with PBS/Tween20 solution for 15 min followed by a second wash with the same solution for 24 h at 4 °C while shaking. Finally, EBs were stored at 4 °C before imaging on glass slides, in FocusClear (CellExplorer Labs, #FC-101) clearing solution. Image acquisition was carried out using Spinning Disk Confocal Olympus.

Transcriptome analysis by RT–qPCR, RNA-seq and QuantSeq
For RT–qPCR experiments, total RNA was extracted from cells using the Qiagen RNeasy Mini Kit with on-column DNase digestion step. Total RNA (500 ng) was reverse transcribed using the SuperScript III Reverse Transcriptase (Invitrogen). RT–qPCR was performed on a CFX96 Real-Time PCR System (Bio-Rad) using the Promega GoTaq qPCR Master Mix (Ref: A6001) with RT–qPCR primers described in Supplementary Table 5. Relative RNA levels were calculated from C_t values according to the ΔC_t method and normalized to *Tbp* messenger RNA levels. For RNA-seq, total RNA was subjected to polyA enrichment using NEBNext Poly(A) mRNA Magnetic Isolation Module followed by library construction using the NEBNext Ultra II RNA Library Prep Kit for Illumina. QuantSeq libraries are generated using QuantSeq 3′ mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen) according to the manufacturer's instructions. RNA-seq libraries were sequenced on Illumina NovaSeq machine with 100 bp single-end sequencing and QuantSeq libraries are sequenced on NextSeq550 machine with 75 bp single-end sequencing.

ChIP
ChIP was performed as previously described with minor modifications to allow quantification using spike-in of human chromatin[81]. Chromatin from human ESCs was prepared according to the mouse ESC protocol, and 1% (relative to mouse chromatin) was spiked into mouse ESC chromatin. Combined chromatin was used for immunoprecipitation (IP) as follows: chromatin equivalent of 50 μg DNA was incubated overnight in 1× IP buffer (final: 50 mM HEPES–KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate and 0.1% SDS) with 5 μl antibody (rabbit anti-H3K4me1/ab8895/abcam, rabbit anti-H3K4me3/ab8580/abcam, mouse anti-H3K9me2/ab1220/abcam, rabbit anti-H3K9me3/ab8898/abcam, rabbit anti-H3K27ac/ab4729/abcam, mouse anti-Flag/F1804/Sigma, mouse anti-HP1γ/Millipore/05-690, rabbit anti-WIZ/NBP1-80586/Novus, goat anti-OCT4/AF1759/R&D systems, goat anti-SOX2/AF2018/R&D systems, rabbit anti-NANOG/ab80892/Abcam, rabbit anti-NR5A2/ABE2867/SIGMA and mouse anti-ESRRB/PP-H6705-00/R&D systems) at 4 °C on a rotating wheel. Subsequent IP, washes and DNA isolation were performed as previously described[81].

ZFP462 and HP1γ ChIP was performed as previously described[79]. Avi-tagged cells cross-linking and chromatin preparation was done as described above. Chromatin equivalent of 100 μg DNA was incubated overnight in 1× IP buffer (final: 50 mM HEPES–KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% DOC and 0.1% SDS) with 40 μl of pre-blocked Streptavidin M-280 Dynabeads (Invitrogen) at 4 °C on a rotating wheel. Beads were subsequently washed two times with 2% SDS in 1× TE, two times with high-salt buffer (final: 50 mM HEPES–KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton-X100, 0.1% DOC, 0.1% SDS), two times with DOC buffer (10 mM Tris pH 8, 0.25 mM LiCl, 1 mM EDTA, 0.5% NP-40 and 0.5% DOC) and one time with 1× TE. The beads were

resuspended in elution buffer, treated with RNase A and proteinase K and reverse crosslinked overnight at 65 °C. The next day, the beads were separated with magnet, supernatant was phenol/chloroform extracted and IP DNA was ethanol precipitated.

GLP ChIP was performed as previously described[82]. Avitag-GLP cells were crosslinked and nuclei preparation was done as described above. Nuclei were washed once in 5 ml NUC buffer (15 mM HEPES pH 7.5, 60 mM KCl, 15 mM NaCl and 0.32 mM sucrose) and resuspended in 1 ml of NUC buffer supplemented with Complete Protease Inhibitors, 3.3 μl of 1 M CaCl₂ and 50 U of micrococcal nuclease (LS004798, Worthington). Enzymatic activity was induced for 15 min at 37 °C and shaking at 1,000 rpm and stopped by addition of 50 μl of STOP solution (250 mM EDTA and 500 mM EGTA) and 110 μl of 10× ChIP buffer (500 mM HEPES–KOH pH 7.5, 1.4 M NaCl, 10 mM EDTA, 10% Triton X-100, 1% DOC and 1% SDS) with a further incubation for 5 min on ice. Nuclei were gently disrupted by sonication in Diagenode 15 ml Falcon tubes for eight cycles (5 s on, 5 s off) in ice-cold water using a Bioruptor Plus. Crude chromatin lysate was clarified by spinning at 20,000*g* at 4 °C for 15 min to separate insoluble debris. Chromatin equivalent of 50 μg DNA was incubated overnight in 1× IP buffer (final: 50 mM HEPES–KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% DOC and 0.1% SDS) with 40 μl of pre-blocked Streptavidin M-280 Dynabeads (Invitrogen) at 4 °C on a rotating wheel. Beads were subsequently washed two times with 2% SDS in 1× TE, two times with high-salt buffer (final: 50 mM HEPES–KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton-X100, 0.1% DOC and 0.1% SDS), two times with DOC buffer (10 mM Tris pH 8, 0.25 mM LiCl, 1 mM EDTA, 0.5% NP-40 and 0.5% DOC) and one time with 1× TE. The beads were resuspended in elution buffer, treated with RNase A and proteinase K and reverse crosslinked overnight at 65 °C. The next day, the beads were separated with magnet, supernatant was phenol/chloroform extracted and IP DNA was ethanol precipitated.

ChIP–qPCR analysis and ChIP–seq libraries preparation
ChIP DNA was subjected to qPCR analysis on a CFX96 Real-Time PCR System (Bio-Rad) using the Promega GoTaq qPCR Master Mix (Ref: A6001) with ChIP–qPCR primers described in Supplementary Table 5. Relative ChIP enrichment was calculated by percent input method. For ChIP–seq sample preparation, library construction was performed using the NEBNext Ultra-II Kit (New England Biolabs) following the manufacturer's recommendations. Libraries were sequenced on Illumina NextSeq550 machines, with 75 bp single-end sequencing.

Protein co-IP
Cells were grown to confluency on 15 cm plates, collected in PBS, pelleted by 3 min centrifugation at 300*g*. For each co-IP, 30 million equivalent cell pellets were then resuspended in 5 ml of buffer 1 (10 mM Tris–HCl at pH 7.5, 2 mM MgCl₂, 3 mM CaCl₂ and Complete Protease Inhibitor (Roche)), incubated for 20 min at 4 °C followed by a centrifugation step. The cell pellet was resuspended in buffer 2 (10 mM Tris–HCl at pH 7.5, 2 mM MgCl₂, 3 mM CaCl₂, 0.5% IGEPAL CA-630, 10% glycerol and Complete Protease Inhibitor (Roche)), incubated for 10 min at 4 °C followed by centrifugation. Isolated nuclei were lysed in buffer 3 (50 mM HEPES–KOH, pH 7.3, 200 mM KCl, 3.2 mM MgCl₂, 0.25% Triton X-100, 0.25% NP-40, 0.1% Na-deoxycholate, 1 mM DTT and Roche Complete Protease Inhibitor), 4 μl of Benzonase (E1014, Millipore) was added and incubated for 1 h at 4 °C. The lysate was cleared by centrifugation at 16,000*g* for 10 min at 4 °C. The lysate was cleared by centrifugation and incubated for 2 h at 4 °C with Tag-specific magnetic beads (Dynabeads M-280 Streptavidin/Invitrogen for IP of Avitag-ZFP462 and Avitag-GLP. Anti-FLAG M2 magnetic beads/Sigma for IP of TetR-3xFLAG-CT-ZFP462). The beads were washed four times for 10 min with buffer 4 (50 mM HEPES–KOH, pH 7.3, 300 mM KCl, 3.2 mM MgCl₂, 0.25% Triton X-100, 0.25% NP-40, 0.1% Na-deoxycholate and 1 mM DTT). Beads are further washed four times with Tris buffer (20 mM Tris pH 7.5 and 137 mM NaCl) and used for mass spectrometry analysis.

Mass spectrometry analysis

Mass spectrometry and identification of co-immunoprecipitated proteins was performed as previously described[83].

ATAC-seq

ATAC-seq was performed on WT, *Zfp462* KO, *Zfp462*-FL and *Zfp462*-NT+Mid rescue ESCs, and WT mouse NSCs following a previously published protocol[84]. The experiment was performed in biological replicates. Libraries were paired-end sequenced using an Illumina NextSeq550/Novaseq sequencer.

DNA methylation analysis by mass spectrometry

One microgram of pure genomic DNA was digested using DNA Degradase Plus (Zymo Research) for 6 h at 37 °C. For determination of the methyl-dC content, samples were analysed with liquid chromatography with tandem mass spectrometry (LC–MS/MS) using a triple quadrupole mass spectrometer, employing the selected reaction monitoring (SRM) mode of the instrument. The following transitions were used in the positive ion mode: dC m/z 228.1 → m/z 112.1 and methyl-dC m/z 242.1 → m/z 126.1. Data were interpreted using the Trace Finder software suite (Thermo Fisher Scientific) and manually validated. Calibration curves of defined mixtures of dC and methyl-dC were acquired using the same procedure and employed for calculating the molar percentage of methyl-dC.

Phylogenetic analysis

ZFP462 orthologues were collected in a convergent PSI-BLAST v2.10.0 search against the NCBI non-redundant protein database using amino acids 632–835 of NP_766455.2 as a query. The obtained sequence set was filtered, removing identical and partial sequences, keeping only sequences with identity below 95%. The remaining sequences were aligned using mafft v7.427, phylogeny was reconstructed with IQ-TREE[85] version 1.6.1 using parameters -alrt 1000 -bb 1000, and the phylogenetic tree was visualized using iTOL v6 (ref. 86).

RNA-seq and QuantSeq data analysis

RNA-seq reads were trimmed using Trim Galore v0.5.0, whereas QuantSeq reads are trimmed using BBDuk v38.06. Trimmed reads mapping to abundant sequences included in the iGenomes Ensembl GRCm38 bundle (mouse rDNA, mouse mitochondrial chromosome, phiX174 genome, adapter) were removed using bowtie2 v2.3.4.1 alignment[87]. Remaining reads were analysed using genome and gene annotation for the GRCm38/mm10 assembly obtained from *Mus musculus* Ensembl release 94. Reads were aligned to the genome using STAR v2.6.0c, and reads in genes were counted with featureCounts (subread v1.6.2) with parameter -s 2 for RNAseq and using strand-specific read counting for QuantSeq samples with parameter -s 1 (ref. 88). Differential gene expression analysis on raw counts and variance-stabilized transformation of count data for heat map visualization were performed using DESeq2 v1.18.1 (ref. 89). Functional annotation enrichment analysis of differentially expressed genes was conducted using clusterProfiler v3.6.0 in R v3.4.1 (ref. 90).

ChIP–seq data analysis

ChIP–seq reads were trimmed using Trim Galore v0.4.4 and thereafter aligned to the mm10 reference genome using BWA MEM v0.7.17. Duplicated reads were marked and excluded with Picard MarkDuplicates v2.23.4 and the obtained bam files were used to generate bigwig files using deepTools bamCoverage v3.5.0 with the parameter -normalizeUsing RPKM. Peaks were called from the sorted BAM using MACS2 v2.1.1 with default settings[91]. To analyse overlapping and non-overleaping peaks, bedtools intersect command was used. Peaks in blacklist regions were identified using bedtools intersect v2.27.1 and mm10.blacklist.bed.gz v1. Spike-in normalization of ChIP data was performed as previously described[92]. Repeat analysis was performed using RepEnrich2 following alignment of trimmed reads with bowtie2 v2.2.9. Heat maps and metaplots are generated using deepTools version 3.1.2 (ref. 93).

ATAC-seq data analysis

ATAC-seq reads were processed using nfcore/atacseq v1.0.0 with standard settings. Reads were trimmed using trim-galore v0.5.0, aligned to the mm10 genome using BWA MEM v0.7.17, and duplicated reads were marked with Picard MarkDuplicates v2.19.0. Alignment was filtered using samtools v1.9, thereby removing multimapping and duplicated reads. Peaks were called from the filtered and sorted BAM using MACS2 v2.1.2 with parameters –broad–BAMPE–keep-dup all–nomodel. Consensus peaks were obtained by merging peak calls using bedtools mergeBed v2.27.1. Heat maps and metaplots were generated using deepTools version 3.1.2. DNA sequence motif analysis in ATAC-seq peaks was performed using HOMER version homer/4.10-foss-2018b (ref. 94).

Statistics and reproducibility

All experiments were independently repeated at least two times (unless otherwise stated) with similar results, and representative results are shown throughout all figures. No statistical method was used to pre-determine sample size. No data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments or outcome assessment.

Reporting summary

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

Data Availability

High-throughput sequencing data produced in this study are deposited at the Gene Expression Omnibus (GEO) database under super series accession number GSE175369 (RNA-seq: GSE176321, QuantSeq: GSE176319, ATAC-seq: GSE176322 and ChIP–seq: GSE177058). G9a/GLP DKO ESC ATAC-seq and RNA-seq data are obtained from GSE138102 (ref. 52). REST and ADNP ChIP–seq data are obtained from GSE27148 (ref. 95) and GSE97945 (ref. 51), respectively. Meso-endoderm cell ATAC-seq data are obtained from GSE116262 (ref. 66). Processed single-cell RNA-seq data of mouse embryo stage E4.5 to E7.0 are obtained from ftp://ftp.ebi.ac.uk/pub/databases/scnmt_gastrulation[43]. Mass spectrometry data have been deposited in ProteomeXchange with the primary accession code PXD037238. All other data supporting the findings of this study are available from the corresponding author on reasonable request. Source data are provided with this paper.

References

78. Blomfield, I. M. et al. Id4 promotes the elimination of the pro-activation factor Ascl1 to maintain quiescence of adult hippocampal stem cells. *Elife* **8**, e48561 (2019).
79. Baubec, T., Ivánek, R., Lienert, F. & Schübeler, D. Methylation-dependent and -Independent genomic targeting principles of the MBD protein family. *Cell* **153**, 480–492 (2013).
80. Paquet, D. et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125–129 (2016).
81. Moussa, H. F. et al. Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing. *Nat. Commun.* **10**, 1931 (2019).
82. Mayer, D. et al. Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2. *EMBO J.* **39**, e102591 (2020).
83. Batki, J. et al. The nascent RNA binding complex SFiNX licenses piRNA-guided heterochromatin formation. *Nat. Struct. Mol. Biol.* **26**, 720–731 (2019).

84. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).

85. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

86. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127–128 (2007).

87. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

88. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

89. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

90. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

91. Zhang, Y. et al. Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9**, R137 (2008).

92. Fursova, N. A. et al. Synergy between variant PRC1 complexes defines Polycomb-mediated gene repression. *Mol. Cell* **74**, e8 (2019).

93. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

94. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

95. Arnold, P. et al. Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. *Genome Res.* **23**, 60–73 (2013).

Acknowledgements

Author contributions

R.Y., K.S. and O.B. initiated and designed the study. R.Y and K.S. generated cell lines, performed CRISPR–Cas9 genetic screen and differentiation assays. R.Y. performed all molecular biology experiments and revisions. S.M. and P.H. provided experimental advice and P.H. carried out immunohistochemistry. R.Y. and M.N analysed transcriptome and epigenome data. J.W. and G.M. analysed CRISPR–Cas9 genetic screen data. U.E., G.M and G.V. provided the CRISPR sgRNA library and helped with the screen design. L.M. and C.P. participated in experiments. L.I. and D.S. shared reagents and experimental expertise. J.B. co-supervised part of the project in his laboratory. O.B. supervised all aspects of the project. The manuscript was prepared by R.Y. and O.B. D.S. and J.B. edited the manuscript. All authors discussed results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41556-022-01051-2.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41556-022-01051-2.

Correspondence and requests for materials should be addressed to Ramesh Yelagandula or Oliver Bell.

Peer review information *Nature Cell Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at www.nature.com/reprints.
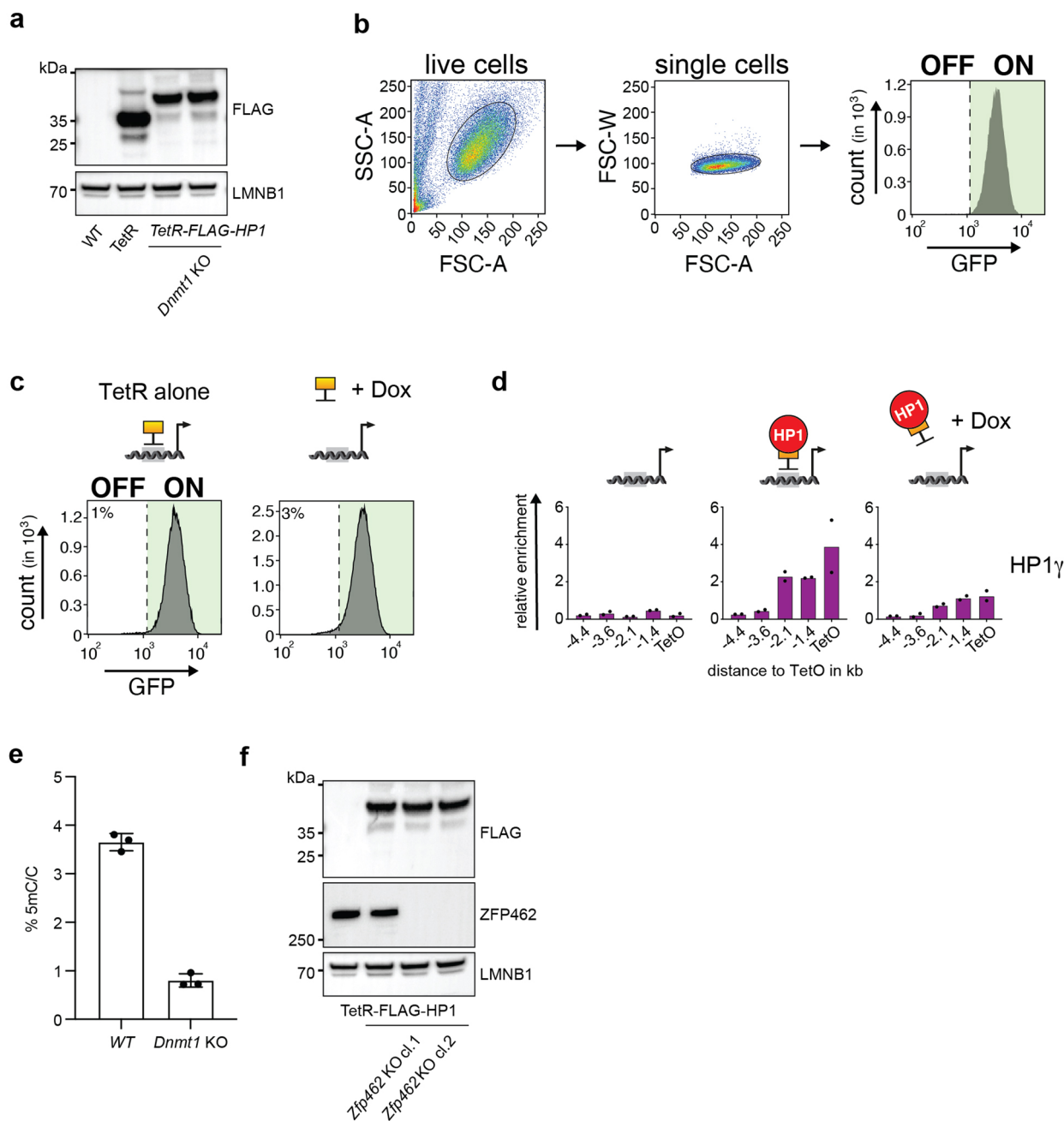
Extended Data Fig. 1 | Characterization of WT and mutant CiA *Oct4* dual reporter cells. a) Western blot confirms expression of TetR-FLAG and TetR-FLAG-HP1 proteins in WT and *Dnmt1* knockout (KO) CiA *Oct4* dual reporter ESCs. LMNB1 is used as protein loading control. **b)** FACS gating strategy used to analyse the GFP fluorescence in cells. **c)** Flow cytometry histograms show GFP expression in CiA *Oct4* dual reporter cells after TetR-FLAG recruitment and after TetR-FLAG release following Dox addition for four days. **d)** ChIP-qPCR shows relative enrichment of HP1γ upstream of the TetO site before tethering, in the presence of TetR-FLAG-HP1 and after Dox-dependent release of TetR-FLAG-HP1 for four days. n = 2 independent biological replicates. **e)** Bar plot shows fraction of cytosine methylation (5mC) in WT and *Dnmt1* KO CiA *Oct4* dual reporter ESCs measured by LC-MS. n = 3 independent biological replicates. Data are presented as mean values +/− SD. **f)** Western blot shows expression of TetR-FLAG-HP1 and ZFP462 in WT and two independent *Zfp462* KO CiA *Oct4* dual reporter ESC lines. LMNB1 is used as loading control.
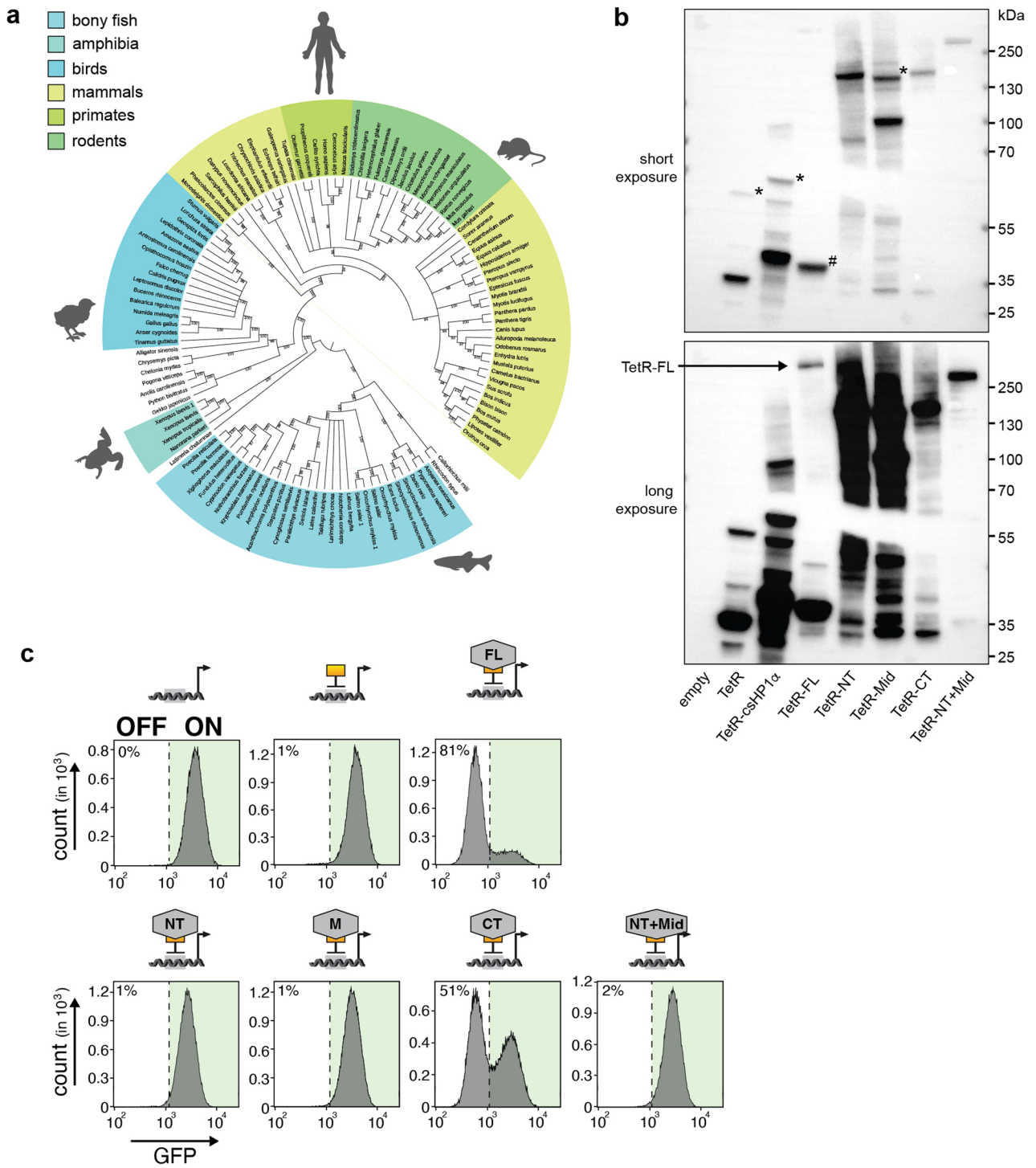
Extended Data Fig. 2 | ZFP462 is conserved across vertebrates and acts as transcriptional repressor. a) Phylogenetic tree of ZFP462 protein orthologues in vertebrate species. Bootstrap values are shown on branches. **b)** Western blot shows expression ZFP462 fusions with TetR-FLAG in CiA *Oct4* dual reporter ESCs. Bottom image shows long exposure. Hashtag indicates cleaved TetR-FLAG protein from TetR-FLAG-ZFP462-FL (full-length). TetR-FLAG fusions initially express mCherry for selection which is cleaved via P2A signal. Asterisk marks P2A uncleaved protein product. **c)** Representative flow cytometry histograms show GFP expression in CiA *Oct4* dual reporter ESCs expressing TetR-FLAG fusions with full-length or truncated ZFP462 proteins. Each histogram is average profile of 100,000 analysed cells.
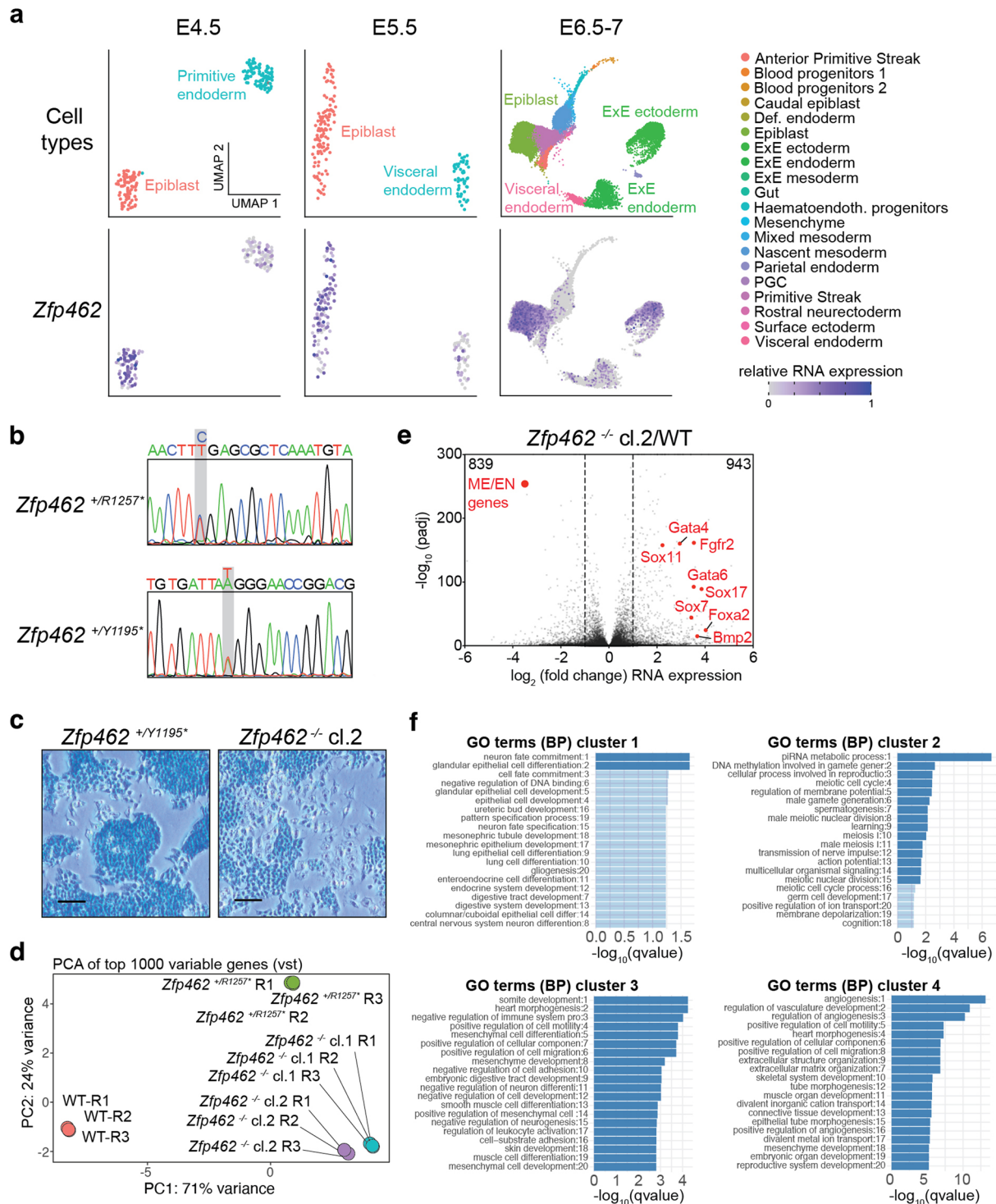
Extended Data Fig. 3 | Zfp462 expression analysis and impact of Zfp462 deletion on ESC morphology and gene expression. a) UMAP plots visualize lineage assignments of cells in mouse embryonic developmental stages (embryonic day (E) E4.5 cells, E5.5 cells and E6.5-7 cells) as previously described (Wolf Reik et.al) (top). UMAP plots of Zfp462 RNA expression at corresponding developmental stages (bottom). **b)** Sanger sequence chromatograms of heterozygous Zfp462 mutants. Heterozygous non-sense mutations are highlighted in grey. **c)** Alkaline phosphatase staining of Zfp462 +/Y1195* and Zfp462 -/- cl.2 ESCs (scale bar = 100 μm). **d)** Correlation plot shows principal component analysis (PCA) of replicate RNA-seq experiments from WT, two Zfp462 +/- and two Zfp462 -/- ESC lines. **e)** Volcano plot shows differential gene expression of Zfp462 -/- cl.2 ESCs compared to WT ESCs. (n = three replicates). **f)** Bar plots show gene ontology (GO) terms enriched in the four clusters of heatmap in Fig. 3e.
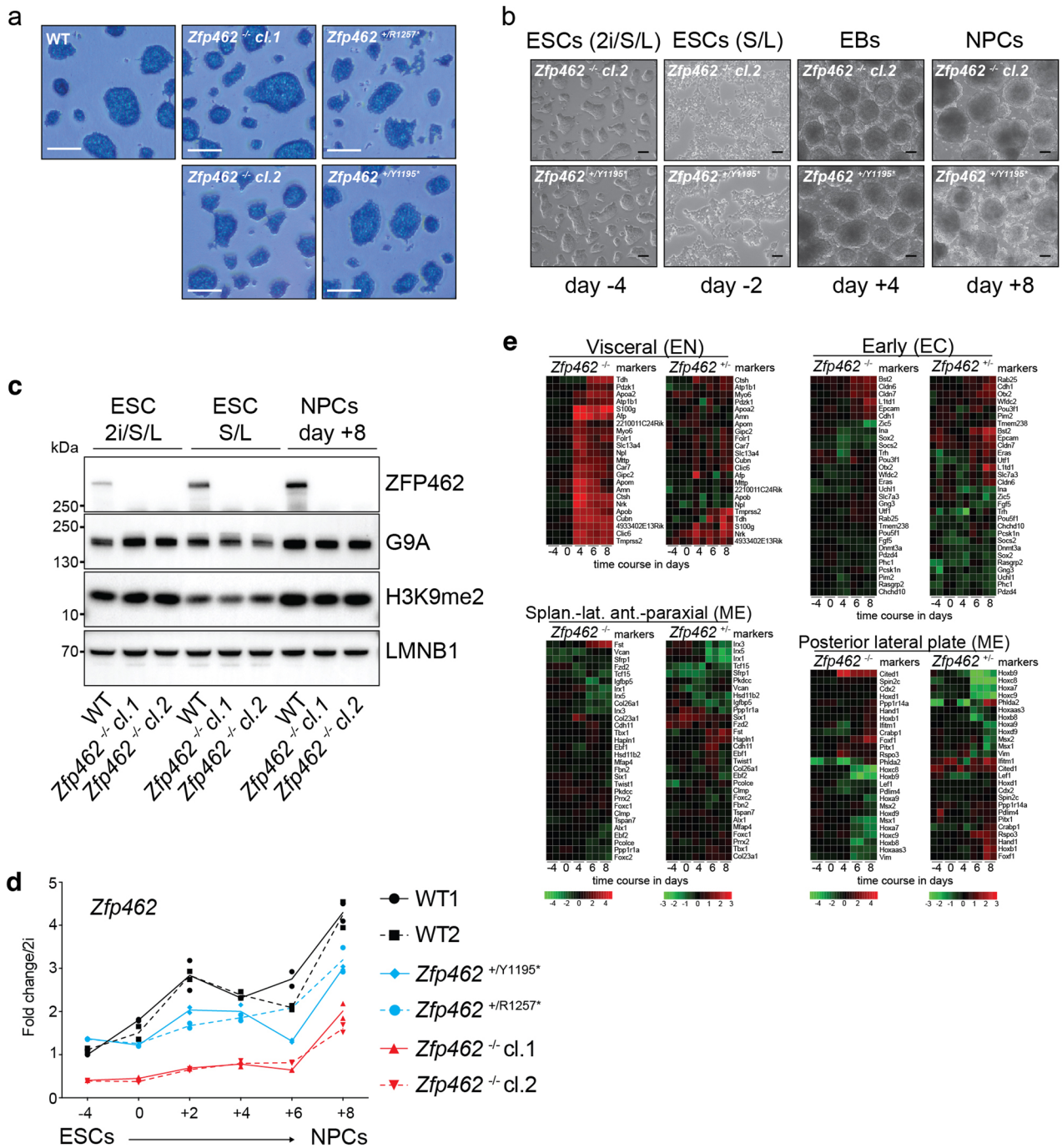
Extended Data Fig. 4 | Lineage-specifying genes are deregulated during neuronal differentiation. a) Alkaline phosphatase staining of WT and *Zfp462* mutant ESCs cultured in 2i/LIF/Serum medium (scale bar = 180 μm). **b)** Representative bright field images show *Zfp462*[+/Y1195*] and *Zfp462*[−/−] cl.2 at corresponding stages of neural differentiation. Scale bar = 100μm. **c)** Western blot analysis shows levels of ZFP462, G9A and H3K9me2 in WT and *Zfp462*[−/−] ESCs during neural differentiation. LMNB1 is used as loading control. **d)** Line plot shows RT-qPCR analysis of *Zfp462* RNA expression during neural differentiation (n = two replicates). Expression level is shown relative to ESCs (2i/S/L). **e)** Heatmaps show differential expression of selected marker genes specific for endodermal (EN), mesodermal (ME) and neural lineages (EC) in heterozygous and homozygous *Zfp462* mutant cells during neural differentiation.
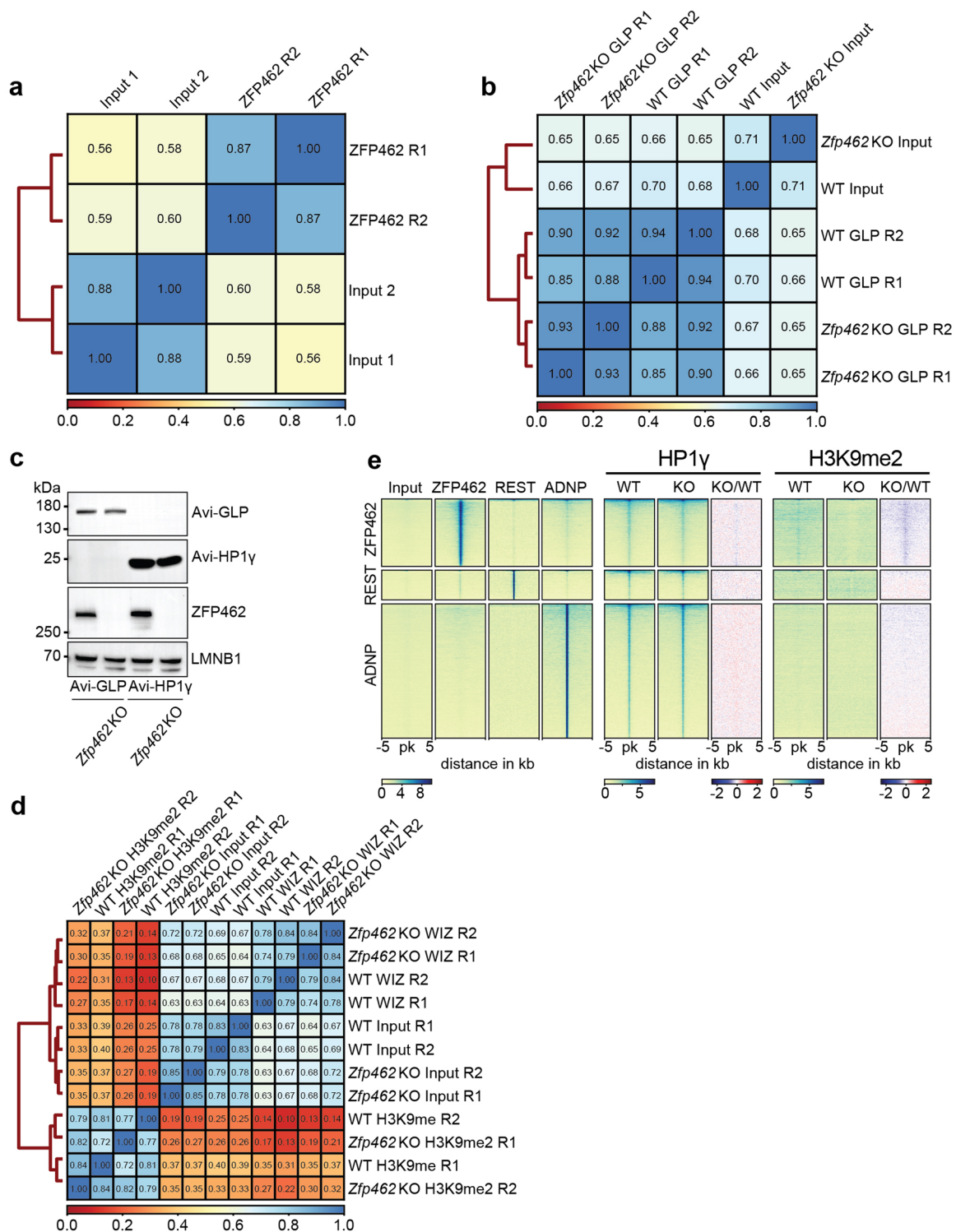
Extended Data Fig. 5 | ChIP-seq profiles of ZFP462, REST, ADNP, HP1γ and H3K9me2 in WT and *Zfp462* KO ESCs. a) Heatmap shows Spearman correlation between input controls and two independent ZFP462 ChIP-seq experiments. **b)** Heatmap shows Spearman correlation between input controls and two independent ChIP-seq experiments of GLP and HP1γ in WT and *Zfp462* KO ESCs. **c)** Western blot shows ZFP462 expression in WT and *Zfp462* KO Avi-FLAG-tagged GLP and HP1γ ESCs. **d)** Heatmap shows Spearman correlation between input controls and two independent ChIP-seq experiments of WIZ and H3K9me2 in WT and *Zfp462* KO ESCs. **e)** Heatmaps of ZFP462, REST, ADNP, HP1γ and H3K9me2 ChIP-seq signals at ZFP462, REST and ADNP peaks in WT and *Zfp462* KO ESCs. Blue-to-red scaled heatmap represents ChIP-seq enrichment ratios between *Zfp462* KO versus WT (KO/WT). Each heatmap represents a 10 kb window centred on peak midpoints, sorted by ZFP462, REST and ADNP signal in their respective clusters. Below are scale bars (n = average distribution of two replicates).
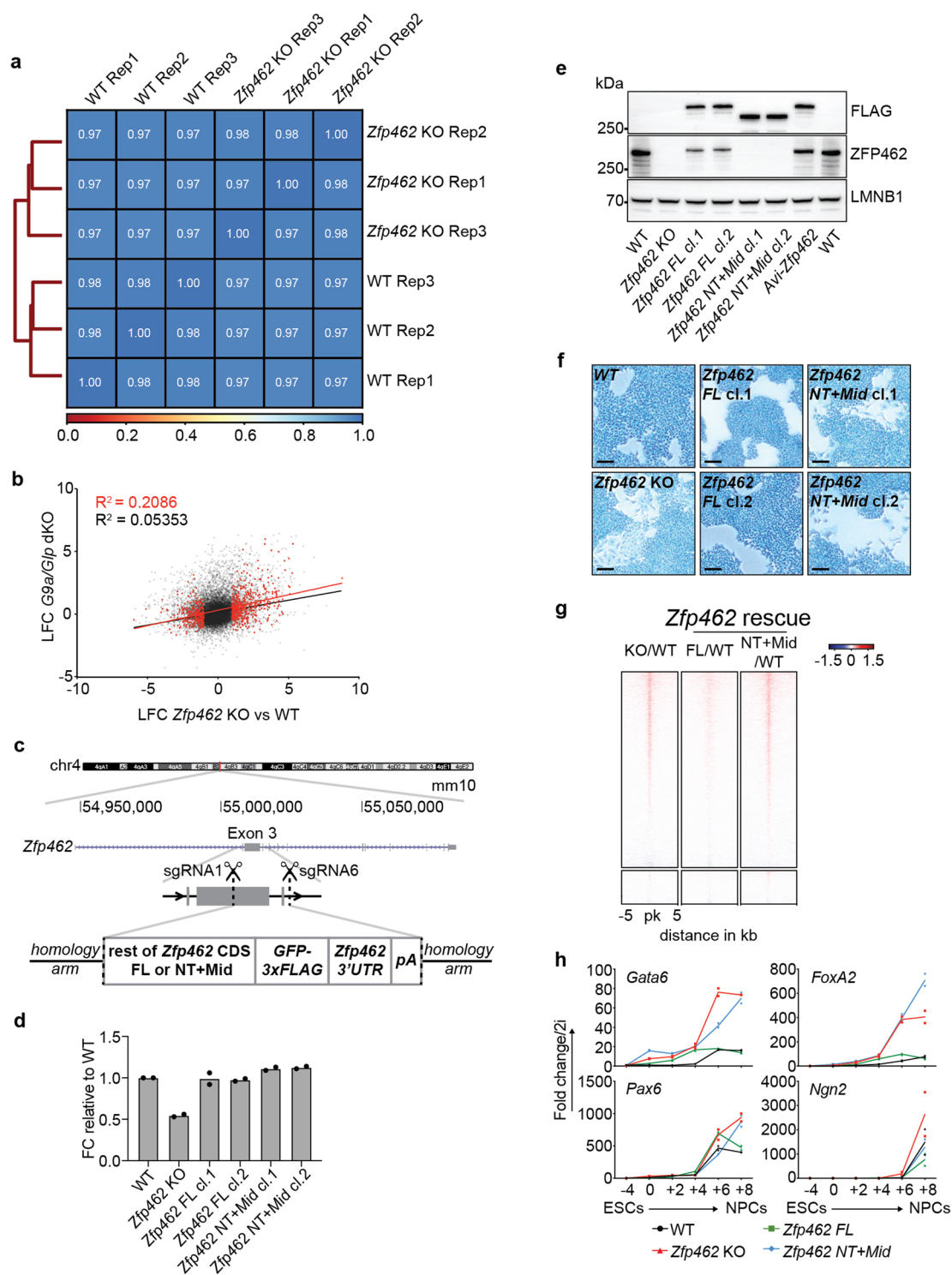
Extended Data Fig. 6 | Rescue of *Zfp462* KO ESCs with ZFP462 full-length and ZFP462 C-terminal truncation. a) Heatmap shows Spearman correlation between three independent ATAC-seq experiments in WT and *Zfp462* KO ESCs. **b)** Scatter plot shows pairwise correlation of gene expression changes between *Zfp462* KO vs WT and *G9a/Glp* dKO vs WT. coefficient of determination (R²): black - all genes, red – differentially regulated genes in *Zfp462* KO vs WT (padj. ≤ 0.05, LFC ≥1). **c)** Cartoon depicts CRISPR strategy to knock-in coding sequences for ZFP462 full-length and ZFP462 NT+Mid proteins at the *Zfp462* gene locus. Insertions were targeted in-frame with exon 3. **d)** Bar plot shows RT-qPCR analysis of *Zfp462* RNA transcript levels in WT, KO and rescue ESCs. n = 2 independent

biological replicates. **e)** Western blot analysis shows ZFP462 protein levels in WT, *Zfp462* KO and rescue ESCs. LMNB1 is used as loading control. **f)** Alkaline phosphatase staining of WT and *Zfp462* KO and rescue ESCs (scale bar = 100μm). **g)** Heatmaps of ATAC-seq signal ratios at ZFP462 and REST peaks of *Zfp462* KO vs WT (KO/WT) and *Zfp462* rescue ESCs (ZFP462 FL/WT) or (ZFP462 NT+Mid/WT). Each row represents a 10 kb window centred on peak midpoints, sorted by KO/WT enrichment ratio (n = average distribution of two ATAC-seq replicates). **h)** Line plots show RT-qPCR analysis of lineage marker expression during neural differentiation (n = two replicates) in WT, *Zfp462* KO and rescue ESCs. Expression levels are shown relative to ESCs (2i/S/L).
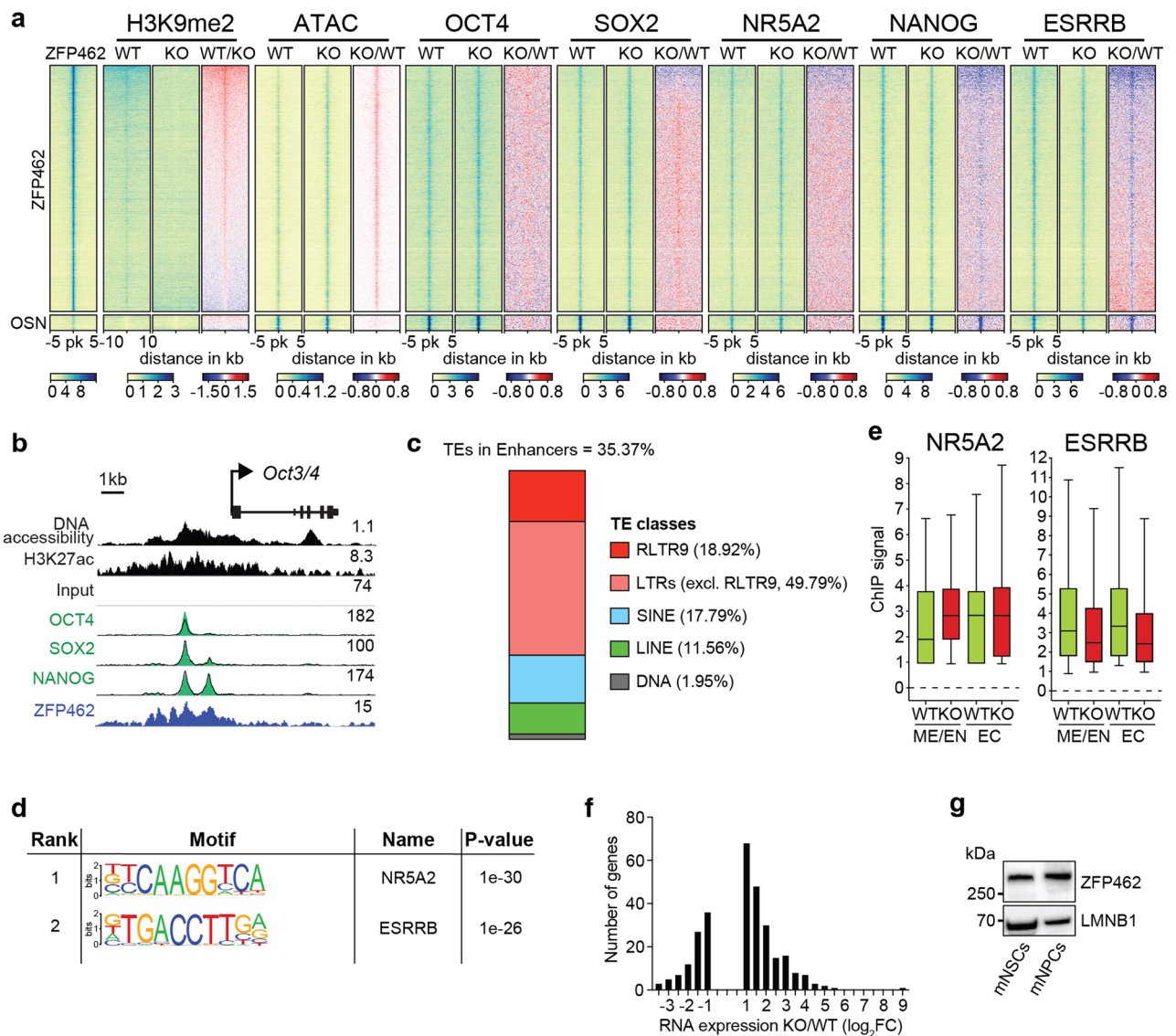
Extended Data Fig. 7 | Correlation between genomic distribution of ZFP462 and pluripotency transcription factors. a) Heatmaps show ATAC-seq and ChIP-seq signal of ZFP462, H3K9me2, OCT4, SOX2, NANOG, ESRRB and NR5A2 at peaks of ZFP462, OSN (shared OCT4-SOX2-NANOG peaks not overlapping with ZFP462 peaks) and REST. Red-to-blue scaled heatmaps represent signal ratios between *Zfp462* KO versus WT (KO/WT). Each row represents a 10 kb window centred on peak midpoints, sorted by H3K9me2 KO/WT ChIP signal loss. (n = average distribution of two ChIP-seq replicates / ATAC-seq three replicates). **b)** Genomic screenshot shows DNA accessibility (ATAC-seq) and ChIP-seq signals of H3K27ac and ZFP462 at the *Oct3/4* locus in WT ESCs. ChIP-seq signals of OCT4, SOX2 and NANOG in WT (black line) and *Zfp462* KO ESCs (green fill) are superimposed. ATAC-seq and ChIP-seq profiles are normalized to library size. **c)** Bar plot shows percentage of ZFP462-bound TE families overlapping with ChromHMM-annotated enhancers in ESCs. ZFP462-bound TEs contribute

a total of 35.37% of ChromHMM-annotated enhancers in ESCs. **d)** HOMER analysis of known DNA sequence motifs enriched at ZFP462 peaks overlapping ChromHMM-annotated ME/EN-specific enhancers. Top ranked DNA sequence motifs and respective significance values are shown in the table. **e)** Box plots shows enrichment of NR5A2 and ESRRB ChIP signal in WT and *Zfp462* KO ESCs at ZFP462 peaks overlapping with Mesoderm (ME)/Endoderm (EN)- and Ectoderm (EC)-specific enhancers. n = 553 (EN/ME), n = 314 (EC). Shown are median (horizontal line), 25th to 75th percentiles (boxes), and 90% (whiskers). Outliers are not shown. **f)** Bar plot shows frequency distribution of significantly up- and down-regulated genes located proximal to ZFP462 peaks annotated as ME/EN-specific enhancers (KO/WT, LFC ≥ 1 and padj. ≤ 0.05). **g)** Western blot shows ZFP462 protein expression in NSCs isolated from mouse brain and NPCs differentiated from WT ESCs. LMNB1 is used as loading control.

Corresponding author(s): Oliver Bell

Last updated by author(s): Oct 10, 2022

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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.

n/a	Confirmed	
☐	☒	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	FACS data was obtained on FACS LSR II Fortessa/BD Biosciences instrument using BD FACSDiva™ Software v8.0. Confocal images were obtained on Spinning Disk Confocal/Olympus. LC-MS data was obtained using Xcalibur™ Software. NGS libraries were sequenced on NextSeq550 and NovaSeq machines/Illumina.
Data analysis	FlowJo 10.4.2, ImageJ 2.1.0/1.53c, PSI-BLAST v2.10.0, IQ-TREE version 1.6.1, iTOL v6, Prism 9, Trim Galore v0.5.0, BBDuk v38.06, bowtie2 v2.3.4.1, STAR v2.6.0c, featureCounts (subread v1.6.2), DESeq2 v1.18.1, clusterProfiler v3.6.0, Trim Galore v0.4.4, BWA MEM v0.7.17, Picard MarkDuplicates v2.23.4, deepTools bamCoverage v3.5.0, MACS2 v2.1.1, bedtools intersect v2.27.1, RepEnrich2 nfcore/atacseq v1.0.0, samtools v1.9, deepTools version 3.1.2 and HOMER version homer/4.10-foss-2018b.

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All sequencing data used for this study have been deposited with NCBI Gene Expression Omnibus (GEO) database under super series accession number GSE175369 (RNA-seq data in GSE176321, QuantSeq data in GSE176319, ATAC-seq data in GSE176322 and ChIP-seq data in GSE177058). G9a/GLP DKO ESCs ATAC-seq and RNA-

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.

☒ 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	No sample size calculation was performed; all experiments in this work build on established experimental schemes in the field of chromatin biology and stem cell biology. Sample sizes are indicated throughout manuscript.
Data exclusions	No data was excluded.
Replication	For each experiment, all attempts at replication were successful.
Randomization	Randomization is not an appropriate aspect of the experimental design presented in this study as measurements and data collection were performed using the same clonal cell lines, nucleic acid and proteins and analyzed with identical methodologies.
Blinding	Blinding is not possible because investigators who performed experiments also analyzed the data.

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

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Western blot analysis: Mouse anti-Flag (F1804/Sigma clone M2), Rabbit anti-ZNF462 (PA5-54585/Invitrogen), Rabbit anti-Lamin B1 (ab16048/Abcam), Rabbit anti-G9A (68851S/Cell signaling), Mouse anti-H3K9me2 (ab1220/Abcam), Anti-Rabbit IgG (H+L) Secondary Antibody-HRP conjugate (65-6120/Invitrogen) and Anti-Mouse IgG (H+L) Secondary Antibody-HRP Conjugate (W4021/Promega). Immunostaining analysis: Goat anti-SOX1 (AF3369/R&D Systems), Rabbit anti-FOXA2 (8186T/Cell Signaling), Donkey anti-Goat IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 488 (Invitrogen, #A32814) and Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 647 (Invitrogen, # A32733). ChIP-qPCR and ChIP-seq analysis: Mouse anti-Flag (F1804/Sigma clone M2), Rabbit anti-H3K4me1(ab8895/abcam), Rabbit anti-H3K4me3(ab8580/abcam), Mouse anti-H3K9me2(ab1220/abcam), Rabbit anti-H3K9me3(ab8898/abcam), Rabbit anti-H3K27ac(ab4729/abcam), Mouse anti-HP1γ(Millipore/05-690), Rabbit anti-WIZ(NBP1-80586/Novus), Goat anti-OCT4(AF1759/R&D systems), Goat anti-SOX2(AF2018/R&D systems), Rabbit anti-NANOG(ab80892/abcam), Rabbit anti-NR5A2(ABE2867/SIGMA) and Mouse anti-ESRRB(PP-H6705-00/R&D systems).
Validation	1) Validation of Mouse anti-FLAG antibody (F1804/Sigma clone M2): The manufacturer states that the antibody has been optimized for single band detection of the FLAG antigen (DYKDDDDK) fused proteins in mammalian, plant and bacterial expression systems. https://www.sigmaaldrich.com/AT/en/product/sigma/f1804 2) Validation of Rabbit anti-ZNF462 antibody(PA5-54585/Invitrogen): The manufacturer states that the specificity of the antibody was

tested by immunocytochemistry of Human cells. https://www.thermofisher.com/antibody/product/ZNF462-Antibody-Polyclonal/PA5-54585
We tested its specificity in mouse cell lysates using western blot analysis.

3) Validation of Rabbit anti-Lamin B1 antibody(ab16048/Abcam): The manufacturer states that the specificity of the antibody was tested by western blot in samples derived from Mouse, Rat and Human. https://www.abcam.com/lamin-b1-antibody-nuclear-envelope-marker-ab16048.html

4) Validation of Rabbit anti-G9A antibody(68851S/Cell signaling): The manufacturer states that the specificity of the antibody was tested by western blot in samples derived from Mouse, Rat, Human and Monkey. https://www.cellsignal.at/products/primary-antibodies/g9a-ehmt2-d5r4r-xp-rabbit-mab/68851

5) Validation of Mouse anti-H3K9me2 antibody(ab1220/Abcam): The manufacturer states that the specificity of the antibody to H3K9me2 was tested " By peptide ELISA ab1220 recognizes di methyl K9, but not unmodified K9, mono methyl K9, tri methyl K9, di methyl K27, tri methyl K27, mono methyl K4, di methyl K4 or tri methyl K4. By Western blot ab1220 is blocked by di methyl K9, but not by unmodified K9, mono methyl K9, tri methyl K9, di methyl K27, tri methyl K27, mono methyl K4, di methyl K4 or tri methyl K4. This indicates the specificity of ab1220 for di methyl K9 of Histone H3". https://www.abcam.com/histone-h3-di-methyl-k9-antibody-mabcam-1220-chip-grade-ab1220.html

6) Validation of Anti-Rabbit IgG (H+L) Secondary Antibody-HRP conjugate(65-6120/Invitrogen):The manufacturer states that the specificity of the antibody was validated by western blot on various mammalian cell lysates. https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Secondary-Antibody-Polyclonal/65-6120

7) Validation of Anti-Mouse IgG (H+L) Secondary Antibody-HRP Conjugate (W4021/Promega): The manufacturer states that the specificity of the antibody was validated by western blot on mammalian cell lysates. https://at.promega.com/en/products/protein-detection/primary-and-secondary-antibodies/anti_mouse-igg-h-and-l-hrp-conjugate/?catNum=W4021#resources

8) Validation of Goat anti-SOX1 (AF3369/R&D Systems): The manufacturer states that the specificity of the antibody has been validated by performing immunofluorescence assays against SOX1 protein in mammalian cells. https://www.rndsystems.com/products/human-mouse-rat-sox1-antibody_af3369

9) Validation of Rabbit anti-FOXA2 (8186T/Cell Signaling): The manufacturer states that the specificity of the antibody has been validated by performing immunofluorescence assays against FOXA2 protein in mammalian cells. https://www.cellsignal.at/products/primary-antibodies/foxa2-hnf3b-d56d6-xp-rabbit-mab/8186

10) Validation of Donkey anti-Goat IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 488 (Invitrogen, #A32814): The manufacturer states that the specificity of the antibody has been validated by performing immunofluorescence assays against a variety of proteins. https://www.thermofisher.com/antibody/product/Donkey-anti-Goat-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A32814

11) Validation of Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor Plus 647 (Invitrogen, # A32733): The manufacturer states that the specificity of the antibody has been validated by performing immunofluorescence assays against a variety of proteins. https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A32733

12) Validation of Rabbit anti-H3K4me1(ab8895/abcam): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.abcam.com/histone-h3-mono-methyl-k4-antibody-chip-grade-ab8895.html

13) Validation of Rabbit anti-H3K4me3(ab8580/abcam): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.abcam.com/histone-h3-tri-methyl-k4-antibody-chip-grade-ab8580.html

14) Validation of Rabbit anti-H3K9me3(ab8898/abcam): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.abcam.com/histone-h3-tri-methyl-k9-antibody-chip-grade-ab8898.html

15) Validation of Rabbit anti-H3K27ac(ab4729/abcam): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.abcam.com/histone-h3-acetyl-k27-antibody-chip-grade-ab4729.html

16) Validation of Mouse anti-HP1γ(Millipore/05-690): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.merckmillipore.com/AT/de/product/Anti-HP1-Antibody-clone-42s2,MM_NF-05-690

17) Validation of Rabbit anti-WIZ(NBP1-80586/Novus): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.novusbio.com/products/wiz-antibody_nbp1-80586

18) Validation of Goat anti-OCT4(AF1759/R&D systems): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.rndsystems.com/products/human-mouse-oct-3-4-antibody_af1759

19) Validation of Goat anti-SOX2(AF2018/R&D systems): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.rndsystems.com/products/human-mouse-rat-sox2-antibody_af2018

20) Validation of Rabbit anti-NANOG(ab80892/abcam): The antibody specificity has been validated for ChIP assay in mouse cells by multiple previously published studies, e.g.: Ziying Liu et.al., 2017, Molecular Cell (https://doi.org/10.1016/j.molcel.2017.01.017). https://www.abcam.com/nanog-antibody-ab80892.html

21) Validation of Rabbit anti-NR5A2(ABE2867/SIGMA): The manufacturer states that the specificity of the antibody has been validated for ChIP assay in mammalian cells. https://www.merckmillipore.com/AT/de/product/Anti-NR5A2-LRH1,MM_NF-ABE2867

22) Validation of Mouse anti-ESRRB(PP-H6705-00/R&D systems): Chronis et.al., (https://doi.org/10.1016/j.cell.2016.12.016) validated the specificity of the antibody for ChIP assay in mouse cells. https://www.rndsystems.com/products/human-err-beta-nr3b2-antibody-h6705_pp-h6705-00

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	1) Mouse Embryonic Stem Cells (mESCs) (TC1/RRID:CVCL_M350). Cells are obtained from the Prof. Crabtree lab at Stanford University, USA. 2) GLP-Avi-tagged mESCs from Mayer, D. et al. "Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2." EMBO J. 39, e102591 (2020). Cells are obtained from Dr. Joerg Betschinger lab at Friedrich Miescher Institute for Biomedical Research, Switzerland. 3) HP1 gamma-Avi-tagged mESCs from Ostapcuk, V. et al. "Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes."" Nature 557, 739–743 (2018). Cells are obtained from Dr. Marc Buehler lab at Friedrich Miescher Institute for Biomedical Research, Switzerland. 4) Neural Stem Cells (NSCs) isolated from subventricular zone of 7-8 week mouse brain. Cells are obtained from Dr. Noelia Urbán lab at Institute of Molecular Biotechnology, Austria.
Authentication	Genotype of cell lines was tested at the level of DNA sequence and protein.
Mycoplasma contamination	All cell lines are free of Mycoplasma contamination. Cells were tested for mycoplasma every other week during culture.
Commonly misidentified lines (See ICLAC register)	Not used in this study.

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.*	ChIP-seq, ATAC-seq, RNA-seq and QuantSeq data produced in this study are deposited at GEO under super series accession number GSE175369. Reviewer access token: wlmtmuactrevveh (please note that same token applies to all sub series accessions). ChIP-seq: GSE177058 (sub series) ATAC-seq: GSE176322 (sub series) RNA-seq: GSE176321 (sub series) Quantseq: GSE176319 (sub series)
Files in database submission	GSM5373686 ChIPseq for ZFP462_Rep1 GSM5373687 ChIPseq for ZFP462_Rep2 GSM5373688 ChIPseq for ZFP462 in NPCs_Rep1 GSM5373689 ChIPseq for ZFP462 in NPCs_Rep2 GSM5373690 Input_for_ZFP462 ChIPseq_Rep1 GSM5373691 Input_for_ZFP462 ChIPseq_Rep2 GSM5373692 ChIPseq for GLP in WT_Rep1 GSM5373693 ChIPseq for GLP in WT_Rep2 GSM5373694 ChIPseq for GLP in Zfp462KO_Rep1 GSM5373695 ChIPseq for GLP in Zfp462KO_Rep2 GSM5373696 Input_for_GLP ChIPseq_in_WT GSM5373697 Input_for_GLP_ChIPseq in Zfp462KO GSM5373698 ChIPseq for HP1gamma in WT_Rep1 GSM5373699 ChIPseq for HP1gamma in WT_Rep2 GSM5373700 ChIPseq for HP1gamma in Zfp462KO_Rep1 GSM5373701 ChIPseq for HP1gamma in Zfp462KO_Rep2 GSM5373702 Input_for_antibody_ChIPs in WT_Rep1 GSM5373703 Input_for_antibody_ChIPs in WT_Rep2 GSM5373704 Input_for_antibody_ChIPs in Zfp462KO Rep1 GSM5373705 Input_for_antibody_ChIPs in Zfp462KO Rep2 GSM5373706 ChIPseq for H3K9me2 in WT_Rep1 GSM5373707 ChIPseq for H3K9me2 in WT_Rep2 GSM5373708 ChIPseq for H3K9me2 in Zfp462KO_Rep1 GSM5373709 ChIPseq for H3K9me2 in Zfp462KO_Rep2 GSM5373710 ChIPseq for H3K27ac in WT_Rep1 GSM5373711 ChIPseq for H3K27ac in WT_Rep2 GSM5373712 ChIPseq for H3K27ac in Zfp462KO_Rep1 GSM5373713 ChIPseq for H3K27ac in Zfp462KO_Rep2 GSM5373714 ChIPseq for WIZ in WT_Rep1 GSM5373715 ChIPseq for WIZ in WT_Rep2

GSM5373716 ChIPseq for WIZ in Zfp462KO_Rep1
GSM5373717 ChIPseq for WIZ in Zfp462KO_Rep2
GSM5373718 ChIPseq for NANOG in WT_Rep1
GSM5373719 ChIPseq for NANOG in WT_Rep2
GSM5373720 ChIPseq for NANOG in Zfp462KO_Rep1
GSM5373721 ChIPseq for NANOG in Zfp462KO_Rep2
GSM5373726 ChIPseq for SOX2 in WT_Rep1
GSM5373727 ChIPseq for SOX2 in WT_Rep2
GSM5373728 ChIPseq for SOX2 in Zfp462KO_Rep1
GSM5373729 ChIPseq for SOX2 in Zfp462KO_Rep2
GSM6395551 ChIPseq for OCT4 in WT_Rep1
GSM6395552 ChIPseq for OCT4 in WT_Rep2
GSM6395553 ChIPseq for OCT4 in Zfp462KO_Rep1
GSM6395554 ChIPseq for OCT4 in Zfp462KO_Rep2
GSM6395555 ChIPseq for NR5A2 in WT_Rep1
GSM6395556 ChIPseq for NR5A2 in WT_Rep2
GSM6395557 ChIPseq for NR5A2 in Zfp462KO_Rep1
GSM6395558 ChIPseq for NR5A2 in Zfp462KO_Rep2
GSM6395559 ChIPseq for ESRRB in WT_Rep1
GSM6395560 ChIPseq for ESRRB in WT_Rep2
GSM6395561 ChIPseq for ESRRB in Zfp462KO_Rep1
GSM6395562 ChIPseq for ESRRB in Zfp462KO_Rep2
GSM6395563 ChIPseq for H3K4me1 in WT_Rep1
GSM6395564 ChIPseq for H3K4me1 in WT_Rep2
GSM6395565 ChIPseq for H3K4me3 in WT_Rep1
GSM6395566 ChIPseq for H3K4me3 in WT_Rep2
GSM5362673 RNA-seq of ESC_WT_Rep1
GSM5362674 RNA-seq of ESC_WT_Rep2
GSM5362675 RNA-seq of ESC_WT_Rep3
GSM5362676 RNA-seq of ESC_Zfp462-/- cl.1_Rep1
GSM5362677 RNA-seq of ESC_Zfp462-/- cl.1_Rep2
GSM5362678 RNA-seq of ESC_Zfp462-/- cl.1_Rep3
GSM5362679 RNA-seq of ESC_Zfp462-/- cl.2_Rep1
GSM5362680 RNA-seq of ESC_Zfp462-/- cl.2_Rep2
GSM5362681 RNA-seq of ESC_Zfp462-/- cl.2_Rep3
GSM5362682 RNA-seq of ESC_Zfp462+/R1257*_Rep1
GSM5362683 RNA-seq of ESC_Zfp462+/R1257*_Rep2
GSM5362684 RNA-seq of ESC_Zfp462+/R1257*_Rep3
GSM5362685 ATAC-seq of wild type mESCs replicate1
GSM5362686 ATAC-seq of wild type mESCs replicate2
GSM5362687 ATAC-seq of wild type mESCs replicate3
GSM5362688 ATAC-seq of Zfp462-/- cl.1 mESCs replicate1
GSM5362689 ATAC-seq of Zfp462-/- cl.1 mESCs replicate2
GSM5362690 ATAC-seq of Zfp462-/- cl.1 mESCs replicate3
GSM5362691 ATAC-seq of wild type mNSCs replicate1
GSM5362692 ATAC-seq of wild type mNSCs replicate2
GSM6395498 ATACseq of wild type mESCs replicate1 (control for rescue experiment)
GSM6395499 ATACseq of wild type mESCs replicate2 (control for rescue experiment)
GSM6395500 ATACseq of Zfp462-KO mESCs replicate1 (control for rescue experiment)
GSM6395501 ATACseq of Zfp462-KO mESCs replicate2 (control for rescue experiment)
GSM6395502 ATACseq of Zfp462-KO_ZFP462-FL-Knock-In mESCs replicate1
GSM6395503 ATACseq of Zfp462-KO_ZFP462-FL-Knock-In mESCs replicate2
GSM6395504 ATACseq of Zfp462-KO_ZFP462-NT-Mid-Knock-In mESCs replicate1
GSM6395505 ATACseq of Zfp462-KO_ZFP462-NT-Mid-Knock-In mESCs replicate2
GSM5331774 QuantSeq of ESC_WT_Serum_LIF_2i_Rep1
GSM5331775 QuantSeq of ESC_WT_Serum_LIF_2i_Rep2
GSM5331776 QuantSeq of ESC_WT_Serum_LIF_Rep1
GSM5331777 QuantSeq of ESC_WT_Serum_LIF_Rep2
GSM5331778 QuantSeq of ESC_WT_4dayEB_Rep1
GSM5331779 QuantSeq of ESC_WT_4dayEB_Rep2
GSM5331780 QuantSeq of ESC_WT_2dayRA_Rep1
GSM5331781 QuantSeq of ESC_WT_2dayRA_Rep2
GSM5331782 QuantSeq of ESC_WT_4dayRA_Rep1
GSM5331783 QuantSeq of ESC_WT_4dayRA_Rep2
GSM5331784 QuantSeq of ESC_Zfp462-/- cl.1_Serum_LIF_2i_Rep1
GSM5331785 QuantSeq of ESC_Zfp462-/- cl.1_Serum_LIF_2i_Rep2
GSM5331786 QuantSeq of ESC_Zfp462-/- cl.1_Serum_LIF_Rep1
GSM5331787 QuantSeq of ESC_Zfp462-/- cl.1_Serum_LIF_Rep2
GSM5331788 QuantSeq of ESC_Zfp462-/- cl.1_4dayEB_Rep1
GSM5331789 QuantSeq of ESC_Zfp462-/- cl.1_4dayEB_Rep2
GSM5331790 QuantSeq of ESC_Zfp462-/- cl.1_2dayRA_Rep1
GSM5331791 QuantSeq of ESC_Zfp462-/- cl.1_2dayRA_Rep2
GSM5331792 QuantSeq of ESC_Zfp462-/- cl.1_4dayRA_Rep1
GSM5331793 QuantSeq of ESC_Zfp462-/- cl.1_4dayRA_Rep2
GSM5331794 QuantSeq of ESC_Zfp462-/- cl.2_Serum_LIF_2i_Rep1
GSM5331795 QuantSeq of ESC_Zfp462-/- cl.2_Serum_LIF_2i_Rep2

GSM5331796 QuantSeq of ESC_Zfp462-/- cl.2_Serum_LIF_Rep1
GSM5331797 QuantSeq of ESC_Zfp462-/- cl.2_Serum_LIF_Rep2
GSM5331798 QuantSeq of ESC_Zfp462-/- cl.2_4dayEB_Rep1
GSM5331799 QuantSeq of ESC_Zfp462-/- cl.2_4dayEB_Rep2
GSM5331800 QuantSeq of ESC_Zfp462-/- cl.2_2dayRA_Rep1
GSM5331801 QuantSeq of ESC_Zfp462-/- cl.2_2dayRA_Rep2
GSM5331802 QuantSeq of ESC_Zfp462-/- cl.2_4dayRA_Rep1
GSM5331803 QuantSeq of ESC_Zfp462-/- cl.2_4dayRA_Rep2
GSM5415700 QuantSeq of ESC_Zfp462+/Y1195*_Serum_LIF_2i_Rep1
GSM5415701 QuantSeq of ESC_Zfp462+/Y1195*_Serum_LIF_2i_Rep2
GSM5415702 QuantSeq of ESC_Zfp462+/Y1195*_Serum_LIF_Rep1
GSM5415703 QuantSeq of ESC_Zfp462+/Y1195*_Serum_LIF_Rep2
GSM5415704 QuantSeq of ESC_Zfp462+/Y1195*_4dayEB_Rep1
GSM5415705 QuantSeq of ESC_Zfp462+/Y1195*_4dayEB_Rep2
GSM5415706 QuantSeq of ESC_Zfp462+/Y1195*_2dayRA_Rep1
GSM5415707 QuantSeq of ESC_Zfp462+/Y1195*_2dayRA_Rep2
GSM5415708 QuantSeq of ESC_Zfp462+/Y1195*_4dayRA_Rep1
GSM5415709 QuantSeq of ESC_Zfp462+/Y1195*_4dayRA_Rep2
GSM5415710 QuantSeq of ESC_Zfp462+/R1257*_Serum_LIF_2i_Rep1
GSM5415711 QuantSeq of ESC_Zfp462+/R1257*_Serum_LIF_2i_Rep2
GSM5415712 QuantSeq of ESC_Zfp462+/R1257*_Serum_LIF_Rep1
GSM5415713 QuantSeq of ESC_Zfp462+/R1257*_Serum_LIF_Rep2
GSM5415714 QuantSeq of ESC_Zfp462+/R1257*_4dayEB_Rep1
GSM5415715 QuantSeq of ESC_Zfp462+/R1257*_4dayEB_Rep2
GSM5415716 QuantSeq of ESC_Zfp462+/R1257*_2dayRA_Rep1
GSM5415717 QuantSeq of ESC_Zfp462+/R1257*_2dayRA_Rep2
GSM5415718 QuantSeq of ESC_Zfp462+/R1257*_4dayRA_Rep1
GSM5415719 QuantSeq of ESC_Zfp462+/R1257*_4dayRA_Rep2

Genome browser session
(e.g. UCSC)

The following files can be uploaded (all at once or selected) to the UCSC genome browser by pasting all (mm10) URLs into "Paste URLs or data" in "add custom tracks".
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ZFP462_ChIP_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ZFP462_ChIP_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ZFP462_ChIP_in_NPCs_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ZFP462_ChIP_in_NPCs_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_ZFP462_ChIP_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_ZFP462_ChIP_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/GLP_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/GLP_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/GLP_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/GLP_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_GLP_ChIP_in_WT.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_GLP_ChIP_in_Zfp462_KO.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/HP1g_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/HP1g_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/HP1g_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/HP1g_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_antibody_ChIPs_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_antibody_ChIPs_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_antibody_ChIPs_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/Input_for_antibody_ChIPs_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K9me2_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K9me2_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K9me2_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K9me2_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K27ac_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K27ac_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K27ac_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/H3K27ac_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/WIZ_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/WIZ_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/WIZ_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/WIZ_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/NANOG_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/NANOG_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/NANOG_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/NANOG_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/SOX2_ChIP_in_WT_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/SOX2_ChIP_in_WT_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/SOX2_ChIP_in_Zfp462_KO_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/SOX2_ChIP_in_Zfp462_KO_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT_Oct4_ChIP_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT_Oct4_ChIP_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-ZFP462KO_Oct4_ChIP_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-ZFP462KO_Oct4_ChIP_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT-NR5a2-Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT-NR5a2-Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-ZFP462KO-NR5a2-Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-ZFP462KO-NR5a2-Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT-ESRRB-Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT-ESRRB-Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-ZFP462KO-ESRRB-Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-ZFP462KO-ESRRB-Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT_H3K4me1_ChIP_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT_H3K4me1_ChIP_Rep2.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT_H3K4me3_ChIP_Rep1.bw
https://data.bioinfo.vbc.ac.at/bell.grp/GSE177058/bw/ESC.TC-WT_H3K4me3_ChIP_Rep2.bw

Methodology

Replicates

All ChIP-seq experiments were performed with two biological replicates.

Sequencing depth

All ChIP-seq samples were sequenced as single-end 75mers using NextSeq550 machine.
Sample total aligned
ChIPseq for ZFP462_Rep1 38,757,876 38,336,954
ChIPseq for ZFP462_Rep2 30,512,660 30,137,545
ChIPseq for ZFP462 in NPCs_Rep1 15,465,330 15,214,983
ChIPseq for ZFP462 in NPCs_Rep2 19,391,843 19,048,300
Input_for_ZFP462 ChIPseq_Rep1 60,121,221 59,742,042
Input_for_ZFP462 ChIPseq_Rep2 66,940,416 66,509,542
ChIPseq for GLP in WT_Rep1 22,448,866 22,139,197
ChIPseq for GLP in WT_Rep2 30,813,085 30,380,176
ChIPseq for GLP in Zfp462KO_Rep1 22,568,461 22,224,665

ChIPseq for GLP in Zfp462KO_Rep2 21,520,558 21,167,539
Input_for_GLP ChIPseq_in_WT 11,738,435 11,616,023
Input_for_GLP_ChIPseq in Zfp462KO 11,566,810 11,437,504
ChIPseq for HP1gamma in WT_Rep1 45,950,905 45,526,890
ChIPseq for HP1gamma in WT_Rep2 44,036,131 43,608,196
ChIPseq for HP1gamma in Zfp462KO_Rep1 41,452,760 41,109,138
ChIPseq for HP1gamma in Zfp462KO_Rep2 38,643,736 38,194,759
Input_for_antibody_ChIPs in WT_Rep1 26,677,163 26,072,878
Input_for_antibody_ChIPs in WT_Rep2 27,194,477 26,564,716
Input_for_antibody_ChIPs in Zfp462KO Rep1 27,106,936 26,537,266
Input_for_antibody_ChIPs in Zfp462KO Rep2 32,548,958 31,762,820
ChIPseq for H3K9me2 in WT_Rep1 43,411,065 42,774,450
ChIPseq for H3K9me2 in WT_Rep2 46,911,111 45,973,624
ChIPseq for H3K9me2 in Zfp462KO_Rep1 31,304,808 30,911,127
ChIPseq for H3K9me2 in Zfp462KO_Rep2 51,447,576 50,550,225
ChIPseq for H3K27ac in WT_Rep1 41,057,915 40,047,284
ChIPseq for H3K27ac in WT_Rep2 45,835,626 44,456,662
ChIPseq for H3K27ac in Zfp462KO_Rep1 43,719,679 42,718,458
ChIPseq for H3K27ac in Zfp462KO_Rep2 34,520,448 33,551,238
ChIPseq for WIZ in WT_Rep1 9,909,557 9,568,667
ChIPseq for WIZ in WT_Rep2 23,530,380 22,706,718
ChIPseq for WIZ in Zfp462KO_Rep1 11,497,525 11,127,325
ChIPseq for WIZ in Zfp462KO_Rep2 27,636,051 26,781,881
ChIPseq for NANOG in WT_Rep1 39,100,435 37,862,147
ChIPseq for NANOG in WT_Rep2 25,097,672 24,229,000
ChIPseq for NANOG in Zfp462KO_Rep1 25,592,904 24,756,062
ChIPseq for NANOG in Zfp462KO_Rep2 26,230,123 25,327,419
ChIPseq for SOX2 in WT_Rep1 20,596,315 19,937,574
ChIPseq for SOX2 in WT_Rep2 43,234,037 41,957,729
ChIPseq for SOX2 in Zfp462KO_Rep1 29,119,458 28,348,809
ChIPseq for SOX2 in Zfp462KO_Rep2 30,105,850 29,313,901
ESC.TC-WT_Oct4_ChIP_Rep1, 36,558,316, 35,571,301
ESC.TC-WT_Oct4_ChIP_Rep2, 31,963,088, 31,043,027
ESC.TC-ZFP462KO_Oct4_ChIP_Rep1, 36,115,090, 35,014,347
ESC.TC-ZFP462KO_Oct4_ChIP_Rep2, 34,925,450, 33,884,716
ESC.TC-WT-NR5A2-Rep1, 16,555,466, 16,218,401
ESC.TC-WT-NR5A2-Rep2, 16,589,819, 16,254,709
ESC.TC-ZFP462KO-NR5A2-Rep1, 16,422,154, 16,105,026
ESC.TC-ZFP462KO-NR5A2-Rep2, 18,010,803, 17,680,508
ESC.TC-WT-ESRRB-Rep1, 16,989,439, 16,479,727
ESC.TC-WT-ESRRB-Rep2, 23,421,095, 22,743,782
ESC.TC-ZFP462KO-ESRRB-Rep1, 35,583,276, 34,337,968
ESC.TC-ZFP462KO-ESRRB-Rep2, 33,122,256, 32,018,885
ESC.TC-WT_H3K4me1_ChIP_Rep1, 25,554,549, 24,564,162
ESC.TC-WT_H3K4me1_ChIP_Rep2, 25,043,656, 23,932,391
ESC.TC-WT_H3K4me3_ChIP_Rep1, 25,799,901, 24,966,052
ESC.TC-WT_H3K4me3_ChIP_Rep2, 26,950,369, 26,118,049

Antibodies	mouse anti-H3K9me2 (ab1220/Abcam), rabbit anti-H3K27ac (ab4729/Abcam), rabbit anti-H3K4me1(ab8895/abcam), rabbit anti-H3K4me3(ab8580/abcam), rabbit anti-WIZ (NBP1-80586/Novus), goat anti-OCT4 (AF1759/R&D Systems), goat anti-SOX2(AF2018/R&D systems), rabbit anti-NANOG (ab80892/Abcam), rabbit anti-NR5A2(ABE2867/SIGMA) and mouse anti-ESRRB(PP-H6705-00/R&D Systems).
Peak calling parameters	ZFP462 peaks were called using MACS2 v2.1.1/default settings. ZFP462-ChIP and Input Sorted BAM files were used for peak calling.
Data quality	We observed good library mapping rates and agreement between replicates. Spearman correlation heatmaps comparing replicates of different ChIP-seq datasets are provided in supplementary files.
Software	ChIP-seq data were aligned using BWA MEM v0.7.17. Coverage tracks were generated using deepTools bamCoverage v3.5.0 .

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation
ES cells are trypsinized and washed prior to preparing single cell suspension for flow analysis either in tubes or 96 well format.

Instrument
FACS LSR II Fortessa/BD biosciences

Software
BD FACSDiva 8 for acquiring data. Flowjo software for visualization.

Cell population abundance
Clonal Oct4 dual reporter mESCs expressing TetR-HP1 or TetR-ZFP462FL/truncations were derived by lentivirus infection and subsequent FACS for mCherry-positive cells (~1 - 20%)

Gating strategy
mESCs were gated for shape (Area FSC-Area SSC), doublets (Area FSC-Height FSC), GFP and mCherry.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.