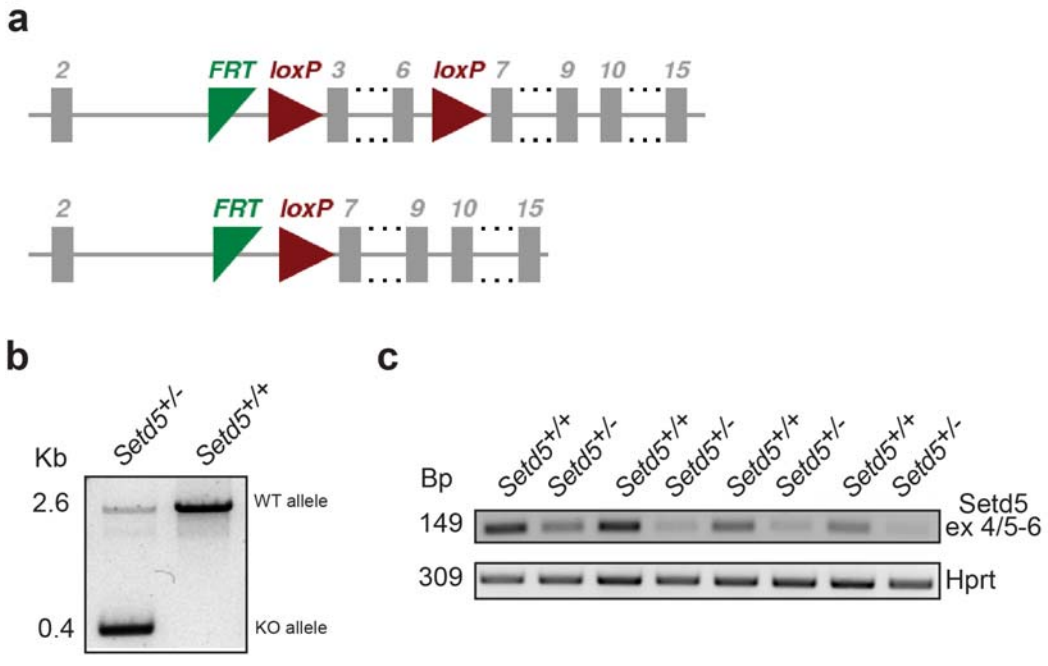
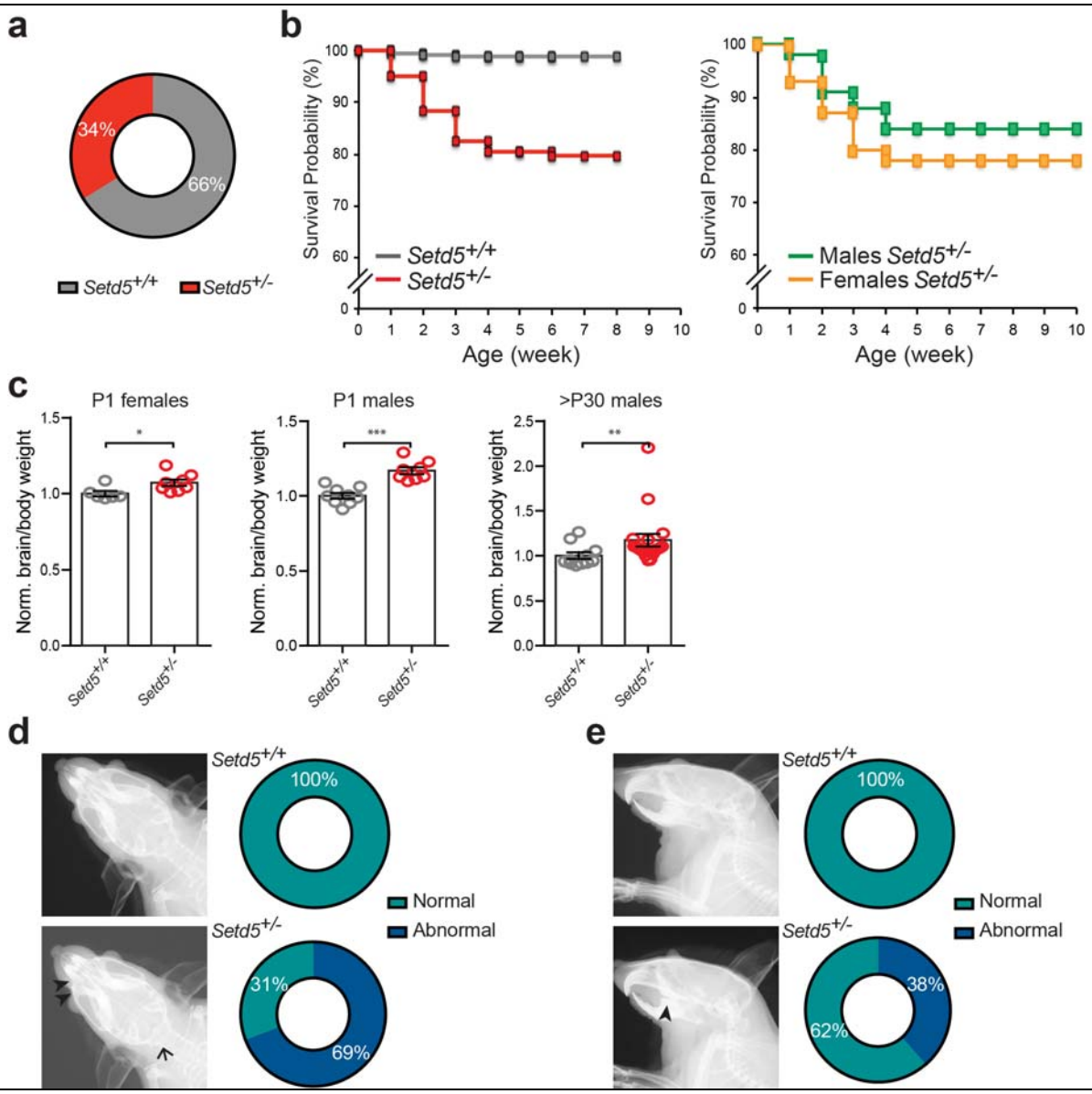


Supplementary Figure 1

***Setd5* gene deletion in mice.**

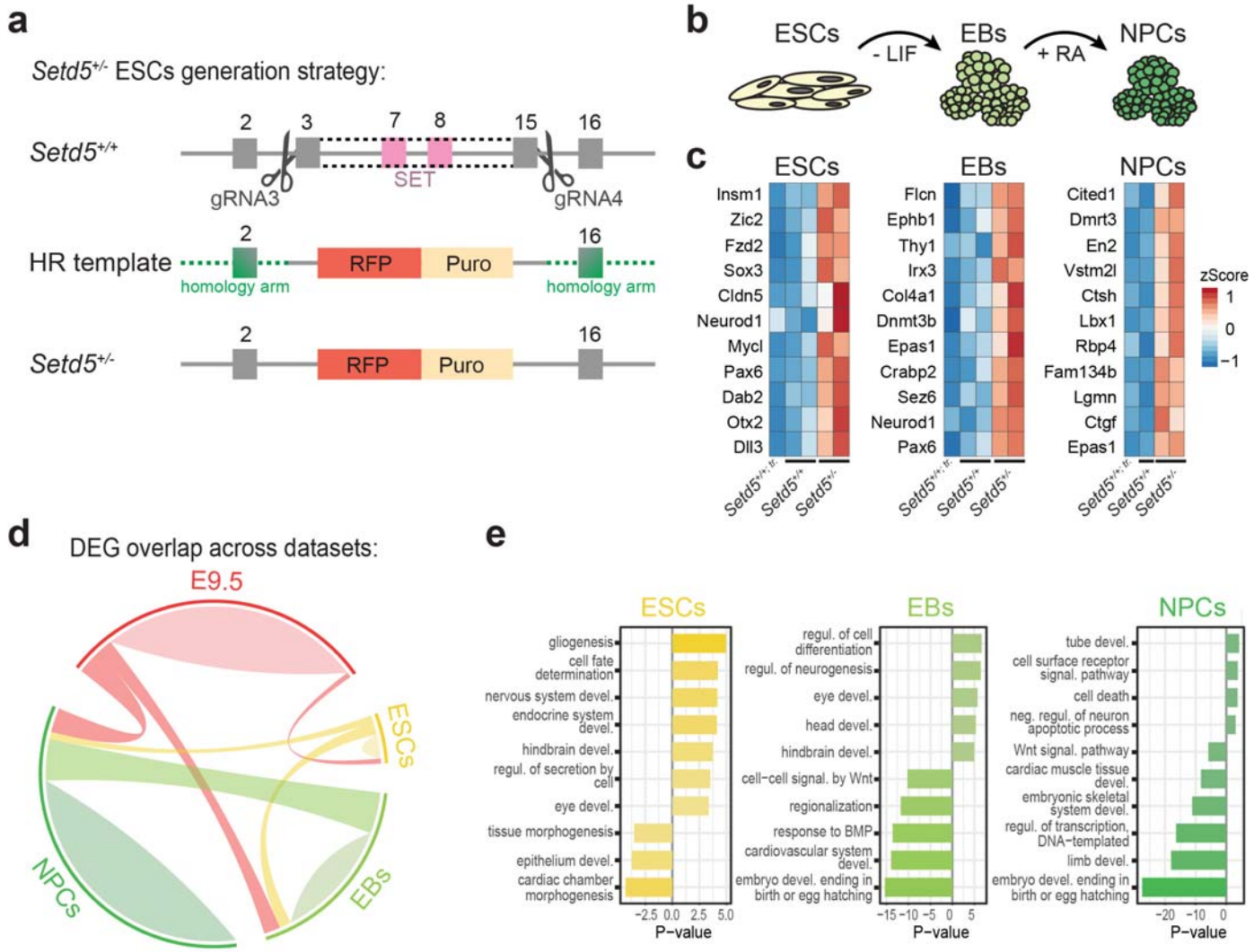
a, Schematic diagram for *Setd5*-floxed allele (top) showing exons 3-6 flanked by loxP sites and the resulting *Setd5*-knockout allele. **b,** Typical genotyping results of $Setd5^{+/-}$ and $Setd5^{+/+}$ mice; the results are shown for single experiment and replication experiments were not performed. **c,** RT-PCR showing reduction of *Setd5* gene expression in the brain of $Setd5^{+/-}$ and $Setd5^{+/+}$ mice (*n(mice)*=4 per genotype); *Hprt* was used as control. Pictures in **b** and **c** represent part of the gels; see **Supplementary Fig. 14** for full-length picture.



Supplementary Figure 2

Lower survival probability and developmental anomalies in *Setd5*$^{+/-}$ animals.

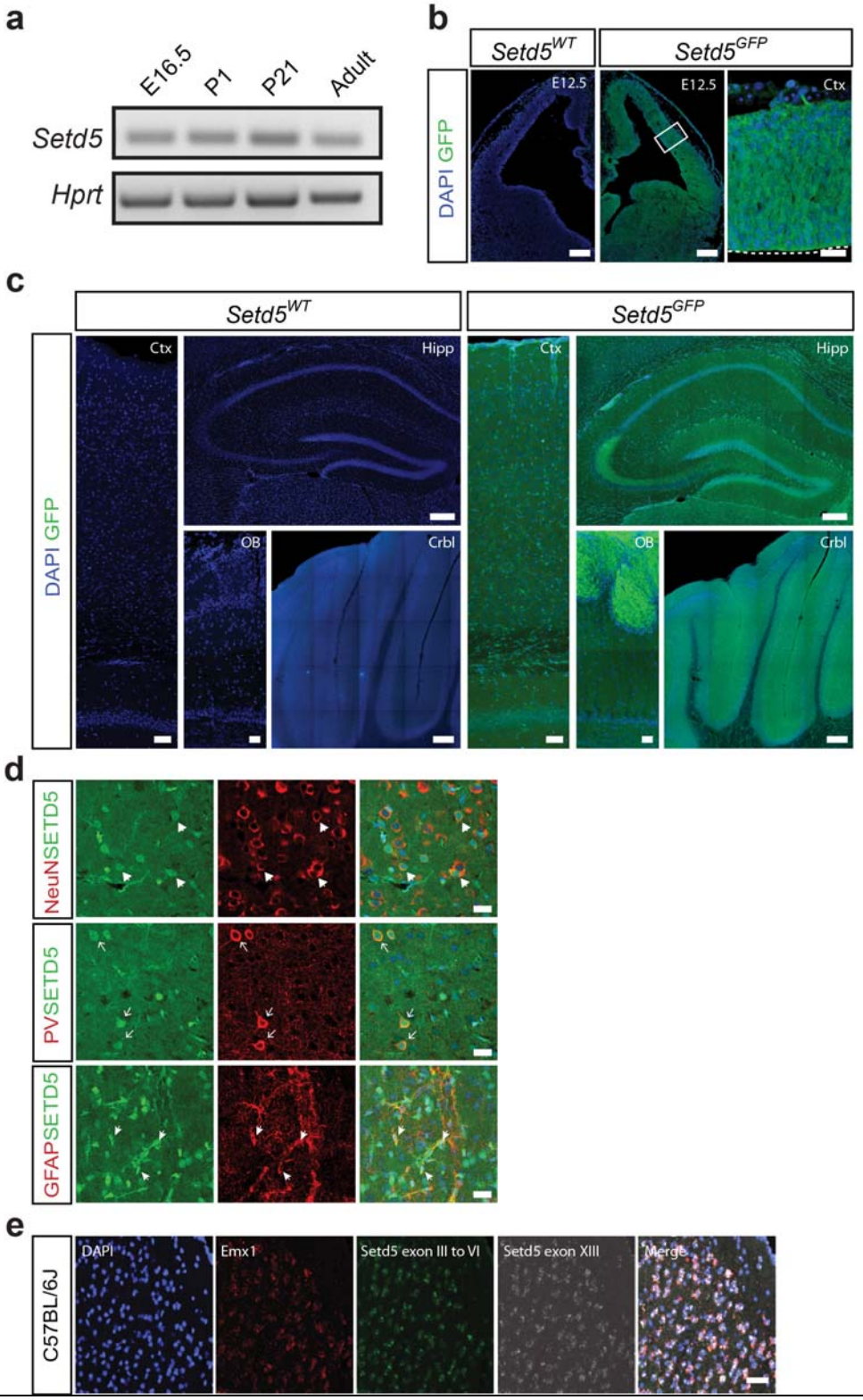
a, Doughnut chart indicating that *Setd5*$^{+/-}$ animals are born at a non-Mendelian ratio (*n*=585 mice, of which 387 *Setd5*$^{+/+}$ and 198 *Setd5*$^{+/-}$). **b,** Kaplan-Meier survival curves showing reduced survival probability of mutant animals (*n* is 387 for *Setd5*$^{+/+}$ and 198 *Setd5*$^{+/-}$), which seems to be slightly more pronounced in the females (*n* is 98 male and 100 female *Setd5*$^{+/-}$). **c,** Increased normalised brain/body weight ratio is observed in P1 male (*n*=9 controls, *n*=8 mutants) and female mutant pups (*n*=6 controls, *n*=8 mutants) as well as in adult male mutants (>P30, *n*=10-17) (values normalised to *Setd5*$^{+/+}$); *P<0.05, **P<0.01, ***P<0.001; data presented as means and SEM, as well as scatter plots (Supplementary Table 1 contains detailed statistics). **d-e,** Typical X-ray images of *Setd5*$^{+/-}$ and *Setd5*$^{+/+}$ adult animals; craniofacial maxillary/pre-maxillary abnormalities (**d**, arrowheads), vertebral fusion abnormalities (**d**, open arrow) and teeth abnormalities (**e**, arrowheads) are observed in *Setd5*$^{+/-}$ mice. The percentage of animals with each of the described phenotype is indicated as doughnut charts (*n*=13 mice of either genotype; data and pictures (as well as top eye picture in Figure 1c) taken with permission from International Mouse Phenotyping Consortium, available at http://www.mousephenotype.org/data/genes/MGI:1920145).



Supplementary Figure 3

Generation and gene expression analysis of *Setd5* mutant ESC lines.

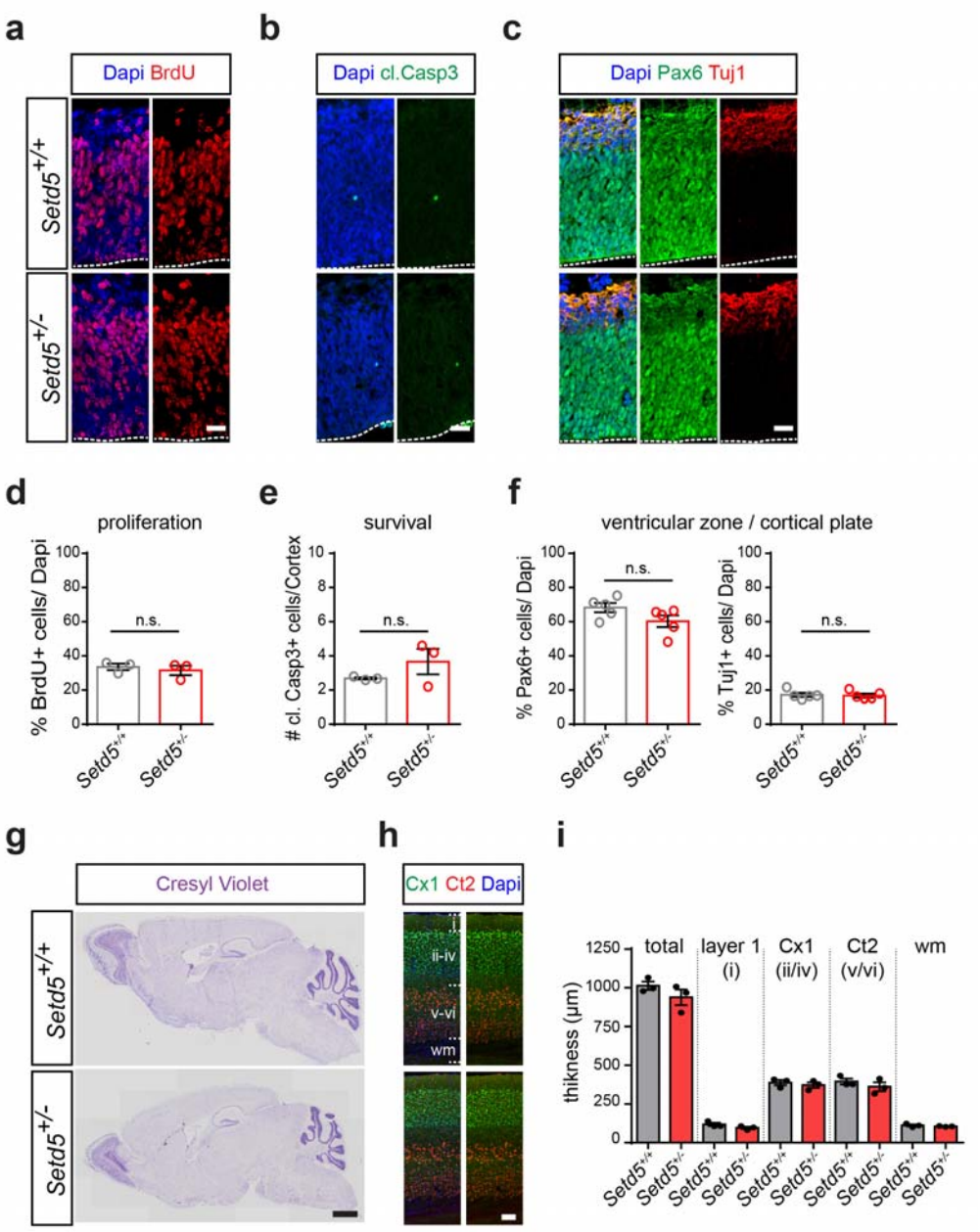
a, Schematics of CRISPR/Cas9 strategy to generate the *Setd5*$^{+/-}$ ESC line; two CRISPR guides were used: one targeting the intronic region upstream of exon 3, the other targeting the intronic region downstream of exon 15. The homologous recombination template carries a Red Fluorescent Protein/Puromycin selection cassette flanked by two 1 kb homology arms identical to the regions upstream or downstream of the CRISPR/Cas9 cut sites (green dashed lines). **b,** Schematic representation of the differentiation process from ESCs to NPCs; LIF: leukemia inhibitory factor; RA: retinoic acid; **c,** Heatmaps of RNA-seq counts (z-Score) for up-regulated neurodevelopment-associated genes in *Setd5*$^{+/-}$ at 3 stages of *in vitro* differentiation. *Setd5*$^{+/+;tr.}$, CRISPR/Cas9 control lines (tr., transfected). **d,** Chord plot illustrating the overlap between the *in vivo* and *in vitro* DEG datasets. Ribbons show overlaps between sets and ribbon width scales with the number of overlapping genes. Remaining areas represent genes unique for each set. **e,** Gene ontology (GO) terms enriched in upregulated (top, bars to the right) and downregulated (bottom, bars to the left) differentially expressed genes (adjusted p-value < 0.05) between wild type and *Setd5* heterozygous ESCs, EBs and NPCs. Bars show the negative decimal logarithm of the enrichment p-value (shown as negative for downregulated genes). (*n*=2 wildtype cell lines, 2 CRISPR control lines, and 2 mutant *Setd5*$^{+/-}$ cell lines, one-sided fisher`s exact test).



Supplementary Figure 4

***Setd5* is expressed throughout development and in multiple brain regions.**

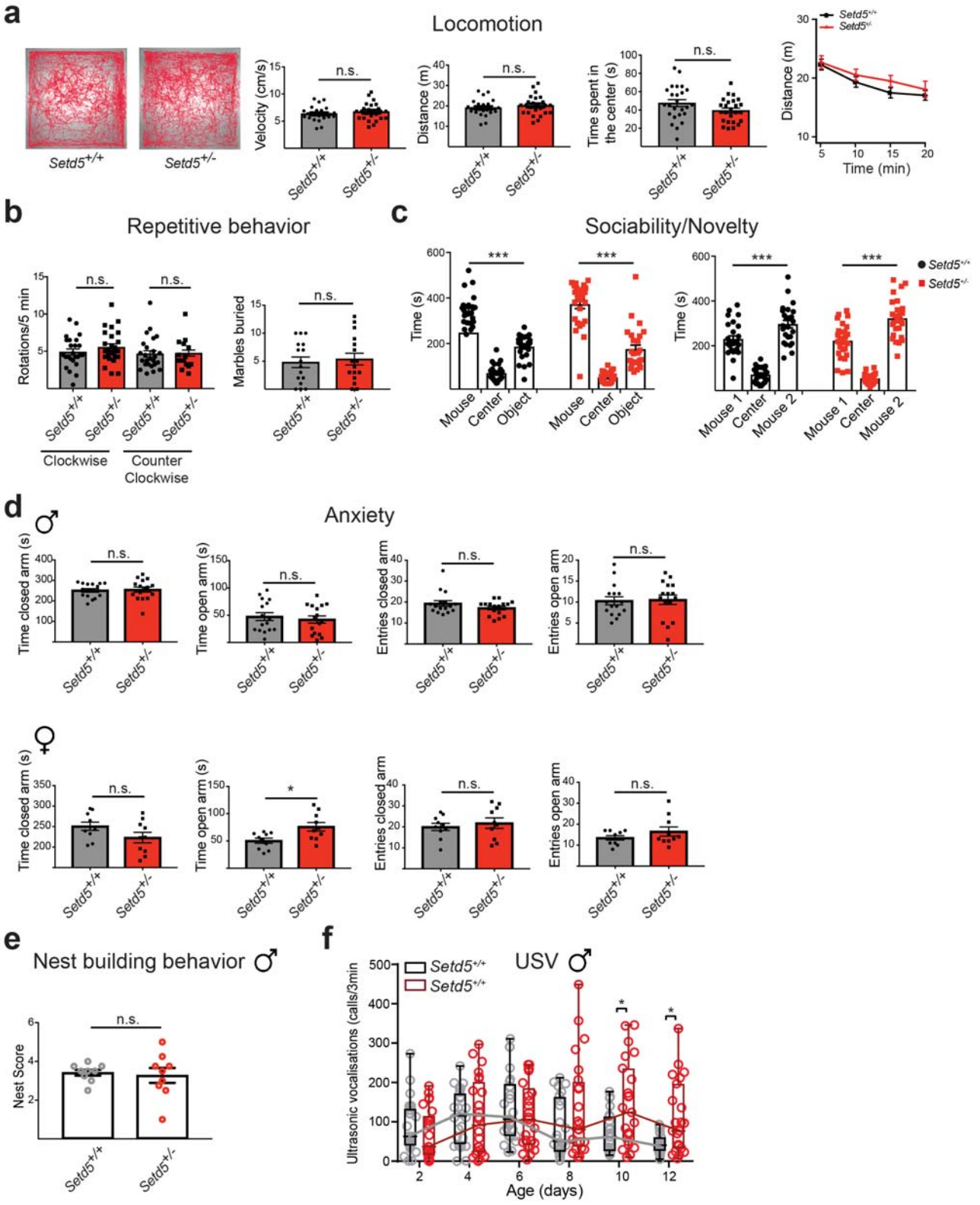
a, *Setd5* gene expression in brains of wild type animals at different developmental stages (E16.5, P1, P21, adult) assessed by RT-PCR; *Hprt* was used as control. **b,** *Setd5-GFP* reporter expression in coronal brain sections of E12.5 $Setd5^{GFP}$ embryos and control ($Setd5^{WT}$), *Setd5-GFP* is expressed throughout the developing cortex (boxed region and close-up; dashed line: ventricular wall). **c,** Representative coronal sections of different brain regions of adult control ($Setd5^{WT}$) and $Setd5^{GFP}$ animals. Broad GFP expression can be observed in several brain regions: cortex (Ctx), hippocampus (Hipp), olfactory bulb (OB) and cerebellum (Crbl). **d,** *Setd5-GFP* positive cells co-localize with NeuN-positive neurons, parvalbumin (PV)-positive interneurons and glial fibrillary acidic protein (GFAP)-positive astrocytes in the adult cortex. **e,** Fluorescent *in-situ* hybridization (*n*= one experiment) employing probes against *Setd5* exon III-VI and *Setd5* exon XIII in adult C57BL/6J sections show nuclear signal and co-localization with Emx1. DAPI nuclear staining (shown in blue) can be seen in (**b-e**). *n*=1 $Setd5^{WT}$ and *n*=3 $Setd5^{GFP}$. Scale bars: 150 µm in (**b**); 25 µm in (**b**, close-up); 50 µm in (**c**) (Ctx); 200 µm in (**c**) (Hipp); 25 µm in (**c**) (OB); 200 µm in (**c**) (Crbl); 20 µm in (**d**) (all) and 50 µm in (**e**). Experiments (**a**-**d**) were repeated at least three times, independently, with similar results, using 3 animals per genotype per developmental stage.



Supplementary Figure 5

Normal brain development and adult brain morphology in *Setd5*$^{+/-}$ mice.

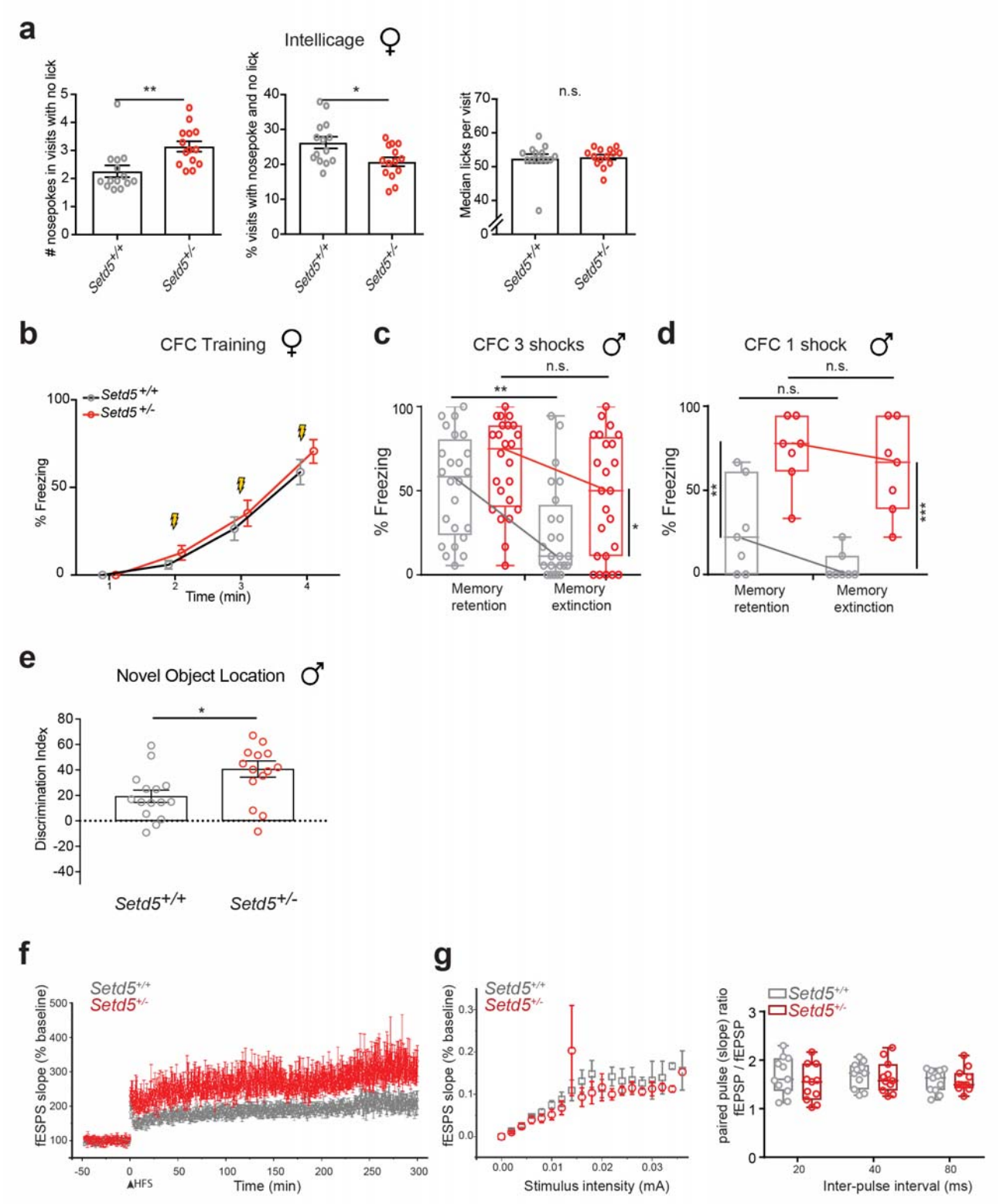
a-f, Representative images (**a**-**c**) and quantifications (**d**-**f**) of coronal forebrain sections and immunofluorescent stainings at E12.5; dashed line: ventricular wall. **a,d**, Two hour BrdU incorporation is similar in *Setd5*$^{+/-}$ and their wild type littermate embryos; **b,e**, Very low numbers of cleaved Caspase3-positive cells can be found in the developing forebrain of both genotypes (*n(*mouse embryos*)*=3 per genotype); **c,f,** The number of Pax6-positive radial glia in the ventricular zone and Tuj1-positive immature neurons in the developing cortical plate is not changed upon heterozygous loss of *Setd5* (*n(*mouse embryos*)*=5 per genotype). **g,** Nissl staining on sagittal sections reveals normal morphology of the adult *Setd5*$^{+/-}$ brain (*n*=3 animals per genotype); **h-i,** Representative images and quantification of the lamination in coronal sections of the somatosensory cortex of *Setd5*$^{+/+}$ and *Setd5*$^{+/-}$ adult brain as assessed by staining for Cux1 positive upper layer and Ctip2 positive lower layer neurons (*n*=3 animals per genotype); data presented as means and SEM, as well as scatter plots; Scale bars: 25 µm in (**a**-**c**); 1 mm in (**g**); 100 µm in (**h**), wm: white matter; n.s., not significant, unpaired, two-sided t test, exact p-values and detailed statistics provided in Supplementary Table 1;



Supplementary Figure 6

Further features of *Setd5* haploinsufficient animals.

a, Normal exploratory behaviour in the open field. Representative trajectories (**a,** left), quantifications (**a,** centre) of velocity, total distance moved, total time spent in the centre (also as measure of anxiety) as well as the distance moved in 5 min bins (**a,** right) by wild type (grey, *Setd5*$^{+/+}$) and mutant (red, *Setd5*$^{+/-}$) mice. **b,** Lack of repetitive behaviours in the open field (**b**, left) and in marble burying within 30 min (**b**, right). **c,** Likewise to the controls, *Setd5*$^{+/-}$ mice prefer to interact more with a mouse than with an object (normal sociability; left) and favour sniffing a novel mouse over a familiar one (normal preference for social novelty; right) in the three chamber sociability test (**a**-**c,** *n*=25 age and sex-matched littermate pairs of *Setd5*$^{+/+}$ and *Setd5*$^{+/-}$); **d,** Performance of male (upper panels) and female (lower panels) *Setd5*$^{+/-}$ mice in the elevated plus maze (total duration 6 min); Quantification of the total time spent in the closed and open arm as well as arm entries point out natural levels of anxiety, largely identical to the wild type littermates (*n*=16 male and 10 female age- and sex-matched littermate pairs of wild type and *Setd5* heterozygous animals); **e,** Nest scores showing normal nesting abilities in *Setd5*$^{+/-}$ males; *n*=9; **f,** Isolation-induced ultrasonic vocalisation profile of *Setd5*$^{+/+}$ and *Setd5*$^{+/-}$ male mouse pups, indicating normal number of calls within the first 6 days, followed by failure of mutants to extinguish this type of communication (*n*=23 male mouse pups per genotype). Data (**a**-**e**) presented as means and SEM and scatter plots; boxplots (**f**) indicate medians (middle line), 25%-75% interquartile range (box) and minimum and maximum data point (whiskers). ***P<0.001; *P<0.05; n.s., not significant (detailed statistics in the Supplementary Table 1); For behavioural tests, in which no inter-genotypic differences were observed, the results of both sexes are presented together, although initially analysed separately. However, for tests in which inter-genotypic differences were discovered, the results are presented for each sex separately.

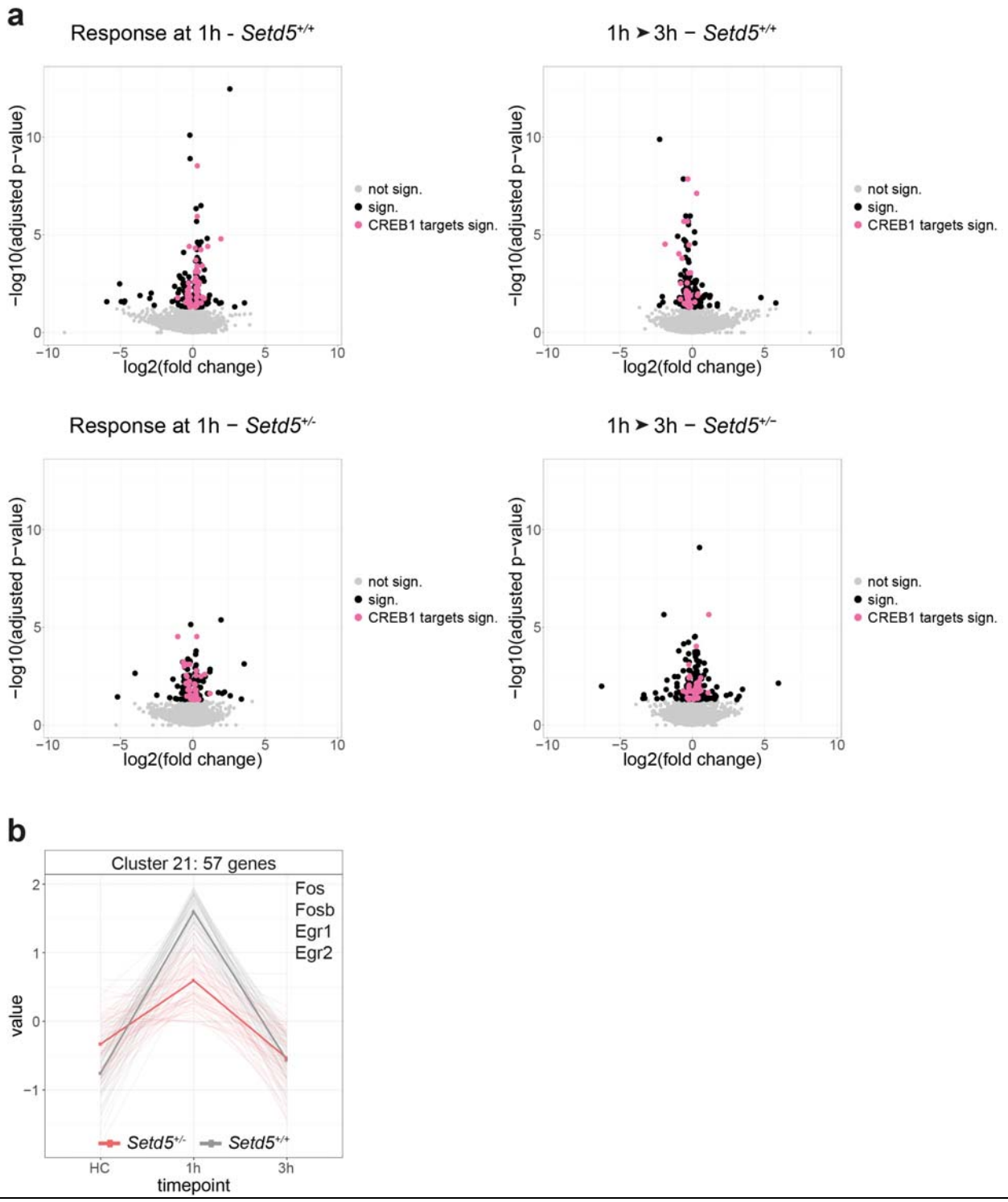


Supplementary Figure 7

Additional behavioural and electrophysiological features of *Setd5*$^{+/-}$ mice.

a, *Setd5*$^{+/-}$ female mice in the Intellicage show more nose pokes per visit (left) but perform significantly less unintentional (no licks) visits

with nose pokes (middle) while displaying similar number of licks per visit (right) as compared with the controls (means, SEM and individual data points, *n*=14 females per genotype). **b,** Similar freezing behaviour curves during fear memory acquisition (training day, 3 shocks) exclude differences in the pain/sensorial perception between the two genotypes (*n*=18 adult females per genotype). **c,** Contextual fear conditioning (CFC) in male mice (strong training, 3 shocks, 5 min total exposure to context, *n*=15 males per genotype) indicates inability of mutants to extinguish the acquired fear memory when the context becomes a safe environment. Boxplots (**c**, **d**) indicate medians (middle line), 25%-75% interquartile range (box) and minimum and maximum data point (whiskers). **d,** Weak training contextual fear conditioning in male mice (1 shock, 150 s total exposure to context, *n*=7 males per genotype) indicates that mutants learn the context-shock association after sub-threshold training, and further fail to extinguish it (10 min extinction training in both weak and strong CFC); **e,** Subthreshold training (3 min) in novel object location memory test induces abnormal memory retention in *Setd5* heterozygous males (*n*=13 males per genotype). Data is presented either as means and SEM, as well as scatter plot; *P<0.05; **P<0.01; ***P<0.001; n.s., not significant; **f,** Averaged traces (mean ± SEM) indicating the dynamics of fEPSP slope before and up to 5h after LTP induction by high frequency stimulation (HFS, min 0): enhanced late-LTP in the mutants; *n*=6 (5h), 8 (4h), 10 (3h), 10 (1h) slices for each genotype; one slice per mouse; males, P21-24. **g,** Similar averaged input-output curves (8 dorsal hippocampus slices from 8 animals per genotype, left panel) and likewise paired-pulse ratios (11 dorsal hippocampus slices from 3 animals per genotype), indicating no difference in the presynaptic function at the hippocampal CA3 to CA1 synapses of P21-24 male *Setd5*$^{+/+}$ and *Setd5*$^{+/-}$ mice. For detailed statistics please refer to Supplementary Table 1.

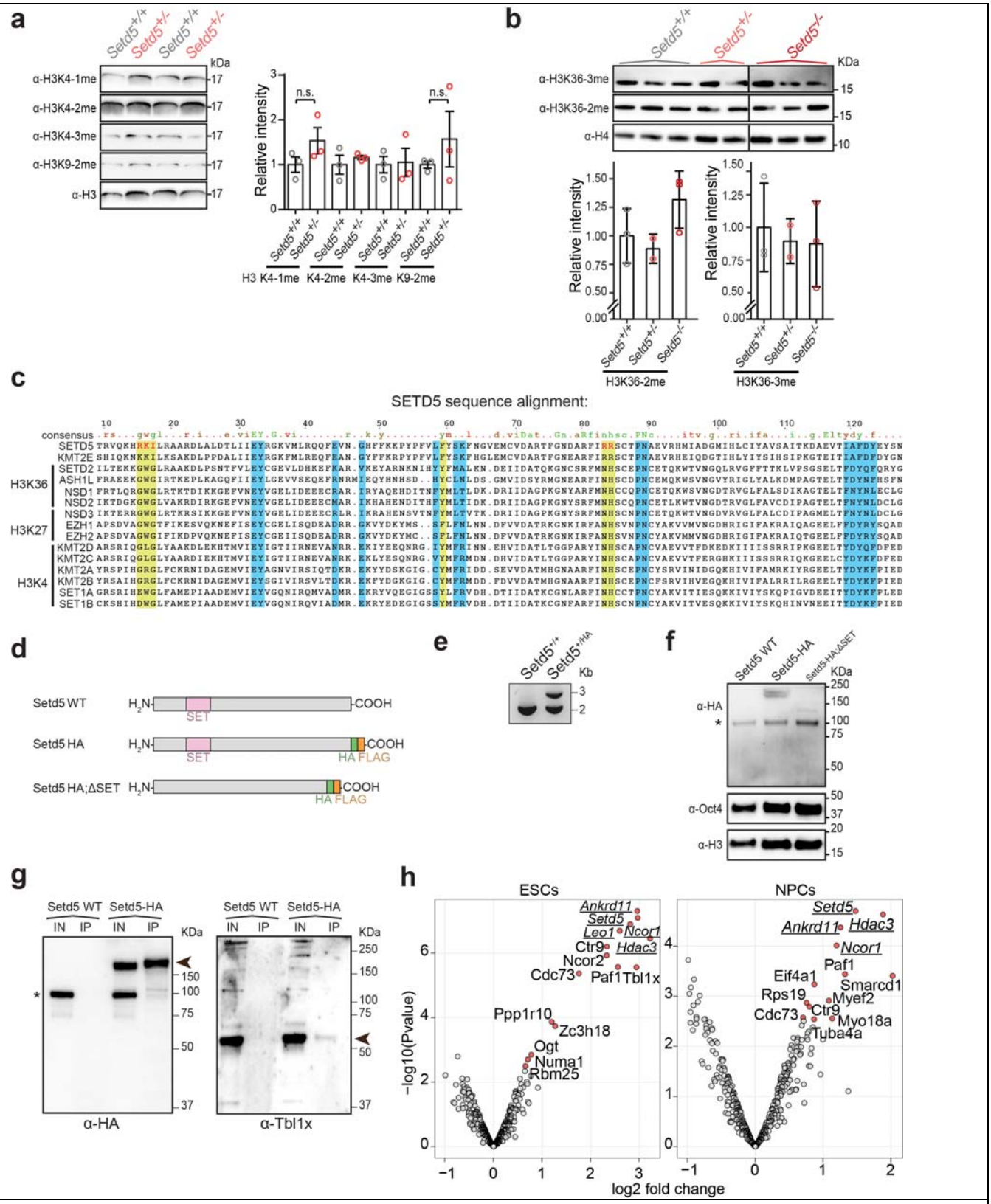


Supplementary Figure 8

Transcriptional response to CFC in *Setd5*$^{+/+}$ and *Setd5*$^{+/-}$ mice.

a, Volcano plots highlighting Creb1 target genes in the gene expression changes 1h or 3h after conditioning in wild type and mutant animals. The y-axis shows the negative decimal logarithm of the adjusted p-value and the binary logarithm of the fold change is shown on the x-axis; pink, differentially expressed Creb1 target genes; black, other differentially expressed genes; grey, non-significant genes

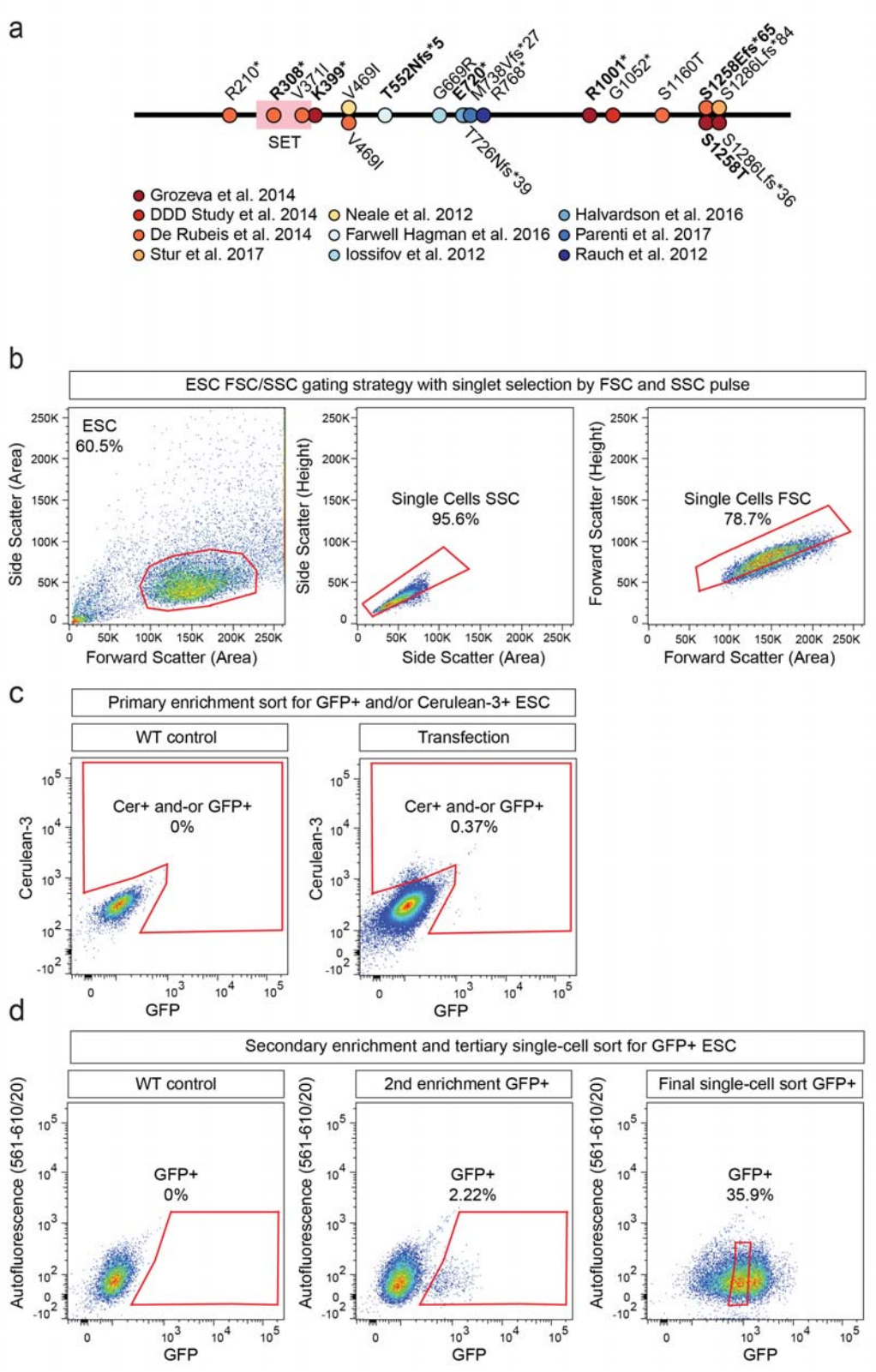
(*n(mice)*=4 control and 5 mutant animals, EdgeR, likelihood ratio test). **b,** Example of a DEG cluster showing the same CFC response in *Setd5*$^{+/+}$ and *Setd5*$^{+/-}$ mice (see Methods). Genes listed in the top right corner are early response genes contained in this cluster. Thick lines, average of all genes per cluster; thin lines, single genes; red, *Setd5*$^{+/-}$; grey, *Setd5*$^{+/+}$; HC, *n*=5; 1h, *n*=4; 3h, *n*=5 per genotype respectively.



Supplementary Figure 9

SETD5 lacks methyltransferase activity.

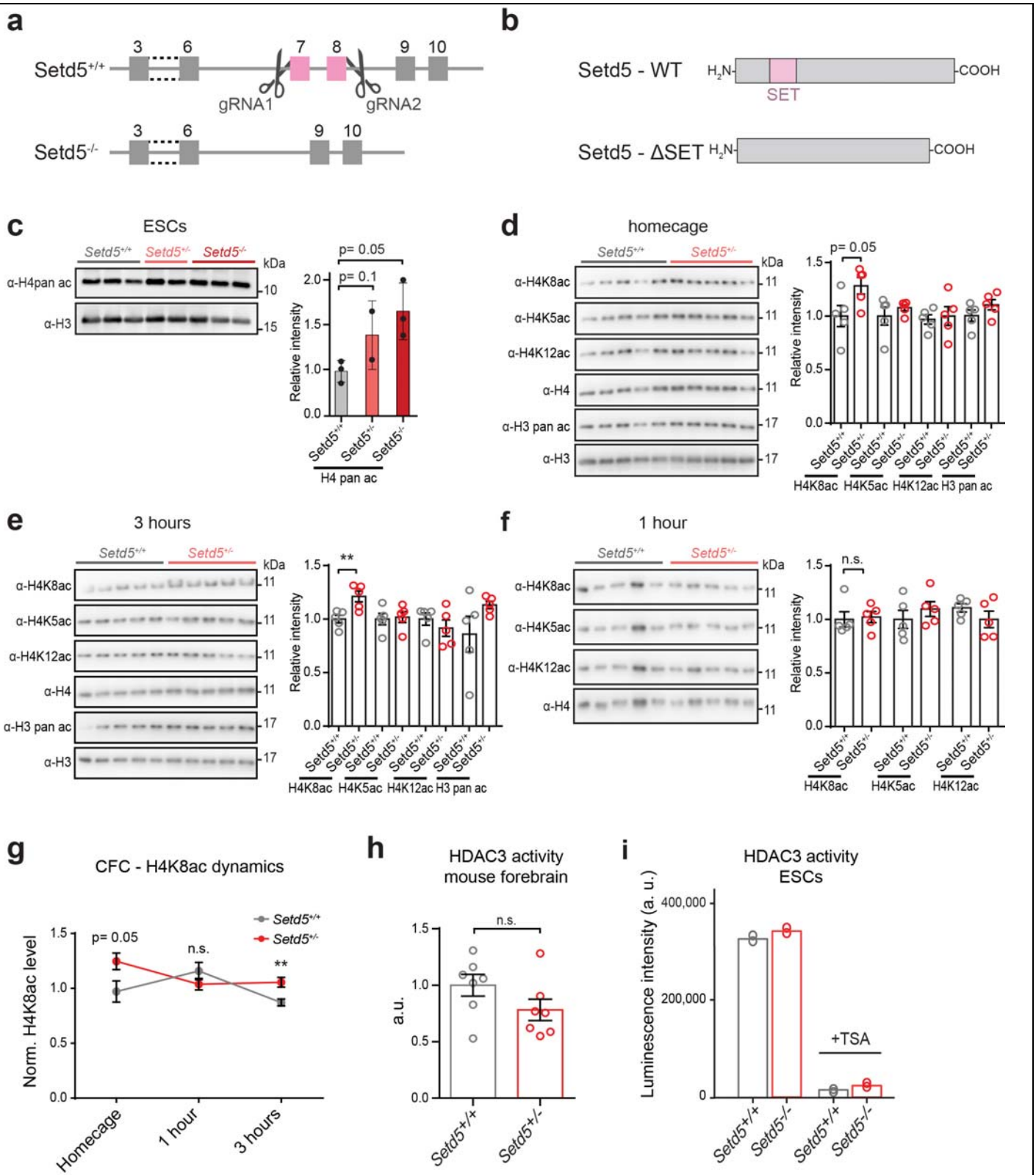
a-**b,** Immunoblots and quantifications evaluating histone H3 methylation states in *Setd5* mutant hippocampal and ESC samples. (**a**) *n(*mice*)*=5, *N*(independent experiments)=3 (**b**) *n(*independent cell lines *Setd5*$^{+/-}$ and *Setd5*$^{-/-}$*)*=2 and 3 respectively, *N(*independent experiments*)*=3. No difference in the levels of H3K4-1me, H3K4-2me, H3K4-3me. H3K9-2me, H3K36-3me or H3K36-2me was observed. Data presented as means and SEM (**a**) or SD (detailed statistics in the Supplementary Table 1) (**b**). **c,** Sequence alignment of human SET domain containing proteins. SETD5 and its closest homolog KMT2E (MLL5) are aligned to 13 characterized histone methyltransferase enzymes. Colour scheme: yellow, amino acids involved in SAM binding; blue, amino acids involved in histone H3 recognition. Amino acids in red are the 5 residues in SETD5 that differ the most from the consensus SAM binding motif. The specific histone H3 lysine residues methylated by the characterised SET domains are indicated on the left. **d,** Schematics of the endogenously C-terminus tagged Setd5. **e,** PCR genotyping of a successful insertion of the FLAG-HA tag construct (*Setd5*$^{+/+;HA}$) compared to a *Setd5*$^{+/+}$ control line; FLAG-HA was inserted only on one allele. Experiments were repeated three times, with similar results. **f,** Western blot results showing the successful tagging of endogenous Setd5. To show that our CRISPR/Cas9 strategy was capable of knocking out *Setd5*, the two single RNA guides used to generate *Setd5*$^{-/-}$ (Supplementary Fig. 11a-b) were used on a *Setd5*$^{+/HA}$ background line and the western blots show complete loss of the Setd5-HA band. Experiments were repeated four times, with similar results. **g,** Western blot of Setd5-HA immunoprecipitation in non-tagged *Setd5*$^{+/+}$ and tagged *Setd5*$^{+/HA}$. Membranes blotted for Setd5-HA (left), Tbl1x (right). Arrowheads: Setd5-HA and Tbl1x respectively; asterisk: non-specific band; IN, input: nuclear extract before immunoprecipitation; IP, immunoprecipitation eluate with triple-FLAG peptide; 0.5% of IN and 25% of IP eluate were loaded. Experiment was performed once. **h**, Volcano plot of proteins identified by quantitative TMT mass spectrometry from Setd5 immunoprecipitation in ESCs and NPCs. Significant proteins (adjusted P value < 0.05 and $\log_2$ Fold Change > 0) in red, not significant proteins in grey. Proteins associated with ASD in underlined italics. (*n*=2 wild type and 2 Setd5 FLAG/HA knock in line, limma statistical analysis).



Supplementary Figure 10

Human SETD5 rescue experiment.

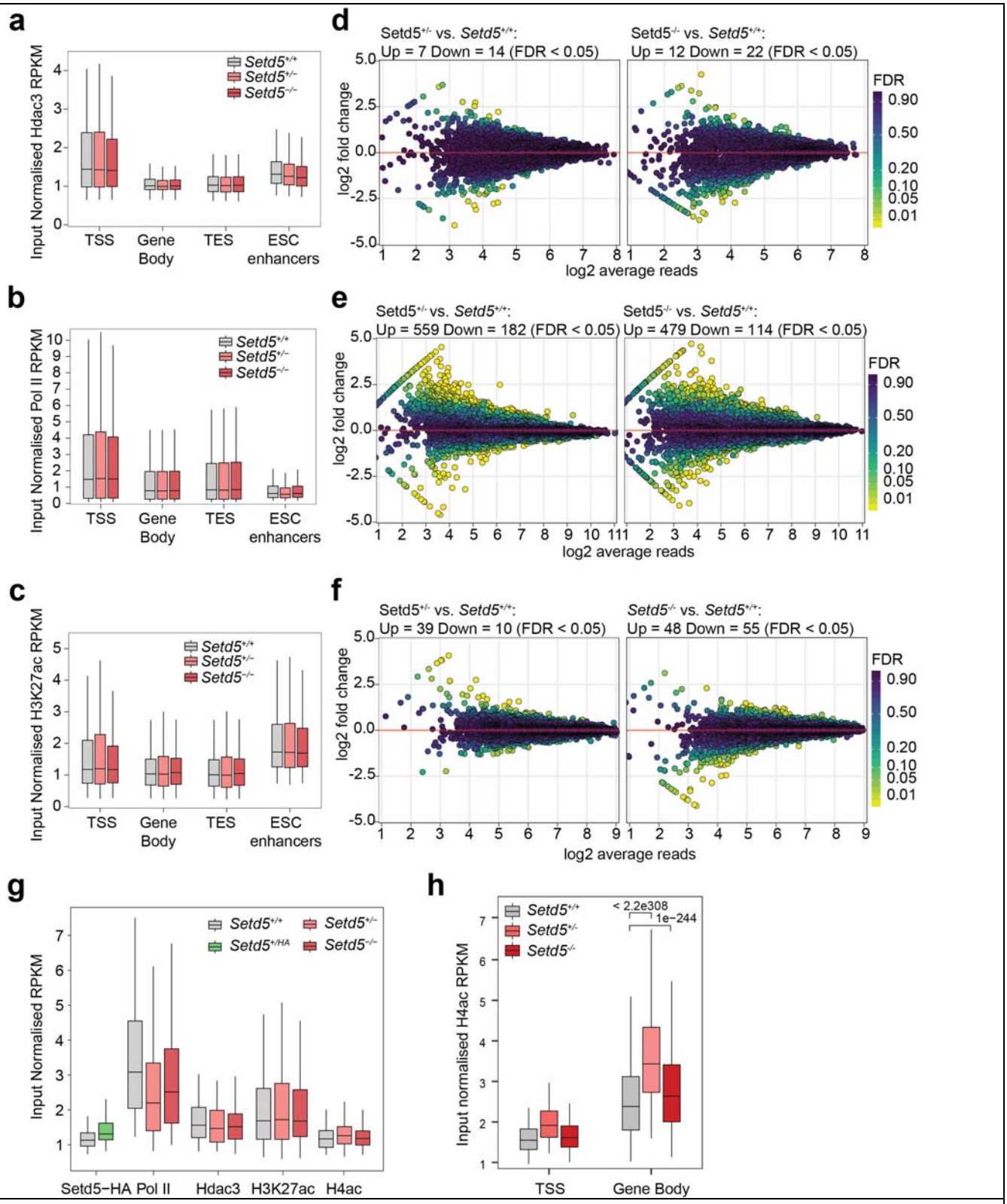
a, Schematics of the localization of *de novo SETD5* mutations in ID and ASD patients; each mutation is colour-coded based on the study reporting the *de novo* mutation; mutations used in this study are highlighted in bold. **b**-**d,** Representative example of the bulk enrichment sort and single cell sorting strategy. Shape analysis scatter plots (**b**). To exclude doublets, live cells are gated on forward scatter/ side scatter (FSC/SSC, left) followed by SSC (middle) and FSC (right). Scatter plots (**c**) of primary bulk enrichment for Setd5-P2A-GFP and CRISPR/Cas9-T2A-Cerulean3 positive cells (right) compared to WT non-transfected control (left). Scatter plots (**d**) showing enrichment for Setd5-P2A-GFP cells over the course of the secondary bulk enrichment sort (middle) and tertiary single cell sort (right). Auto-fluorescence was measured to ensure GFP positive cells.



Supplementary Figure 11

Increased H4 acetylation in *Setd5* mutants.

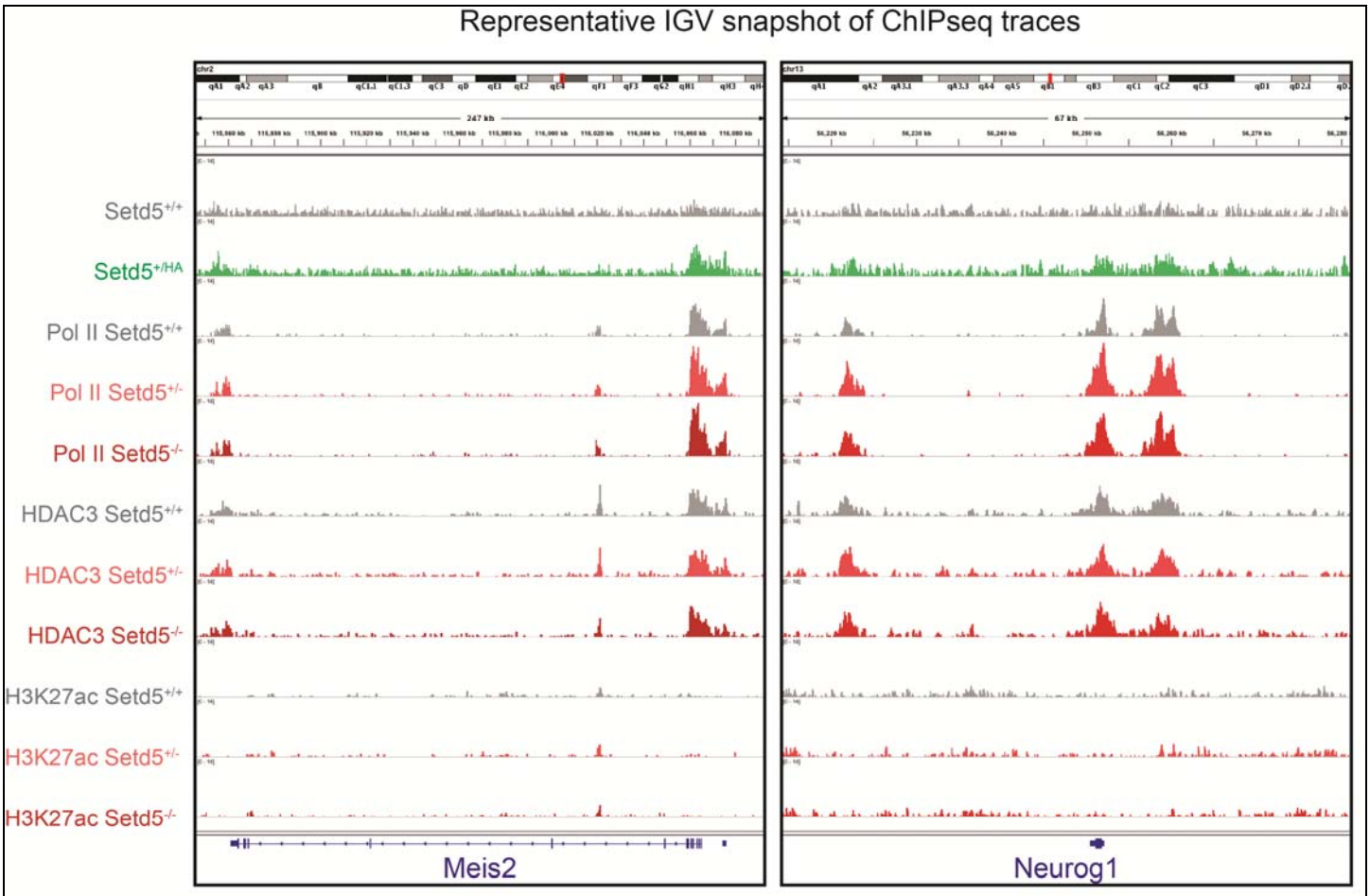
a, Schematics of the CRISPR/Cas9 strategy to generate the ΔSET *Setd5*$^{+/-}$ *and* *Setd5*$^{-/-}$ ESC line; two CRISPR guides were used: one targeting the intronic region upstream of exon 7, the other targeting the intronic region downstream of exon 8. **b,** Schematics of WT and SET domain deleted Setd5 proteins. **c,** Western blots and quantification of histone post-translational modification in *Setd5*$^{+/+}$, *Setd5*$^{+/-}$ and *Setd5*$^{-/-}$ ESCs, showing hyperacetylation of histone H4 in *Setd5*$^{-/-}$; *n(*cell lines)= 3 for *Setd5*$^{+/+}$ and *Setd5*$^{-/-}$ and *n*=2 for *Setd5*$^{+/-}$; **d**-**f,** Western blot images and quantifications showing H4K8, H4K5, H4K12 and H3 pan acetylation in *Setd5*$^{+/-}$ CA hippocampal regions dissected from homecage animals (**d**), 3 hours after CFC training (**e**) and 1 hour after CFC (**f**). The band intensity of each acetylated protein residue was normalized to the corresponding total histone level (*n(*mice*)*=5 per genotype, *N(*independent experiments*)*=3*,* per time point, respectively). **g,** H4K8ac dynamics in CA hippocampal samples obtained from homecage and CFC-trained mice showing H4K8 acetylation is significantly higher in samples from homecage (p-value= 0.054) and 3 hours post-training (p-value= 0.009) animals (levels normalized to WT H4K8ac; *n(*mice*)*=5 per genotype, *n*(independent experiments*)*=3*,* per timepoint respectively; (two-sided t-test for each timepoint) **h,** Hdac3 activity assay in brain samples from *naïve Setd5*$^{+/+}$ and *Setd5*$^{+/-}$ mice (*n*=7). **i,** Bar plots (means) of de-acetylation activity assay using samples from Hdac3 IP in *Setd5*$^{+/+}$ and *Setd5*$^{-/-}$ ESCs. Dots represent individual technical replicates; *n(independent enzymatic reactions)*=2 in DMSO and 3 in presence of Trichostatin A (TSA), broad HDAC inhibitor, used as control to show signal specificity towards de-acetylation, data means are shown). Data in **c**-**h** presented as means and SEM, **P<0.01; n.s., not significant; detailed statistics in the Supplementary Table 1.



Supplementary Figure 12

Chip-seq binding profile and differential binding analysis at TSS of Hdac3, Pol II and H3K27ac.

a-**c,** Boxplots of input normalised ChIP-seq in *Setd5*$^{+/+}$, *Setd5*$^{+/-}$, and *Setd5*$^{-/-}$. Signal for 4 different genomic regions: Transcription Start Sites (TSS, *n*=34159) Gene bodies (*n*=34151), transcription end sites (TES, *n*=34159), intergenic ESCs specific enhancers (*n*=4099) for Hdac3 (**a**), Pol II (**b**), and H3K27ac (**c**). **d**-**f,** MA plot of differential binding analysis for *Setd5*$^{+/-}$ (left) and *Setd5*$^{-/-}$ (right) compared to *Setd5*$^{+/+}$ at TSS using peaks of Hdac3 (**d**), Pol II (**e**), and H3K27ac (**f**). EdgeR likelihood ratio test. **g,** Boxplots of input normalised ChIP-seq RPKM in *Setd5*$^{+/+}$, *Setd5*$^{+/HA}$, *Setd5*$^{+/-}$, and *Setd5*$^{-/-}$. Signals of Setd5-HA, Pol II, Hdac3, H3K27ac and H4 pan acetylation for 182 genes that show significant (FDR < 0.05) decrease RNA-Pol II binding in *Setd5*$^{+/-}$ at TSS. The presence of H3K27ac signal indicates expressed genes. **h,** Boxplots of ChIP-seq RPKM signal in *Setd5*$^{+/+}$, *Setd5*$^{+/-}$, and *Setd5*$^{-/-}$ for 2 different genomic features: TSS having at least one H4 pan acetylation peak (*n*=3018) and H4 pan acetylation peaks within gene bodies (*n*=43448 peaks in 9088 gene bodies). TSS is defined as +/- 1.25 kb around gene start position; one-sided Mann–Whitney *U* test. Boxplots in **a**, **b, c, g,** and **h** represent mean, normalized by input Reads Per Kilobase of genomic regions, per Million mapped reads (RPKM) values of 2 biological replicates from two independent cell lines. Input levels are equal to 1. Outliers, less than 5 or more than 95 percentile of mean RPKM, were removed. The upper, centre, and lower line of the boxplot indicates 75%, 50%, and 25% quantile, respectively. Whiskers extend to the most extreme data point within 1.5-times the interquartile range.



Supplementary Figure 13

Complete Chip-seq tracks of *Meis2* and *Neurog1*.

ChIP-seq tracks of Setd5-HA in non-tagged (*Setd5*$^{+/+}$) and HA-tagged (*Setd5*$^{+/HA}$) ESCs and ChIP-seq tracks of Pol II, Hdac3, and H3K27ac in *Setd5*$^{+/+}$, *Setd5*$^{+/-}$, *Setd5*$^{-/-}$ ESCs at the *Meis2* and *Neurog1* gene locus. Presented signals are input-subtracted, merged replicates, Reads per Genomic Content (RPGC). (*n*=2 independent ChIP biological replicates and *n*=3 independent input biological replicates for Pol II, Hdac3, and H3K27ac ChIP-seq and *n*=1 independent input sample for Setd5-HA ChIP-seq with similar results).

Main Figures

Figure 5b

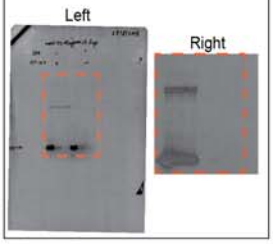


Figure 5c

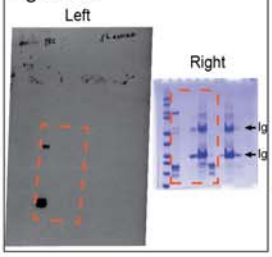


Figure 6b

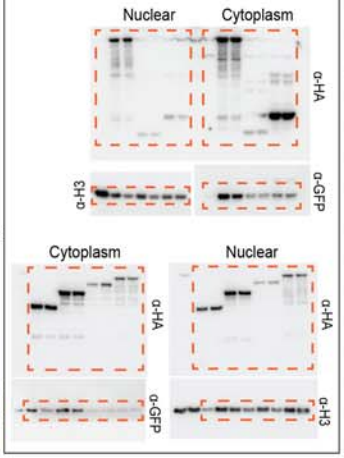
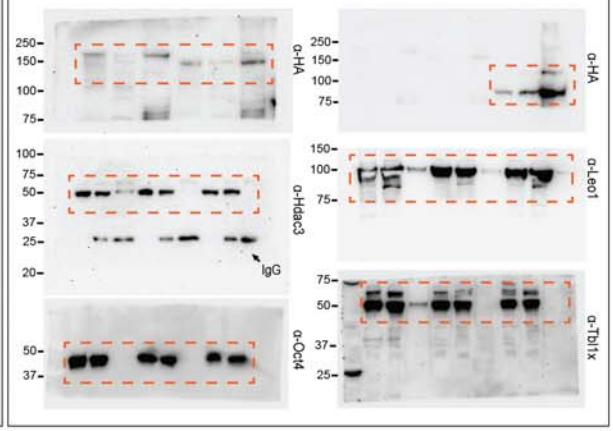
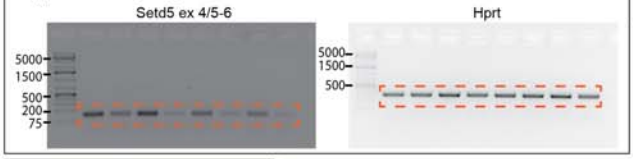


Figure 6c

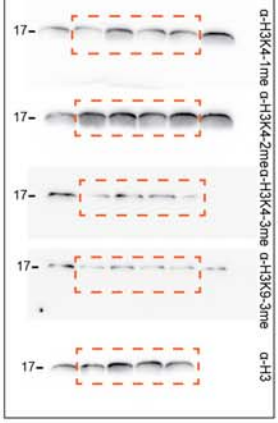


Supplementary Figures

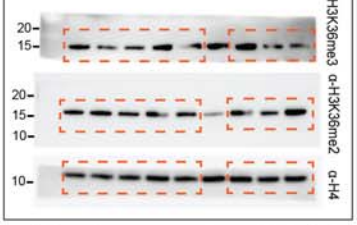
Supplementary Figure 1c



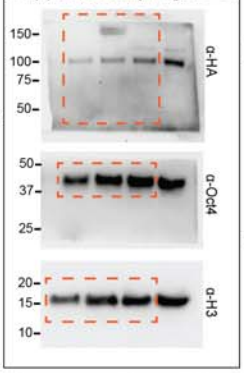
Supplementary Figure 9a



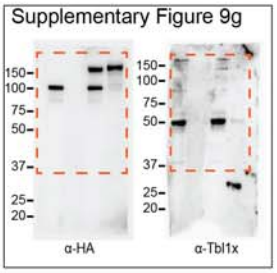
Supplementary Figure 9b



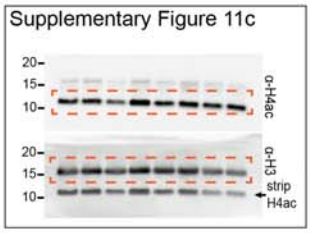
Supplementary Figure 9f



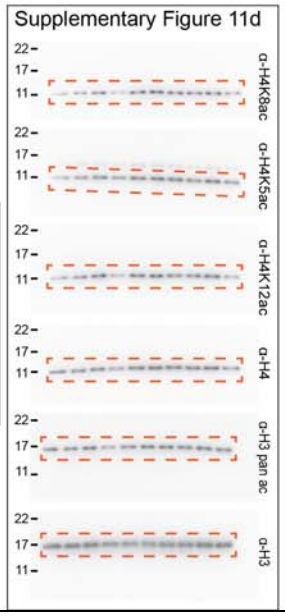
Supplementary Figure 9g



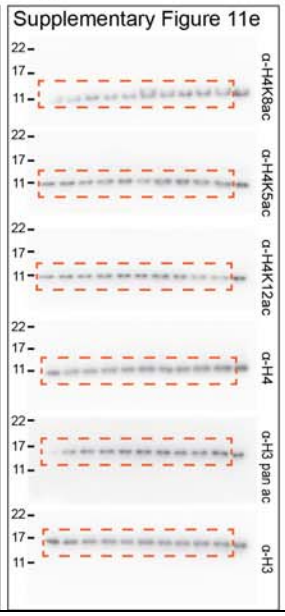
Supplementary Figure 11c



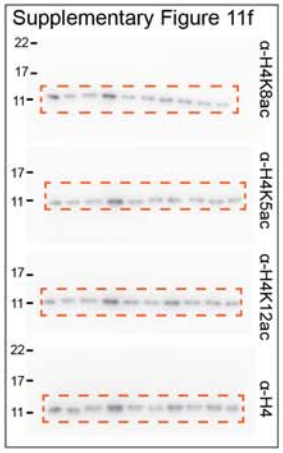
Supplementary Figure 11d



Supplementary Figure 11e



Supplementary Figure 11f



Supplementary Figure 14
Full scans of films, gels and western blots presented in this manuscript.
Dashed orange boxes indicate radioactive film-, gel- and western blot-fragments presented in the respective figure panels.