LETTER

https://doi.org/10.1038/s41586-019-1024-7

Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3

Lorna A. Farrelly[1], Robert E. Thompson[2], Shuai Zhao[3,4], Ashley E. Lepack[1], Yang Lyu[1], Natarajan V. Bhanu[5], Baichao Zhang[3,4], Yong-Hwee E. Loh[1], Aarthi Ramakrishnan[1], Krishna C. Vadodaria[6], Kelly J. Heard[6], Galina Erikson[6], Tomoyoshi Nakadai[7], Ryan M. Bastle[1], Bradley J. Lukasak[2], Henry Zebroski III[8], Natalia Alenina[9], Michael Bader[9], Olivier Berton[1], Robert G. Roeder[7], Henrik Molina[8], Fred H. Gage[6], Li Shen[1], Benjamin A. Garcia[5], Haitao Li[3,4], Tom W. Muir[2] & Ian Maze[1,10]*

Chemical modifications of histones can mediate diverse DNA-templated processes, including gene transcription[1–3]. Here we provide evidence for a class of histone post-translational modification, serotonylation of glutamine, which occurs at position 5 (Q5ser) on histone H3 in organisms that produce serotonin (also known as 5-hydroxytryptamine (5-HT)). We demonstrate that tissue transglutaminase 2 can serotonylate histone H3 tri-methylated lysine 4 (H3K4me3)-marked nucleosomes, resulting in the presence of combinatorial H3K4me3Q5ser in vivo. H3K4me3Q5ser displays a ubiquitous pattern of tissue expression in mammals, with enrichment observed in brain and gut, two organ systems responsible for the bulk of 5-HT production. Genome-wide analyses of human serotonergic neurons, developing mouse brain and cultured serotonergic cells indicate that H3K4me3Q5ser nucleosomes are enriched in euchromatin, are sensitive to cellular differentiation and correlate with permissive gene expression, phenomena that are linked to the potentiation of TFIID[4–6] interactions with H3K4me3. Cells that ectopically express a H3 mutant that cannot be serotonylated display significantly altered expression of H3K4me3Q5ser-target loci, which leads to deficits in differentiation. Taken together, these data identify a direct role for 5-HT, independent from its contributions to neurotransmission and cellular signalling, in the mediation of permissive gene expression.

Serotonergic signalling is critical to an array of biological processes ranging from neurotransmission to gastrointestinal motility and hormone release. Although vesicular packaging of 5-HT is essential for many of its functions, a few studies have demonstrated the presence of extravesicular monoamines in the nucleus and soma of monoamine-producing cells[7,8]. Whereas the presence of extravesicular 5-HT has been suggested to be important during periods of activity-dependent somatic release[7], it has remained unclear whether extravesicular 5-HT might have additional roles within the nucleus.

Serotonin has been shown to form covalent bonds with cytosolic proteins via transamidation by transglutaminase 2 (TGM2), a modification that has been suggested to alter signalling properties of monoaminylated substrates[9–11]. We therefore investigated whether nuclear proteins may be similarly modified. Given that TGM2 may interact with, or even crosslink, histone proteins in vitro[12–14], we began by investigating whether histones might serve as endogenous substrates for serotonylation. We used a bioorthogonal metabolic-labelling approach[15], based on propargylated serotonin (5-PT) that can be delivered to living cells (Extended Data Fig. 1a, b), followed by protein labelling and immunoprecipitation of serotonylated substrates (Fig. 1a, left). Western blotting revealed H3—but not H4, H2A or H2B—as a putative endogenous substrate for serotonylation (Fig. 1a, right and Extended Data Fig. 1c, d).

To assess whether TGM2, a ubiquitously expressed and functionally diverse protein[16] (Extended Data Fig. 2a), is critical for H3 serotonylation in human cells, we performed bioorthogonal-labelling-based immunoprecipitations in the presence or absence of cystamine dihydrochloride, which is an inhibitor of transglutaminase activity. Treatment with cystamine eliminated nearly all of the H3-serotonyl signal (Extended Data Fig. 2b). Whereas TGM2 is likely to be necessary for H3 serotonylation, many transglutaminases exist that may affect deposition of the mark. Therefore, we examined whether TGM2 is sufficient to catalyse H3 serotonylation. We used an alternative human tissue-culture system, HEK 293T cells, which—unlike HeLa cells—do not express TGM2 (Extended Data Fig. 2c). Using a genetic-rescue-like approach, we confirmed that exogenous expression of active nuclear TGM2 in these cells is sufficient to promote deposition of the mark (Extended Data Fig. 2d).

To assess whether H3 is a direct substrate for serotonylation, we performed recombinant TGM2 enzymatic assays in vitro on histone and non-histone substrates (for example, fibrinogen, a protein previously demonstrated to be serotonylated[17]). Monoaminylation assays were performed using the fluorescent monoamine analogue, monodansylcadaverine (MDC), in the presence or absence of cystamine or excess 5-HT. Whereas both fibrinogen and H3 display TGM2-dependent transamidation of MDC—signals that are attenuated by cystamine application and donor competition with 5-HT—no signal was observed for H4 (Fig. 1b). Subsequent radioactivity-based serotonylation assays were performed, and revealed consistent outcomes (Extended Data Fig. 2e).

To identify the site(s) of serotonylation on H3, we next performed targeted liquid chromatography–tandem mass spectrometry (LC–MS/MS) following in vitro TGM2 assays with 5-HT. Peptide tandem mass spectrometry analyses (Extended Data Fig. 2f) revealed Gln5 as a reactive amino acid substrate for the mark. Consistent with this assignment, mutation of Gln5 to Ala—producing H3(Q5A)—leads to loss of transamidation activity by TGM2 (Extended Data Fig. 2g).

Given the close proximity of Q5ser to Lys4 (a residue that, when tri-methylated, provides a critical signature for transcriptional initiation), we examined the potential effect of K4me3 on TGM2-mediated monoaminylation in vitro. Unmodified or H3K4me3 nucleosomes (Extended Data Fig. 3a–e) were subjected to TGM2 assays with 5-HT. We used antibodies raised against the K4me3 and Q5ser marks (see Methods), and found that TGM2 equally monoaminylates unmodified and K4me3 H3 substrates (Fig. 1c and Extended Data Fig. 3f). Likewise, the presence of Q5ser had no effect on the ability of MLL1, an H3K4 methyltransferase, to modify nucleosomes in vitro (Extended Data Fig. 3g).

[1]Fishberg Department of Neuroscience, Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [2]Department of Chemistry, Princeton University, Princeton, NJ, USA. [3]Beijing Advanced Innovation Center for Structural Biology, MOE Key Laboratory of Protein Sciences, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing, China. [4]Tsinghua-Peking Joint Center for Life Sciences, Tsinghua University, Beijing, China. [5]Epigenetics Institute, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [6]Laboratory of Genetics, The Salk Institute for Biological Sciences, La Jolla, CA, USA. [7]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY, USA. [8]The Rockefeller University Proteomics Resource Center, The Rockefeller University, New York, NY, USA. [9]Max-Delbrück-Center for Molecular Medicine (MDC), Berlin, Germany. [10]Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA. *e-mail: ian.maze@mssm.edu

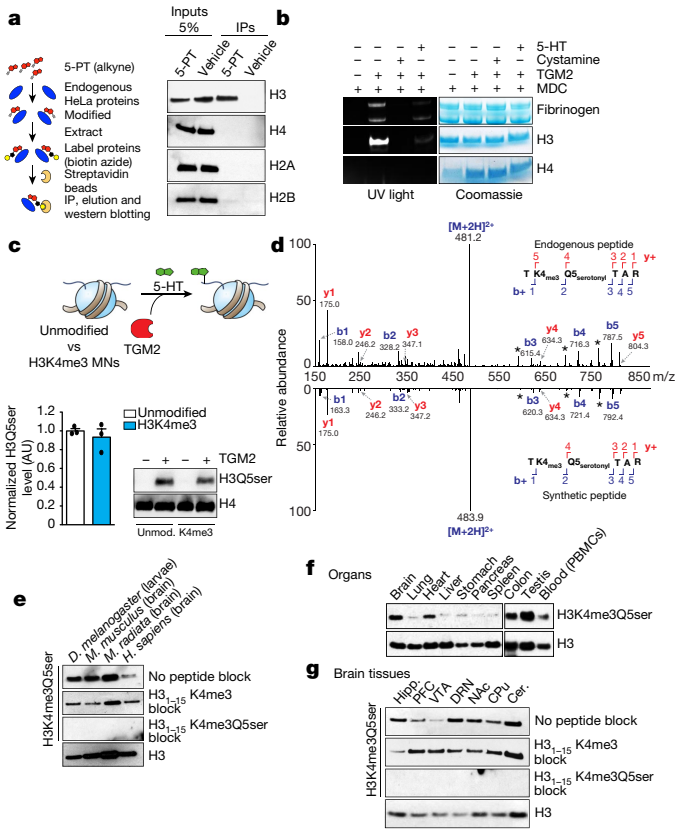


Fig. 1 | Identification of endogenous serotonylated H3 Gln5. **a**, Bioorthogonal labelling of 5-PT-modified histone proteins from nuclear extracts of HeLa cells. H3 results were confirmed in at least three independent experiments. IP, immunoprecipitation. **b**, TGM2 monoaminylation (MDC) assays with histone versus non-histone substrates with or without TGM2 inhibition with cystamine (4 mM) or donor competition with excess 5-HT (500 μM). H3 results confirmed in at least three independent experiments. **c**, TGM2 serotonylation assays on unmodified (unmod.) versus H3K4me3 nucleosomes ($n = 3$ per modification type, two-tailed Student's t-test; $t_4 = 0.7309$, $P = 0.50$). H4 is provided as a loading control. AU, arbitrary units. **d**, Mirror plot representation (aligned by mass/charge ratio (m/z), and magnified 2× on either side of the precursor ion) of MS/MS spectra for endogenous versus heavy synthetic (that is, D5-labelled on the N terminus leading to mass shifts for $b+$ fragment ions exclusively) H3K4me3Q5ser peptides in undifferentiated RN46A-B14 cells. Results confirmed in at least three independent experiments (RN46A-B14 cells pre- versus post-differentiation and in mouse brain). **e**, Multi-species comparison of H3K4me3Q5ser expression ($n = 1$). H3 is provided as a loading control. **f**, **g**, Representative analysis of peripheral organ and blood (**f**) and brain (**g**) cell extracts from mouse for the presence of K4me3Q5ser. PBMC, peripheral blood mononuclear cells; Hipp, hippocampus; PFC, prefrontal cortex; VTA, ventral tegmental area; NAc, nucleus accumbens; CPu, caudate putamen; Cer, cerebellum. $n = 3$ or 4 per tissue or brain region; see Supplementary Table 1 for quantifications. H3 is shown as a loading control. Data are mean ± s.e.m.

To better understand the functional roles of H3 serotonylation in vivo, we raised and validated both single (H3Q5ser) and dual (H3K4me3Q5ser) modification-specific antibodies (Extended Data Fig. 4a–h). LC–MS/MS analyses on material immunoprecipitated with these antibodies in a serotonergic cell line (RN46A-B14 cells) and in adult mouse brain (dorsal raphe nucleus, DRN) unambiguously confirmed the presence of these marks in mammalian cells and tissues (Fig. 1d and Extended Data Fig. 5a, b). Whereas both post-translational modifications were present in culture, only the dual modification was identified in vivo in brain (as well as across a wide range of 5-HT-producing organisms from *Drosophila* to humans (Fig. 1e)). Accordingly, we chose to focus on H3K4me3Q5ser and began by investigating its distribution across mammalian tissues. We identified a ubiquitous pattern of expression, in which the mark is enriched in organs that produce 5-HT (such as brain and colon) and displays more-limited signal in some non-serotonergic organs. Robust signals were also observed in heart, circulating blood (that is, peripheral blood mononuclear cells (PBMCs)) and testes (Fig. 1f). Contrary to initial expectations, in brain, the H3K4me3Q5ser signal is not segregated to regions in which 5-HT is produced (for example, DRN) but is broadly distributed across structures (Fig. 1g; see Supplementary Table 1 for quantifications). Even within DRN, the mark was qualitatively found to not be specific to serotonergic neurons, also displaying signal in non-serotonergic neurons and in non-neuronal cells (Extended Data Fig. 6a–c). This expression is ablated in mice that do not produce serotonin (Extended Data Fig. 6d–f).

To extend our findings to a model of human serotonergic neuronal differentiation, we examined the expression of H3K4me3Q5ser in human pluripotent stem cell (hPSC)-derived 5-HT neurons pre- and post-differentiation[18] (see Fig. 2a–c for cellular validations). We found that differentiation leads to a significant increase in H3K4me3Q5ser levels, with concomitant changes observed in H3K4me3 (Fig. 2d and Extended Data Fig. 4h). To assess genome-wide effects of serotonergic differentiation, we performed chromatin immunoprecipitation with DNA sequencing (ChIP-seq) using our dual-post-translational-modification antibody. In hPSCs, peak-calling revealed negligible enrichment of the mark. However, the total number of peaks for H3K4me3Q5ser increased significantly with differentiation; genomic distribution patterns for the mark in 5-HT neurons revealed a strong bias towards promoters (Fig. 2e and Supplementary Tables 2, 3). To investigate precise genomic loci that display differential regulation of H3K4me3Q5ser during differentiation, we performed diffReps[19] and identified 12,086 protein-coding genes with altered enrichment. A large majority of these changes were found to be increased events occurring within gene promoters (Fig. 2f and Supplementary Table 4), as exemplified by the neuronal gene locus for *ELAVL3* versus *NANOG*, a transcription factor associated with self-renewal in undifferentiated stem cells (Fig. 2g). Further assessments (Enrichr, using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis)[20,21] of genes displaying altered H3K4me3Q5ser enrichment indicated significant associations with pathways involved in neuronal differentiation and development (Supplementary Table 5). To understand the potential roles of H3K4me3Q5ser in the regulation of gene expression, we used RNA-sequencing analysis (RNA-seq) in 5-HT neurons versus hPSCs, and performed Fisher's exact tests comparing genes that contain differential H3K4me3Q5ser sites to those that are differentially expressed. In doing so, we identified correlations between enrichment of the mark and altered expression, particularly in comparison to genes that display increased abundance (Fig. 2h and Supplementary Table 6). KEGG analysis of these overlapping protein-coding genes identified significantly enriched pathways associated with 'axon guidance' and 'serotonergic synapse', among others (Supplementary Table 5).

Consistent results were observed in mouse brain during embryonic development from embryonic day (E)9.5 to E17.5, a time period in which serotonergic neurons become fully differentiated and establish far-reaching innervations[22] (Extended Data Fig. 7a–f and Supplementary Tables 7–14). Similar to results in human 5-HT neurons, neuronal genes that are activated during development exhibited significantly increased enrichment for both H3K4me3 and H3K4me3Q5ser, whereas other genes that remained equally expressed at both time points displayed similar genomic enrichment for the two marks at both developmental stages. These data suggest that the combination of K4me3 and Q5ser on H3 may reinforce patterns of permissive, or inducible, gene expression in vivo. However, given that both marks are induced during differentiation and development in these systems, it remained difficult to separate their specific contributions.

Therefore, to more causally assess the role of Q5ser in these processes, we used a RN46A-B14 cell culture model[23], in which these

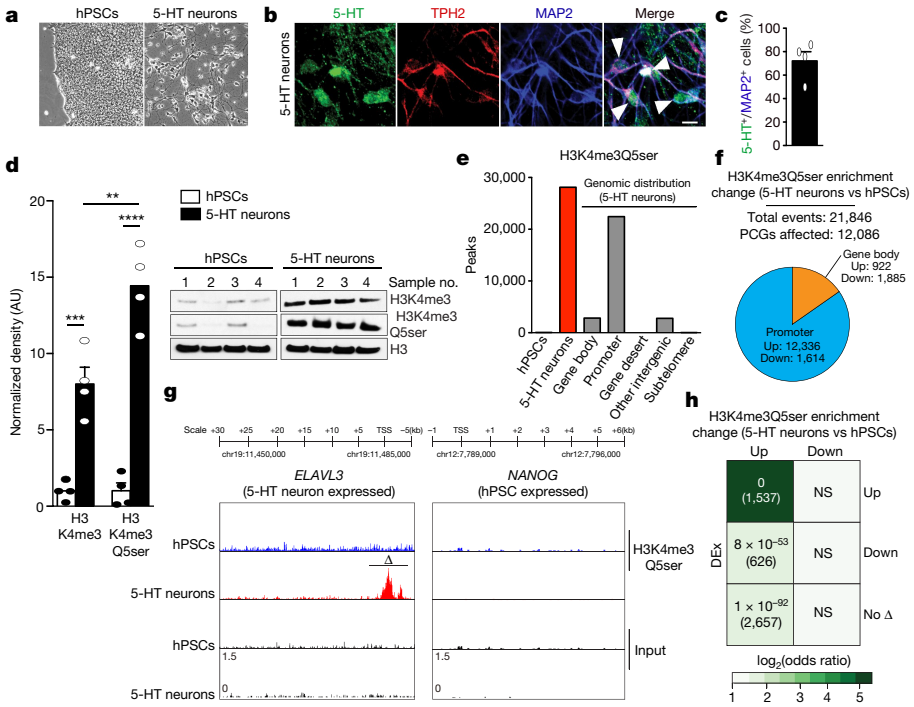


Fig. 2 | H3K4me3Q5ser is responsive to human serotonergic neuronal differentiation. a–c, Bright-field (**a**) and immunofluorescence (**b**) images (line 1, provided as representative) that validate differentiation of 5-HT$^+$TPH2$^+$ neurons from hPSCs with approximately 72% efficiency (**c**; 5-HT$^+$MAP2$^+$) across all four lines examined. Scale bar, 20 μm. **d**, Quantitative western blot analysis of H3K4me3 and H3K4me3Q5ser expression in 5-HT neurons versus hPSCs. Two-tailed Student's *t*-tests; $n = 4$ per group; K4me3: $t_6 = 6.134$, $^{***}P = 0.0009$; K4me3Q5ser: $t_6 = 9.513$, $^{****}P < 0.0001$; K4me3Q5ser versus K4me3 neurons: $t_6 = 3.755$, $^{**}P = 0.0095$. H3 was used as a loading control. **e**, MACS v.2.1.1-based peak-calling for H3K4me3Q5ser ChIP-seq data pre- and post-differentiation (false discovery rate, FDR < 0.05; fold change, FC > 1.2; cut-offs applied after adjusting for multiple comparisons; $n = 4$ lines per pre- versus post-differentiation; normalized to respective inputs). **f**, diffReps analysis of H3K4me3Q5ser differential enrichment pre- versus post-differentiation (FDR < 0.05, FC > 1.2 cut-offs applied after adjusting for multiple comparisons; $n = 4$ lines per pre- versus post-differentiation). Pie chart indicates distribution of genic differential enrichment events for the mark comparing promoter and gene body regions. **g**, Representative genome browser tracks of *ELAVL3* versus *NANOG* loci for H3K4me3Q5ser (versus DNA input) in hPSCs versus 5-HT neurons. $n = 4$ lines per pre- versus post-differentiation). Δ indicates a statistically significant site of differential enrichment for H3K4me3Q5ser. TSS, transcriptional start site. **h**, Odds-ratio analysis of overlapping genes displaying differential H3K4me3Q5ser enrichment (FDR < 0.05, FC > 2.5; $n = 4$ lines per pre- versus post-differentiation) versus differential gene expression (DEx, $n = 3$ lines per pre- versus post-differentiation; lines 1–3, FDR < 0.05 cut-off applied after adjusting for multiple comparisons). Insert numbers indicate respective *P* values for associations, with the number of protein-coding genes overlapping per significant category in paretheses. Data are mean ± s.e.m.

manipulatable, proliferative 5-HT-producing cells can be differentiated in vitro to increase 5-HT production, induce cell-cycle arrest and promote a neurite-outgrowth phenotype that is dependent on brain-derived neurotrophic factor (Fig. 3a). Consistent with data from human 5-HT neurons and developing mouse brain, we found that RN46A-B14 cellular differentiation leads to a significant increase in H3K4me3Q5ser levels (Fig. 3b). However, unlike in the other two models, we did not observe changes in H3K4me3 expression ($n = 3$; $t_4 = 1.188$, $P = 0.30$), a phenomenon that we reasoned might aid in further dissecting specific contributions of H3K4me3Q5ser to gene expression. To validate the usefulness of this system for causal assessments of H3K4me3Q5ser, we assessed genome-wide effects of differentiation by performing ChIP-seq using our dual-post-translational-modification antibody versus an antibody against H3K4me3—note that the H3K4me3 antibody detects the methyl mark equally in both the presence and absence of Q5ser (Extended Data Fig. 8a). Before differentiation, peak-calling revealed roughly equal genome-wide enrichment of the two marks. However, the total number of peaks called for H3K4me3Q5ser increased with differentiation, whereas enrichment for H3K4me3 remained unchanged; genomic distribution patterns for the two post-translational modifications were nearly identical (Fig. 3c, Extended Data Fig. 8b and Supplementary Tables 15–18). To investigate genomic loci that display differential regulation of H3K4me3Q5ser during differentiation, diffReps was performed and identified 6,386 protein-coding genes with altered enrichment; about 60% of all K4me3 enriched genes pre-differentiation displayed such alterations in H3K4me3Q5ser signal. The majority of these changes were found to be increased events occurring within gene promoters (Fig. 3d and Supplementary Tables 19, 20), as exemplified by the vimentin (*Vim*) locus (Fig. 3e). Around 65% of 'bivalent genes' in RN46A-B14 cells (broadly defined as those that display substantial peaks for both H3K4me3 and H3K27me3 pre-differentiation, which account for around 14% of all K4me3 marked protein-coding genes within these cells), were found to overlap with genes that display differentially enriched H3K4me3Q5ser, which indicates that alterations in this mark may act to circumvent bivalent gene repression during differentiation (Supplementary Tables 21–23). Subsequent assessments of gene pathways that display altered H3K4me3Q5ser enrichment in RN46A-B14 cells yielded similar results to those observed in differentiated human serotonergic neurons and in mouse brain, pathways that displayed minimal overlap with those identified in H3K4me3 differential analyses (Supplementary Tables 24, 25). Finally, we performed RNA-seq in these cells and applied Fisher's exact tests, comparing genes with differential H3K4me3Q5ser sites to those that are differentially expressed. We identified distinct (versus H3K4me3; Extended Data Fig. 8c) and positive correlations between enrichment of the two post-translational modifications and differential gene expression (Fig. 3f and Supplementary Table 26), indicating that H3K4me3Q5ser is likely to be a permissive modification.

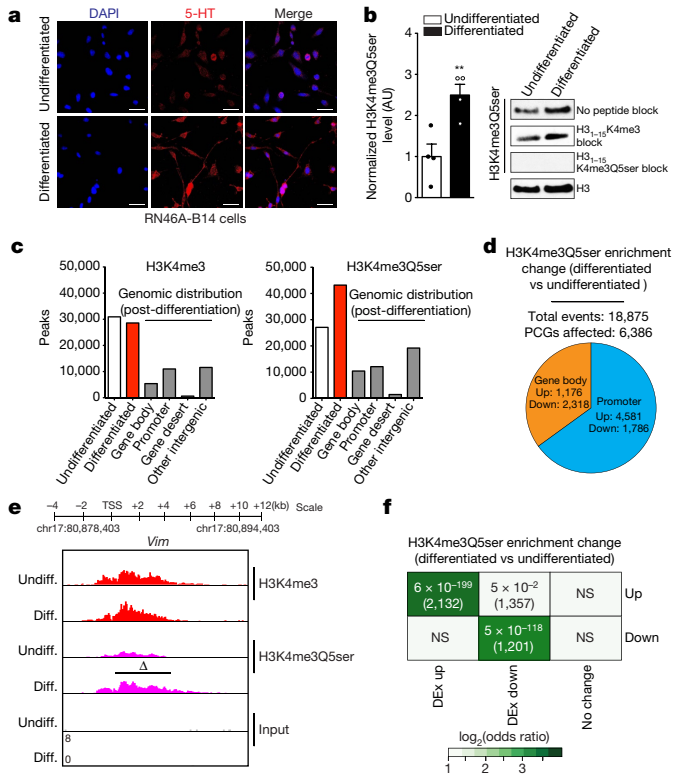


Fig. 3 | H3K4me3Q5ser correlates with permissive gene expression during cellular differentiation independently of H3K4me3. **a**, Immunofluorescence images of 5-HT in RN46A-B14 cells overlaid with a nuclear co-stain (DAPI), pre- and post-differentiation. Scale bar, 20 μm. Results confirmed in at least three independent experiments **b**, Quantitative western blot analysis of H3K4me3Q5ser expression in RN46A-B14 cells pre- versus post-differentiation (two-tailed Student's *t*-test; $n = 4$ per group; $t_6 = 3.736$, $^{**}P = 0.0097$). H3 was used as a loading control. **c**, MACS v.2.1.1-based peak-calling for H3K4me3 versus H3K4me3Q5ser ChIP-seq data pre- and post-differentiation in RN46A-B14 cells (FDR < 0.05, FC > 1.2 cut-offs applied after adjusting for multiple comparisons; $n = 3$ independent biological samples per pre- versus post-differentiation; normalized to respective inputs). **d**, Analysis of H3K4me3Q5ser differential enrichment comparing RN46A-B14 cells pre- and post-differentiation (FDR < 0.05, FC > 1.2 cut-offs applied after adjusting for multiple comparisons; $n = 3$ independent biological samples per pre- versus post-differentiation). Pie chart indicates distribution of genic differential enrichment events for the mark comparing promoter and gene body regions. **e**, Genome browser tracks of the *Vim* locus for H3K4me3 and H3K4me3Q5ser (versus DNA input) in RN46A-B14 cells pre- and post-differentiation. Δ indicates a statistically significant site of differential enrichment for H3K4me3Q5ser as determined by diffReps ($n = 3$ independent biological samples per pre- versus post-differentiation). Diff., differentiated; undiff., undifferentiated. **f**, Odds-ratio analysis of overlapping genes that display differential H3K4me3Q5ser enrichment versus differential gene expression (DEx, $n = 3$ independent biological samples per pre- versus post-differentiation, FDR < 0.05 cut-off applied after adjusting for multiple comparisons). Insert numbers indicate respective *P* values for associations, with the number of protein-coding genes that overlap per significant category in parentheses. Data are mean ± s.e.m.

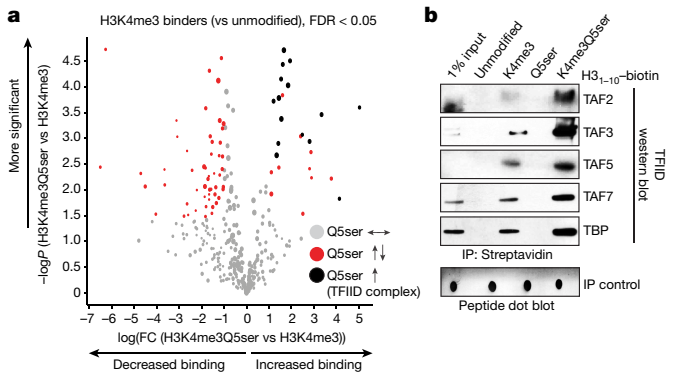


Fig. 4 | H3K4me3Q5ser potentiates TFIID interactions with H3K4me3. Modified $H3_{1–10}$ peptide immunoprecipitations from HeLa nuclear extracts identifying potentiated interactions between TFIID-complex proteins and the H3K4me3 tail in the presence of Q5ser. **a**, LC–MS/MS with label-free intensity-based absolute quantification. $n = 4$ independent biological replicates per peptide; results of *t*-tests (adjusted for multiple comparisons) are illustrated with a Volcano plot. Sizes of the filled circles indicate the approximate amount of a given protein. FDR < 0.05, FC > 2 cut-offs applied. Grey circles, K4me3 binders unaffected by Q5ser; red circles, K4me3 binders potentiated or attenuated by Q5ser; black circles, TFIID-complex proteins altered in their interactions by Q5ser. **b**, Representative western blot images of modified $H3_{1–10}$ peptide immunoprecipitations from HeLa nuclear extracts (quantifications provided in Extended Data Fig. 9a).

To directly validate the putative role of H3K4me3Q5ser in active gene expression, we generated lentiviruses expressing either wild-type H3.3 or H3.3(Q5A) (H3.3 was used to ensure chromatin incorporation as cells become post-mitotic[24])—a mutation that prevents serotonylation (Extended Data Fig. 2g) but does not affect the ability of H3 to be tri-methylated at lysine 4 (Extended Data Fig. 8d)—and transduced them into RN46A-B14 cells during differentiation. Gene expression patterns of H3K4me3Q5ser target versus non-target loci were assessed by RNA-seq (Supplementary Table 27). Our data demonstrate that competing against endogenous H3K4me3Q5ser expression alters levels of H3K4me3Q5ser target genes, with little to no effect on non-target loci (Extended Data Fig. 8e). Furthermore, many of the target genes regulated by H3.3(Q5A) displayed opposite patterns of regulation in comparison to those observed during normal differentiation (Extended Data Fig. 8f), suggesting that attenuating expression of the mark may cause these cells to remain in a more undifferentiated-like state. In line with this idea, KEGG analysis identified axon guidance as the sole enriched pathway affected by H3.3(Q5A) expression (Supplementary Table 28). Next, we applied a specific inhibitor of TGM2 (LDN 27219) to RN46A-B14 cells during differentiation, followed by quantitative PCR with reverse transcription assessments of candidate genes predicted to be regulated by the mark. Of the candidate loci examined, all were found to be downregulated in their expression during differentiation (Extended Data Fig. 8g), consistent with the putative role of H3K4me3Q5ser as a permissive modification. Finally, predicted associations between induction of H3K4me3Q5ser and progression of cellular development were validated through assessments of RN46A-B14 neurite outgrowth during differentiation, in which expression of H3.3(Q5A) was found to significantly reduce neurite length (Extended Data Fig. 8h, i), consistent with the cells remaining in a less differentiated state.

To evaluate potential mechanisms that contribute to the permissive role of H3K4me3Q5ser in gene expression, we investigated the effect of Q5ser on nuclear protein-binding interactions with the N-terminal H3 tail, both in the presence and absence of K4me3. We used synthetically modified and biotinylated H3 peptides in nuclear-extract immunoprecipitations, followed by LC–MS/MS to identify putative binding proteins. Numerous interactions with H3K4me3 versus unmodified H3 were identified, including many well-established H3K4me3 'readers' (Supplementary Table 29). Focusing on H3K4me3-binding proteins, we assessed whether the presence of Q5ser in combination with K4me3 might influence these binding events. In doing so, we found that 77 (around 21%) of these proteins displayed either potentiated or attenuated binding in the presence of Q5ser (Fig. 4a and Supplementary Tables 29, 30). Among those potentiated were 15 members of the general transcription factor complex TFIID (TATA-box-binding protein and its associated factors (TAFs)). Such enhanced interactions between TFIID and H3K4me3Q5ser were subsequently validated via western blotting post-nuclear extract immunoprecipitations (Fig. 4b

and Extended Data Fig. 9a), and via peptide immunoprecipitations against the purified TFIID complex (Extended Data Fig. 9b).

Finally, to epigenomically explore relationships between differential H3K4me3Q5ser enrichment and TFIID recruitment to chromatin, ChIP-seq for TAF3—a proxy for the complex—was performed in pre- and post-differentiation RN46A-B14 cells. TAF3–TFIID–chromatin associations were elevated in response to differentiation, with the bulk of the increased enrichment observed around transcriptional start sites (Extended Data Fig. 9c and Supplementary Tables 31, 32), as exemplified by the *Prrg4* locus (Extended Data Fig. 9d). Further assessments of TAF3–TFIID peak overlap after differentiation with differential H3K4me3 or H3K4me3Q5ser (versus those genes containing differential sites for both or neither of the marks) revealed that about 48% of all TAF3–TFIID-bound genes following transition to a post-mitotic state correlate with altered H3K4me3Q5ser enrichment (Extended Data Fig. 9e). Whereas alterations in both marks were found to correlate with TAF3–TFIID genes, only around 28% of differentially enriched genes for H3K4me3 were found to overlap with TAF3–TFIID after differentiation. By comparison with H3K4me3, greater overlaps were also observed between protein-coding genes that display altered enrichment for both H3K4me3Q5ser and TAF3–TFIID post-differentiation (Extended Data Fig. 9f and Supplementary Table 33).

To our knowledge, serotonylation of H3Q5 represents the first endogenous monoaminyl modification of histones to be described (Extended Data Fig. 9g), as well as the first non-methyl post-translational modification of glutamine residues (for example, H2AQ105me[25]) identified in histone proteins. Much remains to be delineated with respect to the in vivo functions of histone serotonylation. For example, questions remain regarding the stoichiometry of the mark relative to H3K4me3; whether there are specific non-H3K4me3 'readers' or demonoaminylases that control its activities; and whether the mark is responsive to environmental perturbations or pathophysiological states associated with alterations in 5-HT production and/or signalling. Q5ser does appear, however, to be a modification that acts to alter specific protein interactions with H3K4me3 and is likely to be involved in reinforcing permissive patterns of gene expression within selected tissues and/or cell types. The precise mechanisms through which the mark is targeted to genomic loci, however, remain unknown. Furthermore, whereas H3K4me3Q5ser associations with TFIID may facilitate locus-specific gene expression during differentiation, such interactions are likely to explain only a fraction of the role of the mark in gene regulation. Finally, it is clear that alternative histone monoaminylation states may exist in vivo and might act distinctly or cooperatively with H3 serotonylation to regulate gene expression in eukaryotic cells. Thus, these findings may reveal a vast array of mechanisms for future investigations related to this class of histone post-translational modifications.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1024-7.

Received: 6 November 2017; Accepted: 15 February 2019;
Published online 13 March 2019.

1. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
2. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
3. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
4. Lauberth, S. M. et al. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).
5. Vermeulen, M. et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).
6. van Ingen, H. et al. Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure* **16**, 1245–1256 (2008).
7. Colgan, L. A., Putzier, I. & Levitan, E. S. Activity-dependent vesicular monoamine transporter-mediated depletion of the nucleus supports somatic release by serotonin neurons. *J. Neurosci.* **29**, 15878–15887 (2009).
8. Young, A. B., Pert, C. D., Brown, D. G., Taylor, K. M. & Snyder, S. H. Nuclear localization of histamine in neonatal rat brain. *Science* **173**, 247–249 (1971).
9. Walther, D. J. et al. Serotonylation of small GTPases is a signal transduction pathway that triggers platelet α-granule release. *Cell* **115**, 851–862 (2003).
10. Watts, S. W., Priestley, J. R. & Thompson, J. M. Serotonylation of vascular proteins important to contraction. *PLoS ONE* **4**, e5682 (2009).
11. Hummerich, R., Thumfart, J. O., Findeisen, P., Bartsch, D. & Schloss, P. Transglutaminase-mediated transamidation of serotonin, dopamine and noradrenaline to fibronectin: evidence for a general mechanism of monoaminylation. *FEBS Lett.* **586**, 3421–3428 (2012).
12. Ballestar, E., Abad, C. & Franco, L. Core histones are glutaminyl substrates for tissue transglutaminase. *J. Biol. Chem.* **271**, 18817–18824 (1996).
13. Kim, J. H. et al. Histone cross-linking by transglutaminase. *Biochem. Biophys. Res. Commun.* **293**, 1453–1457 (2002).
14. Sileno, S. et al. A possible role of transglutaminase 2 in the nucleus of INS-1E and of cells of human pancreatic islets. *J. Proteomics* **96**, 314–327 (2014).
15. Lin, J. C. et al. Characterization of protein serotonylation via bioorthogonal labeling and enrichment. *J. Proteome Res.* **13**, 3523–3529 (2014).
16. Nurminskaya, M. V. & Belkin, A. M. Cellular functions of tissue transglutaminase. *Int. Rev. Cell Mol. Biol.* **294**, 1–97 (2012).
17. Hummerich, R., Costina, V., Findeisen, P. & Schloss, P. Monoaminylation of fibrinogen and glia-derived proteins: indication for similar mechanisms in posttranslational protein modification in blood and brain. *ACS Chem. Neurosci.* **6**, 1130–1136 (2015).
18. Lu, J. et al. Generation of serotonin neurons from human pluripotent stem cells. *Nat. Biotechnol.* **34**, 89–94 (2016).
19. Shen, L. et al. diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS ONE* **8**, e65598 (2013).
20. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
21. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44** (W1), 90–97 (2016).
22. Daubert, E. A. & Condron, B. G. Serotonin: a regulator of neuronal morphology and circuitry. *Trends Neurosci.* **33**, 424–434 (2010).
23. White, L. A. et al. Distinct regulatory pathways control neurofilament expression and neurotransmitter synthesis in immortalized serotonergic neurons. *J. Neurosci.* **14**, 6744–6753 (1994).
24. Maze, I. et al. Critical role of histone turnover in neuronal transcription and plasticity. *Neuron* **87**, 77–94 (2015).
25. Tessarz, P. et al. Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505**, 564–568 (2014).

Acknowledgements We thank R. Cagan (ISMMS), J. Coplan (SUNY Downstate) and C. Tamminga (UTSW) for providing *Drosophila*, macaque and human samples, respectively, for analysis of H3K4me3Q5ser (Fig. 1e), and G. Johnson (University of Rochester Medical Center) for wild-type and catalytically dead TGM2 constructs. This work was supported by grants from the National Institutes of Health: DP1 DA042078 (I.M.), R01 MH116900 (I.M.), R21 DA044767 (I.M.), P50 MH096890 (I.M.), R37 GM086868 (T.W.M.), P01 CA196539 (T.W.M. and B.A.G.), R01 GM110174 (B.A.G.), R21 DA040837 (B.A.G.), R01 CA129325 (R.G.R.), R01 CA204639 (R.G.R.), T32 DA007135 (R.M.B.), as well as awards from: MQ Mental Health Research Charity, MQ15FIP100011 (I.M.), Alfred P. Sloan Foundation, Fellowship in Neuroscience (I.M.), the JPB Foundation (F.H.G.), the Bob and Mary Jane Engman Foundation (F.H.G.), the Volkswagen Foundation (N.A.) and the National Natural Science Foundation of China (31430020 and 31621092, H.L.).

Reviewer information *Nature* thanks Tatiana Kutateladze and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions I.M. conceived of the project with input from L.A.F., O.B. and T.W.M. L.A.F., F.H.G., B.A.G., H.L., T.W.M. and I.M. designed the experiments and interpreted the data. L.A.F., R.E.T., S.Z., A.E.L., Y.L., N.V.B., B.Z., K.C.V., K.J.H., R.M.B., B.J.L., H.M. and I.M. collected and analysed the data. Y.-H.E.L., A.R., G.E. and L.S. performed the sequencing-based bioinformatics with input from I.M. T.N. and R.G.R. provided the Flag–TBP purified TFIID complex. H.Z.III provided a subset of chemically modified peptides. N.A. and M.B. provided tissues from *Tph1* and *Tph2* knockout mice. L.A.F., H.L., T.W.M. and I.M. wrote 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/s41586-019-1024-7.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1024-7.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to I.M.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

Chemicals. Commonly used reagents were purchased from Sigma Aldrich and Fisher Scientific. Amino acid derivatives and coupling agents for fluorenylmethyloxycarbonyl cholride (Fmoc) solid-phase peptide synthesis (SPPS) were purchased from Novabiochem and from Matrix Innovation. 5-HT hydrochloride (cat. no. H-9523), monodansylcadaverine (cat. no. 30432) and cystamine dihydrochloride (cat. no. C121509) were purchased from Sigma Aldrich. LDN 27219 was purchased from R&D Systems (cat. no. 4602). Hydroxytryptamine creatinine sulfate–5-(1,2-^{3}H(N)) was purchased from Perkin Elmer (cat. no. NET498001MC). Biotin cadaverine was purchased from Invitrogen. 5-PT was prepared in house as a hydrochloride salt, according to reported synthetic procedures[15].

Analytical. Analytical reversed-phase high-performance liquid chromatography (RP-HPLC) was performed using an Agilent 1100 series instrument with a Vydac C18 column (5 μm, 1 ml/min). Semi-preparative reverse-phase high performance liquid chromatography (RP-HPLC) was performed on an Agilent 1200 series system with a Vydac C18 column (semi-preparative column 12 μm, 10 mm × 250 mm, 4 ml/min). Preparative scale RP-HPLC purifications were performed on a Waters prep LC equipped with a Waters 2535 Binary Gradient module, Waters 2489 UV detector and a Vydac C18 column (10 μm, 22 × 250 mm, 18 ml/min). All RP-HPLC used 0.1% trifluoroacetic acid (TFA) in water (HPLC solvent A), and 90% acetonitrile and 0.1% TFA in water (HPLC solvent B) as mobile phases. Mass spectrometry analysis was performed using electrospray ionization (ESI) on a MicrOTOFQ II ESI-Q-TOF mass spectrometer (Bruker Daltonics).

Animals. Mice (C57BL/6J) were purchased from The Jackson Laboratory. Tissues from *Tph1* and *Tph2* double-knockout mice (versus littermate controls) were provided by N. Alenina and M. Bader at the Max-Delbrück Center for Molecular Medicine. Animals were group-housed on a 12-h light:dark cycle (lights on from 07:00 to 19:00) at constant temperature (23 °C) with ad libitum access to food and water. All animal protocols were approved by the IACUC at both the Icahn School of Medicine at Mount Sinai (ISMMS) and the Max-Delbrück Center for Molecular Medicine. Pregnant female mice were euthanized with carbon dioxide and mouse embryonic brains (whole) were collected at E9.5 and E17.5 for subsequent sequencing and western blotting experiments. For PBMC isolation, adult mice were killed by rapid decapitation. Trunk blood (~500 μl) was collected in 1.5-ml heparin–lithium lined Eppendorf tubes (cat. no. 022379208). Trunk blood was subsequently mixed with 2 ml of room temperature RPMI medium in 15-ml conical tubes and layered over 2 ml of RT Ficoll reagent (Ficoll-paque PLUS reagent, GE Healthcare cat. no. 17-1440-02 Lot no. 10260010). The total mixture was centrifuged at 2,200 r.p.m for 15 min at room temperature with a slow brake setting. PBMCs were visible as a cloudy layer. This layer was removed slowly with a 1-ml pipette and added to a new sterile 15-ml conical tube. BEP wash buffer in Dulbecco's PBS (DPBS) (5% bovine serum albumin (BSA), 2 mM EDTA) was subsequently added to the PBMC layer and centrifuged at room temperature for 10 min to wash and pellet the cells.

Tissue culture. *Cell culture materials.* Dulbecco's Modified Eagle Medium (DMEM) (Sigma; cat. no. D8437), 0.05% trypsin (Thermo Fisher; cat. no. 25300120), 10% fetal bovine serum (Omega; cat. no. FB11), 100 μg ml^{-1} streptomycin and 100 U ml^{-1} penicillin (Thermo Fisher; cat. no. 15140122), RPMI-1640 medium (Sigma cat. no. R8758), geneticin (Thermo Fisher; cat. no. 11811031), matrigel (VWR; cat. no. 47743-716), B-27 supplement, (Thermo Fisher; cat. no. 17504-044), 0.1% insulin, (Sigma; cat. no. I6634), 0.1% apo-transferrin, (Sigma; cat. no. T1428), 0.1% putrescine (Sigma; cat. no. P5780). B16 base medium (pH 7.35) (740 ml DMEM (GIBCO; cat. no. 31053-036), 260 ml F-12 nutrients (GIBCO; cat. no. 21700-075), 14 mM sodium bicarbonate, 11 mM HEPES (Sigma; Cat. no. H6147) and 1% BSA (Sigma; Cat. no. A3311) in 1 l MilliQ water).

HEK 293T and HeLa cell culture. HEK 293T and HeLa cell lines were obtained from the American Type Culture Collection (ATCC) and grown in DMEM supplemented with 10% FBS with 100 μg ml^{-1} streptomycin and 100 U ml^{-1} penicillin. Cells were maintained at 37 °C in a 5% CO_2, 95% humidified incubator.

Human PSC and serotonergic neuronal culture. Fibroblasts from healthy individuals were cultured for reprogramming into induced PSCs. Reprogramming from primary human fibroblasts was achieved using Sendai virus within CytoTune-iPS Sendai Reprogramming Kits according to manufacturer's instructions (Invitrogen, A13780-02, A16517, A16518). PSCs were confirmed to be karyotypically normal and were cultured on matrigel-coated plastic plates (BD Biosciences) using a modified recipe of mTeSR1 medium (StemCell Technologies). In brief, hPSCs were maintained as colonies and passaged (line: #1 = p13, #2 = p12, #3 = p14 and #4 = p19) by mechanical and enzymatic dissociation (Dispase). Cells were cultured for one week with daily medium changes before collecting. Serotonergic neurons were generated using a modified version of a previously described method[18]. In brief, floating embyroid bodies were generated by mechanical and enzymatic (collagenase) dissociation of hPSC colonies. Embyroid bodies were cultured for one week with DMEM/F12-Glutamax:neurobasal medium (1:1) including N2 and B27 supplements, 1× non-essential amino acids, SB431542 (2 μM), DMH-1 (2 μM) and CHIR99021 (1.4 μM) (serotonergic medium). Embyroid bodies were plated onto to poly-ornithine/laminin (Sigma)-coated plastic plates in serotonergic medium with the addition of sonic hedgehog protein (Shh, 500 ng/ml) for another week. Rosette-forming embyroid bodies and derived neural precursor cells were selected and cultured as monolayers (~400,000 neural precursor cells per slide) in serotonergic medium containing SHH (500 ng/ml) and FGF-4 (10 ng/ml), and were then maintained at high densities. Serotonergic neurons were derived by culturing for an additional three weeks at lower densities in serotonergic differentiation medium comprised of a base of neurobasal medium containing glial cell-derived neurotrophic factor (10 ng/ml), brain-derived neurotrophic factor (10 ng/ml) (Peprotech), dibutyrl-cyclicAMP (1 mM, Sigma), ascorbic acid (200 nM, Sigma), laminin (1 μg/ml), DAPT (2.5 μM) and TGFβ3 (1 ng/ml). Medium was changed every other day for three weeks before cells were collected for further processing. All procedures for human stem cell work were monitored and approved by the Institutional Review Board of the Salk Institute.

RN46A-B14 cell differentiation. For RN46A-B14 cell differentiation, a modified version of published protocols was used[23,26]. In brief, RN46A-B14 cells (provided by O. Berton, ISMMS) were cultured and maintained in DMEM with 10% FBS, 1% penicillin–streptomycin and 250 μg/ml geneticin at 33 °C with 5% CO_2 on 10-cm cell culture dishes. The day before differentiation, cell culture dishes were coated with 6 ml matrigel–DMEM mixture for 1 h at room temperature. After matrigel polymerization, 0.7–0.8 × 10^6 cells were seeded onto each plate with 10 ml regular growth medium. Plates were then incubated at 33 °C overnight, enabling the cells to adhere to the plate. On the day of differentiation, regular growth medium was removed from plates and 10 ml of differentiation medium (1% penicillin–streptomycin, 1% B-27 supplement, 0.1% insulin, 0.1% apotransferrin, 0.1% putrescine, in B16 base medium) was added to each plate. Cells were then incubated at 39 °C with 5% CO_2, 95% humidity for three days. On the fourth day, differentiation medium was changed conditionally at a 1:1 ratio. Differentiated RN46A-B14 cells were collected on the sixth day with 0.05% trypsin by centrifugation, washed several times and then reconstituted in DPBS on ice.

Other cell lines and authentication. RN46A-B14 cells were provided by O. Berton, ISMMS. hPSCs and hPSC-derived 5-HT neurons were provided by F. Gage, Salk Institute. Sf9 cells were provided by T. Muir, Princeton University. HEK 293T and HeLa cells were authenticated drawing on our vast experience working with these cell lines (based on cell morphology, growth conditions and other factors). RN46A-B14 cells were validated for (1) 5-HT production, and (2) differentiation potential (for example, cell division arrest, neurite formation and alterations in gene expression), as described throughout the manuscript. hPSCs and derived 5-HT neurons were validated as discussed in the main text (Fig. 2). Sf9 cells were validated via their usefulness in the purification of TGM2 protein. All cell lines tested negative for mycoplasma contamination.

Treatments. 5-PT treatment: cells (HEK 293T or HeLa S2) were incubated with 5-PT (500 μM) at a confluency of ~75%. Cells were treated for 3 h, followed by collection with 0.05% trypsin and pelleting in DPBS. Cystamine treatment: where indicated, cells received cystamine dihydrocholoride at a concentration of 4 mM for 3 h in conjunction with 5-PT treatments.

LDN 27219 treatment: cells (RN46A-B14) were incubated with the reversible TGM2 inhibitor LDN 27219 (10 μM) on day 1 of differentiation. On the fourth day, differentiation medium was changed conditionally at a 1:1 ratio with no further inhibitor added. Differentiated RN46A-B14 cells were collected on the sixth day with 0.05% trypsin by centrifugation, washed several times and then reconstituted in DPBS on ice.

Transfections. HEK 293T cells were cultured in six-well plates, and transient transfections were performed at a FuGENE HD Transfection Reagent (Promega; cat. no. E2311) to DNA ratio of 3:1. For transfections, 1 μg of plasmid DNA was transfected per well; 48 h post-transfection, cells were detached with 0.05% trypsin and collected.

Peptide synthesis and generation of recombinant proteins. *Preparation of serotonylated peptides for antibody generation.* Peptide immunogens corresponding to $H3_{1-10}$ (Q5ser or K4me3Q5ser) were prepared using standard Fmoc-strategy solid-phase chemistry on a 2-Cl trityl resin incorporating a C-terminal Cys residue for conjugation to maleimide-activated keyhole limpet haemocyanin (KLH). Orthogonally protected Fmoc-Glu(OAll)-OH was incorporated at position 5 for on-resin serotonylation, and Fmoc-Lys(me3)-OH was coupled at position 4 to generate the dual modified immunogen. Following assembly of the peptide backbones, resin-bound peptides were allyl deprotected by treatment with $Pd(0)(PPh_3)_4$ (0.2 equivalents (equiv.)), *N*,*N*-dimethylbarbituric acid (10 equiv.) in DCM for 30 min (two treatments). Resins were extensively washed with DCM and DMF to remove residual Pd, and the Glu sidechain carboxylate was serotonylated by two treatments with PyAOP (1.1 equiv.) and DIEA (4.0 equiv.) in DMF, followed by direct addition of serotonin hydrochloride (2.0 equiv.) and agitation for 30 min. The crude serotonylated peptides were cleaved from the resin with concomitant global deprotection via treatment with 5% ethanedithiol (EDT), 2.5% iPr_3SiH, 2.5% H_2O in TFA for 2 h at room temperature. Peptides were purified by preparative C-18 RP-HPLC, and

final products were characterized by analytical C-18 RP-HPLC and ESI-MS. The purified peptides were then provided to EMD Millipore for generation of rabbit polyclonal antibodies, as described below.

Preparation of biotinylated $H3_{1-10}$ peptides. Modified peptides corresponding to $H3_{1-10}$ were prepared using standard Fmoc-strategy solid-phase chemistry on a 2-Cl trityl resin incorporating a C-terminal Lys(biotin) residue. Peptides were serotonylated, purified and characterized according to the above conditions.

Preparation of $H3_{1-15}$ peptides. Peptides corresponding to $H3_{1-15}$ were synthesized on a Rink Amide ChemMatrix resin using standard Fmoc-strategy solid-phase chemistry. The methylation state of H3K4 was varied through incorporation of suitably protected lysine residues to access me1 (Fmoc-Lys(Me,Boc)-OH), me2 (Fmoc-Lys(Alloc)-OH)) or Me3 (Fmoc-Lys(me3)-OH). Conversion of the Lys(Alloc) to Lys(me2) was achieved through dual Alloc/Allyl deprotection as outlined above, followed by reductive amination with 37% formaldehyde and $NaBH_3CN$ in PBS/MeOH before serotonylation. Peptides were serotonylated, purified and characterized according to the above conditions.

Recombinant proteins. Human histones H3.2 and H4 were cloned into pRUTH5 to yield an N-terminal His_6-TEV tag. Histones were then expressed in BL21 Star(DE3) cells (Thermo Fisher; cat. no. C6010-03) into inclusion bodies for 3 h at 37 °C. Inclusion bodies were re-solubilized in 6 M guanidine-HCl, 1 M NaCl and 50 mM Tris-HCl (pH 8) before purification on a Ni-NTA affinity column. To generate recombinant H3(Q5A), mutations were introduced into the pRUTH5-H3.2 construct using standard site-directed mutagenesis. Additional histones were purchased from Active Motif (H3.1; cat. no. 31294 and expressed-protein-ligation (EPL)-generated H3K4me3; cat. no. 31278). For some experiments, recombinant human TGM2 was purchased from R&D Systems (cat. no. 4376-TG-050), and guinea pig TGM2 was purchased from Zedira (cat. no. T006); KMT2A (MLL1) complex from Active Motif (cat. no. 31423); and fibrinogen from human plasma from Sigma (cat. no. F4883).

Production of TGM2 in Sf9 cells. TGM2 (Uniprot ID: P21980) was expressed using a baculovirus expression vector system. The coding sequence for TGM2 was obtained from a mammalian expression vector mCer-TG2-YFP (Addgene Plasmid #85471; this plasmid contained two non-native missense mutations (T141A and V224G), which were mutated to the endogenous residues). This gene was subcloned into a pACEBac1 vector using standard restriction enzyme methods with a N-terminal Flag affinity tag introduced for purification of the enzyme yielding pACEBac1-Flag-TGM2. All sequences were confirmed by Sanger sequencing (Genewiz). For generation of the bacmid, pACEBac1-flag-TGM2 was transformed into DH10bac *Escherichia coli* competent cells according to the manufacturer's instructions (ThermoFisher Scientific). Virus was generated as previously described[2]. TGM2 was expressed by adding a 1:500 dilution of the P3 virus to a 2×10^6 cells per ml culture of Sf9 cells. After incubating for 48 h at 27 °C in the dark, cells were collected by centrifugation, washed with PBS and pelleted. Using 40 μl per ml culture, the pellet was resuspended in TEGN (25 mM Tris, 1 mM EDTA, 10% glycerol, 0.02% IGEPAL CA 630, pH 7.5) plus 500 mM NaCl, 1 mM DTT, 0.1 mM PMSF. The cells were lysed by 20 strokes of a tight Dounce homogenizer, followed by sonication. The lysate was cleared by centrifugation (17,000*g* × 30 min), NaCl concentration diluted to 350 mM with TEGN and again cleared by centrifugation. The supernatant was incubated with Flag–agarose beads (Genscript G1 a-DYKDDDDK, 0.5 μl beads per ml culture) for 3 h at 4 °C. Following incubation, the beads were washed three times with TGN (25 mM Tris, 10% glycerol, 0.02% IGEPAL CA 630, pH 7.5) plus 350 mM NaCl, and a final wash with TGN plus 120 mM NaCl. The protein was then eluted from the beads through incubation at 4 °C with Flag peptide (0.25 mg/ml in TGN plus 120 mM NaCl, 1 mM DTT, 3 × 10 min × 1 column volume). The eluted protein was then concentrated using a Vivaspin 500 centrifugal concentrator (MWCO 30,000) and dialysed against TGN plus 120 mM NaCl, 1 mM DTT (2 × 4 h). The dialysed, purified TGM2 was aliquoted and stored at −80 °C.

Generation of H3K4me3 and H3Q5ser nucleosomes. *Peptide synthesis.* Synthetic peptide thioesters used in semisynthesis were manually synthesized on SEA-PS resin (Iris). The coupling of the first glycine residue was performed using 5 equiv. of Fmoc-Gly-OH, 4.9 equiv. of HATU, and 10 equiv. DIEA (2 × 1 h). Following coupling, a capping step using 10% DIEA and 10% Ac_2O in DMF (5 ml) was performed (1 × 10 min). The peptide was then extended using iterative cycles of Fmoc deprotection (3 ml of 20% piperidine in DMF (1 × 1 min, 2 × 8 min)) and amino acid coupling (4.9 equiv. of HOBt/HBTU and 10 equiv. DIEA for 20–45 min) with thorough washing with DMF between each step. Double couplings were used when necessary to ensure complete acylation of the growing peptide. Serotonylated glutamine residues were introduced by incorporation of a Fmoc-L-glutamic acid γ-allyl ester (Chem-Impex International) followed by deprotection and coupling of a serotonin as described above (in Preparation of serotonylated peptides for antibody generation). Following peptide assembly, the crude peptide was cleaved and deprotected using 92.5% TFA, 2.5% EDT, 2.5% iPr_3SiH and 2.5% H_2O. The crude peptide was then precipitated with diethyl ether and allowed to air-dry. The peptides were then converted to C-terminal thioesters according to established protocols[27]. In brief, (for a 0.05 mmol scale synthesis) the crude peptide (bearing the C-terminal bis(sulfanylethyl)amino linker) was incubated in a solution of mercaptopropionic acid (MPA, 0.5 ml) and TCEP (16 equiv. relative to peptide) in 2× PBS (10 ml) pH 4.0 at 37 °C overnight. The reaction mixture was diluted with water (10 ml) and 10% aqueous TFA (5 ml), extracted with Et_2O to remove MPA, lyophilized overnight and then purified by preparative RP-HPLC using a gradient of 0–30% B over 60 min.

Preparation of recombinant histones. Recombinant human histones H2A, H2B and H4 were expressed in *E. coli* and purified according to established[28] protocols with slight modifications. BL21 (DE3) cells were transformed with plasmids for histone expression and grown in Luria-Bertani (LB) medium at 37 °C until an OD_{600} of 0.6. Protein expression was then induced by addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The histones were expressed at 37 °C for 3–4 h before cells were pelleted by centrifugation (4,000*g*, 15 min, 4 °C). The pellets were resuspended in 10 ml lysis buffer per litre cell culture (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 1 mM BME, 0.1% Triton X-100 pH 7.5). Cells were lysed by sonication (30 s on:30 s off, 5 cycles, 45% amplitude) and centrifuged (16,000*g*, 25 min, 4 °C). The inclusion body pellet was washed twice with cold lysis buffer containing no Triton X-100, and extracted with resuspension buffer (7 M urea, 10 mM Tris, 1 mM EDTA, 100 mM NaCl, 1 mM DTT) for 2 h at 4 °C. The inclusion body extract was then cleared by centrifugation (16,000*g*, 25 min, 4 °C) and filtered. Histones were then isolated through ion exchange chromatography using a HiTrap SP FF column (GE Healthcare Life Sciences) using a gradient of 0.1 to 1 M NaCl in resuspension buffer over 60 min. Fractions were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE), and histone-containing fractions were pooled and purified by preparative RP-HPLC using a gradient of 20–73% B over 60 min.

Preparation of truncated histones. BL21 (DE3) cells were transformed with plasmids coding for N-terminally His_6–SUMO tagged histone H3(14-135, K14C). Expression was performed as described above, and cell pellets were resuspended in 20 ml denaturing lysis buffer (6 M Gdn.HCl, 50 mM Na_2HPO_4, 300 mM NaCl, 1% Triton X-100, 5 mM imidazole, 1 mM TCEP, pH 8.0), lysed by sonication and cleared by centrifugation (16,000*g*, 25 min, 4 °C). The His_6-tagged protein was affinity purified from the lysate by Ni-NTA affinity chromatography. The protein was eluted in elution buffer (6 M urea, 50 mM Na_2HPO_4, 250 mM imidazole, 1 mM TCEP pH 7.5) and dialysed stepwise at 4 °C into refolding buffer (50 mM Na_2HPO_4, 300 mM NaCl, 5 mM DTT, pH 7.5) containing 2 M urea for 4 h, 2 M urea for 18 h and 0 M urea for 4 h. ULP1 protease was added before the final dialysis step to cleave off the His_6–SUMO tag. After cleavage, solid Gdn.HCl was added to solubilize the truncated histone that had partially precipitated during proteolysis, and the solubilized mixture was purified directly by preparative RP-HPLC using a gradient of 20–73% B over 60 min.

Preparation of modified semisynthetic histones. Full-length histones were generated by semisynthesis from two pieces, as previously described[29]. In brief, H3 peptide α-thioesters (residues 1–13) were incubated with truncated histone H3 (residues 14–135, K14C) in 6 M Gdn.HCl, 0.1 M Na_2HPO_4, 20 mM TCEP, 0.1 M mercaptophenyl acetic acid at pH 7.5–8.0. Upon completion of the ligation reaction (as judged by HPLC and ESI-MS analysis) the ligation product (H3Q5ser, K14C) was purified by preparative RP-HPLC using a gradient of 20–73% B over 60 min. The purified semisynthetic protein was then treated with a solution of 2-bromoethylammonium bromide (50 mM) in 6 M Gdn.HCl, 1 M HEPES, 10 mM methionine, 20 mM DTT, pH 7.8 at 27 °C to convert the non-native cysteine residue into a lysine mimic at position 14 ($K14K_c$)[30]. Upon full conversion to the alkylated product, the protein was purified by semi-preparative RP-HPLC using a gradient of 20–73% B over 40 min, and its identity and purity validated by HPLC and ESI-MS analyses.

Histone octamer refolding. Histone octamers were formed using established protocols[31]. Histones were first dissolved in unfolding buffer (6 M guanidinium chloride, 20 mM Tris-HCl, 1 mM DTT, pH 7.9) and their concentration determined by UV-visible absorption at 280 nm. The four histones were combined (1:1:1.1:1.1 molar ratio of H3, H4, H2A and H2B) in unfolding buffer to a final protein concentration of 1 mg/ml. The mixture was repeatedly dialysed against a refolding buffer (2 M NaCl, 10 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 7.9) at 4 °C for 4, 18 and 4 h. Histones octamers were then purified from aggregates and dimers by size-exclusion chromatography using a Superdex S200 10/300 increase column (GE Healthcare Life Sciences). The fractions were analysed by SDS–PAGE and pure fractions were pooled, diluted to a final concentration of 50% glycerol and stored at –20 °C.

DNA preparations. An expression plasmid containing multiple copies of the 147-bp Widom 601 DNA sequence[32] flanked by EcoRV sites was transformed into DH5α *E. coli* cells. Following expression and isolation of the plasmid and digestion with EcoRV, monomeric 601 DNA was purified by precipitation with PEG 6000[33].

Nucleosome reconstitution. Modifications were formed using established methods[34]. In brief, histone octamers and 601 DNA were mixed in 2 M TEK (20 mM Tris-HCl,

2 M KCl, 0.1 mM EDTA, 1 mM DTT pH 7.5 at 4 °C) and placed in a Slide-A-Lyzer MINI dialysis device (3.5-kDa molecular mass cut-off, Thermo Fisher Scientific). The solution was dialysed into 200 ml 1.4 M TEK (10 mM Tris-HCl, 1.4 M KCl, 1 mM DTT, 0.1 mM EDTA, pH 7.5 at 4 °C) at 4 °C for 1 h. By use of a peristaltic pump, 330 ml nucleosome was added at a rate of 1 ml/min overnight to 10 mM TEK (10 mM Tris-HCl, 10 mM KCl, 1 mM DTT, 0.1 mM EDTA, pH 7.5 at 4 °C). Two subsequent dialysis steps (2 × 3 h) were performed in 200 ml of 10 mM TEK before transferring samples to 1.5-ml microcentrifugation tubes, centrifuging (17,000*g*, 4 °C, 10 min) and isolating the supernatant. Nucleosome concentrations were quantified by UV absorbance at 260 nm, and nucleosome quality verified by native gel electrophoresis (5% acrylamide, 0.5× TBE, 120 V, 60 min SYBR Gold nucleic acid stain from Life Technologies).

Enzymatic assays. *TGM2–MDC serotonylation assays—peptides and monomeric histones.* For assessments of TGM2-mediated transamidation of MDC/5-HT to peptides/proteins, 10 μg of fibrinogen, BSA or recombinant histones/peptides were incubated with 5 mM MDC/5-HT (5 μM 5-^{3}HT was used in radioactive assays) and 0.25 μg rTGM2 in a final volume of 25 μl enzymatic buffer (25 mM Tris-Cl, pH 8 and 5 mM $CaCl_2$, plus protease inhibitors). Reactions were then incubated at room temperature for 3 h in the dark. Samples were next heated to 98 °C for 8 min in denaturing sample buffer before gel electrophoresis. MDC-labelled protein bands could be visualized by fluorescence under UV light, followed by Coomassie staining to control for equal protein loading. Radioactive 5-HT transfer was monitored by scintillation counting (post-filter capture).

TGM2 serotonylation-biotin–cadaverine assays—mononucleosomes. Monoaminylation assays were performed by incubating 8 pmol of wild-type or H3K4me3 mononucleosomes with TGM2 (40 nM) in 20 μl enzymatic buffer (25 mM Tris, 10 mM $CaCl_2$, 10 mM DTT, 10 mM KCl, 5 mM serotonin or biotin–cadaverine, pH 7.8) for 1 h at 30 °C. Following incubation for 1 h, the reaction was quenched with a 4× SDS loading buffer and analysed by western blotting using the H3Q5Ser antibody (1:500 in 3% milk) or Streptavidin DyLight 800 (1:10,000 in 3% TBST), and a histone H4 antibody (1:1,000 in TBST) was used to control for loading.

MLL1 complex methyltransferase assays—monomeric histones. To assess MLL1-mediated methylation activity, 5 ug of protein (H3 versus H3(Q5A)) was incubated with 1 μg recombinant MLL1 complex in reaction buffer (50 mM TrisCl pH 8.6, 0.02% Triton X-100, 2 mM $MgCl_2$, 1 mM TCEP and 50 μM SAM) for 3 h at room temperature. Methylation status was assessed via immunoblotting with a H3K4me3-specific antibody.

MLL1 complex methyltransferase assays—mononucleosomes. Equal concentrations of unmodified versus Q5ser mononucleosomes were incubated in reaction buffer containing 50 mM TrisCl pH 8.6, 0.02% Triton X-100, 2 mM MgCl2, 1 mM TCEP and 0.2 μM tritiated SAM[3] (PerkinElmer NET 155050UC, LOT 2094503) for 3 h at room temperature. Following incubations, reactions were pipetted onto filter paper squares (1 cm × 1 cm) and allowed to fully dry. Filter papers were subsequently washed three times in ice-cold sodium carbonate buffer (0.1 M; pH 8.5) in a large beaker on a magnetic stir plate. Filter papers were again allowed to air-dry at room temperature before placing into a scintillation vial containing 5 ml of scintillation fluid. Incorporation of tritiated SAM was measured by liquid scintillation counting (counts per million) of tritium (^{3}H).

LC–MS/MS validation of serotonylated $H3_{1-15}$. Samples were analysed by LC–MS/MS (Dionex 3000 coupled to Q-Exactive, Thermo Fisher). Peptides were separated by C-18 chromatography (inner diameter of 75 μm/3 μm particles, Nikkyo Technologies) using a gradient increasing from 1% B to 40% B in 25 min (A: 0.1% formic acid, B: acetonitrile in 0.1% formic acid). The mass spectrometer was operated in parallel reaction monitoring mode[35] (resolution 35,000, automatic gain control (AGC) target, 5e5, maximum injection time of 60 ms and an isolation window of *m*/*z* of 2.0). The mass spectrum was acquired from *m*/*z* 300–1,650; *m*/*z* = 100 was set as the lowest mass for MS/MS. Charge states 2+ to 5+ of the modified and non-modified peptides were targeted. An energy of 25 normalized collision energy was used for the quadruply charged peptide: ARTKQTARKSTGGKA-NH_2, with and without modification. Extended Data Figure 2f represents a high-resolution, high-accuracy tandem mass spectra (sample +5-HT/+TGM2) of the quadruply charged peptide ARTKQTARKSTGGKA-NH2 modified by 5-HT at Gln5 (*m*/*z* 430.4997 (0.12 p.p.m.)). The full amino acid sequence was accounted for. Selected fragment ions used to validate the serotonylated glutamine are annotated (doubly charged y10/y11 and doubly charged b4/b5). We observed a diagnostic fragment ion at *m*/*z* 160.07554, corresponding to loss of serotonin minus ammonia ($C_{10}H_{10}NO$). The findings were further validated by repeating the experiment using deuterated (D_4) serotonin. As expected, both precursor and fragments containing D_4 serotonin shifted by 4.025 Da (data not shown).

Cellular fractionation. *Nuclear, cytosolic and chromatin preparations.* Cells were collected and homogenized in 200 μl ice-cold buffer A (10 mM HEPES, pH 7.9, 10 mM KCL, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 0.5 mM PMSF) using a 1-ml glass dounce homogenizer and tight pestle. Homogenized lysates were transferred to 1.5-ml Eppendorf tubes and Triton-X 100 detergent was added to a final concentration of 0.1%. Samples were next incubated on ice for 30 min with intermittent flicking of the tubes every 5 min. After centrifugation at 1,300*g* for 5 min at 4 °C, the supernatants were removed and separated into tubes labelled 'S1'. The remaining pellets from this first spin were deemed nuclei. The S1 fraction was further clarified by centrifuging at 20,000*g* for 5 min to remove particulate material (the supernatant is cytosol). The pellet from the first spin was then re-suspended in 200 μl of buffer A (without Triton-X 100) and centrifuged for 5 min to wash it. To further fractionate and separate nuclei into chromatin versus soluble nuclear material, the nuclear pellet was resuspended in 100 μl of buffer B (3 mM EDTA, 0.2 mM EGTA) and incubated on ice for 30 min with flicking every 5 min. Following incubation, the samples were centrifuged, and the supernatants removed as soluble nuclear fraction versus the pellets, which represent chromatin. Whole-cell lysates from tissues were prepared in RIPA buffer. Protein concentrations were determined using the DC protein assay (BioRad; cat. no. 5000111).

Soluble nuclear extract preparation from HeLa cells. Nuclear extracts were prepared from HeLa cells according as described[36], or were purchased from CilBiotech. Before immunoprecipitations, extracts were dialysed into the following buffer: 20 mM HEPES, 0.01% NP40, 0.5 mM EDTA, 10% glycerol, 150 mM KCL and 0.4 mM PMSF. Concentrations were determined by DC protein assays.

Immunoprecipitations. *5-PT immunoprecipitations.* 5-PT immunoprecipitations were performed using whole cell, cytosolic, soluble nuclear or chromatin fractions prepared as described (in 'Cellular fractionation'). Using copper-click chemistry, the alkyne-functionalized 5-HT derivative 5-PT was conjugated to a biotin–azide molecule (Thermo Fisher; cat. no. B10184) and used to label serotonylated proteins. Following treatments with 5-PT, cellular lysates were prepared in proprietary CLICKiT reaction buffer (Thermo Fisher; cat. no. C33372). For whole-cell lysates, cell pellets were sonicated in 200 μl of buffer. For nuclear, cytosolic and chromatin fractions, samples were resuspended in ~200 μl of reaction buffer. Prior to the addition of CLICK reagents, samples were centrifuged at 16,000*g* for 5 min to remove insoluble debris. Lysates were then transferred to fresh 1.5-ml centrifuge tubes. In the following order, 0.25 mM biotin azide, 0.1 mM $CuSO_4$, 0.1 mM copper protectant and 50 μl of 1× CLICK additive was added to the lysates, and the final volumes were increased to ~500 μl with CLICK reaction buffer. Samples were allowed to incubate for 2 h at 4 °C. Magnetic streptavidin beads (Thermo Fisher; cat. no. 11206D) were next washed twice in PBS, and lightly vortexed for each wash. The reaction volume was further increased to 1 ml with PBS, with 5% removed for input before the addition of washed beads. For each immunoprecipitation, 20 μl of bead slurry was used. Immunoprecipitates were rotated at room temperature for 3 h. Following incubation, samples were placed on a magnetic stand to allow beads to separate from the supernatants. Using a pipette, the supernatant (flow through) was carefully removed. The beads were then washed by inverting the tubes and lightly vortexing the beads in cold RIPA buffer (10 mM Tris-Cl, pH 8.0, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 1 mM PMSF) containing protease inhibitors. This was repeated six times, followed by six washes in cold DPBS to remove excess detergent. After the last wash, sample buffer was added to beads and boiled for 8 min at 98 °C, followed by gel electrophoresis and incubation with appropriate primary and secondary antibodies.

Histone tail peptide immunoprecipitations and associated mass spectrometry. One hundred micrograms of each peptide was resuspended with 100 μl of pre-washed immobilized streptavidin beads (DynaBeads streptavidin M-280) in 0.01% DPBS/Triton-X 100, followed by incubation overnight rotating at 4 °C. For each peptide immunoprecipitation, 40 μl of 50% peptide–bead slurry was added to dialysed HeLa cell extracts (10 mg per immunoprecipitation, with 2 mM $MgCl_2$ added just before the beads) or purified TFIID complex (immunoprecipitated from Flag-TATA-box-binding protein (TBP)-expressing HeLa cells, as previously described[4] in binding buffer: 150 mM KCl, 50 mM Tris-HCl at pH 8.0, 1% NP40, 1 mM DTT), and each sample rotated for 4 h at 4 °C. The immunoprecipitates were then centrifuged for 1 min at 1,000 r.p.m. to pellet the beads. Beads were then washed six times in binding buffer substituted with 300 mM KCl (and no $MgCl_2$). For immunoblotting analyses, beads were next washed in cold DPBS and were boiled for 8 min in 30 μl of denaturing sample buffer.

For mass spectrometry analyses, proteins bound to peptides were eluted in 8 M urea, 0.1 M ammonium bicarbonate, 10 mM DTT. After reduction, cysteines were alkylated in 30 mM iodoacetamide. Proteins were then digested with LysC (LysC, endoproteinase LysC, Wako Chemicals) in less than 4 M urea, followed by trypsinization (Trypsin Gold, Promega) in less than 2 M urea. Digestions were halted by adding TFA, and the digests were desalted[37] and analysed by reversed phase nano-LC–MS/MS using a Fusion Lumos (Thermo Scientific) operated in high/high mode.

Data were quantified and searched against Uniprot's human database (70,246 sequences, accessed in March 2016) using ProteomeDiscoverer v.1.4.0.288 (Thermo Scientific), combined with Mascot v.2.5.1 (Matrix Science) and

MaxQuant v.1.6.0.13[38]. Oxidation of methionines and protein N-terminal acetylation were allowed as variable modifications, and all cysteines were treated as being carbamidomethylated. Peptide matches were filtered using a Percolator[39], with a calculated FDR of 1%. For quantitative MaxQuant analyses, FDR threshold for peptides were set at 2% and 1% for proteins.

Data analysis: the experiment consisted of three peptide conditions (H3, H3K4me3 and H3K4me3Q5ser), each with four biological replicates per peptide. Data were processed using Perseus v.1.6.0.7[40]. Reverse matches and potential contaminations were removed. $\log_2$-transformed intensity-based absolute quantification values[41] were normalized by median (per sample). For comparisons, it was required that signals be measured in a minimum three of the four replicates for at least one condition. Missing values were imputed (width: 0.3, downshift: 1.8). First, H3 and H3K4me3 were compared using a *t*-test (permutation-based FDR of 0.05). Only significant proteins found to be more abundant in the H3K4me3 versus H3 condition—with a minimum twofold linear difference—were considered for the second analysis (361 proteins): comparison of H3K4me3Q5ser to H3K4Met3 (*t*-test, permutation-based FDR of 0.05 with $S = 0.1$). All measured proteins and signals, as well as fold differences and *P* and *Q*-values are shown in Supplementary Table 29. Primary data, as well as MaxQuant search results, have been uploaded to PRIDE (PXD008106).

LC–MS/MS identification of endogenous H3Q5ser and H3K4me3Q5ser. *Sample preparation and targeted MS/MS.* Chromatin from RN46A-B14 cells post-differentiation (5×10^6 per sample), as well as 16-gauge mouse DRN tissue punches (bilateral, 3 animals pooled per sample), were lysed and sonicated as described below in ChIP experiments, but without crosslinking. Sonicated chromatin was then incubated with one of our two specific antibodies (H3Q5ser or H3K4me3Q5ser, 7.5 μg per sample) and the marks were immunoprecipitated. On beads, unmodified lysines were derivatized using a 1:3 mixture of propionic anhydride and acetonitrile, as previously described[42]. The derivatized histones were then digested with trypsin (enzyme:sample ratio, 1:20, 6 h, room temperature) in 50 mM NH_4HCO_3. The peptides were further derivatized twice to propionylate N termini. Next, the samples were desalted for LC–MS analysis using C-18 stage tips. The acidified samples were separated using a 75-μm internal diameter × 17 cm Reprosil-Pur C18-AQ (3 μm; Dr. Maisch) nano-column mounted on an EASY-nLC nanoHPLC (Thermo Scientific). The HPLC gradient was as follows: 2–26% solvent B (A, 0.1% formic acid; B, 95% MeCN, 0.1% formic acid) over 45 min, from 26–80% solvent B in 5 min, 80% B for 10 min at a flow-rate of 300 nl/min. The nLC was coupled online to an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher) and data were acquired using targeted data-independent acquisition (DIA). Full-scan mass spectra (m/z 300−1,100) were acquired in the Orbitrap with a resolution of 60,000 and an AGC target of 2×10^5. MS/MS was performed in the ion trap with sequential isolation windows of 50 m/z with an AGC target of 1×10^4, a higher-energy collisional dissociation (HCD) collision energy of 32 and a maximum injection time of 50 ms. MS/MS data were collected in centroid mode. Targeted MS/MS was performed for endogenous Q5ser or K4me3Q5ser marks and D5-propionylated heavy spike-in synthetic peptides. Data were manually searched and the low-abundance endogenous peptides were identified using heavy spike-in controls.

Note on the lack of stoichiometry assessments. Obtaining accurate assessments of stoichiometry represents a complex task, as every detection method contains intrinsic biases. Mass spectrometry, despite its high accuracy and sensitivity, is no exception. In fact, almost all mass spectrometry experiments define only relative abundances and relative changes in post-translational-modification state.

To appropriately define stoichiometries for H3Q5ser and H3K4me3Q5ser, one would need to accurately estimate the molarity not only of the modifications themselves, but also molarities for all other modified states of this peptide (for example, unmodified, H3K4me1, H3K4me2, H3K4me3, H3K4ac and all combinations with H3Q5ser). This task is possible when all synthetic peptide equivalents are synthesized and spiked-in at known molarities. However, this also requires that all forms of a given peptide are detectable within the same sample, as histones from different sample types will present with variable backgrounds and different sensitivities. In the case of the serotonyl modifications, such an experiment is prohibitive, because detection of H3Q5ser and H3K4me3Q5ser peptides requires enrichment using site-specific antibodies, thus creating bias in the quantification of other non-enriched modified forms.

For stoichiometric assessments of these marks, one would need to compare the area under the curve from H3Q5ser containing peptide peaks with other low-level detected marks from the same peptide, such as H3K4me3. Generally, non-enriched histone peptides analysed by mass spectrometry in data-(in)dependent acquisition modes give abundances of 0.002–0.01% for the H3K4me3 mark relative to all $H3_{3-8}$ peptides (unmodified + H3K4me1 + H3K4me2 + H3K4me3 + H3K4ac). For RN46A-B14 cells and mouse DRN, we injected 1–3 ug of histone peptides. The smallest total histone peptide amount that one can inject is about 1 μg, as low-abundance marks, such as H3K4me3, are present at only ∼0.00000084% of the total histone peptides in such a mix, and 3 μg is the most that can be injected in a nanoflow LC without clogging the C18 column and creating chromatographic anomalies. We increased the selectivity and sensitivity of detection of serotonyl species by performing targeted mass spectrometry for m/z corresponding to endogenous Q5ser or K4me3Q5ser peptides. Because the signal intensity of the serotonylated peptides was low and often in the range of noise, we increased the specificity of detection by using D5-labelled synthetic versions of these peptides as elution time controls. We confirmed that the low-level signal corresponding to m/z of serotonylated $H3_{3-8}$ peptides obtained in full mass spectrometry was indeed the correct species by looking at MS/MS for the specific fragment ions within a 0.02 tolerance limit. Under these conditions, serotonyl peptides were not be detected in non-enriched samples. Thus, for a representative sample with ∼0.005% abundance of H3K4me3 relative to all other peptides of $H3_{3-8}$, we deduce that H3Q5ser and H3K4me3Q5ser are lower in abundance; however, the degree to which they differ cannot be accurately defined.

When the marks are immunoprecipitated, endogenous Q5ser and K4me3Q5ser peptides are detected passing all criteria discussed above. However, information about other $H3_{3-8}$ peptides is lost throughout this process and relative stoichiometries cannot be inferred. Until technology exists that allows for enrichment of all modified forms of this peptide with equal affinity in a single step, stoichiometry cannot be measured. Finally, although the serotonyl marks are not of high abundance in vivo (at least in comparison to some other forms of this peptide, such as unmodified $H3_{3-8}$), we remain confident in its functional significance as a combinatorial potentiator of binding interactions with H3K4me3, itself a low-abundance modification of great importance.

Western blotting. Proteins were electrophoresed on 18% or 4–12% NuPAGE Bis-Tris protein gels (Invitrogen) and transferred to nitrocellulose membranes. Efficient transfer was always confirmed with direct blue staining (0.1% stock aqueous solution in milliQ water) before being incubated with primary antibodies (listed below) overnight at 4 °C (all membranes were blocked in 5% milk or 5% BSA depending on the primary antibody used). For competition assays, antibodies were pre-incubated with indicated peptides at a 5:1 concentration of peptide to antibody for 1 h at room temperature. Following pre-incubation with peptides, membranes were incubated with antibody/peptide overnight at room temperature.

Membranes were then washed and incubated with peroxidase-labelled secondary antibodies (1:2,000–1:10,000, depending on the primary antibody used). Bands were visualized using Immobilon Western Chemiluminescent HRP Substrate (Millipore). Where appropriate, bands were quantified with NIH ImageJ Software and bands were normalized to total histone H3 to control for equal loading.

Peptide dot blots. Peptides were dotted in titration onto nitrocellulose membranes. Membranes were left to dry for approximately 1 h until subsequently rehydrating in DPBS and staining for equal peptide loading with direct blue. Membranes were then blocked in 5% milk-PBST and incubated with appropriate primary and secondary antibodies.

Antibodies. All primary antibodies used in this study were validated, where appropriate, for use in ChIP, immunoblotting and/or immunocytochemistry or immunohistochemistry experiments: H3 general (Abcam ab1791; lot no. GR293151-1, 1:10,000), H3K4me3 (Abcam ab8580; lot no. GR273043-6, 1:1,000), H3Q5ser (Millipore ABE1791; lot no. RA1690, 1:500), H3K4me3Q5ser (Millipore ABE2580; lot no. RA1608107, 1:500), H3K27me3 (Cell Signaling 9733S; lot no. 9, used for ChIP-seq), TGM2 (Abcam ab2386; lot no. GR312304-3, 1:500), TAF2 (Bethyl A302-132A, 1:500), TAF3 (Abcam ab188332; lot no. GR220332-1, used for blotting, 1:500), TAF3 (Millipore 07-1802; lot no. 3008423, used for ChIP-seq), TBP (Abcam ab8181; lot no. GR300917-3, 1:1,000), H4 (Abcam ab10158; lot no. GR239048-1, 1:1,000), H2A (Abcam ab88770; lot no. GR197571-1, 1:1,000), H2B (Abcam ab1790; lot no. GR188969-2, 1:1,000), haemagglutinin (Abcam ab9134; lot no. GR4307-10, or HA–Alexa Fluor 488 (Santa Cruz SC-805; lot no. K1716, 1:1,000), Tph2 (Abcam ab111828; lot no. GR266431-6, or Millipore, sheep, 1:1,000), NeuN (Millipore MAB377; lot no. 2683919, 1:1,000) and 5-HT (Sigma S5545; lot no. 033M4805, or Protos Biotech, rabbit, 1:1,000), RFP (Abcam ab65856; lot no. GR3181906-1, 1:1,000), MAP2 (Abcam, chicken, 1:1,000). The TAF5 and TAF7 antibodies were provided by R. Roeder (The Rockefeller University, 1:500).

Immunocytochemistry. To assess potential cell-permeabilizing properties of 5-PT, HeLa cells were cultured in 12-well tissue culture plates. Upon reaching ∼80% confluency, growth medium was aspirated and 5-PT (500 μM) was applied in pre-warmed DMEM (10% FBS and 1% penicillin–streptomycin) and allowed to incubate for 3 h at 37 °C. As a negative control, cells that were not treated with 5-PT (vehicle only) were also assessed. After incubation, medium was removed from each well and 500 μl of 3.7% formaldehyde in DPBS was added for 15 min at room temperature. Fixative was then removed and wells washed twice with 3% BSA in DPBS. For cell permeabilization, 500 μl of 0.5% Triton X-100 in DPBS was added to each well and incubated for 20 min at room temperature. Wells were subsequently washed with 3% BSA in DPBS. A reaction cocktail consisting of 5 μM Alexa Fluor picolyl azide (Thermo Fisher; cat. no. C10641) in Click-iT reaction

buffer containing 0.1 mM copper sulfate and copper sulfate protectant premix (50:50) was added to each well including the negative control. The plate was then incubated for 30 min at room temperature and protected from light. The reaction mixture was then removed and wells were washed again with 3% BSA in DPBS. Final wash solution was removed and cells were incubated with DAPI as a nuclear co-stain. hPSC 5-HT neuron and RN46A-B14 cell immunocytochemistry was performed using standard protocols. Immunofluorescence was visualized using a confocal microscope (Zeiss LSM 780 or Zeiss Airy Scan microscope).

Immunohistochemistry. Brains were collected from mice following transcardiac perfusion with sterile PBS and 2% paraformaldehyde. Brains were further post-fixed in 2% paraformaldehyde for 24 h and incubated in 30% sucrose for an additional 2 days. Fixed brains were sectioned at 40 μm using a Leica CM3050 S cryostat. Free-floating DRN sections were washed in 1× TBS and then blocked for 1 h at room temperature. Brain sections were then incubated with the following primary antibodies at room temperature overnight: H3K4me3Q5ser (1:250), NeuN (1:1,000), TPH2 (Abcam ab121013, 1:250) and 5-HT (1:5000). Sections were then washed three times in TBS and incubated with Alexa Fluor secondary antibodies (1:500) overnight at 4 °C. DAPI was used as a nuclear co-stain. Immunofluorescence was visualized using a confocal microscope (Zeiss LSM 780).

RNA isolation and quantitative PCR with reverse transcription. Cell pellets and tissues were homogenized in Trizol (Thermo Fisher; cat. no. 15596026) and processed according to the manufacturer's instructions. RNeasy Microcolumns (Qiagen; cat. no. 74004) were used to further purify RNA, and a nanodrop spectrophotometer confirmed RNA 260/280 and 260/230 ratios to be >1.8. RNA was reverse-transcribed using a BioRad iScript Kit (cat. no. 1708891). cDNA was quantified by quantitative PCR using SYBR green. Each reaction was run in technical duplicate or triplicate and analysed following the $\Delta\Delta C_t$ method, as previously described, using an appropriate normalization control[43]. Primer sequences can be found in Supplementary Table 34.

Chromatin immunoprecipitation. Brain tissues and cell pellets (10×10^6 cells/n) tissues were crosslinked and quenched. Samples were washed thoroughly before being subjected to lysis and sonication, exactly as previously descibed[24]. Samples were then incubated with specific antibodies (7.5 μg per sample) bound to M-280 Dynabeads on a rotator at 4 °C overnight. The following day, immunoprecipitates were washed, eluted and reverse-crosslinked. Following RNA and protein digestion, DNA fragments were purified using a Qiagen PCR purification kit.

Library preparation. Following DNA and RNA purifications, RNA-seq and ChIP-seq libraries were prepared according to Illumina protocols and sequenced on an Illumina HiSeq2500 or 4000 Sequencer_High Output mode v.4.

ChIP-seq peak-calling and differential analysis. For H3K4me3, H3K4me3Q5ser and TAF3 ChIP-seq in RN46A-B14 cells, raw sequencing reads were aligned to the rat genome (rn6) using default settings of Bowtie v.2.2.0. Only uniquely mapped reads were retained. Alignments were filtered using SAMtools v.0.1.19 to remove duplicated reads. Peak-calling—normalized to respective inputs—was performed using MACS v.2.1.1[44] with default settings and filtered for FDR < 0.05 and fold change > 1.2; the window size was set as 300 bp. Raw sequencing reads from mouse brains (H3K4me3 versus H3K4me3Q5ser) were aligned to the mouse genome (mm10) using default settings of HISAT v.0.1.6b. Only uniquely mapped reads were retained. Alignments were filtered using SAMtools v.0.1.19 to remove duplicate reads, and peak-calling—normalized to respective inputs—was performed using MACS v.2.1.1 with default settings and filtered for FDR < 0.05 and fold change > 1.2. Raw sequencing reads from human PSCs versus 5-HT neurons (H3K4me3Q5ser) were aligned to the human genome (hg38) using default settings of HISAT2 v.2.1.0. Again, only uniquely mapped reads were retained. Alignments were filtered using SAMtools v.1.8 to remove duplicated reads, and peak-calling—normalized to respective inputs—was performed using MACS v.2.1.1 with default settings and filtered for FDR < 0.05 and fold change > 1.2. Differential analyses of H3K4me3 and H3K4me3Q5ser (for all cells or tissues examined) were performed using diffReps[19] with a window size of 1 kb. A default Q value cut-off of 0.05 was used, followed by a fold change cut-off of 1.2. Peaks and differential sites were further annotated to nearby genes or intergenic regions using the region analysis tool from the diffReps package.

RNA-seq analysis and correlations with ChIP-seq. Raw sequencing reads from RN46A-B14 cells were mapped to rn6 using TopHat2 v.2.1.0[45]. Counts of reads mapping to genes were obtained using HTSeq-count against Ensembl v.81 annotation. Differential expression analysis was done using the DESeq2 package[46] (v.1.6.3) at an FDR cut-off of 0.05. Raw RNA-seq reads obtained from mouse brains were mapped to mm10 using HISAT v.0.1.6b. StringTie v.1.2.1 was used to obtain gene-level fragments per kilobase of transcript per million mapped reads values against Ensembl v.84. Raw RNA-seq reads obtained from human cells were mapped to the human genome (hg38) using HISAT2 v.2.1.0. Counts of reads mapping to genes were obtained using HTSeq-count against Ensembl v.90 annotation. Differential expression analysis was done using the DESeq2 package (v.1.6.3), again with a FDR <0.05 cut-off. RNA-seq differential lists were compared to ChIP-seq differential lists based on gene names, and heat maps were drawn using the GeneOverlap (https://github.com/shenlab-sinai/geneoverlap/) package. Note that a more conservative fold change cut-off of >2.5 was used to generate odds-ratio heat maps for 5-HT neuron/PSC diffReps versus gene expression to avoid including false positive 'down' events in selected comparisons.

Constructs and viruses. H3.3 (wild type or Q5A)-Flag-HA(C-Ter) was cloned into a pCDH-RFP vector by PCR and enzyme restriction digestion. Purified plasmids were sent to Genewiz for sequence validation. H3.3-pCDH-RFP plasmids were then sent to Cyagen Biosciences for lentiviral packaging. pShooter pCMV-nuc-myc vectors expressing NLS-tagged wildtype or mutant (W241A) human TGM2[47] were provided by G. Johnson (University of Rochester Medical Center).

RN46A-B14 viral infections. Pre-differentiation, RN46A-B14 cells were split and seeded at ~$0.7–0.8 \times 10^6$ density on 10-cm tissue culture plates coated with matrigel and grown in previously described differentiation medium. On day 1, cells were transduced with either lenti-H3.3 wild type or lenti-H3.3(Q5A) (1:1,000 dilution) in growth medium. On day 4, cells underwent a conditioned medium change (1:1). Post-differentiation, RN46A-B14 cells were either collected with 0.05% trypsin by centrifugation, washed several times and reconstituted in calcium-free DPBS on ice for RNA-seq, or fixed on coverslips for microscopy assessments.

RN46A-B14 neurite length assessments. Neurite outgrowth assays were performed on virally infected, fixed cells that were stained using antibodies for RFP, haemagglutinin and DAPI (RFP was used for length assessments). Cells were imaged in a blinded fashion. Because RN46A-B14 cells grow in a linear, rod-like conformation, the total neurite length (that is, the sum of the length of each cell) was measured using ImageJ software.

General statistics and study design. One-way ANOVAs were performed to determine significance for experiments involving more than two treatment conditions (with subsequent Dunnett's post hoc tests). Two-tailed student's t-tests were used for all other statistical comparisons, with the exception of TGM2 inhibitor PCRs in Extended Data Fig. 8g (one-tailed student's t-tests were used). For data presented without statistics, experiments were repeated multiple times (often ≥ 3) to ensure reproducibility, unless otherwise stated. All data involving statistics are presented as mean ± s.e.m.

Sample size, randomization and blinding. Adequate sample sizes were determined based upon inter-sample variability. Throughout the manuscript, we determined the significance of results based on a general confidence interval of 95%. Where appropriate, cells or animals were randomly assigned to groups (segregated by differentiation conditions, chemical compound, drug or viral treatments and/or genotype). Furthermore, tissues collected from animals for ChIP-seq analyses were randomly pooled to provide ample tissue for biochemical procedures and to minimize variance. Where necessary, investigators were blinded to conditions such as genotype, viral treatment conditions and so on.

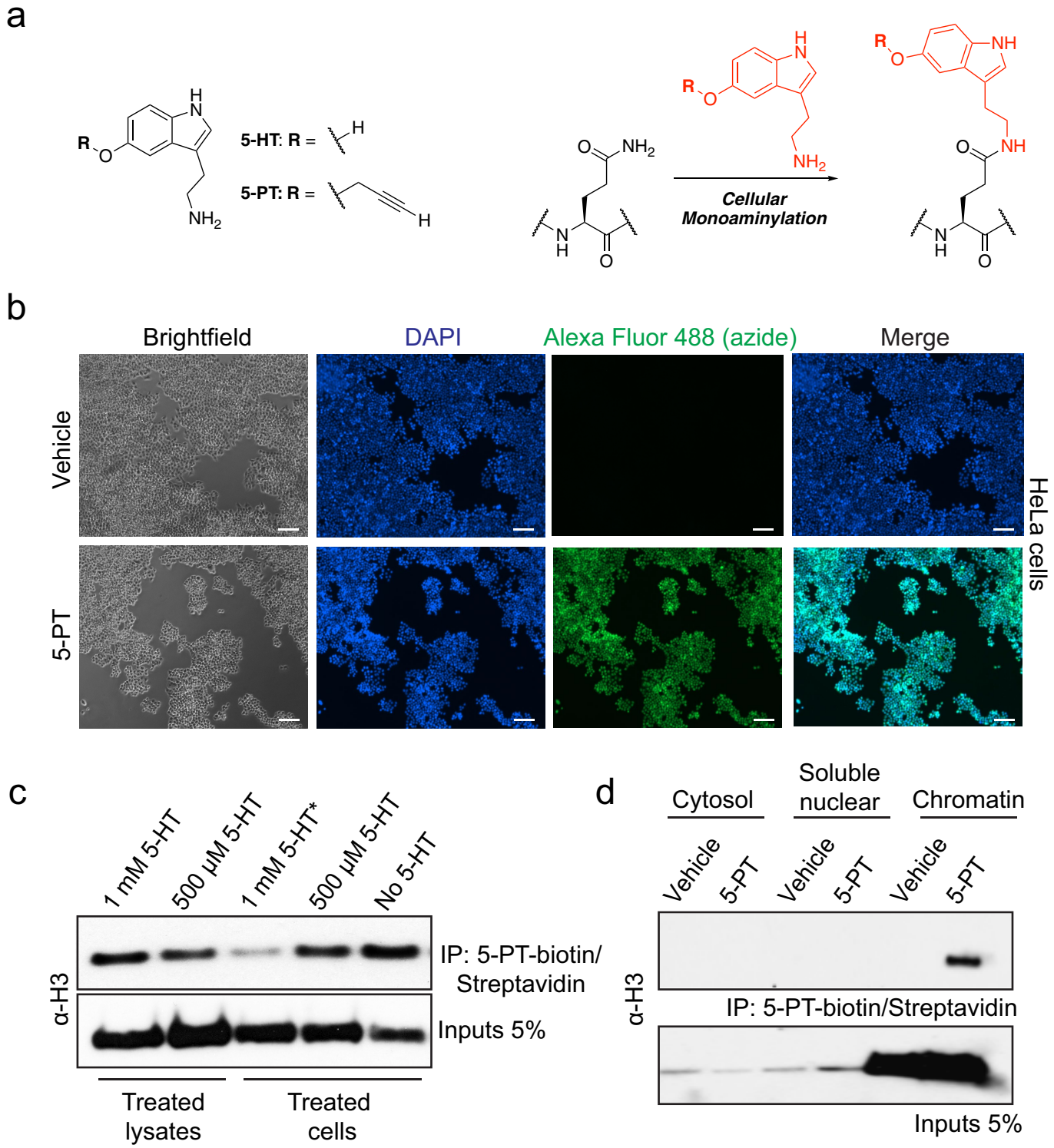
Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

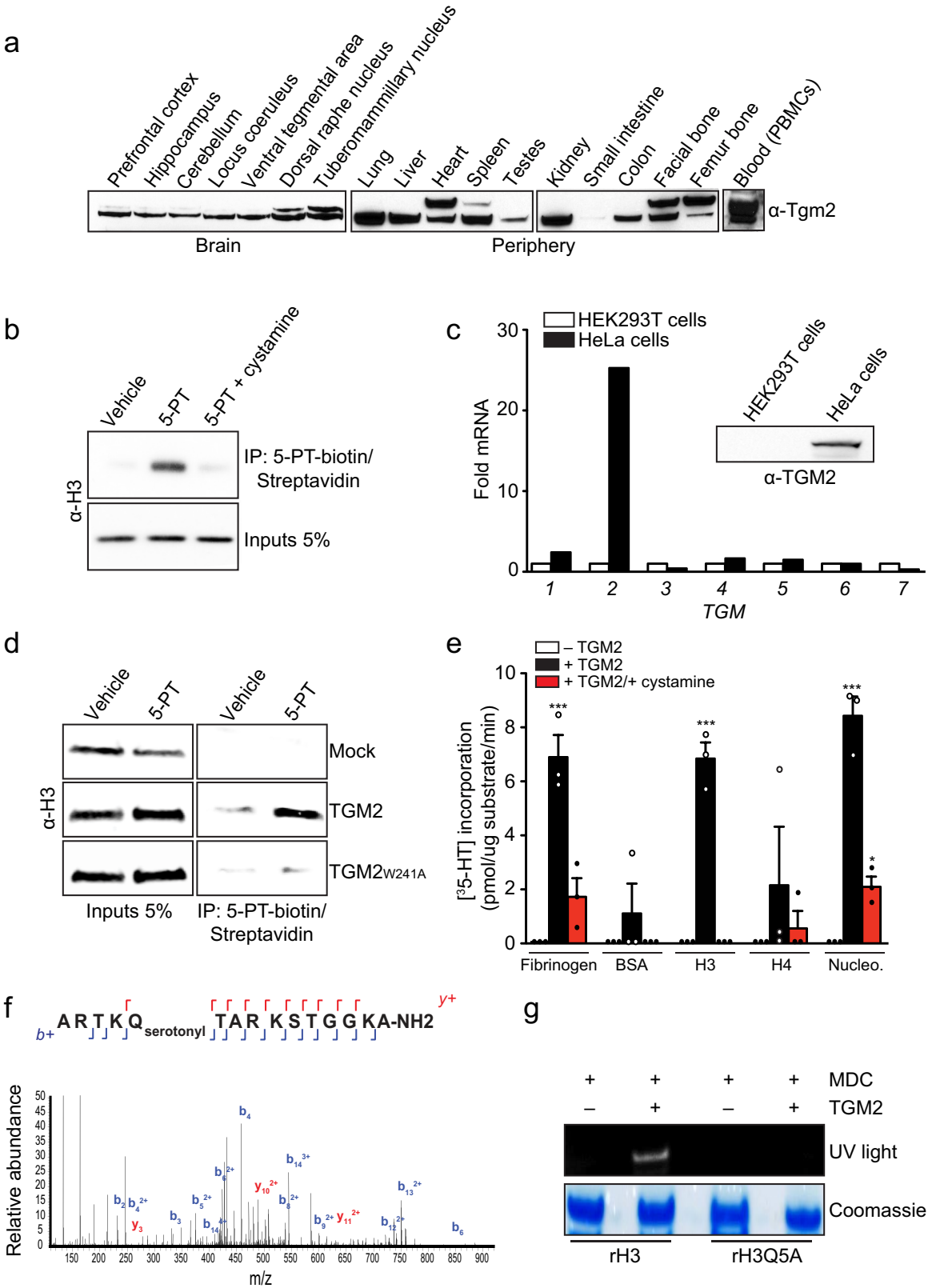
Data from ChIP-seq and RNA-seq experiments have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) database under accession numbers GSE106495 and GSE117910. The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE[48,49] partner repository with the dataset identifier PXD008106. Additional raw data files are available at https://chorusproject.org under project no. 1513. We declare that the data supporting findings for this study are available within the article and Supplementary Information (Supplementary Fig. 1). Related data are available from the corresponding author upon reasonable request. No restrictions on data availability apply.

26. Eaton, M. J. & Whittemore, S. R. Autocrine BDNF secretion enhances the survival and serotonergic differentiation of raphe neuronal precursor cells grafted into the adult rat CNS. *Exp. Neurol.* **140**, 105–114 (1996).
27. Ollivier, N. et al. Tidbits for the synthesis of bis(2-sulfanylethyl)amido (SEA) polystyrene resin, SEA peptides and peptide thioesters. *J. Pept. Sci.* **20**, 92–97 (2014).
28. Dann, G. P. et al. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* **548**, 607–611 (2017).
29. Nguyen, U. T. et al. Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat. Methods* **11**, 834–840 (2014).
30. Simon, M. D. & Shokat, K. M. A method to site-specifically incorporate methyl-lysine analogues into recombinant proteins. *Methods Enzymol.* **512**, 57–69 (2012).
31. Debelouchina, G. T., Gerecht, K. & Muir, T. W. Ubiquitin utilizes an acidic surface patch to alter chromatin structure. *Nat. Chem. Biol.* **13**, 105–110 (2017).
32. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42 (1998).
33. Dyer, P. N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375**, 23–44 (2004).

34. Wojcik, F. et al. Functional crosstalk between histone H2B ubiquitylation and H2A modifications and variants. *Nat. Commun.* **9**, 1394 (2018).
35. Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S. & Coon, J. J. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol. Cell. Proteomics* **11**, 1475–1488 (2012).
36. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489 (1983).
37. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
38. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).
39. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4**, 923–925 (2007).
40. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
41. Schwanhäusser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
42. Sidoli, S., Bhanu, N. V., Karch, K. R., Wang, X. & Garcia, B. A. Complete workflow for analysis of histone post-translational modifications using bottom-up mass spectrometry: from histone extraction to data analysis. *J. Vis. Exp.* **17**, e54112 (2016).
43. Tsankova, N. M. et al. Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nat. Neurosci.* **9**, 519–525 (2006).
44. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
45. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
46. 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).
47. Gundemir, S. & Johnson, G. V. Intracellular localization and conformational state of transglutaminase 2: implications for cell death. *PLoS ONE* **4**, e6123 (2009).
48. Vizcaíno, J. A. et al. A guide to the Proteomics Identifications Database proteomics data repository. *Proteomics* **9**, 4276–4283 (2009).
49. Vizcaíno, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44** (D1), D447–D456 (2016).

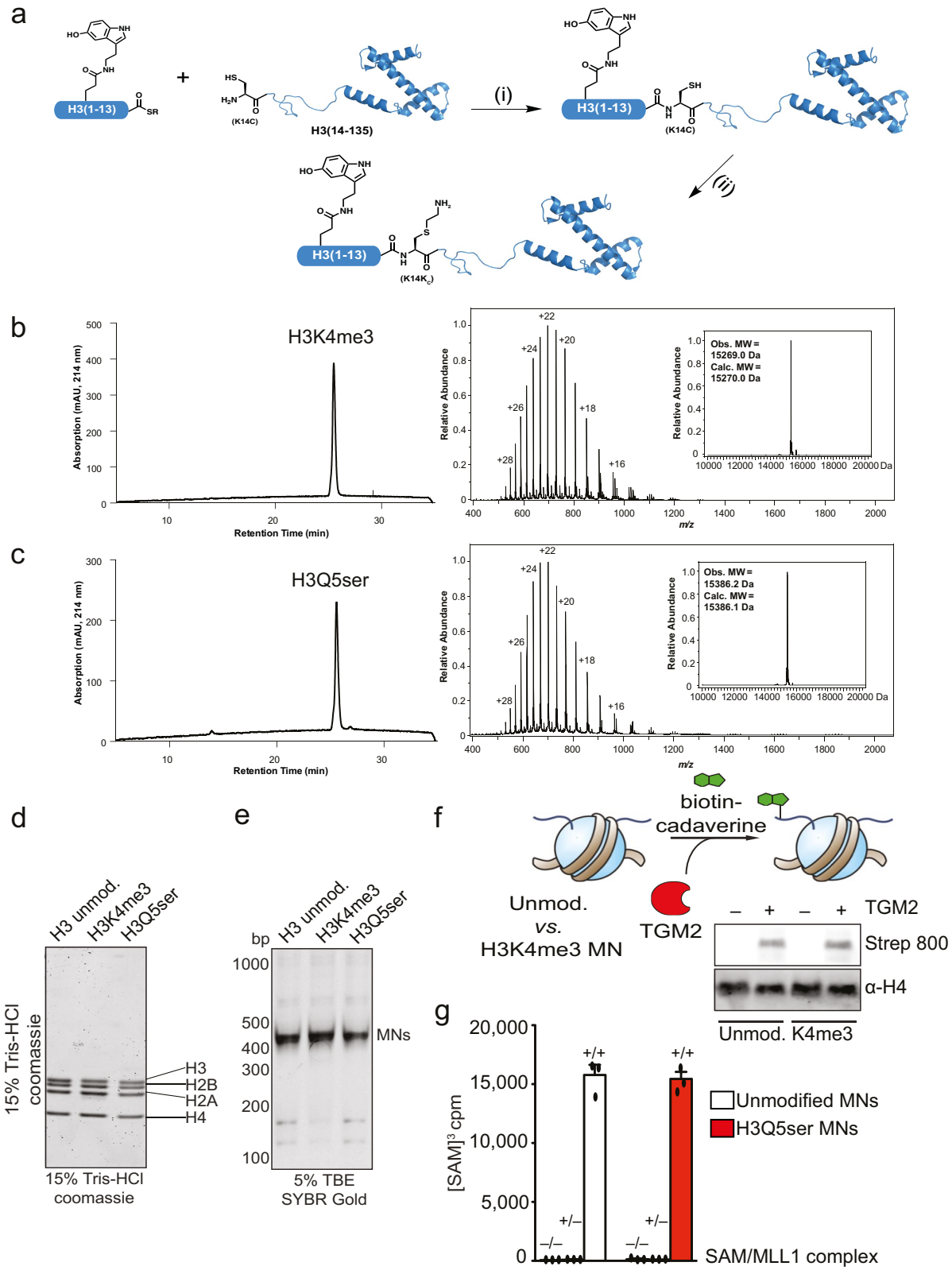


Extended Data Fig. 1 | Bioorthogonal labelling by 5-PT identifies H3 serotonylation in chromatin. a, Structures of 5-HT and 5-PT in transamidation reactions. **b**, Immunofluorescence of intracellular 5-PT in HeLa cells (bottom) after exogenous application of the molecule (versus vehicle, top). Intracellular 5-PT was imaged in fixed cells after chemical labelling with Alexa Fluor 488 azide. DAPI was used as a nuclear co-stain. Results confirmed in at least two independent experiments. Scale bars, 500 μm. **c**, Cell versus lysate donor competition assays indicating that application of excess 5-HT to live HeLa cells, but not to processed lysates, before chemical labelling and 5-PT-based pulldowns results in loss of H3 signal after immunoprecipitation. Input and immunoprecipitate western blots are shown. **d**, Cellular fractionation analysis identifying H3 serotonylation post-immunoprecipitation in HeLa cell chromatin, but not in soluble nuclear or cytosolic fractions. Input and immunoprecipitate western blots are shown. H3 results confirmed in at least three independent experiments.



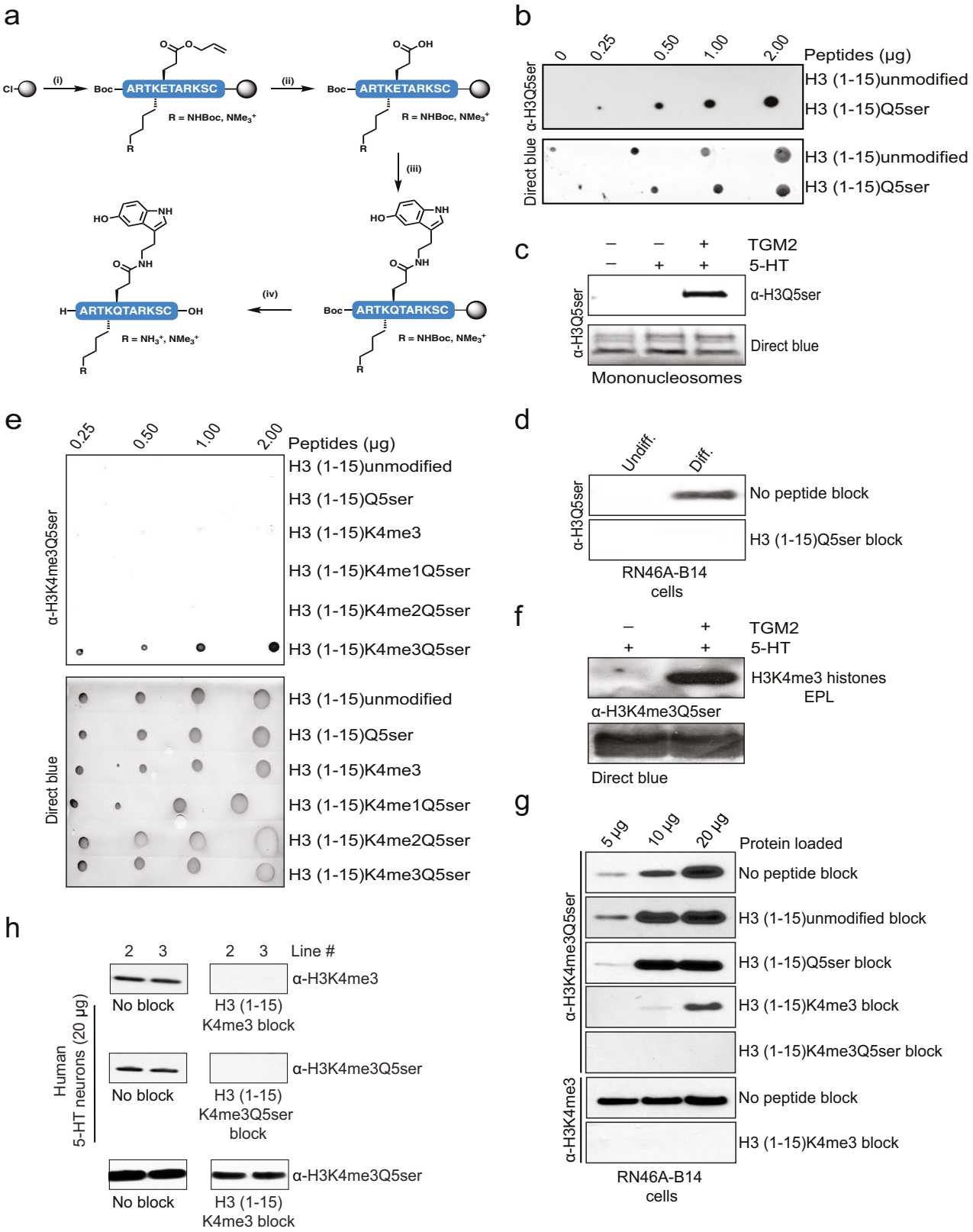
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | TGM2 is both necessary and sufficient to promote H3 serotonylation in cells. a, Western blot analysis of TGM2 expression across mouse brain and peripheral tissues (including blood) validate TGM2 as a ubiquitously expressed enzyme in mammals. Whereas two isoforms of TGM2 exist and are expressed differentially across tissues, the functions of these two variants remain unknown ($n = 1$ per brain region or organ tissue). **b**, Bioorthogonal labelling of 5-PT-modified H3 from HeLa cell extracts in the absence or presence of the transglutaminase inhibitor cystamine (4 mM). Input and immunoprecipitate western blots are shown. Results confirmed in at least two independent experiments. **c**, Quantitative PCR with reverse transcription (RT–qPCR) analysis of TGM genes comparing HeLa and HEK 293T cells ($n = 1$ per cell type). Insert western blot validating that TGM2 is differentially expressed across cell types. **d**, Bioorthogonal labelling of 5-PT-modified H3 from mock (non-transfected) versus nuclear expression of TGM2 in HEK 293T cell extracts. Catalytically dead TGM2(W241A) was used as a negative control. Input and immunoprecipitate western blots are shown. Results confirmed in at least two independent experiments. **e**, Recombinant histones (H3 versus H4) and reconstituted nucleosomes, as well as fibrinogen (positive control) and BSA (negative control), were subjected to radioactive serotonylation (5-^{3}HT) reactions in vitro ($n = 3$), in the absence or presence of cystamine (one-way ANOVA: fibrinogen, $F_{2,6} = 33.47$, $^{***}P = 0.0006$; Dunnett's post hoc test comparing +TGM2 and –TGM2, $P < 0.001$; BSA, $F_{2,6} = 1.000$, $P = 0.4219$; H3, $F_{2,6} = 130.0$, $^{****}P < 0.0001$; Dunnett's post hoc test comparing +TGM2 and –TGM2, $P < 0.001$; H4, $F_{2,6} = 0.7345$, $P = 0.5184$; nucleosomes, $F_{2,6} = 88.52$, $^{****}P < 0.0001$; Dunnett's post hoc test comparing +TGM2 and –TGM2, $P < 0.001$, and comparing +TGM2/+cystamine and –TGM2, $P < 0.05$). **f**, LC–MS/MS analysis of a TGM2-transamidated H3 peptide ($H3_{1–15}$), identifying Gln5 as a substrate for serotonylation. The *y* and *b* series indicate peptide fragments at amide bonds. Results confirmed in at least two independent experiments. **g**, TGM2 monoaminylation assays examining MDC transfer to recombinant wild-type H3 versus H3(Q5A), indicating Q5 as a dominant reactive amino acid substrate for H3. Coomassie staining of loading is shown. H3 results confirmed in at least three independent experiments. Data are mean ± s.e.m.



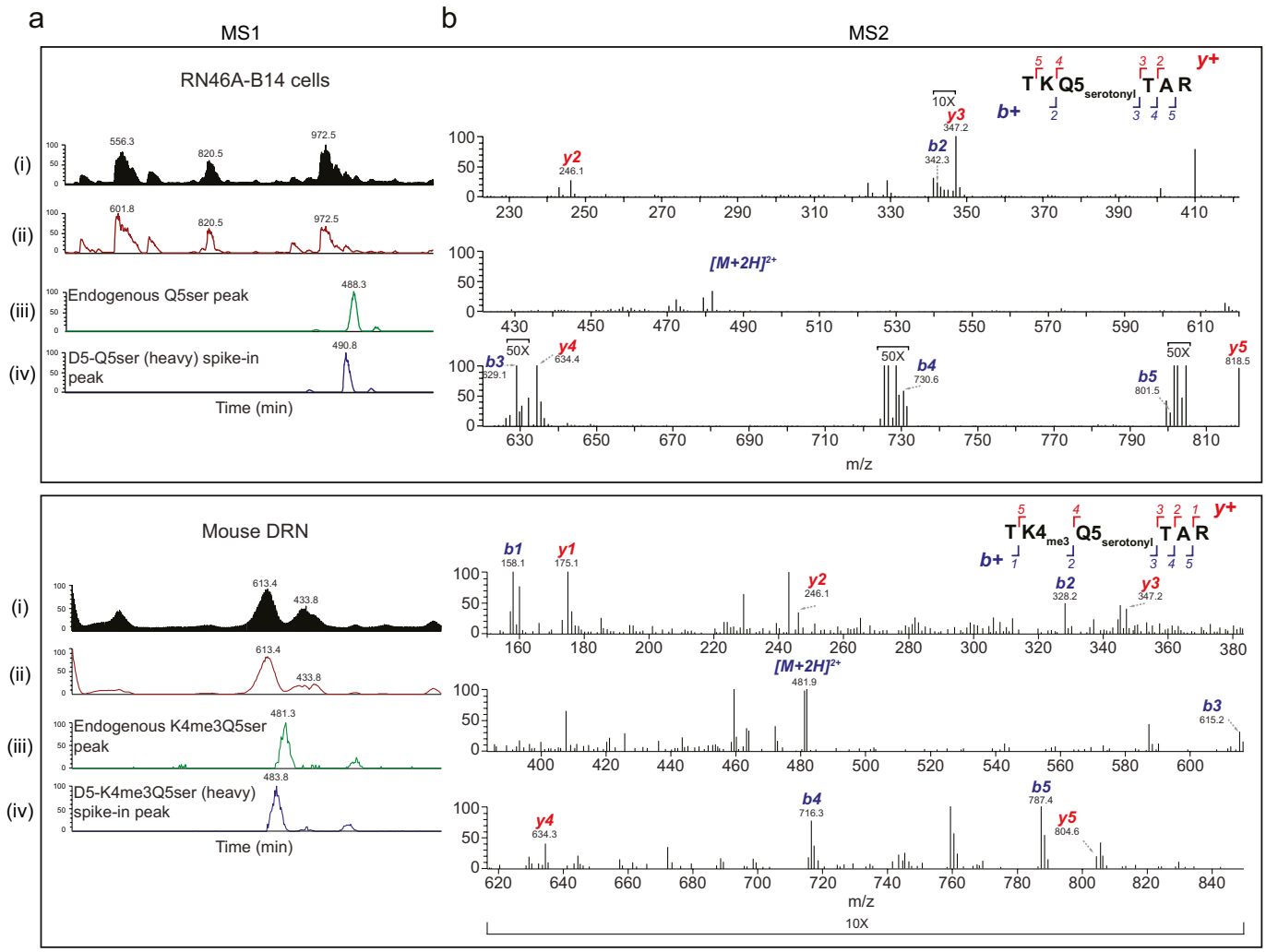
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Histone semisynthesis and nucleosome assembly. **a**, Representative semisynthesis of serotonylated histone H3 comprising (i) native chemical ligation between the serotonylated $H3_{1–13}$ α-thioester and a truncated histone $H3_{14–135}$, followed by (ii) cysteine alkylation to give K14 thialysine. **b**, **c**, RP-HPLC and mass spectrometry characterization of semisynthetic histone proteins, including H3K4me3 (**b**) and H3Q5Ser (**c**). Purified proteins were eluted from a C_{18} RP-HPLC column using a gradient of 0–73% solvent B (0.1% TFA in 9:1 acetonitrile:water) in solvent A (0.1% TFA in water) detecting absorption at 214 nm. Mass spectra of purified proteins were deconvoluted (inset) and observed versus calculated masses are shown. Results confirmed in at least three independent experiments. **d**, **e**, Validation of octamer (**d**) and mononucleosome (**e**) assembly post-semisynthesis of unmodified, K4me3 and Q5ser proteins. Results confirmed in at least three independent experiments. **f**, TGM2 monoaminylation assays on unmodified versus H3K4me3 nucleosomes using biotin cadaverine in place of 5-HT ($n = 3$, two-tailed Student's t-test; $t_4 = 0.500$, $P = 0.64$). H4 is provided as a loading control. **g**, MLL1-complex methyltransferase assays on unmodified versus Q5ser mononucleosmes (−SAM/−MLL1, +SAM/−MLL1 and +SAM/+MLL1; $n = 3$, two-tailed Student's t-test, +/+ versus +/+; $t_4 = 0.3100$, $P = 0.76$). Data are mean ± s.e.m.

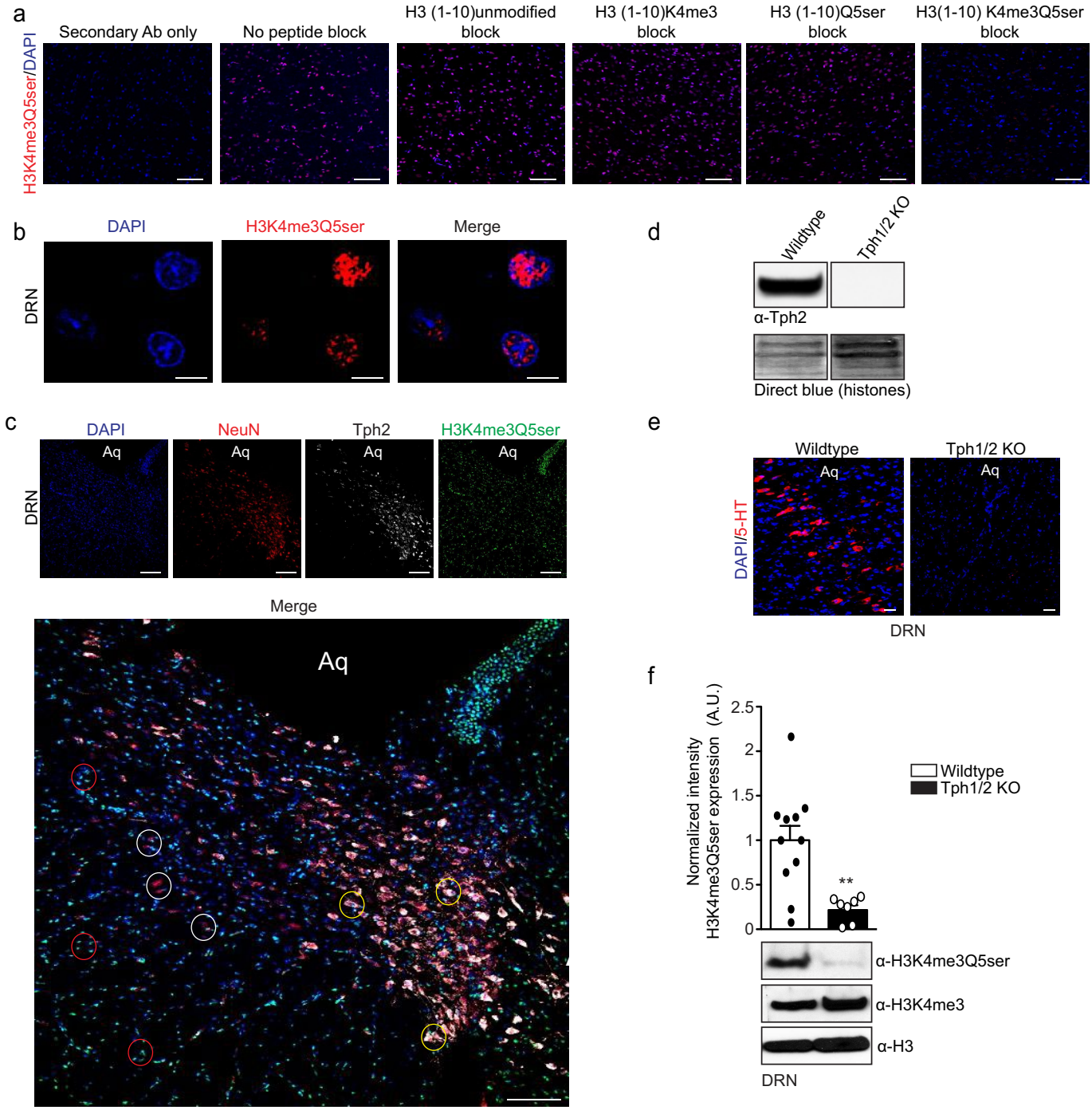


Extended Data Fig. 4 | See next page for caption.

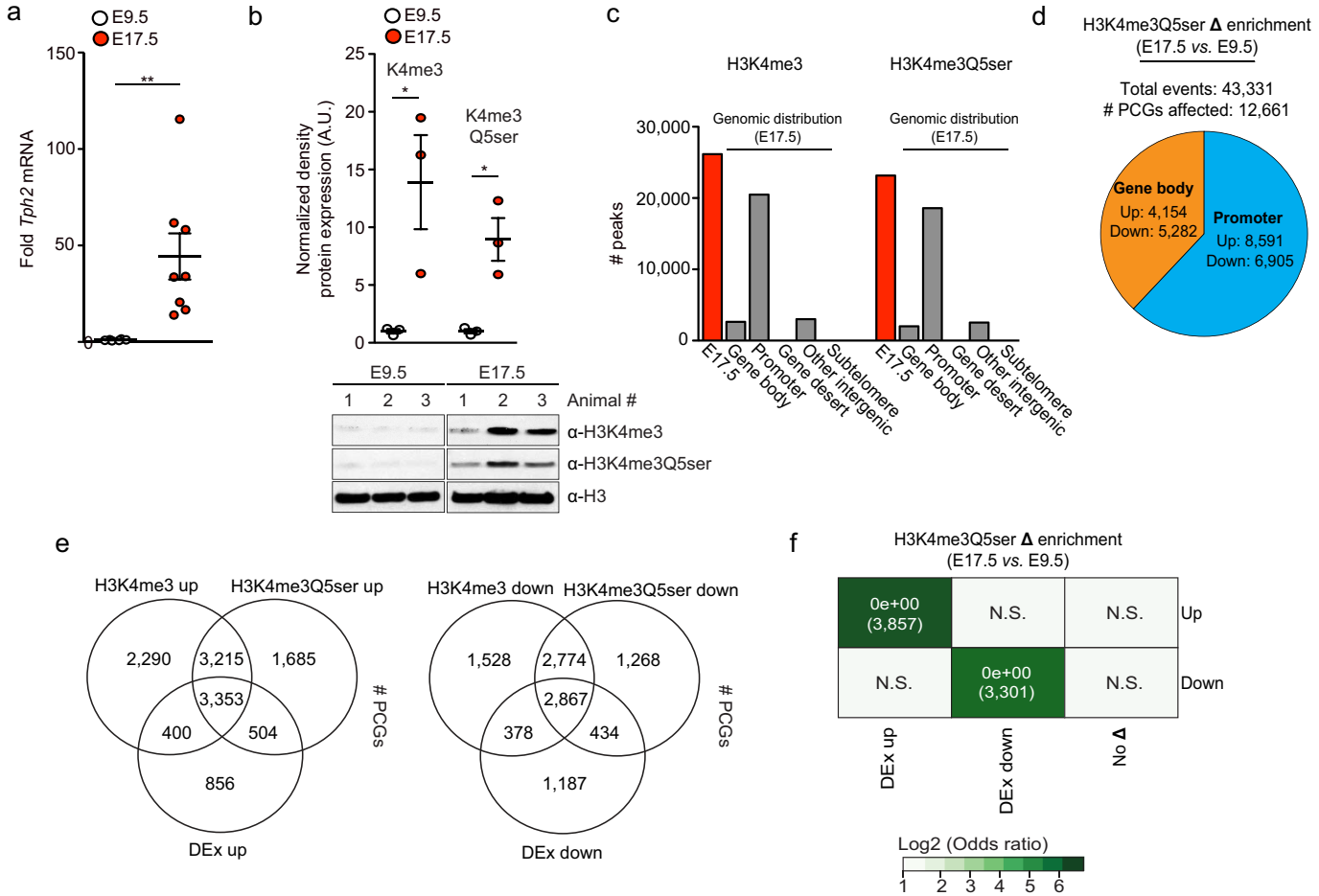
Extended Data Fig. 4 | H3Q5ser and H3K4me3Q5ser antibody validations. a, Synthesis of peptide antigens on 2-Cl trityl resin by (i) iterative Fmoc solid-phase peptide synthesis incorporating Fmoc–Glu(OAII)-OH at position 5 and either Fmoc–Lys(Boc)-OH or Fmoc–Lys(Me_3)-OH at position 4, (ii) followed by Pd(0) deallylation, (iii) 5-HT coupling and (iv) acidolytic cleavage from the resin and global deprotection. Side-chain protecting groups are omitted for clarity. **b**, Peptide dot blot titrations testing the H3Q5ser antibody's reactivity against unmodified versus Q5ser peptides; note that linear signal was observed only with the Q5ser peptide. Direct blue staining was used to control for peptide loading. **c**, Western blot analysis of TGM2 serotonylation assays on unmodified mononucleosomes, which reveals that the H3Q5ser antibody only detects signal when the nucleosomes have been transamidated with 5-HT. Direct blue staining was used to control for protein loading. **d**, Peptide competition western blot analysis of lysates from RN46A-B14 cells indicating the specificity of the H3Q5ser antibody. **e**, Peptide dot blot titrations testing the H3K4me3Q5ser antibody's reactivity against various peptides; note that linear signal was observed only with the K4me3Q5ser peptide. Direct blue staining was used to control for peptide loading. **f**, Western blot analysis of TGM2 serotonylation assays on monomeric EPL K4me3 H3, revealing that the H3K4me3Q5ser antibody only detects signal when the K4me3 histone has been transamidated with 5-HT. Direct blue staining was used to control for protein loading. **g**, Peptide competition western blot analysis of lysates from RN46A-B14 cells indicating the specificity of the H3K4me3Q5ser antibody in vivo; note that a minimum of 20 μg protein was loaded for all other comparisons made throughout the study. **h**, Peptide competition western blot analysis of lysates from hPSC-derived 5-HT neurons indicating the specificity of the H3K4me3Q5ser antibody in human cells. For experiments in **b**–**h**, similar results were confirmed in at least two independent experiments or assays.



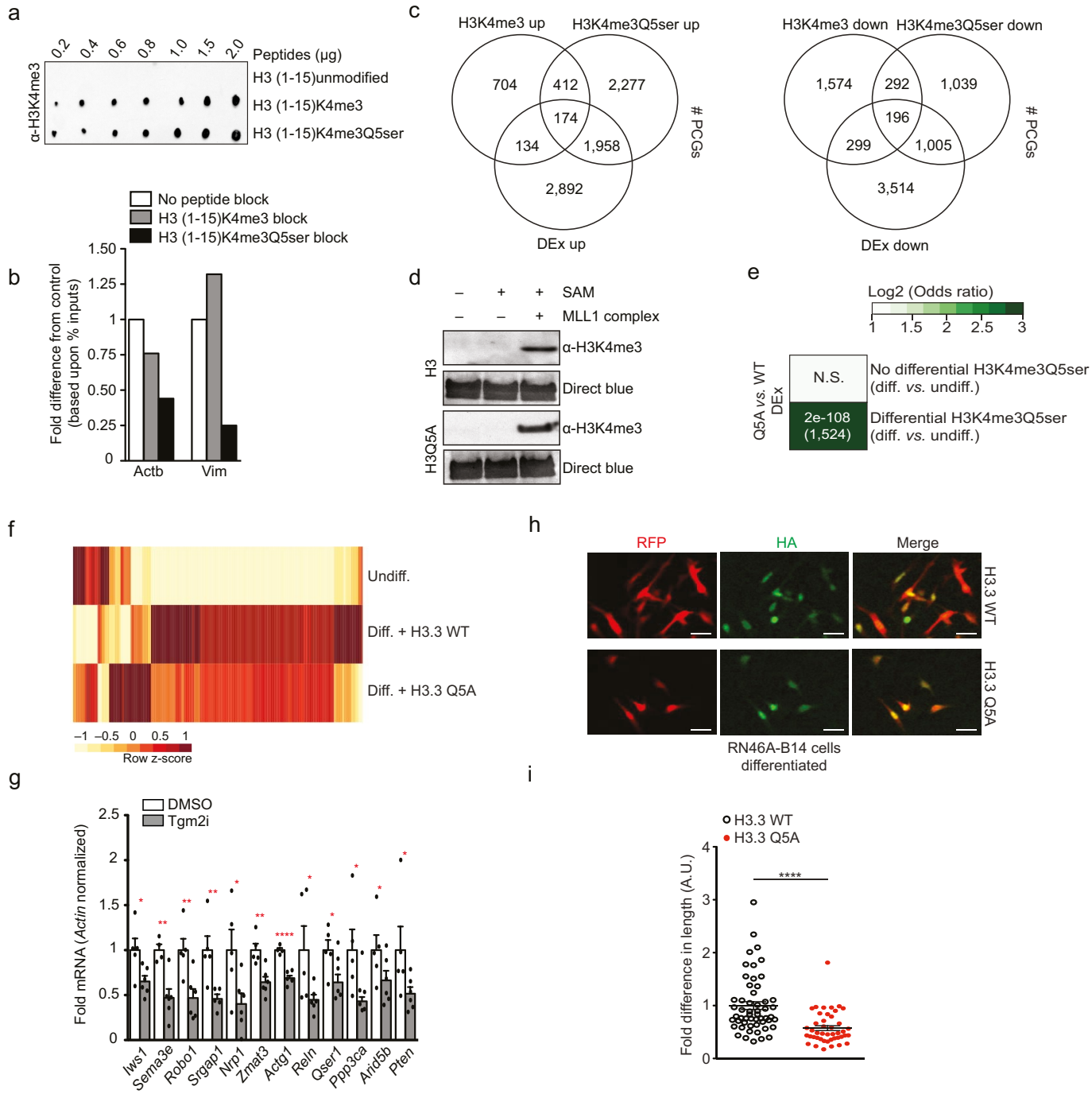
Extended Data Fig. 5 | LC–MS/MS identification of H3 serotonylation in cells and brain. LC–MS/MS analysis of endogenous and synthetic (N-terminally labelled with D5) H3Q5ser or H3K4me3Q5ser peptides in RN46A-B14 cells (post-differentiation, top) and brain (mouse DRN, bottom) following immunoprecipitation using our H3Q5ser and H3K4me3Q5ser antibodies, respectively. **a**, Liquid chromatography–mass spectrometry (LC–MS, MS1) chromatograms showing elution profiles for the samples including the endogenous peptide versus a synthetic peptide standard. Top to bottom: (i) total ion current chromatogram depicting the summed intensity of all signals; (ii) base peak chromatogram showing the intensity of the top signal; (iii) extracted ion chromatogram of the *m*/*z* of the endogenous peptide; (iv) extracted ion chromatogram of the *m*/*z* value of the synthetic, heavy-labelled peptide. **b**, MS/MS (MS2) spectrum (aligned by *m*/*z*, and magnified where indicated on either side of the precursor ion) of the H3Q5ser and H3K4me3Q5ser peptides obtained by HCD fragmentation. *b*+ and *y*+ fragment ions were annotated manually, and the top insets show good coverage. Results confirmed in at least two independent experiments (RN46A-B14 cells pre- versus post-differentiation and/or in mouse brain).



Extended Data Fig. 6 | Euchromatic H3K4me3Q5ser distribution in DRN. a, Immunofluorescence of H3K4me3Q5ser in brain (DRN) with or without peptide competition as indicated. DAPI was used as a nuclear co-stain and a secondary-antibody-only control is also included. Scale bars, 100 μm. **b**, Immunofluorescence of H3K4me3Q5ser in DRN (no peptide block) revealing a euchromatic distribution pattern in the nucleus with near-total exclusion from DAPI-rich chromocentres typical of neurons. Scale bars, 10 μm. **c**, Immunofluorescence images of H3K4me3Q5ser in DRN, counterstained with DAPI, NeuN (a neuronal cell marker) and TPH2 (a marker of serotonergic neurons). Merged images reveal that H3K4me3Q5ser is not only expressed in serotonergic neurons ($NeuN^+TPH2^+$, yellow circles) but also in non-serotonergic neurons ($NeuN^+TPH2^-$, white circles) and in non-neuronal cells ($NeuN^-TPH2^-$, red circles). Aq, aqueduct. Scale bars, 100 μm. **d**, Western blot validation of TPH2 knockout (KO) in DRN of TPH1 and TPH2 KO mice. Direct blue was used to control for loading. **e**, Immunofluorescence validation of 5-HT depletion in DRN of TPH1 and TPH2 knockout mice versus wild-type littermate controls. Scale bars, 100 μm. For all immunofluorescence experiments (**a**–**e**), results were confirmed in at least two independent experiments. **f**, Western blot validation that TPH1 and TPH2 knockout results in loss of H3K4me3Q5ser signal in DRN ($n = 11$ wild type versus $n = 7$ TPH1 and TPH2 knockout, two-tailed Student's t-test; $t_{16} = 3.425$, $^{**}P = 0.0035$); no effects on H3K4me3 or total H3 expression were observed. Data are mean ± s.e.m.

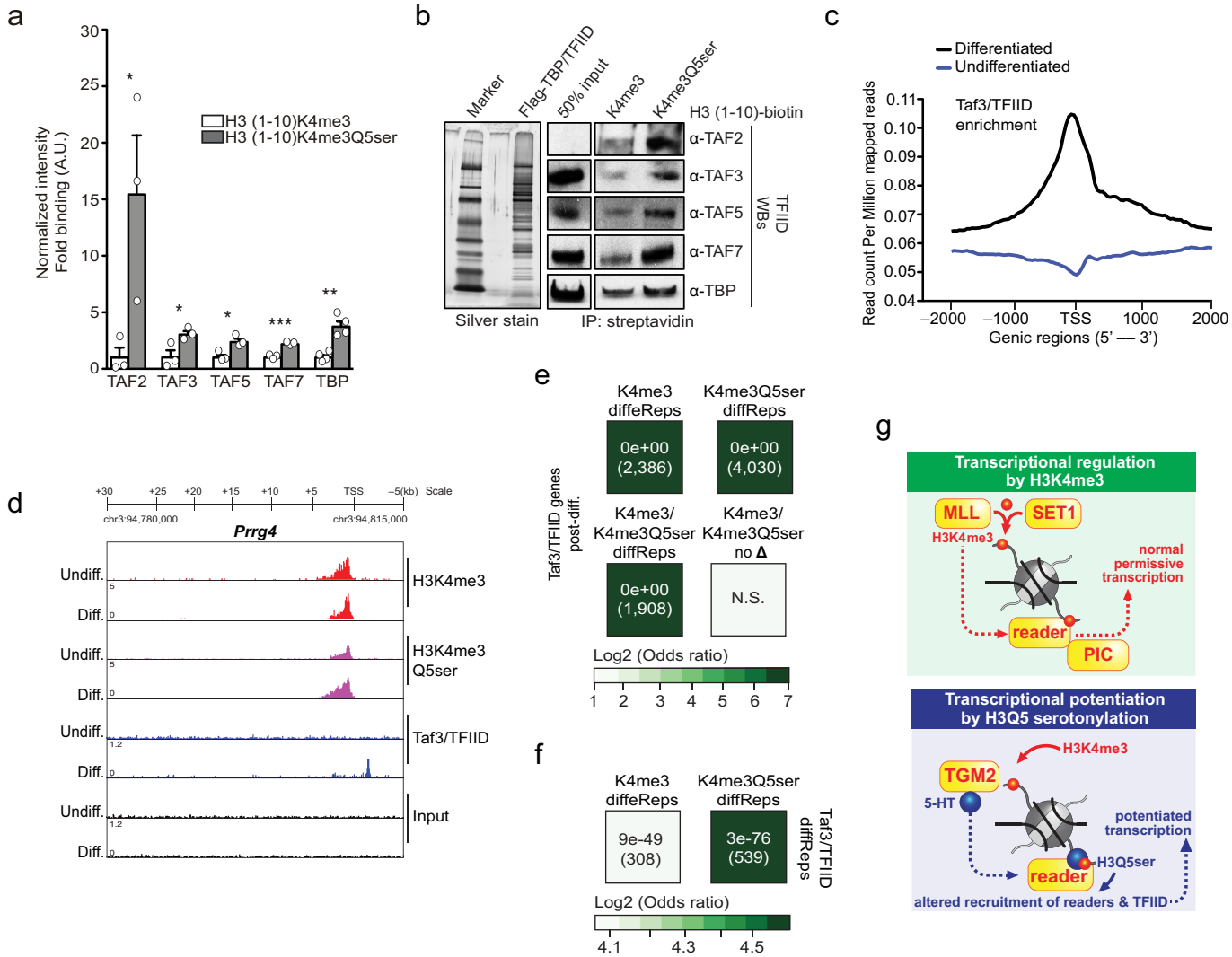


Extended Data Fig. 7 | H3K4me3Q5ser positively correlates with gene expression in developing mouse brain. a, RT–qPCR analysis of *Tph2* gene expression in embryonic mouse brain (E9.5 ($n = 6$) versus E17.5 ($n = 8$), two-tailed Student's *t*-test; $t_{12} = 3.100$, $^{**}P = 0.0092$). *Actb* was used as a normalization control. **b**, Quantitative western blot analysis of H3K4me3 and H3K4me3Q5ser expression in embryonic mouse brain at E9.5 versus E17.5 (two-tailed Student's *t*-test, two-sided, $n = 3$ independent biological replicates per age; K4me3: $t_4 = 3.168$, $^{*}P = 0.0339$; K4me3Q5ser: $t_4 = 4.277$, $^{*}P = 0.0129$). H3 was used as a loading control. **c**, MACS v.2.1.1-based peak-calling for H3K4me3 versus H3K4me3Q5ser ChIP-seq data in embryonic mouse brain at E17.5 (FDR <0.05, FC >1.2 cut-offs applied after adjusting for multiple comparisons; $n = 3$ independent biological replicates per age). **d**, Analysis of H3K4me3Q5ser differential enrichment comparing E17.5 and E9.5 embryonic mouse brain using diffReps (FDR <0.05, FC >1.2 cut-offs applied after adjusting for multiple comparisons; $n = 3$ independent biological replicates per age). Pie chart indicates distribution of genic differential enrichment events for the mark comparing promoter and gene body regions. **e**, Overlap of differential enrichment events at protein-coding genes for H3K4me3 and H3K4me3Q5ser versus differential gene expression during mouse brain development indicates significant associations between the two marks, as well as positive correlations with gene expression. **f**, Odds-ratio analysis (using Fisher's exact tests) of overlapping genes displaying differential H3K4me3Q5ser enrichment versus differential gene expression ($n = 3$ independent biological replicates per age, FDR <0.05 cut-off applied after adjusting for multiple comparisons) indicating positive correlations between mark enrichment and permissive gene expression during development. Insert numbers indicate respective *P* values for associations, followed by the number of protein-coding genes overlapping per significant category. Data are mean ± s.e.m.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Functional validations of the role of H3 serotonyl in permissive gene expression. a, Peptide dot blot titrations testing the ability of the H3K4me3 antibody to detect the mark in the absence or presence of Q5ser; note that Q5ser does not occlude antibody recognition of the mark. Results confirmed in at least two independent experiments. **b**, Peptide competition qChIP validation of H3K4me3Q5ser enrichment at two target loci (*Actb* and *Vim*) identified via ChIP-seq in RN46A-B14 cells (undifferentiated, $n = 1$ per group). **c**, Overlap of differential enrichment events at protein-coding genes for H3K4me3 and H3K4me3Q5ser versus differential gene expression in response to RN46A-B14 cell differentiation indicates positive correlations between H3K4me3Q5ser gene expression in the absence of significant changes in H3K4me3 enrichment (FDR <0.05, FC >1.2 cut-offs applied after adjusting for multiple comparisons. $n = 3$ independent biological replicates per differentiation state per antibody). **d**, Recombinant methylation assays (MLL1-complex-mediated) on wild-type H3 versus H3(Q5A), followed by western blot for H3K4me3, indicating that the Q5A mutation does not affect methylation capacity at K4. Direct blue was used to control for protein loading. Results confirmed in at least two independent experiments. **e**, Odds-ratio analysis (using Fisher's exact tests) of overlapping genes displaying differential H3K4me3Q5ser enrichment (or not) versus differential gene expression (DEx, $n = 5$ (H3.3(WT)) versus $n = 6$ (H3.3(Q5A)) independent biological replicates per virus), FDR <0.05 cut-off applied after adjusting for multiple comparisons) comparing differentiated RN46A-B14 cells expressing either H3.3(WT) or H3.3(Q5A). Insert numbers indicate respective *P* values for associations, followed by the number of protein-coding genes overlapping per significant category. **f**, Heat map of RNA-seq data comparing undifferentiated and differentiated ($n = 5$ H3.3(WT) versus $n = 6$ H3.3(Q5A)-expressing) RN46A-B14 cells using normalized RNA expression values (averaged between replicates) to generate *z*-scores for each row. The genes that displayed differential enrichment during differentiation are represented, along with altered gene expression (see Fig. 3), and along with opposing regulation in the context of H3.3(Q5A). These genes were found to significantly enrich for pathways associated with axon guidance signalling via KEGG analysis (see Supplementary Table 28). **g**, RT–qPCR analysis of candidate gene expression in RN46A-B14 cells ($n = 5$ DMSO versus $n = 6$ LDN 27219/Tgm2i; one-tailed Student's *t*-test; *Iws1*: $t_9 = 2.559$, $^*P = 0.0154$; *Sema3e*: $t_8 = 3.982$, $^{**}P = 0.0020$; *Robo1*: $t_9 = 3.344$, $^{**}P = 0.0043$; *Srgap1*: $t_8 = 3.312$, $^{**}P = 0.0053$; *Nrp1*: $t_9 = 2.452$, $^*P = 0.0183$; *Zmat3*: $t_9 = 3.820$, $^{**}P = 0.0020$; *Actg1*: $t_8 = 7.836$, $^{****}P < 0.0001$; *Reln*: $t_9 = 2.209$, $^*P = 0.0273$; *Qser1*: $t_8 = 2.513$, $^*P = 0.0181$; *Ppp3ca*: $t_8 = 2.418$, $^*P = 0.0210$; *Arid5b*: $t_9 = 1.754$, $^*P = 0.0567$; *Pten*: $t_9 = 1.936$, $^*P = 0.0424$). *Actb* was used as a normalization control. **h**, Immunofluorescence of RN46A-B14 cells infected during differentiation with lentiviruses expressing either wild-type H3.3–HA or H3.3(Q5A)–HA. Results confirmed at least three independent coverslips per viral treatment. Scale bars, 20 μm. **i**, Neurite outgrowth analysis examining RN46A-B14 cellular length post-differentiation ($n = 54$ H3.3(WT) versus $n = 44$ H3.3(Q5A)-expressing cells; two-tailed Student's *t*-test, $t_{96} = 4.664$, $^{****}P < 0.0001$). Data are mean ± s.e.m.



Extended Data Fig. 9 | Genome-wide associations between H3K4me3Q5ser and TFIID. a, Western blot quantifications related to Fig. 4a (Student's *t*-tests, two-sided) of modified $H3_{1-10}$ peptide immunoprecipitations from HeLa nuclear extracts (TAF2: $n = 3$ per peptide, $t_4 = 2.724$, $^*P = 0.05$; TAF3: $n = 3$ per peptide, $t_4 = 2.920$, $^*P = 0.04$; TAF5: $n = 3$ per peptide, $t_4 = 3.685$, $^*P = 0.02$; TAF7: $n = 3$ per peptide, $t_4 = 8.885$, $^{***}P = 0.0009$; TBP: $n = 4$ per peptide, $t_6 = 5.383$, $^{**}P = 0.0017$). **b**, Modified $H3_{1-10}$ peptide immunoprecipitations against purified TFIID from Flag-tagged TBP expressing soluble HeLa nuclear extracts (left, silver stain of the purified complex), followed by western blots (right). Inputs are provided ($n = 1$ per protein examined). **c**, Gene plot (ngs.plot) of TAF3–TFIID enrichment (fold change (expressed in $\log_2$) versus input) comparing signals pre- and post-differentiation ($n = 3$ independent biological replicates per differentiation states). **d**, IGV genome browser tracks of the *Prrg4* locus for H3K4me3, H3K4me3Q5ser and TAF3–TFIID (versus DNA input) in RN46A-B14 cells pre- and post differentiation (example locus was chosen on the basis of MACS v.2.1.1- and diffReps-based statistical comparisons). Odds-ratio analysis (using Fisher's exact tests) of overlapping genes displaying differential H3K4me3 or H3K4me3Q5ser enrichment (or those genes containing differential sites for both or neither of the two marks, FDR <0.05, FC >1.2) versus TAF3 peaks (**e**; $n = 3$ independent biological replicates per differentiation state, MACS v.2.1.1, FDR <0.05, FC >1.2 cut-offs applied after adjusting for multiple comparisons; normalized to respective DNA inputs) or differential TAF3–TFIID enrichment (**f**; $n = 3$ independent biological replicates per differentiation state, diffReps, FDR <0.05, FC >1.2 cut-offs applied after adjusting for multiple comparisons) post-differentiation. Insert numbers indicate respective *P* values for associations, followed by the number of protein-coding genes overlapping per significant category. **g**, Model of the effect of Q5ser on K4me3-mediated gene expression in vivo; our data suggest that the presence of combinatorial H3K4me3Q5ser alters interactions with certain K4me3 reader proteins, such as the TFIID complex, to potentiate or stabilize permissive gene expression in mammalian cells.

nature research

Corresponding author(s): Ian Maze; 2017-11-14775C

Reporting Summary

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

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No custom algorithms were used.
Data analysis	For most statistical comparisons, Prism 5 for Mac OS X was used. ChIP-seq peak calling and differential analysis For H3K4me3, H3K4me3Q5ser and Taf3 ChIP-seq in RN46A-B14 cells, raw sequencing reads were aligned to the rat genome (rn6) using default settings of Bowtiev2.2.0. Only uniquely mapped reads were retained. Alignments were filtered using SAMtoolsv0.1.19 to remove duplicated reads. Peak-calling–normalized to respective inputs–was performed using MACSv2.1.144 with default settings and filtered for FDR < 0.05 and fold change > 1.2; the window size was set as 300 bp. Raw sequencing reads from mouse brains (H3K4me3 vs. H3K4me3Q5ser) were aligned to the mouse genome (mm10) using default settings of HISATv0.1.6b. Only uniquely mapped reads were retained. Alignments were filtered using SAMtoolsv0.1.19 to remove duplicate reads, and peak-calling–normalized to respective inputs–was performed using MACSv2.1.1 with default settings and filtered for FDR < 0.05 and fold change > 1.2. Raw sequencing reads from human PSCs vs. ¬5-HT neurons (H3K4me3Q5ser) were aligned to the human genome (hg38) using default settings of HISAT2v2.1.0. Again, only uniquely mapped reads were retained. Alignments were filtered using SAMtoolsv1.8 to remove duplicated reads, and peak-calling–normalized to respective inputs–was performed using MACSv2.1.1 with default settings and filtered for FDR < 0.05 and fold change > 1.2. Differential analyses of H3K4me3 and H3K4me3Q5ser (for all cells/tissues examined) were performed using diffReps19 with a window size of 1 kb. A default q-value cutoff of 0.05 was used, followed by a fold change cutoff of 1.2. Peaks and differential sites were further annotated to nearby genes or intergenic regions using the region analysis tool from the diffReps package.

RNA-seq analysis and correlations with ChIP-seq
Raw sequencing reads from RN46A-B14 cells were mapped to rn6 using TopHat2v2.1.045. Counts of reads mapping to genes were obtained using HTSeq-count against Ensembl v81 annotation. Differential expression analysis was done using the DESeq2 package46 (v1.6.3) at an FDR cutoff of 0.05. Raw RNA-seq reads obtained from mouse brains were mapped to mm10 using HISAT v0.1.6b. StringTie v1.2.1 was used to obtain gene-level FPKM values against Ensembl v84. Raw RNA-seq reads obtained from human cells were mapped to the human genome (hg38) using HISAT2v2.1.0. Counts of reads mapping to genes were obtained using HTSeq-count against Ensembl v90 annotation. Differential expression analysis was done using the DESeq2 package (v1.6.3), again with a FDR<0.05 cutoff. RNA-seq differential lists were compared to ChIP-seq differential lists based on gene names, and heat maps were drawn using the GeneOverlap (https://github.com/shenlab-sinai/geneoverlap/) package. Note that a more conservative fold change cutoff of > 2.5 was used to generate odds ratio heat maps for 5-HT neuron/PSC diffReps vs. gene expression in order to avoid including false positive 'down' events in selected comparisons.

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

Data

Policy information about availability of data

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

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

Data from ChIP- and RNA-seq experiments have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) database under accession numbers GSE106495 and GSE117910. The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD008106. Additional raw data files are available on https://chorusproject.org under project no. 1513. We declare that the data supporting findings for this study are available within the article and Supplementary Information (please see Supplementary Fig.1 for source data). Related data are available from the corresponding author upon reasonable request. No restrictions on data availability apply.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	Adequate sample sizes are generally determined based upon inter-sample variability. Throughout the manuscript, we determined the significance of results based upon a general confidence interval of 95%. We do not include specific justifications of sample size within the methods (e.g., power analyses), as sample sizes were based on extensive laboratory experience with these endpoints.
Data exclusions	No data were excluded.
Replication	All biological endpoints were reliably reproduced using numerous biological and technical replicates for each experiment. All novel tools used in this study were extensively validated, as demonstrated in the manuscript submission.
Randomization	Where appropriate, cells or animals were randomly assigned to groups (segregated by differentiation conditions, chemical compound/drug/viral treatments and/or genotype). Furthermore, tissues collected from animals for ChIP-seq analyses were randomly pooled to provide ample tissue for biochemical procedures and to minimize variance.
Blinding	Where necessary, investigators were blinded to conditions such as genotype, viral treatment conditions, etc.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants

Methods

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

Unique biological materials

Policy information about availability of materials

Obtaining unique materials

Besides the H3Q5ser and H3K4me3Q5ser antibodies (described below), no unique materials were used. We declare that the data supporting findings for this study are available within the article and Supplementary Information. Related data are available from the corresponding author upon reasonable request. No restrictions on data availability apply. The H3Q5ser and H3K4me3Q5ser antibodies were produced in collaboration with EMD Millipore and will be released upon publication.

Antibodies

Antibodies used

All primary antibodies used in this study were validated, where appropriate, for use in ChIP, immunoblotting and/or immunocytochemistry (ICC)/immunohistochemistry (IHC) experiments:

H3 general (Abcam ab1791; lot#GR293151-1–1:10,000), H3K4me3 (Abcam ab8580; Lot#GR273043-6–1:1,000), H3Q5ser (Millipore ABE1791; lot#RA1690–1:500), H3K4me3Q5ser (Millipore ABE2589; Lot#RA1608107–1:500), H3K27me3 (Cell Signaling 9733S; Lot#8/used for ChIP-seq), TGM2 (Abcam ab2386; Lot#GR312304-3–1:500), TAF2 (Bethyl A302-132A–1:500), TAF3 (Abcam ab188332; Lot#GR220332-1/used for blotting–1:500), TAF3 (Millipore 07-1802; Lot#3008423/used for ChIP-seq), TBP (Abcam ab8181; lot#GR300917-3–1:1,000), H4 (Abcam ab10158; Lot#GR239048-1–1:1,000), H2A (Abcam ab88770; Lot#GR197571-1–1:1,000), H2B (Abcam ab1790; Lot#GR188969-2–1:1,000), HA (Abcam ab9134; Lot#GR4307-10, or HA-488 Alexa Fluor SC-805; lot#K1716–1:1,000), Tph2 (Abcam ab111828; Lot#GR266431-6, or Millipore, sheep–1:1,000), NeuN (Millipore MAB377; Lot#2683919–1:1,000) and 5-HT (Sigma S5545; Lot#033M4805, or Protos Biotech, rabbit–1:1,000), RFP (Abcam ab65856; lot#GR3181906-1–1:1,000), MAP2 (Abcam, chicken–1:1,000). The TAF5 and TAF7 antibodies were kindly provided by Dr. Robert Roeder (The Rockefeller University–1:500).

Validation

H3Q5ser and H3K4me3Q5ser antibody validations (Extended Data Figure 4)
b, Peptide dot blot titrations testing the α-H3Q5ser antibody's reactivity against unmodified vs. Q5ser peptides; note that linear signal was only observed with the Q5ser peptide. Direct blue (DB) staining was used to control for peptide loading. c, WB analysis of TGM2 serotonylation assays on unmodified mononucleosomes revealing that the α-H3Q5ser antibody only detects signal when the nucleosomes have been transamidated with 5-HT. DB staining was used to control for protein loading. d, Peptide competition WB analysis of lysates from RN46A-B14 cells indicating the specificity of our α-H3Q5ser antibody. e, Peptide dot blot titrations testing the α-H3K4me3Q5ser antibody's reactivity against various peptides; note that linear signal was only observed with the K4me3Q5ser peptide. Direct blue (DB) staining was used to control for peptide loading. f, WB analysis of TGM2 serotonylation assays on monomeric K4me3 (EPL) H3 revealing that the α-H3K4me3Q5ser antibody only detects signal when the K4me3 histone has been transamidated with 5-HT. DB staining was used to control for protein loading. g, Peptide competition WB analysis of lysates from RN46A-B14 cells indicating the specificity of our α-H3K4me3Q5ser antibody in vivo; note that a minimum of 20 µg protein was loaded for all other comparisons made throughout the study. h, Peptide competition WB analysis of lysates from hPSC-derived 5-HT neurons indicating the specificity of our α-H3K4me3Q5ser antibody in human cells. For experiments in b-h, similar results were confirmed in ≥ 2 independent experiments/assay.

All other antibodies used in this study (many of which have been commercially validated) were validated in human/rodent cell lines or tissues via immunoblotting or IHC/IF prior to experimentation.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HEK293T and HeLa cell lines were obtained from the American Type Culture Collection (ATCC). RN46A-B14 cells were provided by Dr. Olivier Berton, ISMMS. hPSCs and hPSC-derived 5-HT neurons were provided by Dr. Fred Gage, Salk Institute. Sf9 cells were provided by Dr. Tom Muir, Princeton University.

Authentication

HEK293T and HeLa cells were authenticated drawing upon our vast experience working with these cells lines (based on cell morphology, growth conditions, etc.).

RN46A-B14 cells were validated for 1) 5-HT production, and 2) differentiation potential (e.g., cell division arrest, neurite formation, alterations in gene expression, etc.), as described throughout the manuscript.

hPSCs and derived 5-HT neurons were validated as discussed in the manuscript (Fig. 2).

	Sf9 cells were validated through their usefulness in the purification of TGM2 protein.
Mycoplasma contamination	We can confirm that all cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Adult male mice (Mus musculus, C57BL/6J, 7-10 weeks old) were purchased from The Jackson Laboratory. Embryonic brains were harvested in utero from pregnant dams (Mus musculus, C57BL/6J, male and female, E9.5-E17.5). Tissues from male Tph 1/2-knockout mice (vs. littermate controls, Mus musculus, 7-10 weeks old) were provided by Drs. Natalia Alenina and Michael Bader at the Max-Delbrück-Center for Molecular Medicine (MDC), Berlin Germany. Animals were group housed on a 12-hour light/dark cycle (lights on from 7:00 A.M. to 7:00 P.M.) at constant temperature (23ºC) with ad libitum access to food and water. All animal protocols (including tissue/blood collections) were approved by the IACUC at both the Icahn School of Medicine at Mount Sinai (ISMMS) and the MDC.
Wild animals	The study did not involve samples collected from wild animals.
Field-collected samples	The study did not involve samples collected from the field.

Human research participants

Policy information about studies involving human research participants

Population characteristics	All iPSC lines were derived from fibroblasts obtained from healthy individuals not being treated for any known conditions.
Recruitment	All subjects were females (young adults).

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.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106495 Secure Reviewer token: yxaxwqqgfhcjdch
Files in database submission	ChIPseq_Input_Day_0_1.fastq.gz ChIPseq_Input_Day_0_2.fastq.gz ChIPseq_Input_Day_0_3.fastq.gzz ChIPseq_Input_Day_7_1.fastq.gz ChIPseq_Input_Day_7_2.fastq.gz ChIPseq_Input_Day_7_3.fastq.gz ChIPseq_H3K4me3Q5ser_Day_0_1.fastq.gz ChIPseq_H3K4me3Q5ser_Day_0_2.fastq.gz ChIPseq_H3K4me3Q5ser_Day_0_3.fastq.gz ChIPseq_H3K4me3Q5ser_Day_7_1.fastq.gz ChIPseq_H3K4me3Q5ser_Day_7_2.fastq.gz ChIPseq_H3K4me3Q5ser_Day_7_3.fastq.gz ChIPseq_H3K4me3_Day_0_1.fastq.gz ChIPseq_H3K4me3_Day_0_2.fastq.gz ChIPseq_H3K4me3_Day_0_3.fastq.gz ChIPseq_H3K4me3_Day_7_1.fastq.gz ChIPseq_H3K4me3_Day_7_2.fastq.gz ChIPseq_H3K4me3_Day_7_3.fastq.gz 12_S46_L005_R1_001.fastq.gz 19_S47_L005_R1_001.fastq.gz 4_S48_L005_R1_001.fastq.gz 5_S44_L005_R1_001.fastq.gz 6_S45_L005_R1_001.fastq.gz 7_S49_L005_R1_001.fastq.gz TAF3_5_S7_L002_R1_001.fastq.gz TAF4_6_S8_L002_R1_001.fastq.gz TAF5_12_S9_L002_R1_001.fastq.gz TAF6_19_S10_L002_R1_001.fastq.gz TAF7_4_S11_L002_R1_001.fastq.gz

TAF8_7_S12_L002_R1_001.fastq.gz
INPUT_12_S18_L003_R1_001.fastq.gz
INPUT_16_S19_L003_R1_001.fastq.gz
INPUT_19_S20_L003_R1_001.fastq.gz
INPUT_2_S13_L003_R1_001.fastq.gz
INPUT_4_S14_L003_R1_001.fastq.gz
INPUT_5_S15_L003_R1_001.fastq.gz
INPUT_6_S16_L003_R1_001.fastq.gz
INPUT_7_S17_L003_R1_001.fastq.gz
IP_12_S26_L004_R1_001.fastq.gz
IP_16_S27_L004_R1_001.fastq.gz
IP_19_S28_L004_R1_001.fastq.gz
IP_2_S21_L004_R1_001.fastq.gz
IP_4_S22_L004_R1_001.fastq.gz
IP_5_S23_L004_R1_001.fastq.gz
IP_6_S24_L004_R1_001.fastq.gz
IP_7_S25_L004_R1_001.fastq.gz
Pool1_Dual_12_S9_L002_R1_001.fastq.gz
Pool1_Dual_19_S10_L002_R1_001.fastq.gz
Pool1_Dual_4_S5_L002_R1_001.fastq.gz
Pool1_Dual_5_S6_L002_R1_001.fastq.gz
Pool1_Dual_6_S7_L002_R1_001.fastq.gz
Pool1_Dual_7_S8_L002_R1_001.fastq.gz
Pool2_H3K4me3_12_S15_L003_R1_001.fastq.gz
Pool2_H3K4me3_19_S16_L003_R1_001.fastq.gz
Pool2_H3K4me3_4_S11_L003_R1_001.fastq.gz
Pool2_H3K4me3_5_S12_L003_R1_001.fastq.gz
Pool2_H3K4me3_6_S13_L003_R1_001.fastq.gz
Pool2_H3K4me3_7_S14_L003_R1_001.fastq.gz
Pool4_Input_12_S27_L005_R1_001.fastq.gz
Pool4_Input_19_S28_L005_R1_001.fastq.gz
Pool4_Input_4_S23_L005_R1_001.fastq.gz
Pool4_Input_5_S24_L005_R1_001.fastq.gz
Pool4_Input_6_S25_L005_R1_001.fastq.gz
Pool4_Input_7_S26_L005_R1_001.fastq.gz

H3K4me3Q5ser_MACS2_0d_peaks.txt
H3K4me3Q5ser_MACS2_7d_peaks.txt
H3K4me3_MACS2_0d_peaks.txt
H3K4me3_MACS2_7d_peaks.txt
pre_differentiation_H3K27me3_peaks.txt
post_differentiation_H3K27me3_peaks.txt
annotated_TAF_0_peaks.txt
annotated_TAF_7_peaks.txt
annotated_Dual_9.5_peaks.txt
annotated_Dual_17.5_peaks.txt
annotated_H3K4me3_9.5_peaks.txt
annotated_H3K4me3_17.5_peaks.txt
annotated_IPSC_peaks.txt
annotated_Neuron_peaks.txt

All processed data also provided in Supplementary Data Tables 1-34 (Excel) as part of the revised submission.

Genome browser session
(e.g. UCSC)

N/A

Methodology

Replicates

For all ChIP-seq experiments (in hPSCs/5-HT neurons, RN46A-B14 cells and in brain), an n=3-4 biological replicates were used in each group.

Sequencing depth

Average sequencing depth for RN46A-B14 cell experiments: 45 million reads
Average sequencing depth for mouse brain experiments: 30 million reads
Average sequencing depth for human PSC/5-HT neuron experiments: 50 million reads

Antibodies

For RN46A-B14 cells: TAF3 (Millipore 07-1802; Lot#3008423), H3K4me3 (Abcam ab8580; Lot#GR273043-6), H3K4me3Q5ser (Millipore 4491a; Lot#RA1608107).
For mouse brain samples: H3K4me3Q5ser (Millipore 4491a; Lot#RA1608107), H3K4me3 (Abcam ab8580; Lot#GR273043-6).
For human PSCs/5-HT neuron samples: H3K4me3Q5ser (Millipore 4491a; Lot#RA1608107)

Peak calling parameters

Peak-calling was performed using MACSv2.1.1 with default settings

Data quality

Peaks were filtered for FDR < 0.05 and fold change > 1.2.

Software	MACSv2.1.1 diffReps