

SCIENTIFIC REPORTS



OPEN

Cell Type-Specific Survey of Epigenetic Modifications by Tandem Chromatin Immunoprecipitation Sequencing

Mari Mito^{1,2}, Mitsutaka Kadota³, Kaori Tanaka³, Yasuhide Furuta^{4,5}, Kuniya Abe⁶, Shintaro Iwasaki^{2,7} & Shinichi Nakagawa^{1,8}

The nervous system of higher eukaryotes is composed of numerous types of neurons and glia that together orchestrate complex neuronal responses. However, this complex pool of cells typically poses analytical challenges in investigating gene expression profiles and their epigenetic basis for specific cell types. Here, we developed a novel method that enables cell type-specific analyses of epigenetic modifications using tandem chromatin immunoprecipitation sequencing (tChIP-Seq). FLAG-tagged histone H2B, a constitutive chromatin component, was first expressed in *Camk2a*-positive pyramidal cortical neurons and used to purify chromatin in a cell type-specific manner. Subsequent chromatin immunoprecipitation using antibodies against H3K4me3—a chromatin modification mainly associated with active promoters—allowed us to survey the histone modifications in *Camk2a*-positive neurons. Indeed, tChIP-Seq identified hundreds of H3K4me3 modifications in promoter regions located upstream of genes associated with neuronal functions and genes with unknown functions in cortical neurons. tChIP-Seq provides a versatile approach to investigating the epigenetic modifications of particular cell types *in vivo*.

The tissues of living organisms consist of a wide variety of cell types that orchestrate coordinated tissue-specific biological processes. An extreme case of this complexity can be found in the nervous system of higher vertebrates, which possesses two major cell types: glial cells and neuronal cells. The latter can be further divided into numerous subtypes based on their morphologies, physiological properties, molecular marker expression, and their combinations^{1,2}. The neuronal lineage is generated through a series of differential gene expression patterns during the course of its development, and it is then diversified into each cell type with individual physiological properties^{2,3}. While neurons are a major cell type in the mammalian cortex, they only cover less than half of the total cells at most⁴. Therefore, the gene expression profiles of neurons are typically averaged and masked by this complex mixture of cells in whole-brain tissue samples. Moreover, since certain neuronal subtypes are present only in a small population in particular regions of the brain, the gene expression profiles of minor neurons are hardly accessed. Thus, technical advances that enable sensitive and specific gene expression detection in target cells have been warranted.

The most straightforward approach for cell type-specific gene expression is the isolation of a subpopulation of cells and subsequent RNA sequencing (RNA-Seq). Indeed, FACS, cell panning, or more recently, single-cell

¹RNA Biology Laboratory, RIKEN, 2-1 Hirosawa, Wako, 351-0198, Japan. ²RNA Systems Biochemistry Laboratory, RIKEN, 2-1 Hirosawa, Wako, 351-0198, Japan. ³Phyloinformatics Unit, RIKEN Center for Life Science Technologies, 2-2-3 Minatojima-Minamimachi, Chuo-ku, Kobe, 650-0047, Japan. ⁴Animal Resource Development Unit and RIKEN Center for Life Science Technologies, 2-2-3 Minatojima Minami-machi, Chuo-ku, Kobe, 650-0047, Japan. ⁵Genetic Engineering Team, RIKEN Center for Life Science Technologies, 2-2-3 Minatojima Minami-machi, Chuo-ku, Kobe, 650-0047, Japan. ⁶Technology and Development Team for Mammalian Genome Dynamics, RIKEN BioResource Center, 3-1-1 Koyadai, Tsukuba, Ibaraki, 305-0074, Japan. ⁷Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, 2-1 Hirosawa, Wako, 351-0198, Japan. ⁸RNA Biology Laboratory, Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo, 060-0812, Japan. Correspondence and requests for materials should be addressed to S.I. (email: shintaro.iwasaki@riken.jp) or S.N. (email: nakagawas@pharm.hokudai.ac.jp)

RNA-Seq analyses, have been applied in particular neuronal subtypes^{2,5}. However, during the preparation of singly isolated cells, neurons must be placed under non-physiological conditions for a relatively long time, such that the procedures potentially affect the transcriptome of the cells. In addition, most dendrites and axons, which contain localized RNAs that are often linked to long-term potentiation for learning and memory⁶, are supposedly lost during cell preparation, which probably leads to underrepresentation of these localizing RNAs. Translating ribosome affinity purification (TRAP)^{7,8} and Ribo-Tag approaches⁹ are unique technologies that can circumvent these issues. In these techniques, ribosome-associated mRNAs are recovered by epitope-tagged ribosomal proteins driven under a cell type-specific promoter and analyzed by subsequent RNA-Seq^{7–9}. Although these methods are generally amenable to a wide range of tissues and lineages in principle, they fail to isolate non-protein-coding RNAs (ncRNAs), which do not associate with ribosomes. Considering that ncRNAs occupy a significant portion of the transcriptional output from the genomes of higher eukaryotes and that they regulate a variety of physiological processes^{10–12}, another comprehensive survey, which covers both the coding and non-coding transcriptome, should be developed.

In eukaryotic cells, the epigenetic modifications of chromatin components play essential roles in the control of gene expression^{13,14}. The differential chemical modifications of histones and DNA that are induced by various external stimuli underlie the conditional and specific gene expression patterns in neurons and glial cells¹⁵. Recent studies have reported successful genome-wide analyses of DNA methylation using FACS-isolated neurons from post-mortem human brains^{16,17}, dopaminergic neurons from mouse brains¹⁸, or manually micro-dissected hippocampal granular neurons¹⁹. Although the application of similar approaches for histone modifications would theoretically provide more information on transcriptional states, their implementation must be carried out with the caveat of relatively dynamic histone modifications²⁰, which are potentially affected by cell sorting processes.

Here, we overcame these issues by developing a method termed tandem chromatin immunoprecipitation sequencing (tChIP-Seq) that enables cell type-specific analyses of epigenetic modifications in mice. The exogenous expression of constitutive chromatin component H2B with an epitope tag under the *Camk2a* promoter and subsequent H3K4me3 ChIP-Seq allowed us to address neuron-specific chromatin modification, which is inaccessible when using whole-brain samples. We not only rediscovered known H3K4me3 modifications in promoters associated with neuronal function but also identified the modifications in promoters of novel neuronal RNAs with unknown functions. Thus, tChIP-Seq showed potential for diverse applications in various tissues and cell types and in any type of chromatin modification, including DNA methylation.

Results

We examined the efficiency of chromatin purification from target cells in the pools of mixed cells by an epitope-tagged chromatin component by initially using the human embryonic kidney (HEK) 293 cell line, which stably expresses exogenous FLAG-tagged human histone H2BJ (H2B-FLAG hereafter). The cell lysate was mixed with mouse NIH3T3 cells at various ratios, and the chromatin fraction was subsequently isolated by chromatin immunoprecipitation (ChIP) using an anti-FLAG antibody. Purified human and mouse DNA were quantified by qPCR with primer sets corresponding to human and mouse *Gapdh/GAPDH* and *Cd19/CD19* loci. Compared to the contaminated fraction of mouse chromatin, human chromatin was dominant up to a 1:10 dilution, whereas the mouse and human fractions were comparable in more diluted conditions (Supplemental Figure S1). Therefore, it would be fair to assume that a target cell type representing more than 10% of the total cells can be properly applied in our strategy for investigating epigenetic modifications.

The aforementioned pilot experiments with cultured cells led us to investigate cell type-specific epigenetic modification in tissues from living organisms, especially neurons in the brain. For this, we created an animal model that can ectopically express H2B-FLAG (Fig. 1A). We initially created a knock-in mouse *Rosa26*^{CAG floxed-pA H2B-FLAG}; H2B-FLAG was conditionally expressed upon induction of Cre-mediated recombination from the *Rosa26* locus^{21,22} (Fig. 1B,C). The floxed mice were then crossed with *CamK2a*^{CreERT2}²³ to obtain a mouse line (hereafter referred to as *Camk2a*^{H2B-FLAG}) expressing H2B-FLAG in *Camk2a*-positive neurons upon injection of tamoxifen. As a control, we crossed *Rosa26*^{CAG floxed-pA H2B-FLAG} with *Vasa*^{Cre}²⁴ and established a mouse line (*Rosa*^{H2B-FLAG}) that ubiquitously and constitutively expresses H2B-FLAG throughout development and as an adult. As expected, pyramidal neurons were identified by large, round DAPI-stained nuclei that specifically expressed H2B-FLAG in the cerebral cortex of tamoxifen-injected *Camk2a*^{H2B-FLAG}, whereas almost all the DAPI-positive cells, including glial cells, that possess small nuclei expressed H2B-FLAG in *Rosa*^{H2B-FLAG} mice (Fig. 1D). We tested the ratio of H2B-FLAG and endogenous H2B by preparing nucleus lysate from the cortex of *Rosa*^{H2B-FLAG} mice and performing Western blot analyses. The tagged exogenous H2B was observed as dual bands, one of which is presumably translated from an in-frame leading start codon unexpectedly derived from the loxP sequences, resulting in the addition of 35 amino acid sequences (Fig. 1B,E). The ectopically expressed H2B FLAG consisted of approximately one-sixth of the total amount of H2B in *Rosa*^{H2B-FLAG}, resulting in a 20% increase in H2B compared to H2B in wild-type mice (Fig. 1E,F). Despite the slight increase in H2B protein, both *Camk2a*^{H2B-FLAG} and *Rosa*^{H2B-FLAG} mice displayed no observable abnormalities, suggesting that the expression of tagged H2B neither perturbs normal development nor interferes with physiological processes as previously reported for GFP-tagged H2BJ²⁵. The population of H2B-FLAG-labeled cells (~10%) in the brain passed our technical limitation (~10%, Supplemental Figure S1), which allowed further sequencing of the isolated DNA fractions.

We then asked if H2B-FLAG is evenly distributed throughout the genome by performing chromatin immunoprecipitation sequencing analyses (ChIP-Seq) using an anti-FLAG antibody from the brain tissues of *Rosa*^{H2B-FLAG} mice. Dissected brains were cut into small pieces, rapidly frozen in liquid nitrogen, and mechanically crushed into a fine powder in a pre-chilled tube to enable rapid and uniform fixation of tissue samples. After fixation with 1% paraformaldehyde, we immunoprecipitated chromatins using an anti-FLAG antibody and purified DNA for subsequent deep sequencing. We improved the efficiency of immunoprecipitation and reduced background contamination by using 1 × Denhardt's solution and Blocking Reagent from Roche in the immunoprecipitation

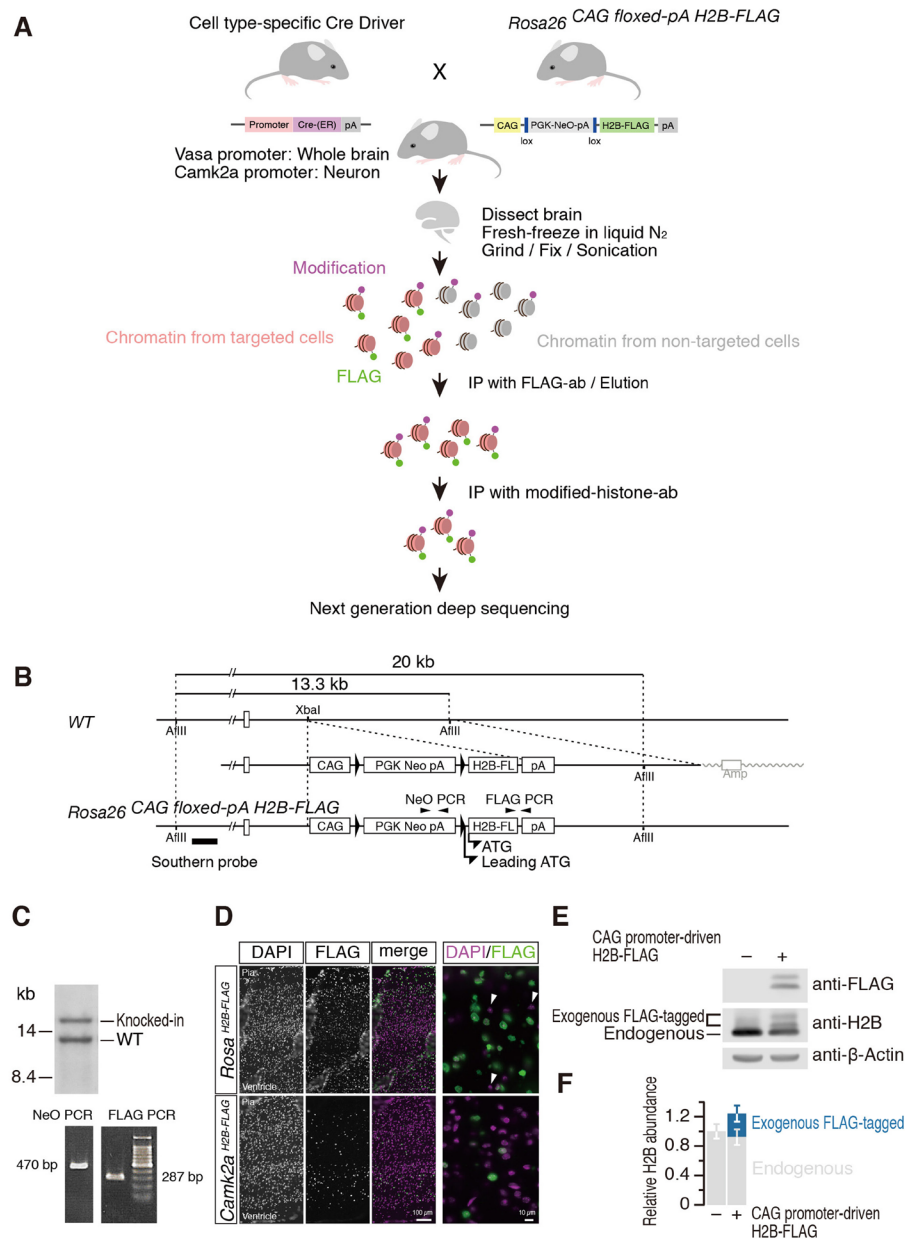


Figure 1. Genetic labeling of cell type-specific chromatin by H2B-FLAG. **(A)** Schematic drawings of the experimental design. Cell type-specific expression of H2B-FLAG was induced by crossing Cre-driver mice that express Cre recombinase in a particular cell type with floxed H2B-FLAG mice that possess an expression cassette of H2B-FLAG upon Cre-mediated recombination. Chromatins from cell types of interest were recovered from cellular lysate prepared from whole tissues by immunoprecipitation using an anti-FLAG antibody. Then, subsequent ChIP with antibodies against a specific chromatin modification, H3K4me3 in this study, was performed. **(B)** Schematic drawings of the targeting strategy to generate *Rosa26*^{stop-floxed CAG-H2B-FLAG} mice. CAG, CAG promoter; PGK, PGK1 promoter; Neo, neomycin resistant gene; H2B-FL, H2B-FLAG; pA, polyadenylation signal of bovine growth hormone. Designed and leading in-frame ATG codons for H2B-FLAG are depicted. Position of the probe used for Southern blot analyses is shown in a bold bar. Positions of primers used for genotyping are indicated by arrowheads. **(C)** Southern blot analyses of the targeted ES cells used for generation of the chimera mice, and the agarose gel electrophoresis images of genotyping PCR. **(D)** Expression of H2B-FLAG in the cortical plates of *Rosa*^{H2B-FLAG} and *Camk2a*^{H2B-FLAG} mice. H2B-FLAG was almost ubiquitously expressed in the cortical cells in *Rosa*^{H2B-FLAG} mice. Occasional and exceptional cells that did not express the tagged protein were indicated by arrow heads. In the cortical plate of *Camk2a*^{H2B-FLAG} mice, H2B-FLAG expression was observed only in large cortical neurons identifiable by large, round nuclei, as expected from a previous report on the Cre-driver mouse²³. **(E)** Expression levels of the exogenous H2B-FLAG in *Rosa*^{H2B-FLAG} mice were compared to the endogenous H2B expression level. Total cell lysates prepared from the cortical plate of *Rosa*^{H2B-FLAG} mice were blotted with anti-FLAG, anti-H2B, and anti- β -actin antibodies. **(F)** Quantitation of the results shown in (E). Data represent the mean and S.D. (n = 3). The exogenous H2B-FLAG consisted of ~16% of the total H2B protein.

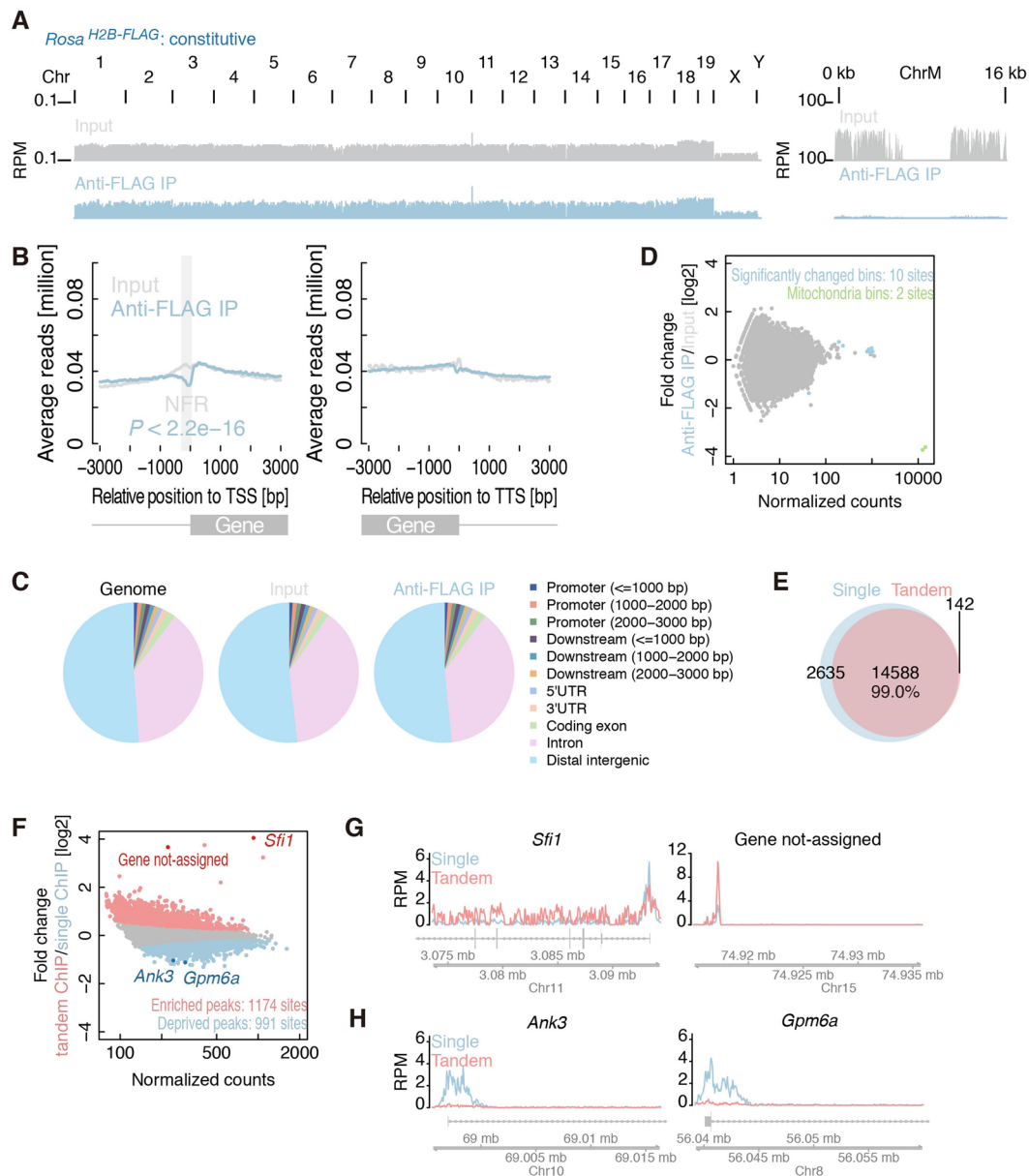


Figure 2. Purification of chromatin by H2B-FLAG. **(A)** Read distributions of input DNA and H2B-FLAG-bound DNA from *Rosa^{H2B-FLAG}* mice along each chromosome. ChrM represents mitochondria genome. **(B)** Enrichments of reads along the transcription start site (shown at 0 on the x-axis). Significance was calculated by the binomial test. TSS: transcription start site, TTS: transcription termination site. **(C)** Assignment of the origin of reads from input DNA and H2B-FLAG-bound DNA from *Rosa^{H2B-FLAG}* mice. **(D)** MA-plot of the mean reads in every 10-kbp bin along entire genomes versus their differential changes in anti-FLAG IP compared to input reads. Significantly changed bins (FDR < 0.01) are highlighted in light blue. Bins from mitochondrial genomes are depicted in light green. **(E)** Venn-diagram of peaks called by single H3K4me3 ChIP-Seq and tandem ChIP-Seq from *Rosa^{H2B-FLAG}* mice. **(F)** MA-plot of the mean read counts in called peaks along entire genomes versus their differential changes in single ChIP compared to tChIP. **(G and H)** Distribution of reads from *Rosa^{H2B-FLAG}* single H3K4me3 ChIP-Seq and tandem ChIP-Seq along the genomic loci, where biases were observed. RPM stands for reads per million.

buffer, which was essential for this study (M.M. and S.N., unpublished observation). The origin of reads from the input DNA and FLAG-purified samples was indistinguishable along the chromosomes, except for the mitochondrial genome, which was free of nucleosomes (Fig. 2A). We also observed a drop in mapped reads in the FLAG-purified samples ~300 bp upstream of the transcription start site (Fig. 2B), which has been reported to be a nucleosome-depleted region²⁶. Apart from these exceptions, the number of reads mapped to various genome coordinates exhibited similar trends (Fig. 2C). To further confirm the distribution of FLAG-purified chromatin more quantitatively, we compared the number of mapped reads in 10-kbp bins over the entire genome (Fig. 2D). Among the countable 247,297 bins, only 10 bins exhibited statistically significant (FDR < 0.01) changes between

the input and FLAG-purified chromatin, and 2 of them were from the mitochondrial genome. These observations suggested that the first immunoprecipitation using the FLAG antibody enabled relatively unbiased purification of chromatin from cells expressing H2B-FLAG. A similar observation has been previously reported for chromatin immunoprecipitation of another constitutive component, Histone H3²⁷.

The isolation of chromatin by H2B-FLAG allowed us to perform subsequent ChIP-Seq for modified histones (or tandem ChIP-Seq) in *Rosa*^{H2B-FLAG} mice. Here, we focused on tri-methylated lysine 4 of histone H3 (H3K4me3), which is a hallmark of active transcription and is enriched at the promoter regions^{28,29}. We performed three independent tChIP-Seqs and simple H3K4me3 ChIP-Seqs (or single ChIP-Seqs) and found that ~80% of the peaks were commonly called in all three experiments (Supplemental Figure S2A), indicating the high reproducibility of our tChIP-Seq approach. As expected, called peaks from both samples were enriched around transcription start sites (Supplemental Figure S2C). We observed extensive overlap of the peaks in single ChIP-Seq and tChIP-Seq; 99% of the peaks were commonly identified in single ChIP-Seq and tChIP-Seq (Fig. 2E), and a comparison of the number of reads mapped to identified peak regions exhibited a strong correlation ($r > 0.93$ for all possible combinations) between single ChIP-Seq and tChIP-Seq (Supplemental Figure S3). Despite this high correspondence, we observed differential enrichments in particular loci, and 1174 and 991 sites displayed statistically significant ($FDR < 0.01$) enrichment and deprivation in the tChIP-Seq samples, respectively, compared to the input single ChIP-Seq samples (Fig. 2F–H). This effect may be due to a certain bias towards longer DNA after H2B purification (Supplemental Figure S2E), which may reflect stable nucleosome selection during the first chromatin purification process. Alternatively, the efficiency of the second immunoprecipitation using anti-H3K4me3 antibodies may be affected by the presence of the FLAG epitope at particular loci, even though the first H2B-FLAG purification did not essentially alter the DNA population (Fig. 2A–D). It is also possible that the H2B variant used in this study, N-terminal extension by tagging, and/or slight increase in the total amount of H2B protein (Fig. 1B,E) may cause these biases as has been observed in core histone variants that display specific genomic distributions^{30,31}. It should also be noted that a significant portion of the FLAG-H2B molecules expressed in mice have extended N-terminal sequences, presumably derived from unexpected in-frame start codon in loxP sequences (Fig. 1E). These molecules may be preferentially deposited in particular genomic loci or affect histone modifications. Therefore, we offset all of these biases by mainly performing subsequent differential and quantitative analyses in a comparison of the tChIP-Seq datasets obtained from *Rosa*^{H2B-FLAG} and *Camk2a*^{H2B-FLAG}; we did not compare the results obtained with single ChIP-Seq and tChIP-Seq. We assumed that this would be the best practice for downstream analyses.

Then, we prepared single and tandem ChIP-Seq libraries from *Camk2a*^{H2B-FLAG} mice to investigate the neuron-specific epigenome. Primary chromatin isolation from *Camk2a*-positive neurons did not show a strong bias along the entire genome (Supplemental Figure S4), suggesting that H2B-FLAG is evenly distributed throughout the genome in *Camk2a*-positive neurons, as is the case for the whole brain samples (Fig. 2D). We again found that their called peaks were highly reproducible (Supplemental Figure S2B) and that they were enriched around transcription start sites/promoter regions (Supplemental Figure S2D). We subsequently compared the number of reads mapped to the peak regions called in both *Camk2a*^{H2B-FLAG} and *Rosa*^{H2B-FLAG} tChIP-Seq and examined the fold change distributions along the called peaks (pink color in Fig. 3A). In the case of single ChIP-Seq of *Camk2a*^{H2B-FLAG} and *Rosa*^{H2B-FLAG}, the data were obtained from two essentially identical brain samples, so the difference between them should not be huge. As expected, most of the peaks from single ChIP-Seqs exhibited neither enrichment nor a decrease, and they accumulated around the $\log_2 = \text{zero}$ region (gray color in Fig. 3A). On the other hand, the fold change exhibited a much wider distribution when comparing the peaks obtained from *Camk2a*^{H2B-FLAG} and *Rosa*^{H2B-FLAG} tChIP-Seq (pink color in Fig. 3A). These clear differences were consistent with the idea that tChIP-Seq can correctly discriminate the origin of cells, which cannot be done with single ChIP-Seq. Strikingly, we observed a strong accumulation of promoter H3K4me3 marks in known neuronal genes, including *Camk2a*, whose promoter defined the target cells in this experiment, and vesicular glutamate transporter 1 (*Vglut1/Slc17a7*), a well-known glutaminergic neuron marker gene³² (Fig. 3B). Concomitant to such an accumulation in representative glutaminergic neuronal genes, non-neuronal marker genes were deprived in the *Camk2a*^{H2B-FLAG} tChIP-Seq data. For example, H3K4me3 peaks were detected in the promoter regions of astrocyte marker *Aldh1l1*³³ and microglia marker *Cd40* in *Rosa*^{H2B-FLAG} tChIP-Seq, but they were depleted in *Camk2a*^{H2B-FLAG} tChIP-Seq (Fig. 3C). We also monitored the reduction of peaks in blood vessel genes, including *Robo4* and *Cdh5*^{34,35} (Fig. 3D). These observations supported that we could successfully investigate chromatin modification in specific cell types.

We further quantitatively investigated the distribution of the peaks enriched in *Camk2a*-positive cells without bias by counting the number of reads mapped to called peaks in *Camk2a*^{H2B-FLAG} tChIP and *Rosa*^{H2B-FLAG} tChIP and performed differential enrichment analysis. Again, the read counting was highly reproducible among replicates (Supplemental Figure S5). We observed significant enrichment of reads in the peaks assigned to neuronal genes (identified GO term “neuron development”) and depletion of those assigned to blood vessel genes (GO term “blood vessel development”) in the tChIP data (Fig. 3E), whereas single ChIP did not exhibit such a difference (Supplemental Figure S6). Moreover, unbiased GO-term analysis resulted in gene accumulations related to “synapse” and “cation channel activity”, which are a key component and function of neurons, respectively (Fig. 3F). Taken together, we concluded that quantitative analysis of *Camk2a*^{H2B-FLAG} tChIP-Seq data successfully captures the genomic positions of chromatin modifications associated with active transcription in neuronal cells.

Other relevant cell type-specific techniques based on RNA-Seq, FACS-sorted RNA-Seq³⁶ and TRAP-Seq^{7,8} have previously been used to explore the transcriptional profiles of particular neurons. Because the H3K4me3 modification is preferentially^{28,29}, if not exclusively³⁷, associated with active promoters, tChIP-Seq enables the prediction of candidates of genes that are actively transcribed in the cells. We thus evaluated the power of tChIP-Seq relative to the techniques that directly profile the transcripts by reanalyzing the transcriptome data obtained in similar conditions^{8,36}. For the consistency of the data analysis, we calculated the neuron enrichments (RNA-Seq

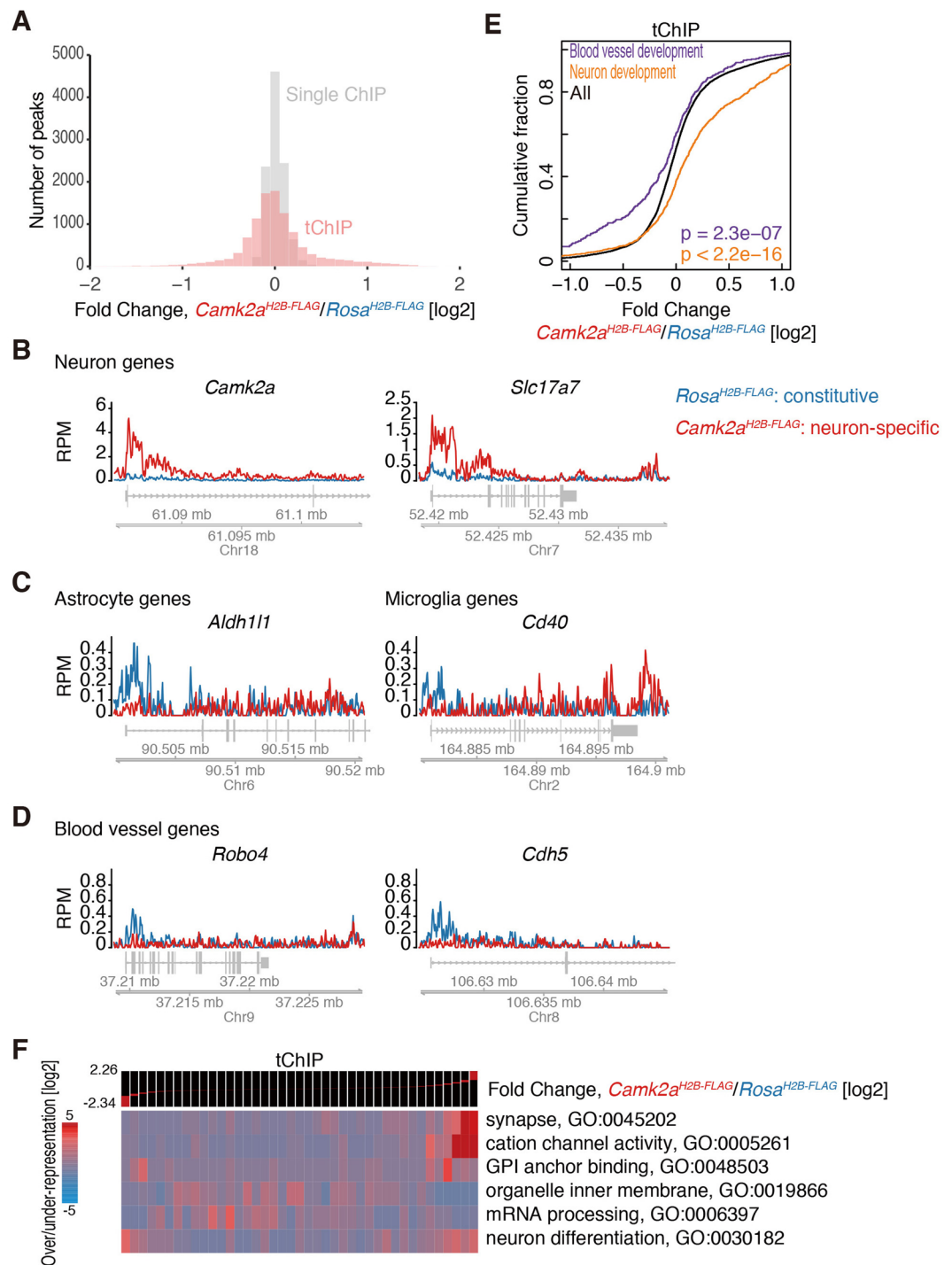


Figure 3. tChIP-Seq from *Camk2a*^{H2B-FLAG} showed accumulation and depletion of reads along known neuronal genes and non-neuronal genes, respectively. (A) Histogram of number of peaks along differential changes of single H3K4me3 ChIP-Seq and tChIP-Seq in *Camk2a*^{H2B-FLAG} compared to those from *Rosa*^{H2B-FLAG}. Bin width is 0.1. (B–D) Distribution of reads from tChIP-Seq of *Rosa*^{H2B-FLAG} and *Camk2a*^{H2B-FLAG} along each gene locus. RPM: reads per million. (E) Cumulative distribution of reads in peaks corresponding to neuron development and blood vessel development genes along differential changes between *Camk2a*^{H2B-FLAG} and *Rosa*^{H2B-FLAG} tChIP-Seq. Significance is calculated by Mann-Whitney U test. (F) GO analysis along differential changes between *Camk2a*^{H2B-FLAG} and *Rosa*^{H2B-FLAG} tChIP-Seq with iPAGE.

from FACS-sorted neurons or from mRNA associated with tagged ribosomes expressed under the *Camk2a* promoter) compared to input RNA (brain RNAs without selection) as we did in tChIP-Seq. Even after considering the differences in the applied approaches, mouse background, and laboratories where the data were collected, we

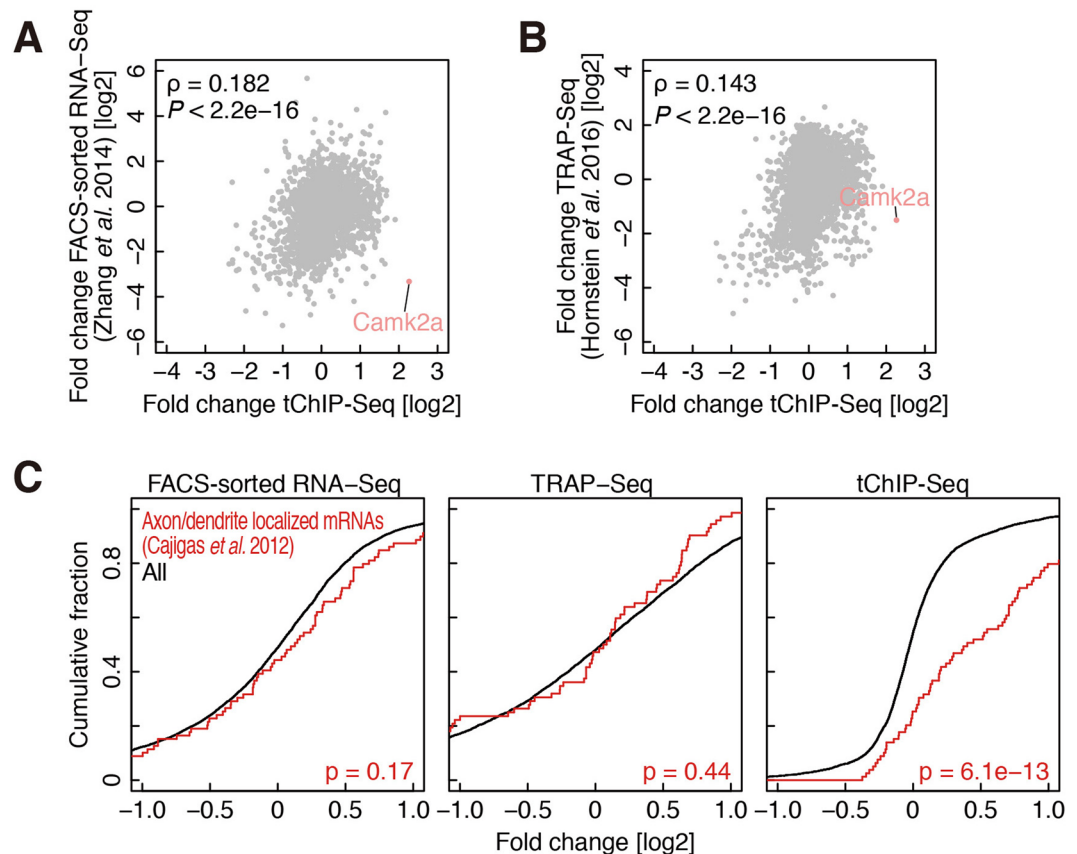


Figure 4. Comparison of tChIP-Seq to previously reported methods for neuron-specific gene expression. (**A** and **B**) Correlation of tChIP-Seq enrichment and FACS-sorted RNA-Seq (**A**) and TRAP-Seq (**B**). ρ : Spearman's rank correlation. P value is calculated by Student's t test. (**C**) Cumulative distribution of axon/dendrite-localized transcripts along FACS-sorted RNA-Seq, TRAP-Seq, and tChIP-Seq. Significance is calculated by Mann-Whitney U test.

observed only a weak correlation of mRNA enrichment between published data sets and tChIP-Seq (Fig. 4A,B, and Supplemental Figure S7). Notably, we did not detect any enrichment of the well-established neural marker *Camk2a* in FACS-sorted RNA-Seq and TRAP-Seq, unlike tChIP-Seq. This finding was even more contradictory in TRAP-Seq, in which the target cells were constrained by the *Camk2a* promoter⁸, which is the same promoter we used for tChIP to enrich cell type-specific chromatin. Given that *Camk2a* mRNA is localized and translated at dendrites reviewed in^{6,37}, we speculate that dendritic and axonal mRNAs were lost in FACS-sorted RNA-Seq or TRAP-Seq for unknown reasons. We also investigated the enrichment of mRNAs previously assigned as dendritic or axonal transcripts by micro-dissection of synaptic neuropils. Based on RNA-Seq³⁸ in FACS-sorted RNA-Seq, TRAP-Seq, and tChIP-Seq data, we found that the enrichment was clearly observed only in tChIP-Seq (Fig. 4C). Although the causes of this difference among techniques remains unclear (see more details in the Discussion), tChIP-Seq uncovered a wide range of H3K4me3 modifications in promoter regions located upstream of the genes reported to be active in neurons.

Our exploration of H3K4me3 modifications by tChIP-Seq was not restricted to protein-coding genes; it was also expanded into ncRNAs. The power of our tChIP-Seq approach was exemplified by clear identification of the H3K4me3 peaks in the promoter regions of micro-RNA precursors of miR137, miR129-2, and miR124a-3, all of which regulate the development of the nervous system and are closely related to neuronal disorders³⁹⁻⁴¹ (Fig. 5A). We also detected an enrichment of the peaks in the promoter region of *Miat/Gomafu*, which is specifically expressed in neuronal cells and controls animal behavior^{42,43} (Fig. 5B). In contrast, the peak was decreased in *Neat1*, which is expressed in glial cells, but not in neurons, under normal conditions^{44,45} (Fig. 5C).

We finally aimed to identify novel H3K4me3 modifications in promoters of neuronal genes using *Camk2a*-tChIP-Seq. Statistical analysis (see Methods and Materials) identified 1584 peaks as neuron-enriched sites (and also 974 sites as deprived) (Fig. 6A, Supplemental Table 1). Among them, we newly identified H3K4me3 modifications in promoters of genes including *Golga7b*, *Fam43b*, and *Fam135b*, which have not previously been reported for neural function (Fig. 6B). As expected, *in situ* hybridization of these transcripts revealed that their transcripts were observed, if not exclusively, in the cortical neuron with large, round nuclei (AP in Fig. 6C), some of which expressed H2B-FLAG in *Camk2a*^{H2B-FLAG} (FISH in Fig. 6C). Although the transcripts of *Fam43b* and *Fam135b* were predominantly found in the nucleus, the cytoplasmic localization of their substantial fractions was evident by active translation, which was observed by ribosome profiling in the mouse brain⁸ (Fig. 6D). Thus, the

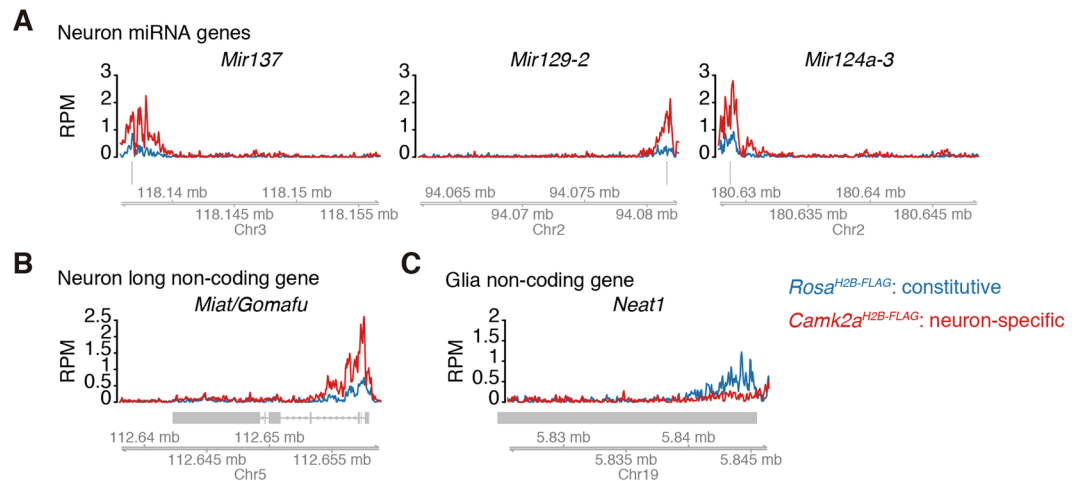


Figure 5. Neuron-specific ncRNAs verified by tChIP-Seq. (A–C) Distribution of reads from tChIP-Seq of *Rosa*^{H2B-FLAG} and *Camk2a*^{H2B-FLAG} mice along neuron-specific miRNA loci (A), long non-coding RNA loci (B), and gliia-specific long non-coding RNA loci (C). RPM: reads per million.

tChIP approach allowed us to discover novel promoters that preferentially possess particular chromatin modifications in neurons.

Discussion

We have thus demonstrated that tChIP-Seq enables cell type-specific analyses of epigenetic modifications using a mouse model. As far as we know, this is the first study to investigate chromatin modifications of particular cell types in tissues of living animals without isolating the cells, which is a process that potentially alters chromatin modification at a higher turnover rate during the preparation of single cells. This method simply relies on the purification of chromatin using FLAG-tagged H2B specifically expressed upon Cre-mediated recombination, and it should be easy to implement for any type of histone modification in any cell type where Cre-driver mice are available. Our approach is also theoretically compatible with other ChIP-Seq-based analyses for various transcription regulators, bisulfite sequencing analyses of DNA methylation, and DNA footprinting strategies in DNase-Seq or ATAC (assay for transposase-accessible chromatin)-Seq for nucleosome survey⁴⁶.

Many previous studies have reported successful profiling of gene expression in particular neurons of interest^{2,5}. These studies, as well as analyses of H3K4me3 modifications in promoter regions by tChIP-Seq, have identified a similar group of genes that regulates neuronal functions; however, substantial differences can be found depending on the method and technical principles that they rely on. tChIP-Seq efficiently detected H3K4me3 modifications in the promoters of genes whose transcripts are known to be exported from the cell body to dendritic synapses or axons reviewed in^{6,38}, whereas FACS-sorted RNA-Seq and TRAP-Seq in neurons underrepresented the transcripts of these genes^{8,36}. We were currently unaware of the molecular basis for this discrepancy, but it might be caused by the different technologies that are used in each study. For example, the loss of dendrites/axons during sample preparation (especially FACS sorting) may lead to a biased isolation of transcripts near the cell body. Alternatively, poly(A)-selection during library preparation, which was conducted in the two RNA-Seq methods, may fail to capture transcripts with shorter poly(A)-tails, which is typical characteristic of dendrite/axon-localized mRNAs⁴⁷. In addition, the differences might reflect the biological outcome of post-transcriptional regulation: a subset of mRNAs may be actively transcribed but rapidly degraded and are not represented in the mRNA profiles. It should also be noted that certain synaptic mRNAs are packaged into ribosome-inaccessible compartments called neuronal granules during transportation⁴⁸ and are therefore underrepresented in ribosome-based analyses. Lastly, it should be stressed that the H3K4me3 modification is not a definitive marker for active transcription, and bivalent modifications of H3K4me3 and H3K27me3 have been proposed to coordinate the expression of developmentally regulated genes to enable tight gene expression regulation in ES cells⁴⁹. Therefore, it is possible that certain fractions of tChIP-identified promoters with H3K4me3 modifications are not active and downstream genes are not transcribed into RNAs. Whatever the case, further analyses are warranted to explore these possibilities. In addition to coding mRNAs, ncRNAs can be surveyed by our tChIP-Seq approach. This feature is highly advantageous over TRAP-Seq^{7,8}, which is unable to capture miRNA precursors and long non-coding RNA in principle, such as *Miat/Gomafu*, which are predominantly localized to the nucleus and do not associate with ribosomes.

While we observed a relatively unbiased H2B-FLAG distribution over the entire genome, ChIP-Seq analyses of FLAG-purified chromatin revealed that H2B-FLAG is specifically enriched or depleted from particular genomic regions. As previously reported for H3, H2B-FLAG is not observed in the mitochondrial genome, and it is less enriched at the upstream of transcription start sites, both of which are regions that are known to be nucleosome-free²⁷. Consequently, tChIP-Seq analyses are not applicable to analyses of these specific regions, which may be one of the major technical drawbacks of this approach. Recent proteomic analyses revealed that different cell types express distinct sets of histone H2B variants, which are deposited on specific genomic regions,

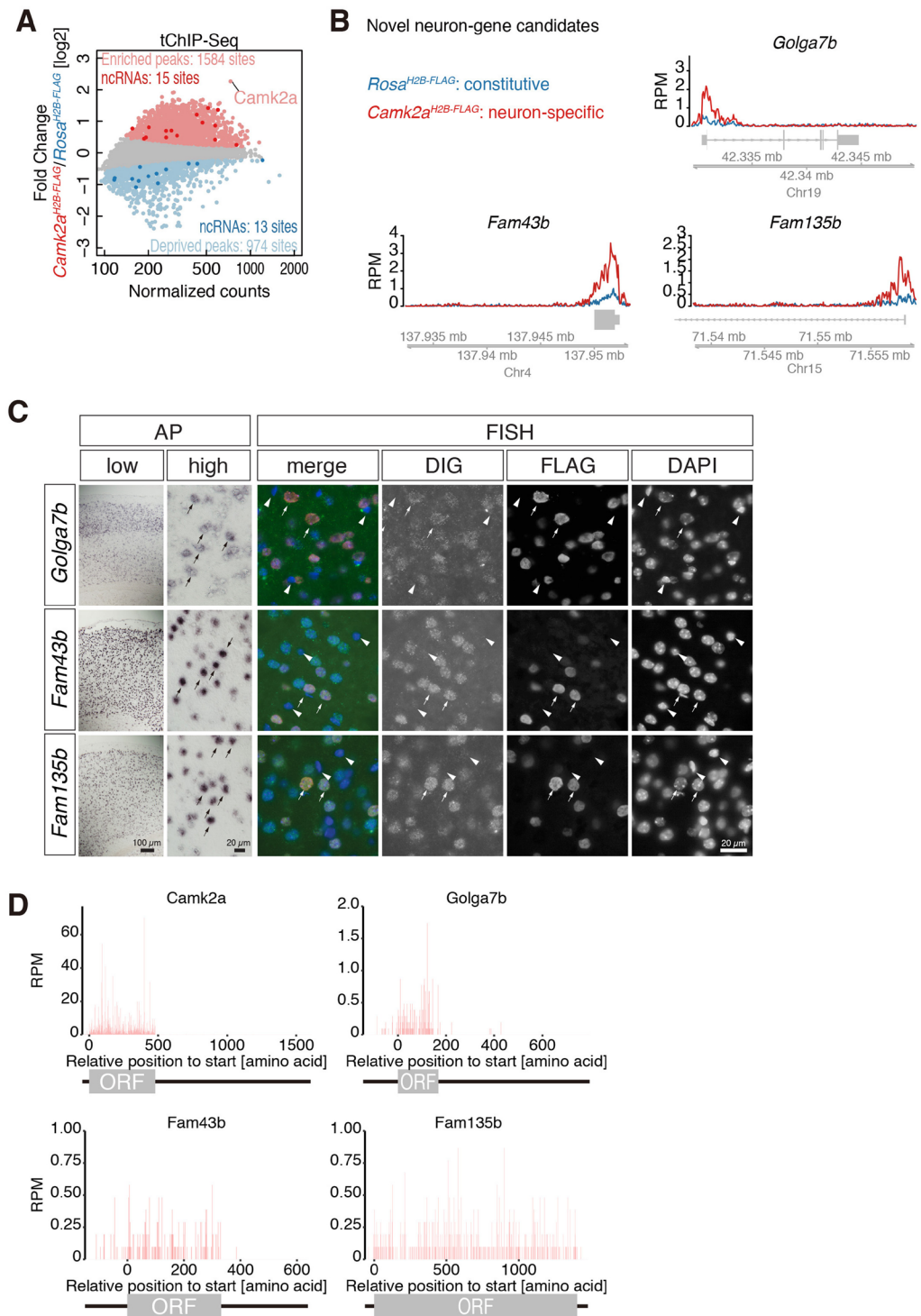


Figure 6. Novel neuron-genes identified by tChIP-Seq. **(A)** MA-plot of mean reads in H3K4me3 peaks and their differential changes in *Camk2a^{H2B-FLAG}* tChIP-Seq, highlighting defined neuron-enriched and neuron-depleted loci. Significantly changed peaks (FDR < 0.01) were highlighted in light pink and light blue. Peaks assigned to ncRNAs were depicted in darker pink and blue. **(B)** Distribution of reads from tChIP-Seq of *Rosa^{H2B-FLAG}* and *Camk2a^{H2B-FLAG}* along novel neuron gene candidates *Golga7b*, *Fam43b*, and *Fam135b*. RPM: reads per million. **(C)** Expression pattern of *Golga7b*, *Fam43b*, and *Fam135b* in the cortex of *Camk2a^{H2B-FLAG}* revealed by *in situ* hybridization. The transcripts were detected using chromogenic reaction of alkaline phosphatase-conjugated secondary antibodies (AP). For simultaneous detection with H2B-FLAG, the transcripts were detected with fluorescent-conjugated secondary antibodies (FISH). Distinct signals were observed in the cortical neurons identified by large, round nuclei (arrows), whereas little or no signals were observed in glial cells with small nuclei (arrowheads). Note that all the H2B-FLAG-positive cells expressed the candidate transcripts. **(D)** Ribosome footprint distribution along the novel neuronal genes. Reads assigned on each codon were depicted.

under specific conditions^{50,51}. The use of FLAG-tagged histone H4, which consists of only one type of molecule, would be a good alternative for the preparation of the input samples for tChIP to avoid these biases.

Another technical limitation of tChIP-Seq in further applications is its sensitivity. Unlike mRNAs, which are usually present in multiple copies, only 2 copies of genomic DNA are present in cells, making it challenging to perform ChIP experiments at a small scale. Our model experiments using cultured cells estimated that at least 10% of the entire cell population should consist of the target cell type to obtain enough signal against background noise from non-target cell types (Supplemental Figure S1). Thus, tChIP-Seq is currently limited to the analyses of non-rare cell types in tissues that are available in relatively large amounts. More detailed optimization of the purification of cell type-specific chromatin is warranted for the analyses of smaller populations of target cells in the future. Further improvement of the signal-to-noise ratio and efforts to use a small amount of initial materials (as in earlier studies^{52–54}) will be demanding tasks to extend our tChIP-Seq approach to a wider variety of cell types.

Taken together, the dataset obtained with our tChIP-Seq approach will be a versatile resource for lists of neuron-specific promoters. The profiling of multiple modalities of epigenetic modifications in particular cell types should further expand our understanding of the molecular basis that produces the functional diversity of cells in the tissues of living animals.

Methods

Generation of Rosa26 (FLAG-H2B) mice. An expression cassette that drives FLAG-tagged H2B under the control of the CAG promoter was knocked-in into the Rosa26 locus as previously described^{22,25}. Briefly, H2B-FLAG was cloned into pENTR2B (Invitrogen) and subsequently cloned into pROSA26-STOP-DEST through the Gateway system to generate the targeting vector. TT2 ES cells were transfected with a targeting vector linearized by Sall digestion and checked for homologous recombination using PCR and Southern blot. Chimera mice were generated using the ES cells to obtain $+/\text{Rosa26}^{\text{floxCAG-H2B-FLAG}}$, which were used for further studies. To induce the expression of H2B-FLAG expression in cortical neurons, $+/\text{Rosa26}^{\text{floxCAG-H2B-FLAG}}$ were crossed with $\text{CaMK}^{\text{CreERT2}}/+$ to obtain double-heterozygotes that carry both of the transgenes. Tamoxifen (#T5648, Sigma-Aldrich) was diluted in sunflower oil at a concentration of 10 mg/ml, and double-heterozygous male mice at the age of 8 weeks were treated with 5 contiguous intraperitoneal injections of tamoxifen (1 mg/day) for 5 days. A month later, the animals were sacrificed for sample preparation. Mice that ubiquitously express H2B-FLAG were prepared by crossing $+/\text{Rosa26}^{\text{floxCAG-H2B-FLAG}}$ (R26R (H2B-FLAG, Accession No. CDB1193K)) with $\text{Vasa-Cre}^{\text{24}}$, and siblings that underwent Cre-mediated recombination ($\text{Rosa26}^{\text{CAG-H2B-FLAG}}$) were maintained as heterozygotes. Genotyping PCR analyses were performed using Quick Taq HS DyeMix (#DTM-101, TOYOBO, Japan) and the following PCR conditions: pre-denaturation at 94 °C for 2 minutes and 30 cycles of denaturation at 94 °C for 30 seconds, annealing at 57 °C for 30 seconds, and extension at 68 °C for 30 seconds.

NeO_FW: AACCCAGATGACTACCTATCCTCC
 NeO_RV: GCGTGCTGAGCCAGACCTCCATCG
 FLAG_FW: CATTACAACAAGCGCTCGAC
 FLAG_RV: TAGAAGGCACAGTCGAGG

All methods were carried out in accordance with relevant guidelines and regulations. All experimental protocols were approved by the safety division of RIKEN (H27-EP071).

Immunohistochemistry. Dissected brains were embedded in Tissue-Tek (Sakura, Japan) and freshly frozen in dry-ice ethanol. Cryo-sections at a thickness of 10 μm were collected on PLL-coated slide glasses, fixed in 4% PFA for 10 minutes at a room temperature, and permeabilized in 100% MetOH at –20 °C for 5 minutes. After rehydration in PBS, the sections were incubated with anti-FLAG primary antibody (#F1804, Sigma-Aldrich) and were subsequently incubated with anti-mouse Cy3-conjugated secondary antibodies (Millipore, AP124C). Fluorescent images were obtained using an epifluorescence microscope (BX51, Olympus) equipped with a CCD camera (DP70).

In situ hybridization. Detailed protocols for the preparation of DIG-labeled RNA probes and *in situ* hybridization are described elsewhere⁵⁵. The following primers were used for preparation of the DNA templates for RNA probes. The reverse primers contain T7 promoter sequences for *in vitro* transcription to prepare DIG-labeled probes.

Golga7b_FW: CTCCCATGAGGCTGATCTGAGTC
 Golga7b_RV: tgacTAATACGACTCACTATAGGGGTATACTCACGCAATACACTTGTG
 Fam43b_FW: GGCCTTTCCAGGACCCAAGTGAGC
 Fam43b_RV: tgacTAATACGACTCACTATAGGGGGATGCTGCTGAGCCGCGGAGCTC
 Fam135b_FW: GCCTCCATAATATCTGATTCAGGC
 Fam135b_RV: tgacTAATACGACTCACTATAGGGTTCCACTGCCTTCAGGCTATCCAG

Tandem chromatin immunoprecipitation. The cerebral cortexes of adult brains were manually cut into small pieces (maximum size was a 5 mm³ cube) and quickly frozen in liquid nitrogen. Frozen tissue pieces from 2 animals were placed in eight 2-ml Eppendorf tubes, with a metal bullet (#SK-100-D5, TOKKEN, Japan) in each tube; all of these components were pre-chilled at –80 °C. The tubes containing the bullet were placed in a holder (#SK-100, TOKKEN) that was placed in liquid nitrogen for 30 seconds just before use, and the tissues were powdered with vigorous shaking for 30 seconds with 60 strokes. After placement in a –20 °C freezer for 15 minutes, 1.8 ml of 1% formaldehyde (28906, Pierce 16% formaldehyde (w/v), methanol-free) in PBS was added to each of

the eight tubes, and the cells were fixed for 10 minutes at 23 °C with gentle rotation. One hundred μ l of 2.5 M glycine was added to each tube to stop the fixative reaction. After centrifugation at 3,000 g for 5 minutes at 4 °C, the pellets were washed twice with PBS, combined into 2 tubes, and suspended in 1.3 ml/tube of lysis buffer 1 [50 mM Hepes-KOH (pH = 7.5), 140 mM NaCl, 1 mM EDTA (pH = 8.0), 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1 \times protease inhibitor cocktail (PI: #25955-24, Nacalai, Japan)]. After incubation for 10 minutes at 4 °C with gentle rotation, the cells were pelleted with centrifugation at 3,000 g for 5 minutes at 4 °C and re-suspended in 1 ml/tube of lysis buffer 2 [10 mM Tris-Cl (pH = 8.0), 200 mM NaCl, 1 mM EDTA (pH = 8.0), 0.5 mM EGTA (pH = 8.0), and 1 \times PI] by vortexing. The cells were further incubated for 10 minutes at 4 °C with gentle rotation, pelleted with centrifugation at 3,000 g for 5 minutes at 4 °C, and re-suspended in 800 μ l/tube of RIPA buffer (#89900, Thermo Fisher Scientific). After centrifugation at 3,000 g for 5 minutes at 4 °C, the cells were washed again with 500 μ l/tube of RIPA buffer, re-suspended in 1 ml/tube of RIPA containing 10 \times PI, and transferred into two 1-ml tubes (#520130, Covaris). Chromatin was fragmented using a focused-ultrasonicator (S220, Covaris) with the following conditions: temperature 4 °C; peak incident power = 175 Watts; duty cycle = 10%; cycle/burst = 200; time = 2,400 seconds). The chromatin lysates from the 2 tubes were mixed and stored at -80 °C for subsequent immunoprecipitation analyses. The concentration of DNA in the prepared chromatin lysate was approximately 50 ng/ μ l, and lysates containing 100 μ g of DNA were used for subsequent immunoprecipitation.

For M2-conjugated bead preparation, 200 μ l of anti-mouse IgG magnetic beads (#11201D, Invitrogen) were washed with PBS and suspended in 1 ml PBS and incubated with 20 μ g monoclonal anti-FLAG antibody M2 (#F1804, Sigma-Aldrich) for 6 hours at 4 °C with gentle rotation. For anti-H3K4me3 monoclonal antibody (#301-34811, Wako, Japan)-conjugated bead preparation, 40 μ l of beads were used to conjugate 4 μ g of the antibody. The conjugates were washed twice with PBS and suspended in 200 μ l of RIPA.

For purification of the chromatin containing FLAG-tagged H2B, 1 ml of chromatin lysates was first incubated with 20 μ l of the M2-conjugated beads for 6 hours at 4 °C in the presence of 1 \times Denhardt's solution and 1 \times blocking reagents (#11096176001, Sigma-Aldrich). The addition of the two blocking substances was essential to reduce background noise and increase the efficiency of immunoprecipitation. The beads were washed 6 times with RIPA buffer and incubated with 200 μ l of ChIP elution buffer [10 mM Tris (pH = 8.0), 300 mM NaCl, 5 mM EDTA (pH = 8.0), 1% SDS] for 15 minutes at room temperature. The eluted chromatin fractions were diluted with 2 ml of RIPA buffer and further incubated with anti-H3K4me3-conjugated beads overnight at 4 °C in the presence of Denhardt's solution and blocking reagents. The beads were washed 6 times with RIPA buffer and eluted with 200 μ l of ChIP elution buffer for 15 minutes at room temperature.

For the isolation of DNA from the chromatin eluent, the solutions were incubated overnight at 65 °C and sequentially treated with 50 μ g/ml RNaseA (#30141-14, Nacalai, Japan) and 0.5 mg/ml Proteinase K (#3115887001, Sigma-Aldrich) at 37 °C for 30 minutes and 60 minutes, respectively. After the addition of 20 μ g of glycogen (#AM9510, Thermo Fisher Scientific), DNA was purified through phenol-chloroform extraction followed by ethanol precipitation. For the isolation of DNA from the first input solution, 20 μ l of the cell lysate was diluted with 180 μ l of ChIP elution buffer, de-cross-linked at 65 °C, and extracted through phenol-chloroform treatment followed by ethanol precipitation. The concentration of the recovered DNA was measured using Qubit (Thermo Fisher Scientific) following the manufacturer's instructions.

Western blot. The cerebral cortex lysate was prepared by the sonication within Laemmli sample buffer. Anti-FLAG (#F1804, Sigma-Aldrich) (1:1000), anti- β -actin (926-42210, LI-COR) (1:1000) and anti-H2B (#ab1790, Abcam) (1:1000) were used as primary antibodies, and IRDye 800CW anti-mouse IgG (926-32210, LI-COR) and IRDye 800CW anti-rabbit IgG (926-32211, LI-COR) were used as secondary antibodies (1:10000). Fluorescent images were captured by Odyssey CLx (LI-COR).

qPCR analyses. We constructed pCDNA5/FRT/TO-H2B-FLAG by inserting a DNA fragment containing H2B-FLAG isolated by BamHI/SalI digestion from pENTR2B-H2B-FLAG (described above) into pCDNA5/FRT/TO (Invitrogen) via BamHI/XhoI sites. The integration of H2B-FLAG into T-Rex-293 (Thermo Fisher Scientific) was performed by the co-transfection of pCDNA5/FRT/TO-H2B-FLAG and pOG44 (Invitrogen) via Hygromycin B-selection according to manufacturer's instructions. The clonal cell line was isolated by limiting dilution, and it was grown with tetracycline to induce H2B-FLAG. Cell lysate from the integrated cells and mouse NIH3T3 cells were prepared as described above and then mixed at the ratio shown in the figure legend (Supplemental Fig. 1). DNA was purified with immunoprecipitation by anti-FLAG antibody as described above. qPCR was performed using SYBR Premix Ex Taq (TAKARA, Japan) and Thermal Cycler Dice TP870 (TAKARA) according to the manufacturer's protocol. The qPCR calculation for each target loci was performed by the absolute method, referring to a standard curve generated using one of the input DNA samples. The qPCR value was normalized to the value corresponding to the input DNA. The following primer sets were used.

Human *GAPDH*:

5'-CCACATCGCTCAGACACCAT-3' and 5'-AGCCACCCGCGAACTCA-3'

Mouse *Gapdh*:

5'-GCCTACGCAGGTCTTGCTGAC-3' and 5'-CGAGCGCTGACCTTGAGGTC-3'

Human *CD19*:

5'-GGCCTGGGAATCCACATG-3' and 5'-CGGCTGGCACAGGTAGAAG-3'

Mouse *Cd19*:

5'-CACGTTCACTGTCCAGGCAG-3' and 5'-AGTGATTGTCAATGTCTCAG-3'

Library construction and deep sequencing. Library preparation was performed as previously described⁵⁶ with modifications. Input DNA and ChIP DNA were selected for the size range of 100-500 bp using

AMPure XP beads (Beckman Coulter, A63881). Libraries were prepared from 20 ng of input DNA, 2 ng of H3K4me3 ChIP DNA, 20 ng of M2 ChIP DNA, and 2 ng of M2/H3K4ME3 ChIP DNA using a KAPA LTP Library Preparation Kit (KAPA Biosystems) with TruSeq index adaptors. The amount of PEG/NaCl solution used for reaction clean-ups was $1.8 \times$ up to the adaptor ligation step and $2 \times$ after the adaptor ligation step. The size selection of DNA was carried out after the end repair reaction by AMPure XP beads. To confirm the quality and quantity of the prepared libraries, their concentration and size distribution were analyzed with a Qubit dsDNA HS Assay kit and an Agilent High Sensitivity DNA kit. Library qPCR was also carried out to calculate the enrichment folds at positive and negative control regions using equal amounts of input and ChIP library DNA as templates in the reaction. EV were calculated as the ratio of quantified concentrations of “ChIP libraries” and “input libraries”. ChIP-Seq libraries were subjected to on-board cluster generation using HiSeq SR Rapid Cluster Kit v2 (Illumina) and sequenced on the Rapid Run Mode of an Illumina HiSeq1500 (Illumina) to obtain single-end 80 nt reads. Base calling was performed with RTA ver1.18.64, and the fastq files were generated with bcl2fastq ver1.8.4 (Illumina).

Bioinformatic analysis. *Initial read processing.* Reads were filtered by their quality and trimmed when an adapter sequence occurred (FASTX Toolkit, version 0.0.14, Hannon lab, http://hannonlab.cshl.edu/fastx_toolkit/commandline.html).

ChIP-Seq. After depletion of the reads mapped to Repeat Masker sequences from the UCSC genome browser (Bowtie2, version 2.1.0), the remaining reads were mapped to the mouse genome (mm9) (Bowtie, version 1.0.0)^{57,58}, and then the duplicated reads were suppressed by samtools v1.2 rmdup⁵⁹. Read annotation (Fig. 2B) was performed by CEAS (version 1.0.2, Liu Lab, <http://liulab.dfci.harvard.edu/CEAS/index.html>). Reproducible peaks called by MACS (version 1.4.2)⁶⁰ in triplicate experiments were used for downstream analysis. Reads mapped to the peaks were counted (bedtools, version 2.17.0)⁶¹. Their differential enrichments and depletions were analyzed by the DESeq package⁶² in R. Significantly altered peaks were defined by an FDR less than 0.01. GO analysis was performed by iPAGE (<https://tavazoelab.c2b2.columbia.edu/iPAGE/>)⁶³.

RNA-Seq/TRAP-Seq/Ribosome profiling. Data were processed as previously described^{8,36,64}. The A-site position along the ribosome footprints was empirically assigned as 15 nt for 25–31 nt footprints.

Code is available upon request.

Data availability. The sequencing data obtained in this study were deposited in NCBI GEO [GSE100628].

References

- Greig, L. C., Woodworth, M. B., Galazo, M. J., Padmanabhan, H. & Macklis, J. D. Molecular logic of neocortical projection neuron specification, development and diversity. *Nat Rev Neurosci* **14**, 755–769, <https://doi.org/10.1038/nrn3586> (2013).
- Johnson, M. B. & Walsh, C. A. Cerebral cortical neuron diversity and development at single-cell resolution. *Curr Opin Neurobiol* **42**, 9–16 (2017).
- Molyneaux, B. J., Arlotta, P., Menezes, J. R. & Macklis, J. D. Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* **8**, 427–437, <https://doi.org/10.1038/nrn2151> (2007).
- Charvet, C. J., Cahalane, D. J. & Finlay, B. L. Systematic, cross-cortex variation in neuron numbers in rodents and primates. *Cereb Cortex* **25**, 147–160, <https://doi.org/10.1093/cercor/bht214> (2015).
- Okaty, B. W., Sugino, K. & Nelson, S. B. Cell type-specific transcriptomics in the brain. *J Neurosci* **31**, 6939–6943, <https://doi.org/10.1523/JNEUROSCI.0626-11.2011> (2011).
- Steward, O. & Schuman, E. M. Protein synthesis at synaptic sites on dendrites. *Annual review of neuroscience* **24**, 299–325, <https://doi.org/10.1146/annurev.neuro.24.1.299> (2001).
- Heiman, M. *et al.* A translational profiling approach for the molecular characterization of CNS cell types. *Cell* **135**, 738–748, <https://doi.org/10.1016/j.cell.2008.10.028> (2008).
- Hornstein, N. *et al.* Ligation-free ribosome profiling of cell type-specific translation in the brain. *Genome Biol* **17**, 149, <https://doi.org/10.1186/s13059-016-1005-1> (2016).
- Sanz, E. *et al.* Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci USA* **106**, 13939–13944, <https://doi.org/10.1073/pnas.0907143106> (2009).
- Hung, T. & Chang, H. Y. Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA Biol* **7**, 582–585 (2010).
- Mercer, T. R. & Mattick, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* **20**, 300–307, <https://doi.org/10.1038/nsmb.2480> (2013).
- Smalheiser, N. R. The RNA-centred view of the synapse: non-coding RNAs and synaptic plasticity. *Philos Trans R Soc Lond B Biol Sci* **369**, <https://doi.org/10.1098/rstb.2013.0504> (2014).
- Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33**(Suppl), 245–254, <https://doi.org/10.1038/ng1089> (2003).
- Margueron, R. & Reinberg, D. Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* **11**, 285–296, <https://doi.org/10.1038/nrg2752> (2010).
- Riccio, A. Dynamic epigenetic regulation in neurons: enzymes, stimuli and signaling pathways. *Nat Neurosci* **13**, 1330–1337, <https://doi.org/10.1038/nn.2671> (2010).
- Kozlenkov, A. *et al.* Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. *Nucleic Acids Res* **42**, 109–127, <https://doi.org/10.1093/nar/gkt838> (2014).
- Iwamoto, K. *et al.* Neurons show distinctive DNA methylation profile and higher interindividual variations compared with non-neurons. *Genome Res* **21**, 688–696, <https://doi.org/10.1101/gr.112755.110> (2011).
- Niwa, M. *et al.* Adolescent stress-induced epigenetic control of dopaminergic neurons via glucocorticoids. *Science* **339**, 335–339, <https://doi.org/10.1126/science.1226931> (2013).
- Guo, J. U. *et al.* Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nat Neurosci* **14**, 1345–1351, <https://doi.org/10.1038/nn.2900> (2011).
- Barth, T. K. & Imhof, A. Fast signals and slow marks: the dynamics of histone modifications. *Trends Biochem Sci* **35**, 618–626, <https://doi.org/10.1016/j.tibs.2010.05.006> (2010).

21. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **21**, 70–71, <https://doi.org/10.1038/50071999> (1999).
22. Abe, T. *et al.* Establishment of conditional reporter mouse lines at ROSA26 locus for live cell imaging. *Genesis* **49**, 579–590, <https://doi.org/10.1002/dvg.20753> (2011).
23. Erdmann, G., Schutz, G. & Berger, S. Inducible gene inactivation in neurons of the adult mouse forebrain. *BMC Neurosci* **8**, 63, <https://doi.org/10.1186/1471-2202-8-63> (2007).
24. Gallardo, T., Shirley, L., John, G. B. & Castrillon, D. H. Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. *Genesis* **45**, 413–417, <https://doi.org/10.1002/dvg.20310> (2007).
25. Shioi, G. *et al.* A mouse reporter line to conditionally mark nuclei and cell membranes for *in vivo* live-imaging. *Genesis* **49**, 570–578, <https://doi.org/10.1002/dvg.20758> (2011).
26. Oszolák, F., Song, J. S., Liu, X. S. & Fisher, D. E. High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* **25**, 244–248, <https://doi.org/10.1038/nbt1279> (2007).
27. Flensburg, C., Kinkel, S. A., Keniry, A., Blewitt, M. E. & Oshlack, A. A comparison of control samples for ChIP-seq of histone modifications. *Front Genet* **5**, 329, <https://doi.org/10.3389/fgene.2014.00329> (2014).
28. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88, <https://doi.org/10.1016/j.cell.2007.05.042> (2007).
29. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311–318, <https://doi.org/10.1038/ng1966> (2007).
30. Koyama, M. & Kurumizaka, H. Structural diversity of the nucleosome. *J Biochem*. <https://doi.org/10.1093/jb/mvx081> (2017).
31. Buschbeck, M. & Hake, S. B. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol* **18**, 299–314, <https://doi.org/10.1038/nrm.2016.166> (2016).
32. Takamori, S., Rhee, J. S., Rosenmund, C. & Jahn, R. Identification of a vesicular glutamate transporter that defines a glutamatergic phenotype in neurons. *Nature* **407**, 189–194, <https://doi.org/10.1038/35025070> (2000).
33. Cahoy, J. D. *et al.* A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* **28**, 264–278, <https://doi.org/10.1523/JNEUROSCI.4178-07.2008> (2008).
34. Park, K. W. *et al.* Robo4 is a vascular-specific receptor that inhibits endothelial migration. *Dev Biol* **261**, 251–267 (2003).
35. Lampugnani, M. G. *et al.* The molecular organization of endothelial cell to cell junctions: differential association of plakoglobin, beta-catenin, and alpha-catenin with vascular endothelial cadherin (VE-cadherin). *J Cell Biol* **129**, 203–217 (1995).
36. Zhang, Y. *et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**, 11929–11947, <https://doi.org/10.1523/JNEUROSCI.1860-14.2014> (2014).
37. Holt, C. E. & Schuman, E. M. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* **80**, 648–657, <https://doi.org/10.1016/j.neuron.2013.10.036> (2013).
38. Cajigas, I. J. *et al.* The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453–466, <https://doi.org/10.1016/j.neuron.2012.02.036> (2012).
39. Mahmoudi, E. & Cairns, M. J. MiR-137: an important player in neural development and neoplastic transformation. *Mol Psychiatry* **22**, 44–55, <https://doi.org/10.1038/mp.2016.150> (2017).
40. Makeyev, E. V., Zhang, J., Carrasco, M. A. & Maniatis, T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell* **27**, 435–448, <https://doi.org/10.1016/j.molcel.2007.07.015> (2007).
41. Zeng, Y. Regulation of the mammalian nervous system by microRNAs. *Mol Pharmacol* **75**, 259–264, <https://doi.org/10.1124/mol.108.052118> (2009).
42. Ip, J. Y. *et al.* Gomafu lncRNA knockout mice exhibit mild hyperactivity with enhanced responsiveness to the psychostimulant methamphetamine. *Sci Rep* **6**, 27204, <https://doi.org/10.1038/srep27204> (2016).
43. Sone, M. *et al.* The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J Cell Sci* **120**, 2498–2506, <https://doi.org/10.1242/jcs.009357> (2007).
44. Nakagawa, S., Naganuma, T., Shioi, G. & Hirose, T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J Cell Biol* **193**, 31–39, <https://doi.org/10.1083/jcb.201011110> (2011).
45. Nishimoto, Y. *et al.* The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol Brain* **6**, 31, <https://doi.org/10.1186/1756-6606-6-31> (2013).
46. Raj, A. & McVicker, G. The genome shows its sensitive side. *Nat Methods* **11**, 39–40, <https://doi.org/10.1038/nmeth.2770> (2014).
47. Weill, L., Belloc, E., Bava, F. A. & Mendez, R. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat Struct Mol Biol* **19**, 577–585, <https://doi.org/10.1038/nsmb.2311> (2012).
48. Anderson, P. & Kedersha, N. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* **10**, 430–436, <https://doi.org/10.1038/nrm2694> (2009).
49. Voigt, P., Tee, W. W. & Reinberg, D. A double take on bivalent promoters. *Genes Dev* **27**, 1318–1338, <https://doi.org/10.1101/gad.219626.113> (2013).
50. Molden, R. C., Bhanu, N. V., LeRoy, G., Arnaudo, A. M. & Garcia, B. A. Multi-faceted quantitative proteomics analysis of histone H2B isoforms and their modifications. *Epigenetics Chromatin* **8**, 15, <https://doi.org/10.1186/s13072-015-0006-8> (2015).
51. Rea, M. *et al.* Quantitative Mass Spectrometry Reveals Changes in Histone H2B Variants as Cells Undergo Inorganic Arsenic-Mediated Cellular Transformation. *Mol Cell Proteomics* **15**, 2411–2422, <https://doi.org/10.1074/mcp.M116.058412> (2016).
52. Adli, M., Zhu, J. & Bernstein, B. E. Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors. *Nat Methods* **7**, 615–618, <https://doi.org/10.1038/nmeth.1478> (2010).
53. Shankaranarayanan, P. *et al.* Single-tube linear DNA amplification (LinDA) for robust ChIP-seq. *Nat Methods* **8**, 565–567, <https://doi.org/10.1038/nmeth.1626> (2011).
54. Jakobsen, J. S. *et al.* Amplification of pico-scale DNA mediated by bacterial carrier DNA for small-cell-number transcription factor ChIP-seq. *BMC Genomics* **16**, 46, <https://doi.org/10.1186/s12864-014-1195-4> (2015).
55. Nakagawa, S. Analysis of the subcellular distribution of RNA by fluorescence *in situ* hybridization. *Methods Mol Biol* **1206**, 107–122, https://doi.org/10.1007/978-1-4939-1369-5_10 (2015).
56. Kadota, M. *et al.* CTCF binding landscape in jawless fish with reference to Hox cluster evolution. *Sci Rep* **7**, 4957, <https://doi.org/10.1038/s41598-017-04506-x> (2017).
57. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, <https://doi.org/10.1186/gb-2009-10-3-r25> (2009).
58. Song, L., Florea, L. & Langmead, B. Lighter: fast and memory-efficient sequencing error correction without counting. *Genome Biol* **15**, 509, <https://doi.org/10.1186/s13059-014-0509-9> (2014).
59. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352> (2009).
60. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, <https://doi.org/10.1186/gb-2008-9-9-r137> (2008).
61. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842, <https://doi.org/10.1093/bioinformatics/btq033> (2010).
62. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106, <https://doi.org/10.1186/gb-2010-11-10-r106> (2010).

63. Goodarzi, H., Elemento, O. & Tavazoie, S. Revealing global regulatory perturbations across human cancers. *Mol Cell* **36**, 900–911, <https://doi.org/10.1016/j.molcel.2009.11.016> (2009).
64. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* **7**, 1534–1550, <https://doi.org/10.1038/nprot.2012.086> (2012).

Acknowledgements

We would like to thank Dr. Shigehiro Kuraku, Ms. Chieko Nashiki, and Kaori Yanaka for technical support, Dr. Rei Yoshimoto for practical advice on sequencing data analyses, Dr. Yukiko Goto for discussion, and the members of the Iwasaki laboratory for critical reading of the manuscript. This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas (#26113005 to S.N. and JP17H05679 to S.I.), a Grant-in-Aid for Young Scientists (A) (JP17H04998 to S.I.) from the Ministry of Education, Science, Sports, and Culture of Japan (MEXT), and the Pioneering Projects “Cellular Evolution” and “Disease and Epigenome” from RIKEN (to S.N. and S.I.).

Author Contributions

M.M. performed the molecular cloning, ChIP, and immunohistochemistry experiments and the data analyses, M.K. and K.T. performed ChIP sequencing, Y.F. created genetically modified mice, K.A. provided the Vasa-Cre mice, S.I. performed and designed the data analyses, and S.N. organized the experiments.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-19494-9>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018