Dilated cardiomyopathy-associated RNA Binding Motif Protein 20 regulates long pre-mRNAs in neurons

Giulia Di Bartolomei[1], Raul Ortiz[1], Dietmar Schreiner[1], Susanne Falkner[1], Esther E. Creemers[2], Peter Scheiffele[1]*

[1]Biozentrum, University of Basel, Spitalstrasse 41, CH-4056, Basel, Switzerland

[2] Department of Experimental Cardiology, Amsterdam Cardiovascular Sciences, Amsterdam UMC, location University of Amsterdam, 1105AZ, Amsterdam, The Netherlands.

*Corresponding author: peter.scheiffele@unibas.ch

Author Contributions: G.D.B., R.O., D.D. and P.S. designed and analyzed experiments, G.D.B. performed molecular and anatomical studies, G.D.B. and R.O. performed computational analysis of gene expression and CLIP data, S.F. performed two-photon imaging experiments, D.S. performed CLIP and molecular biology experiments, E.E.C. provided materials and experimental advice, G.D.B. and P.S. wrote the manuscript with editing provided by all co-authors.

Keywords: alternative splicing, RNA binding proteins, neural development, cardiomyopathy, neurodevelopmental disorder

This PDF file includes:

Main Text
Figures 1 to 5
Supplementary Material

Abstract
Precise coordination of molecular programs and neuronal growth govern the formation, maintenance, and adaptation of neuronal circuits. RNA metabolism has emerged as a key regulatory node of neural development and nervous system pathologies. To uncover cell-type-specific RNA regulators, we systematically investigated expression of RNA recognition motif-containing proteins in the mouse neocortex. Surprisingly, we found RBM20, an alternative splicing regulator associated with dilated cardiomyopathy, to be expressed in cortical parvalbumin interneurons and mitral cells of the olfactory bulb. Genome-wide mapping of RBM20 target mRNAs revealed that neuronal RBM20 binds pre-mRNAs in distal intronic regions. Loss of neuronal RBM20 has only modest impact on alternative splice isoforms but results in a significant reduction in an array of mature mRNAs in the neuronal cytoplasm. This phenotype is particularly pronounced for genes with long introns that encode synaptic proteins. We hypothesize that RBM20 ensures fidelity of pre-mRNA splicing by suppressing non-productive splicing events in long neuronal genes. This work highlights a common requirement for RBM20-dependent transcriptome regulation in cardiomyocytes and neurons and demonstrates that a major genetic risk factor of heart disease impacts neuronal gene expression.

Introduction

Neurons exhibit complex transcriptional programs that instruct the specification of functionally distinct neuronal cell types. Functional and anatomical properties of neurons emerge during development. However, the molecular underpinnings that define cell type-specific properties are only beginning to be elucidated. RNA-binding proteins (RBPs) have emerged as key regulators of neuronal function through modification of mRNA processing, localization, stability, and translation (Ule & Darnell, 2006; Babitzke *et al*, 2009; Vuong *et al*, 2016; Mauger & Scheiffele, 2017; Holt *et al*, 2019; Ule & Blencowe, 2019; Gomez *et al*, 2021). Moreover, RBP dysfunction is a significant contributor to pathologies, including neurodevelopmental and neurodegenerative conditions (Ling *et al*, 2013; Lopez Soto *et al*, 2019; Gebauer *et al*, 2021; Schieweck *et al*, 2021). Here, we discovered an unexpected neuronal function for the SR-related protein, RNA Binding Motif Protein 20 (RBM20). Thus far, RBM20 was considered to be muscle-specific and to represent a key alternative splicing regulator in cardiomyocytes. Mutations in the *RBM20* gene are linked to an aggressive form of dilated cardiomyopathy (Brauch *et al*, 2009; Parikh *et al*, 2019). In cardiomyocytes, RBM20 protein controls alternative exon usage of transcripts encoding key sarcomere components (Titin and Tropomyosin) and proteins involved in calcium signaling, such as CAMK2D and the α1 subunit of the L-type voltage gated calcium channel (CACNA1C) (Guo *et al*, 2012; Maatz *et al*, 2014; van den Hoogenhof *et al*, 2018; Zhu *et al*, 2021). RBM20 contains an RNA Recognition Motif (RRM) domain, two zinc finger motifs, and an extended arginine/serine-rich region. This domain organization is shared with the paralogues Matrin-3 (MATR3) and ZNF638 (Watanabe *et al*, 2018), and is similar to FUS and TDP43, two RNA/DNA binding proteins that regulate various steps of RNA metabolism. Comprehensive RBM20 protein-RNA interaction maps have been defined for cardiomyocytes (Maatz *et al.*, 2014; van den Hoogenhof *et al.*, 2018; Briganti *et al*, 2020). However, RBM20 expression and function in the brain, are unknown. Considering its highly selective expression in specific cell populations and the critical roles in cardiomyocytes, we hypothesized that neuronal RBM20 controls key steps of RNA metabolism and contributes to the regulation of neuronal gene expression.

Results

RBM20 is selectively expressed in specific GABAergic and glutamatergic neurons

To discover regulators of neuronal cell type-specific transcriptomes, we analyzed a neuronal gene expression dataset covering cortical pyramidal cells and the major GABAergic interneuron populations of the mouse cortex (Furlanis *et al*, 2019). We generated a hand-curated list of 234 potential RBPs, selected based on the presence of at least one predicted RRM domain (Fig. 1A and Fig. S1A). Transcripts for 182 of these RRM proteins were significantly expressed in mouse cortical neurons. While some (such as *Eif3g*, *Hnrnpl*) were broadly detected across all neuronal populations, others were

differentially expressed in specific neuron classes. Thus, *Rbm38* was almost exclusively expressed in VIP$^+$ (vasoactive intestinal polypeptide-positive) interneurons and *Ppargc1b and Rbm20* in parvalbumin-positive (PV$^+$) interneurons, indicating that they may play an important role in cell type-specific gene regulation. Identification of *Rbm20* expression in the brain was surprising considering that the RBM20 protein was thought to be muscle-specific. Thus, we focused our further studies on probing potential functions of RBM20 in the mouse brain.

To validate the neuronal expression, we investigated *Rbm20* mRNA distribution in the mouse neocortex by RNA fluorescent *in situ* hybridization (FiSH). This analysis confirmed significant expression of *Rbm20* in PV$^+$ GABAergic interneurons (Fig. S1B-D). Moreover, we detected low *Rbm20* mRNA levels in somatostatin-positive (SST$^+$) interneurons of the somatosensory cortex (Fig. S1B - D). A survey of open-source datasets (Sjostedt *et al*, 2020) revealed that *Rbm20* mRNA expression in PV$^+$ interneurons is evolutionary conserved in human. To examine RBM20 at the protein level, we raised polyclonal antibodies and confirmed RBM20 expression in PV$^+$ interneurons of the somatosensory cortex (Fig. 1B and C). These experiments also uncovered RBM20 expression outside of the neocortex, with particularly high protein levels in the olfactory bulb (Fig. 1D and E). Here, RBM20 was restricted to glutamatergic cells in the mitral cell layer (MCL) and glomerular layer (GL) as opposed to the expression in GABAergic populations in the somatosensory cortex (Fig. 1D). Genetic marking and *in situ* hybridization for markers *Vglut2* and *Tbr2* confirmed the glutamatergic neuron expression of *Rbm20* mRNA in tufted and mitral cells (Fig S1E - G), with only low levels in a small number of GABAergic neurons (Fig S1H). Similarly, we detected RBM20 protein in mitral cells of newborn mice (genetically marked with Tbx^{cre}*::*$Ai9^{tdTOM}$), suggesting that its expression is initiated during development and continues into the adult (Fig. 1F).

Neuronal RBM20 binds distal intronic regions of mRNAs encoding synaptic proteins

In cardiomyocytes, RBM20 acts as an alternative splicing regulator that gates exon inclusion by binding to exon-proximal intronic splicing silencers (Li *et al*, 2013; Maatz *et al.*, 2014; Dauksaite & Gotthardt, 2018). Within the nucleus, RBM20 localizes to specific foci in close proximity to *Titin* transcripts (Guo *et al.*, 2012; Li *et al.*, 2013) and has been proposed to form "splicing factories", sites for coordinated processing of mRNAs derived from multiple genes (Bertero *et al*, 2019). By contrast, neuronal RBM20 did not concentrate in nuclear foci but instead, was distributed throughout the nucleoplasm (Fig. 2A). This difference in the sub-nuclear localization likely arises from the lack of *Titin* and/or other similarly abundant mRNA targets in neuronal cells. To contrast neuronal and cardiomyocyte RNA targets, we mapped transcripts directly bound by RBM20 in the olfactory bulb and in the heart using cross-linking immunoprecipitation followed by sequencing (CLIP-seq) analysis (Ule *et al*, 2003; Van Nostrand *et al*, 2017b). Our RBM20 antibodies lacked sufficient affinity for immunoprecipitation. Thus, we generated a $Rbm20^{HA}$ knock-in mouse line where the endogenous RBM20 protein is tagged with a Histidine-Biotin acceptor peptide and a triple HA-epitope (Fig. 2B and C and Fig. S2A and B). Introduction of this tag did not alter overall RBM20 protein levels or expression pattern (Fig. S2B, C).

RBM20 CLIP-seq analysis on heart and olfactory bulb tissues from P35-P40 $Rbm20^{HA}$ knock-in mice identified significant peaks in 956 and 2707 unique transcripts, respectively. The CLIP tags obtained in biological replicates were highly correlated (two replicates for heart and three for olfactory bulb, Fig. S3A and B). In both tissues, the majority of the identified RBM20 CLIP peaks mapped to introns (80% in the heart and 94% in the olfactory bulb) (Fig. 2D and E). Only a low fraction of peaks was identified in the 3' untranslated region (11% in the heart and 1% in the olfactory bulb) and coding regions (7% in the heart and 3% in the olfactory bulb). RBM20 targets detected in the heart recapitulated all major target mRNAs identified in previous studies, including *Ttn, Tpm, Pdlim5, Scn5a, Camk2d, Cacna1c, Cacna1d* (Guo *et al.*, 2012; Maatz *et al.*, 2014; van den Hoogenhof *et al.*, 2018; Watanabe *et al.*, 2018; Fenix *et al*, 2021). In heart and olfactory bulb, approximately 90% of the intronic RBM20 binding sites localized to distal regions (> 500 bp from splice site) (Fig. 2F). The previously reported UCUU motif (Guo *et al.*, 2012; Upadhyay & Mackereth, 2020) was significantly enriched at cross-link-induced truncation sites (CITS) (Fig. S3C) and *de novo* binding motif enrichment analysis identified a corresponding U-rich motif at RBM20 bound sites (Fig. S2D).

When directly comparing heart and olfactory bulb CLIP-seq datasets, we found transcripts from 363 genes that were commonly bound by RBM20 in both tissues. Other transcripts like *Ttn* are

recovered as RBM20-bound only in one of the two tissues due to tissue-specific expression of the mRNA (Fig. 3A). Interestingly, *Camk2d* and *Cacna1c*, two important targets of RBM20 alternative splicing regulation in the heart, were identified as RBM20-bound in the olfactory bulb, however, binding sites mapped to different introns (Fig. 3A, and see TableS1). Finally, there was a sizable portion of pre-mRNAs (from 2721 genes) commonly expressed in both heart and olfactory bulb but selectively recovered as RBM20-bound in only one of the two tissues (539 genes in the heart and 2182 in the olfactory bulb). This suggests that RBM20 binds to target mRNAs in a tissue-specific manner. A tissue-specific function of RBM20 was further supported by gene ontology (GO) analysis. RBM20 target mRNAs in the heart showed enrichment for terms of muscle fiber components (M-band, Z-disc, T-tubule, Fig. 3B). By contrast, target genes in the olfactory bulb showed enrichment for terms related to pre- and postsynaptic structures, ion channels, and cytoskeletal components (Fig 3C and Table S2). This suggests that neuronal RBM20 plays a role in the regulation of synapse-related mRNA transcripts.

Neuronal RBM20 is required for normal expression of long pre-mRNAs encoding synaptic proteins

To investigate the impact of RBM20 on the neuronal transcriptome, we performed loss-of-function experiments in *Rbm20* global and conditional knock-out mice. *Rbm20* was conditionally inactivated selectively in either parvalbumin-positive interneurons (*Pvalb*^{cre}*::Rbm20*$^{fl/fl}$, referred to as "*Rbm20*$^{\Delta PV}$") or glutamatergic neurons (*Vglut2*^{cre}*::Rbm20*$^{fl/fl}$, referred to as "*Rbm20*$^{\Delta Vglut2}$"). In global *Rbm20* knock-out mice, RBM20 immune-reactivity was abolished in the olfactory bulb and cells of the somatosensory cortex (Fig.S4 and see Fig. 1E for western blots). Similarly, in the conditional knockout mice we observed a loss of RBM20 immune-reactivity in the respective cell populations (Fig. S4A and B). Importantly, neuronal cell types were normally specified in *Rbm20*$^{\Delta Vglut2}$ mice (Fig. S4C-F). We then applied a Translating Ribosome Affinity Purification protocol (RiboTRAP) optimized for small tissue samples (Heiman *et al*, 2014; Sanz *et al*, 2019; Di Bartolomei & Scheiffele, 2022) to uncover the impact of RBM20 loss-of-function on the neuronal transcriptome. Isolations were performed for *Rbm20*$^{\Delta PV}$ and *Rbm20*$^{\Delta Vglut2}$ conditional knock-out and matching *Rbm20*^{WT} mice (postnatal day 35-40). For quality control, we confirmed appropriate enrichment and de-enrichment of cell type-specific markers (Fig. S5A, B). Subsequently, we assessed the transcriptomes by deep RNA sequencing (4-5 replicates per condition, male and female mice, paired-end, 150 base pair reads, >80 million reads per replicate, see TableS3 and Fig. S5C - F for details). Differential gene expression analysis by DESEQ2 did not identify significant differences in the overall transcriptome of *Rbm20*$^{\Delta PV}$ PV^{+} interneurons as compared to wild-type (Fig. 4A and Table S4). However, in the olfactory bulb glutamatergic cells isolated from Rbm20$^{\Delta Vglut2}$ mice we identified de-regulation of 409 genes (FC ≥ 1.5, adjusted p-value < 0.01). 256 of these genes showed decreased expression in the conditional knockout mice as opposed to transcripts from 153 transcripts that were elevated (Fig. 4B). In the heart, RBM20 is considered a major regulator of alternative splicing. Thus, we quantified shifts in alternative exon usage upon *Rbm20* loss-of-function. We identified 859 differentially regulated alternative exons in PV^{+} interneurons and 1924 exons in the *Vglut2*$^{+}$ population in the olfactory bulb (FC in splicing index ≥ 1.5, p-value ≤ 0.05) (Fig. 4C and Table S5). The functions of the gene products with alternative splicing de-regulation were diverse and the only significantly enriched gene ontology term for genes with de-regulated exons was "mitochondrial protein-containing complexes" (p-value = 0.0003, 100 genes identified) for the olfactory bulb (*Vglut2*$^{+}$) population (Table S6).

To uncover transcript isoforms directly regulated by RBM20, we probed the intersection of the RBM20 binding sites identified by CLIP and alternative exon expression in the *Vglut2*$^{+}$ cell population. Of the 1924 exons with differential incorporation in *Rbm20*$^{\Delta Vglut2}$ cells, there were 659 exons with at least one significant RBM20 CLIP peak mapping to the corresponding gene. This frequency was higher than expected for a random distribution of CLIP peaks over all genes expressed in the sample ($p = 5.2 * 10^{-8}$, hypergeometric distribution test for enrichment analysis). Interestingly, the vast majority of CLIP peaks were distant (>10kb) from differentially expressed exons. This applied regardless of the type of alternative exon feature (alternative last exon, alternative transcription start sites, cassette exons or complex alternative splicing events). Thus, we identified only 1 differentially regulated cassette exon

with a proximal RBM20 binding event (< 500bp from the regulated exon) in the gene *Arhgef1*, (Fig. 4E). This suggests that suppression of alternative exons through proximal intronic splicing silencer elements is unlikely to be the primary essential function of neuronal RBM20.

Overall, RBM20 binding sites identified by CLIP distributed throughout the entire length of introns and did not exhibit a particular concentration at exon-intron boundaries (Fig. 5A). When examining the intersection of differential gene expression and CLIP data, we observed that 129 of the 256 down-regulated transcripts contained CLIP peaks (50.4%) whereas only 7 of the 153 up-regulated transcripts were bound by RBM20 (4.5%) (Fig. 5B). This suggests that RBM20 loss-of-function in neuronal cells results in the loss of mRNAs whereas the observed elevation of selective transcripts is likely to be an indirect, compensatory response. This notion is further supported by the distinct gene ontologies of the de-regulated transcripts. Down-regulated genes displayed a significant enrichment in GO terms related to synaptic components and the cytoskeleton (Fig. 5C and Table S6). By contrast, for the up-regulated genes, there was a pronounced enrichment in mitochondrial components and ribosome composition (Fig. 5C and see Table S6).

TDP-43 and MATR3, two proteins with similar domain organization to RBM20, regulate target mRNAs by suppressing aberrant splicing into cryptic exons, a phenotype most significant for genes with long introns (Polymenidou *et al*, 2011; Ling *et al*, 2015; Attig *et al*, 2018). Interestingly, intron length of transcripts showed a significant correlation with differential gene expression and the number of CLIP peaks in *Rbm20*$^{\Delta Vglut2}$ cells (Fig. 5D - F). Importantly, mean intron length was substantially larger for down-regulated transcripts when compared to all detected genes or to up-regulated genes (Fig. 5E). This difference was amplified when selectively examining introns of down-regulated genes with CLIP peaks (Fig. 5F and G). Finally, introns bound by RBM20 were significantly longer than expected by chance as assed with a permutation test. Random regions with the same properties as RBM20 CLIP peak regions were generated on introns from genes expressed in our Ribo-TRAP dataset (see methods). This resulted in a mean expected intron size of 41.5 kb which is substantially smaller than the mean size of RBM20-bound introns (59.0 kb; p-value < 0.0002). Thus, long neuronal mRNAs are preferentially bound by RBM20 and particularly sensitive to loss of neuronal RBM20 protein.

Discussion

Our work establishes an unexpected function for RBM20 in neuronal RNA metabolism. We identified RBM20 expression in two specific neuronal populations in the mouse brain. Interestingly, these two populations are derived from highly divergent lineages: glutamatergic neurons of the olfactory bulb from the rostral telencephalon and PV^{+} GABAergic interneurons in the somatosensory cortex which arise from the medial ganglionic eminence of the subcortical telencephalon. Genetic deletion of RBM20 had only a modest impact on gene expression and transcript isoforms in PV^{+} interneurons but was associated with substantial alterations in glutamatergic cells of the olfactory bulb. This might be due to the higher expression of RBM20 in mitral and tufted cells as compared to PV^{+} interneurons. Moreover, PV^{+} interneurons express high levels of the RBM20 paralogue MATR3 (Fig.S1A) which may have overlapping functions.

In cardiomyocytes, several direct targets of RBM20-dependent alternative splicing regulation contain intronic RBM20 binding sites in close proximity to regulated exons (Maatz *et al.*, 2014; van den Hoogenhof *et al.*, 2018). In neuronal cells, we did not observe a similar regulation of proximal exons by RBM20. Interestingly, a large fraction of transcripts down-regulated in olfactory bulb RBM20 knock-out cells contained intronic RBM20 binding sites. Transcripts with long introns – which are prominently expressed in the mouse brain (Sibley *et al*, 2015; Zylka *et al*, 2015) – were particularly sensitive to RBM20 loss. Thus, in neuronal cells, the splicing repressive function of RBM20 might prevent the recruitment of cryptic splice acceptor sites within long intronic segments. We hypothesize that the reduced expression of RBM20 target mRNAs observed in *Rbm20*$^{\Delta Vglut2}$ neurons arises from aberrant nuclear splicing which ultimately result in the degradation of the transcripts.

The association of *RBM20* mutations with cardiomyopathies have directed substantial research efforts to uncover its function in the heart and to understand the impact of cardiomyopathy-associated mutations (Khan *et al*, 2016; Briganti *et al.*, 2020; Schneider *et al*, 2020; Fenix *et al.*, 2021). As for mice, human *RBM20* mRNA is also expressed in PV^+ interneurons (Sjostedt *et al.*, 2020). There is a growing body of literature highlighting shared genetic etiology of congenital heart disease and neurodevelopmental disorders (Homsy *et al*, 2015; Jin *et al*, 2017; Rosenthal *et al*, 2021). A significant fraction of children with congenital heart disease exhibit autistic traits including problems with theory of mind and cognitive flexibility (Marino *et al*, 2012) and the probability of an individual with congenital heart disease being diagnosed with autism was estimated to be more than twofold higher than in the typically developing population (Gu *et al*, 2023). A homozygous Ser529Arg substitution in the RNA recognition domain of RBM20 was recently identified in two brothers of consanguineous families affected by epilepsy and developmental delay (Badshah *et al*, 2022). The *RBM20* variant segregated with a second mutation in the *CNTNAP2* gene, which encodes a neuronal transmembrane protein implicated in axon-glia interactions. Thus, it remains to be explored whether alterations in neuronal RBM20 function contributes to the clinical characteristics observed in these individuals. However, given the neuronal expression and function of RBM20 identified in our study, a survey of – thus far unexplored – neurological phenotypes in *RBM20* mutation carriers might be warranted.

Materials and Methods

Immunochemistry and imaging

Animals (males and females) from postnatal day 25 to 40 were anesthetized with ketamine/xylazine (100/10 mg/kg *i.p.*) and transcardially perfused with fixative (4% paraformaldehyde). The brains or hearts were post-fixed overnight in the same fixative at 4°C and washed 3 times with 100 mM phosphate buffer. Coronal brain slices were cut at 40 µm with a vibratome (Leica Microsystems VT1000).

For immunohistochemistry, brain sections were processed as previously described (Traunmuller *et al*, 2023). In brief, brain slices were kept for 1.5 h with a PBS-based blocking solution containing 0.1% Triton X-100 and 5% normal donkey serum (NDS) and subsequently incubated with primary antibodies in blocking solution at 4°C overnight. Secondary antibodies were diluted in 5% NDS in PBS containing 0.05% Triton-X100 to a final concentration of 0.5 µg/ml or 1.0 µg/ml and incubated with sections for 2 h at RT.

Image stacks were acquired at room temperature on a laser-scanning confocal microscope (Zeiss LSM700) using a 40x Apochromat objective with numerical aperture 1.3, controlled by Zen 2010 software. Following acquisition, images were processed and assembled using Fiji (Schindelin *et al*, 2012), OMERO, and Adobe Illustrator software.

Surgeries and stereotaxic injections.

Recombinant Adeno-associated viruses with the rAAV2 capsid (Tervo *et al*, 2016) were produced in HEK293T cells. In brief, initially, a DNA mixture consisting of 70ug AAV helper plasmid, 70ug AAV vector, and 200ug pHGTI-adeno1 is prepared in a falcon tube and added at a 1:4 DNA:PEI ratio to each cell plate. After 48-60 hours, cells are collected and debris are collected by centrifugation at 4,000 rpm, 4ºC for 20 minutes. The supernatant containing the virus is subsequently purified through OptiPrep™ Density Gradient Medium, (Sigma, Cat. No D1556). Viral preparations are concentrated in 100K Millipore Amicon columns at 4°C. The virus samples were then suspended in PBS 1X, aliquoted and stored at -80°C. Viral titers were determined by qPCR and were $>10^{12}$ particles/mL.

Mice (postnatal day 24 to 27) were placed on a heating pad in a stereotaxic frame (Kopf Instrument) under isoflurane anesthesia. A small incision (0.5–1 cm) in the skin overlying the area of interest was made, and bilateral viral injections were performed in the posterior Piriform Cortex using a Picospritzer III pressure injection system (Parker) with borosilicate glass capillaries (length 100 mm, OD 1 mm, ID 0.25 mm, wall thickness 0.375 mm). Coordinates: ML = + 2,2 mm, AP = + 2,35 mm, DV = - 3,95 mm from Bregma. A volume of 100 nl of virus was delivered to each side, through repeated puffs over several minutes. Viruses used were rAAV2-CAG-DiO-eGFP or rAAV2-SYN-Cre virus, driving cre-dependent eGFP expression from the chicken beta actin promoter and human synapsin promoter, respectively. Ten days after viral infection, mice were anesthetized and transcardially perfused. Position

of the viral injection site was confirmed on coronal brain slices using DAPI-stained sections on a AxioScan.Z1 Slide scanner (Zeiss) using a 20x objective.

Tissue clearing and anatomical reconstruction of mitral cells

The olfactory bulb and part of the anterior prefrontal cortex of rAAV2-infected mice were dissected after transcardial perfusion and cleared using the Cubic L protocol (Tainaka *et al*, 2018). Olfactory bulbs were placed into a 5 ml Eppendorf tube filled with pre-warmed CUBIC L solution (10% N-butyl-di-ethanolamine and 10% Triton X-100 dissolved in MilliQ water). The tissue was incubated on a shaking plate at 37°C for 48 h. Cleared bulbs were washed in 50 mM PBS 3 times for 10 min and then cut into two coronal halves under a Binocular Stereo Microscope (Olympus #MVX10). The two halves of each bulb (anterior or posterior) were then embedded in 1% agarose in TBE solution in a glass bottom imaging chamber (Ibidi, Cat.No:80426). Z-stacks of GFP^+ neurons with the soma residing in the mitral cell layer of the olfactory bulb were acquired on a Olympus two-photon microscope fitted with a MaiTai eHP laser (Spectra-Physics) and a 25X objective with 1.05 NA. A volume of up to 2 x 2 mm (xy) x 500-700 µm in depth was acquired in tiles up to 4x4, with x = 0.995 µm, y = 0.995 µm and z = 3 µm pixel size. Laser power was linearly adjusted with imaging depth and typically ranged between 0.5 to 20 mW. GFP^+ mitral cells in the two-photon z-stacks were traced neurons using Neurolucida 360 software.

Fluorescent *in situ* hybridization

Fluorescent *in situ* hybridization was performed using the RNAScope Fluorescent Multiplex Kit (Advanced Cell Diagnostics, Catalog Number 320851). P25 mouse brains were snap frozen in liquid nitrogen and 15 µm coronal sections were cut on a cryostat (Microm HM560, Thermo Scientific). Sections were fixed at 4°C overnight with 4% paraformaldehyde in 100 mM PBS, pH 7.4.

Images were acquired at room temperature with an upright LSM700 confocal microscope (Zeiss) using 40X Apochromat objectives (NA=1.3). Stacks of 10-15 µm thickness (0.44 µm interval between image planes) were acquired from layer 5 (L5) of the primary somatosensory area (S1). Genetically-marked cell types were identified based on the presence of transcripts encoding the *tdTomato* marker. Commercially available probes were used to detect *Rbm20* (ACD #549251) and *tdTomato* (ACD #317041). A region of interest (ROI) was drawn to define the area of the cell and dots in the ROI were manually counted throughout the image z-stacks. The number of dots in the ROI were then normalized to the cell area (measured in μm^2). Images were assembled using Fiji and Adobe Illustrator Software.

Biochemical procedures

Mouse tissues were extracted on ice and lysed in 50 mM Tris HCl pH 8.0, 150 mM NaCl, 0.1% SDS, 5mM EDTA, 1% Igepal, and protease inhibitor (Roche complete). The lysate was sonicated on ice (100 Hz Amplitude 0.5 cycles x 10 pulses) and centrifuged for 20 min at 13'000 g at 4°C. Proteins in supernatant were analyzed by gel electrophoresis on 4%–20% gradient polyacrylamide gels (BioRad, 4561093) and transferred onto nitrocellulose membrane (BioRad 1704158). Membranes were blocked with 5% non-fat dry milk (NFDM, PanReac AppliChem cat. no. A0830) blocking buffer in TBS-T 1X for 2h at RT and protein detection was by chemoluminescence with HRP-conjugated secondary antibodies (WesternBright Quantum, Advasta, Cat.no. K-12043 D20, K-12042 D20).

RiboTRAP purification (Heiman *et al*, 2008; Sanz *et al*, 2009) was performed with some modifications as described in (Di Bartolomei & Scheiffele, 2022).

Library preparation and deep sequencing

For all the RNA-seq experiments, the quality of RNA integrity was analyzed using an RNA 6000 Pico Chip (Agilent, 5067-1513) on a Bioanalyzer instrument (Agilent Technologies) and only RNA with an integrity number higher than 7 was used for further analysis. RNA concentration was determined by Fluorometry using the QuantiFluor RNA System (Promega #E3310) and 50 ng of RNA was reverse transcribed for analysis of marker enrichment by quantitative PCR (*see* extended methods).

Up to five biological replicates per neuronal population and genotype were analyzed. Library preparation was performed, starting from 50ng total RNA, using the TruSeq Stranded mRNA Library Kit (Cat# 20020595, Illumina, San Diego, CA, USA) and the TruSeq RNA UD Indexes (Cat# 20022371, Illumina, San Diego, CA, USA). 15 cycles of PCR were performed. Libraries were quality-checked on the

Fragment Analyzer (Advanced Analytical, Ames, IA, USA) using the Standard Sensitivity NGS Fragment Analysis Kit (Cat# DNF-473, Advanced Analytical).
The samples were pooled to equal molarity. The pool was quantified by Fluorometry using the QuantiFluor ONE dsDNA System (Cat# E4871, Promega, Madison, WI, USA) before sequencing. Libraries were sequenced Paired-End 151 bases (in addition: 8 bases for index 1 and 8 bases for index 2) setup using the NovaSeq 6000 instrument (Illumina) and the S1 Flow-Cell loaded at a final concentration in Flow-Lane loaded of 340pM and including 1% PhiX.
Primary data analysis was performed with the Illumina RTA version 3.4.4. On average per sample: 49±4 millions pass-filter reads were collected on this Flow-Cell.

RNA Seq data analysis

Initial gene expression and alternative splicing analysis was done in collaboration with GenoSplice Technology, Paris as previously described (Furlanis *et al.*, 2019). In brief, sequencing, data quality, reads repartition (e.g., for potential ribosomal contamination), and insert size estimation were performed using FastQC, Picard-Tools, Samtools and RSeQC tool packages. Reads were mapped using STAR (v2.4.0) (Dobin *et al*, 2013) on the mm10 Mouse genome assembly. The input read count matrix was the same as used for the splicing analysis. Two samples from the olfactory bulb were excluded from the analysis after the quality control of the data.

Read counts were summarized using featureCounts (Liao *et al*, 2014). For each gene present in the FASTDB v2021_4 annotations, reads aligning on constitutive regions (that are not prone to alternative splicing) were counted. Based on these read counts, normalization and differential gene expression were performed using DESeq2 (values were normalized to the total number of mapped reads of all samples) (Love *et al*, 2014) on R (v.3.5.3). Genes were considered as expressed if their FPKM value is greater than 96% the background FPKM value based on intergenic regions. A gene is considered as expressed in the comparison if it is expressed in at least 50% of samples in at least one of the 2 groups compared. Results were considered statistically significant for adjusted p-values ≤ 0.01 (Benjamini Hochberg for p-value adjustment as implemented in DESeq2) and $\log_2$(FC) ≥ 1.5 or or ≤-1.5. For the principal component analysis, counts were normalized using the variance stabilizing transform (VST) as implemented in DESeq2. The internal normalization factors of DESeq2 were used to normalize the counts for generation of heatmaps. The alternative splicing analysis was performed by calculating a splicing index (SI) which is the ratio of read density on the exon of interest and the read density on constitutive exons of the same gene. The $\log_2$ fold change (FC) and p-value (unpaired Student's t-test) was calculated by pairwise comparisons of the respective Splicing Index (SI) values. Results were considered significantly different for p-values ≤ 0.01 and $\log_2$(FC) ≥1 or ≤-1.
For the generation of volcano plots, exons and genes with *NA* or *Inf* values were removed to prevent bias caused by genes or exons with very low expression. Plots were created in R with the *ggplot2* package.

CLIP and library preparation

The CLIP experiments were performed according to the seCLIP protocol (Van Nostrand *et al.*, 2017b) with some minor modifications (Traunmuller *et al.*, 2023). Olfactory bulbs from seven mice were pooled for each biological replicate and for heart tissue one heart was used per biological replicate. Samples were flash frozen and ground on dry ice first in a metal grinder and a porcelain mortar. The frozen powder was transferred into a plastic Petri dish and distributed in a thin layer. The samples were UV-crosslinked 3 times at 400 mJ/cm^2 on dry ice with a UV-crosslinker (Cleaver Scientific). The powder was mixed and redistributed on the Petri dish before each UV exposure. After crosslinking, the powder was collected in 3.5ml (olfactory bulbs) or 5.5 ml (heart) in lysis buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, complete protease inhibitors, Roche) and 4 U per ml Turbo-DNase (Thermofisher). Samples were further processed as described in the extended methods section.

seCLIP data processing was performed as described (Van Nostrand *et al*, 2016; Van Nostrand *et al*, 2017a; Van Nostrand *et al*, 2020). In brief, raw reads were processed to obtain unique CLIP tags mapped to mm10 using Clipper (https://github.com/YeoLab/clipper; https://github.com/YeoLab/eclip). Reads from replicates 1 and 2 from the olfactory bulb were concatenated. Peak normalization was

performed by processing the SMInput samples using the same peak calling pipeline. Irreproducible discovery rate (IDR) analysis was performed to identify reproducible peaks across biological replicates (Li *et al*, 2011). IDR (https://github.com/nboley/idr) was used ranking seCLIP peaks by the fold change over the size-matched input. Clip peaks were called based on IDR < 0.05. We observed some short highly represented sequences that were not specific to RBM20 seCLIP isolations, which were excluded based on peak shape and width (< 30 bp) using StoatyDive (Heyl & Backofen, 2021).

For motif discovery, crosslinking-induced truncation sites (CITS) were called using the CTK pipeline (Shah et al. 2017). Briefly, unique tags from replicates were combined and CITS were called by requiring FDR < 0.001. Sequences from -10bp to +10 bp from called CITS were used as input sequences for DREME software (Bailey *et al*, 2009; Bailey *et al*, 2015; Nystrom & McKay, 2021). As a control, sequences of the same length coming from 500 bp upstream of the (-510 to -490 bases) from the CITS site were used. Enrichment of the UCUU motif at the CITS sites was calculated.

Analysis of RBM20-Bound Intron Length

For investigating the intron length of RBM20-bound introns, we performed a permutation test to calculate an empirical p-value. We generated 5,000 sets of random genomic coordinates, mirroring the length distribution and quantity of RBM20 seCLIP peaks. These coordinates were confined to the intronic regions of genes identified in the Vglut2 RiboTrap datasets.

In each permutation, we computed the mean length of all introns that included these random regions. The resulting distribution, based on 6,000 mean intron lengths, yielded a mean of ~ 41.4 kb nucleotides and a standard deviation of ~288.The average length of introns containing RBM20 seCLIP peaks was ~ 59.0 kb, corresponding to a z-score of 60.95 and yielding a p-value $1.666667 * 10^{-4}$.This p-value is constrained by the number of permutations conducted.

Gene Ontology analysis

All the gene ontology analysis were performed by using a statistical overrepresentation test and the cellular component function in PANTHER (http://pantherdb.org/). All genes being detected as expressed in the Ribo-TRAP RNA-sequencing data were used as reference. GO cellular component annotation data set was used and Fisher's Exact test and Bonferroni correction for multiple testing was applied. GO terms with at least 5 genes and with P-value <0.05 were considered as significantly enriched. Significant GO terms were plotted in Prism 9.

For seCLIP GO analysis, any gene that had significant peak expression in the CLIP dataset either for olfactory bulb or heart samples was used and all genes being detected as expressed in the seCLIP size-matched input samples were used as reference.

Statistical methods and data availability

Sample sizes were determined based on the 3R principle, past experience with the experiments and literature surveys. Pre-established exclusion criteria were defined to ensure success and reliability of the experiments: for stereotaxic injection, all mice with mis-targeted injections were excluded from analysis (e.g. if no eGFP signal was detected in the mitral cell layer of the OB). Investigators performing image analysis and quantification were blinded to the genotype and/or experimental group. For Ribo-TRAP pull-down experiments, all the samples presenting enrichment of the wrong marker genes were excluded. For the quantification of RBM20 expression in the olfactory bulb statistical analysis was performed with Prism 9 (GraphPad software) using unpaired t-test. Data presented are mean ± SD. Images were assembled using Fiji, Omero (Swedlow *et al*, 2003) and Adobe Illustrator software.

A detailed description of the exclusion criteria for different experiments is included in the respective method sections. Statistical analyses were conducted with GraphPad Prism 9. The applied statistical tests were chosen based on sample size, normality of data distribution and number of groups compared. Details on *n*-numbers, p-values and specific tests are found in the figure legends. All raw data files, excel analysis tables and additional data supporting the findings of this study could not be included in the manuscript due to space constraints but are available from the corresponding author upon reasonable request.

Data availability

The datasets produced in this study are available in the following databases: MassIVE (code: MSV000093344), PRIDE (code: PXD046806), and GEO (code: *in submission*).

Acknowledgements

We thank Caroline Bornmann and Sabrina Innocenti for excellent support with lab organization and experiments, Pawel Pelzcar and the Centre for Transgenic Models at the University of Basel for outstanding advice and services, Geoffrey Fucile (SciCORE) for help with data analysis, the Biozentrum Imaging Core Facility for support with image acquisition and analysis, the Quantitative Genomics Centre of the University of Basel for excellent technical assistance. The Scheiffele Laboratory is an associate member of the NCCR RNA & Disease, funded by the Swiss National Science Foundation. This work was financially supported by funds to P.S. from the Canton Basel-Stadt/University of Basel, the Swiss National Science Foundation (project 179432), a Swiss National Science Foundation Advanced Grant (TMAG-3-209273), a European Research Council Advanced Grant (SPLICECODE), and AIMS-2-TRIALS which are supported by the Innovative Medicines Initiatives from the European Commission.

Figure 1

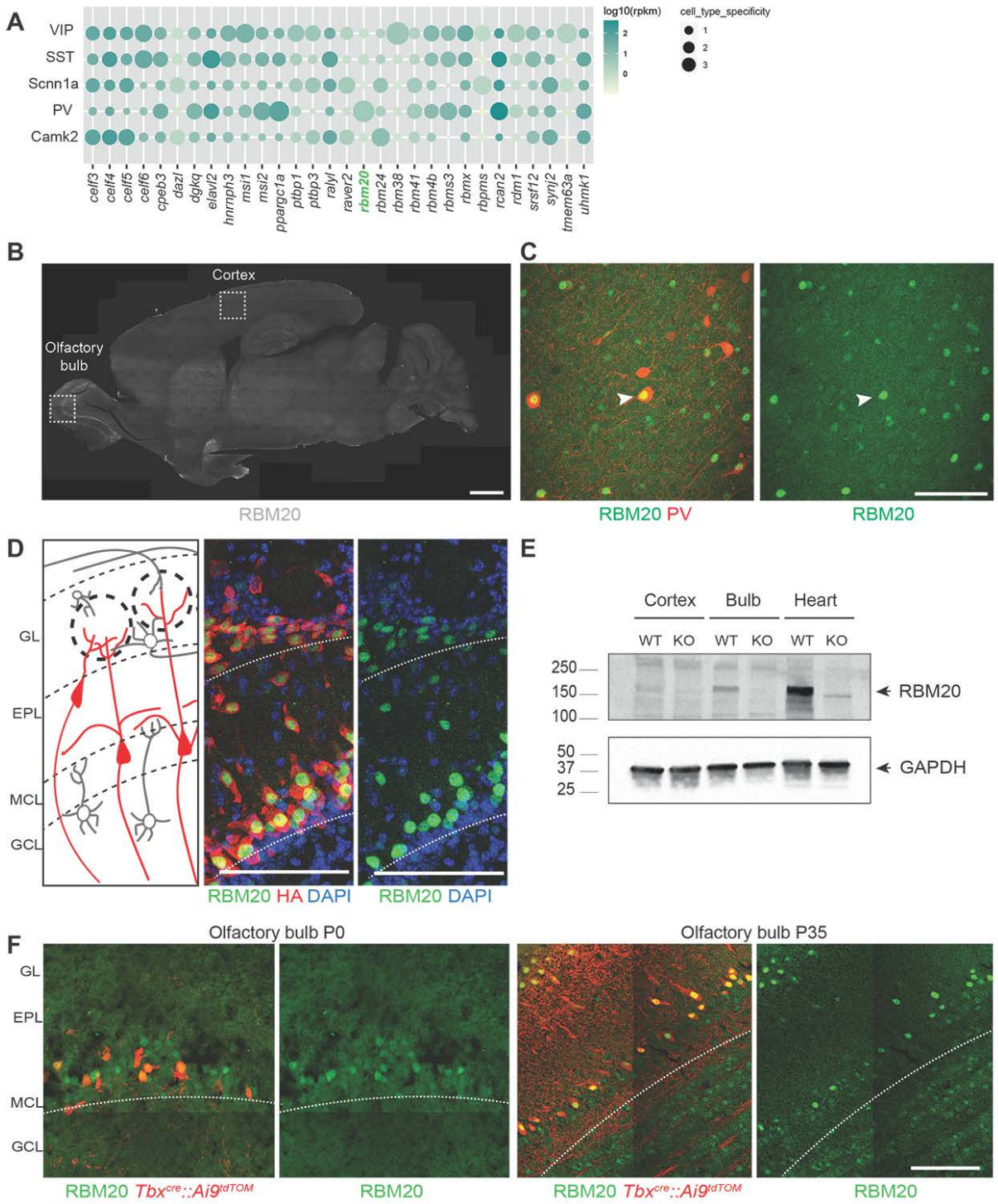

Figure 1. Characterization of *Rbm20* expression in the brain

A. Dot plot of the expression of a hand-curated list of RBPs across different neuronal neocortical populations. RBPs were chosen based on the presence of an RNA recognition motif (RRM) in their sequence and on the ranking of their gini-index value (only the first 20 RBPs displaying the highest gini-index value are displayed (see methods)). RBPs expression was measured by Ribo-TRAP sequencing and expressed as RPKM values normalized over the mean expression across different neuronal populations.

B. Sagittal section of the mouse brain used for immunohistochemistry of RBM20 (grey). The somatosensory cortex and the olfactory bulb regions, where RBM20 is expressed, are highlighted. Scale bar: 1 mm.

C. Immunohistochemistry of RBM20 expression (green) in Parvalbumin positive interneurons (red) in the neocortex.

D. Schematic illustration of the olfactory bulb circuitry and cell types (left). GL: glomerular layer, EPL: external plexiform layer, MCL: mitral cell layer, GCL: granule cell layer, (left). RBM20 expression (green) is specific to the mitral cell layer and glomeruli layer of the olfactory bulb identified in the *Vglut2*^{Cre} *:: Rpl22*^{HA} mouse line (HA staining in red) (middle and right). Scale bar 100 µm.

E. Western blot showing RBM20 expression levels in cortex (CX), olfactory bulb (OB) and heart (HR) samples of WT and constitutive KO mice. RBM20 band at 150-170 kDa is indicated with an arrow. GAPDH detection is used as loading control.

F. Immunofluorescence of RBM20 (green) expression in the mitral cell layer (MCL) of P0 and P35 *Tbx21*^{Cre}*::Ai9*^{tdTomato} mice. Mitral cells and tufted neurons are labelled with the tdTomato reporter (red). Scale bar 100 µm.

Figure 2

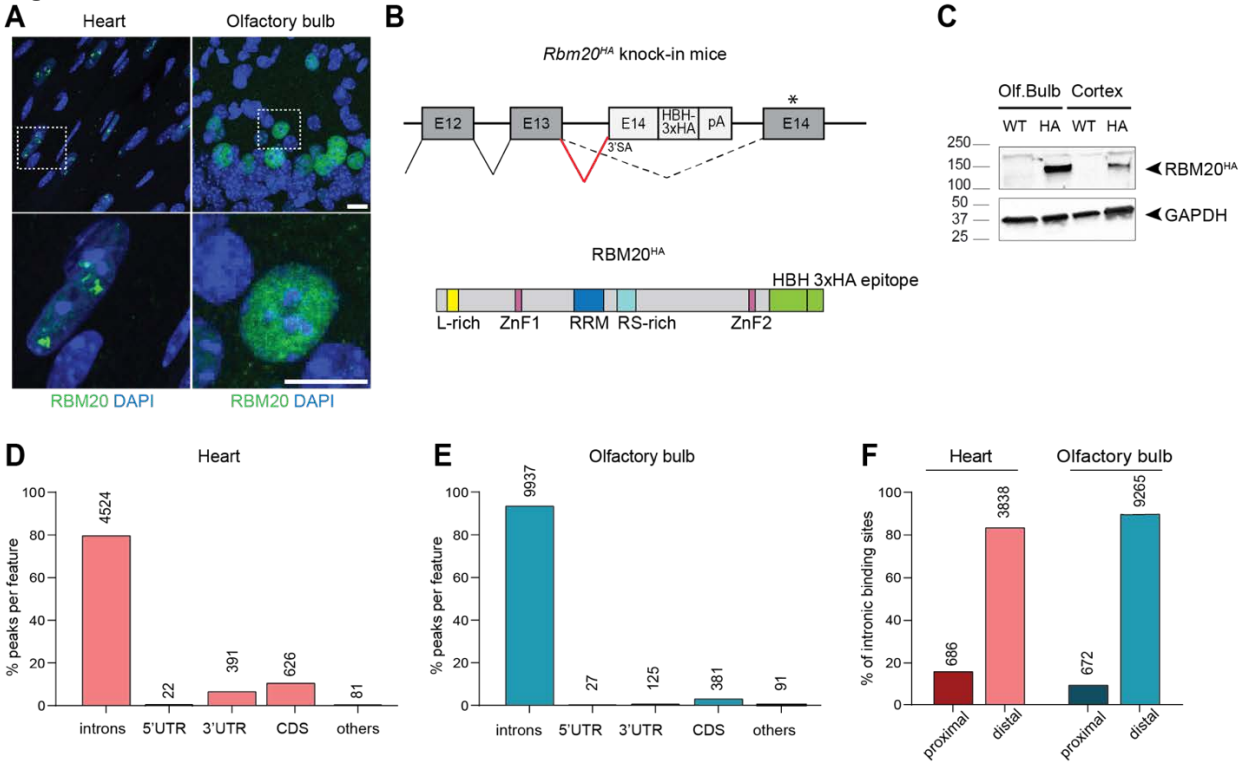

Figure 2. Identification of RBM20 direct targets in the heart and olfactory bulb

A. Sub-nuclear localization of RBM20 (green) in heart cardiomyocytes (left) and mitral cells of the olfactory bulb (right) of WT mice at P35. Scale bar 10 µm.

B. Schematic illustration of HA epitope-tagging of endogenous RBM20 in mice. A cassette was inserted into the *Rbm20* locus containing a strong synthetic 3' splicing acceptor site (3'SA) introduced into sequences derived from *Rbm20* exon 14 and in frame fusion of a histidine-biotin-histidine-3xHA tag, followed by a poly-adenylation signal (up). Schematic representation of the resulting RBM20^{HA} protein where the last exon of the protein is fused to a histidine-biotin-histidine-3xHA tag (down).

C. Western blot showing the validation of RBM20 expression in the olfactory bulb and cortex tissues of WT and *Rbm20^{HA}* tagged mice. GAPDH is used as a loading control.

D, E (D) Quantification of the percentage of peaks identified in the heart (red) and (E) olfactory bulb (blue) tissue in each genomic feature: introns, untranslated regions (3'UTR, 5'UTR), coding sequence (CDS), others (promoters, intergenic regions, non-coding regions). The absolute number of the peaks identified in each genomic feature is reported on top of the corresponding bar.

F. Bar plot showing the percentage of peaks identified in distal (>500 bp) or proximal (< 500 bp) intronic regions in the heart (red) and olfactory bulb (blue). The absolute number of the peaks identified is reported on top of the corresponding bar.

Figure 3

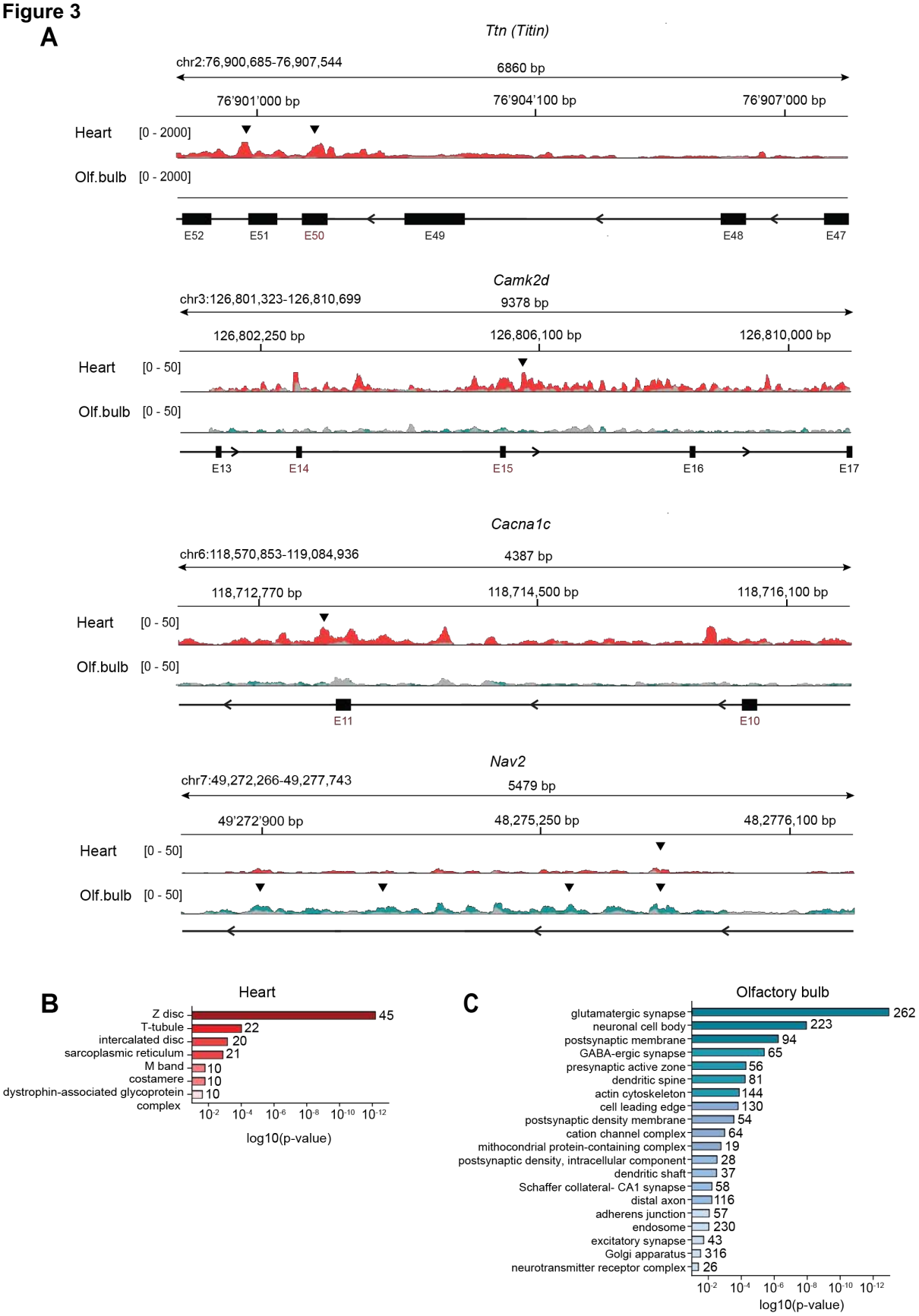

Figure 3. Identification of RBM20 direct mRNA targets in the heart and olfactory bulb

A. Tracks illustrating RBM20 CLIP-seq signal for *Ttn, CamkIId, Cacna1c,* and *Nav2*. Read density obtained for heart samples (red traces) and olfactory bulb (green traces) with the corresponding input samples (overlaid traces in grey). CLIP peaks, considered statistically significant by IDR (score >540, equivalent to IDR<0.05) are marked by black arrowheads. RBM20-dependent

alternative exons previously reported for cardiomyocytes are labeled in red. Note that in the olfactory bulb, RBM20 binding sites are identified on *CamkIId* and *Cacna1c* pre-mRNAs. However, these binding sites are distal (> 500 bp) to the alternative exons. See TableS1 for coordinates.

B. Illustration of the GO categories of RBM20 mRNA targets in the heart (IDR<0.05). Cellular Component analysis with Bonferroni correction (p-value ≤ 0.05). The number of genes found in each category is displayed on top of each bar. Minimal number of genes identified in each category: 5 genes.

C. Illustration of the GO categories of RBM20 mRNA targets in the olfactory bulb (IDR<0.05). Cellular Component analysis with Bonferroni correction (p-value ≤ 0.05). The number of genes found in each category is displayed on top of each bar. Minimal number of genes identified in each category: 5 genes.

Figure 4

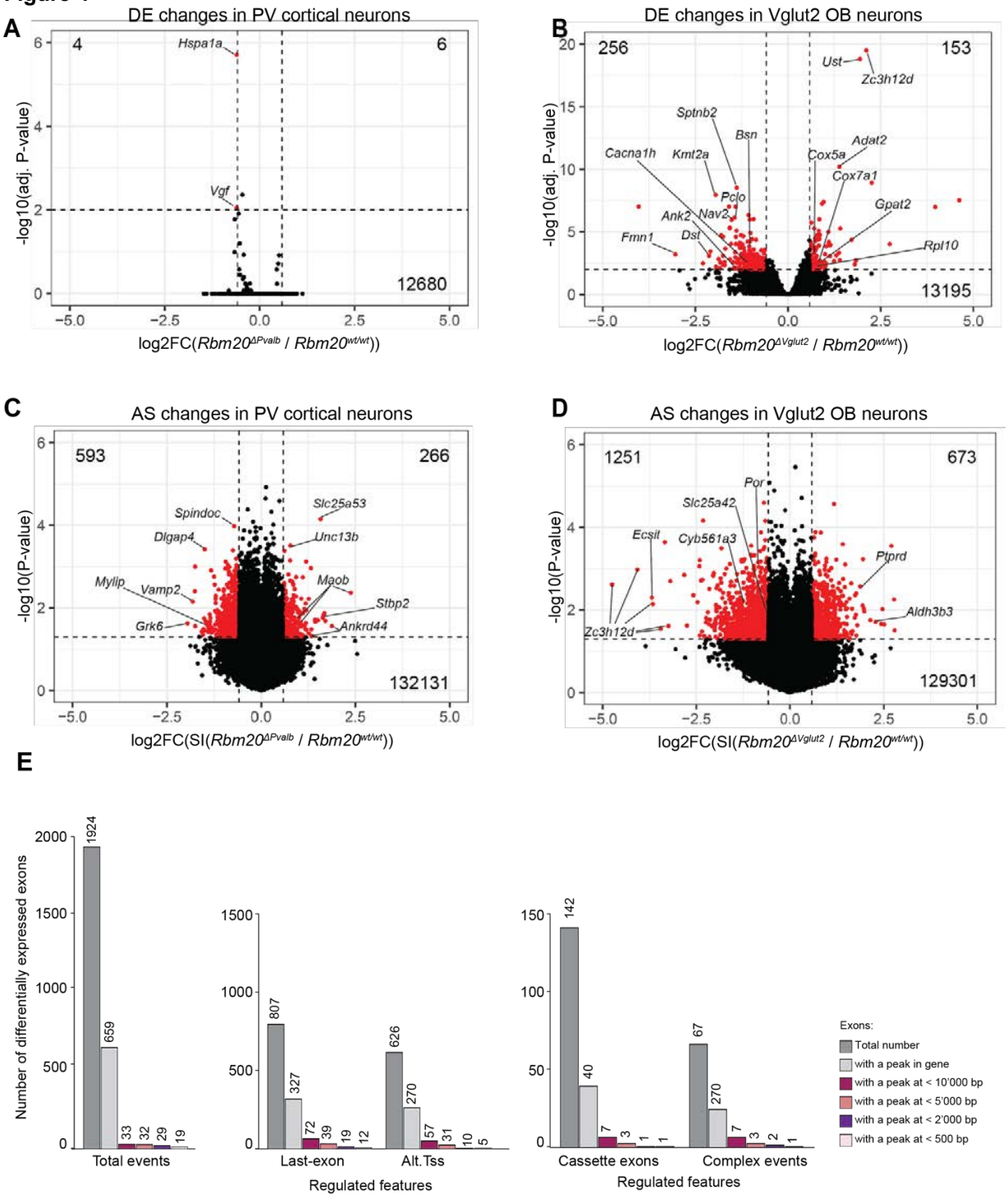

Figure 4. Differential gene expression and alternative exon incorporation rates in *Rbm20* conditional knock-out cells

A. Volcano plot of differential gene expression in RiboTrap-isolated mRNAs from *Rbm20*^{WT} *vs.* *Rbm20*$^{\Delta Vglut2}$ olfactory bulb (P35). Significantly regulated genes shown in red, cut-off fold-change (FC) of 1.5, adjusted p-value <0.01, total number of up- and down-regulated noted on top. Note that *Rbm20* itself is strongly reduced, outside the axis limits, and not represented in this plot (see Table S4).

B. Volcano plot of the differential gene expression in RiboTrap-isolated mRNAs from *Rbm20*^{WT} *vs.* *Rbm20*$^{\Delta PV}$ mouse neocortex (P35) as in panel A. The Y-chromosomal genes *Ddx3y*, *Uty, Kdm5d, Eif2s3y* were highly differentially expressed due to the larger number of *Rbm20* mutant males used in the Ribotrap isolations (3 wild-type females and 1 wild-type male vs. 4 knock-out male mice were used for this experiment). These genes and *Rbm20* itself were excluded from the plot (see Table S4 for complete data).

C. Volcano plot representing differentially-included exons in *Rbm20*$^{\Delta Vglut2}$ RiboTrap-isolated mRNAs from olfactory bulb neurons. The dotted lines correspond to FC values of 1.5 and -1.5 and -log10 (p-value) of 1.3. Significantly regulated exons (FC 1.5 and $p<0.05$) are shown in red.

D. Volcano plot representing differentially-included exons in *Rbm20*$^{\Delta PV}$ RiboTrap-isolated mRNAs from cortical interneurons. The dotted lines correspond to FC values of 1.5 and -1.5 and -log10 (p-value) of 1.3. Significantly regulated exons (FC 1.5 and p-value <0.05) are shown in red.

E. Number of exons differentially expressed in *Vglut2*$^{+}$ cells isolated from the olfactory bulb of *Rbm20*$^{\Delta Vglut2}$ mice and number of exons with significant RBM20 CLIP peaks in binding sites within indicated distances. Equivalent information is provided for exons divided by annotations for specific features of regulation: alternative poly-adenylation (last exons), alternative transcription start sites (TSS), complex events and cassette exons.

Figure 5

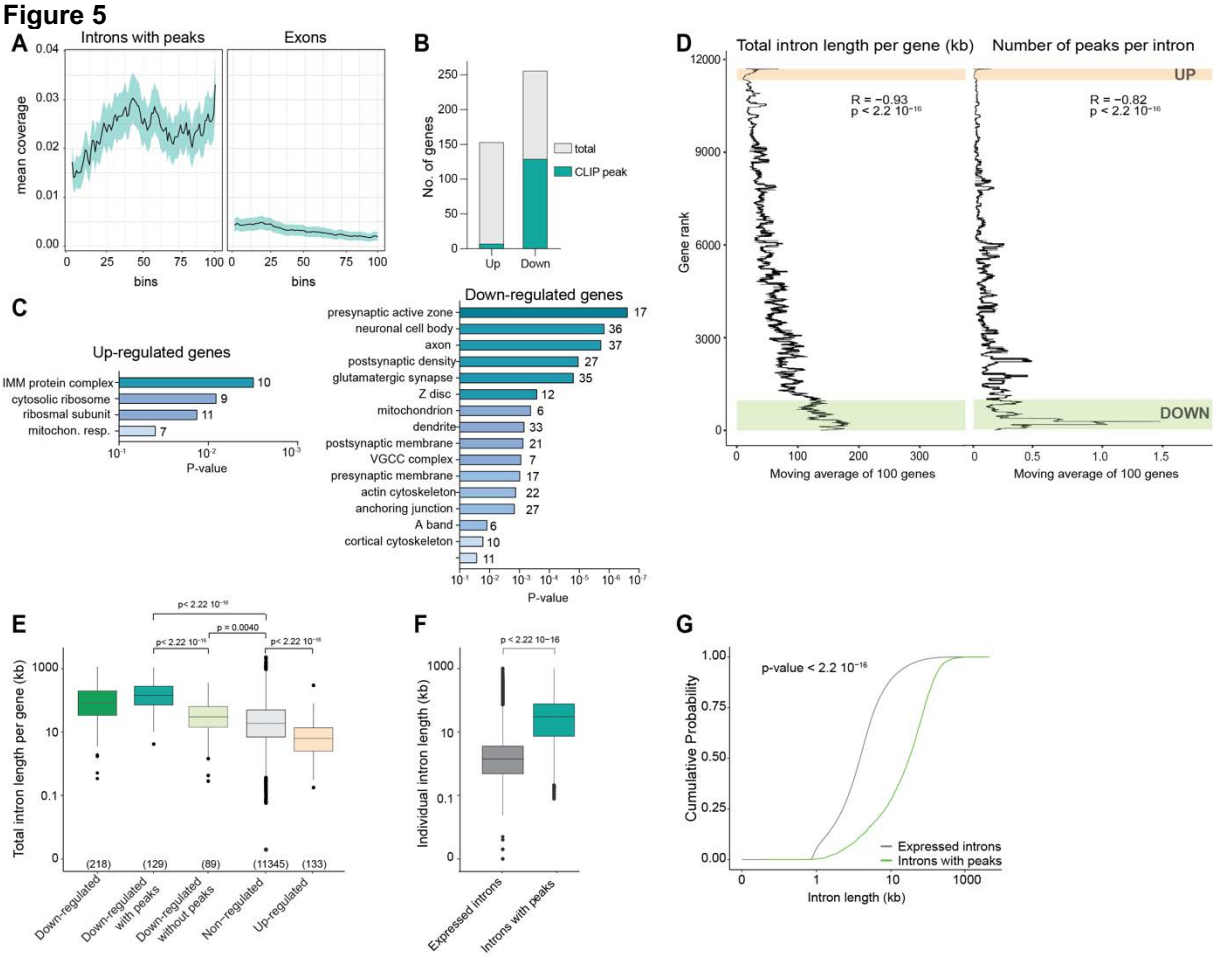

Figure 5. Long pre-mRNAs are depleted in *Rbm20*$^{\Delta Vglut2}$ mitral cells

A. Metagene coverage plots of CLIP peaks across all RBM20-bound introns. Peak density across exons from the same transcripts are shown for comparison.

B. Total number of genes up- or down-regulated (FC > 1.5 and adjusted p-value <0.01) in glutamatergic cells from the olfactory bulb of *Rbm20*$^{\Delta Vglut2}$ mice. The fraction of genes with significant RBM20 CLIP peaks is indicated in blue.

C. Illustration of gene ontologies enriched amongst up- and down-regulated genes (cellular component analysis with Bonferroni correction (p-value 0.05). Minimum number of 5 genes identified per category.

D. Correlation of differential gene expression in *Rbm20*$^{\Delta Vglut2}$ cells, intron length and CLIP-seq data. Genes were ranked by FC in differential gene expression and mean total intron length (left) and mean number of intronic CLIP peaks (right) for blocks of 100 genes were plotted. Ranks of the genes meeting FC cut-off for down- or up-regulation are highlighted in color. Spearman's coefficients and p-values are indicated.

E. Boxplot showing the total intron length per gene (expressed in $\log_{10}$ scale) for categories of downregulated genes (all, with or without RBM20 binding sites), non-regulated genes, and up-regulated genes in *Rbm20*$^{\Delta Vglut2}$ in RiboTrap isolates from the olfactory bulb. Only annotated genes are plotted (number of genes shown at the bottom). P-values from Wilcoxon test are indicated. Medians: all down-regulated genes 83.2 kb, down-regulated genes with peaks: 152.4 kb; down-regulated genes without peaks: ~33.6 kb; all non-regulated genes: 21.5 kb; all up-regulated genes: 7.4 kb.

F. Boxplot ($\log_{10}$ scale) illustrating the total length of introns found in genes identified in our RBM20 Ribo-TRAP dataset (grey) compared to introns presenting RBM20 binding sites (green). RBM20 bound introns exhibit a higher intron length. P-values from Wilcoxon test are indicated. Medians: expressed introns: 1.4 kb; introns with peaks: 30.5 kb.

G. Plot representing the cumulative probability distribution of intron length between the two groups of introns as in panel F. P-value (Kolmogorov-Smirnov test) is indicated.

References

Attig J, Agostini F, Gooding C, Chakrabarti AM, Singh A, Haberman N, Zagalak JA, Emmett W, Smith CWJ, Luscombe NM *et al* (2018) Heteromeric RNP Assembly at LINEs Controls Lineage-Specific RNA Processing. *Cell* 174: 1067-1081 e1017

Babitzke P, Baker CS, Romeo T (2009) Regulation of translation initiation by RNA binding proteins. *Annu Rev Microbiol* 63: 27-44

Badshah N, Mattison KA, Ahmad S, Chopra P, Johnston HR, Ahmad S, Khan SH, Sarwar MT, Cutler DJ, Taylor M *et al* (2022) Novel Missense CNTNAP2 Variant Identified in Two Consanguineous Pakistani Families With Developmental Delay, Epilepsy, Intellectual Disability, and Aggressive Behavior. *Front Neurol* 13: 918022

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37: W202-208

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. *Nucleic Acids Res* 43: W39-49

Bertero A, Fields PA, Ramani V, Bonora G, Yardimci GG, Reinecke H, Pabon L, Noble WS, Shendure J, Murry CE (2019) Dynamics of genome reorganization during human cardiogenesis reveal an RBM20-dependent splicing factory. *Nat Commun* 10: 1538

Brauch KM, Karst ML, Herron KJ, de Andrade M, Pellikka PA, Rodeheffer RJ, Michels VV, Olson TM (2009) Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J Am Coll Cardiol* 54: 930-941

Briganti F, Sun H, Wei W, Wu J, Zhu C, Liss M, Karakikes I, Rego S, Cipriano A, Snyder M *et al* (2020) iPSC Modeling of RBM20-Deficient DCM Identifies Upregulation of RBM20 as a Therapeutic Strategy. *Cell reports* 32: 108117

Dauksaite V, Gotthardt M (2018) Molecular basis of titin exon exclusion by RBM20 and the novel titin splice regulator PTB4. *Nucleic acids research* 46: 5227-5238

Di Bartolomei G, Scheiffele P (2022) An Optimized Protocol for the Mapping of Cell Type-Specific Ribosome-Associated Transcript Isoforms from Small Mouse Brain Regions. *Methods Mol Biol* 2537: 37-49

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15-21

Fenix AM, Miyaoka Y, Bertero A, Blue SM, Spindler MJ, Tan KKB, Perez-Bermejo JA, Chan AH, Mayerl SJ, Nguyen TD *et al* (2021) Gain-of-function cardiomyopathic mutations in RBM20 rewire splicing regulation and re-distribute ribonucleoprotein granules within processing bodies. *Nat Commun* 12: 6324

Furlanis E, Traunmuller L, Fucile G, Scheiffele P (2019) Landscape of ribosome-engaged transcript isoforms reveals extensive neuronal-cell-class-specific alternative splicing programs. *Nature neuroscience* 22: 1709-1717

Gebauer F, Schwarzl T, Valcarcel J, Hentze MW (2021) RNA-binding proteins in human genetic disease. *Nature reviews Genetics* 22: 185-198

Gomez AM, Traunmuller L, Scheiffele P (2021) Neurexins: molecular codes for shaping neuronal synapses. *Nat Rev Neurosci* 22: 137-151

Gu S, Katyal A, Zhang Q, Chung W, Franciosi S, Sanatani S (2023) The Association Between Congenital Heart Disease and Autism Spectrum Disorder: A Systematic Review and Meta-Analysis. *Pediatr Cardiol* 44: 1092-1107

Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, Maatz H, Schulz H, Li S, Parrish AM *et al* (2012) RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nature medicine* 18: 766-773

Heiman M, Kulicke R, Fenster RJ, Greengard P, Heintz N (2014) Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). *Nature protocols* 9: 1282-1291

Heiman M, Schaefer A, Gong S, Peterson JD, Day M, Ramsey KE, Suarez-Farinas M, Schwarz C, Stephan DA, Surmeier DJ *et al* (2008) A translational profiling approach for the molecular characterization of CNS cell types. *Cell* 135: 738-748

Heyl F, Backofen R (2021) StoatyDive: Evaluation and classification of peak profiles for sequencing data. *Gigascience* 10

Holt CE, Martin KC, Schuman EM (2019) Local translation in neurons: visualization and function. *Nat Struct Mol Biol* 26: 557-566

Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, DePalma SR, McKean D, Wakimoto H, Gorham J *et al* (2015) De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science (New York, NY* 350: 1262-1266

Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, Zeng X, Qi H, Chang W, Sierant MC *et al* (2017) Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nat Genet* 49: 1593-1601

Khan MA, Reckman YJ, Aufiero S, van den Hoogenhof MM, van der Made I, Beqqali A, Koolbergen DR, Rasmussen TB, van der Velden J, Creemers EE *et al* (2016) RBM20 Regulates Circular RNA Production From the Titin Gene. *Circ Res* 119: 996-1003

Li Q, Brown JB, Huang H, Bickel PJ (2011) Measuring reproducibility of high-throughput experiments. *The Annals of Applied Statistics* 5: 1752-1779, 1728

Li S, Guo W, Dewey CN, Greaser ML (2013) Rbm20 regulates titin alternative splicing as a splicing repressor. *Nucleic acids research* 41: 2659-2672

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

Ling JP, Pletnikova O, Troncoso JC, Wong PC (2015) TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science (New York, NY* 349: 650-655

Ling SC, Polymenidou M, Cleveland DW (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* 79: 416-438

Lopez Soto EJ, Gandal MJ, Gonatopoulos-Pournatzis T, Heller EA, Luo D, Zheng S (2019) Mechanisms of Neuronal Alternative Splicing and Strategies for Therapeutic Interventions. *J Neurosci* 39: 8193-8199

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15: 550

Maatz H, Jens M, Liss M, Schafer S, Heinig M, Kirchner M, Adami E, Rintisch C, Dauksaite V, Radke MH *et al* (2014) RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. *The Journal of clinical investigation* 124: 3419-3430

Marino BS, Lipkin PH, Newburger JW, Peacock G, Gerdes M, Gaynor JW, Mussatto KA, Uzark K, Goldberg CS, Johnson WH, Jr. *et al* (2012) Neurodevelopmental outcomes in children with congenital heart disease: evaluation and management: a scientific statement from the American Heart Association. *Circulation* 126: 1143-1172

Mauger O, Scheiffele P (2017) Beyond proteome diversity: alternative splicing as a regulator of neuronal transcript dynamics. *Curr Opin Neurobiol* 45: 162-168

Nystrom SL, McKay DJ (2021) Memes: A motif analysis environment in R using tools from the MEME Suite. *PLoS Comput Biol* 17: e1008991

Parikh VN, Caleshu C, Reuter C, Lazzeroni LC, Ingles J, Garcia J, McCaleb K, Adesiyun T, Sedaghat-Hamedani F, Kumar S *et al* (2019) Regional Variation in RBM20 Causes a Highly Penetrant Arrhythmogenic Cardiomyopathy. *Circ Heart Fail* 12: e005371

Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, Ling SC, Sun E, Wancewicz E, Mazur C *et al* (2011) Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nature neuroscience* 14: 459-468

Rosenthal SB, Willsey HR, Xu Y, Mei Y, Dea J, Wang S, Curtis C, Sempou E, Khokha MK, Chi NC *et al* (2021) A convergent molecular network underlying autism and congenital heart disease. *Cell Syst* 12: 1094-1107 e1096

Sanz E, Bean JC, Carey DP, Quintana A, McKnight GS (2019) RiboTag: Ribosomal Tagging Strategy to Analyze Cell-Type-Specific mRNA Expression In Vivo. *Current protocols in neuroscience / editorial board, Jacqueline N Crawley [et al]* 88: e77

Sanz E, Yang L, Su T, Morris DR, McKnight GS, Amieux PS (2009) Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci U S A* 106: 13939-13944

Schieweck R, Ninkovic J, Kiebler MA (2021) RNA-binding proteins balance brain function in health and disease. *Physiological reviews* 101: 1309-1370

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B *et al* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9: 676-682

Schneider JW, Oommen S, Qureshi MY, Goetsch SC, Pease DR, Sundsbak RS, Guo W, Sun M, Sun H, Kuroyanagi H *et al* (2020) Dysregulated ribonucleoprotein granules promote cardiomyopathy in RBM20 gene-edited pigs. *Nature medicine* 26: 1788-1800

Sibley CR, Emmett W, Blazquez L, Faro A, Haberman N, Briese M, Trabzuni D, Ryten M, Weale ME, Hardy J *et al* (2015) Recursive splicing in long vertebrate genes. *Nature*

Sjostedt E, Zhong W, Fagerberg L, Karlsson M, Mitsios N, Adori C, Oksvold P, Edfors F, Limiszewska A, Hikmet F *et al* (2020) An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science (New York, NY* 367

Swedlow JR, Goldberg I, Brauner E, Sorger PK (2003) Informatics and quantitative analysis in biological imaging. *Science* 300: 100-102

Tainaka K, Murakami TC, Susaki EA, Shimizu C, Saito R, Takahashi K, Hayashi-Takagi A, Sekiya H, Arima Y, Nojima S *et al* (2018) Chemical Landscape for Tissue Clearing Based on Hydrophilic Reagents. *Cell reports* 24: 2196-2210 e2199

Tervo DG, Hwang BY, Viswanathan S, Gaj T, Lavzin M, Ritola KD, Lindo S, Michael S, Kuleshova E, Ojala D *et al* (2016) A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. *Neuron* 92: 372-382

Traunmuller L, Schulz J, Ortiz R, Feng H, Furlanis E, Gomez AM, Schreiner D, Bischofberger J, Zhang C, Scheiffele P (2023) A cell-type-specific alternative splicing regulator shapes synapse properties in a trans-synaptic manner. *Cell reports* 42: 112173

Ule J, Blencowe BJ (2019) Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Mol Cell* 76: 329-345

Ule J, Darnell RB (2006) RNA binding proteins and the regulation of neuronal synaptic plasticity. *Curr Opin Neurobiol* 16: 102-110

Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB (2003) CLIP identifies Nova-regulated RNA networks in the brain. *Science (New York, NY* 302: 1212-1215

Upadhyay SK, Mackereth CD (2020) Structural basis of UCUU RNA motif recognition by splicing factor RBM20. *Nucleic acids research* 48: 4538-4550

van den Hoogenhof MMG, Beqqali A, Amin AS, van der Made I, Aufiero S, Khan MAF, Schumacher CA, Jansweijer JA, van Spaendonck-Zwarts KY, Remme CA *et al* (2018) RBM20 Mutations Induce an Arrhythmogenic Dilated Cardiomyopathy Related to Disturbed Calcium Handling. *Circulation* 138: 1330-1342

Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen JY, Cody NAL, Dominguez D *et al* (2020) A large-scale binding and functional map of human RNA-binding proteins. *Nature* 583: 711-719

Van Nostrand EL, Nguyen TB, Gelboin-Burkhart C, Wang R, Blue SM, Pratt GA, Louie AL, Yeo GW (2017a) Erratum to: Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP). *Methods Mol Biol* 1648: E1

Van Nostrand EL, Nguyen TB, Gelboin-Burkhart C, Wang R, Blue SM, Pratt GA, Louie AL, Yeo GW (2017b) Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP). *Methods Mol Biol* 1648: 177-200

Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K *et al* (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* 13: 508-514

Vuong CK, Black DL, Zheng S (2016) The neurogenetics of alternative splicing. *Nature reviews Neuroscience* 17: 265-281

Watanabe T, Kimura A, Kuroyanagi H (2018) Alternative Splicing Regulator RBM20 and Cardiomyopathy. *Front Mol Biosci* 5: 105

Zhu C, Wu J, Sun H, Briganti F, Meder B, Wei W, Steinmetz LM (2021) Single-molecule, full-length transcript isoform sequencing reveals disease-associated RNA isoforms in cardiomyocytes. *Nat Commun* 12: 4203

Zylka MJ, Simon JM, Philpot BD (2015) Gene length matters in neurons. *Neuron* 86: 353-355

Supplementary Information

Dilated cardiomyopathy-associated RNA Binding Motif Protein 20 regulates long pre-mRNAs in neurons

Giulia Di Bartolomei, Raul Ortiz, Dietmar Schreiner, Susanne Falkner, Esther E. Creemers, Peter Scheiffele

Table of Content

- Extended Materials and Methods
- Figure S1, (related to Fig. 1) and figure legend: Characterization of Rbm20 expression in the brain
- Figure S2, (related to Fig. 2 and 3) and figure legend: Generation of RBM20 – HA tagged mouse line
- Figure S3, (related to Fig. 2 and 3) and figure legend: Identification of RBM20 binding sites on transcript mRNAs
- Figure S4, (related to Fig. 4) and figure legend: Normal morphological differentiation of mitral cells in the absence of RBM20.
- Figure S5, (related to Fig. 4) and figure legend: Quality control analysis of Ribo-TRAP RNA-sequencing samples
- Table S1 and table legend (related to Fig. 2 and 3): Identified RBM20 binding sites in the heart and olfactory bulb tissues
- Table S2 and table legend (related to Fig. 2 and 3): Gene Ontology analysis tables of transcripts directly bound by RBM20
- Table S3 and table legend (related to Fig. 4): Number of expressed genes and percentage of mapped and unmapped unique reads in Ribo-TRAP RNA-sequencing experiments
- TableS4 and table legend (related to Fig. 4 and 5): Summary tables resulting from differential gene expression analysis
- TableS5 and table legend (related to Fig. 4 and 5): Summary tables resulting from the alternative exon usage analysis
- Table S6 and table legend (related to Fig. 4 and 5): Gene Ontology analysis tables of de-regulated transcripts and alternatively spliced exons
- Table S7 and table legend (related to Fig. 5): Intron length analysis table for de-regulated and non-regulated transcripts

Extended Material and Methods

Mice

All procedures involving animals were approved by and performed in accordance with the guidelines of the Kantonales Veterinärat Basel-Stadt. Male and female mice were used in this study. All other mice were in C57BL/6J background. *Rpl22*^{HA} (RiboTag) mice (Sanz *et al*, 2009), *CamK2*^{cre} (Tsien *et al*, 1996), *SST*^{cre} (Taniguchi *et al*, 2011), *Pvalb*^{cre} mice (Hippenmeyer *et al*, 2005), *Ai9*^{tdTOM} reporter mice (Madisen *et al*, 2010), and *Vglut2*^{cre} mice (Vong *et al*, 2011) were obtained from Jackson Laboratories (Jax stock no: 011029, 017320, and 007909, 028863, 007914 tdtomato and 016963). *Rbm20*^{floxed} mice and *Rbm20* constitutive KO mice (Khan *et al*, 2016) were backcrossed for at least 4 generations to a C57BL/6J background. *Tbx21*^{cre} mice (Haddad *et al*, 2013), were kindly provided by Dr. R. Datta's Laboratory.

The *Rbm20*-COIN allele knock-in mouse line was generated in collaboration with the Center for Transgenic Mouse (CTM) lines in Basel. The line was generated using the Crispr-Cas9 system. gRNAs targeting the last intron of *Rbm20* and template "*coin allele*" construct were injected together with RNA encoding for Crispr-Cas9 nuclease into C57BL/6J zygotes. The surviving embryos were transferred into recipient females. The *coin* module fused to a histidine-biotin-histidine-3HA tag. The *coin allele* is inserted in an orientation opposite to the gene's direction of transcription. The gRNA used were: 5' TTGAGTCGGGGGTCCCACTG 3'. The 1'311 bp single stranded megamer containing the upstream homology sequence (lower case), *coin module* (upper case letters) and downstream homology sequence (lower case) and the optimized codon usage (lower case) used:
5'ggcgaggctgctgctggagagccctgatttcttctctgtttgactcgcgaattctgaggggataagcgccctgcatatgtatgcattcttctttggg
agcctgcagccaccttcatgcccagtaaggctatgcttactgtgccagatcacccccctgtaggctcacatagagccatgaccagcaacagcat
agcgggatttccagaggcttcactgaggcagctatgacctgctcttgcctcccagggcatCCCCAGTACCGTTCGTATAatgtatgc
TATACGAAGTTATGGGCCCCTCTGCTAACCATGTTCATGCCTTCTTCTTTTTCCTACAGAAGTACC
TGTCTCAGCTGGCAGAGGAGGgactcAAGGAGACGGAGGGGACAGACAGCCCAAGCCCCGAGCG
TGGTGGGATTGGTCCACACTTGGAAAGGAAGAAGCTAGCtGGcCAcCATCACCACCAcCATGGTG
CcGCTGGAAAGGCCGGTGAAGGTGAAATCCCTGCCCCTCTTGCTGGTACaGTTTCTAAGATACTc
GTAAAAGAAGGTGACACTGTTAAAGCTGGTCAAACAGTTCTGGTGCTGGAGGCcATGAAAATGGA
GACAGAAATTAACGCTCCTACTGACGGAAAAGTTGAAAAGGTGTTAGTTAAGGAAAGAGATGCTG
TTCAAGGTGGTCAAGGTCTAATCAAGATCGGCGTTGCAGGTCATCAcCACCAtCATCAcGGcGCcgc
cgggTATCCCTACGATGTGCCTGACTATGCTgctggcTATCCTTACGACGTGCCCGATTATGCAgccggc
TATCCATACGATGTCCCAGATTACGCTgccTAGGATCTTTTTCCCTCTGCCAAAAATTATGGGGACA
TCATGAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCATTGCAATAGTGT
GTTGGAATTTTTTGTGTCTCTCACTCGGAAGGACATATGGGAGGGCAAATCATTTAAAACATCAGA
ATGAGTATACCGTTCGTATAgcatacatTATACGAAGTTATTGGGACCCCCGACTCAAggtctcctgatgaat
gctaactttctaagttgcctgacttgagtcagctggcacctgccctgtgggtcagacttcttcacttttcacacttgtggtttggagtaaagtgggaga
ggctgtagagactgaggcattcattctgccaaggcccctgacagaaacgctacctgagatggctgtggcagaggctcctggctccctgataaa
aggtgtaccagggaaacgtgagctgaggtgggagggagtgagg'. The entire insert sequence is highlighted in capital letters. Activation by Cre-recombinase inverts the *coin* module, resulting in alternative splicing of the tagged exon. All lines were maintained on a C57Bl6/J background. Both males and females were used for all the experiments unless stated otherwise in the respective method sections.

Antibodies

The following commercially available antibodies were used: rat-anti-HA (Roche, #11867431001) and rabbit-anti-GAPDH (Cell Signaling #5174), rabbit-anti-MAP2 (Synaptic Systems #188002), rabbit-anti-RFP (Rockland #600-401-379), chicken-anti-GFP (Aves Labs Inc. #GFP-1020), goat anti-Parvalbumin (Swant, PVG213), rabbit-anti-NeuN (Novus Biologicals cat. n. NBP1-77686SS). Secondary antibodies coupled to horseradish peroxidase (HRP) or fluorescent dyes were from Jackson ImmunoResearch (goat anti-rabbit HRP #111-035-003; goat anti-rat HRP #112-035-143), donkey anti-rat IgG-Cy3 and Cy5 (Jackson ImmunoResearch, 712-165-153, 706-175-148) donkey anti-goat IgG-Cy3 and donkey anti-chicken IgG-Cy3 (Jackson ImmunoResearch, 705-165-147, 703-165-155).

For the generation of polyclonal anti-RBM20 antibodies, a synthetic peptide consisting of the RBM20 C-terminus was used: C+PERGGIGPHLERKKL (n- to c-terminus, C+ indicates a cysteine added to the n-terminus for thiol-mediated coupling). The synthetic peptide was conjugated to keyhole limpet hemocyanin for immunization of rabbits and guinea pigs (Eurogentec, Belgium). Resulting sera were affinity-purified on the peptide antigen and the specificity of the resulting antibodies was confirmed using lysates and tissue sections from *Rbm20* knock-out mice.

Immunohistochemistry

For quantifications of RBM20 positive neurons in the olfactory bulb, tile-scan images from 30 µm slices from the olfactory bulb of P35 mice were acquired. Mean intensity analyses for RBM20 signal were performed in Fiji (Schindelin *et al*, 2012) using a custom-made Python script, as previously described (DOI-https://github.com/imcf-shareables/3D_spots_count/blob/main/README.md). In brief, neuronal

cells were identified based on the nuclear DAPI signal. The mean intensity of RBM20 protein in each nucleus was then measured and the background signal was subtracted.

For the characterization of RBM20 sub-nuclear localization, brain and heart samples from *Rbm20* WT and *Rbm20* cKO mice (P35-P40) were anesthetized with ketamine/xylazine (100/10 mg/kg *i.p.*) and transcardially perfused with fixative (4% paraformaldehyde). The brains and hearts were post-fixed overnight in the same fixative at 4°C and washed 3 times with 100 mM phosphate buffer (PB). Coronal brain slices were cut at 40 µm with a vibratome (Leica Microsystems VT1000). Brain samples were immersed in 15% and subsequently 30% sucrose in 1X PBS for 48 h, cryoprotected with Tissue-Tek optimum cutting temperature (OCT) and frozen at –80° until use. Tissue was sectioned at 40 µm on a cryostat (Microm HM560, Thermo Scientific) and collected in 1X PBS. Immunohistochemistry and imaging were performed as previously described.

Targeted LC-MS sample preparation and analysis

Murine nuclear extracts of heart tissue were lysed in 100 mM Triethylamonium bicarbonate pH 8.5 / 5% SDS / 10 mM Tris (2-carboxyethyl) phosphin using 20 cycles of sonication (30 s on / 30 s off per cycle) on a Bioruptor system (Dianode) followed by heating to 95° C for 10 min. Protein extracts were alkylated using 15 mM iodoacetamide at 25°C in the dark for 30 min. For each sample, 50 µg of protein lysate was captured, digested, and desalted using STRAP cartridges (Protifi, NY, US) following the manufacturer's instructions. Samples were dried under vacuum and stored at –80°C until further use.

For parallel reaction-monitoring (PRM) assays (Peterson *et al*, 2012) three proteotypic peptides derived from RBM20 were selected for assay development (ASPPTESDLQSQACR, QGFGCSCR and SGSPGPLHSVSGYK). A mixture containing 100 fmol of each heavy reference peptide (JPT, Berlin, Germany) including iRT peptides (Biognosys, Schlieren, Switzerland) was used. The setup of the µRPLC-MS system was as described previously (Ahrne *et al*, 2016). Peptides were analyzed per LC-MS/MS run using a linear gradient ranging from 95% solvent A (0.15% formic acid, 2% acetonitrile) and 5% solvent B (98% acetonitrile, 2% water, 0.15% formic acid) to 45% solvent B over 60 min at a flow rate of 200 nl/min. Mass spectrometry analysis was performed on a Q-Exactive HF mass spectrometer equipped with a nanoelectrospray ion source (Thermo Fisher Scientific) as described previously (Hauser *et al*, 2022). The acquired raw-files were database searched against a *Mus musculus* database (Uniprot, download date: 2020/03/21, total of 44'786 entries) by the MaxQuant software (Version 1.0.13.13). To control for variation in sample amounts, the total ion chromatogram (only comprising peptide ions with two or more charges) of each sample was determined by Progenesis QI (version 2.0, Waters) and used for normalization. The datasets of this study are deposited on MassIVE (code: MSV000093344) and PRIDE (code: PXD046806).

In situ hybridization

For quantification of *Vglut2, Tbr2* and *Rbm20* transcripts expression in mitral and tufted neurons of the olfactory bulb, P25 animals were euthanized and the brains were harvested and processed as described above. Stacks of 10-15 µm width (0.44 µm interval between stacks) were acquired from olfactory bulb slices at room temperature with an upright LSM700 confocal microscope (Zeiss) using 40X Apochromat objectives. A ROI was drawn to define the area of each cell residing either in the mitral cell layer or glomeruli layer of the olfactory bulb. Dots in the ROIs were detected automatically throughout the z-stacks for each channel, using a custom-made Python script, as described in (DOI-https://github.com/imcf-shareables/3D_spots_count/blob/main/README.md). The following commercial probes were used: *Rbm20* (549251), *slc17a6* (319171), *Tbr2* (Eomes): (429641). Images from 3 mice were used for the quantification (2 images per slice). *Gad2* (415071) *in situ* hybridization was performed on 15 µm olfactory bulb slices of P25 mice.

Two photon image analysis

A total of 10 neurons (5 neurons per genotype from at least 3 biological replicates) were analyzed. Both apical and lateral dendrites of mitral cells were traced semi-automatically by using the user-guided 3D image detection algorithm. Tracings were checked and corrected manually when needed. Subcellular components, such as spines and other small protrusions were not traced. The following parameters were extracted from the software Neurolucida Explorer®. for each traced neuron: the number of dendrites from different centrifugal orders (*i.e.* primary dendrites, secondary dendrites etc.) and the total dendritic lengt of the glomeruli tufts. Average values were calculated for each neuron analyzed. Graphs and statistical analyses (t-tests) were made using GraphPad Prism.

Quality control of ribotag pulldowns

The enrichment and de-enrichment of markers following neuronal markers were tested: for the olfactory bulb pull-down samples: *Rbm20*, *Vglut2*, *Vglut1*, *Tbr2*, *Pcdh21*, Vgat, *Gfap*, *Gad67*. For pulldowns from cortex of PVCre mice: *Rbm20*, *Pvalb*, *Vgat*, *Gad67*, *Vglut1*, *Gfap*. In both cases, *Gapdh* mRNA was used as a housekeeping gene for normalization. The fold enrichment and de-enrichment values of each

marker were calculated for each cell population in immunoprecipitated RNA, comparing it to input purifications. Only samples that showed correct enrichment or de-enrichment for excitatory or inhibitory neuronal markers and a de-enrichment for glia markers were further used for sequencing. DNA oligonucleotides were used with FastStart Universal SYBR Green Master (Roche, 4913914001) and comparative C_T method. For each assay, three technical replicates were performed and the mean was calculated. RT-qPCR assays were analyzed with the StepOne software. DNA Oligonucleotides used (name and sequence 5'-3' are indicated):
List of primer sequences:

Primer Name	Sequence
Rbm20	Forward (Sense) TGCATGCCCAGAAATGCCTGCT Reverse (AntiSense) AAAGGCCCTCGTTGGAATGGCT
Tbr2	Forward (Sense) ATAAACGGACTCAACCCCACC Reverse (AntiSense) CCCTGCATGTTATTGTCCGC
Pchd21	Forward (Sense) ATCACTGTCAACGACTCAGACC Reverse (AntiSense) GTCAATGGCAGCTGAGTTTTCC
Vglut2	Forward (Sense) GCATGGTCTGGTACATGTTCTG Reverse (AntiSense) GACGGGCATGGATGTGAAAAAC
Gad67	Forward (Sense) GTACTTCCCAGAAGTGAAAC Reverse (AntiSense) GAATAGTGACTGTGTTCTAGG
Gfap	Forward (Sense) CTCGTGTGGATTTGGAGAG Reverse (AntiSense) AGTTCTCGAACTTCCTCCT
Vglut1	Forward (Sense) ACCCTGTTACGAAGTTTAACAC¨ Reverse (AntiSense) CAGGTAGAAGGTCCAGCTG
Vgat	Forward (Sense) CGTGACAAATGCCATTCAG Reverse (AntiSense) AAGATGATGAGGAACAACCC
Pvalb	Forward (Sense) CATTGAGGAGGATGAGCTG Reverse (AntiSense) AGTGGAGAATTCTTCAACCC

CLIP sample preparation

The lysate was transferred into a glass homogenizer and homogenized by 30 strokes on ice. 1 ml aliquots of homogenized tissue were transferred to 2 ml tubes, 10 µl of RNaseI (Thermofisher) diluted in PBS (1:5 -1:40) were added to each tube. Samples were incubated at 37°C with shaking for 5 min at 1200 x rpm and then put on ice. 10 µl RNasin RNase-inhibitor (40 U/µl, Promega) were added to each tube. Sample were mixed and centrifuged at 16.000 x g for 15min at 4°C. The supernatants were transferred to a new tube and 60 µl from each sample were taken and further processed for sized matched INPUT (SMIn). 10 µl HA-magnetic beads (Pierce) was added to each sample and incubated

at 4°C for 4h in a rotating shaker. Following incubation, the beads were washed 2x with a high salt wash buffer (50mM Tris-HCl pH7.5, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate), 2x with the lysis buffer, 2x with low salt wash buffer (20 mM Tris-HCl pH7.5, 10mM $MgCl_2$, 0.2% Tween-20) and 1x with PNK buffer (70 mM Tris-HCl pH6.5, 10 mM $MgCl_2$). Beads were re-suspended in 100 µl PNK-mix (70 mM Tris-HCl pH6.5, 10 mM MgCl2, 1 mM DTT, 100 U RNasin, 1 U TurboDNase, 25 U Polynucleotide-Kinase (NEB)) and incubated at 37°C for for 20 min on a shaking termomixer (1200 x rpm). Upon RNA dephosphorylation, the beads were washed (2x high salt, 2x lysis and 2x low salt buffers as before) and additionally with 1x Ligase buffer (50 mM Tris-HCl pH7.5, 10 mM MgCl2). Beads were then re-suspended in 50 µl ligase mix (50 mM Tris-HCl pH7.5, 10 mM $MgCl_2$, 1 mM ATP, 3% DMSO, 15% PEG8000, 30 U RNasin, 75 U T4 RNA-ligase (NEB)). 10 µl of the beads / ligase mix were transferred to a new tube and 1 µl of pCp-Biotin (Jena Bioscience) were added to validate IP of the RNA-protein-complexes by western blot. 4 µl of the RNA-adaptor mix containing 40 µM of each InvRiL19 & InvRand3Tr3 (IDT) were added to the remaining of the samples (40 µl). Samples were placed at RT for 2 h for adaptor ligation. Samples were washed 2x with high salt, 2x with lysis and 1x with low salt buffers. Finally, beads were re-suspended in 1x LDS sample buffer (Thermofisher) supplemented with 10 uM DTT and incubated for 10 min at 65°C, shaking on a thermomixer at 1200 x rpm. Eluates or inputs were loaded on 4-12% Bis-Tris, 1.5 mm gel (Thermofisher) and separated at 130 V for ~ 1.5 h. Proteins were transferred overnight at 30 V to a nitrocellulose membrane (Amersham). The membranes were placed in a 15 cm Petri dish on ice and an area between 55 and 145 kDa was cut out small pieces and transferred in a 2 ml tube.

RNA extraction, reverse transcription using InvAR17 primer, cDNA clean-up using silane beads (Thermofisher), second adaptor ligation (InvRand3Tr3) and cDNA purification steps were performed as previously described (Van Nostrand *et al*, 2016). The sequencing libraries were amplified using Q5-DNA polymerase (NEB) and i50X/i70X Illumina indexing primers (IDT). Final libraries were amplified with 14 cycles Libraries were purified and concentrated with ProNEX size selective purification system (Promega) using sample/beads ratio of 1/2.4. Samples were loaded on a 2% agarose gel and the area corresponding to the size between 175 bp and 350 bp was cut out. The amplified and purified libraries were then extracted from the gel using gel extraction kit (Machery&Nagel) and eluted with 16 µl.

The concentrations and the size distributions of the libraries were determined on the Fragment analyzer system (Agilent). 75 bp single-end sequencing was performed on the NextSeq500 platform using Mid Output Kit v2.5 (75 cycles).

Adaptor and primer sequences used in this study:

Name	Sequence
InvRi L19	/5Phos/rArGrArUrCrGrGrArArGrArGrCrArCrA rCrGrUrC/3SpC3/
InvRand3 Tr3	/5Phos/NNNNNNNNNNAGATCGGAAGA GCGTCGTGT/3SpC3/
InvA R17	CAGACGTGTGCTCTTCCGA
i501	AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTACACG ACGCTCTTCCGATC*T
i502	AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCTACACG ACGCTCTTCCGATC*T
i503	AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTTTCCCTACACG ACGCTCTTCCGATC*T
i504	AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCTTTCCCTACACG ACGCTCTTCCGATC*T
i701	CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTG CTCTTCCGATC*T
i702	CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTG CTCTTCCGATC*T
i703	CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTGTG CTCTTCCGATC*T
i704	CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTG CTCTTCCGATC*T

X* = Phosphorthioated base

References cited in Supplementary Methods

Ahrne E, Glatter T, Vigano C, Schubert C, Nigg EA, Schmidt A (2016) Evaluation and Improvement of Quantification Accuracy in Isobaric Mass Tag-Based Protein Quantification Experiments. *Journal of proteome research* 15: 2537-2547

Haddad R, Lanjuin A, Madisen L, Zeng H, Murthy VN, Uchida N (2013) Olfactory cortical neurons read out a relative time code in the olfactory bulb. *Nat Neurosci* 16: 949-957

Hauser D, Behr K, Konno K, Schreiner D, Schmidt A, Watanabe M, Bischofberger J, Scheiffele P (2022) Targeted proteoform mapping uncovers specific Neurexin-3 variants required for dendritic inhibition. *Neuron* 110: 2094-2109 e2010

Hippenmeyer S, Vrieseling E, Sigrist M, Portmann T, Laengle C, Ladle DR, Arber S (2005) A developmental switch in the response of DRG neurons to ETS transcription factor signaling. *PLoS Biol* 3: e159

Khan MA, Reckman YJ, Aufiero S, van den Hoogenhof MM, van der Made I, Beqqali A, Koolbergen DR, Rasmussen TB, van der Velden J, Creemers EE *et al* (2016) RBM20 Regulates Circular RNA Production From the Titin Gene. *Circ Res* 119: 996-1003

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR *et al* (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature neuroscience* 13: 133-140

Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ (2012) Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* 11: 1475-1488

Sanz E, Yang L, Su T, Morris DR, McKnight GS, Amieux PS (2009) Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci U S A* 106: 13939-13944

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B *et al* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9: 676-682

Taniguchi H, He M, Wu P, Kim S, Paik R, Sugino K, Kvitsiani D, Fu Y, Lu J, Lin Y *et al* (2011) A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* 71: 995-1013

Tsien JZ, Chen DF, Gerber D, Tom C, Mercer EH, Anderson DJ, Mayford M, Kandel ER, Tonegawa S (1996) Subregion- and cell type-restricted gene knockout in mouse brain. *Cell* 87: 1317-1326

Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K *et al* (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* 13: 508-514

Vong L, Ye C, Yang Z, Choi B, Chua S, Jr., Lowell BB (2011) Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* 71: 142-154

Supplementary Figures
Figure S1

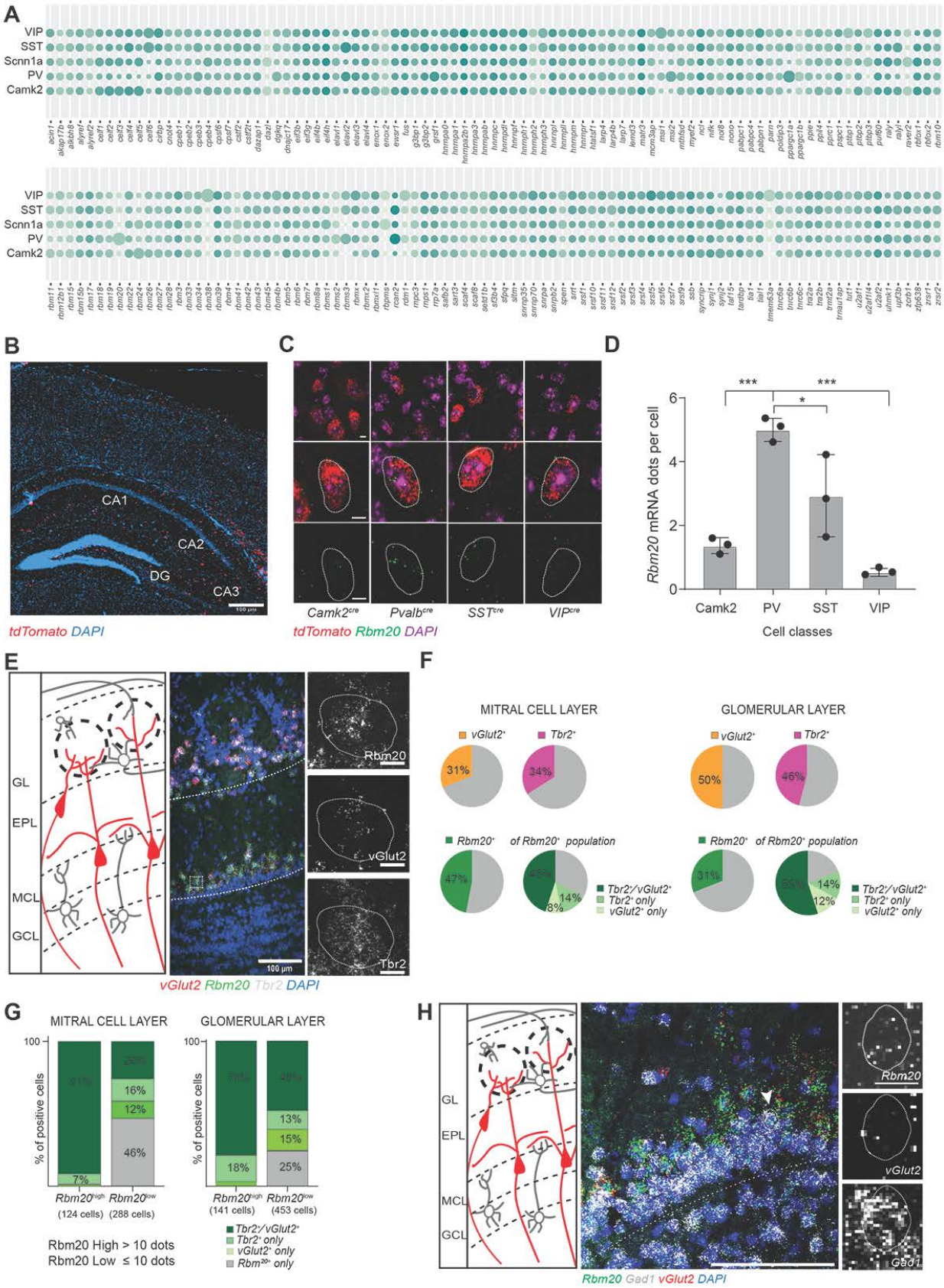

Figure S1: Characterization of Rbm20 expression in the brain

A. Dot plot of the expression of a hand-curated list of RBPs across different neuronal neocortical populations. RBPs were chosen based on the presence of an RNA recognition motif (RRM) in their sequence and their expression in the neocortex. RBPs expression was measured by Ribo-

TRAP sequencing and expressed in RPKM values normalized over the mean expression across different neuronal populations.

B. Fluorescent *in situ* hybridization (FISH) on P25 brain slices from mice with genetic marking of cell populations (*Pvalb*^{cre}*::Ai9* cre-dependent tdTomato expression). Red: *tdTtomato* mRNA, blue: DAPI. Scale bar 100 µm;

C. Fluorescent *in situ* hybridization for *Rbm20* transcripts in *tdTomato* -marked cells for different cell classes (cre-dependent Ai9 tdTomato reporter crossed to the indicated cre-recombinase expressing lines: *Camk2*, *Pvalb*, *Sst*, *Vip*) in P23-26 somatosensory cortex cells in layer 5. Green: *Rbm20* mRNA, red: *tdTtomato* mRNA, blue: DAPI. Scale bar 10 µm.

D. Quantification of *Rbm20* mRNA expression as in C, expressed as mean number of fluorescent dots per cell from three animals per genotype. P-value < 0.01, one-way Anova.

E. Schematic illustration of the olfactory bulb circuitry and cell types. GL: glomerular layer, EPL:external plexiform layer, MCL: mitral cell layer, GCL: granule cell layer, (left). Fluorescent *in situ* hybridization (FISH) on brain slices for *Rbm20* (green), *Tbr2* (gray), *Vglut2* (red) mRNAs (middle). The insets show the example of a cell expressing all three markers (right). Scale bar: 100 µm; scale bar insets: 10 µm.

F. Pie charts indicating the quantification of the percentage of cells of the MCL and GL expressing *Vglut2, Rbm20* and *Tbr2* transcripts. Amongst the Rbm20^{+} cells, the percentage of neurons presenting co-localization with glutamatergic markers was calculated.

G. Quantification of the percentage of neurons of the mitral cell layer (MCL) and glomerular layer (GL) expressing high or low *Rbm20* mRNA levels, co-localizing with *Tbr2* and *Vglut2* markers. The absolute number of high and low *Rbm20* expressing neurons identified in both the mitral cell layer and the glomeruli layer is reported in brackets.

H. Fluorescent *in situ* hybridization (FISH) on brain slices of *Rbm20* (green), *Gad1* (gray), *Vglut2* (red) mRNAs. Scale bar 100 µm. The arrow in the inset on the right of the panel shows an example of a cell expressing low *Rbm20* levels and co-localizing with *Gad1* marker but not with Vg*lut2* marker. Scale bar inset: 10 µm.

Figure S2

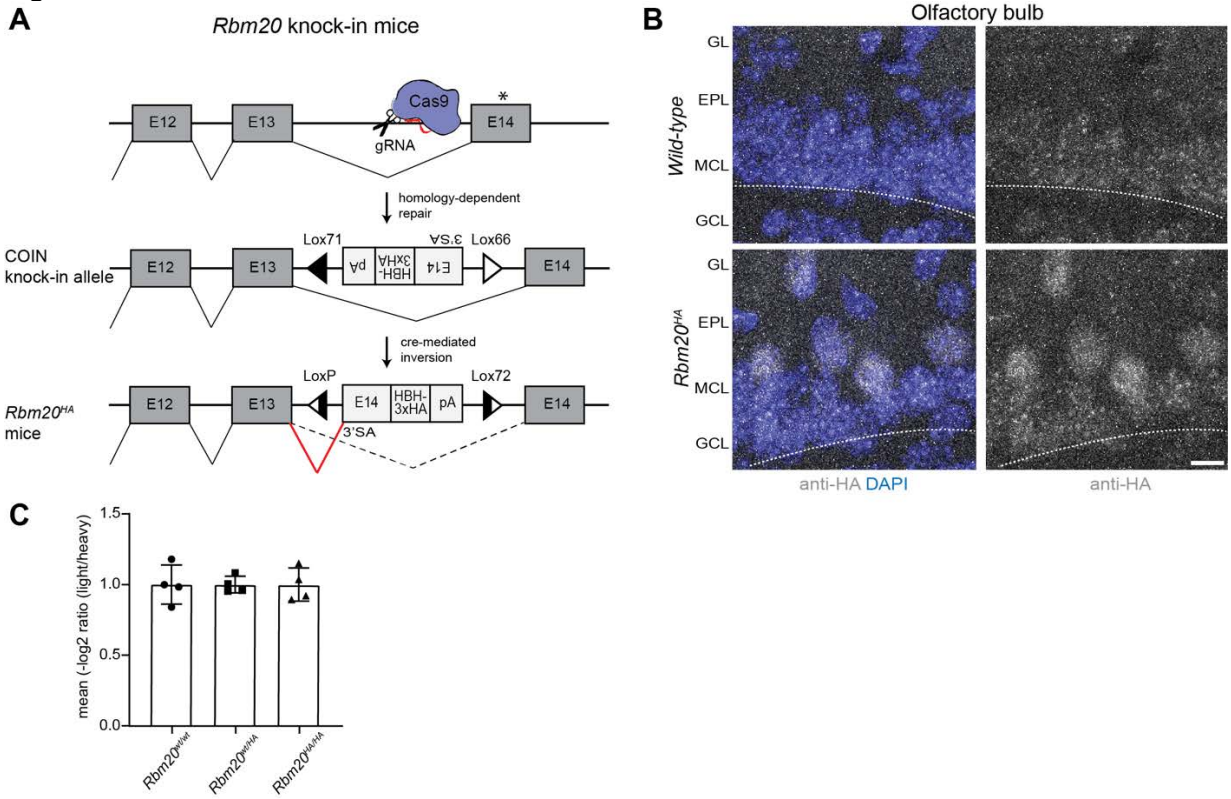

Figure S2: Generation of *Rbm20* – HA tagged mouse line

A. Schematic illustration of the "COIN allele" strategy used for generation of *Rbm20*^{HA} knock-in mice. The COIN allele module was inserted in the *Rbm20* locus with the CRISPR-CAS9 system. Upon Cre-mediated recombination of the loxP sites, the COIN module is inverted and the presence of a strong synthetic 3' splicing acceptor site (3'SA) results in the expression of the tagged *Rbm20* isoform. While the strategy was achieved to obtain conditional tagging in specific cell types the cre-mediated inversion frequency was too low for use *in vivo*.

B. Immunohistochemistry of RBM20^{HA} protein in mitral cells of wild type and *Rbm20*^{HA} mice. HA (gray), DAPI (blue). Scale bar: 10 µm.

C. Targeted proteomic analysis on heart samples of wild type (WT), *Rbm20*$^{WT/HA}$ heterozygous, and *Rbm20*$^{HA/HA}$ homozygous knock-in mice. Three proteotypic peptides were quantified. Plotted results represent means of all three peptides normalized to wild-type samples. Four biological replicates per genotype. The mean of the $-\log_2$ ratio (light/heavy) peptides was calculated and displayed. Note that an assessment of RBM20 expression level in the *Rbm20*^{HA} mice by Western blot was not possible as the C-terminal HA-epitope reduced binding affinity of the anti-RBM20 antibody raised against the C-terminus of the protein.

Figure S3

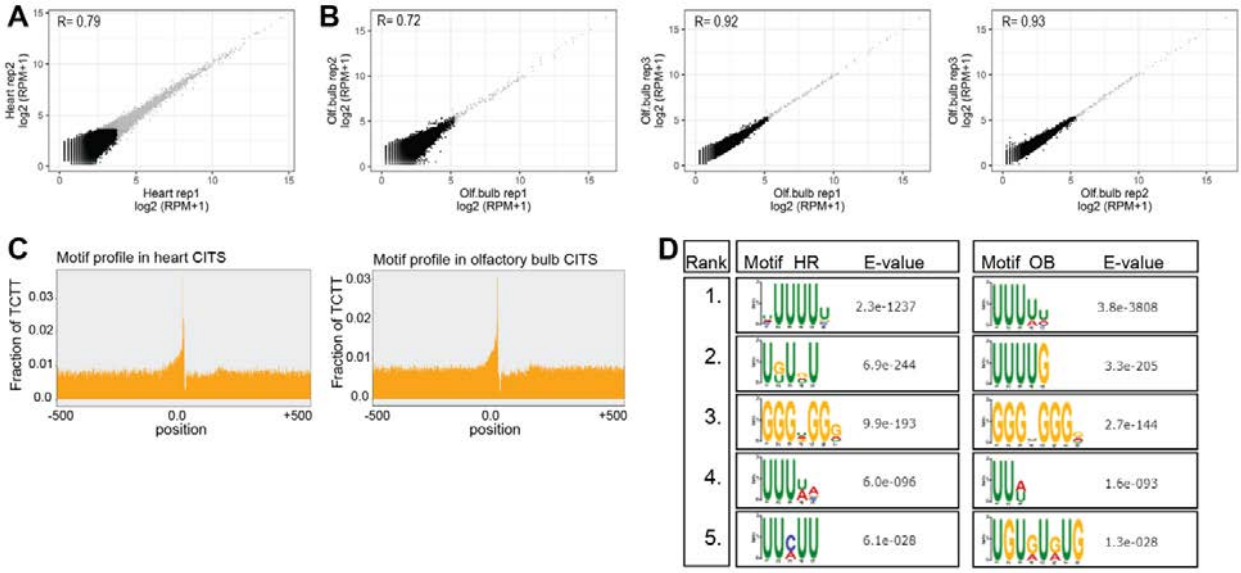

Figure S3: Identification of RBM20 binding sites on transcript mRNAs

A – B Correlation analysis of normalized counts (reads per million [rpm] +1) of called CLIP peaks between seCLIP replicates of the heart (2 samples) and the olfactory bulb (3 samples). Gray shades represent density of the data points. Pearson coefficients are indicated in above the corresponding plots.

C. Enrichment of the TCTT motif at cross-link-induced truncation sites (CITS) in both heart and olfactory bulb tissues. The enrichment is calculated from the frequency of the TCTT motif starting at each position of the inferred cross-linked site, normalized by frequency of the same motif in adjacent flanking regions.

D. Motif finding analysis performed with DREME on heart and olfactory bulb seCLIP datasets. The first 5 statistically significant motifs are reported, ranked based on the enrichment p-value (E-value; right of each panel). The E-value is defined as the p-value times the number of candidate motifs tested. The enrichment p-value is calculated using Fisher's Exact Test for enrichment of the motif in the positive sequences, calculated after erasing sites that match previously found motifs. Note that the G-rich motif (rank 3) is commonly found to be non-specifically recovered in seCLIP datasets.

Figure S4

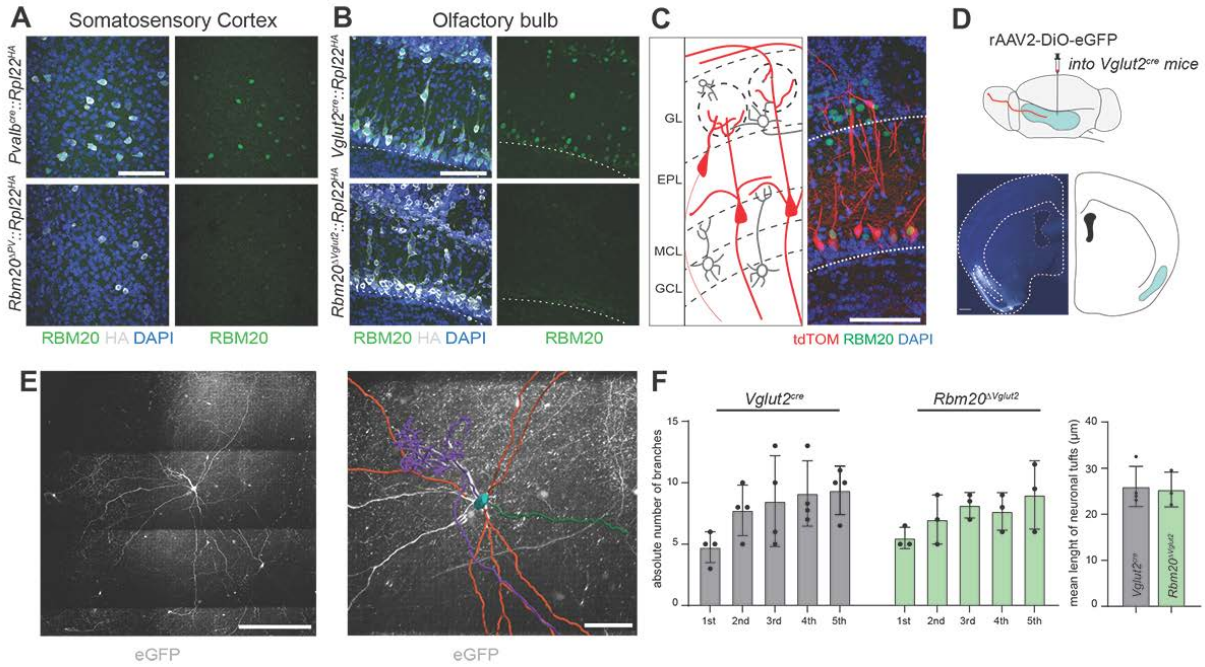

Figure S4. Normal morphological differentiation of mitral cells in the absence of RBM20.

A, B. Immunostaining of RBM20 (green), RPL22-HA (gray) and DAPI (blue) in the cortex and olfactory bulb of P35 *Pvalb*^{Cre}*::Rpl22*$^{HA/HA}$ and (B) *Vglut2*^{Cre}*::Rpl22* $^{HA/HA}$ mice, compared to corresponding littermates carrying the conditional *Rbm20*$^{fl/fl}$ alleles (*Rbm20*$^{\Delta PV}$ and *Rbm20*$^{\Delta Vglut2}$, lower panels).

C. RBM20-positive (green) mitral cells retrogradely labeled through injection of rAAV2-SYN-Cre virus (red) in piriform cortex of a *Ai9*^{tdTOM} mice at P35. Scale bar 100 µm.

D. Schematic illustration of the rAAV2-DIO-eGFP virus injection into the Piriform cortex of Vg*lut2*^{Cre} mice for retrograde labelling of olfactory bulb mitral cells. Representative image of the site of viral injection in the piriform cortex. Scale bar 500 µm.

E. GFP^{+} mitral cell (gray) retrogradely labeled through injection of rAAV2-SYN-DiO-GFP virus in the piriform cortex of Vg*lut2*^{Cre} and *Vglut2cre::Rbm20*$^{fl/fl}$ (*Rbm20*$^{\Delta Vglut2}$) mice. Tracing of the neuronal arborization is displayed on the right. Scale bar 500 µm.

F. Quantification of the absolute number of branches and the mean length of the neuronal tufts in reconstructed mitral cells of Vg*lut2*^{Cre} and *Rbm20*$^{\Delta Vglut2}$ mice.

Figure S5

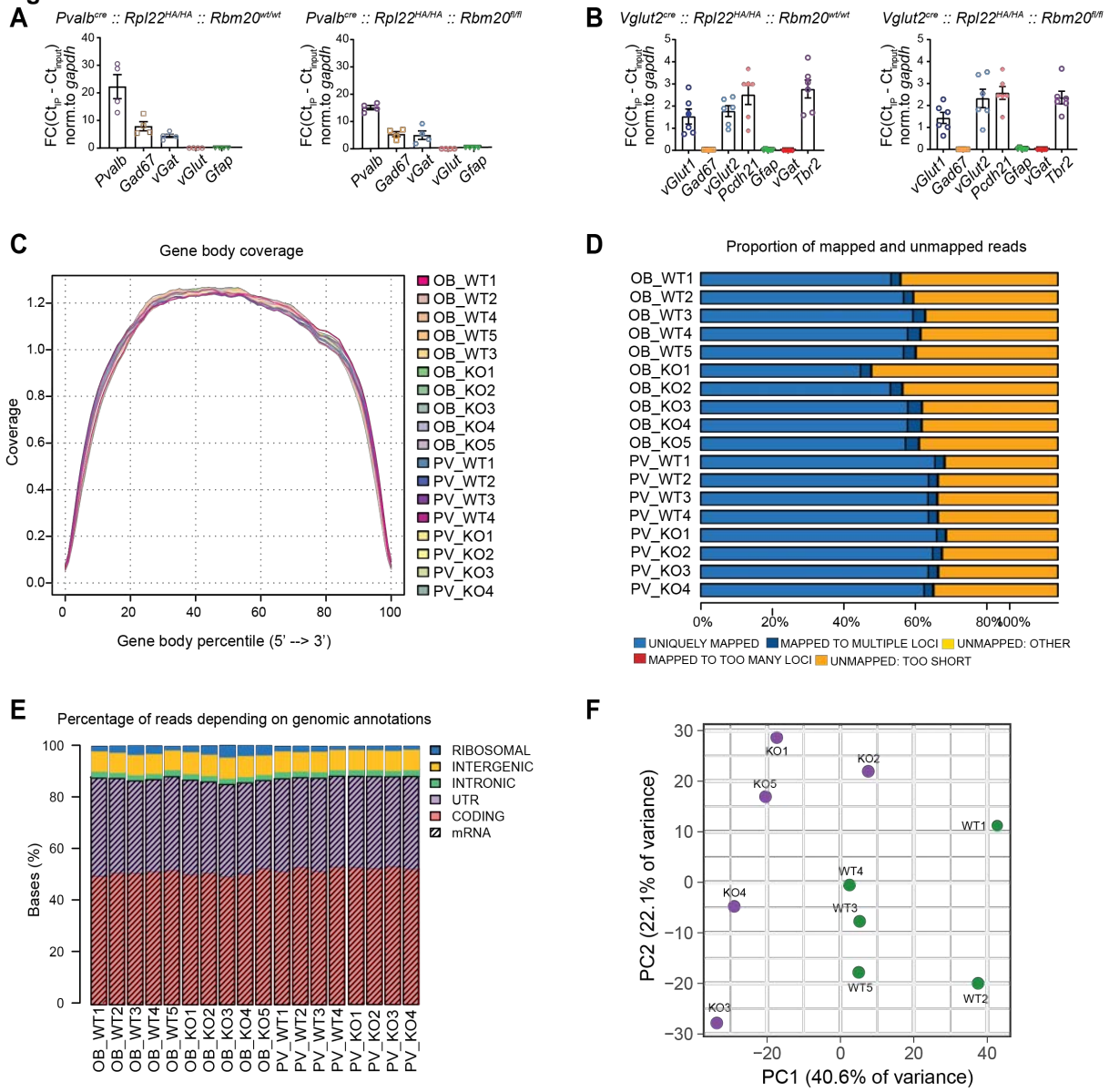

Figure S5: Quality control analysis of Ribo-TRAP RNA-sequencing samples

A – B. Fold-enrichment (FC) of markers specific to inhibitory cortical neurons or glutamatergic neurons for WT and cKO samples. For PV samples, the following markers were tested: *Pvalb*, *Gad67*, *Vgat*, *Vglut1*, *Gfap*. For glutamatergic: *Vglut1*, *Vglut2*, *Gad67*, *Pcdh21*, *Gfap*, *Vgat* and *Tbr2*.

C. Coverage plot indicating the percentage of read bases at a given position of the transcript. No sample displayed 3' or 5' coverage bias across the transcript length.

D. Bar plot of each biological replicate where for each sample the following parameters are indicated: the proportion of number of reads uniquely mapped, mapped to multiple loci, mapped to too many loci or unmapped reads for all the samples. All samples show highly similar values across biological replicates, as well as across brain region and genotype, suggesting a high consistency and homogeneity of the RNA-seq data.

E. Bar plot representing the relative percentage of reads falling on genomic features for all the biological samples.

F. PCA of genes expressed in each olfactory bulb sample (n=5 biologically independent samples per genotype). Variance explained by the principal components 1 and 2 (PC1 and PC2) is indicated. Gene expression values were normalized by Variance Stabilizing Transformation (VST).

Supplementary Tables:

Table S1: Identified RBM20 binding sites in the heart and olfactory bulb tissues. List of RBM20 peaks identified on transcript mRNAs in the heart and in olfactory bulb tissues. Peaks were identified through the peak caller Clipper followed by irreproducible discovery rate (IDR) analysis between replicates. In this table, beyond the standard output parameters produced by IDR, columns containing information about the annotation of the targeted transcript and the position of RBM20 binding site in relation to the exon-intron boundaries (R package 'AnnotatR) are reported. Moreover, the list of read counts summarized using featureCounts for identification of expressed genes in the input samples of heart and olfactory bulb is reported.

Table S2: Gene Ontology analysis of transcripts directly bound by RBM20. Gene ontology analysis results by Panther of mRNA transcripts directly bound by RBM20 RNA-binding protein in both heart and olfactory bulb tissues.

Table S3: Expressed genes and percentage of mapped and unmapped unique reads in Ribo-TRAP RNA-sequencing experiments. Read counts summarized using featureCounts. For each sample from either *Vglut2*$^+$ neurons of the olfactory bulb or PV$^+$ interneurons of the neocortex.

Table S4: Summary of differential gene expression analysis. Expression values (FPKM) of genes identified in Parvalbumin positive and *Vglut2*$^+$ neurons in Ribo-TRAP RNA sequencing experiments and the results of their differential gene expression analysis by DESEQ2 in wild-type vs. *Rbm20* conditional knock-out mice.

Table S5: Summary of alternative exon usage analysis. Expression values (RPKM) of all the exons identified in Parvalbumin positive and Vglut2 positive cells in Ribo-TRAP RNA sequencing experiments and the results of the alternative exon usage analysis in wild-type vs. *Rbm20* conditional knock-out mice.

Table S6: Gene Ontology of de-regulated transcripts and alternatively spliced exons. Gene ontology analysis results by Panther for de-regulated transcripts (all, downregulated and upregulated) and exon usage analysis in olfactory bulb neurons upon RBM20 ablation.

Table S7: Analysis of intron length. Analysis of intron length in de-regulated, non-regulated and RBM20-bound transcripts in olfactory bulb neurons.