SUPPLEMENTARY MATERIALS AND METHODS

HITS-CLIP and RNA-seq

HITS-CLIP was performed essentially as described previously (Licatalosi et al., 2008; Preitner et al., 2014) on E14 Swiss Webster mouse brains. Briefly, embryonic brain tissues were partially dissociated before UV irradiation (UV at 254 nm, 3X 300 mJ/cm2 in a Stratalinker). Cell lysates were sonicated, treated with DNase, and digested with benzonase (Sigma E8263; at a final dilution of 1:2000; 10 min at 37°C) to produce protein-bound RNAs of ~40 nt fragments. Then pre-cleared lysates were mixed with 30 µl Sepharose beads and 30 µg antibody (mouse polyclonal anti-IMP2 antibody, Abnova H00010644-A01; rabbit anti-IMP2 antibody, Proteintech 11601-1-AP), and rotated for 3 hours at 4°C. Immunoprecipitated protein-RNA complexes were stringently washed with SDS- containing buffer and dephosphorylated, and the radiolabeled RNA adapter (5'-UCGUAUGCCGUCUUCUGCUUGU-3') were ligated at the 3' end of the RNA fragments. Protein-RNA complexes were then phosphorylated at the 5' end, separated on a 4-12% denaturing NuPAGE gel (Invitrogen), transferred to a nitrocellulose membrane, and exposed on a film. Narrow bands corresponding to the size of the specific protein-RNA complexes were then excised from the membrane. RNA fragments were extracted, ligated to the 5' adapter (5'-GUUCAGAGUUCUACAGUCCGACGAUC-3') and amplified by RT-PCR (primers: AATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGA, and CAAGCAGAAGACGGCATACGA). The purified PCR products were sequenced on Illumina sequencing platform.

For mRNA-seq, polyadenylated RNAs from E14 mouse brains were isolated, fragmented with benzonase, and ~40 nt RNA fragments isolated on a polyacrylamide gel. RNA fragments were dephosphorylated, 3' linkers were ligated, RNA fragments were gel isolated again,

rephosophorylated with polynucleotide kinase, and 5' linkers were added, before cDNA amplification, libary preparation, and sequencing in the same way as for CLIP.

Processing and alignment of HITS-CLIP reads

The quality of raw sequence reads (36 nt) was analyzed using fastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter sequences were trimmed, and reads with more than 2 undetermined bases (Ns) were removed. The sum of base Phred quality scores of each read was calculated, and reads with the lowest 1% sum of Phred scores were filtered. The rest reads were trimmed from the 3'- end to retain 28 nt of high quality. After filtering and trimming, reads were aligned to mouse genome (NCBI37/mm9 assembly) using Bowtie (Langmead et al., 2009) (http://bowtie-bio.sourceforge.net/index.shtml), allowing up to 2 mismatches. Reads aligned to multiple positions in the genome were filtered. RefSeq mRNA database was used to build the exon junction index to align reads that overlap the exon junction sites. Those reads that were not aligned to genome were aligned to exon junctions, and only uniquely aligned reads were retained. Reads that aligned to the same genomic locations were collapsed to remove potential duplicates resulting from PCR amplifications. The control mRNA-seq data were aligned to the mouse genome and exon junction sites by the same procedure. The aligned reads were annotated based on RefSeq mRNA database. The distribution of reads in mouse genome was visualized on UCSC genome browser.

HITS-CLIP data have been deposited at NCBI Gene Expression Omnibus, and are accessible through GEO Series accession number GSE83822 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE83822).

Peak identification

To identify IMP2 binding sites, we next combined and analyzed reads from all four CLIP replicates. Overlapping reads were grouped into clusters, and each cluster was assigned two values, peak height and fold change. Peak height was defined as the maximum number of overlapping reads in the cluster, and median fold change was calculated by comparing to the randomly sampled background from mRNA-seq data. To differentiate significantly enriched IMP2 binding sites from false positives, we generated mock CLIP data sets by randomly sampling the same number of sequence reads as in the IMP2 CLIP samples from the control mRNA-seq data. The false discovery rate was calculated as the ratio of the number of peaks identified from mock CLIP data to that from CLIP data using the same cutoff. We selected high-confidence peaks on mature mRNAs, based on the following criteria: (a) a minimal peak height of 18 in pooled CLIP samples, which is greater than the maximal peak height in five different sets of mock CLIPs, (b) a minimal median fold change of 2, (c) reproducible peaks requiring that within a peak region, at least 4 overlapping reads be present in the four replicates. Additionally, in order to eliminate occasional mRNAs with high read coverage along the transcript but no prominent peaks, we required that the peak height should be at least 1.2x the highest peak in the introns, and at least 10x the average read coverage in the introns. The 1850 identified IMP2 binding sites mapped to 747 protein-coding genes (from RefSeq database).

Motif searching

Motif analysis was performed using MEME tools (Bailey et al., 2009) (http://meme.nbcr.net/meme/). The sequence logos were generated on WebLogo 2.8.2 (http://weblogo.berkeley.edu). The *Z*-score statistic for motif finding was applied as described (Tompa, 1999).

GO analysis of target mRNAs

The enrichment of gene ontology (GO) terms in IMP2 target mRNAs was analyzed using DAVID online tools (Huang et al., 2009) (http://david.abcc.ncifcrf.gov). Transcript abundance was determined from mRNA-seq data, and transcripts were considered as expressed if the lower bound of the 95% confidence interval of the estimated fragments per kilobase of exon model per million mapped fragments (FPKM) was greater than 0.1 (Trapnell et al., 2010) (http://cufflinks.cbcb.umd.edu). DAVID reported *p*-values, based on a modified Fisher's exact test. The *p*-values reported for this study were adjusted using the Benjamini-Hochberg method.

Functional analysis using IPA

Canonical pathway analysis, disease association, and network analysis were performed using the core analysis module of the Ingenuity Pathway Analysis (IPA) package. Figure 3D shows disorders corresponding to a specific disease and represented by at least two genes, within the IPA neurological disease category. Single genes causing human Mendelian disorders (Table S2) were identified by a systematic manual search in the OMIM database (http://www.ncbi.nlm.nih.gov/omim).

Comparison of IMP2 targets with axonally localized mRNAs

IMP2 target mRNAs were compared to the axonal mRNA list identified in rat dorsal root ganglion (Gumy et al., 2011). The two lists of mRNAs expressed in embryonic axons and adult axons were combined, and the Entrez Gene IDs and official symbols converted from Affymetrix ID were used for the comparison. The overlap between the target set and the axonal mRNAs was evaluated by Fisher's exact test. Venn diagram was drawn to scale using Venn diagram generator available online (http://jura.wi.mit.edu/bioc/tools/venn.php).

Immunolocalization in embryos

Cryostat sections (10 µm) were collected on Superfrost Plus slides (Fisher) and kept at -80°C. Slides were blocked in 1X PBS with 10% normal serum and 0.1% Triton X-100 for 30 min at room temperature (RT), incubated overnight at 4°C with the primary antibody diluted in the blocking solution, washed with PBS, incubated for 1 hour at RT with Alexa Fluor conjugated secondary antibody (Molecular Probes, 1:500) diluted in PBS with 10% normal serum, counterstained with DAPI, washed with PBS, and mounted using Fluoromount G mounting medium (SouthernBiotech). Primary antibodies and dilutions: anti-GFP antibody, Molecular Probes A11120, 1:200; anti-IMP1 antibody, 1:400 anti-IMP2 antibody, 1:1000; anti-IMP3 antibody, 1:1000 anti-neurofilament antibody, Sigma N5264, 1:400). IMP1, 2, and 3 antibodies were obtained from Finn C Nielsen's lab at University of Copenhagen in Denmark. Images were captured using a Nikon 80i upright fluorescence microscope.

DNA constructs for RNAi experiments

Plasmids expressing IMP2 shRNAs were produced by cloning the following sequences under the U6 promoter (sense IMP2 target sequence underlined). As a control for RNAi experiments, a plasmid expressing a non-targeting control shRNA under the U6 promoter was used (based on the HuSH plasmid, Origene, and computationally confirmed not to recognize sequences in the chick transcriptome).

IMP2 shRNA1: GATCCACGGAACAGTGGAGAATGTGAAGCTTGA

CATTCTCCACTGTTCCGTTTTTTTGGAAGC

IMP2 shRNA2: GATCCGGACCTTAATGCCTTCAGAGAAGCTTGT

CTGAAGGCATTAAGGTCCTTTTTTGGAAGC

Control shRNA: GATCCACAAGATGAAGAGCACCAAGAAGCTTG

TTGGTGCTCTTCATCTTGT TTTTTTTGGAAGC

To test the effectiveness and specificity of shRNAs against IMP2, coding sequence constructs of IMP2 mRNA, and also IMP1 and IMP3 as controls, were first amplified by PCR from chick E5 spinal cord cDNA using the oligos shown below, and cloned into vector pCS2-GFP to express GFP-IMP1/2/3 fusion proteins respectively.

IMP1-5'-EcoR1: AAAAGAATTCGCCGCCACCATGAACAAGCTGTACATCGGGAACC

IMP1-3'-Xba1: AAAATCTAGATTTCCTCCTCGCTTGCAGCTGGCC

IMP2-5'-Sal1: AAAAGTCGACAGATCTATGTGTCCTCCGGATGGGGAATATGC

IMP2-3'-Xba1: AAAATCTAGACCTGCTGGACGATTTCCCTGATCTTGC

IMP2-5'-Xho1: AAAACTCGAGGGCCGGGGGCAGCGCAGCATCGC

IMP2-3'-Xba1-2: AAAATCTAGATTACCCTCTCCTGCTCGGGGGTTCCC

IMP3-5'-EcoR1: AAAAGAATTCGCCGCCACCATGAACAAGCTCTACATCGGCAACCT

IMP3-3'-Xba1: AAAATCTAGATTTTCGTCTTGGCTGAGGCTGTCC

To confirm knockdown by Western blot, because only a fraction of commissural neurons become electroporated in vivo, the shRNAs were tested in cultured cells. 0.8 µg of shRNA plasmids (one control shRNA, two IMP2 shRNAs) and 0.1µg of RFP plasmids were transfected into 293 cells (HEK-293, American Type Culture Collection) in 24 well plates in triplicate using lipofectamine 2000 (Invitrogen). After 24 hours, 0.01 µg of each template plasmid (GFP-IMP1, GFP-IMP2, GFP-IMP3) was transfected separately into previously shRNA-transfected cells. After 48 hours, cells were washed with cold PBS, lysed in 100 µl/well SDS-PAGE sample buffer, and sonicated 30 seconds to shear DNA. Then the lysates were heated at 95°C for 10 min and analyzed by SDS-PAGE and Western blot with anti-GFP antibody (1:1000, Invitrogen A11122).

Characterization of spinal commissural axons by ovo electroporation

In ovo electroporation in chicken embryos was performed essentially as described (Reeber et al., 2008; Avraham et al., 2010). Briefly, fertilized white eggs (Charles River Laboratories)

were incubated at 38°C for 64 hours. DNA mix (1-4 µg/µl) containing IMP2 or control shRNA and Math1-GFP plasmid diluted in 1xPBS and trypan blue solution was introduced with a fine glass micropipet into the spinal cord lumen of HH stage 17 chicken embryos, and unilaterally electroporated into the spinal cord. The Math1-GFP plasmid was constructed by replacing the CMV promoter in the plasmid pCS2-GFP with the Math1 promoter region from Math1-Tau-GFP plasmid, a kind gift from Dr. Jane Johnson (UT Southwestern) (Helms and Johnson, 1998; Lumpkin et al., 2003). Square-wave current (five pulses, 25 V, 50 ms ON and 950 ms interval) was generated using CUY21 electroporator (BEX CO., LTD). Windowed eggs were further incubated at 37°C for 64 hrs, then mounted as open book preparations and imaged by confocal microscopy using a Nikon Ti inverted microscope, equipped with a Yokagawa CSU-10 spinning disk confocal, a Hamamatsu ORCA-AG cooled CCD camera, and Metamorph image acquisition software. Axon numbers were counted at the dorsoventral locations indicated in Fig. 4 by an investigator blind to the shRNA used. Results were tested for statistical significance by Student's unpaired two-tailed t-test, with similar variance between the groups being compared.

To assess the effects of IMP2 RNAi on ROBO1 expression, pmath1-GFP and shRNA plasmids were introduced into the spinal cord of HH stage 17 chick embryos in ovo as described above. 64h after electroporation, commissural neuron cultures were prepared by isolating, and dissociating dorsal spinal cord tissue, followed by culture for 24h on glass coverslips coated with poly-D-Lysine and laminin. Cells were then fixed in 4%paraformaldehyde, 3% sucrose, washed four times with PBS, and incubated with rabbit anti-ROBO1 antibodies (Rockland # 600-401-692) in PBS + 1% BSA + 0.1% Triton X-100 overnight at 4C, washed again as above, incubated 1h at room temperature in Alexa flurophore-conjugated anti-rabbit antibody and phalloidin-rhodamine in PBS + 1% BSA, and mounted in Fluoromount G (SouthernBiotech). Images were captured using a Nikon 80i

upright fluorescence microscope. ROBO1 immunolabeling was quantified using the Metamorph software package. Immunolabeling intensities were measured in cell bodies and growth cones traced with Metamorph based on the phalloidin-rhodamine staining. Neurons were selected for quantitation based on GFP expression, blind to the ROBO1 channel. Fig. 4G displays total axon immunofluorescence; comparable effects on ROBO1 were seen when normalizing by axon area or GFP levels. In micrographs for display (Fig. S4B), contrast and brightness were adjusted for visualization; for quantitation, unadjusted raw images were used with no pixels saturated. Results were tested for statistical significance by Student's unpaired two-tailed t-test, with similar variance between the groups being compared.

All animal experiments were performed in compliance with relevant ethical regulations, and were approved by the IACUC at Harvard Medical School.

SUPPLEMENTAL REFERENCES

Avraham, O., Zisman, S., Hadas, Y., Vald, L., and Klar, A. (2010). Deciphering axonal pathways of genetically defined groups of neurons in the chick neural tube utilizing in ovo electroporation, Journal of visualized experiments : JoVE.

Bailey, T., Boden, M., Buske, F., Frith, M., Grant, C., Clementi, L., Ren, J., Li, W., and Noble, W. (2009). MEME SUITE: tools for motif discovery and searching, Nucleic acids research *37*, 8.

Gumy, L.F., Yeo, G.S., Tung, Y.C., Zivraj, K.H., Willis, D., Coppola, G., Lam, B.Y., Twiss, J.L., Holt, C.E., and Fawcett, J.W. (2011). Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization, RNA *17*, 85-98.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M.*, et al.* (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP, Cell *141*, 129-141.

Helms, A., and Johnson, J. (1998). Progenitors of dorsal commissural interneurons are defined by MATH1 expression, Development (Cambridge, England) *125*, 919-928.

Huang, D.W., Sherman, B., and Lempicki, R. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nature protocols *4*, 44-57.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol *10*, R25.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X.*, et al.* (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing, Nature *456*, 464-469.

Lumpkin, E.A., Collisson, T., Parab, P., Omer-Abdalla, A., Haeberle, H., Chen, P., Doetzlhofer, A., White, P., Groves, A., Segil, N.*, et al.* (2003). Math1-driven GFP expression in the developing nervous system of transgenic mice, Gene Expr Patterns *3*, 389-395.

Preitner, N., Quan, J., Nowakowski, D.W., Hancock, M.L., Shi, J., Tcherkezian, J., Young-Pearse, T.L., and Flanagan, J.G. (2014). APC is an RNA-binding protein, and its interactome provides a link to neural development and microtubule assembly, Cell *158*, 368-382.

Reeber, S.L., Sakai, N., Nakada, Y., Dumas, J., Dobrenis, K., Johnson, J.E., and Kaprielian, Z. (2008). Manipulating Robo expression in vivo perturbs commissural axon pathfinding in the chick spinal cord, J Neurosci *28*, 8698-8708.

Tompa, M. (1999). An exact method for finding short motifs in sequences, with application to the ribosome binding site problem, Proceedings / International Conference on Intelligent Systems for Molecular Biology ; ISMB International Conference on Intelligent Systems for Molecular Biology, 262-271.

Trapnell, C., Williams, B., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M., Salzberg, S., Wold, B., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation, Nature biotechnology *28*, 511-515.

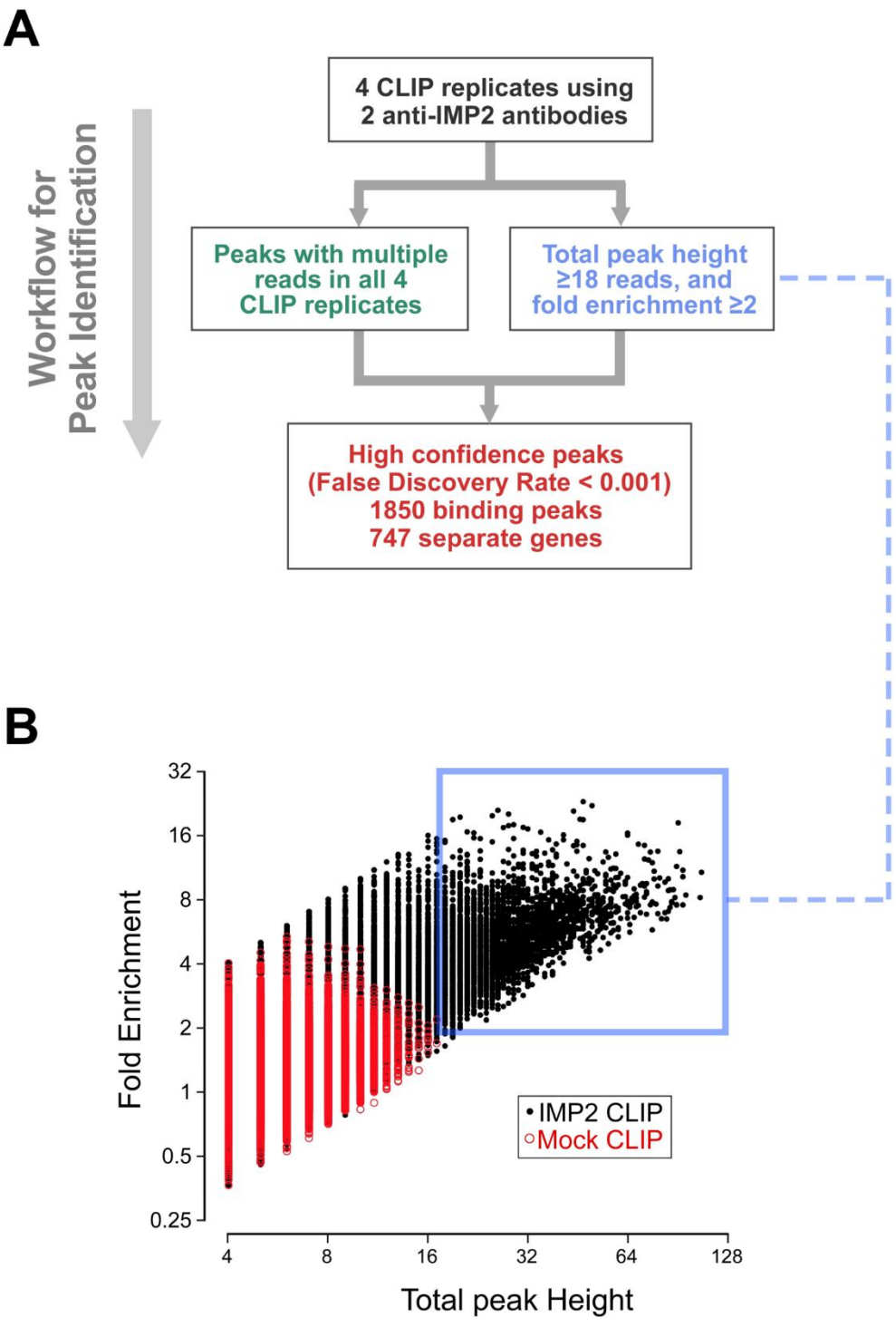

Figure S1. Criteria for selection of high confidence mRNA targets.

(A) Workflow of the analysis. Left box in the flowchart: Only clusters with multiple overlapping reads in all four CLIP replicates were considered: a minimum of four reads in at least three of the four replicates, and a minimum of two reads in the fourth replicate. Right box in the flowchart: For each cluster, reads from all four replicates were pooled and a peak height corresponding to the position of maximum number of overlapping reads within the cluster was calculated. Only peaks with a minimum height of 18, and minimum 2-fold enrichment were further considered based on comparison to the estimated background. **(B)** To estimate background, we generated random samples (mock CLIP), each matching in size the IMP2 CLIP dataset, from control brain mRNA-seq data. A median background number of reads was calculated for each position on the genome, based on repeated sampling of mock CLIP datasets. Enrichment over background (fold change) was then calculated for each cluster, by comparing peak height in IMP2 HITS-CLIP with the median read number of mock CLIP at the corresponding genomic location. Peak height and fold change were then plotted for IMP2 HITS-CLIP as well as for a mock HITS-CLIP sample (lower panel). From this analysis, it was found that choosing a peak height of at least 18, and fold enrichment of at least 2, reduces the false discovery rate to <0.001.

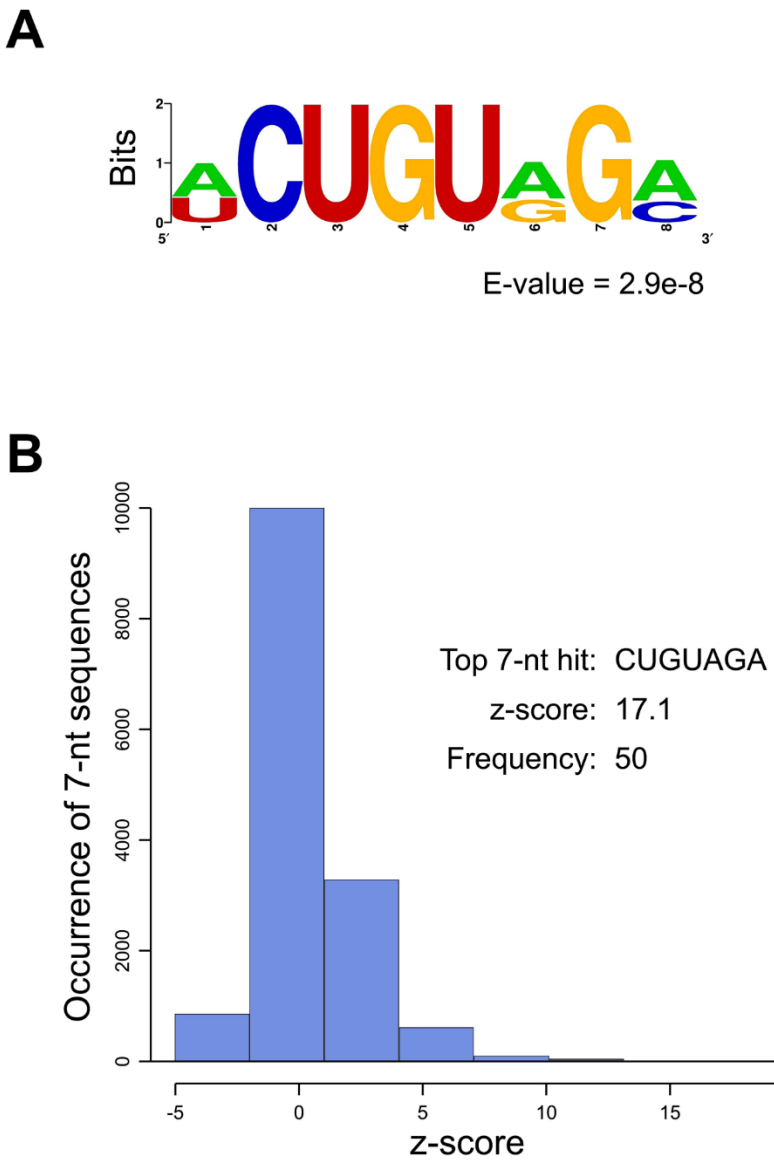

Figure S2. Identification of enriched sequence motif at IMP2 binding sites.

(A) Sequence motif identified by MEME software. (**B)** A preferential binding sequence was also identified using Z-score statistics. Histogram of z-scores indicates the enrichment of 7-nt sequences in all IMP2 CLIP peaks, and the most enriched 7-nt sequence is shown. The two approaches are consistent with one another in identifying an enriched motif. A previous study by Hafner et al. included IMP2 in a validation of the PAR-CLIP approach in cultured HEK-293 cells (Hafner et al., 2010). The use of HITS-CLIP here enabled us to investigate IMP2 mRNA targets in native brain, to identify targets relevant to neural development. Also, this approach assesses native IMP2 (endogenous levels and isoforms), whereas the previous approach used recombinant IMP2 in an overexpression system which can result in formation of aberrant complexes of IMP proteins (Bell et al., 2013).

Disease Associations

Disease clusters	Disease terms	*p*
Psychiatric	Schizophrenia	2e-11
Neurodegeneration	Amyotrophic lateral sclerosis Huntington's disease Alzheimer's disease Parkinson's disease	4e-9 3e-8 4e-6 7e-3
Brain cancer	Glioblastoma Astrocytoma Glioma	3e-8 4e-8 1e-7
Neurodevelopment	Type 1 Lissencephaly Mental retardation	2e-6 2e-4
Other	Gliosis	4e-3

Figure S3: IMP2 target mRNAs encode proteins linked to neurologic diseases.

Associated disease terms were identified by IPA Disease and Disorder Analysis (see Materials and Methods).

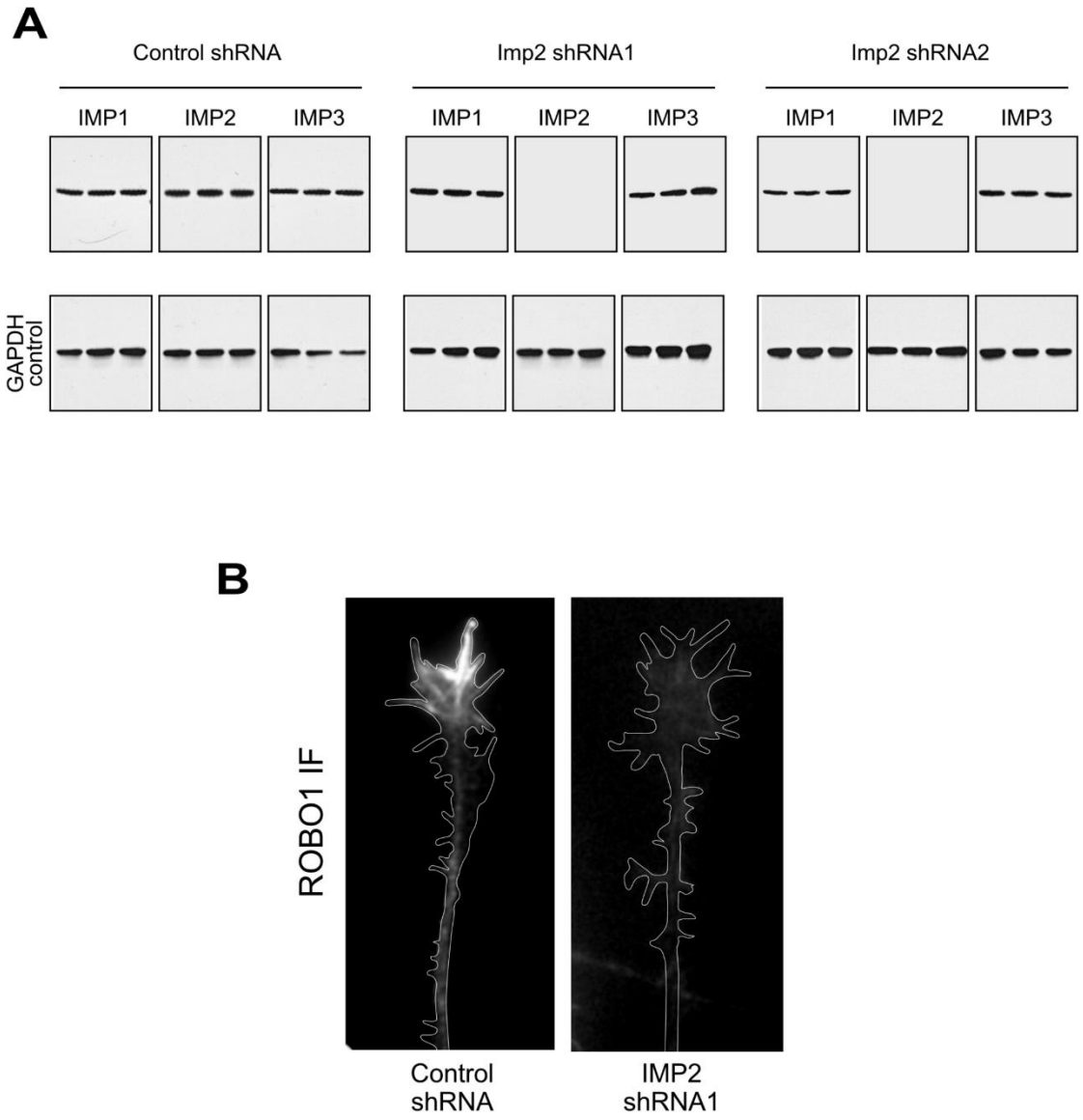

Figure S4: Confirmation of protein knockdown by IMP2 shRNAs, and reduced expression of the IMP2 target *Robo1* on axons. **(A)** Western blot, confirming that two independent IMP2 shRNAs strongly knocked down protein expression from IMP2 mRNA, compared with controls including IMP1 and IMP3 mRNAs, and the metabolic housekeeping enzyme GAPDH. **(B)** ROBO1 immunofluorescence was reduced on commissural axons after in vivo treatment with IMP2 shRNA compared with control shRNA. To visualize individual axons, immunolabeling was performed after dissociating the neurons and placing them in culture. A white line indicates the outer boundary of the growth cone and distal axon shaft based on phalloidin-rhodamine labeling.

Table S1

Click here to Download Table S1

Table S2

Click here to Download Table S2