Supplemental Information Inventory

1. Figure S1, **Identification of MOV10-associated mRNAs**, related to Figure 1
2. Figure S2, **FMRP knockdown and overexpression has no effect on target mRNA levels,** related to Figure 2
3. Figure S3, **Characterization of the AGO and MOV10 binding sites** related to Figure 3
4. Figure S4, **MOV10 modulates regulation in the 3'UTR and mediates microRNA-mediated protein expression** related to Figure 4
5. Figure S5, **FMRP regulates expression of endogenous proteins** related to Figure 5
6. Figure S6, **Identification of miR-328 as regulating MAZ expression in the 3'UTR**, related to Figure 6
7. Supplemental legends
8. Supplemental Tables S1-4 are excel files attached separately
    a. Table S1. Gene ids from RNA-IP, related to Fig. S1
    b. Table S2. iCLIP clusters, related to Fig. 2
    c. Table S3. RNA seq, related to Fig. 4
    d. Table S4. iCLIP clusters found in both C5 and C7, related to Figure 4
9. Supplemental Tables S5-6 are included here
    a. Table S5. Primers used in qRT-PCR, related to Figure 2 and 5 and the PCR reactions described in the Supplemental Experimental Procedures under 'qRT-PCR'.
    b. Table S6. Primers used for 3'UTRs, related to Figures 5 and 6 and the for the section in the Experimental Procedures 'Luciferase Reporter Constructs'

10. Supplemental Experimental Procedures
11. Supplemental References

S1. Identification of MOV10-associated mRNAs, related to Figure 1

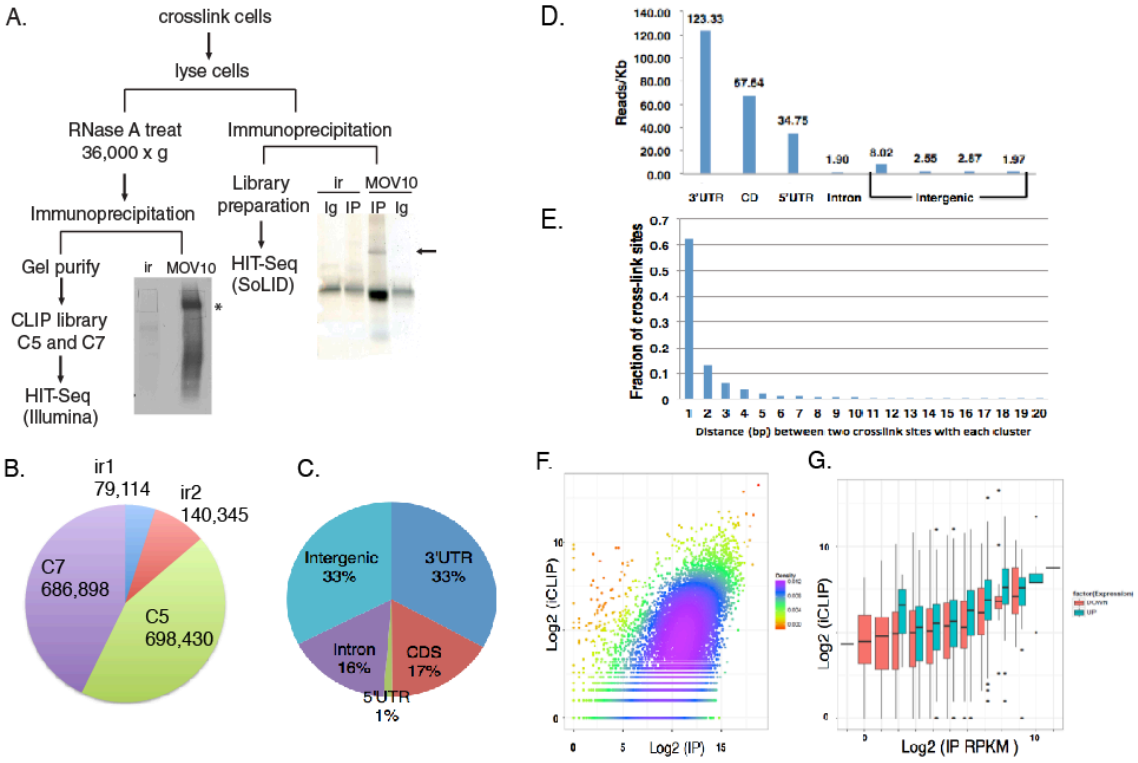


Figure S2, **FMRP knockdown and overexpression has no effect on target mRNA levels,** related to Figure 2

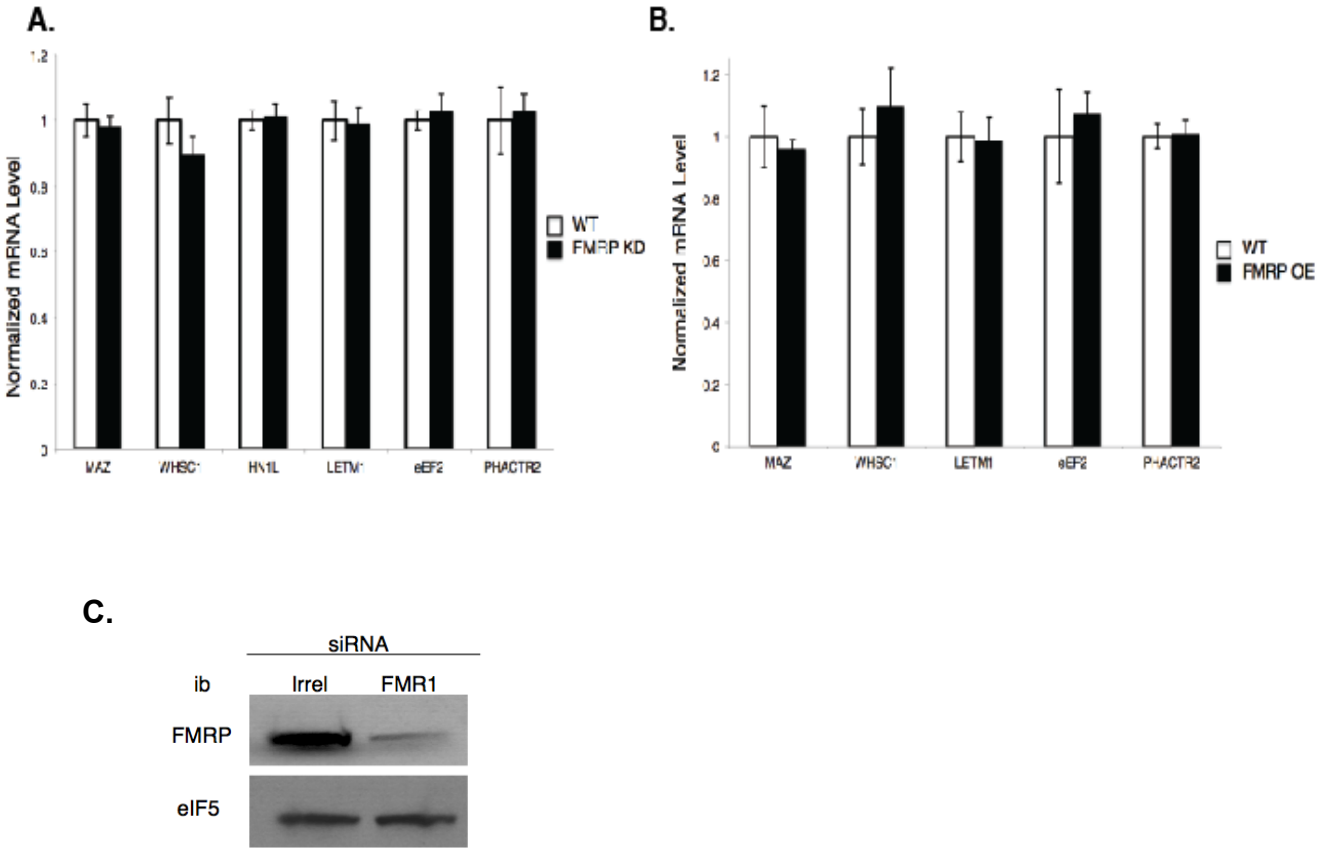
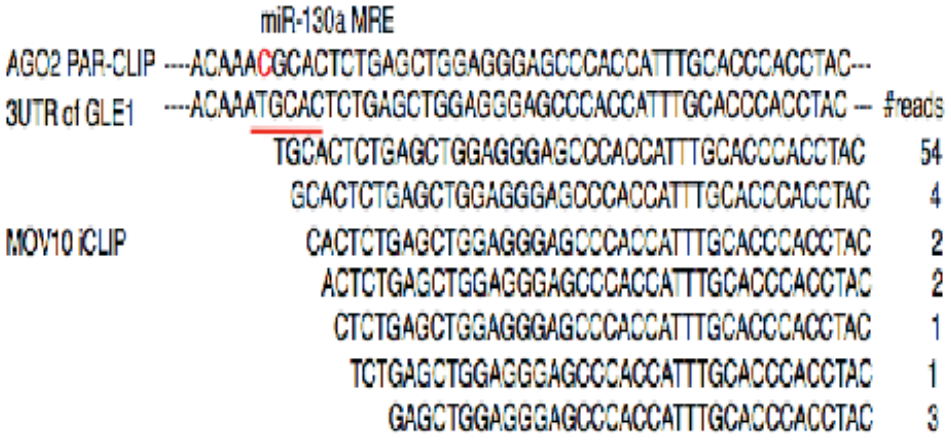
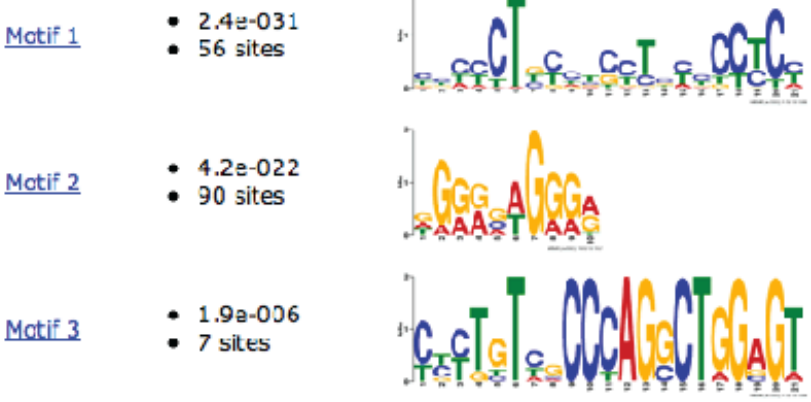


Figure S3, **Characterization of the AGO and MOV10 binding sites** related to Figure 3

A.



B.



D.

Figure S4, **MOV10 modulates regulation in the 3'UTR and mediates microRNA-mediated protein expression** related to Figure 4

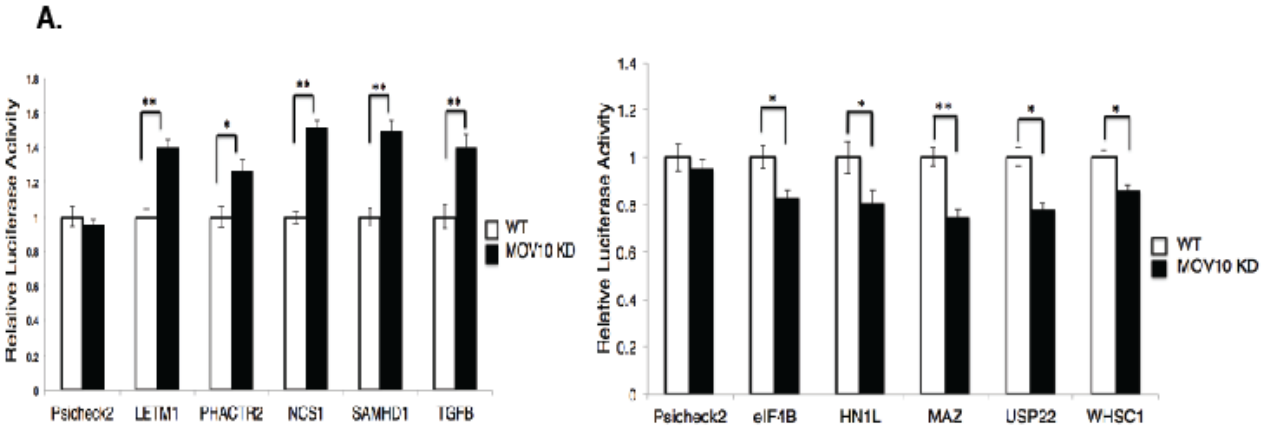


Figure S5, **FMRP regulates expression of endogenous proteins** related to Figure 5

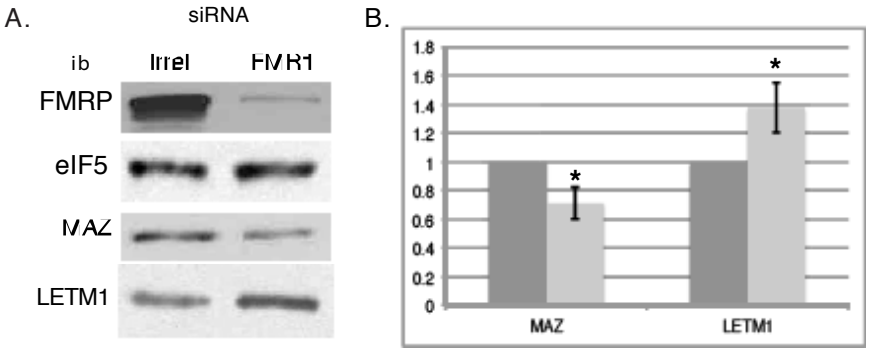
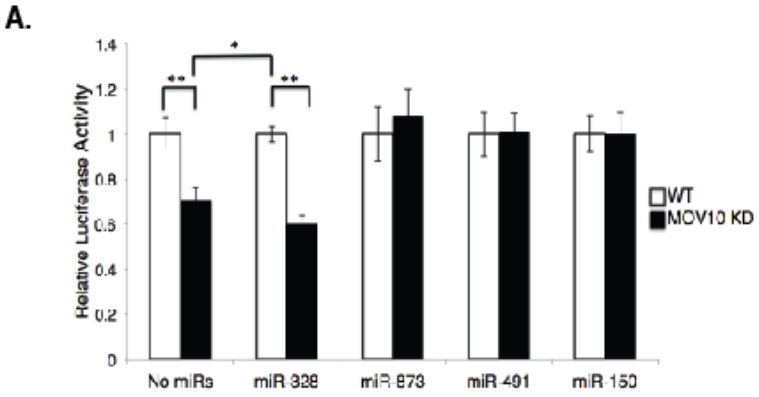


Figure S6, **Identification of miR-328 as regulating MAZ expression in the 3'UTR**, related to Figure 6



Legend for Figures

Fig. S1. Identification of MOV10-associated mRNAs.

A) Schematic of library preparations of MOV10 IPs (RNA-IP) (right) and iCLIP (left). Right, silver stain of UV-cross-linked IPs (IP), MOV10 (arrow); Left, autoradiography of [^{32}P]-labeled MOV10-RNA complex (star), C7 iCLIP library. Ir- irrelevant IP, Ig- IP'ing antibody (Ab)**. B)** Distribution of mapped reads from the iCLIP libraries (C5 and C7) and irrelevant IPs (ir) distinguished by barcodes (see Methods); **C)** Distribution of genome mapped reads from C5 and C7 among gene regions. **D)** Distribution of normalized reads (RPKM) plotted for the gene regions indicated on the x-axis. Numbers on indicate the average coverage for each gene region. **E)** Frequency of nucleotide distance (x-axis) observed between two cross-link sites in clusters identified in MOV10 iCLIP experiments. **F)** Gene

density map showing the correlation between iCLIP (y-axis) and MOV10 RNA-IP (x-axis) transcript read number. Of the 2117 genes found to be enriched in the MOVq0 RNA-IP, 1850 (87.4%) were identified to have significant binding in the iCLIP data; **G)** Boxplot of iCLIP reads (y axis) and RPKM of MOV10 RNA-IP (x-axis) classified by both its RPKMs and its enrichment relative to total RNA in HEK293 cells ('up' – the RNA-IP was enriched relative to total RNA; 'down'-the RNA-IP was diminished relative to total RNA). See Tables S1 and 2.

S2. MOV10 recognizes a subset of RNAs bound by FMRP.

A. Real-time PCR quantification of mRNA in WT and FMR1 knockdown HEK293 cells (n=3). **B)** Real-time PCR quantification of mRNA in WT and FMR1 over-expressed HEK293 cells (n=3). **C)** Western Blot showing FMRP knockdown in HEK293 prior to MOV10 immunoprecipitation (79.5% knockdown, $p<0.01$ by student t-test)

S3. Identification of MOV10-associated mRNAs.

A) Cross-link site of AGO (top red) within the 3'UTR of GLE1 (seed sequence underlined) for mir-130a. Number of MOV10 iCLIP reads and their alignment to the start of the cross-link site is shown. **B)** MEME results of 3'UTR MOV10 iCLIP sites.

S4. MOV10 modulates regulation in the 3'UTR and mediates microRNA-mediated protein expression.

A) The 3'UTR of MOV10 target mRNAs were sub-cloned into a luciferase expression reporter and luciferase was measured in the presence and absence of MOV10. Luciferase expression validated the protein levels observed in Western Blot analysis, with a subset of mRNA levels increasing on MOV10 knockdown (Left panel) and a subset of mRNA levels decreasing on MOV10 knockdown (right panel). **B)** The 3'UTR of MAZ (WT) was sub-cloned into a luciferase reporter and co-transfected with miRNAs, indicated in the presence (white bars) or absence of MOV10 (black bars). Error bars represent standard deviation, p-values obtained by student t-test, * $p < 0.05$, ** $p < 0.01$

S5. FMRP regulates expression of endogenous proteins. HEK293T cells were transfected daily with siRNAs to an irrelevant control (irrel) or FMR1 for 48-72 hours. 12.5 ug of extract (FMRP, LETM1 and eIF5) or 25 ug of extract (MAZ) was analyzed on a 7.5% gel, transferred to PVDF and immunoblotted (ib) with the antibodies indicated. Average knockdown of FMRP in 6 experiments was 80.5%. In the absence of FMRP, MAZ was reduced approximately 30% ($p=0.02$, $n=6$, one-tailed t-test) and LETM1 was increased approximately 38% ($p=0.04$, $n=6$, one-tailed t-test). eIF5 is the loading control

S6. MOV10 modulates miRNA-mediated translational regulation.

A) The 3'UTR of MAZ (WT) was sub-cloned into a luciferase reporter and co-transfected with miRNAs, indicated in the presence (white bars) or absence of MOV10 (black bars). It was determined that miR-328 was responsible for

MOV10-mediated translational suppression. Error bars represent standard deviation, p-values obtained by student t-test, * p <0.05, ** p <0.01

Table S5. Primers used in qRT-PCR, related to Figure 2 and 5 and the PCR reactions described in the Supplemental Experimental Procedures under 'qRT-PCR'.

Gene	Sequence (5' to 3')
CALM3- Fwd	CTTCGACAAGGATGGAGATGG
CALM3- Rev	AACTCTGGGAAGTCAATGGTC
CERS- Fwd	CCATACAGAGCACTATCCCTTCCCAG
CERS- Rev	CCTTTCGCTTGACATCAGAGGCAATG
eEF2- Fwd	ACATTCTCACCGACATCACC
eEF2- Rev	GAACATCAAACCGCACACC
GNB2L1- Fwd	AATACTGTGGGTGTCTGCAAG
GNB2L1- Rev	TTAGCCAGATTCCACACCTTG
GAPDH- Fwd	CTTTGTCAAGCTATTTCCT
GAPDH- Rev	CTTGCTCAGTGTCCTTGC
MAZ- Fwd	GCTTCTCCCGGCCGGAT
MAZ-Rev	GAAAGCTGCCTCACATTTCTCACATTTG
WHSC1	GGAAGAGTCTTCAAAAACGCACTGCAAG
WHSC1	CCGTAAGGCTTATTCACCTTGATGTGC

Table S6. Primers used for 3'UTRs, related to Figures 5 and 6 and the for the section in the Experimental Procedures 'Luciferase Reporter Constructs'

*Gene	Sequence (5' to 3')
eIF4B-Fwd	CATGCTCGAGTGAGGGAGAAGATTATGCCGAATAGACCT
eIF4B-Rev	CATGATTGATGCGGCCGCCTTCAGAGATGGTAGGAGGAGGAGGA
HN1L-Fwd	CATGCTCGAGGGGAGGCAAATCCAGCATCTCC
HN1L-Rev	CATGATTGATGCGGCCGCCACAGGGTCTTGCTCTGTGGTC
LETM1-Fwd	CATGCTCGAGCAGAGGTGAAGAGCTAGAACCACTGG
LETM1-Rev	ATTGATTGTTGCGGCCGCCGATTTCTACAAAAGGTGCTTAACACG GC
NCS1-Fwd	CATGCTCGAGATGCCTGGGAACCACTCACC
NCS1-Rev	CATGATTGATGCGGCCGCGGCGAGGACAAGTCCAGGAAG
USP22-Fwd	CATGCTCGAGGTAGCCTTATCTGCAGCTGGTCAGA
USP22-Rev	CATGATTGATGCGGCCGCTCTTCGGCTGCTGTGTCAATAAAAGGA
WHSC1-Fwd	GCATCTCGAGGCAAATAGCGCCAGGCGG
WHSC1-Rev	GCATTCAGTTGCGGCCGCCCATCGGAGGCATCAGTGCC
*MAZ	All constructs were purchased from GeneWiz, Inc. This includes WT MAZ 3'UTR, and the MAZ-Δ328 mutant.

Supplemental Methods

RNA sedimentation gradient analysis

Two adult male C57/Bl6 mouse brains or 10^8 HEK293F cells were harvested and lysed in buffer containing 150mM NaCl, 50mM Tris, 0.5% Triton X-100. Cell lysates were spun for 10 min at 1000xg. The supernatant (S1) was then spun again at 10,000xg for 10 min to obtain S2. 15-30% linear sucrose gradients were prepared with buffer containing 100mM KCl, 20mM Tris (pH 7.5), 5mM $MgCl_2$. Postnuclear supernatant S2 was overlaid on the gradient and centrifuged at 40,000 rpm for 30 minutes at 4°C. Gradients were fractionated into 500 uL fractions by bottom displacement. Odd numbered fractions were TCA precipitated and analyzed by western blot for FMRP or MOV10, as indicated:

affinity purified rabbit antibody to MOV10 (Bethyl A301-571A) at 1:1000, anti-FMRP (hybridoma supernatant (Devys et al., 1993) HRP-conjugated anti rabbit and anti mouse antibodies from Amersham and Jackson laboratories, respectively. 7G1 (Brown et al., 2001) was used for immunoprecipitating FMRP from brain and anti-FMRP (Abcam ab17722) for immunoprecipitating human FMRP from HEK293 cells.

Co-IP experiments and general procedure for western blotting.

Cultured L-M(TK-) cells (Ceman et al., 1999) were lysed in ice cold lysis buffer (0.15 M NaCl, 50 mM Tris, 0.5% Triton X-100, 30mM EDTA). Postnuclear supernatants were immunoprecipitated with either anti-FLAG coupled beads (Sigma) or anti-myc coupled beads (Sigma). After washing, samples for immunoblotting were prepared in 1x sample buffer, resolved by SDS-PAGE and analyzed by western blotting. Briefly, membranes were blocked using 5% non-fat dry milk in PBS containing 1% TWEEN-20 for 1 hour at room temperature. Primary antibody was applied for 1 hr at room temperature or overnight at 4°C followed by a brief wash in 1% non-fat milk PBS containing 1% TWEEN-20 wash buffer. HRP conjugated secondary antibody was applied at 1:5000 dilution for 1 hour at room temperature and washed 4x15 minutes using wash buffer. HRP signal was detected using ECL and exposed to film. The following antibodies were used for immunoblot: anti-MOV10 ab (Bethyl A301-571A) at 1:1000, anti-FMRP (hybridoma supernatant (Devys et al., 1993)) for western blots (1/10), anti-FMRP (Abcam ab17722) at 1:1000, anti-Phactr2 (Abcam ab85262) at 1:50, anti-

MAZ (Santa Cruz sc-28745) at 1:200, anti SAMHD1 (Thermo Scientific PA5-27898) at 1:2000, anti TGFb1 (Sigma AV37156-100UG) at 1.25ug/uL, anti-eIF5 (Santa Cruz) at 1:10,000. HRP-conjugated anti rabbit and anti mouse antibodies from Amersham and Jackson laboratories, respectively.

For testing protein-protein interaction, recombinant FMRP and MOV10 were prepared as described in (Ascano et.al) from SF9 cells using Nickel NTA columns. Recombinant FMRP and MOV10 were incubated with RNAse A on ice overnight and immunoprecipitated with either the anti-MOV10 antibody or an irrelevant rabbit affinity purified antibody bound to Protein A Dynabeads for 10 min in the cold. Immunoblotting was done using either anti-MOV10 or anti-FMRP as described above.

Polysome analysis

HEK293F cells were treated for 15 min with cycloheximide (100ug/mL), pelleted and washed in PBS followed by lysis in buffer containing 100 mMKCl, 20mM Tris (pH 7.5), 5mM $MgCl_2$ supplemented with 0.3% Igepal CA-630 and pelleted at 20,000xg for 15 minutes. Linear sucrose gradients (15-45%) were prepared with buffer containing 100 mM KCl, 20mM Tris (pH 7.5), 5mM $MgCl_2$ using a gradient maker (BioComp). Postnuclear supernatants were overlaid on the gradient and centrifuged for 75 min at 35,000 rpm at 4°C. Gradients were fractionated into 1mL fractions by bottom displacement. Odd numbered fractions were TCA precipitated from 400-500 uL fractions and analyzed by western blot on 4-20% gradient SDS-PAGE gels.

Identification of MOV10-associated RNAs by cross-linking IP (RNA-IP)

3.2 X 10^9 HEK293F cells were UV-cross linked three times (Stratalinker), lysed in 0.3 M lysis buffer (50 mM Tris 7.5, 300 mM NaCl, 30 mM EDTA, 0.5% triton), cleared by spinning at 30,000 rpms for 35 minutes, precleared with 2X volume of protein A sepharose, immunoprecipitated with a rabbit antibody to an irrelevant antigen, immunoprecipitated with the MOV10 antibody (Bethyl). For all IPs, 30 ug of rabbit antibody was coupled to 100 ul of 50% Protein A sepharose/10^9 cells). Antigen-antibody-coupled beads were washed 1x in 0.3 M lysis buffer, 1x in wash buffer (Ule et al., 2005; Ule et al., 2003), treated with 40 units of RQ1 DNAse (Promega) and 80 units of DNAse (NEB) at 37 degrees for 15 minutes, washed 2x in high salt, 1x in low salt, 5% of the IP was removed for silver stain. The rest was treated with 1 mg of Proteinase K for 30 min at 37 degrees, the Protein A sepharose was removed by a 14,000 rpm spin for 5 minutes and the eluted RNA was isolated by phenol/chloroform extraction followed by ethanol precipitation. 6.6% of the sample was analyzed by Bioanalyzer and subjected to library preparation and sequencing analysis on the LifeScience SOLiD 4 platform and pair-end protocol was used. More than 300M mapped reads were obtained from the RNA-IP and the total RNA. The reads were mapped to the human genome (hg19) by used Bioscope 1.3. RefSeq genes were used and EdgeR was used to identify the MOV10 associated RNAs.

iCLIP analyses

HEK293F cells were grown to ~ $4x10^8$ and UV-cross-linked three times (Stratalinker) with mixing between treatments. A published CLIP protocol was followed (Konig et al., 2010, 2011) with the following exceptions: phosphatase treatment was performed using shrimp alkaline phosphatase (SAP), the irrelevant IP was performed with a rabbit affinity purified antibody to zebra finch FMRP, described in (Winograd et al., 2008), which does not cross-react with mammalian FMRP. The MOV10 IPs were performed with a rabbit affinity purified antibody (Bethyl).

High-throughput sequencing and bioinformatic analysis of iCLIP data

MOV10 iCLIP libraries were sequenced by the UIUC sequencing core facility using the Illumina HiSeq2000 platform. Cutadapt (Martin, 2011) was used to trim adaptors and low quality bases (<20) at the 3` ends of reads. Fastx_collapser from fastx-toolkit was used to collapse reads and filter duplicates resulting from the PCR. Reads were separated into four samples by the barcodes at the 5` ends of reads. After trimming the barcodes, reads with 18 and more were kept and aligned to the human genome (hg19). A similar amount of mapped reads (76%~79%) were obtained using different aligners including Bowtie2, Tmap1 and Novoalign. We used the data mapping by Novoalign with “x 4 g 20” for downstream analysis because it achieved the most mapped reads (79%). RseQC (Wang et al., 2012) was used to evaluate the quality of sequencing and mapping reads. Good alignments with alignment score not less than 10 were kept for downstream analysis. Pyicoclip, a modified False Discovery Rate

method implemented in Pyicos tools (Althammer et al., 2011), was used to call peaks of iCLIP cDNAs within genes. Clusters were defined as significant genomic regions in which multiple reads aligned. In C5 and C7, 32,331 clusters were identified within 8986 genes, with 15,475 of these clusters found in 3'UTRs. Cross-link sites were identified before each mapped read based on iCLIP protocol (Konig et al., 2010, 2011). The cross-link sites with the most mapped reads were selected for MRE analysis. 10K random sites with random genes were selected for control analysis. FMRP binding sites predicted by PARalyzer (Genome biology 2011, 12:R79) were downloaded from GSM97615, in which the cross-link sites were usually in the middle of clusters. RefSeq gene models were used.

To identify MOV10 iCLIP targets that were present at least once in both the C5 and C7 libraries and that were used in the RNA-seq and protein analysis, the following analysis was performed on the C5 and C7 libraries. Raw data was trimmed using the FASTX-toolkit; quality trimming from the 3' end for quality scores lower than 20 and adaptor trimming. Reads longer than 30 nucleotides were then de-duplicated and separated into C5 and C7 bins, for IP and irrelevant in both cases, using the barcodes. These de-duplicated tags were aligned to the human genome (hg19) (2 mismatches) using Novoalign (V2.08.01) and uniquely mapping tags were extracted. Of these, any overlapping tags were then consolidated into longer regions using BEDTools (v2.16.2). For C5 and C7 respectively, any region in the IP bin that overlapped with regions in the irrelevant bin was discarded to identify the MOV10 iCLIP regions. The genomic coordinates

of these regions were used to categorize the features (genes, exons, UTRs etc.) that coincided with the iCLIP regions. The high-specificity iCLIP regions were deemed to be those that were present in both the C5 and C7 experiments. C5 contained 27,126 uniquely mapped de-duplicated (compressed) tags and mapped to 9,436 iCLIP regions. C7 contained 28,136 uniquely mapped de-duplicated (compressed) tags and mapped to 2,186 iCLIP regions. The average iCLIP target size was 38 nucleotides. Where not specified, BEDTools (v2.16.2) and in-house Perl scripts were used for analysis.

Venn Diagram

The 23,284 mouse genes [UCSC mm9, used in (Darnell et al., 2011)] and 23,373 human genes (UCSC hg19) share 15,492 symbols in common, thus, 15,492 was used as the common background to make the comparison between the murine brain FMRP targets and the HEK293 FMRP and MOV10 targets. Of the 842 FMRP-bound murine genes reported (Darnell et al., 2011), 781 were in the common background and were used for the subsequent analysis. Of the 5676 genes bound by FMRP (using both isoforms) in a 5'UTR, exon or 3'UTR in HEK293 (Ascano et al., 2012), 5053 were in the common background and were used in the subsequent analysis. Of the 658 genes with one or more MOV10 targets (in both libraries) in a 3'UTR, 5'UTR or exon, 605 were in the common background and were used in the subsequent analysis. To obtain statistics on the 3way-venn diagram overlaps, a permutation approach was used by randomly selecting gene sets of 791, 5053 and 606 from the 15,492 background and the

numbers of genes in common were counted; this was repeated 50,000 times to build up distributions of overlap values, which were used to empirically derive p-values for our observed overlaps. While only using genes with identical symbols between human and mouse will miss some true homologs, it simplifies the comparison by removing any many-to-one or many-to-many relationships that would be impossible to correctly represent in a Venn diagram. Additionally, we computed Fisher's exact tests between the pairwise overlap of the HEK293 FMRP and the HEK293 MOV10 using both the full 23,373 human genes and the common 15,492 as the background; the full human background had a p = 6.67e-139 whereas the common background had a p = 3.82e-86, so the common background is actually more conservative. All of the overlaps except '9' have p-values = 0, meaning that the number of overlaps was not seen in 50,000 samplings. The 9 genes in the FMRP brain and MOV10 lists but not in the HEK293 is actually significantly FEWER than expected by chance (p = 0.003), but the 9+74 = 83 overlap in a pairwise comparison had a p-value = 0; the unsurprising interpretation is that genes sharing murine FMRP and human MOV10 sites almost always have a human FMRP site as well.

Isolating MOV10-RNA complexes from WT and FMR1-knockout brains.

Total RNA: Mouse Brain mRNA was isolated with Trizol (Life Technology) following manufacturers instructions, treated with DNaseI (Biolabs), recombinant RNasin (Promega) and reverse transcribed with Superscript III (Invitrogen). qRT-PCR was performed with iQ SYBR Green Supermix (Bio-Rad) using a

StepOnePlus RT PCR machine (Applied Biosystems) with gene specific primers, (Table S5) Relative levels of mRNA were determined using GAPDH concentration for normalization. Statistical significance was determined using a Student t-Test on samples performed in triplicate.

Brain IP: Brains were harvested from adult male mice (2 months old) that were either WT (FVB.129) or FMR1 knockout mice (Jackson Laboratory). Briefly, each brain was triturated in HBSS and then UV-cross linked 3 times. Cells were lysed in 0.3 M NaCl lysis buffer (50 mM Tris 7.5, 300 mM NaCl, 30 mM EDTA, 0.5% triton), cleared by ultracentrifugation (30,000 rpms for 35 min at 4°C) and sequentially immunoprecipitated with an irrelevant rabbit polyclonal antibody followed by IP with the MOV10 antibody. Both IPs were washed in 0.3 M lysis buffer, 2x with wash buffer, treated with 500 units of DNase, washed 2x with high salt buffer (50 mM Tris, 1 M NaCl 1mM EDTA 1% NP40, 0.1% SDS, 0.5% sodium deoxycholate). To isolate associated RNA, the IPs were treated with proteinaseK followed by phenol/chloroform extraction and ethanol precipitation. First strand synthesis used oligo dT and M-MuLV followed by PCR with the described primers. The immunoprecipitated mRNA was quantified by RT-pPCR in the same manner as the total brain mRNA (see above). Gnb2l1 was also quantified to determine the effect of non-specific antibody binding during the IP process. See Table S6 for the list of PCR primers.

Mean free energy calculations

Mean free energies were calculated in the 550 nucleotides flanking the 453 MOV10 binding sites present in both libraries in the 3'UTR using the program RNAfold (part of the Vienna RNA Package, version 1.8.5 (Hofacker et al., 1994)). The following parameters were used:-p0 -d0 –noLP. These calculations were performed for 50 nucleotide sliding windows across the 550 nucleotide sequence (step size of 1). "The free energy of folding was calculated for 50 nucleotide windows every 5 nucleotides within these region using the Vienna RNAfold package (v. 2.0.3b) (Lorenz et al., 2011). The GC content and mononucleotide content were also calculated for these 50 nucleotide windows. The resulting profiles were averaged over all MOV10 bound sites. The free energy of folding was calculated at each relative location was averaged for plotting.

RNA capture experiments

Flag-FMRP in pGEMT-Easy vector was *in vitro* transcribed and translated in an RRL kit (Promega) as per manufacturer's instructions. Myc-MOV10 was transiently transfected into HEK293F cells and immunoprecipitated using anti-Myc conjugated beads (Sigma). Protein was peptide eluted with Myc peptide. 80ng of protein was used in binding experiments as described in (Stetler et al., 2006). Briefly, protein and biotinylated sc1 or sc1 mutant RNA were incubated at 30°C and captured on streptavidin coated dynabeads, washed, eluted in sample buffer, resolved on a 7.5% polyacrylamide gel and immunoblotted as described above.

Isolating total RNA from HEK293 for RNA-seq

HEK293F cells were transfected three times at 24 hour intervals with irrelevant (IR) or MOV10 siRNAs (KD). Only experiments in which there was >80% knockdown of MOV10 were analyzed. Overexpression libraries (OE) were generated by transfecting 500 ng of myc-MOV10 for 4 hours using PEI. Cells were lysed after 72 hours for KD experiments and after 24 hours for overexpression experiments RNA was isolated with Trizol (Life Technology) following manufacturer's instructions. Isolated RNA was resuspended in nuclease free water and phenol chloroform extracted. 1 ug of RNA sample was run on a 1% agarose gel to resolve ribosomal subunits and 5 ug was submitted for RNA sequencing analysis. Library preparation involves mRNA selection with oligodT beads, then chemical fragmentation, annealing of random hexamer and first strand synthesis with SuperScript II, second strand, 3'-tailing, adaptor ligation and PCR amplification. The libraries were sequenced using the Illumina HiSeq2000 platform.

RNA-Seq alignment and analysis

Raw FASTQ data was quality-trimmed from the 3' end using the program Trimmomatic (v 0.22; Lohse et al. 2012), using a minimal PHRED quality score of 20 and a minimal length of 30. Sequences were then aligned using TopHat v. 2.0.8 (Trapnell et al., 2009) and Bowtie 2.1.0 (Langmead and Salzberg, 2012) using the following parameters:

tophat2 --coverage-search -p 8 -N 7 --read-edit-dist 7 --library-type fr-firststrand -o <SAMPLE_NAME> <REFERENCE_DATABASE> <TRIMMED_FASTQ>

The genome sequence index (<REFERENCE_DATABASE>) was hg19 from UCSC (http://hgdownload.soe.ucsc.edu/downloads.html#human). The BAM alignments were first changed to SAM format using samtools (v 0.1.18, Li et al. 2009) and then raw read counts were tabulated for each sample at the gene level using the GTF gene model file <GTF_FILE> for hg19 from UCSC and htseq-count, from HTSeq v0.5.3p9 (http://www-huber.embl.de/users/anders/HTSeq/doc/index.html) using the default "exon" feature type, "gene_id" attribute and the following parameters:

htseq-count -s reverse -m intersection-nonempty - <GTF_FILE>

The raw read counts were input into R 3.0.0 (Team, 2013)a for data pre-processing and statistical analysis using packages from Biocondutor (Gentleman RC, 2004) as indicated below. Initial quality control analysis indicated that one of the three KD replicates was an outlier and it was removed from the analysis. Genes without 1 Count Per Million (CPM) mapped reads in at least one of the 8 samples were filtered out due to unreliable data in any sample; 14,615 of the 23,368 genes passed this filter and were analyzed using edgeR 3.2.1 (Robinson et al., 2010). The raw count values were used in a negative bionomial statistical model that accounted for the total library size for each sample and an extra TMM normalization factor (Robinson and Oshlack, 2010) for any biases due to changes in total RNA composition of the samples. Pairwise comparison for KD

vs. IR and OE vs. IR were pulled from the model and separately adjusted for multiple testing using the False Discovery Rate method (Benjamini and Hochberg, 1995); genes with FDR $p < 0.05$ were considered significantly different. For sample clustering and heatmaps, comparable expression values were generated from the read counts using edgeR's modified log2 CPM values. Additional annotation information (gene names, descriptions, Gene Ontology terms and pathways) was obtained from Bioconductor's org.Hs.eg.db package (v. 2.9.0; based on NCBI's Entrez database) using the transcript IDs provided in the GTF gene model file instead of the gene symbols.

To compare with the CLIP data, the genomic regions showing CLIP targets for MOV10 were annotated to the closest gene using gene symbols for hg19 from Ensembl and assigned to a region of the gene (e.g., 3' UTR, exon, intron, etc.). Due to discrepancies between gene symbols from UCSC, NCBI and Ensembl, and multiple gene symbols for some CLIP targets, only 947 of the original 1049 CLIP targets could be matched to genes from the RNA-Seq data. This represented 779 different genes, as some genes had CLIP targets in more than one region and/or multiple CLIP targets within a region. Not counting multiple targets per region but allowing more than one region per gene, there were 360 genes with CLIP targets in the 3' UTR, 269 genes with CLIP targets in exons and 163 genes with CLIP targets in introns. For the KD vs. IR and OE vs. IR comparisons, the number of significantly up-regulated, down-regulated and non-significant genes among the 779 genes with CLIP targets were compared with the rest of the 13,863 genes using a Pearson's chi-squared test. Within each

region, a similar comparison between the numbers of up, down and no-change genes with CLIP targets versus the rest of the genes without targets in that region was done with separate Pearson's chi-squared tests for each region.

Heat Map

Expression levels of the 656 genes with MOV10 iCLIP sites were examined in MOV10 over-expression (OE) and knockdown (KD) experiments: 312 had FDR $p < 0.05$ in KD vs IR (mock treatment) (172 up and 140 down) and 412 had FDR $p < 0.05$ in OE vs IR (123 up and 289 down). Combining these lists yielded 541 genes, whose expression is visualized in the heat map using Weighted Gene Correlation Network Analysis (WGCNA) (Langfelder and Horvath, 2008; Zhang and Horvath, 2005). This approach divided the genes into different “modules”—indicated on the left by color--that share a similar expression pattern. There were 4 modules (1-turquoise, 12-tan, 10-purple and 14-cyan) with a total of 139 genes that were anti-correlated: specifically, 100 genes increased in KD and decreased in OE and 39 decreased in KD and increased in OE.

Luciferase reporter constructs

Luciferase assay constructs were obtained by cloning the 3’UTR of target genes amplified by PCR using gene specific primers (Table S7 and genomic DNA obtained from HEK293T cells. The amplified fragment was cloned into the Psicheck2 vector using the NotI and Xho restriction sites. The TGFB1 in PGL3 construct was obtained from John Martin, Cardiff University. HEK293T cells were

seeded at $5x10^4$ cells into a 24 well plate for 24 hours and transfected with irrelevant or MOV10 specific siRNAs (Dharmacon) using PEI for 6 hours. 24 hours later, a second transfection containing irrelevant or MOV10 specific siRNA, 900 ng eGFP, 100 ng Psicheck luciferase reporter was performed. 500 ng TGFB1 in PGL3 was transfected with 500 ng eGFP, and 10 ng renilla 24 hours post MOV10 knockdown. Luciferase activity was measured using a dual-luciferase reporter assay kit (Promega) on a SynergyTM HT Multi-detection plate reader 24 hours post secondary transfection.

MOV10 IP and qRT-PCR of its bound mRNA in FMRP knockdown/ overexpression HEK293 cells.

Bead preparation: Protein A Sepharose beads were washed with lysis buffer (150 mM NaCl, 50 mM Tris (pH 7.5), 30 mM EDTA, 0.5% TritonX, protease and RNase inhibitor) and 5 μg of Rabbit anti- MOV10 IgG (Bethyl Laboratories, Inc) was added to beads. Beads containing irrelevant IgG were prepared using the same method (negative control).

HEK293T cells were transfected three times at 24 hour intervals with irrelevant (IR) or FMR1 siRNAs (KD). HEK 293T cells were transfected one time, 24 hours prior to lysing the cells, with FMRP-eGFP or eGFP (control) for overexpression. Cells (~1.2 X 10^7) were lysed (20mM Tris pH 7.5, 200 mM Sodium Chloride, 2.5 mM Magnesium Chloride, 0.5% Triton X-100, Protease Inhibitor, RNase Inhibitor) and immunoprecipitated with prepared beads at 4°C for 12 hours. The beads were then washed and treated with DNAse at 37°C for

10 minutes followed by proteinase K at 37°C for 10 minutes. The RNA was extracted with phenol-chloroform and precipitated in ethanol, converted into cDNA using Oligo dT primer and Superscript III Reverse Transcriptase. qRT-PCR was performed with iQ SYBR Green Supermix (Bio-Rad) using a StepOnePlus RT PCR machine (Applied Biosystems) with gene specific primers (Table S5). Relative levels of mRNA were determined against total RNA samples. Statistical significance was determined using a Student t-Test on samples performed in triplicate.

AGO IP and qRT-PCR of its bound mRNA

Bead preparation: Protein A Sepharose beads were washed with 0.1M sodium phosphate (pH 8.0) and 0.12 mg of Rabbit anti- mouse IgG (Jackson Immunoresearch) was added to beads, followed by 10 μg of Ago 2A8 antibody (Millipore Corp).

HEK293T cells were transfected three times at 24 hour intervals with irrelevant (IR) or MOV10 siRNAs (KD). Cells (~1.2 X 10^7) cells were lysed (20mM Tris pH 7.5, 200 mM Sodium Chloride, 2.5 mM Magnesium Chloride, 0.5% Triton X-100, Protease Inhibitor, RNase Inhibitor) and immunoprecipitated with prepared beads at 4°C for 12 hours. The beads were then washed in sodium phosphate pH 8.0 and treated with DNAse at 37°C for 10 minutes followed by proteinase K at 37°C for 10 minutes. The RNA was extracted with phenol-chloroform and precipitated in ethanol, converted into cDNA using Oligo dT

primer and Superscript III Reverse Transcriptase. qRT-PCR was performed with iQ SYBR Green Supermix (Bio-Rad) using a StepOnePlus RT PCR machine (Applied Biosystems) with gene specific primers (Table S5). Relative levels of mRNA were determined using CERS concentration for normalization. Statistical significance was determined using a Student t-Test on samples performed in triplicate.

Supplemental References

Althammer, S., Gonzalez-Vallinas, J., Ballare, C., Beato, M., and Eyras, E. (2011). Pyicos: a versatile toolkit for the analysis of high-throughput sequencing data. Bioinformatics *27*, 3333-3340.

Ascano, M., Mukherjee, N., Bandaru, P., Miller, J.B., Nusbaum, J.D., Corcoran, D.L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M.*, et al.* (2012). FMRP targets distinct mRNA sequence elements to regulate protein expression. Nature *492*, 382-386.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series *B*, 289-300.

Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D.*, et al.* (2001). Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. Cell *107*, 477-487.

Ceman, S., Brown, V., and Warren, S.T. (1999). Isolation of an FMRP-associated messenger ribonucleoprotein particle and identification of nucleolin and the fragile X-related proteins as components of the complex. Mol Cell Biol *19*, 7925-7932.

Darnell, J.C., Van†Driesche, S.J., Zhang, C., Hung, K.Y.S., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W.*, et al.* (2011). FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. Cell *146*, 247-261.

Devys, D., Lutz, Y., Rouyer, N., Bellocq, J.-P., and Mandel, J.-L. (1993). The FMR-1 protein is cytoplasmic, most abundant in neurons, and appears normal in carriers of the fragile X premutation. Nat Genet *4*, 335-340.

Gentleman RC, C.V., Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biology *5*, R80.

Hofacker, I.L., Fontana, W., F., S.P., Bonhoeffer, S., Tacker, M., and Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. Monatshefte f Chemie *125*, 167-188.

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol *17*, 909-915.

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2011). iCLIP - Transcriptome-wide Mapping of Protein-RNA Interactions with Individual Nucleotide Resolution. J Vis Exp, e2638.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Meth *9*, 357-359.

Lorenz, R., Bernhart, S.H., Honer zu Siederdissen, C., Tafer, H., Flamm, C., and Stadler, P.F. (2011). ViennaRNA Package 2.0. Algorithms for Molecular Biology *6*, 26.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal *17*, 10-12.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology *11*, R25.

Stetler, A., Winograd, C., Sayegh, J., Cheever, A., Patton, E., Zhang, X., Clarke, S., and Ceman, S. (2006). Identification and characterization of the methyl arginines in the fragile X mental retardation protein Fmrp. Hum Mol Genet *15*, 87-96.

Team, R.C. (2013). R: A language and environment for statistical computing. In R Foundation for Statistical Computing (Vienna, Austria: ISBN 3-900051-07-0).

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: A method for identifying protein-RNA interaction sites in living cells. Methods *37*, 376.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP Identifies Nova-Regulated RNA Networks in the Brain. Science *302*, 1212-1215.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184-2185.

Winograd, C., Clayton, D., and Ceman, S. (2008). Expression of fragile X mental retardation protein within the vocal control system of developing and adult male zebra finches. Neuroscience *157*, 132-142.

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol *4*, article 17.