Supplemental Information

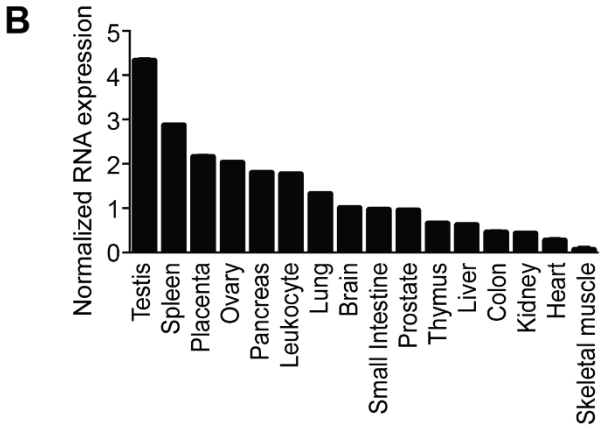
Senataxin Mutation Reveals How R-loops Promote Transcription by Blocking DNA Methylation at Gene Promoters

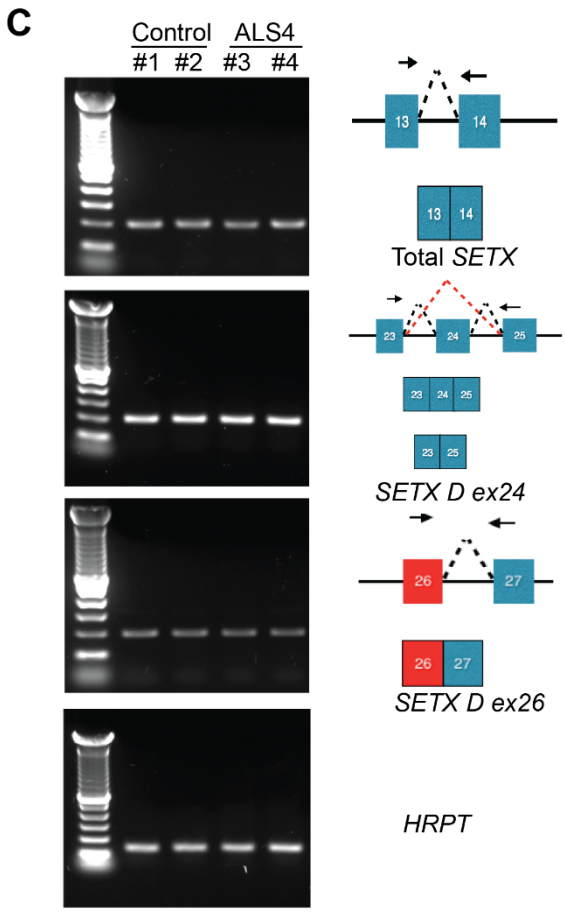
Christopher Grunseich, Isabel X. Wang, Jason A. Watts, Robert D. Guber, Joshua T. Burdick, Zhengwei Zhu, Alan Bruzel, Tyler Lanman, Kelian Chen, Alice B. Schindler, Nancy Edwards, Abhik Ray-Chaudhury, Jianhua Yao, Tanya Lehky, Grzegorz Piszczek, Barbara Crain, Kenneth H. Fischbeck*, Vivian G. Cheung*

* Please send correspondence to KHF (fischbeck@ninds.nih.gov); VGC (vgcheung@med.umich.edu, lead corresponding author)

A

Tissue	FPKM	Fluorescence
Cerebral Cortex	9.7	6.9
Thyroid Gland	30.6	10.6
Adrenal Gland	8.2	2.8
Bone Marrow	8.8	4.5
Spleen	22.7	-
Heart Muscle	5.5	5.7
Skeletal Muscle	2.2	4.3
Smooth Muscle	10.9	5
Lung	13.1	4.7
Liver	7.2	5.1
Gall Bladder	12.8	-
Pancreas	2	3.4
Esophagus	8.6	-
Stomach	9.8	-
Duodenum	7.8	-
Small Intestine	9.7	4.3
Colon	11.4	3.5
Rectum	10.8	-
Kidney	8.9	3.7
Urinary Bladder	12.4	-
Ovary	11	2.9
Testis	29.2	4
Adipose Tissue	10.1	4.5
Skin	8.6	3.5





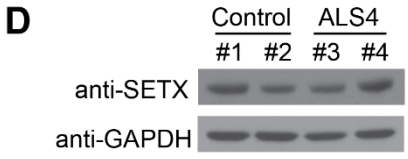


Figure S1. Senataxin is expressed in many cell types. (**A**) Senataxin expression data extracted from the Human Protein Atlas project and BioGPS gene atlas. Gene expression levels indicated as Fragments Per Kilobase of transcript per Million mapped reads (FPKM) and in arbitrary fluorescence units. (**B**) Normalized Senataxin expression from pooled human tissue RNA evaluated by quantitative PCR. Normalized expression in brain set to 1. (**C**) Isoform specific analysis of senataxin expression in control and patient fibroblasts. Gel pictures represent PCR results using the 4 pair primers: 1, primers designed to amplify exons 13 and 14 which are found in all known isoforms, and represent total *SETX* mRNA by RT-PCR. Expression levels of *SETX* are similar in ALS4 patients and controls. 2, RT-PCR amplification of full length exon 24 containing cDNA. Splice variant missing exon 24 was not detected. 3, Splice isoform containing the additional exon 26 detected in all samples evaluated. 4, *HPRT* expression for normalization. (**D**) Western blot shows similar expression of SETX protein in ALS4 patients and controls. Related to Figure 1.

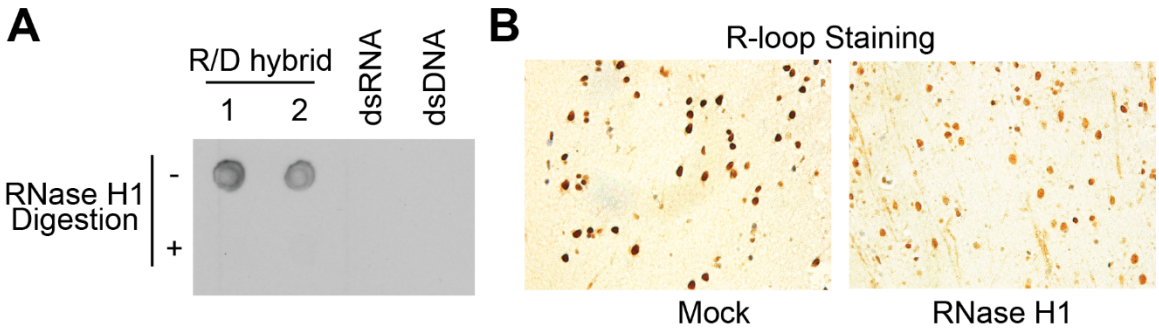


Figure S2. S9.6 antibody specifically recognizes R-loop structure. (**A**) Dot-blot shows S9.6 antibody detects RNA-DNA hybrids formed using oligonucleotide controls with and without *in vitro* RNaseH1-mediated digestion. Double-stranded RNA and DNA included as a loading control. (**B**) S9.6 staining of control motor cortex tissue following either mock enzyme or RNase H1 digestion. Related to Figure 2.

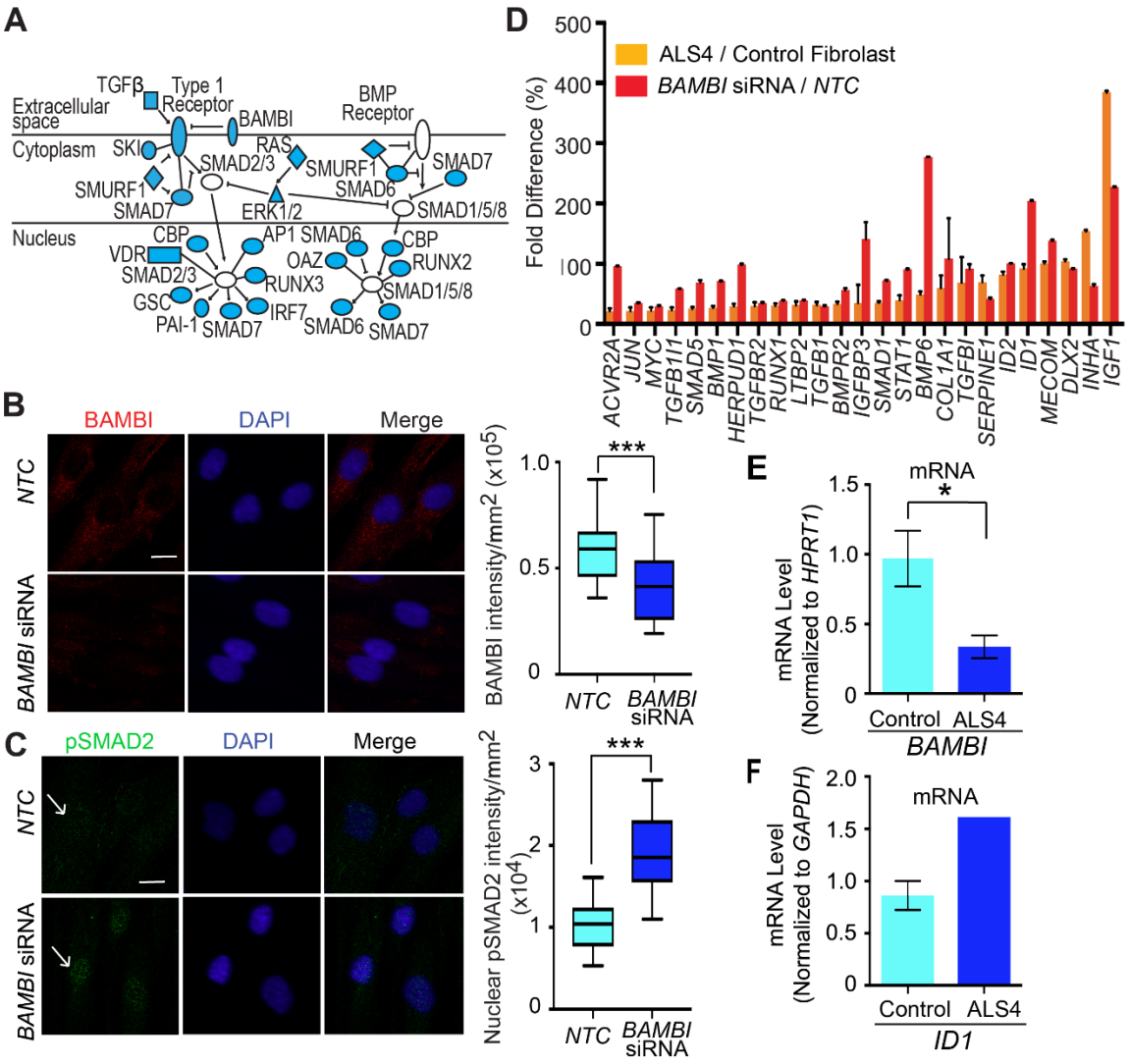


Figure S3. TGFβ activation in ALS4 . (A) RNAseq was performed on 3 control and 3 ALS4 patient fibroblasts and genes were selected with fold differences of (ALS4/control) ≥1.5 or ≤0.67. Genes that are differentially expressed between patients and controls are highlighted in blue. **(B)** Reduced BAMBI immunofluorescence in control fibroblasts following *BAMBI* siRNA mediated knockdown, 2 replicates for each sample, 8 fields per sample, $p < 0.001$, scale bar = 15μm. **(C)** Increased pSMAD2 immunofluorescence staining in control fibroblasts following *BAMBI* siRNA mediated knockdown, 2 replicates for each sample, 8 fields per sample, $p < 0.001$, scale bar = 15μm. Arrows show the nuclear translocation of pSMAD2. **(D)** Expression of upregulated genes (RT-PCR) in the TGFβ pathway from a control fibroblast line following *BAMBI* siRNA knockdown (2 replicates for each sample) (in red), among those genes that were upregulated in the 5 patient and control fibroblasts from Figure 3B (shown here in orange), error bars show SEM. *BAMBI* knockdown resulted in the upregulation (by at least 20%) of 69% (25/36) of the genes previously found to be increased in the patient fibroblasts. Of the 11 remaining genes, 4 were increased between 10-20%, 2 were decreased by at least 10%, and 5 had no change with *BAMBI* knockdown. **(E)** *BAMBI* expression levels are reduced in ALS4 patient fibroblasts. Quantitative PCR analysis from 5 control and 5 ALS4 patient fibroblasts shows a significant ($p<0.05$, t-test) reduction in *BAMBI* expression in the patient samples. mRNA was normalized to *HPRT1* expression. n=2. Error bars = SEM. (**F**) Quantitative PCR shows higher levels of ID1 in an ALS4 lumbar spinal cord compared to controls (n=2) (2 replicates from each sample). mRNA was normalized to *GAPDH*. Related to Figure 3.

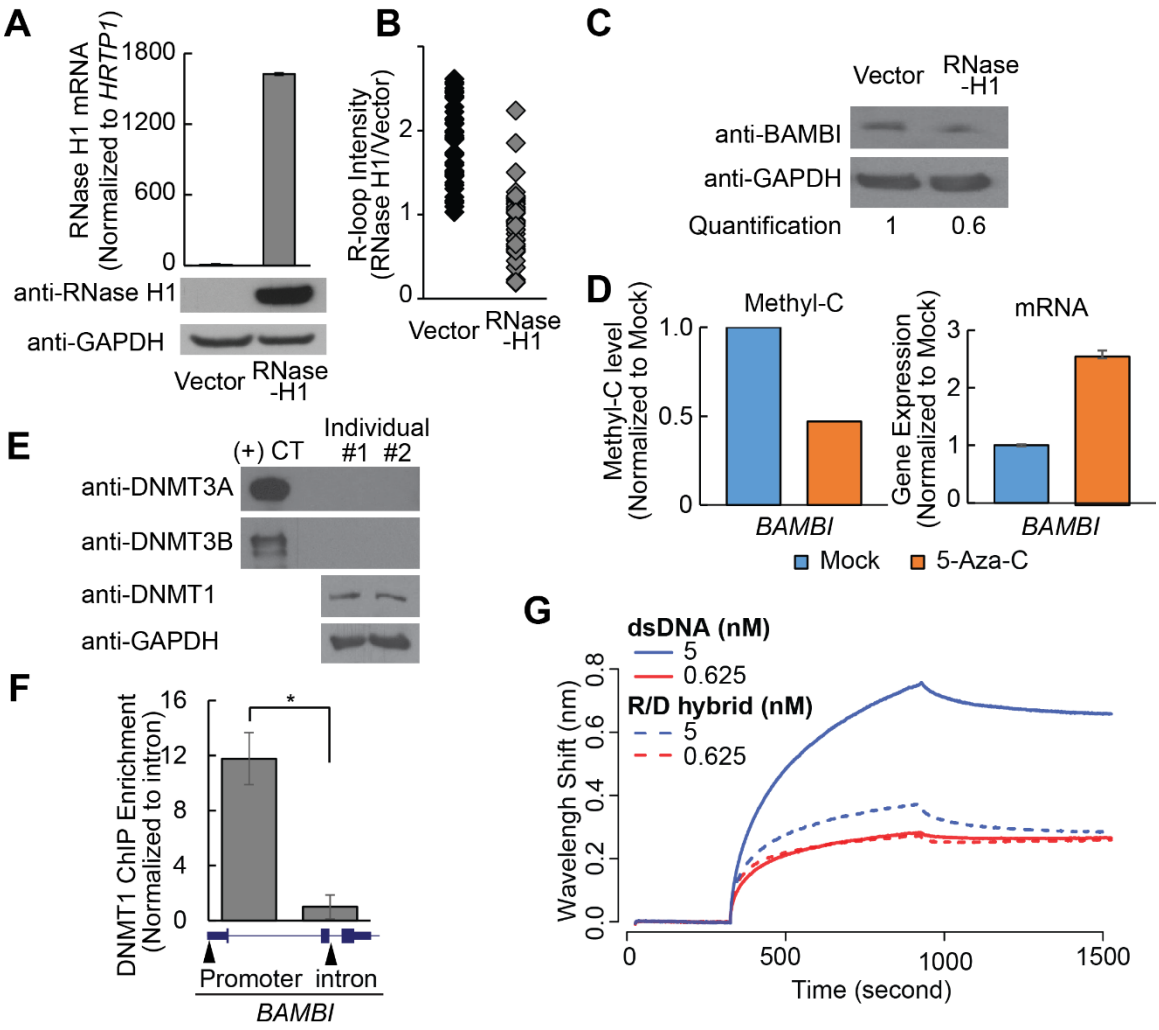


Figure S4. R-loops affect DNMT1 binding which consequently influences *BAMBI* expression. (**A**) Overexpression of human RNaseH1 in primary fibroblasts. mRNA and protein levels of RNaseH1 were measured by quantitative PCR and western blot, respectively. (**B**) Overexpression of RNaseH1 led to reduced R-loop formation. R-loop signal from immunofluorescence staining by S9.6 antibody from 100 nuclei of fibroblasts transfected with RNase H1 or empty vector was quantified and normalized to DAPI staining using ImageJ. (**C**) Overexpression of RNase H1 led to reduced protein expression of BAMBI. Quantification of each band was carried out using Image J, and BAMBI signal was normalized to that of GAPDH from the same sample. (**D**) Treatment by the DNMT inhibitor 5-Aza-C led to reduction of methylation level at *BAMBI* promoter and increased gene expression. Average methyl-C level of all cytosines in *BAMBI* promoter (TSS+/-500bp) is shown. (**E**) DNMT1, but not DNMT3A and DNMT3B, is not expressed in skin fibroblasts. Recombinant DNMT3A and DNMT3B proteins were loaded as positive controls. (**F**) DNMT1 binds to BAMBI promoter in vivo. ChIP-qPCR shows enrichment of DNMT1 in BAMBI promoter. (**G**) Higher association of DNMT1 for dsDNA than RNA/DNA hybrid. Binding of DNMT1 to dsDNA (blue) or RNA/DNA hybrid (red) was assayed by biolayer interferometry. Association of DNMT1 with dsDNA or RNA/DNA hybrid was recorded during 0-300 seconds, followed by dissociation. Related to Figure 4.

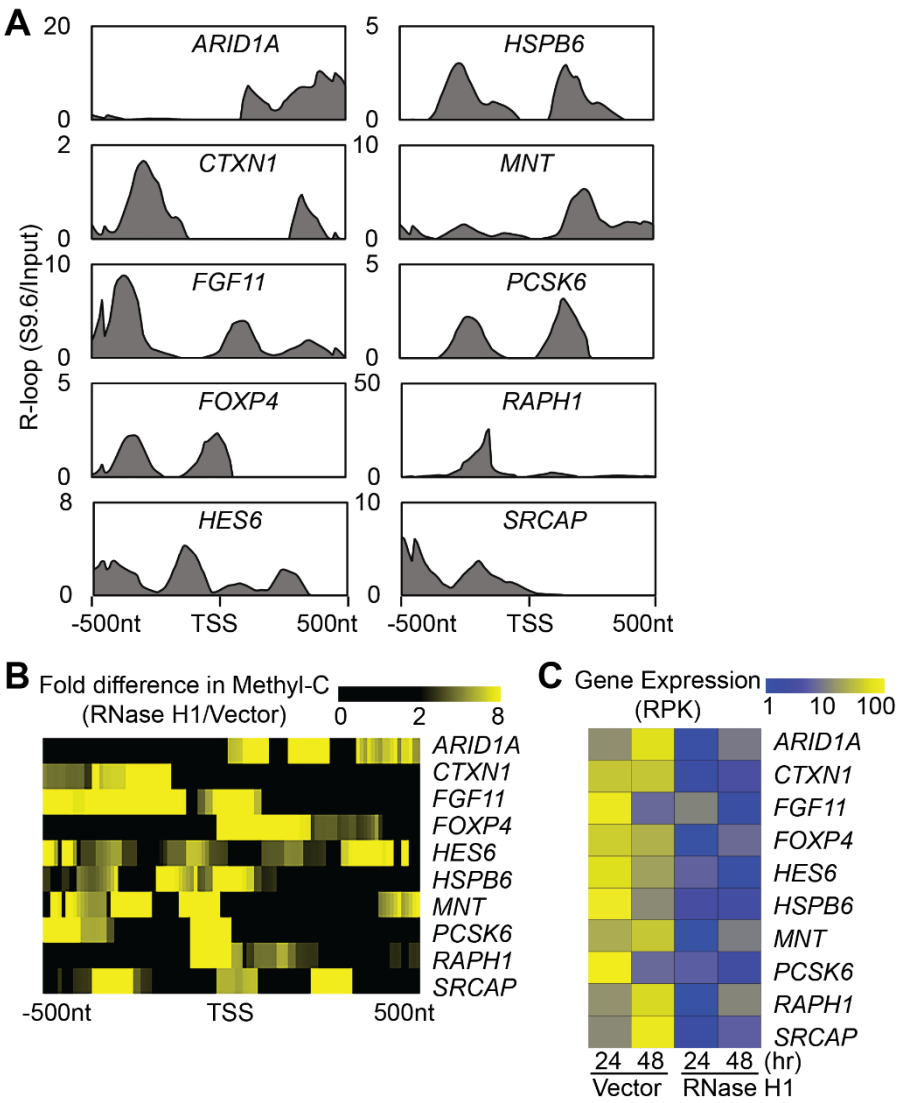


Figure S5. Ten representative genes that are regulated by R-loops. (**A**) Examples of 10 genes (From Fig 7A) that have R-loops identified by DRIP-Seq data in the promoter regions. (**B**) MethylC-Seq data on these 10 genes that showed increased methylation at promoter regions following RNase H1 overexpression. (**C**) The expression levels of these 10 genes decreased following RNase H1 overexpression. Related to Figure 7.